Les hormones thyroïdiennes, leurs récepteurs et l'évolution de la métamorphose chez les Chordés.

Mathilde Paris

THÈSE

Présentée

devant L'ECOLE NORMALE SUPÉRIEURE DE LYON

pour l'obtention

du DIPLÔME DE DOCTORAT

soutenue le 18 décembre 2008

par

MATHILDE PARIS

THYROID HORMONES, THEIR RECEPTORS AND THE EVOLUTION OF METAMORPHOSIS IN CHORDATES

Jury:	Vincent LAUDET	Directeur de thèse
	Béatrice DESVERGNE	Rapportrice
	Detlev ARENDT	Rapporteur
	Philippe JANVIER	Examinateur
	Nicholas HOLLAND	Examinateur
	Frédéric FLAMANT	Examinateur
	Hector ESCRIVA	Membre invité

THYROID HORMONES, THEIR RECEPTORS AND THE EVOLUTION OF METAMORPHOSIS IN CHORDATES

MATHILDE PARIS

Doctorate of Life Science

18 December 2008

ABSTRACT

In an attempt to understand how the regulation of development evolves, particular attention has been put on transcription factors, which regulate gene expression during development. Among transcription factors, nuclear hormone receptors (NRs) have a peculiar status linked to their ligand-dependent activity. However, their role in the evolution of development is still poorly understood. In this context, I studied during my thesis the evolution of the estrogen receptor (ER) and of the thyroid hormone receptor (TR) in chordates (including vertebrates, urochordates like sea squirt and cephalochordates like amphioxus), by focusing on the amphioxus species *Branchiostoma floridae*.

Combined molecular and bioinformatic approaches allowed me to show that the amphioxus ER does not bind estrogen, suggesting that estrogen binding appeared in the vertebrate lineage. In parallel, a large part of my thesis was dedicated to understand how metamorphosis evolved in the chordate lineage, as a paradigm of the evolution of a developmental process. Although most chordates metamorphose, the morphological changes during larva-to-adult transitions vary extensively from one species to another. Does the molecular determinism of metamorphosis in this group reflect this morphological diversity? In the well-studied vertebrates, metamorphosis is triggered by thyroid hormones (THs) binding to their receptor TR, member of the NR superfamily. In order to get better insight into the evolution of the molecular determinism of metamorphosis in chordates, I focused on the most basal chordate amphioxus. Combined biochemical and phylogenetic approaches allowed me to establish that amphioxus produces various THs through metabolic pathways homologous to vertebrate ones. Then I showed that TH-dependent TR activation is essential for metamorphosis induction in amphioxus, like in vertebrates, with the slight difference that the active TH is not T_3, the classical vertebrate TH, but possibly its derivative TRIAC. Consequently the homology of metamorphosis in chordates is revealed by the conservation of its triggering mechanism. This suggests that the evolution of metamorphosis in chordates is marked by the conservation of the couple TH/TR whereas other parts of the regulatory network may change to underlie the morphological diversity observed nowadays.

RÉSUMÉ

Un des principaux objectifs de ma thèse a été de comprendre comment les processus du développement évoluent. Dans ce contexte, les facteurs de transcription sont importants car ils régulent l'expression génique au cours du développement. Parmi les facteurs de transcription, les récepteurs nucléaires (RNs) ont un statut particulier car leur activité est régulée par un ligand. Le rôle qu'ils jouent dans l'évolution du développement et plus généralement leur évolution (leur fonction, leur capacité à lier un ligand...) sont mal compris. Durant ma thèse, je me suis intéressée à l'évolution des RNs chez les chordés (comprenant les vertébrés, les urochordés tels que les tuniciés et les céphalochordés tels que l'amphioxus) en étudiant deux RNs particuliers : le récepteurs aux oestrogènes (ER) et le récepteur aux hormones thyroïdiennes (TR). Les études que j'ai menées ont été réalisées en particulier chez l'amphioxus *Branchiostoma floridae*.

J'ai montré que l'ER d'amphioxus ne lie pas les oestrogènes, suggérant que cette capacité est apparue chez les vertébrés. La majeure partie de ma thèse a été dédiée à l'étude de l'évolution de la métamorphose chez les chordés. Bien que la plupart des chordés métamorphosent, les modifications morphologiques caractéristiques du passage d'une forme larvaire à une forme juvénile sont très variables d'une espèce à l'autre. Cette variabilité morphologique se reflète-telle dans le déterminisme moléculaire de la métamorphose ? Ce dernier est encore mal connu en dehors des modèles vertébrés classiques chez qui la métamorphose est induite par la fixation des hormones thyroïdiennes (HTs) sur TR. Afin de combler ce manque, je me suis intéressée au protochordé amphioxus. Par des approches bioinformatiques et biochimiques, j'ai établi que l'amphioxus produit des HTs par une voie métabolique homologue à celle des vertébrés. J'ai alors montré que les HTs régulent la métamorphose de l'amphioxus, en modulant l'activité de TR, comme chez les vertébrés. Une différence notable est la nature de l'HT active : T_3 chez les vertébrés contre TRIAC, un dérivé de T_3, chez l'amphioxus. J'ai ainsi proposé que l'homologie de la métamorphose chez les chordés est soutenue par la conservation du déterminisme moléculaire de ce processus du développement (le couple HT/TR) alors que les autre parties de la voie de régulation ont été moins conservées au cours de l'évolution et expliquent la diversité morphologique que l'on observe de nos jours.

PUBLICATIONS

Holland N. D., **Paris M.**, and Koop D. (2009) The club-shaped gland of amphioxus: export of secretion to the pharynx in pre-metamorphic larvae and apoptosis during metamorphosis. *Acta Zoologica* **89**, doi: 10.1111/j.1463–6395.(2008)00379.x.

Tocchini-Valentini G. D., Rochel N., Escriva H., Germain P., Peluso-Iltis C., **Paris M.**, Sanglier-Cianferani S., Dorsselaer A. V., Moras D., and Laudet V. (2008) Structural and functional insights into the ligand binding domain of a non-duplicated RXR from the invertebrate chordate amphioxus. *J Biol Chem* **284**, 1938-1948.

Paris M., Brunet F., Markov G., Schubert M., and Laudet V. (2008) The amphioxus genome enlightens the evolution of the thyroid hormone signaling pathway. *Dev Genes Evol* **218**, 667-680.

Paris M., and Laudet V. (2008) The history of a developmental stage: Metamorphosis in chordates. *Genesis* **46**, 657-672.

Schubert M., Brunet F., **Paris M.**, Bertrand S., Benoit G., and Laudet V. (2008) Nuclear hormone receptor signaling in amphioxus. *Dev Genes Evol* **218**, 651-665.

Paris M., Pettersson K., Schubert M., Bertrand S., Pongratz I., Escriva H., and Laudet V. (2008) An amphioxus orthologue of the estrogen receptor that does not bind estradiol: insights into estrogen receptor evolution. *BMC Evol Biol*, **8**, 219.

Paris M., Escriva H., Schubert M., Brunet F., Brtko J., Ciesielski F., Roecklin D., Vivat-Hannah V., Jamin E. L., Cravedi J. P., Scanlan T. S., Renaud J. P., Holland N. D., and Laudet V. (2008). Amphioxus postembryonic development reveals the homology of chordate metamorphosis. *Curr Biol*, **18**, 825–30.

Markov G. V., **Paris M.**, Bertrand S., and Laudet V. (2008) The evolution of the ligand / receptor couple: a long road from comparative endocrinology to compara-

tive genomics. *Molecular and Cellular Endocrinology*, **293**, 5–16.

Pagès M., Calvignac S., Klein C., **Paris M.**, Hugues S., and Hänni C. (2008) Combined analysis of fourteen nuclear genes refines the Ursidae phylogeny. *Mol Phylogenet Evol*, **47**, 73-83.

Fuentes M., Benito E., Bertrand S., **Paris M.**, Mignardot A., Godoy L., Jimenez-Delgado S., Oliveri D., Candiani S., Hirsinger E., D'Aniello S., Pascual-Anaya J., Maeso I., Pestarino M., Vernier P., Nicolas J. F., Schubert M., Laudet V., Geneviere A. M., Albalat R., Garcia Fernandez J., Holland N. D., and Escriva H. (2007) Insights into spawning behavior and development of the european amphioxus (*Branchiostoma lanceolatum*). *J Exp Zool B Mol Dev Evol*, **308**, 484-493.

Bertrand S., Camasses A., **Paris M.**, Holland N.D., and Escriva H. (2006) Phylogenetic analysis of Amphioxus genes of the proprotein convertase family, including aPC6C, a marker of epithelial fusions during embryology. *Int J Biol Sci* **2**, 125-132.

Hassani Z., François J.-C., Alfama G., Dubois G. M., **Paris M.**, Giovannangeli C., and Demeneix B. A. (2007) A hybrid CMV-H1 construct improves efficiency of PEI-delivered shRNA in the mouse brain. *Nucleic Acids Res*, **35**, e65.

Brunet F. G., Crollius H. R., **Paris M.**, Aury J. M., Gibert P., Jaillon O., Laudet V., and Robinson-Rechavi M. (2006) Gene loss and evolutionary rates following whole-genome duplication in teleost fishes. *Mol Biol Evol*, **23**, 1808–16.

REMERCIEMENTS

Je voudrais tout d'abord remercier les membres de mon jury de thèse Béatrice Desvergne, Detlev Arendt, Philippe Janvier, Nicholas Holland et Frédéric Flamant pour avoir accepté de juger mon travail, pour leur indulgence et leurs commentaires précieux.

Je voudrais remercier Vincent Laudet de m'avoir accueillie dans son laboratoire (pendant 5 ans et demi tout de même). On n'a pas assez discuté, on n'a pas toujours été d'accord mais on a toujours fini par un bon compromis. Je rends hommage à son ouverture d'esprit et à sa porte toujours ouverte. Vincent m'a donné un passionnant sujet de thèse, qui a métamorphosé ma vision du développement tardif chez les chordés (non pas que j'en avais une avant la thèse...). Zoupette à fleurs te remercie, chef.

Je voudrais remercier Hector Escriva pour m'avoir si bien encadrée pendant mon DEA et ma première année de thèse. Il était disponible et attentif juste comme il faut et est resté très concerné par mon travail malgré son départ de Lyon. Sa profonde gentillesse s'accompagne d'un esprit vif et d'une efficacité impressionnante. Une grande partie de ce que je sais maintenant vient de lui. Alors un respectueux merci à un brillant chercheur.

Je voudrais également chaudement remercier Linda Holland de m'avoir aidée à travailler sur ce charmant petit ver qu'est l'amphioxus. J'ai été accueillie à Tampa par Linda comme si j'étais de sa propre équipe, j'ai pu partager la joie de la pêche avec les requins et les raies (aïe), le visionnage des Simpson et les nuits blanches avec son équipe. Mille merci pour cette nouvelle variété d'algues Isochrysis qui m'a été d'une utilité que seuls les éleveurs d'amphioxus peuvent apprécier. Je lui suis extrêmement reconnaissante pour tout ce qu'elle a fait pour moi. Je voudrais remercier une fois encore Nick pour m'avoir tant apporté. Sa passion pour la zoologie toujours si vive, son humour et sa disponibilité ont fait de lui un mentor pour moi, et je le salue respectueusement. C'est aussi grâce à lui que j'ai pu oublier dans l'alcool de Rachelle mon angoisse à l'idée d'aller chercher des larves d'amphioxus à la tombée de la nuit dans le garde-manger des requins affamés. Une fausse note toutefois : la conduite de Nick au volant, un chouïa trop napolitaine au goût de mon cardiologue.

Je tiens également à remercier le plus évolutionniste des spécialistes de la leucémie promyélocytaire aigüe (et malheureusement pas l'inverse), Gérard B. Il m'a beaucoup appris, non seulement parce qu'il s'y connaît un max sur les récepteurs nucléaires, mais aussi, et surtout, parce qu'il a empoisonné mes petits poumons pendant plus de 3 ans avec ses Lucky Strike. Il a subi en contre-partie mes jérémiades sur TR qui ne lie pas T3, TR qui lie TRIAC, TR qui n'est pas bien produit en bactéries, TR TR TR... il a été cette oreille

attentive et amicale dont j'avais tant besoin. En plus, il a un goût sûr en matière de comté ("c't'excellent"). Bref un poteau, en plus d'un mentor, qui me manquera beaucoup.

*Un grand merci également à Sophie, pour sa passion de la science (une chercheuse comme il en faudrait d'avantage dans les labos), pour son sourire si lumineux, pour son intérêt et son aide enthousiastes. Grâce à elle, je ne crible plus les clones comme avant. Je voudrais également remercier Marie S. Elle a été une voisine charmante, toujours souriante, et toujours d'une intelligence redoutable. Il faut par contre faire attention avec Marie parce que si on a un problème et qu'on lui en parle, elle passera le temps qu'il faut pour vous le résoudre. A ne pas abuser si l'on veut pouvoir se regarder dans un miroir. Un petit clin d'oeil allumeur à Gérard T, c'est ma façon à moi de le remercier de sa bougritude. Toutes ces années, il a été d'un réconfortant graveleux. Je tire mon chapeau à son expertise biochimique, qui m'a bien aidée ainsi qu'à sa dextérité inégalable dans l'art du jeu de mots. Merci aussi à Pascale qui fait de si bons gâteaux au chocolat et qui m'a fait découvrir les toutes petites souris (*Nannomys *pour les intimes).*

Je voudrais également remercier Florent pour sa bonne humeur quasi permanente, c'est un charmant camarade de paillasse. Alors si tu es polonaise, pas catholique intégriste, que tu as moins de 25 ans, que tu aimes le pâté et les lamproies (mais pas pour les mêmes raisons), et que les vidéos équestres ne te font pas peur, je connais un brillant thésard qui cherche une étudiante à encadrer de près. Je remercie également son chef Michael, qui a bien du mérite, et qui m'a fait profiter de ses talents d'écriture. On a échangé quelques discussions des plus constructives, quelques bonnes engueulades aussi, et quelques bières.

Il y a également tous les autres membres de l'équipe : la douce et chocolatée Mme Christelle Gobet Betoulle, l'enthousiaste et geakesque Alexa, Anne et ses petits mots doux, Laure qui aime tant les poissons (surtout leurs épitaphes), Maria qui perturbe la sexualité du zebrafish, Charles qui travaille à des heures indues, Jasmine qui travaille trop, Maline qui se lève trop tôt le matin, Marie T qui apprécie les ambassades...

Je souhaite ne pas oublier les passés au milieu de cet océan de présents. Alors merci à Stéphanie pour sa gentillesse et sa disponibilité, je lui tire mon chapeau d'avoir fait une thèse, un post-doc et 2 enfants. Je félicite aussi Yann qui a su mener de front 2 carrières (chercheur et stadier). Merci aussi au si discret Shin-Ichi, à la joyeuse Lamia, à François pour son aide du début, à Juliette qui m'a marquée par sa générosité, à Arnaud Martin qui vole vers d'autres horizons évolutifs, à Arnaud Chaumot qui travaillait dans plus de 3 dimensions, à Fred Brunet qui a fait des progrès en braille grâce à moi et à Raquel pour son encouragement au nudisme. Je reste admirative de l'amour d'Ingrid pour l'enseignement-recherche. Je souhaite à Ferdi de s'éclater en post-doc, même s'il y verra moins de poubelles crâmer qu'à Endoume. J'espère qu'il ne perdra jamais sa "roots" attitude version evo-devo.

Un grand merci à Fred Flamant d'avoir aiguisé mes fantasmes laborantins et pour l'intérêt qu'il a porté à mon travail. Merci aux autres thésards d'avoir souffert avec moi et

d'avoir entre autres ouvert mes yeux sur les pigeons mognons. Merci à Barbara Demeneix pour son envie de soniquer des amphioxus ainsi qu'aux membres de l'ENS pour leur aide.

Je n'aurais également pas pu réaliser cette thèse sans le soutien de ma famille et en particulier de mes parents. Ils ont toujours été derrière moi bien qu'ils aient désapprouvé certains de mes choix (ma mère trouvant que je gaspillais les cônes qu'elle m'a vue jeter après chaque utilisation : "quel gâchis". Mon père quelque peu déçu de me voir opter pour la biologie plutôt que pour un hypothétique chapeau). Merci aussi à ma famille d'adoption pour ses encouragements et ses présents vaudous. Merci aussi à mes amis qui ont toujours eu autant de mal à se rappeler le nom de la bestiole sur laquelle je travaillais (j'espère que Samia me pardonnera un jour de ne pas être devenue un vrai *docteur).*

Enfin, je remercie Bastien pour son soutien de roc et de choc. Pendant toutes ces années, il s'est toujours ardemment inquiété de savoir si ma PCR avait marché, si j'avais reçu mes échantillons si précieux, si j'avais lu cet article si intéressant, si j'avais pu discuter avec ce chercheur que je voulais voir mais n'osais aborder. Il n'est pas plus petit que Shane, fait la vaisselle mieux que moi et est tellement brillant. C'est mon héros.

Je voudrais terminer ces (longs) remerciements par un rapide portrait sonore du laboratoire. Parce les souvenirs de ces quelques années passées dans l'équipe de Zoologie Moléculaire passent aussi par les oreilles.

*Ambiance au labo: sonnerie du timer que quelqu'un a oublié sur sa paillasse, centri bien (ou mal) équilibrée, pas d'une personne pressée (un maître de conf' qui a cours ? Une manip oubliée ? Un séminaire qui commence ?), aspiration de la hotte, porte de la chambre froide que l'on ouvre, puis que l'on ferme, cri de désespoir d'un étudiant qui vient de perdre ses embryons d'*in situ*, discussion animée sur l'avenir de la recherche. Pour couronner le tout, certaines personnes ont le mauvais goût d'aimer écouter de la musique : Flamenco pour Stéphanie pendant la rédaction de sa thèse, jazz pour Charles quand il travaille la nuit, Muse pour Juliette se préparant au concert du groupe, Joe Dassin pour Florent (et pourtant il n'a pas de raison !), Scorpions d'Alexa (?????), musique de la petite souris qui croit aux miracles pour Gérard T, chanson réclamant à Benoît de se retourner...*

Ensuite il y a certains sons que chacun fait: le rire de Vincent, les expressions teintées d'accent espagnol d'Hector ("un tipeu", "c'pas mal"), les énervements de Michael ("ça m'énerfe", "la pétite"), l'accent du sud d'Arnaud, le rire crystallin de Sophie, les promesses (non tenues) de fessée de la part de Gérard T, la roulette de souris ou les touches de clavier de Marie S, la canette de soda que Maria boit dans l'après-midi, le bruit de pas genre "force tranquille" de Gérard B, le tripotage de mèches de cheveux de Christelle (je sais, ça fait un bruit limité...), les bracelets d'Ingrid quand elle gigote les bras d'excitation, le silence éloquent de François qui n'en pense pas moins, le "hâaää" de Florent, si difficile à imiter sur support papier, les barbarismes jurassiens de Laure...

Ce sont mes petites madeleines à moi.

ACRONYMS

AR	Androgen receptor
CAR	Constitutive androstane receptor
CDS	Coding DNA sequence
CRE	Cis-regulatory element
CRF	Corticotropin releasing factor
DNA	Deoxyribonucleic acid
DBD	DNA binding domain
ER	Estrogen hormone receptor
ERR	Estrogen-related receptor
FXR	Farnesoid X receptor
GR	Glucocorticoid receptor
GRN	Genetic regulatory network
IOD	Deiodinase
IOP	Iodopanoic acid
LBD	Ligand binding domain
LXR	Liver X receptor
MR	Mineralocorticoid receptor
NR	Nuclear hormone receptor
PERT	Thyroid peroxidase
PPAR	Peroxysome proliferator-activated receptor
PR	Progesterone receptor
PXR	Pregnene X receptor
RAR	Retinoic acid receptor

ROR Retinoic acid orphan receptor

RNA Ribonucleic acid

RXR Retinoid X receptor

T_3 Triiodo-l-thyronine

T_3AM 3-triiodothyronamine

T_4 l-thyroxine

TAAR Trace amine-associated receptor

TETRAC Tetraiodothyroacetic acid

TF Transcription factor

TG Thyroglobulin

TR Thyroid hormone receptor

TRIAC Triiodothyroactic acid

USP Ultraspiracle

CONTENTS

LIST OF FIGURES

LIST OF TABLES

PREFACE

AS TIME GOES BY

In this preface, intended for a lay audience (and especially my family), I will introduce my PhD work and briefly situate it in a general context, in order to explain why I chose to work on this subject. More precisely, I will first try to broadly present evolutionary biology (or at least why it interests me), then I will present the field of evolution of development, and I will finally introduce metamorphosis in the context of evolutionary developmental biology. The last paragraph of the preface describes the content of this manuscript.

One of the main aims of my PhD has been to discuss how metamorphosis evolved in the chordate lineage. There are two questions that one may ask: why metamorphosis and why evolution? Let's answer the second one first.

Although the famous quote of Theodosius Dobzhansky **?**: "*Nothing in biology makes sense except in the light of evolution*" may sound commonplace for some people, it does nonetheless well summarize my motivation to study evolution. Indeed, in the attempt to find mechanistic models that explain biological diversity, the time parameter is of special importance since it explains why distantly related species are more different than closely related species, the way cousins look less alike than brothers and sisters. More generally, as all living organisms share a common ancestry, going from close relatedness, like humans and chimpanzees, to deep ancestry, forming a tree of life (Figure **??**), and as all living organisms differ at some point, then they necessarily evolved through time from this common ancestor to extant biodiversity. Consequently many characteristics of any organism find an explanation in the evolutionary pathway that has been taken from its origins to current days. For instance, why do so many people get backaches and more generally why is our skeleton not that well suited for straight position? The recent origins of mankind from quadrupedal ancestors well explain the resulting bad back that evolution tinkered from an horizontal position (**?**). In other words,

> *There is not a single question in biology that can be answered adequately without a consideration of evolution* (**?**).

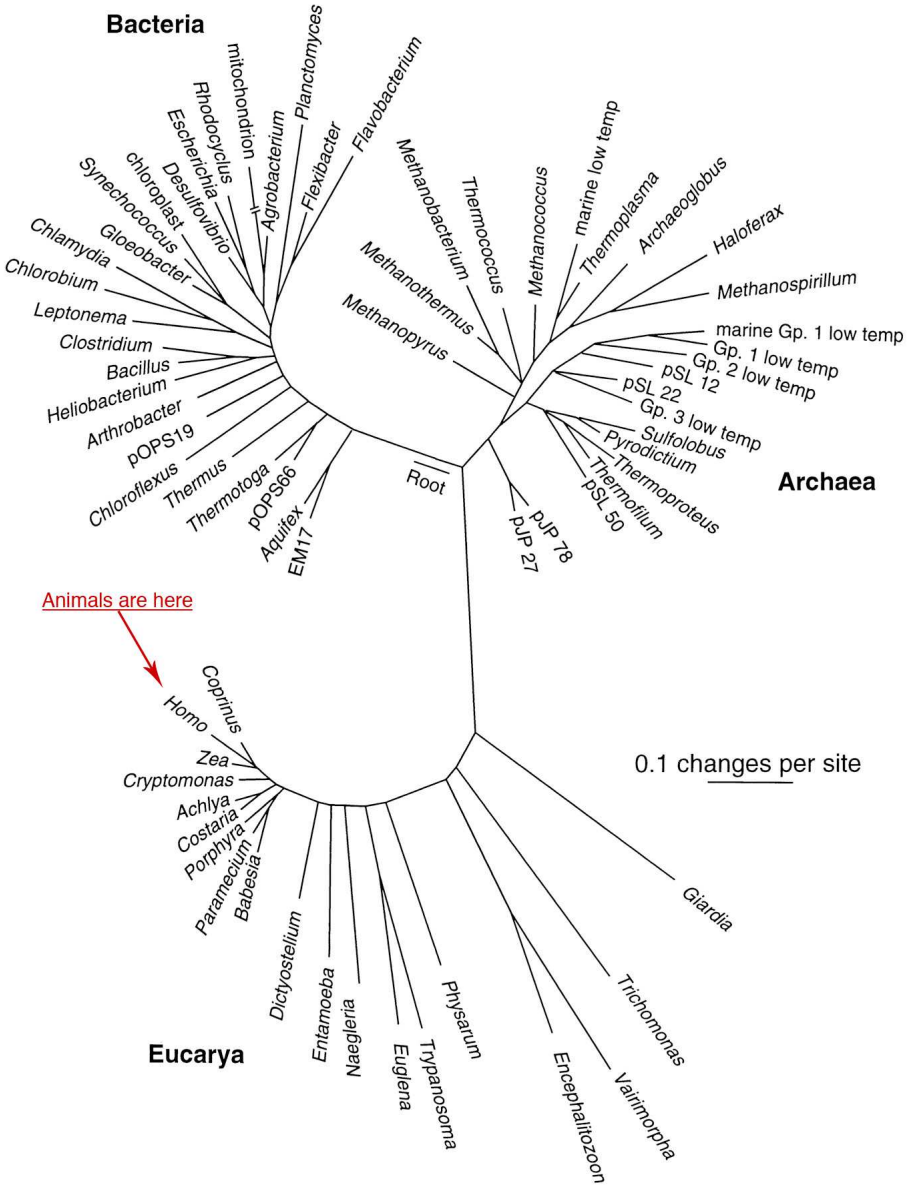


Figure 1: Universal phylogenetic tree, based on SSU rRNA sequences (**?**). This tree illustrates the common ancestry of all living organisms, more than 3 billion years ago. Three kingdoms are clearly visible: the archaeal, eukaryotic and bacterial kingdoms. This tree is presented unrooted because the position of the root is still controversial, although in the bacterial branch is the most admitted hypothesis (for a discussion, see (**?**)).

Evolution is implicit in any subfield of biology. For instance, medicine posits that all organisms of the same species have a similar functioning, and a physician who uses an animal model to test a particular drug implicitly accepts the common ancestry of the animal model with human, evolution from this common ancestor and more specifically inherited similarity (this is homology, see section **??**) of the processes that are scrutinized. Otherwise it would be at the very least immoderate to consider that an efficient medicine in a mouse may be of any potential interest for a human being. In other words, people count on conserved physiological and molecular biological processes between species from a common ancestor to extrapolate rules applicable to humans from their animal model. Evolution makes these similarities decrease with time, explaining why mice are less faithful stand-ins for us than chimpanzees.

While evolutionary concepts are underlying most of biology, the discipline of evolutionary biology is all about understanding the processes by which the living world has changed since the origin of life (**?**). It is this mechanistic approach that gives an explanation to "how" the biodiversity that we have today arose. And to understand the evolution of the shapes of organisms, it is important to study the evolution of development.

EVO-DEVO, A SUBFIELD OF EVOLUTIONARY BIOLOGY

Because evolution unifies all living forms, it can be investigated from virtually all biological areas of studies. Historically, evolution has been associated with systematics (in which species are classified following *relevant* criteria, see section **??**), paleontology (to study the evolution of forms from the fossil record) and population genetics (to study the dynamics of the evolution of mutations in populations of organisms). Although a joint interest in development and evolution dates back to the nineteenth century, it got lost in the first half of the twentieth century, with the building of the modern synthesis. Molecular biology brought evolution and development back together twenty years ago as a new edge of evolutionary biology, called evo-devo, in which evolution is scrutinized in the light of development, and *vice-versa*. Indeed, an individual does not appear already shaped (figure **??**), but comes from a unique cell that develops into a larva or an adult through a very precisely regulated developmental process. Then alterations of development should impact the shape and more generally the phenotype of an organism (e.g. differences in expression of some genes may cause the anterior limb of a mammal to get longer, from mouse-like to bat-like (**?**)). So overall, because the

phenotype is the result of a developmental process, evolution of development is central to understanding the diversity of forms. To cite Hans Spemann (**?**):

> *"We no longer believe, that we first can establish the phylogenetic relations between animals in order to subsequently derive developmental laws. Rather we begin to realize, that we first have to determine these laws, before we can understand or even establish the morphological series that we use to classify organisms."*

Functional approaches are more and more chosen to study issues related to the evolution of development, to unveil genetic mechanisms underlying the evolution of species. This makes evo-devo a very powerful tool to understand the "what" and "how" of multicellular life (**?**).

Figure 2: Reproduction of a homonculus by Nicolas Hartsoecker (1695). It illustrates the "spermid theory" of the origin of babies. According to this theory, individuals contain in their gametes preformed "miniatures" of their descendants, like Russian dolls. The picture represents a "little man" already preformed in the head of a spermatozoid, that simply needs to be deposited in the mother's uterus to further grow. This theory implies that offsprings come from only the father, that there is no fertilization and no subsequent development.

WHY STUDYING THE EVOLUTION OF METAMORPHOSIS?

During my PhD, I have studied the evolution of metamorphosis in chordates (a group containing vertebrates, see section **??**). In this study I define metamorphosis as the larva-to-juvenile transformation like tadpole-to-frog or caterpillar-to-butterfly transitions. More precisely it is a post-embryonic developmental stage allowing a larva to become a juvenile, and which is characterized by drastic ecological, morphological, metabolic and behavioral modifications.

Metamorphosis will be considered in this manuscript in two different and complementary ways: (i) for itself and (ii) as a paradigm for developmental processes.

(i) It is a spectacular developmental process that whets imagination: there is some sort of magic in the abrupt transformation of a tadpole into a frog, or of a maggot into a fly, and indeed this developmental process does trigger interest among

children (the macroscopic scale at which metamorphosis occurs compared to the microscopic scale of most other developmental processes may help). Moreover metamorphosis is so widespread among animals (see chapter **??**) that the evolution of this dramatic event of larva-to-juvenile transition is especially intriguing, considering the theoretical strong obstacles to such a non-linear developmental strategy. In addition, and as will be detailed in later chapters, metamorphosis is hormonally regulated (thyroid hormones in the case of amphibians), which makes it an ideal model to study the role of endocrine systems during development. Overall, how the tadpole-to-frog transformation evolved has remained unclear and therefore deserves more insight.

(ii) Metamorphosis is highly variable among chordates: it is thus difficult to decipher whether it appeared several times independently during chordate evolution or was present in the chordate ancestor (chordate metamorphoses would be homologous) and subsequently diverged widely between species. Metamorphosis can then be studied as an illustration of difficulties in determining homology and more generally it can be considered as an example used to decipher the general rules that constrain the evolution of developmental processes. As explained in the previous section, the mechanisms through which development evolves and impacts species phenotypes are getting increasing attention from evolutionary biologists. My work is a small contribution to this field of evolutionary biology.

During my PhD I have studied the evolution of the genetic network regulating metamorphosis in chordates. This work led me to propose a model in which the evolution of metamorphosis at the phenotypic level is underlain by the evolution of the corresponding genetic network. Before presenting my thesis from the above two axes, I will first present the evolutionary context of my work (chapter **??**). Then I will briefly introduce the notion of metamorphosis in both broad and developmental senses (chapter **??**), with a special focus on chordates and amphioxus. I will then describe the regulatory network of metamorphosis in chordates (and especially amphibians) with an emphasis on the thyroid hormone receptor (chapters **??** and **??**) (pieces of these chapters were modified from **?**).

After this introduction, I will present my PhD work in the form of articles recently published: the main part of my work was dedicated to understanding metamorphosis in amphioxus and especially its molecular determinism. For this purpose, I have sought into the thyroid hormone signaling pathway in amphioxus (chapter **??**), and its role in the regulation of metamorphosis (chapter **??**). To complete this work, unpublished additional experiments were done and are presented in chapter **??**. These articles allowed me to propose a model of the evolution of metamorphosis in chordates, based on alterations of its molecular determinism

(this model is discussed in chapter **??**, and uses modified parts of (**?**)). Contemporaneously I have also studied the evolution of the estrogen receptor in amphioxus, a gene closely related to the thyroid hormone receptor (chapters **??** and **??**. In chapter **??**, the evolution of NRs will also be discussed).

Part I

INTRODUCTION

1 HOMOLOGY AND CHORDATE PHYLOGENY

1.1 DESCENT WITH MODIFICATION AND HOMOLOGY

As briefly stated in the preface, all living organisms share a common ancestry, 3.8 billion year ago. Then the evolutionist's job is to understand how the subsequent 3.8 billion years of evolution gave rise to the extant diversity from the common ancestor. How to determine what criteria to look for in a given set of species?

1.1.1 *"All true classification is genealogical"*

During evolution, characteristics of organisms change, sometimes leaving a more or less easy-to-follow track of their common ancestry (or even no track at all). One of the most famous examples is the vertebrate anterior limb: the human arm, the chicken wing, the dolphin flipper, or the amphibian anterior leg (figure **??**), evolved from the same structure in their common ancestor, to become a grasping, flying, swimming or running organ. Even before evolutionary theories were available, the morphological similitude between body parts had already been noticed by the French naturalist Geoffroy Saint-Hilaire and reformulated by the morphologist Richard Owen who proposed that such structures are *homologous*, *i.e.* they corresponded to:

> *the same organ in different animals under every variety of form and function* (**?**, p379).

Charles Darwin added the evolutionary (temporal) dimension to the notion of homology (**?**, p413):

> [...] *I believe that something more is included* [in our classification, than mere resemblance]*; and that propinquity of descent,—the only known cause of the similarity of organic beings,—is the bond, hidden as it is by various degrees of modification, which is partially revealed to us by our classifications.*

So two characters are considered homologous if they are:

> *derived from the same or equivalent feature of their nearest common ancestor* (**?**).

Sometimes the "*degrees of modification*" are such that the common origin may be difficult to reveal. How to ensure that the structures considered in several animals are homologous, and did not appear several times independently in each species during evolution? Homology is difficult to prove (although statistics can help). In practice, it is inferred based on a concurring set of criteria: for instance a similar position in the organism, same internal structures, similar developments are indications of a common origin. Whereas the existence of only one common criterion suggests homology but does not rule out convergence, the accumulation of common points rules out convergence of the character, because it would be very unlikely that all characters independently appeared several times in each lineage. In the case of the tetrapod anterior limb, the human arm, the chicken wing, the dolphin flipper are located in the same position, they are constituted of the same bones (figure **??**), they harbor the same regulatory gene expression during development (**?**). Thus this body of indices strongly supports the homology of vertebrate anterior limbs.

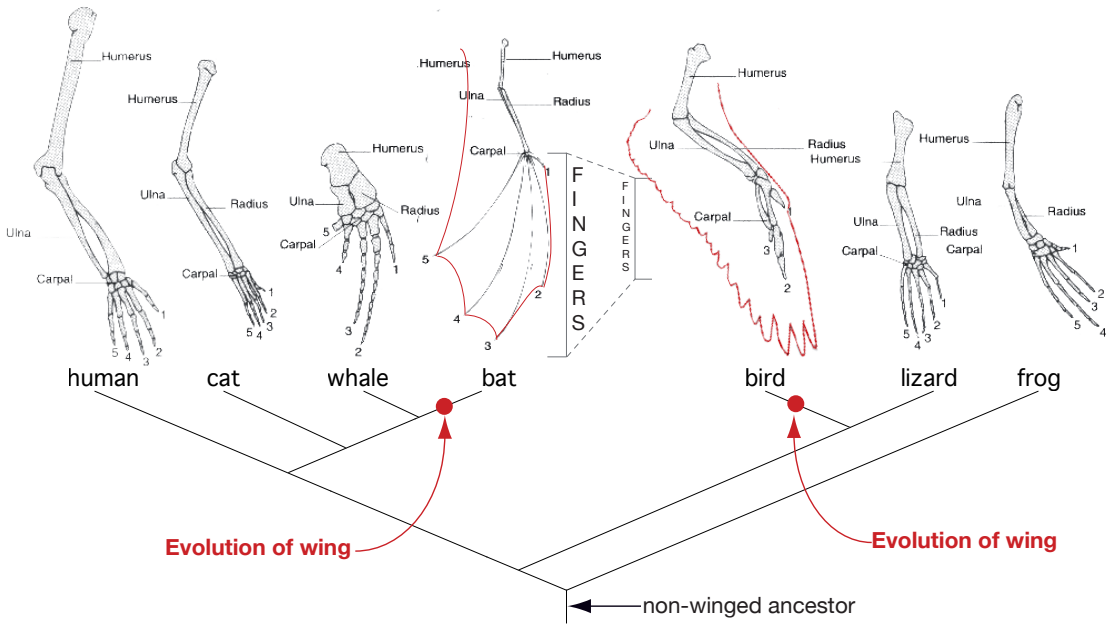


Figure 3: Bone structure of the anterior limb from various tetrapods. They are constituted of the same internal structures, such as bones which are clearly visible on the figure (e.g. carpal, humerus, radius...). They constitute a text-book example of homologous structures. In contrast, the bat and bird wings, drawn in red, are not homologous structures and appeared twice independently during evolution. So although the bat and bird limbs are homologous, their wing structures are not. Modified from (**?**)).

1.1.2 *Convergence*

When comparing a character in different species using a large body of criteria (like position, shape, internal structure, development...), one should be able to detect cases of convergence, for which similar and functionnally redundant structures appeared independently during evolution. If we go back to the example of vertebrate anterior limbs, the bird and bat wings are not considered as homologous. Indeed the flying surface in bird consists of feathers extending all along the arm whereas the flight surface in bats consists of a membrane stretched between the bones of the elongated fingers and the arm (figure **??**, highlighted in red). Accordingly, the 300 million year old reptile-like common ancestor from birds and mammals was most probably non-winged (which is confirmed by fossil record (**?**)).

1.1.3 *Development and homology*

Although homology was originally established for morphological structures, during my PhD, I have studied the homology of developmental processes. Homologous structures are not directly inherited but are recreated anew in each generation from a single cell, through a series of steps during the development of the organism. Then one would guess that the developments of these homologous structures as well as their genetic regulation are homologous as well. This is most often the case. Sometimes, though, the function of some genes involved in the regulation of a developmental process may not have been conserved whereas the genetic network has been. For instance, all insects are organized following an anterio-posterior (A-P) polarity (with the head on one end and the abdomen on the other). However, the *Drosophila bicoid* gene, which is at the top of the genetic hierarchy controlling anterior-posterior (A-P) patterning, is present only in some flies, replacing the function of *Orthodenticle* in other insects, whereas downstream parts of the regulatory network are homologous.

Segmentation (during which the early embryo is divided into a repeating series of segmental structures along the A-P axis) is another example of homologous developmental processes regulated by non-homologous genes. Indeed, although segmentation is ancestral in the insect lineage, the segmental function of the genes *fushi tarazu* (*ftz*) and *even-skipped* (*eve*) appeared in some dipterans (like *Drosophila*), the genes did not have this function in the ancestral insects (as illustrated for *Schistocerca*) (**??**).

Not only can the development of homologous morphological characters be regulated by non-homologous genes, but conversely homologous genes may have

different functions during development. For instance *ftz* mentioned above, is a homeobox gene (the protein contains a homeobox that allows it to bind DNA on specific sequences). As said before, it regulates segmentation in *Drosophila melanogaster*. However, this gene is thought to have displayed a homeotic activity in the arthropod ancestor, *i.e.* it could specify regional identity during development. During insect development, it gained its role in segmentation but also lost its homeotic activity (**?**).

Overall, the homology of development, developmental genes and morphology is not straightforward (**??**). Some have proposed that homology can be better understood by considering the evolution of the whole network regulating a developmental process, rather than considering isolated genes. Then, all parts of the developmental regulatory network do not encounter the same evolutionary dynamics and alterations preferentially affect a subset of a network. From the above example, genes involved in early A-P patterning (*bicoid*, *ftz* or *eve*) are more prone to evolutionary changes than genes involved in later A-P patterning, like *engrailed* which harbors a more conserved expression pattern in all arthropods (**?**).

Similarly, I have investigated during my thesis whether some parts of the molecular determinism of metamorphosis in chordates were better conserved than others, this way revealing a homology of metamorphoses. I focused on the amphioxus and my results are given in chapters **??** and **??**. In chapter **??**, a model of chordate metamorphosis is proposed, based on the differential evolution of parts of the genetic network regulating metamorphosis, with conserved and more rapidly evolving modules (**?**).

1.2 AMPHIOXUS AND CHORDATES IN THE PHYLOGENETIC CONTEXT OF MY PHD

Chordate phylogeny has significantly changed during my thesis. Before summarizing these changes and their impact on my work, I will first briefly introduce phylogenetic concepts and techniques.

1.2.1 *Phylogenetic reconstructions*

The essence of phylogenetic reconstruction is as follows: "relevant" characters are compared between different species; if two species look especially alike (if they share a similar character repertoire), they are probably “close cousins” that diverged quite recently. In contrast, if the "likeness" is less pronounced, then the

species are probably more distantly related. Several parameters come into play: the finding of "relevant" characters and a way of measuring "likeness".

Evolutionary documents

Characters are considered as "relevant" if they are homologous (see section **??** for further details), *i.e.* they were inherited from a common ancestor. This criterion is essential in phylogenetic reconstructions that infer the evolutionary path taken by characters from a common ancestry.

Historically, morphological traits (like the number and position of appendages, their position...) were first compared and used to build phylogenies. At least some of them have the advantage of being found in fossils of extinct species that can be compared to extant ones. However, their number is rather limited and the homology of morphological characters may be difficult to assess. Nowadays, another biological material is available and greatly informative regarding phylogeny: the genome. This source of evolutionary information presents several advantages:

- genes are *directly* and *vertically* inherited from parents to offspring (except in the case of horizontal gene transfer, frequent in prokaryotes but rare in Eukaryotes)
- genes and genomes are very *rich* in information (in some species, up to billions of sites can be considered as characters to approach some questions).
- Several kinds of characters can be used, depending on the question that is asked:
  - the genetic alphabet (A, T, C, G) of nucleic acid
  - the amino acid alphabet of proteins
  - the gene content of a genome
  - the genomic structure (e.g. gene order, exon-intron structure, gaps)
- In our "post-genomic" era, high throughput collection of this genomic material is becoming fast and affordable.
- Accurate models of gene, genome and protein evolution are available.

All these reasons make DNA a valuable evolutionary document (**?**).

Models of phylogenetic recontructions

What about the way of measuring "likeness"? When using morphological characters, the comparative algorithm is usually based on parsimonious reasoning: a phylogenetic tree is constructed that minimizes the number of evolutionary events.

This “parsimonious” tree is considered to be the closest to the real one. This is the reasonning that was used for the inference of the wing-status of the tetrapod ancestor in figure **??**. There are also other more realistic phylogenetic reconstruction algorithms, based on statistical models of character evolution (especially for molecular markers). For instance, likelihood and bayesian methods (**??**) infer the evolution of sequences by estimating the most likely tree or a distribution of the most likely trees of protein or DNA sequences.These methods take into account the properties of the molecular marker (for instance, some site substitutions have more drastic consequences than others) and are thus more accurate than parsimonious methods (**?**). These phylogenetic methods also allow one to infer the sequence of ancestral sequences, this way "resurrecting" sequences, that otherwise are not available to extant evolutionary biologists (**?**).

In order to study the evolution of genes involved in thyroid hormone signaling pathway in chordates (chapters **??** and **??**) and the evolution of the estrogen hormone receptors (chapter **??**), I have used maximum likelihood methods. Moreover, I have also discussed the relevance of ancestral reconstructions, in the particular case of the evolutionary study of steroid hormone receptors (chapter **??**).

1.2.2 *Chordate phylogeny*

Chordate characteristics

The chordate lineage comprises cephalochordates (such as the amphioxus *Branchiostoma lanceolatum*), urochordates (such as the tunicate *Ciona intestinalis*) and vertebrates. The vertebrate group comprises notably agnathans (like lamprey and hagfish), actinopterygian fishes (like *Danio rerio*), mammals (like us), sauropsids (including birds) and amphibians (including anurans like *Xenopus laevis*) (figure **??**).

Chordates are united by several characters:

- a notochord, at least at one stage of their development (hence their name). It is a mesendoderm derivative that stretches along the body between the gut and the nerve tube. The notochord persists for the entire life in cephalochordates and only during larval period in some urochordates.
- a hollow dorsal neural tube.
- a postanal tail
- segmented muscles, at least at one stage.

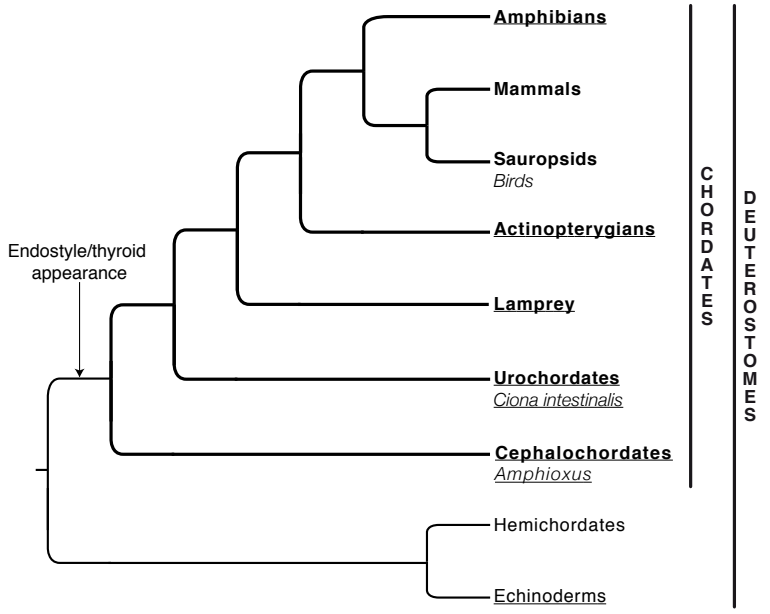


Figure 4: Simplified phylogenetic tree of deuterostomes. Groups where a metamorphosis stage is usually described are underlined. From Paris et al., 2008).

- a thyroid hormone producing gland, called thyroid gland in vertebrates and endostyle in non-vertebrate chordates (I will come back to the endostyle in next chapter, as well as in chapter **??** and in the discussion).

As will be detailed later, there are few cephalochordate species, that have a very similar morphology although having diverged a long time ago (e.g. *Branchiostoma floridae* and *Branchiostoma lanceolatum* diverged 200 million years ago, when the Atlantic ocean opened). In contrast, Urochordates, that comprise more than 2500 species (**?**), display a very large diversity in morphologies, ecological niches, developments, genome sizes... There are 3 major urochordate groups, ascidians, thaliaceans and appendicularians. The major group of ascidians has a biphasic life style with a pelagic and free-swimming "tadpole-like" larva and a benthic "sac-like" adult. Appendicularians (also called larvaceans) retain tadpole structures and stay free-swimming as adults (see section **??** for further details) while thaliaceans have a pelagic lifestyle as adults. This extensive diversity has made urochordate phylogeny rather unclear although appendicularians are considered as most basal (**?**).

Vertebrates are the most successful chordate group, in terms of species numbers (about 45,000 (**?**)). They comprise the gnathostomes and the agnathans, although the monophyly of the latter is still debated (**?**). Vertebrates exploit much more ecological niches than other chordates: fresh and marine waters, terrestrial habitats (that they

are the only animals to have colonized, with insects), air, all in a wide range of temperatures, latitudes, etc. Accordingly, many adaptations to specific ecological niches have evolved (from wings to amphibious life-style through regulation of internal temperature, etc.). What contributed to the success of Vertebrates is thought to be neural crest cells, at the origin of head structures (e.g. advanced sense organs) that enabled vertebrates to become more active predators (**?**) (see next section).

Vagaries of chordate phylogeny

All the shared characters above are widely accepted and help establish chordate monophyly. However, and this is of special interest here, the nature of the most basal chordate has recently caused an intense debate. At the beginning of my PhD, urochordates were considered as sister group to all other chordates and cephalochordates were considered as sister groups to vertebrates (e.g. (**??**)) (figure **??**a). This was supported both by morphology (e.g. segmented musculature) and molecular phylogeny (**?** and references therein)). **?**, soon followed by **?**, contributed to the first phylogenomic support of the basal branching of amphioxus within chordates. Using the recently released genome sequence of the urochordate *Oikopleura dioica* (**?**) (in addition to *C.Intestinalis* genome (**?**)), they suggested that urochordates are the sister group to vertebrates (forming the "olfactores" group (**??**) (figure **??**b). The recent release of the amphioxus genome (**?**), as well as all subsequent independent phylogenomic studies (*e.g.* (**??**)) supported this topology.

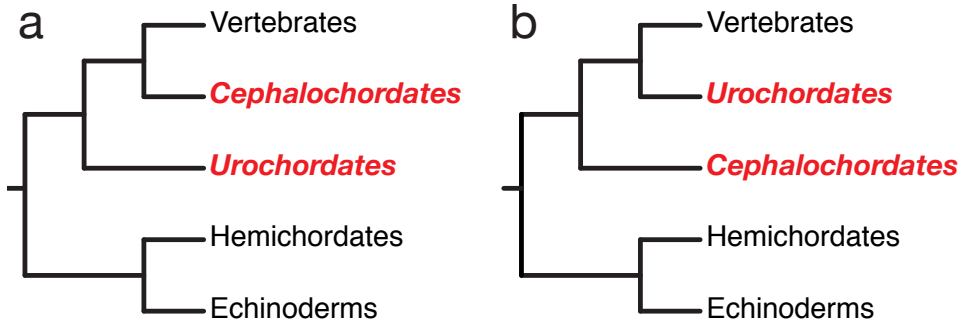


Figure 5: Ancient (**a**) and recent (**b**) views of chordate phylogeny.

The erroneous phylogeny, placing amphioxus at the base of the vertebrate tree, was mainly due to the fast evolution of urochordates, Because of this fast rate of evolution, urochordate sequences harbor high rates of evolution, which biases phylogenetic recontructions (**?**) by the so-called long branch attraction bias: sequences harboring a long branch are artificially grouped together at the base of a

phylogenetic tree. Because of their long branch, urochordate sequences are pushed to the base of the chordate tree whereas slowly-evolving amphioxus and vertebrate sequences are grouped together. Then the increase in the number of available sequences and the improvement of phylogenetic methods allowed to select only slowly evolving genes (that are thus less prone to long branch attraction) and to accurately take into account fast evolving sequences (e.g. (**??**)). In accordance with their fast evolving genome, urochordates have a derived development and body plan (e.g. the larva has no gut, the Hox gene clustering was lost (**?**)), and, as depicted before, display various body shapes. As will be developed later urochordates are thus considered as a highly derived animal model (**?**) whereas cephalochordates are considered as having kept more ancestral features (**?**).

This new phylogeny sheds light on the evolution of some vertebrate-specific structures: the placodes. These structures are supposed to have largely contributed to the evolutionary novelties of the vertebrate head that allowed vertebrates (and especially jawed vertebrates) to adapt to a predator life style, from a filter feeding ancestor (**?**). Placodes are ectodermic cell types that give rise to components of several vertebrate-specific cranial structures like the otic vesicles, adenohypophysis or lens. All placodal cells derive from a primordium located at the neural plate border, as exemplified by the expression pattern of genes involved in placode formation (e.g. *Six1/2* or *Eya1*). Cells with placodal characteristics have been reported in several urochordate species and have been proposed to be homologous to vertebrate olfactory, otic and pituitary placodes: they are sensory cells located at the same position, they are marked by the expression of similar developmental genes like *eya*, *pitx* and *six* (**???**). However urochordate placode-like structures are not migratory. In addition, the developmental control by several genes appeared specifically in vertebrates (**?**). Then it is generally thought that placodes and neural crest cells evolved from pre-existing sensory cells present in the chordate ancestor and were recruited in vertebrate and in urochordates to form specific structures: e.g. the atrial siphons, used to filter-feed in urochordates, and the different neural crest and placodal derivatives in vertebrates (**??**).

Neural crest cells are multipotent ectoderm derivatives and migratory cells that arise at the border of the neural plate and that are at the origin of many vertebrate-specific head structures like cephalic sensory ganglia or part of the cranium. While no migrating neural crest cells have been discovered in amphioxus (**?**), neural crest-like cells were reported in the ascidian tunicate *Ecteinascidia turbinate*: they are migratory cells forming at the border of the neural plate, expressing some neural crest markers and giving rise to pigment cells, like some vertebrate neural crest cells (**?**). These data suggested a more ancient origin of neural crest cells than

previously thought, at the origin of urochordates and vertebrates. However, these cells are not considered as neural crest cells *per se* because they do not display the full multipotency of vertebrate structures and because they do not express the complete genetic regulatory network (**?**). Rather, it is thought that neural crest cells did not suddenly appear in the vertebrate ancestor but rather evolved form preexisting "rudiment" structures present in the chordate ancestor, and that can be found in non-vertebrate chordates, that gained a sub-set of neural crest properties (like migration) in the common ancestor of vertebrates and urochordates (**?**), and was further elaborated during vertebrate evolution by gene co-option (**?**).

Whole genome duplications at the base of the vertebrate clade

It was proposed decades ago that whole genome duplications appeared at the base of the vertebrate clade (**?**). Duplication would be a generous source of genetic material "ready to use", *i.e.* complete functional genes with their open reading frame, regulatory regions and biochemical context. One copy of a duplicated gene could keep the function of the original gene when the other one, freed from any selection pressure linked to the original gene function, could experience another fate by rapidly accumulating mutations and maybe acquire new functions (**?**). In contrast, it seems more difficult for an already functional single-copy gene to acquire a new function without altering its present one. To quote Ohno "*natural selection merely modified, while redundancy created*" and he imagined that gene duplications played an important role in major evolutionary transitions whereas classical Darwinian gradualism was restricted to microevolutionary events. It comes as no surprise that such a model induced a lot of debates and that many authors tried to challenge it, especially when complete genome sequence provided for the first time the data to tackle this issue.

Known as "2R" for "2 Rounds of whole genome duplications", Ohno's hypothesis postulates that 2 events of genome duplication occurred at the origin of Vertebrates. Ohno's theory was ignored by many biologists for two decades until genome sequencing and the appearance of new computational methods allowed confirming this hypothesis (**??**). The exact dating of the duplication events during vertebrate evolution is still unclear: did the two duplications occur before lamprey split (**?**), or one before and one after (**?**)? The upcoming release of the lamprey genome should help solve this issue. One central question is still to know if gene duplications were instrumental in the appearance of the new morphological traits of vertebrates. The split of amphioxus and tunicates before genome duplication surely make both of them interesting models to study the effect of duplication in the evolution of gene function (**?**).

1.2.3 *Deuterostome phylogeny*

Together with Xenoturbellida (like *Xenoturbella bocki*) and the ambulacraria, composed of the echinoderms (like the sea urchin *Strongylocentrotys purpuratus*) and the hemichordates (like *Saccoglossus kowalevskii*), chordates constitute the deuterostomes group (**?**) (figure **??**). Synapomorphic developmental features are for instance the enterocoelous formation of the body cavity or the secondary emergence of the mouth, after the anus. However, the chordate mouth was also proposed to have either independently appeared or to have shifted (**?**) or to have independently evolved (**??**).

The monophyly of echinoderms and hemichordates has important consequences for our understanding of deuterostome evolution. Indeed, hemichordates share some characteristics (like gill slits) with chordates (hence their name). Consequently the structures, previously thought to be derived in hemichordates and chordates, are ancestral in deuterostomes and were secondarily lost in echinoderms. In addition, the inferred deuterostome ancestor is more "amphioxus-like" than previously expected.

The deuterostome monophyly has recently been challenged by a phylogenetic study that proposed that ambulacraria and Xenoturbellida were more closely related to protostomes than to chordates (**??**). Although these results were not statistically supported and need further confirmation, they may lead to interesting conclusions regarding the evolution of bilaterians because they place chordates at the base of bilaterians (with amphioxus at the base of chordates). This would then suggest that the characters common to chordates and hemichordates are basal in bilaterians, which would lead to the conclusion that the very bilaterian ancestor is more "amphioxus-like" than previously expected. Although this hypothesis is very elusive, it once more contradicts that the rampant gradualist view of the evolution of chordates and vertebrates.

1.3 PALEONTOLOGICAL CONTEXT OF CHORDATE ORIGINS

Pikaia gracilens (figure **??**a) may be the most famous proposed fossil of a basal chordate. It is an elongated, amphioxus-like animal, like most fossils of early chordates, and the most important discovery of the Burgess shales according to **?**. It used to be the oldest chordate fossil (about 505 million years old, in the middle Cambrian) (**?**). Although it possesses important chordate characteristics (e.g. a notochord), it is considered to belong to a stem chordate group (*i.e.* an extinct

group of animals more closely related to chordates than to any other extant group), since for instance its pharynx is located in the dorsal position, and not the ventral position, as in all other chordates (**?**). Some paleontologists even consider Pikaia as being a protostome-like fossil (**?**). More ancient and better admitted chordate specimen, like *Cathaymyrus diadexus* (about 10 million years older) (**?**) or the 520 million year old urochordate *Shankouclava* (**?**) have stolen some of Pikaia's superb. More importantly, the discovery of older vertebrate fossils (see below), necessarily pushed earlier the origins of chordates from Early Cambrian.

Although the Lower Cambrian animal *Haikouella* (figure **??**b), from Chengjiang, has previously been interpreted as a basal deuterostome (**?**), it is more probably related to vertebrates since it displays clear vertebrate structures like an eye, gills or possibly sensor capsules (**??**). The Agnathan-like 530 million year old *Haikouichthys* and *Myllokunmingia* are less polemic basal vertebrates, as they exhibit neural crest cell derivatives such as visceral arch skeletons (and maybe a cartilaginous skull). They display an elongated body shape, gills, segmented muscle blocks and a notochord (figure **??**c).

Some particular fossils, that do not display an "amphioxus-like" shape, have also been proposed to be ancestral chordates. The stylophorans are small calcite-plate invertebrate fossils that extend from middle cambrian (~510 million years ago) to upper carboniferous (~ 300 million years ago) (figure **??**). According to the calcichordate theory, some of these fossils exhibit structures interpreted as chordate-specific characters. In particular a series of openings have been interpreted as gill slits (e.g. (**?**)). The calcichordate theory posits that the deuterostome ancestor displayed a bilateral symmetry. During evolution it began to lay down on its right side, which gave rise to an asymmetric ancestral chordate, in which most of the structures located on the right side (especially gill slits) had degenerated (this change of orientation is called dexhiothetism). Accordingly, some stylophoran fossils, the cornutes, that would have descended from this chordate ancestor, display several asymmetries, among which "gill slits" only on the left side. The subsequent chordate descendants would have regained symmetry (giving rise to stylophoran mitrates), some of which would have been at the origin of vertebrates, urochordates and cephalochordates. Amphioxus larvae (see section **??**) would have kept this ancestral asymmetry. However, this evolutionary scenario is very hypothetical for several reasons, notably the interpretation of the chordate-like structures (among which gill slits) is very controversial. An alternative and well accepted scenario proposes that stylophorans are derived echinoderms (**?**). The stylophoran fossils will be further (and briefly) discussed in chapter **??**.

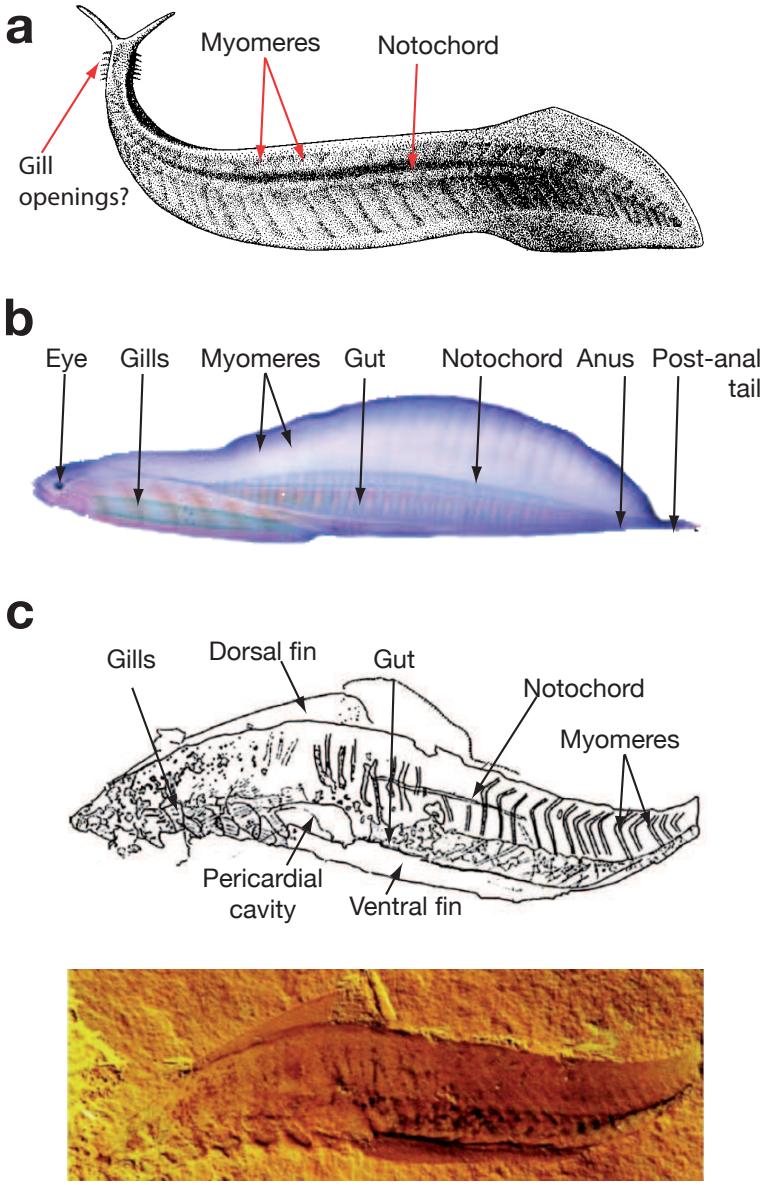


Figure 6: Comparative drawings of some chordates fossils. **a.** Comparative drawing of larvae and juveniles from diverse Chordate groups. **a.** *Pikaia gracilens*, from the Burgess shales. The presence of a notochord clearly indicates a chordate origin. From (**?**). **b.** *Haikouella lanceolatum*, from Chengjiang, has a large compressed body as well as several vertebrate-specific characters, such as eyes, muscle fibers in the myomeres, and a notochord extending into a postanal tail. Modified from (**?**). **c.** Fossil *Myllokunmingia fengjiaoa* from Chengjiang. Different chordate and vertebrate specific organs are indicated. The upper view corresponds to the interpretation of the fossil (lower view).

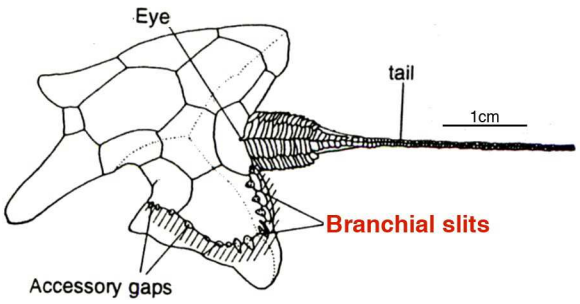


Figure 7: Reconstruction of the cornute *Ceracystis perneri* in a dorsal view, as interpreted by R.P.S. Jefferies. Modified from (**?**).

1.4 THE ANIMAL MODEL AMPHIOXUS

A large part of my PhD was dedicated to understanding amphioxus post-embryonic development. As explained in section **??**, amphioxus is now considered to branch at the base of the chordate tree (figure **??**).

Introduction to the amphioxus

Amphioxus is a ~ 5 cm long marine benthic animal (figure **??**) with a fish-like shape. It can be found in tempered and tropical shallow seas all over the world. As an adult, this filter feeder lives for a few years buried in the sand most of the time (figure **??**a), except at night fall, on a spawning night, when it swims towards the surface, spreads its seeds before passively settling back down to the sand. The larva, on the contrary, is planktonic (figure **??** illustrates development).

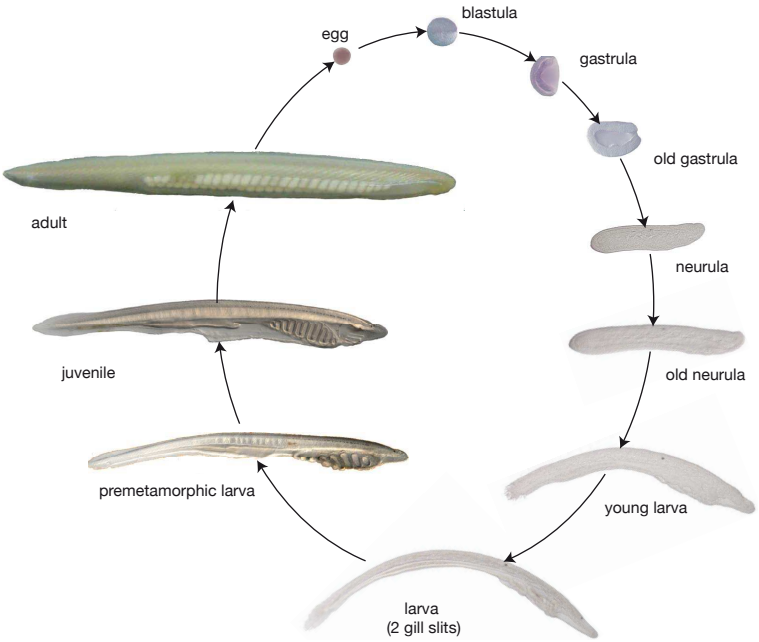


Figure 8: Development and life cycle of *Branchiostoma floridae*. Pictures were modified from (**??**) or courtesy of Florent Campo Paysaa.

The term amphioxus refers to the cephalochordates that comprise only about 30 species, with very close morphologies. The term "cephalochordate" indicates that the notochord (a chordate synapomorphy, see section **??**) goes from *head* (=*cephalon*) to tail (figure **??**b). Whereas the notochord disappears during mammalian development, it is permanent in both larvae and adults. Three genera consitute the cephalochordates: *Branchiostoma*, *Epigonichthys* and *Asymmetron*.

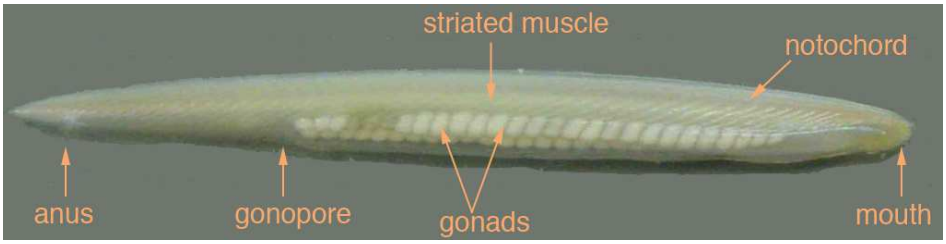


Figure 9: Pictures of adult amphioxus. **a.** The amphioxus is buried in the sand, except the most anterior part, in which the mouth is free to filter water to catch plankton. Courtesy of Hector Escriva. **b.** Lateral view of an amphioxus adult. The series of gonads on each side of the body are clearly visible. Modified from http://www.bioportal.jp/blog/mt-static/FileUpload/pics/amphioxus.jpg.

Amphioxus phylogeny

Branchiostoma is the largest group and comprise all the amphioxus species that are studied on a regular basis. There are mainly 3 model species, the Japanese species *Branchiostoma belcheri*, the European species *Branchiostoma lanceolatum* and the East American species *Branchiostoma floridae*. The European species is the historical model (**?**), whereas the East American species is the current most advanced model, whose genome was recently sequenced (**??**). Although *Branchiostoma* species diverged a long time ago (between 100 and 200 millions years ago (**??**)), they have conserved a very similar morphology. During my thesis I have mostly worked on the species from Florida, and to a lesser extent on the European species. In the manuscript, I will use the term amphioxus to refer to *Branchiostoma*.

The main morphological difference between *Branchiostoma* and other groups is the gonadal arrangement: there are two series of gonads on each side of the body in *Branchiostoma* (figure **??**b), whereas there is only one on the right side of *Epigonichthys* and *Asymmetron*. From morphology, *Epigonichthys* and *Asymmetron* were considered as a unique group and named *Epigonichthys* (with *Asymmetron* as a synonym). The sequencing of the mitochondrial genome from several cephalochor-

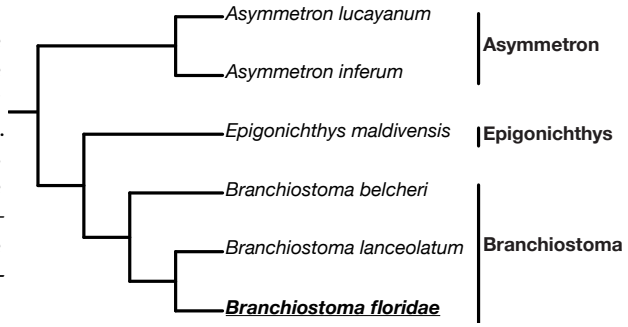


Figure 10: Phylogenetic relationships of some cephalochordate species. The three genus *Branchiostoma*, *Epigonichthys* and *Asymmetron* are represented. Note that the asymmetric genus *Epigonichthys* and *Asymmetron* are not monophyletic. The species *Branchiostoma floridae*, on which I have more specifically worked, is underlined. Modified from (**?**)

date species brought new insight into the evolution of cephalochordates. Indeed, molecular phylogenetic studies have confirmed the monophyly of *Branchiostoma* and have proposed the paraphyly of the asymmetric species: *Branchiostoma* would be closely related to *Epigonichthys*, with *Asymmetron* as an outgroup (**??**). This is interesting regarding the evolution of amphioxus metamorphosis because *Branchiostoma* larvae are asymmetric like *Epigonichthys* or *Asymmetron* and become more symmetric during metamorphosis. Considering the phylogenetic position of *Branchiostoma* in the cephalochordate tree, parsimony suggests that adult asymmetry was ancestral in the Cephalochordata and secondarily lost in the *Branchiostoma* genera, whereas larval asymmetry would have been conserved. However, data are too scarce data for asymmetric species to negate or confirm this hypothesis (for instance, development has been described only in *Branchiostoma*). This point will be discussed in chapter **??**.

Amphioxus as an animal model of evo-devo

Amphioxus is a rather young animal model for developmental genetics. Although the European species used to be the most studied one in zoological experiments (**??**), the East American species was promoted by Linda and Nicholas Holland about 15 years ago (**?**), as an ideal model to study the early evolution of chordates and vertebrates (see below). However, the European species, from the input of Hector Escriva, is making a comeback because of its technical advantages (the animal, and especially the spawning habits, are easier to deal in this species (**??**)). Amphioxus used to be promoted for its previously thought phylogenetic position, as the sister group to vertebrates. However, its current position, as a basal chordate (see previous section **??** for further details) makes it also very interesting because its study allows to get insights into chordate origins. Moreover, amphioxus displays other characteristics which make it a unique model to study the evolution of chordates, and even vertebrates. Its morphology has not changed much since chordates origins about 550 million years ago. Indeed, fossils of primitive chordates like *Pikaia* are

"amphioxus-like" (see section **??** for further details) and amphioxus genome has a slow evolving rate (**?**). In contrast and as explained earlier, the vertebrate sister-group urochordates are morphologically and developmentally much derived (even within the urochordate subphylum, the numerous genera of Urochordates display extensive morphological and developmental differences). So all things considered, this special phylogenetic situation makes amphioxus interesting in two aspects: it is basal *and* slowly evolving (especially morphologically, see section **??**) (**?**). This makes amphioxus a relevant model to study the evolution of chordates (**??**).

Notably, the current phylogenetic position of amphioxus supposes that many conserved features that it shares with vertebrates are ancestral within chordates and that the chordate ancestor was more "vertebrate-like" than previously expected. This will be discussed in this manuscript in the context of the evolution of metamorphosis.

2 INTRODUCTION TO METAMORPHOSIS

The inaugural use of the Greek word "metamorphosis" is attributed to the Latin Ovide (**?**). It is generally associated with the notion of shape change ("meta" meaning change and "morphe" form) of an individual. In biology it refers to the larva-to-juvenile transformation like tadpole-to-frog or caterpillar-to-butterfly transitions (see the following section **??**) whereas the lay audience has applied the term to fantasy or litterature.

2.1 METAMORPHOSIS FOR THE LAY AUDIENCE

Although metamorphosis refers to a *transformation*, the drastic changes that occur at metamorphosis and also the sudden and seemingly magical origin of the modifications make it more than a simple transformation: the association between the pre- and post-metamorphic forms is not straightforward and the extensiveness of the modifications renders mechanisms underlying these changes difficult to conceive. Then supernatural causes are more appealing, but also more convenient because they shrug off the explanation problem to focus on the result of the metamorphosis: two *different* organisms in one. This concept has also been extensively used for less realistic transformations, sometimes physical, sometimes metaphorical, some of which are briefly introduced below.

The very evocative potential of metamorphosis was amply illustrated in countless films and books. Horror or Sci Fi movies have overused the concept to feed audiences with the inexhaustible stories of gentle people transforming into horrible monsters or ugly monsters transforming into even more atrocious other monsters. The notion of metamorphosis has also been illustrated less "literally". Often below the surface of morphological alterations, it can also be a metaphor for inexplicable and important changes that reunite two apparently separate parts of the same individual. Among the plethora of examples, one of the most famous films dealing with metamorphosis, called *The fly* directed by David Cronenberg, tells the story of a man slowly transforming into a fly. The film is as much about slimy and disgusting flesh exhibition as about internal duality of a man, transfigured into his physical appearance. Among books, *The Strange case of Dr Jekyll and Mr. Hyde* by R.L. Stevenson again uses the metaphor of metamorphosis to discuss the

conflict between two irreconcilable parts of a person. Similarly, Kafka centers his novel "*Die Verwandlung*" ("*The Metamorphosis*") around the metamorphosis of a typical average man called Gregor into a giant cockroach. First physical, the transformations slowly reaches his mind and Gregor gradually loses his humanity (especially his dignity) until death. A notable difference with Stevenson's novel is the irreversibility of the metamorphosis: Gregor never goes back to the way he used to be, like metamorphosed animals never "re-metamorphose" to their larval form. Kafka's *Metamorphosis* is not about internal conflict of a person, but rather about how a person reacts to the changes in his relatives' look. Overall each time, metamorphosis does not "distort" a person, but rather "develops" the true and hidden personality, which is hard to accept for both him/her and the others. This notion of "transformation", this denial of permanence is central in the book Metamorphosis by Ovid, in which a series of independent short stories tell the story of different event more or less dealing with metamorphosis. This notion may remind of one of the ecological advantages of animal metamorphosis, that allows one individual to colonize different niches by adapting each time to their particular environmental characteristics (e.g. water for an amphibian tadpole and earth for adult frog).

Even in the geological vocabulary, metamorphosis is used to designate important changes in the rock structure, due to gigantic forces (the term metamorphism is rather used). This notion of super-powerful force that cannot be controled is always underlying when talking about metamorphosis. Even in biology, as will be discussed in chapter **??**, genetic parameters "program" the occurrence of metamorphosis (it *has* to happen), as well as external conditions influence the onset and unfolding of metamorphosis (see chapter **??**), like the moon triggers werewolf transformation.

2.2 METAMORPHOSIS FOR A BIOLOGIST

2.2.1 *Definition*

In biology, metamorphosis has a more restricted definition, based on developmental considerations. Although there are still debates about the exact definition of metamorphosis (**?**), a consensus embraces the notion of drastic morphological changes associated with an ecological transition. In the following chapters, I will define metamorphosis as a short post-embryonic developmental stage allowing a larva to become a juvenile/adult and during which important morphological, ecological and metabolic modifications occur. This definition will be applied to

animals (figure **??**). Plant development won't be discussed here. Metamorphosis is not applicable to unicellular organisms like Bacteria (and even multicellular Bacteria) or Archaea, for which there is no development.

A direct corollary to the occurrence of a metamorphosis stage during development is the existence of larval and juvenile/adult stages. Between the juvenile and the adult occurs sexual maturation (for instance in amphibians, amphioxus, sea urchin, oyster). The latter, during which morphological changes can also occur, should not be confused with metamorphosis because it is not triggered by the same molecular mechanism (see chapter **??**) and most often does not involve an ecological transformation. However, in some species and especially in insects, the post-metamorphic individual is already mature (thus there is no juvenile stage).

2.2.2 *Metamorphosis is widespread among metazoans.*

If one were to pick up any animal and guess whether it metamorphoses during its life time, one should bet on the affirmative answer. Indeed, most species metamorphose, including many insects, mollusks, annelids, cnidarians, sea urchins, some fishes, frogs or sea squirts (of course this list is not exhaustive). The plankton is basically filled with small aquatic larvae from non-insect invertebrates (**??**). The diversity of larval forms can compare to the diversity of names given to the different larvae: e.g. planula larva in some cnidarians, pluteus larva in some echinoderms, caterpillar in some insects, maggot in others, trochophor larva in some mollusks and annelids, tadpole in some amphibians and urochordates, ammocoete in some lampreys, nauplius larva in some crustaceans...

This diversity in metamorphoses extends from full metamorphosis (with extensive remodeling) to no metamorphosis. Indeed, some animals (like us) directly develop into an individual that looks like the adult and apparently do not go through the morphological, ecological and metabolic modifications characteristic of metamorphosis. Actually, there are many ecological strategies among animals, with a morphological remodeling more or less pronounced, ecological shifts more or less drastic, feeding or non feeding larvae... depending on each species situation (see following section). Some species also undergo several morphological and ecological metamorphic transitions (like insects or crustaceans), although the last transition is usually the most dramatic one, in addition to leading to settlement before transformation into the pre-reproductive form (juvenile). Figure **??** summarizes some typical life cycles.

As can be seen on figure **??**, life cycles of metazoans are very diverse. They can be rather "simple" as it is the case for the amphioxus *B.floridae* or the sea

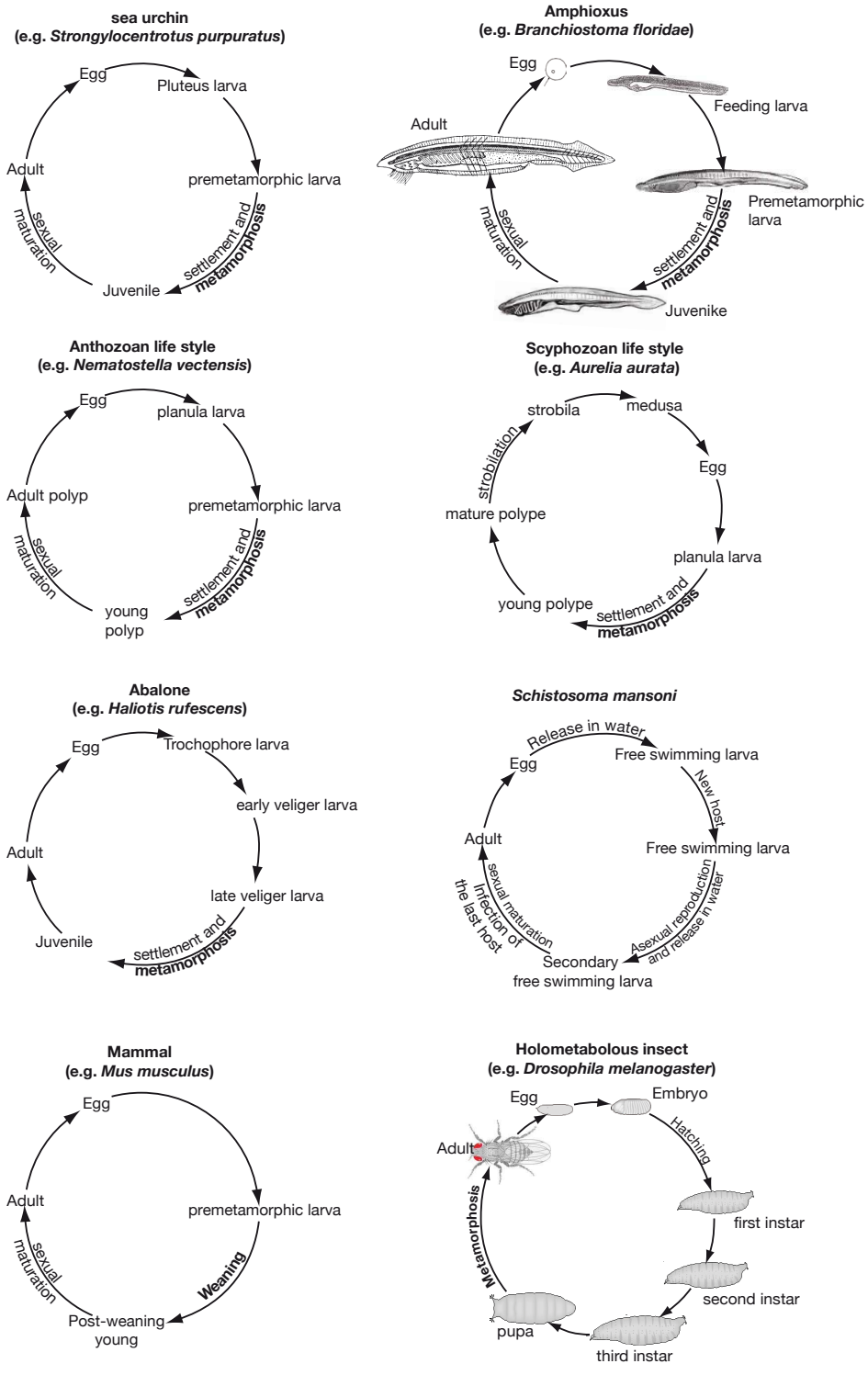


Figure 11: Life cycles of various metazoans, notably chordates. The amphioxus drawings were modified from (**??**) and www.cosmovisions.com/cephalocordes.htm.).

urchin *P.lividus*: a small egg (from the fertilization of gametes released in water) develops into a larva that grows until metamorphosis to the juvenile stage (figure **??**). The latter grows until sexual maturation, reproduces and cycle starts again with another individual.

Other life cycles are more complicated. For instance, cnidarians display a wide diversity of life cycles (figure **??**), with, for instance, the addition of the medusa phase in scyphozoans (the strobilation phase has been sometimes considered as a metamorphosis in some articles, although it is more a phase of asexual reproduction than a metamorphic phase (**??**).

Crustaceans also have a large diversity of life cycles. Some species have a direct development while others harbor very diverse larval forms that follow one another through transitions much like metamorphoses (e.g. with nauplius, zoea, cypris larvae that may hardly look like the adult).

Insects, for which the molecular determinism of metamorphosis has been especially well studied, exhibit a large diversity of postembryonic development strategies. Apterygotes have a direct development whereas the development of all pterygotes is indirect, with a more or less drastic metamorphose. The hemimetabolous species (like the grasshopper) hatch as larvae that resemble the adults and develop through a series of gradual changes until adult condition. On the other hand holometalous insects (like flies, beatles or butterflies) go through much a single more drastic and abrupt changes (see figure **??**). Although the successive moltings that occur during insect growth may be seen as morphological and ecological transitions, according to our definition, only the last one leading to the juvenile stage may be considered as a *bona fide* metamorphosis (figure **??**).

The parasitic life style can also complicate life cycles. For instance the life cyle of Platyhelminthes depends on the different hosts that will be infected. *Shistosoma mansoni* infects human at a sexually mature stage, where it spawns. Early embryos are discharged out of the host and hatch in water as free-swimming later embryos. Then they penetrate a second host (a snail) in which they asexually proliferate and are relased into water. They can also infect another intermediate host or infect the final one in which they spawn (figure **??**).

Differences of postembryonic developmental phases can be highlighted by comparing distantly related species. For instance, both life cycles in amphioxus and sea urchin are rather simple, with a premetamorphic larva grown from the egg and a postmetamorphic juvenile, like-looking the adult. However the morphological modifications have nothing in common. As will be detailed in section **??**, amphioxus metamorphosis is rather mild with the modification of left-right asymmetry in the anterior part of the animal. On the contrary, during sea urchin only a few larval

cells give rise to the adult during metamorphosis whereas all other larval tissues degenerate. The very bauplan of the animal changes (from bilateral symmetry to pentaradial symmetry).

This variability of metamorphic strategies is also conspicuous between more closely related species: even within the amphibian lineage, anurans go through an extensive remodelling while urodeles have a much milder, or even no apparent metamorphosis (**?**). Even animals from the same species (also amphibians) may follow different developmental strategies (**?**). More generally, there are many amphibian as well as many insect species, qualified as paedomorphic, that have lost an indirect development and reproduce as larvae. This change in the timing of development, and more precisely on the relative time of appearance and rate of development for characters already present in ancestors (heterochrony (**?**)) has been proposed to be of significant importance in the course of the evolution of phenotypes. Different types of amphibian paedomorphy will be described in chapter **??** Moreover, how neoteny may evolve in amphibians will be discussed in chapter **??**.

2.2.3 *Selective advantage of indirect development*

The frequent occurrence of larval stages raises the question of the selective advantage of going through the fragile and energetically costly metamorphosis stage. In other words, why has metamorphosis been retained during evolution? In addition to the fact that it may have been evolutionary difficult to lose an indirect life style once it appeared, a multiphasic life cycle offers several advantages, all of which are related to ecological shifts: lowering of competition, better dispersal, allocation of life tasks to different life phases.

The division of the animal life style into the larval and adult ones minimizes intraspecific competition (different location, food...). Most frequently in aquatic species, the larva is pelagic whereas the adult is benthic (like echinoderms, amphioxus, sea squirts, crustaceans or teleost fishes) but the animals may also live in fresh *vs* marine water like lampreys, aquatic *vs* terrestrial environments like some amphibians or dragonflies, or also in two different terrestrial ecosystems like butterflies (whose larva is herbivorous and whose adult is nectarivorous).

In parallel to lowering ecological competition by living in a different environment than adults, larvae may be seen as dispersal tools, especially in the aquatic environment (**?**), whereas the adult harbors a reproduction apparatus. Indeed, they are often free-swimming pelagic microscopic individuals that are produced in large quantities, released in the environment and passively transported over

the course of water currents, so that they can explore an extended area. On the opposite, the adult is bigger (a few centimeters), benthic and has limited ability to scatter (e.g. in ascidians, amphioxus, crustaceans or echinoderms). For instance, the larval amphioxus is covered with cilia and swims continuously towards the light whereas, although the adult is mobile, it lives for several years buried in sand most of the time (only post-metamorphic animals have been reported in sand samples (**?**)).

The larval stage can also be a feeding period, during which the larva grows and accumulates energy. For example, the basal vertebrate lamprey grows as a larva for months until it has reached a certain size. Only then, and under favorable environmental conditions will it metamorphose (**?**).

Among all these briefly described but diverse metamorphoses, I have focused on chordates.

2.3 METAMORPHOSIS IN CHORDATES

Metamorphosis is widespread among chordates (figure **??**). The most classic example is the common frog (like *Rana catesbeiana*), on which an aquatic and herbivore tadpole is transformed into a terrestrial and carnivorous juvenile. In this respect, larvae and adults display specific structures adapted to their environment: for instance the tadpole has gill slits, a swimming tail, a lateral line (for motion detection under water), eyes (and corresponding pigment) adapted for aquatic vision or a cement gland which secretes a mucus that allows the larva to adhere on a support. Loss of the gill slits, shrinkage and disappearance of the tail, development of limbs and lungs, intestine and eye modifications but also switch from larval to adult haemoglobin type are among the most obvious events that can be described (for reviews, see (**????**)). Although a classical textbook example, the tadpole to frog transition is not representative of the variety of metamorphoses one can observe in chordates (figure **??**). In the next paragraphs, I will briefly introduce metamorphic strategies adopted by some chordate species, as illustrations of the morphological and ecological variability that one can found in this group.

The other tetrapods (mammals and sauropsids) are generally considered not to metamorphose. Most actinopterygian fishes have a metamorphosis characterized by an ecological shift crucial for the geographic dispersal of the species and therefore have been very intensively investigated for economic purposes (**?**). For example, most marine fish larvae (such as the flatfishes discussed above) are planktonic and transform into juveniles that either swim in the water column or are benthic (Leis et al., 2006). In freshwater populations, metamorphosis is also morphologically

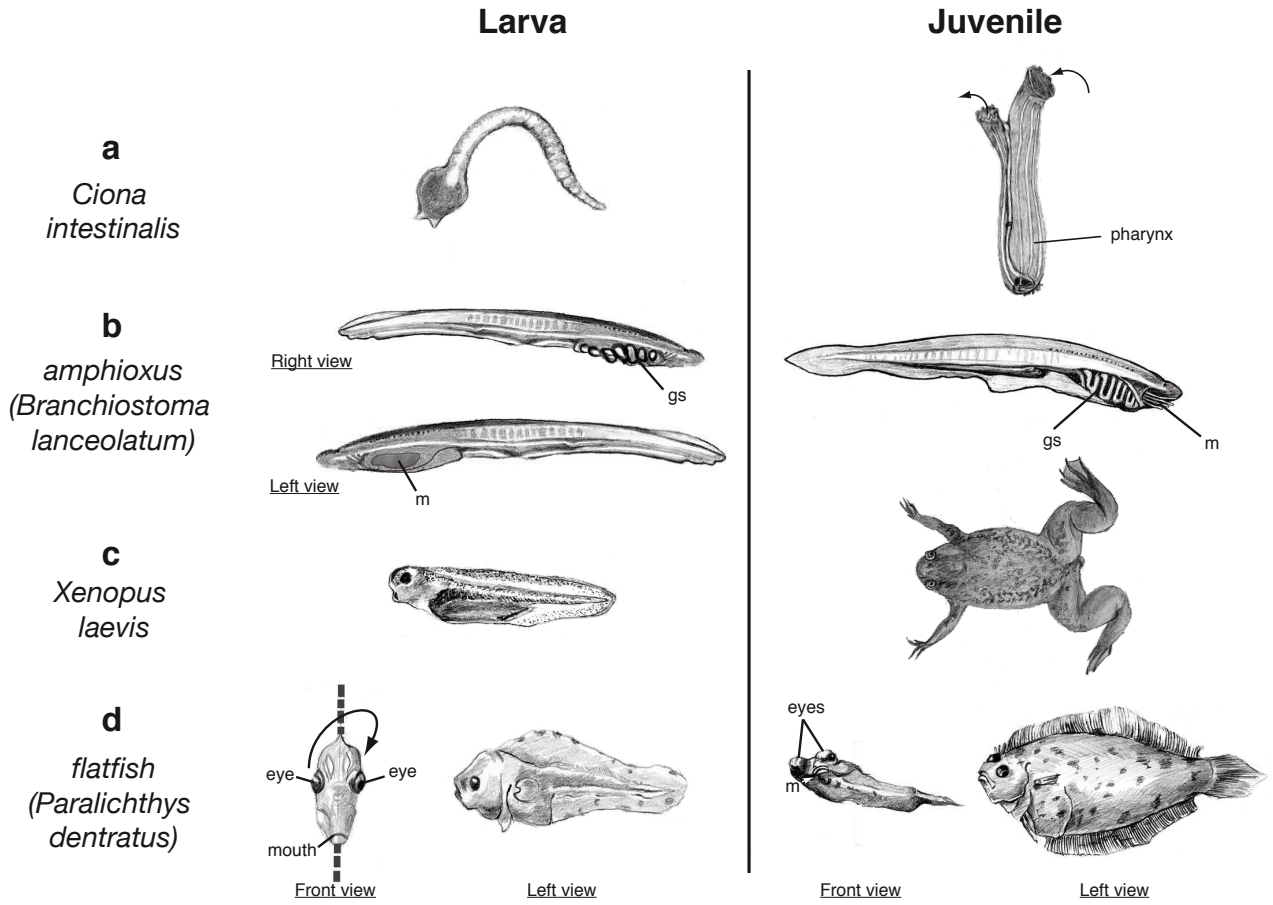


Figure 12: Comparative drawing of larvae and juveniles from diverse Chordate groups. **a.** Among remarkable events, the *Ciona intestinalis* larva changes its body plan. The two arrows indicate the water flow through the two siphons in the adult. **b.** The amphioxus larva displays asymmetric features (e.g. the mouth is located on the left side of the body and the round gill slits are only on the sight side of the body) and gets more symmetric during metamorphosis (e.g. the mouth migrates to a more ventral part and the gill slits get a U shape and migrate to be on both sides). The amphioxus larva is presented from both right and left views. **c.** The tail of *Xenopus laevis* regresses when limbs develop. (d) In the flatfish *Paralichthys dentratus*, the left eye of the symmetric flatfish larva migrates to the right side of the body. The larva and the adult are represented from two angles (front view + left view). The dashed line represents the larval plan of symmetry. An arrow indicates the migration itinerary of the right eye. gs: gill slits, m: mouth. From (**?**).

obvious, as seen in the alteration of the zebrafish fins and gut (**?**), or salmonid smoltification (**?**).

The larva from the basal vertebrate lamprey dwells in the silt of freshwater streams and filter feeds on primary detritus. Larval period lasts several years after which lampreys metamorphose into adults, some of which are (sometimes parasitic) predators. During metamorphosis, several structures are remodeled: the eyes and the gut epithelium are modified, the larval kidney is replaced with a juvenile one, the mouth develops into a specific sucker apparatus and the endostyle is profoundly remodeled into a thyroid gland (**?**). The members of the sister group hagfish apparently have a direct development, although their post-embryonic development remains to be fully studied (**?**).

Outside vertebrates, dramatic morphological remodeling occurs in some urochordates where the very bauplan of the organism is modified between larvae and adults, like in ascidians (figure **??**). Ascidians are the most studied and the largest taxonomic group in urochordates, with *Ciona intestinalis* as a genomic and developmental model system. The larva is pelagic and "tadpole-like" with a typical chordate body plan (a tail -"uro"- with a notochord, a dorsal nerve tube, striated bilateral segmented musculature), whereas the adult is sessile with a different body plan. *Ciona intestinalis*, for instance, has a "sac-like" body with no tail, an hypertrophied pharynx and two siphons allowing water flux (figure **??**). This remodeling is so drastic that some zoologists, including Linnaeus, did not relate the adult to the larva and classified the former as a mollusk. Notably, other ascidian species undergo a slightly less dramatic transition as they do not lose their tails during metamorphosis (**?**). The other two classes of tunicates, Larvacea and Thaliacea, are both planktonic in the adult stage. In addition, in Larvaceans, the adult keeps a tadpole-like body plan and can thus be viewed as paedomorphic. In Thaliacea, the larva does not settle before rapidly metamorphosing into a planktonic sac-like adult (**?**).

Amphioxus metamorphosis

Amphioxus (and more precisely *Branchiostoma*) was described at metamorphosis by **?** in very precise detail more than a century ago. The animal goes into an ecological transition from pelagos to benthos. Accordingly, the larval body is covered with cilia that are used for locomotion and are replaced by muscular swimming at metamorphosis (**??**). However, the most conspicuous change is the transition from an asymmetric larva to a more symmetric juvenile. Indeed the larval pharynx is highly asymmetric: the mouth is on the left side whereas the single gill slit row is located on the right side (figure **??**a). During metamorphosis, the animal will get

more symmetric: in parallel to the appearance of two metapleural folds on each side of the body, that will recover the pharynx and gill slits in the juvenile, a second gill slit appears above the first row (figure **??**b). Soon after, all initially round gills slits get a U shape (figure **??**b) and the first gill slit row migrates towards the left side of the pharynx while the mouth migrates in opposite direction towards a more ventral and anterior position (figures **??**b and c. A close-up of the pharynx from the figure **??**b is given figure **??**). Cirris develop around the mouth that migrates from the left side to an antero-ventral part (figure **??**) (**??**). The gut also changes, with the appearance of a protuberance just posterior to the pharynx (figure **??**c), called gut diverticulum (**?**). All these morphological modifications have been taken into account during the experiments I conducted and participated to, and that are reported in chapters **??**, **??** and in the article reported in the annexe (page **??**).

This asymmetric organization of gill slit development has been proposed to support the calcichordate theory (see section **??**), according to which the ancestral chordate (that would have looked like cornute fossils) was asymmetric and contained gill slits only on one side of the body. Amphioxus larvae would then have kept the ancestral asymmetry. During evolution, symmetry would have been regained, the way amphioxus larvae become symmetric during metamorphosis (e.g. (**?**), discussed in (**?**)). The cornute fossils are briefly discussed in chapter **??**.

Morphologically less conspicuously, the organ endostyle also undergoes profound remodeling from an anterior position and a rather compact shape (figure **??**a) to a ventral position and an elongated shape (figure **??**c). Moreover, an amphioxus-specific organ, called the club-shaped gland, located posteriorly to the endostyle, disappears. The fate of this organ has been debated for decades (does it really disappear? Is it reincorporated into surrounding tissues?) and will be discussed in page **??** (article included in the annexe). There are probably also metabolic and physiological modifications (for instance, is endostyle activity changing during its metamorphosis?) but they have not been looked for yet. At the end of metamorphosis, a much more symmetric juvenile starts its benthic life buried in sand.

Some other amphioxus species from the genus *Asymmetron* and *Epygonichthys* (**?**) stay asymmetric even as adults. As mentioned at the end of the previous section, the post-embryonic development of these species is still unresolved yet, although some modifications of the symmetry may occur between larval adult stages (for instance, the adult mouth is in the midline).

During my PhD, I have sought into the regulatory mechanisms underlying the evolution of metamorphosis in chordates. Metamorphosis has been molecularly well characterized only in amphibians and more precisely in anurans, and to a

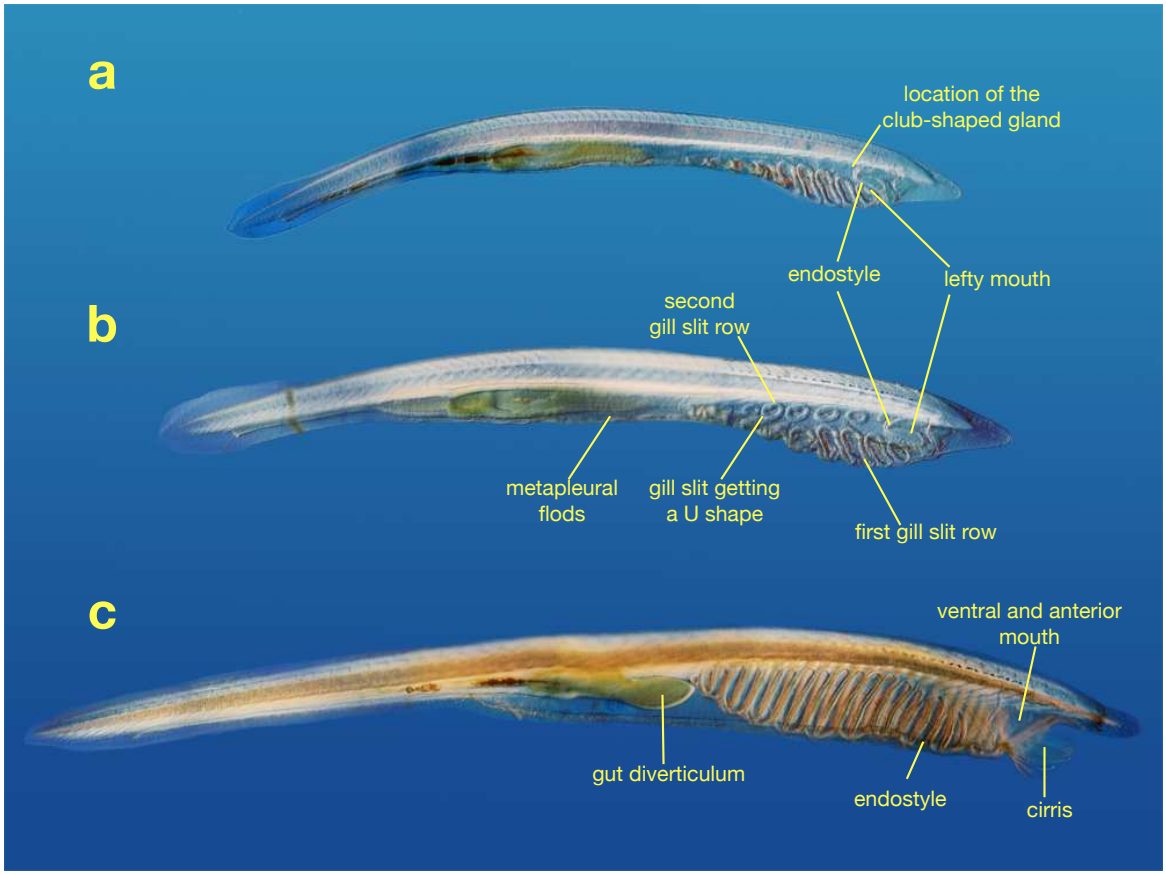


Figure 13: Pictures of amphioxus larvae before (**a**), in the middle of (**b**), and after metamorphosis (**c**). **a.** The premetamorphic amphioxus larva is asymmetric: its mouth is on the left side of the pharynx and the single gill slit row on the right side. The endostyle is clearly visible. **b.** During metamorphosis the animal gets more symmetric. The details of morphological remodelings are given in the main text. **c.** A fully metamorphosed juvenile. Anterior extremity is orientated towards the right. The pictures are not scaled one to the other. This picture appeared in (**??**).

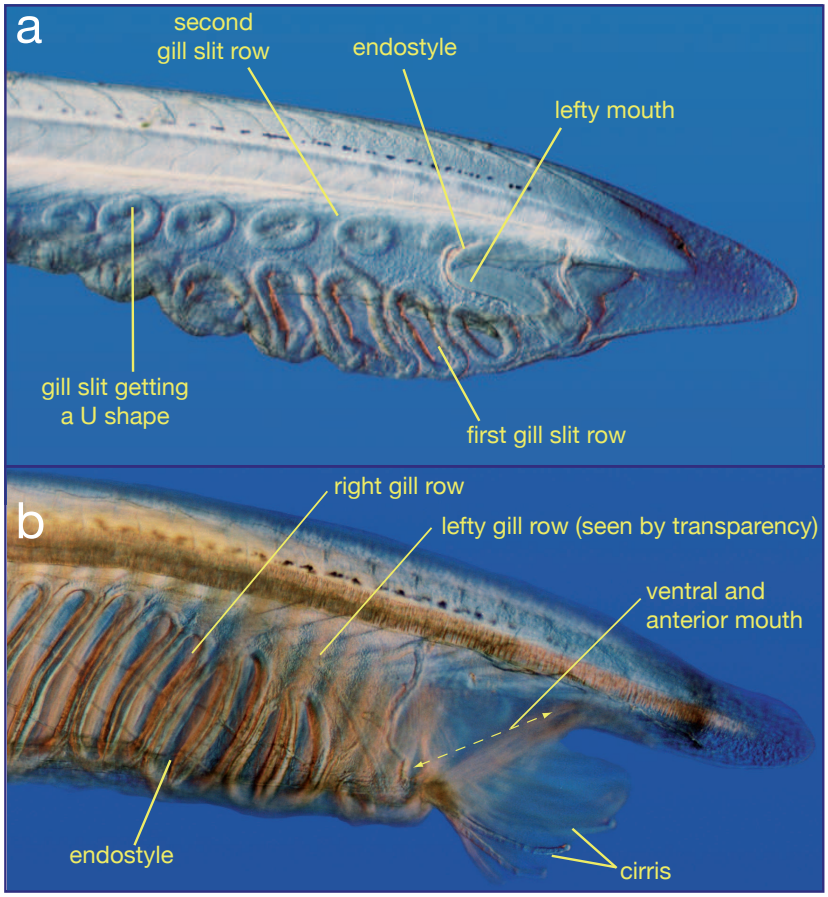


Figure 14: Close-up of the amphioxus pharynx during metamorphosis from figures **??**b (**a**) and **??**c (**b**). The large opening of the mouth, behind the endostyle, is indicated with a dashed line. The pictures are not scaled one to the other.

lesser extent in teleost fishes. The knowledge about the molecular determinism of metamorphosis will be briefly presented in the next chapter of this manuscript.

3

REGULATION OF METAMORPHOSIS IN CHORDATES

Vertebrate metamorphosis has been molecularly well characterized only in anurans, probably for reasons that are historical (**?**), anthropocentric (anurans are the metamorphosing animals closest to humans), epistemological (amphibian metamorphosis is quite spectacular and appealing for the biologist), as well as technical (tadpoles are big and easy to breed in a lab). Thus I will briefly present the knowledge accumulated in anurans, and to a lesser extent in actinopterygian fishes, since most of the experiments dissecting the molecular determinism of metamorphosis have been performed in these groups.

3.1 REGULATION OF AMPHIBIAN METAMORPHOSIS

3.1.1 *Amphibian metamorphosis is triggered by thyroid hormones*

How a horse can transform a tadpole into a frog.

In anurans the main trigger of metamorphosis is known to be endocrine, through production of the precursor thyroid hormone (TH) T_4 (thyroxine) in the thyroid gland and its subsequent transformation into the more active derivative T_3 (triiodothyronine) (figure **??**), although some T_3 is also produced in the thyroid gland. T_3 and T_4 are iodinated tyrosine derivatives: basically they consist of 2 tyrosines coupled and iodinated 3 and 4 times, respectively.

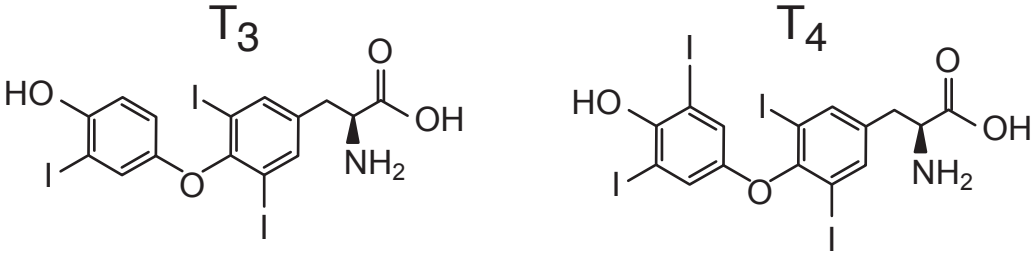


Figure 15: T_3 and T_4 are iodinated tyrosine derivatives

The discovery that THs induce metamorphosis was done by J.F. Gudernatsch almost a century ago. He was studying "the influence of various organic extracts on the development of fish (*Belone*, *Gobius* etc.) and amphibian (*Rana*) eggs" (**?**). Although his experiments on fish did not give conclusive results (because boats could not deliver fresh eggs fast enough), he was more satisfied with the results obtained from frogs. Indeed, as he was feeding tadpoles of *Rana temporaria* and *Rana esculenta* with pieces of horse, calf, rabbit, pigs or even cats, he noticed that tadpoles fed with thyroid extracts metamorphosed earlier and gave rise to "fully developed frogs, though dwarfs in size". So exogenous thyroid extracts prematurely induced metamorphosis (**?**). But was this effect of any biological significance or artificial? In other words, is the active compound of horse thyroid necessary for frog metamorphosis?

The answer came a few years after Gudernatch's discovery, when metamorphosis was reported to be inhibited upon thyroid removal (**?**) and recovered upon subsequent feeding of the larvae with thyroid preparations. These experiments showed that thyroid gland activity is necessary and sufficient to induce frog metamorphosis. Chemical inhibition of TH production by various compounds has also demonstrated its efficiency in preventing metamorphosis and creating giant tadpoles (e.g. (**?**)).

The active molecule present in the thyroid was isolated at the end of the same decade (**?**) from hog thyroid extracts (figure **??**) and identified as an "iodine-containing compound", which was originally called thyroxin (T_4). Much later its more active but less abundant derivative T_3 was identified (**???**).

Tissue-specific metamorphic response to TH induction

Making a tadpole into a frog requires an extensive remodelling mostly related to the water-to-land ecological shift: from loss of the tail to development of limbs, through repositionning of the eyes, remodeling of the intestine, keratinization of the skin, modifications of the central nervous signal, ossification of the previously cartilaginous skull, adaptation of the visual system to fit new light dispersion properties. Then, the range of the molecular response to TH signal is as wide as *de novo* morphogenesis, apoptosis, switching of the hemoglobin subtype, of the visual pigment (from porphyropsin to thodopsin), production of keratin...

Consequently, the response to the hormone is local, tissue-specific and tissue-autonomous. For instance a limited and localized application of TH to a tadpole will induce only the surrouding tisues to undergo metamorphosis (**?**). Moreover exposition of exogenous T_3 to isolated larval tissues will induce the resorption of the tail and the morphogenesis of the hind limb bud. Convincingly, if a larval

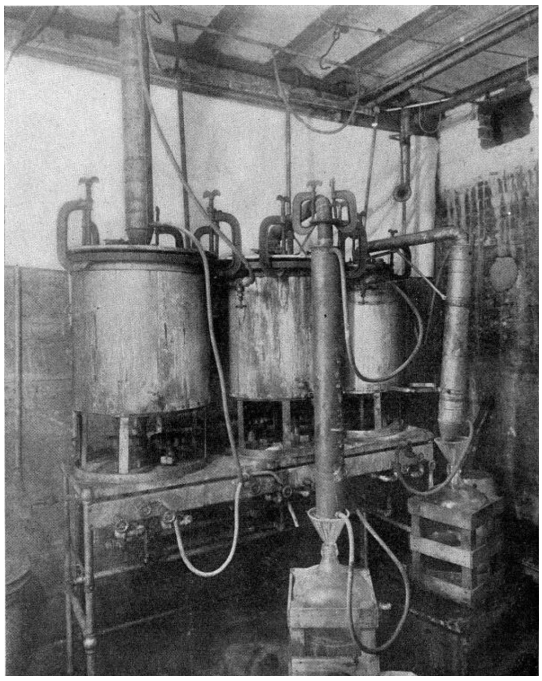


Figure 16: The three tanks in which about 3000 kilograms of hog thyroid glands were treated for the isolation of thyroxin (**?**).

organ (for instance an eye) is transplanted in another part of the organism, whose metamorphic fate is different from the grafted organ (for instance the tail), both structures will "normally" metamorphose (the larval eye will become an adult one whereas the tail will resorb) (reported in (**?**).

3.1.2 *A rapid overview of the TH signalling pathway*

A peak of TH production induces metamorphosis

Experiments on metamorphic effects of exogenous THs have their biological justification in the occurrence of a peak of TH production at the beginning of metamorphosis, which is the key physiological trigger (figure **??**)((**?**) for a review, see (**?**)). TH levels reach a maximum during the climax *i.e.* when the major morphological changes (e.g. tail regression) occur. If the highest TH levels are prevented, late metamorphic events will be prevented (**?**).

THs induce modifications of transcription, through the activation of their receptor TR

What are the molecular consequences of TH signal and how do they lead to metamorphosis? More than 40 years ago, Jamshed Tata proposed that THs induce

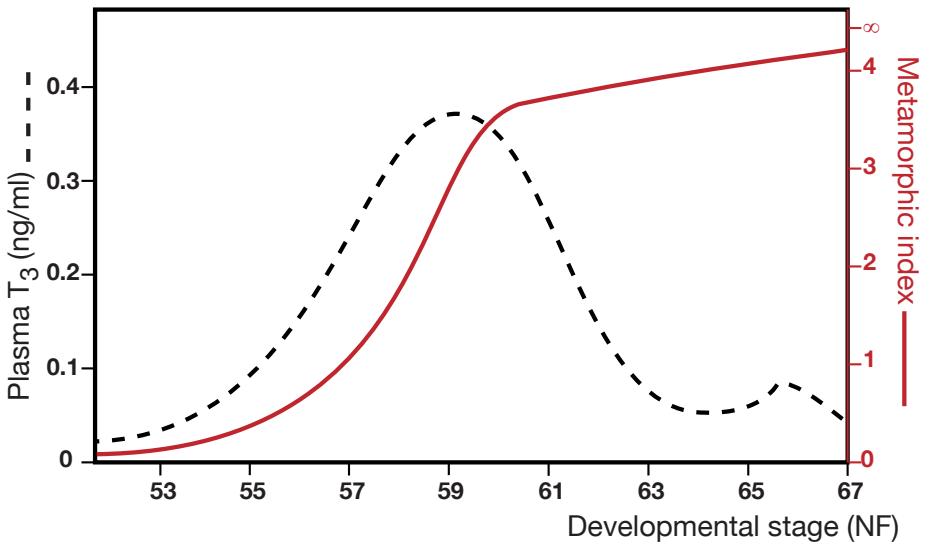


Figure 17: There is a singe transition form larval to adult form in amphibian metamorphosis. The idealized dashed curve depicts the single burst of T_3 (ng/ml) just before the transition from the larval to the adult state in *Xenopus laevis* (red curve). Developmental stages on the abscissa (NF) refer to (**?**). Metamorphic index is ratio of hind limb: tail length. Adapted from (Tata, livre).

mRNA expression (**??**). It is now well established that the signal represented by a peak of TH production is mediated by its two receptors, the thyroid hormone receptors (TRs) TRα and TRβ, issued from a vertebrate duplication. More precisely, they are TH-regulated transcription factor that will modulate the expression of target genes upon activation by TH binding, subsequently leading to the morphological remodeling characteristic of metamorphosis. Notably, TRβ is the main mediator of TH signal, at the onset of metamorphosis, while TRα is involved in earlier development (see section **??**. Chapter **??** will be dedicated to the more of action of TRs).

TH metabolism

TH synthesis in the thyroid gland takes place in the follicular cells called thyrocytes and in the lumen enclosed by the thyrocytes. This lumen is full of the protein thyroglobulin. As previously stated (figure **??**), THs are not gene-encoded peptides but iodinated tyrosine derivatives produced through a metabolic pathway. T_4 is the most abundant precursor that will be subsequently degraded into other THs. Thus TH production can be schematized into iodine fixation of tyrosine that are subsequently to produce T_4 and subsequent degradation into TH derivatives. More precisely, TH production can be decomposed into several intermediate steps, each of them illustrated in figure **??** and briefly depicted hereafter (a more complete

description of TH signalling pathway is given in chapter **??**): (1) Iodine first needs to be concentrated by the specific transmembrane Sodium/Iodine Symporter (SIS), (2) then it is added to tyrosines from the matrix protein thyroglobulin by the Thyroid Peroxidase (TPO) to form monoiodotyrosines or diiodotyrosines (if one or two iodines have been added to a tyrosine, respectively). (3) The iodinated tyrosines are coupled by the TPO to form a TH (mostly T_4 but also T_3). (4) The newly synthesized TH is then separated from the matrix protein, (5) and transported in the blood by various carriers such as TBG (thyroid binding globulin) or TTR (transthyretin). (6) In peripheral tissues, T_4 is deiodinated into the more active form T_3. (7) Both THs can be converted into less active forms, that will be subsequently degraded. (8) There are also alternative TH pathways, like sulfatation and glucuronidation, mostly resulting in production of inactive T_3 derivatives, targeted for degradation. (9) However, some alternative active forms are also produced, like the deaminated form of T_3, TRIAC, which is a very potent, but very unstable TH derivative (see (**?**) (figures **??** and **??**). (10) Finally, within the cell, upon thyroactive compound binding, TR will modulate the expression of target genes. The direct synthesis of T_3 in the thyroid gland, along with the peripheral deiodination of T_4, is an important source of active hormone since cellular uptake is faster for T_3 than for T_4 (**?**).

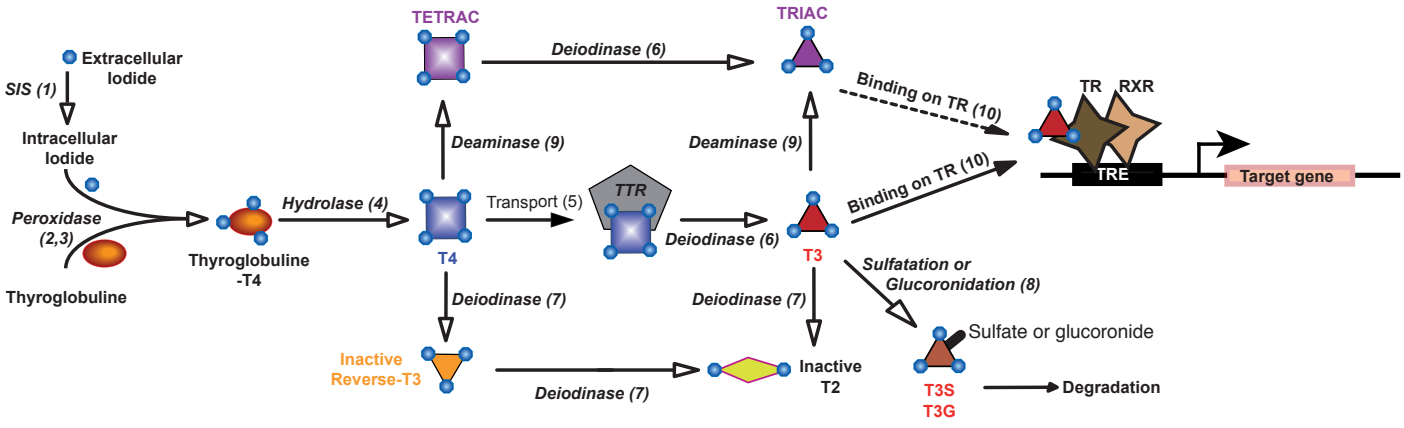


Figure 18: Thyroid hormone signaling pathway. The different steps from iodine concentration to TH binding to TR are indicated. Numbers located between brackets refer to the main text. From (**?**).

3.1.3 *Diversity of TH derivatives*

THs and other iodinated tyrosine derivatives were discovered as thyroid gland products (hence their name). They have been detected in the plasma of all verte-

brates that have been tested (for detailed values, see (**?**)). In mammals and amphibians the two main thyroid products are the stock hormone T_4, or l-thyroxine, and the more active T_3, or 3,3',5-triiodo-l-thyronine (iodinated 4 and 3 times respectively) (figure **??**). However, other kinds of related molecules exist depending on the number and positions of the iodines. Among such compounds, 3,5-T_2, 3,3'-T_2 or 3,3',5'-triiodo-l-thyronine (Reverse-T_3, see figure **??**) are mildly active TH derivatives in mouse (**?**) (see table **??**).

Decarboxylated or deaminated analogues, like TRIAC (triiodothyroacetic acid) or TETRAC (tetraiodothyroacetic acid) have also been described in mammals and in anurans (figure **??**) (**?**). Notably TRIAC, the T_3 deaminated derivative, displays strong thermogenic effects and has a higher affinity for TR than T_3 (**?**). A few decades ago, when TH signaling was not as well characterized as nowadays, TRIAC was even proposed to be the active compound in amphibians, with T_3 as a precursor, like T_4 is the T_3 precursor (**?**). However this hypothesis was rapidly abandoned because TRIAC and TETRAC are kept at low levels (levels of circulating TETRAC are less than 1 % of T_4 ones levels whereas TRIAC levels are about 1/20 of T_3 ones). The low TRIAC and TETRAC levels are due to their high affinity for degradation enzymes like deiodinases (**??**).

Recently, the potential importance of deaminated TH derivatives has been proposed. Indeed, the deiodinased and deaminated T_3 derivative 3-iodothyronamine (T_1AM) was detected in mice and shown to induce rapid and non-genomic effects on metabolism and cardiac rhythm (**?**), by binding to the G protein-coupled trace amine receptor (see section **??**).

Overall, the general biological action of TH derivatives remains vague. In other chordates no exhaustive repertoire is available. This is mainly explained by anthropocentric reasons - and the tendency to look for what is similar/different compared to us; see (**?**) for a discussion of this aspect - but also by the fact that such comprehensive analyses are much more demanding than simple two-by-two comparisons with vertebrates. The evolutionary relevance of an expanded iodinated tyrosine derivatives repertoire, larger than simply the THs T_3 and T_4 in chordates, should not be neglected.

3.1.4 *Modulations of TH action: how to find solutions to the problem of a fluctuating environment?*

Interactions of the TH signalling pathway with the environment

Although THs are the obligate signal for initiating and completing metamorphosis, other hormones and factors, like environmental signals, can influence the onset and

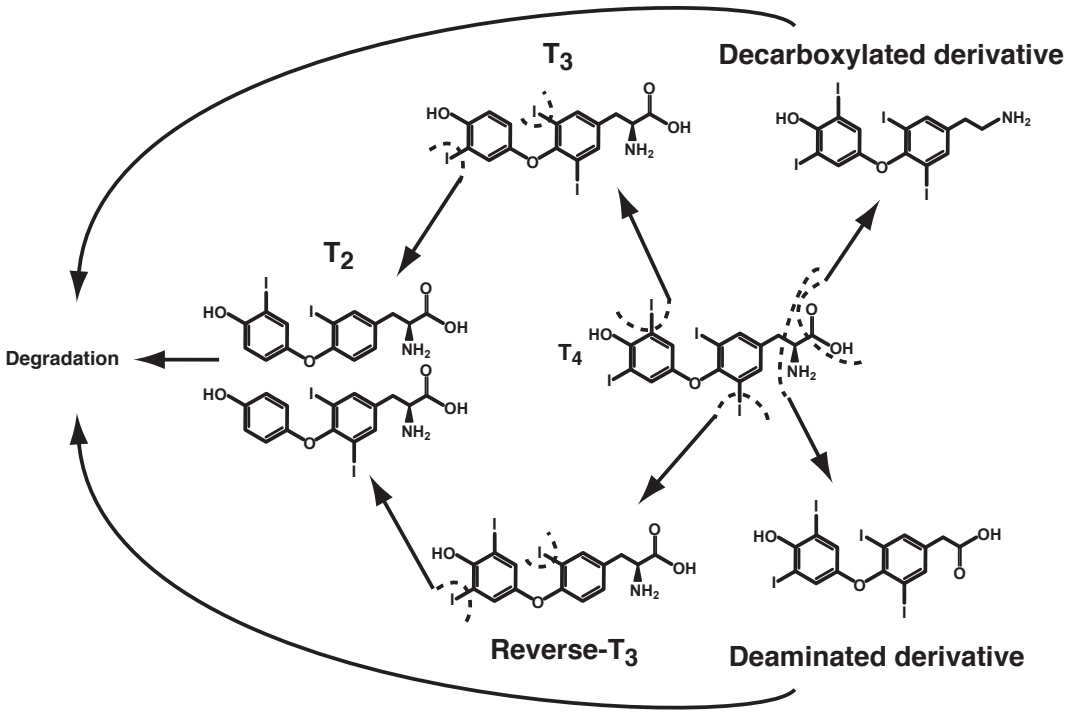


Figure 19: Schematic representation of various TH derivatives. The dashed lines indicate the different alterations to T_4.).

progression of metamorphosis. For instance tadpoles put in a stressful situation (like reduction of water volumes or proximity to a predator) will metamorphose faster (**?**). This ability to respond adequately to variable external stimuli has probably been selected in amphibians living in unstable environment, so that they adapt efficiently to unfavorable conditions: for instance, a tadpole may escape more easily from a drying pond.

The link between the environment and the initiation of metamorphosis could be traced back to the tadpole's neuroendocrine system (hypothalamus and pituitary gland) through the intermediate of hypothalamic TRH (thyrotropin releasing hormone), CRF (corticotropin releasing hormone) and the TSH (thyroid stimulating hormone) from the pituitary (figure **??**. The role of TRH in the regulation of metamorphosis is still unclear). For water volume reduction, increases in levels of the hypothalamus-product CRH (corticotropin-releasing hormone) were shown to occur and to be necessary for precocious metamorphosis (**?**).

Overall, when environmental parameters (e.g. water volume or food availability) modulate the triggering of metamorphosis, it is always through alteration of the TH production which can thus be viewed as the key endogenous inducer of the metamorphosis process ((**?????**), (**?**) for review). The hypothalamo-pituitary axis may then be seen as a monitoring system that can transduce environmental signals

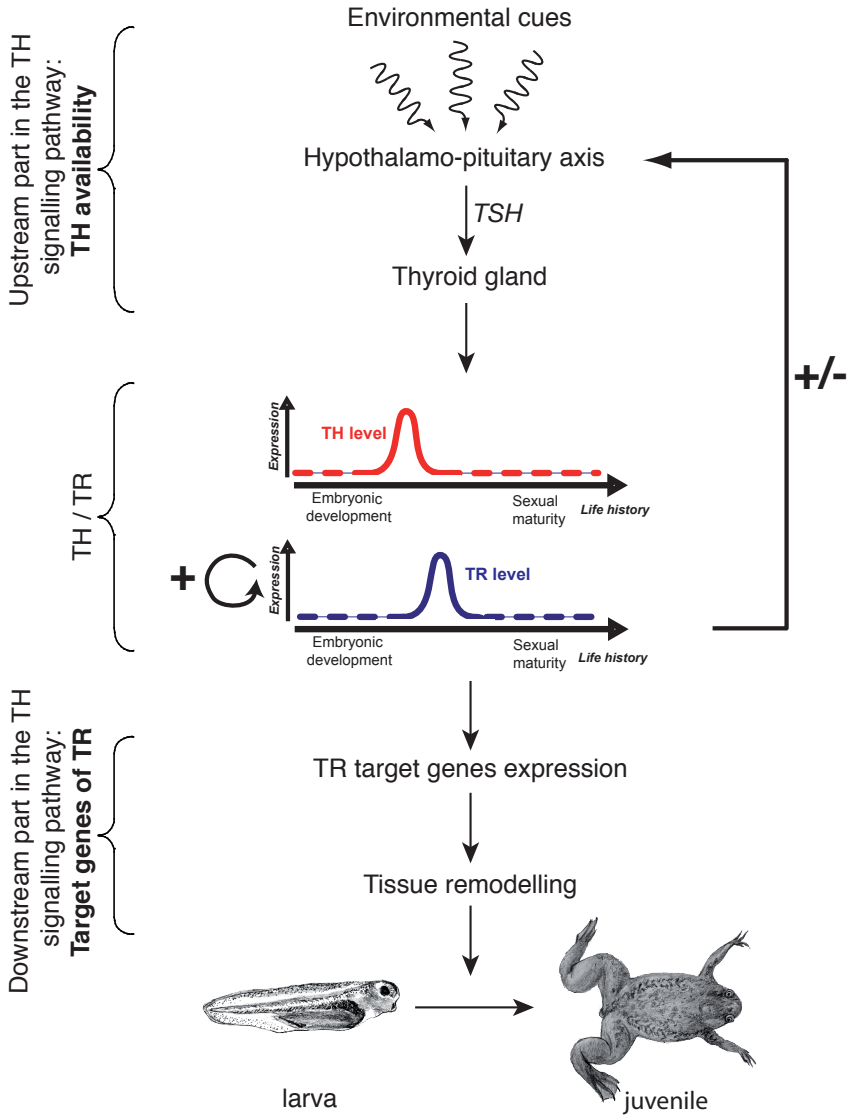


Figure 20: Schematic structure of TH cascade in the anuran organism during metamorphosis. Three parts are highlighted: the couple TR / TH at the center of the pathway is upregulated at the onset of metamorphosis. More upstream in the pathway TH synthesis takes place, and more downstream, the expression of TR target genes is modified upon TR activation. Since TR (more precisely the gene TRβ) itself is one of its own target genes, its autoactivation is critical for the triggering of the climax of metamorphosis. TR also regulates the production of TH, thus contributing to the climax.

into a physiological response that will adapt the developmental course to the environmental stimulus.

Competence

As **?** first reported, metamorphosis can be shifted earlier in time by adding exogenous T_3, which boils down to simulating a precocious TH rise (figure **??**). This responsiveness to hormonal information, called competency, is acquired early during development, when TRβ expression starts (figure **??**), much earlier than the actual beginning of metamorphosis (**?**), when circulating TH levels rise (starting at stage 44-45). At stages 47-48, competency is better acquired, figure **??**). Thus the competency phase is correlated with TR expression (figure **??**): in the absence of TRs, TH signal cannot be mediated and only when the main mediators of TH signals are expressed can THs induce metamorphosis (see (**?**) for an exhaustive discussion of the significance of TH/TR expression at the onset of metamorphosis).

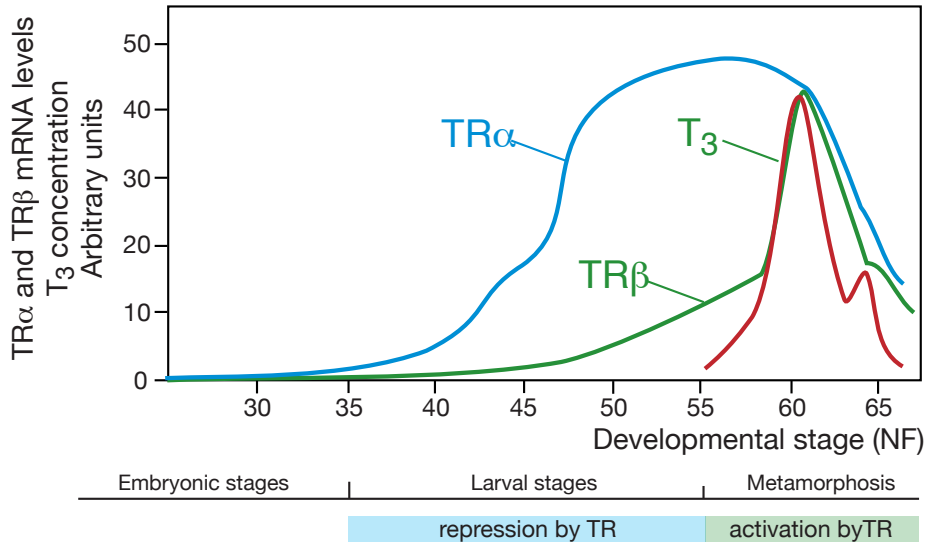


Figure 21: T_3 and TR levels during *Xenopus laevis* development. For clarity, T_3 and mRNA levels are plotted as arbitrary units on different scales. Modified from (**?**)

3.1.5 *A role of TH signalling pathway before metamorphosis?*

As stated above, TR expression starts before THs are produced. This raises the question of the functional role of unliganded TR. As will be detailed in chapter **??**, TRβ is needed for a proper early development, especially for *X. laevis* eye formation (**?**). In addition, THs and TH metabolism (deiodinase activity) were recently detected in *X. laevis* eggs and early embryos (**?**), suggesting a possible

role of THs during early development in *X. laevis*, although not resolved yet. Overall, several labs (e.g. the Demeneix, Sachs and Brown labs) have recently gotten interested in possible roles of TR and/or THs during early development and the understanding of their relative roles is improving.

3.2 TH SIGNALLING PATHWAY IN OTHER CHORDATES

The analysis of species other than *X. laevis* or other anuran amphibians should bring interesting comparative insight into the evolutionary link between THs and metamorphosis in chordates. Several questions then need to be asked: are THs endogenously produced? What is their action on metamorphosis?

3.2.1 *In chordates TH production takes place in the specialized organ thyroid/endostyle.*

All chordates harbor a TH-producing gland, that is found in no other taxon, and which thus constitutes a chordate synapomorphy (see section **??**). It is an endoderm (gut) derivative located in the pharyngeal region, called thyroid in vertebrates and endostyle in invertebrate chordates. The main difference between the thyroid and the endostyle (which justifies the use of different names) is the follicular organization of the TH-producing cells in the thyroid and not in the endostyle (**?**). Some morphological discrepancies exist between the thyroids of different vertebrate groups, due to differential clustering of the numerous follicular units: for example, the thyroid is in one piece in mammals but birds and amphibians have two fully separated lobes on each side of the trachea. In most fishes and in agnathans, the follicles are not aggregated in an organized thyroid and are scattered in the pharyngeal region (figure **??**) (**?**). However, as the functional unit of the thyroid is the follicle, the macroscopic shape of the organ is probably not of fundamental functional and evolutionary significance.

Only recently has the homology between the thyroid and the endostyle been greatly strenghtened. More precisely, only pieces of the endostyle are thought to be homologous to the thyroid follicles ((**?**), p8). Indeed, some parts of the endostyle secrete a mucoprotein that helps gathering food particles from water and are specific to endostyles while others are able to fix radioactive iodine (**??**). It is the smallest, most dorsal and posterior part that does not produce mucus and that is thought to be the homologue of the thyroid follicles. This hypothesis was based on several arguments: (i) the endostyle and the thyroid are located in a similar position in the body plan, (ii) some endostyle cells have thyroid-producing properties such

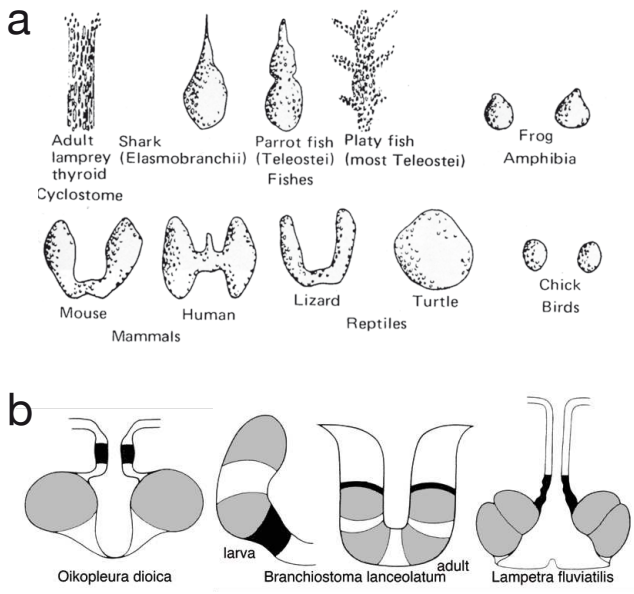


Figure 22: Pitcures of thyroid glands (**a.**) and endostyles (**b.**) from various chordates. In the endostyles, the thyroid-producing regions (black) are always situated dorsal to the glandular regions (grey). Modified from (**?**) and (**?**).

as iodine uptake (**???**) and TH metabolism (**????**), (iii) several thyroid-specific genes are expressed during endostyle formation (Ttf1 in lamprey (**?**)) and (iv) in lamprey, some cells of the endostyle transform into thyroidal follicles during metamorphosis (**?**).

It has been proposed several times (**??**) and recently re-actualized (see chapter **??**) that the common ancestor of chordates had an endostyle, quite like the extant ones: this endostyle had two functions: (i) the production of mucus for food collecting, and (ii) TH synthesis (**?**). The ancestor of all chordates lived in seawater and was most probably filter feeding. Among the species that derived from this ancestor, some became non-filter-feeding animals and thus did not need mucus production anymore. This latter function subsequently degenerated and the endocrine function only was kept in vertebrates (**?**). Lamprey development provides a nice living illustration of this scenario since, at metamorphosis, lamprey switches its mode of feeding from filter feeding to macrophagy - or non-feeding - in parallel to the disappearance of the endostyle and the formation of a thyroid with a typical vertebrate organization (**?**).

3.2.2 *TH metabolism in other chordates*

TH synthesis in other chordates

In mammals and fishes, TH metabolism is very similar to the one in amphibians (for a review, see (**?**)). The "thyroidal" picture in other chordates has been investigated much less thoroughly.

The search for TH derivatives in a wide variety of chordates was intense during the 50's and 60's and significantly calmed down afterwards (**?????**). These experiments were facilitated by the fact that TH synthesis is dependent on exogenous iodide, which has the great advantage of being very specific to THs. Consequently the detection of iradio-labelled odine characterizes and localizes TH synthesis.

THs were found in all chordate animals that were searched for them. T_3 and T_4 (as well as their precursors MIT and DIT) were detected in amphioxus (**??**), *Ciona intestinalis* (**??**), lampreys (**?**), sharks (**?**) and lungfishes (reported in (**?**)). TH metabolism was also found in all chordates examined: peroxidase-dependent iodine fixation (**??**), peroxidase activity (**???**), deiodinase activity (**???**), and a thyroglobulin-like protein were reported in lamprey, hagfish and amphioxus (**???**). More generally, the evolution of the TH signaling pathway within chordates will be discussed in chapter **??**. It is important to keep in mind that TH metabolism is not limited to T_3 and T_4 in mammals or amphibians: other TH derivatives are produced, like deiodinated, decarboxylated or deaminated derivatives, as well as sulfated and glucuronidated conjugates (**?**) (figures **??** and **??**). Similarly, it is most probable that other chordates have a richer iodinated tyrosine derivatives repertoire than just T_3 and T_4. As already mentioned in section **??**, this variability can be of significant evolutionary importance and should not be dismissed.

TH and TH derivative transport in chordates

After synthesis, T_4 needs to reach target organs from the production site (the thyroid gland). The tissue- and cellular-specific expressions of the genes involved in TH transport as well as deiodinases can influence the timing and localization of TH action and therefore modulate the tissue-specific response to TH. Generally in vertebrates, TH plasma transport is similar, although the relative importance of the different carriers varies between species - for example TBG is the main T_4 carrier in human instead of TTR in rodents) (**?**). In lamprey, one TTR was recently cloned from two species, *P. marinus* and *L. appendix* (**?**), and its expression was found to be correlated with metamorphosis. In amphioxus and in *C.intestinalis*, virtually nothing is known about TH transport.

Intracellular transporters contribute to the fine regulation of TH availability to each cell. Unfortunately, due to the relative youth of this field, virtually nothing is known about cellular transport outside mammals. Moreover the molecules are apparently not specific to THs and can also transport other amino-acid derivatives. Inside the cell, free TH levels are also controlled. For instance, CTHBP (Cytosolic Thyroid Hormone Binding Protein) is a cytosolic protein that binds T_3 and thus is likely to modulate the availability of T_3 to TR in the nucleus (**?**).

3.3 THS AND METAMORPHOSIS OUTSIDE ANURAN AMPHIBIANS

3.3.1 *Urodeles and amphibians paedomorphosis*

Studies on amphibian metamorphosis, as described in the previous sections, have been done on anuran amphibians like *Xenopus laevis*. However, not all amphibian species have a similar post-embryonic development. For instance, urodeles are a group of amphibians in which some species, called paedomorphic, are sexually mature at the larval stage. There are several sorts of paedomorphosis. Some species, called facultative paedomorphs, like *Ambystoma tigrinum*, may facultatively metamorphose in nature and may be induced to metamorphose by exogenous THs or even stressful conditions. The defects in the natural triggering of metamorphosis have been attributed, not to TR mutations, but to problems in TSH production e.g. (**?**). Other paedomorphs that never metamorphose in nature can be induced to do so under TH stimulation, again suggesting problems in TH production. On the contrary, permanent paedomorphs (like *Necturus*) cannot be induced to metamorphose, even after TH stimulation. A scenario for the evolution of paedomorphosis is detailed in the discussion (see chapter **??**).

3.3.2 *Metamorphosis in fishes is triggered by THs*

After amphibians, actinopterygian fishes have been best studied for the role of TH in the regulation of metamorphosis. More precisely, the molecular pathway controlling metamorphosis has been mainly elucidated in flatfishes (pleuronectiformes) such as flounders and turbot in which metamorphosis is spectacular, as in anurans. In these fishes a symmetrical planktonic larva transforms into an asymmetric benthic adult which has the two eyes on the same side of the body ((**?**), p442) (figure **??**). The data obtained from these species took advantage of the knowledge accumulated in anurans as a starting point and confirmed the central

role of THs and TRα in the early regulation of metamorphosis (**???**). A schematic structure of the TH signaling pathway involved in metamorphosis can be drawn from those studies in which the TH / TR couple appears to be the key component triggering metamorphosis in anurans and flat fishes (figure **??**). Upstream TH metabolism regulates TH production and availability, therefore TR activity; more downstream target genes whose expression is regulated by the coupled TH/TR control the morphological remodeling leading to metamorphosis (figure **??**).

The role of THs in post-embryonic development has also been documented in other groups. In zebrafish, THs were shown to be important for the larva-to-juvenile transition: exogenous THs induce differentiation of the pectoral fin precociously, in contrast the inhibition of TH synthesis blocks it (**?**). Smoltification in salmonid fishes corresponds to physiological, morphological, metabolic and behavioral modifications prior to migration from freshwater to sea water. It is considered to correspond to metamorphosis: it is a period of high T_4 and T_3 levels (**??**) involving a morphological, physiological and behavioral switch from a freshwater-adapted form to a salt-water-adapted form. At least some of the former modifications were shown to be TH-sensitive (**??**).

3.3.3 *TH in mammal development*

A noticeable difference with amphibians is the apparent lack of metamorphosis in mammals. Nonetheless, THs play central roles during mammalian development. Deficiency in TH production after birth causes severe developmental problems and especially mental retardation in humans. For instance, cretinism is due to problems in TH production (and the huge goiter in cretins' necks is caused by the swelling of the thyroid gland). Similarly, people living in remote mountains, used to have iodine supply deficiencies and mental development problems.

In mouse by suppressing THs or TR, it was shown that the alteration of the TH signalling pathway leads to several defects in eye, bone and intestine developments (for a review, see **?**). For instance mice exhibited delayed maturation of bone and intestine (**?**).

THs were shown to be especially important during weaning in mouse. Suppression of the thyroid gland (by the mutation of the *pax8* gene, involved in thyroid follicles formation (**?**)), causes severe malformation and death at weaning. Indeed, several vital morphological and metabolic modifications occur during that period. It is an ecological shift that is characterized by modifications like neural maturation, intestine remodeling or chondrogenesis that are reminiscent of similar events during amphibian metamorphosis. These modifications occur in correlation with

a peak of TH production (both T_3 and T_4) (**?**) and were shown to be under the control of TH and TR (**???**).

3.3.4 *TH and metamorphosis in other chordates*

Metamorphosis in lamprey

In lamprey, although several factors like temperature and growth advancement (indicated by length and weight, for instance) modulate metamorphosis (**?**), it is now well established that TH endogenous synthesis needs to be tightly regulated for the proper metamorphosis. Interestingly, the role of TH displays a remarkable difference compared to amphibians and teleost fishes: several reports clearly show that it is a drop, instead of an increase, in circulating TH that ends the larval period. This was shown by several means: circulating TH levels were shown to drop at the onset of metamorphosis, exogenous THs blocked metamorphosis and the inhibition of TH signaling with chemical compounds induced premature metamorphosis (**???????**). However, the molecular pathway that mediates TH action has been poorly investigated. T_3 binding was demonstrated in lamprey nucleus (**?**). There are two TRs in lamprey, from a lamprey-specific duplication (**?**), whose function was unknown before my PhD. In chapter **??**, I will present the molecular characterization that I have done for these receptors, in which I show that both receptors are functional (they bind TR DNA response elements, bind T_3 and activate transcription upon T_3 binding), suggesting interesting differences in the molecular events that regulate metamorphosis.

Metamorphosis in urochordates

Another interesting case of divergence of the TH signaling pathway is provided by urochordates. Metamorphosis of *Ciona intestinalis* can be induced by exogenous THs and inhibited by goitrogens known to block TH synthesis in vertebrates (**??**). However, the molecular pathways triggered by THs remain elusive since the only *Ciona intestinalis* TR does not bind T_3 (**?**). Chapter **??** will describe some of the work that I have done on this issue. The urochordates are particularly interesting regarding the evolution of metamorphosis because of the diversity of their post-embryonic developmental strategies. The elucidation of the role of THs and iodinated tyrosine derivatives in general in late development in other species, and in particular in larvaceans should generate interesting data for comparative analysis.

THs and metamorphosis in amphioxus

Chapter **??** presented the animal model amphioxus and chapter **??** described amphioxus metamorphosis, on which most of my work was focused.

Although amphioxus metamorphosis has been known for more than a century (**?**), its molecular basis has remained quite elusive until now. To my knowledge, the only attempt to decipher whether THs are involved in amphioxus metamorphosis was performed by **?**: the author treated premetamorphic larvae of *Branchiostoma lanceolatum* collected from plankton with T_3, T_4 and a TH synthesis inhibitor. He noticed a slight acceleration of metamorphosis by THs and no effect by the inhibitor. However, as the author confessed himself, few and rather old animals were used for the experiment: "*All the larvae were at a late stage of development, mostly with 12-14 gill slits*". Although "*not one was metamorphosing*", the larvae were already advanced in their larval life and were about to metamorphose (starting when the larva have about 12-13 gills, H.Escriva, personnal observation). As hormonal action occurs during an extended period of time, I conducted experiments with younger larvae to elucidate the possible role of THs during amphioxus metamorphosis (chapters **??** and **??**).

As explained in the next chapter, TR is the mediator of TH signal. In order to better understand the evolution of TH signaling pathway in chordates, I focused on TR action in amphioxus. The following chapter will introduce the molecular basis of TR action.

4

TR IS A TH-DEPENDENT TRANSCRIPTION FACTOR

Nuclear hormone receptors (NRs) constitute a rather large family of ligand-dependent transcription factors (figure **??**). Two properties of NRs make them of particular interest: (i) they play a central role in various physiological and pathophysiological processes as diverse as reproduction, development, and the control of homeostasis (**???**) and (ii) their activity is regulated upon ligand binding, and thereby may be manipulated with synthetic compounds, for therapeutical purposes. Because NRs are both intracellular receptors of non-protein compounds and ligand-dependent transcription factors, they have been a burning subject in molecular biology. Moreover, their dependency on ligand makes them convenient to study fundamental mechanisms of transcription. There are 48 NRs in human and 33 in amphioxus (**?**). Among NRs, TRs belong to the subfamily NR1 (boxed in orange in figure **??**) that regroups some of the best studied NRs (e.g. RAR, TR, PPAR).

This chapter does not intend to be exhaustive on TR molecular mechanism and will focus on notions that are relevant for the understanding of the experiments reported in this manuscript.

4.1 TR STRUCTURE / FUNCTION

TRs are nuclear proteins and ligand-dependent transcription factors. In the absence of a ligand, the receptor can bind DNA, interacts specifically with response elements in the regulatory regions of target genes, and actively inhibits transcription. Inhibition results because receptor conformation favors interactions with corepressors that recruit deacetylases able to block transcription (figure **??**). Ligand binding induces a conformational changes that results in the recruitment of coactivators which in turn alter chromatin structure and influence the recruitment of the transcription machinery (figure **??**) (**?**). This recruitment leads to activation of the expression of target genes, subsequently triggering specific biological processes such as metamorphosis (**?**). There are two TRs in vertebrates, TRα and TRβ that play different role in TH signaling and that were issued form a vertebrate specific

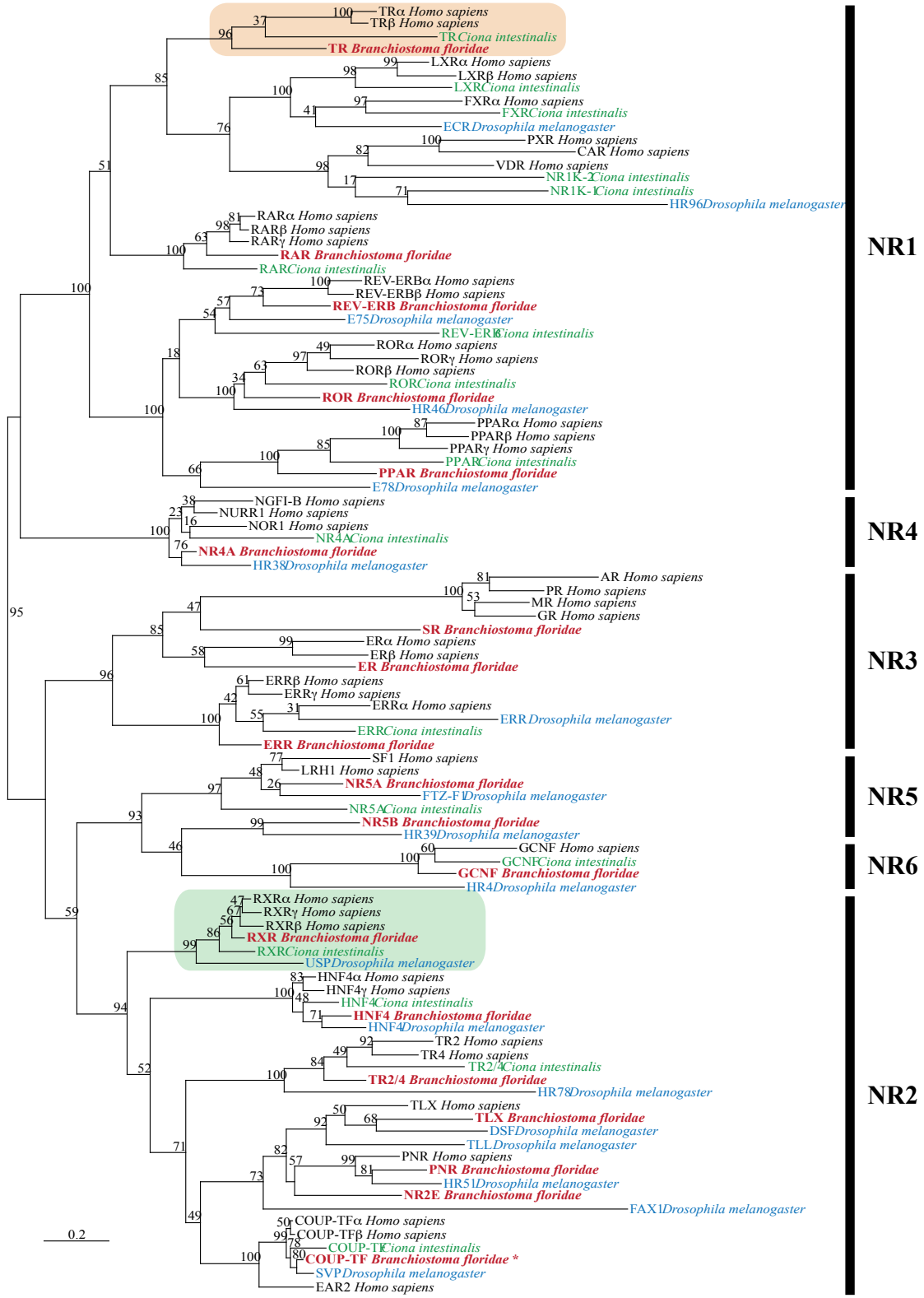


Figure 23: Phylogeny of the nuclear hormone receptor (NR) superfamily. The maximum likelihood tree is based on NRs from humans (*Homo sapiens*), sea squirts (*Ciona intestinalis*), fruit flies (*Drosophila melanogaster*) and amphioxus (*Branchiostoma floridae*). Human NRs are shown in black, sea squirt NRs in green, fruit fly NRs in blue and amphioxus NRs in red. TRs are boxed. Bootstrap support values are given for each branch. Modified from (**?**).

duplication (there have been independent supplementary duplications in fishes and in *X.laevis* (**?**), see section **??** for further details)

Like all NRs, TR displays a modular organization in functional domains (figure **??**): the DNA binding domain (DBD), the ligand binding domain (LBD), the N-terminal A/B domain, the hinge between the DBD and the LBD and the C-terminal F-domain. Only the DBD and the LBD are highly conserved between NRs, even within TRs (**?**) and will focus the attention of this chapter.

4.1.1 *A/B domain*

The A/B domain, at the N-terminal end, contains an activation function (AF-1) in TRα1 and TRβ1 isoforms (**?**). This domain is dependent on cellular context (ref cf Germain) and can also be submitted to post-translational modifications, like phosphorylation, that can alter transcriptional properties (e.g. DNA binding properties, AF-1-dependent activation) (**?**). The AF-1 is highly variable in both length and sequence. This domain displays a especially fast rate of substitution, potentially reflecting a minor evolutionary importance and has been the subject of virtually no evolutionary study.

4.1.2 *C domain: DNA binding domain*

TRs recognize specific sequences (figure **??**a), called thyroid hormone response elements (TREs). The consensus TRE is composed of two (AGGTCA) motifs in different configurations, the consensus binding site being constituted of 2 motifs in the same orientation and separated by 4 nucleotides (DR4, figure **??**a). There are however numerous alternative sites. For instance, an imperfect DR4 in the promoter of TR in *Xenopus laevis* (xDR4 in figure **??**a), is responsible for TR expression autoregulation (figure **??** in the previous chapter) (**?**). TRs may also bind inverted palindromes of the (AGGTCA) motif (figure **??**a). Whether all TREs contain (AGGTCA) motifs is still to be determined On the other hand, many potential sequences are not *bona fide* TREs, as exemplified in the genome-wide determination of binding sites of the estrogen receptor, another NR (**?**).

The DNA binding domain (DBD) is a highly conserved short 66-residue long core. 3D structures of the TR DBD show the same modular organization as other NRs (**?**):

- Two cysteine-rich zinc fingers (figure **??**b) interact with DNA on TREs (figure **??**c).

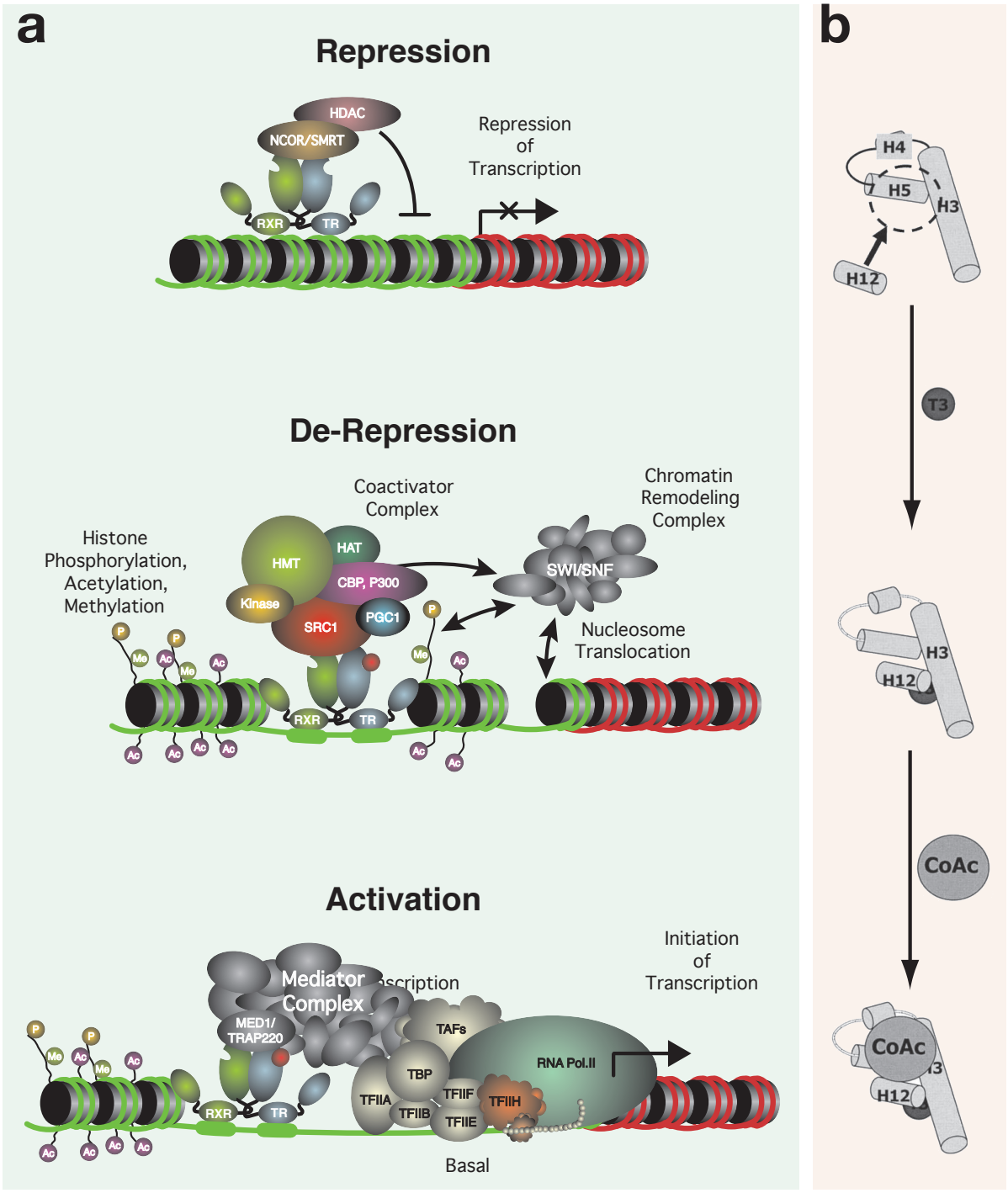


Figure 24: Molecular basis of NR-dependent transcriptional regulation. In absence of a ligand, NRs are bound to DNA to response elements located in the promoter regions of target genes and are associated with a complex containing histone deacetylases (HDACs) and corepressors (NCOR/SMRT) that repress transcription (upper panel). Binding of a ligand induces a conformational change in the LBD resulting in dissociation of corepressors and recruitment of coactivators (NCOA, CBP and P300 and PGC1) that are associated with histone acetyltransferases (HATs), histone methyltransferases (HMTs), kinases and an ATP-dependent chromatin remodeling complex (SWI/SNF), which together decompact and unfold the condensed chromatin (middle panel). Finally, the coactivators dissociate from the NRs and the mediator complex (including MED1/TRAP220) is assembled on the DNA locus. This mediator complex recruits transcription factors of the basal transcription machinery and RNA polymerase II (RNA Pol.II) to the promoter resulting in initiation of transcription (lower panel). Courtesy of Gérard Benoît, adapted from (**?**).

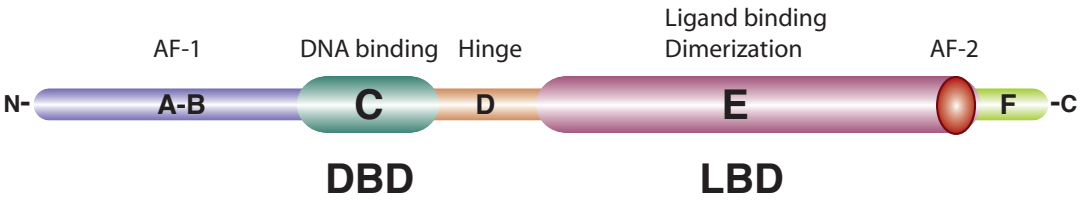


Figure 25: Schematic modular NR organization. Receptors are organized into the A-B domain, displaying a ligand independent activity (AF-1 domain), a DNA binding domain (DBD), a ligand binding domain (LBD), that also plays a role in dimerization and displaying a ligand dependent activity (AF-2 domain), a hinge between the DBD and the LBD, as well as a C-terminal domain.

- Between the two zinc-fingers, the P-box forms an α helix and makes direct contacts with DNA. It is responsible for DNA-binding specificity (**?**).
- The A-box, organized as an α helix, makes contacts with the minor groove of DNA (figure **??**c) with the T-box forming a connexion loop.

TRs can bind TREs as monomers (on one motif), homodimers or heterodimers with another NR, e.g. with VDR (vitamin D receptor) or RAR (retinoic acid receptor), but predominantly with RXR (retinoic X receptor, boxed in green on figure **??**), as illustrated in figure **??**c. Heterodimerization increases the affinity and specificity of binding site (**?**). For instance, the heterodimer TR-RXR binds the DR4 site, whereas VDR-RXR and RAR-RXR bind DR3 and DR5, respectively, although the heterodimer TR-RXR is the most important one (**?**, and references therein). The DBD also encompasses a minor dimerization interface (**?**).

4.1.3 *D domain: hinge*

This region is less conserved than the surrounding DBD and LBD and seems to serve as a "hinge" between the functionally more important regions. It leaves some conformational flexibility to the receptor and allows the DBD and the LBD to adopt different conformations and bind to different TREs (e.g. a DR4, which implies different orientations of the receptor binding the C-terminal half-site (**?**)). It was recently proposed that the D domain (**?**) is more structured (and so more rigid) than previously thought, with α helices (e.g. Ho in figure **??**). The D domain may then adopt in turn more or less rigid conformations and influence

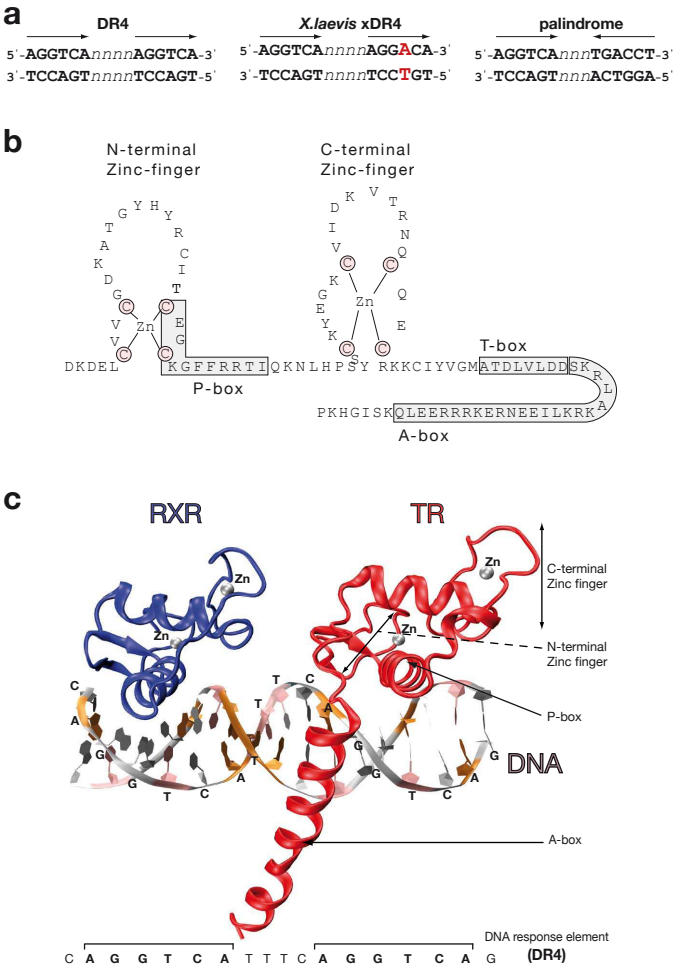


Figure 26: Structure / function of the TR DBD. **a.** DNA sequences of 3 TR response elements: the consensus DR4, the xDR4 site, with one divergent site with DR4, a palindromic binding site (**?**). **b.** Schematic illustration of the TR DBD. Boxes highlight regions involved in response element binding: the P-box, the T-box and the A-box, as seen in **c**. The 8 cysteine residues that direct bonds with zinc atoms have been highlighted. **c.** Structure of the 5′-RXR-TR-3′ heterodimer on a cognate direct repeat response element spaced by four base pairs (DR4). The two Zinc atoms in each DBD are visible. The different modules of the TR DBD are indicated on the right: the two zinc fingers surrounding the Zinc atoms, the P-box (structured as an α-helix) making direct contact with the DNA major groove and responsible for the binding motif specificity, the T-box and the A-box (structured as an α-helix) making minor groove contact. . Based on (**?**).

co-activator recruitment and DNA recognition. This was recently confirmed by the first determination of the 3D structure of the intact heterodimer PPAR/RXR, comprising both the DBD and the LBD (**?**). Indeed, the PPAR hinge was showed to be organized into several helices and to interact with DNA, 5' to the (AGGTCA) motif. Interestingly, the hinge domain of RXR is devoid of secondary structure and seems to be more flexible than the PPAR hinge, which may help RXR to adapt to the different partners it may have (e.g. TR, RAR) (**?**). The hinge also contains a nuclear localization signal that addresses TR to the nuclear compartment (**?**).

4.1.4 *E domain: ligand binding domain*

Conserved structure of NR LBDs

Although more divergent than the DBD, the LBD is rather well conserved among TRs (**?**). It is about 200 amino acids long and is responsible for both dimerization and ligand binding and also bears the main activation domain (AF-2). Crystal structures of numerous NR LBDs show a much better conservation of 3D structure than sequence. A typical LBD is constituted of 12 α helices organized in 3 layers (figure **??**a and **??**b). A hydrophobic pocket, surrounded by helices 3, 6 and β sheets, is called the ligand binding pocket (LBP) because the ligand can fit in (figure **??**a and **??**b) and interact with the receptor to stabilize an active conformation.

The unliganded TR is a transcriptional repressor

In the absence of a ligand in the LBP, TR is in the so-called "apo" form, *i.e.* it is not in the stabilized conformation in which the helix 12 (red in the figure **??**) would be close to helix 3, as described in section **??** (figure **??**b, upper panel, and more generally **??**, left panel). In this conformation, TR interacts with co-repressors, like N-COR (nuclear receptor corepressor) or SMRT (silencing mediator of retinoid and thyroid receptor) (**??**) (figures **??**a and **??**b, upper panels). More precisely, hydrophobic motifs (I/LXXII) in the corepressor can directly interact with a surface constituted by TR amino acids from the helices 3 and 5. This interface (site 1 in figure **??**) seems to be the principal interaction surface. Several other ones have been identified, in the helix 1 and above the helix 11 (sites 2 and 3 in figure **??**) (**?**). Both N-CoR and SMRT recruit in turn a large complex containing histone deacetylases (**???**) that will compact chromatin and block transcription (figure **??**a). So TR actively represses transcription in the absence of TH.

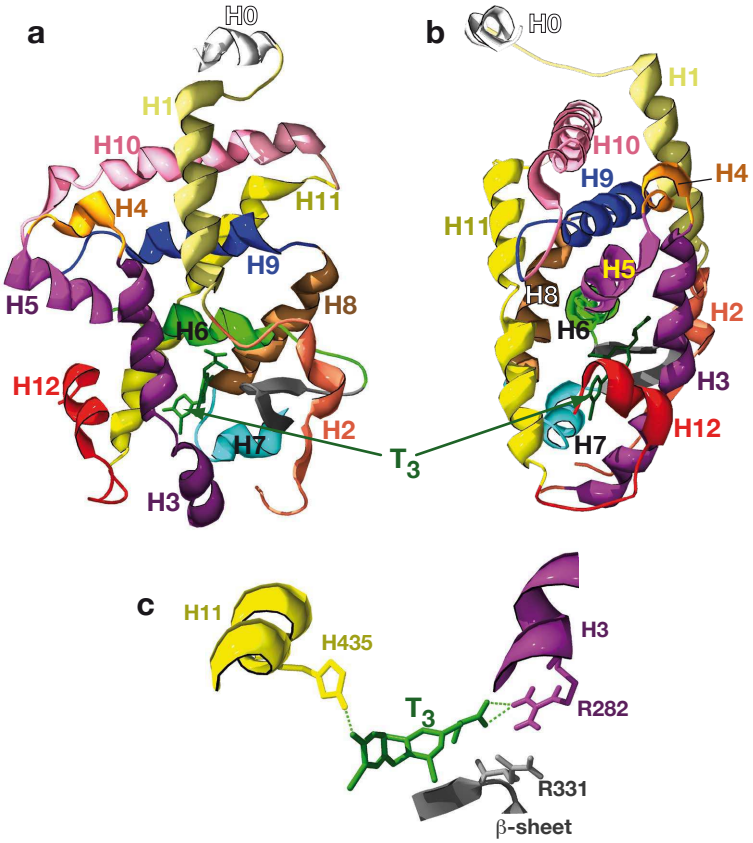


Figure 27: Structure / ligand binding of the TRβ LBD. **a.** and **b.** 2 perspectives of the same 3D structure of the human TRβ. The 12 α helices have been indicated, plus the helix HO at the C-terminal en of the hinge. The ligand T_3 is also indicated in the ligand binding pocket (LBP). **c.** Detail of interactions between TR LBD with T_3. Individual atoms of the side chains of residues arginine (at position 282) in the helix 3, asparagine (at position 331) in the β sheet and histidine (at position 435) in the helix 11, as well as T_3 are shown, plus the surrounding amino acids (represented as helices). Hydrogen bonds are indicated by green dashed lines. [PDB: 2h6w]. Illustrations built with the software SwissPDB-viewer.

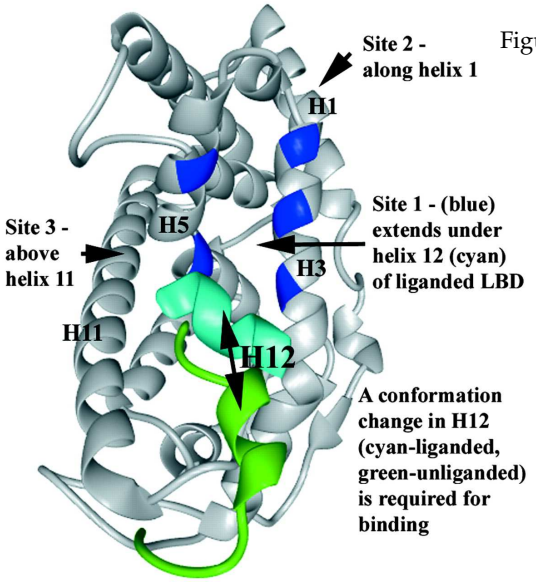


Figure 28: Model of the TR LBD indicating the surfaces and conformation changes regulating corepressor binding. The residues of helices 3 and 5 comprising the inferred N-CoR interaction site 1 are represented in blue. Two different conformations of helix 12 are indicated; the conformation in the hormone-bound TR structure is cyan (?), whereas the conformation modeled from the unliganded PPAR structure is green (?). In the hormone-bound structure some of the residues of Site 1 are buried, whereas in the unliganded model, Site 1 is fully exposed. The mutation-sensitive surface residues of Site 2 occur behind helix 1 on the right side of this view. The mutation-sensitive surface residues of Site 3 occur above helix 11 on the left side of this view. Illustration from (?).

Ligand		Affinity (nM)	
NAME	NATURE	TRα	TRβ
T_3	Agonist	0.058	0.081
T_4	"	14	3
Reverse-T_3	"	0.11	46
TRIAC	"	154	0.02
GC-1	"	0.44	0.067
NH3	Antagonist	20	93

Table 1: Affinities of human TRα and TRβ to some of their ligands (**?**).

Ligands

TR has a high affinity for T_3, as well as other TH derivatives, like the deaminated T_3 derivative TRIAC or the T_3 precursor T_4, as well as synthetic compounds like GC-1. Actually many compounds can interact with TR, with varying affinities. Relative affinities of TH derivatives for human TRβ are TRIAC > T_3 > T_4 > Reverse-T_3 > T_2 (table **??**, see figure **??** in chapter **??** for a description of the diverse compounds). These ligands just fit in the LBP (**?**) and stabilize the "bound" conformation by making direct hydrogen bonds with some amino acids of the pocket (e.g. T_3 in figure **??**). In this conformation, the ligand is then buried within the center of the LBD, in which it forms the hydrophobic core (**?**).

Conformational changes induced by ligand binding

Crystallography and directed mutagenesis reveal how TH binding leads to a series of changes in TR structure and transactivation potential. Indeed, TH binding induces a conformational change of the LBD, with a shift of the helix 12 towards helices 3, 4 and 5 (figure **??**b, middle panel, as exemplified for RXR on figure **??**), leading to the formation of an hydrophobic cleft composed of amino acids of helices 3, 5 and 12. This hydrophobic cleft can interact with a LXXLL motif present in coactivators, such as TIF2, ACTR or SRC-1,members the p160 family (figure **??**a, middle panel) (**?**). In turn, the p160s recruit other coactivators like histone-methyltransferases or histone acetyltransferases like CBP/p300 (REF). After chromatin remodeling by the coactivators, the transcriptional complex may be recruited by the coactivator TRAP220 (figure **??**a). Interestingly, a supplementary regulatory layer may be added, by altering the above complex formation in a cell-context dependent manner. For example, phosphorylation or methylation of complex subunits may lead to modifications of the complex integrity or specificity (**??**). In addition to mRNA transcription initiation, depicted above, some coactivators are thought to have other functions such as mRNA splicing, export from the

nucleus, translation, and posttranslational modifications of the synthesized protein or targeting to the proteasome (**?**).

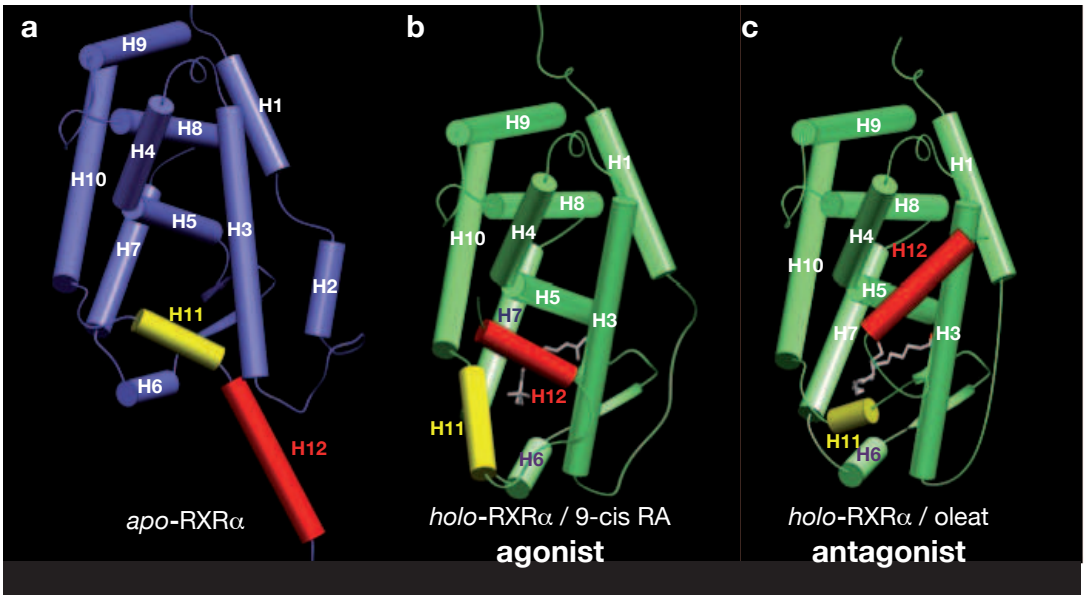


Figure 29: Structure of the RXRα LBD (**a**) in the apo form (*i.e.* in the absence of a ligand) (**?**), (**b**) in the holo form in complex the agonist 9-cis reinoic acid and (**?**) (**c**) (**?**) in complex with the antagonist oleat (**?**). The 12 α helices have been indicated. Ligand binding induces a conformational change of the LBD. Notably, the helix 12 (highlighted in red) displays different positions.

Coactivator recruitment is realized to the expense of corepressors. Indeed, the main corepressor interface (site 1 in figure **??**), which was available in the absence of T_3, becomes hidden by helix 12 in the liganded conformation. This refolding pattern (compared to a "mouse-trap" system) explains most NR ligand-dependent activity but is especially pronounced for RARs and RXRs (figure **??**), whereas the complete extension of the helix 12 in the "apo" conformation does not seem to be the rule for TRs (**?**).

It should be noticed that for other NRs, like ER, ligand binding induces releasing of the chaperone protein HSP and relocalization towards the nucleus (see chapter **??**).

NH3 is an antagonist of TR action

The above ligands, that induce conformational changes and TR transcriptional activation, are called agonists. So for instance, T_3, T_4, TRIAC or GC-1 are more or less potent agonists, with TRIAC and GC-1 as better agonists for TRβ than for TRα (table **??**). However, there are other kinds of ligands, called antagonists, that also fit in the LBP and interact with TR, but that do not induce the same conforma-

tional change (this alternative conformational change induced by antagonists is clearly visible for RXR in figure **??**, right panel). These compounds are intensely investigated for their medical potential (e.g. (**?**)).

Among TR antagonists, NH3 was designed to prevent a proper refolding of the helix 12 (figure **??**) to prevent coactivator recruitment (**?**). Accordingly NH3 does blocks TR activation, and the subsequent biological processes regulated by TR (e.g. amphibian metamorphosis). However, this compound also acts as a partial agonist because it induces the releasing of corepressors as well (**?**). So both activating and repressive potentials of TR are blocked by NH3. During my thesis, I have investigated the effects of NH3 on TR action and metamorphosis in amphioxus. Corresponding results are discussed in chapter **??**.

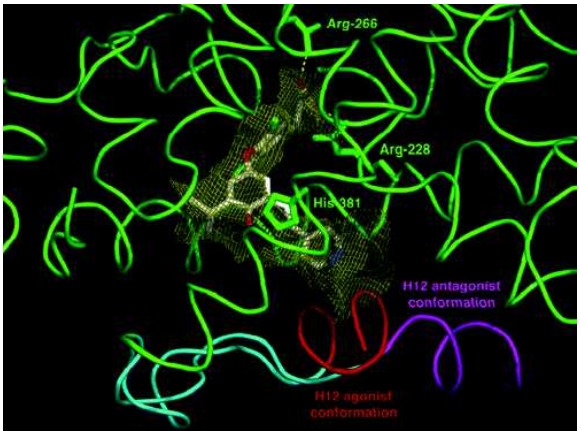


Figure 30: Prediction of the 3D structure of the human TRβ in complex with a designed antagonist docked into the ligand binding domain of TR. The backbone of the protein is displayed as a green tube (except the loop between H11 and H12, which is colored cyan, and antagonist conformation of H12 as magenta, and the agonist conformation of H12 as red). The side chains of residues Arg228, Arg266, and His381 are displayed as green capped sticks, and the ligand is displayed as capped sticks (white = carbon, red = oxygen, blue = nitrogen, green = bromine). Hydrogen bonds between the ligand and protein are represented by dashed yellow lines, and the surface of the ligand binding cavity is represented as a yellow mesh. There is a severe steric clash between the 5'-substituent of 6 and the agonist conformation of H12, which forces H12 into the antagonist conformation. From (**?**).

4.1.5 *Non-genomic effects of THs*

From experiments conducted in mammals, alternative mechanisms may also underlie TH action: upon ligand binding, cytoplasmic TR can interact with members of the phosphatidylinositol 3-kinase pathway (e.g. (**?**)). On the other hand, other pathways involve receptors different from TR, like some integrins (**?**) or the TAAR (trace-amine associated receptor) (**?**). Although these alternative pathways have a mechanism still unclear yet (**??**), they are probably marginal compared to the genomic pathway (as illustrated by the drastic reductions of TH effects in TR knock-out experiments (**?**). Metamorphosis in particular, which is of special interest here, depends on TR (see below). Moreover, receptor families, like TAARs, seem

to have a fast evolution (**?**) It is therefore still very difficult to extrapolate on the function of those genes outside mammals and even outside mouse and human.

However, non-genomic effects of THs stay marginal compared to TR-mediated effects since mice deprived of TH have comparable phenotypes than mice deprived of TR, although quantitative differences exist (see following sections) (**?**).

4.2 THE ROLE OF TR DURING DEVELOPMENT

4.2.1 *TR function during metamorphosis*

As explained earlier, TR is the main mediator of TH signal. Importance of TR during metamorphosis was evidenced by transgenic tadpoles in which a dominant negative form of TR was expressed and which failed to manifest major metamorphic events, like tail and gill slit resorption, intestinal remodeling or limb bud growth did not occur (**?**). The ability of TR to release corepressors and to recruit coactivators upon ligand binding was also demonstrated by several means: (i) metamorphosis was inhibited by TR antagonists (see above) and (ii) transgenesis with mutated TRs unable to release corepressors or mutated coactivators unable to interact with TR inhibited metamorphosis (**?**).

TR activation initiates an extensive transcriptional cascade (**??**). Many transcription factors are regulated early after TR activation, and some have been shown to contain a TRE in their promoter and to be direct TR targets - e.g. TH/bZIP (TH-induced basic leucine zipper transcription factor), BTEP (basic transcription element-binding protein) or TRβ itself (**???**) - that will in turn modulate the expression of their target genes... leading to metamorphosis.

TRα and TRβ have specific roles in the process of amphibian metamorphosis: TRβ is mainly implicated as a molecular switch triggered by the rise of TH levels whereas TRα is instrumental in blocking the metamorphosis program before TH levels increase (**??**). More precisely, TRα is expressed very early, even before TH synthesis, whereas TRβ expression is controlled by and thus correlates with TH levels (see figure **??** in chapter **??**) (**??**). This peak of TRβ expression at the onset of metamorphosis is particularly important because it transmits the signal of the peak of TH: as TRβ is one of its own target genes, the peak of TH production activates a positive loop of TRβ transcription and activation characteristic of the metamorphosis climax (**?**).

4.2.2 *TR and early development*

In amphibians

Attention was long concentrated on the role of TR during metamorphosis because TR was thought to have no major roles when virtually no TH is produced, as during early development (figure **??** in section **??**). Recently however, TR was shown to be important during amphibian premetamorphic development by inhibiting the expression of specific target genes. Accordingly, chromatin immunoprecipitation (ChIP) experiments showed that TR binds DNA on the promoter of target genes in premetamorphic tadpole larvae (**?**), when TH levels are undetectable, and is in a repressive state (**?**). A possible role of TR may then be to actively repress the initiation of metamorphosis. Indeed, in transgenic tadpoles expressing a dominant negative form of the corepressor N-COR, TR was not repressing gene expression, leading to an increase in the expression of its target genes as well as a faster development (**?**). However, TR inhibitory action during early development has more important effects than just preventing metamorphosis. The reciprocal transgenesis of mutated TRs unable to recruit N-COR showed a more drastic phenotype with embryonic lethality for 75 % of the tadpoles and with string defects in cranial and especially eye development for the remaining 25 % (**?**), demonstrating the function of TR before metamorphosis, for instance as a regulator of eye formation.

In mammals

TR activation mechanism, as TH-dependent transcription factor is very similar in mammals and in amphibians. As in amphibians, a role during development has been well documented, mostly through the use of transgenic mice. Comparative experiments were done between mice devoid of TRα, TRβ (e.g. (**??**)) and/or TH (mice KO for the gene *Pax8* do not develop a thyroid gland, (**?**)), with complementary results. From these experiments, TRα seems to mainly regulate early development, before weaning (about 3 weeks after birth). More precisely, the post-natal and pre-weaning period is characterized by important morphological and metabolic modifications like neural maturation, intestine remodeling, or chondrogenesis. These modifications occur in correlation with a peak of TH production (both T_3 and T_4) (**?**), and were shown to be under the control of TH and TRαfl (**???**).

The importance of TR action prior to weaning was exemplified by the severe developmental damages caused by both suppression of TH production or knock-in of a dominant negative TRα (**?**), leading in both cases to death at weaning. On the other hand, mice devoid of TRα suffered from much less severe defects and did not die at weaning. This effect was attributed to the transcriptional inhibitory role of

TRα before TH levels increase: not lifting the transcriptional repression exerted by TRα at the appropriate developmental moment (weaning), is more harmful than lifting TRα transcriptional inhibition during early development (**?**). Accordingly, the suppression of both TH and TRα causes a less severe phenotype (**?**).

4.3 TR PHYLOGENY

TRs were originally described in chicken (**?**) and soon thereafter in mammals (**??**) and in other gnathostomes (e.g. (**?**)). As mentioned above, there are two TRs in gnathostomes (TRα and TRβ) from a vertebrate-specific genome duplication (**??**) (actually there are two TRα and two TRβ in *Xenopus laevis* from a recent species-specific genome duplication (**?**) but the duplicates are functionally redundant). TRs were long thought to be vertebrate-specific, which turned out to be erroneous. Instead TRs have been found in both deuterostomes and protostomes: there are two TRs in lamprey from a lamprey-specific duplication (**?**), there is one TR in *Ciona intestinalis* (**?**), in amphioxus (**?**) (see chapter **??** for further details), in the echinoderm *Strongylocentrotus purpuratus* (**?**). TRs were also found in the platyhelminths *Schistosoma mansoni*, *S. japonium* and *Schmidtea mediterranea*, in a mollusc *Lottia gigantean*, in an arthropod *Daphnia pulex* (**?**) as well as in the annelid *Capitella sp.*. In contrast, TR was lost in ecdysozoans (no TR has been found in Ecdysozoan species whose genome has been sequenced so far) (**?**).

For a long time, attention was focused on vertebrate TRs, which all showed similar properties. Indeed, all vertebrate (and even gnathostome) TRs characterized so far display the same modular organization, the same 3D structure and the same activation process (excluding isoform-specific peculiarities, like mouse $TR\alpha_2$). The only exception, so far, is the TR from *Ciona intestinalis*, in which a 200-amino acid long insertion in the C-terminal end may be at the origin of T_3 binding impairments (**?**).

As will be considered in the discussion of this manuscript (see chapter **??** in the discussion), the ancient origin of TRs in bilaterians raises the question of the ancestral function of this NR, especially in species where no clear TH signaling has been described.

4.4 NR EVOLUTION

The evolution of TR ligand binding can be considered in the perspective of the evolution of NR evolution. Indeed, NRs display various ligand binding properties that seem difficult to reconcile. Several NRs like TR are high-affinity ligand-

dependant transcription factors whereas others like PPAR are low affinity receptors (**?**). However many other (about half of human NRs) are called orphan because they either have an unknown ligand or they do not have any and their activity is regulated by alternative mechanisms (**?**). Moreover NRs closely related bind very different ligands (e.g. TR binds a tyrosine derivative whereas RAR binds a vitamin A derivative).

This disparity in ligand binding properties raises the question of the evolution of NR ligand binding ability evolved (**????**). What was the ligand binding properties of NR ancestors? How did it evolve? These questions raise more general ones about the evolution of endocrine systems. In the chapter **??** and the general discussion (chapter **??**), I will go further into this debate, to which I contributed additional insights on the evolution of the estrogen receptor.

5 ORGANISATION OF THE MANUSCRIPT

My thesis work focused on the evolution of a late developmental process named metamorphosis. In this respect I have explored to what extent the evolution of the thyroid hormone signaling pathway may underlie the evolution of metamorphosis in chordates. The work I am presenting will highlight metamorphosis as a triptych in which endocrinology (TH availability), transcriptomics (TR regulation of gene expression) and morphology (remodeling characteristic of metamorphosis) are intricately linked. During my thesis I have investigated through the scope of these different aspects whether chordate metamorphoses were homologous and how they evolved from a common ancestor to reach extant diversity. I have also studied the evolution of the nuclear receptor by focusing on the estrogen receptor. In the rest of this manuscript, I will follow the following organization:

- In chapter **??** (**?**), I have studied whether amphioxus is able to produce THs in the first place and through what metabolic pathway. Using an *in silico* strategy, I searched in the amphioxus *Branchiostoma floridae* genome for orthologous sequences to vertebrate genes encoding proteins involved in TH metabolism.
- In chapter **??**, I showed that the T_3 derivative TRIAC induces metamorphosis in amphioxus through a molecular mechanism probably homologous to the vertebrate one. Both *In vivo* and *in vitro* experiments allowed me to show that TR function during metamorphosis was similar in amphioxus and in vertebrates, suggesting an ancestral role of TR as the mediator of TH signal (**?**).
- The previous work on TH amphioxus thyroid signaling pathway during metamorphosis was complemented by unpublished experiments described in chapter **??**. First, an *in vivo* analysis of TH metabolism in amphioxus was realized in collaboration with J.P. Cravedi (Toulouse) and H. Escriva. It allowed us to demonstrated that TRIAC is a natural TH product in amphioxus (at least in adults). Second, I also completed the analysis of the importance of THs during metamorphosis by studying the effects of TH synthesis inhibition on amphioxus metamorphosis. Finally, in collaboration with the AliX company, the 3D structure of the amphioxus TR in complex with TRIAC was obtained and allowed us to understand the structural basis of TR activity in amphioxus.

- A last article, described in chapter **??**, was dedicated to the understanding of the evolution of the estrogen receptor, a transcription factor closely related to thyroid hormone recepors (**?**).

- The final part of this manuscript (chapters **??** and **??**) discuss the evolution of NR ligand binding ability, in the light of my thesis work (chapter **??**). In chapter **??**, I will more thoroughly discuss the evolution of metamorphosis in chordates: the conservation of the molecular determinism of chordate metamorphosis as opposed to the intensive variability in tissue remodeling made us propose a model for the evolution of metamorphosis in amphioxus, according to which TH-regulated metamorphosis in chordates is ancestral. I will particularly discuss how my results relate to the current debate on the evolution of developmental regulation.

Part II

ARTICLES

6

THE AMPHIOXUS GENOME ENLIGHTENS THE EVOLUTION OF THE THYROID HORMONE SIGNALING PATHWAY

Before discussing whether THs regulate metamorphosis in amphioxus in chapter **??**, a first issue relates to the origin and evolution of TH signaling itself: is amphioxus able to produce THs? If yes, how?

TH signaling roughly comprises three modules: TH production (and degradation), TH action through TR binding, and TH target genes whose expression is regulated upon TR activation. An important feature of TH signaling is that THs are not proteins encoded by genes but instead are tyrosine derivatives produced through a metabolic pathway (described in section **??**). Although TH production has been previously demonstrated *in vivo* in amphioxus, it is not clear whether this production is homologous to the vertebrate one, and thus is ancestral in the chordate lineage.

The following article deals with this issue, namely the evolution of the TH signaling pathway in chordate. More precisely it is a phylogenetic study of the different members of the TH signaling pathway, utilizing species whose genomes have been sequenced. Notably I will focus on amphioxus, whose genome was recently sequenced (**??**). The evolution of the genes encoding the different proteins of the TH signaling pathway is studied using phylgenetic reconstructions made from maximum likelihood models of sequence evolution. This article will propose that THs are produced in amphioxus through a signaling pathway homologous to the vertebrate one, suggesting an ancestral origin of TH action.

This article was published in the journal Development, Genetics and Evolution, in the special edition about the annotation of the amphioxus genome.

Please notice that I am not presenting this article and the following one (chapter **??**) in the same chronologic order as the publication order (this article was published *after* the one described in chapter **??**, which it refers to). I chose this non-chronological order because I it was more logical to first present the study of TH production first and the one on TH effects *in vivo* second.

Dev Genes Evol
DOI 10.1007/s00427-008-0255-7

ORIGINAL ARTICLE

The amphioxus genome enlightens the evolution of the thyroid hormone signaling pathway

Mathilde Paris · Frédéric Brunet · Gabriel V. Markov · Michael Schubert · Vincent Laudet

Received: 11 July 2008 / Accepted: 18 September 2008

Abstract Thyroid hormones (THs) have pleiotropic effects on vertebrate development, with amphibian metamorphosis as the most spectacular example. However, developmental functions of THs in non-vertebrate chordates are largely hypothetical and even TH endogenous production has been poorly investigated. In order to get better insight into the evolution of the thyroid hormone signaling pathway in chordates, we have taken advantage of the recent release of the amphioxus genome. We found amphioxus homologous sequences to most of the genes encoding proteins involved in thyroid hormone signaling in vertebrates, except the fast-evolving thyroglobulin: sodium iodide symporter, thyroid peroxidase, deiodinases, thyroid hormone receptor, TBG, and CTHBP. As only some genes encoding proteins involved in TH synthesis regulation were retrieved (TRH, TSH receptor, and CRH receptor but not their corresponding receptors and ligands), there may be another mode of upstream regulation of TH synthesis in amphioxus. In accord with the notion that two whole genome duplications took place at the base of the vertebrate tree, one amphioxus gene often corresponded to several vertebrate homologs. However, some amphioxus specific duplications occurred, suggesting that several steps of the TH pathway were independently elaborated in the cephalochordate and vertebrate lineages. The present results therefore indicate that amphioxus is capable of producing THs. As several genes of the TH signaling pathway were also found in the sea urchin genome, we propose that the thyroid hormone signaling pathway is of ancestral origin in chordates, if not in deuterostomes, with specific elaborations in each lineage, including amphioxus.

Keywords *Branchiostoma floridae* · Cephalochordate · Chordate · Development · Evolution · Thyroid hormone · Endostyle

Introduction

In isolated human populations, usually located far from the sea, goiter and mental retardation are efficiently prevented by using iodized table salt. Iodine is mainly used in mammals to synthesize thyroid hormones (THs), so the mental impairments observed among "cretins" indicate the important role that THs play as regulators of development (Flamant and Samarut 2003). Outside mammals, THs can have more drastic effects on development and, for instance, THs are also responsible for the regulation of metamorphosis in amphibians (Shi 2000). In this case again, the main source of TH is endogenous, since inhibition of TH synthesis by chemical means (using goitrogens such as PTU) blocks metamorphosis. The metabolism of TH has been extensively studied in mammals and amphibians, and was found to be very similar. In both cases, the most active form is T_3 (3,3′,5-triiodo-L-thyronine), which can be produced from its precursor T_4 (L-thyroxine). T_4 is an iodinated tyrosine derivative: it is formed by two tyrosines, coupled each with two iodines whereas in T_3 the external

Communicated by J.J. Gibson-Brown

Electronic supplementary material The online version of this article (doi:10.1007/s00427-008-0255-7) contains supplementary material, which is available to authorized users.

M. Paris · F. Brunet · G. V. Markov · M. Schubert · V. Laudet (✉)
Institut de Génomique Fonctionnelle de Lyon, CNRS UMR5242–INRA 1288–ENS–UCBL, IFR128 BioSciences Lyon-Gerland, Ecole Normale Supérieure de Lyon,
46 allée d'Italie,
69364 Lyon Cedex 07, France
e-mail: Vincent.Laudet@ens-lyon.fr

tyrosine lacks an iodine (Fig. 1a). In vertebrates, TH synthesis takes place in the thyroid gland, a gut derivative, which is organized in follicles constituted of single layered epithelial cells called thyrocytes, enclosing a lumen filled with the matrix protein thyroglobulin. The main steps of TH signaling pathway are indicated in Fig. 1b. The first event in TH synthesis involves increasing the concentration of iodine in thyrocytes, through the sodium/iodine symporter (SIS) in their basal membrane. This is an important step, since iodine is present at very low concentrations in food, as well as in blood. In the lumen of the follicles, iodine gets oxidized and transferred to a few specific tyrosine residues on thyroglobulin to form monoiodotyrosine—MIT—if only one iodine is transferred to a tyrosine, or diiodotyrosine—DIT, if two iodines are transferred to a tyrosine. Iodinated tyrosines are further coupled, under the catalysis of the thyroid hormone peroxidase (PERT). The iodinated thyroglobulin is then incorporated back into the thyrocytes and hydrolyzed in lysosomes, allowing the release of T_4 (if two DITs have been coupled), and, to a

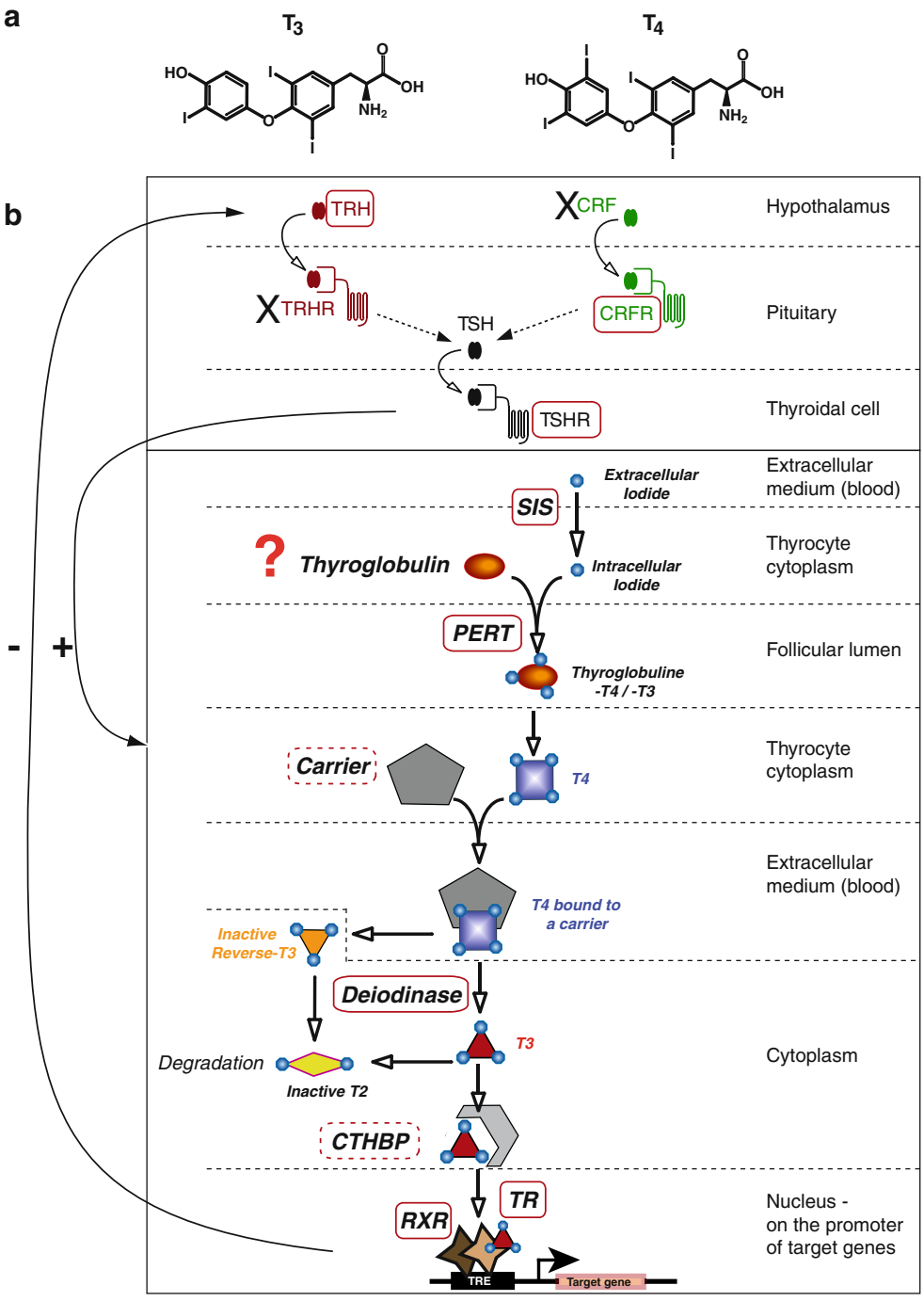


Fig. 1 TH signaling pathway, as it is known in vertebrates. **a** Molecular structure of T_4 and T_3. **b** The different steps of the TH signaling pathway, as it is known in vertebrates, are indicated. The steps of TH production and action have been separated from the hypothalamo-pituitary regulation of TH production (see the *boxes*). The neuroendocrine regulation of TH production is indicated with a "+". The negative feedback loop by TR on this upper regulation is also indicated. The genes encoding members of the TH signaling pathway and studied here are highlighted as follows: (1) the genes for which orthologous sequences were found in the amphioxus genome are *boxed in red* and (2) the genes for which no orthologous sequence was found are indicated with a *cross*. The case of thyroglobulin remains unclear. Given the uncertainty of the carrier (TBG being the only one found) and CTHBP proteins (see the text), the amphioxus sequences are indicated with *dashed lines*

lesser extent, T_3 (if a MIT and a DIT tyrosines have been coupled) (for a review, see Hulbert 2000).

Although THs are produced only in the vertebrate thyroid gland, they are found everywhere in the organism, transported through blood circulation. The hydrophobic benzene rings of TH increase the tendency of THs to partition in plasma membranes (Schreiber and Richardson 1997), so that only a small fraction of T_4 and T_3 is "freely" transported and most of it travels bound to a "carrier" protein, like thyroxine-binding globulin (TBG), transthyretin (TTR), or albumin in mammals. In peripheral tissues, T_4 is deiodinated into the active form T_3 by specific proteins called deiodinases. These enzymes remove iodines from the inner or outer ring of the tyrosine skeleton. There are three different deiodinases in tetrapods (D1, D2, and D3), which have different affinities for the various THs and are responsible for fine-tuning thyroid hormone action jointly with a few alternative TH pathways (like sulfatation and glucoronidation) that produce mostly inactive T_3 derivatives, targeted for fast degradation (Wu et al. 2005). In each target cell, T_3 binds to its receptor, the thyroid hormone receptor (TR), which belongs to the superfamily of nuclear hormone receptors (NRs) (Flamant et al. 2006). As many members of the NR superfamily, TR is a ligand-dependent transcription factor: in the absence of T_3, TR is most often bound to DNA on specific sequences in the promoter of target genes whose expression it inhibits, through the recruitment of transcriptional corepressors. Upon T_3 binding, TR recruits coactivator proteins and activates the expression of target genes that will lead to the biological effects induced by THs (Fig. 1b).

In mammals, whereas deiodinases constitute a peripheral system for controlling TH production, the hypothalamic pituitary thyroid axis allows a central control of TH production by the thyroid. In this axis, the hypothalamus produces TSH-releasing hormone (TRH) that stimulates the secretion by the pituitary of thyroid-stimulating hormone (TSH), which in turn stimulates TH production by TH-producing cells in the thyroid (Fig. 1b). An important aspect of this regulation is the negative feedback loop, through which THs are negatively regulating their own production through the inhibition of TRH and TSH production (Fig. 1b). This TH-dependent axis is a paradigm of integrated exquisite endocrine regulation of hormone production at the level of the organism, because it can integrate external signals (e.g. food, population density), therefore allowing a link between the environment and the endocrine production of THs (Yen 2001). In amphibians, up-regulation of TSH secretion is taken care of by the corticotropin-releasing hormone (CRH), a primary regulator of stress response that can stimulate TSH production. Thus, the production of the CRH peptide by the hypothalamus is dependent on environmental stimuli and integrates stress information that can be transmitted down to the thyroid gland (Denver 1997) (Fig. 1b).

The evolution of this rather complex signaling pathway has been investigated mainly in gnathostomes, where it was found to be well conserved (Hulbert 2000). However, it is still unknown how the capacity to produce THs and their function evolved in the first place. In the basal vertebrate lamprey, THs are produced through a pathway probably homologous to the mammalian one (for instance Manzon et al. 2007) and regulate metamorphosis, as in amphibians, with the difference that it is a drop, and not a peak of TH, that triggers metamorphosis (Manzon et al. 2001). Not much is known about the TH pathways of non-vertebrate chordates (the vertebrate sister group urochordates like the sea squirt *Ciona intestinalis*, and the cephalochordates like amphioxus), except that they possess an organ homologous to the thyroid gland, which is named endostyle, and can produce THs (see Paris and Laudet 2008 for a review). However, biological effects of THs outside vertebrates have been reported several times, and THs have been linked with metamorphosis in urochordates (Patricolo et al. 2001), amphioxus (Paris et al. 2008), and in echinoderms (Heyland and Hodin 2004).

Several lines of evidence, including biochemical studies, demonstrated that there is an active TH metabolism similar to vertebrates in the most basal chordate: amphioxus (reviewed in Eales 1997). First, the endostyle is an amphioxus organ that is widely accepted as being homologous to the follicles of the vertebrate thyroid gland (Ogasawara 2000). Secondly, both T_3 and T_4 have been detected in amphioxus (Covelli et al. 1960). Thirdly, fixation of iodine was reported in the endostyle of amphioxus (Fredriksson et al. 1985) and was shown to be dependent on peroxidase activity. Fourthly, a protein with biochemical properties similar to thyroglobulin has been described in amphioxus (Monaco et al. 1981). Fifthly, deiodinase activity has been indirectly demonstrated by showing that inhibition of T_4 deiodination by chemical means inhibits metamorphosis. Lastly, there is an active TR in amphioxus, that is involved in metamorphosis (Paris et al. 2008). Taken together, these results strongly suggest that there is production of T_3 and T_4 and more generally an active TH signaling pathway in amphioxus. However, only very few members of this signaling pathway have been identified so far in amphioxus. Here we take advantage of the recent release of the amphioxus (*Branchiostoma floridae*) genome (Holland et al. 2008; Putnam et al. 2008) and describe orthologs of the main genes involved in the TH signaling pathway. We propose that in amphioxus THs are produced the same way as in vertebrates, which suggests an ancient origin of the TH signaling pathway within the chordate lineage. Several lines of evidence from

echinoderms suggest an even more ancient origin of TH signaling within deuterostomes.

Materials and methods

Sequence retrieval We used zebrafish or human protein sequences to retrieve Ensembl families (version 45 as of June 2007) of our sequences of interest. Ensembl families are made of genomic sequences as well as Swiss-Prot and TrEMBL data. We subsequently blasted (blastp) the protein sequences of interest (usually a human sequence) against the amphioxus genome (Putnam et al. 2008). Orthologous sequences from other species (e.g. the sea urchin *Strongylocentrotus purpuratus*, the sea anemone *Nematostella vectensis*, and the bee *Apis mellifera*), used as outgroups, were retrieved by blast searches with the amphioxus sequences carried out on the NCBI site (www.ncbi.nlm.nih.gov/). Protein sequences from the lamprey *Petromyzon marinus* were obtained from protein, EST or genomic databases. Protein sequences from the elephant shark *Callorhinchus milii* were obtained from genomic databases (after protein sequence prediction using Genscan). A complete list of retrieved sequences is given Fig. S2. We also did a reverse blast (blastp) onto all Metazoa non-redundant sequences available both at the Swiss-prot group (expasy.org/sprot/) and NCBI (ncbi.nlm.nih.gov/) and constructed phylogenetic trees using selected sites of the sequences. Closely related genes were used as outgroups. Alignments were made using Muscle (Edgar 2004), allelic sequences were removed from the alignments after comparisons of the nucleotide sequences and best conserved sites were selected using GBlock (Castresana 2000) for further phylogenetic analysis.

Phylogenetic analysis For each family, we performed first a neighbor-joining (NJ) analysis, with Poisson law correction and pairwise gap-removal, as implemented in PhyloWin (Galtier et al. 1996) (the relevant data are available upon request to FB), and subsequently we performed a maximum likelihood (ML) phylogenetic analysis using PhyML v2.4.4 (Guindon and Gascuel 2003). For the ML analyses, JTT model with eight rate categories and an estimated gamma shape parameter were used. Robustness was assessed by bootstrap analysis (1,000 repetitions).

Evolutionary rate comparison between vertebrate peroxidases Two urochordate sequences (one—PERT1—from *C. intestinalis* and the other from *Halocynthia roretzi*) were used as outgroups. The sequence Q6NUY7_BRARE, Q640K8_XENLA, and Q9YH34_XENLA were excluded from the analysis because their phylogenetic position does not fall into one of the four peroxidase groups PERT, PERE, PERL, and PERM. The input tree corresponded to the topology described in Fig. 3.

Results

Our strategy was to search for amphioxus sequences from the *B. floridae* genome, that are homologous to vertebrate proteins involved in TH signaling. When possible, sequences from basal vertebrates, urochordates, and sea urchin were also studied for comparative purposes. Our search identified a total of 24 genes in amphioxus (Holland et al. 2008; Putnam et al. 2008; Schubert et al. 2008) plus RXR and TR that were previously identified. For almost all gene families, several orthologs were found, arising from lineage specific duplications. In most cases, the amphioxus genes were not direct orthologs of each vertebrate paralogous gene, but rather branched at the base of a tree of closely related vertebrate genes. The complete list of identified genes, with their corresponding accession numbers assigned by the JGI consortium, is given in Table S1.

SIS

Blast searches allowed us to identify eight predicted genes, for which orthology was tested by a phylogenetic analysis (Fig. 2). We found three groups of closely related vertebrate genes, including SIS, each time with 100% bootstrap (the sodium-iodide symporter SIS, the apical iodide transporter, and the sodium-lactate cotransporter). Two out of the three groups contain genes encoding proteins characterized as sodium-iodine symporters located either on the basolateral membrane (SIS) or on the apical membrane of thyrocytes (apical iodide transporter) in human (Rodriguez et al. 2002) (shown in brackets in Fig. 2). The genes of the third vertebrate group encode proteins identified as sodium-lactate cotransporters (Gopal et al. 2007). Notably, three sea urchin (*S. purpuratus*) genes cluster together at the base of the vertebrate groups (with a bootstrap support of 71%). When going deeper in the tree, eight amphioxus genes cluster together (96% bootstrap support) at the base of this cluster of paralogous genes (with a bootstrap support of 75%). These topologies suggest three independent series of duplication, one at the base of the sea urchin lineage, one in the amphioxus lineage, and the other at the base of the vertebrate lineage, giving rise to the three vertebrate-specific groups. As a sequence from the elephant shark *C. milii* genome was found in all three vertebrate groups (with low bootstrap support, though, Fig. S1), the vertebrate duplications occurred before *C. milii* split, probably during the two rounds of whole genome duplication that occurred at the base of the vertebrate group (Dehal and Boore 2005).

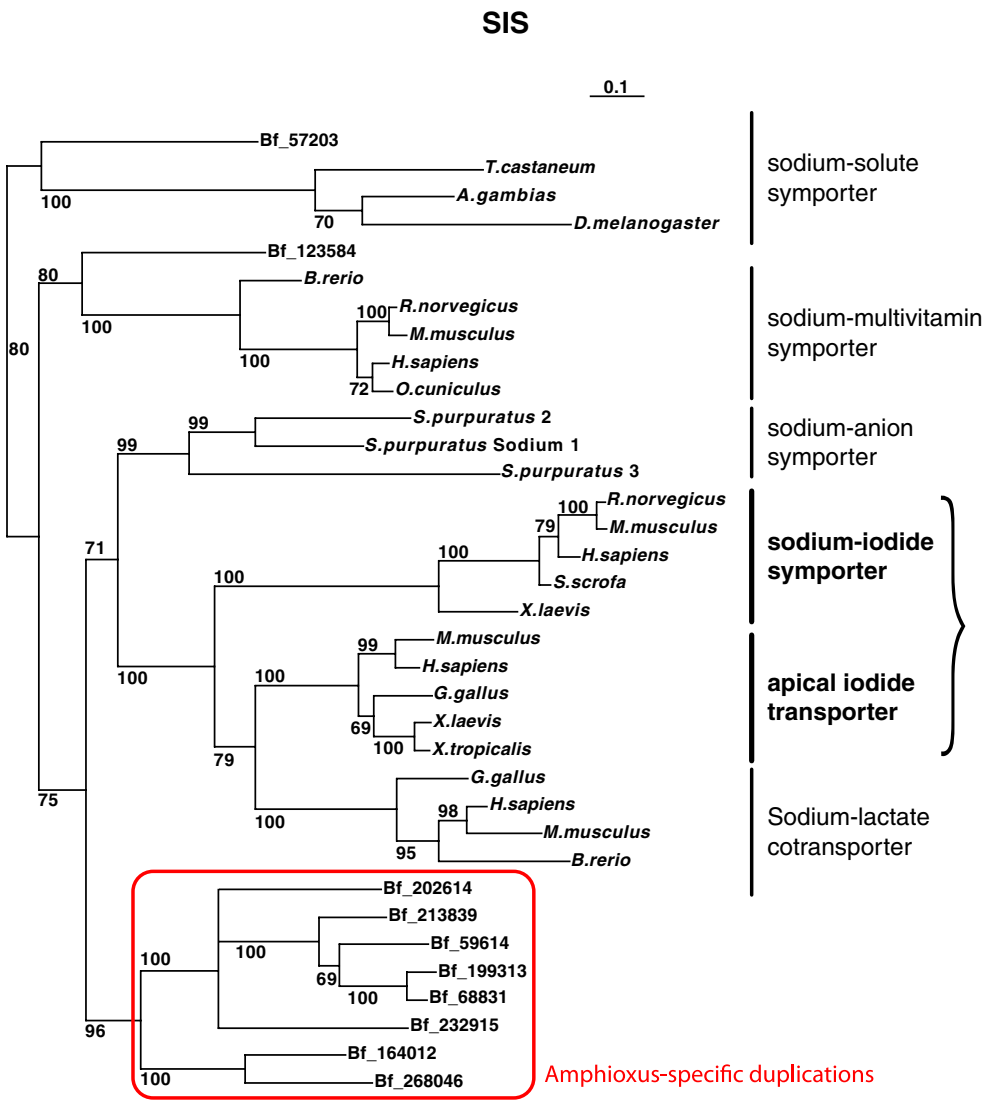


Fig. 2 Phylogenetic tree of SIS and related protein sequences. A maximum likelihood (ML) tree was obtained from analysis of SIS amino acid sequences. Bootstrap percentages obtained after 1,000 replicates are shown. Nodes with bootstrap support below 50% were collapsed. The amphioxus SIS-like sequences have been *boxed in red*. The *scale bar* indicates the number of changes per site. A similar tree including sequences from basal vertebrates is given in Fig. S1

The structures of the amphioxus proteins were then predicted *in silico*, in order to determine if one of them was more likely to be a sodium-iodine symporter. From the TMpred program (http://www.ch.embnet.org/software/TMPRED_form.html), the eight sequences were predicted to harbor 12 to 13 transmembrane domains like the already characterized vertebrate SIS, except the sequence Bf_68831 (Fig. 2), for which 12 supplementary transmembrane domains were predicted in its 200 amino acid longer C-terminal region (data not shown). Moreover, several sites that are known to be important for the general integrity of the function of the transporter and are found to be conserved within sodium symporters, are conserved in the amphioxus sequences (Dohan et al. 2003).

PERT

PERTs are part of a big family of peroxidases that catalyze oxidation of various substrates with hydrogen peroxide (including myeloperoxidase (PERM), eosinophil peroxidase (PERE), thyroid peroxidase (PERT), and lactoperoxidase, (PERL) (Daiyasu and Toh 2000)). Paralogous sequences have been described in urochordates and in the Japanese amphioxus species *Branchiostoma belcheri* (Ogasawara 2000; Ogasawara et al. 1999).

Blast searches allowed us to identify seven predicted genes, for which orthology was tested by phylogenetic analysis (Fig. 3). From our analysis, two groups consisting of three and four amphioxus sequences respectively branch together with high bootstrap support (100%) each. The group of three genes is distantly related to PERTs but still branches within peroxidase-like proteins (and the corresponding proteins may thus display peroxidase activity). The four other amphioxus sequences (red box, Fig. 3), with which the known *B. belcheri* sequence branches (Ogasawara 2000), are more closely related to PERTs and are located at the base of the tree constituted of the vertebrate-specific subgroup and urochordate sequences with a good bootstrap support

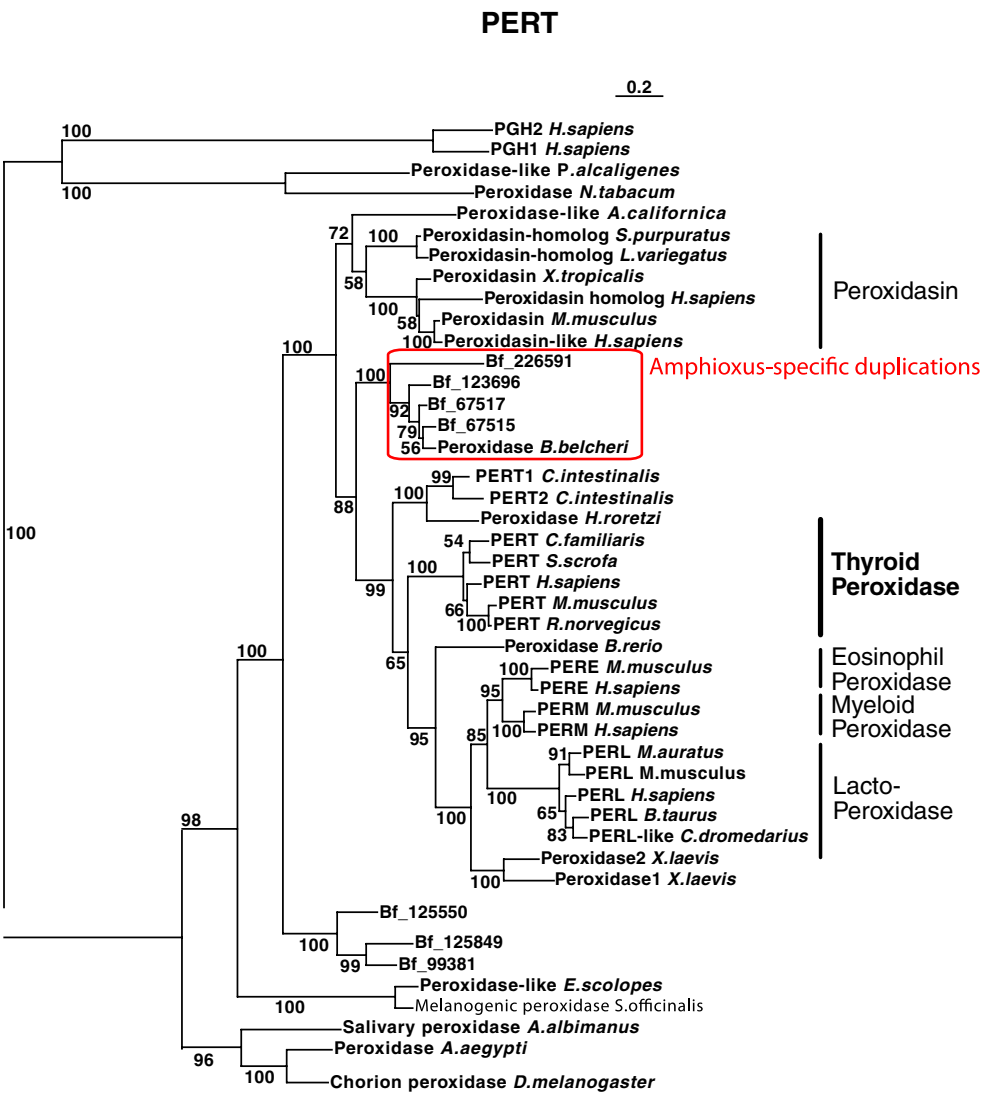


Fig. 3 Phylogenetic tree of PERT and related protein sequences. A maximum likelihood (ML) tree was obtained from analysis of peroxidase amino acid sequences. Bootstrap percentages obtained after 1,000 replicates are shown. The amphioxus PERT-like sequences have been *boxed in red*. PGH: Prostaglandin G/H synthase. The *scale bar* indicates the number of changes per site. A similar tree including sequences from basal vertebrates is given in Fig. S2

(88%). Amphioxus sequences have very short branches, suggesting a low evolutionary rate and a rapid duplication burst. Similarly, PERT genes seem to evolve at a slower rate than the three closely related peroxidases PERM, PERL, and PERE. To test this hypothesis, we compared the evolutionary rates between the different vertebrate peroxidase families, using the relative-rate test on available sequences (Robinson et al. 1998). Results are shown in Table S3, as differences of substitution rate between groups of species. From these analyses, it appears that PERTs have effectively evolved at significantly lower rates than the three other peroxidase groups. This common slow evolutionary rate in amphioxus and vertebrate PERTs may reflect that these proteins have kept an ancestral function (see Discussion).

Thyroglobulin

Thyroglobulin is a large protein (more than 2,700 amino acids in humans) that contains an esterase domain extending to the 500 most C-terminal amino acids, whereas the N-terminal domain houses thyroglobulin type I repeats (TY repeats), in which many PERT-targeted tyrosines have been located. Many hits are retrieved from blasting the human thyroglobulin sequence against the amphioxus genome. However, when they are aligned with esterase domain of thyroglobulins, they branch with esterases and not thyroglobulins in a phylogenetic tree (data not shown) and lack a N-terminal part homologous to the thyroglobulin one. Conversely, TY repeats were recognized in some predicted genes. However, these repeats, although first discovered in thyroglobulins (hence their name), are found in many different proteins (Novinec et al. 2006). Among the amphioxus retrieved sequences, the best hit (Bf_123169) is a very long protein (about 2,400 amino acids). However, nothing else but the TY repeats is homologous to the thyroglobulin (data not shown) and the C-terminal part is not an esterase sequence. Overall, up to now, no unambiguous thyroglobulin sequence was found

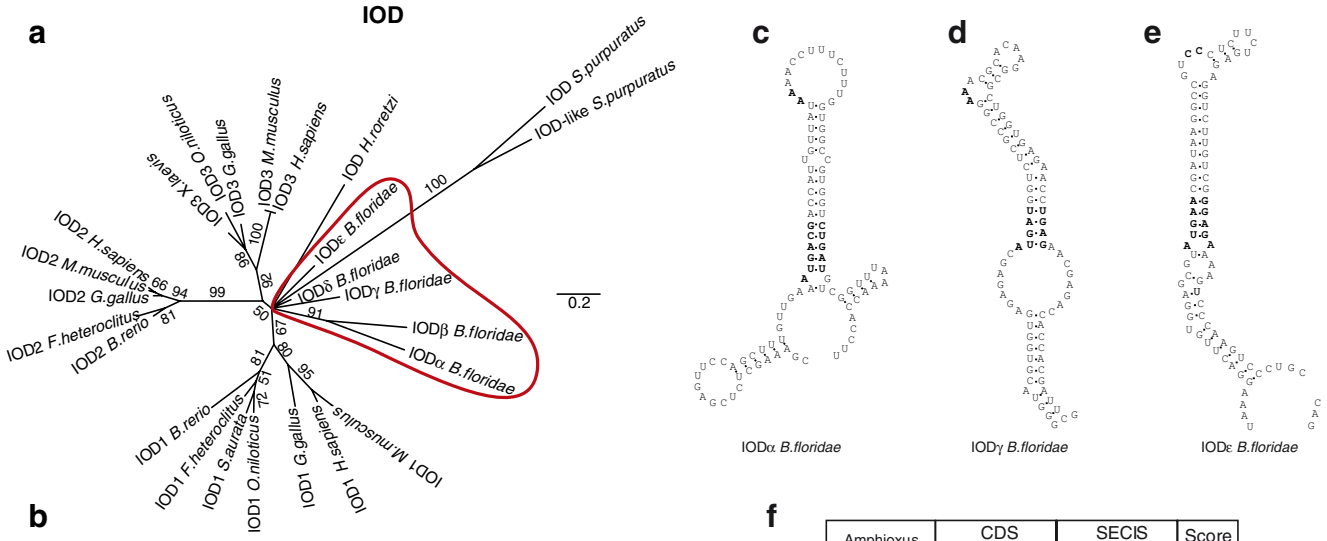
in the amphioxus genome or outside vertebrates, although one rather short sea urchin (*S. purpuratus*) sequence (137 amino acids, accession number 115921343) clusters with vertebrate thyroglobulins.

Transport proteins

Several carriers for THs have been identified in vertebrates, like thyroxine-binding globulin (TBG), transthyretin (TTR), and albumin (with a rather low affinity). Although present in lampreys (Schreiber and Richardson 1997) and sea urchins (*S. purpuratus*) (data not shown), no gene encoding albumin (that is not specific to TH transport) and no TTR gene were found in the amphioxus genome. One gene related to TBG was found but it is not a direct ortholog (Fig. S4). Once in the cell, THs can also bind to cytosolic proteins, like the cytosolic thyroid hormone-binding protein (CTHBP). A gene made of the concatenation of two predicted genes from the genome (Bf_267438 and Bf_110402) branches at the base of a subtree constituted of two vertebrate gene families including CTHBP (Fig. S5). Further studies will be required to assess, which proteins mediate TH transport in amphioxus.

Deiodinases

We retrieved five sequences from the amphioxus genome that cluster with previously described deiodinase sequences (data not shown). However, only one amphioxus sequence (Bf_123596 called IODβ in Fig. 4a) includes a domain known to be important for vertebrate deiodinase function (Bianco et al. 2002) (Fig. 4b). All other predicted sequences were truncated at those positions. Notably, in this activation domain, all vertebrate deiodinases, which are selenopro-



f

Amphioxus sequence	CDS		SECIS		Score (>15)
	Start	End	Start	End	
IODα *B.floridae*	2437	16226	16544	16639	29.8
IODγ *B.floridae*	676	2481	3058	3153	25.8
IODε *B.floridae*	391	2432	2992	3081	22.07

Fig. 4 Phylogenetic and structural analyses of putative deiodinases in amphioxus. **a** A maximum likelihood (ML) tree was obtained from analysis of deiodinase amino acid sequences. As no outgroup was found for deiodinases, the tree presented is unrooted. Bootstrap percentages obtained after 1,000 replicates are shown. Nodes with bootstrap support below 50% were collapsed. The amphioxus sequences have been *circled in red*. The *scale bar* indicates the number of changes per site. A similar tree including sequences from basal vertebrates is given in Fig. S3a. **b** Amino acid sequence alignment of the active catalytic domain of the deduced amino acid sequences for the human deiodinases as well as the sequences found in invertebrates. The site, where a selenocysteine is located, is shown in *red*. In humans and *H. roretzi*, the translation of the TGA codon into a selenocysteine has been experimentally verified and is shown in *bold* (Berry et al. 1991; Curcio et al. 2001; Baqui et al. 2003). In the amphioxus and sea urchin sequences, the homologous TGA is proposed to be translated into a selenocysteine as well (*italicized*). **c**–**e** SECIS elements in the putative 3′UTR of amphioxus IODα (**c**), IODγ (**d**), and IODε (**e**), as predicted by SECISearch analysis (Kryukov et al. 2003). The characteristic adenosine that precedes the quartet of non-Watson-Crick base pairs, a TGA_GA motif in the quartet and two adenosines or cytosines in the apical loop, are highlighted in *bold* (Kryukov et al. 2003). **f** The position of the CDS in shotgun sequences used for retrieving amphioxus IODs as predicted by Genscan are indicated as well as the positions of the SECIS elements and their scores as given by SECISearch

teins (Bianco et al. 2002), have a selenocysteine, encoded by a TGA codon, which is key to the activity of the protein (Bianco et al. 2002) (hence the name selenoprotein). In the amphioxus sequence Bf_123596, a cysteine encoded by a TGC codon is located at the homologous position to the selenocysteine of vertebrates. This is unlikely to be due to sequencing error as the same codon was found in a sequence obtained from an independent cDNA library (accession number BW699364). This difference casts doubts on the functionality of the Bf_123596 protein as an amphioxus deiodinase.

As the TGA codon encoding the selenoprotein is predicted as a stop codon and not a selenocysteine by gene prediction software, genes encoding selenoproteins would be inappropriately predicted as pseudogenes (with a precocious stop codon) or the protein would be predicted as truncated of the part around the amino acid (with the TGA codon predicted as intronic). In order to test this hypothesis, we performed a tblastn search on the shotgun sequence of the amphioxus genome, using the amino acid sequence of the rat IOD2 as a query to avoid the gene prediction problem. Using this method, we retrieved several genes, some of which included a TGA codon at the position corresponding to the selenocysteine in the query sequence. We were able to attribute all of the new sequences to previously found truncated amphioxus sequences (Fig. 4a, the sequences were renamed IODα to IODε). All newly identified sequences have an activation domain that is very well conserved with the activation domains of vertebrate sequences (Fig. 4b). In particular, all but the sequence Bf_123596 (corresponding to IODβ in Fig. 4) have a TGA codon at the "selenocysteine position". To confirm our hypothesis that at least some of the amphioxus sequences are selenoproteins, we searched for selenocysteine insertion sites (SECIS) known to be required for proper translation of the TGA codon into a selenocysteine (Bianco et al. 2002). SECIS were predicted in the putative 3′UTRs of IODα, IODγ, and IODε using the SECISearch program (Kryukov et al. 2003) (Fig. 4c–f). A histidine residue also important for deiodinase activity (Bianco et al. 2002) is conserved in all sequences. We hence predict that IODα, IODγ, and IODε are selenoproteins, and, taken together, these data strongly suggest that there are several active deiodinases in amphioxus.

A similar analysis was performed with the sea urchin genome (*S. purpuratus*) and one sequence was found containing both the activation domain with a putative senelocysteine and another partial sequence, for which the sequence corresponding to the activation domain was not available (Fig. 4a,b). The corresponding 3′UTR sequence was not available. A single gene has previously been described in the urochordate species *H. roretzi* as being a functional deiodinase (Shepherdley et al. 2004).

The monophyly of the three vertebrate deiodinases was recovered in our phylogenetic analysis (Fig. 4a). However, none of the non-vertebrate sequences could be precisely located in the phylogenetic tree with high bootstrap support. They branched at the base of the three vertebrate deiodinase families (IOD1, IOD2, and IOD3), suggesting that at the origin of deuterostomes, there was only one deiodinase that was duplicated independently in the sea urchin, amphioxus, and vertebrate lineages (Fig. 4a).

Genes involved in TH synthesis regulation

We searched the amphioxus genome for genes orthologous to CRF, TRH, TSH, and their respective receptors. Homologs to vertebrate CRFR, TSHR were found (Figs. S7 and S8). An orthologous sequence to the preproTRH was also found (Fig. S9). Interestingly, it contains only one canonical progenitor sequence (Glu-His-Pro-Gly) (in comparison to about five in mammals and birds, Vandenborne et al. 2005 and references therein), but also 21 sequences with the His replaced with a Ser (Glu-Ser-Pro-Gly). The flanking sequence (Lys-Arg) is conserved. No orthologs to TRH receptor, CRF, and TSH were found in the amphioxus genome. As a gene orthologous to CRH was found in insects and since the corresponding protein is rather small, it is likely that CRF is present in amphioxus, but not in the genome sequence.

Protein sequences involved in TH signaling in basal vertebrates

In order to gain insights into the evolution of the genes encoding proteins implicated in TH signaling and as these genes are mostly known from classical vertebrate models (human, mouse, zebrafish, and *Xenopus*), we searched for them in databases of basal vertebrates (e.g. lamprey and cartilaginous fishes). The sequences of two relevant genomes are publicly available and sufficiently annotated to allow a systematic search: the lamprey *P. marinus* (http://genome.ucsc.edu/cgi-bin/hgGateway?clade=other&org=Lamprey&db=) and the elephant shark *C. milii* (http://esharkgenome.imcb.a-star.edu.sg/). The dataset was completed with EST databases. The gene families we found in amphioxus are also present in cartilaginous fishes: SIS (Fig. S1), PERT (Fig. S2), IOD (Fig. S3), TBG (Fig. S4b), CTHBP (Fig. S5b), CRFR (Fig. S7b), TSHR (Fig. S8b), and TRH (Fig. S9). With the exception of IOD and TRH (which is a short protein), we also found representatives of each gene family in *P. marinus*. In addition, TR and RXR have previously been identified in both *P. marinus* and in the shark *Scyliorhinus canicula* (dogfish) (Escriva et al. 2002; Paris et al. 2008). As in amphioxus, we did not find TG in these genomes.

The predicted protein sequences could not be placed reliably in the context of phylogenetic trees. Given the low coverage of the two genomes (1.4 and 5.9, respectively, for *C. milii* and *P. marinus*) and the small size of DNA fragments in the EST and genome databases (less than 3 kb on average), only partial sequences could be retrieved and probably did not provide enough signal to be well positioned in our phylogenetic trees. Nevertheless, these data confirm the presence of the TH signaling pathway in basal vertebrates.

Discussion

An ancestral origin of TH signaling pathway in chordates

In this analysis, we have shown that the genome of the amphioxus *B. floridae* contains genes coding for proteins homologous to most of the genetic equipment necessary to endogenously produce thyroid hormones. SIS, PERT, deiodinases, cytosolic binding proteins and specific nuclear receptors are present in the genome (Fig. 1b). In addition, biochemical reactions necessary for TH synthesis are apparently similar in amphioxus and vertebrates (see the Introduction for more details). Many genes were also found in cartilaginous fishes, lampreys, and urochordates. Consequently, we propose that the TH signaling cascades of extant chordates are homologous, *i.e.* they evolved from an ancestral cascade that was present in the common ancestor of all chordates. Nevertheless, several key components of the pathway are missing (e.g. thyroglobulin) or are different (e.g. independent duplications) in amphioxus when compared to vertebrates.

In all chordates studied so far, THs were shown to regulate some features of post-embryonic development, the most spectacular example being metamorphosis (in amphibians (Tata 2006), in mammals (Flamant and Samarut 2003), in teleost fishes (Power et al. 2001), in lamprey (Manzon et al. 2001), in urochordates (Patricolo et al. 2001), and also in amphioxus (Paris et al. 2008)). The ancestry of the TH signaling pathway suggests that already in the chordate ancestor, metamorphosis was regulated by THs, which were synthesized endogenously. Interestingly, lamprey metamorphosis is inhibited by THs suggesting a specific elaboration of the TH signaling pathway in lampreys when compared to other chordates (Youson 1997).

Evolution of members of the TH signaling pathway

In most cases, several amphioxus sequences branch at the base of a group of several vertebrate sequences. The topology, where an amphioxus group corresponds to several vertebrate paralogs points to gene duplications at the base of the vertebrate lineage, probably due to the two rounds of whole genome duplication (WGD) events that occurred at the base of the vertebrate clade (Dehal and Boore 2005; Putnam et al. 2008). This pattern was observed repeatedly throughout the analysis of the amphioxus genome (about 25% of chordate gene families follow this scenario (Holland et al. 2008; Putnam et al. 2008), the other gene families having most probably undergone loss (Lynch and Conery 2000)). Accordingly, several paralogous genes were found for basal vertebrates (Figs. S1–S9). However, because of the limited genomic resources at this taxonomic position, only partial sequences were retrieved. Therefore, they could not be reliably placed in our phylogenetic trees.

Gene and genome duplications have been proposed to be substantial sources of new genes (Ohno 1970). After duplication, two identical genes would encode proteins able to perform the same function. This functional redundancy would free one of the two copies that could increase its mutation rate (Force et al. 1999). Several scenarios have been proposed: proteins encoded by both copies may share parts of the original function (subfunctionalization), one of the proteins may fulfill the ancestral function while the other one either degenerates (non-functionalization) or gains new functions (neofunctionalization). The study of the evolutionary rate of the peroxidase proteins is in agreement with the last point. Indeed, within the thyroid/myeloid/eosinophil/lymphoid peroxidase (PERT/PERM/PERE/PERL) family, the thyroid peroxidases PERTs, may carry the most ancestral function of the peroxidase quartet. PERTs are evolving significantly more slowly than the three other groups (see Table S3, this low evolutionary rate is illustrated by their short branches in Fig. 3) and the closely related invertebrate chordate sequences (amphioxus and urochordate) also have short branches indicating low mutation rates. Accordingly, thyroid peroxidase activity was reported in lampreys, urochordates, and amphioxus (see Paris and Laudet 2008 for a review). Interestingly, a duplication occurred specifically in the *C. intestinalis* lineage (Fig. 3), and the characterized gene (PERT1) displays only partial functional redundancy with peroxidase activity in the *C. intestinalis* endostyle (Ogasawara et al. 1999), possibly suggesting a partition of PERT function between the two proteins because of a subfunctionalization event (Markov et al. 2008). Further characterization of PERT2 will give further insights into the evolution of thyroid peroxidase activity in *C. intestinalis*. We propose that in the chordate ancestor, there was one peroxidase gene, which was a thyroid peroxidase (or more appropriately called an endostyle peroxidase). During subsequent lineage-specific evolution, this gene duplicated several times independently in the

vertebrate, *C. intestinalis* and in the amphioxus lineage. Among the four amphioxus retrieved sequences, Bf_67515 is the best candidate for encoding an active PERT. Indeed, PERTs are the only transmembrane peroxidases and Bf_67515 is the only amphioxus sequence harboring a putative transmembrane domain (as predicted by TMpred) located at the same position as in vertebrate PERTs (sites 752–776 corresponding to the sites 847 to 871 in the human PERT). Moreover, peroxidase activity has been located in the endostyle of amphioxus larvae in the outer surface of plasma membranes as well as to the inner surface of membranes in cytoplasmic compartments (Fredriksson et al. 1985). Additionally, the orthologous sequence from *B. belcheri* (Fig. 3) is expressed in the thyroid gland homolog, the endostyle during development (Ogasawara 2000). However, it does not contain the transmembrane domain. Whether this corresponds to an alternative splicing or a genomic difference needs further investigation.

Several genes were not retrieved in our analysis whereas biochemical studies suggested a different result. The fact that no gene related to thyroglobulin was found in amphioxus, whereas all the other key members of the signaling pathway are present, is not surprising because thyroglobulin is a long and divergent protein in which only some tyrosines are implied in TH production. A biochemical study revealed the existence of a protein harboring classical thyroglobulin properties in the amphioxus endostyle: a large protein that incorporates iodines through peroxidase activity to produce T_3 and T_4 (Monaco et al. 1981). It is possible that a protein non-homologous with vertebrate thyroglobulins is the source of tyrosine and the amphioxus sequence Bf_123169 is an interesting candidate. The purification of the protein detected by Monaco et al. (1981) and the further cloning of the corresponding gene as well as the cloning of similar proteins in other lineages (lampreys, urochordates, sea urchins) may help to resolve the issue of iodinated tyrosine sources in invertebrates. Similarly, no direct ortholog of TH transporters was found in the amphioxus genome. Only sequences distantly related to TBG were retrieved (no ortholog of TTR or albumin was found). Nonetheless, THs have been shown to be produced in amphioxus (Covelli et al. 1960) and considering the high hydrophobicity of THs, it is most probable that different carriers are involved in TH transport in amphioxus, but their exact nature remains unclear. This illustrates limitations to the approach we chose for studying TH signaling: phylogenetic studies only allow the detection of possible candidates that carry out a biological function known from vertebrates.

The case of deiodinases is interesting with respect to genome annotation. We found several amphioxus sequences that are most probably bona fide deiodinases. However, based on the genome annotation, the predicted sequences lack the activation domain because of a TGA codon very likely wrongly annotated as a stop codon instead of a selenocysteine codon that is always present in vertebrate deiodinase genes (Bianco et al. 2002). By manually annotating the amphioxus genome, we discovered several sequences that are likely to be active deiodinases in amphioxus. We propose that the ancestral deiodinase displayed a T_4-to-T_3 outer-ring deiodination activity because (1) T_4-to-T_3 production probably occurs in amphioxus (Covelli et al. 1960) and (2) the *H. roretzi* deiodinase was shown to have such an activity (Shepherdley et al. 2004). Of course, this hypothesis will require functional evidence with the biochemical characterization of the deiodinases retrieved in this study (especially IODα, IODγ, and IODε) (Fig. 4).

An upper and neuroendocrine regulation of TH production seems to have appeared only recently in the vertebrate lineage. We found only a few genes (TRH, TSHR, and CRFR) in the amphioxus genome that are implicated in a higher regulation of TH production. Orthologs of these genes performing functions not related to TH signaling exist in protostomes questioning the physiological role of the amphioxus sequences we retrieved. Based on our data, and since there is no clear hypothalamic pituitary thyroid axis in amphioxus (Holland et al. 2008), we can conclude that there is probably no homologous higher regulation of TH synthesis in amphioxus comparable to that of vertebrates. It is possible that the ability of a higher regulation by this axis evolved specifically in the vertebrate lineage. In contrast, a TH regulatory system different from the one in vertebrates may exist in amphioxus, since metamorphosis is regulated by THs in amphioxus (Paris et al. 2008), and since environmental conditions probably influence the onset of metamorphosis (crowded animals metamorphose more slowly than animals kept out of close contact from each other, M.P. unpublished observation). A similar trend is observed in anurans, for which stress situations influence TH production through alteration of CRH, TRH, and TSH production (Denver 1997).

Although the upper TH synthesis regulation seems to be divergent in amphioxus, TR, the receptor of TH in the TH signaling pathway, is well conserved within chordates (Schubert et al. 2008). Indeed, the only amphioxus TR (Fig. S6) is responsive to TRIAC, a T_3 derivative, and regulates amphioxus metamorphosis (Paris et al. 2008). Its partner, RXR, also displays functional characteristics that are very well conserved (as a heterodimer partner of several NRs (Schubert et al. 2008)). The functional conservation of TR and RXR not only within chordates, but even in the last deuterostome ancestor, is a plausible hypothesis: the genome of the sea urchin *S. purpuratus* contains only one TR and one RXR (Howard-Ashby et al. 2006), but these two NRs have not been functionally characterized yet.

Specific elaboration of the TH signaling pathway in amphioxus

As discussed above, the amphioxus genome did not undergo the two WGD events that occurred specifically during vertebrate evolution (Dehal and Boore 2005; Putnam et al. 2008). In contrast, specific duplications of genes encoding members of the TH signaling pathway occurred in amphioxus. In the part of the pathway upstream of TR (TH metabolic pathway), we found eight SIS-related genes corresponding to three paralogous vertebrate genes, four PER genes corresponding to three vertebrate paralogs, and five deiodinase genes corresponding to the three deiodinases genes known in vertebrates. This is in sharp contrast to what is observed for the more downstream part of the pathway (*i.e.* the receptor part): there is only one TR (for two vertebrate genes), one RXR (for three vertebrate genes), and only one copy for each of the co-activator/co-repressor paralogous groups (Schubert et al. 2008). This feature suggests that an independent elaboration of the upper part (the TH metabolic pathway) occurred in amphioxus, whereas the lower part (the receptor part) is similar in amphioxus and vertebrates. The reason why such an elaboration occurred remains elusive and will certainly need the precise functional characterization of these duplicates to be fully understood. It is tempting to link it to the central role that TR (and the TR cofactors) play in the TH signaling pathway and which could constrain its evolution, whereas TH availability could evolve more easily (see for instance the plasticity of TH metabolism in amphibians (Boorse and Denver 2002; Callery et al. 2001; Safi et al. 2004)). In each species, TH availability could independently evolve depending on specific selection pressures, which could be reflected in the lineage-specific duplications that are reported here. For instance, the active TH in amphioxus is not T_3, but probably the poorly characterized T_3 derivative TRIAC (Paris et al. 2008). This change may reflect a series of specific alterations in the metabolic pathway controlling TH production. It would thus be very interesting to study, if the specific amphioxus duplicates of deiodinases are differentially involved in the regulation of the iodine content of T_4 and its derivatives. Two main pieces of information are needed to test this model: (1) biochemical characterization of the amphioxus duplicates of SIS, PER, and deiodinases and (2) knowledge of the TH derivatives that are present, and active, in amphioxus. We are currently addressing these issues experimentally in our laboratory. Nonetheless, the amphioxus-specific duplications described here allow us to point out that amphioxus is not "our ancestor", as it is often more or less implied in many gradualist views of chordate evolution, but rather a cousin.

Is the TH signaling pathway conserved in deuterostomes?

TH metabolism and metamorphosis have both been present at the origin of deuterostomes (Paris and Laudet 2008). In chordates, TH signaling seems to be homologous (as mentioned at the beginning of the Discussion). Data on echinoderms suggest an even more ancient origin of TH signaling: at the base of the deuterostome tree. THs are important inducers of metamorphosis in some echinoideas (sea urchins, sea biscuits, and sand dollars), asteroideas (sea stars), and ophiuoideas (brittle stars) (Hodin 2006 and

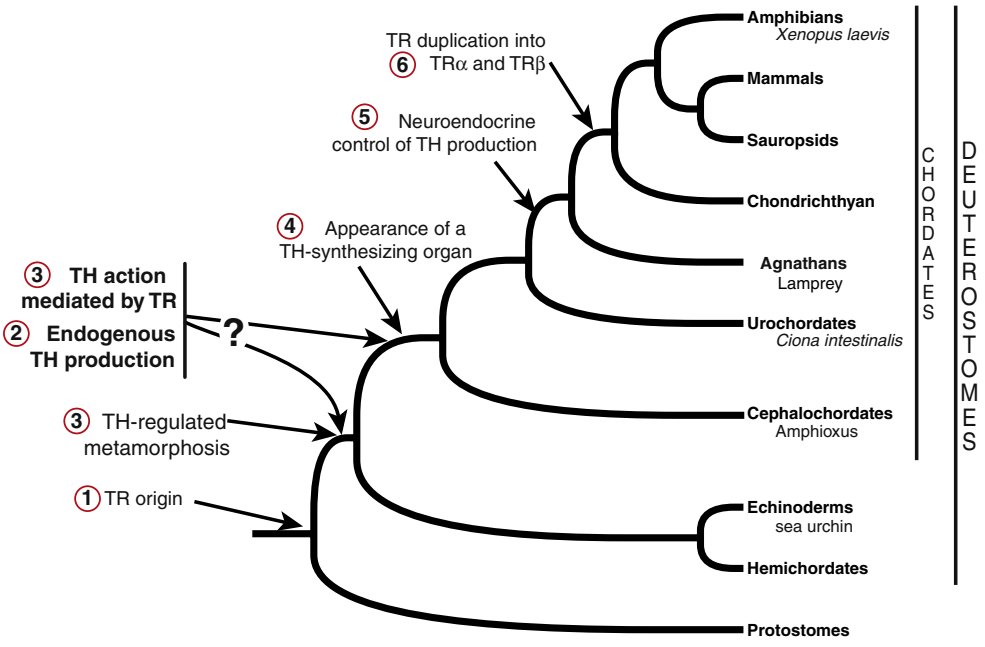


Fig. 5 Evolution of the thyroid signaling pathway. On a simplified bilaterian tree, the main steps of TH signaling pathway elaboration have been indicated. *Numbers* refer to chronological steps of the evolution of the TH signaling pathway in bilaterians and are discussed in the main text. The tree is based on Marletaz et al. (2008). Metamorphosis is to be understood in a broad sense as a TH/TR-regulated post-embryonic developmental phase characterized by ecological, metabolic, and morphological modifications (discussed in Paris et al. 2008)

references therein). Moreover, peroxidase activity-dependent T_4 production was also demonstrated (Heyland et al. 2006). We found in the sea urchin genome many genes orthologous to vertebrate TH signaling members. Consequently, it is very probable that the deuterostome ancestor was capable of endogenous TH production. However, as echinoderms do not have a recognizable endostyle/thyroid, it will be interesting to localize TH production.

There are still many deuterostome groups that have been neglected in terms of developmental and genomic studies. This is, for example, the case for hemichordates: almost nothing is known about TH signaling in this sister group of echinoderms (Fig. 5), although iodine was detected in the pharyngeal area (same body part as the endostyle/thyroid) (Ruppert 2005) supporting the notion of an ancestral origin of TH production in deuterostomes. Other deuterostome taxa that are poorly studied include sea stars, sea cucumber, crinoids as well as divergent urochordates, such as larvaceans and thaliaceans (Brusca and Brusca 2003).

The existence of a TH signaling pathway outside deuterostomes still remains elusive. Indeed, in protostomes or cnidarians, TH metabolism has been poorly studied (Eales 1997), few reports have investigated TH effects on metamorphosis, and, although TRs were cloned in several protostome species, they have not been molecularly characterized (Paris and Laudet 2008). We propose the following scenario regarding the evolution of the TH signaling pathway (Fig. 5): (1) appearance of a TR gene at the base of the bilaterians, whose function remains elusive; (2) the chordate, and likely even the deuterostome, ancestor acquired the ability to endogenously produce THs (TH production has been reported in several deuterostomes and the main members of the TH signaling pathway like SIS, PERT, IODs, TR, and some carrier proteins have been described in vertebrates, amphioxus, and, to a lesser extent, in echinoderms); (3) this process can probably be correlated with the ability to metamorphose under TH/TR regulation; (4) more complete and localized internalization of TH synthesis in a specialized organ with the evolution of a dual-function endostyle in chordates and its subsequent specialization into a thyroid within the vertebrate lineage; (5) in the vertebrate lineage, there is the appearance of a neuroendocrine control of TH synthesis using peptide hormones, such as TRH and TSH, that are not found in the amphioxus genome; (6) further elaboration of the pathway occurred in gnathostomes with the duplication of the ancestral TR into two genes, TRα and TRβ, allowing a further refinement of the pathway (there has also been a lamprey-specific TR duplication (Escriva et al. 2002)). According to this scenario, the downstream part of the pathway, namely the genes regulated by TH and TR, are derived in each species giving rise to the extraordinary diversity of morphological, physiological, and ecological rearrangements observed during metamorphosis (Paris and Laudet 2008).

Conclusions

From our study and previous data, we propose that the amphioxus TH signaling pathway is homologous to the vertebrate TH signaling pathway implying an ancient origin of TH metabolism. However, biochemical investigation on the proteins encoded by the genes described here should be carried out in the future. Outside chordates, much scarcer data are available. In some echinoderms, TH production and biological actions by TH are similar to what has been observed in chordates. Further work on echinoderms will be required to address questions such as where are THs produced and whether TR is involved in TH action in echinoderms. In order to better understand the evolution of the TH signaling pathway and its link to development and especially to metamorphosis, data on a wider range of animals should be obtained. Thus, even if a continuous effort should be maintained to keep improving our understanding of the TH signaling pathway in chordates, the sister group of chordates, the Ambulacraria (regrouping the echinoderms, the hemichordates, and xenoturbella (Marletaz et al. 2008)) and the protostomes are a "thyroidal desert" that may be worth our attention.

Acknowledgements We would like to thank Maria Theodosiou for critical reading of the manuscript. This work was supported by funds from the ANR, CNRS, and the MENRT. This study was further supported by CRESCENDO, a European Union Integrated Project of FP6, and by CASCADE, a Network of Excellence of FP6.

References

Baqui M, Botero D, Gereben B, Curcio C, Harney JW, Salvatore D, Sorimachi K, Larsen PR, Bianco AC (2003) Human type 3 iodothyronine selenodeiodinase is located in the plasma membrane and undergoes rapid internalization to endosomes. J Biol Chem 278:1206–1211

Berry MJ, Banu L, Larsen PR (1991) Type I iodothyronine deiodinase is a selenocysteine-containing enzyme. Nature 349:438–440

Bianco AC, Salvatore D, Gereben B, Berry MJ, Larsen PR (2002) Biochemistry, cellular and molecular biology, and physiological roles of the iodothyronine selenodeiodinases. Endocr Rev 23:38–89

Boorse GC, Denver RJ (2002) Acceleration of Ambystoma tigrinum metamorphosis by corticotropin-releasing hormone. J Exp Zool 293:94–98

Brusca RC, Brusca GJ (2003) Invertebrates. Sinauer Associates, Sunderland

Callery EM, Fang H, Elinson RP (2001) Frogs without polliwogs: evolution of anuran direct development. Bioessays 23:233–241

Castresana J (2000) Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis. Mol Biol Evol 17:540–552

Covelli I, Salvatore G, Sena L, Roche J (1960) Sur la formation d'hormones thyroidiennes et de leurs précurseurs par Branchiostoma lanceolatum. C R Soc Biol Paris 154:1165–1169

Curcio C, Baqui MM, Salvatore D, Rihn BH, Mohr S, Harney JW, Larsen PR, Bianco AC (2001) The human type 2 iodothyronine deiodinase is a selenoprotein highly expressed in a mesothelioma cell line. J Biol Chem 276:30183–30187

Daiyasu H, Toh H (2000) Molecular evolution of the myeloperoxidase family. J Mol Evol 51:433–445

Dehal P, Boore JL (2005) Two rounds of whole genome duplication in the ancestral vertebrate. PLoS Biol 3:e314

Denver RJ (1997) Environmental stress as a developmental cue: corticotropin-releasing hormone is a proximate mediator of adaptive phenotypic plasticity in amphibian metamorphosis. Horm Behav 31:169–179

Dohan O, De la Vieja A, Paroder V, Riedel C, Artani M, Reed M, Ginter CS, Carrasco N (2003) The sodium/iodide symporter (NIS): characterization, regulation, and medical significance. Endocr Rev 24:48–77

Eales JG (1997) Iodine metabolism and thyroid-related functions in organisms lacking thyroid follicles: are thyroid hormones also vitamins? Proc Soc Exp Biol Med 214:302–317

Edgar RC (2004) MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res 32:1792–1797

Escriva H, Manzon L, Youson J, Laudet V (2002) Analysis of lamprey and hagfish genes reveals a complex history of gene duplications during early vertebrate evolution. Mol Biol Evol 19:1440–1450

Flamant F, Samarut J (2003) Thyroid hormone receptors: lessons from knockout and knock-in mutant mice. Trends Endocrinol Metab 14:85–90

Flamant F, Baxter JD, Forrest D, Refetoff S, Samuels H, Scanlan TS, Vennstrom B, Samarut J (2006) International Union of Pharmacology. LIX. The pharmacology and classification of the nuclear receptor superfamily: thyroid hormone receptors. Pharmacol Rev 58:705–711

Force A, Lynch M, Pickett FB, Amores A, Yan YL, Postlethwait J (1999) Preservation of duplicate genes by complementary, degenerative mutations. Genetics 151:1531–1545

Fredriksson G, Öfverholm T, Ericson LE (1985) Electron-microscopic studies of iodine-binding and peroxidase activity in the endostyle of the larval amphioxus (Branchiostoma lanceolatum). Cell Tissue Res 241:257–266

Galtier N, Gouy M, Gautier C (1996) SEAVIEW and PHYLO_WIN: two graphic tools for sequence alignment and molecular phylogeny. Comput Appl Biosci 12:543–548

Gopal E, Umapathy NS, Martin PM, Ananth S, Gnana-Prakasam JP, Becker H, Wagner CA, Ganapathy V, Prasad PD (2007) Cloning and functional characterization of human SMCT2 (SLC5A12) and expression pattern of the transporter in kidney. Biochim Biophys Acta 1768:2690–2697

Guindon S, Gascuel O (2003) A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. Syst Biol 52:696–704

Heyland A, Hodin J (2004) Heterochronic developmental shift caused by thyroid hormone in larval sand dollars and its implications for phenotypic plasticity and the evolution of nonfeeding development. Evolution Int J Org Evolution 58:524–538

Heyland A, Reitzel AM, Price DA, Moroz LL (2006) Endogenous thyroid hormone synthesis in facultative planktotrophic larvae of the sand dollar Clypeaster rosaceus: implications for the evolutionary loss of larval feeding. Evol Dev 8:568–579

Hodin J (2006) Expanding networks: Signaling components in and a hypothesis for the evolution of metamorphosis. Integr Comp Biol 46:719–742

Holland LZ, Albalat R, Azumi K, Benito-Gutierrez E, Blow MJ, Bronner-Fraser M, Brunet F, Butts T, Candiani S, Dishaw LJ, Ferrier DE, Garcia-Fernandez J, Gibson-Brown JJ, Gissi C, Godzik A, Hallbook F, Hirose D, Hosomichi K, Ikuta T, Inoko H, Kasahara M, Kasamatsu J, Kawashima T, Kimura A, Kobayashi M, Kozmik Z, Kubokawa K, Laudet V, Litman GW, McHardy AC, Meulemans D, Nonaka M, Olinski RP, Pancer Z, Pennacchio LA, Pestarino M, Rast JP, Rigoutsos I, Robinson-Rechavi M, Roch G, Saiga H, Sasakura Y, Satake M, Satou Y, Schubert M, Sherwood N, Shiina T, Takatori N, Tello J, Vopalensky P, Wada S, Xu A, Ye Y, Yoshida K, Yoshizaki F, Yu JK, Zhang Q, Zmasek CM, de Jong PJ, Osoegawa K, Putnam NH, Rokhsar DS, Satoh N, Holland PW (2008) The amphioxus genome illuminates vertebrate origins and cephalochordate biology. Genome Res 18:1100–1111

Howard-Ashby M, Materna SC, Brown CT, Chen L, Cameron RA, Davidson EH (2006) Gene families encoding transcription factors expressed in early development of Strongylocentrotus purpuratus. Dev Biol 300:90–107

Hulbert AJ (2000) Thyroid hormones and their effects: a new perspective. Biol Rev Camb Philos Soc 75:519–631

Kryukov GV, Castellano S, Novoselov SV, Lobanov AV, Zehtab O, Guigo R, Gladyshev VN (2003) Characterization of mammalian selenoproteomes. Science 300:1439–1443

Lynch M, Conery JS (2000) The evolutionary fate and consequences of duplicate genes. Science 290:1151–1155

Manzon RG, Holmes JA, Youson JH (2001) Variable effects of goitrogens in inducing precocious metamorphosis in sea lampreys (Petromyzon marinus). J Exp Zool 289:290–303

Manzon RG, Neuls TM, Manzon LA (2007) Molecular cloning, tissue distribution, and developmental expression of lamprey transthyretins. Gen Comp Endocrinol 151:55–65

Markov GV, Paris M, Bertrand S, Laudet V (2008) The evolution of the ligand/receptor couple: a long road from comparative endocrinology to comparative genomics. Mol Cell Endocrinol 293:5–16

Marletaz F, Gilles A, Caubit X, Perez Y, Dossat C, Samain S, Gyapay G, Wincker P, Le Parco Y (2008) Chaetognath transcriptome reveals ancestral and unique features among bilaterians. Genome Biol 9:R94

Monaco F, Dominici R, Andreoli M, Pirro RD, Roche J (1981) Thyroid hormone formation in thyroglobulin synthesized in the amphioxus (Branchiostoma lanceolatum Pallas). Comp Biochem Physiol B 70:341–343

Novinec M, Kordis D, Turk V, Lenarcic B (2006) Diversity and evolution of the thyroglobulin type-1 domain superfamily. Mol Biol Evol 23:744–755

Ogasawara M (2000) Overlapping expression of amphioxus homologs of the thyroid transcription factor-1 gene and thyroid peroxidase gene in the endostyle: insight into evolution of the thyroid gland. Dev Genes Evol 210:231–242

Ogasawara M, Di Lauro R, Satoh N (1999) Ascidian homologs of mammalian thyroid peroxidase genes are expressed in the thyroid-equivalent region of the endostyle. J Exp Zool 285:158–169

Ohno S (1970) Evolution by gene duplication. Springer, Berlin

Paris M, Escriva H, Schubert M, Brunet F, Brtko J, Ciesielski F, Roecklin D, Vivat-Hannah V, Jamin EL, Cravedi JP, Scanlan TS, Renaud JP, Holland ND, Laudet V (2008) Amphioxus postembryonic development reveals the homology of chordate metamorphosis. Curr Biol 18:825–830

Paris M, Laudet V (2008) The history of developmental stages: metamorphosis in chordates. Genesis, in press

Patricolo E, Cammarata M, D'Agati P (2001) Presence of thyroid hormones in ascidian larvae and their involvement in metamorphosis. J Exp Zool 290:426–430

Power DM, Llewellyn L, Faustino M, Nowell MA, Bjornsson BT, Einarsdottir IE, Canario AV, Sweeney GE (2001) Thyroid hormones in growth and development of fish. Comp Biochem Physiol C Toxicol Pharmacol 130:447–459

Putnam NH, Butts T, Ferrier DE, Furlong RF, Hellsten U, Kawashima T, Robinson-Rechavi M, Shoguchi E, Terry A, Yu JK, Benito-Gutierrez EL, Dubchak I, Garcia-Fernandez J, Gibson-Brown JJ, Grigoriev IV, Horton AC, de Jong PJ, Jurka J, Kapitonov VV, Kohara Y, Kuroki Y, Lindquist E, Lucas S, Osoegawa K, Pennacchio LA, Salamov AA, Satou Y, Sauka-Spengler T, Schmutz J, Shin IT, Toyoda A, Bronner-Fraser M, Fujiyama A, Holland LZ, Holland PW, Satoh N, Rokhsar DS (2008) The amphioxus genome and the evolution of the chordate karyotype. Nature 453:1064–1071

Robinson M, Gouy M, Gautier C, Mouchiroud D (1998) Sensitivity of the relative-rate test to taxonomic sampling. Mol Biol Evol 15:1091–1098

Rodriguez AM, Perron B, Lacroix L, Caillou B, Leblanc G, Schlumberger M, Bidart JM, Pourcher T (2002) Identification and characterization of a putative human iodide transporter located at the apical membrane of thyrocytes. J Clin Endocrinol Metab 87:3500–3503

Ruppert EE (2005) Key character uniting hemichordates and chordates: homologies or homoplasies. Can J Zool 83:8–23

Safi R, Bertrand S, Marchand O, Duffraisse M, de Luze A, Vanacker JM, Maraninchi M, Margotat A, Demeneix B, Laudet V (2004) The axolotl (Ambystoma mexicanum), a neotenic amphibian, expresses functional thyroid hormone receptors. Endocrinology 145:760–772

Schreiber G, Richardson SJ (1997) The evolution of gene expression, structure and function of transthyretin. Comp Biochem Physiol B Biochem Mol Biol 116:137–160

Schubert M, Brunet F, Paris M, Bertrand S, Benoit G, Laudet V (2008) Nuclear hormone receptor signaling in amphioxus. Dev Genes Evol, in press

Shepherdley CA, Klootwijk W, Makabe KW, Visser TJ, Kuiper GGJM (2004) An ascidian homolog of vertebrate iodothyronine deiodinases. Endocrinology 145:1255–1268

Shi Y-B (2000) Amphibian metamorphosis: from morphology to molecular biology. Wiley, New York

Tata JR (2006) Amphibian metamorphosis as a model for the developmental actions of thyroid hormone. Mol Cell Endocrinol 246:10–20

Vandenborne K, Roelens SA, Darras VM, Kuhn ER, Van der Geyten S (2005) Cloning and hypothalamic distribution of the chicken thyrotropin-releasing hormone precursor cDNA. J Endocrinol 186:387–396

Wu SY, Green WL, Huang WS, Hays MT, Chopra IJ (2005) Alternate pathways of thyroid hormone metabolism. Thyroid 15:943–958

Yen PM (2001) Physiological and molecular basis of thyroid hormone action. Physiol Rev 81:1097–1142

Youson JH (1997) Is lamprey metamorphosis regulated by thyroid hormones? Amer Zool 37:441–460

7

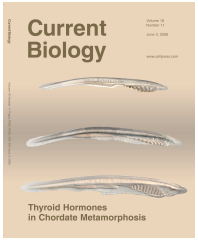
AMPHIOXUS POST-EMBRYONIC DEVELOPMENT REVEALS THE HOMOLOGY OF CHORDATE METAMORPHOSIS

As explained in chapter **??** in the introduction, metamorphoses display a wide variability among chordates, from undetectable, as in mammals, to an extensive tissue remodeling, like in tunicates. This variability suggests that metamorphosis appeared several times independently during evolution. However, the molecular determinism of metamorphosis was well established only in some amphibians, and to a lesser extent in teleost fishes, where THs were shown to be the main players (see chapter **??** in the introduction).

In order to gain better insight into the evolution of metamorphosis in chordates, we sought into the molecular determinism of this post-embryonic developmental stage, by focusing on the amphioxus animal model. Indeed, as explained in chapter **??**, amphioxus both branches at the base of the chordate tree and has evolved relatively slowly since then. It is thus a favorable model to digging into the origins of TH effects on metamorphosis in chordates. However, very few experiments had been conducted on amphioxus before, because of the unavailability of larvae at a late developmental stage. This article reports the first large scale study of the role of THs during amphioxus metamorphosis.

The work described in this section illustrates the difficulty in determining homology for anciently derived or fast evolving developmental processes. Here the morphological component of the evolution of metamorphosis may be underlain by its molecular determinism, where some parts of the regulatory pathway are better conserved than others. See the general discussion (chapter **??**) for further details.

This article was published in Current Biology in June 2008. Please notice that supplemental data are presented after the main article to help understand the main article.



Current Biology *18*, 825–830, June 3, 2008 ©2008 Elsevier Ltd All rights reserved DOI 10.1016/j.cub.2008.04.078

Report

Amphioxus Postembryonic Development Reveals the Homology of Chordate Metamorphosis

Mathilde Paris,[1] Hector Escriva,[2] Michael Schubert,[1] Frédéric Brunet,[1] Julius Brtko,[3] Fabrice Ciesielski,[4] Dominique Roecklin,[4] Valérie Vivat-Hannah,[4] Emilien L. Jamin,[5] Jean-Pierre Cravedi,[5] Thomas S. Scanlan,[6] Jean-Paul Renaud,[4] Nicholas D. Holland,[7] and Vincent Laudet[1,*]

[1]Institut de Génomique Fonctionnelle de Lyon
Université de Lyon
CNRS, INRA, UCB Lyon 1
IFR128 Lyon Biosciences
Ecole Normale Supérieure de Lyon
69364 Lyon Cedex 07
France

[2]Laboratoire Arago
UMR-7628, CNRS
University Pierre and Marie Curie
66651 Banyuls sur Mer
France

[3]Institute of Experimental Endocrinology
Slovak Academy of Sciences
833 06 Bratislava
Slovak Republic

[4]AliX
Boulevard Sébastien Brant
Parc d'Innovation
67400 Illkirch
France

[5]INRA, UMR-1089 Xenobiotiques
31931 Toulouse Cedex 9
France

[6]Department of Physiology and Pharmacology
Oregon Health and Science University
3181 SW Sam Jackson Park Road
Portland, Oregon 97239

[7]Marine Biology Research Division
Scripps Institution of Oceanography
University of California, San Diego
La Jolla, California 92093-0202

Summary

Most studies in evolution are centered on how homologous genes, structures, and/or processes appeared and diverged. Although historical homology is well defined as a concept, in practice its establishment can be problematic, especially for some morphological traits or developmental processes. Metamorphosis in chordates is such an enigmatic character. Defined as a spectacular postembryonic larva-to-adult transition, it shows a wide morphological diversity between the different chordate lineages, suggesting that it might have appeared several times independently. In vertebrates, metamorphosis is triggered by binding of the thyroid hormones (THs) T_4 and T_3 to thyroid-hormone receptors (TRs). Here we show that a TH derivative, triiodothyroacetic acid (TRIAC), induces metamorphosis in the cephalochordate amphioxus. The amphioxus TR (amphiTR) mediates spontaneous and TRIAC-induced metamorphosis because it strongly binds to TRIAC, and a specific TR antagonist, NH3, inhibits both spontaneous and TRIAC-induced metamorphosis. Moreover, as in amphibians, *amphiTR* expression levels increase around metamorphosis and are enhanced by THs. Therefore, TH-regulated metamorphosis, mediated by TR, is an ancestral feature of all chordates. This conservation of a regulatory network supports the homology of metamorphosis in the chordate lineage.

*Correspondence: vincent.laudet@ens-lyon.fr

Results and Discussion

Metamorphosis, when looked at morphologically, differs strikingly in chordates, which comprise urochordates (such as tunicates), cephalochordates (such as amphioxus), and vertebrates: In ascidian tunicates, all larval tissues are drastically remodeled into the adult animal; in amphioxus, the highly asymmetric larva transforms into a relatively symmetric adult; and, conversely, in some flatfish the symmetric larva becomes an asymmetric adult. This morphological diversity raises the question of whether the molecular mechanisms involved in metamorphosis are fundamentally different from one chordate group to the next or are conserved—at least partially, for instance in their triggering events. By now, numerous studies have identified thyroid hormones (THs) and TH receptor (TR) as key components of gene networks controlling metamorphosis in vertebrates [1]. In contrast, the functional roles of such molecules have been poorly studied in ascidians [2] and not at all in amphioxus, which, albeit being morphologically much more vertebrate-like than any urochordate, is now considered to be part of the most basal group in the phylum Chordata [3]. This phylogenetic position and its morphology make amphioxus an excellent model for the study of the origin of metamorphosis in chordates [4]. Therefore, we have been studying the roles of THs and TR on the metamorphosis of the Florida amphioxus (*Branchiostoma floridae*).

Premetamorphic amphioxus larvae are conspicuously asymmetric: For example, the mouth is on the left side and there is a single row of gill slits on the right side (Figure 1A, middle and top). Then, during metamorphosis, approximate bilateral symmetry is established by medial migration of the mouth and the production of a second row of gill slits (Figure 1A, bottom). We studied whether THs could trigger amphioxus metamorphosis by treating premetamorphic larvae in vivo (in batches of 50, Table S1 available online) with 10^{-8} M THs (T_3, T_4, or reverse T_3, a T_4 derivative almost inactive in vertebrates [1]). In early larvae, T_3 promoted metamorphosis with an early onset and a short duration (2.5 days as compared to about 7 days for control animals) (Figure 1B and Figure S1, as well as Table S1). In contrast, T_4 was less potent, because after 3 days of treatment only 45% of the treated larvae were undergoing metamorphosis, as compared to 98% of the T_3-treated larvae. Reverse T_3 was considerably less effective than either T_3 or T_4 in accelerating the onset and condensing the duration of metamorphosis (Figure 1B and Figure S1C). Because in vertebrates T_3 is mainly formed through deiodination

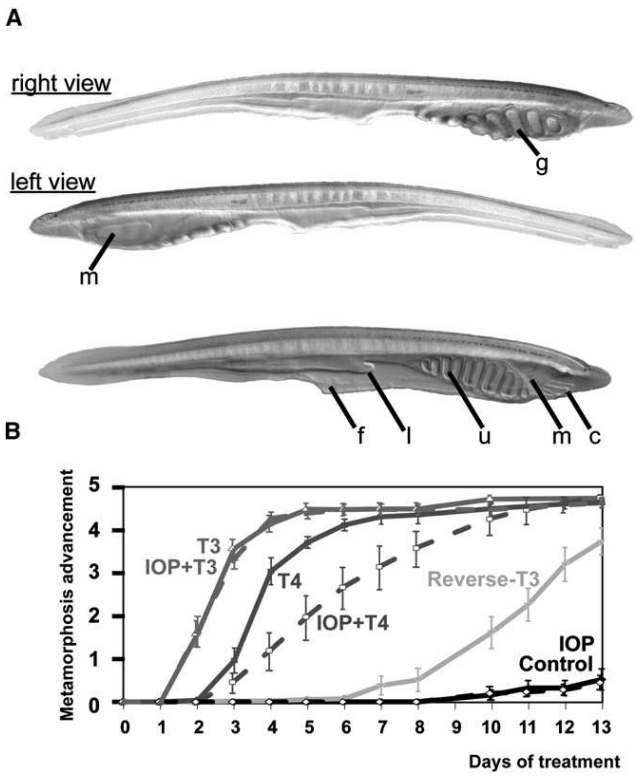


Figure 1. THs Induce Metamorphosis in Premetamorphic Amphioxus

(A) Premetamorphic amphioxus displays characteristic asymmetric features (top and middle) that disappear after metamorphosis (bottom). The following abbreviations are used: c, cirri; f, metapleural folds; g, gill slits; l, modification of the gut; m, mouth; and u, gill slits in U shape.

(B) Effects on metamorphosis of 12-day-old premetamorphic amphioxus larvae of T_3, T_4, and reverse T_3 at 10^{-8} M, complemented or not with IOP at 10^{-5} M, was monitored with five specific morphological criteria (fully described in the Supplemental Experimental Procedures). Error bars correspond to a 5% confidence interval.

of T_4, we tested the effect of a deiodinase inhibitor, iodopanoic acid (IOP), on T_3 and T_4 treatments. The addition of IOP partially inhibited metamorphosis in T_4-treated larvae but had no effect on animals treated with T_3 (Figure 1B). This result suggests that in amphioxus, as in vertebrates, metamorphosis is mainly induced by T_3 synthesized by deiodination of T_4.

The amphioxus endostyle is widely considered to be a homolog of the vertebrate thyroid gland [5, 6]. This organ derived from pharyngeal endoderm of the larva secretes mucoprotein into the gut lumen to help capture particulate food. Both T_3 and T_4 have been detected in the amphioxus endostyle [7–9], and we have confirmed these earlier results on T_3 on adult extracts both by high-performance liquid chromatography (HPLC) coupled to mass spectrometry (25 ng/g amphioxus with a relative standard deviation of 12%) and by radioimmunoassays (data not shown). Moreover, by searching the amphioxus genome, we have found orthologs for many of the vertebrate genes encoding proteins involved in the metabolism of THs—for example, peroxidases and deioninases (Figure S2 and Table S4). Taken together, these results suggest that there is an active TH signaling system in amphioxus. The available amphioxus genome sequence, however, does not include a gene orthologous to the gene encoding the vertebrate thyroglobulin protein (the source of the tyrosines of the thyronine skeleton of THs). This result is surprising because of an earlier biochemical identification of a thyroglobulin protein in amphioxus [9]. It remains possible that the gene was not detected in the amphioxus genome, because the coverage is not complete. Moreover, amphioxus, like many other metazoans, does have several genes corresponding to a subdomain (Tg1) of vertebrate thyroglobulin [10]. The amphioxus Tg1 domain proteins might be the source of the tyrosines required for TH synthesis; alternatively, however, such tyrosines might be derived from another protein(s).

Because vertebrate TH effects are mediated by TRs, we cloned the full-length cDNA of the single *TR* gene, *amphiTR* (Figure S3) to study the molecular mechanisms controlling amphioxus metamorphosis. The gene encodes a protein with a DNA-binding domain (DBD), which shares 74% sequence identity with human TRα (Figure S3B), whereas the ligand-binding domain (LBD) is only 38% identical. Not surprisingly, amphiTR binds to DNA on classical TR response elements (TREs) as either a homo- or heterodimer with the single retinoid X receptor (RXR) from amphioxus (Figure S5). In contrast, in the LBD of amphiTR, two of three residues directly contacting the ligand in human TRα are divergent [11] (Figure S3A).

We compared the transactivation activity of amphiTR with that of other chordate TRs—namely, TRα from the rat, two TRs from a lamprey-specific duplication [12] (Figure S4), and TR from the tunicate *Ciona intestinalis* [2]. In comparison to vertebrate TRs, those of amphioxus and *Ciona* cannot activate reporter-gene expression in transiently transfected mammalian cells after TH stimulation (Figure 2A, Figure S6). Moreover, we did not detect any T_3 binding in nuclear extracts from whole tissues of adult amphioxus (Figure 2D), whereas high-affinity binding was detected in cytosol extracts (data not shown). Because amphiTR is nuclear in transfected mammalian cells (Figure 2E), TR does probably not bind T_3 in vivo in amphioxus. This hypothesis was confirmed in vitro both by limited proteolysis and mass-spectrometry experiments (Figures 2B and 2C and Figure S7), where only a very weak binding of amphiTR to T_3 or T_4 was detected. These results indicate that TR does not bind T_3 in amphioxus (TR does not appear to bind T_3 in *Ciona* either; Figures 2A and 2B), whereas in vertebrates TRs bind T_3, thereby activating TR-dependent transcription (Figure 2A).

The induction of amphioxus metamorphosis by T_3 in the absence of T_3 binding to amphiTR is unexpected. We think that a likely explanation for this paradox may be that some other TH metabolite is active in amphioxus. We thus tested several TH derivatives for their ability to bind amphiTR. One of the compounds tested was triiodothyroacetic acid (TRIAC), which binds mammalian TRs effectively but is present at very low concentrations in mammals and is thus generally considered to have only a minor role in vertebrate TH signaling. In amphioxus, TRIAC was the only TH derivative tested that binds to and significantly activates amphiTR (Figures 3A, 3B, and 3C). Binding was fairly strong because 100% of amphiTR was in complex with TRIAC in mass-spectrometry experiments (Figure 3A). In amphioxus, the difference in amphiTR affinity for T_3 and TRIAC probably results from the divergence of two of the key amino acids contacting directly T_3 [11, 13] (Figure 3D), whereas the binding sites between TRIAC and TR are well conserved between amphioxus and mammals (Figure 3E). In transfected mammalian cells, 10^{-7} M TRIAC activates amphiTR (Figure 3C), and at an even lower concentration (10^{-8} M) TRIAC stimulates recruitment of the coactivator SRC-1 to amphiTR (Figure 3F). The absence of significant release of the corepressor NcoR in a heterologous mammalian system

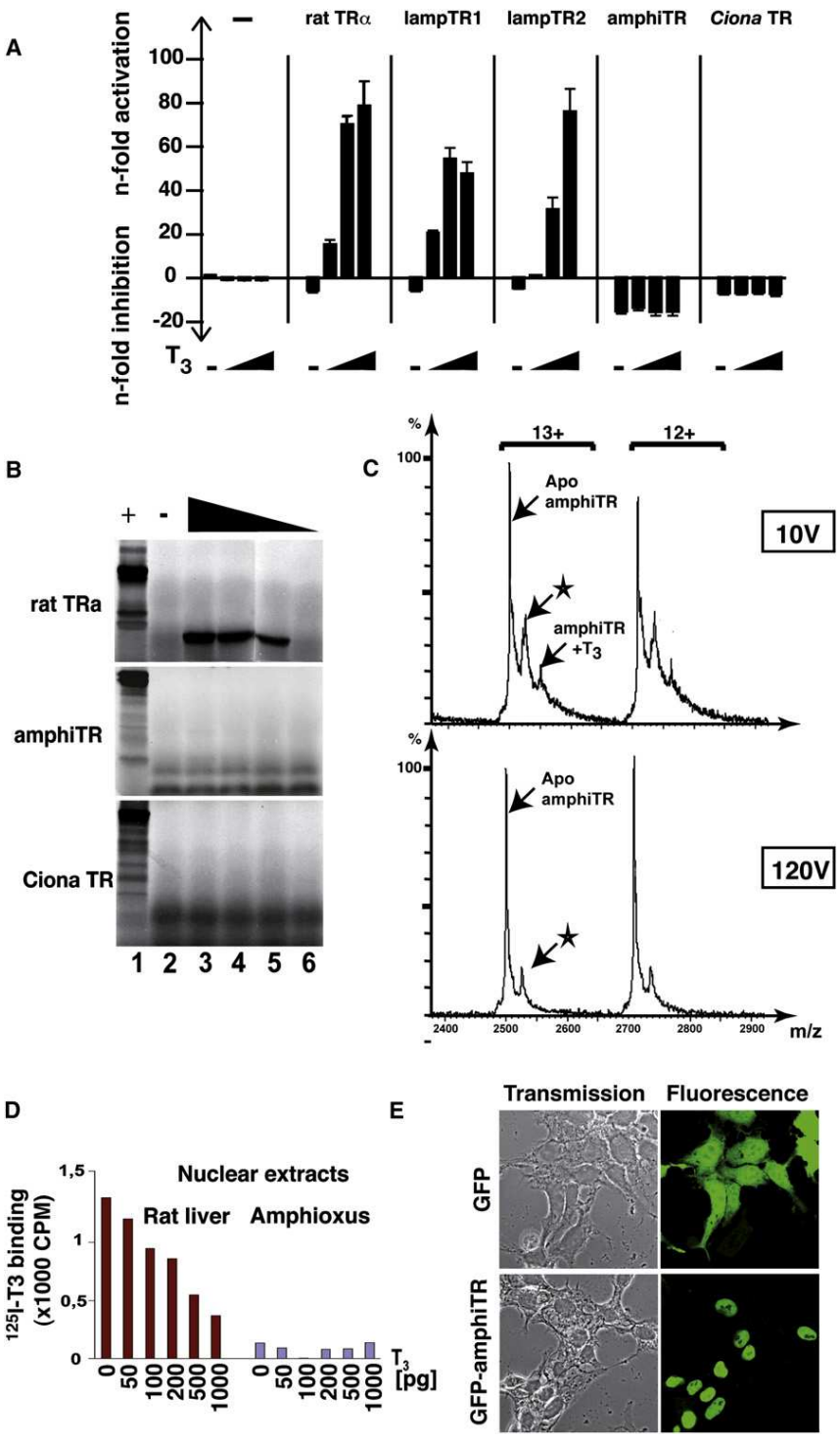


Figure 2. Amphioxus TR Is Not a T_3 Receptor and Is a Strong Constitutive Inhibitor of Transcription

(A) GAL4-LBD constructs from several chordate TRs were tested for their ability to activate a (17 m)5x-G-luc reporter plasmid in the presence of increasing doses of T_3 (10^{-8} M to 10^{-6} M). Error bars represent the standard error of the mean (SEM).

(B) Limited proteolysis of chordate TRs with T_3. Lane 1 shows undigested protein, and lanes 2–6 show digested protein in the absence (lane 2) or presence (lanes 3–6) of T_3 (10^{-3} M to 10^{-6} M).

(C) Nondenaturing electrospray-mass spectrometry (ESI-MS) analysis in the presence of T_3 either at 10V (permissive conditions) or at 120V (stringent conditions). At low Vc (10V), three species were detected and corresponded to the apo form of the amphiTR-LBD (apo), an unidentified adduct with a Dmass of 329 Da (star), and the T_3-bound amphiTR LBD (T_3). The ratio for all three species was 60%:25%:15%, respectively. Binding of T_3 was reversible: At high Vc (120V), the complex between amphiTR-LBD and T_3 was disrupted. Two protonation states are shown (13+ and 12+ charge states).

(D) Nuclear extracts of amphioxus adults were tested for their ability to bind radiolabeled T_3. Nuclear extracts of rat liver were used as a positive control (Kd = 0.95 nM).

(E) A GFP-amphiTR chimera was localized in the nucleus of transfected 293 cells by confocal microscopy.

(Figure 3F) probably explains the low level of amphiTR-mediated transcription in our experiments. We thus tested the effects of TRIAC on amphioxus metamorphosis and found that this compound, like T_3 and T_4, induces larvae to prematurely undergo metamorphosis (Figure 3G and Table S2). In sum, these data strongly suggest that TRIAC or a close derivative is the actual TR ligand in amphioxus. In vivo, a previous study reported the detection of TRIAC in amphioxus adult extracts [8]. For amphioxus larvae entering metamorphosis, it would be desirable to demonstrate a peak of TH production and to identify the active form of the hormone. However, because these larvae are minute in comparison to the adults, such measurements would require an enormous number of larvae and are impractical with current techniques. However, we predict that the physiological TH during amphioxus metamorphosis is a T_3 derivative. Our data and the results from Covelli et al. [8] strengthen the notion that TRIAC is a good candidate.

Previous work in vertebrates has shown that TH activation of TRβ and its subsequent autoinduction are transcriptional events triggering metamorphosis [14]. This raises the question of whether amphioxus metamorphosis is similarly triggered. We therefore tested whether THs influence the developmental expression of *amphiTR* in vivo. Our analysis shows that amphiTR mRNA expression is peaking around metamorphosis (Figure 4A) and is significantly induced by T_3 treatment of premetamorphic larvae and adults (Figures 4B and 4C) as well as by TRIAC treatment of adults (Figure 4C). In amphioxus, as in amphibians, we found a putative TR response element that is upstream of the coding sequence of *amphiTR* and is bound in vitro by the receptor (electromobility shift assay [EMSA] experiments in Figure S9). This suggests an involvement of amphiTR in TH-induced metamorphosis. In order to test this involvement, we studied the ability of NH3, a TR partial

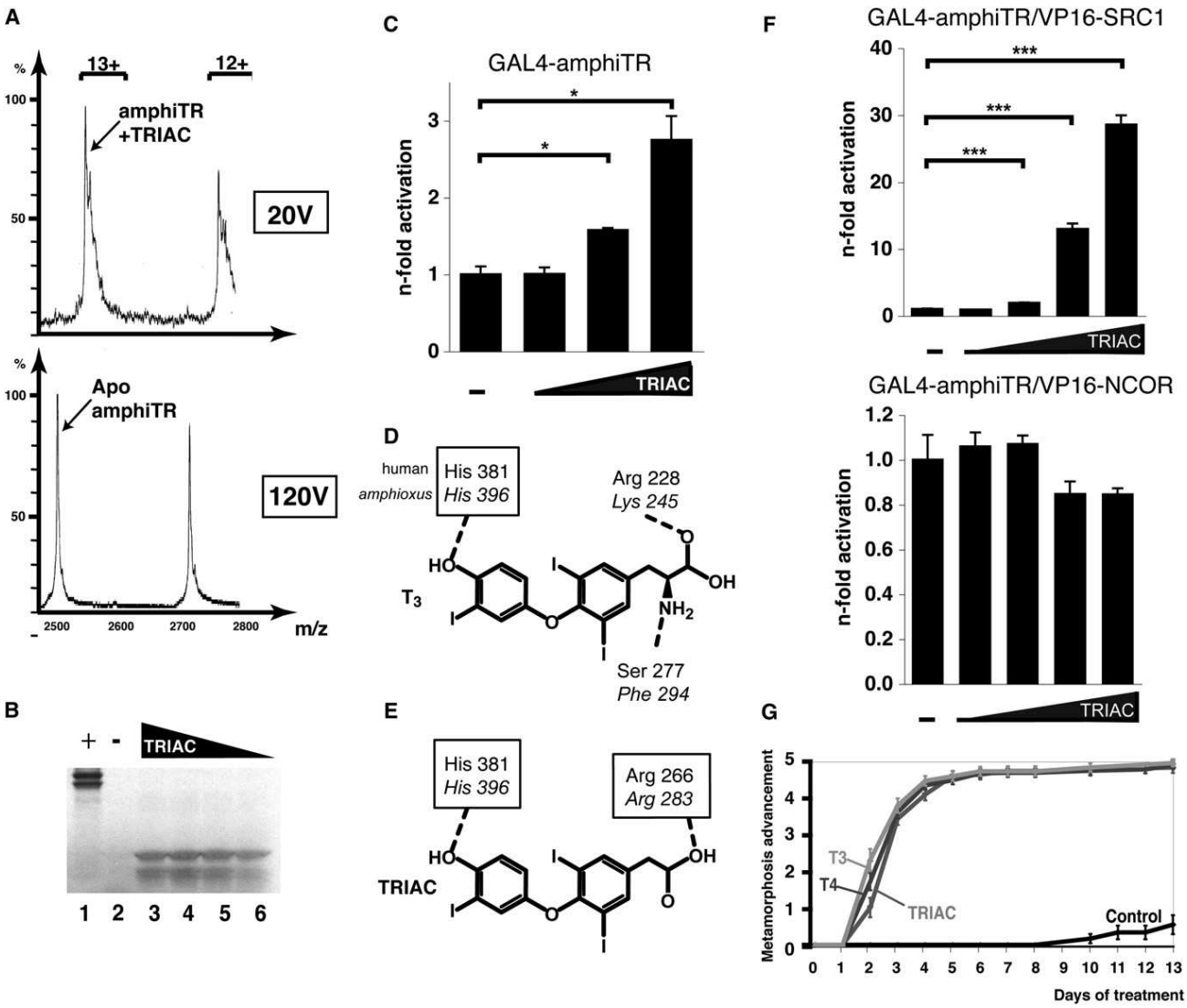


Figure 3. amphiTR Is a TRIAC Receptor

(A) Nondenaturing ESI-MS analysis in the presence of TRIAC binding to the amphiTR LBD. At low Vc (20V), 100% of amphiTR-LBD binds TRIAC. Increasing the Vc to 120V resulted in a complete dissociation of the complexes (in agreement with the reversible binding of TRIAC to the amphiTR LBD); the species detected corresponded to the apo amphiTR LBD. Two protonation states are shown (13+ and 12+ charge states).

(B) Limited proteolysis of amphiTR with TRIAC. Lane 1 shows undigested protein, and lanes 2–6 show digested protein in the absence (lane 2) or presence (lanes 3–6) of TRIAC (10^{-3} M to 10^{-6} M).

(C) The GAL4-amphiTR-LBD chimera was tested for its ability to activate a (17 m)5x-G-luc reporter plasmid in the presence of increasing doses of TRIAC (10^{-8} M to 10^{-6} M). Error bars represent the standard error of the mean (SEM).

(D and E) T_3 (D) and its acetic-acid derivative TRIAC (E) are shown surrounded by the residues known to make direct bonds with rat TRα [11, 13]. The amino acid at each site is shown for rat TRα (above) and amphiTR (below). Residues identical in amphiTR and rat TRα are boxed.

(F) Representation of the mammalian two-hybrid SRC1 and NCOR recruitment assay. The GAL4-amphiTR-LBD chimera was transfected either with the coactivator SRC1 or with the interaction domain of the corepressor NCOR fused to the strong activation domain VP16 in the presence of increasing doses of TRIAC (10^{-9} M to 10^{-6} M). Bonferroni corrected Student's t test: * $p < 0.05$; *** $p < 0.001$. Error bars represent the standard error of the mean (SEM).

(G) The effect of T_3, T_4, and TRIAC on the metamorphosis of 12-day-old premetamorphic amphioxus larvae was monitored. Treatments were carried out at 10^{-7} M. Error bars correspond to a 5% confidence interval.

antagonist [15], to inhibit spontaneous and TRIAC-induced metamorphosis. Interestingly, NH3 is known to inhibit amphibian metamorphosis [16]. We first checked that NH3 functions as an amphiTR antagonist: amphiTR binds NH3 in vitro (Figure S8A), and, in transfected mammalian cells, NH3 prevents TRIAC from activating amphiTR (Figure S8B) and from stimulating the recruitment of the coactivator SRC-1 to amphiTR (Figure S8C). In vivo NH3 was able to inhibit both spontaneous and TRIAC-induced metamorphosis (Figures 4D and 4E and Table S3). It was less efficient on early stages, according to morphological criteria (no conspicuous difference was detected until metamorphosis stage 2; Figures 4D and 4E), suggesting that less TH is required for early than for late metamorphosis, as in amphibians. Consequently this experiment shows that TR mediates TH-induced metamorphosis in amphioxus, as in amphibians. In sum, amphiTR is likely to play a key role in the TH-induced cascade triggering the metamorphosis of amphioxus.

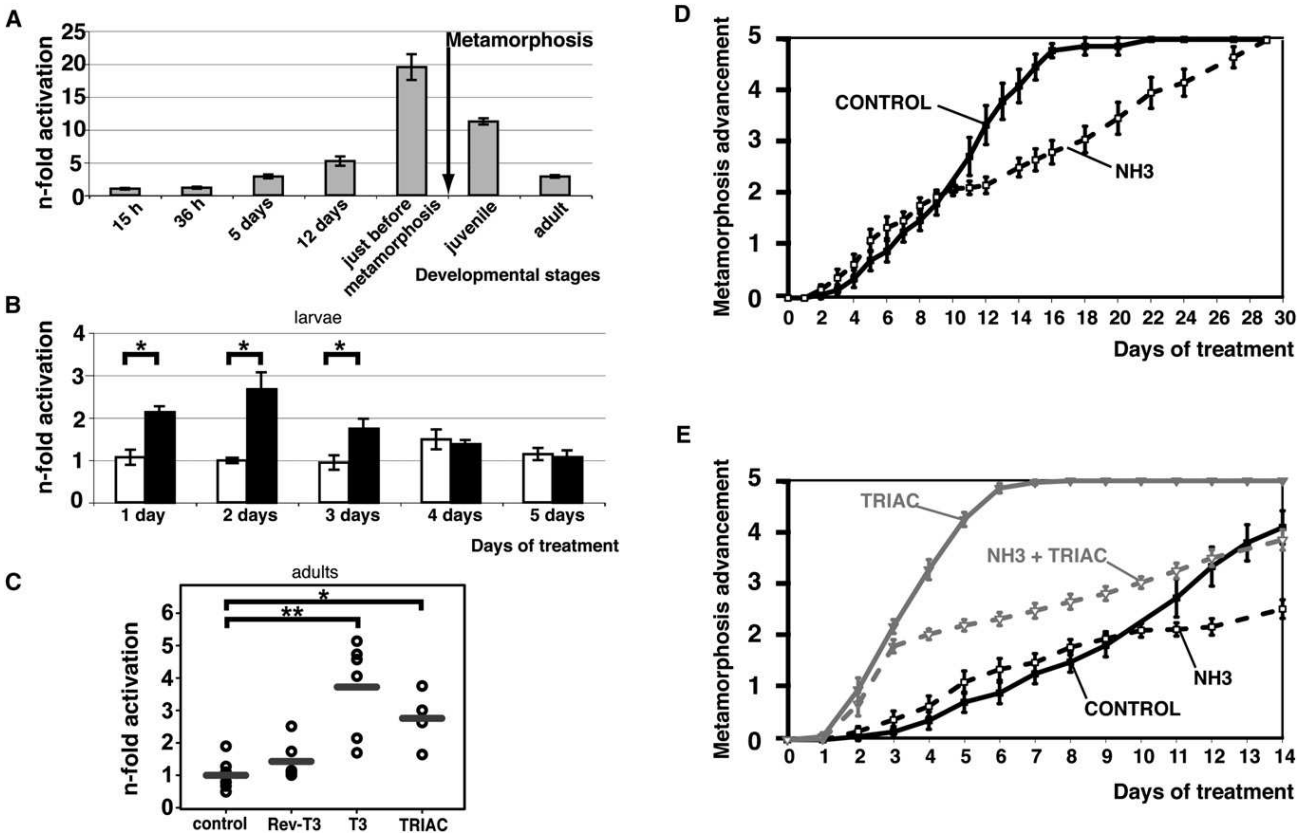


Figure 4. Involvement of amphiTR in TH-Induced Metamorphosis

(A–C) The expression of *amphiTR* was measured by quantitative RT-PCR during development (A), in premetamorphic larvae treated or not with T_3 at 5×10^{-8} M (B) or in adults treated with reverse T_3, T_3, or TRIAC at 10^{-7} M (C). The experiment on larvae started 18 days after fertilization. Bonferrini-corrected Wilcoxon test: * $p < 0.05$; ** $p < 0.01$. Error bars represent the standard error of the mean (SEM).

(D and E) The effect of NH3 at 10^{-6} M on spontaneous (D) and TRIAC-induced metamorphosis (E) of 12-day-old premetamorphic amphioxus larvae was monitored. TRIAC treatments were carried out at 10^{-8} M. Error bars correspond to a 5% confidence interval.

In this study we provide insights into the history of the TH signaling cascade during chordate evolution. Our data suggest that TH-triggered metamorphosis is an ancestral feature of all chordates. We propose that the upregulation of TR in this process, an event that has been found in all vertebrates studied to date [17, 18], is conserved in all chordates and is instrumental for the activation of the regulatory-gene network leading to metamorphosis. Given that THs control aspects of postembryonic development in nonmetamorphosing vertebrates, such as mammals [17], neotenic amphibians [19], and fishes [20], we propose that all chordates share a developmental transition controlled by THs. From this broader point of view, chordate metamorphosis can be considered as a postembryonic process controlled by THs and TR, two components forming the most conserved part of the regulatory network controlling metamorphosis (Figure S10A). The downstream components of such gene hierarchies are more variable and hence account for the diversity of metamorphic processes observed in chordates [19], which following this logic should also include processes like mammalian weaning [17]. Consequently, the distinction between direct and indirect development in chordates should be seen more as a difference in the intensity of the response to TH. Accordingly, even if the couple TH-TR controls postembryonic development in both mouse and anourans, TH-regulated gene expression patterns are highly variable between the two species [21, 22].

Whether TH-regulated metamorphosis occurs in echinoderms, and thus can be considered as a synapomorphy of deuterostomes, is an interesting possibility that remains to be investigated (Figure S10B). Indeed, in echinoderm groups, which undergo a metamorphosis involving profound changes in body symmetry, THs have been reported to induce this event [23], and several genes of the TH signaling pathway have been found in the sea urchin genome (M.P., M.S., F.B. and V.L., unpublished data). The most interesting insight derived from the present study is that TH-induced metamorphosis is an ancient chordate feature that is most usefully defined not morphologically, but genetically—by the presence of a core genetic mechanism based on the two partners TH and TR.

Interestingly, the precise nature of the active TH varies from one lineage to another: Our data suggest that it is not T_3 in amphioxus, and even if THs seem to play a role in the metamorphosis of the urochordate *Ciona intesinalis*, its unique TR does not bind T_3 [2]. This implies that T_3 might not be the ancestral TH in chordates.

In addition, our results strengthen the old idea that basal chordates already had a dual-function thyroid homolog—namely, an endostyle producing both exocrine secretions to capture food and TH to induce metamorphosis [6, 24]. It is likely that, during the evolution of the definitive vertebrate thyroid gland from an ancestral endostyle, the exocrine role

dwindled and disappeared while the endocrine production of THs became the major and finally the only function.

In cephalochordates, TH-dependent regulation of the transition from a highly asymmetric larva to a relatively symmetric juvenile in the genus *Branchiostoma* raises the question of how THs affect metamorphosis in the related genus *Epigonichthys*, which remains conspicuously asymmetric as an adult. Whether this is an instance of neoteny or a different type of metamorphosis remains to be determined. In addition, several enigmatic fossils, including the widely discussed stylophorans [5, 25], also show asymmetric features. Our conclusion that TH-induced metamorphosis is an ancient feature invites the re-evaluation of these fossils, considering that some could be premetamorphic or neotenic.

Supplemental Data

Experimental Procedures, ten figures, and four tables are available at http://www.current-biology.com/cgi/content/full/18/11/825/DC1/.

Acknowledgments

We thank John Lawrence for generously providing laboratory space at the University of South Florida. We would also like to thank Gérard Benoît, François Bonneton, Bastien Boussau, Barbara Demeneix, Frédéric Flamant, and two anonymous reviewers for reading the manuscript and for fruitful discussion, Anne Chabadel for help with the confocal microscope, Georges Delous for technical help with TH measurements of amphioxus extracts, and Maria Danihelova for technical assistance. This work was supported by the Centre National de la Recherche Scientifique, the Ecole Normale Supérieure de Lyon, the Association pour la Recherche sur le Cancer (ARC), and the Ministere de l'Education Nationale, de la Recherche et de la Technologie (MENRT), by CRESCENDO, a European Union Integrated Project of FP6, and also by grants of the Network of Excellence (NoE) Project EC, FOOD-CT-2004-506319 (CASCADE) and VEGA No. 2/0022/08.

Received: October 29, 2007
Revised: April 17, 2008
Accepted: April 25, 2008
Published online: May 29, 2008

References

1. Flamant, F., Baxter, J.D., Forrest, D., Refetoff, S., Samuels, H., Scanlan, T.S., Vennstrom, B., and Samarut, J. (2006). International Union of Pharmacology. LIX. The pharmacology and classification of the nuclear receptor superfamily: Thyroid hormone receptors. Pharmacol. Rev. *58*, 705–711.
2. Carosa, E., Fanelli, A., Ulisse, S., Di Lauro, R., Rall, J.E., and Jannini, E.A. (1998). Ciona intestinalis nuclear receptor 1: A member of steroid/thyroid hormone receptor family. Proc. Natl. Acad. Sci. USA *95*, 11152–11157.
3. Delsuc, F., Brinkmann, H., Chourrout, D., and Philippe, H. (2006). Tunicates and not cephalochordates are the closest living relatives of vertebrates. Nature *439*, 965–968.
4. Schubert, M., Escriva, H., Xavier-Neto, J., and Laudet, V. (2006). Amphioxus and tunicates as evolutionary model systems. Trends Ecol. Evol. *21*, 269–277.
5. Gee, H. (1996). Before the Backbone: Views on the Origin of the Vertebrates (London: Chapman & Hall).
6. Ogasawara, M. (2000). Overlapping expression of amphioxus homologs of the thyroid transcription factor-1 gene and thyroid peroxidase gene in the endostyle: Insight into evolution of the thyroid gland. Dev. Genes Evol. *210*, 231–242.
7. Tong, W., Kerkof, P., and Chaikoff, I.L. (1962). Identification of labeled thyroxine and triiodothyronine in amphioxus treated with 131-I. Biochim. Biophys. Acta *56*, 326–331.
8. Covelli, I., Salvatore, G., Sena, L., and Roche, J. (1960). Sur la formation d'hormones thyroidiennes et de leurs précurseurs par Branchiostoma lanceolatum. C. R. Soc. Biol. Paris *154*, 1165–1169.
9. Monaco, F., Dominici, R., Andreoli, M., De Pirro, R., and Roche, J. (1981). Thyroid hormone formation in thyroglobulin synthesized in the amphioxus (Branchiostoma lanceolatum Pallas). Comp. Biochem. Physiol. B *70*, 341–343.
10. Novinec, M., Kordis, D., Turk, V., and Lenarcic, B. (2006). Diversity and evolution of the thyroglobulin type-1 domain superfamily. Mol. Biol. Evol. *23*, 744–755.
11. Wagner, R.L., Apriletti, J.W., McGrath, M.E., West, B.L., Baxter, J.D., and Fletterick, R.J. (1995). A structural role for hormone in the thyroid hormone receptor. Nature *378*, 690–697.
12. Escriva, H., Manzon, L., Youson, J., and Laudet, V. (2002). Analysis of lamprey and hagfish genes reveals a complex history of gene duplications during early vertebrate evolution. Mol. Biol. Evol. *19*, 1440–1450.
13. Wagner, R.L., Huber, B.R., Shiau, A.K., Kelly, A., Cunha Lima, S.T., Scanlan, T.S., Apriletti, J.W., Baxter, J.D., West, B.L., and Fletterick, R.J. (2001). Hormone selectivity in thyroid hormone receptors. Mol. Endocrinol. *15*, 398–410.
14. Tata, J.R. (2006). Amphibian metamorphosis as a model for the developmental actions of thyroid hormone. Mol. Cell. Endocrinol. *246*, 10–20.
15. Nguyen, N.H., Apriletti, J.W., Cunha Lima, S.T., Webb, P., Baxter, J.D., and Scanlan, T.S. (2002). Rational design and synthesis of a novel thyroid hormone antagonist that blocks coactivator recruitment. J. Med. Chem. *45*, 3310–3320.
16. Lim, W., Nguyen, N.H., Yang, H.Y., Scanlan, T.S., and Furlow, J.D. (2002). A thyroid hormone antagonist that inhibits thyroid hormone action in vivo. J. Biol. Chem. *277*, 35664–35670.
17. Fraichard, A., Chassande, O., Plateroti, M., Roux, J.P., Trouillas, J., Dehay, C., Legrand, C., Gauthier, K., Kedinger, M., Malaval, L., et al. (1997). The T3R alpha gene encoding a thyroid hormone receptor is essential for post-natal development and thyroid hormone production. EMBO J. *16*, 4412–4420.
18. Marchand, O., Duffraisse, M., Triqueneaux, G., Safi, R., and Laudet, V. (2004). Molecular cloning and developmental expression patterns of thyroid hormone receptors and T3 target genes in the turbot (Scophtalmus maximus) during post-embryonic development. Gen. Comp. Endocrinol. *135*, 345–357.
19. Safi, R., Bertrand, S., Marchand, O., Duffraisse, M., de Luze, A., Vanacker, J.M., Maraninchi, M., Margotat, A., Demeneix, B., and Laudet, V. (2004). The axolotl (Ambystoma mexicanum), a neotenic amphibian, expresses functional thyroid hormone receptors. Endocrinology *145*, 760–772.
20. Brown, D.D. (1997). The role of thyroid hormone in zebrafish and axolotl development. Proc. Natl. Acad. Sci. USA *94*, 13011–13016.
21. Yen, P.M., Feng, X., Flamant, F., Chen, Y., Walker, R.L., Weiss, R.E., Chassande, O., Samarut, J., Refetoff, S., and Meltzer, P.S. (2003). Effects of ligand and thyroid hormone receptor isoforms on hepatic gene expression profiles of thyroid hormone receptor knockout mice. EMBO Rep. *4*, 581–587.
22. Das, B., Cai, L., Carter, M.G., Piao, Y.L., Sharov, A.A., Ko, M.S., and Brown, D.D. (2006). Gene expression changes at metamorphosis induced by thyroid hormone in Xenopus laevis tadpoles. Dev. Biol. *291*, 342–355.
23. Heyland, A., Reitzel, A.M., and Hodin, J. (2004). Thyroid hormones determine developmental mode in sand dollars (Echinodermata: Echinoidea). Evol. Dev. *6*, 382–392.
24. Sembrat, K. (1953). Effect of the endostyle of the lancelet (Branchiostoma lanceolatum Pall.) on the metamorphosis of axolotl. Zool. Pol. *6*, 3–19.
25. Ruta, M. (1999). Brief review of the stylophoran debate. Evol. Dev. *1*, 123–135.

Supplemental Data

Amphioxus Postembryonic Development Reveals the Homology of Chordate Metamorphosis

Mathilde Paris, Hector Escriva, Michael Schubert, Frédéric Brunet, Julius Brtko, Fabrice Ciesielski, Dominique Roecklin, Valérie Vivat-Hannah, Emilien L. Jamin, Jean-Pierre Cravedi, Thomas S. Scanlan, Jean-Paul Renaud, Nicholas D. Holland, and Vincent Laudet

Supplemental Experimental Procedures

Embryonic and Larval Culture and Treatment with THs

Embryonic cultures were as previously described [S1]. Treatments with T_3, reverse T_3, T_4, TRIAC (all from Sigma, St. Louis, Missouri), NH3, IOP (TokyoKaiser, Japan) were carried out 12 days after fertilization, when the larvae had six or seven gill slits. Seawater complemented with the appropriate treatment was renewed in the cultures daily.

Monitoring of Amphioxus Metamorphosis on Morphological Criteria

A premetamorphic amphioxus is asymmetric with a single row of gill slits on the right side and the mouth on the left. During metamorphosis, the body becomes more and more symmetrical as a series of rapid morphological changes occurs in the following order: (1) metapleural folds arise along either side of the body; (2) a second row of gill slits appears dorsal to the first row on the right side; (3) tongue bars indent each gill slit, and the first gill-slit row migrates to the left side; (4) the mouth migrates ventrally and cirri appear; and (5) the liver diverticulum becomes conspicuous. The progression of metamorphosis is quantified by counting daily how many of these five criteria are present in each animal. The score then refers to the mean of the scores determined for individual animals in each batch (Figure 1C and Figure S1).

Constructs

We cloned the full-length amphioxus TR and lamprey TR1 and TR2 by polymerase chain reaction (PCR) into a pSG5 vector between EcoR1 sites [S2]. The rat TRα, cloned into a pSG5 vector, was a gift from Barbara Demeneix, and *Ciona intestinalis* TR, cloned in a pSG5 vector as well, was a gift from Emmanuelle Janini [S3]. Chimeras comprising the GAL4 DNA-binding domain fused with the LBD of rat TRα (residues 120 to 410), the LBD of amphiTR (residues 121 to 431), the LBD of the two TRs from lampreys (lampTR1 and lampTR2, residues 85 to 371 and 108 to 395, respectively) and the LBD of *Ciona* TR (residues 137 to 672) were cloned in the pG4MpolyII vector. Chimeras comprising either SRC1 or the interaction domain of NCOR fused to the sequence encoding the activation domain of the VP16 protein from herpes simplex virus were a gift from Gérard Benoit. Chimeras comprising the full-length amphiTR (amphiTR-GFP) were cloned into the pEGFP-C1 vector. The pEGFP-C1 vector was a gift from Irma Machuca.

Cell Culture and Transfections

Human embryonic kidney 293 cell culture and transfections using Lipofectamine Plus reagent (Invitrogen) were done according to the manufacturer's recommendations and as previously described [S4]. Results show the mean ± SEM (n = 3) of representative experiments. All experiments were performed at least four times except the test of activation of the ciona TR by T_3 and the test of release of NcoR by amphiTR under TRIAC stimulation, which were both tested twice. Statistical tests were performed with a Bonferroni-corrected Student's t test: * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$.

Protein Preparation and Nondenaturating Mass Spectrometry

amphiTR LBD (E147–E431) cDNA was cloned in the prokaryotic expression vector pET-15b (Novagen) in frame with the vector N-terminal-located $(His)_6$ tag. Protein was expressed in *E. coli* BL21 strain and subsequently purified by affinity chromatography followed by size-exclusion chromatography. Noncovalent mass spectrometry used electrospray ionization (ESI)-based mass spectrometer (LCT, Micromass, Manchester, UK). Before ESI-MS analysis, buffer exchange was required to replace amphiTR purification buffer by a 50 mM ammonium buffer (NH_4OAc [pH 5.5]). Buffer exchange used NAP5 desalting columns (Sephadex G-25, GE Healthcare) and was performed according to manufacturer's recommendations. Ten micromolars of amphiTR in NH_4OAc was incubated with or without 2 M equivalents of thyroid hormones and derivatives (T_3, MW = 651.0 g/mol; T_4, MW = 776.9 g/mol; TRIAC, MW = 621.9 g/mol) before being infused into the mass spectrometer. Pressure in the interface was adjusted to 6.5 mbar, and the sample cone voltage (Vc, corresponds to the acceleration of the ions in the interface) varied from 120V (disruptive for noncovalent complexes) to 5V (most permissive to observe noncovalent interaction).

Limited Proteolytic Digestion

These assays were done as previously described [S5].

Isolation of Amphioxus Nuclei and Cytosol and Extraction of Nuclear Nonhistone-Protein-Containing Fractions

Whole amphioxus were minced and homogenized in ice-cold SMCT according to Ichikawa et al. [S6]. The cytosol was separated from the mixture of disrupted cells by ultracentrifugation at 104,000 × g. The crude nuclear pellet was then mixed with 2.6 mol/L sucrose and treated by isopycnic ultracentrifugation at 220,000 × g for 75 min. The nuclear nonhistone proteins were obtained directly by extracting nuclei in a high-ionic-strength KMTD buffer containing 0.3 mol/L KCl, 1 mmol/L $MgCl_2$, 10 mmol/L Tris-HCl (pH 8.0), and 1 mmol/L DTT at 0°C for 1 hr and separated from the mixture of disrupted nuclei by ultracentrifugation at 135,000 × g.

Binding of 125I-3,3′,5-L-Triiodothyronine (T3) to Amphioxus Nuclear and cytosolic Fractions

The assay was performed at 22°C in 0.5 ml KMTD buffer (pH 8.0) for nuclear extracts and in 0.5 ml low-ionic-strength SMCT buffer (0.32 mol/L sucrose, 10 mmol/L Tris HCl [pH 7.4], 1 mmol/L $MgCl_2$, 1 mmol/L $CaCl_2$, 1 mmol/L DTT, 0.1 mmol/L PMSF) for cytosolic extracts. The protein samples were incubated with 3×10^{-10} mol/L of 125I-T3 for 2 hr. Nonspecific binding of the labeled ligand was determined by simultaneous incubation with 3×10^{-7} mol/L of nonlabeled T3 per sample. After incubation, 0.5 ml of Dowex1-X8 (80 mg/mL) suspension in SMCT buffer (pH 7.4) for cytosol extract or of KMTD buffer (pH 8.0) for nuclear extract at 0°C to 4°C was added to each sample. The supernatant was collected after short centrifugation at 1000 × g. Then, 0.5 ml of the supernatant was decanted, and its radioactivity was quantified in a Beckman model 4000γ-spectrometer.

Detection of THs in Amphioxus Adult Extracts

Sample preparation for HPLC-MS/MS analyses was achieved from 5 g of amphioxus ground in 50 ml of HCl 0.5 N followed by heating at 100°C for 2 hr. The sample mixture was centrifuged at 5000 rpm during 20 min, and 5 ml of an aqueous solution of NaCl 6%, 10 ml methanol, and 10 ml cyclohexane were added to the supernatant. The organic phase was eliminated by centrifugation, and the aqueous phase was filtered. Three extractions by ethyl acetate were performed on the aqueous phase, and organic phase was evaporated to dryness. Obtained extract was finally dissolved in 500 μL methanol (+NH_4OH 1%) to be analyzed by HPLC-MS/MS.

Chromatographic separation was performed with a Thermo Electron Surveyor liquid-chromatography system equipped with a quaternary pump (Les Ulis, France). THs were separated with a Polaris C18-A column (3 μm, 150.0 × 2.0 mm i.d.), and separation was achieved with a binary mobile phase at a flow rate of 200 μL/min comprising solvent A (CH_3CN/H_2O [v/v, 1:9]) and solvent B (CH_3CN/H_2O [v/v, 9:1]). The gradient elution program was 0% to 100% of B from 0 to 15 min followed by B at 100% until 20 min, and the sample volume injected was 20 μL.

This HPLC system was coupled to a TSQ Quantum triple-quadrupole mass spectrometer (Thermo Electron, Les Ulis, France) equipped with an Electrospray ionization source (ESI). MS instrument was operated in the negative mode with optimized source conditions (spray voltage 3600V, heated capillary temperature 370°C, sheath gas [N2] pressure 13 au, auxiliary gas [N2] pressure 7 au). Optimal MS/MS parameters were obtained in the single ion monitoring (SIM) mode, which was the most sensitive for the THs

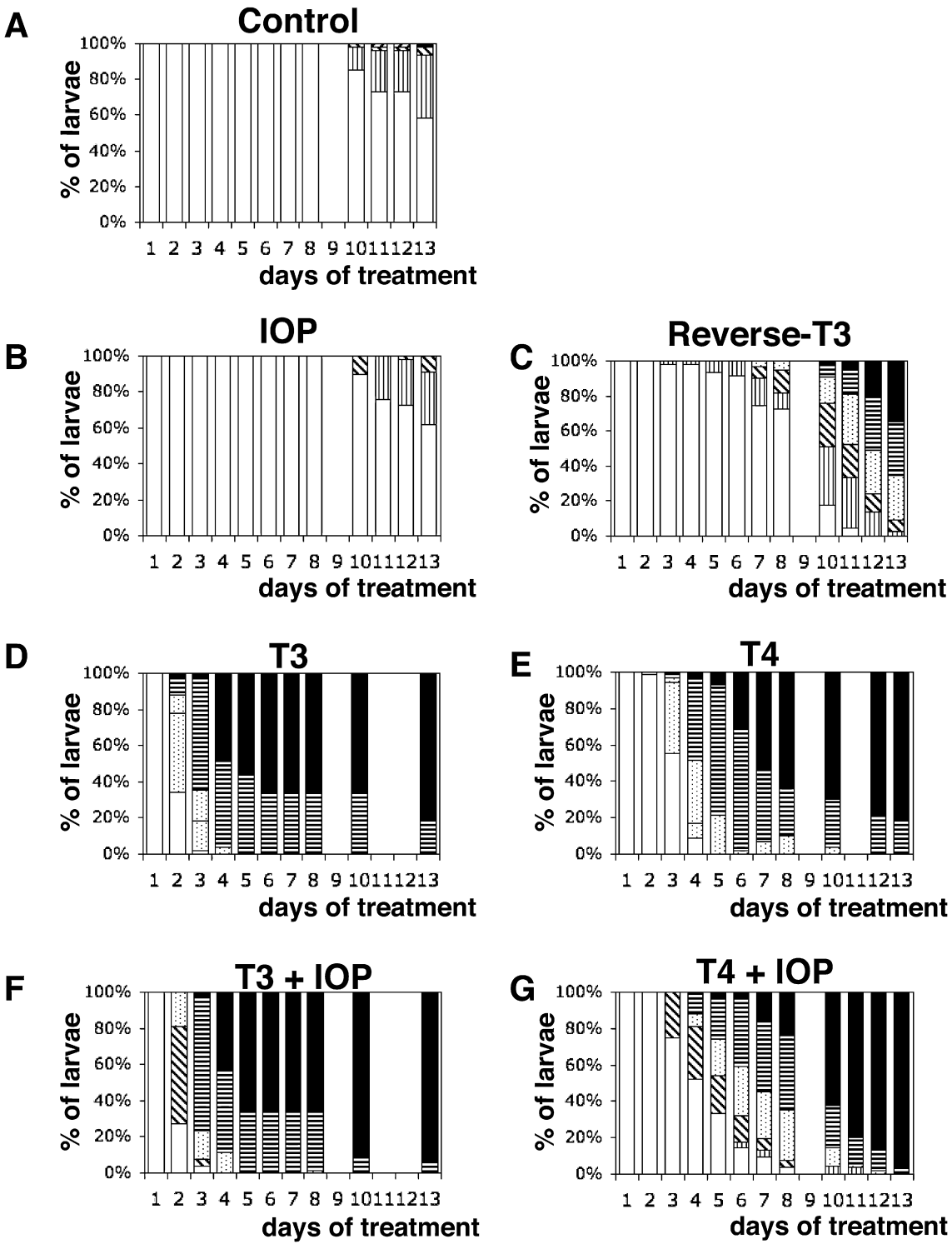


Figure S1. Advancement of Metamorphosis in TH-Treated Premetamorphic Larvae

Distribution of the batches of TH-treated larvae depending on the five metamorphosis stages described in Table S1. The different conditions are as follows: control batch (A), treatment with IOP at 10^{-5} M (B), reverse T_3 at 10^{-8} M (C), T_3 at 10^{-8} M (D), T_4 at 10^{-8} M (E), IOP at 10^{-5} M plus T_3 at 10^{-8} M (F), and IOP at 10^{-5} M plus T_4 at 10^{-8} M (G). Bar shadings are as follows: white bars (stage 0), vertical lines (stage 1), oblique lines (stage 2), dots (stage 3), horizontal lines (stage 4), and black bars (stage 5).

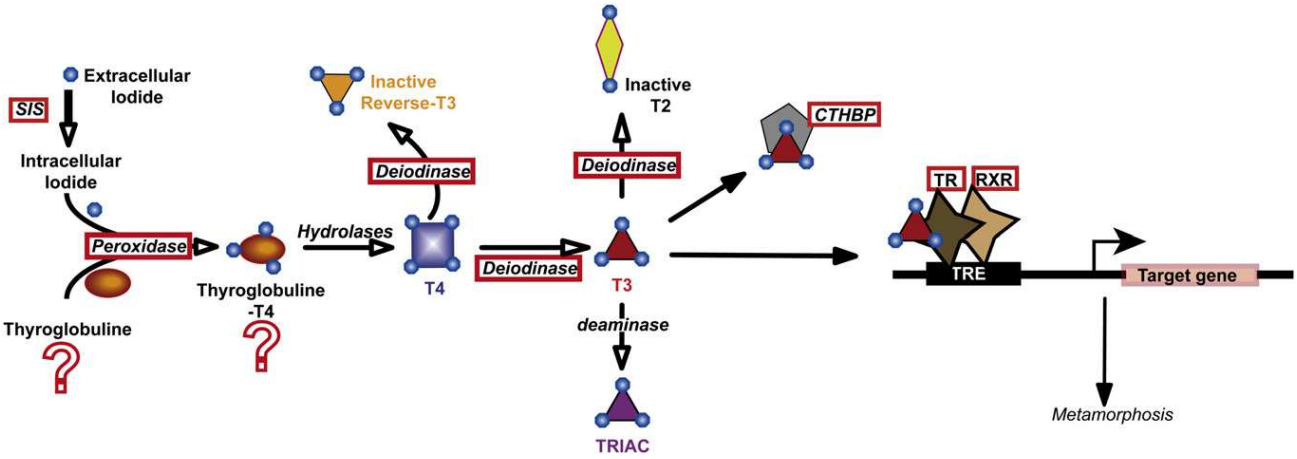


Figure S2. Thyroid-Hormone Signaling Pathway in Amphioxus

Summary of phylogenetic analyses of the different amphioxus genes homologous to vertebrate members of the TH metabolism pathway. Genes surrounded by a red line have been found in the amphioxus genome. Hydrolases catalyzing the production of T_4 from iodinated thyroglobulin and the deaminases catalyzing the production of TRIAC from T_3 are not specific for this pathway and have not been searched for in the genome.

quantification, and in the single reaction monitoring (SRM) mode for the confirmation of the THs identification ([M-H]→[M-$C_2H_3NO_2$-I]−• for T_4 at R_T = 11.4 min, [M-H]→[M-I]−• for TETRAC at R_T = 15.6 min, [M-H]→[M-NH3]− for T_3 at R_T = 10.4 min, and [M-H]→I−• for TRIAC at R_T = 14.0 min). External calibration curves have been constructed by injections of standard solutions diluted at five concentrations ranging from 1 to 1000 pg/μL, and coefficients of determination (r2) were higher than 0.990. Measured limits of quantification were 1 pg/μL for T_3 and 5 pg/μL for TRIAC, TETRAC, and T_4, and no matrix effect was observed in our analysis conditions.

Quantitative RT-PCR

Total RNAs were prepared from whole animals with Trizol (Invitrogen) according to the manufacturer's instructions and reverse-transcribed with random primers and MMLV Reverse Transcriptase (Invitrogen). The cDNA was then used as template for a quantitative real-time PCR assay with the QuantiTect SYBR Green PCR reagents (QIAGEN) and the DNA Engine Opticon system (MJ Research). Expression of amphiTR was normalized to the corresponding 28S expression levels. The experiment on adults was performed with four to eight animals, and experiments on larvae were performed on single samples of pooled animals. Results show the mean ± SEM (n = 4). Statistical tests for the experiments on adults were performed with a Bonferroni-corrected Wilcoxon test because of heteroscedasticity and otherwise with a Student's t test: * $p < 0.05$; ** $p < 0.01$. Efficiency and specificity of the amplification of the different samples was controlled by generating standard curves and by carrying out melting curves and agarose gels of the amplicons. The relative level of each of the RNAs was calculated by 2^{-CT} (CT being the cycle number, in which SYBR green fluorescence exceeds the constant threshold value). All experiments were done at least three times in triplicates. The sequences of forward and reverse primers were as follows: 28S-F (5′-ACTCTGGATAACCCAGCCGAT-3′), 28S-R (5′-TGCCTTCCTTGGATGTGGTAG-3′), amphiTR-F (5′-CCGATGCATGACATGTGAAGG-3′), and amphiTR-R (5′-CACTTCTTGAAGCGACACTCC-3′).

Microscopy

Cells were seeded on sterilized coverslips in 12-well plates. Forty-eight hours after transfection, cells were fixed in 3.7% paraformaldehyde, pH 7.2, for 10 min. Cells were analyzed with a confocal Zeiss LSM510 microscope with a 63× (NA1.4) Plan Neofluor objective.

Retrieval of the Genes Involved in TH Metabolism from the Amphioxus Genome and Subsequent Phylogenetic Analysis

The human genes of interest were used to retrieve Ensembl families along with the Swissprot and Trembl sequences. Each member of the family was used to blast the amphioxus proteins sequences via the Lassap software from GeneIT. For each family, amphioxus best-hit sequences were retrieved and then aligned via ClustalW [S7] followed by manual adjustment, if necessary. To ensure that no amphioxus genes were missed, we performed a BLAST search of some genes of interest directly on the amphioxus web site at JGI (http://genome.jgi-psf.org/Brafl1/Brafl1.home.html). The sea urchin genome sequence was searched and sequences added to the existing alignments. Phylogenetic analyses were performed with neighbor joining via the Poisson method (500 bootstrap replicates) (Phylowin [S7]) and maximum likelihood using PHYML v2.4.4 [S8] (JTT model, Gamma distribution parameter = 2, starting tree bionj, 100 bootstrap replicates, optimized tree topology, branch lengths, and rate parameters).

Phylogenetic Analysis of Deuterostome TRs

Amino acid sequences of deuterostome TRs were aligned with the MUSCLE 3.6 program [S9] and manually corrected in SEAVIEW. Phylogenetic trees were inferred by maximum likelihood with PhyML v2.4.3 under a WAG substitution matrix with a four-category gamma rate correction (α estimated) and with the proportion of invariant sites estimated. The DBD and LBD protein regions were used. Robustness was assessed by bootstrap analyses (1000 replicates) [S10].

Species abbreviations and their groups are indicated as follows. Birds: chicken, *Gallus gallus*. Mammals: human, *Homo sapiens*. Amphibians: African clawed frog, *Xenopus laevis*. Teleost fish: zebrafish, *Danio rerio* and salmon, *Salmo salar*. Chondrichthyans: catshark, *Scyliorhinus canicula*. Cyclostomes: sea lamprey, *Petromyzon marinus*. Urochordates: *Ciona intestinalis* and *Ciona savignyi*. Cephalochordates: *Branchiostoma floridae* and *Branchiostoma lanceolatum*. Echinoderms: sea urchin, *Strongylocentrotus purpuratus*.

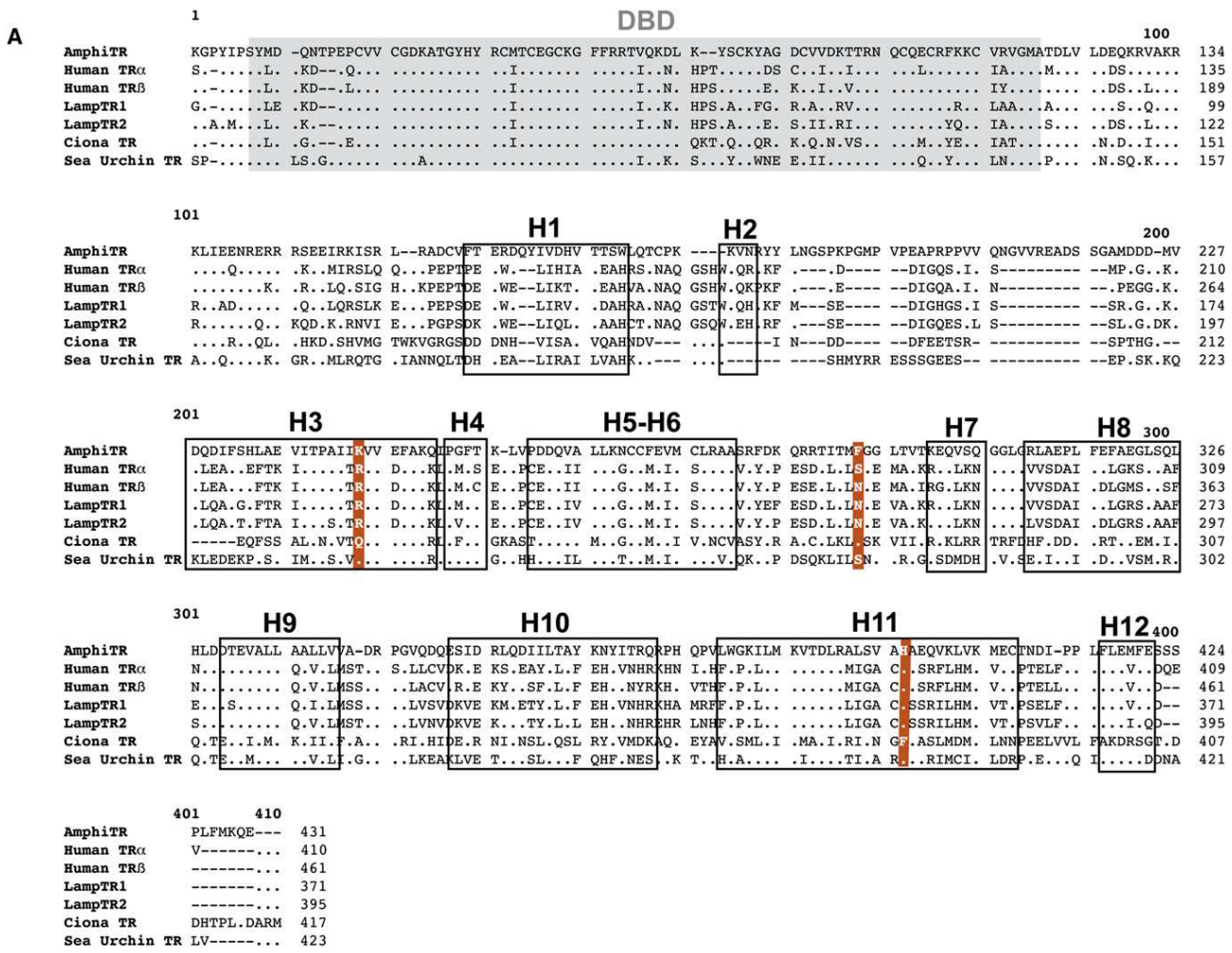
Accession numbers are as follows: *Branchiostoma floridae* RAR, AAM46149; human RARα, CAA29787; human RARβ, CAA68398; human RARγ, P13631; sea urchin TR, XP_001185977; *Ciona intestinalis* TR, NP_001027658; *Ciona savignyi* TR, BAB68358; *Branchiostoma floridae* TR, EF672345; *Branchiostoma lanceolatum* TR, EF672345; lamprey TR1, ABC49722; lamprey TR2, ABC49723; catshark TRα, EF672346; zebrafish TRα, AAA99811; salmon TRα, AAD38689; chicken TRα, CAA68792, human TRα, AAA66021; *Xenopus laevis* TRα, AAA49970; chicken TRβ, CAA35544; human TRβ, CAA28412; *Xenopus laevis* TRβ, AAA49657; zebrafish TRβ, AAF14239; and salmon TRβ, AAL06731.

Electrophoretic Mobility Shift Assay

EMSAs were performed as previously described [S11]. Where indicated, a 10- and 100-fold molar excess of 30 bp unlabeled oligonucleotides (a consensus DR4, the putative DR4, called "aDR4" found upstream of the amphiTR coding sequence, the palindromic consensus sequence HREpal, and a nonrelated probe) were added as competitors.

Supplemental References

S1. Holland, L.Z., and Yu, J.K. (2004). Cephalochordate (amphioxus) embryos: Procurement, culture, and basic methods. Methods Cell Biol. *74*, 195–215.



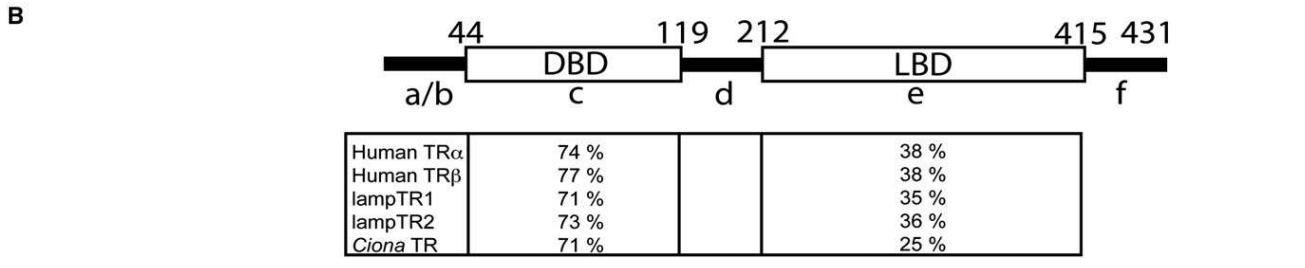


Figure S3. Sequence Analysis of amphiTR

(A) Sequence alignment of several TRs including amphiTR. The DBD is highlighted in light gray. The 12 helices of the LBD are indicated on the basis of the known 3D structure of rat TRα [S12]. The three amino acids of rat TRα making direct hydrogen bonds with T3 are highlighted in red. The more divergent A/B and F domains are omitted from the alignment. However, the numbering of the sites along the alignment starts at the beginning of each protein.
(B) Percent identity of the amphioxus TR with some chordate TRs in the DNA- and ligand-binding domains.

S2. Escriva, H., Manzon, L., Youson, J., and Laudet, V. (2002). Analysis of lamprey and hagfish genes reveals a complex history of gene duplications during early vertebrate evolution. Mol. Biol. Evol. *19*, 1440–1450.

S3. Carosa, E., Fanelli, A., Ulisse, S., Di Lauro, R., Rall, J.E., and Jannini, E.A. (1998). Ciona intestinalis nuclear receptor 1: A member of steroid/thyroid hormone receptor family. Proc. Natl. Acad. Sci. USA *95*, 11152–11157.

S4. Wang, Z., Benoit, G., Liu, J., Prasad, S., Aarnisalo, P., Liu, X., Xu, H., Walker, N.P., and Perlmann, T. (2003). Structure and function of Nurr1 identifies a class of ligand-independent nuclear receptors. Nature *423*, 555–560.

S5. Escriva, H., Bertrand, S., Germain, P., Robinson-Rechavi, M., Umbhauer, M., Cartry, J., Duffraisse, M., Holland, L., Gronemeyer, H., and Laudet, V. (2006). Neofunctionalization in vertebrates: the example of retinoic acid receptors. PLoS Genet *2*, e102.

S6. Ichikawa, K., Brtko, J., DeGroot, L.J., Hashizume, K., and Yamada, T. (1989). Stabilization, accurate determination, and purification of rat liver nuclear thyroid hormone receptor. J. Endocrinol. *120*, 237–243.

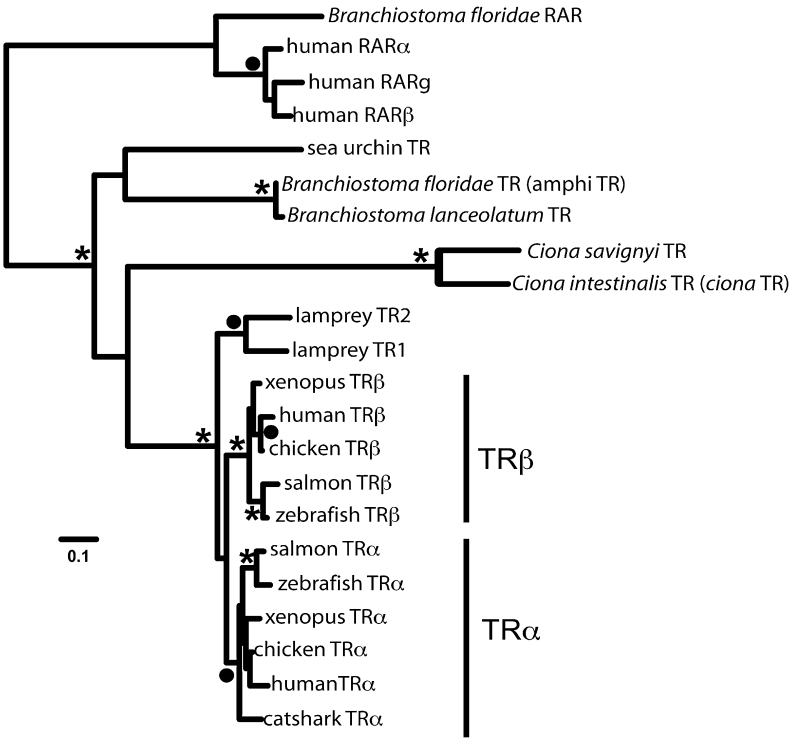


Figure S4. Phylogenetic Tree of TRs

A maximum likelihood (ML) tree was obtained on the basis of the amino acid sequences of the DBD and LBD of several deuterostome TRs via a WAG+γ+I model. Bootstrap percentages obtained after 1000 replicates are shown for selected branches. A star indicates that the bootstrap was maximal (100%), and a black dot indicates that the bootstrap was above 90%. A scale bar indicates the number of changes per site. The tree was rooted with RAR sequences.

S7. Galtier, N., Gouy, M., and Gautier, C. (1996). SEAVIEW and PHYLO_-WIN: Two graphic tools for sequence alignment and molecular phylogeny. Comput. Appl. Biosci. *12*, 543–548.

S8. Guindon, S., and Gascuel, O. (2003). A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. Syst. Biol. *52*, 696–704.

S9. Edgar, R.C. (2004). MUSCLE: Multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res. *32*, 1792–1797.

S10. Felsenstein, J. (1985). Confidence limits on phylogenies: An approach using the bootstrap. Evolution Int. J. Org. Evolution *39*, 783–791.

S11. Wada, H., Escriva, H., Zhang, S., and Laudet, V. (2006). Conserved RARE localization in amphioxus Hox clusters and implications for Hox code evolution in the vertebrate neural crest. Dev. Dyn. *235*, 1522–1531.

S12. Wagner, R.L., Apriletti, J.W., McGrath, M.E., West, B.L., Baxter, J.D., and Fletterick, R.J. (1995). A structural role for hormone in the thyroid hormone receptor. Nature *378*, 690–697.

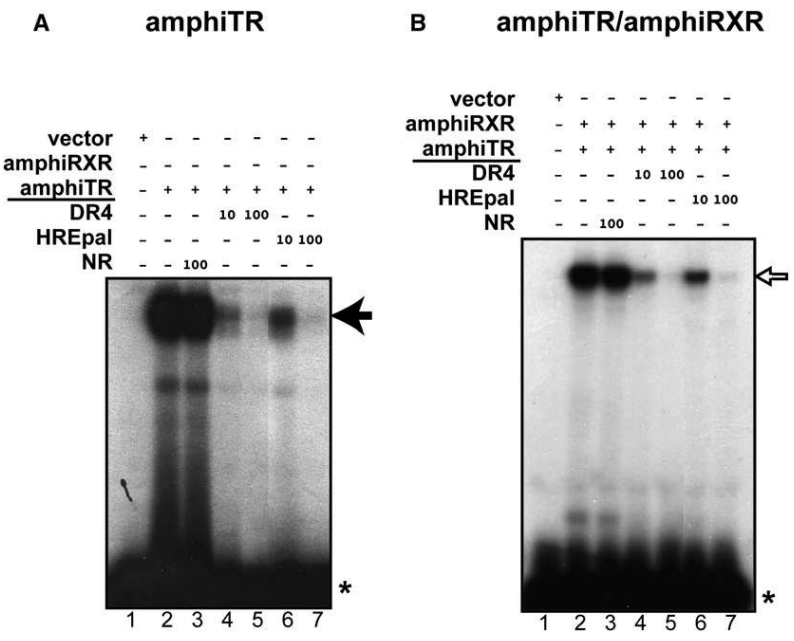


Figure S5. DNA Binding Characterization of amphiTR

The amphiTR and amphiRXR were synthesized in vitro, and either amphiTR alone (A) or amphiTR and amphiRXR (B) was allowed to bind to a ^{32}P-labeled consensus DR4 probe in an EMSA. Lane 1, unprogrammed (pSG5) reticulocytes. Lanes 2–7, programmed lysates. Lanes 3–7, unlabeled nonrelated oligonucleotide, DR4, or HREpal was added at indicated molar excess as competitors to test the specificity of the binding. The asterisk indicates free DR4 probe. The black arrow indicates the probe bound to homodimers of amphiTR, and the white arrow indicates the probe bound to heterodimers of amphiTR and amphiRXR.

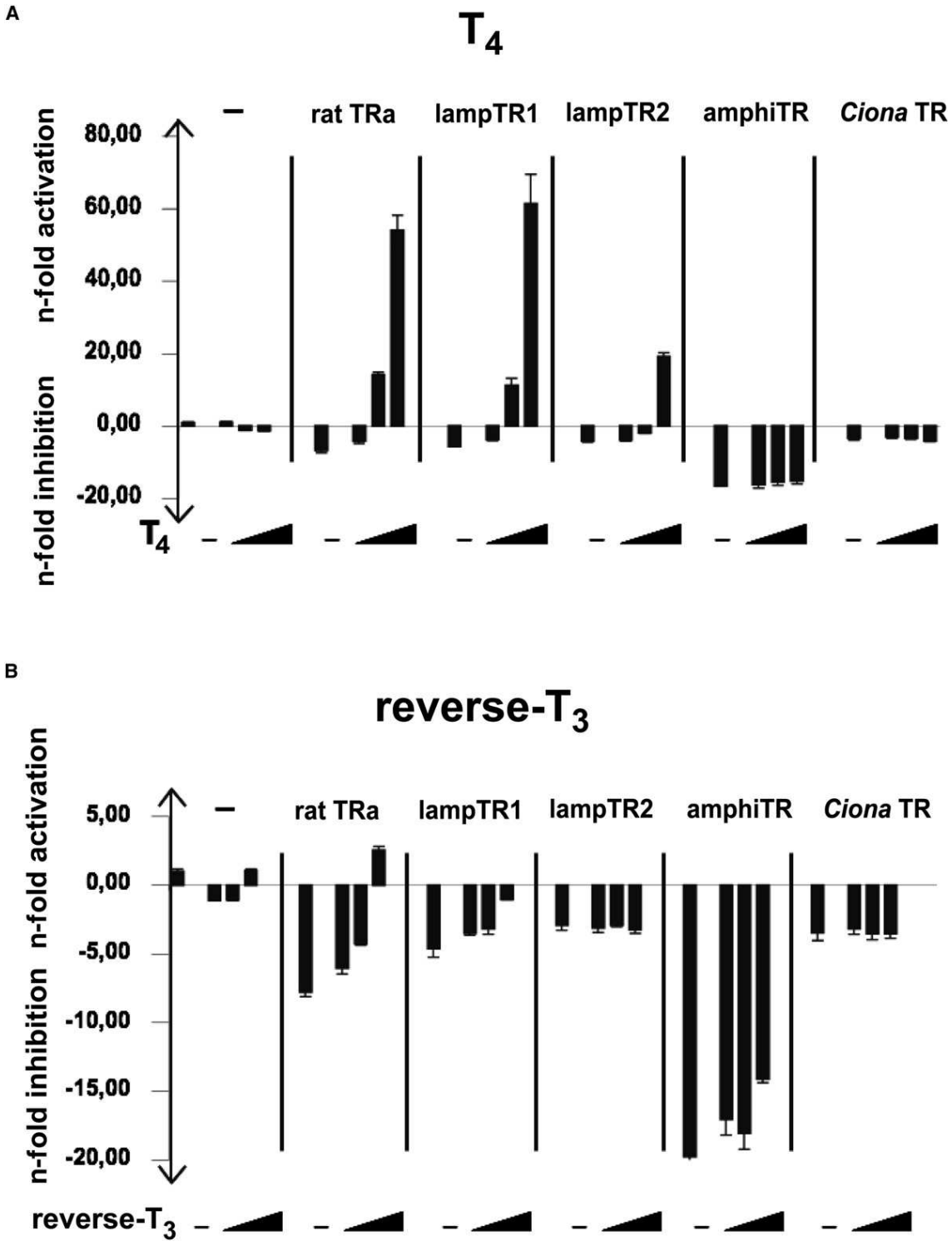


Figure S6. The amphiTR Is Activated neither by T_4 nor by Reverse T_3

Transcriptional activity of chordate TRs after T_4 (A) and reverse T_3 (B) treatment. The GAL4-LBD constructs containing the TRα from rat, the TR1 or the TR2 from lamprey, the TR from amphioxus, or the TR from *Ciona intestinalis* were tested in transfected 293 cells for their ability to activate the cotransfected cognate (17 m)5x-G-luc reporter plasmid. Cells were treated with increasing doses of hormones (10^{-8} M to 10^{-6} M).

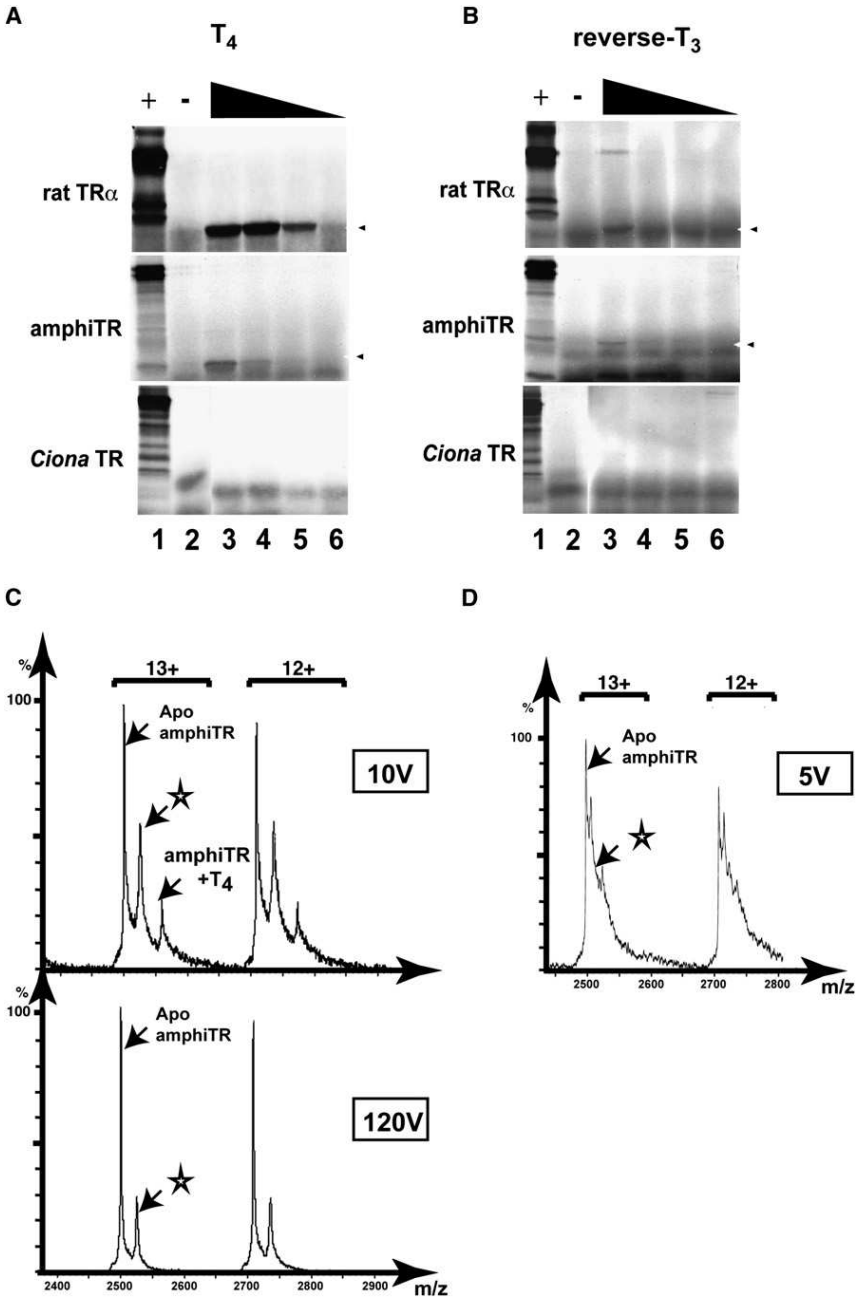


Figure S7. The amphiTR Binds neither T_4 nor Reverse T_3

(A and B) Limited proteolysis of chordate TRs with T_4 (A) and reverse T_3 (B). The first lane represents the undigested protein, and the following lanes represent the digested protein in the absence (lane 2) or presence of T_4 at decreasing concentrations (10^{-3} M to 10^{-6} M, lanes 3–6). Protected bands compared with the negative control (lane 2) are indicated with a black arrowhead.

(C) Nondenaturing ESI-MS analysis of T_4 binding to the amphiTR LBD. Ion acceleration (Vc) varied from 10V (preserves noncovalent complexes) to 120V (disrupts noncovalent interaction). At low Vc (10V), three species were detected: the apo form of the amphiTR-LBD, an unidentified adduct with a Dmass of 329 Da (star), and the T_4-bound amphiTR LBD. The ratio for all three species was 55%:30%:15%, respectively. Binding of T_4 was reversible: at high Vc (120V), the complex between amphiTR-LBD and T_4 was disrupted.

(D) Nondenaturing ESI-MS analysis of reverse T_3 binding to amphiTR LBD. No binding of reverse T_3 was found in the gas phase.

For each mass-spectrometry experiment, two protonation states are shown (13+ and 12+ charge states).

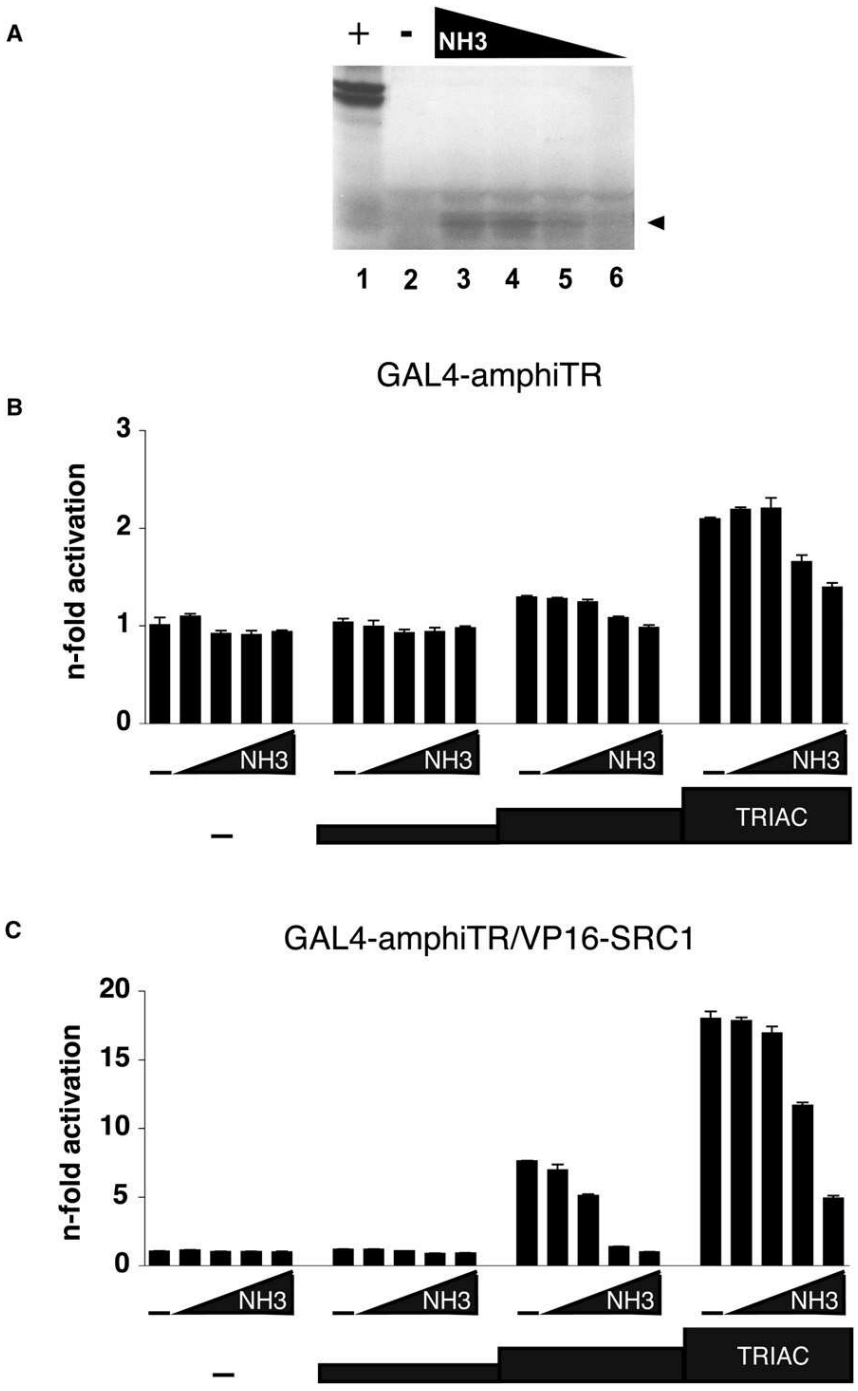


Figure S8. NH3 Acts as an amphiTR Antagonist

(A) Limited proteolysis of amphiTR with NH3. Lane 1, undigested protein; lanes 2–6, digested protein in the absence (lane 2) or presence (lanes 3–6) of TRIAC (4.10^{-4} M, 10^{-4} M, 10^{-5} M, 10^{-6} M). The protected band compared with the negative control (lane 2) is indicated with a black arrowhead.

(B) GAL4-amphiTR-LBD was tested for its ability to activate a (17 m)5x-G-luc reporter plasmid in the presence of increasing doses of NH3 (10^{-8} M, 10^{-7} M, 10^{-6} M, 4.10^{-6} M) in competition with increasing doses of TRIAC (10^{-8} M to 10^{-6} M).

(C) Representation of the competition of NH3 with TRIAC in a mammalian two-hybrid SRC1 recruitment assay. The GAL4-amphiTR-LBD chimera was transfected with the coactivator SRC1 fused to the strong activation domain VP16 in the presence of increasing doses of NH3 (10^{-8} M, 10^{-7} M, 10^{-6} M, 4.10^{-6} M) in competition with TRIAC (10^{-8} M to 10^{-6} M).

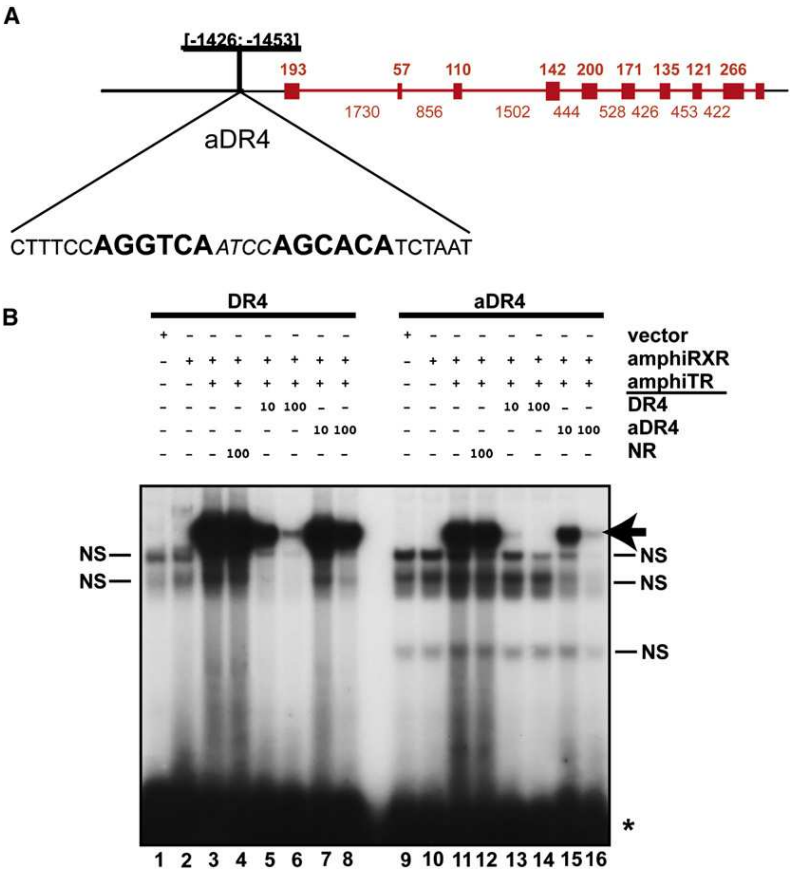


Figure S9. DNA Binding Characterization of amphiTR to the amphiTR Regulatory Region

(A) Schematic representation of the genomic structure of amphiTR. The exon-intron structure is conserved between amphioxus and vertebrates. There is a DR4-like response element (aDR4) 1426 bp upstream of the first ATG, with two mismatches in the second DR, compared with the consensus DR4 (AGGTCAacacAGATGA).

(B) The ability of amphiTR to bind the aDR4 has been tested by EMSA. The amphiTR and amphiRXR were synthesized in vitro and allowed to bind to a ^{32}P-labeled consensus DR4 probe or to the ^{32}P-labeled aDR4 probe in an EMSA (lanes 3–8 and lanes 11–16, respectively). Lanes 1 and 9 show unprogrammed (pSG5) reticulocytes; lanes 2 and 10 show amphiRXR alone. Unlabeled nonrelated oligonucleotide (lanes 4 and 12), DR4 (lanes 5 and 6 and lanes 13 and 14), and aDR4 (lanes 7 and 8 and lanes 15 and 16) were added as competitor to test the specificity of the binding. The star indicates free probe. The black arrow indicates the probe bound to heterodimers of amphiTR and amphiRXR. NS indicates nonspecific binding.

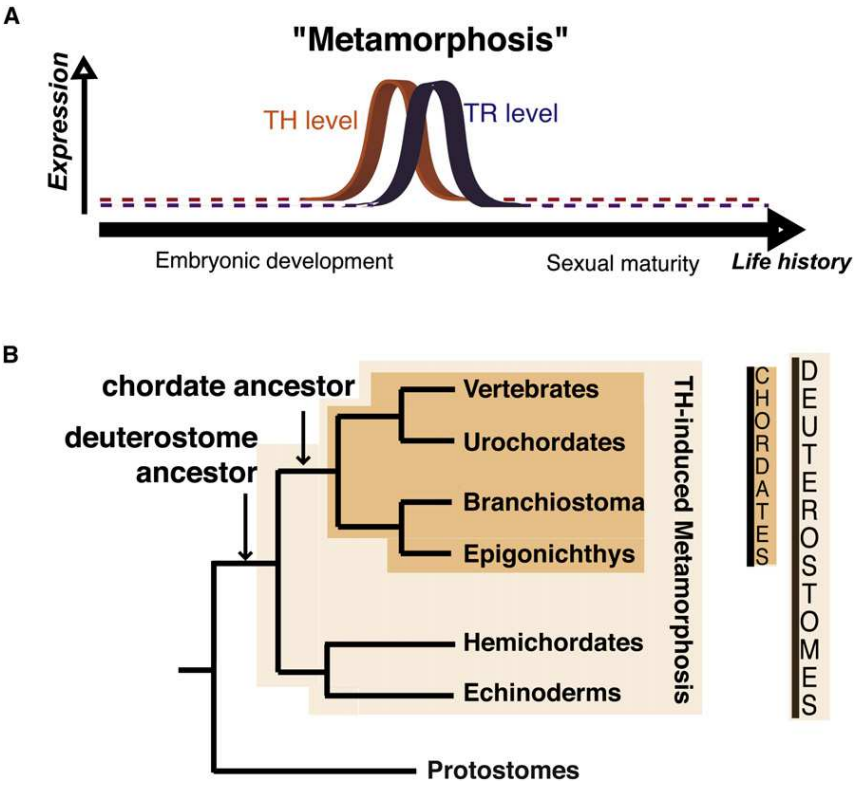


Figure S10. Evolution of the TH Signaling Pathway in Chordates

(A) In all chordates studied so far, a peak of TH production correlates with a peak of TR expression and is a key determinant for triggering metamorphosis. We propose that these features are conserved in all chordates and even constitute a chordate synapomorphy.

(B) Phylogenetic tree summarizing the evolution of TH-induced metamorphosis in deuterostomes.

8

ADDITIONAL EXPERIMENTS

8.1 TH METABOLISM IN AMPHIOXUS

In chapter **??**, we proposed that a T_3 derivative was responsible for triggering metamorphosis in amphioxus, TRIAC being an interesting candidate. However, very few experiments had been performed to determine TH repertoire in amphioxus (only a 50 year-old article reported endogenous production of TRIAC in amphioxus by **?**). We performed some additional experiments to determine what TH products were produced in amphioxus, and more precisely, if TRIAC was a major physiological TH derivative.

In collaboration with Jean-Pierre Cravedi in Toulouse and Hector Escriva in Banyuls, we first made measurements of T_3, T_4, TRIAC and TETRAC by mass spectrometry on adult extracts and were able to detect T_3 at about 25 ng/g amphioxus (see chapter **??**). In order to improve the detection sensitivity in our experiments, and since amphioxus endogenously produces T_3, we studied the fate of exogenous radiolabelled T_3 and T_4 in amphioxus to determine is amphioxus is able to produce TRIAC from T_3 or T_4. For that purpose, and also in collaboration with J.P. Cravedi and H.Escriva, we treated batches of 10 adults with radiolabelled ^{125}I-T_3 and ^{125}I-T_4. Several conditions were used: larvae were either treated with radiolabeled THs at a concentration of 10^{-10}M, or a mix of 10^{-10}M radiolabeled TH and 10.10^{-7}M cold TH. After a 48h treatment, the organic phase was extracted from treated animals and the remaining radiolabelled compounds were separated by high performance liquid chromatography (HPLC). The results are given below for both T_3 and T_4 (figure **??**).

As can be seen on figure **??**, TRIAC is a major T_3 derivative (figures **??**a and b), and can also be obtained from T_4 (figures **??**e and f). Interestingly, there are also TH derivatives that we could not assign to any other TH derivative we tested: 3,5-T_2, reverse-T_3, diiodotyrosine (DIT), monoiodotyrosine (MIT) or the decarboxylated T_3 derivative triiodothyronamine (TR_3AM) (data not shown). This reveals differences in amphioxus TH metabolism compared to mammals. Indeed, in mammals, reverse-T_3 seems to be produced in greater amounts than in amphioxus, in which it was undetectable. Moreover, TETRAC is more abundant than TRIAC

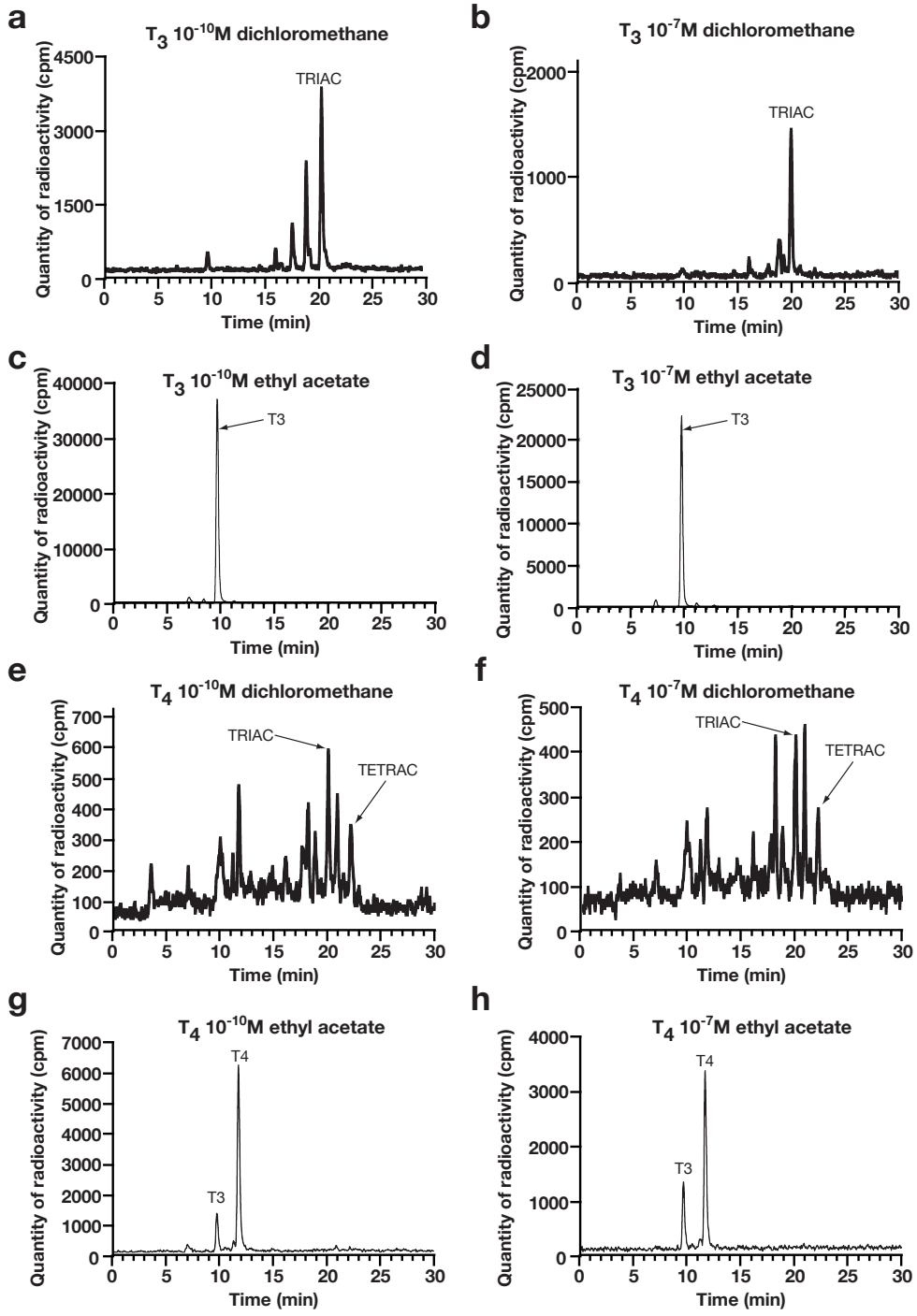


Figure 31: HPLC chromatogram of amphioxus extracts, treated with radiolabelled ^{125}I-T_3 (**a-d**) or ^{125}I-T_4 (**e-h**). Two concentrations were used, either 10^{-10}M (**a,c,e,g**) or 10^{-7}M (**b,d,f,h**). Two extraction methods were also used: dichloromethane (**a-d**) that favors the extraction of TH acetic derivative, and ethyl acetate (**e-h**) that favors extraction of T_4 and its deiodinated derivatives. No T_2, diiodotyrosine (DIT) or monoiodotyrosine (MIT) were detected.

(**?**), whereas the situation seems to be reverse in amphioxus. The other peaks could not be identified. This experiment also suggests that there are active outer-ring deiodinases in amphioxus, since T_3 and TRIAC can be produced from T_4 (figures **??**e-h). Indeed, T_4 to T_3 transformation (prior to TRIAC formation) is done by removing one iodine from the outer ring (figure **??**) of T_4 by deiodinases. As inner ring and outer ring deiodinations are two distinct enzymatic reactions (that may be taken care of by different deiodinases), we can then conclude that amphioxus possesses deiodinases that carry out outer-ring deiodination, which is not surprising as we proposed fron the phylogenetic study of the TH signaling pathway (see chapter **??**), and as in urochordates (**?**).

From this experiment, we can conclude that (i) TRIAC is an important derivative of amphioxus TH metabolism, which strengthens its possible role during amphioxus metamorphosis (see chater **??**), and (ii) although very close, amphioxus TH metabolism probably displays differences with vertebrates, which we anticipated from our genomic analysis of the TH signaling pathway (chapter **??**).

Several important experiments remain to be done. Our conclusions on amphioxus TH metabolism pertain to adult, not on larval metabolism: we have confirmed that amphioxus is *capable* of producing TRIAC but we still do not know whether it does so at larval stage. More generally, we have not brought to light any clue about TH production in general during amphioxus metamorphosis: is there a peak of TH production at the onset of metamorphosis, that accompanies TR expression, as we proposed in chapter **??**? This question will however be difficult to answer. We tried to detect TH metabolism in larvae using a similar protocol as described above for adults. However, we could not collect enough biological material and the results were not conclusive. It will be technically very challenging to collect enough larvae and do such an experiment. An alternative will be to study the expression of the genes involved in TH signaling pathway during development and determine if a peak of expression occurs at metamorphosis (as for deiodinases in amphibians (**?**)). This should allow us to confirm or reject our model described in chapter **??** and discussed in the final discussion (**??**), according to which a peak of TH production and TR expression are the triggering events of metamorphosis in amphioxus.

Alternatively and in order to determine whether a peak of TH production triggers metamorphosis in amphioxus, we studied whether blocking the peak of TH production inhibits amphioxus metamorphosis. This experiment is described in the next section.

8.2 TH SYNTHESIS AND METAMORPHOSIS IN AMPHIOXUS

In chapter **??**, we showed that THs and TH derivatives trigger metamorphosis in amphioxus. In order to confirm the central role of THs in the regulation of amphioxus metamorphosis, we tested the effects of TH synthesis inhibition on metamorphosis. For that purpose, we treated 11 day-old premetamorphic larvae with various compounds known to interfere with TH metabolism in vertebrates and monitored the effect of the compounds on metamorphosis advancement. We thought that this approach would be fruitful because TH production seems to be very similar in amphioxus and in vertebrates (see chapter **??** and the previous section above). We tested the effect of Thiouracil (TU), propylthiouracil (PTU) and 2-mercapto-1-methylimidazole (MMI), known to inferfere with PERT function, as well as potassium perchlorate ($KClO_4$), an anionic competitor of iodine uptake (all these compounds, called goitrogens, were previously used for the basal vertebrate lamprey (**?**)). T_3 and TRIAC were used as positive controls of metamorphosis competence and KCl was used as a negative control (**?**). Results are shown in figure **??**.

As can be seen in figure **??**, the various compounds inhibit metamorphosis with various efficiencies (figures **??**b-f), PTU and MMI being the most efficient ones (figure **??**d and e) and $KClO_4$ having a very mild effect (figure **??**b). As T_3 and TRIAC, but not KCl, induce metamorphosis (figure **??**a), we can conclude that the goitrogen effect reported here is specific. Inteterestingly, exogenous T_3 recovered metamorphosis advancement in treated animals (dashed lines in figures **??**e-f), suggesting that inhibitory effects of goitrogens are effectively due to TH production deficiency.

The inhibition of metamorphosis by goitrogens is only partial, since treated animals seem to eventually metamorphose. The experiment described here was started shortly before natural metamorphosis occurred (metamorphosis started at day 2 in control). It is then possible that goitrogen action was initiated too late for a complete disruption of TH production (e.g. TH levels increase before amphibian metamorphosis starts, figure **??** in chapter **??**). To tackle this issue, we treated younger larvae (4 days after fertilization) with goitrogens: metamorphosis was inhibited with a better efficiency, but it was not completely blocked (data not shown), suggesting that either THs are not the only triggering factor, either goitrogens do not well accumulate in amphioxus tissues, or they do not block PERT of SIS function as well as in vertebrates. It is difficult to decide between these different hypotheses without further biochemical experiments on PERT and SIS function. However, and this is of special interest here, goitrogens inhibit

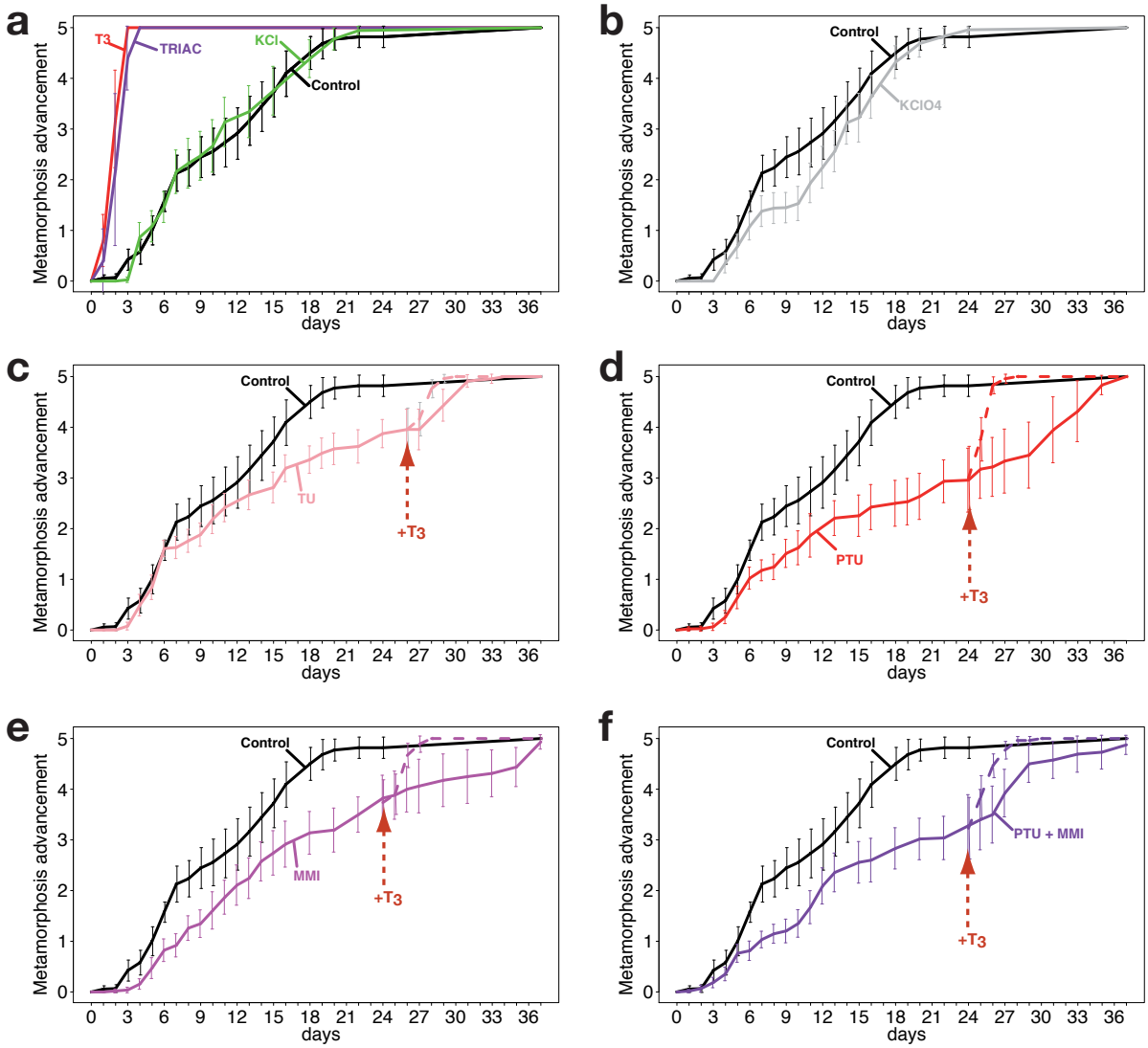


Figure 32: Effects of goitrogens on amphioxus metamorphosis. (**a.**) Vehicle alone (DMSO diluted 100.000 times) or the negative control KCl at 5.10^{-4}M, the positive controls T_3 and TRIAC at 10^{-8}M, (**b.**) $KClO_4$ at 5.10^{-4}M, (**c.**) TU at 5.10^{-4}M, (**d.**) PTU at 5.10^{-4}M, (**e.**) MMI at 5.10^{-4}M, or (**f.**) PTU at 5.10^{-4}M and MMI at 5.10^{-4}M. The effect on metamorphosis was monitored with five specific morphological criteria from no metamorphosis (0) to full metamorphosis (5) (fully described in chapter **??**). After 24 or 26 days, batches were split in two and T_3 was added to one half of each batch, to determine if metamorphosis advancement could be rescued (indicated in dashed lines). Error bars correspond to a 5% confidence interval.

metamorphosis, thus confirming our results from chapter **??** that THs are important regulators of metamorphosis in amphioxus. It also suggests that a peak of TH production is an important event at the onset of metamorphosis.

8.3 STRUCTURAL BASIS OF TR ACTION

We previously showed that THs are important regulators of metamorphosis in amphioxus, with TRIAC as the potential active form. More generally, we concluded that TH signaling pathway is similar in amphioxus and in vertebrates, except that the active TH is different: T_3 in vertebrates and a T_3 derivative, potentially TRIAC, in amphioxus (chapters **??**, **??** and above). This difference reflects in the differences of TR functioning in both groups: in vertebrates (and especially mammals), TR is activated by T_3 and TRIAC (although *in vivo* TRIAC levels are kept very low), whereas the TR from amphioxus (amphiTR) is activated by TRIAC but not by T_3 (**?**). To better understand the structural basis of this difference, the AliX company from Strasbourg reconstructed the 3D structure of the amphioxus TR LBD in complex with TRIAC.

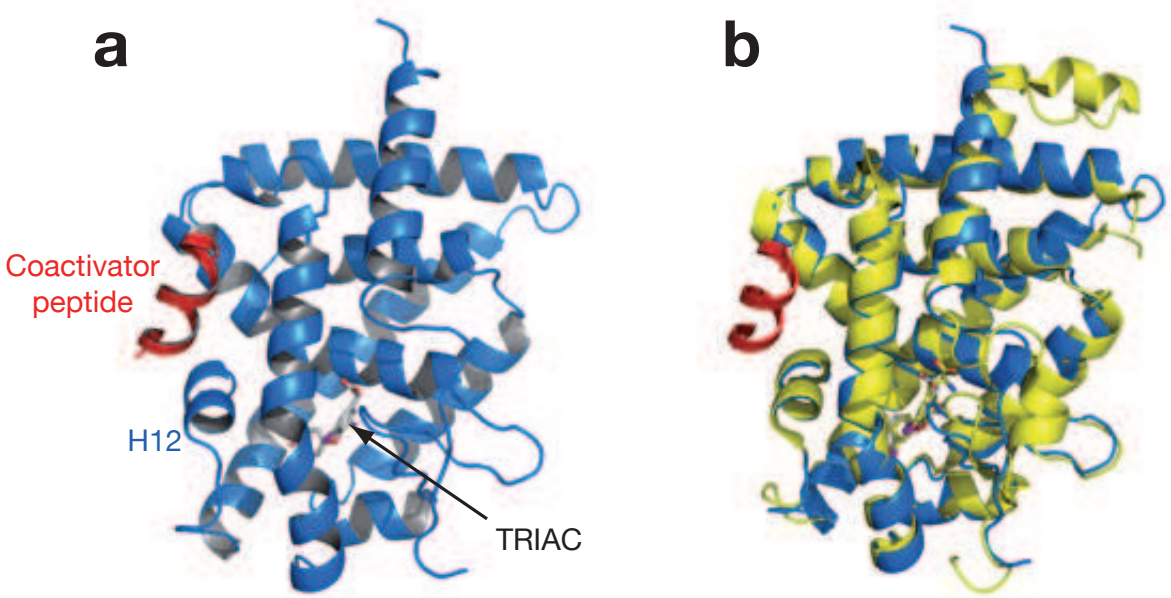


Figure 33: 3D structure of amphiTR LBD in complex with TRIAC. (**a.**) The amphiTR LBD is organized into 12 α helices enclosing a ligand binding pocket that can contain TRIAC. A peptide comprising the interaction interface of the coactivator NCOA2 (in red) was added to stabilize the holo conformation. (**b.**) The amphiTR LBD-TRIAC (blue) almost perfectly superimposes with the human TRα LBD in complex with T_3 (in yellow). The 3D structure was obtained by the AliX company (Strasbourg).

As expected, amphiTR LBD displays the classical NR organization in 3 layers of α helices and contains a ligand binding pocket, in which TRIAC is located. Morever, amphiTR is in a holo conformation, with the helix 12 close to helices 11, 5 and 3 (figure **??**a). The conservation of the 3D structure is more conspicuous in figure **??**b, in which the structures of amphiTR in complex with TRIAC almost perfectly superimposes with the one of human TRα in complex with T_3 [ACC:2H79], despite a high sequence divergence (less than 40 % identity with human TRs (**?**)). Homologous sites in the LBP interact with the ligand in both amphiTR and human TRα: the histidine H397 interacts with the alcohol group of the outer ring (compared to H381 for human TRα), and R283 interacts with the alcohol group in the lateral chain (compared to R266 for human TRα) (figure **??**). Several other amino acids (F294 and R279) interact with the latter alcohol group and stabilize the interaction between amphiTR and TRIAC (figure **??**). This confirms the similar ligand binding behavior or amphiTR and human TRα.

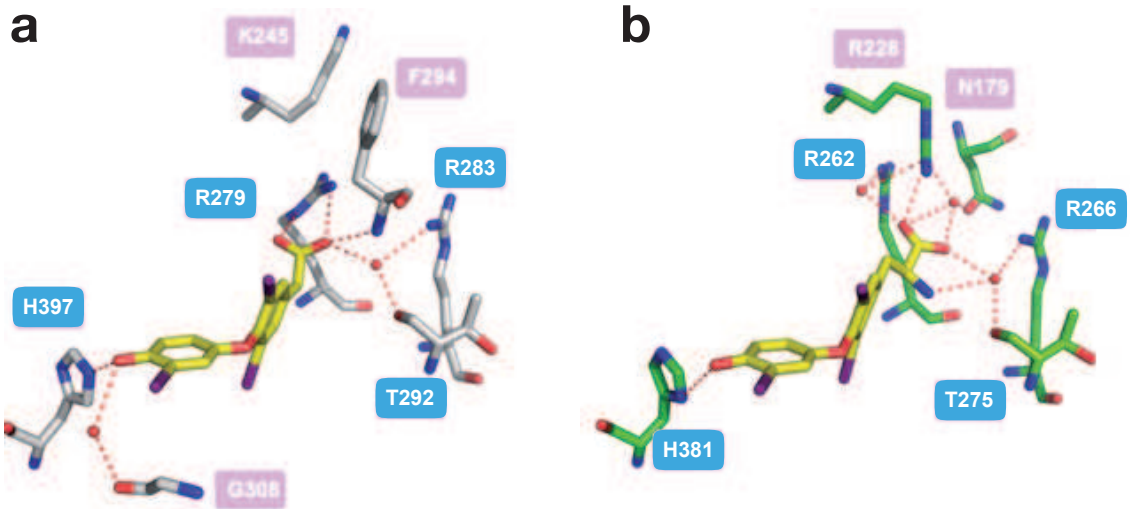


Figure 34: 3D view of the amphiTR (in grey, **a**) and human TRα (in green, **b**) residues that participate to interaction with TRIAC (yellow in **a**) and T_3 (yellow in **b**), respectively. Hydrogen bonds are indicated by red dashed lines. Conserved site between human TRα and amphioxus TR are highlighted in blue whereas divergent sites are highlighted in pink. The 3D structure of the amphioxus TR was obtained by the AliX company (Strasbourg).

The main difference between amphiTR and human TRα lies in the size of the ligand binding pocket. Notably, the homologous site to a histidine at position 177 in the helix 3 of human TRα is a tryptophane in amphiTR (at position 176, figure **??**). The big aminoacid W176 induces a shift of R279 in helix 3 towards the LBP in amphiTR, compared to R262 in human TRα (figure **??**), leading to a reduction of the LBP (figure **??**). The longer side chain of T_3 compared to TRIAC makes it

too large to fit in amphiTR LBP and thus strongly decreases amphiTR affinity to T_3. On the opposite, the shorter TRIAC may enter amphiTR LBP and may interact with some residues to stabilize amphiTR active conformation.

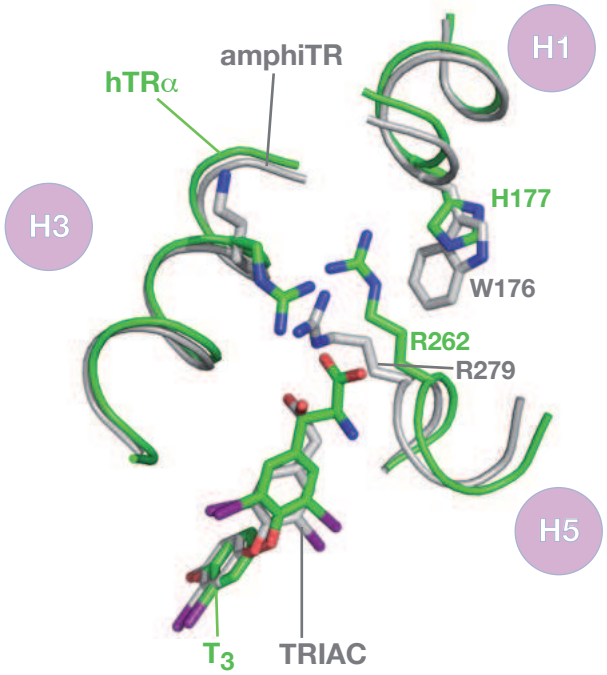


Figure 35: 3D view of superimposition of the amphiTR (in grey) and human TRα (in green) in interaction with TRIAC (in grey) and T_3 (in green), respectively. Only parts of helices 1, 3 and 5 have have been represented, as well as some of the residues forming the LBP and explaining differences of amphiTR and human TRα in their T_3 binding affinities. The 3D structure was obtained by the AliX company (Strasbourg).

In sum, these data confirm the conservation of amphiTR LBD structure and function with vertebrate TRs. However, specific divergent sites explain the better affinity of amphiTR to TRIAC than to T_3 and confirm our previous results *in vivo* (chapter **??**). Interestingly, the unique TR from the sea urchin *Strongylocentrotus purpuratus* is vertebrate-like at most sites of its LBD, that are important for ligand binding (**?**), suggesting that it may also display a similar ligand binding pattern and be able to interact with T_3.

9

AN AMPHIOXUS ORTHOLOGUE OF THE ESTROGEN RECEPTOR THAT DOES NOT BIND ESTRADIOL: INSIGHTS INTO ER EVOLUTION

Nuclear hormone receptors have their activity regulated upon ligand binding, a phenomenon that allows them to interact with the environment and also to regulate transcription. How this dual function arose during evolution deserves some insight. There is little doubt that all NRs are homologous (**?**) and evolved from a common ancestor to the diversity encountered now (**?**). If one considers NR modular organization (as depicted in chapter **??**), two domains, the DBD and the LBD are of particular interest.

The evolution of DNA binding ability is rather straightforward: almost all NRs have very close DNA binding properties. Indeed all DBDs interact with DNA through zinc-fingers and bind very similar combinations of more or less degenerate (AGGTCA) motifs (see section **??**). Differences lie in binding site specificity, whose alteration may occur through point mutations (which are rather limited considering the similarities between specificities between different NRs). It was proposed that this zinc-finger DBD, that is specific to NRs, may be homologous with the DNA binding elements from LIM and GATA proteins, that are found in both metazoans and yeast (**??**). Moreover, NRs typically bind DNA as dimers, which confers most of DNA binding specificity. Then dimerization is probably ancestral and its evolution may be of greater importance than the evolution the the DBD regarding the evolution of DNA binding ability. Interestingly, some NRs are able to interact with several other members of the family (e.g. TR can interact with VDR or RXR, or itself). This conservative evolution suggests that the ancestral NR was a zinc-finger transcription factor that bound DNA as a dimer (as a homodimer or a heterodimer) on motifs around the canonical (AGGTCA) one. Moreover, as most NRs are transcription factors, the ancestral NRs were most probably transcription factors.

In contrast, the evolution of ligand binding ability and more generally of transcriptional activation, is more elusive and has been discussed in the past decade (e.g. (**?????**)). LBDs function in diverse ways, so it is difficut to elaborate upon the evolution of ligand binding. First, and as mentioned in the introduction (section **??**), a large portion of NRs are still orphan, meaning that no known ligand has

been identified so far. Some NRs may have a ligand *in vivo* that has not been identified yet (despite extensive search, like for ERR (**?**)). However the activity of others may be regulated by ligand-independent mechanisms. For instance Nurr1 has no ligand binding pocket (some may talk about a "filled" pocket), and its activity is regulated by post-translational modifications like phosphorylation (**?**). Other orphan receptors, such as some USPs (the insect RXR orthologue), do have a pocket but a USP-specific loop interacts with the LBP and prevents ligand binding (**?**). Second, ligand-activated receptors show important ligand binding differences. Some NRs, like TR, display a high affinity and a highly selectivity for their ligand. Accordingly, they have a small and constrained ligand binding pocket. Alternatively, other NRs like LXR have larger pockets combined to a wider repertoire of low affinity ligands (**?**). Notably, the NR PXR binds to a large diversity of foreign compounds and is involved in detoxification (**?**). A last category of NRs have a non-exchangeable ligand (referred to as constitutive, e.g. RORα and cholesterol). Consequently transcriptional activity the receptor is modulated by other means (e.g. expression, post-translational modifications...) (**?**). How did this diversity evolve, and what were the ligand binding properties of the ancestral receptors?

The above diversity of ligand binding profiles should be put in a phylogenetic perspective. Orphan receptors, high-affinity and low-affinity NRs are not grouped together but instead are scattered along the NR tree (figure **??**). Moreover, the chemical nature of ligands does not fit NR phylogeny either. Indeed, closely related NRs may bind chemically very different ligands (e.g. TR and RAR) whereas distantly related NRs may bind similar ligands (ecdysteroid receptors and other sex steroid receptors, figure **??**). Overall, considering NR phylogeny does not help to understand the distribution of NR ligand binding characteristics.

How this ligand binding situation evolved is then far from clear. Several hypotheses have been presented. One proposes that the ancestral receptor was a high affinity receptor that subsequently evolved new binding affinities after gene duplication and divergence (**??**). In contrast, ligand binding may also have been secondarily acquired from an orphan ancestor (**??**) that subsequently gained the ability to be regulated upon ligand binding.

The following article is in accord with previous works on the evolution of NR ligand binding. More precisely, it deals with the evolution of estrogen binding in the NR3 subfamily (see figure **??**, as well as figure **??** in chapter **??**). The article is divided into two parts. First, we used a comparative approach and focused on the lamprey and amphioxus estrogen receptors (ER): we characterized their DNA and ligand binding as well as their transactivation abilities to show that lamprey ER behaves as a "classical" vertebrate ER whereas the amphioxus ER is not an

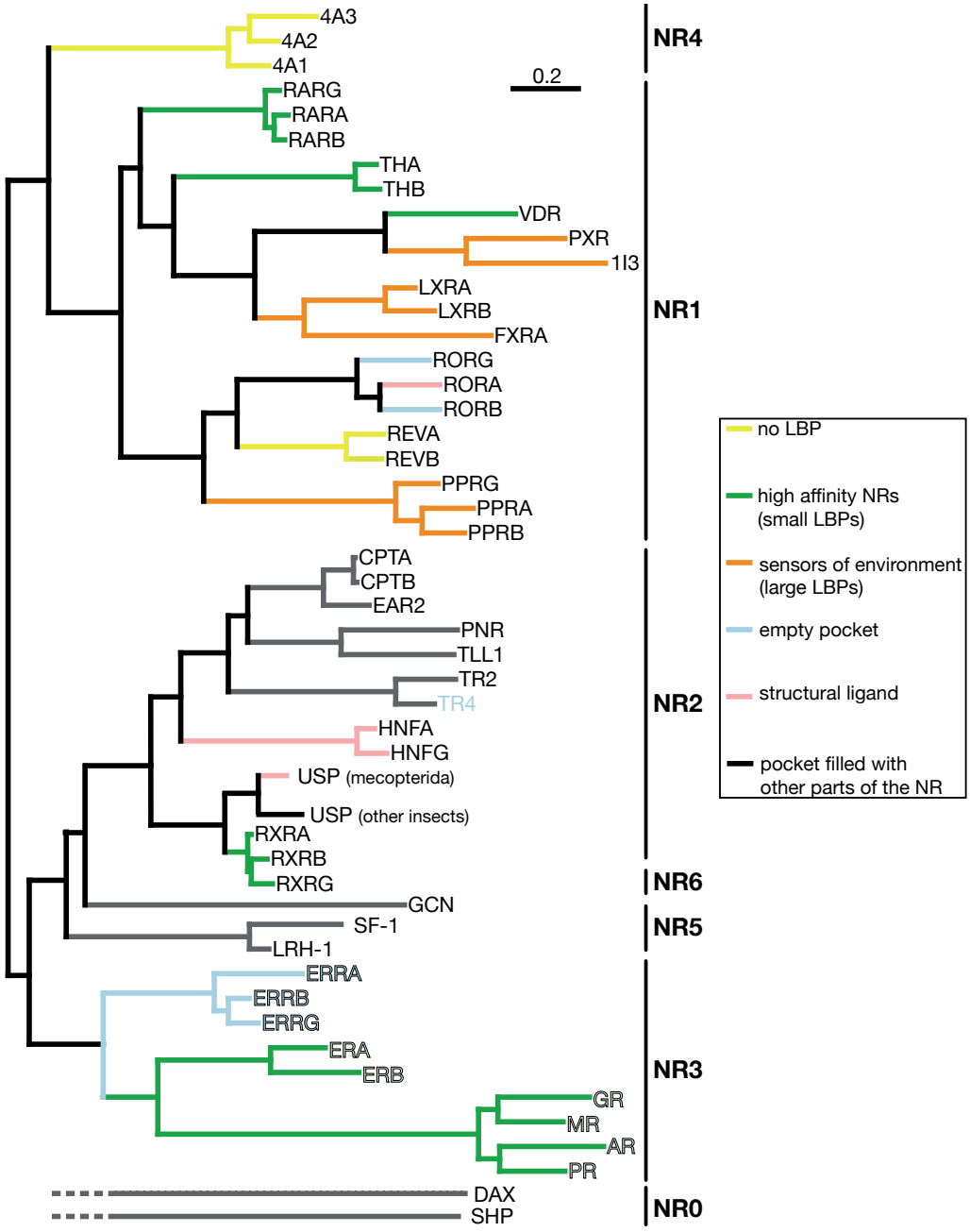


Figure 36: Phylogenetic tree of human NRs, color-coded with ligand binding properties. Some NRs are high affinity NRs (green, it is still unclear which ligand is the natural RXR ligand (**?**)). Some NRs have large pockets and are considered as "sensors" of nutritional environment (orange) (**?**). Other NRs have a structural ligand which is not involved in modulation of activity (pink). Some NRs are kept in an unliganded conformation, and have an empty pocket (blue), a pocket filled with other parts of the NR (USP in black) or no ligand binding pocket (yellow). The situation is still unclear for many other NRs (grey).

estrogen receptor. In a second time, we used computational methods to discuss the pros and cons of an approach different from comparative endocrinology: instead of comparing ERs from extant species to understand the evolution of ER ligand binding ability, one may predict *in silico* ancestral sequences, "resurrect" and molecularly characterize them, to infer biochemical properties of the real ancestral ERs.

This article was published in BMC Evolutionary Biology.

BMC Evolutionary Biology

Research article

An amphioxus orthologue of the estrogen receptor that does not bind estradiol: Insights into estrogen receptor evolution

Mathilde Paris[1], Katarina Pettersson[2], Michael Schubert[1], Stephanie Bertrand[3], Ingemar Pongratz[2], Hector Escriva[3] and Vincent Laudet*[1]

Address: [1]Institut de Génomique Fonctionnelle de Lyon, Molecular Zoology team, Université de Lyon, Ecole Normale Supérieure de Lyon, Université Lyon 1, CNRS, INRA, Institut Fédératif 128 Biosciences Gerland Lyon Sud, France, [2]Karolinska Institutet, Department of Biosciences and Nutrition, S-141 57 Huddinge, Sweden and [3]CNRS, UMR 7628, Modèles en Biologie Cellulaire et Evolutive, Observatoire océanographique, F-66651, Banyuls/mer, France

Email: Mathilde Paris - mathilde.paris@ens-lyon.fr; Katarina Pettersson - kape@biosci.ki.se; Michael Schubert - michael.schubert@ens-lyon.fr; Stephanie Bertrand - stephanie.bertrand@obs-banyuls.fr; Ingemar Pongratz - ingemar.pongratz@biosci.ki.se; Hector Escriva - hector.escriva@obs-banyuls.fr; Vincent Laudet* - vlaudet@ens-lyon.fr

* Corresponding author

Published: 25 July 2008

BMC Evolutionary Biology 2008, **8**:219 doi:10.1186/1471-2148-8-219

Received: 3 January 2008
Accepted: 25 July 2008

This article is available from: http://www.biomedcentral.com/1471-2148/8/219

Abstract

Background: The origin of nuclear receptors (NRs) and the question whether the ancestral NR was a liganded or an unliganded transcription factor has been recently debated. To obtain insight into the evolution of the ligand binding ability of estrogen receptors (ER), we comparatively characterized the ER from the protochordate amphioxus (*Branchiostoma floridae*), and the ER from lamprey (*Petromyzon marinus*), a basal vertebrate.

Results: Extensive phylogenetic studies as well as signature analysis allowed us to confirm that the amphioxus ER (amphiER) and the lamprey ER (lampER) belong to the ER group. LampER behaves as a "classical" vertebrate ER, as it binds to specific DNA Estrogen Responsive Elements (EREs), and is activated by estradiol (E_2), the classical ER natural ligand. In contrast, we found that although amphiER binds EREs, it is unable to bind E_2 and to activate transcription in response to E_2. Among the 7 natural and synthetic ER ligands tested as well as a large repertoire of 14 cholesterol derivatives, only Bisphenol A (an endocrine disruptor with estrogenic activity) bound to amphiER, suggesting that a ligand binding pocket exists within the receptor. Parsimony analysis considering all available ER sequences suggest that the ancestral ER was not able to bind E_2 and that this ability evolved specifically in the vertebrate lineage. This result does not support a previous analysis based on ancestral sequence reconstruction that proposed the ancestral steroid receptor to bind estradiol. We show that biased taxonomic sampling can alter the calculation of ancestral sequence and that the previous result might stem from a high proportion of vertebrate ERs in the dataset used to compute the ancestral sequence.

Conclusion: Taken together, our results highlight the importance of comparative experimental approaches vs ancestral reconstructions for the evolutionary study of endocrine systems: comparative analysis of extant ERs suggests that the ancestral ER did not bind estradiol and that it gained the ability to be regulated by estradiol specifically in the vertebrate lineage, before lamprey split.

Background

Hormone signaling is a very important feature in metazoans, allowing communication between cells or organs within the organism. Two components of these signaling systems are of particular importance, the hormone and its receptor. The nuclear hormone receptor (NR) superfamily includes ligand dependent transcription factors that play a central role in various physiological processes as diverse as reproduction, development, and control of homeostasis [1,2]. They share a common structural organization and exhibit a highly conserved DNA binding domain (DBD) and a moderately conserved ligand-binding domain (LBD). Some members of this superfamily are liganded receptors (24 among the 48 genes encoding NRs in the human genome) but many lack identified ligand and are therefore called "orphan" [3]. Some orphan receptors are 'true' orphans in the sense that they do not possess a *bona fide* ligand-binding pocket (LBP), like the members of the NR4 subfamily (for instance, NURR1, DHR38 or NGFI-B. For review, see [4]), and are regulated by other mechanisms [4]. Alternatively, the crystal structures of several orphan receptors such as HNF4 were found to have a phospholipid constitutively bound to a large ligand binding pocket [5,6]. The functional and evolutionary implications of these constitutive ligands remain discussed. Other orphan nuclear receptors have a ligand binding pocket and thus have the potential to bind compounds. It is still not known whether those receptors have natural ligands, still to be discovered. Undoubtedly, the existence of such orphan receptors with physiological or developmental activities constitutes both a major challenge for understanding nuclear receptor evolution and a potential opportunity for pharmacology [1].

The existence of orphan and liganded members in the NR family raises the question of the evolution regarding their ligand binding ability. Whether the ancestral NR was liganded or orphan and more generally how NR ligand binding ability evolved has been recently debated [7-14]. In general, it is still unclear if there is a correlation between the evolution of the hormone repertoire and NRs. Moreover the mechanisms underlying this coevolution are of particular interest [7,12,15-19].

Among the scenarios of NR evolution that have been proposed, one suggests that the ancestral NR was a ligand-independent transcription factor which acquired the ability to be regulated by ligands several times during evolution [7,18-20]. This hypothesis was based on the observation that compounds of similar chemical nature bind to divergent NRs and on the contrary compounds of very different nature bind to closely related receptors. For instance, orphan receptors are found in all families of NRs, and steroid receptors are not monophyletic but are located in two different subfamilies within the NR superfamily: the ecdysteroid as well as the sex steroid receptors. Interestingly, the evolution of sex steroid hormone receptors has also been used as an argument for an alternative hypothesis, the ligand exploitation model [8,11] (for an alternative view, see [21,22]). Phylogenetic trees show that sex steroid hormone receptors are grouped with ERRs as the NR3 subfamily, following the official nomenclature [23]. They contain receptors that bind estradiol (ERs), that form the NR3A group as well as mineralocorticoids (MRs), glucocorticoids (GRs), progesterone (PRs), and androgen (ARs) that form the NR3C group. All known ligands in this subfamily can be seen as variations around the archetypical sterol skeleton. Consequently, Thornton et al. suggested in the ligand exploitation model that the ancestral steroid receptor was a high affinity estradiol receptor [8,11] and the other steroid receptors that originated later on, experienced, following gene duplication, shifts in their binding affinities to eventually bind to their extant ligand. The model in fact suggests that the newly duplicated receptors (here NR3C) exploit as ligands chemical species that serve as intermediary compounds in the "ancestral ligand" synthesis pathway (here the estradiol synthesis pathway) [8]. According to this view, orphan receptors, like ERRs, secondarily lost the ability to have their activity regulated by a ligand and became orphan. Interestingly, within the NR3 family, two receptor subfamilies, ERRs and ERs, appear to be ancient since they are found in a wide variety of metazoans including deuterostomes and protostomes, whereas, up to now, MRs, GRs, PRs and ARs have been found only in vertebrates. The only non-vertebrate ERs that have been described so far were from mollusks and were shown to be unable to bind estradiol [11,12,24-26]. Since the ligand exploitation model implies an ancestral estradiol-binding ER and since all liganded ER found so far come from vertebrates, and to improve taxonomic sampling, the ER orthologues from the basal vertebrate lamprey and the invertebrate chordate amphioxus were characterized here. Indeed, lamprey and amphioxus are located at key positions in the chordate phylum [27-30]. Moreover, amphioxus (*Brachiostoma floridae*) is much less derived than urochordates in its morphology as well as in its genome organization [30]. Indeed, amphioxus and vertebrates share a similar general body plan whereas urochordate morphology is more derived. For instance, during metamorphosis of some urochordates, the tadpole-like larva transforms into an adult that looks so different that it was first considered as a mollusk [31]. Moreover the urochordate genome is fast evolving [27], with for instance the loss of the clustering of the hox genes [32]. There is no ER in the sequenced genome of *Ciona intestinalis* [33] or in the sea urchin [34], one ER was previously cloned in lamprey [8], only one ER was found in the amphioxus genome [35]. These reasons make lamprey and amphioxus excellent models to study the evolution of estrogen signaling pathway at the origin

of vertebrates. In this study, we cloned the unique ER from amphioxus (amphiER) and characterized it, as well as the previously cloned but uncharacterized lamprey ER (lampER). AmphiER is an orphan receptor, showing no affinity to the estrogen hormone estradiol, when in contrast, the lamprey ER behaves as a "classical" vertebrate ER. As no ER from invertebrates studied so far binds estradiol, we propose that the ancestral ER (and the ancestral steroid receptor) was not a receptor for estradiol and gained later on during evolution the ability to bind the hormone.

Results

Cloning of the ER from amphioxus (amphiER)

Using degenerate primers designed to match motifs in the most conserved part of vertebrate ERs in the DNA binding domain, a single gene fragment from total RNA of an adult *Branchiostoma floridae* was amplified, cloned and sequenced. Rapid amplification of cDNA ends (RACE) was utilized to obtain the full-length cDNA. From this sequence, a new set of specific primers were designed and used to amplify the full length open reading frame of this gene. The obtained cDNA [GenBank: ACF16007] is 2118 bp long and encodes a 705 aa long putative protein (Figure 1) that harbors the classical features of an ER with the 5 main functional domains (Figure 2A), among which a highly conserved DNA binding domain (DBD) and a less conserved ligand binding domain (LBD). The DBD shares an 82% sequence identity with the human ERα one (83% with human ERβ) and much less with the other NR3 receptors (<62%). The same pattern is observed for the LBD, although this domain is less conserved since it exhibits only 34% amino acid identity with human ERα (35% for human ERβ) and about 20% with other steroid receptors (Figure 2A). The three other domains, namely the A/B region in the N-terminal part, the hinge between DBD and LBD, and the short C-terminal end of the protein, are more divergent, which is a general pattern for NRs [2] (Figure 2A). The recent release of the amphioxus genome confirmed the presence of a single ER gene [35]. In contrast the previously described lamprey ER is more similar to the human ERα with its DBD sharing a 93% sequence identity (93% for human ERβ) and its LBD sharing 55% sequence identity (56% for human ERβ) [8].

Phylogenetic analysis of ERs

The orthology relationships of the amphioxus and lamprey ER sequences were studied in a phylogenetic analysis of the NR3 family using an exhaustive dataset comprising 69 members of the NR3 subfamily as well as sequences of RXRs as an outgroup. The dataset included the 6 currently known mollusk ER sequences (from *Nucella lapillus*, *Crassostrea gigas*, *Marisa cornuarietis*, *Thais clavigera*, *Octopus vulgaris*, *Aplysia californica*), as well as the 2 NR3 sequences previously known from amphioxus (1 ERR [GenBank: AAU88062] and 1 NR3C [JGI: 201600], retrieved from a previous work [36] or from the complete genome sequence [35]). In the resulting phylogenetic tree, the sequence of lamprey ER branches within the ER clade with a high bootstrap support (95%), at the expected position before the split of vertebrate ERα and ERβ (Figure 2B, and for a tree presenting all sequences, see Additional file 1A), as previously shown [8]. The sequence of amphiER branches within the ERs and is located at the base of the vertebrate estrogen receptor group, before the split of ERα and ERβ but after the split of the mollusk ERs (bootstrap value of 81%, Figure 2B). However its precise position within the ER group is poorly supported (bootstrap value of 42%).

Such low bootstrap supports reveal either the weakness of the phylogenetic signal contained in ER proteins, or the presence of two incompatible signals in the data, one supporting the observed position of amphiER within ERs, and the other supporting another position. Whereas the weakness of the signal is not testable, the long branch leading to amphiER in the tree suggests that sites that have undergone a large number of substitutions may account for one of the two signals. Such sites may be saturated to the point that phylogenetic methods are not able to correctly recover their evolution, a situation leading to the long branch attraction artifact [37]. It is therefore important to correctly characterize sites that support the ER position of amphiER: if only fast-evolving sites support this hypothesis, it is probably due to long branch attraction, and an alternative branching should be favored. Alternatively, if slowly-evolving sites support this position, one can confidently identify amphiER as a *bona fide* ER. To characterize sites with respect to their evolutionary rates and the amphiER position they favor, both site likelihoods and site evolutionary rates were computed for all possible positions of amphiER.

First, AmphiER was pruned from the tree shown in Figure 2B, and then re-grafted in all 149 remaining branches. This yielded 149 topologies, for which site likelihoods and site evolutionary rates could be computed using PhyML-aLRT. This allowed us to obtain an evolutionary rate per site averaged over all possible positions of amphiER, and therefore independent from the precise position of amphiER in the tree. Additionally, as likelihoods were computed for each of the 149 positions, these positions could be compared according to the Approximately Unbiased test (AU test, implemented in Consel [38]). Out of all the 149 resulting trees, 26 could not be distinguished with the AU test and had a likelihood significantly better than all the other ones (p-value > 0,05). Of these 26 topologies, all but three place amphiER within the ER clade ("ER" trees). The remaining topologies ("alter-ER" trees) place amphiER either at the base of the NR3C clade (comprising the ARs, PRs, MRs and GRs), within the NR3C or at the

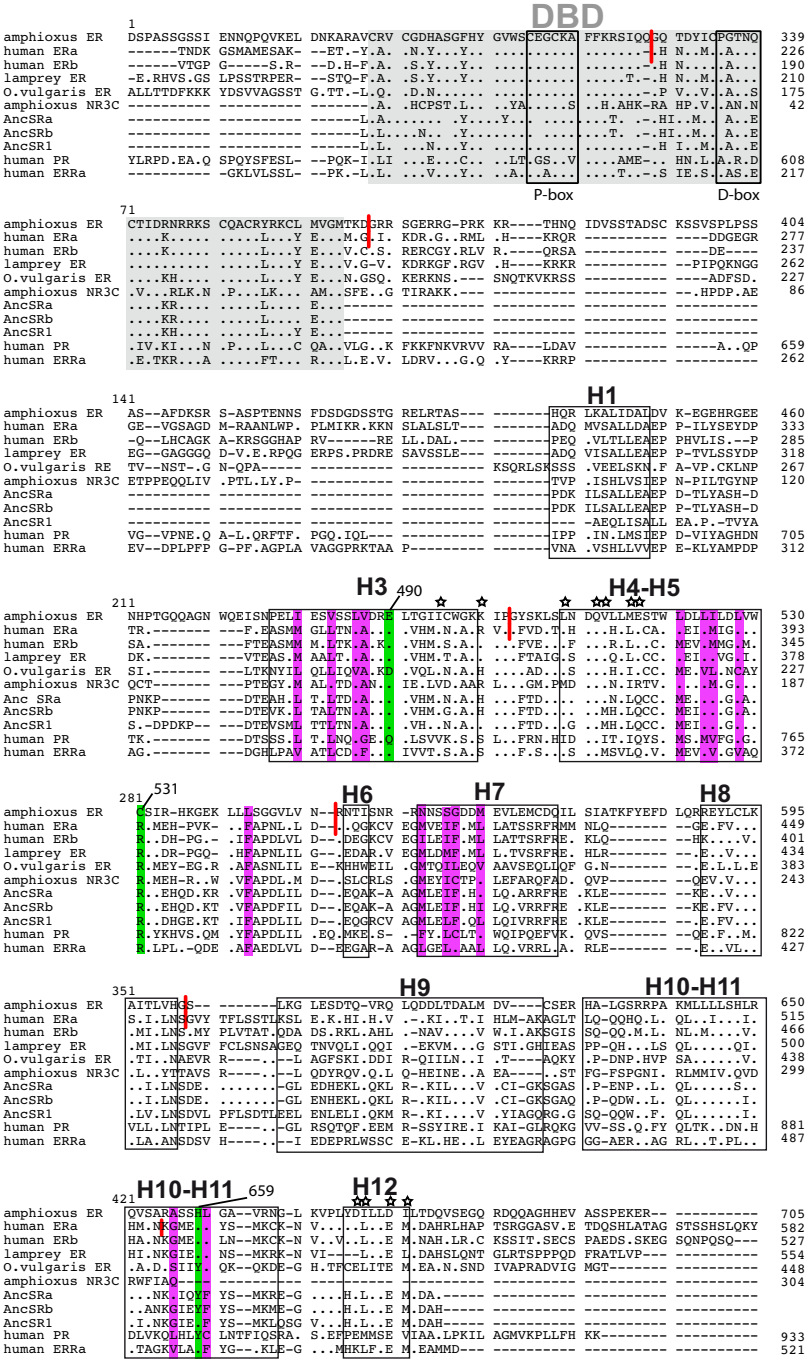


Figure 1

Sequence alignment of several ERs including the amphioxus ER and the lamprey ER, as well as ancestral steroid receptors. The DBD is highlighted with light grey. The 12 helices from the LBD are indicated, based on the known 3D structure of human ERα [49]. Amino acids from human ERα making direct hydrogen bonds with E_2 are indicated in green. Amino acids making hydrophobic bonds with E_2 are highlighted in purple. Amino acids known to be involved in co-activator interaction have been indicated with a star on top of each site [55]. The more divergent A/B domain as well as the F domain have been omitted from the alignment. However, the numbering of the sites along the alignment starts at the beginning of each protein. The exon-intron limits of amphiER and humanERα have been indicated with small red strokes. The sequences of AncSRa and AncSRb have been inferred in this study. The sequence of AncSR1 was retrieved from a previous analysis [11].

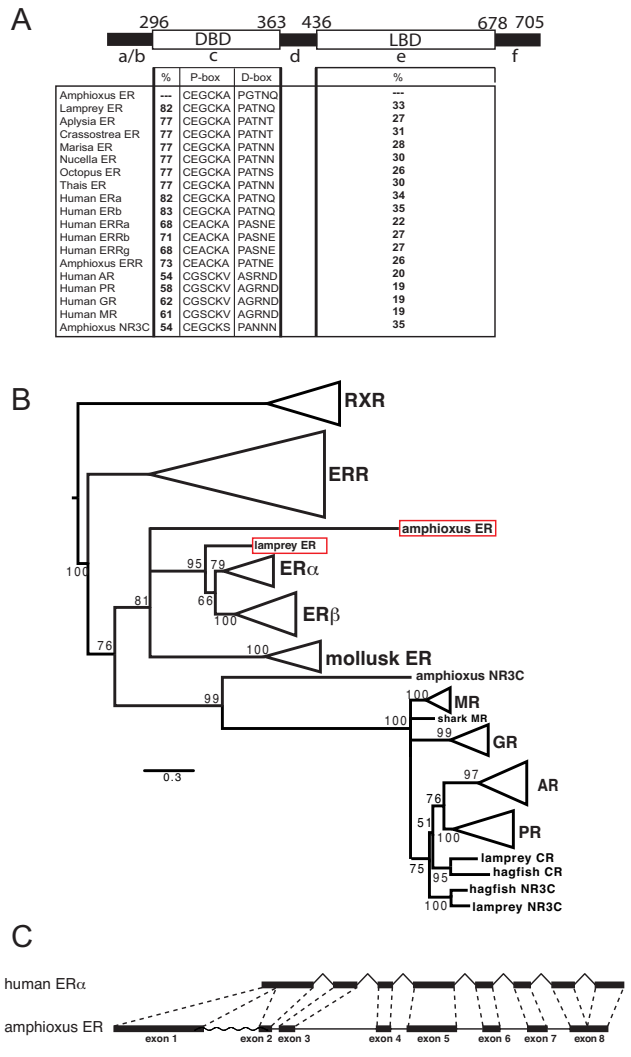


Figure 2

There is a single ER in amphioxus. (A) Schematic representation of the different domains of amphiER. Percent identity of the amphioxus ER with other sequences from the NR3 subfamily in the DNA- and ligand-binding domains is indicated. Amino acid sequence of the highly conserved P-box and D-box in the DBD are shown. (B) Maximum likelihood (ML) tree obtained from the analysis of the amino acid sequences of the DBD and the LBD of a wide range of NR3 under a JTT+γ +i model. Bootstrap percentages obtained after 1,000 ML replicates are shown above selected branches. Scale bar indicates number of changes per site. The tree was rooted by selected RXR sequences. (C) The exon-intron structure of amphiER is conserved with that of human ERα, except two minor differences: the first human exon corresponds to the first two amphioxus exons and the last two human exons correspond to the last amphioxus exon.

base of (ER, NR3C) (Figure 3A). Because site evolutionary rates had been computed, sites having a higher likelihood for the "alter-ER" trees could be compared with sites favoring the "ER" topologies with respect to their evolutionary rates. Interestingly, the sites pleading for the "alter-ER" trees evolve significantly faster than the sites pleading for the "ER" trees (mean evolutionary rates of 1.20 and 0.90, p-value < 10^{-5} with a Wilcoxon-test or $p < 0.001$ with a an unpaired t-test). This suggests that the "alter-ER" signal in the alignment is probably due to long branch attraction to the NR3C subtree, which might also be at the origin of the low bootstrap support (42%) for the position of amphiER. Conversely, this suggests that amphiER should be considered as an ER, as the signal at the origin of this position does not seem to be artifactual.

An additional test can be run to further confirm this hypothesis, and consists in reestimating the phylogeny using only slowly-evolving sites. For that purpose, the distribution of expected relative evolutionary rates across sites of the alignment was plotted, as found by phymlaLRT [39,40] (Figure 3B). Fastest-evolving sites were removed from the dataset based on three different rate thresholds (2.5, 2 or 1.5, Figure 3B and 3C), and trees were reconstructed based on the alignments containing only the remaining slowly-evolving sites. These operations did not impact the monophyly of ERs (Figure 3C) or the statistical support. This shows that the clustering of amphiER with vertebrate ERs does not come from saturated sites, which argues against long branch attraction being at the origin of this position [41]. Accordingly, complementary phylogenetic analyses with different methods (bayesian, Neighbor-joining, parsimony) gave similar results (see Additional file 1). From these studies we conclude that amphiER does indeed belong to the ER subfamily, which is confirmed by the general conservation of the exon-intron structure of amphiER with human ERα, especially at two exon-intron splice sites in the DBD and in the LBD after helix 3 (Figure 2C and short red strokes in Figure 1) [42].

Chordate ERs, including amphioxus ER and lamprey ER, are able to bind estrogen specific response elements (ERE)

To test whether the lamprey ER and the amphioxus ER are able to bind DNA on specific estrogen response elements (EREs), electrophoresis mobility shift assays were performed using a radiolabeled consensus ERE sequence (see Additional file 2). These experiments show that, like vertebrate ERs, amphiER and lampER are able to bind DNA specifically on a consensus ERE. This binding is specific, since a 100-fold excess of non-specific DNA was not able to compete for binding, whereas a 100-fold excess of cold ERE completely suppressed it (see Additional file 2, compare lanes 15 and 17, as well as lanes 19 and 21). ERs contain two major conserved signatures in the DBD, the P-box (CEGCKA), responsible for the binding specificity to response elements, and the D-box, also involved in the

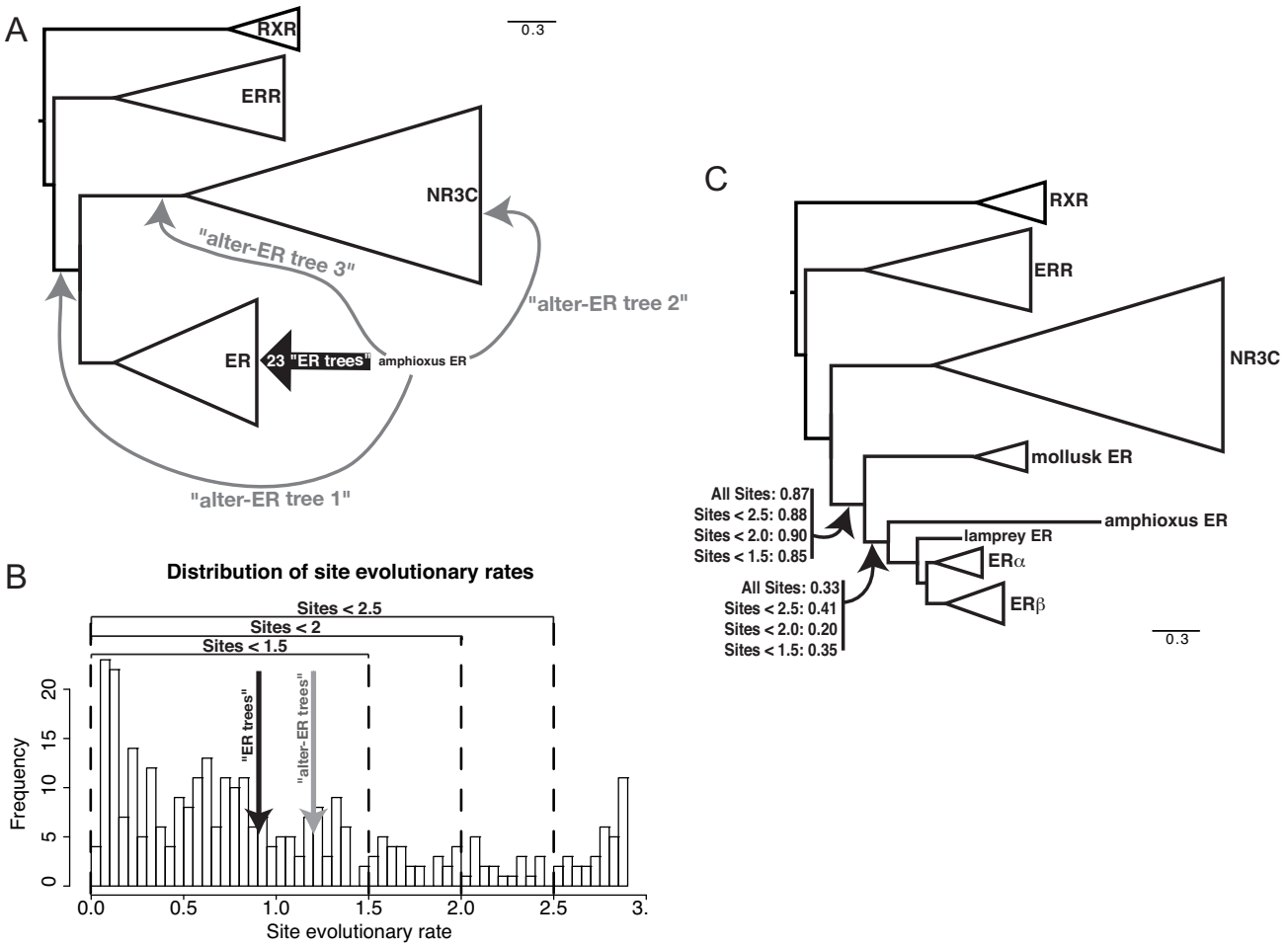


Figure 3

Phylogenetic analysis of amphiER. (A) Summary of the 26 best placements of amphiER within a phylogenetic tree comprised of 69 NR3 and 8 RXR sequences. 23/26 topologies ("ER-tree") place amphiER within the ER clade, the 3 remaining topologies ("alter-ER tree") place amphiER either at the base of (ER, NR3C) or within the NR3C family (close to an AR) or at the base of the NR3 family. The mean evolutionary rate of the sites supporting one of the 23 "ER-tree" topologies (0,9) or for the "alter-ER tree" topologies (1,2) are indicated in (B). (B) Distribution of the site relative evolutionary rates. Rates were estimated using an 8 class discretized gamma distribution. The vertical dotted lines correspond to different tentative threshold (2.5, 2, 1.5), above which sites have been discarded due to their high evolutionary rate, before reestimating the phylogeny of the consecutive alignment. (C) Estimation of the minimum of Chi2-based and SH-like supports, available in the aLRT-PHYML software, for the branches defining the monophyly of ERs as well as the position of amphiER. 4 trees were inferred using an alignment on which the fastest evolving sites were removed (no site removed, 34, 53 and 82 sites removed out of 323, with a mean evolutionary rate threshold above 2.5, 2.0 and 1.5, respectively).

DNA binding specificity of the ER dimers (Figure 2A). The P-box is highly conserved in all known ERs, including amphiER and lamprey ER and is different from other NR3 members. AmphiER and lamprey ER also have a well conserved D-box, amphiER D-box containing just a few conservative mutations, (*e.g.* a mutation of an alanine in glycine, Figure 2A). Since the three characterized mollusk ERs (from *A. californica*, *O. vulgaris* and *Thais clavigera* [11,12,24]) also bind EREs and since the P-box and D-box are well conserved in all known ERs, including those from mollusks, ERE binding appears to be a feature specific to all ERs.

Lamprey ER, but not amphioxus ER, is able to induce transactivation of a reporter gene in response to estradiol stimulation

The transactivation ability of lamprey ER and amphiER was then compared with that of human ERα. AmphiER failed to induce transcription of a reporter construct containing a consensus ERE in front of a minimal promoter

in transfected mammalian cells after stimulation by the natural vertebrate ER ligand, estradiol (E_2) as well as a wide variety of other vertebrate ER ligands (the natural agonist 3β-androstenediol [43], and the phytoestrogens resveratrol [44] and enterolactone [45]) (Figure 4A and 4B). In order to improve the detection sensitivity, was also tested the transactivation capacity of amphiER in response to E_2 as a construct containing only the LBD fused to the GAL4 DNA-binding domain. In this case again, no activation was detected (Figure 4C). In agreement with this result, no recruitment of the coactivator SRC1 (an homologue of which is present in the amphioxus genome, see Discussion) was detected in mammalian two-hybrid assay (Figure 4D). However lampER is activated by E_2, with an intensity comparable to humanERα (Figure 4C), which suggests that the lamprey ER is a high affinity E_2-dependant transcription factor.

Since amphiER is able to bind DNA but is unable to activate transcription of a reporter gene, the dominant negative capacity of the amphioxus protein was tested. A dose-dependent decrease in the reporter gene activity was clearly visible in 2 different cell lines when increasing amounts of the amphiER plasmid were added together

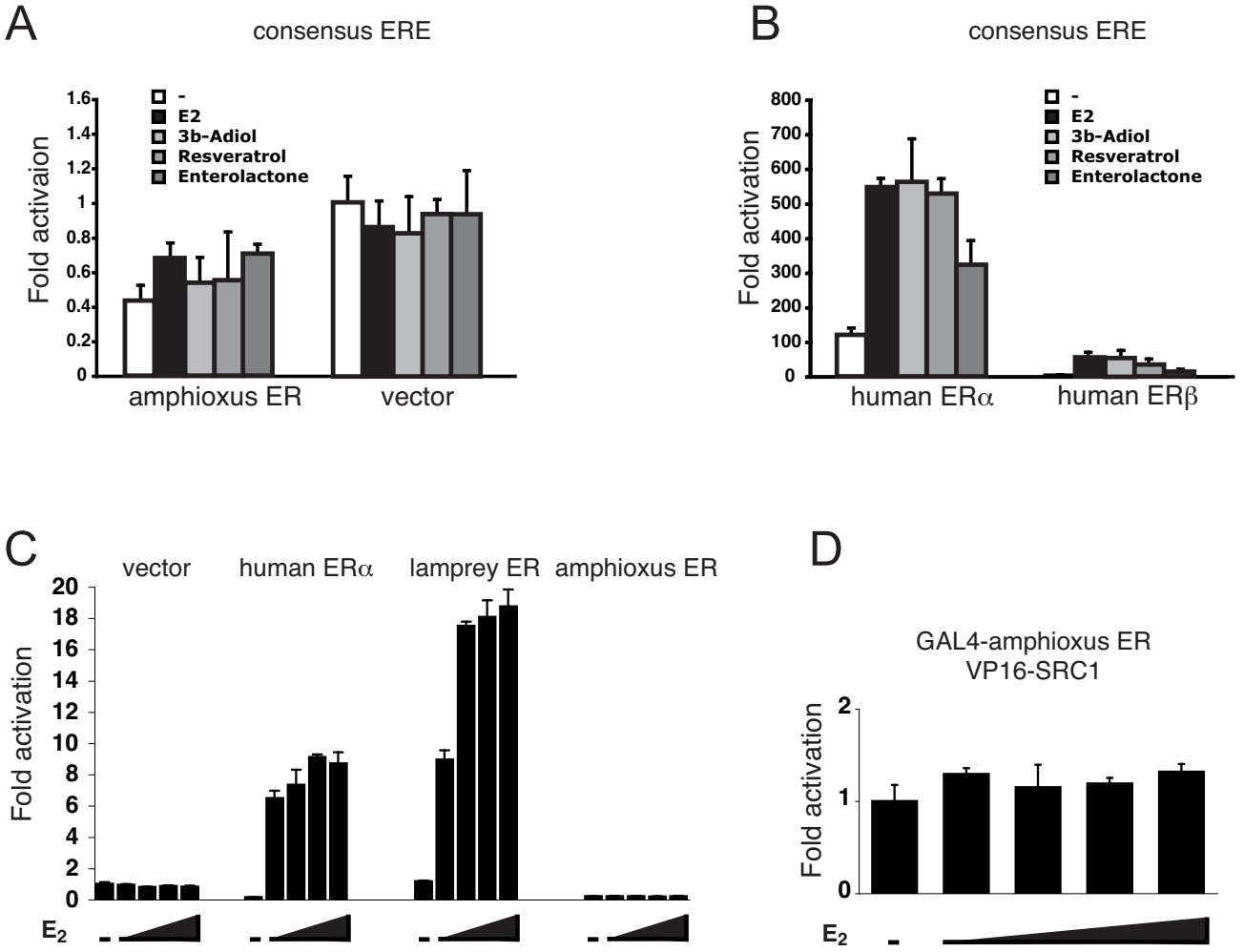


Figure 4

LampER is activated by E_2 whereas amphiER is not activated by ER agonists. The pSG5 constructs containing either amphiER (A), human ERα or human ERβ (B), were tested in transfected Cos 7 cells for their ability to activate the co-transfected cognate ERE-luc reporter plasmid after E_2 stimulation (10^{-6}M). (C) GAL4-LBD constructs from several chordate ERs were tested in transfectec 293 cells for their ability to activate a (17m)5x-G-luc reporter plasmid in the presence of increasing doses of E_2 (10^{-9}M to 10^{-6}M). (D) Mammalian two-hybrid SRC1 recruitment assay. The GAL4-amphiER-LBD chimera was used with the coactivator SRC1 fused to the strong activation domain VP16 to transfect 293 cells in the presence of increasing doses of E_2 (10^{-9}M to 10^{-6}M).

with constant amounts of human ERα or ERβ in transient transfection experiments. This decrease was observed both with synthetic consensus EREs (Figures 5A and 5B. See also Additional file 3) and with the natural ERE present in the classical ER pS2 target gene (Figure 5C). Apparently, amphiER is able to compete with human ERα or ERβ for binding to the ERE sites present in the reporter constructs, and in doing so, prevents ERα and ERβ from inducing transcription, which results in a decrease in reporter gene activity. Thus, in contrast to *Aplysia*, *Octopus* or *Thais* ER [11,12,24], amphiER does not display constitutive transcriptional activity under our experimental conditions and rather exhibits an inhibitory effect (Figure 4). This clearly shows that the absence of transcriptional activity observed here is not an artifact linked to a poor expression of the construct but rather reflects the inability of amphiER to activate transcription in mammalian cells.

LampER is an estradiol receptor whereas amphiER is not able to bind ER ligands except the synthetic compound Bisphenol A

In order to confirm that lamprey ER is an E_2 receptor and to better understand the molecular basis behind the inability of the amphiER to become transcriptionally activated by estradiol stimulation, E_2 binding by lamprey ER and amphiER was tested *in vitro*. For that purpose, limited proteolysis assay allows to assess whether addition of different putative ligands can induce a conformational change in amphiER [13]. Using this method the ligand induced conformational change of the LBD is revealed by the alteration of the receptor sensitivity toward proteolytic digestion by trypsin. As expected, E_2 was able to protect human ERα from proteolysis (Figure 6A). Interestingly lamprey ER was also protected from proteolysis by E_2 even at the lowest concentration tested (Figure 6A), thus confirming the results of the transactivation assays that the lamprey ER is a high affinity E_2 receptor. In contrast no protection of amphiER by estradiol was observed, even at very high ligand concentrations (10^{-3}M) (Figure 6A). Since estradiol does not protect amphiER from proteolysis, several other classical ER ligands were tested, such as the synthetic ER agonists diethylstilbestrol [46], 4-hydroxy-tamoxifene [46] or bisphenol A (BPA) [47], the natural agonist 3β-androstenediol [43], the phytoestrogen enterolactone [45] or the synthetic ER antagonist ICI-182780 [46]. All compounds were able to bind to human ERα (Figure 6B to 6G) as expected, whereas none but BPA was able to bind to amphiER (Figure 6G). However, BPA did not induce transactivation by amphiER in mammalian cells reporter assay, and did not induce recruitment of the coactivator SRC1 either (see Additional file 4).

In order to rule out the possibility that amphiER is activated by a compound related to E_2, a large panel of 14 other steroids and cholesterol derivatives were tested for

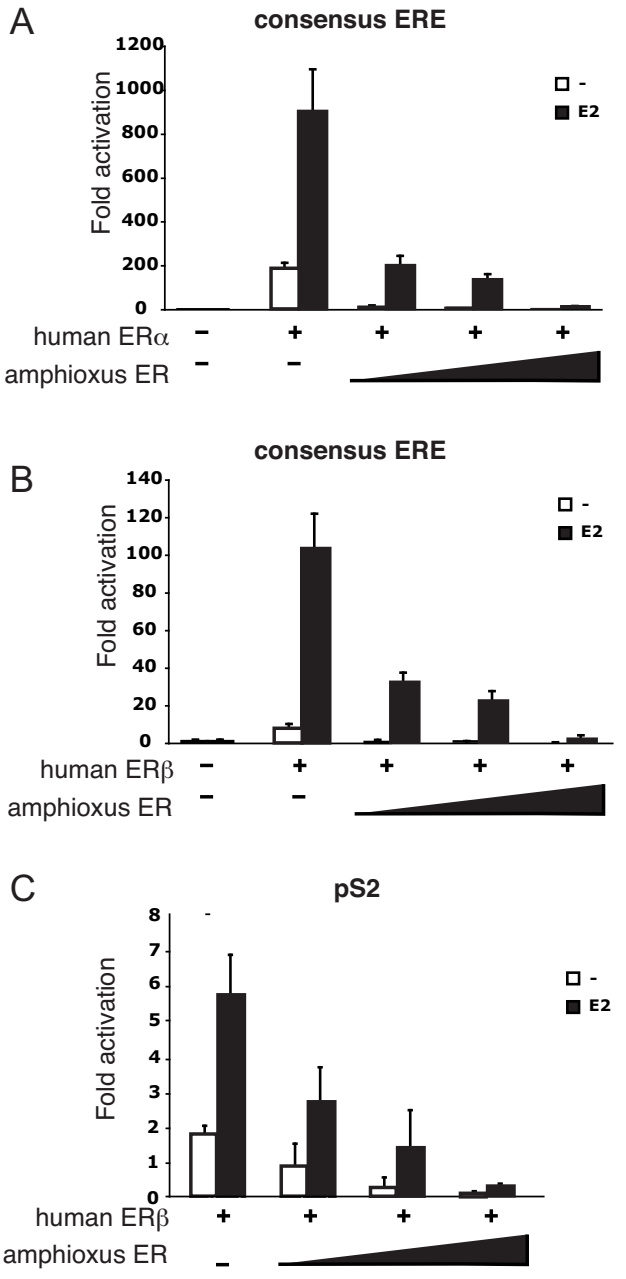


Figure 5
The amphioxus ER acts as a dominant negative estrogen receptor. A pSG5 construct containing human ERα (A) or human ERβ (B) was tested in transfected HeLa cells for its ability to activate the co-transfected cognate ERE-luc reporter plasmid after E_2 stimulation (10^{-6}M) in presence of increasing doses of the amphiER construct. (C) A pSG5 construct containing human ERβ was tested in transfected HeLa cells for its ability to activate the co-transfected ps2 promoter after E_2 stimulation (10^{-6}M) in the presence of increasing doses of the amphiER construct.

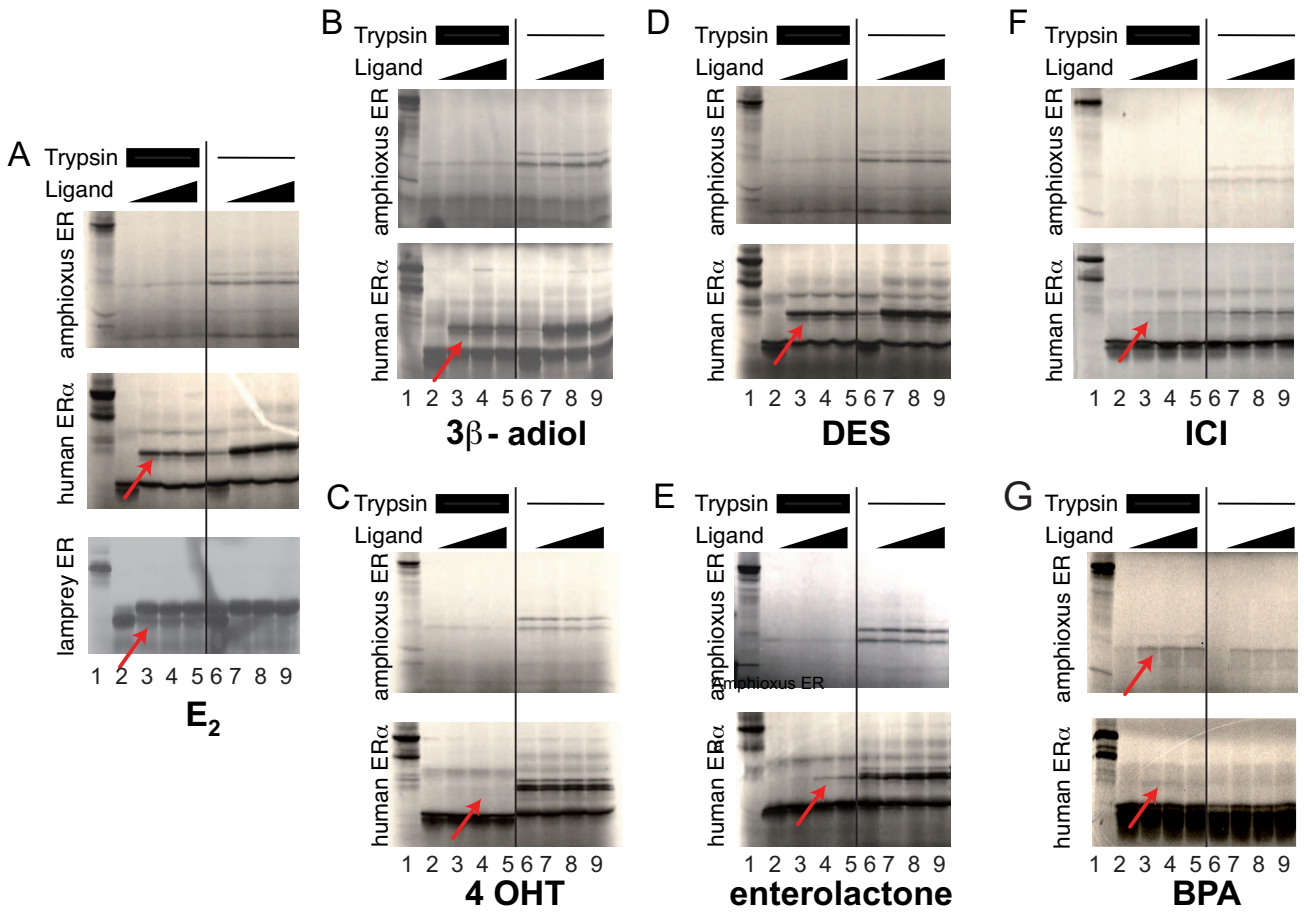


Figure 6

Limited proteolysis of lampER with E_2 and of amphiER with various ER ligands. Human ERα was used as a positive control. lane 1: undigested protein, lanes 2–5, 6–9: digested protein in the absence (lane 2 and 6) or presence (lanes 3–5 and 7–9) of ligand (10^{-3}M to 10^{-5}M). 2 different trypsine doses are shown, indicated by thick or thin bars above each panel. The ligands are (A) estradiol, (B) 3β-Androstane-diol, (C) 4-hydroxytamoxifen, (D) diethylstibestrol, (E) enterolactone, (F)ICI-182780 and (G) bisphenol A.

their ability to bind and activate amphiER. None of the tested compounds, even at high doses, had any effect on amphiER transcription activity (see Additional file 5A). Accordingly, no recruitment of the coactivator SRC1 by amphiER was detected in mammalian two hybrid assays (See Additional file 5B). The most probable explanation is thus the lack of binding by those compounds to amphiER (See Additional file 6).

Taken together these results show that the ER from lamprey behaves as a "classical" ER since it binds DNA on a classical ERE and is activated by binding E_2. On the other hand, though the single ER from amphioxus is able to bind the ERE, it does not bind any tested ER ligand and cholesterol derivative, except bisphenol A. However, no transcriptional activity was detected upon stimulation by any of the tested ligands. Since none of the mollusk ERs sequenced up to now binds E_2 either, our data suggest that E_2-binding by ER is restricted to vertebrates, implying that vertebrates specifically gained the ability to be regulated by E_2 (see Discussion).

Ancestral reconstruction of steroid receptors

In previous analyses that discussed the evolution of ERs, it was argued that estradiol binding was an ancient function of all sex steroid receptors (SRs, comprised of ERs and NR3C members) and that the binding to other steroids was more recent, with estradiol binding ability getting restricted to ERs [8,11]. Those conclusions were based on the reconstruction of the ancestral SR sequence [11]. Since the finding that estradiol binding is not shared by all ERs but restricted to vertebrate ERs contradicts this hypothesis, and to get better insight into this apparent contradiction, the sequence of the ancestral steroid receptor was reesti-

mated. When the ancestor (AncSR1) was first "resurrected", only one non-vertebrate sequence was available [8,11]. The impact of more non-vertebrate sequences (including amphiER) was thus tested on the reconstruction of the ancestor of steroid receptors. The sequence of the ancestral steroid receptor (AncSRa), at the node grouping ERs and NR3C, was inferred using PAML 4 [48], from the alignment described previously (the study was restricted to DBD and LBD) and the topology shown in Figure 2B. The predicted sequence resembles AncSR1 (Figure 1) with 12 out of 18 amino acids involved in ligand binding [49] being ER-like (Figure 1). However, important differences were noticed between AncSR1 and AncSRa. First one of the 3 amino acids making direct contacts with E_2 is different in AncSR1 and AncSRa: at this position, AncSR1 is vertebrate-ER like (a His residue is present at position 524 of humanERα, located in helix H10–H11, in green in the alignment, Figure 1) whereas the amino acid is different in AncSRa and is mollusk-ER like (Tyr instead of His) and mutations at this site have been shown to impair ERα activity in human [50,51]. Second, when a phylogenetic tree is built including both ancestral sequences AncSR1 and AncSRa as well as various NR3 sequences, AncSRa branches deep in steroid receptors as expected since it was built on the same dataset (Figure 7. For a complete tree presenting all leaves, see Additional file 7) but AncSR1 branches close to the vertebrate ERs, which is surprising. In order to determine whether the taxonomic sampling, and not details of the alignment, was responsible for these differences, we calculated a second SR ancestral sequence (AncSRb) using a smaller dataset: taxon sampling was reduced by removing most of the non-vertebrate steroid receptor sequences from the alignment (the amphioxus ER, the amphioxus NR3C as well as 5 out of the 6 mollusk ER sequences were removed to obtain a dataset closer to the one used in [8]). In this case, AncSRb branches next to AncSR1, closer to the ER clade than to AncSRa (Figure 7). This result clearly shows that the reconstruction of ancestral sequences is influenced by the set of sequences available and that restricted taxonomic sampling biases ancestral SR sequences towards vertebrate ER sequences. Therefore conclusions based on such an analysis (specifically that the ancestral ER was able to bind estradiol) should be considered as only tentative, since taxonomic sampling of available steroid sequences is very much vertebrate-centered. Overall, taken together, our data do no support the hypothesis that the ancestral steroid receptor was an estradiol receptor.

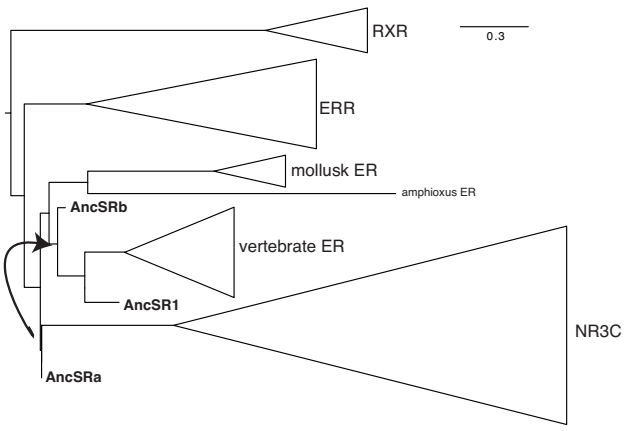


Figure 7
The reconstruction of the ancestral sequence of steroid receptors is sensitive to taxonomic sampling. The ancestral sequence of ER and NR3C was inferred using either a complete dataset (AncSRa) or a partial dataset (AncSRb) where 5 mollusk ER sequences as well as amphiER and amphiNR3C were omitted. The position of those sequences within the phylogenetic tree calculated with the complete dataset was compared. The position of a previously described ancestor (AncSR1) is indicated as well. Triangles represent the different NR clades. For the complete tree, presenting all the 80 sequences present in the tree, see Additional file 7.

Discussion

The amphioxus ER does not bind estradiol

In this paper we cloned and functionally characterized the lamprey and amphioxus orthologues of the human estrogen receptors. Our results show that lampER binds estradiol whereas amphiER does not. We propose that 3 types of ERs can be distinguished, depending on their ligand binding properties: vertebrate ERs (including lamprey) are the only *bona fide* estradiol receptors, mollusk ERs do not bind estradiol and are constitutively active transcription factors and amphiER does not bind estradiol and is transcriptionally silent in mammalian cells. This is supported by two points: (i) the experimental approach developed here is biologically relevant since the binding of bisphenol A (BPA) to amphiER was observed using the same experimental conditions as for E_2 suggesting that amphiER is correctly folded and that a ligand binding pocket is likely to be present. (ii) One of the three key amino acid positions within the LBP of amphiER (Cys 531, located between the helixes H5 and H6, in green in Figure 1) diverges from vertebrate ERs (Arg 394 in human ERα), whilst the two other key positions (Glu 490 and His 659 in amphiER, located in helix H3 and H10–H11 respectively, in green in Figure 1) are conserved with the vertebrate ERs (amino acids corresponding to Glu 353 and His 524 in human ERα), suggesting that potential contacts between amphiER and estradiol are impaired. Accordingly, a recent *in silico* study of amphiER ligand binding ability confirmed an "unusual ligand recognition in amphioxus ER" [52].

It was unexpected that no effect of the synthetic ER agonist BPA was detected in the transactivation assay of the recep-

tor in mammalian cells since BPA induces a conformational change of amphiER. This apparent absence of coactivator recruitment (see Additional file 4B) resulting in no transcriptional activity in response to BPA can be interpreted in several ways: (i) because of the different geometry of the ligand-binding pocket in amphiER, BPA behaves as an antiestrogen (partial agonist or even a partial antagonist) and blocks the transcriptional activation properties of amphiER, for instance by inducing a conformational change that does not allow coactivator recruitment (like human ERα and 4-raloxifen, [53]) or by excluding amphiER from nucleus (like ICI-182,780 with human ERα, [54]). (ii) Alternatively, the coactivator interface of amphiER does not fit with mammalian coactivators, resulting in artifactual loss of activation. However, the conservation of the amino acids involved in co-activator interaction, compared to human ERα does not support this hypothesis (sites indicated with a star in Figure 1, as described in [55]). Among the divergent sites, at a position implicated in the charge clamp necessary for coactivator contact, amphiER contains an aspartate (D677) instead of a glutamate in humanERα (E542). Importantly, the divergence (E->D) is conservative and preserves the negative charge of the amino acid, which is important for interaction with the lysine from helix 3 (conserved in amphiER) to form this charge clamp [56]. In addition, a unique orthologue of the p160 family of coactivators was found in the amphioxus genome [35] and its overall conservation with its 3 human ohnologues (genes that have been duplicated during the two rounds of whole genome duplications in the chordate lineage [57]) is good. (iii) Interaction between mammalian chaperones like HSP90 and amphiER is impaired, leading to improper binding to the hormone [58]. Taking these results into account, it will be interesting to test the effect of BPA on the subcellular localization of amphiER and to study if other related compounds are able to bind and/or activate amphiER. In addition, it will be important, when cell cultures from amphioxus are available, to check the activity of amphiER in a monospecific transient transfection assay. It should be remembered that some orphan receptors such as ERRs are thought to have no natural ligands even if they are able to bind synthetic compounds [59]. More generally, the precise status of amphiER in terms of ligand binding remains an open question. It is nevertheless clear, and this is an important issue for the current evolutionary debate, that amphiER is not able to bind estradiol.

Is there any receptor for estradiol in amphioxus?

The observation that amphiER does not bind E_2 is indeed a surprising observation since E_2 was detected in amphioxus by RIA, the hormonal production being correlated with breeding season [15]. Several aspects of steroid metabolism were described in amphioxus [60] and the homologues of many enzymes necessary for estradiol synthesis in mammals were cloned from amphioxus ovaries [15,61]. Of particular interest is the report of an aromatase gene (CYP19) in amphioxus, which suggests that the crucial step in estradiol synthesis is indeed possible in amphioxus. These experimental data were recently confirmed by the analysis of the complete amphioxus genome sequence [35]. It may be that, in amphioxus, the active sex hormone is an E_2-derivative [62] or another sex hormone, like in the case of androgens in lamprey [63], and this derivative is still to be discovered. In a similar way, we recently demonstrated that the amphioxus TR orthologue does not bind T_3 or T_4, the classical thyroid hormones, but deaminated derivatives TRIAC and TETRAC, which are able to induce amphioxus metamorphosis [64].

A second possibility is that E_2 itself has a central role in sex maturation in amphioxus, and that the functional estrogen receptor in amphioxus is different from amphiER. Several candidates are possible. First, there is another steroid receptor in amphioxus (amphiNR3C in Figures 1 and 2) [35] that exhibits several ER-like features. Its P- and D-boxes are closer to ERs than to vertebrate NR3C (Figure 2A). The sequence identity of its LBD with human ERα (37%) and with NR3C members (35%) are similar. Moreover, most of the amino acids involved in ligand binding are more ER-like than AR-, PR- or MR-like (Figure 1). However it is the only NR3C receptor (orthologous to AR, PR, MR and GR) found in the amphioxus genome. Thus if amphiNR3C plays the role of an estradiol receptor, this suggests an absence of a "classical" steroid receptor able to bind testosterone, progesterone or corticoids. Alternatively, a non-nuclear receptor could mediate E_2 action in amphioxus. Indeed, several non-genomic effects of estradiol were reported in mammals involving GPCRs (for reviews see [65-67]). For instance, very recently, a high affinity receptor for the steroid androstenedione linked to the membrane, was described in lamprey [68] and a GPCR with high affinity for progestines was isolated from sea trout [69].

Implications for the evolution of ERs

The absence of E_2 binding by the amphioxus estrogen receptor has interesting consequences for the evolution of SRs and ERs. Indeed, only the well characterized gnathostome ERs and the lamprey ER (studied here) have been shown to mediate E_2 action. Outside vertebrates, all the ERs studied so far (in mollusks and amphioxus) do not bind E_2 [11,12,24-26]. Parsimony implies that the function of estradiol in the bilaterian ancestor was not mediated by ER and that ER had another function. Only later during evolution, in the vertebrate lineage, ER would then have gained the ability to be activated by E_2 and to mediate the hormonal action of this compound (Figure 8). The alternative scenario (ancestral E_2 binding and independ-

ent loss of either ER itself or E_2 binding to ER in mollusks and invertebrate deuterostomes) is more costly in terms of evolutionary events, even if the hypothesis of an NR3C orthologue binding E_2 is taken into account. Thus, taken together, our results do not support previous scenarios of steroid receptor evolution based on a reconstruction of the ancestral steroid hormone receptor AncSR1 [8,11].

To describe the evolution of a protein, being able to study ancestral sequences at different nodes of a phylogeny would obviously provide historically relevant information that is not available otherwise [70]. However such sequences have disappeared long ago and can only be statistically estimated. The accuracy and bias of these estimations therefore need to be investigated. Indeed, functional studies of ancestral sequences are of any value only if the ancestral reconstruction is reliable enough. The confidence associated with the previously published ancestral steroid receptor is quite low. Indeed, the overall accuracy of the reconstruction of the LBD (AncSR1) was only 62% [11]. This is similar for the ancestor inferred here (AncSRa) on an enriched dataset, with an overall accuracy of the DBD+LBD of 70%. Moreover, amino acid uncertainty was high at many sites of AncSRa and AncSR1: more than 60 sites have more than 1 possible amino acid with a probability superior to 0.2. If one were to make an exhaustive study, one would need to reconstruct and test more than 10^{24} potential proteins (if all possible combinations of amino acids with probability > 0.2 were tested). In fact several of the sites involved in ligand binding have low probabilities. Examples of more reliable reconstructions of nuclear receptors have been published, *e.g.* the ancestor of MR and GR (mineralocortoid and glucocorticoid receptors) in which the overall accuracy of the LBD was above 99%, with no disrupting mutation at any site [71]. The reconstructed ancestor of RARs also showed a high average confidence (99% [13]). This discrepancy between results obtained on the ancestor of all steroid receptors or merely of MR and GR for instance, can be explained by the higher sequence divergence observed among all SRs than simply among subfamilies MR and GR (see branch length in Figure 2B). Consequently, the uncertainty associated to the sequence of the ancestral steroid receptor as estimated with nowadays methods is probably too high to provide a firm basis for evolutionary conclusions. Moreover, the phylogenetic reconstruction of ancestral sequences has been shown to be biased towards the most frequent (and more stable) amino acids, resulting in an under-estimation of the less frequent amino acids (the stability of the ancestral protein is then over-estimated [72]). In reconstructions of ancient proteins, where the evolutionary signal has been lost due to a high number of substitutions, such biases might be problematic. Thus, current reconstruction methods do not seem powerful enough to infer a biologically meaningful ancestral steroid receptor given the amount of divergence between sequences.

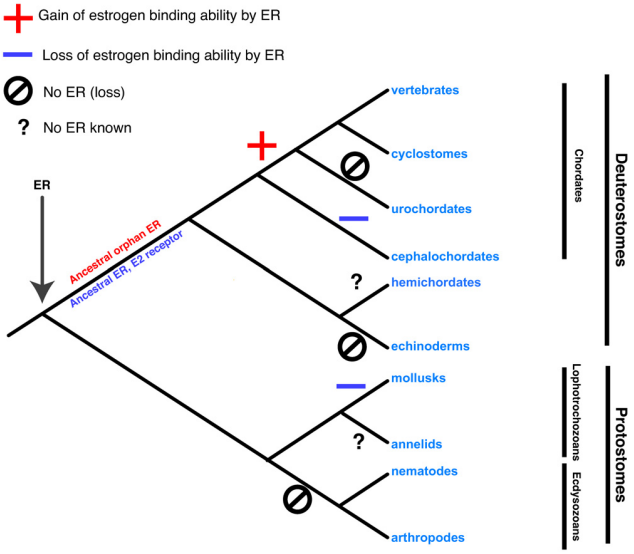


Figure 8
Model of evolution of the ligand binding ability of ERs. On a classical phylogenetic tree of bilaterians, data available on the binding ability of all known ERs have been indicated. Two hypotheses are compared in terms of parsimony, whether the ancestral ER was liganded (in blue) or not (*ie* an orphan receptor) (in red). This result displays different costs in terms of parsimony: one unique event of gain specifically in vertebrate for the "ancestral ER orphan" hypothesis against at least two parallel events of loss of binding for the "ancestral ER binding E_2" hypothesis. In addition, three events of loss of the ER gene in urochordates, echinoderms and insects+nematodes are implied by the current distribution of the gene across metazoans.

Nonetheless, all these reservations put aside, it is surprising that the previously reconstructed ancestral SR, is vertebrate ER-like. As almost all the extant sequences used as matrix for the reconstruction came from vertebrates and led to the estimation of a "vertebrate-like" ancestral sequence, the same ancestral steroid receptor as previously published [11] was estimated, but adding more sequences from various taxa. This reconstruction was done using a phylogeny equivalent to the one previously published [11]. Using this approach, the new AncSRa is more divergent from vertebrate ERs than AncSR1. Interestingly AncSR1 was shown to bind E_2 with a very low affinity (250 times lower affinity than human ERα [11]), suggesting that AncSRa may be an even worse estradiol receptor.

The bias of AncSR1 towards vertebrate ERs is explained by a lack of non-vertebrate sequences used for the reconstruction. Indeed, removing some non-vertebrate sequences

from our dataset leads to an estimation of an ancestral steroid receptor that is more "vertebrate ER"-like (AncSRb in Figure 7). The clustering of AncSRb with the ER clade and the exclusion of AncSRa from the ER clade were supported with good statistical values (minimum of Chi2-based and SH-like supports of 0.83 and 0.89, respectively). Those data show that AncSR1 reconstruction was probably sensitive to the vertebrate bias in the data set (Figure 7). Overall, we suggest that analysis based on ancestral reconstructions should be taken as tentative, especially in case of low statistical confidence and limited taxonomic sampling. In case of the ancestral steroid receptor, even if exhaustive taxonomic sampling is necessary, phylogenetic signal is weak and the resulting confidence is quite low. Thus we think that even if the ancestral sequence built here is biologically more relevant than previously calculated ones (because of better taxonomic sampling), it remains quite uncertain. Consequently conclusions regarding the ancestral steroid receptor should be based mostly on comparative characterization of extant receptors. In that case, all the data based on invertebrate ER receptors (from mollusks and amphioxus) support an ancestor of steroid receptors that was not able to bind estradiol. This conclusion will obviously require the functional characterization of ERs from other protostome phyla in order to carefully check if this observation is general. Thus, available data converge towards a re-evaluation of the ancestral status of estrogen receptors.

Sequence conservation reflects functional constraints: ligand binding ability is more recent than DNA-binding ability

From our and previous studies, only vertebrate ERs are able to bind and activate transcription under estradiol stimulation [11,12,24-26]. The LBD of amphiER is more divergent from its vertebrate counterparts (ca. 34% amino acid identities) than the LBD of other liganded amphioxus nuclear receptors such as amphiRAR (ca. 58%), which has been shown to bind the same ligand as its vertebrate homologue [7,13]. This suggests that a conserved functional feature (e.g. binding to the same ligand) is reflected in the sequence conservation of the LBD.

The same observation can be done concerning the DBD since all ERs, including amphiER, have a highly conserved DBD and are able to bind EREs. Thus, for this domain also, a conservation of the function is reflected in sequence conservation.

Accordingly with this notion, the LBD of invertebrate ERs is highly divergent but their DNA binding domain, as well as other functionally important domains not directly linked to ligand binding such as the dimerization interface, or the amino acids responsible for interaction with the co-activators [55] are well conserved. This is true for amphiER as well as mollusk ERs. This strongly suggests that amphiER is a *bona fide* NR regulating ERE-containing genes in an E_2-independent manner. Post-translational modifications such as phosphorylation or the presence/absence of other receptor-interacting proteins such as transcriptional coactivators have been shown to regulate unliganded nuclear receptors [73]. Whether one of these mechanisms acts to regulate the activity of invertebrate ERs or if those receptors have unknown ligands still to be identified remains to be explored. Anyway our observations strongly suggest that for ERs, the DNA binding function of the receptor as well as its interaction with co-regulators have been conserved due to selective pressure. Interestingly, when studying the AncSRa, the P- and D-boxes in the DBD are ER/ERR-like (Figure 1), suggesting that ER/ERR DNA binding ability is ancestral, in accordance with the fact that these are the only receptors of the NR3 family found in invertebrates. This difference in the selection pressure between DBD and LBD has been proposed to be a general evolutionary pattern for the whole NR family [7]. The plasticity of the ligand binding ability of NRs was recently illustrated in the case of RXR-USP where the ability of the receptor to be regulated by a ligand was suggested to have been subject to several successive episodes of gain and loss during evolution [74].

Evolution of endocrine systems: refinement of the ligand exploitation model

The ligand exploitation model hypothesizes how new hormones and new receptors appear during evolution. It suggests that the ancestral ligand is the last metabolite of a synthesis pathway [8]. According to this model, the ancestral steroid ligand was estradiol (and the ancestral SR bound estradiol). During evolution, other steroid receptors appeared by duplication of the ancestral ER and gained the ability to bind other steroids, intermediate in the synthesis pathway (like testosterone or progesterone).

Our findings on the evolution of ERs do not support the ligand exploitation model, since our data strongly suggest that the ancestral ER did not bind estradiol. However, as estradiol has been detected in deuterostomes as well as protostomes (for instance in vertebrates, amphioxus, echinoderms, mollusks, for review, see [75]), steroid signaling may have been already present in bilaterian ancestor. However, up to now, the ancestral steroid molecule remains to be determined. If estradiol is an ancient hormone, it then probably bound another receptor and later on ER gained the ability to recognize it, as did other steroid receptors for their extant ligand. Thus the evolution of steroid system intermingles two distinct processes, the evolution of the receptor on one hand, and the evolution of the ligand on the other.

The receptor can evolve by point mutations and change its affinity for a ligand towards another. This idea was convincingly exemplified in the case of corticoid receptors

(the ancestor of MR and GR) for which it was recently demonstrated that ability of the ancestral vertebrate corticoid receptor to bind gnathostome-specific hormone aldosterone (a MR ligand) was a by-product of its ability to bind the ancestral ligand 11-deoxycorticosterone (DOC) [71]. GR gained the ability to bind cortisol only in the gnathostome lineage, in parallel to endogenous synthesis of the hormone [16,71]. This detailed study shows that a receptor binding a given ligand can acquire affinity for compounds present in the cell that are structurally close to its natural ligand: this refines the ligand exploitation model, since new ligands are not necessarily precursors of ancient ligands, simply compounds present in the cell and structurally close to the ancestral ligand. Similar conclusions were drawn previously in the case of RAR evolution [13]. It has to be emphasized that the pool of available compounds is also subject to evolutionary changes in parallel. For instance, the spatiotemporal production of estradiol is variable in the different vertebrate groups (reviewed in [76]). Glucocorticoids differ in mouse (cortisol) and in human (corticosterone), with both hormones being GR ligands [77]. There are several androgens in teleost fishes, with 11-ketotestosterone being teleost-specific [78]. As there are 2 androgen receptors (ARs) in teleost fishes, from a whole genome duplication [79], the study of the ligand-binding ability of those ARs is a potentially interesting case for the evolution of endocrine systems. As highlighted by Bridgham et al. (2006), lamprey does not produce cortisol [71]. In accordance, their genomes do not contain the sequence corresponding to the enzyme responsible for cortisol synthesis (11b-hydroxylase) and in general classical steroids except estradiol are rarely found in lamprey. This suggests that the steroids actually found in lamprey are different from the ones found in mammals (reviewed in [80]). Those cases exemplify the largely underestimated diversity of endocrine systems: except for lamprey and some teleost fishes, the hormonal pool of animals remains largely unknown. As proposed for the study of steroid receptors, a comparative approach should be applied to determine the metabolism of steroids in poorly studied animals. Indeed, the hormonal pool of such animals is usually evaluated from the presence/absence of putative orthologues of mammalian enzymes. As the enzymatic machinery involved in hormonal metabolism has a very labile activity (reviewed in [81,82]), equating orthology with functional identity might be unreliable.

The evolution of steroid receptors can be replaced in the more general context of ligand-nuclear receptor co-evolution. The evolution of the NR1H subfamily, that includes receptors for other steroidal compounds, like the major transcriptional regulator of bile salt synthesis farnesoid × receptor (FXR), the pregnane × receptor (PXR), the vitamin D receptor (VDR) or liver × receptor (LXR)/ecdysone receptor (Ecr), has been extensively studied and is not in line with the ligand exploitation model [83-85]. For instance, comparative functional studies of FXRs from various chordate species showed that the vertebrate FXRs bind "late" cholesterol derivatives (from a complex synthesis pathway) but are thought to have evolved from an ancestral FXR that bound early cholesterol derivatives (from a simpler synthesis pathway) [83].

In other cases, the evolution of ligand binding is more "chaotic" with close orthologs having a selective ligand binding ability that varies extensively (vertebrate VDRs are very well conserved when PXRs have the widest ligand repertoire of all NRs) [83].

These complex histories are probably linked to specific function of some of those NRs, considered as xenotoxic compounds "sensors". This tight relationship with the unstable environment probably makes receptors like FXR and especially PXR more prone to fast evolution [86]. Yet they illustrate the impressive variety of scenarios of NR evolution.

Conclusion

In this article, we demonstrated that vertebrate ERs (including lamprey ER) are estradiol receptors whilst non-vertebrate ER (including amphioxus ER) are not. The most parsimonous scenario proposes that the ancestral ER was not able to bind estradiol and that it had another function. It later gained the ability to be regulated by estradiol, specifically in the vertebrate lineage. However, additional critical data remains to be discovered in poorly studied taxa [62]. To fully understand the evolution of steroid signaling pathway, a larger number of taxa need to be targeted for detailed comparative studies. More precisely, ERs and other steroid receptors should be cloned from widely distributed taxa, especially in protostomes. Enzymes involved in steroidogenesis should also be cloned and characterized, to understand the evolution of steroid availability. In order to avoid the blinders of a "vertebrate-centered" view, it is of particular importance to establish the steroid hormone repertoire of an enlarged animal panel, including more protostomes. The description of various endocrine systems will certainly be relevant to the early evolution of hormone signaling.

Methods

Cloning of amphiER

An initial piece of amphiER was obtained by degenerate PCR on different RT reactions from total RNA extracted either from developing B. floridae embryos and larvae (at 13 h–15 h, 28 h, 36 h, 48 h or 3 d–4 d of development) or from B. floridae adults. The oligonucleotides used were

as follows: forward primer 5'-TGYGARGGITGYAARGCIT-TYTT-3' and reverse primer 5'-GTRCAYTSRTTIGTIGCIG-GRCA-3'.

The touchdown PCR program used was as follows:

5' 94 degrees

5× (30" 94 degrees, 1' 55 degrees, 1' 72 degrees)

5× the same cycle, but at 50 degrees annealing temperature

5× the same cycle, but at 45 degrees annealing temperature

5× the same cycle, but at 40 degrees annealing temperature

25× the same cycle, but at 37 degrees annealing temperature

7' 72 degrees

All degenerate PCRs irrespective of the RT reaction template used yielded a 83 bp fragment of amphiER. The fragment was sequenced on both strands and used for the design of oligonucleotides for 5' and 3' RACE experiments with the Invitrogen GeneRacer Kit. The template for the RACE experiments was pooled total RNA from 13 h–15 h B. floridae embryos and from B. floridae adults. In addition to the oligonucleotides provided by the kit, for the 3' RACE, the following primers were used:

3' RACE, 1st PCR: 5'-AACGGAGCATTCAGCAAGGTC-3'

3' RACE, 2nd PCR: 5'-GCATTCAGCAAGGTCAGACAG-3'

5' RACE, 1st strand cDNA synthesis: 5'-ATGTAATCT-GTCTGACCTTGC-3'

5' RACE, 1st PCR: 5'-CTGTCTGACCTTGCTGAATGC-3'

5' RACE, 2nd PCR: 5'-TCTGACCTTGCTGAATGCTCC-3'

The protocols for the 1st and 2nd round of PCR experiments are given in the Invitrogen GeneRacer Kit. The 3' and 5' RACE products were subsequently sequenced on both strands and used for the design of oligonucleotides for the full-length cloning of amphiER: forward primer 5'-CGGCGAAGCGAAGAAGATCGAG-3' and reverse primer 5'-CTTAACCGATACTAACGGAACAG-3'. The full-length amphiER was obtained by PCR on pooled RT reactions from total RNA extracted from B. floridae 13 h–15 h embryos, 3 d–4 d larvae and B. floridae adults. The PCR protocol used was as follows:

10' 94 degrees

5× (30" 94 degrees, 30" 55 degrees, 2' 72 degrees)

35× the same cycle, but at 50 degrees annealing temperature

10' 72 degrees

The full-length amphiER clone resulting from this PCR is 2279 bp long, was cloned into the pCR2.1 vector (Invitrogen) and subsequently sequenced on both strands.

Plasmid constructs and reagents

Full length amphiER were amplified by polymerase chain reaction (PCR) and the obtained fragments were inserted into a pSG5 vector between EcoR1 sites. Lamprey ER was a generous gift from JW Thornton. Human pSG5-ERα and pSG5-ERβ and the 3xERE-Luc luciferase reporter construct have been described previously [87]. The pS2-Luc reporter construct encompasses an 1100 bp estrogen-responsive region of the human pS2 promoter inserted into the pGL3 basic vector (Promega). Chimeras comprising the GAL4 DNA-binding domain fused with the LBD of the human ERα (residues 251 to 595), the LBD of amphiER (residues 364 to 705), the LBD of lampER (residues 234 to 554) have been cloned in the pG4MpolyII vector. 17β-estradiol, genistein, 3β-androstenediol, resveratrol, cholesterol, cholic acid, chenodeoxycholic acid, 22®-hydroxycholesterol, 20-Hydroxyecdysone, pregnenolone, trans-Dehydroandrosterone (DHEA), corticosterone, progesterone, 4-androstene-3,17-dione, estrone, testosterone, 5α-androstan-17β-ol-3-one and 1a,25-Dihydroxyvitamin D3 (calcitriol) were purchased from Sigma. Enterolactone was a generous gift from Dr Sari Mäkelä [88].

Phylogenetic analysis of NR3

Protein sequences of NR3 family members were obtained from GenBank by BLAST search using *Homo sapiens* ERα as a query. Eight additional sequences from the closely related RXR group were also obtained to serve as outgroup sequences. For accession numbers of the sequences used, see Additional file 8.

The retrieved sequences were aligned using the muscle 3.6 program [89] and the resulting alignment was manually corrected with SEAVIEW [90]. Phylogenetic tree was calculated by maximum likelihood as implemented in PhyML version 2.4.3 under a JTT substitution matrix plus a eight-category gamma rate correction (α estimated) and with the proportion of invariant sites estimated. Both the DBD

and the LBD were used. Robustness was assessed by bootstrap analysis (1,000 repetitions) [91].

The Bayesian inference was done using the program MrBayes 3.1.2 [92]. Two simultaneous independent runs were performed. For each run, one chain was sampled every 100 generations for 1,000,000 generations after the burn-in cycles, until the average SD of split frequencies was <0.01; additionally, the potential scale reduction factors of the parameters were close to or equal to 1, which indicates that the runs had most probably converged. The neighbour-joining (Poisson correction) and maximum parsimony trees were done with Phylo_win [90].

Likelihood-based tests of alternative topologies placing amphiER at all possible positions in the tree

The 149 trees were built by reconnecting amphiER from the maximum likelihood tree, into the 149 possible positions. The branch length and the different parameters of the obtained trees were re-estimated using PhyML. Likelihood-based tests of the 149 alternative topologies were calculated using CONSEL: site-wise log-likelihood values, available as output of PhyML, were used to calculate the P-values of the different positions according to the AU test with the software R.

Ancestral sequence reconstruction

The aminoacid sequence of the ancestral AncSRa and AncSRb was inferred only for the most conserved part of the alignment, *i.e.* the DBD and LBD (defined as in Figure 1). The ancestral sequences were reconstructed by maximum likelihood as implemented in PAML [48], under the JTT substitution model and a gamma distribution with 8 categories of rates across sites, using the tree described in Additional file 1A for AncSRa and a the same topology truncated of the mollusk ER sequences and the amphioxus NR3 sequences, after reestimation of the branch lengths using phyml (JTT+γ) for AncSRb.

Electrophoretic Mobility Shift Assay (EMSA)

EMSAs were performed as previously described [93]. Where indicated, a 10- and 100-fold molar excess of 30-bp unlabeled oligonucleotides (a consensus ERE and a non-related probe) were added as competitors. The sequence of the probe containing the consensus ERE is 5'-CGGGCCGAGGTCACAGTGACCTCGGCCCGT-3' and the sequence of the non-related probe is 5'-CTAGTC-CTAGGTCTAGAGAATTCA-3'.

Cell culture and transfections

Human embryonic kidney 293 cell culture and transfections using Lipofectamine Plus reagent (Invitrogen) were done according to the manufacturer's recommendations and as previously described [94]. Briefly, 200 ng of the chimeras comprising the GAL4 DNA-binding domain fused with the LBD of either human ERα, lampER or AmphiER (or LBD for the control) were co-transfected together with 100 ng of reporter plasmid and 10 ng of a β-galactosidase expression vector, included as a control for transfection efficiency. For the mammalian double hybrid assays, the GAL4-amphiER-LBD chimera was transfected with 200 ng of the coactivator SRC1 fused to the strong activation domain VP16. Three to five hours post-transfection, serum and hormones (as indicated in the figures) were added to the cells which were incubated for an additional 48 hours before harvest and luciferase and β-galactosidase activities were determined. Results show the mean ± s.e.m. (n = 3) of representative experiments. Human HeLa cervical cancer cells and CV-1 green monkey kidney cells were routinely maintained in Dulbecco's modified eagle's medium (Invitrogen), supplemented with 10% fetal bovine seum, 1% v/v L-glutamine and 1% v/v penicillin/streptomycin. Cells were seeded in 12 or 24 well plates one day prior to transfection. Transient transfections were carried out using the Lipofectamine Plus reagent according to instructions of the manufacturer (Invitrogen) in culture media devoid of serum, phenol-red and antibiotics. Briefly, 1 ng of ERα, ERβ or AmphiER expression vectors were co-transfected together with 100 ng 3xERE-Luc (or 200 ng pS2-Luc where indicated) and 20 ng of a β-galactosidase expression vector, included as a control for transfection efficiency. In the co-expression experiments, AmphiER was co-transfected together with ERα or ERβ in ratios of 0.5:1, 1:1 and 1:5, respectively. Three hours post-transfection, serum and hormones (as indicated in the figures) were added to the cells which were thereafter incubated for an additional 48 hours before harvest and luciferase and β-galactosidase activities were determined. Figures represent results from at least three independent experiments performed in duplicates. Data is presented as mean +/- SD of fold induction of relative luciferase values corrected against β-galactosidase activity, where activity obtained from transfected reporter plasmid alone and treated with vehicle, was arbitrarily set to 1.

Limited proteolytic digestion

These assays were done as previously described [13].

Authors' contributions

MP, HE, MS, SB and VL contributed to the conception and design of the study. MS cloned amphiER. MP performed the EMSA, limited proteolysis experiments, part of the transactivaton assays and the bioinformatics study. KP and IP performed the rest of the transactivation assays. MP and VL wrote the manuscript.

Additional material

Additional file 1

Phylogenetic analysis of NR3 sequences using several methods. *Phylogenetical trees of an alignment comprising 69 NR3 sequences as well as RXR sequences were inferred using the maximum likelihood method (ML) (A), Bayesian analysis (B), neighbour-joining method (C) and maximum parsimony method (MP) (D) based on an elision alignment of the DBD and LBD of 77 NR3 and RXRs (accession numbers are given in Additional file 8). Labels above each branch show percentages of bootstrap values after 1000 replicates (A), posterior probabilities (B), percentages of bootstrap values after 500 replicates (C) or 100 replicates (D). The fastest evolving sites (with an evolutionary rape above 2, as indicated in the Figure 3A) were removed from the alignment before computing phylogeny by maximum parsimony, to preserve the branching of mollusk ERs within the ER clade. In (A) nodes with bootstrap values below 50% are presented as polytomies, as in the Figure 2B.*
Click here for file
[http://www.biomedcentral.com/content/supplementary/1471-2148-8-219-S1.pdf]

Additional file 2

DNA binding characterization of chordate ERs. *Various chordate members of the NR3 family, namely human ERα, human ERβ, mouse ERRα, amphiER and lamprey ER, were synthesized* in vitro *and allowed to bind to a ^{32}P-labeled consensus ERE probe in an EMSA. Lane 1, empty vector (pSG5) reticulocytes lysates. Lanes 2–5, human ERα. Lanes 6–9, human ERβ. Lanes 10–13, mouse ERRα. Lanes 14–17, amphiER. Lanes 18–21, lamprey ER. Lanes 3–5, 7–9, 11–13, 15–17, 19–21, unlabeled non-specific oligonucleotide (NS) or ERE were added at indicated molar excess as competitors to test the specificity of the binding. The arrows indicated the gel shift induced by amphiER binding the ERE probe. The asterisk indicates free ERE probe.*
Click here for file
[http://www.biomedcentral.com/content/supplementary/1471-2148-8-219-S2.pdf]

Additional file 3

The amphioxus ER acts as a dominant negative estrogen receptor in CV1 cells. *A pSG5 construct containing human ERα (A) or human ERβ (B) was tested in transfected CV1 cells for its ability to activate the co-transfected cognate ERE-luc reporter plasmid after E_2, genistein or β-Androstane-diol stimulation (10^{-6}M) in presence of increasing doses of the amphiER construct.*
Click here for file
[http://www.biomedcentral.com/content/supplementary/1471-2148-8-219-S3.pdf]

Additional file 4

The amphioxus ER is not activated by BPA. *(A) GAL4-LBD constructs from several chordate ERs were tested in transfected 293 cells for their ability to activate a (17 m)5x-G-luc reporter plasmid in the presence of increasing doses of BPA (10^{-9}M to 10^{-6}M). (B) Representation of the mammalian two-hybrid SRC1 recruitment assay. The GAL4-amphiER-LBD chimera was used with the coactivator SRC1 fused to the strong activation domain VP16 to transfect 293 cells in the presence of increasing doses of BPA (10^{-9}M to 10^{-6}M).*
Click here for file
[http://www.biomedcentral.com/content/supplementary/1471-2148-8-219-S4.pdf]

Additional file 5

amphiER is not activated by cholesterol derivatives. *(A) The GAL4-amphiER-LBD chimera was tested in transfected 293 cells for its ability to activate a (17 m)5x-G-luc reporter plasmid in the presence of various cholesterol derivatives at a high concentration (1 μM) (black). The empty vector (white) was used as a negative control and the GAL4-humanERα-LBD in the presence of E_2 was used as a positive control (B) Representation of the mammalian two-hybrid SRC1 recruitment assay. The GAL4-amphiER-LBD chimera was used with the coactivator SRC1 fused to the strong activation domain VP16 to transfect 293 cells in the presence of various cholesterol derivatives at 1 μM. The empty vector (white) was used as a negative control.*
Click here for file
[http://www.biomedcentral.com/content/supplementary/1471-2148-8-219-S5.pdf]

Additional file 6

Limited proteolysis of amphiER with various cholesterol derivatives. *lane 1: undigested protein, lanes 2–4, 5–7: digested protein in the absence (lane 2 and 5) or presence (lanes 3–4 and 6–7) of ligand (10^{-3}M and 10^{-4}M). 2 different trypsine doses are shown, indicated by thick or thin bars above each panel. The ligands are cholic acid (A), Chenodeoxycholic acid (B), 22R-OH-cholesterol (C), cholesterol (D), 4-androstene-3,17-dione (E), DHEA (F), corticosterone (G), progesterone (H), pregnenolone (I), estrone (J), testosterone (K), 5α-androstane-dione (L), 20-hydroxyecdysone (M) and calcitriol (N).*
Click here for file
[http://www.biomedcentral.com/content/supplementary/1471-2148-8-219-S6.pdf]

Additional file 7

Phylogenetic tree of NR3 sequences as well as ancestral sequences. *Complete tree corresponding to the simplified one presented in the figure 7. The ancestral sequence of ER and NR3C was inferred using either a complete dataset (AncSRa) or a partial dataset (AncSRb) where 5 mollusk ER sequences as well as amphiER and amphiNR3C were omitted. The position of those sequences within the phylogenetic tree calculated with the complete dataset was compared. The position of a previously described ancestor (AncSR1) is indicated as well. Minimum of Chi2-based and SH-like supports are shown for each branch.*
Click here for file
[http://www.biomedcentral.com/content/supplementary/1471-2148-8-219-S7.pdf]

Additional file 8

Accession number of sequences used for phylogenetic analyses. *AR: androgen receptor; ER: estrogen receptor; ERR: estrogen related receptor; GR: glucocorticoid receptor; MR: mineralocorticoid receptor; PR: progesterone receptor; RXR: retinoid × receptor.*
Click here for file
[http://www.biomedcentral.com/content/supplementary/1471-2148-8-219-S8.pdf]

Acknowledgements

We thank Bastien Boussau for help with phylogenetic analysis and for comments on the manuscript. We are grateful to Joseph Thornton for the gift of the lamprey ER clone, Gerard Benoit and Gabriel Markov for critical reading of the manuscript. This work was supported by the EU funded CASCADE network of excellence, CNRS, UCB Lyon 1, ENS de Lyon and MENRT.

References

1. Gronemeyer H, Gustafsson JA, Laudet V: **Principles for modulation of the nuclear receptor superfamily.** *Nat Rev Drug Discov* 2004, **3:**950-964.
2. Laudet V, Gronemeyer H: **The Nuclear Receptor FacsBook.** London: Academic Press;; 2005.
3. Germain P, Staels B, Dacquet C, Spedding M, Laudet V: **Overview of nomenclature of nuclear receptors.** *Pharmacological reviews* 2006, **58:**685-704.
4. Benoit G, Cooney A, Giguere V, Ingraham H, Lazar M, Muscat G, Perlmann T, Renaud JP, Schwabe J, Sladek F, Tsai MJ, Laudet V: **International Union of Pharmacology. LXVI. Orphan nuclear receptors.** *Pharmacological reviews* 2006, **58:**798-836.
5. Dhe-Paganon S, Duda K, Iwamoto M, Chi YI, Shoelson SE: **Crystal structure of the HNF4 alpha ligand binding domain in complex with endogenous fatty acid ligand.** *The Journal of biological chemistry* 2002, **277:**37973-37976.
6. Wisely GB, Miller AB, Davis RG, Thornquest AD Jr, Johnson R, Spitzer T, Sefler A, Shearer B, Moore JT, Miller AB, Willson TM, Williams SP: **Hepatocyte nuclear factor 4 is a transcription factor that constitutively binds fatty acids.** *Structure* 2002, **10:**1225-1234.
7. Escriva H, Delaunay F, Laudet V: **Ligand binding and nuclear receptor evolution.** *Bioessays* 2000, **22:**717-727.
8. Thornton JW: **Evolution of vertebrate steroid receptors from an ancestral estrogen receptor by ligand exploitation and serial genome expansions.** *Proceedings of the National Academy of Sciences of the United States of America* 2001, **98:**5671-5676.
9. Escriva H, Bertrand S, Laudet V: **The evolution of the nuclear receptor superfamily.** *Essays in biochemistry* 2004, **40:**11-26.
10. Krylova IN, Sablin EP, Moore J, Xu RX, Waitt GM, MacKay JA, Juzumiene D, Bynum JM, Madauss K, Montana V, Lebedeva L, Suzawa M, Williams JD, Williams SP, Guy RK, Thornton JW, Fletterick RJ, Willson TM, Ingraham HA: **Structural analyses reveal phosphatidyl inositols as ligands for the NR5 orphan receptors SF-1 and LRH-1.** *Cell* 2005, **120:**343-355.
11. Thornton JW, Need E, Crews D: **Resurrecting the ancestral steroid receptor: ancient origin of estrogen signaling.** *Science* 2003, **301(5640):**1714-1717.
12. Keay J, Bridgham JT, Thornton JW: **The *Octopus vulgaris* estrogen receptor is a constitutive transcriptional activator: evolutionary and functional implications.** *Endocrinology* 2006, **147:**3861-3869.
13. Escriva H, Bertrand S, Germain P, Robinson-Rechavi M, Umbhauer M, Cartry J, Duffraisse M, Holland L, Gronemeyer H, Laudet V: **Neofunctionalization in vertebrates: the example of retinoic acid receptors.** *PLoS genetics* 2006, **2:**e102.
14. Baker ME: **Evolution of adrenal and sex steroid action in vertebrates: a ligand-based mechanism for complexity.** *Bioessays* 2003, **25:**396-400.
15. Mizuta T, Kubokawa K: **Presence of sex steroids and cytochrome P450 (CYP) genes in amphioxus.** *Endocrinology* 2007.
16. Ortlund EA, Bridgham JT, Redinbo MR, Thornton JW: **Crystal structure of an ancient protein: evolution by conformational epistasis.** *Science* 2007, **317(5844):**1544-1548.
17. Fox JE, Bridgham JT, Bovee TF, Thornton JW: **An evolvable oestrogen receptor activity sensor: development of a modular system for integrating multiple genes into the yeast genome.** *Yeast (Chichester, England)* 2007, **24:**379-390.
18. Escriva H, Safi R, Hanni C, Langlois MC, Saumitou-Laprade P, Stehelin D, Capron A, Pierce R, Laudet V: **Ligand binding was acquired during evolution of nuclear receptors.** *Proceedings of the National Academy of Sciences of the United States of America* 1997, **94:**6803-6808.
19. Laudet V: **Evolution of the nuclear receptor superfamily: early diversification from an ancestral orphan receptor.** *Journal of molecular endocrinology* 1997, **19:**207-226.
20. Baker ME, Chandsawangbhuwana C: **Analysis of 3D models of octopus estrogen receptor with estradiol: evidence for steric clashes that prevent estrogen binding.** *Biochem Biophys Res Commun* 2007, **361:**782-788.
21. Baker ME: **Amphioxus, a primitive chordate, is on steroids: evidence for sex steroids and steroidogenic enzymes.** *Endocrinology* 2007, **148:**3551-3553.
22. Baker ME: **Recent insights into the origins of adrenal and sex steroid receptors.** *Journal of molecular endocrinology* 2002, **28:**149-152.
23. committee NRN: **A unified nomenclature system for the nuclear receptor superfamily.** *Cell* 1999, **97:**161-163.
24. Kajiwara M, Kuraku S, Kurokawa T, Kato K, Toda S, Hirose H, Takahashi S, Shibata Y, Iguchi T, Matsumoto T, Miyata T, Miura T, Takahashi Y: **Tissue preferential expression of estrogen receptor gene in the marine snail, Thais clavigera.** *Gen Comp Endocrinol* 2006, **148:**315-326.
25. Matsumoto T, Nakamura AM, Mori K, Akiyama I, Hirose H, Takahashi Y: **Oyster estrogen receptor: cDNA cloning and immunolocalization.** *Gen Comp Endocrinol* 2007, **151:**195-201.
26. Bannister R, Beresford N, May D, Routledge EJ, Jobling S, Rand-Weaver M: **Novel estrogen receptor-related Transcripts in Marisa cornuarietis; a freshwater snail with reported sensitivity to estrogenic chemicals.** *Environmental science & technology* 2007, **41:**2643-2650.
27. Delsuc F, Brinkmann H, Chourrout D, Philippe H: **Tunicates and not cephalochordates are the closest living relatives of vertebrates.** *Nature* 2006, **439:**965-968.
28. Marletaz F, Holland LZ, Laudet V, Schubert M: **Retinoic acid signaling and the evolution of chordates.** *International journal of biological sciences* 2006, **2:**38-47.
29. Blair JE, Hedges SB: **Molecular phylogeny and divergence times of deuterostome animals.** *Mol Biol Evol* 2005, **22:**2275-2284.
30. Schubert M, Escriva H, Xavier-Neto J, Laudet V: **Amphioxus and tunicates as evolutionary model systems.** *Trends in ecology & evolution (Personal edition)* 2006, **21:**269-277.
31. Holland LZ, Gibson-Brown JJ: **The *Ciona intestinalis* genome: when the constraints are off.** *Bioessays* 2003, **25:**529-532.
32. Ikuta T, Saiga H: **Organization of Hox genes in ascidians: present, past, and future.** *Dev Dyn* 2005, **233:**382-389.
33. Dehal P, Satou Y, Campbell RK, Chapman J, Degnan B, De Tomaso A, Davidson B, Di Gregorio A, Gelpke M, Goodstein DM, Harafuji N, Hastings KE, Ho I, Hotta K, Huang W, Kawashima T, Lemaire P, Martinez D, Meinertzhagen IA, Necula S, Nonaka M, Putnam N, Rash S, Saiga H, Satake M, Terry A, Yamada L, Wang HG, Awazu S, Azumi K, *et al.*: **The draft genome of *Ciona intestinalis*: insights into chordate and vertebrate origins.** *Science* 2002, **298:**2157-2167.
34. Howard-Ashby M, Materna SC, Brown CT, Chen L, Cameron RA, Davidson EH: **Gene families encoding transcription factors expressed in early development of *Strongylocentrotus purpuratus*.** *Developmental biology* 2006, **300:**90-107.
35. Holland LZ, Albalat R, Azumi K, Benito-Gutiérrez E, Blow MJ, Bronner-Fraser M, Brunet F, Butts T, Candiani S, Dishaw LJ, Ferrier DE, Garcia-Fernàndez J, Gibson-Brown JJ, Gissi C, Godzik A, Hallböök F, Hirose D, Hosomichi K, Ikuta T, Inoko H, Kasahara M, Kasamatsu J, Kawashima T, Kimura A, Kobayashi M, Kozmik Z, Kubokawa K, Laudet V, Litman GW, Mchardy AC, Meulemans D, Nonaka M, Olinski RP, Pancer Z, Pennacchio LA, Rast JP, Rigoutsos I, Robinson-Rechavi M, Roch G, Saiga H, Sasakura Y, Satake M, Satou Y, Schubert M, Sherwood N, Shiina T, Takatori N, Tello J, Vopalensky P, Wada S, Xu A, Ye Y, Yoshida K, Yoshizaki F, Yu JK, Zhang Q, Zmasek CM, de Jong PJ, Osoegawa K, Putnam NH, Rokhsar DS, Satoh N, Holland PW: **The amphioxus genome illuminates vertebrate origins and cephalochordate biology.** *Genome Res* 2008, **18:**1100-1111.
36. Bardet PL, Schubert M, Horard B, Holland LZ, Laudet V, Holland ND, Vanacker JM: **Expression of estrogen-receptor related receptors in amphioxus and zebrafish: implications for the evolution of posterior brain segmentation at the invertebrate-to-vertebrate transition.** *Evolution & development* 2005, **7:**223-233.
37. Felsenstein J: **Cases in which parsimony or compatibility methods will be positively misleading.** *Systematic zoology* 1978, **27:**401-410.
38. Shimodaira H, Hasegawa M: **CONSEL: for assessing the confidence of phylogenetic tree selection.** *Bioinformatics (Oxford, England)* 2001, **17:**1246-1247.
39. Anisimova M, Gascuel O: **Approximate likelihood-ratio test for branches: A fast, accurate, and powerful alternative.** *Systematic biology* 2006, **55:**539-552.
40. Guindon S, Gascuel O: **A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood.** *Systematic biology* 2003, **52:**696-704.

41. Lartillot N, Brinkmann H, Philippe H: **Suppression of long-branch attraction artefacts in the animal phylogeny using a site-heterogeneous model.** *BMC evolutionary biology* 2007, **7(Suppl 1):**S4.
42. Zhang Z, Burch PE, Cooney AJ, Lanz RB, Pereira FA, Wu J, Gibbs RA, Weinstock G, Wheeler DA: **Genomic analysis of the nuclear receptor family: new insights into structure, regulation, and evolution from the rat genome.** *Genome research* 2004, **14:**580-590.
43. Kuiper GG, Carlsson B, Grandien K, Enmark E, Haggblad J, Nilsson S, Gustafsson JA: **Comparison of the ligand binding specificity and transcript tissue distribution of estrogen receptors alpha and beta.** *Endocrinology* 1997, **138:**863-870.
44. Gehm BD, McAndrews JM, Chien PY, Jameson JL: **Resveratrol, a polyphenolic compound found in grapes and wine, is an agonist for the estrogen receptor.** *Proceedings of the National Academy of Sciences of the United States of America* 1997, **94:**14138-14143.
45. Wang LQ: **Mammalian phytoestrogens: enterodiol and enterolactone.** *Journal of chromatography* 2002, **777:**289-309.
46. Dahlman-Wright K, Cavailles V, Fuqua SA, Jordan VC, Katzenellenbogen JA, Korach KS, Maggi A, Muramatsu M, Parker MG, Gustafsson JA: **International Union of Pharmacology. LXIV. Estrogen receptors.** *Pharmacological reviews* 2006, **58:**773-781.
47. Wetherill YB, Akingbemi BT, Kanno J, McLachlan JA, Nadal A, Sonnenschein C, Watson CS, Zoeller RT, Belcher SM: **In vitro molecular mechanisms of bisphenol A action.** *Reprod Toxicol* 2007, **24:**178-198.
48. Yang Z: **PAML 4: phylogenetic analysis by maximum likelihood.** *Mol Biol Evol* 2007, **24:**1586-1591.
49. Tanenbaum DM, Wang Y, Williams SP, Sigler PB: **Crystallographic comparison of the estrogen and progesterone receptor's ligand binding domains.** *Proceedings of the National Academy of Sciences of the United States of America* 1998, **95:**5998-6003.
50. Ekena K, Weis KE, Katzenellenbogen JA, Katzenellenbogen BS: **Identification of amino acids in the hormone binding domain of the human estrogen receptor important in estrogen binding.** *The Journal of biological chemistry* 1996, **271:**20053-20059.
51. Aliau S, Mattras H, Richard E, Bonnafous JC, Borgna JL: **Differential interactions of estrogens and antiestrogens at the 17 beta-hydroxyl or counterpart hydroxyl with histidine 524 of the human estrogen receptor alpha.** *Biochemistry* 2002, **41:**7979-7988.
52. Baker ME, Chandsawangbhuwana C: **Motif analysis of amphioxus, lamprey and invertebrate estrogen receptors: toward a better understanding of estrogen receptor evolution.** *Biochem Biophys Res Commun* 2008, **371:**724-728.
53. Brzozowski AM, Pike AC, Dauter Z, Hubbard RE, Bonn T, Engstrom O, Ohman L, Greene GL, Gustafsson JA, Carlquist M: **Molecular basis of agonism and antagonism in the oestrogen receptor.** *Nature* 1997, **389:**753-758.
54. Pick H, Jankevics H, Vogel H: **Distribution plasticity of the human estrogen receptor alpha in live cells: distinct imaging of consecutively expressed receptors.** *Journal of molecular biology* 2007, **374:**1213-1223.
55. Shiau AK, Barstad D, Loria PM, Cheng L, Kushner PJ, Agard DA, Greene GL: **The structural basis of estrogen receptor/coactivator recognition and the antagonism of this interaction by tamoxifen.** *Cell* 1998, **95:**927-937.
56. Nolte RT, Wisely GB, Westin S, Cobb JE, Lambert MH, Kurokawa R, Rosenfeld MG, Willson TM, Glass CK, Milburn MV: **Ligand binding and co-activator assembly of the peroxisome proliferator-activated receptor-gamma.** *Nature* 1998, **395:**137-143.
57. Ohno S: **Evolution by gene duplication.** Berlin, New York,: Springer-Verlag;; 1970.
58. Pratt WB, Toft DO: **Steroid receptor interactions with heat shock protein and immunophilin chaperones.** *Endocrine reviews* 1997, **18:**306-360.
59. Greschik H, Wurtz JM, Sanglier S, Bourguet W, van Dorsselaer A, Moras D, Renaud JP: **Structural and functional evidence for ligand-independent transcriptional activation by the estrogen-related receptor 3.** *Molecular cell* 2002, **9:**303-313.
60. Callard GV, Pudney JA, Kendall SL, Reinboth R: ***In vitro* conversion of androgen to estrogen in amphioxus gonadal tissues.** *Gen Comp Endocrinol* 1984, **56:**53-58.
61. Castro LF, Santos MM, Reis-Henriques MA: **The genomic environment around the Aromatase gene: evolutionary insights.** *BMC evolutionary biology* 2005, **5:**43.
62. Baker ME: **Co-evolution of steroidogenic and steroid-inactivating enzymes and adrenal and sex steroid receptors.** *Mol Cell Endocrinol* 2004, **215:**55-62.
63. Bryan MB, Young BA, Close DA, Semeyn J, Robinson TC, Bayer J, Li W: **Comparison of synthesis of 15 alpha-hydroxylated steroids in males of four North American lamprey species.** *Gen Comp Endocrinol* 2006, **146:**149-156.
64. Paris M, Escriva H, Schubert M, Brunet F, Brtko J, Ciesielski F, Roecklin D, Vivat-Hannah V, Jamin EL, Cravedi J-P, Scanlan TS, Renaud J-P, Holland ND, Laudet V: **Amphioxus Postembryonic Development Reveals the Homology of Chordate Metamorphosis.** *Current Biology* 2008, **18:**825-830.
65. Losel R, Wehling M: **Nongenomic actions of steroid hormones.** *Nature reviews* 2003, **4:**46-56.
66. Wehling M, Losel R: **Non-genomic steroid hormone effects: membrane or intracellular receptors?** *J Steroid Biochem Mol Biol* 2006, **102:**180-183.
67. Losel RM, Falkenstein E, Feuring M, Schultz A, Tillmann HC, Rossol-Haseroth K, Wehling M: **Nongenomic steroid action: controversies, questions, and answers.** *Physiological reviews* 2003, **83:**965-1016.
68. Bryan MB, Scott AP, Li W: **The sea lamprey (*Petromyzon marinus*) has a receptor for androstenedione.** *Biology of reproduction* 2007, **77:**688-696.
69. Thomas P, Pang Y, Dong J, Groenen P, Kelder J, de Vlieg J, Zhu Y, Tubbs C: **Steroid and G protein binding characteristics of the seatrout and human progestin membrane receptor alpha subtypes and their evolutionary origins.** *Endocrinology* 2007, **148:**705-718.
70. Thornton JW: **Resurrecting ancient genes: experimental analysis of extinct molecules.** *Nat Rev Genet* 2004, **5:**366-375.
71. Bridgham JT, Carroll SM, Thornton JW: **Evolution of hormone-receptor complexity by molecular exploitation.** *Science* 2006, **312(5770):**97-101.
72. Williams PD, Pollock DD, Blackburne BP, Goldstein RA: **Assessing the accuracy of ancestral protein reconstruction methods.** *PLoS computational biology* 2006, **2:**e69.
73. Volakakis N, Malewicz M, Kadkhodai B, Perlmann T, Benoit G: **Characterization of the Nurr1 ligand-binding domain co-activator interaction surface.** *Journal of molecular endocrinology* 2006, **37:**317-326.
74. Iwema T, Billas IM, Beck Y, Bonneton F, Nierengarten H, Chaumot A, Richards G, Laudet V, Moras D: **Structural and functional characterization of a novel type of ligand-independent RXR-USP receptor.** *The EMBO journal* 2007, **26:**3770-3782.
75. Lafont R, Mathieu M: **Steroids in aquatic invertebrates.** *Ecotoxicology (London, England)* 2007, **16:**109-130.
76. Lange IG, Hartel A, Meyer HH: **Evolution of oestrogen functions in vertebrates.** *J Steroid Biochem Mol Biol* 2002, **83:**219-226.
77. Lu NZ, Wardell SE, Burnstein KL, Defranco D, Fuller PJ, Giguere V, Hochberg RB, McKay L, Renoir JM, Weigel NL, Wilson EM, McDonnell DP, Cidlowski JA: **International Union of Pharmacology. LXV. The pharmacology and classification of the nuclear receptor superfamily: glucocorticoid, mineralocorticoid, progesterone, and androgen receptors.** *Pharmacological reviews* 2006, **58:**782-797.
78. Olsson PE, Berg AH, von Hofsten J, Grahn B, Hellqvist A, Larsson A, Karlsson J, Modig C, Borg B, Thomas P: **Molecular cloning and characterization of a nuclear androgen receptor activated by 11-ketotestosterone.** *Reprod Biol Endocrinol* 2005, **3:**37.
79. Jaillon O, Aury JM, Brunet F, Petit JL, Stange-Thomann N, Mauceli E, Bouneau L, Fischer C, Ozouf-Costaz C, Bernot A, Nicaud S, Jaffe D, Fisher S, Lutfalla G, Dossat C, Segurens B, Dasilva C, Salanoubat M, Levy M, Boudet N, Castellano S, Anthouard V, Jubin C, Castelli V, Katinka M, Vacherie B, Biemont C, Skalli Z, Cattolico L, Poulain J, *et al.*: **Genome duplication in the teleost fish *Tetraodon nigroviridis* reveals the early vertebrate proto-karyotype.** *Nature* 2004, **431:**946-957.
80. Bryan MB, Scott AP, Li W: **Sex steroids and their receptors in lampreys.** *Steroids* 2008, **73:**1-12.
81. Markov G, Paris M, Bertrand S, Laudet V: **The evolution of the ligand/receptor couple: A long road from comparative endocrinology to comparative genomics.** *Molecular and Cellular Endocrinology* 2008 in press.

82. Markov G, Lecointre G, Demeneix B, Laudet V: **The «street light syndrome», or how protein taxonomy can bias experimental manipulations.** *Bioessays* 2008, **30:**349-357.
83. Reschly EJ, Ai N, Ekins S, Welsh WJ, Hagey LR, Hofmann AF, Krasowski MD: **Evolution of the bile salt nuclear receptor FXR in vertebrates.** *Journal of lipid research* 2008.
84. Reschly EJ, Ai N, Welsh WJ, Ekins S, Hagey LR, Krasowski MD: **Ligand specificity and evolution of liver × receptors.** *J Steroid Biochem Mol Biol* 2008.
85. Reschly EJ, Bainy AC, Mattos JJ, Hagey LR, Bahary N, Mada SR, Ou J, Venkataramanan R, Krasowski MD: **Functional evolution of the vitamin D and pregnane × receptors.** *BMC evolutionary biology* 2007, **7:**222.
86. Iyer M, Reschly EJ, Krasowski MD: **Functional evolution of the pregnane × receptor.** *Expert opinion on drug metabolism & toxicology* 2006, **2:**381-397.
87. Delaunay F, Pettersson K, Tujague M, Gustafsson JA: **Functional differences between the amino-terminal domains of estrogen receptors alpha and beta.** *Molecular pharmacology* 2000, **58:**584-590.
88. Penttinen P, Jaehrling J, Damdimopoulos AE, Inzunza J, Lemmen JG, Saag P van der, Pettersson K, Gauglitz G, Makela S, Pongratz I: **Diet-derived polyphenol metabolite enterolactone is a tissue-specific estrogen receptor activator.** *Endocrinology* 2007, **148:**4875-4886.
89. Edgar RC: **MUSCLE: multiple sequence alignment with high accuracy and high throughput.** *Nucleic acids research* 2004, **32:**1792-1797.
90. Galtier N, Gouy M, Gautier C: **SEAVIEW and PHYLO_WIN: two graphic tools for sequence alignment and molecular phylogeny.** *Comput Appl Biosci* 1996, **12:**543-548.
91. Felsenstein J: **Confidence limits on phylogenies: An approach using the bootstrap.** *Evolution* 1985, **39:**783-791.
92. Swofford DL, Waddell PJ, Huelsenbeck JP, Foster PG, Lewis PO, Rogers JS: **Bias in phylogenetic estimation and its relevance to the choice between parsimony and likelihood methods.** *Systematic biology* 2001, **50:**525-539.
93. Wada H, Escriva H, Zhang S, Laudet V: **Conserved RARE localization in amphioxus Hox clusters and implications for Hox code evolution in the vertebrate neural crest.** *Dev Dyn* 2006, **235:**1522-1531.
94. Wang Z, Benoit G, Liu J, Prasad S, Aarnisalo P, Liu X, Xu H, Walker NP, Perlmann T: **Structure and function of Nurr1 identifies a class of ligand-independent nuclear receptors.** *Nature* 2003, **423:**555-560.

Part III

DISCUSSION

10

EVOLUTION OF NR LIGAND BINDING ABILITY

During my thesis I focused on the role of the TH signaling pathway in the evolution of metamorphosis in chordates. Chapters **??**, **??** and **??** were dedicated to this subject, whereas chapter **??** focused on the evolution of NRs in general and ERs in particular. In the present chapter, I will discuss the evolution of NRs and, in a second time, I will discuss the evolution of chordates in light of my thesis work (chapter **??**).

In human, a thyroid hormone receptor usually refers to TR, as an estradiol receptor usually refers to ER. However, this assessment is not relevant in species distantly related to mammals. As explained in chapter **??**, amphioxus ER does not bind estradiol and therefore is not an estradiol receptor. Consequently, the evolution of the gene and the function performed by the gene are not identical. For instance, TRs were present in the ancestral bilaterian. However, TH binding by TR has been demonstrated only in chordates (chapter **??**). Since the discussion below, which concerns the evolution of NRs takes into account these two aspects (the gene and its function), it is important to keep the distinction in mind.

10.1 EVOLUTION OF THYROID HORMONE RECEPTORS

Thyroid hormone receptors (TRs) have been found in all major groups of bilaterians. However, their function has been well studied only in chordates, where it was shown to be well conserved throughout chordate evolution. Indeed, we have shown that all TRs (except that of *Ciona intestinalis*) are TH-dependent transcription factors. Accordingly, the 3D structure of TRs has also been well conserved (compare the 3D structures of the amphioxus TR and the human TRα, on figure **??**b in chapter **??**). Consequently, TR ligand binding has been quite stable since chordate origins. Whether this statement may be generalized to all deuterostomes and even bilaterians, whose genome contains an uncharacterized TR, will need further investigation (see section **??** below). Interestingly, properties of coactivator recruitment have been well conserved (chapters **??** and **??**) whereas corepressor releasing seems to show some discrepencies (chapter **??**). Whether this peculiarity can be applied generally to TRs is still unclear.

10.2 EVOLUTION OF ESTROGEN RECEPTORS

Chapter **??** showed that the estrogen receptor (ER) from amphioxus does not bind estradiol, like all non-vertebrate ERs cloned so far and unlike all vertebrate (including lamprey) ERs. Parsimony reasoning leads us to conclude that estradiol binding was acquired during vertebrate evolution (**?**). However, this conclusion raises several questions since estradiol may be present in amphioxus (e.g. estradiol detection by RIA, cloning of enzymes involved in estradiol synthesis, see (**?**) and references therein). Then what receptor may mediate the potential estradiol signal? We and **?** proposed that an interesting candidate was the amphioxus NR3C, homologous to vertebrate mineralocorticoid, glucocorticoid, progesterone and androgen receptors (MR, GR, PR and AR, figure **??**). Accordingly, **?** reported that the amphioxus NR3C binds estradiol and regulated transcription upon estradiol binding, in addition to confirming our results on amphioxus ER ligand binding properties. Moreover, amphioxus NR3C is expressed in germ cells, suggesting a potential role in the regulation of reproduction cycle. This finding is not very surprising, as we suggested in the discussion of chapter **??**, because amphioxus NR3C is ER-like (**?**). **?** proposed that the common ancestor of NR3A (ERs) and NR3C (MR, GR, AR and PR) was an estradiol receptor (figure **??**). According to this scenario, amphioxus ER independently lost the ancestral ability to bind estradiol. This scenario, of an independent loss of estradiol binding by amphioxus ER, is supported by sequence analysis. Indeed, amphioxus ER displays some divergent sites, that are sufficient to explain alterations in estradiol binding (e.g. a cysteine at position 531 instead of an arginine) (**?**). Notably **?** showed that some point mutations of amphioxus NR3C amino-acids into amino-acids of the amphioxus ER were sufficient to suppress estradiol binding. As these sites are amphioxus ER-specific, whereas other ERs, including mollusk ones, are "vertebrate-like" at these positions, this result suggests that amphioxus ER may have specifically lost the ability to bind estradiol from an estradiol binding. Here it is important to understand that the evolution of ERs does not necessarily equal the evolution of the function carried out by the receptor (e.g. there are ERs in many bilaterians, but only vertebrate ERs bind estradiol).

The article by **?** does put the results and the conclusions of our work in a different perspective (**?**). From our results and previous ones, no estradiol receptor could be found outside vertebrates. However, the finding of **?**, that amphioxus NR3C binds estradiol, may push back to our bilaterian ancestor the appearance of estradiol binding by steroid receptors. According to this scenario, estradiol was bound by two receptors in the common ancestor of bilaterians until vertebrate split. Then

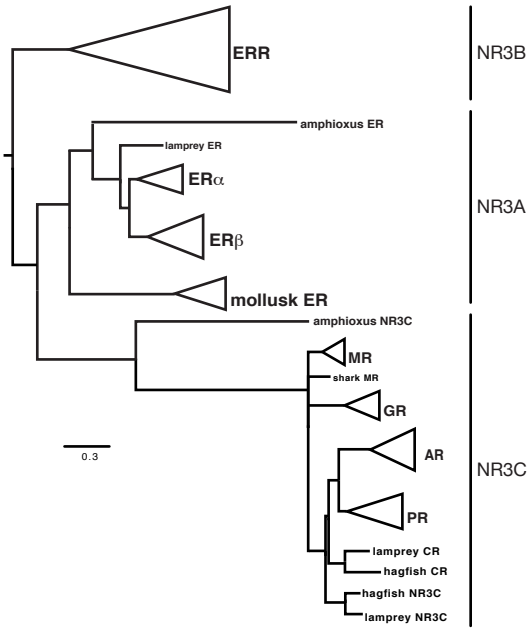


Figure 37: Reminder of NR3 phylogeny. The NR3 subfamily is constituted of the NR3A (ERs), NR3B (ERRs) and NR3C (other sex steroid receptors). NR3Bs branch at the root of this tree. The NR3A contains ERs from amphioxus, lamprey, mollusks as well as ERα and ERβ from a vertebrate specific duplication. NR3Cs contain AR, MR, GR, AR and PR, issued from vertebrate specific duplications, except for basal vertebrates (lamprey, hagfish) that diverged before some of the vertebrate-specific duplications. Amphioxus has one representative of each group.

NR3C was lost in protostomes while NR3A was also lost in ecdysozoans and became unresponsive to estradiol in mollusks (figure **??**). In amphioxus, NR3C became the only estradiol receptor while specifically in the amphioxus lineage ER lost its ability to bind estradiol and in the vertebrate lineage, the NR3Cs underwent point mutations leading to mineralocorticoid, glucocorticoid, androgen and progesterone binding (**???**). A subsequent question would then be the relative functions of the two apparently redundant receptors during early bilaterian history. In addition and until now, there is no estradiol receptor in mollusks whereas estradiol effects have been reported. The nature of the corresponding receptor needs to be answered ((**?**) and references therein). A last interrogation regards the absence of non-estradiol steroid receptor in amphioxus: it is important to determine whether non-estradiol steroids are active in amphioxus, and if so, through which receptor. These different questions need to be answered for a better understanding the evolution of estradiol binding and more generally of steroid binding.

Overall, several conclusions can be drawn from our study (**?**) and the one from **?**: (i) the amphioxus ER is not an estradiol receptor. (ii) whether the ancestral NR3 bound estradiol or did not, the evolution of steroid binding has been dynamic with several events of loss, gain and modifications of ligand binding, (iii) although computational predictions are informative, they should be considered as only indicative and (iv) the comparative approach, used in both articles, is an especially powerful tool to study evolution. (v) However, almost all we know about steroid hormone binding comes from studies in chordates. Many lineages remain

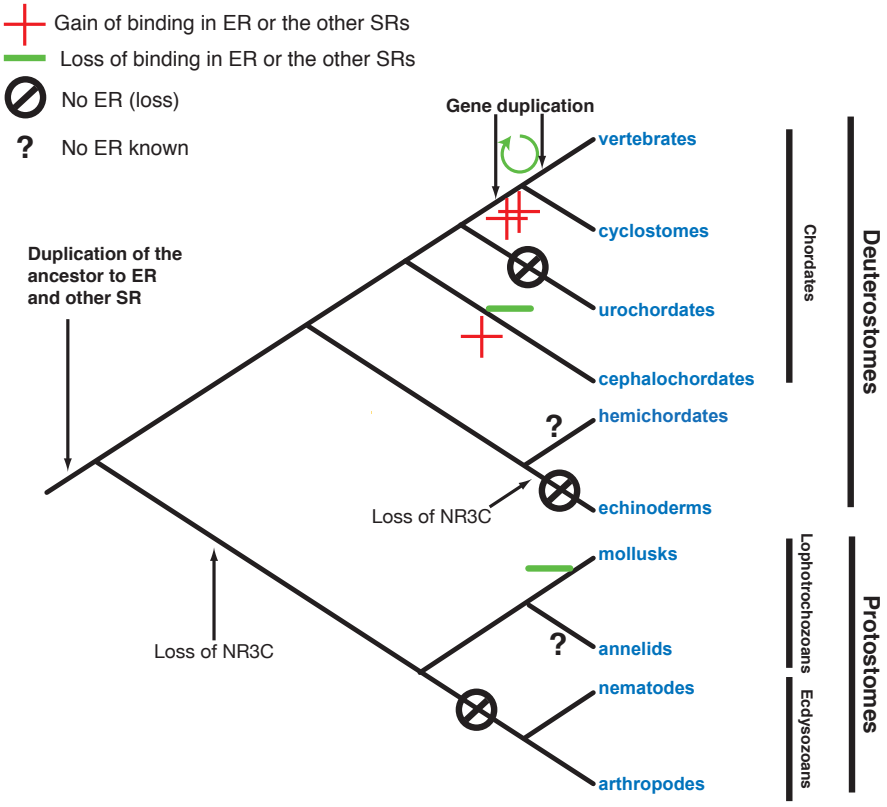


Figure 38: Scenarios of estradiol binding in bilaterians. Red signs describe a scenario according to which estradiol binding was recently acquired during evolution. Green signs describe a scenario according to which estradiol binding is ancestral in bilaterians.

to be studied, especially in protostomes. (vi) Finally, and related to the former statement, steroidogenesis should be much more thoroughly investigated, from a non-anthropocentric point of view, in order to better grasp steroid hormone signaling pathway outside vertebrates, and thus to better understand its evolution (**?**).

10.3 EVOLUTION OF NR LIGAND BINDING

The evolution of NRs has been seen as modular, with the appearance of ligand binding as a control switch of NR transcriptional activity. However, it is still unclear how this ability appeared and evolved. Although TR ligand binding has been investigated with a rather small taxonomic sampling, it seems to have been fairly well conserved, since all chordate TRs tested so far (to the exception of the *Ciona intestinalis* TR) are TH-dependent transcription factors. The only difference is the exact nature of the ligand: T_3 in vertebrates and a T_3 derivative, possibly TRIAC, in amphioxus. This difference is conservative since vertebrate TRs do bind TRIAC as well. Similarly, although ERs experienced several events of gain and loss of ligand binding ability, the evolution of ligand binding in the NR3 family in general was confined to refinements of affinities to steroid hormones (figure **??**) and members of the NR4 subfamily appear to lack a ligand binding pocket (in yellow on the figure **??**). However, this conservative evolution cannot be generalized to all families. For instance, some NR1 members are high affinity or low affinity receptors of very different ligands (TR, RAR, VDR, PPAR, Rev-erb), while others have constitutive ligands (RORα) or are still orphan (RORγ) (figure **??**). In the NR2 subfamily, our lab has contributed to bring to light the understanding of RXR/USP evolution. Indeed, the ancestral RXR (in the common ancestor of bilaterians and cnidarians) is supposed to display similar properties as the human RXR for 2 reasons: first, mollusk USP (**?**), and cnidarian USP (**?**) bind 9*cis*-retinoic acid and activate transcription upon ligand binding; second sequence analysis revealed that LBPs from all RXRs (except arthropod ones) have experienced a conserved evolution (**??**). In the arthropod lineage, RXR-USP became orphan and in mecopterida (including insects and lepidopterans), USP ligand affinity changed to phospholipid (**?**). Notably, in all arthropods, USP is found in an antagonist conformation. Overall this RXR-USP example illustrates the plasticity of NR ligand binding evolution. At the scale of the whole family, It is still very obscure how this diversity in ligand binding arose.

Although it is not unlikely (although not clear yet) that the ancestral NR3 was a steroid receptor (**??**), the nature of the ancestor of NR1, of NR2... or of all NRs

is much more elusive. The structure of different LBDs revealed a common 3D structure but different strategies of activation, *i.e.* a similar active conformation but different activation processes (see the introduction of chapter **??** for further details). This diversity suggests that NR ligand binding evolution has necessarily been non-linear and is difficult to assess. Indeed, ligand binding properties are fast evolving (for instance human LRH-1 is a ligand-dependent transcription factor whereas mouse LRH-1 activity is believed to be ligand-independent with a large empty pocket (**?**)). As **?** highlighted, NRs are evolutionary plastic, which may have blurred the signal of the ancestral state to a point where it becomes difficult to unearth.

Although ancestral reconstructions should be considered as only indicative (see chapter **??**), what may these approaches teach us about the evolution of NRs? In order to try and estimate the ancestral NR state, it is possible to map on an NR phylogenetic tree the characteristics of each NR and to use statistics to infer the ancestral state. I have thus inferred the 2 following characteristics of ancestral NRs: (**a**) the size of the ligand binding pocket as well as (**b**) ligand binding properties. The different states that were tested were (i) specific and high affinity endocrine of paracrine receptors (like TR or ER) ; (ii) poorly specific and low-affinity nutritional sensors (like PPAR or LXR) ; (iii) receptors with structural ligands (like some USP), (iv) receptors with an empty pocket (like ERR); or (v) no pocket (like Nurr1) (**?**). This experiment was done with the software Mesquite (**?**)

When one follows this procedure, the ancestral NR is estimated to have (**a**) a rather small pocket (~ 420 $Å^3$, figure **??**), like high affinity endocrine or paracrine NRs. In addition, using a model where all transitions between the different tested ligand binding properties are equiprobable, (**b**) 3 states are considered as significantly better than the other ones: "endocrine or paracrine", "no pocket" and "sensor" states were all likely (P-values of 0.67, 0.11 and 0.14, respectively).

The experiments of reconstruction of ancestral LBP volumes and ligand binding properties gave similar results: the ancestral NR is inferred to be a high affinity receptor. Indeed, endocrine or paracrine receptors (that are high affinity receptors) usually have a small LBP that just fits the ligand (e.g. the volume of the RARγ LBP is 412 $Å^3$ (**?**) whereas the volume of its ligand 9-cis retinoic acid is 291 $Å^3$). Following this hypothesis, the nature of the ancestral ligands remains an enigma that will be difficult to assess. A complementary approach may be to estimate the gene-content of the ancestral organism, to determine which metabolic pathway was present, that may reflect the endogenous ligand. However, one should keep in mind that the evolutionary origin of a ligand may have been exogenous (like a vitamin) and its endogenous synthesis subsequently acquired during evolution (**?**).

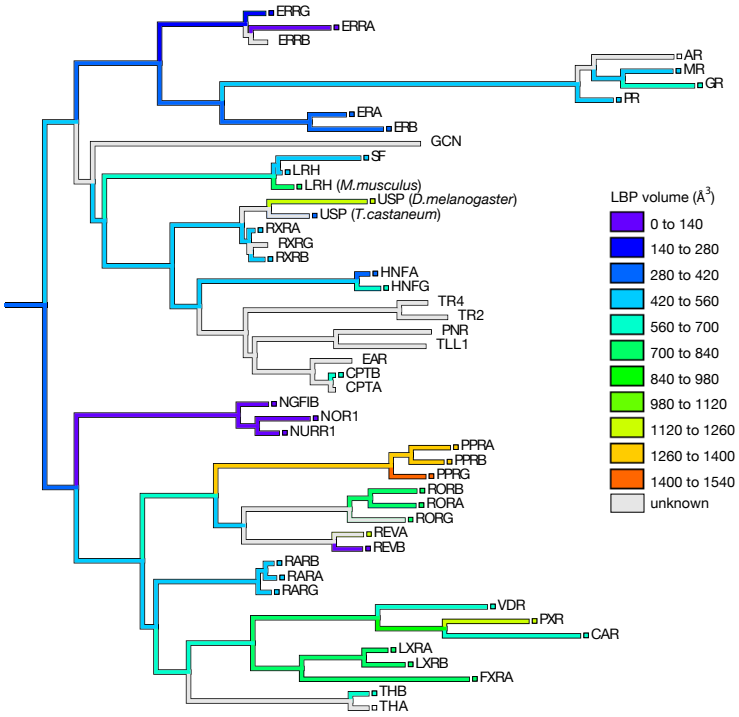


Figure 39: Estimation of the size of NR ligand binding pockets. A phylogenetic tree of human NRs (plus LRH-1 from *M. musculus* and 2 insect USPs) was built using PhyML (**?**). Then the size of each NR LBP was plotted on this tree, the sizes of the ancestral sequences were estimated by maximum likelihood and color-coded on the phylogenetic tree, using the software Mesquite (**?**).

One of the possible ancestral states inferred from our analysis, an orphan receptor with an empty pocket is in agreement with our estimation of the ancestral LBP size, since some NRs with a small pocket are transcriptionally active in the absence of a ligand (e.g. ERR). This scenario of an ancestral orphan receptor, that was a ligand-independent transcription factor has previously been proposed (**??**). This hypothesis was based on the observation that NR phylogeny does not reflect the distribution of ligands in the tree (e.g. steroid receptors are paraphyletic) and the size of the LBP. This hypothesis also allows tan easier interpretation of the apparently complex evolution of ligand binding since an ancestral small LBP could have adapted to various evolutionary scenarios without no loss of function. Indeed the ancestral NR could have evolved from a state of no binding to independent grains of ligand binding by parallel point mutations.

The third hypothesis of an ancestral "sensor" receptor is appealing for several reasons. First, the NR family appeared in the ancestral metazoan (**?**), which was most probably a multicellular animal with few cellular types and displaying a limited level of integration. Thus an endocrine or paracrine function seems rather useless in such conditions. Intuitively, the ancestral NR may have served as an environmental "sensor", that could have interacted with many compounds (like extant "sensors" (**?**)) and would thus have helped reacting in an appropriate way to external conditions (abundant food? toxic compounds? etc...). Indeed, such a transcription factor which activity is regulated upon metabolic or foreign compounds is found in humans (e.g. the NRs PPAR and PXR) but also in yeast (e.g. Oaf1 and Pdr1p). In humans PPAR has a large LBP (figure **??**) that binds products of fatty acid metabolism (**?**) and PXR, also with a large LBP, binds a wide range of foreign compounds and is involved in detoxification (**?**). In yeast, Oaf1 is known to induce fatty acid metabolism. Although its proposed homology with NRs is doubtful, the protein was predicted to have a NR LBP-like 3D structure, it was also shown to interact with oleate and to activate transcription upon ligand-binding (**?**). Another study highlighted the analogous function of the human sensor PXR and the yeast TF Pdr1p: they both bind structurally diverse compounds with low affinity (μM order) and regulate gene activation upon ligand binding (**?**). Although there is no clear evidence that the above yeast TFs and NRs are homologous, these examples illustrate the biological role that NRs may have played in our ancient animal ancestor. Second, The hypothesis of an ancestral "sensor" explains the existence of a LBP in so many NRs and provides an evolutionary framework to the evolution of the many activation strategies that appeared in the NR family. Indeed, if the ancestral NR had a large LBP in which many compounds may have fit in it, then it could have secondarily easily adapted to one specific ligand,

like for endocrine NRs, or lose its dependancy to ligand binding, like in NR4. This hypothesis of an ancestral role of NRs in the regulation of xenobiotic action has been previously proposed (**?**) and is supported by a phylogenetic analysis of genes involved in steroidogenesis recently performed in our lab, that suggests that extant NR ligands are the evolutionary end-product of metabolic pathway linked to xenobiotics degradation (environmental chemicals) (**?**). However, although this hypothesis is attractive, it is still very speculative. Moreover, an ancestral "sensor" NR would probably have a large LBP (like LXR), which is not supported by the estimation of the ancestral LBP volume (figure **??**).

Overall, the above experiment seems to favor a ligand-binding state, to the opposite of ligand-independent state. However, the resolution is still very poor, since only 2 states out of 5 were rejected with the model used. It is possible that a model where all transitions are equiprobable is not realistic. On the other hand, the taxonomic sampling used here may be biased. Indeed human NRs were almost exclusively used because almost all functional assays were done on them. It will be very interesting in the next future to broaden our knowledge of NR ligand binding ability (especially in NR2) but most importantly to perform functional studies on NRs from other species. In this respect, NRs have been reported in all bilaterians (e.g. TR (**?**)), and in other metazoan groups like cnidarians (e.g. RXR (**?**)), placozoans (e.g. ERR (**?**)) or sponges (e.g. RXR (**?**)). In these species that diverged from human a long time ago, NRs may have evolved totally different characteristics that should be informative regarding ancestral NR states. Consequently, a larger taxonomic sampling is absolutely necessary to better understand the evolution of NRs.

10.4 CONCLUSIONS ON THE EVOLUTION OF NRS

Overall, the evolution of NR may be seen as follows: the ancestral NRs were transcription factors that bind DNA on combinations of (AGGTCA) motifs. How their transcriptional activity was originally regulated is still elusive. Was it ligand-binding dependent? Independent? As discussed above, more work needs to be done. What can be concluded here, is that the subsequent evolution was chaotic, with several episodes of gain, loss and alterations of ligand binding.

Although the evolution of ligand binding has focused the attention of evolutionary biologists, other NR features may be worth digging into, since the mode of action of NRs is better and better understood. For instance, how does the target gene repertoire evolve? A better understanding of how NR cis-regulatory sequences evolve should contribute to shed light on the evolution of NR function and more

generally of the regulatory network they are involved in. Such large-scale studies of cis-regulation are now technically possible and are appearing in the litterature (**?**), some of which focus on the evolutionary component of cis-regulation (**?**). Another issue of NR evolution may relate to the evolution of cofactors interaction and more generally of NR-protein interaction, which has been overlooked, except for few studies (e.g. (**?**)). These issues should help us understand the evolution of NR function in a broader context than the biochemical properties of the receptor alone and maybe illustrate the evolution of transcription factors in general.

11

EVOLUTION OF METAMORPHOSIS IN CHORDATES

As described in the introduction, because chordate metamorphoses are morphologically very diverse, they were proposed by several authors to have appeared independently several times in the chordate lineage (**????**). Recent data, including the work I have done during my PhD, have nevertheless shown that, in chordates like amphibians and actinopterygians, THs and other thyroactive compounds are playing a key role in the control of metamorphosis. The conservation of the molecular determinism of metamorphosis suggests a unique and ancient origin for this key post-embryonic developmental process in the chordate lineage. More precisely, we proposed that the chordate ancestor had an indirect development and metamorphosed under iodothyronine derivative stimulation (**?**). This hypothesis unifies all the postembryonic developmental stages regulated by THs and their derivatives since it implies that they are all homologous. This model also provides an evolutionary framework through which metamorphosis has diversified within chordates. We speculate that this hypothesis can even be extended to the deuterostome super-phylum, although key data are still missing to unambiguously demonstrate that TH-regulated metamorphosis is shared by all deuterostomes.

In the next sections, I will explain our model and give some examples of evolution of metamorphosis from alterations of TH signaling pathway (the following text was modified from (**?**)).

11.1 THE EVOLUTION OF METAMORPHOSIS IS GOVERNED BY THE EVOLUTION OF THE TH SIGNALING PATHWAY.

Our hypothesis that metamorphosis is an ancestral feature of chordates is supported by the conservation of its molecular determinism. Indeed THs, other iodinated tyrosine derivatives and their receptor TR have been systematically involved in metamorphosis in all chordates in which their role has been tested (as detailed in chapter **??** of the introduction, and in chapters **??**, **??** and **??** of this manuscript), even in animals located at very distant branches of the chordate phylogenetic tree (amphibians, actinopterygian fishes and amphioxus). Then parsimony reasoning

suggests that metamorphosis was already regulated by a iodothyronine compound and TR in the common ancestor of all chordates. Consequently the different larva-to-adult transitions described above should be considered as homologous, based on the conservation of their shared molecular determinism (and more precisely, the central part of the TH signaling pathway formed by the coupled thyroactive compound/TR, as represented in figure **??**). More precisely, we propose that in most chordates studied to date the onset of metamorphosis is characterized by a peak of a thyroactive compound, activating TR that modifies the expression of target genes and leads to morphological remodeling characteristic of the larva-to-juvenile transition (figure **??**a).

I argue that, in all chordates, the key component of TH signaling pathway involved in metamorphosis is the coupled TH / TR (as represented in figure **??**b). Thus TR activation by TH (T_3, T_4 or a more active TH derivative) can be viewed as the master switch controlling the post-embryonic larva-to-juvenile transition in all chordates. In this scenario, the coupled TH/TR is the most central part of the TH-metamorphosis network, and thus the most conserved. In contrast, the more upstream part (composed of genes involved in TH metabolism that regulate TH availability and TR activity) and the more downstream part (mainly TR target genes, whose expression is regulated by the coupled TH/TR and which control morphological remodeling leading to metamorphosis) are less constrained and more prone to diverge (figure **??**b). This model suggests the existence of a post-embryonic period during which TH/TR controls a key transition, molecularly conserved but morphologically extremely diverse. The differences that exist between the different chordate metamorphoses should be seen as derived rather than ancestral.

The evolution of metamorphosis and thus its wide diversity at the morphological level result from the evolution of TH signaling. Modifications of phenotypes linked to the larva-to-adult transition are effected by alterations of the TH signaling pathway, preferentially acting upon the thyroactive compound availability or the target genes, rather than on thyroactive compound as the trigger or TR as the mediator of that initial signal. Consequently, the differences that exist between these developmental processes should be seen as derived rather than ancestral. In the following sections, I discuss some cases of variability of metamorphosis, based on alterations of the TH signaling pathway at different levels.

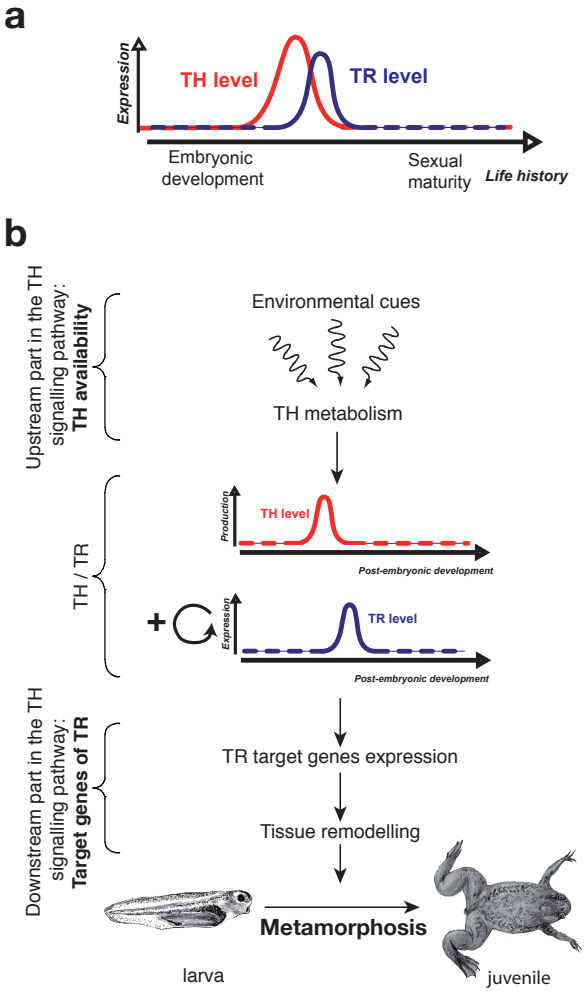


Figure 40: Chordate metamorphosis is regulated by THs and TR. **a.** Chordate metamorphosis may be defined as a short post-embryonic developmental stage, allowing a larva to become a juvenile through morphological, metabolic and ecological remodeling and which is triggered by a peak of TH production, mediated by a peak of TR expression. **b.** Schematic view of TH signaling pathway in chordates, separated in 3 modules (simplified version of figure **??**). The most central module (the couple TH/TR) was conserved throughout chordate evolution while the more peripheral modules evolved faster, and account for metamorphosis plasticity.

11.2 CHORDATE METAMORPHOSIS AND ALTERATIONS OF TH SIGNALING: SOME EXAMPLES

11.2.1 *Heterochrony and evolution of metamorphosis in amphibians: alteration of the TH signaling pathway.*

In amphibians, metamorphosis can be spectacular, as seen in the anuran *Xenopus laevis*. Still, amphibians exhibit a wide range of variations in the intensity of the morphological remodeling during metamorphosis. The modifications can be relatively inconspicuous, as for urodeles and apodes, or even reduced to the point of vanishing, as in neotenic amphibians. On the other hand the larval stage can be suppressed, as in direct developing frogs, where little frogs directly hatch. However, as will be exemplified, in all cases there is a TH sensitive period of morphological remodeling. TH signaling was studied in the entirely terrestrial and direct developing frog *Eleutherodactylus coqui* in an attempt to understand the molecular basis of the apparent suppression of the larval period. It was shown that *E. coqui* does metamorphose under TH stimulation but the main difference with "classic" amphibian metamorphosis is the very early triggering of this metamorphosis, as it occurs inside the egg (**?**). More precisely, during late developmental stages, there is a TH-dependent period of morphological remodeling (including cartilage remodeling, tail regression or limb elongation) correlated with a peak of TR? expression (**?**). Consequently, *E. coqui* metamorphosis is rather similar to the "classical anuran", except that the timing of TH production and TR expression have probably been shifted earlier during development. It is also possible that the TR target genes have diverged as well, since the morphological remodeling is not as drastic as in anurans. *E. coqui* development is a clear example of heterochrony (species-specific variation in the timing of the development, leading to changes in the phenotype of the species).

Alternative changes of TH signaling lead to metamorphoses that are very mild and make adults look too much like the larvae to be described as metamorphic. These animals are called paedomorphic, because the adults have retained larval traits, and constitute another case of heterochrony. As described in the introduction, there are several cases of paedomorphosis: in facultative paedomorphic species, like the Mexican axolotl (*Ambystoma mexicanum*), the tadpole responds to exogenous TH by metamorphosing, but does not spontaneously metamorphose (**?**). In fact the axolotl is a spectacular example of a series of species of the genus Ambystoma in which several types of unusual post-embryonic life history strategy occur (**?**). Facultative neoteny has been attributed especially to low levels of circulating

THs due to weak activation of the thyroid axis by the pituitary gland, when the rest of TH signaling pathway is "normally" functional (**??**). Obligatory neotenic amphibians (such as the perennibranchiates Urodeles *Necturus* or *Proteus*), long thought to be insensitive to THs, were in fact shown to respond to exogenous TH but with limited amplitude: for instance expression of the typical metamorphosis genes TR? and stromelysin are induced by TH (**?**). Rather, the lack of response to TH is probably caused by the loss of TH regulation by some target genes such as collagenase 3.

These examples of "alternative" metamorphoses illustrate the important plasticity of the metamorphosis process in different amphibians. Each time, TR was shown to be functional (**??**). Rather, the genetic variations thought to be responsible for phenotypic differences rely on the production of TH (an early production in *E. coqui*, or a default of TH production in the axolotl) or on the target genes (as in the *Necturus*) or both. These extreme examples show that metamorphosis triggered by TH is always present in amphibians, even if in some cases it is weakened or shifted during late development stages.

11.2.2 *Evolution of metamorphosis in other vertebrates: why direct development is not so direct in chordates.*

Under the hypothesis of an ancestral metamorphosis in the chordate lineage, then descending species should have divergent variants of this developmental stage. According to our definition of metamorphosis as a post-embryonic developmental stage characterized by a thyroactive compound-dependent morphological remodeling, groups that were previously thought to have a direct development may in fact be seen as animals with indirect development and a mild metamorphosis. Indeed, similarly to neotenic amphibians having a mild metamorphosis, mammals metamorphose. In these lineages, the metamorphosis stage does not coincide with birth since the latter is not marked by an intense TH signaling. Sexual maturity does not qualify as metamorphosis either since it is regulated by steroids and is often asynchronous with metamorphosis. On the contrary, mammalian weaning gathers all characteristics of metamorphosis as defined above as a period of tissue remodeling governed by TH (**?**). It is an ecological shift that is characterized by modifications like neural maturation, intestine remodeling or chondrogenesis that are reminiscent of similar events during amphibian metamorphosis. These modifications occur in correlation with a peak of TH production (both T_3 and T_4) (**?**) and were shown to be under the control of TH and TR (**???**). Consequently, weaning should be viewed as homologous to classical metamorphosis even if the precise

role of THs in controlling this event and the regulatory hierarchy controlled by these hormones should be better analyzed. In addition weaning is poorly studied outside classical mammalian models (e.g. mouse, rat, human) and it should be analyzed in a wider taxonomic sampling.

As discussed above, in actinopterygians, and especially in flat fishes, THs are important regulators of metamorphosis (**?**), the best known case being flounder, for which T_4 and T_3 were reported to induce precocious metamorphosis (**?**). But the role of THs in post-embryonic development has also been documented in other groups. In zebrafish, THs were shown to be important for the larva-to-juvenile transition: exogenous THs induce differentiation of the pectoral fin precociously, in contrast the inhibition of TH synthesis blocks it (**?**). Smoltification in salmonid fishes is considered to correspond to metamorphosis: it is a period of high T_4 and T_3 levels (**??**) involving a morphological, physiological and behavioral switch from a freshwater-adapted form to a salt-water-adapted form. At least some of the former modifications were shown to be TH-sensitive (**??**). The larval to adult transition is extremely important in terms of recruitment in many fish species that are living close to the substratum, such as coral fishes. Although a direct role of thyroid hormones in regulating metamorphosis has not yet been clearly demonstrated, in some species like the grouper, existing evidence clearly points in this direction (**?**).

In lamprey, although it is now well established that TH endogenous synthesis needs to be tightly regulated for proper metamorphosis, the role of TH displays a remarkable difference compared to other chordates. Strikingly, several reports clearly show that it is a drop, instead of an increase, in circulating TH that ends the larval period (**???????**). Among other factors, temperature and growth advancement (indicated by length and weight, for instance) are important parameters: lampreys enter metamorphosis when they have reached a certain size (**?**). Interestingly even if the inversion of TH action is unique to lampreys, TR is still likely to play a role in this TH-mediated action since nuclear TH binding was reported (**?**) and the TRs from lamprey are activated by TH (**?**). Given that it is a drop of TH levels that triggers metamorphosis, TR's role in controlling metamorphosis must be different in lamprey than in other vertebrates. It is possible that in lamprey, TR induces the expression of genes that inhibit metamorphosis. According to this scenario, when TH levels drops, TR will stop inducing the expression of these inhibitors, then triggering metamorphosis. However, many other hypotheses can be proposed, since in vivo functional experiments are still to be completed. The use of specific TR agonists and antagonists (**???**) could help deciphering the specific role of TR in this inverted situation.

Another interesting case of divergence on the TH signaling pathway is provided by urochordates. Metamorphosis of *Ciona intestinalis* can be induced by exogenous THs and inhibited by goitrogens known to block TH synthesis in vertebrates (**??**). However, the molecular pathways triggered by THs remain elusive since the only TR of *Ciona intestinalis* (**?**) does not bind any iodinated tyrosine derivative tested (**??**). The urochordates are particularly interesting regarding the evolution of metamorphosis because of the diversity of their post-embryonic developmental strategies. The elucidation of the role of THs and iodinated tyrosine derivatives in general in late development in other species, and in particular in larvaceans should bring interesting data for comparative analysis.

11.2.3 *TH signaling pathway and development outside chordates*

The role of THs in metamorphosis might even be extended to all deuterostomes. Indeed, most echinoderms metamorphose, although their life history is quite variable from one species to another (**?**). Interestingly, effects of TH on the metamorphosis of echinoderm larvae have now been reported in several families, including species with feeding and nonfeeding development: Echinoidea (sea urchins, sea biscuits, and sand dollars) asteroids (sea stars) and ophiuoidea (brittle stars) ((**?**) and references therein). Moreover, outside chordates, echinoderms are the only animals that display clear endogenous TH metabolism: T_4 production was demonstrated in sea urchin and sea biscuits (**??**). Interestingly orthologs of many proteins involved in mammal TH metabolism were found in the sea urchin genome (**??**). This suggests an even more ancestral origin of TH endogenous synthesis in deuterostomes. One TR was found in the genome of the sea urchin *Strongylocentrotus purpuratus* (**?**). It is not as divergent as the TR from *Ciona intestinalis*, making it a potential TH receptor. It will be interesting to test in vitro if this receptor can bind any iodinated tyrosine derivative and regulate transcription of TRE-containing genes in response to this hormone. As THs seem to play a similar role in echinoderm metamorphosis as in chordates, the involvement of TR in this process would be of particular interest.

TR genes were recently described in some protostomes (mollusks, flatworms and crustaceans) (**?**), implying that the gene was present in the common ancestor of all bilaterians and was secondarily lost in insects and nematodes. It will be very interesting to decipher the function of these genes. The encoded proteins are apparently able to bind similar DNA sequences and to heterodimerize with the classical nuclear receptor partner RXR (**?**) but are they active transcription factors? More particularly are they responsive to THs? Attempts to determine if THs had

any effects on protostome development are scant and relatively preliminary. One study reported that THs accelerated metamorphosis in the mollusk abalone (**?**) and another one in a crustacean (a prawn (**?**)). In contrast THs did not seem to have any effect on the crustacean *Daphnia* (**?**) in accordance with the observation that no endogenous TH is present in this species (**?**). Consequently, hypotheses on the roles of THs or other derivatives and TR in protostomes can only be tentative: a potential responsiveness to THs would imply a very ancient function of TR as the TH receptor or, on the other hand, the establishment of an alternative function would imply an evolutionary shift. In any case, it is necessary to better characterize the TH signaling pathway in protostomes: are they able to produce iodinated tyrosine derivatives, and if they do so, how? Are the TRs functional? What are the molecular events triggered by iodinated tyrosine derivatives? To what extent do they regulate metamorphosis? Those questions need to be addressed, in order to fully understand the evolution of TH signaling in bilaterian and its role in late development.

11.3 PLASTICITY OF TH SIGNALING AND PLASTICITY OF METAMORPHOSIS

11.3.1 *The evolution of metamorphosis is tightly linked to environmental conditions.*

We proposed a new definition of metamorphosis highlighting the role of THs and TH derivatives as the triggering element and TR as the direct mediator of TH signal in chordates. This definition is based on the hypothesis of an ancestral origin of TH and thyroactive compound-induced metamorphosis in chordates. From this definition, all chordates metamorphose under thyroactive compound stimulation. We postulate that the impressive diversity of metamorphic phenotypes is constrained by the divergence of the TH signaling in the chordate groups. Interestingly metamorphosis is intrinsically linked to the environment, since it allows an animals to change its ecological niche. Alterations of the TH signaling pathway may then allow a species to adapt to a constantly changing environment (see chapter **??**). The fast changing environment may constrain a fast adapting TH signaling and a fast evolving metamorphosis.

An elegant example of the fast evolution of amphibian metamorphosis has recently been demonstrated by **?**, who used geographical data to clarify the apparently complex evolution of metamorphosis in marsupial frogs. Marsupial frogs are a group of amphibians in which females possess a pouch on their back, where

they carry their eggs. In some frogs, these eggs hatch directly into froglets, in others, they hatch into tadpoles. Direct-developping and biphasic frogs do not seem to constitute monophyletic groups, as shown by a phylogenetic tree based on the analysis of several loci. Notably, a group of biphasic frogs seems to be nested inside a clade of direct-developping frogs. Two hypotheses could explain this phylogenetic distribution: either (a) direct development has appeared independently many times, or (b) a tadpole phase has reappeared from direct-developing ancestors. This second hypothesis seems *a priori* unlikely, as many examples of transitions from biphasic to direct development are known, and almost no example shows the other kind of transition. The authors reconstructed the evolution of frog development along the phylogeny using two different probabilistic models of character evolution. A simple model of evolution provides support for hypothesis (b), but a more complex model supports hypothesis (a). This second model is significantly more likely, so this would appear the case is closed, if geographical data had nothing more to say. In fact, analysing the geographical distribution of marsupial frogs, the authors found that biphasic frogs live at a higher altitude than direct-developing ones. They then used a probabilistic model to infer the altitudes of various marsupial frog ancestors. They found that ancestors of the biphasic group of frogs lived at an altitude more consistent with a direct-developing lifestyle than with a biphasic lifestyle, in agreement with hypothesis (b).

Overall, this study finely illustrates the adaptability of metamorphosis to environmental changes. It would be interesting to study the evolution of TH signaling in the different above species, as differences of TH signaling between species may reflect the heterochrony. It was proposed that amphibian paedomorphy is an adaptation of amphibian post-embryonic development to specific environmental conditions, in which the aquatic niche is more stable than the terrestrial one. As Garstang wrote in a poem (**?**):

> "But when a lake's attractive, nicely aired, and full of food
> They climb to youth perpetual, and rear a tadpole brood[...]
> They think aquatic life is bliss, terrestrial is curse"

More prosaically, species that live in a favorable aquatic environment and whose larval stage is extended will be have a selective advantage compared to other species that do not benefit as much from the favorable environment. Larval prolongation may reach sexual maturity, causing animals to reproduce without metamorphosing (**?**). This sort of paedomorphy, called neoteny, may appear rapidly through alterations of endocrine regulation by the TH signaling pathway (see section **??**) and thus allows to give a rapid answer to specific environmental conditions.

11.3.2 *Modulation of TH availability*

The plasticity of the TH pathway allows evolution of the regulation of metamorphosis within the chordates, as proposed in the above example. The action of thyroactive compounds, mediated by TR, is dependent on the availability of the compound to each cell, as well as on the list of potential target genes. In amphibians and mammals, this control is observed at several levels: global TH level is regulated by the hypothalamo-pituitary axis while the local levels are tightly regulated by membrane transporters (**?**) and deiodinases (**?**). The situation is potentially similar in all chordates where some orthologous genes to TH transporters have been isolated (like the *C.intestinalis* orthologous gene to Oatp14, involved in transport of T_3 through the blood-brain barrier (**??**)). Of course, the expression pattern of TR is a central parameter, but the role of its partners (coactivators and corepressors) should not be neglected. A crucial aspect in the diversification of the metamorphosis pattern observed in vertebrates is the plasticity of the TH-regulated gene network. In Xenopus it has been shown that THs control a specific gene regulatory hierarchy through TR activation. It is likely that, as observed in obligate paedomorphic species, this gene network is modified in different species, resulting in a distinct outcome in terms of metamorphosis. This importance of the gene regulatory cascade controlled by a relatively invariant master regulator is reminiscent of the situation observed in other cases such as eye development (with pax6 being the master gene (**?**)) or antero-posterior development of arthropods (with *Hox* genes being the key selector genes (**?**) and more generally (**?**)). Indeed, TH-regulated gene expression pattern is highly variable from one species to another, as shown in *X. laevis* and the mouse (**??**). It should be interesting to compare gene expression in different vertebrate groups. One of the only common pathways is the TH-induced positive upregulation of one of the TR genes occurring during amphibian, fish or amphioxus metamorphosis (**??**) and mammalian weaning (TRα, (**??**)) in concert with a peak of TH production (**?**). The TR gene activated in these various groups is not the same though: TRβ is the key target in *X. laevis* whereas in actinopterygians and mammals TRα appears to be more important.

11.4 PERSPECTIVES FOR THE STUDY OF THE EVOLUTION OF CHORDATE METAMORPHOSIS

One should keep in mind that the model of metamorphosis evolution presented in this chapter is based on data accumulated in a few model species that represent a very restricted taxonomic sampling of vertebrates. Weaning in mammals,

metamorphosis in urochordates or larval to adult transitions in actinopterygian fishes (not to mention the post-embryonic development in sauropsids (that include birds, lizards, snakes and turtles, figure **??** in chapter **??**) are known from work on very few model systems. It will be very important in the future to study the role of THs and their derivatives and TR in the coordination of this neglected period of the vertebrate post-embryonic development. We nevertheless believe that the model proposed in this paper offers a coherent view of chordate post-embryonic development in general and of chordate metamorphosis in particular.

GENERAL CONCLUSIONS ABOUT THE EVOLUTION OF TR AND CHORDATE METAMORPHOSIS

During my PhD, I sought the role of the transcription factor TR in the evolution of metamorphosis in chordates. I would like to take the opportunity in this conclusion to place my work in the more general context of the role of developmental gene regulation in the evolution of phenotypes.

Evo-devo studies have focused on the role of transcription factors (TFs) in the evolution of phenotypes. As they are master regulators of early development, they are of particular interest in such investigations, as these proteins are key to the establishment of body axes and body shape. They are at the core of a gene regulatory network that governs early development, so that any alteration of their function should have massive consequences on development and on body shape (**?**). There has been an intense debate recently about which component of the developmental regulatory network are most important in evolution: are modifications of coding fractions of the genome (and more precisely in TF sequences) responsible for the differences between species? Are alterations in non-coding fractions of the genome responsible for the spatio-temporal expression of genes more prone to underlie evolution (e.g. by altering the TF expression pattern or the list of its target genes), while TFs have been kept fairly conserved throughout animal evolution? The discovery that most animals share a common set of developmental regulatory genes has drawn considerable attention on the cis-regulatory sequences as raw material for evolutionary changes (**?**). However, as illustrated on the figure **??**, the molecular mechanisms through which the evolution of TFs underlies the evolution of body shape is still sharply debated (e.g. (**?????**), see (**?**) for a recent summary).

More generally, the evolution of TFs is then to be considered in the context of the evolution of the genetic regulatory network (GRNs) TFs are involved in, and which govern development. It was proposed that alterations of GRNs underlie the evolution of morphology (**??**). Following this model, how may GRNs evolve?

Figure 41: Greg Wray (left) and Jerry Coyne promoted their take on the genetic basis of evolution with custom T-shirts at the the 2008 Oregon symposium on evo-devo. Picture taken by John Postlethwait and published in (**?**).

In agreement with the important role that cis-regulation may play in the evolution of developmental processes, **?** proposed an elegant model of the evolution of GRNs, according to which evolution of homologous developmental processes may depend on reorganization of corresponding genetic regulatory networks, with some features retained through time while others evolved faster (**??**). In other words, as **?** proposed,

> "*Regulatory change in the control systems of regulatory genes is the raw driver of developmental evolution*".

The authors discovered that the central core of the genetic regulatory network responsible for endomesoderm specification in echinoderms had been well conserved since echinoderm origins whereas more upstream and downstream parts diverged.

Our work is in line with this model of evolution of developmental processes through evolution of expression regulation. Indeed, we proposed that the main transcription factor regulating metamorphosis in chordate was conserved throughout chordate evolution for more than 500 million years, whereas the more upstream modules (TH metabolism and TR expression) and the more downstream modules (TR target genes repertoire) may evolve faster. It remains to be determined if this model will be confirmed by the analysis of the regulation of metamorphosis in other species, and if the evolution of regulatory networks controlling metamorphosis may be considered as representative of the evolution of GRNs in general.

A better understanding of how GRNs evolve will require genomic-scale studies in more species representing a great range of evolutionary divergences. This is now technically possible (**???**) and should provide solid bases to better understand mechanisms underlying the evolution of forms.

Part IV

BIBLIOGRAPHY

BIBLIOGRAPHY

Allen B. M. 1916. The results of extirpation of the anterior lobe of the hypophysis and of the thyoid of *Rana pipiens* larvae. *Science*, **44**(1143), 755–758.

Apriletti J. W., Ribeiro R. C., Wagner R. L., Feng W., Webb P., Kushner P. J., West B. L., Nilsson S., Scanlan T. S., Fletterick R. J., & Baxter J. D. 1998. Molecular and structural biology of thyroid hormone receptors. *Clin Exp Pharmacol Physiol Suppl*, **25**, S2–11.

Arendt D., & Nübler-Jung K. 1994. Inversion of dorsoventral axis? *Nature*, **371**(6492), 26.

Baker K. D., Shewchuk L. M., Kozlova T., Makishima M., Hassell A., Wisely B., Caravella J. A., Lambert M. H., Reinking J. L., Krause H., Thummel C. S., Willson T. M., & Mangelsdorf D. J. 2003. The *Drosophila* orphan nuclear receptor DHR38 mediates an atypical ecdysteroid signaling pathways. *Cell*, **113**(6), 731–42.

Baker M. E. 2005. Xenobiotics and the Evolution of Multicellular Animals: Emergence and Diversification of Ligand-Activated Transcription Factors. *Integr Comp Biol*, **45**(1), 172–178.

Baker M. E. 2008. *Trichoplax*, the simplest known animal, contains an estrogen-related receptor but no estrogen receptor: Implications for estrogen receptor evolution. *Biochem Biophys Res Commun*, **375**(4), 623–627.

Barnett P., Tabak H. F., & Hettema E. H. 2000. Nuclear receptors arose from pre-existing protein modules during evolution. *Trends Biochem Sci*, **25**(5), 227–228.

Barrington E. J., & Thorpe A. 1963. Comparative observations on iodine binding by *Saccoglossus horsti Brambell* and *Goodhart*, and by the tunic of *Ciona intestinalis (L). Gen Comp Endocrinol*, **3**, 166–75.

Barrington E. J., & Thorpe A. 1965. The Identification of Monoiodotyrosine, Diiodotyrosine and Thyroxine in Extracts of the Endostyle of the Ascidian, Ciona Intestinalis L. *Proc R Soc Lond B Biol Sci*, **163**, 136–49.

Barrington E. J., & Thorpe A. 1968. Histochemical and biochemical aspects of iodine binding in the tunic of the ascidian, *Dendrodoa grossularia (van Beneden)*. *Proc R Soc Lond B Biol Sci*, **171**(22), 91–109.

Bassham S., & Postlethwait J. H. 2005. The evolutionary history of placodes: a molecular genetic investigation of the larvacean urochordate Oikopleura dioica. *Development*, **132**(19), 4259–4272.

Bassham S., Cañestro C., & Postlethwait J. H. 2008. Evolution of developmental roles of Pax2/5/8 paralogs after independent duplication in urochordate and vertebrate lineages. *BMC Biol*, **6**, 35.

Bastien J., & Rochette-Egly C. 2004. Nuclear retinoid receptors and the transcription of retinoid-target genes. *Gene*, **328**(Mar), 1–16.

Baxter J. D., Goede P., Apriletti J. W., West B. L., Feng W., Mellstrom K., Fletterick R. J., Wagner R. L., Kushner P. J., Ribeiro R. C., Webb P., Scanlan T. S., & Nilsson S. 2002. Structure-based design and synthesis of a thyroid hormone receptor (TR) antagonist. *Endocrinology*, **143**(2), 517–24.

Benoit G., Cooney A., Giguere V., Ingraham H., Lazar M., Muscat G., Perlmann T., Renaud J. P., Schwabe J., Sladek F., Tsai M. J., & Laudet V. 2006. International Union of Pharmacology. LXVI. Orphan nuclear receptors. *Pharmacol Rev*, **58**(4), 798–836.

Benoit G., Malewicz M., & Perlmann T. 2004. Digging deep into the pockets of orphan nuclear receptors: insights from structural studies. *Trends Cell Biol*, **14**(7), 369–376.

Bertrand S., Brunet F. G., Escriva H., Parmentier G., Laudet V., & Robinson-Rechavi M. 2004. Evolutionary genomics of nuclear receptors: from twenty-five ancestral genes to derived endocrine systems. *Mol Biol Evol*, **21**(10), 1923–37.

Bishop C. D., & Brandhorst B. P. 2003. On nitric oxide signaling, metamorphosis, and the evolution of biphasic life cycles. *Evol Dev*, **5**(5), 542–50.

Bishop C. D., Erezyilmaz D. F., Flatt T., Georgiou C. D., Hadfield M. G., Heyland A., Hodin J., Jacobs M. W., Maslakova S. A., Pires A., Reitzel A. M., Santagata S., Tanaka K., & Youson J. H. 2006. What is metamorphosis? *Integr Comp Biol*, **46**(6), 655–661.

Blair J. E., & Hedges S. B. 2005. Molecular phylogeny and divergence times of deuterostome animals. *Mol Biol Evol*, **22**(11), 2275–84.

Boelen A., Wiersinga W. M., & Fliers E. 2008. Fasting-Induced Changes in the Hypothalamus-Pituitary-Thyroid Axis. *Thyroid*, **18**, 123–129.

Boeuf G., & Gaignon J. 1989. Effects of rearing conditions on growth and thyroid hormones during smolting of atlantic salmon, *Salmo salar L.* *Aquaculture*, **82**, 29–38.

Bone Q. 2001. Chordata (Chordates). *Encyclopedia of life science*.

Bonneton F., Zelus D., Iwema T., Robinson-Rechavi M., & Laudet V. 2003. Rapid divergence of the ecdysone receptor in Diptera and Lepidoptera suggests coevolution between ECR and USP-RXR. *Mol Biol Evol*, **20**(4), 541–53.

Bonneton F., Brunet F. G., Kathirithamby J., & Laudet V. 2006. The rapid divergence of the ecdysone receptor is a synapomorphy for Mecopterida that clarifies the Strepsiptera problem. *Insect Mol Biol*, **15**(3), 351–62.

Boorse G. C., & Denver R. J. 2002. Acceleration of Ambystoma tigrinum metamorphosis by corticotropin-releasing hormone. *J Exp Zool*, **293**(1), 94–8.

Boorse G. C., & Denver R. J. 2003. Endocrine Mechanisms Underlying Plasticity in Metamorphic Timing in Spadefoot Toads. *Integr Comp Biol*, **43**(5), 646–657.

Bourguet W., Ruff M., Chambon P., Gronemeyer H., & Moras D. 1995a. Crystal structure of the ligand-binding domain of the human nuclear receptor RXR-alpha. *Nature*, **375**(6530), 377–382.

Bourguet W., Ruff M., Bonnier D., Granger F., Boeglin M., Chambon P., Moras D., & Gronemeyer H. 1995b. Purification, functional characterization, and crystallization of the ligand binding domain of the retinoid X receptor. *Protein Expr Purif*, **6**(5), 604–608.

Bourguet W., Vivat V., Wurtz J. M., Chambon P., Gronemeyer H., & Moras D. 2000. Crystal structure of a heterodimeric complex of RAR and RXR ligand-binding domains. *Mol Cell*, **5**(2), 289–298.

Bourlat S. J., Juliusdottir T., Lowe C. J., Freeman R., Aronowicz J., Kirschner M., Lander E. S., Thorndyke M., Nakano H., Kohn A. B., Heyland A., Moroz L. L., Copley R. R., & Telford M. J. 2006. Deuterostome phylogeny reveals monophyletic chordates and the new phylum Xenoturbellida. *Nature*, **444**(7115), 85–8.

Bouton D., Escriva H., de Mendonca R. L., Glineur C., Bertin B., Noel C., Robinson-Rechavi M., de Groot A., Cornette J., Laudet V., & Pierce R. J. 2005. A conserved retinoid X receptor (RXR) from the mollusk *Biomphalaria glabrata* transactivates transcription in the presence of retinoidsbroo. *J Mol Endocrinol*, **34**(2), 567–82.

Braverman L. E., & Utiger R. D. 1996. *The thyroid. A Fundamental and Clinical Text*. Philaderphia, New York: Lippincott-Raven.

Bridgham J. T., Carroll S. M., & Thornton J. W. 2006. Evolution of hormone-receptor complexity by molecular exploitation. *Science*, **312**(5770), 97–101.

Bridgham J. T., Brown J. E., Rodríguez-Marí A., Catchen J. M., & Thornton J. W. 2008. Evolution of a new function by degenerative mutation in cephalochordate steroid receptors. *PLoS Genet*, **4**(9), e1000191.

Bringhurst R. 2002. *The Elements of Typographic Style*. Version 2.5. Point Roberts, WA, USA: Hartley & Marks, Publishers.

Brooks A. R., Sweeney G., & Old R. W. 1989. Structure and functional expression of a cloned *Xenopus* thyroid hormone receptor. *Nucleic Acids Res*, **17**(22), 9395–9405.

Brown D. D. 1997. The role of thyroid hormone in zebrafish and axolotl development. *Proc Natl Acad Sci U S A*, **94**(24), 13011–6.

Brown D. D. 2005. The role of deiodinases in amphibian metamorphosis. *Thyroid*, **15**(8), 815–21.

Brown D. D., & Cai L. 2007. Amphibian metamorphosis. *Dev Biol*, **306**(1), 20–33.

Buchholz D. R., Hsia S. C., Fu L., & Shi Y. B. 2003. A dominant-negative thyroid hormone receptor blocks amphibian metamorphosis by retaining corepressors at target genes. *Mol Cell Biol*, **23**(19), 6750–8.

Buchholz D. R., Paul B. D., Fu L., & Shi Y. B. 2006. Molecular and developmental analyses of thyroid hormone receptor function in *Xenopus laevis*, the African clawed frogs. *Gen Comp Endocrinol*, **145**(1), 1–19.

Butler A. B., & Saidel W. M. 2000. Defining sameness: historical, biological, and generative homology. *BioEssays*, **22**(9), 846–853.

Cai L., & Brown D. D. 2004. Expression of type II iodothyronine deiodinase marks the time that a tissue responds to thyroid hormone-induced metamorphosis in Xenopus laevis. *Dev Biol*, **266**(1), 87–95.

Callery E. M., & Elinson R. P. 2000a. Opercular development and ontogenetic re-organization in a direct-developing frog. *Dev Genes Evol*, **210**(7), 377–81.

Callery E. M., & Elinson R. P. 2000b. Thyroid hormone-dependent metamorphosis in a direct developing frog. *Proc Natl Acad Sci U S A*, **97**(6), 2615–20.

Canestro C., Albalat R., Hjelmqvist L., Godoy L., Jornvall H., & Gonzalez-Duarte R. 2002. Ascidian and amphioxus *Adh* genes correlate functional and molecular features of the ADH family expansion during vertebrate evolution. *J Mol Evol*, **54**(1), 81–9.

Carosa E., Fanelli A., Ulisse S., Di Lauro R., Rall J. E., & Jannini E. A. 1998. Ciona intestinalis nuclear receptor 1: a member of steroid/thyroid hormone receptor family. *Proc Natl Acad Sci U S A*, **95**(19), 11152–7.

Carroll S. B. 2005. Evolution at two levels: on genes and form. *PLoS Biol*, **3**(7), e245.

Carroll S. B. 2008. Evo-devo and an expanding evolutionary synthesis: a genetic theory of morphological evolution. *Cell*, **134**(1), 25–36.

Carroll S. B., Grenier J. K., & Weatherbee S. D. 2005. *From DNA to diversity : molecular genetics and the evolution of animal design*. 2nd edn. Malden, MA: Blackwell.

Carroll S. M., Bridgham J. T., & Thornton J. W. 2008. Evolution of hormone signaling in elasmobranchs by exploitation of promiscuous receptors. *Mol Biol Evol*, Sep.

Chambon J. P., Soule J., Pomies P., Fort P., Sahuquet A., Alexandre D., Mangeat P. H., & Baghdiguian S. 2002. Tail regression in *Ciona intestinalis* (Prochordate) involves a Caspase-dependent apoptosis event associated with ERK activation. *Development*, **129**(13), 3105–14.

Chambon J. P., Nakayama A., Takamura K., McDougall A., & Satoh N. 2007. ERK- and JNK-signalling regulate gene networks that stimulate metamorphosis and apoptosis in tail tissues of ascidian tadpoles. *Development*, **134**(6), 1203–19.

Chandra V., Huang P., Hamuro Y., Raghuram S., Wang Y., Burris T. P., & Rastinejad F. 2008. Structure of the intact PPAR-gamma-RXR-alpha nuclear receptor complex on DNA. *Nature*, Oct.

Chen J. D., & Evans R. M. 1995. A transcriptional co-repressor that interacts with nuclear hormone receptors. *Nature*, **377**(6548), 454–457.

Chen J.-Y., Huang D.-Y., Peng Q.-Q., Chi H.-M., Wang X.-Q., & Feng M. 2003. The first tunicate from the Early Cambrian of South China. *Proc Natl Acad Sci U S A*, **100**(14), 8314–8318.

Chiellini G., Apriletti J. W., Yoshihara H. A., Baxter J. D., Ribeiro R. C., & Scanlan T. S. 1998. A high-affinity subtype-selective agonist ligand for the thyroid hormone receptor. *Chem Biol*, **5**(6), 299–306.

Christiaen L., Jaszczyszyn Y., Kerfant M., Kano S., Thermes V., & Joly J.-S. 2007. Evolutionary modification of mouth position in deuterostomes. *Semin Cell Dev Biol*, **18**(4), 502–511.

Clarke N. D., & Berg J. M. 1998. Zinc fingers in *Caenorhabditis elegans*: finding families and probing pathways. *Science*, **282**(5396), 2018–2022.

Covelli I., Salvatore G., Sena L., & Roche J. 1960. Sur la formation d'hormones thyroidiennes et de leurs précurseurs par *Branchiostoma lanceolatum*. *C R Soc Biol Paris*, **154**, 1165–9.

Cowen R. K., Paris C. B., & Srinivasan A. 2006. Scaling of Connectivity in Marine Populations. *Science*, **311**(5760), 522–527.

Cretekos C. J., Wang Y., Green E. D., Martin J. F., Rasweiler J. J., & Behringer R. R. 2008. Regulatory divergence modifies limb length between mammals. *Genes Dev*, **22**(2), 141–151.

Darwin C. 1859. *On the origin of species by means of natural selection, or the preservation of favoured races in the struggle for life.* London: John Murray.

Das B., Cai L., Carter M. G., Piao Y. L., Sharov A. A., Ko M. S., & Brown D. D. 2006. Gene expression changes at metamorphosis induced by thyroid hormone in *Xenopus laevis* tadpoles. *Dev Biol*, **291**(2), 342–55.

Davidson E. H., & Erwin D. H. 2006. Gene regulatory networks and the evolution of animal body plans. *Science*, **311**(5762), 796–800.

Davis P. J., Davis F. B., & Cody V. 2005. Membrane receptors mediating thyroid hormone action. *Trends Endocrinol Metab*, **16**(9), 429–435.

Dawes R., Dawson I., Falciani F., Tear G., & Akam M. 1994. *Dax*, a locust *Hox* gene related to *fushi-tarazu* but showing no pair-rule expression. *Development*, **120**(6), 1561–1572.

de Jesus E. G. T., Toledo J. D., & Simpas M. S. 1998. Thyroid Hormones Promote Early Metamorphosis in Grouper (*Epinephelus coioides*) Larvae. *Gen Comp Endocrinol*, **112**(1), 10–16.

Debuire B., Henry C., Bernissa M., Biserte G., Claverie J. M., Saule S., Martin P., & Stehelin D. 1984. Sequencing the *erbA* gene of avian erythroblastosis virus reveals a new type of oncogene. *Science*, **224**(4656), 1456–1459.

Dehal P., & Boore J. L. 2005. Two rounds of whole genome duplication in the ancestral vertebrate. *PLoS Biol*, **3**(10), e314.

Dehal P., Satou Y., Campbell R. K., Chapman J., Degnan B., De Tomaso A., Davidson B., Di Gregorio A., Gelpke M., Goodstein D. M., Harafuji N., Hastings K. E., Ho I., Hotta K., Huang W., Kawashima T., Lemaire P., Martinez D., Meinertzhagen I. A., Necula S., Nonaka M., Putnam N., Rash S., Saiga H., Satake M., Terry A., Yamada L., Wang H. G., Awazu S., Azumi K., Boore J., Branno M., Chin-Bow S., DeSantis R., Doyle S., Francino P., Keys D. N., Haga S., Hayashi H., Hino K., Imai K. S., Inaba K., Kano S., Kobayashi K., Kobayashi M., Lee B. I., Makabe K. W., Manohar C., Matassi G., Medina M., Mochizuki Y., Mount S., Morishita T., Miura S., Nakayama A., Nishizaka S., Nomoto H., Ohta F., Oishi K., Rigoutsos I., Sano M., Sasaki A., Sasakura Y., Shoguchi E., Shin-i T., Spagnuolo A., Stainier D., Suzuki M. M., Tassy O., Takatori N., Tokuoka M., Yagi K., Yoshizaki F., Wada S., Zhang C., Hyatt P. D., Larimer F., Detter C., Doggett N., Glavina T., Hawkins T., Richardson P., Lucas S., Kohara Y., Levine M., Satoh N., & Rokhsar D. S. 2002. The draft genome of *Ciona intestinalis*: insights into chordate and vertebrate origins. *Science*, **298**(5601), 2157–67.

Delsuc F., Brinkmann H., Chourrout D., & Philippe H. 2006. Tunicates and not cephalochordates are the closest living relatives of vertebrates. *Nature*, **439**(7079), 965–8.

Delsuc F., Tsagkogeorga G., Lartillot N., & Philippe H. 2008. Additional molecular support for the new chordate phylogeny. *Genesis*, **46**(11), 592–604.

Denver R. J. 1997. Environmental stress as a developmental cue: corticotropin-releasing hormone is a proximate mediator of adaptive phenotypic plasticity in amphibian metamorphosis. *Horm Behav*, **31**(2), 169–79.

Denver R. J. 1998. The molecular basis of thyroid hormone-dependent central nervous system remodeling during amphibian metamorphosis. *Comp Biochem Physiol C Pharmacol Toxicol Endocrinol*, **119**(3), 219–28.

Denver R. J. 2008. Chordate metamorphosis: ancient control by iodothyronines. *Curr Biol*, **18**(13), R567–9.

Desvergne B. 1994. How do thyroid hormone receptors bind to structurally diverse response elements? *Mol Cell Endocrinol*, **100**(1-2), 125–131.

Dobzhansky T. 1964. Biology, Molecular And Organismic. *Am Zool*, **4**(Nov.), 443–52.

Dominguez P., Jacobson A. G., & Jefferies R. P. S. 2002. Paired gill slits in a fossil with a calcite skeleton. *Nature*, **417**(6891), 841–844.

Donoghue P. C. J., Graham A., & Kelsh R. N. 2008. The origin and evolution of the neural crest. *BioEssays*, **30**(6), 530–541.

Dubois A., & Drogi P. 2008. *Métamorphose*.

Eales J. G. 1997. Iodine metabolism and thyroid-related functions in organisms lacking thyroid follicles: are thyroid hormones also vitamins? *Proc Soc Exp Biol Med*, **214**(4), 302–17.

Egea P. F., Mitschler A., Rochel N., Ruff M., Chambon P., & Moras D. 2000. Crystal structure of the human RXRalpha ligand-binding domain bound to its natural ligand: 9-cis retinoic acid. *EMBO J*, **19**(11), 2592–2601.

Escriva H., Safi R., Hanni C., Langlois M. C., Saumitou-Laprade P., Stehelin D., Capron A., Pierce R., & Laudet V. 1997. Ligand binding was acquired during evolution of nuclear receptors. *Proc Natl Acad Sci U S A*, **94**(13), 6803–8.

Escriva H., Delaunay F., & Laudet V. 2000. Ligand binding and nuclear receptor evolution. *BioEssays*, **22**(8), 717–27.

Escriva H., Manzon L., Youson J., & Laudet V. 2002a. Analysis of lamprey and hagfish genes reveals a complex history of gene duplications during early vertebrate evolution. *Mol Biol Evol*, **19**(9), 1440–50.

Escriva H., Holland N. D., Gronemeyer H., Laudet V., & Holland L. Z. 2002b. The retinoic acid signaling pathway regulates anterior/posterior patterning in the nerve cord and pharynx of amphioxus, a chordate lacking neural crest. *Development*, **129**(12), 2905–16.

Escriva H., Bertrand S., Germain P., Robinson-Rechavi M., Umbhauer M., Cartry J., Duffraisse M., Holland L., Gronemeyer H., & Laudet V. 2006. Neofunctionalization in vertebrates: the example of retinoic acid receptors. *PLoS Genet*, **2**(7), e102.

Everts M. E., Verhoeven F. A., Bezstarosti K., Moerings E. P., Hennemann G., Visser T. J., & Lamers J. M. 1996. Uptake of thyroid hormones in neonatal rat cardiac myocytes. *Endocrinology*, **137**(10), 4235–42.

Felsenstein J. 1978. Cases in which parsimony or compatibility methods will be positively misleading. *Syst zool*, **27**, 401–10.

Flamant F., & Samarut J. 2003. Thyroid hormone receptors: lessons from knockout and knock-in mutant mice. *Trends Endocrinol Metab*, **14**(2), 85–90.

Flamant F., Poguet A. L., Plateroti M., Chassande O., Gauthier K., Streichenberger N., Mansouri A., & Samarut J. 2002. Congenital hypothyroid *Pax8*(-/-) mutant mice can be rescued by inactivating the *TRalpha* genes. *Mol Endocrinol*, **16**(1), 24–32.

Flamant F., Baxter J. D., Forrest D., Refetoff S., Samuels H., Scanlan T. S., Vennstrom B., & Samarut J. 2006. International Union of Pharmacology. LIX. The pharmacology and classification of the nuclear receptor superfamily: thyroid hormone receptors. *Pharmacol Rev*, **58**(4), 705–11.

Force A., Lynch M., Pickett F. B., Amores A., Yan Y. L., & Postlethwait J. 1999. Preservation of duplicate genes by complementary, degenerative mutations. *Genetics*, **151**(4), 1531–45.

Forrest D., Erway L. C., Ng L., Altschuler R., & Curran T. 1996. Thyroid hormone receptor beta is essential for development of auditory function. *Nat Genet*, **13**(3), 354–357.

Fraichard A., Chassande O., Plateroti M., Roux J. P., Trouillas J., Dehay C., Legrand C., Gauthier K., Kedinger M., Malaval L., Rousset B., & Samarut J. 1997. The T3R alpha gene encoding a thyroid hormone receptor is essential for post-natal development and thyroid hormone production. *EMBO J*, **16**(14), 4412–20.

Fredriksson G. 1988. *Thyroid-like systems in endostyles: a study of morphology, function and evolution in "primitive" chordates.* Ph.D. thesis, University of Stockholm.

Fredriksson G., Ofverholm T., & Ericson L. E. 1985. Electron-microscopic studies of iodine-binding and peroxidase activity in the endostyle of the larval amphioxus (*Branchiostoma lanceolatum*). *Cell Tiss Res*, **241**(2), 257–266.

Friesema E. C., Jansen J., Heuer H., Trajkovic M., Bauer K., & Visser T. J. 2006. Mechanisms of disease: psychomotor retardation and high T3 levels caused by mutations in monocarboxylate transporter 8. *Nat Clin Pract Endocrinol Metab*, **2**(9), 512–23.

Fuentes M., Schubert M., Dalfo D., Candiani S., Benito E., Gardenyes J., Godoy L., Moret F., Illas M., Patten I., Permanyer J., Oliveri D., Boeuf G., Falcon J., Pestarino M., Fernandez J. G., Albalat R., Laudet V., Vernier P., & Escriva H. 2004. Preliminary observations on the spawning conditions of the European amphioxus (*Branchiostoma lanceolatum*) in captivity. *J Exp Zool B Mol Dev Evol*, **302**(4), 384–91.

Fuentes M., Benito E., Bertrand S., Paris M., Mignardot A., Godoy L., Jimenez-Delgado S., Oliveri D., Candiani S., Hirsinger E., D'Aniello S., Pascual-Anaya

J., Maeso I., Pestarino M., Vernier P., Nicolas J. F., Schubert M., Laudet V., Geneviere A. M., Albalat R., Garcia Fernandez J., Holland N. D., & Escriva H. 2007. Insights into spawning behavior and development of the european amphioxus (*Branchiostoma lanceolatum*). *J Exp Zool B Mol Dev Evol*.

Fukazawa H., Hirai H., Hori H., Roberts R. D., Nukaya H., Ishida I., & Tsuji K. 2001. Induction of abalone larval metamorphosis by thyroid hormones. *Fisheries Sciences*, **67**, 985–988.

Furlow J. D., & Brown D. D. 1999. *In vitro* and *in vivo* analysis of the regulation of a transcription factor gene by thyroid hormone during *Xenopus laevis* metamorphosis. *Mol Endocrinol*, **13**(12), 2076–89.

Furlow J. D., & Kanamori A. 2002. The transcription factor basic transcription element-binding protein 1 is a direct thyroid hormone response gene in the frog *Xenopus laevis*. *Endocrinology*, **143**(9), 3295–3305.

Gans C., & Northcutt R. G. 1983. Neural Crest and the Origin of Vertebrates: A New Head. *Science*, **220**(4594), 268–273.

Garstang W. 1951. *Larval forms and other zoological verses.* Chicago Press.

Gauthier K., Chassande O., Plateroti M., Roux J. P., Legrand C., Pain B., Rousset B., Weiss R., Trouillas J., & Samarut J. 1999. Different functions for the thyroid hormone receptors TRalpha and TRbeta in the control of thyroid hormone production and post-natal development. *EMBO J*, **18**(3), 623–31.

Gee H. 1996. *Before the backbone : views on the origin of the vertebrates*. London: Chapman & Hall.

Germain P., Staels B., Dacquet C., Spedding M., & Laudet V. 2006a. Overview of nomenclature of nuclear receptors. *Pharmacol Rev*, **58**(4), 685–704.

Germain P., Chambon P., Eichele G., Evans R. M., Lazar M. A., Leid M., Lera A. R. D., Lotan R., Mangelsdorf D. J., & Gronemeyer H. 2006b. International Union of Pharmacology. LXIII. Retinoid X receptors. *Pharmacol Rev*, **58**(4), 760–772.

Gilbert L. I., Tata J. R., & Atkinson B. G. 1996. *Metamorphosis : postembryonic reprogramming of gene expression in amphibian and insect cells*. Cell biology. San Diego, Calif.: Academic Press.

Glass C. K. 1994. Differential recognition of target genes by nuclear receptor monomers, dimers, and heterodimers. *Endocr Rev*, **15**(3), 391–407.

Gould S. J. 1977. *Ontogeny and phylogeny*.

Gould S. J. 1990. *Wonderful Life: The Burgess Shale and the Nature of History*. WW Norton & Company, London.

Grasse P.-P. 1948. *Traité de zoologie. Tome XI: Échinodermes-Stomodés-Prochordés*. Libraires de l'académie de médecine.

Gronemeyer H., Gustafsson J. A., & Laudet V. 2004. Principles for modulation of the nuclear receptor superfamily. *Nat Rev Drug Discov*, **3**(11), 950–64.

Gross J., & Pitt-Rivers R. 1953a. 3:5:3,l-Triiodothyronine. 1. Isolation from thyroid gland and synthesis. *J Biochem*, **53**, 645–652.

Gross J., & Pitt-Rivers R. 1953b. 3:5:3,l-Triiodothyronine. 2. Physiological activity. *J Biochem*, **53**, 652–657.

Gudernatsch J. F. 1912. Feeding Experiments on Tadpoles. *Wilhem Roux Arch Entwicklungsmech Org*, **35**, 457–84.

Guindon S., & Gascuel O. 2003. A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. *Syst Biol*, **52**(5), 696–704.

Hadfield M. G. 2000. Why and how marine-invertebrate larvae metamorphose so fast. *Semin Cell Dev Biol*, **11**(6), 437–43.

Hadj-Sahraoui N., Seugnet I., Ghorbel M. T., & Demeneix B. 2000. Hypothyroidism prolongs mitotic activity in the post-natal mouse brain. *Neuroscience Letters*, **280**(2), 79–82.

Hashiguchi Y., & Nishida M. 2007. Evolution of trace amine associated receptor (TAAR) gene family in vertebrates: lineage-specific expansions and degradations of a second class of vertebrate chemosensory receptors expressed in the olfactory epithelium. *Mol Biol Evol*, **24**(9), 2099–2107.

Hauenschild A., Ringrose L., Altmutter C., Paro R., & Rehmsmeier M. 2008. Evolutionary Plasticity of Polycomb/*Trithorax* Response Elements in *Drosophila* Species. *PLoS Biology*, **6**(10), e261.

Havis E., Le Mevel S., Morvan Dubois G., Shi D. L., Scanlan T. S., Demeneix B. A., & Sachs L. M. 2006. Unliganded thyroid hormone receptor is essential for *Xenopus laevis* eye developments. *EMBO J*, **25**(20), 4943–51.

Hennemann G. 1986. *Thyroid hormone metabolism*. Marcel Dekker, Inc.

Heyland A., Hodin J., & Reitzel A. M. 2005. Hormone signaling in evolution and development: a non-model system approach. *BioEssays*, **27**(1), 64–75.

Heyland A., Reitzel A. M., Price D. A., & Moroz L. L. 2006a. Endogenous thyroid hormone synthesis in facultative planktotrophic larvae of the sand dollar *Clypeaster rosaceus*: implications for the evolutionary loss of larval feeding. *Evol Dev*, **8**(6), 568–79.

Heyland A., Price D. A., Bodnarova-Buganova M., & Moroz L. L. 2006b. Thyroid hormone metabolism and peroxidase function in two non-chordate animals. *J Exp Zool B Mol Dev Evol*, **306**(6), 551–66.

Hinchliffe J. R. 2002. Developmental basis of limb evolution. *Int J Dev Biol*, **46**(7), 835–845.

Hinman V. F., & Davidson E. H. 2007. Evolutionary plasticity of developmental gene regulatory network architecture. *Proc Natl Acad Sci U S A*, **104**(49), 19404–19409.

Hinman V. F., Nguyen A. T., Cameron R. A., & Davidson E. H. 2003. Developmental gene regulatory network architecture across 500 million years of echinoderm evolution. *Proc Natl Acad Sci U S A*, **100**(23), 13356–13361.

Hiroi Y., Kim H.-H., Ying H., Furuya F., Huang Z., Simoncini T., Noma K., Ueki K., Nguyen N.-H., Scanlan T. S., Moskowitz M. A., Cheng S.-Y., & Liao J. K. 2006. Rapid nongenomic actions of thyroid hormone. *Proc Natl Acad Sci*, **103**(38), 14104–14109.

Hodin J. 2006. Expanding networks: Signaling components in and a hypothesis for the evolution of metamorphosis. *Integr Comp Biol*, **46**(6), 719–742.

Hoekstra H. E., & Coyne J. A. 2007. The locus of evolution: evo devo and the genetics of adaptation. *Evolution*, **61**(5), 995–1016.

Holland L. Z., & Gibson-Brown J. J. 2003. The *Ciona intestinalis* genome: when the constraints are off. *BioEssays*, **25**(6), 529–32.

Holland L. Z., & Holland N. D. 1998. Developmental gene expression in amphioxus: New insights into the evolutionary origin of vertebrate brain regions, neural crest, and rostrocaudal segmentation. *Am Zool*, **38**(4), 647–658.

Holland L. Z., & Holland N. D. 2001. Evolution of neural crest and placodes: amphioxus as a model for the ancestral vertebrate? *J Anat*, **199**(Pt 1-2), 85–98.

Holland L. Z., Albalat R., Azumi K., Benito-Gutierrez E., Blow M. J., Bronner-Fraser M., Brunet F., Butts T., Candiani S., Dishaw L. J., Ferrier D. E., Garcia-Fernandez J., Gibson-Brown J. J., Gissi C., Godzik A., Hallbook F., Hirose D., Hosomichi K., Ikuta T., Inoko H., Kasahara M., Kasamatsu J., Kawashima T., Kimura A., Kobayashi M., Kozmik Z., Kubokawa K., Laudet V., Litman G. W., McHardy

A. C., Meulemans D., Nonaka M., Olinski R. P., Pancer Z., Pennacchio L. A., Pestarino M., Rast J. P., Rigoutsos I., Robinson-Rechavi M., Roch G., Saiga H., Sasakura Y., Satake M., Satou Y., Schubert M., Sherwood N., Shiina T., Takatori N., Tello J., Vopalensky P., Wada S., Xu A., Ye Y., Yoshida K., Yoshizaki F., Yu J. K., Zhang Q., Zmasek C. M., de Jong P. J., Osoegawa K., Putnam N. H., Rokhsar D. S., Satoh N., & Holland P. W. 2008. The amphioxus genome illuminates vertebrate origins and cephalochordate biology. *Genome Res*, **18**(7), 1100–11.

Holland N. D. 2005. Chordates. *Curr Biol*, **15**(22), R911–R914.

Holland P. W., Holland L. Z., Williams N. A., & Holland N. D. 1992. An amphioxus homeobox gene: sequence conservation, spatial expression during development and insights into vertebrate evolution. *Development*, **116**(3), 653–61.

Horard B., & Vanacker J.-M. 2003. Estrogen receptor-related receptors: orphan receptors desperately seeking a ligand. *J Mol Endocrinol*, **31**(3), 349–357.

Howard-Ashby M., Materna S. C., Brown C. T., Chen L., Cameron R. A., & Davidson E. H. 2006. Gene families encoding transcription factors expressed in early development of *Strongylocentrotus purpuratus*. *Dev Biol*, **300**(1), 90–107.

Huang E. Y., Zhang J., Miska E. A., Guenther M. G., Kouzarides T., & Lazar M. A. 2000. Nuclear receptor corepressors partner with class II histone deacetylases in a Sin3-independent repression pathway. *Genes Dev*, **14**(1), 45–54.

Hulbert A. J. 2000. Thyroid hormones and their effects: a new perspective. *Biol Rev Camb Philos Soc*, **75**(4), 519–631.

Hörlein A. J., Näär A. M., Heinzel T., Torchia J., Gloss B., Kurokawa R., Ryan A., Kamei Y., Söderström M., & Glass C. K. 1995. Ligand-independent repression by the thyroid hormone receptor mediated by a nuclear receptor co-repressor. *Nature*, **377**(6548), 397–404.

Inui Y., & Miwa S. 1985. Thyroid hormone induces metamorphosis of flounder larvae. *Gen Comp Endocrinol*, **60**(3), 450–4.

Ito Y., Okada R., Mochida H., Hayashi H., Yamamoto K., & Kikuyama S. 2004. Molecular cloning of bullfrog corticotropin-releasing factor (CRF): effect of homologous CRF on the release of TSH from pituitary cells in vitro. *Gen Comp Endocrinol*, **138**(3), 218–27.

Iwema T., Billas I. M., Beck Y., Bonneton F., Nierengarten H., Chaumot A., Richards G., Laudet V., & Moras D. 2007. Structural and functional characterization of a novel type of ligand-independent RXR-USP receptor. *EMBO J*, **26**(16), 3770–82.

Jacob F. 1977. Evolution and tinkering. *Science*, **196**(4295), 1161–1166.

Janvier P. 1996. *Early Vertebratess*. Oxford Science Publications.

Janvier P. 2003. Vertebrate characters and the Cambrian vertebrates. *CRPaléontol*, **2**, 523–531.

Jefferies R. 1971. Some comments on the origin of chordates. *J Paleontol*, **45**, 910–912.

Jeffery W. R., Swalla B. J., Ewing N., & Kusakabe T. 1999. Evolution of the ascidian anural larva: evidence from embryos and molecules. *Mol Biol Evol*, **16**(5), 646–54.

Jeffery W. R., Strickler A. G., & Yamamoto Y. 2004. Migratory neural crest-like cells form body pigmentation in a urochordate embryo. *Nature*, **431**(7009), 696–699.

Jenner R. A., & Wills M. A. 2007. The choice of model organisms in evo-devo. *Nat Rev Genet*, **8**(4), 311–319.

Jeyakumar M., Webb P., Baxter J. D., Scanlan T. S., & Katzenellenbogen J. A. 2008. Quantification of ligand-regulated nuclear receptor corepressor and coactivator binding, key interactions determining ligand potency and efficacy for the thyroid hormone receptor. *Biochemistry*, **47**(28), 7465–7476.

Jorgensen J. M. 1998. *The biology of hagfishes*. 1st edn. London ; New York: Chapman & Hall.

Joss J. M. 2006. Lungfish evolution and development. *Gen Comp Endocrinol*, **148**(3), 285–9.

Kallen J. A., Schlaeppi J.-M., Bitsch F., Geisse S., Geiser M., Delhon I., & Fournier B. 2002. X-ray structure of the hRORalpha LBD at 1.63 A: structural and functional data that cholesterol or a cholesterol derivative is the natural ligand of RORalpha. *Structure*, **10**(12), 1697–1707.

Kao H. Y., Ordentlich P., Koyano-Nakagawa N., Tang Z., Downes M., Kintner C. R., Evans R. M., & Kadesch T. 1998. A histone deacetylase corepressor complex regulates the Notch signal transduction pathway. *Genes Dev*, **12**(15), 2269–2277.

Kashian D. R., & Dodson S. 2004. Effects of vertebrate hormones on development and sew determination in Daphnia Magna. *Env Toxicol Chem*, **23**, 1282–1288.

Kawahara A., Baker B. S., & Tata J. R. 1991. Developmental and regional expression of thyroid hormone receptor genes during *Xenopus* metamorphosis. *Development*, **112**(4), 933–43.

Keay J., Bridgham J. T., & Thornton J. W. 2006. The *Octopus vulgaris* estrogen receptor is a constitutive transcriptional activator: evolutionary and functional implications. *Endocrinology*, **147**(8), 3861–9.

Kendall E. 1919. Isolation of the iodine compound which occurs in the thyroid. *J Biol chem*, **39**, 125–147.

Kluge B., Renault N., & Rohr K. B. 2005. Anatomical and molecular reinvestigation of lamprey endostyle development provides new insight into thyroid gland evolution. *Dev Genes Evol*, **215**(1), 32–40.

Kobayashi H., Tsuneki K., Akiyoshi H., Kobayashi Y., Nozaki M., & Ouji M. 1983. Histochemical distribution of peroxidase in ascidians with special reference to the endostyle and the branchial sac. *Gen Comp Endocrinol*, **50**(2), 172–87.

Koehler K., Gordon S., Brandt P., Carlsson B., Bäcksbro-Saeidi A., Apelqvist T., Agback P., Grover G. J., Nelson W., Grynfarb M., Färnegårdh M., Rehnmark S., & Malm J. 2006. Thyroid receptor ligands. 6. A high affinity "direct antagonist" selective for the thyroid hormone receptor. *J Med Chem*, **49**(23), 6635–6637.

Koehrle J., Auf'mkolk M., Rokos H., Hesch R. D., & Cody V. 1986. Rat liver iodothyronine monodeiodinase. Evaluation of the iodothyronine ligand-binding site. *J Biol Chem*, **261**(25), 11613–11622.

Kon T., Nohara M., Yamanoue Y., Fujiwara Y., Nishida M., & Nishikawa T. 2007. Phylogenetic position of a whale-fall lancelet (Cephalochordata) inferred from whole mitochondrial genome sequences. *BMC Evol Biol*, **7**, 127.

Kostrouch Z., Kostrouchova M., Love W., Jannini E., Piatigorsky J., & Rall J. E. 1998. Retinoic acid X receptor in the diploblast, *Tripedalia cystophora*. *Proc Natl Acad Sci U S A*, **95**(23), 13442–7.

Kozmik Z. 2008. The role of *Pax* genes in eye evolution. *Brain Res Bull*, **75**(2-4), 335–9.

Kroiher M. 2000. Morphological chimeras of larvae and adults in a hydrozoan–insights into the control of pattern formation and morphogenesis. *Int J Dev Biol*, **44**(8), 861–866.

Krylova I. N., Sablin E. P., Moore J., Xu R. X., Waitt G. M., MacKay J. A., Juzumiene D., Bynum J. M., Madauss K., Montana V., Lebedeva L., Suzawa M., Williams J. D., Williams S. P., Guy R. K., Thornton J. W., Fletterick R. J., Willson T. M., & Ingraham H. A. 2005. Structural analyses reveal phosphatidyl inositols as ligands for the NR5 orphan receptors SF-1 and LRH-1. *Cell*, **120**(3), 343–55.

Kuraku S., Meyer A., & Kuratani S. 2008. Timing of genome duplications relative to the origin of the vertebrates: did cyclostomes diverge before, or after? *Mol Biol Evol*, Oct.

Lartillot N., & Philippe H. 2008. Improvement of molecular phylogenetic inference and the phylogeny of Bilateria. *Philos Trans R Soc Lond B Biol Sci*, **363**(1496), 1463–1472.

Laudet V. 1997. Evolution of the nuclear receptor superfamily: early diversification from an ancestral orphan receptor. *J Mol Endocrinol*, **19**(3), 207–26.

Laudet V., & Gronemeyer H. 2002. *The Nuclear Receptor FacsBook*. London: Academic Press.

Lee Y., & Mahdavi V. 1993. The D domain of the thyroid hormone receptor alpha 1 specifies positive and negative transcriptional regulation functions. *J Biol Chem*, **268**(3), 2021–2028.

Lee Y.-H., Coonrod S. A., Kraus W. L., Jelinek M. A., & Stallcup M. R. 2005. Regulation of coactivator complex assembly and function by protein arginine methylation and demethylimination. *Proc Natl Acad Sci U S A*, **102**(10), 3611–3616.

Leloup J., & Buscaglia M. 1977. La triiodothyronine: hormone de la metamorphose des amphibians. *Comptes rendus de l'académie des sciences*, **284**, 2261–2263.

Leloup J., & Lebel J. M. 1993. Triiodothyroine is necessary for the action of growth hormone in acclimatation to seawater of brown (Salmo trutta) and rainbow trout (*Oncorhynchus mykiss*). *Fish Physiology and Biochemistry*, **11**, 165–173.

Levine M., & Tjian R. 2003. Transcription regulation and animal diversity. *Nature*, **424**(6945), 147–51.

Li X. Y., MacArthur S., Bourgon R., Nix D., Pollard D. A., Iyer V. N., Hechmer A., Simirenko L., Stapleton M., Luengo Hendriks C. L., Chu H. C., Ogawa N., Inwood W., Sementchenko V., Beaton A., Weiszmann R., Celniker S. E., Knowles D. W., Gingeras T., Speed T. P., Eisen M. B., & Biggin M. D. 2008. Transcription factors bind thousands of active and inactive regions in the *Drosophila* blastoderm. *PLoS Biol*, **6**(2), e27.

Lim W., Nguyen N. H., Yang H. Y., Scanlan T. S., & Furlow J. D. 2002. A thyroid hormone antagonist that inhibits thyroid hormone action in vivo. *J Biol Chem*, **277**(38), 35664–70.

Lin C.-Y., Vega V. B., Thomsen J. S., Zhang T., Kong S. L., Xie M., Chiu K. P., Lipovich L., Barnett D. H., Stossi F., Yeo A., George J., Kuznetsov V. A., Lee Y. K., Charn T. H., Palanisamy N., Miller L. D., Cheung E., Katzenellenbogen B. S., Ruan Y., Bourque G., Wei C.-L., & Liu E. T. 2007. Whole-genome cartography of estrogen receptor alpha binding sites. *PLoS Genet*, **3**(6), e87.

Lintlop S. P., & Youson J. H. 1983. Binding of triiodothyronine to hepatocyte nuclei from sea lampreys, *Petromyzon marinus L.*, at various stages of the life cycle. *Gen Comp Endocrinol*, **49**(3), 428–36.

Lonard D. M., & O'Malley B. W. 2006. The expanding cosmos of nuclear receptor coactivators. *Cell*, **125**(3), 411–414.

Löhr U., Yussa M., & Pick L. 2001. Drosophila fushi tarazu. a gene on the border of homeotic function. *Curr Biol*, **11**(18), 1403–1412.

Maddison W., & Maddison D. 2008. *Mesquite: a modular system for evolutionary analysis, v.2.5*.

Mallatt J., Chen J., & Holland N. D. 2003. Comment on "A new species of *yunnanozoan* with implications for deuterostome evolution". *Science*, **300**(5624), 1372; author reply 1372.

Mangin L. 2008. La métamorphose, un phénomène partagé. *Pour la science*, **369**, 18.

Mansouri A., Chowdhury K., & Gruss P. 1998. Follicular cells of the thyroid gland require *Pax8* gene functions. *Nat Genet*, **19**(1), 87–90.

Manzon R. G., & Youson J. H. 1997. The effects of exogenous thyroxine (T4) or triiodothyronine (T3), in the presence and absence of potassium perchlorate, on the incidence of metamorphosis and on serum T4 and T3 concentrations in larval sea lampreys (*Petromyzon marinus L.*). *Gen Comp Endocrinol*, **106**(2), 211–20.

Manzon R. G., & Youson J. H. 1999. Temperature and KClO(4)-induced metamorphosis in the sea lamprey (*Petromyzon marinus*). *Comp Biochem Physiol C Pharmacol Toxicol Endocrinol*, **124**(3), 253–7.

Manzon R. G., & Youson J. H. 2002. KClO(4) inhibits thyroidal activity in the larval lamprey endostyle in vitro. *Gen Comp Endocrinol*, **128**(3), 214–23.

Manzon R. G., Eales J. G., & Youson J. H. 1998. Blocking of KC1O4-induced metamorphosis in premetamorphic sea lampreys by exogenous thyroid hormones (TH); effects of KC1O4 and TH on serum TH concentrations and intestinal thyroxine outer-ring deiodination. *Gen Comp Endocrinol*, **112**(1), 54–62.

Manzon R. G., Holmes J. A., & Youson J. H. 2001. Variable effects of goitrogens in inducing precocious metamorphosis in sea lampreys (*Petromyzon marinus*). *J Exp Zool*, **289**(5), 290–303.

Manzon R. G., Neuls T. M., & Manzon L. A. 2007. Molecular cloning, tissue distribution, and developmental expression of lamprey transthyretins. *Gen Comp Endocrinol*, **151**(1), 55–65.

Marcellini S., & Simpson P. 2006. Two or four bristles: functional evolution of an enhancer of scute in Drosophilidae. *PLoS Biol*, **4**(12), e386.

Marchand O., Duffraisse M., Triqueneaux G., Safi R., & Laudet V. 2004. Molecular cloning and developmental expression patterns of thyroid hormone receptors and T3 target genes in the turbot (Scophtalmus maximus) during post-embryonic development. *Gen Comp Endocrinol*, **135**(3), 345–57.

Marimuthu A., Feng W., Tagami T., Nguyen H., Jameson J. L., Fletterick R. J., Baxter J. D., & West B. L. 2002. TR surfaces and conformations required to bind nuclear receptor corepressor. *Mol Endocrinol*, **16**(2), 271–286.

Markov G., Tavares R., Dauphin-Villemant C., Baker M., B. D., & V. L. submitted. Independent elaborations of steroid hormone signaling pathways in Metazoans. *Proc Natl AcadSci USA*.

Markov G. V., Paris M., Bertrand S., & Laudet V. 2008a. The evolution of the ligand / receptor couple: a long road from comparative endocrinology to comparative genomics. *Mol Cell Endocrinol*, **293**, 5–16.

Markov G. V., Lecointre G., Demeneix B., & Laudet V. 2008b. The "street light syndrome", or how protein taxonomy can bias experimental manipulations. *BioEssays*, **30**(4), 349–357.

Marletaz F., Gilles A., Caubit X., Perez Y., Dossat C., Samain S., Gyapay G., Wincker P., & Le Parco Y. 2008. Chaetognath transcriptome reveals ancestral and unique features among bilaterians. *Genome Biol*, **9**(6), R94.

Mayr E. 2001. *What evolution is*. New York: Basic Books.

Mazet F., Hutt J. A., Milloz J., Millard J., Graham A., & Shimeld S. M. 2005. Molecular evidence from *Ciona intestinalis* for the evolutionary origin of vertebrate sensory placodes. *Dev Biol*, **282**(2), 494–508.

McComb D. M., Gelsleichter J., Manire C. A., Brinn R., & Brown C. L. 2005. Comparative thyroid hormone concentration in maternal serum and yolk of the

bonnethead shark (*Sphyrna tiburo*) from two sites along the coast of Florida. *Gen Comp Endocrinol*, **144**(2), 167–73.

Meulemans D., & Bronner-Fraser M. 2007. The amphioxus SoxB family: implications for the evolution of vertebrate placodes. *Int J Biol Sci*, **3**(6), 356–364.

Michalik L., Auwerx J., Berger J. P., Chatterjee V. K., Glass C. K., Gonzalez F. J., Grimaldi P. A., Kadowaki T., Lazar M. A., O'Rahilly S., Palmer C. N. A., Plutzky J., Reddy J. K., Spiegelman B. M., Staels B., & Wahli W. 2006. International Union of Pharmacology. LXI. Peroxisome proliferator-activated receptors. *Pharmacol Rev*, **58**(4), 726–741.

Miwa S., & Inui Y. 1985. Effects of l-thyroxine and ovine growth hormone on smoltification of amago salmon (*Oncorhynchus rhodurus*). *Gen Comp Endocrinol*, **58**(3), 436–442.

Mochizuki K., Yagi E., Sakaguchi N., Mochizuki H., Takabe S., Kuranuki S., Suzuki T., Shimada M., & Goda T. 2007. The critical period for thyroid hormone responsiveness through thyroid hormone receptor isoform alpha in the postnatal small intestine. *Biochim Biophys Acta*, **1770**(4), 609–16.

Monaco F., Andreoli M., La Posta A., & Roche J. 1978. Thyroglobulin biosynthesis in a larval (ammocoete) and adult freshwater lamprey (*Lampetra planeri Bl.*). *Comp Biochem Physiol B*, **60**(1), 87–91.

Monaco F., Dominici R., Andreoli M., PirroR. D., & Roche J. 1981. Thyroid hormone formation in thyroglobulin synthesized in the amphioxus (*Branchiostoma lanceolatum Pallas*). *Comp Biochem Physiol B*, **70**, 341–3.

Moore D. D., Kato S., Xie W., Mangelsdorf D. J., Schmidt D. R., Xiao R., & Kliewer S. A. 2006. International Union of Pharmacology. LXII. The NR1H and NR1I receptors: constitutive androstane receptor, pregnene X receptor, farnesoid X receptor alpha, farnesoid X receptor beta, liver X receptor alpha, liver X receptor beta, and vitamin D receptor. *Pharmacol Rev*, **58**(4), 742–759.

Moreno M., Kaptein E., Goglia F., & Visser T. J. 1994. Rapid glucuronidation of tri- and tetraiodothyroacetic acid to ester glucuronides in human liver and to ether glucuronides in rat liver. *Endocrinology*, **135**(3), 1004–9.

Morris S. C. 1979. The Burgess shale (Middle cambrian) fauna. *Ann Rev Ecol Syst*, **10**, 327–349.

Morvan Dubois G., Sebillot A., Kuiper G. G., Verhoelst C. H., Darras V. M., Visser T. J., & Demeneix B. A. 2006. Deiodinase activity is present in *Xenopus laevis* during early embryogenesis. *Endocrinology*, **147**(10), 4941–9.

Morvan-Dubois G., Demeneix B. A., & Sachs L. M. 2008. *Xenopus laevis* as a model for studying thyroid hormone signalling: From development to metamorphosis. *Mol Cell Endocrinol*, **293**(1-2), 71–79.

Nascimento A. S., Dias S. M. G., Nunes F. M., Aparício R., Ambrosio A. L. B., Bleicher L., Figueira A. C. M., Santos M. A. M., de Oliveira Neto M., Fischer H., Togashi M., Craievich A. F., Garratt R. C., Baxter J. D., Webb P., & Polikarpov I. 2006. Structural rearrangements in the thyroid hormone receptor hinge domain and their putative role in the receptor function. *J Mol Biol*, **360**(3), 586–598.

Nelson J. S. 2006. *Fishes of the world*. 4th edn. Hoboken, N.J.: John Wiley.

Newman R. A. 1998. Ecological constraints on amphibian metamorphosis: interactions of temperature and larval density with responses to changing food level. *Oecologia*, **115**(1), 9–16.

Nguyen N. H., Apriletti J. W., Cunha Lima S. T., Webb P., Baxter J. D., & Scanlan T. S. 2002. Rational design and synthesis of a novel thyroid hormone antagonist that blocks coactivator recruitment. *J Med Chem*, **45**(15), 3310–20.

Nielsen C. 1998. Origin and evolution of animal life cycles. *Biol Rev Camb Philos Soc*, **73**(2), 125–155.

Nieuwkoop P. D., & Faber F. 1994. *Normal Table of Xenopus laevis*. Garland, New York.

Nohara M., Nishida M., Miya M., & Nishikawa T. 2005. Evolution of the mitochondrial genome in cephalochordata as inferred from complete nucleotide sequences from two epigonichthys species. *J Mol Evol*, **60**(4), 526–537.

Nolte R. T., Wisely G. B., Westin S., Cobb J. E., Lambert M. H., Kurokawa R., Rosenfeld M. G., Willson T. M., Glass C. K., & Milburn M. V. 1998. Ligand binding and co-activator assembly of the peroxisome proliferator-activated receptor-gamma. *Nature*, **395**(6698), 137–43.

Nomura T., Khan M. M., Kaul S. C., Dong H. D., Wadhwa R., Colmenares C., Kohno I., & Ishii S. 1999. Ski is a component of the histone deacetylase complex required for transcriptional repression by Mad and thyroid hormone receptor. *Genes Dev*, **13**(4), 412–423.

Nübler-Jung K., & Arendt D. 1996. Enteropneusts and chordate evolution. *Curr Biol*, **6**(4), 352–353.

Ogasawara M. 2000. Overlapping expression of amphioxus homologs of the *thyroid transcription factor-1* gene and *thyroid peroxidase* gene in the endostyle: insight into evolution of the thyroid glands. *Dev Genes Evol*, **210**(5), 231–42.

Ohno S. 1970. *Evolution by gene duplication*. Berlin, New York,: Springer-Verlag.

Onate S. A., Tsai S. Y., Tsai M. J., & O'Malley B. W. 1995. Sequence and characterization of a coactivator for the steroid hormone receptor superfamily. *Science*, **270**(5240), 1354–1357.

Ortlund E. A., Bridgham J. T., Redinbo M. R., & Thornton J. W. 2007. Crystal structure of an ancient protein: evolution by conformational epistasis. *Science*, **317**(5844), 1544–8.

Owen R. 1843. *Lectures on the comparative anatomy and physiology of the invertebrate animals*. Longman, Brown, Green, and Longmans.

Pace N. R. 1997. A molecular view of microbial diversity and the biosphere. *Science*, **276**(5313), 734–740.

Paris M., Pettersson K., Schubert M., Bertrand S., Pongratz I., Escriva H., & Laudet V. 2008a. An amphioxus orthologue of the estrogen receptor that does not bind estradiol: insights into estrogen receptor evolution. *BMC Evol Biol*, **8**(1), 219.

Paris M., Escriva H., Schubert M., Brunet F., Brtko J., Ciesielski F., Roecklin D., Vivat-Hannah V., Jamin E. L., Cravedi J. P., Scanlan T. S., Renaud J. P., Holland N. D., & Laudet V. 2008b. Amphioxus postembryonic development reveals the homology of chordate metamorphosis. *Curr Biol*, **18**(11), 825–30.

Paris M., & Laudet V. 2008. The history of a developmental stage: Metamorphosis in chordates. *Genesis*, **46**, 657–672.

Paris M., Brunet F., Markov G., Schubert M., & Laudet V. 2008c. The amphioxus genome enlightens the evolution of the thyroid hormone signaling pathway. *Dev Genes Evol*, Nov.

Patel N. H., Ball E. E., & Goodman C. S. 1992. Changing role of even-skipped during the evolution of insect pattern formation. *Nature*, **357**(6376), 339–342.

Patricolo E., Ortolani G., & Cascio A. 1981. The effect of l-thyroxine on the metamorphosis of Ascidia malaca. *Cell Tissue Cell Res*, **214**(2), 289–301.

Patricolo E., Cammarata M., & D'Agati P. 2001. Presence of thyroid hormones in ascidian larvae and their involvement in metamorphosis. *J Exp Zool*, **290**(4), 426–30.

Pearson J. C., Lemons D., & McGinnis W. 2005. Modulating *Hox* gene functions during animal body patterning. *Nat Rev Genet*, **6**(12), 893–904.

Pennisi E. 2008. Evolutionary biology. Deciphering the genetics of evolution. *Science*, **321**(5890), 760–3.

Phelps C., Gburcik V., Suslova E., Dudek P., Forafonov F., Bot N., MacLean M., Fagan R. J., & Picard D. 2006. Fungi and animals may share a common ancestor to nuclear receptors. *Proc Natl Acad Sci U S A*, **103**(18), 7077–7081.

Philippe H., Zhou Y., Brinkmann H., Rodrigue N., & Delsuc F. 2005a. Heterotachy and long-branch attraction in phylogenetics. *BMC Evol Biol*, **5**, 50.

Philippe H., Lartillot N., & Brinkmann H. 2005b. Multigene analyses of bilaterian animals corroborate the monophyly of Ecdysozoa, Lophotrochozoa, and Protostomia. *Mol Biol Evol*, **22**(5), 1246–1253.

Plateroti M., Chassande O., Fraichard A., Gauthier K., Freund J. N., Samarut J., & Kedinger M. 1999. Involvement of T3Ralpha- and beta-receptor subtypes in mediation of T3 functions during postnatal murine intestinal development. *Gastroenterology*, **116**(6), 1367–78.

Power D. M., Llewellyn L., Faustino M., Nowell M. A., Bjornsson B. T., Einarsdottir I. E., Canario A. V., & Sweeney G. E. 2001. Thyroid hormones in growth and development of fish. *Comp Biochem Physiol C Toxicol Pharmacol*, **130**(4), 447–59.

Prud'homme B., Gompel N., & Carroll S. B. 2007. Emerging principles of regulatory evolution. *Proc Natl Acad Sci U S A*, **104 Suppl 1**, 8605–12.

Putnam N. H., Butts T., Ferrier D. E., Furlong R. F., Hellsten U., Kawashima T., Robinson-Rechavi M., Shoguchi E., Terry A., Yu J. K., Benito-Gutierrez E. L., Dubchak I., Garcia-Fernandez J., Gibson-Brown J. J., Grigoriev I. V., Horton A. C., de Jong P. J., Jurka J., Kapitonov V. V., Kohara Y., Kuroki Y., Lindquist E., Lucas S., Osoegawa K., Pennacchio L. A., Salamov A. A., Satou Y., Sauka-Spengler T., Schmutz J., Shin I. T., Toyoda A., Bronner-Fraser M., Fujiyama A., Holland L. Z., Holland P. W., Satoh N., & Rokhsar D. S. 2008. The amphioxus genome and the evolution of the chordate karyotype. *Nature*, **453**(7198), 1064–71.

Ranjan M., Wong J., & Shi Y. B. 1994. Transcriptional repression of *Xenopus* TR beta gene is mediated by a thyroid hormone response element located near the start site. *J Biol Chem*, **269**(40), 24699–24705.

Rastinejad F., Perlmann T., Evans R. M., & Sigler P. B. 1995. Structural determinants of nuclear receptor assembly on DNA direct repeats. *Nature*, **375**(6528), 203–211.

Reisz R. R. 1997. The origin and early evolutionary history of amniotes. *Trends in Ecology & Evolution*, **12**(6), 218–222.

Renaud J. P., Rochel N., Ruff M., Vivat V., Chambon P., Gronemeyer H., & Moras D. 1995. Crystal structure of the RAR-gamma ligand-binding domain bound to all-trans retinoic acid. *Nature*, **378**(6558), 681–689.

Ribeiro R. C., Apriletti J. W., Yen P. M., Chin W. W., & Baxter J. D. 1994. Heterodimerization and deoxyribonucleic acid-binding properties of a retinoid X receptor-related factor. *Endocrinology*, **135**(5), 2076–2085.

Roche J., Lissitzky S., & Michel R. 1952a. Sur la présence de triiodothyronine dans la thyroglobuline. *C R Acad Scie*, **234**, 1228–1230.

Roche J., Lissitzky S., & Michel R. 1952b. Sur la triiodothyronine, produit intermediaire de la transformation de la diiodothyronine en thyroxine. *C R Acad Scie*, **234**, 997–998.

Ronquist F., & Huelsenbeck J. P. 2003. MrBayes 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics*, **19**(12), 1572–1574.

Roustaian P., & Gaik L. A. 2006. Effect of thyroxine immersion on larval survival, growth and postlarvae production of freshwater prawn, Macrobrachium rosenbergii (de Man). *Aquaculture Research*, **37**, 1378–1380.

Ruta M. 1999. Brief review of the stylophoran debate. *Evol Dev*, **1**(2), 123–35.

Sachs L. M., & Shi Y. B. 2000. Targeted chromatin binding and histone acetylation in vivo by thyroid hormone receptor during amphibian development. *Proc Natl Acad Sci U S A*, **97**(24), 13138–13143.

Sachs L. M., Jones P. L., Havis E., Rouse N., Demeneix B. A., & Shi Y. B. 2002. Nuclear receptor corepressor recruitment by unliganded thyroid hormone receptor in gene repression during *Xenopus laevis* development. *Mol Cell Biol*, **22**(24), 8527–38.

Safi R., Bertrand S., Marchand O., Duffraisse M., de Luze A., Vanacker J. M., Maraninchi M., Margotat A., Demeneix B., & Laudet V. 2004. The axolotl (*Ambystoma mexicanum*), a neotenic amphibian, expresses functional thyroid hormone receptors. *Endocrinology*, **145**(2), 760–72.

Safi R., Vlaeminck-Guillem V., Duffraisse M., Seugnet I., Plateroti M., Margotat A., Duterque-Coquillaud M., Crespi E. J., Denver R. J., Demeneix B., & Laudet V. 2006. Pedomorphosis revisited: thyroid hormone receptors are functional in *Necturus* maculosus. *Evol Dev*, **8**(3), 284–92.

Sap J., Muñoz A., Damm K., Goldberg Y., Ghysdael J., Leutz A., Beug H., & Vennström B. 1986. The c-erb-A protein is a high-affinity receptor for thyroid hormone. *Nature*, **324**(6098), 635–640.

Sato Y., Buchholz D. R., Paul B. D., & Shi Y. B. 2007. A role of unliganded thyroid hormone receptor in postembryonic development in *Xenopus laevis*. *Mech Dev*, **124**(6), 476–488.

Satoh N. 1994. *Developmental biology of ascidians*. Cambridge University Press, Cambridge, UK.

Sauka-Spengler T., & Bronner-Fraser M. 2008. Evolution of the neural crest viewed from a gene regulatory perspective. *Genesis*, **46**(11), 673–682.

Sauka-Spengler T., Meulemans D., Jones M., & Bronner-Fraser M. 2007. Ancient evolutionary origin of the neural crest gene regulatory network. *Dev Cell*, **13**(3), 405–420.

Scanlan T. S., Suchland K. L., Hart M. E., Chiellini G., Huang Y., Kruzich P. J., Frascarelli S., Crossley D. A., Bunzow J. R., Ronca-Testoni S., Lin E. T., Hatton D., Zucchi R., & Grandy D. K. 2004. 3-Iodothyronamine is an endogenous and rapid-acting derivative of thyroid hormone. *Nat Med*, **10**(6), 638–42.

Schlosser G. 2005. Evolutionary origins of vertebrate placodes: insights from developmental studies and from comparisons with other deuterostomes. *J Exp Zoolog B Mol Dev Evol*, **304**(4), 347–99.

Schlosser G. 2008. Do vertebrate neural crest and cranial placodes have a common evolutionary origin? *BioEsays*, **30**(7), 659–672.

Schubert M., Escriva H., Xavier-Neto J., & Laudet V. 2006. Amphioxus and tunicates as evolutionary model systems. *Trends Ecol Evol*, **21**(5), 269–77.

Schubert M., Brunet F., Paris M., Bertrand S., Benoit G., & Laudet V. 2008. Nuclear hormone receptor signaling in amphioxus. *Dev Genes Evol*, Sep.

Schueler P. A., Schwartz H. L., Strait K. A., Mariash C. N., & Oppenheimer J. H. 1990. Binding of 3,5,3′-triiodothyronine (T3) and its analogs to the in vitro translational products of c-erbA protooncogenes: differences in the affinity of the alpha- and beta-forms for the acetic acid analog and failure of the human testis and kidney alpha-2 products to bind T3. *Mol Endocrinol*, **4**(2), 227–234.

Sea Urchin Genome Sequencing C., Sodergren E., Weinstock G. M., Davidson E. H., Cameron R. A., Gibbs R. A., Angerer R. C., Angerer L. M., Arnone M. I., Burgess D. R., Burke R. D., Coffman J. A., Dean M., Elphick M. R., Ettensohn

C. A., Foltz K. R., Hamdoun A., Hynes R. O., Klein W. H., Marzluff W., McClay D. R., Morris R. L., Mushegian A., Rast J. P., Smith L. C., Thorndyke M. C., Vacquier V. D., Wessel G. M., Wray G., Zhang L., Elsik C. G., Ermolaeva O., Hlavina W., Hofmann G., Kitts P., Landrum M. J., Mackey A. J., Maglott D., Panopoulou G., Poustka A. J., Pruitt K., Sapojnikov V., Song X., Souvorov A., Solovyev V., Wei Z., Whittaker C. A., Worley K., Durbin K. J., Shen Y., Fedrigo O., Garfield D., Haygood R., Primus A., Satija R., Severson T., Gonzalez-Garay M. L., Jackson A. R., Milosavljevic A., Tong M., Killian C. E., Livingston B. T., Wilt F. H., Adams N., Belle R., Carbonneau S., Cheung R., Cormier P., Cosson B., Croce J., Fernandez-Guerra A., Geneviere A.-M., Goel M., Kelkar H., Morales J., Mulner-Lorillon O., Robertson A. J., Goldstone J. V., Cole B., Epel D., Gold B., Hahn M. E., Howard-Ashby M., Scally M., Stegeman J. J., Allgood E. L., Cool J., Judkins K. M., McCafferty S. S., Musante A. M., Obar R. A., Rawson A. P., Rossetti B. J., Gibbons I. R., Hoffman M. P., Leone A., Istrail S., Materna S. C., Samanta M. P., Stolc V., *et al.* . 2006. The Genome of the Sea Urchin *Strongylocentrotus purpuratus*. *Science*, **314**(5801), 941–952.

Sembrat K. 1953. Effect of the endostyle of the lancelet (*Branchiostoma lanceolatum Pall.*) on the metamorphosis of axolotls. *Zool Polon*, **6**, 3–19.

Seo H. C., Kube M., Edvardsen R. B., Jensen M. F., Beck A., Spriet E., Gorsky G., Thompson E. M., Lehrach H., Reinhardt R., & Chourrout D. 2001. Miniature genome in the marine chordate Oikopleura dioica. *Science*, **294**(5551), 2506.

Seo H. C., Edvardsen R. B., Maeland A. D., Bjordal M., Jensen M. F., Hansen A., Flaat M., Weissenbach J., Lehrach H., Wincker P., Reinhardt R., & Chourrout D. 2004. Hox cluster disintegration with persistent anteroposterior order of expression in Oikopleura dioica. *Nature*, **431**(7004), 67–71.

Shepherdley C. A., Klootwijk W., Makabe K. W., Visser T. J., & Kuiper G. G. J. M. 2004. An Ascidian Homolog of Vertebrate Iodothyronine Deiodinases. *Endocrinology*, **145**(3), 1255–1268.

Shi Y.-B. 2000. *Amphibian metamorphosis : from morphology to molecular biology*. New York: John Wiley.

Shu D.-G., Morris S. C., & Zhang X.-L. 1996. A Pikaia-like chordate from the Lower Cambrian of China. *Nature*, **384**(6605), 157–158.

Shu D., Morris S. C., Zhang Z. F., Liu J. N., Han J., Chen L., Zhang X. L., Yasui K., & Li Y. 2003. A new species of yunnanozoan with implications for deuterostome evolution. *Science*, **299**(5611), 1380–1384.

Sly B. J., Snoke M. S., & Raff R. A. 2003. Who came first–larvae or adults? origins of bilaterian metazoan larvae. *Int J Dev Biol*, **47**(7-8), 623–32.

Somorjai I., Bertrand S., Camasses A., Haguenauer A., & Escriva H. 2008. Evidence for stasis and not genetic piracy in developmental expression patterns of *Branchiostoma lanceolatum* and *Branchiostoma floridae*, two amphioxus species that have evolved independently over the course of 200 Myr. *Dev Genes Evol*, Oct.

Spangenberg D. B. 1971. Thyroxine induced metamorphosis in Aurelia. *J Exp Zool*, **178**(2), 183–94.

Spemann H. 1915. *Zur Geschichte und Kritik des Begriffs der Homologies*. Wilhelm Engelmann.

Stokes M. 1997. Larval locomotion of the lancelet. *J Exp Biol*, **200**(11), 1661–1680.

Stokes M. D. 1996. Larval settlement, post-settlement growth and secondary production of the Florida lancelet (= amphioxus) *Branchiostoma floridae*. *Mar Ecol Prog Ser*, **130**, 71–84.

Stokes M. D., & Holland N. D. 1995. Ciliary Hovering in Larval Lancelets (=Amphioxus). *Biol Bull*, **188**(3), 231–233.

Strickberger, Monroe W. 2000. *Evolution*. Jones & Bartlett Publishers, Sudbury.

Sutija M., Longhurst T. J., & Joss J. M. 2003. Deiodinase type II and tissue specific mRNA alternative splicing in the Australian lungfish, *Neoceratodus forsteri*. *Gen Comp Endocrinol*, **132**(3), 409–17.

Sutija M., Longhurst T. J., & Joss J. M. 2004. Deiodinase type III in the Australian lungfish, *Neoceratodus forsteri*. *Gen Comp Endocrinol*, **136**(2), 152–61.

Swalla B. J., Cameron C. B., Corley L. S., & Garey J. R. 2000. Urochordates are monophyletic within the deuterostomes. *Syst Biol*, **49**(1), 52–64.

Tata J. R. 1966. Requirement for RNA and protein synthesis for induced regression of the tadpole tail in organ culture. *Dev Biol*, **13**(1), 77–94.

Tata J. R. 1968. Early metamorphic competence of *Xenopus* larvae. *Dev Biol*, **18**(5), 415–40.

Tata J. R. 1993. Gene expression during metamorphosis: an ideal model for post-embryonic development. *BioEssays*, **15**(4), 239–48.

Tata J. R. 2000. Autoinduction of nuclear hormone receptors during metamorphosis and its significance. *Insect Biochem Mol Biol*, **30**(8-9), 645–51.

Tata J. R. 2006. Amphibian metamorphosis as a model for the developmental actions of thyroid hormone. *Mol Cell Endocrinol*, **246**(1-2), 10–20.

Tata J. 1998. *Hormonal signaling and postembryonic development*. Springer-Verlag, Berlin Heidelberg New York.

Tata J., & Widnell C. 1966. Ribonucleic acid synthesis during the early action of thyroid hormones. *Biochem J*, **98**, 604–620.

Thakur J. K., Arthanari H., Yang F., Pan S.-J., Fan X., Breger J., Frueh D. P., Gulshan K., Li D. K., Mylonakis E., Struhl K., Moye-Rowley W. S., Cormack B. P., Wagner G., & Näär A. M. 2008. A nuclear receptor-like pathway regulating multidrug resistance in fungi. *Nature*, **452**(7187), 604–609.

Thomas I. M. 1956. The accumulation of radioactive iodine by amphioxus. *J Mar Biol Ass UK*, **35**, 203–10.

Thornton J. M. 2001a. From genome to function. *Science*, **292**(5524), 2095–7.

Thornton J. W. 2001b. Evolution of vertebrate steroid receptors from an ancestral estrogen receptor by ligand exploitation and serial genome expansions. *Proc Natl Acad Sci U S A*, **98**(10), 5671–6.

Thornton J. W. 2004. Resurrecting ancient genes: experimental analysis of extinct molecules. *Nat Rev Genet*, **5**(5), 366–75.

Thornton J. W., Need E., & Crews D. 2003. Resurrecting the ancestral steroid receptor: ancient origin of estrogen signaling. *Science*, **301**(5640), 1714–7.

Tinnikov A., Nordström K., Thorén P., Kindblom J. M., Malin S., Rozell B., Adams M., Rajanayagam O., Pettersson S., Ohlsson C., Chatterjee K., & Vennström B. 2002. Retardation of post-natal development caused by a negatively acting thyroid hormone receptor alpha1. *EMBO J*, **21**(19), 5079–5087.

Tomura H., Lazar J., Phyillaier M., & Nikodem V. M. 1995. The N-terminal region (A/B) of rat thyroid hormone receptors alpha 1, beta 1, but not beta 2 contains a strong thyroid hormone-dependent transactivation function. *Proc Natl Acad Sci U S A*, **92**(12), 5600–5604.

Tong W., & Chaikoff I. L. 1961. 131I utlization by the aquarium snail and the cockroach. *Biochim Biophys Acta*, **48**, 347–351.

Tong W., Kerkof P., & Chaikoff I. L. 1962. Identification of labeled thyroxine and triiodothyronine in amphioxus treated with 131-I. *Biochim Biophys Acta*, **56**, 326–31.

Tsuneki K., Kobayashi H., & Ouji M. 1983. Histochemical distribution of peroxidase in amphioxus and cyclostomes with special reference to the endostyle. *Gen Comp Endocrinol*, **50**(2), 188–200.

Tzagarakis-Foster C., & Privalsky M. L. 1998. Phosphorylation of thyroid hormone receptors by protein kinase A regulates DNA recognition by specific inhibition of receptor monomer binding. *J Biol Chem*, **273**(18), 10926–10932.

Umesono K., & Evans R. M. 1989. Determinants of target gene specificity for steroid/thyroid hormone receptors. *Cell*, **57**(7), 1139–1146.

Varsamos S., Nebel C., & Charmantier G. 2005. Ontogeny of osmoregulation in postembryonic fish: a review. *Comp Biochem Physiol A Mol Integr Physiol*, **141**(4), 401–29.

Voss S. R., & Shaffer H. B. 1997. Adaptive evolution via a major gene effect: Paedomorphosis in the Mexican,Ä†axolotl. *Proc Natl Acad Sci USA*, **94**(25), 14185–14189.

Wagner G. P. 1989. The biological homology concept. *Ann Rev Ecol Syst*, **20**, 51–59.

Wagner G. P. 2007. The developmental genetics of homology. *Nat Rev Genet*, **8**(6), 473–9.

Wagner R. L., Apriletti J. W., McGrath M. E., West B. L., Baxter J. D., & Fletterick R. J. 1995. A structural role for hormone in the thyroid hormone receptor. *Nature*, **378**(6558), 690–7.

Wang Z., & Brown D. D. 1993. Thyroid hormone-induced gene expression program for amphibian tail resorption. *J Biol Chem*, **268**(22), 16270–8.

Wang Z., Benoit G., Liu J., Prasad S., Aarnisalo P., Liu X., Xu H., Walker N. P., & Perlmann T. 2003. Structure and function of Nurr1 identifies a class of ligand-independent nuclear receptors. *Nature*, **423**(6939), 555–60.

Weinberger C., Thompson C. C., Ong E. S., Lebo R., Gruol D. J., & Evans R. M. 1986. The *c-erb-A* gene encodes a thyroid hormone receptor. *Nature*, **324**(6098), 641–646.

Weirich R. T., Schwartz H. L., & Oppenheimer J. H. 1987. An analysis of the interrelationship of nuclear and plasma triiodothyronine in the sea lamprey, lake trout, and rat: evolutionary considerations. *Endocrinology*, **120**(2), 664–677.

Wickstead J. H. 1967. *Branchiostoma lanceolatum* larvae: some experiments on the effects of thiouracil on metamorphosis. *J Mar Biol Ass UK*, **47**, 49–59.

Wiens J. J., Kuczynski C. A., Duellman W. E., & Reeder T. W. 2007. Loss and re-evolution of complex life cycles in marsupial frogs: does ancestral trait reconstruction mislead? *Evolution*, **61**(8), 1886–1899.

Willey A. 1894. *Amphioxus and the ancestry of the vertebrates*. New York, London,: Macmillan.

Winchell C. J., Sullivan J., Cameron C. B., Swalla B. J., & Mallatt J. 2002. Evaluating Hypotheses of Deuterostome Phylogeny and Chordate Evolution with New LSU and SSU Ribosomal DNA Data. *Mol Biol Evol*, **19**(5), 762–776.

Wray G. A. 2007. The evolutionary significance of cis-regulatory mutations. *Nat Rev Genet*, **8**(3), 206–16.

Wu S. Y., Green W. L., Huang W. S., Hays M. T., & Chopra I. J. 2005. Alternate pathways of thyroid hormone metabolism. *Thyroid*, **15**(8), 943–58.

Wu W., Niles E. G., & LoVerde P. T. 2007. Thyroid hormone receptor orthologues from invertebrate species with emphasis on *Schistosoma mansoni*. *BMC Evol Biol*, **7**, 150.

Yamano K., & Miwa S. 1998. Differential gene expression of thyroid hormone receptor alpha and beta in fish development. *Gen Comp Endocrinol*, **109**(1), 75–85.

Yen P. M., Feng X., Flamant F., Chen Y., Walker R. L., Weiss R. E., Chassande O., Samarut J., Refetoff S., & Meltzer P. S. 2003. Effects of ligand and thyroid hormone receptor isoforms on hepatic gene expression profiles of thyroid hormone receptor knockout mice. *EMBO Rep*, **4**(6), 581–7.

Young G., Bjornsson B. T., Prunet P., Lin R. J., & Bern H. A. 1989. Smoltification and seawater adaptation in Coho salmon (*Oncorhyncus kisutch*): plasma prolactin, growth hormone, thyroid hormone, and cortisol. *Gen Comp endocrinol*, **74**, 335–345.

Youson J. H., Manzon R. G., Peck B. J., & Holmes J. A. 1997. Effects of exogenous thyroxine (T4) and triiodothyronine (T3) on spontaneous metamorphosis and serum T4 and T3 levels in immediately premetamorphic sea lampreys, *Petromyzon marinus*. *J Exp Zool*, **279**(2), 145–55.

Youson J. H. 1997. Is Lamprey Metamorphosis Regulated by Thyroid Hormones? *Amer. Zool.*, **37**(6), 441–460.

Zeitlinger J., Zinzen R. P., Stark A., Kellis M., Zhang H., Young R. A., & Levine M. 2007. Whole-genome ChIP-chip analysis of Dorsal, Twist, and Snail suggests

integration of diverse patterning processes in the *Drosophila* embryo. *Genes Dev*, **21**(4), 385–390.

Zhaxybayeva O., Lapierre P., & Gogarten J. P. 2005. Ancient gene duplications and the root(s) of the tree of life. *Protoplasma*, **227**(1), 53–64.

Zuckerkandl E., & Pauling L. 1965. Molecules as documents of evolutionary history. *J Theor Biol*, **8**(2), 357–366.

Part V

APPENDIX

ADDITIONAL ARTICLES

In addition to the articles described in the main chapters of this manuscript, I participated in other studies. A brief description of the different subjects as well as my contribution to each article are described below. The articles are given in the following pages.

- I wrote with Vincent Laudet a review that discusses the evolution of metamorphosis in light of the evolution of its molecular determinism. Most of this review was included after modifications in the introduction (chapters **??** and **??**) and the discussion (chapter **??**). This article was published in Genesis (page **??**)

- I participated in a study focusing on morphological aspects of amphioxus metamorphosis (page **??**). In this study, the fate of the club-shaped gland during metamorphosis was scrutinized in an attempt to understand whether it disappears or is, at least partly, reincorporated into other adjacent organs. I contributed to the amphioxus culture, as well as metamorphosis monitoring. This article was published in Acta Zoologica (page **??**).

- I participated in the phylogenetic study of the NR superfamily. In this study, published in the Development, Genes and Evolution special edition on the amphioxus genome, the complete set of amphioxus NRs as well as their corepressors, were searched in the amphioxus genome, and their sequence evolution was studied through phylogenetic analyses. I contributed to the study of the NR3 subfamily (page **??**).

- I participated in the writing of a review that discusses pros and cons of current approaches used in comparative endocrinology to study the evolution of endocrine systems and that pleads for a multi-species approach. This article was published in Molecular and cellular endocrinology (page **??**).

- I participated in a collective study of the reproductive behavior of the amphioxus species *Branchiostoma floridae* that enables to deepen our knowledge of inducible spawning in amphioxus. This article was published in Journal of experimental zoology (part B) (page **??**).

- I participated in a phylogenetic study of Ursidae using genes involved in the thyroid hormone signaling pathway. I was involved in the choice of

the markers and I participated in the phylogenetic studies. This article was published in Molecular Phylogenetics and Evolution (page **??**).

- I participated in the improvement of a method of shRNA injection in mouse brain. This technique was applied to the TR gene. My contribution was technical, as I participated in some experiments of cell transfection and measurements of mRNA expression by quantitative RT-PCR. This article was published in Nucleic Acid Research (page **??**)

- I participated in the study of the convertase family in amphioxus by contributing to the phylogenetic analysis of this family. This article was published in International Journal of Biological Sciences (page **??**).

- I participated in a structural analysis of the amphioxus RXR. My contribution consisted on making a phylogenetic analysis of RXRs in a large sample of metazoans. This article was published in Journal of Biological Chemistry (page **??**).

- I participated in an analysis of the fate of duplicate genes in fishes after the fish-specific genome duplication. This article was published in Molecular Biology and Evolution (page **??**).

REVIEW

The History of a Developmental Stage: Metamorphosis in Chordates

Mathilde Paris and Vincent Laudet*

Institut de Génomique Fonctionnelle de Lyon, Université de Lyon, CNRS, INRA, UCB Lyon 1, IFR128 Lyon Biosciences, Ecole Normale Supérieure de Lyon, 69364 Lyon Cedex 07, France

Received 15 April 2008; Revised 1 September 2008; Accepted 3 September 2008

Summary: Metamorphosis displays a striking diversity in chordates, a deuterostome phylum that comprises vertebrates, urochordates (tunicates), and cephalochordates (amphioxus). In anuran amphibians, the tadpole loses its tail, develops limbs, and undergoes profound changes at the behavioral, physiological, biochemical, and ecological levels. In ascidian tunicates, the tail is lost and the head tissues are drastically remodeled into the adult animal, whereas in amphioxus, the highly asymmetric larva transforms into a relatively symmetric adult. This wide diversity has led to the proposal that metamorphosis evolved several times independently in the different chordate lineages during evolution. However, the molecular mechanisms involved in metamorphosis are largely unknown outside amphibians and teleost fishes, in which metamorphosis is regulated by the thyroid hormones (TH) T_3 and T_4 binding to their receptors (thyroid hormone receptors). In this review, we compare metamorphosis in chordates and then propose a unifying definition of the larva-to-adult transition, based on the conservation of the role of THs and some of their derivatives as the main regulators of metamorphosis. According to this definition, all chordates (if not, all deuterostomes) have a homologous metamorphosis stage during their postembryonic development. The intensity and the nature of the morphological remodeling varies extensively among taxa, from drastic remodeling like in some ascidians or amphibians to more subtle events, as in mammals. genesis 0:1–16, 2008. **© 2008 Wiley-Liss, Inc.**

Key words: thyroid hormones; metamorphosis; chordates; endostyle; evolution

Animal metamorphosis, the possibility for an organism to transform its form during life, has fascinated biologists for years. In developmental biology, it generally refers to a postembryonic developmental stage of an animal that is characterized by intense changes at the morphological, physiological, biochemical, behavioral, and ecological levels (Bishop *et al.*, 2006). More precisely, it can be viewed as a transition from a larva to a juvenile that is morphologically distinct ("meta" meaning "change" and "morph" meaning "form" in Greek) and exploits a different ecological environment. It should be distinguished from the transition to sexual maturity in many species, since the latter is not triggered by the same molecular mechanism and most often does not involve an ecological transformation.

When talking about metamorphosis, the caterpillar-to-butterfly and tadpole-to-frog transitions immediately come to mind. However, most metazoan phyla metamorphose: for example, in cnidarians, flat worms or mollusks, although the morphological remodeling varies quite extensively between lineages (Brusca and Brusca, 2003). The study of this spectacular transformation in insects as well as in vertebrates has shown the important role that hormones play during development (Gilbert *et al.*, 1996). In this review, we will discuss the notion of metamorphosis in the chordate phylum, which comprises cephalochordates (such as the amphioxus *Branchiostoma*), urochordates (such as the tunicate *Ciona intestinalis*), and vertebrates (Fig. 1). Within vertebrates, we will mostly present the agnathan lamprey, actinopterygian fishes, mammals, sauropsids (including birds), and amphibians in which metamorphosis has been studied to various extents (Fig. 1).

Vertebrate metamorphosis has been molecularly well characterized only in anurans, probably for historical (Gudernatsch, 1912), anthropocentric (anurans are the metamorphosing animals closest to humans), epistemological (amphibian metamorphosis is quite spectacular and appealing for the biologist), as well as technical reasons (tadpoles are big and easy to breed in a lab). First, we will briefly present the knowledge accumulated in

*Correspondence to: Vincent Laudet, Institut de Génomique Fonctionnelle de Lyon, Université de Lyon, CNRS, INRA, UCB Lyon 1, IFR128 Lyon Biosciences, Ecole Normale Supérieure de Lyon, 69364 Lyon Cedex 07, France. E-mail: Vincent.Laudet@ens-lyon.fr or vlaudet@ens-lyon.fr

Contract grant sponsors: MENRT, CNRS

Published online in Wiley InterScience (www.interscience.wiley.com).

DOI: 10.1002/dvg.20443

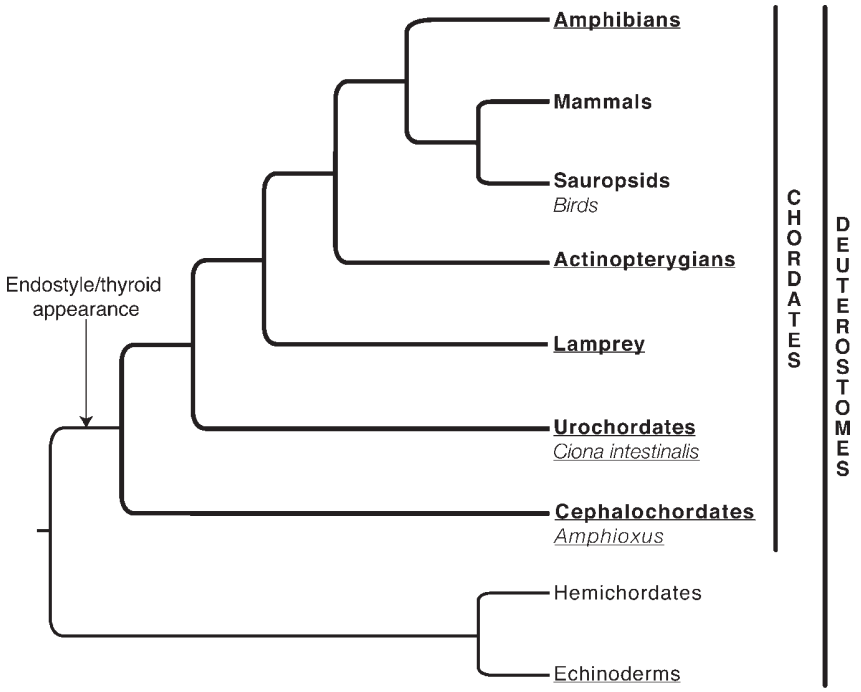


FIG. 1. Simplified phylogenetic tree of deuterostomes. Groups where a metamorphosis stage is usually described are underlined.

anurans, and to a lesser extent in actinopterygian fishes, since most of the experiments dissecting the molecular determinism of metamorphosis have been performed in those groups.

In anurans the main trigger of metamorphosis is known to be endocrine, through production of the precursor thyroid hormone (TH) T_4 (thyroxine) in the thyroid gland and its subsequent transformation into the more active derivative T_3 (triiodothyronine). Some T_3 is also produced in the thyroid (thus both T_3 and T_4 will be referred as THs). Indeed, exogenous T_3 or T_4 prematurely induces metamorphosis (Gudernatsch, 1912), whereas inhibition of endogenous T_3 or T_4 synthesis blocks metamorphosis and produces giant tadpoles (Allen, 1916; Buchholz *et al.*, 2006). More precisely, there is a peak of TH production at the beginning of metamorphosis, which is the key physiological trigger [(Leloup and Buscaglia, 1977) for review, see (Tata, 2006)]. TH levels reach a maximum during the climax, that is, when the major morphological changes (e.g., tail regression) occur. When environmental parameters (e.g., water volume or food availability) modulate the triggering of metamorphosis, it is always through alteration of the TH production that can thus be viewed as the key endogenous inducer of the metamorphosis process (Boelen *et al.*, 2008; Boorse and Denver, 2003; Denver, 1998; Ito *et al.*, 2004; Newman, 1998; Tata, 2006) (Fig. 2). We will restrict the term TH to both T_3 and T_4, as they are the only compounds known to be produced in the thyroid in significant quantities. Other iodinated compounds derived from T_3 or T_4 that display a TH-like activity will be called "thyroactive compounds" or more generally "TH derivatives" in the present review.

THs activate signaling pathways through specific binding to the high-affinity thyroid hormone receptor (TR). TR is a TH-regulated transcription factor that belongs to the superfamily of nuclear receptors that plays a central role in various physiological processes as diverse as reproduction, development, and the control of homeostasis (Germain *et al.*, 2006; Gronemeyer *et al.*, 2004; Laudet and Gronemeyer, 2005). In the absence of ligand, the receptor is located on specific response elements in the regulatory regions of target genes and actively inhibits its transcription. Ligand binding activates the receptor which then modulates the transcription of target genes, subsequently causing the morphological remodeling characteristic of metamorphosis (Buchholz *et al.*, 2006). There are two TRs in anurans (TRα and TRβ) from a vertebrate-specific genome duplication (Dehal and Boore, 2005; Escriva *et al.*, 2002) [actually there are two TRα and two TRβ in *Xenopus laevis* from a recent species-specific genome duplication (Escriva *et al.*, 2002), but the duplicates are functionally redundant]. TRα and TRβ have specific roles in the process of amphibian metamorphosis: TRβ is mainly implicated as a molecular switch triggered by the rise of TH levels, whereas TRα is instrumental in blocking the metamorphosis program before TH levels increase (Brown and Cai, 2007; Havis *et al.*, 2006). More precisely, TRα is expressed very early, even before TH synthesis, whereas TRβ expression is controlled by and thus correlates with TH levels [(Kawahara *et al.*, 1991), review in (Buchholz *et al.*, 2006)]. This peak of TRβ expression at the onset of metamorphosis

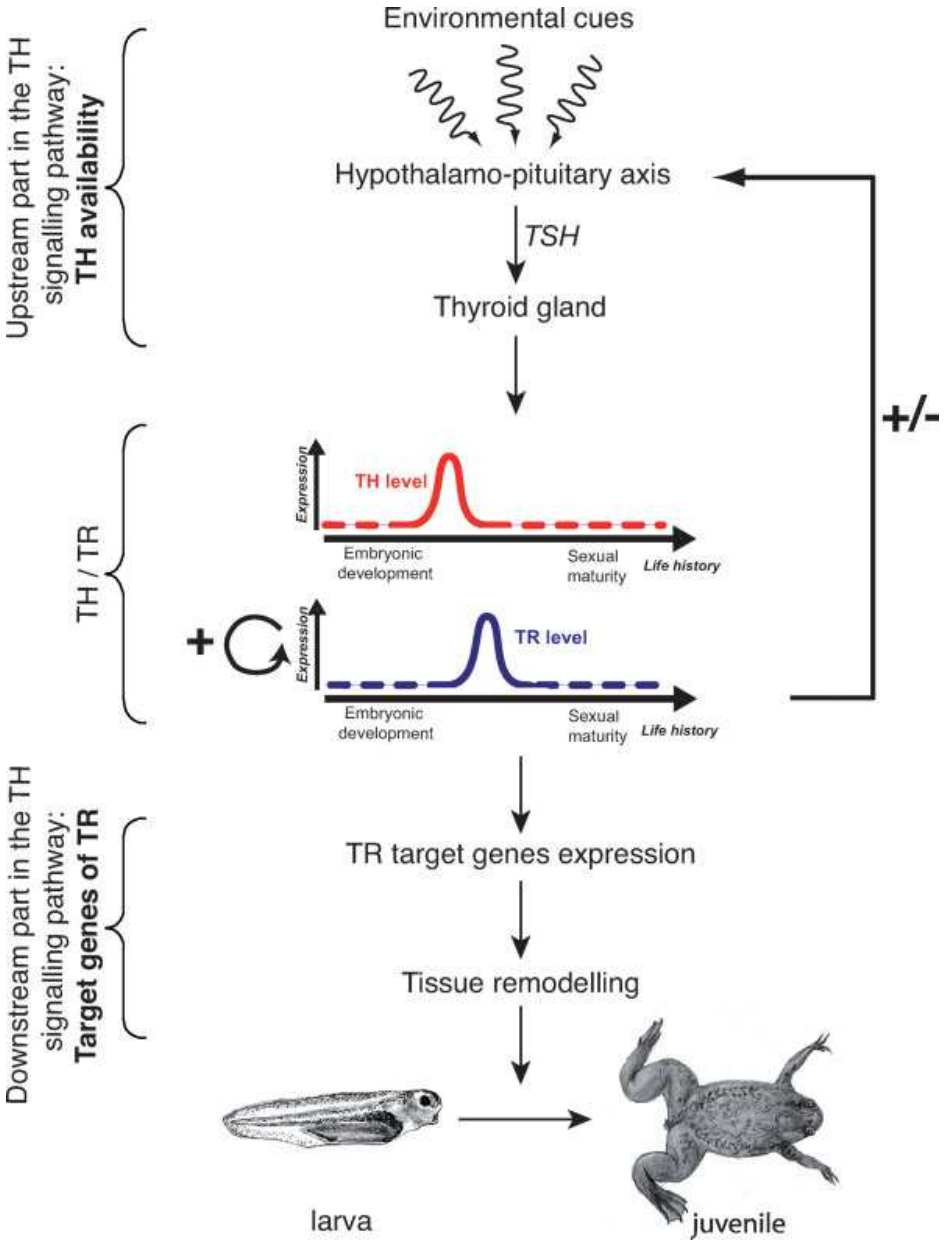


FIG. 2. Schematic structure of TH cascade in the anuran organism during metamorphosis. Three parts are highlighted: the couple TR/TH at the center of the pathway is upregulated at the onset of metamorphosis. More upstream in the pathway takes place TH synthesis, and more downstream, the expression of TR target genes is modified upon TR activation. Since TR (more precisely the gene TRβ) itself is one of its own target genes, its autoactivation is critical for the trigger of the climax of metamorphosis. TR also regulates the production of TH, thus contributing to the climax. [Color figure can be viewed in the online issue, which is available at www.interscience.wiley.com.]

is particularly important, because it transmits the signal of the peak of TH: as TRβ is one of its own target genes, the peak of TH production activates a positive loop of TRβ production and activation characteristic of the metamorphosis climax (Tata *et al.*, 1993).

In actinopterygian fishes, the molecular pathway controlling metamorphosis has been mainly studied in flatfishes (pleuronectiformes) such as flounders and turbot in which metamorphosis is spectacular, as in anurans. In these species, a symmetrical planktonic larva transforms into an asymmetric benthic adult that harbors two eyes on the same side of the body [(Nelson, 2006), p442] (Fig. 3d). The data obtained from these species took advantage of the knowledge accumulated in anurans as a starting point, and confirmed the central role of THs and their receptor in the early regulation of metamorphosis (Marchand *et al.*, 2004; Power *et al.*, 2001; Yamano and Miwa, 1998). A schematic structure of the TH signaling

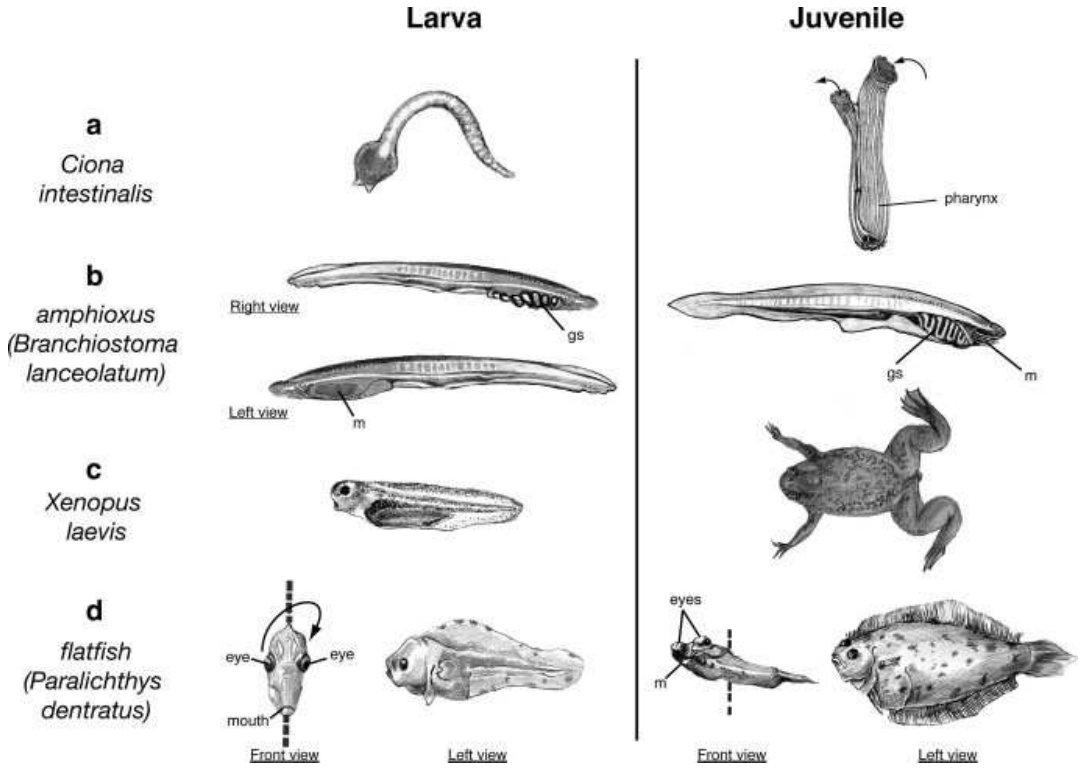


FIG. 3. Comparative drawing of larvae and juveniles from diverse Chordate groups. (**a**) Among remarkable events, the *Ciona intestinalis* larva changes its body plan. The two arrows indicate the water flow through the two siphons in the adult. (**b**) The amphioxus larva displays asymmetric features (e.g., the mouth is located on the left side of the body and the round gill slits are only on the sight side of the body) and gets more symmetric during metamorphosis (e.g., the mouth migrates to a more ventral part and the gill slits get a U shape and migrate to be on both sides). The amphioxus larva is presented from both right and left views. (**c**) The tail of *Xenopus laevis* regresses when limbs develop. (**d**) In the flatfish *Paralichthys dentratus*, the left eye of the symmetric flatfish larva migrates to the right side of the body. The larva and the adult are represented from two angles (front view + left view). The dashed line represents the larval plan of symmetry. An arrow indicates the migration itinerary of the right eye. gs: gill slits, m: mouth.

pathway involved in metamorphosis can be drawn from those studies in which the TH/TR couple appears to be the key component triggering metamorphosis in anurans and flat fishes (Fig. 2). Upstream TH metabolism regulates TH production and availability, therefore TR activity; more downstream target genes whose expression is regulated by the coupled TH/TR control the morphological remodeling leading to metamorphosis (Fig. 2).

Because metamorphoses in chordates are morphologically very diverse, several authors proposed that metamorphosis has appeared independently several times in the chordate lineage (Bishop and Brandhorst, 2003; Heyland *et al.*, 2005; Nielsen, 1998; Sly *et al.*, 2003). Recent data have nevertheless shown that, in a number of different lineages of chordates, THs and other thyroactive compounds are playing a key role in the control of metamorphosis, as reviewed above for amphibians and actinopterygians,. In this review, we highlight the common features uniting all metamorphoses in chordates and we propose a unique ancient origin for this key postembryonic developmental process based on the conservation of its molecular determinism. More precisely, we propose that the chordate ancestor had an indirect development and metamorphosed under thyroactive compound stimulation. This hypothesis unifies all the postembryonic developmental stages regulated by THs and their derivatives since it implies that they are all homologous. This model also provides an evolutionary framework through which metamorphosis has diversified within chordates. We speculate that this hypothesis can even be extended to the deuterostome phylum, although key data are still missing to unambiguously demonstrate that TH-regulated metamorphosis is shared by all deuterostomes.

To discuss this hypothesis, we will tackle several questions: (i) What are the animals that metamorphose and what are the main events occurring during the larva-to-adult transition? (ii) To what extent are thyroactive compounds available in chordates? Is TR involved in its action? (iii) How does evolution of TH signaling constrain the evolution of metamorphosis in chordates? We will show that, although the effects of THs and their derivatives during late development vary between chor-

date taxa, these differences should be considered as derived rather than ancestral.

DIVERSITY OF CHORDATE METAMORPHOSES

In the two classical anuran models *Xenopus laevis* and the bullfrog *Rana catesbeiana*, an aquatic and herbivore tadpole is transformed into a terrestrial and carnivorous predator juvenile. Loss of the gill slits, shrinkage and disappearance of the tail, development of limbs and lungs, intestine and eye modifications but also switch from larval to adult haemoglobin type are among the most obvious events that can be described (for review, see Brown and Cai, 2007; Gilbert *et al*., 1996; Shi, 2000; Tata, 2006). Although a classical textbook example, the tadpole to frog transition is not necessarily representative of the variety of metamorphoses one can observe in chordates (Fig. 3).

The other tetrapods (mammals and sauropsids) are generally considered not to metamorphose. Most actinopterygian fishes have a metamorphosis characterized by an ecological shift crucial for the geographic dispersal of the species, and therefore have been very intensively investigated for economic purposes (Cowen *et al*., 2006). For example, most marine fish larvae (such as the flatfishes discussed above) are planktonic and transform into juveniles that either swim in the water column or are benthic (Leis, 2006). In freshwater populations, metamorphosis is also morphologically obvious, as seen in the alteration of the zebrafish fins and gut (Brown, 1997), or salmonid smoltification (Varsamos *et al*., 2005).

At the base of the vertebrate clade in lampreys, the larva dwells in the silt of freshwater streams and filter feeds on primary detritus. Larval periods lasts several years after which lampreys metamorphose into adults, some of which are (sometimes parasitic) predators. During metamorphosis, several structures are remodeled: the eyes and the gut epithelium are modified, the larval kidney is replaced with a juvenile one, the mouth develops into a specific sucker apparatus, and the endostyle is profoundly remodeled into a thyroid gland (Youson, 1997). The members of the sister group myxine apparently have a direct development, although their postembryonic development remains to be fully studied (Jørgensen, 1998).

Outside vertebrates, dramatic morphological remodeling occurs in some urochordates, where the very *bauplan* of the organism is modified between larvae and adults, like in ascidians (Fig. 3a). Ascidians are the most studied and the largest taxonomic group in urochordates, with *C. intestinalis* as a genomic and developmental model system. The larva is pelagic and "tadpole-like" with a typical chordate body plan (a tail—"uro"—with a notochord, a dorsal nerve tube, striated bilateral segmented musculature), whereas the adult is sessile with a different body plan. *C. intestinalis*, for instance, has a "sac-like" body with no tail, a hypertrophied pharynx, and two siphons allowing water flux (Fig. 3a). This remodeling is so drastic that some zoologists, including Linnaeus, did not relate the adult to the larva and classified the former as a mollusk. However, other ascidian species undergo a slightly less dramatic transition as they do not lose their tails during metamorphosis (Jeffery *et al*., 1999). In the other two classes of tunicates, Larvacea and Thaliacea that are both planktonic in the adult stage, the situation is different. In Larvaceans, the adult keeps a tadpole-like body plan and can thus be viewed as paedomorphic. In contrast, in Thaliacea no clear larval stage is visible, a situation reminiscent of a direct development strategy (Brusca and Brusca, 2003).

Amphioxus undergoes a very different metamorphosis. The most conspicuous change is the transition from a pelagic asymmetric larva to a benthic and more symmetric juvenile: the left mouth migrates to an antero-ventral part when the gill slits duplicate and split up from the right side in the larva to both sides in the juvenile (Fig. 3b) (Stokes and Holland, 1995; Willey, 1894). Some other amphioxus species from the genus *Asymmetrons* and *Epygonichthys* (Kon *et al*., 2007) stay asymmetric even as adults. However, very little is known about their postembryonic development.

Overall, this rapid zoological presentation shows that most chordates display an ecological larva-to-adult transition, characterized by an adaptation from the larval ecological niche to the adult one. Interestingly, the intestine is in most (if not all) cases remodeled to adapt to the new feeding mode. However, these modifications vary between one taxon to another to the point that their common ancestry seems quite difficult to reconcile. In contrast, the molecular events triggering metamorphosis exhibit interesting similarities that reveal a common origin. In the next section, we will briefly establish that all chordates can produce THs and then discuss the role of THs in the evolution of metamorphosis.

CONSERVATION OF THE TH SIGNALING PATHWAY IN CHORDATES

In Chordates TH Production Takes Place in the Specialized Organ Thyroid/Endostyle

All chordates harbor a TH-producing gland, that is found in no other taxon, and which thus constitutes a chordate synapomorphy (Fig. 1). It is an endoderm (gut) derivative located in the pharyngeal region, called thyroid gland in vertebrates and endostyle in invertebrate chordates. The main difference between the thyroid gland and the endostyle (which justifies the use of different names) is the follicular organization of the TH-producing cells in the thyroid gland and not in the endostyle (Fredriksson, 1988).

Some morphological discrepancies exist between the thyroid glands of different vertebrate groups, due to differential clustering of the numerous follicular units: for example, the thyroid gland is in one piece in mammals but birds and amphibians have two separated lobes on each side of the trachea. In most fishes and in agnathans, the follicles are not aggregated in an organized thyroid

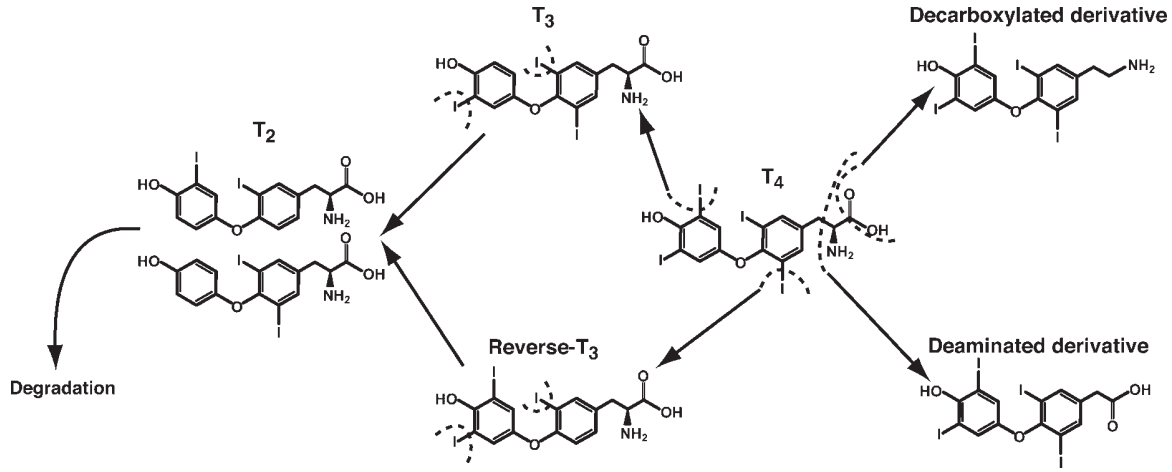


FIG. 4. Pathways involved in the formation of various TH derivatives. The structure of the TH derivatives is indicated. The diverse modifications of T3, reverse-T3 and the precursor T_4 are highlighted with dashed lines.

gland and are scattered in the pharyngeal region (Braverman and Utiger, 1996). However, as the functional unit of the thyroid is the follicle, the macroscopic shape of the organ is probably not of fundamental functional and evolutionary significance.

Only rather recently has the homology between the thyroid gland and the endostyle been clearly established. More precisely, only pieces of the endostyle are thought to be homologous to the thyroid follicles (Fredriksson, 1988, p8). Indeed, some parts of the endostyle secrete a mucoprotein that helps gathering food particles from water and are specific to endostyles (Fredriksson, 1988; Ogasawara, 2000). It is the mucus nonproducing, smallest, most dorsal and posterior part that is thought to be the homologue of the thyroid follicles. This hypothesis was based on several arguments: (i) the endostyle and the thyroid gland are located in a similar position in the body plan, (ii) some endostyle cells have thyroid-producing properties such as iodine uptake (Fredriksson *et al.*, 1985; Manzon and Youson, 2002; Thomas, 1956) and TH metabolism (Barrington and Thorpe, 1968; Covelli *et al.*, 1960; Kluge *et al.*, 2005; Manzon and Youson, 2002), (iii) several thyroid-specific genes are expressed during endostyle formation [*Ttf1* in lamprey (Kluge *et al.*, 2005)], and (iv) in lamprey, some cells of the endostyle transform into thyroidal follicles during metamorphosis (Fredriksson, 1988).

It has been proposed several times (Ogasawara, 2000; Sembrat, 1953) and recently reactualized (Paris *et al.*, 2008) that the common ancestor of chordates had an endostyle, quite like the extant ones: this endostyle had two functions: (i) the production of mucus for food collecting, and (ii) TH synthesis (Ogasawara, 2000). The ancestor of all chordates most probably lived in seawater and was filter feeding. Among the species that derived from this ancestor, some became nonfilter-feeding animals and thus did not need mucus production anymore. This latter function subsequently degenerated and the endocrine function was kept only in vertebrates (Fredriksson, 1988). Lamprey development provides a nice living example of this scenario since, at metamorphosis, lamprey switches its mode of feeding from filter feeding to macrophagy (or nonfeeding) in parallel to the disappearance of the endostyle and the formation of a thyroid with a typical vertebrate organization.

Diversity of Thyroactive Compounds

THs and other iodinated tyrosine derivatives were discovered as thyroid gland products (hence their name). They have been detected in the plasma of all vertebrates that have been tested (for detailed values, see Hulbert, 2000). In mammals and amphibians, the two main thyroid products are the stock hormone T_4, or l-thyroxine, and the more active T_3, or 3,3′,5-triiodo-l-thyronine (iodinated 4 and 3 times, respectively) (Fig. 4). However, other kinds of related molecules exist depending on the number and positions of the iodines. Among such compounds, 3,5-T_2, 3,3′-T_2 or 3,3′,5′-triiodo-l-thyronine (Reverse-T_3, see Fig. 4) are mildly active TH derivatives in mammals (Flamant *et al.*, 2006). Decarboxylated or deaminated analogues have also been described in mammals and in anurans (Fig. 4) and were shown to be active: for instance TRIAC, the T_3 deaminated derivative displays strong themogenic effects and has a higher affinity for TR than T_3. This compound is less stable than T_3, but its production represents 14% of T_3 metabolism (Moreno *et al.*, 2008; Wu *et al.*, 2005). Nonetheless, the general biological action of those TH derivatives remains vague (Wu *et al.*, 2005). In other chordates, no exhaustive repertoire is available. This is mainly explained not only by anthropocentric reasons [and the tendency to look for what is similar/different compared to us; see (Markov *et al.*, 2008) for a discussion of this aspect], but also by the fact that such comprehensive analyses are much more demanding than simple two-by-two compari-

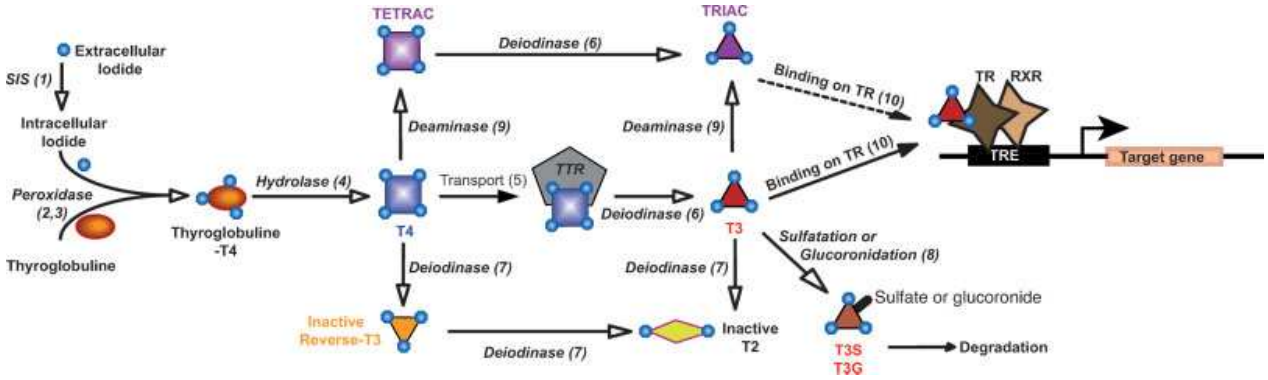


FIG. 5. Thyroid hormone signaling pathway. The different steps from iodine concentration to TH binding to TR are indicated. Numbers located between brackets refer to the main text.

sons with vertebrates. The evolutionary relevance of an expanded iodinated tyrosine derivatives repertoire, larger than simply the THs T_3 and T_4 in chordates, will be discussed later.

TH and TH Derivative Synthesis in Tetrapods

In the amphibian X*enopus laevis*, in mouse and in human, TH production looks quite similar. The most active TH is T_3 and, to a lesser extent, its precursor T_4. A compound is called active when its binding to TR stabilizes a conformational change of the receptor, that leads to the recruitment of coactivators (like SRC1) that are required for the subsequent transcription of target genes (Flamant *et al*., 2006). Note that in some cases, the binding of the hormone induces gene repression probably via corepressor recruitment (Guissouma *et al*., 2005).This review does not intend to be exhaustive about TH signaling in mammals or in amphibians, so for further information see (Yen, 2001; Zhang and Lazar, 2000) for general reviews on TH action and more specifically (Brown and Cai, 2007; Buchholz *et al*., 2006; Tata, 2006) about amphibians.

TH synthesis in the thyroid gland takes place in the follicular cells called thyrocytes and in the lumen enclosed by the thyrocytes. This lumen is full of the protein thyroglobulin. TH production can be decomposed into several intermediate steps, each of them is illustrated in Figure 5 and briefly described hereafter: (1) Iodine first needs to be concentrated by the specific transmembrane Sodium/Iodine Symporter. (2) Then it is added to tyrosines from the matrix protein thyroglobulin by the thyroid peroxidase (TPO). (3) These iodinated tyrosines are coupled by the TPO to form a TH (mostly T_4 but also T_3). (4) The newly synthesized TH is then separated from the matrix protein, (5) and transported in the blood by various carriers such as transthyretin (TTR). (6) In peripheral tissues, T_4 is deiodinated into the more active form T_3. (7) Both THs can be converted into less active forms, which will be subsequently degraded. (8) There are also alternative TH pathways, like sulfatation and glucuronidation, mostly resulting in production of inactive T_3 derivatives, targeted for degradation. (9) However, some alternative active forms are also produced, like the deaminated form of T_3, TRIAC, which is a very potent, but very unstable TH derivative (see Wu *et al*., 2005) (Figs. 4 and 5). (10) Finally, within the cell, upon thyroactive compound binding, TR will modulate the expression of target genes. The direct synthesis of T_3 in the thyroid gland, along with the peripheral deiodination of T_4, is an important source of active hormone since cellular uptake is faster for T_3 than for T_4 (Everts *et al*., 1996).

TH and TH Derivative Synthesis in Other Chordates

The search for TH derivatives in a wide variety of chordates was intense during 1950s and 1960s, since the precise TH metabolism in those animals was not known (Barrington and Thorpe, 1963; Covelli *et al*., 1960; Gross and Pitt-Rivers, 1953a,b; Tong *et al*., 1961). These experiments were facilitated by the fact that TH synthesis is dependent on exogenous iodide, which has the great advantage of being very specific to THs. Consequently, the detection of iodine characterizes and localizes TH synthesis. Moreover, iodide is highly reactive, and radiolabeled iodide has been used for a long time as a marker of endogenous TH synthesis.

THs were found in all chordate animals that were searched for them. T_3 and T_4 (as well as their precursors MIT and DIT) were detected in amphioxus (Covelli *et al*., 1960; Tong *et al*., 1962), *C. intestinalis* (Barrington and Thorpe, 1965; Patricolo *et al*., 2001), lampreys (Manzon and Youson, 1997), sharks (McComb *et al*., 2005), and lungfishes (reported in Joss, 2006). TH metabolism was also found in all chordates examined: peroxidase-dependent iodine fixation (Manzon and Youson, 2002; Tong *et al*., 1961), peroxidase activity (Eales *et al*., 1997; Kobayashi *et al*., 1983; Tsuneki *et al*., 1983), deiodinase activity (Shepherdley *et al*., 2004; Sutija *et al*., 2003, 2004), and a thyroglobulin-like protein were also reported in lamprey, hagfish, and amphioxus (Manzon and Youson, 2002; Monaco *et al*.,

1978, 1981). Accordingly, most orthologs of the genes involved in tetrapod TH metabolism were found in amphioxus and *C. intestinalis* genomes (Holland *et al*., 2008).

It is important to keep in mind that TH metabolism is not limited to T_3 and T_4 in mammals or amphibians: other TH derivatives are produced, like deiodinated, decarboxylated or deaminated derivatives, as well as sulfated and glucuronidated conjugates (Wu *et al*., 2005) (Figs. 4 and 5). Similarly, it is most probable that other chordates have a richer iodinated tyrosine derivatives repertoire than just T_3 and T_4. For instance, amphioxus synthesizes deaminated TH derivatives (Covelli *et al*., 1960) and it is probably a T_3 derivative, and not T_3, which is the active form in amphioxus (Paris *et al*., 2008). This variability can be of significant evolutionary importance and should not be dismissed.

TH and TH Derivative Transport in Chordates

THs are either secreted from the thyroid/endostyle or incorporated through the gut, so they have to reach their target cells. Benzene rings in the tyrosine structure increase the tendency of THs to partition in plasma membranes. Consequently, only a minute fraction of TH is "freely" transported and most TH travels bound to a "carrier" protein. Three main TH plasma carriers are known in mammals: Thyroxine binding globulin (TBG), TTR, and albumin. These transporters not only have various binding affinities to THs (TBG has the highest affinity—around 10^{-10} M, when the albumin has the lowest—around 10^{-5} M), but also various plasma concentrations, which allows a finely tuned TH delivery to the different organs.

Generally, in vertebrates, TH plasma transport is similar, although the relative importance of the different carriers varies between species (e.g., TBG is the main T_4 carrier in human instead of TTR in rodents) (Power *et al*., 2001). In lamprey, one TTR was recently cloned from two species, *P. marinus* and *L. appendix* (Manzon *et al*., 2007), and expression correlated with metamorphosis. This suggests a deeper ancestry for TTR than previously thought (Eneqvist *et al*., 2003). In amphioxus and in *Ciona*, virtually nothing is known about TH transport. One gene related to TBG (but no ortholog to TTR) was found in the amphioxus genome (Brunet F., personal communication). It will be interesting in the future not only to look for TTR in other genomes, like sea squirt or sea urchin, but also to clone and functionally characterize the expression of these genes found in the genomes.

Once distributed close to each cell, THs need to pass the plasma membrane. As other nuclear receptor ligands, these hormones were first thought to passively enter the cell. In fact, cellular TH uptake is now known to be facilitated by organic anion transporters, including organic anion transporters like NTCP (sodium taurocholate cotransporting polypeptide) or OATP1 (Na-independent organic anion cotransporting polypeptide 1), and amino acid transporters like MCT8 (monocarboxylate transporter 8) (Friesema *et al*., 2005). Transporters display various tissue and ligand selectivity, allowing a tight refinement of TH availability in cells (Friesema *et al*., 2005). Unfortunately, due to the relative youth of this field, virtually nothing is known about cellular transport outside of mammals. Moreover, the molecules are apparently not specific to THs and can also transport other amino-acid derivatives.

Even inside the cell, free TH levels are controlled. For instance, CTHBP (Cytosolic Thyroid Hormone Binding Protein) is a cytosolic protein that binds T_3, and thus is likely to modulate the availability of T_3 to TR in the nucleus (Shi, 2000). One gene orthologous to CTHBP was found in the amphioxus genome (F. Brunet, personal communication), but its function has not been investigated.

The tissue- and cellular-specific expressions of the genes involved in TH transport as well as deiodinases can influence the timing and localization of TH action, and therefore modulate metamorphosis. However, almost no data are available outside amphibians and/or mammals about such genes, which makes their evolutionary study in a TH context more complicated.

TR Is the TH Receptor in Chordates

The mode of action of TR as an iodinated tyrosine derivative-dependent transcription factor has been well studied in chordates. The two "classical" vertebrate TRs, TRα and TRβ, that have diverged from the vertebrate-specific genome duplication (Escriva *et al*., 2002) have been quite extensively studied in mammals and amphibians (Buchholz *et al*., 2006; Flamant *et al*., 2006). They were shown to preferentially bind T_3 and to be the main mediators of T_3 action. The same conclusions were drawn for actinopterygian fishes: their TRs are high affinity T_3-dependent transcription factors. This was shown both in flatfishes that exhibit drastic metamorphic changes (Marchand *et al*., 2001, 2004), as well as in some other fishes (Essner *et al*., 1999; Yamano and Miwa, 1998). In the most basal chordate lamprey, the two known TRs are not orthologous to TRα and TRβ, and instead are derived from a lamprey-specific duplication (Escriva *et al*., 2002). Both receptors are activated by T_3 (Paris *et al*., 2008), like other vertebrate TRs.

The conservation of TR function, as a TH derivative-dependent transcription factor can be extended to chordates. Indeed, the only TR recently cloned from amphioxus was shown to be activated not by T_3, but by its T_3 derivative TRIAC (Paris *et al*., 2008). Consequently, the genome of the common ancestor of amphioxus and vertebrates (i.e., the chordate ancestor) probably already contained a gene encoding a TH derivative-receptor, although the nature of the ancestral active compound is more difficult to assess. Interestingly, the unique *Ciona* TR does not bind any TH derivative tested so far (Carosa *et al*., 1998; Paris *et al*., 2008). As it has a highly diver-

gent sequence compared to other chordate TRs, its unresponsiveness to such compounds is probably derived.

Alternative mechanisms may also be involved in TH action in chordates. Particularly, nongenomic TH effects were reported in mammals, involving TH receptors other than TR (Davis *et al*., 2007; Scanlan *et al*., 2004). However, the mode of action of those receptors is not well understood yet (Apriletti *et al*., 1998; Scanlan *et al*., 2004). It is therefore still very difficult to extrapolate on the function of those genes outside mammals. We can conclude from these studies that TH derivatives are produced by all chordates, and that TH and a TH derivative signal is predominantly mediated by TR. Consequently, all chordates have the basic toolkit to trigger and regulate metamorphosis using TH signaling. But is this what occurs?

EVOLUTION OF METAMORPHOSIS IN CHORDATES

The Evolution of Metamorphosis Is Governed by the Evolution of the TH Signaling Pathway

Our hypothesis that metamorphosis is an ancestral feature of chordates is supported by the molecular determinism of this developmental process. Indeed THs, other iodinated tyrosine derivatives and their receptor TR have been systematically involved in metamorphosis in all chordates in which their role has been tested. We thus propose that in most chordates studied to date, the onset of metamorphosis is characterized by a peak of a thyroactive compound, activating TR that modifies the expression of target genes and leads to morphological remodeling characteristic of the larva-to-juvenile transition.

We argue that, in all chordates, the key component of TH signaling pathway involved in metamorphosis is the coupled TH/TR (Fig. 2). Thus, TR activation by TH (or a more active TH derivative) can be viewed as the master switch controlling the postembryonic larva-to-juvenile transition in all chordates. In this scenario, the coupled TH/TR is the most central part of the TH-metamorphosis network, and thus the most conserved. In contrast, the more upstream part (composed of genes involved in TH metabolism that regulate TH availability and TR activity) and the more downstream part (mainly TR target genes, whose expression is regulated by the coupled TH/TR and which control morphological remodeling leading to metamorphosis) are less constrained and more prone to diverge (Fig. 2). This model suggests the existence of a postembryonic period during which TH/TR controls a key transition, molecularly conserved but morphologically extremely diverse.

The evolution of metamorphosis, and thus its wide diversity at the morphological level, results from the evolution of TH signaling. Modifications of phenotypes linked to the larva-to-adult transition are allowed by alterations of the TH signaling pathway, preferentially acting upon the thyroactive compound availability or the target genes, rather than on thyroactive compound as the trigger or TR as the mediator of that initial signal. Before we discuss this evolutionary potential in various chordates, let us examine the data that support this view of an ancient pathway acting in all chordates.

Metamorphosis Is Regulated by THs and/or TH Derivatives in Basal Chordates

In amphioxus, because of technical difficulties in handling larvae until metamorphosis, very few experiments have been performed to understand how metamorphosis is regulated. However, very recently various iodinated tyrosine derivatives were shown to control metamorphosis in amphioxus, as in other chordates (Paris *et al*., 2008). The importance of TR in the triggering of metamorphosis was demonstrated by a pharmacological approach, since a specific TR antagonist inhibited both spontaneous and thyroactive compound-induced amphioxus metamorphosis, as in anurans.

As metamorphosis is regulated by THs and other iodinated tyrosine derivatives and TR in animals located at very distant branches of the chordate phylogenetic tree (amphibians, actinopterygian fishes, and amphioxus), parsimony reasoning suggests that metamorphosis was already regulated by a thyroactive compound and TR in the common ancestor of all chordates. Consequently, the different larva-to-adult transitions described above should be considered as homologous, based on the conservation of their shared molecular determinism (and more precisely, the central part of the TH signaling pathway formed by the coupled thyroactive compound/TR, as represented in Fig. 2). More generally, our model assumes that all chordates metamorphose, with an intensity that varies quite extensively between taxa. The differences that exist between those developmental processes should be seen as derived rather than ancestral. We will thus now discuss some cases of variability of metamorphosis, based on alterations of the TH signaling pathway at different levels.

Heterochrony and Evolution of Metamorphosis in Amphibians: Alteration of the TH Signaling Pathway

In amphibians, metamorphosis can be spectacular, as seen in the anuran *Xenopus laevis*. Still, amphibians exhibit a wide range of variations in the intensity of the morphological remodeling during metamorphosis. The modifications can be of mild magnitude, as for urodeles and apodes, or it can be reduced to the point that it becomes unnoticed, as in neotenic amphibians. On the other hand, the larval stage can be suppressed, as in direct developing frogs, where little frogs directly hatch. However, as will be exemplified, in all cases there is a TH sensitive period of morphological remodeling.

TH signaling was studied in the entirely terrestrial and direct developing frog *Eleutherodactylus coqui*, in an attempt to understand the molecular basis of the appa-

rent suppression of the larval period. It was shown that *E. coqui* does metamorphose under TH stimulation, but the main difference with "classic" amphibian metamorphosis is the very early triggering of this metamorphosis, as it occurs inside the egg (Callery and Elinson, 2000). More precisely, during late developmental stages, there is a TH-dependent period of morphological remodeling (including cartilage remodeling, tail regression or limb elongation) correlated with a peak of TRβ expression (Callery and Elinson, 2000). Consequently, *E. coqui* metamorphosis is rather similar to the "classical anuran", except that the timing of TH production and TR expression have probably been shifted earlier during development. It is also possible that the TR target genes have diverged as well, since the morphological remodeling is not as drastic as in anurans. *E. coqui* development can be seen as a remarkable documented example of heterochrony (species-specific variation in the timing of the development, leading to changes in the phenotype of the species).

Alternative alterations of TH signaling lead to metamorphoses that are very mild and make adults look too much like the larvae to be described as metamorphic. These animals are called paedomorphic, because the adults have retained larval traits, and constitute another case of heterochrony. There are several cases of paedomorphosis: In facultative paedomorphic species, like the Mexican axolotl (*Ambystoma mexicanum*), the tadpole responds to exogenous TH by metamorphosing, but does not spontaneously metamorphose (Voss and Shaffer, 1997). In fact, the axolotl is a spectacular example of a series of species of the genus *Ambystoma* in which several types of unusual postembryonic life history strategy occur (Voss and Shaffer, 1997). Facultative neoteny has been attributed especially to low levels of circulating THs due to weak activation of the thyroid axis by the pituitary gland, when the rest of TH signaling pathway is "normally" functional (Boorse and Denver, 2002; Safi *et al*., 2004). Obligatory neotenic amphibians (such as the perennibranchiates Urodeles *Necturus* or *Proteus*), long thought to be insensitive to THs, were in fact shown to respond to exogenous TH but with limited amplitude: for instance, expression of the typical metamorphosis genes TRβ and stromelysin are induced by TH (Safi *et al*., 2006). Rather, the lack of response to TH is probably caused by the loss of TH regulation by some target genes such as collagenase 3.

These examples of "alternative" metamorphoses illustrate the important plasticity of the metamorphosis process in different amphibians. Each time, TR was shown to be functional (Safi *et al*., 2004, 2006). Rather, the genetic variations thought to be responsible for phenotypic differences rely on the production of TH (an early production in *E. coqui*, or a default of TH production in the axolotl) or on the target genes (as in the *Necturus*), or both. These extreme examples show that metamorphosis triggered by TH is always present in amphibians, even if in some cases it is weakened or shifted during late development stages.

Evolution of Metamorphosis in Other Vertebrates: Why Direct Development Is Not So Direct in Chordates

We argued that metamorphosis is ancestral in the chordate lineage. Consequently, all descending species should have divergent variants of this developmental stage. According to our definition of metamorphosis as a postembryonic developmental stage characterized by a thyroactive compound-dependent morphological remodeling, groups that were previously thought to have a direct development may in fact be seen as animals with indirect development and a mild metamorphosis. Indeed, similarly to neotenic amphibians having a mild metamorphosis, mammals metamorphose. In these lineages, the metamorphosis stage does not coincide with birth since the latter is not marked by an intense TH signaling. Sexual maturity does not qualify as metamorphosis either since it is regulated by steroids and is often asynchronous with metamorphosis. On the contrary, mammalian weaning gathers all characteristics of metamorphosis as defined above as a period of tissue remodeling governed by TH (Fraichard *et al*., 1997). Indeed, it is an ecological shift that is characterized by modifications like neural maturation, intestine remodeling, or chondrogenesis that are reminiscent of similar events during amphibian metamorphosis. These modifications occur in correlation with a peak of TH production (both T_3 and T_4) (Hadj-Sahraoui *et al*., 2000), and were shown to be under the control of TH and TR (Flamant *et al*., 2002; Gauthier *et al*., 1999; Plateroti *et al*., 1999). Consequently, weaning should be viewed as homologous to classical metamorphosis even if the precise role of THs in controlling this event and the regulatory hierarchy controlled by these hormones should be better analyzed. In addition, weaning is poorly studied outside classical mammalian models (e.g., mouse, rat, human) and it should be analyzed in a wider taxonomic sampling.

As discussed above, in actinopterygians, and especially in flat fishes, THs were shown to be important regulators of metamorphosis (Power *et al*., 2001), the best known case being flounder, for which T_4 and T_3 were reported to induce precocious metamorphosis (Inui and Miwa, 1985). But the role of THs in postembryonic development has also been documented in other groups. In zebrafish, THs were shown to be important for the larva-to-juvenile transition: exogenous THs induce differentiation of the pectoral fin precociously, in contrast, the inhibition of TH synthesis blocks it (Brown, 1997). Smoltification in salmonid fishes is considered to correspond to metamorphosis: it is a period of high T_4 and T_3 levels (Boeuf and Gaignon, 1989; Young *et al*., 1989) involving a morphological, physiological, and behavioral switch from a freshwater-adapted form to a salt-water-adapted form. At least some of the former modifications were shown to be TH-sensitive (Leloup and Lebel, 1993; Miwa and Inui, 1985). The larval to adult transition is extremely important in terms of recruitment in many fish species that are living close to the substratum, such

as coral fishes. Although a direct role of THs in regulating metamorphosis has not yet been clearly demonstrated, in some species like the grouper, existing evidence clearly points in this direction (de Jesus *et al*., 1998).

In lamprey, although it is now well established that TH endogenous synthesis needs to be tightly regulated for the proper metamorphosis, the role of TH displays a remarkable difference compared to other chordates. Indeed, several reports clearly show that it is a drop, instead of an increase, in circulating TH levels that ends the larval period (Manzon and Youson, 1997, 1999, 2002; Manzon *et al*., 1998, 2001; Youson, 1997; Youson *et al*., 1997). Among other factors, temperature and growth advancement (indicated by length and weight, for instance) are important parameters: lampreys enter metamorphosis when they have reached a certain size (Youson, 1997). Interestingly even if the reversion of TH action is unique to lampreys, TR is still likely to play a role in this TH-mediated action since nuclear TH binding was reported (Lintlop and Youson, 1983), and the TRs from lamprey are activated by TH (Paris *et al*., 2008). Given that it is a drop of TH levels that triggers metamorphosis, TR's role in controlling metamorphosis must be different in lamprey than in other vertebrates. It is possible that in lamprey, TR induces the expression of genes that inhibit metamorphosis. According to this scenario, when TH levels drops, TR will stop inducing the expression of these inhibitors, then triggering metamorphosis. However, many other hypotheses can be proposed, since in vivo functional experiments are still to be completed. The use of specific TR agonists and antagonists (Chiellini *et al*., 1998; Lim *et al*., 2002; Nguyen *et al*., 2002) could help deciphering the specific role of TR in this inverted situation.

Another interesting case of divergence on the TH signaling pathway is provided by urochordates. Metamorphosis of *C. intestinalis* can be induced by exogenous THs and inhibited by goitrogens known to block TH synthesis in vertebrates (Patricolo *et al*., 1981, 2001). However, the molecular pathways triggered by THs remain elusive since the only TR of *C. intestinalis* (Dehal *et al*., 2002) does not bind any iodinated tyrosine derivative tested (Carosa *et al*., 1998; Paris *et al*., 2008). The urochordates are particularly interesting regarding the evolution of metamorphosis because of the diversity of their postembryonic developmental strategies. The elucidation of the role of THs and iodinated tyrosine derivatives in general in late development in other species, and in particular in larvaceans should bring interesting data for comparative analysis.

Modulation of TH Availability

The plasticity of the TH pathway allows evolution of the regulation of metamorphosis within the chordates. Indeed the action of thyroactive compounds, mediated by TR, is dependent on the availability of the compound to each cell, as well as on the list of potential target genes. In amphibians and mammals, this control is observed at several levels: global TH level is regulated by the hypothalamo-pituitary axis, while the local levels are tightly regulated by membrane transporters (Friesema *et al*., 2006) and deiodinases (Brown, 2005). The situation is potentially similar in all chordates, where some orthologous genes to TH transporters have been isolated (like the *Ciona* orthologous gene to Oatp14, involved in transport of T_3 through the blood–brain barrier) (Chambon *et al*., 2002, 2007). Of course, the expression pattern of TR is a central parameter, but the role of its partners (coactivators and corepressors) should not be neglected. A crucial aspect in the diversification of the metamorphosis pattern observed in vertebrates is the plasticity of the TH-regulated gene network. In *Xenopus*, it has been shown that THs control a specific gene regulatory hierarchy through TR activation. It is likely that, as observed in obligate paedomorphic species, this gene network is modified in different species, resulting in a distinct outcome in terms of metamorphosis. This importance of the gene regulatory cascade controlled by a relatively invariant master regulator is reminiscent of the situation observed in other cases such as eye development [with pax6 being the master gene (Kozmik, 2008)] or antero-posterior development of arthropods [with Hox genes being the key selector genes (Pearson *et al*., 2005) and more generally (Carroll *et al*., 2005)]. Indeed, TH-regulated gene expression pattern is highly variable from one species to another, as shown in *X. laevis* and the mouse (Das *et al*., 2006; Yen *et al*., 2003). It should be interesting to compare gene expression in other vertebrates. One of the only common pathways is the TH-induced positive upregulation of one of the TR genes occurring during amphibian, fish, or amphioxus metamorphosis (Marchand *et al*., 2004; Paris *et al*., 2008) and mammalian weaning (Flamant *et al*., 2002; Mochizuki *et al*., 2007) in concert with a peak of TH production (Flamant and Samarut, 2003). The TR gene activated in these various groups is not the same though: TRβ is the key target in Xenopus, whereas in actinopterygians and mammals TRα appears to be more important.

CONCLUSIONS

In this review, we proposed a new definition of metamorphosis highlighting the role of THs and TH derivatives as the triggering element and TR as the direct mediator of TH signal in chordates. This definition is based on the hypothesis of an ancestral origin of TH and thyroactive compound-induced metamorphosis in chordates. From this definition, all chordates metamorphose under thyroactive compound stimulation. We postulate that the impressive diversity of metamorphic phenotypes is constrained by the divergence of the TH signaling in the chordate groups. One should keep in mind that this model is based on data accumulated in a few model species that represent a very restricted taxonomic sampling of vertebrates. Weaning in mammals, metamorphosis in

urochordates or larval to adult transitions in actinopterygian fishes (not to mention the postembryonic development in sauropsids [that include birds, lizards, snakes, and turtles, (Fig. 1)] are known from work on very few model systems. It will be very important in the future to study the role of THs and their derivatives and TR in the coordination of this neglected period of the vertebrate postembryonic development. We nevertheless believe that the model proposed in this article offers a coherent view of chordate postembryonic development in general and of chordate metamorphosis in particular.

The role of THs in metamorphosis might even be extended to all deuterostomes. Indeed, most echinoderms metamorphose, although their life history is quite variable from one species to another (Heyland *et al*., 2005). For instance, the bilaterial sea urchin larva transforms into a pentameral juvenile (e.g., *Paracentrotus* or *Strongylocentrotus*, the two classical models). Interestingly, effects of TH on the metamorphosis of echinoderm larvae have now been reported in several families, including species with feeding and nonfeeding development: Echinoidea (sea urchins, sea biscuits, and sand dollars) asteroids (sea stars) and ophiuoidea (brittle stars) [(Hodin, 2006) and references therein]. Moreover, outside chordates, echinoderms are the only animals that display clear endogenous TH metabolism: T_4 production was demonstrated in sea urchin and sea biscuits (Heyland *et al*., 2006a,b). Interestingly, orthologs of many proteins involved in mammal TH metabolism were found in the sea urchin genome (Holland *et al*., 2008; Sea Urchin Genome Sequencing *et al*., 2006). This suggests an even more ancestral origin of TH endogenous synthesis in deuterostomes. One TR was found in the genome of the sea urchin *Strongylocentrotus purpuratus* (Howard-Ashby *et al*., 2006). It is not as divergent as the TR from *C. intestinalis*, making it a potential TH receptor. It will be interesting to test in vitro if this receptor can bind any iodinated tyrosine derivative and regulate transcription of TRE-containing genes in response to this hormone. As THs seem to play a similar role in echinoderm metamorphosis as in chordates, the involvement of TR in this process would be of particular interest.

TR genes were recently described in some protostomes (mollusks, flat worms, and crustacean) (Wu *et al*., 2007), implying that the gene was present in the common ancestor of all bilaterians and was secondarily lost in insects and nematodes. It will be very interesting to decipher the function of these genes. The encoded proteins are apparently able to bind similar DNA sequences and to heterodimerize with the classical nuclear receptor partner RXR (Wu *et al*., 2007), but are they active transcription factors? More particularly, are they responsive to THs? Attempts to determine if THs had any effects on protostome development are scant and relatively preliminary. One study reported that THs accelerated metamorphosis in the mollusk abalone (Fukazawa *et al*., 2001) and another one in a crustacean [a prawn (Roustaian and Gaik, 2006)]. In contrast, THs did not seem to have any effect on the crustacean *Daphnia* (Kashian and Dodson, 2004) in accordance with the observation that no endogenous TH is present in this species (Eales, 1997). Consequently, hypotheses on the roles of THs or other derivatives and TR in protostomes can only be tentative: a potential responsiveness to THs would imply a very ancient function of TR as the TH receptor or, on the other hand, the establishment of an alternative function would imply an evolutionary shift. In any case, it is necessary to better characterize the TH signaling pathway in protostomes: are they able to produce iodinated tyrosine derivatives, and if they do so, how? Are the TRs functional? What are the molecular events triggered by iodinated tyrosine derivatives? To what extent do they regulate metamorphosis? Those questions need to be tackled, in order to fully understand the evolution of TH signaling in bilaterian and its role in late development.

ACKNOWLEDGMENTS

The authors are extremely grateful to Françoise Boussau for her contribution on the drawings of chordate larvae and juveniles. They thank Bastien Boussau for critical reading of the manuscript.

LITERATURE CITED

Allen BM. 1916. The results of extirpation of the anterior lobe of the hypophysis and of the thyoid of Rana pipiens larvae. Science 44:755–758.

Apriletti JW, Ribeiro RC, Wagner RL, Feng W, Webb P, Kushner PJ, West BL, Nilsson S, Scanlan TS, Fletterick RJ, Baxter JD. 1998. Molecular and structural biology of thyroid hormone receptors. Clin Exp Pharmacol Physiol Suppl 25:S2–S11.

Barrington EJ, Thorpe A. 1963. Comparative observations on iodine binding by Saccoglossus horsti Brambell and Goodhart, and by the tunic of *Ciona intestinalis* (L). Gen Comp Endocrinol 3:166–175.

Barrington EJ, Thorpe A. 1965. The identification of monoiodotyrosine, diiodotyrosine and thyroxine in extracts of the endostyle of the ascidian, *Ciona intestinalis* L. Proc R Soc Lond B Biol Sci 163:136–149.

Barrington EJ, Thorpe A. 1968. Histochemical and biochemical aspects of iodine binding in the tunic of the ascidian, *Dendrodoa grossularia* (van Beneden). Proc R Soc Lond B Biol Sci 171:91–109.

Bishop CD, Brandhorst BP. 2003. On nitric oxide signaling, metamorphosis, and the evolution of biphasic life cycles. Evol Dev 5:542–550.

Bishop CD, Erezyilmaz DF, Flatt T, Georgiou CD, Hadfield MG, Heyland A, Hodin J, Jacobs MW, Maslakova SA, Pires A, Reitzel AM, Santagata S, Tanaka K, Youson JH. 2006. What is metamorphosis? Integr Comp Biol 46:655–661.

Boelen A, Wiersinga WM, Fliers E. 2008. Fasting-induced changes in the hypothalamus-pituitary-thyroid axis. Thyroid 18:123–129.

Boeuf G, Gaignon JP. 1989. Effects of rearing conditions on growth and thyroid hormones during smolting of atlantic salmon, *Salmo salar* L. Aquaculture 82:29–38.

Boorse GC, Denver RJ. 2002. Acceleration of *Ambystoma tigrinum* metamorphosis by corticotropin-releasing hormone. J Exp Zool 293:94–98.

Boorse GC, Denver RJ. 2003. Endocrine mechanisms underlying plasticity in metamorphic timing in spadefoot toads. Integr Comp Biol 43:646–657.

Braverman LE, Utiger RD. 1996. The thyroid. A fundamental and clinical text. Philadelphia, NY: Lippincott-Raven.

Brown DD. 1997. The role of thyroid hormone in zebrafish and axolotl development. Proc Natl Acad Sci USA 94:13011–13016.

Brown DD. 2005. The role of deiodinases in amphibian metamorphosis. Thyroid 15:815–821.

Brown DD, Cai L. 2007. Amphibian metamorphosis. Dev Biol 306:20–33.

Brusca RC, Brusca GJ. 2003. Invertebrates, 2nd ed. Sunderland, MA: Sinauer Associates. xix, 936 p.

Buchholz DR, Paul BD, Fu L, Shi YB. 2006. Molecular and developmental analyses of thyroid hormone receptor function in *Xenopus laevis*, the African clawed frog. Gen Comp Endocrinol 145:1–19.

Callery EM, Elinson RP. 2000. Thyroid hormone-dependent metamorphosis in a direct developing frog. Proc Natl Acad Sci USA 97:2615–2620.

Carosa E, Fanelli A, Ulisse S, Di Lauro R, Rall JE, Jannini EA. 1998. *Ciona intestinalis* nuclear receptor, Part 1: A member of steroid/thyroid hormone receptor family. Proc Natl Acad Sci USA 95:11152–11157.

Carroll SB, Grenier JK, Weatherbee SD. 2005. From DNA to diversity: Molecular genetics and the evolution of animal design, 2nd ed. Malden, MA: Blackwell. ix, 258p.

Chambon JP, Nakayama A, Takamura K, McDougall A, Satoh N. 2007. ERK- and JNK-signalling regulate gene networks that stimulate metamorphosis and apoptosis in tail tissues of ascidian tadpoles. Development 134:1203–1219.

Chambon JP, Soule J, Pomies P, Fort P, Sahuquet A, Alexandre D, Mangeat PH, Baghdiguian S. 2002. Tail regression in *Ciona intestinalis* (Prochordate) involves a Caspase-dependent apoptosis event associated with ERK activation. Development 129:3105–3114.

Chiellini G, Apriletti JW, Yoshihara HA, Baxter JD, Ribeiro RC, Scanlan TS. 1998. A high-affinity subtype-selective agonist ligand for the thyroid hormone receptor. Chem Biol 5:299–306.

Covelli I, Salvatore G, Sena L, Roche J. 1960. Sur la formation d'hormones thyroidiennes et de leurs précurseurs par Branchiostoma lanceolatum. C R Soc Biol Paris 154:1165–1169.

Cowen RK, Paris CB, Srinivasan A. 2006. Scaling of connectivity in marine populations. Science 311:522–527.

Das B, Cai L, Carter MG, Piao YL, Sharov AA, Ko MS, Brown DD. 2006. Gene expression changes at metamorphosis induced by thyroid hormone in *Xenopus laevis* tadpoles. Dev Biol 291:342–355.

Davis PJ, Leonard JL, Davis FB. 2007. Mechanisms of nongenomic actions of thyroid hormone. Front Neuroendocrinol 29:211–218.

de Jesus EGT, Toledo JD, Simpas MS. 1998. Thyroid hormones promote early metamorphosis in grouper (*Epinephelus coioides*) larvae. Gen Comp Endocrinol 112:10–16.

Dehal P, Boore JL. 2005. Two rounds of whole genome duplication in the ancestral vertebrate. PLoS Biol 3:e314.

Dehal P, Satou Y, Campbell RK, Chapman J, Degnan B, De Tomaso A, Davidson B, Di Gregorio A, Gelpke M, Goodstein DM, Harafuji N, Hastings KE, Ho I, Hotta K, Huang W, Kawashima T, Lemaire P, Martinez D, Meinertzhagen IA, Necula S, Nonaka M, Putnam N, Rash S, Saiga H, Satake M, Terry A, Yamada L, Wang HG, Awazu S, Azumi K, Boore J, Branno M, Chin-Bow S, DeSantis R, Doyle S, Francino P, Keys DN, Haga S, Hayashi H, Hino K, Imai KS, Inaba K, Kano S, Kobayashi K, Kobayashi M, Lee BI, Makabe KW, Manohar C, Matassi G, Medina M, Mochizuki Y, Mount S, Morishita T, Miura S, Nakayama A, Nishizaka S, Nomoto H, Ohta F, Oishi K, Rigoutsos I, Sano M, Sasaki A, Sasakura Y, Shoguchi E, Shin-i T, Spagnuolo A, Stainier D, Suzuki MM, Tassy O, Takatori N, Tokuoka M, Yagi K, Yoshizaki F, Wada S, Zhang C, Hyatt PD, Larimer F, Detter C, Doggett N, Glavina T, Hawkins T, Richardson P, Lucas S, Kohara Y, Levine M, Satoh N, Rokhsar DS. 2002. The draft genome of *Ciona intestinalis*: Insights into chordate and vertebrate origins. Science 298:2157–2167.

Denver RJ. 1998. Hormonal correlates of environmentally induced metamorphosis in the Western spadefoot toad, *Scaphiopus hammondii*. Gen Comp Endocrinol 110:326–336.

Eales JG. 1997. Iodine metabolism and thyroid-related functions in organisms lacking thyroid follicles: are thyroid hormones also vitamins? Proc Soc Exp Biol Med 214:302–317.

Eales JG, Holmes JA, McLeese JM, Youson JH. 1997. Thyroid hormone deiodination in various tissues of larval and upstream-migrant sea lampreys, *Petromyzon marinus*. Gen Comp Endocrinol 106:202–210.

Eneqvist T, Lundberg E, Nilsson L, Abagyan R, Sauer-Eriksson AE. 2003. The transthyretin-related protein family. Eur J Biochem 270:518–532.

Escriva H, Manzon L, Youson J, Laudet V. 2002. Analysis of lamprey and hagfish genes reveals a complex history of gene duplications during early vertebrate evolution. Mol Biol Evol 19:1440–1450.

Essner JJ, Johnson RG, Hackett PB Jr. 1999. Overexpression of thyroid hormone receptor α1 during zebrafish embryogenesis disrupts hindbrain patterning and implicates retinoic acid receptors in the control of hox gene expression. Differentiation 65:1–11.

Everts ME, Verhoeven FA, Bezstarosti K, Moerings EP, Hennemann G, Visser TJ, Lamers JM. 1996. Uptake of thyroid hormones in neonatal rat cardiac myocytes. Endocrinology 137:4235–4242.

Flamant F, Baxter JD, Forrest D, Refetoff S, Samuels H, Scanlan TS, Vennstrom B, Samarut J. 2006. International Union of Pharmacology. LIX. The pharmacology and classification of the nuclear receptor superfamily: Thyroid hormone receptors. Pharmacol Rev 58:705–711.

Flamant F, Poguet AL, Plateroti M, Chassande O, Gauthier K, Streichenberger N, Mansouri A, Samarut J. 2002. Congenital hypothyroid Pax8(−/−) mutant mice can be rescued by inactivating the TRα gene. Mol Endocrinol 16:24–32.

Flamant F, Samarut J. 2003. Thyroid hormone receptors: Lessons from knockout and knock-in mutant mice. Trends Endocrinol Metab 14:85–90.

Fraichard A, Chassande O, Plateroti M, Roux JP, Trouillas J, Dehay C, Legrand C, Gauthier K, Kedinger M, Malaval L, Rousset B, Samarut J. 1997. The T3Rα gene encoding a thyroid hormone receptor is essential for post-natal development and thyroid hormone production. EMBO J 16:4412–4420.

Fredriksson G. 1988.Thyroid-like systems in endostyles: A study of morphology, function and evolution in "primitive" chordates. In: Department of Zoology, PhD thesis. Stockholm: University of Stockholm. p 33.

Fredriksson G, Öfverholm T, Ericson LE. 1985. Electron-microscopic studies of iodine-binding and peroxidase activity in the endostyle of the larval amphioxus (Branchiostoma lanceolatum). Cell Tissue Res 241:257–266.

Friesema EC, Jansen J, Milici C, Visser TJ. 2005. Thyroid hormone transporters. Vitam Horm 70:137–167.

Friesema EC, Kuiper GG, Jansen J, Visser TJ, Kester MH. 2006. Thyroid hormone transport by the human monocarboxylate transporter 8 and its rate-limiting role in intracellular metabolism. Mol Endocrinol 20:2761–2772.

Fukazawa H, Hirai H, Hori H, Roberts RD, Nukaya H, Ishida I, Tsuji K. 2001. Induction of abalone larval metamorphosis by thyroid hormones. Fish Sci 67:985–988.

Gauthier K, Chassande O, Plateroti M, Roux JP, Legrand C, Pain B, Rousset B, Weiss R, Trouillas J, Samarut J. 1999. Different functions for the thyroid hormone receptors TRα and TRβ in the control of thyroid hormone production and post-natal development. EMBO J 18:623–631.

Germain P, Staels B, Dacquet C, Spedding M, Laudet V. 2006. Overview of nomenclature of nuclear receptors. Pharmacol Rev 58:685–704.

Gilbert LI, Tata JR, Atkinson BG. 1996. Metamorphosis: Postembryonic reprogramming of gene expression in amphibian and insect cells. San Diego, CA: Academic Press. xvi, 687 p.

Gronemeyer H, Gustafsson JA, Laudet V. 2004. Principles for modulation of the nuclear receptor superfamily. Nat Rev Drug Discov 3: 950–964.

Gross J, Pitt-Rivers R. 1953a. 3:5:3′-Triiodothyronine. I. Isolation from thyroid gland and synthesis. J Biochem 53:645–652.

Gross J, Pitt-Rivers R. 1953b. 3:5:3′-Triiodothyronine. II. Physiological activity. J Biochem 53:652–657.

Gudernatsch JF. 1912. Feeding experiments on tadpoles. Wilhem Roux Arch Entwickl Mech Org 35:457–484.

Guissouma H, Dupré SM, Demeneix BA. 2005. Transcriptional repression of the TRH gene. Med Sci 21:854–859.

Hadj-Sahraoui N, Seugnet I, Ghorbel MT, Demeneix B. 2000. Hypothyroidism prolongs mitotic activity in the post-natal mouse brain. Neurosci Lett 280:79–82.

Havis E, Le Mevel S, Morvan Dubois G, Shi DL, Scanlan TS, Demeneix BA, Sachs LM. 2006. Unliganded thyroid hormone receptor is essential for *Xenopus laevis* eye development. EMBO J 25:4943–4951.

Heyland A, Hodin J, Reitzel AM. 2005. Hormone signaling in evolution and development: A non-model system approach. Bioessays 27: 64–75.

Heyland A, Price DA, Bodnarova-Buganova M, Moroz LL. 2006a. Thyroid hormone metabolism and peroxidase function in two non-chordate animals. J Exp Zoolog B Mol Dev Evol 306: 551–566.

Heyland A, Reitzel AM, Price DA, Moroz LL. 2006b. Endogenous thyroid hormone synthesis in facultative planktotrophic larvae of the sand dollar *Clypeaster rosaceus*: Implications for the evolutionary loss of larval feeding. Evol Dev 8:568–579.

Hodin J. 2006. Expanding networks: Signaling components in and a hypothesis for the evolution of metamorphosis. Integr Comp Biol 46:719–742.

Holland LZ, Albalat R, Azumi K, Benito-Gutierrez E, Blow MJ, Bronner-Fraser M, Brunet F, Butts T, Candiani S, Dishaw LJ, Ferrier DE, Garcia-Fernandez J, Gibson-Brown JJ, Gissi C, Godzik A, Hallbook F, Hirose D, Hosomichi K, Ikuta T, Inoko H, Kasahara M, Kasamatsu J, Kawashima T, Kimura A, Kobayashi M, Kozmik Z, Kubokawa K, Laudet V, Litman GW, McHardy AC, Meulemans D, Nonaka M, Olinski RP, Pancer Z, Pennacchio LA, Pestarino M, Rast JP, Rigoutsos I, Robinson-Rechavi M, Roch G, Saiga H, Sasakura Y, Satake M, Satou Y, Schubert M, Sherwood N, Shiina T, Takatori N, Tello J, Vopalensky P, Wada S, Xu A, Ye Y, Yoshida K, Yoshizaki F, Yu JK, Zhang Q, Zmasek CM, de Jong PJ, Osoegawa K, Putnam NH, Rokhsar DS, Satoh N, Holland PW. 2008. The amphioxus genome illuminates vertebrate origins and cephalochordate biology. Genome Res 18:1100–1111.

Howard-Ashby M, Materna SC, Brown CT, Chen L, Cameron RA, Davidson EH. 2006. Gene families encoding transcription factors expressed in early development of *Strongylocentrotus purpuratus*. Dev Biol 300:90–107.

Hulbert AJ. 2000. Thyroid hormones and their effects: A new perspective. Biol Rev Camb Philos Soc 75:519–631.

Inui Y, Miwa S. 1985. Thyroid hormone induces metamorphosis of flounder larvae. Gen Comp Endocrinol 60:450–454.

Ito Y, Okada R, Mochida H, Hayashi H, Yamamoto K, Kikuyama S. 2004. Molecular cloning of bullfrog corticotropin-releasing factor (CRF): Effect of homologous CRF on the release of TSH from pituitary cells in vitro. Gen Comp Endocrinol 138:218–227.

Jeffery WR, Swalla BJ, Ewing N, Kusakabe T. 1999. Evolution of the ascidian anural larva: Evidence from embryos and molecules. Mol Biol Evol 16:646–654.

Jørgensen JM. 1998. The biology of hagfishes, 1st ed. London, New York: Chapman & Hall. xix, 578 p.

Joss JM. 2006. Lungfish evolution and development. Gen Comp Endocrinol 148:285–289.

Kashian DR, Dodson S. 2004. Effects of vertebrate hormones on development and sew determination in *Daphnia Magna*. Environ Toxicol Chem 23:1282–1288.

Kawahara A, Baker BS, Tata JR. 1991. Developmental and regional expression of thyroid hormone receptor genes during Xenopus metamorphosis. Development 112:933–943.

Kluge B, Renault N, Rohr KB. 2005. Anatomical and molecular reinvestigation of lamprey endostyle development provides new insight into thyroid gland evolution. Dev Genes Evol 215:32–40.

Kobayashi H, Tsuneki K, Akiyoshi H, Kobayashi Y, Nozaki M, Ouji M. 1983. Histochemical distribution of peroxidase in ascidians with special reference to the endostyle and the branchial sac. Gen Comp Endocrinol 50:172–187.

Kon T, Nohara M, Yamanoue Y, Fujiwara Y, Nishida M, Nishikawa T. 2007. Phylogenetic position of a whale-fall lancelet (Cephalochordata) inferred from whole mitochondrial genome sequences. BMC Evol Biol 7:127.

Kozmik Z. 2008. The role of Pax genes in eye evolution. Brain Res Bull 75:335–339.

Laudet V, Gronemeyer H. 2005. The Nuclear Receptor FacsBook. London: Academic Press.

Leis JM. 2006. Are larvae of demersal fishes plankton or nekton? In: Sims DW, Southward AJ, editors. Advances in marine biology. Plymouth: Academic Press. pp 57–141.

Leloup J, Buscaglia M. 1977. La triiodothyronine: hormone de la metamorphose des amphibians. Comptes rendus de l'académie des sciences 284:2261–2263.

Leloup J, Lebel JM. 1993. Triiodothyroine is necessary for the action of growth hormone in acclimatation to seawater of brown (*Salmo trutta*) and rainbow trout (*Oncorhynchus mykiss*). Fish Physiol Biochem 11:165–173.

Lim W, Nguyen NH, Yang HY, Scanlan TS, Furlow JD. 2002. A thyroid hormone antagonist that inhibits thyroid hormone action in vivo. J Biol Chem 277:35664–35670.

Lintlop SP, Youson JH. 1983. Binding of triiodothyronine to hepatocyte nuclei from sea lampreys, *Petromyzon marinus* L., at various stages of the life cycle. Gen Comp Endocrinol 49:428–436.

Manzon RG, Eales JG, Youson JH. 1998. Blocking of KC1O4-induced metamorphosis in premetamorphic sea lampreys by exogenous thyroid hormones (TH); effects of KC1O4 and TH on serum TH concentrations and intestinal thyroxine outer-ring deiodination. Gen Comp Endocrinol 112:54–62.

Manzon RG, Holmes JA, Youson JH. 2001. Variable effects of goitrogens in inducing precocious metamorphosis in sea lampreys (*Petromyzon marinus*). J Exp Zool 289:290–303.

Manzon RG, Neuls TM, Manzon LA. 2007. Molecular cloning, tissue distribution, and developmental expression of lamprey transthyretins. Gen Comp Endocrinol 151:55–65.

Manzon RG, Youson JH. 1997. The effects of exogenous thyroxine (T4) or triiodothyronine (T3), in the presence and absence of potassium perchlorate, on the incidence of metamorphosis and on serum T4 and T3 concentrations in larval sea lampreys (*Petromyzon marinus* L.). Gen Comp Endocrinol 106:211–220.

Manzon RG, Youson JH. 1999. Temperature and KClO(4)-induced metamorphosis in the sea lamprey (*Petromyzon marinus*). Comp Biochem Physiol C Pharmacol Toxicol Endocrinol 124: 253–257.

Manzon RG, Youson JH. 2002. KClO(4) inhibits thyroidal activity in the larval lamprey endostyle in vitro. Gen Comp Endocrinol 128:214–223.

Marchand O, Duffraisse M, Triqueneaux G, Safi R, Laudet V. 2004. Molecular cloning and developmental expression patterns of thyroid hormone receptors and T3 target genes in the turbot (*Scophtalmus maximus*) during post-embryonic development. Gen Comp Endocrinol 135:345–357.

Marchand O, Safi R, Escriva H, Van Rompaey E, Prunet P, Laudet V. 2001. Molecular cloning and characterization of thyroid hormone receptors in teleost fish. J Mol Endocrinol 26:51–65.

Markov G, Paris M, Bertrand S, Laudet V. 2008. The evolution of the ligand/receptor couple: A long road from comparative endocrinology to comparative genomics. Mol Cell Endocrinol 293:5–16.

McComb DM, Gelsleichter J, Manire CA, Brinn R, Brown CL. 2005. Comparative thyroid hormone concentration in maternal serum and yolk of the bonnethead shark (*Sphyrna tiburo*) from two sites along the coast of Florida. Gen Comp Endocrinol 144:167–173.

Miwa S, Inui Y. 1985. Effects of -thyroxine and ovine growth hormone on smoltification of amago salmon (*Oncorhynchus rhodurus*). Gen Comp Endocrinol 58:436–442.

Mochizuki K, Yagi E, Sakaguchi N, Mochizuki H, Takabe S, Kuranuki S, Suzuki T, Shimada M, Goda T. 2007. The critical period for thyroid hormone responsiveness through thyroid hormone receptor isoform α in the postnatal small intestine. Biochim Biophys Acta 1770:609–616.

Monaco F, Andreoli M, La Posta A, Roche J. 1978. Thyroglobulin biosynthesis in a larval (ammocoete) and adult freshwater lamprey (*Lampetra planeri* Bl.). Comp Biochem Physiol B 60:87–91.

Monaco F, Dominici R, Andreoli M, PirroR. D, Roche J. 1981. Thyroid hormone formation in thyroglobulin synthesized in the amphioxus (Branchiostoma lanceolatum Pallas). Comp Biochem Physiol B 70:341–343.

Moreno M, de Lange P, Lombardi A, Silvestri E, Lanni A, Goglia F. 2008. Metabolic effects of thyroid hormone derivatives. Thyroid 18:239–253.

Nelson JS. 2006. Fishes of the world, 4th ed. Hoboken, NJ: Wiley. xix, 601 p.

Newman RA. 1998. Ecological constraints on amphibian metamorphosis: Interactions of temperature and larval density with responses to changing food level. Oecologia 115:9–16.

Nguyen NH, Apriletti JW, Cunha Lima ST, Webb P, Baxter JD, Scanlan TS. 2002. Rational design and synthesis of a novel thyroid hormone antagonist that blocks coactivator recruitment. J Med Chem 45:3310–3320.

Nielsen C. 1998. Origin and evolution of animal life cycles. Biol Rev 73:125–155.

Ogasawara M. 2000. Overlapping expression of amphioxus homologs of the thyroid transcription factor-1 gene and thyroid peroxidase gene in the endostyle: Insight into evolution of the thyroid gland. Dev Genes Evol 210:231–242.

Paris M, Escriva H, Schubert M, Brunet F, Brtko J, Ciesielski F, Roecklin D, Vivat-Hannah V, Jamin EL, Cravedi J-P, Scanlan TS, Renaud J-P, Holland ND, Laudet V. 2008. Amphioxus postembryonic development reveals the homology of chordate metamorphosis. Curr Biol 18:825–830.

Patricolo E, Cammarata M, D'Agati P. 2001. Presence of thyroid hormones in ascidian larvae and their involvement in metamorphosis. J Exp Zool 290:426–430.

Patricolo E, Ortolani G, Cascio A. 1981. The effect of l-thyroxine on the metamorphosis of *Ascidia malaca*. Cell Tissue Res 214:289–301.

Pearson JC, Lemons D, McGinnis W. 2005. Modulating Hox gene functions during animal body patterning. Nat Rev Genet 6:893–904.

Plateroti M, Chassande O, Fraichard A, Gauthier K, Freund JN, Samarut J, Kedinger M. 1999. Involvement of T3Rα- and β-receptor subtypes in mediation of T3 functions during postnatal murine intestinal development. Gastroenterology 116:1367–1378.

Power DM, Llewellyn L, Faustino M, Nowell MA, Bjornsson BT, Einarsdottir IE, Canario AV, Sweeney GE. 2001. Thyroid hormones in growth and development of fish. Comp Biochem Physiol C Toxicol Pharmacol 130:447–459.

Roustaian P, Gaik LA. 2006. Effect of thyroxine immersion on larval survival, growth and postlarvae production of freshwater prawn, *Macrobrachium rosenbergii* (de Man). Aquac Res 37:1378–1380.

Safi R, Bertrand S, Marchand O, Duffraisse M, de Luze A, Vanacker JM, Maraninchi M, Margotat A, Demeneix B, Laudet V. 2004. The axolotl (*Ambystoma mexicanum*), a neotenic amphibian, expresses functional thyroid hormone receptors. Endocrinology 145:760–772.

Safi R, Vlaeminck-Guillem V, Duffraisse M, Seugnet I, Plateroti M, Margotat A, Duterque-Coquillaud M, Crespi EJ, Denver RJ, Demeneix B, Laudet V. 2006. Pedomorphosis revisited: Thyroid hormone receptors are functional in *Necturus maculosus*. Evol Dev 8:284–292.

Scanlan TS, Suchland KL, Hart ME, Chiellini G, Huang Y, Kruzich PJ, Frascarelli S, Crossley DA, Bunzow JR, Ronca-Testoni S, Lin ET, Hatton D, Zucchi R, Grandy DK. 2004. 3-Iodothyronamine is an endogenous and rapid-acting derivative of thyroid hormone. Nat Med 10:638–642.

Sea Urchin Genome Sequencing C, Sodergren E, Weinstock GM, Davidson EH, Cameron RA, Gibbs RA, Angerer RC, Angerer LM, Arnone MI, Burgess DR, Burke RD, Coffman JA, Dean M, Elphick MR, Ettensohn CA, Foltz KR, Hamdoun A, Hynes RO, Klein WH, Marzluff W, McClay DR, Morris RL, Mushegian A, Rast JP, Smith LC, Thorndyke MC, Vacquier VD, Wessel GM, Wray G, Zhang L, Elsik CG, Ermolaeva O, Hlavina W, Hofmann G, Kitts P, Landrum MJ, Mackey AJ, Maglott D, Panopoulou G, Poustka AJ, Pruitt K, Sapojnikov V, Song X, Souvorov A, Solovyev V, Wei Z, Whittaker CA, Worley K, Durbin KJ, Shen Y, Fedrigo O, Garfield D, Haygood R, Primus A, Satija R, Severson T, Gonzalez-Garay ML, Jackson AR, Milosavljevic A, Tong M, Killian CE, Livingston BT, Wilt FH, Adams N, Belle R, Carbonneau S, Cheung R, Cormier P, Cosson B, Croce J, Fernandez-Guerra A, Geneviere A-M, Goel M, Kelkar H, Morales J, Mulner-Lorillon O, Robertson AJ, Goldstone JV, Cole B, Epel D, Gold B, Hahn ME, Howard-Ashby M, Scally M, Stegeman JJ, Allgood EL, Cool J, Judkins KM, McCafferty SS, Musante AM, Obar RA, Rawson AP, Rossetti BJ, Gibbons IR, Hoffman MP, Leone A, Istrail S, Materna SC, Samanta MP, Stolc V, Tongprasit W, Tu Q, Bergeron K-F, Brandhorst BP, Whittle J, Berney K, Bottjer DJ, Calestani C, Peterson K, Chow E, Yuan QA, Elhaik E, Graur D, Reese JT, Bosdet I, Heesun S, Marra MA, Schein J, Anderson MK, Brockton V, Buckley KM, Cohen AH, Fugmann SD, Hibino T, Loza-Coll M, Majeske AJ, Messier C, Nair SV, Pancer Z, Terwilliger DP, Agca C, Arboleda E, Chen N, Churcher AM, Hallbook F, Humphrey GW, Idris MM, Kiyama T, Liang S, Mellott D, Mu X, Murray G, Olinski RP, Raible F, Rowe M, Taylor JS, Tessmar-Raible K, Wang D, Wilson KH, Yaguchi S, Gaasterland T, Galindo BE, Gunaratne HJ, Juliano C, Kinukawa M, Moy GW, Neill AT, Nomura M, Raisch M, Reade A, Roux MM, Song JL, Su Y-H, Townley IK, Voronina E, Wong JL, Amore G, Branno M, Brown ER, Cavalieri V, Duboc V, Duloquin L, Flytzanis C, Gache C, Lapraz F, Lepage T, Locascio A, Martinez P, Matassi G, Matranga V, Range R, Rizzo F, Rottinger E, Beane W, Bradham C, Byrum C, Glenn T, Hussain S, Manning G, Miranda E, Thomason R, Walton K, Wikramanayke A, Wu S-Y, Xu R, Brown CT, Chen L, Gray RF, Lee PY, Nam J, Oliveri P, Smith J, Muzny D, Bell S, Chacko J, Cree A, Curry S, Davis C, Dinh H, Dugan-Rocha S, Fowler J, Gill R, Hamilton C, Hernandez J, Hines S, Hume J, Jackson L, Jolivet A, Kovar C, Lee S, Lewis L, Miner G, Morgan M, Nazareth LV, Okwuonu G, Parker D, Pu L-L, Thorn R, Wright R. 2006. The genome of the sea urchin *Strongylocentrotus purpuratus*. Science 314:941–952.

Sembrat K. 1953. Effect of the endostyle of the lancelet (Branchiostoma lanceolatum Pall.) on the metamorphosis of axolotl. Zool Pol 6:3–19.

Shepherdley CA, Klootwijk W, Makabe KW, Visser TJ, Kuiper GGJM. 2004. An ascidian homolog of vertebrate iodothyronine deiodinases. Endocrinology 145:1255–1268.

Shi Y-B. 2000. Amphibian metamorphosis: From morphology to molecular biology. New York: John Wiley. xiv, 288p.

Sly BJ, Snoke MS, Raff RA. 2003. Who came first—larvae or adults? Origins of bilaterian metazoan larvae. Int J Dev Biol 47:623–632.

Stokes MD, Holland ND. 1995. Ciliary hovering in larval lancelets (=Amphioxus). Biol Bull 188:231–233.

Sutija M, Longhurst TJ, Joss JM. 2003. Deiodinase type II and tissue specific mRNA alternative splicing in the Australian lungfish, *Neoceratodus forsteri*. Gen Comp Endocrinol 132:409–417.

Sutija M, Longhurst TJ, Joss JM. 2004. Deiodinase type III in the Australian lungfish, *Neoceratodus forsteri*. Gen Comp Endocrinol 136: 152–161.

Tata JR. 2006. Amphibian metamorphosis as a model for the developmental actions of thyroid hormone. Mol Cell Endocrinol 246:10–20.

Tata JR, Baker BS, Machuca I, Rabelo EM, Yamauchi K. 1993. Autoinduction of nuclear receptor genes and its significance. J Steroid Biochem Mol Biol 46:105–119.

Thomas IM. 1956. The accumulation of radioactive iodine by amphioxus. J Mar Biol Assoc U K 35:203–210.

Tong W, Kerkof P, Chaikoff IL. 1961. 131I utilization by thyroid tissue of the hagfish. Biochim Biophys Acta 52:299–304.

Tong W, Kerkof P, Chaikoff IL. 1962. Identification of labeled thyroxine and triiodothyronine in amphioxus treated with 131-I. Biochim Biophys Acta 56:326–331.

Tsuneki K, Kobayashi H, Ouji M. 1983. Histochemical distribution of peroxidase in amphioxus and cyclostomes with special reference to the endostyle. Gen Comp Endocrinol 50:188–200.

Varsamos S, Nebel C, Charmantier G. 2005. Ontogeny of osmoregulation in postembryonic fish: A review. Comp Biochem Physiol A Mol Integr Physiol 141:401–429.

Voss SR, Shaffer HB. 1997. Adaptive evolution via a major gene effect: Paedomorphosis in the Mexican axolotl. Proc Natl Acad Sci USA 94:14185–14189.

Willey A. 1894. Amphioxus and the ancestry of the vertebrates. New York, London: Macmillan. xiv, 316 p.

Wu SY, Green WL, Huang WS, Hays MT, Chopra IJ. 2005. Alternate pathways of thyroid hormone metabolism. Thyroid 15:943–958.

Wu W, Niles EG, LoVerde PT. 2007. Thyroid hormone receptor orthologues from invertebrate species with emphasis on *Schistosoma mansoni*. BMC Evol Biol 7:150.

Yamano K, Miwa S. 1998. Differential gene expression of thyroid hormone receptor α and β in fish development. Gen Comp Endocrinol 109:75–85.

Yen PM. 2001. Physiological and molecular basis of thyroid hormone action. Physiol Rev 81:1097–1142.

Yen PM, Feng X, Flamant F, Chen Y, Walker RL, Weiss RE, Chassande O, Samarut J, Refetoff S, Meltzer PS. 2003. Effects of ligand and thyroid hormone receptor isoforms on hepatic gene expression profiles of thyroid hormone receptor knockout mice. EMBO Rep 4:581–587.

Young G, Bjornsson BT, Prunet P, Lin RJ, Bern HA. 1989. Smoltification and seawater adaptation in Coho salmon (*Oncorhyncus kisutch*): Plasma prolactin, growth hormone, thyroid hormone, and cortisol. Gen Comp Endocrinol 74:335–345.

Youson JH. 1997. Is lamprey metamorphosis regulated by thyroid hormones? Am Zool 37:441–460.

Youson JH, Manzon RG, Peck BJ, Holmes JA. 1997. Effects of exogenous thyroxine (T4) and triiodothyronine (T3) on spontaneous metamorphosis and serum T4 and T3 levels in immediately premetamorphic sea lampreys, *Petromyzon marinus*. J Exp Zool 279: 145–155.

Zhang J, Lazar MA. 2000. The mechanism of action of thyroid hormones. Annu Rev Physiol 62:439.

Acta Zoologica (Stockholm) **89**: 000–000 (July 2008) doi: 10.1111/j.1463-6395.2008.00379.x

The club-shaped gland of amphioxus: export of secretion to the pharynx in pre-metamorphic larvae and apoptosis during metamorphosis

Nicholas D. Holland,[1] Mathilde Paris[2] and Demian Koop[1]

[1]Marine Biology Research Division, Scripps Institution of Oceanography, La Jolla, CA 92093-0202, USA, [2]Institute de Génomique Fonctionnelle de Lyon, Université de Lyon, CNRS, INRA, UCB Lyon 1, IFR 128 Lyon Biosciences, Ecole Normale Supérieure de Lyon, 46 Allée d'Italie, 69364, Lyon Cedex 07, France

Keywords:
Cephalochordata, *Branchiostoma*, lancelet, TUNEL, endostyle, suspension feeding

Accepted for publication: 15 September 2008

Abstract

Holland, N.D., Paris, M. and Koop, D. 2008. The club-shaped gland of amphioxus: export of secretion to the pharynx in pre-metamorphic larvae and apoptosis during metamorphosis. — *Acta Zoologica* (Stockholm) **00**: 000–000

In amphioxus larvae, the club-shaped gland is a tube connecting the pharyngeal lumen with the external environment. The functions of the gland and its fate during the larva-to-juvenile metamorphosis have long been controversial. Here we use a fixative including ruthenium red to preserve extracellular secretions (presumably glycoproteins) in late pre-metamorphic larvae. This procedure reveals reddish, fibrogranular material in the lumen of the club-shaped gland and in the pharynx adjacent to the gland's inner opening. This finding strengthens the idea that secretions of the club-shaped gland are exported to the pharyngeal lumen to help form a mucous trap for capturing food particles entering the mouth. We also use the terminal desoxynucleotidyl transferase-mediated dUTP nick end labelling (TUNEL) assay to study apoptosis in the tissues of metamorphosing larvae. One of the earliest events of metamorphosis is the massive apoptotic destruction of the club-shaped gland. Therefore, despite some previous opinions to the contrary, the cells of the gland do not survive to participate in the genesis of the definitive endostyle or any other post-larval structures.

Nicholas D. Holland, Marine Biology Research Division, Scripps Institution of Oceanography, La Jolla, CA 92093-0202, USA.
E-mail: nholland@ucsd.edu

Introduction

The club-shaped gland of amphioxus was initially discovered in pre-metamorphic larvae by Schulze (1851) and was later found to be limited to that developmental stage (Kowalevsky 1867, 1877). In 1882, Hatschek showed that the structure arises as an evagination from the pharyngeal endoderm in late embryos and soon develops into a tube open at either end. The inner opening of the tube communicates with the pharyngeal lumen, and the outer opening is a pore in the epidermis near the anteroventral rim of the mouth. The cells comprising most of the tube are secretory (Lacalli 2008), but the region of the tube near its external end comprises non-secretory duct cells. Several homologies have been proposed between the amphioxus club-shaped gland and structures in other animals on the basis of morphology (Table 1), although none of these proposals has been widely accepted.

Table 1 also summarizes the diverse functions that have been attributed to the amphioxus club-shaped gland. The long-held majority view was that the gland produces secretions exiting via its external pore. More recently, however, dye experiments and direct observations of particles moving in the gland lumen strongly indicate that the secretions of the gland are transported inward, into the pharyngeal lumen (Olsson 1983; Gilmour 1996; Holland and Yu 2002). The inward transport of secretions is also indicated by transmission electron micrographs showing that cilia projecting into the lumen of the gland are orientated to induce an inward flow (Lacalli 2008).

It is generally agreed that the club-shaped gland can no longer be detected soon after the larva begins metamorphosis. However, there have been diverse opinions about the exact fate of the cells comprising the gland during the course of metamorphosis. Willey (1891) thought such cells enter the

Table 1 Speculations on the functions and homologies of the amphioxus club-shaped gland

Reference	Proposed function in larvae	Proposed homology
Van Beneden and Julin (1886)	–	Ascidian tunicate intestine
Willey (1891)	Outlet for noxious material	Gill slit of vertebrate[1]
Neal (1898)	–	Part of vertebrate mouth
Goldschmidt (1905)	Secretes feeding mucus into pharynx	–
Hatschek (1892)	–	Endostyle/thyroid
Orton (1914)	Outlet of adhesive or flotation mucus	–
Garstang (1928)	Early outlet for adhesive; later secretes feeding mucus into pharynx	Ascidian tunicate coelom
Goodrich (1930)	Outlet for unspecified secretions	–
Conklin (1932)	Exports secretion against eating sand	–
Bone (1958)	Outlet for secretion for food capture	–
Fuller (1958)	Produces hormone for metamorphosis[2]	–
Wickstead (1967)	Produces hormone for metamorphosis[2]	–
Webb (1969)	Outlet for secretion to aid food capture	–
Olsson (1983)	May secrete enzymes into pharynx	Endostyle/thyroid
Berrill (1987)	Outlet for adhesive	–
Gilmour (1996)	Secretes feeding mucus into pharynx	–
Yasui *et al.* (2000)	–	Ambulacrarian coelom
Godoy *et al.* (2006)	Source of antimicrobial secretion	–
Lacalli (2008)	Secretes feeding mucus into pharynx	–

[1]This idea was also a favourite of van Wijhe (1914) and others who will not be listed here.
[2]Unlikely because of the lack of iodine binding by the club-shaped gland (Fredriksson *et al.* 1984).

pharyngeal lumen and are destroyed there. Hatschek (1892) implied that the club-shaped gland, in its entirety, is converted into the definitive endostyle, a glandular organ (presumed to be homologous to the vertebrate thyroid) running along the ventral side of the pharynx in post-metamorphic specimens of amphioxus. Klaatsch (1898) proposed that the cells of the gland transdifferentiate into cartilage-like cells supporting the preoral cirri (projections fringing the vestibule after metamorphosis). Van Wijhe (1914) thought that the gland cells degenerate – some in the gut lumen and others within the haemal channels. Olsson (1983) and Gilmour (1996) suggested that the cells of the club-shaped gland might be incorporated into the post-metamorphic endostyle, although Fredriksson *et al.* (1984) thought that such incorporation was unlikely.

The controversies about the functions of the amphioxus club-shaped gland and uncertainty about the ultimate fate of its component cells prompted the present study. To gain insights into the functions of the gland, we preserved its secretions *in situ* in a fixative including ruthenium red. In addition, to determine the fate of the gland cells at metamorphosis, we studied apoptosis in the larval tissues using the TUNEL (terminal desoxynucleotidyl transferase mediated dUTP nick end *labelling*) assay. Elucidation of apoptosis was facilitated by the recent findings of Paris *et al.* (2008) that amphioxus larvae can be stimulated to enter metamorphosis synchronously by the administration of triiodothyronine (T_3), a thyroid hormone. Taken together, our results demonstrate first that the amphioxus club-shaped gland exports secretions to the pharyngeal lumen to help capture food particles and second that the cells of the gland are rapidly destroyed by a pervasive apoptosis soon after metamorphosis begins.

Materials and methods

Collection and culture of amphioxus

Ripe specimens of the Florida amphioxus (*Branchiostoma floridae*) were collected using shovel and sieve at a depth of about 1 m in Tampa Bay, FL. Eggs were obtained from electrically stimulated females and were fertilized according to Holland and Holland (1993). The embryos and larvae were raised in filtered seawater in deep Petri dishes (100 mm × 25 mm) in an incubator at 30° ± 0.5 °C under ambient light conditions. Each Petri dish contained between 100 and 150 larvae in 100 mL filtered seawater. The larvae were cleaned once a day by transferring them individually by pipette to seawater in a new Petri dish. The larvae were fed twice a day on a mixture of live algae (60% *Isochrysis* + 40% *Monochrysis* for the first 5 days and subsequently 50% *Isochrysis* + 30% *Monochrysis* + 10% *Dunaliella* + 10% *Tetraselmis*). The algae, after having been centrifuged free of their medium, were added to each larval culture until the dish looked light green when viewed from the side (approximately 10^5 to 10^6 algal cells per mL).

Fixation of club-shaped gland secretions

To conserve the secretions of the club-shaped gland, we fixed three pre-metamorphic larvae (12 days old; with six gill slits) for 1 h in a mixture of 2% formaldehyde and 1% glutaraldehyde in 0.1 M cacodylate buffer (pH 7.3) containing 0.35 M NaCl and 1% ruthenium red (a modification of the method of Flood 1991, suggested to us by T. H. J. Gilmour). The fixation and the subsequent steps were performed at room temperature. After fixation, the specimens were rinsed in three 10-min changes of 0.1 M cacodylate buffer with 0.35 M NaCl and 1% ruthenium red and then post-fixed for 30 min in 0.1 M cacodylate buffer with 0.35 M NaCl and 1% osmium tetroxide. The specimens were then rinsed briefly in distilled water, dehydrated in an ethanol series and embedded in Spurr's resin. Sections were cut 4-μm thick with glass knives, counterstained with 0.1% aqueous azure A and mounted in immersion oil.

PFA fixation, hormone treatment, staging metamorphosis

As a starting point for the study of metamorphosis, 20 larvae (12 days old – most with six, a few with seven, gill slits) were fixed before hormone administration. Fixation was in freshly prepared 4% paraformaldehyde in 0.1 M 3-(*n*-morpholino) propanesulphonic acid (MOPS) buffer (pH 7.5) containing 1 mM ethyleneglycol tetraacetic acid, 2 mM $MgSO_4$ and 0.5 M NaCl (referred to below simply as PFA). After fixation overnight in the refrigerator, the specimens were rinsed briefly in distilled water and then stored in 70% ethanol at –20 °C until further processing. Metamorphosis was initiated by administering triiodothyronine (T_3) (Sigma, St Louis, MO) to approximately 200 12-day-old larvae. Without T_3 treatment, the larvae would not have entered metamorphosis for approximately another week and would have done so asynchronously (Paris *et al*. 2008). The larvae to be treated were isolated in 150 mL filtered seawater in a 140-mm × 20-mm Petri dish, and T_3 (dissolved in 1.5 µL of dimethyl sulphoxide) was added to give a final hormone concentration of 10^{-8} M. Paris *et al*. (2008) previously showed that this concentration of dimethyl sulphoxide has no adverse effects and that larvae stimulated to undergo metamorphosis with this concentration of T_3 develop into healthy juvenile animals. We renewed the T_3 daily for the duration of the experiment, during which the larvae were fed and maintained at 30 °C.

After the start of T_3 administration, stage 1 (metapleural folds just beginning) was reached between 1 and 1.5 days, stage 2 (advent of second row of gill slits) was reached between 1.5 and 2 days, stages 3/4 (gill slits becoming U-shaped, mouth migrating, and preoral cirri appearing) was reached between 2 and 3 days, and stage 5 (juvenile with elongating liver) was reached after 4 days. At each of the following time-points after the start of T_3 treatment, the following numbers of larvae were fixed in PFA: 26 h (25), 33 h (13), 40 h (30), 47 h (27), 54 h (30), and 72 h (38). Following overnight fixation in the refrigerator, the larvae were stored in 70% ethanol at –20 °C until further processing.

TUNEL assay for apoptosis

For visualizing apoptosis by the TUNEL assay, the larvae were labelled with the ApopTag Peroxidase In Situ Apoptosis Detection Kit (Chemicon, Temecula, CA) according to the manufacturer's instructions as modified by Bayascas *et al*. (2002). After the TUNEL reaction had developed, it was stopped in PFA and the larvae were mounted in 80% glycerol in sodium phosphate-buffered saline. For some of the whole mounts, the best-focused parts of a z-series were blended together with an IMAGE-PRO PLUS 5.1 image analysis program (Media-Cybernetics, Silver Spring, MD). After being photographed as whole mounts, the specimens were demounted, washed briefly in distilled water, dehydrated in an ethanol series, embedded in Spurr's resin, cut as 7-µm cross-sections with glass knives, mounted (without counterstaining) in immersion oil, and observed with differential interference contrast illumination.

Results

Pre-metamorphic larvae (six gill slits); secretion by the club-shaped gland)

After fixation of late pre-metamorphic larvae in a solution including ruthenium red, the lumen of the duct region of the club-shaped gland appears empty (Fig. 1A; dc). In contrast, the lumen of the glandular region of the club-shaped gland contains reddish fibrogranular material (Fig. 1B,C), presumably comprising polyanionic glycoproteins (Luft 1971), There can be little doubt that this ruthenium red-positive material is secreted by the cells surrounding the lumen of the club-shaped gland. Moreover, the entry of this secretion into the pharyngeal lumen is strongly suggested by the conspicuous concentration of such material immediately adjacent to the inner opening of the club-shaped gland (Fig. 1B, arrow).

In addition to the secretion apparently entering from the club-shaped gland, the pharyngeal lumen is traversed by a thin sheet of similar material. These two components are contiguous, although the sheet is more extensive, occupying a rostrocaudal zone delimited by the anterior and posterior lips of the mouth. The fibrogranular sheet slants ventrally at an angle of approximately 35° from the upper lip (Fig. 1A, twin asterisks) to the lower limb of the endostyle (Fig. 1A, single asterisk). The origin(s) of the fibrogranular sheet are uncertain, although it could well include a mixture of secretions from the club-shaped gland, endostyle and upper lip. Elsewhere in the larvae, the fixative with ruthenium red preserved no detectable secretions – either in the gut lumen or on the epidermal surface including the rostral mucus gland and the preoral organ (data not shown).

Distribution of TUNEL-positive cells in pre-metamorphic and metamorphic larvae

Stage 0, pre-metamorphic larvae. Most of the relatively few TUNEL-positive nuclei present in pre-metamorphic (six-gill-slit) larvae are distributed in the epidermis at the anterior tip of the animal (Fig. 2A,B), in the tail fin, and around the opening of the anus (Fig. 2E). Elsewhere, only a very few TUNEL-positive nuclei occur in the lateral epidermis (Fig. 2C, arrow) and in the lips around the mouth (Fig. 2D).

Early stage 1 larvae. In the first sample after the start of T_3 administration (26 h), most of the larvae show a striking increase in the frequency of TUNEL-labelled nuclei in the lips of the mouth (Fig. 2F,G) and in epidermal cells (arrowed

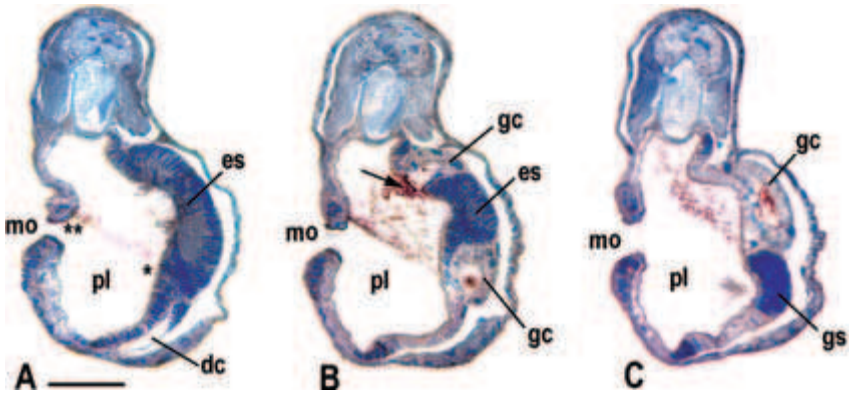


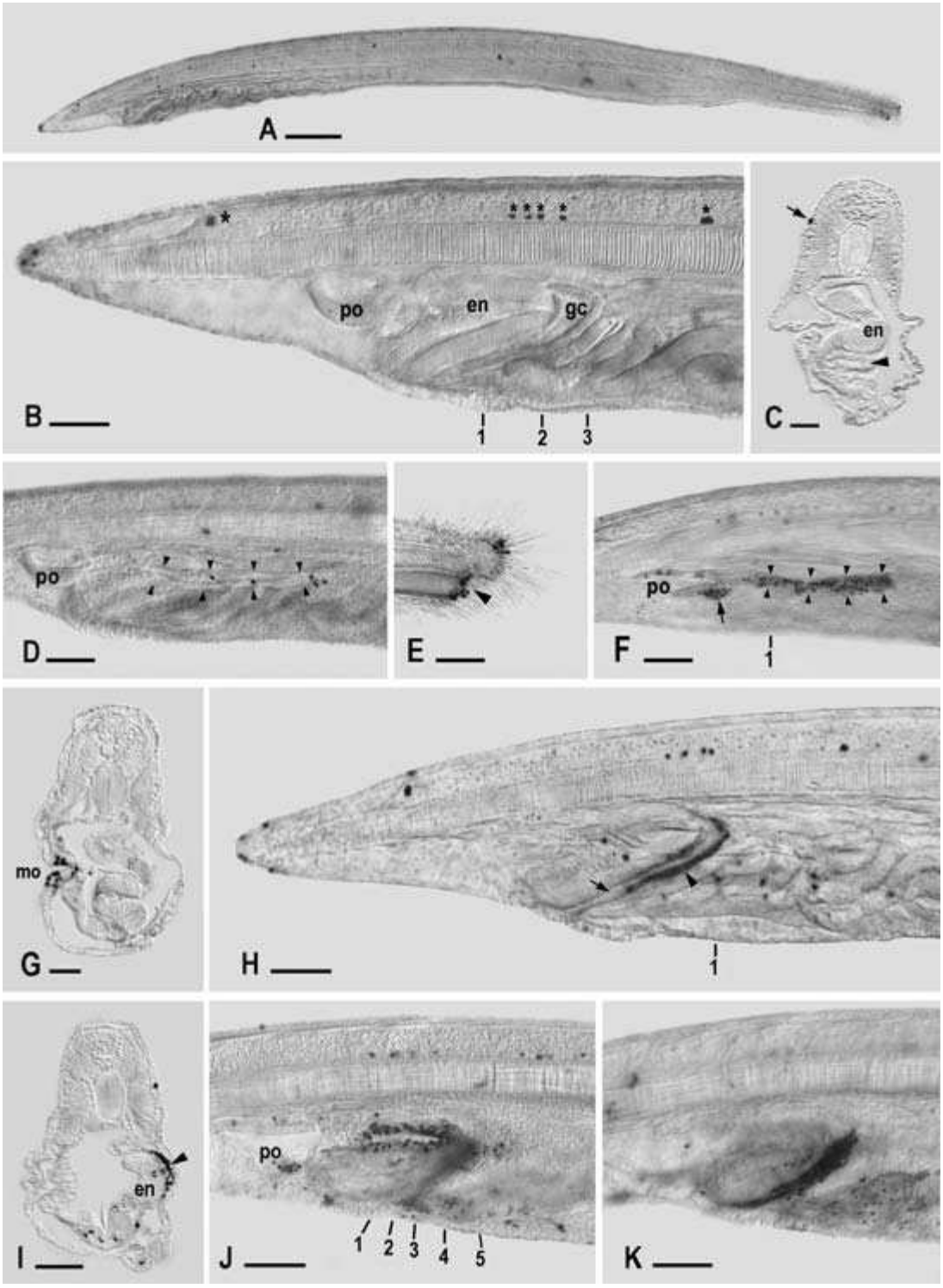
Fig. 1—Pre-metamorphic larva of the Florida amphioxus (12 days old; with six gill slits); cross-sections (as seen from the tail end) through levels where the mouth (mo) opens into the pharyngeal lumen (pl); stained for extracellular polyanions with ruthenium red and counterstained blue with azure A. Scale bar: 25 μm. —**A**. Cross-section through level 1 in Fig. 2B, showing the endostyle (es), the duct of the club-shaped gland (dc), and a delicate sheet of ruthenium red-positive material (twin asterisk) running from the upper lip of the mouth to the ventral limb of the endostyle (single asterisk). —**B**. Section through level 2 in Fig. 2B, showing the posterior angle of the endostyle (es) and the glandular part of the club-shaped gland (gc); the dorsal, right region of the pharyngeal lumen (pl) contains ruthenium red-stained material that is most conspicuous just outside the inner opening (arrow) of the club-shaped gland. —**C**. Section through level 3 in Fig. 2B, showing the anterior rim of the first gill slit (gs) and the glandular part of the club-shaped gland (gc); ruthenium red-stained material is present in the club-shaped gland lumen and in the dorsal, right region of the pharyngeal lumen (pl).

in Fig. 2F) at the ventroposterior corner of the preoral organ (po). These zones of apoptosis are presumably related to the extensive remodelling and migration of the preoral organ and mouth that soon follows.

Late stage 1 larvae. In addition to the numerous apoptotic cells associated with the mouth and preoral organ, a second conspicuous region of TUNEL-positive cells appears in the glandular part of the club-shaped gland (indicated by the arrowheads in Fig. 2H,I). In contrast, at this stage, the cells comprising the duct of the gland (Fig. 2H, arrow) show no signs of apoptosis.

Stage 2 larvae. The intense apoptosis in the lips of the (now shrinking) mouth is still present (Fig. 2J), and the lips are beginning to sink inward (Fig. 3A, asterisks) toward the midline of the animal. Elsewhere, virtually all the cells of the club-shaped gland, including those of its duct, are TUNEL-positive (Figs 2K, 3B–D,K). In addition, numerous apoptotic cells are present in the lips of the most anterior gill slit (Fig. 3E, asterisks), which is destined to close up and disappear during the next stage of metamorphosis. In contrast, the lips of the other gill slits of the left side and the newly forming ones of the right side include only relatively few TUNEL-positive cells (data not shown).

Fig. 2—Pre-metamorphic and metamorphic larvae of the Florida amphioxus: TUNEL reactions without counterstaining. The whole mounts are side views with the anterior toward the left; and the cross-sections are seen from the tail end. —**A**. Overview of a pre-metamorphic larva (12 days old, with six gill slits). Scale bar: 200 μm. —**B**. Anterior end of (A) showing the preoral organ (po), endostyle (en) and glandular region of the club-shaped gland (gc); the dark spots at the anterior tip are TUNEL-positive nuclei in the ectoderm, but the dark spots (marked by asterisks) in the nerve cord are pigment cells. Scale bar: 50 μm. —**C**. Cross-section through level 1 in (B); only one TUNEL-positive nucleus is visible (arrowed in epidermis); the other tissues, including the endostyle (en) and club-shaped gland (arrowhead), show no labelling. Scale bar: 25 μm. —**D**. Surface view near the anterior end of (A) showing the preoral organ (po) and the mouth (artefactually closed during fixation and so appearing as an anteroposteriorly extended slit, indicated by arrowheads); only a few TUNEL-positive nuclei are present in the epidermal cells of the lips. Scale bar: 50 μm. —**E**. Posterior end of (A) showing TUNEL-positive nuclei in the tail epidermis and associated with the anus (arrowhead). Scale bar: 50 μm. —**F**. Surface view of the anterior end of an early stage 1 metamorphic larva showing numerous epidermal cells with TUNEL-positive nuclei (arrowed) at the posteroventral margin of the preoral organ (PO) and the lips of the mouth (arrowheads). Scale bar: 50 μm. —**G**. Cross-section through level 1 in (F); several TUNEL-positive nuclei are conspicuous in epidermal cells of the upper and lower lips of the mouth (mo). Scale bar: 25 μm. —**H**. Late metamorphic stage 1 larva. TUNEL-positive nuclei are conspicuous in the glandular portion of the club-shaped gland (arrowhead), but not in its duct region (arrow). Scale bar: 50 μm. —**I**. Cross-section through level 1 in (H) showing the endostyle (en) and the TUNEL-positive cells in the glandular region of the club-shaped gland (arrowhead). Scale bar: 25 μm. —**J**. Surface view of a stage 2 metamorphic larva; TUNEL-positive epidermal cells are still abundant in the preoral organ (PO) and in the lips of the moth, which are now much contracted. Scale bar: 50 μm. —**K**. Deeper focal plane of the larva in (J) showing TUNEL-positive cells throughout the club-shaped gland. Scale bar: 50 μm.



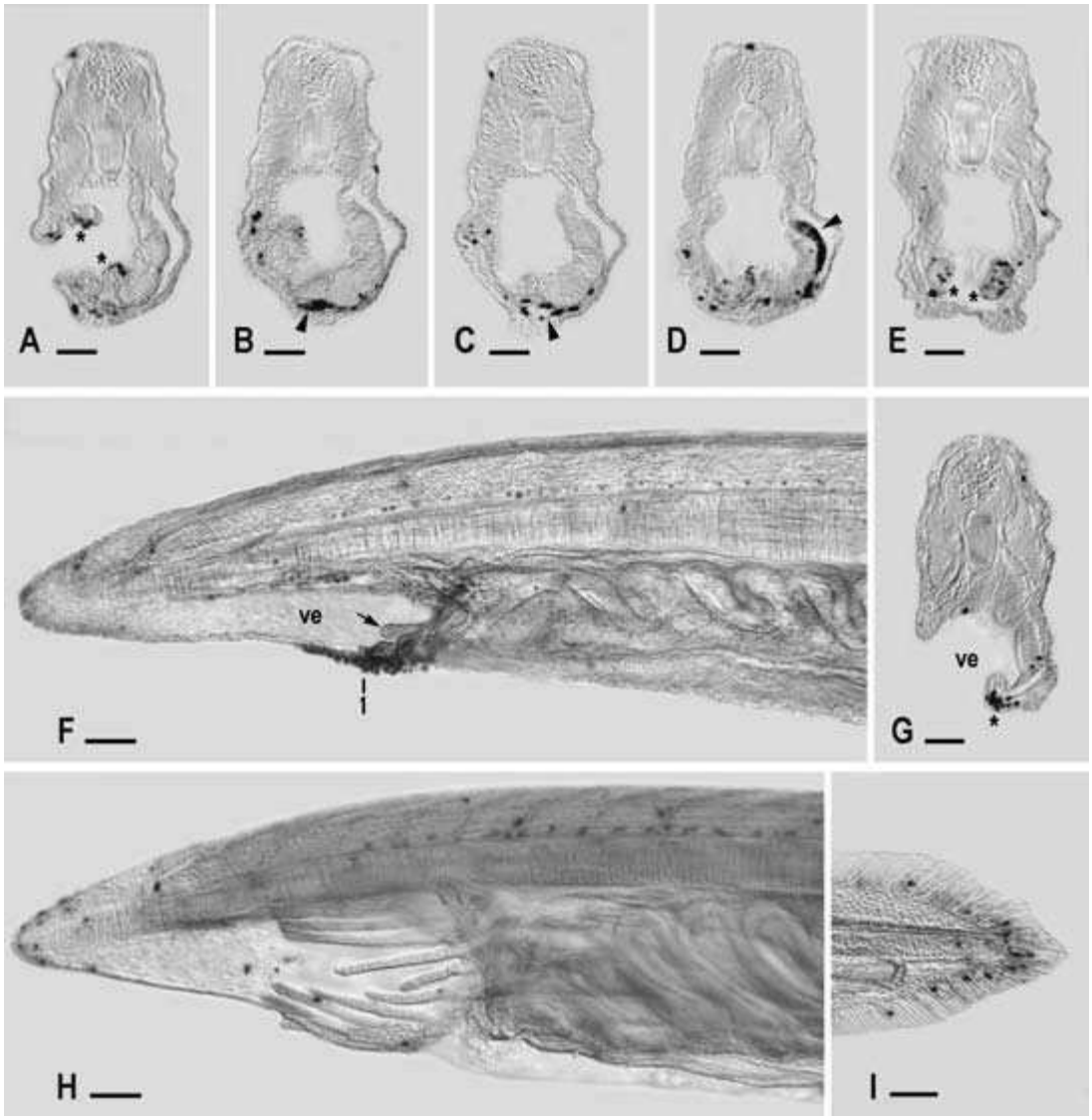


Fig. 3—Metamorphic larvae and juvenile of the Florida amphioxus: TUNEL reactions without counterstaining. The whole mounts are side views with the anterior toward the left; the cross-sections are seen from the tail end. —**A**. Stage 2 metamorphic larva; cross-section through level 1 in Fig. 2J, showing TUNEL-positive cells in the lips (indicated by asterisks) of the now in-sunken mouth. Scale bar: 25 μm. —**B**,**C**. Stage 2 metamorphic larva; cross-section through levels 2 and 3, respectively, of Fig. 2J showing TUNEL-positive cells in the duct region of the club-shaped gland (arrowhead). Scale bar: 25 μm. —**D**. Stage 2 metamorphic larva; cross-section through level 4 of the larva in Fig. 2(J) showing numerous TUNEL-positive cells in the glandular region of the club-shaped gland (arrowhead). Scale bar: 25 μm. —**E**. Stage 2 metamorphic larva; cross-section through level 5 of Fig. 2(J) showing conspicuous TUNEL-positive cells in the lips (asterisks) of the most anterior left gill slit. Scale bar: 25 μm. —**F**. Anterior end of a stage 3/4 larva showing many TUNEL-positive nuclei in the epidermis along the right margin of the vestibule (ve). The arrow indicates a nascent oral cirrus. Scale bar: 50 μm. —**G**. Cross-section through level 1 of the larva in (F) showing TUNEL-positive epidermal cells along the right margin (asterisk) of the vestibule (ve). Scale bar: 25 μm. —**H**. Anterior end of a juvenile with a few TUNEL-positive nuclei in the epidermis. Scale bar: 50 μm. —**I**. Posterior end of a juvenile showing a few TUNEL-positive nuclei in the tail epidermis. Scale bar: 50 μm.

Stage 3/4 larvae. By these metamorphic stages, the club-shaped gland is no longer visible. Its remnants may be represented by a few TUNEL-positive cells scattered in the haemal space on the right side of the head; alternatively, however, such cells might be apoptotic haemocytes. In addition, an intense zone of TUNEL-positive epidermal cells is now present along the right margin of the vestibule (Fig. 3F,G, asterisk), although apoptotic cells are rare in the nascent oral cirri (Fig. 3F, arrow).

Stage 5 (juvenile). By the beginning of the juvenile stage, no conspicuous zones of TUNEL-positive cells remain in any of the tissues. As in the late pre-metamorphic larvae, most of the apoptotic cells are scattered in the epidermis at the anterior tip (Fig. 3H) and tail (Fig. 3I) of the animal.

Discussion

The function(s) of the club-shaped gland

In the past, many students of amphioxus believed that the club-shaped gland emits secretions (with a variety of proposed functions) from its outer opening (Table 1). In contrast, the results of the present study support the minority view that the gland's secretions exit its inner opening into the pharyngeal lumen (Goldschmidt 1905; Garstang 1928; Olsson 1983; Gilmour 1996; Holland and Yu 2002; Lacalli 2008). The best insights on how the club-shaped gland's secretions function within the pharynx of amphioxus larvae have come from Gilmour (1996), who videotaped the paths of particles in the gut lumen and preserved the presumed mucus trap with ruthenium red-containing fixatives. He concluded that the endostyle and the club-shaped gland cooperate to produce a thin vertical curtain of mucus (orientated mid-sagittally in the pharyngeal lumen) to capture ingested particles and then transport them toward the more posterior gut regions. Subsequently, Lacalli (2008) elaborated on this scheme by proposing that the mucus sheet is produced predominantly by the endostyle and is transported posteriorly by the ciliated peripharyngeal bands. Our present results are in general agreement with those of Gilmour (1996) and Lacalli (2008), although we differ from them in finding that the particle-trapping surface of the mucus may be orientated not vertically near the midline of the pharynx, but at a 35° angle slanting down from the upper lip to the lower limb of the endostyle on the opposite wall of the pharynx. Moreover, in our preparations, the edges of the slanted sheet are not closely associated with the peripharyngeal bands. Certainly, in the fixed larvae studied to date, some features of the mucus trap might be distorted as the result of the shrinkage, extraction and displacement that extracellular materials tend to suffer during histological fixation. A satisfactory understanding of the structure of the particle trap and its detailed functions will require further work.

Apoptosis in metamorphic larvae of amphioxus: fate of the club-shaped gland

Bayascas *et al.* (2002), who previously studied apoptosis in amphioxus larvae (in very early ones that were only 2 days old), described an abundance of TUNEL-positive cells in the ectoderm and endoderm, especially along the left anterior side, where the mouth is opening, and in the differentiating tail. We repeated and confirmed the results of Bayascas *et al.* (2002) on early larvae (our unpublished data). In contrast, our study of later (12-day-old) pre-metamorphic larvae demonstrates little apoptosis in any of their tissues. Soon after the start of metamorphosis, however, numerous epidermal cells around the mouth rim and near the preoral organ become TUNEL-positive. The apoptosis of cells at the mouth rim is evidently related to the shrinkage of the mouth and its anteromedial migration to become the small aperture surrounded by the velum of the juvenile.

Following the initial burst of apoptosis associated with the mouth, virtually all the cells of the glandular portion of the club-shaped gland become TUNEL-positive, and, a few hours later, the massive apoptosis spreads to the cells comprising the duct region. By the time the gill slits of the metamorphic larvae are assuming a U-shape, the apoptotic club-shaped gland appears to disintegrate into scattered TUNEL-positive cells that evidently disperse in the haemal channels and soon disappear, as proposed by van Wijhe (1914). There is no evidence that such cells enter the pharyngeal lumen as suggested by Willey (1891). Moreover, our results conclusively show that the amphioxus club-shaped gland does not transdifferentiate into cartilage-like cells (Klaatsch 1898) and is not converted into any components of the post-metamorphic endostyle (Hatschek 1892; Olsson 1983; Gilmour 1996).

In sum, the club-shaped gland of amphioxus is present throughout the larval stage and appears to supply a component of the mucus that traps food particles in the pharynx. It is likely that the endostyle also produces some of the feeding mucus during the larval stage in addition to its likely fixation of iodine, presumably for the synthesis of thyroid hormones (Fredriksson *et al.* 1984). Following the apoptotic destruction of the club-shaped gland during metamorphosis, the endostyle of juvenile and adult amphioxus presumably becomes the major (or even exclusive) source of mucus for capturing food particles in the pharyngeal lumen. In addition, the post-larval endostyle continues to be the major site for iodination (Ericson *et al.* 1985) possibly because of the continued synthesis of thyroid hormones influencing tissue growth and maintenance in the juvenile and adult amphioxus.

Even with the new information presented here, the homology between the club-shaped gland of amphioxus and structures in other animals remains enigmatic. None of the morphology-based homologies listed in Table 1 has been widely accepted (the comparisons with the chordate mouth/foregut or a chordate gill slit remain the best of an inconclusive lot). Finally, a recent review of developmental genetic data (Jackman *et al.* 2004) gave no clear insights into homologies between the club-shaped gland and structures in other animals, and the authors suggested that the amphioxus club-shaped gland might be an evolutionary novelty that has arisen within the cephalochordate lineage.

Acknowledgements

We are grateful to Linda Holland for her criticism of the manuscript, to John Lawrence and Susan Bell of the Biology Department of the University of South Florida for their generous provision of laboratory space, and to Tom Gilmour for his technical advice. This work was supported in part by NSF grant IOB-04-16292.

References

Bayascas, J. R., Yuste, V. J., Benito, E., Garcia-Fernàndez, J. and Comella, J. X. 2002. Isolation of AmphiCASP-3/7, an ancestral caspase from amphioxus (*Branchiostoma floridae*). Evolutionary considerations for vertebrate caspases. – *Cell Death and Differentiation* **9**: 1078–1089.

Berrill, N. J. 1987. Early chordate evolution. Part 2. Amphioxus and ascidians. To settle or not to settle. – *International Journal of Invertebrate Reproduction and Development* **11**: 15–28.

Bone, Q. 1958. Observations upon the living larva of amphioxus. – *Pubblicazioni della Stazione Zoologica di Napoli* **30**: 458–471.

Conklin, E. G. 1932. The embryology of amphioxus. – *Journal of Morphology* **54**: 69–152.

Ericson, L. E., Fredriksson, G. and Overholm, T. 1985. Ultrastructural localization of the iodination centre in the endostyle of the adult amphioxus (*Branchiostoma lanceolatum*). – *Cell and Tissue Research* **241**: 267–273.

Flood, P. R. 1991. Architecture of, and water circulation and flow rate in, the house of the planktonic tunicate *Oikopleura labradoriensis*. – *Marine Biology* **111**: 95–111.

Fredriksson, G., Ericson, L. E. and Olsson, R. 1984. Iodine binding in the endostyle of larval *Branchiostoma lanceolatum* (Cephalochordata). – *General and Comparative Endocrinology* **56**: 177–184.

Fuller, A. S. 1958. Studies on some Cephalochordata, with special reference to the amphioxides problem. PhD Dissertation, University of London.

Garstang, W. 1928. The morphology of the Tunicata, and its bearings on the phylogeny of the Chordata. – *Quarterly Journal of Microscopical Science* **72**: 51–187.

Gilmour, T. H. J. 1996. Feeding methods of cephalochordate larvae. – *Israel Journal of Zoology* **42**(Suppl.): 87–95.

Godoy, L., Gonzàlez-Duarte, R. and Albalat, R. 2006. *S*-nitrosoglutathinone reductase activity of amphioxus ADH3: insights into the nitric oxide metabolism. – *International Journal of Biological Sciences* **2**: 117–124.

Goldschmidt, R. 1905. Amphioxides. – *Wissenschaftliche Ergebnisse der Deutschen Tiefsee-Expedition auf dem Dampfer 'Valdivia'* **12**: 1–93 + pl. I–X.

Goodrich, E. S. 1930. The development of the club-shaped gland in amphioxus. – *Quarterly Journal of Microscopical Science* **74**: 155–164.

Hatschek, B. 1882. Studien über Entwicklung des Amphioxus. – *Arbeiten aus den Zoologischen Institute der Universität Wien und der Zoologischen Station in Triest* **4**: 1–88 + I–IX.

Hatschek, B. 1892. Die Metamerie des Amphioxus und des Ammocoetes. – *Verhandlungen der Anatomische Gesellschaft* **6**: 136–161.

Holland, N. D. and Holland, L. Z. 1993. Embryos and larvae of invertebrate deuterostomes. In Stern, C. D. and Holland, P.W. H. (Eds): *Essential Developmental Biology: a Practical Approach*, pp. 21–32. IRL Press, Oxford.

Holland, N. D. and Yu, J. K. 2002. Epidermal receptor development and sensory pathways in vitally stained amphioxus (*Branchiostoma floridae*). – *Acta Zoologica (Stockholm)* **83**: 309–319.

Jackman, W. R., Mougey, J. M., Panopoulou, G. D. and Kimmel, C. B. 2004. *crabp* and *maf* highlight the novelty of the amphioxus club-shaped gland. – *Acta Zoologica (Stockholm)* **85**: 91–99.

Klaatsch, H. 1898. Ueber den Bau und die Entwickelung des Tentakelapparates des Amphioxus. – *Anatomischer Anzeiger* **14**: 184–195.

Kowalevsky, A. 1867. Entwickelungsgeschichte des *Amphioxus lanceolatus*. – *Mémoires de l'Académie Impériale des Sciences de St.-Pétersbourg (Série VII)* **11**(4): 1–7 + pl. I–III.

Kowalevsky, A. 1877. Weitere Studien über die Entwickelungsgeschichte des Amphioxus lancelolatus, nebst einem Beitrage zur Homologie des Nervensystems der Würmer und Wirbelthiere. – *Archiv für Microskopische Anatomie* **13**: 181–204 + pl. XV–XVI.

Lacalli, T. 2008. Mucus secretion and transport in amphioxus larvae: organization and ultrastructure of the food trapping system, and implications for head evolution. – *Acta Zoologica Stockholm* **89**: 219–230.

Luft, J. H. 1971. Ruthenium red and violet. I. Chemistry, purification, methods of use for electron microscopy and mechanism of action. – *Anatomical Record* **171**: 347–368.

Neal, H.V. 1898. The segmentation of the nervous system in *Squalus acanthias*. A contribution to the morphology of the vertebrate head. – *Bulletin of the Museum of Comparative Zoology Harvard* **31**: 145–294 + pl. I–IX.

Olsson, R. 1983. Club-shaped gland and endostyle in larval *Branchiostoma lanceolatum* (Cephalochordata). – *Zoomorphology* **103**: 1–13.

Orton, J. H. 1914. On a hermaphrodite specimen of amphioxus with notes on experiments in rearing amphioxus. *Journal of the Marine Biological Association of the United Kingdom* **10**: 506–512.

Paris, M., Escriva, H., Schubert, M., Burnet, F., Brtko, J., Ciesielski, F. *et al*. 2008. Amphioxus post-embryonic development reveals the homology of chordate metamorphosis. – *Current Biology* **18**: 825–830.

Schulze, M. 1851. Beobachtung junger Exemplare von Amphioxus. – *Zeitschrift für Wissenschaftliche Zoologie* **3**: 416–419 + pl. XII.

Van Beneden, E. and Julin, C. 1886. Recherches sur la morphologie des tuniciers. – *Archives de Biologie* **6**: 237–476 + pl. VII–XVI.

Van Wijhe, J. W. 1914. Studien über Amphioxus. I. Mund und Darmkanal während der Metamorphose. – *Verhandelingen der Koninklijke Akademie van Wetenschappen te Amsterdam (Section 2)* **18**: 1–84 + pl. I–V.

Webb, J. E. 1969. On the feeding and behaviour of the larva of *Branchiostoma lanceolatum*. – *Marine Biology* **3**: 58–72.

Wickstead, J. H. 1967. *Branchiostoma lanceolatum* larvae: some experiments on the effect of thiouracil on metamorphosis. – *Journal of the Marine Biological Association of the United Kingdom* **47**: 49–59.

Willey, A. 1891. The later larval development of amphioxus. – *Quarterly Journal of Microscopical Science* **32**: 183–234 + pl. XIII–XV.

Yasui, K., Zhang, S. C., Uemura, M. and Saiga, H. 2000. Left–right asymmetric expression of *BbPtx*, a Ptx-related gene, in a lancelet species and the developmental left-sidedness of deuterostomes. – *Development* **127**: 187–195.

Dev Genes Evol
DOI 10.1007/s00427-008-0251-y

ORIGINAL ARTICLE

Nuclear hormone receptor signaling in amphioxus

Michael Schubert · Frédéric Brunet · Mathilde Paris · Stéphanie Bertrand · Gérard Benoit · Vincent Laudet

Received: 11 July 2008 / Accepted: 5 September 2008

Abstract The nuclear hormone receptors (NRs) form a superfamily of transcription factors unified by conserved protein structure and mode of function. While most members of this superfamily are activated by ligands, such as thyroid hormones, steroids, vitamin D or retinoic acid, other NRs are called orphan receptors because they have no known ligand. NR-dependent signaling is crucial for vertebrate development with the majority of receptors being expressed in the developing embryo. Due to massive gene duplications during vertebrate diversification, there are usually more NRs in vertebrates than in invertebrates. In this study, we examine the evolutionary diversification of the NR superfamily and of NR-dependent signaling in chordates (vertebrates, tunicates, and amphioxus). We take advantage of the unique features of the genome of the invertebrate amphioxus, which is characterized by a vertebrate-like gene content without having undergone massive duplications, to assess the NR signaling complement (NRs and NR coregulators) of the ancestral chordate. We find 33 NRs in amphioxus, which are more NRs than originally anticipated. This increase is mainly due to an amphioxus-specific duplication of genes encoding receptors of the NR1H group. In addition, there are three heterologous NRs in amphioxus that could not be placed within the framework of the NR superfamily. Apart from these exceptions, there is usually one amphioxus NR or NR signaling coregulator for each paralogous group of two, three, or four human receptors suggesting that the ancestral chordate had a set of 22 different NRs plus one copy of each NR coregulator.

Keywords *Branchiostoma floridae* · Cephalochordate · Chordate · Development · Evolution · Invertebrate-to-vertebrate transition · Nuclear hormone receptor coregulator

Michael Schubert and Frédéric Brunet are equal first authors.

Communicated by J.J. Gibson-Brown

Electronic supplementary material The online version of this article (doi:10.1007/s00427-008-0251-y) contains supplementary material, which is available to authorized users.

M. Schubert · F. Brunet · M. Paris · G. Benoit · V. Laudet (✉)
CNRS UMR5242-INRA 1288-ENS-UCBL,
Institut de Génomique Fonctionnelle de Lyon, IFR128
BioSciences Lyon-Gerland, Ecole Normale Supérieure de Lyon,
46 allée d'Italie,
69364 Lyon Cedex 07, France
e-mail: Vincent.Laudet@ens-lyon.fr

S. Bertrand
CNRS UMR7628, UPMC Univ Paris 06,
Observatoire Océanographique,
66651 Banyuls-sur-Mer, France

Introduction

Nuclear hormone receptors (NRs) constitute a superfamily of DNA-binding transcription factors that are involved in a wide variety of physiological processes both during development and in the adult (reviewed in Gronemeyer et al. 2004). NRs are best known for their functions as ligand-activated transcription factors that mediate the response to hormones, such as thyroid hormones, steroids, vitamin D or retinoic acid (reviewed in Gronemeyer et al. 2004). In addition to this role, there is also a non-negligible number of NRs that either do not bind a ligand at all or for which a ligand has not yet been described. These NRs without known ligand are called orphan receptors (reviewed in Benoit et al. 2006; Germain et al. 2006c). Phylogenetic

analyses have shown that NRs arose very early in the metazoan lineage long before the divergence of protostomes and deuterostomes and diversified through complex series of gene duplications and gene losses (Escriva et al. 1997; Bertrand et al. 2004; reviewed in Escriva et al. 2000).

NR proteins are characterized by a modular structure consisting of six domains (reviewed in Germain et al. 2006c). At the N-terminus, the A and B domains contain a transcriptional activation unit called activation function 1 (AF-1). The C region of the NR is the DNA-binding domain (DBD), which is required for the interaction of the NR protein with the DNA. The DBD is the most highly conserved domain of NRs and consists of cysteine-rich zinc finger motifs, α-helices, and a C-terminal extension. The D domain acts as a hinge between the C region (the DBD) and the E region, which corresponds to the ligand-binding domain (LBD) and which includes a second activation domain (AF-2) crucial for interactions of NRs with coactivators and corepressors and often also a nuclear localization signal. Finally, the F region, which is not present in all NRs, is not very well characterized and has no clear known function.

Interactions of NRs with coactivators and corepressors are required for defining the transcriptional activities of NRs, which can function both as activators and repressors of gene transcription depending on the target gene and on the presence or absence of a ligand (reviewed in Bastien and Rochette-Egly 2004; Germain et al. 2006c). Although transcriptional regulation by NRs is very complex and diverse, in the absence of ligand, a liganded NR (such as the retinoic acid receptor, RAR, or the thyroid hormone receptor, TR) is bound to DNA predominantly as heterodimer with the NR retinoid X receptor (RXR) and is associated with a corepressor complex that represses the transcriptional activity of the NRs (reviewed in Hu and Lazar 2000; Bastien and Rochette-Egly 2004; Goodson et al. 2005; Germain et al. 2006c). In some cases, NRs can also bind to DNA as homodimers (e.g. the hepatocyte nuclear factor 4, HNF4) or as monomers (e.g. the nerve growth factor-induced clone—B, NGFI-B) (reviewed in Benoit et al. 2006; Germain et al. 2006c). Ligand binding induces a conformational change in the NR leading to the dissociation of the corepressors and to the binding of a coactivator complex (reviewed in Bastien and Rochette-Egly 2004; Germain et al. 2006c). The activity of this coactivator complex leads to unfolding of the condensed chromatin (reviewed in Bastien and Rochette-Egly 2004; Germain et al. 2006c; Lonard and O'Malley 2006). After this relaxation of the chromatin structure, the coactivators are replaced by a mediator complex that recruits the basal transcription machinery to the DNA locus for initiation of transcription (reviewed in Bastien and Rochette-Egly 2004; Germain et al. 2006c).

Most NRs are expressed during vertebrate development (Bertrand et al. 2007) and have important functions in the developing embryo (reviewed in Laudet and Gronemeyer 2002). For example, NRs are involved in survival and proper development of the early embryo (PPAR, HNF4, ERR, NOR1, LRH1, GCNF), in patterning and formation of the central nervous system (TR, RAR, REV-ERB, ROR, TR2/4, TLX, PNR, COUP-TF, EAR2, ER, NURR1, NOR1, SF1), in organogenesis of, among others, the heart, lung, kidney, liver, pancreas, and the immune and reproductive systems (DAX1, RAR, PPAR, ROR, HNF4, RXR, COUP-TF, GR, PR, AR, SF1), and in the development of the visual system (RAR, ROR, TLX, PNR) (reviewed in Benoit et al. 2006; Dahlman-Wright et al. 2006; Flamant et al. 2006; Germain et al. 2006a; Germain et al. 2006b; Lu et al. 2006; Michalik et al. 2006; Moore et al. 2006). Moreover, in vertebrates (most prominently in amphibians and teleost fish) and in invertebrate chordates (such as amphioxus), the thyroid hormone receptor (TR) controls metamorphosis, the postembryonic transition from larva to adult (Paris et al. 2008a). Along the same lines, the functions of another NR, the retinoic acid receptor (RAR), in early development are also conserved between invertebrate chordates and vertebrates. Thus, it has been shown that in amphioxus and vertebrates, RAR controls expression of *Hox* genes, which are crucial mediators of anteroposterior axial patterning in development (Schubert et al. 2006b; reviewed in Marlétaz et al. 2006; Campo-Paysaa et al. 2008).

In humans, there are a total of 48 NRs. In contrast, the invertebrates studied so far generally have fewer genes encoding NR proteins. For example, there are 21 NRs in the genome of the fruit fly *Drosophila melanogaster* and only 17 NRs in the genome of the sea squirt *Ciona intestinalis* (Yagi et al. 2003; Bertrand et al. 2004). This difference in the total number of NRs between invertebrates and vertebrates can be attributed to massive whole genome duplications that accompanied the diversification of the vertebrates (Holland et al. 1994; Dehal and Boore 2005; Holland et al. 2008; Putnam et al. 2008; reviewed in Schubert et al. 2006a). Phylogenetic analyses of the complete NR superfamily have shown that the vast majority of NRs from metazoan animals fall into six distinct subfamilies (numbered NR1 through NR6). In addition, there is a seventh "subfamily" (called NR0), which contains all receptors lacking the typical NR protein structure (Fig. 1) (Nuclear Receptors Nomenclature Committee 1999; reviewed in Germain et al. 2006c). Within each subfamily, the NRs are further subdivided into several groups (Fig. 1) (reviewed in Germain et al. 2006c). To better understand the evolutionary diversification of the NR superfamily and of NR-dependent signaling in chordates (i.e. in vertebrates, tunicates, and amphioxus), we decided to characterize the complement of NRs and of key

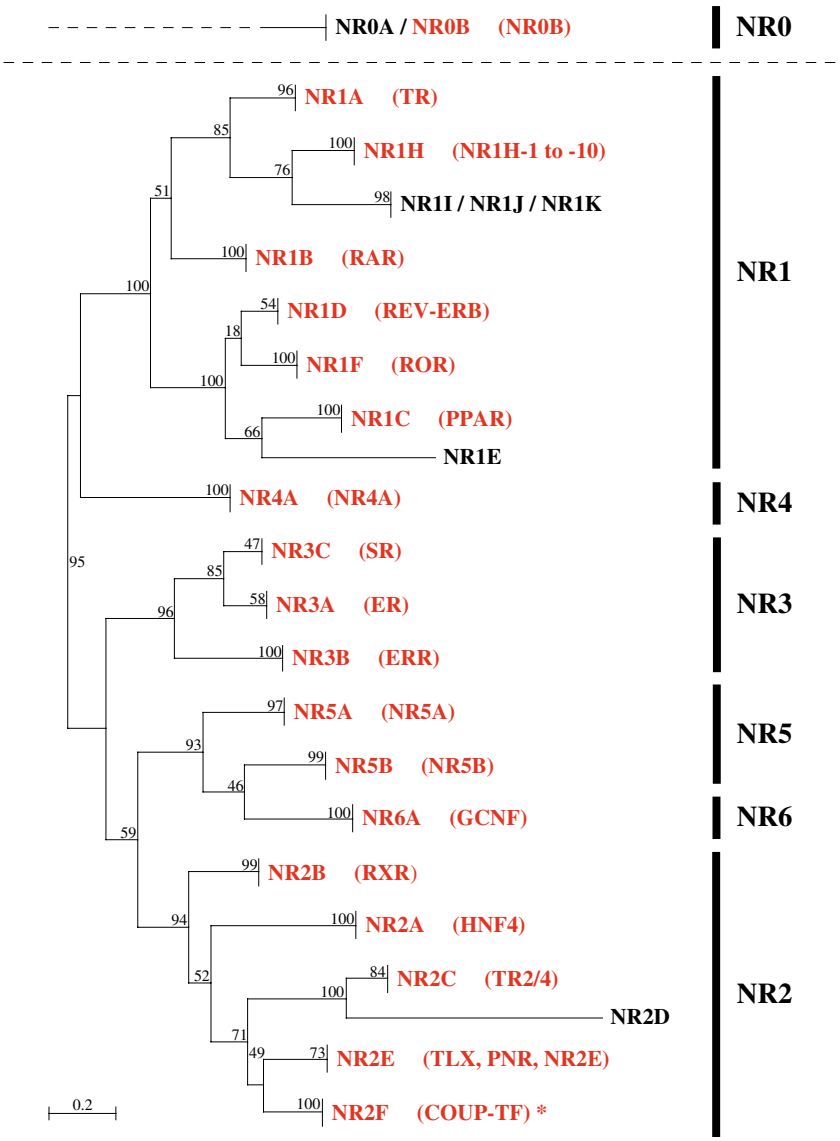


Fig. 1 Diagrammatic phylogeny of the nuclear hormone receptor (NR) superfamily. The overall topology and bootstrap support values are derived from a NR superfamily tree including NRs from humans, sea squirts, fruit flies, and amphioxus. NR0 has been added for illustration purposes. In *red* are NR groups that include amphioxus sequences. The names of the respective amphioxus NRs are indicated in *brackets*. In the NR1H group, there are ten amphioxus NRs called NR1H-1, NR1H-2, NR1H-3, NR1H-4, NR1H-5, NR1H-6, NR1H-7, NR1H-8, NR1H-9, and NR1H-10. * indicates that a COUP-TF receptor has previously been described in amphioxus that could not be identified in the genome

components of NR signaling in the genome of the cephalochordate amphioxus (*Branchiostoma floridae*). Of the two invertebrate chordate groups (tunicates and amphioxus), we chose amphioxus as a model for our studies, because it is now widely regarded as the best available stand-in for the chordate ancestor (Yu et al. 2007; Holland et al. 2008; Putnam et al. 2008). Moreover, amphioxus is characterized by an overall body plan and a genome that are vertebrate-like, but simpler (reviewed in Schubert et al. 2006a; Garcia-Fernàndez et al. 2007). For example, due to whole genome duplications in the vertebrate lineage, there are often two, three, or four vertebrate genes, where amphioxus only has a single ortholog (Holland et al. 2008; Putnam et al. 2008). Moreover, unlike the genomes of representatives of the other invertebrate chordate group, the tunicates (e.g., the sea squirt *Ciona intestinalis*), the amphioxus genome has not experienced any major reorganization by lineage-specific gene losses or duplications (Dehal et al. 2002; Holland et al. 2008; Putnam et al. 2008).

Taking advantage of these unique features of the amphioxus genome, we have identified the complete set of NRs in amphioxus. In addition, we have characterized the most important coregulators of the NR signaling machinery in amphioxus. We found that amphioxus has a total of 33 NRs, which is significantly more than expected (Holland et al. 2008). This high number of NRs in amphioxus is mainly due to a lineage-specific duplication of NR1H receptors, which could be related to an elaboration in amphioxus of protection mechanisms against xenobiotics and toxins. Apart from this amphioxus-specific duplication, there is generally one amphioxus receptor for two, three, or four human NRs and NR coregulators. For example, we identified one ortholog of the estrogen receptor (ER) and one ortholog of the steroid receptor

(SR) in amphioxus, while there are two ERs and four SRs in humans. Thus, both ER and SR have already been present in the last common ancestor of all chordates and have secondarily been lost in sea squirts and other tunicates. In sum, the data on NRs and NR coregulators in the amphioxus genome suggest that the ancestral chordate had a set of 22 different NRs and a single copy of each NR coregulator.

Materials and methods

Data mining

The NR data set was based on sequence alignments of previous studies (Bertrand et al. 2004). The alignment file already included all NRs from humans (*Homo sapiens*), sea squirts (*Ciona intestinalis*), and fruit flies (*Drosophila melanogaster*). In contrast, the NR coregulator data set was based on the Ensembl families of the NR coregulators of interest (NCOA, CBP and P300, PGC1, NCOR/SMRT, MED1/TRAP220), which are available at the Ensembl genome browser (www.ensembl.org/index.html). We used Ensembl version v48 for this analysis. All NR and coregulator sequences were BLASTed against the amphioxus genome (version v1.0), which is available at the JGI webpage (genome.jgi-psf.org/Brafl1/Brafl1.home.html). The best hits were retrieved and aligned using MUSCLE (Edgar 2004) followed by manual refinement. We created both nucleotide and amino acid alignments. All alignment files are available from the authors. The nucleotide alignments were used to identify and eliminate allelic polymorphs of the different NR and NR coregulator sequences retrieved from the amphioxus genome. Amphioxus allelic polymorphs are characterized by nucleotide sequences that differ almost exclusively in third codon positions, which allows their identification and elimination (Holland et al. 2008; Putnam et al. 2008).

Phylogenetic analyses

The amino acid alignments of amphioxus NR and NR coregulator sequences were used for phylogenetic tree reconstruction experiments. Phylogenetic analyses on the comprehensive NR data set including all NRs from humans, sea squirts, fruit flies and amphioxus were carried out using both the neighbor joining (NJ) and maximum likelihood (ML) method. Likewise, the different NR coregulator alignments were analyzed by phylogenetic tree reconstruction using both the NJ and ML method. The NJ analyses, carried out with the program Phylo_Win (Galtier et al. 1996), were based on a Poisson correction and pairwise gap removal. The ML trees were calculated using PhyML (Guindon and Gascuel 2003) with default parameters (a JTT model with a gamma distribution parameter of 2). Bootstrap support values were calculated in 100 bootstrap replicates. The NJ and ML analyses yielded identical results and, hence, only the ML trees are discussed.

Results

In this study, we have annotated from the amphioxus genome all NR genes and the genes of some of the main coregulators of NR signaling. Gene annotation was based on phylogenetic analyses of the identified sequences in the context of the NR superfamily or of the coregulator protein families (Supplementary Figs. S1–S6). The sequence IDs in the amphioxus genome of the amphioxus NRs and of the NR coregulators are given in the supplementary material (Supplementary Tables S1, S2).

Amphioxus NRs

In total, there are 33 NR genes in amphioxus, 32 of which we were able to identify directly in the amphioxus genome (Fig. 1, Table 1). Although amphioxus COUP-TF had previously been cloned from amphioxus (Langlois et al. 2000), we were not able to identify this receptor in the genome probably because it is located in an unassembled region of the amphioxus genome (Putnam et al. 2008). In the following, we will discuss the different NR subfamilies according to the overall structure and nomenclature of the superfamily (Fig. 1).

In the NR1 subfamily, there are a total of 15 amphioxus NRs (Table 1). As expected, there is a single amphioxus ortholog in most NR1 groups. For example, there is a single amphioxus thyroid hormone receptor (TR), retinoic acid receptor (RAR), REV-ERB receptor, retinoid-related orphan receptor (ROR) and peroxisome proliferator-activated receptor (PPAR; Fig. 2a). The amphioxus TR and RAR have previously been cloned and functionally characterized (Escriva et al. 2002; Paris et al. 2008a). In contrast, we were not able to find any amphioxus representatives of the NR1I/J/K group, which includes, for example, the vertebrate vitamin D receptors (VDRs), pregnane X receptors (PXRs) and constitutive androstene receptors (CARs). In the NR1H group (Fig. 2b), which contains the vertebrate liver X receptors (LXRs) and farnesoid X receptors (FXRs), the amphioxus receptors have experienced a lineage-specific duplication resulting in ten NR1H-like receptors (Table 1). Since two of these independent amphioxus duplicates are very short, containing only the DBD (NR1H-9 and NR1H-10), they could not be included in our phylogenetic analyses of the whole NR superfamily

Table 1 List of nuclear hormone receptors (NRs) in humans (*Homo sapiens*), sea squirts (*Ciona intestinalis*), fruit flies (*Drosophila melanogaster*) and amphioxus (*Branchiostoma floridae*)

NR group	NR name in humans	NR name in sea squirts	NR name in fruit flies	NR name in amphioxus
NR0A			KNI, KNRL, EG	
NR0B	DAX1, SHP			NR0B
NR1A	TRα, TRβ	TR		TR
NR1B	RARα, RARβ, RARγ	RAR		RAR
NR1C	PPARα, PPARβ, PPARγ	PPAR		PPAR
NR1D	REV-ERBα, REV-ERBβ	REV-ERB	E75	REV-ERB
NR1E			E78	
NR1F	RORα, RORβ, RORγ	ROR	HR46	ROR
NR1H	LXRα, LXRβ, FXRα	LXR, FXR	ECR	NR1H-1
NR1H	LXRα, LXRβ, FXRα	LXR, FXR	ECR	NR1H-2
NR1H	LXRα, LXRβ, FXRα	LXR, FXR	ECR	NR1H-3
NR1H	LXRα, LXRβ, FXRα	LXR, FXR	ECR	NR1H-4
NR1H	LXRα, LXRβ, FXRα	LXR, FXR	ECR	NR1H-5
NR1H	LXRα, LXRβ, FXRα	LXR, FXR	ECR	NR1H-6
NR1H	LXRα, LXRβ, FXRα	LXR, FXR	ECR	NR1H-7
NR1H	LXRα, LXRβ, FXRα	LXR, FXR	ECR	NR1H-8
NR1H	LXRα, LXRβ, FXRα	LXR, FXR	ECR	NR1H-9[a]
NR1H	LXRα, LXRβ, FXRα	LXR, FXR	ECR	NR1H-10[a]
NR1I	VDR, PXR, CAR			
NR1J			HR96	
NR1K		NR1K-1, NR1K-2		
NR2A	HNF4α, HNF4γ	HNF4	HNF4	HNF4
NR2B	RXRα, RXRβ, RXRγ	RXR	USP	RXR
NR2C	TR2, TR4	TR2/4		TR2/4
NR2D			HR78	
NR2E	TLX		TLL, DSF	TLX
NR2E	PNR		HR51	PNR
NR2E				NR2E
NR2E			FAX1	
NR2F	COUP-TFα, COUP-TFβ, EAR2	COUP-TF	SVP	COUP-TF[b]
NR3A	ERα, ERβ			ER
NR3B	ERRα, ERRβ, ERRγ	ERR	ERR	ERR
NR3C	GR, MR, AR, PR			SR
NR4A	NGFI-B, NURR1, NOR1	NR4A	HR38	NR4A
NR5A	SF1, LRH1	NR5A	FTZ-F1	NR5A
NR5B			HR39	NR5B
NR6A	GCNF	GCNF	HR4	GCNF
NRa				NRa
NRb				NRb
NRc				NRc[a]
Total	48 NRs in humans	17 NRs in sea squirts	21 NRs in fruit flies	33 NRs in amphioxus

[a] DNA-binding domain only.
[b] COUP-TF has been cloned from amphioxus (*Branchiostoma floridae*), but could not be identified in the genome sequence.

(Fig. 2b). We have, thus, verified that these two amphioxus receptors (NR1H-9 and NR1H-10) are members of the NR1H group by constructing a phylogenetic tree based only on the DBD regions of the NR1A, NR1H, and NR1I/J/K sequences (data not shown). Interestingly, in the amphioxus genome NR1H-1, NR1H-2, NR1H-3, and NR1H-4 are located on the same scaffold (on scaffold 153), as are NR1H-5 and NR1H-7 (on scaffold 291) and NR1H-8 and NR1H-9 (on scaffold 91). This linkage pattern strongly suggests that the diversity of amphioxus NR1H receptors is the result of a series of lineage-specific tandem duplication events. Within the NR1 subfamily, the branch support for the association of the amphioxus receptors with their respective NR group is very strong except for amphioxus REV-ERB, where bootstrap support is only at 73%. This bootstrap value increases to 82% in a tree based

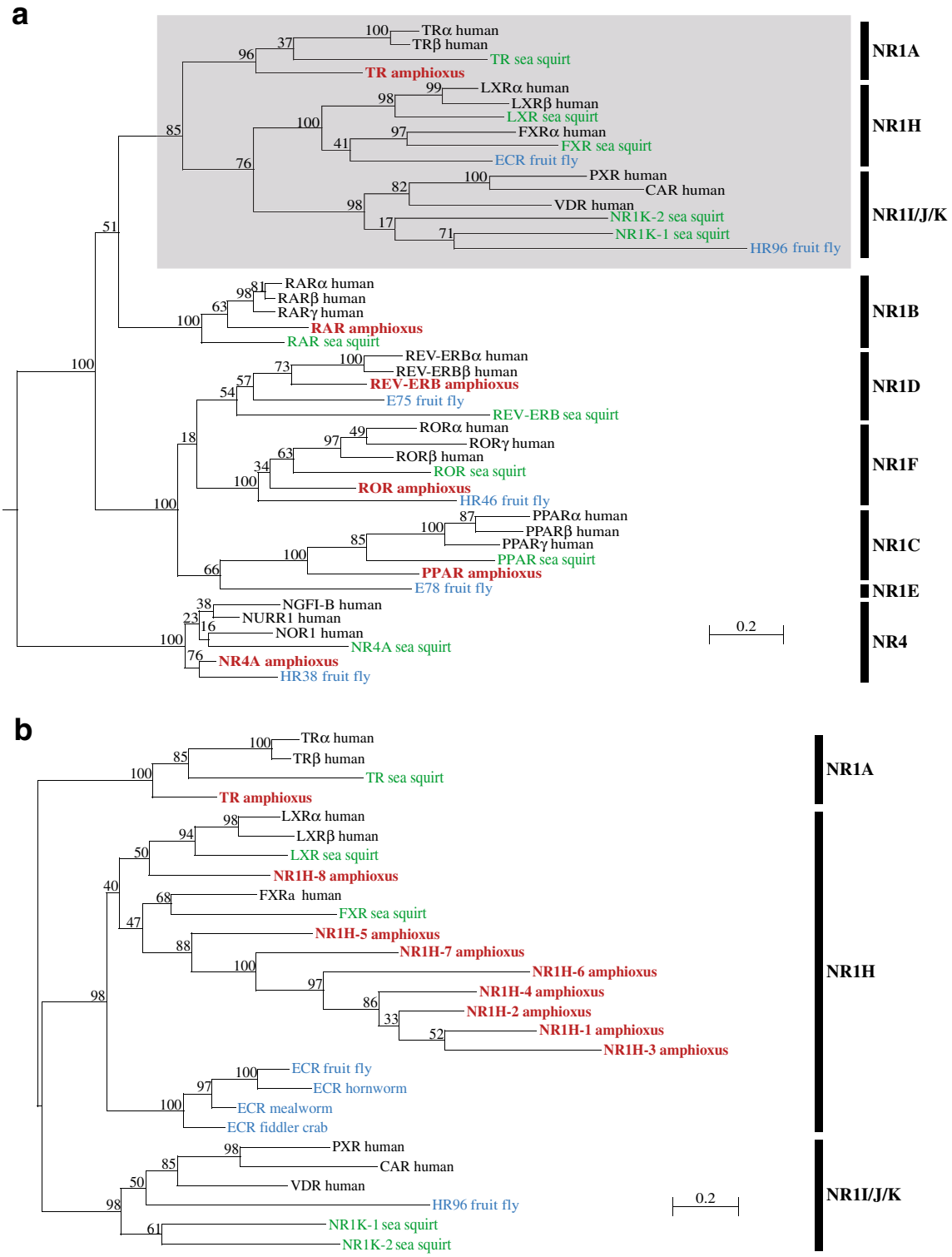


Fig. 2 Nuclear hormone receptor (NR) phylogeny of the NR1 and NR4 subfamilies. **a** The NR1 and NR4 branches of the overall NR superfamily tree based on NRs from humans, sea squirts, fruit flies, and amphioxus. The *grey shading* highlights the NR1 groups used for calculating the phylogeny shown in (**b**). Human NRs are shown in *black*, sea squirt NRs in *green*, fruit fly NRs in *blue*, and amphioxus NRs in *red*. Bootstrap support values are given for each branch. **b** Phylogenetic analysis of the NR1A, NR1H, and NR1I/J/K groups highlighting the lineage-specific duplication of NR1H receptors in amphioxus. Human NRs are shown in *black*, sea squirt NRs in *green*, fruit fly NRs in *blue*, and amphioxus NRs in *red*. Bootstrap support values are given for each branch

only on sequences of the NR1C, NR1D, NR1E, and NR1F groups (data not shown). We are, hence, confident that the amphioxus genome contains a true REV-ERB ortholog.

We found a total of seven amphioxus receptors of the NR2 subfamily (Table 1). In the amphioxus genome, there is a single gene encoding retinoid X receptor (RXR), hepatocyte nuclear factor 4 (HNF4), testicular receptor 2/4 (TR2/4) and tailless (TLX) (Fig. 3). Moreover, there is a single amphioxus photoreceptor-specific nuclear receptor (PNR) and an additional amphioxus receptor that is associated with the NR2E group (Fig. 3). It is very likely that this second amphioxus NR2E receptor is the result of a lineage-specific duplication, although our phylogeny does not provide conclusive evidence for this hypothesis (Fig. 3). The amphioxus NR2 receptors strongly branch within their respective NR groups, the weakest being amphioxus TLX with a branch support of 92%. In addition to RXR and TR2/4, a chicken ovalbumin upstream promoter—transcription factor (COUP-TF) ortholog has previously been described in amphioxus (Langlois et al. 2000; Escriva et al. 2002). Thus, although we were not able to find amphioxus COUP-TF in the amphioxus genome sequence, this receptor does exist in amphioxus.

In the NR3 subfamily, there is a single amphioxus receptor for each group (Table 1). Thus, there is one amphioxus steroid receptor (SR), one estrogen receptor (ER) and one estrogen receptor-related receptor (ERR), all of which have previously been described (Horard et al. 2004; Bardet et al. 2005; Paris et al. 2008b). Although the branch support for the association of amphioxus SR with

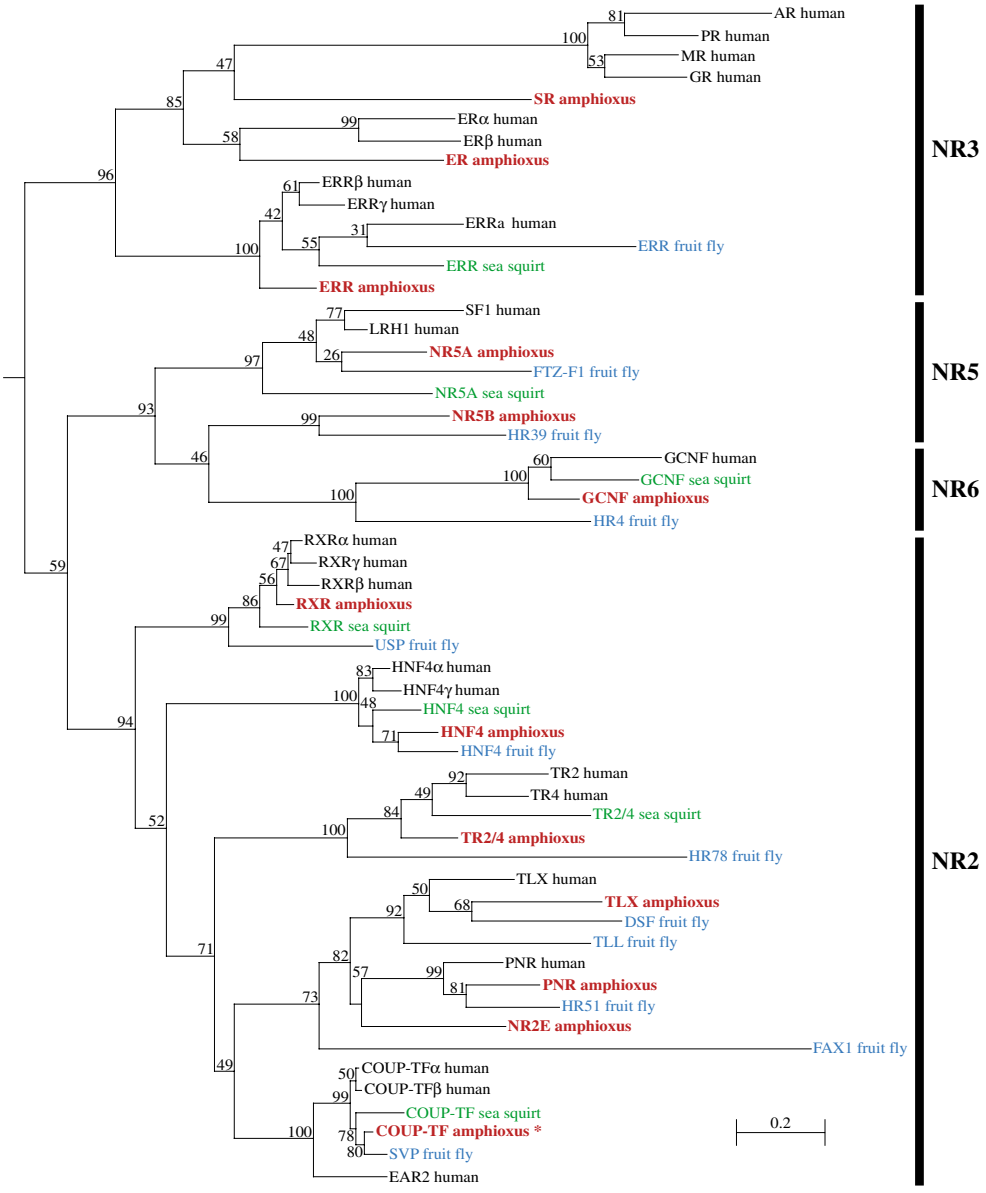


Fig. 3 Nuclear hormone receptor (NR) phylogeny of the NR2, NR3, NR5, and NR6 subfamilies. The tree shows the NR2, NR3, NR5, and NR6 branches of the overall NR superfamily tree calculated with NRs from humans, sea squirts, fruit flies, and amphioxus. Human NRs are shown in *black*, sea squirt NRs in *green*, fruit fly NRs in *blue*, and amphioxus NRs in *red*. Bootstrap support values are given for each branch. *Asterisk* indicates that a COUP-TF receptor has previously been described in amphioxus that could not be identified in the genome

the vertebrate SRs (i.e. with the glucocorticoid receptor, GR, the mineralocorticoid receptor, MR, the androgen receptor, AR, and the progesterone receptor, PR) and of amphioxus ER with vertebrate ERs is not very strong (bootstrap of 47% and 58%, respectively), the overall association of these two amphioxus receptors with the NR3 subfamily is very well supported (bootstrap of 96%) (Fig. 3). In a phylogeny of the NR3 subfamily using the NR2B group (i.e. the RXRs) as outgroup, the bootstrap support for the branching of amphioxus SR with vertebrate SRs increases to 86%, while the grouping of amphioxus ER with its vertebrate orthologs is still only weakly supported (bootstrap of 57%) (data not shown). Similar conclusions have been reached in a recent detailed phylogenetic analysis of the NR3 subfamily (Paris et al. 2008b). To further characterize the three amphioxus NR3 receptors, we compared the amino acid sites required for DNA and ligand binding in human NR3 receptors with those of amphioxus (Table 2). Unfortunately, the amphioxus SR sequence we identified in the genome is incomplete and lacks the C-terminus. We were, hence, unable to analyze three sites in the C-terminal region of the LBD of amphioxus SR (Table 2). Within the DBD, the P-box of amphioxus ER is identical to that of human ERs. In contrast, the LBD sequences of amphioxus ER and the human ERs are less conserved with one of the three amino acids involved in ligand binding in human ERs being different in amphioxus ER (Table 2). Our comparison also indicates that the amphioxus SR is more ER-like than SR-like, both as far as DNA and ligand binding is concerned (Table 2). Thus, the P-box and the D-box in the DBD of the amphioxus SR are more like human ERs and of the seven key amino acid positions in the LBD, four are ER-like (T141, E147, L181, and I218) and none is SR-like (two amino acids, R188 and F199, are conserved with both human ERs and SRs and one, A138, is divergent in amphioxus SR) (Table 2).

In humans, the NR4 subfamily comprises three receptors: the nerve growth factor-induced clone—B (NGFI-B), the NUR-related factor 1 (NURR1) and the neuron-derived orphan receptor 1 (NOR1). As expected, there is only a single NR4 subfamily member in amphioxus, which is called NR4A (Table 1). This amphioxus receptor groups very strongly within this NR subfamily (100% bootstrap support) (Fig. 2a).

In our phylogenetic analysis, some NR5 subfamily members are weakly associated with the NR6 subfamily (Fig. 3). Thus, one of the two amphioxus NR5 subfamily members (NR5A) robustly groups (with a bootstrap support of 97%) with the other members of the NR5A group, which includes the human steroidogenic factor 1 (SF1) and liver receptor homolog 1 (LRH1), while the second amphioxus NR5 receptor (NR5B), together with the fruit fly hormone receptor-like in 39 (HR39), is weakly associated with the

Table 2 Amino acids important for DNA (DBD) and ligand (LBD) binding in NR3 subfamily receptors

Nuclear receptor	DBD		LBD							
	P-box	D-box								
Human ERα	CEGCKA	PATNQ	G (344)	T (247)	**E (353)**	L (387)	**R (394)**	F (404)	I (424)	**H (524)**	L (525)	M (528)
Human ERβ	CEGCKA	PATNQ	M (296)	T (299)	**E (305)**	L (339)	**R (346)**	F (356)	I (376)	**H (475)**	L (476)	M (479)
Amphioxus ER	CEGCKA	PGTNQ	E (481)	S (484)	E (490)	L (524)	C (531)	L (542)	M (570)	H (659)	L (660)	V (663)
Human AR	CGSCKV	ASRND	S (702)	**N (705)**	**Q (711)**	S (745)	**R (752)**	F (764)	Q (783)	F (876)	**T (877)**	L (780)
Human GR	CGSCKV	AGRND	T (561)	**N (564)**	**Q (570)**	M (604)	**R (611)**	F (623)	**Q (642)**	Y (735)	C (736)	**T (739)**
Human MR	CGSCKV	AGRND	S (767)	**N (770)**	**Q (776)**	**S (810)**	**R (817)**	**F (829)**	L (848)	F (941)	**C (942)**	**T (945)**
Human PR	CGSCKV	AGRND	S (716)	**N (719)**	**Q (725)**	S (759)	**R (766)**	F (778)	L (797)	Y (890)	C (891)	T (894)
Amphioxus SR	CEGCKS	PANNN	A (138)	T (141)	E (147)	L (181)	R (188)	F (199)	I (218)	–	–	–
Human ERRα	CEACKA	PASNE	A (323)	C (326)	E (332)	L (366)	R (373)	F (383)	L (402)	H (495)	F (496)	C (499)
Human ERRβ	CEACKA	PASNE	T (241)	C (244)	E (250)	L (284)	R (291)	Y (301)	L (320)	H (408)	F (409)	C (412)
Human ERRγ	CEACKA	PASNE	T (243)	C (246)	E (252)	L (286)	R (293)	Y (303)	L (322)	H (410)	F (411)	I (414)
Amphioxus ERR	CEACKA	PATNE	T (289)	C (292)	E (298)	L (332)	R (339)	F (349)	L (358)	H (456)	F (457)	I (460)

Amino acids known to interact with the ligand in human ERα (Tanenbaum et al. 1998), ERβ (Pike et al. 1999), AR (Pereira de Jésus-Tran et al. 2006), GR (Bledsoe et al. 2002), MR (Bledsoe et al. 2005) and PR (Williams and Sigler 1998) are shown in *bold*. Amino acid positions are given in parentheses.

NR6 subfamily (Fig. 3). A phylogenetic analysis based on the NR5 and NR6 subfamily sequences and using the NR2B group (i.e. the RXRs) as outgroup recovers the integrity of the NR5 and NR6 subfamilies, with strong support for the NR6 subfamily (97% bootstrap support) and weak support for the NR5 subfamily (46% bootstrap support) (data not shown). It is interesting to note that there are no NR5B receptors in humans and sea squirts (Table 1) suggesting that members of the NR5B group might have secondarily been lost in the lineages leading to extant sea squirts and vertebrates. In the NR6 subfamily that includes the vertebrate germ cell nuclear factor (GCNF), there is only a single amphioxus receptor (Table 1), which branches very robustly within this subfamily (Fig. 3).

The NR superfamily also includes receptors that are not characterized by the typical NR protein structure. These receptors are grouped together in the NR0 "subfamily" (Nuclear Receptors Nomenclature Committee 1999; reviewed in Germain et al. 2006c). In the amphioxus genome, we found only one member of this NR0 "subfamily", a member of the NR0B group that is orthologous to human dosage-sensitive sex reversal–adrenal hypoplasia congenita critical region of the X chromosome protein 1 (DAX1) and small heterodimer partner (SHP) (Table 1). Since the amphioxus NR0B protein lacks a DBD, it could not be included in the overall NR superfamily phylogeny, but an alignment of human DAX1 and SHP with the single amphioxus NR0 subfamily member clearly identifies this amphioxus receptor as a member of the NR0B group (data not shown). Amphioxus NR0B shares almost identical levels of sequence identity with the two human NR0B receptors DAX1 (38.4%) and SHP (37.9%).

Apart from the amphioxus NRs that we were able to categorize into existing NR subfamilies, we also found three divergent amphioxus NRs (called NRa, NRb, and NRc) that we failed to place within the existing framework of the NR superfamily (Fig. 1, Table 1). Thus, while NRa is characterized by the typical organization of a NR receptor, NRb has a conserved DBD, but very divergent C-terminal LBD region, and NRc is very short consisting only of a putative DBD region. Since we were unable to reliably place the three divergent amphioxus NRs in the context of a phylogenetic tree of the NR superfamily, we carried out BLAST analyses of the protein sequences to obtain an idea about the sequence similarities of amphioxus NRa, NRb, and NRc with representatives of the different NR subfamilies. The BLAST results indicate that NRa can loosely be associated with the NR2 or NR3 subfamilies, NRb with the NR5 or NR6 subfamilies and NRc with the NR1 subfamily (data not shown). Interestingly, an analysis of the three divergent amphioxus receptors in the context of the sea urchin NR complement (Howard-Ashby et al. 2006; Sea Urchin Genome Sequencing Consortium 2006) revealed that amphioxus NRa robustly branches with the sea urchin receptor Sp-nr2C (data not shown). This result suggests that with the availability of more genomic data (both from amphioxus as well as from other species), these three amphioxus receptors can hopefully be further characterized and categorized in the near future.

Amphioxus coregulators

Coregulators are very important components of NR signaling. Thus, a search of NR coregulators in the amphioxus genome allows important insights into the conservation (but also into possible differences) of the NR signaling machinery in amphioxus and vertebrates. We have focused on three classes of NR signaling coregulators: coactivators, corepressors, and components of the mediator complex (Table 3). We identified amphioxus orthologs of the classical vertebrate nuclear receptor coactivator (NCOA), of CREB-binding protein (CBP) and histone acetyltransferase p300 (P300), both of which are members of the same coregulator family, as well as of peroxisome proliferator-activated receptor gamma coactivator 1 (PGC1). For each of these coactivator families, we only found a single amphioxus ortholog (Table 3): amphioxus NCOA (Supplementary Fig. S2), amphioxus CBP/P300 (Supplementary Fig. S3) and amphioxus PGC1 (Supplementary Fig. S4). For corepressors, we searched the amphioxus genome for orthologs of the classical vertebrate nuclear receptor corepressor/silencing mediator for retinoid and thyroid hormone receptors (NCOR/SMRT) and identified only a single amphioxus NCOR (Table 3; Supplementary Fig. S5). Finally, as component of the mediator

Table 3 List of nuclear hormone receptor coregulators (NRCs) in humans (*Homo sapiens*) and amphioxus (*Branchiostoma floridae*)

Coregulator class	Coregulator name in humans	Coregulator name in amphioxus
Coactivators	NCOA1, NCOA2, NCOA3	NCOA
	CBP, P300	CBP/P300
	PRGC1/PGC1α, PRGC2/PGC1β, PPRC1/PGC1-related	PGC1
Corepressors	NCOR1, NCOR2/SMRT	NCOR
Mediator complex	MED1/TRAP220	TRAP220

complex, we looked for orthologs of the vertebrate mediator complex subunit 1/thyroid hormone receptor-associated protein complex 220 kDa component (MED1/TRAP220). Again, we identified only a single ortholog in the amphioxus genome, which we called amphioxus TRAP220 (Table 3; Supplementary Fig. S6).

Discussion

In this study, we identified all NRs and some key components of the NR signaling machinery in the invertebrate chordate amphioxus, the best available stand-in for the last common ancestor of all chordates. In comparison to 48 NRs in humans, 17 in sea squirts and 21 in fruit flies, there are 33 NRs in amphioxus (32 NRs were identified in the amphioxus genome and COUP-TF has previously been cloned from amphioxus), which is more than initially expected. For NR signaling coregulators, we found a single amphioxus ortholog for each coregulator family analyzed. By discussing this amphioxus NR and NR coregulator data set we will now infer the NR signaling complement of the ancestral chordate.

Amphioxus as a simple model to study NR signaling evolution in chordates

There are a large number of NR groups, with a single amphioxus NR and two, three, or four human receptors. For these NR groups, amphioxus is a very good system to study the evolution of NR function in chordates, which has already been shown for NRs belonging to three different subfamilies (NR1, NR2, and NR3). In the NR1 subfamily, the thyroid hormone receptor, TR, and the retinoic acid receptor, RAR, have already been studied in amphioxus. As in some vertebrates, most prominently in amphibians, TR controls metamorphosis in amphioxus (Paris et al. 2008a). Whether, like in vertebrates, the amphioxus TR also has other roles in late development, for example in the anterior central nervous system or in the intestine (reviewed in Flamant et al. 2006), remains to be established. Amphioxus RAR is conspicuously expressed in hindbrain and spinal cord of developing amphioxus as well as in the middle third of the embryo (Escriva et al. 2002), which has been hypothesized to closely resemble the expression pattern of RAR in the last common ancestor of all chordates (Escriva et al. 2006). Moreover, it has been shown that RAR is crucial for anteroposterior regionalization of the developing amphioxus embryo, where it acts upstream of *Hox* genes in the patterning of the amphioxus endoderm, central nervous system and general ectoderm (reviewed in Marlétaz et al. 2006; Campo-Paysaa et al. 2008). This role of RAR in anteroposterior patterning of the embryo seems to be largely conserved between amphioxus and vertebrates (reviewed in Marlétaz et al. 2006; Campo-Paysaa et al. 2008).

In the NR2 subfamily, the single amphioxus orthologs of the human retinoid X receptor (RXR), testicular receptor 2/4 (TR2/4) and COUP-TF have already been studied (Langlois et al. 2000; Escriva et al. 2002). In amphioxus, RXR is very broadly expressed in the developing embryo (Escriva et al. 2002), which is not surprising, because it has important functions as heterodimeric binding partner of other NRs in amphioxus and vertebrates (Escriva et al. 2002) and ubiquitous availability of this NR is, thus, a crucial component of NR-dependent signaling (reviewed in Germain et al. 2006b). It will be very interesting to further study the functions of this single RXR in amphioxus, both as a proper transcription factor and as heterodimeric binding partner of other NRs. In vertebrates, as in amphioxus, RXRs have broad or even ubiquitous expression patterns. Moreover, loss-of-function experiments in mice have shown that RXR function is required for development of the vertebrate heart and eyes (reviewed in Germain et al. 2006b). In contrast to RXR, amphioxus TR2/4 expression is detectable only in the anterior amphioxus central nervous system and in the anterior and posterior thirds of the endoderm (Escriva et al. 2002). In vertebrates, TR2/4 is expressed in the developing kidney, gut endoderm and in neuronal precursors (reviewed in Benoit et al. 2006) suggesting that at least some of the functions of TR2/4 in endoderm and central nervous system might be conserved between amphioxus and vertebrates.

Although we were not able to identify a COUP-TF ortholog in the amphioxus genome, this receptor has previously been cloned from amphioxus (Langlois et al. 2000) suggesting that the COUP-TF locus is located in a region of the amphioxus genome that contains gaps and, hence, could not be properly assembled (Putnam et al. 2008). Developmental expression of amphioxus COUP-TF has been assessed. In developing amphioxus, COUP-TF expression is limited to hindbrain and anterior spinal cord of late larvae (Langlois et al. 2000). In contrast to this relatively restricted expression in amphioxus, vertebrate COUP-TFs are expressed in various different tissues during development and are required for central nervous system patterning as well as for the development of the heart and of the circulatory system (reviewed in Benoit et al. 2006). Thus, it is possible that at least some of the functions of COUP-TF in the developing central nervous system might have evolved in the last common ancestor of amphioxus and vertebrates (Langlois et al. 2000).

The NR3 subfamily includes the estrogen receptors (ERs), steroid receptors (SRs) and the estrogen receptor-related receptors (ERRs). All three amphioxus representatives of this subfamily (amphioxus ER, SR, and ERR) have already been studied (Horard et al. 2004; Bardet et al. 2005; Paris et al.

2008b). For example, it has been shown that the amphioxus ER is unable to bind estradiol, the classical ligand of vertebrate ERs, as well as a variety of other known natural and synthetic ligands of vertebrate ERs (Paris et al. 2008b). In vertebrates, ERs are crucial for the development of the reproductive organs and of the brain (reviewed in Dahlman-Wright et al. 2006). In contrast, ERRs in vertebrates are important for early development, which has been shown in both mice and zebrafish (Bardet et al. 2005; reviewed in Benoit et al. 2006). In zebrafish, for example, ERRs control the morphogenetic movements during gastrulation (Bardet et al. 2005). Interestingly, in both zebrafish and amphioxus, ERRs are expressed in a segmented manner in hindbrain and muscular somites suggesting that ERRs might have already been required for the development of these two tissue layers in the last common ancestor of all chordates (Bardet et al. 2005; Bertrand et al. 2007).

For the NR coregulator families we studied, there is always a single amphioxus ortholog for each paralogous group of two, three, or four human coregulator family members. Hence, we identified the unique representatives of key components of the NR signaling machinery in the amphioxus genome including classical NR coactivators (nuclear receptor coactivator, NCOA, CREB-binding protein, CBP, and histone acetyltransferase p300, P300, as well as peroxisome proliferator-activated receptor gamma coactivator 1, PGC1), classical NR corepressors (nuclear receptor corepressor/silencing mediator for retinoid and thyroid hormone receptors, NCOR/SMRT) and an important component of the NR mediator complex (mediator complex subunit 1/thyroid hormone receptor-associated protein complex 220-kDa component, MED1/TRAP220) (reviewed in Hu and Lazar 2000; Bastien and Rochette-Egly 2004; Goodson et al. 2005; Germain et al. 2006c; Lonard and O'Malley 2006). For each of these coregulator families analyzed, there are multiple representatives in vertebrates (reviewed in Hu and Lazar 2000; Goodson et al. 2005; Germain et al. 2006c; Lonard and O'Malley 2006). As mentioned above, in amphioxus we only isolated a single ortholog in each of these coregulator families: there is only one coactivator of the NCOA, CBP/P300, and PGC1 families (amphioxus NCOA, CBP/P300 and PGC1, respectively), one corepressor of the NCOR/SMRT family (amphioxus NCOR) and one MED1/TRAP220 component of the mediator complex (amphioxus TRAP220). This coregulator complement in amphioxus represents a vertebrate-like set with the important difference that the amphioxus NR coactivators, corepressors, and mediator complex components have not been duplicated. It is, hence, very likely that the ancestral chordate had a NR coregulator complement that very closely resembled that of amphioxus, and that in the vertebrate lineage this basic coregulator set has been expanded by the massive whole genome duplications that characterized the evolution of vertebrates (Holland et al. 1994; Holland et al. 2008; Putnam et al. 2008; reviewed in Schubert et al. 2006a). These data also show that the amphioxus NR signaling machinery is vertebrate-like, but simpler. This relative simplicity in amphioxus highlights the usefulness of amphioxus as a model to study the evolution and elaboration of NR signaling in chordates. For example, experimental targeting of a single one of these amphioxus factors will provide information that, in vertebrates, can only be obtained by targeting two, three, or even four paralogous genes.

Gene duplication and loss accompanied the diversification of chordate lineages

In the amphioxus genome, we found 33 NRs, which is more than we initially expected. This elevated number of NRs is mainly due to a duplication in amphioxus of the receptors of the NR1H group. In this NR group, instead of single liver X receptor (LXR) and farnesoid X receptor (FXR) orthologs, as is the case in sea squirts (Dehal et al. 2002; Bertrand et al. 2004), there are ten amphioxus members. These NR1H duplicates most likely arose from an amphioxus-specific duplication. In vertebrates, NR1H receptors control lipid and cholesterol metabolism and, together with NR1I/J/K receptors (the vitamin D receptor, VDR, the pregnane X receptor, PXR, and the constitutive androstene receptor, CAR), are involved in the metabolism of toxins and xenobiotics (reviewed in Moore et al. 2006). The duplication of members of the NR1H group in amphioxus suggests that the signaling networks controlled by these receptors might have also been elaborated in amphioxus. It is, hence, conceivable that this elaboration lead to the evolution of novel mechanisms for degradation of and protection from various kinds of toxins and xenobiotics derived from marine microorganisms and the environment surrounding the sedentary, filter-feeding amphioxus adult. In this context, it is also very interesting to note that there is no NR1I/J/K member in amphioxus, while representatives of this group exist in humans (VDR, PXR, and CAR), sea squirts (NR1K-1 and NR1K-2), and even fruit flies (the hormone receptor-like in 96, HR96). Thus, the amphioxus ortholog of this group is either located in a region of the genome that contains gaps and could not be assembled or has been lost in the amphioxus lineage. In the case of a lineage-specific loss, since members of the NR1I/J/K group are very closely related to the NR1H receptors that have extensively been duplicated in amphioxus, it is possible that in amphioxus one or several of the NR1H duplicates have assumed the functions of the lost NR1I/J/K receptor.

Another example of an amphioxus-specific duplication can be found in the NR2E group. In vertebrates, the two members of this NR group, tailless (TLX) and the photoreceptor-specific nuclear receptor (PNR), have impor-

tant roles in brain and retina development (reviewed in Benoit et al. 2006). In amphioxus, in addition to TLX and PNR, there is an additional NR2E group member. From the phylogenetic analysis, it can be assumed that this receptor, called NR2E, is the result of a lineage-specific duplication event in amphioxus. Since both TLX and PNR function during development of the visual system in vertebrates, it is tempting to suggest that the additional NR2E receptor in amphioxus functions alongside TLX and PNR in the relatively complex visual system of amphioxus with its frontal eye, pineal eye, and multiple photoreceptors (Glardon et al. 1998; reviewed in Lacalli 2001). Finally, it is very interesting to note that there is no NR2E receptor in the sea squirt genome (Dehal et al. 2002; Bertrand et al. 2004) suggesting that both TLX and PNR have been lost in the tunicate lineage, which could be related to the tendency in tunicates towards modification and reduction of the visual system (reviewed in Lacalli 2001). Functional analyses of the three amphioxus NR2E receptors will provide very interesting insights into the evolution of novel (neofunctionalization) and shared (subfunctionalization) roles for gene copies after lineage-specific duplication.

In the recent past, members of the NR3 subfamily (ERs, SRs, and ERRs) have been used repeatedly as molecular models to understand the evolution of ligand-binding capacities within the NR superfamily (Escriva et al. 2000; Baker 2003; Thornton et al. 2003; Bridgham et al. 2006). In humans, there are two ERs, four SRs (the glucocorticoid receptor, GR, the mineralocorticoid receptor, MR, the androgen receptor, AR, the progesterone receptor, PR), and three ERRs. Interestingly, of the three NR3 groups, only ERRs have been identified in the genomes of sea squirts and fruit flies (Bertrand et al. 2004). In amphioxus, there seems to be a single ortholog in each of the three NR3 groups, although the branch support for the grouping of the amphioxus ER and SR with their respective group is not very good. In contrast, the amphioxus ERR groups very robustly with the ERRs from humans, sea squirts, and fruit flies. Sequence comparisons have shown that both the LBD and the DBD of the amphioxus ER and SR are more similar to vertebrate ERs than to vertebrate SRs (Holland et al. 2008; Paris et al. 2008b). In addition, the amphioxus ER is unable to bind estradiol, the classical ligand of vertebrate ERs, as well as a variety of other known natural and synthetic ligands of vertebrate ERs (Paris et al. 2008b). In its lack of ability to bind estradiol, the amphioxus ER is, hence, more similar to ERs of other invertebrates, such as mollusks (Thornton et al. 2003; Paris et al. 2008b). Since the amphioxus SR has not yet been functionally described, it remains to be established, whether the amphioxus SR is capable to bind steroids (for example estrogens) and to activate transcription in a ligand-dependent manner. In vertebrates, such as mice, ERs and SRs are crucial for the development of the reproductive organs (reviewed in Dahlman-Wright et al. 2006; Lu et al. 2006). Moreover, ERs are also required, for example, for neuronal migration in the embryonic mouse brain and SRs, such as GR, are important for lung and liver formation in mice (reviewed in Dahlman-Wright et al. 2006; Lu et al. 2006). To better understand the evolution of ER and SR function in chordates, it will be necessary to carefully assess the functions of ER and SR in developing and adult amphioxus, in particular because ERs and SRs are absent from the sea squirt genome (Dehal et al. 2002; Bertrand et al. 2004) suggesting that these NR3 receptors have secondarily been lost in the tunicate lineage.

Within the NR5 subfamily, there are two human receptors, two from amphioxus and fruit flies and one from sea squirts (Bertrand et al. 2004). The two human members, steroidogenic factor 1 (SF1) and the liver receptor homolog 1 (LRH1), both belong to the NR5A group, which also includes the single sea squirt NR5, fruit fly fushi tarazu-transcription factor 1 (FTZ-F1) and amphioxus NR5A. Interestingly, the second NR5 subfamily member in amphioxus clusters very strongly with the fruit fly hormone receptor-like in 39 (HR39) suggesting that this amphioxus receptor is a member of the NR5B group. We, hence, called this amphioxus receptor NR5B. In fruit flies, HR39 is required for proper female reproductive tract development and function (Allen and Spradling 2008) and, during post-embryonic development, probably also acts early in the ecdysone cascade that triggers metamorphosis (Bonneton et al. 2008; reviewed in King-Jones and Thummel 2005). It is intriguing to speculate that the amphioxus ortholog of fruit fly HR39, NR5B, might also have roles in reproduction and metamorphosis of amphioxus, which would establish a molecular link of these two processes between arthropods and chordates. In contrast to amphioxus, there are no NR5B receptors in humans and sea squirts. It is, thus, likely that a NR5B receptor was present in the last common ancestor of all chordates and that the members of this receptor group have secondarily been lost in the lineage leading to extant tunicates and vertebrates (Bertrand et al. 2004).

In the NR0 "subfamily", there are two representatives in humans, one in amphioxus and none in sea squirts. While NR0 receptors have apparently been lost in sea squirts, both the two human NR0 receptors and the single amphioxus NR0 are members of the NR0B group (Dehal et al. 2002; Bertrand et al. 2004). The two human NR0 receptors, dosage-sensitive sex reversal–adrenal hypoplasia congenita critical region of the X chromosome protein 1 (DAX1) and small heterodimer partner (SHP), lack the C region (the DBD) and are, thus, unable to bind to classical NR DNA response elements (reviewed in Benoit et al. 2006). Based on sequence alignments and on the overall length of the sequence, it is very likely that the amphioxus NR0B is also unable to bind

DNA, although DNA-binding experiments with the amphioxus NR0B will be necessary to validate this hypothesis.

In addition to the amphioxus NRs that group within existing NR subfamilies, we also identified three divergent NRs in the amphioxus genome (called NRa, NRb, and NRc). These three receptors did not branch reliably with any NR subfamily. It is possible that these three receptors initially originated from an amphioxus-specific duplication. After duplication, the coding sequences of the duplicates diverged possibly resulting in neofunctionalization or subfunctionalization. Alternatively, these three NRs could also be remnants of ancient NR subfamilies, whose representatives have secondarily been lost in most species, for which the NR repertoire has already been assessed. The fact that one of the three divergent amphioxus NRs (NRa) robustly groups with a sea urchin receptor (Sp-nr2C) lends support to the latter hypothesis and suggests that with more genomic data and whole genome sequences, it might become feasible to classify these three divergent amphioxus NRs.

Conclusions

In this study, we identified the NR and NR coregulator complement of the invertebrate chordate amphioxus, which is located at the base of the chordates and as such is the best available stand-in for the last common ancestor of all chordates. Our data suggest that the chordate ancestor had 22 NRs, more specifically one NR0 (NR0B), eight NR1s (TR, RAR, PPAR, REV-ERB, ROR, LXR, FXR, NR1I), six NR2s (HNF4, RXR, TR2/4, TLX, PNR, COUP-TF), three NR3s (ER, ERR, SR), one NR4 (NR4A), two NR5s (NR5A, NR5B), and one NR6 (GCNF). In addition, the ancestral chordate had a single copy of the NR coactivators NCOA, CBP/P300, and PGC1, of the NR corepressor NCOR, and of the NR mediator complex component TRAP220. Thus, the last common ancestor of chordates was already characterized by a very elaborate set of NRs and NR coregulators that have independently been modified and expanded both in amphioxus and vertebrates. In sum, this work highlights the utility of the amphioxus genome for understanding the evolution of chordates and also shows that the amphioxus genome has by no means been fixed in an ancestral state: this genome has undergone various lineage-specific gene duplications and losses after the split from the lineage leading to extant tunicates and vertebrates some 550 million years ago.

Acknowledgments The authors would like to thank Michael E. Baker for critical reading of the manuscript. This work was supported by grants from the ANR and the CNRS to M.S., the MENRT to M.S. and V.L. and by CRESCENDO, a European Union Integrated Project of FP6, and by CASCADE, a Network of Excellence of FP6. S.B. is funded by an ARC postdoctoral fellowship.

References

Allen AK, Spradling AC (2008) The *Sf1*-related nuclear hormone receptor *Hr39* regulates *Drosophila* female reproductive tract development and function. Development 135:311–321

Baker ME (2003) Evolution of adrenal and sex steroid action in vertebrates: a ligand-based mechanism for complexity. Bioessays 25:396–400

Bardet PL, Schubert M, Horard B, Holland LZ, Laudet V, Holland ND, Vanacker JM (2005) Expression of estrogen-receptor related receptors in amphioxus and zebrafish: implications for the evolution of posterior brain segmentation at the invertebrate-to-vertebrate transition. Evol Dev 7:223–233

Bastien J, Rochette-Egly C (2004) Nuclear retinoid receptors and the transcription of retinoid-target genes. Gene 328:1–16

Benoit G, Cooney A, Giguere V, Ingraham H, Lazar M, Muscat G, Perlmann T, Renaud JP, Schwabe J, Sladek F, Tsai MJ, Laudet V (2006) International Union of Pharmacology. LXVI. Orphan nuclear receptors. Pharmacol Rev 58:798–836

Bertrand S, Brunet FG, Escriva H, Parmentier G, Laudet V, Robinson-Rechavi M (2004) Evolutionary genomics of nuclear receptors: from twenty-five ancestral genes to derived endocrine systems. Mol Biol Evol 21:1923–1937

Bertrand S, Thisse B, Tavares R, Sachs L, Chaumot A, Bardet PL, Escriva H, Duffraisse M, Marchand O, Safi R, Thisse C, Laudet V (2007) Unexpected novel relational links uncovered by extensive developmental profiling of nuclear receptor expression. PLoS Genet 3:e188

Bledsoe RK, Montana VG, Stanley TB, Delves CJ, Apolito CJ, McKee DD, Consler TG, Parks DJ, Stewart EL, Willson TM, Lambert MH, Moore JT, Pearce KH, Xu HE (2002) Crystal structure of the glucocorticoid receptor ligand binding domain reveals a novel mode of receptor dimerization and coactivator recognition. Cell 110:93–105

Bledsoe RK, Madauss KP, Holt JA, Apolito CJ, Lambert MH, Pearce KH, Stanley TB, Stewart EL, Trump RP, Willson TM, Williams SP (2005) A ligand-mediated hydrogen bond network required for the activation of the mineralocorticoid receptor. J Biol Chem 280:31283–31293

Bonneton F, Chaumot A, Laudet V (2008) Annotation of *Tribolium* nuclear receptors reveals an increase in evolutionary rate of a network controlling the ecdysone cascade. Insect Biochem Mol Biol 38:416–429

Bridgham JT, Carroll SM, Thornton JW (2006) Evolution of hormone-receptor complexity by molecular exploitation. Science 312:97–101

Campo-Paysaa F, Marlétaz F, Laudet V, Schubert M (2008) Retinoic acid signaling in development: tissue-specific functions and evolutionary origins. Genesis, in press

Dahlman-Wright K, Cavailles V, Fuqua SA, Jordan VC, Katzenellenbogen JA, Korach KS, Maggi A, Muramatsu M, Parker MG, Gustafsson JÅ (2006) International Union of Pharmacology. LXIV. Estrogen receptors. Pharmacol Rev 58:773–781

Dehal P, Satou Y, Campbell RK, Chapman J, Degnan B, De Tomaso A, Davidson B, Di Gregorio A, Gelpke M, Goodstein DM, Harafuji N, Hastings KE, Ho I, Hotta K, Huang W, Kawashima T, Lemaire P, Martinez D, Meinertzhagen IA, Necula S, Nonaka M, Putnam N, Rash S, Saiga H, Satake M, Terry A, Yamada L, Wang HG, Awazu S, Azumi K, Boore J, Branno M, Chin-Bow S, DeSantis R, Doyle S, Francino P, Keys DN, Haga S, Hayashi H, Hino K, Imai KS, Inaba K, Kano S, Kobayashi K, Kobayashi M, Lee BI, Makabe KW, Manohar C, Matassi G, Medina M, Mochizuki Y, Mount S, Morishita T, Miura S, Nakayama A, Nishizaka S, Nomoto H, Ohta F, Oishi K, Rigoutsos I, Sano M, Sasaki A, Sasakura Y, Shoguchi E, Shin-i T, Spagnuolo A,

Stainier D, Suzuki MM, Tassy O, Takatori N, Tokuoka M, Yagi K, Yoshizaki F, Wada S, Zhang C, Hyatt PD, Larimer F, Detter C, Doggett N, Glavina T, Hawkins T, Richardson P, Lucas S, Kohara Y, Levine M, Satoh N, Rokhsar DS (2002) The draft genome of *Ciona intestinalis*: insights into chordate and vertebrate origins. Science 298:2157–2167

Dehal P, Boore JL (2005) Two rounds of whole genome duplication in the ancestral vertebrate. PLoS Biol 3:e314

Edgar RC (2004) MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res 32:1792–1797

Escriva H, Safi R, Hänni C, Langlois MC, Saumitou-Laprade P, Stehelin D, Capron A, Pierce R, Laudet V (1997) Ligand binding was acquired during evolution of nuclear receptors. Proc Natl Acad Sci USA 94:6803–6808

Escriva H, Delaunay F, Laudet V (2000) Ligand binding and nuclear receptor evolution. Bioessays 22:717–727

Escriva H, Holland ND, Gronemeyer H, Laudet V, Holland LZ (2002) The retinoic acid signaling pathway regulates anterior/posterior patterning in the nerve cord and pharynx of amphioxus, a chordate lacking neural crest. Development 129:2905–2916

Escriva H, Bertrand S, Germain P, Robinson-Rechavi M, Umbhauer M, Cartry J, Duffraisse M, Holland LZ, Gronemeyer H, Laudet V (2006) Neofunctionalization in vertebrates: the example of retinoic acid receptors. PLoS Genet 2:e102

Flamant F, Baxter JD, Forrest D, Refetoff S, Samuels H, Scanlan TS, Vennström B, Samarut J (2006) International Union of Pharmacology. LIX. The pharmacology and classification of the nuclear receptor superfamily: thyroid hormone receptors. Pharmacol Rev 58:705–711

Galtier N, Gouy M, Gautier C (1996) SEAVIEW and PHYLO_WIN: two graphic tools for sequence alignment and molecular phylogeny. Comput Appl Biosci 12:543–548

Garcia-Fernàndez J, D'Aniello S, Escriva H (2007) Organizing chordates with an organizer. Bioessays 29:619–624

Germain P, Chambon P, Eichele G, Evans GM, Lazar MA, Leid M, de Lera AR, Lotan R, Mangelsdorf DJ, Gronemeyer H (2006a) International Union of Pharmacology. LX. Retinoic acid receptors. Pharmacol Rev 58:712–725

Germain P, Chambon P, Eichele G, Evans GM, Lazar MA, Leid M, de Lera AR, Lotan R, Mangelsdorf DJ, Gronemeyer H (2006b) International Union of Pharmacology. LXIII. Retinoid X receptors. Pharmacol Rev 58:760–772

Germain P, Staels B, Dacquet C, Spedding M, Laudet V (2006c) Overview of nomenclature of nuclear receptors. Pharmacol Rev 58:685–704

Glardon S, Holland LZ, Gehring WJ, Holland ND (1998) Isolation and developmental expression of the amphioxus *Pax-6* gene (*AmphiPax-6*): insights into eye and photoreceptor evolution. Development 125:2701–2710

Goodson M, Jonas BA, Privalsky MA (2005) Corepressors: custom tailoring and alterations while you wait. Nucl Recept Signal 3: e003

Gronemeyer H, Gustafsson JÅ, Laudet V (2004) Principles for modulation of the nuclear receptor superfamily. Nat Rev Drug Discov 3:950–964

Guindon S, Gascuel O (2003) A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. Syst Biol 52:696–704

Holland LZ, Albalat R, Azumi K, Benito-Gutiérrez E, Blow MJ, Bronner-Fraser M, Brunet F, Butts T, Candiani S, Dishaw LJ, Ferrier DEK, Garcia-Fernàndez J, Gibson-Brown JJ, Gissi C, Godzik A, Hallböök F, Hirose D, Hosomichi K, Ikuta T, Inoko H, Kasahara M, Kasamatsu J, Kawashima T, Kimura A, Kobayashi M, Kozmik Z, Kubokawa K, Laudet V, Litman GW, McHardy AC, Meulemans D, Nonaka M, Olinski RP, Pancer Z, Pennacchio LA, Pestarino M, Rast JP, Rigoutsos I, Robinson-Rechavi M, Roch G, Saiga H, Sasakura Y, Satake M, Satou Y, Schubert M, Sherwood N, Shiina T, Takatori N, Tello J, Vopalensky P, Wada S, Xu A, Ye Y, Yoshida K, Yoshizaki F, Yu JK, Zhang Q, Zmasek CM, de Jong PJ, Osoegawa K, Putnam NH, Rokhsar DS, Satoh N, Holland PWH (2008) The amphioxus genome illuminates vertebrate origins and cephalochordate biology. Genome Res 18:1100–1111

Holland PWH, Garcia-Fernàndez J, Williams NA, Sidow A (1994) Gene duplications and the origins of vertebrate development. Development Suppl 1994:125–133

Horard B, Castet A, Bardet PL, Laudet V, Cavailles V, Vanacker JM (2004) Dimerization is required for transactivation by estrogen-receptor-related (ERR) orphan receptors: evidence from amphioxus ERR. J Mol Endocrinol 33:493–509

Howard-Ashby M, Materna SC, Brown CT, Chen L, Cameron RA, Davidson EH (2006) Gene families encoding transcription factors expressed in early development of *Strongylocentrotus purpuratus*. Dev Biol 300:90–107

Hu X, Lazar MA (2000) Transcriptional repression by nuclear hormone receptors. Trends Endocrinol Metab 11:6–10

King-Jones K, Thummel CS (2005) Nuclear receptors—a perspective from *Drosophila*. Nat Rev Genet 6:311–323

Lacalli TC (2001) New perspectives on the evolution of protochordate sensory and locomotory systems, and the origin of brains and heads. Philos Trans R Soc Lond B 356:1565–1572

Langlois MC, Vanacker JM, Holland ND, Escriva H, Queva C, Laudet V, Holland LZ (2000) Amphicoup-TF, a nuclear orphan receptor of the lancelet *Branchiostoma floridae*, is implicated in retinoic acid signalling pathways. Dev Genes Evol 210:471–482

Laudet V, Gronemeyer H (2002) The nuclear receptor facts book. Academic Press, San Diego

Lonard DM, O'Malley BW (2006) The expanding cosmos of nuclear receptor coactivators. Cell 125:411–414

Lu NZ, Wardell SE, Burnstein KL, Defranco D, Fuller PJ, Giguere V, Hochberg RB, McKay L, Renoir JM, Weigel NL, Wilson EM, McDonnell DP, Cidlowski JA (2006) International Union of Pharmacology. LXV. The pharmacology and classification of the nuclear receptor superfamily: glucocorticoid, mineralocorticoid, progesterone, and androgen receptors. Pharmacol Rev 58:782–797

Marlétaz F, Holland LZ, Laudet V, Schubert M (2006) Retinoic acid signaling and the evolution of chordates. Int J Biol Sci 2:38–47

Michalik L, Auwerx J, Berger JP, Chatterjee VK, Glass CK, Gonzalez FJ, Grimaldi PA, Kadowaki T, Lazar MA, O'Rahilly S, Palmer CN, Plutzky J, Reddy JK, Spiegelman BM, Staels B, Wahli W (2006) International Union of Pharmacology. LXI. Peroxisome proliferator-activated receptors. Pharmacol Rev 58:726–741

Moore DD, Kato S, Xie W, Mangelsdorf DJ, Schmidt DR, Xiao R, Kliewer SA (2006) International Union of Pharmacology. LXII. The NR1H and NR1I receptors: constitutive androstane receptor, pregnene X receptor, farnesoid X receptor α, farnesoid X receptor β, liver X receptor α, liver X receptor β, and vitamin D receptor. Pharmacol Rev 58:742–759

Nuclear Receptors Nomenclature Committee (1999) A unified nomenclature system for the nuclear receptor superfamily. Cell 97:161–163

Paris M, Escriva H, Schubert M, Brunet F, Brtko J, Ciesielski F, Roecklin D, Vivat-Hannah V, Jamin EL, Cravedi JP, Scanlan TS, Renaud JP, Holland ND, Laudet V (2008a) Amphioxus postembryonic development reveals the homology of chordate metamorphosis. Curr Biol 18:825–830

Paris M, Pettersson K, Schubert M, Bertrand S, Pongratz I, Escriva H, Laudet V (2008b) An amphioxus orthologue of the estrogen receptor that does not bind estradiol: insights into estrogen receptor evolution. BMC Evol Biol 8:219

Pereira de Jésus-Tran K, Côté PL, Cantin L, Blanchet J, Labrie F, Breton R (2006) Comparison of crystal structures of human androgen receptor ligand-binding domain complexed with various agonists reveals molecular determinants responsible for binding affinity. Protein Sci 15:987–999

Pike AC, Brzozowski AM, Hubbard RE, Bonn T, Thorsell AG, Engström O, Ljunggren J, Gustafsson JÅ, Carlquist M (1999) Structure of the ligand-binding domain of oestrogen receptor beta in the presence of a partial agonist and a full antagonist. EMBO J 18:4608–4618

Putnam NH, Butts T, Ferrier DEK, Furlong RF, Hellsten U, Kawashima T, Robinson-Rechavi M, Shoguchi E, Terry A, Yu JK, Benito-Gutiérrez È, Dubchak I, Garcia-Fernàndez J, Gibson-Brown JJ, Grigoriev IV, Horton AC, de Jong PJ, Jurka J, Kapitonov V, Kohara Y, Kuroki Y, Lindquist E, Lucas S, Osoegawa K, Pennacchio LA, Salamov AA, Satou Y, Sauka-Spengler T, Schmutz J, Shin- IT, Toyoda A, Bronner-Fraser M, Fujiyama A, Holland LZ, Holland PWH, Satoh N, Rokhsar DS (2008) The amphioxus genome and the evolution of the chordate karyotype. Nature 453:1064–1072

Schubert M, Escriva H, Xavier-Neto J, Laudet V (2006a) Amphioxus and tunicates as evolutionary model systems. Trends Ecol Evol 21:269–277

Schubert M, Holland ND, Laudet V, Holland LZ (2006b) A retinoic acid-*Hox* hierarchy controls both anterior/posterior patterning and neuronal specification in the developing central nervous system of the cephalochordate amphioxus. Dev Biol 296:190–202

Sea Urchin Genome Sequencing Consortium (2006) The genome of the sea urchin *Strongylocentrotus purpuratus*. Science 314:941–952

Tanenbaum DM, Wang Y, Williams SP, Sigler PB (1998) Crystallographic comparison of the estrogen and progesterone receptor's ligand binding domains. Proc Natl Acad Sci USA 95:5998–6003

Thornton JW, Need E, Crews D (2003) Resurrecting the ancestral steroid receptor: ancient origin of estrogen signaling. Science 301:1714–1717

Williams SP, Sigler PB (1998) Atomic structure of progesterone complexed with its receptor. Nature 393:392–396

Yagi K, Satou Y, Mazet F, Shimeld SM, Degnan B, Rokhsar D, Levine M, Kohara Y, Satoh N (2003) A genomewide survey of developmentally relevant genes in *Ciona intestinalis*. III. Genes for Fox, ETS, nuclear receptors and NFκB. Dev Genes Evol 213:235–244

Yu JK, Satou Y, Holland ND, Shin-I T, Kohara Y, Satoh N, Bronner-Fraser M, Holland LZ (2007) Axial patterning in cephalochordates and the evolution of the organizer. Nature 445:613–617

Molecular and Cellular Endocrinology xxx (2008) xxx–xxx

Contents lists available at ScienceDirect

Molecular and Cellular Endocrinology

journal homepage: www.elsevier.com/locate/mce

Review

The evolution of the ligand/receptor couple: A long road from comparative endocrinology to comparative genomics

Gabriel V. Markov[a,b], Mathilde Paris[a], Stéphanie Bertrand[c], Vincent Laudet[a,*]

[a] Molecular Zoology Team, Institut de Génomique Fonctionnelle de Lyon, Université de Lyon, Ecole Normale Supérieure de Lyon, Université Lyon 1, CNRS, INRA, Institut Fédératif 128 Biosciences Gerland Lyon Sud, France
[b] USM 501/UMR CNRS 5166-Evolution des Régulations Endocriniennes, Muséum National d'Histoire Naturelle, Paris, France
[c] Departament de Genètica, Facultat de Biologia, Universitat de Barcelona, Av. Diagonal 645, edifici annex, 1a planta, 08028 Barcelona, Spain

ARTICLE INFO

Article history:
Received 11 February 2008
Received in revised form 14 May 2008
Accepted 11 June 2008

Keywords:
Nuclear receptors
Comparative endocrinology
Gene duplication
Phylogeny
Orthology

ABSTRACT

Comparative endocrinology considers the evolution of bioregulatory systems and the anatomical structures and molecules that constitute the neuroendocrine and endocrine systems. One aim of comparative endocrinology is to trace the origins of the main endocrine systems. The understanding of the evolution of the ligand/receptor couple is central to this objective. One classical approach to tackle this question is the characterization of receptors and ligands in various types of non-model organisms using as a starting point the knowledge accumulated on classical models such as mammals (mainly human and mouse) and arthropods (with *Drosophila* among other insects). In this review we discuss the potential caveats associated to this two-by-two comparison between a classical model and non-model organisms. We suggest that the use of an evolutionary approach involving comparisons of several organisms in a coherent framework permits reconstruction of the most probable scenarios. The use of the vast amount of genomic data now available, coupled to functional experiments, offers unprecedented possibilities to trace back the origins of the main ligand/receptor couples.

Contents

Comparative endocrinology, a very active branch of endocrinology, is mainly interested in the origin and diversification of hormonal systems in living organisms. Given the medically oriented knowledge that is a salient feature of modern endocrinology research, one basic focus of comparative endocrinologists is to trace back the origin of the human major endocrine systems and to understand the main events that have prompted the diversification

* Corresponding author at: Molecular Zoology Team, Institut de Génomique Fonctionelle de Lyon, UMR 5242 du CNRS, INRA, IFR128 BioSciences Lyon-Gerland, Université de Lyon, Ecole Normale Supérieure de Lyon, 46 Allée d'Italie, 69364 Lyon Cedex 07, France. Tel.: +33 4 72 72 81 90; fax: +33 4 72 72 80 80.
E-mail address: Vincent.Laudet@ens-lyon.fr (V. Laudet).

0303-7207/$ – see front matter © 2008 Elsevier Ireland Ltd. All rights reserved.
doi:10.1016/j.mce.2008.06.011

of these systems (Norris, 2007). This is not an easy task, since this in fact tackles one of the major questions in comparative sciences, that is the origin of complex systems, here the cell–cell communication systems that are acting at the level of the whole organism. Comparative endocrinology, like any other field of life science with an evolutionary component relies on the comparison itself to infer the existence of a given process, anatomical structure, or molecule, back in time. Indeed, since it is usually impossible to have direct information on ancestor species (with the, still anecdotical, but promising exception of ancient DNA research, see Lalueza-Fox et al., 2007) it is the observation that a given feature is conserved in two distant species that allows to conclude that this feature was indeed present in the common ancestor of these two species. Thus a major aspect of research projects in comparative endocrinology is the characterization of endocrine systems in non-model organisms, using model organisms (mainly human but also mouse or arthropods such as *Drosophila*) as a starting point.

The purpose of this short review is double. First, we will emphasize that despite its apparent conceptual simplicity the comparative approach in endocrinology is paved with methodological difficulties. This clearly suggests that if artefactual inferences are to be avoided the traditional comparative approach should be replaced by multi-disciplinary evolutionary and functional studies. Second, we will detail how the impressive amount of data generated by recent genomic analysis offers unprecedented possibilities to carry out this type of research, thus placing comparative endocrinology in front of a major shift in its methods and approaches. Of course such a short review can only provide rapid glances on this burgeoning field. We will thus illustrate this paper by several examples taken from the recent literature without being exhaustive.

1. The ligand/receptor couple

The evolution of the ligand/receptor couple is a question that attracts considerable debate and theoretical discussion of complex experimental approaches. In fact hormones and receptors are central in the understanding of endocrine systems and, their origin as well as their parallel variation through co-evolution is a major evolutionary question. Indeed, divergence of proteins in different species requires ligand and receptor(s) coevolution to improve binding affinity and/or specificity. Coevolution is thus a ubiquitous process that is responsible for the parallel adaptative evolution of hormone/receptors couples in the broadest sense.

On this aspect of coevolution of ligand/receptor pairs the field is sharply cut into two parts given the chemical nature of the ligand. All the ligands that are peptides or proteins, i.e. that are encoded by genes provide conceptually relatively simple cases of ligand/receptor coevolution, with continuous adaptation across time. Protein–protein interaction in general is, from the coevolutionary point of view, not basically different from the interaction between a given receptor and its ligand (Waddell et al., 2007). In such cases, it is believed, and it has been demonstrated in several specific cases, that the genes encoding the ligand and the receptors are undergoing parallel evolution. One of the first examples of such coevolution is the one of the receptors for LH and FSH, which suggests that indeed the specificity of each ligand/receptor pair is maintained in divergent species (Moyle et al., 1994). More recently the case of prolactin receptors in mammals showed how episodes of adaptative evolution have modified the genes encoding the receptor and the ligand (Li et al., 2005, and references therein). In most mammals the prolactin gene evolved very slowly but this near-stasis was interrupted by bursts of rapid changes during the evolution of several mammals orders such as artiodactyls, primates or rodents. Since prolactin has to bind its receptor to fulfill its function, it was anticipated that the gene encoding the prolactin receptor should be subjected to selective pressure in the same mammals. This has been shown to be the case and the correlation between the evolutionary rates of the ligand and the receptor is effectively indicative of such coevolution. Similar examples including G protein-coupled receptors (GPCRs) such as the receptors for PRXamides (Park and Palczewski, 2005) or the secretin (PACAP and VIP) and their receptors (Cardoso et al., 2007) support this theory of ligand–receptor coevolution. Thus, conceptually, the existence of evolutionary couples is relatively well understood and provides a coherent framework for functional evolution studies (Dean and Thornton, 2007).

This situation contrasts with the second, that of receptors for which the ligand is not a peptide or a protein but rather a small molecule. In such cases the ligand is not a gene product but is derived from a biochemical pathway that starts from an inactive precursor, sometimes derived from an external source such as food, which is transformed into the active molecule (see Simões-Costa et al., 2008 for a recent illustration on retinoic acid metabolism). This is the case for some GPCRs but also for many nuclear receptors (NRs), for which the ligands are the products of complex biochemical pathways. In most cases these pathways contain a rate-limiting step, producing the active compound. This critical step is most often the one that is physiologically regulated. In addition it contains a catabolic part that is responsible for the degradation of the ligand and that is also subjected to precise regulation (see You, 2004; Bélanger et al., 1998, for a review on steroids). In these cases of ligand/receptor pairs a simple coevolution mechanism obviously cannot operate. In the case of NRs, several models such as ligand exploitation (Thornton, 2001) or refinement of ligand-binding specificity by mutations (Escriva et al., 2006) have recently been proposed to explain how changes of specificity can take place during evolution. Nevertheless even if the situation for these ligand/receptor couples is more complex, the existence of coevolution is still possible. Indeed, a recent report on cannabinoid receptors suggests that the evolution of cannabinoid receptors is correlated with the evolution of diacylglycerol lipase, an enzyme implicated in the metabolism of anandamide and 2-arachidonyl glycerol (2-AG) the two endogenous ligands of cannabinoid receptors (McPartland et al., 2007). It remains to be investigated if such example can be generalized to other receptor systems.

It is clear nevertheless that any discussion on the evolution of a ligand/receptor couple should be confronted with experimental data, which consists in the characterization of ligands and receptors in various species, distantly related to the most classical models. It has also to be emphasized that receptors and ligands structures are often more conserved than their physiological function which depends on the expression patterns of receptors. This often blurs the recognition of orthology between related ligand/receptor pairs.

2. The phylogenetic framework: metazoan evolution

Any analysis of the evolution of a given ligand/receptor couple should take into account the phylogeny of the organisms from which these various pairs are coming from. Therefore, it is important at that level to rapidly present here the framework on which these comparative approaches are developed. Since most of these studies are done at the scale of animals we will limit this rapid presentation to the case of metazoans.

Fig. 1 depicts a simplified version of the currently endorsed evolutionary tree for metazoans. It is striking that even if the distribution of model organisms is very much biased towards two of the main clades of metazoans (namely ecdysozoans with *Drosophila* and *Caenorhabditis elegans* and deuterostomes with human, mouse,

Please cite this article in press as: Markov, G.V., et al., The evolution of the ligand/receptor couple: A long road from comparative endocrinology to comparative genomics. Mol. Cell. Endocrinol. (2008), doi:10.1016/j.mce.2008.06.011

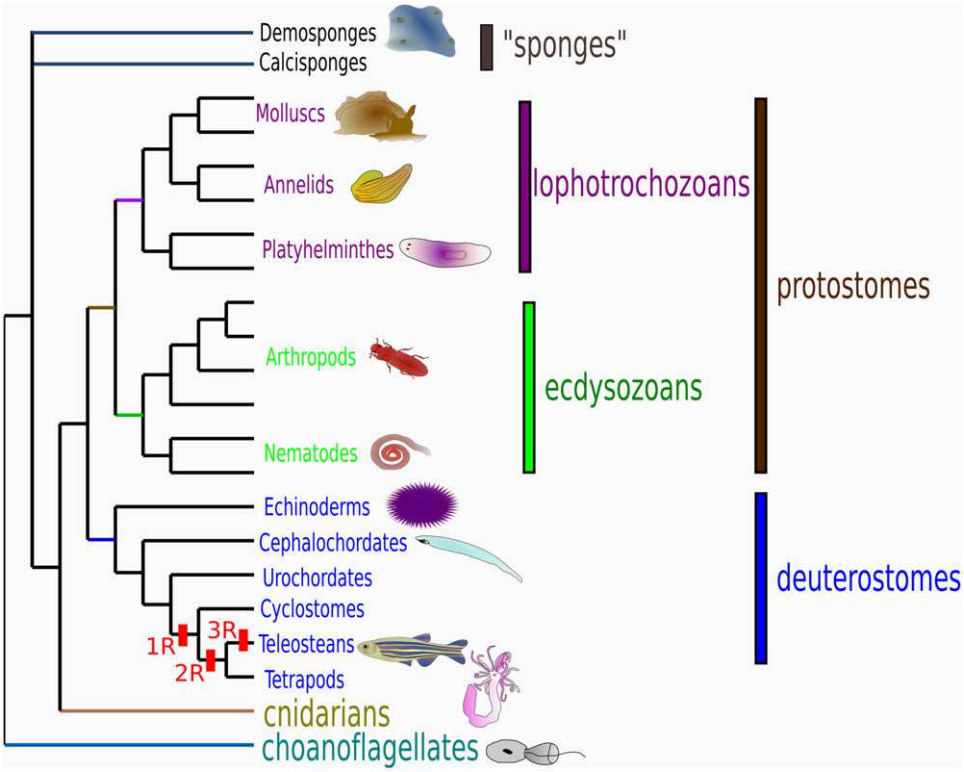


Fig. 1. An updated view on metazoan phylogeny. 1R, 2R, and 3R indicate the whole genome duplication events that occurred in the vertebrate lineage.

chicken, xenopus and zebrafish, among others) genomic data are now available for all the main metazoan clades. For example, looking at cnidarians and sponges, at the base of the metazoan tree, the genome of the sea anemone *Nematostella vectensis* (Putnam et al., 2007) is now available whereas the one of the sponge – *Amphimedon queenslandica* – is available as trace archives. These genomes are bringing crucial information that now allow one to make inferences on early metazoan gene families (Simionato et al., 2007).

From the phylogenetic tree depicted in Fig. 1, the bilaterians, that are the metazoans with three embryonic layers (ectoderm, endoderm and mesoderm) and with a clear antero-posterior axis, are divided into three main clades: on one hand, the lophotrochozoans and the ecdysozoans that together form the protostomes and, on the other hand, the deuterostomes. The monophyly of these three clades is relatively well accepted as is their branching order even if some analyses based on large genomic data sets tend to propose alternative schemes (for example the Coelomata hypothesis that groups arthropods and deuterostomes, excluding nematodes; see for example Rogozin et al., 2007).

The topology of the metaozan tree may have a major impact on our inferences regarding the origin and diversification of endocrine genes. This influence is well illustrated by a recent analysis of NR genes distribution and phylogeny in metazoans in which are compared the scenarios implied by the Ecdysozoan or Coelomata hypothesis on the evolutionary history of nuclear receptors (Bertrand et al., 2004). The use of alternative topologies of the metazoan tree (as the Coelomata hypothesis used for the analysis of the Forkhead family; Carlsson and Mahlapuu, 2002) may affect the conclusions drawn relative to the ancestry and evolution of specific genes. At the present stage of our phylogenetic knowledge, this should be regarded with caution.

Another important observation, developed below, is that the number of genes present in the genomes of several classical model organisms tends to be extremely variable because of large scale events such as whole genome duplication, lineage specific expansion of specific genes or, alternatively gene loss (see Panopoulou and Poustka, 2005 for a review on vertebrates). It is now widely recognized that extremely important models such as *Drosophila*, *C. elegans*, or the urochordate *Ciona intestinalis* have experienced extensive gene loss (see Bertrand et al., 2004 for references). This may have important implications in terms of endocrine gene evolution. For example, the estrogen receptor has been found in vertebrates but not in invertebrate chordates such as *Ciona*, *Drosophila* or *C. elegans*. This has led to the proposal that this receptor was a key innovation of vertebrates (Laudet, 1997; Escriva et al., 2000). In fact the observation that an estrogen receptor orthologous gene is present in several mollusks (Thornton et al., 2003; Keay et al., 2006) as well as in cephalochordates (Paris et al., 2008a) shows that it is in fact much more ancient than expected and that the gene was lost independently in ecdysozoans and urochordates (Bertrand et al., 2004; Escriva et al., 2004). A very similar situation was found for the thyroid hormone receptor (Bertrand et al., 2004; Wu et al., 2006, 2007). It is now widely accepted that no conclusion can be reached on the presence or absence of a given gene in the common ancestor of all bilaterians if data from the three main lineages of metazoans (that is lophotrochozoans, ecdysozoans and deuterostomes) are not available. Given that deuterostomes plus insects and nematodes contain the most dominant model organisms this conclusion should be strongly re-emphasized: no safe conclusion can be drawn on the ancestry of a given gene family without data from lophotrochozoans and/or cnidarians.

3. Two by two comparisons of receptors, proceed at your own risk!

A classical approach in comparative endocrinology is to identify in various organisms, for example the cephalochordate amphioxus, an orthologue of a given receptor known in a classical model organ-

Please cite this article in press as: Markov, G.V., et al., The evolution of the ligand/receptor couple: A long road from comparative endocrinology to comparative genomics. Mol. Cell. Endocrinol. (2008), doi:10.1016/j.mce.2008.06.011

isms like human. This type of analysis may be important to clearly show if this receptor, and by extension, the corresponding signaling pathway is effectively present in the organism of interest. We will highlight below that, even if of course the conclusions reached from this type of analysis are interesting, this two-by-two comparative approach has several caveats that renders it quite risky. Thus, the conclusions and evolutionary models reached by these traditional comparative approaches, based on the assumption that one can extend to their zoological groups the knowledge accumulated in human or *Drosophila*, should be put into a larger perspective. It is only through their independent confirmation based on large-scale evolutionary analysis that these conclusions will be firmly assessed.

3.1. Gene duplication

An important result generated over the last 10 years of genomic analysis is that genes, and even genomes, duplications once believed to be relatively rare events are in fact a major evolutionary mechanism that has been instrumental in shaping the current biodiversity (Ohno, 1970, and see also Volff, 2005 for a recent review in teleosts). Gene duplication is an important mechanism of gene diversification that can be observed in nearly all organisms. Of course tandem duplication of individual genes, the simplest case of gene duplication, occurs quite often. But two more global processes that have broad implications for the functional anatomy of genomes should be emphasized. The first is the whole genome duplication, the importance of which was first highlighted in the 1970s by Susumu Ohno (Ohno, 1970) and later revealed by the study of invertebrate chordates and early vertebrates (Garcia-Fernández and Holland, 1994) as well as fishes (Wittbrodt et al., 1998). It is now well established that in several cases, such as at the base of the vertebrate tree (Dehal and Boore, 2005) or early on during actinopterygian fish evolution (Jaillon et al., 2004) whole genome duplication took place and impacted strongly on the appearance and diversification of complex features, including the endocrine systems (Holland et al., 2008). The fate of duplicated genes (non-, sub- or neo-functionalization, see Force et al., 1999 that is, respectively loss of one copy, sharing of the ancestral function between the two duplicates or acquisition of a new function by one of the copy), and its link with the origin of evolutionary novelties, is currently a major and fruitful research area and the study of endocrine genes should be placed in this context. A very convincing example of such an analysis in the context of actinopterygian fish whole genome duplication has been recently published for the MyoD family (Macqueen and Johnston, 2008).

The other extreme case of gene duplication is lineage-specific duplication events that can sometimes be extensive. This is the mechanism at the origin of the 270 nuclear receptor genes that are present in the genome of the nematode *C. elegans* (Sluder and Maina, 2001). It has been shown that most of these genes correspond in fact to orthologues of a unique gene encoding the orphan receptor HNF-4, which was massively duplicated in *C. elegans* and related species (Robinson-Rechavi et al., 2005). The functional significance of this burst of duplication is still under investigation, but it is clear that such an event should have consequences on the physiology of the animal and may completely modify the evolutionary scenarios constructed at the level of an individual gene family. GPCRs (Cardoso et al., 2006), but also receptors with tyrosine kinase activities (Rikke et al., 2000), the heterotrimeric G protein a-subunit (O'Halloran et al., 2006) and the *Hedgehog*-related genes (*Hog*) (Aspöck et al., 1999) also provide cases of lineage-specific expansion in nematodes.

The presence of these duplication events has a major consequence for the comparative endocrinologist, as it reinforces the

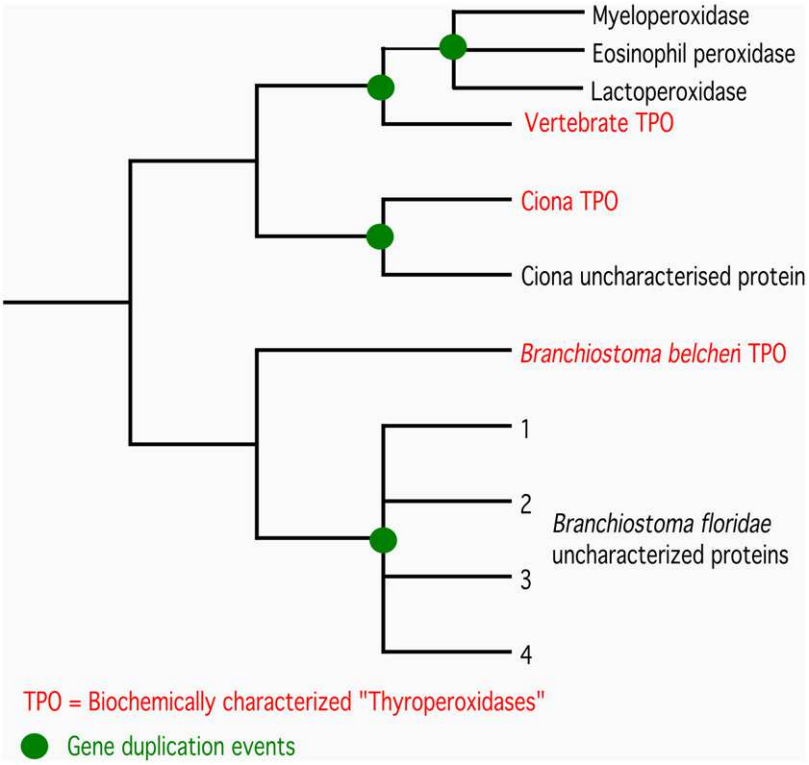


Fig. 2. Phylogenetic relationships in the chordate thyroperoxidase family. Duplication events are indicated by green spots. Note that myeloperoxidase, lactoperoxidase and eosinophil peroxidases are the mammalian members of a gene family where duplications occurred at various levels in the vertebrate lineage. General tree topology is based on Heyland, 2006, and completed with data from *Ciona* complete genome (http://www.treefam.org/cgi-bin/TFinfo.pl?ac=TF314416) and from Holland et al., 2008, for *Branchiostoma floridae* sequences. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of the article.)

Please cite this article in press as: Markov, G.V., et al., The evolution of the ligand/receptor couple: A long road from comparative endocrinology to comparative genomics. Mol. Cell. Endocrinol. (2008), doi:10.1016/j.mce.2008.06.011

need for an accurate phylogenetic study of any protein of interest in order to avoid simplified, and often erroneous interpretations. Fig. 2 shows the example of the thyroperoxidase (TPO) gene, which is implicated in the synthesis of vertebrate thyroid hormones. The vertebrate TPO is a member of a multigenic family, being the orthologue of a large group of peroxidases that experienced a complex series of duplication events within the vertebrate lineages. In mammals we find not only thyroid peroxidase but also lactoperoxidase, myelo-peroxidase and eosinophil peroxidase, all these genes coming from a unique ancestor gene that underwent duplications at various levels within the vertebrate lineages (Heyland et al., 2006). Using primers corresponding to a various set of metazoan peroxidases, a gene encoding a peroxidase was cloned in the urochordates *C. intestinalis* and *Halocynthia roretzi*. Since in vertebrates four closely related types of peroxidases are found, one cannot conclude that this *Ciona* gene is indeed a TPO (see our recent review on the importance of gene nomenclature, Markov et al., 2008). It is only with additional data, here through the analysis of the expression pattern and the comparison to the previously reported biochemical activity that the conclusion that the gene encodes a TPO can be reached. Interestingly, a more extensive genomic analysis (Fig. 2) shows that this protein is the product of only one of the two duplicated genes in the *Ciona* genome. The original "*Ciona* TPO" was shown to be expressed in a domain that does not overlap that of its classical regulator TTF1 in vertebrates, and that its expression domain was restricted to the endostyle zone 7, whereas previous histochemical studies reported a peroxidase activity also in other parts of the endostyle (zones 8 and 9) (Ogasawara et al., 1999). The presence of the second gene suggests that a more complete study is needed in order to test if it too has also a TPO-activity, if it is expressed in zones 8 and 9 of the endostyle and if the two genes have overlapping or complementary expression patterns and thus to reconstruct the detailed history of these genes. Since the duplication event that gave rise to the two *Ciona* genes is independent of the duplications that occurred in vertebrates, only functional analysis can determine whether the *Ciona* genes have activities related to thyroid, myelo-, lacto- or eosinophile peroxidases. The situation is even more complex in cephalochordates. Another gene encoding a TPO was characterised in the Chinese amphioxus *Branchiostoma belcheri* (Ogasawara, 2000) but the analysis of the *Branchiostoma floridae* genomes (Holland et al., 2008) shows that there are four orthologues of this gene, once again arising from an independent series of duplications from the ones that occurred in *Ciona* or vertebrates. Thus, starting from a unique gene with an unknown activity, the peroxidase gene family has been independently elaborated three times in vertebrates, urochordates and cephalochordates. Since proteins that undergo duplication are prone to subfunctionalisation or neofunctionalisation events (Force et al., 1999), this should be taken into account when comparing the functions of two proteins in two different organisms, because this can have major effects at the physiological level. This example illustrates that one should take into account the full set of genes and their complex history in order to infer their ancestral functions.

3.2. Gene loss

Gene loss is an often neglected aspect that, due to the current interest in whole genome sequences, has recently been shown to be a frequent and important evolutionary mechanism that contributed significantly to the emergence of divergent animal lineages (Danchin et al., 2006). The analysis of the presence of NRs genes in complete genome sequences of metazoans shows that the NR complement of different animal models is extremely variable and that gene loss was effectively frequent, as discussed above in the case of *Drosophila*, nematodes and urochordates (Bertrand et al., 2004). The case of nematodes is particularly puzzling since the events of gene loss are hidden by the massive lineage specific expansion of the HNF4 gene discussed above (Robinson-Rechavi et al., 2005) thus illustrating the complexity of the individual gene family history (Fig. 3).

But gene loss is not only revealed by comparison of distant evolutionary organisms. One example, taken from the evolution of NRs illustrates the importance of taking gene loss into account. When we compared, through careful phylogenetical analysis, the NR complement present in mammals and teleost fishes we were surprised to find cases that could not be explained by the classical "more genes in fish" scenario (Crollius and Weissenbach, 2005; Bertrand et al., 2004). This is well exemplified by the case of the Rev-erb genes, orphan nuclear receptors. In all known mammals, two paralogous Rev-erb genes, called Rev-erbα and Rev-erbβ, corresponding to a unique orthologue in *Drosophila* (called E75) and amphioxus are known (Laudet, 1997). Given the actinopterygian fish whole genome duplication event discussed above, we were expecting to find a maximum of four Rev-erb genes in zebrafish, namely two Rev-erbα and two Rev-erbβ. During our systematic analysis of NR genes expression patterns in zebrafish we were thus surprised to find in fact five Rev-erb genes in zebrafish whereas four are effectively present in pufferfishes (tetraodon and fugu) (Bertrand et al., 2007).

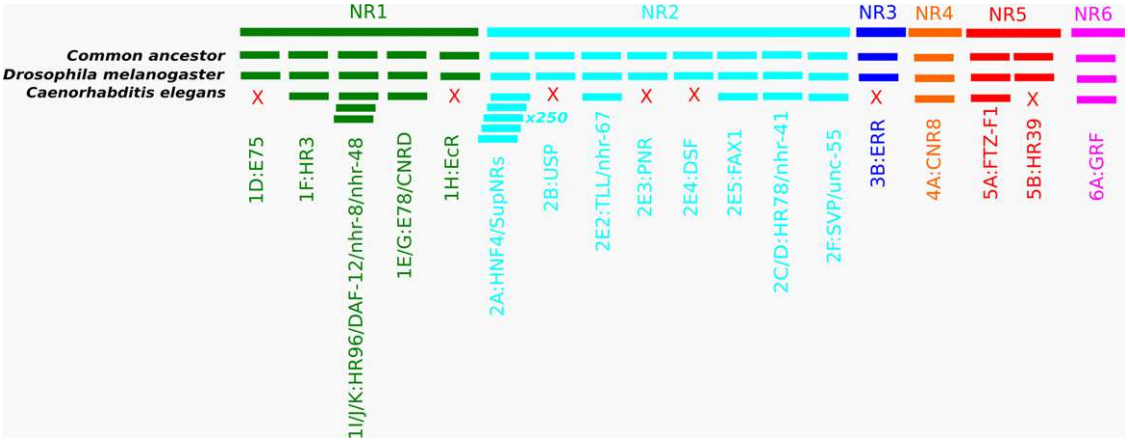


Fig. 3. Gene duplications and losses in *Caenorhabditis elegans*: the example of the nuclear receptor superfamily. The figure shows the loss of seven genes and duplications of two of them in *C. elegans*, in comparison with the *Drosophila* gene set. Data are from Bertrand et al. (2004). Note that the gene set of the "common ancestor" presented here is minimal, and coherent with the comparison of only these two species. A more complete analysis would show that some other genes were also lost independently in *Drosophila* and *C. elegans*.

Please cite this article in press as: Markov, G.V., et al., The evolution of the ligand/receptor couple: A long road from comparative endocrinology to comparative genomics. Mol. Cell. Endocrinol. (2008), doi:10.1016/j.mce.2008.06.011

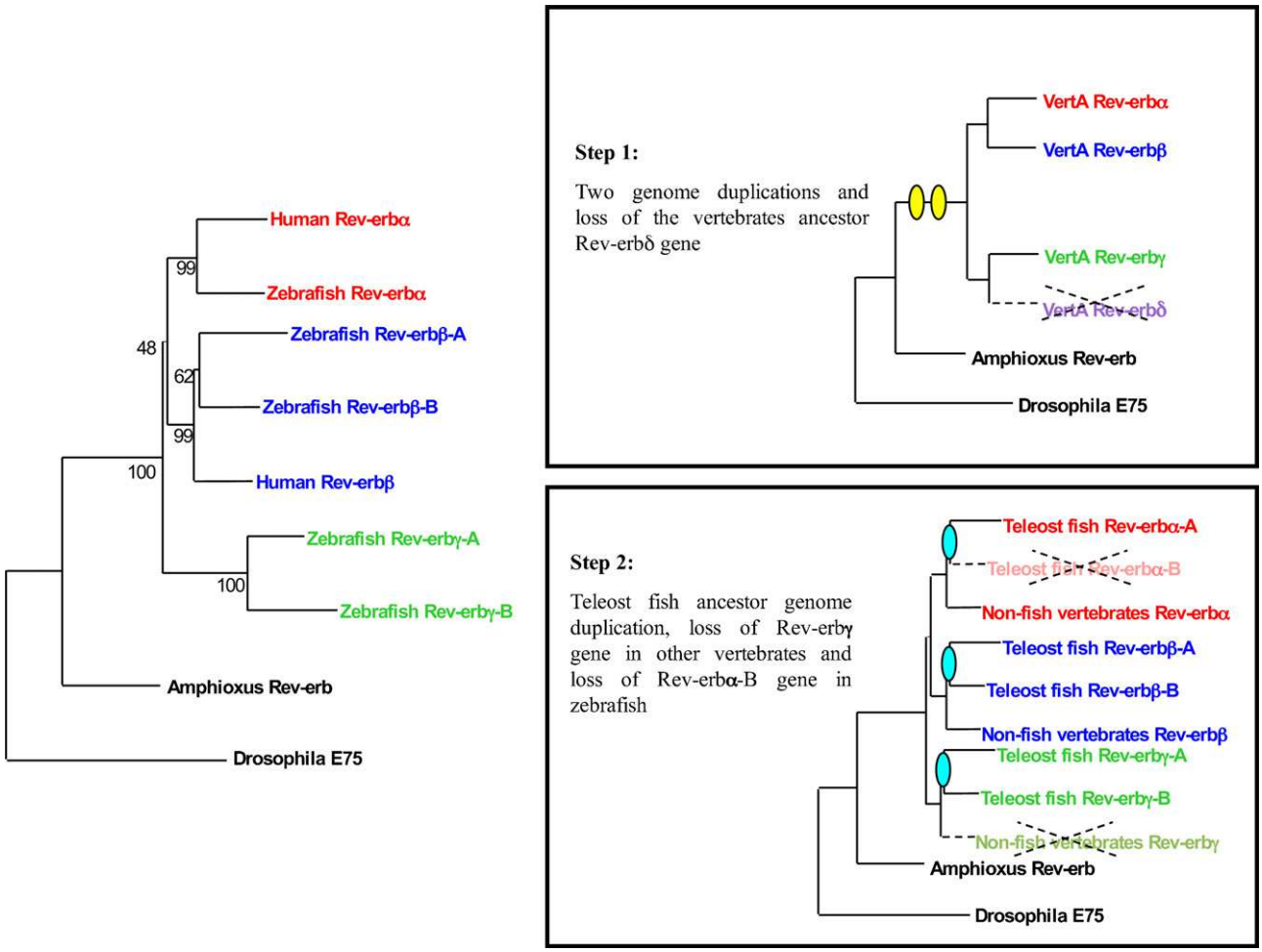


Fig. 4. Phylogenetical tree of the Rev-erb subfamily in zebrafish and human and the most probable evolutionary scenario. Branch-lengths are arbitrary. The two whole genome duplications that took place at the origin of vertebrates and the whole genome duplication at the origin of teleost fish are schematised by yellow and light-blue spots, respectively. VertA means vertebrate ancestor. The most parsimonious scenario explaining the actual evolutionary relationships between human and zebrafish Rev-erb genes can be separated in two steps. The first one corresponds to the loss of one Rev-erb paralogue (here named Rev-erbδ) in the ancestor of vertebrates after the two whole genome duplications. The second step is the loss of Rev-erbγ paralogue in the ancestor of the vertebrates that diverged after fish divergence, and the loss of one Rev-erbα duplicate (here named Rev-erbα-B) in zebrafish after the whole genome duplication that took place at the base of teleost fish lineage. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of the article.)

By phylogenetical analysis (Fig. 4) we were able to deduce that in fact three paralogous Rev-erb genes (Rev-erbα, β and a third gene called γ) were present in the ancestor of vertebrates. In tetrapods this third gene, Rev-erbγ was lost. In fishes, the three ancestral Rev-erb genes were duplicated giving rise to six genes. The differential loss of these fish-specific paralogues explains the different numbers of Rev-erb genes in various fish species: zebrafish lost one Rev-erbα copy whereas pufferfish lost the two Rev-erbα genes. The analysis of the expression patterns strongly suggests that if these genes share a similar function each of them has specific implications in distinct processes, probably linked to circadian rhythm regulation (Bertrand et al., 2007; Kakizawa et al., 2007). This case is certainly not an isolated one and taking only the example of NR genes several other cases of specific gene loss in mammals were found (e.g. ERRs, COUP-TFs, SF-1) whereas, in contrast, one case of fish specific loss (CAR) was found (Bertrand et al., 2007). It is now known that these gene losses occurred in other gene families in mammals but also in other zoological groups (Danchin et al., 2006; Wyder et al., 2007).

All these examples illustrate that the evolutionary history of any given gene cannot be reconstructed without a careful phylogenetic analysis based on the use of adequate methods and that the importance of mechanisms such as gene duplication and gene loss should be adequately tested before concluding on the presence or absence of a given signaling pathway. Classical comparative work based on the two-by-two comparison of human and another animal will never match the quality of information reached by an in depth phylogenetical analysis.

3.3. Evolutionary shifts

Of course genes are not only duplicated or lost. Many genes are conserved as unambiguous orthologues in a wide variety of organisms, but even in such cases one should be careful when studying gene evolution since evolutionary shifts can have dramatic impacts in terms of gene function, as well as at the level of phylogeny. Long branch attraction is a well-known artefact in molecular phylogeny that can artificially group rapidly evolving clades together in a position clearly not compatible with the known phylogeny of species (Delsuc et al., 2005).

RXR evolution in insects provides an illustration of this phenomenon (Fig. 5). The known orthologue of RXR in *Drosophila* is the USP gene that plays an important role as a common heterodimeric partner for several NRs, including the ecdysone receptor, EcR. When we did phylogenetic studies of USP in several insects we saw a

Please cite this article in press as: Markov, G.V., et al., The evolution of the ligand/receptor couple: A long road from comparative endocrinology to comparative genomics. Mol. Cell. Endocrinol. (2008), doi:10.1016/j.mce.2008.06.011

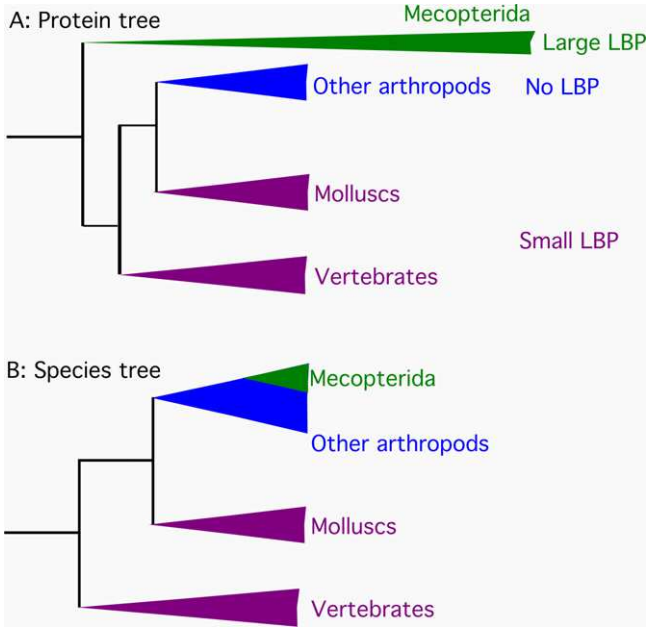


Fig. 5. Long branches attraction and functional shifts: the example of USP-RXR. Adapted from Bonneton et al. (2003) and Iwema et al. (2007). (A) USP-RXR protein tree, with an abnormal position of Mecopterida at the base of the bilaterians. Further analysis showed that this long branch, corresponding to an acceleration of the evolutionary rate in Mecopterida USP, correlates with changes in the ligand-binding abilities. (B) Species tree based on classical neutral markers such as ribosomal RNA genes), showing the real position of Mecopterida within arthropods. The three different colours refer to the different binding abilities of the ligand-binding pocket (LBP). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of the article.)

very striking tree topology since all the USPs found in Diptera and Lepidoptera were not clustered with other arthropod sequences as expected, but at the base of all bilaterian RXRs (Bonneton et al., 2003). At the level of sequence identity this example is quite spectacular since, in the ligand-binding domain the RXR from a beetle (e.g. *Tribolium*) is more closely related to the human RXR than to the *Drosophila* USP. We found that in fact this corresponds to a very strong acceleration of evolutionary rate that occurred during the evolution of mecopterida, a monophyletic group with approximately 25% of extant insect species, including Trichoptera (caddisflies), Mecoptera (scorpionflies) and Siphonaptera (fleas) in addition to Diptera and Lepidoptera (Bonneton et al., 2006). Careful structural and functional analysis of RXR from *Tribolium* and *Drosophila* allowed us to reconstruct a complex and dynamic evolutionary history shaped by several functional shifts during which several events of loss and gain of ligand-binding occurred (Iwema et al., 2007). Interestingly, this event is not restricted to RXR since several similar shifts were also recently observed in other nuclear receptors, all acting in the ecdysone cascade (Bonneton et al., 2008). This example illustrates that such evolutionary shifts are not just complicating factors for phylogenetic analysis, they also provide very fruitful cases in which the evolution of ligand/receptor couples can be scrutinized. In this case, this reveals an unsuspected evolutionary flexibility of NRs in terms of ligand binding.

The literature now contains several cases of evolutionary shifts that all provide illuminating examples of the power of natural selection acting at the level of endocrine pathways. Among these, the nuclear receptors provide two recent examples of receptors, ERs and TRs, that were believed to be recent but for which orthologues were recently found in non-model protostomes. Interestingly, in both cases these protostome receptors do not bind the *bonafide* vertebrate ligand suggesting a more complex history than anticipated. Indeed the ligand binding evolution of estrogen receptor is still far from understood (see contrasting views in Baker, 2003; Thornton et al., 2003; Paris et al., 2008b) and similar questions are suggested by the recent characterization of thyroid hormone receptors in platyhelminthes, molluscs or even in a deuterostome, the cephalochordate amphioxus (Wu et al., 2007; Paris et al., 2008a). At a much smaller taxonomic level the careful analysis of the ligand-binding specificity of PXR in mammals also provides an example of shifts in the ligand-binding ability of a nuclear receptor (Krasowski et al., 2005; Reschly et al., 2007). For other receptor systems, the interested reader could, among a long potential list, refer to the cases of the Growth Hormone receptor in primates (Liu et al., 2001; Wallis, 2001), the TSH receptor which exhibits striking functional differences between mammals and actinopterygian fishes, but also among mammals (Farid and Szkudlinski, 2004) or the genes of the ectodysplasin pathway in vertebrates (Pantalacci et al., 2008).

3.4. Recombination events

Genomic data now available greatly facilitates assessment of gene orthology. In some multigenic families, the duplication rate is so high and variable for different genes that even a cautious phylogenetic analysis is not sufficient to establish orthology relationships. Such a problem occurs when examining the vertebrate CYP2 family that contains genes encoding xenobiotic metabolising enzymes, which have undergone many duplications events at various taxonomic levels (e.g. mammal-specific, rodent-specific, and mouse-specific duplications, with primate-specific or/and rat-specific duplications in parallel) (Nelson, 2005; Thomas, 2007). In such cases, data on the chromosomic localisation can provide very useful information since the analysis of neighboring genes allows one to assess the orthology of the whole syntenic region and, by extension, of the genes of interest. This was indeed the case of the CYP2 genes since some of the genes are located in genomic clusters, where it is believed that recombination events occur quite often, giving rise to the wide diversity of CYP2 genes (Thomas, 2007). Such events will erase the phylogenetic signal and make the functional comparisons between proteins that underwent different duplication events meaningless.

Another example of careful analyses of syntenic regions that led to large scale conclusions on complex evolutionary scenario are the evolutionary analyses of the major histocompatibility cluster (MHC) by Pontarotti's group (Danchin et al., 2003). In the context of fish genome duplication the example of MyoD gene family evolution also illustrates how the use of synteny could help to infer orthology (Macqueen and Johnston, 2008). This type of analysis is particularly useful when studying genes encoding short peptide hormones, which are often difficult to unambiguously identify at a large evolutionary scale. The recent analyses of the POMC/AGPR/MCR gene repertoire in fugu (Klovins et al., 2004) and the study of the origin and coevolution between NPY receptor and prolactin-releasing hormone-receptors in vertebrates (Lagerström et al., 2005), also show the utility of complementary use of data from phylogenetic analysis coupled with careful synteny analysis. The construction of the "gene rosace" diagram that visualizes the respective relationships between genes located in different regions facilitates this type of analysis (see Jaillon et al., 2004, for an explanatory illustration). As for gene analyses, such studies take profit of in depth phylogenomic knowledge about the analysed genes, and also are greatly improved by comparisons between more than two genomes (for detailed examples see Kasahara et al., 2007; Muffato and Crollius, 2008).

When possible, phylogenetic approach should integrate in a known phylogenetic species and gene trees all the information possibly available at the leave of the trees. Information are of course the

Please cite this article in press as: Markov, G.V., et al., The evolution of the ligand/receptor couple: A long road from comparative endocrinology to comparative genomics. Mol. Cell. Endocrinol. (2008), doi:10.1016/j.mce.2008.06.011

sequences but also functional information such as binding properties and other physiological data. In some cases it is possible to calculate the probability to have this information at the node of the tree and from the information available at the node propagate dawn it to the non-annotated leaves (see Thornton, 2004, for a detailed review). But even the finest probabilistic approaches are useless if the sampling is inappropriate.

4. The elusive ligands

The previous examples may give the impression that working on the evolution of receptor genes is full of obstacles that complicate analysis and interpretation. Thus, one solution to study evolution of a specific endocrine signaling pathway could be to test directly if the ligand of interest is present in a wide variety of organisms. This attractive solution is nevertheless also paved with obstacles that are not easily solved and for which a comparative genomics approach does not always provide clues. The two main lines of evidence for the existence of a given ligand in a given animal are (i) the detection of an effect of this ligand in a given physiological or developmental process, most often considered from the known effect of this ligand in mammals and (ii) the direct detection of the ligand through genomic or analytical chemistry analysis.

4.1. Effects

The rationale behind these studies, which are commonly done in traditional comparative endocrinology is that if a given molecule has an effect in a given species this should be taken as a clear indication that there is a receptor for this molecule and thus a signaling pathway and a "physiological role". Somehow mirroring the original receptor definition by Paul Ehrlich, for which a receptor was defined as any molecule that is able to bind exogenous elements in human cells (reviewed in Prüll, 2003), people tend to consider that any molecule of human origin acting on a living system is effectively proof that this molecule binds to another one, which can be called a "receptor". But even if this assumption is true, it cannot be used as an indication that this "receptor" is related in anyway to the human receptor for this molecule. The problem is that many authors have applied this concept without taking into account this important caveat.

A striking example is the case of the arboreal mycorrhizes fungi of the order Glomerales, such as *Glomus intraradices*, that live in symbiosis with many trees and play crucial roles at the root/soil interface. It has been shown that the roots of the plant have a positive effect in stimulating the growth of the fungal hyphe during pre-symbiosis steps (see Requena et al., 2007, for a review). The precise mechanism how plant signals are perceived by the fungi is still unknown but it has been shown that flavonoids have a strong chimio-tactic effect and that this effect can be mimicked by estrogens and blocked by antiestrogens (Poulin et al., 1997; Scervino et al., 2005). This was taken as an indication that the genome of these fungi contained an estrogen receptor and that a specific receptor/interacting protein with a binding site for flavonoid or structurally related compounds (estrogens and antiestrogens) exists in these fungi (Requena et al., 2007; Poulin et al., 1997; Catford et al., 2006). However, our current knowledge of NR evolution indicates that the genome of *Glomus* (which is currently being sequenced) should be devoid of NR genes (Escriva et al., 1997). Among the possible candidates for mediating the effect of estrogens in *Glomus*, one of the most probable would be an orthologue of the estrogen binding protein already found in other fungi, including *Candida albicans* (Madani et al., 1994; Cheng et al., 2006). This protein, well known under the name of "Old yellow enzyme" is one of the most ancient enzymatic systems characterized, and is a cytoplasmic oxydoreductase whose activity can be regulated by estrogen binding (Cheng et al., 2006 and references therein). This example demonstrates that whereas there is an effect of estrogen in this species, this could not be taken as an indication that the endogenous receptor is in any way evolutionary related to the nuclear estrogen receptor found in human.

RXRs provide here again an example of the difficulties associated with the identification of endogenous ligands. These receptors were first described in the early 1990s as orphan receptors whose activity could be modulated by all-*trans*-retinoic acid (see Laudet and Gronemeyer, 2005, for references). It was later shown that 9-*cis*-retinoic acid, an isomeric derivative of all-*trans*-retinoic acid, is able to bind with a high affinity to RXR. Indeed several rexinoids, specific ligands of RXRs, were developed as pharmacological compounds for treatment of diseases such as diabetes and insulin resistance (Pinaire and Reifel-Miller, 2007). Ironically the detection of 9-*cis*-retinoic acid *in vivo* has proven to be difficult, casting doubts on the *in vivo* relevance of this ligand (see Mic et al., 2003). This has been recently confirmed by genetic evidence in the mouse (Calléja et al., 2006). It is now thought that RXR acts rather as a fatty acid or fatty acid derivative sensor, although the exact identity and relevance of its ligand(s) remains a matter of debate (see references in Iwema et al., 2007). This example depicts that even in human, it is not because a molecule has an effect that it indicates its endogenous existence in a physiologically relevant manner!

These examples may appear artificial, but they illustrate a type of assumption that has been commonly used in some comparative endocrinology studies. The estrogen receptor of several molluscs such as *Aplysia*, *Octopus*, and others have been cloned and convincingly shown to be unable to bind to estrogens (Thornton et al., 2003; Keay et al., 2006; Kajiwara et al., 2006; Matsumoto et al., 2007; Bannister et al., 2007). Nevertheless, effects of estrogens are reported in these animals (e.g. sex reversal in several mollusc species) and this is taken as evidence in favor of the existence of a classical nuclear estrogen receptor (see Lafont and Mathieu, 2007 for references). However, the actual data suggests that if such a receptor exists it is not the unique ER orthologue already characterized. Among the several possible explanations we propose that: (i) the molecule used for the treatment is metabolized to another compound that is the active one. Indeed we recently discovered such a situation in amphioxus, where thyroid hormones are metabolized to a compound that binds the TR (Paris et al., 2008a); (ii) a NR that binds estrogen exists but this is not an orthologue of ER; (iii) the effect of estrogen is mediated by another receptor system, e.g. a transmembrane receptor (see Revankar et al., 2005 for an example) or a cytoplasmic protein such as in the case of *Glomus* discussed above. Thus one should treat claims of the existence of a signaling pathway based on treatment with human derived compound with much caution. Such data provide indications, but only indirect evidence that should be verified by more functional approaches including the cloning and characterisation of the relevant receptor.

4.2. Detection

Detecting a given ligand in one's favorite non-classical model organism is probably an excellent option for a comparative endocrinologist. One interesting case to illustrate the complexity of ligand evolution in different organisms is the one of the lamprey steroid hormones. For a long time, it was thought that as a vertebrate, lamprey should produce the classical steroids, estradiol, testosterone, and progesterone, and indeed, these were effectively identified by radioimmunoassays (RIA) and used for physiological (for an example see Bolduc and Sower, 1992) or functional (Thornton, 2001) studies. Nevertheless, in the 1980s, chromatographic tech-

Please cite this article in press as: Markov, G.V., et al., The evolution of the ligand/receptor couple: A long road from comparative endocrinology to comparative genomics. Mol. Cell. Endocrinol. (2008), doi:10.1016/j.mce.2008.06.011

niques and more recently, blood steroid profiles analyzed by high performance liquid chromatography (HPLC), contradicted this view and showed that the main circulating steroids are 15-alpha-hydroxylated (Kime and Callard, 1982; Lowartz et al., 2003). The actual nature of the active steroid hormones and their mechanism of action in lamprey are still a matter of debate (reviewed in Bryan et al., 2008). This example shows that the diversity of ligands in non-model species is often underestimated. In this case the presence of the classical steroids is not questioned, but they are not the physiologically relevant compounds.

Many reports have shown the presence of a typical human hormone in early vertebrates or even in invertebrates, steroids being the compounds for which the number of such reports is the highest. Most often the detection of these products is based on the use of antibodies, generated from human compounds and widely used on human tissue samples for medical application. These antibodies are mainly used for immunohistochemical staining and/or radioimmunoassay and two recent examples illustrate the caveats and uncertainties of such approaches. We have nevertheless to strongly insist on the fact that these examples were not selected because they are particularly dubious but rather because they illustrate numerous studies done in classical comparative endocrinology.

The first example concerns the detection of immunoreactivity for progesterone in the giant Rohde cells of the amphioxus *Branchiostoma belcheri* (Takeda et al., 2003). Using rabbit polyclonal antibodies against progesterone conjugated to bovine serum albumin from two commercial suppliers, the authors performed immunohistochemical staining on adult amphioxus and show staining in giant neurons known as the Rohde cells. The authors cautiously conclude that "progesterone-like" substances are likely to be present in these neurons, suggesting the existence of neurosteroids in this species. They performed a number of controls to avoid problems due to non-specific binding: they replaced the primary antibody by normal rabbit serum, they omitted the primary antibody from the staining reaction and they used PBS instead of the primary antibody. Furthermore, they also performed absorption tests with the antibodies, progesterone, and BSA, respectively, corroborating the fact that the staining is specific. The main problem with this type of paper is that we have no information on the actual specificity of the antibodies. Are these reagents able to recognize only progesterone itself or do they show cross-reactivity with other closely related molecules such as pregnenolone, deoxycorticosterone or 17β-hydroxyprogesterone? What if the amphioxus contains no progesterone but related steroids, such as those found in lamprey? The authors of this study were cautious and spoke of "progesterone-like substances". It is interesting to note that the amphioxus genome contains several, but not all the enzymes implicated in steroidogenesis (e.g. a CYP21 orthologue is missing) and that several of these enzymes are duplicated (e.g. 3β-HSD for which six genes were found in amphioxus, compared to three in human and the SDR family that contains enzymes with 17β-HSD and 11β-HSD activities for which 31 genes are known in human for more than 100 in amphioxus) (Holland et al., 2008). However, only one steroid receptor, located at the base of the vertebrate GR, MR, PR and AR, is present in the amphioxus genome, and, to date, nothing is known about its specificity. Thus it would not be surprising if amphioxus steroidogenesis has been elaborated independently from vertebrate steroidogenesis, and searching for human compounds in amphioxus may simply be unfruitful.

The second example concerns the presence of estrogens in molluscs, and more precisely in the cephalopod *Octopus vulgaris*, an interesting case since it is much referred to in the debate on the ligand-binding evolution of estrogen receptors (Thornton et al., 2003; Keay et al., 2006; Paris et al., 2008b). It has been shown that, in this species, 17β-estradiol and progesterone are found in oviduct and ovarian tissues, and that the concentration of these hormones in females correlates with phases of the reproductive cycle (Di Cosmo et al., 2001; D'Aniello et al., 1996). The doses of hormones in these tissues are measured using a radioimmunoassay and we are faced with the same uncertainties about the specificity of the detection method in this study as in the previous one. In addition, the level of hormones detected both at the level of radioimmunoassay or by HPLC is very close to the lower limits of detection (D'Aniello et al., 1996). From all the data accumulated in *Octopus*, one can conclude that steroid hormone metabolism does exist in this species, but it is difficult to be more conclusive on the precise identity of the steroids that will be found (see below).

These two examples emphasize the difficulties associated with detection based on the use of antibodies. In fact, before the genome era the same situation was found for NR and many reports claimed for the detection of steroid receptors in a wide variety of non-metazoans organisms such as plants or fungi (see Agarwal et al., 1994; Milanesi et al., 2001; Milanesi and Boland, 2006). Despite the fact that, as in the above mentioned studies, all the controls were correct, all these studies displayed artefacts: there is no gene related to steroid receptors outside metazoans (Escriva et al., 1997, 2004). We believe that basically the same situation holds for the immunological detection of NR ligands in invertebrates: perhaps some of these reports are correct but in the absence of more firm evidence this cannot be confirmed (Lafont and Mathieu, 2007). Thus the use of antibodies as an evolutionary probe, even if this is attractive because these experiments are technically easy to perform, should be avoided.

4.3. Other lines of evidence

Other methods of analysis for the presence of specific compounds can be used. This is for example the ability of a given tissue to metabolize a given compound, i.e. to carry out a given reaction (e.g. hydroxylation at the position 3β indicative of a 3β-HSD activity see Di Cosmo et al., 2001). Once again the problem in these experiments is the conclusions that can reasonably be drawn. That a given type of activity exists *in vivo* is interesting but given the diversity of steroidogenic enzymes (or enzymes using different substrates) in metazoans one cannot take for granted that because a reaction occurs with a given labelled intermediate it actually uses this very substrate *in vivo*.

Another approach often used is searching for the presence of the gene encoding an enzyme responsible for a critical step in the synthesis pathway of a given ligand. For example the existence of an aromatase in amphioxus (Castro et al., 2005) argues for the existence of estrogens in this species. This is convincing in this particular case since the phylogeny of the isolated clone was carefully assessed including synteny analysis. This is much less clear in other cases, such as the recent description of putative CYP11A and CYP17 in amphioxus (Mizuta and Kubokawa, 2007). The phylogeny of these clones, taking into account a more exhaustive dataset, with sequences from non-chordate species, shows that the phylogenetical inferences of this report were inexact and that in both cases, the cloned gene are orthologous to a new group of deuterostome CYPs, that was lost in vertebrates (Markov et al., in preparation).

Our critical arguments may well in turn be criticised as, we agree, it is easy to operate previous publications without offering positive alternatives. It seems clear that the best studies are those done with careful analytical chemistry methods. For example HPLC coupled with mass spectrometry allows the clear detection and identification of compounds and offers a very strong and rigorous alternative (Bridgham et al., 2006 see also Lafont and Mathieu, 2007 for a similar conclusion). The problem with this method is

Please cite this article in press as: Markov, G.V., et al., The evolution of the ligand/receptor couple: A long road from comparative endocrinology to comparative genomics. Mol. Cell. Endocrinol. (2008), doi:10.1016/j.mce.2008.06.011

of course that it is much more demanding and expensive than antibody detection. But they have the unique advantage of unambiguously assessing the presence and diversity of the compounds found in a given organism. The recent characterization of steroids present in the nematode *C. elegans* provides an interesting example of the variety of compounds that can be revealed through such a detailed analysis (Motola et al., 2006).

5. Conclusion—evolutionary vs. comparative approaches: an evolutionary shift for comparative endocrinology

Most often, endocrinology studies are realized with a long-term medical objective. This leads to a deformation of the evolutionary perspective, which may be too much "human-centered". Doing comparative endocrinology should therefore first be accompanied by an effort to go again this natural anthropocentric view and to consider equally the evidences coming from different taxa. This may seems obvious, but in fact "rampant" anthropocentrism and a graded view of evolution (e.g. that evolution is progressing toward complexity) is often difficult to avoid as nicely pointed out by Gould (1996).

In fact a key solution to these methodological and interpretation difficulties relies in the comparative strategy itself that should incorporate evidences coming from various organisms and scientific approaches in an integrated manner. An interesting example of the power of such an approach is the case of the suiform aromatases (Gaucher et al., 2004). In this study, by combining bioinformatic, molecular evolution, paleontology, cladistics, structural biology and organic chemistry analysis, the authors propose that the conservation of three subfunctionalised aromatase paralogues in pigs is the result of a selection for Suoidea with larger litter than their ancestors. This selective event has allowed their survival during the global climatic shift that began in the Eocene. Bioinformatic analyses (estimation of divergence times, detection of positive selection by K_a/K_s analysis) were correlated with the presence of residues that were subject to positive selection in the substrate-binding site and with previous experimental data about different substrate specificities for these enzymes. Additionally, a detailed examination of the palaeontological record and of the number of pups in modern artiodactyls as well, correlated with data on global climate changes, led to this quite audacious hypothesis, that is consistent with data of many different research fields. It would be interesting – and possible, thanks to the availability of genome data – to check if such a correlation occurs in other vertebrates and could be statistically significant.

We argue that future comparative endocrinology studies should combine large-scale evolutionary analysis, with several standardised phylogenetic and chemical methods to ensure the robustness of the conclusions. The classical two-by-two comparison that is prone to artefactual conclusions based on partial analysis and the use of poorly refined detection methods, such as those based on antibodies should be replaced by such multidisciplinary studies, even if these approaches are experimentally much more difficult. Comparative endocrinology is thus now facing a real challenge: to perform multidisciplinary evolutionary approaches that will effectively offer solutions, using rigorous technical and conceptual basis, to the long-standing questions of the origin of the endocrine signalling pathways.

Acknowledgements

We are grateful to ARC, CNRS, EMBO, MENRT, and Region Rhone-Alpes for financial support. Work from our laboratory is also funded by the EU Cascade Network of Excellence and the integrated project Crescendo. We thank Barbara Demeneix, François Bonneton and Hector Escriva for critical reading of the manuscript.

References

Agarwal, M.K., Mirshahi, M., Braq, S., Jullienne, A., Leblanc, N., Stibon, F., Guern, J., 1994. *Nicotiana tabacum* contains a putative mineralocorticoid receptor. Biochem. Biophys. Res. Commun. 200, 1230–1238.
Aspöck, G., Kagoshima, H., Niklaus, G., Bürglin, T.R., 1999. *Caenorhabditis elegans* has scores of hedgehog-related genes: sequence and expression analysis. Genome Res. 9, 909–923.
Baker, M.E., 2003. Evolution of adrenal and sex steroid action in vertebrates: a ligand-based mechanism for complexity. Mol. Cell. Endocrinol. 25, 396–400.
Bannister, R., Beresford, N., May, D., Routledge, E.J., Jobling, S., Rand-Weaver, M., 2007. Novel estrogen receptor-related transcripts in *Marisa cornuarietis*, a freshwater snail with reported sensitivity to estrogenic chemicals. Environ. Sci. Technol. 41, 2643–2650.
Bélanger, A., Hum, D.W., Beaulieu, M., Lévesque, E., Guillemette, C., Tchernof, A., Bélanger, G., Turgeon, D., Dubois, S., 1998. Characterization and regulation of UDP-glucuronosyltransferases in steroid target tissues. J. Steroid Biochem. Mol. Biol. 65, 301–310.
Bertrand, S., Brunet, F.G., Escriva, H., Parmentier, G., Laudet, V., Robinson-Rechavi, M., 2004. Evolutionary genomics of nuclear receptors: from twenty-five ancestral genes to derived endocrine systems. Mol. Biol. Evol. 21, 1923–1937.
Bertrand, S., Thisse, B., Tavares, R., Sachs, L., Chaumot, A., Bardet, P., Escrivà, H., Duffraisse, M., Marchand, O., Safi, R., Thisse, C., Laudet, V., 2007. Unexpected novel relational links uncovered by extensive developmental profiling of nuclear receptor expression. PLoS Genet. 3, e188.
Bolduc, T.G., Sower, S.A., 1992. Changes in brain gonadotropin-releasing hormone, plasma estradiol 17-beta, and progesterone during the final reproductive cycle of the female sea lamprey, *Petromyzon marinus*. J. Exp. Zool. 264, 55–63.
Bonneton, F., Zelus, D., Iwema, T., Robinson-Rechavi, M., Laudet, V., 2003. Rapid divergence of the ecdysone receptor in Diptera and Lepidoptera suggests coevolution between ECR and USP-RXR. Mol. Biol. Evol. 20, 541–553.
Bonneton, F., Brunet, F.G., Kathirithamby, J., Laudet, V., 2006. The rapid divergence of the ecdysone receptor is a synapomorphy for Mecopterida that clarifies the Strepsiptera problem. Insect Mol. Biol. 15, 351–362.
Bonneton, F., Chaumot, A., Laudet, V., 2008. Annotation of *Tribolium* nuclear receptors reveals an increase in evolutionary rate of a network controlling the ecdysone cascade. Insect Biochem. Mol. Biol. 38, 416–429.
Bridgham, J.T., Carroll, S.M., Thornton, J.W., 2006. Evolution of hormone–receptor complexity by molecular exploitation. Science 312, 97–101.
Bryan, M.B., Scott, A.P., Li, W., 2008. Sex steroids and their receptors in lampreys. Steroids 73, 1–12.
Calléja, C., Messaddeq, N., Chapellier, B., Yang, H., Krezel, W., Li, M., Metzger, D., Mascrez, B., Ohta, K., Kagechika, H., Endo, Y., Mark, M., Ghyselinck, N.B., Chambon, P., 2006. Genetic and pharmacological evidence that a retinoic acid cannot be the RXR-activating ligand in mouse epidermis keratinocytes. Genes Dev. 20, 1525–1538.
Cardoso, J.C.R., Pinto, V.C., Vieira, F.A., Clark, M.S., Power, D.M., 2006. Evolution of secretin family GPCR members in the metazoa. BMC Evol. Biol. 6, 108.
Cardoso, J., de Vet, E., Louro, B., Elgar, G., Clark, M., Power, D., 2007. Persistence of duplicated PAC1 receptors in the teleost, *Sparus auratus*. BMC Evol. Biol. 7, 221.
Carlsson, P., Mahlapuu, M., 2002. Forkhead transcription factors: key players in development and metabolism. Dev. Biol. 250, 1–23.
Castro, L.F.C., Santos, M.M., Reis-Henriques, M.A., 2005. The genomic environment around the Aromatase gene: evolutionary insights. BMC Evol. Biol. 5, 43.
Catford, J.G., Staehelin, C., Larose, G., Piché, Y., Vierheilig, H., 2006. Systemically suppressed isoflavonoids and their stimulating effects on nodulation and mycorrhization in alfalfa split-root systems. Plant Soil 285, 257–266.
Cheng, G., Yeater, K.M., Hoyer, L.L., 2006. Cellular and molecular biology of *Candida albicans* estrogen response. Eukaryot. Cell. 5, 180–191.
Crollius, H.R., Weissenbach, J., 2005. Fish genomics and biology. Genome Res. 15, 1675–1682.
D'Aniello, A., Cosmo, A.D., Cristo, C.D., Assisi, L., Botte, V., Fiore, M.M.D., 1996. Occurrence of sex steroid hormones and their binding proteins in *Octopus vulgaris* lam. Biochem. Biophys. Res. Commun. 227, 782–788.
Danchin, E.G.J., Abi-Rached, L., Gilles, A., Pontarotti, P., 2003. Conservation of the MHC-like region throughout evolution. Immunogenetics 55, 141–148.
Danchin, E., Gouret, P., Pontarotti, P., 2006. Eleven ancestral gene families lost in mammals and vertebrates while otherwise universally conserved in animals. BMC Evol. Biol. 6, 5.
Dean, A.M., Thornton, J.W., 2007. Mechanistic approaches to the study of evolution: the functional synthesis. Nat. Rev. Genet. 8, 675–688.
Dehal, P., Boore, J.L., 2005. Two rounds of whole genome duplication in the ancestral vertebrate. PLoS Biol. 3, e314.
Delsuc, F., Brinkmann, H., Philippe, H., 2005. Phylogenomics and the reconstruction of the tree of life. Nat. Rev. Genet. 6, 361–375.
Di Cosmo, A., Cristo, C.D., Paolucci, M., 2001. Sex steroid hormone fluctuations and morphological changes of the reproductive system of the female of *Octopus vulgaris* throughout the annual cycle. J. Exp. Zool. 289, 33–47.
Escriva, H., Safi, R., Hänni, C., Langlois, M.C., Saumitou-Laprade, P., Stehelin, D., Capron, A., Pierce, R., Laudet, V., 1997. Ligand binding was acquired dur-

Please cite this article in press as: Markov, G.V., et al., The evolution of the ligand/receptor couple: A long road from comparative endocrinology to comparative genomics. Mol. Cell. Endocrinol. (2008), doi:10.1016/j.mce.2008.06.011

ing evolution of nuclear receptors. Proc. Natl. Acad. Sci. U.S.A. 94, 6803–6808.
Escriva, H., Delaunay, F., Laudet, V., 2000. Ligand binding and nuclear receptor evolution. Mol. Cell. Endocrinol. 22, 717–727.
Escriva, H., Bertrand, S., Laudet, V., 2004. The evolution of the nuclear receptor superfamily. Essays Biochem. 40, 11–26.
Escriva, H., Bertrand, S., Germain, P., Robinson-Rechavi, M., Umbhauer, M., Cartry, J., Duffraisse, M., Holland, L., Gronemeyer, H., Laudet, V., 2006. Neofunctionalization in vertebrates: the example of retinoic acid receptors. PLoS Genet. 2, e102.
Farid, N.R., Szkudlinski, M.W., 2004. Minireview: structural and functional evolution of the thyrotropin receptor. Endocrinology 145, 4048–4057.
Force, A., Lynch, M., Pickett, F.B., Amores, A., Yan, Y.L., Postlethwait, J., 1999. Preservation of duplicate genes by complementary, degenerative mutations. Genetics 151, 1531–1545.
Garcia-Fernández, J., Holland, P.W., 1994. Archetypal organization of the amphioxus Hox gene cluster. Nature 370, 563–566.
Gaucher, E.A., Graddy, L.G., Li, T., Simmen, R.C.M., Simmen, F.A., Schreiber, D.R., Liberles, D.A., Janis, C.M., Benner, S.A., 2004. The planetary biology of cytochrome P450 aromatases. BMC Biol. 2, 19.
Gould, S.J., 1996. Full House. Harmony Books, New York.
Heyland, A., Price, D.A., Bodnarova-Buganova, M., Moroz, L.L., 2006. Thyroid hormone metabolism and peroxidase function in two non-chordate animals. J. Exp. Zool. B Mol. Dev. Evol. 306, 551–566.
Holland, L.Z., Albalat, R., Azumi, K., Benito-Gutierrez, E., Bronner-Fraser, M., Brunet, F., Butts, T., Candiani, S., Dishaw, L.J., Garcia-Fernandez, J., Ferrier, D.E.K., Gibson-Brown, J.J., Gissi, C., Godzik, A., Hallbook, F., Hirose, D., Hosomichi, K., Ikuta, T., Inoko, H., Kasahara, M., Kasamatsu, J., Kawashima, T., Kimura, A., Kobayashi, M., Kozmik, Z., Kubokawa, K., Laudet, V., Litman, G.W., McHardy, A.C., Meulemans, D., Nonaka, M., Olinski, R.P., Pancer, Z., Pestarino, M., Rast, J.P., Rigoutsos, I., Roch, G., Saiga, H., Sasakura, Y., Satake, M., Satou, Y., Schubert, M., Sherwood, N., Shiina, T., Takatori, N., Tello, J., Vopalensky, P., Wada, S., Xu, A., Ye, Y., Yoshida, K., Yoshizaki, F., Yu, J.K., Zhang, Q., Zmasek, C.M., Putnam, N.H., Rokhsar, D.S., Satoh, N., Holland, P.W.H., 2008. The amphioxus genome illuminates vertebrate origins and cephalochordate biology. Genome Res. 18, 1100–1111.
Iwema, T., Billas, I.M.L., Beck, Y., Bonneton, F., Nierengarten, H., Chaumot, A., Richards, G., Laudet, V., Moras, D., 2007. Structural and functional characterization of a novel type of ligand-independent RXR-USP receptor. EMBO J. 26, 3770–3782.
Jaillon, O., Aury, J., Brunet, F., Petit, J., Stange-Thomann, N., et al., 2004. Genome duplication in the teleost fish *Tetraodon nigroviridis* reveals the early vertebrate proto-karyotype. Nature 431, 946–957.
Kajiwara, M., Kuraku, S., Kurokawa, T., Kato, K., Toda, S., Hirose, H., Takahashi, S., Shibata, Y., Iguchi, T., Matsumoto, T., Miyata, T., Miura, T., Takahashi, Y., 2006. Tissue preferential expression of estrogen receptor gene in the marine snail, *Thais clavigera*. Gen. Comp. Endocrinol. 148, 315–326.
Kakizawa, T., Nishio, S., Triqueneaux, G., Bertrand, S., Rambaud, J., Laudet, V., 2007. Two differentially active alternative promoters control the expression of the zebrafish orphan nuclear receptor gene Rev-erbalpha. J. Mol. Endocrinol. 38, 555–568.
Kasahara, M., Naruse, K., Sasaki, S., Nakatani, Y., Qu, W., Ahsan, B., Yamada, T., Nagayasu, Y., Doi, K., Kasai, Y., Jindo, T., Kobayashi, D., Shimada, A., Toyoda, A., Kuroki, Y., Fujiyama, A., Sasaki, T., Shimizu, A., Asakawa, S., Shimizu, N., Hashimoto, S., Yang, J., Lee, Y., Matsushima, K., Sugano, S., Sakaizumi, M., Narita, T., Ohishi, K., Haga, S., Ohta, F., Nomoto, H., Nogata, K., Morishita, T., Endo, T., Shin-I, T., Takeda, H., Morishita, S., Kohara, Y., 2007. The medaka draft genome and insights into vertebrate genome evolution. Nature 447, 714–719.
Keay, J., Bridgham, J.T., Thornton, J.W., 2006. The *Octopus vulgaris* estrogen receptor is a constitutive transcriptional activator: evolutionary and functional implications. Endocrinology 147, 3861–3869.
Kime, D.E., Callard, G.V., 1982. Formation of 15 alpha-hydroxylated androgens by the testis and other tissues of the sea lamprey, *Petromyzon marinus*, in vitro. Gen. Comp. Endocrinol. 46, 267–270.
Klovins, J., Haitina, T., Fridmanis, D., Kilianova, Z., Kapa, I., Fredriksson, R., Gallo-Payet, N., Schiöth, H.B., 2004. The melanocortin system in Fugu: determination of POMC/AGRP/MCR gene repertoire and synteny, as well as pharmacology and anatomical distribution of the MCRs. Mol. Biol. Evol. 21, 563–579.
Krasowski, M.D., Yasuda, K., Hagey, L.R., Schuetz, E.G., 2005. Evolution of the pregnane × receptor: adaptation to cross-species differences in biliary bile salts. Mol. Endocrinol. 19, 1720–1739.
Lafont, R., Mathieu, M., 2007. Steroids in aquatic invertebrates. Ecotoxicology 16, 109–130.
Lagerström, M.C., Fredriksson, R., Bjarnadóttir, T.K., Fridmanis, D., Holmquist, T., Andersson, J., Yan, Y., Raudsepp, T., Zoorob, R., Kukkonen, J.P., Lundin, L., Klovins, J., Chowdhary, B.P., Postlethwait, J.H., Schiöth, H.B., 2005. Origin of the prolactin-releasing hormone (PRLH) receptors: evidence of coevolution between PRLH and a redundant neuropeptide Y receptor during vertebrate evolution. Genomics 85, 688–703.
Laudet, V., 1997. Evolution of the nuclear receptor superfamily: early diversification from an ancestral orphan receptor. J. Mol. Endocrinol. 19, 207–226.
Laudet, V., Gronemeyer, H., 2005. The Nuclear Receptor FactsBook. Academic Press, London.
Lalueza-Fox, C., Römpler, H., Caramelli, D., Stäubert, C., Catalano, G., et al., 2007. Melanocortin 1 receptor allele suggests varying pigmentation among Neanderthals. Science 318, 1453–1455.
Li, Y., Wallis, M., Zhang, Y., 2005. Episodic evolution of prolactin receptor gene in mammals: coevolution with its ligand. J. Mol. Endocrinol. 35, 411–419.
Liu, J.C., Makova, K.D., Adkins, R.M., Gibson, S., Li, W.H., 2001. Episodic evolution of growth hormone in primates and emergence of the species specificity of human growth hormone receptor. Mol. Biol. Evol. 18, 945–953.
Lowartz, S., Petkam, R., Renaud, R., Beamish, F.W.H., Kime, D.E., Raeside, J., Leatherland, J.F., 2003. Blood steroid profile and in vitro steroidogenesis by ovarian follicles and testis fragments of adult sea lamprey, *Petromyzon marinus*. Comp. Biochem. Physiol. A Mol. Integr. Physiol. 134, 365–376.
Macqueen, D.J., Johnston, I.A., 2008. An update on MyoD evolution in teleosts and a proposed consensus nomenclature to accommodate the tetraploidization of different vertebrate genomes. PLoS ONE 3, e1567.
Madani, N.D., Malloy, P.J., Rodriguez-Pombo, P., Krishnan, A.V., Feldman, D., 1994. *Candida albicans* estrogen-binding protein gene encodes an oxidoreductase that is inhibited by estradiol. Proc. Natl. Acad. Sci. U.S.A. 91, 922–926.
Markov, G., Lecointre, G., Demeneix, B., Laudet, V., 2008. The «street light syndrome», or how protein taxonomy can bias experimental manipulations. Bioessays 30, 349–357.
Matsumoto, T., Nakamura, A.M., Mori, K., Akiyama, I., Hirose, H., Takahashi, Y., 2007. Oyster estrogen receptor: cDNA cloning and immunolocalization. Gen. Comp. Endocrinol. 151, 195–201.
McPartland, J.M., Norris, R.W., Kilpatrick, C.W., 2007. Coevolution between cannabinoid receptors and endocannabinoid ligands. Gene 397, 126–135.
Mic, F.A., Molotkov, A., Benbrook, D.M., Duester, G., 2003. Retinoid activation of retinoic acid receptor but not retinoid X receptor is sufficient to rescue lethal defect in retinoic acid synthesis. Proc. Natl. Acad. Sci. U.S.A. 100, 7135–7140.
Milanesi, L., Monje, P., Boland, R., 2001. Presence of estrogens and estrogen receptor-like proteins in *Solanum glaucophyllum*. Biochem. Biophys. Res. Commun. 289, 1175–1179.
Milanesi, L., Boland, R., 2006. Presence of vitamin D3 receptor (VDR)-like proteins in *Solanum glaucophyllum*. Physiol. Plant. 128, 341–350.
Mizuta, T., Kubokawa, K., 2007. Presence of sex steroids and cytochrome P450 genes in amphioxus. Endocrinology 148, 3554–3565.
Motola, D.L., Cummins, C.L., Rottiers, V., Sharma, K.K., Li, T., Li, Y., Suino-Powell, K., Xu, H.E., Auchus, R.J., Antebi, A., Mangelsdorf, D.J., 2006. Identification of ligands for DAF-12 that govern dauer formation and reproduction in *C. elegans*. Cell 124, 1209–1223.
Moyle, W.R., Campbell, R.K., Myers, R.V., Bernard, M.P., Han, Y., Wang, X., 1994. Co-evolution of ligand–receptor pairs. Nature 368, 251–255.
Muffato, M., Crollius, H.R., 2008. Paleogenomics in vertebrates, or the recovery of lost genomes from the mist of time. Bioessays 30, 122–134.
Nelson, D.R., 2005. Gene nomenclature by default, or BLASTing to Babel. Hum. Genom. 2, 196–201.
Norris, D.O., 2007. Vertebrate Endocrinology. Elsevier Academic Press, San Diego.
O'Halloran, D.M., Fitzpatrick, D.A., McCormack, G.P., McInerney, J.O., Burnell, A.M., 2006. The molecular phylogeny and functional significance of a nematode specific clade of heterotrimeric G-protein α-subunit genes. J. Mol. Evol. 63, 87–94.
Ogasawara, M., Lauro, R.D., Satoh, N., 1999. Ascidian homologs of mammalian thyroid peroxidase genes are expressed in the thyroid-equivalent region of the endostyle. J. Exp. Zool. 285, 158–169.
Ogasawara, M., 2000. Overlapping expression of amphioxus homologs of the thyroid transcription factor-1 gene and thyroid peroxidase gene in the endostyle: insight into evolution of the thyroid gland. Dev. Genes. Evol. 210, 231–242.
Ohno, S., 1970. Evolution by Gene Duplication. Springer-Verlag, Heidelberg.
Panopoulou, G., Poustka, A.J., 2005. Timing and mechanism of ancient vertebrate genome duplications—the adventure of a hypothesis. Trends Genet. 21, 559–567.
Pantalacci, S., Chaumont, C., Benoît, G., Sadier, A., Delsuc, F., Douzery, E., Laudet, V., 2008. Conserved features and evolutionary shifts of the EDA signaling pathway involved in vertebrate skin appendage development. Mol. Biol. Evol. 25, 912–928.
Paris, M., Escriva, H., Schubert, M., Brunet, F., Brtko, J., Cieselski, F., Roecklin, D., Vivat-Hannah, V., Cravedi, J.-P., Scanlan, T.S., Renaud, J.-P., Holland, N.D., Laudet, V., 2008a. Amphioxus metamorphosis and the origin of the thyroid hormone signaling pathway. Curr. Biol. 18, 825–830.
Paris, M., Pettersson, K., Schubert, M., Bertrand, S., Pongratz, I., Escriva, H., Laudet, V., 2008b. An amphioxus orthologue of the estrogen receptor that does not bind estradiol: insight into estrogen receptor evolution. BMC Evol. Biol., in press.
Park, P.S., Palczewski, K., 2005. Diversifying the repertoire of G protein-coupled receptors through oligomerization. Proc. Natl. Acad. Sci. U.S.A. 102, 8793–8794.
Pinaire, J.A., Reifel-Miller, A., 2007. Therapeutic potential of retinoid × receptor modulators for the treatment of the metabolic syndrome. PPAR Res., 941–956.
Poulin, M.J., Simard, J., Catford, J.G., Labrie, F., Piche, Y., 1997. Response of symbiotic endomycorrhizal fungi to estrogens and antiestrogens. Mol. Plant Microb. Interact., 10.
Prüll, C.-R., 2003. Part of a scientific master plan? Paul Ehrlich and the origins of his receptor concept. Med. Hist. 47, 332–356.
Putnam, N.H., Srivastava, M., Hellsten, U., Dirks, B., Chapman, J., Salamov, A., Terry, A., Shapiro, H., Lindquist, E., Kapitonov, V.V., Jurka, J., Genikhovich, G., Grigoriev, I.V., Lucas, S.M., Steele, R.E., Finnerty, J.R., Technau, U., Martindale, M.Q., Rokhsar, D.S., 2007. Sea anemone genome reveals ancestral eumetazoan gene repertoire and genomic organization. Science 317, 86–94.
Reschly, E.J., Bainy, A.C.D., Mattos, J.J., Hagey, L.R., Bahary, N., Mada, S.R., Ou, J., Venkataramanan, R., Krasowski, M.D., 2007. Functional evolution of the vitamin D and pregnane X receptors. BMC Evol. Biol. 7, 222.

Please cite this article in press as: Markov, G.V., et al., The evolution of the ligand/receptor couple: A long road from comparative endocrinology to comparative genomics. Mol. Cell. Endocrinol. (2008), doi:10.1016/j.mce.2008.06.011

Requena, N., Serrano, E., Ocón, A., Breuninger, M., 2007. Plant signals and fungal perception during arbuscular mycorrhiza establishment. Phytochemistry 68, 33–40.
Revankar, C.M., Cimino, D.F., Sklar, L.A., Arterburn, J.B., Prossnitz, E.R., 2005. A transmembrane intracellular estrogen receptor mediates rapid cell signaling. Science 307, 1625–1630.
Rikke, B.A., Murakami, S., Johnson, T.E., 2000. Paralogy and orthology of tyrosine kinases that can extend the life span of *Caenorhabditis elegans*. Mol. Biol. Evol. 17, 671–683.
Robinson-Rechavi, M., Maina, C.V., Gissendanner, C.R., Laudet, V., Sluder, A., 2005. Explosive lineage-specific expansion of the orphan nuclear receptor HNF4 in nematodes. J. Mol. Evol. 60, 577–586.
Rogozin, I.B., Wolf, Y.I., Carmel, L., Koonin, E.V., 2007. Analysis of rare amino acid replacements supports the coelomata clade. Mol. Biol. Evol. 24, 2594–2597.
Scervino, J.M., Ponce, M.A., Erra-Bassells, R., Vierheilig, H., Ocampo, J.A., Godeas, A., 2005. Arbuscular mycorrhizal colonization of tomato by Gigaspora and Glomus species in the presence of root flavonoids. J. Plant. Physiol. 162, 625–633.
Simionato, E., Ledent, V., Richards, G., Thomas-Chollier, M., Kerner, P., Coornaert, D., Degnan, B.M., Vervoort, M., 2007. Origin and diversification of the basic helix–loop–helix gene family in metazoans: insights from comparative genomics. BMC Evol. Biol. 7, 33.
Simões-Costa, M.S., Azambuja, A.P., Xavier-Neto, J., 2008. The search for non-chordate retinoic acid signaling: lessons from chordates. J. Exp. Zool. B Mol. Dev. Evol. 310, 54–72.
Sluder, A.E., Maina, C.V., 2001. Nuclear receptors in nematodes: themes and variations. Trends Genet. 17, 206–213.
Takeda, N., Kubokawa, K., Matsumoto, G., 2003. Immunoreactivity for progesterone in the giant Rohde cells of the amphioxus, *Branchiostoma belcheri*. Gen. Comp. Endocrinol. 132, 379–383.
Thomas, J.H., 2007. Rapid birth-death evolution specific to xenobiotic cytochrome P450 genes in vertebrates. PLoS Genet. 3, e67.
Thornton, J.W., 2001. Evolution of vertebrate steroid receptors from an ancestral estrogen receptor by ligand exploitation and serial genome expansions. Proc. Natl. Acad. Sci. U.S.A. 98, 5671–5676.
Thornton, J.W., Need, E., Crews, D., 2003. Resurrecting the ancestral steroid receptor: ancient origin of estrogen signaling. Science 301, 1714–1717.
Thornton, J.W., 2004. Resurrecting ancient genes: experimental analysis of extinct molecules. Nat. Rev. Genet. 5, 366–375.
Volff, J., 2005. Genome evolution and biodiversity in teleost fish. Heredity 94, 280–294.
Waddell, P.J., Kishino, H., Ota, R., 2007. Phylogenetic methodology for detecting protein interactions. Mol. Biol. Evol. 24, 650–659.
Wallis, M., 2001. Episodic evolution of protein hormones in mammals. J. Mol. Evol. 53, 10–18.
Wittbrodt, J., Meyer, A., Schartl, M., 1998. More genes in fish? Mol. Cell. Endocrinol. 20, 511–512.
Wu, W., Niles, E.G., El-Sayed, N., Berriman, M., LoVerde, P.T., 2006. *Schistosoma mansoni* (Platyhelminthes, Trematoda) nuclear receptors: sixteen new members and a novel subfamily. Gene 366, 303–315.
Wu, W., Niles, E.G., LoVerde, P.T., 2007. Thyroid hormone receptor orthologues from invertebrate species with emphasis on *Schistosoma mansoni*. BMC Evol. Biol. 7, 150.
Wyder, S., Kriventseva, E., Schröder, R., Kadowaki, T., Zdobnov, E., 2007. Quantification of ortholog losses in insects and vertebrates. Genome Biol. 8, R242.
You, L., 2004. Steroid hormone biotransformation and xenobiotic induction of hepatic steroid metabolizing enzymes. Chem. Biol. Interact. 147, 233–246.

Please cite this article in press as: Markov, G.V., et al., The evolution of the ligand/receptor couple: A long road from comparative endocrinology to comparative genomics. Mol. Cell. Endocrinol. (2008), doi:10.1016/j.mce.2008.06.011

JOURNAL OF EXPERIMENTAL ZOOLOGY (MOL DEV EVOL) 308B:484–493 (2007)

Insights Into Spawning Behavior and Development of the European Amphioxus (*Branchiostoma lanceolatum*)

MICHAEL FUENTES[1], ELIA BENITO[2], STEPHANIE BERTRAND[2], MATHILDE PARIS[4], AURELIE MIGNARDOT[5], LAURA GODOY[2], SENDA JIMENEZ-DELGADO[2], DIANA OLIVERI[3], SIMONA CANDIANI[3], ESTELLE HIRSINGER[6], SALVATORE D'ANIELLO[2], JUAN PASCUAL-ANAYA[2], IGNACIO MAESO[2], MARIO PESTARINO[3], PHILIPPE VERNIER[5], JEAN-FRANÇOIS NICOLAS[6], MICHAEL SCHUBERT[4], VINCENT LAUDET[4], ANNE MARIE GENEVIERE[1], RICARD ALBALAT[2], JORDI GARCIA FERNANDEZ[2], NICHOLAS D. HOLLAND[7], AND HECTOR ESCRIVA[1]*

[1]*Laboratoire Arago, UMR7628 CNRS and University Pierre and Marie Curie-Paris6 Banyuls sur Mer, F-66650 France*

[2]*Departament de Genetica, Facultat de Biologia, Universitat de Barcelona, 08007 Barcelona, Spain*

[3]*Dipartimento di Biologia, Università di Genova, 16126 Genova, Italy*

[4]*Institut de Génomique Fonctionnelle de Lyon, IFR128 BioSciences Lyon-Gerland, Ecole Normale Supérieure de Lyon, Lyon Cedex, France*

[5]*Developpement, Evolution et Plasticite du Systeme Nerveux, UPR 2197, Institut de Neurobiologie A. Fessard, CNRS Gif-sur-Yvette, France*

[6]*Département de Biologie du Développement, Institut Pasteur, 757242 Paris Cedex, France*

[7]*Marine Biology Research Division, Scripps Institution of Oceanography, University of California San Diego, La Jolla, California, USA, 92093*

ABSTRACT The cephalochordate amphioxus (*Branchiostoma* sp.) is an important animal model for studying the evolution of chordate developmental mechanisms. Obtaining amphioxus embryos is a key step for these studies. It has been shown that an increase of 3–4°C in water temperature triggers spawning of the European amphioxus (*Branchiostoma lanceolatum*) in captivity, however, very little is known about the natural spawning behavior of this species in the field. In this work, we have followed the spawning behavior of the European amphioxus during two spawning seasons (2004 and 2005), both in the field and in captivity. We show that animals in the field spawn approximately from mid-May through early July, but depending on the year, they show different patterns of spawning. Thus, even if temperature has a critical role in the induction of the spawning in captivity, it is not the major factor in the field. Moreover, we report some improvements on the methodology for inducing spawning in captivity (e.g. in maintenance, light cycle control and induction of spawning in a laboratory without running sea water system). These studies have important implications for amphioxus animal husbandry and for improving

Grant sponsors: ANR BLAN NT_NV_49; Postdoctoral fellowship from the European Community's Human Potential Programme HPRN-CT-2002-00263 ''Neurogenome''; MENRT, CNRS, ARC and the Région Rhône-Alpes; MEC BFU2005-00252; European project Embryomics (NEST012916).

*Correspondence to: Hector Escriva, Laboratoire Arago, UMR 7628, CNRS and University Pierre and Marie Curie, BP 44, F-66651 Banyuls sur Mer, France. E-mail: hescriva@obs-banyuls.fr

Received 5 March 2007; Revised 24 April 2007; Accepted 30 April 2007

Published online 22 May 2007 in Wiley InterScience (www.interscience.wiley.com). DOI: 10.1002/jez.b.21179

laboratory techniques to develop amphioxus as an experimental animal model. *J. Exp. Zool. (Mol. Dev. Evol.) 308B:484–493, 2007.*

How to cite this article: Fuentes M, Benito E, Bertrand S, Paris M, Mignardot A, Godoy L, Delgado SJ, Oliveri D, Candiani S, Hirsinger E, D'Aniello S, Pascual-Anaya J, Maeso I, Pestarino M, Vernier P, Nicolas JF, Schubert M, Laudet V, Geneviere AM, Albalat R, Garcia-Fernàndez J, Holland ND, Escriva H. 2007. Insights into spawning behavior and development of the european amphioxus (*Branchiostoma lanceolatum*). J. Exp. Zool. (Mol. Dev. Evol.) 308B:484–493.

The invertebrate chordate amphioxus is one of the closest living relatives of the vertebrates. Amphioxus is similar to vertebrates in many morphological and genomic respects, but simpler. Amphioxus and vertebrates posses, at some stage in their life, a hollow nerve cord dorsal to a notochord, a postanal tail, an endostyle (the homolog of the vertebrate thyroid gland) and pharyngeal gill slits. However, amphioxus lacks vertebrate-specific characters such as paired sensory organs or a cartilaginous or bony skeleton (Shimeld and Holland, 2005; Schubert et al., 2006). In addition, the amphioxus genome has representatives of most vertebrate gene families, but it has not undergone the extensive gene duplications characteristic of vertebrates (Abi-Rached et al., 2002; Dehal and Boore, 2005). For example, amphioxus has a single cluster of Hox genes whereas mammals have four (Garcia-Fernandez and Holland, '94) and a single retinoic acid receptor as compared with three in mammals (Escriva et al., 2002).

Owing to morphological similarities, amphioxus used to be classified as the sister group of the vertebrates, whereas tunicates were considered to be the sister group of the clade comprising amphioxus plus the vertebrates. However, recent studies demonstrate that amphioxus diverged early during chordate evolution and it represents the sister group of both tunicates and vertebrates (Bourlat et al., 2006; Delsuc et al., 2006). The vertebrate-like body plan and the close phylogenetic relationship to vertebrates already qualifies amphioxus as the best-available model for the proximate ancestor of the vertebrates and as a remarkable animal model for studying the evolution of developmental mechanisms (Evo-Devo). However, the new phylogenetic position of cephalochordates makes amphioxus not only a model for studying the evolution of vertebrates, but also for better understanding the elaboration of the entire chordate lineage.

The availability of living amphioxus, especially embryos, is a prerequisite for using amphioxus in studies on the evolution of developmental mechanisms. Today, three amphioxus species are used for these studies, the East Asian *Branchiostoma belcheri*, the American-Caribbean *Branchiostoma floridae* and the European *Branchiostoma lanceolatum*. We recently performed a preliminary study of the natural spawning behavior of the European species and we described a reliable method to induce spawning of mature adults in captivity (Fuentes et al., 2004). From the three amphioxus species studied today, *B. lanceolatum* is the only one for which spawning can be induced on a daily basis. In this report, we compare the spawning behavior of the European amphioxus during two spawning seasons in the field (i.e. 2004 and 2005). We also compare the developmental timing of amphioxus embryos between the Florida and the European species and we show several improvements on the previously described technique for inducing spawning in captivity. Our results, both from the field as well as from laboratory experiments will help develop efficient techniques for long-term amphioxus aquaculture.

MATERIALS AND METHODS

Sampling of amphioxus and field observations

During May and June 2004 and 2005, adult amphioxus (*B. lanceolatum*) were collected daily at the previously described site in Argeles-sur-Mer, France, and according to previously described methods (Fuentes et al., 2004). Collected animals were counted and classified as full versus empty (i.e. animals with visible gonads versus post-spawning stage) depending on their gonadal status.

Maintenance of amphioxus in the laboratory

Following collection, adult animals were kept in tanks with filtered seawater. Incoming seawater was filtered through four 1-mm grain-size sand filters of $1\,m^3$. Water was changed twice a week

and the animals were subjected to a 14-hr-light/10-hr dark cycle.

Animals with empty gonads (post-spawning stage) were maintained at 18–19°C temperature in 15-L rectangular plastic tanks (100–200 animals/tank) without sand. Water was changed twice a day and food was always added just after the water change. Animals were fed with crushed Tetramicrofood (0.5 g/feeding) (www.tetra-online.-com) supplemented with 15.000–40.000 cells/mL/feeding of mixed algae (equal parts of *Dunaliella tertiolecta*, *Isochrysis galbana* and *Tetraselmis suecica*). Algal cultures were grown in f/2 medium as described in Guillard et al. (Guillard and Ryther, '62). The f/2 medium was eliminated by centrifugation before feeding the algae to the amphioxus. The animals used in temperature shock experiments were maintained in starvation at 18–19°C temperature in 15-L plastic tanks without sand in a room with ambient light. The water on these animals was changed twice a week.

Temperature measurements in the field

A thermo data logger (HOBO Pendant Temp/Light Intensity Data logger, ONSET Computer Corporation) was used to measure continuous changes of seawater temperature in the field. This battery-powered thermometer can report 6,500 consecutive data records. The initial settings, including measurement intervals and data acquisition, were set with a computer using the software platform HOBOware® (V. 2.1.1_18), ONSET Computer Corporation. For this study, measurements were taken at 30-min intervals. At this setting, about 135 days of temperature data can be recorded and the battery power could last for about 4 years. After initialization, we anchored the data loggers approximately 5 cm above the bottom at the amphioxus collection site at Argeles-sur-Mer, France, at a depth of 4 m.

Video recordings

To observe the spawning behavior of amphioxus adults, we placed ten amphioxus adults with full gonads (previously maintained at 19°C for at least 7 days) individually in 12.5-cm^2 cell culture plastic bottles (BD FalconTM Cell Culture Flasks, BD Biosciences) filled with seawater. The flasks were placed in a temperature-controlled bath and the temperature was raised to 23°C to initiate spawning 36 hr later. Throughout this 36-hr interval, the 14-hr light/10-hr dark cycle was maintained while the locomotory behavior of the amphioxus adults was recorded with an infrared video camera (Sony DCR-TRV70 MiniDV Camcorder, Sony). The infrared light permitted observation in otherwise dark conditions (a modification of the technique of Kubokawa et al. [2003]). The movements of the animals were subsequently counted and plotted on a graph.

Maintenance and spawning of amphioxus in a laboratory without running seawater

Amphioxus adults collected at Argeles-sur-Mer, France, were transported both to the University of Barcelona, Spain and to the University of Genova, Italy, where culturing and spawning were performed. At the University of Barcelona, Spain, the amphioxus were kept at 17°C in 5-L glass tanks with continuous aeration in filtered seawater obtained from the Laboratoire Arago, Banyuls-sur-Mer, France. Seawater was changed daily and animals were kept in an artificial 14-hr light/10-hr dark cycle in phase with the natural light cycle. Animals were fed with Marine Snow food supplemented with 15.000–40.000 cells/mL/feeding of mixed algae (1:1:1 parts of *D. tertiolecta*, *I. galbana* and *T. suecica*). Spawning was induced by a 36-hr thermal shock at 21°C without aeration. At the University of Genova, Italy, the amphioxus were kept at 19°C in 50-L glass tanks with filtered seawater obtained locally in a 14-hr light/10-hr dark artificial cycle. Animals were fed according to Fuentes et al. (2004). Spawning was induced by a 36-hr thermal shock at 23°C without aeration.

Inversion of the day/night cycle

Amphioxus adults with full gonads were collected in Argeles-sur-Mer, France, and brought to the University of Barcelona, Spain. Adults were kept at 17°C to prevent temperature stress and undesired spawning events. Starting with their arrival in the laboratory, the amphioxus adults were exposed to a 14-hr light/10-hr dark cycle that was phase shifted relative to the natural photoperiod such that the light period started at 10 PM and ended at noon. After 3 weeks on this phase-shifted light regime, the animals were induced to spawn by the temperature shock method of Fuentes et al. (2004).

DNA staining

Embryos were fixed for 10 min in a methanol-glycerol solution (75:25 v:v) containing 0.1 μg/mL 4,6-diamidino-2-phenylindole (DAPI) (4,6-diamidino-2-phenylindole, stock concentration was at 2.5 μg/mL). The suspension was further completed with 1 volume of phosphate-buffered saline

glycerol (50:50 v:v). The embryos were finally mounted on pre-washed glass plates. Cells were observed with a fluorescence microscope (the excitation wavelength was 350 nm and the emission wavelength was 460 nm).

RESULTS

Natural behavior

The breeding season of *B. lanceolatum* has been reported to occur from May to July (Fuentes et al., 2004). We have previously described the spawning behavior of *B. lanceolatum* during 2003 spawning season (Fuentes et al., 2004) (Fig. 1A). Here, we have expanded our observations on the spawning behavior to the 2004 and 2005 breeding seasons (Fig. 1B and C). We also report the water temperature at the collection site (for the 2004 season) (Fig. 1B). In our previous studies, we have classified living amphioxus into five ripeness categories depending on the size of their gonads (Fuentes et al., 2004). However, our subsequent observations with captive animals show that even animals with small gonads are able to spawn (data

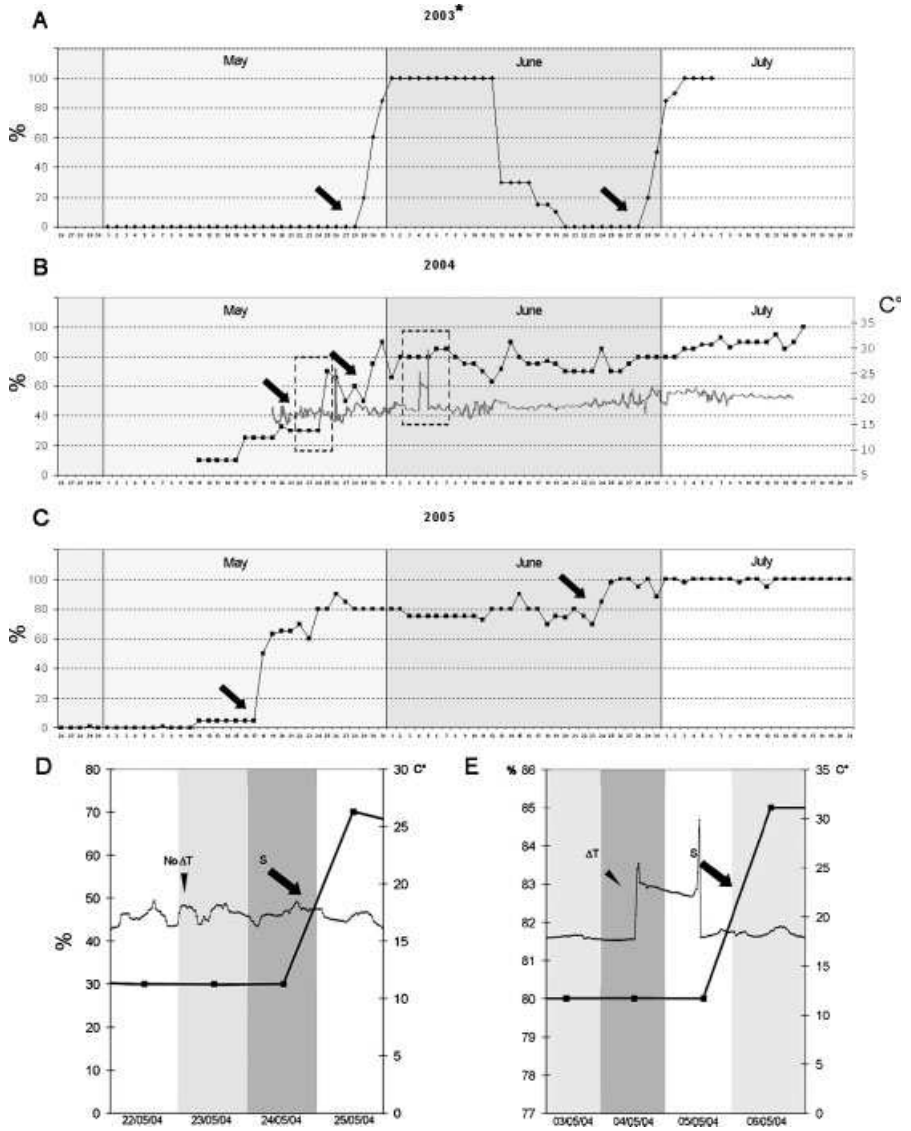


Fig. 1. Graphical representation of the percentage of empty animals at Argeles-sur-Mer between late April and July (**A**) 2003, (**B**) 2004 and (**C**) 2005. In (**B**), the water temperature at the collection site in Argeles-sur-Mer is also shown. Arrows represent spawning days, arrowheads in (**D**) and (**E**) represent time points where a change in water temperature would be required to induce the spawning. (**D**) and (**E**) show, respectively, the spawning on May 25, 2004, and June 5, 2004. While the temperature change on June 4, 2004 is followed by a small variation of the percentage of empty animals on June 5, 2004, the variation of the percentage of empty animals recorded on May 25, 2004 is not preceded by a temperature change.

not shown). Thus, here we classify living animals into only two ripeness categories, namely empty (post-spawning stage) and full (i.e. all animals with visible gonads). We collected at least 50 animals per day during May, June and July of 2004 and 2005 and classified them into these two ripeness categories. From this classification, we calculated the percentage of empty animals in the field (Fig. 1).

The spawning behavior of different species of amphioxus shows conspicuous variability. Thus, breeding in *B. belcheri* is limited to a few contiguous days per year (Wu et al., '94). In *B. floridae*, multiple synchronous spawnings within one breeding season have been observed with a 1–2-week interval between two consecutive spawnings (Stokes and Holland, '96). For *B. lanceolatum*, in the 2003 breeding season, we have shown that ripe animals spawn with synchronous episodes with a 1-month interval between them (Fuentes et al., 2004) (Fig. 1A). Our present data from the 2004 and 2005 breeding seasons show that the amount of empty animals in the population of *B. lanceolatum* in Argeles-sur-Mer has increased progressively during these two breeding seasons. In 2004, this progression showed two clear spawning events at the end of May (25 and 30 May), during which 40 and 20%, respectively, of the animals spawned (Fig. 1B, arrows). In 2005, we observed 2 days with a clear change in the percentage of empty animals: on 18 May about 60% spawned, and on 24 June, about 20% spawned (Fig. 1C, arrows). The percentage of empty animals between 18 May and 24 June remains constant. Whether amphioxus does not spawn during these days or some spawn while others fill their gonads thus maintaining the whole percentage of empty animals constant is unknown. These results show clearly that the breeding season of *B. lanceolatum* spans from May to July. The first and the last spawning days of the season however vary between early to late May and late June to early July, respectively. A second major observation is that the natural spawning behavior of *B. lanceolatum* is highly variable. In some years, spawnings are synchronous with an interval of about 1 month between two consecutive spawnings (e.g. in 2003), whereas in other years the whole population spawns progressively (e.g. in 2004) or exhibits a behavior in between these two extremes (e.g. in 2005). The physicochemical and/or biological factors that are responsible for these different spawning behaviors remain unknown and further studies both in the field and in captivity will be necessary to understand how this natural spawning behavior is regulated and controlled.

As a first step to understand some of these factors, we tried to characterize the role of temperature variations in the spawning behavior of amphioxus in the field. We have previously reported that a temperature change can induce spawning of amphioxus adults in captivity (Fuentes et al., 2004). During the 2004 breeding season, we measured the water temperature at the collection site every 30 min from late May to mid-July to study the possible correlation between temperature and natural spawnings of *B. lanceolatum* (Fig. 1B). The spawning of 25 May (Fig. 1B and D) was not preceded by an obvious temperature change. However, two temperature changes occurred several days later, a temperature spike on 26 May (from 16 to 25°C) and 4 June (from 18 to 22°C). The temperature spike on 26 May lasts only a few hours and does not seem to have an effect on the spawning behavior 36 hr later. The temperature change on 4 June, lasts 24 hr and precedes a small change in the percentage of empty animals on 6 June that could be indicative of a small natural spawning event (Fig. 1B and E). These observations suggest that even if temperature played a role in the natural spawning behavior of *B. lanceolatum*, it is not the only and probably not the major environmental variable controlling the spawning behavior of the European amphioxus.

Spawning behavior in the laboratory

The impact of temperature on spawning of *B. lanceolatum* in captivity has already been shown (Stach, '99; Fuentes et al., 2004). A constant temperature change (from 18 to 23°C) initiated at 6 PM induces the natural spawning of mature animals about 36 hr later. Using this technique regularly during the 2004 and 2005 breeding seasons, we have obtained an average of spawning animals of 60 and 70%, respectively (Fig. 2). These spawnings were independent of the major natural spawning days in the field (Figs. 1 and 2), thus demonstrating that this temperature change method is highly efficient and reproducible and is sufficient for inducing spawning in mature *B. lanceolatum* adults.

In our previous work, we have also shown that animals in captivity by feeding them with a mixture of different algae can mature their gonads. Moreover, we have demonstrated that these animals can subsequently spawn in captivity after changing the water temperature (Fuentes et al., 2004). However, the animals develop their

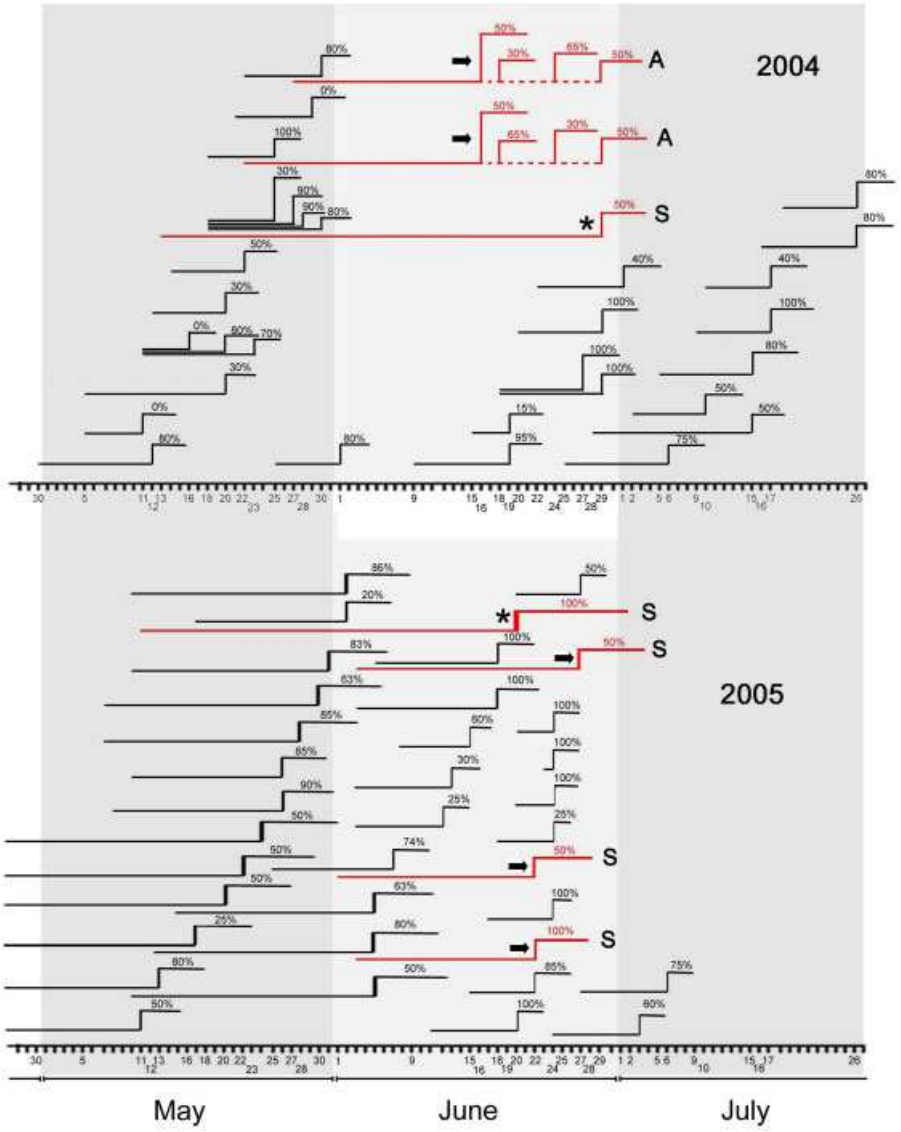


Fig. 2. Induction of spawning in laboratory-maintained amphioxus by temperature shock in 2004 and 2005. Each horizontal line represents a single amphioxus culture maintained at a constant temperature of 18°C. Horizontal lines always span from the first day of each amphioxus adult culture to the temperature shock (shown as inflexion point in the line that represents an increase in the temperature from 18 to 23°C). The percentage of animals that spawned is indicated above the line for each temperature shock. Red lines represent animals that were fed and induced to spawn after they developed the gonads in culture. Cultures where gonads developed in more than 1 month are labeled with a star and those where gonads developed in less than 1 month are highlighted with an arrow. Cultures with synchronous development of gonads are labeled with an ''S'' and those where they developed gonads asynchronously are marked with an ''A.'' These asynchronous cultures were shocked several times (represented by several inflexion points in the lines): once animals in these cultures became full, they were shocked, whereas empty animals were kept and shocked only when they had developed gonads.

gonads slower in captivity than in the field. We have repeated these feeding experiments several times during the 2004 and 2005 seasons (red lines in Fig. 2). We observed that different animal cultures develop their gonads at different rates, with some cultures needing more than 1 month to develop mature gonads (star in Fig. 2) and others developing gonads in about 20 days (arrow in Fig. 2). A second observation is that gonads develop synchronously in some cultures (i.e. 95–100% of the animals in these cultures develop full gonads at the same speed and all of them can thus be shocked together) (labeled ''S'' in Fig. 2), but asynchronously in other cultures (i.e. the pace with which the gonads are developing varies between individuals in these cultures; thus,

since only full animals are shocked, animals from these asynchronous cultures can be shocked on different days) (labeled A in Fig. 2). Since all the animals fed in captivity were subjected to the same conditions (water temperature, light, food), these differences in gonadal development might be because of biological differences (age, size, weight of the animals) between animals placed in different aquarium tanks.

To obtain the exact time course of the *B. lanceolatum* spawning event, we recorded the spawning with an infrared camera (Fig. 3A). During the 36 hr of the temperature shock, the animals were not very agitated and did not conspicuously swim around. Some swimming episodes were detected late during the first night of the temperature shock (Fig. 3B). Just after the sunset (8:45 PM) of the second temperature shock day, the animals became very active with vigorous irregular swimming occurring at two separate time points, the first one about 1 hr after the sunset and the second 1 hour thereafter (Fig. 3B).

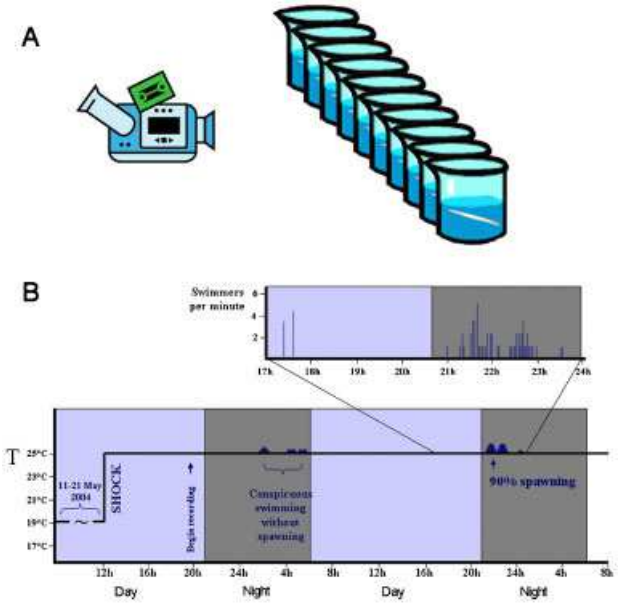


Fig. 3. Video recording of amphioxus spawning in captivity. (**A**) Relative position of the infrared camera and the individualized amphioxus during the experiment. (**B**) Graphical representation of the behavior of temperature-shock amphioxus. The blue hills on the black line represent the number of amphioxus swimmers per minute counted during 36 hr. Daytime is represented by light gray and the night by dark gray. The spawning moment is magnified showing that two main activity moments occur during the night, the first one about 1 hr after the sunset and the second one about 2 hr after the sunset. Following the first activity moment, 90% of the amphioxus spawned and 100% of them were empty after the second activity moment.

Following the first of these two episodes of high activity, 90% of the amphioxus adults spawned, whereas after the second activity episode the remaining adults emptied their gonads.

Improvement of the spawning method

As described previously (Fuentes et al., 2004), we have developed a simple method to induce spawning of mature amphioxus adults in captivity. Animals with full gonads kept at 18°C for 1 week are placed in a water bath at 23°C in the afternoon of day n. The spawning then commences at least 45 min after the sunset of day $n+1$. However, some further modifications of the method have been necessary: (i) to obtain unfertilized eggs allowing the researcher to fertilize in a synchronous way in vitro, (ii) to store and maintain amphioxus in a laboratory without running seawater, and (iii) to change the time and timing of the spawning.

(i) *Obtaining unfertilized eggs*. To obtain unfertilized eggs, we adapted a method previously used for the Florida species (i.e. separating individual mature adults in plastic cups with a small volume of filtered seawater a few hours before sunset of the spawning night) (Holland and Yu, 2004). Mature adults, temperature shocked and separated into individual plastic cups several hours before sunset, spawned normally after the sunset. To fertilize the eggs in a cup containing a female, we added a drop of sperm from the cup of a spawned male and hence obtained a culture of synchronously developing embryos.

(ii) *Storage and maintenance of amphioxus in a laboratory without running seawater*. Amphioxus adults were kept in laboratories without running seawater (e.g. University of Barcelona and University of Genova) for multiple months and spawning of these animals was successfully induced in both places using the temperature change method (see Materials and Methods). Percentage of spawning was similar in Barcelona, Genova and the Laboratoire Arago, France.

(iii) *Modification of the spawning time*. It is well known that amphioxus adults spawn after sunset and that this spawning behavior is not favorable for numerous types of in vivo experiments that can be performed on developing amphioxus. Thus, we have tried to alter the timing of the spawning by manipulating the light/dark cycles to obtain unfertilized eggs during the day. A phase shifted light/dark cycle with 14-hr light/10-hr dark was set up as follows: light 10:00 PM–12:00 PM, dark

12:00 PM–10:00 PM. Subsequently, the animals were maintained 3 weeks under these conditions before inducing spawning with the temperature change method. The continuous temperature shock was initiated at 10 PM of day *n* and about 60% of the animals spawned 36 hr later, at about 1 PM. The percentage of spawning is comparable with the percentage obtained from animals that were kept in a natural light cycle.

Developmental timing

In developmental studies of amphioxus it is normal to refer to a given developmental stage by using the time that has passed since fertilization. However, biological and physical factors alter this developmental timing. For example, different amphioxus species might actually develop at different rates at the same temperature. We have compared the developmental timing of the Florida amphioxus (*B. floridae*) with that of the European amphioxus (*B. lanceolatum*) to facilitate future comparisons in developmental studies. This comparison (Table 1, Fig. 4) from fertilization to metamorphosis shows that development of both species at a similar temperature (23–25°C) is comparable (Table 1, Fig. 4). However, the Florida species will develop much faster at a higher temperature (30°C), where embryos and larvae of the European species will die. In contrast, the European species can develop at lower temperatures than the Florida amphioxus, which does not develop normally if the temperature drops below 20°C (Fig. 4).

As a preliminary step toward understanding the micro-structural events controlling cell divisions in amphioxus, following the fertilization of *B. lanceolatum* eggs, we have precisely monitored the progression of the first cell division (at 19°C) (Fig. 5). Ten minutes after fertilization the male chromatin is visible inside the cytoplasm, whereas the second polar body begins to form. Ten minutes later, the egg has left meiosis phase II and the female and male pronuclei are migrating toward each other to fuse about 30 min after fertilization. The first mitotic metaphase is observed at 40 min, anaphase at 50 min and cytokinesis is completed 60 min post-fertilization. When *B. lanceolatum* are cultured at 19°C, meiosis and first mitosis is significantly slower than in *B. floridae* embryos cultured at 24°C, where first cleavage takes place 30 min after fertilization. Unfortunately, because the *B. lanceolatum* and *B. floridae* embryos used for these experiments were allowed to develop at different temperatures (at 19 and 24°C, respectively) (Holland and Holland, '92), direct comparisons of the first cell cycles in these two species are currently impossible. Hence, to work out potential parallels and differences of the first cell cycle in *B. lanceolatum* and *B. floridae*, studies performed at similar temperatures in both species will be necessary.

Concluding remarks

About 20 years ago, the first studies with amphioxus embryos of the Florida species established a new animal model for studying the

TABLE 1. Developmental timing of two Branchiostoma species

	B. lanceolatum		*B. floridae*	
Stage	Time at 19°C	Time at 23°C	Time at 25°C	Time at 30°C
First cleavage	1 hr	ND	30 min	ND
Blastula	4 hr	3 hr	4 hr	2.5 hr
Early gastrula (onset of gastrulation)	6 hr	4 hr	4.5 hr	3.5 hr
Cup-shaped gastrula	9 hr	5–6 hr	6 hr	4.5 hr
Neurula (ciliated begins rotating inside fertilization envelope)	12–13 hr	8 hr	8.5 hr	6 hr
Hatching	14–15 hr	10 hr	9.5 hr	6.5 hr
Anterior somites visible	16 hr	11 hr	10.5 hr	7.5 hr
Late neurula (onset of muscular movement)	35–40 hr	22 hr	20 hr	12 hr
Mouth-first gill slit	50–60 hr	30–35 hr	30–32 hr	23–24 hr
2nd gill slit	10 days	3–4 days	3 days	36 hr
3rd gill slit	16–18 day	7 days	6 days	3 days
Metamorphosis	ND	45–50 days	41–49 days	26–32 days

The developmental timing for *B. floridae* has been obtained from Holland and Holland ('92) and Holland and Yu (2004).
ND, not determined.

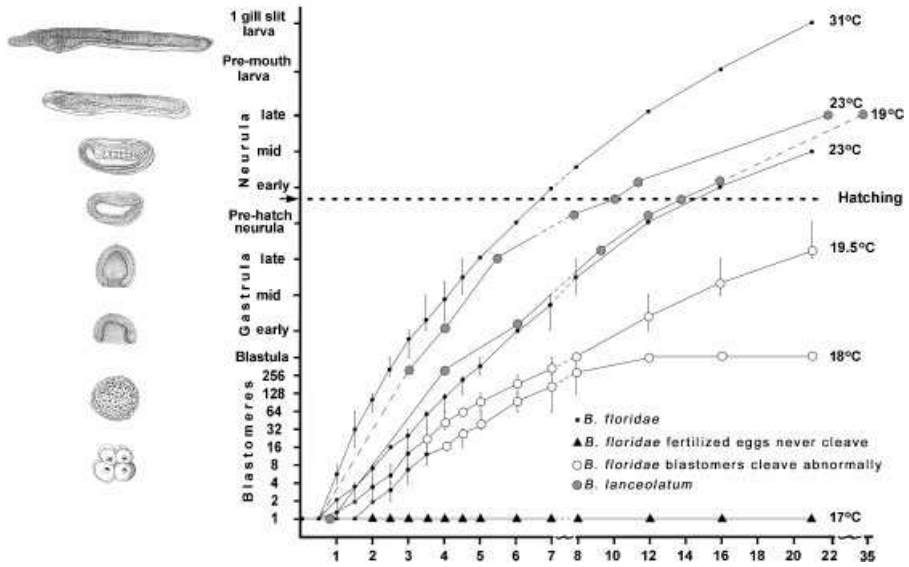


Fig. 4. Growth curves of developing *Branchiostoma floridae* (black dots, black triangles, empty circles) and *Branchiostoma lanceolatum* (gray circles) at different temperatures. The y-axis shows the different developmental stages of amphioxus, whereas the x-axis represents developmental time in hours. The drawings of amphioxus embryos and larvae have been modified from Conklin ('32).

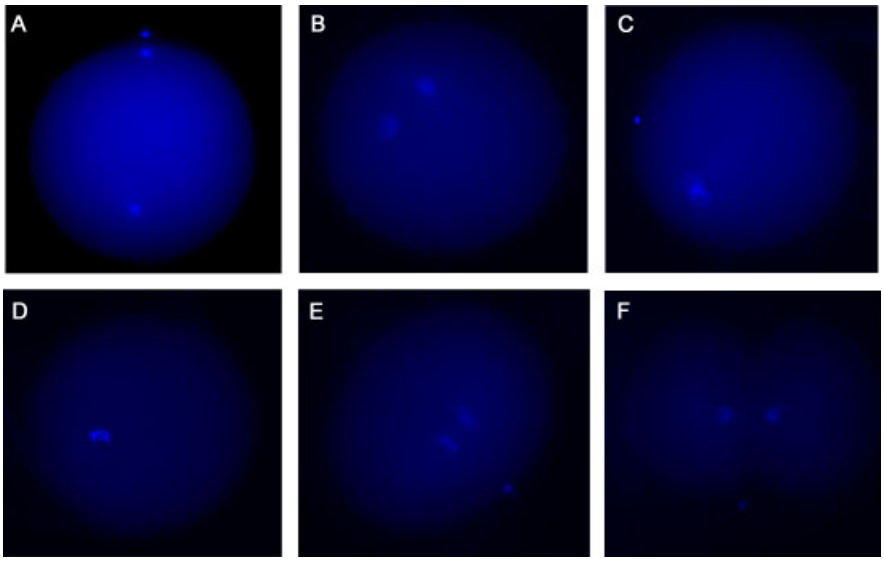


Fig. 5. DAPI-stained fertilized *Branchiostoma lanceolatum* eggs. (**A**) 10 min post-fertilization, the first polar body is at the animal pole inside the fertilization membrane, whereas the egg is completing meiosis phase II and the sperm has penetrated into the cytoplasm. (**B**) By 20 min, male and female chromatin has decondensed and both pronuclei are migrating. (**C**) Syngamy is at 30 min after fertilization. (**D**) By 40 min, the zygote is in metaphase. (**E**) Anaphase is observed at 50 min. (**F**) By 60 min, cytokinesis is taking place.

evolution of vertebrates from an invertebrate ancestor. Since then, steady progress has been made by the scientific community in developing new resources and techniques for this invertebrate chordate, including the sequencing of the amphioxus genome. Nonetheless, continuous laboratory cultures of amphioxus adults have not yet been developed. Our studies show for the first time that amphioxus adults can be maintained in a laboratory without running seawater and that these adults can be used to obtain embryos on a daily basis. This special knowledge will provide the

basis for the development of new experimental approaches and will certainly make amphioxus a much more attractive animal model for studies of the evolution of the chordate and vertebrate body plan.

ACKNOWLEDGMENTS

We are very grateful to Neus Abella for fruitful discussions about the day/night cycle inversion experiments. We also thank the fishermen of the ''Nereis'' for their help with collection of adult amphioxus. HE is supported by ANR BLAN NT_NV_49. SDA is supported by a postdoctoral fellowship from the European Community's Human Potential Programme HPRN-CT-2002-00263 ''Neurogenome''. EH and JFN are supported by the European project Embryomics NEST 012916. SB is an EMBO postdoctoral fellow. We are also grateful for funding from MENRT, CNRS, ARC and the Région Rhône-Alpes as well as FPU, CIRIT, and FPI fellowships of SJD, JPA and IM respectively.

LITERATURE CITED

Abi-Rached L, Gilles A, Shiina T, Pontarotti P, Inoko H. 2002. Evidence of en bloc duplication in vertebrate genomes. Nat Genet 31:100–105.

Bourlat SJ, Juliusdottir T, Lowe CJ, Freeman R, Aronowicz J, Kirschner M, Lander ES, Thorndyke M, Nakano H, Kohn AB, Heyland A, Moroz LL, Copley RR, Telford MJ. 2006. Deuterostome phylogeny reveals monophyletic chordates and the new phylum Xenoturbellida. Nature 444:85–88.

Conklin EG. 1932. The embryology of amphioxus. J Morphol 54:69–151.

Dehal P, Boore JL. 2005. Two rounds of whole genome duplication in the ancestral vertebrate. PLoS Biol 3:e314.

Delsuc F, Brinkmann H, Chourrout D, Philippe H. 2006. Tunicates and not cephalochordates are the closest living relatives of vertebrates. Nature 439:965–968.

Escriva H, Holland ND, Gronemeyer H, Laudet V, Holland LZ. 2002. The retinoic acid signaling pathway regulates anterior/posterior patterning in the nerve cord and pharynx of amphioxus, a chordate lacking neural crest. Development 129:2905–2916.

Fuentes M, Schubert M, Dalfo D, Candiani S, Benito E, Gardenyes J, Godoy L, Moret F, Illas M, Patten I, Permanyer J, Oliveri D, Boeuf G, Falcon J, Pestarino M, Fernandez JG, Albalat R, Laudet V, Vernier P, Escriva H. 2004. Preliminary observations on the spawning conditions of the European amphioxus (*Branchiostoma lanceolatum*) in captivity. J Exp Zool B Mol Dev Evol 302: 384–391.

Garcia-Fernandez J, Holland PWH. 1994. Archetypal organization of the amphioxus Hox gene cluster. Nature 370: 563–566.

Guillard RRL, Ryther JH. 1962. Studies of marine planktonic diatoms. I. Cyclotella nana Hustedt and Detonula confervacea Cleve. Can J Microbiol 8:229–239.

Holland LZ, Holland ND. 1992. Early development in the lancelet (= Amphioxus) *Branchiostoma floridae* from sperm entry through pronuclear fusion: presence of vegetal pole plasm and lack of conspicuous ooplasmic segregation. Biol Bull 182:77–96.

Holland LZ, Yu JK. 2004. Cephalochordate (amphioxus) embryos: procurement, culture, and basic methods. Methods Cell Biol 74:195–215.

Kubokawa K, Mizuta T, Morisawa M, Azuma N. 2003. Gonadal state of wild amphioxus populations and spawning success in captive conditions during the breeding period in Japan. Zool Sci 20:889–895.

Schubert M, Escriva H, Xavier-Neto J, Laudet V. 2006. Amphioxus and tunicates as evolutionary model systems. Trends Ecol Evol 21:269–277.

Shimeld SM, Holland ND. 2005. Amphioxus molecular biology: insights into vertebrate evolution and developmental mechanisms. Can J Zool 83:90–100.

Stach T. 1999. The ontogeny of the notochord of *Branchiostoma lanceolatum*. Acta Zoologica (Stockholm) 80:25–33.

Stokes MD, Holland ND. 1996. Reproduction of the Florida lancelet (*Branchiostoma floridae*): spawning patterns and fluctuations in gonad indexes and nutritional reserves. Invert Biol 115:349–359.

Wu XH, Zhang SC, Wang YY, Zhang BL, Qu YM, Jiang XJ. 1994. Laboratory observation on spawning, fecundity and larval development of amphioxus (*Branchiostoma belcheri tsingtauense*). Chin J Oceanol Limnol 12:289–294.

Available online at www.sciencedirect.com
ScienceDirect
Molecular Phylogenetics and Evolution 47 (2008) 73–83
MOLECULAR PHYLOGENETICS AND EVOLUTION
www.elsevier.com/locate/ympev

Combined analysis of fourteen nuclear genes refines the Ursidae phylogeny

Marie Pagès [a,b], Sébastien Calvignac [a], Catherine Klein [a], Mathilde Paris [c], Sandrine Hughes [a], Catherine Hänni [a,*]

[a] *Paléogénétique et Evolution Moléculaire, Université de Lyon, Institut de Génomique Fonctionnelle, INRA, CNRS, Université Lyon 1, Ecole Normale Supérieure, 46 allée d'Italie, 69364 Lyon cedex 07, France*
[b] *Laboratoire de Paléontologie, Phylogénie et Paléobiologie, Institut des Sciences de l'Evolution de Montpellier, UMR CNRS 5554, Université Montpellier II, Place Eugène Bataillon, 34095 Montpellier Cedex 05, France*
[c] *Zoologie Moléculaire, Université de Lyon, Institut de Génomique Fonctionnelle, INRA, CNRS, Université Lyon 1, Ecole Normale Supérieure, 46 allée d'Italie, 69364 Lyon cedex 07, France*

Received 22 February 2007; revised 11 October 2007; accepted 17 October 2007
Available online 5 November 2007

Abstract

Despite numerous studies, questions remain about the evolutionary history of Ursidae and additional independent genetic markers were needed to elucidate these ambiguities. For this purpose, we sequenced ten nuclear genes for all the eight extant bear species. By combining these new sequences with those of four other recently published nuclear markers, we provide new insights into the phylogenetic relationships of the Ursidae family members. The hypothesis that the giant panda was the first species to diverge among ursids is definitively confirmed and the precise branching order within the *Ursus* genus is clarified for the first time. Moreover, our analyses indicate that the American and the Asiatic black bears do not cluster as sister taxa, as had been previously hypothesised. Sun and sloth bears clearly appear as the most basal ursine species but uncertainties about their exact relationships remain. Since our larger dataset did not enable us to clarify this last question, identifying rare genomic changes in bear genomes could be a promising solution for further studies.

Keywords: Molecular phylogeny; Nuclear DNA; Hibernation; Ursidae

1. Introduction

Among Carnivores, the family of Ursidae comprises eight extant species, including the giant panda (*Ailuropoda melanoleuca*), the spectacled bear (*Tremarctos ornatus*) and six ursine species: the polar bear (*Ursus maritimus*), the brown bear (*Ursus arctos*), the American black bear (*Ursus americanus*), the Asiatic black bear (*Ursus thibetanus*), the sun bear (*Helarctos malayanus*) and the sloth bear (*Melursus ursinus*). Although the belonging of the giant panda to the Ursidae is not a riddle anymore (Davis, 1964; O'Brien et al., 1985; Van Valen, 1986; Mayr, 1986), the phylogenetic relationships among ursids still remain unclear. The fact that ursine species underwent a rapid diversification 5 Myr[1] ago (Kurtén, 1968; Wayne et al., 1991; Waits et al., 1999) could explain why resolving this phylogeny represents a tricky challenge. Despite previous molecular investigations, mainly based on mitochondrial DNA (Zhang and Ryder, 1994; Vrana et al., 1994; Talbot and Shields, 1996a; Waits et al., 1999; Yu et al., 2004; Fulton and Strobeck, 2006), three thorny questions are still pending: (Q1) what is the exact relationship between the giant panda and the spectacled bear at the base of the Ursidae

* Corresponding author. Address: Institut de Génomique Fonctionnelle, INRA, CNRS, Université Lyon 1, Ecole Normale Supérieure, F-69364, France. Fax: +33 0 4 72 72 80 80.
E-mail address: Catherine.Hanni@ens-lyon.fr (C. Hänni).

[1] *Abbreviations used:* Myr, million years; kb, kilobases.

1055-7903/$ - see front matter © 2007 Elsevier Inc. All rights reserved.
doi:10.1016/j.ympev.2007.10.019

radiation? (Q2) is the supposed sistership of the two black bears, *Ursus americanus* and *Ursus thibetanus* true?; (Q3) what is the branching order among the six ursine species?

A possible explanation of the inability to resolve these uncertainties is that highly variable molecular markers such as mitochondrial ones are too homoplasic (Talbot and Shields, 1996a; Waits et al., 1999). On the contrary, slowly evolving sequences, such as the only two nuclear markers available for all Ursidae species (transthyretin intron 1 and interphotoreceptor retinoid binding protein exon 1), are too poorly informative to allow robust phylogenetic inferences (Yu et al., 2004). However, concatenating numerous slowly evolving sequences is likely to overcome the lack of phylogenetic signal (Gadagkar et al., 2005). This strategy has already solved problematic relationships inside recently diversified families of other carnivores (see Bardeleben et al., 2005). Y-linked markers could in particular be valuable targets. Under the hypothesis of a male driven evolution, Y-linked genetic markers are supposed to be less variable than mitochondrial markers (and so, less homoplasic) but more variable than X-linked or autosomal genes (Haldane, 1947; Lessells, 1997; Pecon-Slattery and O'Brien, 1998). These markers already proved to be powerful to reconstruct phylogenetic trees concerning mammals (Pecon-Slattery and O'Brien, 1998; Nishida et al., 2003; Moreira, 2002; Pecon-Slattery et al., 2004).

In this paper, we provide the sequences of nine additional autosomal gene fragments and one Y-linked marker. Phylogenetic analyses performed on a concatenated dataset composed of these new sequences (nearly 6 kb) and four other recently published markers (nearly 4 kb) (two Y-linked and two autosomal ones; Yu et al., 2004; Pagés et al., submitted for publication) allow us to address each of the three still unresolved questions (Q1, Q2 and Q3) concerning the evolutionary history of the Ursidae.

2. Materials and methods

2.1. Samples

Samples were selected from a variety of sources including blood, tissues, and hairs from free ranging bears or individuals from zoos. A total number of 9 female and 13 male bears, representing all the eight species, were sampled (Table 1).

2.2. Choice of the nuclear genes for phylogenetic analyses

First, we selected three nuclear genetic markers that were recently used to resolve phylogenies of carnivore families that diversified at the same time as Ursidae (15–20 Myr, Waits et al., 1999): TRSP (selenocysteine tRNA) and FES (feline sarcoma protooncogene) were proved to be valuable for the resolution of the Canidae phylogeny (Canidae, 12 Myr, Bardeleben et al., 2005), and UBE1Y (ubiquitin activating enzyme E1 on Y) for the Felidae phylogeny (12–15 Myr, Pecon-Slattery et al., 2004). We also chose to sequence another marker, the vWF (von Willebrand Factor), recently used to infer the phylogeny of a rodent family (Echimyidae; Galewski et al., 2005).

Table 1
Ursid samples used in this study

Subfamily	Scientific name	Common name	Sex	Number of samples	Type	Sample source
Ailuropodinae	*Ailuropoda melanoleuca*	Giant panda	Male	1	Tissue	Collection of tissues conserved in alcohol, MNHN (France)
Tremarctinae	*Tremarctos ornatus*	Spectacled bear	Male	1	Hairs	Parc de la Tête d'Or (Lyon, France)
Ursinae			Female	1	Tissue	F. Catzeflis (University of Montpellier, France)
	Ursus thibethanus	Asiatic black bear	Male	1	Hairs and tissue	Réserve Africaine de Sigean (France)
	Melursus ursinus	Sloth bear	Male	1	Hairs	Moscow Zoo (Russia)
			Male	1	Hairs	Leipzig Zoo (Germany)
			Female	1	Hairs	Leipzig Zoo (Germany)
	Helarctos malayanus	Sun bear	Male	1	Hairs	Parc zoologique de Paris (France)
			Female	1	Hairs	Zoo de Saint Martin La Plaine (France)
	Ursus americanus	American black bear	Male	1	Tissue	F. Catzeflis (University of Montpellier, France)
			Male	1	DNA extract	I. Delisle (University of Alberta, Canada)
			Female	1	Hairs and blood	Zoo de Peaugres (France)
			Female	1	DNA extract	I. Delisle (University of Alberta, Canada)
	Ursus arctos	Brown bear	Male	1	DNA extract	I. Delisle (University of Alberta, Canada)
			Male	2	DNA extract	P. Taberlet (LECA, Grenoble, France)
			Unknown	1	Hairs and tissue	P. Taberlet (LECA, Grenoble, France)
			Female	3	DNA extract	P. Taberlet (LECA, Grenoble, France)
	Ursus maritimus	Polar bear	Male	1	Hairs	Zoo de la Palmyre (France)
			Male	1	DNA extract	I. Delisle (University of Alberta, Canada)
			Female	2	Hairs	Zoo de la Palmyre (France)
			Female	1	DNA extract	I. Delisle (University of Alberta, Canada)

Taxonomic denomination follows the classification of Wozencraft (1993).

Six supplemental genes, all involved into the thyroid hormone pathway, were targeted: DI1 (iodothyronine deiodinase type I), DI2 (iodothyronine deiodinase type II), TG (thyroglobulin), SIS (sodium iodide symporter), TRH (thyrotropin-releasing hormone), TSHβ (thyroid stimulating hormone beta-subunit). These markers were initially sequenced to determine whether hibernating and non-hibernating bear species had been submitted to different selective pressures on this pathway. As no positive selection was detected and since these markers appeared to be under neutral evolution (data not shown), we used these six nuclear markers for a phylogenetic purpose.

Finally, four nuclear markers already available for all ursid species were used in this study: SRY (sex determining region of the Y chromosome), ZFY (zinc finger protein on Y), TTR (transthyretin) and IRBP (interphotoreceptor retinoid binding protein) (Yu et al., 2004; Pagés et al., submitted for publication).

In this way, a total of 14 nuclear genes were considered for the phylogenetic analyses (see Table 2 for accession numbers).

2.3. DNA isolation, PCR amplification and sequencing

To avoid contaminations, pre-amplification procedures and post-amplification analyses were performed in independent rooms. DNA was extracted from blood and hair samples with QiAmp DNA Mini Kit (Qiagen) and from tissue with DNEasy Tissue Kit (Qiagen), both in accordance with the manufacturer's instructions.

Primer sets used to amplify the UBE1Y, TRSP, FES and vWF genes were found in the literature (Table 3). UBE1Y primers were tested on female bear samples in order to ensure their specificity for the Y copy. All amplifications were carried out in 25 μL containing about 30 ng of DNA extract, 0.2 mg/mL BSA (Roche, 1 mg/mL), 250 μM of each dNTP, 0.2 μM of each primer, 1 unit of Perkin Elmer Gold Taq polymerase (Applied Biosystems), 2.5 μL of 10× buffer, 1.5 mM of MgCl2. Cycling conditions were as follows: one activation step at 94 °C for 5 min followed by 40 cycles of denaturation at 94 °C for 30 s, annealing at 48–60 °C depending on the primers (Table 3) for 30 s, elongation at 72 °C for 45 s–1′30 min depending on the length of the target, and a final extension at 72 °C for 7 min.

To amplify the genes implicated in the thyroid pathway, we used a semi-nested PCR approach. For each fragment, two sets of degenerate primers were designed from conserved regions determined on the alignment of orthologous sequences of other mammals available in databank (Table 3). Reaction mixes were identical to that mentioned above. Two successive touch-down PCR were performed. To amplify each gene for each species, two different touch-down thermal cycle programs were used. It consisted of seven steps described in the Supplementary Data S1.

When PCR products proved to be difficult to sequence, they were cloned by using Topo TA Cloning for Sequencing kit (Invitrogen) and several clones were analysed. Clones and PCR products were sequenced by a service provider (Genome Express, France). All the sequences generated in this study were deposited in EMBL under the Accession Nos. AM748289 to AM748384 (Table 2).

2.4. Phylogenetic analyses

2.4.1. Sequence analyses

The nuclear sequences generated in this study and the already published ones (Table 2) were aligned by eye using

Table 2
Accession numbers of nuclear gene sequences used in this study

	Ursus maritimus	*Ursus arctos*	*Ursus americanus*	*Helarctos malayanus*	*Melursus ursinus*	*Ursus thibetanus*	*Tremarctos ornatus*	*Ailuropoda melanoleuca*	*Canis familiaris*
vWF	AM748289[d]	AM748290[d]	AM748291[d]	AM748292[d]	AM748293[d]	AM748294[d]	AM748295[d]	AM748296[d]	AF099154
ZFY	AM748297[c]	AM748298[c]	AM748299[c]	AM748300[c]	AM748301[c]	AM748302[c]	AM748303[c]	AM748304[c]	AF393756
SRY	AM748305[c]	AM748306[c]	AM748307[c]	AM748308[c]	AM748309[c]	AM748310[c]	AM748311[c]	AM748312[c]	AF107021[d]
IRBP	AY303843[b]	AY303842[b]	AY303837[b]	AY303839[b]	AY303838[b]	AY303841[b]	AY303840[b]	AY303836[b]	XM_546201
TTR	AY303848[b]	AF039741[a]	AY303844[b]	AY303846[b]	AY303845[b]	AY303847[b]	AF039740[a]	AF039738[a]	AY50579
TRSP	AM748313[d]	AM748314[d]	AM748315[d]	AM748316[d]	AM748317[d]	AM748318[d]	AM748319[d]	AM748320[d]	AY609084
FES	M748321[d]	AM748322[d]	AM748323[d]	AM748324[d]	AM748325[d]	AM748326[d]	AM748327[d]	AM748328[d]	AY885365
UBE1Y	AM748329[d]	AM748330[d]	AM748331[d]	AM748332[d]	AM748333[d]	AM748334[d]	AM748335[d]	AM748336[d]	
DI1	AM748337[d]	AM748338[d]	AM748339[d]	AM748340[d]	AM748341[d]	AM748342[d]	AM748343[d]	AM748344[d]	XM_536701
DI2	AM748345[d]	AM748346[d]	AM748347[d]	AM748348[d]	AM748349[d]	AM748350[d]	AM748351[d]	AM748352[d]	ENSCAFT00000027378
TG	AM748353[d]	AM748354[d]	AM748355[d]	AM748356[d]	AM748357[d]	AM748358[d]	AM748359[d]	AM748360[d]	ENSCAFT00000001727
SIS	AM748361[d]	AM748362[d]	AM748363[d]	AM748364[d]	AM748365[d]	AM748366[d]	AM748367[d]	AM748368[d]	XM_541946
TSH beta	AM748369[d]	AM748370[d]	AM748371[d]	AM748372[d]	AM748373[d]	AM748374[d]	AM748375[d]	AM748376[d]	U15644
TRH	AM748377[d]	AM748378[d]	AM748379[d]	AM748380[d]	AM748381[d]	AM748382[d]	AM748383[d]	AM748384[d]	XM_541757

Missing data: *Canis familiaris*, UBEIY, SIS INTRON 6, TSH*β* intron 1; *Ailuropoda melanoleuca*, UBE1Y intron 18.

[a] Flynn and Nedbal, 1998.
[b] Yu et al., 2004.
[c] Pagès et al., submitted for publication.
[d] This study.

Table 3
Abbreviation, name, sequence and original reference of primers used in this study

Designation	Gene name and amplified region	Nucleotide sequence 5′ → 3′	Annealing temperature	Original publication
vWF	von Willebrand Factor gene exon 28			
vWF-A		CTGTGATGGTGTCAACCTCACCTGTGAAGCCTG	60 °C	Porter et al., 1996
vWF-B		TCGGGGGAGCGTCTCAAAGTCCTGGATGA		
TRSP	Selenocyteine tRNA gene 5′ gene-flanking and tRNA coding regions			
TRSP-1F		GGGCTTCTGAAAGCCGACTT	50 °C	Bardeleben et al., 2005
TRSP-1R		CCGCCCGAAAGGTGGAATTG		
FES	Feline sarcoma protooncogene exon 13–intron 13–exon 14			
FESF		GGGGAACTTTGGCGAAGTGTT	50 °C	Venta et al., 1996
FESR		TCCATGACGATGTAGATGGG		
UBE1Y	Ubiquitin activating enzyme E1 on Y intron 17–exon 18–intron 18			
UBEF		GCTCTTCAAGCAGTCAGCTGAA	48 °C	Chang and Li, 1995
UBER		ATCCAGATGTAGGGG		
UBE-exon-forl[a]		AGACTCCAAKTTTGTGGAACG		This study
UBE-exon-revl[a]		TGKTCAGGTGGAAAGTTGTGC		
DI1[c]	Iodothyronine deiodinase type I exon 1			
DIl_forT2		CCCDGYYAGYGCTGTGG		
DIl_forl3		TCTGGGTSCTCTTKSAGGT		
Dllrevll		TCCCARATGYTGCACCTCTG	TC	This study
DIl_revl2		SYKGACCACVGGRCAGT		
DIlurs-Fl[b]		GGCCATGCACGTGGCCG		This study
DIlurs-Rl[b]		TCCTGAGAGCYGGACCAC		
DI2[c]	Iodothyronine deiodinase type II exon 2			
DI2-F21		GTGAAAYTGGGTGARGATGC		
DI2-F22		CCCAATTCCMGYGTGGTGC	TC	This study
DI2-R21		CTCTTGCTGAAATTCTTCTCC		
DI2-R22		GCCAAYGCCGGACTTCTTG		
TG[c]	Thyroglobulin exon 9			
TG-F91		STCMGARAGGMRGCAGGC		
TG-F92		GMCCTCVGGCTACTTCAG	TC	This study
TG-R91		SRCTTCGAGTTCCAGGAAT		
TG-R92		CACTCTGAGTTRAAGCACTG		
SIS[c]	Sodium iodide symporter exon 6–intron 6–exon 7			
SISF61		CYKGACCCGMGGMGCCG		
SIS_F62		GCCGCTAYACMTTCTGGAC	TC	This study
SISR71		BVARSAKRGGGTCRCAGTC		
SIS_R72		CTGGTCCGGGGCAGAGAT		
TRH[c]	Thyrotropin-releasing hormone exon 2			
TRH_for2		GTCAGCATCCWGGCAAAAG		
TRH_for22		GGYTCKCCAARCGTCAGC	TC	This study
TRH_rev2		CCTCCAGSGGYTCSCTG		
TRH_rev22		TGVCYYCKRCTCAGGTCA		
TSHB[c]	Thyroid stimulating hormone beta-subunit complete gene (exon 1–intron 1–exon 2)			
TSHb-for11		ACTGCTMYCTWYYTGATGTC		
TSHb-forl2		VTTTTTGGCCTDGCATGTGG	TC	This study
TSHb-rev21		CYAMCABATMRGACTTCTG		
TSHb-rev22		TTGYYTTRATGGCYTCATGT		

TC, touch down. (a) Used to amplify the exon 18 of UBE1Y of the taxa *Ailuropoda melanoleuca*; (b) the gene DI1 of the taxa *Helarctos malayanus*. (c) genes amplified by semi-nested PCR.

SEAVIEW (Galtier et al., 1996). Ambiguous regions due to nucleotide repetition were excluded from the alignment before performing the phylogenetic analyses.

To compare evolutionary rates between coding and non-coding sequences, pairwise divergence values were calculated for each part of the genes by computing uncorrected distances in PAUP v4.0b10 (Swofford, 1998) within the ursid species. GC-content at the third position for coding sequences and total GC-content for the non-coding sequences were computed using PHYLO_WIN (Galtier et al., 1996).

2.4.2. Datasets used

All the genes were concatenated and different datasets were used depending on the chosen outgroup. Outgroups were selected according to the phylogenetic issues to solve (Table 4). To address question (Q1), *Canis familiaris* was used as an outgroup for the analyses. Following this first analysis, the giant panda was then used as an outgroup to investigate other family member relationships (Q2 and Q3). Finally, since we did not succeed in amplifying the UBE1Y intron 18 for the giant panda, the spectacled bear was used as an outgroup in analyses where this intron was considered. Thus, three different datasets were generated (Table 4).

2.4.3. Phylogenetic analyses

The appropriate model of evolution was first determined for each region of each gene and then, for the three concatenated datasets, using MrModeltest 2.0 (Nylander et al., 2004), a Modeltest reduced version adapted to options available with Bayesian analysis softwares (Posada and Crandall, 1998). The selected models for each partition and dataset are shown in supplementary data S2. Maximum Likelihood (ML) analyses were performed with PHYML (Guindon and Gascuel, 2003) using the online interface http://atgc.lirmm.fr/phyml/(Guindon et al., 2005). For each analysis, the transition/transversion ratio, the proportion of invariable sites as well as the gamma distribution parameter (if necessary) were estimated and the starting tree was determined by BioNJ analysis of the datasets (default settings). Using optimization options, 1000

Table 4
Schematic representation of the three combined datasets used in this study to infer ursid phylogeny

Genes included in the datasets		Dataset 1 Ql *Out: Canis familiaris*	Dataset 2 Q2/Q3 *Out: Ailuropoda melanoleuca*	Dataset 3 Q2/Q3 *Out: Tremarctos ornatus*
vWF	exon 28	1076	1076	1076
ZFY	exon	350	397	397
SRY	5′ flanking region	417	438	438
	coding region	662	666	666
	3′ flanking region	152	167	168
IRBP	exon 1	1274	1274	1274
TTR	intron 1	946	984	987
TRSP	5′ flanking region	287	299	299
	tRNA	66	66	66
FES	exon 13	98	98	98
	intron 13	277	294	294
	exon 14	64	64	64
UBE1Y	intron 17		206	215
	exon 18		171	172
	intron 18			807
DI1	exon 1	192	192	192
DI2	exon 2	491	491	491
TG	exon 9	742	742	742
SIS	exon 6	117	117	117
	intron 6		80	80
	exon 7	111	111	111
TSHbeta	exon 1	115	115	115
	intron 1		389	389
	exon 2	166	166	167
TRH	exon 2	360	366	366
Total	in base pairs	7963	8969	9791
	% Coding regions	73.9	68.1	62.4
	% Non-coding regions	26.1	31.9	37.6
	% Y-linked regions	19.9	22.8	29.2

Each partition is associated to an outgoup (Out). Missing data for the outgroups are shaded. The ursid homologous sequences were consequently excluded from the partition. For each gene or part of gene, the number of aligned sites is given. Percentages of coding, non-coding and Y-linked regions are indicated for each dataset.

bootstrap (Bp) replicates were performed. ML analyses were first performed independently on each locus. Since the 14 genes yielded consistent, compatible topologies, analyses of the concatenated sequences were then carried out using the three different datasets described above.

Bayesian analyses (BA) were performed using MrBayes v3.1 (Ronquist and Huelsenbeck, 2003). Four independent runs of 5,000,000 generations each were performed applying appropriate independent models of evolution to each gene. A burn-in period of 100,000 generations was determined graphically using Tracer1.2 (Rambaud and Drumond, 2003), a software that allows an easy plotting of all parameters against the number of generations. For each dataset, all runs gave similar tree topologies and posterior probability (pp) values.

Alternative topologies were finally tested for significance using the Kishino-Hasegawa test (KH test) (Kishino and Hasegawa, 1989) and the Shimodaira-Hasegawa test (SH test) (Shimodaira and Hasegawa, 1999) (RELL option, 1000 Bp replicates) in PAUP* v4.0b10 (Swofford, 1998).

3. Results and discussion

3.1. Ursidae sequences analyses

The sequences of the ten targeted genes were generated for all the ursids, except one part of the UBE1Y fragment (intron 18) missing for the *A. melanoleuca* specimen. Information concerning these sequences is given in Table 5.

Proportionally to the sequence length, the number of variable sites within Ursidae ranged from 0.6% (1/166, TSHβ exon 2) to 5.1% (34/666, SRY coding sequence) among coding sequences, and from 3.3% (10/299, 5′ flanking region of TRSP; 13/389, TSHβ intron 1) to 6.3% (13/206, UBE1Y intron 17) among non-coding sequences. Y-chromosome fragments contained the highest number of variable sites either among coding or non-coding regions (5.1% SRY coding region; 6.3% UBE1Y intron 17).

For each gene used in this study, no significant difference in GC-contents was found between ursid sequences (data not shown).

Table 5
Characteristics of the nuclear Ursidae sequences used in this study

	Coding sequences				Non-coding sequences		
	NS	D	Vs	Pi	D	Vs	Pi
ZFY	397	2	8	3			
SRY							
5′untranslated region	438				2.5	16	3
Coding region	666	3.6	34	3			
3′untranslated region	167				4.1	8	1
UBE1Y							
Intron 17	206				5.7	13	2
Exon 18	171	1.7	3	0			
*Intron 18**	807				1.9	25	2
vWF	1076	3.1	45	9			
IRBP	1274	2	38	4			
TTR	984				3.3	48	13
TRSP							
5′ flanking region	299				3.3	10	2
tRNA coding region	66	3	2	0			
FES							
Exon 13	98	2	3	1			
Intron 13	294				2.7	11	1
Exon 14	64	3.1	2	0			
DI1	192	2.6	5	0			
DI2	491	1.2	6	2			
TG	742	1.7	19	3			
SIS							
Exon 6	117	2.6	3	1			
Intron 6	80				4.9	4	2
Exon 7	111	3.6	5	2			
TSHβ							
Exon 1	115	2.6	5	0			
Intron	389				1.8	13	1
Exon 2	166	0.6	1	0			
TRH	366	2.4	10	1			
Total	8969						
Coding regions	6112	2.4	189	29			
Non coding regions	2857				3.4	148	27

NS, number of aligned sites; D, highest uncorrected distance among Ursidae representatives (%); Vs, number of variable sites among Ursidae; Pi, number of parsimony informative sites. The asterisk indicates that all these parameters were computed without giant panda sequences.

3.2. New lights on the Ursidae evolutionary history

From the analysis of the 14 nuclear genes, we obtained the phylogenetic tree of the Ursidae family given in Fig. 1. BA and ML analyses of the three datasets gave identical topologies. Most relationships are well resolved (supports of 79–100% for Bp; 1.00 for pp), the only exceptions concern the phylogenetic positions of the sloth bear, *M. ursinus* and the sun bear, *H. malayanus*. This tree allows us to discuss the three last questions about the Ursidae phylogeny (Q1, Q2, Q3).

3.2.1. Q1: Basal ursid radiation

Three alternative hypotheses were initially proposed: (1) the giant panda, *Ailuropoda melanoleuca* may be the earliest species to diverge from other ursids, (2) alternatively, this may be the spectacled bear, *Tremarctos ornatus*, (3) or a common ancestor of the giant panda and the spectacled bear, i.e. these two bears may be a sister group. All these possibilities received strong to moderate supports in previous mitochondrial DNA analyses (Talbot and Shields, 1996a; Waits et al., 1999; Table 6). In the phylogenetic reconstructions inferred from our nuclear dataset 1 (*Canis familiaris* as outgroup), we observed that the giant panda unambiguously appears as the first ursid to diverge, followed by the spectacled bear (see supplementary data S3). This is supported by the highest values of bootstrap or posterior probabilities (ML 100% Bp; BA 1.00 pp; Table 6). In addition, topology tests performed on the dataset 1 reject the two other alternative hypotheses (KH test and SH test, $P < 0.03$). These findings are congruent with the recently published molecular phylogeny of the Arctoidea (Fulton and Strobeck, 2006) based on four nuclear

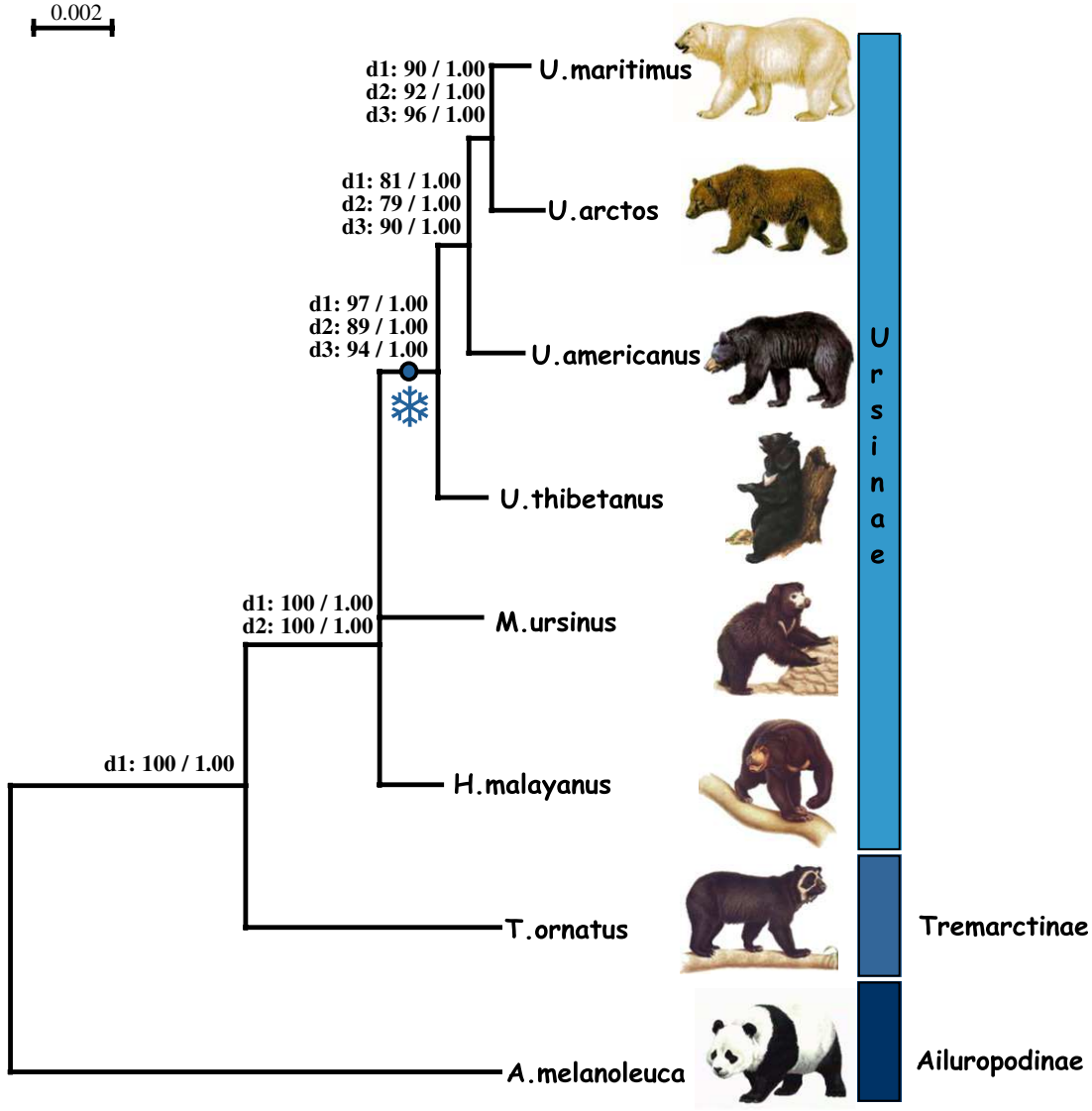


Fig. 1. Phylogenetic tree of the Ursidae family based on the analyses of 14 combined nuclear genes (dataset 2; see text) and reconstructed following Bayesian methods. BA and ML analyses of the 3 datasets gave an identical topology. Numbers above branches reflect supports (Bp/pp) obtained from the analysis of the three datasets (d1, d2, d3). The snow-flake stands for the putative acquisition of hibernating abilities by the common ancestor of the *Ursus* genus lineage.

sequence-tagged sites (STS) and IRBP exon 1 (68% Maximum Parsimony Bp, BA 0.98 pp).

3.2.2. Q2: Presumed sistership of the two black bears

The association of the two black bears, *Ursus americanus* and *Ursus thibetanus* as sister taxa was first proposed by Talbot and Shields using the entire cytochrome *b* sequences (1996a) (Table 6). However, it was shown that alternative hypotheses could not be excluded using larger mitochondrial DNA datasets (Waits et al., 1999). More recently, a sister branching of the black bears was once again suggested by the analysis of the TTR intron 1, but with a weak statistical support (40 and 52% Bp; Table 6). We show here, that topology tests performed on these datasets do not favour any of the hypotheses ($P > 0.05$; Table 7).

Our data are clearly not consistent with this sistership hypothesis. Indeed, all the three combined nuclear datasets gave identical results (Fig. 1): the American black bear clustered with the polar/brown bear clade and the Asiatic black bear is placed as the sister taxon to the polar/brown/American black bears. To test the reliability of this finding, we considered alternative hypotheses concerning Q2 (Table 7): (1) the Asiatic black bear groups with the polar/brown bear clade and the American black bear is placed as sister taxon to this former clade, and (2) the two black bears are clustered as sister taxa. KH and SH-tests show that trees where the two black bears are associated as sister taxa are significantly worse than trees where there are not clustered as sister taxa (e.g. $P < 0.05$ for KH and SH-tests based on dataset 3; Table 7). However, these tests failed to find significant differences between the alternative branching orders of the two black bears ($P > 0.05$ for KH and SH-tests based on datasets 2 and 3; Table 7). Nevertheless, the analyses of the three datasets

Table 6
Support values found in this study and in previous molecular analyses concerning the alternative hypotheses of the three unresolved ursid phylogenetic questions

		Previous published analysis					Multi nuclear genes analyses					
		Based on combined mitochondrial DNA		Based on exon (IRBP)	Based on intron (TTR)	Based on the 2 combined nuclear genes (IRBP/TTR)	Partition 1		Partition 2		Partition 3	
							Complete (7963 bp)	TTR excluded	Complete (8969 bp)	TTR excluded	Complete (9791 bp)	TTR excluded
Phylogenetic issue	alternative topologies	Talbot and Shields, 1996	Waits et al., 1999	Yu et al., 2004	Yu et al., 2004	Yu et al., 2004	This study		This study		This study	
Q1	(O, ((A.m, T.o), U))	92	—				—	—				
	(O, A.m, (T.o, U))	*71*	67/99	Not assessed	Not assessed	Not assessed	**100**/1.00	**100**/1.00	Not possible	Not possible	Not possible	Not possible
	(O, T.o, (A.m, U))	54	—				—	—				
Q2	(O, (U.am, U.thib), U′)	*58*/*62*/ 75	36/**38**		**40**/52	52/**58**/1.00	—	—	—		—	
	(O, U.thib, (U.am, U′))	—	—	Unresolved	—	—	**81**/1.00	**82**/1.00	**79**/1.00	**75**/1.00	**90**/1.00	**85**/1.00
	(O, U.am, (U.thib, U′))	—	—		—	—	—	—	—		—	
Q3	(O, (H.m, M.u)U″)			—	92/**98**	86/**86**/1.00	0.42	—	**50**		—	
	(O, H.m, (M.u, U″))	Unresolved	Unresolved	63/**64**	—	—	**55**	**71**/0.78	0.96	**62**/0.99	**50**/0.98	**75**/0.99
	(O, M.u, (H.m, U″))			—	—	—	—	—	—		—	

Support values correspond to the node highlighted in bold and represent the values observed for the best trees. Underlined numbers indicate posterior probabilities. NJ bootstrap values are in italics, ML bootstrap in bold and other numbers correspond to Maximum Parsimony bootstrap values.
Abbreviations: O, outgroup; A.m, *Ailuropoda melanoleuca*; T.o, *Tremarctos ornatus*; U, Ursinae; U.am, *Ursus americanus*; U.thib, *Ursus. thibetanus*, U′, (*Ursus maritimus, Ursus arctos*); H.m, *Helarctos malayanus*; M.u, *Melursus ursinus*; U″, *Ursus* genus representatives.
Q1: basal position of the Ursidae radiation?; Q2: sistership of the two black bears?; Q3: branching order among ursine group, i.e. positions of H.m and M.u?

Table 7
Likelihood ratio tests for the alternative hypotheses concerning phylogenetic relationships of the two black bears

Alternative topologies	IRBP Yu et al. (2004)			TTR Yu et al. (2004)			IRBP + TTR Yu et al. (2004)			Dataset 2 This study			Dataset 3 This study		
	Δ-lnL	KH-test	SH-test	Δ-lnL	KH-test	SH-test	Δ-lnL	KH-test	SH-test	Δ-lnL	KH-test	SH-test	Δ-lnL	KH-test	SH-test
((((U.mar, U.arct), U.am), U.thib), O1)	5.28	0.225	0.225	2.30	0.234	0.234	3.47	0.172	0.224	2.46	0.27	0.63	0.57	0.303	0.685
((((U.mar, U.arct), U.thib), U.am), O1)	5.28	0.225	0.225	2.30	0.234	0.234	3.47	0.172	0.224	15.63	0.095	0.117	14.05	0.088	0.092
((((U.mar, U.arct), (U.am, U.thib)), O1)	5.28	0.225	0.225	(1672)	—	—	1.29	0.275	0.577	20.28	0.035*	0.041*	16.76	0.039*	0.044*
((((U.mar, U.arct), U.am), U.thib), O2)	(1974)	—	—	2.30	0.234	0.234	2.39	0.227	0.333	(14803)	—	—	(15151)	—	—
((((U.mar, U.arct), U.thib), U.am), O2)	0	0.472	0.795	2.30	0.234	0.234	2.39	0.227	0.333	13.4	0.122	0.172	13.78	0.092	0.101
(((U.mar, U.arct), (U.am, U.thib)), O2)	0	0.347	0.790	0	0.506	0.804	(3705)	—	—	18.17	0.042*	0.067	16.57	0.042*	0.043*

Six topologies were tested representing all the possible topologies for the relationships between Ursinae. The association of brown and polar bears as sister taxa is considered as well-established. KH/SH-tests were performed on nuclear datasets provided by Yu et al., 2004 and on the two larger datasets out of the three generated in this study. Values in parenthesis correspond to the log-likelihood of the best tree; −ΔlnL, to the log-likelihood difference as compared to the best tree. An asterisk denotes that the test is significant at $P < 0.05$ (the topology is significantly worse than the best one). U.mar = *Ursus maritimus*; U.arct = *Ursus arctos*; U.am = *Ursus americanus*; U.thib = *Ursus thibetanus*; O1 = "*H. malayanus*), *M. ursinus*), *T. ornatus*), *A. melanoleuca*)"; O2 = "*M. ursinus*), *H. malayanus*), *T. ornatus*), *A. melanoleuca*)".

with two different methods of phylogenetic inferences sustain the hypothesis of the Asiatic black bear as the first black bear to diverge (Fig. 1) with strong supports (e.g. 90% Bp, 1.00 pp for the larger dataset 3; Table 6).

Consequently, with our larger nuclear datasets, the branching order inside the genus *Ursus* is clarified. Moreover, it should be noted that these analyses identified *Ursus* representatives as a monophyletic group and were in agreement on this point with phylogenetic reconstructions using nuclear TTR intron 1 (Yu et al., 2004). These findings are particularly interesting regarding the acquisition of hibernating capabilities by bears. As *Ursus* genus, including the extinct cave bear species (Hänni et al., 1994; Loreille et al., 2001), consists only in hibernating species (Stirling, 1993), their monophyly indicates that the hibernating capability likely appeared only once along the evolutionary history of the ursids (Fig. 1). Based on the current range distribution of the ursine species (Herreo et al., 1999), the acquisition of this peculiar feature by a common Asiatic ancestor could be the cause or the consequence of the colonisation by the *Ursus* representatives of the northern regions of the North hemisphere. Alternatively, hibernation could also represent a symplesiomorphic feature shared by the *Ursus* representatives, and lost by the other bear species.

3.2.3. Q3: The Ursinae radiation

The six ursine species underwent such a rapid radiation around 5 Myr ago (Kurtén, 1968; Wayne et al., 1991). Thus, determining their branching order is a difficult challenge. Molecular phylogenies based either on mitochondrial data (Zhang and Ryder, 1994; Vrana et al., 1994; Talbot and Shields, 1996a; Waits et al., 1999) or on nuclear data (Yu et al., 2004; Fulton and Strobeck, 2006) failed to clarify this problem. Up to now, the only well-established relationship within Ursinae was the sistership of brown and polar bears (Taberlet and Bouvet, 1992; Talbot and Shields, 1996a; Waits et al., 1999; Yu et al., 2004). More precisely, analyses of mitochondrial markers have concluded to closer relationships between polar bears and brown bear population of the ABC islands leading to the paraphyly of the brown bear species on mitochondrial data (Talbot and Shields, 1996b). Moreover, ancient DNA studies based on mitochondrial markers elucidated the sistership of the extinct cave bear with the brown-polar bear clade (Hänni et al., 1994; Loreille et al., 2001). We clarified above the relationships among the genus *Ursus* (Q2), so the last point to elucidate relates to the branching order of *M. ursinus* and *H. malayanus*.

Our data clearly placed these two species at the base of the ursine radiation and were excluded from the monophyletic group of *Ursus* representatives. However, ML reconstructions based on our 3 datasets do not discriminate between two of the three possible topologies given the weak Bp supports (Bp $\leqslant$55%; Table 6). Nevertheless, BA performed on datasets 2 and 3 give stronger support to

M. ursinus clustered with the representatives of the genus *Ursus* (pp values of 0.96 and 0.98).

Combined genes could lead to an unresolved topology if the genes give contradictory results. In this case, Jeffroy et al. (2006) showed that it could be better to remove the problematic genes from the analyses. In their study combining IRBP and TTR, Yu et al. (2004) observed that both genes taken separately gave contradictory results: IRBP sustained the hypothesis of *M. ursinus* clustered with the representatives of the genus *Ursus*, whereas TTR supported the *M. ursinus* and *H. malayanus* sistership (Table 6). In order to assess whether TTR is responsible for the low supports observed, we performed the same phylogenetic analyses on our three datasets but excluding this marker. In this case, a single topology is well supported: *M. ursinus* clusters with the representatives of the genus *Ursus*. The exclusion of TTR increases all the support values for this hypothesis in both ML and BA analyses (e.g. Bp from 50–75% for dataset 3; Table 6). Nevertheless, further work is required to confirm the positions of the sloth and sun bears within the ursid tree.

3.2.4. Further investigations

In this study, we resolved the main ambiguities among the phylogeny of Ursidae (Q1 and Q2) and clarified the last one (Q3). Even the combination of 14 nuclear gene fragments was not informative enough to solve definitively the relationships between all the representatives of the Ursinae sub-family (Q3). This lack of resolution is certainly due to the rapid burst of ursine species during bear evolutionary history. A solution to circumvent this difficulty could be to identify Rare Genomic Changes (RGCs) in the Ursidae genomes (Springer et al., 2004). RGCs would be perfect arbiters and will certainly be needed, as there is little chance that increasing the nuclear sequence dataset will provide any further information. Searching RGCs could be directed by two theoretical and empirical points: (1) insertions, deletions, and retrotransposon integrations occur more likely in non-coding regions (Springer et al., 2004); (2) as other restricted regions of the genome with limited recombination, the non-recombining region of the Y chromosome is expected to accumulate more retrotransposon elements (Pecon-Slattery and O'Brien, 1998; Pecon-Slattery et al., 2004). So, scanning Y-linked chromosome intron sequences to find RGCs in bear genomes could prove to be useful to overcome the lack of phylogenetic signal among ursids. We are confident that these additional independent markers will consolidate the Ursidae tree.

Acknowledgments

We thank all the institutions that gave samples for this study: the MNHN (Paris), the Mammals Research Institute of Poland and Zoos of Moscow, Leipzig, Berlin, and in France: La Palmyre, Peaugres, Saint Martin La Plaine, Vincennes, Parc de la Tête d'Or de Lyon, Réserve Africaine de Sigean and more particularly I. Delisle (University of Alberta, Canada), G. Véron (MNHN, France), F. Catzeflis (CNRS, ISEM, France), and P. Taberlet (LECA, France). We are particularly grateful to M. Duffraisse for help in the lab, to P. Chevret, L. Orlando, A. Chaumont and V. Laudet for helpful discussions and critical reading, and to J.J. Jaeger for constant support. We wish to acknowledge J. Carew for English corrections. We finally thank the CNRS, Université Lyon 1, ENS de Lyon and MNRT for financial supports. M.P. holds a fellowship from the ARC Foundation.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.ympev.2007.10.019.

References

Bardeleben, C., Moore, R.L., Wayne, R.K., 2005. A molecular phylogeny of the Canidae based on six nuclear loci. Mol. Phylogenet. Evol. 37, 815–831.

Chang, B.H., Li, W.H., 1995. Estimating the intensity of male-driven evolution in rodents by using X-linked and Y-linked Ube 1 genes and pseudogenes. J. Mol. Evol. 40, 70–77.

Davis, D., 1964. The giant panda: a morphological study of evolutionary mechanisms. Fieldiana. Zool. Mem. 3, 1–339.

Flynn, J.J., Nedbal, M.A., 1998. Phylogeny of the Carnivora (Mammalia): congruence vs incompatibility among multiple data sets. Mol. Phylogenet. Evol. 9, 414–426.

Fulton, T.L., Strobeck, C., 2006. Molecular phylogeny of the Arctoidea (Carnivora): effect of missing data on supertree and supermatrix analyses of multiple gene data sets. Mol. Phylogenet. Evol. 41, 165–181.

Gadagkar, S.R., Rosenberg, M.S., Kumar, S., 2005. Inferring species phylogenies from multiple genes: concatenated sequence tree versus consensus gene tree. J. Exp. Zoolog. B Mol. Dev. Evol. 304, 64–74.

Galewski, T., Mauffrey, J.F., Leite, Y.L., Patton, J.L., Douzery, E.J., 2005. Ecomorphological diversification among South American spiny rats (Rodentia; Echimyidae): a phylogenetic and chronological approach. Mol. Phylogenet. Evol 34, 601–615.

Galtier, N., Gouy, M., Gautier, C., 1996. SEAVIEW and PHYLO_WIN: two graphic tools for sequence alignment and molecular phylogeny. Comput. Appl. Biosci. 12, 543–548.

Guindon, S., Gascuel, O., 2003. A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. Syst. Biol. 52, 696–704.

Guindon, S., Lethiec, F., Duroux, P., Gascuel, O., 2005. PHYML Online—a web server for fast maximum likelihood-based phylogenetic inference. Nucleic Acids Res. 33, W557–W559.

Haldane, J.B.S., 1947. The mutation rate of the gene for haemophilia, and its segregation ratios in males and females. Ann. Eugen. 13, 262–271.

Hänni, C., Laudet, V., Stehelin, D., Taberlet, P., 1994. Tracking the origins of the cave bear (*Ursus spelaeus*) by mitochondrial DNA sequencing. Proc. Natl. Acad. Sci. USA 91, 12336–12340.

Herreo, S. Compiled by Servheen, C., Herrero, H., Peyton, B. and the IUCN/SSC Bear and Polar Bear Specialist Groups (1999). Introduction, In: Bears: Status Survey and Conservation Action Plan, pp. 1–6. IUCN.

Jeffroy, O., Brinkmann, H., Delsuc, F., Philippe, H., 2006. Phylogenomics: the beginning of incongruence? Trends Genet. 22, 225–231.

Kishino, H., Hasegawa, M., 1989. Evaluation of the maximum likelihood estimate of the evolutionary tree topologies from DNA sequence data, and the branching order in Hominoidea. J. Mol. Evol. 29, 170–179.

Kurtén, B., 1968. The Pleistocene Mammals of Europe. Aldine, Chicago, III.

Lessells, K., 1997. More mutations in males. Nature 390, 236–237.

Loreille, O., Orlando, L., Patou-Mathis, M., et al., 2001. Ancient DNA analysis reveals divergence of the cave bear, *Ursus spelaeus*, and brown bear, *Ursus arctos*, lineages. Curr. Biol. 11, 200–203.

Mayr, E., 1986. Uncertainty in science: is the giant panda a bear or a raccoon? Nature 323, 769–771.

Moreira, M.A., 2002. SRY evolution in Cebidae (Platyrrhini: Primates). J. Mol. Evol. 55, 92–103.

Nishida, S., Pastene, A.L., Goto, M., Koike, H., 2003. SRY gene structure and phylogeny in the cetacean species. Mammal Study 28, 57–66.

Nylander, J.A., Ronquist, F., Huelsenbeck, J.P., Nieves-Aldrey, J.L., 2004. Bayesian phylogenetic analysis of combined data. Syst. Biol. 53, 47–67.

O'Brien, S.J., Nash, W.G., Wildt, D.E., Bush, M.E., Benveniste, R.E., 1985. A molecular solution to the riddle of the giant panda's phylogeny. Nature 317, 140–144.

Pagès, M., Maudet, C., Bellemain, E., Taberlet, P., Hughes, S., Hänni, C., submitted for publication. A system for sex-determination from degraded DNA: a useful tool for paleogenetics and conservation genetics of Ursids.

Pecon-Slattery, J., O'Brien, S.J., 1998. Patterns of Y and X chromosome DNA sequence divergence during the Felidae radiation. Genetics 148, 1245–1255.

Pecon-Slattery, J., Pearks Wilkerson, A.J., Murphy, W.J., O'Brien, S.J., 2004. Phylogenetic assessment of introns and SINEs within the Y chromosome using the cat family felidae as a species tree. Mol. Biol. Evol. 21, 2299–2309.

Porter, C.A., Goodman, M., Stanhope, M.J., 1996. Evidence on mammalian phylogeny from sequences of exon 28 of the von Willebrand factor gene. Mol. Phylogenet. Evol. 5, 89–101.

Posada, D., Crandall, KA., 1998. MODELTEST: testing the model of DNA substitution. Bioinformatics 14, 817–818.

Rambaud A., Drumond A., 2003. Tracer, version v1.2.

Ronquist, F., Huelsenbeck, J.P., 2003. MrBayes 3: Bayesian phylogenetic inference under mixed models. Bioinformatics 19, 1572–1574.

Shimodaira, H., Hasegawa, M., 1999. Multiple comparisons of log-likelihoods with applications to phylogenetic inference. Mol. Biol. Evol. 16, 1114–1116.

Springer, M.S., Stanhope, M.J., Madsen, O., de Jong, W.W., 2004. Molecules consolidate the placental mammal tree. Trends Ecol. Evol. 19, 430–438.

Stirling, I., 1993. Bears-majestic creatures of the wild. Rodale Press Emmaus, Pennsylvania.

Swofford, D.L., 1998. PAUP*. Phylogenetic Analysis using Parsimony (*and other methods). Version 4. Sinauer Associates, Sunderland, Massachusetts.

Taberlet, P., Bouvet, J., 1992. Génétique de l'Ours brun des Pyénées (*Ursus arctos*): premiers résultats. CR Acad. Sci. III. 314, 15–21.

Talbot, S.L., Shields, G.F., 1996a. A phylogeny of the bears (Ursidae) inferred from complete sequences of three mitochondrial genes. Mol. Phylogenet. Evol. 5, 567–575.

Talbot, S.L., Shields, G.F., 1996b. Phylogeography of brown bears (*Ursus arctos*) of Alaska and paraphyly within the Ursidae. Mol. Phylogenet. Evol. 5, 477–494.

Van Valen, L.M., 1986. Palaeontological and molecular views of panda phylogeny. Nature 319, 428.

Venta, P.J., Brouillette, J.A., Yuzbasiyan-Gurkan, V., Brewer, G.J., 1996. Gene-specific universal mammalian sequence-tagged sites: application to the canine genome. Biochem. Genet. 34, 321–341.

Vrana, P.B., Milinkovitch, M.C., Powell, J.R., Wheeler, W.C., 1994. Higher level relationships of the arctoid Carnivora based on sequence data and "total evidence". Mol. Phylogenet. Evol. 3, 47–58.

Waits, L.P., Sullivan, J., O'Brien, S.J., Ward, R.H., 1999. Rapid radiation events in the family Ursidae indicated by likelihood phylogenetic estimation from multiple fragments of mtDNA. Mol. Phylogenet. Evol. 13, 82–92.

Wayne, R.K., Van Valkenburgh, B., O'Brien, S.J., 1991. Molecular distance and divergence time in carnivores and primates. Mol. Biol. Evol. 8, 297–319.

Wozencraft, W.C., 1993. Order Carnivora. In: Wilson, D.E., Reeder, D.M. (Eds.), Mammal Species of the World: A Taxonomic and Geographic Reference. Smithsonian Institution Press, Washington, pp. 279–349.

Yu, L., Li, Q.W., Ryder, O.A., Zhang, Y.P., 2004. Phylogeny of the bears (Ursidae) based on nuclear and mitochondrial genes. Mol. Phylogenet. Evol. 32, 480–494.

Zhang, Y., Ryder, O., 1994. Phylogenetic relationships of bears (the Ursidae) inferred from mitochondrial DNA sequences. Mol. Phylogenet. Evol. 3, 351–359.

Published online 10 April 2007

Nucleic Acids Research, 2007, Vol. 35, No. 9 ***e65***
doi:10.1093/nar/gkm152

A hybrid CMV-H1 construct improves efficiency of PEI-delivered shRNA in the mouse brain

Zahra Hassani[1,†], Jean-Christophe François[2,3,†], Gladys Alfama[1], Ghislaine Morvan Dubois[1], Mathilde Paris[4], Carine Giovannangeli[2,3] and Barbara A. Demeneix[1,*]

[1]CNRS, UMR 5166, 7 rue Cuvier CP32 Paris Cedex 05, F-75231 France; MNHN, USM 501, Département de ''Régulations, développement et diversité moléculaire'', Evolution des Régulations Endocriniennes, 7 rue Cuvier CP32 Paris Cedex 05, F-75231 France, [2]INSERM, U565, Acides nucléiques: dynamique, ciblage et fonctions biologiques, 57 rue Cuvier, CP26, Paris Cedex 05, F-75231, France, [3]MNHN, USM503, Département de ''Régulations, développement et diversité moléculaire'', Laboratoire des Régulations et dynamique des génomes, 57 rue Cuvier, CP26, Paris Cedex 5, F-75231, France; CNRS, UMR5153, Acides nucléiques : dynamique, ciblage et fonctions biologiques, 57 rue Cuvier, CP26, Paris Cedex 5, F-75231, France and [4]ENS Lyon, Laboratoire de Biologie Moléculaire de la Cellule, UMR CNRS 5161, 46 allée d'Italie, 69364 Lyon, France

Received July 26, 2006; Revised February 26, 2007; Accepted February 27, 2007

ABSTRACT

RNA-interference-driven loss of function in specific tissues *in vivo* should permit analysis of gene function in temporally and spatially defined contexts. However, delivery of efficient short hairpin RNA (shRNA) to target tissues *in vivo* remains problematic. Here, we demonstrate that efficiency of polyethylenimine (PEI)-delivered shRNA depends on the regulatory sequences used, both *in vivo* and *in vitro*. When tested *in vivo*, silencing of a luciferase target gene by shRNA produced from a hybrid construct composed of the CMV enhancer/promoter placed immediately upstream of an H1 promoter (50%) exceeds that obtained with the H1 promoter alone (20%). In contrast, in NIH 3T3 cells, the H1 promoter was more efficient than the hybrid construct (75 versus 60% inhibition of target gene expression, respectively). To test CMV-H1 shRNA efficiency against an endogenous gene *in vivo*, we used shRNA against thyroid hormone receptor α1 (TRα1). When vectorized in the mouse brain, the hybrid construct strongly derepressed CyclinD1-luciferase reporter gene expression, CyclinD1 being a negatively regulated thyroid hormone target gene. We conclude that promoter choice affects shRNA efficiency distinctly in different *in vitro* and *in vivo* situations and that a hybrid CMV-H1 construct is optimal for shRNA delivery in the mouse brain.

INTRODUCTION

RNA interference (RNAi) is now a well-described mechanism by which a double-stranded RNA (dsRNA) leads to the sequence-specific inhibition of its homologous gene. First described in plants (1), it has since been applied to numerous invertebrate and vertebrate models (2). In mammalian cells, the introduction of long dsRNAs into mammalian cells activates protein kinase PKR and RNase L, leading to an interferon response and hence to the non-specific extinction of genes resulting in cell death (3,4). This non-specific effect of long dsRNAs into mammalian cells can be bypassed by using small RNA duplexes of 19–21 nt, which are sufficient to trigger specific RNAi in mammalian cells without activating the interferon response (5).

This highly efficient and specific technology opens up a broad spectrum of experimental and therapeutic possibilities (6,7). However, the lack of suitable delivery systems for short interfering RNAs (siRNAs) *in vivo* has hampered the advance of such applications.

To date two main methods have been developed to obtain sufficient intracellular levels of RNAi for gene silencing in mammals *in vivo*. The first method involves direct vectorization of siRNA into the target cell. The alternative is to vectorize plasmids containing short hairpin RNA (shRNA) cassettes, which leads to shRNA transcription, followed by Dicer-dependent loop cleavage and siRNA production. Most often, the promoters used to direct the expression of the shRNAs are H1 (8) and U6 (9–11) or CMV (12). So far these promoter systems have been tested with a variety of viral delivery systems

*To whom correspondence should be addressed. Tel: +33 1 40 79 36 07; Fax: +33 1 40 79 36 18; Email: demeneix@mnhn.fr
†The authors wish it to be known that, in their opinion, the first and second authors contributed equally to this work.

including adenoviruses (12,13), retroviruses (14) and lentiviruses (15). However, there are only few reports of their use with non-viral delivery of shRNA-carrying plasmids, neither *in vitro* nor *in vivo*.

We previously demonstrated that cationic lipids could provide efficient delivery of siRNAs into the brain of newborn mice brains, producing >80% inhibition of an exogenous gene with only 5 pmol of siRNAs (16). However, generally speaking cationic lipoplexes are less efficient than polyplexes for plasmid vectorization *in vivo*. For instance in the adult mammalian brain, the cationic polymer polyethylenime (PEI), and particularly linear 22-kDa PEI (L-PEI), easily outclasses other lipid- or polymer-based vectors (16) [for review see (17)]. We thus chose to use PEI vectorization to evaluate the capacity of two different constructs: either H1 alone or a hybrid CMV enhancer and minimal promoter/H1 promoter construct (hereafter called the CMV/H1 construct). Each construct was tested for its capacity to direct the expression of shRNA-carrying plasmids *in vitro* and *in vivo* in the brains of newborn and adult mice.

We show that, when tested *in vitro* on two different cell lines, the efficiency of target silencing is independent of the promoter used. In contrast, when tested *in vivo* in either the developing or the adult brain, the efficiency of the shRNAs depends markedly on the promoter used to drive production. We found that the H1-shRNA construct provided only a slight decrease in luciferase target gene expression. However, surprisingly, we found that the hybrid CMV-H1 construct directing shRNA production provided a significant inhibition of the co-transfected luciferase gene in both *in vivo* situations tested.

Further, we tested the efficiency of a shRNA directed against endogenous thyroid hormone α1 (shTRα1) in cell cultures and in the newborn mouse brain. We first demonstrated that the shTRα1 efficiently inhibits the expression of endogenous TRα1 in cell culture. Then we show that the presence of the shTRα1-carrying plasmid in the newborn mouse brain strongly increases the expression of a co-transfected CyclinD1-luciferase reporter gene, CyclinD1 being a negatively regulated thyroid hormone target gene in our experimental paradigm. These data demonstrate that shRNAs under the control of a hybrid CMV-H1 construct efficiently inhibit endogenous gene expression *in vivo*, and that this technology can be used for further gene function analysis in the brain of newborn mice.

METHODS

Plasmid constructions and siRNAs

The *Renilla reniformis* (RLluc)- and *Photinus pyralis* (PPluc)-luciferases carrying plasmids are respectively pRL-CMV (Promega) and a pGL2-basic vector with a cytomegalovirus (CMV) promoter inserted in the multiple cloning site referred in the following as pGL2-CMV. The shRNA sequence directed against pGL2 luciferase (shLuc in the text) is the following: 5′CGTACGCGGAATACT TCGATTCAAGAGATCGAAGTATTCCGCGTACG3′. The H1-shLuc construction based on a pSUPERbasic backbone (8) was kindly provided by Dr A Harel-Bellan (Institut André Lwoff, Villejuif, France).

For the CMV-H1-shLuc construction, the H1-shLuc plasmid was digested by BamH1 and HindIII. The 280-bp insert containing the shLuc and the H1 promoter sequences was inserted into the luciferase-free pRL-CMV vector obtained after the digestion of pRL-CMV with PstI and XbaI. This digestion releases the *Renilla* luciferase gene sequence but conserves the full CMV enhancer and early promoter. The selected clones were sequenced using the CMV forward primer (MWG Biotech). One clone, clone 6 was found to carry both the shLuc sequence and the H1 promoter just following the CMV enhancer/early promoter (see scheme, Figure 1).

The CMV-shLuc construction was obtained by ligation of a synthetic double-stranded oligonucleotide shLuc carrying PstI and XbaI ends (Eurogentec) inside the PstI/XbaI sites of pRL-CMV.

Two 21-nt-long double-stranded siRNA targeted to firefly luciferases siLuc and siLucmt were synthesized according to the sequences previously described (5,16). For the inhibition of TRα1, we designed a 21-bp sequence directed against a specific sequence for TRα1 gene (siTRα1). The sequence of the siTRα1 was the following: siTRα1(up): 5′CGCUCUUCCUGGAGGUC UUTT3′; siTRα1(down): 5′AAGACCUCCAGGAAGA GCGTT3′. For the CMV-H1-shTRα1 construction, we designed a 100-bp palindromic sequence based on the siTRα1 sequence and containing an 8-bp loop (TCAAGAG). This 100-bp sequence was designed with Bsa1 and Not1 cohesive ends, which allowed its insertion inside the Bsa1/Not1 site of CMV-H1-shLuc plasmid where the shLuc sequence was deleted previously.

Preparation of complexes and gene transfer *in vivo*

All animal experiments were conducted in accordance with the principles and procedures described in Guidelines for Care and Use of Experimental Animals.

DNA was complexed L-PEI 22 kDa with a 6N/P charge ratio (where N = the positive charges carried by protonable amines of PEI and P = the negative charges carried by DNA phosphates). For this, 0.18 µl of L-PEI 100 mM were added per µg of DNA used. All solutions were prepared in 5% glucose.

The protocol of stereotaxic injection into the lateral ventricles of newborn mice brains was as described previously. Briefly, 2 µl of a solution containing 0.1 µg of each luciferase (i.e. pGL2-CMV and pRL-CMV) and 0.2–0.8 µg of either the shLuc-containing plasmid tested or a plasmid carrying an irrelevant shRNA under a CMV enhancer/promoter, were stereotaxically injected into the lateral ventricles of newborn mice brains. Here, 24–72 h post-transfection, the animals were dissected, the brains were removed and the dual luciferase assay was performed to detect the activity of *Photinus* (pGL2) and *Renilla* (pRL) luciferases separately. Five animals were injected per group, and the two hemispheres were analysed separately for reporter gene expression (i.e. $n = 10$ hemispheres per group). Each experiment was performed at

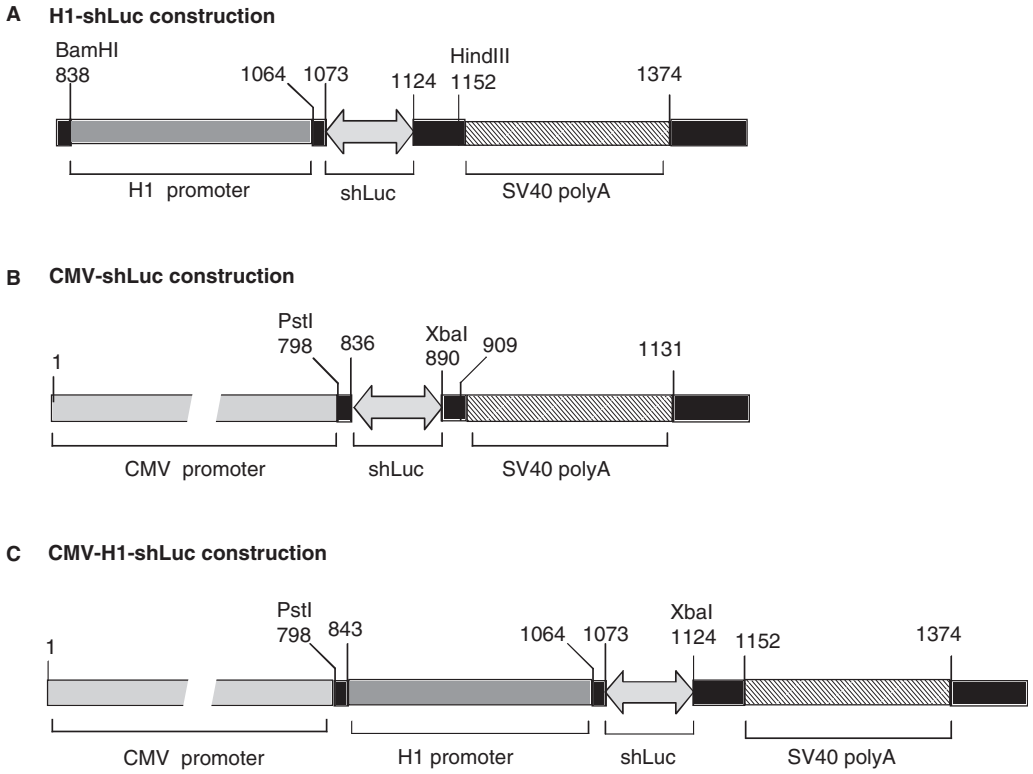


Figure 1. Scheme of the different promoter constructions driving shLuc. The three promoter constructs tested for driving-shLuc mediated inhibition after PEI-based transfection in the brain are shown. In (**A**) the H1 promoter alone; (**B**) CMV enhancer/promoter used alone and (**C**) a hybrid CMV-H1 promoter. In each construction, the shLuc contains a TTTTT sequence required for the H1-transcription arrest and is followed by a SV40 polyadenylation sequence transcription arrest for RNA pol II in CMV- and CMV-H1 promoter constructs.

least three times providing similar results. Means are shown ± SEM. Unpaired Student's *t*-test or a non-parametric Wilcoxon test were used for statistical comparison between the control and treated groups.

For analysis of CyclinD1 expression, we used a luciferase reporter gene under the control of a CyclinD1 promoter. In each group, 0.5 μg of CyclinD1-luciferase plasmid was injected per hemisphere, along with either 0–200 nM of siTRα1, or 0.2–0.5 μg of CMV-H1-shTRα1 plasmid. Note that the total amount of nucleic acids transfected was constant in each experiment, as appropriate quantities of siGFP (siRNA directed against green fluorescent protein) were added bringing the final siRNA concentration in all cases to 200 nM. Likewise, in the experiments using CMV-H1-shTRα1, the total amount of plasmids injected was 1 μg per hemisphere. The irrelevant plasmid used to complete the total amount of nucleic acids in the different groups of the shTRα1 experiment was the empty pcDNA3 (Invitrogen). In these experiments, 10 animals were injected per group and the two hemispheres were quantified separately for reporter gene expression (i.e. $n = 20$ per group). Each experiment was performed at least twice providing similar results. Means ± SEM are shown. Statistical analysis was done using unpaired Student's *t*-test or a nonparametric Wilcoxon test for comparison between the control and treated groups.

Cell culture studies

Human carcinoma cells (HeLa, HeLaX1/5 and HeLa2053) and mouse fibroblastic cells (NIH 3T3) were cultured in complete growth medium (Dulbecco's modified Eagle's medium supplemented with 10% (v/v) foetal calf serum (FCS), 100 μg/ml penicillin and streptomycin) at 37°C with 5% CO_2. One day prior to transfection, 20×10^4 cells were plated in 48-well plates. The next day, the medium was replaced by reduced serum medium (Opti-MEM, Invitrogen). Each well was incubated with a 100 μl transfection mix containing pRL-CMV and pGL2-CMV (both at 30 ng), and 0.37 μl Lipofectamine 2000 reagent (Invitrogen), either with 250 ng pshRNA plasmids or 70 ng siRNAs in 400 μl Opti-MEM for 4 h. HeLaX1/5 cells, which stably expressed firefly luciferase protein under the control of TetOFF-inducible promoter (18) was co-transfected similarly using shRNA carrying plasmids (250 ng) and pRL-CMV (0.1 μg). After the incubation period, the transfection was stopped by adding 500 μl of complete medium with 20% FCS. After 24 h incubation, cells were washed with PBS and lysed by adding passive lysis buffer (Promega). An aliquot of the lysate was assayed for protein concentration (BioRad) and for luciferase activities using Dual luciferase kits (Promega). The firefly luciferase activity was then divided by the *Renilla* luciferase activity in order to normalize for

transfection efficiencies of HeLa and NIH 3T3. All experiments were done in triplicate.

For siTRα1 transfections, the same protocol was adapted to transfect siTRα1 into NIH 3T3 cells stably expressing both TRα1 and TRα2 isoforms. Three concentrations of siTRα1 were tested ranging from 5 to 100 nM (500 µl/well). In this experiment, the transfection vector used was jetSI/DOPE, using the complexation protocol described previously (Preparation of complexes and gene transfer *in vivo* section). One day post-transfection, cells were harvested and total RNAs extracted. For all experiments, reverse transcription was performed with MMLV-RT (Invitrogen) using 1 µg of total RNA mixed with 500 ng of hexamer oligonucleotides according to the manufacturer's instructions. Here, 1% of cDNAs was used for quantitative PCR (Q-PCR) using specific primers (as listed below) and SYBR-Green PCR kit (Qiagen). Q-PCR reactions were performed at least three times in duplicate on a DNA engine Opticon system (MJ Research). Each point was repeated three times (i.e. three wells for each condition) and each well was measured in triplicate. Data were analysed by normalizing TRα1 and TRα2 expression to that of a control gene 36B4. The sequences of the oligonucleotides used are as follows: TRα1 forward: 5′CAGAGGGTGTGCGGAGCTGGT3′; TRα1 reverse: 5′CCTGTCCAAGGGCTGGAGGGT3′; TRα2 forward: 5′GCATGTTGTTCAGGGTCCGCAG GT3′; TRα2 reverse: 5′GGGCTCTTCGGGCTCTGGT GCT3′; 36B4 forward: 5′ACCTCCTTTCCAGGCTTT3′, 36B4 reverse: 5′CCCACCTTGTCTCCAGTCTTT3′. Specificity of the amplification was optimized by determining melting curves of the amplicons.

RESULTS AND DISCUSSION

Our objectives in this study were to compare the capacities of different constructs to direct efficient production of shRNA *in vitro* and *in vivo* in the mammalian brain. The expression cassettes of shRNA delivery vectors typically exploit RNA polymerase III (Pol III) promoters (8–11), (19–20), whilst some authors have suggested that use of a Pol II promoter can be used successfully in certain cell contexts using non-viral (21) or lentiviral vectors (22). However, the use of Pol II promoters for shRNA production is subject to certain constraints. For instance, some authors have demonstrated a need for both a very short distance (6 bp) between the Pol II promoter and the shRNA sequence as well as a short polyadenylation signal (21). Others have shown that the presence of an intron between the Pol II promoter and the shRNA sequence is necessary for efficient production (22). These constraints do not apply to Pol III constructs. We therefore tested the efficiency of a Pol III construct (H1) in different *in vivo* and *in vitro* situations. However, we found that the H1 promoter alone was not very efficient *in vivo*. So we also tried a hybrid CMV enhancer minimal promoter/H1 promoter construct (CMV/H1 construct). We chose this strategy because this enhancer had previously been shown to increase promoter efficiency *in vivo*, such as for the PDGF beta promoter (23) or to increase U6-mediated transcription of shRNAs in plasmids (24). Other fusion promoters have also been described to increase the shRNA production efficiency by recombinant SV40 derivative viruses (25) or by lentiviral vectors (26). In our experimental set up, a construct with CMV enhancer/promoter (CMV-shRNA) alone was used as a control for the enhancement of the H1 promoter, knowing that in our construct the CMV promoter was too far from the hairpin to be efficient in producing shRNA, as shown by Xia *et al.* (12). Given that one of the major advantages of using an shRNA approach is that it should be able to be exploited to provide tissue-specific and developmental-stage-specific gene knockdown, we examined the efficiencies of these constructs in the context of the developing and the adult mouse brain.

Design of different promoter constructs driving the transcription of shLuc

Three different promoter constructs were tested for the production of shLuc (Figure 1). The first was a classic H1 promoter transcribing shLuc from a pSUPER vector backbone (8). In this construction, the type III RNA polymerase synthesizes the shRNA and is stopped by a five-T stretch at the 3′ end of the shRNA (Figure 1a). In this H1-shLuc construct, the first nucleotide of the shLuc sequence is located 30 bp from the TATA box of the H1 promoter (27). Second, in order to improve H1-mediated production of shRNAs, we also constructed a CMV-H1 hybrid construct driving the production of the shRNA. In this construction, the full enhancer and promoter sequence of CMV has been integrated upstream of the H1 promoter (CMV-H1-shLuc). Finally, as a control to validate that in the CMV-H1 construct, the production of the hairpin is generated by the H1 promoter and not by the CMV promoter, we designed a CMV-shLuc plasmid where the CMV enhancer/promoter is too far from the hairpin sequence to allow efficient shRNA production. We thus did not expect the CMV-shLuc plasmid to provide any inhibition of the target luciferase (Figure 1c).

In cell culture, both H1-shLuc and CMV-H1-shLuc trigger efficient target gene knockdown

We tested the efficiency of H1-shLuc to trigger the inhibition of a co-transfected luciferase target gene by RNAi in a mouse fibroblastic cell line (NIH 3T3), as this construct has been shown to be effective in certain culture conditions (28). We then compared the H1-shLuc efficiency to that of the CMV-H1-shLuc in the same cultures.

As described in the Methods section, two luciferase reporter genes (pGL2-CMV and pRL-CMV at 30 ng each) were co-transfected in NIH 3T3 cells (Figure 2a), along with the different constructions described in Figure 1. Quantification of luciferase expression revealed strong inhibition efficiency for both H1-shLuc and CMV-H1-shLuc (0.25 µg–100 µl transfection mixture). This 80% inhibition of the target gene with H1- and CMV-H1-shLuc was stronger than obtained with siLuc used at 70 ng (Figure 4a). Similar results were obtained using HeLa cells instead of NIH 3T3 cells

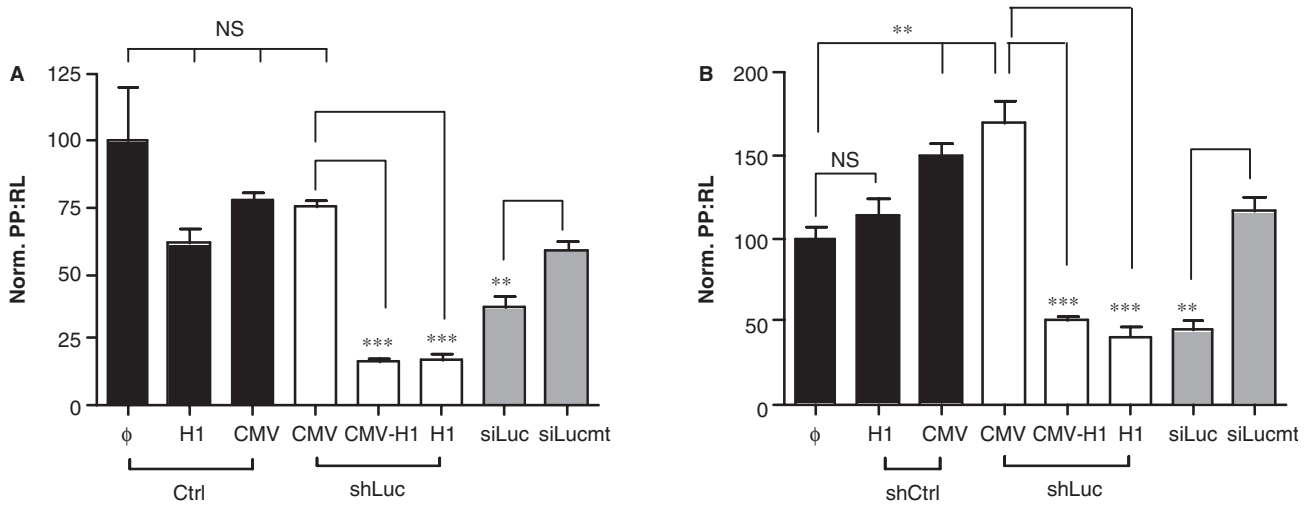


Figure 2. H1- and CMV-H1-shLuc show equal efficiency in cell cultures. (**A**) Plasmids containing shLuc (white bars) or control sequences (black bars) or siRNAs against PP luciferase (grey bars) were transfected in NIH 3T3 cells as described in the methods section, along with two reporter genes (pGL2-CMV and pRL-CMV). Only H1-shLuc and CMV-H1-shLuc lead to an inhibition, like siLuc. (**B**) HeLa X1/5 cells stably expressing firefly luciferase were transfected by the different plasmid constructions described in (a) or with siRNAs. As in NIH 3T3 cells, only H1-shLuc, CMV-H1-shLuc and siLuc provided significant inhibition of the endogenous luciferase expression. Means ± SEM are shown. NS = 'not significant'; $^{*}P \leq 0.05$; $^{**}P \leq 0.01$; $^{***}P \leq 0.001$ ($n = 3$ samples per group). This graph is for a single representative experiment, which was repeated three times giving similar results.

(data not shown). The control CMV-shLuc did not provide inhibition of PP:RL ratio. Even though the CMV-shLuc construct was conceived so as not to produce hairpins we tried to verify its eventual ability to synthesize hairpins by northern blotting. Using northern blot protocols adapted for small size RNA inoculated from transfected 3T3 cells, we found that the CMV-H1-shLuc construct produced a visible 60-nt hairpin sequence, whereas no signal of a functional hairpin detection with the CMV-shLuc construction was seen (data not shown).

The same inhibition of an endogenous luciferase was obtained with H1-shLuc or CMV-H1-shLuc on HeLa cells stably expressing firefly luciferase, HeLa X1/5 (Figure 2b) and HeLa 2053 cell lines which expressed PPluc under the control of a tetracycline responsive promoter (Tet-ON) (data not shown). As shown in Figure 2, we observed a strong inhibition of PP:RL ratio in the presence of H1- or CMV-H1-shLuc, both constructs being able to inhibit endogenous gene expression.

H1-shLuc provides only limited inhibition of target gene expression in the newborn mouse brain

The first construction tested *in vivo* was the H1-shLuc plasmid, which was tested in the newborn mouse brain model. This model provides a useful paradigm for assessing the performance of vectors and plasmids in an *in vivo* context (29,30). After co-transfection of this construction along with target (PPluc) and control (RRluc) luciferase reporter genes in the lateral ventricles of newborn mice, we found that the H1-shLuc-driven inhibition of PPluc was very limited (Figure 3a). In fact, inhibition of target gene expression never exceeded 25% at any time point tested (from 16 to 72 h) (Figure 3b).

This result contrasts with the high efficiency of pSUPER-p53 as first described by Brummelkamp *et al.*, where an H1-mediated transcription of a p53-targeting shRNA leads to the strong knockdown of target gene expression in cell culture (8). Thus, it is worth noting that in our experiments in HeLa cells, the PP:RR ratio was decreased by >80% in the presence of H1-shLuc, a level that is ~2.5 times greater than the inhibition obtained in the mouse brain (25%) with the identical construction. This divergence suggests that the efficiency of shRNA-carrying plasmids in terms of target gene inhibition is cell context-dependent and supports the idea that each tissue and condition (*in vivo* versus *in vitro*) requires optimization for the use of shRNAs-triggered RNAi. Two hypotheses could explain this difference in the efficiency of H1-shRNAs in the mouse brain and in cell cultures. The first could be a lower copy number of shRNAs produced by H1 in the brain compared to cell cultures. The second possibility could be the lower efficiency of Dicer-induced processing of the hairpin in the brain that would lead to less efficient production of shRNA molecules for RNAi triggering.

The hybrid CMV-H1 construct provides the best inhibition of target gene expression in the mammalian brain

We next tested the CMV-H1-shLuc construction in the newborn mouse brain and found that it triggered >50% inhibition of PP:RRluc ratio (Figure 4a). The maximal inhibition obtained with this hybrid promoter construction was observed at 50 h post-transfection (Figure 4b) and was still visible at 72 h post-transfection (data not shown).

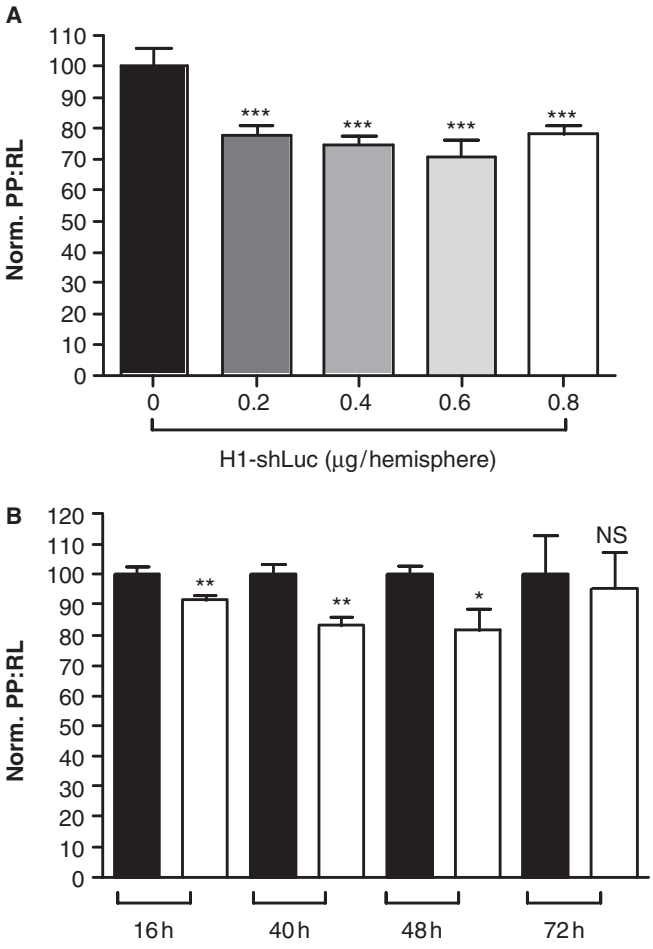


Figure 3. H1-shLuc provides only limited inhibition of a co-transfected luciferase reporter gene *in vivo*: dose dependence and time course. (**A**) Two luciferase reporter genes-containing plasmids (pGL2-CMV and pRL-CMV) were co-transfected by stereotaxic injection into the lateral ventricles of newborn mice brains, along with H1-shLuc construction at different concentrations (grey bars) ranging from 0.1 to 0.4 µg/µl, or with an irrelevant H1-shRNA as a control (black bar). The sequence of the shLuc hairpin is directed against its target gene PP-luc and has no homology with RL-luc, which serves as a control for transfection. The graph represents the normalization of PP-luc against RL-luc expression. In the presence of H1-shLuc at all doses tested, we observed a decrease of PP:RL ratio of ~25%. (**B**) The time course of H1-shLuc efficiency shows that inhibition of target gene expression does not vary between 16 and 48 h. At 72 h, no significant inhibition is found. Black and white bars correspond to 0.86 µg/hemisphere of irrelevant H1-shRNA and H1-shluc constructions, respectively. Means ± SEM are shown. NS = 'not significant'; $^{*}P \leq 0.05$; $^{**}P \leq 0.01$; $^{***}P \leq 0.001$. $n = 10$ per group.

However, when tested in the adult mouse brain, the level of inhibition was not as strong as that observed in the newborn brain. In this situation, the CMV-H1-shLuc construct provided ~25% inhibition of PP:RRluc ratio compared to the control group. This inhibition was constant from 48 to 96 h post-transfection (Figure 4c). Note that in the adult mouse brain, the presence of the H1-shLuc construction did not produce any significant inhibition of its target luciferase gene (Figure 4c grey columns). So in the newborn and in the adult mouse brain context, the presence of the CMV enhancer upstream of the H1 type III promoter driving the synthesis of a hairpin sequence provides a significant improvement in terms of inhibitory effect induced by the shRNA.

Whilst this work was in progress, a paper appeared describing the use a PolIII promoter appended to a CMV enhancer (31), which is similar but distinct from our CMV-H1-shLuc construction, in which the entire CMV enhancer/promoter is located upstream of the H1-promoter. A further difference is that Ong *et al.* (2005) used the hybrid promoter first in plasmid constructs, *in vitro*, then in baculovirus-mediated infection of shRNAs *in vivo* in rat brains. In contrast, our results are obtained using non-viral constructs both *in vitro* and *in vivo*. Thus, given the generally greater facility of creating and using non-viral constructs, our findings that the hybrid promoter is optimal *in vivo* with plasmid-based constructs will open up new possibilities for applying shRNA technology in new settings.

CMV-shLuc is inefficient in both the newborn and in the adult mouse brain

In order to test whether the increased inhibitory efficiency of the CMV-H1 hybrid promoter was due to the CMV enhancer/promoter alone, we injected CMV-shLuc along with the two luciferase reporter genes in the brains of newborn mice. We found no inhibition of the target gene in the CMV-shLuc groups at any dose tested, compared to the control groups (Supplementary Figure 1). This construction was also tested in the adult mouse brain, where no inhibition was visible (data not shown). This result supports the idea that in the CMV-H1 hybrid promoter, the CMV only plays the role of an enhancer, and that the transcription of the shRNA is mediated by the type III promoter H1 alone and not by the CMV enhancer/promoter [as observed in (31)]. The position of the CMV promoter start site, >100 bp from the hairpin sequence, explains the lack of efficient shRNA production, as demonstrated by Xia *et al.* (12).

SiRNA or shRNA knockdown of thyroid hormone receptor α1 relieves T3-dependent repression of CyclinD1-luciferase transcription

After having demonstrated that the presence of a CMV enhancer/promoter upstream of an H1 promoter strongly increases the inhibition efficiency of shRNAs, we exploited this technology to follow physiological regulations in the mouse brain *in vivo*. Previous studies revealed both the role of thyroid hormones (TH) in NSC cycling *in vivo* (32), and the specific role of TRα1 as the mediator of this regulation. We hypothesized that the action of T3 on NSC cycling may implicate CyclinD1 and assessed whether this regulation was mediated or not by TRα1.

Indeed, CyclinD1 is known to be a T3-regulated gene. But T3-dependent regulations can vary as a function of tissue and developmental stage. For instance, T3 up-regulates CyclinD1 in the pancreas (33) and liver (34), while T3 represses CyclinD1 transcription in neuroblastoma cells (35). In order to assess CyclinD1 regulation in neurogenic areas of the brains of newborn

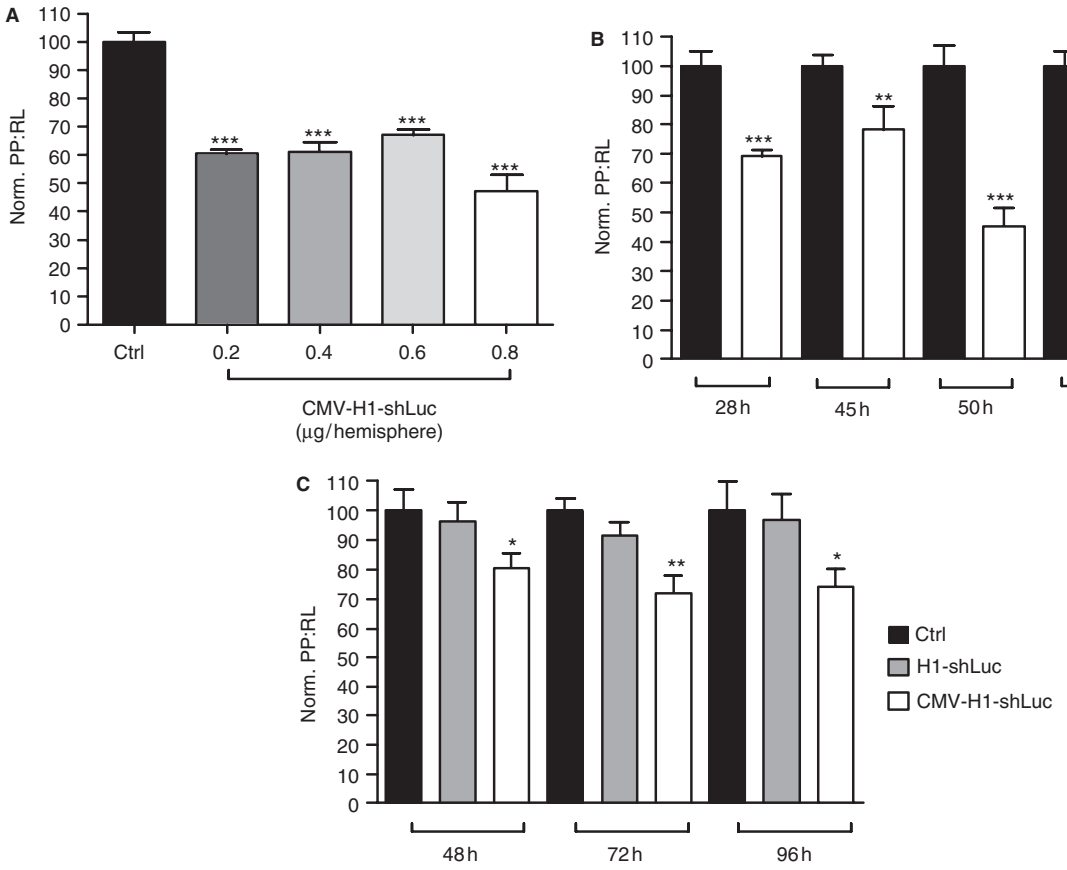


Figure 4. A CMV-H1 hybrid construct driving shLuc provides enhanced inhibition of a co-transfected target gene in the newborn (a and b) and in the adult (c) mouse brain. (**A**) Dose dependence of CMV-H1-shLuc efficiency. The inhibition efficiency of CMV-H1-shLuc was tested at different doses ranging from 0.1 to 0.4 µg/µl. After co-transfection of pGL2-CMV and pRL-CMV along with 0.4 µg/µl of CMV-H1-shLuc (i.e. 0.8 µg/hemisphere), we observed up to 50% inhibition of the targeted luciferase expression. This level of inhibition was obtained at 48 h post-transfection. (**B**) Time course efficiency of CMV-H1-shLuc. Significant inhibition of the target gene with 0.4 µg/µl of CMV-H1-shLuc was seen at all times tested. The maximal level of inhibition (50%) was seen at 50 h post-transfection. (**C**) In the adult brain, H1-shLuc provided no inhibition of PP:RL ratio (grey bars) compared to controls (black bars). CMV-H1-shLuc leads to 25% inhibition of the target gene at 72 h post-transfection (white bars) and up to 112 h post-transfection (data not shown). Means ± SEM are shown. NS = 'not significant'; $^{*}P \leq 0.05$; $^{**}P \leq 0.01$; $^{***}P \leq 0.001$. $n = 10$ injected hemispheres per group.

mice, a CyclinD1-luciferase reporter gene was injected into the lateral ventricles of hypothyroid animals. Luciferase expression of the reporter gene was followed 18 h after T3 treatment compared to controls (Figure 5b). We observed a significant decrease of CyclinD1-luciferase expression in T3-treated animals. Thus, CyclinD1 transcription is negatively regulated in neurogenic areas lining the lateral ventricles of newborn mice brains.

In order to assess whether TRα1 is involved in this T3-mediated repression of CyclinD1 expression, we designed siRNAs directed against TRα1, and checked their efficiency and their specificity on 3T3 culture cells stably expressing TRα1 and TRα2 (Figure 5a). Using quantitative real-time RT-PCR, the siTRα1 designed was shown to specifically inhibit the expression of TRα1 and did not affect the expression of TRα2 at any dose tested (Figure 5a). By co-transfecting these siTRα1 along with the CyclinD1-luciferase reporter gene in the brain of euthyroid newborn mice, we observed a significant and dose-dependent increase in CyclinD1-luciferase expression (Figure 5c). The same result was obtained when using plasmids coding for shTRα1 under the control of a CMV-H1 construct: we observed a significant up-regulation of CyclinD1-luciferase expression in CMV-H1-shTRα1-injected mice brains (Figure 5d).

These data suggest first, that PEI-vectorized shRNA technology can be used to examine physiological regulations *in vivo* and second, that TRα1 may play a role in the regulation of CyclinD1 by thyroid hormones. Note that we observed the same increase in CyclinD1-luciferase expression following co-injection of either 200 nM of siTRα1 (Figure 5c) or by 0.5 µg of CMV-H1-shTRα1 plasmid per hemisphere (Figure 5d).

CONCLUDING REMARKS

The results presented here show that PEI-based delivery of shLuc to the newborn or the adult mouse brain is optimal when used with a plasmid construct bearing a hybrid CMV enhancer–H1 promoter. The system was most efficient in the newborn brain, reaching 50% inhibition

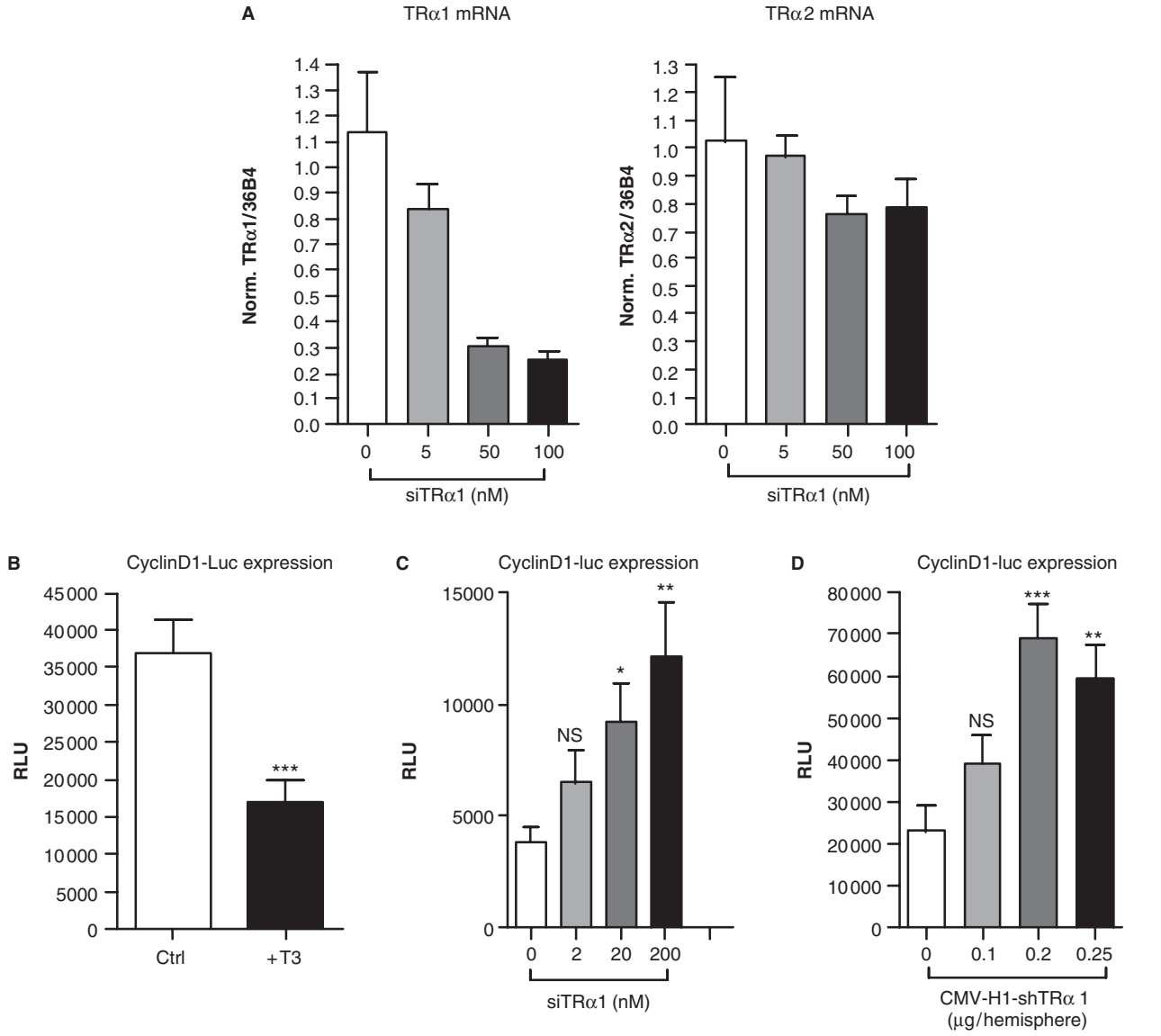


Figure 5. CyclinD1 transcriptional repression by T3 implicates thyroid hormone receptor α1 (TRα1). (**A**) Quantification of TRα1 and TRα2 mRNAs in 3T3 cells transfected by 5–100 nM of siTRα1 (500 µl per well). The total amount of TRα1 and TRα2 mRNA in each condition was normalized to that of 36B4 (a control gene) in the same group. Note that only TRα1, and not TRα2 expression, is inhibited by shTRα1 (**B**) CyclinD1-luciferase reporter gene expression *in vivo* is down-regulated in the presence of T3. CyclinD1-luciferase complexed by L-PEI was injected into the lateral ventricles of hypothyroid newborn mice. Then either a saline solution (NaCl 0.9%, control) or T3 (2.5 mg/kg body weight) was injected subcutaneously (see Methods section). Luciferase assays were performed 18 h later. T3 treatment induces a strong decrease in cyclinD1-luciferase expression compared to the control group. (**C**) Different doses (ranging from 2 to 200 nM) of siRNAs directed against TRα1 (siTRα1) were co-injected along with the CyclinD1-luciferase plasmid in the lateral ventricles of euthyroid newborn mice, inducing a strong activation of CyclinD1-luciferase expression. (**D**) The same experiment as described in (c) was performed by using plasmids coding for shTRα1 instead of siTRα1. The different doses of shTRα1-coding plasmid tested ranged from 0.1 to 0.25 µg/hemisphere. We observed a similar result to that seen with siTRα1, i.e. a strong de-repression of CyclinD1-luciferase following knockdown of TRα1. Means ± SEM are shown, $n = 20$ injected hemispheres per group. Each experiment was performed at least twice, providing similar results. NS = 'not significant'; $^{*}P \leq 0.05$; $^{**}P \leq 0.01$; $^{***}P \leq 0.001$ ($n = 3$ samples per group).

of target gene expression. This level of inhibition is less than that obtained in the newborn mouse brain when siRNA against luciferase is delivered with cationic lipids (16). However, as cationic-lipid-delivered siRNA in the adult brain is inefficient (data not shown), we propose that this PEI-based delivery of shLuc under a hybrid promoter will become a useful tool for performing gene knockdown in the brain at different developmental stages. As this non-viral delivery technology combines high efficiency with the use of non-immunogenic components, it could also prove to be useful in terms of therapeutic applications. Indeed, the current demonstration that it can be used to dissect the role of different transcription factors in regulating specific genes, already opens up new possibilities for reverse

genetic studies focused on specific target tissues at defined developmental stages.

ACKNOWLEDGEMENTS

Funding to pay the open access publication charge was provided by the CNRS. ZH was a fellow of the Association Française contre les Myopathies (AFM) and the Fondation pour la Recherche Medicale. This work was supported by the AFM and EU contract, LSHM CT 2005-018652, Crescendo.

Conflict of interest statement. None declared.

REFERENCES

1. Napoli,C., Lemieux,C. and Jorgensen,R. (1990) Introduction of a chimeric chalcone synthase gene into petunia results in reversible co-suppression of homologous genes in trans. *Plant Cell*, **2**, 279–289.
2. Hutvagner,G. and Zamore,P.D. (2002) RNAi: nature abhors a double-strand. *Curr. Opin. Genet. Dev.*, **12**, 225–232.
3. Manche,L., Green,S.R., Schmedt,C. and Mathews,M.B. (1992) Interactions between double-stranded RNA regulators and the protein kinase DAI. *Mol. Cell. Biol.*, **12**, 5238–5248.
4. Minks,M.A., West,D.K., Benvin,S. and Baglioni,C. (1979) Structural requirements of double-stranded RNA for the activation of 2',5'-oligo(A) polymerase and protein kinase of interferon-treated HeLa cells. *J. Biol. Chem.*, **254**, 10180–10183.
5. Elbashir,S.M., Harborth,J., Lendeckel,W., Yalcin,A., Weber,K. and Tuschl,T. (2001) Duplexes of 21-nucleotide RNAs mediate RNA interference in cultured mammalian cells. *Nature*, **411**, 494–498.
6. Song,E., Lee,S.K., Wang,J., Ince,N., Ouyang,N., Min,J., Chen,J., Shankar,P. and Lieberman,J. (2003) RNA interference targeting Fas protects mice from fulminant hepatitis. *Nat. Med.*, **9**, 347–351.
7. Dorn,G., Patel,S., Wotherspoon,G., Hemmings-Mieszczak,M., Barclay,J., Natt,F.J., Martin,P., Bevan,S., Fox,A. *et al.* (2004) siRNA relieves chronic neuropathic pain. *Nucleic Acids Res.*, **32**, e49.
8. Brummelkamp,T.R., Bernards,R. and Agami,R. (2002) A system for stable expression of short interfering RNAs in mammalian cells. *Science*, **296**, 550–553.
9. Paul,C.P., Good,P.D., Winer,I. and Engelke,D.R. (2002) Effective expression of small interfering RNA in human cells. *Nat. Biotechnol.*, **20**, 505–508.
10. Sui,G., Soohoo,C., Affarel,B., Gay,F., Shi,Y. and Forrester,W.C. (2002) A DNA vector-based RNAi technology to suppress gene expression in mammalian cells. *Proc. Natl Acad. Sci. USA*, **99**, 5515–5520.
11. Yu,J.Y., DeRuiter,S.L. and Turner,D.L. (2002) RNA interference by expression of short-interfering RNAs and hairpin RNAs in mammalian cells. *Proc. Natl. Acad. Sci. USA*, **99**, 6047–6052.
12. Xia,H., Mao,Q., Paulson,H.L. and Davidson,B.L. (2002) siRNA-mediated gene silencing in vitro and in vivo. *Nat. Biotechnol.*, **20**, 1006–1010.
13. Uprichard,S.L., Boyd,B., Althage,A. and Chisari,F.V. (2005) Clearance of hepatitis B virus from the liver of transgenic mice by short hairpin RNAs. *Proc. Natl. Acad. Sci. USA*, **102**, 773–778.
14. Gurzov,E.N. and Izquierdo,M. (2006) RNA interference against Hec1 inhibits tumor growth in vivo. *Gene Ther.*, **13**, 1–7.
15. Raoul,C., Abbas-Terki,T., Bensadoun,J.C., Guillot,S., Haase,G., Szulc,J., Henderson,C.E. and Aebischer,P. (2005) Lentiviral-mediated silencing of SOD1 through RNA interference retards disease onset and progression in a mouse model of ALS. *Nat. Med.*, **11**, 423–428.
16. Hassani,Z., Lemkine,G.F., Erbacher,P., Palmier,K., Alfama,G., Giovannangeli,C., Behr,J.P. and Demeneix,B.A. (2005) Lipid-mediated siRNA delivery down-regulates exogenous gene expression in the mouse brain at picomolar levels. *J. Gene Med.*, **7**, 198–207.
17. Lemkine,G.F. and Demeneix,B.A. (2001) Polyethylenimines for in vivo gene delivery. *Curr. Opin. Mol. Ther.*, **3**, 178–182.
18. Gossen,M. and Bujard,H. (1992) Tight control of gene expression in mammalian cells by tetracycline-responsive promoters. *Proc. Natl. Acad. Sci. USA*, **89**, 5547–5551.
19. Boden,D., Pusch,O., Lee,F., Tucker,L., Shank,P.R. and Ramratnam,B. (2003) Promoter choice affects the potency of HIV-1 specific RNA interference. *Nucleic Acids Res.*, **31**, 5033–5038.
20. Kawasaki,H. and Taira,K. (2003) Short hairpin type of dsRNAs that are controlled by tRNA(Val) promoter significantly induce RNAi-mediated gene silencing in the cytoplasm of human cells. *Nucleic Acids Res.*, **31**, 700–707.
21. Zhou,H., Xia,X.G. and Xu,Z. (2005) An RNA polymerase II construct synthesizes short-hairpin RNA with a quantitative indicator and mediates highly efficient RNAi. *Nucleic Acids Res.*, **33**, e62.
22. Yang,S. and Hecht,N.B. (2004) Translin associated protein X is essential for cellular proliferation. *FEBS Lett.*, **576**, 221–225.
23. Liu,B.H., Wang,X., Ma,Y.X. and Wang,S. (2004) CMV enhancer/human PDGF-beta promoter for neuron-specific transgene expression. *Gene Ther.*, **11**, 52–60.
24. Xia,X.G., Zhou,H., Ding,H., Affarel,B., Shi,Y. and Xu,Z. (2003) An enhanced U6 promoter for synthesis of short hairpin RNA. *Nucleic Acids Res.*, **31**, e100.
25. Strayer,D.S., Feitelson,M., Sun,B. and Matskevich,A.A. (2005) Paradigms for conditional expression of RNA interference molecules for use against viral targets. *Methods Enzymol.*, **392**, 227–241.
26. Unwalla,H.J., Li,H.T., Bahner,I., Li,M.J., Kohn,D. and Rossi,J.J. (2006) Novel Pol II fusion promoter directs human immunodeficiency virus type 1-inducible coexpression of a short hairpin RNA and protein. *J. Virol.*, **80**, 1863–1873.
27. Hannon,G.J., Chubb,A., Maroney,P.A., Hannon,G., Altman,S. and Nilsen,T.W. (1991) Multiple cis-acting elements are required for RNA polymerase III transcription of the gene encoding H1 RNA, the RNA component of human RNase P. *J. Biol. Chem.*, **266**, 22796–22799.
28. Wyatt,C.A., Geoghegan,J.C. and Brinckerhoff,C.E. (2005) Short hairpin RNA-mediated inhibition of matrix metalloproteinase-1 in MDA-231 cells: effects on matrix destruction and tumor growth. *Cancer Res.*, **65**, 11101–11108.
29. Boussif,O., Lezoualc'h,F., Zanta,M.A., Mergny,M.D., Scherman,D., Demeneix,B. and Behr,J.P. (1995) A versatile vector for gene and oligonucleotide transfer into cells in culture and in vivo: polyethylenimine. *Proc. Natl. Acad. Sci. USA*, **92**, 7297–7301.
30. Schwartz,B., Benoist,C., Abdallah,B., Scherman,D., Behr,J.P. and Demeneix,B.A. (1995) Lipospermine-based gene transfer into the newborn mouse brain is optimized by a low lipospermine/DNA charge ratio. *Hum. Gene. Ther.*, **6**, 1515–1524.
31. Ong,S.T., Li,F., Du,J., Tan,Y.W. and Wang,S. (2005) Hybrid cytomegalovirus enhancer-h1 promoter-based plasmid and baculovirus vectors mediate effective RNA interference. *Hum. Gene. Ther.*, **16**, 1404–1412.
32. Lemkine,G.F., Raj,A., Alfama,G., Turque,N., Hassani,Z., Alegria-Prevot,O., Samarut,J., Levi,G. and Demeneix,B.A. (2005) Adult neural stem cell cycling in vivo requires thyroid hormone and its alpha receptor. *FASEB J.*, **19**, 863–865.
33. Ledda-Columbano,G.M., Perra,A., Pibiri,M., Molotzu,F. and Columbano,A. (2005) Induction of pancreatic acinar cell proliferation by thyroid hormone. *J. Endocrinol.*, **185**, 393–399.
34. Alisi,A., Demori,I., Spagnuolo,S., Pierantozzi,E., Fugassa,E. and Leoni,S. (2005) Thyroid status affects rat liver regeneration after partial hepatectomy by regulating cell cycle and apoptosis. *Cell Physiol. Biochem.*, **15**, 69–76.
35. Garcia-Silva,S. and Aranda,A. (2004) The thyroid hormone receptor is a suppressor of ras-mediated transcription, proliferation, and transformation. *Mol. Cell. Biol.*, **24**, 7514–7523.

International Journal of Biological Sciences
ISSN 1449-2288 www.biolsci.org 2006 2(3):125-132

Research paper

Phylogenetic analysis of Amphioxus genes of the proprotein convertase family, including aPC6C, a marker of epithelial fusions during embryology

Stéphanie Bertrand[1], Alain Camasses[2], Mathilde Paris[1], Nicholas D. Holland[3], Hector Escriva[2]

1. Laboratoire de Biologie Moléculaire de la Cellule, UMR 5161 CNRS/ENS de Lyon, Ecole Normale Supérieure de Lyon, IFR128 BioSciences Lyon-Gerland, 46, Allée d'Italie, 69364 Lyon Cedex 07, France.
2. Laboratoire Arago, UMR 7628, CNRS and University Pierre and Marie Curie, BP 44, F-66651 Banyuls sur Mer, France.
3. Marine Biology Research Division, Scripps Institution of Oceanography, University of California San Diego, La Jolla, CA 92093, USA

Corresponding address: Hector Escriva, Laboratoire Arago, UMR 7628, CNRS and University Pierre and Marie Curie, BP 44, F-66651 Banyuls sur Mer, France. hector.escriva@obs-banyuls.fr.

Received: 2006.01.27; Accepted: 2006.04.22; Published:

The proprotein convertases (PCs) comprise a family of subtilisin-like endoproteases that activate precursor proteins (including, prohormones, growth factors, and adhesion molecules) during their transit through secretory pathways or at the cell surface. To explore the evolution of the PC gene family in chordates, we made a phylogenetic analysis of PC genes found in databases, with special attention to three PC genes of the cephalochordate amphioxus, the closest living invertebrate relative to the vertebrates. Since some vertebrate PC genes are essential for early development, we investigated the expression pattern of the C isoform of the amphioxus PC6 gene (aPC6C). In amphioxus embryos and larvae, aPC6C is expressed at places where epithelia fuse. Several kinds of fusions occur: ectoderm-to-ectoderm during neurulation; mesoderm-to-ectoderm during formation of the preoral ciliated pit; and endoderm-to-ectoderm during formation of the mouth, pharyngeal slits, anus, and external opening of the club-shaped gland. Presumably, at all these sites, aPC6C is activating proteins favoring association between previously disjunct cell populations.

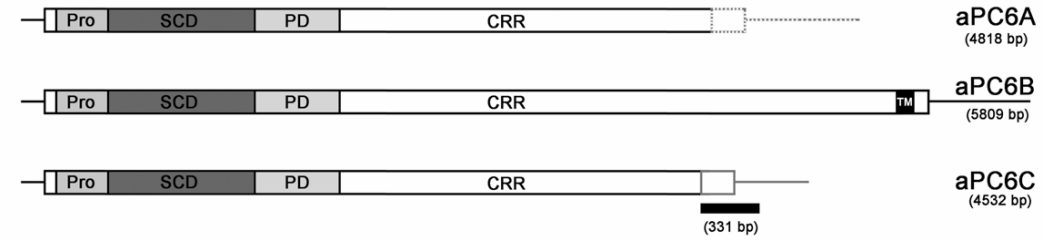
Key words: amphioxus, gene expression pattern, prohormone convertases, PC6

1. Introduction

The function of proteins destined for export or for the plasma membrane surface in eukaryotic cells is usually regulated by a specific proteolysis mechanism of inactive precursors which generates biologically active molecules. This process of activation is usually performed by a family of calcium-dependent endoproteases related to the prokaryotic subtilisin and the yeast kexin enzymes that are called proprotein convertases (PCs) [1, 2]. PCs (also called kexins) are members of the subtilisin-like proteases superfamily, together with five other protein families (i.e. subtilisins, thermitases, proteinase K, lantibiotic peptidases and pyrolysins) [3]. Seven PC subfamilies have been identified, including furin/PACE [4], PC1/PC3 [5], PC2 [6, 7], PC4 [8], PACE4 [9], PC5/PC6 [10], and PC7/LPC/PC8 [11]. PCs have been found not only in vertebrates but also in yeast [12-14], Hydra [15], protostomes [16-19], and in the invertebrate chordate amphioxus [20, 21], which places the origin of the family very early during eukaryotic evolution. PCs share a common structure of five domains including an N-terminal signal peptide followed by a propeptide of 80-90 residues terminating with the activation canonical cleavage site R-X-R/R-R, a catalytic domain of approximately 240 residues related to subtilisin, a conserved P or Homo B domain of approximately 150 residues which function is still unknown, and a variable C-terminal domain [1, 2], (Fig. 1). The catalytic and P-domains are highly conserved, and the C-terminal domain is variable, both in length and sequence, but contains a conserved Cys rich region and a transmembrane domain in some members of the PC family. Depending on the presence or absence of this transmembrane domain, PCs function in different subcellular compartments. The members with transmembrane domains (e.g. furin, PC6B, or PC7) appear to function primarily in the *trans*-Golgi network (TGN) and/or in constitutive vesicles derived from it. The other members (e.g. PC2, PC1/PC3, PACE4, PC4, PC6A) are sorted into dense-core vesicles in the regulated secretory pathway where they process a variety of prohormones or proneuropeptides [1, 2]. Two members of the PC family are expressed exclusively in neuroendocrine tissues (PC1 and PC2) and one is restricted to reproductive organs (PC4). The remaining PCs are expressed in many tissues of mammalian embryos and adults, including neuroendocrine system, liver, gut, and brain [22]. The colocalization of a PC and its substrate both at the tissue and subcellular level likely contributes to substrate selectivity. Some PCs, like PC6, undergo tissue-specific alternative splicing that in the case of PC6 generates soluble and membrane bound isoforms which are involved in the regulated secretory pathway or in the TGN compartment

respectively. This indicates that alternative splicing plays a very important role in the control of the PC6 gene function.

Figure 1. Schematic representation of the three different splicing variants described for the amphioxus PC6 gene, aPC6A, aPC6B and aPC6C [20]. Boxes represent the coding sequence including propeptide (Pro), catalytic domain (SCD), P domain (PD), the 3' Cys rich domain (CRR) and the transmembrane domain of the B isoform (TM). The three isoforms differ in the 3' part of the CRR domain and are indicated by a grey dashed line for the A isoform or a grey line for the C isoform. The position of the 331 bp sequence of aPC6C used as a probe for the *in toto in situ* hybridizations is indicated with a black box.



In chordates, including the cephalochordate amphioxus, the physiological roles of PACE4 and PC6 genes during development are poorly understood, even if they are critical for a correct developmental process [23]. Amphioxus is the best available stand-in for the extinct proximate invertebrate ancestor of the vertebrates. At the anatomical level, amphioxus is vertebrate-like but simpler. In many respects amphioxus is like a stripped-down, generalized vertebrate. Indeed, it has pharyngeal gill slits, a dorsal hollow nerve cord and notochord, but lacks paired eyes, ears, limbs and neural crest. Moreover, at the genetic level, amphioxus genome also shares many characteristics with vertebrate genomes, but is less complex. The amphioxus genome has not undergone the two waves of gene duplications, that took place during vertebrate evolution and which are responsible for the presence of several duplicated vertebrate genes whereas only one pro-orthologue is present in amphioxus [24].

To gain insight into the evolution of the PC gene family, we have studied: (i) the phylogenetic relationships of known PC genes in eukaryotes, and (ii) we have characterized the expression pattern of the isoform C of the invertebrate chordate amphioxus (*Branchiostoma lanceolatum*) PC6 gene (i.e. the pro-ortholoue of vertebrate PC6 and PACE4 genes). We show first, that the PC genes family is not divided into seven subfamilies as previously published [3, 20], but only into five or six groups, and second, by using *in toto in situ* hybridization, we show that aPC6C is expressed in regions where there is a contact between different embryonic layers like in mouth, exterior opening of the club-shaped gland, and in anus (i. e. endoderm-ectoderm associations), in ciliated pit (ectoderm-mesoderm), or in the sealing of the ectoderm mid-dorsally after neurulation (i.e. ectoderm-ectoderm associations). These results suggest that aPC6C could be involved in epithelial fusions during embryology.

2. Materials and methods

Phylogenetic analysis

Amino acid sequences were aligned using the CLUSTAL W program [25] and manually corrected with SEAVIEW [26]. Phylogenetic trees were inferred by (1) the Neighbor-Joining method [27] with Poisson-corrected distances on amino acids, implemented in PHYLO_WIN [26]; and (2) with PHYML [28], a fast and accurate maximum likelihood heuristic, under the JTT substitution model [29], with a gamma distribution of rates between sites (six categories, parameter alpha estimated by PHYML). Amino acid sites with gaps in any sequence were excluded from the calculations. The bootstrap analysis (1000 repetitions), was carried out by the method of Felsenstein [30]. Divergent sequences for which the alignment was uncertain were excluded.

Embryo collection, in situ hybridization, and histology

Ripe animals of the Mediterranean amphioxus (*Branchiostoma lanceolatum*) were collected in Argelès-sur-Mer (France), and gametes were obtained by heat stimulation [31]. A 331 bp fragment of *B. lanceolatum* aPC6 cDNA was used for the synthesis of antisense riboprobes. This fragment includes the coding sequence of the last 64 aminoacids and extends from nucleotide position 4058 to position 4389 of the previously published sequence for aPC6c [20], and is specific to the c isoform of aPC6. Fixation, whole-mount *in situ* hybridization and histological sections were performed as described by Holland et al (1996) [24].

3. Results and discussion

Phylogenetic analysis of the PC gene family

To examine the relationships between invertebrate PC and vertebrate PC genes, we constructed a phylogenetic tree with PC amino acid sequences obtained from GenBank, including all the invertebrate sequences that we found for each PC subfamily, the yeast kexins and selected vertebrate

sequences from at least one mammalian, one amphibian and one fish representative. Phylogenetic trees were constructed with two different methods (ML and NJ), and were rooted with a group of invertebrate and vertebrate sequences of the the subtilisin related protein Site-1 (membrane-bound transcription factor site-1) [32]. The results obtained with both methods were similar and those obtained with the NJ method are shown in Fig. 2. From these analyses we define six orthologous PC subfamilies supported by high bootstrap values: (i) a subfamily containing vertebrate PC7 and yeast kexins; (ii) the PC2 subfamily; (iii) the furins and PACE sequences; (iv) the PC4 subfamily; (v) a family containing PC5/PC6 and PACE4 sequences and (vi) the PC1/PC3 subfamily. Different authors have divided the PC family into seven groups, since the PC5/PC6/PACE4 subfamily was previously described as two different groups. However, the use of invertebrate sequences (and particularly the invertebrate chordate amphioxus sequence of PC6), clearly shows that PACE4 and PC5/PC6 groups appeared specifically in vertebrates. These paralogous genes should have arisen from the genome duplications that occurred between the chordata and the vertebrata radiations [33].

Developmental expression of aPC6C

Three members of the PC family are known in the cephalochordate amphioxus (*B. californiense*), PC2, PC3 and PC6 [20, 21]. Moreover, three different splicing variants have been described for the amphioxus PC6 gene, named aPC6A, aPC6B and aPC6C (Fig 1) [20]. However, the function of the amphioxus PCs is not yet known. As a first step in the study of the function of amphioxus PCs, we have determined the pattern of expression of the aPC6C isoform during the embryonic development of *Branchiostoma lanceolatum* by *in situ* hybridization of whole mount animals (Fig. 3). We first isolated a 331 bp cDNA fragment specific for the C isoform that includes the coding sequence for the last 64 amino acids and 192 nucleotides of the 3′ non-coding region (Fig. 1). The nucleotide sequence identity between the published *B. californiense* sequence (accession number AAF26302) and the *B. lanceolatum* one is 100% (data not shown), including the 3′ non-coding part. We will refer to this gene as aPC6C instead of AmphiPC6C (the customary way of referring to genes from *B. floridae*) because Oliva et al (2000) [20] originally used this abbreviation, even if they probably worked on *B. floridae* due to some confusion about the species by their biological supply company.

The gene expression pattern of the aPC6C isoform was examined from mid-blastula through the early larval stage. aPC6C transcripts are ubiquitously distributed up until late gastrula, but become spatially restricted thereafter. In late gastrula, aPC6C transcripts are found dorsally in the ectoderm and posteriorly around the blastopore. In mid-neurula, aPC6C is expressed in the ectoderm in the most anterior and posterior parts of the embryo as well as in ectodermal cells fusing in the dorsal midline. In the early larvae (before the mouth opens), aPC6C is still expressed in ectodermic cells posteriorly, in cells bordering the neuropore, and the ventral ectoderm of the anterior part of the body. In later larvae (with the open mouth and first gill slit), in the anterior part of the body aPC6C expression is detected in cells bordering the neuropore, in ventral ectoderm, in the cells of the ciliated pit, in the pharyngeal endoderm, in some ectodermal cells around the mouth, in cells of the club-shaped gland near its external opening and in ectodermal cells just outside the opening of the first gill slit. In the posterior part of the body aPC6C transcripts can be detected in the cells just within the anus as well as in elongated ectoderm cells comprising the tail. This expression pattern suggests that the isoform C of the amphioxus PC gene may play an important role in epithelial fusions during development. Thus, in all the body openings where ectoderm-endoderm contacts are very important (e.g. mouth, exterior opening of club-shaped gland and anus), as well as in regions where there are ectoderm-ectoderm and ectoderm-mesoderm associations (e.g. the sealing of the ectoderm mid-dorsally after neurulation and the ciliated pit respectively), aPC6C is expressed.

In vertebrates the two orthologs of the amphioxus PC6 gene (PC6 and PACE4, see fig 2) are expressed in several tissues or organs homologs to the tissues where the amphioxus PC6C gene is expressed. However, many other expression domains have also been described. Thus, in *Xenopus*, xPACE4 is expressed in a completely different way than the amphioxus PC6C isoform. xPACE4 is a maternal RNA unequally distributed in the oocyte. Later on, a localized expression is detected in the notochord, the brain and a subset of endodermal precursors [34]. For the second paralogue of aPC6 in *Xenopus,* xPC6, gene expression was studied by using a probe that recognizes all the different isoforms of the gene. Like in amphioxus, xPC6 transcripts are ubiquitously distributed until the end of gastrulation but they exhibit a dynamic expression pattern shortly thereafter. Indeed, xPC6 is mainly expressed in ectoderm-derived tissues (e.g. neural folds, neural crest, eyes, nasal placode, lateral line, otic vesicle and brain), but also in mesoderm (e.g. pronephric duct, notochord) and other mixed structures like pharyngeal arches [35]. In mouse, PC6 is first expressed in extraembryonic tissues and in the distal region of the primitive streak, a homolog of the amphioxus blastopore. Later on, expression is observed in the somites, the dorsal surface ectoderm, the vertebral cartilage primordia and in the apical ectodermal ridge of limb buds [36].

The differences in the expression pattern between aPC6C and the vertebrate PACE4 and PC6 genes may either represent a secondary derivation in amphioxus or a function that has been coopted in vertebrates after the vertebrate-specific genes duplications. However, it is not excluded that other

splicing variants, not yet characterized either in vertebrates or in amphioxus, can show closely related expression patterns different from the ones already described. A very important step for the comprehension of PC genes function during chordate development in the future will be, first, the complete characterization of the expression pattern of each splicing variant both in vertebrates as well as in amphioxus.

As we have proposed above, aPC6C could be involved in epithelial fusions during embryology. In the same way, the mouse PC6 gene shows a prominent expression in the site of fusion of the two decidual lobes, a extraembryonic structure, and in the distal region of the primitive streak [36]. Moreover, another closely related proprotein convertase in vertebrates, the furin (see fig. 2), plays a role in the ventral closure of the cardiogenic mesoderm [37], indicating that the function of PCs in the association of different tissues may not be entirely restricted to the PC6 paralogue. Since PCs are proteases that activate many key regulatory molecules, our results suggest that similar regulators implicated in the association between different tissues can be controlled through proteolytic cleavage by using different members of the PC gene family. In vertebrates, the natural substrates of furin, PACE4 and PC6 remain poorly determined, as well as the extent to which their substrate specificities overlap under physiological conditions. However, *in vitro* experiments have shown that furin could be implicated in remodeling of the extracellular matrix by processing metalloproteinases [38, 39] and furin, as well as PC5/6, in the regulation of cell adhesion by processing integrins [40, 41]. The study, both *in vitro* and *in vivo,* of the substrates for each PC (including their splicing variants) under physiological conditions, as well as the degree in which different PC genes can process similar substrates in order to control perfectly the local concentration of active molecules at the place where they are required will be extremely important in the future for the comprehension of the role of different PC genes during chordate development.

Figure 2. Phylogeny of the PC gene family. Tree was constructed by the distance neighbor-joining method with 1000 bootstrap replicates in order to test the robustness of the branches. Bootstrap values (in %) are indicated on each branch of the tree. Root was placed using an outgroup of several subtilisin related proteins Site-1, including sequences from *Arabidopsis thaliana* (NP_197467), *Drosophila melanogaster* (NP_649337), *Anopheles gambiae* (XP_320328), chicken (XP_414071), human (NP_003782), and mouse (NP_062683). Accession numbers of the sequences used are: P29146, XP_541276, NP_006191, XP_520079, Q04592, XP_424841, XP_355911, NP_002561, Q9NJ15, XP_542201, NP_032819, NP_001021543, NP_727963, AAA87006, XP_424712, NP_058787, NP_038656, CAA46031, AAA87005, XP_419332, NP_002585, NP_032818, NP_001023732, NP_477318, XP_308012, NP_594835, NP_014161, NP_001025528, NP_004707, NP_032820, AAW83023, AAH84090, AAH94153, CAA47118, AAT99304, CAF99544, CAG03450, CAG06456, CAF98620, CAG01397, CAG08431, BAD11989, XP_850069, BAC97793, XP_784245, BAC05491, XP_393918, AAA27768, XP_545820, CAA92109, NP_002560, AAA37643, XP_545820, XP_585571, AAW83025, NP_060043, AAA41814, Q28193, BAA00877 and XP_542201. Invertebrate sequences are in colored boxes, yeast in blue, diploblasts in magenta, protostomes in green and deuterostomes (including amphioxus) in red. The different subfamilies are clustered with the same colored branches. Abreviations are: Sp, *Schizosaccharomyces pombe*; Sc, *Saccharomyces cerevisiae*; Hv, *Hydra vulgaris*; Ce, *Caenorhabditis elegans*; Dm, *Drosophila melanogaster*; Ag, *Anopheles gambiae*; Bm, *Bombix mori*; Am, *Apis melifera*; Ac, *Aplysia californica*; Hr, *Halocynthia roretzi*; Sp, *Strongylocentrotus purpuratus*; Amphioxus, *Branchiostoma californiense*; Gg, *Gallus gallus*; Xl, *Xenopus laevis*; Tn, *Tetraodon nigroviridis*; Ol, *Oryzias latipes*; Hs, *Homo sapiens*; Mm, *Mus musculus*; Cf, *Canis familiaris*; Bt, *Bos taurus*; Rn, *Rattus norvegicus*.

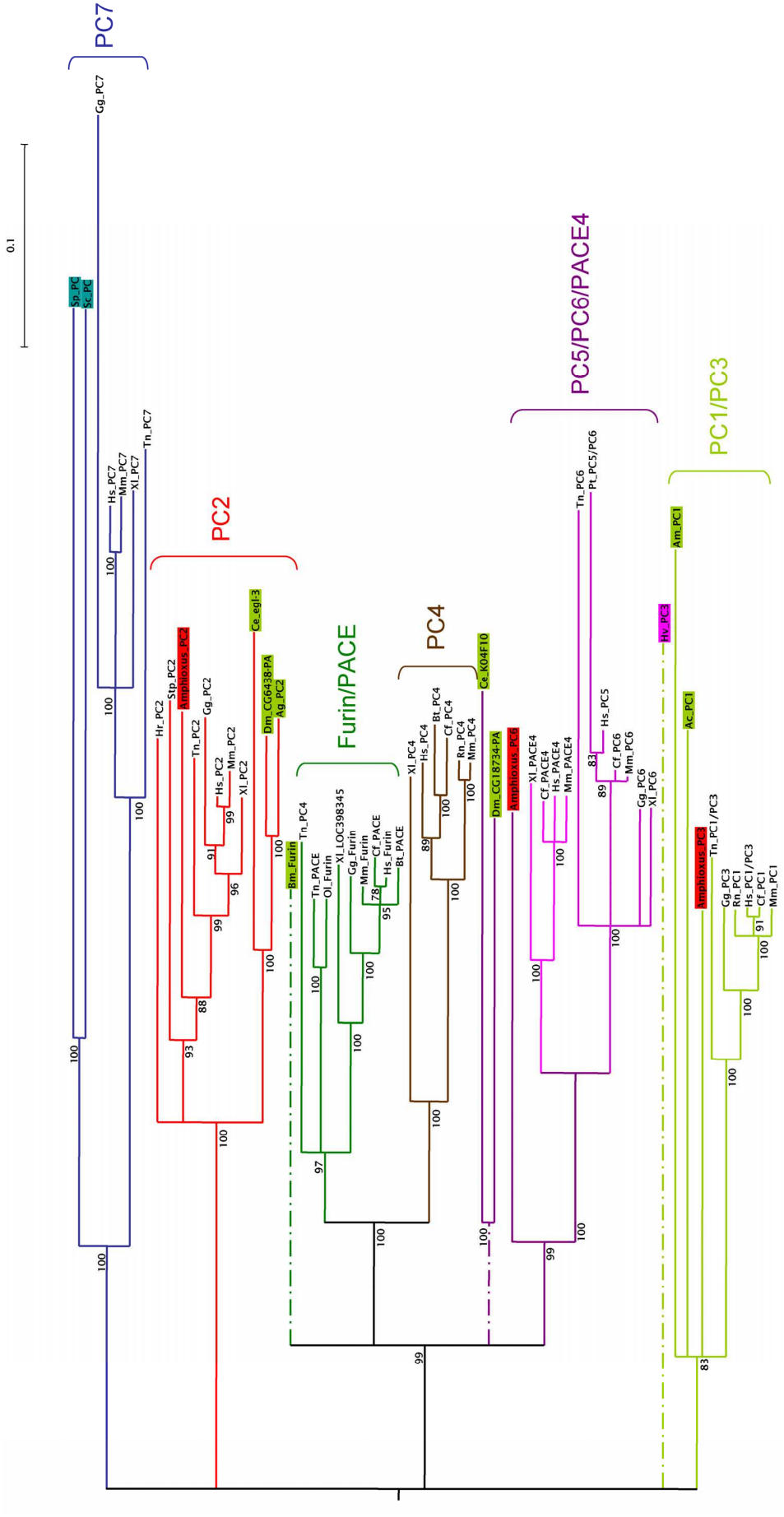
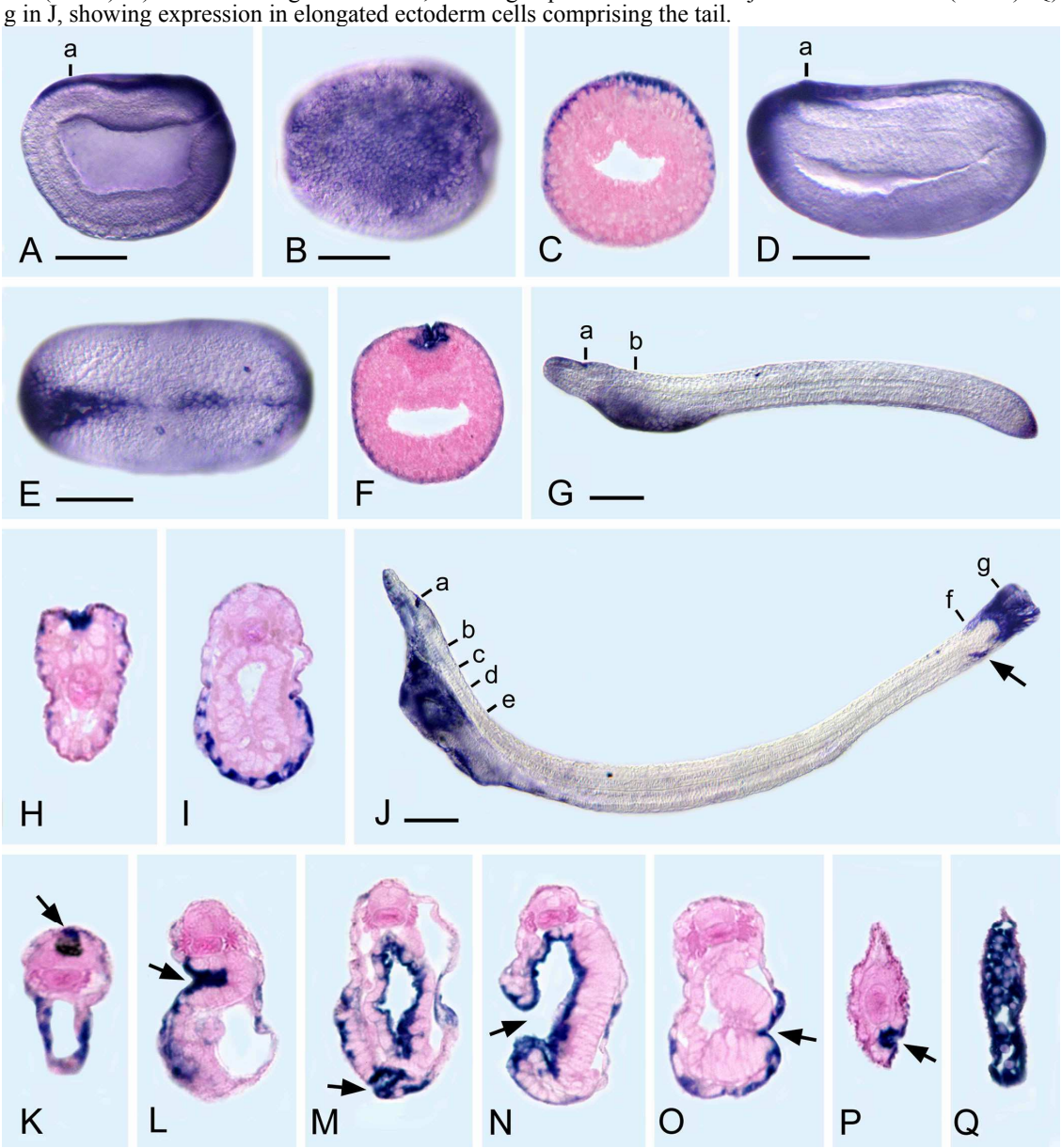

PC7
PC2
Furin/PACE
PC4
PC5/PC6/PACE4
PC1/PC3
0.1
Cg_PC7
Sp_PC
Sc_PC
Hs_PC7
Mm_PC7
Xl_PC7
Tn_PC7
Hr_PC2
Stp_PC2
Amphioxus_PC2
Tn_PC2
Gg_PC2
Hs_PC2
Mm_PC2
Xl_PC2
Ce_egl-3
Dm_CG6438-PA
Ag_PC2
Bm_Furin
Tn_PC4
Tn_PACE
Ol_Furin
Xl_LOC398345
Gg_Furin
Mm_Furin
Cf_PACE
Hs_Furin
Bt_PACE
Xl_PC4
Hs_PC4
Bt_PC4
Cf_PC4
Rn_PC4
Mm_PC4
Ce_K04F10
Dm_CG18734-PA
Amphioxus_PC6
Xl_PACE4
Cf_PACE4
Hs_PACE4
Mm_PACE4
Tn_PC6
PL_PC5/PC6
Hs_PC5
Cf_PC6
Mm_PC6
Gg_PC6
Xl_PC6
Hv_PC3
Am_PC1
Ac_PC1
Amphioxus_PC3
Tn_PC1/PC3
Gg_PC3
Rn_PC1
Hs_PC1/PC3
Cf_PC1
Mm_PC1

Figure 3. *AmphiPC6* expression in embryos and larvae of amphioxus. All whole mounts with anterior toward left and 50 µm scale lines. Sections, counterstained pink, as seen from posterior end of body. A) Side view of late gastrula with expression in ectoderm dorsally and posteriorly. B) Dorsal view of A. C) Section through level a in A. D) Side view of early neurula with expression in ectoderm dorsally, anteriorly, and posteriorly. E) Dorsal view of D. F) Section through level a in D, showing expression in ectoderm cells fusing in the dorsal midline. G) Side view of larva shortly before the mouth opens, showing ectodermal expression in the neuropore, ventroanteriorly, and posteriorly. H) Section through level a in G showing expression in the ectoderm cells bordering the neuropore. I) Section through level b in G, showing ectodermal expression on the ventral side of the body. J) Side view of larva with an open mouth and first gill slit; the arrow indicates the position of the anus. K) Section through level a in J, showing expression in ectoderm cells ventrally and also dorsally (arrow) near the neuropore. L) Section through level b in J, showing strong expression in the cells of the ciliated pit (arrow). M) Section through c in J, showing expression in the pharyngeal endoderm, in some ectoderm cells and in cells of the club-shaped gland near its external opening (arrow). N) Section through level d in J, showing the open mouth (arrow) and expression in the pharyngeal endoderm. O) Section through level e in J, showing ectodermal expression just outside the opening of the first gill slit (arrow). P) Section through level f in J, showing expression in the cells just within the anus (arrow). Q) Section through g in J, showing expression in elongated ectoderm cells comprising the tail.



Acknowledgements

S.B. is supported by HRA Pharma, Fondation Mérieux, Fondation par la recherche Médicale and Association de Recherche contre le Cancer.

Conflict of interest

The authors have declared that no conflict of interest exists.

References

1. Rouille Y, et al. Proteolytic processing mechanisms in the biosynthesis of neuroendocrine peptides: the subtilisin-like proprotein convertases. Front Neuroendocrinol, 1995. 16(4): 322-61.
2. Zhou A, et al. Proteolytic processing in the secretory pathway. J Biol Chem, 1999. 274(30): 20745-8.
3. Siezen R.J and Leunissen J.A. Subtilases: the superfamily of subtilisin-like serine proteases. Protein Sci, 1997. 6(3): 501-23.
4. Van den Ouweland A.M, et al. Nucleotide sequence analysis of the human fur gene. Nucleic Acids Res, 1989. 17(17): 7101-2.
5. Smeekens S.P and D.F Steiner. Identification of a human insulinoma cDNA encoding a novel mammalian protein structurally related to the yeast dibasic processing protease Kex2. J Biol Chem, 1990. 265(6): 2997-3000.
6. Seidah N.G, et al. cDNA sequence of two distinct pituitary proteins homologous to Kex2 and furin gene products: tissue-specific mRNAs encoding candidates for pro-hormone processing proteinases. DNA Cell Biol, 1990. 9(10): 789.
7. Seidah N.G, et al. cDNA sequence of two distinct pituitary proteins homologous to Kex2 and furin gene products: tissue-specific mRNAs encoding candidates for pro-hormone processing proteinases. DNA Cell Biol, 1990. 9(6): 415-24.
8. Nakayama K, et al. Identification of the fourth member of the mammalian endoprotease family homologous to the yeast Kex2 protease. Its testis-specific expression. J Biol Chem, 1992. 267(9): 5897-900.
9. Kiefer M.C, et al. Identification of a second human subtilisin-like protease gene in the fes/fps region of chromosome 15. DNA Cell Biol, 1991. 10(10): 757-69.
10. Lusson J, et al. cDNA structure of the mouse and rat subtilisin/kexin-like PC5: a candidate proprotein convertase expressed in endocrine and nonendocrine cells. Proc Natl Acad Sci U S A, 1993. 90(14): 6691-5.
11. Seidah N.G, et al. cDNA structure, tissue distribution, and chromosomal localization of rat PC7, a novel mammalian proprotein convertase closest to yeast kexin-like proteinases. Proc Natl Acad Sci U S A, 1996. 93(8): 3388-93.
12. Davey J, et al. Isolation and characterization of krp, a dibasic endopeptidase required for cell viability in the fission yeast Schizosaccharomyces pombe. Embo J, 1994. 13(24): 5910-21.
13. Julius D, et al. Isolation of the putative structural gene for the lysine-arginine-cleaving endopeptidase required for processing of yeast prepro-alpha-factor. Cell, 1984. 37(3): 1075-89.
14. Mizuno K, et al. Yeast KEX2 genes encodes an endopeptidase homologous to subtilisin-like serine proteases. Biochem Biophys Res Commun, 1988. 156(1): 246-54.
15. Chan S.J, et al. Conservation of the prohormone convertase gene family in metazoa: analysis of cDNAs encoding a PC3-like protein from hydra. Proc Natl Acad Sci U S A, 1992. 89(15): 6678-82.
16. Chun J.Y, et al. The function and differential sorting of a family of aplysia prohormone processing enzymes. Neuron, 1994. 12(4): 831-44.
17. Peters K, McDowall J., and Rose A.M. Mutations in the bli-4 (I) locus of Caenorhabditis elegans disrupt both adult cuticle and early larval development. Genetics, 1991. 129(1): 95-102.
18. Roebroek A.J, et al. cDNA sequence of a Drosophila melanogaster gene, Dfur1, encoding a protein structurally related to the subtilisin-like proprotein processing enzyme furin. FEBS Lett, 1991. 289(2): 133-7.
19. Smit A.B, et al. Structural characterization of a Lymnaea putative endoprotease related to human furin. FEBS Lett, 1994. 343(1): 27-31.
20. Oliva A.A Jr, Chan S.J, and Steiner D.F. Evolution of the prohormone convertases: identification of a homologue of PC6 in the protochordate amphioxus. Biochim Biophys Acta, 2000. 1477(1-2): 338-48.
21. Oliva A.A Jr, Steiner D.F., and Chan S.J. Proprotein convertases in amphioxus: predicted structure and expression of proteases SPC2 and SPC3. Proc Natl Acad Sci U S A, 1995. 92(8): 3591-5.
22. Taylor N.A, Van De Ven W.J., and Creemers J.W. Curbing activation: proprotein convertases in homeostasis and pathology. Faseb J, 2003. 17(10): 1215-27.
23. Essalmani R, et al. Deletion of the gene encoding proprotein convertase 5/6 causes early embryonic lethality in the mouse. Mol Cell Biol, 2006. 26(1): 354-61.
24. Holland L.Z, Holland P.W.H., and Holland N.D. Revealing homologies between body parts of distantly related animals by in situ hybridization to developmental genes: Amphioxus versus vertebrates. In: Ferraris J.D and Palumbi S.R., Editors. Molecular Zoology: Advances, Strategies, and Protocols. Wiley-Liss. 1996: 267-282.
25. Thomson J.D, Higgins D.G., and Gibson T.J. Clustal W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Research, 1994. 22: 4673-4680.
26. Galtier N, Gouy M., and Gautier C. SEAVIEW and PHYLO_WIN: two graphic tools for sequence alignment and molecular phylogeny. Computer Applications in the Biosciences, 1996. 12: 543-548.
27. Saitou N. and Nei M. The neighbor-joining method: a new method for reconstructing phylogenetic trees. Molecular biology and evolution, 1987. 4: 406-425.
28. Guindon S. and Gascuel O. A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. Syst. Biol, 2003. 52: 696–704.
29. Jones D.T, Taylor W.R., and Thornton J.M. The rapid generation of mutation data matrices from protein sequences. Comput. Appl. Biosci, 1992. 8: 275–282.
30. Felsenstein J. Confidence limits on phylogenies: an approach using the bootstrap. Evolution, 1985. 39: 783-791.
31. Fuentes M, et al. Preliminary observations on the spawning conditions of the European amphioxus (Branchiostoma lanceolatum) in captivity. J Exp Zoolog Part B Mol Dev Evol, 2004. 302(4): 384-91.
32. Seidah N.G, et al. Mammalian subtilisin/kexin isozyme SKI-1: A widely expressed proprotein convertase with a unique cleavage specificity and cellular localization. Proc Natl Acad Sci U S A, 1999. 96(4): 1321-6.
33. Dehal P and Boore J.L. Two rounds of whole genome duplication in the ancestral vertebrate. PLoS Biol, 2005. 3(10): e314.
34. Birsoy B, et al. XPACE4 is a localized pro-protein convertase required for mesoderm induction and the cleavage of specific TGFbeta proteins in Xenopus development. Development, 2005. 132(3): 591-602.
35. Nelsen S, et al. Proprotein convertase genes in Xenopus development. Dev Dyn, 2005. 233(3): 1038-44.
36. Constam D.B, Calfon M., and Robertson E.J. SPC4, SPC6, and the novel protease SPC7 are coexpressed with bone morphogenetic proteins at distinct sites during embryogenesis. J Cell Biol, 1996. 134(1): 181-91.
37. Roebroek A.J, et al. Failure of ventral closure and axial rotation in embryos lacking the proprotein convertase Furin. Development, 1998. 125(24): 4863-76.
38. Pei D and Weiss S.J. Furin-dependent intracellular activation of the human stromelysin-3 zymogen. Nature, 1995. 375(6528): 244-7.
39. Leighton M and Kadler K.E. Paired basic/Furin-like proprotein convertase cleavage of Pro-BMP-1 in the trans-Golgi network. J Biol Chem, 2003. 278(20): 18478-84.

40. Lehmann M, et al. Lack of integrin alpha-chain endoproteolytic cleavage in furin-deficient human colon adenocarcinoma cells LoVo. Biochem J, 1996. 317: 803-9.
41. Stawowy P, et al. Endoproteolytic activation of alpha(v) integrin by proprotein convertase PC5 is required for vascular smooth muscle cell adhesion to vitronectin and integrin-dependent signaling. Circulation, 2004. 109(6): 770-6.

THE JOURNAL OF BIOLOGICAL CHEMISTRY VOL. 284, NO. 3, pp. 1938–1948, January 16, 2009

Structural and Functional Insights into the Ligand-binding Domain of a Nonduplicated Retinoid X Nuclear Receptor from the Invertebrate Chordate Amphioxus*[S]

Received for publication, July 24, 2008, and in revised form, October 17, 2008 Published, JBC Papers in Press, November 4, 2008, DOI 10.1074/jbc.M805692200

Giuseppe D. Tocchini-Valentini[‡1,2], Natacha Rochel[‡1], Hector Escriva[§3], Pierre Germain[¶], Carole Peluso-Iltis[‡], Mathilde Paris[§], Sarah Sanglier-Cianferani[‖], Alain Van Dorsselaer[‖], Dino Moras[‡4], and Vincent Laudet[§5]

From the [‡]Département de Biologie et de Génomique Structurales and [¶]Département Biologie du Cancer, Institut de Génétique et de Biologie Moléculaire et Cellulaire, UMR7104 CNRS, U596 INSERM, Université Louis Pasteur, 67400 Illkirch, France, [§]Institut de Génomique Fonctionnelle de Lyon, Ecole Normale Supérieure de Lyon, Université Lyon 1, CNRS, INRA, Institut Fédératif 128 Biosciences Gerland, 69364 Lyon Sud, France, and [‖]Laboratoire de Spectrométrie de Masse Bio-Organique, UMR/CNRS 7178, Institut Pluridisciplinaire Hubert CURIEN, Université Louis Pasteur, Ecole de Chimie, Polymères et Matériaux, 67087 Strasbourg, France

Retinoid X nuclear receptors (RXRs), as well as their insect orthologue, ultraspiracle protein (USP), play an important role in the transcription regulation mediated by the nuclear receptors as the common partner of many other nuclear receptors. Phylogenetic and structural studies have shown that the several evolutionary shifts have modified the ligand binding ability of RXRs. To understand the vertebrate-specific character of RXRs, we have studied the RXR ligand-binding domain of the cephalochordate amphioxus (*Branchiostoma floridae*), an invertebrate chordate that predates the genome duplication that produced the three vertebrates RXRs (α, β, and γ). Here we report the crystal structure of a novel apotetramer conformation of the AmphiRXR ligand-binding domain, which shows some similarity with the structures of the arthropods RXR/USPs. AmphiRXR adopts an apo antagonist conformation with a peculiar conformation of helix H11 filling the binding pocket. In contrast to the arthropods RXR/USPs, which cannot be activated by any RXR ligands, our functional data show that AmphiRXR, like the vertebrates/mollusk RXRs, is able to bind and be activated by RXR ligands but less efficiently than vertebrate RXRs. Our data suggest that amphioxus RXR is, functionally, an intermediate between arthropods RXR/USPs and vertebrate RXRs.

The nuclear receptor (NR)[6] superfamily is specific to animals and performs an abundance of functions, from embryonic development to metamorphosis and from homeostasis of various physiological functions to the control of metabolism (for a review, see Ref. 1). Nuclear receptors are ligand-activated transcription factors. Many members of the superfamily thus bind major hormones, such as steroids, thyroid hormones, or retinoids. These occupy a special position in gene regulation by providing a direct link between the ligand, which they bind, and the target gene, whose expression they regulate. There are also many nuclear receptors that are not known to bind any ligand and are thus called "orphan nuclear receptors" (2–4). Some nuclear receptors which were originally discovered as orphans have since been shown to bind small hydrophobic molecules, such as fatty acids, but others appear to be true orphans (4).

Nuclear receptors are composed of several functional domains. The N-terminal A/B domain is highly variable in length and sequence and contains a constitutively active transactivation function AF-1. Most conserved are domains C and E, corresponding grossly to the DNA-binding domain and the ligand-binding domain (LBD), respectively. The LBD contains the ligand-dependent transactivation function AF-2. The DNA-binding domain and LBD are connected via a flexible hinge (domain D). NRs act *in vivo* and *in vitro* as ligand-dependent transcriptional transregulators through binding, most often as heterodimers, to *cis*-acting response elements present in cognate target genes. Activation of gene transcription occurs after binding of ligand, which results in a conformational change of the LBD, leading to release of corepressor and binding of coactivator to the LBD (5).

* This study was supported by CNRS, INSERM, Université Louis Pasteur, European Commission on Structural Proteomics in Europe (SPINE) Grant QLG2-CT-220-0098, and SPINE2-Complexes Grant LSHG-CT-2006-031220 under the integrated program, Quality of Life and Management of Living Resources. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

[S] The on-line version of this article (available at http://www.jbc.org) contains supplemental Table 1 and Figs. 1 and 2.

[1] Both of these authors contributed equally to this work.

[2] Present address: Istituto di Biologia Cellulare, Consiglio Nazionale delle Ricerche, Campus A, Buzzati-Traverso, Via Ramarini 32, Monterotondo Scalo, 00016 Rome, Italy and Istituto di Genetica Molecolare, Consiglio Nazionale delle Ricerche, Via Abbiategrasso 207, I-27100 Pavia, Italy.

[3] Present address: Laboratoire Arago, UMR 7628, CNRS and University Pierre and Marie Curie, BP 44, F-66651 Banyuls sur Mer, France.

[4] To whom correspondence may be addressed: Laboratoire de Biologie et Génomique Structurales, IGBMC, 1 Rue Laurent Fries, BP 10142, Illkirch 67404, France. Tel.: 33-388-653-220; Fax: 33-388-653-276; E-mail: moras@igbmc.u-strasbg.fr.

[5] To whom correspondence may be addressed: Institut de Génomique Fonctionnelle de Lyon, Molecular Zoology team, Ecole Normale Supérieure de Lyon 1, Institut Fédératif 128 Biosciences Gerland, 69364 Lyon Sud, France. Tel.: 33-472-728-190; Fax: 33-472-728-080; E-mail: vincent.laudet@ens-lyon.fr.

[6] The abbreviations used are: NR, nuclear receptor; RXR, retinoid X nuclear receptor; AmphiRXR, amphioxus RXR; HsRXR, *Homo sapiens* RXR; MmRXR, *Mus musculus* RXR; Bg, RXR, *B. glabrata* RXR; LBD, ligand-binding domain; LBP, ligand-binding pocket; USP, ultraspiracle protein; BtUSP, *B. tabaci*; TcUSP, *T. castaneum* USP; H*x*, helix *x*; RA, retinoic acid; DHA, docohexanoic acid; ESI, electrospray ionization; MS, mass spectrometry.

To date, the crystal structures of ~40 different NR LBDs have been solved, mainly mammalian nuclear receptors. All structures are very similar, sharing a common fold containing 11–12 helices (6–7). The majority of the LBD crystals were obtained in the presence of a ligand (holo-LBD) and only a few without ligand (apo-LBD). Comparison of apo- and holo-LBD structures led to the so-called "mouse trap model" of the ligand-dependent conformational switch (8). In apo-LBD, different positions of the C-terminal helix 12 (H12; also called AF-2 helix) highlight the large flexibility of this region (7, 9). In holo-LBD, H12 is folded back on the core of the LBD, closing the LBP. This holo-position of H12 is stabilized by direct or indirect interactions with the ligand. The tightened position of H12, together with H3 and H4, results in the formation of a hydrophobic groove in which coactivator proteins can bind through a conserved L*XX*LL motif, forming an α-helix. The ligand-induced conformational change concomitantly leads to destruction of the surface that binds the corepressor, resulting in its release from the LBD.

The conserved domains present in NRs allow relatively easy identification of nuclear receptors in genomic sequences (10) and also robust phylogenetic reconstruction on the scale of the superfamily (11, 12). Such analysis shows that NRs for similar ligands do not group in the tree but are interspersed with receptors for totally different ligands, whereas orphans are widely distributed in the tree. This observation has led to the proposal that the superfamily evolved from an orphan receptor that gained the capacity of ligand binding several times independently (13–15). The case of retinoid receptors is interesting in that respect. Indeed, at least two distinct types of nuclear receptors, the RARs (α (NR1B1), β (NR1B2), and γ (NR1B3); activated by all-*trans*-retinoic acid (RA)) and the RXRs (α (NR2B1), β (NR2B2) and γ (NR2B3); activated by 9-*cis*-RA) apparently transduce the retinoid message (16–18). It has thus been proposed that the capacity to bind, for example, retinoic acids, would have been acquired several times by mutation of an orphan nuclear receptor, allowing the establishment of nuclear retinoic acid signaling pathways (RARs and RXRs). The case of RXR and USP, its arthropod homologue, is especially interesting in that context, because this receptor plays a pivotal role inside the NR superfamily as a ubiquitous heterodimer partner for class II NRs (19).

It is noteworthy that structural analysis can provide more insight into the evolution of a protein domain (*e.g.* through the comparative structural analysis of domains from animals located at critical positions in the evolutionary tree). In that respect, the study of RXR LBD is important, since RXR is an ancient gene that was subjected to several functional shifts during evolution. In addition, the status of RXR as a *bona fide* liganded receptor and the identity of its endogenous ligand are still under discussion (18, 20). Several experiments suggest that there is a very small amount of 9-*cis*-RA *in vivo*, casting doubts on its physiological relevance as an RXR ligand (21). In addition, it has been suggested, from *in vivo* detection of RXR-activating molecules, that docohexaneoic acid (DHA) could be a ligand for RXR (22). Thus, it may be possible that 9-*cis*-RA is only a pharmacological activator of RXR.

A large number of homologues of RXR have been identified in a variety of animals including early metazoans, such as sponges and cnidarians, but also arthropods, mollusks, and invertebrate chordates. In cnidarians, RXR was shown to bind 9-*cis*-RA (23), and it activates transcription[7] in response to it. In contrast, in insects, the orthologue of RXR, called USP, is unable to bind 9-*cis*-RA and experienced a very high evolutionary rate in dipterans and lepidopterans (see Ref. 24 and references therein). Detailed comparative structural analysis of USPs from insects have shown that these can be separated into two different types of receptors: (i) in basal insects, such as hemipterans and coleopterans, the USP ligand-binding pocket is partly filled by residues of unfolded H6 and H11 helices and is thus likely to be an orphan receptor; (ii) in mecopterida (a terminal group of insects that includes dipteras and lepidopteras among others), a phospholipid is present as a structural ligand inside the LBD that may be indicative of the existence of an endogenous ligand yet to be identified (28). In mollusks, RXR can be activated by 9-*cis*-RA and DHA, at least using transient transfection experiments (25). The case of amphioxus is especially interesting, since this species is clearly located in the lineage giving rise to vertebrates but before the complete genome duplications of vertebrates (26) (see Fig. 7 for a phylogenetic tree). Thus, AmphiRXR is closely related to the three vertebrate RXR genes but represents a nonduplicated version of these genes. Consequently, determining the structure of the LBD encoded by the unique amphioxus AmphiRXR gene and the comparison of this structure with its vertebrate orthologues will infer the main structural features of RXR in the common ancestor of vertebrates (*i.e.* before gene duplication events).

In this paper, we report the crystal structure of the apo form AmphiRXR LBD, revealing a stable antagonist conformation not driven by either a fortuitous or an antagonist ligand. A detailed analysis of the comparison with previous crystal structures of HsRXR (29–34), of nonvertebrate RXR (25) and USPs (27, 28) in different conformational states allows us to better understand the evolution and functional dynamics of these orphan receptors and to identify the structural basis of their functional differences.

EXPERIMENTAL PROCEDURES

Cloning and Purification—The cDNA coding for the *Branchiostoma floridae* RXR (accession number AF378829), AmphiRXR (35) ligand-binding domain (residues 266–484) was PCR-amplified from pSG5 vector and cloned by homologous recombination into the vector pET15b. The residue's position number given here corresponds to the AmphiRXR sequence with the corresponding human sequence numbers within parentheses, (HsAA-*xxx*), for some residues. The recombinant protein was overproduced and purified in two steps. Briefly, the LBD of AmphiRXR (residues 266–484) was cloned in pET15b expression vector to obtain an N-terminal hexahistidine-tagged fusion protein and was overproduced in *Escherichia coli* BL21 (DE3). Cells were grown in LB medium and subsequently induced for 6 h at 20 °C with 1 mM isopropyl thio-β-D-galactoside. Protein purification included a metal

[7] H. Escriva and V. Lauder, unpublished results.

affinity chromatography step on a nickel-Hitrap column and a gel filtration on Superdex 75 16/60. The cell pellets were resuspended in binding buffer (20 mM Tris-HCl, pH 8.0, 300 mM NaCl, 10% glycerol, 5 mM imidazole), lysed by sonication, and clarified by centrifugation. The cleared supernatant was loaded on a nickel-Hitrap column (GE Healthcare). The protein was eluted using 150 and 300 mM imidazole, pH 8.0. The protein was subjected to gel filtration on a Superdex 75 16/60 column (GE Healthcare) in 20 mM Tris, pH 8, 200 mM NaCl, and 5 mM dithiothreitol. The protein was then concentrated to 10 mg/ml using a 10K Centricon device. The purity and homogeneity of the protein were assessed by SDS-PAGE and native PAGE and electrospray ionization mass spectrometry. The protein is in equilibrium between homotetramer and homodimer according to analytical gel filtration and native PAGE. The homodimer dissociates into a monomer when characterized by mass spectrometry studies in 50 mM ammonium acetate.

Transient Transfection Assays—293T cells were maintained in Dulbecco's modified Eagle's medium supplemented with 5% of charcoal-treated fetal calf serum. The cells were transfected at 70% confluence in 24-well plates using 4 μl of ExGen 500 (Euromedex) with 1.0 μg of total DNA, including 0.1 μg of reporter plasmid (17m)x5-tk-luc, and 10 ng of cytomegalovirus-β-galactosidase as an internal control to account for variations of transfection efficiency. The culture medium was changed 6 h after transfection, and when appropriate, ligands dissolved in EtOH were added to different final concentrations (10^{-9} to 10^{-6} M). Cells were lysed 24 h after transfection and assayed for luciferase activity. For the test of TIF2 (transcriptional intermediary factor 2) action, 10–500 ng of TIF2 expression vector were assessed.

Pull-down Assays—DNAs of TIF2.1 and TIF2.1m123, kindly provided by Hinrich Gronemeyer (36), were both transcribed and translated *in vitro* using the TNT T7-coupled reticulocyte lysate system (Promega) following the instructions of the manufacturer. Batches of 50 μM Ni^{2+}-nitrilotriacetic acid-agarose gel (Qiagen) washed and equilibrated in buffer A (50 mM Tris-HCl (pH 8.0), 100 mM NaCl, 10 mM $MgCl_2$, 5 mM imidazole, 10% glycerol, and 0.1% Nonidet P-40) were loaded with 2 μg of *E. coli*-purified His-tagged LBDs (HsRXRα or AmphiRXR). Incubation with TIF2.1 or TIF2.1m123 was done in a final volume of 50 μl of buffer A containing 2 μl of *in vitro* translated protein and 10 μg of bovine serum albumin and in the presence of 2 μM ligand or 1% of ethanol. After incubation for 1 h at 4 °C and three washing steps in buffer A, bound proteins were recovered with SDS sample buffer and revealed by autoradiography of 12% SDS-polyacrylamide gels.

Electrospray Ionization Mass Spectrometry—Before mass analysis, amphioxus and human RXR LBDs purification buffers were exchanged against 50 mM ammonium acetate on two successive gel filtration columns (NAP-10; GE Healthcare) and then concentrated on Centricon 10-kDa devices. ESI-MS measurements were performed on an electrospray time-of-flight mass spectrometer (LCT; Waters, Manchester, UK). The mass measurements of the noncovalent complexes were performed in ammonium acetate (50 mM, pH 6.8). Samples were diluted to 20 μM and continuously infused into the ESI ion source at a flow rate of 6 μl/min through a Harvard syringe pump. Titration experiments were done by adding 2.5 mol eq of ligands to HsRXR or AmphiRXR, followed by a 15-min incubation at 4 °C. Competition experiments were performed by adding an equimolar mixture of two ligands (each in a 2.5-fold molar excess) to the protein, followed by a 15-min incubation. A 2.5-fold molar excess of ligand was used as a standard in our mass spectrometry experiments for two reasons: (i) to be able to detect ligands even with poor affinities (50–100 μM) and (ii) to assess the specificity of the binding (with a 2.5-fold molar excess, nonspecific binding would also lead to the appearance of bis-adducts, which could not be seen with only a 1:1 ratio). The ligands tested were synthetic (BMS649 and CD3254) or natural (DHA, 9-*cis*-RA, and oleic acid) agonist ligands of HsRXR. 9-*cis*-RA, all-*trans*-RA, oleic acid, and DHA were diluted in 10 mM EtOH. BMS649 and CD3254 were diluted in 10 mM EtOH/DMSO (90:10, v/v). To prevent dissociation in the gas phase during the ionization and desorption process, the cone voltage was optimized to 5 V. Mass data were acquired in the positive ion mode on the m/z 1000–5000 mass range. Calibration of the instrument was performed using the multiply charged ions produced by a separate injection of horse heart myoglobin diluted to 2 μM in 1:1 water/acetonitrile (v/v) acidified with 1% (v/v) formic acid. The relative abundance of the different species present on ESI mass spectra were measured from their respective peak heights (37). All experiments were reproduced at least three times.

An important feature concerning the noncovalent MS approach is whether ESI mass spectra obtained in nondenaturing conditions in the gas phase of the mass spectrometer faithfully reflect direct in solution binding. Several control experiments were performed in order to argue in favor of structurally specific interactions. A series of "successive in solution competitions" strongly argue in favor of site-specific interactions. Instead of the simultaneous addition of a mixture of two ligands, as realized in the direct in solution competitions, successive ligand addition was achieved; a first ligand, L1, is bound to the protein (formation of a 1:1 RXR·L1 complex), and a second one, L2, serves to displace the equilibrium toward the formation of the RXR·L2 complex. Such experiments allow us to unambiguously assess that both ligands compete for the same binding site and that ESI-MS-observed interactions do not result from any artifactual gas phase aggregation. Finally, experiments with different binding partners (either different ligands or human *versus* amphioxus proteins) were also performed, leading to subsequent changes on ESI mass spectra.

Structure Determination—The homodimeric protein fractions were used for crystallization trials. The purified protein conditioned in 20 mM Tris, pH 8.0, 200 mM NaCl, 5 mM dithiothreitol and concentrated to 5 mg/ml. Crystals were grown by the hanging drop diffusion method and appeared in 4 days at 17 °C. The initial drop contained 5 mg/ml protein, 1.5 M sodium acetate, 50 mM *N*-(2-acetamido)iminodiacetic acid, pH 6.5, and was equilibrated *versus* a reservoir containing 1.8 M sodium acetate, 100 mM *N*-(2-acetamido)iminodiacetic acid, pH 6.5. Two different crystals forms, hexagonal and orthorhombic, were obtained in the same crystallization conditions. Only the orthorhombic form resulted in being suitable for structural determination. The solvent content of the crystals is 52.2% with

two LBD dimers (homotetramer) per asymmetric unit. Thus, the homodimeric protein may reassociate with time and salt concentration of the precipitant agent. Crystals were flash cooled in ethane after cryoprotection in a solution containing the reservoir solution plus 25% glycerol. A 2.7 Å resolution data set was obtained merging high and low resolution data sets from only one crystal at the BM29 Synchrotron line of the European Synchrotron Radiation Facility facility in Grenoble. Data were integrated and scaled using the DENZO-SCALEPACK package (38). The program CNS-SOLVE was used throughout the structure determination and refinement calculations (39). An initial phase estimate was obtained by molecular replacement using a partial structure of HsRXR dimer. After a rigid body refinement, the solution resulted in R_{cryst} of 34% and R_{free} of 35%. The model obtained was used to determine the NCS matrix of the four monomer molecules in the asymmetric unit in order to include averaging and solvent flattening procedure of density modification to calculate a first electron density map. The map unequivocally showed the missing parts of the model and allowed the complete manual building of the structure. The phase bias of the initial model was reduced by torsion angle molecular dynamics simulated annealing refinement. Subsequent refinement cycles alternated least square minimization and model building using the program O (40). The NCS was constrained in the first round of refinement and subsequently restrained and released. *B* factors were refined using two groups per residue. In the last cycle, 27 water molecules were located from a difference Fourier map. The loop between H1 and H3 (residues 281–289) and the lateral chains of the residues at positions 271–274, 299, 478, and 482–484 could not be constructed because of poor electron density. The quality of the final model quality was assessed with Procheck. Because of the absence of electron density for the loop between H1 and H3, the integrity of the protein in the crystals was checked by SDS-PAGE (data not shown). For the structure comparison, Cα traces of the models were superimposed using the *lsq* commands of O and default parameters. The figures were generated with Pymol. Crystallographic data and refinement statistics are summarized in supplemental Table 1. The chemical structures of all of the ligands described in this paper are shown in supplemental Fig. 1.

RESULTS

Residues That Contact 9-cis-Retinoic Acid Are All Conserved between AmphiRXR and HsRXR—Previous phylogenetical analysis clearly shows that, as expected, the amphioxus RXR sequence is positioned before the duplications giving rise to the three vertebrate genes (35) (supplemental Fig. 2). In addition, amphioxus RXR was shown to heterodimerize with amphioxus retinoic acid nuclear recepter or thyroid nuclear receptor and to bind to the relevant direct repeat sequence (41, 42).

To determine the crystal structure of the ligand-binding domain we selected a region that encompasses residues 266–484, leaving out the last 30 amino acid of the molecule that form an F-domain with no sequence identity with other NRs. Fig. 1 represents the structure-based sequence alignment of RXR LBDs from different organisms. The alignment reveals a high degree of sequence identity except for the length and composition of the loop between H1 and H3. Based on the structure of human RXRα, the residues that contact the 9-*cis*-RA are all conserved. The ability to bind this ligand seems thus to be preserved, and this is supported by mass spectrometry studies as well as by transient transfection assays (see below). The same degree of conservation is encountered also in the residues that form the hydrophobic groove in contact with the L*XX*LL motif activator helix (43). The only exception is a Val residue that in amphioxus RXR substitutes at position 306 a Phe, forming Van der Waals contacts with the coactivator helix. The amino acids of the dimerization interface are strictly conserved (19).

AmphiRXR Is Activated by 9-cis-RA in Cultured Cells but at Higher Concentrations than Vertebrate RXRs—Since the main residues implicated in ligand binding are conserved between AmphiRXR and vertebrates RXRs, we tested whether 9-*cis*-RA and DHA were able to activate AmphiRXR in transient tranfection assays in mammalian cells. To avoid any interference with endogenous RXR expressed in 293T cells, we used fusions between the Gal4 DNA-binding domain and RXR-LBD. As shown in Fig. 2*A*, both 9-*cis*-RA and all-*trans*-RA activate 2–4-fold Gal4-MmRXRα and Gal4-AmphiRXR. The effect of all-*trans*-RA at the concentration we used can be explained by its isomerization in 9-*cis*-RA (46). Nevertheless, in the experimental condition we used, we did not observe any reproducible effect of DHA (*lanes 4* and *8*). Of note, using strictly the same conditions as those described in the original paper (22), we observed that DHA activates MmRXRα but not AmphiRXR (not shown). Dose response experiments (Fig. 2*B*) show that 9-*cis*-RA activates AmphiRXR in a dose-dependent manner but with a range of concentration higher than mouse RXR. A weak activation by 9-*cis*-RA is also observed for the mollusk *Biomphalaria glabrata* RXR (45). Given that amphioxus is located at the base of chordates in phylogenetical tree, we can conclude from these data that the ability of RXR to bind and to be activated by 9-*cis*-RA is thus conserved in all chordates. This is in accordance with the conservation in amphioxus RXR of the amino acid residues implicated in ligand binding in vertebrate RXRs.

We thus explored the interaction between RXR and coactivators, such as TIF2. The activity of both Gal4-MmRXRα and Gal4-AmphiRXR was potentiated by the cotransfection of a human TIF2 expression vector (Fig. 2*B*). These findings were confirmed by pull-down assays (Fig. 2*C*) showing that *in vitro* amphioxus RXR, like human RXRα, can interact with TIF2 in a ligand-dependent manner. Interestingly, specific rexinoids, such as BMS649, induce the recruitment of TIF2, but not its mutant TIF2.1m123, mutated in the domain known to interact with NRs. Rexinoids are ligands that selectively bind and activate RXRs but not RARs in contrast to retinoids. This is in accordance with our observation that the recently sequenced amphioxus genome contains a unique homologue of the p160 co-activators (51). All of these observations suggest that (i) the mechanisms by which RXR activates transcription are conserved between amphioxus and mammals and (ii) the binding interface between AmphiRXR and TIF2 is functionally conserved.

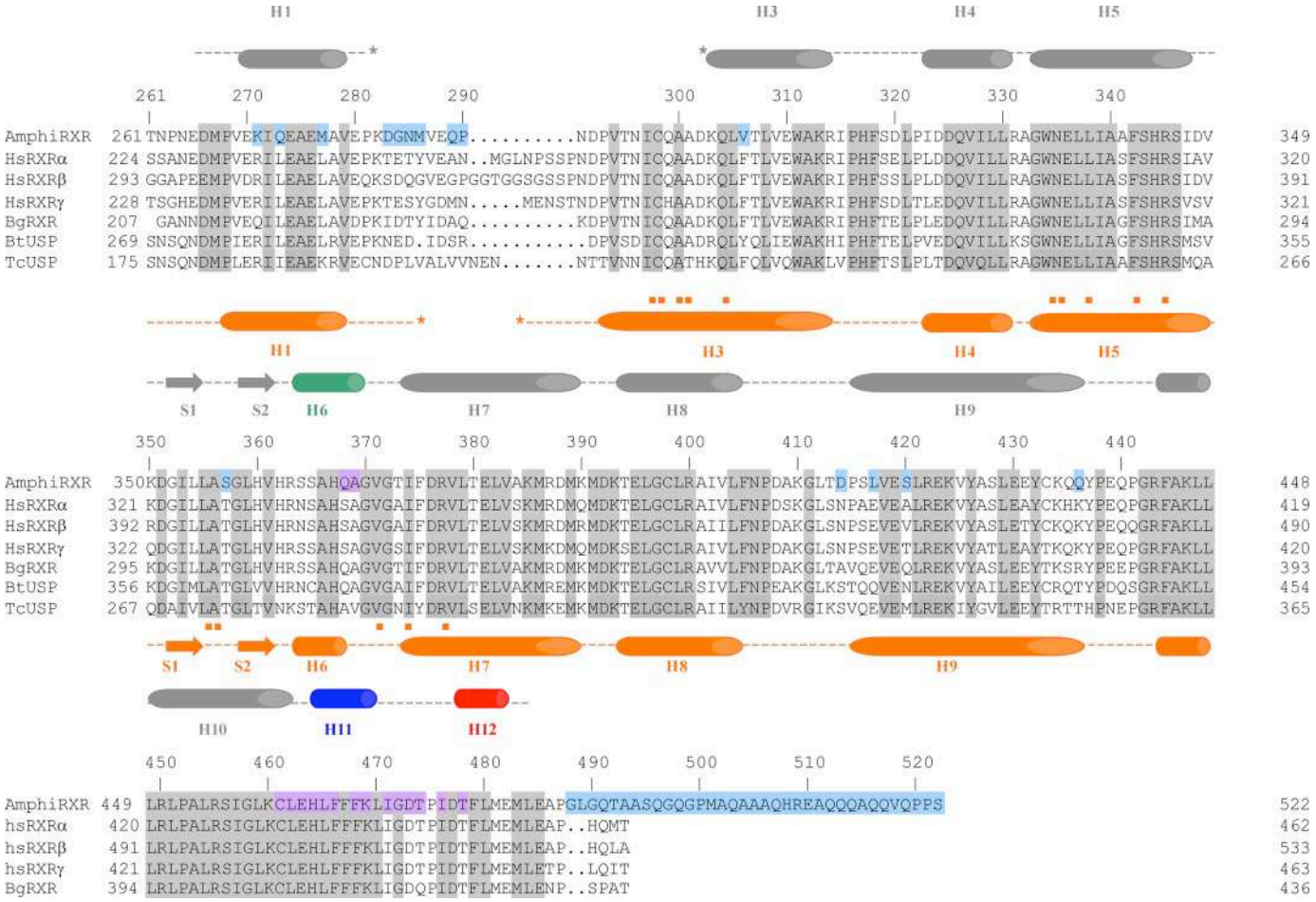


FIGURE 1. **Structure-based sequence alignment of the LBD domain of AmphiRXR.** The sequence of the AmphiRXR LBD is compared with those of human RXRα, RXRβ, and RXRγ as well as the USP from *B. tabaci*, *B. glabatra*, and *T. castaneum*. The residues in the *blue box* are unique to AmphiRXR. The secondary structure of the AmphiRXR and that of 9-*cis*-RA HsRXRα are reported on the *top* (in *gray*) and *bottom* (in *orange*) *lines*, respectively. Helix 6, helix 11, and loop 11–12 and helix 12 of AmphiRXR are shown in *green*, *blue*, and *red*, respectively. The residues forming the pocket in HsRXRα are indicated by the *orange dots*. The residues of AmphiRXR involved in the tetramer interface are shown in *purple*.

AmphiRXR Binds Retinoid Ligands in Vitro—To complement transfection assays, direct evidence for ligand binding to AmphiRXR was given by ESI-MS performed in nondenaturing conditions. Indeed, this particular MS application has proven its efficiency for the characterization of protein-ligand complexes in general and NR LBDs-ligand complexes especially (46–49). No fortuitous ligand was copurified with AmphiRXR. In order to evaluate relative affinities of 9-*cis*-RA, DHA, BMS649, oleic acid, and CD3254 (supplemental Fig. 1) for AmphiRXR LBD in solution titration and competition experiments were monitored by ESI-MS. After fine tuning of the instrumental set-up (see "Experimental Procedures") in order to avoid dissociation of weak/hydrophobic AmphiRXR/ligand assemblies, titration experiments involving increasing ligand concentrations revealed that all previously mentioned molecules bind to AmphiRXR LBD. Fig. 3 depicts the ESI mass spectra obtained after the addition of a 2.5-fold molar excess (50 μM) of oleic acid (*B*) 9-*cis*-RA (*C*) or BMS649 (*D*) to AmphiRXR LBD (20 μM). Peak heights of the different species observed on the ESI mass spectra were considered in order to evaluate relative bound/free protein ratios (37) (see Fig. 3*E*). In strictly identical experimental and instrumental conditions, only 30% of the AmphiRXR·9-*cis*-RA complexes were detected, whereas up to 80% of the AmphiRXR-ligand complexes were observed for BMS649 and CD3254. Increasing the incubation time from 15 min to 24 h does not change the relative abundances. In the presence of all-*trans*-RA, no ligand binding was observed (data not shown). These results indicate that the synthetic rexinoids CD3254 and BMS649 have higher affinities for AmphiRXR than 9-*cis*-RA or oleic acid or DHA, and a relative affinity ranking can be deduced (Fig. 3*E*). For comparison, similar experiments were performed with HsRXRα LBD, leading always to higher bound/free protein ratios (Fig. 3*E*). The fact that 9-*cis*-RA presents a higher affinity than DHA for HsRXRα is in agreement with the observations of Lengqvist *et al.* (48, 50). Affinity rankings established by titration experiments for both RXR homologues were further confirmed by direct in solution competition experiments involving a mixture of equimolar amounts of ligands (10 μM each) and 10 μM RXR (data not shown). Altogether, these data demonstrate that retinoid ligands that bind to HsRXR also bind to AmphiRXR but with a lower affinity, which is also a proof of the plasticity of the ligand binding property during evolution. The weakest ESI-MS-derived affinity for the tested ligands for AmphiRXR corresponds also to the weakest activation observed in *in vitro* activity assays, which suggests that those ligands are not the natural

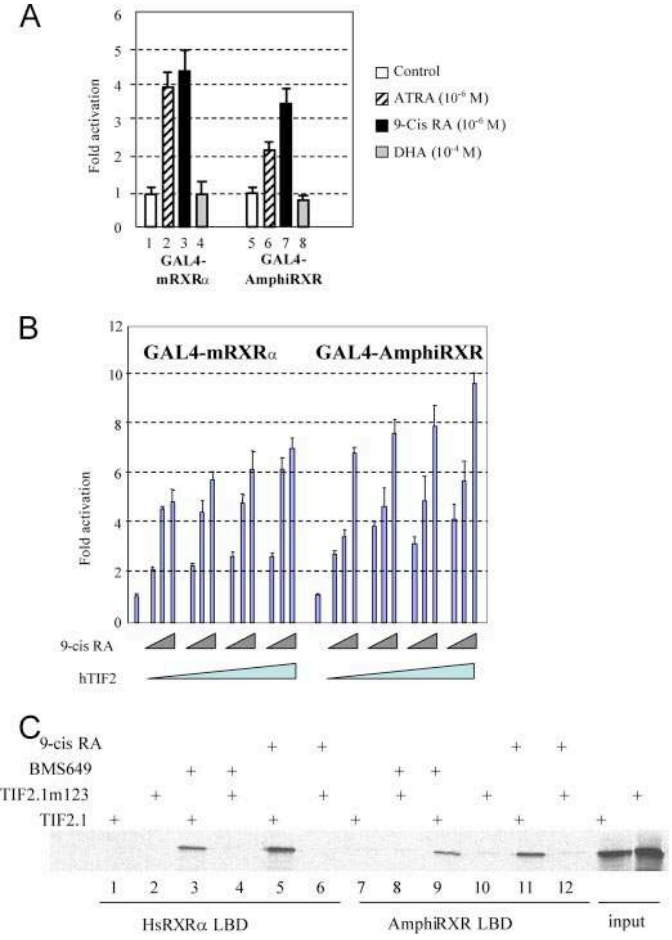


FIGURE 2. **AmphiRXR is activated by 9-*cis*-RA in cultured cells.** *A*, cells were transfected with Gal4-MmRXRα or Gal4-AmphiRXR along with 17m5-tk-luc reporter plasmid and induced with ligands. *B*, cells were cotransfected with Gal4-MmRXRα or Gal4-AmphiRXR and increasing amounts of hTIF2 (10, 100, 250, and 500 ng), along with the 17m5-tk-luc reporter plasmid, and induced with increasing amounts of ligands (10^{-9}, 10^{-8}, and 10^{-7} M for MmRXRα; 10^{-8}, 10^{-7}, and 10^{-6} M for AmphiRXR). *C*, purified His-HsRXRα or His-AmphiRXR was incubated in the presence of 2 μM BMS649 or 2 μM 9-*cis*-RA and with *in vitro* translated [^{35}S]methionine TIF2.1 or TIF2.1m123, as indicated. Input corresponds to 100% of the TIF2 used in the assays.

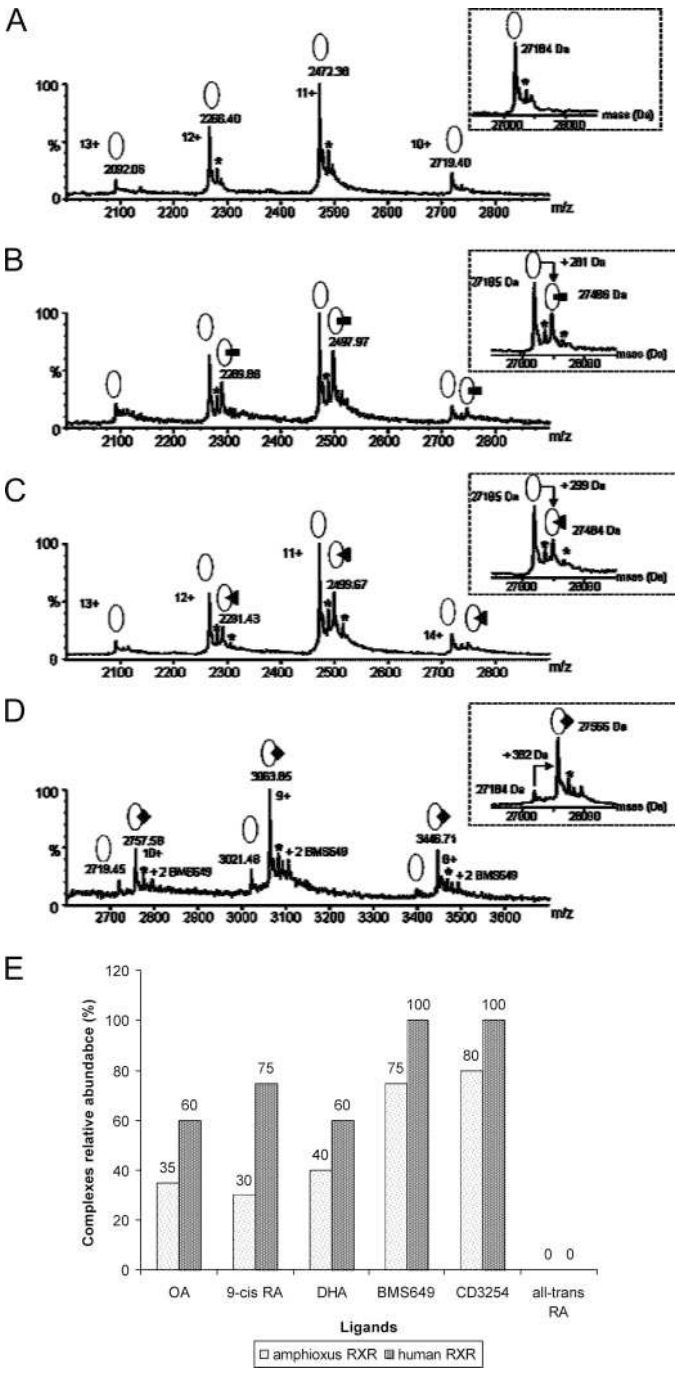


FIGURE 3. **AmphiRXR LBD binds retinoid ligands.** ESI mass spectra of AmphiRXR LBD protein (*A*) apo and after incubation with a 2.5-fold molar excess of oleic acid (*B*), of 9-*cis*-RA (*C*), and of BMS649 (*D*). Deconvulated mass spectra are presented for each ligand as an *inset*. *E*, distribution plot of the relative proportion of bound/free protein for AmphiRXR and HsRXRα. Specific binding of both ligands to AmphiRXR is observed with higher affinities for the synthetic ligands CD3254 and BMS649.

ligand, if any, for AmphiRXR. No explanation of the difference in the affinity between amphioxus and human RXR can be found by looking at the residues forming the ligand-binding pocket.

Oligomeric Structure—The asymmetric unit contains two dimers related by a nearly perfect noncrystallographic 2-fold axis (Fig. 4). Within each dimer, the classical dimerization interface (Fig. 4*B*) is identical to that observed in homodimeric structures of the apo- and holo-HsRXRs (29–34) or USPs (27, 28). It involves helix 10 and, to a lesser extent, helix 9 and the loop between helices 7 and 8 of each monomer. The buried surface of each monomer is 1162 Å^2 in comparison with 1147 Å^2 for HsRXR. The tetramer is stabilized by a bottom to bottom organization of dimers reminiscent of the previously reported crystal structures of unliganded RXR (32) but without the intermolecular exchange of helices 12. In the apo-HsRXR (32) and in the apo ascidian RXR tetramer (Protein Data Bank code 2Q60), in addition to the canonical dimer interface, the tetramer interface involves interaction between H3-H3, H11-H11, and H12 that protrudes from one monomer into the coactivator binding site of an adjacent monomer. In the holotetramer structure of the mollusk BgRXR (25), within each dimer, one monomer is in open conformation and one is in a closed holo conformation. The open monomers show a swapping of H12 as observed in the apo-HsRXR tetramer, and additional tetramer interfaces are formed between H11 of each closed monomer and H6 of the open monomer in the adjacent canonical dimer (32). In

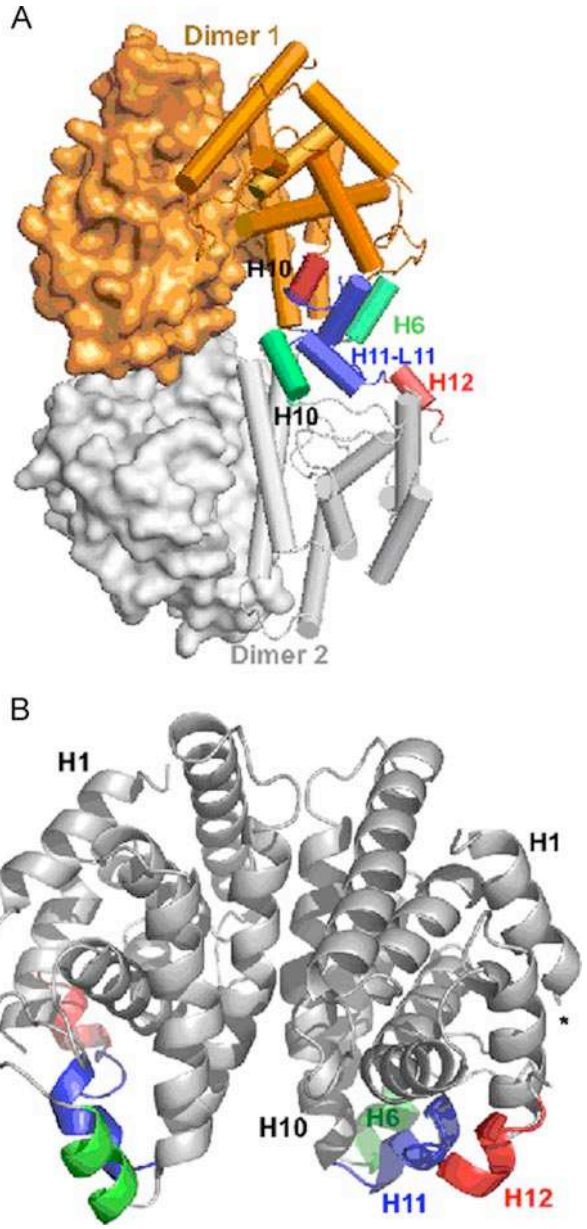


FIGURE 4. **Structure of the AmphiRXR LBDs.** *A*, structure of the tetramer with one dimer shown in *orange* and the other one in *gray*. Within each canonical dimer, one monomer is shown in a *surface representation*, and in the other one, the helices and the β-sheet are represented by *cylinders* and *arrows*, respectively. Helix 6 is shown in *green*, helix 11 and loop 11–12 are in *blue*, and helix 12 is in *red*. The same *color code* is used in all figures. *B*, the AmphiRXR homodimer H12 (*red*) is positioned in the coactivator binding cleft. The missing loop connecting H1 to H3 not visible in the electron density map is shown by *stars*.

AmphiRXR tetramer, the tetrameric interface formed between the two dimers is constituted by residues spanning between helix 10 and helix 12 of one dimer interacting with residues H11 and of H6 of the other. This interface involves mainly Van der Waals contacts and one hydrogen bond between Asp-473 and Lys-468 in the loop H11-H12 of another monomer (Fig. 4). It buries 653 Å^2 of the surface of each monomer, in contrast to 1215 Å^2 per monomer for HsRXR, suggesting that the tetramer might actually not be very stable in solution.

AmphiRXR Is in an Inactive Conformation—Interestingly, the LBD of AmphiRXR exhibits an antagonist-like conformation with a distorted H12 bound in the coactivator groove. Mass spectrometry showed that AmphiRXR does not copurify with any bound ligand. The crystal structure confirms this result, showing no density for a putative small molecule bound. The LBP, as observed in holo-HsRXR, is filled with the partially unfolded H11 (Fig. 5*B*). As a consequence of the conformation of H11, the N-terminal part of H3 is released from its contact to H11. This relaxed structure results in a more flexible H1-H3 connection that translates in a local disorder and explains the absence of clear density for this connection.

Comparison with Mammalian RXRs—The superimposition of the crystal structures of apo- and holo-HsRXRα and AmphiRXR LBDs (Fig. 5, *A* and *B*) highlights the unfolding and the reposition of H11 that occupies most of the LBP. In apo-HsRXRα, H11 results from a kink at position 434 (HsGlu-434 in HsRXRα) inducing an 80° twist of the C terminus with respect to the H10 helical axis. As a consequence of this movement, HsLeu-441, HsPhe-437, and HsPhe-438 fill the fraction of the LBP cavity normally occupied by the β-ionone ring of the 9-*cis*-RA ligand (30). In the case of AmphiRXR, H11 is partially unfolded and occupies a larger section of the same pocket (Figs. 5*B* and 6). The main differences between the two structures are a consequence of the new position of the shorter H11, which in AmphiRXR begins at position Glu-463 (HsGlu-434). The first 6 or 7 residues are not too far in space and present similar orientations of their side chains with Phe-466, Phe-467, Leu-470, and Ile-471, pointing toward the inside core of the protein (Fig. 5). In AmphiRXR, Leu-470 carbonyl forms a hydrogen bond with Gln-304 at the C terminus of H3. The Ile-471 side chain contacts Gln-304 and H5. These interactions are stabilized by a network of hydrogen bonds involving main chain carbonyl and amino groups in a pseudohelical conformation (Leu-465–Phe-468, Phe-466–Lys-469, and Phe-467–Leu-470). The C-terminal ends of H11 and H12 have lost their helical conformation to reach the coactivator binding cleft between H3 and H4. As a result, the stabilizing interaction formed with the N-terminal part of H3 disappears. The first 10 residues of H3 and the connection H1-H3 are not visible in the electron density map. This could be the consequence of a static disorder in the crystal, due to a higher flexibility of these fragments and/or a partial unfolding of H3.

A comparison with the holo form of HsRXRα bound to 9-*cis*-RA (Fig. 5*B*) underlines some conformational changes that may be associated with an eventual ligand binding. In the liganded HsRXRα, five residues of H5 contact the ligand; all of these residues are conserved in AmphiRXR. Three of these residues keep the same conformation in the absence of the ligand (Asn-335 (HsAsn-306), Leu-338 (HsLeu-309), and Phe-342 (HsPhe-313)) when Trp-334 (HsTrp-305) and Arg-345 (HsArg-316) point in the opposite direction with respect to the pocket (Fig. 6). Arg-345 (HsArg-316) that in holo-HsRXRα forms ionic interactions with the carboxylate group of ligands now points to the solvent. The repositioned lateral chain of Trp-334 is part of a network of interactions that involves Ile-471 of H11, Asp-473 peptidic chain, and Thr-474. These two side chains would easily reorient toward the "bound ligand" pocket.

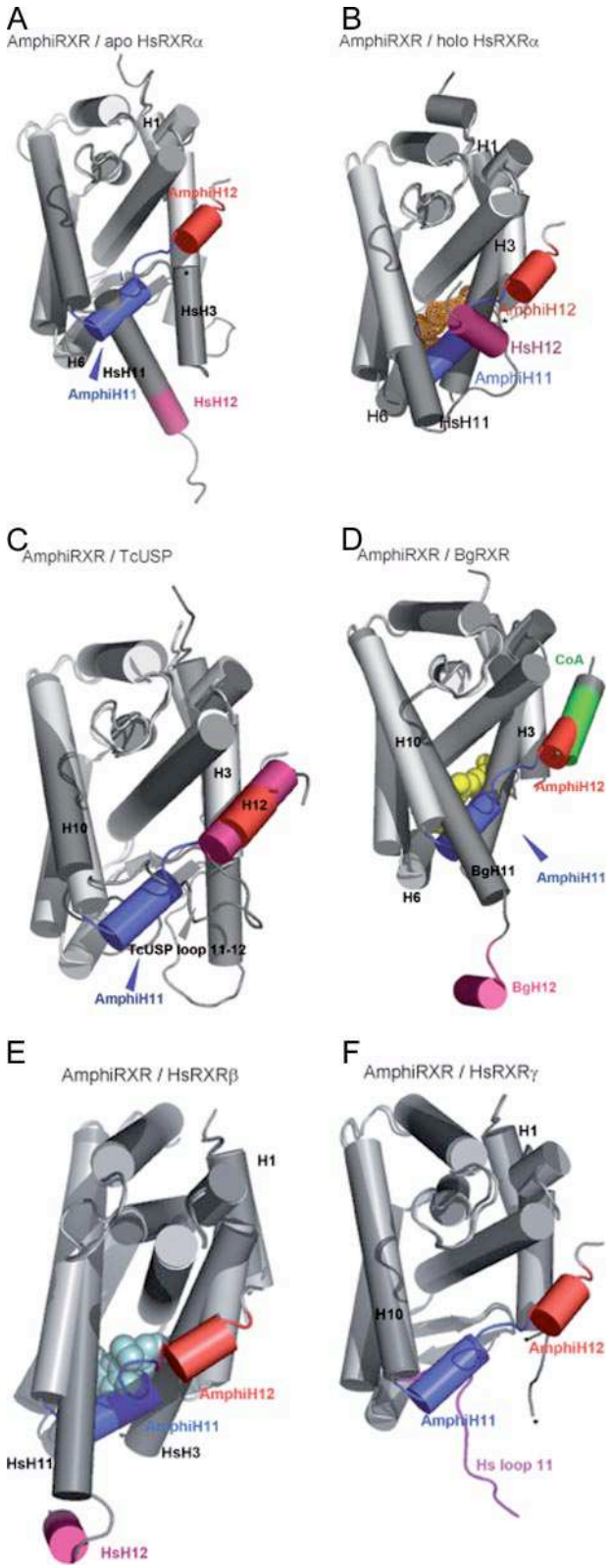


FIGURE 5. **Stable apo antagonist conformation of AmphiRXR.** Shown is the superimposition of AmphiRXR (in *gray*) over apo-HsRXRα (in *dark gray*) (*A*), holo-HsRXRα (in *dark gray*) (*B*), and TcUSP (in *dark gray*) (*C*), showing the close three-dimensional similarity of AmphiRXR and TcUSP, notably the conformations of H11 and H12, the open monomer BgRXR (in *dark gray*) (*D*), HsHXRβ-LG100268 (Protein Data Bank code 1H9U) (in *dark gray*) (*E*), and HsRXRγ (Protein Data Bank code 2GL8) (in *dark gray*) (*F*). The missing loop 1–3 of AmphiRXR is shown by *stars*. Helices 11 and 12 of AmphiRXR are shown in *blue* and *red*, respectively. Helices H12 of HsRXRα, TcUSP, BgRXR, and HsRXRβ and loop 11–12 of HsRXRγ are *colored* in *pink*. The ligands in holo-HsRXRα, -BgRXR, and -HsRXRβ are shown by *orange surface*, and *yellow* and *blue spheres*, respectively. The coactivator peptide in BgRXR is shown in *green*.

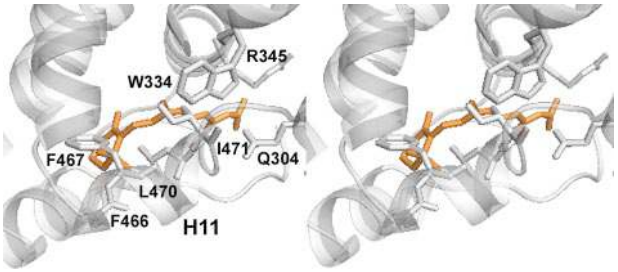


FIGURE 6. **Ligand binding cavity of HsRXRα.** Shown is a *stereoview* of the residues filling the AmphiRXR pocket superposed to 9-*cis*-RA (in *orange*), as observed in the holo-HsRXRα crystal structure.

The two other isotypes of HsRXR have also been crystallized, HsRXRβ as a homodimer in complex with the rexinoid LG100268 (33) (Fig. 5*E*) or a heterodimer with LXRα and in complex with methoprene acid and coactivator peptide (34) and HsRXRγ as an apotetramer (Protein Data Bank code 2GL8) (Fig. 5*F*). In the homodimer HsRXRβ structure, the C-terminal activation helix H12 does not adopt the agonist active conformation and protrudes outside the core of the LBD stabilized by contacts with another dimer in the crystal packing, whereas the rest of the structure is similar to the holo agonist HsRXRα structure. The apo-HsRXRγ tetramer shows an unfolding of H11 and H12 and as a consequence a released N-terminal H3 and absence of density of H1–H3 and the beginning of H3 similar to AmphiRXR. The variability in the position of H12 underlines its flexibility and the lack of a stable "functional" conformation in absence of ligands or cofactors.

Comparison with Invertebrates RXRs—Another invertebrate RXR, the RXR from the mollusk *B. glabrata* has been crystallized in complex with 9-*cis*-retinoic acid and coactivator peptide as a tetramer. In the BgRXR tetramer, within each dimer, one monomer is in open conformation, and one is in a closed holo conformation (25). Both monomers interact similarly with 9-*cis*-RA and coactivator peptide. The open monomers show a swapping of H12, as observed in the apo-HsRXRα tetramer. Although the closed conformation of BgRXR monomer is similar to that of the holo-HsRXRα structure, in the open conformation of BgRXR, the end of H11 and H12 protrudes outside of the core of the LBD despite the presence of ligand and coactivator peptide (Fig. 5*C*). The main differences with the apo-AmphiRXR structure are the position of the C-terminal H11 and H12 and also the position of H3, which in BgRXR contacts the 9-*cis*-retinoic acid and closes the ligand-binding pocket.

The superimposition of AmphiRXR on USPs from *Bemisia tabaci* (BtUSP) (27) and from *Tribolium castaneum* (TcUSP) (28) emphasizes some striking similarities with these two molecules. In all cases, H11 occupies part of the LBP (Fig. 5*D*). The main differences concern H6 and H11. In BtUSP and TcUSP, H6 occupies most of the pocket, whereas in AmphiRXR, the LBP is partially occupied by H11. Initially, the absence of H1 and the N terminus of H3 due to proteolysis in BtUSP (27) could provide an explanation for the observed conformational changes of H11 and H12. However, the structure of TcUSP (28)

and the present structure clearly establish the functional relevance and the more general character of this apo structure.

DISCUSSION

Although RXR is an early component of the mechanism of regulation of metazoans, the retinoid signaling though RXR is still enigmatic (18, 20). The existence and the necessity of specific endogenous ligands that activate RXR are still controversial (18, 20, 21, 22, 28). Several crystal structures of mammalian RXRs and invertebrate RXRs have revealed an inherent flexibility of H12. Natural retinoid ligands have been shown to be inefficient to reposition H12 in HsRXRβ or BgRXR (25, 33, 34). Thus, H12 is only weakly recruited to the canonical active position upon binding of ligands. To learn more about this receptor and to understand the vertebrate-specific character of RXRs, we have studied the RXR ligand-binding domain of the cephalochordate amphioxus, an invertebrate chordate evolutionarily close to vertebrates. Here we report the crystal structure of a novel apotetramer conformation of AmphiRXR LBD. This structure shows some similarities with the crystal structures of the arthropod RXR/USPs that were not observed in mammalian RXR structures. Our functional data demonstrate that AmphiRXR shows weak binding and activation by retinoid ligands. Like the invertebrate BgRXR, AmphiRXR is not activated by retinoids at physiological concentrations. These observations support the proposal of a functional role of this stable apo conformation of AmphiRXR. It is of interest that an AmphiRXR-like conformation is present in insect USPs in functional heterodimers with the ecdysone nuclear receptor (BtEcR/BtUSP and TcEcR/TcUSP) in the agonist conformation and the activation RXR helix in the coactivator binding cleft (antagonist conformation) underlying a silent role of RXR. The capability of AmphiRXR to bind with low affinity retinoids and coactivators, a capability observed for the other invertebrate BgRXR and not observed for TcUSP or BtUSP, may reflect the possibility of a sensor role of the retinoid ligand in the molecular activation of the heterodimeric partner. In that respect, AmphiRXR is similar to BgRXR, and given the current knowledge of the metazoan phylogeny, this may suggest that this corresponds to an ancestral function of RXR (that should be found, for example, in *Tripedalia*) that has been modified in vertebrates and insects (Fig. 7). The role of RXR-specific signaling thus needs to be addressed in further studies. Of note, the amphioxus is a particularly suitable model system for such a study, since it allows treatments with specific agonist or antagonist compounds that will allow dissecting the *in vivo* functional role of this receptor (26). In addition, in amphioxus, several partners of RXR, such as retinoic acid nuclear receptor or thyroid nuclear receptor, have been characterized both in terms of *in vivo* function and in terms of pharmacology (35, 42). Finally the amphioxus genome contains only one gene for each type of receptor (*i.e.* only one RXR, one retinoic acid nuclear receptor, and one thyroid nuclear receptor), thus providing a much simpler system than vertebrates in which genetic redundancy and paralogue-specific sequence divergence impede the understanding of the basal function of a given receptor type (51). Thus, it should be possible using amphioxus to study in more detail the molecular dialogue that may exist between

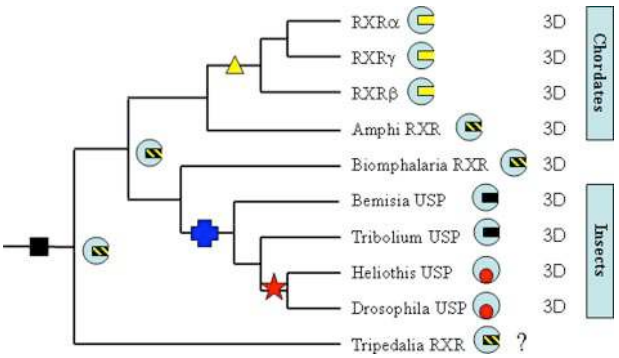


FIGURE 7. **Schematized tree of RXRs highlighting the evolutionary and structural plasticity of the ligand-binding domain.** Vertebrate RXRs are depicted with an empty binding pocket (*yellow*) and can readily be activated by low doses of ligands. For AmphiRXR (this study) and mollusk RXR (*Biomphalaria*; see Ref. 25), the ligand-binding pocket is partially filled with H11 residues, but the receptor can still be able to bind a ligand at high doses. A loss of this ligand binding ability occurred at the base of insects (28), giving rise to non-*Mecopterida* USPs (in *Bemisia* and *Tribolium*) that do not bind ligand. A further evolutionary change in *Mecopterida* (*Drosophila* and *Heliothis*) linked to rapid sequence divergence produced receptors exhibiting a large binding pocket filled with a phospholipid. We propose that the situation found in amphioxus and *Biomphalaria* will also be found in cnidarian RXR (*Tripedalia*; see Ref. 23), but this remains to be shown. *Colored symbols* indicate functional shifts. *Black square*, early event of gain of ligand binding (15). *Blue cross*, loss of ligand binding at the base of insects (28) (note that we still ignore precisely when this event occurred). *Red star*, gain of a structural ligand in *Mecopterida* (28). *Yellow triangle*, shift of H11 and acquisition of high affinity ligand binding in vertebrates (this study). *3D*, the available crystal structure of LBDs.

RXR, its heterodimeric partners, and the transcriptional coactivators.

In this paper, we show that the amphioxus RXR LBD exhibits a striking organization, since it adopts an apo antagonist conformation with a peculiar conformation of helix H11 filling the ligand-binding pocket and shows some similarities with the non-*Mecopterida* insect RXRs in which unfolded H6 and H11 fill the ligand-binding pocket. In addition, our functional data show that AmphiRXR, like the vertebrates and mollusk RXRs, is able to bind and be activated by RXR ligands but less efficiently than vertebrate RXRs. The AmphiRXR LBD is thus structurally similar to the insect USP LBDs, whereas it is functionally related to vertebrate RXRs, since it can be regulated by the same ligands, although with a weaker efficiency. Thus, our data suggest that amphioxus RXR is, functionally, an intermediate between arthropod RXR/USPs and vertebrate RXRs. We propose that in most metazoans, RXR is similar to the situation found in amphioxus or mollusk, and from this situation, two different types of RXR evolved: (i) the ligand-independent form in non-*Mecopterida* insects (further modified later in *Mecopterida*) and (ii) the vertebrate receptors that reinforced their ability to be regulated by ligand binding (*yellow star* in Fig. 7). Such an event is likely to have occurred before the whole genome duplication events that produced the three paralogous receptors in vertebrates, RXRα, -β, and -γ, since the three paralogous receptors have identical ligand-binding pockets and exhibit the same ligand binding activities (Fig. 7). Nevertheless, it has to be emphasized that almost no functional data are available from RXRs in nonmammalian vertebrates, such as fishes, or even early vertebrates, such as lamprey and hagfishes. It is only through careful structural and functional analysis of these

receptors that a precise model of ligand binding evolution of chordate RXRs will be determined.

According to the strong conservation of the residues lining the putative canonical LBP, one would predict that all RXR-like LBD, including those from insects, such as the hemipteran *Bemisia* (BtUSP) or the coleopteran *Tribolium* (TcUSP), would bind most of the known ligands of HsRXR. This contrasts with the experimental observations in *Tribolium* USP, which does not bind and is not activated by any RXR ligands (25). Of note, these observations have been reinforced by an evolutionary analysis of the pattern of substitution rate in the ligand-binding domains of RXRs and USPs from a wide variety of metazoans (25). Indeed, these data indicate that in TcUSP or BtUSP, residues that would belong to a potential RXR-like LBP exhibit high evolutionary rates, suggesting a relaxation of the evolutionary pressure in this region that may be related to a loss in ligand binding capability. AmphiRXR appears to be in a paradoxical situation when compared with the receptors from insects, since it adopts a similar apo conformation (with H12 in an antagonist position and the ligand-binding pocket filled by H11 residues), whereas it is able to bind to RXR-specific ligands, such as 9-*cis*-retinoic acid, BMS649, or oleic acid. Of note, *in silico* modeling analysis cannot provide a simple steric explanation of this observation, since a homology model that would accommodate ligands like 9-*cis*-RA or fatty acids can easily be built for both the insect receptors and AmphiRXR. Additional experiments are necessary to elucidate the mechanism of ligand binding and more generally the question of the necessity of a ligand for this orphan receptor. RXR provides a very clear example showing that sequence conservation is a necessary prerequisite for the prediction of functional properties, such as ligand binding, but is not sufficient alone to reach a conclusion.

Taken together, the data available on RXRs illustrate the remarkable evolutionary plasticity of this LBD that can adapt to different functional shifts, such as changes in the ligand binding abilities, together with a conservation of other functions, such as heterodimerization. This certainly calls for other comparative structural and functional studies of this receptor that will serve to illustrate basic concepts in NR function, such as the role of ligand regulation, the definition of endogenous ligands, or the structural basis of ligand affinity.

REFERENCES

1. Gronemeyer, H., Gustafsson, J. A., and Laudet, V. (2004) *Nat. Rev. Drug Discov.* **3,** 950–964
2. Gustafsson, J. A. (1999) *Science* **284,** 1285–1286
3. Kliewer, S. A., Lehmann, J. M., and Willson, T. M. (1999) *Science* **284,** 757–760
4. Benoit, G., Cooney, A., Giguere, V., Ingraham, H., Lazar, M., Muscat, G., Perlmann, T., Renaud, J. P., Schwabe, J., Sladek, F., Tsai, M. J., and Laudet, V. *Pharmacol. Rev.* (2006) **58,** 798–836
5. Laudet, V., and Gronemeyer, H. (2002) *The Nuclear Receptor Facts Book*, Academic Press, London
6. Wurtz, J. M., Bourguet, W., Renaud, J. P., Vivat, V., Chambon, P., Moras, D., and Gronemeyer, H. (1996) *Nat. Struct. Biol.* **3,** 87–94
7. Renaud, J. P., and Moras, D. (2000) *Cell Mol. Life Sci.* **57,** 1748–1769
8. Renaud, J. P., Rochel, N., Ruff, M., Vivat, V., Chambon, P., Gronemeyer, H., and Moras, D. (1995) *Nature* **378,** 681–689
9. Kallenberger, B. C., Love, J. D., Chatterjee, V. K., and Schwabe, J. W. (2003) *Nat. Struct. Biol.* **10,** 136–140
10. Robinson-Rechavi, M., Carpentier, A. S., Duffraisse, M., and Laudet, V. (2001) *Trends Genet.* **17,** 554–556
11. Laudet, V., Hänni, C., Coll, J., Catzeflis, F., and Stéhelin, D. (1992) *EMBO J.* **11,** 1003–1013
12. Escriva, H., Bertrand, S., and Laudet, V. (2004) *Essays Biochem.* **40,** 11–26
13. Escriva, H., Safi, R., Hänni, C., Langlois, M. C., Saumitou-Laprade, P., Stehelin, D., Capron, A., Pierce, R., and Laudet, V. (1997) *Proc. Natl. Acad. Sci. U. S. A.* **94,** 6803–6808
14. Laudet, V. (1997) *J. Mol. Endocrinol.* **19,** 207–226
15. Escriva, H., Delaunay, F., and Laudet, V. (2000) *BioEssays* **22,** 717–727
16. Germain, P., Staels, B., Dacquet, C., Spedding, M., and Laudet, V. (2006) *Pharmacol. Rev.* **58,** 685–704
17. Germain, P., Chambon, P., Eichele, G., Evans, R. M., Lazar, M. A., Leid, M., De Lera, A. R., Lotan, R., Mangelsdorf, D. J., and Gronemeyer, H. (2006) *Pharmacol. Rev.* **58,** 712–725
18. Germain, P., Chambon, P., Eichele, G., Evans, R. M., Lazar, M. A., Leid, M., De Lera, A. R., Lotan, R., Mangelsdorf, D. J., and Gronemeyer, H. (2006) *Pharmacol. Rev.* **58,** 760–772
19. Brelivet, Y., Kammerer, S., Rochel, N., Poch, O., and Moras, D. (2004) *EMBO Rep.* **5,** 423–429
20. Calléja, C., Messaddeq, N., Chapellier, B., Yang, H., Krezel, W., Li, M., Metzger, D., Mascrez, B., Ohta, K., Kagechika, H., Endo, Y., Mark, M., Ghyselinck, N. B., and Chambon, P. (2006) *Genes Dev.* **20,** 1525–1538
21. Mic, F. A., Molotkov, A., Benbrook, D. M., and Duester, G. (2003) *Proc. Natl. Acad. Sci. U. S. A.* **100,** 7135–7140
22. de Urquiza, A. M., Liu, S., Sjöberg, M., Zetterström, R. H., Griffiths, W., Sjövall, J., and Perlmann, T. (2000) *Science* **290,** 2140–2144
23. Kostrouch, Z., Kostrouchova, M., Love, W., Jannini, E., Piatigorsky, J., and Rall, J. E. (1998) *Proc. Natl. Acad. Sci. U. S. A.* **95,** 13442–13447
24. Bonneton, F., Zelus, D., Iwema, T., Robinson-Rechavi, M., and Laudet, V. (2003) *Mol. Biol. Evol.* **20,** 541–553
25. de Groot, A., de Rosny, E., Juillan-Binard, C., Ferrer, J. L., Laudet, V., Pierce, R. J., Pebay-Peyroula, E., Fontecilla-Camps, J. C., and Borel, F. (2005) *J. Mol. Biol.* **354,** 841–853
26. Schubert, M., Escriva, H., Xavier-Neto, J., and Laudet, V. (2006) *Trends Ecol. Evol.* **21,** 269–277
27. Carmichael, J. A., Lawrence, M. C., Graham, L. D., Pilling, P. A., Epa, V. C., Noyce, L., Lovrecz, G., Winkler, D. A., Pawlak-Skrzecz, A., Eaton, R. E., Hannan, G. N., and Hill, R. J. (2005) *J. Biol. Chem.* **280,** 22258–22269
28. Iwema, T., Billas, I. M., Beck, Y., Bonneton, F., Nierengarten, H., Chaumot, A., Richards, G., Laudet, V., and Moras, D. (2007) *EMBO J.* **26,** 3770–3782
29. Bourguet, W., Ruff, M., Chambon, P., Gronemeyer, H., and Moras, D. (1995) *Nature* **375,** 377–382
30. Egea, P. F., Mitschler, A., Rochel, N., Ruff, M., Chambon, P., and Moras, D. (2000) *EMBO J.* **19,** 2592–2601
31. Bourguet, W., Vivat, V., Wurtz, J. M., Chambon, P., Gronemeyer, H., and Moras, D. (2000) *Mol Cell.* **5,** 289–298
32. Gampe, R. T., Jr., Montana, V. G., Lambert, M. H., Wisely, G. B., Milburn, M. V., and Xu, H. E. (2000) *Genes Dev.* **14,** 2229–2241
33. Love, J. D., Gooch, J. T., Benko, S., Li, C., Nagy, L., Chatterjee, V. K., Evans, R. M., and Schwabe, J. W. (2002) *J. Biol. Chem.* **277,** 11385–11391
34. Svensson, S., Ostberg, T., Jacobsson, M., Norström, C., Stefansson, K., Hallén, D., Johansson, I. C., Zachrisson, K., Ogg, D., and Jendeberg, L. (2003) *EMBO J.* **22,** 4625–4633
35. Escriva, H., Holland, N. D., Gronemeyer, H., Laudet, V., and Holland, L. Z. (2002) *Development* **12,** 2905–2916
36. Voegel, J. J., Heine, M. J., Zechel, C., Chambon, P., and Gronemeyer, H. (1996) *EMBO J.* **15,** 3667–3675
37. Peschke, M., Verkerk, U. H., and Kebarle, P. (2004) *J. Am. Soc. Mass Spectrom.* **15,** 1424–1434
38. Otwinowski, Z., and Minor, W. (1997) *Methods Enzymol.* **276,** 307–326
39. Brünger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J. S., Kuszewski, J., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T., and Warren, G. L. (1998) *Acta Crystallogr. Sect. D* **54,** 905–921
40. Jones, T. A., Zou, J. Y., Cowan, S. W., and Kjeldgaard, M. (1991) *Acta Crystallogr. Sect. A* **47,** 110–119
41. Escriva, H., Bertrand, S., Germain, P., Robinson-Rechavi, M., Umbhauer, M., Cartry, J., Duffraisse, M., Holland, L., Gronemeyer, H., and Laudet, V.

(2006) *PLoS Genet.* **2,** e102
42. Paris, M., Escriva, H., Schubert, M., Brunet, F., Brtko, J., Ciesielski, F., Roecklin, D., Vivat-Hannah, V., Jamin, E. L., Cravedi, J. P., Scanlan, T. S., Renaud, J. P., Holland, N. D., and Laudet, V. (2008) *Curr. Biol.* **18,** 825–830
43. Egea, P. F., Mitschler, A., and Moras, D. (2002) *Mol Endocrinol.* **16,** 987–997
44. Deleted in proof
45. Bouton, D., Escriva, H., de Mendonça, R. L., Glineur, C., Bertin, B., Noël, C., Robinson-Rechavi, M., de Groot, A., Cornette, J., Laudet, V., and Pierce, R. J. (2005) *J. Mol. Endocrinol.* **34,** 567–582
46. Bovet, C., Wortmann, A., Eiler, S., Granger, F., Ruff, M., Gerrits, B., Moras, D., and Zenobi, R. (2007) *Protein Sci.* **16,** 938–946
47. Sanglier, S., Bourguet, W., Germain, P., Chavant, V., Moras, D., Gronemeyer, H., Potier, N., and Van Dorsselaer, A. (2004) *Eur. J. Biochem.* **271,** 4958–4967
48. Lengqvist, J., Mata De Urquiza, A., Bergman, A. C., Willson, T. M., Sjövall, J., Perlmann, T., and Griffiths, W. (2004) *J. Mol. Cell Proteomics* **3,** 692–703
49. Bitsch, F., Aichholz, R., Kallen, J., Geisse, S., Fournier, B., and Schlaeppi, J. M. (2003) *Anal. Biochem.* **323,** 139–149
50. Lengqvist, J., Alvélius, G., Jörnvall, H., Sjövall, J., Perlmann, T., and Griffiths, W. J. (2005) *J. Am. Soc. Mass Spectrom.* **16,** 1631–1640
51. Holland, L. Z., Albalat, R., Azumi, K., Benito-Gutiérrez, E., Blow, M. J., Bronner-Fraser, M., Brunet, F., Butts, T., Candiani, S., Dishaw, L. J., Ferrier, D. E., Garcia-Fernàndez, J., Gibson-Brown, J. J., Gissi, C., Godzik, A., Hallböök, F., Hirose, D., Hosomichi, K., Ikuta, T., Inoko, H., Kasahara, M., Kasamatsu, J., Kawashima, T., Kimura, A., Kobayashi, M., Kozmik, Z., Kubokawa, K., Laudet, V., Litman, G. W., McHardy, A. C., Meulemans, D., Nonaka, M., Olinski, R. P., Pancer, Z., Pennacchio, L. A., Pestarino, M., Rast, J. P., Rigoutsos, I., Robinson-Rechavi, M., Roch, G., Saiga, H., Sasakura, Y., Satake, M., Satou, Y., Schubert, M., Sherwood, N., Shiina, T., Takatori, N., Tello, J., Vopalensky, P., Wada, S., Xu, A., Ye, Y., Yoshida, K., Yoshizaki, F., Yu, J. K., Zhang, Q., Zmasek, C. M., de Jong, P. J., Osoegawa, K., Putnam, N. H., Rokhsar, D. S., Satoh, N., and Holland, P. W. (2008) *Genome Res.* **18,** 1100–1111

Supplementary Figure1: Chemical structures of some retinoids.

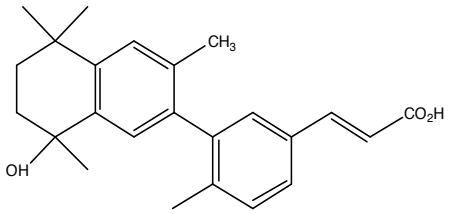
Supplementary Figure2: Phylogenetic tree of RXR and USP. A maximum likelihood (ML) tree was obtained from the analysis of the amino acid sequences of the DBD and LBD of all metazoan RXR and USP sequences available from Genbank with PhyML under a JTT + γ + I model. Bootstrap percentages after 1000 replicates are shown. Nodes with bootstrap support below 50 % were collapsed. The AmphiRXR sequence from Branchiostoma floridae is indicated in red. The tree was rooted with other NR2 (HNF4 and COUP-TF) as well as some NR3 (ER) sequences. The bar indicates the number of changes per sites.

Supplementary Table 1

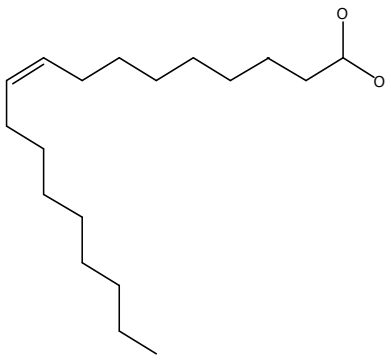
Crystallographic data and refinement statistics

Space group	$P2_12_12_1$
Cell Dimensions (Å)	a=78.72, b=96.12, c=131.34
Resolution (last shell) (Å)	15-2.79 (2.89-2.79)
Completeness (last shell)	90.3% (75%)
Rsym (last shell)	0.082 (0.20)
<I/σ(I)> (last shell)	5.4 (2.5)
Redundancy	4.5
Rwork (no. of reflections)	20.2% (21644)
Rfree (no. of reflections)	25.7% (1167)
No. of protein atoms	6207
No. of solvent molecules	49
r.m.s.d. bond length (Å)	0.016
r.m.s.d. bond angle (°)	1.67
Average B factor (Å^2)	56.5

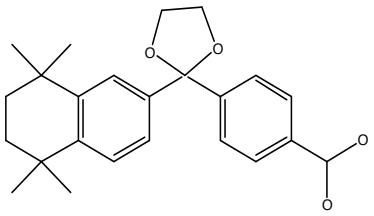
Figure S1: Chemical structures of some retinoids.



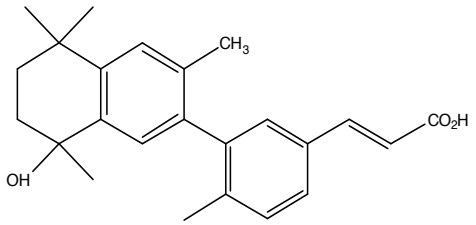
9cis retinoic acid



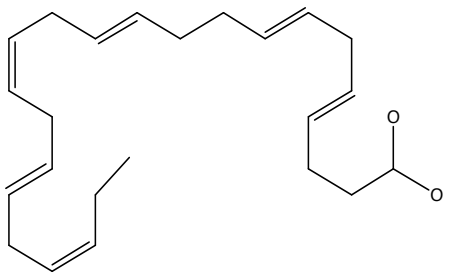
Oleic acid



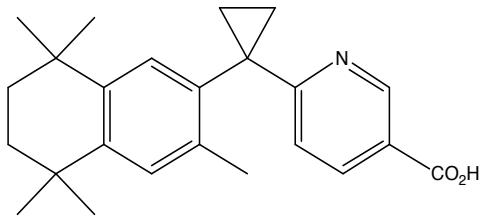
BMS649



CD3254



DHA

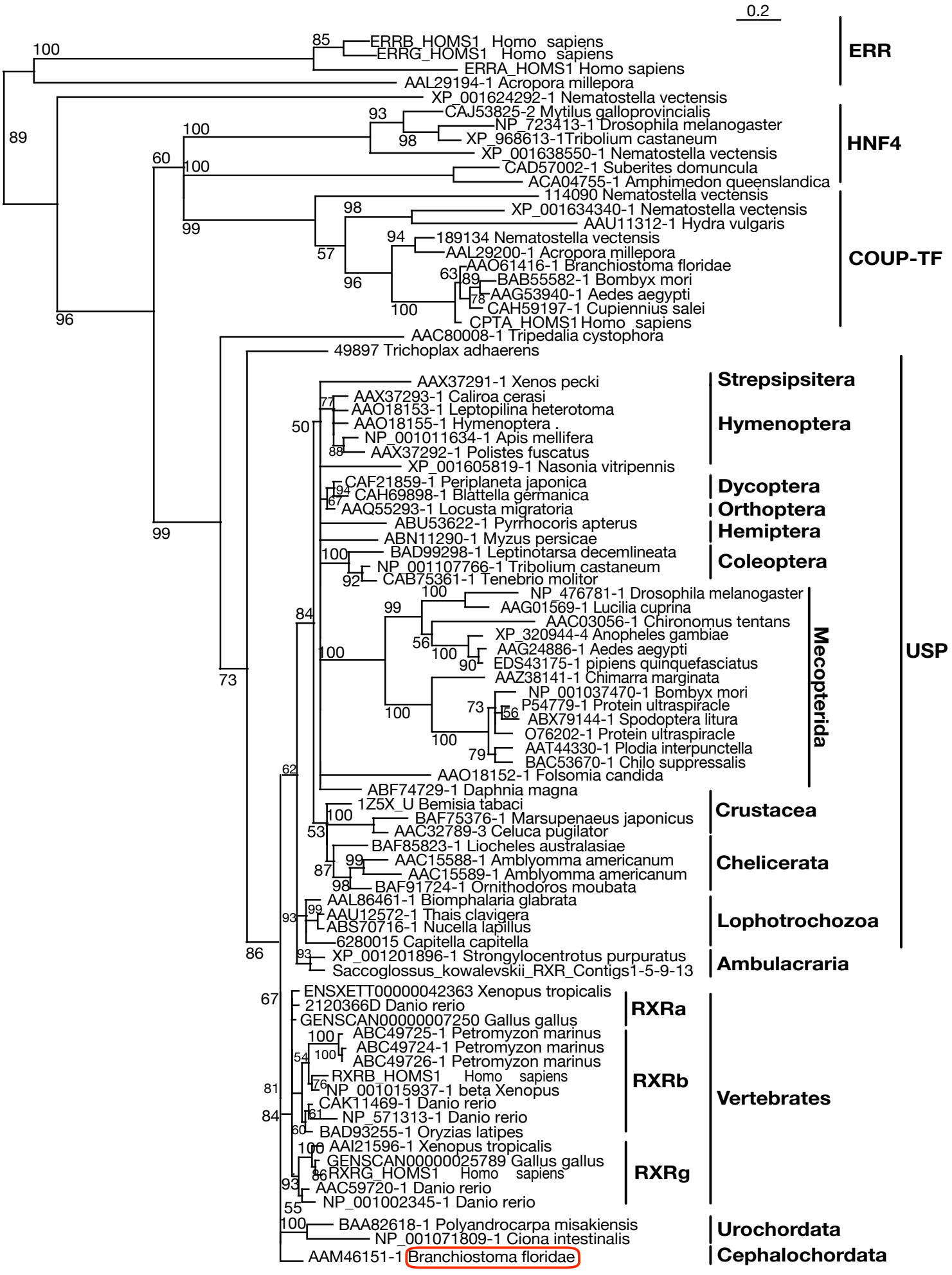


LG268

BMS649: 4-[2-(5,5,8,8-tetramethyl-5,6,7,8-tetrahydro- naphthalen-2-yl)-[1,3]dioxolan-2-yl]-benzoic acid

CD3254: 3-[4-Hydroxy-3-(3,5,5,8,8-pentamethyl-5,6,7,8-tetrahydro-naphthalen-2-yl)-phenyl]-acrylic acid

LG268: 6-[1-(3,5,5,8,8-pentamethyl-5,6,7,8-tetrahydronaphtalen-2-yl)cyclopropyl]pyridine-3-carboxylic acid


0.2
ERRB_HOMS1 Homo sapiens
ERRG_HOMS1 Homo sapiens
ERRA_HOMS1 Homo sapiens
AAL29194-1 Acropora millepora
XP_001624292-1 Nematostella vectensis
CAJ53825-2 Mytilus galloprovincialis
NP_723413-1 Drosophila melanogaster
XP_968613-1Tribolium castaneum
XP_001638550-1 Nematostella vectensis
CAD57002-1 Suberites domuncula
ACA04755-1 Amphimedon queenslandica
114090 Nematostella vectensis
XP_001634340-1 Nematostella vectensis
AAU11312-1 Hydra vulgaris
189134 Nematostella vectensis
AAL29200-1 Acropora millepora
AAO61416-1 Branchiostoma floridae
BAB55582-1 Bombyx mori
AAG53940-1 Aedes aegypti
CAH59197-1 Cupiennius salei
CPTA_HOMS1 Homo sapiens
ERR
HNF4
COUP-TF
AAC80008-1 Tripedalia cystophora
49897 Trichoplax adhaerens
AAX37291-1 Xenos pecki
AAX37293-1 Caliroa cerasi
AAO18153-1 Leptopilina heterotoma
AAO18155-1 Hymenoptera .
NP_001011634-1 Apis mellifera
AAX37292-1 Polistes fuscatus
XP_001605819-1 Nasonia vitripennis
CAF21859-1 Periplaneta japonica
CAH69898-1 Blattella germanica
AAQ55293-1 Locusta migratoria
ABU53622-1 Pyrrhocoris apterus
ABN11290-1 Myzus persicae
BAD99298-1 Leptinotarsa decemlineata
NP_001107766-1 Tribolium castaneum
CAB75361-1 Tenebrio molitor
NP_476781-1 Drosophila melanogaster
AAG01569-1 Lucilia cuprina
AAC03056-1 Chironomus tentans
XP_320944-4 Anopheles gambiae
AAG24886-1 Aedes aegypti
EDS43175-1 pipiens quinquefasciatus
AAZ38141-1 Chimarra marginata
NP_001037470-1 Bombyx mori
P54779-1 Protein ultraspiracle
ABX79144-1 Spodoptera litura
O76202-1 Protein ultraspiracle
AAT44330-1 Plodia interpunctella
BAC53670-1 Chilo suppressalis
AAO18152-1 Folsomia candida
ABF74729-1 Daphnia magna
1Z5X_U Bemisia tabaci
BAF75376-1 Marsupenaeus japonicus
AAC32789-3 Celuca pugilator
BAF85823-1 Liocheles australasiae
AAC15588-1 Amblyomma americanum
AAC15589-1 Amblyomma americanum
BAF91724-1 Ornithodoros moubata
AAL86461-1 Biomphalaria glabrata
AAU12572-1 Thais clavigera
ABS70716-1 Nucella lapillus
6280015 Capitella capitella
XP_001201896-1 Strongylocentrotus purpuratus
Saccoglossus_kowalevskii_RXR_Contigs1-5-9-13
ENSXETT00000042363 Xenopus tropicalis
2120366D Danio rerio
GENSCAN00000007250 Gallus gallus
ABC49725-1 Petromyzon marinus
ABC49724-1 Petromyzon marinus
ABC49726-1 Petromyzon marinus
RXRB_HOMS1 Homo sapiens
NP_001015937-1 beta Xenopus
CAK11469-1 Danio rerio
NP_571313-1 Danio rerio
BAD93255-1 Oryzias latipes
AAI21596-1 Xenopus tropicalis
GENSCAN00000025789 Gallus gallus
RXRG_HOMS1 Homo sapiens
AAC59720-1 Danio rerio
NP_001002345-1 Danio rerio
BAA82618-1 Polyandrocarpa misakiensis
NP_001071809-1 Ciona intestinalis
AAM46151-1 Branchiostoma floridae
Strepsipsitera
Hymenoptera
Dycoptera
Orthoptera
Hemiptera
Coleoptera
Mecopterida
Crustacea
Chelicerata
Lophotrochozoa
Ambulacraria
RXRa
RXRb
RXRg
Vertebrates
Urochordata
Cephalochordata
USP

Gene Loss and Evolutionary Rates Following Whole-Genome Duplication in Teleost Fishes

Frédéric G. Brunet, Hugues Roest Crollius,† Mathilde Paris,* Jean-Marc Aury,‡ Patricia Gibert,§ Olivier Jaillon,‡ Vincent Laudet,* and Marc Robinson-Rechavi‖*

*Laboratoire de Biologie Moléculaire de la Cellule, INRA LA 1237, CNRS UMR5161, IFR 128 BioSciences Lyon-Gerland, Ecole Normale Supérieure de Lyon, Lyon, France; †Dyogen Group, CNRS UMR8541, Ecole Normale Supérieure, Paris, France; ‡Structure et évolution des génomes, CNRS UMR8030, Genoscope, Evry, France; §Laboratoire de Biométrie et Biologie Evolutive, CNRS UMR5558, Université Claude Bernard Lyon 1, Villeurbanne, France; and ‖Department of Ecology and Evolution, Biophore, University of Lausanne, Lausanne, Switzerland

Teleost fishes provide the first unambiguous support for ancient whole-genome duplication in an animal lineage. Studies in yeast or plants have shown that the effects of such duplications can be mediated by a complex pattern of gene retention and changes in evolutionary pressure. To explore such patterns in fishes, we have determined by phylogenetic analysis the evolutionary origin of 675 *Tetraodon* duplicated genes assigned to chromosomes, using additional data from other species of actinopterygian fishes. The subset of genes, which was retained in double after the genome duplication, is enriched in development, signaling, behavior, and regulation functional categories. The evolutionary rate of duplicate fish genes appears to be determined by 3 forces: 1) fish proteins evolve faster than mammalian orthologs; 2) the genes kept in double after genome duplication represent the subset under strongest purifying selection; and 3) following duplication, there is an asymmetric acceleration of evolutionary rate in one of the paralogs. These results show that similar mechanisms are at work in fishes as in yeast or plants and provide a framework for future investigation of the consequences of duplication in fishes and other animals.

Introduction

Whole-genome duplication (WGD) is expected to have a large impact on the evolution of lineages in which it has occurred. Understanding this impact necessitates unraveling a complex network of causes and consequences, including which genes are retained in duplicate, and how the duplication modifies their evolutionary constraints. The most progress on these and related issues has been made in the yeast *Saccharomyces cerevisiae* (Kellis et al. 2004; Davis and Petrov 2005; Scannell et al. 2006) as well as in the plant *Arabidopsis thaliana* (Blanc and Wolfe 2004; Maere et al. 2005). Studies in animals are less advanced, may be because the species that are the easiest to study did not experience such genome duplication (fruit flies and nematodes).

Three WGD events have been proposed in ancient vertebrate history: 2 at the origin of the group and a third specific to fishes (Meyer and Van de Peer 2005). Although the 2 rounds of duplication at the origin of vertebrates remain controversial (but see Dehal and Boore 2005), fishes have provided the first clear demonstration of ancient genome duplication in vertebrate evolution (Jaillon et al. 2004; Woods et al. 2005). More recent tetraploids are known in many vertebrate lineages (Otto and Whitton 2000), but they do not give insight into the longterm consequences of duplication and cannot explain the eventual success of tetraploid lineages, in the way that more ancient duplications do. Thus, fishes provide a unique platform to study genome duplication in vertebrates.

A genome duplication was originally suggested in ray-finned fishes based on the finding that zebrafish and medaka possess 7 Hox clusters (Amores et al. 1998; Wittbrodt et al. 1998; Naruse et al. 2000), against 4 in mammals and 1 in most invertebrates, as well as by comparative mapping (Postlethwait et al. 2000), but it was only confirmed with the release of 2 pufferfishes' genome sequences. First, 2 studies identified many short duplicated groups of linked genes (paralogons) in the *Takifugu rubripes* genome and dated duplications with a molecular clock to a window between divergence of ray-finned fishes from tetrapods and diversification of teleost fishes (Christoffels et al. 2004; Vandepoele et al. 2004). Second, all chromosomes of *Tetraodon nigroviridis* were assigned to paralogons, demonstrating the genome scale of the duplication, and each pair of paralogons was shown to be homologous to 1 (nonduplicated) human chromosomal region (Jaillon et al. 2004). It has been estimated that 76% to 80% of paralogs were secondarily lost after the WGD (Jaillon et al. 2004; Woods et al. 2005).

When did this duplication occur? Direct dating of fish gene duplications has relied on molecular clock methods with large error bars, using 2 closely related Tetraodontiformes (*Tetraodon* and *Takifugu*). Molecular clock dating can be biased by saturation of synonymous changes at the timescales considered as well as by changes in evolutionary rates between mammals and fishes and in duplicate genes (Robinson-Rechavi and Laudet 2001; Van de Peer et al. 2001; Jaillon et al. 2004; Jordan et al. 2004; Venkatesh et al. 2005). Less sensitive to these problems, phylogenies of a few tens of gene families have shown a high frequency of gene duplications to be a common feature among sampled teleosts, but not among other fishes (Robinson-Rechavi et al. 2001; Hoegg et al. 2004), confirmed by the distribution of Hox complexes (Crow et al. 2006). Finally, comparative mapping shows that paralogons are homologous between *Tetraodon* and zebrafish (a Cypriniforme), confirming that the WGD occurred before the divergence of these 2 lineages of teleosts (Woods et al. 2005).

The aim of this study is to provide a framework for understanding the consequences of WGD in teleost

Key words: genome duplication, gene loss, neofunctionalization, Gene Ontology, evolutionary rates, selection.

E-mail: marc.robinson-rechavi@unil.ch.

Mol. Biol. Evol. 23(9):1808–1816. 2006
doi:10.1093/molbev/msl049
Advance Access publication June 29, 2006

Table 1
Distribution of *Tetraodon* Gene Duplications by Phylogenetic Relative Dating

Chromosome Distribution	Ancestral Vertebrate Duplication	Ancestral Fish-Specific Duplication			Recent Duplication	Total
		Strong Evidence[a]	Unrooted[b]	Weak Support[c]		
Same	14	12	11	2	10	49
Different	247	213	120	20	26	626
All	261		378		36	675

NOTE.—Classification of all genes that have at least 2 paralogs in the *Tetraodon nigroviridis* genome, which share the same human ortholog, and for which both paralogs are mapped to *T. nigroviridis* chromosomes. Duplications are considered ''ancestral to vertebrates'' if they occurred in vertebrates before the actinopterygian/tetrapod divergence (fig. 1*A*). Duplications are considered ''ancestral to fishes'' if at least 1 order of fishes outside percomorphs shares the duplication (fig. 1*B*). There are 447 gene phylogenies that could not be classified because of low phylogenetic resolution or because only *Tetraodon* and *Takifugu* sequences were available among fishes, making phylogenetic dating impossible.

[a] Phylogenies with a clear outgroup, Neighbor-Joining bootstrap support over 75%, and consistent results with different methods.

[b] Phylogenies with good support, as in the footnote ''a,'' but no good outgroup sequences available for rooting.

[c] Phylogenies with low bootstrap support (under 75%) in Neighbor-Joining or different results depending on methods.

fishes. Duplicate genes were characterized by phylogeny, evolutionary rates, and functional classification. We used the 2 published fish genomes as well as other available fish data.

Materials and Methods

We combined predicted genes from genome sequences and gene sequences from databases. Genes from the freshwater pufferfish *T. nigroviridis* were from the published version of the genome (Jaillon et al. 2004) (8.2X coverage). Genes from the pufferfish *T. rubripes* (Aparicio et al. 2002) and the unpublished genome of the zebrafish *Danio rerio* were from Ensembl (Hubbard et al. 2002), versions 22.2c and 22.3b, respectively. These genomic data were complemented with all actinopterygian fish proteins from the Hovergen "clean prot" database version 44 (Duret et al. 1994). This database classifies vertebrate genes in homologous families, excluding partial sequences with less than 80% coverage of the alignment of complete proteins, and provided us directly with protein alignments to compare with the genes from genomes. In addition, human and mouse proteins were recovered from Swissprot and its complement TrEMBL (Boeckmann et al. 2003), as outgroups to actinopterygian fishes. Amino acid sequences were aligned using ClustalW (Thomson et al. 1994), followed by manual adjustments if needed.

We defined groups of paralogs as 2 or 3 *Tetraodon* genes that have the same human best hit from Swissprot, using Blast with the settings used for "exofish" gene detection (Roest Crollius et al. 2000; Jaillon et al. 2004); "best hit" was defined according to Blast score (Altschul et al. 1997). To reinforce phylogenetic with mapping analysis, we used only groups with at least 2 genes mapped to *Tetraodon* chromosomes. This defines a set of 704 pairs with both paralogs assigned to chromosomes, plus 419 triplets with at least 2 paralogs assigned to chromosomes, making a total of 1,123 gene families to investigate. For each corresponding amino acid alignment, a preliminary phylogenetic tree was estimated by Neighbor-Joining (Saitou and Nei 1987) with rate heterogeneity between sites corrected for by a gamma law of parameter alpha set to 0.8 and 500 bootstrap replicates. Phylogenies were then refined with regard to species sampling and partial sequences in Phylo_win (Galtier et al. 1996) and interpreted manually. All phylogenies are available in the Supplementary Material online.

The rate of gene loss after genome duplication was computed using the subset of genes that are mapped to *Tetraodon* chromosomes and have a clear orthology relation between human and *Tetraodon*. Orthology was defined either from the phylogeny or by reciprocal best Blast hit, for cases of one-to-one homologs. There are 2,371 such human genes, of which 364 have 2 *Tetraodon* co-orthologs, which were shown in our analysis to be duplicated in the fish-specific genome duplication. This gives a rate of gene retention after the duplication of 15% (364/2,371). The figure of 378 in table 1, instead of 364, includes fish genes for which human orthology is not clear.

For evolutionary rate computations, amino acid alignments were extended to the nucleotide coding sequences. For each pair, the number of synonymous substitutions per synonymous site, Ks, and the number of nonsynonymous substitutions per nonsynonymous site, Ka (Li 1993; Pamilo and Bianchi 1993), were calculated using the "diverge" option of the GCG software. We removed pairs with null Ka or Ks values and pairs for which Ka or Ks were not computable because sites were too few. We also manually removed predicted proteins containing clear frameshifts. Thus, evolutionary rates were computed for 254 WGD and 1,666 singleton genes (categories defined in fig. 2). For evolutionary rate computations, we updated the predicted *Takifugu* genes to version 31.2 from Ensembl, and we added human and mouse genes from Ensembl version 36.

Human orthologs of mapped duplicated *Tetraodon* genes, as determined by phylogenetic analysis, were compared for over- or underrepresentation of Gene Ontology (GO) terms (Ashburner et al. 2000), using GOToolBox (Martin et al. 2004), with the hypergeometric test, and the Benjamini and Hochberg (1995) correction for multiple testing. The reference data set was composed of all human orthologs of mapped *Tetraodon* genes.

Results

Phylogenetic and Chromosomal Classification of Duplicated Genes

We identified 1,123 pairs or triplets of *Tetraodon* paralogs with the same human best hit and for which at

least 2 genes were mapped to *Tetraodon* chromosomes. We conducted a phylogenetic analysis of each of the corresponding gene families, of which 675 could be classified with respect to gene duplication in fishes. Phylogenetic reconstruction was used to date duplication events, relatively to speciation events (fig. 1), as opposed to molecular clock dating. Phylogenies of duplicate genes can be biased by rate differences when few species and simple methods are used (Van de Peer et al. 2002; Fares et al. 2006). Although we can never guarantee the accuracy of all reconstructions, care was taken to maximize taxonomic sampling, to take into account the rate heterogeneity among sites, and to check automatic alignment results. Altogether, we analyzed sequences from 112 species of ray-finned fishes, sampling 13 lineages (Percomorpha + 12 actinopterygian orders). The most notable contribution was from zebrafish, which is included in 90% of the phylogenies; 79% of phylogenies include only zebrafish and the 2 Tetraodontiformes. We also checked for long-branch attraction a la Fares et al. (2006) by comparing evolutionary rates: there is no significant difference between the "ancestral vertebrate" paralogs and the "ancestral fish-specific" paralogs (not shown). Among paralogs that diverged in the teleost fishes (table 1), most duplicated before the divergence of main teleost fish lineages (e.g., fig. 1*B*). We take these pairs to be the best markers of the teleost fish–specific WGD in the following analyses.

A previous classification of *Tetraodon* paralogs, based on a simple synonymous molecular clock (Jaillon et al. 2004), provided evidence for paralogons, that is, large duplicated regions that often comprise whole chromosomes. We repeated this analysis, but using exclusively the paralogs whose period of duplication has been phylogenetically verified to be fish specific (Supplementary Figure I, Supplementary Material online). This confirms the coverage of the *Tetraodon* genome by paralogons in a robust manner. Two-thirds of all ancient fish-specific paralogs are part of a duplicated block of at least 5 genes. Conversely, recent duplicates are the only class of paralogs that are frequently neighbors on the same chromosome (table 1).

Consistency between chromosomal order and phylogeny supports the validity of our phylogenetic analysis in identifying the paralogs from WGD (Fares et al. 2006). This also provides a verification of the "ancestral vertebrate" paralogs (table 1): 80% are not included in fish-specific paralogons, consistent with a separate origin. Interestingly, these "ancestral vertebrate" paralogs define 6 duplicated blocks (of at least 5 genes), which are distinct from the fish-specific paralogons; these may be remnants of ancestral vertebrate genome duplication (Dehal and Boore 2005).

We verified orthology relations between zebrafish and *Tetraodon* chromosomes (Supplementary Figure II, Supplementary Material online): most zebrafish duplicated linkage groups are indeed orthologous to *Tetraodon* paralogons. The duplicated blocks from 1 species for which we did not identify an homologous duplicated block in the other species are those defined by the smallest number of gene pairs. Thus, the combined phylogenetic and mapping analysis confirms that the paralogons covering *Tetraodon* and zebrafish chromosomes find their origin in a WGD that took place before the divergence of these species. This is consistent with the findings of Woods et al. (2005), based on linkage mapping.

Genome duplication in fishes was followed by massive gene loss because most genes in the *Tetraodon* or *Takifugu* genomes are not found in characteristic gene pairs. Out of 2,371 human genes with a clear orthology relation to mapped *Tetraodon* genes, 364 have a phylogenetic profile typical of the WGD in *Tetraodon* (see Materials and Methods). This gives an estimate of 15% of gene pairs retained after the WGD, that is, loss of one of the paralogs in 85% of pairs. This is slightly higher than the estimates of 76% to 80% reached by more approximate methods (Jaillon et al. 2004; Woods et al. 2005). It is possible that 85% is still an underestimate because we neglect the fast-evolving genes for which a human ortholog is not detected and which may be more likely to lose a paralog after duplication. The exact figure of gene loss remains to be determined, ideally by comparison to a closer nonduplicated genome, such as that of nonteleost fish. To improve our understanding of the process of gene loss and retention, we compared the characteristics of paralogs from the WGD to those of nonduplicated genes.

Evolutionary Rates of Duplicate Genes

We first compared the substitution rates between *Tetraodon* and *Takifugu* orthologs. As they all diverged at the time of the *Tetraodon/Takifugu* speciation, we can directly convert differences between evolutionary distances (Ka and Ks) into differences between evolutionary rates. We see 2 clear trends (fig. 2*A*): WGD paralogs differ from singleton genes and there is asymmetry between the 2 paralogs from WGD. The most significant differences are in selective pressure, as measured by the Ka/Ks ratio. But these differences can be due either to changes in selective patterns after duplication, or to bias in the genes kept after duplication, or to a combination of both. To distinguish these possible causes, we tested rates that are phylogenetically independent of the duplication.

Comparing the evolutionary rates between mouse and human orthologs of the fish genes, we find results similar to the slower WGD paralogs in pufferfishes, with notably lower Ka in the orthologs of pufferfish genes that were kept in double (fig. 2*A* and *B*). Mammalian orthologs of fish WGD pairs evolve 23% slower than orthologs of singletons. These differences in evolutionary rates between orthologs of singletons and those of duplicates cannot be due to the fish WGD. Instead, they show that the fish genes that were already under the strongest selective pressure were retained as duplicates after the genome duplication.

Another feature to take into account is that teleost fish genes seem to evolve faster, under weaker constraint, than mammalian genes (Robinson-Rechavi and Laudet 2001; Jaillon et al. 2004; Steinke et al. 2006). Indeed, in our data, the Ka/Ks ratio is systematically higher in pufferfishes than in mammalian orthologs (fig. 2*A*). The difference is highly significant for singleton genes and fast-evolving paralogs (paired *t*-test: $P < 2.2 \times 10^{-16}$) but not for slow-evolving paralogs. The latter are thus under relatively high selective pressure, typical of mammalian rather than pufferfish genes. Of note, the divergence time

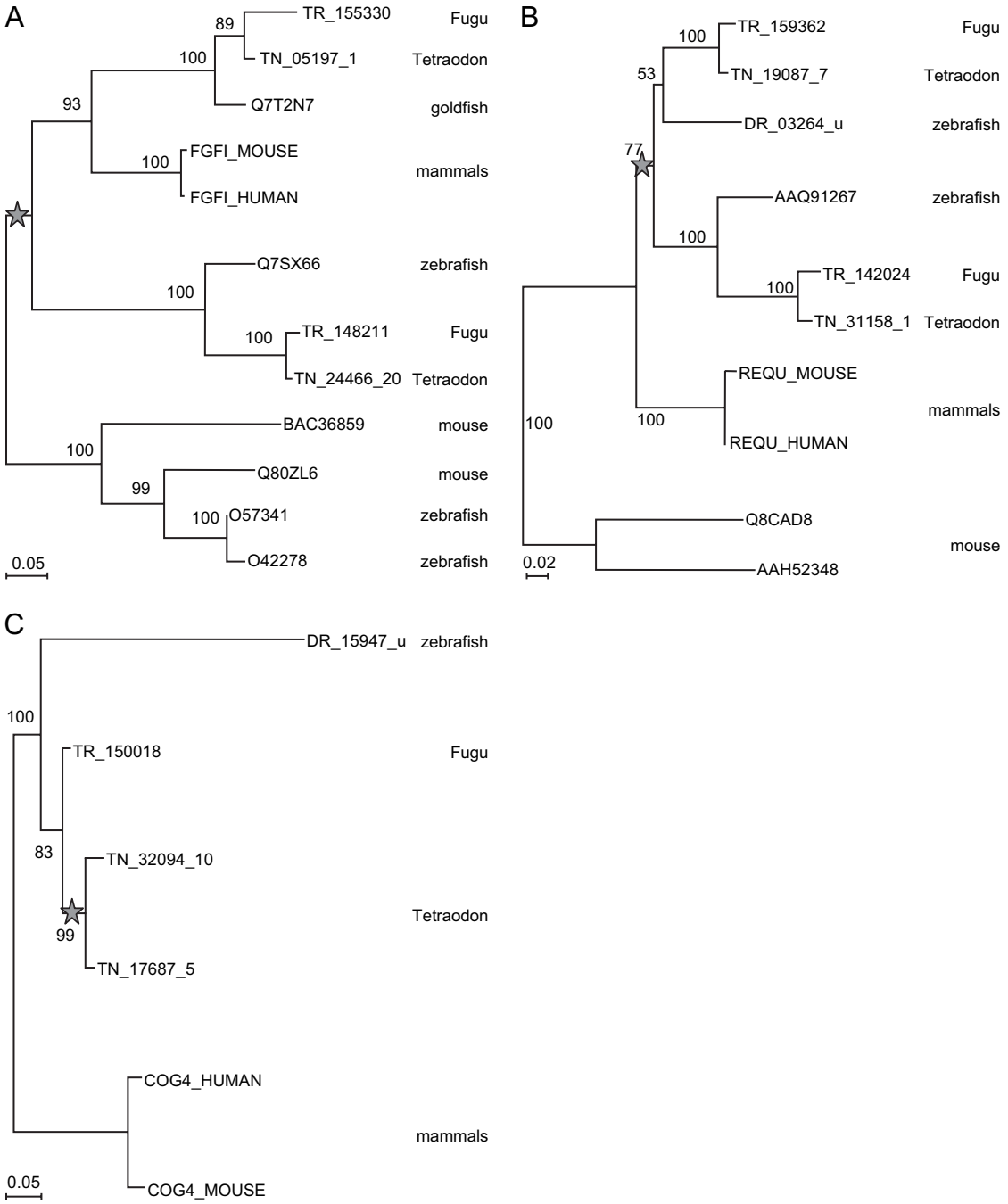


FIG. 1.—Examples of relative dating of gene duplications by phylogeny. Representative examples of the main phylogenetic scenarios observed. Neighbor-Joining phylogenetic trees, with substitutions corrected for heterogeneity between sites by a gamma law ($\alpha = 0.8$); Maximum Likelihood gives identical topologies (not shown). Branch length is proportional to the number of substitutions per site (see measure bar for each tree); numbers at nodes are support in percentage of 500 bootstrap replicates. The stars represent the duplication that led to 2 *Tetraodon* paralogs in each case. (*A*) Fibroblast growth factor-18 precursor, duplicated before the tetrapod–actinopterygian split (ancestral vertebrate duplication), with secondary loss of 1 copy in mammals, leading to 2 copies in fishes but only 1 in human; clade BAC36859–O42278 is a more distant paralog used as an outgroup. (*B*) Zinc-finger protein ubi-d4 (Requiem), duplicated anciently in fishes; clade Q8CAD8–AAH52348 is a more distant paralog used as an outgroup. (*C*) Conserved oligomeric Golgi complex component 4, duplicated specifically in *Tetraodon*, after the divergence with other fishes.

between *Tetraodon* and *Takifugu* was recently reassessed at 85 MYA (Yamanoue et al. 2006), similar to the divergence between human and mouse. This affects neither relative rate tests (Robinson-Rechavi and Laudet 2001; Steinke et al. 2006) nor comparisons of Ka/Ks ratios (this study), which are both independent of divergence time.

Once we have established contributions to evolutionary rate, which are not a direct result of duplication (biased retention and species-specific differences), can we also measure the direct effect of duplication? Lower purifying selection is expected on redundant genes after duplication. In principle, both paralogs could experience released

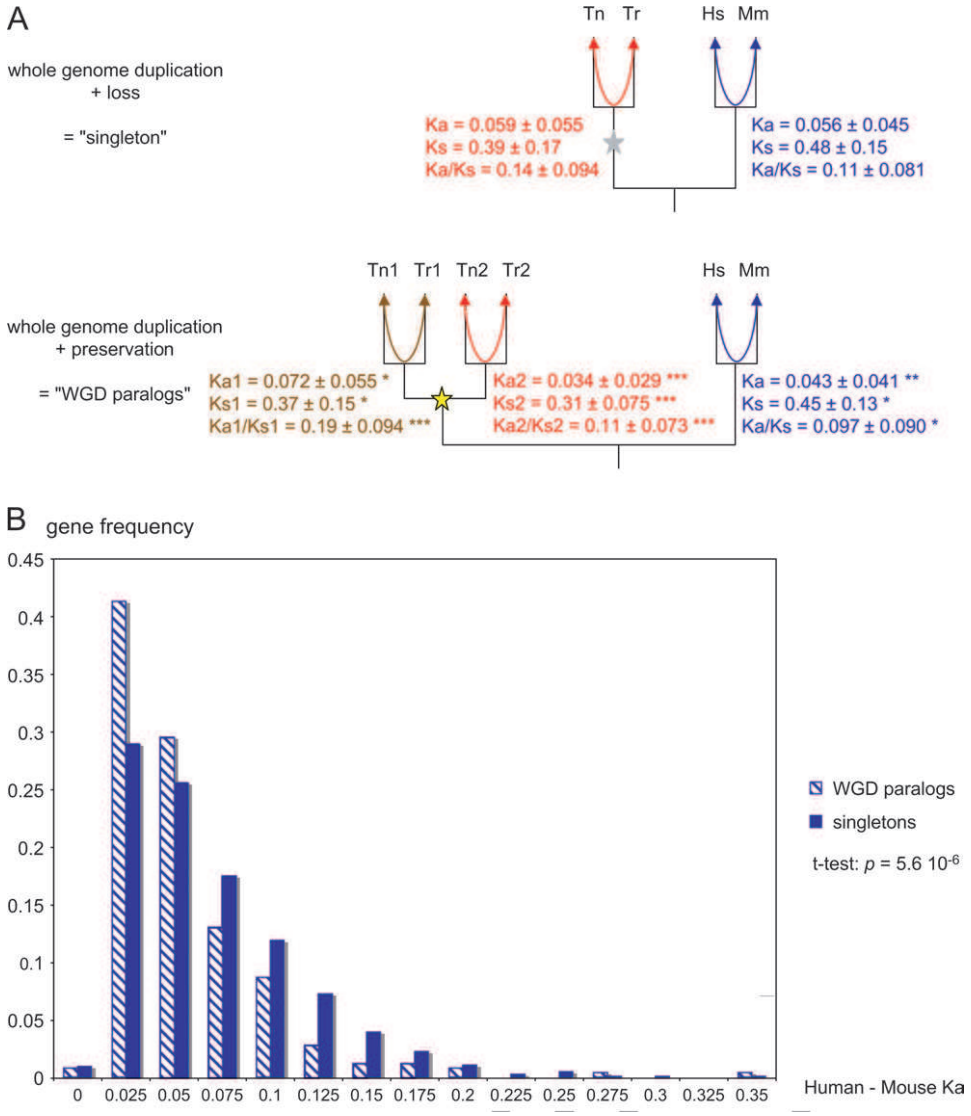


FIG. 2.—Evolutionary rates of singleton and duplicate genes in *Tetraodon nigroviridis*. (*A*) Genes are classified according to duplication history. Tn = *T. nigroviridis*; Tr = *Takifugu rubripes*; Hs = *Homo sapiens*; Mm = *Mus musculus*. Ka = mean number of nonsynonymous substitutions per nonsynonymous site, measuring protein evolutionary rate; Ks = mean number of synonymous substitutions per synonymous site, measuring neutral evolutionary rate; Ka/Ks = mean ratio of nonsynonymous to synonymous substitutions, measuring selective pressure. All means are ± standard deviation. The arrows indicate which sequences were compared. In red and brown, the comparison between pufferfish orthologs; for WGD paralogs, we report separately the values for ortholog pairs with highest Ka (in brown) and with lowest Ka (in red); stars indicate a significant difference to the pufferfish singleton values (unpaired *t*-test: * $P < 0.05$; ** $P < 10^{-4}$; *** $P < 10^{-6}$). In blue, the comparison between mouse and human orthologs; stars indicate a significant difference to the mammalian singleton values. (*B*) Frequency distributions of Ka values between human and mouse orthologous genes, according to their duplication history in fishes.

selection and higher rates, or the relaxation could be restricted to only one paralog, whereas the other would conserve ancestral constraints and evolutionary rates. We first tested whether the observed difference between WGD paralogs is meaningful. Indeed 2 homologous gene pairs will never have exactly the same rate because of stochastic as well as mechanistic factors, such as variations in mutation rates along the genome. For this, we contrast variation in Ka to variation in Ks, used as a proxy for neutral variation in evolutionary rate. The relative variation in Ka is significantly larger than the variation in Ks (figs. 2*A* and 3), supporting a selective difference between the paralogs. A striking 36% of paralogous pairs have higher Ka variation than predicted by the neutral distribution (ΔKa higher than the 97.5th percentile of ΔKs distribution). And this is a conservative estimate, considering that for 81% of paralogs, ΔKa is higher than ΔKs, indicating potential selective differences. The results are similar if we exclude pairs with mean Ka < 0.05, to account for possible stochastic effects on small numbers of substitutions (not shown). Thus, the rate differences between the 2 paralogs can be largely interpreted in terms of selective differences.

If we compare the rates of these paralogs to those of mammalian orthologs on the one hand and to singleton pufferfish genes on the other, it seems that the slower copy has retained the strong constraint of its ancestor (Ka2/Ks2

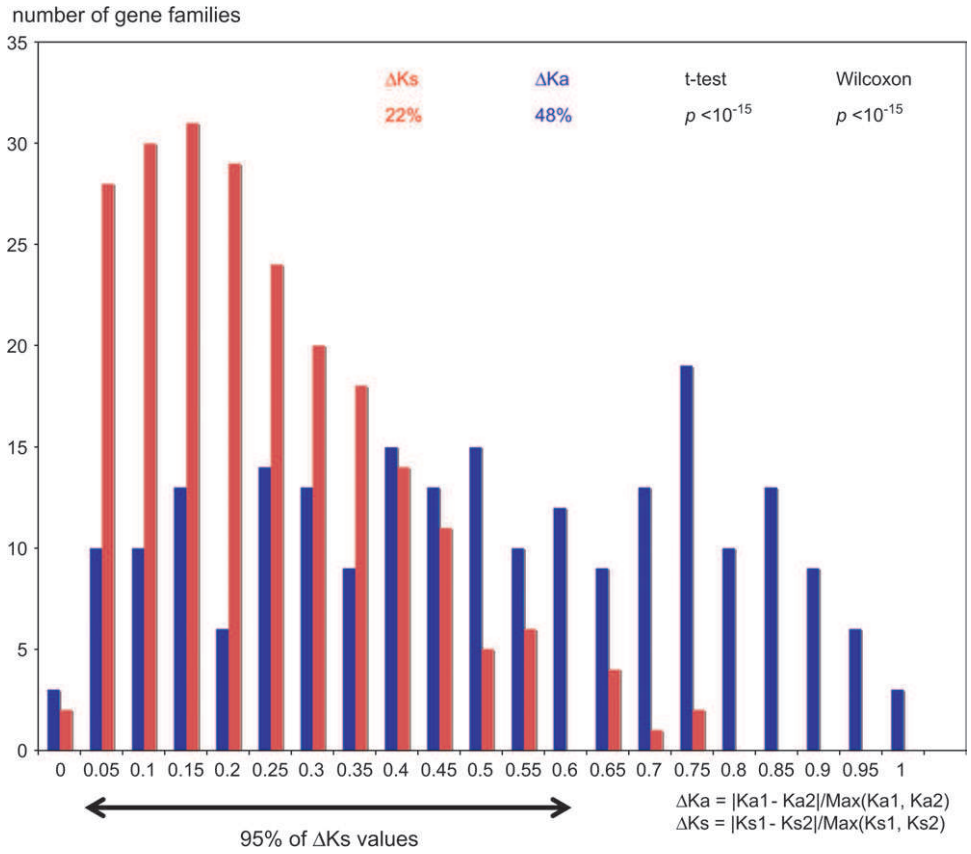


FIG. 3.—Variation in evolutionary rates after WGD in Tetraodontiformes. Distributions of ΔKs (red) and ΔKa (blue) values between WGD paralogs. Ka1, Ka2, Ks1, and Ks2 as defined in figure 2*A*. The arrow shows the range of variation in ΔKs for 95% of gene pairs, defining a neutral expectation of rate variation between paralogs.

not significantly different from mammalian ortholog Ka/Ks), whereas the other copy has experienced a large relaxation of selective constraint (Ka1/Ks1 highest of any subset of data, $P < 10^{-6}$ in all comparisons).

We have only identified 36 unambiguous recently duplicated genes in *Tetraodon*, and Ka and Ks can be computed for both fishes and mammals for only 23 of them. This low number limits statistical power, but we observe that on average they have intermediary evolutionary rates and selective pressures, between those of paralogs from the WGD and those of singletons (i.e., Ka of mammalian orthologs = 0.052). Interestingly, there appear to be selective differences between these recent paralogs, as between WGD paralogs, although sample size limits significance (ΔKa = 39%; ΔKs = 22%; $P = 0.056$).

Functional Classification of Duplicate Genes

We examined the possibility that the genome duplication may have retained duplicate genes in specific functional classes, by using the GO classification (Ashburner et al. 2000) of human orthologs of *Tetraodon* genes. Orthologs of WGD paralogs are significantly enriched in terms that relate to development, signaling, behavior, and regulation (table 2). This enrichment is confirmed if we consider the less detailed level 2 GO terms (not shown), with notably an excess of genes associated to the biological process "development" ($P = 0.0065$). Conversely, they are depleted in terms describing fundamental processes of the cell, such as metabolism or catalytic activity. For the most part, the same terms are enriched or depleted in the recently duplicated genes, but low sample size again limits significance (not shown).

Discussion

We interpret our results in light of the now classical "duplication–degeneration–complementation" (DDC) model for the functional consequences of gene duplication (Force et al. 1999). The DDC model proposes 3 possible fates for the new paralogs: nonfunctionalization, in which 1 copy is lost and the situation reverts to its preduplication state; neofunctionalization, in which 1 copy acquires a new function while the other keeps the ancestral function; and subfunctionalization, in which each copy loses part of the ancestral function, both copies being required then to keep the full function active. It is clear from genomic studies that nonfunctionalization is the main fate of duplicate genes: gene loss appears to have been 88% in about 80 Myr since genome duplication in yeasts (Kellis et al. 2004), 70% in ≤86 Myr in *Arabidopsis* (Bowers et al. 2003), and 79% in about 61–67 Myr in cereals (Paterson et al. 2004). The figure of 85% we obtain for gene loss in *Tetraodon* is similar to the observations in other lineages,

Table 2
Functional Classification of *Tetraodon* Duplicated Genes

Ontology	GO Term	Observed Gene Number	Enrichment/ Depletion	*P* Value
Biological process	Cell communication	43	×1.79	3.7×10^{-06}
	Organismal physiological process	24	×2.04	1.2×10^{-04}
	Death	10	×2.86	9.5×10^{-04}
	Locomotory behavior	6	×3.21	0.0054
	System development	9	×2.14	0.012
	Mesoderm development	3	×5.15	0.012
	Tissue development	5	×2.67	0.023
	Regulation of cellular process	29	×1.24	0.037
	Response to stimulus	19	×1.35	0.039
	Regulation of physiological process	28	×1.22	0.042
	Metabolism	54	÷1.31	4.4×10^{-05}
	Cellular physiological process	77	÷1.06	0.050
Molecular function	Pattern binding	6	×3.16	0.0059
	Neurotransmitter binding	4	×3.74	0.013
	Structural constituent of eye lens	2	×8.42	0.014
	Ion binding	30	×1.33	0.017
	Carbohydrate binding	6	×2.53	0.018
	Channel or pore class transporter activity	4	×2.59	0.045
	Catalytic activity	41	÷1.28	0.0045
	Nucleic acid binding	13	÷1.54	0.018
	Oxidoreductase activity	4	÷2.23	0.027
	Ligase activity	1	÷4.87	0.029
	Hydrolase activity	13	÷1.45	0.031
	Nucleotide binding	13	÷1.43	0.034
Cellular component	Connexon complex	3	×8.22	0.0018
	Extracellular matrix (sensu Metazoa)	9	×2.54	0.0038
	Membrane	35	×1.28	0.017
	Membrane-bound organelle	39	÷1.30	0.0029
	Intracellular organelle	45	÷1.23	0.0051
	Intracellular	58	÷1.13	0.015

NOTE.—Statistics on human orthologs of *Tetraodon* genes, which were classified according to duplication history as in figure 2. The observed gene number is the number of human genes annotated for each term. Only GO terms of level 3 or higher with a significant enrichment or depletion in paralogs from WGD are reported ($P \leq 0.05$ after correction for multiple tests). Enrichment ($\times n$) and depletion ($\div n$) are reported as the ratio between expected and observed frequency of terms, based on expectations from human genes that were defined as orthologs of *Tetraodon* genes in the combined phylogenetic and mapping analysis.

despite the greater age of the event; even less gene loss has been reported based on similarity hits without phylogenetic analysis (Jaillon et al. 2004; Woods et al. 2005). This is probably because most gene loss occurs rapidly after duplication (Lynch and Conery 2000; Jaillon et al. 2004; Scannell et al. 2006), so that subsequent evolution does not change the figure significantly.

Slowly evolving genes are more likely to be found duplicated in the yeast *S. cerevisiae* and the nematode worm *Caenorhabditis elegans* (Davis and Petrov 2004). The bias is similar in both yeast and worm and is maintained over evolutionary time, indicating that gene retention was also biased after the WGD in yeast. Jordan et al. (2004) also showed that genes that have paralogs evolve slower than singletons, in several eukaryotic and prokaryotic genomes, but did not use a phylogenetic framework with independent estimates of duplication and evolutionary rate. Using phylogenetically independent estimates, we show that such a bias exists in duplicated genes in a vertebrate genome (fig. 2): nonduplicated mammalian orthologs of gene pairs retained from a WGD in fishes evolve 23% slower in mammals than orthologs of singleton genes. This is a comparable figure to observations for nematode and yeast genes (Davis and Petrov [2004] computed differences relative to the slower rate; this gives 30% in fishes, compared to 25% in nematode and 50% in yeast). Because these gene pairs result from 1 WGD event, where all genes were duplicated, the bias must come from a differential retention of the duplicates, rather than a biased generation.

Unraveling the contributions of biased retention, different rates in fishes and mammals, and duplication itself, we find evidence for asymmetric acceleration of protein evolution after WGD. A recent study found evidence for different rates of evolution between fish-specific paralogs in 24 genes as well as for increased rates of radical amino acid substitution (Steinke et al. 2006), although it is not clear to what extent this constitutes evidence for positive selection. In yeast, differences in evolutionary rate between WGD paralogs have been found to be pervasive and associated to some evidence for positive selection on amino acid sites (Fares et al. 2006). In that study, it was estimated that 19% of duplicate genes have fixed amino acids by positive selection. We find considerably higher differences in Ka than in Ks between paralogs, which also indicates a change in selective pressure, for at least 36% of gene pairs. The difference appears due to acceleration of nonsynonymous rate in 1 paralog of the pair, whereas the other keeps the ancestral constraint. Therefore, as in yeast, fly, and nematode (Conant and Wagner 2003; Kellis et al. 2004), results are suggestive of neofunctionalization of duplicated genes in vertebrates, insofar as this can be indicated by evolutionary rates. This would

be in agreement with the original evolutionary theory of Ohno (1970) as well as with a more recent model of subfunctionalization followed by neofunctionalization (He and Zhang 2005).

Retention of duplicates is also biased with regard to function (table 2). The excess of development and signaling functions supports the much discussed idea that genome duplication may have played an important role in establishing developmental diversity of fishes (Vogel 1998; Wittbrodt et al. 1998; Meyer and Schartl 1999; Venkatesh 2003; Postlethwait et al. 2004; Meyer and Van de Peer 2005; Volff 2005). The duplication may also have been important in establishing behavioral diversity. Interestingly, by a quite different approach, Steinke et al. (2006) found an excess of transcription factors among genes with accelerated evolution or specific duplication in fishes. The enrichment in communication and developmental genes is consistent with observations in insects and yeasts (Jordan et al. 2004) and *Arabidopsis* (Maere et al. 2005), although another study found an enrichment of metabolic genes in *Arabidopsis* (Blanc and Wolfe 2004), and contrasting results were found in mammals (Nembaware et al. 2002; Jordan et al. 2004).

In conclusion, the consequences of WGD in fishes have been mediated both by biased gene loss, resulting in enrichment of development, signaling, or behavioral genes compared with mammals, and by changes in selective pressure, asymmetric between duplicates. The functional consequences of these changes remain to be explored.

Supplementary Material

Supplementary Table I and Figures I and II are available at Molecular Biology and Evolution online (http://www.mbe.oxfordjournals.org/).

Acknowledgments

We are grateful to the Sanger Institute for making zebrafish genome data publicly available. Laurent Duret and Simon Penel helped with fish data from Hovergen. F.G.B. and V.L. acknowledge support from Région Rhone-Alpes, Ministére de l'éducation nationale, de la recherche et de la technologie, European Union, and Institut national de la recherche agronomique.

Funding to pay the Open Access publication charges for this article was provided by Etat de Vaud.

Literature Cited

Altschul SF, Madden TL, Schaffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ. 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res 25:3389–402.

Amores A, Force A, Yan YL, et al. (13 co-authors). 1998. Zebrafish *hox* clusters and vertebrate genome evolution. Science 282:1711–4.

Aparicio S, Chapman J, Stupka E, et al. (41 co-authors). 2002. Whole-genome shotgun assembly and analysis of the genome of Fugu rubripes. Science 297:1301–10.

Ashburner M, Ball CA, Blake JA, et al. (20 co-authors). 2000. Gene ontology: tool for the unification of biology. Nat Genet 25:25–9.

Benjamini Y, Hochberg Y. 1995. Controlling the false discovery rate: a practical and powerful approach to multiple testing. J R Stat Soc 57:289–300.

Blanc G, Wolfe KH. 2004. Functional divergence of duplicated genes formed by polyploidy during Arabidopsis evolution. Plant Cell 16:1679–91.

Boeckmann B, Bairoch A, Apweiler R, et al. (12 co-authors). 2003. The SWISS-PROT protein knowledgebase and its supplement TrEMBL in 2003. Nucleic Acids Res 31:365–70.

Bowers JE, Chapman BA, Rong J, Paterson AH. 2003. Unravelling angiosperm genome evolution by phylogenetic analysis of chromosomal duplication events. Nature 422:433–8.

Christoffels A, Koh EG, Chia JM, Brenner S, Aparicio S, Venkatesh B. 2004. Fugu genome analysis provides evidence for a whole-genome duplication early during the evolution of ray-finned fishes. Mol Biol Evol 21:1146–51.

Conant GC, Wagner A. 2003. Asymmetric sequence divergence of duplicate genes. Genome Res 13:2052–8.

Crow KD, Stadler PF, Lynch VJ, Amemiya C, Wagner GP. 2006. The "fish-specific" Hox cluster duplication is coincident with the origin of teleosts. Mol Biol Evol 23:121–36.

Davis JC, Petrov DA. 2004. Preferential duplication of conserved proteins in eukaryotic genomes. PLoS Biol 2:e55.

Davis JC, Petrov DA. 2005. Do disparate mechanisms of duplication add similar genes to the genome? Trends Genet 21:548–51.

Dehal P, Boore JL. 2005. Two rounds of whole genome duplication in the ancestral vertebrate. PLoS Biol 3:e314.

Duret L, Mouchiroud D, Gouy M. 1994. HOVERGEN: a database of homologous vertebrate genes. Nucleic Acids Res 22:2360–5.

Fares MA, Byrne KP, Wolfe KH. 2006. Rate asymmetry after genome duplication causes substantial long-branch attraction artifacts in the phylogeny of Saccharomyces species. Mol Biol Evol 23:245–53.

Force A, Lynch M, Pickett FB, Amores A, Yan Y-l, Postlethwait J. 1999. Preservation of duplicate genes by complementary, degenerative mutations. Genetics 151:1531–1545.

Galtier N, Gouy M, Gautier C. 1996. SEAVIEW and PHYLO_WIN: two graphic tools for sequence alignment and molecular phylogeny. Comput Appl Biosci 12:543–8.

He X, Zhang J. 2005. Rapid subfunctionalization accompanied by prolonged and substantial neofunctionalization in duplicate gene evolution. Genetics 169:1157–64.

Hoegg S, Brinkmann H, Taylor JS, Meyer A. 2004. Phylogenetic timing of the fish-specific genome duplication correlates with the diversification of teleost fish. J Mol Evol 59: 190–203.

Hubbard T, Barker D, Birney E, et al. (35 co-authors). 2002. The Ensembl genome database project. Nucleic Acids Res 30: 38–41.

Jaillon O, Aury J-M, Brunet F, et al. (61 co-authors). 2004. Genome duplication in the teleost fish Tetraodon nigroviridis reveals the early vertebrate proto-karyotype. Nature 431: 946–57.

Jordan IK, Wolf YI, Koonin EV. 2004. Duplicated genes evolve slower than singletons despite the initial rate increase. BMC Evol Biol 4:22.

Kellis M, Birren BW, Lander ES. 2004. Proof and evolutionary analysis of ancient genome duplication in the yeast Saccharomyces cerevisiae. Nature 428:617–24.

Li WH. 1993. Unbiased estimation of the rates of synonymous and nonsynonymous substitution. J Mol Evol 36:96–9.

Lynch M, Conery JS. 2000. The evolutionary fate and consequences of duplicate genes. Science 290:1151–5.

Maere S, De Bodt S, Raes J, Casneuf T, Van Montagu M, Kuiper M, Van de Peer Y. 2005. Modeling gene and genome

duplications in eukaryotes. Proc Natl Acad Sci USA 102: 5454–9.
Martin D, Brun C, Remy E, Mouren P, Thieffry D, Jacq B. 2004. GOToolBox: functional analysis of gene datasets based on Gene Ontology. Genome Biol 5:R101.
Meyer A, Schartl M. 1999. Gene and genome duplications in vertebrates: the one-to-four (-to-eight in fish) rule and the evolution of novel gene functions. Curr Opin Cell Biol 1:699–704.
Meyer A, Van de Peer Y. 2005. From 2R to 3R: evidence for a fish-specific genome duplication (FSGD). Bioessays 27:937–45.
Naruse K, Fukamachi S, Mitani H, et al. (20 co-authors). 2000. A detailed linkage map of medaka, Oryzias latipes: comparative genomics and genome evolution. Genetics 154:1773–84.
Nembaware V, Crum K, Kelso J, Seoighe C. 2002. Impact of the presence of paralogs on sequence divergence in a set of mouse-human orthologs. Genome Res 12:1370–6.
Ohno S. 1970. Evolution by gene duplication. Heidelberg: Springer-Verlag.
Otto SP, Whitton J. 2000. Polyploid incidence and evolution. Annu Rev Genet 34:401–37.
Pamilo P, Bianchi NO. 1993. Evolution of the Zfx and Zfy genes: rates and interdependence between the genes. Mol Biol Evol 10:271–81.
Paterson AH, Bowers JE, Chapman BA. 2004. Ancient polyploidization predating divergence of the cereals, and its consequences for comparative genomics. Proc Natl Acad Sci USA 101:9903–8.
Postlethwait J, Amores A, Cresko W, Singer A, Yan Y-L. 2004. Subfunction partitioning, the teleost radiation and the annotation of the human genome. Trends Genet 20:481–90.
Postlethwait JH, Woods IG, Ngo-Hazelett P, Yan YL, Kelly PD, Chu F, Huang H, Hill-Force A, Talbot WS. 2000. Zebrafish comparative genomics and the origins of vertebrate chromosomes. Genome Res 10:1890–902.
Robinson-Rechavi M, Laudet V. 2001. Evolutionary rates of duplicate genes in fish and mammals. Mol Biol Evol 18:681–3.
Robinson-Rechavi M, Marchand O, Escriva H, Bardet P-L, Zelus D, Hughes S, Laudet V. 2001. Euteleost fish genomes are characterized by expansion of gene families. Genome Res 11:781–8.
Roest Crollius H, Jaillon O, Bernot A, et al. (12 co-authors). 2000. Estimate of human gene number provided by genome-wide analysis using Tetraodon nigroviridis DNA sequence. Nat Genet 25:235–8.
Saitou N, Nei M. 1987. The neighbor-joining method: a new method for reconstructing phylogenetic trees. Mol Biol Evol 4:406–25.
Scannell DR, Byrne KP, Gordon JL, Wong S, Wolfe KH. 2006. Multiple rounds of speciation associated with reciprocal gene loss in polyploid yeasts. Nature 440:341–5.
Steinke D, Salzburger W, Braasch I, Meyer A. 2006. Many genes in fish have species-specific asymmetric rates of molecular evolution. BMC Genomics 7:20.
Thomson JD, Higgins DG, Gibson TJ. 1994. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res 22:4673–80.
Van de Peer Y, Frickey T, Taylor JS, Meyer A. 2002. Dealing with saturation at the amino acid level: a case study based on anciently duplicated zebrafish genes. Gene 295:205–11.
Van de Peer Y, Taylor JS, Braasch I, Meyer A. 2001. The ghosts of selection past: rates of evolution and functional divergence in anciently duplicated genes. J Mol Evol 53:434–44.
Vandepoele K, De Vos W, Taylor JS, Meyer A, Van de Peer Y. 2004. Major events in the genome evolution of vertebrates: paranome age and size differ considerably between ray-finned fishes and land vertebrates. Proc Natl Acad Sci USA 101:1638–43.
Venkatesh B. 2003. Evolution and diversity of fish genomes. Curr Opin Genet Dev 13:588–92.
Venkatesh B, Tay A, Dandona N, Patil JG, Brenner S. 2005. A compact cartilaginous fish model genome. Curr Biol 15:R82–3.
Vogel G. 1998. Doubled genes may explain fish diversity. Science 281:1119–21.
Volff JN. 2005. Genome evolution and biodiversity in teleost fish. Heredity 94:280–94.
Wittbrodt J, Meyer A, Schartl M. 1998. More genes in fish? Bioessays 20:511–5.
Woods IG, Wilson C, Friedlander B, Chang P, Reyes DK, Nix R, Kelly PD, Chu F, Postlethwait JH, Talbot WS. 2005. The zebrafish gene map defines ancestral vertebrate chromosomes. Genome Res 15:1307–14.
Yamanoue Y, Miya M, Inoue JG, Matsuura K, Nishida M. 2006. The mitochondrial genome of spotted green pufferfish Tetraodon nigroviridis (Teleostei: Tetraodontiformes) and divergence time estimation among model organisms in fishes. Genes Genet Syst 81:29–39.

Kenneth Wolfe, Associate Editor

Accepted June 23, 2006

COLOPHON

This thesis was typeset with LaTeX 2ε using Hermann Zapf's *Palatino* and *Euler* type faces (Type 1 PostScript fonts *URW Palladio L* and *FPL* were used). The listings are typeset in *Bera Mono*, originally developed by Bitstream, Inc. as "Bitstream Vera". (Type 1 PostScript fonts were made available by Malte Rosenau and Ulrich Dirr.)

The typographic style was inspired by **?** as presented in *The Elements of Typographic Style* (**?**). It is available for LaTeX via CTAN as "`classicthesis`".

Final Version as of March 5, 2009 at 11:14.