Biodiversité du microbiome cutané des organismes marins : variabilité, déterminants et importance dans l’écosystème
Marlène Chiarello

To cite this version:

HAL Id: tel-01693132
https://tel.archives-ouvertes.fr/tel-01693132
Submitted on 25 Jan 2018

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L’archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.
Délivré par l’Université de Montpellier

Préparée au sein de l’école doctorale GAIA
Et de l’unité de recherche MARBEC (Centre pour la Biodiversité marine, l’Exploitation et la Conservation)

Spécialité : Écologie Fonctionnelle et Sciences Agronomiques

Présentée par Marlène Chiarello

Biodiversité du microbiome cutané des organismes marins : variabilité, déterminants et importance dans l’écosystème

Soutenue le 29 novembre 2017 devant le jury composé de

M. Thierry BOUVIER, DR, CNRS
M. Frédéric DELSUC, DR, CNRS
M. Pierre GALAND, DR, CNRS
M. Fabrice NOT, DR, CNRS
M. Loïc PELLISIER, Asst. Prof., ETH Zürich
M. Téléphore SIME-NGANDO, DR, CNRS
Mme Eve TOULZA, MCF, Université de Perpignan
M. Sébastien VILLÉGER, CR, CNRS

Directeur
Examinateur
Invité
Invité
Rapporteur
Rapporteur
Examinateur
Co-encadrant
Résumé


Mots-clefs : microbiome, diversité phylogénétique, Séquençage haut-débit, communautés microbiennes, peau, Teleostei, Odontoceti, écosystèmes coralliens
Abstract

Oceans contain thousands of microbial species playing crucial roles for the functioning of the marine ecosystem. These microorganisms are present everywhere in the water column. Some microorganisms also colonize the surface and the digestive tract of marine macro-organisms, forming communities called microbiomes. These microbiomes have positive effects for their host’s fitness. The diversity of these marine animal surface microbiome is still largely understudied, despite recent progress in molecular biology that now permits to fully assess its different facets of biodiversity, i.e. taxonomic, phylogenetic and functional. The goal of this thesis is therefore to describe the diversity of the surface microbiome of marine animals, to assess its variability at different levels, as well as its determinants, and the significance of such diversity at the ecosystem’s scale. Firstly, I have assessed the efficiency of various diversity indices to detect ecological signals in the specific case of microbial communities. Secondly, I have described the surface microbiome of major marine animal clades (teleostean fishes, cetaceans and several classes of invertebrates). I found that these microbiomes are highly distinct from the surrounding planktonic communities. I demonstrated that these microbiomes are variable both between individuals from the same species and between species, but do not show a phylosymbiosis pattern. Last, I assessed the contribution of surface microbiomes to the global microbial community at the scale of a coral reef ecosystem. I demonstrated that marine animal surfaces host almost twenty times more microbial species than the water column, and 75% of the phylogenetic richness present in the ecosystem. In a context of massive erosion of marine macroscopic organisms, it is therefore urgent to exhaustively assess marine microbial biodiversity and its vulnerability facing anthropic pressures.

Key-words : microbiome, phylogenetic diversity, Next Generation Sequencing (NGS), microbial communities, skin, Teleostei, Odontoceti, coral reef ecosystems

Ensuite je remercie mes deux co-directeurs de thèse pour leur encadrement de haut vol et leur soutien tout au long de cette thèse. Travailler avec vous deux a été un plaisir, et j’espère pouvoir continuer à collaborer avec vous lorsque je volerai de mes propres ailes.

Merci à Thierry pour la confiance qu’il m’a accordée depuis le début, en acceptant de me co-encadrer durant mon stage de M1, puis en me proposant un stage de M2, et en me permettant de continuer en thèse. Merci pour les échanges toujours constructifs que nous avons eu tout au long de ces 5 années. Merci également pour ton écoute et ton enthousiasme, mais aussi pour tes doutes, qui m’ont fait me poser les bonnes questions. Merci de m’avoir permis de m’épanouir tout au long de cette thèse.

Cette thèse ne serait pas ce qu’elle est sans Sébastien, de loin mon co-auteur le plus impliqué, et mon relecteur le plus sévère. Il fallait qu’il y ait un seul poisonologue qui ait envie de faire pondre un papier sur les indices de diversité à une pipétologue, et que je tombe dessus ! Séb, rien que cela prouve que pour toi, la science n’a pas de frontières. Merci de t’être mis à la microbio, de t’être arraché les cheveux avec moi quand les manip ne fonctionnaient pas et quand mes discussions d’articles ressemblaient à du porridge mal cuít. Séb, c’est également toi qui m’a le plus aidée à organiser mes campagnes de terrain, dont (et surtout !) celle de Mayotte. Tu t’es mouillé au sens propre comme au sens figuré pour cette thèse, et tu m’a permis de me faire des souvenirs impérissables, de récits coralliens, tortues, et poissons de toutes les couleurs, mais également de kiki en chaussettes et sandales, scotch autour des mollets, sur la plage au coucher du soleil. J’en ris encore. Bref. Sans déconner, merci !

Merci ensuite à tous ceux qui m’ont aidée tout au long de cette thèse. Merci aux personnes qui m’ont accompagnée sur le terrain, et en particulier à Thomas Claverie et Elliott Sucré, sans qui l’échantillonnage à Mayotte n’aurait pas été possible. Merci à Jean-Christophe Auguet, relecteur attentif qui m’a dépatouillée plus d’une fois de mes galères bioinformatiques. Merci à Corinne Bouvier pour m’avoir formée à la biomol, et pour les références musicales de très haut niveau ;-) Merci à Fabien Rillieuvineuve d’avoir passé tes journées, et parfois tes soirées avec moi à faire des extractions d’ADN et des PCR ; travailler avec toi a été un vrai plaisir. Merci également à toutes les personnes de MARBEC et du bâtiment 24, qui ont enrichies ces trois années de conversations toujours intéressantes, et qui m’ont filé un coup de pouce à l’occasion ; je pense notamment à Yvan, Yannick, Jérémie, Franck, Fabien, David, Yunne, Béa, Patricia, Aude, Julien, Hahn, Claudia, Inês, François… J’en oublie certainement.

Merci à toutes les personnes avec qui j’ai passé de très bons moments en dehors de la fac: Théo (et son macro-organisme associé, Félise), Fabien, Anne-Sophie, Eva, Charlotte, Amandine (et son macro-organisme associé, Romain), Mariam, Timothée, Quentin, Laure et Clara. Certains épisodes, que par discretion je n’évoquerai pas ici, resteront gravés dans le marbre :). En plus d’être de bons collègues, vous êtes de bons amis. J’espère que nous resterons en contact par la suite. Merci à mes amis de longue date, Cassy et Sébastien, auprès de qui je tiens à m’excuser de ne pas avoir beaucoup donné de nouvelles durant cette troisième année de thèse.

Enfin, la partie la plus sensible, donc je serais brève : merci à toute ma famille, et en particulier à mes thésards de parents, sans qui je ne me serais pas lancée dans cette aventure. Merci de votre soutien sans faille durant toutes ces années. Merci à ma maman (<3) d’avoir corrigé les quelques coquilles de cette thèse. Merci aussi à mon papa, tu ne liras certainement pas cette thèse, mais t’es quand même le meilleur des papa <3. Merci à mon frère et ma sœur, qui ont fait que je ne me suis jamais sentie seule. Merci à Landry d’avoir mesuré la longueur totale de 138 poissons au cours de cette thèse, et d’avoir payé l’abonnement Netflix qui a sauvé mes soirées fatiguées. Merci à Adèle (^_^). Merci à ma tante, Hildegarde, et mes cousins (Roland, Guylain et Gauthier) d’être aussi pas sortables et géniaux à la fois ! Merci à mes grands-parents, et à ma grand-mère Hélène, de m’avoir laissé autant de souvenirs.

Enfin, je remercie mon macro-organisme associé, Joseph <3
Table des matières

1. INTRODUCTION GENERALE : MICROBIOMES ET ORGANISMES MARINS ................................................. 9
   1.1 LE MICROBIOME DES ANIMAUX ........................................................................................................ 9
       1.1.1 Perspective historique et définitions ......................................................................................... 9
       1.1.2 Diversité et échelles de variabilité du microbiome associé aux animaux .............................. 11
       1.1.3 Mécanismes à l’origine de la variabilité des microbiomes ......................................................... 16
   1.2 CAS DES MICROBIOMES ASSOCIÉS AUX ORGANISMES MARINS .................................................. 28
       1.2.1 Composition et rôles du microbiome chez les invertébrés marins ....................................... 29
       1.2.2 Composition et rôles du microbiome chez les vertébrés marins .......................................... 32
       1.2.3 Rôles des microbiomes associés aux animaux dans les écosystèmes marins .......................... 38
   1.3 OBJECTIFS DE LA THESE .................................................................................................................... 42

2. METHODES D’ANALYSE DE LA DIVERSITE DU MICROBIOME CUTANE DES ANIMAUX MARINS .................................................................................................................. 49
   2.1 HISTORIQUE DES METHODES UTILISEES POUR L’ETUDE DU MICROBIOME ........................ 49
       2.1.1 Méthodes culturelles .................................................................................................................. 49
       2.1.2 Méthodes moléculaires basées sur un gène marqueur ................................................................ 50
       2.1.3 Approches métagénomiques ....................................................................................................... 51
   2.2 METHODES UTILISEES LORS DE LA THESE .................................................................................... 52
       2.2.1 Campagnes de prélèvement ........................................................................................................ 52
       2.2.2 Campagne 1 .................................................................................................................................. 52
       2.2.3 Campagne 2 .................................................................................................................................. 54
       2.2.4 Campagne 3 .................................................................................................................................. 55
       2.2.5 Campagne 4 .................................................................................................................................. 57
       2.2.6 Campagne 5 .................................................................................................................................. 58

3. REFLEXION METHODOLOGIQUE : CHOIX D’INDICES DE DIVERSITE APPLIQUES AU MONDE MICROBIEN ........................................................................................................................................ 65
   3.1 FACETTES ET COMPOSANTES DE LA BIODIVERSITE ..................................................................... 65
   3.2 PARTICULARITES DES COMMUNAUTES MICROBIENNES .................................................................. 71
   3.3 OBJECTIFS DU CHAPITRE ET PRINCIPAUX RESULTATS ............................................................... 73
   3.4 MANUSCRIT A ...................................................................................................................................... 75

4. FACTEURS DE VARIATION DU MICROBIOME CUTANE DES VERTEBRES MARINS ........................ 99
   4.1 ECHelles DE VARIABILITÉ ET FACTEURS ENVIRONNEMENTAUX ...................................................... 99
   4.2 MANUSCRIT B ..................................................................................................................................... 102
   4.3 MANUSCRIT C ..................................................................................................................................... 115
   4.4 MANUSCRIT D ..................................................................................................................................... 137

5. IMPORTANCE ET VULNERABILITÉ DU MICROBIOME CUTANE DES ANIMAUX DANS L’ECOSYSTEME MARIN .................................................................................................................. 151
   5.1 MICROBIOME CUTANE ET COMMUNAUTES PLANCTONIQUES ....................................................... 151
   5.2 VULNERABILITE DES MACRO-ORGANISMES ET DE LEURS MICROBIOMES ............................... 152
   5.3 OBJECTIFS DU CHAPITRE ET PRINCIPAUX RESULTATS ............................................................... 153
   5.4 MANUSCRIT E ..................................................................................................................................... 155

6. CONCLUSIONS ET PERSPECTIVES : VARIABILITE, DETERMINANTS ET IMPORTANCE DU MICROBIOME CUTANE DES ANIMAUX MARINS ........................................................................... 179
   6.1 MESURES DE LA DIVERSITE ET DE LA VARIABILITE DU MICROBIOME ........................................ 179
   6.2 ÉCHelles DE VARIABILITE ET DETERMINANTS .............................................................................. 181
       6.2.1 Variabilité intraspécifique du microbiome et déterminants potentiels .................................. 181
       6.2.2 Spécificité du microbiome et phylosymbiose ......................................................................... 184
       6.2.3 Quelles sont les échelles temporelles de variation du microbiome cutané ? ........................... 188
   6.3 ROLES DES MICROBIOMES CUTANES DANS LES ECOSYSTEMES MARINS .................................. 190
       6.3.1 Rôles fonctionnels potentiels ....................................................................................................... 190
       6.3.2 Contribution du microbiome à la diversité microbienne marine .............................................. 191
6.3.3  Vulnérabilité de la diversité microbienne face aux changements globaux................. 193

7. REFERENCES BIBLIOGRAPHIQUES ............................................................................. 197

8. ANNEXES.................................................................................................................. 233
   8.1  SYNTHESE BIBLIOGRAPHIQUE SUR LES DETERMINANTS DU MICROBIOME DIGESTIF DU POISSON..... 233
   8.2  PROTOCOLES EXPERIMENTAUX........................................................................... 236
       8.2.1  Protocole d'extraction basée sur le kit Blood & Tissue (Qiagen).......................... 236
       8.2.2  Protocole basé sur le kit PowerSoil® (MoBio).................................................. 239
       8.2.3  Protocole basé sur le kit Maxwell® Buccal Swab LEV DNA (Promega)............. 241
   8.3  INFORMATIONS SUPPLEMENTAIRES AU MANUSCRIT A .......................................... 247
   8.4  INFORMATIONS SUPPLEMENTAIRES AU MANUSCRIT C .......................................... 256
   8.5  INFORMATIONS SUPPLEMENTAIRES ET METHODES DU MANUSCRIT D......................... 268
       8.5.1  Matériel et méthodes employées dans le manuscrit D.................................. 268
       8.5.2  Informations supplémentaires au manuscrit D........................................... 277
   8.6  INFORMATIONS SUPPLEMENTAIRES AU MANUSCRIT E .......................................... 301
   8.7  COLLABORATIONS ............................................................................................... 313
1. Introduction générale : microbiomes et organismes marins

... and then to my great surprise [I] perceived that the aforhead matter [dental plaque] contained very many small living Animals, which moved themselves very extravagantly.

Antoni van Leeuwenhoek, 1683, Philosophical Transactions, 2017, 14(155-166), 568-574

1.1 Le microbiome des animaux

1.1.1 Perspective historique et définitions

L’étude des micro-organismes vivant sur et à l’intérieur des macro-organismes remonte aux premiers pas de la microbiologie. La première observation d’organismes microscopiques a été réalisée à la fin du XVIIe siècle par Antoni van Leeuwenhoek sur ses propres échantillons de salive et de fèces. Grâce à un microscope de sa fabrication, il a été le premier à démontrer que le tube digestif humain contient de nombreux micro-organismes, qu’il nomme « animalcules ». Van Leeuwenhoek a également noté des différences entre les micro-organismes de ces deux habitats, et entre des individus malades et sains (Leeuwenhoek 1684; Dobell 1920).

Par la suite, les progrès de la microbiologie ont démontré que certains de ces micro-organismes causent des maladies, notamment grâce aux postulats de Koch publiés en 1890. S’ensuit un demi-siècle de recherches focalisées essentiellement sur les pathogènes humains (e.g. Herter 1906; Repaci 1910). Ce n’est que dans les années 1950 que l’étude de ces microorganismes (pathogènes, commensaux ou bénéfiques) dans leur ensemble, appelés collectivement « microflore » (ou parfois « flore bactérienne »), s’est développée (e.g. Roine and Elvehjem 1950).

A l’origine le terme de « microflore » était utilisé dans les ouvrages de botanique pour désigner les micro-algues (Zacharias 1896; Bachmann 1901). Il a été longtemps conservé pour désigner les communautés de bactéries, virus et micro-eucaryotes associés au système digestif ou à la peau humaine, à cause de la difficulté à classer ces microbes dans le règne animal ou végétal. Bien qu’encore utilisé, il est progressivement remplacé par le terme de « microbiote », désignant l’ensemble des micro-organismes vivant dans ou sur un hôte animal ou végétal.

Tout comme pour l’étude des microorganismes environnementaux, l’étude du microbiote humain s’est essentiellement reposée sur des méthodes microscopiques ou culturales jusque dans les années 1980 (Hiergeist et al. 2015). Ceci oblige de travailler sur des isolats ou
des cultures pures de bactéries qui ne représentent que 0.5% à 30% de l’abondance microbienne totale dans la plupart des environnements (Amann, Ludwig, and Schleifer 1995; Langendijk et al. 1995). Ce n’est que plus récemment, dans les années 1990, puis 2000, que la diversité du microbiote a pu être appréhendée, grâce au développement de techniques d’empreinte moléculaire, puis du séquençage à haut débit (NGS) qui permettent le séquençage de gènes spécifiques (Figure 1.1). Ces méthodes, associées à une augmentation de l’effort de séquençage, ont révélé la grande diversité du microbiote, estimant que le microbiote digestif humain contiendrait 1000 à 1150 espèces bactériennes (Qin et al. 2010), dont 80% n’avaient jamais été cultivées (Lagier et al. 2015).

L’essor des NGS a poussé à définir la totalité des gènes portés par le microbiote sous le terme de « microbiome » (Turnbaugh et al. 2007). Aujourd’hui ces deux derniers termes sont souvent utilisés comme des synonymes dans la littérature scientifique. Il existe une autre définition de ce terme, donnée par Lederberg (2001), définissant le microbiome comme « la communauté de micro-organismes commensaux, pathogènes et symbiotiques qui partagent littéralement notre corps », c’est-à-dire vivant sur et à l’intérieur de notre corps. Cette définition est plus proche de la notion de « microbiote » et est proche du sens courant du terme « microbiome ». Dans cette thèse, nous utiliserons principalement le terme de « microbiome » qui fera référence à la fois au microbiote et aux gènes portés par celui-ci.

Figure 1.1 : Nombre de publications traitant du microbiome humain depuis 1965. Les courbes représentent respectivement le nombre de publications comportant les mots-clefs (Microbiome OR Microbiota OR Microflora) ; (Microbiome OR Microbiota OR Microflora) NOT Human ; et (Microbiome OR Microbiota OR Microflora) NOT Human AND (sea OR marin*). La recherche a été faite à partir de la plateforme ISI Web of Knowledge (www.webofknowledge.com).
Le microbiome est principalement étudié chez l’Homme et chez des vertébrés terrestres. Je commencerais donc par résumer les principales connaissances sur la composition et le fonctionnement du microbiome chez ces animaux, puis je développerai les résultats moins nombreux chez les animaux marins.

1.1.2 Diversité et échelles de variabilité du microbiome associé aux animaux

A cause de son importante cruciale pour la santé humaine, l’essentiel des recherches concernant le microbiome des animaux terrestres ont été réalisées chez l’Homme et chez le modèle souris.

Le microbiome humain comporte une grande diversité d’eucaryotes, archées, bactéries et virus. Il a été démontré que nous hébergeons autant de cellules bactériennes à l’intérieur de notre tube digestif que de cellules humaines dans tout notre corps (Sender, Fuchs, and Milo 2016). Le nombre de gènes portés par le microbiote intestinal est au moins 150 fois plus important que le génome humain (Zhu, Wang, and Li 2010), et contribue à un large panel de fonctions biochimiques et métaboliques qui ne sont pas codées par le génome humain, si bien que certains chercheurs considèrent le microbiote comme un organe à part entière vital pour la survie de son hôte à tous les stades de sa vie (Neish 2009). Les cellules microbiennes composant le microbiome interagissent de multiples manières avec les cellules de l’hôte, jouant notamment un rôle crucial dans le développement du système immunitaire (Belkaid and Hand 2014). Un grand nombre de maladies humaines sont liées à un déséquilibre de la composition du microbiome, telles que des maladies intestinales, des maladies du foie ou encore de l’asthme (Carding et al. 2015). Le microbiome varie le long de la vie de son hôte. Chez l’homme, il est variable à la fois entre individus et au sein d’un même individu. Les principales interactions entre le microbiome intestinal humain et son hôte sont indiquées figure 1.2.
Figure 1.2 : Principales interactions entre le microbiome digestif humain et son hôte, et influences sur la santé humaine. Les cercles montrent l’évolution du microbiome avec l’âge et les différents facteurs influençant sa composition : facteurs liés aux interactions hôte-microbiome, y compris les potentielles interactions avec le microbiome maternel lors du développement intra-utérin (Axe Symbiosis), et facteurs externes tels que la prise d’antibiotiques ou le régime alimentaire (Axe Dysbiosis). Ces derniers facteurs peuvent déséquilibrer le microbiome digestif, entraînant des maladies, telles que (A) le syndrome du colon irritable (Inflammatory Bowel Disease, IBD), des ulcères gastriques, stéatoses hépatiques du foie ou d’autres maladies métaboliques comme l’obésité (B) mais aussi l’asthme ou l’hypertension (C). Enfin le microbiome digestif peut influencer le fonctionnement du cerveau (D). Figure extraite et légende adaptée de (Nicholson et al. 2012)

1.1.2.1 Echelles de variation du microbiome humain

Echelle Intra-individuelle
Les différents organes du corps humain sont colonisés par des dizaines de phylums bactériens, et des centaines d’espèces différentes. Au sein d’un individu, ces communautés sont très différentes à la fois en termes d’abondance, de richesse et de composition, entre les organes, mais aussi à l’intérieur d’un seul et même organe (Figure 1.3). Dans le cas du microbiome cutané par exemple, elles varient selon la zone du corps, principalement à causes de différences de pH, de température et d’humidité entre les différents microhabitats cutanés (Oh et al. 2014). Les microbiomes digestifs et cutanés sont des communautés dynamiques, évoluant constamment depuis la naissance de l’individu jusqu’à sa mort. Ainsi, différents facteurs exogènes (i.e. liés à des effets de l’environnement) et
endogènes (i.e. liés à l’hôte) peuvent influencer la composition du microbiome au cours de la vie de l’individu.

Le facteur alimentaire, par exemple, peut avoir des effets à court terme sur le microbiome digestif : un changement rapide d’habitudes alimentaires (David et al. 2014), ou la prise d’antibiotiques (Raymond et al. 2016) modifient fortement la structure du microbiome digestif. Cependant des études ont relevé la résilience du microbiome digestif, qui retrouve généralement un équilibre assez rapidement après ce type de perturbations (Dethlefsen et al. 2008), même si le rétablissement du microbiome initial n’est pas toujours atteint (Raymond et al. 2016).

Un facteur endogène important modifiant le microbiome de l’individu est son état de santé. Certaines maladies modifient en effet profondément le microbiome intestinal, modifications qui peuvent être à l’origine de complications de ces maladies (e.g. cas de la cirrhose du foie (Bajaj et al. 2014)). De même, la dérégulation du système immunitaire liée à la progression du VIH entraîne des profondes modification du microbiome intestinal, modifications qui progressent avec la maladie (Vujkovic-Cvijin et al. 2013).

![Figure 1.3 : Composition du microbiome bactérien dans les différentes parties du corps humain. Données extraites de (Grice and Segre 2011), (Fierer et al. 2012), (Hilty et al. 2010) et (Spor, Koren, and Ley 2011).](image)

**Échelle Inter-individuelle**

Des études menées sur plus de 200 individus ont démontré qu’il existe également une certaine variabilité inter-individuelle de la composition des micro-organismes intestinal, cutané et vaginal (The Human Microbiome Project Consortium 2012). Différents facteurs sont à l’origine de cette variabilité inter-individuelle, principalement documentée dans le cas du microbiome digestif. Les facteurs considérés à ce jour comme les plus importants selon la littérature sont les facteurs exogènes, qui peuvent influencer le microbiome à court et à long terme. Le premier facteur est le régime alimentaire, qui induit de fortes différences entre
populations humaines (Schnorr et al. 2014; Clemente et al. 2015) et au sein des populations ainsi, entre individus ayant des préférences alimentaires contrastées (Wu et al. 2011). Certaines études suggèrent que l’on peut retrouver au sein des populations humaines différents types de microbiomes digestifs, appelés « entérotypes », qui restent stables au cours de la vie des individus (Wu et al. 2011).

Des facteurs endogènes ont également des conséquences à long terme sur la structure et la composition du microbiome digestif. Le premier facteur est le facteur génétique. Ainsi, des individus d’une même famille possèdent un microbiome digestif plus proche que des individus non apparentés, et partagent un certain nombre de souches microbiennes identiques (Schwarz et al. 2008; Turnbaugh et al. 2009). Il a par exemple été démontré que des jumeaux monozygotes, génétiquement identiques, présentent un microbiome plus proche que des jumeaux dizygotes (Goodrich, Davenport, Beaumont, et al. 2016). D’autres études récentes basées sur des méthodes génomiques ont pu corréler les micro-variantes existant dans le génome entier des individus (Single Nucleotide Polymorphism, SNPs) avec leur microbiome cutané et digestif (Blekhman et al. 2015).

Dans le cas du microbiome cutané humain, les facteurs influençant sa variabilité inter-individuelle sont moins connus. Grice et Segre (2011) ont fait l’inventaire des différents types de facteurs pouvant influencer la composition du microbiome cutané (Figure 1.4). Suite à cette revue parue il y a 6 ans, des études plus récentes ont confirmé ces premiers résultats, à savoir l’influence de facteurs exogènes, tels que l’hygiène ou l’utilisation de cosmétiques et la zone du corps, et de facteurs endogènes tels que l’âge et le sexe de la personne. D’autres études ont permis d’importantes avancées dans la compréhension de l’assemblage du microbiome cutané, démontrant notamment l’impact de la génétique de l’hôte et du fonctionnement de son système immunitaire sur sa composition.
Figure 1.4: Principaux facteurs influençant la composition du microbiome cutané. Figure adaptée de Grice et Segre (2011). Les facteurs relevés par Grice et Segre qui ont été confirmés ou nuancés par la suite ont été colorés en bleu. Les facteurs publiés dans la littérature postérieurement à la publication de la revue ont été colorés en rouge. 1 Zeeuwen et al. (2012) ont analysé la composition du microbiome cutané après une lésion superficielle de la peau. Ils ont montré que le microbiome se reformait sur la cicatrice d’une composition proche de celle du microbiome colonisant les couches profondes de l’épiderme. Ils ont également retrouvé un effet du sexe des participants à l’étude dans toutes les couches de l’épiderme. 2 Ying et al. (2015) ont comparé le microbiome cutané provenant de 7 zones du corps de 71 personnes vivant dans les zones rurales et urbaines de la région de Shangai. Ils ont confirmé l’impact de la zone du corps, de l’âge et du sexe sur la composition du microbiome, ainsi qu’un effet de la zone dans laquelle vivent les participants. 3 Oh et al. (2012) ont comparé le microbiome cutané sur 4 zones du corps chez des volontaires âgés de 2 à 40 ans. Ils ont confirmé le fort impact de l’âge sur le microbiome cutané, dont la composition subit de profonds changements entre l’enfance et l’âge adulte. 4 Oh et al. (2016) ont analysé la composition du microbiome chez 12 participants sur 17 zones du corps et sa stabilité à long- et à court terme. Ils ont confirmé la forte spécificité du microbiome cutané en fonction de la zone du corps, qui reste stable au cours du temps, à la fois au niveau des phylums bactériens, mais également au niveau des souches microbiennes. 5 Leung et al. (2015) ont comparé la composition du microbiome cutané d’individus chinois, américains et Tanzaniens et ont montré qu’il existait de fortes différences entre ces populations. Ils ont également montré que des personnes occupant un même foyer possédaient un microbiome plus proche que des personnes ne vivant pas ensemble. 6 Perez et al. (2016) ont nuancé l’impact des populations humaines sur le microbiome cutané, montrant que des personnes de différents groupes ethniques vivant dans la même ville présentent peu de différences. 7 Chehoud et al. (2013) ont confirmé l’importance du système immunitaire dans la modulatio du microbiome cutané, en montrant que le complément module la composition du microbiome cutané ; ils ont également montré que le microbiome cutané module lui-même l’expression de gènes impliqués dans le complément. 8 et 9 Oh et al. (2013) et Smeekens et al. (2014) ont montré que des patients immunodéficients présentent un microbiome cutané déséquilibré, avec une augmentation de l’abondance relative de pathogènes opportunistes. 10 Blekhman et al. (2015) ont corrélaté la composition du microbiome cutané avec les variations génétiques entre individus. Ils ont montré que des changements d’allèles de gènes impliqués dans les processus immunitaires étaient corrélés à la variabilité du microbiome sur 10 des 15 zones du corps étudiées. 11 Bouslimani et al. (2015) ont réalisé une cartographie moléculaire et microbienne de deux volontaires. Ils ont montré que des variations de concentration de molécules utilisées dans les savons et les cosmétiques étaient corrélées avec la composition locale du microbiome.

**Echelle inter-spécifique**


Les mécanismes à l’origine des différents niveaux de variabilité décrits plus haut ne sont pas encore entièrement démêlés ; dans le cas de la variabilité intra-individuelle, la corrélation entre la présence de certaines espèces microbienennes et caractéristiques physico-chimiques du micro-habitat suggèrent un effet de filtre environnemental fort sur la composition de la communauté microbienne locale. A l’échelle inter-individuelle, des mécanismes de transferts verticaux et horizontaux, interagissant avec des mécanismes de filtres environnementaux liés à l’alimentation par exemple, peuvent engendrer une individualité forte du microbiome digestif. Enfin, à l’échelle interspécifique, les patrons de phylosymbiose pourraient s’expliquer par des phénomènes de coévolution entre l’hôte et son microbiome. Dans la section suivante, j’aborderai les mécanismes d’assemblage des communautés pouvant être à l’origine des différentes échelles de variabilité observées.

### 1.1.3 Mécanismes à l’origine de la variabilité des microbiomes

Le microbiome forme une communauté complexe de micro-organismes ayant des phénotypes variés, qui interagissent à la fois entre eux et avec les cellules de leur hôte. En ce sens, le corps humain peut être considéré comme un écosystème à part entière, et la santé humaine comme une résultante partielle des différents services écosystémiques rendus par les communautés qui l’habitent (Costello et al. 2012).

Le microbiome d’un individu est constitué d’un ensemble de différentes communautés, inféodées chacune à un organe ou une partie du corps. Par exemple, comme expliqué plus haut, les communautés colonisant la peau du visage, des bras, ou l’intérieur de la bouche sont très différentes. Cependant ces communautés ne sont pas isolées : des microorganismes peuvent être transférés entre ces différentes zones du corps et avec le milieu ambiant (air, sol, eau, autres organismes). Ces communautés forment donc une métacommunauté, définie comme un ensemble de communautés interconnectées entre elles par des phénomènes de dispersion des organismes (Leibold et al. 2004).

La composition d’une métacommunauté sera donc dépendante de la composition de ses communautés locales, qui peut être influencée par les capacités de dispersion des microorganismes, des processus de *species sorting*, ainsi que des *processus stochastiques* (Costello et al. 2012).

### 1.1.3.1 Species Sorting

Le processus de *species Sorting* regroupe l’ensemble des facteurs abiotiques de l’environnement local et les interactions biotiques entraînant un « tri » des espèces microbiennes dans un milieu donné, tri qui dépendra de l’adéquation entre la niche écologique de l’espèce et les conditions du milieu (Leibold et al. 2004). Dans le cas du microbiome associé aux animaux, le *species Sorting* agit de trois manières complémentaires.

Tout d’abord, on peut voir le micro-environnement constitué par la petite surface de peau ou de muqueuse comme un environnement abiotique présentant un ensemble de conditions physico-chimiques particulières, appelés filtres environnementaux, dans lesquels ne survivront que les cellules microbiennes possédant des exigences écologiques compatibles (Costello et al. 2012; Kraft et al. 2015). Dans le cas du microbiome digestif humain, par exemple, le principal filtre environnemental est l’oxygène : on sait qu’il existe un gradient d’oxygène dans l’intestin, favorisant la croissance de bactéries anaérobies facultatives dans les parties antérieures (début de l’intestin grêle), et celle de bactéries
anaérobies strictes dans le côlon qui est quasiment anoxique (Donaldson, Lee, and Mazmanian 2015).

D’un autre côté, les cellules microbiennes et les cellules de l’hôte peuvent être vues comme une seule et même communauté présentant des interactions. Dans ce cas, l’hôte macroscopique n’est pas un support passif, mais possède au contraire un rôle actif dans le façonnement des communautés microbiennes qui l’habitent (Bevins and Salzman 2011). De multiples exemples démontrent l’existence de ces interactions directes entre les cellules de l’hôte et les cellules du microbiome, notamment avec les cellules immunitaires. Par exemple, les récepteurs de type Toll, portés par les cellules du système immunitaire inné, et permettant normalement de reconnaître des micro-organismes pathogènes et d’enclencher une inflammation, reconnaissent également des motifs portés par de nombreuses bactéries bénéfiques colonisant notre intestin. Dans ce cas, la réaction déclenche au contraire la production par l’hôte de facteurs favorisant la protection des tissus de l’intestin et l’homéostasie du tube digestif (Rakoff-Nahoum et al. 2004). Ce type d’interaction peut notamment expliquer les variations interindividuelles du microbiome liées aux micro-variations existant dans le génome des individus, notamment dans les gènes du système immunitaire (Blekhman et al. 2015).

Enfin, en plus des processus de sélection « passive » (rôle de filtre environnemental de l’hôte) et « active » (interactions directe entre les cellules de l’hôte et son microbiome), un troisième mécanisme intervenant dans le species sorting est l’interaction entre les espèces microbiennes elles-mêmes, qui peut entrainer l’extinction de certaines espèces, ou favoriser la croissance d’autres. Les différents types d’interactions directes sont fournis dans le tableau 1.1. Cependant, une étude focalisée sur le microbiome digestif humain a démontré que ces interactions biotiques (à l’exception des relations de compétition) seraient moins importantes dans la structuration du microbiome que les filtres environnementaux (Levy and Borenstein 2013).
<table>
<thead>
<tr>
<th>Interaction</th>
<th>Définition</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Compétition</strong></td>
<td>Deux espèces possédant une niche écologique similaire ne peuvent cohabiter durablement si une ressource est limitante. Celles-ci vont entrer en compétition, et l’espèce la plus adaptée à la niche sera favorisée au cours du processus de sélection naturelle. Celle-ci conduit à la diminution des effectifs voire à la disparition de certaines espèces, ou à la spécialisation des espèces pour certaines ressources. Les interactions compétitives sont des forces majeures dans la modulation de la composition des communautés microbiennes. Dans le cas du microbiome digestif humain, des chercheurs ont démontré à l’aide de simulations que les interactions compétitives entre les espèces qui le composent permettraient de réduire sa diversité et d’augmenter sa stabilité.</td>
</tr>
<tr>
<td><strong>Prédation</strong></td>
<td>Autrement appelée « régulation top-down », la prédation est un mode de nutrition consistant à tuer d’autres organismes pour se nourrir de leur substance. Les prédateurs de bactéries (bactéries, protozoaires, virus bactériophages) ont un impact important sur la diversité des communautés bactériennes, car ils écrétent préférentiellement les espèces les plus abondantes, laissant ainsi une place aux espèces moins compétitives (hypothèse kill the winner). Ces prédateurs ont notamment été mis en évidence dans le mucus du corail et pourraient participer au contrôle des pathogènes.</td>
</tr>
<tr>
<td><strong>Amensalisme</strong></td>
<td>On parle d’amensalisme quand l’action d’une espèce nuit à une autre, sans que celle-ci ne semble en tirer un bénéfice direct. Cette relation est observée notamment lorsque certaines espèces modifient localement leur milieu, suite à des activités enzymatiques nécessaires à leur métabolisme, mais aussi par la synthèse de composés inhibiteurs (e.g. synthèses d’antibiotiques ou d’antifongiques). Ces interactions antagonistes pourraient jouer un rôle important dans le microbiome associé aux animaux, en éliminant les micro-organismes pathogènes.</td>
</tr>
<tr>
<td><strong>Mutualisme</strong></td>
<td>Les interactions mutualistes sont des interactions bénéfiques aux deux partenaires. Elles sont fréquemment observées lorsque deux partenaires microbiens dépendent l’un de l’autre pour la dégradation de certaines ressources trophiques (syntrophie). Dans le microbiome intestinal humain, il existe par exemple une relation syntrophique entre des bactéries et des archées métanogènes et acétogènes permettant la digestion des glycogènes. Un autre exemple d’interaction coopérative est l’organisation des cellules en microcolonies ou en biofilms à la surface des muqueuses ou de la peau, permettant notamment la résistance aux antibiotiques et à la prédation. Cependant, si de nombreux exemples de coopération existent, le rôle relatif des interactions coopératives par rapport aux autres types d’interaction est débattu dans le cas des communautés microbiennes.</td>
</tr>
</tbody>
</table>

**Tableau 1.1 : Les différentes interactions entre micro-organismes pouvant modifier la composition des communautés microbiennes associées aux macro-organismes.** Références : 1 (Gause 1934) ; 2 (Whittaker 1965) ; 3 (Foster and Bell 2012) ; 4 (Coyte, Schluter, and Foster 2015) ; 5 (Thingstad 2000) ; 6 (Welsh et al. 2016) ; 7 (Efrony et al. 2007) ; 8 (Lowrey et al. 2015) ; 9 (Koropatkin, Cameron, and Martens 2012) ; 10 (Macfarlane and Dillon 2007) ; 11 (Oliveira, Niehus, and Foster 2014).
1.1.3.2 Dispersion

La dispersion est le mouvement des micro-organismes d’un patch (*i.e.* de l’habitat d’une communauté locale) à un autre. Dans le cas des communautés de micro-organismes, ces phénomènes de dispersion sont très majoritairement passifs à cause de leur petite taille (Trousselier et al. 2017). Dans le cas du microbiome humain, l’essentiel de ces phénomènes de dispersion se fait certainement par contacts cutanés. Ces mécanismes de transfert ont un impact majeur sur la composition du microbiome, et peuvent avoir lieu horizontalement (*i.e.* entre individus, ou entre individus et leur milieu environnant) ou verticalement (*i.e.* entre un individu et sa progéniture). Les différents mécanismes de transferts sont discutés ci-après. Dans cette partie, je ne traiterais pas des phénomènes de transmission verticale des endosymbiontes, qui sont intracellulaires et qui présentent des mécanismes de transfert particuliers (Bright and Bulgheresi 2010).

**Mécanismes de transfert vertical**

Le transfert vertical désigne le passage de micro-organismes de la mère ou du père à la descendance. Ces transferts sont permis par différents types de mécanismes, dépendant principalement du mode de reproduction de l’organisme (Figure 1.5). Si quelques études démontrent que ces transferts existent chez tous les êtres vivants (Funkhouser and Bordenstein 2013), leur mécanisme et leur dynamique ont principalement été étudiés dans le cas du microbiome digestif de l’homme. De plus, l’ensemble de ces études concerne le microbiome digestif, ou le microbiome interne. La possibilité d’un transfert vertical associé à la surface des macro-organismes n’a, à ma connaissance, jamais été explorée.
Figure 1.5 : Différentes voies possibles pour la transmission verticale du microbiome. Les éponges et les coraux peuvent transmettre des cellules microbiennes via leurs ovocytes. Chez les organismes ovipares, il existe une contamination de l’intérieur de l’œuf. Au moment de l’éclosion il peut exister une contamination du juvénile par les micro-organismes se trouvant à la surface de l’œuf, pouvant provenir des parents si l’œuf est couvé. Chez les organismes vivipares, le juvénile est contaminé par le microbiome maternel lors de l’accouchement. Après la naissance, des inocula du microbiome parental peuvent encore exister, via l’allaitement, les soins parentaux, ou encore la proximité physique entrainant des transferts horizontaux.

- Mécanismes transovariens ou via l’ovocyte. Chez les invertébrés marins, notamment les éponges, il existe une transmission verticale du microbiome interne (Webster et al. 2010), via l’ovocyte, ou directement entre les tissus adultes et l’embryon (Thacker and Freeman 2012). Chez les coraux, une transmission verticale du microbiome a également été démontrée, mais semble dépendre du mode de fécondation (Bourne, Morrow, and Webster 2016). Chez les mollusques, la transmission transovarienne de cellules microbiennes existe également (Cary and Giovannoni 1993). On retrouve également ce type de transmission chez des animaux ovipares, où l’intérieur même de l’œuf peut être contaminé lors de sa formation par des microsporidies et des particules virales, comme cela a été démontré chez certaines espèces de poissons (Funkhouser and Bordenstein 2013). Cependant l’importance relative de ce type de transfert par rapport à d’autres mécanismes est inconnue.
• **Mécanismes intra-utérins.** Chez les animaux placentaires, on a longtemps cru que le placenta était une barrière stérile empêchant la contamination du fœtus par des bactéries pathogènes. Les avancées du séquençage à haut débit ont permis de montrer qu’au contraire, le placenta était colonisé par des micro-organismes diversifiés (Aagaard et al. 2014), présents en faible abondance. Ces micro-organismes pourraient être transmis au fœtus et ainsi assurer la transmission verticale de bactéries bénéfiques (Wassenaar et Panigrahi 2014).

• **Mécanismes quia-partum.** Lors de la naissance chez les organismes vivipares, le tube digestif du nourrisson, quasiment stérile *in-utero*, est colonisé par le microbiome maternel, vaginal, ou cutané en fonction du mode d’accouchement (Bäckhed et al. 2015). Chez les animaux non mammifères, l’importance et les mécanismes de ces transferts sont loin d’être éucidés. Chez une espèce de lézard vivipare, il a été démontré que des petits séparés de leur mère dans les 24h suivant leur naissance partageaient un tiers d’espèces bactériennes intestinales avec leur mère, ce qui souligne l’importance de la contribution maternelle via l’inoculum initial, mais ne démèlent pas les mécanismes à l’origine du transfert (Kohl et al. 2017). Chez les organismes ovipares, les juvéniles peuvent également être colonisés par les micro-organismes présents à la surface de l’œuf, qui peuvent provenir du microbiome maternel ou paternel (Funkhouser et Bordenstein 2013).

• **Mécanismes post-partum.** Chez les mammifères, des bactéries sont également transmises de la mère à son petit via l’allaitement (Funkhouser et Bordenstein 2013). Chez l’Homme, cet inoculum maternel (*i.e.* lié au mode d’accouchement et à l’allaitement) va préparer les conditions anaérobies dans le tube digestif du nouveau né pour ensuite permettre la colonisation par d’autres souches microbiennes après la phase d’allaitement, puis sa stabilisation progressive (Bäckhed et al. 2015). Ce processus de maturation du microbiome va permettre de mettre en place un microbiome digestif ressemblant à celui d’un adulte un an après l’accouchement (Bäckhed et al. 2015). Malgré les différentes successions microbiennes intervenant dans le tube digestif du nourrisson et la disparition apparente de l’inoculum maternel, on peut retrouver cet effet maternel chez l’adulte. Une étude menée sur quatre générations de souris a ainsi montré que des souris apparentées avaient un microbiome plus proche que des souris non apparentées (Ley et al. 2005). Si la persistance de souches maternelles liées à cet inoculum initial (*i.e.* non liée à des inoculations ultérieures) chez l’adulte n’a pas été prouvée, d’autres facteurs liés à l’allaitement peuvent favoriser un certain conservatisme vertical du microbiome, notamment via la fourniture d’anticorps qui vont potentiellement contrôler la mise en place de son système immunitaire (Bordenstein et Theis 2015). Chez un poisson d’eau douce, le discus commun (*Symphysodon aequifasciata*), un mécanisme de transfert particulier a été mis en évidence, provenant à la fois des
microbiomes maternel et paternel (Sylvain and Derome 2017). En effet ce poisson présente la particularité de sécrèter un mucus cutané particulier, dont vont se nourrir les juvéniles après la naissance, entraînant un changement important de la composition du microbiome digestif. Cependant ce comportement est assez anecdotique chez les poissons, et est partagé par une trentaine d’autres espèces seulement (Noakes 1979). Dans le cas d’autres espèces de poissons, la possibilité d’un transfert vertical du microbiome digestif ou cutané n’a jamais été étudiée. D’autres mécanismes post-partum de transfert du microbiome peuvent exister. C’est le cas notamment de tous les phénotypes de soin aux jeunes (e.g. lécher le museau chez les ruminants), ou de cohabitation dans un même terrier ou un même nid (Spor, Koren, and Ley 2011). Chez certaines espèces, le juvénile se nourrit des fèces parentales, ce qui assure la colonisation de son tube digestif par les micro-organismes nécessaires à sa survie. C’est le cas de nombreux mammifères terrestres, tels que le koala, l’éléphant et le panda géant (Spor, Koren, and Ley 2011).

Une étude récente, basée sur les métagénomomes et métabolomes de microbiome digestif de 30 souches de souris, a démontré que l’environnement auquel sont exposées les souris à la naissance, ainsi que les différences génétiques entre celles-ci, sont des facteurs prépondérants dans la composition du microbiome, et qu’un changement d’environnement au cours de la vie de la souris ne modifie pas substantiellement la composition du microbiome. Cela suggère que l’inoculum maternel a une influence très importante sur la composition du microbiome, qui subsiste à l’âge adulte (Snijders et al. 2016). On parle dans ce cas de priority effect.

Mécanismes de transfert horizontal
Il existe deux types de transferts horizontaux, le premier s’effectue entre l’organisme et son environnement immédiat, et le deuxième a lieu entre organismes.

- Transferts depuis l’environnement. En plus des mécanismes de transfert verticaux, une part de la transmission du microbiome semble se faire de manière horizontale chez certains animaux ; chez des invertébrés marins tels que des moules, des éponges ou des coraux, une partie du microbiome n’est pas transmise verticalement, mais horizontalement depuis l’eau de mer (Hentschel et al. 2012; Nyholm and McFall-Ngai 2004). Les juvéniles du calamar Euprymna scolopes utilisé comme organisme modèle en en biologie à cause de ses relations symbiotiques avec des bactéries, acquiert son symbionte bioluminescent (Vibrio fischeri) depuis l’eau de mer à chaque génération (Nyholm and McFall-Ngai 2004).

- Transferts entre animaux. La plupart des transferts horizontaux intra-spécifiques ont lieu entre individus apparentés, physiquement ou géographiquement proches, ou
encore faisant partie du même groupe social. En ce sens, les transferts horizontaux favorisent également la transmission verticale du microbiome. Chez les insectes sociaux par exemple, la transmission du microbiome digestif par trophallaxie ou coprophagie est un phénomène important, permettant le transfert du microbiome de génération en génération (Koch and Schmid-Hempel 2011; Brune and Dietrich 2015). D’autres animaux ne présentant pas de comportements coprophages présentent des transferts horizontaux du microbiome digestif. Chez l’Homme, il a été démontré que des personnes vivant sous le même toit avaient des microbiomes digestif, cutané et oral plus proches que des personnes ne vivant pas ensemble (Song et al. 2013). Il a notamment été démontré que nous échangions en moyenne $8.8 \times 10^8$ bactéries lors d’un baiser, et que ce transfert modifiait transitoirement la composition du microbiome buccal (Kort et al. 2014). Dans la même étude, les chercheurs ont également montré que ces échanges réguliers de microorganismes entre les partenaires avaient également des conséquences à long terme sur le microbiome, puisque des individus en couple avaient en moyenne des communautés microbiennes buccales plus semblables qu’avec d’autres individus. Ces transferts de microbiome peuvent également exister en dehors du cercle familial, à la fois par des contacts directs (Meadow et al. 2013) et indirects, par exemple via le partage d’un même environnement de travail (Ross and Neufeld 2015).

L’échange de micro-organismes entre individus ou entre le milieu et un hôte peut donc modifier les densités ou la diversité microbienne chez l’un ou l’autre des partenaires d’échange et entraîner des phénomènes de mass effect allant jusqu’à un rescue effect. Le phénomène de mass effect est une dynamique spatiale dans laquelle une différence de nombre ou de densité des organismes entre deux habitats entraîne un flux net d’individus vers l’un des deux habitats et augmente l’abondance d’espèces qui ne sont normalement pas adaptées à ce deuxième habitat (Leibold et al. 2004). De manière similaire, le phénomène de rescue effect décrit une dynamique dans laquelle l’extinction d’une espèce dans une communauté locale est empêchée par les phénomènes d’immigration (Brown and Kodric-Brown 1977).

### 1.1.3.3 Processus stochastiques ou ecological drift

En plus des changements liés à l’immigration des espèces et à leur sélection par le milieu, l’abondance relative des espèces microbiennes au sein d’une communauté peut varier suite à des processus stochastiques, dénommés ecological drift, qui peuvent entraîner l’extinction locale d’espèces présentes en faible abondance. Par exemple, des espèces spécialistes dépendantes d’un substrat présent en faible quantité dans le milieu, ou encore des espèces venant d’immigrer dans la communauté (Costello et al. 2012).

Un exemple extrême d’ecological drift est la théorie neutre (Hubbell 2001), assumant que l’abondance locale des espèces est le résultat de processus stochastiques liés à la dispersion et à la mortalité des individus, les espèces ayant des capacités égales de croissance dans un
milieu donné. La théorie neutre, ignorant les processus de species sorting, permet d’expliquer les dynamiques de certaines communautés, où les espèces sont fonctionnellement proches (Leibold et al. 2004). Le microbiome intestinal humain, qui présente une forte redondance fonctionnelle (Lozupone et al. 2012), pourrait constituer une bonne application de cette théorie (Fierer et al. 2012). Dans une étude récente basée sur des méthodes de métagénomique, les chercheurs ont tenté de quantifier l’importance relative des processus neutres et des processus de niche écologique dans la structuration de la diversité du microbiote intestinal humain. Même si la structure des communautés prédite par la théorie neutre reflétait assez bien la distribution des espèces observées dans l’intestin, les modèles basés sur les processus de niche écologique faisaient de meilleures prédictions (Jeraldo et al. 2012). L’influence relative de ces phénomènes stochastiques est en fait dépendante du stade de vie de l’individu. Ainsi, une étude menée chez le poisson zèbre (Danio rerio) a montré que la composition du microbiome digestif des alevins était essentiellement aléatoire, démontrant l’importance des phénomènes stochastiques lors de la colonisation initiale du tube digestif. Cependant au fur et à mesure de la maturation de l’hôte, la composition de son microbiome s’écartait des prédictions faites sous une théorie neutre, suggérant l’influence croissante de processus de niche au fur et à mesure de la maturation du système immunitaire et du tube digestif de l’hôte (Burns et al. 2016).

1.1.3.4 Evolution
Les mécanismes de transferts du microbiome, à la fois verticaux et horizontaux entre proches parents, peuvent, s’ils sont reproduits fidèlement à chaque génération, engendrer un processus de coévolution entre l’hôte et les différentes espèces microbiennes qu’il héberge (Bordenstein and Theis 2015). Ce mécanisme pourrait expliquer les patrons de phylosymbiose observés à l’échelle interspécifique. Ce mécanisme n’est cependant pas obligatoire pour l’établissement d’une phylosymbiose (Brooks et al. 2016). En effet, le terme de phylosymbiose repose sur l’hypothèse que les variations phénotypes des hôtes au cours de l’évolution vont entraîner une certaine congruence de leurs microbiomes associés chez des espèces proches, qui peuvent apparaître même si le microbiome est acquis depuis l’environnement à chaque génération (Brucker and Bordenstein 2012; Brooks et al. 2016). Ainsi des espèces phylogénétiquement proches auront des physiologies, des systèmes immunitaires et un métabolisme plus proche que des espèces non apparentées. Par simple effet de filtre environnemental, elles vont donc sélectionner des communautés bactériennes plus proches que des espèces plus éloignées. De plus, des espèces phylogénétiquement proches auront tendance également à posséder les mêmes niches écologiques, potentiellement habitées par des communautés microbiennes semblables ; elles auront donc tendance à être colonisées par le même type d’inoculum微观ien (Douglas and Werren 2016). La figure 1.6 résume les différents mécanismes pouvant engendrer un patron de phylosymbiose.
Figure 1.6 : Différents mécanismes à l'origine de la phylosymbiose. Une proximité de la composition du microbiome de deux espèces proches peut être liée à une proximité physiologique ou écologique, favorisant la colonisation et/ou la croissance des mêmes clades bactériens. Les transferts horizontaux entre proches parents peuvent favoriser les phénomènes de coévolution entre l’hôte et les différents membres du microbiome qui lui est associé. Récemment, une étude menée sur des insectes a également démontré que ces phénomènes de coévolution pourraient être favorisé par une mort des hybrides enclenchée par le microbiome digestif (Brucker and Bordenstein 2013).

Le terme de phylosymbiose ne présuppose pas de transmission verticale du microbiome, et certains auteurs le préfèrent donc aux notions d’holobionte et d’hologénome, qui cherchent à expliquer les mêmes patrons sous un angle évolutif. Le terme d’holobionte est utilisé pour désigner un hôte multicellulaire et l’ensemble des micro-organismes qui lui sont associés, et a été à l’origine utilisé pour désigner les coraux et leur microbiome (Rohwer et al. 2002). L’hologénome désigne l’ensemble de l’information génétique de l’hôte et de son microbiome associé. Ces deux termes sont le fondement d’une théorie dénommée *hologenome theory of evolution*, proposée en 2007 puis explicitée en 2008 par Ilana et Eugene Rosenberg (Zilber-Rosenberg and Rosenberg 2008), qui considèrent l’holobionte et son hologénome comme une seule et même entité soumise à la sélection naturelle. Cette théorie est basée sur plusieurs hypothèses : (i) que tous les êtres vivants sont associés à des
communautés microbiennes, (ii) que ces micro-organismes sont transmis fidèlement de génération en génération, (iii) que les associations entre l’hôte et son microbiome affectent la fitness globale de l’holobionte, qui peut être sélectionné ou au contraire disparaître, (iv) que des processus de diversification se font également à l’échelle de l’hologénome, et que des variations peuvent avoir lieu à la fois dans le microbiome et dans le génome de l’hôte.

Cette théorie est très débattue (Moran and Sloan 2015; Theis et al. 2016; Douglas and Werren 2016); en effet, s’il ne fait aucun doute que la très grande majorité, si ce n’est la totalité des animaux et des végétaux sont colonisés par des micro-organismes, que ceux-ci peuvent avoir un rôle très important dans leur survie et/ou leur reproduction, le fait de considérer l’hologénome comme une seule et même entité soumise à la sélection pose question : tout d’abord, cela suppose une transmission verticale fidèle du microbiome (Bordenstein and Theis 2015; Douglas and Werren 2016). Comme vu précédemment, ces transferts existent, mais ils ne concernent certainement pas l’ensemble du microbiome, dont une partie est acquise depuis l’environnement. D’autre part, si effectivement le microbiome influence la fitness de son hôte, la réciproque, dans beaucoup de cas, n’est pas vraie (Douglas and Werren 2016). De nombreux micro-organismes colonisant des êtres vivants sont capables de persister, voire de se reproduire dans l’environnement. Par conséquent, si le macro-organisme meurt, cela n’affecte pas forcément négativement la fitness des bactériens qui colonisaient son tube digestif.

Cependant, les notions d’holobionte et d’hologénome peuvent également être appréhendées dans un sens plus souple ; incorporant des composantes du microbiome qui sont transmises verticalement avec leur hôte, affectent la fitness de leur hôte et évoluent avec lui, et des composantes moins dépendantes de l’hôte qui peuvent ou non affecter sa fitness (Figure 1.7) (Theis et al. 2016).
Figure 1.7 : Différentes composantes de l’holobionte. L’holobionte désigne un hôte macroscopique et l’ensemble des micro-organismes qui lui sont associé. Il comprend des membres qui affectent (de manière bénéfique ou non) la fitness de l’holobionte, qui peuvent co-évoluer avec leur hôte (en bleu) ou non (en rouge). D’autres n’affectent pas le phénomène de l’holobionte (en noir). A noter que les gènes de l’hôte sont également soumis à cette définition : cela permet de distinguer les gènes neutres des gènes bénéfiques ou détrimentaux, et également de distinguer les gènes conservés au cours des générations, et ceux qui subissent des mutations ou des recombinations chromosomiques. De plus, dans cette figure, les auteurs distinguent la notion de microbiome, qui englobe également la partie variable du microbiome, qui peut être acquise de l’environnement et ne fait pas partie du « core microbiome » de l’espèce. Figure tirée de (Theis et al. 2016).

1.2 Cas des microbiomes associés aux organismes marins


En plus de ces organismes eucaryotes, les océans présentent une importante diversité de procaryotes. Les dernières expéditions d’échantillonnage de la biodiversité procaryotique marine (expéditions Tara Oceans) ont retrouvé plus de 35 000 espèces dans 243 échantillons.
pélagiques provenant de tous les océans, et de différentes profondeurs (Sunagawa et al. 2015). Un très grand nombre d’études ont montré depuis maintenant plus de 30 ans que ces microorganismes planctoniques ont des fonctions essentielles dans tous les grands cycles biogéochimiques de la planète et sont de ce fait obligatoires pour le bon fonctionnement de l’écosystème global (Kirchman 2010).


D’autre part, une érosion de la diversité des macro-organismes aurait par nature des répercussions sur les espèces microbiennes associées. Aujourd’hui les changements globaux ont déjà des conséquences alarmantes sur la diversité macroscopique des écosystèmes terrestres, et ont entraîné une diminution de l’habitat et de la taille des populations de plus de 8 000 espèces de vertébrés (Ceballos et al. 2015). Les écosystèmes marins sont aussi en première ligne face aux perturbations anthropiques (Thorne-Miller 1999). S’il est plus difficile d’appréhender l’ampleur des extinctions des animaux marins, on sait qu’elles altèreront le fonctionnement des écosystèmes océaniques (Worm et al. 2006). Les écosystèmes coralliens, notamment, sont particulièrement vulnérables au changement climatique et aux pressions de pêche (McClanahan, Graham, and Darling 2014); le réchauffement climatique ayant entraîné des épisodes de blanchiment sans précédent et à l’échelle mondiale ces deux dernières années (Hughes et al. 2017). Les pertes de diversité des microbiomes associés sont aujourd’hui inconnues.

1.2.1 Composition et rôles du microbiome chez les invertébrés marins
Les invertébrés marins sont représentés par plus de 30 phylums (Lesser 2011). Ce sont les animaux les plus anciens ; les premiers fossiles de cnidaires datant d’il y a environ 580 millions d’années (Norse 1993). Par la suite, ils ont évolué très rapidement durant le Mésozoïque (-252 à -66 Ma, (Alroy et al. 2008)), et présentent aujourd’hui une grande diversité de plans d’organisations, de morphologies et de physiologies.
Parmi ces invertébrés, les coraux (sensu Classe Anthozoa) sont de loin les plus étudiés, du fait de leur rôle d’espèces ingénieuses dans les écosystèmes coralliens, fournissant un habitat à plus de 4 000 espèces de poissons (Kaiser 2011), et dont la production primaire supporte environ 25% de toutes les espèces marines connues (Bertness, Gaines, and Hay 2001). Ces écosystèmes sont vulnérables, du fait de la sensibilité des coraux face au changement climatique (Kaiser 2011). En effet, en cas d’augmentation de la température de l’eau, beaucoup d’espèces de coraux perdent leurs algues symbiotiques (Symbiodinium spp) leur permettant de fixer le carbone ; ce blanchiment, bien que réversible, peut aboutir à la mort des coraux s’il se prolonge dans le temps (Kaiser 2011).

En plus des algues photosynthétiques, les coraux abritent une diversité importante de micro-organismes procaryotes, eucaryotes, mais également des virus (Nguyen-Kim et al. 2014; Bourne, Morrow, and Webster 2016). Ces communautés sont souvent spécifiques de l’espèce de corail, mais varient selon la physiologie de l’hôte, le micro-habitat à l’intérieur de l’hôte (Figure 1.8), le stade biologique et l’environnement du corail (Bourne, Morrow, and Webster 2016; Ainsworth et al. 2015). Les organismes procaryotes hébergés par le corail participent à la fixation et la dégradation du carbone, et réalisent les réactions de fixation de l’azote, de nitrification et de dénitrification essentielles à la survie de leur hôte. Ils participent également à la protection de leur hôte contre les pathogènes (Krediet et al. 2013). Ces communautés sont sensibles aux stress environnementaux, qui entraînent leur déséquilibre, l’altération leur fonctionnement et contribuent au blanchiment du corail et à sa plus grande sensibilité aux maladies (Rosenberg and Ben-Haim 2002; Nguyen-Kim et al. 2015; Bettarel et al. 2015; Zaneveld et al. 2016).
Les éponges sont également des membres importants de la faune benthique marine (Bell 2008), en particulier dans les récifs coralliens (Hentschel et al. 2012). Malgré leur activité de filtration de l’eau environnante, les communautés microbiennes associées aux éponges présentent des espèces typiques, qui sont très rares dans l’eau de mer (Taylor et al. 2013). Ces communautés sont très diversifiées et présentent des densités de populations extrêmement élevées: il a été estimé qu’un centimètre cube de tissu héberge plus de $10^5$ cellules microbiennes ; ensemble, ces cellules microbiennes constituent 35% de la biomasse de l’éponge (Taylor et al. 2007). Ces micro-organismes sont principalement retrouvés dans la matrice extracellulaire entourant le choanoderme, appelée en anglais mesohyl, et constituant la majeure partie du corps de l’éponge (Hentschel et al. 2012) (Figure 1.8). Les communautés microbiennes associées aux éponges présentent une diversité phylogénétique importante, qui dépend de l’espèce hôte (Hentschel et al. 2012). Une étude récente focalisée sur le microbiome de 804 espèces d’éponges échantillonnées dans les océans
Atlantique et Indien, et dans les mers Méditerranée et Rouge a prouvé que la grande majorité des espèces bactériennes retrouvées sur les éponges était spécifique de l’espèce hôte, et a également mis en évidence un patron de phylosymbiose (voir partie 1.1.4) (Thomas et al. 2016). Cette découverte, associée au fait que les éponges sont capables de transmettre verticalement certains de leurs symbiontes, suggère une coévolution possible entre ces organismes et leur microbiome.


La composition du microbiome interne et externe des milliers d’espèces n’appartenant pas aux classes d’invertébrés cités ci-dessus est encore inconnue, notamment celle des échinodermes (holothuries, oursin, étoiles de mer), des cnidaires comme les méduses, ou des crustacés.

### 1.2.2 Composition et rôles du microbiome chez les vertébrés marins

Le microbiome des vertébrés marins (poissons téléostéens, chondrichthyens, tortues et mammifères marins), est assez peu étudié (Figure 1.1) en comparaison des invertébrés marins tels que les coraux ou les éponges. Pourtant, ceux-ci constituent la biomasse macroscopique la plus importante dans les écosystèmes marins, et ont des fonctions essentielles dans les écosystèmes, telles que le contrôle de l’abondance des proies et le recyclage des nutriments. Tout comme chez l’Homme et les autres modèles mentionnés précédemment, leur état physiologique est en partie conditionné par leur microbiome, colonisant leur tube digestif, leur peau, leur appareil respiratoire ainsi que leurs appareils génitaux et urinaires. Parmi le petit nombre d’études sur ces organismes, l’essentiel s’est focalisé sur le microbiome digestif.

#### 1.2.2.1 Microbiome digestif

Une étude a par exemple montré un effet important du stade de vie du saumon (*Salmo salar*) sur son microbiome, qui était lié aux périodes de transition entre l’eau de mer et l’eau douce (Llewellyn et al. 2016). Chez une espèce de poisson euryhaline (*Poecilia sphenops*) élevée en conditions contrôlées, un changement progressif de la salinité a entraîné un turn-over important des bactéries dominantes colonisant son tube digestif, turn-over qui n’était pas lié à la modification des communautés planctoniques environnantes ; ce qui suggère que les modifications liées à l’eau de mer ne sont pas liées aux communautés planctoniques, mais bien à une modification de la physiologie du poisson.

Des études ont également démontré une spécificité du microbiome pour l’espèce hôte, qui s’explique en partie par des différences de régime alimentaire, mais pas seulement. Ainsi, une étude récente basée sur des poissons coralliens (famille *Acanthuridae*) a démontré qu’en plus de l’alimentation, le microbiome était en partie corrélé avec la phylogénie du poisson, soulignant l’impact possible d’autres facteurs physiologiques caractéristiques à l’espèce (*e.g.* système immunitaire) (Miyake, Ngugi, and Stingl 2015). De plus, une étude basée sur des méthodes d’empreinte moléculaire (GDDE, voir chapitre 2) et une mesure de la diversité métabolique potentielle (Biolog Ecoplate®) du microbiome digestif de 15 espèces de poissons a également suggéré une importante redondance fonctionnelle du microbiome entre poissons ayant des régimes alimentaires contrastés (Mouchet et al. 2012).

L’ensemble des études présentées dans le Tableau 2 suggère qu’il existe, tout comme dans le cas du microbiome digestif humain, une combinaison de facteurs physiologiques et environnementaux conditionnant le microbiome digestif des poissons télopéristéens. Cependant, l’influence relative de ces deux facteurs, et leurs conséquences sur le fonctionnement de l’écosystème intestinal et la santé de l’hôte restent encore à démêler.
<table>
<thead>
<tr>
<th>Facteur de variabilité du microbiome testé</th>
<th>Nombre d’espèces total</th>
<th>Proportion des études ayant trouvé un effet significatif</th>
<th>Détails</th>
</tr>
</thead>
<tbody>
<tr>
<td>Espèce</td>
<td>79</td>
<td>13/13</td>
<td>-</td>
</tr>
<tr>
<td>Position dans la phylogénie des hôtes</td>
<td>12</td>
<td>1/1</td>
<td>Sur des poissons coralliens; effet faible par rapport à celui du régime alimentaire</td>
</tr>
<tr>
<td>Antibiotiques</td>
<td>4</td>
<td>4/4</td>
<td>Réduction de la diversité du microbiome, modifications de sa composition</td>
</tr>
<tr>
<td>Pré- ou probiotiques</td>
<td>3</td>
<td>3/3</td>
<td>Modification de la composition du microbiome</td>
</tr>
<tr>
<td>Régime alimentaire ou changement d’alimentation</td>
<td>35</td>
<td>15/15</td>
<td>Effet d’un changement d’alimentation dépendant de l’espèce; les différences de régimes alimentaires expliquent les différences spécifiques</td>
</tr>
<tr>
<td>Géographie (lieu de capture)</td>
<td>3</td>
<td>1/3</td>
<td>Différence de composition du microbiome entre poissons provenant de différents cours d’eau; pas d’effet de la géographie sur le saumon Salmo salar</td>
</tr>
<tr>
<td>Salinité</td>
<td>32</td>
<td>3/3</td>
<td>Les phases de migration entraînent un shift du microbiome chez les espèces anadromes. Chez les espèces marines, il existe un fort effet de la salinité de l’eau environnant les poissons</td>
</tr>
<tr>
<td>Température de l’eau ou saisonnalité</td>
<td>14</td>
<td>3/3</td>
<td>Modification de la composition du microbiome</td>
</tr>
<tr>
<td>Caractéristiques de l’individu hôte (Génotype, Phénotype)</td>
<td>7</td>
<td>6/7</td>
<td>Différences marquées entre des poissons à croissance lente et à croissance rapide; différences entre différents écotypes d’une même espèce</td>
</tr>
<tr>
<td>Micro-habitat dans le tube digestif</td>
<td>9</td>
<td>8/8</td>
<td>Différences de composition du microbiome entre muqueuse et le contenu du tube digestif; différences entre les zones antérieures et postérieures du tube</td>
</tr>
<tr>
<td>Stade de vie de l’hôte</td>
<td>16</td>
<td>4/5</td>
<td>Successions microbiennes entre les stades larvaire et adulte, augmentation de la diversité chez l’adulte</td>
</tr>
</tbody>
</table>

**Tableau 1.2 : Synthèse de 51 études du microbiome digestif de poissons téloéostéens,** réalisées en eau douce ou en eau de mer. Le tableau complet contenant l’ensemble des facteurs, des espèces et des références est disponible en Annexe 8.1.

Tout comme chez les vertébrés terrestres, le microbiome digestif joue un rôle crucial pour la fitness des poissons marins, notamment via la protection contre les organismes pathogènes (production de molécules antimicrobiennes, modulation du système immunitaire (Lowrey et al. 2015; Gómez and Balcázar 2008)), la dégradation et l’assimilation des nutriments, et la maturation et la différenciation des cellules épithéliales de l’intestin (Nayak 2010). Chez les poissons à consommation humaine, beaucoup d’études sont focalisées sur la perte de l’équilibre du microbiome digestif suite à un stress (changement d’alimentation, antibiotiques), pouvant entraîner des pertes économiques, et au rétablissement de cet équilibre via l’administration de probiotiques (Balcázar et al. 2006).
Chez les mammifères marins, des études du microbiome fécal publiées sur 9 espèces appartenant aux ordres Carnivora (phoques, otaries) et Sirenia (lamantin, dugong) ont mis en évidence un effet de l’âge de l’animal et du régime alimentaire de l’espèce, suggérant que certains patrons retrouvés chez les mammifères terrestres pourraient être généralisables au microbiome digestif des mammifères marins (Nelson et al. 2015). Une étude focalisée sur le microbiome digestif de deux odontocètes, le cachalot pygmée (Kogia breviceps) et le cachalot nain (K. sima) a démontré que les principales bactéries présentes sont des membres du phylum Bacteroidetes, présent également dans le tube digestif d’autres vertébrés terrestres et marins (Figures 1.3 et 1.9), et qui pourraient jouer un rôle dans la dégradation de molécules complexes telles que des glycanes (Erwin et al. 2017).

Enfin, une étude de métagénomique a comparé le microbiome de différentes espèces de cétacées et a confirmé un fort effet du régime alimentaire. De plus, les mysticètes (baleines filtreuses) ont un microbiome différent de celui d’autres mammifères carnivores et herbivores (Figure 1.9). En effet celui-ci présente à la fois des activités fermentatives caractéristiques des microbiomes des mammifères herbivores, et des activités liées au catabolisme des protéines, plus proches de celles des microbiomes des mammifères carnivores (Sanders et al. 2015).
<table>
<thead>
<tr>
<th>Famille/ordre</th>
<th>Microbiome digestif</th>
<th>Microbiome cutané</th>
</tr>
</thead>
<tbody>
<tr>
<td>Poissons téléostéens</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Perciformes</td>
<td><img src="image1" alt="Perciformes" /></td>
<td><img src="image2" alt="Perciformes" /></td>
</tr>
<tr>
<td>Mugilliformes</td>
<td><img src="image3" alt="Mugilliformes" /></td>
<td><img src="image4" alt="Mugilliformes" /></td>
</tr>
<tr>
<td>Cyprinodontiformes</td>
<td><img src="image5" alt="Cyprinodontiformes" /></td>
<td><img src="image6" alt="Cyprinodontiformes" /></td>
</tr>
<tr>
<td>Pleuronectiformes</td>
<td><img src="image7" alt="Pleuronectiformes" /></td>
<td><img src="image8" alt="Pleuronectiformes" /></td>
</tr>
<tr>
<td>Mammifères</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dugongidae</td>
<td><img src="image9" alt="Dugongidae" /></td>
<td><img src="image10" alt="Dugongidae" /></td>
</tr>
<tr>
<td>Phocidae</td>
<td><img src="image11" alt="Phocidae" /></td>
<td><img src="image12" alt="Phocidae" /></td>
</tr>
<tr>
<td>Otaridae</td>
<td><img src="image13" alt="Otaridae" /></td>
<td><img src="image14" alt="Otaridae" /></td>
</tr>
<tr>
<td>Mysticeti</td>
<td><img src="image15" alt="Mysticeti" /></td>
<td><img src="image16" alt="Mysticeti" /></td>
</tr>
<tr>
<td>Odontici</td>
<td><img src="image17" alt="Odontici" /></td>
<td><img src="image18" alt="Odontici" /></td>
</tr>
<tr>
<td>Reptiles</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Chelonioidae</td>
<td><img src="image19" alt="Chelonioidae" /></td>
<td><img src="image20" alt="Chelonioidae" /></td>
</tr>
<tr>
<td>Requins</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Chondrichthyes</td>
<td><img src="image21" alt="Chondrichthyes" /></td>
<td><img src="image22" alt="Chondrichthyes" /></td>
</tr>
</tbody>
</table>

Chez les autres groupes de vertébrés marins, tout reste encore à faire. L’inventaire de la bibliographie ne m’a permis de trouver qu’une seule étude inventoriant le microbiome fécal et digestif de 8 individus de la même espèce de tortue de mer, dont la composition en termes de phylums bactériens est proche de celle du microbiome de vertébrés terrestres, à l’exception des Gammaprotéobactéries, qui sont plutôt une caractéristique des organismes marins (Figure 1.9) (Abdelrhmam et al. 2016). Enfin, une seule étude a inventorié le microbiome digestif d’individus sains de 3 espèces de requins (Givens et al. 2015), démontrant la spécificité du microbiome pour l’espèce hôte et une distinction importante avec le microbiome digestif des poissons téléostéens. D’autres études ont également démontré l’existence d’infections bactériennes chez les requins (e.g. (Janse and Kik 2012)).

1.2.2.2 Microbiome cutané

A ce jour moins de 20 études se sont intéressées au microbiome cutané des vertébrés marins (Figure 1.9). Contrairement au tube digestif, la peau est en contact direct avec le milieu marin, présentant un grand nombre et une grande diversité de bactéries (Kirchman 2010). On pourrait donc s’attendre à ce que celui-ci ait une composition proche de celles des communautés planctoniques environnantes. Les études publiées à ce jour sur divers taxons (requin, baleine à bosse, poissons téléostéens) démontrent cependant qu’au contraire ces communautés, comme celles du tube digestif, sont très différentes des communautés planctoniques (Larsen et al. 2013; Apprill et al. 2014). La plupart des poissons téléostéens sont recouverts d’un mucus synthétisé par les cellules muqueuses au niveau de l’épithélium cutané (Ángeles Esteban 2012), qui constitue la première barrière de défense contre les microorganismes pathogènes; sa texture en gel emprisonne les microorganismes, qui sont ensuite emportés avec lui lors de son renouvellement (Ángeles Esteban 2012), mais il peut également servir de ressource nutritive pour les communautés microbiennes cutanées (Bordas et al. 1998). Le microbiome cutané des poissons joue très certainement un rôle important dans cette fonction de barrière, en sécrétant un grand nombre de molécules antimicrobiennes (Hellio et al. 2002; Lowrey et al. 2015) et en modulant le fonctionnement du système immunitaire (Kelly and Salinas 2017). Une étude menée sur le poisson chat d’eau douce (Ictalurus punctatus) a d’ailleurs démontré qu’un déséquilibre de ce microbiome cutané engendrait une augmentation de la susceptibility à l’infection par Flavobacterium columnare, un pathogène externe fréquent (Mohammed and Arias 2015).

La surface cutanée des vertébrés marins présente une importante diversité : présence d’écailles ou non, sécrétion et nature du mucus cutané, présence de poils, etc. Cette diversité peut entraîner une importante variabilité du microbiome cutané, à la fois entre grands clades de macro-organismes, mais également à l’intérieur d’un clade. Une étude portant sur six espèces de poissons téléostéens (la seule sur plus de 2 espèces) prélevés dans l’Atlantique a en effet démontré qu’il existait une importante spécificité entre le microbiome
et son espèce hôte (Larsen et al. 2013). Cependant d’autres facteurs, tels que l’état physiologique de l’individu, ou encore sa localisation géographique, influencent également la composition du microbiome, et entrainent une variabilité inter-individuelle importante de celui-ci, ce qui a aussi été observé chez la baleine à bosse (Apprill et al. 2014).

Les mécanismes à l’origine de ces patrons sont encore complètement inconnus. La spécificité du microbiome pour son espèce hôte ou pour un individu donné pourrait être liée à des différences métaboliques ou immunitaires, différents états de santé, ou des différences écologiques : par exemple, les contacts avec d’autres individus, des proies, des plantes ou algues, ou des substrats abiotiques. De plus, les vertébrés marins sont des animaux mobiles ; le microbiome cutané peut donc à la fois être influencé par les masses d’eau présentes autour du poisson, mais aussi celles expérimentées auparavant. Ceci soulève la nécessité de décrire dans un premier temps la variabilité du microbiome cutané en conditions contrôlées afin de s’affranchir au mieux de cette variable environnementale, accompagnée d’une analyse fine des différentes échelles de variabilité du microbiome.

D’autre part, une étude récente ayant analysé le microbiome cutané du requin renard commun (Alopias vulpinus) a mis en évidence des patrons de variations très différents de ceux observés sur la baleine et les poissons (Doane et al. 2017), présentant une dissimilarité faible entre individus et un potentiel enzymatique particulier. Cela suggère que les patrons de variation du microbiome cutané ne sont potentiellement pas tous généralisables entre clades de macro-organismes, et démontre ainsi la nécessité d’investigations plus systématiques des différents microbiomes des vertébrés.

De nombreux pathogènes ont été isolés de la surface des vertébrés marins (Alvarez-Perez et al. 2010; Benhamed et al. 2014; Florio et al. 2016). Chez les poissons téléostéens, la peau et les branchies sont des voies préférentielles d’entrée pour de nombreux micro-organismes pathogènes (Austin and Austin 2016), ce qui souligne la nécessité d’une meilleure compréhension des interactions entre ces pathogènes et les microorganismes bénéfiques ou commensaux présents sur la peau de ces animaux.

1.2.3 Rôles des microbiomes associés aux animaux dans les écosystèmes marins

1.2.3.1 Diversité des microbiomes vs. Communautés planctoniques
Les communautés microbiennes planctoniques marines (définies ici comme les cellules microbiennes « libres », mais aussi celles associées à la neige marine) sont très diversifiées et jouent des rôles prépondérants dans le fonctionnement des océans, en tant qu’acteurs des grands cycles biogéochimiques (C, N, P). Depuis 10 ans, la baisse des coûts de séquençage a permis l’étude des communautés planctoniques à de très larges échelles spatio-temporelles. A une échelle mondiale, plus de 35 000 espèces microbiennes ont été
décrites, nombre qui atteint un plateau lors de l’augmentation de la taille d’échantillonnage (Sunagawa et al. 2015). Les mêmes chercheurs ont conclu que la diversité microbienne hébergée par les océans était plus élevée que celle hébergée par l’homme (Figure 1.10).

![Diagram](image)

**Figure 1.10 : Courbes d’accumulation des espèces ou génomes microbiens en fonction du nombre d’échantillons dans les écosystèmes marins.** A gauche, graphique tiré de Sunagawa et al. (2015), comparant le nombre de gènes assemblés à partir des échantillons issus de l’expédition Tara Océans, et ceux obtenus à partir d’échantillons contenant un nombre équivalent de séquences dans le microbiome humain (Li et al. 2014). L’asymptote inférieure obtenue à partir des échantillons humains suggère que le microbiome digestif humain est moins diversifié que celui des océans. Par analogie, on peut s’attendre à ce que la diversité du microbiome des espèces marines soit plus faible que celle des communautés marines. Cependant si l’on considère l’ensemble des microbiomes associés aux divers organismes vivant dans les océans (à droite), et en faisant l’hypothèse que chaque espèce héberge un microbiome relativement unique, l’asymptote des microbiomes cumulés devrait être plus élevée que celle de l’eau de mer.

Le tube digestif et la surface des macro-organismes constituent un habitat important pour les micro-organismes : le tube digestif humain contient jusqu’à $10^{13}$ cellules.mL$^{-1}$ (Sender, Fuchs, and Milo 2016), représentant entre 1000 et 1200 espèces (Qin et al. 2010). De même, la peau humaine est colonisée jusqu’à 1,2 millions de cellules microbiennes par centimètre carré (estimation via des méthodes culturales) (Reichel, Heisig, and Kampf 2011), appartenant à environ 1 000 espèces microbiennes (Pennisi 2008). Ce microbiome est loin d’être partagé par tous les êtres vivants. Ainsi chez les mammifères, il existe une spécificité du microbiome pour l’espèce hôte, suggérant une diversité globale du microbiome des vertébrés beaucoup plus élevée (Ley et al. 2008; Muegge et al. 2011; Groussin et al. 2017).

Etant donné l’importante diversité des macro-organismes marins présentant des physiologies, des plans d’organisation, et des systèmes immunitaires très différents, on peut supposer que les microbiomes associés à chaque macro-organisme soient très différents entre eux, et de ce fait pourraient augmenter la diversité microbienne globale associée aux
macro-organismes. Par conséquent, il est possible que la biodiversité microbienne associée à tous les animaux marins constitue une part significative de la diversité globale de l’écosystème océanique (Figure 1.10). Ce niveau maximum dépendra de la dissimilarité des microbiomes, à la fois entre macro-organismes d’une même espèce, mais également entre macro-organismes d’espèces différentes (Taylor et al. 2004).

1.2.3.2 Rôle de réservoirs de la diversité marine
Les microbiomes associés aux organismes marins ont une composition très différente de celle des communautés planctoniques, et présentent des espèces microbiennes qui ne sont pas détectées dans l'eau environnante (Frias-Lopez et al. 2002). C’est notamment le cas des éponges marines, où de nombreuses études ont rapporté l’existence de clusters de séquences de gène ribosomal 16S (gène codant la sous unité 16S du ribosome procaryote) dans les éponges, qui ne sont pas détectées dans l’eau de mer (Hentschel et al. 2002; Taylor et al. 2007; Simister et al. 2012).

Récemment, une méta-analyse portant sur des clusters de gène ribosomal 16S de microbiome d’éponge a comparé ces séquences à celles obtenues à partir d’un grand nombre d’échantillons d’eau de mer, de sédiments, de cheminées hydrothermales et de coraux obtenus dans toutes les mers du globe (Taylor et al. 2013). Les chercheurs ont retrouvé ces bactéries dans tous les types d’environnements ; démontrant que ces micro-organismes ne sont pas hébergés uniquement par les éponges. Cependant ces bactéries présentaient des abondances relatives faibles dans ces environnements. De plus, la comparaison de ces clusters avec des données de métatranscriptomique (basées sur l’ARN, reflétant l’activité des cellules) n’a permis d’identifier ces clusters « éponges-spécifiques » que dans un petit nombre d’échantillons, suggérant que la majorité de ces bactéries sont inactives dans l’eau de mer.
Ce type d’étude suggère que les macro-organismes marins pourraient fournir un habitat pour des micro-organismes appartenant à la fraction dite rare des micro-organismes vivant dans l’eau de mer, majoritairement pauvre en nutriments (à l’exception de la neige marine). Une étude récente a même suggéré que les macro-organismes mobiles pourraient participer à la dispersion de ces microorganismes, permettant d’expliquer leur grande répartition géographique (Troussellier et al. 2017). Le potentiel rôle de réservoir et de disperseur des macro-organismes pourrait ainsi permettre d’engendrer des phénomènes de rescue effect, permettant d’expliquer la persistance et la répartition des micro-organismes faiblement abondants dans l’eau de mer.

1.2.3.3 Rôles liés au maintien de la santé de l’hôte
Les microorganismes jouent un rôle très important dans la santé de leur hôte, notamment en assurant sa nutrition et sa protection contre les microorganismes pathogènes. Le microbiome associé au tube digestif des poissons herbivores joue également très
certainement un rôle essentiel dans la dégradation de composés difficilement digestibles (Clements, Raubenheimer, and Choat 2009). Nous avons vu également que les baleines filtreuses présentent un microbiome très particulier, adapté à la consommation de krill, très riche en chitine (Sanders et al. 2015). Par conséquent, le microbiome associé à ces animaux, en maximisant la quantité de nutriments extraits de leur bol alimentaire, participe indirectement au bon fonctionnement de tout le réseau trophique.

En plus de ses fonctions cruciales dans la digestion, le microbiome digestif participe à la biosynthèse d’acides aminés essentiels1, favorise le stockage de triglycérides par le foie, et produit des vitamines nécessaires à la vie de l’animal (Morowitz, Carlisle, and Alverdy 2011). Les cellules microbiennes colonisant le tube digestif et la peau des animaux sont en interaction constante entre elles, et avec les cellules immunitaires de l’hôte, exerçant ainsi des fonctions de protection contre les pathogènes, à la fois via l’éducation et la régulation du système immunitaire, mais également en éliminant directement les micro-organismes pathogènes (Nayak 2010).

Le microbiome est également probablement impliqué dans la tolérance aux perturbations environnementales. Chez les coraux, une étude récente a démontré que la présence de certains taxons bactériens présents dans le mucus était corrélée à la réponse du corail à une élévation rapide de la température, suggérant que ceux-ci pourraient être impliqués dans la tolérance du corail à la chaleur (Ziegler et al. 2017).

Dans ce contexte, le microbiome associé aux animaux marins joue potentiellement un rôle prépondérant dans le maintien et le bon fonctionnement des populations d’animaux à tous les niveaux trophiques, dans un écosystème soumis à de nombreuses pressions anthropiques, et devrait être davantage pris en compte dans les études de conservation (Bahrndorff et al. 2016).

Les animaux jouent de nombreux rôles dans les écosystèmes marins. Les coraux ont un rôle important de fixation du carbon et de production primaire, fonction réalisée en interaction avec leurs symbiontes microbiens. Les poissons marins, quant à eux, (ici au sens large, i.e. poissons téléostéens et chondrichtyens) présentent des régimes alimentaires très variés, certains étant consommateurs primaires (i.e. herbivores, détritivores) d’autres ayant des rôles de consommateurs secondaire (e.g. invertivore, le requin baleine, se nourrissant de zooplancton) ou top-prédateurs (e.g. le barracuda, le requin tigre) (Levinton and Levinton 2017). Ils contrôlent ainsi les autres compartiments de l’écosystème et régulent le cycle des nutriments en consommant des organismes des niveaux trophiques les plus bas aux plus haut. Par exemple, dans les récifs coralliens, les coraux sont consommés par des poissons papillons (Chaetodontidae), des poissons perroquets (Scaridae) et l’étoile de mer Acanthaster, permettant de réinjecter le carbone organique dans les autres maillons du réseau trophique. Ces poissons permettent également de contrôler la quantité d’algues

---
1 Chez les mammifères non ruminants, environ 20% de la lysine et de la thréonine circulant dans le plasma proviennent des activités de biosynthèse du microbiome digestif (Metges 2000; Metges and Petzke 2005).
(Lirman 2001), qui se développent sur les coraux. Hors ces algues bloquent l’accès à la lumière des coraux (McCork 1999), augmentent leur mortalité en stimulant l’activité microbienne à leur surface (Smith et al. 2006), et empêchent la re-colonisation du milieu par les larves de coraux (Kuffner et al. 2006). Indirectement, ces poissons permettent donc de favoriser la production primaire.


Par conséquent, la disparition ou même la baisse de la diversité et/ou de la biomasse des animaux marins entraînerait une perte de nombreuses fonctions écosystémiques qu’ils portent (e.g. en tant que proie, minéralisateur, prédateur, brouteur, ou responsable de bio-érosion et bioturbation).

1.3 Objectifs de la thèse
L’ensemble des travaux cités ci-dessus souligne que les organismes marins sont colonisés par une grande diversité de micro-organismes, qui sont supposés jouer des fonctions essentielles pour leur hôte. Ils mettent aussi en évidence que la diversité du microbiome cutané et les différents facteurs pouvant l’influencer n’ont été que très peu étudiés avant 2014. Ces microbiomes constituent donc un tout nouveau champs d’étude, dont les premières découvertes suggèrent qu’il pourrait contribuer de manière significative à la diversité et au fonctionnement de l’écosystème global.

L’érosion sans précédent de la biodiversité macroscopique – certains auteurs affirment que la Terre est entrée dans sa sixième extinction de masse (Ceballos et al. 2015) – pourrait avoir des conséquences importantes sur la diversité microbienne marine, compte tenu de la variabilité importante du microbiome associé aux macro-organismes. Par un effet de cascade, elle aurait des répercussions sur le fonctionnement des écosystèmes océaniques.

Dans ce contexte, ma thèse visait à répondre à 3 objectifs i) mesurer la diversité des microbiomes associés aux macro-organismes marins, ii) mettre en évidence ses facteurs de variabilité et ses déterminants, et iii) quantifier sa contribution à la biodiversité microbienne marine et sa vulnérabilité face aux changements globaux. Ces différents objectifs sont représentés par les questions en gras dans la figure 1.11, et détaillés ci-après.

Figure 1.11 : Schéma représentatif des différentes questions auxquelles j’ai tenté de répondre durant ma thèse. Chacune des questions correspond à un objectif de thèse, traité dans les manuscrits indiqués en rouge clair.

Les trois objectifs majeurs de ce travail sont synthétisés ci-dessous et développés dans leur intégralité en introduction des chapitres correspondants.
Objectif 1 (chapitres 2 et 3)
Comme nous l’avons vu précédemment, peu d’études ont été publiées sur le microbiome cutané des animaux marins. Par conséquent, il n’existait pas de méthodes reconnues et largement utilisées pour le prélèvement, et l’analyse de ces microbiomes. Le premier objectif de ma thèse a eu pour but de mettre au point une méthodologie pour inventorier le microbiome associé à la peau des animaux marins. Il se décompose en deux parties. La première a consisté à valider les méthodes de prélèvement du microbiome, de l’extraction de son ADN et des méthodes d’amplification pour les différents types d’organismes échantillonnés au cours de cette thèse. Cette étape a nécessité deux campagnes de tests, ainsi qu’un travail en laboratoire qui sont explicitées dans le chapitre 2 avant la présentation synthétique des 3 campagnes de terrain. La deuxième partie a consisté à comparer différents types d’indices de diversité destinés à mesurer la diversité et la variabilité des communautés microbiennes, et sera détaillé dans le chapitre 3 sous forme d’un manuscrit en cours de soumission.

Objectif 2 (chapitre 4)
Chez les mammifères terrestres, dont l’Homme, de nombreux facteurs exogènes et endogènes influencent la composition des microbiomes cutané et digestif, tels que l’alimentation et le système immunitaire. Chez les animaux marins, ces facteurs sont encore mal identifiés. De plus, dans la plupart des cas, l’influence relative de ces deux types de facteurs (endogènes et exogènes) n’est pas mesurée, ce qui constitue un frein à l’élaboration d’hypothèses sur les mécanismes à l’origine de l’effet de ces facteurs. L’objectif général de ce chapitre est donc de distinguer l’influence relative des facteurs exogènes et endogènes sur la composition du microbiome cutané des animaux marins. Pour cela, je me suis affranchie des facteurs environnementaux pour spécifiquement (i) mesurer le niveau de variabilité « basal » de la structure du microbiome cutané à différentes échelles (intra- et inter-individuelle, et interspécifique), (ii) déterminer les échelles de variabilité majeures du microbiome, et (iii) identifier les déterminants pouvant expliquer la variabilité de ce microbiome à ces différentes échelles. Ces objectifs ont été réalisés sur 3 types de modèles animaux : deux poissons téléostéens élevés en conditions contrôlées (dorade et loup de mer), deux mammifères marins maintenus en captivité (orque et grand dauphin), et 44 espèces de poissons colonisant le même récif corallien. Les résultats de cet objectif sont explicités dans le chapitre 4 sous la forme de 3 manuscrits scientifiques.

Objectif 3 (chapitre 5)
Il existe une importante diversité de macro-organismes dans les écosystèmes marins, dont les surfaces présentent des caractéristiques contrastées (composition de la peau, du mucus cutané, type d’écailles) pouvant engendrer une importante variabilité des microbiomes cutanés entre grands clades d’animaux. Dans le cas d’un écosystème très diversifié tel qu’un récif corallien, on pourrait donc s’attendre à ce que la diversité cumulée des microbiomes
associés aux divers animaux soit très importante. Le troisième objectif de ma thèse, détaillé dans le chapitre 5, est de quantifier la contribution de la diversité des microbiomes associés aux macro-organismes invertébrés et vertébrés à la biodiversité microbienne totale à l’échelle d’un écosystème. Enfin, les écosystèmes coralliens étant particulièrement vulnérables au changement climatique, l’objectif final était de replacer l’importance des microbiomes dans un contexte de changement global, et de mesurer leur vulnérabilité face aux pressions anthropiques.
2. Méthodes d’analyse de la diversité du microbiome cutané des animaux marins

Le héros puise de l’eau et y lave ses mains victorieuses, et pour que le gravier ne blesse pas la tête aux cheveux de serpents, il amollit le sol avec un lit de feuilles, y étend des tiges nées sous la mer et y dépose la tête de Méduse, fille de Phorcys. Une tige récemment coupée et encore vivante, imbibée de sève, capta la puissance du monstre, se durcit à son contact et sentit dans ses rameaux et son feuillage une rigidité nouvelle. Alors les nymphes de la mer tentent de reproduire ce miracle sur de nombreux rameaux, et ont la joie d’obtenir le même effet, quand elles jettent dans l’eau des semences venant de ces tiges. Maintenant encore les coraux présentent la même propriété : ils n’acquièrent leur dureté qu’au contact de l’air, et leur tige souple dans la mer devient de la pierre quand elle en sort.

Ovide, Métamorphoses, livre IV (753-803), 1er siècle

2.1 Historique des méthodes utilisées pour l’étude du microbiome

2.1.1 Méthodes culturales

Les premières études explorant la diversité des communautés microobiennes ont reposé sur des méthodes culturales. Elles consistaient à prélever un échantillon, puis à réaliser des étalements sur des milieux gélosés. Ensuite, des colonies étaient isolées par des repiquages successifs puis étaient identifiées en fonction de leurs propriétés enzymatiques, métaboliques et morphologiques. Il est désormais acquis qu’en milieu marin, ces méthodes ne permettent de cultiver, et donc d’accéder qu’à une très faible fraction de la diversité bactérienne, souvent <1% de l’abondance totale (Ferguson, Buckley, and Palumbo 1984).

Récemment, de nombreux efforts ont été realisés pour la mise en culture de souches microobiennes intestinales humaines, notamment en combinant des méthodes métagénomiques (voir ci-après) avec de multiples techniques de mise en culture (Lagier et al. 2015; Browne et al. 2016). Ces méthodes ont permis de séquencer jusqu’à 39% de la diversité bactérienne intestinale publiée par le Human Microbiome Consortium (The Human Microbiome Project Consortium 2012) et ont mis en évidence des espèces microobiennes présentes en trop faible abondance (<10^6 cellules par gramme de fèces) pour être détectées par les méthodes classiques de séquençage haut débit (Lagier et al. 2012). Dans le cas du microbiome cutané humain, les premières études basées sur des méthodes culturales décrivaient moins d’une vingtaine de micro-organismes (Marples 1965). Aujourd’hui, grâce aux méthodes moléculaires, on considère que le microbiome cutané humain comporte environ 1000 espèces microobiennes (Sender, Fuchs, and Milo 2016).
2.1.2 Méthodes moléculaires basées sur un gène marqueur

2.1.2.1 Méthodes d’empreinte moléculaire
Les méthodes d’empreinte moléculaire fournissent une information de base sur la structure des communautés microbiennes. Elles consistent à séparer les fragments d’ADN par migration sur un gel ou dans un capillaire en fonction de leur taille et/ou de leur composition en bases nucléotidiques (Kuczynski et al. 2012). Certaines de ces méthodes (e.g. la DGGE) permettent également d’identifier les espèces correspondant aux différents fragments par excision des bandes du gel et séquençage (Huybens, Mainil, and Marlier 2009). Le nombre de bandes, ou OTUs, varie de 20 à 50 selon les échantillons. Il s’agit donc d’une représentation partielle de la diversité, et qui correspond en général aux OTUs les plus abondantes de la communauté étudiée (Lalande, Villemur, and Deschênes 2013). Cependant, les patrons de distribution de la diversité bactérienne déterminés avec ces méthodes d’empreinte moléculaire sont cohérents avec ceux obtenus par séquençage haut débit (Gobet, Boetius, and Ramette 2013).

2.1.2.2 Clonage et séquençage d’un gène marqueur
Une autre approche consiste à séparer les différents fragments d’ADN par insertion dans un vecteur de clonage. Les constructions obtenues sont ensuite introduites dans des cellules compétentes cultivables, et les clones recombinants sont sélectionnés sur un milieu adapté. Le fragment d’intérêt est amplifié par PCR à partir de colonies isolées, puis séquencé afin d’identifier les espèces présentes dans l’échantillon de départ (DeSantis et al. 2007). Cette méthode, longue, et nécessitant de nombreuses étapes de manipulation, a été peu utilisée et a été assez vite remplacée par des méthodes de séquençage haut débit.

2.1.2.3 Séquençage haut débit
Les méthodes précédentes sont aujourd’hui très peu utilisées, à l’exception des méthodes culturales, qui restent indispensables pour certaines applications. Les méthodes d’empreinte moléculaire ont en effet été remplacées par de nouvelles méthodes de séquençage à haut débit, appelées collectivement Next Generation Sequencing (NGS) permettant de séquerclonage simultanément des milliers, voire des millions de gènes en un seul run (Buermans and den Dunnen 2014). Cette démocratisation des NGS s’explique par une diminution du coût d’analyse, une facilité à l’accès des serveurs et à des outils bioinformatiques simplifiés (e.g. GALAXY) et leur fort rendement en termes d’information génétique (jusqu’à 600 Giga bases
par run en HiSeq® Illumina). Aujourd’hui, les NGS sont d’autant plus accessibles qu’une partie significative du processus est externalisable vers des sociétés extérieures (de l’extraction de l’ADN à la table d’OTUs). Ces méthodes consistent à séquencer directement une courte région du gène codant l’ARNr 16S, en général une à trois régions hypervariables, en fonction du séquenceur utilisé. Elles sont actuellement utilisées dans la quasi-totalité des études impliquant l’analyse de la diversité taxonomique du microbiome, et les technologies évoluent très rapidement. Elles présentent cependant des limitations. En effet, tout comme les méthodes précédentes, elles nécessitent une amplification par PCR du gène étudié, entraînant des biais dans l’estimation des abundances microbiennes (Pinto and Raskin 2012). De plus, comme elles ne permettent de séquencer qu’un petit fragment d’ADN (jusqu’à 600 pb dans le cas du séquençage MiSeq de Illumina) (Buermans and den Dunnen 2014), elles ont souvent une résolution moins importante qu’avec un séquençage classique (séquençage de Sanger) de la totalité du gène ADNr 16S ou du génome bactérien, car elles ne permettent pas de différencier des espèces proches. De plus, la qualité de l’identification sera souvent dépendante de la région hypervariable utilisée (Petrosino et al. 2009). Enfin, les techniques de séquençage haut débit font des erreurs, dont on peut limiter le nombre en utilisant des méthodes de séquençage séquençant les deux brins d’un même fragment d’ADN (e.g. Illumina MiSeq en 2*250 nucléotides) et en utilisant des logiciels de nettoyage de séquences performants (Kozich et al. 2013).

Les progrès en bioinformatique qui ont suivi de près les développements de ces méthodes ont permis la mise en place d’outils de routine performants, permettant de stocker et d’analyser les énormes masses de données (e.g. outils Mothur, Qiime), de corriger et d’éliminer les erreurs commises par les séquenceurs, et d’améliorer nettement leur résolution malgré la petite taille du fragment d’ADN séquencé. Pour toutes ces raisons, ce sont les méthodes que j’ai le plus utilisé tout au long de ma thèse.

2.1.3 Approches métagénomiques
La métagénomique regroupe les approches visant à une analyse complète des génomes présents dans un échantillon (métagénome). Ceci est réalisé par séquençage dit Shotgun, consistant à séquencer aléatoirement l’ensemble de l’ADN extrait. Les études métagénomiques permettent d’estimer les fonctions portées par les gènes du microbiome, et donc d’accéder à la diversité fonctionnelle du microbiome, ce que les méthodes NGS ne permettent pas de faire (Abubucker et al. 2012). Elles restent cependant aujourd’hui encore très chères (environ 600€ par échantillon) et nécessitent un travail de nettoyage et d’alignement des séquences très important (Roumpeka et al. 2017).
2.2 Méthodes utilisées lors de la thèse

Dans le chapitre 1 j’ai évoqué le manque d’investigations systématiques du microbiome dans les différents grands clades des animaux marins ainsi que les limites de la généralisation des mécanismes d’assemblage et de transfert du microbiome entre les différents types d’organismes, et entre les microbiomes digestifs et cutanés. Il nous paraissait donc nécessaire d’étudier à la fois les poissons téloéostéens, qui constituent la biomasse macroscopique la plus importante dans les océans, mais également les grands clades d’invertébrés notamment récifaux et les mammifères marins.

2.2.1 Campagnes de prélèvement

J’ai réalisé 5 campagnes de prélèvement au cours de ma thèse, dont 2 campagnes tests ayant permis de valider les protocoles de prélèvement du microbiome et d’extraction de l’ADN. Un résumé de l’ensemble des échantillons prélevés, des espèces ciblées et des méthodes utilisées est disponible dans le tableau 2.1.

<table>
<thead>
<tr>
<th>Campagne</th>
<th>Lieu</th>
<th>Année</th>
<th>Nb d’échantillons</th>
<th>Nb d’espèces</th>
<th>Méthodes utilisées</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 : Poissons</td>
<td>SMEL, Sète</td>
<td>2014</td>
<td>22</td>
<td>2</td>
<td>P : Découpe nageoire</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>E : Extraction indirecte + kit Qiagen</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>A : pyrosequençage 454</td>
</tr>
<tr>
<td>2 : Roussettes</td>
<td>SMEL, Sète</td>
<td>2014</td>
<td>11</td>
<td>1</td>
<td>P : mousse DNA-free (Whatman)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>E : kits Blood &amp; Tissue® (Qiagen) + PowerSoil® (MoBio)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>A : Dosage de l’ADN</td>
</tr>
<tr>
<td>3 : Cétacés</td>
<td>Marineland, Antíbes</td>
<td>2014</td>
<td>43</td>
<td>2</td>
<td>P : mousse DNA-free (Whatman)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>E : kit Qiagen</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>A : MiSeq Illumina</td>
</tr>
<tr>
<td>4 : Poissons, invertébrés</td>
<td>Etang de Thau</td>
<td>2015</td>
<td>91</td>
<td>≥ 11</td>
<td>P : Coton-tige SK-2S (Isolex), 2 types de préservation</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>E : kits PowerSoil® et Maxwell® (Promega)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>A : Dosage de l’ADN</td>
</tr>
<tr>
<td>5 : Poissons, invertébrés</td>
<td>Mayotte</td>
<td>2015</td>
<td>314</td>
<td>≥ 83</td>
<td>P : Coton-tiges Isolex</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>E : kit Maxwell®</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>A : MiSeq Illumina</td>
</tr>
</tbody>
</table>

Tableau 2.1 : Récapitulatif des 5 campagnes de prélèvement effectuées dans le cadre de ma thèse, précisant l’année de prélèvement, le nombre d’échantillons ainsi que le nombre minimal d’espèces prélevés, et les méthodes utilisées. P : méthode(s) de prélèvement du microbiome, E : méthode(s) d’extraction de l’ADN, A : méthode(s) d’analyse. Les campagnes de test ayant permis de valider les méthodes de prélèvement et d’extraction du microbiome sont indiquées sur fond gris.

2.2.2 Campagne 1

L’objectif de la campagne 1 était de mesurer simultanément les variabilités intra- et inter-individuelle, et interspécifique du microbiome cutané de deux espèces de poissons téloéostéens (Sparus aurata et Dicentrarchus labrax) élevées en conditions identiques et contrôlées (en bassin aquacole). Ces deux espèces ont été prélevées à la Station
Méditerranéenne de L’Environnement Littoral (SMEL, Sète). Cette campagne a été réalisée durant mon master 2, mais l’analyse approfondie des données et la rédaction du manuscrit à partir de ces échantillons a été réalisée durant ma première année de thèse, et a fait l’objet de la rédaction du manuscrit B.

2.2.2.1 Prélèvement
Le microbiome cutané a été étudié chez 4 individus de chacune des deux espèces, et sur 4 zones du corps pour chacun des individus. Afin de ne pas contaminer sa surface, chaque poisson a été prélevé à la pêche à la ligne dans le bassin, puis immédiatement mis à mort selon les législations en vigueur. L’anesthésie, comme méthode alternative à l’abattage, a été considérée afin d’immobiliser temporairement le poisson durant le prélèvement du microbiome. Cependant l’effet des principaux anesthésiques pour poisson (Isoeugénol, Benzocaïne, et le MS222 qui sont des antimicrobiens potentiels) sur le microbiome cutané est inconnu. De plus, le mucus cutané joue un rôle primordial pour la survie du poisson (Shephard 1994). Un grattage du mucus cutané aurait donc pu être délétère pour le poisson. Par conséquent la mise à mort rapide immédiatement après capture a été préférée dans l’ensemble des manipulations impliquant des poissons téléostéens.

La variabilité intra-individuelle du poisson a été analysée sur les zones du corps les plus contrastées possibles. Afin de comparer des surfaces fonctionnellement équivalentes, nous avons prélevé l’entièreté de 4 nageoires d’intérêt : 2 nageoires proches des zones d’excrétion (nageoire anale et nageoire pectorale, proche des branchies) et 2 nageoires potentiellement pauvres en nutriments et soumises à différents régimes hydrodynamiques (la nageoire caudale et la nageoire dorsale) (Figure 2.1).

Figure 2.1 : Différentes nageoires prélevées sur la dorade royale (*Sparus aurata*) et le loup (*Dicentrarchus labrax*) lors de la campagne 1.

2.2.2.2 Extraction de l’ADN et analyses
L’extraction de l’ADN à partir des nageoires a été réalisée à partir d’un protocole adapté d’Amalfitano et Fazi (2008) et validée par une DGGE en amont du séquençage, fourni en Annexe 8.2. La méthode d’analyse choisie par la suite a été le pyroséquençage 454, principalement pour des raisons de coût au moment de l’analyse. L’amplification de la région hypervariable V1-V3 du gène codant l’ARN ribosomique 16S a été réalisée à l’aide des
amorces bactériennes 27F et 519R (voir tableau 2.2) par un laboratoire extérieur, qui a également réalisé le séquençage.

<table>
<thead>
<tr>
<th>Campagne</th>
<th>Amorce Sens</th>
<th>Amorce Antisens</th>
<th>Références</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>27F 5’-AGR GTT TGA TCM TGG CTC AG-3’</td>
<td>519R 5’-GTN TTA CNG CGG CKG CTG-3’</td>
<td>1 et 2</td>
</tr>
<tr>
<td>3</td>
<td>341F 5’-ACG GRA GGC AGC AG-3’</td>
<td>784R 5’-TAC CAG GGT ATC TAA TCC T-3’</td>
<td>3 et 4</td>
</tr>
<tr>
<td>5</td>
<td>515F 5’-GTG CCA GCM GCC GCG GTA A-3’</td>
<td>806R 5’- GGA CTA CNV GGG TWT GTA AT-3’</td>
<td>5 et 6</td>
</tr>
</tbody>
</table>


2.2.3 Campagne 2
L’objectif de la campagne 2 était de tester et de valider les méthodes de prélèvement du microbiome cutané afin de les appliquer lors de la campagne 3 réalisée sur des cétacées. Les cétacées présentent une structure de peau particulièr, faite de microsillons de faible dimension (0.1-1.2 µm), associée à la sécrétion d’un gel particulier contenant des molécules hydrolysat les molécules bactériennes permettant l’attachement des bactéries (Scardino and de Nys 2011), limitant le micro- et le macrofouling. Pour ces raisons, nous nous attendions à des faibles densités microbiennes sur ces animaux. Il était donc nécessaire de tester nos méthodes d’extraction de l’ADN sur des organismes également faiblement colonisés par les micro-organismes ; nous avons choisi un requin, la rousselette, qui était élevée à la SMEL.

2.2.3.1 Prélèvement
Ces animaux servant à d’autres manipulations, ils ont été brièvement anesthésiés à l’aide de 2-phénoxyéthanol dosé à 300 ppm. Le microbiome a été prélevé sur le dos et sur la queue de trois mâles du même âge, élevés dans des conditions identiques. Après un rinçage bref de la peau de l’animal au NaCl 9% stérile, le microbiome a été prélevé à l’aide d’un applicateur DNA-free à pointe mousse (Whatman), par frottement doux dans le sens des écailles sur un cercle de 2 cm de diamètre durant 30 secondes sur chaque côté de la mousse. 3 échantillons ont été prélevés sur chacun des animaux, au niveau des parties antérieure, médiane, et postérieure du dos. Afin de pouvoir comparer l’efficacité d’extraction du microbiome de l’animal avec celle d’autres types de communautés, 2 échantillons de biofilm ont été prélevés sur les sucres permettant la distribution de l’eau dans le bassin.
**Extraction d’ADN et analyses**

Deux kits d’extraction d’ADN ont été testés, le kit Blood & Tissue® (Qiagen, Pays-Bas) et le kit PowerSoil® (MoBio, EU), à l’aide des protocoles fournis en Annexe 8.2. La quantité et la qualité de l’ADN extrait a été mesurée par spectrophotométrie (NanoDrop 1000, Thermo Fisher Scientific, EU) et fluorimétrie (Qbit, Invitrogen, EU) et sont présentées dans le tableau 2.3.

<table>
<thead>
<tr>
<th>Type d’échantillon</th>
<th>Kit</th>
<th>Concentration Nanodrop (ng.µL⁻¹)</th>
<th>A260/A280</th>
<th>A260/A230</th>
<th>Concentration Qbit® (ng.µL⁻¹)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Roussette</td>
<td>Powersoil®</td>
<td>2.13</td>
<td>2.19</td>
<td>0.40</td>
<td>&lt; 1</td>
</tr>
<tr>
<td>Sucre</td>
<td>Powersoil®</td>
<td>2.38</td>
<td>9.38</td>
<td>1.17</td>
<td>-</td>
</tr>
<tr>
<td>Roussette</td>
<td>Blood &amp; Tissue</td>
<td>12.99</td>
<td>1.44</td>
<td>0.58</td>
<td>&lt; 1</td>
</tr>
<tr>
<td>Sucre</td>
<td>Blood &amp; Tissue</td>
<td>2.23</td>
<td>1.00</td>
<td>0.67</td>
<td>-</td>
</tr>
</tbody>
</table>

**Tableau 2.3** : Concentrations d’ADN obtenues lors de la campagne n°2. Après vérification des concentrations d’ADN obtenues sur les roussettes à l’aide d’un dosage Qbit, il s’est avéré que le plus fort rendement obtenu sur les roussette à l’aide du kit Qiagen étaient vraisemblablement liées à la présence de contaminants.

Les deux kits ont présenté des rendements équivalents et faibles. Pour des raisons de coût c’est le kit Qiagen qui a été choisi pour l’extraction de l’ADN des microbiomes cutanés des cétacés lors de la campagne 3.

**2.2.4 Campagne 3**

L’objectif de la campagne 3 était de mesurer simultanément les variabilités intra- et inter-individuelle, et interspécifique du microbiome cutané de deux espèces de cétacés, l’orque (*Orcinus Orca*) et le grand dauphin (*Tursiops truncatus*) en conditions identiques. Ces deux espèces sont maintenues en captivité dans deux bassins séparés dans le parc de loisirs Marineland® d’Antibes. Cette campagne a été réalisée durant ma première année de thèse en collaboration avec Marineland® et la Fondation Marineland, et a fait l’objet de la rédaction du manuscrit C.

**2.2.4.1 Prélèvement**

Le prélèvement a été réalisé en une journée sur 4 individus orques et 4 individus dauphins, grâce à l’aide de leurs soigneurs. Afin d’étudier la variabilité intra-individuelle du microbiome cutané, 4 zones du corps ont été ciblées, les plus distantes possibles sur le corps de l’animal, tout en évitant la tête, fréquemment touchée par les soigneurs : la nageoire dorsale, la partie supérieure des nageoires pectorale et caudale, ainsi que la zone anale (Figure 2.2).
Les soigneurs ont demandé à chacun des individus de lever hors de l’eau une partie de leur corps, comme ils le font régulièrement lors des examens vétérinaires. Ensuite, un des soigneurs a réalisé le prélèvement du microbiome en frottant une surface circulaire calibrée de 7 cm de diamètre durant deux fois 30 secondes de chaque côté d’une mousse DNA-free (Whatman) (Figure 2.3). Dans le même temps, 6 échantillons de 200 mL d’eau ont été prélevés dans chacun des bassins et filtrés sur une membrane 47 mm de diamètre 0.2 μm. Les échantillons ont été placés dans des cryotubes, et congelés immédiatement à -196°C dans l’azote liquide, puis stockés à -80°C avant analyse.

2.2.4.2 Extraction d’ADN et analyses

L’extraction de l’ADN à partir des mouches a été réalisée à l’aide du protocole basé sur le kit Qiagen (Annexe 8.2). La méthode d’analyse choisie par la suite a été le séquençage MiSeq Illumina des régions hypervariables V3-V4 du gène codant l’ARN ribosomique 16S. L’amplification a été réalisée à l’aide des amorces bactériennes 341F et 784R modifiées pour le séquençage Illumina (Tableau 2.2), et le séquençage a été réalisé par un laboratoire...

2.2.5 Campagne 4
L’objectif de la campagne 4 était de tester et de valider les méthodes de prélèvement et d’extraction du microbiome de surface des vertébrés et des invertébrés marins, afin de les appliquer lors de la campagne 5 qui a été réalisée dans le récif de Mayotte. Pour cela nous avons choisi un écosystème côtier facilement accessible comportant une diversité macroscopique suffisante: l’étang de Thau.

2.2.5.1 Prélèvement
91 échantillons ont été prélevés provenant de 4 espèces de poissons et 6 types d’invertébrés (ascidie, éponges, oursins, anémones, aplysies, étoile de mer), mais également des échantillons d’eau et de biofilms épilithiques. Les poissons ont été capturés à la pêche à la ligne ou à l’aide d’une nasse appâtée. Les invertébrés ont été prélevés à l’épuisette ou à la main. Tous les échantillons ont été manipulés avec des gants. Le prélèvement des microbiomes de surface a été réalisé à l’aide des coton tiges Isohelix SK-2S (Cell Projects LTD, RU) conçus pour des prélèvements buccaux. Le prélèvement des communautés planctoniques a été réalisé par filtration de 100mL d’eau de surface à travers une membrane de 0.2 μm de largeur de pore.
Deux méthodes de préservation des échantillons ont été testées. La première méthode était la congélation immédiate dans l’azote liquide. La deuxième méthode de préservation consistait à immerger complètement l’échantillon dans du Rnalater® (Thermo Fisher Scientific, EU) avant le transport dans une glacière contenant des pains de glace.

2.2.5.2 Extraction d’ADN et analyses
Deux kits d’extraction d’ADN ont été testés, le kit PowerSoil (MoBio), et le kit Maxwell® Buccal Swab LEV DNA (Promega, EU) à l’aide du protocole fourni en annexe 8.2. La quantité et la qualité de l’ADN extrait a été mesurée par spectrophotométrie (NanoDrop 1000, Thermo Fisher Scientific, EU) et est présentée dans le tableau 2.4.
### Tableau 2.4 : Concentrations d’ADN obtenues sur les échantillons prélevés lors de la campagne 4, à l’aide de deux kits d’extraction différents et deux méthodes de préservation des échantillons.

<table>
<thead>
<tr>
<th>Kit d’extraction</th>
<th>Mode de préervation</th>
<th>Concentration (ng/µL)</th>
<th>A260/A280</th>
<th>A260/A230</th>
<th>Nombre d’échantillons</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kit MoBio</td>
<td>Azote liquide</td>
<td>23,48 ± 17,6</td>
<td>2.3 ± 0.7</td>
<td>1.3 ± 0.5</td>
<td>12</td>
</tr>
<tr>
<td></td>
<td>RNA Later</td>
<td>17,45 ± 13,1</td>
<td>1.9 ± 0.5</td>
<td>0.5 ± 0.1</td>
<td>10</td>
</tr>
<tr>
<td>Kit Maxwell</td>
<td>Azote liquide</td>
<td>74,0 ± 49,5</td>
<td>1.9 ± 0.2</td>
<td>1.9 ± 0.3</td>
<td>7</td>
</tr>
<tr>
<td></td>
<td>RNA Later</td>
<td>88,6 ± 103</td>
<td>1.7 ± 0.2</td>
<td>1.6 ± 0.4</td>
<td>8</td>
</tr>
</tbody>
</table>

Le kit Maxwell® présentait un excellent rendement comparativement au kit PowerSoil®. Le kit PowerSoil®, on l’a vu précédemment, présentait une efficacité équivalente au kit Blood & Tissue® (Qiagen) lorsqu’il était testé sur des échantillons de microbiome cutané de roussette. Par conséquent je n’ai pas testé ce dernier kit sur ces échantillons. J’ai directement choisi le kit Maxwell® pour l’extraction de l’ADN des échantillons prélevés lors de la campagne 5.

En revanche, la préservation de l’échantillon à l’aide de RNA Later® n’a pas significativement amélioré le rendement ou la qualité de l’ADN, voire a augmenté la quantité de contaminants dans l’ADN extrait, puisque le ratio A260/A230 était en moyenne 2.6 et 1.2 fois inférieur, respectivement pour les kits PowerSoil® et Maxwell®, lorsque l’échantillon était conservé dans du RNA Later®. Par conséquent, et également pour des questions de coûts et de gestion à bord du bateau, nous avons choisi de préserver les échantillons au froid sans l’ajout de conservateurs.

#### 2.2.6 Campagne 5

L’objectif de la campagne 5 était de mesurer la contribution des microbiomes de surface des divers macro-organismes présents à la diversité microbienne globale à l’échelle d’un écosystème corallien. Nous avons choisi le lagon de l’île de Mayotte dans l’océan indien (Figure 2.4). Deux types d’habitats ont été étudiés, situés au sud-ouest de l’île et loin des zones urbanisées, et présentant une belle couverture corallienne (Figure 2.5) : le récif frangeant le long de la côte de l’île, et la pente interne du récif barrière bordant le lagon.
**Figure 2.4 : Zones d’échantillonnage de la campagne 5.** L’île de Mayotte est au sud de l’archipel des Comores, dans le canal du Mozambique. Le récif barrière échantillonné (1) se situe à 15 km du récif frangeant (2).

**Figure 2.5 : Sites d’échantillonnage.** Les deux sites (Barrière à gauche, et frangeant, à droite) présentaient une importante couverture corallienne et une grande abondance et diversité de poissons téléostéens. Photos prises par Sébastien Villéger.
L’échantillonnage a été réalisé du 17 au 27 novembre 2015, et visait à collecter la plus grande diversité macroscopique possible retrouvée dans le lagon. Au total 138 poissons ont été échantillonnés, appartenant à 44 espèces et 22 familles typiques des écosystèmes coralliens (e.g. *Chaetodontidae*, *Acanthuridae*, *Scaridae*, *Labridae*). 82 colonies d’anthozoaïres ont été échantillonnées, appartenant à 13 genres de coraux durs, 3 genres de coraux mous et une anémone. 43 autres invertébrés (crustacés, échinodermes, mollusques et éponges) ont également été prélevés. Le tableau récapitulatif de l’ensemble des prélèvements de macro-organismes réalisés lors de cette campagne est fourni en Annexe 8.6.

2.2.6.1 Prélèvement
Afin de ne pas contaminer la surface des poissons, ceux-ci ont été capturés à l’aide d’un harpon ou à la pêche à la ligne. Après la mise à mort de l’animal, celui-ci a été couché sur un côté et son microbiome cutané a été prélevé par frottois d’un coton-tige Isohelix sur la totalité de la surface cutanée opposée, à l’exception de la tête, par laquelle les pêcheurs ont maintenu le poisson pour l’amener à bord.
Le microbiome de surface des anthozoaïres a été prélevé par exposition d’un spécimen de 1 à 10 cm en dehors de l’eau, provoquant la production de mucus, contenant le microbiome de surface (Leruste, Bouvier, et Bettarel 2012). Dans les cas des anthozoaïres ne produisant pas de mucus, le mucus de surface a été prélevé par frottois du spécimen entier durant 30 secondes à l’aide d’un coton tige Isohelix. Les autres invertébrés (notamment concombre de mer et oursins) ont été prélevés à l’aide de pinces ou avec des gants, et l’entièreté de leur surface a été frottée à l’aide d’un coton-tige. Dans le même temps, 35 échantillons de 200 mL d’eau ont été collectés à 50 cm de la surface (17 échantillons), et à 50 centimètres au dessus du fond (18 échantillons) et ont été filtrés sur une membrane polycarbonate de porosité de 0,2 μm.

Les échantillons prélevés au cours d’une journée d’échantillonnage ont été stockés dans une glacière électrique maintenue à -5°C tout au long de la journée, puis rapatriés le soir au laboratoire et stockés à -80°C. A la fin de la campagne d’échantillonnage, les échantillons ont été rapatriés par avion dans la glacière électrique refroidie à -20°C et remplie de pains de glace.

2.2.6.2 Extraction d’ADN et analyses
L’ADN contenu dans les différents types d’échantillons a été extrait à l’aide du kit Maxwell® *Buccal Swab LEV DNA* selon le protocole fourni en Annexe 8.2 et utilisé lors de la campagne 4. La méthode d’analyse choisie par la suite a été le séquençage MiSeq Illumina des régions hypervariables V3-V4 du gène codant l’ARN ribosomique 16S. L’amplification a été réalisée à l’aide des amorces bactériennes 515F et 806R modifiées pour le séquençage Illumina (Tableau 2.2), permettant de cibler l’ensemble de la diversité procaryotique (bactéries et
3. Réflexion méthodologique : choix d’indices de diversité appliqué au monde microbien

Distance is to microbial ecology as energy is to physics!

Ruth E. Ley, Rob Knight et Jeffrey I. Gordon, 2007

3.1 Facettes et composantes de la biodiversité

Une étape indispensable à l’étude de la diversité des communautés microbiennes est le choix d’indices de diversité adéquats pour répondre à la problématique de recherche. Il existe de nombreux indices de diversité, décrivant chacun une ou plusieurs facettes de la biodiversité (taxonomique, phylogénétique, fonctionnelle), répertoriés dans le tableau 3.1. La mesure de la biodiversité la plus simple, et la plus utilisée, est le nombre d’espèces présentes, la plupart du temps des OTUs dans les études d’écologie microbienne (Fuhrman 2009). Cette mesure est la richesse taxonomique de la communauté. Dans le cas des communautés microbiennes, la capacité d’échantillonnage est limitée soit par la profondeur de séquençage, soit par le nombre de clones utilisés dans les méthodes de clonage-séquençage (voir chapitre 2). Des versions corrigées de cet indice de diversité sont disponibles. Elles sont basées sur l’extrapolation des courbes de raréfaction (Haegeman et al. 2013), ou sur des modèles mathématiques de distribution des abondances bactériennes (Hill et al. 2003). Dans les communautés de macro- et de micro-organismes, on considère généralement que toutes les espèces (ou les OTUs) n’ont pas la même importance pour le fonctionnement des écosystèmes. Les espèces les plus abondantes auraient un impact majeur sur les fonctions globales exercées par la communauté (Fuhrman 2009) ; mais (Campbell et al. 2011) et (Lynch and Neufeld 2015) nuancent ce postulat général, démontrant qu’une partie des OTUs microbiennes les moins abondantes présentent une activité importante, et ont des fonctions uniques dans l’écosystème (e.g. nitrification). En fonction de la question posée, on peut donc pondérer la richesse par les abondances relatives des espèces ; on mesure alors la diversité taxonomique de la communauté, qui est dépendante à la fois du nombre d’espèces présentes et de l’équité des leurs abondances. Plusieurs indices existent pour mesurer cette diversité, dont les indices de Shannon (Shannon 1948) et de Gini-Simpson (Simpson 1949), qui sont les plus utilisés.

La mesure de la facette taxonomique (richesse ou diversité) de la biodiversité n’est pas suffisante, car elle considère toutes les espèces ou toutes les OTUs microbiennes de manière équivalente indépendamment de leur rôle fonctionnel ou de leur histoire évolutive (Lozupone et Knight 2005). En conséquence une communauté présentant un grand nombre d’OTUs appartenant à un petit nombre de genres bactériens proches phylogénétiquement (ce qui peut exister dans un bioréacteur par exemple (Singleton, Richardson, et Aitken 2011)) aura une diversité taxonomique équivalente à une
communauté composée du même nombre d’OTUs présentant des abondances relatives équivalentes, mais appartenant à différents phylums microbiens (e.g. la majorité des communautés microbiennes « naturelles », y compris le microbiome associé aux animaux, voir chapitre 1). Pour palier à cette difficulté, il est nécessaire de mesurer la **facette phylogénétique** de la biodiversité, en prenant en compte la topologie et les longueurs des branches de l’arbre phylogénétique reliant les espèces. Cette facette peut également être pondérée ou non par les abondances relatives des espèces. On parle alors respectivement de **richesse phylogénétique**, mesurée par le PD de Faith qui n’est pas pondéré par les abondances (Faith 1992)), et de **diversité phylogénétique**, qui peut être mesurée par l’entropie de Allen (Allen, Kon, and Bar-Yam 2009), prenant en compte des abondances des espèces.

La troisième facette de la diversité est la **facette fonctionnelle**, renseignant sur la diversité des fonctions réalisées par les espèces dans la communauté. Celle-ci, bien que la plus explicative du fonctionnement des écosystèmes, est rarement mesurée dans le cas des communautés de micro-organismes. Elle nécessite en effet de mesurer les traits fonctionnels de toutes les espèces présentes dans la communauté étudiée, ce qui n’est possible qu’en isolant et en cultivant les espèces dans différents types de conditions (e.g. systèmes BIOLOG® ou Omnilog®). Les communautés microbiennes contiennent un très grand nombre d’espèces, dont la plupart ne sont pas cultivées ou cultivables (Barer and Harwood 1999; Connon and Giovannoni 2002), ce qui rend impossible une mesure exhaustive des traits fonctionnels liés aux espèces. De plus, pour les micro-organismes présentant une importante plasticité phénotypique, la mesure d’une activité en conditions de laboratoire par une souche type ne présage pas nécessairement l’existence d’une telle activité par des souches de la même espèce dans des conditions naturelles (Torsvik and Øvreås 2002). Pour appréhender la diversité fonctionnelle chez les communautés microbiennes, on mesure donc généralement la diversité des gènes fonctionnels contenus dans le génome des espèces présentes (diversité fonctionnelle potentielle), dans leur transcriptome (exprimée)², ou encore la diversité des protéines produites par les micro-organismes (réalisée) (Franzosa et al. 2015). Cependant, si ce type de méthode indique quelles fonctions sont présentes, elles ne répondent pas à la question « Qui fait quoi ? » et ne permettent donc pas de calculer les indices de diversité fonctionnelle utilisés chez les macro-organismes³ (e.g. (Villéger, Mason, and Mouillot 2008))

---

² Ces diversités fonctionnelles potentielles et exprimées sont mesurées par le séquençage shotgun, permettant de séquencer l’ensemble du métagénome ou du métatranscriptome (voir chapitre 2)

³ Il existe des méthodes pour savoir « qui fait quoi », nécessitant l’utilisation complémentaire de sondes espèces-spécifiques fluorescentes (FISH) et la détection d’utilisation de substrats marqués (e.g. Stable Isotope Probing, SIP), difficilement applicables à une communauté entière (Torsvik and Øvreås 2002).
Figure 3.1 : Facettes et composantes de la biodiversité. Les différents figurés correspondent à différentes espèces de micro-organismes dans une communauté. Ces espèces sont plus ou moins éloignées phylogénétiquement, et sont regroupées en deux groupes fonctionnels distincts, ici les organotrophes et les phototrophes.

Ces trois facettes (taxonomique, phylogénétique et fonctionnelle) et composantes de la biodiversité (pondérée par les abondances relatives – composante structurelle – ou non pondérée par les abondances relatives – composante compositionnelle – des espèces, des lignées phylogénétiques, ou des fonctions) peuvent être mesurées à plusieurs échelles. La diversité alpha est la diversité mesurée à l’échelle d’une communauté microbienne. La diversité gamma est mesurée à l’échelle d’une métacommunauté (i.e. ensemble de communautés microbiennes ayant des interactions via les processus de dispersion-colonisation, voir chapitre 1). Enfin, la diversité beta est la mesure de dissimilarité entre deux communautés, et est généralement mesurée par le ratio entre la diversité gamma et la moyenne des diversités alpha des deux communautés considérées (Whittaker 1965) (Figure 3.2). La diversité beta peut-être mesurée entre différentes communautés locales à un instant t (spatial), ou dans une même communauté étudiée à des moments différents (temporel).
Figure 3.2 : Différentes échelles (alpha, beta et gamma) de la diversité du microbiome mesurée chez une population de 3 poissons de la même espèce (en haut). La diversité alpha correspond à la diversité du microbiome d’un individu. La diversité beta correspond à la dissimilarité existant entre les microbiomes de deux des individus considérés, et la diversité gamma correspond à la diversité du microbiome de l’ensemble de la population considérée. En bas, illustration des deux composantes de la diversité beta phylogénétique. A gauche, les deux individus considérés hébergent 5 espèces microbiennes en commun contre 2 espèces propres au microbiome de chacun des individus. La diversité beta, ou dissimilarité, entre les deux microbiomes est alors plutôt faible. A droite, la diversité beta entre les deux individus est pondérée par les abondances des espèces microbiennes. Dans ce cas, elle est plutôt forte, car les espèces propres à chacun des deux individus, bien que peu nombreuses, présentent des effectifs plus forts que les espèces communes.

La majorité des indices cités précédemment ont récemment été regroupés dans des cadres d’étude « unifiés » permettant de mesurer différentes facettes, composantes, et échelles de la biodiversité à l’aide d’un seul formalisme mathématique incluant les caractéristiques des espèces (taxonomique, phylogénétique ou fonctionnelle) et leurs importance (identique pour les indices de composition ou dominance pour ceux de structure) (Chen et al. 2012; Veresoglou et al. 2014; Chao, Chiu, and Jost 2014).

La grande quantité d’indices disponibles reflète donc la diversité des facettes, composantes et échelles de la biodiversité. Il n’existe pas d’indices meilleurs que d’autres, car leur application dépend de la question posée. Cependant il existe des indices différents mais qui mesurent la même facette, composante et échelle de la biodiversité. C’est le cas de la généralisation de l’indice Unifrac (Chen et al. 2012), et de la version phylogénétique de la
diversité beta de Chao et al (Box1, (Chao, Chiu, and Jost 2010)) . Toutes deux mesurent des diversités beta phylogénétiques, mais leur efficacité respective n’a pas été évaluée.

Les indices de diversité sont principalement utilisés pour détecter des différences entre communautés dans le temps et/ou dans l’espace, et de regrouper des communautés semblables. Ils permettent également de détecter des structures d’assemblage non-aléatoires, par exemple du phylogenetic clustering, c’est à dire lorsqu’une communauté est composée d’espèces phylogénétiquement plus proches qu’attendu par hasard étant donné le pool régional d’espèces et sa richesse spécifique, par exemple en raison d’un filtre abiotique sélectionnant des espèces possédant des traits similaires (e.g. présence d’un antibiotique) (Cadotte and Tucker 2017).
Box 1 : les méthodes unifiées – Exemple du framework de Chao et al. (2014)

Le formalisme de Chao et al. repose sur 2 concepts : la déclinaison de chaque facette de la diversité avec une pondération de l’importance des espèces (q) et l’expression de tous ces indices en nombres équivalents d’espèces, ou Hill numbers (Jost 2007). Le concept de « nombre équivalent » correspond au nombre d’espèces, de même abondance et également dissimilaires qui produirait la même valeur d’indice observée dans la communauté :

Communauté A « réelle »

Communauté B « théorique »

\[ qPD(T) = 4 \text{ équivalents d’espèces} \]

\[ qPD(T) = 4 \text{ équivalents d’espèces} \]

Dans une communauté composée de \( S \) espèces \( i=1,\ldots, S \) ayant des abondances relatives (ou tout autre métrique de pondération de leur importance écologique) notées \( p_i \), la diversité taxonomique peut être formalisée selon la formule suivante :

\[ qD = \left( \sum_{i=1}^{S} (p_i^q) \right)^{\frac{1}{1-q}} \quad ; \quad q \geq 0 ; \quad q \neq 1 \]

Lorsque \( q=0 \), les abondances relatives ne sont pas prises en compte et le nombre équivalent d’espèces est égal à la richesse spécifique \( S \) de la communauté. Lorsque \( q \) tend vers \( 1 \), les espèces sont pondérées proportionnellement à leurs abondances et la formule tend vers l’exponentielle de l’entropie de Shannon (i.e. exprimée en nombres équivalents d’espèces (Jost 2007)) :

\[ 1D = \lim_{q \to 1} qD = \exp \left( - \sum_{i=1}^{S} p_i \log p_i \right) \]

Lorsque \( q=2 \), les espèces abondantes auront plus de poids, tandis que les espèces rares auront peu d’impact sur la valeur de l’indice, qui correspondra à l’entropie de Simpson exprimée en nombres équivalents (Jost 2007).

Cette formule a été généralisée pour être applicable au calcul de différentes facettes de la biodiversité. Dans le schéma ci-dessus, les distances phylogénétiques entre espèces ont été intégrées dans la valeur de la diversité, calculée de la manière suivante :

\[ qPD(T) = \left( \sum_{i=1}^{a_T} L_i \left( \frac{a_i}{T} \right)^q \right)^{\frac{1}{1-q}} \quad ; \quad q \geq 0 ; \quad q \neq 1 ; \]

où \( L_i \) est la longueur d’une branche de l’arbre \( B_T \) et \( a_i \) est l’abondance totale issue de la branche \( i \) de l’arbre phylogénétique. \( T \) est la hauteur totale de l’arbre.

Ce formalisme a de nombreux avantages, en grande partie car il est basé sur les nombres équivalents d’espèces, qui permettent une interprétation très intuitive des valeurs de diversité. De plus, ces indices respectent la doubling property, stipulant que si l’on réunit deux assemblages de taille égale, sans aucune espèce en commun, la valeur de la diversité de l’assemblage obtenu doit être deux fois supérieure à celle des assemblages initiaux (Chao, Chiu, et Jost 2014). Cette propriété est nécessaire à la partition non biaisée de la diversité en composantes alpha et beta, et n’est pas respectée par la plupart des indices de diversité classiques, tels que l’indice de Shannon ou l’indice de Simpson (Jost 2007).

Cette méthode permet également le calcul d’indices de diversité beta, via une décomposition multiplicative de la diversité (Chiu, Jost, et Chao 2013), en divisant la diversité gamma mesurée à l’échelle des deux assemblages par la moyenne des diversités alpha mesurées dans chacun des assemblages :

\[ qD_\beta(T) = \frac{qPD_\gamma(T)}{qPD_\alpha(T)} \]
3.2 Particularités des communautés microbiennes

La plupart des indices cités ci-dessus ont été conçus pour l’étude de communautés de macro-organismes. Certains sont pourtant fréquemment utilisés en écologie microbienne, même si les méthodes unifiées impliquant l’évaluation de la biodiversité par plusieurs indices complémentaires sont encore peu appliquées, probablement pour des raisons pratiques (absence de scripts R publiés ou de logiciels). Par exemple, il est courant de voir des études basées sur la diversité alpha taxonomique et de la diversité beta phylogénétique, alors que les indices existent pour calculer la facette phylogénétique à ces deux échelles (e.g. (Lowrey et al. 2015) basée sur l’entropie de Shannon et l’indice Unifrac). Il est important de noter que les propriétés des indices n’ont pas encore été évaluées sur des communautés microbiennes alors que celles-ci présentent un certain nombre de caractéristiques uniques dans le monde vivant, explicitées figure 3.3, qui pourraient affecter la sensibilité des indices à détecter des règles d’assemblages non-aléatoires dans les communautés ou des différences significatives entre elles.

Une étude récente menée sur plus de 35 000 échantillons contenant plus de 5 millions d’espèces micro- et macroscopiques, a ainsi démontré que les communautés microbienne présentaient une richesse plus importante, une équité plus faible et consécutivement un plus grand nombre d’espèces rares que les communautés de macro-organismes. La même étude a estimé à 10^{12} le nombre total d’espèces microbienne sur Terre (Lacey et Lennon 2016), occupant l’essentiel de l’arbre du vivant (Figure 3.3). Ainsi, on retrouve environ millier d’OTUs bactériennes dans 10 litres d’eau de mer, alors que l’on peut dénombrer jusqu’à une vingtaine d’espèces le long d’un transect de 5×25 m en plongée dans un récif corallien (e.g. (Friedlander and Parrish 1998)).

Ces propriétés se surajoutent à la non exhaustivité des méthodes d’échantillonnage et de séquençage (chapitre 2), qui si elles ne sont pas prises en compte, peuvent entrainer une interprétation biaisée des indices de diversité classiques. Par exemple, il a été démontré qu’il est impossible d’estimer correctement la richesse d’une communauté microbienne, à cause du très grand – et surtout imprédictible – nombre d’espèces rares non échantillonnées (Hill et al. 2003; Haegeman et al. 2013).

Figure 3.3 : Spécificités des communautés microbiennes. Les communautés microbiennes présentent (A) une richesse très élevée ; ce qui est visible lorsque l’on compare les asymptotes des courbes d’accumulation du pico-nanoplancton eucaryote océanique (courbe bleu ciel) et celles obtenues à partir du microplancton et du mésoplancton eucaryotes, qui sont visibles à l’œil nu. Figure adaptée de Vargas et al. (2015). (B) une structure d’abondance très déséquilibrée, avec quelques OTUs hyper-dominantes et un très grand nombre d’espèces en très faible abondance. Figure adaptée de Pommier et al. (2010). (C) une diversité phylogénétique recouvrant l’essentiel de l’arbre de la vie (procaryotes dont archées et eucaryotes). Figure de Gaby D’Alessandro / © AMNH.
3.3 Objectifs du chapitre et principaux résultats

L’objectif de ce chapitre est de comparer les différentes méthodes d’analyse de la diversité dans le cas particulier des communautés microbiennes (voir ci-dessus). Le but est de mettre en évidence les avantages et/ou inconvénients des différents indices disponibles à partir de données issues de séquençage haut débit.

Pour la réalisation de cet objectif, nous avons souhaité répondre à deux questions complémentaires :
- Quelle est la capacité des différents indices à détecter les processus d’assemblage des communautés microbiennes ?
- Quelle est la capacité des indices à détecter des différences entre des communautés microbiennes?

Pour répondre à ces questions, nous avons créé artificiellement des communautés microbiennes. Pour répondre au 1er objectif, nous avons simulé des communautés microbiennes selon deux règles d’assemblage sur l’arbre phylogénétique (i.e. assemblage aléatoire et phylogenetic clustering) en faisant varier pour chaque type la richesse en OTUs (de 40 à 2000) et l’équitéabilité de leurs abondances (i.e. l’abondance relative de l’OTU la plus abondante variant entre 0.2 et 37%). A partir de ces communautés, les 6 indices de diversité phylogénétique les plus couramment utilisés ont été calculés. La capacité de chaque indice à détecter le phylogenetic clustering a été mesurée par comparaison avec leur valeur obtenue lors d’une répartition aléatoire des espèces sur l’arbre.

Pour répondre au deuxième objectif, 5 types de communautés plus ou moins contrastés sur l’arbre phylogénétique ont été simulés. 12 indices de diversité beta, incluant les indices les plus utilisés en microbiologie (Unifrac, Lozupone and Knight 2005) ainsi que ceux inclus dans les méthodes unifiées (Framework de Chao et al. (2014)), ont été calculés entre chacune des paires de communautés. La capacité des différents indices à détecter des différences entre ces communautés a été quantifiée par une PERMANOVA.

Dans les deux cas, nous avons relevé une plus grande efficacité des indices pondérés par les abondances relatives des espèces microbiennes. Les indices de diversité beta les plus efficaces étaient ceux basés sur les distances phylogénétiques entre espèces.
3.4 Manuscrit A

Contrasted performance of taxonomic and phylogenetic diversity indices to detect assembly rules and dissimilarities between microbial communities

Chiarello M.¹, Bouvier T.¹ and Villéger S.¹

¹Marine Biodiversity, Exploitation and Conservation (MARBEC), Université de Montpellier, CNRS, IRD, IFREMER, Place Eugène Bataillon, Case 093, 34 095 Montpellier Cedex 5, France

En préparation

Des informations supplémentaires complètent ce manuscrit, disponibles en section 7.3
Abstract
Many diversity indices have been proposed for measuring facets and components of biodiversity. Most of these metrics were designed for and tested on macrobial communities, but microbial communities have features that distinguish them from the macrobial ones, e.g. very high species richness and a low evenness of species abundances, that could impact the performance of diversity indices to reveal ecological signals.

Here, we simulated microbial communities to test the performance of (1) six phylogenetic diversity indices to efficiently detect the assembly rule that shaped communities, and (2) twelve beta-diversity indices to detect differences between communities. Indices accounting for abundance of species (e.g. Chao’s PDq indices with q=1 or 2) were the most efficient to detect clustering assembly. The distinction between two communities was improved with indices accounting for phylogenetic relatedness between species (e.g. weighted Unifrac).

Overall results highlighted that the choice of an index can strongly impact the conclusions of a microbial ecological study.
INTRODUCTION

Biodiversity is a multivariate concept that gathers complementary facets accounting for components measurable at different scales. The taxonomic facet of biodiversity considers species as independent units, while the phylogenetic diversity includes their phylogenetic distinctiveness. Both facets can include two components: considering only the presence-absence of species or phylogenetic lineages leads to the ‘compositional’ diversity (giving an equal weight to all species in a community), while accounting for their abundance refers to the ‘structural’ diversity. These facets and components measure the diversity within a community (alpha-diversity) or the dissimilarity between communities (beta-diversity). These alpha- and beta-diversities are the scales of biodiversity.

The use of complementary indices is required to have a comprehensive assessment of biodiversity as well as to understand its ecological determinants (e.g., Lozupone et al. 2007; Tucker et al. 2016)). Many biodiversity indices have been proposed for the last decades towards these aims (see a non-exhaustive list Table 1) with several indices currently available for measuring facets and components of alpha- and beta-diversity with even several indices for a given facet × component (e.g. weighted-Unifrac and Chao’s phylogenetic beta-diversity index βPD(q=1), Table 1).

Numerous studies using simulations and/or case studies have classified indices to look for complementarity or redundancy (Cadotte et al. 2010; Tucker et al. 2016; Miller, Farine, and Trisos 2016), while other studies have evaluated their performance to reveal non-random assembly processes such as phylogenetic clustering, i.e. when species present in a local community are less phylogenetically diverse than expected given species richness and species regional pool. (Kembel 2009; Mazel et al. 2015; Miller, Farine, and Trisos 2016) or to detect structural or compositional differences between communities, i.e. when clustering habitats (Koleff, Gaston, and Lennon 2003; Barwell, Isaac, and Kunin 2015; Baselga and Leprieur 2015). However, these studies focused mainly on communities with small richness values (i.e. less than 100 species), which typically occur in case of temperate communities of macro-organisms. In addition, the influence of a high unevenness of species abundances, and of the subsequent high proportion of rare species, which is typical from microbial communities, has not been assessed (but see e.g. (Hill et al. 2003; Haegeman et al. 2013)).
Recent advances in genomics (e.g. Next Generation Sequencing, NGS) have revolutionized the assessment of microbial biodiversity, both in terms of accuracy and completeness (Fierer et al. 2007; Logares et al. 2014; Frade et al. 2016). Massive sampling through space and/or time (see e.g. > 200 samples in world oceans (Yutin et al. 2007; Sunagawa et al. 2015) has revealed (i) the tremendous microbial richness in soil and water (e.g. more than 2 000 OTUs in 1g of soil (Roesch et al. 2007)), (ii) the highly skewed distribution of abundance among species, i.e. that a few species dominate in each community while many species are represented by a few cells (e.g. 99% of OTUs contributing to less than 0.1% of total abundance in soil (Baldrian et al. 2012)), and (iii) the presence of distant phylogenetic lineages within most microbial communities (e.g. Lactobacillus and Actinobacteria that frequently co-occur in human vagina diverged more than 3 billions years ago) (Hedges, Dudley, and Kumar 2006; Ma, Forney, and Ravel 2012).

<table>
<thead>
<tr>
<th>Biodiversity facet</th>
<th>Taxonomic diversity</th>
<th>Phylogenetic diversity</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Weight given to relative abundances</strong></td>
<td>$q=0$</td>
<td>$q=1$</td>
</tr>
<tr>
<td>Alpha diversity</td>
<td>$\text{TD}(q=0)$</td>
<td>$\text{TD}(q=1)$</td>
</tr>
<tr>
<td>Richness</td>
<td>$=_{\text{TD}}$</td>
<td>$=_{\text{TD}}$</td>
</tr>
<tr>
<td>Shannon</td>
<td>$=_{\text{TD}}$</td>
<td>$=_{\text{TD}}$</td>
</tr>
<tr>
<td>Simpson</td>
<td>$=_{\text{TD}}$</td>
<td>$=_{\text{TD}}$</td>
</tr>
<tr>
<td>Beta diversity</td>
<td>$\beta\text{TD}(q=0)$</td>
<td>$\beta\text{TD}(q=1)$</td>
</tr>
<tr>
<td>Jaccard</td>
<td>Sorensen</td>
<td>$=_{\text{PD}}$</td>
</tr>
<tr>
<td>Bray-Curtis</td>
<td>Unifrac</td>
<td>W-Unifrac</td>
</tr>
</tbody>
</table>

\[ g\text{-Unifrac}(d\in[0,1]) = W\text{-Unifrac} \]

| g-Unifrac(d=1) = W-Unifrac |

Table 1: Diversity indices tested in this study. Chao’s and Leinster & Cobbold’s unified frameworks were respectively indicated as $D(q)$ (corresponding to the $^qD$ in (Chao, Chiu, and Jost 2014)) and $D_{\text{LC}}(q)$ (corresponding to the $^qD^2$ in (Leinster and Cobbold 2012)), with prefixes T- or P- indicating respectively the taxonomic or phylogenetic version of both of them. The parameter $q$ indicates weight given to species abundance. Similarly, $d$ parameter of g-Unifrac indices permits indicates the importance given to relative abundances of OTUs. Some indices of both frameworks are related to other alpha diversity metrics when the latter are converted into Hill numbers; this equivalence was indicated by a ‘=’ sign.
Only a few studies have yet assessed indices behavior in the specific case of microbial communities and they all focused on comparing different components and/or facets of biodiversity (e.g. comparing the abundance-weighted and unweighted versions of Unifrac or Leinster & Cobbold’s measures (Lozupone et al. 2007; Veresoglou et al. 2014)), without further intra-facet of intra-component comparisons. Therefore, the performance of diversity indices to detect non-random assembly rules in microbial communities or differences between them remains unknown, while there is an urgent need for validating the ability of diversity indices to detect such patterns.

Here, using simulated realistic communities with various characteristics (i.e. varying levels of richness and evenness), we compared the performance of the most used taxonomic and phylogenetic diversity indices for (i) detecting the assembly rule that shaped a microbial community, and (ii) revealing ecological differences between a set of communities.

We hypothesized that the richness of microbial OTUs as well as the unevenness of their abundance affect the performance of diversity indices to detect assembly rules and dissimilarities between communities.

**MATERIALS AND METHODS**

We aimed to test intrinsic performance of diversity indices to detect patterns in microbial communities, *i.e.* independently from potential biases estimates of OTUs abundances due to sampling and/or sequencing (Lundin et al. 2012). We thus simulated microbial communities, not NGS outputs. We simulated a stochastic pure birth phylogenetic tree of 6 000 OTUs and used this tree and the corresponding species as the regional pool for all subsequent analyzes. Local communities were simulated according to 16 combinations of species richness and abundance distribution. We considered 4 levels of OTU richness (*S*=40, 200, 1 000 and 2 000 OTUs), which represent the range of values commonly observed in samples of natural microbial communities. We simulated 4 levels of unevenness of OTUs abundance for each level of species richness. More precisely, abundances were randomly sorted from a lognormal distribution of unit means, which standard deviations were of σ=0.25, 0.5, 1 or 2 respectively. The higher this parameter σ, the higher is the unevenness of abundances, *i.e.* the higher is the difference between most abundant and least abundant OTUs (see S2).
1. Performance of alpha-diversity indices for detecting phylogenetic clustering

1.1 Simulation of microbial communities according to contrasted assembly rules

We first tested the performance of phylogenetic alpha-diversity indices to detect whether phylogenetic clustering shaped a community. Phylogenetic clustering is detected when species in a community are less distantly related to each other than expected from a random sorting of species from the regional pool (Mouquet et al. 2012). Phylogenetic clustering is often driven by environmental filtering of phylogenetically conserved biological features such as tolerance to a stress (Horner-Devine and Bohannan 2006), sometimes in interaction with other mechanisms (Cadotte and Tucker 2017). We did not aim to simulate a drastic phylogenetic clustering affecting phylogenetic composition but rather a more realistic situation where marked abiotic or biotic constraints favor microbes that have the ecological characteristics needed to dominate in a community while OTUs that do not have such characteristics could occur (for instance through immigration from other source habitats) but remain at low abundance (Gibbons et al. 2013; Lynch and Neufeld 2015).

We simulated 40 types of communities (see supplementary information S1) differing in their richness and unevenness of OTU abundances (i.e. the level of dominance in the community), and their phylogenetic assembly rules (2 modalities, clustering and random assembly).

Then, for each of these simulated distributions of OTUs abundances (i.e. for each $S \times \sigma$ scenario) we simulated two microbial communities according to two assembly rules. First, a community was simulated under a random assembly process, by assigning simulated abundances to OTUs randomly picked along the phylogenetic tree. Then, a second community with the same structure of abundance (i.e. with the same $S \times \sigma$), resulting from a hypothetical phylogenetic clustering, was simulated using a three-steps procedure: (i) an OTU randomly picked of the phylogenetic tree was attributed the highest relative abundance, (ii) phylogenetic distances between this dominant OTU and all other OTUs of the tree were computed, and (iii) the remaining $S-1$ abundance values sorted in a decreasing order were successively attributed to OTUs using a random sampling (no replacement) with odds proportional to the inverse of the distance with the dominant OTU. This way, OTUs closer to the dominant OTU on the phylogenetic tree had a higher probability to be sampled earlier during the process, and therefore to have higher abundances. This procedure ensured that dominant OTUs were phylogenetically more similar than expected at random, while allowing the many rare OTUs to be distant from these dominant OTUs. Simulation of random and phylogenetically clustered communities was independently replicated 500 times for each level.
of species richness and unevenness of abundances. Simulated communities thus do not account for potential non-random patterns in species frequency and regional total abundance observed in some metacommunities (e.g. mass-effect processes) or trait-evolution (Kraft et al. 2007; Miller, Farine, and Trisos 2016).

Simulation process yielded a total of 16,000 microbial communities (4 levels of OTUs richness $S \times 4$ levels of abundance unevenness $\sigma \times 2$ assembly rules $\times 500$ replicates) (S1).

1.2 Selection and computation of alpha-diversity indices

Numerous indices are currently available in the literature (see e.g. Tucker et al (2016) for phylogenetic indices). The purpose of this paper was not to evaluate all of them, especially because many of them are highly correlated (Tucker et al. 2016). We selected two sets of indices that are based on Hill numbers and allow to measure taxonomic and phylogenetic diversity with the same unit (i.e. equivalent number of species, or Hill numbers, (see (Jost 2007; Chao, Chiu, and Jost 2014), Table 1). The first set of indices comes from Chao et al. framework. This unified framework is based on a single formula that account for different biological entities (species or length of phylogenetic branches), while permitting to choose the relative weight given to species dominance (e.g. abundance, biomass) (parameter $q$) ( (Chao, Chiu, and Jost 2014) ). When $q=0$ the index is insensitive to dominance, and thus measures compositional diversity. When $q=1$, the dominance of entities and their phylogenetic distinctiveness have the same contribution to the index value. When $q=2$, abundances dominance of entities have twice more weight than the phylogenetic distances. We chose Chao’s indices because they are related to well-known alpha-diversity metrics, when converted into Hill numbers (Jost 2007). For instance, phylogenetic Chao’s diversity indices with $q=0$, 1, and 2 are related to the widely used Faith’s PD, Allen’s and Rao’s indices, respectively (Table 1) (Rao 1982; Faith 1992; Allen, Kon, and Bar-Yam 2009). The second set of indices tested comes from the unified framework of (Leinster and Cobbold 2012), which is also based on a single parameter modulating the importance given to dominant entities, but with a similarity-based approach instead of the dissimilarity-based approach used by Chao et al. (Chao, Chiu, and Jost 2010). Contrary to Chao’s indices, phylogenetic diversity indices proposed by Leinster & Cobbold ($\text{PD}_{\text{LC}}(q=0, 1 \text{ or } 2)$) do not have equivalents among classical phylogenetic diversity metrics. These 6 phylogenetic alpha-diversity indices were computed using the entropart R-package (Marcon and Hérault 2014).
1.3 Assessing indices performance to detect phylogenetic clustering

The ability of each phylogenetic index to detect phylogenetic clustering for each scenario of OTU richness and unevenness of OTU abundances was assessed by comparing the diversity of each assemblage simulated under the clustering assembly rule with the diversity of all the communities simulated under a random process. More precisely, we computed the P-value as the rank of the observed diversity value in the clustering assembled community among the increasingly sorted diversity values of all \( r = 500 \) replicates of randomly assembled communities corrected by \( (r-1) \). We considered that a diversity index efficiently detected phylogenetic clustering when \( P<0.05 \), meaning that the diversity value in the observed community was lower than the 5% lowest diversity values in the random communities. This null-model approach corresponds to the “richness-like” category in Gotelli (2000), i.e. simulated communities have the same species richness and same abundance distribution than the observed community and regional pool is the same for all communities (e.g. (Mazel et al. 2015)).

2. Performance of beta-diversity indices to detect differences between communities

2.1 Datasets simulation

The objective was to use this set to assess the performance of beta-diversity indices to detect differences between community types, such as the ones expected when studies communities from environments experiencing different abiotic and/or biotic conditions. We defined a set of 5 hypothetical community types from A to E with type B being as type A (to test for false positive detection) and types C, D and E being increasingly different from A (see S3 for a comprehensive scheme of the method). For each community type, 6 communities were randomly simulated using the same clustering procedure as for alpha-diversity (see above). A total of 50 sets of such 30 communities were simulated.

Communities were simulated following successive steps. First, a ‘localizer OTU’ was randomly sampled on the phylogenetic tree, and was used to create all the communities, but was not included in any of them. Phylogenetic distances between this localizer and all other OTUs on the tree were computed and ranked from the smallest to the largest distance value, and separated into 400 quantiles. Then, a ‘reference OTU’ was randomly sampled in the quantiles for each type, using the rule described below, and kept constant for all communities of the given type, but not included in any of them:
For A and B community types, the same reference OTU was taken, and was sampled within the distances to the localizer OTU lower than the third quantile (0.5% closest OTUs); these communities were used to quantify erroneous detection of a difference between communities of a same type (i.e. false positives).

The reference OTU of community type C was randomly sampled among the OTUs located between the 1.75 to 2.25% closest OTUs (i.e. among 7th, 8th and 9th quantiles);

- The reference OTUs from community types D and E were respectively randomly sampled among the 4.75-5.25% and 9.75-10.25% closest OTUs (i.e. among 19th, 20th and 21th quantiles, and 39th, 40th and 41th quantiles, respectively).

Then, phylogenetic distances between each reference OTU and all other OTUs on the tree were computed, and to ensure that OTUs that would be sampled in the next steps would be close enough to their reference OTU, only OTUs with a distance below the median of all distance values were retained for the next step (i.e. among 3000 OTUs, which is higher than the maximum richness of communities we tested). The dominant OTU of a given community was then sampled on the tree with a probability proportional to the inverse of the distance to the reference OTU, and it received the highest simulated abundance. Finally the remaining S-1 abundances values were attributed to the OTUs following the clustering procedure used before. This procedure ensured that communities within a type were more similar to each other than with communities from other types, while allowing for rare species being phylogenetically similar between community types.

2.2 Selection and computation of beta-diversity indices

We computed a total of 12 beta-diversity indices from three frameworks. The first set of indices grouped classic Jaccard and Bray-Curtis indices. Jaccard is a presence/absence beta-diversity measure, correlated to the Sorensen (Sorensen= Jaccard / (2- Jaccard)). The Bray-Curtis index is the generalization of the Sorensen index, accounting for the relative abundances of species (Bray and Curtis 1957). These two indices were computed using the betapart R-package (Baselga and Orme 2012).

The second set of indices came from the Unifrac framework that extended the Jaccard index to account for phylogenetic information (C. Lozupone and Knight 2005), and which is frequently used in microbial ecology (e.g. (Pontarp et al. 2012; Clemente et al. 2015)). Unweighted Unifrac (U-Unifrac), measuring the fraction of unshared branch length between
species from two communities on a common phylogenetic tree, is insensitive to OTUs abundances (C. Lozupone and Knight 2005). This index was then extended to include the relative abundance of phylogenetic lineages (Weighted-Unifrac, (C. A. Lozupone et al. 2007)). Chen et al. (Chen et al. 2012) proposed a set of generalized Unifrac distances (g-Unifrac), permitting to imbalance the importance given to abundances and the phylogenetic distances between OTUs using a single parameter \( d \) varying between 0 and 1: g-Unifrac\((d=0)\) gives a low weight to OTUs relative abundances (while being distinct from U-Unifrac), while g-Unifrac\((d=1)\) gives a stronger weight to OTU abundances and is identical to the Weighted-Unifrac \((W\text{-Unifrac})\). We computed g-Unifrac indices with \( d = 0, 0.5 \) and 1, as well as the U-Unifrac, using the \textit{GUniFrac} package (Chen and ORPHANED 2012).

The third set of indices came from Chao et al. framework, based on a multiplicative decomposition of diversity (Chao, Chiu, and Jost 2014). In other words, beta-diversity between two communities is computed as the ratio between the gamma-diversity (diversity of the pooled communities) and the average of the alpha-diversity of the two communities (Chao et al. 2014). Taxonomic and phylogenetic beta-diversity indices were computed using the formula accounting for species or phylogenetic lineages, respectively, giving weights of \( q=0, 1 \) and 2 to their relative abundances. We then scaled the beta-diversity values between 0 and 1 as proposed by Villéger et al. (Villéger et al. 2012). These beta-diversity indices were calculated using personal functions based on the \textit{entropart} R-package, provided online: https://github.com/marlenec/chao, and in S4. Note that some of Chao’s beta-diversity indices are closely related to other indices obtained by a multiplicative decomposition of diversity (Chao et al. 2014). For instance, Chao’s taxonomic beta-diversity index with \( q=0 \) \((\beta\text{TD}(q=0))\) is related to Sorensen index (Table 1).

All these indices vary from 0, when communities share the same phylogenetic lineages (in case of U-Unifrac) or relative abundances of OTUs (in case of Bray-Curtis) to 1, when communities are composed of distinct phylogenetic lineages or are dominated by different OTUs.

### 2.3 PERMANOVAs

For each of the 16 \((S\times\sigma)\) scenarios, all the beta-diversity indices were computed for the 66 pairs of communities compared, that is 30 pairs within each community type (\textit{i.e.} 15 combinations of 2 communities among the 6 from the same type) plus 36 pairs of
communities belonging to distinct community types (S3). Then, for each index and each of
the 4 comparisons of community types (A vs. B, A vs. C, A vs. D and A vs. E) a
PERMANOVA was performed on a matrix containing all 66 values of beta-diversity using
the ‘adonis’ function in Vegan R-package (Dixon 2003). The explanatory variable used was
the type of each community, and the distinction between the compared community types was
considered successful when the PERMANOVA raised a P-value<0.05. Note that as
communities from type A and B were simulated using the same parameters, they are expected
to be not significantly different and thus this simulation allows to measure false-positive rate.
The entire simulation procedure was replicated 50 times, i.e. from the random sorting of the
community types to the PERMANOVA. Finally we computed the percentage of successful
PERMANOVAs as a measure of index performance.

3. Statistical analyses
To disentangle the relative influence of the simulation scenarios (S×σ) and the chosen alpha
or beta-diversity index on the success of detection of clustering or environmental difference
we ran Boosted Regression Trees (BRT) using P-values from the clustering experiment or the
PERMANOVAs. For the clustering detection test, a single BRT model was constructed using
the 48 000 P-values (4 S × 4 σ × 6 indices × 500 replicates) with the richness S, the
unevenness σ, and alpha-diversity index as explanatory variables. For the experiment
focusing on dissimilarity between communities, BRT models were constructed separately on
each facet of biodiversity (i.e. on P-values raised from PERMANOVAs performed
respectively on taxonomic and phylogenetic beta-diversity indices) and on each pair of
community types (A/C, A/D and A/E), making a total of 6 BRT models. Taxonomic and
phylogenetic BRTs were constructed respectively on a total of 4 000 P-values (4 S × 4 σ × 5
indices) and 5 600 P-values (4 S × 4 σ × 7 indices). Both taxonomic and phylogenetic BRTs
were constructed using the richness, unevenness, and beta-diversity index as explanatory
variables. BRT were fitted using the ‘gbm.step’ function provided in the dismo R-package
(Hijmans, Phillips, and Elith 2016), with a bag fraction of 0.75, a learning rate of 0.001 and a
tree complexity of 3, following the recommendations of (Elith, Leathwick, and Hastie 2008).
All BRT models were fitted using more than 1 000 trees.
RESULTS

1. Performance of phylogenetic alpha-diversity indices to detect clustering assembly

Phylogenetic Chao’s and Leinster & Cobbold’s alpha-diversity indices showed contrasted patterns with varying levels of $q$ (Fig. 1 & S5). As expected, Chao’s PD($q=0$) was independent of abundance unevenness, while PD($q=1$) and PD($q=2$) decreased as unevenness increased. Contrary to the Chao’s index, PD$_{LC}(q=0)$ varied with unevenness, and ranged from 6.2 to 5.4 for a community composed of 1 000 OTUs. Similarly, PD$_{LC}(q=1)$ and PD$_{LC}(q=2)$ decreased as the unevenness increased, but with a lower slope than Chao’s indices (S5).

Overall, clustering detection efficiency ranged between 2% to 100%. The used index had the strongest impact on the clustering detection (BRT model, relative influences of 75.1%, 22.7% and 2.2%, respectively for the used alpha-diversity index, the species richness and abundance unevenness). Abundance-weighted indices (including PD$_{LC}(q=0)$, which give a small weight to abundances) performed up to 7 times better than indices taking only account of OTUs presence/absence (Fig. 1 and S6). For instance, PD($q=0$), which is independent from OTUs abundances, detected clustering in only 8.4 to 17.2% of cases, while PD($q=2$), which is weighted by OTUs abundances, reached a detection efficiency of 18 to 100%. The most efficient indices of Chao’s and Leinster & Cobbold’s frameworks (i.e. PD($q=1$ and 2) and PD$_{LC}(q=1$ and 2)) did not significantly differ for all assemblage tested (Kruskal Wallis, $P>0.05$).
Figure 1: Efficiency of clustering detection by Chao’s indices. Values are expressed as percentage of significant P-values (<0.05) across the 500 replicates, for each level of OTUs richness (S, titles of panels) and each level of unevenness of their abundances (σ, abscisses).

2. Efficiency to detect differences between communities

As expected, the detection efficiency of the difference between community types was higher for very distinct types, averaging 61.3±33.4% and 74.7±33.0% (respectively for taxonomic and phylogenetic diversity indices) when comparing communities from A and E types, and 44.6±32.7% and 55.2±35.8% when comparing A and C types (Fig. 2, Fig. 3 and S7 & S8). The detection of significant differences was however variable for a given pair of community types, depending mostly on the used index (Fig. 2 and Fig. 3).
Figure 2: Efficiency of taxonomic diversity indices to detect difference between communities belonging to different community types, measured as the percentage of significant P-values (PERMANOVA, $P<0.05$) across the 50 replicates, for communities composed of 200 and 1 000 OTUs (S, panels titles) and each level of unevenness ($\sigma$, abscissa). For results obtained with 40- and 2 000-OTUs communities, see Supplementary Information S7.
Figure 3: Efficiency of phylogenetic diversity indices to detect the difference between communities belonging to the different community types, measured as the percentage of significant P-values (PERMANOVA, $P<0.05$) across the 50 replicates, for communities composed of 200 and 1000 OTUs (S, panels titles) and each level of unevenness ($\sigma$, abscissae). For results obtained with 40- and 2000-OTUs communities, see Supplementary Information S8.
First, abundance-weighted indices detected a significant difference between community types in 83.3±24.6% of cases, while composition-based indices detected a difference in only 21.6±15.0% of cases, whatever the pair of community types and the facet of biodiversity considered. Detection of difference also varied between biodiversity facets. For taxonomic diversity, both the identity of the beta-diversity index and intrinsic features of the community (richness and, in a lesser extent, unevenness) influenced P-values raised from PERMANOVAs (Table 2). For instance, the percentage of difference detected was higher in communities composed of more than 40 OTUs, especially in case of relative abundance-weighted indices (Fig. 2 and S7). By contrast, the identity of phylogenetic index explained 90 to 93% of variability of PERMANOVA P-values (Table 2), with a higher performance of abundance-weighted indices than composition-based indices (Fig. 3 and S8).

<table>
<thead>
<tr>
<th>Types</th>
<th>Factor</th>
<th>Relative influence (%)</th>
<th>Relative influence (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>A vs C</td>
<td>Beta-diversity Index</td>
<td>68.7</td>
<td>90.6</td>
</tr>
<tr>
<td></td>
<td>Richness (S)</td>
<td>22.0</td>
<td>6.0</td>
</tr>
<tr>
<td></td>
<td>Uneveness (σ)</td>
<td>9.3</td>
<td>3.4</td>
</tr>
<tr>
<td>A vs D</td>
<td>Beta-diversity Index</td>
<td>59.0</td>
<td>92.6</td>
</tr>
<tr>
<td></td>
<td>Richness (S)</td>
<td>33.4</td>
<td>5.9</td>
</tr>
<tr>
<td></td>
<td>Uneveness (σ)</td>
<td>7.6</td>
<td>1.5</td>
</tr>
<tr>
<td>A vs E</td>
<td>Beta-diversity Index</td>
<td>51.5</td>
<td>91.8</td>
</tr>
<tr>
<td></td>
<td>Richness (S)</td>
<td>38.3</td>
<td>5.9</td>
</tr>
<tr>
<td></td>
<td>Uneveness (σ)</td>
<td>10.2</td>
<td>2.3</td>
</tr>
</tbody>
</table>

Table 2: Relative influence of the used index, OTUs richness and unevenness on the efficiency of community types detection. The relative influence values of each tested factor are based on the number of times the factor is selected for splitting while constructing each tree, weighted by the squared improvement to the model resulting from each split and averaged over all trees. BRT models were performed on P-values raised from PERMANOVAs testing the difference between community types, separately for taxonomic and phylogenetic beta-diversity indices.

Moreover, considering only the efficiency of abundance-weighted indices, phylogenetic diversity indices were on average more efficient than the composition-based ones, the former averaging 83.34±17% to 97±8.3% of environmental detection (for A/C to A/E comparisons), and the latter averaging 62.96±30.3% to 77.6±30.4% of detection (see Fig. 2, Fig. 3, S7 and
S8). Finally, the proportion of erroneous detection of a difference between identical communities (i.e. A vs. B types) was overall low (3.8±3.6%) (S9). However, in case of moderately skewed rich communities (S=1000, sigma=0.5), all taxonomic and phylogenetic indices raised erroneous detection of a difference in more than 6% of cases. False-detection rate higher than 10% were found for the communities with only 40 species and high evenness of abundance (S=40, sigma=0.25) for all phylogenetic diversity indices (14.5±4.4% across tested indices).

Indices measuring the same facet and component of biodiversity generally reached similar efficiencies. For instance the Bray-Curtis index had intermediate behavior between βTD(q=1) and βTD(q=2), and the Jaccard index reached similar efficiencies as βTD(q=0) (Fig. 2 & S7). Similarly, W-Unifrac efficiencies were comparable to that of βPD(q=2) (Fig. 3 & S8).

**DISCUSSION**

Chao’s and Leinster & Cobbold’s frameworks are both based on Hill numbers (Leinster and Cobbold 2012; Chao, Chiu, and Jost 2014). Their taxonomic diversity indices are numerically identical. However their phylogenetic diversity indices showed contrasted patterns, especially at q=0 (S5). While Chao’s PD(q=0) was independent of the unevenness of abundance in the community, its equivalent from Leinster & Cobbold framework (PD_{LC}(q=0)) decreased as the unevenness increased. Such a sensitivity to OTUs abundances prevent from considering this index as a relevant compositional index of biodiversity (Leinster and Cobbold 2012). Consequently, Chao et al. framework is the only one taking account of the compositional component of phylogenetic diversity, i.e. the diversity of phylogenetic lineages present, independently of their respective dominance (Chao, Chiu, and Jost 2010).

Considering Chao’s framework, phylogenetic clustering was hardly detected by its composition-based index, as PD(q=0) detected it in less than 20% of simulated cases (Fig. 1). Abundance-weighted Chao’s indices, by contrast, were much more efficient, detecting clustering in ca. 86.3±23.3% of cases. This low efficiency of composition-based indices is related to their key property that is giving the same importance to all the OTUs present, whatever their relative abundance in the community. Consequently, composition-based indices are unlikely to distinguish between a clustered community and a randomly assembled one, due to the many rare OTUs that blur the phylogenetic closeness between dominant OTUs. Yet, detecting clustering in natural communities using composition-based indices is
unlikely but not impossible when phylogenetic clustering is drastic and is thus affecting composition. Horner and Bohannan (Horner-Devine and Bohannan 2006) identified clustering in 55% of microbial assemblages studied from distinct environments (sediments, soil, freshwater mesocosms). (Miller, Farine, and Trisos 2016) showed a high performance of Faith’s PD, \( i.e. \ PD(q=0) \) in our study, to reveal habitat filtering using simulated communities and richness-like null model. They also showed a slightly lower efficiency of abundance-weighted indices to reveal this pattern. The differences between their results and ours may be related to the difference in simulation of phylogenetically clustered communities since Miller et al (2016) simulated species abundances independently from community composition (that was simulated based on trait-evolution) while we simulated both processes simultaneously. (Freilich and Connolly 2015) already argued that environmental filtering should impact both species occurrence and abundance, as species that are the best adapted to a given environment would also be the most abundant, and that other species that do not have the physiological traits necessary to grow in a given environment would not be abundant but could still be present. Accordingly, they showed a higher performance of abundance-weighted indices to detect environmental filtering compared to indices based only on species composition. Abundance-weighted indices are thus more likely to detect phylogenetic clustering, and therefore should be used in priority.

The efficiency of clustering detection was dependent on the intrinsic features of the communities (richness, unevenness). Clustering was poorly detected in poor communities (40 OTUs), as even the most efficient indices (\( PD(q=2) \) and \( PD_{lc}(q=2) \)) never reached 80% efficiency. This efficiency was particularly low when the unevenness was high, dropping the detection to ~20%. In communities with very few OTUs and a few ones contributing mostly to the cumulated OTU abundance \( i.e. 40\% \) of OTUs totalizing more than 95% of total abundance, S3), values of abundance-weighted indices were very low in both clustered and randomly assembled communities \( \text{e.g. average } PD(q=2) \text{ of } 1.6\pm0.4 \text{ and } 4.1\pm1.4, \text{ respectively} \), making difficult to disentangle these two patterns. This suggests that for such species-poor communities, it may be challenging to detect clustering. It also underlines that low sampling and sequencing effort of OTUs-rich communities that results in low observed OTUs richness because of the absence of the rarest species (Hill et al. 2003; Bent and Forney 2008; McCoy and Matsen IV 2013; Haegeman et al. 2013) could prevent from unraveling ecological processes. In our simulation, we did not consider the sampling issues that could rise from NGS data, as insufficient sampling, PCR artifacts and sequencing errors. However,
considering that these biases mainly impact the rarest members of the community (Pinto and Raskin 2012). Overall, our results highlight the need of a first analysis of community features (richness, unevenness) before further analyses and that abundance-weighted indices should be preferred to the only use of composition-based indices to detect non-random assembly rules. Last, previous studies based on simulated species-poor communities (i.e. less than 200 species) demonstrated that efficiency of clustering detection depends on tree shape (Kraft et al. 2007; Mazel et al. 2015). Additionally, Mazel et al. tested different indices on mammal gut microbial assemblages and showed contrasted sensitivity of indices to length of basal or terminal branches of the tree. Further studies are therefore needed to test for effect of phylogenetic tree shape, especially for presence of many long basal branches, typical of microbial phyla that diverge more than 3 billions years ago.

The weaker performance of composition-based indices was also outlined for detecting differences between distinct community types, since composition-based beta-diversity indices detected such differences in only ~20% of cases (e.g. Unifrac, Fig. 3). Contrastingly, abundance-weighted indices reached 70% to more than 90% efficiency, respectively for taxonomic and phylogenetic diversity indices (Fig. 2, Fig. 3, S7 and S8). The weight given to the q and d values for Chao’s and g-Unifrac abundance-weighted indices did not impact the detection, in all but two cases. The first one was the g-Unifrac(d=0), which gives a very small weight to the abundances of lineages, and had an intermediate efficiency between abundance-weighted and composition-based indices (Fig. 3). Its efficiency was particularly low in moderately distinct communities (from type A vs. types D and C). This may be due to the low sensitiveness of g-Unifrac(d=0) to differences in abundances of phylogenetic lineages between communities that have similar phylogenetic composition. The second case was for closer communities (A vs. C types), where W-Unifrac and βPD(q=2) were slightly more efficient than their equivalents that were less weighted by abundances (i.e. respectively g-Unifrac(d=0 and 0.5) and βPD(q=1), Fig. 3). For instance, βPD(q=2) was more sensitive to difference in OTUs abundances compared to βPD(q=1) when communities hosed the same phylogenetic lineages. However when communities share the same abundant OTUs while having different rare species abundance-weighted indices would hardly detect differences. For instance, Chen et al. (Chen et al. 2012), showed a higher sensitivity of U-Unifrac and g-Unifrac(d=0.5) to variations of the rare and the moderately abundant lineages, respectively. This result
highlights the need of comparing indices accounting differently for OTUs abundance to
disentangle differences between communities driven by dominant and rare species.

Besides the higher performance of abundance-weighted indices, our simulations also
demonstrated that accounting for taxonomic or phylogenetic facet of biodiversity had a strong
impact on the detection of differences between communities. The performance of taxonomic
diversity indices to detect the difference between community types was highly dependent on
the intrinsic features of the communities (richness and unevenness), while these features had
much less impact on the performance of phylogenetic diversity indices (Table 2). Accordingly, the efficiency of abundance-weighted phylogenetic diversity indices was less
variable than the abundance-weighted taxonomic diversity indices (i.e. respectively
interquartile ranges of 78-100% and 39.5-100% efficiency), and suggests that phylogenetic
diversity indices would be more reliable. Abundance-weighted phylogenetic diversity indices
had on average 20% higher detection rates than the abundance-weighted taxonomic ones.
This higher efficiency of phylogenetic diversity indices was particularly striking for
phylogenetically similar communities (A vs. C types), while for more distant communities (A
vs. E types), abundance-weighted taxonomic diversity indices reached nearly the same
efficiency as the abundance-weighted phylogenetic diversity indices (Fig. 2 & 3 and S7 &
S8). This smaller gap between taxonomic and phylogenetic diversity indices in this latter case
was driven by the very few OTUs in common between communities belonging to more
distant community types (A vs. E) that enhanced the detection of the difference without the
use of phylogenetic information.

Several studies highlighted the need of considering the phylogenetic relationships between
organisms to assess the biodiversity of assemblages, which contain more biological
information than OTU membership (C. Lozupone and Knight 2005; Chen et al. 2012). This
was confirmed by Doll et al. (2013), who used both real sets of communities and simulated
assemblages to reveal the benefits of adding phylogenetic information into alpha-diversity
metrics. However, to our knowledge, the present study is the first that demonstrated the
higher efficiency of phylogenetic beta-diversity indices for detecting differences between
communities.

We also demonstrated that efficiency of detection of phylogenetic clustering and differences
was affected by the intrinsic features of the communities (richness, unevenness). Indeed,
detection of difference between communities was facilitated in communities with moderately to highly skewed ($\sigma = 1$ or 2) distributions of OTUs abundances, especially in those containing less than 1000 OTUs, as abundance-weighted phylogenetic diversity indices were ~20% more efficient than in less skewed communities. In poorer communities, less information (i.e. less phylogenetic lineages) is indeed available to discriminate communities. In these conditions, if the compared communities are dominated by a few OTUs, which likely differ, they should be more easily discriminated by abundance-weighted indices. In most cases diversity indices provided false positive rate lower than 5% (i.e. detected significant differences between communities belonging to the same type, S9). False positive rate exceeded 5% in only a few scenarios: all taxonomic and phylogenetic beta-diversity indices in species-rich communities with high evenness abundance (i.e. $S=1000$, $\sigma = 0.5$) and only phylogenetic beta-diversity indices for communities with only 40 species and high evenness abundance ($S=40$, $\sigma = 0.25$), where false positive rate raised 18%. However, in simulations close to microbial communities, i.e. $>1000$ OTUs having highly-skewed abundances ($\sigma \geq 1$), phylogenetic beta-diversity indices raised false positives in less than 6% of cases testifying for their relevance. 

To conclude, our simulations demonstrated that phylogenetic diversity indices accounting only for species presence/absence (i.e. Faith’s PD) should be used with caution when aiming to detect non-random assembly processes while Rao’s phylogenetic diversity index (as modified by Chao, i.e. PD(q=2)) performed well in most situations. Similarly, phylogenetic beta-diversity indices weighted by OTUs relative abundance performed much better than composition-based phylogenetic diversity indices and than to all taxonomic diversity indices. W-Unifrac reached similar efficiencies to Chao’s $\beta$PD(q=2) but Chao’s diversity indices allow computing facets and components of biodiversity within a unique framework, while g-Unifrac indices (including W-Unifrac) are restricted to measuring phylogenetic beta-diversity. Analyzing complementary indices describing the taxonomic and phylogenetic facets including presence and abundance of OTUs as well as their phylogenetic relationships is certainly the most pragmatic way to fully understand the processes that shaped the multifaceted microbial biodiversity, i.e. there is no single perfect index. Finally, properties of indices highlighted here with simulated microbial communities are valid when studying any type of species-rich communities, including tropical assemblages of plants or vertebrates.
4. Facteurs de variation du microbiome cutané des vertébrés marins

4.1 Echelles de variabilité et facteurs environnementaux

Nous avons vu lors du chapitre 1 que les quelques études menées sur les vertébrés marins ont montré qu’ils sont recouverts par une importante diversité de micro-organismes et que le microbiome cutané est à la fois distinct des communautés présentes dans la colonne d’eau (Apprim et al. 2014; Doane et al. 2017), mais également spécifique de l’espèce du macro-organisme hôte (Larsen et al. 2013).

Ces quelques études menées sur plusieurs espèces de vertébrés de mers tempérées suggèrent aussi que le microbiome de surface des animaux marins présente plusieurs niveaux de variabilité, à la fois entre espèces et entre individus d’une même espèce (Larsen et al. 2013; Apprim et al. 2014). Ces différences peuvent être liées à la fois à des facteurs propres à la l’individu ou à l’espèce (fonctionnement du système immunitaire, écologie ou comportement), facteurs dit endogènes, mais également à des facteurs exogènes liés à des effets directs de l’environnement sur le microbiome de surface (e.g. ajout d’un désinfectant dans l’eau, Mohammed et Arias 2015) ou indirects (e.g. effets d’un changement de la salinité environnante sur la physiologie de l’hôte, entrainant des modifications du microbiome, Schmidt et al. 2015). Ces différents facteurs exogènes et endogènes peuvent avoir un impact à la fois à l’échelle de l’individu, entre individus, mais également entre espèces, et sont indiqués en Figure 4.1.
La plupart des études ayant testé l’existence de ces différents facteurs sur le microbiome cutané des vertébrés marins se sont focalisées sur le test d’un seul facteur, mesuré à un seul niveau de variabilité (Larsen et al. 2013; Aprrill et al. 2014; Doane et al. 2017). De plus, la totalité des études menées sur ces animaux ont été réalisées en conditions naturelles, permettant potentiellement d’avoir une vision du microbiome « naturel » des espèces ciblées, mais ne permettant pas de distinguer l’influence relative des facteurs exogènes et endogènes sur la variabilité du microbiome, car les individus peuvent avoir fréquenté des habitats différents plus ou moins longtemps avant leur capture. Pour distinguer la part des facteurs endogènes, il est en effet nécessaire de travailler dans un environnement contrôlé, ou au minimum uniforme et connu entre individus et entre espèces.

De plus, bien que la totalité du corps des poissons et des cétacés soit recouverte du même tégument (Angeles Esteban 2012), les différentes parties du corps de ces animaux pourraient, tout comme chez l’Homme, présenter des conditions physico-chimiques contrastées. Ces différences pourraient être dues à (i) des variations de la composition ou de la quantité de nutriments présents à la surface de la peau ou dans le mucus des poissons téléostéens, (ii) une exposition variable à l’excrétion de nutriments (ammonium par les branchies, phosphates et urée par la papille urogénitale et l’anus) et/ou (iii) une exposition variable aux courants d’eau durant la nage. Ces différents facteurs pourraient engendrer la coexistence de microhabitats distincts et donc entraîner une variabilité intra-individuelle (i.e. entre les parties du corps) du microbiome cutané. Cependant, la diversité intra-individuelle du microbiome cutané n’a jamais été mesurée dans le cas des macro-organismes marins, et donc son importance relative par rapport à la variabilité observée entre individus et entre espèces n’a jamais été évaluée.

L’objectif de ce chapitre est donc de quantifier l’importance relative des différents niveaux de variabilité du microbiome cutané des animaux marins, prélevé dans un même environnement. Pour cela, j’ai travaillé à 3 échelles spatiales complémentaires (bassins d’élevage, parc marin, lagon).

Dans un premier temps, j’ai quantifié simultanément la variabilité du microbiome aux échelles inter- et intra-individuelle de deux espèces de poissons téléostéens (Dicentrarchus labrax et Sparus aurata) élevés en conditions contrôlées dans des bassins (Manuscrit B). J’ai mis en évidence une spécificité du microbiome cutané pour son espèce hôte. J’ai également mis en évidence une forte variabilité à la fois entre individus d’une même espèce, mais également entre zones du corps, mais ces différences n’étaient consistantes ni entre individus, ni entre zones du corps.

Ensuite, j’ai quantifié pour la première fois la variabilité du microbiome chez deux espèces de cétacés odontocètes, le grand dauphin (Tursiops truncatus), et l’orque (Orcinus orca) élevées au parc marin Marineland d’Antibes (http://www.marineland.fr/fr) dans des conditions semblables (Manuscrit C). Le microbiome cutané de ces deux espèces était très
diversifié, et présentait des fonctions distinctes de celles des communautés planctoniques. Il était également spécifique de son espèce hôte. Contrairement aux cas des poissons téléotéens, le microbiome de ces mammifères présentait une forte individualité, tant en termes de structure taxonomique que phylogénétique. J’ai également démontré que le microbiome de ces deux cétacés est taxonomiquement plus proche de celui d’autres animaux marins, y compris les poissons téléostéens, que d’autres mammifères terrestres, suggérant un impact fort du milieu marin sur le microbiome.

Enfin, je me suis intéressée plus spécifiquement à la variabilité interspécifique du microbiome chez 44 espèces de poissons téléostéens prélevées dans le lagon de Mayotte (Manuscrit D). J’ai démontré que chaque espèce présentait un microbiome cutané propre, qui ne s’expliquait pas par la phylogénie du poisson. J’ai également testé l’effet de 6 traits écologiques pouvant expliquer cette variabilité interspécifique (position dans la colonne d’eau, mobilité, régime alimentaire, taille, période d’activité, comportement d’agrégation). De manière surprenante, le seul trait expliquant les différences entre espèces était le régime alimentaire. Enfin, j’ai démontré que la richesse phylogénétique présente à la surface des poissons (i.e. le nombre de lignées phylogénétiques, voir chapitre 3) était plus élevée chez les espèces sensibles à la surpêche.

4.2 Manuscrit B
RESEARCH ARTICLE

High diversity of skin-associated bacterial communities of marine fishes is promoted by their high variability among body parts, individuals and species

Marlène Chiarello∗, Sébastien Villégé, Corinne Bouvier, Yvan Bettarel and Thierry Bouvier

UMR 9190 Marine Biodiversity, Exploitation and Conservation (MARBEC), Université de Montpellier, CNRS, IRD, IFREMER, Place Eugène Bataillon, Case 093, 34 095 Montpellier Cedex 5, France

∗Corresponding author: UMR 9190 Marine Biodiversity, Exploitation and Conservation (MARBEC), Université de Montpellier, CNRS, IRD, IFREMER, Place Eugène Bataillon, Case 093, 34 095 Montpellier Cedex 5, France. Tel: +334-67-14-40-93; E-mail: marlene.chiarello@univ-montp2.fr

One sentence summary: Skin-associated bacterial communities of marine fishes are different from surrounding bacterioplankton with an overall high diversity ascribed to its variability among body parts, individuals and species.

Editor: Julian Marchesi

ABSTRACT

Animal-associated microbiotas form complex communities, which are suspected to play crucial functions for their host fitness. However, the biodiversity of these communities, including their differences between host species and individuals, has been scarcely studied, especially in case of skin-associated communities. In addition, the intraindividual variability (i.e. between body parts) has never been assessed to date. The objective of this study was to characterize skin bacterial communities of two teleostean fish species, namely the European seabass (Dicentrarchus labrax) and gilthead seabream (Sparus aurata), using a high-throughput DNA sequencing method. In order to focus on intrinsic factors of host-associated bacterial community variability, individuals of the two species were raised in controlled conditions. Bacterial diversity was assessed using a set of four complementary indices, describing the taxonomic and phylogenetic facets of biodiversity and their respective composition (based on presence/absence data) and structure (based on species relative abundances) components. Variability of bacterial diversity was quantified at the interspecific, interindividual and intraindividual scales. We demonstrated that fish surfaces host highly diverse bacterial communities, whose composition was very different from that of surrounding bacterioplankton. This high total biodiversity of skin-associated communities was supported by the important variability, between host species, individuals and the different body parts (dorsal, anal, pectoral and caudal fins).

Keywords: Sparus aurata; Dicentrarchus labrax; skin microbiome; next generation sequencing; phylogenetic diversity
INTRODUCTION

All animals host at their surface and in several internal organs consortia of microorganisms, namely bacteria, archaea, fungi and viruses, collectively called microbiotas. These microbiotas form diversified communities and play critical roles for their host, as they facilitate nutrient absorption, regulate metabolism and defend against pathogen invasion (Sekirov et al. 2010).

Skin habitat is a unique interface, influenced both by surrounding environment (air, water, soil) and host-associated factors (health state, mobility, excretion of wastes and mucus, and immune molecules secretion). These interacting factors lead to a patchy physical and chemical environment at the surface of an individual and to contrasted environments between individuals (Shephard 1994; Grice and Segre 2011). Therefore, a variability of skin microbiome in terms of abundance and diversity is expected at both interindividual and intraindividual scales. Human skin microbiota has been particularly well studied, especially since the launch of the Human Microbiome Project in 2007 (Turnbaugh et al. 2007). These studies highlighted the high diversity of human skin microbiota (Schommer and Gallo 2013). They also evidenced that human skin-associated bacterial communities were highly variable between body parts, and between individuals (Fierer et al. 2010). These interindividual and intraindividual variations have been related to individual physiology (e.g. age, sex, health state, immune system), personal habits (e.g. hygiene, cosmetic use, clothing) and local-scale parameters (e.g. pH, temperature, humidity), even if the specific impact of each of these drivers, and the underlying interactions at a microbial scale were not systematically demonstrated (see Grice and Segre 2011 for a comprehensive review).

In contrast to human, skin microbiota of animals are yet still largely unknown. Among them, marine vertebrates, which represent more than 10 000 species on Earth (www.iobis.org), were only occasionally investigated during the last two decades (Larsen et al. 2013). In addition, most of the recent studies on marine vertebrates focused on the gastrointestinal microbiome (Mouchet et al. 2012; Xing et al. 2013), and revealed tight interactions between the host and its gut microbial communities (Pérez et al. 2010). Bacterial epibionts of marine vertebrates remain largely understudied, yet they are believed to play major roles in maintaining host health (Boutin et al. 2012). The few reports published to date found that the bacterial community composition was different among six Atlantic teleostean fish species, and highly different from that of surrounding planktonic communities (Larsen et al. 2013). Similarly, a recent study, focusing on wild humpback whale skin-associated bacterial communities, evidenced that despite individuals share a core set of species, bacterial community composition was variable between individuals because of differences in host physiology (Apprill et al. 2014). Moreover, while the entire fish body, including the head, trunk and also the fins, is recovered by the same integument, body parts of marine fishes may harbor contrasted local conditions due to (i) disparate epidermal mucus composition throughout body’s surface (Ángeles Esteban 2012), (ii) variable exposure to nutrient excretion fluxes through gills and vent, and (iii) variable water flow during swimming. These environmental variations at fish surface may drive variations of skin-associated bacterial communities between body parts. Such differences in skin microbial diversity between body parts have never been assessed to date on marine animals. Additionally, the only studies that assessed marine animals skin microbial diversity focused on wild individuals or fish kept in situ cages, making difficult to disentangle the effects of past and current environmental conditions experienced by the animals from their intrinsic characteristics (e.g. physiology, behavior) at the sampling time. Skin surface and hence bacterial epibionts are indeed directly exposed to the external biotic and abiotic components from the surrounding water column, while marine vertebrates are vertically and horizontally very mobile, which induces spatio-temporal variability (i.e. effects of seasonality and geographical location) in the composition of skin microbiota (Le Nguyen et al. 2008; Wilson, Danilowicz and Meijer 2008). To quantify the variability of this parameter among individuals and species, it is necessary to use animals raised in the same environment.

Another current gap in the description of biodiversity of skin microbial communities is the lack of simultaneous assessment of both taxonomic (i.e. based on species or OTUs) and phylogenetic (i.e. based on phylogenetic lineages) diversity facets (Escalas et al. 2013). Indeed, phylogenetic diversity has been proposed to be a better predictor of community functioning than taxonomic diversity because it accounts for complementarities among species (Zavarzin, Stackebrandt and Murray 1991; Fierer, Bradford and Jackson 2007). For instance, using marine bacterial species, Gravel et al. (2012) experimentally showed that the phylogenetic diversity of planktonic bacterial communities strongly explained the productivity of the community, suggesting functional complementarity of different phylogenetic lineages (even if functional conservatism along phylogenetic lineages is a debated issue; see Achenbach and Coates 2010; Xing et al. 2013). Moreover, communities composed of distantly related bacterial species stabilize community production when they are exposed to perturbations (Awasthi et al. 2014). Changes in the phylogenetic diversity of skin microbiome may therefore change its functions, and may thus disturb its homeostatic relations with the host and favor disease. The phylogenetic diversity should then be considered when assessing the level and variability of skin microbiome diversity. For example, two communities dominated by different OTUs, i.e. having a high taxonomic structural dissimilarity will have a low phylogenetic dissimilarity if abundant OTUs are phylogenetically close.

In addition, each diversity facet (taxonomic and phylogenetic) should be assessed accounting not only for composition (species presence/absence) but also for the structure of community by considering species relative abundances. Indeed, two communities can appear to be highly dissimilar in terms of phylogenetic composition (i.e. they host phylogenetically very distant species) only because of their rare species, and thus be similar in terms of phylogenetic structure (i.e. when taking account of species relative abundances) (Escalas et al. 2013).

In this study, we assessed the interspecific, intraspecific and intraindividual variability of the taxonomic and phylogenetic diversity of skin bacterial communities of two marine fish species, namely the European seabass (Dicentrarchus labrax) and gilthead seabream (Sparus aurata), bred in controlled environmental conditions. Our first objective was to determine whether bacterial diversity differed between the skin-associated bacterial communities and the surrounding bacteriaplankton. Our second aim was to test whether two fish species host different skin bacterial communities. Finally, we assessed the variability of bacterial diversity between individuals per fish species, and, within individuals, between different parts of the body (i.e. anal, caudal, dorsal and pectoral fins), and compared it with the interspecific difference. As mentioned above, an effect of host species has already been evidenced in other wild teleostean species. We expected that this important variability should persist in
controlled conditions between seabass and seabream-associated bacterial communities, because of intrinsic physiological differences between these species. Additionally, we expected a high level of intraindividual variability of fish skin bacterial associates, due to differences in habitats between the different body parts studied (Angeles Esteban 2012).

MATERIALS AND METHODS

Sampling

Four European seabass (D. labrax) and four giltthead seabreams (S. aurata) were sampled at the Marine Station of University of Montpellier (Sète, France). After larval stage, the two species were raised in the same conditions in two monospecific tanks (5 m³) for 2.5 and 7 years, respectively. The two tanks were connected to the same water filtration system (activated carbon filter, no sterilization) in a closed circulating water system, and tanks were regularly filtered with subsurface water of the Thau lagoon (renewal of 2% vol. per day). Physicochemical conditions were almost identical in the two tanks at the time of sampling (see data S1, Supporting Information). Individuals of the two species were fed with the same commercial pellets and received no anti- or probiotic treatment during their entire life. Individually tagged, suspended in air by the hook shaft, stunned and killed by cervical dislocation by a certified animal manipulator (following the European directive 2010/63/UE on the protection of animals used for scientific purposes). This protocol was chosen to avoid contacts between fish surface and other surfaces (tank wall, soil or hands of experimenters). Immediately after death, dorsal, caudal, left pectoral and anal fins were collected with ethanol-rinsed scissors and surgical pliers and stored in 100 mL of tank water were collected in each tank and filtered through a 47 mm 0.2 µm polycarbonate membrane (Whatman, Clifton, USA). The four filters were then placed in sterile cryotubes. All samples were snap-frozen at −196°C in liquid nitrogen, transported to the lab and stored at −80°C for 1 week before being analyzed.

DNA extraction, amplification and sequencing

Bacterial DNA recovery from fin surface was adapted from Amalfitano and Fazi (2008) for complex matrices. Each fin was immersed into 6 mL of a PBS solution containing 0.5% of tween 20 (vol/vol) and vortexed at maximum speed during 10 min (Vortex genie 2, Scientific Industries, Bohemia, USA). The solution was then filtered through a 47 mm 0.2 µm polycarbonate membrane (Whatman, Clifton, USA). Bacterial DNA was extracted by using the DNeasy® Blood & Tissue kit (Qiagen, Venlo, Netherlands), Bacterial DNA was extracted by using the Qubit dsDNA BR Assay kit (Invitrogen, Carlsbad, USA) and the manufacturer’s protocol facilitating lysis of Gram-positive bacteria. DNA was eluted in 100 µL of buffer AE and quantified by fluorescence using the Qubit dsDNA BR Assay (Invitrogen, Carlsbad, USA) and the manufacturer’s protocol facilitating lysis of Gram-positive bacteria. DNA quality was assessed by spectrophotometry (Nanodrop 1000, Wilmington, USA). Values of A260/280 and A260/230 nm averaged 2.3 ± 0.2 and 4.6 ± 0.5, respectively. All DNA samples were then diluted to 10 ng·µL⁻¹. An external laboratory (Research and Testing Laboratory, Lubbock, USA) performed PCR Amplification of the V1–V3 region of the 16S rRNA gene using universal bacterial primers 27F (5'-AGGTTTGTATCMTGGCTCAG-3') and 519R (5'-GTNTTACNCGGGCKGCTC-3') (Vergin et al. 1998; Ishah et al. 2013) and the HotStarTaq Plus Master Mix (Qiagen, Venlo, Netherlands) as follows: initial denaturation at 94°C for 3 min, followed by 28 cycles of 94°C for 30s, 53°C for 40s and 72°C for 1 min, ending with a final extension at 72°C for 5 min. After amplification, equimolar amounts of DNA were mixed, purified (Ampure beads kit, Agencourt Bioscience Corporation, USA) and sequenced using a Roche 454 FLX titanium pyrosequencer. We obtained a total of 104 548 (~200 bp) reads from the sequencing of PCR amplicons from the 36 samples. The nucleotide sequence data reported are available in the NCBI SRA database under the accession number SRP050454.

Sequence processing and taxonomic classification

Sequences were processed following the SOP analysis pipeline of Schloss, Gevers and Westcott (2011; http://www.mothur.org/wiki/454_SOP, 02/2014) using Mothur (Schloss et al. 2009). Briefly, anormal flows (homopolymers >8bp, >1 mismatch to the barcode, or >2 mismatches to the primer), and very short flows (<200 bp) were discarded. Then, sequences were determined using a maximum-likelihood approach using PyroNoise (Quince et al. 2011). Chimera were detected and eliminated using UCHIME (Edgar et al. 2011). Up to 71 744 unique sequences with an average length of 244 bp were retained. Sequences presenting more than 97% identity were clustered, and a representative sequence (i.e. the closest sequence of all other sequences) for each cluster was selected. Using these sequences, clusters were classified using the Ribosomal Database Project II Classifier (Wang et al. 2007). Non-prokaryotes and mitochondrial clusters were excluded. The number of sequences varied between samples (data S2, Supporting Information) and these differences may not reflect true difference in richness and biomass of bacterial communities but rather difference in sampling effort (e.g. mucus volume) and/or efficiency of amplification and sequencing. To correct for this uneven number of sequences, we calculated taxonomic and phylogenetic diversities on bootstrapped samples (Bryant et al. 2008). More precisely, we considered 1000 randomized subsamples of 113 sequences (the minimal number of sequences among the 36 samples) for each community. We then only considered the mean of diversity indices among the 1000 bootstrapped samples because their variances were negligible.

Phylogenetic analyses

All representative sequences were aligned using MAFFT v7 (FFT-NS2) (Katoh et al. 2002) and a phylogenetic tree was reconstructed using FastTree 2 (Price, Dehal and Arkin 2010), implemented in QIME software (Caporaso et al. 2010). The tree was rooted using a set of eight archaeal 16S rRNA gene sequences obtained from SILVA database (Quast et al. 2013). A chronogram was then adjusted on the phylogenetic tree using the ‘chronos’ function (discrete model, 20 evolution rates) provided in the R-package ape (Paradis, Claude and Strimmer 2004). This function provides a dated ultrametric tree using a maximum-likelihood algorithm and calibration points, provided in data S3 (Supporting Information).

Alpha diversity computation

Alpha diversity was described using a set of four complementary indices, describing taxonomic and phylogenetic compositional diversity (i.e. taxonomic and phylogenetic richness based,
respectively, on presence/absence of OTUs and phylogenetic lineages), and taxonomic and phylogenetic structural diversity (i.e. taking account of relative abundances of OTUs and phylogenetic lineages, respectively).

Taxonomic richness (S) was assessed as the number of different OTUs in each community. Phylogenetic richness (Faith’s PD), based on the sum of branch lengths of the phylogenetic tree grouping OTUs present in the sample, was calculated using the Picante R-package (Kembel et al. 2019). Taxonomic structural diversity was assessed using Shannon alpha diversity (Shannon 1948). Phylogenetic structural diversity was assessed using Allen alpha diversity (Allen, Kon and Bar-Yam 2009). Allen index of diversity is similar to the Shannon diversity, excepted that it is based on phylogenetic branch lengths instead of OTUs. Allen index was calculated using the “ChaoFD” function of entropart package (Marron and Hérault 2014). These two indices were expressed in equivalent number of species, as recommended by Jost (2007). This transformation allows direct comparisons between diversity values (Chao, Chiu and Jost 2014).

Beta diversity computation

Alpha diversity indices describe diversity at a local scale. To fully assess bacterial diversity, it is also necessary to measure beta diversity, i.e. the dissimilarity between communities. Similarly to alpha diversity computation, dissimilarity was assessed using a set of four indices describing each facet (phylogenomic and taxonomic) and component (compositional or structural) of diversity.

Compositional (i.e. based on presence/absence matrices) taxonomic and phylogenetic beta diversities were assessed by the Sorensen (Sørensen 1948; Koleff, Gaston and Lennon 2003) and phyloSor dissimilarity indices (Bryant et al. 2008; Leprieur et al. 2012), respectively, using the betapart R-package (Baselga and Orme 2012). PhyloSor is similar to the Sorensen index, excepted that it is calculated on branch lengths. These two beta diversity measurements are scaled between 0 (when communities share the same OTUs or phylogenetic lineages) and 1 (when communities have no OTU or phylogenetic lineages in common).

Structural (i.e. accounting for entities relative abundances) taxonomic and phylogenetic beta diversities were calculated using the multiplicative decomposition of Shannon and Allen indices, respectively, following the general framework proposed by Chao, Chiu and Jost (2014). These two beta diversity measures were scaled between 0 (when, in case of taxonomic beta diversity, communities share the same OTUs at the same abundances) and 1 (when communities have no OTU in common) as suggested by Villéger et al. (2012), and were therefore directly comparable to Sorensen and PhyloSor indices (Chao, Chiu and Jost 2014). These four beta diversity indices were calculated at the intra-individual and inter-individual, and interspecies scale.

Statistical analyses

Phylogenetic and taxonomic richness tend to be correlated as the increasing the number of OTUs increase the probability of covering more phylogenetic lineages. Consequently, we computed the standardized effect size of the PD index (SES.PD) comparing the observed PD value and its expected value under a null model maintaining sample species richness, using the SES.PD function of the Picante package (Kembel 2009). A positive/negative SES.PD value indicates a phylogenetic over/underdispersion, i.e. OTUs found in the sample are more/less phylogenetically distant than expected.

RESULTS

Alpha diversity

Alpha diversity in water and on fish skin
Alpha diversity patterns differed, depending on the facet (taxonomic or phylogenetic) and component (composition or structure) considered. Indeed, the taxonomic richness (related to species composition) was significantly higher in each water replicate (averaging ca. 46 OTUs ± 2.3 after bootstrap subsampling, n = 4 water replicates) than on fish skin (ca. 22 OTUs ± 7.2 per bootstrapped sample, n = 32; but note there were 73 ± 5.4 OTUs per individual) (Kruskal–Wallis test, KW, P < 0.05, Fig. 1a). On the contrary, taxonomic alpha diversity (accounting for relative abundances of OTUs) was not significantly different between the two types of samples (Shannon alpha diversity, KW P > 0.05, Fig. 1c). The high taxonomic richness of planktonic communities was indeed mainly due to rare OTUs (OTUs accounting for <1% of total abundance), representing about 89.9 ± 0.8% of the OTUs present in water. On the contrary, OTU abundances were more evenly distributed in skin communities, with fewer rare OTUs (4.0 ± 7.3% of present OTUs). Among all identified OTUs, only 7% were found in water replicates, while all of them were found in at least one fish sample.

Additionally, while phylogenetic richness (PD) did not significantly differ between water and fish (KW, P > 0.05, Fig. 1b), phylogenetic alpha diversity (Allen alpha diversity, based on the relative abundances of phylogenetic lineages) was significantly lower in water than in fish skin samples (KW, P < 0.05, Fig. 1d). Besides, planktonic communities were significantly phylogenetically underdispersed, i.e. the OTUs forming the community were clustered on the phylogenetic tree (SES.PD = −2.95 ± 1.8, P = 0.02 ± 0.01), while fish communities were neither underdispersed nor overdispersed (SES.PD = −0.36 ± 0.8, P = 0.39 ± 0.2).

Alpha diversity patterns in fish skin samples

For all facets and components of the alpha diversity, there was no significant difference between (i) the two fish species (interspecific), (ii) the individuals of each species (interindividual) and (iii) body parts (intraisindividual) (KW, P > 0.05, Fig. 1).

Beta diversity

Beta diversity between planktonic replicates
Taxonomic composition of planktonic communities presented an important level of variability between water replicates, as shown by Sørensen’s dissimilarity index averaging 0.65 ± 0.01, n = 4 (Fig. 2). When considering phylogenetic proximity between OTUs, dissimilarity dropped by 30% (phyloSor dissimilarity index, 0.46 ± 0.01). Values of dissimilarity taking
account of relative abundances of OTUs (Shannon beta diversity) or that of phylogenetic lineages (Allen beta diversity) decreased (0.27 ± 0.01 and 0.06 ± 0.003, respectively). This indicated a strong homogeneity of water replicates when taking into account abundant OTUs and their phylogenetic relatedness.

**Beta diversity between water and fish skin**

All values of dissimilarity between water and fish skin were high, for each facet and component of diversity considered. Taxonomic compositional and structural beta diversity was almost maximal (Sørensen’s dissimilarity index 0.97 ± 0.03; Shannon beta diversity 0.94 ± 0.07) (data S4, Supporting Information Fig. 2). Bacterial dissimilarity between those two habitats decreased by 31 and 45% when we accounted phylogenetic distance alone or associated with OTUs relative abundances (respectively, for phylSor 0.66 ± 0.05 and Allen beta diversity 0.49 ± 0.1). Nevertheless, whatever the facets and components of diversity considered, the planktonic and fish skin-associated bacterial communities differed significantly (Table 1, PERMANOVA, P < 0.05).

**Beta diversity of skin-associated bacterial community within each species**

Interindividual and intraindividual dissimilarity values were particularly high, as taxonomic and phylogenetic compositional and structural variability of skin communities for the two fish species was comparable to that observed at the interspecific scale (S4, Supporting Information Fig. 2). However, despite these high differences among individuals and body parts, they were not higher than expected in a null model with hypothesis of no effects of host species effect (PERMANOVA, P < 0.05).

**Beta diversity of skin-associated bacterial communities between fish species**

Interspecific dissimilarity was high in terms of taxonomic composition and structure, averaging, respectively, 0.81 ± 0.03 (Sørensen dissimilarity, S4, Supporting Information) and 0.69 ± 0.11 (Shannon beta diversity, Fig. 2). However, when considering phylogenetic relationships between lineages, interspecific compositional and structural dissimilarity values halved (phylSor, 0.43 ± 0.07, S4, and Allen beta diversity, 0.27 ± 0.07, Fig. 2) compared respectively to Sørensen and to Shannon beta diversity. Skin-associated bacterial communities are thus phylogenetically more similar than taxonomically. However, these relatively low interspecific dissimilarities compared to the taxonomic ones were still higher than expected in a null expectation model with hypothesis of no effects of host species effect (Table 1, PERMANOVA, P < 0.05).

---

**Figure 1.** Richness and alpha diversity values of bacterial communities, at all scales (interspecific, inter- and intraindividual) studied. Points and error bars indicate respectively mean and confidence interval limits (5th and 95th centiles of values obtained from the 1000 bootstrapped subsamples) of diversity indices. To facilitate graphs interpretation, phylogenetic richness (PD, graph (b)) was respectively scaled to the total branch lengths of the phylogenetic tree. (c) Taxonomic structural alpha diversity for each bacterial community, calculated on relative abundance of OTUs using Shannon index. (d) Phylogenetic structural alpha diversity, calculated on relative abundances of terminal branches of the chronogram. On each graph, an asterisk indicates a significant difference of richness and alpha diversity values between water and fin samples (KW, P < 0.05).
predictable by individuals nor by body parts. However, in each individual, skin-associated bacterial communities differed between fins, as some OTUs were unique to certain fin samples (Fig. 3).

**Dominant phylogenetic groups**

Planktonic communities were essentially dominated by the phyla Proteobacteria (38–54% of sequences obtained from water samples) and Bacteroidetes (41–52% of sequences) (Fig. 4). Skin communities were mainly composed of members of the phyla Proteobacteria (30–85% of sequences obtained from fin samples), Actinobacteria (2–53%), Bacteroidetes (0.4–27%) and Firmicutes (0.4–12%) (Fig. 4). Actinobacteria and Firmicutes lineages were not detected in the planktonic communities. At finer taxonomic levels, there were also disparities between epibiotic and planktonic communities. Planktonic Proteobacteria were mainly composed of Gammaproteobacteria and Alphaproteobacteria, whereas skin-associated Proteobacteria amount to a large fraction of Betaproteobacteria. The same disparity was observed for the Bacteroidetes, mostly comprised of Flavobacteria in planktonic communities, and of Sphingobacteria and Flavobacteria in skin-associated communities. The two fish species were dominated by the same major bacterial clades (Fig. 4), and the relative abundances of these dominant clades did not significantly differ between the two species (KW, \( P < 0.05 \)).

**DISCUSSION**

Seabass and seabream skin harbored more diverse bacterial communities than water

Skin-associated bacterial communities were particularly diverse, since as many as ca. 73 (±5.4) OTUs were detected on each individual, while only ca. 46 (±2.3) OTUs were detected in 100 mL of water (Fig. 1). Additionally, OTU abundances were particularly uneven in planktonic communities, with few dominant and a great number of rare ones. Such an uneven distribution of OTU abundances in seawater has been previously reported in the Mediterranean sea by a study using amplicon-based sequencing (Pommier et al. 2010). As expected, planktonic communities were composed of typical marine classes such as the Alphaproteobacteria and Gammaproteobacteria and Flavobacteria (Barbenn and Casamayor 2010) (Fig. 4). They likely originated from the marine water that was used to fill up the experimental tanks. In contrast, bacterial OTUs of skin samples had much more abundances distribution than the planktonic communities, and when considering whole individuals, phylogenetic diversity of skin-associated bacterial communities was
50% more diverse than the one of bacterioplankton (Fig. 4). Even distribution of OTU abundances and high diversity of phylogenetic lineages (Fig 1) of skin-associated bacterial communities could be related to the particular nutritive conditions existing at the fish surface. Most of teleostean fishes secrete mucus, which is constituted by a high diversity of gel-forming glycoproteins, glycosaminoglycans and proteins (Shephard 1994). Such components can serve as nutrient sources for epibiotic bacteria (Bordas et al. 1998), which thus provide a mix of different resource niches. Environmental complexity and resource partitioning has been shown to favor rich bacterial diversity (Ramette and Tiedje 2007; Schauer et al. 2009). By contrast, in a closed water-circulated system, particles are trapped by a filtration system, potentially inducing particularly homogeneous environmental conditions. In such situation, water may be considered as a desert from a nutritional viewpoint as it provides less nutrients and less nutrient types (Azam and Malfatti 2007). This may explain the significant phylogenetic clustering observed in water samples. The phylogenetic diversity of skin-associated bacterial communities recorded in our study is greater than that had been evidenced by rDNA 16S cloning and sequencing approach on six fish species of the Atlantic ocean, where a total of only five different phyla were identified (while as many as 15 phyla were detected on fishes in our study) (Larsen et al. 2013). Such difference could be of methodological order, as clone libraries are known to underestimate bacterial richness (e.g. comparative study on ant microbiome (Kautz et al. 2013).

Beyond their high diversity, skin bacterial communities exhibited a very different composition from that of their planktonic counterparts, as indicated by the low number of common OTUs (21 OTUs i.e. 3% of all detected OTUs, Fig. 3) and high dissimilarity values (S4, Supporting Information Fig. 2) between these two types of habitats, whatever the facets and components of diversity considered. This agrees with recent investigations on wild or outdoor farmed teleostean fishes (Wang et al. 2010; Larsen et al. 2013), although such studies did not permit disentangling intrinsic drivers of microbiome diversity from confounding extrinsic factors (e.g. fish that experienced different water masses due to their mobility and/or change in water masses). Here fishes were raised in controlled conditions during their entire lives, which likely greatly minimizes such artifacts. This line of evidence from different fish species and life histories highlights that the specificity of skin biotope promotes a specific signature of skin bacterial community compared to planktonic cells.

**Phylogenetic homogeneity of skin-associated bacterial communities**

The diversity of bacterial epibionts was highly variable across all scales (i.e. interspecific, interindividual and intraindividual, S4, Supporting Information Fig. 2). However, among the 15 phyla detected on fish skin, more than 95% of OTUs belonged to only four clades, namely the phyla Actinobacteria and Firmicutes and the classes Sphingobacteria and α-proteobacteria (Fig. 4). This dominance of a core set of phyla drove the low values of phylogenetic structural and compositional dissimilarity (i.e. phyloSor and Allen indices, S4, Supporting Information and Fig. 2) compared to taxonomic structural and compositional dissimilarity values. This detected skin bacterial clades are similar to those identified in the gut microbiota of wild and reared *S. aurata* and *D. labrax*, but harbored different genera (Carda-Díéguez, Mira and Fouz 2014; Kormas et al. 2014), potentially depicting a specific character of the skin habitat. Interestingly, these skin-associated clades have already been reported in skin microbiota of other vertebrates as teleosts (Wang et al. 2010; Larsen et al. 2013), marine mammals (Apprill et al. 2014), amphibians (Walke et al. 2014) and human (Grice and Segre 2011). Moreover, a recent review focusing on marine macroscopic algae and invertebrates surfaces also revealed the same core of high-level bacterial clades (Wahl et al. 2012). To explore such apparent conservatism, we compared the representative sequence of predominant OTUs belonging to each previously cited core phyla using BLASTn (http://blast.ncbi.nlm.nih.gov/) (Fig. 5). We observed joint OTUs with other marine vertebrates associated microbiota, water and soil bacterial communities, but the most unforeseen result was that as high as 40% of these fish skin-specific sequences were highly similar (99–100% identity, 97–100% coverage) with sequences of OTUs previously isolated from healthy human skin surface. These OTUs were identified as

![Figure 3. Venn diagrams of detected OTUs (a) on skin surface of the two fish species and in water samples, and (b) on each fish fin type, summed with all individuals of the two fish species. Venn diagram (a) is scaled to actual numbers of OTUs and overlaps. Venn diagram (b) is not proportional to numbers of OTUs and overlaps.](http://femsec.oxfordjournals.org/)
Propionibacterium sp. (representing 1.7–28.8% of sequences in fish samples), followed by Corynebacterium sp. (0–9.7%), Ochrobactrum sp. (0–9.3%), Geobacillus sp. (0–8.9%) and Staphylococcus sp. (0–7.0%). This similarity between human and fish skin microbiomes was unexpected, as employees of the breeding station were systematically wearing latex gloves for the maintenance of the aquaculture system. In addition, all OTUs belonging to these core clades and recorded in human skin were not detected in water (either because they were absent or sufficiently rare and below the detection threshold, Fig. 5). Shared bacterial species between fish and human is surprising but supports the hypothesis that some biotic surface characteristics, which do not exist in the aqueous surrounding environments, may drive skin-associated community structure and that these clades may present phylogenetically conserved traits permitting their growth on living surface. Here, further studies are needed to determine whether such similarities at skin surface exist in other animal clades, or if it is a characteristic of teleostean fishes or a result of our particular experimental design. Moreover, for an exhaustive assessment of microbial diversity, such analysis should be extended to viruses, Archaea and microeukaryotes, as that has been done with human skin microbiome (excepted in the case of viruses) (Findley et al. 2013; Probst, Auerbach and Moissl-Eichinger 2013).

**Bacterial skin-associated community structure is dependent on fish species**

OTU composition in skin-associated bacterial communities differed strongly between the two fish species, as around 70% of OTUs detected at the surface of each species were not detected in the other species (Fig. 3). The high taxonomic structural beta diversity (ca. 70%), which is more informative because it takes into account the relative abundances of OTUs, indicated that among shared species, the bacterial associates of the two fish species were dominated by different OTUs. However, the drop of beta diversity values when taking account for phylogenetic affiliation of OTUs (phylogenetic structural beta diversity ca. 30%, Fig 2) demonstrated that, while both species were dominated by different OTUs, these OTUs belonged to phylogenetically close bacterial clades. In addition, this moderate level of phylogenetic structural beta diversity was marked by a significant interspecific difference (Table 1), suggesting that species host close but distinct phylogenetic lineages. Such lines of evidence for host species specificity were recently observed on six teleostean species by Larsen et al. (2013). In this study, authors suggested that several physiological species-specific factors (e.g. skin mucus composition, antimicrobial properties) could lead to such pattern. However, as their observations were made from
wild animals, it is not easy to partition these effects from others, related to species ecology (i.e. mobility, food, contact with biotic and abiotic surfaces), and environmental variations (i.e. food availability, physicochemical conditions of surrounding water masses). Here, hosts were raised under the same conditions (e.g. same food, no antibiotic treatments, same physicochemical conditions). We therefore can assert that this host species specificity is primarily due to intrinsic physiologic factors.

Larsen et al. (2013) suggested that skin microbiota selection from host species could be considered (i) as an active selection by the host, or/and (ii) as a passive selection of bacterial species able to grow on fish surface. Indeed, while general mechanisms of innate immune system are largely conserved among fishes and other vertebrates, the species-related variability of skin immune components has been proven for fish (Ángeles Esteban 2012). For example, lysozyme activity, as well as the nature of antimicrobial peptides secreted in the mucus layer, was found to be different between several freshwater fish species (Ángeles Esteban 2012; Nigam et al. 2013).

In addition, the structure of bacterial communities could be affected by their specific capacities to adhere and grow on skin mucus, as on a culture medium. Fish skin mucus remains poorly documented, yet some species-related variation in mucus characteristics has been reported, such as mucin composition and hydration (i.e. mucin concentration) (Roberts and Powell 2005). Variations in mucin hydration induce changes in its viscoelasticity, which may influence bacterial attachment (Ángeles Esteban 2012). However, no study to date compared the skin mucus chemical composition between gilthead seabream and European seabass.

**Skin bacterial communities are variable among and within individuals**

**Among individuals**

The high level of interindividual variability observed (Fig. 2) has already been reported between individuals of wild or in situ captive teleosts and cetaceans, and was partly explained by the geographical location (Le Nguyen et al. 2008; Larsen et al. 2013; Apprill et al. 2014). In our study, however, and as discussed before, animals of each species were raised rigorously in the same conditions from birth to sampling, which precludes any extrinsic environmental influence on bacterial skinfish associates. This high level of variability between individuals is therefore an intrinsic feature of the studied fish skin bacterial community. Consequently, the effect of geographical location reported above may indeed exceed the intrinsic variability.

**Within individuals**

The bacterial community composition and structure differed between body parts, as shown by the high values of intraspecific beta diversity (S4, Supporting Information and Fig. 2), which was as high as the one between the two fish species. Indeed, 53–61% of OTUs detected on each fin were not detected on other ones (Fig. 3), suggesting niche specialization within individuals. However, this intraspecific variability was not explained by fin type when considering the four individuals in each species because of the large interindividual variability (Table 1, Fig. 2). In other words, there is always a difference in bacterial community composition and structure between the four fins in each individual, but when comparing two individuals, a single fin type host different communities. This is in contradiction with studies on the human skin microbiome, which reported that bacterial communities were primarily shaped by skin parts rather than individuals or time (Costello et al. 2009; Grice et al. 2009).

Here the absence of a common community composition pattern between fin types or individuals, coupled with a huge variability at these two scales, suggests that (i) the body parts studied and individuals did not particularly differ in terms of habitat quality, and therefore (ii) bacterial communities composition is unpredictable, either because being dependent of stochastic events of bacterial colonization and extinction or dependent on a large number of interacting factors (e.g. local release of nutrients or antimicrobial molecules). This therefore underlines the need to explore surface-associated communities variability across the entire body, and especially around oral gape, gills and lateral line, which may harbor more contrasted microenvironmental conditions. Further studies are also needed to determine if such pattern exists in other marine or terrestrial vertebrates.

Another possibility is that skin microbiome may temporarily vary independently in each individual and body parts. Indeed, a strong temporal variability of skin-associated bacterial communities has been recently reported in human (Costello et al. 2009). In the case of marine teleosts, studies about temporal dynamics of skin microbiota are very scarce. While a few studies evidenced changes in the composition of cutaneous bacterial communities due to seasonality and diet changes (Larsen et al. 2013; [110x582 to 483x719])

![Figure 5. Results of GenBank search upon the 50 most abundant OTUs of main phyla recovered in skin samples. For each OTU, the representative sequence was compared to Genbank sequences using BLASTn. For each sequence, the best match (≥97% coverage, ≥99% identity) was selected, and the origin of the samples was noted. OTUs with no good matches were grouped in the “Unisolated” class.](http://femsec.oxfordjournals.org/)
Landeira-Dabarca, Sieiro and Álvarez 2013), an intrinsic temporal dynamic has never been assessed. Assessing the intraindividual variability of the microbiome for additional fish species and for other aquatic and terrestrial animals is therefore needed to confirm our findings and to identify the drivers of this intraindividual variability.

To conclude, our study demonstrated that fish surface is colonized by a skin-specialized community of bacteria, composed by clades that are not detected in the overlying water. Skin associates were characterized by a high diversity, which seems to be promoted by their important variability between species, individuals and body parts. As diversity is generally positively correlated to microbial communities functioning and stability (Whiteboile et al. 2009; De Schryver and Vadstein 2014), and negatively correlated with susceptibility to invaders (De Roy et al. 2013; De Schryver and Vadstein 2014), this unique biodiversity may favor fish resistance to pathogen invasion through the skin (Wang et al. 2010). Testing the relationship between microbial diversity and infection by pathogens becomes therefore an urgent challenge.

SUPPLEMENTARY DATA

Supplementary data are available at FEMSEC online.

ACKNOWLEDGEMENTS

We thank C. Amiel and G. Sposito from the SMEL station, Sète, for providing fishes and for their help during the sampling process. We thank D. Kalenitchenko and P.E. Galand for their help with genetic data analysis, and L. Dejouy and F. Rieuville for their help during the sampling process. We are grateful to three anonymous reviewers and the editor for their comments that helped us to improve our article.

Conflict of interest. None declared.

REFERENCES

Achenbach LA, Coates JD. Disparity between bacterial phylogeny and physiology—comparing 16S rRNA sequences to assess relationships can be a powerful tool, but its limitations need to be considered. ASM News 2000;66:714-5.
Ishak HD, Plowes R, Sen R, et al. Bacterial diversity in solenopsis invicta and solenopsis geminata ant colonies...
Landeira-Dabarca A, Sieiro C, Álvarez M. Change in food inges-
Leprieur F, Albouy C, DeBortoli J, et al. Quantifying phyloge-
Marcon E, Hérault B. entreptor: An R Package to Measure and Par-
Maccel L, Rodgers P. eulerAPE: drawing area-proportional 3-
Mouchet MA, Bouvier C, Bouvier T, et al. Genetic difference but functional similarity among fish gut bacterial communities through molecular and biochemical fingerprints. FEMS Microb-
Pérez T, Balcázar JL, Ruiz-Zarzauela I, et al. Host–microbiota inter-
Roberts SD, Powell MD. The viscosity and glycopolypeptide biochem-
Schauer R, Bienhold C, Ramette A, et al. Bacterial diversity and biogeography in deep-sea surface sediments of the South At-
tlantic Ocean. ISME J 2009;4:159–70.
Schloss PD, Westcott SL, Ryabin T, et al. Introducing mothur: open-source, platform-independent, community-supported software for describing and comparing microbial communi-
Sørensen TJ. A Method of Establishing Groups of Equal Amplitude in Plant Sociology Based on Similarity of Species Content and Its Application to Analyses of the Vegetation on Danish Commons. København: I kommission hos E. Munksgaard, 1948.
Wang W, Zhou Z, He S, et al. Identification of the adherent micro-
bacteria on the gills and skin of poly-cultured gibel carp (Carassius auratus gibelio) and bluntnose black bream (Megalobrama amblycephala Yih). Aquat Res 2010;41:872–83.
Wellington EM, Berry A, Krssek M. Resolving functional diversity in relation to microbial community structure in soil: exploit-


Captive bottlenose dolphins and killer whales harbor a species-specific skin microbiota that varies among individuals

Chiarello M.*, Villéger S.*, Bouvier C.*, Auguet JC.*, and Bouvier T.*

1 Marine Biodiversity, Exploitation and Conservation (MARBEC), Université de Montpellier, CNRS, IRD, IFREMER, Place Eugène Bataillon, Case 093, 34 095 Montpellier Cedex 5, France

En révision chez Scientific Reports

Des informations supplémentaires complètent ce manuscrit, disponibles en section 7.4
ABSTRACT

Marine animals surfaces host diverse microbial communities, which play major roles for host’s health. Most inventories of marine animal surface microbiota have focused on corals and fishes, while cetaceans remain overlooked. The few studies focused on wild cetaceans, making difficult to distinguish intrinsic inter- and/or intraspecific variability in skin microbiota from environmental effects. We used high-throughput sequencing to assess the skin microbiota from 4 body zones of 8 bottlenose dolphins (Tursiops truncatus) and killer whales (Orcinus orca), housed in captivity (Marineland park, France). Overall, cetacean skin microbiota is more diverse than planktonic communities and is dominated by different phylogenetic lineages and functions. In addition, the two cetacean species host different skin microbiotas. Within each species, variability was higher between individuals than between body parts, suggesting a high individuality of cetacean skin microbiota. Overall, the skin microbiota of the assessed cetaceans related more to the humpback whale and fishes’ than to microbiotas of terrestrial mammals.
INTRODUCTION

Marine animals’ surfaces are associated with highly diverse microbial communities, which play major roles for their health, including protection against macrofouling, and pathogens (Wahl et al. 2012; Bourne, Morrow, and Webster 2016). These surface microbiota were shown to be both distinct from surrounding planktonic samples (Wahl et al. 2012), and host-species specific (Bourne, Morrow, and Webster 2016), suggesting that they could have coevolved with their animal hosts (McFall-Ngai et al. 2013). In addition, marine animal surface microbiota are dynamic assemblages (Glasl, Herndl, and Frade 2016), with composition of microbial Operational Taxonomic Units (OTUs) as well as their relative abundance varying between host life stages (Lema, Bourne, and Willis 2014), surrounding environmental conditions (Zaneveld et al. 2016) and geographical location (Salerno, Bowen, and Rappé 2016). However, most of these findings have been reported from marine invertebrates, and especially corals. Whether these observations could be generalized to marine vertebrates, which constitute the most important biomass fraction of macroorganisms in the global ocean, is barely unknown (but see recent work on fishes (Larsen et al. 2013; Chiarello et al. 2015) and whales (Apprill et al. 2014)). Among marine vertebrates, mammals are represented by more than 100 species belonging to three clades (pinnipeds, cetaceans and sirenians) which respective ancestors were terrestrial. Marine mammals hence have biological features, including skin structure, similar to terrestrial mammals. Therefore, assessing the composition of skin microbiota of marine mammals could shed light on the importance of evolutionary legacies and adaptation to marine environment in shaping skin microbiota of animals.

The only marine mammal skin microbiota described to date is the one of the free-ranging humpback whale from the North Pacific. Apprill et al. (Apprill et al. 2011, 2014) showed that individuals share a core skin microbiota and that variability in taxonomic and phylogenetic diversity of skin microbiota among individuals is driven by geographical location and the health state of the whale. However, such studies on wild animals do not allow disentangling individual-driven variation of skin microbiota from the effect of environmental conditions. Animals housed in controlled environment offer the opportunity to measure the interspecific and inter-individual variability of animals skin microbiota independently from environmental variability, and to assess the intra-individual variability of their microbiota (Bik et al. 2016).
Besides assessing the taxonomic and phylogenetic facets of skin microbiota, describing its functional role is fundamental to understand the link between microbiota and host health. Indeed, skin is the first line of defense from pathogen infections in mammals with skin microbiota closely interacting with its host cells from the epidermis to the deep dermis (Nakatsuji et al. 2013), to modulate immunity (Wang et al. 2016; Belkaid and Tamoutounour 2016), and support antagonistic effects against pathogens (Cogen et al. 2010). However the functional diversity of the skin microbiota of marine mammals has never been assessed, as well as its congruence with its phylogenetic diversity has never been assessed (Nelson et al. 2015).

Recent advances in bioinformatics (e.g. PICRUSt (Langille et al. 2013)) allow predicting metagenome functional content from 16S rDNA data and hence to assess simultaneously the taxonomic, phylogenetic and potential functional diversities of microbial communities.

Here, using high-throughput sequencing, we assessed the taxonomic and phylogenetic diversities of the skin microbiota from 4 body zones (i.e. the dorsal, anal and pectoral fins, and its anal zone) of 8 individuals of two emblematic Odontoceti (toothed whales) species, the bottlenose dolphin (*Tursiops truncatus*) and the killer whale (*Orcinus orca*), housed in controlled conditions. We also predicted the functional facet of microbiota diversity using PICRUST software. We first measured the similarity between the microbiota of the two species. Second, we quantified the magnitude of intraspecific variability of microbiota, *i.e.* between individuals of each species and between their body parts. Third, we analyzed the similarity between the skin microbiota of cetaceans and the one of terrestrial mammals and non-mammal vertebrates.

**MATERIAL AND METHODS**

**Sampling of skin and planktonic microbiotas**

We sampled skin microbiota of four killer whales (*Orcinus orca*) and four bottlenose dolphins (*Tursiops truncatus*) housed at Marineland (Antibes, France) in accordance with European laws (Directive EC 1999/22 and EU CITES 338/97). Animals were manipulated by their caretakers, in accordance with internal practices of the park. Sampling was done using a non-invasive method (swabbing a small surface for 1 minute). All manipulations were approved by Marineland’s scientific committee.

Killer whales and dolphins were aged from 13 to more than 30 years at the time of sampling.
Contrary to dolphins, killer whales were affiliated, with the younger ones being siblings or half-siblings, and the older one (Freya) being the mother of the older male (Valentin). All animals but one (i.e. Valentin which received an antifungal treatment that ended two weeks before the day of sampling) did not receive any antibiotics during the 6 months before sampling. Individuals of the two species were kept in two separated pools, which are filled by the same seawater circulation system. Seawater is pumped from 600-meters offshore and 68-meters deep in the Mediterranean Sea, and filtered through sand. Water flux is set so that the water of each pool is renewed every 2 hours.

**Table 1: Animals included in this study.** The age of each animal at sampling time is indicated, as well as their kinship, when known. Animals older than 30 years old were captured from the wild during the early 1980s; therefore their exact age is unknown.

<table>
<thead>
<tr>
<th>Species</th>
<th>Individual</th>
<th>Age (years)</th>
<th>Sex</th>
<th>Complementary Information</th>
</tr>
</thead>
<tbody>
<tr>
<td>K. whales</td>
<td>Freya</td>
<td>&gt;30</td>
<td>Female</td>
<td>Valentin’s mother</td>
</tr>
<tr>
<td></td>
<td>Valentin</td>
<td>18</td>
<td>Male</td>
<td>Antifungal treatment ended 15 days before sampling</td>
</tr>
<tr>
<td></td>
<td>Wikie</td>
<td>13</td>
<td>Female</td>
<td>Valentin’s half sister, sister of Inouk</td>
</tr>
<tr>
<td></td>
<td>Inouk</td>
<td>15</td>
<td>Male</td>
<td>Valentin’s half brother, brother of Wikie</td>
</tr>
<tr>
<td></td>
<td>Sharki</td>
<td>&gt;30</td>
<td>Female</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Lotty</td>
<td>&gt;30</td>
<td>Female</td>
<td></td>
</tr>
<tr>
<td>Dolphins</td>
<td>Dam</td>
<td>17</td>
<td>Male</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Rocky</td>
<td>15</td>
<td>Male</td>
<td></td>
</tr>
</tbody>
</table>

The day of sampling, each animal was asked by its caretaker to raise successively 4 body zones (i.e. the dorsal, caudal and pectoral fins, and anal zone) outside of water. These four zones could be considered as distinct patches for microbiotas (i.e. distant to each other by > 30 cm) and experience different micro-environmental conditions (e.g. the anal zone because of release of feces and urine). After briefly rinsing the skin using 100-mL autoclaved
seawater, skin microbiota was sampled by swabbing a 63-cm² circular surface using sterile foam-tipped applicators from Whatman (GE Healthcare) during 30 seconds on each side of the swab. For the caudal and pectoral fins, only the upper side of the fin was sampled. We then cut the tip of the swab using ethanol-rinsed scissors and placed the sponge part of the swab into sterile cryotubes.

For each species, three 100-mL pool and input water (i.e. exit of pipe from filtering system) samples were collected and filtrated through a 47 mm diameter, 0.2 µm pore size, polycarbonate membrane (Whatman, Clifton, USA). The membranes were then placed in sterile cryotubes. All samples were immediately snap-frozen in liquid nitrogen, transported to the laboratory and stored at -80°C before DNA extraction.

**16S rDNA amplification and sequencing**

DNA was extracted using the DNeasy Blood & Tissue kit (Qiagen, ID 69504) using the manufacturer’s protocol with a few modifications. Briefly, swabs were placed in 2mL sterile microtubes, and 260 µL of enzymatic lysis buffer were added. After a 30-minutes incubation at 37°C, 50 µL of proteinase K and 200 µL of AL buffer were added before the incubation at 56°C for 30 minutes. The elution step was done twice in 100 µL of elution buffer. The two eluates were pooled to obtain a single 200 µL DNA sample per swab. DNA quality and quantity was assessed by spectrophotometry (NanoDrop 1000, Thermo Fisher Scientific, USA).

The V3-V4 region of the 16S rDNA gene was amplified using bacterial primers modified for Illumina sequencing 341F (5’-CTTCCCTACACGAGCTCTTCCGATCT-ACGGRAGGCAGCAG-3’) (Liu et al. 2007) and 784R (5’-GGAGTTCACGTGTGCTCTTCCGATCT-TACCAGGGGTATCTAATCCT-3’) (Andersson et al. 2008). Amplification was very difficult due to the low DNA concentration, and possible contamination by keratinocytes in skin samples. Consequently skin and water samples were amplified using two different PCR kits and conditions, which are provided in Supplementary Information S1. Both sample types were amplified in triplicates. After PCR, the success of amplification was verified by migration on agarose gels, and equal volumes of three PCR products were pooled for each sample. After pooling, final concentration measured by Nanodrop (Wilmington, USA) averaged 14 ng. µL⁻¹ (±17, n=43). After amplification, equimolar amounts of all PCR products were pooled and cleaned up using calibrated
Ampure XP beads by an external laboratory (MR DNA, Shallowater, USA) and sequenced on a single run of Illumina platform using the 2 x 250 bp MiSeq chemistry. To check biases induced by the two different PCR protocols, we amplified 2 water DNA samples using both PCR kits and compared them after sequencing. They showed similar community structure (see S1). The nucleotide sequence data is available in the NCBI SRA database under the biosample numbers SAMN07278850-SAMN07278894.

**Sequence processing and phylogenetic analyses**

Assembly of paired reads was performed by the sequencing platform. All subsequent steps of sequence processing were performed following the SOP of Kozich et al for MiSeq (Kozich et al. 2013), [https://www.mothur.org/wiki/MiSeq_SOP](https://www.mothur.org/wiki/MiSeq_SOP) (2016) using Mothur (Schloss et al. 2009). After removing sequences with an irregular length (i.e. outside a range of 420-460 pb), sequences were aligned along the SILVA reference database (Quast et al. 2013) (release 123). Unaligned sequences were removed from the final alignment during this process. Chimeras were removed using UCHIME (Edgar et al. 2011). Filtered sequences were then classified using the SILVA reference taxonomy and the non-bacterial reads were removed. After these steps, we obtained a total of 2,198,758 sequences from our 43 samples, with 51,133±20,883 (expressed as Mean±SD) sequences per sample. The number of sequences read for each sample is unlikely correlated with total abundance of bacteria in sample, while it could bias assessments of microbial biodiversity. Therefore, to ensure that further diversity assessments were not biased by the uneven sequencing efficiency among samples, 10,000 sequences were sub-sampled within each sample (Supplementary Information S2). Non-parametric Chao’s coverage estimator was computed in each community to assess effect of subsampling level using “Coverage” function provided in [entropart](https://cran.r-project.org/package=entropart) R-package (Marcon and Hérault 2014). This index averaged 0.98±0.008 among microbial communities testifying for the accuracy of further diversity analyzes.

Sequences were then clustered into OTUs with 99% sequence identity, and the dominant sequence for each OTU was selected as reference and aligned against the SILVA reference database using Mothur for subsequent phylogenetic tree reconstruction. An outgroup was defined using a set of archaeal sequences obtained from SILVA database and re-aligned against the previous alignment of reference sequences using the MAFFT v7 with –add option (Katoh et al. 2002) before tree reconstruction using Fasttree (Price, Dehal, and Arkin 2010).
To estimate the potential functions of microbial OTUs based on 16S rDNA data, we used PICRUST software (Langille et al. 2013) on reference sequences, using KEGG orthologs (Kanehisa et al. 2004) grouped into pathways (function `categorize_by_function.py`, level=3). A matrix containing 329 pathways was obtained. We then removed all eukaryotic functions, for instance genes related to cardiovascular diseases and categories grouped as “organismal systems”. NSTI values averaged 0.04±0.02 and 0.12±0.02 respectively in skin-associated and planktonic communities, indicating that OTUs sequences were close enough to the nearest 16S rDNA of reference genomes to infer functions.

Investigating the possible presence of pathogens

Two additional phylogenetic analyses were performed separately for the two genera *Staphylococcus* and *Streptococcus* to look for putative pathogenic bacteria on cetacean skin. Near full-length 16S rDNA sequences of well-known pathogenic and non-pathogenic species of these genera were downloaded from the SILVA database (ACC number provided in Supplementary S3). Reference sequences of the most abundant OTUs belonging to these two genera (*i.e.* 35 Staphylococci and 31 Streptococci sequences), as well as the SILVA sequences were aligned against the SILVA reference database using Mothur, and added into the SILVA reference phylogenetic tree using ARB software (Ludwig et al. 2004). The full phylogenetic tree was then pruned using the `ape` R-package to remove all but the added sequences, while keeping the topology of the tree. We then visualized the phylogenetic tree to determine if OTUs from this study were close to the pathogenic species considered.

Assessing diversity of and dissimilarity between skin microbiotas

Four complementary diversity indices were computed to assess the taxonomic and phylogenetic facets of diversity, including their respective compositional and structural components (Chiarello et al. 2015). The compositional diversity accounts only for the presence/absence of OTUs or phylogenetic lineages (here defined as subsets of the phylogenetic tree, containing OTUs and their associated branch lengths). Compositional taxonomic diversity was measured by counting the number of OTUs in a sample (OTUs or functional richness). The phylogenetic compositional diversity (*i.e.* the phylogenetic richness) was measured as Faith’s PD (Faith 1992) divided by the total PD of the tree (to scale values between 0 and 1). The structural diversity accounts for the relative abundance of OTUs or phylogenetic lineages, based on the number of sequences represented by each OTU.
The taxonomic structural diversity was computed using the Shannon index (Shannon 1948), expressed in Hill numbers (Jost 2007) on abundance of OTUs. The phylogenetic structural diversity was measured using the Allen index (Allen, Kon, and Bar-Yam 2009). All diversity indices were computed using R software. The taxonomic alpha diversity indices were computed using our own functions (available at https://github.com/marlenec/chao), while the Faith PD and Allen index were calculated respectively using the picante and entropart packages (Kembel et al. 2010; Marcon and Hérault 2014).

Similarly, we used four complementary beta-diversity indices to assess the taxonomic and phylogenetic dissimilarity between pairs of microbiotas, according to their composition or structure. The compositional taxonomic dissimilarity was assessed based on presence/absence of OTUs, using the Sorensen index (Sørensen 1948) computed with betapart package (Baselga and Orme 2012). The structural taxonomic dissimilarity, taking into account the relative abundance of OTUs, was measured using the multiplicative decomposition of the Shannon index (Chiarello et al. 2015). The phylogenetic compositional and structural dissimilarities were computed using the unweighted and weighted versions of the Unifrac index (Lozupone and Knight 2005; Lozupone et al. 2007), respectively, from the GUniFrac package (Chen and ORPHANED 2012).

Kruskal-Wallis tests (KW) were performed on alpha-diversity indices to assess the effect of sample type (i.e. water vs. skin samples), species, individual, sex, or body zone on microbial alpha-diversity. When significant, the KW was followed by post-hoc pairwise comparisons among groups using the pgirmess package, which includes the correction for multiple tests from Siegel and Castellan (Siegel and Castellan 1988; Giraudoux 2011). The correlation between the age of the individual and its associated alpha-diversity was assessed using a Spearman’s correlation test using stats R-package. Beta-diversity values were visualized on PCoA plots using the ape package (Paradis, Claude, and Strimmer 2004). The effect of sample type, species, individual, age, sex, and body zone on the structure and composition of microbial communities was assessed by performing separated one-factor PERMANOVAs with 999 permutations on beta-diversity values using vegan package (Dixon 2003). The number of identical OTUs between skin microbiota and planktonic communities was analyzed using an Euler Diagramm computed with venneuler R-package (Wilkinson and Urbanek 2011). To assess how each microbial clade contributed to the dissimilarity between
planktonic and skin microbiotas, as well as between microbiotas of cetacean species, we performed a LefSe analysis (Segata et al. 2011). LefSe provides Linear Discriminant Analysis (LDA) scores for the bacteria clades contributing the most to the differences between cetacean species.

Comparing skin microbiota of cetaceans and other vertebrates

The skin microbiota of dolphins and killer whales was compared to the published skin microbiota of 11 terrestrial and marine vertebrates, namely Human (Grice et al. 2009; Staudinger, Pipal, and Redl 2011; Oh et al. 2014), pig (McIntyre et al. 2016), humpback whale (Apprill et al. 2014) and eight teleostean fish species (Larsen et al. 2013; Chiarello et al. 2015). Due to the different primers that were used for these different species, we could not directly reanalyze sequences from studies to assess OTUs abundance. Therefore, we extracted clades relative abundance from published figures and averaged across all individuals (i.e. 36 humans, 4 pigs and 57 humpback whales) for each mammalian species. In the case of marine fishes, as individual data was not available for all species, we chose to average clades relative abundances of all species to make a single “fish” category. The most abundant clades colonizing the animals were averaged for each animal; and a Bray-Curtis dissimilarity index (Bray and Curtis 1957) (BC) between the different microbiotas was computed based on the relative abundance of the different clades. A BC index of 1 indicates that microbiotas are maximally dissimilar, i.e. that they are dominated by different clades while a BC=0 indicates that the two microbiotas have the same taxonomic structure (i.e. same clades with same abundances).

RESULTS

Diversity of skin and planktonic microbiotas

We recovered a total of 7,287 OTUs among our 43 samples, with OTU richness ranging from 210 to 606 across samples. Water samples (481±64 OTUs, n=11 samples) were significantly richer than skin samples (332±84 OTUs, n=32 samples) (Kruskal-Wallis, P<0.0001, Fig. 1). However when considering the relative abundance of OTUs, skin samples were significantly more diverse than water samples (KW, P=0.001), with a Shannon index of 30.4±23.6 and 9.2±2.5 equivalent number of species, respectively (Supplementary Information S4). Phylogenetic richness did not significantly differ between planktonic and skin-associated communities (KW, P=0.06, S4). However, when taking into account the relative abundance of phylogenetic lineages, skin samples were significantly more diverse than water samples (KW,
P(<0.0001), with Allen index being ca. 1.8 times higher in skin-associated communities than in the planktonic ones (Fig. 1).

Skin-associated microbial alpha-diversity did not significantly differ between species (KW, P>0.05). At intraspecific level, there was no effect of individuals or body zones on OTUs richness and phylogenetic richness (KW, P>0.05, Fig. 1 and S4). However a significant effect of individual on taxonomic diversity was found for both species (Shannon index, KW, P=0.02 and 0.01 for dolphins and killer whales, respectively), which was not explained by age (Spearman’s correlation test, P>0.05) or sex (KW, P>0.05). Post-hoc pairwise comparisons showed that the dolphins Sharki and Rocky hosted significantly different levels of taxonomic diversity; and that Valentin hosted significantly lower taxonomic diversity than Freya, Inouk and Wiki (P<0.05, S4). The microbiotas of these individuals also have contrasted levels of phylogenetic diversity (P<0.05).
Fig. 1: Biodiversity of skin microbial communities from captive killer whales (A and D) and common bottlenose dolphins (B and E), and planktonic communities (C and F). The first row of plots (A-C) illustrates taxonomic richness, i.e. number of OTUs observed in a sample and the second row of plots (D-F) illustrates phylogenetic diversity, measured using Allen’s index (i.e. accounting for relative abundance of phylogenetic lineages). Total diversity of each individual (i.e. accounting for all body zones sampled) is illustrated with larger light-gray bars. Bars on panels C and F represent the mean (and associated standard deviation) of OTU richness and phylogenetic diversity for planktonic communities \((n=3 \text{ water samples})\). “Pool” refers to animal’s surrounding water, and “Input” refers to the water sampled at the exit of filtering system.

Dissimilarity between microbiotas

Water communities and skin-associated microbial communities significantly differed for all facets of biodiversity considered (Table 2, Fig. 2, Supplementary Information S6). For instance, only 12% of OTUs found on cetacean were also present in surrounding planktonic communities (Fig. 3). Taxonomic \((0.61\pm0.17)\) and phylogenetic \((0.43\pm0.19)\) dissimilarities reached their higher level between planktonic and skin-associated communities (S5).

Dolphin- and killer whale-associated communities significantly differed for both taxonomic and phylogenetic diversity, as revealed by PERMANOVA (Table 2). Within each species, individuals had skin microbiotas with significantly different phylogenetic structure (Table 2).
The age of the individual, and in a lesser extent, its sex, had a significant effect on diversity of dolphin skin microbiota when considering relative abundances of OTUs or phylogenetic lineages (PERMANOVA, Table 2). For killer whales, neither age nor sex had a significant effect on skin-associated microbiota. Microbiotas from the four studied body zones were not significantly different (PERMANOVA, Table 2).

**Fig. 2**: Phylogenetic dissimilarity between microbial communities illustrated along the two first axes from Principal Coordinates Analyses computed on weighted Unifrac dissimilarity index. All samples (i.e. skin-associated communities of 4 body zones of 4 captive common bottlenose dolphins and of 4 killer whales, and planktonic communities from respective pools and exit of filtering system) are present on panel (A). Panels (B) and (C) represent only captive killer whales or common bottlenose dolphin skin-associated communities with the 4 body zones of each individuals being delimited by a polygon.
Table 2: Determinants of biodiversity of skin microbial communities.

<table>
<thead>
<tr>
<th>Biodiversity facets</th>
<th>Taxonomic</th>
<th>Phylogenetic</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Sorensen</td>
<td>Beta-Shannon</td>
</tr>
<tr>
<td><strong>Dissimilarity indices</strong></td>
<td>R²</td>
<td>P</td>
</tr>
<tr>
<td><strong>Factor</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Biotic vs. abiotic</td>
<td>0.09</td>
<td><strong>0.001</strong></td>
</tr>
<tr>
<td>Dolphins vs. K. whales</td>
<td>0.06</td>
<td><strong>0.001</strong></td>
</tr>
<tr>
<td><strong>Killer whales</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Individuals</td>
<td>0.23</td>
<td><strong>0.010</strong></td>
</tr>
<tr>
<td>Body zones</td>
<td>0.19</td>
<td>0.683</td>
</tr>
<tr>
<td>Age</td>
<td>0.07</td>
<td>0.105</td>
</tr>
<tr>
<td>Sex</td>
<td>0.08</td>
<td>0.092</td>
</tr>
<tr>
<td><strong>Dolphins</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Individuals</td>
<td>0.22</td>
<td><strong>0.035</strong></td>
</tr>
<tr>
<td>Body zones</td>
<td>0.20</td>
<td>0.569</td>
</tr>
<tr>
<td>Age</td>
<td>0.07</td>
<td>0.211</td>
</tr>
<tr>
<td>Sex</td>
<td>0.07</td>
<td>0.332</td>
</tr>
</tbody>
</table>

Effect of each factor was tested using permutational ANOVAs (PERMANOVAS, 999 permutations) on dissimilarity matrices with for each facet of biodiversity, two indices: one index accounting only for composition of OTUs or phylogenetic lineages (i.e. Sorensen or Unweighted Unifrac), and one index accounting for relative-abundance of OTUs or phylogenetic lineages (i.e. beta-Shannon or weighted Unifrac). Bold P-values (<0.05) indicate a significant effect of the tested factor. Partial R-squared (R²) is the proportion of variation in the dissimilarity matrix explained by the tested factor.

Fig. 3: Euler diagram representing the number of skin-associated OTUs from each species (killer whales and common bottlenose dolphins) and planktonic communities that are shared between the three conditions and unique ones.
Composition of bacterial communities

Planktonic communities were dominated by Alphaproteobacteria (95.9±0% of sequences), especially Hyphomonadaceae, Rhodospirillaceae and Rhodobiaceae (71.4±0.1%, 7.6±0.1% and 6.6±0.1% of all Alphaproteobacteria, respectively) (Fig 4 and S6). Ca. 90% of planktonic OTUs could not be identified at genus level, excepted Anderseniella sp. [Alphaproteobacteria], which contributed to 6.3±0.1% of sequences in water samples (S6).

Cetacean skin microbiota was mostly composed of Gammaproteobacteria (57.5±27%), Alphaproteobacteria (22.4±18.8%), Actinobacteria (7.9±0.1%) and Bacilli (7.3±0.1%) (Fig 4). The most abundant genus on both species was Psychrobacter sp. [Gammaproteobacteria], which dominated skin samples, making 30.1±31% of total abundance on killer whales’ skin and 45.2±28% on dolphins’ skin, while they represented a small fraction of sequences (1.1±0.2%) in planktonic communities (S6). Other genera were punctually abundant in skin samples, including Enhydrobacter (8.4±11.9% in both species), Staphylococcus (4.6±7.5%), Sphingomonas (3.5±5.9%), Paracoccus (2.8±4.9%) and Gardnerella (0.4±0.7%) (S6).

Families and genera revealed by LefSe analysis for the two host species were mostly scarce and are not visible in S6. Four biomarkers were found for killer whales’ skin, belonging to Alphaproteobacteria: Phyllobacteriaceae (log10 effect size=4.5), Rubellimicrobiunm and Ruegeria (Rhodobacteraceae) (3.6 and 3.5), and Microvirga (Methylobacteriaceae) (3.4). Three biomarkers were significantly more abundant on dolphin’s skin: Nocardiaceae (3.4), Enterobacteriaceae (3.2) and Caulobacter (Caulobacteraceae) (3.1).

Among OTUs from skin microbiotas, 237 were identified as Staphylococcus sp., with the most abundant one averaging 4.4% of sequences in dolphin-associated communities. The phylogenetic analysis of the 35 most abundant ones showed that none of them was related to the recognized marine mammal pathogen Staphylococcus delphini or other pathogenic staphylococci (Supplementary S3). The most abundant Staphylococcus was closely related to the opportunist Staphylococcus warneri. 45 OTUs identified as Streptococcus were recovered in skin-associated communities, with the most abundant one averaging 0.5% of sequences of both species’ microbiotas. None of the 40 most abundant OTUs were related to the pathogenic hemolytic Streptococci.
Fig. 4: Mean relative abundance of bacterial classes in skin-associated communities of common bottlenose dolphin and killer whales, and planktonic communities. P: upper side of pectoral fin, D: dorsal fin, C: upper side of caudal fin, A: anal zone. “Pool” refers to animal’s surrounding water, and “Input” refers to the water sampled from the exit of pipe from filtering system.

Potential functional diversity of planktonic and skin microbiota

LefSe analysis identified 19 functional biomarkers of planktonic communities, with strongest effect sizes for pathways involved in environmental information processing (Supplementary Information S7). Other functional biomarkers of planktonic communities were pathways related to cellular processes, especially those related to motility (flagellar assembly pathway and motility proteins, being respectively twice and 60% more abundant in planktonic communities) and cell cycle, principally reflected by *Caulobacter* cell cycle pathways.

Cetacean skin microbiota was characterized by functions involved in genetic information processing, especially pathways related to DNA repair and recombination proteins (Supplementary Information S7), DNA replication and translation (especially from 0.9 to 1.3% of proteins involved in ribosome biogenesis in skin communities).
Comparison of cetacean microbiota with other vertebrate microbiotas

Predominant clades in skin microbiota of dolphins and killer whales sampled for this study were distinct from skin microbiota reported for terrestrial and marine vertebrates (Fig. 5). Dissimilarity in relative abundance of major microbial clades between the two toothed whale species studied here was twice lower than dissimilarity between the toothed whales and the baleen whale (free-ranging Humpback whales) or the teleostean fishes (Fig. 5). Humpback whale hosts higher proportions of Bacteroidetes than toothed whales, while fishes host more Firmicutes and Betaproteobacteria. Skin microbiota of toothed whale was highly dissimilar from the skin microbiota of pig and human (Bray-Curtis >0.95, Fig. 5).

Fig. 5: Comparison of cetacean microbiota with microbiotas of other vertebrates. Mean relative abundance of predominant microbial clades in marine and terrestrial animals (A) and associated pairwise Bray-Curtis dissimilarity computed on mean relative abundance of these clades (B). In panel B, the error bars associated to the 3 top bars are the standard deviation across the two Bray-Curtis values obtained from the separated comparison of dolphins and killer whales skin-associated microbiota with human, fish and pig’s microbiota, respectively.
DISCUSSION

The skin microbiotas of the captive dolphins and killer whales were distinct from their surrounding planktonic communities. Indeed, while 100 mL of water contained nearly 1.5 times more OTUs than a single sample corresponding to a 63 cm² swabbing of an animal’s skin, the taxonomic diversity, accounting for OTUs relative abundances, was three times higher in skin samples than in water samples (S4). Therefore, skin of cetacean host less OTUs than the surrounding seawater, but due to a higher evenness of OTUs abundances, skin microbiota is indeed more diverse than planktonic communities. Accordingly, phylogenetic diversity, taking into account the relative abundance of phylogenetic lineages, was also up to three times higher in skin-associated communities. Hence, OTUs dominating skin microbiota were distributed among distant phylogenetic lineages while OTUs dominating planktonic communities were clustered into a few lineages. These contrasted patterns were reported for teleostean fishes raised in controlled conditions (Chiarello et al. 2015) and free-ranging humpback whales (Apprill et al. 2014).

Besides differences in level of diversity, planktonic and skin microbial communities also host different OTUs and phylogenetic lineages (Table 2, Fig. 2-4). Planktonic communities were indeed dominated by *Hyphomonadaceae* and *Rhodospirillaceae*, which contain several genera typical of marine environments (Abraham and Rohde 2014; Baldani et al. 2014), and *Anderseniella* (*Rhodobacteriaceae*), that was firstly isolated from marine sediment, and is present in marine aerosols (Brettar et al. 2007; Seifried, Wichels, and Gerdts 2015). By contrast, skin-associated microbial communities were dominated by *Psychrobacter sp*. [*Gammaproteobacteria*], a genus that was previously shown to be predominant on the skin of humpback whales (Apprill et al. 2014), and was also isolated from the skin and muscle biopsies of Weddel seals (Mellish et al. 2010), and from the skin of teleostean fishes (Lowrey et al. 2015). Since *Psychrobacter sp*. could act as an opportunistic pathogen in skin lesions of sea lions (Alvarez-Perez et al. 2010), it should be looked for on skin of other marine animals, and particularly on endangered mammals.

Other predominant genera which were found on both killer whales and dolphins skin (*Enhydrobacter*, *Staphylococcus*, *Sphingomonas*, *Paracoccus*) are known commensals of the human skin (Perez et al. 2016). Such genera were not detected in healthy free-ranging whales (Apprill et al. 2014), suggesting transfer from caretakers (*e.g.* during medical examinations or training). The skin of marine mammals has key similarities with the skin of terrestrial mammals, characterized by a very thick epidermidis composed of keratinocytes (Berta, Sumich, and Kovacs 2015), with specificities, for instance an incomplete process of
cornification, referred as ‘parakeratosis’, which also naturally occurs in mammalian mucosa (Cozzi, Huggenberger, and Oelschläger 2016; Delport et al. 2016). These skin characteristics may explain the presence of such human-associated genera on skin of captive dolphins and killer whales’ skin through transfer from their caretakers, as it has been observed in other animals maintained in captivity (Delport et al. 2016; Clayton et al. 2016). However, skin microbiota from free-ranging dolphin and killer whale has to be analyzed before excluding the possibility that such genera might naturally occur in Odontoceti.

Additionally, cetacean skin surface is covered by a biogel that smoothed its surface and prevents the attachment of settling organisms (Baum et al. 2001). This property, together with animal’s behavior (swimming and jumping which favoring particles detachments), and skin sloughing, may induce a constant shedding of skin-associated microorganisms. Captive animals may perform these behaviors less frequently than wild animals, which may ultimately favor the growth of opportunistic bacteria. Moreover, while all but one animal did not receive antibiotics for at least 6 months (15 days in the case of one killer whale) the occasional use of antibiotics on such captive animals may also modify their microbiota. For instance, killer whales and dolphins inhabiting industrialized coastal zones, hence likely confronted to antibiotics released to the sea through wastewater, were shown to host antibiotic-resistant bacteria in their pulmonary system and gut (Kueneman et al. 2014; Council et al. 2016). However, long-term effects of occasional antibiotic use on skin microbiota, especially in marine mammals, still need to be investigated.

The different phylogenetic lineages present in planktonic and skin-associated communities could perform different functions (Table 2), with a higher proportion of biochemical pathways related to motility and membrane transport in planktonic communities. A metagenomics approach in surface seawater communities showed a similar trend towards dominance of flagellum assembly pathway and of membrane transporters, which is consistent with the motile heterotrophic lifestyle of surface planktonic communities competing for nutrients (DeLong et al. 2006). By contrast, skin-associated communities contained higher proportions of functions involved in protein folding, DNA replication, reparation and translation. Such functions may be driven by the need for skin-associated bacterial cells to grow rapidly on the skin of the animal to counter sloughing (Apprill et al. 2014). However, in this study, microbial functions were estimated with the PICRUST software using phylogenetic affiliation of OTUs and a reference genome database. This assessment is limited to previously annotated genes (ignoring undiscovered functional genes), and do not account
for potential differences in gene expression. Therefore, the high similarity in functional
diversity between animals found in this study should be confirmed by further metagenomics
and metatranscriptomics studies.

Skin microbiota was species-specific (Table 2). Host-species specificity of skin microbiota
had already been evidenced in other marine (Larsen et al. 2013) and terrestrial animals
(Kueneman et al. 2014; Council et al. 2016; Avena et al. 2016). The major contributors in this
interspecific difference were several Alphaproteobacterial families, namely
Phyllobacteriaceae and Rhodobacteraceae, that were more abundant on killer whale’s skin,
and Nocardiaceae [Actinobacteria] and Enterobacteriaceae [Gammaproteobacteria] that had
higher abundances on dolphin’s skin (Fig 4 and S6).

This difference of microbial lineages between host species living in similar conditions has
also been found for amphibians in natural pounds (McKenzie et al. 2012), and fishes raised in
controlled conditions (Chiarello et al. 2015). These findings reinforce the hypothesis that even
in aquatic environment were microbes are highly abundant and diverse; the unique features of
animal’s skin shape its microbiota. Further studies are needed to determine which factors (e.g.
differences in immune system, skin structure, and/or pH and body temperature) on cetacean
skin may promote this species effect. A likely important one is the differential expression of
antimicrobial peptides between species, that were found to be secreted in the skin of several
Delphinidae (Meyer and Seegers 2004), and which were shown to determine interspecific
microbial differences in invertebrate model species (Franzenburg et al. 2013).

Within each species, individuals showed contrasted levels of OTUs and phylogenetic
diversity (Fig 1 & S4) as well as dissimilarities in abundances of taxonomic clades and
phylogenetic lineages (Fig 4 & S6, Table 2). Hence, individual features seem to play an
important role for shaping diversity of skin microbiota even when individuals have been
living in the same environment and have frequent social interactions including direct skin
contact (Dudzinski 1998; Dudzinski et al. 2009), which should favor homogenization of skin
microbiota among individuals (as observed for humans (Meadow et al. 2013)). This is well
illustrated by the mother killer whales Freya and her young son Valentin (Table 1). Both are
in continual contact, but did not have closer skin-associated microbial structures (Fig. 2).

Inter-individual variability was already documented for the pulmonary microbiota of
bottlenose dolphins housed in SeaWorld (QLD, Australia) (Lima et al. 2012) and was shown
to be consistent over time. Moreover, in the same study, authors confirmed this variability in
a total of 24 free-ranging dolphins of two species (*Tursiops truncatus* and *T. aduncus*), without being able to detect any influence of age or sex of the animals. Our results for captive animals highlight the importance of intraspecific variability of skin microbiome as well as correlation between these differences and individual traits, *e.g.* sex and age as shown for dolphins. Further studies are needed to unravel the proximal drivers, such as immunity or physiology, of skin microbiota variability within a species.

Body zones did not have replicable different skin microbiota (Table 2, Fig 2), contrary to the patterns observed in humans (Grice et al. 2009; Perez et al. 2016). This absence of difference in skin microbiota between body zones suggests that environmental conditions are more homogeneous throughout the body of cetaceans than humans, probably due to several reasons that are not mutually exclusive: the absence of hair follicles and sebaceous and sweat glands in the dermis of cetaceans (Cozzi, Huggenberger, and Oelschläger 2016), the absence of moist vs. dry microenvironments differentiation due to the aquatic habitat, and/or the leaching effect of swimming that would homogenize physicochemical conditions at skin surface. However, marine mammals could also harbor unique skin microbiome in other micro-niches that we did not sample in our study, and which may provide different nutrient sources and protection, by the presence of mucus in the eyelids (Berta, Sumich, and Kovacs 2015) or pulmonary surfactant in the blowhole (Berta, Sumich, and Kovacs 2015), or which may retain particles and microbes more easily, *e.g.* at fin folding.

Studies focusing on gut microbiota of insects (Brucker and Bordenstein 2012; Brooks et al. 2016) and terrestrial mammals (Phillips et al. 2012; Brooks et al. 2016; Groussin et al. 2017) found a correlation between hosts phylogeny and microbiota and suggested that microbiota result from ‘phylosymbiosis’ (Brucker and Bordenstein 2012; Brooks et al. 2016), partly due to co-speciation of hosts and microbes that are vertically transmitted (Groussin et al. 2017). In the case of skin-associated microbiota, and more importantly in the case of those of animals living in seawater, there is still no test if this microbiota is vertically or horizontally transmitted between individuals, and if differences in skin microbiota are correlated to host’s phylogeny. Here, using previously published microbiota of 11 vertebrate species (Fig. 5), we showed that the microbiota of captive bottlenose dolphins and killer whales is twice closer to humpback whale and marine teleostean fishes than to terrestrial mammals (human and pig). This suggests that the marine environment has a strong impact on the composition of skin microbiota compared to evolutionary legacies within mammals. The specificity of marine
skin microbiotas may be related to the aqueous conditions (Torsvik, Øvreås, and Thingstad 2002), as well as the salinity experienced by skin-associated microbial cells, which are major structuring factors of prokaryotic communities in other environments (Lozupone and Knight 2007; Auguet, Barberan, and Casamayor 2009). Assessing skin microbiotas of more marine vertebrates, including fishes from several orders, pinnipeds and sirenians as well as reptiles, is thus needed to confirm this hypothesis. Such assessments should include both metagenomics and metatranscriptomics approaches to unravel drivers and roles of skin microbiotas.

ACKNOWLEDGEMENTS
We acknowledge the “Fondation Marineland” for its financial support, and Marineland staff for their technical assistance during sampling. We also acknowledge three anonymous reviewers who helped improving this manuscript.

AUTHORS CONTRIBUTIONS
MC, TB and SV conceived the study. MC, SV and CB collected samples. MC performed DNA extraction and PCR, and MC and JCA analyzed results. MC drafted the manuscript, and all authors (MC, TB, SV, CB and JCA) contributed to the final version.

ADDITIONAL INFORMATION
This project was funded by the “Marineland Foundation” (3,000€). The funders had no role in study design, data collection and analysis decision to publish, or preparation of the manuscript. Authors declare no conflicts of interests.
4.4 Manuscrit D

Skin microbiome of coral reef fishes is diversified, species-specific, not phylogenetically conserved and vulnerable to human activities

Marlène Chiarello¹, Jean-Christophe Auguet¹, Yvan Bettarel¹, Corinne Bouvier¹, Thomas Claverie¹, Nicholas AJ Graham², Fabien Rieuvilleneuve¹, Elliot Sucré¹, Thierry Bouvier¹, and Sébastien Villéger¹

¹Marine Biodiversity, Exploitation and Conservation (MARBEC), Université de Montpellier, CNRS, IRD, IFREMER, Place Eugène Bataillon, Case 093, 34 095 Montpellier Cedex 5, France
²Lancaster Environment Centre, Lancaster University, Lancaster, UK

En préparation

Des informations supplémentaires complètent ce manuscrit, disponibles en section 7.5
SUMMARY
All animals are associated with microbial communities, called microbiomes (Grice et al. 2009; Wahl et al. 2012; Larsen et al. 2013; Apprill et al. 2014; Bourne, Morrow, and Webster 2016). Studies on terrestrial vertebrates found that this microbiome varies across species, and that these differences are related to host ecological traits (e.g. diet), and evolutionary legacy (i.e. related species tend to host more similar microbiome, a pattern called “phylosymbiosis”) (Ochman et al. 2010; Phillips et al. 2012; Groussin et al. 2017). However, these studies focused on intestinal microbiome of terrestrial vertebrates. The existence of such patterns on skin microbiome, which is crucial for host fitness in marine environment (Efrony et al. 2007; Gómez and Balcázar 2008; Krediet et al. 2013; Lowrey et al. 2015), is underexplored. More specifically, the skin microbiome of marine fishes, which constitute the most diverse group of vertebrates (Nelson, Grande, and Wilson 2016), remain largely unknown, while many species are under increasing threat especially on coral reefs (Graham et al. 2011).

We analyzed the prokaryotic microbiome of 44 fish species of the Western Indian Ocean. We found that prokaryotes living on fish skin are highly diverse, contain numerous unknown clades, and are highly variable between species. However, these interspecific differences were not coupled to phylosymbiosis and only weakly explained by fish diet. Finally, we found that fish species hosting the highest microbial diversity are also the most vulnerable to fishing.

RESULTS AND DISCUSSION
Skin of reef fishes is a microbial hotspot
We sampled the skin microbiome of 138 individuals of 44 species from the coral reefs around Mayotte Island (France) (Fig 1). These species represented 5 orders and 22 families and the main ecological groups dominating coral reefs (see supplementary Information S1).

Biodiversity of microbiome was assessed using 16S rDNA high-throughput sequencing, using universal prokaryotic primers (Caporaso et al. 2011; Apprill et al. 2015, 2011) and applying two complementary indices of phylogenetic diversity, the first one based on presence/absence data, which is an index of phylogenetic richness (Faith’s PD), and the second one taking account for microbial abundances and which is referred as phylogenetic entropy (Allen index). We assessed the dissimilarity between microbial communities considering only the presence of OTUs (U-Unifrac) or accounting for their abundance (W-Unifrac) (See supplementary methods for more details).
Fig 1: Phylogenetic tree relating all 44 fish species included in this study (A) adapted from Rabosky et al., along with the mean class-level composition of their skin associated microbiome (B). The 13 most abundant classes in all communities are represented with different colors, and other classes are in white. The mean composition of planktonic communities is indicated at the top. The mean phylogenetic richness of each community (C) is indicated by proportional circles and a color scale from light gray to black. The panel (D) represents the vulnerability to overfishing index as in Cheung et al.
A total of 10,430 OTUs were found on fishes, representing 34 archaeal and bacterial classes and 19 phyla, while 2,210 OTUs representing 17 classes and 11 microbial phyla were found in planktonic communities (Figure 1). Fish skin microbiome was highly variable, as no OTU was recovered in all individuals. Fish skin microbiome was significantly distinct from surrounding planktonic communities, both in term of composition in microbial lineages and of their abundances (Table 1).

Table 1: Determinants affecting microbial community structure

<table>
<thead>
<tr>
<th>Factor</th>
<th>Unweighted Unifrac</th>
<th>Weighted Unifrac</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fish vs. water</td>
<td>0.001</td>
<td>0.001</td>
</tr>
<tr>
<td>Fish species</td>
<td>0.001</td>
<td>0.001</td>
</tr>
</tbody>
</table>

Effect of each factor was tested using two separated permutational ANOVAs (PERMANOVAS, 999 permutations) on relative abundance weighted and unweighted Unifrac dissimilarity matrices.

Fish skin microbiome was significantly enriched in *Gammaproteobacteria* (14±12% of abundance in water column vs. 38±24% on fish surface), especially *Vibrionaceae* (1±3% vs. 7±11%) and *Altermonodales* (8±10% vs. 30±12%), *Rhizobiales* (0.01±0.03% vs. 3±5%) and *Clostridiales* (0.03±0.04% vs. 3±4%) compared to planktonic communities that were enriched in *Cyanobacteria* (24±12% of abundance in water column vs. 4±8% on fish surface), *Rhodobacteraceae* (7±4% vs. 6±9%) and *Flavobacteriaceae* (9±4% vs. 5±7%) (Fig 1 and S2). Bacteria largely dominated both planktonic and skin-associated communities, as the 75 OTUs identified as Archaea averaged 1.1% of abundance in planktonic communities and 0.8% in skin-associated communities. 37% of Archaeal OTUs were affiliated to the phylum *Thaumarchaeota*, 17% to the phylum *Euryarchaeota*, and all other OTUs remained unclassified. *Thaumarchaeota* were mostly affiliated to the Marine Group I (9 OTUs out of 28 Thaumarchaeota) and South African Gold Mine Group 1 (4 OTUs). *Euryarchaeota* were mostly classified into *Thermoplasmata* (5 OTUs out of 13 Euryarchaeota), *Methanomicrobia* (4 OTUs) and *Halobacteria* (3 OTUs). While not represented in the figure in S2 because of their small effect size, *Thaumarchaeota*, *Thermoplasmata* and *Halobacteria* were significantly more abundant in fish samples (see S2 for their respective effect sizes). These archaeal clades were already found in seawater and in sponges (Polónia et al. 2016) and in
surface mucus of scleratinian corals (Kellogg 2004). To our knowledge, this is one of the first
records of Archaea on teleost fish surface.

Fish skin microbiome and planktonic communities harboured high proportions of unclassified
microbial taxa. Using the Mothur taxonomic affiliation method, as many as 60% of the
11,583 recovered OTUs recovered in both fish skin microbiome and planktonic communities
could not be classified at class level (Fig 1) and 46% even could not be classified at phylum
level. These OTUs ranged from 0 to 34% of total abundance in a sample. We refined the
taxonomic affiliation of the most frequent unclassified OTUs (i.e. 93 OTUs that were
recovered in at least 15 samples, see Supplementary Methods) using the Arb parsimony
insertion tool and the Silva backbone tree (v128). Twenty-three of them belonged to classes
that were not detected during OTU classification by Mothur, so the actual diversity of tropical
microbiome is likely higher than the one reported here (Fig 1, S3). These findings on only a
small subset of fish species suggests that the thousands of tropical fishes certainly hosts a
large number of unknown microbes, which highlights the need for larger screening of marine
vertebrate microbiome.

In addition to the differences between fish skin microbiome and planktonic communities in
terms of composition and structure, phylogenetic richness present on a fish species
(accounting only for presence/absence of microbial phylogenetic lineages) was on average 1.2
times higher on the skin of one fish than in 200-mL of seawater (Kruskal Wallis test on
Faith’s PD, S4). This higher richness of skin-associated microbiome compared to surrounding
planktonic communities is contrasting from other studies on farmed teleostean fishes
(Chiarello et al. 2015) and cetaceans (Chiarello et al., in revision; Apprill et al. 2014), which
found higher phylogenetic richness in seawater than in skin microbiome. Phylogenetic
entropy (measured with Allen index which accounts for the relative abundance of prokaryotic
phylogenetic lineages) of fish skin microbiome was also 1.4 times higher than phylogenetic
entropy of planktonic communities (S4), and suggests that skin-associated microbiome of
tropical fishes was not only richer, but also characterized by a higher evenness of OTUs
abundances than surrounding planktonic communities.

Phylogenetic richness of skin microbiome varied significantly among fish species (S4). In
addition, phylogenetically related fish species had on average similar levels of prokaryotic
phylogenetic richness at their surface (Moran’s I autocorrelation tests, P<0.05 in 96% of
subsamples, n=999, see Supplementary Methods and S4) For instance, *Chaetodon falcula* and *Chaetodon auriga* shared very similar levels of phylogenetic richness at their surface (Fig 1, respectively 1.9±0.8 and 1.7±1.3, making a 10% difference), while *Pterois miles* and *Pterois radiata* hosted a phylogenetic richness of 5.7±1.49 and 5.88±0.00 respectively (also 10% of difference between both *Pterois*). Phylogenetic entropy of microbiomes, which accounts for dominance of phylogenetic lineages, was also significantly different among species (S4) and these differences were also related to phylogenetic distances between fishes, with a weaker signal (Moran’s I, P<0.05 in 88% of subsamples, n=999, S4). For instance the difference between phylogenetic entropies hosted *Chaetodon falcula* and *Chaetodon auriga* was higher than the one between phylogenetic richness (38% of difference), while *Pterois miles* and *Pterois radiata* hosted very similar levels of phylogenetic richness, with a 2.4% difference between them.

Therefore, fish species host different levels of microbial diversity, including richness of microbial phylogenetic lineages, as well as diversity of the most abundant ones, and these differences are related to their phylogeny.

### Assessing dissimilarity between fish skin microbiome and looking for its determinants

Phylogenetic dissimilarity between individuals’ microbiome from the same species was overall high (0.76±0.07 for U-Unifrac, and 0.56±0.15 for W-Unifrac, respectively; S5) especially when compared to the dissimilarity between water samples (S5). Such high intraspecific variability of skin microbiome confirms findings reported for other fish species (Larsen et al. 2013; Chiarello et al. 2015). Nevertheless, variability in phylogenetic structure of microbiome was significantly higher between fish species than between individuals from the same species (PERMANOVA, Table 1, S5). This confirms that fish species host distinct microbial phylogenetic lineages, as it was previously reported for *Teleostei* (Larsen et al. 2013; Chiarello et al. 2015).

However, these high interspecific differences in skin microbiome where not significantly correlated with phylogenetic distance between fish species (Mantel tests on W-Unifrac and U-Unifrac, P<0.05 in 8.4 and 0% of subsamples, respectively, n=999, Fig 2, S6). This absence of phylosymbiosis was persistent even when considering clustering levels for defining OTUs higher than 97% (see Supplementary Methods and S7). Hence, two fish species that are phylogenetically close did not host more similar microbiomes than two distant ones (Fig 2). For instance, the dissimilarity values between the two angelfishes *Pygoplites diacanthus* and
Pomacanthus imperator (Perciformes) (U-Unifrac=0.85±0.11 and W-Unifrac=0.80±0.17) were higher to the ones obtained between Pygoplites diacanthus and the scorpionfish Pterois miles (Scorpaeniformes) (U-Unifrac=0.84±0.07 and W-Unifrac=0.73±0.10). Even within the order Perciformes, which contained most species included in this study, no phylosymbiosis was detected and correlation between phylogenetic distance of hosts and phylogenetic dissimilarity of microbiome was even lower than 0.1 (Mantel tests performed on W-Unifrac & U-Unifrac P<0.05 in 0.7 and 0% of subsamples respectively, n=999, Fig 2 & S6). For instance, the dissimilarity values between Chaetodon falcata and the phylogenetically distant Pterocaesio trilienata (U-Unifrac=0.83±0.06 and W-Unifrac=0.75±0.06) were nearly identical to the dissimilarity values between the two phylogenetically close species Chaetodon falcata and Chaetodon lunula (U-Unifrac=0.84±0.06 and W-Unifrac=0.74±0.17).

**Fig 2:** Phylogenetic dissimilarity between skin-associated microbiomes of fishes belonging to the same taxonomic order (dots) and from fishes belonging to different orders (+' sign), along with the divergence times between fish species. For clarity, only one individual per fish species is represented, and thus intraspecific dissimilarities are not shown.
This absence of phylosymbiosis in fish skin microbiome contrasts with previous studies that reported phylosymbiosis for gut microbiome of terrestrial animals (mammals (Groussin et al. 2017), hominids (Ochman et al. 2010), insects (Brooks et al. 2016)) and for microbiome of tropical sponges (Easson and Thacker 2014). However, whether gut microbiome is correlated with host phylogeny has never been tested for teleostean fishes.

Teleostean fishes are the most diverse groups of vertebrates and comprise 40 orders, 448 families, and more than 30,000 species (Nelson, Grande, and Wilson 2016). According to times estimates from Rabosky et al. (Rabosky et al. 2013), the main fish families included in our study diverged from 53 Mya to 129 Mya. Such divergence times are much larger than those of clades in studies that detected a phylosymbiosis signal, e.g. terrestrial mammals (0.8 to 150 Mya) (Groussin et al. 2017), hominids (15-20 Mya) (Ochman et al. 2010), and other terrestrial clades (0.2-100 Mya) (Brooks et al. 2016), suggesting that phylosymbiosis pattern may be more detectable in phylogenetically close organisms. Still, phylosymbiosis signal was detected among sponge species that diverged 680 Mya (Easson and Thacker 2014).

The absence of a significant phylogenetic signal among the marked interspecific differences in fish skin microbiome could be explained by the role of three drivers: (i) environmental factors, (ii) fish ecological traits, and/or (iii) fish immune system. First, fishes included in this study were sampled at less than 15 km of distance away from each other, and reef type (fringing vs barrier) explained less than 3% of dissimilarity in skin microbiome of fishes found in both habitats, while fish species explained around 30% (PERMANOVAs performed on fish species recovered in both sites, P=0.001 for fish species and 0.002 for reef type, S8). For comparison, microbial phylogenetic entropy of plankton showed higher variability between reef types (PERMANOVA performed on W-Unifrac, P=0.001 R^2=0.27). Hence the variability in skin microbiome found within and among species is unlikely driven by environmental factors and is thus rather driven by host-specific drivers.

We tested whether phylogenetic structure of skin microbiome could be predicted by key ecological traits of fishes, including body size, schooling, period of activity, diet, mobility, and position in the water column (S9 and Supplementary Methods). The only trait that yielded a significant, although weak explanatory effect (R^2=0.18 when tested with W-Unifrac) was diet (Table 2 and S10). Such an effect was not due to a transfer of microbial cells from sessile invertebrates to sessile invertebrates-eating species (S11). Although it has been proven that diet shapes gut microbiome of other vertebrates, including teleostean fishes, at both
interspecific (Ley et al. 2008; Muegge et al. 2011; Sullam et al. 2012; Sanders et al. 2015; Groussin et al. 2017) and intraspecific scales (Wu et al. 2011; Muegge et al. 2011; Sullam et al. 2015; Williams et al. 2016), this is the first report of an effect of species diet on its skin microbiome. An explanation would be a transfer from fishes’ feces to their skin, explaining an indirect effect from their diet. However gut microbiome of coral fish are still largely unknown (but see (Miyake, Ngugi, and Stingl 2015) on Acanthuridae from Red Sea).

Table 2: Effect of fish ecological traits on the structure of skin microbial communities

<table>
<thead>
<tr>
<th>Ecological Trait</th>
<th>Unweighted Unifrac</th>
<th>Weighted Unifrac</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>% of sign. P-values</td>
<td>Mean R²</td>
</tr>
<tr>
<td>Diet</td>
<td>97.9</td>
<td>0.16 ± 0.01</td>
</tr>
<tr>
<td>Size</td>
<td>23.2</td>
<td>0.10 ± 0.00</td>
</tr>
<tr>
<td>Schooling</td>
<td>5.1</td>
<td>0.09 ± 0.00</td>
</tr>
<tr>
<td>Mobility</td>
<td>4.7</td>
<td>0.05 ± 0.00</td>
</tr>
<tr>
<td>Position</td>
<td>8.8</td>
<td>0.05 ± 0.00</td>
</tr>
<tr>
<td>Activity</td>
<td>0</td>
<td>0.04 ± 0.00</td>
</tr>
</tbody>
</table>

Results of the PERMANOVAs-based analysis assessing the effect of fish ecological traits performed on 999 subsampling replicates containing one individual per fish species (see Supplementary Methods for more details). Mean and associated standard deviation of R-squared values on 999 subsamples are provided, as well as the percentage of subsamples where the correlation was significant (P<0.05).

Third, fish-associated microbial communities closely interact with immune system of their host, eliciting immune responses and educating immune system (Kelly and Salinas 2017). Malmstrom et al. (Malmstrøm et al. 2016) recently revealed that the number of copies of histo-incompatibility genes MHCI and MHCII, which encode proteins that detect non-self antigens and trigger an immune response, varies drastically among teleost fishes and do not perfectly fit phylogeny. In addition, differences in skin immunology could occur between individuals (e.g. between starved and nourished individuals (Caruso et al. 2010), between healthy and infected individuals (Lindenstrøm, Secombes, and Buchmann 2004), and between juveniles and adults (Grøntvedt and Espelid 2003)). Therefore, difference in immune system of fish could explain the high levels of both intra- and interspecific variability in skin microbiome as well as the absence of phylogenetic signal.
Microbial diversity hosted by fishes is vulnerable to global change

Climate change and fishing pressure have been the main drivers of decline of marine fish populations for the last 30 years (Jones et al. 2004; Graham et al. 2008; Bellwood et al. 2004; Myers et al. 2007). Therefore, any microbial diversity associated to fish species that are vulnerable to these threats is vulnerable too. We found that phylogenetic richness of skin microbiome was significantly correlated to host’s vulnerability to fishing (as estimated by (Cheung, Pitcher, and Pauly 2005) ; Spearman’s correlation test, P<0.05 in 95.1% of subsamples, rho=0.43± 0.08, n=999). Hence, the 10% most vulnerable fish species included in our study (vulnerability score ranging from 36 for Pterois radiata to 79 for Sphyraena barracuda) hosted a phylogenetic richness 1.5 times higher than the 10% less vulnerable species (vulnerability ranging from 10 for Corythoichthys flavofasciatus to 25 for Myripristis violacea). However, phylogenetic entropy of microbiome was not significantly correlated to vulnerability to fishing (P<0.05 in 47.9% of subsamples, rho=-0.33±0.07, n=999). Thus, the fish species that are the most threatened by fishing host OTUs that are phylogenetically unique and rare. Decline of these fishes would therefore be accompanied by the loss of such microbial diversity from coral reef. Neither phylogenetic richness, nor phylogenetic entropy, was significantly correlated to vulnerability to habitat loss (S12).

We here report that the high macroscopic fish biodiversity in a coral reef supports a high biodiversity of microbial species because each fish species host a high and unique diversity of microbes. In addition, the absence of phyllosymbiosis pattern related to these differences has important consequences for conservation of microbial diversity associated to fishes since protecting a few species of each clade does not prevent from losing a unique fraction of microbial diversity. Such a loss is even more likely under current fishing effort, which target species with the highest phylogenetic richness of microbes. These findings raise the need for a comprehensive assessment of the whole microbial biodiversity associated to coral reefs that are vanishing at an accelerated rate.

Authors’ contribution statement

Conflict of Interest

Authors have no conflict of interest to disclose.

Additional Information

This project was funded by the TOTAL Foundation. The funders had no role in study design, sampling and analysis decision to publish, or preparation of the manuscript.
5. Importance et vulnérabilité du microbiome cutané des animaux dans l’écosystème marin

*Many species of bacteria grow tenaciously attached to solid surfaces. [...] Exclusive of exotic species from the terrestrial environment along the littoral zone, most of the bacteria found in the sea appear to be associated with solid surfaces.*

Claude E. Zobell, 1943

---

5.1 Microbiome cutané et communautés planctoniques

Dès 1933, le microbiologiste américain, Arthur Henrici, écrivait, dans un contexte où prévalaient les cultures pures en milieu liquide, qu’« il est évident que la plupart des bactéries aquatiques ne sont pas des organismes flottants, mais croissent sur les surfaces immergées » (Henrici 1933). Dix ans plus tard, Claude Zobell a constaté que des bactéries introduites dans une bouteille en verre contenant un milieu liquide disparaissaient de la phase liquide, et que les parois de la bouteille présentaient une densité bactérienne croissante, démontrant les propriétés d’adhésion et de croissance des cellules microbiennes sur les surfaces (Zobell 1943).

Il est aujourd’hui reconnu que beaucoup d’espèces microbiennes sont capables de passer d’un mode de vie sessile à un mode de vie attaché à des surfaces sous forme de microcolonies ou de biofilms (Grossart 2010). Les biofilms sont des assemblages de microorganismes pouvant comprendre de nombreux taxons, englouis dans une matrice polysaccharidique sécrétée par les cellules elles-mêmes (Stoodley et al. 2002). La synthèse de ces polymères, associée à une différenciation des cellules, entraîne l’émergences de propriétés écologiques propres au système biofilm, telles que l’augmentation des transferts horizontaux de gènes, de la résistance au stress nutritionnel, aux molécules antimicrobiennes et à la prédation (Flemming et al. 2016). Depuis la publication de Zobell en 1943, de nombreux auteurs ont démontré expérimentalement que l’essentiel de l’abondance et de l’activité microbienne marine se situe au niveau des biofilms (*e.g.* (Haglund et al. 2002)). D’autres ont également démontré que les biofilms présentent généralement une diversité importante, que ce soit sur des surfaces inertes, ou à la surface d’animaux (Egan, Thomas, and Kjelleberg 2008; Sweet, Croquer, and Bythell 2011; Zettler, Mincer, and Amaral-Zettler 2013; Larsen et al. 2013; Apprill et al. 2014; Chiarello et al. 2015; Apprill 2017).

Malgré cela, l’essentiel des études portant sur la « diversité microbienne marine » se concentre sur la colonne d’eau (Sunagawa et al. 2015), ou dans une moindre mesure, sur le sédiment (Zinger et al. 2011). Ainsi, comme nous l’avons vu tout au long des chapitres 1 et 4,
les communautés microbiennes vivant à la surface des animaux marins sont encore largement méconnues. Il est nécessaire aujourd’hui d’amplifier l’effort actuel de recherche sur ces communautés de surface. En effet, elles jouent un rôle tout aussi important que les cellules planktoniques, sinon plus, dans le fonctionnement de l’écosystème marin. Non seulement elles présentent des taux métaboliques et reproductifs importants (e.g. biofilms, Grossart 2010), mais elles participent à la vie et à la santé des animaux peuplant l’écosystème marin (e.g. nutrition, résistance), et donc à leur potentiel de reproduction et d’adaptation.


5.2 Vulnérabilité des macro-organismes et de leurs microbiomes

Comme nous l’avons vu lors du chapitre 1, nous assistons aujourd’hui à une crise environnementale sans précédent au cours de laquelle les océans risquent d’être en première ligne, principalement à cause des pressions anthropiques, i.e. le réchauffement climatique, la surpêche et la pollution. Les modèles climatiques récents prédissent une augmentation moyenne de la température de l’océan de 2,7 degrés au cours du 21e siècle, ainsi qu’une acidification de l’eau, une diminution de l’oxygène dissous et une baisse de la productivité primaire (Bopp et al. 2013).


Chez les macro-organismes, l’effet des changements globaux est beaucoup mieux étudié. Des études ont reporté une diminution drastique de l’abondance (80 à 90%) des très grands organismes marins tels que les baleines, ainsi qu’une diminution de la répartition
géographique des grands prédateurs pélagiques (McCauley et al. 2015). Les écosystèmes coralliens sont particulièrement touchés par les changements globaux (Ban, Graham, and Connolly 2014), entrainant des épisodes de blanchiment massifs des coraux (Hughes et al. 2017), et des diminutions de l’abondance de nombreuses espèces de poissons dépendantes de ces habitats ont déjà été reportées (Wilson et al. 2006). Dans certaines régions, on enregistre même l’apparition de *dead zones*, vides de tout macro-organisme à cause du manque d’oxygène (Diaz and Rosenberg 2008).

Que deviennent alors les microorganismes associés aux espèces macroscopiques disparues ? Quelle est la conséquence de la perte de diversité macroscopique sur la diversité microscopique de l’écosystème ? Ces questions sont aujourd’hui sans réponse.

**5.3 Objectifs du chapitre et principaux résultats**

L’objectif de ce chapitre était donc de mesurer pour la première fois la contribution des communautés microbiennes associées à la surface des macro-organismes marins à la diversité microbienne totale d’un écosystème. Pour cela, nous nous sommes focalisés sur un écosystème corallien au sud de l’île de Mayotte, France. En une campagne d’échantillonnage, nous avons prélevé le microbiome de surface des espèces les plus abondantes, soit 138 poissons téléostéens, 84 anthozoaires, et 43 autres invertébrés tels que des mollusques, des échinodermes et des crustacés (voir chapitre 2).

J’ai tout d’abord comparé la diversité microbienne portée par ces animaux à celle présente dans l’eau les environnant. J’ai démontré que le microbiome des animaux marins présente la plus grande richesse phylogénétique, et que celle-ci comporte des lignées phylogénétiques qui ne sont pas retrouvées dans l’eau de mer. Au contraire, les communautés planctoniques présentent une très faible diversité propre, et partagent donc l’essentiel de leur diversité avec les autres compartiments de l’écosystème.

Ensuite, j’ai démontré que la diversité totale du microbiome associé aux animaux était favorisée par une importante dissimilarité entre taxons, et qu’il n’existait pas de microbiome « typique » des poissons ou des coraux.

Enfin, j’ai étudié l’impact d’une perte de diversité macroscopique sur la diversité microbienne globale en simulant la disparition des espèces de coraux durs les plus sensibles au réchauffement climatique, et celle des espèces de poissons les plus sensibles à la surpêche et à la perte d’habitat liée à la perte des coraux. J’ai démontré qu’une perte totale de ces espèces entraînait *in fine* une perte de 46% de la diversité microbienne totale de l’écosystème.

Ces résultats sont détaillés dans le manuscrit E.
5.4 Manuscrit E

Macroscopic biodiversity of coral reefs supports a high, unique and vulnerable microscopic biodiversity

Chiarello M.¹, Auguet J-C.¹, Claverie T.¹,², Sucré E.¹,², Bouvier C.¹, Rilleuvineuve F.¹, Nicholas AJ Graham³, Bettarel Y¹, Villéger S.¹ and Bouvier T.¹

¹Marine Biodiversity, Exploitation and Conservation (MARBEC), Université de Montpellier, CNRS, IRD, IFREMER, Place Eugène Bataillon, Case 093, 34 095 Montpellier Cedex 5, France
²Centre Universitaire de Formation et de Recherche de Mayotte, Route nationale 3, BP53, 97660 Dembeni, France
³Lancaster Environment Centre, Lancaster University, Lancaster, UK

En préparation

Des informations supplémentaires complètent ce manuscrit, disponibles en section 7.6
ABSTRACT

Marine environments host a huge diversity and biomass of microorganisms especially within the water column as well as on surfaces and in gut of macro-organisms. While microbial communities inhabiting living surfaces have crucial roles for host’s fitness, most of their diversity, including variability between species and animals groups, is still unknown especially in species-rich ecosystems such as coral reefs. Here, we assessed the contribution of various living surfaces, consisting in 138 fish samples, 84 hard, soft corals and anemones, and 43 other mobile and sessile invertebrates on the total microbial diversity in a well-preserved coral reef ecosystem, as well as vulnerability of this microbial diversity to loss of macroscopic diversity due to anthropic changes. We demonstrated that animal surfaces host most of total reef microbial diversity because surface microbiomes are rich and species-specific. Hence, a removal of all macro-organisms included in our study would induce a loss of more than 80% of prokaryotic phylogenetic richness in the ecosystem.
INTRODUCTION

Marine ecosystems provide habitats supporting a high abundance and high diversity of micro-organisms, including viruses, Archaea, bacteria and microeucaryotes (Kirchman 2010). The numerous studies of eukaryotic and prokaryotic microbial diversities from the water column (Bork et al. 2015) have demonstrated that marine microbes are functionally diverse, including auto- and heterotrophic organisms, primary producers and grazers and viruses, and are consequently key players of biogeochemical cycles occurring in the oceans (Kirchman 2010).

However marine ecosystems also include diverse biotic microhabitats were microbes are abundant, that are surface of living animals, including sessile invertebrates such as corals, and mobile vertebrates like fishes (McFall-Ngai et al. 2013). Scleratinian corals, for instance, besides their microalgal symbionts, host from $10^6$ to $10^8$ microbial cells per milliliter of their surface mucus, which can represent thousands of bacterial Operational Taxonomic Units (OTUs) (Bourne, Morrow, and Webster 2016). Corals also host Archaea and viruses, demonstrating the high phylogenetic diversity of coral-associated microbial communities (Nguyen-Kim et al. 2015; Frade et al. 2016). Those communities play crucial roles for their hosts, especially mediating resistance to pathogens (Krediet et al. 2013) and tolerance to environmental perturbations (Glasl, Herndl, and Frade 2016), likely including thermal stress (Ziegler et al. 2017). Moreover, members of coral-associated microbial communities include several ecosystem-significant genera involved in N, C and S cycling, which are participating to coral’s nutrition, and hence to marine ecosystem functioning (Bourne, Morrow, and Webster 2016).

Despite the potentially significant roles of animal surfaces-associated communities in marine ecosystems, they have been scarcely investigated to date compared to planktonic and benthic microbial communities. In addition, most studies focused on corals and sponges (Sunagawa, Woodley, and Medina 2010; Thomas et al. 2016), and to a lesser extent to some mollusks of economical interest (e.g. (Lokmer et al. 2016)). The microbiome of other invertebrates, i.e. echinoderms, crustaceans and cnidarians other than scleratinian corals, is still largely unknown. Similarly, if several studies have investigated the microbiome of teleostean fishes (Ghanbari, Kneifel, and Domig 2015), most of them focused on the gut microbiome (Miyake, Ngugi, and Stingl 2015; Givens et al. 2015). By contrast, the surface microbiome of fishes has been poorly investigated (e.g. (Larsen et al. 2013; Chiarello et al. 2015) for temperate ecosystems) and never for species inhabiting tropical reefs. Like surface microbiome of
marine invertebrates, surface microbiome of teleostean fishes is distinct from planktonic communities (Chiarello et al. 2015), diverse and species-specific (Larsen et al. 2013).

Most of marine ecosystems host high biodiversity of marine macro-organisms, which have diverse biology (presence of scales or not, nature of the skin, composition of the mucus, release of nutrients and antimicrobial compounds) and ecology (e.g. mobility, diet), and hence contrasted habitats for micro-organisms which could eventually promote a global high surface-associated microbial diversity at the ecosystem level. Therefore, it is urgent to test which proportion of marine microbial biodiversity is actually living on macro-organisms. This statement is particularly true for coral reefs, which host more than two thirds of the marine teleostean biodiversity (Kulbicki et al. 2013; Nelson, Grande, and Wilson 2016), and >1 millions invertebrates species (Stella et al. 2011). These ecosystems play crucial roles for global oceanic ecosystem services, generating primary production sustaining more than 25% of all marine species and providing food and protection for millions of people (Peterson and Lubchenko 2012). These ecosystems are vulnerable to anthropic changes, especially ocean warming and acidification, which induced massive coral bleaching events all around the world during the last two years (Hughes et al. 2017). While a few studies have investigated the possible direct effects of anthropic pressures on host-associated microbiome in coral reefs (especially in corals, e.g. (Webster et al. 2016; Zaneveld et al. 2016)), the effect of an erosion of macroscopic biodiversity on ecosystem’s microbial biodiversity is still unknown.

The objective of this study was therefore to assess the contribution of host-associated surface microbiomes to the total microbial diversity of a healthy coral reef and the potential effects of an erosion of macroscopic diversity due to anthropic pressures on reef microbial diversity.

MATERIAL AND METHODS

Study area

Sampling was conducted on November 2015 (17th to 27th) on coral reefs around Mayotte Island (France), located in the Western part of the Indian Ocean. Mayotte lagoon is the third largest lagoon in the world and houses 195 km of coral reefs and more than 700 fish species (Wickel et al. 2014). Fish were sampled from two sites in the South West of the lagoon: a fringing reef (S12°54'17.46'', E44°58'15.72''), and the inner slope of the barrier reef (S12°57'33.72'', E45°04'49.38''). Both sites are far from cities, were at a good ecological state at the time of sampling with more than >50% of coral cover and abundant fish
communities including predators such as groupers and barracudas.

**Sampling procedure**

Sampling objective was to target the most abundant organisms from each of main animal groups within a radius of 50m around each site. Sampling was authorized by the Mayotte’s maritime authority (permit N°12/UTM/2015).

In order to avoid contamination during sampling; fishes were caught using speargun and hookline and killed immediately after capture by cervical dislocation (following the European directive 2010/63/UE). Fishes were handled by the mouth using a clamp and all participants wear gloves. After death, the fish was laid down, and skin microbiome was sampled by swabbing the entire untouched side of the animal using buccal swabs (SK-2S swabs, Isohelix, UK), with the exception of the head. Surface microbiomes of *Anthozoa* were sampled by taking out a 1 to 10 cm fragment of a colony out of the water and exposing to air, which causes mucus secretion. This mucus dripping from the coral surface contains the microbiota and was directly collected (Leruste, Bouvier, and Bettarel 2012). For the few corals that don’t secrete mucus when exposed to air, surface microbiome was sampled by swabbing the entire untouched surface of the specimen. Other invertebrates were caught on the bottom using gloves, and entirely swabbed on the boat for 1 minute. In the case of clams and hermit crabs, the swab was carefully inserted inside the shell before gently rubbing the surface of the living tissues of the animal for 30 s. All invertebrates were released after microbiome sampling.

A total of 138 fishes from 44 species and 22 families were sampled on the two sites. 84 colonies of hard and soft corals, gorgonians and anemones (referred collectively as *Anthozoa*) were sampled, belonging to 14 scleratinian genera, 3 octocorallians (soft corals and gorgonians) and one anemone species. We also sampled 43 other invertebrates, which belonged to crustaceans, echinoderms, mollusks and sponges. The number of individuals (or colony) sampled in each site varies from 1 to 9 (see Table 1 and Supplementary Information S1).
Table 1: Microbial communities sampled

<table>
<thead>
<tr>
<th>Coral reef compartment</th>
<th>Animal group</th>
<th>Number of species or taxa</th>
<th>Number of samples</th>
<th>Details</th>
</tr>
</thead>
<tbody>
<tr>
<td>Animal surface</td>
<td>Teleost fish</td>
<td>44</td>
<td>138</td>
<td>5 orders and 22 families</td>
</tr>
<tr>
<td></td>
<td>Anthozoa</td>
<td>18</td>
<td>84</td>
<td>14 hard coral genera, 2 soft coral genera, 1 gorgonidae and 1 anemone</td>
</tr>
<tr>
<td></td>
<td>Other invertebrates</td>
<td>12</td>
<td>43</td>
<td>2 crustaceans, 7 echinoderms, 1 mollusk, and 2 sponges</td>
</tr>
<tr>
<td>Planktonic</td>
<td>-</td>
<td>-</td>
<td>35</td>
<td>17 surface samples, and 18 samples from the bottom of the water column</td>
</tr>
</tbody>
</table>

Total 300

Details about animal taxa and number of individuals sampled for each taxon are provided in Supplementary Information S1.

Fishes were identified at species level. Corals were identified at genus level. Classification of other invertebrates was made at species level when possible (for crustacean, starfish, urchin, anemone), or at higher taxonomic levels for a few clades (one sea cucumber, comatule, hermit crab, brittle stars and giant clams).

Each day of sampling 200-mL seawater samples were collected at sea surface and at 30 cm from the bottom of the sea, stored in an electric cooler, and filtrated at the end of each day through a 47 mm 0.2 µm polycarbonate membrane (Whatman, Clifton, USA). The membranes were then placed in sterile cryotubes. A total of 35 water samples were collected during the campaign.

All samples were stored at -5°C in an electric cooler during the day and remained frozen until DNA extraction.

16S rDNA amplification and sequencing

Swabs and water membranes were incubated during 30 minutes at 37°C in 570 µL of lysis buffer from Maxwell® Buccal Swab LEV DNA kits (Promega Corporation, Madison, USA) and 2 µL of 37.5-KU,µL⁻¹ Ready-Lyse lysozyme™ (Epicentre Technologies, Madison, USA). Then, 30 µl of proteinase K (from manufacturer’s kit) were added and tubes were incubated
overnight at 56°C. The totality of the solution was then placed in the kit for extraction. In the case of coral mucus, two 570-µL mucus samples were incubated separately with 570 µL of lysis buffer, following the same protocol and extracted using two kits before pooling. DNA extraction was performed using the Maxwell® 16 Bench-top extraction system following manufacturer’s instructions, and eluted in 50 µL of elution buffer. The V3-V4 region of the 16S rDNA gene was amplified using the prokaryotic primers modified for Illumina sequencing 515F (5'-C TTT CCC TAC ACG ACG CTC TTC CGA TCT - GTG CCA GCM GCC GCG GTA A- 3') (Caporaso et al. 2011) and the modified version of 806R by Apprill et al. (Apprill et al. 2015) (5’ – G GAG TTC AGA CGT GTG CTC TTC CGA TCT - GGA CTA CNV GGG TWT CTA AT - 3’), with PuRe Taq Ready-To-Go PCR Beads (Amersham Biosciences, Freiburg, Germany) using 1 µL of extracted DNA and 10 µM µL⁻¹ of each primer as follows: initial denaturation at 94°C for 1 min, followed by 35 cycles of 94°C for 1 min, 55°C for 1 min and 72°C for 1 min, ending with a final extension at 72°C for 10 min. Equimolar amounts of sample DNA extracted from each reef type were separately pooled and sequenced in two separated runs by an external laboratory (INRA GeT-PlaGE platform, Toulouse, France) on an Illumina platform using the 2x 250 bp MiSeq chemistry. 7 PCR blanks were included in each sequencing run in order to assess the presence of contaminants, which were removed during subsequent steps of sequence processing.

Sequence processing and phylogenetic analyses
Sequence processing was performed following the Standard Operating Procedure of Kozich et al for MiSeq (Kozich et al. 2013), https://www.mothur.org/wiki/MiSeq_SOP, 2017) using Mothur (Schloss et al. 2009). After assembly of paired reads from each run, sequences of both runs were merged and sequences with an abnormal length (outside a range of 250-300 pb) were removed. Sequences were aligned along the SILVA reference database (Quast et al. 2013) release 128. Chimeras were removed using UCHIME (Edgar et al. 2011). Sequences were then classified using the SILVA reference taxonomy and the non-prokaryotic ones were removed. 5,828,480 sequences from 300 samples were kept after cleaning process, ranging from 2,443 to 43,504 sequences per sample. After this, 2,000 sequences were sub-sampled within each sample (as done in (Caporaso et al. 2011)) in order to correct the uneven sequencing efficiency among samples and hence carry further analyzes on unbiased estimates of microbial diversity. Sequences were then grouped into OTUs using a 97% cutoff parameter.
to obtain 18,486 OTUs from 300 samples. Chao’s non-parametric coverage computed using entropart R-package averaged 93±5% across all samples (Chao et al. 1988; Marcon and Hérault 2014).

The most frequent sequence for each OTU was selected as reference and added to the SILVA reference phylogenetic tree using ARB software (Ludwig et al. 2004). The full phylogenetic tree was then pruned using the ape R-package to remove all but the added sequences, while keeping the topology of the reference SILVA tree. A chronogram was then adjusted to the phylogenetic tree using PATHd8 (Britton et al. 2007). The divergence time between Archaea and Bacteria was fixed at 3.8 Ga. The minimum divergence time between Euryarchaeota and other Archaea was set to 2.7 Ga (Blank 2009), and the maximum age of apparition of Thermoplasmatales was set to 2.32 Ga (Blank 2009). The minimum age of apparition of Cyanobacteria was set to 2.5 Ga (Schirrmeister et al. 2013). The minimum divergence time between Rickettsiales and the rest of Alphaproteobacterial sequences was set at 1.6 Ga, as done by Groussin et al. (Groussin et al. 2017). Finally the divergence times between Chromatiaceae and other Gammaproteobacteria, was set to minimum 1.64 Ga (Brocks et al. 2005).

Assessing microbial biodiversity

The overall surface microbiome of each fish species, Anthozoa genus, or other invertebrates’ taxa, was computed as the mean relative abundance of each OTU across individuals from the same taxon (see S1). Phylogenetic diversity of each microbiome was described with two complementary indices (Chao, Chiu, and Jost 2010). Phylogenetic richness of each species was measured with Faith’PD which accounts only for position of OTUs on the phylogenetic tree (Faith 1992) using picante R-package (Kembel et al. 2010), and then corrected by the total height of the prokaryotic tree in order to scale values between 0 and 1. Phylogenetic entropy, taking into account the relative abundance of OTUs as well as their position on the phylogenetic tree, was measured using Allen’s index (Allen, Kon, and Bar-Yam 2009), which is a phylogenetic extension of Shannon’s entropy. Allen index was computed using our own R-function (https://github.com/marlenec/chao, q=1) adapted from entropart R-package (Marcon and Hérault 2014). Faith’s index increases when communities are composed of OTUs that are phylogenetically distant, while Allen’s index increases when their most abundant OTUs are phylogenetically distant.
Phylogenetic dissimilarity between all pairs of microbiomes was assessed using two complementary indices: the relative abundance weighted and unweighted versions of the Unifrac index using the *GUniFrac* R-package (Chen and ORPHANED 2012). Abundance unweighted phylogenetic dissimilarity indices ranges from 0 when assemblages share the same phylogenetic lineages to 1 when assemblages are made of phylogenetically distant OTUs, while weighted dissimilarity indices tend to 0 when assemblages share the same dominant phylogenetic lineages and tend to 1 when assemblages are dominated by phylogenetically distant OTUs.

**Statistical tests**

The comparison of phylogenetic richness and entropy values obtained from each ecosystem compartment (planktonic and animal surface microbiome) was done using a Kruskal-Wallis test (999 permutations) in *vegan* R-package. To assess the contribution of each compartment to global microbial diversity in the ecosystem, two analyses were performed. The first analysis consisted in selecting OTUs that were unique from each compartment, *i.e.* recovered in only in planktonic or animal surface-associated communities, and calculating the phylogenetic richness associated to these OTUs. The second analysis consisted in accumulation curves, which were obtained by randomly picking from 1 to 75 animal taxa and 1 to 35 planktonic communities and computing the total phylogenetic richness for each set of communities. This procedure was performed 100 times and permitted to compare level of phylogenetic richness for a given number of communities. Additionally to this visual analysis, we also estimated the total richness of planktonic and animal surface-associated communities, by computing the Chao1 estimator on the pooled OTUs raw counts across all communities in both compartments (Chao 1984) using *fossil* R-package.

Significant difference in microbial structure and composition between planktonic and animal surface-associated microbial communities was assessed using two Permutational multivariate ANOVAs (PERMANOVAs) performed respectively on abundance unweighted phylogenetic dissimilarities (U-Unifrac) and weighted dissimilarities (W-Unifrac). To assess if planktonic communities were more variable than surface-associated communities, a permutation analysis of dispersion (PERMDISP) was performed on both weighted and unweighted dissimilarities among planktonic communities and among animal surface-associated communities using the *vegan* R-package.
Significant differences between surface microbiomes associated to the three major animal types (44 teleostean fish species, 24 Anthozoan genera and 12 other invertebrates’ taxa) were assessed using PERMANOVAs performed on both unweighted and weighted dissimilarities. Last, the effect of animal’s taxon to its associated surface microbiome was assessed on the full dataset of a given animal type (138 fishes, 84 Anthozoan colonies, or 43 individuals in “other invertebrates” group) using a separated PERMANOVAs for each animal type (teleostean fishes, Anthozoa and other invertebrates) performed on both unweighted and weighted dissimilarities.

In order to identify bacterial clades that are different between planktonic and animal surface associated communities, and consistent among water samples or animal taxa (i.e. consistently retrieved in all subclasses of the same class), we performed a LefSe analysis (Segata et al. 2011), using respectively water samples and animal taxa as subclasses. Then, to identify biomarkers for main animal groups (teleostean fishes, Anthozoa and other vertebrates), a LefSe analysis was performed using these groups as main classes, and the different animal taxa belonging to these groups (Fish species, Anthozoan genera and other invertebrates’ taxa) as subclasses.

Tests performed on abundance weighted and unweighted diversity and dissimilarity indices revealed overall congruent results (see S2).

Vulnerability of microbial diversity scenarios

To assess the vulnerability of reef microbial diversity to the loss of macro-organisms we simulated an extinction scenario combining the effects of global warming and overfishing. More precisely, macro-organisms species were removed proportionally to their respective vulnerability to heat stress for hard corals, and their vulnerability to habitat loss (due to coral bleaching) plus their vulnerability to fishing for fishes, using 100 replicates. Vulnerability to overfishing (Cheung, Pitcher, and Pauly 2005) was obtained from Fishbase (http://www.fishbase.org/, 2017). Cheung’s vulnerability to fishing is based on fish’s total length, life history traits (including age at first maturity, longevity, growth rate, fecundity, spatial behaviour), and geographical range. Vulnerability to habitat loss due to global change in coral reefs was computed as done in (Graham et al. 2011) using data from Fishbase.org and expert knowledge for input variables, which were diet specialization, habitat specialization, recruitment specialization for live coral and body size.
Vulnerability of scleratinian coral genus to global warming was based on their bleaching response in the Western Indian Ocean (McClanahan et al. 2007). The genus *Isopora*, for which we had no data, was excluded from the extinction scenario. Similarly, as no data was available for the other *Anthozoa*, as well as all other invertebrates, we excluded them from the scenario as well. Therefore, the vulnerability of microbial diversity reported hereafter is based on microbiome of 44 fish species and 13 genera of scleratinian corals.

At each level of extinction (5 to 100% of the most vulnerable scleratinian an fishes removed), the percentage of remaining diversity was compared to the one obtained from a random loss of the same number of corals and fishes, computed 100 times. A significant deviation from this random loss was assessed by computing P-value, calculated as the rank of the mean diversity value in the 100 replicates of the extinction scenario, among the increasingly sorted diversity values of all 100 replicates of the random scenario. We considered that the deviation from the random scenario was significant when $P<0.05$, meaning that the diversity value in the observed community was lower than the 5% lowest diversity values following the random loss of animal taxa.

**RESULTS**

**Comparison of animal surface microbiomes and planktonic communities**

On average, we recovered 201±33 (mean±SD) and 253±159 OTUs in planktonic and animal surface-associated communities, respectively. Animal surface-associated communities were the phylogenetically richer ones (Kruskal-Wallis, $P<0.001$ and associated post-hoc tests $P<0.05$) with one individual containing on average 6.6±1.2% of the total branch length of the phylogenetic tree grouping all OTUs recovered in the ecosystem, and planktonic community recovered in one 200-mL water sample contained and 3.2±1.4% of the total height of the tree, respectively (Figure 1 A, S2-Fig 1).
**Fig 1: Phylogenetic richness of reef microbial communities**

(A): Phylogenetic richness measured using Faith’s PD (expressed as proportion of total phylogenetic richness present in all samples) of animal surface-associated microbiomes (e.g. on one fish or coral colony) and of planktonic communities (i.e. in each water sample) Boxes represent the interquartile range of phylogenetic richness among communities. Thick bars represent the median of phylogenetic richness, and vertical segments extend to the most extreme data point that is no more than 1.5 times the length of the box away from the box. (B): total phylogenetic richness of all animal surface-associated microbiomes (i.e. after pooling all samples) and of all planktonic communities, and contribution of the unique (i.e. those recovered only on animals or only in water samples) and shared OTUs to phylogenetic richness.

The pool of all animal surface-associated communities contained higher OTU richness and phylogenetic richness than all planktonic communities combined (Figure 1 B, accumulation curves, S3). Chao’s richness estimator reached 29,321 OTUs in animal surface microbiome, which was 4.6 times the estimation for planktonic communities (6,395.4).

More importantly, the 13,825 OTUs that could not be found in planktonic communities made 75% percent of the total phylogenetic richness (Fig 1B, Unique fraction). By contrast, the 938 OTUs found only in planktonic communities made 8.8% of the total phylogenetic richness, which accounted for half of planktonic phylogenetic richness (Fig 1B). 2.5% of OTUs that were found only in animal surface-associated microbiomes belonged to the 5% most abundant OTUs in the same compartment, while 11.1% were moderately abundant OTUs (belonging to the 5-20% most abundant OTUs), and 56.5% were rare OTUs (belonging to the 50% less abundant OTUs). By contrast, none of the OTUs found only in planktonic communities belonged to the 5% most abundant ones. 3.1% belonged to the 5-20% most abundant OTUs, and 65.2% belonged to the 50% less abundant OTUs (Fig 2-A).
Fig 2: Dominance and taxonomic identity of OTUs unique to animal surface-associated microbiomes or planktonic communities. (A) Rank of abundance of unique OTUs from each compartment. Each bar represents the proportion of unique OTUs within each group of dominance (rank of abundance in decreasing order). (B) Heatmap representing the relative abundance of unique OTUs in each compartment (Animal surface and plankton). OTUs were clustered at family-level, and only families that made more than 1% abundance in at least one compartment are displayed.

OTUs unique to planktonic communities were affiliated to Flavobacteriaceae (0.15%), Rhodobacteraceae (0.06%), and the cyanobacterial Family I (0.01%). OTUs unique to animal surface-associated microbiomes were mostly affiliated to Flavobacteriaceae (1.61%), Rhodobacteraceae (0.58%) and Shewanellaceae (0.57%) (Fig 2-B). Overall most of OTUs unique to a type of community remained unclassified given current reference database, even at class level (65% of unique OTUs).

Variability among planktonic communities (U-Unifrac = 0.61±0.05) was significantly lower than variability among animal surface-associated communities (U-Unifrac=0.78±0.05, PERMDISP, P<0.001, and Fig 3 and S2).

Despite this high variability between taxa, the composition of animal surface-associated communities significantly differed from the planktonic ones (PERMANOVA, animal surfaces...
vs. planktonic communities, P=0.001, R²=0.15) (Fig 3-A). Subsequent LefSe analyses identified 101 significant biomarkers discriminating planktonic vs. animal surfaces-associated communities. Animal surface were especially enriched in unclassified Gammaproteobacteria, Alphaproteobacteria, and Firmicutes, while planktonic communities showed significantly higher abundances of Cyanobacteria (S4).

Fig 3: Phylogenetic dissimilarity between microbial communities
(A) Microbial communities plotted on the two first axes from a PCoA computed on pairwise unweighted Unifrac dissimilarities between all communities, the closer two points the lowest the dissimilarity in phylogenetic composition of microbial communities, and (B) average intra-and inter-group variability of biotic surfaces associated microbiome, measured with unweighted Unifrac. See S2-Fig 2 for analyzes of dissimilarity in phylogenetic structure using abundance-weighted Unifrac index.

Phylogenetic richness and dissimilarity between reef organisms
Variability among microbiomes of each group of macro-organisms averaged 0.76±0.05 between different fish species, 0.74±0.04 between Anthozoa genera, and 0.79±0.05 between other invertebrates’ taxa (U-Unifrac, KW and associated post-hoc tests, P<0.05 in all pairwise comparisons). Among each main animal group, there was a significant effect of fish species, Anthozoa genera and, to a lesser extent, other invertebrates’ taxa on surface-associated microbial community structure and composition (PERMANOVA, Table 2). Microbiomes of fishes, Anthozoa and other invertebrates were significantly different when tested by PERMANOVA (P=0.001, R²=0.06) (Fig3-A). However, the variability between fish- Anthozoa- and other invertebrates-associated microbiomes was only 1.02 times higher than
the variability between microbiomes from the same group of animals (U-Unifrac, Fig 3-B and W-Unifrac, S2-Fig 2). Accordingly LefSe could not identify any significant biomarkers for the surface microbiomes of these three main animal groups, which shared similar class-level composition (S5).

Table 2: Phylogenetic dissimilarity of surface-associated microbiomes

<table>
<thead>
<tr>
<th>Factor</th>
<th>U-Unifrac</th>
<th>W-Unifrac</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>P</td>
<td>R^2</td>
</tr>
<tr>
<td>Fish species</td>
<td>0.001</td>
<td>0.39</td>
</tr>
<tr>
<td>Anthozoa Genera</td>
<td>0.001</td>
<td>0.39</td>
</tr>
<tr>
<td>Other invertebrates</td>
<td>0.001</td>
<td>0.27</td>
</tr>
</tbody>
</table>

The effect of fish species, Anthozoa genera, and other invertebrates taxa were tested using three separated Permutational Multivariate ANOVAS (PERMANOVAs, 999 permutations) on surface-associated microbiomes of (i) the 138 teleostean fish samples, (ii) the 82 Anthozoa samples, and (iii) the 43 other invertebrates’ samples. Bold P-values indicate a significant effect of the tested factor.

Unraveling the core surface microbiome of reef organisms

No OTU was common to all animal taxa studied. Among the 17,483 OTUs found on animal surfaces, 62% were recovered in the surface microbiome of only one taxon, and only 121 OTUs (0.7% of OTUs) were found in at least half of animal taxa (38 taxa or more). These “core” OTUs were mostly affiliated to Gammaproteobacteria, especially Alteromonadaceae, Halomonadaceae and Vibrionaceae, Alphaproteobacteria, especially Rhodobacteraceae, and Flavobacteriaceae (S6), and ranged between 0 and 42% of abundance in a taxon. The 5 most abundant “core” OTUs across taxa were identified as Catenococcus (Vibrionaceae) (0-23% of abundance in a taxon), Alteromonadales including Aestuariibacter (0-17%), Cyanobacteria (0-42%) and unclassified Gammaproteobacteria (0-29%).

Vulnerability of reef microbial diversity to loss of macro-organisms

The most vulnerable fish species to the combined overfishing and habitat loss pressures were Chaetodon trifascialis, Chaetodon meyeri and Sphyraena barracuda. The most vulnerable coral genus to ocean warming in our dataset was Montastrea (McClanahan et al. 2007). There was no correlation between habitat loss of fish species and their surface OTU richness, phylogenetic richness or phylogenetic entropy (Spearman’s correlation test, P=0.77, 0.57 and 0.44, respectively). There was a marginally significant correlation between vulnerability to
fishing and fish surface microbial diversity (P=0.09, 0.09 and 0.07 and rho=0.25, 0.25 and 0.27 respectively for OTU richness and phylogenetic richness and entropy). However, there was no correlation between the combined vulnerability of fishing and habitat loss on fish surface microbial diversity (P=0.36, 0.20 and 0.70, respectively for OTU richness and phylogenetic richness and entropy). There was no correlation between coral’s vulnerability to warming and its associated OTU richness, and phylogenetic richness and entropy (P= 0.25, 0.48 and 0.64, respectively).

Extinction of the 5% most vulnerable coral and fish species would erase 2.6±1.2% of prokaryotic OTUs (S7) and 1.9±0.9% of microbial phylogenetic richness (Figure 4). A loss of the 50% most vulnerable coral and fish species would induce a loss of 22.4±3.3% of OTUs and 17±2.5% of phylogenetic richness. The slope of microbial diversity extinction increased slightly with increasing the loss of species (Fig 4 and S7). The local extinction of the 13 Scleratinians included in the analysis and the 44 fish species would lead to a remaining microbial OTU richness of 45% and a phylogenetic richness of 54% of their respective initial value at the ecosystem level (Fig 4 and S7). Simulated loss of microbial diversity following loss of the most vulnerable animals was however not significantly different from loss expected under random extinction of fish species and corals’ genera (Fig 4 and S7).
Fig 4: Vulnerability of coral reef microbial diversity to loss of fishes and scleratinian corals. (A) Black points and bars represent mean (±SD) of remaining microbial phylogenetic richness on the studied coral reef for a given proportion of coral and fish taxa lost. Fishes and corals loss were simulated according to their decreasing vulnerability to habitat loss and fishing, and bleaching, respectively (100 replicates). When 100% of coral and fish species are lost, the remaining microbial phylogenetic richness corresponds to the one of planktonic communities, and communities associated to the surface of animal taxa that were not included in the extinction scenario because their vulnerability was unknown (*i.e.* the scleratinian *Isopora*, soft corals, gorgonians, anemone and all non-anthozoan invertebrates). scenario simulating a random loss of fishes and corals (*i.e.* species randomly removed independently from their vulnerability), is illustrated with the grey area that represent range of remaining phylogenetic richness among 100 replicates.

(B) Slope of the loss in phylogenetic richness for the “warming+fishing” scenario, is represented as the mean ± SD of slopes calculated at each level of extinction on all 100 replicates.
DISCUSSION

Microbial diversity hosted by animal surfaces contributed to 93% of the whole prokaryotic OTUs richness found on studied coral reefs. Animal surface-associated communities reached a ca. 6-fold higher phylogenetic richness than planktonic communities (Fig 1-B). Prokaryotic OTUs found only on biotic surfaces covered 75% of the phylogenetic tree grouping all OTUs recovered from the ecosystem and were abundant on animal surfaces, while OTUs found only in seawater were mostly rare in the corresponding communities and phylogenetically redundant (Figure 2-A).

These results confirmed that the variety of reef animals’ surfaces provide habitats for a high prokaryotic diversity including for microbial lineages that are not detected in the water column.

Microbial communities covering abiotic surfaces (e.g. rocks) are usually distinct from those present in surrounding water column (Roth-Schulze et al. 2016). Indeed, colonizing surfaces requires several features, such as capacity of detecting molecular signals from such surfaces and motility to reach them (Dunne 2002), capacity of attachment to a surface (pili or extracellular polysaccharides synthesis) (Van Houdt and Michiels 2005), and capacity to grow as a biofilm, involving cell-to-cell signaling and differential gene expression (Vlamakis et al. 2013). Many bacterial species are able to switch between free-living and attached lifestyles (Grossart 2010), and some authors even argue that virtually all bacteria are able to grow as biofilms, which allows a higher activity compared to the planktonic phenotype (Donlan 2002; Dunne 2002; McDougald et al. 2011).

Marine animals have developed strategies to limit microfouling, using specific skin texture limiting bio-adhesion (Scardino and de Nys 2011), secreting a constantly renewing mucus that embed microbial cells and particles (Scardino and de Nys 2011; Ángeles Esteban 2012), or a gel that smooths the surface of the animal and dissolves bacterial adhesins (Baum et al. 2001), sloughing (Amy Apprill et al. 2014), and secreting antimicrobial compounds (Rakers et al. 2013). Therefore, to establish on a biotic surface, prokaryotes need to pass through these defenses from host, e.g. by degrading host’s surface mucus (Balebona et al. 1998), biding specific interactions (Grzeškowiak et al. 2011) and by evading host’s immune defense (Ryu et al. 2014). These characteristics could select specific prokaryotic clades, adapted to the interaction with their macroscopic host. Accordingly, prokaryotic communities recovered on animal surfaces in our study were significantly distinct from planktonic communities (Fig3-...
A). They were mostly enriched in *Gammaproteobacteria*, especially *Alteromonadales* and *Vibrionales* (S4). *Vibrionales* and *Alteromonadales* (in our study mostly *Aestuariibacter*, S4) are both marine bacteria, commonly found on corals, fish, crustaceans, shellfish and sponges (Nicole S Webster et al. 2010; Lee et al. 2012; Givens et al. 2013; Lokesh and Kiron 2016; Lokmer et al. 2016). The *Vibrionales*, particularly, are able to stick to and degrade chitin and mucin (Szabady et al. 2011; Boyd et al. 2015), both molecules that are frequent on shellfish and crustacean surfaces, and in fish and coral mucus.

Additionally to their overall high phylogenetic diversity, animal surface-associated communities were overall 1.3 times more variable than planktonic communities (Fig 3-A). The three major animal groups sampled, *i.e.* teleostean fishes, *Anthozoa* and other invertebrates hosted significantly different prokaryotic communities at their surface, but this factor only explained 6% of abundance unweighted phylogenetic dissimilarity (U-Unifrac) (see S2 for W-Unifrac). As a consequence no significant biomarker was identified for each of these groups, indicating that there is no typical microbiome for corals or teleostean fishes (Fig 3-A). For instance the two corals *Goniopora* and *Echinopora* were as dissimilar (U-Unifrac=0.76 and W-Unifrac=0.60) as *Echinopora* and the parrotfish *Scarus russelii* (U-Unifrac=0.75 and W-Unifrac=0.59). Nevertheless, despite high variability, fish species explained 39% of prokaryotic composition (tested on U-Unifrac) and 41% of prokaryotic structure (tested on W-Unifrac). Similarly, *Anthozoa* genera explained 39% of prokaryotic composition and 46% of prokaryotic structure (Table 2).

Hence, despite an absence of a specific microbiome for the three main groups of animal sampled, within each group, species did have significantly different microbiome. While species-specificity of surface microbiome has been already reported for species from the same groups of animals, including corals and teleostean fishes (Larsen et al. 2013; Frade et al. 2016), here we report for the first time that this interspecific variability within a clade is as high as between species from different clades that diverged ~800 Mya (divergence time estimate between *Teleostei* and *Anthozoa* according to http://www.timetree.org/). For instance, abundance weighted dissimilarity between the two parrotfishes *Chlorurus sordidus* and *Scarus caudofasciatus* (W-Unifrac=0.47) was higher than the one between *Chlorurus sordidus* and the scleratinian coral *Platygyra* sp (W-Unifrac=0.41).

Accordingly to the high variability of animal surface microbiome, rarefaction curve based on animal surface microbiomes showed that sampled microbial diversity is far from reaching its
asymptote (S3). Hence, reef microbial diversity is certainly even higher than the one reported here based on a small subset of macroscopic diversity living on the studied coral reef ecosystem (Mayotte lagoon hosts more than 700 fish species (Wickel et al. 2014)).

The loss of an animal species due to environmental disturbances at a given location will induce the loss of its associated microbial diversity. Scenario of microbial diversity erosion following macroscopic extinctions revealed that the loss of the 5% most vulnerable fish and scleratinian corals would induce a loss of only ~2% of the total reef prokaryotic OTUs and phylogenetic richness. A loss of the 50% most vulnerable coral and fish species induced equivalent losses of OTUs and phylogenetic richness (18 and 21%), indicating a moderate phylogenetic redundancy between animal surface-associated OTUs. A loss of all scleratinian corals (excepted *Isopora*) and fishes would leave 54% of total prokaryotic phylogenetic richness, corresponding to prokaryotic communities associated with seawater and invertebrates other than hard corals. These levels of prokaryotic diversity loss were not significantly different from the one expected under a random extinction scenario (S7). Indeed, the most vulnerable macro-organisms to the combined “fishing+warming” pressure, do not host a higher proportion of microbial diversity, while the most vulnerable fishes to fishing host phylogenetically richer microbiomes. For instance, the 10 most vulnerable species to fishing and habitat loss hosted a phylogenetic richness only slightly higher than the 10 less vulnerable fishes (7.7±3.9 vs. 6.9±4.1).

To our knowledge, no quantified vulnerability measures exist for invertebrates other than sessile *Anthozoa*, so we did not include them in our extinction scenario. Due to their wide phylogenetic range, spanning from crustaceans, echinoderms, mollusks, to sponges, microbiomes of invertebrates other than *Anthozoa* showed the highest variability (Fig3-B and S2), which suggests that loss of warming sensitive taxa of sponges, starfishes and sea urchin (Przeslawski et al. 2008; Byrne et al. 2009), or taxa targeted by humans such as crustaceans and mollusks (Polovina et al. 1995), would induce a more severe erosion of microbial diversity than the one simulated here. For instance in our system, a total removal of all living species would result to a remaining 16% of initial microbial phylogenetic richness in the reef, corresponding to the microbial diversity hosted by planktonic communities.

While the overall functional diversity supported by animal surface microbiome still needs to be addressed, it should be highlighted that some major members of their core microbiomes...
have important functional roles, as marine sulfur cycling (Catenococcus, (Podgorsek and Imhoff 1999; Sorokin 2003)), hydrocarbon degrading (Alteromonadaceae, (López-Pérez and Rodriguez-Valera 2014)) and photosynthesis (Cyanobacteria). Further metagenomic studies are therefore needed to investigate the functional roles of these communities associated to biotic surfaces in the ecosystem, and the potential impacts of macroscopic diversity loss to global microbial functional diversity.

Acknowledgments

We thank Emily Darling and Jérémie Vidal-Dupiol for helping us for identifying hard corals, and Frederic Ducarme for his help identifying echinoderms.

Conflict of Interest

Authors have no conflict of interest to disclose.

Additional Information

This project was funded by the TOTAL Foundation. The funders had no role in study design, sampling and analysis decision to publish, or preparation of the manuscript.
6. Conclusions et perspectives : variabilité, déterminants et importance du microbiome cutané des animaux marins

_Tu le sauras quand tu sauras que tu ne sais pas et que tu attendras de savoir._

Amadou Hampâté Bâ, 1969, à partir de récits traditionnels peuls

6.1 Mesures de la diversité et de la variabilité du microbiome


Chez les communautés de macro-organismes, les indices de diversité beta taxonomique et phylogénétique sont très corrélés, comme l’ont démontré Tucker et al. (2016) sur des simulations de paysages comprenant des communautés de 64 espèces. Durant ma thèse, sur des simulations de communautés microbiennes comportant de 40 à 2000 espèces, j’ai démontré qu’il existait des différences entre les indices de diversité, phylogénétiques et taxonomiques, prenant en compte ou non les abondances relatives des OTUs.

J’ai démontré que l’importante richesse des communautés microbiennes facilite à la fois la détection des règles d’assemblage, mais également la différence entre les communautés. Sur la figure 6.1 comparant l’efficacité de la détection des différences entre deux communautés proches, on observe en effet que la plupart des indices augmentent en efficacité chez les communautés de « type microscopique », à l’exception des indices phylogénétiques ne prenant pas en compte les abondances (e.g. l’indice Unifrac (Lozupone and Knight 2005)). Cette meilleure sensibilité des indices de diversité taxonomique et phylogénétique pour les communautés microbiennes s’explique vraisemblablement par le très grand nombre d’espèces, augmentant la quantité d’information incorporée dans le calcul de l’indice.

Globalement, chez les communautés de « type microscopique », les indices de diversité phylogénétique pondérés par les abondances relatives sont les plus performants, car ce sont ceux qui sont le moins impactés par les caractéristiques des communautés (richesse, équilibre des abondances). Dans le cas des indices de richesse phylogénétique ne prenant pas en compte les abondances, un très grand nombre d’espèces rares et partagées entre les
communautés peut par exemple empêcher de détecter les différences entre communautés de composition proche, mais de structure différente.

Ces résultats sont en accord avec d’autres publications ayant démontré que les indices non pondérés par les abondances ne permettent pas une bonne estimation de la diversité microbienne (Haegeman et al. 2013; Hill et al. 2003). De manière similaire, d’autres études avaient également souligné l’intérêt de prendre en compte les distances phylogénétiques entre OTUs pour l’étude des communautés microbiennes, sans réellement mesurer le gain en termes de sensibilité (Lozupone and Knight 2005; Veresoglu et al. 2014), bien qu’il atteigne tout de même 20% dans mon étude (Manuscrit A).

**Figure 6.1:** Pourcentage de détection d’une différence significative entre deux types de communautés proches sur l’arbre phylogénétique ; communautés composées respectivement de 40 OTUs (graphe du haut « Type Macroscopique » et 2 000 OTUs (graphe du bas, « Type Microscopique »). En abscisse, le déséquilibre des abondances des OTUs composant les communautés, indiquées en log(sd) de la loi log-normale utilisée pour la simulation des communautés. Plus le log(sd) est grand, plus le déséquilibre est important, et plus la proportion d’espèces rares est grande. Dans les communautés microbienues « naturelles », le log(sd) est généralement proche de 2 (Shoemaker, Locey, and Lennon 2017).
Les analyses portant sur les jeux de données issus de mes campagnes d’échantillonnage soulignent presque systématiquement des différences entre les indices de diversité phylogénétique pondérés ou non pondérés par les abondances relatives, différences qui m’ont permis d’accéder à une meilleure compréhension des patrons existants. Par exemple, dans le Manuscrit B, la prise en compte de l’abondance relative des OTUs dans le calcul des dissimilarités entre les microbiomes cutanés de poissons en captivité m’a permis de déterminer une différence entre les deux espèces étudiées, différence indécelable avec des indices non pondérés par les abondances (e.g. Unifrac ou Jaccard). A contrario, j’ai mis en évidence dans mon étude sur le récif de Mayotte (Manuscrits D et E), que les espèces de poissons les plus vulnérables à la surpêche présentaient une richesse phylogénétique microbienne plus importante que les espèces les moins vulnérables, et que ce patron n’était pas retrouvé lorsque l’on prenait en compte les abondances des OTUs. Ceci signifie que les poissons les plus vulnérables présentent des lignées phylogénétiques éloignées les unes des autres, dont les plus éloignées sont présentes en très faible abondance, soulignant la particularité de leur microbiome cutané.

L’ensemble de ces résultats confirme que pour avoir une vue exhaustive des communautés microbiennes, il est indispensable de combiner l’analyse de différents indices de diversité. Dans ce contexte, les approches unifiées proposées par Chao et al. (Chao, Chiu, and Jost 2014) et Chen et al. (Chen et al. 2012) sont particulièrement utiles car elles permettent de mesurer simultanément différentes facettes et/ou composantes de la diversité. La méthode de Chao, notamment, est basée sur les nombres équivalents d’espèces (Jost 2007), permettant la comparaison directe des valeurs entre les facettes taxonomique et phylogénétique de la biodiversité, et la partition de chaque indice en gamma-, alpha-, et beta-diversité. Cependant cette méthode est encore récente, et le manque d’implémentation d’outils pour calculer et interpréter ses résultats limite considérablement son utilisation. En particulier, cette méthode nécessite l’utilisation d’un arbre phylogénétique ultramétrique reliant les espèces microbiennes, ce qui nécessite une étape supplémentaire et délicate d’adaptation d’un chronogramme à l’arbre phylogénétique avant l’analyse (Paradis 2013). Cette étape prend énormément de temps sur les très grands arbres, caractéristiques des communautés microbiennes (selon ma propre expérience, de quelques jours à deux semaines sur un ordinateur portable, pour des arbres de 5 000 à 10 000 espèces) Cette limitation souligne la nécessité du développement d’outils phylogénétiques adaptés au monde microbiens, notamment en termes de temps de calcul et de mémoire nécessaire à la manipulation de ces énormes jeux de données.

6.2 Échelles de variabilité et déterminants

6.2.1 Variabilité intraspécifique du microbiome et déterminants potentiels

Dans cette thèse, j’ai décrit pour la première fois la variabilité intra-individuelle du microbiome cutané des poissons téléostéens et des cétacés. Celle-ci était presque aussi
élevée que la variabilité entre individus d’une même espèce (Tableau 6.2), en particulier chez les poissons. Chez les deux modèles testés, aucun effet reproductible de la zone du corps n’a été mis en évidence. Ce résultat contraste fortement avec ceux obtenus sur la peau humaine, où la structure du microbiome est très différentes entre les parties du corps et s’explique par des paramètres physico-chimiques du micro-environnement (i.e. principalement le pH et l’humidité) (Oh et al. 2014). Ce résultat suggère donc que contrairement à la peau humaine, le tégument des cétacés et la surface du poisson, sans être homogène, ne présenterait pas de micro-niches très différenciées comme c’est le cas chez l’Homme, ou alors que ces micro-niches existent, mais n’ont pas été échantillonnées.


<table>
<thead>
<tr>
<th>Manuscrit</th>
<th>Modèle</th>
<th>Ratio dissimilarité interspécifique / dissimilarité inter-individuelle</th>
<th>Ratio dissimilarité inter-individuelle / dissimilarité intra-individuelle</th>
</tr>
</thead>
<tbody>
<tr>
<td>B</td>
<td>Poissons téléostéens (captivité)</td>
<td>1.01</td>
<td>1.03±0.07</td>
</tr>
<tr>
<td>C</td>
<td>Cétacés (captivité)</td>
<td>1.03</td>
<td>1.5±0.14</td>
</tr>
<tr>
<td>D et E</td>
<td>Poissons téléostéens</td>
<td>1.3</td>
<td>Non mesurée</td>
</tr>
<tr>
<td>E</td>
<td>Anthozoaires</td>
<td>1.2</td>
<td>Non mesurée</td>
</tr>
</tbody>
</table>

**Tableau 6.2** : Ratio entre la dissimilarité phylogénétique moyenne pondérée par les abondances (voir chapitre 3) entre les espèces de vertébrés ou les genres de coraux comparées au cours des différentes campagnes d’échantillonnage réalisées durant la thèse, et la diversité beta moyenne entre individus d’une même espèce ou d’un même genre.
Au sein de tous les modèles que j’ai étudiés, j’ai trouvé une diversité intra-taxon importante (i.e. entre individus d’un même taxon). En effet, la dissimilarité entre taxons en captivité était à peine supérieure à la dissimilarité moyenne entre individus d’un même taxon (Tableau 6.2). Dans le cas des cétacés en captivité, la variabilité intra-spécifique était expliquée par l’âge et le sexe de l’individu chez les dauphins, mais restait inexpliquée chez l’orque. Dans le cas des poissons téléröstéens en captivité, aucun effet, à la fois de l’individu et de son sexe n’a été mis en évidence, mais une forte variabilité entre les zones du corps a été relevée.

Chez les individus prélevés dans même milieu naturel (Manuscrits D et E), j’ai également observé une variabilité inter-individuelle forte, avec un ratio variabilité interspécifique : variabilité inter-individuelle plus faible qu’en conditions contrôlées (e.g. ratio 1,3:1 dans le cas des poissons téléröstéens) ; suggérant que les conditions naturelles pourraient augmenter les différences interspécifiques du microbiome cutané, par exemple via des différences de régime alimentaire, qui n’existent pas en conditions contrôlées.


Cette importante variabilité inter-individuelle n’est donc pas une caractéristique du microbiome cutané, et pourrait être liée à des fortes variations de la réponse immunitaire selon les individus, par exemple de l’activité du lysozyme, une molécule du système immunitaire inné, qui chez les poissons téléröstéens dépend de nombreux paramètres dont l’âge, le sexe, le pH et la température de l’eau, mais également les facteurs de stress et la présence d’infections (Saurabh and Sahoo 2008). Chez l’homme, des chercheurs ont également démontré que des micro-variations génétiques, principalement dans des gènes du système immunitaire, étaient corrélées à la variabilité des microbiomes cutanés et digestifs (Blekman et al. 2015). Cette hypothèse pourrait se vérifier sur les cétacés en captivité, qui présentent des degrés d’apparentement différents ; elles nécessiteraient la recherche d’une corrélation entre de petites modifications du génome et les variations du microbiome. Une telle étude permettrait également d’identifier les gènes les plus corrélés à ces modifications, et donc de formuler des hypothèses concernant les facteurs proximaux qui conditionneraient la structure du microbiome cutané chez ces animaux.
6.2.2 Spécificité du microbiome et phylosymbiose

Dans l’ensemble des modèles étudiés (poissons et cétacés en captivité, poissons et invertébrés récifaux en milieu naturel), un effet important de l’espèce (ou du genre) hôte a été mis en évidence (Manuscrit B à E). Ce résultat confirme les études existantes, menées sur les poissons téléostéens (Larsen et al. 2013) et les coraux (Bourne, Morrow, and Webster 2016), démontrant une spécificité du microbiome de surface pour son hôte.

**Figure 6.2 :** Dissimilarité phylogénétique moyenne (indice UniFrac pondéré par les abondances des OTUs) en fonction de l’ancienneté de la divergence phylogénétique entre les espèces de poissons prélevées lors de la campagne S (A) ; et PCoA basée sur le même indice de dissimilarité, montrant l’ensemble des communautés microbiennes prélevées, i.e. planctoniques, et microbiomes de surface moyens par espèce de poisson, par genre d’anthrozoaire et par taxon d’invertébré non anthrozoaire (B). Sur le graphique (A), on peut voir que deux espèces de poissons ayant divergé il y a 40 millions d’années (ici les poisson-anges *Pygoplites diacanthus* et *Pomacanthus imperator*, W-Unifrac=0.80±0.17) portent des microbiomes aussi dissimilaires que deux poissons d’ordres taxonomiques différents (*Pygoplites diacanthus* et *Pterois miles*, W-Unifrac=0.73±0.10). Sur le graphique B, on observe qu’il existe un recouvrement partiel des différents types de microbiomes cutanés, faisant que certains poissons (ici *Chlorurus sordidus*) ont un microbiome plus proche de certains coraux (ici *Platygyra sp*) que d’autres poissons phylogénétiquement proches (ici *Scarus caudofasciatus*).

Cependant, j’ai également démontré sur 44 espèces de poissons téléostéens partageant le même milieu naturel (Manuscrit D) que cet effet de l’espèce hôte n’est pas lié à sa phylogénie, c’est à dire que deux espèces proches phylogénétiquement (par exemple deux poissons-anges) présentent des microbiomes aussi distincts que deux poissons phylogénétiquement éloignés (par exemple un poisson-ange et une rascasse) (Figure 6.2-A). Plus important, même s’il existe une différence statistique entre les grands clades d’animaux prélevés lors de la même campagne (poissons, anthrozoaires et autres types d’invertébrés), celle-ci est faible, et il n’existe pas vraiment de microbiome caractéristique de chacun des grands clades d’animaux marins. J’ai notamment relevé que les microbiomes de deux coraux...
n’étaient pas toujours plus proches que les microbiomes d’un corail et d’un poisson (Figure 6.2-B).

Ces résultats montrent des résultats sensiblement différents de ceux d’autres publications menées sur le microbiome digestif de divers vertébrés affirmant la présence d’un patron de phylosymbiose. Ils suggèrent que ce patron pourrait ne pas être universel à tous les clades de vertébrés, ou ne pas s’appliquer au microbiome de surface. Il faut cependant souligner que malgré le grand nombre de publications utilisant le terme de phylosymbiose, peu ont testé ce patron de manière non biaisée (Tableau 6.1).

Parmi ces études, une seule, focalisée sur les microbiomes digestifs de 33 espèces de mammifères terrestres (Groussin et al. 2017), a directement testé la corrélation entre les dissimilarités du microbiome et les distances phylogénétiques entre les hôtes. D’autres ont testé ce patron à l’aide de méthodes basées sur la reconstruction d’un arbre récapitulant les dissimilarités entre microbiomes (e.g. (Brooks et al. 2016)), qui souvent ne conservent pas fidèlement les dissimilarités entre communautés microbiennes (Maire et al. 2015). Certaines se sont basées sur un nombre insuffisant d’espèces (Brucker and Bordenstein 2012), ou encore sont imprécises quant aux méthodes utilisées (e.g. (Esson and Thacker 2014)). Enfin, certaines études (non répertoriées dans le tableau 6.1) rapportent un effet de la phylogénie de l’hôte sur le microbiome, sans avoir réalisé de réel test statistique (e.g. (Nelson et al. 2013)).

De plus, les mécanismes expliquant la présence d’un tel patron de phylosymbiose sont encore très loin d’être expliqués, en particulier chez les poissons. Dans la littérature, les patrons de phylosymbiose sont associés aux phénomènes de coévolution, alors que la coévolution implique des patrons et des processus différents (Brooks et al. 2016). La phylosymbiose décrit une corrélation entre les dissimilarités entre microbiomes et la phylogénie des hôtes. La coévolution fait référence à une congruence entre les distances phylogénétiques des hôtes et de tout ou partie de leurs symbiontes. Cependant, lorsque les dissimilarités entre microbiomes sont calculées à l’aide d’indices phylogénétiques, ces dissimilarités peuvent être considérées comme une distance phylogénétique moyenne entre tous les symbiontes composant le microbiome. Ainsi les études testant les patrons de phylosymbiose à partir de méthodes typiques des tests de coévolution (i.e. congruence d’arbres) mesurent une coévolution moyenne de tous les clades microbiens présents. Or il peut ne pas exister de patron de phylosymbiose alors même qu’une partie des symbiontes composant le microbiome co-évoluent avec leur hôte, si par exemple une partie importante des symbiontes est acquise de manière stochastique. A l’inverse, un patron de phylosymbiose n’implique pas nécessairement un phénomène de coévolution : le phénotype des hôtes est plus distinct s’ils sont très éloignés phylogénétiquement, et peut donc sélectionner des espèces microbiennes plus distinctes que chez des hôtes phylogénétiquement proches (Brooks et al. 2016).
<table>
<thead>
<tr>
<th>Modèle hôte</th>
<th>Clades</th>
<th>Analyses et résultats</th>
<th>Remarques</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mammifères terrestres</td>
<td>33 espèces</td>
<td>Test de corrélation de Mantel entre distance phylogénétique des hôtes et dissimilarités entre microbiomes : significatif.</td>
<td>L’intensité du signal phylogénétique clade par clade a pu être estimée via une reconstitution des microbiomes ancestraux, comparée à celle obtenue à partir d’un modèle nul.</td>
</tr>
<tr>
<td>Hominidés</td>
<td>5 espèces</td>
<td>Construction d’un arbre représentatif des dissimilarités entre microbiomes par une méthode de parcimonie, comparaison visuelle avec la phylogénie des hôtes : résultat positif.</td>
<td>Problème des méthodes basées sur la reconstruction d’un arbre récapitulant les distances entre microbiomes. Les méthodes de parcimonie sont sensibles à la saturation du signal et produisent des topologies parfois peu fiables.</td>
</tr>
<tr>
<td>Insectes et</td>
<td>5 clades, 31 espèces</td>
<td>Tests de congruence par permutation (basés sur des mesures de congruence) entre la phylogénie de chaque clade animal et cladogramme représentatif des dissimilarités entre microbiomes : significatif.</td>
<td>Problème des méthodes basées sur la reconstruction d’un arbre récapitulant les distances entre microbiomes.</td>
</tr>
<tr>
<td>Hominidés</td>
<td></td>
<td>Test de corrélation de Mantel entre les distances phylogénétiques entre hôtes et dissimilarité des microbiomes : significatif</td>
<td>Pas d’explication de la gestion des dissimilarités intra-spécifiques du microbiome, qui peuvent entraîner des faux positifs lors du test de corrélation.</td>
</tr>
<tr>
<td>Eponges marines</td>
<td>20 espèces</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pikas</td>
<td>4 populations</td>
<td>Comparaison visuelle entre le cladogramme des populations d’hôtes et cladogramme représentatif des dissimilarités entre microbiomes : résultat positif.</td>
<td></td>
</tr>
</tbody>
</table>

Chez les poissons, les mécanismes de mise en place du microbiome restent largement méconnus. Certaines études menées sur le microbiome digestif du poisson zèbre (*Danio rerio*) suggèrent que des phénomènes neutres de colonisation et de dispersion depuis le milieu environnant et entre larves expliqueraient principalement la composition du microbiome, puis que celui-ci se spécialiserait chez les larves plus âgées au fur et à mesure de la maturation de leur système immunitaire (Burns et al. 2016). Au sein de ce patron général, la place pour une éventuelle transmission verticale de certains symbiontes, malgré la présence de micro-organismes dans les œufs de certaines espèces de poissons (Brock and Bullis 2001), est inconnue. La recherche d’une telle transmission serait pourtant théoriquement possible, via par exemple l’introduction de bactéries recombinantes (marquées à l’aide d’une molécule fluorescente, *e.g.* la *Green Fluorescent Protein*) naturelles du microbiome de ces animaux dans des œufs par micro-injection (Goetz et al. 2001), puis le suivi de leur éventuelle croissance dans l’embryon, puis dans les tissus ou à la surface des larves et des adultes.

Pour l’instant, la possibilité d’une transmission verticale, et par extension, d’une possibilité de coévolution entre les lignées de microbes cutanés et leur hôte reste complètement inconnue. L’absence de signal phylogénétique malgré l’effet fort de l’hôte sur son microbiome peut être lié à différentes causes.

Tout d’abord, les facteurs liés au milieu naturel pourraient entrainer une hétérogénéité dans les conditions environnementales expérimentées par les animaux et brouiller la détection d’un signal phylogénétique. Cet effet pourrait être d’autant plus fort que le microbiome cutané se situe à l’interface entre son hôte et le milieu extérieur, et peut donc subir les effets directs (*e.g.* acidification du milieu) et indirects de celui-ci (*e.g.* modification de la physiologie du poisson suite à l’acidification du milieu). Dans le cas des poissons télésotéens que j’ai étudiés, cette hypothèse semble fragile car le site sur lequel ont été prélevés les animaux n’avait que peu d’impact sur leur microbiome cutané (manuscrit D). De plus, parmi les études ayant détecté un patron de phyllosymbiose dans le microbiome digestif, certaines ont travaillé sur des animaux élevés dans des conditions différentes (Brooks et al. 2016; Groussin et al. 2017), ce qui pourrait suggérer que les conditions environnementales ne masquent pas les patrons de phyllosymbiose.

Ensuite, le microbiome cutané pourrait dépendre de caractéristiques du poisson qui ne sont pas fortement liées à la phylogénie, ou qui évoluent très rapidement. Par exemple, des modifications rapides et importantes de la biologie de l’hôte (*e.g.* la taille moyenne à maturité, qui a beaucoup évolué ces 10 derniers millions d’années (Floeter et al. 2017)) peuvent entrainer des modifications profondes de leurs interactions avec les micro-organismes qui leurs sont associés, et ainsi entrainer des dissimilarités entre microbiomes de poissons d’espèces proches phylogénétiquement plus importantes qu’attendu étant donné leur temps de divergence. Dans le cas de notre étude, malgré une large gamme de temps de divergence (8 et 129 Ma), la majorité des espèces étaient en réalité assez éloignées...
phylogénétiquement (entre 80 et 129 Ma, voir figure 6.2). La prise en compte de davantage de paires d’espèces ayant divergé il y a moins de 10 millions d’années, présentant des traits biologiques plus proches, pourrait donc éventuellement faciliter la détection d’un patron de phylosymbiose chez ces poissons.

Dans le cadre de notre étude, nous avons testé l’impact des principaux traits écologiques des poissons, et n’avons relevé qu’un effet marginal du régime alimentaire sur la structure du microbiome. Cependant, d’autres traits pourraient entrer en jeu, tels que des modifications récentes du système immunitaire, qui n’ont pas été prises en compte dans notre étude. Un bon moyen de tester cette hypothèse serait de séquencer les principaux gènes du système immunitaire des poissons échantillonnés (e.g. gènes codant les récepteurs de l’immunité innée, les peptides antimicrobiens, ou encore des immunoglobulines) et de réaliser une phylogénie à partir des séquences obtenues, puis de tester la corrélation entre ces « distances phylogénétiques immunitaires » et le microbiome cutané.

En conclusion si des patrons de phylosymbiose existent chez de nombreux clades animaux, le microbiome cutané des poissons téléostéens marins semble être une exception. Afin d’écarter complètement l’hypothèse d’une variabilité environnementale, il serait cependant nécessaire de tester ce patron en conditions contrôlées, tout en s’assurant d’un nombre suffisant d’espèces pour pouvoir tester un tel patron. De plus, il serait intéressant d’inclure dans ce type d’analyse les poissons d’eau douce, présentant des particularités physiologiques, telles que la taille de leurs œufs, qui sont en moyenne plus gros que chez les espèces marines apparentées (Hardie and Hebert 2004), et des soin parentaux en moyenne plus importants (Breder 1966), pouvant favoriser les phénomènes de transferts verticaux des souches microbiennes. De plus, les capacités de dispersion des poissons d’eau douce sont plus restreintes que celles des animaux marins, suggérant qu’ils pourraient être colonisés par les mêmes espèces microbiennes (i.e. celles présentes dans le lac ou le cours d’eau) génération après génération, favorisant potentiellement les phénomènes de coévolution.

6.2.3 Quelles sont les échelles temporelles de variation du microbiome cutané ?

Le microbiome cutané des animaux marins est un assemblage dynamique, dont la composition et les activités varient au cours du temps. Ainsi, le microbiome de surface du corail présente des variations au fil des saisons (Li et al. 2014), mais également à l’échelle de quelques jours (Sweet et al. 2017). La partie variable du microbiome est souvent mise en opposition avec une partie stable que l’on appelle le core microbiome, et définie comme l’ensemble des microorganismes associés de manière récurrente à tous les individus d’une espèce hôte extraite d’un même habitat (Hernandez-Agreda, Gates, and Ainsworth 2017). Dans le cas des coraux, une étude a ainsi démontré que ce core microbiome ne concerne qu’une petite partie des espèces microbiennes retrouvées dans le squelette du corail (Hester et al. 2016).
Dans le cadre cette thèse, toutes les campagnes de prélèvement ont consisté en des échantillonnages ponctuels du microbiome cutané. En effet, l’objectif de cette thèse, face à un manque cruel de données dans la littérature sur le microbiome cutané des animaux marins, était de décrire le microbiome du plus large panel d’organismes possibles, et de mesurer l’ensemble des échelles de variation de ce microbiome (entre parties du corps, entre individus, entre espèces) à un moment donné. Nous avons donc fait le choix d’ignorer la variabilité temporelle du microbiome.

Les parties variable et stable du microbiome répondent certainement à des dynamiques très différentes. Ainsi, la partie variable du microbiome pourrait permettre à l’holobiontique de supporter les variations environnementales, via des ajustements des abondances des cellules microbiennes (Hester et al. 2016). Le core microbiome, au contraire, serait constitué de symbiontes adaptés à des niches à l’intérieur ou à la surface de l’hôte qui ne sont pas dépendantes des variations de l’environnement. Ces symbiontes ne semblent pas être inféodés à seule espèce de corail (Hester et al. 2016) ; cependant leur stabilité suggère que s’il existe une transmission verticale du microbiome répétée au fil des générations, il faudra certainement s’intéresser de plus près à ce core microbiome.

Dans le cas des poissons et des invertébrés prélevés dans le cadre de ma thèse, l’absence de suivi temporel du microbiome de ces animaux m’a obligée à considérer le microbiome en entier, i.e. sans distinction de la partie variable du microbiome, par rapport à sa partie stable (core microbiome). Cette absence de distinction a pu participer à l’importante variabilité de ce microbiome entre les individus d’un même taxon.

Les prochaines études portant sur les échelles de variabilité et les déterminants ne pourront plus faire l’économie d’un suivi temporel. Ainsi, lors d’une étude à laquelle j’ai collaboré (Résumé fourni en Annexe 8.7) menée sur des huîtres (Crassostrea gigas) infectées par le virus OsHV-1 qui cause d’importantes pertes dans les exploitations conchylicoles, les collaborateurs ont pu observer expérimentalement que cette infection causait une altération de la physiologie de l’huître et une immunodéficience, entrainant des modifications importantes de son microbiome, et ceci en l’espace de seulement quelques heures.

Ainsi des solutions techniques devront être développées pour l’étude des microbiomes cutanés en s’affranchissant de cette variabilité temporelle. En milieu naturel, elles peuvent impliquer des techniques de marquage, afin de repérer et prélever le microbiome du même individu plusieurs fois, et/ou des techniques de capture et d’immobilisation de l’animal ne perturbant pas le microbiome cutané. Ainsi, dans le cas des poissons, une étude de l’impact des principaux anesthésiques pour poissons sur le microbiome cutané sera nécessaire, car ils ont potentiellement des effets antimicrobiens (tricaine mesilate MS 222, benzocaïne et isoeugénol (Morrow and Berry 1988; Devi et al. 2010)). Enfin, dans le cas des mammifères
marins, qui sont difficiles à approcher, dont l’échantillonnage est particulièrement contraignant (Nelson et al. 2015), et dont la poursuite répétée en bateau peut perturber les populations naturelles (New et al. 2015), l’étude des animaux nés en captivité dans les différents parc existants pourrait être une bonne manière d’étudier la variation temporelle de leur microbiome. En effet, dans ces conditions, les animaux sont suivis médicalement et sont donc habitués à être manipulés par l’Homme ; de plus les conditions de captivité permettent de s’affranchir des variations environnementales sur le microbiome cutané, qui peuvent être très importantes en raison de leur grande mobilité (Durban and Pitman 2012).

6.3 Rôles des microbiomes cutanés dans les écosystèmes marins

6.3.1 Rôles fonctionnels potentiels

l’atmosphère, entrainant notamment la formation de nuages et participant à la régulation locale du climat (Bourne et al. 2013).

Cependant, les quelques études métagénomiques menées sur le microbiome de surface de coraux en milieu naturel se sont focalisées essentiellement sur le rôle du microbiome pour son hôte, et non pour le fonctionnement de l’écosystème (Wegley et al. 2007; Kimes et al. 2010).

Dans le cadre de ma thèse, j’ai démontré que la diversité phylogénétique associée à la surface des macro-organismes contribuait fortement à la diversité microbienne globale de l’écosystème corallien. Des études de métagénomique et de métatranscriptomique sont nécessaires pour déterminer si ce patron se retrouve également pour la diversité fonctionnelle, et si oui, quelles sont les fonctions uniquement portées par les microbes vivant sur les macro-organismes.

6.3.2 Contribution du microbiome à la diversité microbienne marine

Durant toutes les campagnes réalisées durant ma thèse, j’ai relevé une importante diversité du microbiome cutané des animaux (poissons, invertébrés, cétacés). Dans le cas des poissons télérosteens élevés en conditions contrôlées (Manuscrit B), j’ai démontré que les communautés microbiennes prélevées sur un seul poisson (i.e. correspondant à 4 nageoires) étaient équivalentes en termes de diversité phylogénétique à celles contenues dans 200 millilitres d’eau de mer. En captivité, les poissons ne baignent cependant pas dans une eau de mer « naturelle » : celle-ci est filtrée, ce qui peut réduire la densité et la diversité des communautés microbiennes planctoniques (Manes et al. 2011; Wold et al. 2014).

J’ai donc répété ce type de comparaisons en environnement naturel, comparant les communautés vivant à la surface de 44 espèces de poissons, 18 genres d’anthozoaires (coraux dur et mous, et anémones) et 12 autres types d’invertébrés. J’ai alors relevé que le microbiome présent à la surface d’un animal présentait une diversité du même ordre de grandeur que celle retrouvée dans 200 millilitres d’eau de mer. Lorsque l’on accumule l’ensemble de la diversité des microbiomes associés aux animaux, la diversité des microbiomes dépasse largement la diversité planctonique. En effet les communautés cutanées, en plus de leur richesse, sont aussi très dissimilaires entre elles, alors que les échantillons d’eau sont très similaires entre eux et contiennent peu de lignées phylogénétiques qui leur sont propres (Figure 6.2-B). Ainsi à l’échelle de l’écosystème corallien, mes résultats suggèrent que la majorité de la biodiversité microbienne se trouve à la surface des animaux. Les récifs coralliens sont des hot-spots de diversité macroscopique et donc des hot-spots de diversité microscopique.

Une précédente étude avait déjà fait le constat que des communautés microbiennes provenant de différents océans étaient plus semblables entre elles que le microbiome
prélevé au niveau d’une même partie du corps (bouche, estomac, anus, système pulmonaire) de mammifères (*Tursiops truncatus* et *Zalophus californianus*) vivant dans une même zone géographique (Bik et al. 2016). La forte dissimilarité des microbiomes entre individus, y compris en environnement contrôlé, associée à leur dynamique temporelle potentiellement contrastée, semblent souligner que les conditions fournies par les macro-organismes sont variables entre espèces et au cours du temps, et ne sont vraisemblablement liées qu’en partie à des variations environnementales.

La surface des macro-organismes fournit un habitat complexe, présentant en particulier de nombreux types de nutriments. Cet habitat va favoriser la croissance d’une grande diversité d’espèces microbiennes d’exigences trophiques contrastées, mais également favoriser leur diversification via le développement de mécanismes de défense contre les molécules antimicrobiennes de l’hôte (anticorps, lysozyme). Ces microbes, à cause du renouvellement de la peau ou du mucus, seront relargués dans l’eau de mer, qui contrairement à la surface cutanée, est un milieu oligotrophe. Beaucoup de ces micro-organismes présenteront alors une activité et une croissance réduite, alimentant potentiellement la banque d’OTUs dites « rares » dans le milieu marin (Troussellier et al. 2017) ; OTUs qui pourront alors coloniser d’autres organismes s’ils rencontrent cette opportunité, ou au contraire déperir. Le faible pourcentage de richesse phylogénétique associé à l’eau de mer que j’ai mis en évidence dans ma thèse semble accréditer cette hypothèse, la richesse phylogénétique associée spécifiquement aux matrices biotiques représentant 75% de la richesse phylogénétique totale de l’écosystème. Pour aller plus loin il serait nécessaire de rechercher dans l’eau de mer les taxons microbiens spécifiques de la surface des animaux par PCR quantitative (q-PCR, plus sensible que la PCR classique), afin de confirmer leur présence (leur absence totale militerait pour l’hypothèse de la transmission verticale de ce taxon), mais également de mesurer leur activité via une comparaison entre l’abondance du gène cible et de son transcrit, en réalisant des q-PCR sur l’ADN et l’ARN extraits d’un même échantillon.

Ainsi, en partant de l’hypothèse que les taxons microbiens spécifiques à la surface des animaux sont en réalité bien présents dans l’eau de mer, il serait intéressant de savoir s’ils font partie de la part « active » de la biosphère rare océanique, ou si au contraire il font partie de la part « dormante » des cellules microbiennes (Lynch and Neufeld 2015), et pourraient développer une activité lors de la rencontre avec un hôte macroscopique. Au contraire, une activité de ces taxons microbiens disproportionnée par rapport à leurs abondances constituerait un argument supplémentaire à la nécessité de prendre en compte les microbiomes des animaux dans les études de la biodiversité et du fonctionnement des communautés microbiennes marines.

S’il est clair que les macro-organismes hébergent une diversité microbienne importante à leur surface, il n’existe que peu d’études ayant mesuré leur importance en termes d’abondance dans l’écosystème. Chez les coraux, des études en microscopie confocale ou à épifluorescence ont permis de dénombrer 10^6-10^8 cellules par millilitre de mucus cutané, et
de l’ordre de $10^5$ cellules par centimètre carré de corail (Garren and Azam 2012; Nguyen-Kim et al. 2015; Thu et al. 2015).

Chez les poissons, des méthodes culturales ont permis de dénombrer de $10^2$ à $10^4$ bactéries par centimètre carré de peau (Benhamed et al. 2014), couverte par une couche de mucus de 30 à 100 µm d’épaisseur (Shephard 1994), ce qui donne une concentration bactérienne de l’ordre de $10^4$ à $10^6$ bactéries cultivables par millilitre de mucus cutané, soit du même ordre de grandeur que la densité microbienne totale retrouvée dans l’eau de mer (Kirchman 2010). Considérant que la proportion de bactéries cultivables est généralement faible, l’abondance réelle devrait être encore plus importante. Ceci suggère que le microbiome de surface des animaux marins pourrait contribuer de manière majeure à l’abondance microbienne totale existant dans les océans, qu’il reste à estimer via des méthodes microscopiques.

6.3.3 Vulnérabilité de la diversité microbienne face aux changements globaux

Le microbiome cutané des animaux présente une diversité forte et unique dans l’écosystème. Dans le récif corallien étudié au cours de ma thèse, les animaux présentaient une richesse microbienne vingt fois plus élevée que celle associée aux communautés planctoniques, et 75% de la richesse phylogénétique de l’écosystème n’était pas retrouvée dans l’eau de mer. Dans le cas des poissons téléostéens (manuscrit D), j’ai démontré que les espèces les plus sensibles à la pêche étaient également celles hébergeant la plus grande richesse phylogénétique. Par la suite, j’ai étudié l’effet d’une érosion de la diversité macroscopique sur la diversité microscopique de l’écosystème, et ait démontré qu’une élimination des coraux durs et des poissons entraîne une perte de la moitié de la richesse phylogénétique microbienne à l’échelle de l’écosystème, soulignant le fait que les différentes taxons étudiés ne présentent qu’une redondance partielle de leurs microbiomes.

La diversité globale du microbiome de surface des animaux est donc supportée par une importante dissimilarité entre les taxons étudiés.

Ce constat pourrait être renforcé par l’étude du microbiome cutané d’autres espèces présentées dans ces écosystèmes coralliens, en particulier les espèces les plus sensibles aux pressions anthropiques, telles que des mammifères, des requins, raies, dugongs et tortues. Ces espèces, ayant des stratégies écologiques particulières (e.g. très mobiles, consommateurs secondaires ou prédateurs) pourraient présenter des microbiomes particuliers dans l’écosystème. Un tel constat appuierait encore l’urgence d’une protection efficace de la biodiversité marine. Enfin, les microbiomes de surface des animaux marins jouent des fonctions essentielles dans la protection contre les infections et les perturbations environnementales (Krediet et al. 2013; Aprill et al. 2014; Lowrey et al. 2015; Glasl, Herndl, and Frade 2016; Mohamed Alipliah et al. 2016; Ziegler et al. 2017), et peuvent être déséquilibrées par des modifications anthropiques (Zaneveld et al. 2016; Webster et al.
Les espèces marines menacées sont également vulnérables aux maladies infectieuses, d’autant plus qu’elles présentent des petites tailles de populations (e.g. (Gaydos et al. 2004; Raverty et al. 2017)). Dans le cas des coraux, l’élévation des températures entraîne également l’apparition de micro-organismes opportunistes ou pathogènes pouvant potentiellement participer au blanchiment et à la mort du corail (Rosenberg et al. 2008). Dans un contexte de changement global, il paraît donc essentiel de mesurer les impacts des perturbations anthropiques sur l’abondance, l’activité et la diversité du microbiome cutané et leurs conséquences sur la physiologie de l’hôte et sur les populations d’animaux peuplant les écosystèmes marins.

L’objectif de cette thèse était d’explorer une diversité microbienne marine jusque là méconnue : celle associée à la surface des animaux marins. Pour répondre à cet objectif, j’ai étudié le microbiome de surface des animaux à plusieurs échelles. J’ai notamment étudié pour la première fois simultanément les variabilités intra- et inter-individuelle, et interspécifique du microbiome cutané des poissons téléostéens et des cétacés. J’ai démontré que ce microbiome de surface est dépendant de son espèce hôte, même lorsque les hôtes sont élevés en conditions contrôlées ou proviennent d’un même environnement, mais qu’il est également très variable entre individus et entre parties du corps. Enfin, j’ai démontré que les microbiomes de surface des animaux marins pourraient contribuer de manière significative à la diversité microbienne marine. De futures études expérimentales pourront évaluer le rôle que peuvent jouer ces microbiomes cutanés, à la fois pour leur hôte, mais également pour le fonctionnement de l’écosystème dans un contexte de changement global.
7. Références bibliographiques

A


Austin, Brian, and Dawn A. Austin. 2016. *Bacterial Fish Pathogens: Disease of Farmed and Wild Fish*. Springer.


Breder, Charles M. 1966. Modes of Reproduction in Fishes. 1st edition. Published for the American Museum of Natural History by the Natural History Press, Garden City, N.Y.


C


E


F


Garren, Melissa, and Farooq Azam. 2012. “Corals Shed Bacteria as a Potential Mechanism of


Graham, Nicholas A. J., Pascale Chabanet, Richard D. Evans, Simon Jennings, Yves Letourneur, M. Aaron MacNeil, Tim R. McClanahan, Marcus C. Öhman, Nicholas V. C.


Hentschel, Ute, Jörn Hopke, Matthias Horn, Anja B. Friedrich, Michael Wagner, Jörg Hacker,


J


K


L


doi:10.1007/978-3-642-38922-1_233.


M


Q


**R**


Roine, P., and Ca Elvehjem. 1950. “Significance of the Intestinal Flora in Nutrition of the


S


Salerno, Jennifer L., Brian W. Bowen, and Michael S. Rappé. 2016. “Biogeography of


Sørensen, Thorvald Julius. 1948. *A Method of Establishing Groups of Equal Amplitude in Plant Sociology Based on Similarity of Species Content and Its Application to Analyses of the Vegetation on Danish Commons*. København: I kommission hos E. Munksgaard.


Stella, Jessica S., Morgan S. Pratchett, Pat A. Hutchings, and Geoffrey P. Jones. 2011. “Coral-
Associated Invertebrates: Diversity, Ecological Importance and Vulnerability to Disturbance.” In Oceanography and Marine Biology: An Annual Review. CRC Press.


T


Taylor, Michael W., Peter Tsai, Rachel L. Simister, Peter Deines, Emmanuelle Botte, Gavin Ericson, Susanne Schmitt, and Nicole S. Webster. 2013. “‘Sponge-Specific’ Bacteria Are Widespread (but Rare) in Diverse Marine Environments.” The ISME Journal 7 (2): 438–43. doi:10.1038/ismej.2012.111.


V


Wickel, Julien, Alain Jamon, Mathieu Pinault, Patrick Durville, and Chabanet Pascale. 2014. “Composition et Structure Des Peuplements Ichtyologiques Marins de L’île de Mayotte (Sud-Ouest de L’océan Indien).” *Société Française D’Ichtyologie*.


Y


Z


8. Annexes

8.1 Synthèse bibliographique sur les déterminants du microbiome digestif du poisson

Cette synthèse comprend 51 études entre 2004 et 2017 ayant recherché l’effet de facteurs exogènes et/ou endogènes sur la structure ou la composition du microbiome digestif d’un total de 99 espèces de poissons téléostéens. Les résultats ont été synthétisés dans le tableau 1.2 du chapitre 1, et représentés dans leur intégralité ci-après. Pour chaque étude et pour chaque facteur, un « oui » signifie que le facteur avait un impact significatif sur le microbiome digestif des espèces concernées. Un « non » signifie que le facteur n’avait pas d’effet significatif. Un signe « - » signifie que le facteur n’a pas été testé. Pour plus d’information sur les techniques d’études utilisées, consulter le chapitre 2.

<table>
<thead>
<tr>
<th>1er Auteur</th>
<th>Date</th>
<th>Espèces</th>
<th>Technique d’étude</th>
</tr>
</thead>
<tbody>
<tr>
<td>Jensen</td>
<td>2004</td>
<td>Hippoglossus hippoglossus</td>
<td>Empreinte moléculaire + séquençage ADNr 165</td>
</tr>
<tr>
<td>Moran</td>
<td>2005</td>
<td>Kyphus sydneyanus</td>
<td>Empreinte moléculaire</td>
</tr>
<tr>
<td>Skrodenyte-Arbaciauskiene</td>
<td>2006</td>
<td>Ditrema termincki, Girella punctata, Pseudolabrus japonicus, Sebastes pachycephalus, Takifugu niphobles, Thalassoma cupido</td>
<td>Cultures + séquençage ADNr 165</td>
</tr>
<tr>
<td>Shina</td>
<td>2006</td>
<td>Salmo trutta fario</td>
<td>Cultures + clonage/séquençage ADNr 165</td>
</tr>
<tr>
<td>Navarrete</td>
<td>2006</td>
<td>Oncorhynchus kisutch (juvenile)</td>
<td>Cultures et empreinte moléculaire + séquençage ADNr 165</td>
</tr>
<tr>
<td>Ringa</td>
<td>2006</td>
<td>Gadus morhua</td>
<td>Cultures + Séquençage ADNr 165</td>
</tr>
<tr>
<td>Martin-Antonio</td>
<td>2007</td>
<td>Solea senegalensis</td>
<td>Cultures + séquençage ADNr 165</td>
</tr>
<tr>
<td>Kim</td>
<td>2007</td>
<td>Oncorhynchus mykiss</td>
<td>Empreinte moléculaire + Cultures + séquençage ADNr 165</td>
</tr>
<tr>
<td>Skrodenyte-Arbaciauskiene</td>
<td>2008</td>
<td>Salmo salar (juvénile), S. trutta (juvénile)</td>
<td>Cultures + séquençage ADNr 165</td>
</tr>
<tr>
<td>Sun</td>
<td>2009</td>
<td>Epinephelus coioides</td>
<td>Cultures + séquençage ADNr 165</td>
</tr>
<tr>
<td>Navarrete</td>
<td>2009</td>
<td>Salmo salar (juvénile)</td>
<td>Empriente moléculaire + séquence ADNr 165</td>
</tr>
<tr>
<td>He</td>
<td>2010</td>
<td>Oreochromis sp.</td>
<td>Empriente moléculaire + séquence ADNr 165</td>
</tr>
<tr>
<td>Tapia-Paniagua</td>
<td>2010</td>
<td>Solea senegalensis</td>
<td>Empriente moléculaire + séquence ADNr 165</td>
</tr>
<tr>
<td>Navarrete</td>
<td>2010</td>
<td>Salmo salar</td>
<td>Empriente moléculaire</td>
</tr>
<tr>
<td>Smriga</td>
<td>2010</td>
<td>Acanthurus nigricans, Chlorurus sordidus, Lutjanus bohar</td>
<td>Microscopie à épifluorescence</td>
</tr>
<tr>
<td>Mansfield</td>
<td>2010</td>
<td>Oncorhynchus mykiss</td>
<td>Empriente moléculaire + séquence ADNr 165</td>
</tr>
<tr>
<td>Roeseiders</td>
<td>2011</td>
<td>Dario rerio</td>
<td>Empriente moléculaire</td>
</tr>
<tr>
<td>Auteur</td>
<td>Année</td>
<td>Genre</td>
<td>Espèce</td>
</tr>
<tr>
<td>--------</td>
<td>-------</td>
<td>-------</td>
<td>--------</td>
</tr>
<tr>
<td>Cantas</td>
<td>2011</td>
<td>Salmo salar</td>
<td>(juvenile)</td>
</tr>
<tr>
<td>Silva</td>
<td>2011</td>
<td>Carassius auratus</td>
<td>Sparus aurata</td>
</tr>
<tr>
<td>Semova</td>
<td>2012</td>
<td>Dario rerio</td>
<td></td>
</tr>
<tr>
<td>Wu</td>
<td>2012</td>
<td>Ctenopharyngodon idellus</td>
<td></td>
</tr>
<tr>
<td>Wu</td>
<td>2012</td>
<td>Pelteobagrus fulvidraco</td>
<td></td>
</tr>
<tr>
<td>Zhou</td>
<td>2012</td>
<td>Gadus morhua</td>
<td></td>
</tr>
<tr>
<td>Geraylou</td>
<td>2012</td>
<td>Aciptenprer baeri</td>
<td></td>
</tr>
<tr>
<td>Liu</td>
<td>2012</td>
<td>Carassius auratus</td>
<td>Ctenopharyngodon idellus</td>
</tr>
<tr>
<td>Li</td>
<td>2012</td>
<td>Carassius auratus</td>
<td>Hypophthalmichthys molitrix</td>
</tr>
<tr>
<td>Sullam</td>
<td>2012</td>
<td>Notothenia coriiceps</td>
<td>Oncorhynchus mykiss</td>
</tr>
<tr>
<td>Sullam</td>
<td>2012</td>
<td>Pelteobagrus fulvidraco</td>
<td>Poccella reticulata</td>
</tr>
<tr>
<td>Navarrete</td>
<td>2012</td>
<td>Notothenia coriiceps</td>
<td>Oncorhynchus mykiss</td>
</tr>
<tr>
<td>Geraylou</td>
<td>2013</td>
<td>Aciptenprer baeri</td>
<td></td>
</tr>
<tr>
<td>Star</td>
<td>2013</td>
<td>Gadus morhua</td>
<td></td>
</tr>
<tr>
<td>Li</td>
<td>2013</td>
<td>Cyprinus carpio</td>
<td></td>
</tr>
<tr>
<td>Kühlwein</td>
<td>2013</td>
<td>Cyprinus carpio L.</td>
<td></td>
</tr>
<tr>
<td>Green</td>
<td>2013</td>
<td>Salmo salar</td>
<td></td>
</tr>
<tr>
<td>Ye</td>
<td>2014</td>
<td>Dorosoma cepedianum</td>
<td>Hypophthalmichthys molitrix</td>
</tr>
<tr>
<td>Larsen</td>
<td>2014</td>
<td>Ictalurus punctatus</td>
<td>Lepomis macrochirus</td>
</tr>
<tr>
<td>Xia</td>
<td>2014</td>
<td>Lates calcarifer</td>
<td></td>
</tr>
<tr>
<td>Li</td>
<td>2014</td>
<td>Ctenopharyngodon idellus</td>
<td>Megalobrama ambycephalo Carassius auratus</td>
</tr>
<tr>
<td>Givens</td>
<td>2015</td>
<td>Bairdiella chrysoura</td>
<td>Caranx hippos</td>
</tr>
<tr>
<td>-------------</td>
<td>------</td>
<td>----------------------</td>
<td>--------------</td>
</tr>
<tr>
<td>Schmidt</td>
<td>2015</td>
<td>Paeclila sphenops</td>
<td>-</td>
</tr>
<tr>
<td>Llewellyn</td>
<td>2016</td>
<td>Salmo salar</td>
<td>-</td>
</tr>
<tr>
<td>Parris</td>
<td>2016</td>
<td>Acanthochromis polycanthus</td>
<td>Apogonidae unknown</td>
</tr>
<tr>
<td>Hennersdorf</td>
<td>2016</td>
<td>Atule mate</td>
<td>Epinephelus fuscoguttatus</td>
</tr>
<tr>
<td>Bolnik</td>
<td>2016</td>
<td>Gasterosteus aculeatus</td>
<td>Perca fluviatilis</td>
</tr>
<tr>
<td>Sullam</td>
<td>2016</td>
<td>Paeclila reticula</td>
<td>-</td>
</tr>
<tr>
<td>Eichmiller</td>
<td>2016</td>
<td>Aplodinotus grunniens</td>
<td>Cyprinus carpio</td>
</tr>
<tr>
<td>Nielsen</td>
<td>2017</td>
<td>Siganus fuscens</td>
<td>-</td>
</tr>
<tr>
<td>Carlson</td>
<td>2017</td>
<td>Gambusia affinis</td>
<td>-</td>
</tr>
<tr>
<td>Michl</td>
<td>2017</td>
<td>Oncorhynchus mykiss (juvenile)</td>
<td>-</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Nombre d'études avec effet positif</th>
<th>Nombre total d'études</th>
</tr>
</thead>
<tbody>
<tr>
<td>13</td>
<td>13</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>15</td>
<td>15</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>6</td>
<td>6</td>
</tr>
<tr>
<td>8</td>
<td>8</td>
</tr>
<tr>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>13</td>
<td>13</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>15</td>
<td>15</td>
</tr>
<tr>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>7</td>
<td>7</td>
</tr>
<tr>
<td>8</td>
<td>8</td>
</tr>
<tr>
<td>5</td>
<td>5</td>
</tr>
</tbody>
</table>

Séquençage haut débit ADNr 165
8.2 Protocoles expérimentaux

8.2.1 Protocole d’extraction basée sur le kit Blood & Tissue (Qiagen)
Méthode adaptée de (Amalfitano and Fazi 2008) pour l’extraction de microbiome de surface de nageoires (Campagne 1)
Immerger la nageoire entière (si nécessaire découpée en morceaux) dans 6mL de PBS contenant 0.5% de Tween 20 (vol/vol). Vortexer à vitesse maximum durant 10 minutes (Vortex genie 2, Scientific Industries, Bohemia, USA). Filtrer la solution à travers une membrane polycarbonate 47 mm de diamètre et 0.2 µm de largeur de pore. Extraire l’ADN a partir du filtre obtenu à l’aide du kit Blood & Tissue® (Qiagen, Venlo, Netherlands), en suivant le protocole modifié pour la lyse des bactéries Gram positif fourni ci-après, et en éluant l’ADN obtenu dans 100 µL de tampon AF.

Méthode utilisée pour l’extraction du microbiome des cétacés et de la roussette (Campagnes 2 et 3)
- Cas des membranes polycarbonates pour les échantillons d’eau : suivre le protocole fourni ci-après et éluer l’ADN dans 100 µL de tampon AF.
- Cas des pointes mousse (Whatman) : la mousse absorbant beaucoup de liquide, il a fallu doubler les volumes des solutions initiales.
  Placer la pointe mousse dans 260 µL de tampon de lyse ;
  Après les 30 min d’incubation, ajouter 50mL de protéinase K et 200 µL de tampon AL.
  Puis suivre le protocole ci-après.
  Éluer l’ADN dans 100 µL de tampon AF.

1- Préparation du tampon de lyse pour les bactéries Gram positif :

For pretreatment of gram-positive bacteria (page 45)

- Enzymatic lysis buffer:
  - 20 mM Tris-Cl, pH 8.0
  - 2 mM sodium EDTA
  - 1.2% Triton® X-100
  - Immediately before use, add lysozyme to 20 mg/ml
2-Protocole Qiagen Blood & Tissue® modifié pour la lyse des bactéries Gram positif

Protocol: Pretreatment for Gram-Positive Bacteria
This protocol is designed for purification of total DNA from Gram-positive bacteria, such as Corynebacterium spp. and B. subtilis. The protocol describes the preliminary harvesting of bacteria and incubation with lysozyme to lyse their cell walls before DNA purification.

Important points before starting
- See “Quantification of starting material”, page 17, for details of how to collect and store samples, and how to determine the number of cells in a bacterial culture.
- Ensure that ethanol has not been added to Buffer AL (see “Buffer AL”, page 18). Buffer AL can be purchased separately (see page 56 for ordering information).
- This pretreatment protocol has not been thoroughly tested and optimized for high-throughput DNA purification using the DNeasy 96 Blood & Tissue Kit. As a general guideline, we recommend to decrease the amount of starting material when using this protocol with the DNeasy 96 Blood & Tissue Kit.

Things to do before starting
- Prepare enzymatic lysis buffer as described in “Equipment and Reagents to Be Supplied by User”, page 14.
- Preheat a heating block or water bath to 37°C for use in step 3.

Procedure
1. Harvest cells (maximum 2 x 10^8 cells) in a microcentrifuge tube by centrifuging for 10 min at 5000 x g (7500 rpm). Discard supernatant.
2. Resuspend bacterial pellet in 180 µl enzymatic lysis buffer.
3. Incubate for at least 30 min at 37°C.
   After incubation, heat the heating block or water bath to 56°C if it is to be used for the incubation in step 5.
   Note: Do not add proteinase K directly to Buffer AL.
   Ensure that ethanol has not been added to Buffer AL (see “Buffer AL”, page 18). Buffer AL can be purchased separately (see page 56 for ordering information).
5. Incubate at 56°C for 30 min.
   Optional: If required, incubate at 95°C for 15 min to inactivate pathogens. Note that incubation at 95°C can lead to some DNA degradation.

6. Add 200 µl ethanol (96–100%) to the sample, and mix thoroughly by vortexing.
   It is important that the sample and the ethanol are mixed thoroughly to yield a homogeneous solution.
   A white precipitate may form on addition of ethanol. It is essential to apply all of the precipitate to the DNeasy Mini spin column. This precipitate does not interfere with the DNeasy procedure.
4. Pipet the mixture from step 3 (including any precipitate) into the DNeasy Mini spin column placed in a 2 ml collection tube (provided). Centrifuge at \( \geq 6000 \times g \) (8000 rpm) for 1 min. Discard flow-through and collection tube.*

5. Place the DNeasy Mini spin column in a new 2 ml collection tube (provided), add 500 µl Buffer AW1, and centrifuge for 1 min at \( \geq 6000 \times g \) (8000 rpm). Discard flow-through and collection tube.*

6. Place the DNeasy Mini spin column in a new 2 ml collection tube (provided), add 500 µl Buffer AW2, and centrifuge for 3 min at 20,000 x g (14,000 rpm) to dry the DNeasy membrane. Discard flow-through and collection tube.

It is important to dry the membrane of the DNeasy Mini spin column, since residual ethanol may interfere with subsequent reactions. This centrifugation step ensures that no residual ethanol will be carried over during the following elution.

Following the centrifugation step, remove the DNeasy Mini spin column carefully so that the column does not come into contact with the flow-through, since this will result in carryover of ethanol. If carryover of ethanol occurs, empty the collection tube, then reuse it in another centrifugation for 1 min at 20,000 x g (14,000 rpm).

7. Place the DNeasy Mini spin column in a clean 1.5 ml or 2 ml microcentrifuge tube (not provided), and pipet 200 µl Buffer AE directly onto the DNeasy membrane. Incubate at room temperature for 1 min, and then centrifuge for 1 min at \( \geq 6000 \times g \) (8000 rpm) to elute.

Elution with 100 µl (instead of 200 µl) increases the final DNA concentration in the eluate, but also decreases the overall DNA yield (see Figure 2, page 21).

8. Recommended: For maximum DNA yield, repeat elution once as described in step 7.

This step leads to increased overall DNA yield.

A new microcentrifuge tube can be used for the second elution step to prevent dilution of the first eluate. Alternatively, to combine the eluates, the microcentrifuge tube from step 7 can be reused for the second elution step.

Note: Do not elute more than 200 µl into a 1.5 ml microcentrifuge tube because the DNeasy Mini spin column will come into contact with the eluate.

* Flow-through contains Buffer AE or Buffer AW1 and is therefore not compatible with bleach. See page 8 for safety information.
8.2.2 Protocole basé sur le kit PowerSoil® (MoBio)

- Cas des pointes mousse (Whatman) : la mousse étant de grande taille et absorbant beaucoup de liquide, il a fallu adapter le début du protocole.
  Placer la mousse dans un eppendorf 5mL. Transférer dans cet eppendorf le contenu, billes et solution de lyse, de deux tubes bead beating fournis dans le kit. Vortexer.
  Ajouter 120 μL de solution C1, homogénéiser puis suivre le protocole ci-après à partir de l’étape 5 jusqu’à l’étape 13.
  A l’étape 14, diviser le volume du surnageant le surnageant dans 4 tubes différents, et ajouter 300 μL de solution C4 dans chaque tube.
  A l’étape 15, passer successivement le contenu des 4 tubes sur colonne, centrifuger et jeter l’éluat à chaque passage.

- Cas des coton-tiges isoHelix : placer directement l’embout coton dans le tube bead beating fourni et suivre le protocole ci-après.
Protocole PowerSoil® MoBio

EXPERIENCED USER PROTOCOL
PowerSoil® DNA Isolation Kit
Catalog No. 12888-50 & 12888-100

Please wear gloves at all times

1. To the PowerBead Tubes provided, add 0.25 grams of soil sample.

2. Gently vortex to mix.

3. Check Solution C1. If Solution C1 is precipitated, heat solution to 60°C until dissolved before use.

4. Add 60 µl of Solution C1 and invert several times or vortex briefly.

5. Secure PowerBead Tubes horizontally using the MO BIO Vortex Adapter tube holder for the vortex (MO BIO Catalog# 13000-V1-24) or secure tubes horizontally on a flat-bed vortex pad with tape. Vortex at maximum speed for 10 minutes.

Note
If you are using the 24 place Vortex Adapter for more than 12 preps, increase the vortex time by 5-10 minutes.

6. Make sure the PowerBead Tubes rotate freely in your centrifuge without rubbing. Centrifuge tubes at 10,000 x g for 30 seconds at room temperature. CAUTION: Be sure not to exceed 10,000 x g or tubes may break.

7. Transfer the supernatant to a clean 2 ml Collection Tube (provided).

Note
Expect between 400 to 500 µl of supernatant. Supernatant may still contain some soil particles.

8. Add 250 µl of Solution C2 and vortex for 5 seconds. Incubate at 4°C for 5 minutes.

9. Centrifuge the tubes at room temperature for 1 minute at 10,000 x g.

10. Avoiding the pellet, transfer up to, but no more than, 600 µl of supernatant to a clean 2 ml Collection Tube (provided).

11. Add 200 µl of Solution C3 and vortex briefly. Incubate at 4°C for 5 minutes.

12. Centrifuge the tubes at room temperature for 1 minute at 10,000 x g.

13. Avoiding the pellet, transfer up to, but no more than, 750 µl of supernatant into a clean 2 ml Collection Tube (provided).

14. Shake to mix Solution C4 before use. Add 1200 µl of Solution C4 to the supernatant and vortex for 5 seconds.
15. Load approximately 675 μl onto a Spin Filter and centrifuge at 10,000 x g for 1 minute at room temperature. Discard the flow through and add an additional 675 μl of supernatant to the Spin Filter and centrifuge at 10,000 x g for 1 minute at room temperature. Load the remaining supernatant onto the Spin Filter and centrifuge at 10,000 x g for 1 minute at room temperature.

Note
A total of three loads for each sample processed are required.

16. Add 500 μl of Solution C5 and centrifuge at room temperature for 30 seconds at 10,000 x g.
17. Discard the flow through.
18. Centrifuge again at room temperature for 1 minute at 10,000 x g.
19. Carefully place spin filter in a clean 2 ml Collection Tube (provided). Avoid splashing any Solution C5 onto the Spin Filter.
20. Add 100 μl of Solution C6 to the center of the white filter membrane. Alternatively, sterile DNA-Free PCR Grade Water may be used for elution from the silica Spin Filter membrane at this step (MO BIO Catalog# 17000-10).
21. Centrifuge at room temperature for 30 seconds at 10,000 x g.
22. Discard the Spin Filter. The DNA in the tube is now ready for any downstream application. No further steps are required.

We recommend storing DNA frozen (-20° to -80°C). Solution C6 contains no EDTA. To concentrate the DNA see the Hints & Troubleshooting Guide.

8.2.3 Protocole basé sur le kit Maxwell® Buccal Swab LEV DNA (Promega)
Cas des membranes polycarbonate et des coton-tiges : placer la moitié du filtre ou la pointe du coton-tige dans la « clearing column » comme indiqué ci-après. Ajouter 570 μL de tampon de lyse et 2 μL de lysozyme (37.5-KU.μL$^{-1}$ Ready-Lyse lysozyme™, Epicentre Technologies, Madison, USA) et incuber à 37°C pendant 30 min. Ensuite ajouter 30 μL of proteinase K et incuber sur la nuit à 56°C. Suivre les instructions du protocole à partir de l’étape 8 et réaliser une élution dans 100 μL de tampon d’élution.
Protocole Maxwell® Buccal Swab LEV DNA

3.B. Preparation of Buccal Swab Samples (Cat. # AS1295)

Buccal Swab Sample Processing Capacity

The total yield of genomic DNA from buccal swab samples depends on the cellular material contained on the swabs. Each cartridge supplied in the Maxwell® 16 Buccal Swab LEV DNA Purification Kit is designed to purify genomic DNA from 1 or 2 buccal swabs.

Note: Customers have performed this protocol successfully with samples collected using Pur-Wraps Sterile Polyester Tipped Applicators, Puritan Medical Products Cat. # 25-806 1PD or Pur-Wraps Sterile Cotton Tipped Applicators, Puritan Medical Products Cat. # 25-806 1 PC. We also recommend using ClickFit Microtube, 1.5ml (Cat. # V4741) for sample preparation.

1. Collect samples with a standard buccal swab collection procedure.
2. Assemble a Clearing Column/microtube for each sample.
3. Cut the head off the applicator stick. Add dried swab head to the Clearing Column/microtube assembly.
4. In a separate tube, mix 300µl Lysis Buffer + 30µl Proteinase K for each sample.
5. Add 300µl of Lysis Buffer/Proteinase K to swab head in the Clearing Column/microtube assembly.
6. Close tube over the Clearing Column and vortex for 10 seconds.
   Note: If using tubes other than the recommended Cat. # V4741, the tube may not close.
7. Incubate for 20 minutes at 56°C.
   Note: Some flowthrough from the column may be observed in the microtube after incubation. This is normal.
8. Centrifuge the Clearing Column/microtube assembly with swab for 2 minutes at maximum speed.
9. Remove the Clearing Column with swab head and discard.
10. Add flowthrough to well #1 of the Maxwell® 16 LEV cartridge (see Section 3.C for cartridge preparation).
12. Once the extraction is complete, remove and cap each elution tube. Store appropriately until use.

3.C. Maxwell® 16 Cartridge Preparation

1. Change gloves before handling cartridges, LEV Plungers and Elution Tubes. Place the cartridges to be used in the Maxwell® 16 LEV Cartridge Rack (Cat. # AS1251). Place each cartridge in the rack with the label side facing away from the Elution Tubes. Press down on the cartridge to snap it into position. Carefully peel back the seal so that all plastic comes off the top of the cartridge. Ensure that all sealing tape and any residual adhesive are removed before placing cartridges in the instrument.
2. Place one plunger into well #8 of each cartridge.
3. Place an empty Elution Tube into the Elution Tube position for each cartridge in the Maxwell® 16 LEV Cartridge Rack. Add 50μl of Elution Buffer to the bottom of each Elution Tube.

Notes:
1. If you are processing fewer than 16 samples, center the cartridges on the platform.
2. Specimen or reagent spills on any part of the Maxwell® 16 LEV Cartridge Rack should be cleaned with a detergent-water solution, followed by a bacteriocidal spray or wipe, then water. Do not use bleach on any instrument parts.

![Diagram of Maxwell® 16 LEV DNA Purification Cartridge]

Figure 1. Maxwell® 16 LEV DNA Purification Cartridge. This figure shows the contents of a cartridge. In all cases, lysate sample is added to well #1.

![Diagram of Maxwell® 16 LEV Cartridge Rack]

Figure 2. Setup and configuration in the Maxwell® 16 LEV Cartridge Rack. Elution Buffer is added to the Elution Tubes as indicated.

Setup for AS2000 Maxwell® 16 Instruments

Refer to the Maxwell® 16 Instrument Operating Manual #TM295 for more detailed information.

To run the “Blood” protocol, you must have Maxwell® 16 firmware version 4.71 or higher installed on your instrument.

1. Turn on the Maxwell® 16 Instrument. The instrument will power up, display the firmware version number, proceed through a self-check and home all moving parts.

2. Verify that the instrument settings indicate an “LEV” hardware configuration and “Rack” operational mode setting.

3. Select “Run” on the Menu screen, and press the Run/Stop button to start the method.

4. Select “DNA” on the menu screen, then select “OK” at the Verification screen.

5. Select “Blood” on the Menu screen, then select “OK” at the Verification screen. “Blood” Method is used for both kits (Cat. # AS1290 and AS1295).

6. Open the door when prompted to do so on the screen. Press the Run/Stop button to extend the platform.

**Warning:** Pinch point hazard.

7. Transfer the Maxwell® 16 LEV Cartridge Rack containing the prepared cartridges on the Maxwell® 16 Instrument platform. Ensure that the rack is placed in the Maxwell® 16 Instrument with the Elution Tubes closest to the door. The rack will only fit in the instrument in this orientation. If you have difficulty fitting the rack on the platform, check that the rack is in the correct orientation. Ensure that the cartridge rack is level on the instrument platform.

**Note:** Hold the Maxwell® 16 LEV Cartridge Rack by the sides to avoid dislodging cartridges from the rack.

8. Verify that samples were added to well #1 of the cartridges, cartridges are loaded on the instrument. Elution Tubes are present with 50µl of Elution Buffer and LEV Plungers are in well #8.

9. Press the Run/Stop button. The platform will retract. Close the door.

**Warning:** Pinch point hazard.

10. The Maxwell® 16 Instrument will immediately begin the purification run. The screen will display the steps performed and the approximate time remaining in the run.

**Notes:**

1. Pressing the Run/Stop button or opening the door will pause the run.

2. If the run is abandoned before completion, the instrument will wash the particles off the plungers and eject the plungers into well #8 of the cartridge. The sample will be lost.

11. When the automated purification run is complete, the LCD screen will display a message that the method has ended.

---

Promega Corporation - 2100 Woods Hollow Road, Madison, WI 53711-5380 USA - Toll Free in USA 800-356-0526 - 608-274-4380 - Fax 608-277-2516
TM295 - Revised 7/11
www.promega.com
End of Run

12. Follow on-screen instructions at the end of the method to open door. Verify that plungers are located in well #8 of the cartridge at the end of the run. If plungers are not removed from the magnetic plunger bar, push them down gently by hand to remove them.

13. Press the Run/Stop button to extend the platform out of the instrument.

14. Remove the Maxwell® 16 LEV Cartridge Rack from the instrument. Remove Elution Tubes containing DNA, and close the tubes.

  Note: Following the automated purification procedure, the LEV Cartridge Rack will be warm. It will not be too hot to touch. To remove the rack from the instrument platform, hold onto the sides of the rack.

15. Remove the cartridges and plungers from the Maxwell® 16 LEV Cartridge Rack, and discard as hazardous waste. Do not reuse reagent cartridges, LEV Plungers or Elution Tubes.

Setup for AS3000 Maxwell® 16 MDx Instruments

Refer to the Maxwell® 16 MDx Instrument Technical Manual #TM320 for detailed information. To run the “Blood” protocol, you must have the Maxwell® 16 Firmware version 1.10 or higher installed on your instrument

1. Turn on the Maxwell® 16 MDx Instrument. The instrument will power up, display the firmware version number, and proceed through a self-check and home all moving parts.

2. Verify that the Home screen indicates “LEV” and that the LEV hardware is present. Press “Run” to continue.

3. Enter user and PIN, if this option is enabled.

4. At the Protocols screen, select “DNA”.

5. At the Method screen, select “Blood”.

6. On the next screen, verify that the correct user was chosen. The protocol should read “DNA”. Select “Run/Stop” to continue.

7. Open the door when prompted on the screen, then select “Run/Stop”.

  Warning: Pinch point hazard.

8. Follow on-screen instructions for bar code reader input if this option is enabled.

9. Transfer the Maxwell® 16 LEV Cartridge Rack containing the prepared cartridges on the Maxwell® 16 Instrument platform. Ensure that the rack is placed in the Maxwell® 16 Instrument with the Elution Tubes closest to the door. The rack will only fit in the instrument in this orientation. If you have difficulty fitting the rack on the platform, check that the rack is in the correct orientation. Ensure the rack is level on the instrument platform.

  Note: Hold the Maxwell® 16 LEV Cartridge Rack by the sides to avoid dislodging cartridges from the rack.
Setup for AS3000 Maxwell® 16 MDx Instruments (continued)

10. Verify that samples were added to well #1 of the cartridges, cartridges are loaded on the instrument, Elution Tubes are present with 50μl of Elution Buffer and LEV Plungers are in well #8.

11. Press the Run/Stop button. The platform will retract. Close the door.

Warning: Pinch point hazard.

The Maxwell® 16 Instrument will immediately begin the purification run. The screen will display the approximate time remaining in the run.

Notes:
1. Pressing the Run/Stop button or opening the door will pause the run.
2. If the run is abandoned before completion, the instrument will wash the particles off the plungers and eject the plungers into well #8 of the cartridge. The samples will be lost.

12. When the automated purification run is complete, follow instructions on the screen for data transfer. For detailed instructions, refer to the Maxwell® 16 MDx Instrument Technical Manual #TM320 and Maxwell® Sample Track Software Technical Manual #TM314.

End of Run

13. Follow on-screen instructions at the end of the method to open door. Verify that plungers are located in well #8 of the cartridge at the end of the run. If plungers are not removed from the magnetic plunger bar, push them down gently by hand to remove them.

14. Press the Run/Stop button to extend the platform out of the instrument.

15. Remove the Maxwell® 16 LEV Cartridge Rack from the instrument. Remove Elution Tubes containing DNA, and cap the tubes.

Note: Following the automated purification procedure, the LEV Cartridge Rack will be warm. It will not be too hot to touch. To remove the rack from the instrument platform, hold onto the sides of the rack.

16. Remove the cartridges and plungers from the Maxwell® 16 LEV Cartridge Rack, and discard as hazardous waste. Do not reuse spent cartridges, LEV Plungers or Elution Tubes.

For the Maxwell® 16 MDx Instrument, ensure samples are removed before the UV light treatment to avoid damage to the nucleic acid.

5. Reference

Contrasted performance of taxonomic and phylogenetic diversity indices to detect assembly rules and dissimilarities between microbial communities

Chiarello M.*, Bouvier T. and Villéger S.
Supplementary Information S1: Illustration of the method used to assess the performance of alpha-diversity indices to detect phylogenetic clustering. Clustering detection was performed by comparing $D_{clustering}$: value of the tested index of a given clustered community, to the distribution of all 500 $D_{random}$: values of the tested index in randomly assembled communities.
**Supplementary Information S2:** Properties of the communities simulated, with 4 richness values (S, panels titles) and 4 levels of unevenness (σ, abscisses). The level of unevenness was controlled by the log(sd) (σ) parameter of the log-normal distribution (see Material & Methods) and resulted in a strong difference in terms of community structure. Communities computed with a σ=2 were characterized by high dominances, while communities computed with a σ of 0.25 had a very even distribution of abundances. *Left panels:* Relative abundance of the most abundant OTU of the community, expressed in percentage; *right panels:* proportion of OTUs needed to make more than 95% of cumulated relative abundance.
Supplementary Information S3: Illustration of the method used to assess the performance of beta-diversity indices to detect the difference between community types. For each ($S^*\sigma$) scenario, 5 community types were defined, composed of 6 communities each, with a certain level of intra-environmental variability. For clarity, only two communities were shown for each site. Dots represent OTUs defined on tips of the phylogenetic tree; the size of each dot is proportional to the relative abundance of the OTU. We assessed the performance of detection of significant differences between communities belonging to different types by calculating beta-diversity values between all pair of communities in the two compared types, and performing a one-way PERMANOVA.

Supplementary Information S4: R function to calculate Chao’s indices. Available at https://github.com/marlenec/chao.
Supplementary Information S5: Phylogenetic alpha-diversity values for Chao’s (left panels) and Leinster & Cobbold’s indices (right panels) in randomly assembled communities, for each richness ($S$, panels titles) and unevenness values ($\sigma$, in abscisses).
Supplementary Information S6: Efficiency of clustering detection by Leinster & Cobbold’s indices. Values are expressed as percentage of significant P-values (<0.05) across the 500 replicates, for each level of OTUs richness (S, titles of panels) and each level of unevenness of their abundances (σ, abscises). For 1,000 and 2,000-OTUs communities, PD_{TC}(q=1) and PD_{TC}(q=2) are superposed and thus only PD_{TC}(q=2) is visible.
Supplementary Information S7: Efficiency of taxonomic diversity indices to detect difference between communities belonging to different community types, measured as the percentage of significant P-values (PERMANOVA, \( P<0.05 \)) across the 50 replicates, for communities composed of 40 and 2,000 OTUs (\( S \), panels titles) and each level of unevenness (\( \sigma \), abscissas).
Supplementary Information S8: Efficiency of phylogenetic diversity indices to detect the difference between communities belonging to the different community types, measured as the percentage of significant P-values (PERMANOVA, $P<0.05$) across the 50 replicates, for communities composed of 40 and 2,000 OTUs ($S$, panels titles) and each level of unevenness ($\sigma$, abscissae).
SUPPLEMENTARY INFORMATION

Captive bottlenose dolphins and killer whales harbor a species-specific skin microbiota that varies among individuals

Chiarello M., Villéger S., Bouvier C., Auguet JC., and Bouvier T.
**Supplementary Information S1: Description of the two PCR protocols used in this study and comparison of bacterial composition on water samples**

<table>
<thead>
<tr>
<th>Kit</th>
<th>Skin samples</th>
<th>Water samples</th>
</tr>
</thead>
<tbody>
<tr>
<td>Phusion High-Fidelity</td>
<td>PuRe Taq Ready-To-Go PCR Beads</td>
<td></td>
</tr>
<tr>
<td>Total vol. (µL)</td>
<td>20</td>
<td>25</td>
</tr>
<tr>
<td>DNA vol. (µL)</td>
<td>2</td>
<td>5</td>
</tr>
<tr>
<td>Initial denaturation</td>
<td>1 min 98°C</td>
<td>2 min 94°C</td>
</tr>
<tr>
<td>PCR cycle</td>
<td>1 min 94°C; 40s 57.8°C; 30s 72°C</td>
<td>1 min 94°C; 40s 57.8°C; 30s 72°C</td>
</tr>
<tr>
<td>Nb. of cycles</td>
<td>35</td>
<td>35</td>
</tr>
<tr>
<td>Final extension</td>
<td>10 min 72°C</td>
<td>10 min 72°C</td>
</tr>
</tbody>
</table>

**S1-Table 1:** PCR reagents and conditions used for the two sample types studied. Skin DNA and water DNA were respectively amplified using the Phusion High-Fidelity DNA polymerase (Biolabs, Ipswich, USA) and PuRe Taq Ready-To-Go PCR Beads (Amersham Biosciences, Freiburg, Germany) following manufacturer’s instructions.
S1-Fig 1: Most abundant classes and families in planktonic communities analyzed using Phusion and Ready-To-Go kits. Both PCR types were performed on the same DNA extracted from animals’ surrounding water. Class-level bacterial composition was very similar between both PCR types.
S1-Fig 2: PCoAs based on Weighted Unifrac, showing planktonic communities analyzed using both PCR types. On (A) panel, all samples included in this study plus water replicates that could be amplified using Phusion kit. On (B) panel, only planktonic communities were displayed. Water replicates from both animals’ pools clustered together, whatever the PCR type used for amplification.
Supplementary Information S2: Assessing the effect of sequences subsampling

S2-Figure 1: Rarefaction curves obtained on raw sequencing outputs (i.e. with uneven number of sequences read for each sample) (A) and after random subsampling of 10,000 reads for each sample, and (B) number of OTUs at each rarefaction size. Subsampling did not affect the ranking of OTUs richness between planktonic and skin-associated communities (See results in main text).
Supplementary Information S3: phylogenetic analysis of *Staphylococcus sp.*

S3-Fig 1: Phylogenetic relationships of sequences affiliated to *Staphylococcus sp.*

Sequences from this study are labelled as « OTUXX », while other sequences were downloaded from SILVA database. Macrococcus sp and Salinicoccus sp were used as outgroup. Accession numbers from downloaded sequences are available in the table below.
**S3-Table 1:** Accession numbers of sequences downloaded from SILVA database used to make phylogenetic analysis of Staphylococcus-affiliated sequences.

<table>
<thead>
<tr>
<th>Species</th>
<th>Accession</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>Macrococcus bovicus</em></td>
<td>Y15714.1.1546</td>
</tr>
<tr>
<td><em>Macrococcus caseolyticus</em></td>
<td>KP191046.1.1466</td>
</tr>
<tr>
<td><em>Salinicoccus albus</em></td>
<td>EF177692.1.1478</td>
</tr>
<tr>
<td><em>Salinicoccus alkalphilus</em></td>
<td>AF275710.1.1459</td>
</tr>
<tr>
<td><em>Staphylococcus agnetis</em></td>
<td>HM484980.1.1409</td>
</tr>
<tr>
<td><em>Staphylococcus arlettae</em></td>
<td>AB009933.1.1494</td>
</tr>
<tr>
<td><em>Staphylococcus aureus</em></td>
<td>L36472.3194.4748</td>
</tr>
<tr>
<td><em>Staphylococcus aureus</em> subsp. <em>anaerobius</em></td>
<td>D83355.1.1476</td>
</tr>
<tr>
<td><em>Staphylococcus auricularis</em></td>
<td>D83358.1.1475</td>
</tr>
<tr>
<td><em>Staphylococcus capitis</em> subsp. <em>capitis</em></td>
<td>L37599.1.1469</td>
</tr>
<tr>
<td><em>Staphylococcus caprae</em></td>
<td>AB009935.1.1492</td>
</tr>
<tr>
<td><em>Staphylococcus carnosus</em> subsp. <em>carnosus</em></td>
<td>B009934.1.1493</td>
</tr>
<tr>
<td><em>Staphylococcus chromogenes</em></td>
<td>D83360.1.1475</td>
</tr>
<tr>
<td><em>Staphylococcus cohnii</em> subsp. <em>cohnii</em></td>
<td>D83361.1.1477</td>
</tr>
<tr>
<td><em>Staphylococcus condimenti</em></td>
<td>Y15750.1.1545</td>
</tr>
<tr>
<td><em>Staphylococcus delphini</em></td>
<td>AB009938.1.1493</td>
</tr>
<tr>
<td><em>Staphylococcus devriesei</em></td>
<td>FJ389206.1.1537</td>
</tr>
<tr>
<td><em>Staphylococcus epidermidis</em></td>
<td>D83363.1.1475</td>
</tr>
<tr>
<td><em>Staphylococcus equorum</em></td>
<td>AB009939.1.1494</td>
</tr>
<tr>
<td><em>Staphylococcus felis</em></td>
<td>D83364.1.1475</td>
</tr>
<tr>
<td><em>Staphylococcus gallinarum</em></td>
<td>D83366.1.1477</td>
</tr>
<tr>
<td><em>Staphylococcus haemolyticus</em></td>
<td>X66100.1.1544</td>
</tr>
<tr>
<td><em>Staphylococcus hominis</em> subsp. <em>hominis</em></td>
<td>X66101.1.1544</td>
</tr>
<tr>
<td><em>Staphylococcus hyicus</em></td>
<td>D83368.1.1476</td>
</tr>
<tr>
<td><em>Staphylococcus intermedia</em></td>
<td>D83369.1.1476</td>
</tr>
<tr>
<td><em>Staphylococcus lentus</em></td>
<td>D83370.1.1480</td>
</tr>
<tr>
<td><em>Staphylococcus lugdunensis</em></td>
<td>AB009941.1.1492</td>
</tr>
<tr>
<td><em>Staphylococcus massiliensis</em></td>
<td>EU707796.1.1477</td>
</tr>
<tr>
<td><em>Staphylococcus muscae</em></td>
<td>FR733703.1.1537</td>
</tr>
<tr>
<td><em>Staphylococcus pasteuri</em></td>
<td>AB009944.1.1494</td>
</tr>
<tr>
<td><em>Staphylococcus petrasii</em> subsp. <em>jettensis</em></td>
<td>JN092118.1.1444</td>
</tr>
<tr>
<td><em>Staphylococcus piscifermentans</em></td>
<td>Y15754.1.1544</td>
</tr>
<tr>
<td><em>Staphylococcus pseudointermedius</em></td>
<td>AJ780976.1.1512</td>
</tr>
<tr>
<td><em>Staphylococcus saccharolyticus</em></td>
<td>L37602.1.1527</td>
</tr>
<tr>
<td><em>Staphylococcus saprophyticus</em> subsp. <em>saprophyticus</em></td>
<td>AP008934.743716.745270</td>
</tr>
<tr>
<td><em>Staphylococcus simiae</em></td>
<td>AY727530.1.1478</td>
</tr>
<tr>
<td><em>Staphylococcus simulans</em></td>
<td>D83373.1.1476</td>
</tr>
<tr>
<td><em>Staphylococcus succinii</em> subsp. <em>succinii</em></td>
<td>AF004220.1.1548</td>
</tr>
<tr>
<td><em>Staphylococcus warneri</em></td>
<td>L37603.1.1470</td>
</tr>
<tr>
<td><em>Staphylococcus xylosus</em></td>
<td>D83374.1.1477</td>
</tr>
</tbody>
</table>
Supplementary Information S4: Alpha-diversity recovered from samples.

S4-Fig 1: Diversity of skin-associated communities of killer whales (A and D), dolphins (B and E), and of planktonic communities (C and F). The first row of plots (A-C) illustrates the phylogenetic richness observed in a sample, measured with relative Faith PD, i.e. the sum of the height of all branches from the phylogenetic tree recovered in the sample corrected by the total height of the phylogenetic tree, and converted into percentage. The second row of plots (D-F) illustrates taxonomic diversity, measured by the index of Shannon, which is based on the relative abundances of OTUs recovered in a sample. Total diversity of each individual (i.e. accounting for all body zones sampled) is illustrated with larger light-gray bars. Bars on panels C and F represent the mean (and associated standard deviation) of OTU richness and phylogenetic diversity for planktonic communities (n=3 water samples). “Pool” refers to animal’s surrounding water, and “Input” refers to the water sampled at the exit of filtering system.
Supplementary Information S5: Dissimilarity between microbiotas

S5-Fig 1: Average of pairwise dissimilarity between sets of microbiotas for two facets and two components of biodiversity. A: Taxonomic (based on OTUs) compositional (insensitive to relative abundances of OTUs) dissimilarity. B: taxonomic structural (weighted by OTUs relative abundances) dissimilarity. C-D: phylogenetic compositional and structural dissimilarities. Error bars represent the standard deviation of dissimilarity indices among pairs of microbiotas.
Supplementary Information S6: Structure of microbiotas

S6-Fig 1: Mean relative abundance of bacterial families in skin-associated communities of common bottlenose dolphin and killer whales, and planktonic communities. P: upper side of pectoral fin, D: dorsal fin, C: upper side of caudal fin, A: anal zone. “Pool” refers to animal’s surrounding water, and “Input” refers to the water sampled at the exit of pipe from filtering system.

S6-Fig 2: Mean relative abundance of bacterial genera in skin-associated communities of common bottlenose dolphin and killer whales, and planktonic communities. P: upper side of pectoral fin, D: dorsal fin, C: upper side of caudal fin, A: anal zone. “Pool” refers to animal’s surrounding water, and “Input” refers to the water sampled at the exit of pipe from filtering system.


a: Moraxellaceae; b: Staphylococcaceae; c: Sphingomonadaceae; d: Rhodobacteraceae; e: Bifidobacteriaceae; f: Rhizobiaceae; g: Pasteurellaceae; h: Micrococcaceae; i: Fusobacteriaceae; j: Pseudomonadaceae; k: Streptococcaceae; l: Corynebacteriaceae; m: Flavobacteriaceae; n: Actinomycetaceae; o: Pseudoalteromonadaceae; p: Halomonadaceae; q: Alteromonadaceae.
Supplementary Information S7: LefSe analysis performed on KEGG pathways

S7-Fig 1: LefSe analysis performed on KEGG pathways, showing the most discriminating pathways (higher LDA score) between skin-associated and planktonic communities. Functional categories associated with highest LDA score for each type of microbial communities are underlined, and pathways belonging to these categories are illustrated with dots of corresponding colors.
Supplementary Methods

Skin microbiome of coral reef fishes is diversified, species-specific, not phylogenetically conserved and vulnerable to global change

Marlène Chiarello, Jean-Christophe Auguet, Yvan Bettarel, Corinne Bouvier, Thomas Claverie, Nicholas AJ Graham, Fabien Rieuvilleneuve, Elliot Sucré, Thierry Bouvier, and Sébastien Villéger

Sampling area
Fish sampling was conducted on November 2015 (17th to 27th) on coral reefs around Mayotte Island (France), located in the Western part of the Indian Ocean. Mayotte lagoon is the third largest lagoon in the world and houses 195 km of coral reefs and more than 700 fish species (Wickel et al. 2014). Fish were sampled from two sites in the South West of the lagoon: a fringing reef (S12°54‘17.46”, E44°58‘15.72”), and the inner slope of the barrier reef (S12°57‘33.72”, E45°04‘49.38”). Both sites are far from cities, were at a good ecological state at the time of sampling with more than >50% of coral cover and abundant fish communities including predators such as groupers and barracudas.

Sampling procedure
Sampling objective was to target the most abundant species of major ecologically and phylogenetically contrasted fish families present in each site (within a radius of 50m), including Acanthuridae, Balistidae, Chaetodontidae, Labridae, Pomacanthidae, Pomacentridae, Scaridae and Scorpaenidae. In order to take into account the interspecific variability of skin microbiome, up to 3 adult individuals of each species were sampled in each site.
In order to avoid contamination during sampling, fishes were caught using speargun and hook line and killed immediately after capture by cervical dislocation (following the European directive 2010/63/UE). Fish were handled only by the mouth using a clamp and all participants wear gloves. After death, the fish was laid down on one side along a ruler, measured, and skin microbiome was sampled by swabbing the entire untouched side of the body (from back of operculum to caudal peduncle, i.e. head not included) using buccal swabs (SK-2S swabs, Isohelix, UK).

A total of 138 fishes were sampled for their skin microbiome. They belonged to 44 species with 29 species represented by at least three individuals (Supplementary Information S1) and 10 species represented by a single individual. Species belonged to 5 orders and 22 families, with 35 species belonging to Perciformes (S1).

Fish Sampling was authorized by the Mayotte’s directorate of maritime affairs (permit N°12/UTM/2015).

To assess microbial diversity in the water column of the two sites, a total of thirty six 200-mL seawater samples were collected at sea surface (9 samples) and at 30 cm from the bottom of the sea (9 samples), stored in an electric cooler, and filtrated at the end of the day through a 47 mm 0.2 µm polycarbonate membrane (Whatman, Clifton, USA). The membranes were then placed in sterile cryotubes. One surface water sample taken on fringing reef could not be amplified during subsequent steps, and was removed, making a total of 35 water samples included in this study.

All samples were stored in an electric cooler during the day, which was cooled to -20°C before each sampling day, and cooled again during the day if the temperature inside rose over -5°C. At the end of the day, all samples were transported in the lab and stored at -80°C during two weeks, and then transported back by plane in mainland France in the electric cooler for DNA extraction.

16S rDNA amplification and sequencing
Swabs and water membranes were incubated during 30 minutes at 37°C in 570 µL of lysis buffer from Maxwell® Buccal Swab LEV DNA kits (Promega Corporation, Madison, USA) and 2 µL of 37.5-KU µL⁻¹ Ready-Lyse lysozyme™ (Epicentre Technologies, Madison, USA).
Then, 30 µl of proteinase K (from manufacturer’s kit) were added and tubes were incubated overnight at 56°C. The totality of the solution was then placed in the kit for extraction. DNA extraction was performed using the Maxwell® 16 Bench-top extraction system following manufacturer’s instructions, and eluted in 50 µL of elution buffer. The V3-V4 region of the 16S rDNA gene was amplified using the prokaryotic primers modified for Illumina sequencing 515F (5’-C TTT CCC TAC ACG ACG CTC TTC CGA TCT - GTG CCA GCM GCC GCG GTA A- 3’) (Caporaso et al. 2011) and the modified version of 806R by Apprill et al. (Apprill et al. 2015) (5’ – G GAG TTC AGA CGT GTG CTC TTC CGA TCT - GGA CTA CNV GGG TWT CTA AT - 3’), with PuRe Taq Ready-To-Go PCR Beads (Amersham Biosciences, Freiburg, Germany) using 1µL of extracted DNA and 10µM.µL⁻¹ of each primer as follows: initial denaturation at 94°C for 1 min, followed by 35 cycles of 94°C for 1 min, 55°C for 1 min and 72°C for 1 min, ending with a final extension at 72°C for 10 min. Equimolar amounts of sample DNA extracted from each sample site were separately pooled and sequenced in two separated runs by an external laboratory (INRA GeT-PlaGE platform, Toulouse, France) on an Illumina platform using the 2x 250 bp MiSeq chemistry. 7 PCR blanks were included in each sequencing run in order to assess the presence of contaminants, which were removed during subsequent steps of sequence processing.

**Sequence processing to define OTUs and their phylogenetic relationships**

Sequence processing was performed following the SOP of Kozich et al for MiSeq (2013), [https://www.mothur.org/wiki/MiSeq_SOP](https://www.mothur.org/wiki/MiSeq_SOP), 2017) using Mothur (Schloss et al. 2009). After assembly of paired reads in each run, sequences of both runs were merged and sequences with an abnormal length (outside a range of 250-300 pb) were removed. Sequences were aligned along the SILVA reference database (Quast et al. 2013) release 128, with removing of sequences abnormally aligned. Chimeras were removed using UCHIME (Edgar et al. 2011). Filtered sequences were then classified using the SILVA reference taxonomy and the non-prokaryotic ones were removed. 10,877 sequences from 173 samples were kept after cleaning process, ranging from 2,450 to 43,306 sequences per sample. After this, 2,000 sequences were sub-sampled within each sample in order to correct the uneven sequencing efficiency among samples. Sequences were then grouped into Operational Taxonomic Units (OTUs) using a 97% cutoff parameter, and the relative abundance of all OTUs was computed using number of sequences. Relative abundances of OTUs recovered from blank samples were then subtracted.
to their respective relative abundance in all other samples. Rarefaction curves obtained from all samples are provided in S13.

The dominant sequence for each OTU was selected as reference and added into the SILVA reference phylogenetic tree (release 128) using ARB software (Ludwig et al. 2004). The full phylogenetic tree was then pruned using the ape R-package to remove all but the added sequences, while keeping the topology of the tree. A chronogram was then adjusted to the phylogenetic tree using PATHd8 (Britton et al. 2007). The divergence time between Archaea and Bacteria was fixed at 3.8 Ga. The minimum divergence time between Euryarchaeota and other Archaea was set to 2.7 Ga (Blank 2009), and the maximum age of apparition of Thermoplasmatales was set to 2.32 Ga (Blank 2009). The minimum age of apparition of Cyanobacteria was set to 2.5 Ga (Schirrmeister et al. 2013). The minimum divergence time between Rickettsiales and the rest of Alphaproteobacterial sequences was set at 1.6 Ga, as done by Groussin et al. (Groussin et al. 2017). Finally the divergence times between Chromatiaceae and other Gammaproteobacteria, was set to minimum 1.64 Ga (Brocks et al. 2005).

60% of the recovered OTUs were not classified by Mothur at class level using SILVA reference alignment. We refined the taxonomic affiliation of the most frequent unclassified OTUs (i.e. 93 OTUs that were recovered in at least 15 samples, see Supplementary Methods) using the Arb parsimony insertion tool and the Silva backbone tree. Results are provided in S3.

**Computing phylogenetic diversity**

For each microbial assemblage (i.e. skin microbiome of one fish individual or planktonic microbes from one water sample) we assessed two complementary phylogenetic diversity indices. First we measured phylogenetic richness which accounts only for presence of microbial lineages using Faith’PD index (Faith 1992) computed with *picante* R-package (Kembel et al. 2010), and then corrected by the sum of all branches lengths in the prokaryotic tree in order to scale values between 0 and 1. Faith’s PD index increases when OTUs are phylogenetically distant. Second, we measured phylogenetic entropy accounting for the relative abundance of OTUs, using the Allen’s index (Allen, Kon, and Bar-Yam 2009), which is a phylogenetic extension of Shannon’s taxonomic entropy index. Allen index was
computed using our own R-function ([https://github.com/marlenec/chao](https://github.com/marlenec/chao), q=1) based on entropart R-package (Marcon and Hérault 2014). Allen’s index increases when the most abundant OTUs are phylogenetically distant.

Phylogenetic dissimilarity between pairs of microbial assemblages was assessed using two indices accounting only for OTUs presence or their dominance, *i.e.* unweighthed and abundance-weighted versions of the Unifrac index, respectively, both computed using the GUniFrac R-package (Chen and ORPHANED 2012). Phylogenetic dissimilarity indices accounting for composition (or structure) ranges from 0 when assemblages share the same (dominant) phylogenetic lineages to 1 when assemblages are made of (or dominated by) phylogenetically distant OTUs.

**Fish phylogeny and ecological traits**

Phylogenetic relationships between studied fish species were extracted from a published time-calibrated phylogeny containing 7,822 fish species, covering all Actinopterygian orders (Rabosky et al. 2013). Out of our 44 fish species, 13 were not initially present in the phylogenetic tree, and were manually grafted next to their closest species accordingly to literature. One species (*Cephalopholis argus*) was incorrectly branched next to Scaridae in the initial tree, and was therefore also grafted next to its closest relative (see S14).

The ecology of the 44 species was described using a set of 6 categorical traits describing body size at fish maturity, mobility, period of activity, schooling behaviour, position in water column and diet. Values were taken from a global database of functional traits for 6,316 tropical reef fishes (Mouillot et al. 2014). The distribution of trait values among the 44 studied species is described in S9.

**Fish vulnerability**

Vulnerability of each fish species to fishing (Cheung, Pitcher, and Pauly 2005) was obtained from Fishbase ([http://www.fishbase.org/](http://www.fishbase.org/), 2017). Cheung’s vulnerability to fishing is based on fish’s total length, life history traits (including age at first maturity, longevity, growth rate, fecundity, spatial behaviour), and geographical range. A vulnerability close to 100 means that the vulnerability to extinction is very high, while vulnerability close to 0 means that the vulnerability to extinction is low. Vulnerability to habitat loss due to global change in coral reefs was computed as in (Graham et al. 2011) based on species diet specialization, habitat specialization, recruitment specialization for live coral and body size estimated from data from Fishbase.org and expert knowledge.
Determinants of microbial alpha-diversity

The comparison of both diversity facets (phylogenetic richness and phylogenetic entropy) obtained in planktonic communities and fish skin microbiome, as well as between fish species was done using separated Kruskal-Wallis tests (999 permutations) in vegan R-package. To test if closely related fish species have more similar levels of phylogenetic richness and entropy values than expected by chance, we computed Moran’s I which is used as an autocorrelation measure of trait variation along a phylogenetic tree. We thus used the inverse of patristic distances between fish species as a measure of phylogenetic proximity (Münkemüller et al. 2012) and replicate the test on 999 random subsamples of 44 individuals (1 per species) to account for intraspecific variability. For each test, Moran’s I values were than compared to the ones obtained when shuffling diversity values 999 times on the phylogenetic tree using adephylo R-package.

To test whether microbial phylogenetic richness and entropy on fish skin was correlated with the vulnerability of fish species to anthropogenic threats, we performed a Spearman’s correlation test between each phylogenetic diversity index (richness and entropy) and each fish vulnerability score (to overfishing and to habitat loss). These analyzes were performed on 999 subsamples of individuals to account for intraspecific variability in skin microbiome.

Determinants of dissimilarity between skin microbiomes

The comparison of the structure of fish-associated microbial communities and the planktonic ones was performed on the full dataset using a permutational ANOVA (PERMANOVA) performed on dissimilarity values (Weighted and Unweighted Unifrac) using vegan R-package (Table 1). To assess how each microbial clade contributed to the dissimilarity between planktonic and skin-associated microbial communities, we performed a LefSe analysis (Segata et al. 2011). LefSe provides Linear Discriminant Analysis (LDA) scores for the bacterial clades contributing the most to the differences between communities (S2).

The assessment of the effect of fish species was done using a PERMANOVA (Table 1). To assess the effect of reef type on fish skin microbiome, we performed a PERMANOVA on the 16 species for which we sampled at least one representative on both reef types (total of 74 individuals). To compare the effects of reef type and fish species on its microbiome, both factors, as well as the interaction between them, were included in the analysis (S8). The correlation between dissimilarities between microbiomes and hosts’ phylogeny (phylosymbiosis) was measured using Mantel tests based on Pearson’s coefficient, using
vegan R-package and 999 permutations. In order to take into account the effect of intraspecific variability 999 random subsamples containing one individual per fish species were considered. To check if the phylogenetic signal was blurred due to very long phylogenetic branches between fish species, separated Mantel tests were also computed on the fish order Perciformes only.

In order to assess the effects of host phylogeny at higher phylogenetic scales than OTUs, we used the Beta Diversity Through Time (BTTD) approach developed by Groussin et al. (Groussin et al. 2017), which computes various beta-diversity indices at different time periods (slices) along the bacterial phylogenetic tree. We went back in time this way until 900 Mya, which corresponds approximately to divergence between bacterial orders, and computed Sorensen and Bray-Curtis indices at each slice of 100 Mya. At each slice, correlation between pairwise beta-diversity values and host phylogeny was tested using a Mantel test based on Pearson’s coefficient and 999 permutations. This analysis was performed using 500 subsamples of 44 individuals (1 per species).

The effect of fish ecological traits was assessed using PERMANOVAs, using 999 random subsamples containing one individual by fish species. The ecological traits were ordered in the model according to their independent contribution (greatest to least) to the total variability.

As for all analyses involving dataset subsampling, results were reported as the percentage of significant P-values obtained in all subsamples, together with the number of subsamples used in the analysis, and when useful, the mean standard deviation of the statistic among all subsamples.

REFERENCES


Wickel, Julien, Alain Jamon, Mathieu Pinault, Patrick Durville, and Chabanet Pascale. 2014. “Composition et Structure Des Peuplements Ichtyologiques Marins de L’île de Mayotte (Sud-Ouest de L’océan Indien).” *Société Française d’Ichtyologie.*
8.5.2 Informations supplémentaires au manuscrit D

Supplementary Information

Skin microbiome of coral reef fishes is diversified, species-specific, not phylogenetically conserved and vulnerable to global change

Marlène Chiarello, Jean-Christophe Auguet, Yvan Bettarel, Corinne Bouvier, Thomas Claverie, Nicholas AJ Graham, Fabien Rieuville, Elliot Sucré, Thierry Bouvier, and Sébastien Villéger
**Supplementary Information S1:**

<table>
<thead>
<tr>
<th>Order</th>
<th>Family</th>
<th>Species</th>
<th>Number of individuals</th>
</tr>
</thead>
<tbody>
<tr>
<td>Beryciformes</td>
<td>Holocentridae</td>
<td>Myripristis murdjan</td>
<td>2</td>
</tr>
<tr>
<td>Beryciformes</td>
<td>Holocentridae</td>
<td>Myripristis violacea</td>
<td>1</td>
</tr>
<tr>
<td>Perciformes</td>
<td>Acanthuridae</td>
<td>Acanthurus leucosternon</td>
<td>5</td>
</tr>
<tr>
<td>Perciformes</td>
<td>Acanthuridae</td>
<td>Acanthurus lineatus</td>
<td>3</td>
</tr>
<tr>
<td>Perciformes</td>
<td>Acanthuridae</td>
<td>Ctenochaetus striatus</td>
<td>7</td>
</tr>
<tr>
<td>Perciformes</td>
<td>Acanthuridae</td>
<td>Naso unicornis</td>
<td>1</td>
</tr>
<tr>
<td>Perciformes</td>
<td>Caesionidae</td>
<td>Pterocaesio tile</td>
<td>1</td>
</tr>
<tr>
<td>Perciformes</td>
<td>Caesionidae</td>
<td>Pterocaesio trilinenata</td>
<td>2</td>
</tr>
<tr>
<td>Perciformes</td>
<td>Carangidae</td>
<td>Caranx melampygus</td>
<td>3</td>
</tr>
<tr>
<td>Perciformes</td>
<td>Chaetodontidae</td>
<td>Chaetodon auriga</td>
<td>2</td>
</tr>
<tr>
<td>Perciformes</td>
<td>Chaetodontidae</td>
<td>Chaetodon falcula</td>
<td>6</td>
</tr>
<tr>
<td>Perciformes</td>
<td>Chaetodontidae</td>
<td>Chaetodon lunula</td>
<td>4</td>
</tr>
<tr>
<td>Perciformes</td>
<td>Chaetodontidae</td>
<td>Chaetodon meyeri</td>
<td>3</td>
</tr>
<tr>
<td>Perciformes</td>
<td>Chaetodontidae</td>
<td>Chaetodon trifascialis</td>
<td>3</td>
</tr>
<tr>
<td>Perciformes</td>
<td>Ephippidae</td>
<td>Platax orbicularis</td>
<td>5</td>
</tr>
<tr>
<td>Perciformes</td>
<td>Ephippidae</td>
<td>Platax teira</td>
<td>1</td>
</tr>
<tr>
<td>Perciformes</td>
<td>Kyphosidae</td>
<td>Kyphosus vaigiensis</td>
<td>3</td>
</tr>
<tr>
<td>Perciformes</td>
<td>Labridae</td>
<td>Cheilinus fasciatus</td>
<td>3</td>
</tr>
<tr>
<td>Perciformes</td>
<td>Labridae</td>
<td>Hemigymnus fasciatus</td>
<td>3</td>
</tr>
<tr>
<td>Perciformes</td>
<td>Labridae</td>
<td>Thalassoma hebraicum</td>
<td>3</td>
</tr>
<tr>
<td>Perciformes</td>
<td>Lethrinidae</td>
<td>Monotaxis grandoculis</td>
<td>5</td>
</tr>
<tr>
<td>Perciformes</td>
<td>Mullidae</td>
<td>Parupeneus cyclostomus</td>
<td>2</td>
</tr>
<tr>
<td>Perciformes</td>
<td>Mullidae</td>
<td>Parupeneus trifasciatis</td>
<td>3</td>
</tr>
<tr>
<td>Perciformes</td>
<td>Pinguipedidae</td>
<td>Parapercis hexophtalma</td>
<td>3</td>
</tr>
<tr>
<td>Perciformes</td>
<td>Pomacanthidae</td>
<td>Pomacanthus imperator</td>
<td>3</td>
</tr>
<tr>
<td>Perciformes</td>
<td>Pomacanthidae</td>
<td>Pygopliches diacanthus</td>
<td>6</td>
</tr>
<tr>
<td>Perciformes</td>
<td>Pomacentridae</td>
<td>Abudeduf sexfasciatus</td>
<td>3</td>
</tr>
<tr>
<td>Perciformes</td>
<td>Pomacentridae</td>
<td>Abudeduf sparoides</td>
<td>4</td>
</tr>
<tr>
<td>Perciformes</td>
<td>Pomacentridae</td>
<td>Amphiprion akallopisos</td>
<td>3</td>
</tr>
<tr>
<td>Perciformes</td>
<td>Scaridae</td>
<td>Chlorurus sordidus</td>
<td>6</td>
</tr>
<tr>
<td>Perciformes</td>
<td>Scaridae</td>
<td>Scarus caudofasciatus</td>
<td>4</td>
</tr>
<tr>
<td>Perciformes</td>
<td>Scaridae</td>
<td>Scarus russelii</td>
<td>1</td>
</tr>
<tr>
<td>Perciformes</td>
<td>Serranidae</td>
<td>Cephalopholis argus</td>
<td>6</td>
</tr>
<tr>
<td>Perciformes</td>
<td>Serranidae</td>
<td>Cephalopholis boenak</td>
<td>1</td>
</tr>
<tr>
<td>Perciformes</td>
<td>Sphyraenidae</td>
<td>Sphyraena barracuda</td>
<td>1</td>
</tr>
<tr>
<td>Perciformes</td>
<td>Zanclidae</td>
<td>Zanclus cornutus</td>
<td>6</td>
</tr>
<tr>
<td>Scorpaeniformes</td>
<td>Scorpaenidae</td>
<td>Pterois miles</td>
<td>2</td>
</tr>
<tr>
<td>Scorpaeniformes</td>
<td>Scorpaenidae</td>
<td>Pterois radiata</td>
<td>1</td>
</tr>
<tr>
<td>Syngnathiformes</td>
<td>Syngnathidae</td>
<td>Corythoichthys flavofasciatus</td>
<td>3</td>
</tr>
<tr>
<td>Tetraodontiformes</td>
<td>Balistidae</td>
<td>Balistapus undulatus</td>
<td>3</td>
</tr>
<tr>
<td>Tetraodontiformes</td>
<td>Balistidae</td>
<td>Sufflaman chrysopterus</td>
<td>6</td>
</tr>
<tr>
<td>Tetraodontiformes</td>
<td>Diodontidae</td>
<td>Arothron nigropunctatus</td>
<td>1</td>
</tr>
<tr>
<td>Tetraodontiformes</td>
<td>Monacanthidae</td>
<td>Cantherhines pardalis</td>
<td>1</td>
</tr>
</tbody>
</table>

**Total Individuals** 138

*S1: Fish individuals included in this study.* Classification is from Nelson, 2016.
Supplementary Information S2:

S2-Figure: **LefSe analysis showing the most differential microbial taxa** between fish skin-associated and planktonic communities. Only microbial taxa that raised an LDA score > 4 are shown. See supplementary methods for information and references about LefSe.
S2-Table: Full results of LefSe analysis, showing all significant biomarkers detected in fish surface and in seawater

<table>
<thead>
<tr>
<th>Prokaryotic clades enriched in fish skin</th>
<th>log10 LDA Score</th>
<th>Prokaryotic clades enriched in seawater</th>
<th>log10 LDA Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Proteobacteria</td>
<td>5.12</td>
<td>Cyanobacteria</td>
<td>5.02</td>
</tr>
<tr>
<td>Gammaproteobacteria</td>
<td>5.04</td>
<td>unclassified</td>
<td>4.71</td>
</tr>
<tr>
<td>Alphaproteobacteria</td>
<td>4.42</td>
<td>Rhodobacteraceae</td>
<td>4.34</td>
</tr>
<tr>
<td>Vibrionaceae</td>
<td>4.41</td>
<td>Flavobacteriaceae</td>
<td>4.28</td>
</tr>
<tr>
<td>Vibrionales</td>
<td>4.41</td>
<td>Flavobacteriales</td>
<td>4.24</td>
</tr>
<tr>
<td>Firmicutes</td>
<td>4.31</td>
<td>Flavobacteriia</td>
<td>4.24</td>
</tr>
<tr>
<td>Catenococcus</td>
<td>4.28</td>
<td>Proteobacteria</td>
<td>4.18</td>
</tr>
<tr>
<td>Alteromonadales</td>
<td>4.17</td>
<td>Bacteroidetes</td>
<td>4.14</td>
</tr>
<tr>
<td>Rhizobiales</td>
<td>4.15</td>
<td>Rhodobacteraceae</td>
<td>4.09</td>
</tr>
<tr>
<td>Clostridia</td>
<td>4.09</td>
<td>Rhodobacteriales</td>
<td>4.09</td>
</tr>
<tr>
<td>Clostridiales</td>
<td>4.09</td>
<td>Aestuarilibacter</td>
<td>3.88</td>
</tr>
<tr>
<td>Alteromonadales</td>
<td>4.00</td>
<td>Alteromonadaceae</td>
<td>3.88</td>
</tr>
<tr>
<td>Methylobacteriaceae</td>
<td>3.94</td>
<td>Rhodospirillaceae</td>
<td>3.59</td>
</tr>
<tr>
<td>Methylobacterium</td>
<td>3.94</td>
<td>Rhodospirillales</td>
<td>3.35</td>
</tr>
<tr>
<td>Clostridiales</td>
<td>3.91</td>
<td>BD2_7</td>
<td>3.04</td>
</tr>
<tr>
<td>Halomonadaceae</td>
<td>3.89</td>
<td>Cellvibrionales</td>
<td>3.02</td>
</tr>
<tr>
<td>Oceanospirillales</td>
<td>3.89</td>
<td>OM182_clade</td>
<td>2.66</td>
</tr>
<tr>
<td>Betaproteobacteria</td>
<td>3.84</td>
<td>Betaproteobacteria</td>
<td>2.38</td>
</tr>
<tr>
<td>Burkholderiales</td>
<td>3.82</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Breoghania</td>
<td>3.75</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bradyrhizobiaceae</td>
<td>3.73</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Psychrobium</td>
<td>3.73</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Shewanellaceae</td>
<td>3.73</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fusobacteriaceae</td>
<td>3.70</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fusobacteria</td>
<td>3.70</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fusobacteria</td>
<td>3.70</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fusobacteriales</td>
<td>3.70</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Aliivibrio</td>
<td>3.68</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ilyobacter</td>
<td>3.68</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Firmicutes</td>
<td>3.67</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Planctomycetales</td>
<td>3.59</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Planctomycetaceae</td>
<td>3.59</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Planctomycetes</td>
<td>3.59</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Planctomycetacia</td>
<td>3.59</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bacilli</td>
<td>3.56</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Campylobacterales</td>
<td>3.55</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Epsilonproteobacteria</td>
<td>3.55</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Actinobacteria</td>
<td>3.54</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Campylobacteraceae</td>
<td>3.50</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nesiotobacter</td>
<td>3.50</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Oxalobacteraceae</td>
<td>3.43</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Campylobacteraceae</td>
<td>3.43</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bacillales</td>
<td>3.37</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Burkholderiales</td>
<td>3.33</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kingdom</td>
<td>Order</td>
<td>Weight</td>
<td></td>
</tr>
<tr>
<td>-------------------------</td>
<td>-------------------</td>
<td>--------</td>
<td></td>
</tr>
<tr>
<td>Sphingobacteriia</td>
<td></td>
<td>3.32</td>
<td></td>
</tr>
<tr>
<td>Sphingobacteriales</td>
<td></td>
<td>3.32</td>
<td></td>
</tr>
<tr>
<td>Rhodospirillum</td>
<td></td>
<td>3.31</td>
<td></td>
</tr>
<tr>
<td>Planctomycetales</td>
<td></td>
<td>3.31</td>
<td></td>
</tr>
<tr>
<td>Vibrionaceae</td>
<td></td>
<td>3.31</td>
<td></td>
</tr>
<tr>
<td>Comamnonadaceae</td>
<td></td>
<td>3.31</td>
<td></td>
</tr>
<tr>
<td>Burkholderiales</td>
<td></td>
<td>3.30</td>
<td></td>
</tr>
<tr>
<td>Chitinophagaceae</td>
<td></td>
<td>3.29</td>
<td></td>
</tr>
<tr>
<td>Hydrotalea</td>
<td></td>
<td>3.26</td>
<td></td>
</tr>
<tr>
<td>Azorhizobium</td>
<td></td>
<td>3.26</td>
<td></td>
</tr>
<tr>
<td>Rhodospirillales</td>
<td></td>
<td>3.23</td>
<td></td>
</tr>
<tr>
<td>Microcccales</td>
<td></td>
<td>3.20</td>
<td></td>
</tr>
<tr>
<td>Bacillales</td>
<td></td>
<td>3.17</td>
<td></td>
</tr>
<tr>
<td>Clostridacies_1</td>
<td></td>
<td>3.13</td>
<td></td>
</tr>
<tr>
<td>Roseimaritima</td>
<td></td>
<td>3.13</td>
<td></td>
</tr>
<tr>
<td>FamilyI</td>
<td></td>
<td>3.11</td>
<td></td>
</tr>
<tr>
<td>Acetobacteraceae</td>
<td></td>
<td>3.07</td>
<td></td>
</tr>
<tr>
<td>Capnocytophaga</td>
<td></td>
<td>3.06</td>
<td></td>
</tr>
<tr>
<td>Bacilli</td>
<td></td>
<td>3.06</td>
<td></td>
</tr>
<tr>
<td>Flavihuimbacter</td>
<td></td>
<td>3.05</td>
<td></td>
</tr>
<tr>
<td>Actinobacteria</td>
<td></td>
<td>3.04</td>
<td></td>
</tr>
<tr>
<td>Clostridium_sensu_stricto</td>
<td></td>
<td>3.04</td>
<td></td>
</tr>
<tr>
<td>Clostridacies_2</td>
<td></td>
<td>3.04</td>
<td></td>
</tr>
<tr>
<td>Proteinivorax</td>
<td></td>
<td>3.02</td>
<td></td>
</tr>
<tr>
<td>Family_XIV</td>
<td></td>
<td>3.02</td>
<td></td>
</tr>
<tr>
<td>Bacillaceae</td>
<td></td>
<td>2.98</td>
<td></td>
</tr>
<tr>
<td>Methanomicrobales</td>
<td></td>
<td>2.98</td>
<td></td>
</tr>
<tr>
<td>Lachnospiraceae</td>
<td></td>
<td>2.96</td>
<td></td>
</tr>
<tr>
<td>Corynebacteriales</td>
<td></td>
<td>2.96</td>
<td></td>
</tr>
<tr>
<td>Acetobacteraceae</td>
<td></td>
<td>2.95</td>
<td></td>
</tr>
<tr>
<td>Dermabacteraceae</td>
<td></td>
<td>2.95</td>
<td></td>
</tr>
<tr>
<td>Ornithobacterium</td>
<td></td>
<td>2.94</td>
<td></td>
</tr>
<tr>
<td>Clostridacies_1</td>
<td></td>
<td>2.93</td>
<td></td>
</tr>
<tr>
<td>Hyphomonadaceae</td>
<td></td>
<td>2.93</td>
<td></td>
</tr>
<tr>
<td>Xanthobacteraceae</td>
<td></td>
<td>2.93</td>
<td></td>
</tr>
<tr>
<td>Saccharococcus</td>
<td></td>
<td>2.93</td>
<td></td>
</tr>
<tr>
<td>Halobacteriales</td>
<td></td>
<td>2.92</td>
<td></td>
</tr>
<tr>
<td>Haematobacter</td>
<td></td>
<td>2.91</td>
<td></td>
</tr>
<tr>
<td>Caulobacterales</td>
<td></td>
<td>2.91</td>
<td></td>
</tr>
<tr>
<td>Lachnospiraceae</td>
<td></td>
<td>2.91</td>
<td></td>
</tr>
<tr>
<td>Halobacteriales</td>
<td></td>
<td>2.90</td>
<td></td>
</tr>
<tr>
<td>Ruminococcaceae</td>
<td></td>
<td>2.89</td>
<td></td>
</tr>
<tr>
<td>Halobacteria</td>
<td></td>
<td>2.88</td>
<td></td>
</tr>
<tr>
<td>Desulfotignum</td>
<td></td>
<td>2.88</td>
<td></td>
</tr>
<tr>
<td>Sulphurovum</td>
<td></td>
<td>2.88</td>
<td></td>
</tr>
<tr>
<td>Deltaproteobacteria</td>
<td></td>
<td>2.88</td>
<td></td>
</tr>
<tr>
<td>Hyunsoonleella</td>
<td></td>
<td>2.87</td>
<td></td>
</tr>
<tr>
<td>Microcccaseae</td>
<td></td>
<td>2.86</td>
<td></td>
</tr>
<tr>
<td>Thaumarchaeota</td>
<td></td>
<td>2.86</td>
<td></td>
</tr>
<tr>
<td>Carnobacteriaceae</td>
<td></td>
<td>2.85</td>
<td></td>
</tr>
<tr>
<td>Family/Monogram</td>
<td>Abundance</td>
<td></td>
<td></td>
</tr>
<tr>
<td>----------------------------------------</td>
<td>-----------</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Micrococcales</td>
<td>2.85</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Micrococcales</td>
<td>2.85</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sungkyunkwania</td>
<td>2.85</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Coriobacteriia</td>
<td>2.85</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Coriobacteriales</td>
<td>2.84</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Planctopirus</td>
<td>2.84</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Auritidibacter</td>
<td>2.83</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pibocella</td>
<td>2.83</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Coriobacteriaceae</td>
<td>2.82</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lactobacillales</td>
<td>2.80</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Clostridium_sensu_stricto_17</td>
<td>2.78</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Telluria</td>
<td>2.77</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sphingomonadales</td>
<td>2.77</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Croceicoccus</td>
<td>2.77</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Erythrobacteraceae</td>
<td>2.77</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Zunongwangia</td>
<td>2.74</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nonlabens</td>
<td>2.73</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bacillaceae</td>
<td>2.73</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Helicobacteraceae</td>
<td>2.73</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Glaciecola</td>
<td>2.72</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Desulfobacteraceae</td>
<td>2.71</td>
<td></td>
<td></td>
</tr>
<tr>
<td>DEV007</td>
<td>2.71</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cytophagales</td>
<td>2.71</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cytophagia</td>
<td>2.70</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Arcobacter</td>
<td>2.70</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Prevotellaceae</td>
<td>2.70</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Prevotella_7</td>
<td>2.69</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Neisseriaceae</td>
<td>2.69</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Stenoxymbacter</td>
<td>2.69</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cenarchaeales</td>
<td>2.68</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bacteroidia</td>
<td>2.68</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cenarchaeaceae</td>
<td>2.68</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Neisseriales</td>
<td>2.68</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Methanomicrobia</td>
<td>2.68</td>
<td></td>
<td></td>
</tr>
<tr>
<td>DEV007</td>
<td>2.68</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bacteroidales</td>
<td>2.68</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Anaerosporobacter</td>
<td>2.67</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cenarchaeum</td>
<td>2.67</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Enterobacteriales</td>
<td>2.66</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cyclobacteriaceae</td>
<td>2.66</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Moraxellaceae</td>
<td>2.66</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Moraxella</td>
<td>2.66</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pseudomonadales</td>
<td>2.66</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pasteurellaceae</td>
<td>2.66</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Solirubrobacterales</td>
<td>2.65</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Enterobacteriaceae</td>
<td>2.65</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pasteurellaceae</td>
<td>2.65</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rhodococcus</td>
<td>2.65</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Marine_Group_I</td>
<td>2.65</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Corynebacteriales</td>
<td>2.64</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Taxon</td>
<td>Rank</td>
<td></td>
<td></td>
</tr>
<tr>
<td>-------------------------------</td>
<td>----------</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Algoriphagus</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pasteurellales</td>
<td>2.64</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Deltaproteobacteria</td>
<td>2.63</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Epilithonimonas</td>
<td>2.62</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Thermoleophilia</td>
<td>2.62</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nocardiaceae</td>
<td>2.60</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Chiayiivirga</td>
<td>2.60</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Corynebacterium_1</td>
<td>2.58</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Chitinophagaceae</td>
<td>2.56</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Thaumarchaeota</td>
<td>2.54</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SS1_B_09_64</td>
<td>2.54</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dolosigranulum</td>
<td>2.54</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rickettsiaceae</td>
<td>2.53</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Candidatus_Cryptoprodotis</td>
<td>2.53</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rickettsiiales</td>
<td>2.53</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Thaumarchaeota</td>
<td>2.53</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Desulfovibionales</td>
<td>2.53</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SS1_B_09_64</td>
<td>2.52</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SEEP_SR4</td>
<td>2.50</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Chlamydiae</td>
<td>2.50</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Chlamydiiales</td>
<td>2.47</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Simkaniaiaceae</td>
<td>2.47</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Verrucomicrobiales</td>
<td>2.47</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Simkania</td>
<td>2.46</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sphingobacteriaceae</td>
<td>2.46</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Actinobacteria</td>
<td>2.45</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Desulfobulbaceae</td>
<td>2.45</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Klugiella</td>
<td>2.44</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Verrucomicrobiae</td>
<td>2.44</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Microbacteriaceae</td>
<td>2.43</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pedobacter</td>
<td>2.42</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Desulfovibionales</td>
<td>2.41</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rhizobium</td>
<td>2.41</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Euryarchaeota</td>
<td>2.39</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rhizobiaceae</td>
<td>2.38</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mollicutes_RF9</td>
<td>2.38</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mollicutes</td>
<td>2.37</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Thermoplasmata</td>
<td>2.37</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Thermoplasmatales</td>
<td>2.37</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Thermoplasmatales_Incertae_Sedis</td>
<td>2.37</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Spirochaetales</td>
<td>2.36</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mollicutes_RF9</td>
<td>2.36</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Spirochaetaceae</td>
<td>2.36</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tenericutes</td>
<td>2.36</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Spirochaetes</td>
<td>2.36</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Frankiales</td>
<td>2.35</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Treponema_2</td>
<td>2.35</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Geodermatophilaceae</td>
<td>2.34</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Xanthomonadales</td>
<td>2.34</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Xanthomonadaceae</td>
<td>2.34</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mollicutes_RF9</td>
<td>2.33</td>
<td></td>
<td></td>
</tr>
<tr>
<td>------------------------</td>
<td>------</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lentisphaeraceae</td>
<td>2.30</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lentisphaera</td>
<td>2.28</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lentisphaeraceae</td>
<td>2.28</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lentisphaerales</td>
<td>2.28</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lentisphaeria</td>
<td>2.27</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Elusimicrobium</td>
<td>2.26</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Elusimicrobiales</td>
<td>2.26</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Elusimicrobia</td>
<td>2.26</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Elusimicrobiaceae</td>
<td>2.26</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
**Supplementary Information S3:**

<table>
<thead>
<tr>
<th>Kingdom</th>
<th>Phylum</th>
<th>Class</th>
<th>% of uncl. OTUs</th>
<th>Range of rel. abundance (%)</th>
<th>Range of number of hosts</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bacteria</td>
<td>Verrucomicrobia</td>
<td>Verrucomicrobiae</td>
<td>10.75</td>
<td>0-19.0</td>
<td>16-46</td>
</tr>
<tr>
<td></td>
<td></td>
<td>WCHB1-41*</td>
<td>1.08</td>
<td>0-10.2</td>
<td>36</td>
</tr>
<tr>
<td></td>
<td></td>
<td>R76_B128*</td>
<td>1.08</td>
<td>0-0.6</td>
<td>27</td>
</tr>
<tr>
<td></td>
<td>Bacteroidetes</td>
<td>Bacteroidia*</td>
<td>9.68</td>
<td>0-5.3</td>
<td>15-38</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Cytophagia</td>
<td>9.68</td>
<td>0-11.4</td>
<td>16-53</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Sphingobacteriia</td>
<td>3.23</td>
<td>0-8.3</td>
<td>23-53</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Flavobacteriia</td>
<td>2.15</td>
<td>0-0.5</td>
<td>18-25</td>
</tr>
<tr>
<td></td>
<td>Firmicutes</td>
<td>Clostridia</td>
<td>9.68</td>
<td>0-4.2</td>
<td>15-28</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Erysipelotrichia*</td>
<td>1.08</td>
<td>0-2.3</td>
<td>22</td>
</tr>
<tr>
<td></td>
<td>Proteobacteria</td>
<td>Alphaproteobacteria</td>
<td>8.60</td>
<td>0-12.6</td>
<td>19-51</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Betaproteobacteria</td>
<td>1.08</td>
<td>0-4.1</td>
<td>15</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Deltaproteobacteria</td>
<td>8.06</td>
<td>0-1.0</td>
<td>15-35</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Epsilonproteobacteria</td>
<td>1.08</td>
<td>0-4.7</td>
<td>15</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Gammaproteobacteria</td>
<td>2.15</td>
<td>0-1.2</td>
<td>27-48</td>
</tr>
<tr>
<td></td>
<td>Cyanobacteria</td>
<td>Chloroplast</td>
<td>7.53</td>
<td>0-16.5</td>
<td>15-32</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Cyanobacteria</td>
<td>2.15</td>
<td>0-1.0</td>
<td>24-24</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Melainabacteria*</td>
<td>1.08</td>
<td>0-0.6</td>
<td>16</td>
</tr>
<tr>
<td></td>
<td></td>
<td>ML635J-21*</td>
<td>1.08</td>
<td>0-0.5</td>
<td>20</td>
</tr>
<tr>
<td></td>
<td>Planctomycetes</td>
<td>Phycisphaera*</td>
<td>3.23</td>
<td>0-1.7</td>
<td>17-79</td>
</tr>
<tr>
<td></td>
<td></td>
<td>OM190*</td>
<td>2.15</td>
<td>0-0.8</td>
<td>23-23</td>
</tr>
<tr>
<td></td>
<td></td>
<td>vadinHA49*</td>
<td>1.08</td>
<td>0-0.7</td>
<td>19</td>
</tr>
<tr>
<td></td>
<td>Actinobacteria</td>
<td>Acidimicrobiia*</td>
<td>1.08</td>
<td>0-4.7</td>
<td>50</td>
</tr>
<tr>
<td></td>
<td>Parcubacteria</td>
<td>Candidatus_Campbellbacteria*</td>
<td>1.08</td>
<td>0-1.6</td>
<td>19</td>
</tr>
<tr>
<td></td>
<td>Lentisphaerae</td>
<td>Oligosphaeria*</td>
<td>1.08</td>
<td>0-10.9</td>
<td>28</td>
</tr>
<tr>
<td></td>
<td>unclassified</td>
<td>unclassified</td>
<td>4.30</td>
<td>0-10.0</td>
<td>15-70</td>
</tr>
<tr>
<td></td>
<td>Archaea</td>
<td>Euryarchaeota</td>
<td>Thermoplasmata</td>
<td>4.30</td>
<td>0-1.0</td>
</tr>
</tbody>
</table>

**S3-Table: Class-level putative affiliation of the most common unclassified OTUs**

(unclassified OTUs at class level and found in at least 15 samples) found using Arb parsimony insertion tool. % of uncl. OTUs: Percentage of the selected OTUs belonging to a given class. Range of rel. abundance (%) is the minimum and the maximum relative abundance values of OTU(s) belonging to this class. Range of number of hosts is the minimum and the maximum number of individuals where OTUs belonging to this class were recovered. Single values in this last column indicate the number of individual where the only OTU belonging to this class was recovered. *Asterisks indicate classes that were initially not identified by Mothur in the whole dataset.
Supplementary Information S4:

S4-Fig: Microbial phylogenetic richness recovered in water samples and on fish surfaces. Boxes represent the interquartile range of alpha diversity values among water samples (blue), among all fish skin microbiomes (white) and among individuals of a given fish family (colored according to the fish order). Thick bars represent the median of alpha-diversity values, and horizontal segments extend the most extreme data point which is no more than 1.5 times the length of the box away from the box. In the case of Sphyraenidae, Monacanthidae and Diodontidae, only one individual was sampled, and is represented by a single dot.
S4-Table 1: Determinants of microbial diversity recovered in water and on fish surfaces

<table>
<thead>
<tr>
<th>Factor</th>
<th>Phylogenetic richness</th>
<th>Phylogenetic entropy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Water vs. Fish</td>
<td>1</td>
<td>0.003</td>
</tr>
<tr>
<td>Fish Species</td>
<td>43</td>
<td>&lt;0.001</td>
</tr>
</tbody>
</table>

Effect of sample type (water vs. fish skin) and fish species on each facet of microbial phylogenetic diversity (richness and entropy) were tested using two separated non-parametric Kruskal-Wallis tests on all 173 (fish skin + water) samples, and on all 138 fish individuals, respectively.

S4-Table 2: Effect of fish phylogeny on the diversity of its skin microbiome

<table>
<thead>
<tr>
<th>Index of alpha-diversity</th>
<th>Moran’s I</th>
<th>% of sign. P-values (&lt;0.05)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Phylogenetic richness (Faith’s PD)</td>
<td>0.03±0.02</td>
<td>96.2</td>
</tr>
<tr>
<td>Phylogenetic entropy (Allen’s index)</td>
<td>0.02±0.02</td>
<td>88.9</td>
</tr>
</tbody>
</table>

The effect of fish phylogeny on its skin-associated microbial diversity was assessed using Moran’s I index of autocorrelation scaling between -1 and 1, 1 meaning that there is a strong correlation between phylogeny and the trait value tested. I values were compared to I values obtained from a null model were diversity values were randomly shuffled on fish phylogeny (999 replicates). Mean and associated standard deviation of Moran’s I on 999 subsamples are provided, as well as the percentage of subsamples where the correlation was significant (P<0.05).
Supplementary Information S5:

S5-Fig: Abundance-unweighted (left) and weighted (right) phylogenetic dissimilarity values, among water samples (n=35 samples), between fish samples and water samples (n=173), between individuals of the same fish species (n=34 species with more than 1 individual), and among individuals from different species (n=44 species). Boxes represent the interquartile range dissimilarity values. Thick bars represent the median of dissimilarity values, and horizontal segments extend the most extreme data point which is no more than 1.5 times the length of the box away from the box.
Supplementary Information S6:

**S6-Fig: Weighted-Unifrac between fish skin-associated microbiomes and divergence times between fish species.** One individual of each of the 44 fish species was randomly selected; each point representing one dissimilarity value between two individuals. ‘Intra-order’ means individuals belonging to species from the same taxonomic order; ‘inter-order’ means individuals belonging to species from different taxonomic orders.

**S6-Table: Test for phylosymbiosis**

<table>
<thead>
<tr>
<th>Fish species included</th>
<th>Dissimilarity index</th>
<th>P-value</th>
<th>% of sign. P-values</th>
</tr>
</thead>
<tbody>
<tr>
<td>All species</td>
<td>Unweighted Unifrac</td>
<td>0.28 ± 0.19</td>
<td>8.4</td>
</tr>
<tr>
<td></td>
<td>Weighted Unifrac</td>
<td>0.43 ± 0.18</td>
<td>0</td>
</tr>
<tr>
<td>Perciformes only</td>
<td>Unweighted Unifrac</td>
<td>0.48 ± 0.23</td>
<td>0.7</td>
</tr>
<tr>
<td></td>
<td>Weighted Unifrac</td>
<td>0.68 ± 0.22</td>
<td>0</td>
</tr>
</tbody>
</table>

*Phylosymbiosis was tested using a Mantel test (999 permutations) for correlation between hosts’s phylogeny and dissimilarity values between microbiomes.* Mean and associated standard deviation of P-values obtained on 999 subsamples are provided, as well as the percentage of subsamples where the correlation was significant (P<0.05).
Supplementary Information S7:

S7-Table: Results of BDTT (Beta-Diversity Through Time) assessing the correlation between microbiome variability and hosts’ phylogeny across bacterial phylogenetic tree

<table>
<thead>
<tr>
<th>Time (Mya)</th>
<th>Sorensen Index</th>
<th>Bray-Curtis Index</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>% of sign. P-values</td>
<td>$R^2$</td>
</tr>
<tr>
<td>0</td>
<td>2.2</td>
<td>0.005±0.01</td>
</tr>
<tr>
<td>100</td>
<td>2.8</td>
<td>0.006±0.01</td>
</tr>
<tr>
<td>200</td>
<td>4.8</td>
<td>0.006±0.01</td>
</tr>
<tr>
<td>300</td>
<td>9.6</td>
<td>0.006±0.01</td>
</tr>
<tr>
<td>400</td>
<td>9.8</td>
<td>0.007±0.01</td>
</tr>
<tr>
<td>500</td>
<td>12.2</td>
<td>0.007±0.01</td>
</tr>
<tr>
<td>600</td>
<td>8.6</td>
<td>0.006±0.01</td>
</tr>
<tr>
<td>700</td>
<td>2.8</td>
<td>0.005±0.01</td>
</tr>
<tr>
<td>800</td>
<td>2.2</td>
<td>0.004±0.01</td>
</tr>
<tr>
<td>900</td>
<td>1.6</td>
<td>0.004±0.01</td>
</tr>
</tbody>
</table>

This method is from Groussin et al. Nature Communication 2016, (see Supplementary Methods) for 9 periods of 100 millions of year from present time (t=0) to 900 Mya in the past. At each period of time, BDTT computes microbiome variability using both abundance-weighted (Bray Curtis dissimilarity) and unweighted dissimilarity indices (Sorensen index) and assesses their correlation with hosts’ divergence times using a Mantel test. Results of the Mantel tests show that even when considering deep prokaryotic phylogenetic clades (i.e. 900 Mya correspond roughly to the apparition of Bilaterians and divergence of bacterial orders), there is no correlation between host phylogeny and the structure and composition of its associated skin microbiome. Mean and associated standard deviation of R-squared values obtained on 999 subsamples are provided, as well as the percentage of subsamples where the correlation was significant (P<0.05).
Supplementary Information S8:

S8-Table: Effect of reef type and fish species on fish skin microbial community structure

<table>
<thead>
<tr>
<th>Factor</th>
<th>Unweighted Unifrac</th>
<th>Weighted Unifrac</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>P</td>
<td>R²</td>
</tr>
<tr>
<td>Reef type</td>
<td>0.001</td>
<td>0.02</td>
</tr>
<tr>
<td>Fish species</td>
<td>0.001</td>
<td>0.27</td>
</tr>
<tr>
<td>Reef type * Fish species</td>
<td>0.020</td>
<td>0.20</td>
</tr>
</tbody>
</table>

The effect of reef type versus the one of fish species was assessed by computing a permutational ANOVAs (PERMANOVAs) on abundance weighted Unifrac and unweighted Unifrac dissimilarities between the 74 individuals that belonged to the 16 fish species that were sampled on both sites.
Supplementary Information S9:

S9-Fig: Distribution of functional trait values among fish phylogeny. Traits are encoded as ordinal variables as indicated in S8-Table.

S9-Table: Encoding of functional traits tested in this study

<table>
<thead>
<tr>
<th>Functional traits</th>
<th>Encoding</th>
</tr>
</thead>
<tbody>
<tr>
<td>Size at maturity</td>
<td>1: 7.1-15 cm; 2: 15.1-30cm; 3: 30.1-50cm; 4: 30.1-50cm; 5: 50.1-80cm; 6: &gt;80cm</td>
</tr>
<tr>
<td>Mobility</td>
<td>1: sedentary; 2: mobile within a reef; 3: highly mobile, i.e. between reefs</td>
</tr>
<tr>
<td>Activity</td>
<td>1: diurnal; 2: diurnal and nocturnal; 3: nocturnal</td>
</tr>
<tr>
<td>Schooling</td>
<td>1: solitary; 2: pairing; 3: small group; 4: medium group; 5: large group</td>
</tr>
<tr>
<td>Position</td>
<td>1: sea bottom; 2: above sea bottom; 3: pelagic</td>
</tr>
</tbody>
</table>

FC: Pelagic macro-organisms, i.e. large organisms living in the water column, as well as benthic fishes
IM: Mobile invertebrates, i.e. all-size mobile benthic invertebrates
IS: Sessile invertebrates, i.e. corals, sponges, ascidians and all other sessile benthic invertebrates

PK: Planktivorous fishes are fishes eating planktonic and small benthic organisms that can migrate in the water column (copepods, crustacean larvae…)
HM: Herbivorous macro-algal fishes are fishes eating macro-algae and sea grass
HD: Herbivorous-detritivorous fishes eat undefined or detrital material
OM: Omnivorous fishes regroup fishes that are both herbivorous or detritivorous, and carnivorous
**Supplementary Information S10:**

**S10-Fig1: Minimum Spanning Trees (MST) showing fish diets** based on Unweighted-Unifrac dissimilarity values (A), or abundance weighted-Unifrac dissimilarity values (B) calculated on mean microbiome per fish species. Species were colored according to fish main prey (S8). The closer the species are on the tree, the more similar are their microbiomes. Complete names of each species are provided in Fig 1 of the manuscript, and in S8.

**S10-Fig2: PCoA plots** representing all individuals included in this study, based on (A) unweighted-Unifrac dissimilarity values and (B) abundance weighted-Unifrac dissimilarity values between microbiomes. Individuals were colored according to their main prey (S8).
Supplementary Information S11: Investigation of an eventual transfer from sessile invertebrates-associated cells to fish skin microbiome

At the same time of sampling and on the same sites, 83 sessile invertebrates were sampled, representing a wide range of clades, i.e. scleratinian corals, soft corals, sponges, anemones and gorgonians. Surface microbiome was sampled, either by sampling of surface mucus, or swabbing, and was analyzed using the same methods as for fish skin swabs. We computed the phylogenetic dissimilarity between surface microbiome of these sessile organisms, and fish skin surface microbiome, depending on fish diet.

S11-Fig: Abundance unweighted-Unifrac (A) and weighted-Unifrac (B) dissimilarity values between sessile invertebrates’ microbiomes and fish skin microbiomes, separated by fish diets. Both phylogenetic dissimilarities were significantly different between different fish diets (Kruskal-Wallis test, P<0.001 for both U- and W-Unifrac). Post-hoc pairwise comparisons of dissimilarities among different diets are represented by letters, two different letters indicating a significant difference between dissimilarity values between diets (post-hoc pairwise comparisons available in pgirmess package, P<0.05). Fishes eating sessile invertebrates hosted a microbiome not significantly closer to the one of sessile invertebrates, than fishes having contrasting diets.
Supplementary Information S12: Relation between fish vulnerability to anthropic pressures and microbiome diversity

S12-Table: Spearman’s correlation tests between phylogenetic diversity of microbiome and vulnerability of the 44 fish species

<table>
<thead>
<tr>
<th>Anthropic pressure</th>
<th>Phylogenetic richness</th>
<th>Phylogenetic entropy</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>rho</td>
<td>% of sign. P-values</td>
</tr>
<tr>
<td>Overfishing</td>
<td>0.43 ± 0.08</td>
<td>95.1</td>
</tr>
<tr>
<td>Habitat loss</td>
<td>-0.33 ± 0.07</td>
<td>69.8</td>
</tr>
</tbody>
</table>

999 correlation tests were performed on 999 subsamples of one individual per fish species (See Supplementary Methods). Mean and associated standard deviation of Spearman’s rho values on 999 subsamples are provided, as well as the percentage of subsamples where the correlation was significant (P<0.05).

S12-Fig: Mean phylogenetic richness (measured by relative Faith’s PD) recovered on fish species (n=44 species, 1 individual randomly subsampled), and fish vulnerability to overfishing (left) and to habitat loss (right).
Supplementary Information S13: Rarefaction curves

S13-Fig1: Rarefaction curves obtained from each water replicate. Sample names starting by ‘B-’ were taken on barrier reef, while sample names starting by ‘F-’ were taken on fringing reef.
S13-Fig2: Rarefaction curves for all fish individuals. The most undersampled species (showing a non-flat curve) were also the phylogenetically richer ones (*Pterois miles*, *Pterois radiata*, *Parupeneus cyclostomus*, *Chlorurus sordidus* and *Cephalophis boenak*, see Fig 1).
Supplementary Information S14: Host phylogenetic tree and branching of species not initially included in Rabosky et al. (2013) phylogenetic tree

A. Rabosky et al. (2013) modified phylogenetic tree with added species in Newick format

(\textit{Supplementary Information S14: Host phylogenetic tree and branching of species not initially included in Rabosky et al. (2013) phylogenetic tree})

\textbf{A. Rabosky et al. (2013) modified phylogenetic tree with added species in Newick format}

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{newick_tree.png}
\caption{Newick format of the modified phylogenetic tree.}
\end{figure}
### B. Branching information of added species and references

13 species were not initially present in Rabosky et al. phylogenetic tree, and one 
(*Cephalophilis argus*) was incorrectly branched between Scaridae and Caesionidae. To define the branching emplacement of these species, we searched for resolved phylogenies of fish families, and searched for the species’ closest relative already included in Rabosky’s tree. We then placed our species in replacement to its closest relative. When possible, we favored studies using genetic data based on multiple loci. No published phylogenetic tree of Caesionidae and Mullidae based on genetic data could be found; therefore we used studies based on morphological and anatomical traits:

<table>
<thead>
<tr>
<th>Species added</th>
<th>Branching emplacement</th>
<th>Type of data</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>Acanthurus leucosternon</em></td>
<td>In replacement of <em>Acanthurus nigricans</em></td>
<td>Genetic</td>
<td>(Sorenson et al. 2013)</td>
</tr>
<tr>
<td><em>Acanthurus lineatus</em></td>
<td>On a separated branch, connected to the common node of all other <em>Acanthurus sp.</em></td>
<td>Genetic</td>
<td>(Sorenson et al. 2013)</td>
</tr>
<tr>
<td><em>Caesio xanthonota</em></td>
<td>In replacement of <em>Caesio caerulea</em></td>
<td>Morphology</td>
<td>(Carpenter 1990)</td>
</tr>
<tr>
<td><em>Cantherhines pardalis</em></td>
<td>In replacement of <em>Cantherinhes pullus</em></td>
<td>Genetic</td>
<td>(Santini, Sorenson, and Alfaro 2013)</td>
</tr>
<tr>
<td><em>Corythoichthys flavofasciatus</em></td>
<td>In replacement of <em>Corythoichthys intestinalis</em></td>
<td>No data*</td>
<td></td>
</tr>
<tr>
<td><em>Cephalopholis argus</em></td>
<td>In replacement of <em>Cephalopholis nigri</em></td>
<td>Genetic</td>
<td>(Craig and Hastings 2007)</td>
</tr>
<tr>
<td><em>Myripristis murdjan</em></td>
<td>Branched to the node separating <em>Myripristis violacea</em> from all other species</td>
<td>Genetic</td>
<td>(Dornburg et al. 2012)</td>
</tr>
<tr>
<td><em>Parapercis hexophtalma</em></td>
<td>In replacement of <em>Parapercis clathrata</em></td>
<td>No data*</td>
<td></td>
</tr>
<tr>
<td><em>Parupeneus cyclostomus</em></td>
<td>On a separated branch, connected to the common node of all <em>Upeneichthys sp.</em></td>
<td>Morphology</td>
<td>(Kim 2002)</td>
</tr>
<tr>
<td><em>Parupeneus trifasciatus</em></td>
<td>Branched to the node separating <em>Parupeneus cyclostomus</em> from all other species</td>
<td>Morphology</td>
<td>(Kim 2002)</td>
</tr>
<tr>
<td><em>Pterocaesio tile</em></td>
<td>In replacement of <em>Pterocaesio digramma</em></td>
<td>Morphology</td>
<td>(Carpenter 1990)</td>
</tr>
<tr>
<td><em>Pterocaesio trillienata</em></td>
<td>In replacement of <em>Pterocaesio marri</em></td>
<td>Morphology</td>
<td>(Carpenter 1990)</td>
</tr>
<tr>
<td><em>Scarus caudofasciatus</em></td>
<td>In replacement of <em>Scarus spinus</em></td>
<td>Genetic</td>
<td>(Choat et al. 2012)</td>
</tr>
<tr>
<td><em>Scarus russelii</em></td>
<td>In replacement of <em>Scarus flavipectoralis</em></td>
<td>Genetic</td>
<td>(Choat et al. 2012)</td>
</tr>
</tbody>
</table>

*To our knowledge, no phylogenetic tree of Syngnathidae including *Corythoichthys flavofasciatus* has been published yet. Therefore, we placed our species in replacement of one of its congenerics. The same problem appeared for *Parapercis hexophtalma*, which was
placed in replacement of its only congeneric that also occurs in Indian ocean (Taquet and Diringer 2007)

C. References


SUPPLEMENTARY INFORMATION

Macroscopic biodiversity of coral reefs supports a high, unique and vulnerable microscopic biodiversity

Chiarello M., Auguet J-C., Claverie T., Sucré E., Nicholas AJ Graham, Rilleuvineuve F., Bouvier C., Villéger S. and Bouvier T.
**Supplementary Information S1:**

<table>
<thead>
<tr>
<th>Taxonomic affiliation</th>
<th>Identification</th>
<th>Nb of samples</th>
<th>Sample Type</th>
</tr>
</thead>
<tbody>
<tr>
<td>Teleostei</td>
<td>Acanthurus leucosternon</td>
<td>5</td>
<td>Swab</td>
</tr>
<tr>
<td>Teleostei</td>
<td>Acanthurus lineatus</td>
<td>3</td>
<td>Swab</td>
</tr>
<tr>
<td>Teleostei</td>
<td>Ctenochaetus striatus</td>
<td>7</td>
<td>Swab</td>
</tr>
<tr>
<td>Teleostei</td>
<td>Naso unicornis</td>
<td>1</td>
<td>Swab</td>
</tr>
<tr>
<td>Teleostei</td>
<td>Balistapus undulatus</td>
<td>3</td>
<td>Swab</td>
</tr>
<tr>
<td>Teleostei</td>
<td>Sufflamen chrysopterum</td>
<td>6</td>
<td>Swab</td>
</tr>
<tr>
<td>Teleostei</td>
<td>Pterocaesio tile</td>
<td>1</td>
<td>Swab</td>
</tr>
<tr>
<td>Teleostei</td>
<td>Pterocaesio trilienata</td>
<td>2</td>
<td>Swab</td>
</tr>
<tr>
<td>Teleostei</td>
<td>Caranx melampygus</td>
<td>3</td>
<td>Swab</td>
</tr>
<tr>
<td>Teleostei</td>
<td>Chaetodon auriga</td>
<td>2</td>
<td>Swab</td>
</tr>
<tr>
<td>Teleostei</td>
<td>Chaetodon falcula</td>
<td>6</td>
<td>Swab</td>
</tr>
<tr>
<td>Teleostei</td>
<td>Chaetodon lunula</td>
<td>4</td>
<td>Swab</td>
</tr>
<tr>
<td>Teleostei</td>
<td>Chaetodon meyeri</td>
<td>3</td>
<td>Swab</td>
</tr>
<tr>
<td>Teleostei</td>
<td>Chaetodon trifascialis</td>
<td>3</td>
<td>Swab</td>
</tr>
<tr>
<td>Teleostei</td>
<td>Forcipiger flavissimus</td>
<td>3</td>
<td>Swab</td>
</tr>
<tr>
<td>Teleostei</td>
<td>Arothron nigropunctatus</td>
<td>1</td>
<td>Swab</td>
</tr>
<tr>
<td>Teleostei</td>
<td>Platax orbicularis</td>
<td>5</td>
<td>Swab</td>
</tr>
<tr>
<td>Teleostei</td>
<td>Platax teira</td>
<td>1</td>
<td>Swab</td>
</tr>
<tr>
<td>Teleostei</td>
<td>Myripristis mardjan</td>
<td>2</td>
<td>Swab</td>
</tr>
<tr>
<td>Teleostei</td>
<td>Myripristis violacea</td>
<td>1</td>
<td>Swab</td>
</tr>
<tr>
<td>Teleostei</td>
<td>Kyphosus vaigiensis</td>
<td>3</td>
<td>Swab</td>
</tr>
<tr>
<td>Teleostei</td>
<td>Cheilinus fasciatus</td>
<td>3</td>
<td>Swab</td>
</tr>
<tr>
<td>Teleostei</td>
<td>Hemigymnus fasciatus</td>
<td>3</td>
<td>Swab</td>
</tr>
<tr>
<td>Teleostei</td>
<td>Thalassoma hebraicum</td>
<td>3</td>
<td>Swab</td>
</tr>
<tr>
<td>Teleostei</td>
<td>Monotaxis grandoculis</td>
<td>5</td>
<td>Swab</td>
</tr>
<tr>
<td>Teleostei</td>
<td>Cantherhines pardalis</td>
<td>1</td>
<td>Swab</td>
</tr>
<tr>
<td>Teleostei</td>
<td>Parupeneus cyclostomus</td>
<td>2</td>
<td>Swab</td>
</tr>
<tr>
<td>Teleostei</td>
<td>Parupeneus trifasciatus</td>
<td>3</td>
<td>Swab</td>
</tr>
<tr>
<td>Teleostei</td>
<td>Parapercis hexophtalma</td>
<td>3</td>
<td>Swab</td>
</tr>
<tr>
<td>Teleostei</td>
<td>Pomacanthus imperator</td>
<td>3</td>
<td>Swab</td>
</tr>
<tr>
<td>Teleostei</td>
<td>Pygoplites diacanthus</td>
<td>6</td>
<td>Swab</td>
</tr>
<tr>
<td>Teleostei</td>
<td>Abudefdaf sexfasciatus</td>
<td>3</td>
<td>Swab</td>
</tr>
<tr>
<td>Teleostei</td>
<td>Abudefdaf sparoide</td>
<td>4</td>
<td>Swab</td>
</tr>
<tr>
<td>Teleostei</td>
<td>Amphiprion akallopis</td>
<td>3</td>
<td>Swab</td>
</tr>
<tr>
<td>Teleostei</td>
<td>Chlorurus sordidus</td>
<td>6</td>
<td>Swab</td>
</tr>
<tr>
<td>Teleostei</td>
<td>Scarus caudofasciatus</td>
<td>4</td>
<td>Swab</td>
</tr>
<tr>
<td>Teleostei</td>
<td>Scarus russelli</td>
<td>1</td>
<td>Swab</td>
</tr>
<tr>
<td>Teleostei</td>
<td>Pterois miles</td>
<td>2</td>
<td>Swab</td>
</tr>
<tr>
<td>Teleostei</td>
<td>Pterois radiata</td>
<td>1</td>
<td>Swab</td>
</tr>
<tr>
<td>Teleostei</td>
<td>Cephalopholis argus</td>
<td>6</td>
<td>Swab</td>
</tr>
<tr>
<td>Teleostei</td>
<td>Cephalopholis boenak</td>
<td>1</td>
<td>Swab</td>
</tr>
<tr>
<td>Kingdom</td>
<td>Family</td>
<td>Species</td>
<td>Quantity</td>
</tr>
<tr>
<td>------------</td>
<td>----------------------</td>
<td>--------------------------</td>
<td>----------</td>
</tr>
<tr>
<td>Teleostei</td>
<td>Sphyraena barracuda</td>
<td>1</td>
<td>Swab</td>
</tr>
<tr>
<td>Teleostei</td>
<td>Coryphaenichthys flavofasciatus</td>
<td>3</td>
<td>Swab</td>
</tr>
<tr>
<td>Teleostei</td>
<td>Zanclus cornutus</td>
<td>6</td>
<td>Swab</td>
</tr>
<tr>
<td><strong>Total Teleostei</strong></td>
<td></td>
<td><strong>138</strong></td>
<td></td>
</tr>
<tr>
<td>Anthozoa</td>
<td>Acropora cf. clathrata</td>
<td>2</td>
<td>Mucus</td>
</tr>
<tr>
<td>Anthozoa</td>
<td>Acropora cf. cytherea</td>
<td>3</td>
<td>Mucus</td>
</tr>
<tr>
<td>Anthozoa</td>
<td>Acropora cf. elseyi</td>
<td>3</td>
<td>Swab</td>
</tr>
<tr>
<td>Anthozoa</td>
<td>Acropora cf. intermedia</td>
<td>3</td>
<td>Mucus</td>
</tr>
<tr>
<td>Anthozoa</td>
<td>Acropora cf. latistella</td>
<td>3</td>
<td>Mucus</td>
</tr>
<tr>
<td>Anthozoa</td>
<td>Acropora cf. maricata</td>
<td>3</td>
<td>Mucus</td>
</tr>
<tr>
<td>Anthozoa</td>
<td>Acropora cf. nasuta</td>
<td>2</td>
<td>Mucus</td>
</tr>
<tr>
<td>Anthozoa</td>
<td>Echinopora sp</td>
<td>2</td>
<td>Mucus</td>
</tr>
<tr>
<td>Anthozoa</td>
<td>Favia sp</td>
<td>4</td>
<td>Mucus</td>
</tr>
<tr>
<td>Anthozoa</td>
<td>Favites sp</td>
<td>2</td>
<td>Mucus</td>
</tr>
<tr>
<td>Anthozoa</td>
<td>Fungia sp</td>
<td>6</td>
<td>Mucus</td>
</tr>
<tr>
<td>Anthozoa</td>
<td>Goniopora sp</td>
<td>3</td>
<td>1 mucus 2 swabs</td>
</tr>
<tr>
<td>Anthozoa</td>
<td>Herpolitha sp</td>
<td>3</td>
<td>Mucus</td>
</tr>
<tr>
<td>Anthozoa</td>
<td>Isopora sp</td>
<td>3</td>
<td>1 mucus 2 swabs</td>
</tr>
<tr>
<td>Anthozoa</td>
<td>Lobophyllia sp</td>
<td>2</td>
<td>Swab</td>
</tr>
<tr>
<td>Anthozoa</td>
<td>Montastrea sp</td>
<td>3</td>
<td>Mucus</td>
</tr>
<tr>
<td>Anthozoa</td>
<td>Montipora sp</td>
<td>3</td>
<td>2 mucus 1 swab</td>
</tr>
<tr>
<td>Anthozoa</td>
<td>Platygyra sp</td>
<td>6</td>
<td>Mucus</td>
</tr>
<tr>
<td>Anthozoa</td>
<td>Porites sp (branching)</td>
<td>3</td>
<td>Mucus</td>
</tr>
<tr>
<td>Anthozoa</td>
<td>Porites sp (massive)</td>
<td>5</td>
<td>3 mucus 2 swabs</td>
</tr>
<tr>
<td>Anthozoa</td>
<td>Gorgonidae</td>
<td>5</td>
<td>Swab</td>
</tr>
<tr>
<td>Anthozoa</td>
<td>Sarcophyton</td>
<td>9</td>
<td>5 mucus 4 swabs</td>
</tr>
<tr>
<td>Anthozoa</td>
<td>Sinularia</td>
<td>4</td>
<td>Swab</td>
</tr>
<tr>
<td>Anthozoa</td>
<td>Heteractis magnifica</td>
<td>2</td>
<td>Swab</td>
</tr>
<tr>
<td><strong>Total Anthozoa</strong></td>
<td></td>
<td><strong>84</strong></td>
<td></td>
</tr>
<tr>
<td>Malacostraca</td>
<td>Paguroidea sp</td>
<td>5</td>
<td>Swab</td>
</tr>
<tr>
<td>Malacostraca</td>
<td>Panulirus versicolor</td>
<td>3</td>
<td>Swab</td>
</tr>
<tr>
<td>Crinoidea</td>
<td>Comatulida sp</td>
<td>3</td>
<td>Swab</td>
</tr>
<tr>
<td>Holothuroidea</td>
<td>Bohadschia sp</td>
<td>3</td>
<td>Swab</td>
</tr>
<tr>
<td>Echinoidea</td>
<td>Echinothrix diadema</td>
<td>2</td>
<td>Swab</td>
</tr>
<tr>
<td>Holothuroidea</td>
<td>Holothuria atra</td>
<td>3</td>
<td>Swab</td>
</tr>
<tr>
<td>Asteroidea</td>
<td>Linckia laevigata</td>
<td>6</td>
<td>Swab</td>
</tr>
<tr>
<td>Asteroidea</td>
<td>Linckia multifora</td>
<td>2</td>
<td>Swab</td>
</tr>
<tr>
<td>Ophiuroidea</td>
<td>Ophiaroidea sp</td>
<td>6</td>
<td>Swab</td>
</tr>
<tr>
<td>Bivalvia</td>
<td>Tridacna sp</td>
<td>3</td>
<td>Swab</td>
</tr>
<tr>
<td>Demospongia</td>
<td>Spheciospongia cf. inconstans</td>
<td>3</td>
<td>Swab</td>
</tr>
<tr>
<td>Demospongia</td>
<td>Cliona cf. vastifica</td>
<td>4</td>
<td>Swab</td>
</tr>
<tr>
<td><strong>Total Other Invertebrates</strong></td>
<td></td>
<td><strong>43</strong></td>
<td></td>
</tr>
<tr>
<td><strong>TOTAL SAMPLES</strong></td>
<td></td>
<td><strong>300</strong></td>
<td></td>
</tr>
</tbody>
</table>
**S1:** Description of all samples included in this study. In the “identification” column, we indicated the finest level of identification we could reach for each animal.
Supplementary Information S2:

S2-Fig 1: Phylogenetic entropy of reef microbial communities

Phylogenetic entropy measured using Allen’s index of animal surface-associated microbiomes (e.g. on one fish or coral colony) and of planktonic communities (i.e. in each water sample) Boxes represent the interquartile range of phylogenetic richness among communities. Thick bars represent the median of phylogenetic richness, and vertical segments extend to the most extreme data point that is no more than 1.5 times the length of the box away from the box. Differences between both compartments (animal surfaces and planktonic communities) are congruent with those found with phylogenetic richness (Kruskal-Wallis, P<0.05; Fig 1A).
S2-Fig 2: Phylogenetic dissimilarity between microbial communities

(A) Microbial communities plotted on the two first axes from a PCoA computed on pairwise weighted Unifrac dissimilarities between all communities, the closer two points the lowest the dissimilarity in phylogenetic composition of microbial communities, and (B) average intra- and inter-group variability of animal surfaces-associated microbiome, measured with weighted Unifrac. See S2-Fig 2 for analyzes of dissimilarity in phylogenetic structure using unweighted Unifrac index.

Results are similar to those presented in Figure 3. Mean intra-compartment (within animal surface-associated microbiomes or within planktonic communities) variability averaged 0.34±0.12 and 0.64±0.10, respectively. Planktonic communities showed significantly less variability than animal surface-associated microbiomes (W-Unifrac, PERMDISP, P<0.0001). Inter-group (between fishes, Anthozoa and other invertebrates) variability was 1.06 times higher than the intra-group one (within fishes, Anthozoa and other invertebrates, W-Unifrac). Intra-group dissimilarity averaged 0.63±0.11 among teleostean fishes, 0.57±0.09 among Anthozoa, and 0.67±0.10 among other invertebrates (W-Unifrac, KW and associated post-hoc tests, P<0.05 in all pairwise comparisons).
Supplementary Information S3:

S3: Accumulation curves showing the OTU richness (top) and the phylogenetic richness (bottom) obtained from animal surfaces-associated microbiomes and planktonic communities, depending on the number of randomly selected planktonic communities or animal taxa. Richness values are represented as the mean ± the standard deviation across the 100 replicates.
Supplementary Information S4:

S4: LefSe analysis showing significant biomarkers distinguishing animal surface-associated microbiomes and planktonic communities. Only microbial taxa that raised an LDA score > 4 are shown.
S5: Class-level taxonomic structure of microbial assemblages from surface of fishes, Anthozoa, and other invertebrates and from planktonic communities.
Supplementary Information S6:

S6: Composition of OTUs that are recovered in at least half of animal taxa (“Core” OTUs). OTUs were grouped depending on their respective family and class, and the percentage of OTUs belonging to each taxonomic clade is indicated.
Supplementary Information S7:

(A) Black points and bars represent mean (±SD) of remaining microbial OTU richness on the studied coral reef for a given proportion of coral and fish taxa lost. Fishes and corals loss were simulated according to their decreasing vulnerability to habitat loss and fishing, and bleaching, respectively (100 replicates). When 100% of coral and fish species are lost, the remaining microbial phylogenetic richness corresponds to the one of planktonic communities, and communities associated to the surface of animal taxa that were not included in the extinction scenario because their vulnerability was unknown (*i.e.* the scleratinian *Isopora*, soft corals, gorgonians, anemone and all non-anthozoan invertebrates). The scenario simulating a random loss of fishes and corals (*i.e.* species randomly removed independently from their vulnerability), is illustrated with the grey area that represent range of remaining phylogenetic richness among 100 replicates.

S7: Vulnerability of coral reef microbial diversity to loss of fishes and scleratinian corals. (A) Black points and bars represent mean (±SD) of remaining microbial OTU richness on the studied coral reef for a given proportion of coral and fish taxa lost. Fishes and corals loss were simulated according to their decreasing vulnerability to habitat loss and fishing, and bleaching, respectively (100 replicates). When 100% of coral and fish species are lost, the remaining microbial phylogenetic richness corresponds to the one of planktonic communities, and communities associated to the surface of animal taxa that were not included in the extinction scenario because their vulnerability was unknown (*i.e.* the scleratinian *Isopora*, soft corals, gorgonians, anemone and all non-anthozoan invertebrates). The scenario simulating a random loss of fishes and corals (*i.e.* species randomly removed independently from their vulnerability), is illustrated with the grey area that represent range of remaining phylogenetic richness among 100 replicates.

311
(B) Slope of the loss in phylogenetic richness for the “warming+fishing” scenario, is represented as the mean ± SD of slopes calculated at each level of extinction on all 100 replicates.
ABSTRACT

Crassostrea gigas, the main oyster species exploited worldwide, suffers from devastating mortality outbreaks, whose severity has dramatically increased since 2008 in Europe. It particularly affects juvenile stages decimating up to 90% of spat in some farms. A series of works published this last decade have underscored the complexity of this pathosystem which remained to be fully characterized. By combining ecologically realistic experimental procedures and thorough molecular analysis of oyster full-sib families with contrasted susceptibilities to the disease, we identified here the mechanisms of the pathogenesis affecting juvenile oysters. We showed that OsHV-1 infection leads to an immunodeficiency that favors a secondary bacterial infection, which is fatal. In addition, we identified some molecular determinants of the resistance that should give solutions to improve oyster culture in the near future.

Mots-clés : microbiome, diversité phylogénétique, Séquençage haut-débit, communautés microbiennes, peau, Teleostei, Odontoceti, écosystèmes coralliens

Abstract : Oceans contain thousands of microbial species playing crucial roles for the functioning of the marine ecosystem. These microorganisms are present everywhere in the water column. Some microorganisms also colonize the surface and the digestive tract of marine macroorganisms, forming communities called microbiomes. These microbiomes have positive effects for their host’s fitness. The diversity of these marine animal surface microbiome is still largely understudied, despite recent progress in molecular biology that now permits to fully assess its different facets of biodiversity, i.e. taxonomic, phylogenetic and functional. The goal of this thesis is therefore to describe the diversity of the surface microbiome of marine animals, to assess its variability at different levels, as well as its determinants, and the significance of such diversity at the ecosystem’s scale. Firstly, I have assessed the efficiency of various diversity indices to detect ecological signals in the specific case of microbial communities. Secondly, I have described the surface microbiome of major marine animal clades (teleostean fishes, cetaceans and several classes of invertebrates). I found that these microbiomes are highly distinct from the surrounding planktonic communities. I demonstrated that these microbiomes are variable both between individuals from the same species and between species, but do not show a phyllosymbiosis pattern. Last, I assessed the contribution of surface microbiomes to the global microbial community at the scale of a coral reef ecosystem. I demonstrated that marine animal surfaces host almost twenty times more microbial species than the water column, and 75% of the phylogenetic richness present in the ecosystem. In a context of massive erosion of marine macroscopic organisms, it is therefore urgent to exhaustively assess marine microbial biodiversity and its vulnerability facing anthropic pressures.

Key-words : microbiome, phylogenetic diversity, Next Generation Sequencing (NGS), microbial communities, skin, Teleostei, Odontoceti, coral reef ecosystems