



**HAL**  
open science

# Caractérisation génétique et moléculaire d'une incompatibilité nucléo-cytoplasmique chez *Arabidopsis* *Thaliana*

Nicolas Gobron

► **To cite this version:**

Nicolas Gobron. Caractérisation génétique et moléculaire d'une incompatibilité nucléo-cytoplasmique chez *Arabidopsis Thaliana*. Sciences agricoles. Université Paris Sud - Paris XI, 2012. Français. NNT : 2012PA112077 . tel-00713647

**HAL Id: tel-00713647**

**<https://theses.hal.science/tel-00713647>**

Submitted on 2 Jul 2012

**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.

N° D'ORDRE

\*\*\*\*\*

UNIVERSITE DE PARIS-SUD  
U.F.R. SCIENTIFIQUE D'ORSAY

THESE  
présentée  
pour obtenir le grade de

DOCTEUR EN SCIENCES  
DE L'UNIVERSITE PARIS 11, ORSAY  
Discipline : Sciences du végétal

par

Nicolas GOBRON

**CARACTERISATION GENETIQUE ET MOLECULAIRE  
D'UNE INCOMPATIBILITE NUCLEO-CYTOPLASMIQUE  
CHEZ *ARABIDOPSIS THALIANA***

*Soutenue le 27 Avril 2012*

Jury :

Joëlle Ronfort  
Pierre Saumitou-Laprade  
Matthieu Falque  
Jacqui Shykoff  
Françoise Budar

Rapporteur  
Rapporteur

Directrice de thèse

# REMERCIEMENTS

Je tiens avant tout à remercier Françoise Budar pour m'avoir accueilli dans son équipe « Organites et reproduction ». Merci pour ton entière disponibilité, pour ton encadrement et pour tout ce que tu m'as transmis durant ces années.

Je remercie également Hakim Mireau et les autres membres de l'équipe pour leur bonne humeur, leur soutien et leur disponibilité au quotidien.

Merci à tout ceux de la SGAP et notamment à Philippe Guerche pour m'avoir accueilli dans son laboratoire de l'INRA de Versailles, ainsi qu'à Olivier Loudet, pour ses conseils très précieux lors de ce travail.

J'exprime également toute ma gratitude à Joëlle Ronfort, Pierre Saumitou-Laprade, Matthieu Falque et Jacqui Shykoff pour avoir accepté de juger ce travail.

# ABREVIATIONS - UNITES - NOMENCLATURE

$^{\circ}\text{C}$ : degré Celsius	RIL : recombinant inbred line
A : Adénine	RM : Rétrocroisement Maternel
ADN : Acide désoxyribonucléique	ROS : Reactive Oxygen Species
ADNmt : Acide désoxyribonucléique mitochondrial	RuBisCO : Ribulose-1,5-diphosphate carboxylas oxygénase :
AOX : Alternative oxidase	s : secondes
ARN : acide ribonucléique	S : Stérile
ARNm : Acide Ribonucléique Messenger	SMNC : Stérilité Mâle Nucléo-Cytoplasmique
ARNr : Acide Ribonucléique Ribosomique	SNP : Single Nucleotide Polymorphism
ARNt : Acide Ribonucléique Transfert	SRC : Signalisation Rétrograde Chloroplastique
ATP : Adénosine-5'-TriPhosphate	SRM : Signalisation Rétrograde Mitochondriale
AV : Accession Versailles	T : Thymine
BDM : Bateson-Dobzhansky-Muller	U : Uracile
C : Cytosine	uL : microlitre
C <sub>6</sub> H <sub>12</sub> O <sub>6</sub> : glucose	uM : micromolaire
Ca <sup>2+</sup> : ion calcium	UTR : UnTranslated Region
CAPS : Cleaved Amplified Polymorphic Sequence	V : Volt
Ch : Chromosome	
CM : Cercle Maître	
CO <sub>2</sub> : dioxygène	
COX : Cytochrome c oxidase	
cRT-PCR : Circular Reverse Transcriptase-Polymerase Chain Reaction	
F : Fertile	
G : Guanine	
GFP : Green Fluorescent Protein	
GWAS : Genome-Wide Association Study	
H <sub>2</sub> O : eau	
H <sub>2</sub> O <sub>2</sub> eau oxygénée	
IDS : Inducteur de Stérilité	
Kb : kilo-base	
M : molaire	
Ma : millions d'années	
Mb : méga-base	
mM : millimolaire	
MSAT : Microsatellite	
NADP : Nicotinamide Adénine Dinucléotide Phosphate	
NADPH : Nicotinamide Adénine Dinucléotide Phosphate-oxidase	
NEP : Nuclear-encoded Polymerase	
NF : Norflurazon	
NIDS : Non Inducteur de Stérilité	
O <sub>2</sub> : oxygène	
O <sub>2</sub> <sup>-</sup> : ion superoxyde	
ORF : Open Reading Frame	
pb : paire de bases	
PC : Population de cartographie	
PCR : Polymerase Chain Reaction	
PEP : Plastid-encoded Polymerase	
pH : potentiel hydrogène	
PPR : Pentatricopeptide repeat	
PSI : Photosystem I	
PSII : Photosystem II	
Rf : Restaurateur de fertilité	
RFLP : Restriction Fragment Length Polymorphism :	

# SOMMAIRE

<b>SOMMAIRE</b> .....	4
<b>INTRODUCTION</b> .....	7
I. Les organites chloroplastiques et mitochondriaux de la cellule végétale.....	8
I.1 Fonctions biologiques des chloroplastes et des mitochondries.....	8
I.2 Origines de la cellule eucaryote végétale.....	10
I.3 Héritéité des chloroplastes et des mitochondries chez les angiospermes.....	10
II. Génome chloroplastique des plantes actuelles.....	12
II.1 Taille et contenu en gènes chloroplastiques.....	12
II.2 Structure et organisation du génome chloroplastique.....	15
II.3 Expression des gènes chloroplastiques.....	17
III. Génome mitochondrial des plantes actuelles.....	24
III.1 Taille et contenu en gènes mitochondriaux.....	24
III.2 Structure et organisation du génome mitochondrial.....	27
III.3 Expression des gènes mitochondriaux.....	31
IV. Interactions nucléo-cytoplasmiques dans la cellule végétale.....	40
IV.1 La famille des protéines à motifs PPR.....	40
IV.2 Coordination entre compartiments génétiques de la cellule végétale.....	44
IV.2.1 La signalisation antérograde.....	44
IV.2.2 La signalisation rétrograde.....	44
V. Coévolution nucléo-cytoplasmique.....	51
V.1 Origine évolutive des génomes cytoplasmiques et nucléaire.....	51
V.2 Co-adaptation des génomes cytoplasmiques et nucléaire.....	52
V.3 Modèle de Bateson-Dobzhansky-Muller (BDM) adapté aux interactions nucléocytoplasmiques.....	54
V.4 Exemples de rupture de co-adaptation nucléo-cytoplasmique chez les eucaryotes.....	56
VI. Conflits génomiques entre compartiments nucléaire et cytoplasmiques.....	61
VI.1 Les stérilités mâles nucléocytoplasmiques (SMNC) .....	61
VI.2 La SMNC à l'échelle de la population : la gynodioécie.....	63
VI.3 La SMNC à l'échelle de l'individu : la gamétogenèse mâle.....	66
VI.4 Aspects moléculaires des stérilités mâles nucléo-cytoplasmiques.....	68
VI.4.1 Facteurs mitochondriaux inducteurs de stérilité.....	68
VI.4.2 Facteurs nucléaires restaurateurs de fertilité.....	75
VII. La plante modèle <i>Arabidopsis thaliana</i> .....	79
VII.1 Présentation.....	79
VII.2 Histoire évolutive de <i>Arabidopsis thaliana</i> .....	82
VII.3 Diversité du génome nucléaire des accessions de <i>Arabidopsis thaliana</i> .....	83
VII.4 Diversité des génomes cytoplasmiques des accessions de <i>Arabidopsis thaliana</i> .....	85

<b>PRESENTATION ET OBJECTIFS DE LA THESE</b> .....	92
<b>RESULTATS</b> .....	93
I Cartographie génétique des mainteneurs de stérilité.....	96
I.1 Résultats préliminaires.....	99
I.2 Stratégie et plans de croisements pour la création de populations de cartographie.....	101
I.2.a Introgression des allèles Mr-0 mainteneurs de stérilité dans un fond Sha.....	101
I.2.b Fixation des allèles Mr-0 mainteneurs de stérilité.....	101
I.2.b.a Fixation des allèles Mr-0 au locus ch1.....	101
I.2.b.b Fixation des allèles Mr-0 au locus ch3.....	103
I.2.c Production de populations de cartographies.....	105
I.3 Cartographie génétique du locus mainteneur de stérilité au chromosome 1.....	108
I.3.a Analyse des populations de cartographie.....	108
I.3.b Analyse des recombinants au locus ch1.....	110
I.3.c Analyse génétique des distortions de ségrégation au chromosome 1.....	116
I.4 Cartographie génétique du locus mainteneur de stérilité au chromosome 3.....	119
I.4.a Analyse des populations de cartographie.....	119
I.4.b Analyse des recombinants au locus ch3.....	124
I.4.c Analyse moléculaire de l'intervalle contenant le locus ch3 chez Sha et Mr-0.....	128
II Caractérisation phénotypique de la production de pollen dans certaines combinaisons génétiques.....	130
II.1 Caractérisation phénotypique de la stérilité mâle nucléo-cytoplasmique Sha/Mr-0.....	130
II.1.a Observations de la viabilité des grains de pollen.....	130
II.1.b Observations des anthères en développement.....	132
II.2 Caractérisation phénotypique des distortions de ségrégation Sha/Mr-0.....	135
II.2.a Distorsions de ségrégation au chromosome 1.....	136
II.2.b Distorsions de ségrégation au chromosome 3.....	137
III Répartition des facteurs de la stérilité mâle nucléo-cytoplasmique Sha x Mr-0 au sein de la diversité naturelle de l'espèce.....	137
III.1 Recherche d'accessions mainteneuses de la stérilité mâle nucléo-cytoplasmique induite par le cytoplasme de Sha.....	140
III.2 Recherche d'accessions dont le cytoplasme est inducteur d'une stérilité mâle en présence des allèles de Mr-0.....	143
IV Recherche du compartiment cytoplasmique responsable de l'induction de stérilité.....	145
IV.1 Analyse du chlorotype des accessions utilisées en croisement avec Mr-0.....	145
IV.2 Identification de marqueurs mitochondriaux spécifiques des cytoplasmes IDS.....	145
IV.3 Recherche d'un gène mitochondrial candidat facteur d'IDS.....	153
IV.4 Influence des allèles mainteneurs Mr-0 sur l'expression des gènes mitochondriaux de Sha.....	155

<b>DISCUSSION</b> .....	158
Bilan critique des résultats.....	159
Le déterminant cytoplasmique de stérilité mâle.....	159
Les locus nucléaires mainteneurs de stérilité mâle.....	163
Le locus mainteneur chr3.....	163
Le locus mainteneur ch1.....	165
Les locus distorateurs de ségrégation.....	166
Les observations cytologiques.....	167
Quelle est la signification évolutive de la SMNC Sha x Mr-0 ? .....	169
Les barrières génétiques dans le croisement Sha x Mr-0 .....	174
Incompatibilités alléliques inter-locus dans le croisement Sha x Mr-0.....	174
Incompatibilités alléliques intra-locus dans le croisement Sha x Mr-0.....	176
Les barrière(s) génétique(s) dans le croisement Sha x Etna-1.....	178
<b>MATERIELS ET METHODES</b> .....	181
MATERIEL.....	182
I. Plantes.....	182
II. Oligonucléotides.....	182
METHODES.....	183
I. Culture des plantes et croisements.....	183
II Biologie moléculaire.....	183
II.1 Extraction d'ADN en tubes.....	183
II.2 Extraction d'ADN en plaques.....	184
II.3 Extraction d'ARN en tubes.....	185
II.4 Définition des amorces.....	186
II.5 Amplification par PCR (Polymerase Chain Reaction) .....	186
II.6 Circularisation des ARN et préparation d'ADNc pour les cRT-PCR.....	188
II.7 Électrophorèse.....	189
II.8 Séquençage d'acides nucléiques.....	190
III. Microscopie.....	190
III.1 Coloration d'Alexander.....	191
III.2 Coloration à l'iodure de propidium.....	191
<b>BIBLIOGRAPHIE</b> .....	198
<b>ANNEXES</b> .....	219
Annexe 1.....	220
Annexe 2.....	226
Annexe 3.....	231

# **INTRODUCTION**

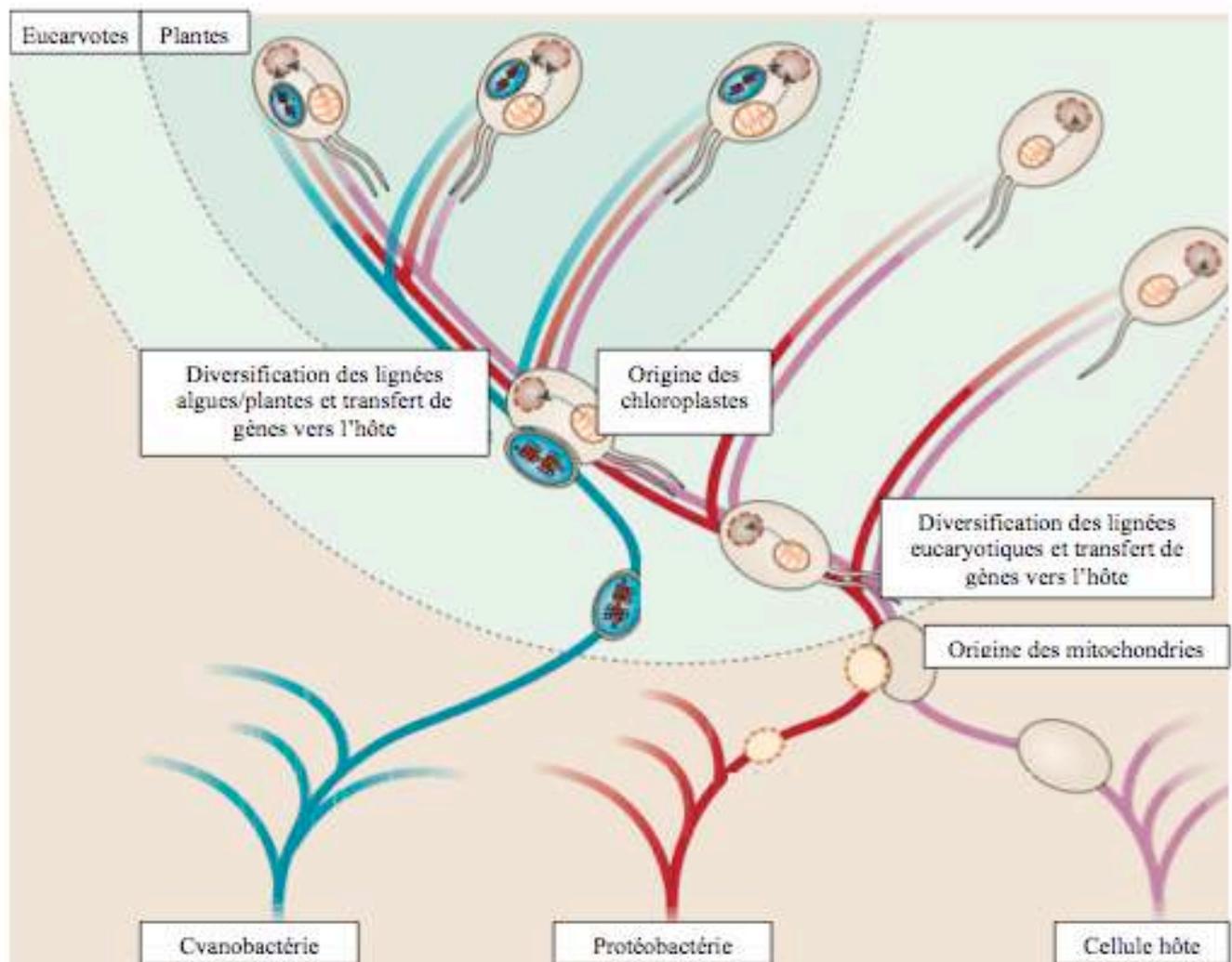
## I. Les organites chloroplastiques et mitochondriaux de la cellule végétale

### I.1 Fonctions biologiques des chloroplastes et des mitochondries

Les organites sont des structures spécialisées présentes dans le cytoplasme des cellules eucaryotes. Ils sont séparés de l'environnement cellulaire par au moins une membrane ; la cellule eucaryote est donc compartimentée. Les chloroplastes et les mitochondries, présents dans le cytoplasme des cellules eucaryotes végétales, sont des organites qui possèdent leurs propres génomes. Leurs fonctions sont vitales pour la cellule.

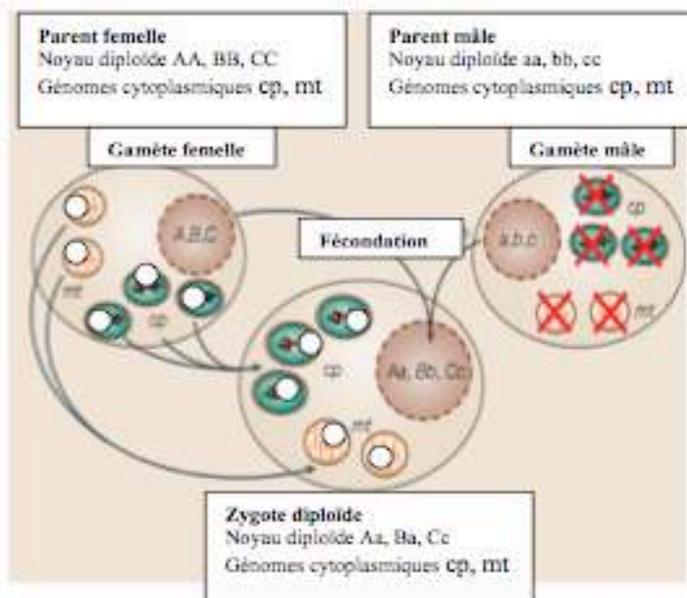
Les chloroplastes assurent la photosynthèse : c'est un processus bioénergétique, composé de deux phases, permettant aux plantes de synthétiser de la matière carbonée organique ( $C_6H_{12}O_6$ ) et du dioxygène ( $O_2$ ) à partir d'eau ( $H_2O$ ), de dioxyde de carbone ( $CO_2$ ) et d'énergie solaire. La photosynthèse se déroule en deux phases distinctes : la première phase, aussi appelée phase des réactions photochimiques, permet à partir de l'énergie solaire, d' $H_2O$  et de NADP, de produire de l' $O_2$ , du NADPH et de l'ATP. Les réactions de la première phase nécessitent des complexes situés dans la membrane des thylacoïdes : le photosystème II (PSII), la plastoquinone, le cytochrome b6f, le photosystème I (PSI), un complexe ATP synthase ainsi que des molécules comme la plastocyanine. La seconde phase, ou phase de fixation du carbone, est indépendante de la lumière et est assurée dans le stroma par le cycle de Calvin. L'ATP et le NADPH, synthétisés lors de la première phase, vont être utilisés lors de la seconde phase pour fixer le  $CO_2$  et générer du glucose  $C_6H_{12}O_6$  exporté vers le cytosol ou stocké dans le chloroplaste sous forme d'amidon. L'enzyme permettant de fixer le  $CO_2$  est la ribulose 1,5 bisphosphate carboxylase/oxygénase (Rubisco). La photosynthèse constitue l'activité principale des chloroplastes. Néanmoins, il existe également un métabolisme non photosynthétique du chloroplaste : certaines réactions impliquées dans l'assimilation de l'azote, du soufre et dans la biosynthèse d'acides gras et d'acides aminés sont localisées dans les chloroplastes.

Les mitochondries sont le siège de la respiration cellulaire. Le rôle des mitochondries dans la synthèse d'ATP grâce au cycle de Krebs et à la chaîne de phosphorylation oxydative est bien établi (Saraste, 1999). La chaîne de phosphorylation oxydative est composée de complexes protéiques ancrés dans la membrane interne mitochondriale. Ces complexes produisent un gradient transmembranaire de protons exploité par l'ATP synthase pour produire de l'ATP. Les mitochondries jouent un rôle central dans de nombreux autres processus métaboliques dont la biosynthèse de précurseurs d'acides aminés, de nucléotides, de cofacteurs vitaminiques et de groupements fer-soufre (Mackenzie et McIntosh, 1999 ; Bowsher et Tobin, 2001). Les



**Figure A1 : origines de la cellule eucaryote**

Les mitochondries ont pour ancêtre une alpha protéobactérie et les chloroplastes, une cyanobactérie. Le schéma est un diagramme de l'évolution des eucaryotes, présentant l'incorporation des mitochondries et des chloroplastes dans la lignée eucaryotique par endosymbioses successives, puis la co-évolution des génomes nucléaires et des génomes des organites lors de laquelle a lieu le transfert de gènes vers le génome de l'hôte. Cet hôte semble avoir été un être unicellulaire sans organite. La nature exacte de cette cellule hôte est discutée par les évolutionnistes (d'après Timmis et al., 2004).



**Figure A2 : hérédité maternelle des organites**

Les génomes nucléaires des organismes eucaryotiques sont très souvent diploïdes (comme celui d'*Arabidopsis thaliana*). Un gène nucléaire va donc avoir deux allèles qui sont potentiellement différents. Les gamètes produits lors de la méiose sont haploïdes et portent seulement un allèle (ici, les allèles A, B et C proviennent du parent femelle et les allèles a, b et c du parent mâle). Le zygote obtenu après la fécondation hérite d'un allèle nucléaire de chaque gène provenant de chaque parent. Le zygote est ainsi diploïde et hétérozygote Aa, Bb, Cc. Au contraire, les génomes des organites sont hérités d'un seul parent. Ici est représenté une hérédité des organites qui est maternelle : les organites présents dans la descendance proviennent du parent femelle. Dans le schéma, les croix rouges représentent la dégradation des organites du parent mâle avant la fécondation et les ronds blancs, la transmission maternelle des organites (d'après Timmis et al., 2004).

mitochondries régulent le métabolisme cellulaire en contrôlant le flux de facteurs, comme les ions, les acides aminés, les carbohydrates, entre elles et le cytosol. Les mitochondries sont également impliquées dans la signalisation cellulaire, en régulant l'homéostasie cellulaire du calcium (Vandecasteele et al., 2001 ; Logan et Knight, 2003), et dans la mort cellulaire programmée (Jones, 2000 ; Youle et Karbowski, 2005). Intéressons-nous maintenant à l'origine des chloroplastes et des mitochondries de la cellule végétale.

## I.2 Origines de la cellule eucaryote végétale

Au moins deux événements distincts d'endosymbiose sont à l'origine de la cellule eucaryote végétale que nous connaissons actuellement (Taylor, 1979 et Margulis, 1993). La symbiose est définie comme une relation bénéfique mutuelle qui implique au moins deux partenaires biologiques. L'endosymbiose fait référence au cas particulier où l'un des partenaires, l'endosymbiote, est présent dans le cytoplasme de l'autre partenaire (Dyall et al., 2004). L'évolution des cellules eucaryotes est schématisée dans la figure A1. En 1885, Schimper suggère que les chloroplastes pourraient provenir de microorganismes symbiotiques (cité dans Palenik, 2002) et Mereschkowsky émet l'hypothèse, en 1905, que les chloroplastes dérivent de cyanobactéries, un groupe de bactéries photosynthétiques ayant de nombreuses similarités avec les chloroplastes actuels (Mereschkowsky, 1905). Les mitochondries sont des organites plus petits que les chloroplastes et ne sont pas colorées naturellement, contrairement aux chloroplastes. La découverte des mitochondries s'est donc faite plus tardivement que les chloroplastes et c'est Wallin qui, en 1927, propose que les mitochondries ont une origine bactérienne (Wallin, 1927). Aujourd'hui, des analyses phylogénétiques, structurales et biochimiques confirment que les mitochondries dérivent de formes d'alpha-protéobactéries captées par la cellule, nucléée (eucaryote) ou de type archée, il y a au moins 1,5 milliard d'années. Une cyanobactérie, ancêtre du chloroplaste actuel, se serait ensuite ajoutée, il y a entre 1,2 et 1,5 milliards d'années, aux mitochondries déjà présentes dans le cytoplasme d'une cellule eucaryote (pour revue, Timmis et al., 2004).

## I.3 Hérité des chloroplastes et des mitochondries chez les angiospermes

Peu de temps après la redécouverte des lois de Mendel en 1900, Correns et Baur découvrent simultanément l'existence d'une hérédité non-Mendélienne chez des espèces appartenant à la famille des angiospermes. Correns observa chez *Mirabilis jalapa* que, suite à des croisements

entre des plantes (ou des branches) ayant des feuilles de couleur jaune ou verte, le caractère « couleur des feuilles » était hérité seulement du parent maternel (hérédité maternelle). Les branches vertes produisent toujours des graines qui donnent des plantes vertes et les branches jaunes produisent toujours des graines qui donnent des plantes jaunes. Le parent qui donne le pollen n'a pas d'influence sur le caractère « couleur des feuilles » de la descendance (Correns, 1909). Baur identifia un autre type d'hérédité non mendélienne : en réalisant des croisements réciproques avec des plantes de *Pelargonium zonale* ayant des feuilles de couleurs différentes, il mit en évidence que le caractère « couleur des feuilles » était hérité des deux parents (hérédité biparentale) de façon non mendélienne (Baur, 1909). La découverte de l'existence des chloroplastes puis l'identification de la présence d'ADN dans les chloroplastes ont pu matérialiser cette hérédité non mendélienne. Pour *Mirabilis jalapa* et *Pelargonium zonale*, les chloroplastes, qui portent les traits associés à la couleur des feuilles, sont hérités de façon maternelle et biparentale respectivement (revue de Hagemann, 2000). Chez *Arabidopsis thaliana*, comme chez la grande majorité des végétaux supérieurs, la transmission des chloroplastes à la descendance est maternelle. À noter qu'une très faible fréquence de chloroplastes d'origine paternelle peut être transmis à la descendance chez des espèces comme *Arabidopsis thaliana* (Azhagiri et Maliga, 2007) et *Nicotiana tabacum* (Horlow et al., 1990). Néanmoins, ces événements sont extrêmement rares et ont été observés uniquement en conditions de laboratoire (Azhagiri et Maliga, 2007). Plusieurs mécanismes existent pour contrôler le mode de transmission des chloroplastes chez les angiospermes. Avant la fécondation, des événements cellulaires peuvent bloquer ou faciliter la transmission des chloroplastes (Mogensen, 1996). Du côté mâle, une distribution extrêmement polarisée des chloroplastes dans les microspores lors de la première division de méiose induit une absence de chloroplastes dans les cellules génératives ; la non transmission peut également se faire par une dégénérescence ou une exclusion des chloroplastes dans les cellules génératives et/ou végétatives. Ces événements cellulaires bloquent la transmission mâle des chloroplastes chez les végétaux à hérédité maternelle alors que dans le cas des chloroplastes hérités de façon biparentale, ce sont des événements cellulaires contraires qui facilitent la transmission des chloroplastes d'origine paternelle (Hu et al., 2008). Chez certaines espèces, l'hérédité des chloroplastes peut être affectée par des événements cellulaires dans le zygote. Chez *Medicago sativa*, la transmission des chloroplastes paternels dépend du génotype du parent femelle (Smith, 1989) qui influence la localisation des chloroplastes soit dans la portion apicale de la cellule oeuf, qui forme l'embryon, soit dans la portion basale qui forme le suspenseur. Seuls les chloroplastes présents dans la portion apicale seront transmis (Rusche et al., 1995).

Concernant les mitochondries des angiospermes, leurs génomes sont transmis par le parent femelle, à l'exception de certains cas, comme chez *Cucumis melo* où la transmission est paternelle (Mogensen, 1996). Dans le cas général où l'hérédité des mitochondries est maternelle, les événements cellulaires qui empêchent la transmission de mitochondries d'origine paternelle ne sont pas les mêmes que dans le cas de l'hérédité des chloroplastes. La distribution des mitochondries dans la microspore est non polarisée. Contrairement à la distribution des chloroplastes, la cellule générative (qui est le précurseur des cellules spermatiques) reçoit des mitochondries. Ces mitochondries restent structurellement intactes dans les cellules génératives (Sodmergen et al., 2002). La dégradation de l'ADN mitochondrial dans les cellules des gamètes mâles a été observée par microscopie chez des angiospermes comme *Medicago truncatula* (Matsushima et al., 2008). Cette dégradation n'est pas observée chez les angiospermes qui ont une hérédité paternelle ou biparentale (Nagata et al., 1999). Chez *Arabidopsis thaliana*, les mitochondries et les chloroplastes sont transmis à la descendance par le parent femelle. Cette transmission maternelle des organites est schématisée dans la figure A2. Nous allons maintenant nous intéresser en détail aux génomes actuels des chloroplastes et des mitochondries.

## II. Génome chloroplastique des plantes actuelles

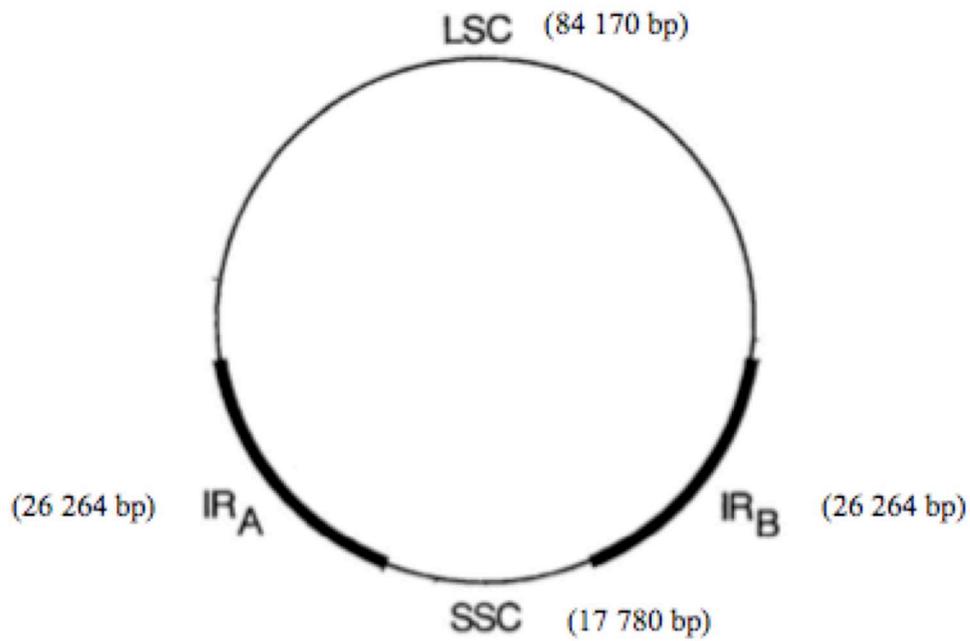
### II.1 Taille et contenu en gènes chloroplastiques

Les premiers génomes de plantes à avoir été séquencés sont ceux de génomes chloroplastiques appartenant à *Marchantia polymorpha* et *Nicotiana tabacum*, (Ohyama et al., 1986 ; Shinozaki et al., 1986). En juillet 2011, 213 génomes chloroplastiques étaient séquencés et disponibles dans les bases de données ; Parmi eux sont disponibles 182 génomes de végétaux appartenant aux *Viridiplantae* qui regroupent notamment les végétaux terrestres et les algues vertes. Chez les *Viridiplantae*, la taille des génomes varie de 38 kb chez *Helicosporidium sp.*, un parasite non photosynthétique détecté dans divers groupes d'arthropodes et d'algues (Tartar et Boucias, 2004 ; Bläske-Lietze et Boucias, 2005 ; de Koning et Keeling, 2006), à 521 kb chez *Floydiella terrestris*, une algue verte (Brouard et al., 2010). Au sein des *Spermatophyta*, ou plantes à graines, 139 génomes chloroplastiques ont été séquencés dont 14 pour des plantes appartenant aux *Coniferophyta* (ou conifères) et 121 pour des plantes appartenant aux *Magnoliophyta* (ou plantes à fleurs). Parmi les génomes chloroplastiques de conifères séquencés, on peut citer ceux de *Pinus contorta* et *Pinus thunbergii* qui font 120 kb et 119kb respectivement. Au sein des plantes à fleurs, le génome

chloroplastique fait 134 kb chez *Oriza sativa*, 140 kb chez *Zea mays*, 150 kb chez *Helianthus annuus* et *Phaseolus vulgaris*, et 153 kb chez *Brassica rapa*. Le seul génome chloroplastique d'*Arabidopsis thaliana* disponible a été séquencé en 1999 sur l'accession Columbia (Col-0) et a une taille de 154 kb (Sato et al., 1999). D'après la classification phylogénétique, les organismes les plus proches d'*Arabidopsis* et dont les génomes chloroplastiques ont été séquencés, sont *Capsella bursa-pastoris* (154 kb), *Crucihimalaya wallichii* (155 kb), et *Olimarabidopsis pumila* (154 kb) : les tailles de ces génomes sont presque identiques à celle du génome d'*Arabidopsis thaliana*. En règle générale, la taille des génomes chloroplastiques varie peu d'une plante terrestre à l'autre.

Chez les *Viridiplantae*, le génome chloroplastique code en moyenne 85 protéines et 43 ARNt et ARNr. Chez *Arabidopsis thaliana*, 129 gènes codent 85 protéines, 37 ARNt et 7 ARNr alors que 130 gènes codent 85 protéines, 37 ARNt et 8 ARNr chez *Capsella bursa-pastoris*. En règle générale, le contenu en gènes est bien conservé d'une espèce végétale à l'autre. Néanmoins, il existe des exceptions à cette conservation : le génome chloroplastique de *Rhizanthella gardneri*, une orchidée souterraine non photosynthétique d'Australie, code seulement 20 protéines, 4 ARNr et 9 ARNt (Delannoy et al., 2011). Au contraire, le génome du pin de Corée, *Pinus koraiensis*, code 273 protéines et 40 ARNr et ARNt (Noh et al., soumission directe à GenBank). Les gènes supplémentaires prédits chez *Pinus koraiensis* et chez d'autres plantes sont en fait des cadres ouverts de lecture, sans fonction connue, et appelés *ycf* (pour *hYpothetical Chloroplast open reading Frames*) (Hallick and Bairoch 1994, Maier et al. 1995). Chez les végétaux supérieurs, les gènes chloroplastiques peuvent être classés en trois catégories. La première regroupe ceux qui codent des protéines de l'appareil photosynthétique : le PSI, le PSII, le cytochrome b6f, l'ATP synthase, la grande sous-unité de la Rubisco et des protéines du complexe NAD(P)H déshydrogénase. La seconde catégorie de gènes chloroplastiques code des ARN et des protéines impliquées dans l'expression des gènes chloroplastiques : ARNt, ARNr, protéines de l'ARN polymérase, protéines ribosomiques et d'autres protéines comme la maturase K, qui intervient dans l'épissage des introns chloroplastiques. La troisième catégorie contient des gènes qui codent des protéines n'appartenant pas aux deux premières catégories : c'est le cas de *CemA*, une protéine de fonction inconnue présente dans la membrane interne du chloroplaste (Sasaki et al., 1993).

À quelques exceptions près, les génomes chloroplastiques présentent donc une taille et un contenu en gènes relativement constants d'une espèce à l'autre. Néanmoins, comme nous



**Figure A3 : structure du génome chloroplastique d'*Arabidopsis thaliana***

Elle consiste en une grande région simple copie (LSC) et une petite région simple copie (SSC). Les deux séquences inversées et répétées (IRA et IRB) sont représentées en lignes plus épaisses. Les tailles des régions et des séquences répétées sont indiquées entre parenthèses. Cette structure est valable pour les génomes chloroplastiques des autres organismes possédant des chloroplastes (d'après Sato et al., 1999).

l'avons vu, certaines espèces ont des génomes chloroplastiques de taille très petite ou très grande. Une observation importante est que la taille du génome chloroplastique n'est pas corrélée au nombre de gènes qu'il contient : en effet, malgré le très faible nombre de gènes chloroplastiques présents chez *Rhizanthella gardneri*, son génome de 59,2 kb est plus grand que celui de *Helicosporidium sp.*, alors que ce dernier code 26 protéines et 28 ARNr et ARNt. (Delannoy et al., 2011). Pour comprendre ces variations de taille, il faut savoir que la molécule d'ADN chloroplastique, généralement représentée sous forme circulaire, contient deux copies inversées d'une séquence répétée qui séparent une grande et une petite région présentes en simple copie. Les causes principales de la variation de taille des génomes chloroplastiques entre différentes espèces sont la perte et le transfert de gènes vers le noyau, mais aussi la variation de taille de ces séquences répétées. Les gènes situés dans les séquences répétées sont présents en duplicats (une copie par répétition). Par exemple, la taille des séquences répétées inversées de l'ADN chloroplastique de *Pinus thunbergii* est extrêmement petite (495 pb) et celui-ci possède un petit génome chloroplastique de 119 kb. De plus, ce génome a perdu les gènes codants le complexe NADH déshydrogénase. A l'autre extrême existe le génome chloroplastique de *Pelargonium hortorum*, qui possède des séquences répétées inversées d'environ 76 kb pour une taille globale de 217 kb (Palmer et al., 1987 ; Chumley et al., 2006). Le petit génome chloroplastique de *Rhizanthella gardneri* contient deux séquences répétées de 9767 bp (Delannoy et al., 2011). Le génome chloroplastique de l'accession Columbia d'*Arabidopsis thaliana* contient une paire de séquences répétées de 26264 bp, séparées par une petite et une grande régions de 17780 bp et 84170 bp respectivement (Sato et al., 1999). La structure du génome chloroplastique d'*Arabidopsis thaliana* est représentée dans la figure A3.

## II.2 Structure et organisation du génome chloroplastique

Au début des années 70, Kolodner et Tewar observent, par microscopie électronique, des molécules d'ADN circulaires en grande proportion dans des échantillons de chloroplastes. D'autres molécules d'ADN, linéaires cette fois, sont également observées, mais sont imputées à la fragmentation des molécules circulaires survenues lors des traitements d'extraction de l'ADN chloroplastique (Kolodner et Tewar, 1972). Ces molécules linéaires vont être considérées comme des artefacts durant des dizaines d'années jusqu'à l'arrivée de nouveaux traitements plus doux utilisés sur les échantillons pour observer la structure des génomes chloroplastiques. Progressivement, les molécules linéaires ne vont plus être considérées

comme des artéfacts mais bel et bien comme des structures d'ADN présentes dans les chloroplastes. Aujourd'hui, une proportion seulement de l'ADN chloroplastique total est considérée comme circulaire, les autres formes étant linéaires, ramifiées ou plus complexes (Bendich, 2004).

Outre la présence de différentes formes moléculaires d'ADN chloroplastique, on sait aussi qu'une cellule peut contenir plusieurs chloroplastes, que chaque chloroplaste peut contenir plusieurs copies d'ADN chloroplastique et que la quantité de cet ADN peut changer lors du développement des tissus. Des recherches menées chez *Arabidopsis thaliana* ont montré que dans les feuilles matures, l'ADN chloroplastique, visualisé par de la microscopie à fluorescence, est absent de la plupart des chloroplastes avant même la sénescence des tissus foliaires. La quantité d'ADN chloroplastique par chloroplaste diminue au cours du développement de la feuille (Rowan et al., 2004). La même équipe a examiné et comparé l'ADN chloroplastique de feuilles extrait à différents stades de développement chez *Nicotiana tabacum*, *Medicago truncatula* et *Pisum sativum* ; d'après les résultats, au fur et à mesure du vieillissement des tissus foliaires, la quantité d'ADN chloroplastique décline et le pourcentage de chloroplastes sans ADN détectable augmente chez *Medicago truncatula* et *Pisum sativum*, alors que la majorité des chloroplastes présents dans les feuilles matures de *Nicotiana tabacum* conservent leur ADN chloroplastique (Shaver et al., 2006).

L'équipe de Bendich a également comparé la structure moléculaire et la quantité d'ADN chloroplastique dans différents tissus de *Zea mays* et à différents stades de développement. D'après eux, la structure et la quantité d'ADN chloroplastique peuvent varier en fonction des tissus et des stades de développement des chloroplastes (Oldenburg et Bendich, 2004). L'hypothèse pour expliquer la variation de la quantité d'ADN chloroplastique au cours du développement des feuilles est la suivante : la réplication de l'ADN chloroplastique a lieu dans les cellules méristématiques (Kuroiwa, 1991 ; Fujie et al., 1994) ; lors du développement de la feuille, l'augmentation du nombre de chloroplastes entraînerait une diminution de la quantité d'ADN chloroplastique par chloroplaste, comme c'est le cas chez le pois (Lamppa et al. 1980). Une autre hypothèse serait que l'ADN chloroplastique soit dégradé comme c'est le cas chez le riz (Sodmergen et al., 1991).

Aujourd'hui, il est admis que l'ADN chloroplastique a une structure complexe et multigénomique dont la quantité et la structure varient selon les tissus et les stades de développement des chloroplastes. Néanmoins, l'organisation et la structure des génomes

chloroplastiques restent généralement stables et conservées entre espèces végétales. Ce qui ne veut pas dire qu'il n'y a jamais de réarrangements dans ces génomes. En effet, des inversions (Kim et al. 2005), des translocations et des insertions/délétions (Chumley et al., 2006) ont été identifiées dans certains génomes chloroplastiques. Par exemple, plusieurs membres de la famille des *Poaceae* sont connus pour posséder trois inversions d'ADN chloroplastique par rapport à l'ADN chloroplastique de la plupart des plantes à fleurs (Doyle et al., 1992).

### II.3 Expression des gènes chloroplastiques

Les gènes chloroplastiques et leurs séquences régulatrices peuvent être organisés en opérons. Cette organisation est caractéristique des génomes des procaryotes (Jacob et Monod, 1961). Toutefois, les transcrits chloroplastiques polycistroniques (ou précurseurs) doivent être maturés pour devenir fonctionnels (Stern et al., 2010). Par exemple, l'opéron *rrn* chloroplastique code les quatre molécules d'ARNr (*rrn16*, *rrn23*, *rrn4.5* et *rrn5*) et deux ARNt (*trnI* et *trnA*). Le transcrit précurseur est maturé par un ensemble d'évènements de maturation post-transcriptionnelle pour finalement aboutir à l'obtention d'ARN matures (Baumgartner et al., 1993; Chun et al., 2001). À l'heure actuelle, les études disponibles semblent indiquer que le contrôle de l'expression des gènes chloroplastiques se situe principalement au niveau transcriptionnel et post-transcriptionnel chez les végétaux supérieurs. Les contrôles traductionnel et post-traductionnel de l'expression des gènes des organites ne seront donc pas détaillés dans cette thèse. L'expression des gènes chloroplastiques est influencée par des signaux environnementaux comme la lumière, le rythme circadien et différents stress, mais également par l'étape de développement du chloroplaste et du tissu dans lequel il se trouve.

Différentes études montrent que l'expression des gènes chloroplastiques peut être contrôlée à l'étape de la transcription. Par exemple, la transcription des gènes chloroplastiques liés à la photosynthèse est très faible, voire inexistante, dans les étapes initiales du développement (lorsque le chloroplaste n'est encore qu'un proplastide présent dans la cellule méristématique). Puis la lumière induit une augmentation générale de l'activité transcriptionnelle des gènes liés à la photosynthèse, rendant les chloroplastes compétents à l'activité photosynthétique (Baumgartner et al., 1993; Bisanz-Seyer et al., 1989, Mullet, 1993). Une régulation tissu spécifique existe également : la plupart des gènes liés à la photosynthèse ne sont pas transcrits dans les tissus non photosynthétiques comme les racines par exemple (Isono et al., 1997). L'activité transcriptionnelle des gènes chloroplastiques dépend également de l'état

physiologique du chloroplaste. Par exemple, l'activité transcriptionnelle du gène *psaA* est contrôlée de façon réversible par l'état redox du chloroplaste (Pfannschmidt et al., 1999; Link, 2003).

Les chloroplastes des végétaux supérieurs utilisent au moins deux ARN polymérases, appelées NEP (pour Nuclear-encoded Polymerase) et PEP (pour Plastid-encoded Polymerase) qui sont globalement responsables de la transcription des gènes de ménage (ou gènes constitutifs) et des gènes reliés à l'activité photosynthétique respectivement. Ces deux enzymes sont non seulement responsables de la transcription de lots distincts de gènes lors du développement du chloroplaste mais sont aussi actives à différentes étapes du développement chloroplastique. NEP assure la majorité de l'activité transcriptionnelle dans les premières étapes du développement du chloroplaste puis cette activité est majoritairement prise en charge par PEP lorsque le chloroplaste se développe pour permettre la transcription des gènes reliés à la photosynthèse (Liere et Maliga, 2001). Chaque ARN polymérase se fixe sur des séquences promotrices particulières : NEP reconnaît des promoteurs de type phagique T3/T7 et PEP reconnaît des régions promotrices dont les séquences consensus sont du type sigma 70 d' *Escherichia coli*. Certains gènes chloroplastiques transcrits tout au long du développement de la plante ont des séquences promotrices qui peuvent être reconnues par les deux polymérases, ce qui permet une expression constitutive de ces gènes (Swiatecka-Hagenbruch et al., 2007).

L'ARN polymérase NEP ressemble à l'ARN polymérase monomérique du bactériophage T3/T7 (Hess and Börner, 1999). NEP est une enzyme codée par le génome nucléaire (Hajdukiewicz et al., 1997) puis importée dans les chloroplastes. Deux types de NEP assurent la transcription des gènes chloroplastiques : le produit du gène nucléaire *RpoTp* est exclusivement importé dans les chloroplastes alors que *RpoTmp* code une autre polymérase NEP adressée vers les chloroplastes, mais aussi vers les mitochondries. Ces deux gènes ont été clonés chez *Arabidopsis thaliana* (Hedtke et al., 2000; Hedtke et al., 2002). Actuellement, les analyses phylogénétiques des ARN polymérases de type phagique suggèrent que les polymérases NEP ont été produites par duplication du gène codant l'ARN polymérase mitochondriale après la séparation des angiospermes et des gymnospermes (Kabeya et al., 2002,2003). Ces duplications auraient ensuite acquis les séquences d'adressage vers les chloroplastes (Hedtke et al., 1997).

L'ARN polymérase PEP est une enzyme homologue de l'ARN polymérase d'*Escherichia coli* qui est composée de deux éléments principaux : le cœur de l'enzyme, catalysant l'activité de

synthèse de l'ARN, et le facteur sigma échangeable qui est principalement responsable de la reconnaissance du promoteur (Ishihama, 2000). Chez les végétaux photosynthétiques, le génome chloroplastique contient 4 gènes *rpo* (*rpoA*, *rpoB*, *rpoC1* et *rpoC2*) qui codent les sous unités protéiques du cœur de la polymérase PEP (Lopez-Juez, 2007). Les gènes codant les facteurs sigma ne se trouvent pas dans le génome chloroplastique mais sont localisés dans le génome nucléaire. Les facteurs sigma sont nécessaires à la polymérase pour reconnaître les promoteurs de type bactérien et pour permettre une initiation correcte de la transcription des gènes chloroplastiques (Suzuki et al., 2004). Le génome nucléaire d'*Arabidopsis thaliana* contient 6 gènes (*AtSIG1* à *AtSIG6*) codant des facteurs sigma dont les séquences N-terminales possèdent des extensions riches en sérine et en thréonine caractéristiques d'un signal d'adressage vers les chloroplastes (Hakimi et al., 2000; Fujiwara et al., 2000). Certains facteurs sigma sont responsables de la reconnaissance de promoteurs spécifiques par PEP (Ishizaki et al., 2005) en fonction du développement chloroplastique et de l'environnement. Il existe une demi-douzaine de régions promotrices différentes reconnues par l'ARN polymérase PEP, en fonction du facteur sigma associé, dans le génome chloroplastique (Allison, 2000).

L'expression des gènes chloroplastiques peut aussi nécessiter la fixation de facteurs de transcription, codés par le noyau et adressés à des séquences promotrices spécifiques dans le génome chloroplastique. La séquence comprise entre -16 et -102 nucléotides en amont du site d'initiation de la transcription du gène *rbcL* (gène qui code la grande sous-unité de la Rubisco) est un site de fixation pour un facteur de transcription chloroplastique (CDF1) codé par le génome nucléaire (Shiina et al., 1998). Chez *Arabidopsis thaliana*, ce facteur de liaison à l'ADN chloroplastique semble jouer un rôle important dans la spécificité tissulaire de la transcription du gène *rbcL*, mais également vis-à-vis de la dépendance à la lumière qui caractérise l'activité du promoteur de ce gène (Chun et al., 2001).

Suite à la transcription, différents processus post-transcriptionnels interviennent pour maturer les transcrits. Parmi eux existent l'édition, l'épissage des introns et la maturation des extrémités 5' et 3'.

L'édition est un processus de conversion post-transcriptionnelle de résidus cytidine (C) en résidus uridine (U), nécessaire à l'expression de gènes des organites en « corrigeant » le transcrit codé par le génome (Takenaka et al., 2008). La majorité des modifications post-transcriptionnelles qui sont liées à l'édition chloroplastique sont des changements de résidu C->U; les changements U->C existent, mais n'ont été observés que chez peu d'espèces et

concernent peu de sites (Yoshinaga et al., 1996 ; Kugita et al., 2003 ; Wolf et al., 2004). L'édition a été découverte pour la première fois en 1986 dans les transcrits *cox2* de la mitochondrie des trypanosomes (un groupe de protistes parasites) où quatre résidus U permettent de rétablir un cadre de lecture qui était absent du gène (Benne et al., 1986). L'édition des ARN chloroplastiques a été décrite pour la première fois en 1991 chez *Zea mays* où un codon d'initiation AUG est restauré grâce à l'édition C->U du codon ACG situé à l'extrémité 5' du transcrit de *rpl2* (Hoch et al., 1991). A noter que des éléments de séquence en amont du site édité peuvent être nécessaires pour l'édition de l'ARN (Schmitz-Linneweber et Barkan, 2007). Le génome chloroplastique d'*Arabidopsis thaliana* Col-0 contient 34 sites d'édition (Chateigner-Boutin et Small, 2007). Parmi les 11 gènes chloroplastiques *ndh* (qui codent les sous unités du complexe NAD(P)H déshydrogénase) d'*Arabidopsis thaliana*, *ndhA*, *ndhB*, *ndhD* et *ndhF* contiennent respectivement un, huit, quatre et un sites d'édition (Lutz et Maliga, 2001). Chez *Arabidopsis*, la protéine PPR CRR4, codée par le noyau, est nécessaire à l'édition d'un site spécifique dans les chloroplastes. En effet, le mutant *crr4* est spécifiquement altéré dans un site d'édition qui, chez le sauvage, crée un codon d'initiation dans le transcrit *ndhD*. Le phénotype du mutant *crr4* est une baisse de la quantité de protéines du complexe NA(P)DH déshydrogénase chloroplastique (Kotera et al., 2005).

L'épissage des ARN est une étape de maturation lors de laquelle les introns vont être excisés de l'ARN et les exons « raboutés » entre eux pour donner un ARN mature. Les introns présents dans les gènes chloroplastiques sont classés en deux familles : les introns du groupe I et ceux du groupe II. Cette classification se fait en fonction de la structure primaire et secondaire de l'intron, mais aussi en fonction du mécanisme d'épissage. Les introns du groupe I sont caractérisés par une structure secondaire composée de dix domaines où chacun des domaines a un rôle spécifique dans le repliement de l'intron. Les introns du groupe II sont composés de six domaines et le mécanisme d'épissage est différent de celui des introns du groupe I. Les génomes chloroplastiques des végétaux supérieurs contiennent une vingtaine d'introns du groupe II et un intron du groupe I (pour revue, Stern et al., 2010). Trois familles de protéines jouent un rôle dans l'épissage des introns chloroplastiques : les maturases, les protéines à domaine CRM (pour Chloroplast RNA splicing and ribosome Maturation) et les protéines PPR (pour revue, Eckardt, 2007 ; de Longevialle et al., 2010). La maturase MatK, dont le gène est localisé dans l'intron du gène chloroplastique *trnK*, est fortement soupçonnée de jouer un rôle dans l'épissage des introns chloroplastiques des gènes *trnK*, *trnA*, *trnL*, *trnV*, *rpl2*, *rps12* et *atpF* (Zoschke et al., 2010). Une autre maturase, codée par le génome nucléaire,

a été identifiée dans les génomes d'*Arabidopsis thaliana* et d'*Oryza sativa* (Mohr et Lambowitz, 2003). Cette maturase (AtnMat4) est adressée aux chloroplastes et aux mitochondries en fonction du codon d'initiation utilisé pour sa traduction : l'isoforme la plus longue est adressée aux mitochondries alors que l'isoforme la plus courte est adressée aux chloroplastes. Les mutants *AtnMat4* sont embryon létaux (Keren et al., 2009). Les protéines à domaines CRM (pour Chloroplast RNA splicing and ribosome Maturation), toutes codées par le génome nucléaire, sont impliquées dans le métabolisme des introns chez les organites. Les protéines CRM sont classées en fonction du nombre de domaines CRM qu'elles contiennent. Le premier gène codant une protéine CRM a été identifié chez *Zea mays* et a été nommé *CRS1* pour *Chloroplast RNA Splicing 1* (Jenkins et al., 1997). La protéine CRS1 est requise pour l'épissage de l'intron (appartenant au groupe II) du gène *atpF*. D'autres protéines à domaines CRM nécessitent des protéines supplémentaires pour former un complexe capable d'assurer l'épissage des introns. Par exemple, la protéine CRS2 est associée, dans les chloroplastes, à la protéine CAF1 (pour CRS2-associated factor 1) pour former un complexe CRS2-CAF1 nécessaire à l'épissage d'introns du groupe II chez *Zea mays* (Ostheimer et al., 2003). Certaines protéines à domaines CRM sont localisées à la fois dans les mitochondries et dans les chloroplastes, comme c'est le cas de la protéine AtCFM3a (Asakura et al., 2008). Le mutant *AtCFM2* a un phénotype albinos et est spécifiquement altéré dans l'épissage des introns chloroplastiques (Asakura et Barkan, 2007). Dans les chloroplastes, au moins deux protéines appartenant à la famille des PPR jouent un rôle dans l'épissage d'ARN spécifiques. Il s'agit de PPR4, identifiée chez *Zea mays* et d'OTP51 (pour Organellar Transcript Processing 51) identifiée chez *Arabidopsis thaliana*. Il a été montré que PPR4 s'associe au premier intron de *rps12*, un intron du groupe II (transcrits en segments et épissés en *trans*) et que PPR4 est nécessaire dans l'épissage en *trans* de cet intron : dans les mutants *ppr4*, cet intron n'est pas épissé (Schmitz-Linneweber et al., 2006). Il a également été montré que la protéine OTP51 est impliquée dans l'épissage de l'ARN du deuxième intron de *ycf3* ; OTP51 semble aussi influencer l'épissage de plusieurs ARNt et de *atpF* (de Longevialle et al., 2008). Deux autres protéines de la famille PPR, HCF152 et PPR5, semblent également impliquées dans l'épissage des ARN *psbB-psbT-psbH-petB-petD* et *trnG-UCC* respectivement (Meierhoff et al., 2003; Cushing et al., 2005; Beick et al., 2008).

La maturation des extrémités 5' et 3' est une étape post-transcriptionnelle essentielle pour l'obtention d'ARN fonctionnels. Différents mécanismes sont proposés pour expliquer la maturation des extrémités 5' des ARNm chloroplastiques. Il est vraisemblable qu'une

protéine, la RNaseJ, puisse assurer cette étape de maturation par ses activités endoribonucléasique et exoribonucléasique 5'→3' (Mathy et al., 2007). Dans ce modèle, un premier processus de clivage endoribonucléasique des résidus triphosphorylés présents aux extrémités 5' est suivi par un rapide « switch » de cette protéine vers une activité exoribonucléasique 5'→3'. Il a été montré que l'activité exoribonucléasique de la RNaseJ est inhibée par des résidus triphosphorylés (de la Sierra-Gallay et al., 2008), d'où la nécessité de les cliver pour permettre l'activation secondaire de l'activité exoribonucléasique. D'autres facteurs peuvent empêcher l'activité exoribonucléasique de ces enzymes comme des structures secondaires de la molécule d'ARNm et/ou des protéines spécifiques de liaison à l'ARNm. La RNase E est un second candidat pour cette étape de maturation (Walter et al., 2010). Tout comme les extrémités 5', les extrémités 3' des ARN chloroplastiques sont générées principalement par des processus post-transcriptionnels qui impliquent des réactions enzymatiques endoribonucléasique et exoribonucléasique. Aux extrémités 3' de l'ARN, des structures secondaires engendrées par des répétitions inversées, ou des protéines dont certaines protéines PPR, peuvent empêcher la digestion exoribonucléasique 3'→5' d'avoir lieu. Pour que l'activité de maturation exoribonucléasique 3'→5' ait lieu, ces structures secondaires et/ou la liaison de protéines protectrices peuvent être supprimées directement par clivage endoribonucléasique, ou indirectement via une étape de polyadénylation de l'extrémité 3' qui va permettre de restimuler l'activité exoribonucléasique pour « surpasser » ces facteurs et permettre la maturation de l'extrémité 3' (Hayes et al., 1999 ; Kudla et al., 1996). Les acteurs principaux de cette maturation 3' des transcrits sont les suivants : l'exoribonucléase PNPase, qui semblent être l'exoribonucléase majoritaire pour les ARNm, et l'exoribonucléase RNR1 ayant un rôle secondaire dans la maturation des extrémités 3' des ARNm mais un rôle majeur pour celle des ARNr. Un mutant *rnr1* d'*Arabidopsis* présente des défauts dans quatre ARNr chloroplastiques et aucun défaut visible dans le métabolisme des ARNm (Bollenbachet al., 2005). Deux endoribonucléases, RNase E et RNase J déjà fortement soupçonnées d'être impliquées dans la maturation des extrémités 5', sont également des candidats pour la maturation des extrémités 3' des ARNm ainsi qu'une RNase III pour la maturation d'ARNr (Drider et al., 2004). Des protéines portant des motifs RRM (pour RNA Recognition motifs) peuvent également être des facteurs en trans de la maturation des extrémités 3' (Tillich et al., 2009).

Les demi-vies des ARN chloroplastiques varient de 6 heures pour l'ARNm codé par le gène *psaA* à plus de 40 heures pour l'ARNm du gène *psbA* (Kim et al., 1993 ; Klaff et Gruissem ;

1991). Le processus de dégradation des ARN chloroplastiques se fait en trois étapes : un clivage endoribonucléasique, une polyadénylation et une dégradation exoribonucléasique. Dans un premier temps, le clivage endoribonucléasique va inactiver biologiquement le transcrit. Rnase J et Rnase E, ainsi que CSP41a/b sont des candidats pour le clivage interne et le clivage des extrémités des transcrits chloroplastiques (Yang et al., 1997). Le clivage endoribonucléasique va créer un substrat pour la polyadénylation des fragments d'ARN. Deux enzymes, PNPase et PAP, sont étudiées pour leur capacité à ajouter cette queue poly(A). Le rôle de PNPase dans cette étape est confirmé (Yechudai et al., 2003), alors que celui de PAP est encore hypothétique (Zimmer et al., 2009). Pour finir, les fragments de transcrits polyadénylés vont être dégradés par différentes enzymes : RNase II, RNR1 mais aussi PNPase qui a également une activité exoribonucléasique en plus de l'activité de polyadénylation (Hayes et al., 1996; Kishine et al., 2004; Perrin et al., 2004). Une famille de protéines, les RNP chloroplastiques (pour RiboNucleoProtéines), est particulièrement abondante dans le stroma des chloroplastes et est impliquée dans la stabilisation des ARN chloroplastiques. Ces protéines se caractérisent par la présence de motifs RRM (Lorkovic et Barta., 2002) et par une régulation dépendante de la lumière (Nakamura et al., 2004). Certains membres de cette famille sont impliqués dans la maturation des extrémités 3' des ARN (Hayes et al, 1996 ; Schuster et Gruissem, 1991), dans la stabilité des ARN (Nakamura et al., 2001), et dans l'édition des ARN (Hirose et Suguira., 2001). Le plus souvent, il semblerait que les RNP chloroplastiques s'associent aux ARN de façon non spécifique dans le chloroplaste. Ainsi, les protéines RNP impliquées dans la stabilité des ARN chloroplastiques sont le plus souvent des stabilisateurs non spécifiques ayant un rôle de protection contre la dégradation des ARN (Nakamura et al., 2004). Néanmoins, il semblerait que certaines protéines RNP aient un mode d'action plus spécifique sur certains transcrits chloroplastiques : l'étude génétique de *CP31A* et *CP31B*, deux gènes nucléaires codant des protéines RNP, semble indiquer que ces protéines RNP sont nécessaires à l'édition et à la stabilité d'ARN chloroplastiques spécifiques. Les protéines RNP forment certainement des complexes avec d'autres protéines pour assurer les fonctions liées au métabolisme des ARN chloroplastiques (Tillich et al., 2009).

### III. Génome mitochondrial des plantes actuelles

#### III.1 Taille et contenu en gènes mitochondriaux

En juillet 2011, 2496 génomes mitochondriaux d'eucaryotes étaient séquencés et disponibles dans les bases de données. Les plus petits génomes mitochondriaux connus sont ceux de protistes parasites appartenant au genre *Plasmodium* : ils font environ 6kb (Feagin et al., 2000). Chez les vertébrés, 1553 génomes mitochondriaux séquencés sont disponibles ; parmi eux, ceux d'*Homo sapiens sapiens* et d'*Homo sapiens neanderthalensis* font respectivement 16569 pb et 16565 pb.

Chez les plantes, la taille du génome mitochondrial est plus grande que chez les autres eucaryotes. Le premier génome mitochondrial de plante à avoir été séquencé, celui de *Marchantia polymorpha*, fait 186 kb (Oda et al., 1992). 47 génomes mitochondriaux séquencés, appartenant aux *Viridiplantae*, sont disponibles (32 chez les *Streptophyta* et 15 chez les *Chlorophyta*) ; chez les *Brassicaceae*, le génome mitochondrial d'*Arabidopsis thaliana*, séquencé en 1997, fait 367 kb (Unsel et al, 1997) et celui de *Brassica napus* fait 222 kb (Handa et al., 2003). Des génomes mitochondriaux de variants intraspécifiques sont également disponibles pour *Arabidopsis thaliana* (accessions Col et Ler), *Zea mays*, *Beta vulgaris* et *Oryza sativa*. Par exemple, la sous-espèce *mays* de *Zea mays* a un génome mitochondrial de 569 kb alors que celui de la sous-espèce *parviglumis* a une taille de 680 kb. Pour *Oryza sativa*, deux génomes de variants intraspécifiques sont également disponibles : ceux de *Oryza sativa Japonica* (490 kb) et de *Oryza sativa Indica* (491 kb) (Notsu et al., 2002). Une particularité des génomes mitochondriaux de plantes est la grande variabilité de leur taille entre différentes espèces végétales (ce qui n'est pas le cas pour les génomes chloroplastiques), mais aussi en intraspécifique. La taille moyenne des génomes mitochondriaux de *Viridiplantae* est de 282 kb avec un minimum de 13 kb pour *Polytomelle capuana*, une algue verte. Dans la famille des *Cucurbitaceae*, la taille des génomes mitochondriaux varie de 379 kb chez *Citrullus lanatus* à 983 kb chez *Cucurbita pepo*. Celui de *Cucumis melo*, également de la famille des *Cucurbitaceae*, est prédit pour avoir une taille de 2900 kb, mais ce génome mitochondrial n'a pas été séquencé (Kubo et Newton, 2008). Le plus grand génome mitochondrial séquencé disponible, celui de *Silene conica*, a une taille de 11,3 Mb (Sloan et al., 2012).

Les gènes mitochondriaux des eucaryotes codent 3 types de molécules : des protéines, des ARN ribosomiques (ARNr) et des ARN de transfert (ARNt). Chez les métazoaires

(organismes eucaryotes multicellulaires, mobiles et hétérotrophes), le nombre de gènes mitochondriaux codant des protéines varie de 7 chez *Anopheles funestus*, un moustique vecteur de la malaria, à 18 chez différentes espèces de calmars comme *Dosidicus gigas*. En règle générale, le nombre de gènes mitochondriaux codant des protéines est relativement constant chez les métazoaires avec environ 13 gènes par génome mitochondrial (comme c'est le cas chez *Homo sapiens sapiens*). Le nombre de gènes codant des ARNt et des ARNr est également constant chez les métazoaires avec environ 24 gènes par génome mitochondrial.

Chez les *Viridiplantae*, les *Streptophyta* ont un nombre de gènes mitochondriaux codant des protéines qui varie de 30 chez *Silene latifolia* à plus de 100 chez la sous espèce *mays* de *Zea mays*. Néanmoins, parmi ces 163 gènes, 125 sont des gènes de fonction inconnue, prédits par bioinformatique. Le nombre de gènes codant des ARNt et des ARNr varie de 13 chez *Silene latifolia* à 39 chez *Oriza sativa indica*. Chez les *Viridiplantae*, les *Chlorophyta* ont un nombre de gènes mitochondriaux codant des protéines qui varie de 7 chez *Polytomella capuana* à 72 chez *Pseudendoclonium akinetum*, deux espèces d'algues ; le nombre de gènes codant des ARNt et des ARNr varie de 12 chez *Dunaliella salina* à 40 chez *Micromonas sp*. Chez les *Viridiplantae*, le nombre de gènes mitochondriaux est moins constant d'une espèce à l'autre que chez les métazoaires.

Chez la plupart des angiospermes, les gènes mitochondriaux codent des sous-unités protéiques des complexes respiratoires I (NADH déshydrogénase), III (ubiquinone-cytochrome c réductase), IV (cytochrome oxydase) et V (ATP synthase), ainsi que des protéines qui participent à la biogenèse du cytochrome c. Les sous-unités protéiques du complexe II (succinate déshydrogénase) sont toutes codées par le génome nucléaire chez *Arabidopsis thaliana*, mais des sous-unités de ce complexe peuvent être codées par le génome mitochondrial chez d'autres espèces comme *Cucurbita pepo*. Chez cette espèce, les sous-unités 3 et 4 de la succinate déshydrogénase sont codées par le génome mitochondrial. Les gènes mitochondriaux codent également des protéines ribosomiques, des ARNr et des ARNt. La dégénérescence du code génétique fait qu'il existe souvent plusieurs codons pour un même acide aminé ; en conséquence, il existe plusieurs ARNt spécifiques du même acide aminé mais pourvus d'anticodons différents. Par exemple, dans le génome mitochondrial d'*Arabidopsis*, 5 gènes codent des ARNt pour l'acide aminé Ser. Le génome mitochondrial ne code pas tous les ARNt nécessaires à la traduction de ses ARNm. Les gènes codants les ARNt pour les acides aminés Ala, Arg, Leu, Thr et Val sont absents des génomes mitochondriaux

des plantes supérieures ; cette liste peut s'agrandir en fonction des espèces. Dans le génome mitochondrial d'*Arabidopsis thaliana*, il manque en plus les gènes codants les ARNt pour les acides aminés Phe, Met et Trp. Ces ARNt n'étant pas codés par le génome mitochondrial, les mitochondries doivent importer ces molécules d'ARNt à partir du cytosol (Marienfeld et al., 1999 ; Duchene et Marechal-Drouard, 2001). Le nombre et l'identité des ARNt importés dans les mitochondries diffèrent même entre espèces phylogénétiquement proches. Le génome mitochondrial d'*Arabidopsis thaliana* code 21 ARNt pour 13 acides aminés (Asn, Asp, Cys, Gln, Glu, Gly Ile, Lys, Met, Pro, Ser, Trp et Tyr). Celui de *Brassica napus* code 15 ARNt pour 12 acides aminés. Par comparaison aux ARNt codés par *Arabidopsis*, il manque, chez *Brassica napus*, un ARNt pour la Lys, deux pour la Ser et deux pour la Tyr. Un ARNt pour l'His est présent dans le génome mitochondrial de *Brassica napus*, alors qu'il est absent dans celui d'*Arabidopsis thaliana*. Chez *Oryza sativa*, le génome mitochondrial de la variété *Japonica* contient un gène codant un ARNt pour Val qui est absent de la variété *Indica*. Les deux gènes codants les ARNr de la petite sous-unité (18S) et de la grande sous-unité (26S) du ribosome sont présents dans tous les génomes mitochondriaux. Chez les végétaux terrestres, ainsi que chez certaines algues et chez *Reclinomonas*, le gène codant l'ARNr 5S est également présent dans le génome mitochondrial (Oudot-Le Secq et al., 2001 ; Bullerwell et al., 2003). Chez *Arabidopsis thaliana*, le génome mitochondrial code donc 3 ARNr (5S, 18S et 26S). Tout comme les gènes codant les sous-unités du complexe II succinate déshydrogénase et ceux codant les ARNt, les gènes codant les protéines ribosomiques peuvent, selon les espèces végétales, être localisés dans le génome mitochondrial ou dans le génome nucléaire. Le génome mitochondrial d'*Arabidopsis thaliana* code 7 protéines ribosomiques et celui de *Physcomitrella patens*, 14 protéines ribosomiques. Le génome mitochondrial contient également des gènes qui codent d'autres protéines de fonction connue. Par exemple, le gène mitochondrial *MatR* code une maturase (Unsel et al., 1997). Chez les plantes, *MatR* est situé dans le quatrième intron du gène *nad1* et, selon toute vraisemblance, joue un rôle dans l'épissage de certains ARN mitochondriaux (de Longevialle et al., 2010). Le gène *MatR* est présent dans les génomes mitochondriaux de nombreuses espèces végétales comme *Zea mays* et *Oryza sativa*. Néanmoins, *MatR* est absent du génome mitochondrial de *Physcomitrella patens*.

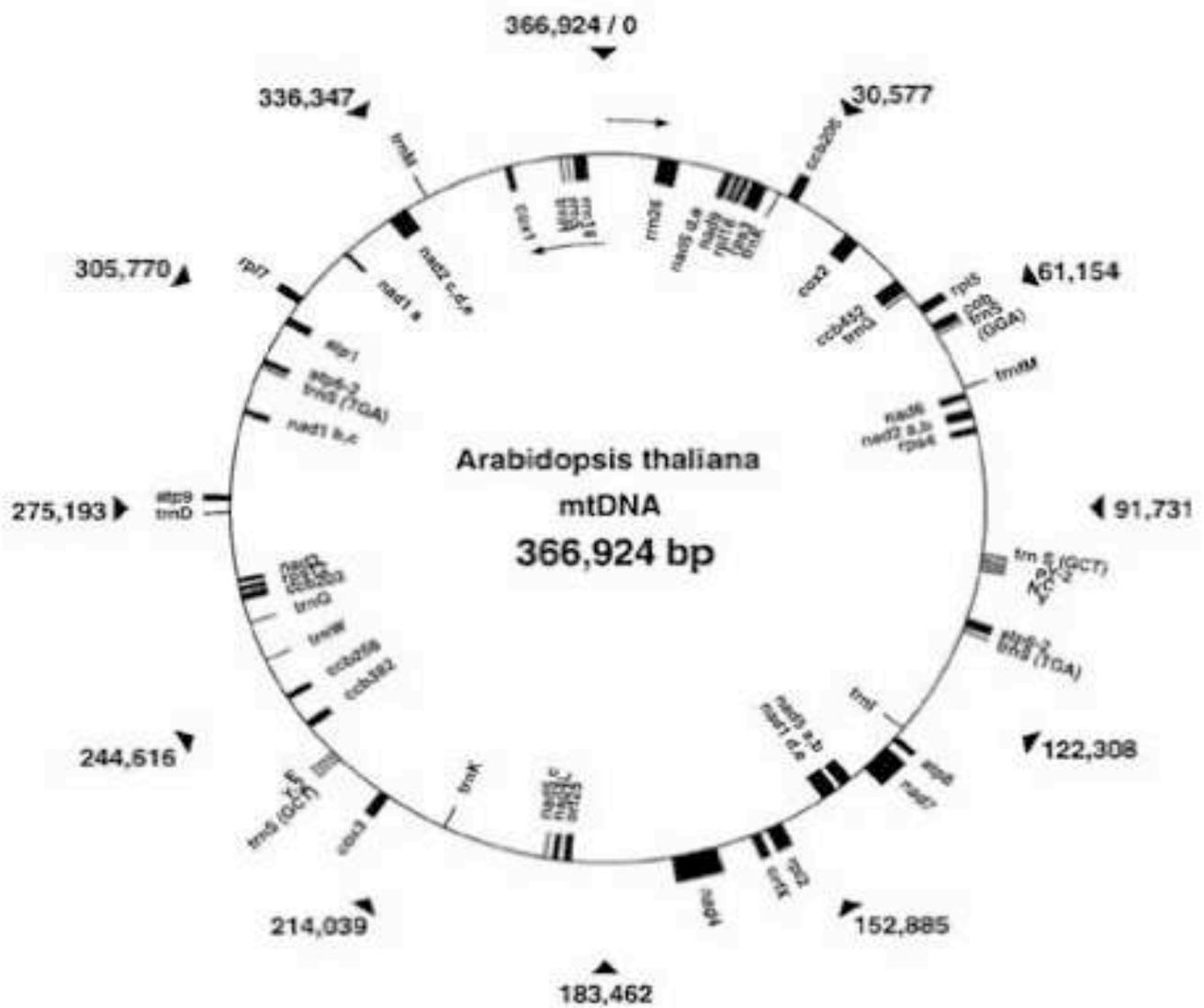
Les génomes mitochondriaux des plantes supérieures contiennent des ORF (pour Open Reading Frame : séquence ouverte de lecture) de fonction inconnue. Certaines sont conservées entre espèces, et donc supposées avoir une fonction conservée, d'autres ne sont

pas conservées et ne trouvées dans le génome mitochondrial d'une seule espèce, voire un cytotype particulier. Un nombre important de protéines hypothétiques, sans fonction connue, sont prédites à partir des séquences des génomes mitochondriaux (chez *Arabidopsis thaliana*, plus de 80 protéines hypothétiques sont prédites). Les ORF conservées codent certainement des protéines dont les fonctions ne sont pas encore identifiées : par exemple, les gènes *atp4* et *atp8* étaient autrefois annotés *orf25* et *orfB*. Aujourd'hui, nous savons qu'ils codent des sous-unités protéiques du complexe ATP synthase (Heazlewood et al., 2003). Concernant les ORF non conservées, nous verrons dans le chapitre suivant que les génomes mitochondriaux de plantes subissent des recombinaisons fréquentes et que les réarrangements de l'ADN mitochondrial dans des lignées particulières peuvent être à l'origine de l'apparition d'ORF non conservées. Nous verrons également dans le paragraphe VI.4.1 que certaines de ces ORF non conservées sont impliquées dans les phénotypes de stérilité mâle nucléo-cytoplasmique (Hanson et Bentolila, 2004).

Comme nous l'avons vu, les génomes mitochondriaux des végétaux sont plus grands et plus variables en taille que ceux des génomes mitochondriaux d'autres eucaryotes. Dans le cas où on ne considère pas les *orf* non conservées, cette grande taille n'est pas corrélée à une plus grande capacité codante de protéines de fonction connue. Il semblerait que ce soient les régions intergéniques, plus grandes dans les génomes mitochondriaux de végétaux que quoi ?, qui sont responsables de cette différence de taille. L'analyse des régions intergéniques des mitochondries végétales montre qu'elles sont constituées de séquences chloroplastiques (1,6-6,2% du génome), de séquences nucléaires (0,1-13,4% du génome) et de séquences d'origine inconnue (Kubo et Mikami, 2007 ; Allen et al., 2007).

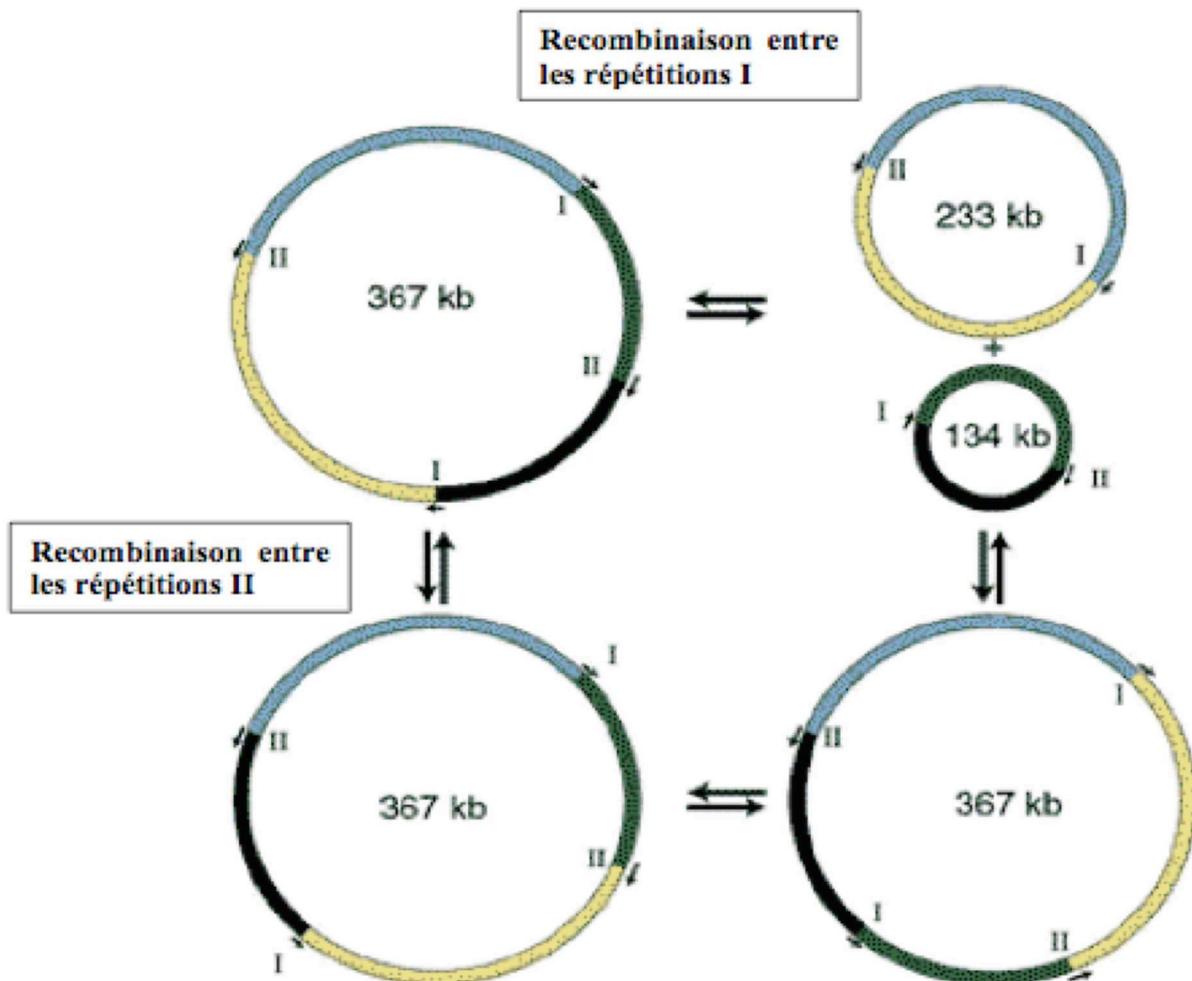
### III.2 Structure et organisation du génome mitochondrial

D'après Lonsdale, le génome mitochondrial des angiospermes se présente sous la forme d'un chromosome circulaire double brin nommé « Cercle Maître » (CM), contenant tous les gènes mitochondriaux et les séquences non codantes (Lonsdale, 1988, revue de Scheffler, 1999). Les tailles des génomes mitochondriaux, mentionnées dans le paragraphe III.1, font référence à la taille des CM. Cependant, les molécules circulaires de la taille d'un CM ont rarement été détectées directement dans les préparations d'ADN mitochondrial de plantes, et leur présence n'est validée que par le clonage moléculaire (Kitazaki and Kubo, 2010), par des marqueurs génétiques et par l'assemblage de fragments d'ADN (Bendich et Drlica, 2000). Le CM du



**Figure A4 : distribution des gènes identifiés dans le CM du génome mitochondrial d'*Arabidopsis thaliana***

33 protéines de fonction connue, 3 ARNr et 20 ARNt sont codés par le génome mitochondrial d'*Arabidopsis thaliana* (Unsel et al., 1997 ; Duchene et Marechal-Drouard, 2001). Les boîtes noires représentent les gènes mitochondriaux (les noms des gènes sont indiqués à côté) (D'après Dombrowski et al., 1998).



**Figure A5 : modèle représentant la structure multipartite du génome mitochondrial d'*Arabidopsis thaliana*.**

L'ADN mitochondrial contient des répétitions directes (I) et des répétitions inversées (II). Lorsque deux copies de la même répétition sont présentes en orientation directe (répétitions I) l'une par rapport à l'autre sur le CM (de 367 kb), ces copies induisent des recombinaisons intragénomiques à l'origine de la formation de cercles sub-génomiques (de 233 kb et de 134 kb). Lorsque deux copies de la même répétition sont présentes en orientation inverse sur le CM (répétitions II), elles induisent la formation de formes isomériques présentant des inversions par rapport au CM (Palmer and Shields, 1984 ; Shedge et al., 2007). Schéma adapté de Rasmusson et al., Plant physiology online Fifth Edition (<http://5e.plantphys.net/article.php?ch=0&id=151>).

génomique mitochondrial d'*Arabidopsis thaliana* est présenté dans la figure A4.

Une forte activité de recombinaison intragénomique caractérise les génomes mitochondriaux des angiospermes (Mackenzie, 2007). Les recombinaisons se font grâce à des séquences répétées présentes dans l'ADN mitochondrial. Le nombre et la taille de ces séquences répétées varient en fonction des espèces (Alverson et al., 2011) : le génome mitochondrial d'*Arabidopsis thaliana* contient 69 séquences répétées qui ont une taille supérieure à 50 pb alors que celui de *Vitis vinifera* en contient 114 (Kurtz et al., 2001 ; Marechal et Brisson, 2010). En règle générale, parmi ces séquences répétées de plus de 50 bp, chaque génome mitochondrial contient au moins une paire de grandes séquences répétées en orientation directe et une paire de grandes séquences répétées en orientation indirecte. Lorsque deux copies de la même répétition sont présentes en orientation directe l'une par rapport à l'autre sur le CM, ces copies induisent des recombinaisons intragénomiques à l'origine de la formation de cercles sub-génomiques. Lorsque deux copies de la même répétition sont présentes en orientation inverse sur le CM, elles induisent la formation de formes isomériques présentant des inversions par rapport au CM (Palmer and Shields, 1984 ; Shedje et al., 2007). Les petites séquences répétées favorisent les recombinaisons intergénomiques pouvant être à l'origine de la formation de nouvelles ORF (Vedel et al., 1994). La fréquence de recombinaison est proportionnelle à la taille des séquences répétées : les grandes séquences répétées (supérieures à 1 kb) recombinent à forte fréquence ; les séquences répétées de taille intermédiaire (de 100 à 1000 pb) recombinent de façon sporadique et les petites séquences répétées (inférieures à 100 pb) recombinent rarement (Arrieta-Montiel et al., 2009 ; André et al., 1992 ; Marechal et al., 2010). L'organisation du génome mitochondrial chez les végétaux, est donc considérée comme multipartite et dynamique où coexistent des génomes qui sont au même niveau de stœchiométrie que le CM mais qui ont des configurations différentes (Palmer et Shields, 1984 ; Fauron et al., 1995). Un modèle de l'organisation multipartite du génome mitochondrial d'*Arabidopsis thaliana* est proposé dans la figure A5.

L'évolution de la structure des génomes mitochondriaux de plantes est rapide (Palmer et Herbon, 1988) et influencée par des recombinaisons intra et intermoléculaires générant un groupe hétérogène de configurations génomiques au sein d'un individu (Palmer et Shields, 1984 ; Marechal et Brisson, 2010). Des recombinaisons sporadiques entre de petites répétitions (supérieures à 100 pb et inférieures à 1 kb) peuvent produire des molécules d'ADN présentes à un niveau sous-stœchiométrique, appelées sublimons. Des événements de

recombinaison impliquant ces sublimons seraient impliqués dans l'évolution des génomes mitochondriaux de plante (Small et al., 1987). Les génomes mitochondriaux des végétaux supérieurs étant très stables en séquence et caractérisés par un faible taux de mutation, les différentes formes et nombre de copies des génomes favoriseraient ainsi la variabilité génétique mitochondriale (Palmer, 1988 ; Burger et al., 2003 ; Kubo et Mikami, 2007).

### III.3 Expression des gènes mitochondriaux

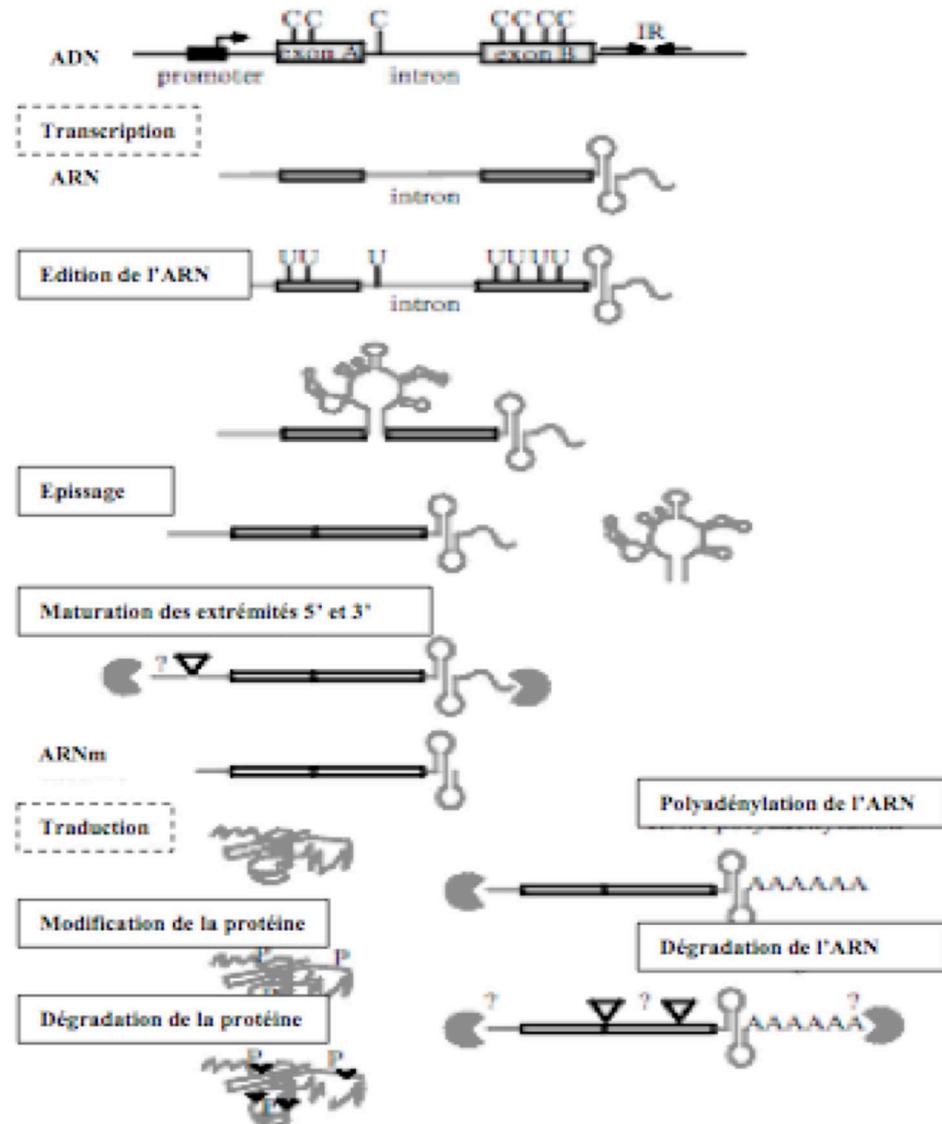
Chez la majorité des eucaryotes, la transcription des gènes mitochondriaux est réalisée par une ou deux ARN polymérases de type phagique codées par le noyau. Chez les algues vertes, comme *Chlamydomonas*, chez le lycophyte *Selaginella* et chez les céréales, un seul gène nucléaire code une ARN polymérase adressée aux mitochondries (Yin et al., 2010). Chez les angiospermes et chez *Physcomitrella patens*, plusieurs ARN polymérases de type phagique sont codées par une petite famille de gènes nucléaires. Le génome nucléaire de *Physcomitrella patens* possède 3 gènes codant des ARN polymérases de type phagique ; deux d'entre elles (RpoTmp1 et RpoTmp2) sont adressés aux mitochondries et aux chloroplastes et une (RpoTm) est adressée seulement aux mitochondries (Richter et al., 2002 ; Kabeya et Sato, 2005). Deux ARN polymérases, RpoTm et RpoTmp, codées par le noyau, sont adressées aux mitochondries et responsables de la transcription des gènes mitochondriaux chez *Arabidopsis thaliana* (Cahoon et Stern, 2001). RpoTm est adressée uniquement aux mitochondries et RpoTmp est adressée à la fois aux mitochondries et aux chloroplastes. RpoTm et RpoTmp semblent toutes les deux importantes pour la transcription des gènes mitochondriaux lors de la gamétogenèse, mais ces deux polymérases auraient des rôles distincts lors du développement de la plante (Tan et al., 2010). Chez le protiste *Reclinomonas americana*, l'ARN polymérase mitochondriale est codée par le génome mitochondrial (Lang et al., 1997).

Peu de données sont disponibles sur d'éventuels cofacteurs impliqués dans l'initiation de la transcription des gènes mitochondriaux (Kühn et al., 2005). Néanmoins, chez les plantes, les ARN polymérases adressées aux mitochondries pourraient nécessiter, *in vivo*, des facteurs supplémentaires pour assurer une initiation efficace et précise de la transcription des gènes mitochondriaux. Des co-facteurs candidats ont été étudiés : chez le blé, une protéine de 69 kDa, de la famille des protéines à motifs PPR, stimule, *in vitro*, la transcription du promoteur *cox2* (Ikeda et Gray, 1999 ; Schmitz-Linneweber et Small, 2008). Des tests *in vitro*, réalisés chez *Arabidopsis*, ont montré que d'autres protéines PPR sont nécessaires à une initiation

correcte de la transcription mitochondriale. Elles se lieraient de façon non spécifique et sans interagir avec RpoTm ou RpoTmp, à des fragments de promoteur de gènes mitochondriaux (Kühn et al., 2007). Des facteurs sigma, codés par le noyau et adressés aux chloroplastes, semblent également localisés dans les mitochondries. C'est le cas notamment pour les facteurs sigma AthSig5 chez *Arabidopsis thaliana* (Fujiwara et al., 2000) et ZmaSig2B chez le maïs (Beardslee et al., 2002). De plus, par des expériences de biochimie, il a été montré que ZmaSig2B co-purifie avec RpoTm (Beardslee et al., 2002), ce qui laisse penser que ce facteur sigma est impliqué dans la transcription des gènes mitochondriaux au moins chez *Zea mays*.

Concernant les régions promotrices des gènes mitochondriaux, des tests *in vitro* et des analyses de séquence ont identifié des sites d'initiation de la transcription et des motifs de promoteurs chez de nombreuses espèces dont *Pisum sativum* (Hoffmann et Binder, 2002) et *Nicotiana tabacum* (Edqvist et Bergman, 2002). Chez les dicotylédones, certains promoteurs mitochondriaux ont des motifs consensus de type CRTA (Weihe, 2004), ATTA ou RGTA qui sont reconnus par les ARN polymérases. Plusieurs séquences d'ADN mitochondrial, identifiées comme étant des régions promotrices, ne possèdent pas de motif consensus (Kühn et al., 2005). Il est possible que d'autres motifs consensus de régions promotrices existent mais que n'aient pas encore été identifiés. La présence, pour un seul gène mitochondrial, de plus d'un promoteur et de multiples sites d'initiation de la transcription est une caractéristique commune aux génomes mitochondriaux des eudicotylédones et des monocotylédones (Tracy et Stern, 1995 ; Kühn et al., 2005). C'est par exemple le cas du gène *atp9* qui, chez *Zea mays*, possède 6 sites d'initiation de la transcription (Kühn et al., 2005). Il est possible que la présence de multiples promoteurs pour un seul gène puisse permettre d'assurer la transcription en dépit des réarrangements fréquents du génome mitochondrial (Kühn et al., 2005). Les réarrangements intra et intergénomiques peuvent modifier l'activité promotrice d'un gène mitochondrial : l'activité des promoteurs de *cox2* est par exemple dépendante du contexte génomique dans lequel ils se trouvent (Lupold et al., 1999). Néanmoins, il apparaît que la transcription, dans les mitochondries de plantes, est un processus « relâché » qui, en général, est peu contrôlé et peu modulé (Holec et al., 2006). Des processus post-transcriptionnels semblent être responsables du contrôle de l'expression des gènes mitochondriaux.

Après l'étape de transcription, les transcrits codés par les génomes mitochondriaux peuvent être maturés pour devenir fonctionnels. Les principaux processus utilisés pour la maturation



**Figure A6 : étapes de la maturation post-transcriptionnelle des ARN dans les mitochondries de plantes**

Plusieurs étapes de maturation du transcrit primaire sont nécessaires avant la traduction de l'ARNm mature. Ces étapes sont l'édition des ARN, l'épissage des introns et la maturation des extrémités 5' et 3'. Le contrôle de l'expression des gènes peut se faire à chacune de ces étapes. La dégradation de l'ARNm peut être favorisée par une polyadénylation de l'extrémité 3' UTR (d'après Binder et Brennicke, 2003).

des transcrits mitochondriaux connus à l'heure actuelle sont la modification des extrémités 5' et 3', l'édition et l'épissage des introns (figure A6). Quelques protéines impliquées dans les étapes post-transcriptionnelles du métabolisme des ARN ont été identifiées et/ou caractérisées dans les mitochondries de plantes.

L'observation des extrémités des ARNm mitochondriaux chez *Arabidopsis thaliana* a révélé que, à partir d'une même séquence d'ADN, plusieurs ARN ayant des extrémités 5' différentes pouvaient être générés. L'hypothèse de la génération de multiples extrémités 5' lors de l'initiation de la transcription à différents sites d'initiation ne peut pas être rejetée formellement. Il semble néanmoins que la majorité des extrémités 5' des ARNm matures soit générée par un ou des processus post-transcriptionnels (Forner et al., 2007). Très peu d'acteurs de la maturation des extrémités 5' des transcrits mitochondriaux ont été décrits. Aucune activité 5'→3' exoribonucléasique n'a été détectée dans les mitochondries, ce qui semble indiquer que la maturation de ces extrémités se ferait plutôt par des activités endoribonucléasiques (Forner et al., 2007 ; Forner et al., 2008). Dans les mitochondries de plantes, les régions 3' et 5' non codantes des ARNm peuvent contenir des structures secondaires, appelées éléments t, comparables à celles des ARNt. Les endonucléases RNaseZ et RNaseP, dont on sait qu'elles sont impliquées dans la maturation des ARNt mitochondriaux, semblent également pouvoir cliver des éléments t présents dans certains ARNm mitochondriaux (Canino et al., 2009). Des facteurs codés par le noyau et adressés aux mitochondries sont impliqués dans la maturation des extrémités 5' des transcrits mitochondriaux. Chez *Arabidopsis thaliana*, le gène nucléaire *RPF2* (pour *RNA PROCESSING FACTOR2*) code une protéine à motifs PPR qui influence la formation des extrémités 5' des transcrits codés par les gènes mitochondriaux *nad9* et *cox3* (Jonietz et al., 2010). Le véritable rôle de *RPF2* dans la maturation des extrémités 5' des ARNm codés par *nad9* est pour l'instant inconnu. Son rôle est plus clair en ce qui concerne la maturation des extrémités 5' des ARNm codés par *cox3*. Chez le sauvage, les ARNm matures de *cox3* ont une taille de 1,5 kb et ont une extrémité 5' située à 380 bases en amont du site d'initiation de la traduction. L'inactivation de *RPF2* provoque l'accumulation de transcrits *cox3* supplémentaires plus longs, ayant des extrémités 5' localisées à 448 et 540 bases en amont du site d'initiation de la traduction. L'accumulation, chez le mutant *rpf2*, de ces ARN supplémentaires s'accompagne d'une diminution de la quantité des transcrits matures (extrémité 5' à -380) ; cela indique que les molécules d'ARNm supplémentaires observées chez le mutant sont certainement des intermédiaires dans la génération d'ARNm *cox3*

matures. Une des hypothèses concernant le rôle de RPF2 est que cette protéine puisse guider les endonucléases RNaseZ ou RNaseP vers les éléments t des transcrits *nad9* et *cox3* pour cliver ces éléments t et ainsi maturer les extrémités 5' des transcrits *nad9* et *cox3* (Jonietz et al., 2010). Un autre gène nucléaire, *RPF1* (pour *RNA PROCESSING FACTOR 1*) code une protéine à motifs PPR essentielle à la maturation de l'extrémité 5' de l'ARNm *nad4* chez *Arabidopsis thaliana*. Dans les plantes qui contiennent une protéine RPF1 fonctionnelle, les ARNm *nad4* sont majoritairement maturés avec une extrémité 5' à -228 ; quand RPF1 est non fonctionnelle, les ARNm *nad4* ont majoritairement une extrémité 5' à -390. Ces observations peuvent suggérer que les ARN ayant une extrémité 5' à -390 sont des intermédiaires de la maturation de l'ARNm *nad4*. Ces ARN seraient convertis en ARNm matures (extrémités 5' à -228) grâce à l'assistance de RPF1. Comme RPF2, RPF1 semble être un facteur qui améliore le ou les processus post-transcriptionnels à l'origine de la maturation de l'extrémité 5' d'ARNm mitochondriaux (Hölzle et al., 2011). RPF1 et RPF2 sont deux protéines de la famille des *Rf*-like, décrite dans le chapitre VI. 4. 2.

Des études fonctionnelles ont montré que deux exoribonucléases 3'→5' ont un rôle crucial dans la maturation en 3' des ARNm mitochondriaux. Il s'agit d'une PNPase (pour polynucleotide phosphorylase) et d'une RNase R mitochondriale (Perrin et al., 2004). Chez *Arabidopsis thaliana*, ces deux protéines sont appelées AtmtPNPase et AtmtRNaseII. Lorsque AtmtPNPase est sous exprimée, il apparaît une accumulation d'ARN *atp9* qui ont des régions 3' non codantes beaucoup plus longues que les ARNm *atp9* matures chez la plante sauvage. Lorsque AtmtRNaseII est sous exprimée, environ la moitié des ARN *atp9* présentent de courtes extensions en 3' comparés aux ARN matures de la plante sauvage. Ces petites extensions peuvent être supprimées par une action de l'AtmtRNaseII *in vitro*. Les extrémités 3' des ARNm *atp9* mitochondriaux d'*Arabidopsis* semblent donc maturées *via* au moins deux étapes post-transcriptionnelles : AtmtPNPase va tout d'abord retirer de grandes extensions en 3' de l'ARNm (pouvant atteindre plusieurs kb), puis AtmtRNaseII va dégrader les courtes extensions restantes pour générer les extrémités 3' caractéristiques de l'ARNm mature (Perrin et al., 2004a). La protéine AtmtPNPase semble aussi impliquée dans le métabolisme de l'ARNr 18S dans les mitochondries d'*Arabidopsis thaliana* (Perrin et al., 2004b).

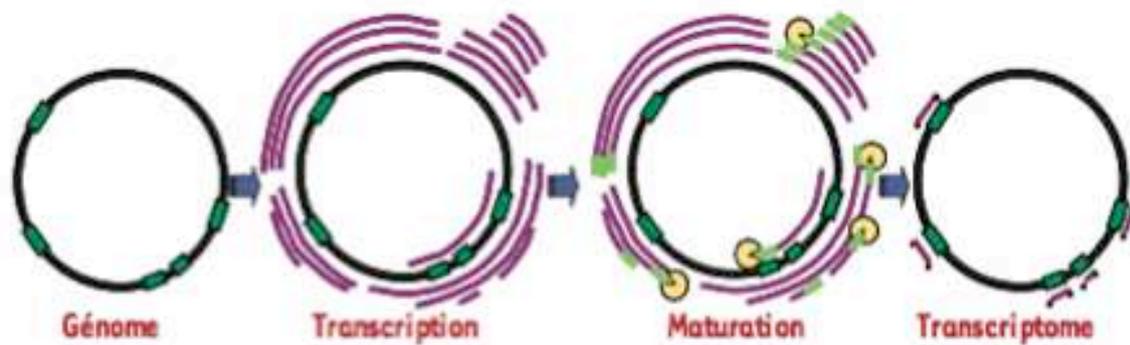
Tout comme les transcrits des gènes chloroplastiques, les transcrits mitochondriaux subissent le processus d'édition des ARN. Chez *Arabidopsis thaliana*, plus de 500 sites sont édités dans les transcrits mitochondriaux. La très grande majorité de ces sites est située dans les cadres

ouverts de lecture. L'édition d'un résidu cytidine est déterminée par des éléments en *cis*, situés à proximité du site d'édition, qui sont spécifiquement reconnus par des facteurs de liaison agissant en *trans* pour faciliter l'accès d'une enzyme hypothétique capable d'éditer l'ARN (Okuda et al., 2006). Les facteurs de liaison identifiés dans ce processus appartiennent à la famille des protéines à motifs PPR (Hammani et al., 2011). Plusieurs membres de cette famille ont été identifiés comme ayant un rôle dans l'édition des ARN mitochondriaux de végétaux. En voici quelques exemples : chez *Physcomitrella patens*, la protéine PpPPR\_71 est requise pour l'édition du site *ccmF-2* présent dans le transcrit mitochondrial *ccmFc*. PpPPR\_71 se lie spécifiquement dans la région -40/+5 couvrant le site d'édition (Tasaki et al., 2010). La protéine PPR OTP87 est requise pour l'édition d'au moins deux sites, *nad7-C24* et *atp1-C1178*, dans les mitochondries d'*Arabidopsis* (Hammani et al., 2011). Chez *Arabidopsis thaliana*, la protéine PPR MEF14 (pour Mitochondrial Editing Factor) est un facteur en *trans* requis pour l'édition du site *matR-1895* dans les mitochondries (Verbitskiy et al., 2011). Chez *Arabidopsis thaliana*, l'inactivation de PPR596 augmente l'édition de plusieurs sites d'ARN mitochondriaux (Doniwa et al., 2010). Comme pour l'édition des ARN chloroplastiques, les conséquences de l'édition des ARN mitochondriaux sont variables en fonction du résidu modifié : sur la première et la seconde position du codon, l'édition va induire un changement d'acide aminé dans la protéine alors que l'édition de la troisième position ne va pas induire de changement d'acide aminé dans la plupart des cas. L'apparition de codons d'initiation et de terminaison suite à l'édition a également été observée (Giegé et Brennicke, 1999). Parmi les mutants d'*Arabidopsis* qui sont spécifiquement altérés dans l'édition d'un ou de plusieurs sites cytidine de transcrits mitochondriaux, peu de mutants (à l'exception notamment du mutant *OTP87* (Hammani et al., 2011) présentent des phénotypes sévères.

Les gènes mitochondriaux de plantes, de protistes et de champignons, qui codent des ARNr, des ARNt et des ARNm, peuvent contenir des introns. L'expression des gènes mitochondriaux dépend du retrait, par épissage des transcrits, des séquences introniques. Ces introns sont classés en deux groupes : les introns du groupe I et les introns du groupe II. Les introns du groupe I ont une structure secondaire composée de dix domaines qui ont chacun un rôle spécifique dans le repliement de l'ARN. La structure secondaire des introns du groupe II contient six domaines hélicoïdaux, relativement conservés, autour d'un noyau central dont les variations de séquence et de structure permettent de distinguer deux sous groupes, IIA et IIB, à l'intérieur du groupe II (Barkan, 2004). Le génome mitochondrial d'*Arabidopsis* contient 25 introns du groupe II (Bonen, 2008). La majorité des introns du groupe II se trouvent dans les

gènes codant les sous-unités du complexe 1 ; les autres introns du groupe II se trouvent dans les régions codantes des gènes *ccmFc*, *cox2*, *rpl2* et *rps3* (Unsold et al., 1997). Un intron du groupe I, présent dans le gène *cox1* des mitochondries de plantes, a également été identifié (Cho et al., 1998). Les deux groupes d'introns se distinguent principalement par leur structure et leur biochimie. Le mode d'épissage des introns du groupe I se fait par transesterification et les introns linéaires épissés sont circularisés. L'épissage des introns du groupe II se fait principalement par deux étapes de transestérification amenant les introns épissés à avoir une forme de lasso. De plus, les introns du groupe I ne possèdent pas de sites consensus d'épissage alors que ceux du groupe II en contiennent (pour revue, Bonen et Vogel, 2001). La nature dynamique des génomes mitochondriaux de plantes, *via* de nombreuses recombinaisons génomiques, peut induire la fragmentation de gènes dont les exons sont dispersés dans le génome. Ces exons sont transcrits indépendamment et l'ARNm peut ensuite être généré grâce à un épissage *en trans* des introns morcelés (Bonen, 1993).

Le retrait des introns est dépendant de l'activité de nombreux facteurs d'épissage (pour revue, Bonen et Vogel, 2001 ; Barkan, 2004 ; Gagliardi et Binder, 2007 ; Bonen 2008). Certains introns des deux groupes codent des ARN catalytiques, mobiles et leurs propres machineries distinctes d'épissage et de mobilité : ces ORF introniques codent des fonctions pour l'épissage (ARN maturase) et la mobilité (endonucléase « autoguidée »). Par exemple, l'épissage des introns du groupe II dans les mitochondries de levures est facilité par des maturases qui sont codées par les introns mitochondriaux. Chez ces organismes, les maturases agissent spécifiquement dans l'épissage de l'intron qui les a codés (Robart et Zimmerly, 2005). Les introns du groupe II dans les génomes mitochondriaux de plantes ont certaines particularités : ils sont fortement dégénérés ; ils ne codent plus, pour la majorité d'entre eux, les maturases nécessaires à leur épissage et ils sont incapables de s'auto-épisser *in vitro* (Keren et al., 2009). La seule maturase présente dans le génome mitochondrial des plantes est MATR ; son rôle dans l'épissage des introns reste à confirmer (Keren et al., 2009). Ces particularités suggèrent que l'épissage des introns du groupe II, dans les mitochondries de plantes, nécessite des facteurs codés par le noyau : chez *Nicotiana sylvestris*, une mutation nucléaire (*nms1*) perturbe l'épissage du premier intron du transcrit *nad4* (Brangeon et al., 2000). Chez *Arabidopsis thaliana*, une protéine PPR (OPT43) est impliquée dans l'épissage du premier intron de *nad1* (de Longevialle et al., 2007). De plus, l'analyse du génome nucléaire d'*Arabidopsis thaliana* et d'*Oryza sativa* a révélé l'existence de cinq gènes nucléaires très apparentés à des maturases codées par des introns du groupe II chez d'autres organismes (Mohr et Lambowitz, 2003 ; Keren et al., 2009). Enfin, il a été montré récemment



**Figure A7 : modèle présentant la polyadénylation des extrémités 3' des transcrits et la dégradation de ces transcrits par la PNPase**

Le génome est représenté par un cercle, les gènes par des rectangles verts. La transcription mitochondriale donne lieu à une population de transcrits primaires, représentés par des traits violets. Après la polyadénylation, représentée par des ajouts verts, la PNPase, en jaune, dégrade les ARN aberrants et mature les précurseurs. Il en résulte un transcriptome constitué de transcrits matures, dont les extrémités sont ici marquées par des points (Holec et Gagliardi, 2006).

qu'un lien existe entre l'édition et l'épissage des introns : en effet, des événements d'édition existent dans des introns mitochondriaux, à des positions importantes pour la structure secondaire et tertiaire des introns. Une modification de l'édition à ces positions pourrait influencer la structure de l'intron, ainsi que son épissage (Castandet et al., 2010).

Une famille de protéines, les protéines DEAD-box qui sont considérées comme des hélicases de l'ARN (pour revue, Rocak et Linder, 2004 ; Linder 2006), peut théoriquement être impliquée dans tous les processus liés aux ARN : synthèse de l'ARN, modification de l'ARN, clivage de l'ARN, dégradation de l'ARN, biogenèse des ribosomes et initiation de la traduction (Linder, 2006). Chez *Arabidopsis thaliana*, 58 gènes nucléaires qui codent des protéines DEAD-box ont été identifiés (Boudet et al., 2001 ; Mingam et al., 2004 ; Matthes et al., 2007). Deux protéines DEAD-box adressées aux mitochondries, PMH1 et PMH2 (pour PUTATIVE MITOCHONDRIAL RNA HELICASE) ont été identifiées comme ayant un rôle dans le métabolisme des ARN mitochondriaux chez *Arabidopsis thaliana*. Chez des mutants *pmh2*, l'efficacité de l'épissage d'introns mitochondriaux est réduite et les niveaux de plusieurs ARNm matures mitochondriaux sont également réduits sans que la transcription ne soit affectée. Cela semble indiquer que PMH2 agit au niveau post-transcriptionnel, avec peu de spécificité ; PMH2 pourrait être une protéine chaperonne d'ARN, nécessaire à la formation ou au maintien de structures secondaires de l'ARN. La fonction précise de PMH1 dans le métabolisme des ARN mitochondriaux est inconnue (Köhler et al., 2010).

L'abondance des ARN a un impact important pour le fonctionnement correct de la cellule. La présence de chaque ARN est le résultat d'un équilibre entre la synthèse et la dégradation des ARN. Dans les mitochondries de plantes, l'ajout d'une série d'adénosines à l'extrémité 3' de l'ARN, appelée polyadénylation (Dreyfus et Régnier, 2002), est un signal qui favorise la dégradation des ARN mitochondriaux par la PNPase (exoribonucléase 3'->5') (Perrin et al., 2004). Au moins chez *Arabidopsis thaliana*, le transcriptome mitochondrial dépend du contrôle de la stabilité des ARN qui s'appuie sur la polyadénylation et la dégradation par la PNPase. Ce modèle est présenté dans la figure A7. Il a aussi été proposé que des structures en tige-boucle, situées aux extrémités 3' des ARNm, ou des facteurs spécifiques codés par le noyau, puissent permettre à ces transcrits d'être stables en les protégeant contre l'activité des exoribonucléases 3'->5' (Dombrowski et al., 1997 ; Bellaoui et al., 1997 ; Xiao et al., 2006).

## IV. Interactions nucléo-cytoplasmiques dans la cellule végétale

### IV.1 La famille des protéines à motifs PPR

Comme nous l'avons vu dans les chapitres II et III, l'expression des gènes chloroplastiques et mitochondriaux est très largement contrôlée par des protéines codées par le noyau. En 2000, une nouvelle famille de gènes est identifiée chez les plantes par des approches de bioinformatique (Small and Peeters, 2000 ; Aubourg et al., 2000). Ces gènes codent des protéines, nommées protéines à motifs PPR (pour Pentatricopeptide Repeat), composées de motifs dégénérés de 35 acides aminés répétés en tandem de 2 à 26 fois (Small et Peeters, 2000). Les protéines à motifs PPR sont présentes chez tous les eucaryotes avec néanmoins une expansion de la famille chez les végétaux supérieurs (pour revue, O'Toole et al., 2008), ce qui laisse supposer un rôle particulier de ces protéines chez ces organismes. Le nombre de gènes codant des protéines à motifs PPR a explosé lors de l'évolution des plantes terrestres (Aubourg et al., 2000). On compte 450 gènes codant ces protéines chez *Arabidopsis thaliana*, 477 chez *Oriza sativa* (Lurin et al., 2004) et environ 600 chez *Populus trichocarpa* et *Vitis vinifera*. Par comparaison, il n'existe que 103 gènes codants des protéines à motifs PPR chez la mousse *Physcomitrella patens*, 11 chez *Chlamydomonas reinhardtii*, 5 chez *Saccharomyces cerevisiae* et 6 chez *Homo sapiens* (Schmitz-Linneweber and Small, 2008). L'organisme non végétal qui contient le plus de gènes codant des protéines à motifs PPR est un protozoaire parasite *Trypanosoma brucei* avec 28 gènes prédits. Les protéines à motifs PPR semblent absentes chez les procaryotes comme la protéobactérie *Rickettsia prowazekii* et la cyanobactérie *Synechocystis sp.* (Pusnik et al., 2007), ce qui laisse penser qu'elles sont spécifiques des eucaryotes. L'analyse de la séquence N-terminale des protéines à motifs PPR indique qu'une forte proportion de ces protéines est probablement adressée aux organites (Small et al., 2004). D'après des prédictions bioinformatiques, 54% des protéines à motifs PPR sont adressées aux mitochondries et 19% aux chloroplastes chez *Arabidopsis thaliana* (Lurin et al., 2004). Quelques rares protéines à motifs PPR adressées au noyau ou adressées conjointement au noyau et aux mitochondries ont été décrites dans la littérature (Mili et Pinol-Roma, 2003 ; Ding et al., 2006). Le motif PPR est variable en séquence et en taille, ce qui permet de distinguer trois catégories de motifs PPR nommés P, L et S. Le motif P est commun à tous les eucaryotes, alors que les motifs S (pour « Short ») de 31 acides aminés et L (pour « Long ») d'au moins 35 acides aminés sont spécifiques des plantes (Lurin et al., 2004). La combinaison de ces motifs, au sein d'une protéine, permet de classer les protéines à motifs PPR en deux sous-familles : la sous-famille de type P contient des protéines

constituées uniquement de motifs P et la sous-famille PLS contient des protéines constituées des trois motifs P, L et S (Lurin et al., 2004 ; Rivals et al., 2006). De plus, une distinction entre les protéines de la sous-famille PLS peut se faire par la présence ou non d'extensions à l'extrémité C-terminale : certaines protéines PPR se terminent par un domaine DYW, avec une extension de 100 acides aminés en C terminal, ou par une région E, importante pour l'activité de la protéine (Lurin et al., 2004). Dans certaines protéines, les motifs PPR peuvent être associés à des motifs supplémentaires comme le motif SMR (pour « MutS related motif ») identifié notamment dans des protéines de réparation de l'ADN, le motif bHLH (pour « basic helix-loop-helix ») caractéristique des facteurs de transcription, et le motif RRM (pour « RNA recognition motif ») qui est un domaine de reconnaissance et de liaison aux ARN (pour revue, Saha et al., 2007). Il existe également des protéines PPR qui se terminent par un domaine DYW, avec une extension de 100 acides aminés en C terminal, ou par une région E importante

Les propriétés biochimiques et le mode d'action des protéines à motifs PPR sont très peu connus. Chaque motif PPR forme deux hélices alpha anti-parallèles, et la répétition en tandem des motifs PPR formerait une super hélice contenant une « rainure » capable d'interagir avec une cible nucléique. Plusieurs protéines à motifs PPR ont été décrites comme étant capables de se lier directement et spécifiquement à des cibles ARN (Meierhoff et al., 2003 ; Okuda et al., 2006 ; Gillman et al., 2007 ; Williams-Carrier et al., 2008 ; Pfalz et al., 2009), alors que peu d'exemples de protéines à motifs PPR capables de se lier à l'ADN ont été décrits dans la littérature (Ikeda et Gray, 1999). Les domaines mis en place par la répétition des motifs PPR ne s'apparentent à aucun motif catalytique de fonction connue. On sait néanmoins que les protéines à motifs PPR s'associent souvent à des protéines dans des complexes protéiques : chez *Arabidopsis*, HCF152 forme un homodimère dans les chloroplastes (Nakamura et al., 2003) ; dans le stroma des chloroplastes, PPR2 est présent dans un complexe protéique hétérogène (Williams et al., 2003) ; CRP1 se trouve dans un complexe composé de plusieurs sous unités (Fisk et al., 1999) et PPR592 a été retrouvé dans des fractions de gel filtration contenant des complexes de haut poids moléculaire (Gillman et al., 2007). La liaison PPR-ARN serait un pré requis pour l'établissement ou l'empêchement d'activités enzymatiques précises sur les ARN (Lurin et al., 2004). Les activités enzymatiques seraient assurées non pas par les protéines à motifs PPR, mais par d'autres protéines dont le recrutement sur l'ARN serait facilité par les protéines à motifs PPR. Pour les protéines à motifs PPR qui ont des extensions en C-terminal, il est possible que l'activité soit assurée par l'extension C-

terminale. Grâce à l'étude de mutants, on sait que les protéines à motifs PPR jouent un rôle crucial à toutes les étapes de l'expression des gènes des organites : la transcription, l'édition, la maturation et la traduction des ARN chloroplastiques et mitochondriaux (Schmitz-Linneweber et Small, 2008).

En 2000, lorsque la famille des protéines PPR est identifiée, trois gènes codants des protéines contenant ces motifs avaient déjà été étudiés : *CRP1* chez *Zea mays* (Fisk et al., 1999), *PET309* chez *Saccharomyces cerevisiae* (Manthey et McEwen, 1995), et *CYA-5* chez *Neurospora crassa* (Coffin et al., 1997). Ces trois gènes codent des protéines impliquées dans la maturation ou la traduction d'ARNm d'origine chloroplastique ou mitochondriale. Chez *Arabidopsis thaliana*, la protéine MRL1 est nécessaire à la production et/ou la stabilisation des ARNm *rbcL* en agissant très certainement comme une barrière aux dégradations exoribonucléolytiques (Johnson et al., 2010). Chez le maïs, PPR10 joue un rôle dans la stabilisation et l'accumulation de transcrits chloroplastiques. En effet, cette protéine se fixe sur des sites situés en 5' et 3' d'ARNm chloroplastiques et la liaison protéine-ARN semble stopper les activités exoribonucléolytiques aux extrémités de ces transcrits. (Pfalz et al., 2009). Chez *Chlamydomonas reinhardtii*, la protéine PPR MCA1 est nécessaire à la stabilisation de l'ARNm *petA* (Lown et al., 2001). D'autres protéines à motifs PPR sont associées aux clivages endoribonucléolytiques (Hashimoto et al., 2003) : le mutant *ppr2* d'*Arabidopsis* est affecté dans l'accumulation du complexe chloroplastique NAD(P)H déshydrogénase (NDH). Chez le sauvage, l'ARN polycistronique *rps12-rps7-ndhB* est clivé à différents sites pour obtenir les monocistrons. Le clivage entre *rps7* et *ndhB* nécessite deux événements de coupure indépendants dont l'un permet la formation de l'extrémité 5' de *ndhB*. Cet événement de coupure est absent chez le mutant *ppr2* (Hashimoto et al., 2003). Un autre exemple est celui de la protéine HCF152 impliquée dans le clivage endoribonucléolytique de l'ARN *psbH-petB* (Meierhoff et al., 2003). Des protéines à motifs PPR sont également associées à l'épissage des introns. Par exemple, la protéine PPR531 de *Physcomitrella patens* participe à l'épissage de l'intron du transcrit mitochondrial *clpP* (Hattori et al., 2007) alors que l'inactivation du gène nucléaire *OTP70* chez *Arabidopsis thaliana* entraîne des défauts dans l'épissage du transcrit chloroplastique *rpoC1* (Schmitz-Linneweber et al., 2006 ; de Longevialle et al., 2007 ; Chateigner-Boutin et al., 2011). Les protéines à motifs PPR peuvent également jouer un rôle dans l'édition des ARN. Par exemple, CRR4 (Kotera et al., 2005) et CRR21 (Okuda et al., 2007) sont nécessaires à l'édition de sites indépendants présents sur le transcrit chloroplastique *ndhD* d'*Arabidopsis* (Okuda et al., 2010). Des protéines à motifs PPR sont également impliquées dans la traduction des ARNm codés par les organites : dans

les chloroplastes de maïs, la protéine CRP1 est nécessaire pour la traduction de l'ARNm *petA* qui code une sous-unité du complexe cytochrome b6f ; ce complexe ne s'accumule pas chez les mutants *crp1* (Barkan et al., 1994 ; Fisk et al., 1999). De plus, CRP1 est nécessaire pour la traduction de l'ARNm *psaC* qui code une sous-unité du PS1 : il a été montré que CRP1 est associé à la région 5'UTR de l'ARNm *psaC* *in vivo* (Schmitz-Linneweber et al., 2005) et que, dans les mutants *crp1*, le complexe PS1 diminue en abondance (Barkan et al., 1994). Ces résultats indiquent que des protéines à motifs PPR, comme CRP1, peuvent influencer la traduction d'ARNm spécifiques par des interactions PPR-ARNm (Schmitz-Linneweber et al., 2005). Chez *Arabidopsis thaliana*, PPR336 est une protéine de la membrane interne mitochondriale impliquée dans la traduction d'ARNm mitochondriaux. Des analyses de fractions polysomiales en gradient de saccharose ont montré que PPR336 s'associe aux polysomes dans les mitochondries (Uyttewaal et al., 2008a ; Pusnik et al., 2007). PNM1 code une protéine à motifs PPR adressée aux mitochondries et au noyau, selon des expériences de fusions GFP et d'immunodétection. Dans les mitochondries, PNM1 est associé à des polysomes et il est probable que cette protéine joue un rôle dans la traduction d'ARNm mitochondriaux. Dans le noyau, PNM1 interagit avec le facteur de transcription TCP8 qui peut se lier au promoteur de PNM1 (Hammani et al., 2011)..

Des protéines à motifs PPR ont également été isolées grâce à leur capacité à se lier aux séquences promotrices actives de gènes chloroplastiques ou mitochondriaux. Chez le blé, la protéine PPR p63 joue un rôle dans la transcription du gène mitochondrial *cox2* : en effet, p63 a été caractérisée comme étant une protéine de liaison à l'ADN ayant une affinité pour le promoteur de *cox2*. De plus, p63 a été co-purifiée avec des régions promotrices mitochondriales et p63 augmente la quantité de transcrits *cox2* dans un système de transcription *in vitro* (Ikeda et Gray., 1999). Des protéines à motifs PPR particulières, détaillées dans le chapitre VI, jouent un rôle dans la suppression de l'expression de gènes mitochondriaux inducteurs de stérilité dans les systèmes de stérilité nucléo-cytoplasmique (Desloire et al., 2003; Gillman et al., 2007).

## IV.2 Coordination entre compartiments génétiques de la cellule végétale

### IV.2.1 La signalisation antérograde

Les protéines qui composent les sous-unités des complexes présents dans les organites sont codées par plusieurs génomes compartimentés, ce qui nécessite une expression coordonnée des gènes pour produire les concentrations correctes de complexes protéiques et pour maintenir la fonction des organites. Les protéines impliquées dans l'expression des gènes des organites (comme les facteurs de transcription et les protéines de liaison aux ARN de type protéines à motifs PPR) sont des protéines codées par le noyau (Jung et Chory, 2010, détails dans les chapitres II.3, III.3 et IV.1) qui peuvent nécessiter une expression coordonnée pour le développement et le maintien des organites.

Dans le noyau, l'expression de gènes liés à la photosynthèse est co-régulée avec la transcription de gènes chloroplastiques et de facteurs de traduction (Biehl et al., 2005). Les profils d'expression des gènes nucléaires et chloroplastiques qui codent des protéines ribosomiques chloroplastiques, montrent que les niveaux d'expression de ces gènes augmentent simultanément dans les deux compartiments lors du développement des graines chez *Nicotiana tabacum* (MacLean et al., 2008). Les mécanismes antérogrades coordonnent l'expression de gènes dans les organites en réponse à des stimuli environnementaux et endogènes qui sont perçus par le noyau (Woodson et Chory, 2008). Ces stimuli peuvent être des signaux de nature développementale, hormonale ou environnementale (Rodermel, 2001 ; Leister, 2005 ; Nott et al., 2006).

La différenciation des chloroplastes est également le résultat d'une régulation antérograde. Les profils protéiques du type sauvage et de mutants affectés dans le développement des chloroplastes chez *Arabidopsis* (Rutschow et al., 2008), ainsi que les profils protéiques à différents stades de développement des chloroplastes chez le riz (Kleffmann et al., 2007), le pois (Kanervo et al., 2008) et le maïs (Lonosky et al., 2004) ont été étudiés. Lors de la transition des étioplastés en chloroplastes par exemple, le niveau d'expression des gènes codant les protéines liées à la traduction des protéines plastidiques et à la photosynthèse, ainsi que la quantité de ces protéines, augmentent (Kleffmann et al., 2007 ; Lonosky et al., 2004).

### IV.2.2 La signalisation rétrograde

Des changements dans les activités métaboliques ou l'état fonctionnel des organites sont connus pour affecter l'expression de gènes nucléaires. Cette régulation fait intervenir la

signalisation rétrograde qui est importante pour redéfinir l'expression de gènes nucléaires en réponse à un état fonctionnel ou métabolique des chloroplastes (Woodson et Chory, 2008). La signalisation rétrograde permet aux cellules de moduler la signalisation antérograde (Butow et Avadhani, 2004 ; Pesaresi et al., 2007). Dans la littérature, plusieurs termes existent pour faire référence aux modifications de l'expression des gènes nucléaires suite à des changements dans les activités métaboliques ou l'état fonctionnel des organites : communication rétrograde, signalisation rétrograde ou régulation rétrograde (Rhoads et Subbiah, 2007). Il existe une signalisation rétrograde mitochondriale (SRM) et une signalisation rétrograde chloroplastique (SRC).

Les modifications de l'expression des gènes nucléaires peuvent se faire *via* la transmission de signaux produits par les organites. Ces signaux permettent d'informer le noyau de stress cellulaires, de communiquer l'état physiologique (de développement et de fonctionnement) des organites au noyau, et d'induire une réponse adaptative à la fluctuation des conditions de l'environnement (Woodson et Chory ; 2008). Pogson et al. (2008) classent les signaux rétrogrades en deux catégories. La première catégorie de signaux rétrogrades regroupe les signaux qui contrôlent le développement de la biogenèse des organites (contrôle biogénique). Les sous-unités des complexes protéiques et les cofacteurs ont besoin d'être présents en stoechiométrie correcte pour un assemblage précis : le contrôle biogénique permet de réguler la synthèse et l'assemblage des sous-unités protéiques présentes dans les organites. Le processus de « contrôle par épistasie de synthèse » de l'expression des gènes nucléaires repose sur le fait que la synthèse, par un compartiment, d'une sous-unité protéique est dépendante de la synthèse, par un autre compartiment, d'une ou plusieurs autres sous unités du même complexe (Wollman et al., 1999 ; Wostrikoff et al., 2004). Une fois que la capacité optimale de complexes protéiques est atteinte, les sous-unités codés par les organites ne sont plus synthétisées. La levée du blocage de synthèse peut avoir lieu si les sous-unités d'origine nucléaire sont de nouveau importées. De la même façon, les cofacteurs qui s'accumulent en excès par rapport à leurs protéines de liaison peuvent réguler à la fois leur propre synthèse, mais aussi celles des protéines auxquelles elles se lient (Pogson et al., 2008). Le contrôle par épistasie de synthèse a été démontré pour plusieurs sous unités des complexes du photosystème présents dans les chloroplastes de *Chlamydomonas*, mais jamais chez les plantes (Pogson et al., 2008). La deuxième catégorie de signaux rétrogrades regroupe les signaux qui réalisent des ajustements rapides du métabolisme énergétique en réponse à des contraintes environnementales et de développement (contrôle opérationnel) ; ceci afin de

maintenir une production optimale et pour limiter et réparer les dommages induits par les différents stress : le contrôle opérationnel permet principalement d'optimiser la performance photosynthétique des chloroplastes (Pogson et al., 2008) En réponse à l'état physiologique des chloroplastes, les niveaux de transcription des gènes nucléaires qui codent des facteurs antioxydants, des protéines de réponse aux stress, des protéines impliquées dans la mort cellulaire programmée et des protéines de la photosynthèse vont être régulés (Woodson et Chory, 2008).

La SRC est composée de plusieurs voies de signalisation induites par des stress, par la lumière et par des signaux de développement. L'état redox de la chaîne photosynthétique de transport des électrons, l'accumulation de Mg-protoporphyrine IX dans le chloroplaste (la Mg-protoporphyrine IX est un intermédiaire de la biosynthèse de chlorophylle) et l'inhibition de l'expression de gènes chloroplastiques semblent être des stimuli qui induisent la SRC ou des éléments de cette signalisation.

Les changements de l'état redox de la chaîne de transport des électrons et des changements de niveau de ROS ( $O_2$ ,  $O_2^-$  et  $H_2O_2$ ) semblent être des acteurs majeurs de la SRC soit directement, en agissant sur l'expression de gènes nucléaires, soit indirectement en activant des messagers de la SRC dont certains commencent à être connus (revue de Pesaresi et al., 2007 ; Jung et Chory., 2010). Les protéines EXECUTER (EX1 et EX2) sont à la fois dépendantes des niveaux d' $O_2^-$  et de  $H_2O_2$ , et impliquées dans la régulation de gènes nucléaires. Ces deux protéines jouent un rôle lors du transfert des signaux des chloroplastes vers le noyau (Lee et al., 2007).

L'existence d'une SRC dépendante de l'état des chloroplastes (contrôle biogénique) peut être observée suite à un traitement des plantes avec l'herbicide Norflurazon (NF). Le NF est un inhibiteur de la phytoène désaturase qui est une enzyme de la voie de biosynthèse des caroténoïdes. L'absence de caroténoïdes cause des dommages photo-oxydatifs sur les chloroplastes lorsque les plantes traitées sont exposées longtemps à la lumière blanche. Le traitement au NF entraîne une réduction importante de l'expression de gènes nucléaires, comme *LHCB* (pour *Light Harvesting complex of PSII*) et d'autres gènes impliqués dans la photosynthèse (Oelmüller et Mohr, 1986 ; Gray et al., 2003). Des mutants d'*Arabidopsis*, insensibles au traitement NF, ont été identifiés : les mutants *gun* (pour *genomes uncoupled*) (Susek et al., 1993). La majorité des mutants *gun* sont mutés dans des gènes qui codent des enzymes de la voie du tétrapyrrole. Cette voie de biosynthèse est à l'origine de la production

des hèmes et des chlorophylles a et b. Les études de génétique moléculaire réalisées chez *Chlamydomonas reinhardtii* et *Arabidopsis thaliana* suggèrent que des perturbations dans la voie du tétrapyrrole peuvent générer des signaux de la SRC. Pour identifier des gènes nucléaires sous le contrôle de la SRC, Koussevitzky et al. (2007) ont analysé, par biopuces, l'expression des gènes nucléaires chez des plantes sauvages et chez des mutants *gun* en présence du traitement NF. Dans un premier temps, ils ont identifié plus de 1000 gènes nucléaires différentiellement exprimés dans les mutants *gun* par rapport aux plantes sauvages dans ces conditions. Parmi eux, 330 gènes nucléaires sont exprimés dans les mutants *gun* et réprimés dans les plantes sauvages. En résumé, la perte de fonction des gènes *GUN* entraîne une augmentation de l'expression des gènes nucléaires associés à la photosynthèse lorsque le développement des chloroplastes est bloqué par le NF. La nature du signal rétrograde est sujet à controverse : d'après Strand et al., (2003), il peut s'agir d'un intermédiaire de la voie du tétrapyrrole, comme la Mg-protoporphyrine IX ; pour d'autres, la nature du signal rétrograde n'est pas la Mg-protoporphyrine IX mais plutôt un signal secondaire provenant du chloroplaste (Mochizuki et al., 2008). Il est également possible que chez les mutants *gun*, une perturbation de la voie de biosynthèse du tétrapyrrole puisse produire localement des ROS ou changer l'état rédox des chloroplastes (Moulin et al., 2008). Récemment, Woodson et al., (2011) ont présenté des résultats obtenus sur un mutant d'*Arabidopsis thaliana* (*gun6-1D*) qui surexprime une enzyme de la voie du tétrapyrrole : la ferrochélatase 1. L'expression des gènes nucléaires associés à la photosynthèse est plus élevée chez le mutant *gun6-1D* que chez le sauvage. Ces données suggèrent que l'hème, spécifiquement produit par la ferrochélatase 1, peut être utilisé comme un signal rétrograde permettant de coordonner l'expression des gènes nucléaires associés à la photosynthèse avec le développement des chloroplastes. L'équilibre rédox de la plastoquinone, coenzyme localisée dans la membrane des thylacoïdes et qui participe au transfert d'électrons du PSII au cytochrome b6f, a également été proposé comme point de départ de la SRC (Pfannschmidt et al., 1999). Mais d'autres résultats tendent à prouver que l'expression des gènes nucléaires associés à la photosynthèse n'est pas couplée au statut rédox de la plastoquinone (Frigerio et al., 2007), ce qui fait douter de l'implication du statut rédox de la plastoquinone dans une voie de SRC (Alboresi et al., 2011).

L'existence d'une SRC dépendante de l'état de développement des chloroplastes a pu également être observée chez le mutant *immutans* d'*Arabidopsis* (Aluru et al., 2009). Ce mutant a un phénotype de feuilles panachées avec des secteurs de feuilles qui sont albinos et d'autres qui sont verts. Les profils d'expression des gènes nucléaires ont été observés et comparés à partir de prélèvements réalisés dans les secteurs blancs (chloroplastes dont le

développement est arrêté) et dans les secteurs verts (chloroplastes sauvages). Les résultats montrent des changements d'expression significatifs des gènes nucléaires dans les secteurs blancs par rapport aux secteurs verts. Dans les secteurs blancs, l'expression des gènes nucléaires reliés à la photosynthèse est réprimée alors que celle des gènes reliés à la réponse aux stress et au métabolisme de l'énergie est induite (Aluru et al., 2009). Ces analyses indiquent que la SRC est dépendante de l'état de développement des chloroplastes, et que la SRC régule un grand nombre de gènes nucléaires (Jung et Chory, 2010).

Contrairement aux études qui portent sur la SRC dépendante de l'état de développement des chloroplastes, les études qui portent sur la SRC dépendante de l'état fonctionnel des chloroplastes utilisent des échantillons contenant des chloroplastes intacts de feuilles matures auxquelles on a appliqué des conditions environnementales particulières afin de modifier l'état fonctionnel des chloroplastes (forte lumière, lumière spécifique au photosystème I ou II, inhibiteurs du transport d'électrons) (Rossel et al., 2002 ; Fey et al., 2005; Adamiec et al., 2008). Une forte lumière va induire une modification de l'état rédox de tous les composants photosynthétiques du chloroplaste, ainsi que des changements de l'expression des gènes nucléaires (Rossel et al., 2002 ; Kleine et al., 2007 ; Adamiec et al., 2008). L'état rédox des chloroplastes est également modifié en utilisant des sources lumineuses qui vont exciter spécifiquement le photosystème I ou le photosystème II (Fey et al., 2005). L'utilisation d'inhibiteurs (Gray et al., 2003) de la traduction chloroplastique ou de mutants de gènes nucléaires affectant négativement la traduction dans le chloroplaste (Pesaresi et al., 2001 ; Pesaresi et al., 2006) peut aussi permettre de mettre en évidence des voies de signalisation de la SRC. Dans ces cas, l'inhibition de la traduction chloroplastique va avoir pour conséquence de réprimer des gènes nucléaires impliqués dans la photosynthèse (Gray et al., 2003).

Les changements d'état physiologique des mitochondries peuvent être induits par des stress (de nature biotique et abiotique), par des signaux développementaux, mais également par des mutations nucléaires ou mitochondriales (Kushnir et al., 2001 ; Gomez-Casati et al., 2002 ; Rhoads et Subbaiah, 2007). Ces changements d'état physiologique des mitochondries vont induire des stimuli en direction du noyau. Le nombre de stimuli identifiés dans la SRM est plus faible que dans le cas de la SRC. Parmi ces stimuli, existent les ROS comme l'O<sub>2</sub>, H<sub>2</sub>O<sub>2</sub> et l'O<sub>2</sub><sup>-</sup>. (Rhoads et Subbaiah, 2007). Dans les mitochondries végétales, l'inhibition de la voie du cytochrome entraîne une production de ROS (Maxwell et al., 1999 ; Yao et al., 2002) et une modification de l'expression des gènes nucléaires (Saisho et al., 1997 ; Yu et al., 2001 ; Karpova et al., 2002 ; Maxwell et al., 2002). Les chaînes de transport des électrons des

mitochondries sont des sites majeurs de production de ROS. L'état physiologique des organites va avoir une incidence sur les niveaux des ROS dans la cellule. Les ROS peuvent être utiles en tant que composants de la signalisation intracellulaire. La formation de ROS par les mitochondries existe en conditions de respiration normale, mais cette formation peut être augmentée en réponse à des conditions de stress (Rhoads et al., 2006). Les niveaux de ROS vont agir sur l'expression de gènes nucléaires, notamment par l'intermédiaire de voies de signalisation dont les acteurs sont très mal connus.

En réponse à l'état physiologique des mitochondries, les niveaux de transcription des gènes nucléaires qui codent des facteurs de transcription impliqués dans la respiration, la biogenèse des peroxisomes et les réponses aux stress oxydatifs vont être régulés (Liu et Butow, 2006 ; Butow et Avadhani, 2004 ). Les mitochondries, et plus précisément l'AOX et le complexe IV cytochrome oxydase, utilisent de l'oxygène. Dans les cellules végétales, les variations de la disponibilité en oxygène sont ressenties en premier par les mitochondries entraînant une régulation rétrograde de gènes nucléaires (Rhoads et Subbaiah, 2007). Les mécanismes de détection de l'oxygène, ainsi que le rôle des mitochondries dans les voies de signalisation associées à l'oxygène sont peu connus chez les plantes (Bailey-Serres et Chang, 2005). Il a été montré, dans un système de culture cellulaire de maïs, qu'une privation en oxygène déclenche une augmentation immédiate et réversible de la concentration en  $\text{Ca}^{2+}$  cytosolique (Subbaiah et al., 1994). Un marqueur spécifique du  $\text{Ca}^{2+}$  mitochondrial a ensuite confirmé que l'augmentation du  $\text{Ca}^{2+}$  cytosolique était due à du  $\text{Ca}^{2+}$  d'origine mitochondriale. De plus, en état de privation d'oxygène, l'utilisation du rouge de ruthénium, un composé qui bloque les canaux  $\text{Ca}^{2+}$  mitochondriaux, abolit les changements de concentration en  $\text{Ca}^{2+}$  cytosolique observé lors d'une privation en oxygène (Subbaiah et al., 1994a,b). Une élévation de la quantité de ROS mitochondriaux lors de l'inhibition de l'activité AOX semble être un déclencheur de la libération de  $\text{Ca}^{2+}$  mitochondrial dans le cytosol (Rhoads et Subbaiah ; 2007). Comment une augmentation de la concentration en  $\text{Ca}^{2+}$  cytosolique peut-elle induire des changements de l'expression des gènes nucléaires ? Il semblerait que des protéines kinases dépendantes du calcium, des facteurs de transcription et des protéines de régulation comme les membres de la famille 14-3-3 (Lu et al., 1994 ; Trewavas, 1999 ; Sanders et al., 2002 ; Hardingham et al., 1997) puissent à la fois se lier au  $\text{Ca}^{2+}$  cytosolique et réguler l'expression de gènes nucléaires. Une autre hypothèse serait que le  $\text{Ca}^{2+}$  cytosolique lui-même pourrait diffuser dans le noyau. Des études d'imagerie montrent qu'une augmentation de la concentration en  $\text{Ca}^{2+}$  cytoplasmique s'accompagne d'une augmentation de la concentration en  $\text{Ca}^{2+}$  nucléaire en condition d'anoxie (Subbaiah et al., 1994a,b). Chez Arabidopsis, des

analyses de transcriptome et de protéome indiquent que les transcrits et les protéines des enzymes du cycle de Krebs, ainsi que des composants de la chaîne de phosphorylation oxydative, sont régulés à la baisse suite à une privation des plantes en oxygène. Après réoxygénation, il y a un retour à la normale (c'est-à-dire identique à celui existant avant le stress anoxique) du niveau d'expression des gènes nucléaires (Gonzali et al., 2005 ; Huang et al., 2005). En anoxie, la concentration en  $\text{Ca}^{2+}$  est donc une composante de la SRM essentielle chez de nombreuses espèces dont Arabidopsis, le riz et l'orge (Nie et al., 2006 ; Sedbrook et al., 1996 ; Tsuji et al., 2000). En résumé, la privation en oxygène pourrait inhiber l'activité de complexes mitochondriaux, ce qui déclencherait une élévation de la quantité de ROS d'origine mitochondriale et la libération de  $\text{Ca}^{2+}$  mitochondrial dans le cytosol. Ces changements pourraient induire une modification de l'expression de gènes nucléaires liés au fonctionnement des chloroplastes. Chez les végétaux, la concentration en  $\text{Ca}^{2+}$  semble être une composante de la SRM également lors d'autres stress comme le stress oxydatif (Evans et al., 2005 ; Price et al., 1994 ; Rentel et Knight, 2004), le stress osmotique et le stress « froid » (Logan et Knight, 2003). Le stress oxydatif se déclare lorsqu'il y a un fort déséquilibre entre la production de ROS et le système de défense contre les ROS, entraînant des dommages cellulaires (Halliwell et Gutteridge, 1999). Le stress oxydatif peut entraîner des dysfonctionnements mitochondriaux chez les plantes (Moller, 2001). Les enzymes impliquées dans le cycle de Krebs et les composants de la chaîne mitochondriale de transport des électrons (AOX et composants permettant la phosphorylation oxydative) sont particulièrement sensibles aux dommages oxydatifs. On sait que l'inhibition de la chaîne mitochondriale de transport des électrons chez des mutants nucléaires *mrrd* (pour *mitochondrial retrograde regulation deficient*), incapables d'induire la production de l'AOX, entraîne des dommages cellulaires causés par les ROS (Zarkovic et al., 2005). Le complexe AOX a la capacité de réduire la production mitochondriale de ROS (Rhoads et al., 2006). L'inhibition de la chaîne mitochondriale de transport des électrons permet de mimer des dysfonctionnements mitochondriaux. L'antimycine A (AA) et le myxothiazol inhibent le complexe III, alors que le cyanure (KCN) et l'azide inhibent le complexe IV (COX). Cette inhibition induit une augmentation de la production de ROS dans la cellule (comme c'est le cas lors des stress abiotiques et biotiques). La présence d'une forte concentration de ROS, globalement néfaste pour la cellule, va induire une augmentation de l'expression des gènes nucléaires codant l'AOX (Ito et al., 1997). L'inhibition du complexe III par l'AA induit l'expression de gènes nucléaires codant l'AOX chez le tabac (gène *NtAox1*), le soja (gène *GmAox1*), le maïs (gènes *ZmAox1*, *ZmAox2* et *ZmAox3*) et chez Arabidopsis (gène *AtAox1a*)

(Clifton et al., 2005 ; Djajanegara et al., 2002 ; Karpova et al., 2002 ; Saisho et al., 1997 ; Vanlerberghe et McIntosh, 1994).

## V. Coévolution nucléo-cytoplasmique

### V.1 Origine évolutive des génomes cytoplasmiques et nucléaire

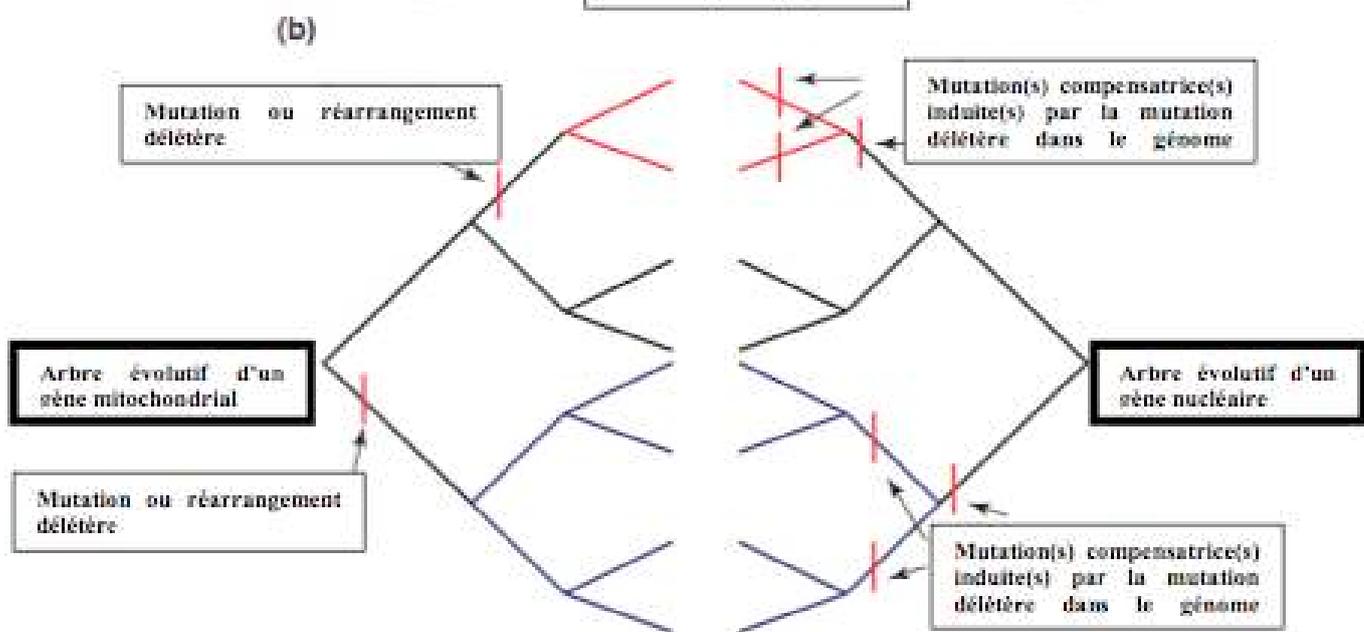
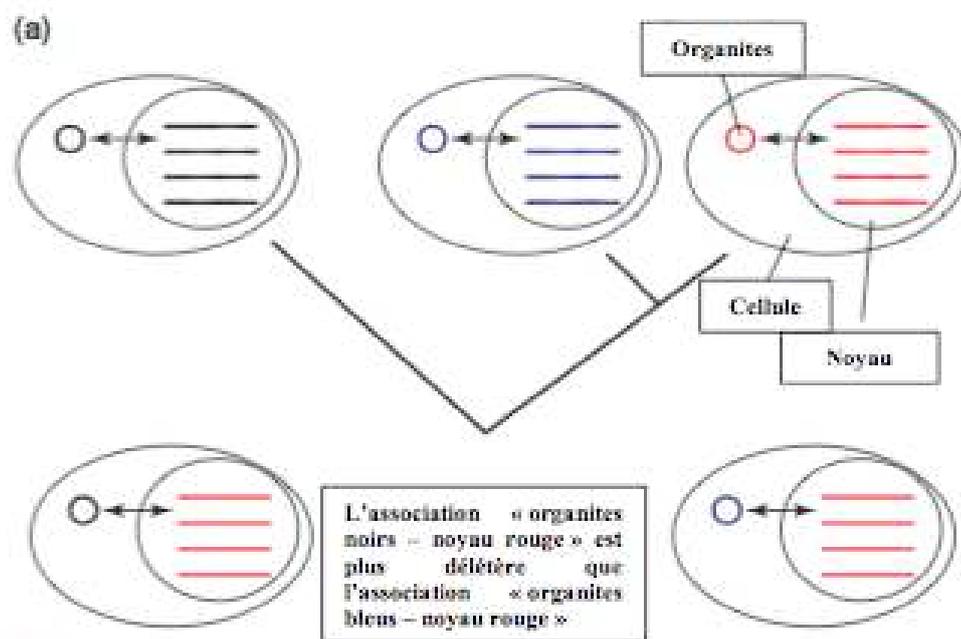
Futuyma a défini la coévolution comme un « changement génétique réciproque dans des espèces en interaction, en raison de la sélection naturelle imposée par chacune des espèces sur l'autre » (Futuyma et al., 1998). La coévolution nucléo-cytoplasmique peut être définie de la même manière en remplaçant « espèces » par « compartiments génétiques » (Rand et al., 2004). Comme nous l'avons vu dans le chapitre I.2, deux évènements distincts d'endosymbiose, l'un impliquant une alpha-protéobactérie et l'autre impliquant une cyanobactérie, sont à l'origine de la cellule végétale que nous connaissons aujourd'hui. Au cours de l'intégration génétique (ou transfert directionnel de l'information génétique), des changements génétiques dans le génome de la cellule hôte et dans le génome des ancêtres bactériens sont apparus. Un des changements génétiques les plus marquants dans la transition des ancêtres bactériens vers les organites est une réduction de la capacité codante des génomes cytoplasmiques (pour revue, Timmis et al., 2004). La capacité codante des chloroplastes et des mitochondries actuels (d'une centaine et d'une cinquantaine de protéines respectivement chez les végétaux) correspondrait à 5% de la capacité codante des génomes des ancêtres bactériens ancestraux estimée à 3000-7000 protéines (Kaneko et al., 1996 ; Wu et al., 2004 ; Gray et al., 1999). Les génomes des alpha-protéobactéries actuelles *Mesorhizobium loti* (7 Mb) et *Bradyrhizobium japonicum* (9,1 Mb) codent plus de 6700 et 8300 protéines respectivement. Les génomes des cyanobactéries actuelles *Nostoc PCC 7,200* (6,4 Mb) et *Nostoc punctiforme* (9Mb) codent plus de 5400 et 7200 protéines respectivement (Martin et al., 2002). Aujourd'hui, le génome mitochondrial d'*Arabidopsis thaliana* ne code que 4% de la quantité de protéines codées par le génome de *Rickettsia prowazekii*, une alpha-protéobactérie (Martin et al., 2002).

Cette réduction importante de la capacité codante des génomes cytoplasmiques peut s'expliquer par le transfert de gènes des génomes des endosymbiotes vers le génome de la cellule hôte et/ou par perte de gènes redondants (avec ceux du génome de la cellule hôte) ou non fonctionnels (Gray et al., 1999). En effet, on sait que chez les eucaryotes, de nombreuses protéines sont codées par le génome nucléaire et adressés aux chloroplastes et aux

mitochondries (Bogorad, 1975 ; Ellis, 1981). De plus, chez *Arabidopsis*, 18% du protéome nucléaire a une origine cyanobactérienne (Martin et al., 2002) et les génomes nucléaires des eucaryotes contiennent des séquences d'ADN d'origine mitochondriale (Ricchetti et al., 1999 ; Mourier et al., 2001). Cette intégration génétique a forgé les génomes nucléaires et cytoplasmiques des eucaryotes actuels. Chez les angiospermes, des évènements de transfert fonctionnels de gènes peuvent encore être observés avec, par exemple, un transfert de gène chloroplastique vers le noyau pour 16 000 gamètes chez *Nicotiana tabacum* (Huang et al., 2003). Le gène mitochondrial *rps10* est connu pour avoir été transféré plusieurs fois vers le noyau, de façon indépendante (Adams et al., 2000 ; 2003). Il en est de même pour le gène *cox2* transféré récemment dans le génome nucléaire de plusieurs espèces végétales (Huang et al., 2004). Si le transfert d'un gène d'une mitochondrie ou d'un chloroplaste vers le noyau devient fonctionnel, il peut y avoir une redondance génétique (une copie du gène est présente dans l'organite, et une autre dans le noyau), d'où la perte d'une pression de sélection en faveur du maintien par l'organite de la copie génétique correspondante. L'évolution réductive des génomes des organites pourrait s'expliquer au moins en partie par ce processus de transfert-redondance-suppression. La nature du vecteur permettant ce transfert de l'information génétique est sujette à discussions et peut se réaliser *via* des molécules d'ADN ou d'ARN (Dyall et al., 2004 ; Huang et al., 2004 ; Timmis et al., 2004). Enfin, une copie génétique transférée dans le génome nucléaire doit, pour devenir fonctionnelle, acquérir différents éléments génétiques comme des signaux d'expression et une région codant un peptide d'adressage vers les organites (Adams et Palmer 2003).

## V.2 Co-adaptation des génomes cytoplasmiques et nucléaire

La co-adaptation est définie par Blier et al (2001) de la manière suivante : « sélection favorisant l'accumulation, dans un pool génétique, de gènes en interactions harmonieuses. Les forces sélectives qui agissent sur l'un de ces gènes vont également affecter les partenaires qui interagissent avec ce gène ». Au niveau moléculaire, la co-adaptation a pour résultat que des mutations dans un gène A, qui code une protéine en interaction avec une protéine codée par un gène B, seront compensées par des mutations dans le gène B (Rand et al., 2004). Les contraintes fonctionnelles et structurales des protéines en interaction, codées par des génomes différents, peuvent théoriquement avoir une influence sur la co-adaptation des génomes en faisant appel à une évolution darwinienne positive appliquée aux interactions (Rand et al., 2004). Par exemple, le système mitochondrial de transport des électrons contient des



### Figure A8 : modèle de la coadaptation nucléocytoplasmique

La co-adaptation a lieu lorsque la sélection naturelle agit sur l'évolution des gènes en interaction. (a) Des complexes de gènes co-adaptés peuvent être perturbés lorsque les organites d'une lignée sont mis en présence d'un fond nucléaire provenant d'une lignée différente. Le degré de perturbation peut dépendre du temps de divergence entre les deux lignées. (b) Au niveau moléculaire, ce modèle de coadaptation nucléocytoplasmique prédit qu'une mutation ou un réarrangement délétère dans l'un des gènes en interaction sera compensé par une ou plusieurs mutations compensatrices dans le ou les autres gènes en interaction. Ce schéma représente l'évolution de deux gènes coadaptés, celui d'un gène mitochondrial (à gauche) et celui d'un gène nucléaire, pour différents lignages (d'après Rand et al., 2004).

complexes dont les sous-unités protéiques sont codées par les génomes mitochondrial ou nucléaire. Pour maintenir une fonction physiologique optimale des complexes protéiques, la sélection naturelle va continuellement favoriser une co-adaptation entre les protéines en interaction et donc sélectionner les gènes qui codent des protéines les plus à même d'interagir entre elles de façon optimale. Le fort degré d'interaction entre les protéines codées par les gènes nucléaires et cytoplasmiques peut nécessiter une évolution coordonnée en réponse à des conditions de l'environnement (Rand et al., 2004). Enfin, les différences d'hérédité entre les organites (hérédité maternelle) et le noyau (hérédité biparentale) pourraient générer les conditions pour une évolution cytoplasmique « égoïste » qui entraînerait une coévolution nucléaire (Hurst et al., 1996 ; Birky, 2001 ; Burt et Trivers, 2006 ; Barr et Fishman, 2009).

### V.3. Modèle de Bateson-Dobzhansky-Muller adapté aux interactions nucléocytoplasmiques

Dans les années 1930, Dobzhansky et Muller proposent un modèle d'incompatibilité entre gènes nucléaires pour expliquer les barrières des flux de gènes qui apparaissent dans la descendance de croisements interspécifiques. D'après eux, les létalités et les stérilités observées chez certains hybrides sont causées par des interactions épistatiques négatives entre deux ou plusieurs locus nucléaires qui ont fonctionnellement divergé entre les deux espèces utilisées pour obtenir les hybrides (Dobzhansky, 1937 ; Muller, 1942). Les interactions de type Dobzhansky-Muller semblent être une source importante d'incompatibilité hybride dans beaucoup de systèmes. À part pour quelques systèmes bien étudiés notamment chez la drosophile (Presgraves et al., 2003) et Arabidopsis (pour revue, Bomblies et Weigel, 2010), la fonction et l'histoire évolutive des gènes impliqués dans les interactions de type Dobzhansky-Muller sont mal connus (Orr et Presgraves, 2000).

Certaines incompatibilités nucléocytoplasmiques peuvent être considérées comme des formes particulières d'incompatibilités de type Dobzhansky-Muller causées par des interactions inappropriées entre les génomes nucléaire et cytoplasmiques (Chou et al., 2010). D'après les prédictions du modèle de Dobzhansky-Muller adapté aux interactions nucléocytoplasmiques (figure A8), l'interaction fonctionnelle entre des produits de gènes co-adaptés peut être perturbée lorsque qu'un organite d'une lignée est mis en présence d'un fond nucléaire provenant d'une autre lignée. Le degré de perturbation observé chez les hybrides serait fonction du temps de divergence entre les deux lignées parentales (Rand et al., 2004). Chez

les animaux, les forts taux de mutation des génomes mitochondriaux (Wolfe et al., 1987 ; Palmer et Herbon ; 1998), couplés à des faibles populations effectives ( $N_e$ ) (Wright 1931,1938) peuvent permettre à des cytotypes légèrement délétères d'augmenter en fréquence dans la population (Birky, 2001). Chez les plantes, l'ADN mitochondrial est plus stable en séquence, mais des réarrangements peuvent également induire des modifications fonctionnelles. En théorie, l'apparition de ces cytotypes favoriserait des allèles nucléaires ayant des effets compensateurs sur les effets délétères induits par les mutations mitochondriales (Rand et al., 2004). Une mutation délétère dans un locus A établirait une pression de sélection positive pour la fixation d'une mutation compensatrice dans le locus B en interaction avec le locus A. Les gènes nucléaires, plus à même de muter que les gènes des organites, semblent être plus susceptibles de contribuer aux mutations compensatrices que les gènes des organites. L'accumulation de mutations légèrement délétères dans les gènes mitochondriaux d'une population peut donc être toléré grâce à l'apparition de mutations compensatrices dans les gènes nucléaires qui codent des protéines interagissant avec les protéines codées par les gènes mitochondriaux (Rand et al., 2004). Ce modèle pourrait expliquer le maintien d'états co-adaptés entre les génomes de la cellule eucaryote. (Rand et al., 2004).

Plusieurs méthodes permettent de rompre la co-adaptation entre les compartiments cytoplasmiques et nucléaire. Dans le cas où les individus A et B sont capables de se reproduire sexuellement par des croisements réciproques et que l'hérédité des organites est maternelle, il est possible de créer des hybrides réciproques en F1 ou des lignées alloplasmiques. Dans le cas des hybrides réciproques en F1 : la première génération obtenue en utilisant le parent A comme femelle et le parent B comme mâle sera hétérozygote A/B et aura hérité des organites du parent A. Dans la F1 obtenue grâce au croisement réciproque (parent B comme femelle et parent A comme mâle), les individus seront également hétérozygotes A/B mais auront cette fois-ci hérité des organites du parent B. Ainsi, la comparaison des phénotypes obtenus dans les deux F1 permettra d'observer des plantes qui possèdent le même matériel génétique nucléaire, mais un matériel génétique cytoplasmique différent. À noter que la création de F2 réciproques (obtenues suite à une génération d'autofécondation des deux F1 réciproques) est souvent nécessaire pour créer des ruptures de co-adaptation : une seule copie allélique coadaptée au compartiment cytoplasmique est souvent suffisante pour maintenir une coadaptation nucléocytoplasmique fonctionnelle ; certaines F2 étant homozygotes, l'absence de copie coadaptée au compartiment

cytoplasmique permettra de révéler une rupture de co-adaptation. Des rétrocroisements récurrents avec le parent mâle peuvent produire des lignées alloplasmiques lorsque les deux parents originels sont d'espèces différentes (par exemple, des rétrocroisements récurrents avec le parent B vont finalement donner des lignées dont le cytoplasme sera celui du parent A et le noyau, celui du parent B). Si les deux parents sont de la même espèce, on ne parle plus de lignées alloplasmiques mais de cytolignées (Budar et Roux, 2011).

Pour créer des ruptures de co-adaptation à l'aide de deux individus A et B qui ne sont pas capables de se reproduire sexuellement, la création de cybrides, ou hybrides cytoplasmiques, peut contourner ce problème. Les cybrides peuvent être obtenus par fusion de protoplastes. Si deux génotypes différents A et B (possédant des cytoplasmes et des noyaux différents) sont considérés, en l'absence de fusion nucléaire, un cybride aura le noyau de A ou de B, dans un cytoplasme qui est hybride entre A et B (Pelletier et al., 1985)

#### V.4 Exemples de rupture de co-adaptation nucléo-cytoplasmique chez les eucaryotes

En 1929, Renner décrit des phénotypes de pigmentation aberrants chez des hybrides interspécifiques du genre *Oenothera* (Renner, 1928 ; Stubbe, 1989). Des études similaires réalisées sur d'autres plantes ont montré que des hybrides, des lignées alloplasmiques ou des cybrides pouvaient présenter des phénotypes anormaux ayant pour origine des incompatibilités nucléocytoplasmiques. Ces phénotypes comprennent des déficiences chlorophylliennes, des modifications morphologiques des fleurs, des stérilités et des létalités.

Aujourd'hui, les incompatibilités nucléocytoplasmiques de type Bateson-Dobzhansky-Muller ont été observées en interspécifique non seulement chez beaucoup de plantes, mais aussi chez les primates (Barrientos et al., 1998), les amphibiens (Liepins et Hennen, 1977), les mouches (Sackton et al., 2003), les guêpes (Ellison et al., 2008) et les levures (Spirek et al., 2000) ; et en intraspécifique, chez des crustacés copépodes marins (Ellison et Burton, 2006) et chez le pois (Bogdanova et al., 2009). En effet, un cas de barrière génétique résultant d'une rupture de co-adaptation entre les compartiments mitochondrial et nucléaire a été décrit chez une espèce de crustacé (Ellison et Burton, 2006), et un cas de rupture de co-adaptation entre les compartiments chloroplastique et nucléaire a été observé chez *Pisum sativum* (Bogdanova et al., 2009).

Une co-adaptation nucléo-cytoplasmique a été mise en évidence par l'équipe de Burton lors

de croisements entre populations de crustacés copépodes marins (*Tigriopus californicus*) (Burton et al., 1999 ; Willett et Burton, 2003). Ces populations de *Tigriopus californicus* isolées géographiquement ont de forts niveaux de divergence à plusieurs locus nucléaires et mitochondriaux (Burton et Lee, 1994). Lors de croisements inter populationnels, les générations F1 sont en général vigoureuses. Par contre, les générations F2 présentent fréquemment une réduction de la performance de plusieurs caractères reliés à la valeur sélective (Burton, 1990 ; Edmands, 1999). Ces réductions sont attribuées à des anomalies d'une ou de plusieurs fonctions cellulaires qui peuvent avoir des conséquences sur la physiologie des individus. Selon le modèle de mutations compensatrices proposé par Rand et al., (2004), les hybrides F1 contiennent une copie des gènes nucléaires coadaptés aux variants mitochondriaux, ce qui évite une perte de la valeur sélective. Dans certaines F2, le manque de mutations compensatrices au niveau des gènes nucléaires entraînerait une réduction de la valeur sélective chez ces hybrides (Ellison et Burton, 2008). Parmi les populations de copépodes, de forts niveaux de substitutions en acides aminés ont été identifiés dans des protéines du complexe IV mitochondrial de transport des électrons (cytochrome c oxydase) : deux de ces protéines sont codées par les gènes mitochondriaux *coxI* et *coxII* (Burton et Rawson, 1999 ; Rawson et Burton, 2002). Par des rétrocroisements répétés, les gènes nucléaires d'une population ont été introgressés dans le fond cytoplasmique de l'autre population. Ces combinaisons nucléocytoplasmiques ont montré une réduction de l'activité du complexe IV, probablement due à des interactions suboptimales entre les sous unités COX codées par les génomes mitochondriaux et les autres sous unités du complexe codées par le génome nucléaire (Edmands et Burton ; 1999). L'équipe de Burton a également trouvé des variants de la protéine, codée par le noyau, qui compose le cytochrome c. Ces variants ont été isolés à partir de deux populations différentes de *Tigriopus californicus*. Des résultats de reconstitution *in vitro* de l'activité cytochrome c oxydase indiquent que ces variants ont une plus grande activité lorsqu'ils sont en présence des variants du complexe IV qui proviennent de la même population (Rawson et Burton, 2002). Ces résultats suggèrent que les gènes nucléaires et cytoplasmiques qui codent les protéines du système de transport des électrons, ont co-évolué au sein de chacune des populations et que les nouvelles combinaisons nucléocytoplasmiques ont créé des incompatibilités fonctionnelles qui peuvent avoir des conséquences sur la valeur sélective des hybrides. Les pertes de valeur sélective observées dans les hybrides de *Tigriopus californicus* ont finalement été définitivement attribuées à des interactions génétiques entre les compartiments mitochondriaux et nucléaire (Ellison et Burton, 2008).

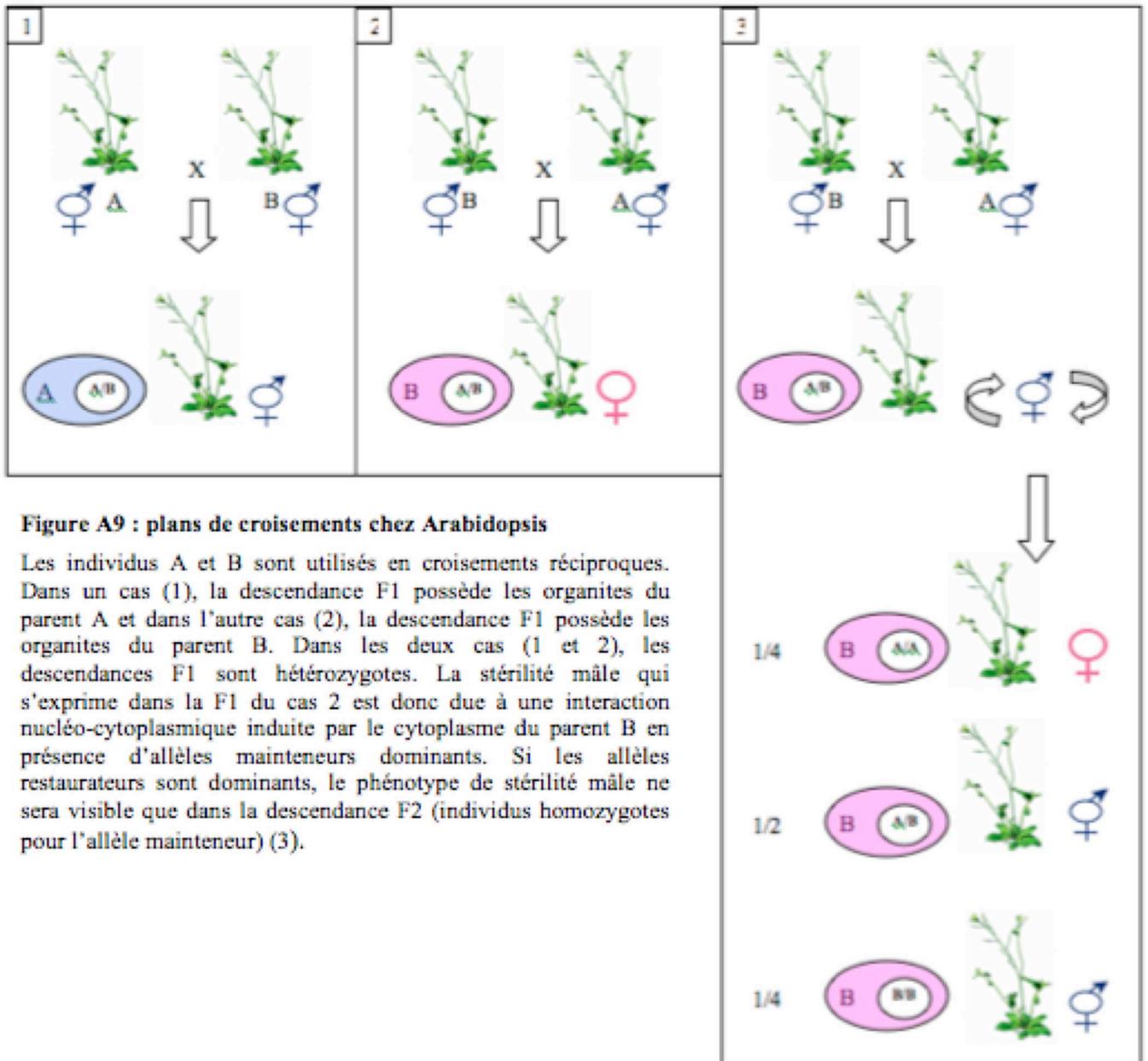
Il existe également un cas de rupture de co-adaptation entre les compartiments chloroplastique et nucléaire révélé par des croisements entre sous espèces de *Pisum sativum*. En croisant la sous espèce sauvage d'origine palestinienne *elatius*, utilisée comme parent femelle avec la sous espèce domestiquée *sativum*, utilisée comme parent mâle, Bogdanova et al (2009) ont obtenu une descendance F1 présentant une forte stérilité, un déficit en chlorophylle et une réduction des stipules. La F1 réciproque semble normale. Le phénotype de déficience chlorophyllienne fait penser à un dysfonctionnement des chloroplastes. De plus, les phénotypes associés à l'incompatibilité disparaissent lors de la transmission biparentale accidentelle des chloroplastes dans le croisement, indiquant que l'incompatibilité a une origine nucléo-chloroplastique (Bogdanova, 2007). La cartographie semble indiquer que deux locus nucléaires non liés provenant de *sativum* sont responsables, en présence du cytoplasme d'*elatius*, des phénotypes observés en F1 (Bogdanova et al., 2009).

La majorité des incompatibilités nucléocytoplasmiques de type Dobzhansky-Muller ont été observées à la suite de croisements interspécifiques. Deux lignées de *Drosophila simulans* (*siI* et *siII*), et une lignée de *Drosophila mauritiana* (*maI*) ont été utilisées pour produire des génotypes utiles à l'étude des coadaptations nucléocytoplasmiques. Les deux lignées de *Drosophila simulans* diffèrent par leur ADN mitochondrial. Les ADN mitochondriaux de *siI* et *siII* ont été respectivement introgressés, par rétrocroisements, dans un fond nucléaire *siII* et *siI*, ainsi que dans un fond nucléaire *maI*. L'activité de la cytochrome c oxydase COX a ensuite été mesurée dans ces différents génotypes. Les résultats montrent que l'activité de COX est réduite chez les génotypes possédant l'ADN mitochondrial de *Drosophila simulans* en présence des allèles nucléaires de *Drosophila mauritiana*. Plusieurs hypothèses sont émises pour expliquer les perturbations de l'activité COX chez ces génotypes. L'assemblage du complexe COX et/ou la coadaptation entre le complexe COX et le cytochrome c, peuvent être altérés chez ces génotypes : les protéines COX constituées de sous unités nucléaires et mitochondriales d'origines différentes sont moins efficaces pour transférer les électrons du cytochrome c à l'oxygène. Une hypothèse alternative propose que la diminution d'activité de COX ait pour origine une diminution de la densité de ces complexes, ce qui sous-entend que la biogenèse des complexes COX soit influencée par les interactions entre les sous-unités codées par le noyau et par les mitochondries. Les tests d'activité des complexes COX sont dépendants de la densité des enzymes COX dans la membrane interne mitochondriale, de l'aire de la surface de ces membranes, et de la densité des mitochondries (Sackton et al.,

2003).

Peu de cas de coadaptation nucléocytoplasmique ont été analysés au niveau moléculaire. Une incompatibilité nucléocytoplasmique, à l'origine d'une isolation reproductive, a été mise en évidence chez des hybrides obtenus par croisements entre deux espèces de levures *Saccharomyces bayanus* et *Saccharomyces cerevisiae* (Lee et al., 2008). Ces hybrides présentent un défaut de la respiration, ce qui entraîne une incapacité à sporuler. La construction de lignées hybrides dans lesquelles un ou deux chromosomes proviennent de *Saccharomyces bayanus*, et les autres chromosomes de *Saccharomyces cerevisiae*, a permis d'identifier des facteurs génétiques responsables de la stérilité des hybrides F2. Un hybride a particulièrement été étudié : il contient le chromosome 13 de *Saccharomyces bayanus*, les autres chromosomes et les mitochondries de *Saccharomyces cerevisiae*. Le gène *AEP2* de *Saccharomyces bayanus*, localisé sur le chromosome 13, code une protéine nécessaire à la traduction d'un gène mitochondrial (*OLII*) codant une sous-unité du complexe ATP synthase (Finnegan et al., 1991). L'interaction de la protéine, AEP2, avec la région 5' UTR de l'ARNm *OLII*, est nécessaire à la traduction de cet ARNm. Une analyse de séquence a montré que les régions 5'UTR du gène *OLII* ont fortement divergé entre les deux espèces de levure. L'incompatibilité est donc probablement due à l'incapacité de la protéine AEP2 de *Saccharomyces bayanus* à reconnaître la région 5'UTR de l'ARNm *OLII* codé par *Saccharomyces cerevisiae* (Lee et al., 2008).

Schmitz-Linneweber et al. (2005) ont analysé des hybrides et des cybrides obtenus par fusion de cellules somatiques provenant de *Nicotiana tabacum* et d'*Atropa belladonna* (ces deux espèces de la famille des Solanaceae sont sexuellement incompatibles). Les cybrides Ab(Nt) ont un génome nucléaire provenant d'*Atropa belladonna* et des chloroplastes provenant de *Nicotiana tabacum*. Les plantes Ab(Nt) ont un phénotype albinos (Kushnir et al., 1991). Schmitz-Linneweber et al. (2005) ont montré que les défauts de pigmentation des plantes Ab(Nt) étaient dus principalement à un défaut d'édition de l'ARNm codant la sous-unité alpha de l'ATPase chloroplastique. Un facteur nucléaire, nécessaire à l'édition du transcrit *atpA* codé par le génome chloroplastique de *Nicotiana tabacum*, est absent du génome nucléaire d'*Atropa belladonna*. Chez *Atropa belladonna*, le site n'est pas édité. Chez les plantes Ab(Nt), le défaut d'édition du site 264 du transcrit *atpA* est principalement responsable du phénotype observé. Néanmoins, d'autres défauts d'édition peuvent avoir un rôle mineur dans l'incompatibilité observée chez les cybrides Ab(Nt) (Schmitz-Linneweber et al., 2005).



**Figure A9 : plans de croisements chez Arabidopsis**

Les individus A et B sont utilisés en croisements réciproques. Dans un cas (1), la descendance F1 possède les organites du parent A et dans l'autre cas (2), la descendance F1 possède les organites du parent B. Dans les deux cas (1 et 2), les descendance F1 sont hétérozygotes. La stérilité mâle qui s'exprime dans la F1 du cas 2 est donc due à une interaction nucléo-cytoplasmique induite par le cytoplasme du parent B en présence d'allèles mainteneurs dominants. Si les allèles restaurateurs sont dominants, le phénotype de stérilité mâle ne sera visible que dans la descendance F2 (individus homozygotes pour l'allèle mainteneur) (3).

## VI. Conflits génomiques entre compartiments nucléaire et cytoplasmiques

### VI.1 Les stérilités mâles nucléocytoplasmiques (SMNC)

Le conflit génomique peut être défini comme une situation où la sélection a des effets antagonistes sur les compartiments génétiques. La différence d'hérédité entre les compartiments nucléaire et cytoplasmiques peut provoquer un conflit génomique concrétisé chez les plantes par les stérilités mâles nucléocytoplasmiques (ou SMNC). La SMNC est définie comme une incapacité héritée de façon maternelle à produire du pollen fonctionnel (pour revue Budar et al., 2003).

Les SMNC sont des systèmes faisant intervenir deux composantes : une composante mitochondriale et une composante nucléaire. L'incapacité à produire du pollen fonctionnel est généralement gouvernée par une ORF mitochondriale inhabituelle qui est le plus souvent de nature chimérique (Hanson et Bentolila, 2004). Dans tous les exemples documentés, l'effet du cytoplasme inducteur de stérilité mâle peut être contrecarré par un ou des gènes nucléaires, généralement dominants, appelés des restaurateurs de fertilité (*Rf*) (Budar et al., 2003). Les copies alléliques des gènes *Rf* ne restaurant pas la fertilité mâle sont appelées des mainteneurs de stérilité, généralement récessifs. Autrement dit, les systèmes de SMNC déterminent la biologie reproductive des plantes, par le biais d'une interaction entre un cytoplasme inducteur de stérilité mâle et un ou plusieurs gènes nucléaires *Rf*.

Chez une espèce hermaphrodite, un individu est femelle s'il porte le facteur cytoplasmique inducteur de stérilité en présence de gènes mainteneurs, c'est-à-dire en l'absence de gènes restaurateurs. Un individu peut être hermaphrodite dans deux cas de figure : s'il porte un cytoplasme non inducteur de stérilité, ou s'il porte un cytoplasme inducteur de stérilité en présence de restaurateurs nucléaires de fertilité. Les acteurs moléculaires mitochondriaux et nucléaires des systèmes de SMNC seront détaillés dans le paragraphe VI.4. L'hérédité du déterminant inducteur de stérilité mâle étant maternelle, les SMNC peuvent être mises en évidence par des croisements réciproques : dans le cas général où les restaurateurs sont dominants, le phénotype de stérilité mâle ne sera visible que dans la F2 des croisements réciproques. Si les restaurateurs sont récessifs (ou les mainteneurs dominants), la descendance F1 sera mâle stérile dans un sens de croisement, et fertile dans l'autre sens (figure A9).

Le polymorphisme sexuel visible dans les systèmes de SMNC peut être considéré comme le

résultat d'un conflit génomique entre deux génomes qui diffèrent par leur mode de transmission (Cosmides et Tooby, 1981). L'hérédité des facteurs cytoplasmiques est maternelle alors que l'hérédité des gènes nucléaires est biparentale. Cela génère une situation où les intérêts des génomes sont en opposition lors de la production de pollen : les facteurs mitochondriaux peuvent favoriser l'apparition de la stérilité mâle alors que les facteurs nucléaires vont favoriser la fertilité des sexes mâle et femelle (Burt et Trivers, 2006).

Les systèmes de SMNC sont étudiés pour plusieurs raisons : en sélection, ils sont exploités pour produire des semences hybrides (Frankel et Galun, 1977 ; Mackenzie, 2004). Une plante mâle stérile est incapable de s'autoféconder. Chaque graine collectée sur une femelle est donc le résultat d'une pollinisation croisée. Ces graines, obtenues par croisement, vont donner des plantes hybrides potentiellement capables d'exprimer l'hétérosis. Si le fruit ou la graine est le produit désiré de l'hybride, la transmission du ou des gènes restaurateurs de fertilité, par le parent mâle du croisement, est essentielle pour conférer la capacité d'autofécondation des plantes hybrides. Pour les espèces chez qui le produit désiré est l'organe végétatif ou la fleur, la fertilité des hybrides n'est pas nécessaire et les gènes *Rf* dispensables dans la stratégie de production des semences hybrides. En recherche fondamentale, les SMNC sont des modèles moléculaires permettant d'étudier les relations génétiques et évolutives entre les génomes mitochondrial et nucléaire. Les SMNC intéressent deux communautés scientifiques différentes : les généticiens des populations, qui étudient les SMNC présentes dans certaines populations naturelles dites gynodioïques, où coexistent des individus femelles (mâle stériles) et des individus hermaphrodites dans une même population naturelle (Darwin, 1877), et les généticiens moléculaires, qui essaient d'identifier, principalement chez des espèces cultivées, les gènes impliqués dans les systèmes de SMNC.

Les SMNC ont été décrites chez plus de 140 espèces et se produisent dans deux situations différentes : dans la descendance de croisements entre des individus hermaphrodites provenant de différentes espèces ou de différentes populations de la même espèce, ou dans des populations gynodioïques. Dans la descendance de croisements entre des hermaphrodites, la progéniture mâle stérile est parfois observée dans la première génération ou dans les générations suivantes. Chez les ancêtres de l'individu utilisé comme parent femelle pour le croisement, il est possible qu'un gène mitochondrial inducteur de stérilité ait entraîné la sélection d'un ou de plusieurs gènes *Rf* (ce qui a permis de rétablir l'hermaphrodisme). Lors du croisement de cet individu avec un individu ne possédant pas les gènes *Rf*, la descendance

va présenter une dissociation du cytoplasme inducteur de stérilité et des restaurateurs nucléaires de fertilité, aboutissant ainsi à un phénotype de stérilité mâle. La seconde situation dans laquelle la SMNC peut être observé est celle des populations gynodioïques.

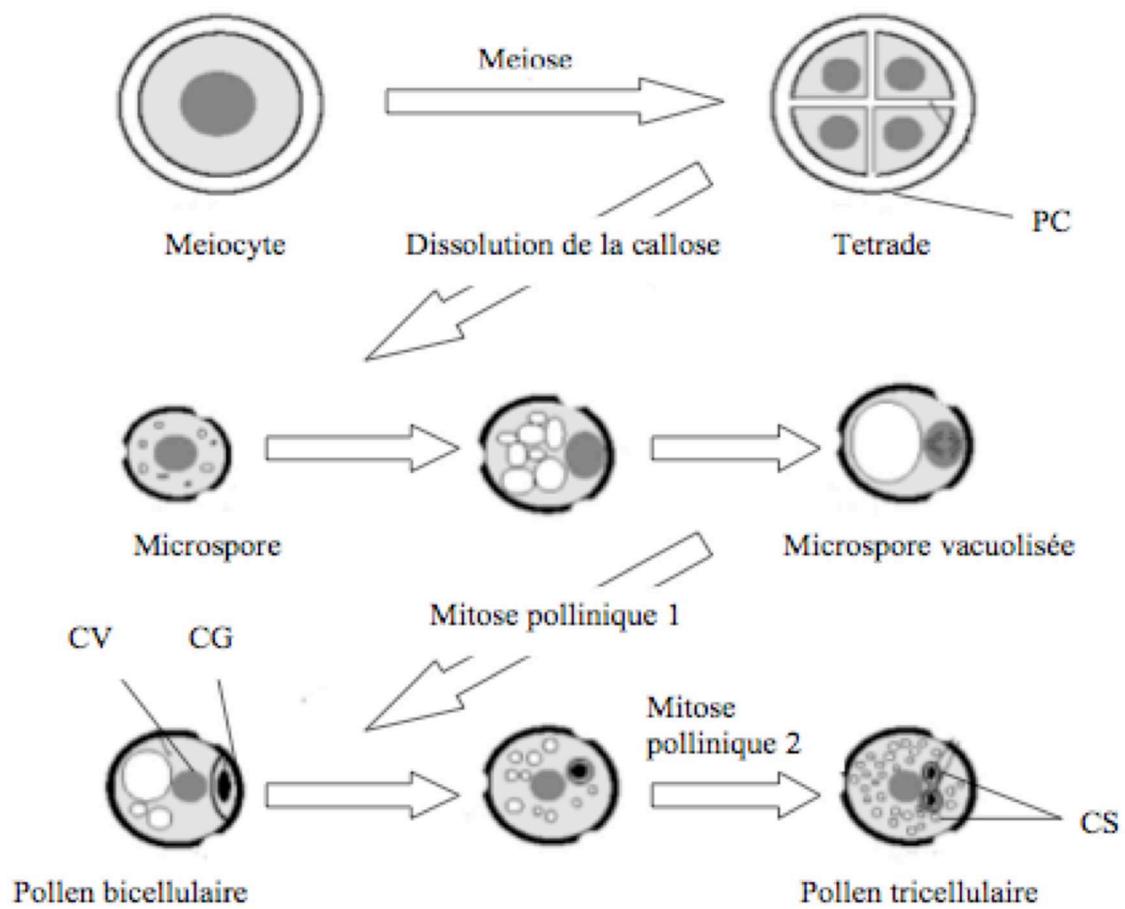
## VI.2 La SMNC à l'échelle de la population : la gynodioécie

La gynodioécie est l'un des systèmes de reproduction les plus fréquents, après l'hermaphroditisme, chez les angiospermes (présent chez 7% des espèces. Richards, 1997). Même s'il existe des cas de gynodioécie dans lesquels la détermination du sexe est sous contrôle uniquement nucléaire, la SMNC est le système de détermination du sexe le plus souvent à l'origine du polymorphisme sexuel observé dans les populations gynodioïques (Ashman, 2003 ; Bailey et Delph, 2007). La gynodioécie est considérée comme une étape intermédiaire dans l'évolution de la dioécie à partir d'ancêtres hermaphrodites (Darwin, 1877 ; Lloyd, 1974a,b). Plus généralement, l'étude des conflits génomiques qui conduisent à la gynodioécie permet de mieux comprendre la transmission et l'évolution du génome mitochondrial dans le cadre d'une coévolution nucléocytoplasmique (McCauley et Bailey, 2003).

Le maintien du polymorphisme sexuel dans des populations gynodioïques nucléocytoplasmiques a depuis longtemps intrigué les évolutionnistes : la SMNC peut se répandre dans la population grâce à un avantage femelle obtenu par une réallocation des ressources non utilisées pour la production de pollen vers la production de graines. Cet avantage des individus femelles, par rapport aux individus hermaphrodites, est aussi appelé compensation femelle (Darwin, 1877) et est considéré comme le résultat d'un compromis (« trade-off » en anglais) entre la production de pollen et la production de graines. Plusieurs équipes ont essayé de quantifier la compensation femelle dans des espèces gynodioïques notamment chez *Thymus vulgaris* (Gigord et al., 1999). Lorsque la compensation femelle existe, la proportion d'individus femelles qui portent le facteur cytoplasmique inducteur de stérilité mâle dans la population augmente à chaque génération *via* la sélection naturelle (Lewis, 1941). Les génomes nucléaires associés à un cytoplasme inducteur de stérilité étant désavantagés dans leur propagation, tout gène nucléaire capable de restaurer la fertilité mâle a un avantage sélectif en présence du cytoplasme stérilisant. Ceci devrait conduire à la fixation des gènes restaurateurs dans la population et donc à une population composée uniquement d'individus hermaphrodites. Par exemple, chez *Mimulus*, le

cytoplasme IDS est fixé dans la population *M. guttatus* Iron Mtn alors que cette population est entièrement hermaphrodite, ce qui indique que les restaurateurs nucléaires sont fixés dans cette population (Case et Willis, 2008). Yamagishi et Terachi (1994) ont trouvé des radis sauvages qui possèdent le cytoplasme Ogura IDS et qui ont une fertilité mâle normale, déduisant ainsi que les restaurateurs sont présents dans la population sauvage : ces SMNC sont dites dans un état « cryptique ». Néanmoins, des individus femelles sont observés dans les populations gynodioïques. L'existence des individus femelles peut s'expliquer de deux façons : si ces individus femelles produisent des graines en quantité et/ou en qualité très supérieures à celles qui sont produites par les individus hermaphrodites (compensation femelle élevée), les individus femelles vont pouvoir se maintenir dans la population, même si des restaurateurs nucléaires sont présents. Au contraire, si la compensation femelle est faible, la SMNC peut se maintenir dans la population malgré la présence de gènes *Rf* par un effet pléiotrope négatif des allèles restaurateurs. Cet effet est appelé coût de la restauration (Charlesworth, 1981 ; Delannay et al., 1981 ; Frank, 1989 ; Gouyon et al., 1991 ; Bailey et al., 2003). D'après Bailey et al. (2003), le coût de la restauration, qui peut affecter soit la production de pollen, soit la production de graines, peut s'exprimer de trois façons : un coût « constitutif » qui est indépendant de la nature du cytoplasme (cytotypes stérilisant et non stérilisant) ; un coût « silencieux » pour lequel seuls les restaurateurs qui n'affectent pas la détermination du sexe (en présence d'un cytotype non stérilisant) induisent un coût ; un coût « exprimé », pour lequel les restaurateurs qui affectent la détermination du sexe induisent un coût (en présence d'un cytotype stérilisant).

Un coût peut être dominant (intervient chez les individus qui ont une seule copie de l'allèle restaurateur) ou récessif (intervient seulement chez les individus homozygotes pour l'allèle restaurateur). Dans le cas où le coût de la restauration serait dominant, les individus ayant un seul allèle restaurateur ont une production de graines et/ou de pollen qui est relativement plus faible que celle des individus sans allèle restaurateur : le polymorphisme au locus restaurateur est maintenu par un avantage du rare ; au contraire, un coût de la restauration qui est récessif provoque une surdominance du locus restaurateur, ce qui permet de maintenir le polymorphisme au locus restaurateur : les individus hétérozygotes ont une fitness mâle (pollen) plus élevée que les individus homozygotes. Les individus hétérozygotes sont hermaphrodites dans le cas où la restauration est dominante ; cependant, ces individus n'expriment pas de coût de la restauration car ce dernier est récessif (Bailey et al., 2003 ; Delph et al., 2007). Des études ont essayé de quantifier le coût de la restauration chez des



**Figure A10 : développement du gamétophyte mâle chez Arabidopsis**

PC paroi de callose. CV : cellule végétative. CG : cellule génératrice. CS : cellules spermatiques.

espèces gynodioïques : chez *Lobelia siphilitica*, un coût de la restauration sur la viabilité des grains de pollen a été identifié (Bailey, 2002). Dans des systèmes de SMNC chez le riz et chez le maïs, les restaurateurs altèrent les profils de transcription de gènes mitochondriaux essentiels dans des cytoplasmes inducteurs ou non inducteurs de stérilité mâle, ce qui peut être considéré comme un coût de la restauration (Delph et al., 2007).

La compensation femelle et le coût de la restauration peuvent expliquer le maintien du polymorphisme sexuel dans des populations gynodioïques nucléocytoplasmiques (Delannay et al., 1981 ; Bailey et al., 2003). Néanmoins, d'autres facteurs peuvent intervenir et sont parfois pris en compte dans les modèles, comme la structure de la population, le degré d'autofécondation des individus hermaphrodites, la limitation (ou la disponibilité) du pollen ainsi que le nombre de cytotypes stérilisants dans la population (Delph et al., 2007).

### VI.3 La SMNC à l'échelle de l'individu : la gamétogenèse mâle

Chez les angiospermes, le développement de l'anthere peut être divisé en deux phases : la première phase comprend des divisions et des différenciations cellulaires qui conduisent à la mise en place des tissus de l'anthere : il s'agit de la phase sporophytique. La seconde phase, ou phase gamétophytique, est une phase de développement des microspores en grains de pollen matures qui sont finalement libérés lors de la déhiscence des anthers. Chez des angiospermes comme *Arabidopsis*, la phase sporophytique est constituée de 8 étapes au cours desquelles il y a la formation des quatre couches cellulaires sporophytiques (épiderme, endothécium, couche moyenne et tapis). Ces couches entourent le sac pollinique qui contient les cellules mères du pollen, ou méiocytes. Lors de la phase gamétophytique, les microspores se développent en grains de pollen matures. Parallèlement, le tapis assure des fonctions nutritives cruciales pour le développement du pollen, puis subit une dégénérescence programmée à un stade précis. Lors de la déhiscence des anthers (ou anthèse), les grains de pollen matures sont libérés.

La phase gamétophytique mâle (figure A10) commence par la formation d'une tétrade à partir d'un méiocyte (ou cellule mère du pollen) et se termine à l'anthèse avec la libération des grains de pollen matures. La phase gamétophytique peut être divisée en deux étapes : la microsporogenèse et la microgamétogenèse. La microsporogenèse commence par deux divisions méiotiques, aboutissant à la formation d'une tétrade constituée de quatre

microspores haploïdes enveloppées de callose. Ensuite un dépôt de primexine initie la formation de la paroi du grain de pollen. Le tapis va ensuite sécréter des callases pour dégrader la callose de la tétrade et libérer les microspores dans le locule de l'anthere. La microgametogenèse commence à cette étape. Les vacuoles de la microspore vont alors fusionner pour donner une seule grande vacuole. Cette vacuolisation va entraîner une polarisation du noyau de la microspore (Boavida et al., 2005) qui subit ensuite la première mitose pollinique, ou division asymétrique, pour produire un grain de pollen comprenant une grande cellule végétative et une petite cellule générative internalisée dans la cellule végétative. Les deux cellules sont séparées par une paroi cellulaire constituée de callose. La cellule végétative nourrit la cellule générative en développement, et formera, lors de la germination du pollen, le tube pollinique. Chez *Arabidopsis*, après fission de la vacuole et migration de la cellule générative, la cellule générative subit la deuxième mitose pollinique pour former deux gamètes mâles identiques (Borg et al., 2009) ce qui aboutit à la formation d'un grain de pollen trinué. Les deux gamètes mâles vont participer à la double fécondation : avec le gamète femelle (ou oosphère) pour former le zygote et avec les noyaux polaires pour former le tissu de réserve triploïde ou albumen. Le développement embryonnaire pourra commencer.

L'analyse du transcriptome du gamétophyte mâle d'*Arabidopsis* suggère que parmi les 3500 gènes exprimés lors du développement pollinique, plus de 1400 semblent être préférentiellement ou spécifiquement exprimés dans le grain de pollen (Honys and Twell, 2004; Pina et al., 2005). Peu de ces gènes ont été caractérisés. Parmi eux, les gènes *GSL* (pour *Glucan Synthase-Like*), impliqués dans la synthèse de composants de la paroi cellulaire comme la callose, jouent un rôle important lors de la gamétogenèse mâle. La callose est un composant structural important du gamétophyte mâle (McCormick, 1993). Un dérèglement de la production et/ou du dépôt de callose lors de la gamétogenèse mâle peut entraîner une stérilité mâle (Töller et al., 2008). Un autre exemple de gènes impliqués dans la production de pollen est celui des gènes *AGP6* et *AGP11* (pour *ArabinoGalactanes-Protéines*) sont des macromolécules extracellulaires des plantes et sont constituées d'une faible partie protéique et d'une très grande partie glucidique. Ce sont notamment des constituants de la paroi végétale. Chez *Arabidopsis*, *AGP6* et *AGP11* sont fortement et spécifiquement exprimés dans le grain de pollen (Pereira et al., 2006 Levitin et al., 2008) et des analyses de doubles mutants *agp6-agp11* ont montré qu'ils étaient importants dans la formation de grains de pollen viables (Coimbra et al., 2009).

Dans les systèmes de SMNC, le phénotype de stérilité mâle est dû à un défaut gaméophytique ou sporophytique où, selon les systèmes, les organes mâles sont totalement absents, la production de pollen est déficiente (l'avortement du pollen peut intervenir à différents stades de la gamétogenèse), ou il y a une absence de déhiscence des anthères (Laser et Lersten, 1972 ; Budar et Pelletier, 2001). En fonction des systèmes de SMNC, les événements cytologiques les plus prononcés ont lieu au niveau du tapis de l'anthère. Chez les plantes mâle stériles, le tapis peut être anormalement vacuolisé, les cellules du tapis peuvent être fusionnées ou la mort cellulaire programmée du tapis peut être dérégulée (Shi et al., 2010).

#### VI.4 Aspects moléculaires des stérilités mâles nucléo-cytoplasmiques

##### VI.4.1 Facteurs mitochondriaux inducteurs de stérilité

Lors des premières études sur les systèmes de SMNC, il était connu que le gène inducteur de stérilité mâle était à hérédité maternelle, mais on ne savait pas quel organite, le chloroplaste ou la mitochondrie, était porteur de ce gène. L'hypothèse d'une localisation chloroplastique des gènes inducteurs de stérilité a été écarté par des expériences de fusion de protoplastes issus de tabac mâle fertile et mâle stérile : dans la descendance des produits de fusion, l'ADN chloroplastique du parent mâle fertile était présent dans des cybrides mâle stériles et inversement (Belliard et al., 1979). Des études portant sur la SMNC de type Texas chez le maïs ont également montré que la mitochondrie était à l'origine de l'induction de la stérilité mâle (Dewey et al., 1986, 1987). Depuis, tous les gènes responsables de la stérilité mâle dans les systèmes de SMNC se sont avérés être localisés dans le génome mitochondrial (Schnable et Wise, 1998 ; Chase, 2007). À noter le cas exceptionnel d'une stérilité mâle qui n'a pas une origine mitochondriale : dans la stérilité « 447 » chez la fève, l'inducteur de stérilité est une particule cytoplasmique de type viroïde portant une molécule d'ARN double brin (Lefebvre et al., 1990).

Les deux approches classiquement utilisées pour identifier les gènes mitochondriaux inducteurs de stérilité sont les suivantes : la première approche consiste à comparer les génomes mitochondriaux de plantes stériles à ceux de plantes révertantes fertiles qui ont éliminé, par recombinaison, la séquence inductrice de stérilité mâle. Cette approche a permis d'identifier le gène mitochondrial inducteur de stérilité dans les systèmes de SMNC Ogura et Texas (détaillé dans le chapitre VI.5). La seconde approche consiste à comparer l'expression

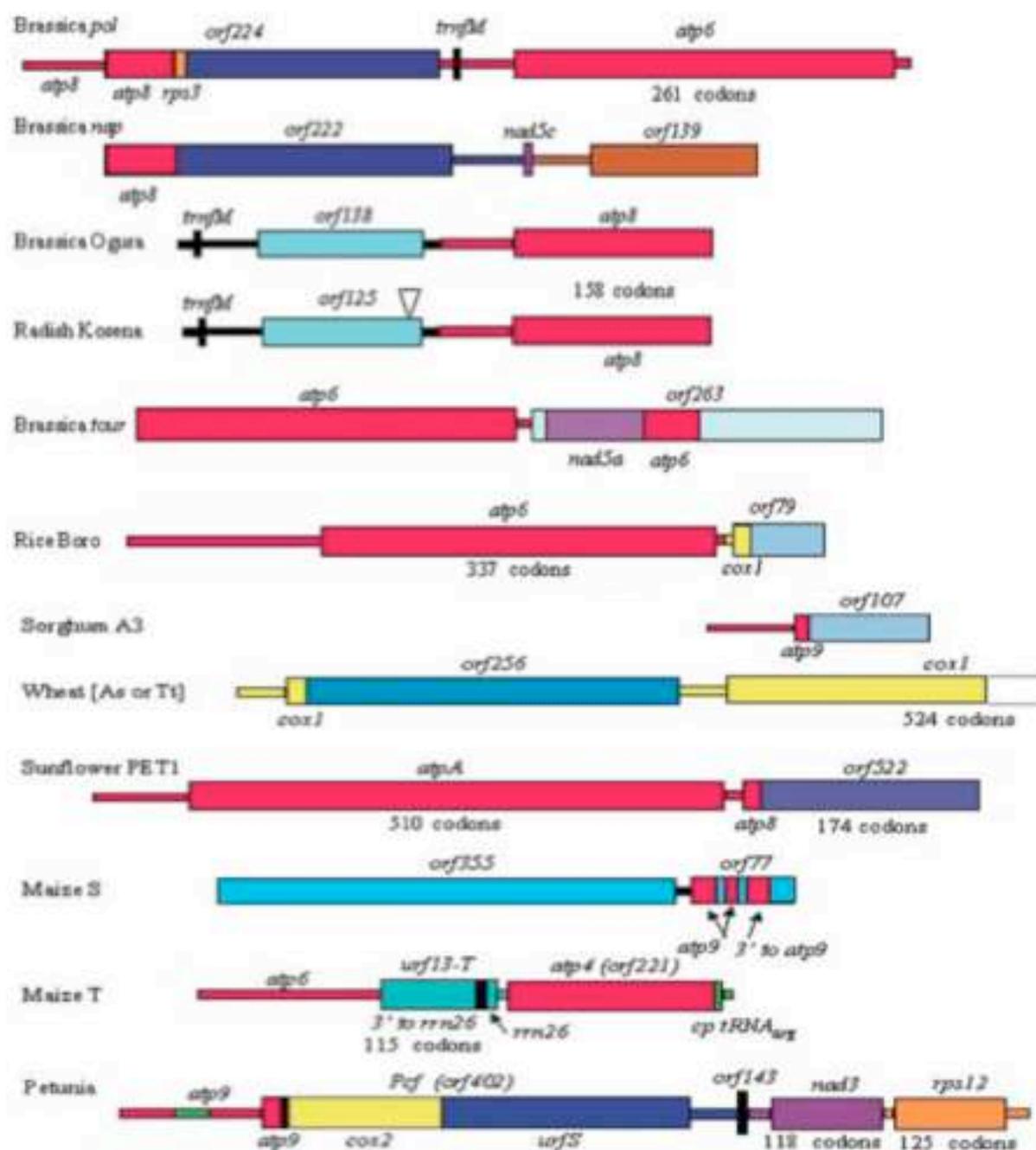
différentielle des génomes mitochondriaux inducteurs de stérilité dans un contexte nucléaire mainteneur ou restaurateur de fertilité (détaillé dans le chapitre VI.4.2). Cependant, le génome nucléaire ayant une grande influence sur l'expression des gènes mitochondriaux, il est nécessaire, dans cette approche, d'utiliser des lignées quasi-isogéniques qui diffèrent seulement à un petit nombre de locus. De plus, les gènes restaurateurs peuvent avoir une influence sur l'expression de plusieurs gènes mitochondriaux, comme dans des systèmes de SMNC chez le riz et le maïs (Delph et al., 2007), ce qui complique également l'identification formelle des gènes liés aux phénotypes de stérilité.

Les deux premiers locus mitochondriaux associés à une induction de stérilité mâle dans les systèmes de SMNC ont été identifiés dans le système Texas du maïs et chez le pétunia.

La séquence d'ADN mitochondrial identifiée chez le maïs Texas est *urf13* (Dewey et al., 1986), alors que chez le pétunia, c'est le gène *pcf* (pour *petunia CMS-associated fused*) qui est responsable de l'induction de la stérilité mâle (Young et Hanson, 1987). L'identification de *urf13* s'est faite grâce à l'utilisation de sondes d'hybridation, qui ont révélé la présence d'une séquence d'ADN mitochondrial fortement transcrite dans le cytoplasme Texas par rapport au cytoplasme non inducteur de stérilité (Dewey et al., 1986), mais également en étudiant des descendants stériles et fertiles (révertants) d'un même cybride, donc vraisemblablement très proches les uns des autres, sauf pour le déterminant de la stérilité (Bonhomme et al., 1991). Des synthèses *in organello* avaient déjà permis de détecter la protéine URF-13 chez des plantes stériles porteuses du cytoplasme Texas (Forde et al., 1978).

Chez le pétunia, la région d'ADN mitochondrial associée à la SMNC a été identifiée par des analyses de lignées hybrides somatiques créées par fusion de protoplastes provenant de lignées mâle stériles et de lignées mâle fertiles (Boeshore et al., 1983). Ces expériences de fusions de protoplastes ont généré des cybrides fertiles et stériles contenant des génomes mitochondriaux ayant subi des recombinaisons entre les génomes des deux parents de fusion (Boeshore et al., 1983 ; Rothenberg et al., 1985). Les génomes mitochondriaux recombinés ont rendu possible la recherche d'une région d'ADN mitochondrial corrélée avec le phénotype de stérilité dans les hybrides somatiques. Cette région a été identifiée (Boeshore et al., 1985) puis nommée « locus *S-pcf* » (Hanson et al., 1989). Parmi les trois gènes contenus dans le locus *S-pcf*, le gène *pcf* code une protéine dont l'abondance est fortement réduite dans des lignées de pétunia restaurées (Nivison et Hanson, 1989).

L'identification de *urf13* et de *pcf* a permis d'établir une caractéristique importante des locus



**Figure A11 : gènes chimériques associés à différents systèmes de SMNC**

En rouge : gènes codant des sous-unités de l'ATP synthase. En bleu : orf inconnues et non conservées dans les régions associées aux stérilités. En jaune : gènes codant des sous unités de la cytochrome oxydase. En orange : gènes codants des protéines ribosomiques. En marron : orf inconnues mais conservées dans l'ADN mitochondrial des plantes vasculaires. En vert : séquences d'ADN chloroplastique. (D'après Hanson et Bentolila, 2004).

associés aux SMNC : les deux locus sont formés par des réarrangements dans les génomes mitochondriaux. Les nouvelles *orf* créées (voir le paragraphe III.2) sont considérées comme des mutations gain de fonction (pour revue Schnable et Wise, 1998 ; Hanson et Bentolila, 2004 ; Fujii et al., 2010). Les nouvelles *orf*, associées aux phénotypes de stérilité, sont des *orf* chimériques, composées de séquences d'ADN d'origine inconnue et de séquences dérivées de fragments de gènes connus. Les fragments de gènes connus dérivent souvent de gènes mitochondriaux codant des sous-unités des complexes de la chaîne de phosphorylation oxydative comme le complexe ATP synthase ou la cytochrome oxydase. Par exemple, chez le pétunia, un fragment du gène codant la sous-unité ATP9 et un fragment du gène *cox2* sont présents dans le gène *pcf* (Young et Hanson, 1987). Des séquences d'origine chloroplastique et des séquences qui dérivent de gènes codant des protéines ribosomiques ont également été détectées dans les locus liés aux stérilités, notamment à proximité du gène *pcf* de pétunia (Hanson et Bentolila, 2004). Le gène mitochondrial *urf13* est composé de séquences non codantes normalement trouvées en 3' du gène *rrn26* et des séquences d'origine inconnue (Wise et al., 1999). Aujourd'hui, une douzaine de régions d'ADN mitochondrial associées de façon plus ou moins convaincante à la stérilité mâle ont été identifiées (figure A11). Les gènes liés au phénotype de stérilité mâle présentent pas ou peu d'homologie de séquence d'un système de SMNC à l'autre. Néanmoins, il existe des exceptions comme chez *Brassica napus* où l'*orf222*, responsable de la SMNC *nap* et l'*orf224*, responsable de la SMNC *polima* codent des protéines possédant 79% de similarité (L'Homme et al., 1997).

Les *orf* chimériques inductrices de stérilité sont généralement fusionnées ou situées à proximité de séquences promotrices (pour revue, Vedel et al., 1994 ; Bellaoui et al., 1997). Les *orf* chimériques sont co-transcrites avec les séquences de gènes connus puis traduites à partir des ARNm bicistroniques produits (Budar et Pelletier, 2001). L'association et la co-expression des gènes de stérilité avec les séquences de gènes connus semblent essentielles pour un maintien durable de ces gènes dans les génomes mitochondriaux (Budar et Pelletier, 2001 ; Budar et al., 2003).

L'étude des séquences des gènes chimériques associés aux SMNC a montré que plusieurs d'entre elles codaient potentiellement des protéines avec des domaines transmembranaires, potentiellement capables de s'intégrer dans les membranes internes des mitochondries. La structure et la biochimie de quelques protéines codées par les gènes de stérilité ont été étudiées expérimentalement (pour revue, Hanson et Bentolila, 2004). En particulier, la protéine T-URF13, responsable, chez le maïs, de la stérilité associée au cytoplasme Texas

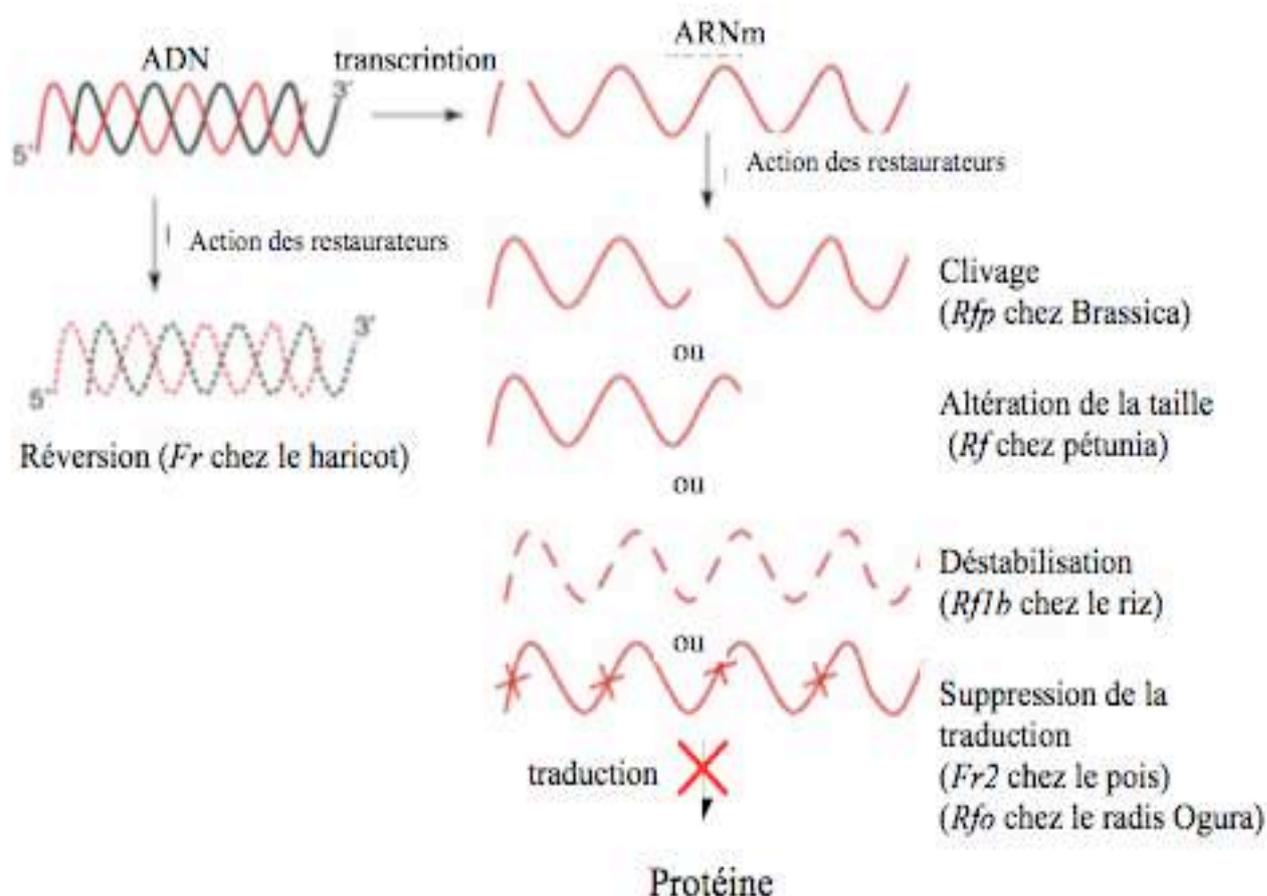
(Wise et al., 1999 ; Rhoads et al., 1995) ; la protéine ORF138, responsable de la stérilité Ogura chez des espèces de *Brassica* et *Raphanus* (Bonhomme et al., 1992 ; Grelon et al., 1994 ; Krishnasamy et Makaroff, 1993 ; Duroc et al., 2005) ont été étudiées en détail.

Dans le système Texas, T-URF13 possède des domaines hydrophobes et forme des oligomères. La structure qui en résulte constitue un pore, formé de 6 à 8 monomères (Siedow et al., 1995) qui s'ouvre lorsque des toxines spécifiques se lient aux résidus exposés de T-URF13. L'ouverture du pore provoque le découplage de la phosphorylation oxydative puis la mort des bactéries *E. coli* dans lesquelles T-URF13 est exprimée. Le phénotype associé à T-URF13 est non seulement une stérilité mâle, mais également un phénotype de sensibilité à des toxines de champignon comme *Byopolaris maydis*, *Phyllostica maydis* et à l'insecticide méthomyl (Sabar et al., 2003).

La SMNC Ogura est un des systèmes de SMNC les plus étudiés. À l'origine, un cytoplasme inducteur de stérilité a été identifié chez *Raphanus sativus* (Ogura, 1968). Afin d'obtenir du colza mâle stérile, un croisement sexuel entre *Raphanus sativus*, utilisé comme parent femelle et *Brassica napus* utilisé comme parent mâle, suivi par des rétrocroisements répétés, a permis d'obtenir des lignées alloplasmiques où le génome de *Brassica* a été introduit dans le cytoplasme du radis Ogura (Bannerot et al., 1974). Les lignées alloplasmiques étaient mâle stériles, mais possédaient également des déficiences chlorophylliennes, des déformations florales et une faible fertilité femelle (Rousselle, 1981 ; Pelletier et al., 1983 ; Mesquida et Renard, 1978). Pour supprimer ces caractères indésirables, attribués à l'interaction du cytoplasme de *Raphanus sativus* avec le génome nucléaire de *Brassica napus*, des fusions entre des protoplastes issus de ces plantes alloplasmiques et des protoplastes de colza fertile ont été effectuées (Pelletier et al., 1983). Ces fusions ont permis de créer des cybrides où les gènes des organites de *Raphanus sativus*, responsables des caractères indésirables, ont été remplacés par ceux de *Brassica napus*, tout en conservant le facteur cytoplasmique inducteur de stérilité provenant de *Raphanus sativus* et le fond nucléaire de *Brassica napus*. Finalement, des cybrides de *Brassica napus* mâle stériles présentant peu de défauts ont été sélectionnés (Pelletier et al., 1983). La SMNC Ogura, identifiée chez *Raphanus sativus* puis transférée chez *Brassica napus*, est contrôlée par le locus mitochondrial *orf138-orfB* qui s'exprime sous forme d'un ARN bicistronique (Bonhomme et al., 1991, 1992). *Orf138* est le gène inducteur de stérilité (Bonhomme et al., 1992) et *orfB* code la sous-unité 8 du complexe ATP synthase (Gray et al., 1998). Le mécanisme par lequel ORF138 interfère avec la production de pollen

n'est pas complètement élucidé, mais plusieurs éléments d'information sur la protéine ORF138 ont été obtenus (Grelon et al., 1994 ; Duroc et al., 2005). Dans le système Ogura, les plantes mâle stériles ne semblent pas être sensibles aux toxines. ORF138, qui s'oligomérisse, s'associe aussi à la membrane cytoplasmique des bactéries *E. coli* et à la membrane interne des mitochondries. ORF 138 est toxique à faible concentration et inhibe la croissance d'*E. coli*. Contrairement à URF-13, ORF138 ne semble pas agir par inhibition de la respiration bactérienne et il est peu probable que la toxicité résulte de la formation de pores dans la membrane interne de *E. coli*. (Duroc et al., 2005). Récemment, Duroc et al (2009) ont montré que la protéine ORF138 était le principal composant, sinon l'unique, d'un complexe formé dans la membrane interne des mitochondries des plantes stériles ; ils ont également montré que les mitochondries isolées de plantes stériles avaient une plus forte consommation d'oxygène que celles de plantes fertiles, ce qui suggère un découplage partiel de la respiration mitochondriale en présence du complexe ORF138. Un modèle a été proposé, où le découplage partiel induit par ORF138 serait compensé au niveau des tissus végétatifs, mais pas au niveau des organes reproductifs mâles (Duroc et al., 2009). D'autres protéines mitochondriales inductrices de stérilité mâle ont été étudiées comme par exemple la protéine ORF522 chez le tournesol. Des plantes mâle stériles de tournesol, exprimant la protéine ORF522, présentent une diminution de l'activité ATPasique par rapport aux plantes isogéniques fertiles (Sabar et al., 2003).

Dans la plupart des systèmes de SMNC étudiés, la protéine responsable du phénotype de stérilité est présente dans tous les organes des plantes mâles stériles. Pourtant, le plus souvent aucun défaut morphologique ou respiratoire n'est détecté dans les tissus végétatifs de ces plantes (Farineau et al., 1990 ; Hanson et Bentolila, 2004). Toutes ces protéines induisent un dysfonctionnement mitochondrial dont l'effet semble donc limité à un petit nombre de types cellulaires dans les anthères. Les mitochondries jouent un rôle important dans le tapis de l'anthère (le tissu nourricier du grain de pollen), dans les cellules méiotiques mâles et/ou dans le développement des microspores. Le développement normal des cellules tapétales et du pollen est accompagné d'une augmentation de 20 à 40 fois du nombre de mitochondries par cellule (pour revue Linke et Borner, 2005). Chez les plantes mâle stériles, une diminution de la production d'ATP et/ou un dysfonctionnement des mitochondries peut expliquer le développement anormal des grains de pollen. En effet, la production de grains de pollen fonctionnels nécessite une forte demande en énergie ou en produits particuliers d'origine mitochondriale. Néanmoins, il existe des arguments qui semblent en défaveur de cette



**Figure A12 : effets des restaurateurs nucléaires sur l'expression des gènes mitochondriaux inducteurs de stérilité**

Les restaurateurs empêchent l'accumulation de protéines associées à la stérilité par différents processus : le restaurateur *Fr* du haricot élimine la portion de l'ADN mitochondrial qui contient l'ORF associée à la stérilité ; *Rfp* agit en clivant l'ARNm ; *Rf* modifie la taille de l'ARNm alors que *Rflb* agit en le déstabilisant ; *Fr2* supprime la traduction de l'ARNm. Dans le système Ogura (Brassica et Raphanus), *Rfo* pourrait bloquer l'initiation de la traduction de l'ORF138 associée à la stérilité (Uyttewaal et al, 2008b).

hypothèse énergétique : d'autres processus biologiques, qui nécessitent également beaucoup d'ATP (comme l'embryogenèse), ne sont pas affectés chez les plantes mâle stériles.

#### VI.4.2 Facteurs nucléaires restaurateurs de fertilité

L'effet des protéines mitochondriales inductrices de stérilité mâle peut être contrecarré par les produits d'un ou de plusieurs gènes nucléaires *Rf* (Budar et al., 2003). La suppression des anomalies associées à la stérilité mâle permet un retour à la normale des processus métaboliques associés à l'organogenèse mâle, au développement du pollen et à sa maturation (Kotchoni et al., 2010).

La restauration de la fertilité résulte, le plus souvent, d'une diminution de l'expression du gène mitochondrial associé à la stérilité. L'exception à cette règle est le gène *Rf2* du système Texas de maïs. *Rf2* code une aldéhyde déshydrogénase qui agit en compensant le défaut métabolique engendré par T-URF13 (Liu et al., 2001). Dans les autres systèmes de restauration étudiés, les gènes *Rf* codent des protéines qui interfèrent à différentes étapes de l'expression des gènes mitochondriaux inducteurs de stérilité. Certains gènes *Rf* semblent pouvoir restaurer la fertilité en agissant directement sur la séquence de l'ADN mitochondrial contenant le locus inducteur de stérilité. Le locus *Fr* du haricot est associé à la perte partielle d'une région d'ADN mitochondrial particulière associée à la stérilité mâle (la région *pvs*). Le locus *pvs* est réduit à un niveau sous-stœchiométrique en présence du locus *Fr* (Janska et al., 1998). Le(s) gène(s) et le mode d'action du locus *Fr* sont encore inconnus.

Contrairement au gène *Rf2*, la majorité des gènes *Rf* identifiés sont des gènes qui codent des protéines à motifs PPR (voir paragraphe IV.1) capables d'affecter l'expression du locus mitochondrial de stérilité en affectant le profil des ARNm et/ou l'accumulation de la protéine de stérilité (figure A12).

L'analyse des ARNm de la région mitochondriale *pol*, associée à une SMNC chez *Brassica napus*, a révélé que le restaurateur *Rfp* favorise le clivage de l'ARN bicistronique inducteur de stérilité *orf224/atp6* (Li et al., 1998 ; Brown, 1999). Ces résultats indiquent que *Rfp* est impliqué dans le clivage de l'ARN *orf224/atp6* ; la co-expression du gène de stérilité *orf224* avec la séquence de gène *atp6* est certainement essentielle pour que la stérilité mâle puisse s'exprimer. Le clivage, par *Rfp*, de l'ARN *orf224/atp6* empêcherait leur co-expression, permettant ainsi de restaurer la fertilité mâle. Il est également possible que la restauration de

la fertilité soit due à un retour à la normale de la quantité du transcrit *atp6* lors du clivage de l'ARN *orf224/atp6* (Hanson et Bentolila, 2004).

Le locus *Rf* de pétunia est un locus complexe contenant deux gènes PPR (*Rf-591* et *Rf-592*). Le gène *Rf592* code une protéine adressée aux mitochondries qui empêche l'expression du gène de stérilité. La protéine RF592 s'associe à l'extrémité 5'UTR de l'ARNm *pcf* *in vivo*, à l'origine d'une réduction du niveau d'ARNm *pcf* (Hanson et al., 1999 ; Gillman et al., 2007). *Rf-592* contient 14 copies d'un motif PPR (Bentolila et al., 2002). Le gène *Rf-591* code une protéine de fonction inconnue, qui contient des motifs PPR et qui est très semblable à la protéine RF-592. À noter qu'il est possible que le locus *Rf* agisse aussi sur le niveau de protéine PCF (Hanson et al., 1999).

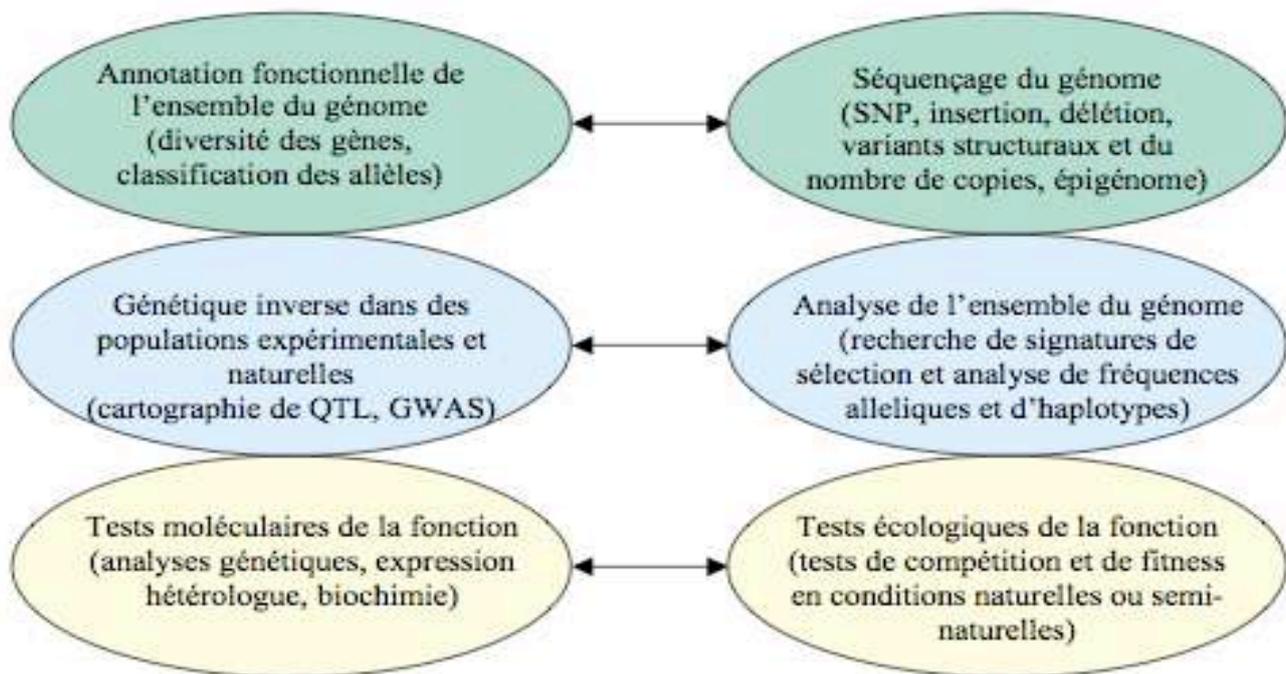
Chez le riz, deux gènes *Rf* (*Rf1a* et *Rf1b*) présents dans le même locus, peuvent supprimer l'expression de l'*orf79*, associée à la SMNC Boro II par différents mécanismes (Wang et al., 2006 ; Kazama et al., 2008). La protéine RF1A empêche la synthèse de la protéine ORF79 en clivant l'ARN bicistronique *B-atp6/orf79*, ce qui affecte à la fois la stabilité et l'activité traductionnelle des ARN *orf79*. Indépendamment de l'activité de clivage, il semblerait également que RF1A favorise l'édition de l'ARNm *atp6*. RF1B agit en favorisant la dégradation des ARN *B-atp6/orf79*. À noter qu'une modification des profils des ARNm, issus des locus inducteurs de stérilité, peut être le résultat indirect d'un défaut de la traduction de ces ARNm car les ARNm mitochondriaux peuvent parfois être déstabilisés lorsqu'ils ne sont plus traduits (Costanzo et Fox, 1990).

La restauration de la SMNC Ogura a été particulièrement étudiée : elle est assurée par le locus *Rfo* identifié chez *Brassica* et *Raphanus* (Brown et al., 2003 ; Desloire et al., 2003). La SMNC Ogura et le locus *Rfo* restaurateur de fertilité ont été introgressés du *Raphanus sativus* vers *Brassica napus* (Pelletier et al., 1983 ; Heyn, 1976). Cependant, l'introgession du locus *Rfo* a également introduit des caractéristiques génétiques délétères entraînant une perte d'information génétique de *Brassica napus* (Delourme et al., 1998) Des méthodes de sélection classique ont été utilisées pour améliorer ces lignées avec un succès limité par le manque de recombinaison homologue entre les régions homéologues de *Raphanus sativus* et de *Brassica napus* (Delourme et al., 1998). Le locus *Rfo* est organisé en un cluster de trois gènes très similaires organisés en tandem, nommés *PPR-A*, *PPR-B* et *PPR-C*. *PPR-B*, génétiquement défini et confirmé par transformation génétique comme étant le gène restaurateur de fertilité (Desloire et al., 2003, Brown et al., 2003 ; Koizuka et al., 2003) est code une protéine prédite qui contient 17 motifs PPR (Uyttewaal et al., 2008b). *PPR-A* et

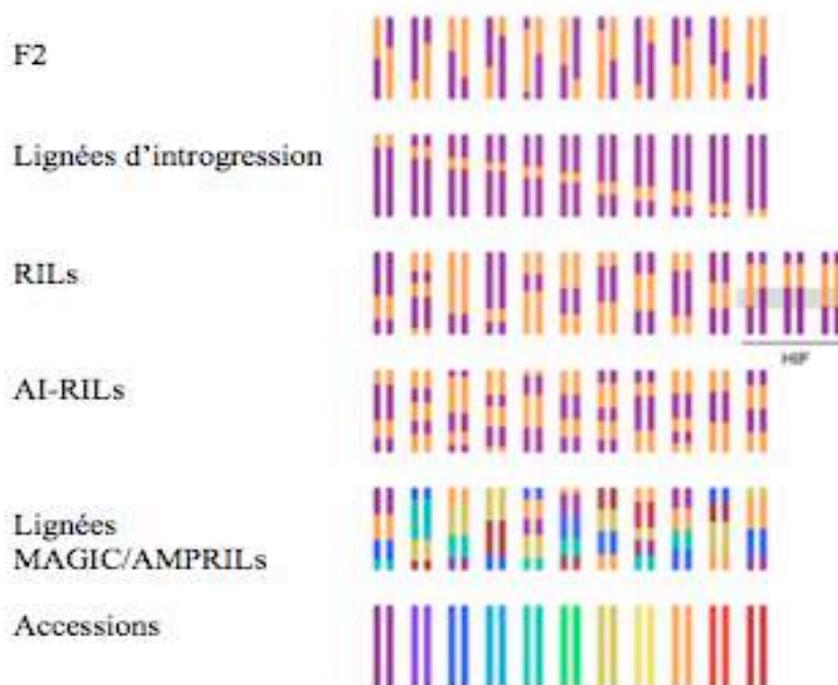
*PPR-B* possèdent 87 % d'identité au niveau des acides aminés codés. *PPR-A* contient une délétion de 12 nucléotides dans la troisième répétition PPR empêchant certainement la protéine PPR-A d'avoir une activité. Le gène *PPR-C* contient une autre délétion qui induit un décalage du cadre de lecture et l'apparition d'un codon stop au milieu du gène. *PPR-C* est un pseudogène (Desloire et al., 2003 ; Uyttewaal et al., 2008b). Il a tout d'abord été montré que la restauration de la fertilité mâle était corrélée à une faible accumulation de la protéine ORF138, sans avoir d'effet sur le niveau d'accumulation de l'ARNm *orf138* (Bellaoui et al., 1997, Uyttewaal et al., 2008b). Puis, en étudiant l'association de PPR-A et PPR-B avec l'ARNm *orf138*, Uyttewaal et al., (2008b) ont montré que la restauration de la fertilité dans le système Ogura était corrélée *in vivo* à l'association de PPR-B avec l'ARNm *orf138*. L'association de PPRB pourrait être à l'origine d'un blocage de l'initiation de la traduction de l'ARNm *orf138* (Uyttewal et al., 2008b).

Parfois, les locus restaurateurs peuvent affecter non seulement l'expression des locus mitochondriaux inducteurs de stérilité, mais également l'expression d'autres locus mitochondriaux non associés à l'induction de la stérilité. Dans le système de SMNC S du maïs, la comparaison de plantes stériles et de plantes restaurées a montré que le locus restaurateur *Rf3* coségrège non seulement avec la présence d'ARN plus courts provenant du locus inducteur de stérilité *orf355/orf77*, mais également avec des ARN plus courts codés par les gènes *cob* et *atp6* (Wen et Chase, 1999a). Cet effet du restaurateur *Rf3* sur de multiples ARNm peut être le résultat de l'action de la même protéine codée par *Rf3*. Il est également possible que le locus *Rf3* soit constitué de plusieurs gènes dont les produits agissent sur des cibles ARN différentes (Hanson et Bentolila, 2004).

Fujii et al (2011) ont montré que la plupart des angiospermes, où les données de séquences génomiques sont disponibles, contiennent de multiples gènes PPR similaires à des gènes *Rf* : ces gènes *Rf-like* sont regroupés en cluster (Barr et Fishman, 2010) et subissent une sélection diversifiante (Fujii et al., 2011). Même si les gènes *Rf-like* ont des séquences similaires à celles des gènes *Rf*, ils ont des caractéristiques propres qui permettent de les différencier des autres PPR *Rf* et leurs fonctions, si elles existent, restent à démontrer.



**Figure A13 : relation entre les approches utilisées pour étudier la variation génétique**  
 Les flèches relient les approches qui sont complémentaires (d'après Weigel 2012).



**Figure A14 : ressources biologiques disponibles chez *Arabidopsis thaliana***

Les couleurs indiquent la contribution génétique de différentes accessions parentales. Seule une paire de chromosome est représentée pour chaque individu. Les individus HIF dérivent des lignées RIL. Les lignées d'insertion T-DNA ne sont pas représentées (d'après Weigel 2012).

## VII. La plante modèle *Arabidopsis thaliana*

### VII.1 Présentation

*Arabidopsis thaliana*, ou Arabette des Dames, est une plante annuelle appartenant à la famille des Brassicaceae, anciennement nommées Crucifères, qui comprend plus de 3200 espèces réparties en 350 genres : c'est une des plus grande famille d'angiospermes, ou plantes à fleurs, apparues il y a 100 millions d'années sur Terre. Parmi les Brassicaceae figurent des plantes d'intérêt agronomique comme le chou, le navet, le colza et la moutarde. Les caractéristiques morphologiques d'*Arabidopsis* sont les suivantes : la racine est pivotante, les feuilles forment une rosette et les inflorescences sont organisées en corymbe. La fleur blanche a une symétrie bilatérale et comprend 4 sépales, 4 pétales, 6 étamines et 2 carpelles.

*Arabidopsis thaliana* a été adoptée comme plante modèle par les généticiens pour ses caractéristiques biologiques favorables à l'expérimentation génétique : son génome diploïde est petit (de 125 à 150 Mb) et réparti sur cinq chromosomes (environ 30 000 gènes codent des protéines) ; le temps de génération est court (environ trois mois) et un cycle de développement peut s'accomplir entièrement *in vitro* ; sa petite taille (30 à 40 cm de hauteur après floraison) est utile pour des expérimentations nécessitant l'étude de milliers d'individus ; la production de graines par autofécondation est importante (une cinquantaine de graines par fruit ou silique, et plusieurs dizaines de milliers de graines par individu) ; cette plante tolère l'allofécondation, ou fécondation croisée, ce qui permet de faire des croisements entre génotypes. Les taux d'allofécondation dans des populations naturelles d'*Arabidopsis* ont été calculés à différents niveaux : global, régional et local. Ces taux peuvent varier selon les niveaux et les populations étudiées (Bomblies et al., 2010) : ils sont compris entre 0.3 et 2,5% (Abbott et Gomes 1989 ; Picó et al., 2008). À noter qu'aucune population naturelle gynodioïque d'*A. thaliana* n'a été mise en évidence à ce jour.

Les caractéristiques biologiques d'*Arabidopsis* ont amené les généticiens à étudier cette plante dans le cadre de leurs disciplines spécifiques, selon différentes approches résumées dans la figure A13. *Arabidopsis thaliana* peut par exemple permettre de détecter des marques de sélection (approches évolutives), d'identifier des locus, des gènes ou des polymorphismes responsables d'une variation phénotypique (approches de cartographie et de génomique fonctionnelle), et aussi de comprendre les interactions entre les plantes et leur environnement (approches écologiques) (pour revue, Trontin et al., 2011 ; Weigel, 2011 ; voir aussi Weigel et Mott, 2009 ; Koornneef et Meinke, 2010).

De nombreuses ressources génétiques, développées lors des approches mentionnées ci-dessus, sont maintenant disponibles chez *Arabidopsis*. Ces ressources sont des accessions, des populations, des lignées (figure A14) et des mutants. Les accessions d'*Arabidopsis* sont des plantes très majoritairement homozygotes issues d'individus récoltés en milieu naturel, dont les allèles ont été fixés par des autofécondations répétées sur plusieurs générations (méthode SSD pour « Single Seed Descent »).

En plus de populations F1 ou F2, de nombreuses lignées sont disponibles : des lignées d'introgression dans lesquelles un locus parental est fixé dans le fond nucléaire de l'autre individu parental (par rétrocroisement récurrent) ; une soixantaine de populations de plusieurs centaines de lignées RIL (pour Recombinant Inbred Lines) (<http://www.inra.fr/internet/Produits/vast/RILs.htm>, <http://www.arabidopsis.org/> et <http://www.arabidopsis.info/>) dans lesquelles des chromosomes recombinants ont été fixés par autofécondation (Reiter et al., 1992 ; Lister et Dean, 1993) ; des lignées AI-RILs (pour Advanced Intercross-Recombinant Inbred Lines) obtenues à partir d'individus, issus de croisements entre les individus F1 ou F2, dont les génotypes sont ensuite fixés par plusieurs générations d'autofécondation (Balasubramanian et al., 2009) ; des lignées HIF (pour Heterogeneous Inbred Family) qui dérivent de lignées RIL dans lesquelles une petite portion du génome ségrége pour les deux allèles parentaux (Tuinstra et al., 1997) ; les lignées NIL (pour Near Isogenic Lines) sont homozygotes pour l'un des allèles parentaux à un locus dans un fond génétique provenant de l'autre parent. Les lignées NIL dérivent soit de lignées HIF, soit de lignées d'introgression (obtenues par rétrocroisements répétés). D'autres approches plus élaborées impliquent l'utilisation de plusieurs parents comme dans le cas des populations MAGIC (pour Multiple Advanced Generation Intercross) et des populations AMPRIL (pour *Arabidopsis* Multiparent RIL). De nombreuses banques de mutants ont également été créées : pour la grande majorité des gènes, des lignées d'insertion T-DNA dans le fond génétique de référence Col-0 ou Ws sont par exemple disponibles (<http://arabidopsis.org>, <http://arabidopsis.info> ; pour revue, Alonso et Ecker, 2006).

Une stratégie expérimentale, appelée GWAS (pour Genome-Wide Association Study) permet en principe de détecter indirectement un polymorphisme responsable d'un phénotype en comparant les variants génétiques de ce polymorphisme chez différents individus et ainsi voir si le variant est associé au trait étudié (Myles et al., 2009 ; Atwell et al., 2010).



**Figure A15 : distribution géographique d'*Arabidopsis thaliana***

Les zones vertes correspondent aux zones de distribution de l'espèce. Les origines géographiques des accessions collectées disponibles dans les centres de ressources sont indiquées par des points rouges. Fin 2011, 7000 accessions provenant de plantes collectées dans la nature étaient disponibles dans les centres de stockage (d'après Koornneef et al., 2004 ; pour revue, Weigel et al., 2012).

## VII.2 Histoire évolutive d'*Arabidopsis thaliana*

*Arabidopsis thaliana* est originaire d'Asie centrale et d'Europe/Afrique du Nord (revue de Hoffmann 2002 ; Al-Shehbaz et O'Kane, 2002) ; elle est maintenant présente et largement distribuée dans l'hémisphère Nord, de la Scandinavie du Nord à l'Equateur (dans les montagnes de Tanzanie et du Kenya). Elle a également été introduite sur le continent américain et au Japon ; des populations ont aussi été identifiées en Corée du Sud et dans plusieurs pays africains (Alonso-Blanco et Koornneef, 2000) (figure A15). Les populations d'*Arabidopsis* ont dû se disperser après la dernière glaciation du Pléistocène, il y a environ 12 000 ans, à partir de plusieurs habitats locaux utilisés comme refuges lors de la dernière glaciation. Au contraire des refuges d'Europe où les populations d'*Arabidopsis* ont contribué à la colonisation post-glaciaire de l'Europe de l'ouest et de l'Europe du nord, (Picó et al., 2010), plusieurs populations dans les refuges d'Asie sont restées génétiquement isolées des autres populations après la fin de la glaciation. Ainsi, des mutations fixées se sont accumulées spécifiquement dans ces populations. Beck et al., (2008) suggèrent que la différenciation est particulièrement forte entre les refuges d'Asie et les autres zones du globe, ce qui suggère un flux de gènes limité entre l'Asie Centrale et les autres endroits du globe et/ou une forte dérive génétique dans cette région. En se basant sur les données d'ADN nucléaire et chloroplastique, Beck et al (2008) proposent que la zone Caucasienne soit une zone ancestrale pour *Arabidopsis* et suggèrent 4 possibilités pour l'origine des populations d'Asie de l'Est. Beck et al (2008) ont également trouvé que les composants maternels d'*A. suecica* sont très similaires à ceux des métapopulations d'Asie, spécialement en Asie centrale et en Chine, ce qui ajoute un argument supplémentaire pour dire que l'Asie est une zone ancestrale pour *Arabidopsis*.

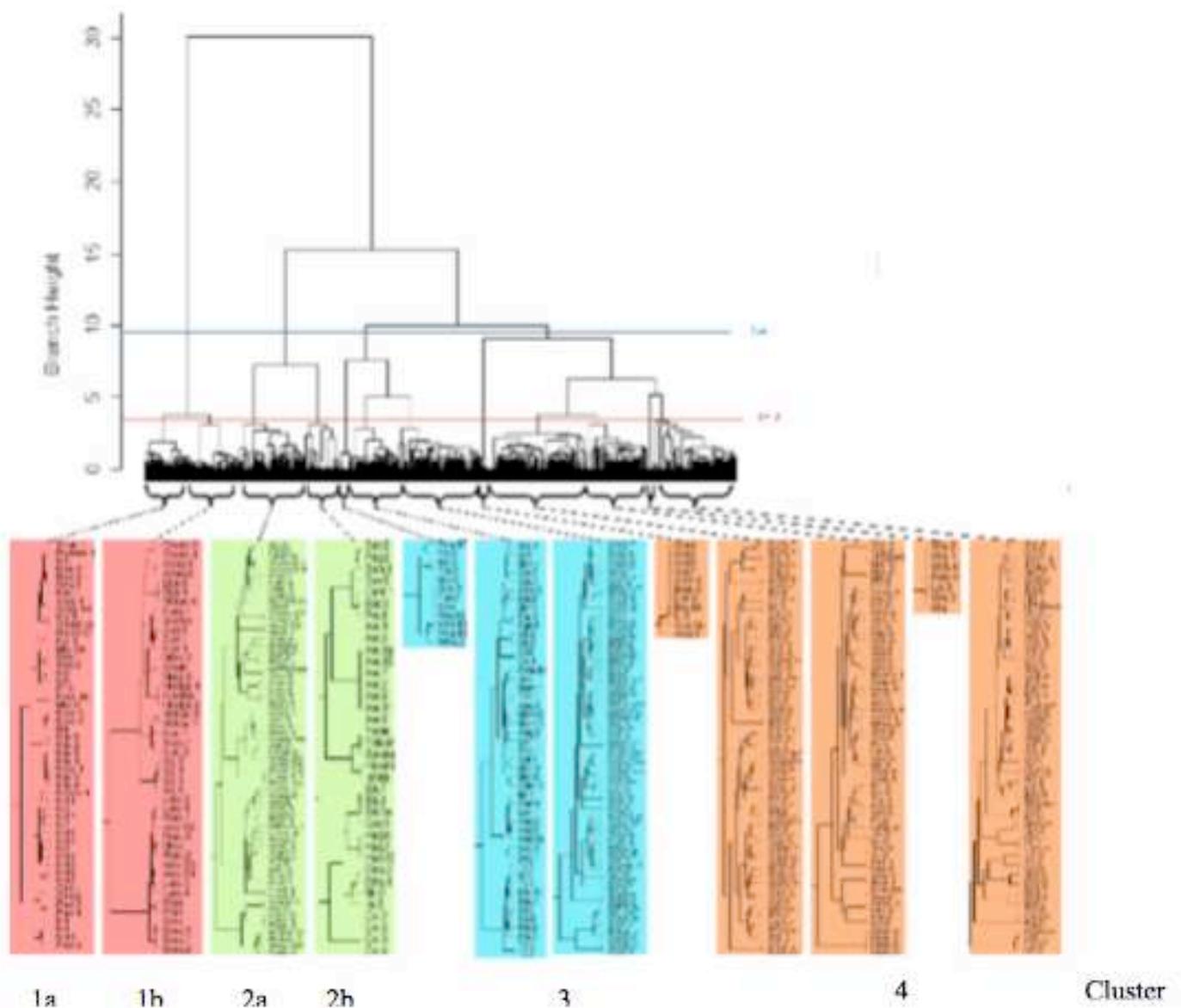
Les accessions d'*Arabidopsis thaliana* sont utilisées pour estimer la diversité génétique chez cette espèce à plusieurs niveaux. À un niveau mondial, des analyses ont mis en évidence une forte structure des populations ainsi qu'une forte isolation par la distance entre les différentes régions du monde (Sharbel et al., 2000 ; Schmid et al., 2006 ; Beck et al., 2008 ; Platt et al., 2010). Plusieurs laboratoires ont initié le développement de nouvelles collections d'*Arabidopsis* pour étudier la variation génétique à un niveau régional dans des zones d'où semble être originaire *Arabidopsis*, comme en Europe (Le Corre 2005 ; Stenoien et al., 2005 ; Bakker et al., 2006 ; François et al., 2008) et en Asie (Schmid et al., 2006 ; He et al., 2007 ; Pico et al., 2008), mais aussi dans des zones où *Arabidopsis* aurait été introduite récemment comme au Japon (Todokoro et al., 1996) et en Amérique du Nord (Jorgensen et Mauricio

2004 ; Bakker et al., 2006). Des études de populations d'*Arabidopsis*, à un niveau local, ont également été menées (Bomblies et al., 2010). Jusqu'à présent, une corrélation significative entre les distances génétiques et géographiques ont été observées uniquement en Asie et en Europe (He et al., 2007 ; Pico et al., 2008). Au contraire, dans les autres régions du monde, ces corrélations n'existent pas, ce qui est interprété soit comme une conséquence d'une colonisation récente et d'une expansion à partir de plusieurs refuges, soit comme la conséquence de colonisations multiples (Todokoro et al., 1996 ; Jorgensen et Mauricio 2004).

### VII.3 Diversité du génome nucléaire des accessions d'*Arabidopsis thaliana*.

Le séquençage des génomes est un moyen efficace de décrire la diversité des génomes. *Arabidopsis* est la première plante dont le génome nucléaire, celui de l'accèsion Col-0, a été entièrement séquencé en 2000 ; des génomes d'autres accessions ont été depuis séquencés (Gan et al., 2011) et le projet 1001 génomes permettra bientôt d'avoir accès aux variations de séquence pour un millier d'accèsions (Cao et al., 2011). Lors de la première phase de ce projet, les génomes nucléaires de 80 accessions d'*Arabidopsis* ont été séquencés, dont ceux d'accèsions d'Europe et d'Asie centrale, ce qui a permis de mettre en évidence des SNP (pour Single Nucleotide Polymorphism), des insertions/délétions et des variations dans les éléments transposables. Cao et al (2011) ont identifié, parmi ces 80 accessions, 12 468 SNP répartis sur 6197 gènes, capables d'altérer des codons d'initiation, d'introduire des codons stop, d'étendre une ORF déjà existante ou d'affecter les sites accepteurs et donneurs d'épissage. Dans cette même étude, 27 167 sites d'insertion/délétion ont été identifiés et ont le potentiel d'induire un décalage du cadre de lecture (cependant, plusieurs événements de ce type peuvent se combiner pour restaurer le bon cadre de lecture). Finalement, 4 525 délétions d'au moins 50pb, dans des séquences codantes de 2 247 gènes, ont été mises en évidence. Des régions dupliquées et des séquences non référencées ont également été identifiées chez *Arabidopsis* (Cao et al., 2011). Ces données indiquent qu'il existe une grande diversité nucléaire chez *Arabidopsis thaliana*. Des 'core-collections' ont été créés au sein d'*A. thaliana*. Une core-collection regroupe une petite partie d'une collection plus étendue ; elle est définie pour contenir le maximum de diversité génétique possible avec le minimum de répétitions (McKhann et al., 2004). Cela permet d'étudier la diversité génétique et phénotypique au sein de la collection avec un petit pool d'individus représentatifs.

Dans notre laboratoire, Simon et al (2011) ont utilisé 341 marqueurs SNP nucléaires pour



**Figure A16 : résultats de clustering d'accessions d'*A. thaliana* réalisé avec des marqueurs nucléaires SNP (Simon et al., 2011)**

Dendrogramme obtenu à partir d'une matrice de 598 lignées et 341 marqueurs SNP. La ligne bleue coupe les clades pour un nombre de clusters  $K = 4$ . La ligne rouge coupe les clades pour un nombre de clusters  $K = 12$ , ce qui correspond à une analyse plus détaillée du clustering. Le clustering des accessions contenues dans chacun des 12 groupes sont détaillées dans les cadres colorés.

génotyper 598 accessions d'Arabidopsis. Le résultat du clustering réalisé à partir de ces données, est présenté dans la figure A16. Son analyse a confirmé la forte structuration de la population d'Arabidopsis où la plupart des lignées qui ont des génotypes très proches ont des origines géographiques communes (Platt et al., 2010). Une interface web (ANATool) a été créée (<https://www.versailles.inra.fr/ijpb/crb/anatool>) où sont disponibles les données de variations moléculaires pour les accessions analysées.

#### VII.4 Diversité des génomes cytoplasmiques des accessions d'*Arabidopsis thaliana*.

Alors qu'une quantité impressionnante de données sont disponibles sur la diversité génomique nucléaire de cette espèce, peu de données ont été publiées sur la diversité cytoplasmique. Ce sont principalement des analyses des génomes chloroplastiques qui ont été entreprises pour étudier la diversité cytoplasmique chez Arabidopsis. En essayant de connaître l'origine maternelle d'*A. suecica*, Sall et al (2003) ont séquencé 12 régions d'ADN chloroplastique de 25 accessions *A. thaliana* qui proviennent principalement de Scandinavie. Ils ont trouvé, comme attendu, une variation considérable parmi les séquences intergéniques du génome chloroplastique d'Arabidopsis. Le polymorphisme chloroplastique a également été utilisé dans le cadre d'études phylogéographiques de populations de la Péninsule Ibérique (Pico et al., 2008) et de Chine (He et al., 2007 ; Yin et al., 2010). Yin et al. (2010) ont analysé les séquences d'ADN chloroplastique de 77 accessions, dont 19 accessions provenant des collections de Chine, ce qui leur a permis d'identifier 49 haplotypes. Par exemple, les populations chinoises en bordure de la rivière Yang-Tse sont regroupées dans 5 haplotypes et les 2 accessions collectées dans les montagnes Altaï de Chine présentent 6 sites variables spécifiques. Dans une autre étude (Beck et al., 2008), les régions intergéniques *trnL-trnF* de 475 individus en provenance de 167 populations d'Arabidopsis ont été séquencées et 16 haplotypes ont été identifiés.

L'organisation complexe des génomes mitochondriaux empêche d'envisager la même stratégie que celle adoptée pour les chloroplastes pour révéler du polymorphisme. La diversité mitochondriale des accessions d'*Arabidopsis thaliana* a été étudiée selon différentes approches : Ullrich et al. (1997) ont étudié par Southern le polymorphisme mitochondrial de 23 accessions d'Arabidopsis. Des profils différents d'une accession à l'autre ont été obtenus : ces différences sont de nature stœchiométrique ou dues à des réarrangements mitochondriaux spécifiques comme par exemple pour l'accession *Ler* où il manque une région mitochondriale

Région	MatK-trnK	NdhC-trnV	rbcL-accD	ndhF-rpl32	Total
Coordonnées	3378-4340	50289-51242	56283-57204	112547-113489	
Taille (pb)	850	840	810	830	3330
SNP	5	18	6	10	39
In/Del	2	7	1	4	14
Inversion	0	0	1	0	1
MSAT	2	3	3	6	14
Total	9	28	11	20	68

**Tableau A1 : sites polymorphes chloroplastiques identifiés chez *A. thaliana* (Moison et al., 2010)**

Les quatre régions chloroplastiques de 95 accessions ont été séquencées par Moison et al. (2010). Les coordonnées se réfèrent aux régions de l'ADN chloroplastique de Col-0.

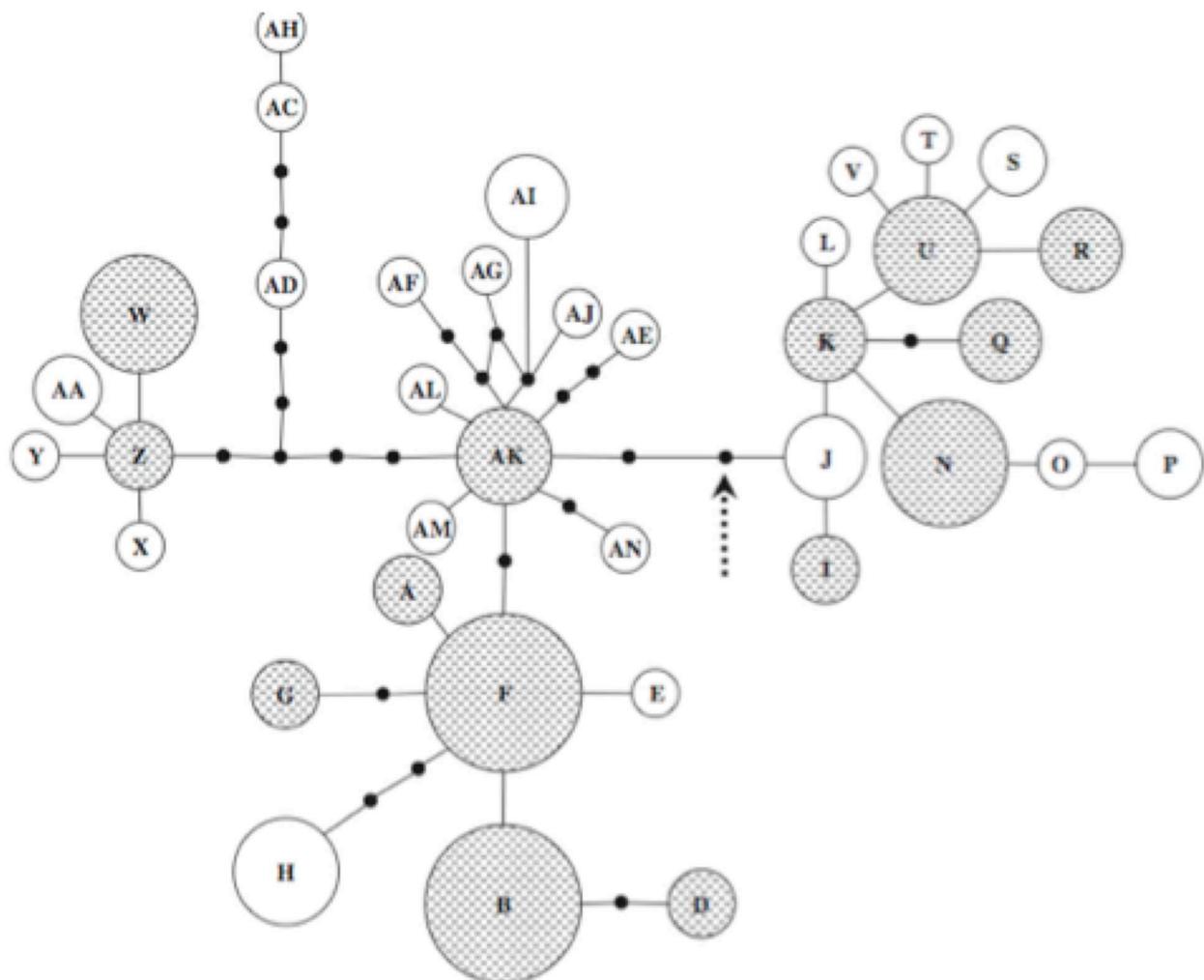
Région	Coordonnées	Type de polymorphisme	Référence	Stratégie
nad5-rnr26	11944-13795	Recombinaison	(Ullrich et al., 1997)	PCR directe
nad5-rnr26	16606-18591	Recombinaison	(Ullrich et al., 1997)	PCR directe
orf131	16965-17203	Présence/absence	(Moison et al., 2010)	PCR directe
orf315	16959-17451	Présence/absence	(Moison et al., 2010)	PCR directe
nad9-rpl16	24147-25309	In/Del 6 pb	(Moison et al., 2010)	PCR-RFLP
cox2	39529-40374	SNP	(Moison et al., 2010)	PCR-RFLP
atp8-orf107c	130231-131125	SNP	(Moison et al., 2010)	PCR-RFLP
orf240a	203703-204264	Présence/absence	(Moison et al., 2010)	PCR directe
orf120	207558-207869	Présence/absence	(Moison et al., 2010)	PCR directe
orf107d	209506-209789	Présence/absence	(Moison et al., 2010)	PCR directe
cox3/atp9	217884/278985	Réarrangement	(Forner et al., 2008)	PCR directe
ccmC	240858-241488	Multiplés SNP	(Forner et al., 2008 ; Moison et al., 2010)	PCR-RFLP + séquençage
orf262	279543-280170	Présence/absence	(Moison et al., 2010)	PCR directe
atp6-2	297337-298283	Présence/absence	(Forner et al., 2008)	PCR directe
orf111b	301020-301272	Présence/absence	(Moison et al., 2010)	PCR directe

**Tableau A2 : polymorphismes testés dans l'ADN mitochondrial des accessions d'*A. thaliana* (Moison et al., 2010)**

Le nom des 15 marqueurs mitochondriaux est décrit dans la première colonne. La seconde colonne donne la position de ces marqueurs sur le génome mitochondrial de l'accèsion C24 et la troisième colonne, le type de polymorphisme observé pour chacun des marqueurs. La stratégie utilisée pour révéler le polymorphisme à chaque marqueur est décrite dans la dernière colonne.

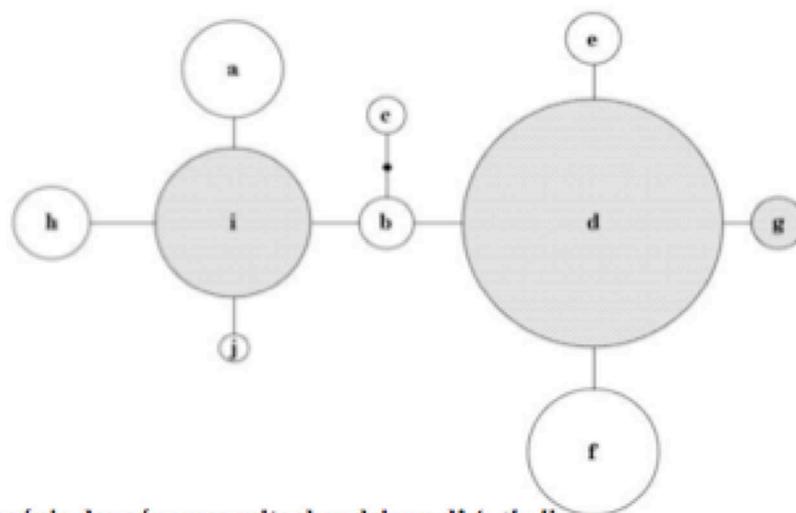
de 5 kb (Ullrich et al., 1997). Binder et Forner se sont particulièrement intéressés aux polymorphismes des ARNm mitochondriaux chez différentes accessions d'*Arabidopsis* (Forner et al., 2005 ; Forner et al., 2008). Dans un premier temps, l'environnement génomique et les transcrits du gène mitochondrial *cox3* ont été étudiés chez Col-0, C24 et *Ler* (Forner et al., 2005) : une insertion de 1,8 kb, présente dans la région en amont du gène *cox3* chez Col-0, est absente chez C24. De plus, des analyses en Southern blot, en utilisant des sondes qui correspondent à la partie 5' du cadre de lecture de *cox3*, ont permis d'identifier des environnements génomiques en amont de *cox3* différents entre les accessions. Pour étudier l'influence de ces réarrangements sur l'expression, les niveaux d'ARN du gène *cox3* ont été analysés par Northern blot et ont révélé que la taille des ARNm *cox3* pouvait varier en fonction des accessions, principalement du fait de la longueur des régions 5' non traduites qui peuvent être de tailles différentes selon les accessions analysées (Forner et al., 2005). Dans une étude plus récente, Forner et al (2008) ont comparé les extrémités 5' et 3' des transcrits mitochondriaux de tous les gènes qui codent des protéines mitochondriales chez Col-0, C24 et *Ler*. En plus des transcrits du gène *cox3* (Forner et al., 2005), les extrémités 5' des transcrits des gènes *nad4*, *nad9*, *ccmB*, *rpl5*, *ccmC*, *atp6-2* diffèrent selon les accessions étudiées (Forner et al., 2008). Ces profils différents peuvent être dus à des facteurs mitochondriaux ou à des facteurs nucléaires différents entre les accessions. Des croisements réciproques ont permis de savoir si les phénotypes moléculaires (profils des transcrits) observés étaient dus à des facteurs cytoplasmiques ou à des facteurs nucléaires : pour les gènes *nad4*, *nad9*, *ccmB* et *rpl5*, les profils des transcrits sont identiques entre les F1 des croisements réciproques respectifs, ce qui indique que les polymorphismes des transcrits de ces quatre gènes ont une hérédité biparentale, typique de facteurs nucléaires. Pour *ccmC* et *atp6-2*, les profils des transcrits diffèrent dans les F1 des croisements réciproques respectifs. Il semblerait que les polymorphismes observés pour ces ARNm soient hérités de façon maternelle, et dus à un élément en *cis* situé dans l'ADNmt qui diffère selon les accessions (pour *ccmC*) ou à l'absence du gène en question dans l'une des accessions parentales (pour *atp6-2*) (Forner et al., 2008).

Plus récemment, dans une étude réalisée au sein du laboratoire, Moison et al (2010) ont décrit et évalué la diversité chloroplastique et mitochondriale de 95 accessions dont 48 accessions provenant de la plus grande core-collection constituée par McKhann et al. (2004). Pour explorer la diversité cytoplasmique des 95 accessions d'*Arabidopsis*, Moison et al (2010) ont utilisés des marqueurs chloroplastiques (tableau A1) et mitochondriaux (tableau A2). Les



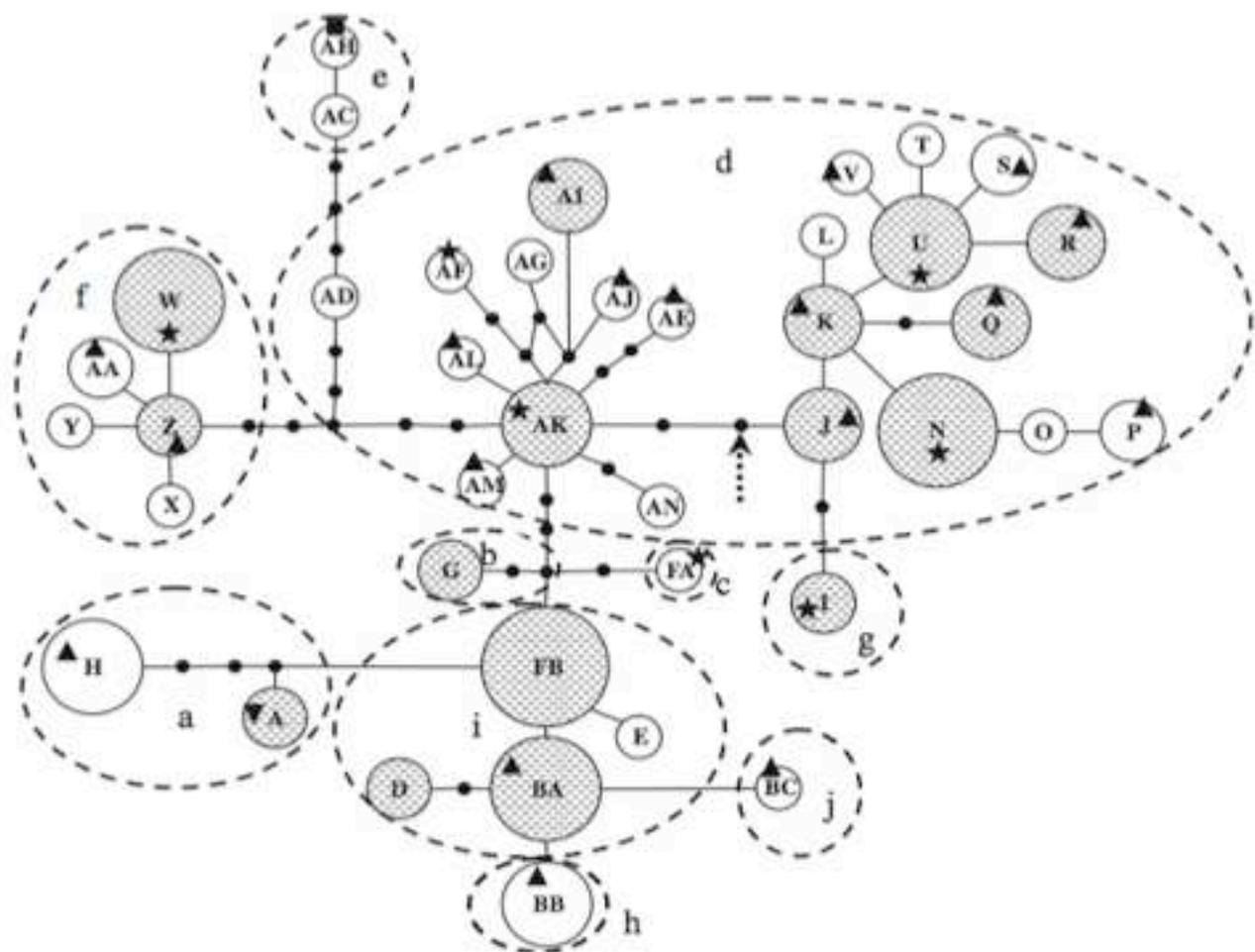
**Figure A17 : phylogénie des génomes chloroplastiques d'*A. thaliana***

Chaque chlorotype est représenté par un cercle dont l'aire est proportionnelle au nombre d'accèsions appartenant au chlorotype dans l'étude de Moison et al (2010). Chaque chlorotype est nommé avec une lettre capitale. Les cercles grisés correspondent aux chlorotypes regroupant plusieurs haplotypes qui diffèrent seulement par la longueur de certains MSAT. Chaque segment entre les cercles représente un polymorphisme ; la longueur du segment n'a pas de signification particulière. Les points noirs représentent des chlorotypes intermédiaires hypothétiques qui n'ont pas été observés. La flèche pointillée indique la racine du réseau (d'après Moison et al., 2010).



**Figure A18 : phylogénie des génomes mitochondriaux d'*A. thaliana***

La légende est la même que dans la figure A17 en remplaçant « chlorotype » par « mitotype ». Chaque mitotype est nommé avec une lettre minuscule. Les cercles grisés correspondent aux mitotypes regroupant plusieurs haplotypes qui diffèrent par l'état stœchiométrique d'un réarrangement (d'après Moison et al., 2010).



**Figure A19 : phylogénie des génomes cytoplasmiques d'*A. thaliana***

La légende est la même que dans la figure 17. Chaque cytotype est nommé avec une lettre capitale qui correspond au chlorotype (sauf pour le chlorotype F, divisé en FA et FB, et le chlorotype B, divisé en BA, BB et BC). Les cercles grisés contiennent plusieurs haplotypes chloroplastiques ou mitochondriaux, selon les mêmes critères que dans les figures 17 et 18. Les cercles pointillés regroupent les cytotypes qui ont le même mitotype. Les étoiles indiquent les groupes qui contiennent au moins une accession de la petite core-collection (n=8) de McKhann et al. (2004). Les triangles indiquent les groupes qui contiennent au moins une accession de la grande core-collection (n=48) de McKhann et al. (2004). Le triangle orienté pointe vers le bas indique le groupe qui contient l'accèsion Col-0 (référence pour le génome chloroplastique) et le carré indique le groupe qui contient C24 (référence pour le génome mitochondrial) (d'après Moison et al., 2010).

marqueurs chloroplastiques ont été identifiés en utilisant quatre couples d'oligos qui amplifient les régions intergéniques *ndhF/rpl*, *matK/trnK*, *ndhC/trnV* et *rbcL/accD* du génome chloroplastique. Les séquences de ces régions ont permis à Moison et al (2010) de révéler 57 sites polymorphes entre les accessions, de type insertion/délétion, SNP et inversion. Les polymorphismes de ce type ont permis de définir 38 chlorotypes et de créer un réseau de relations phylogénétiques entre ces chlorotypes, par la méthode de parcimonie et avec une limite de connexion entre les chlorotypes de cinq étapes (figure A17). Les polymorphismes de type MSAT ont été écartés de l'analyse car ils varient trop fréquemment par rapport à l'échelle de temps considérée (Provan et al., 2001). Pour enraciner le réseau phylogénétique, Moison et al. (2010) ont séquencé les mêmes régions intergéniques de deux espèces apparentées à *A. thaliana* (*A. lyrata* et *A. arenosa*). Le chlorotype J apparaît être le plus proche du génome ancestral chloroplastique d'*A. thaliana*.

Pour analyser la diversité mitochondriale des accessions d'Arabidopsis, 15 marqueurs mitochondriaux, présentés dans le tableau A2, ont été utilisés par Moison et al (2010). Deux types de marqueurs ont été utilisés : des marqueurs obtenus à partir d'amorces dessinées sur des régions conservées de l'ADN mitochondrial (8 couples d'amorces), et des marqueurs obtenus à partir d'amorces dessinées pour amplifier des ORF non conservés (7 couples d'amorces). Seuls les polymorphismes de type SNP, insertion/délétion et réarrangements génomiques qui ne présentent pas de variations stœchiométriques lors des analyses ont été pris en compte pour la définition des mitotype et la construction d'un réseau phylogénétique. Comme pour le réseau chloroplastique, un réseau mitochondrial, composé de 10 mitotypes, a été réalisé (figure A18). Dans la majorité des cas, les accessions qui possèdent le même chlorotype ont le même mitotype. Les réseaux chloroplastiques et mitochondriaux sont très congruents, ce qui indique que les génomes cytoplasmiques sont en fort déséquilibre de liaison, comme cela a déjà été vu chez *Beta vulgaris* (Desplanque et al., 2010). Ce résultat était attendu car les deux génomes sont à hérédité maternelle et semblent former une seule unité évolutive. Les informations des deux génomes ont donc été combinées pour créer un réseau phylogénétique cytoplasmique présenté dans la figure A19. À noter que Moison et al (2010) observent une absence de structuration géographique avec les marqueurs cytoplasmiques utilisés. Pour estimer si une co-adaptation entre les génomes nucléaire et cytoplasmique existe chez *A. thaliana*, ils ont testé la capacité germinative de 27 paires de populations réciproques F2 et ont trouvé que le cytoplasme donneur pouvait avoir un effet significatif sur la capacité germinative ; ceci est interprété comme une co-adaptation

nucléocytoplasmique, pour des traits liés à la germination, chez *Arabidopsis thaliana* (Moison et al., 2010).

## PRESENTATION ET OBJECTIFS DE LA THESE

Un croisement intra-spécifique entre deux accessions d'*Arabidopsis thaliana* (Sha et Mr-0) a donné 100% d'individus mâles stériles dans la descendance F1 lorsque Sha est utilisé comme parent maternel et Mr-0 comme parent mâle (Sha x Mr-0). Dans le croisement réciproque (Mr-0 x Sha), les individus F1 sont mâles fertiles. Étant donné que l'hérédité du cytoplasme chez *Arabidopsis thaliana* est maternelle (non Mendélienne), cela indique que cette stérilité mâle résulte d'une interaction entre le cytoplasme de Sha et les allèles nucléaires de Mr-0. Une cartographie génétique d'une population en ségrégation a également permis, au cours d'un travail préliminaire à cette thèse (Sophie Hiard et Cezary Waszczak), de localiser les locus nucléaires sur les chromosomes 1 et 3. La présence des allèles Mr-0 au deux locus est nécessaire pour observer le phénotype de stérilité mâle. Les objectifs de thèse sont de décrire le phénotype de stérilité mâle, d'identifier les facteurs cytoplasmique(s) et nucléaires, et de commencer l'étude des mécanismes moléculaires sous-jacents à l'origine de cette stérilité mâle

# **RESULTATS**

**a**

**Chr1**

Nom du marqueur

Position du marqueur (Mb)

T1G11	MSAT1.14	C1W12	MSAT1.4	NGA128	MSAT1.13
1,24	5,02	9,62	14,16	20,63	25,83

**Chr2**

Nom du marqueur

Position du marqueur (Mb)

MSAT2.5	MSAT 2.28	MSAT 2.17	CZSOD 2	MSAT 2.10	MSAT2.22
0,21	6,41	10,73	12,02	18,02	19,63

**Chr3**

Nom du marqueur

Position du marqueur (Mb)

MSAT3.02052	MSAT3.32	MSAT3.58	MSAT3.70
2,05	11,21	18,66	23,45

**Chr4**

Nom du marqueur

Position du marqueur (Mb)

MSAT4.00986	MSAT4.35	C1W7	MSAT4.14
0,98	7,55	11,52	15,21

**Chr5**

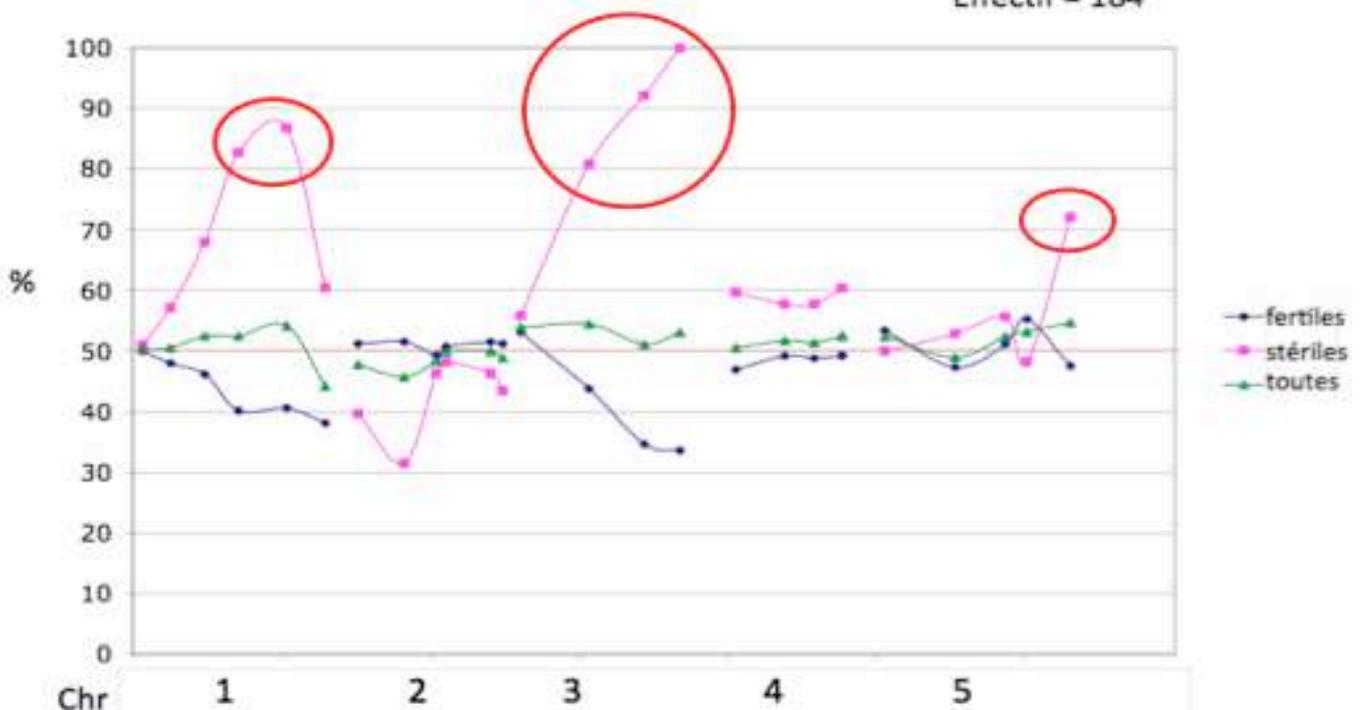
Nom du marqueur

Position du marqueur (Mb)

ATHCTR1A	NGA76	MSAT 5.8	MSAT5.20037	MSAT5.19
0,98	10,42	17,16	20,03	25,92

**b**

Effectif = 184



**Figure B1 : analyse de la population « rétrocroisement maternel 1 » (RM1)**

a- Noms et les positions des marqueurs utilisés pour le génotypage, par chromosome.

b- Représentation graphique des résultats de génotypage. Le pourcentage d'individus hétérozygotes aux marqueurs analysés, est indiqué pour l'ensemble de la population (N=188) en vert ; pour les plantes fertiles (N= 134) en bleu ; pour les plantes stériles (N= 54) en rose. Les cercles rouges entourent les résultats s'écartant significativement ( $\alpha=0,05$ ) de la proportion attendue.

**a****Chr1**

Nom du marqueur

Position du marqueur (Mb)

T1G11	NGA63	CIW12	MSAT 1.4	CIW 1	NGA128	MSAT1.13
1,24	3,22	9,62	14,2	18,37	20,63	25,83

**Chr2**

Nom du marqueur

Position du marqueur (Mb)

MSAT 2.5	MSAT 2.28	MSAT 2.17	CZSOD 2	MSAT 2.10	MSAT 2.22
0,21	6,41	10,73	12,02	18,02	19,63

**Chr3**

Nom du marqueur

Position du marqueur (Mb)

MSAT 3.0205	MSAT 3.32	MSAT 3.58	MSAT 3.2043	MSAT 3.70
2,05	11,21	18,66	20,43	23,45

**Chr4**

Nom du marqueur

Position du marqueur (Mb)

MSAT 4.0098	MSAT 4.35	C1W7	MSAT 4.18	MSAT 4.14
0,98	7,55	11,52	11,97	15,21

**Chr5**

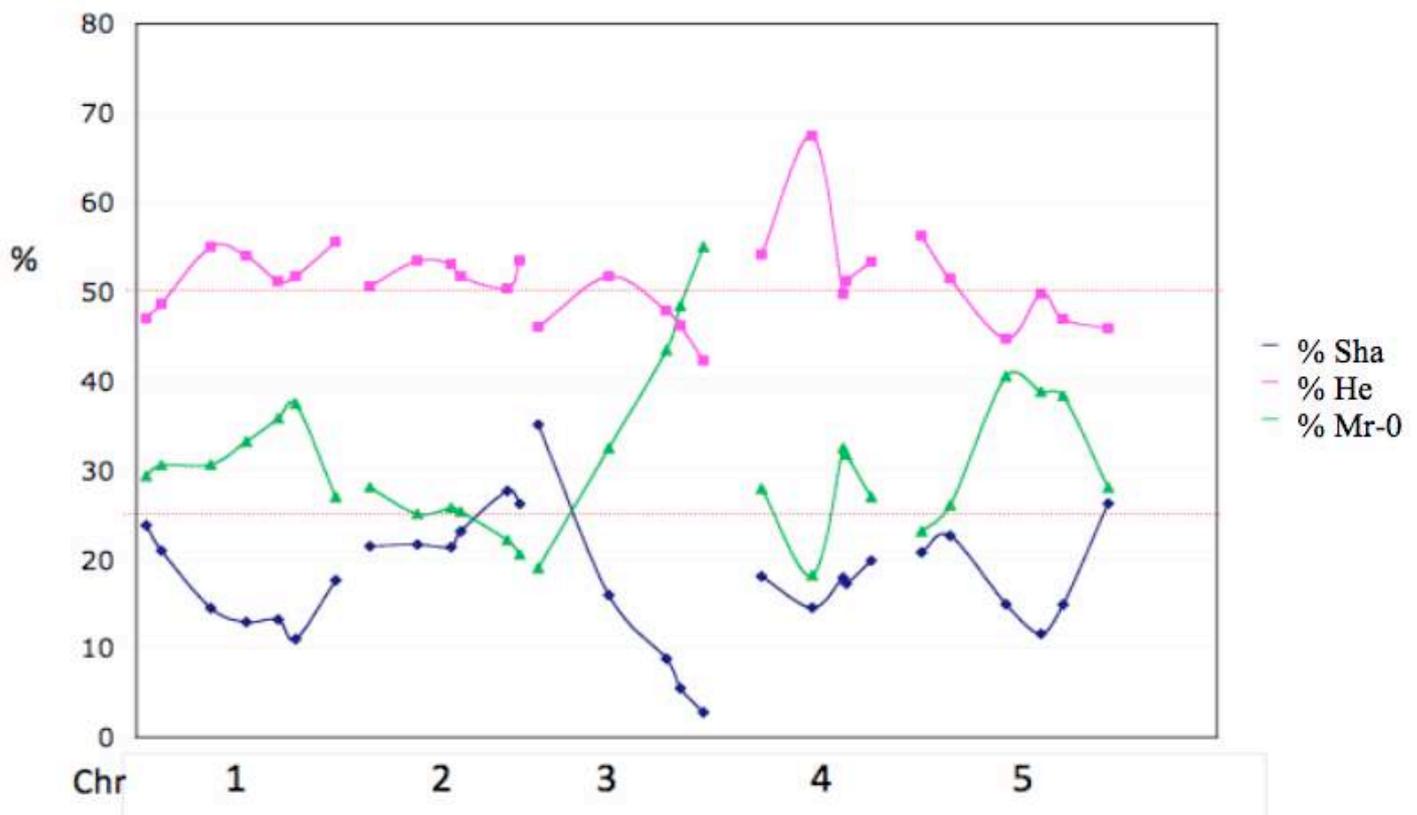
Nom du marqueur

Position du marqueur (Mb)

NGA 158	ICE1	MSAT 5.1261	MSAT 5.8	MSAT 5.2003	MSAT 5.19
1,7	5,4	12,61	17,16	20,03	25,92

**b**

Effectif = 184

**Figure B2 : analyse de la population F2-63RV**

a. Noms et les positions des marqueurs utilisés pour le génotypage, par chromosome.

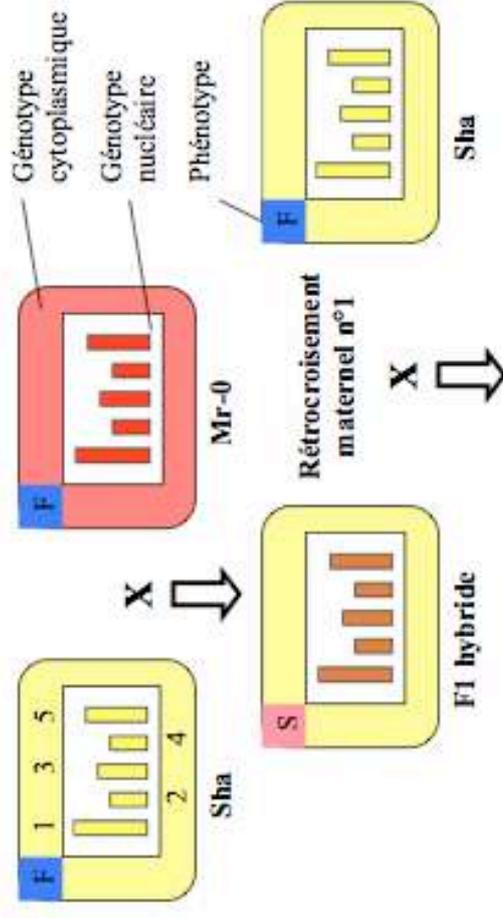
b. Représentation graphique des résultats de génotypage. Le pourcentage de chaque génotype est indiqué (Mr-0 : bleu, Sha : vert, hétérozygote : rose) pour chaque marqueur selon la position physique.

## I Cartographie génétique des mainteneurs de stérilité

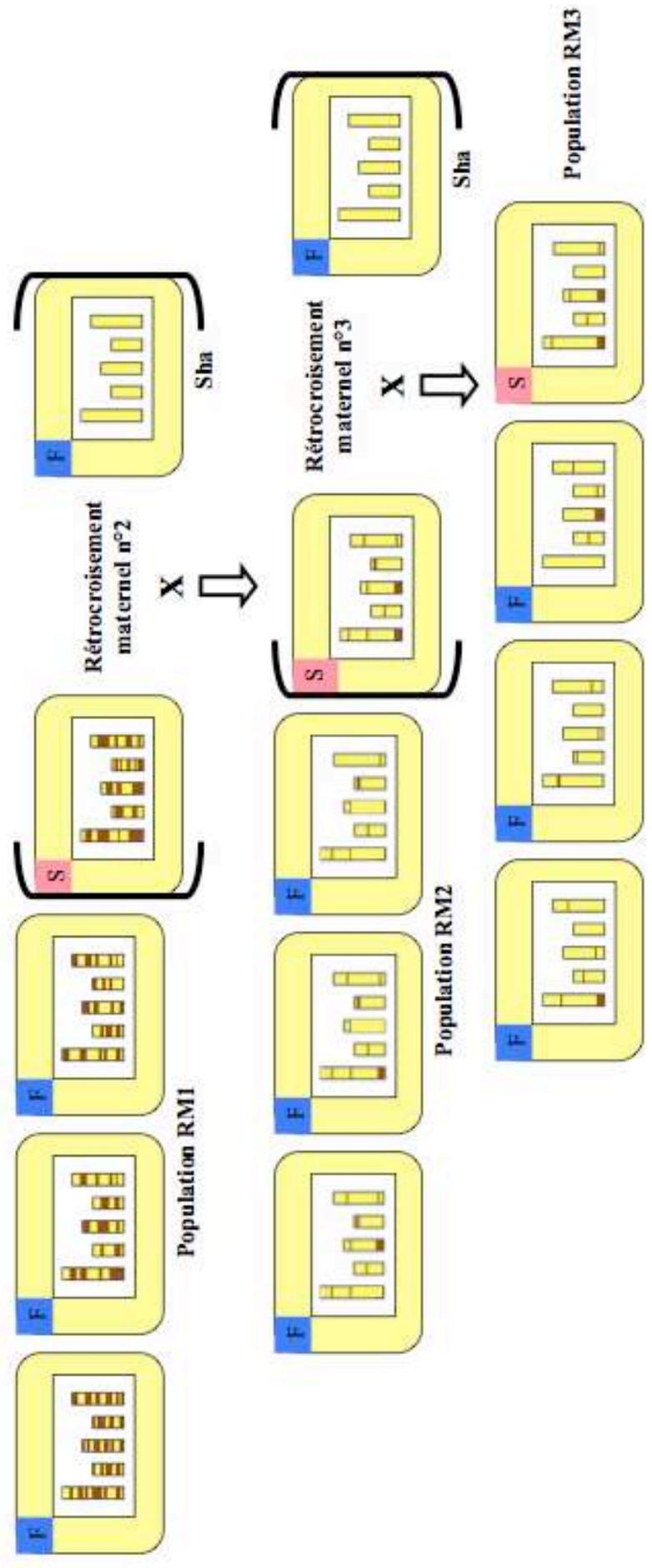
### I.1 Résultats préliminaires

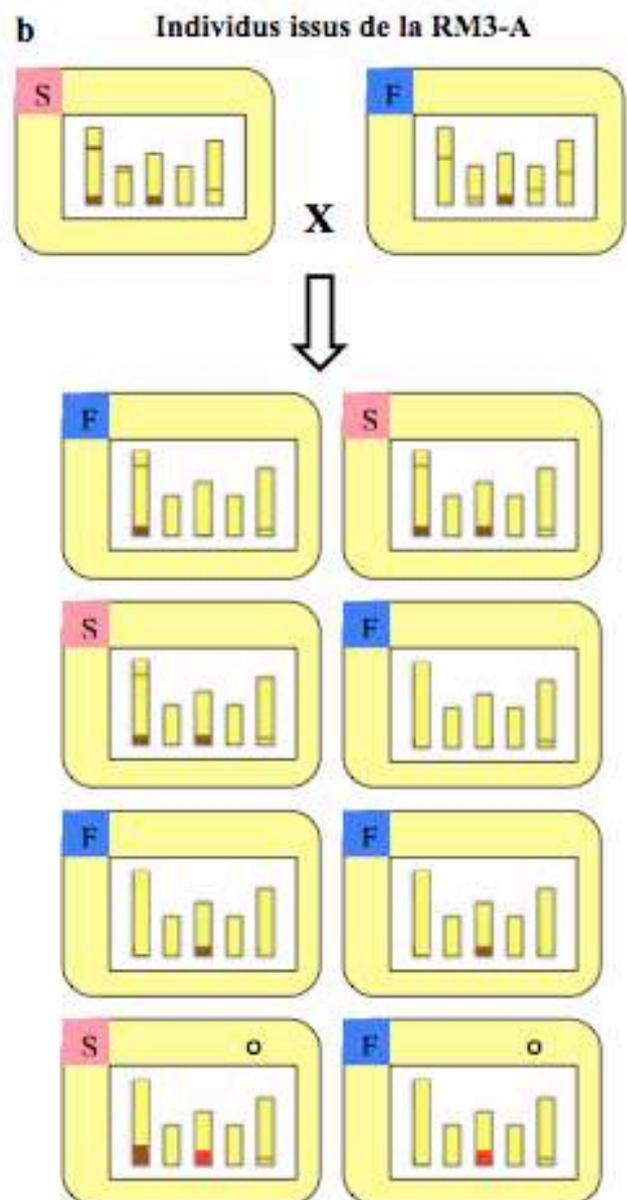
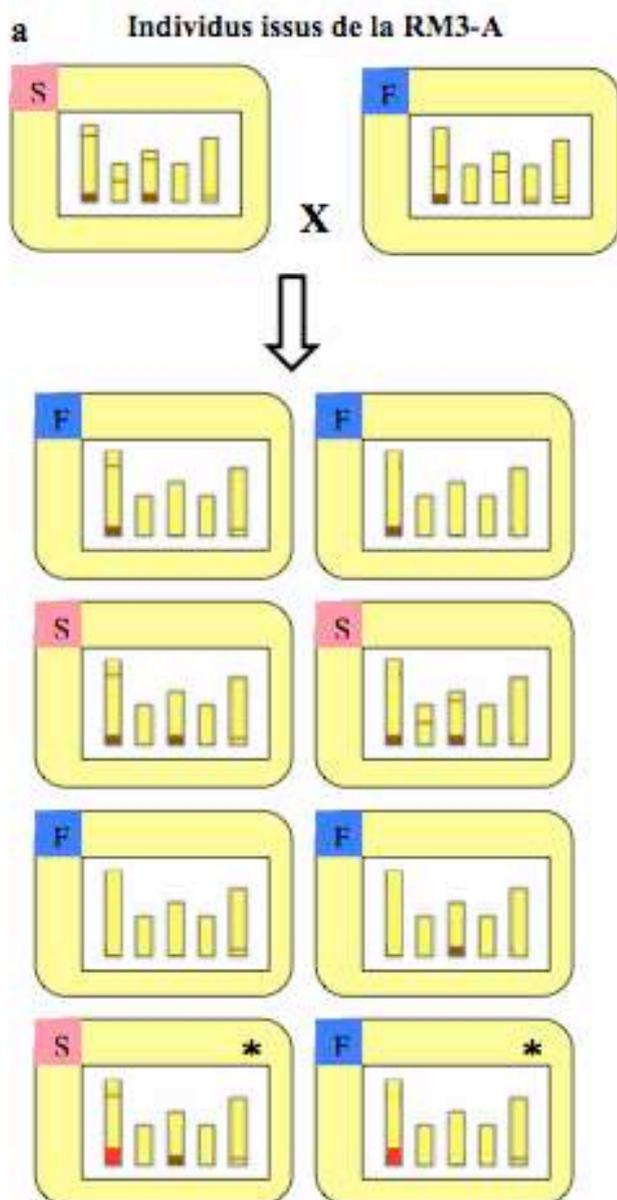
Lorsque j'ai commencé ce travail, des résultats préliminaires avaient été obtenus par Sophie Hiard et Cezary Waszczak sur l'analyse génétique des mainteneurs de stérilité : un premier rétrocroisement des individus F1 mâle stériles par le parent maternel Sha a révélé une proportion d'environ  $\frac{3}{4}$  d'individus fertiles (F) et  $\frac{1}{4}$  d'individus stériles (S) dans la population RM1 (pour rétrocroisement maternel 1), ce qui suggère que deux locus nucléaires indépendants sont nécessaires au maintien de la stérilité. Les allèles mainteneurs sont dominants aux deux locus. Afin de définir les régions génomiques contenant les deux locus d'intérêt, un génotypage des individus de la population RM1 phénotypée (184 individus) a été réalisé à l'aide de 25 marqueurs MSAT localisés sur les 5 chromosomes (figure B1a). La figure B1b présente le pourcentage d'hétérozygotes, pour chaque marqueur, selon le phénotype dans la population RM1. Ils ont constaté que la majorité des plantes stériles étaient hétérozygotes au sud du chromosome 1 (près de 90 % au marqueur NGA128) et hétérozygotes au sud du chromosome 3 (92% et 100% aux marqueurs MSAT3.58 et MSAT3.70 respectivement). Ces résultats semblent indiquer que le premier locus, est situé sur le bras sud du chromosome 1, et que le second est situé sur le bras sud du chromosome 3. Ils ont été appelés respectivement locus ch1 et locus ch3. Il existe également un excès de plantes hétérozygotes au dernier marqueur du chromosome 5 (MSAT5.19) parmi les plantes stériles. Une analyse plus approfondie a montré qu'il existe un déséquilibre de liaison entre le bas du chromosome 5 et le bas du chromosome 3.

La descendance en autofécondation de la F1 fertile du croisement réciproque Mr-0 x Sha, nommée F2-63RV, a également été semée et génotypée (figure B2). Cette population présente des distorsions de ségrégation à certains marqueurs, c'est-à-dire, une déviation de la fréquence observée par rapport aux ratios mendéliens attendus dans cette population. Dans la population F2-63RV, la ségrégation des allèles à certains marqueurs s'écarte significativement de celle attendue ( $\frac{1}{4}$  homozygotes Sha,  $\frac{1}{2}$  hétérozygotes et  $\frac{1}{4}$  homozygotes Mr-0) : on observe un déficit des génotypes homozygotes Sha à des marqueurs situés au sud des chromosomes 1 et 3 et au chromosome 5, et un excès d'hétérozygotes au chromosome 4 (figure B2). De plus, l'analyse de ces résultats confirme un déséquilibre de liaison entre le marqueur MSAT5.19 (dont la ségrégation n'est pas biaisée dans cette population) et le marqueur MSAT3.70. Il est donc probable que plusieurs incompatibilités alléliques entre 2 ou plusieurs locus contribuent à biaiser la transmission des allèles parentaux dans cette



**Figure B3 : introgression des allèles Mr-0 mainteneurs de stérilité mâle**  
Plan de croisement utilisé pour introgresser les allèles Mr-0 mainteneurs de stérilité dans un fond génétique Sha : trois rétrocroisements par le parent récurrent Sha ont permis de produire trois populations « rétrocroisement maternel » (nommées RM1, RM2 et RM3). Dans la RM3, les individus stériles ont, en théorie, les deux locus mainteneurs de stérilité à l'état hétérozygote et se trouvent dans un fond nucléaire et cytoplasmique très majoritairement Sha. Chaque boîte blanche correspond à un génotype. Les cinq barres verticales dans le rectangle blanc représentent les 5 chromosomes. La couleur du fond de chaque boîte indique le cytotypage. En rouge : génotype Mr-0 ; en jaune, génotype Sha ; en marron : génotype hétérozygote. Le phénotype de chaque plante est également indiqué : F = fertile, bleu et S = stérile, rose.





**Figure B4 : fixations des allèles Mr-0 à l'état homozygote au locus ch1 ou au locus ch3.**

(a) Pour créer une population « fixation Mr-0 au locus ch1 », deux individus hétérozygotes au locus ch1 ont été sélectionnés dans la population « rétrocroisement maternel n°3 » (RM3-A) : l'un des individus est stérile (S), car hétérozygote aux locus ch1 et ch3, et l'autre est fertile (F) car Sha au locus ch3. La descendance de ce croisement, appelée population « fixation Mr-0 au locus ch1 » contient donc des individus Mr-0 au locus ch1 et hétérozygote ou Sha au locus ch3 : ces individus (identifiés par \* dans la figure) seront utilisés comme parents femelles et comme parents mâles respectivement pour générer des populations de cartographie au locus ch3.

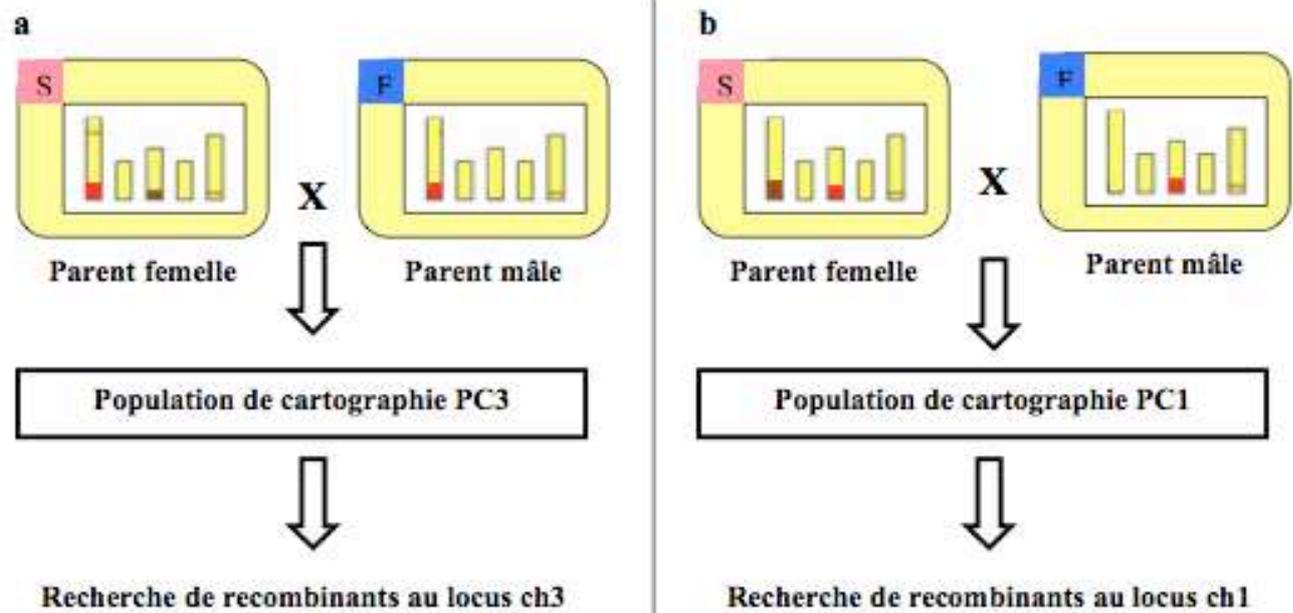
(b) La même stratégie a été effectuée afin d'éviter la ségrégation du locus ch3 mainteneur de stérilité lors de la recherche de recombinants au locus ch1. Les individus sélectionnés dans la population « fixation Mr-0 au locus ch3 » pour produire des populations de cartographie pour le locus ch1 sont identifiés par ° dans la figure. Chaque boîte correspond à un génotype. Les cinq barres verticales dans le rectangle blanc représentent les 5 chromosomes. La couleur du fond de chaque boîte indique le cytotype. En rouge : génotype Mr-0 ; en jaune, génotype Sha ; en marron : génotype hétérozygote. Le phénotype de chaque plante est également indiqué : F = fertile, bleu et S = stérile, rose.

population. Mon travail de thèse a commencé en ayant connaissance de ces résultats préliminaires.

## I.2 Stratégie et plans de croisements pour la création de populations de cartographie

Les incompatibilités alléliques entre les allèles Sha et Mr-0 peuvent interférer avec la cartographie des gènes impliqués dans le maintien de la stérilité mâle, ce qui rendrait difficile l'analyse génétique des mainteneurs de stérilité dans certaines populations. Pour supprimer ces incompatibilités alléliques, il a été décidé d'introgresser, par trois rétrocroisements successifs, les allèles Mr-0 mainteneurs de stérilité dans un fond nucléaire majoritairement Sha. L'accession Sha a été choisie comme parent récurrent car, contrairement à Mr-0, c'est une accession précoce : cela a donc permis de diminuer le temps nécessaire à la réalisation de populations RM3. Le second intérêt d'utiliser Sha comme parent récurrent est d'éviter les éventuelles incompatibilités, non liées à la stérilité mâle que nous souhaitons étudier, entre des allèles nucléaires Mr-0 et le cytoplasme Sha. Afin de cartographier les locus ch1 et ch3, deux stratégies ont été envisagées : la première était de produire des populations de cartographie en utilisant, comme parent femelle, un parent mâle stérile issu de la population RM3, et comme parent mâle, un individu Sha. Mais cela présentait une difficulté due à la nécessité des deux mainteneurs pour l'expression du phénotype stérile : l'interprétation des recombinants à un locus est dépendante du génotype à l'autre locus. Par exemple, si un recombinant est identifié au locus ch3, il ne pourra être exploité que s'il porte au moins un allèle Mr-0 au locus ch1, soit une probabilité de  $\frac{1}{2}$ . Pour contourner cette difficulté, j'ai adopté la seconde stratégie : afin que tous les recombinants soient exploitables à chaque locus, les deux locus ont été cartographiés indépendamment. Pour cela, l'allèle Mr-0 a été fixé à un des deux locus de maintien dans la population où l'autre locus est cartographié. Ainsi, deux types de populations de cartographie indépendantes ont été produites : des populations de cartographie appelées PC3 où les allèles Mr-0 sont fixés au locus ch1 pour cartographier le mainteneur au locus ch3, et des populations de cartographie appelées PC1 où les allèles Mr-0 sont fixés au locus ch3 afin de cartographier le mainteneur au locus ch1.

Le plan de croisement réalisé pour produire des populations de cartographie indépendantes adaptées à la cartographie de chaque locus, est donc composé de trois étapes : l'introgession des allèles mainteneurs de stérilité Mr-0 dans un fond nucléaire Sha afin d'homogénéiser et de stabiliser le contexte génétique de la cartographie (figure B3), la fixation des allèles Mr-0 pour cartographier indépendamment les deux locus mainteneurs (figure B4) et la production



**Figure B5 : production de populations de cartographie indépendantes**

(a) Production des populations de cartographie PC3 adaptées à la cartographie du locus ch3. Les parents proviennent des croisements réalisés pour fixer le locus ch1 (figure B4a).

(b) Production des populations de cartographie PC1 à partir d'individus de la population « Fixation Mr-0 au locus ch3 » (figure B4b). Chaque boîte correspond à un génotype. Les cinq barres verticales dans le rectangle blanc représentent les 5 chromosomes. La couleur du fond de chaque boîte indique le cytotypage. En rouge : génotype Mr-0 ; en jaune, génotype Sha ; en marron : génotype hétérozygote. Le phénotype de chaque plante est également indiqué : F = fertile, bleu et S = stérile, rose.

**a**

Population RM3-A	Chromosome 1			Chromosome 3				
	CIW1	NGA128	ATHGENEA	MSAT3.58	MSAT3.20435	ATHFUS8	MSAT3.23388	MSAT3.70
Nom du marqueur	18,37	20,63	22,40	18,66	20,43	22,64	23,38	23,45
Position physique (Mb)	44	47	49	44	43	44	46	41
Nombre d'individus Sha	38	40	41	44	44	45	42	46
Nombre d'individus He	82	87	90	88	87	89	88	87
Total	0,44	0,56	0,71	0,00	0,01	0,01	0,16	0,29
Chi2 génotypes (1 : 1)	0,29	0,38	0,47	0,00	0,01	0,01	0,12	0,19
Chi2 allèles (3 : 1)								

**b**

Population RM3-B	Chromosome 1			Chromosome 3				
	CIW1	NGA128	ATHGENEA	MSAT3.58	MSAT3.20435	ATHFUS8	MSAT3.23388	MSAT3.70
Nom du marqueur	18,37	20,63	22,40	18,66	20,43	22,64	23,38	23,45
Position physique (Mb)	143	143	99	113	121	99	117	127
Nombre d'individus Sha	101	105	73	121	127	109	114	129
Nombre d'individus He	344	248	172	254	248	208	231	258
Total	7,25	3,62	3,69	0,27	0,16	0,48	0,04	0,00
Chi2 génotypes (1 : 1)	4,40	3,86	2,62	0,16	0,10	0,32	0,03	0,01
Chi2 allèles (3 : 1)								

**Tableau B1 : ségrégation des marqueurs dans les populations RM3-A et RM3-B.**

Les tableaux indiquent, pour les populations RM3-A (a) et RM3-B (b), le nom et la position physique des marqueurs génotypés et les effectifs d'individus hétérozygotes et Sha dans ces populations. Les valeurs de Chi2 ont été calculées, pour chaque marqueur, pour la répartition des génotypes dans la population (attendue à 1 : 1) et la transmission des allèles parentaux (attendue à 3 : 1). Les valeurs de Chi2 significatives ( $\alpha=0,05$ ) sont indiquées en rouge.

de populations de cartographie indépendantes (figure B5).

### I.2.a Introgression des allèles Mr-0 mainteneurs de stérilité dans un fond Sha

Deux rétrocroisements successifs par Sha ont donc été réalisés sur des individus stériles de la population RM1 et hétérozygotes aux marqueurs NGA128 et MSAT3.70, liés respectivement aux locus ch1 et ch3, afin d'obtenir des populations RM2 puis RM3. Pour augmenter nos chances d'avoir une population à transmission normale, deux populations RM3 ont été créées : une population appelée RM3-A (92 individus) et une population RM3-B (276 individus), obtenues respectivement par les croisements « RM2 n°91 x Sha » et « RM2 n°104 x Sha ». Les individus stériles des populations RM3 ont les deux locus mainteneurs de stérilité à l'état hétérozygote qui se trouvent dans un fond nucléaire et cytoplasmique majoritairement Sha (figure B3). Un génotypage est réalisé sur ces deux populations RM3 à l'aide de trois marqueurs situés au sud du chromosome 1 (CIW1, NGA128 et ATHGENEA) et de cinq marqueurs situés au sud du chromosome 3 (MSAT 3.58, MSAT3.20435, ATHFUS6, MSAT3.23388 et MSAT3.70). Ces 8 marqueurs sont hétérozygotes chez les parents mâle stériles RM2 n°91 et RM2 n°104. Le tableau B1 résume le génotypage des populations RM3-A et RM3-B : la ségrégation des allèles aux marqueurs cités précédemment est normale dans la population RM3-A (tableau B1a). Cette population présente une ségrégation phénotypique normale, proche des valeurs attendues : il y a trois fois plus d'individus fertiles que d'individus stériles ( $\chi^2 = 0,06$ ). Dans la population RM3-B, il y a un déficit d'individus hétérozygotes observé aux trois marqueurs du chromosome 1 (tableau B1b), ainsi qu'une proportion d'individus stériles plus faible que celle attendue ( $\chi^2 = 5,72$ ). La faible proportion d'individus stériles est cohérente avec la faible transmission des allèles Mr-0 aux marqueurs du locus ch1. La présence résiduelle d'incompatibilités alléliques dans un fond génétique pas totalement homogénéisé, pourrait expliquer ces observations. Les biais de ségrégation peuvent compliquer la stratégie de cartographie des locus mainteneurs. Face à ces observations, seuls des individus de la population RM3-A ont été sélectionnés afin de réaliser la suite du plan de croisement.

### I.2.b Fixation des allèles Mr-0 mainteneurs de stérilité

#### I.2.b.a Fixation des allèles Mr-0 au locus ch1

Pour fixer l'allèle Mr-0 à l'état homozygote sur le locus ch1, des plantes RM3-A stériles,

**a**

Nom du marqueur	CIW1	NGA128		MSAT3.58	MSAT3.20435
Position physique (Mb)	18,37	20,63		18,66	20,43
Nombre d'individus Sha	22	10		20	20
Nombre d'individus He	30	30		37	40
Nombre d'individus Mr-0	20	23		0	0
Total	79	70		75	78
Chi2 allèles (1 : 1)	0,089	1,256	Chi2 allèles (3/4 : 1/4)	0,008	0,034
Chi2 génotypes (1/2 : 1/4 : 1/2)	0,21	1,256	Chi2 génotypes (1 : 1)	0,013	0,051

**b**

Nom du marqueur	CIW1	NGA128		MSAT3.20435
Position physique (Mb)	18,37	20,63		20,43
Nombre d'individus Sha	20	15		72
Nombre d'individus He	43	40		18
Nombre d'individus Mr-0	19	28		0
Total	88	91		91
Chi2 allèles (1 : 1)	1,114	3,714	Chi2 allèles (3/4 : 1/4)	20,374
Chi2 génotypes (1/2 : 1/4 : 1/2)	1,150	3,089	Chi2 génotypes (1 : 1)	20,888

**Tableau B2 : ségrégation des marqueurs dans les populations « fixation Mr-0 au ch1 »**

a- Population issue du croisement « RM3-A CC3 x RM3-A CA4 ».

b- Population issue du croisement « RM3-A CG4 x RM3-A CH10 ». Les tableaux indiquent le nom et la position physique des marqueurs génotypés et les effectifs d'individus de chaque génotype à chaque marqueur dans ces populations. Les valeurs de Chi2 ont été calculées, pour chaque marqueur, pour la répartition des génotypes dans la population et la transmission des allèles parentaux. Les proportions attendues dans chaque cas sont indiquées entre parenthèses. Les valeurs de Chi2 significatives ( $\alpha=0,05$ ) sont indiquées en rouge.**a**

Nom du marqueur	CIW1	NGA128		MSAT3.58	MSAT3.20435
Position physique (Mb)	18,37	20,63		18,66	20,43
Nombre d'individus Sha	48	47		8	5
Nombre d'individus He	42	42		46	41
Nombre d'individus Mr-0	0	0		32	38
Total	90	89		86	84
Chi2 allèles (3/4 : 1/4)	0,267	0,187	Chi2 allèles (1 : 1)	13,395	28,809
Chi2 génotypes (1 : 1)	0,4	0,281	Chi2 génotypes (1/2 : 1/4 : 1/2)	13,814	29,978

**b**

Nom du marqueur	CIW1	NGA128		MSAT3.20435
Position physique (Mb)	18,37	20,63		20,43
Nombre d'individus Sha	55	43		7
Nombre d'individus He	37	42		53
Nombre d'individus Mr-0	0	0		28
Total	92	85		88
Chi2 allèles (3/4 : 1/4)	2,348	0,006	Chi2 allèles (1 : 1)	16,023
Chi2 génotypes (1 : 1)	3,522	0,012	Chi2 génotypes (1/2 : 1/4 : 1/2)	16,706

**Tableau B3 : ségrégation des marqueurs dans les populations « fixation Mr-0 au ch3 »**

a- Population issue du croisement « RM3-A CB2 X RM3-A CF8 ».

b- Population issue du croisement « RM3-A CH8 X RM3-A CG5 ». Les tableaux indiquent le nom et la position physique des marqueurs génotypés et les effectifs d'individus de chaque génotype à chaque marqueur dans ces populations. Les valeurs de Chi2 ont été calculées, pour chaque marqueur, pour la répartition des génotypes dans la population et la transmission des allèles parentaux. Les proportions attendues dans chaque cas sont indiquées entre parenthèses. Les valeurs de Chi2 significatives ( $\alpha=0,05$ ) sont indiquées en rouge.

hétérozygotes aux deux locus, sont utilisées comme parents femelles et croisées avec des plantes RM3-A fertiles, Sha au locus ch3 et hétérozygotes au locus ch1 (figure B4a). Deux couples ont été formés pour générer deux populations appelées populations « fixation Mr-0 au locus ch1 » : les couples « RM3-A CC3 X RM3-A CA4 » et « RM3-A CG4 X RM3-A CH10 ». Le génotypage des individus obtenus à partir du couple « RM3-A CC3 X RM3-A CA4 » a révélé des proportions d'individus Sha, Mr-0 et hétérozygotes aux marqueurs CIW1 et NGA128 proches des proportions attendues ( $\frac{1}{4}$  d'individus Sha,  $\frac{1}{4}$  d'individus Mr-0 et  $\frac{1}{2}$  d'individus hétérozygotes, tableau B2a). De même, la proportion d'individus Sha et hétérozygotes aux marqueurs MSAT3.58 et MSAT3.20435 est proche de la proportion attendue ( $\frac{1}{2}$  d'individus Sha et  $\frac{1}{2}$  d'individus hétérozygotes, tableau B2a). Il existe un biais de ségrégation au sud du chromosome 3 dans la population issue du croisement « RM3-A CG4 X RM3-A CH10 » (tableau B2b) : les proportions d'individus Sha et hétérozygotes au marqueur MSAT3.20435 sont éloignées des proportions attendues ( $\frac{1}{2}$  d'individus Sha et  $\frac{1}{2}$  d'individus hétérozygotes avec un  $\chi^2 = 30,868$ ) ; l'allèle Mr-0 est contre sélectionné à ce marqueur ( $\chi^2 = 20,579$ ).

Certains individus stériles de la population « fixation Mr-0 au locus ch1 » qui ne présente pas de biais de ségrégation sont homozygotes Mr-0 au locus ch1 (à CIW1 et NGA128) et hétérozygotes au locus ch3 (à MSAT3.58 et MSAT3.20435) : ils seront utilisés comme parents femelles pour produire des populations de cartographie adaptées à la cartographie du locus ch3. Des individus fertiles sont homozygotes Mr-0 au locus ch1 (à CIW1 et NGA128) et Sha au locus ch3 (à MSAT3.58 et MSAT3.20435) : ils seront utilisés comme parents mâles pour produire ces populations de cartographie.

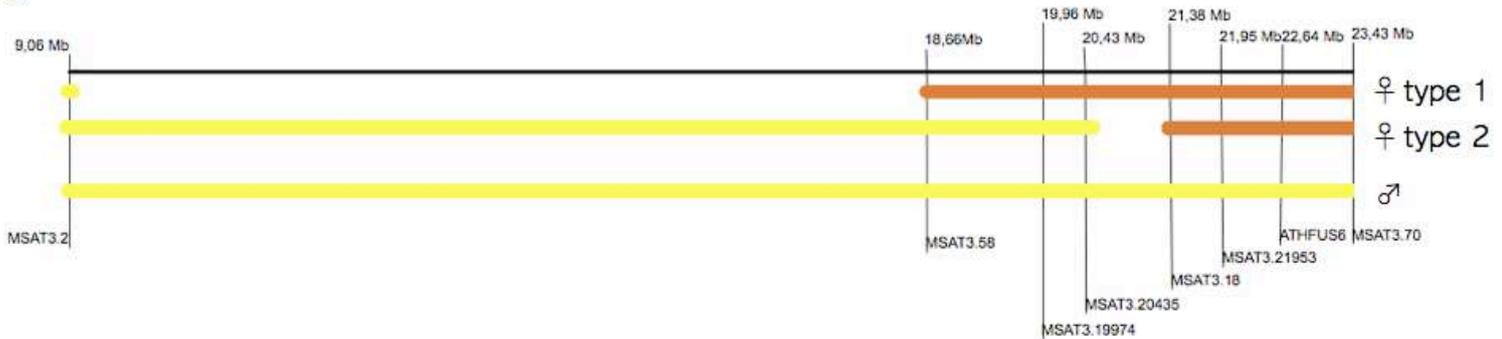
### I.2.b.b Fixation des allèles Mr-0 au locus ch3

Pour fixer l'allèle Mr-0 à l'état homozygote sur le locus ch3, des plantes RM3-A stériles, hétérozygotes aux deux locus, sont utilisées comme parents femelles et croisées avec des plantes RM3-A fertiles Sha au locus ch1 et hétérozygotes au locus ch3 (figure B4b). Au total, deux couples ont été formés pour générer deux populations appelées populations « fixation Mr-0 au locus ch3 » : les couples « RM3-A CB2 x RM3-A CF8 » et « RM3-A CH8 x RM3-A CG5 ». Le génotypage des individus des populations « fixation Mr-0 au locus ch3 » a révélé des proportions d'individus Sha, Mr-0 et hétérozygotes aux marqueurs MSAT3.58 et MSAT3.20435 éloignées des proportions attendues ( $\frac{1}{4}$  d'individus Sha,  $\frac{1}{4}$  d'individus Mr-0 et  $\frac{1}{2}$  d'individus hétérozygotes, tableau B3) : on observe un fort déficit d'allèles Sha à ces

a

Nom de la population	Croisement	Taille de la Population	Nombre de recombinants	
PC3	♀ type 1	40BCm3[CC3xCA4]C2 x 40BCm3[CC3xCA4]F3	644	7 entre MSAT3.19974 et MSAT3.20642 (668 Kb)
PC3.1	♀ type 1	40BCm3[CC3xCA4]C3 x 40BCm3[CC3xCA4]D8	828	11 entre MSAT3.19974 et MSAT3.20642 (668 Kb)
PC3.2	♀ type 1	40BCm3[CC3xCA4]H8 x 40BCm3[CC3xCA4]A8	1012	28 entre MSAT3.19974 et MSAT3.20642 (668 Kb)
PC3.3	♀ type 1	40BCm3[CC3xCA4]F7 x 40BCm3[CC3xCA4]D8	552	28 entre MSAT3.19974 et MSAT3.70 (3475 Kb)
PC3.4	♀ type 2	PC3.3 n°99 x 40BCm3[CC3xCA4]D8 (AF)	1932	32 entre MSAT3.21869 et MSAT3.23388 (1519 Kb)
PC3.5	♀ type 2	PC3.3 n°92 x 40BCm3[CC3xCA4]D8 (AF)	644	16 entre MSAT3.21869 et MSAT3.23388 (1519 Kb)
PC3.6	♀ type 1	40BCm3[CC3xCA4]B5 x 40BCm3[CC3xCA4]F3	575	7 entre ATHFUS6 et MSAT3.23388 (765 Kb)
PC3.7	♀ type 1	40BCm3[CC3xCA4]E6 x 40BCm3[CC3xCA4]G6	161	1 entre ATHFUS6 et MSAT3.23388 (765 Kb)
PC3.8	♀ type 1	40BCm3[CC3xCA4]E9 x 40BCm3[CC3xCA4]F3	529	4 entre ATHFUS6 et MSAT3.23388 (765 Kb)
PC3.9	♀ type 1	40BCm3[CC3xCA4]B8 x 40BCm3[CC3xCA4]A8	184	0 entre ATHFUS6 et MSAT3.23388 (765 Kb)
PC3.10	♀ type 1	40BCm3[CC3xCA4]B4 x 40BCm3[CC3xCA4]G6	207	0 entre ATHFUS6 et MSAT3.23388 (765 Kb)
PC3.11	♀ type 1	40BCm3[CC3xCA4]B4 x 40BCm3[CC3xCA4]A8	368	4 entre ATHFUS6 et MSAT3.23388 (765 Kb)
PC3.12	♀ type 2	PC3.3 n°92 (CD11) x 40BCm3[CC3xCA4]D8 (AF)	368	1 entre ATHFUS6 et MSAT3.23388 (765 Kb)
PC3.13	♀ type 2	PC3.3 n°80 (CA12) x 40BCm3[CC3xCA4]D8 (AF)	184	0 entre ATHFUS6 et MSAT3.23388 (765 Kb)

b



**Figure B6 : caractéristiques des populations de cartographie du locus ch3 et génotypes des parents de ces populations**

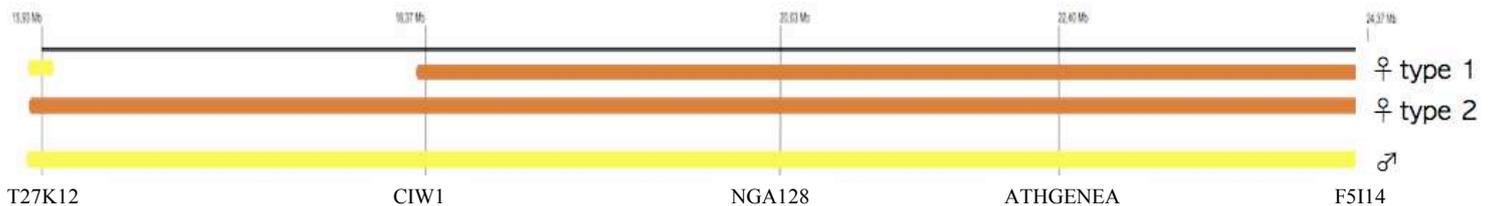
a. Caractéristiques des populations de cartographie PC3.

b Représentation des génotypes des parents de populations de cartographie du locus ch3. Les traits verticaux représentent la position des marqueurs utilisés pour génotyper ces plantes. Leurs noms et positions physiques sur le chromosome 3 sont indiqués au-dessus et au-dessous. En jaune, génotype Sha ; en marron : génotype hétérozygote. Deux types de parent femelle (type 1 ou type 2) sont distingués selon l'étendue de la zone d'hétérozygotie au chromosome 3.

a

Nom de la population	Croisement	Taille de la Population	Nombre de recombinants	
PC1	♀ type 2	40BCm3[CB2xCF8]A4 x 40BCm3[CB2xCF8]A6	644	27 entre CIW1 et NGA128 (2,26 Mb)
PC1.1	♀ type 1	40BCm3[CB2xCF8]F9 x 40BCm3[CB2xCF8]G8	138	6 entre CIW1 et NGA128 (2,26 Mb)
PC1.2	♀ type 2	40BCm3[CB2xCF8]D6 x 40BCm3[CB2xCF8]H1	552	52 entre CIW1 et F5114 (6 Mb)

b



**Figure B7 : caractéristiques des populations de cartographie du locus ch1 et génotypes des parents de ces populations**

a Caractéristiques des populations de cartographie.

b Représentation des génotypes des parents de populations de cartographie du locus ch1. Les traits verticaux représentent la position des marqueurs utilisés pour génotyper ces plantes. Leurs noms et positions physiques sur le chromosome 1 sont indiqués. En jaune, génotype Sha ; en marron : génotype hétérozygote. Deux types de parent femelle (type 1 ou type 2) sont distingués selon l'étendue de la zone d'hétérozygotie au chromosome 1.

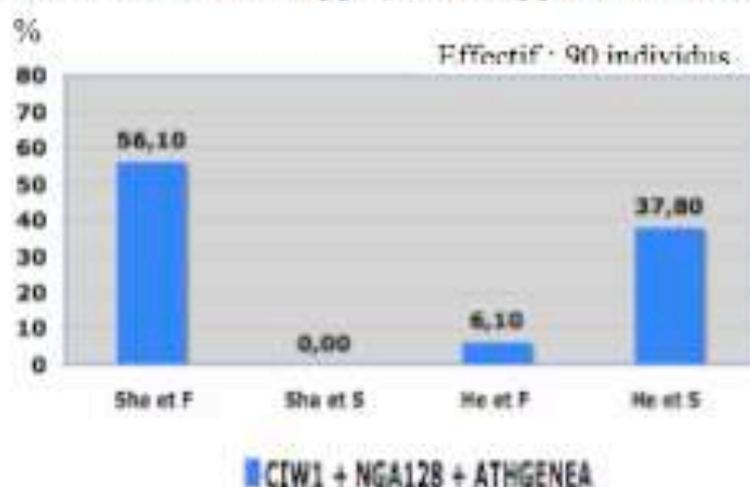
marqueurs dans les populations « fixation Mr-0 au locus ch3 ». Une contre sélection des allèles Sha à ces marqueurs avait également été observée dans la population F2 63RV (figure B2b). Certains individus stériles des populations « fixation Mr-0 au locus ch3 » sont homozygotes Mr-0 au locus ch3 (à MSAT3.58 et MSAT3.20435) et hétérozygotes au locus ch1 (à CIW1 et NGA128). Ces individus seront utilisés comme parents femelles pour produire des populations de cartographie adaptées à la cartographie du locus ch1. Des individus fertiles des populations « fixation Mr-0 au locus ch3 » sont homozygotes Mr-0 au locus ch3 (à MSAT3.58 et MSAT3.20435) et Sha au locus ch1 (à CIW1 et NGA128). Ils seront utilisés comme parents mâles pour produire ces populations de cartographie.

### I.2.c Production de populations de cartographies

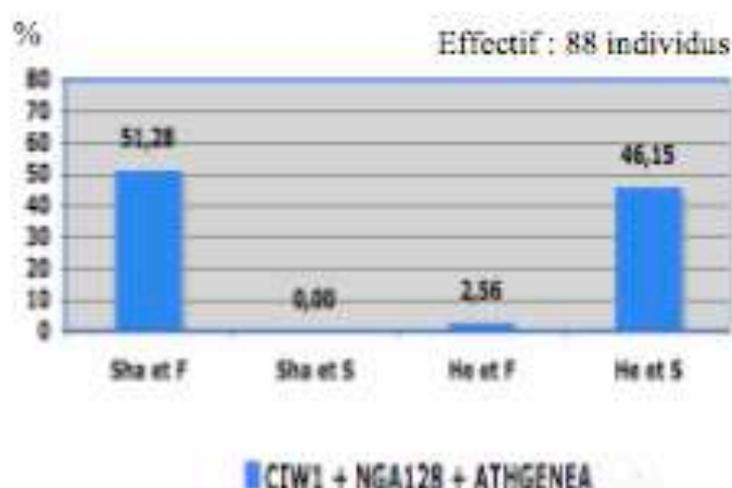
La troisième étape est la production des populations de cartographie (figure B5). Pour produire des populations de cartographie PC3, adaptées à la cartographie du locus ch3, 10 couples de plantes, issues des populations « fixation Mr-0 au locus ch1 » RM3-A CC3 X RM3-A CA4 et qui sont homozygotes Mr-0 au locus ch1 (aux marqueurs CIW1, NGA128 et ATHGENEA), ont été choisies après génotypage et croisées pour produire des populations de cartographie PC3. De plus, au cours du travail de cartographie, 3 individus mâle stériles issus de la PC3.3 (PC3.3 n°80, PC3.3 n°92 et PC3.3 n°99) ont été utilisés comme parents femelles pour produire 4 autres populations de cartographies (PC3.4, PC3.5, PC3.12 et PC3.13) ; la PC3.5 et la PC3.12 sont issues du même lot de graines mais ont été analysées à des moments différents. Ces individus mâle stériles ont été croisés avec un individu fertile, descendant d'autofécondation du parent mâle D8 des populations PC3.1 et PC3.3, homozygote Mr-0 au locus ch1 (aux marqueurs CIW1, NGA128 et ATHGENEA). Pour chaque couple, le parent femelle de la population de cartographie est stérile et hétérozygote au locus ch3 et le parent mâle de la population est homozygote Sha au locus ch3. Les caractéristiques des PC3 et les génotypes des parents de PC3 sont présentés dans la figure B6.

Pour produire des populations appelées PC1, adaptées à la cartographie du locus ch1, 3 couples de plantes issues de croisements entre RM3-A et qui sont homozygotes Mr-0 au locus ch3 (aux marqueurs MSAT3.58 et MSAT3.23388), ont été choisis après génotypage puis croisés pour produire des populations de cartographie PC1. Pour chaque couple, le parent femelle de la population de cartographie est stérile et hétérozygote au locus ch1 et le parent mâle de la population est homozygote Sha au locus ch1. Les caractéristiques des PC1 et les génotypes des parents de PC1 sont présentés dans la figure B7.

### Corrélation Phénotype/Génotype dans la PC1



### Corrélation Phénotype/Génotype dans la PC1.1



### Corrélation Phénotype/Génotype dans la PC1.2

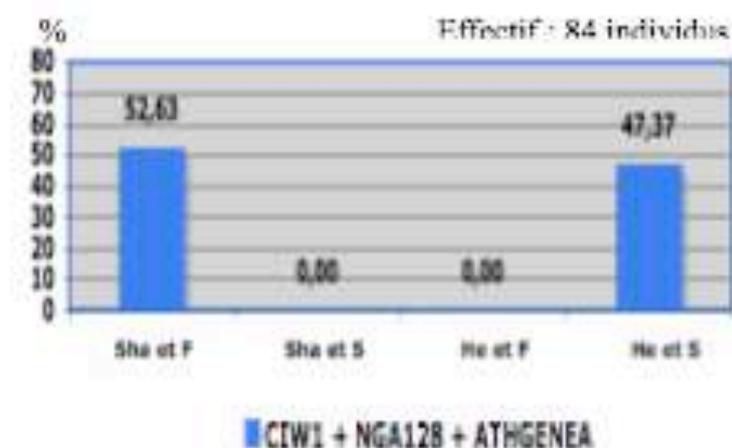


Figure B8 : liaison du génotype aux marqueurs encadrant ch1 avec les phénotypes de stérilité (S) et de fertilité (F)

Pour chacune des populations de cartographie, le diagramme représente la distribution (en %) des plantes de la plaque phénotypée selon leur phénotype (stérile, S ou fertile, F) et leur génotype aux trois marqueurs génotypés. Les plantes recombinantes entre deux marqueurs n'ont pas été prises en compte pour cette analyse. Les effectifs concernés sont indiqués au-dessus de chaque diagramme.

	PC1	PC1.1	PC1.2
Nombre d'individus stériles	36	39	37
Nombre d'individus fertiles	54	45	47
Chi2 (1:1)	7,201	0,428	1,100

**Tableau B4 : ségrégation des phénotypes dans les populations de cartographie du locus ch1.**

Le Chi2 a été calculé pour une proportion égale de plantes stériles et fertiles, attendue dans ces populations. La valeur du Chi2 significative ( $\alpha=0,05$ ) est indiquée en rouge.

	PC1	PC1.1	PC1.2
Nom du marqueur	CIW1	CIW1	CIW1
Nombre d'individus Sha	304	133	268
Nombre d'individus He	303	123	250
Chi2 géotypes (1:1)	0,003	0,724	1,223

**Tableau B5 : ségrégation du marqueur CIW1 dans les populations de cartographie du locus ch1.**

Le Chi2 a été calculé pour une proportion égale de plantes hétérozygotes et Sha, attendue dans ces populations.

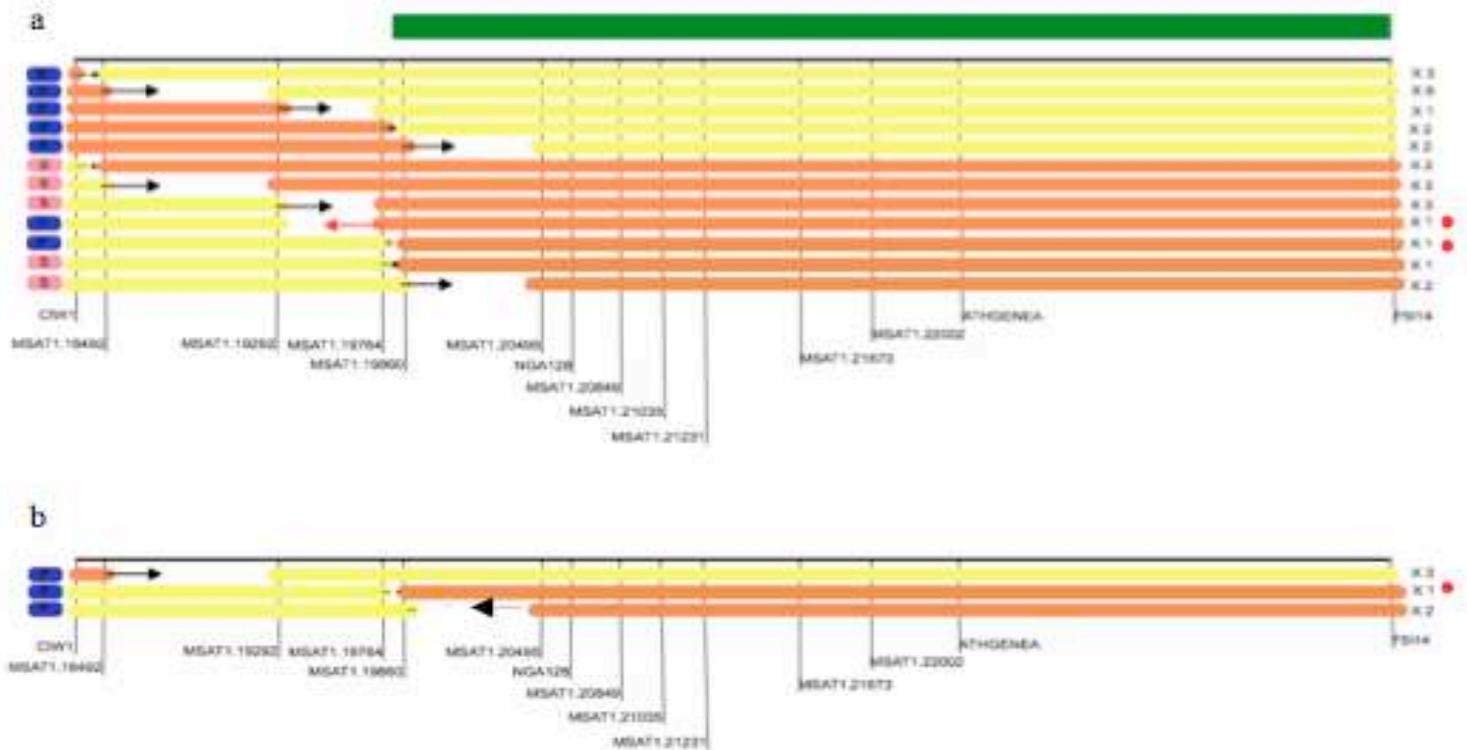
### I.3 Cartographie génétique du locus mainteneur de stérilité au chromosome 1

#### I.3.a Analyse des populations de cartographie

Trois populations de cartographie (notées PC1, PC1.1 et PC1.2) ont été semées. Pour chaque population de cartographie, une plaque de 92 individus a été phénotypée afin de contrôler la corrélation entre le phénotype et le génotype aux marqueurs CIW1, NGA128, ATHGENEA situés au sud du chromosome 1 (figure B8). Le phénotypage des trois plaques, réalisé deux fois au cours de la floraison, a révélé plusieurs catégories de phénotypes : des plantes clairement fertiles (siliques allongées) ou stériles (siliques courtes), mais également parfois des plantes stériles qui ont quelques siliques allongées (plantes S/F) et des plantes fertiles avec quelques siliques courtes (plantes F/S). Les plantes S/F ont été classées dans les plantes S et les plantes F/S ont été classées dans les plantes F (figure B8). À noter que des siliques allongées ont été observées en fin de floraison chez des individus stériles hétérozygotes aux trois marqueurs cités précédemment et que pour de rares individus stériles, le phénotype des hampes florales primaires est différent de celui des hampes florales secondaires, où des siliques allongées sont apparues.

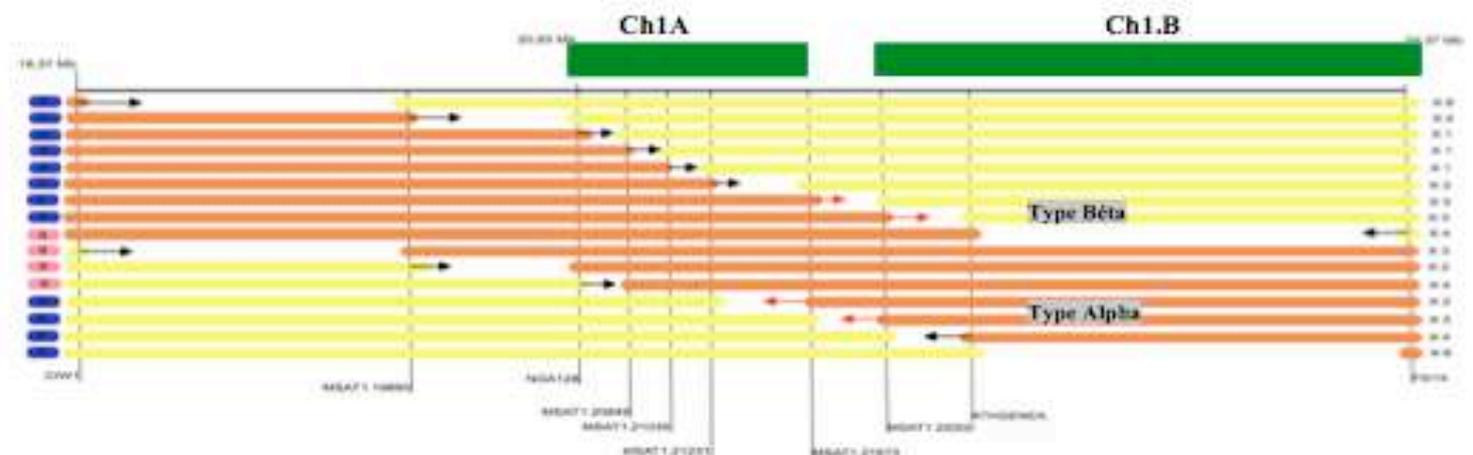
On peut constater dans la figure B8 que le génotype des plantes aux trois marqueurs est très corrélé à leur phénotype dans ces populations : si une plante est hétérozygote à ces trois marqueurs, elle est stérile. Au contraire, si une plante de ces populations est Sha à ces marqueurs, elle est fertile. Cependant, il existe dans la PC1 et la PC1.1, des plantes qualifiées de « hors-types » : ce sont des plantes dont le phénotype observé est différent du phénotype prédit par le génotypage aux marqueurs considérés. Par exemple, dans la PC1, 6,10 % des plantes hétérozygotes au locus ch1 sont fertiles. Dans la PC1.1, 2,56 % des plantes dont l'allèle Mr-0 est présent au locus ch1 sont fertiles. Ces plantes sont appelées plantes « hors-types fertiles ». Aucune plante « hors-type stérile » (plantes qui auraient été stériles en absence de l'allèle Mr-0 au locus ch1) n'a été identifiée dans ces populations de cartographie.

Dans les populations de cartographie du locus mainteneur de stérilité au chromosome 1, la proportion des phénotypes est, en théorie, de  $\frac{1}{2}$  d'individus stériles et  $\frac{1}{2}$  d'individus fertiles. Comme nous pouvons le voir dans le tableau B4, la proportion observée d'individus stériles et fertiles est proche des proportions attendues dans la PC1.1 et la PC1.2 mais est significativement différente dans la population PC1 où il apparaît un déficit en individus stériles. Ce déficit en individus stériles est cohérent avec la présence de hors-types fertiles (6,10 %) dans la PC1. La ségrégation des génotypes (au marqueur CIW1), dans les populations de cartographie du locus mainteneur de stérilité, est très proche des proportions



**Figure B9 : génotypes des recombinants identifiés dans la PCI (a) et la PCI.1 (b).**

La colonne de gauche indique le phénotype des recombinants (F = fertile et S = stérile) et la colonne de droite indique le nombre de recombinants pour un génotype donné. Les traits verticaux représentent la position des marqueurs utilisés pour génotyper ces plantes. Leurs noms et positions physiques sur le chromosome 1 sont indiqués. En jaune, génotype Sha ; en marron : génotype hétérozygote. Le cadre vert représente l'intervalle génétique contenant le locus ch1, défini par les recombinants stériles. Il existe des recombinants fertiles en contradiction avec cette localisation génétique. Ils sont indiqués par des points rouges.



**Figure B10 : génotypes des recombinants identifiés dans la PCI.2.**

La colonne de gauche indique le phénotype des recombinants (F = fertile et S = stérile) et la colonne de droite indique le nombre de recombinants pour un génotype donné. Les traits verticaux représentent la position des marqueurs utilisés pour génotyper ces plantes. Leurs noms et positions physiques sur le chromosome 1 sont indiqués. En jaune, génotype Sha ; en marron : génotype hétérozygote. Les cadres verts indiquent les locus ch1A et ch1B déduits de l'analyse de cette population. Les génotypes recombinants utilisés pour tester l'hypothèse des deux locus sont indiqués. Type bêta : hétérozygotes au locus ch1A et Sha au locus ch1B ; type alpha : Sha au locus ch1A et hétérozygote au locus ch1B.

attendues ( $\frac{1}{2}$  d'hétérozygotes et  $\frac{1}{2}$  de Sha). Ces résultats sont présentés dans le tableau B5.

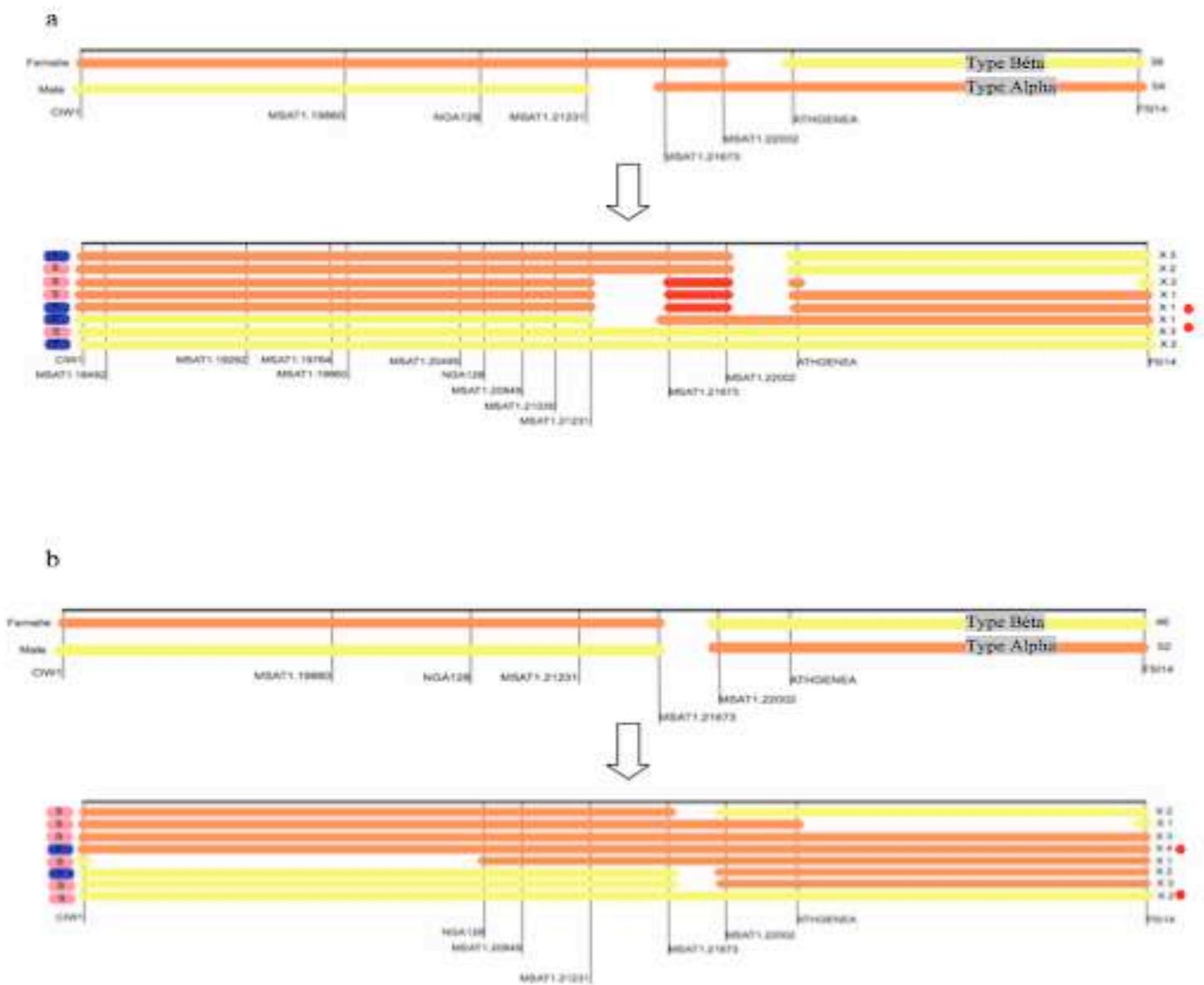
### I.3.b Analyse des recombinants au locus ch1

Le génotypage de la PC1 a été réalisé avec les marqueurs CIW1 et NGA128. Au total, 27 recombinants ont été identifiés, entre ces marqueurs, dans la PC1 (figure B9a). Parmi ces recombinants, deux recombinants stériles, qui sont Sha à MSAT1.19860 et hétérozygotes à MSAT1.20495, et deux recombinants fertiles, qui sont hétérozygotes à MSAT1.19860 et Sha à MSAT1.20495, indiquent que le locus d'intérêt est situé à droite de MSAT1.19860. Néanmoins, dans la PC1, deux plantes fertiles contredisent cette hypothèse et semblent indiquer que le locus ch1 est situé à gauche de MSAT1.19860. Cependant la population PC1 a présenté 6,10 % d'individus hors-types fertiles (figure B8) ; j'ai donc choisi de ne considérer que les recombinants stériles de cette population pour la définition de l'intervalle génétique contenant le locus ch1.

Le génotypage de la PC1.1 a été réalisé avec les marqueurs CIW1 et NGA128. Au total, ce sont 6 recombinants qui ont été identifiés entre ces marqueurs dans la PC1.1 (Figure B9b). Aucun recombinant stérile n'a été identifié dans la PC1.1. Parmi les recombinants fertiles, un recombinant Sha à MSAT1.19764 et hétérozygote à MSAT1.19860 est en contradiction avec l'hypothèse d'une localisation à droite de MSAT1.19860 du locus ch1 et semblent indiquer que le locus ch1 est situé à gauche de MSAT1.19860, alors que les autres recombinants sont cohérents avec l'intervalle défini précédemment. Ce recombinant, comme les recombinants hors-type fertiles de la PC1, pourrait être un hors-type fertile pouvant appartenir aux 2,56 % constatés dans la PC1.1 (figure B8) ; il n'a donc pas été pris en compte dans la définition de l'intervalle génétique contenant le locus ch1.

Après avoir analysé les recombinants identifiés dans la PC1 et la PC1.1, il s'est avéré qu'aucun recombinant stérile ne permet de borner le locus ch1 à droite. Ainsi, il a été décidé, pour la PC1.2, d'agrandir l'intervalle génétique dans lequel on recherche des recombinants. De plus, la présence de hors-types fertiles dans ces populations m'a amené, lors de la cartographie, à éliminer les régions génomiques seulement si au moins un recombinant stérile exclut l'une des bornes de l'intervalle considéré.

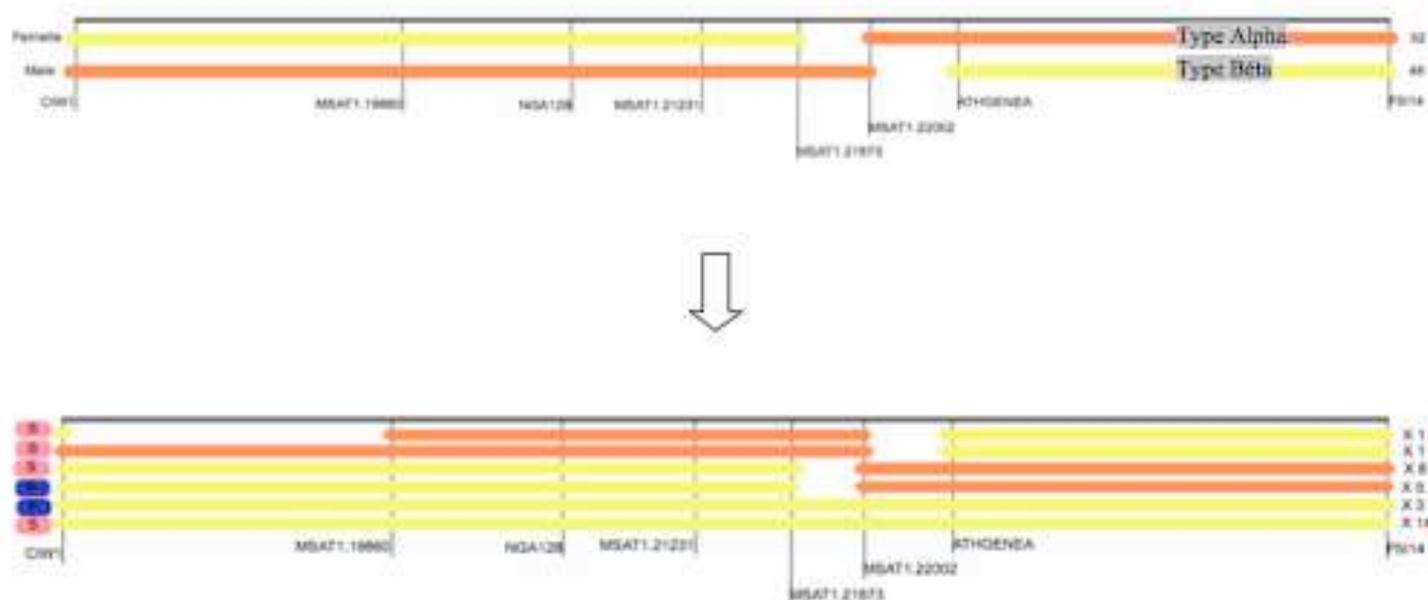
La recherche des recombinants dans la PC1.2 a donc été réalisée avec les marqueurs CIW1 et ATHGENEA (figure B10). La plaque phénotypée dans la PC1.2 ne contient aucune plante hors type avec les marqueurs CIW1, NGA128 et ATHGENEA. 45 recombinants ont été



**Figure B11 : génotypes des descendants issus des croisements entre recombinants PC1.2.**  
 (a) Croisement RecPC1.2 n°39 X RecPC1.2 n°54. (b) Croisement RecPC1.2 n°46 X RecPC1.2 n°52. Pour les génotypes des parents, la colonne de gauche indique si l'individu a été utilisé comme parent mâle ou comme parent femelle ; la colonne de droite indique le numéro de la plante et le type alpha ou bêta de la plante. Dans la descendance des croisements, la colonne de gauche indique le phénotype des descendants (F = fertile et S = stérile) et la colonne de droite indique le nombre de descendants pour un génotype donné. Les traits verticaux représentent la position des marqueurs utilisés pour génotyper ces plantes. Leurs noms et positions physiques sur le chromosome 1 sont indiqués. En jaune, génotype Sha ; en marron : génotype hétérozygote. Les plantes dont le phénotype est contradictoire avec les résultats de la PC1.2 sont indiquées par des points rouges.

identifiés entre les marqueurs CIW1 et ATHGENEA, puis génotypés à l'aide de marqueurs internes. Après avoir phénotypé ces recombinants PC1.2, il apparaît clairement que le locus ch1 est situé à droite de NGA128 (4 recombinants stériles sont Sha à NGA128 et hétérozygotes à MSAT1.20849 et 1 recombinant fertile est hétérozygote à NGA128 et Sha à MSAT1.20849). En utilisant le marqueur ATHGENEA, aucun recombinant stérile ne permet de limiter la borne de droite ; l'intervalle génétique dans lequel on cherche des recombinants a donc été une nouvelle fois agrandi, à l'aide du marqueur F5I14. Il apparaît que le locus ch1 est situé à gauche de F5I14 (4 recombinants stériles sont Sha à F5I14 et hétérozygotes à ATHGENEA et 6 recombinants fertiles sont hétérozygotes à F5I14 et Sha à ATHGENEA). Les 18 recombinants identifiés entre MSAT1.20849 et ATHGENEA sont fertiles (figure B10). Il n'y a aucun recombinant stérile entre ces deux marqueurs. De plus, les recombinants fertiles dans cet intervalle se contredisent : trois recombinants semblent indiquer que le locus ch1 est situé à gauche de MSAT1.22002 alors que trois autres recombinants semblent indiquer que le locus ch1 est situé à droite de MSAT1.22002. De même, deux recombinants fertiles semblent indiquer que le locus ch1 est situé à gauche de MSAT1.21673 alors que trois autres recombinants semblent indiquer que le locus ch1 est situé à droite de MSAT1.21673.

Pour expliquer l'absence de recombinants stériles et les contradictions entre les recombinants fertiles dans la région génomique [MSAT1.21231-ATHGENEA], l'hypothèse suivante a été émise : il est possible que le locus ch1 soit constitué de deux locus, appelés ch1A et ch1B, qui doivent chacun porter l'allèle Mr-0 afin de maintenir la stérilité mâle. Le locus ch1A serait situé dans la région génomique ]NGA128 – MSAT1.21673[, ce qui expliquerait la fertilité des cinq recombinants qui sont Sha à MSAT1.21231 et hétérozygotes à MSAT1.22002. Le locus ch1B serait situé dans la région génomique ]MSAT1.22002–F5I14[, ce qui expliquerait la fertilité des six recombinants qui sont Sha à ATHGENEA et hétérozygotes à MSAT1.1.21673. Afin de tester cette hypothèse, j'ai réalisé des croisements entre ces recombinants fertiles. En effet, si notre hypothèse est correcte, la descendance d'un croisement entre un recombinant fertile hétérozygote au locus ch1A et Sha au locus ch1B (recombinant de type bêta), et un recombinant fertile Sha au locus ch1A et hétérozygote au locus ch1B (recombinant de type alpha), devrait produire des individus stériles car hétérozygotes aux deux locus. Les génotypes des recombinants fertiles utilisés comme parents pour tester cette hypothèse, ainsi que les génotypes de leurs descendants sont présentés dans la figure B11. Les deux populations, obtenues par les croisements RecPC1.2 n°39 (type bêta) X RecPC1.2 n°54 (type alpha) et RecPC1.2 n°46 (type bêta) X RecPC1.2 n°52 (type alpha) sont de petite taille (une quinzaine d'individus par population). Il apparaît, dans ces deux



**Figure B12 : génotypage d'une population issue du croisement RecPC1.2 n°52 X RecPC1.2 n°48.**

Pour les génotypes des parents, la colonne de gauche indique si l'individu a été utilisé comme parent mâle ou comme parent femelle ; la colonne de droite indique le numéro de la plante et le type alpha ou bêta de la plante. Dans la descendance des croisements, la colonne de gauche indique le phénotype des descendants (F = fertile et S = stérile) et la colonne de droite indique le nombre de descendants pour un génotype donné. Les traits verticaux représentent la position des marqueurs utilisés pour génotyper ces plantes. Leurs noms et positions physiques sur le chromosome 1 sont indiqués. En jaune, génotype Sha ; en marron : génotype hétérozygote. Il existe des recombinants en contradiction avec la localisation génétique du locus ch3 définie dans les figures B9 et B10. Ils sont indiqués par des points rouges.

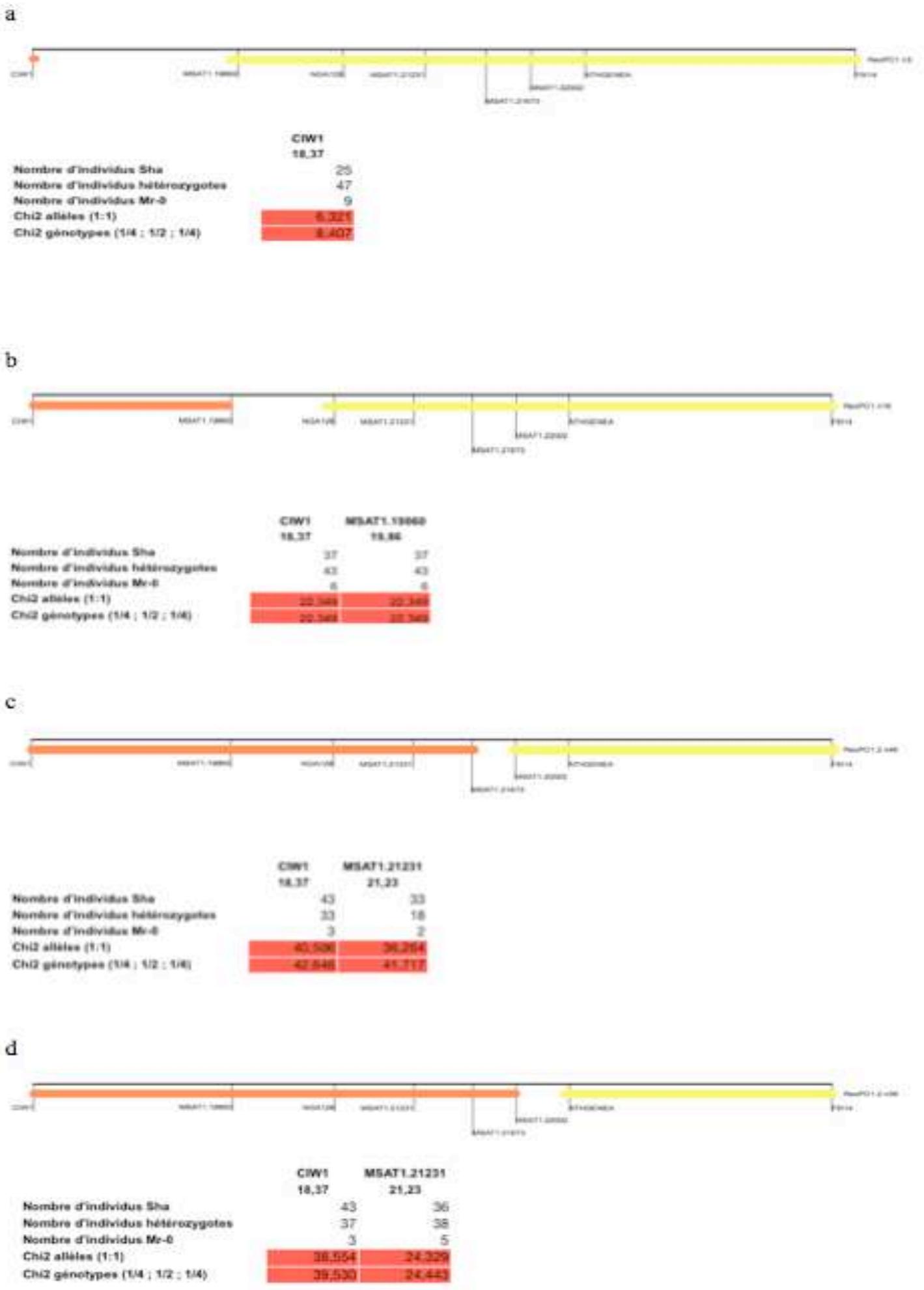
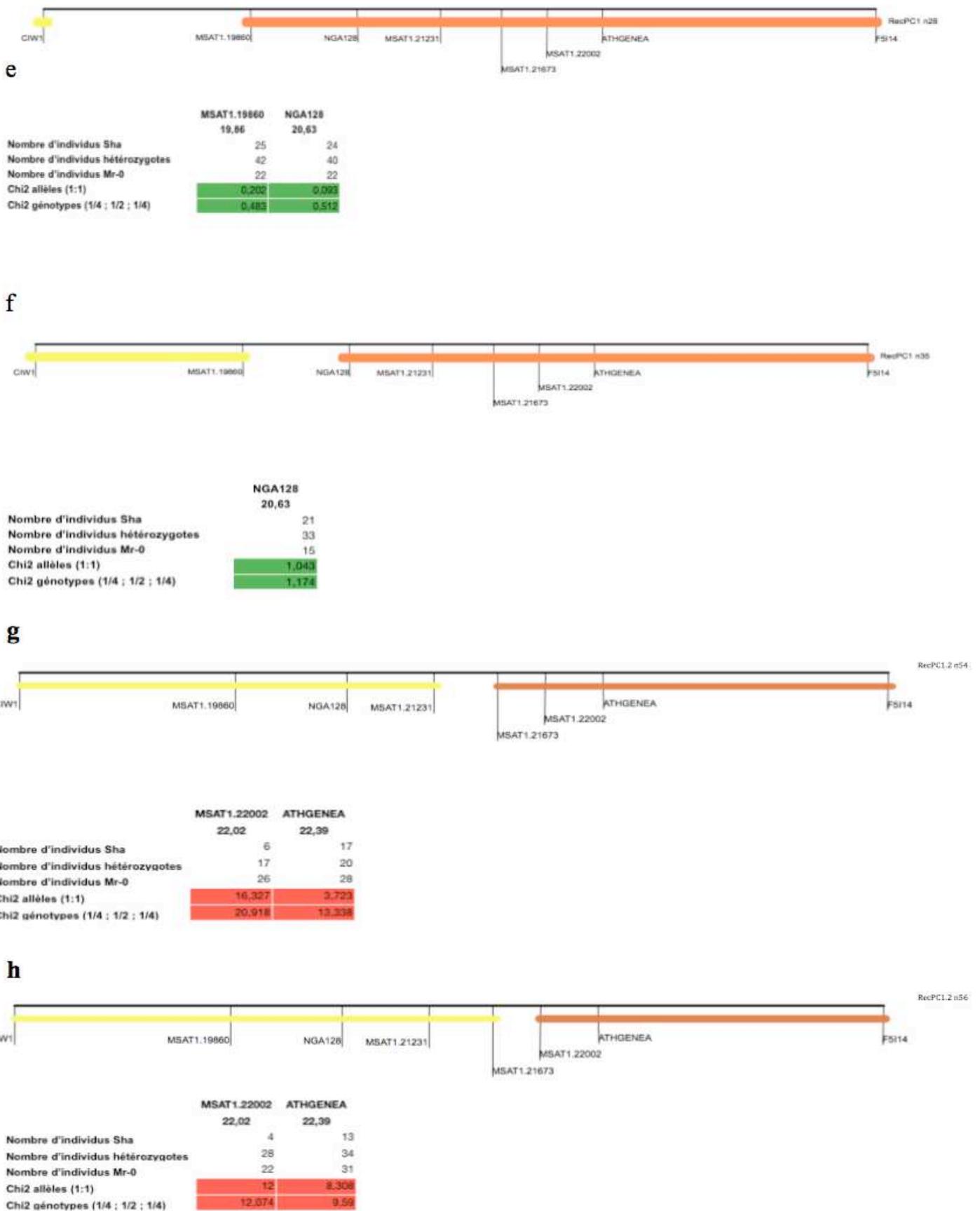


Figure B13 (1/2) : ségrégation des allèles parentaux dans des descendance d'autofécondation de plantes hétérozygotes au chromosome 1



**Figure B13 (2/2) : ségrégation des allèles parentaux dans des descendance d'autofécondation de plantes hétérozygotes au chromosome 1**

Pour chaque plante dont la descendance a été analysée (a à h), le diagramme décrit son génotype dans la région considérée au chromosome 1. Le nom de la plante mère est indiqué à droite. Les traits verticaux représentent la position des marqueurs utilisés pour génotyper ces plantes. Leurs noms et positions physiques sur le chromosome 1 sont indiqués. En jaune, génotype Sha ; en marron : génotype hétérozygote. La ségrégation du ou des marqueurs dans la descendance d'autofécondation de chaque plante est indiquée dans le tableau sous chaque diagramme. Les Chi2 ont été calculés, pour chaque marqueur, pour la répartition des génotypes et des allèles dans la population. Les proportions attendues dans chaque cas sont indiquées entre parenthèses. Les valeurs de Chi2 significatives ( $\alpha=0,05$ ) sont indiquées en rouge et les valeurs de Chi2 non significatives ( $\alpha=0,05$ ) sont indiquées en vert.

populations, des plantes mâles stériles en grande proportion (plus de la moitié des individus). Néanmoins, des plantes hors-types apparaissent également : certains individus sont stériles et Sha au locus ch1, alors que d'autres individus sont fertiles et hétérozygotes au locus ch1. De plus, des plantes ayant un génotype identique à celui des parents fertiles du croisement sont stériles. Ces résultats sont donc difficiles à interpréter, et ces expériences devront être refaites.

### I.3.c Analyse génétique des distortions de ségrégation au chromosome 1

Des croisements ont également été réalisés entre les recombinants PC1.2 de type alpha et bêta en utilisant les recombinants PC1.2 de type bêta comme parents mâles. Dans ces croisements, nous avons observé un très fort biais contre la transmission de l'allèle Mr-0 au locus ch1A (figure B12, 2 individus sur 32) alors que cela n'avait pas été observé dans les croisements où le parent de type bêta était utilisé comme femelle (9 individus sur 15 et 11 individus sur 18, figure B11). Les expériences préliminaires réalisées avant le début de ma thèse, y compris celles réalisées dans la F2-63RV, avaient montré également des biais de ségrégation dans cette région chromosomique. Mes croisements semblent indiquer que ce biais ne se produit que lorsque la plante hétérozygote est utilisée comme mâle. Ceci est confirmé par le fait qu'aucun biais de transmission des marqueurs dans cette région chromosomique n'a été observé dans les populations de cartographie dont les mères sont hétérozygotes dans cette région (alors que les pères sont fixés).

Pour confirmer ce biais de transmission dans la descendance de plantes hétérozygotes dans la région ch1, huit descendances d'autofécondation de recombinants fertiles des PC1 et PC1.2, ont été phénotypées et génotypées avec 8 marqueurs localisés sur le chromosome 1 entre 18,37 Mb et 24,37 Mb (de CIW1 à F5I14) (figure B13). Ces recombinants ont été choisis en fonction de la zone pour laquelle ils sont hétérozygotes dans l'intervalle génétique encadré par les marqueurs CIW1 et F5I14.

L'analyse de ces résultats montre que les marqueurs MSAT1.19860, MSAT1.21231, MSAT1.21673 et MSAT1.22002 présentent une forte distorsion de ségrégation lorsque la plante mère est à la fois hétérozygote à ces marqueurs et hétérozygote au marqueur CIW1 : dans ce cas, l'allèle Mr-0 est fortement contre sélectionné. Dans la descendance d'autofécondation du recombinant PC1 n°3, qui est hétérozygote à CIW1 et Sha à droite de MSAT1.19860, l'allèle Mr-0 est contre sélectionné ( $\chi^2 = 6,3$ ). Dans la descendance des recombinants PC1 n°28 et PC1 n°35, hétérozygotes au sud du chromosome 1 et Sha à gauche de MSAT1.19860, il n'y a pas de distorsion de ségrégation et les  $\chi^2$  sont proches de 0 (0,2

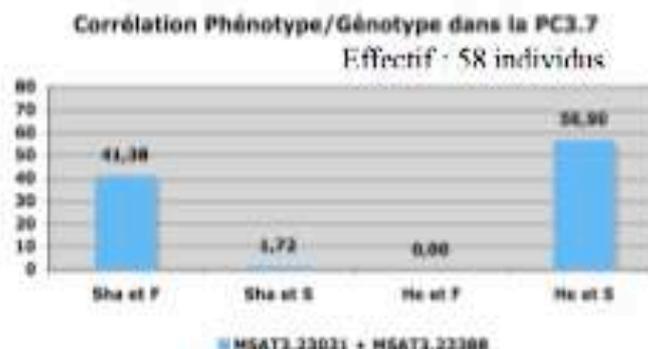
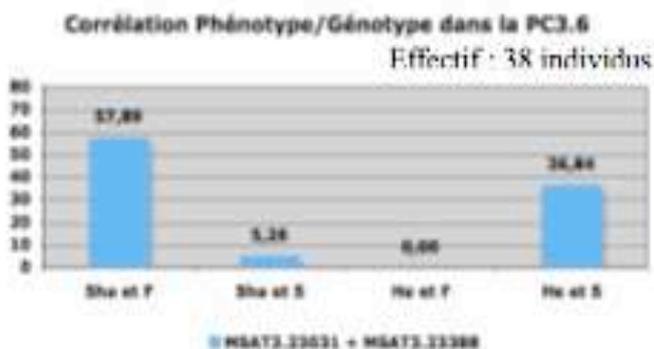
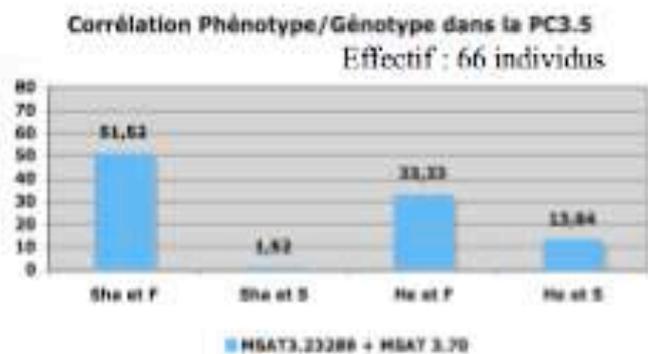
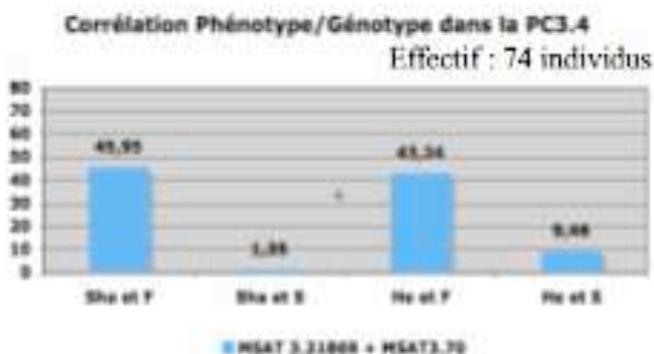
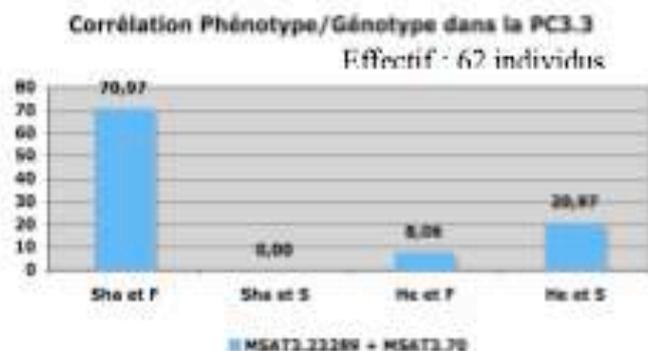
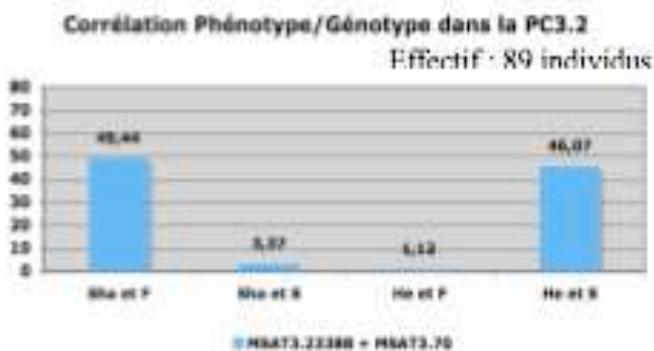
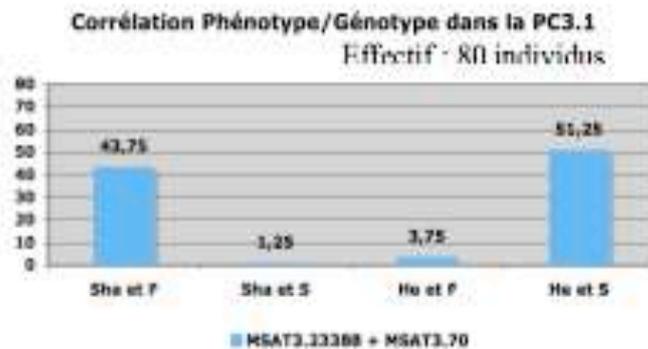
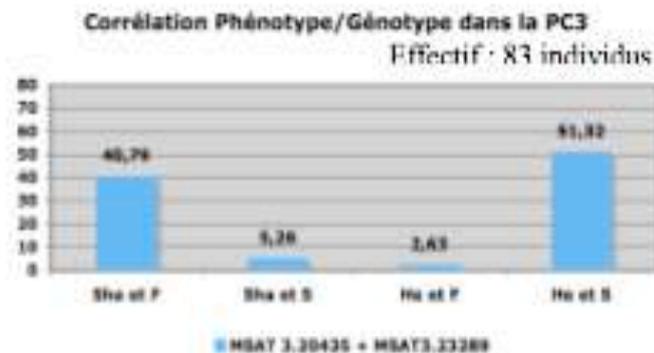
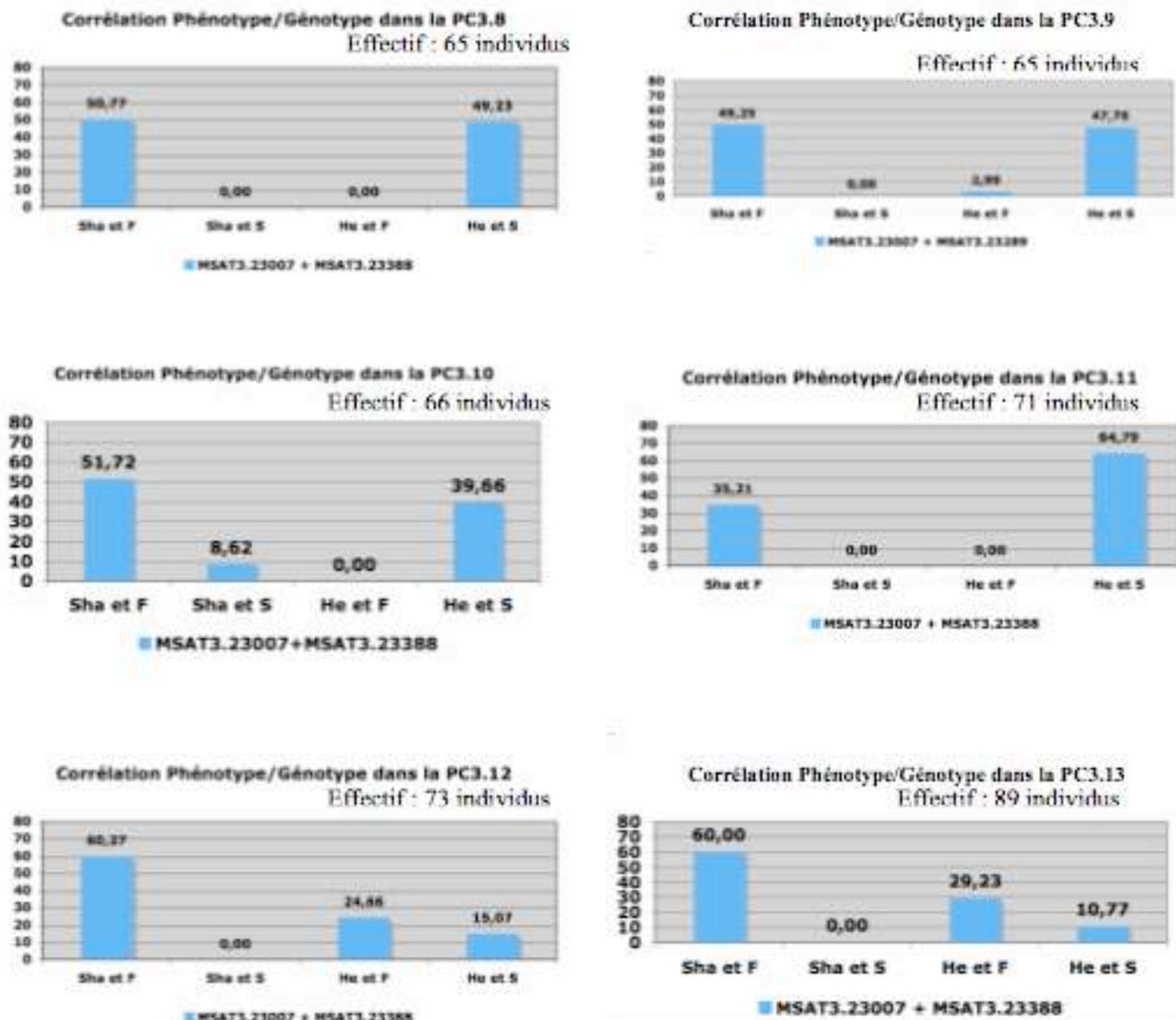


Figure B14 (1/2) : liaison du génotype aux marqueurs encadrant le locus ch3 avec les phénotypes de stérilité (S) et de fertilité (F)



**Figure B14 : liaison du génotype aux marqueurs encadrant ch3 avec les phénotypes de stérilité (S) et de fertilité (F)**

Pour chacune des populations de cartographie, le diagramme représente la distribution (en %) des plantes de la plaque phénotypée selon leur phénotype (stérile, S ou fertile, F) et leur génotype aux deux marqueurs génotypés. Les plantes recombinantes entre deux marqueurs n'ont pas été prises en compte pour cette analyse. Les effectifs concernés sont indiqués au-dessus de chaque diagramme

et 1,04 respectivement). Dans la descendance d'autofécondation des recombinants PC1.2 n°54 (hétérozygote à droite de MSAT1.21231) et PC1.2 n°56 (hétérozygote à droite de MSAT1.21673), l'allèle Sha est fortement contre sélectionné.

Ces résultats indiquent qu'il existe au moins deux locus distorceurs de ségrégation : un locus qui contre sélectionne l'allèle Mr-0, et qui est situé à gauche de MSAT1.19860. Il n'y a pas de co-localisation entre ce locus distorceur de ségrégation, situé à gauche de MSAT1.19860, et les locus ch1A et ch1B mainteneurs de stérilité, situés à droite de MSAT1.19860. Un second locus distorceur de ségrégation, situé à droite de MSAT1.21673, contre sélectionne l'allèle Sha. À ce stade des résultats, ce locus co-localise avec le locus ch1B. Des travaux supplémentaires devront être effectués afin de savoir si ce distorceur a un lien fonctionnel avec le maintien de la stérilité mâle par le locus ch1B.

#### I.4 Cartographie génétique du locus mainteneur de stérilité au chromosome 3

##### I.4.a Analyse des populations de cartographie

14 populations de cartographie (notées de PC3 à PC3.13), soit près de 8200 plantes, ont été semées. Tous les individus de ces populations de cartographie sont homozygotes Mr-0 au locus ch1 (aux marqueurs CIW1, NGA128, ATHGENEA et F5I14). Pour chaque population de cartographie, une plaque de 92 individus a été phénotypée afin de contrôler la liaison entre le phénotype et le génotype aux marqueurs utilisés pour la recherche de recombinants. Ainsi, comme on peut le voir dans la figure B14, le génotype des plantes sur le bras sud du chromosome 3 est lié au phénotype dans la plupart des populations de cartographie : si une plante est hétérozygote à ces marqueurs, la plante est stérile. Au contraire, si une plante est Sha à ces marqueurs, la plante est fertile. Cette liaison est parfaite dans les PC3.8 et PC3.11 où aucune plante hors-type n'a été détectée. Néanmoins, des plantes hors-types existent en plus ou moins grand nombre selon les populations de cartographie : il existe un faible pourcentage (<5%) de plantes hors-types fertiles et hors-types stériles dans les PC3.1, PC3.2, PC3.7 et PC3.9. Dans d'autres populations de cartographies, le pourcentage de plantes hors-types stériles est compris entre 5 et 10%, comme c'est le cas dans la PC3 (5,26 %), la PC3.6 (5,26 %) et la PC3.10 (8,62%). La population PC3.3 contient 8,06 % de hors-types fertiles et aucun hors-type stérile. Enfin, les PC3.4, PC3.5, PC3.12 et PC3.13, ont respectivement 43,23 %, 33,33 %, 24,66 % et 29,23 % de plantes hors-types fertiles, associé à un fort déficit en individus stériles dans ces quatre populations de cartographie. La proportion des phénotypes est, en théorie, de  $\frac{1}{2}$  d'individus stériles et  $\frac{1}{2}$  d'individus fertiles. Comme nous pouvons le

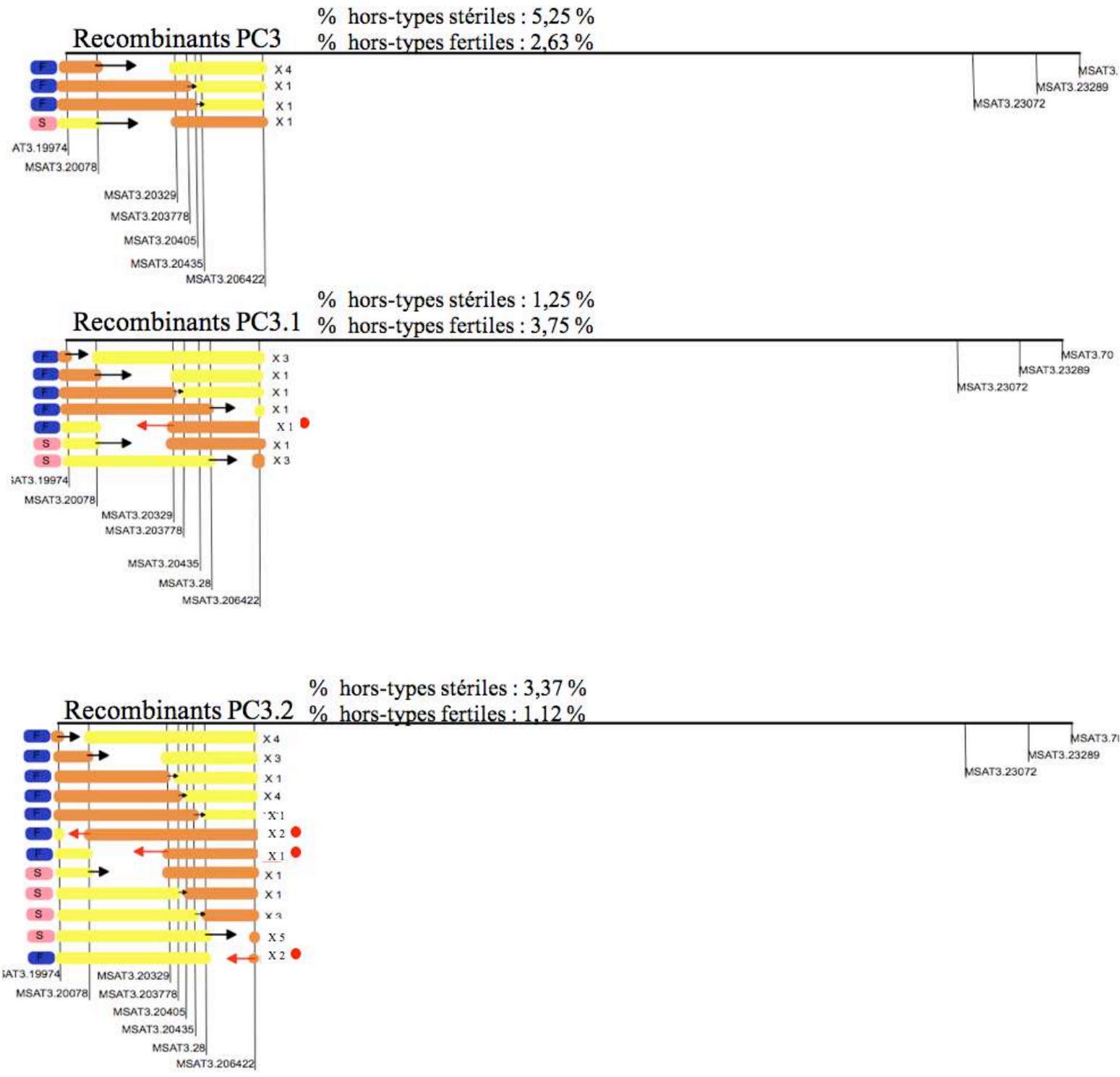


Figure B15 (1/3) : génotypes des recombinants identifiés dans les différentes PC3

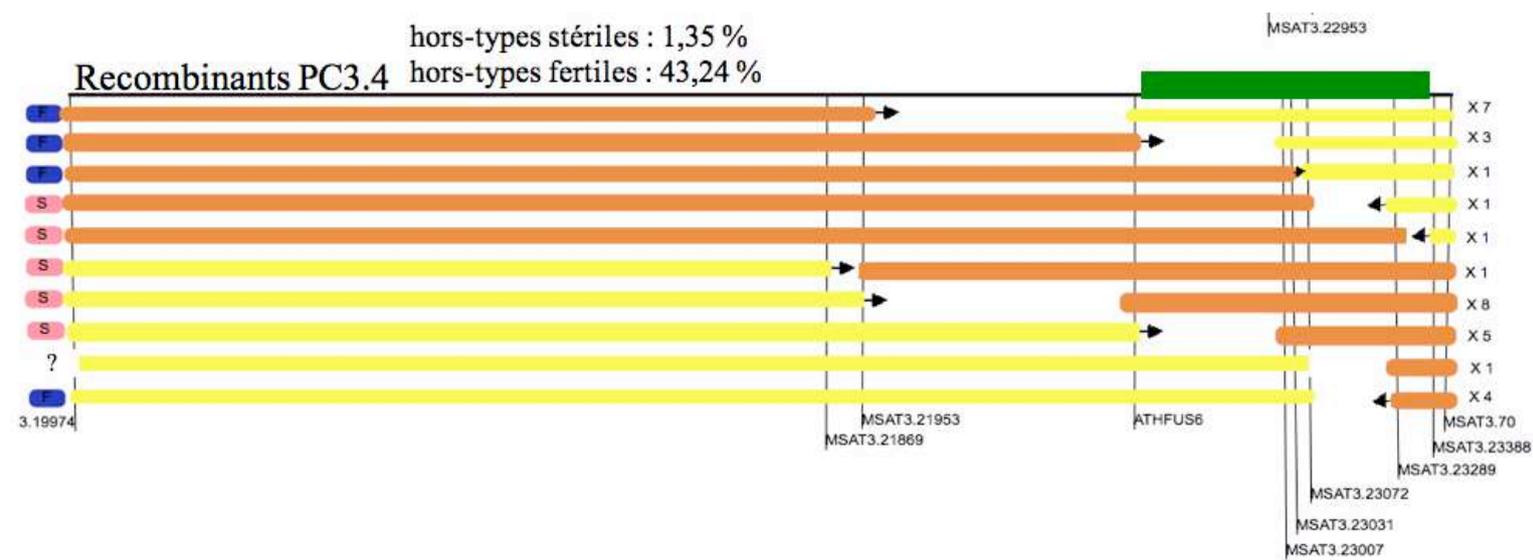
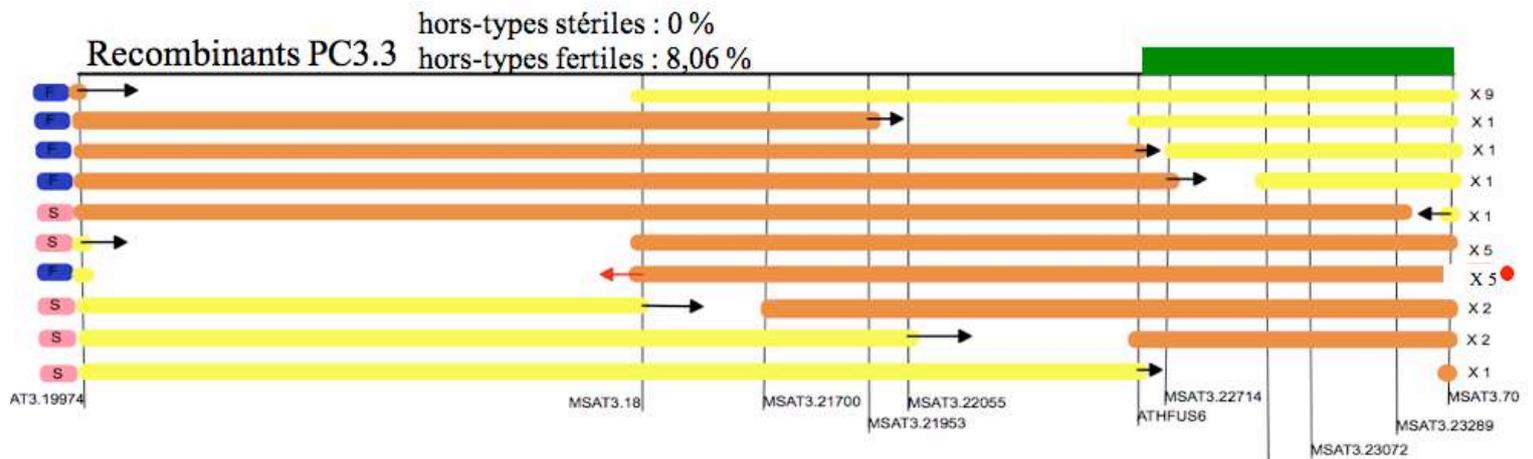
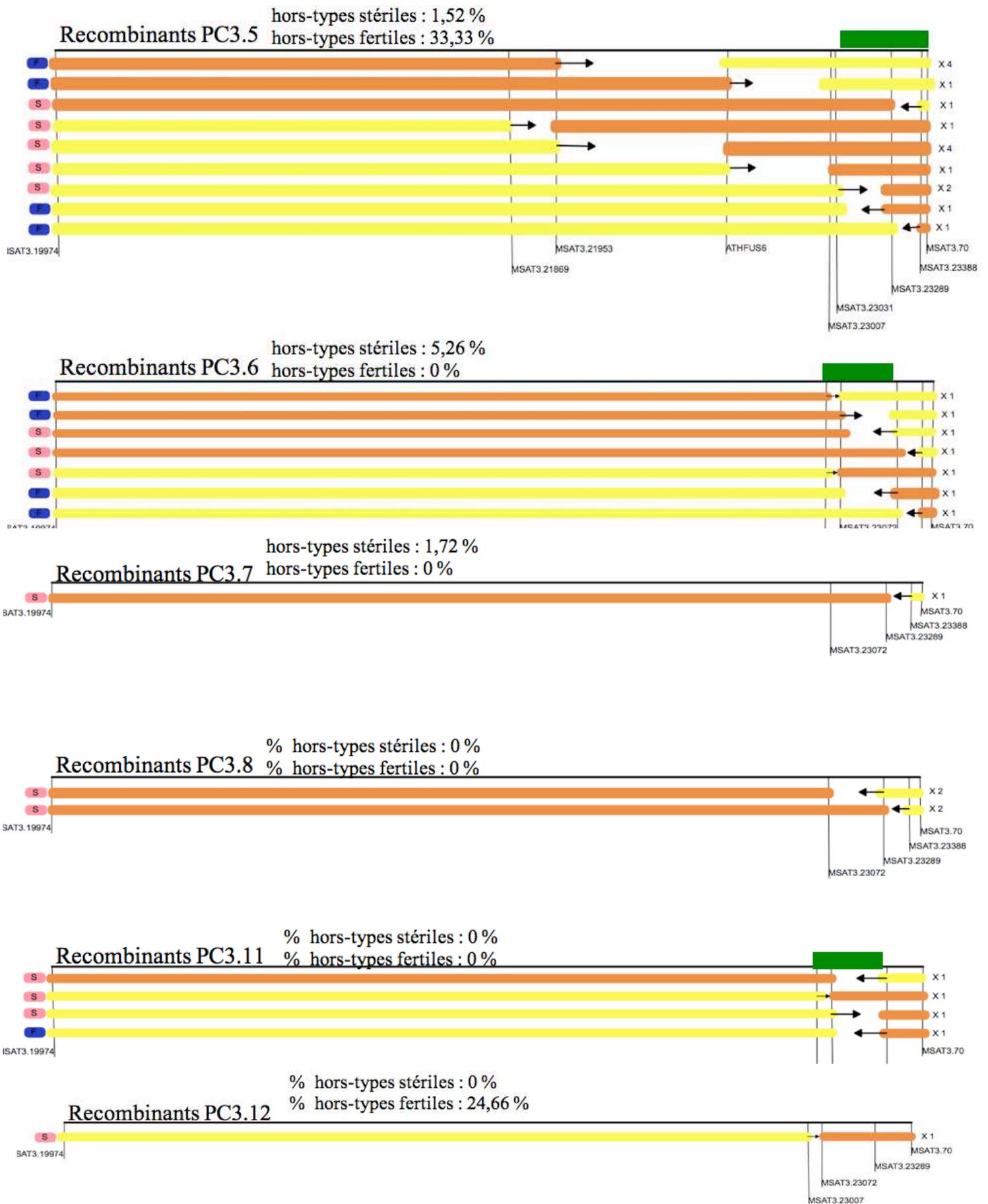


Figure B15 (2/3) : génotypes des recombinants identifiés dans les différentes PC3



**Figure B15 (3/3) : génotypes des recombinants identifiés dans les différentes PC3**

La colonne de gauche indique le phénotype des recombinants (F = fertile , S = stérile et ? = phénotype incertain) et la colonne de droite indique le nombre de recombinants pour un génotype donné. Les traits verticaux représentent la position des marqueurs utilisés pour génotyper ces plantes. En jaune, génotype Sha ; en marron : génotype hétérozygote. Le carré vert représente l'intervalle (contenant le locus ch3) encadré par les recombinants stériles. Les % de plantes hors-types observés dans l'échantillon phénotype de chaque population sont indiqués. Il existe des recombinants fertiles en contradiction avec la localisation génétique du locus ch3 encadré par les recombinants stériles : ils sont indiqués par des points rouges.

	PC3	PC3.1	PC3.2	PC3.3	PC3.4	PC3.5	PC3.6
Nombre d'individus stériles	47	45	46	16	13	12	29
Nombre d'individus fertiles	35	39	45	53	70	75	43
Chi2 génotypes (1:1)	3,51	0,85	0,022	39,681	78,289	91,241	5,444

	PC3.7	PC3.8	PC3.9	PC3.10	PC3.11	PC3.12	PC3.13
Nombre d'individus stériles	46	35	41	32	46	11	3
Nombre d'individus fertiles	31	33	42	34	28	63	61
Chi2 génotypes (1:1)	5,844	0,117	0,024	0,121	8,75	73,081	105,125

### Tableau B6 : ségrégation des phénotypes dans les populations de cartographie du locus ch3

Le Chi2 a été calculé pour une proportion égale de plantes stériles et fertiles, attendue dans ces populations. Les valeurs du Chi2 significatives ( $\alpha=0,05$ ) sont indiquées en rouge.

	PC3	PC3.1	PC3.2	PC3.3	PC3.4	PC3.5	PC3.6
Nom du marqueur	MSAT3.20435	MSAT3.206422	MSAT3.206422	MSAT3.70	MSAT3.70	MSAT3.70	MSAT3.23031
Nombre d'individus Sha	267	378	456	250	489	245	234
Nombre d'individus He	217	394	429	237	492	243	224
Chi2 génotypes (1:1)	5,2	0,837	1,647	0,347	0,009	0,008	0,218

	PC3.7	PC3.8	PC3.9	PC3.10	PC3.11	PC3.12	PC3.13
Nom du marqueur	MSAT3.23388	MSAT3.23031	MSAT3.23007	MSAT3.23007	MSAT3.23388	MSAT3.23388	MSAT3.23388
Nombre d'individus Sha	77	189	60	84	179	149	77
Nombre d'individus He	73	192	64	61	173	160	82
Chi2 génotypes (1:1)	0,106	0,023	0,129	3,64	0,102	0,392	0,159

### Tableau B7 : ségrégation d'un marqueur dans les populations de cartographie du locus ch3

Le Chi2 a été calculé pour une proportion égale de plantes hétérozygotes et Sha, attendue dans ces populations. La valeur du Chi2 significative ( $\alpha=0,05$ ) est indiquée en rouge.

voir dans le tableau B6, les proportions observées d'individus stériles et fertiles sont significativement différentes des proportions théoriques dans les populations PC3.3 (Chi<sup>2</sup> = 39,6), PC3.4 (Chi<sup>2</sup> = 78), PC3.5 (Chi<sup>2</sup> = 91), PC3.6 (Chi<sup>2</sup>=5,4), PC3.7 (Chi<sup>2</sup>=5,8), PC3.11 (Chi<sup>2</sup>=8,75) PC3.12 (Chi<sup>2</sup> = 73) et PC3.13 (Chi<sup>2</sup> = 105). Les populations PC3.4, PC3.5, PC3.12 et PC3.13 ont été produites à partir de génotypes femelles de type 2 (figure B6) où le parent femelle (mâle stérile) a une zone d'hétérozygotie plus petite que le parent femelle de type 1. De plus, ces quatre populations de cartographies ont été obtenues à partir de parents femelles (mâle stériles) issus de la PC3.3, contrairement aux autres populations de cartographies qui ont été obtenues à partir de parents femelles issus des populations « fixation Mr-0 au locus ch1 » RM3-A CC3 X RM3-A CA4. Cette différence de génotype des parents femelles des populations PC3.4, PC3.5, PC3.12 et PC3.13 est peut-être liée au très fort déficit en plantes stériles et au grand nombre de hors-types fertiles observés dans ces populations par rapport aux autres populations de cartographies : il pourrait exister un locus de maintien/restauration dans cette région dont l'effet est mineur et/ou dépendant de ch1 ou ch3. Toutefois, cette hypothèse n'est pas valide pour expliquer le déficit en plantes stériles dans les populations PC3.3, PC3.6, PC3.7 et PC3.11. Ces résultats surprenants nous ont conduit à ne considérer que les recombinants stériles dans les populations de cartographie du locus ch3 mainteneur de stérilité. À noter que, dans ces populations, la ségrégation des allèles aux marqueurs génotypés est très proche des proportions attendues (½ d'hétérozygotes et ½ de Sha), excepté dans la PC3 où un déficit de l'allèle Mr-0 est observé (Chi<sup>2</sup>=5,2) (tableau B7).

#### I.4.b Analyse des recombinants au locus ch3

Au total, plus de 80 recombinants ont été identifiés entre les marqueurs MSAT3.19974 et MSAT3.70 dans les différentes PC3 (figure B15). Aucun recombinant n'a été identifié dans les PC3.9, PC3.10 et PC3.13. Dans un premier temps, les marqueurs MSAT3.19974 et MSAT3.206422 ont été utilisés pour rechercher des recombinants dans les PC3, PC3.1 et PC3.2. Un, quatre et dix recombinants stériles ont été identifiés respectivement dans les populations PC3, PC3.1 et PC3.2. Huit recombinants stériles identifiés dans les populations PC3.1 et PC3.2 indiquent que le locus ch3 est situé à droite de MSAT3.28. La faible proportion de hors-types stériles les populations PC3.1 et PC3.2 (1,25 % et 3,37 %, respectivement), nous autorise à définir la borne de gauche de la région génomique contenant le locus ch3 à MSAT3.28. Cependant, le marqueur MSAT3.19974, plus facile à utiliser que MSAT3.28, sera préféré dans les cartographies suivantes. Un recombinant stérile a été

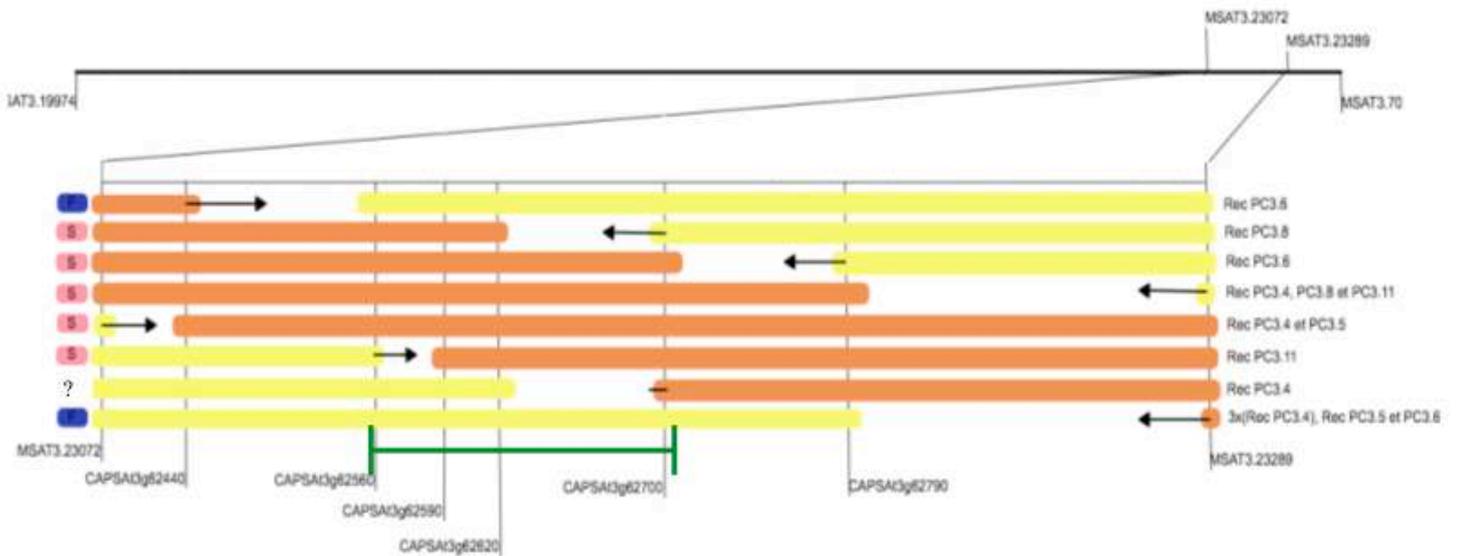
identifié dans la PC3 mais compte tenu de la proportion de plantes hors-types stériles (5,26 %) dans cette population, il n'a pas été pris en compte pour la cartographie du locus ch3.

À cette étape, nous ne pouvons pas définir avec certitude la borne de droite de la région génomique contenant le locus ch3. C'est pour cette raison qu'il a été décidé d'élargir l'intervalle à droite et d'utiliser le marqueur MSAT3.70, situé plus au sud, pour rechercher des recombinants dans les PC3.3, PC3.4 et PC3.5. Des recombinaisons dans la région génomique [MSAT3.19974 – MSAT3.70] ont donc été recherchées et 76 recombinants ont été identifiés dans ces trois populations de cartographie. Parmi ces recombinants, un recombinant stérile dans la PC3.3, cinq recombinants stériles dans la PC3.4 et un recombinant stérile dans la PC3.5, indiquent que le locus ch3 est situé à droite du marqueur ATHFUS6. En ce qui concerne la borne de droite de la région génomique contenant le locus ch3, trois recombinants stériles (un dans chacune des trois populations de cartographie) indiquent que le locus ch3 est situé à gauche de MSAT3.70. Très peu de plantes hors-types stériles sont présentes dans ces populations (0 %, 1,35 % et 1,54 % respectivement).

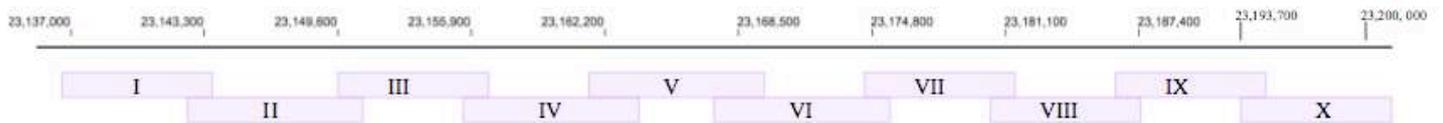
À cette étape, le locus ch3 semble être localisé, d'après les recombinants stériles, dans la région génomique [ATHFUS6-MSAT3.70]. Des marqueurs internes à cet intervalle ont ensuite été utilisés sur les recombinants identifiés dans la région : dans la PC3.4, un recombinant stérile indique que le locus ch3 est situé à droite de MSAT3.23072 et, dans la PC3.5, deux recombinants stériles indiquent que ce locus est situé à droite de MSAT3.23031. Un recombinant stérile de la PC3.4 indique que ch3 serait à gauche de MSAT3.23289, mais un seul recombinant ne nous a pas paru suffisant pour rétrécir l'intervalle de cartographie à ce stade.

Les marqueurs MSAT3.23007 et MSAT3.70 ont ensuite été utilisés pour identifier de nouveaux recombinants dans les PC3.6 à PC3.13. Ainsi, dans la PC3.11, un recombinant stérile indique que le locus ch3 est situé à droite de MSAT3.23072. En ce qui concerne la borne de droite, un recombinant stérile non informatif dans la PC3.6 (car cette population contient 5 % de plantes hors-type stériles), deux recombinants stériles dans la PC3.8 et un recombinant stérile dans la PC3.11 indiquent que le locus ch3 est situé à gauche de MSAT3.23289. Ainsi, la région génomique contenant le locus ch3 est encadré par les marqueurs MSAT3.23072 et MSAT3.23289. Aucun recombinant stérile n'est en contradiction avec la présence du locus mainteneur de stérilité ch3 dans l'intervalle génétique encadré par les marqueurs MSAT3.23072 et MSAT3.23289.

Cet intervalle correspond à une séquence de 217 kb sur la séquence génomique de Col-0. Aucun des microsatellites testés sur cette séquence n'est polymorphe entre Sha et Mr-0. C'est



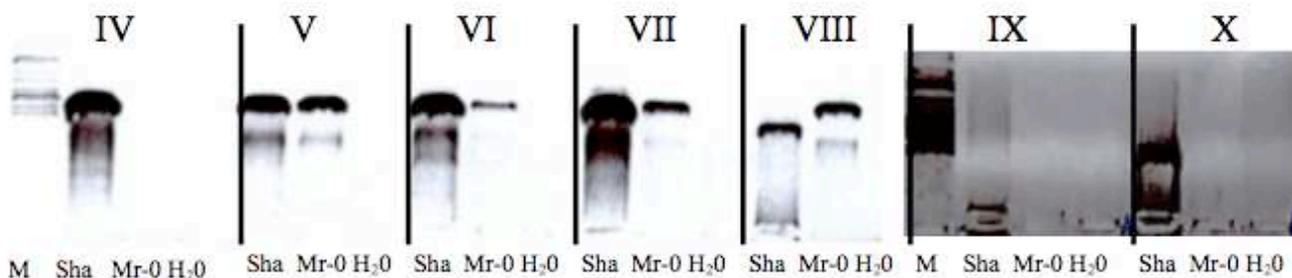
**Figure B16 : cartographie fine du mainteneur de stérilité au locus ch3 à l'aide de marqueurs CAPS.** Génotypes des recombinants dans l'intervalle MSAT3.23072 et MSAT3.23289. La colonne de gauche indique le phénotype des recombinants (F = fertile, S = stérile et ? = phénotype incertain) et la colonne de droite indique le nombre et l'origine des recombinants pour un génotype donné. Les traits verticaux représentent la position des marqueurs utilisés. En jaune, génotype Sha ; en marron : génotype hétérozygote. Le crochet vert indique la position du locus ch3 [CAPSA3g62560-CAPSA3g62700]



N° des fragments	Col-0 (attendu en pb)	Col-0 (en pb)	Sha (en pb)	Mr-0 (en pb)
I	6907	6000 - 8000	6000 - 8000	6000 - 8000
II	6543	6000 - 8000	6000 - 8000	6000 - 8000
III	6479	6000 - 8000	6000 - 8000	6000 - 8000
IV	6727	6000 - 8000	6000 - 8000	Faible amplification à 6000 - 8000
V	6997	6000 - 8000	6000 - 8000	6000 - 8000
VI	6715	6000 - 8000	6000 - 8000	6000 - 8000
VII	6936	6000 - 8000	6000 - 8000	6000 - 8000
VIII	6917	6000 - 8000	>10000	6000 - 8000
IX	6605	6000 - 8000	Faible amplification à 6000 - 8000	Pas d'amplification
X	6483	6000 - 8000	6000 - 8000	Pas d'amplification

**Figure B17 : vérification de la conservation de la région génomique [At3g62560-At3g62700].**

10 couples d'amorces ont été dessinés pour amplifier par PCR, chez Col-0, Sha et Mr-0, des fragments de 6000 à 8000 pb dans cet intervalle. Les résultats pour les trois accessions sont présentés dans le tableau et les positions des 10 fragments chez Col-0 sont représentées sur le schéma.



**Figure B18 : gel d'électrophorèse des produits PCR amplifiant les fragments IV à X de l'ADN Sha et Mr-0**

Les ADNs génomiques des accessions parentales ont été amplifiés avec des amorces dessinées pour amplifier les fragments de la figure B17. Ils ont été séparés par électrophorèse sur un gel d'agarose 1%. M : échelle (1kb DNA ladder Fermentas). Les produits d'amplification des fragments IV à X sont présentés ici.



**Figure B19 : gel d'électrophorèse de confirmation des produits PCR amplifiant le fragment VIII de l'ADN Sha et Mr-0.**

Les ADNs génomiques des accessions parentales ont été amplifiés avec des amorces dessinées pour la région VIII et les produits d'amplification séparés par électrophorèse sur un gel d'agarose 1%. 1 : Sha ; 2 : Mr-0 ; 3 : H<sub>2</sub>O ; M : échelle (1kb DNA ladder Fermentas). Le produit d'amplification obtenu à partir de l'ADN Sha (>10 000 pb) est plus grand que celui obtenu avec l'ADN Mr-0 (6000 pb -8000 pb).

1 2 3 M

pour cette raison que des marqueurs CAPS ont été dessinés dans cet intervalle : six marqueurs CAPS se sont révélés polymorphes entre les parents. Ces CAPS ont ensuite été utilisés pour affiner la cartographie du locus ch3 dans la région génomique [MSAT3.23072 - MSAT3.23289] en utilisant l'ADN de 15 recombinants dans cette région (figure B16). Un recombinant RecPC3.11 stérile est hétérozygote à CAPSAt3g62590 et Sha à CAPSAt3g62560, ce qui exclut la région génomique située à gauche du marqueur CAPSAt3g62560 ; un recombinant RecPC3.8 stérile est Sha à CAPSAt3g62700 et hétérozygote à CAPSAt3g62620, ce qui exclut la région génomique située à droite du marqueur CAPSAt3g62700. Une localisation du locus mainteneur de stérilité ch3 dans l'intervalle génétique ]CAPSAt3g62560-CAPSAt3g62700[, de 53 kb (de 23,137 Mb à 23,190 Mb), n'est contredit par aucun des 15 recombinants fertiles ou stériles.

#### I.4.c Analyse moléculaire de l'intervalle contenant le locus ch3 chez Sha et Mr-0

Afin de vérifier la conservation de la région génomique comprise entre 23, 137 Mb et 23, 190 Mb chez Sha et Mr-0, dix couples d'amorces ont été dessinés pour amplifier par PCR chez trois accessions (Col-0, Sha et Mr-0), des fragments de 6000 à 8000 pb dans cet intervalle (figure B17). Chez Col-0, les fragments amplifiés ont des tailles qui correspondent aux tailles attendues, ce qui indique la validité des amorces et des conditions de PCR (figure B17). Pour Sha, tous les fragments ont les mêmes profils que chez Col-0 à l'exception du fragment IX, qui semble amplifié avec moins d'efficacité que chez Col-0. De façon intéressante, le fragment VIII est plus grand lorsqu'on l'amplifie à partir de l'ADN génomique de Sha, par rapport à sa taille chez Mr-0 (figure B18 et B19). Une insertion pourrait donc se trouver chez Sha entre 23,18 Mb et 23,19 Mb, (coordonnées de la séquence de Col-0). Chez Mr-0, le fragment VIII est de la même taille que chez Col-0. Les fragments IX et X n'ont pas été amplifiés chez Mr-0 et le fragment III a été faiblement amplifié. Les autres fragments ont été correctement amplifiés, à la taille attendue, chez Mr-0. De plus, le séquençage des extrémités des fragments correctement amplifiés chez Sha et Mr-0 correspond bien aux séquences d'ADN prédites chez Col-0.

a

Nom du marqueur	Chromosome 1		Chromosome 3	
	MSAT1.21231	ATHGENEA	MSAT3.23072	MSAT3.23289
Position du marqueur	21,23 Mb	22,40 Mb	23,07 Mb	23,28 Mb
Type 1	Sha	Sha	Sha	Sha
Type 2	He	He	He	He

b

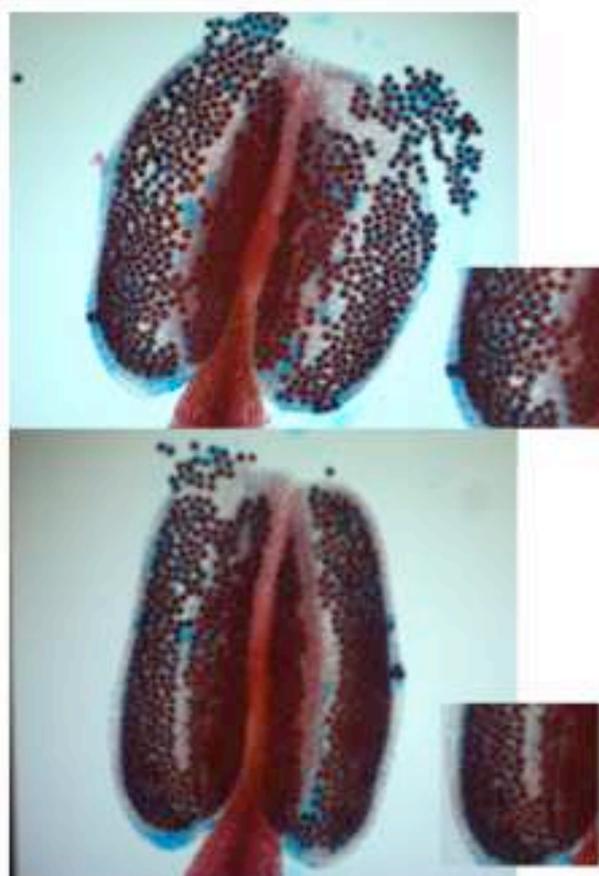


Type 1

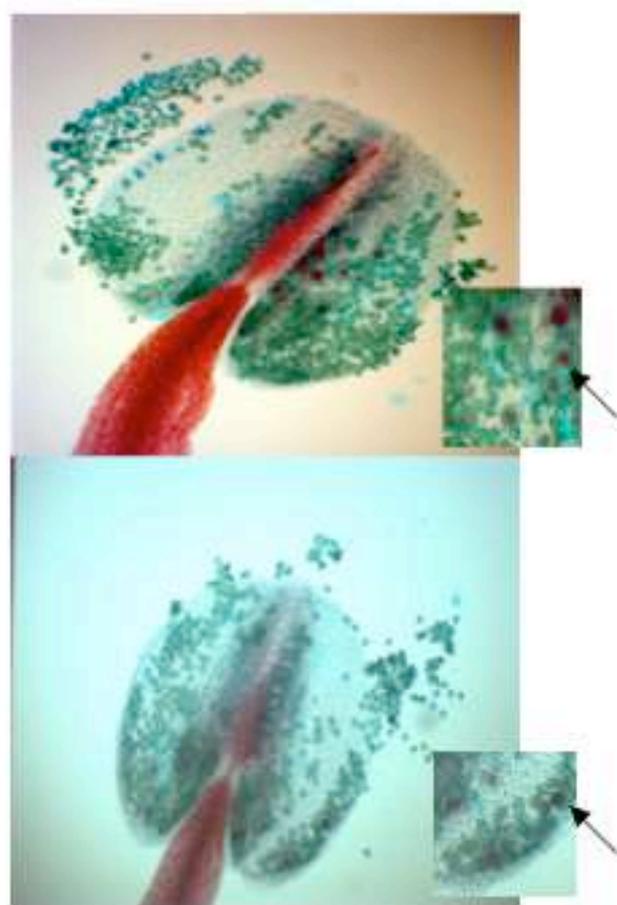


Type 2

c



Type 1



Type 2

**Figure B20 : caractérisation phénotypique de l'incompatibilité nucléo-cytoplasmique Sha/Mr-0**

- a- Génotypage des plantes observées au locus mainteneurs de stérilité. Sha : homozygote Shahdara ; He : hétérozygote.
- b- Hampes florales des plantes matures.
- c- Anthères colorées à l'Alexander. Les grains de pollen sont colorés en rouge s'ils sont viables ou en bleu-vert s'ils sont non viables. L'agrandissement d'une zone du champ est montré en encart. Les rares grains de pollen colorés en rouge visibles dans les anthères des plantes de type 2 sont signalés par des flèches.

## II Caractérisation phénotypique de la production de pollen dans certaines combinaisons génétiques.

### II.1 Caractérisation phénotypique de la stérilité mâle nucléo-cytoplasmique Sha/Mr-0

La caractérisation phénotypique de la stérilité mâle nucléo-cytoplasmique Sha/Mr-0 a été faite en observant la viabilité des grains de pollen et les anthères en développement d'individus fertiles et stériles. Cette caractérisation a été réalisée sur des individus issus d'une population « rétrocroisement maternel 4 » ou RM4. Les graines de cette population proviennent du croisement d'un individu mâle stérile de la RM3-A (RM3-A-CE10) (hétérozygote aux locus ch1A, ch1B et ch3), avec une plante Sha. La population RM4 (25 individus) a été génotypée avec deux marqueurs situés au sud du chromosome 1 (MSAT 1.21231 et ATHGENEA) et deux marqueurs situés au sud du chromosome 3 (MSAT3.23072 et MSAT3.23289). La population RM4 est composée de 17 individus fertiles et de 8 individus stériles : la ségrégation des phénotypes ne s'écarte pas significativement de celle attendue ( $\chi^2(3/1)=0,65$ ), ce qui est également le cas pour la ségrégation des allèles aux quatre marqueurs. Ces résultats, identiques à ceux obtenus dans la population RM3-A, confirment la ségrégation correcte des phénotypes et des marqueurs liés aux locus mainteneurs de stérilité lors des rétrocroisements lorsque les zones d'hétérozygotie sont portées par la mère.

#### II.1.a Observations de la viabilité des grains de pollen

On souhaite caractériser le développement des grains de pollen chez des individus mâle stériles qui possèdent des allèles mainteneurs de stérilité aux trois locus ch1A, ch1B et ch3.

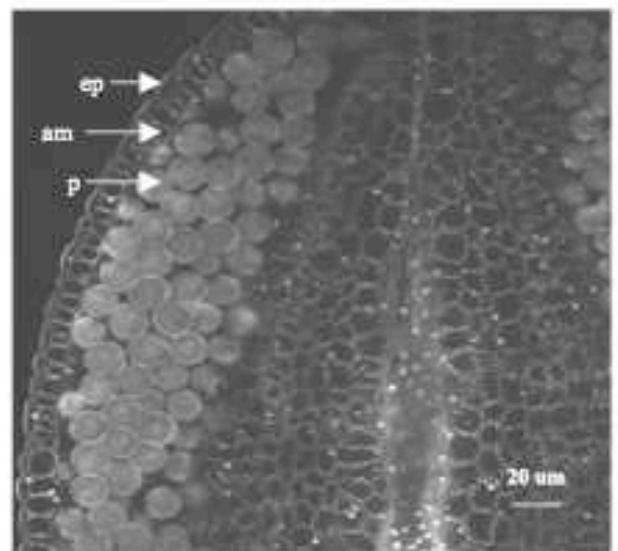
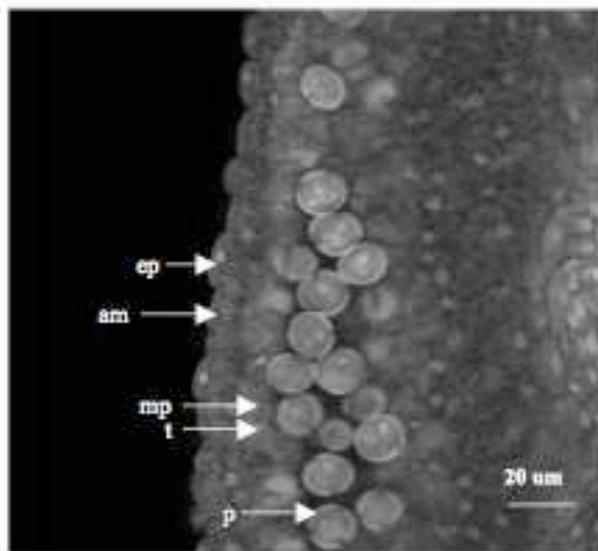
Deux types d'individus ont été sélectionnés dans la population RM4 : des individus de type 1, fertiles, caractérisés par des siliques allongées et homozygotes Sha aux quatre marqueurs génotypés, qui constitueront nos témoins fertiles ; des individus de type 2, mâle stériles, caractérisés par des siliques courtes et hétérozygotes aux quatre marqueurs (figures B20a et B20b). Pour chaque type, deux individus ont été sélectionnés et pour chacun de ces individus, des boutons floraux provenant d'une hampe florale principale et d'une hampe florale secondaire ont été prélevés. La viabilité des grains de pollen a ensuite été observée, en microscopie optique, après coloration d'Alexander. Des résultats types de ces observations sont présentés dans la figure B20c.

Les individus de type 1, qui constituent pour cette expérience la référence « fertilité normale », produisent, comme attendu, des grains de pollen colorés en rouge, donc viables, d'une taille et

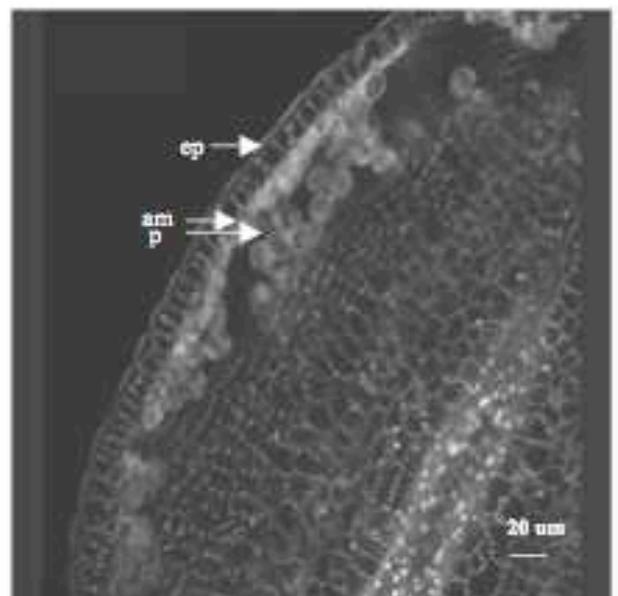
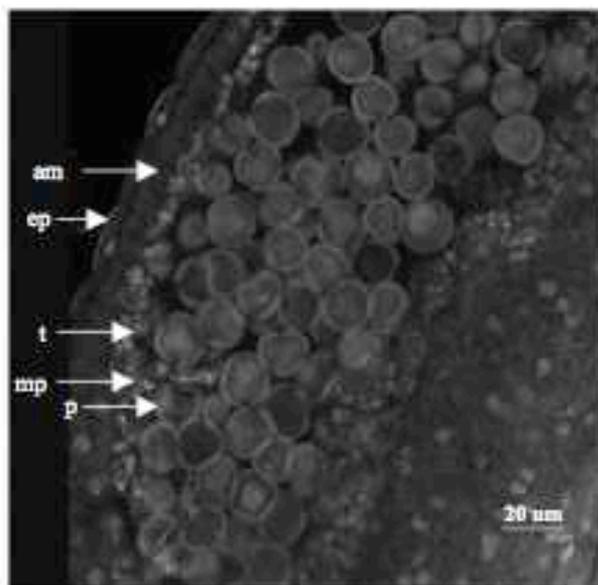
## Stade pollen jeune

## Stade pollen mûr

### Type 1



### Type 2



**Figure B21 : observations cytologiques d'anthères de plantes des types 1 et 2 au cours du développement du pollen**

Les anthères en développement, d'individus de type 1 et de type 2, ont été colorées à l'iodure de propidium et observées au microscope confocal. Deux stades de développement des anthères, un stade précoce (pollen jeune) et un stade tardif (pollen mûr), ont été choisis. L'iodure de propidium colore particulièrement l'ADN, faisant apparaître les noyaux, et les parois cellulaires. Un marquage ponctué (mp) est visible au niveau du tapis de l'anthère. Epiderme (ep), assise mécanique (am), tapis (t), grain de pollen (p).

d'une forme normale (figure B20c). Ces observations sont cohérentes avec le phénotype fertile de ces plantes qui présentent des siliques allongées (figure B20b). Chez ces individus, l'absence des mainteneurs de stérilité aux trois locus ch1A, ch1B et ch3 empêche la stérilité mâle de s'exprimer ; autrement dit, la compatibilité nucléo-cytoplasmique est rétablie chez ces individus.

Chez les individus de type 2, les grains de pollen sont colorés en bleu-vert, ce qui indique qu'ils sont non viables (figure B20c). Ces observations sont cohérentes avec le phénotype stérile de ces plantes qui présentent des siliques courtes (figure B20b). Les allèles Mr-0 mainteneurs de stérilité présents aux locus ch1A, ch1B et ch3, provoquent une incompatibilité avec le cytoplasme de Sha, ce qui aboutit à la mort des grains de pollen. Les grains de pollen des individus de type 2 ont une taille normale, comparable aux grains de pollen des individus de type 1 ; il semblerait donc que les grains de pollen ont avorté à un stade tardif du développement. De plus, nous pouvons observer de rares grains de pollen colorés en rouge, donc probablement viables, dans les anthères des plantes de type 2 (figure B20c). Les plantes observées étant parfaitement stériles, on peut s'interroger sur la capacité de ces grains de pollen à germer et assurer la fécondation d'un sac embryonnaire. On peut aussi rapprocher cette observation de celle, dans plusieurs populations de cartographie, d'individus stériles ayant présenté des siliques allongées en fin de floraison ou sur des hampes florales secondaires. Quel que soit le type d'individu (type 1 ou 2), je n'ai pas observé de différence entre les anthères prélevées sur des hampes florales primaires et secondaires. Il est possible que la pénétrance du phénotype stérile ne soit pas complète dans certaines conditions environnementales et/ou physiologiques. Une vérification de cette hypothèse nécessiterait la conduite d'expériences en conditions contrôlées de génotypes bien caractérisés.

### II.1.b Observations des anthères en développement

Des anthères d'individus de type 1 et de type 2, prélevées à deux stades du développement du pollen, ont été colorées à l'iodure de propidium, un intercalant de l'ADN utilisé pour visualiser les tissus au microscope confocal. Des résultats types de ces observations sont présentés dans la figure B21.

Chez les individus de type 1, les différents tissus qui composent les anthères sont visibles et correctement formés. Au stade « pollen jeune », un marquage ponctué est visible au niveau du tapis de l'anthère : il est possible qu'il s'agisse de la coloration de l'ADN des mitochondries qui sont présentes en grande quantité dans ce tissu et à ce stade de développement. Le tapis

a

	Chromosome 1				Chromosome 3	
	CIW1	MSAT1_19860	MSAT1_21231	ATHGFNFA	MSAT3_23072	MSAT3_23289
	18,37 Mb	19,86 Mb	21,23 Mb	22,40 Mb	23,07 Mb	23,28 Mb
Type 3	He	He	He	He	Sha	Sha
Type 4	Mr-0	Mr-0	Mr-0	Mr-0	Sha	Sha

b



Type 3

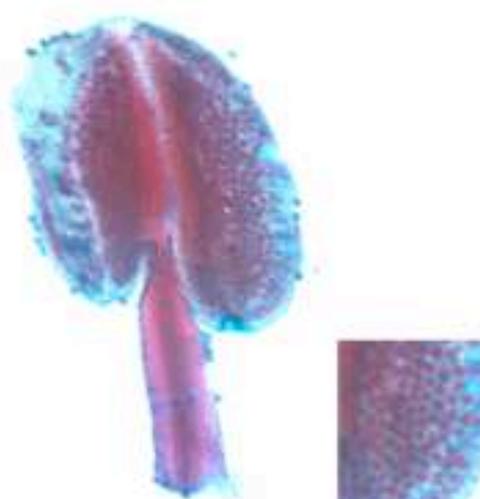
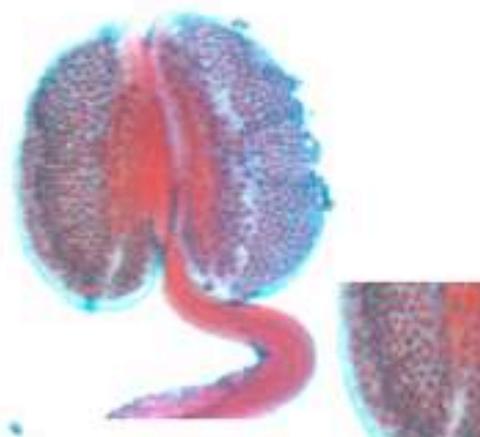


Type 4

c



Type 3



Type 4

Figure B22 : caractérisation phénotypique du distorreur de ségrégation du chromosome 1

- a- Génotypage des plantes observées au locus mainteneurs de stérilité. Sha : homozygote Shahdara ; He : hétérozygote.
- b- Hampes florales des plantes matures.
- c- Anthères colorées à l'Alexander. Les grains de pollen sont colorés en rouge s'ils sont viables ou en bleu-vert s'ils sont non viables. L'agrandissement d'une zone du champ est montré en encart. Les grains de pollen colorés en bleu-vert visibles dans les anthères des plantes de type 3 sont signalés

a

	Chromosome 1			Chromosome 3		
	CIW1	MSAT1.19860	MSAT1.21231	ATHGENEA	MSAT3.23072	MSAT3.23289
	18,37 Mb	19,86 Mb	21,23 Mb	22,40 Mb	23,07 Mb	23,28 Mb
Type 5	Sha	Sha	Sha	Sha	He	He

b



Type 5

c



Type 5

**Figure B23 : caractérisation phénotypique du distorateur de ségrégation du chromosome 3**

- a- Génotypage des plantes observées au locus mainteneurs de stérilité. Sha : homozygote Shahdara ; He : hétérozygote.
- b- Hampes florales des plantes matures.
- c- Anthères colorées à l'Alexander. Les grains de pollen sont colorés en rouge s'ils sont viables ou en bleu-vert s'ils sont non viables. L'agrandissement d'une zone du champ est montré en encart. Les grains de pollen colorés en bleu-vert visibles dans les anthères des plantes de type 5 sont signalés par une flèche.

dégénère au cours de la maturation des grains de pollen pour finalement disparaître à un stade « pollen mûr » du développement de l'anthère. À ce stade tardif, les grains de pollen semblent correctement formés, ce qui est cohérent avec la viabilité des grains de pollen constatée à l'Alexander (figure B20c) et le phénotype de siliques allongées (figure B20b).

Chez les individus de type 2, mâle stériles, les observations montrent une dégénérescence générale des grains de pollen au stade « pollen mûr » (figure B21), ce qui est cohérent avec l'avortement de ces grains de pollen (figure B20c) et l'absence d'allongement des siliques (figure B20b) constatés chez ces individus. Au stade de développement plus précoce que j'ai observé, les anthères de type 2 contiennent des grains de pollen sombres, signe d'une vacuolisation anormale, et déformés par rapport à ceux de type 1. Cependant, certains grains semblent normalement formés et comparables à ceux observés chez les individus de type 1 (figure B21). Ces observations tendent à confirmer que les grains de pollen des plantes stériles avortent à un stade tardif de leur développement. De plus, il semblerait que le marquage ponctué, présent au niveau du tapis de l'anthère au stade « pollen jeune », soit plus intense chez les individus de type 2 que chez les individus de type 1 (figure B21). Nous savons que les mitochondries ont un lien avec la mort cellulaire programmée du tapis. Il se peut qu'un dysfonctionnement de la mort cellulaire programmée puisse expliquer ces observations : un retard de la dégénérescence du tapis peut causer une stérilité mâle (Wu et Cheung, 2000). Cela reste une observation préliminaire qui demande confirmation, notamment par marquage des mitochondries.

## II.2 Caractérisation phénotypique des distortions de ségrégation Sha/Mr-0

Nous savons que des locus situés sur les chromosomes 1 et 3, à proximité des locus mainteneurs de stérilité, étaient capables, à l'état hétérozygote, d'induire une distorsion de ségrégation par la voie mâle. Le locus distorateur de ségrégation du chromosome 1, situé à gauche du marqueur MSAT1.19860, induit un biais de transmission par une contre sélection de l'allèle Mr-0 dans la descendance alors que celui du chromosome 3 induit une contre sélection de l'allèle Sha dans la descendance. Pour observer le pollen dans des plantes présentant ces distorateurs de ségrégation à l'état hétérozygote, des individus de la population RM4 ont été sélectionnés. Des individus de type 3, hétérozygotes au sud du chromosome 1 et Sha au sud du chromosome 3 (figure B22a) ont été utilisées pour observer l'effet du distorateur de ségrégation du chromosome 1 ; des individus de type 5, Sha au sud du chromosome 1 et hétérozygotes au sud du chromosome 3 (figure B23a) ont été utilisées pour observer l'effet du

distorteur de ségrégation du chromosome 3. Chez les individus de type 3 et de type 5, la stérilité mâle nucléo cytoplasmique n'est théoriquement plus maintenue ils sont homozygotes Sha à au moins un locus mainteneur : l'observation éventuelle de grains de pollen avortés ne pourra donc pas être imputée à un effet de l'autre locus mainteneur de stérilité . Pour chaque type, deux individus ont été sélectionnés et pour chacun de ces individus, des boutons floraux provenant d'une hampe florale primaire et d'une hampe florale secondaire ont été prélevés.

## II.2.a Distorsions de ségrégation au chromosome 1

Deux distorsions de ségrégation avaient été observées au sud du chromosome 1 (chapitre I.3.c) : un locus est situé à gauche de MSAT1.19860 et induit une contre sélection des allèles Mr-0 dans la descendance par la voie mâle ; un second locus, situé à droite de MSAT1.21673, induit une contre sélection des allèles Sha dans la descendance par la voie mâle. Ce second locus n'a été observé que dans des plantes fixées Sha à gauche du marqueur MSAT1.21231 (figure B13).

Les individus de type 3 sont fertiles et présentent des siliques normalement allongées (figure B22b), comparables aux siliques des individus de type 1 (figure B20b). Après coloration à l'Alexander, nous observons de nombreux grains de pollen bleu-vert, ce qui indique qu'une proportion significative du pollen produit par les individus de type 3 est non-viable (figure B22c). Si on se réfère aux résultats de génétique obtenus avec des plantes dont le génotype au chromosome 1 était comparable à celui des individus de type 3, mais étaient fixées Mr-0 au ch3 (Figure B13), , on attend une proportion d'environ 50% de pollen avorté dans leurs siliques. L'estimation de cette proportion est assez difficile à faire, et n'a pas pu être réalisée sur mes colorations. Pour le locus distorteur de ségrégation situé à gauche de MSAT1.19860 sur le chromosome 1, il est probable qu'il soit à l'état hétérozygote chez les individus de type 3. Nous avons vu que ce locus, lorsqu'il est à l'état hétérozygote, induit une forte contre sélection des allèles Mr-0 dans la descendance par la voie mâle ; cette contre sélection peut s'expliquer par la mort des grains de pollen qui portent l'allèle Mr-0 chez des individus hétérozygotes au locus distorteur de ségrégation. Les grains de pollen non viables observés chez les individus de type 3 (figure B22c) pourraient être des grains de pollen qui portent l'allèle Mr-0 au locus distorteur de ségrégation du chromosome 1. La question que l'on peut se poser est la suivante : les plantes homozygotes pour l'allèle Mr-0 au chromosome 1 sont également fertiles, mais ne produisent-elles vraiment que du pollen normal ? Pour répondre à cette question, des anthères d'individus de type 4, homozygotes Mr-0 pour la plupart des

marqueurs du chromosome 1 et Sha au chromosome 3 (figure B22a), obtenus par autofécondation de l'individu A8 issu du croisement RM3-A CC3 X RM3-A CA4, ont également été observés après avoir été colorés à l'Alexander : les grains de pollen sont tous colorés en rouge et donc parfaitement viables (figure B22c). L'allèle Mr-0 est donc transmis normalement par la voie mâle chez un individu homozygote Mr-0 au locus distorateur de ségrégation. Ainsi, les individus homozygotes Sha ou Mr-0 au sud du chromosome 1 produisent des grains de pollen qui sont tous viables (type 1 et type 4 respectivement). Ces observations suggèrent que les grains de pollen portant l'allèle contre-sélectionné au distorateur de ségrégation actif dans les plantes de type 3 n'avortent que s'ils sont produits par une plante hétérozygote, mais pas s'ils sont produits par une plante homozygote.

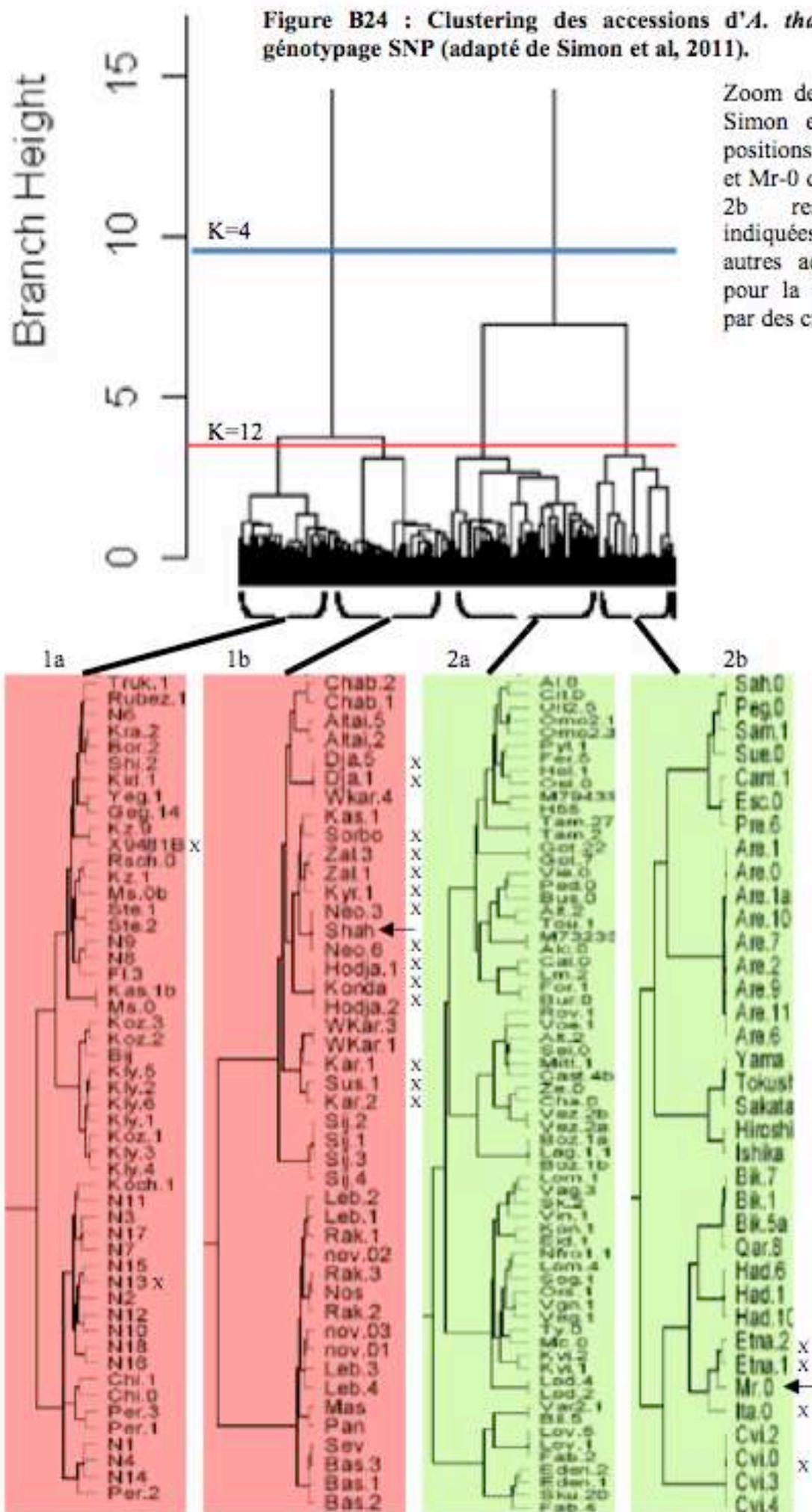
### II.2.b Distorsions de ségrégation au chromosome 3

Les individus de type 5 ont également des siliques normalement allongées (figure B23b), comparables aux siliques des individus de type 1 (figure B20b). Cependant, la coloration d'Alexander montre qu'environ la moitié des grains de pollen produits par ces plantes est non-viable (figure B23c). Nous avons observé un fort déficit de la transmission de l'allèle Sha au sud du chromosome 3 dans la descendance par la voie mâle de plantes hétérozygotes ; cette contre sélection peut s'expliquer par la mort des grains de pollen qui portent l'allèle Sha chez des individus hétérozygotes au locus distorateur de ségrégation. Les grains de pollen avortés observés chez les individus de type 5 (figure B23c) pourraient donc être des grains de pollen qui portent l'allèle Sha au locus distorateur de ségrégation du chromosome 3. Or les plantes homozygotes Sha au sud du chromosome 3, comme par exemple les plantes de type 1, produisent du pollen parfaitement viable (Figure B20c). Ces observations indiquent que les grains de pollen porteurs de l'allèle Sha au sud du chromosome 3 avortent s'ils sont produits par une plante hétérozygote, mais pas s'ils sont produits par une plante homozygote.

### III Répartition des facteurs de la stérilité mâle nucléo-cytoplasmique Sha x Mr-0 au sein de la diversité naturelle de l'espèce

L'apparition d'un phénotype particulier dans la F1 d'un croisement entre deux accessions pourrait être anecdotique, spécifique de ce croisement (due par exemple à une mutation qui serait survenue dans l'une des accessions depuis sa mise en collection). Elle pourrait au contraire refléter une situation issue de l'histoire évolutive des accessions concernées. Dans

Figure B24 : Clustering des accèsions d'*A. thaliana* d'après un génotypage SNP (adapté de Simon et al, 2011).

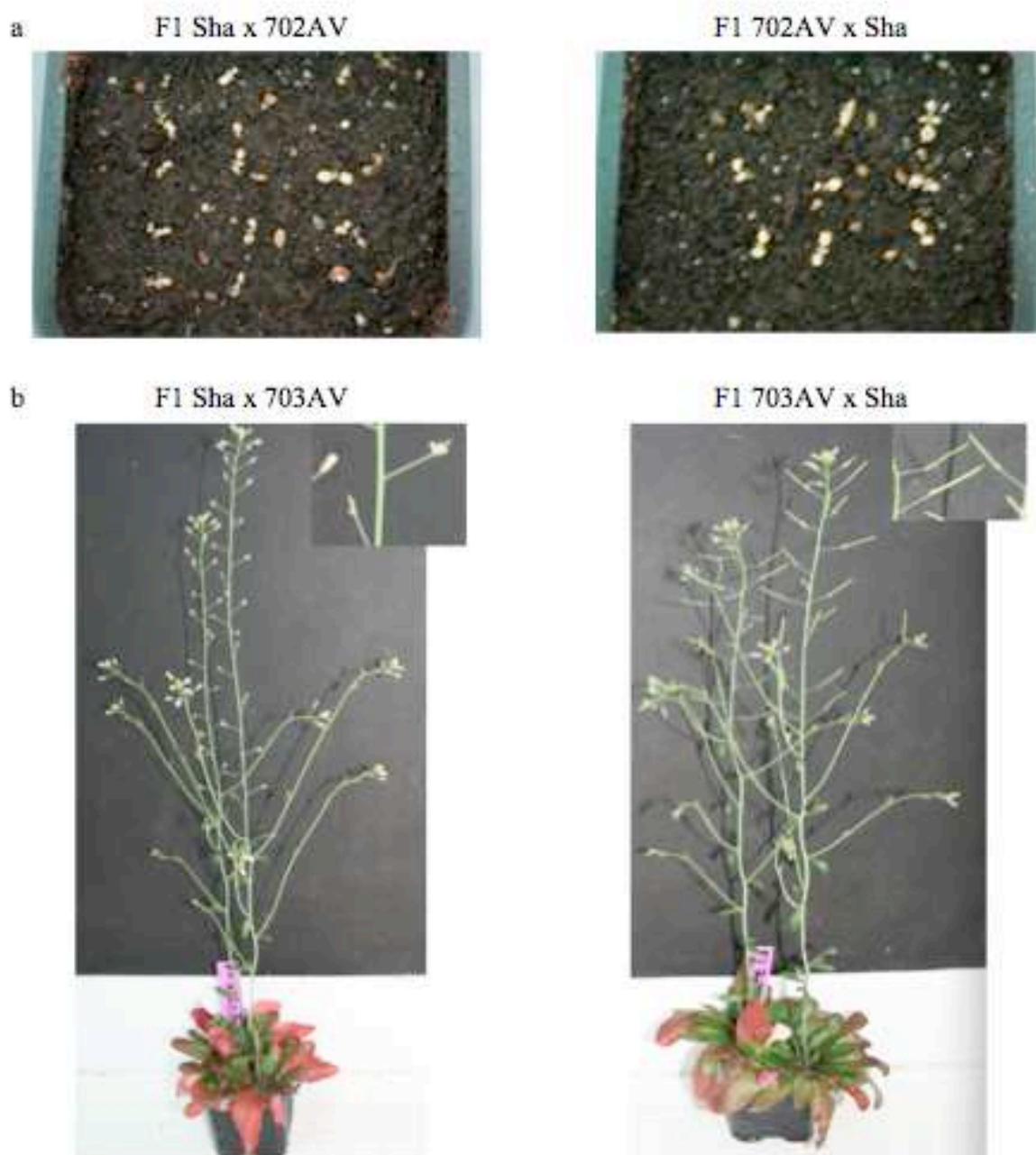


Zoom des clusters 1 et 2 de Simon et al. (2011) : Les positions des accèsions Sha et Mr-0 dans les clusters 1b et 2b respectivement sont indiquées par des flèches. Les autres accèsions pertinentes pour la thèse sont indiquées par des croix.

Accession name	Versailles identification number	Country of origin	Locality of origin	F1 Sha x Acc	F1 Acc x Sha	Croisements	Cluster (Simon et al., 2011)
Mr-0	148 AV	Italy	Monterosso	S	F	MDT 06	2b
Etna-2	703 AV	Italy	Mt. Etna	S	F	NG juin 11	2b
Etna-1	702 AV	Italy	Mt. Etna	Létalité	Létalité	NG juin 11	2b

**Tableau B8 : caractéristiques des accessions 702AV, 703AV et Mr-0**

Ces accessions (colonnes 1 et 2) ont été utilisées comme parents lors des croisements réciproques avec Sha. Les croisements ont été réalisés par Mylène Durant-Tardif (MDT) et par moi-même (NG). Les phénotypes des F1 obtenues par les croisements Sha x Accession (Acc) et Acc x Sha sont indiqués dans les colonnes 5 et 6 respectivement. S = Stérile et F = fertile. Les trois accessions appartiennent au cluster nucléaire 2b (figure B24) d'après Simon et al. (2011).



**Figure B25 : phénotype des F1 obtenues par des croisements réciproques entre Sha et les accessions 702AV et 703AV**

a- Plantules F1 issues de croisements réciproques entre Sha et 702AV environ 10 jours après semis.  
 b- Plantes F1 issues de croisements réciproques entre Sha et 703AV.

ce dernier cas, les facteurs génétiques impliqués dans le phénotype observé devraient également être présents dans d'autres accessions, et leur répartition constituera une information pertinente pour comprendre l'origine évolutive du phénomène. Pour répondre à cette question, différentes accessions de la collection ont été utilisées pour générer des F1 dont les phénotypes ont été observés.

### III.1 Recherche d'accessions mainteneuses de la stérilité mâle nucléo-cytoplasmique induite par le cytoplasme de Sha

Des croisements ont été réalisés entre Sha et d'autres accessions que Mr-0 afin d'identifier d'éventuelles accessions capables de maintenir la stérilité mâle en présence du cytoplasme stérilisant de Sha. Pour choisir les accessions à croiser avec Sha, nous avons recherché des accessions dont le génome nucléaire est proche de celui de Mr-0. Pour cela, nous avons exploité les résultats obtenus par Simon et al. (2011), sur le génotypage par 341 marqueurs SNP de l'ensemble de la collection de Versailles (figure B24). D'après ces résultats, parmi les accessions les plus proches de Mr-0, figurent Ita-0, Cvi-0 et deux accessions italiennes Etna-1 (702AV) et Etna-2 (703AV). Les F1 réciproques de Sha avec Ita-0 et Cvi-0 avaient été obtenues auparavant par l'équipe du Centre de Ressources d'Arabidopsis et étaient fertiles. J'ai utilisé les accessions 702AV et 703AV pour réaliser des croisements réciproques avec l'accession Sha. Les croisements sont résumés dans le tableau B8 et les phénotypes des populations F1 présentés dans la figure B25.

Les F1 réciproques obtenues avec les accessions Sha et 702AV présentent un phénotype inattendu : les graines germent normalement puis les plantules deviennent blanches et meurent (figure B25a). Étant donné que le phénotype est visible dans les deux F1 réciproques, il ne semble pas être dû à une interaction nucléo-cytoplasmique mais plus probablement à des interactions entre allèles parentaux. Nous ne pouvons donc pour l'instant pas conclure sur la présence d'allèles de maintien de la stérilité mâle dans l'accession 702AV.

La F1 du croisement Sha x 703AV est mâle stérile alors que celle du croisement réciproque est fertile (figure B25b) : c'est un résultat identique à celui qui avait été obtenu en utilisant Sha et Mr-0 comme parents. Un rétrocroisement avec le parent Sha a été réalisé sur deux individus F1 stériles pour générer deux populations RM1 (2 X 92 individus) réparties sur deux plaques. Les deux populations ont été phénotypées : sur l'ensemble des deux plaques, 43

a

### Chr1

Nom du marqueur

Position du marqueur (en Mb)

CIW12	NGA128	ATHGENEA	MSAT1.13
9,62	20,63	22,40	25,83

### Chr2

Nom du marqueur

Position du marqueur (en Mb)

MSAT 2.003	MSAT 2.22
0,30	19,63

### Chr3

Nom du marqueur

Position du marqueur (en Mb)

MSAT 3.0205	MSAT 3.1	MSAT 3.65
2,05	12,17	19,84

### Chr4

Nom du marqueur

Position du marqueur (en Mb)

MSAT 4.35	MSAT 4.15
7,55	15,30

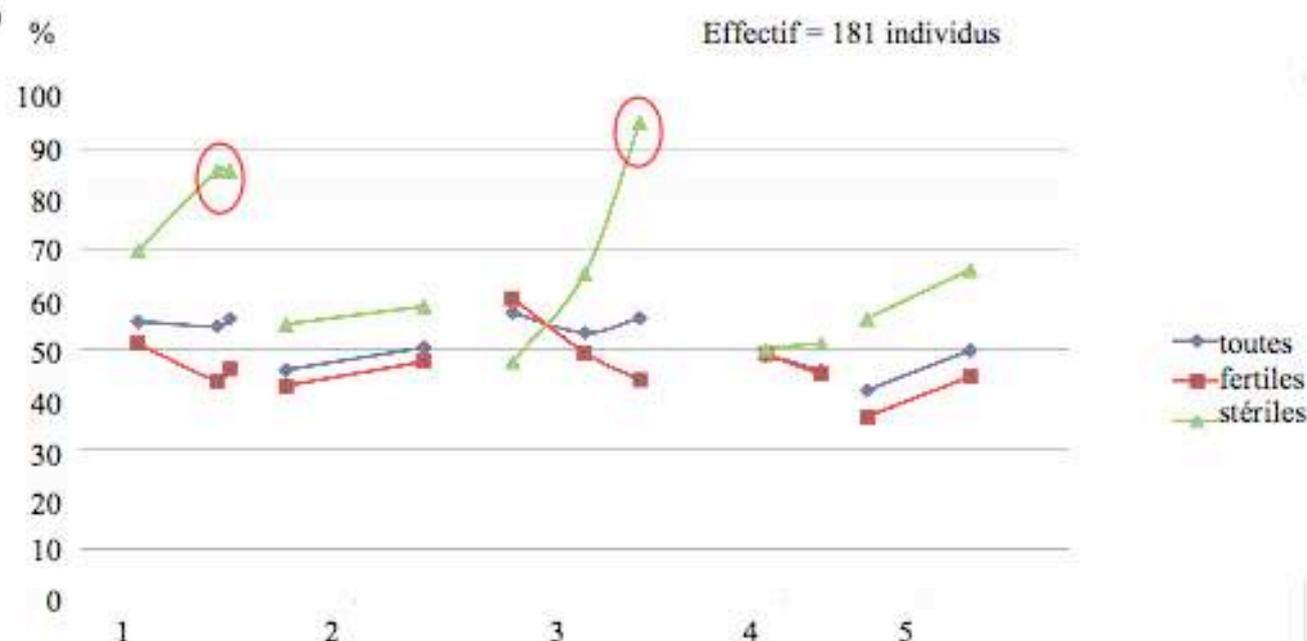
### Chr5

Nom du marqueur

Position du marqueur (en Mb)

NGA 158	ICE3
1,7	16,14

b



**Figure B26 : analyse de la population « rétrocroisement maternel I » (RM1) du croisement Sha X 703AV.**

a- Noms et les positions des marqueurs utilisés pour le génotypage, par chromosome.

b- Représentation graphique des résultats de génotypage. Le pourcentage d'individus hétérozygotes aux marqueurs analysés, est indiqué pour l'ensemble de la population en bleu ; pour les plantes fertiles en rouge; pour les plantes stériles en vert. Les cercles rouges entourent les résultats s'écartant significativement ( $\alpha=0,05$ ) de la proportion attendue.

Nom de l'accession	Numéro d'identification Versailles	Chlorotype	Cluster nucléaire (Simon et al., 2011)	Ville	Pays	Croisements	F1 AccxMr-0	F1 Mr-0+Acc
Db-1	132 AV	J	4b	Tenne/Ts.	Germany	CW	F	F
Wil-1	72 AV	J	4e	Vilnius	Lithuania	CW	F	F
N13	266 AV	J	1a	Konchezero	Russia	CW	F	F
Kas-2	555 AV	H *	1a	Kashmir	Inde	NG	F	F
Kondara	190 AV	S *	1b	Kondara	Tadjikistan	NG	F	F
Hodja•Obi•Garm	203 AV	S	1b	Khodja•Obi•Garm	Tadjikistan	CW	F	F
Sorbo	238 AV	F *	1b	Sorbo	Tadjikistan	NG	F	F
Hodja-Obi-Garm bis	270 AV	S *	1b	Hodja-Obi-Garm	Tadjikistan	NG	F	F
Neo-6	540 AV	U	1b	Neo-Shahdara	Tadjikistan	CW	F	F
Kz-9	403 AV	I *	1a	Atasu	Kazakhstan	NG	F	F
Kz-1	402 AV	J *	1a	Atasu	Kazakhstan	NG	F	F
Lenk2-1	584 AV	M *	1a	Lankaran-Lenk	Azerbadjan	NG	F	F
9481B	261 AV	I	1a	Kazakhstan	Kazakhstan	CW	S	F
Kar-1	531 AV	I *	1b	Karakol river	Kyrgyzstan	NG	S	F
Kar-2	532 AV	I *	1b	Karakol river	Kyrgyzstan	NG	S	F
Sus-1	533 AV	I *	1b	Susamyr	Kyrgyzstan	NG	S	F
Dja-1	534 AV	I *	1b	Djarly	Kyrgyzstan	NG	S	F
Dja-5	535 AV	I *	1b	Djarly	Kyrgyzstan	NG	S	F
Zai-1	536 AV	I	1b	Zalisky	Kyrgyzstan	NG	S	F
Zai-3	537 AV	I *	1b	Zalisky	Kyrgyzstan	NG	S	F
Kyr-1	538 AV	I *	1b	Kyrgyzata	Kyrgyzstan	NG	S	F
Neo-3	539 AV	I	1b	Neo-Shahdara	Tadjikistan	CW	S	F
Shahdara	271 AV	I *	1b	Shahdara River	Tadjikistan	CW	S	F

**Tableau B9 : résultats des croisements réciproques de différentes accessions avec Mr-0**

Les noms et N° dans la collection de Versailles des accessions utilisées en croisements avec Mr-0 sont indiqués dans les deux premières colonnes. La colonne 3 indique le chlorotype de l'accession (les petites croix indiquent les accessions dont j'ai défini les chlorotypes ; les autres chlorotypes ont été définis par Moison et al (2010). La colonne 4 précise le cluster nucléaire (d'après Simon et al, 2011) auquel appartient l'accession. Les colonnes 5 et 6 précisent leur origine géographique. La colonne 7 indique l'auteur du croisement : Nicolas Gobron, (NG) ou Cezary Waszczak (CW). Les colonnes 8 et 9 indiquent le phénotype de la F1 produite avec Mr-0 en utilisant celle-ci comme père ou comme mère, respectivement. F= fertile ; S= stérile.

individus sont stériles et 138 sont fertiles, ce qui veut dire qu'environ  $\frac{3}{4}$  des individus sont fertiles et  $\frac{1}{4}$  sont stériles ( $\chi^2(3:1)=0,149$ ), ce qui semble indiquer que deux locus indépendants sont nécessaires au maintien du phénotype de stérilité, comme c'est le cas pour la stérilité Sha/Mr-0. De plus, Stéphane Boivin (stage M1) a génotypé ces populations avec 12 marqueurs MSAT répartis sur les 5 chromosomes (figure B26a) et ses résultats (figure B26b) indiquent que plus de 85% des individus stériles sont hétérozygotes au sud du chromosome 1 (au marqueur ATHGENEA situé à 22,4 Mb) et que plus de 95% des individus stériles sont hétérozygotes au sud du chromosome 3 (au marqueur MSAT3.65 situé à 19,84 Mb). Les locus mainteneurs de stérilité de l'accession 703AV sont donc portés par les mêmes régions génomiques que ceux de l'accession Mr-0, ce qui suggère que les mêmes gènes nucléaires pourraient être impliqués dans les stérilités Sha/Mr-0 et Sha/703AV.

### III.2 Recherche d'accessions dont le cytoplasme est inducteur d'une stérilité mâle en présence des allèles de Mr-0

Différentes accessions ont été utilisées en croisement avec Mr-0 afin d'identifier celles dont le cytoplasme est inducteur de stérilité en présence des allèles de Mr-0. Le choix des accessions à croiser a été motivé par deux critères : l'origine géographique des accessions et la proximité génétique de leur cytotype, lorsqu'il était connu, avec celui de Sha. L'accession Sha provient d'Asie centrale (Tadjikistan) et possède le cytotype noté I (Moison et al, 2010). Les accessions porteuses d'un cytotype I ou J (le plus proche de I dans le réseau phylogénétique des cytotypes d'Arabidopsis, cf figure A19 de l'introduction) dans le panel étudié par Moison et al (2010), et toutes les accessions originaires d'Asie centrale disponibles dans la collection de Versailles (20 accessions), ont été utilisées dans des croisements réciproques avec Mr-0. Les résultats sont présentés dans le tableau B9. Certains de ces croisements ont été réalisés par Cezary Waszczak avant le début de cette thèse. Les polymorphismes chloroplastiques étant plus rapides et faciles à génotyper que les polymorphismes mitochondriaux, et beaucoup plus informatifs, nous avons assimilé le cytotype des accessions à leur chlorotype en première analyse. Comme nous ne connaissons pas le cytotype de certaines de ces accessions, j'ai établi leur chlorotype (voir §IV.1), qui est également indiqué dans le tableau B9.

Douze accessions d'Asie centrale ont un cytoplasme inducteur de stérilité mâle et ceux de 9 accessions de cette même zone géographique n'induisent pas de stérilité mâle. Trois accessions (72AV, 132AV et 266AV) non originaires d'Asie centrale et porteuses du



chlorotype J n'induisent pas de stérilité mâle dans les populations F1. Ces résultats montrent que la faculté d'induire une stérilité mâle en présence d'allèles de Mr-0 est partagée par le cytoplasme d'un certain nombre d'accessions d'Asie centrale.

#### IV Recherche du compartiment cytoplasmique responsable de l'induction de stérilité

Le second objectif principal de cette thèse a été d'identifier le déterminant inducteur de stérilité présent dans le cytoplasme des accessions inductrices de stérilité (IDS). Au début, nous ne savions pas si le compartiment IDS était chloroplastique ou mitochondrial. Nous sommes donc partis à la recherche d'un éventuel polymorphisme génétique mitochondrial et/ou chloroplastique spécifique des cytoplasmes IDS. Il a fallu pour cela analyser la diversité génétique des cytoplasmes des accessions IDS et non inductrices de stérilité (NIDS) identifiées lors des croisements réciproques avec Mr-0.

##### IV.1 Analyse du chlorotype des accessions utilisées en croisement avec Mr-0

Les quatre régions intergéniques chloroplastiques (tableau A1 de l'introduction) utilisées pour établir la phylogénie des cytoplasmes d'*Arabidopsis thaliana* (Moison et al, 2010) ont donc été séquencées sur les accessions utilisées en croisement avec Mr-0, sauf lorsque cette information était déjà disponible. Les résultats sont présentés dans le tableau B10. Les résultats de ce génotypage m'ont permis d'assigner un chlorotype à chaque accession (reporté dans le tableau B9). Tous les cytoplasmes IDS possèdent un chlorotype I : elles portent des séquences très similaires aux quatre régions intergéniques examinées, à la taille des MSAT près. Il existe aussi dans ce groupe chloroplastique I, une accession au cytoplasme NIDS (403AV), également originaire d'Asie centrale. Ainsi, le chlorotype I est partagé par tous les cytoplasmes IDS, mais n'est pas spécifique de ces cytoplasmes. À cette étape, aucun marqueur spécifique des cytoplasmes IDS n'est donc disponible.

##### IV.2 Identification de marqueurs mitochondriaux spécifiques des cytoplasmes IDS

J'ai utilisé les marqueurs mitochondriaux développés pour l'analyse phylogénétique des cytoplasmes d'*Arabidopsis* (Moison et al, 2010) afin de rechercher s'il était possible de distinguer de façon univoque les cytoplasmes IDS des cytoplasmes NIDS. J'ai analysé le profil de 10 accessions IDS et 5 accessions NIDS, dont 403AV, avec 15 marqueurs

AV code	Chlorotype	nad5-rnm26 Recl	nad5-rnm26 Recl1	nad9/ rpl16	cox2	atp8- orf107c	atp6-2	atp9- cox3	CcmC	orf315	orf131	orf26S2	orf111b	orf 240a	orf 120	orf107d
72AV	J	1	0	1	2	1	1	1	2	SS	SS	1	1	1	1	1
132AV	J	1	0	1	2	1	0	1	2	nd	0	0	0	1	1	1
186AV	A	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
183AV	AH	1	1	1	2	1	1	0	2	1	1	1	1	1	1	1
403AV	I	1	1	1	2	1	0	1	2	1	1	0	0	1	0	0
236AV	I	1	1	1	2	1	0	1	2	1	1	0	0	0	0	0
271AV	I	1	1	1	2	1	0	1	2	1	1	0	0	0	1	0
532AV	I	1	1	1	2	1	0	1	2	1	1	0	0	0	1	0
533AV	I	1	1	1	2	1	0	1	2	1	1	0	0	0	nd	0
534AV	I	1	1	1	2	1	0	1	2	1	1	0	0	0	nd	0
535AV	I	1	1	1	2	1	0	1	2	1	1	0	0	0	1	0
536AV	I	1	1	1	2	1	0	1	2	1	1	0	0	0	nd	0
538AV	I	1	1	1	2	1	0	1	2	1	1	0	0	0	nd	0
539AV	I	1	nd	1	2	1	0	1	2	1	1	0	SS	0	1	0
261AV	I	1	1	1	2	1	0	1	2	1	1	0	SS	0	0	0

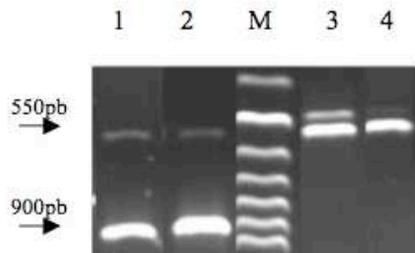
**Tableau B11 : résultats du polymorphisme observé dans l'ADN mitochondrial des accessions d'*A. thaliana***

Les 15 marqueurs du tableau 2 (cf introduction) ont été utilisés pour observer le polymorphisme mitochondrial de 15 accessions IDS (en rose) ou NIDS (en bleu clair). SS : sublimons ; nd : non déterminé ; 0 = pas d'amplification ; 1 = amplification profil 1 ; 2 = amplification profil 2 (les profils sont disponibles dans Moison et al. 2010).

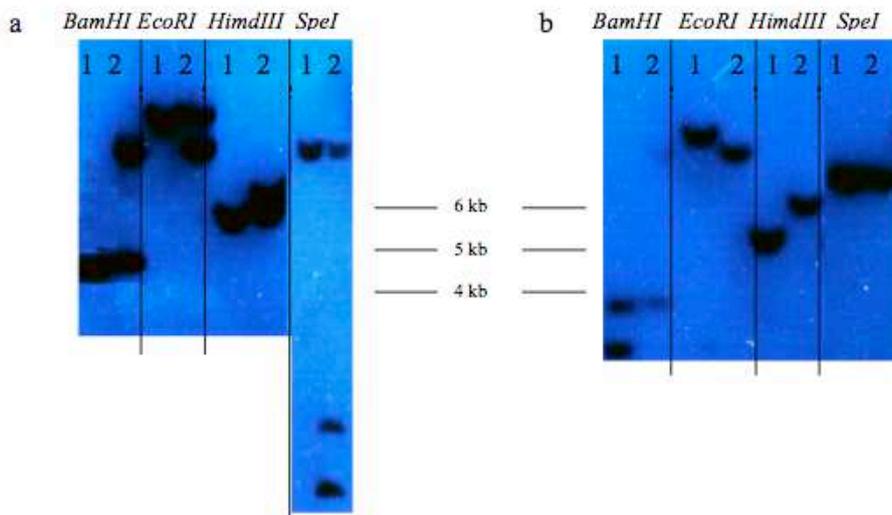


Combinaisons	Références (C24 et Col-0)	NIDS	IDS
L+U	Bande à 830pb	Bande à 830pb	Bande à 830pb
L'+U'	Bande à 901pb	Bande à 901pb	Bande à 901pb
L+U'	Bande à environ 900pb	Bande à environ 900pb	Pas d'amplification
L'+U	Bande majoritaire à 900pb	Bande majoritaire à 900pb	Bande majoritaire à 550pb

**Tableau B12 : résumé des amplifications par PCR obtenus en utilisant les amorces *trnK***  
 Résultats des amplifications par PCR en utilisant des combinaisons d'amorces utilisées pour amplifier les deux copies *trnK* à partir d'ADN d'accessions inductrices de stérilité (IDS) et non inductrices de stérilité (NIDS). La position des amorces sur le génome mitochondrial de C24 est indiquée dans la figure B27.



**Figure B28 : marqueur « TrnK » spécifique des accessions IDS**  
 Photo de gel 1% obtenu en utilisant les amorces L'+U. 1 : Col-0, 2 : 403AV, 3 : 534AV et 4 : Sha. Les accessions 1 et 2 sont NIDS et les accessions 3 et 4 sont IDS. M : échelle (50 pb DNA ladder Fermentas)



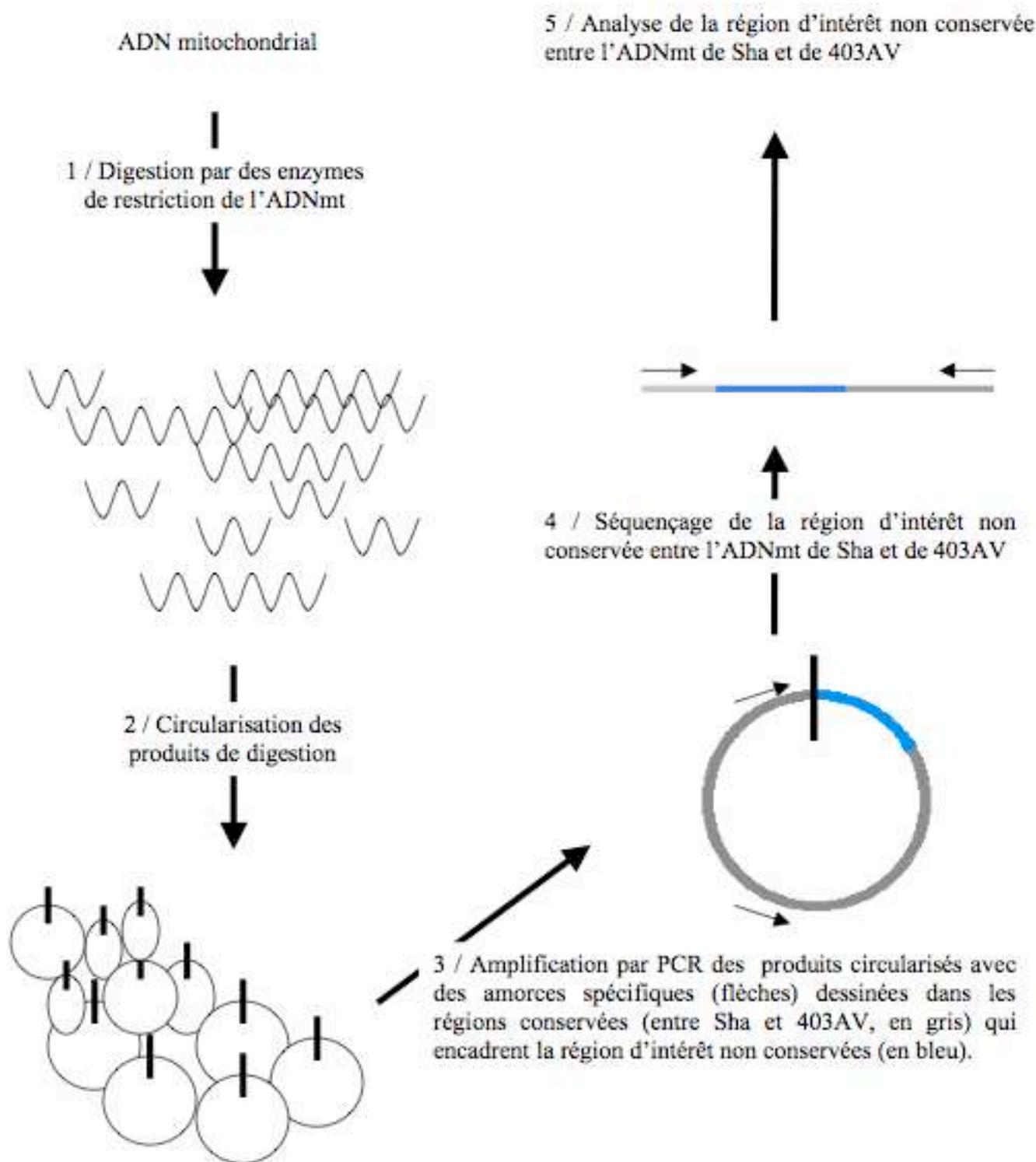
II

enzyme	Sonde <i>trnK</i>	Sonde <i>orf122c</i>
<i>Bam</i> HI	<b>12 kb*</b> + 3,4 kb	<b>12 kb*</b>
<i>Eco</i> RI	<b>8,7 kb*</b> + 13,3 kb	<b>8,7 kb*</b>
<i>Hind</i> III	<b>5,9 kb*</b> + 5 kb	<b>5,9 kb*</b>
<i>Spe</i> I	<b>1,2 kb + 1,5 kb + 6,5 kb + 7,1 kb</b>	7,2 kb

**Figure B29 : southern blot obtenus avec des sondes mitochondriales**  
 I/ L'ADN de Sha (1) et de 403AV (2) a été digéré par des enzymes de restriction dont les noms sont indiqués sur la figure. Les fragments d'ADN digéré ont été séparés par électrophorèse en gel d'agarose, transférés sur membrane et hybridés avec une sonde *trnK* (a) et une sonde *orf122c* (b). Ces southern ont été réalisés par F. Budar.  
 II/ Tableau des tailles attendues en hybridation (figure 29) d'après la séquence de référence. Les fragments notés d'un \* sont les mêmes pour les deux sondes. Les fragments en **gras** sont ceux qui ne sont pas retrouvés ou ne sont pas à la taille attendue chez Sha.

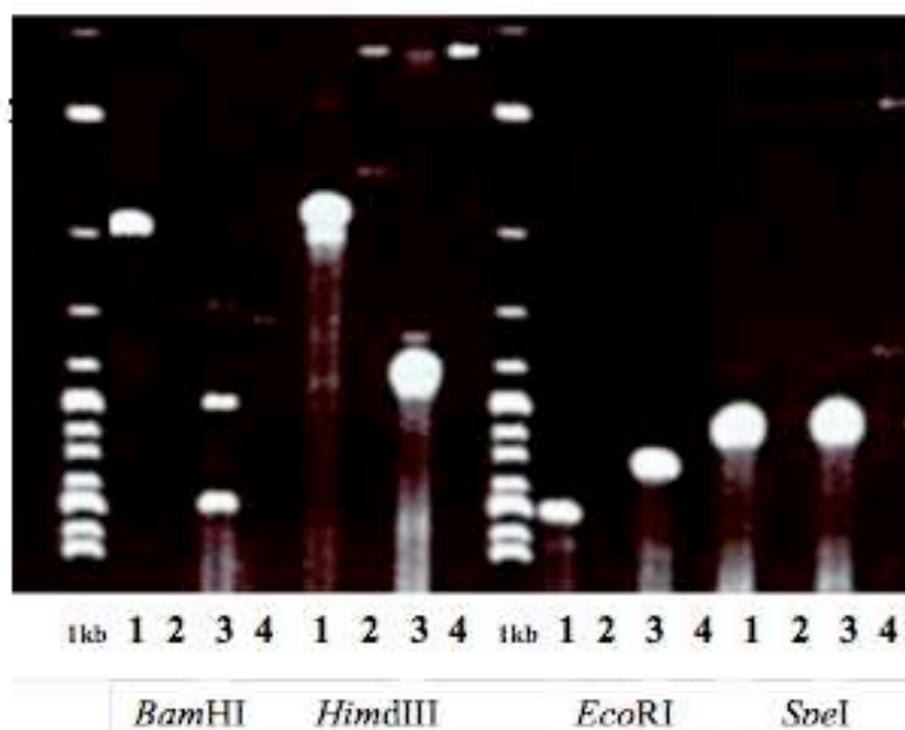
mitochondriaux (tableau A2 de l'introduction). Les résultats sont présentés dans le tableau B11. D'après l'analyse de diversité cytoplasmique réalisée dans l'équipe, le cytotype I, porté par Sha se distingue du cytotype le plus proche (J) par deux polymorphismes : un polymorphisme chloroplastique et un polymorphisme mitochondrial (Moison et al, 2010). Ce dernier correspond à un réarrangement qui a pour effet de supprimer, dans Sha, trois gènes non conservés et adjacents sur le génome mitochondrial de référence : *orf240a*, *orf120* et *orf107d* (Moison et al, 2010). Les 11 accessions porteuses du chlorotype I analysées, dont 403AV, ont toutes les mêmes profils que Sha sauf pour deux de ces marqueurs : *orf240a* et *orf120*. Pour l'*orf120* le profil n'est pas spécifique du caractère IDS du cytoplasme examiné. En revanche, le marqueur *orf240a*, permet de différencier le cytoplasme de 403AV des cytoplasmes IDS. Les 4 autres cytoplasmes NIDS, qui n'appartiennent pas au groupe chloroplastique I (72AV, 132AV, 183AV et 186AV), sont distinguables des cytoplasmes IDS par plusieurs marqueurs, y compris *orf240a*, ce qui était attendu. Le résultat obtenu avec le marqueur *orf240a* est spécifique du caractère IDS ou NIDS du cytoplasme : il n'est jamais amplifié avec l'ADN des accessions IDS ; on observe toujours une amplification de ce marqueur avec l'ADN des accessions NIDS dont celui de l'accession 403AV (95 accessions ont été testées pour ce marqueur par Moison et al., 2010). Ces résultats suggèrent que le déterminant cytoplasmique de la stérilité est mitochondrial et nous procurent un marqueur moléculaire pour tester la capacité des cytoplasmes à induire la stérilité. Toutefois, il s'agit d'un marqueur « négatif » où c'est l'absence d'amplification qui indique la nature IDS du cytoplasme, ce qui est parfois source d'ambiguïtés dans l'interprétation des résultats expérimentaux.

Un autre marqueur mitochondrial, spécifique des accessions IDS, a été identifié par la stratégie suivante : un gène *trnK*, situé à proximité de l'*orf240a* sur le génome mitochondrial C24 d'*A. thaliana*, est dupliqué dans le génome mitochondrial : une autre copie se trouve dans la région *rpl16/nad9/rps3* chez C24 (figure B27). D'après mes résultats (tableau B12), en utilisant des amorces spécifiques de chaque copie du gène *trnK*, la perte de l'*orf240a* chez les accessions IDS ne s'accompagne pas de la perte de la copie du gène *trnK* située à proximité de l'*orf240a* : la copie du gène *trnK* associée à l'*orf240a* est étonnamment amplifiée chez Sha, alors qu'aucun des deux gènes encadrant cette copie (*orf240a* et *orf120*) n'est présent chez Sha. Par ailleurs, des expériences d'hybridation utilisant une sonde *trnK* montrent que deux copies sont présentes dans 403AV alors qu'une seule l'est dans Sha (figure B29). Les tailles des fragments hybridés indiquent que Sha possède la copie de la région *rps3-ccb206*, alors que 403AV porte les deux copies du génome de référence. Il est possible que la copie



**Figure B30 : stratégie utilisée pour analyser la région d'ADNmt d'intérêt, non conservée entre Sha et 403AV**

Cette stratégie a été réalisée à partir d'ADN de Sha et de 403AV.



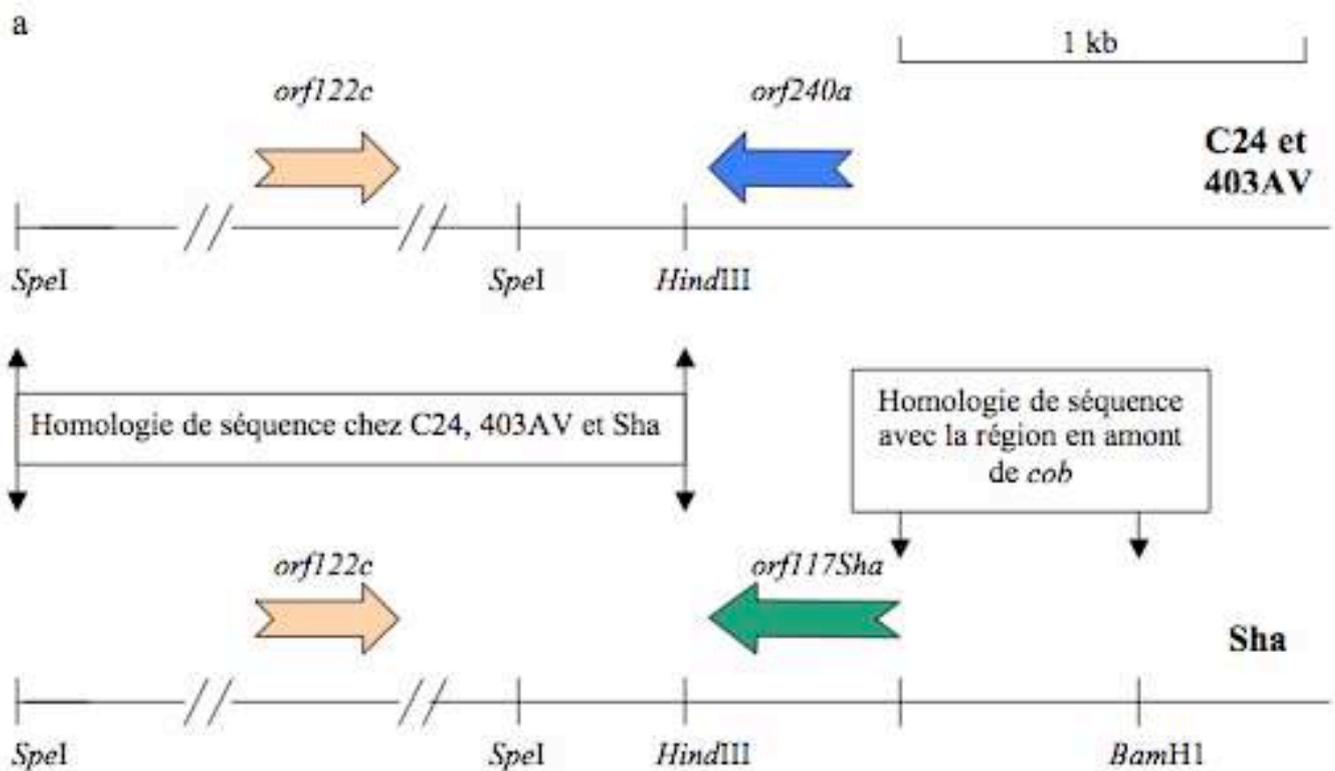
**Figure B31 : gel d'électrophorèse de produits PCR mitochondriaux amplifiés et digérés à partir d'ADNt de Sha et 403AV**

L'ADN total de Sha et 403AV a été digéré avec les enzymes indiquées puis les fragments de digestion ont été circularisés et amplifiés avec des amorces mitochondriales en amont de la région génomique qui diffère entre Sha et 403AV. 1 : Sha L+, 2 : Sha L-, 3 : 403AV L+, 4 : 403AV L-. L+ = ligation, L- = pas de ligation. Cette expérience a été réalisée par F. Budar.

<i>Bam</i> HI	Sha	bande à 1,5 kb
	403AV	2 bandes à 7 kb et 3,5 kb
<i>Hind</i> III	Sha	Bande à 1,3 kb
	403AV	Bande à 2,7 kb
<i>Eco</i> RI	Sha	Bande à 6 kb
	403AV	Bande à 4 kb
<i>Spe</i> I	Sha	Bande à 2,7 kb
	403AV	

**Tableau B13: résumé des résultats de la figure B31**

Les profils moléculaires mitochondriaux obtenus diffèrent entre Sha et 403AV avec les enzymes *Bam*HI, *Hind*III et *Eco*RI. et sont identiques avec l'enzyme *Spe*I.



b

ORF117Sha	12	TINSLSQNIKELQNNR---ISSDQPSNLVYETRAANAVTKTIEHLFHRHNLTLPPRWIHEQ
ORF108	3	TI SLSQ+IE+L+ N SDQ N + A+T T++ L HN+ LP WTH Q
ORF117Sha	69	ILQUILGDULTLTNVTSVLKNIITQLGLASDEFHQVLNVLHLITGGG
ORF108	63	I+ODILGDD+TL NV SVL NITQLGLASDEFHO+ N+++LTGGG
		IVQUILGDUMTLGNVASVLSNIITQLGLASDEFHQIFNIINLITGGG

**Figure B32 : structure de la région génomique de l'orf117Sha.**

a- Schéma représentatif de la région mitochondriale d'intérêt chez C24, 403AV et Sha.

b- Alignement des séquences en acides aminés de ORF117Sha et ORF108. La séquence ORF117Sha a été blastée sur NCBI : 58% de cette séquence est identique à celle d'ORF108 (*Brassica juncea*) et ces deux ORF ont 71% de similarité.

*trnK* soit à l'état de sublimon dans les mitochondries de Sha ce qui expliquerait qu'elle soit détectable en amplification par PCR mais pas en southern. Les quatre amorces qui amplifient les deux copies du gène *trnK* ont ensuite été testées en associant une amorce spécifique d'une copie avec une amorce spécifique de l'autre copie : une de ces combinaisons donne un produit d'amplification majeur de taille spécifique avec les cytoplastes IDS (tableau B12 et figure B28). Cette combinaison d'amorces nous procure un marqueur mitochondrial, positif et probablement spécifique des accessions IDS. Il n'a pas été exploité plus avant, au vu des résultats décrits ci-dessous (§IV.3).

Il est clair à cette étape que la région génomique contenant les *orf240a*, *orf120* et *orf107d* diffère entre les accessions NIDS et IDS, certainement à cause d'un réarrangement spécifique des accessions IDS. De plus, c'est la seule région mitochondriale pour laquelle nous avons détecté une différence entre 403AV (NIDS) et les génomes IDS, ce qui en fait une région candidate pour porter le gène responsable de la stérilité mâle.

#### IV.3 Recherche d'un gène mitochondrial candidat facteur d'IDS

Afin d'identifier les différences génomiques qui existent dans cette région d'intérêt (figure B27b) entre l'ADNmt de 403AV et celui de Sha, des Southern ont été réalisés sur l'ADN de ces deux accessions par Françoise Budar en utilisant des sondes spécifiques aux gènes de la région d'intérêt et plusieurs enzymes de restriction. Ils ont permis d'identifier des enzymes de restriction capables de révéler du polymorphisme de type RFLP, dans la région d'intérêt, entre l'ADNmt de 403AV et celui de Sha (exemple de southern dans la figure B29). Ces résultats mettent en évidence du polymorphisme mitochondrial entre l'ADNmt de Sha (IDS) et de 403AV (NIDS) dans la région de l'*orf240a*.

Cette première analyse a permis de choisir les enzymes utilisés pour une PCR inverse. L'ADN des accessions Sha et 403AV a été digéré par ces enzymes puis circularisé. Les produits circularisés ont ensuite été amplifiés avec des amorces divergentes dessinées dans la région d'ADNmt commune aux deux accessions, en amont du point de divergence (figure B30). Une partie des résultats observés après migration (figure B31) est résumée dans le tableau B13. Les produits d'amplification ont ensuite été séquencés, ce qui a permis d'identifier la séquence de la région en aval de l'*orf122c* chez Sha (figure B32a). Sur cette séquence, une séquence ouverte de lecture de 117 codons, nommée *orf117Sha*, a été

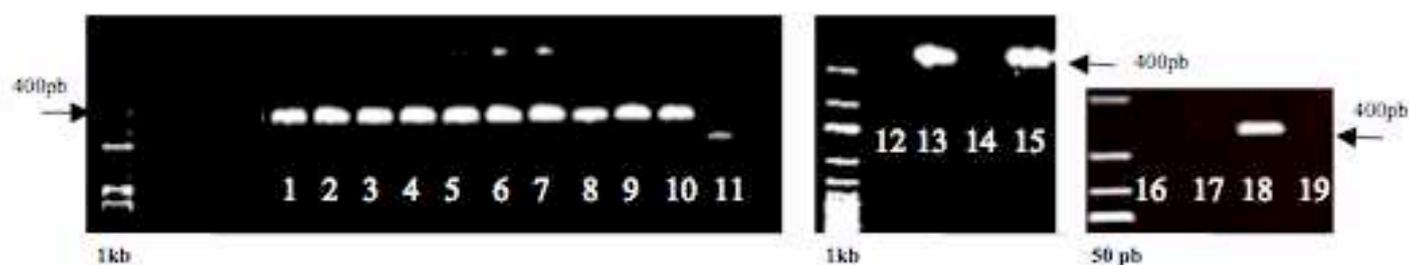


Figure B33 : gel d'électrophorèse de produits PCR de l'*orf117Sha* à partir de l'ADNt des accessions d'*Arabidopsis thaliana*.

La légende des numéros est indiquée dans le tableau B14. Tous les produits d'amplification ont une taille d'environ 400pb.

N°	Accession	Amplification
1	271AV	OK
2	531AV	OK
3	532AV	OK
4	533AV	OK
5	534AV	OK
6	535AV	OK
7	536AV	OK
8	537AV	OK
9	538AV	OK
10	539AV	OK
11	540AV	Faible amplification
12	403AV	Pas d'amplification
13	271AV	OK
14	403AV	Pas d'amplification
15	261AV	OK
16	265AV	Pas d'amplification
17	270AV	Pas d'amplification
18	Sha	OK
19	H2O	Pas d'amplification

Tableau B14 : résumé des résultats de la figure B33.

L'*orf117Sha* est spécifiquement amplifié chez les accessions IDS (en rose). En bleu, les accessions NIDS.

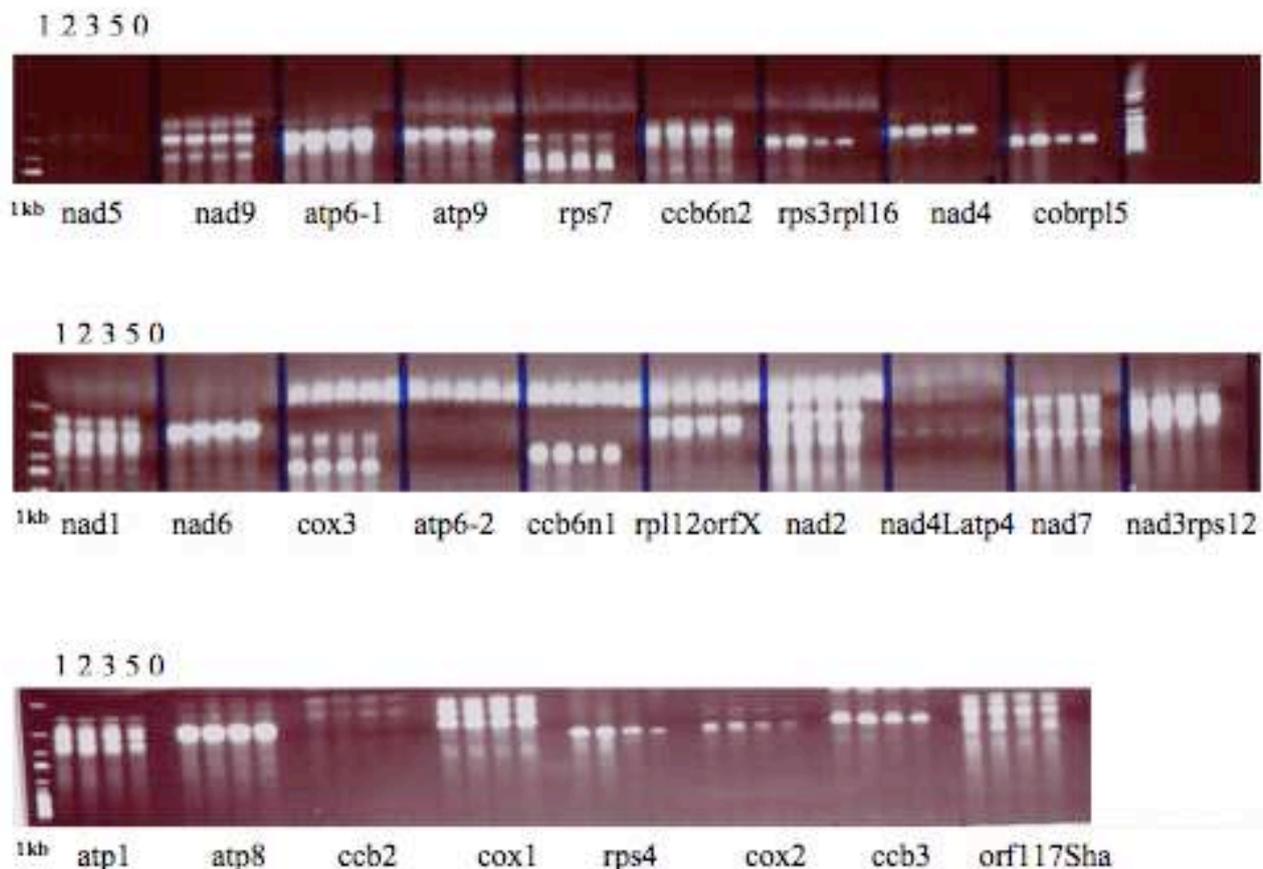
identifiée. Elle ne présente aucune similarité significative avec des séquences d'ADN connues dans les bases de données (Genbank). Cependant, la recherche par BLASTP d'homologues du produit prédit ORF117Sha dans les bases de données protéiques a révélé que la séquence en acides aminés du produit putatif de ce gène présente 58% d'identité, et 71% de similarité avec celle de l'ORF108 décrite par Ashutosh et al., (2008) comme une protéine candidate pour induire la stérilité mâle chez des plantes de *Brassica juncea* porteuses du cytoplasme de *Moricandia arvensis* (figure B32b).

Pour savoir si l'*orf117Sha* est spécifique ou non des accessions IDS, des amorces spécifiques de l'*orf117Sha* ont été utilisées pour amplifier de l'ADN provenant d'accessions IDS et NIDS : les résultats sont présentés dans la figure B33. Le tableau B14 résume les résultats de la figure B33. L'*orf117Sha* est amplifié spécifiquement avec l'ADN des accessions IDS. Dans certains cas, nous avons obtenus des amplifications de faible intensité avec des ADN d'accessions dont le cytoplasme est NIDS, probablement dues à la présence de l'*orf117Sha* sur des sublimons. Ce dernier résultat reste à consolider.

#### IV.4 Influence des allèles mainteneurs Mr-0 sur l'expression des gènes mitochondriaux de Sha

L'*orf117-Sha* est un bon candidat facteur de stérilité mitochondrial, mais il n'est peut-être pas le seul : en effet, les marqueurs mitochondriaux testés ne couvrent pas tout le génome mitochondrial et nous ne pouvons pas exclure qu'une région d'ADN mitochondrial non couverte par notre analyse porte le facteur de stérilité.

Afin d'identifier sans *a priori* d'éventuels autres candidats mitochondriaux à l'IDS des analyses de l'expression des gènes mitochondriaux de Sha, dans un contexte nucléaire mainteneur ou restaurateur, ont été effectuées. La maturation des extrémités 5' et 3' des ARNm mitochondriaux est une étape post-transcriptionnelle importante pour l'expression des gènes mitochondriaux (voir le chapitre II.3 de l'introduction). Nous avons à disposition au laboratoire, un outil de phénotypage moléculaire composé de couples d'amorces, mises au point par Marie Verzaux, qui permet d'observer et d'étudier l'expression des gènes chloroplastiques et mitochondriaux chez *A. thaliana*. Des profils moléculaires comme l'épissage des introns, l'édition et la maturation des extrémités 5' et 3' des transcrits peuvent être observés avec cet outil. Sachant que les allèles mainteneurs et restaurateurs peuvent influencer l'expression du locus mitochondrial de stérilité en affectant le profil d'expression



**Figure B34 : gels d'électrophorèse de produits cRT-PCR.**

Les couples d'amorces dessinés par Marie Verzaux, permettant d'amplifier les extrémités de 26 transcrits mitochondriaux ont été utilisés sur de l'ADN provenant de 4 types d'individus issus d'une population RM4. Les gènes ciblés dans chaque amplification sont indiqués en dessous des gels. Les génotypes des individus sont ceux de type 1 (Sha au sud des chromosomes 1 et 3), type 2 (He au sud des chromosomes 1 et 3), type 3 (He au sud du chromosome 1 et Sha au sud du chromosome 3) et type 5 (Sha au sud du chromosome 1 et He au sud du chromosome 3) définis dans le chapitre II (indiqués au dessus des gels). Pour chaque profil obtenu, le puits 0 contient le contrôle H2O.

des ARNm mitochondriaux (Singh et al., 1996) et que les gènes inducteurs de stérilité sont souvent co-transcrits avec des gènes mitochondriaux conservés de fonction connue, nous avons décidé de réaliser une exploration de la taille des extrémités 5' et 3' des transcrits mitochondriaux de Sha, en présence/absence des allèles Mr-0 mainteneurs de stérilité et des allèles Sha restaurateurs de fertilité pour les gènes de fonction connue. Une population RM4 (92 individus) a été phénotypée et génotypée aux marqueurs MSAT1.21231 et ATHFUS6 situés au sud des chromosomes 1 et 3. Cette population a les mêmes caractéristiques que la population RM4 analysée pour la caractérisation phénotypique dans le chapitre II. Les individus ont été classés selon les types 1, 2, 3 et 5 définis au chapitre II. Les inflorescences d'une dizaine d'individus appartenant au même type ont été poolées (pour chacun des quatre types), et les ARNs totaux ont été extraits puis analysés en cRT-PCR. J'ai utilisé 27 couples d'amorces (cf Matériel et Méthodes) pour amplifier les extrémités des ARNm codés par le génome mitochondrial. J'ai également inclus dans cette analyse un couple d'amorces permettant d'amplifier les extrémités de l'ARNm *orf117Sha*. Les profils d'amplification obtenus sont présentés dans la figure B34 : pour un couple d'amorces donné, quel que soit le génotype analysé, les profils d'amplification sont identiques d'un type à l'autre. Néanmoins, le profil d'amplification de type 1, obtenu avec le couple d'amorces qui amplifie les extrémités de l'ARNm *ccb6n2*, semble être différent des profils obtenus avec les génotypes de type 2, 3 et 5. Ce dernier résultat devra être confirmé.

# **DISCUSSION**

J'ai organisé cette discussion en réalisant, dans une première partie, le bilan critique de mes résultats. Dans une seconde partie, j'essayerai, en me servant des résultats obtenus au cours de cette thèse et dans l'équipe, et en m'appuyant sur des connaissances établies dans d'autres systèmes d'hybrides stériles, d'émettre des hypothèses sur la signification évolutive de la SMNC Sha x Mr-0 chez *Arabidopsis thaliana* ; je terminerai par un bilan des barrières génétiques mises en évidence dans le croisement Sha x Mr-0.

## **Bilan critique des résultats**

### **Le déterminant cytoplasmique de stérilité mâle**

Les phénotypes de stérilité mâle observés dans les F1 des croisements réciproques réalisés entre l'accession Mr-0 et d'autres accessions ont montré que douze accessions d'Asie centrale possèdent un cytoplasme inducteur de stérilité mâle (IDS). Dans la limite des accessions testées, l'Asie centrale est la seule région géographique où le cytoplasme IDS a été détecté.

L'utilisation des marqueurs chloroplastiques et mitochondriaux, développés dans des travaux précédents, sur les cytoplasmes IDS et non inducteur de stérilité (NIDS) identifiés lors des croisements réciproques avec Mr-0 a permis d'identifier, dans un premier temps, le compartiment cytoplasmique probablement impliqué dans l'induction de stérilité. Lors de cette analyse, l'objectif était de placer les accessions IDS et NIDS dans le réseau de cytotypes de Moison et al. (2010). Lorsque les cytoplasmes IDS et NIDS ont été caractérisés avec les marqueurs chloroplastiques, il est apparu que tous les cytoplasmes IDS portaient le même chlorotype (I) et qu'aucun autre chlorotype n'était associé à l'induction de stérilité. Toutefois, le cas de l'accession Kz-9 (403AV) était particulier puisque possédant un chlorotype I, son cytoplasme est NIDS. Cette accession, également originaire d'Asie Centrale, a été très utile pour aller plus loin dans l'identification du déterminant cytoplasmique de la stérilité. Les résultats de chlorotypage m'ont conduit à analyser les accessions utilisées en croisement pour les marqueurs mitochondriaux dont nous disposons dans l'équipe. Un premier polymorphisme mitochondrial spécifique des cytoplasmes IDS a ainsi été identifié (absence de l'*orf240a*). Ce sont ces résultats qui nous ont dirigés vers le compartiment mitochondrial

comme source de facteur stérilisant chez Sha. La stratégie adoptée a consisté à rechercher les différences entre les génomes mitochondriaux de Sha et Kz-9 en faisant l'hypothèse que ces différences sont peu nombreuses et qu'elles couvrent le déterminant cytoplasmique de la stérilité mâle. La région de l'*orf240a* a été explorée et le gène *orf117Sha*, dont la présence est corrélée à l'induction de stérilité mâle, a été mis en évidence.

Depuis, des analyses conduites dans l'équipe par F. Budar ont confirmé que les génomes cytoplasmiques de Sha et Kz-9 sont très proches. Sur la base des données de reséquençage disponibles pour Sha et Kz-9 (1001genomes.org), toutes les régions où des variations non synonymes ont été prédites dans des gènes mitochondriaux ou chloroplastiques ont été reséquencées dans les deux accessions et les deux génotypes ont présenté les mêmes variations de séquence. Le gène mitochondrial prédit *orf117Sha* est le seul candidat identifié à ce jour comme pouvant être le facteur cytoplasmique IDS. Plusieurs résultats expérimentaux permettent de soutenir cette hypothèse : l'*orf117Sha* est spécifique des cytoplasmes IDS ; il semble qu'il y ait très peu d'autres différences entre les génomes cytoplasmiques de Sha et Kz-9 ; de plus, la séquence en acides aminés du produit putatif de ce gène présente 57% d'identité et 71% de similarité avec celle de l'ORF108 décrite par Ashutosh et al., (2008) comme étant impliquée dans la SMNC de *Brassica juncea* portant le cytoplasme de *Moricandia arvensis*. Ces auteurs ont remarqué que les fleurs des individus mâle stériles avaient des ARNm *atpA* plus longs que ceux des individus mâles fertiles. Chez les premiers, le gène non conservé *orf108* est co-transcrit avec le gène *atpA*. Chez les individus fertiles restaurés, le transcrit *atpA-orf108* est clivé dans la séquence codante de l'*orf108*, ce qui aboutit à la production d'ARNm *atpA* monocistroniques (Ashutosh et al., 2008).

Toutefois, d'après les résultats que j'ai obtenus par cRT-PCR, les tailles des ARN portant l'*orf117Sha* ne semblent pas varier selon le contexte mainteneur ou restaurateur du génome nucléaire (confirmé par F. Budar en northern). Étant donné que nous n'avons pas observé de différences entre les profils d'ARN mitochondriaux de plantes fertiles et stériles, le contrôle de l'expression du gène *orf117Sha* dans la SMNC Sha x Mr-0, s'il existe, est différent de celui de l'*orf108* dans la SMNC décrite par Ashutosh et al. (2008). De plus, Kumar et al (2012) ont exprimé l'*orf108* dans des plantes d'*A. thaliana* transformées en utilisant un promoteur anthère spécifique, celui du gène *Atprx18*. L'expression de l'*orf108* chez les plantes transformées aboutit à la mort de 50% des grains de pollen. Ces résultats indiquent que la protéine ORF108 est létale au stade gamétophytique. Il semblerait que l'*orf117Sha*

agisse plutôt à un stade sporophytique étant donné que son expression induit la mort de 100 % des grains de pollen dans la F1 Sha x Mr-0.

En parallèle à l'exploration de la région de l'*orf117Sha* dans le génome de Sha, j'ai entrepris une recherche basée sur l'hypothèse que le contexte génétique, et la présence des allèles mainteneurs de Mr-0 en particulier, pourrait avoir une incidence sur la maturation de l'ARNm d'une séquence mitochondriale liée à la SMNC. Cette recherche a été réalisée sans a priori sur la région génomique pouvant porter le déterminant de stérilité et a consisté à explorer systématiquement les profils en cRT-PCR des gènes mitochondriaux conservés. J'ai mis en évidence une différence de taille des transcrits *ccb6n2* selon le contexte nucléaire. Si cette région est porteuse du facteur de stérilité, on s'attend qu'elle diffère entre les accessions Sha et Kz-9. Des Southern blot sur des ADN de Sha et Kz-9, réalisés par F. Budar, ont montré que les deux accessions ont la même organisation génomique de cette région mitochondriale. De plus le séquençage des produits de cRT-PCR a révélé que Sha possède une séquence identique à la référence dans les régions transcrites non traduites du gène et que les ARNm plus longs détectés ont une 5' UTR étendue ne portant aucune orf candidate. Les différences de profil en cRT-PCR ne sont probablement pas liées à l'IDS, mais révèlent sans doute l'implication d'au moins un gène porté par la région du ch1 sur la maturation de ce gène mitochondrial. Il est intéressant de remarquer que cette région comprend le cluster de gènes codant des PPR *Rf*-like dont certains ont été impliqués dans la maturation de transcrits mitochondriaux (Jonietz et al, 2010 ; Jonietz et al, 2011).

L'un des objectifs de la thèse, qui était d'identifier un ou plusieurs candidats cytoplasmiques impliqués dans le phénotype de stérilité mâle, a été rempli. L'*orf117Sha* mitochondrial est le seul candidat IDS sérieux identifié. Cependant, aucune influence des allèles mainteneurs ou restaurateurs sur l'expression de l'*orf117Sha* n'a été pour le moment mise en évidence : d'après mes résultats et ceux obtenus par d'autres membres de l'équipe, le profil d'expression de l'*orf117Sha* est complexe (avec de nombreuses espèces moléculaires de tailles différentes) et semble identique dans un fond mainteneur et restaurateur. Selon toute vraisemblance, la stratégie adoptée pour identifier des candidats cytoplasmiques IDS a été payante : l'*orf117Sha* mitochondrial est un bon candidat qui reste cependant à confirmer car l'effet du fond génétique (mainteneur ou restaurateur) sur son expression doit encore être prouvé.

Dans les systèmes de SMNC, la différence entre les fonds génétiques mainteneurs et

restaurateurs se manifeste le plus souvent dans l'expression du facteur IDS. La restauration de la fertilité résulte, dans l'immense majorité des cas, d'une diminution de l'expression du gène mitochondrial associé à la stérilité (cf partie VI.4.2 de l'introduction). La majorité des gènes *Rf* identifiés sont des gènes qui codent des protéines capables d'affecter l'expression du locus mitochondrial de stérilité en affectant le profil des ARNm comme c'est le cas dans le système de SMNC de *Brassica juncea* portant le cytoplasme de *Moricandia arvensis* (Ashutosh et al., 2008) ou dans la SMNC *pol* chez *Brassica napus*, où le gène *Rf*, favorise le clivage de l'ARN bicistronique inducteur de stérilité *orf224/atp6* (Li et al., 1998 ; Brown, 1999).

D'après mes résultats, ce type de mécanisme de contrôle post-transcriptionnel de l'expression du facteur IDS ne semble pas être impliqué dans le système Sha x Mr-0. Toutefois, il se peut que le mécanisme de contrôle de l'expression de l'*orf117Sha* se situe bien à un niveau post-transcriptionnel mais que ce mécanisme soit tissu spécifique. L'extraction d'ARN à partir d'inflorescences, comme cela a été le cas pour réaliser les cRT-PCR, ne permettrait pas d'observer d'éventuelles différences d'expression de l'*orf117Sha* entre des plantes fertiles et stériles car le tissu ciblé par l'influence du locus mainteneur serait trop dilué dans ces échantillons. Ces éventuelles différences d'expression pourraient être observées en réalisant des hybridations *in situ* avec une sonde spécifique de l'ARNm *orf117Sha* sur des sections de boutons floraux, ce qui permettrait de déterminer si la quantité de l'ARNm *orf117Sha* dans les tissus reproductifs varie entre génotypes mainteneur et restauré. Une expérience similaire a été réalisée au sein de l'équipe lors de la caractérisation de l'ORF138 responsable de la SMNC dans le système Ogura (Uyttewaal et al, 2008b). Il se peut aussi que le mécanisme de contrôle de l'expression de l'*orf117Sha* agisse sur la traduction de l'ARNm, comme c'est le cas dans le système Ogura du radis (cf partie VI.4.2 de l'introduction) où l'association de PPRB entraîne un blocage de la traduction de l'ARNm *orf138* (Uyttewaal et al., 2008b). Pour tester cette hypothèse, il faudrait obtenir un anticorps anti-ORF117SHA et l'utiliser sur des extraits protéiques de plantes fertiles et stériles pour estimer l'effet du contexte nucléaire sur l'accumulation de la protéine ORF117SHA. Cet effet peut éventuellement être tissu spécifique, auquel cas, il faudrait visualiser, par immunolocalisation, la protéine ORF117SHA dans les tissus des anthères en contexte restaurateur ou mainteneur.

## Les locus nucléaires mainteneurs de stérilité

Au début de la thèse, les résultats de l'équipe indiquaient que deux locus, un situé au sud du chromosome 1 (locus ch1) et un situé au sud du chromosome 3 (locus ch3), étaient hétérozygotes chez la majorité des plantes mâle stériles d'une population RM1 (issue du croisement F1(Sha x Mr-0) x Sha). Mes résultats sur les populations RM3 confirment que la présence d'allèles Mr-0 dans les régions sud de ces deux chromosomes est nécessaire et suffisante pour maintenir la SMNC.

### Le locus mainteneur chr3

Le locus ch3 est encadré, grâce à la cartographie génétique réalisée pendant ce travail, par deux marqueurs, CAPSAt3g62560 et CAPS At3g62700 qui délimitent un intervalle de 53 kb dans la séquence de Col-0.

La stratégie adoptée pour cartographier ce locus s'est révélée efficace et permet maintenant d'envisager une approche « gènes candidats » pour identifier l'allèle Mr-0 mainteneur de stérilité au locus ch3. Chez Col-0, il existe 15 gènes entre *At3g62560* (position physique TAIR10 : 23 137 426 pb – 23 139 045 pb) et *At3g62700* (position physique TAIR10 : 23 190 267 pb – 23 195 762 pb). Parmi les 15 gènes présents chez Col-0, *At3g62620* est le seul dont les logiciels MitoProtII (v1.101) et Predotar prédisent qu'il code une protéine importée par les mitochondries (probabilités de 0,86 et 0,47 respectivement). Le gène *At3g62620* code une protéine similaire à une sucrose-phosphatase qui pourrait être impliquée dans le métabolisme de l'amidon ; la biosynthèse de l'amidon est cruciale durant les phases finales de la maturation du pollen : l'amidon constitue une source de réserve énergétique et sert de « checkpoint » pour la maturation du pollen (Datta et al., 2002). Par exemple, dans la SMNC-S de maïs, de nature gamétophytique, la mort des grains de pollen est associée à une déficience en amidon (Wen et Chase, 1999a). Les gènes *At3g62640* et *At3g62660* sont fortement exprimés dans les grains de pollen (logiciels eFP Browser et Genevestigator). Le gène *At3g62640* n'est pas caractérisé ; *At3g62660* (appelé *GATL7*) code une protéine similaire à une galacturonosyl transférase qui pourrait être impliquée dans la synthèse d'homogalacturonane, un polysaccharide pectique présent dans la paroi des cellules végétales. Ces trois gènes sont des candidats séduisants pour intervenir dans une interaction avec les mitochondries et/ou dans le développement du pollen. Cependant, chacun des 15 gènes

présents au locus ch3 chez Col-0 reste un candidat, s'il est présent à ce locus chez Mr-0, pour le maintien de la stérilité. À noter que deux recombinants stériles dans l'intervalle du locus ch3 sont encore exploitables pour diminuer la taille du locus ch3 et diminuer ainsi le nombre de candidats. Une validation fonctionnelle des gènes candidats pourra être faite par transformation classique pour induire une surexpression de l'allèle Mr-0 (du gène candidat), dans des génotypes fertile fixés Mr-0 au locus ch1 dans un fond Sha (y compris au locus ch3). Si des plantes stériles apparaissent dans la descendance de ces transformants, cela voudra dire que l'allèle Mr-0 testé est bien l'allèle mainteneur du locus ch3. Des contrôles devront être réalisés sur les mêmes génotypes mais avec des vecteurs contenant les allèles Sha ou Col-0 : aucune plante stérile ne devrait apparaître dans la descendance de ces plantes transformées.

D'autres gènes peuvent être présents dans cette région génomique chez Mr-0 mais absents du génome de Col-0. Il est possible qu'un gène supplémentaire mainteneur de stérilité, spécifiquement présent au locus ch3 chez Mr-0 et Etna-2, et favorisant l'expression du déterminant cytoplasmique IDS de Sha, soit présent au locus ch3 chez ces deux accessions (et absent chez Sha). Cependant, aucune insertion de grande taille n'a été détectée chez Mr-0 au locus mainteneur ch3. À cette étape, je ne peux cependant pas exclure totalement qu'un gène soit présent dans cet intervalle chez Mr-0 et Etna-2 et absent des accessions non mainteneuses. D'autres variations structurales, comme des inversions et des petites insertions/délétions, peuvent également exister entre Sha et Mr-0 au locus ch3.

Pour identifier le ou les gènes candidats à tester par transgénése, il faudrait analyser les polymorphismes génétiques entre Sha et Mr-0 au locus ch3, et déterminer ceux dont les allèles Mr-0 sont partagés par les autres accessions mainteneuses identifiées et si possible spécifiques du caractère mainteneur. Des fragments de régions codantes et non codantes de gènes ont été séquencées au locus ch3 (lors des analyses CAPS et lors des amplifications de longs fragments). Des polymorphismes non synonymes de type SNP ou de type « petites insertion/délétion » dans des séquences codantes ont été détectés entre Sha et Mr-0 au locus ch3 ; sans validation fonctionnelle, ces polymorphismes ne peuvent être liés avec certitude à la capacité de maintenir la stérilité. De plus, seule une petite partie de la séquence de cette région est disponible chez Mr-0 et il est possible que d'autres polymorphisme non identifiés existent entre Sha et Mr-0 et que l'un d'entre eux soit responsable du maintien de la stérilité à ce locus chez Mr-0. Il convient donc de continuer à explorer les polymorphismes entre les deux parents dans cette région.

Seuls les locus mainteneurs ont été cartographiés lors de ce travail et il n'est pas certain que les allèles Sha restaurateurs et les allèles Mr-0 mainteneurs soient les allèles d'un même gène au locus ch3. Cependant, nous pouvons faire l'hypothèse que les allèles restaurateurs Sha sont également présents au locus ch3. D'après mes résultats d'amplification sur les accessions Sha et Mr-0 dans cette région, le locus ch3 contient des réarrangements et/ou des variants structuraux dans l'accession Sha par rapport à Col-0, mais aussi par rapport à Mr-0. Dans cette région, le fragment VIII, amplifié à partir de l'ADN génomique de Sha, est de taille supérieure à celle attendue chez Col-0 (figure B19 des résultats). Chez Mr-0, le fragment VIII est de la même taille que chez Col-0 où cette région contient les gènes *PRP3* et *ATL5*. Ces gènes codent respectivement une protéine (riche en proline) de structure de la paroi cellulaire et une protéine zinc finger. Une insertion se trouverait donc dans cette région chez Sha entre 23,18 Mb et 23,19 Mb. L'insertion chez Sha dans cette région pourrait contenir un gène supplémentaire (le gène restaurateur) présent chez Sha et absent chez Mr-0 ; dans ce cas, il faudrait imaginer, du fait du caractère récessif de la restauration, que les allèles Mr-0 agissent en altérant l'expression du gène supplémentaire restaurateur de Sha.

### **Le locus mainteneur ch1**

Le locus ch1 est un locus qui est apparemment plus complexe que le locus ch3. Au début de la cartographie, ce locus mainteneur était considéré comme unique : lorsque l'intervalle de 3,74 Mb, situé à droite de NGA128 (20,63 Mb) et à gauche de F5I14 (24,37 Mb) est hétérozygote, les individus qui sont dans un fond génétique Sha (hormis au locus mainteneur ch3) et cytoplasmique IDS, sont stériles. Nous nous sommes ensuite aperçu, en voulant diminuer la taille du locus ch1, que tous les recombinants dans cet intervalle étaient fertiles. Pour expliquer l'absence de recombinants stériles et les contradictions entre les recombinants fertiles dans cette région génomique, l'hypothèse selon laquelle il est possible que le locus ch1 soit constitué en réalité de deux locus, appelés ch1A et ch1B, qui doivent chacun porter l'allèle Mr-0 afin de maintenir la stérilité mâle, a été formulée. Même s'ils doivent être confirmés, les résultats de phénotypage et de génotypage des populations issues des croisements réciproques entre des recombinants fertiles dans l'intervalle ont montré qu'on peut obtenir des individus stériles dans ces populations, ce qui conforte l'hypothèse.

Les croisements réciproques entre les recombinants fertiles ont également permis d'identifier des locus distorceurs de ségrégation dans la région concernée. Ces locus, dont les phénotypes

induits, la localisation génétique et la signification évolutive seront commentés plus loin, devront être cartographiés indépendamment des locus mainteneurs impliqués dans le système de SMNC.

Concernant les facteurs nucléaires impliqués dans le maintien de la SMNC, l'objectif était d'identifier des gènes nucléaires candidats mainteneurs. Ce travail a permis d'identifier les locus génétiques qui les portent. Le nouvel intervalle contenant le locus ch3 est maintenant assez petit, d'après la séquence de Col-0, pour initier une approche « gène candidat » et confirmer le facteur génétique mainteneur de stérilité à ce locus par une approche de validation fonctionnelle. La cartographie des locus ch1 n'est pas terminée et nécessitera encore beaucoup de travail avant de pouvoir proposer des candidats mainteneurs à ce locus. Cette cartographie pourra nécessiter une fixation des individus à un locus pour rechercher des recombinants à l'autre locus lié. Les recombinants que j'ai obtenus dans cette région pourront être utilisés pour la construction de ces nouvelles populations de cartographie. Une fois les différents locus mainteneurs identifiés dans des intervalles de taille raisonnable, ils pourront faire l'objet d'une approche « gène candidat » et/ou d'une analyse des variants entre Sha et Mr-0 aux locus ch1 mainteneurs.

### **Les locus distorceurs de ségrégation**

Les résultats de génotypage et de phénotypage des huit descendances d'autofécondation de recombinants fertiles des PC1 et PC1.2, et des descendants de croisements entre recombinants fertiles de ces populations, semblent indiquer qu'il existe au moins deux locus distorceurs de ségrégation sur le chromosome 1, à proximité de la région portant le(s) locus de maintien. L'un des locus distorceurs, situé à gauche de MSAT1.19860 (19,86 Mb), entraîne une contre sélection de l'allèle Mr-0 dans la descendance par la voie mâle de plantes hétérozygotes. Un second locus, situé à droite de MSAT1.21673 (21,67 Mb), entraîne une contre sélection de l'allèle Sha dans la descendance par la voie mâle de plantes hétérozygotes. À ce stade des résultats, ce second locus co-localise avec le locus mainteneur ch1B. Des travaux supplémentaires devront être effectués afin de savoir si ce distorceur a un lien fonctionnel avec le maintien de la stérilité mâle par le locus ch1B.

La présence des locus distorceurs de ségrégation au chromosome 1 dans le croisement Sha x Mr-0 a été mise en évidence tardivement au cours de la thèse. Ces locus n'étaient pas attendus

au début de ce travail, même si un locus de ce type avait déjà été détecté au bas du chromosome 3. Une cartographie plus fine de ces locus aurait nécessité le développement et l'expérimentation de nouveaux plans de croisements, ainsi que des analyses supplémentaires de génotypage et de phénotypage. Je ne me suis pas plus intéressé à ces locus distorateurs de ségrégation car leur étude ne faisait pas partie de mes objectifs. D'après une analyse globale, réalisée par F. Budar, de tous les résultats de génétiques obtenus dans l'équipe pour cette région, au moins quatre locus distorateurs de ségrégation, dont les deux identifiés au cours de mon travail, sont localisés tout le long du chromosome 1 dans les descendances de croisements entre Sha et Mr-0.

### **Les observations cytologiques**

La caractérisation phénotypique de la stérilité mâle nucléo-cytoplasmique Sha/Mr-0 a été faite en observant la viabilité des grains de pollen et les anthères en développement d'individus fertiles (génotype Sha aux locus mainteneurs ch1A, ch1B et ch3) et stériles (génotype hétérozygote aux locus mainteneurs ch1A, ch1B et ch3).

Les individus fertiles produisent des grains de pollen viables et des siliques allongées ; les individus stériles ne produisent pas de grains pollen viables et sont donc incapables de produire des siliques allongées. Le fait que les grains de pollen avortés des plantes stériles ont une taille normale, comparable à celle des grains de pollen des individus fertiles, et que quelques grains de pollen viables au test d'Alexander ont été observés chez ces individus stériles nous font penser que les grains de pollen ont avorté à un stade tardif de leur développement. Chez les individus stériles, le développement des anthères semble normal, comparable aux individus fertiles, hormis le fait qu'ils sont incapables de produire du pollen viable.

En ayant connaissance de la présence de locus distorateurs à proximité de ce locus, nous avons caractérisé des individus hétérozygotes dans une région et pas dans l'autre, et qui sont dans un fond Sha (figure B22). Ces individus sont fertiles et produisent des siliques allongées. Cependant, nous nous sommes rendu compte que la viabilité des grains de pollen était affectée (environ 50% de grains colorés) chez ces individus hétérozygotes dans seulement une des deux régions nécessaires au maintien de la SMNC (et Sha à l'autre). Ainsi, des individus qui semblaient parfaitement fertiles à l'oeil nu ont montré un défaut de viabilité des

grains de pollen, certainement dû à l'état hétérozygote des locus distorceurs de ségrégation présents au sud du chromosome 1 ou 3. Même si une plante présente des siliques normalement allongées à l'oeil nu, il se peut que la production de pollen soit quand même affectée lorsqu'on l'observe en microscopie. Environ 50% de pollen viable sont suffisants pour produire des siliques normalement allongées. Lors de la cartographie, des locus distorceurs de ségrégation ont été mis en évidence au sud des chromosomes 1 et 3. La présence des distorceurs de ségrégation dans ces régions peuvent expliquer les observations en coloration d'Alexander qui suggèrent que les grains de pollen n'avortent que s'ils sont produits par une plante hétérozygote au sud du chromosome 1 (ou du chromosome 3), mais pas s'ils sont produits par une plante homozygote Mr-0 ou Sha à ces régions.

Beaucoup d'objectifs définis en début de thèse sur la caractérisation phénotypique de la SMNC n'ont pas été atteints. Un grand nombre de locus identifiés lors de la cartographie ont montré qu'ils étaient capables d'affecter la viabilité des grains de pollen. Il est nécessaire, pour observer le(s) phénotype(s) induit(s) par un locus d'intérêt, d'avoir à disposition le matériel génétique adéquat pour son étude, c'est à dire des génotypes où seuls les allèles au locus étudié ségrégent (les autres locus étant fixés pour l'allèle n'entraînant pas de phénotype) : je n'ai pas eu le temps d'obtenir ce matériel et de réaliser les observations. Pour cette raison, je n'ai pas pu atteindre les objectifs initiaux concernant la caractérisation phénotypique de la SMNC du croisement Sha x Mr-0.

Le peu de temps consacré à la caractérisation phénotype m'a quand même permis de confirmer l'effet des locus mainteneurs Mr-0 sur la viabilité des grains de pollen. Il serait intéressant de confirmer ces résultats en observant des génotypes où seuls les allèles mainteneurs sont présents (et où les locus distorceurs de ségrégation seraient fixés homozygotes dans ces génotypes). Le colorant d'Alexander m'a aussi permis de faire le lien entre les biais de ségrégation observés dans la descendance de plantes fertiles qui n'étaient hétérozygotes que sur une des deux régions chromosomiques concernées et l'avortement d'une proportion des grains de pollen chez ces individus. Ce type d'observation pourra être utilisé pour phénotyper, en vue de cartographie dans des populations adéquates, la position des locus distorceurs de ségrégation.

Étant donné qu'un dysfonctionnement mitochondrial est souvent impliqué dans les systèmes de SMNC connus, il serait intéressant d'observer les mitochondries présentes dans les tissus

des anthères d'individus stériles et de les comparer, à différentes étapes du développement des grains de pollen, aux mitochondries de plantes fertiles Sha. Pour cela, il est possible de marquer les mitochondries avec un anticorps comme l'anti-HSP60. HSP60 est une protéine chaperonne mitochondriale codée par le génome nucléaire et localisée dans les mitochondries ; elle participe à l'importation, dans les mitochondries, de protéines codées par le génome nucléaire (Cheng et al., 1989). On pourrait également réaliser le marquage des mitochondries par une protéine fluorescente comme la GFP adressée aux mitochondries (Duroc et al., 2005). Ainsi, les mitochondries seraient visibles en microscopie à fluorescence et l'on pourrait observer une concomitance entre la mort tardive des grains de pollen et un changement de l'état des mitochondries dans les tissus de l'anthère (suspecté par des résultats préliminaires qui restent à confirmer).

### **Quelle est la signification évolutive de la SMNC Sha x Mr-0 ?**

D'après les phénotypes des F1 issues des croisements réciproques avec Mr-0, douze accessions d'Asie centrale ont un cytoplasme inducteur de stérilité mâle (IDS). Ces résultats montrent que la faculté d'induire une stérilité mâle en présence d'allèles de Mr-0 est partagée par le cytoplasme de plusieurs accessions d'Asie centrale. Étant donné que toutes les accessions porteuses du cytoplasme inducteur de stérilité appartiennent au même cytotype I, il est raisonnable de penser que c'est le même déterminant inducteur de stérilité qui agit dans toutes les F1 stériles observées. De même, les allèles mainteneurs de stérilité sont partagés par au moins deux accessions : Mr-0 et 703AV. Ces deux accessions, capables de maintenir la SMNC induite par le cytoplasme de Sha, proviennent de la même région d'Italie, et sont génétiquement très proches, d'après le génotypage réalisé par Simon et al (2011). Le fait que d'autres accessions soient capables d'induire (comme Sha) et de maintenir (comme Mr-0) la stérilité mâle nous confirme que cette SMNC reflète une situation issue de l'histoire évolutive de l'espèce : le phénotype de stérilité n'est pas induit par des variations génétiques qui seraient apparues depuis la collecte de ces accessions.

Il est possible que la stérilité mâle étudiée au cours de cette thèse soit l'expression d'une SMNC cryptique, active dans des populations ancestrales d'*Arabidopsis* il y a quelques milliers d'années. Une SMNC est dite dans un état « cryptique » lorsqu'elle ne peut s'exprimer dans une population naturelle du fait de la fixation des restaurateurs dans cette population où tous les individus sont hermaphrodites. Dans les modèles évolutifs de SMNC,

la présence du gène mitochondrial IDS induit une pression de sélection sur le génome nucléaire pour la fixation de restaurateurs capables de permettre la production de pollen en présence du cytoplasme IDS. La fixation des restaurateurs dans les génotypes nucléaires d'une population (ou d'une espèce) aboutit à une « disparition », au niveau phénotypique, de la SMNC, même si le facteur cytoplasmique IDS est toujours présent. Il est possible que de nombreuses SMNC n'aient pas encore été observées dans les populations naturelles hermaphrodites à cause de leur état cryptique. Peu de SMNC ont été décrites comme étant des SMNC cryptiques : une SMNC cryptique (à restauration dominante) a été révélée dans des hybrides F2 *Mimulus guttatus* X *M. nasutus* porteurs du cytoplasme de *M. guttatus* (Fishman et Willis, 2006). D'après eux, cette stérilité reflète la présence d'un système de restauration fixé dans les populations *M. guttatus* hermaphrodites. Comme seules des populations composées d'individus hermaphrodites ont été identifiées chez *A. thaliana*, il faut imaginer, dans l'hypothèse d'une SMNC cryptique, que les restaurateurs nucléaires sont aujourd'hui fixés dans les populations d'Asie centrale IDS. Dans mon modèle d'étude, et contrairement à la plupart la plupart des systèmes de SMNC où les allèles restaurateurs sont dominants, ce sont les allèles mainteneurs qui sont dominants. Ainsi, la présence des allèles Mr-0 pourrait avoir un effet direct en favorisant l'expression du déterminant IDS, Dans ce cas, on peut imaginer qu'il manque, chez les accessions restauratrices, un ou plusieurs allèles nécessaires à l'expression de *orf117Sha*. Ces allèles seraient absents de la majorité des accessions, dont celles d'Asie centrale, et présents uniquement dans les accessions Mr-0 et Etna-2. Dans le cas d'un effet indirect, les allèles mainteneurs pourraient agir en induisant un dérèglement du mécanisme de restauration fixé dans les accessions IDS (qui supprime l'expression de *orf117Sha*). Dans ce cas, une duplication, un réarrangement particulier ou un gène supplémentaire, présent uniquement dans le génome de Mr-0 et d'Etna-2, et capable de supprimer l'effet des restaurateurs, permettrait au déterminant IDS de s'exprimer dans la F1 Sha x Mr-0. Il est possible que le gène de maintien ait un rôle différent dans Mr-0 et se trouverait fortuitement favoriser l'expression de la stérilité mâle par un effet direct, par exemple en promouvant la traduction du gène *orf117Sha*. Aucun phénotype lié à un mécanisme épigénétique dépendant du cytoplasme n'a été décrit dans la littérature. Cependant, ce type de mécanisme pourrait expliquer la dominance du phénotype de stérilité mâle dans le croisement Sha x Mr-0 : dans la F1 stérile, l'expression des allèles restaurateurs Sha serait supprimée par un mécanisme épigénétique (par exemple une méthylation de l'allèle restaurateur Sha par la présence de l'allèle mainteneur Mr-0), permettant ainsi à *orf117Sha* de s'exprimer et d'induire le phénotype de stérilité mâle. Toujours est-il que *orf117Sha*

présente des caractéristiques communes avec les déterminants IDS présents dans d'autres systèmes de SMNC : cette orf non conservée du génome mitochondrial ressemble par exemple au déterminant IDS *orf108* décrit chez plusieurs espèces de la famille des Brassicaceae (Ashutosh et al., 2008 ; Kumar et al., 2011). De plus, dans de nombreux exemples bien caractérisés de SMNC, les plantes mâle stériles présentent un phénotype végétatif et floral normal (excepté la stérilité mâle), comme chez les SMNC du pétunia et du tournesol (Hanson et Bentolila, 2004). C'est également le cas pour les individus stériles étudiés pendant la thèse : mise à part l'incapacité à produire du pollen, les individus stériles ont des organes reproducteurs et végétatifs normaux, au moins dans les conditions de culture que nous avons utilisées.

Si le croisement Sha x Mr-0 révèle bien une SMNC cryptique, alors il faut supposer que des populations gynodioïques ont existé dans le passé chez *A. thaliana*. Cela implique un certain niveau d'allogamie pour assurer la pollinisation des individus stériles de ces populations par leurs congénères hermaphrodites. Aujourd'hui, *A. thaliana* est une espèce principalement autogame et ce mode de reproduction ne favorise pas, *a priori*, l'existence d'un système de SMNC dans les populations naturelles. En réalité, nous savons que les populations d'*A. thaliana* ne sont pas 100% autogames et qu'il existe un taux de fécondation croisée (allofécondation) variant de 0,3 à 2,5% (Abbott et Gomes 1989 ; Pico et al., 2008) en fonction du niveau d'étude (global, régional ou local) et selon les populations étudiées (Bomblies et al., 2010). De plus, nous savons que le mode de reproduction d'*A. thaliana* n'a pas toujours été le même et qu'il a certainement évolué d'un mode de reproduction strictement allogame à un mode préférentiellement autogame ; cette transition est estimée à un million d'années et serait la conséquence d'une perte des systèmes d'auto-incompatibilité chez *A. thaliana* (Bechsgaard et al., 2006 ; Tang et al., 2007). L'auto-incompatibilité résulte de systèmes génétiques empêchant l'autofécondation des individus hermaphrodites par reconnaissance et rejet de leur propre pollen ; la perte de ce système peut ainsi favoriser l'autofécondation et l'apparition d'un mode de reproduction préférentiellement autogame. Plusieurs événements indépendants de perte du système d'auto-incompatibilité sont à l'origine des populations autogames observées en milieu naturel (pour revue, Barrett 2002). Des événements indépendants de perte du système d'auto-incompatibilité ont eu lieu chez *A. thaliana* (Boggs et al., 2009). Aujourd'hui, le système d'auto-incompatibilité semble inactif chez cette espèce. Cependant, chez *A. lyrata*, l'espèce la plus proche d'*A. thaliana* (la séparation de ces deux espèces remonte à environ 5 millions d'années (Koch et al., 2001 ; Charlesworth et

Vekemans, 2005), ce système est encore actif et les individus sont auto-incompatibles (et donc allogames). Le locus responsable de l'auto-incompatibilité a été caractérisé chez différentes espèces et est appelé « locus S ». Chez les Brassicaceae, les principaux composants du locus S, situé sur le chromosome 4, sont les gènes *SCR* (pour *S*-locus cysteine rich protein) et *SKR* (pour *S*-locus receptor kinase) qui codent respectivement les déterminants mâle et femelle de l'auto-incompatibilité (Nasrallah, 2000). La transformation d'*A. thaliana* avec les allèles du locus S d'*A. lyrata* permet de restaurer l'auto-incompatibilité chez certaines accessions d'*A. thaliana* (Nasrallah et al., 2004). À l'origine, il est donc très probable qu'*A. thaliana* était une espèce auto-incompatible et donc allogame. Ainsi, à cette époque, l'allogamie aurait été compatible avec l'apparition d'une SMNC et la formation d'une population gynodioïque. Suite à la fixation des restaurateurs, la SMNC serait devenue cryptique. Même si aucune population gynodioïque n'a été identifiée à ce jour chez *A. thaliana*, il ne paraît pas impossible que des populations gynodioïques aient pu exister chez cette espèce.

Pour expliquer le phénotype de stérilité de la F1 Sha x Mr-0, une hypothèse alternative à la SMNC cryptique serait que la stérilité soit due à une incompatibilité génétique de type Bateson-Dobzhansky-Muller (BDM). Ce modèle propose que les phénotypes de stérilité ou de létalité observés chez les hybrides résultent d'interactions épistatiques négatives entre locus (pour revue, Orr, 2005). Le modèle peut être adapté aux interactions nucléocytoplasmiques : des gènes co-adaptés, présents dans le noyau et l'organite (qui codent par exemple des sous-unités d'un même complexe protéique) peuvent être perturbés lorsque qu'un organite d'une lignée est mis en présence d'un fond nucléaire provenant d'une autre lignée. Le degré de perturbation observé chez les hybrides serait fonction du temps de divergence entre les deux lignées parentales (Rand et al., 2004). Selon l'hypothèse d'une incompatibilité nucléocytoplasmique de type BDM, des gènes nucléaires et cytoplasmiques qui codent des produits en interaction fonctionnelle, ont coévolué au sein de chacune des populations ; les nouvelles combinaisons nucléocytoplasmiques ont créé des incompatibilités fonctionnelles qui peuvent avoir des conséquences sur la valeur sélective des hybrides. Par exemple, un cas de barrière génétique intraspécifique résultant d'une coadaptation entre les compartiments mitochondrial et nucléaire a été décrit chez une espèce de crustacé (Ellison et Burton, 2006), et un cas de rupture de coadaptation, visible dès la F1, entre les compartiments chloroplastique et nucléaire a été observé chez *Pisum sativum* (Bogdanova et al., 2009). Ces exemples sont détaillés dans le chapitre V.4 de l'introduction. Dans le cas des

incompatibilités génétiques de type BDM, de nombreux phénotypes sont visibles, comme des déficiences chlorophylliennes ou des nécroses. Par exemple, dans le cas du pois (Bogdanova et al., 2009), le phénotype de stérilité s'accompagne d'autres anomalies (notamment chlorophylliennes), ce qui n'est pas le cas dans mon modèle d'étude où seule la production de pollen est affectée chez les individus mâle stériles. De plus, l'incompatibilité semble avoir une origine nucléo-chloroplastique dans le cas du pois (Bogdanova et al., 2009) alors qu'elle semble plutôt, d'après mes résultats, avoir une origine nucléo-mitochondriale dans mon modèle d'étude. Selon l'hypothèse d'une incompatibilité de type BDM, seuls les locus mainteneurs des accessions Mr-0 et Etna-2 auraient suffisamment divergé de ceux de Sha (et des autres accessions IDS) pour être capables d'induire une rupture de co-adaptation nucléo-cytoplasmique chez les hybrides F1 stériles. Ainsi, les accessions qui ne maintiennent pas la stérilité dans les F1 réciproques avec Sha auraient des facteurs nucléaires compatibles avec le facteur cytoplasmique de Sha, car trop peu divergents de ceux de Sha pour être capables de créer une rupture de co-adaptation nucléo-cytoplasmique. Les facteurs nucléaires et cytoplasmique pourraient, selon cette hypothèse, former par exemple un complexe protéique indispensable et spécifique à la production de pollen ; ainsi, une modification de l'interaction entre les facteurs nucléaires Mr-0 et le facteur cytoplasmique Sha, due à la divergence des locus qui codent ces facteurs, serait à l'origine du phénotype observé.

Cependant, et comme pour l'hypothèse d'une SMNC, il est surprenant d'observer les individus stériles dès la F1 Sha x Mr-0 dans le cas d'une incompatibilité de type BDM. En effet, même s'il existe des exceptions comme dans le cas du pois (Bogdanova et al., 2009) où le phénotype de stérilité est dominant, une seule copie allélique coadaptée au compartiment cytoplasmique est souvent suffisante pour maintenir une coadaptation nucléocytoplasmique fonctionnelle. Le plus souvent, seule l'absence de copie coadaptée au compartiment cytoplasmique permet de révéler une rupture de co-adaptation (la création de F2 réciproques, obtenues suite à une génération d'autofécondation des deux F1 réciproques, est souvent nécessaire pour créer ces ruptures de co-adaptation).

Je trouve que nous avons plus d'arguments en faveur d'une SMNC cryptique que d'une incompatibilité nucléo-cytoplasmique de type BDM : l'hypothèse d'une SMNC cryptique révélée dans la F1 Sha x Mr-0, sera privilégiée dans la suite des travaux qui seront menés au sein de l'équipe sur ce sujet.

## **Les barrières génétiques dans le croisement Sha x Mr-0**

Au cours de l'étude des hybrides issus de Sha et Mr-0 et de leurs descendances, nous avons observé plusieurs phénomènes pouvant être interprétés comme des barrières génétiques entre ces accessions : outre la SMNC Sha x Mr-0, plusieurs locus distorceurs de ségrégation induisant l'avortement du pollen porteur de certains allèles parentaux chez l'hétérozygote, ou des incompatibilités alléliques, ont été mis en évidence dans ce travail ou au cours de travaux préliminaires ou concomitants réalisés dans l'équipe.

L'accession Sha provient d'une région géographique particulière pour Arabidopsis, l'Asie centrale où il est connu que plusieurs populations sont restées génétiquement isolées des autres populations dans des refuges après la fin de la dernière glaciation (il y a 12000 ans). Les accessions Mr-0 (148AV), Etna-1 (702AV) et Etna-2 (703AV), sont originaires de Sicile en Italie. Du fait de cette localisation géographique insulaire, ces trois accessions, génétiquement très proches les unes des autres, ont pu subir des pressions d'adaptation et un isolement reproducteur, avec des effets sur la fixation d'allèles particuliers (par exemple par dérive génétique) dans leur génome. La confrontation des génomes dans la F1 Sha x Mr-0 aurait non seulement « réveillé » une SMNC cryptique, mais également induit des incompatibilités alléliques intra ou inter locus impliquées dans des phénotypes reproductifs (ou pas) élevant ainsi des barrières génétiques entre les deux parents.

## **Incompatibilités alléliques inter-locus dans le croisement Sha x Mr-0**

Une incompatibilité allélique inter-locus se manifeste chez des génotypes qui ont une combinaison allélique délétère impliquant deux locus génétiques. Par exemple, la combinaison d'un des allèles parentaux à l'état homozygote à un locus avec l'allèle de l'autre parent, aussi à l'état homozygote à l'autre locus, peut induire un phénotype délétère et constitue donc une incompatibilité génétique. On peut tester les incompatibilités alléliques inter-locus dans des familles présentant un réassortiment des allèles fixés de deux parents, comme des F2 ou des RIL. Si une incompatibilité de ce type existe dans une population, alors une association non aléatoire (c'est-à-dire une corrélation négative) entre des marqueurs situés à des locus différents devrait apparaître lors du génotypage de cette population où les individus porteurs de la combinaison délétère devraient être sous représentés.

Dans des RILs dérivées du croisement entre les accessions Col-0 et Bur-0, Vlad et al (2010) ont identifié plusieurs régions qui impactent la croissance des plantes. Un des locus identifiés a un effet majeur sur la croissance seulement lorsqu'un génotype spécifique est présent à un second locus. La raison pour laquelle il existe cette interaction épistasique entre ces deux locus est que l'allèle fonctionnel a été transposé d'un locus à l'autre chez Bur-0, par rapport à Col-0 (Vlad et al., 2010). Des incompatibilités intraspécifiques, dues à des duplications de gènes, ont également été identifiées par Bikard et al (2009) qui ont caractérisé les acteurs moléculaires d'une incompatibilité génétique présente dans des populations issues du croisement entre les accessions Col-0 et Cvi. La cartographie indique que deux locus, situés sur des chromosomes différents, sont représentés chez Col-0 par des copies dupliquées d'un gène codant une histidinol-phosphate amino-transferase où l'une des copies est transcriptionnellement inactive. Chez Cvi, cette copie est active mais l'autre est inactive. Dans les populations F2, 1/16 des individus héritent des deux copies inactives à l'état homozygote et meurent au stade embryonnaire (Bikard et al., 2009)

Chez *A. thaliana*, il a également été montré que dans des populations RILs issues d'un croisement entre les accessions Sha et Col-0, deux locus situés sur le chromosome 4 (K4) et le chromosome 5 (K5) ne ségrègent pas indépendamment (Durand et al., 2012). En effet, l'une des quatre combinaisons alléliques homozygote (Col-0 au K4 et Sha au K5) est rare (3%) dans ces RILs par rapport à la proportion attendue (25%); cette combinaison allélique est qualifiée d'incompatible. Chez Col-0, le gène *AtFOLT1* est situé au locus K5 alors qu'il existe une copie supplémentaire de ce gène chez Sha (nommée *AtFOLT2*) et situé au locus K4 impliqué dans l'incompatibilité, où aucun gène homologue n'existe chez Col-0. Le locus Sha *AtFOLT2* a une organisation complexe et contient notamment des séquences *AtFOLT* répétées et des séquences homologues du génome chloroplastique. *AtFOLT1* est exprimé chez Col-0 mais pas chez Sha alors que *AtFOLT2*, absent de Col-0, est exprimé chez Sha. Aucune expression de ces deux gènes n'a été détectée chez les individus Col-0 au locus K4 et Sha au locus K5 où le manque de transcrits *AtFOLT* est responsable de la diminution de la fertilité. Ils ont également montré que l'absence d'ARNm d'*AtFOLT1* chez Sha résulte d'une répression de l'expression de ce gène par méthylation de l'ADN à des sites spécifiques. Des petits ARNs de 21 et 24 nucléotides capables de reconnaître les séquences *ATFOLT* sont détectés chez Sha mais pas chez les accessions dépourvues du gène paralogue *AtFOLT2* comme Col-0. Le locus *AtFOLT2* synthétise de petits ARNs qui initient la méthylation en *trans* des séquences homologues, comme le gène *AtFOLT1*, par la voie RdDM (pour RNA-

directed DNA methylation) (Simon et Meyers, 2011). Ces petits ARNs ne sont pas nécessaires pour maintenir une transmission stable de la méthylation et de la répression d'*AtFOLTI* au cours des générations (Durand et al., 2012).. Des variations structurales entre accessions peuvent jouer un rôle important dans la formation d'épiallèles spécifiques qui, chez des hybrides, induisent des incompatibilités pouvant affecter la production de pollen. Chez les plantes, des mécanismes épigénétiques peuvent en effet contribuer à des barrières reproductives entre espèces (Ishikawa et Kinoshita, 2009 ; Tarutani et al., 2010). Un petit nombre de gènes, qui présentent des profils de méthylation différents entre accessions, ont été décrits comme responsables de différents phénotypes chez les plantes. Par exemple, chez le maïs, le locus *bl*, qui affecte la pigmentation des plantes, est un cas classique de paramutation, c'est-à-dire de transfert d'information épigénétique entre deux allèles d'un gène, entraînant des changements d'expression héréditaires du gène en question (Chandler, 2007). Chez *Arabidopsis*, des études récentes ont décrit des changements épigénétiques transgénérationnels de profils de méthylation de l'ADN. Les nouveaux épiallèles pourraient altérer la transcription de certains gènes et contribuer à la variation phénotypique (Becker et al., 2011).

Des incompatibilités alléliques inter-locus pourraient expliquer le biais de ségrégation observé pour le marqueur situés au sud du chromosomes 5 (MSAT5.19) chez les individus stériles de la population RM1 du croisement Sha x Mr-0 (figure B1). Il existe dans cette population un excès de plantes stériles hétérozygotes à ce marqueur alors que cette région ne semble pas impliquée dans le phénotype. Une analyse approfondie des résultats de génotypage de cette population a montré qu'il existe un déséquilibre de liaison entre marqueur MSAT5.19 et les marqueurs du sud du chromosome 3. Ce déséquilibre de liaison est probablement dû à des combinaisons délétères à au moins deux locus entre les allèles Sha et Mr-0.

### **Incompatibilités alléliques intra-locus dans le croisement Sha x Mr-0**

Dans le croisement Sha x Mr-0, des distorsions de ségrégation ont été observées chez des individus ayant un fond génétique majoritairement Sha, donc en grande partie débarrassés de la majorité des incompatibilités alléliques. Une distorsion de ségrégation se manifeste par une déviation de la fréquence observée pour un marqueur génétique en ségrégation, par rapport aux ratios mendéliens attendus dans la population. Ces distorsions de ségrégation peuvent être dues à des locus appelés « locus distorceurs ». Dans ce modèle d'interaction allélique le locus

distorteur à l'état hétérozygote dans la plante mère entraîne la mort sélective des gamètes qui portent un des allèles parentaux, entraînant la faible transmission de cet allèle dans la descendance (Long et al., 2008). Chez l'individu hétérozygote au locus distorteur, un allèle parental (l'allèle distorteur) est favorisé dans sa transmission à la descendance par rapport à l'autre allèle (l'allèle délétère). Les allèles distorceurs peuvent être qualifiés d'égoïstes car ils empêchent, chez l'hétérozygote, la production des gamètes qui ne les contiennent pas.

Plusieurs études, notamment chez le tabac (Cameron et Moav, 1957), le blé (Loegering et Sears, 1963) et le riz (Sano, 1990), ont montré qu'une distorsion de ségrégation pouvait être due à la mort des grains de pollen qui portent l'allèle ou le variant délétère ; l'autre allèle, qui se répand dans la population au détriment de l'allèle délétère sous représenté, étant l'allèle distorteur. La stérilité d'individus hybrides chez les plantes peut être induite par des interactions alléliques intra-locus qui dérivent de populations divergentes. Par exemple, le riz asiatique *O. sativa* L. est classé en deux types majeurs (ou sous-espèces), *indica* et *japonica*, car les hybrides de ces deux sous-espèces sont, entre autres, faiblement fertiles. Long et al (2008) ont cloné un locus de stérilité mâle, *Sa*, chez des hybrides *indica-japonica*, comprenant deux gènes adjacents, *SaM* et *SaF*, qui codent respectivement une protéine E3 ligase-like et une protéine F-box. La plupart des cultivars *indica* ont les copies alléliques *SaM*<sup>+</sup> et *SaF*<sup>+</sup>, alors que les cultivars japonais ont les allèles *SaM*<sup>-</sup> et *SaF*<sup>-</sup> qui divergent par des variations de séquence. Ils ont montré que la stérilité mâle des hybrides était contrôlée par des interactions directes et indirectes entre les différents allèles des gènes de ce locus (Long et al., 2008).

Des populations d'autofécondation de certains génotypes, issus du croisement Sha x Mr-0, ont permis d'identifier des biais de transmission à deux locus distorceurs de ségrégation situés sur le chromosome 1. Ces génotypes, qui ont un fond génétique Sha et sont hétérozygotes au sud du chromosome 1, sont affectés dans la production de pollen et produisent, par autofécondation, des populations qui présentent des distorsions de ségrégation à des marqueurs situés sur le chromosome 1. Des croisements ont permis de confirmer que seule la voie mâle était affectée chez ces génotypes. Le premier locus distorteur identifié semble entraîner la mort sélective des gamètes qui portent l'allèle Mr-0 et donc la forte transmission de l'allèle Sha à ce locus dans la descendance. Le second locus distorteur entraîne la forte transmission de l'autre allèle parental, l'allèle Mr-0. L'allèle distorteur n'est donc pas le même pour les deux locus : au premier locus, il s'agit de l'allèle Sha alors que pour le second

locus, l'allèle distorateur est l'allèle Mr-0. Selon le locus qui est hétérozygote, l'allèle parental favorisé dans sa transmission n'est pas le même chez une plante qui a un fond génétique Sha. Cela reflète des mécanismes d'interactions alléliques complexes et spécifiques à chaque locus. Nul doute que des incompatibilités alléliques intra-locus, comparables à celles observées chez le riz (Long et al., 2008) sont impliquées dans des phénotypes reproductifs chez des hybrides Sha/Mr-0.

Les intervalles hétérozygotes identifiés contenant ces locus sont encore grands (plusieurs Mb) et d'autres locus distorateurs sont, d'après F. Budar, présents le long du chromosome 1 et ailleurs dans le génome (notamment sur le chromosome 3). Il n'est donc pas exclu que plusieurs locus hétérozygotes interagissent par des interactions épistatiques complexes au sein des intervalles définis au cours de cette thèse. De plus, même si ce type d'incompatibilités alléliques délétères au sein d'un locus semble également affecter la viabilité des grains de pollen dans la F1 hétérozygote Mr-0 x Sha (qui n'est pas soumise à la SMNC), il est difficile de savoir avec certitude si l'effet de ces locus distorateurs est dépendant ou non du cytoplasme, car les combinaisons génétiques parfaitement réciproques n'ont pas été étudiées. À noter qu'aucun locus distorateur dépendant du cytoplasme n'a été, à ma connaissance, décrit dans la littérature. Il n'est pas exclu que certains locus distorateurs de ségrégation soient actifs quelque soit la nature du cytoplasme, mais agissent en contre sélectionnant l'allèle parental en fonction de la nature du cytoplasme.

### **Les barrière(s) génétique(s) dans le croisement Sha x Etna-1**

L'accession Etna-1, originaire de la même zone géographique que Mr-0 et Etna-2 et très proche de ces accessions, induit un phénotype de létalité dans les F1 réciproques avec Sha. Ce phénotype n'est donc pas lié, *a priori*, à une interaction nucléo-cytoplasmique. Des locus indépendants du cytoplasme et impliqués dans des phénotypes de létalité ont déjà été observés chez certains hybrides inter- ou intra-spécifiques de plantes. Différents phénotypes délétères résultent d'interactions épistatiques déficientes entre allèles à un ou plusieurs locus (Bomblies et Weigel, 2007). Bomblies et al (2007) ont utilisés quelques 300 accessions d'*A. thaliana* pour réaliser 1500 croisements afin identifier des hybrides anormaux : une vingtaine d'hybrides nécrotiques ont été observés dans les populations F1. Deux locus qui agissent de façon dominante ont été identifiés comme responsables de la majorité des phénotypes de nécrose : *DMI* (pour *DANGEROUS MIX1*) et *DM2* qui sont des gènes homologues aux gènes

de résistance aux pathogènes de la classe *TIR-NB-LRR* (les protéines NB-LRR sont impliquées dans l'activation de la signalisation en cascade permettant de déclencher la mort cellulaire programmée (pour revue, Bomblies et Weigel, 2010)). Par exemple, un phénotype de nécrose, dû à une activation inappropriée des réponses immunitaires de la plante, est observé chez des hybrides F1 issus du croisement entre les accessions Uk-1 et Uk-3 (Bomblies et al., 2007). Il est possible que les pressions de sélection exercées par les pathogènes sur ce type de gène puissent entraîner une évolution rapide des variants de ces gènes qui est bénéfique pour la lignée parentale mais pouvant causer de sérieux problèmes dans les descendances hybrides à cause des incompatibilités alléliques. Dans cette même étude, 2% des combinaisons testées en croisements réciproques avec des centaines d'accessions (pour 861 combinaisons uniques) présentent des phénotypes nécrotiques (Bomblies et al., 2007). Ils observent également un phénotype pléiotrope (petite taille, excroissances foliaires...) identique dans les deux F1 des croisements réciproques réalisés avec les accessions Sha et Bla-1. Le gène responsable de ce phénotype, un gène *RLK* (pour *receptor-like kinase*) appelé *OAK* (pour *OUTGROWTH-ASSOCIATED PROTEIN KINASE*) se trouve dans un cluster dont la structure est hypervariable. Lors de la duplication du gène ancestral *RLK*, des séquences codantes ont été recrutées pour former un promoteur avec un nouveau domaine d'expression. La divergence dans le domaine extracellulaire de la protéine entraîne une évolution des allèles qui, lorsqu'ils interagissent dans les hybrides Bla-1/Sha produisent des phénotypes non visibles chez les accessions parentales (Smith et al., 2011). Le phénotype de nécrose observé dans le croisement Sha x Etna-1 n'a pas été étudié plus avant car il ne semble pas dépendre d'une interaction nucléo-cytoplasmique. Cependant, par analogie avec les cas décrits dans la littérature, le phénotype de nécrose observé dans les deux F1 des croisements réciproque Sha x Etna-1 pourrait être dû à une interaction allélique déficiente à un ou plusieurs locus, ou entre locus hétérozygotes Sha/Etna-1. Les accessions Etna-1 et Mr-0 sont génétiquement très proches (Simon et al, 2011) ; pourtant, la létalité des hybrides n'a été observée qu'avec Etna-1. Il se peut que les allèles mainteneurs de la SMNC soient également présents chez Etna-1, mais leur l'effet n'aurait pas pu être observé à cause du phénotype de létalité (cette hypothèse est en cours de test dans l'équipe).

La stérilité hybride est une forme commune de mécanisme d'isolement postzygotique entre espèces et sous-espèces divergentes. Elle peut initier un isolement génétique, éventuellement une spéciation, et joue un rôle important dans le maintien de l'identité des espèces. Dans le croisement Sha x Mr-0, plusieurs locus génétiques sont impliqués dans des phénotypes de

stérilité : certains sont indépendants du cytoplasme, comme les locus distorceurs, et ne reflèteraient *a priori* pas de co-évolution nucléo-cytoplasmique ; d'autres locus sont dépendants du cytoplasme, comme les locus mainteneurs de la SMNC cryptique, et reflètent certainement une co-évolution nucléo-cytoplasmique. On peut alors se demander si les barrières génétiques qui affectent la fertilité mâle dans le croisement Sha x Mr-0 ne révèlent pas un processus évolutif en cours qui pourrait aboutir à terme à des événements de spéciation au sein d'*Arabidopsis thaliana*.

# **MATERIELS ET METHODES**

## MATERIEL

### I. Plantes

Toutes les accessions utilisées dans ce travail ont été fournies sous forme de graines par le Centre de Ressources *Arabidopsis thaliana* pour la génomique de Versailles ([http://www-ijpb.versailles.inra.fr/fr/cra/cra\\_accueil.htm](http://www-ijpb.versailles.inra.fr/fr/cra/cra_accueil.htm)). Leurs caractéristiques sont indiquées dans le tableau C1 et dans les fiches descriptives de l'annexe 1 téléchargées sur le site VNAT (Variation Naturelle chez *Arabidopsis thaliana* - <http://dbsgap.versailles.inra.fr/vnat/>).

### II. Oligonucléotides

Les séquences des oligonucléotides utilisés pour amplifier des MSAT nucléaires sont données dans le tableau C2 pour les MSAT localisés au sud du chromosome 1, et dans le tableau C3 pour les MSAT localisés au sud du chromosome 3. Ces amorces permettent d'amplifier des MSAT polymorphes entre les allèles parentaux Sha et Mr-0. Les autres couples d'amorces testés mais non polymorphes sont indiqués dans l'annexe 2. Les couples d'amorces utilisés en multiplex, et qui sont différents de ceux utilisés en simplex, sont indiqués dans le tableau C4. Les amorces utilisées pour amplifier les marqueurs CAPS polymorphes, dans la cartographie fine de ce locus, sont indiquées dans le tableau C5 avec les caractéristiques des enzymes de restriction utilisées, et celles utilisées pour amplifier les longs fragments recouvrant le locus du chromosome 3 sont indiquées dans le tableau C6.

Les séquences des oligonucléotides qui permettent d'amplifier les régions intergéniques du génome chloroplastique (pour définir les chlorotypes des accessions) sont indiquées dans le tableau C7 et celles permettant d'amplifier l'ADN mitochondrial (pour définir les mitotypes) sont indiquées dans le tableau C8. Les amorces utilisées lors des cRT-PCR (annexe 3), pour analyser la taille des transcrits de gènes mitochondriaux, ont été développées par Marie Verzaux (stagiaire – élève ingénieur de l'ESA) dans le cadre d'une étude réalisée sur les processus post-transcriptionnels qui affectent l'expression des gènes des organites chez *Arabidopsis*.

Les séquences des oligonucléotides qui permettent d'amplifier spécifiquement chaque copie du gène mitochondrial *trnK* sont indiquées dans le tableau C9 ;

## METHODES

### I. Culture des plantes et croisements

Les plantes sont cultivées en serre, en chambre de culture ou en chambre froide selon les besoins. En serre et en chambre de culture, les conditions de culture sont les suivantes : jours longs avec 18h de lumière par jour et des températures qui varient entre 22°C le jour et 18°C la nuit. En chambre froide, la température est de 6°C avec un temps de lumière de 6h/24h (jours courts). Les graines sont semées sur un terreau standard et quotidiennement arrosées à l'eau et une fois par semaine à la solution nutritive.

Des semis sur nitrate ont parfois été utilisés pour la levée de dormance de certaines graines (surtout pour Mr-0). La composition du milieu est la suivante : 100ul de KNO<sub>3</sub> 70uM + 1 mL d'agarose 0,1% d'agarose par boîte. Chaque boîte est laissée 2-3 jours en chambre froide puis les graines sont semées sur terreau universel.

Deux types de croisements sont réalisés : des croisements entre plantes fertiles ; des croisements en utilisant un parent mâle stérile. La castration manuelle est nécessaire seulement dans le premier cas, pour supprimer l'autofécondation et rendre artificiellement mâle stérile l'un des parents. La castration se fait en supprimant les anthères avant leur déhiscence en ouvrant des boutons sur le point de s'épanouir. Les castrations et les croisements sont réalisés sous la loupe binoculaire.

### II Biologie moléculaire

#### II.1 Extraction d'ADN en tubes

Un fragment de feuille de rosette est congelé dans l'azote liquide et broyé avec du sable de Fontainebleau dans un tube eppendorf de 1,5 mL. Un volume de 300 uL de tampon d'extraction préchauffé à 60°C est ajouté. L'ensemble est agité puis centrifugé à 13000 g pendant 6 min. La phase aqueuse (surnageant) est prélevée et transférée dans un nouvel eppendorf contenant 250 uL d'isopropanol. Après une centrifugation de 13000 g pendant 10 min, le surnageant est retiré et le culot est séché à l'air libre sur la pailleuse pendant une demi-heure. Le culot est repris dans 300 uL d'eau stérile. Après une centrifugation de 13000 g pendant 10 min, le surnageant est transféré dans un nouvel eppendorf contenant 30 uL d'acétate de sodium 3M à pH 5,2 et 825 uL d'éthanol 95%. Après une centrifugation d'une demi-heure à 13000 g, le surnageant est éliminé et le culot laissé à sécher sur la pailleuse au

moins une demi-heure. L'ADN est finalement resuspendu dans 100 uL d'eau stérile.

#### Tampon d'extraction

Tris-HCl : 200 mM

NaCl : 250 mM

EDTA : 25 mM

SDS : 0,5% (p/v)

pH 7,5

## II.2 Extraction d'ADN en plaques

Pour le génotypage d'un grand nombre de plantes, les extractions d'ADN ont été effectuées en plaques de 96 puits, où quatre puits sont réservés pour les contrôles (92 plantes analysées par plaque). Un fragment de feuille de rosette est placé dans un des 96 tubes d'une plaque, chacun contenant une bille de métal (diamètre 4 mm). Les échantillons sont passés à -80°C et ensuite lyophilisés (pendant une nuit), puis broyés à l'aide d'un broyeur à vibrations 3 min à 30 vibrations/seconde. Après une centrifugation de 5 min à 3000 g, un volume de 150 uL de tampon d'extraction est ajouté dans chaque tube. La plaque est ensuite centrifugée 15 min à 3000 g puis 100 uL de surnageant sont prélevés et transférés dans une plaque de 96 puits en polypropylène, chacun contenant 100 uL d'isopropanol. Le contenu des puits est homogénéisé par aller-retour dans la pipette puis incubé à température ambiante pendant 5 min. Les plaques sont ensuite recouvertes d'un film plastique puis centrifugées 15 min à 3000 g. Le surnageant est vidé et le culot est rincé avec 100 uL d'éthanol 75 %. La plaque est centrifugée 5 min à 13000 g puis l'éthanol est supprimé. Les plaques sont laissées toute la nuit sur la paillasse pour que le reste de l'éthanol s'évapore. 60 uL d'eau stérile sont ajoutés dans chaque puits pour resuspendre l'ADN.

#### Tampon d'extraction

Tris-HCl : 200 mM

NaCl : 250 mM

EDTA : 25 mM

SDS : 0,5% (p/v)

pH 7,5

### II.3 Extraction d'ARN en tubes

La manipulation des ARN nécessite de prendre certaines précautions pour éviter toute contamination avec des RNAses. Il faut ainsi travailler au froid (en chambre froide ou sur la glace) et utiliser, pour sa manipulation, des gants, un Benchcoat et du matériel stérile, ainsi que des produits exclusivement réservés à la manipulation des ARN.

Des inflorescences sont prélevées sur plusieurs plantes de même génotype (un tube stérile de 20 mL est rempli en totalité) avec des pinces et des ciseaux préalablement flambés à l'alcool. Les échantillons sont immédiatement congelés dans l'azote liquide. Ils peuvent être conservés à -80°C en attendant l'extraction.

Les ARN totaux des échantillons ont été extraits en utilisant le Trizol reagent en suivant le protocole préconisé par le fournisseur (Chomezynski et Sacchi 1987).

La première phase de préparation est une phase d'homogénéisation, où les tissus congelés sont broyés en fine poudre dans un mortier stérile avec de l'azote liquide. La poudre est transférée encore congelée dans un tube Eppendorf pesé à vide. 100 mg de poudre sont pesés puis 1mL de Trizol est ajouté dans le tube sous une sorbonne. Le Trizol est une solution monophasique de phénol et de guanidine isothiocyanate, qui maintient l'intégrité des ARN. Le tube est agité pour homogénéiser le contenu.

Les phases aqueuses et organiques sont séparées après ajout de chloroforme. 200 uL de chloroforme sont ajoutés sous sorbonne puis le tube est agité vigoureusement pendant 15 s. La phase de séparation se fait par une incubation de 5 min à température ambiante pour permettre la dissociation complète des complexes nucléoprotéiques. Après une centrifugation à 4°C pendant 15 min à 13000 g, la phase inférieure rougeâtre contient le phénol/chloroforme, les protéines et l'ADN, et la phase supérieure aqueuse incolore contient l'ARN.

La précipitation des ARN se fait en transférant la phase aqueuse dans un nouveau tube contenant 0,5 mL d'isopropanol ; le tube est incubé 10 min à température ambiante puis centrifugé à 4°C pendant 10 min à 12000 g. Un culot s'est formé au fond du tube.

Le lavage du culot se fait en éliminant le surnageant puis en ajoutant 1 mL d'éthanol 75%. Le tube est centrifugé à 4°C pendant 5 min à 7500 g. Finalement, le surnageant est éliminé et les

culots sont laissés à sécher partiellement pendant 10 min à l'air libre. Le culot est finalement dissous dans 50 uL d'eau « RNase free ». Les échantillons sont incubés 10 min à 55°C pour dissolution puis conservés à -80°C.

#### II.4 Définition des amorces

Le logiciel Primer3 (<http://frodo.wi.mit.edu/>) a été utilisé pour définir les amorces selon les paramètres suivants : la taille des amorces est comprise entre 18 et 22 pb et la température d'hybridation des amorces est de 58°C (+/-2 °C). Pour les MSAT, les produits d'amplification doivent être inférieurs à 200 pb pour visualiser un éventuel polymorphisme.

#### II.5 Amplification par PCR (Polymerase Chain Reaction)

Les réactions d'amplification sont menées dans un volume de 25 uL composé de :

- 0,25 uL de dNTP (mélange à 25 mmol/L des 4 désoxynucléotides)
- 0,25 uL d'une polymérase ADN thermostable (Taq produite au laboratoire à 2U/uL)
- 1,25 uL de chaque amorce diluée à 10 uM
- 2,5 uL de tampon 10X
- 3 uL de MgCl<sub>2</sub> (25 mM)
- 2 uL d'ADN matrice (50 à 100 ng d'ADN extrait)

Le tout est complété à 25 uL avec de l'eau stérile.

##### Tampon 10X

KCl : 100 mM  
(NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub> : 100 mM  
Tris-HCl : 200 mM  
MgSO<sub>4</sub> : 20 mM  
Triton X-100 0,1% (v/v), pH 8,8

À noter que la réaction PCR en multiplex utilisée pour amplifier les MSAT est menée dans un volume de 25 uL composé de la même façon que pour la réaction en simplex en ajoutant 1,25 uL (10 uM) de chaque amorce du second couple d'amorce et en ajustant le volume d'eau en conséquence.

Le programme utilisé pour amplifier les MSAT (en simplex et en multiplex) est le suivant :

- 1/ 2 min de dénaturation à 94°C
- 2/ Dénaturation de l'ADN à 94°C pendant 15 s
- 3/ Hybridation des amorces sur la matrice pendant 15 s à une température de 57°C (cette température peut varier selon la taille et la teneur en GC des amorces). Diminution de 0,5°C à chaque cycle
- 4/ Élongation pendant 30 s à 72°C
- 5/ Cycle à partir de l'étape 2 pendant 7 fois
- 6/ 94°C pendant 15s
- 7/ 53°C pendant 15s (cette température peut varier selon la taille et la teneur en GC des amorces)
- 8/ 72°C pendant 30s
- 9/ Cycle à partir de l'étape 6 pendant 25 fois
- 10/ Élongation finale pendant 2 min à 72°C
- 11/ 10°C pendant 10 min

La réaction PCR pour amplifier les fragments de 6000 à 8000 pb est menée dans un volume de 23 uL composé de :

- 1 uL d'un mélange (5/1) d'une Taq polymérase et d'une polymérase Pfu qui améliore la processivité de la réaction. Le mélange est à 2 U/uL) ;
- 2 uL de chaque amorce diluée à 10 uM ;
- 2 uL de tampon Jeff 10X;
- 2 uL d'ADN matrice (50 à 100 ng d'ADN extrait);
- 14 uL d'eau stérile.

#### Tampon Jeff 10 X

- 450 mM Tris-HCl (pH 8.8)
- 110 mM (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub>
- 45 mM MgCl<sub>2</sub>
- 67 mM 2-mercaptoethanol,
- 44 µM EDTA
- 8 mM dATP
- 8 mM dCTP
- 8 mM dGTP
- 8 mM dTTP
- 1.13 mg/ml BSA.

Le programme utilisé pour amplifier les fragments d'ADN de 6000 à 8000 pb est le suivant :

- 1/ 2 min de dénaturation à 94°C
- 2/ Dénaturation de l'ADN à 94°C pendant 20 s
- 3/ Hybridation des amorces sur la matrice pendant 30 s à une température de 65°C (cette température peut varier selon la taille et la teneur en GC des amorces)
- 4/ Élongation pendant 7 min et 30 s à 68°C
- 5/ Cycle à partir de l'étape 2 pendant 35 fois
- 6/ Élongation finale pendant 12 min à 68°C
- 7/ 10°C pendant 10 min

Les réactions et les programmes PCR pour amplifier les fragments chloroplastiques et mitochondriaux sont les mêmes que ceux utilisés pour amplifier les MSAT.

Les réactions PCR pour amplifier les produits de cRT-PCR sont les mêmes que ceux utilisés pour amplifier les MSAT. Le programme PCR est le suivant :

- 1/ 2 min de dénaturation à 94°C
- 2/ Dénaturation de l'ADN à 94°C pendant 20 s
- 3/ Hybridation des amorces sur la matrice pendant 30 s à une température de 53°C
- 4/ Élongation pendant 1 min et 30 s à 72°C
- 5/ Cycle à partir de l'étape 2 pendant 35 fois
- 6/ Élongation finale pendant 15 min à 72°C
- 7/ 10°C pendant 10 min

## II.6 Circularisation des ARN et préparation d'ADNc pour les cRT-PCR

La première étape de la cRT-PCR est une étape de circularisation des ARN. Le volume réactionnel est de 50 uL et est composé de :

- 5 uL de tampon T4 RNA 10X (Fermentas)
- 5 uL de BSA (1mg/uL)
- 5 uL d'ATP à 10 mM
- 4 uL de T4 RNA ligase (10 U/uL)
- 1 uL d'un inhibiteur de RNase (Fermentas - 40 U/uL)
- 5 ug d'ARN (volume selon les échantillons de 4 à 25 uL)

Le tout est complété à 50 uL avec de l'eau RNase free puis incubé à 37°C pendant 2 h. Après avoir circularisé les ARN, 150 uL d'eau RNase free et 200 uL de phénol (qui permet de dissocier les protéines liées à l'acide nucléique) sont ajoutés. Après avoir vortexé les échantillons jusqu'à l'apparition d'une émulsion, une centrifugation de 5 min à 13000 g est réalisée. 180 uL de phase aqueuse sont transférés dans un tube eppendorf contenant 18 uL d'acétate de sodium 3M pH 5,2 et 500 uL d'éthanol absolu RNase free. Les ARN circularisés sont stockés à -25°C dans l'éthanol. Après une centrifugation de 13000 g pendant 30 min, le surnageant est supprimé et le culot séché pendant une demi-heure à l'air libre. Un lavage du culot est réalisé en ajoutant 200 uL d'éthanol RNase free 75 %. Après une centrifugation de 13000 g pendant 10 min, le surnageant est supprimé et le culot de nouveau séché pendant une demi-heure à l'air libre. Pour réaliser la transcription inverse, le culot d'ARN circularisé est remis en solution dans un volume réactionnel de 25 uL. Ce volume est composé de :

5 uL de tampon MMLV 5X (tampon de la transcriptase inverse)

0,5 uL de dNTP (25 uM)

2 uL d'oligonucléotides (hexamères aléatoires, 100 uM)

1 uL d'un inhibiteur de RNase (Promega) (40 U/uL)

2 uL de l'enzyme M MLV RT (transcriptase inverse (Promega) à 200 U/uL)

14,5 uL d'eau stérile

Chaque tube est ensuite placé pendant 3h à 40°C dans un thermocycleur. Les échantillons d'ADNc sont stockés à -20°C.

## II.7 Électrophorèse

Les échantillons d'ADN migrent dans des gels d'agarose (de concentration variable en fonction de la taille des fragments : 3 % pour des fragments taille < 500 pb et 1% pour des fragments de taille > 500 pb), fondu dans du TBE 0,5 X (Tris 50 mmol/L, acide borique 45 mmol/L, EDTA 0,5 mmol/L). Les échantillons d'ARN migrent dans des gels d'agarose 1% fondu dans du TBE 0,5 X (contrôle de la qualité des ARNs extraits).

Les acides nucléiques sont visualisés grâce à l'ajout dans l'agarose fondu de bromure d'éthidium (0,1ug/mL), un intercalant qui fluoresce aux UV. Les échantillons déposés sur le gel sont au préalable mélangés à 2 uL d'un bleu de charge (bleu de bromophénol 0,25%, xylène cyanol 0,25%, glycérol 30 % et EDTA 50 mmol/L). Après avoir déposé 8 uL du

mélange (échantillon + bleu de charge), la migration s'effectue dans des bacs d'électrophorèse contenant du TBE 0,5 X à environ 140 V pendant une période de temps qui varie en fonction de la concentration en agarose et de la taille des fragments d'ADN. Un marqueur de poids moléculaire est également déposé pour estimer la taille des fragments (1kb DNA ladder ou 50bp DNA ladder (Qiagen)).

Un marqueur quantitatif a parfois été utilisé pour quantifier, avant l'envoi en séquençage, la quantité de produit de PCR : une quantité connue de plasmide pDONR207 (Invitrogen) digéré par l'enzyme *Tat1*. La fluorescence de la matrice à doser est comparée à celle des fragments, de tailles connues, du marqueur.

Des dépôts décalés ont parfois été réalisés pour faciliter la recherche de recombinants lors du génotypage des populations de cartographie. Pour cela, 8 uL de produits PCR amplifiés à l'aide d'un premier couple d'amorces (qui amplifient un MSAT) sont dans un premier temps déposés sur le gel. Une migration d'environ 10 min à 140 V est effectuée puis un second dépôt est réalisé en déposant 8 uL de produits PCR amplifiés à l'aide d'un second couple d'amorces (qui amplifient un autre MSAT). La migration est ensuite reprise à 140 V.

## II.8 Séquençage d'acides nucléiques

Les séquençages sont effectués par la société Genoscreen. Les matrices et les amorces spécifiques sont préparées selon leurs recommandations : 100ng par matrice de taille inférieure à 500pb ; 200ng par matrice de taille comprise entre 500pb et 1000pb ; 400ng par matrice de taille supérieure à 1000pb ; 10 uL d'amorce à 5 ou 10uM par réaction de séquence. Après réception, les séquences sont analysées et alignées avec le logiciel GenalysCarbon 2.8.3 beta développé par Metrowerks CodeWarrior Pro.

## III. Microscopie

### III.1 Coloration d'Alexander

Cette coloration est utilisée pour tester la viabilité des grains de pollen (Alexander, 1969). Des boutons floraux sur le point de s'ouvrir, sont préférentiellement prélevés. 2 lames sont rincées à l'alcool, une pour disséquer et une autre avec 20 uL de colorant d'Alexander, Les anthères sont séparées avec des aiguilles sur la première lame puis déposées sur la seconde lame. Une

lamelle est appliquée délicatement et pressée à proximité des anthères. Les préparations sont laissées 1 h sur la paillasse puis du vernis est appliqué autour de la lamelle pour la sceller. Les préparations sont observées au microscope optique.

#### Colorant d'Alexander

10 ml d'éthanol 95%  
1ml de vert de Malachite (1% dans 95% d'éthanol)  
5 ml de fuchsine acide (1% dans l'eau)  
0,5 ml d'orange G (1% dans l'eau)  
5 g de phenol  
5 g d'hydrate de chlore  
2 ml d'acide acétique glacial  
25 ml de glycerol  
50 ml d'eau distillée

#### III.2 Coloration à l'iodure de propidium

L'iodure de propidium est utilisé pour colorer spécifiquement l'ADN. Il marque également les parois des cellules végétales et permet de révéler la structure des tissus au microscope confocal grâce à sa fluorescence.

Une dizaine de boutons floraux est prélevée puis les anthères sont disséquées. Les échantillons sont fixés dans le PFA (5 vol acide propionique ; 10 vol de formaldéhyde 37 % ; 70 vol éthanol absolu) dans un tube pendant 2h à température ambiante. Un traitement à la RNase (20 ug/ml – Rnase A bovine pancréas Boehringer) est effectué dans du NTE 1X (NTE 5X : NaCl (2,5M), Tris (50 mM) et EDTA (5mM)) pendant 30 min à 37°C. La coloration se fait en ajoutant 475 uL d'arginine pH 12,4 à 0,1M et 25 uL d'iodure de propidium (Sigma) pendant une nuit à 4°C. Un rinçage à l'arginine pH 8 à 0,1 M est réalisé puis les échantillons sont montés sur lame dans du glycérol et du PBS (Citifluor). Les échantillons sont observés en microscopie confocale (Argon/Krypton Laser (LEICA TCS NT) avec une excitation à 568 nm en utilisant un filtre d'émission DD488/568 et un filtre de détection LP590.

Accession name	Versailles identification number	City	Country	Cluster nucléaire (Simon et al., 2011)	Cytotype (référence)
Dh-1	132 AV	Tenne/Ts.	Germany	4b	J (NG)
Wj-1	72 AV	Vilnius	Lithuania	4c	J (NG)
N13	266 AV	Konchezero	Russia	1a	J (Moïsson et al., 2010)
Kondara	190 AV	Kondara		1b	S (Moïsson et al., 2010)
Hodja-Obi-Garm	203 AV	Khodja-Obi-Garm		1b	S (Moïsson et al., 2010)
Sorbo	238 AV	Sorbo		1b	F (Moïsson et al., 2010)
Hodja-Obi-Garm bis	270 AV	Hodja-Obi-Garm	Tadjikistan	1b	S (Moïsson et al., 2010)
Neo-3	539 AV	Neo-Shabdara		1b	I (NG)
Neo-6	540 AV	Neo-Shabdara		1b	I (Moïsson et al., 2010)
Shabdara	771 AV	Shah dara		1b	I (NG)
9481B	261 AV	Kazakhstan	Kazakhstan	1a	I (Moïsson et al., 2010)
Kz-9	403 AV	Atasu		I bis	I (NG)
Kz-1	402 AV			1a	J (NG)
Kz-2	555 AV	Kashmir	Inde	1a	II (Moïsson et al., 2010)
Lenk2-1	584 AV	Lankaran-Lenk	Azerbadjan	1a	M (Moïsson et al., 2010)
Kar-1	531 AV			1b	I (Moïsson et al., 2010)
Kar-2	Karakol	Karacol river		1b	I (NG)
Sun-1	533 AV	Sunamyr		1b	I (NG)
Dja-1	534 AV	Djaly	Kyrgyzstan	1b	I (NG)
Dja-5	535 AV			1b	I (NG)
Zal-1	536 AV	Zaliaky		1b	I (NG)
Zal-3	537 AV			1b	I (NG)
Kyt-1	538 AV	Kyrgyzita		1b	I (NG)
Mt-0	148 AV	Montrosso		2b	AN (Moïsson et al., 2010)
Etma-2	703 AV	Mt Etma	Italy	2b	nd
Etma-1	702 AV			2b	nd

**Tableau C1 : description des accessions d'*Arabidopsis thaliana* utilisées lors de la thèse (d'après <http://dbgap.versailles.inra.fr/vmat/>)**

Dans la dernière colonne, NG : résultat obtenu par ce travail ; nd : non déterminé.

Nom du marqueur	Position (Col-0)	Séquence de l'amorce «Forward» (5'→3')	Séquence de l'amorce «Reverse» (5'→3')	Origine
CIW1	18,37	ACATTTTCTCAATCCTTACTC	GAGAGCTTCTTTATTGTGAT	base MSAT
MSAT1.18748	18,75	TTATATACCATTCTCGTCGTGTTA	GCAAAAATGAAATGTTAAGGC	CW
MSAT 1.197642	19,76	CTCATCCCTCCGACCTATCA	TCATCTCTGATCTACCCCTTTG	NG
MSAT1.19860A	19,86	CACGGCCCATTAGTGTTCTT	GAGAATGCAGAAGCACTGACA	NG
MSAT1.20097	20,10	ATGCTAGATAGATTTTAGATG	TATTACAAGGTGAGTTTATCA	CW
MSAT 1.20495A	20,49	GGGTTGCAAGTCGTAAGTT	TGGCCCTCAGTTTTCATTT	NG
NGA128	20,63	GGTCTGTTGATGTCGTAAGTCG	ATCTTGAAACCTTTAGGGAGGG	base MSAT
MSAT1.20849	20,84	GCATAGGAGTGTGATGGGTTTT	TTTGCCTACCTTGGGAAGCAC	NG
MSAT1.21035	21,03	GGTGGGGATAAGTATGTGCATT	TGGATATCGCAGACAACAGC	NG
MSAT1.21231	21,22	TCCAAATGGAAAAACATCTCC	TTGTGAACCCCTTGTGATGGA	NG
MSAT1.21235	21,23	CTTTTACAAGAGTTGGGAGTTAAGA	TCTCCACACAATCCCATCAA	NG
MSAT1.21456	21,45	AATGCCCATGGCTTTTATGT	GCATTTGGGTCCAATTTGATA	NG
MSAT1.21673	21,67	TGATTTTGTTCATTCGTTGAAA	CGAATCGGGAAACCAATAAT	NG
MSAT1.22002	22,00	AGTCCACGGCCAAAAGAGTT	TTCCGCTCACGTTCTTTAT	NG
MSAT1.22118	22,11	TTTTGGGGACTTACTGTTAGTTTTG	CAATGAAAAGTTGGCAGACG	NG
MSAT1.22301	22,30	ACCCGATCTATGGTGGTTCA	GCAACAAAAATTGGACATGG	NG
ATHGENEA	22,40	ACCATGCATAGCTTAAACTTCTT	ACATAACCACAAAATAGGGGTG	base MSAT
F5I14	24,37	CTGCCTGAAATTGTCGAAAC	GGCATCACAGTTCTGATTCC	base MSAT

**Tableau C2 : oligonucléotides utilisés pour amplifier les MSAT localisés au sud du chromosome 1**  
Ces couples d'amorces amplifient des marqueurs MSAT polymorphes entre les allèles parentaux Sha et Mr-0. Dans la dernière colonne, base MSAT : collection de marqueurs développés pour amplifier des MSAT chez *A. thaliana* (la collection est mise à jour quotidiennement par les équipes de la SGAP) ; NG : Nicolas Gobron ; CW : Cezary Waszczak.

Nom de l'amorce	Position (Col-0)	Séquence de l'amorce «Forward» (5'→3')	Séquence de l'amorce «Reverse» (5'→3')	Origine
MSAT3.02052	2,05	AAGGTTACTCTTATCGTTTAT	GGTTTCAACTTAGACGTTT	base MSAT
MSAT3.19	8,81	TAATTCGATCCAATTGACAT	TGGCTTGGCACAAAC	base MSAT
MSAT3.2	9,06	AAGGTACGGCGGTGGATATTG	CGGGGATTCTTCTTCTCTGTG	base MSAT
MSAT3.32	11,21	GCACTTGCAGCTTAACTT	CGTGACTGTCAAACCG	base MSAT
MSAT3.58	18,66	GGTTGAAACTCTATGGTAT	CGCATTTCTTGATTTACTA	base MSAT
MSAT 3.18782A	18,76	CACACAACCTTGGCTTAGGA	AGTTGGGAAAGTTGTGGTTGA	NG
MSAT 3.18776A	18,77	TTCTTGCAACAACAAAAGGA	GGGCGTTACAAGAGAATGGA	NG
F24M12-TGF	19,07	GTTCTCTGCATTCCACACATACTCT	CTTGGGTATTCTGAAGAGCATAAAT	base MSAT
MSAT 3.19208A	19,21	TGGGCCACTCTGTATTGTCA	TAGGCTTGCAGATTGGTTCA	NG
MSAT 3.19575A	19,57	TGTTCAAACCTGCATGGAA	GAAGGCAAGGCATTGTTTACT	NG
MSAT 3.19789A	19,79	CGGTGTCTTCAACCTGCTATT	AAGCAGGACCAACCAAGAAA	NG
MSAT 3.199222	19,92	ACGCGATAAAAAGGTCACCA	TAACGATCGCTCGTGATTT	NG
MSAT 3.19974A	19,97	AGCGGAGGAAGGAGATGATT	CGGTTTATTCAGTTATGGTTGC	NG
MSAT3.20078	20,08	GTATTTTTCATCCCTACTCTCT	GATGCTTCAGTTCAATCTGG	CW
MSAT3.20329A	20,32	CCTAACTTTCGTCTCCCAACC	TTCAGGGACAAGTCTGAGAGG	NG
MSAT 3.20377B	20,38	CATTCTCTGCACGCGAGTTA	ACGTGTCACTCGTTCATTCC	NG
MSAT3.20405	20,41	AGGAGGAGGAAGTTGCAGTA	GAGAGAAAGATGAAAGAGACCT	CW
MSAT3.20435	20,43	TCCACTCTTGTTCGTGATTTTC	CATTATTAGTGTCTCGTATCGTA	CW
MSAT3.28	20,47	TACAAGTCATAATAGAGGC	GGGTTTAGCATTTAGC	base MSAT
MSAT3.206422	20,64	CGGACTCATGACGACAGTCT	ACGCGTGTGTCCATAACTT	NG
MSAT3.206563	20,65	CAAATCCATCTTCCACCAC	CCTCCCCAAGTATCCTTTT	NG
MSAT3.18	21,39	TACCTCAAAGAGCAAACA	TCATACCTACATATTGCCCT	base MSAT
ATHFUS6	22,64	TCGTTACACTGGCTTGCCTTG	TTCCCTGATCAGATTGGTCCG	base MSAT
MSAT3.22714	22,71	TGAAGCTCGTCTTCGTTTTT	TGACGACTTTGGAAATTGAAA	EW
MSAT3.22953	22,95	GAGATGGGCAGAGACAAAAA	AACAATCAATGCATCAGCAA	EW
MSAT3.23007	23,00	AAAAGCGTGATTGTGTTGTG	TGGAATGTTCCATCTTCAACT	EW
MSAT3.23388	23,39	GCGTACAAAAATGGGCAAAA	GAGATTTTCTTGTCTCAATTCC	base MSAT
MSAT3.70	23,45	CACCCATAGACACACAACC	TCAAAACCTCCAACCATAC	base MSAT

**Tableau C3 : oligonucléotides utilisés pour amplifier les MSAT localisés au sud du chromosome 3**  
Ces couples d'amorces amplifient des marqueurs MSAT polymorphes entre les allèles parentaux Sha et Mr-0. Dans la dernière colonne, base MSAT : collection de marqueurs développés pour amplifier des MSAT chez *A. thaliana* (la collection est mise à jour quotidiennement par les équipes de la SGAP) ; NG : Nicolas Gobron ; CW : Cezary Waszczak. EW : Estelle Wenes.

	Nom de l'amorce	Position en Mb	Séquence de l'amorce (5'→3')
a	CIW1.1F	18,37	TGGGTAAAAGATGACCGAGTCT
	CIW1.1R		AACCCCCACTTTTACTTCTCAT
	ATHGENEA.F	22,40	ACCATGCATAGCTTAAACTTCTT
	ATHGENEA.R		ACATAACCACAAATAGGGGTG
b	MSAT3.19974A.2-L	19,97	ATTCCCTCTCTCCGTTTCGT
	MSAT3.19974A.2-R		TCGGTTTATTTCAGTTATGGTTGC
	MSAT3.206422.3-L	20,62	ACTTTCGCTAATACTGAATCCAA
	MSAT3.206422.3-R		ACGCGTGTGTCCATAACTT

**Tableau C4 : oligonucléotides utilisés pour amplifier les MSAT en multiplex**  
a – couples d'amorces utilisés pour amplifier deux MSAT du chromosome 1  
b – couples d'amorces utilisés pour amplifier deux MSAT du chromosome 3

Nom de l'amorce	Séquence de l'amorce (5'→3')	Enzyme de restriction
At3g62440L20	CGCCTGCACAACGATAGTAA	AclI
At3g62440U20	GGATCGGATCAGCAATCTTC	
At3g62560L21	CAGTGGTGAAACTGGTGAGGC	ClaI
At3g62560U21	ATGTTCTGTGGATTGGTTC	
At3g62590L20	CTCTCGGATGGTTGCAGAAT	XhoI
At3g62590U20	CAACCCTGCTGTTGCTTACA	
At3g62620L20	TTGGCTTGGAGAATGAAACC	Bsp1407I
At3g62620U20	TCCATGGAAGACACAGCTCA	
At3g62700L20	CTCTTGAACCAATGCAGCAA	SspI
At3g62700U20	TTCGAAGGAACCGTAAGGAG	
At3g62790L21	TGTTTGACGAGTAGCCAAAGG	DraI
At3g62790U20	CTGGATGGATTTTCAGCGAAT	

**Tableau C5 : oligonucléotides utilisés pour amplifier les marqueurs CAPS au sud du chromosome 3**

Dans la première colonne, le nom de l'amorce contient le nom du gène où un polymorphisme SNP existe entre les allèles Sha et Mr-0. Ce polymorphisme est exploité grâce aux enzymes de restriction indiquées dans la dernière colonne.

Nom de l'amorce	Séquence de l'amorce (5'->3')
IV-3b	GCAATCTACCGGGAAGAACA
IV-5a	TTCATCGCTTAGCTGTGCTG
V-4b	CAAGCCCGCTAAAGTAGTGC
V-6a	AAGATCCAACAACCCGTGAG
VI-5b	ATGTCCAGAGAGGGACATCG
VI-7a	CTGCCATGTTTCCATCATTG
VII-6b	GGATTTACGAGCTGGGATCA
VII-8a	GATAATCACAAGGCCGAGCTA
VIII-7b	CCTTAACCTGTCCGTGCCTA
VIII-9a	GAAAGACCACGCTTGAGGAG
IX-10a	TGTCGCTGTGGAGATCAAAG
IX-8b	TGCCAATTTTCAAACCATCA
X10-b	TCAAGCTTTTGCACCTTGTG
X-0h	CGCGCTAAAGACATCATCAA

**Tableau C6 : oligonucléotides utilisés pour amplifier les fragments de 6000 à 8000 pb au sud du chromosome 3**  
 Les chiffres romains dans la première colonne correspondent aux numéros des fragments amplifiés d'après la figure B18.

Nom de l'amorce	Séquence de l'amorce (5'->3')
KAtL22	TAActCAACGGTAGAGTACTCG
MatKAtU22	AGCGTTTCGTAATTAGTGAAct
ndhFU25	GTCGGAAAAAGGAGAAGTCCAActC
rbcLAtU22	TGAAATTATCCGTGAAGCTTGC
accDAAtL22	CATATATAAAACGGTCTTGACT
rpl32L25	CGCTTTTTCGAGATAGAAGTACGTT
trnVL23	CTCTTTTACCGAGCAGGTCTACG
ndhCU23	AAATGCCAAAACAGGAATAGCAC

**Tableau C7 : oligonucléotides utilisés pour amplifier les régions intergéniques du génome chloroplastique**  
 D'après Moison et al., 2010.

Nom de l'amorce	Séquence de l'amorce (5' → 3')	Utilisation
Aimt:11944U21	GTCCAAAGAGCGCAGACTAGC	Amplification de la région nad5-rm26 region pour détecter l'événement de recombinaison Rec1
Aimt:13775L21	CATGGGATATGATCACAAGAG	Amplification de la région nad5-rm26 region pour détecter l'événement de recombinaison Rec1
Aimt:16606U21	CTACCGAATTCGCCACTGACTA	Amplification de la région nad5-rm26 region pour détecter l'événement de recombinaison Rec2
Aimt:18571L21	AACCACATTTTCATTCAGGAGC	Amplification de la région nad5-rm26 region pour détecter l'événement de recombinaison Rec2
Aimt:24085U21	GTTTGTTCGCTCTGCTTGT	Amplification de la région ccmC
Aimt:241468L21	AAGGAAAGTGGCTCAAAACGC	Amplification de la région ccmC
Aimt:278985U21	GCTGCTATCGGTTATGGAAAC	Amplification du réarrangement atp9-cox3
Aimt:217884L21	GTATTAAACCTTCCCTCCAC	Amplification du réarrangement atp9-cox3
Aimt:297337U21	CGGAAAGACCACCTATTTGT	Amplification spécifique du gène atp6-2
Aimt:298263L21	GGGTACGGGCGCTCTATCAT	Amplification spécifique du gène atp6-2
Aimt:24147U20	CTATCCCATGTTCCGATCT	Amplification de la région intergénique nad9-rpl16
Aimt:25290L20	CGAAGAAATGGTAAGATATG	Amplification de la région intergénique nad9-rpl16
Aimt:130231U20	GAGAAATAAGTGGCTCACGA	Amplification de la région intergénique atp8-orf107c
Aimt:131105L21	CAATCCATCTTTGATACTCC	Amplification de la région intergénique atp8-orf107c
Aimt:39529U21	GAATAGCTATCCCGTCGTACT	Amplification de la région flanquante en aval de la région cox2
Aimt:40354L21	TGCTACGAAAGCTCCGCTCCC	Amplification de la région flanquante en aval de la région cox2
Aimt:209506U17	ATTTGGTTTCAACCACG	Amplification de orf107d
Aimt:209769L21	GTCAAATAAGATCCTGGTGTG	Amplification de orf107d
Aimt:301020U20	GATATCCCTGGCACACACAGA	Amplification de orf111b
Aimt:301254L19	TTTCTCGATTAATGCCTG	Amplification de orf111b
Aimt:207558U20	CTTTATCTAGGTGGCTT	Amplification de orf120
Aimt:207849L21	GTGGACTTCCCAAGACTTTCA	Amplification de orf120
Aimt:16965U19	CTCTTACGGTCACTATAT	Amplification de orf131
Aimt:17187L17	GATCAACCATTTGGTGA	Amplification de orf131
Aimt:203703U18	ATATCTATCAACGCCCTGA	Amplification de orf240a
Aimt:204244L21	ATTTGGAGAGTACACGGAAAC	Amplification de orf240a
Aimt:279543U21	CGTTCCCGTTTCAATCTTAT	Amplification de orf262
Aimt:280150L20	GCGAGCTGTTCCGAAAGTTGC	Amplification de orf262
Aimt:16959U19	AAATGCCCTCTTACGGTCAT	Amplification de orf315
Aimt:17436L16	CCCTGTTGGTACCCTCC	Amplification de orf315

**Tableau C8 : oligonucléotides utilisés pour typer l'ADN mitochondrial**

Ces couples d'amorces ont été mis au point par Moison et al., 2010.

Nom de l'amorce	Séquence de l'amorce (5'->3')	Informations
Atrmt:28529U21 (=U')	CCAGTGGTCTTCTAGCGTTAG	PCR spécifique de la copie trnK située entre rps3 et ccb206
Atrmt:29430L21 (=L')	GTCAATCCGGCCCCGAGACGCA	
Atrmt:204001U21 (=U)	GCCCCAAACAAAATGTATTA	PCR spécifique de la copie trnK située entre l'orf240a et l'orf120
Atrmt:204831L21 (=L)	GTGATGCCCTTCCTGATTA	

**Tableau C9 : oligonucléotides utilisés pour amplifier spécifiquement chaque copie du gène trnK du génome mitochondrial d'*A.thaliana***

Ces couples d'amorces ont été dessinés par Françoise Budar. La position de ces oligonucléotides sur le génome mitochondrial d'*Arabidopsis* est indiquée dans la figure B27.

# **BIBLIOGRAPHIE**

- Abbott RJ, Gomes MF. Population genetic structure and outcrossing rate of *Arabidopsis thaliana* (L.) Heynh. *Heredity* (1989) 62 : 411–418.
- Adamiec M, Drath M, Jackowski G. Redox state of plastoquinone pool regulates expression of *Arabidopsis thaliana* genes in response to elevated irradiance. *Acta Biochim Pol* (2008) 55 : 161–173
- Adams and Palmer. Evolution of mitochondrial gene content: gene loss and transfer to the nucleus. *Mol Phylogenet Evol* (2003) 29 (3) : 380-395
- Adams, K.L., Daley, D.O., Qiu, Y.-L., Whelan, J. and Palmer, J.D. Repeated and recent transfers of a mitochondrial gene to the nucleus with diverse activation mechanisms. *Nature* (2000) 408, 354±357.
- Alboresi A, Dall'Osto L, Aprile A, Carillo P, Roncaglia E, Cattivelli L, Bassi R. Reactive oxygen species and transcript analysis upon excess light treatment in wild-type *Arabidopsis thaliana* vs a photosensitive mutant lacking zeaxanthin and lutein. *BMC Plant Biology* (2011) 11, 62.
- Alexander MP (1969) Differential staining of aborted and nonaborted pollen. *Stain Technol* 44 117–122.
- Allen JO, Fauron CM, Minx P, Roark L, Oddiraju S, Lin GN, Meyer L, Sun H, Kim K, Wang C, Du F, Xu D, Gibson M, Cifrese J, Clifton SW, Newton KJ. Comparisons Among Two Fertile and Three Male-Sterile Mitochondrial Genomes of Maize. *Genetics* (2007) 177 (2) : 1173-1192
- Allen. Effect of Teosinte Cytoplasmic Genomes on Maize Phenotype. *Genetics* (2005) 169 (2) : 863-880.
- Allison. The role of sigma factors in plastid transcription. *Biochimie* (2000) 82 : 537–548
- Alonso, J.M. and J.R. Ecker. Moving forward in reverse: genetic technologies to enable genome-wide phenomic screens in *Arabidopsis*. *Nat. Rev. Genet.* (2006) 7:524-
- Alonso-Blanco and Koornneef. Naturally occurring variation in *Arabidopsis*: an underexploited resource for plant genetics. *trends in plant science Reviews* (2000) 5 (1) : 1360
- Al-Shehbaz, I.A. et S.L. O\_Kane, Jr. Taxonomy and phylogeny of *Arabidopsis* (Brassicaceae). (2002) In C.R. Somerville and E.M Meyerowitz (eds.), *The Arabidopsis Book*, doi/10.119/tab.001,
- Aluru MR, Zola J, Foudree A, Rodermel SR. Chloroplast photooxidation-induced transcriptome reprogramming in *Arabidopsis immutans* white leaf sectors. *Plant Physiol* (2009) 150: 904–92
- Alverson AJ, Wei X, Rice DW, Stern DB, Barry K, Palmer JD. Insights into the Evolution of Mitochondrial Genome Size from Complete Sequences of *Citrullus lanatus* and *Cucurbita pepo* (Cucurbitaceae). *Molecular Biology and Evolution* (2010) 27 (6) : 1436-1448
- Alverson AJ, Zhuo S, Rice DW, Sloan DB, Palmer JD. The mitochondrial genome of the legume *Vigna radiata* and the analysis of recombination across short mitochondrial repeats. *PLoS One* (2011) 6(1):e16404.
- Andre, C., Levy, A., and Walbot, V. Small repeated sequences and the structure of plant mitochondrial genomes. *Trends Genet.* (1992) 8: 128–132
- Andrès C , Lurin C, Small ID . The multifarious roles of PPR proteins in plant mitochondrial gene expression. *Physiol Plant* (2007) 129 (1) : 14-22
- Annemarie Matthes, Stephanie Schmidt-Gattung, Daniela Köhler, Joachim Forner, Steffen Wildum, Monika Raabe, Henning Urlaub and Stefan Binder. Two DEAD-Box Proteins May Be Part of RNA-Dependent High-Molecular-Mass Protein Complexes in *Arabidopsis* Mitochondria. *Plant Physiology* (2007) 145:1637-1646
- Arrieta-Montiel MP, Shedge V, Davila J, Christensen AC, Mackenzie SA. Diversity of the *Arabidopsis* mitochondrial genome occurs via nuclear-controlled recombination activity. *Genetics* (2009) 183 (4) : 1261-8
- Asakura and Barkan. A CRM domain protein functions dually in group I and group II intron splicing in land plant chloroplasts. *Plant Cell* (2007) 19: 3864-75.
- Asakura and Barkan. *Arabidopsis* orthologs of maize chloroplast splicing factors promote splicing of orthologous and species-specific group II introns. *Plant Physiol.* (2006) 142: 1656–1663.
- Asakura Y, Bayraktar OA, Barkan A. Two CRM protein subfamilies cooperate in the splicing of group IIB introns in chloroplasts. *RNA.* (2008) 14(11):2319-32.
- Ashman T-L (2003). Constraints on the evolution of dioecy and sexual dimorphism: Field estimates of quantitative genetic parameters for reproductive traits in three populations of gynodioecious *Fragaria virginiana*. *Evolution*, 57, 2012–25.
- Ashutosh, Kumar P, Dinesh Kumar V, Sharma PC, Prakash S, Bhat SR. (2008). A novel orf108 co-transcribed with the *atpA* gene is associated with cytoplasmic male sterility in *Brassica juncea* carrying *Moricandia arvensis* cytoplasm. *Plant & Cell Physiology* 49: 284 – 289.
- Atwell S, Huang YS, Vilhjálmsson BJ, Willems G, Horton M, Li Y, Meng D, Platt A, Tarone AM, Hu TT, Jiang R, Mulyati NW, Zhang X, Amer MA, Baxter I, Brachi B, Chory J, Dean C, Debieu M, de Meaux J, Ecker JR, Faure N, Kniskern JM, Jones JD, Michael T, Nemri A, Roux F, Salt DE, Tang C, Todesco M, Traw MB, Weigel D, Marjoram P, Borevitz JO, Bergelson J, Nordborg M. Genome-wide association study of 107 phenotypes in *Arabidopsis thaliana* inbred lines, *Nature* (2010) 465:627-31.
- Aubourg S, Boudet N, Kreis M, Lecharny A. In *Arabidopsis thaliana*, 1% of the genome codes for a novel

- protein family unique to plants. *Plant Mol Biol* (2000) 42 (4) : 603-13.
- Azhagiri et Maliga. Exceptional paternal inheritance of plastids in *Arabidopsis* suggests that low-frequency leakage of plastids via pollen may be universal in plants. *The Plant Journal* (2007) 52 (5) : 817-823.
- Babiychuk E, Kushnir S, Belles-Boix E, Van Montagu M, Inzé D. *Arabidopsis thaliana* NADPH oxidoreductase homologs confer tolerance of yeasts toward the thiol-oxidizing drug diamide. *J Biol Chem* (1995) 270: 26224–26231.
- Bailey and Delph. Sex-Ratio Evolution in Nuclear-Cytoplasmic Gynodioecy When Restoration Is a Threshold Trait. *Genetics* (2007) 176 (4) : 2465-2476.
- Bailey MF, Delph LF. A field guide to models of sex-ratio evolution in gynodioecious species. (2007) *Oikos*, 116, 1609–1617.
- Bailey MF, Delph LF, Lively CM Modeling gynodioecy: novel scenarios for maintaining polymorphism. (2003). *Am Nat* 161: 762–776.
- Bailey PC, Martin C, Toledo-Ortiz G, Quail PH, Huq E, Heim MA, Jakoby M, Werber M, Weisshaar B. Update on the basic helix-loop-helix transcription factor gene family in *Arabidopsis thaliana*. *Plant Cell*. (2003) 15 (11) : 2497-502.
- Bailey-Serres and Chang. Sensing and signalling in response to oxygen deprivation in plants and other organisms. *Ann Bot* (2005) 96 (4) : 507-18.
- Balasubramanian S, Schwartz C, Singh A, Warthmann N, Kim MC, Maloof JN, Loudet O, Trainer GT, Dabi T, Borevitz JO, Chory J, Weigel D. QTL Mapping in New *Arabidopsis thaliana* Advanced Intercross-Recombinant Inbred Lines. *PLoS ONE* (2009) 4 (2) : e4318.
- Bannerot H., Bouldard L., Cauderon Y., Tempe J., (1974) Cytoplasmic male sterility transfer from *Raphanus* to *Brassica*. Proc. Eucarpia Meeting Cruciferae Vegetable. Crops Sect., 25-27th Sept. 1974,52-54.
- Barkan A, Walker M, Nolasco M and Johnson D (1994) A nuclear mutation in maize blocks the processing and translation of several chloroplast mRNAs and provides evidence for the differential translation of alternative mRNA forms. *EMBO J*, 13, 3170–3181.
- Barkan, A. Intron splicing in plant organelles. In *Molecular Biology and Biotechnology of Plant Organelles*, H. Daniell and C. Chase, eds (Dordrecht, The Netherlands: Kluwer Academic Publishers) (2004) : 281–308.
- Barr and Fishman. The nuclear component of a cytonuclear hybrid incompatibility in *Mimulus* maps to a cluster of pentatricopeptide repeat genes. *Genetics* (2010) 184 (2) : 455-65
- Barrett, S.C.H. (2002). The evolution of plant sexual diversity. *Nature Reviews Genetics* 3: 274–284.
- Barrientos, A., Kenyon, L. & Moraes, C. T. (1998) Human xenomitochondrial cybrids. Cellular models of mitochondrial complex I deficiency. *J. Biol. Chem.* 273, 14 210–14 217.
- Baumgartner BJ, Rapp JC, Mullet JE. Plastid genes encoding the transcription/translation apparatus are differentially transcribed early in barley (*Hordeum vulgare*) chloroplast development: evidence for selective stabilization of psbA mRNA. *Plant Physio* (1993) 101 : 781–791.
- Baur. Das Wesen und die Erblichkeitsverhältnisse der « Varietates albomarginatae hort von *Pelargonium zonale* ». *Ztschr f ind Abstamm u Vererbungsl* (1909) 1 : 330-351.
- Beardslee T, Chowdhury SR, Chang CC, Jaiswal P, Lerbs-Mache S, Stern D, Allison LA (2002) A nucleus-encoded maize protein with sigma factor activity accumulates in both chloroplasts and mitochondria. *Plant J* 31: 199–209.
- Bechsgaard, J. S., V. Castric, D. Charlesworth, X. Vekemans and M. H. Schierup, (2006) The transition to self-compatibility in *Arabidopsis thaliana* and evolution within S-haplotypes over 10 Myr. *Mol. Biol. Evol.* 23: 1741–1750.
- Beck JB, Schmutz H, Schaal BA. Native range genetic variation in *Arabidopsis thaliana* is strongly geographically structured and reflects Pleistocene glacial dynamics. *Mol Ecol* (2008) 17 (3) : 902-15.
- Becker C, Hagmann J, Müller J, Koenig D, Stegle O, Borgwardt K, Weigel D (2011) Spontaneous epigenetic variation in the *Arabidopsis thaliana* methylome. *Nature* 480: 245–249
- Beick, S., Schmitz-Linneweber, C., Williams-Carrier, R., Jensen, B., and Barkan, A. The pentatricopeptide repeat protein PPR5 stabilizes a specific tRNA precursor in maize chloroplasts. *Mol. Cell. Biol.* (2008) 28: 5337–5347.
- Bellaoui M, Pelletier G, Budar F. The steady-state level of mRNA from the Ogura cytoplasmic male sterility locus in *Brassica* cybrids is determined post-transcriptionally by its 3' region. *EMBO J* (1997) 16 (16) : 5057-68.
- Belliard, G., Vedel, F. and Pelletier, G. (1979). Mitochondrial recombination in cytoplasmic hybrids of *Nicotiana tabacum* by protoplast fusion. *Nature, London* 281,401-403.
- Bendich and Drlica (2000), Prokaryotic and eukaryotic chromosomes: what's the difference? *Bioessays* 22: 481-486.
- Bendich. Circular chloroplast chromosomes: the grand illusion. *Plant Cell* (2004) 16 (7) : 1661-6.

- Benne R, Van den Burg J, Brakenhoff JP, Sloof P, Van Boom JH, Tromp MC. Major transcript of the frameshifted *coxII* gene from trypanosome mitochondria contains four nucleotides that are not encoded in the DNA. *Cell*, (1986) 46 : 819–826.
- Bentolila S, Alfonso AA, Hanson MR. A pentatricopeptide repeat-containing gene restores fertility to cytoplasmic male-sterile plants. *Proc Natl Acad Sci USA* (2002) 99 (16) : 10887-92.
- Biehl A, Richly E, Noutsos C, Salamini F, Leister D (2005) Analysis of 101 nuclear transcriptomes reveals 23 distinct regulons and their relationship to metabolism, chromosomal gene distribution and co-ordination of nuclear and plastid gene expression. *Gene* 344: 33–41.
- Bikard D, Patel D, Le Mette C, et al. 2009. Divergent evolution of duplicate genes leads to genetic incompatibilities within *A. thaliana*. *Science* 323: 623 – 626.
- Bisanz-Seyer C, Li YF, Seyer P and Mache R. The components of the plastid ribosome are not accumulated synchronously during the early development of spinach plants. *Plant Mol Biol* (1989) 12 : 201-211.
- Bläske-Lietze et Boucias. Pathogenesis of *Helicosporidium* sp. (Chlorophyta: Trebouxiophyceae) in susceptible noctuid larvae. *Journal of Invertebrate Pathology* (2005) 90 : 161–168.
- Blier PU, Dufresne F, Burton RS. Natural selection and the evolution of mtDNA-encoded peptides: evidence for intergenomic co-adaptation. *TRENDS in Genetics* (2001) 17 (7) : 400-406.
- Boavida L. C., Vieira A. M., Becker J.D and Feijo J.A, (2005) Gametophyte interaction and sexual reproduction: how plants make a zygote. *Int J Dev Biol* 49: 615-632.
- Boeshore, M.L., Hanson, M.R., and Izhar, S. (1985). A variant mitochondrial DNA arrangement specific to *Pefunia* stable sterile somatic hybrids. *Plant MOI. Biol.* 4, 125-132. 132.
- Boeshore, M.L., Lifshitz, I., Hanson, M.R., and Izhar, S. (1983). Novel composition of mitochondrial genomes in *Pefunia* somatic hybrids derived from cytoplasmic male sterile and fertile plants. *MOI. Gen. Genet.* 190, 459-467.
- Bogdanova VS, Galieva ER, Kosterin OE. Genetic analysis of nuclear-cytoplasmic incompatibility in pea associated with cytoplasm of an accession of wild subspecies *Pisum sativum* subsp. *elatius* (Bieb.) Schmahl. *Theor Appl Genet* (2009) 118 (4) : 801-809.
- Boggs, N.A., J.B. Nasrallah, and M.E. Nasrallah. (2009). Independent S-locus mutations caused self-fertility in *Arabidopsis thaliana*. *PLoS Genet.* 5: e1000426.
- Bogorad. Evolution of organelles and eukaryotic genomes. (1975) *Science* 188:891- 898.
- Bollenbach TJ, Lange H, Gutierrez R, Erhardt M, Stern DB, Gagliardi D (2005) RNR1, a 30 -50 exoribonuclease belonging to the RNR superfamily, catalyzes 30 maturation of chloroplast ribosomal RNAs in *Arabidopsis thaliana*. *Nucleic Acids Res* 33:2751–2763
- Bombliès, K., and Weigel, D. (2007) Hybrid necrosis: Autoimmunity as a potential gene flow barrier for plant species. *Nature Reviews Genetics* 8, pp. 382-389.
- Bombliès and Weigel. *Arabidopsis* and relatives as models for the study of genetic and genomic incompatibilities. *Philosophical Transactions of the Royal Society B: Biological Sciences* (2010) 365 (1547) : 1815-1823.
- Bombliès, K., Lempe, J., Epple, P., Warthmann, N., Lanz, C., Dangl, J. L., and Weigel, D. (2007) Autoimmune response as a mechanism for a Bateson-Dobzhansky-Muller-type incompatibility syndrome in plants. *PLoS Biol.* 5, e236.
- Bombliès K, Yant L, Laitinen RA, Kim ST, Hollister JD, Warthmann N, Fitz J, Weigel D. Local-Scale Patterns of Genetic Variability, Outcrossing, and Spatial Structure in Natural Stands of *Arabidopsis thaliana*. *PLoS Genet* (2010) 6 (3) : e1000890
- Bonen and Vogel. The ins and outs of group II introns. *TRENDS in Genetics* (2001) 17 (6) : 322-331.
- Bonen L.(1993) Trans-splicing of pre-mRNA in plants, animals, and protists. *FASEB J.* 7:40–46.
- Bonen, L. Cis- and trans-splicing of group II introns in plant mitochondria. *Mitochondrion* (2008) 8: 26-34,.
- Bonhomme S, Budar F, Férault M, Pelletier G (1991). A 2.5 kb NcoI fragment of *Ogura* radish mitochondrial DNA is correlated with cytoplasmic male-sterility in *Brassica* cybrids. *Curr Genet* 19: 121-127.
- Bonhomme S, Budar F, Lancelin D, Small I, Defrance M-C, Pelletier G (1992). Sequence and transcript analysis of the Nco2.5 *Ogura*-specific fragment correlated with cytoplasmic male sterility in *Brassica* cybrids. *Mol Gen Genet* 235: 340-348.
- Borg, M., Brownfield, L., and Twell, D. (2009) Male gametophyte development: A molecular perspective. *J. Exp. Bot.* 60: 1465–1478
- Boudet N, Aubourg S, Toffano-Nioche C, Kreis M, Lecharny A (2001) Evolution of intron/exon structure of DEAD helicase family genes in *Arabidopsis*, *Caenorhabditis*, and *Drosophila*. *Genome Res* 11: 2101–2114
- Bowsher et Tobin. Compartmentation of metabolism within mitochondria and plastids. *Journal of Experimental Botany* (2001) 52 : 513-527.
- Branco-Price. Genome-wide Analysis of Transcript Abundance and Translation in *Arabidopsis* Seedlings Subjected to Oxygen Deprivation. *Annals of Botany* (2005) 96 (4) : 647-660

- Brangeon J, Sabar M, Gutierrez S, Combettes B, Bove J, Gendy C, Chetrit P, des Francs-Small CC, Pla M, Vedel F, et al. (2000) Defective splicing of the first *nad4* intron is associated with lack of several complex I subunits in the *Nicotiana sylvestris* NMS1 nuclear mutant. *Plant J* 21:269–280
- Brouard J-S., Otis C., Lemieux C. and Turmel M.. The Exceptionally Large Chloroplast Genome of the Green Alga *Floeydiella terrestris* Illuminates the Evolutionary History of the Chlorophyceae. *Genome Biology and Evolution* (2010) 2 (0) : 240-256
- Brown S, Hu N, Hombria JC. Novel level of signalling control in the JAK/STAT pathway revealed by in situ visualisation of protein-protein interaction during *Drosophila* development. *Development* (2003) 130 (14) : 3077-3084
- Brown. Unique aspects of cytoplasmic male sterility and fertility restoration in *Brassica napus* (1999) *The American Genetic Association* 90:351–356.
- Budar and Roux. The role of organelle genomes in plant adaptation: Time to get to work!. *Plant signaling & behavior* (2011) 6 (5) :
- Budar F and Pelletier G (2001) "Male sterility in plants : occurrence, determinism, significance and use." *CR Acad Sci Paris, Sciences de la vie/Life Sciences* 324:1-8.
- Budar F, Touzet P, De Paepe R. The nucleo-mitochondrial conflict in cytoplasmic male sterilities revisited. *Genetica* (2003) 117 (1) : 3-16
- Budar, F., and G. Pelletier. (2001). Male sterility in plants: occurrence, determinism, significance and use. *C R Acad Sci III* 324:543-550.
- Bullerwell CE, Leigh J, Forget L, Lang BF. A comparison of three fission yeast mitochondrial genomes. *Nucl. Acids Res.* (2003) 31 (2): 759-768.
- Burger G, Gray MW, Lang BF. Mitochondrial genomes: anything goes. *Trends Genet* (2003) 19 (12) : 709-16
- Burt and Trivers. *Genes in conflict: the biology of selfish genetic elements* (2006) Harvard University Press.
- Burton, R.S., P.D. Rawson, and S. Edmands. (1999). Genetic architecture of physiological phenotypes: empirical evidence for coadapted gene complexes. *Am. Zool.* 39:451–462. cumulation in organelle genomes. *Genetica* 103:29–39.
- Burton, R.S.. Hybrid breakdown in developmental time in the copepod *Tigriopus californicus*. *Evolution*, (1990) 44: 1814-1822.
- Burton, R.S. and B.N. Lee.. Nuclear and mitochondrial gene genealogies and allozyme polymorphism across a major phylogeographic break in the copepod *Tigriopus californicus*. *Proc. Natl. Acad. Sci. USA*, (1994) 91: 5197-5201
- Butow and Avadhani. Mitochondrial Signaling:: The Retrograde Response. *Mol Cell* (2004) 14 (1) : 1-15
- Butow, R. A., and N. G. Avadhani, Mitochondrial signalling: the retrograde response. *Mol. Cell* (2004) 14: 1–15
- Cameron, D. R., and R. **Moav**, (1957) Inheritance in *Nicotiana tabacum* XXVII. Pollen killer, an alien genetic locus inducing abortion of microspores not carrying it. *Genetics* **42**: 326- 335.
- Cahoon and Stern Plastid transcription: a *menage à trois*? *Trends in Plant Science*. (2001) 6:45-46.
- Canino G, Bocian E, Barbezier N, Echeverria M, Forner J, Binder S, Marchfelder A (2009) *Arabidopsis* encodes four tRNase Z enzymes. *Plant Physiol* 150: 1494–1502
- Cao J, Schneeberger K, Ossowski S, Günther T, Bender S, Fitz J, Koenig D, Lanz C, Stegle O, Lippert C, Wang X, Ott F, Müller J, Alonso-Blanco C, Borgwardt K, Schmid KJ, Weigel D. Whole-genome sequencing of multiple *Arabidopsis thaliana* populations. *Nature Genetics* (2011) 43 : 956–963
- Case and Willis. Hybrid male sterility in *Mimulus* (Phrymaceae) is associated with a geographically restricted mitochondrial rearrangement. *Evolution* (2008) 62 (5) : 1026-1039
- Castandet B, Choury D, Bégu D, Jordana X, Araya A. Intron RNA editing is essential for splicing in plant mitochondria. *Nucleic Acids Res.* (2010) 38:7112–21
- Chandler, V. L., (2007) Paramutation: from maize to mice. *Cell* 128: 641-645.
- Charlesworth B. (1978) Model for evolution of Y chromosomes and dosage compensation. *Proc. Natl Acad. Sci. USA* 75, 5618–5622.
- Charlesworth, D. (1981). A further study of the problem of the maintenance of female in gynodioecious species. *Heredity*, 46, 27–39
- Charlesworth D. and Vekemans X. (2005) "How and when did *A. thaliana* become highly self-fertilising". *BioEssays*, 27: 472-476
- Chase. Cytoplasmic male sterility: a window to the world of plant mitochondrial-nuclear interactions. *TRENDS in Genetics* (2007) 23 (2) : 81-90
- Chateigner-Boutin AL, des Francs-Small CC, Delannoy E, Kahlau S, Tanz SK, de Longevialle AF, Fujii S, Small I. OTP70 is a pentatricopeptide repeat protein of the E subgroup involved in splicing of the plastid transcript rpoC1. *The Plant Journal* (2011) 65 (4) : 532-542
- Chateigner-Boutin and Small. A rapid high-throughput method for the detection and quantification of RNA

- editing based on high-resolution melting of amplicons. *Nucleic Acids Research* (2007) 35 : e114
- Cheng MY, Hartl FU, Martin J, Pollock RA, Kalousek F, Neupert W, Hallberg EM, Hallberg RL, Horwich AL. *Mitochondrial heat-shock protein hsp60 is essential for assembly of proteins imported into yeast mitochondria* [archive], *Nature*, (1989) 337:620–625.
- Cho, R. J., Campbell, M. J., Winzeler, E. A., Steinmetz, L., Conway, A., Wodicka, L., Wolfsberg, T. G., Gabrielian, A. E., Landsman, D., Lockhart, D. J., and Davis, R. W. (1998). A genome-wide transcriptional analysis of the mitotic cell cycle. *Mol Cell*, 2(1):65-73.
- Cho, Y., Qiu, Y.L., Kuhlman, P., and Palmer, J.D. (1998). Explosive invasion of plant mitochondria by a group I intron. *Proc. Natl. Acad. Sci. USA* 95, 14244–14249
- Chomezynski, P. and Sacchi, N. (1987) Single-step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction. *Analytical Biochemistry* 162, 156±159.
- Chou JY, Hung YS, Lin KH, Lee HY, Leu JY. Multiple Molecular Mechanisms Cause Reproductive Isolation between Three Yeast Species. *Plos Biol* (2010) 8 (7) : e1000432
- Chou K C, Shen H B (2010)b. Plant-mPLoc: A Top-Down Strategy to Augment the Power for Predicting Plant Protein Subcellular Localization. *PLoS ONE*. 5 : e11335.
- Chumley TW, Palmer JD, Mower JP, Fourcade HM, Calie PJ, Boore JL, Jansen RK. The complete chloroplast genome sequence of *Pelargonium x hortorum*: organization and evolution of the largest and most highly rearranged chloroplast genome of land plants. *Mol Biol Evol* (2006) 23 : 2175–2190
- Chun L, Kawakami A, Christopher DA. Phytochrome A mediates blue light and UV-A-dependent chloroplast gene transcription in green leaves. *Plant Physiol* (2001) 125 : 1957-1966
- Chung HJ, Ferl RJ. *Arabidopsis* alcohol dehydrogenase expression in both shoots and roots is conditioned by root growth environment. *Plant Physiol*. (1999) 121(2):429-36.
- Clifton, R., Lister, R., Parker, K.L., Sappl, P.G., Elhafez, D., Millar, A.H., Day, D.A., and Whelan, J. (2005). Stress-induced co-expression of alternative respiratory chain components in *Arabidopsis thaliana*. *Plant Mol. Biol.* 58, 193–212.
- Coffin JM, Hughes SH, Varmus HE (1997). *Retroviruses*. Cold Spring Harbor Laboratory
- Coimbra S, Costa M, Jones B, Mendes MA, Pereira LG (2009) Pollen grain development is compromised in *Arabidopsis agp6 agp11* null mutants. *J Exp Bot* 60: 3133–3142
- Correns, C. Vererbungsversuche mit blass (gelb) grünen und buntblättrigen Sippen bei *Mirabilis*. *Urtica und Lunaria* (1909) 1 : 291–329.
- Costanzo, M.C., and Fox, T.D. (1990). Control of mitochondrial gene expression in *Saccharomyces cerevisiae*. *Annu. Rev. Genet.* 24, 91–113.
- Cushing, D.A., Forsthoefel, N.R., Gestaut, D.R., and Vernon, D.M. *Arabidopsis emb175* and other *ppr* knockout mutants reveal essential roles for pentatricopeptide repeat (PPR) proteins in plant embryogenesis. *Planta* (2005) 221, 424–436
- Darwin, C. R. (1877). *The various contrivances by which orchids are fertilised by insects*. London: John Murray. 2d edition.
- De Koning et Keeling. The complete plastid genome sequence of the parasitic green alga *Helicosporidium* sp. is highly reduced and structured. *BMC Biology* (2006) 4 (1) : 12
- de la Sierra-Gallay I, Zig L, Jamalli A, Putzer H (2008) Structural insights into the dual activity of RNase J. *Nat Struct Mol Biol* 15:206–212.
- de Longevialle A.F., Hendrickson L., Taylor N.L., Delannoy E., Lurin C., Badger M., Millar A.H., Small I. The pentatricopeptide repeat gene OTP51 with two LAGLIDADG motifs is required for the cis-splicing of plastid *ycf3* intron 2 in *Arabidopsis thaliana*. *Plant J.* (2008) 56: 157–168
- de Longevialle AF, Meyer EH, Andrés C, Taylor NL, Lurin C, Millar AH, Small ID. The pentatricopeptide repeat gene OTP43 is required for trans-splicing of the mitochondrial *nad1* Intron 1 in *Arabidopsis thaliana* (2007) *Plant Cell* 19: 3256–3265
- de Longevialle AF, Small ID, Lurin C. Nuclearly Encoded Splicing Factors Implicated in RNA Splicing in Higher Plant Organelles. *Molecular Plant* (2010) 3 (4) : 691-705
- Delannay X, Gouyon PH, Valdeyron G (1981). Mathematical study of the evolution of gynodioecy with cytoplasmic inheritance under the effect of a nuclear gene. *Genetics* 99: 169–181.
- Delannoy E, Fujii S, Colas des Francs-Small C, Brundrett M, Small I. Rampant gene loss in the underground orchid *Rhizanthella gardneri* highlights evolutionary constraints on plastid genomes. *Molecular Biology and Evolution* (2011) 28 (7) : 2077–2086
- Delourme R., Budar, F. (1999) Male sterility *In* *Biology of Brassica Ceonospecies*. C. Gomez-Campo (Ed). Elsevier Science , 185-216.
- Delourme R., Foisset N., Horvais R., Barret P., Champagne G., Cheung W.Y., Landry B.S., Renard M. (1998) Characterisation of the radish introgression carrying the *Rfo* restorer gene for the Ogu-INRA cytoplasmic male sterility in rapeseed (*Brassica napus* L.). *Theor Appl Genet* 97: 129-134.
- Delph LF, Touzet P, Bailey MF. (2007). Merging theory and mechanism in studies of gynodioecy. *Trends in*

- Ecology and Evolution 22: 17–24.
- Desloire S, Gherbi H, Laloui W, Marhadour S, Clouet V, Cattolico L, Falentin C, Giancola S, Renard M, Budar F, Small I, Caboche M, Delourme R, Bendahmane A. Identification of the fertility restoration locus, Rfo, in radish, as a member of the pentatricopeptide-repeat protein family. *EMBO Rep* (2003) 4 (6) : 588-594
- Desplanque B, Viard F, Bernard J, Forcioli D, Saumitou-Laprade P, Cuguen J, Van Dijk H. The linkage disequilibrium between chloroplast DNA and mitochondrial DNA haplotypes in *Beta vulgaris* ssp. *maritima* (L.): the usefulness of both genomes for population genetic studies. *Molecular Ecology* (2000) 9 : 141–154
- Dewey RE, Levings CS 3rd, Timothy DH. Novel recombinations in the maize mitochondrial genome produce a unique transcriptional unit in the Texas male-sterile cytoplasm. *Cell* (1986) 44 (3) : 439-49
- Dewey RE, Timothy DH, Levings CS. A mitochondrial protein associated with cytoplasmic male sterility in the T cytoplasm of maize. *Proc Natl Acad Sci USA* (1987) 84 (15) : 5374-8
- Ding YH, Liu NY, Tang ZS, Liu J, Yang WC. Arabidopsis GLUTAMINE-RICH PROTEIN23 Is Essential for Early Embryogenesis and Encodes a Novel Nuclear PPR Motif Protein That Interacts with RNA Polymerase II Subunit III. *The Plant Cell Online* (2006) 18 (4) : 815-830
- Djajanegara I, Finnegan PM, Mathieu C, McCabe T, Whelan J, Day DA (2002) Regulation of alternative oxidase gene expression in soybean. *Plant Mol Biol* 50: 735-742
- Dobzhansky, T. 1937. *Genetics and the Origin of Species*. New York: Columbia University Press. (2nd ed., 1941; 3rd ed., 1951)
- Dombrowski S, Brennicke A. and Binder S. 3' inverted repeats in plant mitochondrial mRNAs are processing signals rather than substrates generated by *in vitro* transcription. *EMBO J.* (1997) 16, 5069–5076.
- Doniwa Y, Ueda M, Ueta M, Wada A, Kadowaki K, Tsutsumi N. The involvement of a PPR protein of the P subfamily in partial RNA editing of an Arabidopsis mitochondrial transcript. *Gene* (2010) 454 (1-2) : 39-46
- Doyle JJ, Davis JI, Soreng RJ, Garvin D, Anderson MJ. Chloroplast inversions and the origin of the grass family (Poaceae). *Proc. Natl. Acad. Sci. USA* (1992) 89 : 7722–7726.
- Dreyfus and Régnier. The Poly (A) Tail of mRNAs:: Bodyguard in Eukaryotes, Scavenger in Bacteria. *Cell* (2002) 111 (5) : 611-613
- Drider D, Condon C. The continuing story of endoribonuclease III. *J Mol Microbiol Biotechnol.* (2004) 8(4):195-200.
- Duchêne A-M. and Maréchal-Drouard L.. The chloroplast derived trnW and trnM genes are not expressed in Arabidopsis mitochondria. *Biochem. Biophys. Res. Com.*, 285:1213-1216, 2001.
- Durand S, Bouché N, Perez Strand E, Loudet O, and Camilleri C. *Rapid Establishment of Genetic Incompatibility through Natural Epigenetic Variation.* *Current Biology* (2012), doi:10.1016/j.cub.2011.12.054
- Duroc Y, Gaillard C, Hiard S, Defrance MC, Pelletier G, Budar F. Biochemical and functional characterization of ORF138, a mitochondrial protein responsible for Ogura cytoplasmic male sterility in Brassicaceae. *Biochimie* (2005) 87 (12) : 1089-100
- Duroc, Y., S. Hiard, N. Vrielynck, S. Ragu and F. Budar (2009). The Ogura sterility-inducing protein forms a large complex without interfering with the oxidative phosphorylation components in rapeseed mitochondria. *Plant Mol Biol* 70(1-2) :123-37.
- Dyall SD., Mark T. Brown and Patricia J. Johnson. Ancient Invasions: From Endosymbionts to Organelles. *Science* (2004) 304 (5668) : 253-257.
- Dyall. Ancient Invasions: From Endosymbionts to Organelles. *Science* (2004) 304 (5668) : 253-257
- Eckardt. Chloroplast Intron Splicing Mechanisms. *Plant Cell* (2007) 19(2) : 3838
- Edmunds, S., (1999) Heterosis and outbreeding depression in interpopulation crosses spanning a wide range of divergence. *Evolution* 53: 1757-1768.
- Edmunds, S., and R. S. Burton. (1999) Cytochrome c oxidase activity in interpopulation hybrids of a marine copepod: a test for nuclear-nuclear or nuclear-cytoplasmic coadaptation. *Evolution* 53:1972–1978.
- Edqvist and Bergman. Nuclear identity specifies transcriptional initiation in plant mitochondria. *Plant Mol Biol* (2002) 49:59–68
- Ellis R.J. Chloroplast proteins: Synthesis, transport and assembly (1981) *Annu. Rev. Physiol.*, 32, 11-137.
- Ellison and Burton. Interpopulation hybrid breakdown maps to the mitochondrial genome. *Evolution* (2008) 62 (3) : 631-638
- Ellison CK, Burton RS (2006). Disruption of mitochondrial function in interpopulation hybrids of *Tigriopus californicus*. *Evolution* 60: 1382–1391.
- Ellison CK, Niehuis O, Gadau J. Hybrid breakdown and mitochondrial dysfunction in hybrids of *Nasonia* parasitoid wasps. *J Evol Biol.* (2008) 21:1844–1851

- Ellison, C.K. and R.S. Burton. Genotype-dependent variation of mitochondrial transcriptional profiles in interpopulation hybrids. *Proc Nat. Acad. Sci. USA* (2008) 105: 15831-15836.
- Ellison, C.K. and R.S. Burton. Interpopulation hybrid breakdown maps to the mitochondrial genome. *Evolution* (2008) 62:631-638.
- Evans AM, Mustard KJ, Wyatt CN, Peers C, Dipp M, Kumar P, Kinnear NP, Hardie DG. Does AMP-activated Protein Kinase Couple Inhibition of Mitochondrial Oxidative Phosphorylation by Hypoxia to Calcium Signaling in O<sub>2</sub>-sensing Cells ? (2005) *J. Biol. Chem.* 280, 41504
- Farineau, J., L. Pascal and G. Pelletier (1990). "Study of respiratory and photosynthetic activities in several cytoplasmic hybrids of rapeseed with cytoplasmic male sterility." *Plant Physiol Biochem* 28(3): 333-342.
- Fauron C, Casper M, Gao Y, Moore B (1995) The maize mitochondrial genome: dynamic, yet functional. *Trends Genet* 11: 228–235
- Feagin JE. Mitochondrial genome diversity in parasites. *Int J Parasitol.* (2000) 30:371-390.
- Fey V, Wagner R, Bräutigam K, Pfannschmidt T. Photosynthetic redox control of nuclear gene expression. *J Exp Bot* (2005) 56 (416) : 1491-1498
- Finnegan, P. M., M. J. Payne, E. Keramidaris and H. B. Lukins, Sanchirico, M. E., T. D. Fox and T. L. Mason, (1998) Characterization of a yeast nuclear gene, AEP2, required of mitochondrially synthesized *Saccharomyces cerevisiae* Cox2p and for accumulation of mitochondrial mRNA encoding subunit 9 of the ATP synthase. *Curr. Genet.* 20: 53–61.
- Finnegan, P.M., Payne, M.J., Keramidaris, E., and Lukins, H.B. (1991). Characterization of a yeast nuclear gene, AEP2, required for accumulation of mitochondrial mRNA encoding subunit 9 of the ATP synthase. *Curr. Genet.* 20, 53–61.
- Fishman et Willis. **A cytonuclear incompatibility causes anther sterility in *Mimulus* hybrid.** *Evolution* (2006) vol. 60 (7) pp. 1372
- Fisk, D.G., Walker, M.B., and Barkan, A. (1999). Molecular cloning of the maize gene *crp1* reveals similarity between regulators of mitochondrial and chloroplast gene expression. *EMBO J.* 18, 2621–2630.
- Forde BG, Oliver RJ, Leaver CJ. Variation in mitochondrial translation products associated with male-sterile cytoplasms in maize. *Proc Natl Acad Sci U S A.*(1978) 75(8):3841–3845.
- Forner J, Hölzle A, Jonietz C, Thuss S, Schwarzländer M, Weber B, Meyer RC, Binder S. Mitochondrial mRNA polymorphisms in different *Arabidopsis* accessions. *PLANT PHYSIOLOGY* (2008) 148 (2) : 1106-16
- Forner J, Weber B, Thuss S, Wildum S, Binder S. Mapping of mitochondrial mRNA termini in *Arabidopsis thaliana*: t-elements contribute to 5' and 3' end formation. *Nucleic Acids Research* (2007) 35 (11) : 3676-3692
- Forner J, Weber B, Wiethölter C, Meyer RC, Binder S. Distant sequences determine 5' end formation of *cox3* transcripts in *Arabidopsis thaliana* ecotype C24. *Nucleic Acids Research* (2005) 33 (15) : 4673-82
- Frank SA (1989). The evolutionary dynamics of cytoplasmic male sterility. *Am Nat* 133: 345–376.
- Frankel R, Galun E. Pollination mechanisms, reproduction, and plant breeding (1977). Heidelberg: Springer-Verlag
- Frigerio, S., Campoli, C., Zorzan, S., Fantoni, L.I., Crosatti, C., Drepper, F., Haehnel, W., Cattivelli, L., Morosinotto, T. & Bassi, R. (2007). Photosynthetic Antenna Size in Higher Plants Is Controlled by the Plastoquinone Redox State at the Post-transcriptional Rather than Transcriptional Level. *The Journal of Biological Chemistry* (2007) Vol. 282, No. 40pp. 29457-29469,
- Fujie M, Kuroiwa H, Kawanos S, Mutoh S and Kuroiwa T. Behavior of organelles and their nucleoids in the shoot apical meristem during leaf development in *Arabidopsis thaliana* L. *Planta* (1994) 194 : 395-405.
- Fujii S, Kazama T, Yamada M, Toriyama K. Discovery of global genomic re-organization based on comparison of two newly sequenced rice mitochondrial genomes with cytoplasmic male sterility-related genes. (2010) : 1-15
- Fujii S, Bond CS, Small ID. (2011). Selection patterns on restorer-like genes reveal a conflict between nuclear and mitochondrial genomes throughout angiosperm evolution. *Proceedings of the National Academy of Sciences, USA* 108: 1723–1728.
- Fujiwara M, Nagashima A, Kanamaru K, Tanaka K, Takahashi H. Three new nuclear genes, *sigD*, *sigE* and *sigF*, encoding putative plastid RNA polymerase sigma factors in *Arabidopsis thaliana* *FEBS Lett* (2000) 481 : 47–52
- Futuyma, D. (1998) *Evolutionary Biology*, Sinauer Associates
- Gagliardi and Binder (2007) Expression of the plant mitochondrial genome. *In* D Logan, ed, *Plant Mitochondria*. Blackwell Publishing, Ames, IA, pp 50–95
- Gan X, Stegle O, Behr J, Steffen JG, Drewe P, Hildebrand KL, Lyngsoe R, Schultheiss SJ, Osborne EJ, Sreedharan VT, Kahles A, Bohnert R, Jean G, Derwent P, Kersey P, Belfield EJ, Harberd NP, Kemen E, Toomajian C, Kover PX, Clark RM, Ratsch G, Mott R. Multiple reference genomes and transcriptomes

- for *Arabidopsis thaliana*. *Nature* (2011) 477 (7365) : 419-423
- Giegé and Brennicke RNA editing in *Arabidopsis* mitochondria effects 441 C to U changes in ORFs. *Proc Natl Acad Sci USA* (1999) 96: 15324–15329
- Gigord L, Lavigne C, Shykoff J.A, Atlan A. Evidence for effects of restorer genes on male and female reproductive functions of hermaphrodites in the gynodioecious species *Thymus vulgaris* L. (1999) *J. E Biol.* 12:596–604.
- Gillman, J.D., Bentolila, S., and Hanson, M.R. (2007). The petunia restorer of fertility protein is part of a large mitochondrial complex that interacts with transcripts of the CMS-associated locus. *Plant J.* 49: 217–227
- Gomez-Casati, D.F., Busi, M.V., Gonzalez-Schain, N., Mouras, A., Zabaleta, E.J., and Araya, A. (2002). A mitochondrial dysfunction induces the expression of nuclear-encoded complex I genes in engineered male sterile *Arabidopsis thaliana*. *FEBS Lett.* 532, 70–74.
- Gonzali S, Loreti E, Novi G, Poggi A, Alpi A, Perata P (2005) The use of microarrays to study the anaerobic response in *Arabidopsis*. *Ann Bot (Lond)* 96: 661–668
- Gouyon PH, Vichot F, Van Damme JMM (1991). Nuclear- cytoplasmic male sterility: single point equilibria versus limit cycles. *Am Nat* 137: 498–514.
- Gray MW, Lang BF, Cedergren R, Golding GB, Lemieux C, Sankoff D, Turmel M, Brossard N, Delage E, Littlejohn TG, Plante I, Rioux P, Saint-Louis D, Zhu Y, Burger G. (1998) Genome structure and gene content in protist mitochondrial DNAs. *Nucleic Acids Res.*, 26, 865–878
- Gray, J. C., Sullivan, J. A., Wang, J. H., Jerome, C. A. & MacLean, D. Coordination of plastid and nuclear gene expression. *Philos. Trans. R. Soc. Lond., B, Biol. Sci.* (2003) 358, 135–145
- Gray, M.W., Burger, G., Lang, B.F., (1999). Mitochondrial evolution. *Science* 283, 1476–1481.
- Grelon M, Budar F, Bonhomme S, Pelletier G (1994). Ogura cytoplasmic male-sterility (CMS)-associated orf138 is translated into a mitochondrial membrane polypeptide in male-sterile *Brassica* hybrids. *Mol Genet* 243: 540-547.
- Hagemann, R. Erwin Bour or Carl Correns: who really created the theory of plastid inheritance. *J. Hered.* (2000) 91: 435–440.
- Hajdukiewicz PT, Allison LA, Maliga P. The two RNA polymerases encoded by the nuclear and the plastid compartments transcribe distinct groups of genes in tobacco plastids. *EMBO J* (1997) 16 : 4041–4048
- Hakimi MA, Privat I, Valay JG, Lerbs-Mache S. Evolutionary conservation of C-terminal domains of primary sigma 70-type transcription factors between plants and bacteria. *J Biol Chem* (2000) 275 : 9215–9221
- Hallick et Bairoch. Proposals for the naming of chloroplast genes. III. Nomenclature for open reading frames encoded in chloroplast genomes. *Plant Mol Biol Reporter* (1994) 12 S29–S30
- Halliwell, B. and Gutteridge, J. (1999). *Free Radicals in Biology and Medicine*. New York: Oxford University
- Hammani K, Gobert A, Hleibieh K, Choulier L, Small I, Giegé P. An *Arabidopsis* Dual-Localized Pentatricopeptide Repeat Protein Interacts with Nuclear Proteins Involved in Gene Expression Regulation. *The Plant Cell Online* (2011) : 1-12
- Handa. The complete nucleotide sequence and RNA editing content of the mitochondrial genome of rapeseed (*Brassica napus* L.): comparative analysis of the mitochondrial genomes of rapeseed and *Arabidopsis thaliana*. *Nucleic Acids Research* (2003) 31 (20) : 5907-5916
- Hanson, M. R. and S. Bentolila (2004). "Interactions of mitochondrial and nuclear genes that affect male gametophyte development." *Plant Cell* 16 Suppl: S154-69.
- Hanson, M.R., Pruitt, K.D., and Nivison, H.T. (1989). Male sterility loci in plant mitochondrial genomes. In *Oxford Surveys of Plant Molecular and Cell Biology*, 6, B.J. Mifflin, ed (Oxford: Clarendon Press), : 61-85.
- Hanson, M.R., Wilson, R.K., Bentolila, S., Kohler, R.H., and Chen, H.C. (1999). Mitochondrial gene organization and expression in petunia male fertile and sterile plants. *J. Hered.* 90, 362–368.
- Hardingham, G.E., S. Chawla, C.M. Johnson, et H. Bading. (1997). Distinct functions of nuclear and cytoplasmic calcium in the control of gene expression. *Nature.* 385:260-5.
- Hashimoto M, Endo T, Peltier G, Tasaka M, Shikanai T. A nucleus-encoded factor, CRR2, is essential for the expression of chloroplast *ndhB* in *Arabidopsis*. *The Plant Journal* (2003) 36 (4) : 541-549
- Hattori, M., Miyake, H. and Sugita, M. (2007). A pentatricopeptide repeat protein is required for RNA processing of *clpP* pre-mRNA in moss chloroplasts. *J. Biol. Chem.* 282: 10773-10782.
- Hayes R, Kudla J and Gruissem W (1999) Degrading chloroplast mRNA: the role of polyadenylation. *Trends Biochem Sci*, 24, 199–202
- Hayes R, Kudla J, Schuster G, Gabay L, Maliga P and Gruissem W (1996) Chloroplast mRNA 3'-end processing by a high molecular weight protein complex is regulated by nuclear encoded RNA binding

- proteins. *EMBO J*, 15, 1132–1141.
- He F, Kang D, Ren Y, Qu LJ, Zhen Y, Gu H. Genetic diversity of the natural populations of *Arabidopsis thaliana* in China. *Heredity* (2007) 99 (4) : 423-431
- Heazlewood JL, Howell KA, Millar AH. Mitochondrial complex I from *Arabidopsis* and rice: orthologs of mammalian and fungal components coupled with plant-specific subunits. (2003) *Biochim Biophys Acta*. (3):159-69.
- Hedtke B, Börner T and Weihe A. One RNA polymerase serving two genomes. *EMBO* (2000) 15 : 435–440
- Hedtke B, Börner T, Weihe A. Mitochondrial and Chloroplast Phage-Type RNA Polymerases in *Arabidopsis* *Science* (1997) 277 no. 5327 : 809-811
- Hedtke B, Legen J, Weihe A, Herrmann RG, Börner T. Six active phage-type RNA polymerase genes in *Nicotiana tabacum*. *Plant J* (2002) 30 : 625–637
- Hess and Börner. *Int Rev Cytol. Organellar RNA polymerases of higher plants* (1999) 190 : 1-59.
- Heyn F. W., (1976). Transfer of restorer genes from *Raphanus* to cytoplasmic male-sterile *Brassica napus*. *Cruciferae Newsletter*, 1, 15.
- Hirose, T., and M. Sugiura (2001). Involvement of a site-specific trans-acting factor and a common RNA-binding protein in the editing of chloroplast mRNAs: development of a chloroplast in vitro RNA editing system. *EMBO J*. 20:1144–1152.
- Hoch B, Maier RM, Appel K, Igloi GL, Kössel H. Editing of a chloroplast mRNA by creation of an initiation codon. *Nature* (1991) 353 : 178–180
- Hoffmann and Binder. Functional importance of nucleotide identities within the pea *atp9* mitochondrial promoter sequence. *JMolBiol* (2002) 320:943–50.
- Hoffmann. Biogeography of *Arabidopsis thaliana* (L.) Heynh.(Brassicaceae). *Journal of Biogeography* (2002),29, 125-134
- Holec S, Lange H, Kühn K, Alioua M, Börner T, Gagliardi D. Relaxed transcription in *Arabidopsis* mitochondria is counterbalanced by RNA stability control mediated by polyadenylation and polynucleotide phosphorylase. *Molecular and Cellular Biology* (2006) 26 (7) : 2869-76
- Hölzle A, Jonietz C, Törjek O, Altmann T, Binder S, Forner J (2011) A RESTORER OF FERTILITY-like PPR gene is required for 5'-end processing of the *nad4* mRNA in mitochondria of *Arabidopsis thaliana*. *Plant J* 65: 737-744
- Honys and Twell. Transcriptome analysis of haploid male gametophyte development in *Arabidopsis*. *Genome Biol* (2004) 5 (11) : R85
- Horlow, C., Goujoud, J., Lepingle, A., Missonier, C. and Bourgin, J.P. Transmission of paternal chloroplasts in tobacco (*Nicotiana tabacum*). *Plant Cell Rep.* (1990) 9, 249–252.
- Hu Y, Zhang Q, Rao G, Sodmergen. Occurrence of plastids in the sperm cells of Caprifoliaceae: biparental plastid inheritance in Angiosperms is unilaterally derived from maternal inheritance. *Plant Cell Physiol* (2008) 49:958–968.
- Huang CY, Ayliffe MA, Timmis JN (2003) Direct measurement of the transfer rate of chloroplast DNA into the nucleus. *Nature* 422: 72–76
- Huang J, Mullapudi N, Lancto CA, Scott M, Abrahamsen MS, Kissinger JC. (2004). Phylogenomic evidence supports past endosymbiosis, intracellular and horizontal gene transfer in *Cryptosporidium parvum*. *Genome Biol* 5(11): R88
- Huang SB, Greenway H, Colmer TD, Millar AH (2005) Protein synthesis by rice coleoptiles during prolonged anoxia: implications for glycolysis, growth and energy utilization. *Ann Bot (Lond)* 96: 703–715
- Hurst, L. D., A. Atlan, and B. Bengtsson. 1996. Genetic conflicts. *Quarterly Review of Biology* 71:317–364.
- Ikeda and Gray. Characterization of a DNA-binding protein implicated in transcription in wheat mitochondria. *Molecular and Cellular Biology* (1999) 19 (12) : 8113-22
- Irhar, S., Schlichter, M., and Swartberg, D. (1983). Sorting out of cytoplasmic elements in somatic hybrids of *Petunia* and the prevalence of the heteroplasmion through cycles. *MOI. Gen. Genet.* 190, 468-474.
- Ishihama. Functional modulation of *Escherichia coli* RNA polymerase *Annu Rev Microbiol* (2000) 54 : 809-811
- Ishikawa R, Kinoshita T. Epigenetic programming: the challenge to species hybridization (2009) *Mol Plant*. 2(4):589-99.
- Ishizaki Y, Tsunoyama Y, Hatano K, Ando K, Kato K, Shinmyo A, Kobori M, Takeba G, Nakahira Y, Shiina T. A nuclear-encoded sigma factor, *Arabidopsis* SIG6, recognizes sigma-70 type chloroplast promoters and regulates early chloroplast development in cotyledons *Plant J* (2005) 42 : 133–144
- Isono K, Shimizu M, Yoshimoto K, Niwa Y, Satoh K, Yokota A, Kobayashi H. Leaf-specifically expressed genes for polypeptides destined for chloroplasts with domains of sigma 70 factors of bacterial RNA polymerases in *Arabidopsis thaliana* (1997) *Proc Natl Acad Sci USA* 94 : 14948–14953
- Ito, Y., Saisho, D., Nakazono. M., Tsutsumi, N., Hirai. A., (1997). Transcript levels of tandem- arranged

- alternative oxidase genes in rice are increased by low temperature. *Gene*, 203: 121-129.
- Janska H, Sarria R, Woloszynska M, Arrieta-Montiel M, and Mackenzie SA, (1998). Stoichiometric shifts in the common bean mitochondrial genome leading to male sterility and spontaneous reversion to fertility. *Plant Cell* 10:1163–1180.
- Jenkins BD, Kulhanek DJ and Barkan A. Nuclear mutations that block group II RNA splicing in maize chloroplasts reveal several intron classes with distinct requirements for splicing factors. *Plant Cell*, (1997) 9, 283–296.
- Johnson X, Wostrickoff K, Finazzi G, Kuras R, Schwarz C, Bujaldon S, Nickelsen J, Stern DB, Wollman FA, Vallon O. MRL1, a Conserved Pentatricopeptide Repeat Protein, Is Required for Stabilization of rbcL mRNA in *Chlamydomonas* and *Arabidopsis*. *The Plant Cell Online* (2010) 22 (1) : 234-248
- Jones. Does the plant mitochondrion integrate cellular stress and regulate programmed cell death ? *Trends Plant Science* (2000) 5 : 225-230.
- Jonietz C, Forner J, Hölzle A, Thuss S, Binder S. RNA PROCESSING FACTOR2 Is Required for 5' End Processing of nad9 and cox3 mRNAs in Mitochondria of *Arabidopsis thaliana*. *The Plant Cell Online* (2010) 22 (2) : 443-453
- Jorgensen and Mauricio. Neutral genetic variation among wild North American populations of the weedy plant *Arabidopsis thaliana* is not geographically structured. *Mol Ecol* (2004) 13 : 3403–3413
- Jung and Chory. Signaling between Chloroplasts and the Nucleus: Can a Systems Biology Approach Bring Clarity to a Complex and Highly Regulated Pathway?. *Plant Physiology* (2010) 152 (2) : 453-459
- Kabeya and Sato. Unique translation initiation at the second AUG codon determines mitochondrial localization of the phage-type RNA polymerases in the moss *Physcomitrella patens*. *Plant Physiol* (2005) 138:369–82 ;
- Kabeya Y, Hashimoto K, Sato N. Identification and characterization of two phage-type RNA polymerase cDNAs in the moss *Physcomitrella patens*: implication of recent evolution of nuclear-encoded RNA polymerase of plastids in plants. *Plant Cell Physiol* (2002) 43 : 245-255
- Kabeya, H., S. Maruyama, Y. Morita, M. Kubo, K. Yamamoto, S. Arai, T. Izumi, Y. Kobayashi, Y. Katsube, and T. Mikami. Distribution of *Arcobacter* species among livestock in Japan. *Vet. Microbiol.* (2003) 93:153–158.
- Kaneko T, Sato S, Kotani H, Tanaka A, Asamizu E, Nakamura Y, Miyajima N, Hirosawa M, Sugiura M, Sasamoto S, Kimura T, Hosouchi T, Matsuno A, Muraki A, Nakazaki N, Naruo K, Okumura S, Shimpo S, Takeuchi C, Wada T, Watanabe A, Yamada M, Yasuda M, Tabata S. Sequence analysis of the genome of the unicellular cyanobacterium *Synechocystis* sp. strain PCC6803. II. Sequence determination of the entire genome and assignment of potential protein-coding regions. (1996) *DNA Res.*;3(3):109-36.
- Kanervo E, Singh M, Suorsa M, Paakkanen V, Aro E, Battchikova N, Aro EM (2008) Expression of protein complexes and individual proteins upon transition of etioplasts to chloroplasts in pea (*Pisum sativum*). *Plant Cell Physiol* 49: 396–410
- Karpova OV, Kuzmin EV, Elthon TE, Newton KJ. Differential expression of alternative oxidase genes in maize mitochondrial mutants. *The Plant Cell Online* (2002) 14 (12) : 3271
- Kazama T, Nakamura T, Watanabe M, Sugita M, Toriyama K. Suppression mechanism of mitochondrial ORF79 accumulation by Rf1 protein in BT-type cytoplasmic male sterile rice. (2008) *Plant J* 55: 619–628
- Keren I, Bezawork-Geleta A, Koltun M, Maayan I, Belausov E, Levy M, Mett A, Gidoni D, Shaya F, Ostersetzer-Biran O. AtnMat2, a nuclear-encoded maturase required for splicing of group-II introns in *Arabidopsis* mitochondria. *RNA* (2009) 15 (12) : 2299-2311
- Kim KJ, Choi KS, Jansen RK. Two chloroplast DNA inversions originated simultaneously during early evolution in the sunflower family. *Mol Biol Evol* (2005) 22 : 1783–1792.
- Kim, M.K., Christopher, D.A. and Mullet, J.E. (1993) Direct evidence for selective modulation of psbA, rpoA, rbcL and 16S RNA stability during barley chloroplast development. *Plant Mol. Biol.* 22: 447–463.
- Kishine M, Takabayashi A, Munekage Y, Shikanai T, Endo T, Sato F (2004) Ribosomal RNA processing and an RNase R family member in chloroplasts of *Arabidopsis*. *Plant Mol Biol* 55: 595–606
- Kitazaki and Kubo. Cost of Having the Largest Mitochondrial Genome: Evolutionary Mechanism of Plant Mitochondrial Genome. *Journal of Botany* (2010) 2010 : 1-12
- Klaff and Gruissem. Changes in Chloroplast mRNA Stability during Leaf Development. (1991) *The Plant Cell*, Vol 3, Issue 5 517-529
- Kleffmann T, von Zychlinski A, Russenberger D, Hirsch-Hoffmann M, Gehrig P, Gruissem W, Baginsky S (2007) Proteome dynamics during plastid differentiation in rice. *Plant Physiol* 143: 912–923
- Kleine, T., Kindgren, P., Benedict, C., Hendrickson, L., and Strand, A. (2007). Genome-wide gene expression analysis reveals a critical role for CRYPTOCHROME1 in the response of *Arabidopsis* to high irradiance. *Plant Physiol.* 144 1391–1406.
- Köhler D, Schmidt-Gattung S, Binder S. The DEAD-box protein PMH2 is required for efficient group II intron splicing in mitochondria of *Arabidopsis thaliana* *Plant Mol Biol.* (2010) 72(4-5):459-67.

- Koch M, Haubold B, Mitchell-Olds T. 2001. Molecular systematics of the Brassicaceae: evidence from coding plastidic matK and nuclear Chs sequences. *Am J Bot.* 88:534–544.
- Koizuka, N., Imai, R., Fujimoto, H., Hayakawa, T., Kimura, Y., Kohno-Murase, J., Sakai, T., Kawasaki, S., and Imamura, J. (2003). Genetic characterization of a pentatricopeptide repeat protein gene, orf687, that restores fertility in the cytoplasmic male-sterile kosen radish. *Plant J.* 34, 407–415.
- Kolodner et Tewari. Molecular size and conformation of chloroplast deoxyribonucleic acid from pea leaves. *J Biol Chem.* (1972) 247(19) : 6355– 6364
- Koornneef and Meinke. The development of Arabidopsis as a model plant. *The Plant Journal* (2010) 61 (6) : 909-921
- Kotchoni SO, Jimenez-Lopez JC, Gachomo EW, Seufferheld MJ. A New and Unified Nomenclature for Male Fertility Restorer (RF) Proteins in Higher Plants. *PLoS ONE* (2010) 5 (12) : e15906
- Kotera E, Tasaka M, Shikanai T. A pentatricopeptide repeat protein is essential for RNA editing in chloroplasts. *Nature* (2005) 433 : 326–330
- Koussevitzky S, Nott A, Mockler TC, Hong F, Sachetto-Martins G, Surpin M, Lim J, Mittler R, Chory J. Signals from chloroplasts converge to regulate nuclear gene expression. *Science* (2007) 316, 715–719
- Kubo and Mikami. Organization and variation of angiosperm mitochondrial genome. *Physiol Plant* (2007) 129 (1) : 6-13
- Kubo and Newton. Angiosperm mitochondrial genomes and mutations. (2008) *Mitochondrion* 8(1):5-14
- Kudla J, Hayes R, Gruissem W. Polyadenylation accelerates degradation of chloroplast mRNA. *The EMBO Journal* (1996) 15 (24) : 7137-46
- Kudla J, Hayes R, Gruissem W. Polyadenylation accelerates degradation of chloroplast mRNA. *The EMBO Journal* (1996) 15 (24) : 7137-46
- Kugita M, Yamamoto Y, Fujikawa T, Matsumoto T, Yoshinaga K. RNA editing in hornwort chloroplasts makes more than half the genes functional. *Nucleic Acids Res* (2003) 31 : 2417– 2423
- Kühn, K., Bohne, A.V., Liere, K., Weihe, A., and Börner, T. (2007). *Arabidopsis* phage-type RNA polymerases: Accurate in vitro transcription of organellar genes. *Plant Cell* 19: 959–971
- Kühn, K., Weihe, A., and Börner, T. (2005). Multiple promoters are a common feature of mitochondrial genes in Arabidopsis. *Nucleic Acids Res.* 33: 337–346
- Kuroiwa, T. The replication, differentiation, and inheritance of plastids with emphasis on the concept of organelle nuclei. *Int. Rev. Cytol.* (1991) 128 : 1–61.
- Kurtz S, Choudhuri JV, Ohlebusch E, Schleiermacher C, Stoye J, Giegerich R (2001) REPUTER: the manifold applications of repeat analysis on a genomic scale. *Nucleic Acids Res* 29:4633– 4642
- Kumar et al. (2011) An evolutionarily conserved mitochondrial orf108 is associated with cytoplasmic male sterility in different alloplasmic lines of Brassica juncea and induces male sterility in transgenic Arabidopsis thaliana *J Exp Bot* 2012 0: err459v1-err459.
- Kushnir S, Babychuk E, Storozhenko S, Davey MW, Papenbrock J, De Rycke R, Engler G, Stephan UW, Lange H, Kispal G, et al (2001) A mutation of the mitochondrial ABC transporter Stal leads to dwarfism and chlorosis in the Arabidopsis mutant starik. *Plant Cell* 13: 89–100
- Kushnir, S., Babychuk, E., Bannikova, M., Momot, V., Komarnitsky, I., Cherep, N., and Gleba, Y. (1991). Nucleo-cytoplasmic incompatibility in cybrid plants possessing an *Atropa* genome and a *Nicotiana* plastome. *Mol. Gen. Genet.* 225, 225–230
- Lamppa GK, Elliot LV and Bendich AJ. Changes in chloroplast number during pea leaf development. An analysis of a protoplast population *Planta* (1980) 148 : 437–433.
- Lang BF, Burger G, O'Kelly CJ, Cedergren R, Golding GB, Lemieux C, Sankoff D, Turmel M, Gray MW. An ancestral mitochondrial DNA resembling a eubacterial genome in miniature. *Nature* (1997) 387:493–7.
- Laser, K.D., and Lersten, N.R. (1972). Anatomy and cytology of microsporogenesis in cytoplasmic male sterile angiosperms. *Bot. Rev.* 38, 425-454.
- Lee HY, Chou JY, Cheong L, Chang NH, Yang SY, Leu JY. Incompatibility of Nuclear and Mitochondrial Genomes Causes Hybrid Sterility between Two Yeast Species (2008) *Cell* 135, 1065–1073
- Lee KP, Kim C, Landgraf F, Apel K. EXECUTER1- and EXECUTER2-dependent transfer of stress-related signals from the plastid to the nucleus of Arabidopsis thaliana (2007) *Proc Natl Acad Sci U S A.*;104(24):10270-5.
- Lee Y, Kim ES, Choi Y, Hwang I, Staiger CJ, Chung YY, Lee Y. The Arabidopsis Phosphatidylinositol 3-Kinase Is Important for Pollen Development. *PLANT PHYSIOLOGY* (2008) 147 (4) : 1886-1897
- Lefebvre, A., Scalla, R. and Preirfer, P. (1990). The double-stranded RNA associated with the '447' cytoplasmic male sterility in *Vicia faba* is packaged together with its replicase in cytoplasmic membranous vesicles. *Plant Molecular Biology* 14, 477-490.
- Leister. Genomics-based dissection of the cross-talk of chloroplasts with the nucleus and mitochondria in Arabidopsis. *Gene* (2005) 354 : 110-116

- Levitin B, Richter D, Markovich I, Zik M (2008) Arabinogalactan proteins 6 and 11 are required for stamen and pollen function in *Arabidopsis*. *Plant J* 56: 351–363
- Lewis, D. Male sterility in natural populations of hermaphrodite plants. (1941) *New Phytologist*, 40 : 50-63.
- L'Homme Y, Stahl RJ, Li XQ, Hameed A, Brown GG. Brassica nap cytoplasmic male sterility is associated with expression of a mitDNA region containing a chimeric gene similar to the pol CMS-associated *orf224* gene. *Curr Genet* (1997) 31 : 325-335.
- Li XQ, Jean M, Landry BS, Brown GG. Restorer genes for different forms of Brassica cytoplasmic male sterility map to a single nuclear locus that modifies transcripts of several mitochondrial genes. *Proc Natl Acad Sci USA* (1998) 95 (17) : 10032
- Li, X.-Q., M. Jean, B.S. Landry, and G.G. Brown. (1998). Restorer genes for different forms of *Brassica* cytoplasmic male sterility map to a single nuclear locus that modifies transcripts of several mitochondrial genes. *Proc. Natl. Acad. Sci. USA* 95: 10032-10037.
- Liepins, A., Hennen, S., (1977). Cytochrome oxidase deficiency during development of amphibian nucleocytoplasmic hybrids. *Developmental Biology* 57, 284–292.
- Liere and Maliga. Plastid RNA polymerases in higher plants. *Regulation of Photosynthesis*, Anderson B, Aro EM (eds) (2001) : 29–39. Rotterdam: Kluwer.
- Linder. Dead-box proteins: a family affair--active and passive players in RNP-remodeling. *Nucleic Acids Research* (2006) 34 (15) : 4168-4180
- Link. Redox regulation of chloroplast transcription. *Antioxid Redox Signal*. (2003) 5 : :79-87.
- Linke and Börner. Mitochondrial effects on flower and pollen development. *Mitochondrion* (2005) 5 (6) : 389-402
- Lister C. and Dean C. (1993). Recombinant inbred lines for mapping RFLP and phenotypic markers in *Arabidopsis thaliana*. *Plant Journal* 4 745-750
- Liu and Butow. Mitochondrial retrograde signaling. *Annu. Rev. Genet.* (2006) 40 : 159-185
- Liu F, Cui X, Horner HT, Weiner H, Schnable PS. Mitochondrial aldehyde dehydrogenase activity is required for male fertility in maize. *Plant Cell* (2001) 13(5) 1063-78
- Lloyd, D.G. (1974a). Theoretical sex ratios of dioecious and gynodioecious angiosperms. *Heredity* 32, 11-34.
- Lloyd, D.G. (1974b). Female predominant sex ratios in angiosperms. *Heredity* 32, 35-44.
- Loefering W. Q. and E. R. Sears, (1963) Distorted inheritance of stem-rust resistance in timstein wheat caused by a pollen-killing gene. *Can. J. Genet. Cytol.* 5: 65-72.
- Logan DC, Knight MR (2003) Mitochondrial and cytosolic calcium dynamics are differentially regulated in plants. *Plant Physiology* 133: 21-24.
- Long et al (2008) Hybrid male sterility in rice controlled by interaction between divergent alleles of two adjacent genes *Proc Natl Acad Sci U S A* 105:18871-18876.
- Lonosky PM, Zhang X, Honavar VG, Dobbs DL, Fu A, Rodermeil SR. A Proteomic Analysis of Maize Chloroplast Biogenesis. *PLANT PHYSIOLOGY* (2004) 134 (2) : 560-574
- Lonsdale DM, Brears T, Hodge TP, Melville SE, Rottmann WH (1988) The plant mitochondrial genome: homologous recombination as a mechanism for generating heterogeneity. *Philos Trans R Soc Lond B Biol Sci* 319: 149–163
- López-Juez. Plastid biogenesis, between light and shadows *J Exp Bot* (2007) 58(1) pp 11-26.
- Lorkovic, Z.J., Barta, A. (2002). Genome analysis: RNA recognition motif (RRM) and K homology (KH) domain RNA-binding proteins from the flowering plant *Arabidopsis thaliana*. *Nucleic Acids Research* 30(3), 623-35.
- Lown, F.J., Watson, A.T., and Purton, S. (2001). *Chlamydomonas* nuclear mutants that fail to assemble respiratory or photosynthetic electron transfer complexes. *Biochem. Soc. Trans.* 29, 452–455.
- Lu G, Sehnke PC, Ferl RJ (1994b) Phosphorylation and calcium binding properties of an *Arabidopsis* GF14 brain protein homolog. *Plant Cell* 6:501–510.
- Lupold DS, Caoile AG, Stern DB. Genomic context influences the activity of maize mitochondrial *cox2* promoters. *Proc Natl Acad Sci USA* (1999) 96 (20) : 11670
- Lurin C, Andrés C, Aubourg S, Bellaoui M, Bitton F, Bruyère C, Caboche M, Debast C, Gualberto J, Hoffmann B, Lecharny A, Le Ret M, Martin-Magniette ML, Mireau H, Peeters N, Renou JP, Szurek B, Taconnat L, Small I. Genome-wide analysis of *Arabidopsis* pentatricopeptide repeat proteins reveals their essential role in organelle biogenesis. *Plant Cell* (2004) 16 (8) : 2089-103
- Lutz and Maliga. Lack of conservation of editing sites in mRNAs that encode subunits of the NAD(P)H dehydrogenase complex in plastids and mitochondria of *Arabidopsis thaliana*. *Curr. Genet.* (2001) 40 : 214–219
- Mackenzie et McIntosh. Higher plant mitochondria. *The Plant Cell* (1999) 11 : 571–585.
- Mackenzie, S. (2004). The influence of mitochondrial genetics in crop breeding strategies. *Plant Breed. Rev.* 25, 115–138.
- Mackenzie, S. A., (2007) The unique biology of mitochondrial genome instability in plants, : 36–46 in *Plant*

- Mitochondria, edited by D. Logan. Blackwell Publishing, Oxford.
- Mackenzie, S., A., and C. D. Chase, (1990) Fertility restoration is associated with loss of a portion of the mitochondrial genome in cytoplasmic-male sterility common bean. *Plant Cell* 2: 905- 912.
- MacLean D, Jerome CA, Brown APC, Gray JC (2008) Co-regulation of nuclear genes encoding plastid ribosomal proteins by light and plastid signals during seedling development in tobacco and Arabidopsis. *Plant Mol Biol* 66: 475–490
- Maier RM, Neckermann K, Igloi GL, Kössel H. Complete sequence of the maize chloroplast genome-gene content, hotspots of divergence and fine-tuning of genetic information by transcript editing. *J Mol Biol* (1995) 251 :614–628
- Manthey, G.M., and McEwen, J.E. (1995). The product of the nuclear gene PET309 is required for translation of mature mRNA and stability or production of intron-containing RNAs derived from the mitochondrial COX1 locus of *Saccharomyces cerevisiae*. *EMBO J.* 14, 4031–4043.
- Marechal and Brisson. Recombination and the maintenance of plant organelle genome stability. *New Phytologist* (2010) 186 (2) : 299-317
- Margulis. Microbial Communities in the Archean and Proterozoic Eons. New York : W. H. Freeman & Co. Symbiosis in Cell Evolution (1993).
- Marienfeld J, Unseld M, Brennicke A. The mitochondrial genome of Arabidopsis is composed of both native and immigrant information. *Trends Plant Sci* (1999) 4 (12) : 495-502
- Martin W, Rujan T, Richly E, Hansen A, Cornelsen S, Lins T, Leister D, Stoebe B, Hasegawa M, Penny D. Evolutionary analysis of Arabidopsis, cyanobacterial, and chloroplast genomes reveals plastid phylogeny and thousands of cyanobacterial genes in the nucleus. *Proc Natl Acad Sci USA* (2002) 99 (19) : 12246-51
- Mathy N, Benard L, Pellegrini O, Daou R, Wen T, Condon C (2007) 5'-to-3' exoribonuclease activity in bacteria: Role of RNase J1 in rRNA maturation and 5' stability of mRNA. *Cell* 129:681–692.
- Matsushima R, Hamamura Y, Higashiyama T, Arimura S, Sodmergen, Tsutsumi N, Sakamoto W. Mitochondrial dynamics in plant male gametophyte visualized by fluorescent live imaging. *Plant Cell Physiol.* (2008) 49 : 1074–1083.
- Matthes A, Schmidt-Gattung S, Köhler D, Forner J, Wildum S, Raabe M, Urlaub H, Binder S. Two DEAD-Box Proteins May Be Part of RNA-Dependent High-Molecular-Mass Protein Complexes in Arabidopsis Mitochondria. *PLANT PHYSIOLOGY* (2007) 145 (4) : 1637-1646
- Maxwell DP, Nickels R, McIntosh L (2002) Evidence of mitochondrial involvement in the transduction of signals required for the induction of genes associated with pathogen attack and senescence. *Plant J* 29: 269–279
- Maxwell DP, Wang Y, McIntosh L (1999) The alternative oxidase lowers mitochondrial reactive oxygen production in plant cells. *Proc Natl Acad Sci USA* 96: 8271–8276
- Maxwell, K. L., Mittermaier, A. K., Forman-Kay, J. D. and Davidson, A. R. (1999). A simple in vivo assay for increased protein solubility. *Protein Sci* 8, 1908-11.
- McCormick, S. (1993) Male gametophyte development. *Plant Cell*, 5, 1265–1275.
- McKhann HI, Camilleri C, Bérard A, Bataillon T, David JL, Reboud X, Le Corre V, Caloustian C, Gut IG, Brunel D. Nested core collections maximizing genetic diversity in *Arabidopsis thaliana*. *The Plant Journal* (2004) 38 (1) : 193-202
- Meierhoff K, Felder S, Nakamura T, Bechtold N, Schuster G. HCF152, an Arabidopsis RNA binding pentatricopeptide repeat protein involved in the processing of chloroplast psbB-psbT-psbH-petB-petD RNAs. *Plant Cell* (2003) 15: 1480-1495
- Mereschkowsky. *Über Natur und Ursprung der Chromatophoren im Pflanzenreiche* (1905) traduit par Martin et Kowallik. *Eur. J. Phycol.* (1999) 34 : 287–295.
- Mesquida J., Renard M., The entomophilous pollination of male sterile strains of winter Rapeseed (*Brassica napus* L. var. *oleifera* METZGER ) and a preliminary study of alternating devices. *Proc. 4th. Int. Symp. Pollination, Maryland*, (1978) 49-57.
- Mili, S., and Pinol-Roma, S. (2003). LRP130, a pentatricopeptide motif protein with a noncanonical RNA-binding domain, is bound in vivo to mitochondrial and nuclear RNAs. *Mol. Cell. Biol.* 23, 4972–4982.
- Millar A, Eubel H, Jansch L, Kruff V, Heazlewood J, Braun H (2004) Mitochondrial cytochrome c oxidase and succinate dehydrogenase complexes contain plant specific subunits. *Plant Mol Biol* 56: 77–90
- Mingam A, Toffano-Nioche C, Brunaud V, Boudet N, Kreis M, Lecharny A (2004) DEAD-box RNA helicases in *Arabidopsis thaliana*: establishing a link between quantitative expression, gene structure and evolution of a family of genes. *Plant Biotechnol J* 2: 401–415
- Mochizuki N, Tanaka R, Tanaka A, Masuda T, Nagatani A. The steady-state level of Mg-protoporphyrin IX is not a determinant of plastid-to-nucleus signaling in *Arabidopsis*. *Proc Natl Acad Sci U S A.* (2008) 105(39): 15184–15189.
- Mogensen. The hows and whys of cytoplasmic inheritance in seed plants. (1996) *Am. J. Bot.* 83 : 383-404.

- Mohr and Lambowitz. Putative proteins related to group II intron reverse transcriptase/maturases are encoded by nuclear genes in higher plants. *Nucleic Acids Res* (2003) 31:647–652.
- Moison M, Roux F, Quadrado M, Duval R, Ekovich M, Lê DH, Verzaux M, Budar F. Cytoplasmic phylogeny and evidence of cyto-nuclear co-adaptation in *Arabidopsis thaliana*. *Plant J* (2010) 63 (5) : 728-38
- Moller (2001), Plant mitochondria and oxidative stress. Electron transport, NADPH turnover and metabolism of reactive oxygen species. *Annu. Rev. Plant Physiol. Plant Mol. Biol.*, 52,561-591.
- Moulin M, McCormac AC, Terry MJ, Smith AG. Tetrapyrrole profiling in *Arabidopsis* seedlings reveals that retrograde plastid nuclear signalling is not due to Mg-protoporphyrin IX accumulation. *Proc Natl Acad Sci USA*. (2008) 105:15178–15183.
- Mourier T, Hansen AJ, Willerslev E, Arctander P (2001) The Human Genome Project reveals a continuous transfer of large mitochondrial fragments to the nucleus. *Mol Biol Evol* 18: 1833–1837
- Muller HJ (1942). Isolation mechanisms, evolution and temperature. *Biol Symp* 6: 71–125.
- Mullet. Dynamic regulation of chloroplast transcription. *Plant Physiol* (1993) 103 : 309–313
- Myles S, Peiffer J, Brown PJ, Ersoz ES, Zhang Z, Costich DE, Buckler ES. (2009) Association mapping: Critical considerations shift from genotyping to experimental design. *Plant Cell* 21:2194–2202. –
- Nagata N, Saito C, Sakai A, Kuroiwa H, Kuroiwa T. The selective increase or decrease of organellar DNA in generative cells just after pollen mitosis one controls cytoplasmic inheritance. *Planta* (1999) 209 : 53–65.
- Nakamura T, Meierhoff K, Westhoff P, Schuster G. RNA-binding properties of HCF152, an *Arabidopsis* PPR protein involved in the processing of chloroplast RNA. (2003) *Eur J Biochem*. 270(20):4070-81.
- Nakamura T, Schuster G, Sugiura M. and Sugita M. Chloroplast RNA-binding and pentatricopeptide repeat proteins. *Biochemical Society Transactions* (2004) 32 : 571–574 .
- Nasrallah, M., Yogeewaran, K., Snyder, S., Nasrallah, J.B., (2000). *Arabidopsis* species hybrids in the study of species differences and evolution of amphiploidy in plants. *Plant Physiol*. 124, 1605–1614.
- Nasrallah, M.E., Liu, P., Sherman-Broyles, S., Boggs, N.A., Nasrallah, J.B., (2004). Natural variation in expression of self incompatibility in *Arabidopsis thaliana*: implications for the evolution of selfing. *Proc. Natl. Acad. Sci. USA* 101, 16070–16074.
- Nivison and Hanson. Identification of a mitochondrial protein associated with cytoplasmic male sterility in *petunia*. *Plant Cell* (1989) 1 (11) : 1121-30
- Noh et al. Non publié
- Notsu Y, Masood S, Nishikawa T, Kubo N, Akiduki G, Nakazono M, Hirai A, Kadowaki K. The complete sequence of the rice (*Oryza sativa* L.) mitochondrial genome: frequent DNA sequence acquisition and loss during the evolution of flowering plants. *Molecular Genetics and Genomics* (2002) 268 (4) : 434-445
- Nott A, Jung HS, Koussevitzky S, Chory J. Plastid-to-nucleus retrograde signaling. *Annu. Rev. Plant. Biol.* (2006) 57 (1) : 739-759
- Oda K, Yamato K, Ohta E, Nakamura Y, Takemura M, Nozato N, Akashi K, Kanegae T, Ogura Y, Kohchi T, et al (1992) Gene organization deduced from the complete sequence of liverwort *Marchantia polymorpha* mitochondrial DNA. A primitive form of plant mitochondrial genome. *J Mol Biol* 223: 1–7
- Oelmüller and Mohr. Photooxidative destruction of chloroplasts and its consequences for expression of nuclear genes. (1986) *Planta*, 167, 106-113.
- Ogura H., (1968). Studies of the new male-sterility in Japanese radish, with special référence to the utilization of this sterility towards the practical raising of hybrid seeds. *Mern. Fac. Agric. Kagoshima Univer.*, 6, 39-78.
- Ohyama K., Fukuzawa H., Kohchi T., Shirai H., Sano T., Sano S., Umesono K., Shiki Y., Takeuchi M., Chang Z., Aota S-I, Inokuchi H. and Ozeki H. Chloroplast gene organization deduced from complete sequence of liverwort *Marchantia polymorpha* chloroplast DNA. *Nature* (1986) 322 : 572-574.
- Okuda K, Myouga F, Motohashi R, Shinozaki K, Shikanai T. Conserved domain structure of pentatricopeptide repeat proteins involved in chloroplast RNA editing. *Proc Natl Acad Sci USA* (2007) 104 (19) : 8178-83
- Okuda K, Nakamura T, Sugita M, Shimizu T, Shikanai T (2006) A pentatricopeptide repeat protein is a site-recognition factor in chloroplast RNA editing. *J Biol Chem* 281: 37661–37667
- Okuda, K., Hammani, K., Tanz, S. K., Peng, L., Fukao, Y., Myouga, F., Motohashi, R., Shinozaki, K., Small, I. und Shikanai, T. (2010). The pentatricopeptide repeat protein OTP82 is required for RNA editing of plastid *ndhB* and *ndhG* transcripts. *Plant J*. 61: 339-349.
- Oldenburg and Bendich. Changes in the structure of DNA molecules and the amount of DNA per plastid during chloroplast development in maize. *J Mol Biol* (2004) 344 (5) : 1311-30.
- Orr, D. 2005. Armageddon versus extinction. *Conservation Biology* 19:290–292.
- Ostheimer, G., Williams-Carrier, R., Belcher, S., Osborne, E., Gierke, J., and Barkan, A. Group II intron splicing factors derived by diversification of an ancient RNA binding module. *EMBO J*. (2003) 22: 3919–3929.

- O'Toole N, Hattori M, Andres C, Iida K, Lurin C, Schmitz-Linneweber C, Sugita M, Small I. On the Expansion of the Pentatricopeptide Repeat Gene Family in Plants. *Molecular Biology and Evolution* (2008) 25 (6) : 1120-1128
- Oudot-Le Secq MP, Fontaine JM, Rousvoal S, Kloareg B, Loiseaux-De Goër S. The complete sequence of a brown algal mitochondrial genome, the ectocarpale *Pylaiella littoralis* (L.) Kjellm. (2001) *J Mol E* 2001 (2):80-8.
- Palenik. The genomics of symbiosis : hosts keep the baby and the bath water. *Proc. Natl. Acad. Sci.* (2002) 99 : 11996–11997.
- Palmer and Herbon. Plant mitochondrial DNA evolves rapidly in structure but slowly in sequence (1989) *Journal of Molecular Evolution* 28(1-2):87-97.
- Palmer and Shields. (1984). Tripartite structure of the *Brassica campestris* mitochondrial genome. *Nature* 307:437-440
- Palmer JD, Nugent JM, Herbon LA. Unusual structure of geranium chloroplast DNA: a triple-sized repeat, extensive gene duplications, multiple inversions, and new repeat families. *Proc Natl Acad Sci U S A.* (1987) 84 :769–773.
- Palmer, J. D., and Herbon, L. A. (1988). Plant mitochondrial DNA evolves rapidly in structure but slowly in sequence. *J. Mol. E* 28, 87–97.
- Pelletier G, Primard C, Vedel F, Chétrit P, Rémy R, Rousselle P, Renard M (1983). Intergeneric cytoplasmic hybridization in Cruciferae by protoplast fusion. *Mol Gen Genet* 191: 244-250.
- Pelletier G, Vedel F. and Belliard G. Cybrids in genetics and breeding. (1985) *Hereditas* V103 (s3) (49-56
- Pereira LG, Coimbra S, Oliveira H, Monteiro L, Sottomayor M. (2006). Expression of arabinogalactan proteins in pollen tubes of *Arabidopsis thaliana*. *Planta* 223, 374-380
- Perrin R, Meyer EH, Zaepfel M, Kim YJ, Mache R, Grienenberger JM, Gualberto JM, Gagliardi D. Two Exoribonucleases Act Sequentially to Process Mature 3'-Ends of *atp9* mRNAs in *Arabidopsis* Mitochondria. *Journal of Biological Chemistry* (2004) 279 (24) : 25440-25446
- Pesaresi P, Schneider A, Kleine T, Leister D (2007) Interorganellar communication. *Curr Opin Plant Biol* 10: 600–606
- Pesaresi, P., Varotto, C., Meurer, J., Jahns, P., Salamini, F., and Leister, D. (2001). Knock-out of the plastid ribosomal protein L11 in *Arabidopsis*: Effects on mRNA translation and photosynthesis. *Plant J.* 27, 179–189
- Pesaresi. Nuclear Photosynthetic Gene Expression Is Synergistically Modulated by Rates of Protein Synthesis in Chloroplasts and Mitochondria. *The Plant Cell Online* (2006) 18 (4) : 970-991
- Pfalz J, Bayraktar OA, Prikryl J, Barkan A. Site-specific binding of a PPR protein defines and stabilizes 5' and 3' mRNA termini in chloroplasts. *EMBO J* (2009) 28 (14) : 2042-2052
- Pfannschmidt T, Nilsson A, Tullberg A, Link G, Allen JF. Direct Transcriptional Control of the Chloroplast Genes *psbA* and *psaAB* Adjusts Photosynthesis to Light Energy Distribution in Plants. *IUBMB life* (1999) 48 : 271-276
- Picó FX, Méndez-Vigo B, Martínez-Zapater JM, Alonso-Blanco C. Natural Genetic Variation of *Arabidopsis thaliana* Is Geographically Structured in the Iberian Peninsula. *Genetics* (2008) 180 (2) : 1009-1021
- Pina C, Pinto F, Feijo JA, Becker JD (2005) Gene family analysis of the *Arabidopsis* pollen transcriptome reveals biological implications for cell growth, division control, and gene expression regulation. *Plant Physiol* 138: 744–756
- Platt A, Horton M, Huang YS, Li Y, Anastasio AE, Mulyati NW, Agren J, Bossdorf O, Byers D, Donohue K, Dunning M, Holub EB, Hudson A, Le Corre V, Loudet O, Roux F, Warthmann N, Weigel D, Rivero L, Scholl R, Nordborg M, Bergelson J, Borevitz JO. The Scale of Population Structure in *Arabidopsis thaliana*. *PLoS Genet* (2010) 6 (2) : e1000843
- Pogson BJ, Woo NS, Förster B, Small ID. Plastid signalling to the nucleus and beyond. *Trends in plant science* (2008) 13 (11) : 602-9
- Presgraves DC, Balagopalan L, Abmayr SM, Orr HA. Adaptive evolution drives divergence of a hybrid inviability gene between two species of *Drosophila*. *Nature* (2003) 423 (6941) : 715-719
- Price AH, Taylor A, Ripley SJ, Griffiths A, Trewavas AJ, Knight MR (1994) Oxidative signals in tobacco increases cytosolic calcium. *Plant Cell* 6: 1301–1310
- Provan J, Powell W, Hollingsworth PM. Chloroplast microsatellites: new tools for studies in plant ecology and evolution. *Trends in Ecology & Evolution* (2001) 16 (3) : 142-147
- Pusnik M, Small I, Read LK, Fabbro T, Schneider A. Pentatricopeptide Repeat Proteins in *Trypanosoma brucei* Function in Mitochondrial Ribosomes. *Molecular and Cellular Biology* (2007) 27 (19) : 6876-6888
- Rand DM, Haney RA, Fry AJ. Cytonuclear coevolution: the genomics of cooperation. *Trends in Ecology & Evolution* (2004) 19 (12) : 645-653
- Rawson, P.D. and R.S. Burton. Functional coadaptation between cytochrome c and cytochrome c oxidase

- within allopatric populations of a marine copepod. (2002). Proc Nat. Acad. Sci. USA 99:12955-12958.
- Reiter, R.S., J.G.K. Williams, K.A. Feldmann, J.A. Rafalski, S.V. Tingey and P.A. Scolink, (1992). Global and local genome mapping in *Arabidopsis thaliana* by using recombinant inbred lines and random amplified polymorphic DNAs. Proc. Nat. Acad. Sci. USA., 89: 1477-1481.
- Renner. Ueber Koppelungswechsel bei *Oenothera*. Proceedings of the 5th International Congress, (1928) Supplementband of the Z. indukt. Abstamm.-u. VererbLehre. 2: 1216-1220.
- Rentel and Knight. Oxidative Stress-Induced Calcium Signaling in Arabidopsis. *Plant Physiology* (2004) 135:1471-1479
- Rhoads and Subbaiah. Mitochondrial retrograde regulation in plants. *Mitochondrion* (2007) 7 (3) : 177-94
- Rhoads DM, Umbach AL, Subbaiah CC, Siedow JN. Mitochondrial reactive oxygen species: contribution to oxidative stress and interorganellar signaling. *Plant Physiology* (2006);141:357-36
- Rhoads, Charles S. Levings and James N. Siedow. URF13, a ligand-gated, pore-forming receptor for T-toxin in the inner membrane of *cms-T* mitochondria. (1995) JOURNAL OF BIOENERGETICS AND BIOMEMBRANES. Volume 27, Number 4, 437-445,
- Ricchetti M, Fairhead C, Dujon B (1999) Mitochondrial DNA repairs double-strand breaks in yeast chromosomes. *Nature* 402: 96–100
- Richards AJ. (1997). *Plant breeding systems*. 2nd ed. London: Chapman and Hall, London
- Richter U, Kiessling J, Hedtke B, Decker E, Reski R, Börner T, Weihe A. Two RpoT genes of *Physcomitrella patens* encode phage-type RNA polymerases with dual targeting to mitochondria and plastids. *Gene* (2002) 290:95–105
- Rivals E, Bruyere C, Toffano-Nioche C, Lecharny A. (2006). Formation of the Arabidopsis pentatricopeptide repeat family. *Plant Physiol.* 141:825–839.
- Robart AR, Zimmerly S. Group II intron retroelements: function and diversity. *Cytogenet Genome Res.* (2005) 110(1-4):589-97
- Rocak and Linder. DEAD-box proteins: the driving forces behind RNA metabolism. (2004) *Nat. Rev. Mol. Cell. Biol.* 5, 232–241.
- Rodermel. Regulatory interactions between the nucleus and the plastid: plastid-to-nucleus signaling. *Trends Plant Sci.* (2001) 6:471–478
- Rossel JB, Wilson IW, Pogson BJ (2002) Global changes in gene expression in response to high light in Arabidopsis. *Plant Physiology* 130, 1109–1120
- Rothenberg, M., Boeshore, M.L., Hanson, M.R., and Irhar, S. Intergenomic recombination of mitochondria in a somatic hybrid plant. (1985). *Curr. Genet.* 9, 615-618.. Acad.
- Rousselle P., (1981). Étude de systèmes d'androstérilité chez le colza, *Brassica napus*. Thèse de Doctorat d'Ingénieur, ENSA Rennes, 109 p.
- Rowan BA, Oldenburg DJ, Bendich AJ. The demise of chloroplast DNA in Arabidopsis. *Curr Genet.* (2004) 46(3) : 176-181
- Rusche M.L., Mogensen H.L., Zhu T., Smith S.E. The zygote and proembryo of alfalfa: quantitative, three-dimensional analysis and implications for biparental plastid inheritance. *Protoplasma* (1995) 189 : 88–100.
- Rutschow H, Ytterberg AJ, Friso G, Nilsson R, van Wijk KJ (2008) Quantitative proteomics of a chloroplast SRP54 sorting mutant and its genetic interactions with CLPC1 in Arabidopsis. *Plant Physiol* 148: 156–175
- Sabar M, Gagliardi D, Balk J, Leaver CJ. ORFB is a subunit of F1FO-ATP synthase: insight into the basis of cytoplasmic male sterility in sunflower. *EMBO Rep* (2003) 4 (4) : 381-386
- Sackton, T. B., R. Haney, and D. M. Rand. (2003). Cytonuclear coadaptation in *Drosophila*: Disruptions of cytochrome c oxidase activity in backcross genotypes. *Evolution* 57:2315-2325
- Saha D., A. M. Prasad and R. Srinivasan, (2007). Pentatricopeptide repeat protein and their emerging roles in plants. *Plant Physiol. Biochem.* 45: 521-534.
- Saisho D, Nambara E, Naito S, Tsutsumi N, Hirai A, Nakazono M Expression of the gene encoding the alternative oxidase is induced by treatment of rice with the bleaching herbicide norflurazon (1997) *Plant Mol Biol* 35:585–596
- Saisho, D., Nakazono. M., Tsutsumi, N., Hirai. A., (1997). Transcript levels of tandem- arranged alternative oxidase genes in rice are increased by low temperature. *Gene*, 203: 121 129.
- Säll T, Jakobsson M, Lind-Halldén C, Halldén C. Chloroplast DNA indicates a single origin of the allotetraploid Arabidopsis suecica. *Journal of evolutionary biology* (2003) 6 (5) : 1019–1029,
- Sanders D, Pelloux J, Brownlee C, Harper JF (2002) Calcium at the crossroads of signaling. *Plant Cell* 14 : S401–S417
- Sano, Y. (1990). The genetic nature of gametic eliminator in rice. *Genetics* 125: 183±191.
- Saraste. Oxidative Phosphorylation at the fin de siècle. *Science* (1999) 283(5407) : 1488-1493.
- Sasaki Y, Sekiguchi K, Nagano Y, Matsuno R Chloroplast envelope protein encoded by chloroplast genome. *FEBS Lett* (1993) 316: 93–98.

- Sato S, Nakamura Y, Kaneko T, Asamizu E, Tabata S. Complete structure of the chloroplast genome of *Arabidopsis thaliana*. *DNA Res* (1999) 6 (5) : 283-90
- Scheffler, I. E. (1999). *Mitochondria*. Wiley-Liss, New York. 367 :370
- Schmid KJ, Törjék O, Meyer R, Schmutz H, Hoffmann MH, Altmann T. Evidence for a large-scale population structure of *Arabidopsis thaliana* from genome-wide single nucleotide polymorphism markers. *Theor Appl Genet* (2006) 112 (6) : 1104-1114
- Smith LM, Bomblies K, Weigel D (2011) Complex evolutionary events at a tandem cluster of *Arabidopsis thaliana* genes resulting in a single-locus genetic incompatibility. *PLoS Genet* 7: e1002164
- Schmitz-Linneweber and Barkan. RNA splicing and RNA editing in chloroplasts. *Plastid Development. Topics in Current Genetics* (2007) (R. Bock., ed.) Springer, Berlin, Heidelberg 19 : 213-248
- Schmitz-Linneweber and Small. Pentatricopeptide repeat proteins: a socket set for organelle gene expression. *Trends in plant science* (2008) 13 (12) : 663-670
- Schmitz-Linneweber C, Kushnir S, Babiyshuk E, Poltnigg P, Herrmann RG, and Maier RM. (2005). The failure to edit the plastid *atpA* mRNA is the primary cause of nucleo/plastid incompatibility in *Atropa/Nicotiana* cybrids. *eingereicht. Plant Cell*. 17:1815-1828.
- Schmitz-Linneweber C, Williams-Carrier RE, Williams-Voelker PM, Kroeger TS, Vichas A, Barkan A. A Pentatricopeptide Repeat Protein Facilitates the trans-Splicing of the Maize Chloroplast rps12 Pre-mRNA. *The Plant Cell online* (2006) 18 (10) : 2650-2663
- Schnable PS and Wise RP, (1998). The molecular basis of cytoplasmic male sterility and fertility restoration. *Trends Plant Sci* 3:175–180.
- Schuster G and Gruissem W (1991) Chloroplast mRNA 3' end processing requires a nuclear-encoded RNA-binding protein. *EMBO J*, 10, 1493–1502.
- Sedbrook JC, Kronebusch PJ, Borisy GG, Trewavas AJ, Masson PH (1996) Transgenic aequorin reveals organ-specific cytosolic Ca<sup>2+</sup> responses to anoxia in *Arabidopsis thaliana* seedlings. *Plant Physiol* 111: 243–257
- Sharbel TF, Haubold B, Mitchell-Olds T. Genetic isolation by distance in *Arabidopsis thaliana*: biogeography and postglacial colonization of Europe. *Molecular Ecology* (2000) 9 (12) : 2109-18
- Shaver JM, Oldenburg DJ, Bendich AJ. Changes in chloroplast DNA during development in tobacco, *Medicago truncatula*, pea, and maize. *Planta* (2006) 224 (1) : 72-82.
- Shedge V, Arrieta-Montiel M, Christensen AC, Mackenzie SA.. Plant Mitochondrial Recombination Surveillance Requires Unusual RecA and MutS Homologs. *The Plant Cell Online* (2007) 19 (4) : 1251-1264
- Shi S, Ding D, Mei S, Wang J. A comparative light and electron microscopic analysis of microspore and tapetum development in fertile and cytoplasmic male sterile radish (2010) Volume 241, Numbers 1-4, 37-49
- Shiina T, Allison L, Maliga P. *rbcl* transcript levels in tobacco plastids are independent of light : reduced dark transcription rate is compensated by increased mRNA stability. *Plant Cell* (1998) 10 : 1713–1722
- Shinozaki K., Ohme M., Tanaka M., Wakasugi T., Hayashida N., Matsubayashi T., Zaita N., Chunwongse J., Obokata J., Yamaguchi-Shinozaki K., Ohto C., Torazawa K., Meng B. Y., Sugita M., Deno H., Kamogashira T., Yamada K., Kusuda J., Takaiwa F., Kato A., Tohdoh N., Shimada H., and Sugiura M.. The complete nucleotide sequence of the tobacco chloroplast genome: its gene organization and expression. *Molec. gen. Genet.* (1986) 202 : 1–5
- Siedow, J.N., Umbach, A.L. and Moore, A.L. (1995) The active site of the cyanide-resistant oxidase from plant mitochondria contains a binuclear iron center. *FEBS Lett.* 362: 10-14.
- Simon SA, Meyers BC. Small RNA- mediated epigenetic modifications in plants. *Curr Opin Plant Biol.* (2011) 14(2):148-55.
- Simon M, Simon A, Martins F, Botran L, Tisné S, Granier F, Loudet O, Camilleri C. DNA fingerprinting and new tools for fine-scale discrimination of *Arabidopsis thaliana* accessions. *The Plant Journal* (2011) : doi: 10.1111/j.1365-313X.2011.04852.x
- Sloan DB, Alverson AJ, Chuckalovcak JP, Wu M, McCauley DE, Palmer JD, Taylor DR. (2012). Rapid evolution of enormous, multichromosomal genomes in flowering plant mitochondria with exceptionally high mutation rates. *PLoS Biol.* 10: e1001241.
- Small and Peeters. The PPR motif-a TPR-related motif prevalent in plant organellar proteins. *Trends in biochemical sciences* (2000) 25 (2) : 46-47
- Small, I., P.G. Isaac, and C.J. Leaver. (1987). Stoichiometric differences in DNA molecules containing the *atpA* gene suggest mechanisms for the generation of mitochondrial genome diversity in maize. *EMBO J* 6:865–869.
- Small, I., Peeters, N., Legeai, F., and Lurin, C. (2004). Predotar: A tool for rapidly screening proteomes for N-terminal targeting sequences. *Proteomics* 4, 1581–1590.
- Smith. Influence of parental genotype on plastid inheritance in *Medicago sativa*. *Journal of Heredity* (1989)

80 : 214–217.

- Sodmergen Kawano S, Tano S, Kuroiwa T. Degradation of chloroplast DNA in second leaves of rice (*Oryza sativa*) before leaf yellowing. *Protoplasma* (1991) 160 : 89–98.
- Sodmergen, Zhang Q, Zhang Y, Sakamoto W, Kuroiwa T. Reduction in amounts of mitochondrial DNA in the sperm cells as a mechanism for maternal inheritance in *Hordeum vulgare*. *Planta* (2002) 216 :235–244.
- Spirek M, *et al.* (2000) Functional co-operation between the nuclei of *Saccharomyces cerevisiae* and mitochondria from other yeast species. *Curr Genet* 38(4):202-7
- Stenøien HK, Fenster CB, Tonteri A, Savolainen O. Genetic variability in natural populations of *Arabidopsis thaliana* in northern Europe. *Molecular Ecology*, (2005) 14 : 137–148.
- Stern DB, Goldschmidt-Clermont M, Hanson MR. Chloroplast RNA metabolism. *Annu Rev Plant Bio* (2010) 61 : 125–155.
- Stern, D.B., and Kindle, K.L. (1993). 3' End maturation of the *Chlamydomonas reinhardtii* chloroplast *atpB* mRNA is a two-step process.
- Strand A, Asami T, Alonso J, Ecker JR, Chory J. Chloroplast to nucleus communication triggered by accumulation of Mg-protoporphyrin IX. *Nature*. (2003);421:79–83.
- Stubbe W (1989) *Oenothera* — an ideal system for studying the interactions of genome and plastome. *Plant Mol Biol Rep*, 7, 245–257.
- Subbaiah, C.C., Bush, D.S. and Sachs, M.M. (1994a) Elevation of cytoplasmic calcium precedes anoxic gene expression in maize suspension-cultured cells. *Plant Cell* 6: 1747–1762.
- Subbaiah, C.C., Zhang, J. and Sachs, M.M. (1994b) Involvement of intracellular calcium in anaerobic gene expression and survival of maize seedlings. *Plant Physiol.* 105: 369–376.
- Susek RE, Ausubel FM, Chory J (1993) Signal transduction mutants of *Arabidopsis* uncouple nuclear CAB and RBCS gene expression from chloroplast development. *Cell* 74: 787–799
- Suzuki, J.Y., Ytterberg, A.J., Beardslee, T.A., Allison, L.A., Wijk, K.J. and Maliga, P. Affinity purification of the tobacco plastid RNA polymerase and in vitro reconstitution of the holoenzyme. *Plant J.* (2004) 40(1), 164–172.
- Swiatecka-Hagenbruch M, Liere K, Börner T. High diversity of plastidial promoters in *Arabidopsis thaliana*. *Mol. Genet. Genomics* (2007) 277 : 725–734
- Takenaka M, Verbitskiy D, van der Merwe JA, Zehrmann A, Brennicke A. The process of RNA editing in plant mitochondria. *Mitochondrion* (2008) 8 : 35–46
- Tan XY, Liu XL, Wang W, Jia DJ, Chen LQ, Zhang XQ, Ye D. Mutations in the *Arabidopsis* nuclear-encoded mitochondrial phage-type RNA polymerase gene *RPOTm* led to defects in pollen tube growth, female gametogenesis and embryo-genesis. *Plant Cell Physiol* (2010) 51:635–49.
- Tang et al. The Evolution of Selfing in *Arabidopsis thaliana*. *Science* (2007) vol. 317 (5841) pp. 1070–1072
- Tartar and Boucias. The non-photosynthetic, pathogenic green alga *Helicosporidium* sp. has retained a modified, functional plastid genome. *FEMS Microbiol. Lett.* (2004) 233, : 153–157.
- Tarutani, Y. et al. (2010). Trans-acting small RNA determines dominance relationships in *Brassica* self-incompatibility. *Nature*. 466(7309): 983–6.
- Tasaki E, Hattori M, Sugita M. The moss pentatricopeptide repeat protein with a DYW domain is responsible for RNA editing of mitochondrial *ccmF* transcript. *Plant J.* (2010) 62:560–570.
- Taylor. Symbioticism revisited : a discussion of the evolutionary impact of intracellular symbioses. *Proc. R. Soc. Lond. B. Biol. Sci.* (1979) 204 : 267–286.
- Tillich M, Hardel SL, Kupsch C, Armbruster U, Delannoy E, Gualberto JM, Lehwark P, Leister D, Small ID, Schmitz-Linneweber C. Chloroplast ribonucleoprotein CP31A is required for editing and stability of specific chloroplast mRNAs. *Proc Natl Acad Sci USA* (2009) 106 (14) : 6002–7
- Timmis JN, Ayliffe MA, Huang CY, Martin W. Endosymbiotic gene transfer : organelle genomes forge eukaryotic chromosomes. *Nature Reviews Genetics* (2004) 5 (2) : 123–135.
- Todokoro S, Terauchi R, Kawano S. Micro-satellite polymorphisms in natural populations of *Arabidopsis thaliana* in Japan. *Jap J Genet* (1995) 70 : 543–554.
- Töller A, Brownfield L, Neu C, Twell D, Schulze-Lefert P (2008) Dual function of *Arabidopsis* glucan synthase-like genes *GSL8* and *GSL10* in male gametophyte development and plant growth. *Plant J* 54: 911–923
- Tracy and Stern. Mitochondrial transcription initiation: promoter structures and RNA polymerases. *Curr Genet* (1995) 28:205–16.
- Trewavas A (1999) Le calcium, c'est la vie: Calcium makes waves. *Plant Physiol* 20:1–6.
- Trontin C, Tisné S, Bach L, Loudet O. What does *Arabidopsis* natural variation teach us (and does not teach us) about adaptation in plants?. *Current Opinion in Plant Biology* (2011) 14 (3) : 225–231
- Tsuji, Y., Ayaki, H., Whitman, S. P., Morrow, C. S., Torti, S. V., and Torti, F. M. (2000). Coordinate transcriptional and translational regulation of ferritin in response to oxidative stress. *Mol. Cell Biol.* 20,

- 5818–5827.
- Tuinstra, M.R., G. Ejeta, and P.B. Goldsbrough. (1997) Heterogeneous inbred family (HIF) analysis: An approach for developing near-isogenic lines that differ at quantitative trait Theoretical and Applied Genetics 95:1005-1011
- Ullrich H, Lattig K, Brennicke A, Knoop V (1997). Mitochondrial DNA variations and nuclear RFLPs reflect different genetic similarities among 23 *Arabidopsis thaliana* ecotypes. *Plant Mol Biol* 33: 37–45.
- Unsel M, Marienfeld JR, Brandt P, Brennicke A (1997) The mitochondrial genome of *Arabidopsis thaliana* contains 57 genes in 366,924 nucleotides. *Nat Genet* 15: 57–61
- Uyttewaal M, Mireau H, Rurek M, Hammani K, Arnal N, Quadrado M, Giegé P. PPR336 is associated with polysomes in plant mitochondria. *J Mol Biol* (2008a) 375 (3) : 626-36
- Uyttewaal M, Arnal N, Quadrado M, Martin-Canadell A, Vrielynck N, Hiard S, Gherbi H, Bendahmane A, Budar F, Mireau H. Characterization of *Raphanus sativus* Pentatricopeptide Repeat Proteins Encoded by the Fertility Restorer Locus for *Ogura* Cytoplasmic Male Sterility. *The Plant Cell Online* (2008b) 20 (12) : 3331-3345
- Vandecasteele G., Szabadkai G., Rizzuto R. Mitochondrial calcium homeostasis: mechanisms and molecules. *IUBMB Life* (2001) 52 : 213-219.
- Vanlerberghe GC, McIntosh L (1994) Mitochondrial electron transport regulation of nuclear gene expression: studies with the alternative oxidase gene of tobacco. *Plant Physiol* 105: 867–874
- Vanlerberghe, G.C., McIntosh, L. (1994). Mitochondrial electron transport regulation of nuclear gene expression: Studies with the alternative oxidase gene of tobacco. *Plant Physiol.* 105, 867-874.
- Vedel F, Pla M, Vitart V, Gutierrez S, Chetrit P and De Paep R. Molecular basis of nuclear and cytoplasmic male sterility in higher plants. *Plant Physiol. Biochem.* (1994) 32 : 601-618.
- Verbitskiy D, Härtel B, Zehrmann A, Brennicke A, Takenaka M. The DYW-E-PPR protein MEF14 is required for RNA editing at site matR-1895 in mitochondria of *Arabidopsis thaliana*. *FEBS Lett* (2011) 585 (4) : 700-704
- Vlad, D., F. Rappaport, M. Simon and O. Loudet, (2010) Gene transposition causing natural variation for growth in *Arabidopsis thaliana*. *PLoS Genet.* 6: e1000945.
- Wallin. *Symbiogenesis and the Origin of Species* (1927) Williams & Wilkins, Baltimore, MD.
- Walter M, Piepenburg K, Schöttler MA, Petersen K, Kahlau S, Tiller N, Drechsel O, Weingartner M, Kudla J, Bock R (2010) Knockout of the plastid RNase E leads to defective RNA processing and chloroplast ribosome deficiency. *Plant J* 64: 851–863
- Wang Z, Zou Y, Li X, Zhang Q, Chen L, Wu H, Su D, Chen Y, Guo J, Luo D, Long Y, Zhong Y, Liu YG. Cytoplasmic male sterility of rice with boro II cytoplasm is caused by a cytotoxic peptide and is restored by two related PPR motif genes via distinct modes of mRNA silencing. *The Plant Cell Online* (2006) 18 (3) : 676
- Weigel and Mott. The 1001 Genomes Project for *Arabidopsis thaliana*. *Genome Biol* (2009) 10 (5) : 107
- Weigel. Natural variation in *Arabidopsis thaliana*: from molecular genetics to ecological genomics. *PLANT PHYSIOLOGY* (2011) :
- Weihe. The transcription of plant organelle genomes. In: Daniell H, Chase CD, editors. *Molecular biology and biotechnology of plant organelles*. Dordrecht: Kluwer Academic Publishers (2004) 213–37.
- Wen, L., and Chase, C.D. (1999a). Pleiotropic effects of a nuclear restorer-of-fertility locus on mitochondrial transcripts in male-fertile and S male-sterile maize. *Curr. Genet.* 35, 521–526
- Willett, C. S., and R. S. Burton, (2003). Environmental influences on epistatic interactions: viabilities of cytochrome c genotypes in interpopulation crosses. *Evolution* 57:2286–2292.
- Williams L, Zhao J, Morozova N, Li Y, Avivi Y, Grafi G. Chromatin reorganization accompanying cellular dedifferentiation is associated with modifications of histone H3, redistribution of HP1, and activation of E2F-target genes. *Developmental dynamics* (2003) 228 (1) : 113-120
- Williams-Carrier R, Kroeger T, Barkan A . Sequence-specific binding of a chloroplast pentatricopeptide repeat protein to its native group II intron ligand. *RNA* (2008) 14 (9) : 1930-41
- Wise, R.P., Gobelman-Werner, K., Pei, D., Dill, C.L., and Schnable, P.S. (1999). Mitochondrial transcript processing and restoration of male fertility in T-cytoplasm maize. *J. Hered.* 90, 380–385.
- Wolf PG, Rowe CA, Hasebe M. High levels of RNA editing in a vascular plant chloroplast genome : analysis of transcripts from the fern *Adiantum capillus-veneris*. *Gene* (2004) 339 : 89–97
- Wolfe KH, Li WH, Sharp PM (1987) Rates of nucleotide substitution vary greatly among plant mitochondrial, chloroplast, and nuclear DNAs. *Proc Natl Acad Sci USA* 84 : 9054-9058
- Wollman FA, Minai L, Nechushtai R (1999) The biogenesis and assembly of photosynthetic proteins in thylakoid membranes. *Biochim Biophys Acta* 1141: 21–85
- Woodson and Chory. Coordination of gene expression between organellar and nuclear genomes. *Nature Reviews Genetics* (2008) 9 (5) : 383-395

- Woodson JD, Perez-Ruiz JM, Chory J. Heme synthesis by plastid ferrochelatase I regulates nuclear gene expression in plants. *Curr. Biol* (2011) 21:897-903.
- Wostrikoff, K., Girard-Bascou, J., Wollman, F.A., and Choquet, Y.(2004). Biogenesis of PSI involves a cascade of translational autoregulation in the chloroplast of *Chlamydomonas*. *EMBO J*. 23: 2696–2705.
- Wright. Evolution in Mendelian populations. (1931) *Genetics* 16 :97-159
- Wu, H., and Cheung, A.Y. (2000). Programmed cell death in plant reproduction. *Plant Mol. Biol.* 44, 267–281.
- Wu M, Sun LV, Vamathevan J, Riegler M, Deboy R, et al. 2004 Phylogenomics of the Reproductive Parasite *Wolbachia pipientis* wMel: A Streamlined Genome Overrun by Mobile Genetic Elements. *PLoS Biol* 2(3): e69.
- Xianzhou Nie, Douglas C. Durnin, Abir U. Igamberdiev and Robert D. Hill. Cytosolic calcium is involved in the regulation of barley hemoglobin gene expression. *PLANTA* (2006) 223 (3) : 542-549.
- Xiao S., Hsieh J., Nugent R.L., Coughlin D.J., Fierke C.A., Engelke D.R. Functional characterization of the conserved amino acids in Pop1p, the largest common protein subunit of yeast RNases P and MRP. *RNA*. (2006) 12:1023–1037.
- Yang J, Stern DB (1997) The spinach chloroplast endoribonuclease CSP41 cleaves the 3'-untranslated region of petD mRNA primarily within its terminal stem-loop structure. *J. Biol. Chem.* 272: 12874-12880.
- Yao N, Tada Y, Sakamoto M, Nakayashiki H, Park P, Tosa Y, Mayama S (2002) Mitochondrial oxidative burst involved in apoptotic response in oats. *Plant J* 30: 567–579
- Yehudai-Resheff S, Portnoy V, Yogev S, Adir N, Schuster G. Domain analysis of the chloroplast polynucleotide phosphorylase reveals discrete functions in RNA degradation, polyadenylation, and sequence homology with exosome proteins. *Plant Cell*. (2003) 15(9):2003-19.
- Yin C, Richter U, Börner T, Weihe A. Evolution of plant phage-type RNA polymerases: the genome of the basal angiosperm *Nuphar advena* encodes two mitochondrial and one plastid phage-type RNA polymerases. *BMC Evol Biol* (2010) 10 (1) : 379
- Yin P, Kang J, He F, Qu LJ, Gu H. The origin of populations of *Arabidopsis thaliana* in China, based on the chloroplast DNA sequences. *BMC Plant Biol*. (2010) : 10-22
- Yoshinaga K, Iinuma H, Masuzawa T, Uedal K. Extensive RNA editing of U to C in addition to C to U substitutions in the *rbcL* transcript of hornwort chloroplasts and the origin of RNA editing in green plants. *Nucleic Acids Res* (1996) 24 : 1008–1014.
- Youle et Karbowski. Mitochondrial fission in apoptosis. *Nature Rev Mol Cell Biol* (2005) 6 : 657-663.
- Young EG, Hanson MR. A fused mitochondrial gene associated with cytoplasmic male sterility is developmentally regulated. *Cell*. (1987) Jul 3;50(1):41–49.
- Yu J, Nickels R, McIntosh L (2001) A genome approach to mitochondrial-nuclear communication in *Arabidopsis*. *Plant Physiol Biochem* 39: 345–353
- Zarkovic J, Anderson SL, Rhoads DM (2005) A reporter gene system used to study developmental expression of alternative oxidase and isolate mitochondrial retrograde regulation mutants in *Arabidopsis*. *Plant Mol Biol* 57: 871–888
- Zimmer SL, Schein A, Zipor G, Stern DB, Schuster G. Polyadenylation in *Arabidopsis* and *Chlamydomonas* organelles: the input of nucleotidyltransferases, poly(A) polymerases and polynucleotide phosphorylase. (2009) *The Plant Journal* Volume 59, Issue 1, pages 88–99,
- Zoschke R, Nakamura M, Liere K, Sugiura M, Börner T, Schmitz-Linneweber C (2010) An organellar maturase associates with multiple group II introns. *Proc Natl Acad Sci USA* 107:3245–3250

# **ANNEXES**

Annexe 1 : fiches descriptives d'accessions d'*Arabidopsis*

Accession Mr-0																																																					
<p><b>IDENTIFICATION :</b></p> <p style="text-align: right;">Versailles Number : 148                      AV                      Original Stock Center : N1372                      Arabidopsis Identification Stock number                      (AIS) : 150</p>																																																					
<p><b>COLLECTION DATA :</b></p>	<p><b>LOCATION :</b></p> <p style="text-align: right;">Country : Italy                      City : Monte/Tosso                      Altitude : 1000-1500                      Latitude : N 44°30'                      Longitude : E 9°30'                      Habitat :</p>																																																				
<p><b>GENERAL DESCRIPTION OF ACCESSION (just for information, these data should be verified ) :</b></p> <p>Mr-0, ecotype from Monte/Tosso (Italy); large rosette, leaf margins serrated, late flowering, height = 50-52 cm; Altitude (m): 1000-1500; Longitude/Latitude: E9-E10/N44-N45; Collector: Staatssammlung (1952); single seed stock#: CS6795.</p>																																																					
<p><b>CLIMATOLOGY :</b></p>																																																					
 <table border="1" style="margin: 5px auto; font-size: small;"> <tr><th>Month</th><td>J</td><td>F</td><td>M</td><td>A</td><td>M</td><td>J</td><td>J</td><td>A</td><td>S</td><td>O</td><td>N</td><td>D</td></tr> <tr><th>Hours</th><td>35</td><td>42</td><td>42</td><td>47</td><td>51</td><td>55</td><td>64</td><td>61</td><td>55</td><td>48</td><td>33</td><td>31</td></tr> </table> <p>MONTHLY HOURS OF SUNSHINE (hr)</p>	Month	J	F	M	A	M	J	J	A	S	O	N	D	Hours	35	42	42	47	51	55	64	61	55	48	33	31	 <table border="1" style="margin: 5px auto; font-size: small;"> <tr><th>Month</th><td>J</td><td>F</td><td>M</td><td>A</td><td>M</td><td>J</td><td>J</td><td>A</td><td>S</td><td>O</td><td>N</td><td>D</td></tr> <tr><th>Precipitation</th><td>89</td><td>89</td><td>101</td><td>93</td><td>117</td><td>79</td><td>42</td><td>94</td><td>100</td><td>143</td><td>109</td><td>115</td></tr> </table> <p>MONTHLY PRECIPITATION (mm)</p>	Month	J	F	M	A	M	J	J	A	S	O	N	D	Precipitation	89	89	101	93	117	79	42	94	100	143	109	115
Month	J	F	M	A	M	J	J	A	S	O	N	D																																									
Hours	35	42	42	47	51	55	64	61	55	48	33	31																																									
Month	J	F	M	A	M	J	J	A	S	O	N	D																																									
Precipitation	89	89	101	93	117	79	42	94	100	143	109	115																																									
 <table border="1" style="margin: 5px auto; font-size: small;"> <tr><th>Month</th><td>J</td><td>F</td><td>M</td><td>A</td><td>M</td><td>J</td><td>J</td><td>A</td><td>S</td><td>O</td><td>N</td><td>D</td></tr> <tr><th>Rain Days</th><td>7</td><td>6</td><td>8</td><td>8</td><td>10</td><td>7</td><td>5</td><td>5</td><td>7</td><td>8</td><td>9</td><td>8</td></tr> </table> <p>MEAN MONTHLY NUMBER OF RAIN DAYS</p>	Month	J	F	M	A	M	J	J	A	S	O	N	D	Rain Days	7	6	8	8	10	7	5	5	7	8	9	8	 <table border="1" style="margin: 5px auto; font-size: small;"> <tr><th>Month</th><td>J</td><td>F</td><td>M</td><td>A</td><td>M</td><td>J</td><td>J</td><td>A</td><td>S</td><td>O</td><td>N</td><td>D</td></tr> <tr><th>Temperature</th><td>1.7</td><td>3.1</td><td>6.3</td><td>10.2</td><td>14</td><td>18.3</td><td>21</td><td>20.4</td><td>17</td><td>11.9</td><td>7</td><td>3.3</td></tr> </table> <p>MEAN MONTHLY TEMPERATURE (°C)</p>	Month	J	F	M	A	M	J	J	A	S	O	N	D	Temperature	1.7	3.1	6.3	10.2	14	18.3	21	20.4	17	11.9	7	3.3
Month	J	F	M	A	M	J	J	A	S	O	N	D																																									
Rain Days	7	6	8	8	10	7	5	5	7	8	9	8																																									
Month	J	F	M	A	M	J	J	A	S	O	N	D																																									
Temperature	1.7	3.1	6.3	10.2	14	18.3	21	20.4	17	11.9	7	3.3																																									
 <table border="1" style="margin: 5px auto; font-size: small;"> <tr><th>Month</th><td>J</td><td>F</td><td>M</td><td>A</td><td>M</td><td>J</td><td>J</td><td>A</td><td>S</td><td>O</td><td>N</td><td>D</td></tr> <tr><th>Temp Range</th><td>5.7</td><td>6.8</td><td>8.2</td><td>8.9</td><td>9.4</td><td>9.8</td><td>10.4</td><td>9.9</td><td>9</td><td>7.2</td><td>5.4</td><td>5.1</td></tr> </table> <p>MEAN MONTHLY TEMP RANGE (°C)</p>	Month	J	F	M	A	M	J	J	A	S	O	N	D	Temp Range	5.7	6.8	8.2	8.9	9.4	9.8	10.4	9.9	9	7.2	5.4	5.1																											
Month	J	F	M	A	M	J	J	A	S	O	N	D																																									
Temp Range	5.7	6.8	8.2	8.9	9.4	9.8	10.4	9.9	9	7.2	5.4	5.1																																									

## Accession Hodja-Obi-Garm bis

### IDENTIFICATION :

Versailles Number : 270 AV  
Original Stock Center : CS6179  
Arabidopsis Identification Stock number  
(AIS) :

### COLLECTION DATA :

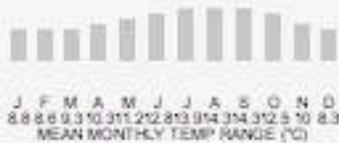
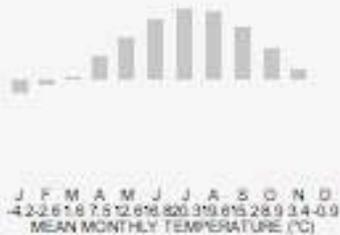
### LOCATION :

Country : Tadjikistan  
City : Hodja-Obi-Garm  
Altitude : 2000  
Latitude : N 38°55'  
Longitude : E 68°47'  
Habitat :

### GENERAL DESCRIPTION OF ACCESSION (just for information, these data should be verified) :

Ecotype from Tadjikistan; single seed propagation from CS922; Collector:  
K. Khurmatov.

### CLIMATOLOGY :



Accession 9481B

IDENTIFICATION :

Versailles Number : 261 AV  
Original Stock Center : CS22460  
Arabidopsis Identification Stock  
number (AIS) :

COLLECTION DATA :

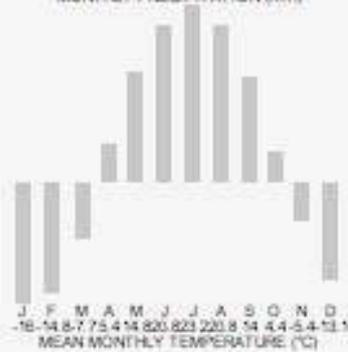
LOCATION :

Country : Kazakhstan  
City : Kazakhstan  
Altitude :  
Latitude : N 48°00'  
Longitude : E 68°00'  
Habitat :

GENERAL DESCRIPTION OF ACCESSION (just for information, these data should be verified ) :

Ecotype from Kazakhstan. Large rosettes, wide leaves with rounded tips and elongated petioles, leaf margins slightly serrated, variability in flowering time, height = 38-45 cm; original seeds collected from single plant.

CLIMATOLOGY :



## Accession Sorbo

### IDENTIFICATION :

Versailles Number : 238  
AV  
Original Stock Center : CS931  
Arabidopsis Identification Stock number  
(AIS) :

### COLLECTION DATA :

### LOCATION :

Country : Tajikistan  
City : Sorbo  
Altitude : 2100-2200  
Latitude : N 38°49'  
Longitude : E 69°28'  
Habitat : mountains

### GENERAL DESCRIPTION OF ACCESSION (just for information, these data should be verified) :

Ecotype from Tajikistan (Southern slopes of Gisserky (Hissar ?) mountain ridge (West Pamiro-Alay) ; possibly late flowering; Habitat: mountains; Collector: K. Khurmatov

### CLIMATOLOGY :



## Accession Kondara

### IDENTIFICATION :

Versailles Number : 190  
 AV  
 Original Stock Center : N916  
 Arabidopsis Identification Stock number  
 (AIS) :

### COLLECTION DATA :

Core collection : 40

### LOCATION :

Country : Tadjikistan  
 City : Kondara  
 Altitude : 1000-1100  
 Latitude : N 38°48'  
 Longitude : E 68°49'  
 Habitat :

### GENERAL DESCRIPTION OF ACCESSION (just for information, these data should be verified) :

Ecotype from Tadjikistan (Kondara : 38.48°43'N/68.49°42'E ; Southern slopes of Gisserky (Hissar ?) mountain ridge (West Pamiro-Alay) ; dark green rosette, leaf margins serrated, height = 33 cm; Collector: K. Khumatov; single seed stock#: CS6175.

### CLIMATOLOGY :



J F M A M J J A S O N D  
 37 32 28 34 40 70 86 86 69 48 40  
 MONTHLY HOURS OF SUNSHINE (h)



J F M A M J J A S O N D  
 105 105 104 174 107 32 9 6 4 35 89 106  
 MONTHLY PRECIPITATION (mm)



J F M A M J J A S O N D  
 10 11 13 12 11 7 3 2 1 4 7 10  
 MEAN MONTHLY NUMBER OF RAIN DAYS



J F M A M J J A S O N D  
 -4.2 -2.8 1.6 7.5 12.6 20.3 21.9 21.5 21.9 21.4 14.0 0.9  
 MEAN MONTHLY TEMPERATURE (°C)



J F M A M J J A S O N D  
 8.8 8.8 9.3 10.9 11.2 12.8 13.9 14.3 14.3 12.5 10.8 8.3  
 MEAN MONTHLY TEMP RANGE (°C)

## Accession Wil-1

### IDENTIFICATION :

Versailles Number : 72 AV  
Original Stock Center : N1594  
Arabidopsis Identification Stock number  
(AIS) : 221

### COLLECTION DATA :

### LOCATION :

Country : Lithuania  
City : Vilnius  
Altitude : 100-200  
Latitude : N 54°41'  
Longitude : E 25°19'  
Habitat : near Towniskaya

### GENERAL DESCRIPTION OF ACCESSION (just for information, these data should be verified) :

Wil-1, ecotype from Wilna/Litauen (Russia); light green plant, leaf margins serrated, height = 36-41 cm, habitat: near Towniskaia; Altitude (m): 100-200; Longitude/Latitude: E25/N55; Daily Centigrade Temperature (Spr/Aut): <0-2/9-10; Collector: Johuber; single seed stock#: CS6888.

### CLIMATOLOGY :



Annexe 2 : amorces, dessinées par moi-même et qui amplifient des produits non polymorphes entre les allèles Sh: La position du MSAT sur le génome de Col. à est indiquée dans le nom de l'amorce (par exemple, MSAT3.187621 amplifie un MSAT situé sur le chromosome 3 à 18,7621 Mb (d'après TAIR9))

Couples d'amorces pour amplifier les MSAT du chromosome 3

Nom de l'amorce	Séquence de l'amorce (5'→3')	Nom de l'amorce	Séquence de l'amorce (5'→3')
MSAT 3.187621-F	AGTTGGGAAAGTTGTGGTTGA	MSAT 3.198514-F	CTGCACGTTTTCTGCCATA
MSAT 3.187621-R	CACACAACCTTGGCTTAGGA	MSAT 3.198514-R	TCATGCACACCATTTTCGTC
MSAT 3.187764-F	GGGCGTTACAAGAGAATGGA	MSAT 3.198573-F	GAGGCAACCCAATACTTCCA
MSAT 3.187764-R	TTCTTGCAACAACAAAAAGGA	MSAT 3.198573-R	GGTTTTGGTGATACCCAAATGA
MSAT 3.188656-F	GGGGGTCAAACCTAACTGTTG	MSAT 3.199222-F	TAACGATCGCCTCGTGATTT
MSAT 3.188656-R	CAATTTAAGGCTGGTCTGTTC	MSAT 3.199222-R	ACGCGATAAAAAGGTCACCA
MSAT 3.188810-F	ACAGCTCAAAATATGGAATCG	MSAT 3.199749-F	TGAAAACGTCGGTTTATTCAGT
MSAT 3.188810-R	TTTGTGATTTGATGGGCTTTTT	MSAT 3.199749-R	AGCGGAGGGAAGGAGATGATT
MSAT 3.189364-F	GGTTTACCCTGCGAGACGA	MSAT 3.201438-F	ATTGAAGCGGTGCTTTTGT
MSAT 3.189364-R	TTTCTCAACTGAATGTCCAACA	MSAT 3.201438-R	GCATACTAAGCCGGTGGATG
MSAT 3.189718-F	TCCAAGGCCAAAGCCAATAC	MSAT 3.202057-F	CATGACGTTTTCTGTCTCTC
MSAT 3.189718-R	GGAATTGAAAAGAAAGCGAAA	MSAT 3.202057-R	CGTCCAAAATCCCTGAGAA
MSAT 3.194659-F	TGAAAACCGACGGAACATAGA	MSAT 3.202479-F	AGAAGCTCAGGTTGCTGCTC
MSAT 3.194659-R	GGTGGCTTCTCCCTTTCTCT	MSAT 3.202479-R	AAGCTCGAAGCTTGGAAAA
MSAT 3.194894-F	GCCGCTTTTTCTTTGTTTCTT	MSAT 3.203055-F	TCCAATGAGTTTTCTTTGAGA
MSAT 3.194894-R	GCCATTGAGTAAGGCTTGGGA	MSAT 3.203055-R	TGGGGGAGAAGCTTTTCAGA
MSAT 3.195493-F	GGTAAGCCCGAAAGTGTGTC	MSAT 3.203698-F	ATCTTGCAAAGAGGCTTCAA
MSAT 3.195493-R	CCATTTGAGCTCGAAGACAA	MSAT 3.203698-R	TCCACATAAGCAAAGGCAAT
MSAT 3.195754-F	GAAGGCAAGGCATTGTTTTACT	MSAT 3.203778-F	ACGTGTCACTCGTTCATTCG
MSAT 3.195754-R	TGTTCAAACCTGCATGGAA	MSAT 3.203778-R	CATTCTCTGCACGCGAGTTA
MSAT 3.195835-F	GCAAACCGTGCATGAAGTTA	MSAT3.18762A-F	AGTTGGGAAAGTTGTGGTTGA
MSAT 3.195835-R	GCAATGGATAATAACCGACGA	MSAT3.18762A-R	CACACAACCTTGGCTTAGGA
MSAT 3.196784-F	GCAAATGGTGGTAGAAAAACG	MSAT3.18763A-F	CGCTATTGCCATTTGGATTT
MSAT 3.196784-R	TTTTCGCTTTCTGTTGACGA	MSAT3.18763A-R	CGTCGTTTGACCAATAACCA
MSAT 3.196901-F	AGCGATGATCAGAAACAGCA	MSAT3.18776A-F	GGTGATATCGAAGGGCGTTA
MSAT 3.196901-R	TGACGTCAACCTAAAGCCAAT	MSAT3.18776A-R	GGATCTTTCAAATACGGAGGA
MSAT 3.197532-F	GGGAAATGAAGAAGAGTGATGG	MSAT3.18A-L	TTGTGTAACAACATACATAAACAC
MSAT 3.197532-R	ACCCTGCAAAAGCTGACACT	MSAT3.18A-R	AAAGAAATATTGTTTTCCATGCAC
MSAT 3.197685-F	CTGGATCTTCGGGAGTCAAA	MSAT3.18B-L	AGAGGCCAAATGAGCTTGA
MSAT 3.197685-R	GCTGGTGAAGCATGCAACTA	MSAT3.18B-R	AAAACGAAACGGTCCCAAAT

MSAT3.18C-L	AAAGGGAGAAAGAGGCCAAA
MSAT3.18C-R	AAAACGAAACGGTCCCAAAT
MSAT3.18-F	TACCTCAAAGAGCAAACA
MSAT3.18-R	TCATACCTACATATTGCCCT
MSAT3.19003A-F	CCTCCACACTTGATTTCTTGG
MSAT3.19003A-R	TCCTAGCTTGGTTAAACCTTCA
MSAT3.19033A-F	GAGCTCCTTTCCGTTCTGTT
MSAT3.19033A-R	CACACGGTCCAAAATCATTG
MSAT3.19176A-F	CACCACATGGATGTGTACGA
MSAT3.19176A-R	CCATATTGGACCACACACACA
MSAT3.19177A-F	CCACTTTCGCTTTCTAGCTTT
MSAT3.19177A-R	AAAGCGAAGTCTGTCCACGA
MSAT3.19206A-F	GAGGATATGGGTCCGCTAAA
MSAT3.19206A-R	AGTGACAGGGGTTGTTACTCTG
MSAT3.19208A-F	TAGGCTTGCAGATGGTTTCA
MSAT3.19208A-R	TGGGCCACTCTGATTTGTCA
MSAT3.19404A-F	GAAACCCGGAAGAGAAATCC
MSAT3.19404A-R	TTGGACCCCTTCTTCTTCAA
MSAT3.19558A-F	CAAAGGGAGGCCATGAAATA
MSAT3.19558A-R	CAACATGGCGAGAGAGAAAA
MSAT3.19575A-F	GAAGGCAAGGCATTGTTTTACT
MSAT3.19575A-R	TGTTCAAACCTGCATGGAA
MSAT3.19610A-F	AGACAACAATATGGTGGGACA
MSAT3.19610A-R	GCATCTTAAACTTTTCCGCATC
MSAT3.19765A-F	CGGGACGACTAAAATCGAGA
MSAT3.19765A-R	CCAAATCATGTCCCTTGTCA
MSAT3.19768A-F	CTGGATCTTCGGGAGTCAAA
MSAT3.19768A-R	GCTGGTGAAGCATGCAACTA
MSAT3.19789A-F	AAGCAGGACCAACCAAGAAA
MSAT3.19789A-R	CGGTGTCTTCAACCTGCTATT
MSAT3.19969A-F	GTGCCACCACATCCTTTGAT
MSAT3.19969A-R	TTCAAACGCAATGTAGCAGA
MSAT3.19974A.1-L	ATTCCTCTCTCCGTTTCGT
MSAT3.19974A.1-R	TGAAAACGTCGGTTTATTCAGTT
MSAT3.19974A.2-R	TCGGTTTATTCAGTTTGGTTGC
MSAT3.19974A.2-L	ATTCCTCTCTCCGTTTCGT

MSAT3.19974A.3-L	ATTCCTCTCTCCGTTTCGT
MSAT3.19974A.3-R	ATGCCGATTTTAGTCGATG
MSAT3.19974A.4-R	CATGCCGATTTTAGTCGAT
MSAT3.19974A.4-L	ATTCCTCTCTCCGTTTCGT
MSAT3.19974A-F	CGGTTTATTCAGTTATGGTTGC
MSAT3.19974A-R	AGCGGAGGGAAGGAGATGATT
MSAT3.199898-L	CTGCTCACGCTTCCATTAT
MSAT3.199898-R	TCTCAACGAAATGGATAAGGTC
MSAT3.199938-L	TGTCAACGAAATCAAATGG
MSAT3.199938-R	AAAGCCACATCACATGGACA
MSAT3.1999L	CCTAACCCACATCCACGAG
MSAT3.1999R	CACATTTGCATTAAGGGATCG
MSAT3.2000L	AATAATGGGAAGCGTGAGCA
MSAT3.2000R	TGCGATGATTTATTTGGTAGGAC
MSAT3.2004L	TTCTCCAAATTCGTCCTGTCTCT
MSAT3.2004R	TGTGCTCGAAGCTACTGTGG
MSAT3.200654-L	GCTGTCTTTACGACGTTCCGAC
MSAT3.200654-R	AAATGGGAAGAATTTGATGAGA
MSAT3.2006L	GGTGGCGAAAAGAAATGAAA
MSAT3.2006R	CCTATTTTTGGGCGTCAAG
MSAT3.20073L	GCAAAGAAATGAGATGTTTTGG
MSAT3.20073R	AGAGAGCTCGAGCGATGAAG
MSAT3.200889-L	CGTCTTCTCTTCCACGAAA
MSAT3.200889-R	TCTTGTCTCTTCCGCTTG
MSAT3.2014L	AACACCATCCAATCCAACG
MSAT3.2014R	CAACAGCTGGAAATCGAACA
MSAT3.2015L	ACTCCACACGTTCTCTGCT
MSAT3.2015R	TTGCTCTCCATCCACAAAAA
MSAT3.20162A-F	ATGTTCCATGGGAAAGCCAAT
MSAT3.20162A-R	CCTTGAGGACAAGGCTGTTT
MSAT3.20175A-F	CCTGACTGCTCCTTTCTTGTG
MSAT3.20175A-R	GTTTCGAATGGGGGAAAAAC
MSAT3.20175L	GTTTCGAATGGGGGAAAAAC
MSAT3.20175R	GGTTGCGGTGTTAGAGAGGA
MSAT3.20205A-F	CATGACGTTTTCTGTCTCTC
MSAT3.20205A-R	CGTCCAAAATCCCTGAGAA

MSAT3.20213A-F	TCGTGACGAAATCGAGAGTC
MSAT3.20213A-R	CGGTTTAAAATTTGTTGGATCA
MSAT3.20213L	CAACACATAAACGGATTAAGACG
MSAT3.20213R	AACGGAATATTCTGGCATCG
MSAT3.20218A-F	CTTTTTCGAAGGTCATCCTGA
MSAT3.20218A-R	CCGAACCTGTGTATCAAATCA
MSAT3.20218L	GGATTCTCGGTTTGAGAGCA
MSAT3.20218R	GCTTTTGGGAAAAGCGTTG
MSAT3.20295A-R	TGTTCAAGAAACACTACATTTACA
MSAT3.20295L	GGAAATTTTTGCTCTTCATGGATT
MSAT3.20295R	GCTTAAGGAAACATGACTCCAG
MSAT3.20298A-F	TTTAGCTTGTGCGCAAAACA
MSAT3.20298A-R	TTCTTTGTATTTTATTGTGAGGAT
MSAT3.20298L	TTCTCTCTCTCTTTTGTCTTTCA
MSAT3.20298R	TTTAGCTTGTGCGCAAAACA
MSAT3.20304A-F	TATTGGGCCAAAAGAGCTTC
MSAT3.20304A-R	CACATTTTACTTTTCTGTTCCAG
MSAT3.20304L	CACATTTTACTTTTCTGTTCCAG
MSAT3.20304R	TATTGGGCCAAAAGAGCTTC
MSAT3.20306L	TCCCCTAGGACGATGTTTTCT
MSAT3.20306R	TGCACAGTGTTCTCCACAAA
MSAT3.20307L	ATCTTTGGCCCACCTTTTTT
MSAT3.20307R	GGTGGAGGCATTTATTTT
MSAT3.20329A-F	TTTCCAGGACAAGTCTGAGAGG
MSAT3.20329A-R	CCTAACTTTTCTGCTCCCAACC
MSAT3.20329L	CCTAACTTTTCTGCTCCCAACC
MSAT3.20329R	ATCCCCAGCTTGGAGAATG
MSAT3.203541A-F	CATGACGCGTGTGAAAGTT
MSAT3.203541A-R	TGGTCCAACAAGAGAAACATC
MSAT3.203541-L	GGCCAAGAATTTGACCATC
MSAT3.203541-R	GACAATGAAGCTGCATGACG
MSAT3.203701-L	AAGGAATAATCCCCAAAACGA
MSAT3.203701-R	ACGAAGACGAATCCGTGAAG
MSAT3.203745-L	AAACGGACGATGTTTTCAAG
MSAT3.203745-R	AAGAATAACTTTTATCCAGTTCAAT
MSAT3.20402-L	TTTGTCAAACCTTATTGAGAAAAAGA
MSAT3.20402-R	TGGGCTTATAAACTTAGTGGTGA

MSAT3.204078-L	TCGACACATAGTTGCCCGTA
MSAT3.204078-R	AAATATGGCCAAAGTATCAATGC
MSAT3.20435 F	TCCACTCTTGTTCGTGTATTTCT
MSAT3.20435 R	CATTATTTAGTGTCTCGTATCGTA
MSAT3.20435.1F	TCCTTGCTGTTTCGGGAGTAA
MSAT3.20435.1R	TCGTGTATTTCAACCCCAACA
MSAT3.20435-L	TCCACTCTTGTTCGTGTATTTCT
MSAT3.20435-R	CATTATTTAGTGTCTCGTATCGTA
MSAT3.204559-L	ATCGATCGTCTCGTGTCTCT
MSAT3.204559-R	CACCATTGGATGTGGCTTTA
MSAT3.204600-L	TCATGAAGCAAAAACCTCAAA
MSAT3.204600-R	GAGACAATCACAAGAACATCCAA
MSAT3.205261-L	GGAGGTGCCTAGGGTTTCTC
MSAT3.205261-R	ATTCACTCCTTGTCCGATCC
MSAT3.205265-L	GGATCGGACAAGGAGTGAAT
MSAT3.205265-R	AAGCAAGTAATAATTTGGAGCATC
MSAT3.205371-L	GGAGGTGCCTAGGGTTTCTC
MSAT3.205371-R	ATTCACTCCTTGTCCGATCC
MSAT3.205581-L	GAATCTGCTCCACCCAAAAA
MSAT3.205581-R	CCAAAAAGATCACGGATTTGA
MSAT3.205691-L	GAATCTGCTCCACCCAAAAA
MSAT3.205691-R	CCAAAAAGATCACGGATTTG
MSAT3.205844-L	TGAATAATGAAAAGGGTTTGTAG
MSAT3.205844-R	TGCTTCCCTTGTGTTGAGCTT
MSAT3.205888-L	ATCTGACCCACTGCTCCAAG
MSAT3.205888-R	TCTTCCAAAAACAAGGTTCC
MSAT3.206089-L	TGTTGCTGACGCACTAGCTT
MSAT3.206089-R	GAGACGGTTTTGGATTTTGA
MSAT3.206381-L	ACAGGGTACCAACACGCGAGT
MSAT3.206381-R	CACGTAGCTTGTCCCTGAA
MSAT3.206422.1-L	CGGACTCATGACGACAGTCT
MSAT3.206422.1-R	CATGGCAAACGACTAGCAGA
MSAT3.206422.2-L	CGGACTCATGACGACAGTCT
MSAT3.206422.2-R	TCATGGCAAACGACTAGCAG
MSAT3.206422.4-L	TGTAATCTTAGACACGCGATGTT
MSAT3.206422.4-R	ACGCGTGTGTCCATAACTT
MSAT3.206422-3-L	ACTTTGCTAATACTGAATCAA

MSAT3.206422-3-R	ACGCGTGTGTCCATAACTT
MSAT3.206422-L	CGGACTCATGACGACAGTCT
MSAT3.206422-R	ACGCGTGTGTCCATAACTT
MSAT3.206563-L	CAAATCCATCTTCCCACCAC
MSAT3.206563-R	CCCTCCCAACTGATCCTTTTT
MSAT3.206853-L	AGCCTTGGTGTTTTCCAATG
MSAT3.206853-R	TCTAAAAAGTTGGCACAATAGGG
MSAT3.206900-L	AAGTCAAAGGAGACGCGATGA
MSAT3.206900-R	GGATATCTCGACGAAAACAATG
MSAT3.21392-L	GCATCTCAGTGCCAACTCAA
MSAT3.21392-R	GCCTCTCCCAAGAAAGAAC
MSAT3.21418-L	GAAACCGGTATGAATGCCATC
MSAT3.21418-R	GAATGTGGTGGCTGTGTCAG
MSAT3.21492-L	TCATTTTAACCGCTCCCTCT
MSAT3.21492-R	GAAGAAAAGCCGCAATCAG
MSAT3.21694-L	TGGTGGTGGTAATCTTATGCTTT
MSAT3.21694-R	TTTCCACAAAATGATACAAAATG
MSAT3.21700-L	CCAACATGGTTGAAAGGTTG
MSAT3.21700-R	TGTTGGAATTAGGGATCTAGGAA
MSAT3.21761-L	ATTAGCTCGCATGTGCTCT
MSAT3.21761-R	CCAGTACCACCACCACCTTCA
MSAT3.21774-L	GGGATGACTGAAACCATCAA
MSAT3.21774-R	GCAITGGCGTATGAAACACA
MSAT3.21869-L	GCCGGTCAAATTTTACAGTA
MSAT3.21869-R	GCCGGTCAAATTTTACAGTA
MSAT3.21869-L	AAATATGAATTATCTTGTGAACAAA
MSAT3.21869-R	AAATATGAATTATCTTGTGAACAAA
MSAT3.21953B-F	TGATCTTCTACACCTAAAACGGAA
MSAT3.21953B-R	TTGGTACATTTCTCCACTATGCTTT
MSAT3.21953-L	CCCTTTGGTACATTTCTCCAC
MSAT3.21953-R	CCCTTTGAGCAAAGACACTACTGAA
MSAT3.21969-L	GCCCTTTGGGGAAGAGGTAAC
MSAT3.21969-R	TCTCGCTCTTTTGAACCTATGA
MSAT3.21975-L	CGGCTTTCTCCTTTCTTTTCC
MSAT3.21975-R	TGAGCTTCTTGGAAAGATTTTCAAT
MSAT3.22055-L	GAAAACAAATCCAATGAGTCAA
MSAT3.22055-R	TCTTCTTATTTGTTGTCAAAGCA

MSAT3.22444-L	TTTTTGACTCCAAATCCTGAAAA
MSAT3.22444-R	CCCATATTACGAATTTACGATGG
MSAT3.22540-L	TTTGGTCTTTAGTCGAGGCAAT
MSAT3.22540-R	TTGAGATCACAAAAGTTTTAGATC
MSAT3.23007MP-L	TCAATTTCTGCTCCGAGGAACAC
MSAT3.23007MP-R	CTTTCTTCTGATCTAATCTGGG
MSAT3.23031-L	AAGGAAGCGAATCCGAAAAT
MSAT3.23031-R	AAGGAAGCGAATCCGAAAAT
MSAT3.23031-R	GTGGGTCCCTTGGATTTCTT
MSAT3.23031-R	GTGGGTCCCTTGGATTTCTT
MSAT3.23048-L	TTGTCCACCAAACTCACACA
MSAT3.23048-R	TGAATGAGTGAATGCGATGA
MSAT3.23053-L	ATTCGGCGGATATCCAA
MSAT3.23053-R	GAATTTGGGCCATCTCAAAG
MSAT3.23065-L	CAGCACTTGCAGGACCAGTA
MSAT3.23065-R	TGGATTGAATCATTTGTCTTTGTG
MSAT3.23068-L	AAAAGGTAGCCTAAACCTGCAA
MSAT3.23068-R	TGGTTAGAAGAGACAAAAGGTTT
MSAT3.23072-L	TTCTCCATTTCTTTTGAACG
MSAT3.23072-R	TCAAAGTCAATAATTTCAATCTCC
MSAT3.23094-L	TCCCTGAAGACCCAAAACAG
MSAT3.23094-R	TGCACGTAAGTGTCTCCAC
MSAT3.23096B-L	ATGCGAAGCTGTTTGTGTTGA
MSAT3.23096B-R	TCTTCATTGATTCACAAAATTAG
MSAT3.23096C-L	ATGCGAAGCTGTTTGTGTTGA
MSAT3.23096C-R	TGATTCACAAAATTAGGTACAAA
MSAT3.23096-L	ATGCGAAGCTGTTTGTGTTGA
MSAT3.23096-R	GTGCAAAATTCGAGGTGTTTT
MSAT3.23110-L	GCTGATGACCATGAAGCTGA
MSAT3.23110-R	ATCCTTGCATAGGCGATCT
MSAT3.23117-L	TTCTTATCCCATGCACTCT
MSAT3.23117-R	TTGTTAATGTAATTCGCCACAA
MSAT3.231183-L	TCCTCAATTCCTGGGAATA
MSAT3.231183-R	CAGCAAACAAATGGGAAGT
MSAT3.231187-L	TTTGTTTTTAGGGTAATGACAA
MSAT3.231187-R	TGAATGAGTAGATGAAAGAGAAG
MSAT3.231194-L	TTTTGAACCTCACCTTGTTTTT

MSAT3.231194-R	ATTAGGGTGCAAATGGTTGG
MSAT3.231197-L	AACGAAAAACGTTGCATCTGA
MSAT3.231197-R	TTGGGTTTATACACGGAACC
MSAT3.231198-L	GGTTCOGTGATAAAACCCAAAT
MSAT3.231198-R	GCCAAAGTCGAAAACCTTCCCTTT
MSAT3.23124-L	TTGGTGGGCTTTTAGCAAAT
MSAT3.23124-R	TAAATGCCAGGCCTAAGA
MSAT3.231303-L	GGTGCACGGCGTATATGTTA
MSAT3.231303-R	AGACACGTTTCGTGCCATGTA
MSAT3.23130-L	TGGGTTATCTAATGTTGTGAAAAA
MSAT3.23130-R	GTGCACCCTCTGCAAATTA
MSAT3.23131-L	CGATTGCTCGAAACACTCAA
MSAT3.23131-R	AAAGCGCAAACGCATAAAGT
MSAT3.23136-L	TCAAGTGGGTTTTGCCTAGC
MSAT3.23136-R	TGCTCTCTTCGCTTTCGTCT
MSAT3.23141-L	AAAGGCTACCATCCCTTTCGT
MSAT3.23141-R	TGGTCTAAAATCAATGCAATGTC
MSAT3.23142-L	GGGTCAAAGCTGTCATTGCT
MSAT3.23142-L	GGGTCAAAGCTGTCATTGCT
MSAT3.23142-R	GTCCGCGTTATGGACAGAAT
MSAT3.23142-R	GTCCGCGTTATGGACAGAAT
MSAT3.23145-L	AAAGGACGGTAAGGTACTCGAC
MSAT3.23145-R	TGAAGACGTTTTATTCTCACTGGA
MSAT3.23146-L	GTGGTTTGCAATTTTAGTGGA
MSAT3.23146-R	CCAAGATGGTGGATAAGAGTGG
MSAT3.231528-L	GCCAAAACCATGTTTCGTAT
MSAT3.231528-R	TGAGGACTTTTTGTAGGCATGT
MSAT3.23152-L	TCTCTTCCAGGAGCTGGTT
MSAT3.23152-R	GTTCCAGGCGATGTTTTGGTT
MSAT3.23154B-L	TGAGCAGTATGATTAGAAACAAGTCA
MSAT3.23154B-R	AGCTGGGTTCGTCTAGTGTC
MSAT3.23154-L	AACTGAATCCGAATGAGCAG
MSAT3.23154-R	AAGTCGTAGGCGCTCAT
MSAT3.23155-L	CGGTTCAATTATACACCATGAGTT
MSAT3.23155-R	GGCAATTTGGTCTGAATTATCA
MSAT3.23158-L	TCAGAGAACCCCATAGAGAGAA
MSAT3.23158-R	TGGCAAGGCTCCTATTAAGC

MSAT3.231653-L	AGCGAGGAGAGAAAGAAGCA
MSAT3.231653-R	CATGCATACAAAACATGTCAGG
MSAT3.23165-L	TCGGATCAGTAAACGATTTCAA
MSAT3.23165-R	CGCCCTCGTTTTATTTCCAAT
MSAT3.23168-L	GCAGCTGCATAGACAGTGGGA
MSAT3.23168-L	GCAGCTGCATAGACAGTGGGA
MSAT3.23168-R	TCTTCGGCGTTATCTCCATC
MSAT3.23168-R	TCTTCGGCGTTATCTCCATC
MSAT3.23173-L	GCGCCTTTCTTCTCTCTCTT
MSAT3.23173-R	CCGCCAGAGTTTTCTAGGTG
MSAT3.23174-L	TCTCGCTGATGATGATTGATG
MSAT3.23174-R	CGAAACAATCCAACCAAAA
MSAT3.2317506-L	TTTTTGTCTCTAGTGCTTGAC
MSAT3.2317506-R	CCAGAGGAGGACAAAACAATC
MSAT3.231754-L	ACGGAAGCTTCGTTTGATT
MSAT3.231754-R	CTCGAATACTCATTTGGGTCAA
MSAT3.23175-L	TCGGTCTCTTTGCTACTTCA
MSAT3.23175-R	CCAACCAGAGGAGGACAAAA
MSAT3.23176-L	TACCCACGCCCTAAATCAAC
MSAT3.23176-R	CCTCAAACGCGAATAGCAA
MSAT3.231780-L	AAGCATAACCATAGACCCGCAAT
MSAT3.231780-R	GGTGCAGATCCAATTTCGTC
MSAT3.23178-L	CCCTAACCAACCTTAATTCGTC
MSAT3.23178-R	GGTGCAGATCCAATTTCGTC
MSAT3.23179-L	CATTGCCGCGAATCACTA
MSAT3.23179-R	GAGTGCCAACCTGGATGATGA
MSAT3.23184B-L	CCAATGCAATCTATCCTTTTGG
MSAT3.23184B-R	TTTTCTCATATAAAAGTAGTCGAT/
MSAT3.23184C-L	TCCGATTCATTTGATTCCA
MSAT3.23184C-R	TTTTCTCATATAAAAGTAGTCGAT/
MSAT3.23184-L	TCCGATTCATTTGATTCCA
MSAT3.23184-R	TCTCGTTTTCTCGCAGCTTT
MSAT3.23186-L	GCTAAGAGGGGAAAATATCACAV
MSAT3.23186-R	ACAAGTGGGTTTTAAAGCTCAAG
MSAT3.23187-L	TCTGGCGGCTCTTTATTTCT
MSAT3.23187-L	AATGCATGGTCCAACCAAT
MSAT3.23187-L	AATGCATGGTCCAACCAAT

MSAT3.23187-R	CGCCAGATAGGGACAGTCTT
MSAT3.23188-L	CGGTGGTCCATTTTTGTAGA
MSAT3.23190-L	GCATTGCCAAAGGAATCTTG
MSAT3.23190-R	TGCCGGAATAGAGACCACT
MSAT3.23197-L	TCCCTCTGCGTATGAATGA
MSAT3.23197-R	AAAACGTTAAGTTTTGCTCTCCTAAA
MSAT3.23200-L	CCGGTATCATACGGATTGCT
MSAT3.23200-R	CACAAGTTACCGGTGGTGTG
MSAT3.23203-L	TGTGGGCTTAGATTTGATCTTC
MSAT3.23203-L	TGTGGGCTTAGATTTGATCTTC
MSAT3.23203-R	TAGGCACAAACGTTGCACTC
MSAT3.23203-R	TAGGCACAAACGTTGCACTC
MSAT3.23211-L	CGCCACTTCTGCTTACCAAG
MSAT3.23211-R	CCTTCCCACTATTTCCACCA
MSAT3.23213-L	GATTCTTCAATTCACACCAAAAA
MSAT3.23213-R	TGTGCACAGCACCAATGTA
MSAT3.23216-L	TTCCAGCAGCACTAACAAA
MSAT3.23216-R	CCAGTGGCTGTCCAAGAAG
MSAT3.23221-L	TTGCTTTACGATTCAGATAAACTCC
MSAT3.23221-R	TTATCCTTCAAATAAACAATAGGAAG
MSAT3.23264-L	GCAGCTGCCTTCTTCATCTC
MSAT3.23264-R	TGGATAAATGGCAATACCTCTT
MSAT3.23265-L	GGAGTGGGTCAGATATGGA
MSAT3.23265-R	AGCAGAGATTTGGTGGTGGT
MSAT3.23277-L	TGCGACCTTTCTAGTTTTCTTTTT
MSAT3.23277-R	TTTCTGAGCACTTGCTTGCT
MSAT3.23280-L	TTCTCACACATACATTTGTTCC
MSAT3.23280-R	TGACCATGTAGCCCATCAAC
MSAT3.23289-L	CCGTACGGCTCAATTTTGTT
MSAT3.23289-L	CCGTACGGCTCAATTTTGTT
MSAT3.23289-R	CGACGGAGGACTTTGTAAGC
MSAT3.23309-L	AGAGAAATGAGGCCTGAACCA
MSAT3.23309-L	AGAGAAATGAGGCCTGAACCA
MSAT3.23309-R	ATGAGCAAGTGCTTGTGTGC
MSAT3.23309-R	ATGAGCAAGTGCTTGTGTGC
MSAT3.23330-L	GCATTTGAAGAAATTTGGCAAGA
MSAT3.23330-L	GCATTTGAAGAAATTTGGCAAGA

MSAT3.23330-R	GCCAGCTGAGCTAGATGATTT
MSAT3.23330-R	GCCAGCTGAGCTAGATGATTT
MSAT3.23352-L	TCACGAACCTTAACATATTTTTGC
MSAT3.23352-L	TCACGAACCTTAACATATTTTTGC
MSAT3.23352-R	GGTGCCGATATTTCTTCTCT
MSAT3.23352-R	GGTGCCGATATTTCTTCTCT
MSAT3.23357-L	CTGTCCAAACCCCAACCTGAT
MSAT3.23357-L	CTGTCCAAACCCCAACCTGAT
MSAT3.23357-R	AAAATGTCAACAAAGCCTGGTT
MSAT3.23357-R	AAAATGTCAACAAAGCCTGGTT
MSAT3.23362-L	TCACAAACCATGGAACAACAA
MSAT3.23362-L	TCACAAACCATGGAACAACAA
MSAT3.23362-R	TCTCTGCCCATCCTCCTCTA
MSAT3.23362-R	TCTCTGCCCATCCTCCTCTA
MSAT3.23388-L	GGCTCTACGTTACTTAAGCGAGC
MSAT3.23388MP-R	CGATTCGTCCGAGAGCGGCAAA
MSAT3.28L	TACAAGTCATAATGAGGCG
MSAT3.28R	GGGTTTAGCATTTAGC
MSAT3.58.1F	GCTGAGCCAAAGCAAAGAC
MSAT3.58.1R	TGGGACAAAGTGGTGTATCCT
MSAT3.70A-L	CGTTTTTCTCTCCAAACGA
MSAT3.70A-R	TATTTTTCGGGCCTCAATCA
MSAT3.70B-L	TCGTTTTTCTCTCCAAACG
MSAT3.70B-R	TTATTTTTCGGGCCTCAATC
MSAT3.70C-L	CGTTTTTCTCTCCAAACGA
MSAT3.70C-R	ATTCGAGACACCACCACACA
MSAT3.70D-L	ATTGCAGCATCGACTCACAC
MSAT3.70D-R	ATTCGAGACACCACCACACA
MSAT3.70E-F	TATTTTTCGGGCCTCAATCA
MSAT3.70E-R	AGCATCGACTCACACACCAC
MSAT3.70-F	CACCCATAGACACACAACC
MSAT3.70L	CACCCATAGACACACAACC
MSAT3.70R	TCAAAACCTCCAACCATAC
MSAT3.70-R	TCAAAACCTCCAACCATAC

Couples d'amorces pour amplifier les MSAT du chromosome 1

Nom de l'amorce	Séquence de l'amorce (5'→3')	Nom de l'amorce	Séquence de l'amorce (5'→3')
MSAT 1.184319-F	AAGACTCTACGCGACCGAAA	MSAT 1.195448-R	GTGTTTTTCCCGTTGTTGCT
MSAT 1.184319-R	TCAAGCCCATGATCAAAGCT	MSAT 1.197467-F	GCCCTACGCAAAGCTTAGAA
MSAT 1.184344-F	CTGGTGGTTGTATACGTTCTG	MSAT 1.197467-R	CCATCCAAGTCGGACACTA
MSAT 1.184344-R	GGAAGGAAGACGATGGACCT	MSAT 1.197642-F	TCATCTCCTGATCTACCCCTTG
MSAT 1.18492B-L	TTTCCTCAGGCTTACAACC	MSAT 1.197642-R	CTCATCCCTCCGACCTATCA
MSAT 1.18492B-R	CTTGCGTTTTCCCCAACAT	MSAT 1.199612-F	CGACACGAAAATGACACGTT
MSAT 1.185580-F	TGGTTTAAGAGCGACTCCTG	MSAT 1.199612-R	ATGCCTTCCCAAGTTTGTGT
MSAT 1.185580-R	GGGCAGCTAACTTCCTGATG	MSAT 1.201377-F	GTCCAAATATCAGGCATCG
MSAT 1.185700-F	GAAAAATGGAGCGCAGAGAC	MSAT 1.201377-R	CGGCCCAAAAAGACAACCTT
MSAT 1.185700-R	TGCTAGCTCACCTCATGCAA	MSAT 1.202112-F	AATCAGCATCCGTTCCAGGAC
MSAT 1.186312-F	GGAACGTTTCATCGCTTCTC	MSAT 1.202112-R	ATGACCGAGCACGGTATTTT
MSAT 1.186312-R	TTAAGTGGAGTCCATGACG	MSAT 1.202352-F	CCAAAACCAATGGTGACACT
MSAT 1.186507-F	ATGCAGAGGAATGTGAAGC	MSAT 1.202352-R	GCGTACTTGCCATTGAGTGA
MSAT 1.186507-R	GCTCGTTTACATCCACGAAGA	MSAT 1.203115-F	GCTGTGCTATATTTCCGTTGG
MSAT 1.18748-L	TTATATACCATTCTCGTCGTGTTA	MSAT 1.203115-R	TGAGCTACATCTTCCCTCAGGT
MSAT 1.18748-R	GCAAAAATGAAATGTTAAGGC	MSAT 1.204285-F	CCTTTACAACACTACTTGCCCTCA
MSAT 1.189086-F	CGTGCCATAAACAGTGGGAGA	MSAT 1.204285-R	CATCACCTGTTCCAGTCGGATT
MSAT 1.189086-R	AGCTCTCTTTCCAGCAATGG	MSAT 1.204498-F	CACCAACACGAGCAGTTTTGT
MSAT 1.189237-F	AGAAGGCACCTTGACGTCGTT	MSAT 1.204498-R	GTGTAAGGGCAGGCAACCTA
MSAT 1.189237-R	AGGTCACACGATTCCTTTTG	MSAT 1.205030-F	GGGGGAGAAGATTCATTTT
MSAT 1.190113-F	TTGGTTTCGTCGTTTGAGC	MSAT 1.205030-R	GCAACTCACCCTATGGTCAAA
MSAT 1.190113-R	CATCCGATTCCTAGTACCC	MSAT 1.205291-F	TGATAGGGTATGACGAAGATCC
MSAT 1.190178-F	TGCCATGTGACTTTTTGACC	MSAT 1.205291-R	CGAATCTCCTAACAAGCAGTCA
MSAT 1.190178-R	TGAATCATCCGTCCTGACAA	MSAT1.18401A-F	CTGAACCCGAATCTCATCGT
MSAT 1.191395-F	CTGTTCTCTGCGTGTGTCC	MSAT1.18401A-R	TGTCATGCATTGCACATACG
MSAT 1.191395-R	GGAACTTGCCAAAACAAGGA	MSAT1.18492A-F	CAATCTTTGGTGGTGGAAA
MSAT 1.191862-F	TGGAAAGCAGCATACCAAGA	MSAT1.18492A-R	TCAGGCTTCACAACCCAAAT
MSAT 1.191862-R	TCCAGGCTTGAGTGATGAAA	MSAT1.18748A-F	AGGCCAAAATACAGCTGACG
MSAT 1.195220-F	AAAAGTGGCCCAGGATGATT	MSAT1.18748A-R	CGGAGGGGAAGATGAAGAAC
MSAT 1.195220-R	TCACTCAACTGGTTCGTTCCG	MSAT1.18862A-F	GGCATTATTCCTCTGTTTCA
MSAT1.19083A-F	GATAAGGATTTCTGCCATTCCG	MSAT1.20851-R	GTCTTTACTGCGACGCTCAC
MSAT1.19083A-R	ACCGGGTTCAACTGTGAAG	MSAT1.20890-L	TTCCGACAACACTTGCAATTT
MSAT1.19270A-F	CCGAATTTCCGAGGAAAACAA	MSAT1.20890-R	ACGATGCTGCTGATGTTGAC
MSAT1.19270A-R	TCTTTCCAGTTTCCAGCATGG	MSAT1.20761-L	GCTGACAAGGATCTTGACGTG
MSAT1.19454A-F	CGTGTTTCCTTATCCGGTGT	MSAT1.20761-R	CCGTGAAAAGGAAACGAAG
MSAT1.19454A-R	TCACATCAAGTTTGTCCCA	MSAT1.20798-L	GGCAACGACCCCTTCAATA
MSAT1.19746A-F	GCCCTACGCAAAGCTTAGAA	MSAT1.20798-R	TCAGAAAACGCTTCTCAT
MSAT1.19746A-R	TTGGGAGTTCCGTATTAGCC	MSAT1.20849-L	GCAATAGGAGTGTGATGGGTTTT
MSAT1.19748A-F	TGGCTCAGAGTGATGAGAAAT	MSAT1.20849-R	TTTGCCCTACCTTGGAAAGCAC
MSAT1.19748A-R	TCTTCCCCAACTCTTGGTCA	MSAT1.20958-L	AAGGCAATATCGTCGTTCTCA
MSAT1.19860A-F	GAGAATGCAGAAGCACTACA	MSAT1.20958-R	TTGATTAATACCAAGGGTTTTATC
MSAT1.19860A-R	CACGGCCATTAGTGTCTTT	MSAT1.20987-L	AGGCCCAATTAACGGTGTGT
MSAT1.19928A-F	CGATGAATCGGTGTGATGA	MSAT1.20987-R	GGATGTGGAGAAACAGATTTCA
MSAT1.19928A-R	GGCTCGCATTTGGATTTAGAA	MSAT1.21035-L	GGTGGGGATAAGTATGTCATT
MSAT1.19935A-F	CCTTGCACAATCGTTACTCG	MSAT1.21035-R	TGGATATCGCAGACAACAGC
MSAT1.19935A-R	AGAGGCTTGTGAAGAAATCAGA	MSAT1.21082-L	TTTTCTAAGTTTGAATGTTTCTATC
MSAT1.20085A-F	GACCCGGATCACTGCATATT	MSAT1.21082-R	TGAAATGAAATGCTTTGCTCA
MSAT1.20085A-R	GTCAAGCGGATCGTCTACAA	MSAT1.21153-L	TGGAACCTCAACCAAGACA
MSAT1.20097A-F	TCGTACGTACATCTTTGCGATA	MSAT1.21153-R	TTGGAATTAGGCAAGGAAATG
MSAT1.20097A-R	GATTTTCGGAATCGCTCCAA	MSAT1.21220-L	TTTAGCCTTTTAGGATTACACTG
MSAT1.20282A-F	TCACTTGAAGACGAATCATCG	MSAT1.21220-R	TTCCGACTGTCCAAGCAAATG
MSAT1.20282A-R	GAGGATGGGATTATGATCAGTG	MSAT1.21226-L	ATGGCTGTGCTTGGCCATAA
MSAT1.20311A-F	GCTGTGCTATATTCGGTTGG	MSAT1.21226-R	TGCGTGACTGAGGTCTGAAG
MSAT1.20311A-R	TGAGCTACATCTTCCCTCAGGT	MSAT1.21228-L	CGAAAACAAAGAAAACAAAACAA
MSAT1.20331A-F	TGCTTTCCGAGTCAATCA	MSAT1.21228-R	GGTTTCGGCTAGAAATCTCC
MSAT1.20331A-R	TGTTCCAAAGTCCAAGCCAGAT	MSAT1.21231-L	TCCAAATGAAAACATCTCC
MSAT1.20483A-F	TCCCAAATCGAGAACAAGC	MSAT1.21231-R	TTGTGAACCTTGTGATGGA
MSAT1.20483A-R	TGGTTGATGACAATGGAAGC	MSAT1.21235-L	CTTTTACAAGAGTTGGGAGTTAA
MSAT1.20484A-F	CACGTTGGTGTGGAAAGTCA	MSAT1.21235-R	TCTCCACACAATCCCATCAA
MSAT1.20484A-R	TGGTGTGTGGTTGCCATAGTAA	MSAT1.21266-L	AAATTTTGCATGACAGGTTTTG
MSAT1.20495A-F	TGGCCCTCAGTTTTTCAATTT	MSAT1.21266-R	GCTTGTAAGGTCCCTCGTTGG
MSAT1.20495A-R	GGGTTGCAAGTCGTCAGTT	MSAT1.21340-L	TCTTGATTATAAACCTAAACCTTT
MSAT1.20502A-F	GGGGGAGAAGATTCATTTT	mSAT1.21340-R	TCCAAATTTGGGACAAACG
MSAT1.20502A-R	GCAACTCACCCTATGGTCAAA	MSAT1.21362-L	TGCTCATGTTCCAGGCTATT
MSAT1.20528A-F	TGATAGGGTATGACGAAGATCC	MSAT1.21362-R	AGGAGGGTTTTAAGGGGGTTT
MSAT1.20528A-R	CGAATCTCCTAACAAGCAGTCA	MSAT1.213755-L	TTGCAAGCTCAAGAAAATAACA
MSAT1.20851-L	GCCCATCTCACCATCCATTA	MSAT1.213755-R	CTCCGCATATTCGATTGACA

MSAT1.21375-L	ATCAATGTAAGAATTACATGCATA
MSAT1.21375-R	CCACATAAATTTGAAGGCTTTTTGT
MSAT1.21408-L	GGCATTGGCATCAAACAGAA
MSAT1.21408-R	CATCTAAAACCTTTTAAACAGTCTG
MSAT1.21421-L	AAGTCATGCTTTACATTTGAAACGA
MSAT1.21421-R	TTCCATCAATCGTCTGAAAACCTT
MSAT1.21422-L	CCGTCAGTAATTTACCGTTCA
MSAT1.21422-R	TGTGGAGACCAAAATCTGGA
MSAT1.21456-L	AATGCCCATGGCTTTTATGT
MSAT1.21456-R	GCAATTTGGGTCCAATTTGATA
MSAT1.21457-L	GGGGTCTCTCGCTTTTCTCT
MSAT1.21457-R	TGTTGGAAAAGTATGGGATGC
MSAT1.21539A-L	AACGCCATTTTTGAGGAGAA
MSAT1.21539A-R	CAAATGAGGAATCACTTTGCAC
MSAT1.21539-L	TGATCAAGGCAACAAAACGA
MSAT1.21539-R	CAAAGACCTGAGCAAAAACCTGA
MSAT1.21600-L	CATCAACGCAAAACATCCAC
MSAT1.21600-R	AGGCCTGTAGCTGAGAGCAA
MSAT1.21603-L	CATTTCAACAATTAATCTATTTAC
MSAT1.21603-R	GGGTTTATTTCCATCACCATGT
MSAT1.21636-L	TCGTAGATCTCAATCCATGTTTTT
MSAT1.21636-R	CCAAATTTGGAAGTGAGACGA
MSAT1.21656-L	TTCGAATTTCAATGGCCTTT
MSAT1.21656-R	TTTCGAAATAGCTTCACCAGTAA
MSAT1.21673-L	TGATTTTTGTTCAATCGTTGAAA
MSAT1.21673-L	TGATTTTTGTTCAATCGTTGAAA
MSAT1.21673-R	CGAATCGGGAAACCAATAAT
MSAT1.21673-R	CGAATCGGGAAACCAATAAT
MSAT1.21694-L	GCGTCGAGGAACCTCTTCTGT
MSAT1.21694-R	TTGGTGACCACAGAACTCG
MSAT1.21853-L	TTGTCCGTCTCTATGTAACAG
MSAT1.21853-R	GGGGTACATCATATTTGAGGATAAA
MSAT1.21946-L	TTTAACTTCCGAAATCTGAAAAGT
MSAT1.21946-R	TTTGCTCTTTGTGTGTTTGGA
MSAT1.21988-L	TGGTGAACCTGCAATACCAA
MSAT1.21988-R	CATGTCTTTGGCAAAATGCAA
MSAT1.22002-L	AGTCCACGGCCAAAGAAGTT

MSAT1.22002-L	AGTCCACGGCCAAAGAAGTT
MSAT1.22002-R	TTCGCCCTCACGTTCTTTAT
MSAT1.22002-R	TTCGCCCTCACGTTCTTTAT
MSAT1.22011-L	TTTTTCTGCCATTTCTTGC
MSAT1.22011-R	ATGGCTTTGTTCTGTTCTTTGG
MSAT1.22063-L	TTCGAAAAGGGGCAATCTA
MSAT1.22063-R	AAAACCTTTCTTCAAAGCACTTAA
MSAT1.22068-L	ACCACCATGCTCACTCCTTC
MSAT1.22068-R	TGCACTCACTCTCGTTGACC
MSAT1.22109-L	TCGTCCTTGGTTCAAAAATATCA
MSAT1.22109-R	GGACTGTCTAAACACCACCA
MSAT1.22110-L	GGCAGATACAATAAAAACCTAGGTC
MSAT1.22110-R	TGAAAAGAAATAGTCTTCGCACA
MSAT1.22118-L	TTTTGGGGACTTACTGTTAGTTT
MSAT1.22118-R	CAATGAAAAGTTGGCAGACG
MSAT1.22162-L	CGAGTGACGTTTGAGACGAA
MSAT1.22162-R	TTCTATTTTACCTGCTGGATTCAA
MSAT1.22191-L	AACGTTCCCAATGAAAACACT
MSAT1.22191-R	CTCCATTTGGGGAAAATGTTG
MSAT1.22225-L	CCAAGACTCAAATTTGGTCAGC
MSAT1.22225-R	TTCACAGCCTGAGAGCCAAAT
MSAT1.22270-L	ACCCCTATTGAGGGTGCTCT
MSAT1.22270-R	TGTTAATCATTGTCTGTTTTATTCA
MSAT1.22301-L	ACCCGATCTATGGTGGTTCA
MSAT1.22301-R	GCAACAAAATTTGGACATGG
MSAT1.22372-L	TCTGCAGCTCAGATTTCTCCA
MSAT1.22372-R	ACGATACCGTCTGGGTGAAA
MSAT1.22606-L	CACAAAATTAAGGAAAATAGCAG
MSAT1.22606-R	CAAAATCTCACGAAACTGTGC
MSAT1.22932-L	GACCCGATTAACCGACAAA
MSAT1.22932-R	CGCATGCAACCATTTAAAGTC
MSAT1.22939-L	GGAGGCGTGAGTGAGAGAAG
MSAT1.22939-R	GCCCATGATGTTCTAATTTATGC
MSAT1.23172-L	AATGTAAGGCCACGACTGC
MSAT1.23172-R	TCTGCTTCGTGAATTTGAG
MSAT1.23197-L	TGCAGAGAACTTTGGAGTAATG
MSAT1.23197-R	GAGAACTTAGCCATCTCACCA

MSAT1.23512-L	CAGATTCATGTACGGTGACCA
MSAT1.23512-R	TGCATGTCAAATTTAACGTAATAGA
MSAT1.23562-L	TGACTGTGCTTCTTGAAATGC
MSAT1.23562-R	CCCCAACACATTTTCTGATTC
MSAT1.23830-L	TTTGGTAGGCACAAATGAAGAA
MSAT1.23830-R	GAGGCTTTGGCATAACAATATCA
MSAT1.23877-L	TTTTGGACATAGCCACATCG
MSAT1.23877-R	TCGTAATGAATTTACGGTCTC
MSAT1.24125-L	TAAGGCCCTTGCTGTTCTTA
MSAT1.24125-R	CACCAACAAAATGCGCTTA
MSAT1.24136-L	CTGACTACCCAACCTGTTTGTCC
MSAT1.24136-R	AAAAGAATTTGTTTCCACCCTACC
MSAT1.24404-L	GTGGTTGCGATATGAGCCTA
MSAT1.24404-R	TCTCTCCATCCCAAATGACA
MSAT1.24698-L	CAAAAATGCCCCAGTTGTTT
MSAT1.24698-R	GAAGGTTCTCACCCTGGACA
MSAT1.24709-L	GGAACATATATATTACGCGTGGA
MSAT1.24709-R	CGAAGGTGCAACAATAGATGA
MSAT1.24925-L	TGAGAGGAGGACGATTTTCA
MSAT1.24925-R	CGTCGCATCCCTCTATAGCTT
NGA128.1F	TCACCAATTTGTTGGGTTCTG
NGA128.1R	CTTCGGTCTGTTGATGTCGT

**Annexe 3 : Amorces dessinées par Marie Verzaux et utilisés pour amplifier les produits de cRT-PCR**

<b>Gene</b>	<b>Nom de l'oligonucléotide</b>	<b>Sequence de l'oligonucléotide</b>
<b>Nad1</b>	mnad1M2	TTTCCACGATATCGTTATGA
	mand1M1	GTTGCACAAAAGCCATTAC
<b>Nad2</b>	mnad2M1	TACAACCTCCATGAATGAGC
	mnad2M2	CCTCTCCTTTGTTCTCAGTTA
<b>Nad4</b>	mNad4M2	CGTTCGGATGGGTGTTTAC
	mNad4M1	AATCGTATCGGTCGTATTCTT
<b>Nad5</b>	mNad5M1	ATCAAACATTTCCGATGAGA
	mnad5M2	GCCGAGCGAATAAGTC
<b>Nad6</b>	mNad6M1	CTGAAGTGTCGCGAAAGA
	mNad6M2	GATTGGGGCTATAGTACTGA
<b>Nad7</b>	mnad7M2	CGATTCTATGTCCAAACATCA
	mnad7M1	AGGATGTTGAGGTCCAAA
<b>Nad9</b>	mNad9M1	CTATTCCCATGTTCCGATCT
	mNad9M2	CCCAAGAATTCGCTATTC
<b>cox 1</b>	mcox1M2	CAGCAATGGCACCGAAAAT
	mcox1M1	ATAACAAAAGATGTGCTCCGA
<b>cox2</b>	mcox2M1	CTGCGTCTTGAGATCCTAAT
	mcox2M2	CGGGTATCCAATCAAT
<b>cox3</b>	mcox3M2	TGTGGTATTCGGCAATATCTT
	mcox3M1	TGAGGCCCAAACCTAGAAAG
<b>Atp1</b>	matp1M1	CCGATCTCATCCACTTGAA
	matp1M2	GGTGGATTAACCTAACGAAAGA
<b>Atp6-1</b>	matp6-1M2	ACCGGTCTGGAATTAGGTGTA
	matp6-1M1	CGGACGCTTGACCACTAGAAA
<b>Atp6-2</b>	matp6-2M1	TCCGCGGGAGTATGAT
	matp6-2M2	ATTAACCGGTCTGGAATTAGG
<b>Atp8</b>	matp8M1	GAGAAATAAGTGGCTCACGA
	matp8M2	AGAAGAAAAGGCATAACCAG
<b>Atp9</b>	matp9M2	CGCGAAATCCATCATTGGCTA
	matp9M1	CAGCTCCGGCACCTATTGA
<b>ccb6-n2</b>	mccb6n-2M1	AAACCACCAACATGTCAAGAT
	mccb6n-2M2	CGCGAAGTACTCTTGT
<b>ccb2</b>	mccb2M2	TGTTCTTTCATCGATTGG
	mccb2M1	GTGGAGGAGAAGATCAGTTTA
<b>ccb3</b>	mccb3M1	AGCTTCTGGTCTTTGACATCA
	mccb3M2	CGTATCTTGTTTGTCTGGAA
<b>ccb6n1</b>	mccb6n-1M1	TGGTGGTTGTTTCTTGTTGT
	mccb6n-1M2	CGCAAATGCTATCCTG
<b>ccb452</b>	mccb452M1	GGGAACATCTCTACTGACGAA
	mccb452M2	GGCTGAAGTGGCTATACAT
<b>rps4</b>	mrps4M1	CGGCGTTGTATTATTGTCAGT
	mrps4M2	CAAACCTTCCTCTTCGGAGC
<b>rps7</b>	mrps7M2	AAAGAGGGGAAGTGCACGTA
	mrps7M1	TTGATTACATCGCGTTCA
<b>nad3 rps12</b>	mnad3rps12M1	GGAAAAGGAACACCGAG
	mnad3rps12M2	GAAGAAGAGGCAGATCAA
<b>nad4L atp4</b>	matp4nad4LM1	GGCATTGACATAATAGGAA
	matp4nad4LM2	GAGAGAGCTTGGTGGTCTTA
<b>cob rpl5</b>	mrpl5cobM2	TCTTCTTGTTCTTTGCCATAA
	mrpl5cobM1	CGGTTTGAGCAACGGATCT
<b>rps3 rpl16</b>	mrps3rpl16M1	AAATAAGATCTCAGATTGACA
	mrps3rpl16M2	CGAAGAAATGGTAAGATATG
<b>rpl2 orfX</b>	mrpl2orfXM2	AAGAGGGCTGGACGAGTG
	mrpl2orfXM1	TACTCTTACAATGCCCATAG

## Résumé

Un croisement intra-spécifique entre deux accessions d'*Arabidopsis thaliana* (Sha et Mr-0) a donné 100% d'individus mâles stériles dans la descendance F1 lorsque Sha est utilisé comme parent maternel et Mr-0 comme parent mâle (Sha x Mr-0). Dans le croisement réciproque (Mr-0 x Sha), les individus F1 sont mâles fertiles. Étant donné que l'hérédité du cytoplasme chez *Arabidopsis thaliana* est maternelle (non Mendélienne), cela indique que cette stérilité mâle résulte d'une interaction entre le cytoplasme de Sha et les allèles nucléaires de Mr-0. Une cartographie génétique d'une population en ségrégation a permis, au cours d'un travail préliminaire à cette thèse (Cezar Sophie Hiard et Cezary Waszczak), de localiser les locus nucléaires sur les chromosomes 1 et 3 ; de plus, la présence des allèles Mr-0 au deux locus est nécessaire pour observer le phénotype de stérilité mâle. Les objectifs de thèse étaient de décrire le phénotype de stérilité mâle, d'identifier les facteurs cytoplasmique(s) et nucléaires, et de commencer l'étude des mécanismes moléculaires sous-jacents à l'origine de cette stérilité mâle. Le clonage positionnel, stratégie adoptée pour identifier les facteurs nucléaires, a permis de diminuer les intervalles aux deux locus et, pour l'un d'entre eux, de pouvoir proposer une approche « gènes candidats ». En parallèle, des croisements réciproques supplémentaires, entre Mr-0 et d'autres accessions que Sha, ont été réalisés et ont permis d'identifier d'autres accessions inductrice de stérilité qui proviennent de la même région géographique que Sha (Asie Centrale) ; la détermination des cytotypes chez les nouvelles accessions testées a montré que la stérilité mâle en F1 apparaît lorsque le parent maternel possède un cytoplasme du même groupe que celui de Sha. Concernant le facteur cytoplasmique inducteur de stérilité, nous avons identifié un seul candidat, l'*orf117Sha* mitochondrial, qui est spécifique des accessions inductrices de stérilité comme Sha et qui présente beaucoup de similarités avec des facteurs inducteurs de stérilité identifiés dans d'autres systèmes de stérilité mâle cytoplasmique chez les Angiospermes. La caractérisation du phénotype de stérilité mâle par une approche cytologique classique a été réalisée sur des génotypes où les allèles Mr-0 aux locus mainteneurs sont introgressés dans un fond génétique Sha ; selon les génotypes, les allèles mainteneurs Mr-0 sont présents simultanément ou uniquement à l'un des deux locus identifiés, afin de pouvoir observer indépendamment l'influence des allèles Mr-0. Des observations d'anthères prélevées sur ces génotypes et colorées à l'Alexander ont permis de confirmer l'effet des allèles Mr-0 aux locus mainteneurs (introgressés dans un fond génétique homogène) sur la viabilité des grains de pollen, ainsi que le déclenchement tardif de la stérilité mâle nucléo-cytoplasmique au cours du développement pollinique. L'effet, sur la viabilité des grains de pollen, d'autres locus que les locus mainteneurs, appelés locus distorceurs de ségrégation et non liés à la SMNC, a également été observé.

De nombreuses incompatibilités génétiques s'expriment dans la F1 Sha x Mr-0 ; elles créent une barrière reproductive post-zygotique en affectant la production de pollen dans la descendance. Le croisement Sha x Mr-0 pourrait révéler un processus évolutif en cours pouvant aboutir à terme à des événements de spéciation au sein d'*Arabidopsis thaliana*.

## Abstract

The genome of plants is compartmentalized between nucleus and organelles (chloroplasts and mitochondria). These compartments have different inheritance patterns (nuclear genes have Mendelian inheritance; organelles genes are maternally inherited) and this has consequences for their co-adaptation, which is under the influence of two forces: on the one hand, a functional co-adaptation whose failure leads to Bateson-Dobzhanski-Muller incompatibilities, on a second hand a genomic conflict, sometimes revealed by nucleo-cytoplasmic male sterility. The wealth of genetic resources at our disposal for *Arabidopsis thaliana* makes it an interesting model to study the co-evolution between genetic compartments. As part of this problem, this thesis is devoted to a male sterility presenting a nucleo-cytoplasmic determinism and which emerged as a result of a cross between two accessions (Sha x Mr-0) while the reciprocal cross (Mr-0 x Sha) produced male fertile F1 individuals.

Results preliminary to this thesis have shown that male sterility results from an interaction between the cytoplasm of Sha and dominant alleles carried by chromosomes 1 and 3. The objectives of this thesis were to describe the phenotype of the male sterility, as well as to identify both cytoplasmic and nuclear factors, and also to start the study of the underlying molecular mechanisms, which are responsible of this male sterility. I followed a strategy of genetic mapping to identify involved nuclear factors by positional cloning. I localized the maintainer locus on chromosome 3 in an interval of 53 kb and began a molecular analysis of this region in the two parental accessions. This now allows a candidate gene approach for this maintainer. I also showed that the maintainer locus on chromosome 1 probably consists of two closely linked loci, whose alleles Mr-0 are both required for the expression of the male sterility. In parallel, reciprocal crosses involving either Mr-0 or Sha with accessions of the collection allowed us to identify at least one other accession with maintainer alleles (Etna-2), and a dozen accessions carrying the male sterility-inducing cytoplasm (all from Central Asia). A comparative analysis of cytoplasmic genomes of Sha and Kz-9, which has a cytoplasm very close to Sha but not inducing sterility, revealed a mitochondrial gene, candidate for male sterility and called *orf117Sha*. It is specific to accessions inducing sterility and its putative product shows 57% identity and 71% similarity with *orf108*, a gene involved in cytoplasmic male sterility in alloplasmic Brassicas (Ashutosh et al, 2008 ; Kumar et al, 2012).

During the mapping and cytological observations, I have highlighted genetic incompatibilities that cause distortions of segregation through the male descent in the progeny of heterozygous plants. These incompatibilities, probably independent of the cytoplasm, create genetic barriers between the two parental accessions of the cross that I studied. I propose that the cross Sha x Mr-0 reveals an ongoing evolutionary process that can eventually lead to speciation events in *Arabidopsis thaliana*.

