



HAL
open science

Recherche des gènes impliqués dans le développement sexué du champignon *Podospora anserina*

Jinane Ait Benkhali Belmanaa

► To cite this version:

Jinane Ait Benkhali Belmanaa. Recherche des gènes impliqués dans le développement sexué du champignon *Podospora anserina*. Sciences agricoles. Université Paris Sud - Paris XI, 2012. Français. NNT : 2012PA112092 . tel-00806575

HAL Id: tel-00806575

<https://theses.hal.science/tel-00806575>

Submitted on 2 Apr 2013

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.

UNIVERSITÉ PARIS-SUD

ÉCOLE DOCTORALE : *GENES, GENOMES, CELLULES*
Laboratoire de différenciation sexuée et méiose des champignons

DISCIPLINE BIOLOGIE

THESE DE DOCTORAT

Soutenu le 08/06/2012

Par

Jinane AIT BENKHALI

**RECHERCHE DES GÈNES IMPLIQUÉS DANS LE
DÉVELOPPEMENT SEXUÉ DU CHAMPIGNON
*PODOSPORA ANSERINA***

Directeur de thèse : Robert DEBUCHY DR2 et Chef d'équipe

Composition du jury :

<i>Président du jury :</i>	Pierre CAPY	Professeur
<i>Rapporteurs :</i>	Cécile FAIRHEAD	Professeur
	Sven SAUPE	Directeur de recherche (CNRS)
	Marc-Henri LEBRUN	Directeur de recherche (CNRS)
<i>Examineurs :</i>	Philippe SILAR	Professeur

*Je dédie ce travail à mes parents
Fatima et Ahmed*

Remerciements

Durant mes années de thèse, j'ai appris beaucoup tant sur le plan professionnel que personnel, et nombreux sont celles et ceux qui m'ont aidée à réaliser ces travaux et que je dois remercier.

Tout d'abord, je tiens à remercier sincèrement Mr Robert DEBUCHY mon directeur de thèse de m'avoir accueillie dans son équipe et son laboratoire, ainsi que pour sa disponibilité et sa veille au déroulement des travaux avec beaucoup d'intérêt et d'écoute.

Je voudrais également, remercier infiniment Frédérique et Evelyne pour leur aide et leur soutien qui m'étaient très précieux durant toute la période que nous avons passée ensemble.

Mes remerciements s'adressent plus particulièrement à Sandrine, Sophie et Karine pour m'avoir accueillie avec gentillesse et bonne humeur pendant plusieurs semaines au sein de leur équipe au CNRS.

Un grand merci à Pierre CAPY et Mohamed CHAKIR pour leurs nombreux conseils et encouragements. Ma reconnaissance la plus profonde à leur disponibilité tout au long de mon parcours depuis ma licence pour toutes les questions de choix et d'orientation.

J'adresse mes remerciements à Cécile FAIRHEAD, Sven SAUPE et Marc-Henri LEBRUN pour avoir accepté d'être rapporteurs de ma thèse et à Annie SAINSAARD, Pierre CAPY et Philippe SILAR d'être membres de jury.

Pour finir, merci à toutes les personnes qui me sont chères et qui m'ont soutenue au cours de ces années de thèse contribuant à la réalisation de ce travail dans les meilleures conditions. Je remercie plus particulièrement mon père Ahmed, mon mari Ali, Mes frères Samir et Fayçal et ma sœur Ibtissam pour leur soutien inestimable.

Résumé

Le champignon filamenteux *Podospora anserina* possède deux types sexuels, *mat+* et *mat-*, caractérisés chacun par une séquence spécifique. La séquence *mat+* contient un seul gène *FPR1*; la séquence *mat-* contient trois gènes : *FMR1*, *SMR1* et *SMR2*. La fonction moléculaire de *SMR1* est inconnue, les autres gènes codent des facteurs de transcription qui contrôlent la fécondation (reconnaissance intercellulaire), et le passage d'un syncytium à un hyphes spécialisé binucléé contenant un noyau *mat+* et un noyau *mat-* (reconnaissance internucléaire). Il n'y a pas eu d'analyse exhaustive des gènes impliqués dans la reconnaissance intercellulaire et le mécanisme de la reconnaissance internucléaire est encore inconnu. Afin de déterminer les cibles de *FPR1* et *FMR1*, et les différents mécanismes impliqués, nous avons utilisé une approche microarray. Le profil transcriptomique des souches *mat+* et *mat-* compétentes pour la fécondation a permis d'identifier 157 gènes cibles, et l'analyse transcriptomique des souches mutantes *fpr1⁻* et *fmr1⁻* a révélé que ces cibles peuvent être soit réprimées, soit activées par *FMR1* ou *FPR1*, ou être sous le contrôle de ces deux facteurs. Ces expériences ont aussi détecté l'existence de 10 gènes activés ou réprimés au même niveau dans *mat+* et *mat-*. La délétion de 32 gènes choisis parmi ces 167 gènes cibles n'a permis de mettre en évidence que deux gènes impliqués dans la fécondation. Les comparaisons des gènes cibles des facteurs de transcription MAT de *Gibberella moniliformis* et *Sordaria macrospora* avec ceux de *P. anserina* révèlent un nombre significatif de gènes cibles communs entre ces espèces, mais ces gènes ont des profils transcriptomiques différents, soulevant la question du rôle de ces gènes cibles. La recherche des gènes cibles de *FPR1*, *FMR1* et *SMR2* impliqués dans la reconnaissance internucléaire a été effectuée en comparant le transcriptome des périthèces issus de deux croisements, l'un n'exprimant que les gènes spécifiques *mat+*, l'autre que les gènes spécifiques *mat-*. Les résultats ont été interprétés selon le modèle d'identité nucléaire et le modèle de ségrégation aléatoire. Le premier modèle a conduit à l'identification de 27 gènes cibles, tandis que 154 gènes cibles ont été identifiés en appliquant le deuxième modèle. Au total 46 souches mutantes ont été construites. Cependant aucune délétion n'a affecté le développement sexué. En parallèle de ces expériences transcriptomiques, nous avons invalidé tous les gènes à HMG-box de *P. anserina*. Les résultats montrent que ces derniers ont un rôle très important dans le développement sexué, particulièrement *Pa_I_13940* qui code un régulateur des gènes des types sexuels, le premier identifié chez les Pezizomycotina.

Mots clés : Champignon filamenteux, *Podospora anserina*, types sexuels, fécondation, reconnaissance internucléaire, transcriptomique, gène à HMG-box.

Abstract

The filamentous fungus *Podospora anserina* has two mating-type idiomorphs, *mat+* and *mat-*. The *mat+* sequence contains one gene *FPR1*, while *mat-* contains three genes: *FMR1*, *SMR1* and *SMR2*. The molecular function of *SMR1* is unknown, *FPR1*, *FMR1* and *SMR2* encode transcriptional regulators which control the fertilization (intercellular recognition) and the transition from a syncytium to a specialized dikaryotic hypha which contains one *mat+* and one *mat-* nucleus (internuclear recognition). No exhaustive analysis is available for the genes involved in the intercellular recognition, while the mechanism of the internuclear recognition is unknown. In order to understand the mechanism of these events and to identify the target genes of mating-type transcription factors, we used a microarray approach. The transcriptomic profiles of the *mat+* and *mat-* strains that are competent for fertilization revealed 157 differentially transcribed genes, and transcriptomic analysis of *fmr1*⁻ and *fpr1*⁻ mutant strains was used to determine the regulatory actions exerted by FMR1 and FPR1 on these differentially transcribed genes. All possible combinations of transcription repression and/or activation by FMR1 and/or FPR1 were observed. Furthermore, 10 additional mating-type target genes were identified that were up- or down-regulated to the same level in *mat+* and *mat-* strains. Of the 167 genes identified, 32 genes were selected for deletion, which resulted in the identification of two genes essential for the sexual cycle. A comparison with similar data set from the two ascomycetes, *Gibberella moniliformis* and *Sordaria macrospora*, reveals significant numbers of orthologous pairs, although transcriptional profiles were not conserved between species, questioning the function of these target genes. Internuclear recognition was investigated by the transcriptomic analysis of perithecia from two crosses expressing *mat+* and *mat-* genes, respectively. The two internuclear recognition models: nuclear identity and random segregation, were used to interpret our results. According to the former model, 27 target genes have been identified, while 154 target genes were identified with the latter model. A total of 46 mutant strains were constructed. However, these strains showed no defects in sexual development. Besides this microarray experiences, we have invalidated all HMG-box genes of *P. anserina*. The results show that the HMG-box genes have a very important role in sexual development, especially *Pa_1_13940*, which encodes the first identified regulator of Pezizomycotinan mating-type genes.

Key words: Filamentous fungus, *Podospora anserina*, mating type, fertilization, internuclear recognition, microarray, HMG-box gene.

Abreviations

Acides nucléiques et nucléotides

ADN	Acide désoxy-ribonucléique
ARN	Acide ribonucléique
ADNc	ADN complémentaire
ARNm	ARN messenger
ADNg	ADN génomique
dNTP	Désoxyribonucléotide triphosphate
CTP	Cytidine triphosphate
A	Adénosine
T	Thymidine
C	Cytidine
G	Guanosine

Unités

k = kilo- (10^3) ; m = milli- (10^{-3}) ; μ = micro- (10^{-6}) ; n = nano- (10^{-9})

%	pourcent
°C	degré Celsius
h	heure
min	minute
sec	seconde
l	litre
g	gramme
M	molaire
pb	paire de base
rpm	rotation par minute

Réactifs chimiques

EDTA	Ethylène Diamine TétraAcétate
SDS	Sodium Dodecyl Sulfate
CTAB	Bromure d'hexadécyltriméthylammonium
NaOH	Hydroxyde de sodium
NaCl	Chlorure de sodium
PEG	Polyéthylène glycol
Tris	Tris (hydroxyméthyl)-aminométhane
DAPI	4',6'-diamidino-2-phénylindole
DEPC	Diéthylpyrocarbonate

Hcl Chlorure d'hydrogène

Acides aminés

A	Ala	alanine
C	Cys	cystéine
D	Asp	acide aspartique
E	Glu	acide glutamique
F	Phe	phénylalanine
G	Gly	glycine
H	His	histidine
I	Ile	isoleucine
K	Lys	lysine
L	Leu	leucine
M	Met	méthionine
N	Asn	asparagine
P	Pro	proline
Q	Gln	glutamine
R	Arg	arginine
S	Ser	sérine
T	Thr	thréonine
V	Val	valine
W	Trp	tryptophane
Y	Tyr	tyrosine

Divers

DO _x	Densité optique à x nm
Cq	Quantification cycle
pH	Potentiel Hydrogène
PCR	Polymérase Chain Réaction
qPCR	PCR quantitative
RT	Reverse Transcription
UV	Ultra Violet
YPG	Yeast / Peptone / Glucose
LB	Bouillon Lysogène
ORF	Open Reading Frame
CDS	Coding DNA Sequence

Table des matières

Remerciements	3
Résumé	5
Abstract	7
Abreviations	9
Table des matières	11
Liste des tableaux	15
Liste des tableaux	15
Liste des figures	17
Liste des annexes	19
Introduction générale	21
Chapitre I : Introduction bibliographique	23
1 Importance des champignons	25
2 Les gènes des types sexuels	27
2.1 Introduction	27
2.2 La reproduction sexuée chez le champignon filamenteux <i>Podospora anserina</i>	28
2.2.1 <i>Podospora anserina</i>	28
2.2.2 Cycle de vie de <i>Podospora anserina</i>	31
2.2.3 Caractérisation moléculaire et fonctionnelle du locus <i>mat</i> chez <i>Podospora anserina</i>	32
2.2.4 Le système phéromones/récepteurs chez <i>Podospora anserina</i>	39
2.3 La reproduction sexuée chez les levures	41
2.3.1 Introduction	41
2.3.2 La reproduction chez <i>Saccharomyces cerevisiae</i>	43
2.3.3 La reproduction chez <i>Schizosaccharomyces pombe</i>	45
3 Les gènes à HMG-box	48
3.1 Introduction	48
3.2 Structure et fonctions des protéines à domaine HMG	50
3.2.1 Structure du domaine HMG	50
3.2.2 Localisation des protéines HMGB dans la cellule	54
3.2.3 Fonction des protéines à domaine HMG	54
4 Evolution des HMG et du locus <i>MAT</i>	65
5 Objectif de la thèse	68
Chapitre II : Matériels et Méthodes	69
1. Les souches de <i>Podospora anserina</i> , <i>Escherichia coli</i> et levure	71
2. Préparation du matériel biologique et extraction des ARN	71

2.1	Cultures et croisements	71
2.2	Extraction des ARN	72
3.	Microarray et analyse des données.....	73
3.1	Etapas de marquage.....	73
3.2	Qualification des cibles	74
3.3	Etapas d'hybridation	74
3.4	Scan, analyses des images et extraction des données brutes.....	75
3.5	Normalisation des données.....	75
3.6	Analyse différentielle	75
4.	Annotation des gènes et recherche des orthologues.....	76
5.	Stratégies de délétion	76
6.	Analyses phénotypiques	77
7.	Séquençage des souches.....	78
7.1	Cultures	78
7.2	Extraction de l'ADN	78
7.3	Purification de l'ADN en gradient de chlorure de césium.....	79
7.4	Construction des banques d'ADN et séquençage	79
8.	Acquisition, analyse et alignement des séquences	80
9.	L'analyse phylogénique	80
10.	PCR quantitative	80
10.1	RT-PCR.....	80
10.2	Choix des amorces	80
10.3	Contrôle de l'efficacité et spécificité de la PCR	81
10.4	RT-qPCR.....	81
11.	Annotation structurale de <i>Pa_1_13940</i>	81
Chapitre III : Article "Genome-Wide Gene Expression Profiling of Fertilization		
Competent Mycelium in Opposite Mating Types in the Heterothallic Fungus		
<i>Podospora anserina</i>"		
82		
1.	Introduction	84
2.	Résumé des résultats et additions non publiées	84
3.	Discussion	87
Chapitre IV : Recherche transcriptomique des gènes cibles des facteurs de transcription		
contrôlant la reconnaissance internucléaire chez <i>Podospora anserina</i>.....		
92		
1.	Introduction	94
2.	Résultats et discussion.....	97

2.1	Choix du temps de prélèvement des croisements pour l'analyse de la reconnaissance internucléaire	97
2.2	Recherche des gènes impliqués dans la reconnaissance internucléaire selon le modèle dit « d'identité nucléaire »	97
2.2.1	Principe de comparaison	97
2.2.2	Analyse transcriptomique et identification des gènes candidats	98
2.2.3	Fonctions des gènes.....	102
2.2.4	Construction de mutants de délétion des gènes cibles	103
2.3	Recherche des gènes impliqués dans la reconnaissance internucléaire selon le modèle dit de « ségrégation aléatoire ».....	103
2.3.1	Principe de comparaison et analyse des gènes	103
2.3.2	Fonctions des gènes.....	104
2.3.3	Analyse phénotypique et délétion des gènes cibles	105
3.	Conclusion.....	106
Chapitre V : Etude des gènes à HMG box chez <i>Podospora anserina</i>		108
1.	Introduction	110
2.	Résultats	111
2.1	Caractérisation des protéines à domaine HMG chez <i>Podospora anserina</i>	111
2.2	Analyse phylogénique des protéines à domaine HMG chez <i>Podospora anserina</i>	116
2.3	Analyse phénotypique des mutants HMG chez <i>Podospora anserina</i>	123
2.4	Analyse des mutants HMG affectés dans le développement sexué	128
2.4.1	Identification des relations de régulation entre les gènes à HMG-box	128
2.4.2	Profils d'expression des gènes <i>mat</i> et de leurs cibles dans les mutants HMG.....	129
2.5	Caractérisation de <i>Pa_1_13940</i>	132
2.5.1	Structure de <i>Pa_1_13940</i>	132
2.5.2	Le mutant ΔPa_1_13940	132
2.5.3	Site de fixation potentiel du facteur de transcription codé par <i>Pa_1_13940</i>	135
3.	Discussion	136
3.1	La majorité des protéines à domaine HMG sont impliquées dans le développement sexué de <i>Podospora anserina</i>	136
3.2	ΔPa_1_13940 possède un phénotype particulier de développement sexué.....	138
3.3	<i>Pa_1_13940</i> régule le développement sexué chez <i>Podospora anserina</i>	139
3.4	Les protéines à domaine HMG forment un réseau de regulation complexe	140
4.	Conclusion.....	142
Conclusion générale		144

Annexes	146
Article	178
Références bibliographiques	196

Liste des tableaux

Tableau 1: Nomenclature des protéines HMG.....	49
Tableau 2: Les gènes spécifiques des cellules <i>M</i> et les gènes spécifiques des cellules <i>P</i> régulés par Ste11p.....	62
Tableau 3: Comparaison des motifs TR des gènes régulés par Ste11p.....	64
Tableau 4 : Les conditions de culture des souches de <i>P. anserina</i> destinées à l'extraction des ARN.	72
Tableau 5: Comparaison entre les gènes spécifiques impliqués dans la fécondation chez <i>S. pombe</i> , <i>S. cerevisiae</i> et <i>P. anserina</i>	89
Tableau 6: Classification des 43 gènes différentiellement exprimés entre le croisement ♂ <i>fprI</i> ⁻ X ♀ <i>WT mat</i> ⁻ et le croisement ♂ <i>fmrI smr2</i> ⁻ X ♀ <i>WT mat</i> ⁺	100
Tableau 7: Liste des gènes différentiellement exprimés par un effet femelle dans la comparaison ♂ <i>fprI</i> ⁻ X ♀ <i>WT mat</i> ⁻ /♂ <i>fmrI smr2</i> ⁻ X ♀ <i>WT mat</i> ⁺ (liste 1) et dans la comparaison ♂ <i>WT mat</i> ⁺ X ♀ <i>WT mat</i> ⁻ /♂ <i>WT mat</i> ⁻ X ♀ <i>WT mat</i> ⁺ (liste 4).	100
Tableau 8 : Gènes régulés par FPR1 chez <i>mat</i> ⁺ et par FMR1 et SMR2 chez <i>mat</i> ⁻	101
Tableau 9: Les gènes sélectionnés pour la délétion dans la liste des gènes nécessaires à la formation des hyphes ascogènes.	105
Tableau 10: Caractéristiques des protéines à domaine HMG de <i>P. anserina</i>	114
Tableau 11: Les protéines à domaine HMG utilisées dans la construction du phylogramme.	120
Tableau 12: Phénotypes induits par les délétions des gènes à HMG-box chez <i>P. anserina</i> . .	124
Tableau 13: Résultats de RT-qPCR des gènes <i>Pa_6_4110</i> , <i>Pa_7_7190</i> , <i>Pa_1_14230</i> et <i>Pa_1_13940</i> dans les mutants HMG ΔPa_6_4110 , ΔPa_7_7190 , ΔPa_1_14230 et ΔPa_1_13940 par rapport au sauvage.....	128
Tableau 14 : Résultats de RT-qPCR des gènes <i>mat</i> et certains de leurs cibles dans les mutants HMG.....	131
Tableau 15: La TR-box est présente dans les régions 5'UTR (1000pb avant l'ATG) de <i>Pa_1_13940</i> , des gènes des types sexuels et des cibles de FMR1 et FPR1.....	136

Liste des figures

Figure 1: La phylogénie des ascomycètes.....	29
Figure 2: Structures végétatives organes reproducteurs de <i>P. anserina</i>	30
Figure 3: Le cycle sexué de <i>P. anserina</i>	32
Figure 4: Structure génétique des idiomorphes <i>mat-</i> et <i>mat+</i> de <i>P. anserina</i>	33
Figure 5 : Régulation de la fécondation par FPR1, FMR1 et SMR2.	34
Figure 6: Régulation de la reconnaissance internucléaire par les gènes <i>mat</i>	36
Figure 7: Comparaison entre le croisement sauvage et les croisements mutant x sauvage à partir du stade de fécondation jusqu'à la formation des spores.	37
Figure 8 : Régulation de l'expression des phéromones et de leurs récepteurs par les facteurs de transcription MAT.	41
Figure 9: Cycle de vie de <i>Saccharomyces cerevisiae</i> et de <i>Schizosaccharomyces pombe</i>	43
Figure 10: Le locus <i>mat</i> et les protéines essentielles dans la conjugaison chez <i>S. cerevisiae</i> et <i>S. pombe</i>	44
Figure 11: Régulation des types sexuels chez <i>S. cerevisiae</i>	44
Figure 12 : La transduction du signal phéromone chez <i>S. pombe</i>	47
Figure 13 : Alignement des séquences protéiques des HMG-box.	51
Figure 14: Structure en L d'une boîte HMG.	52
Figure 15 : Modèle de fonctionnement des protéines HMGB.	55
Figure 16: La protéine Ste11p.	59
Figure 17 : Régulation de la conjugaison chez <i>S. pombe</i>	61
Figure 18 : Modèle de fixation Ste11p-ADN.....	64
Figure 19: La structure des types sexuels chez les champignons.	67
Figure 20: Organigramme de la classification des gènes différentiellement exprimés entre les souches <i>mat+</i> et <i>mat-</i>	85
Figure 21: Organigramme de la classification des gènes contrôlés de la même façon par FPR1 et FMR1.....	86
Figure 22: Modèle de l'identité nucléaire.	95
Figure 23: Modèle de ségrégation aléatoire.	96
Figure 24: Plan de comparaison des croisements utilisés dans l'étude de la reconnaissance internucléaire selon le modèle1.....	98
Figure 25 : Plan de comparaison des croisements utilisés dans l'étude de la reconnaissance internucléaire selon le modèle 2.....	104
Figure 26: Les caractéristiques de la protéine mtHMGB1.	111

Figure 27: Alignement du consensus HMG avec les protéines à domaine HMG de <i>P. anserina</i>	115
Figure 28: Phylogramme non raciné des domaines HMG	123
Figure 29: Croisements hétérozygotes entre les souches mutantes <i>mat+</i> et la souche sauvage <i>mat-</i>	126
Figure 30 : Comparaison de la distribution des périthèces entre les croisements sauvages et les croisements homozygotes mutants ΔPa_1_11050	127
Figure 31: la voie de régulation des gènes à HMG-box.....	129
Figure 32: Le gène <i>Pa_1_13940</i> et la position des introns. alignement des domaines HMG de <i>Pa_1_13940</i> et ses orthologues FMF-1 de <i>N. crassa</i> et Ste11p de <i>S. pombe</i>	133
Figure 33 : Caractéristiques de la protéine <i>Pa_1_13940</i>	134
Figure 34 : Développement des périthèces au point de confrontation entre le mutant ΔPa_1_13940 et le sauvage.	134
Figure 35: Réseaux de régulation entre les protéines à domaine HMG chez <i>S. pombe</i> , <i>U.</i> <i>maydis</i> , les mammifères et <i>P. anserina</i>	141

Liste des annexes

Annexe 1: Composition de la référence utilisée dans les expériences de microarray.....	148
Annexe 2: Protocole de préparation des ARN (Méthode QIAGEN).....	149
Annexe 3 : La technique de délétion des gènes chez <i>P. anserina</i>	149
Annexe 4: Protocole de préparation de la banque « paire end » selon Illumina.....	161
Annexe 5 : Sites d'internet utilisé au cours de cette étude.....	167
Annexe 6 : Liste des gènes différentiellement exprimés dans la comparaison <i>mat+ fprI</i> chez <i>P. anserina</i> et adjacentes aux orthologues des gènes différentiellement exprimés dans la comparaison <i>MAT1-2-1/ΔMAT1-2-1</i>	168
Annexe 7 : Liste des gènes différentiellement exprimés dans la comparaison du croisement ♂ <i>fprI</i> X ♀ <i>WT mat-</i> par rapport au croisement ♂ <i>fmrI-smr2</i> X ♀ <i>WT mat+</i>	171
Annexe 8 : Liste des gènes communs entre les comparaisons ♂ <i>WT mat+</i> X ♀ <i>WT mat- ♂ fprI</i> X ♀ <i>WT mat-</i> (liste2) et ♂ <i>WT mat-</i> X ♀ <i>WT mat+ ♂ fmrI-smr2-</i> X ♀ <i>WT mat+</i> (liste3).....	172
Annexe 9: Les couleurs des résidus présentent dans les alignements des acides aminés	176
Annexe 10: Résultats de la quantification de la transcription des gènes dans les mutants HMG par rapport au sauvage.....	177

Introduction générale

La régulation de la reproduction sexuée est l'un des processus fondamentaux dans le cycle de vie des eucaryotes. Chez les champignons ascomycètes hétérothalliques, la reproduction sexuée n'est possible qu'entre deux individus de types sexuels opposés. Ces types sexuels sont déterminés par un locus existant sous deux formes alternatives *MAT1-1* et *MAT1-2*. Les gènes présents au niveau du locus *MAT* codent des facteurs de transcription régulateurs des gènes impliqués dans le développement sexué. Le système de régulation par les gènes des types sexuels est largement étudié chez les deux levures *Saccharomyces cerevisiae* et *Schizosaccharomyces pombe*. Chez ces deux espèces, les gènes cibles des facteurs de transcription codés par les gènes *mat* ont été identifiés, dont la plupart ont une fonction connue dans la reproduction sexuée. Chez les ascomycètes filamenteux, la reproduction sexuée est beaucoup plus complexe que celle des ascomycètes unicellulaires. En effet, contrairement aux levures, les Pezizomycotina développent des organes femelles complexes et certains champignons hétérothalliques produisent également des cellules mâles qui fécondent les organes femelles de type sexuel opposé. En plus, cette étape de fécondation n'est pas immédiatement suivie par la caryogamie: entre ces deux événements s'intercalent toute une série d'étapes qui vont permettre d'obtenir plusieurs centaines de méioses indépendantes à partir d'un événement de fécondation. La reproduction sexuée chez les ascomycètes filamenteux est bien décrite sur le plan cytologique. Par contre, les mécanismes moléculaires mis en jeu au cours des différentes étapes restent mal ou même complètement inconnus, en particulier pour les plus originaux d'entre eux. En effet, jusqu'à une date récente, les seuls gènes identifiés comme cibles de ces facteurs de transcription *MAT* ont été les précurseurs des phéromones et leurs récepteurs. Les approches génétiques, destinées à rechercher d'autres gènes cibles ont toutes échoué, probablement en raison de la complexité du système de régulation. Pour cela, une étude globale des gènes régulés par les facteurs de transcription *MAT* a été réalisée chez le champignon filamenteux *P. anserina* avec une approche microarray. *P. anserina* est un champignon hétérothallique, appartenant au groupe des Pezizomycotina. Il possède deux types sexuels, *mat+* et *mat-*, contrôlés chacun par une séquence spécifique. La séquence *mat-* contient trois gènes, *FMRI*, *SMR1* et *SMR2*. La séquence *mat+* contient un seul gène : *FPRI*. La fonction moléculaire de *SMR1* est inconnue et les autres gènes codent des facteurs de transcription à domaine HMG. Ces protéines contrôlent au moins trois étapes au cours du cycle sexué de *P. anserina*: la reconnaissance intercellulaire entre *mat+* et *mat-* pendant la fécondation, la reconnaissance internucléaire entre les noyaux *mat+* et *mat-* pendant la transition des cellules plurinucléées aux cellules dicaryotiques, et la formation des hyphes ascogènes. Le domaine HMG des facteurs de

transcription MAT appartient au groupe MATA_HMG dans le cas de FPR1 et SMR2, et au groupe MAT α _HMG dans le cas de FMR1. Les protéines de ces groupes sont caractérisées par la présence d'un seul domaine HMG et une fixation spécifique à l'ADN. Le groupe MATA_HMG inclut aussi Ste11p, le facteur clé du développement sexué de *S. pombe*. Les protéines à domaine HMG semblent avoir un rôle majeur dans le développement sexué des champignons.

Dans ma thèse j'ai étudié plus particulièrement la régulation de la reconnaissance intercellulaire et la reconnaissance internucléaire en identifiant les cibles des facteurs de transcription MAT. En plus j'ai invalidé toutes les protéines à domaine HMG afin de déterminer leur rôle dans le développement sexué de *P. anserina*

Chapitre I : Introduction bibliographique

1 Importance des champignons

Les champignons représentent l'un des plus importants groupes d'organismes, à la fois sur le plan quantitatif et qualitatif (Mueller and Bills 2004). On estime à 1,5 millions le nombre d'espèces de champignons vivant sur Terre (Hawksworth 2001). Cependant, seulement 14 % de ces organismes ont été découverts, ce qui peut être expliqué par la difficulté de cultiver ces organismes en conditions de laboratoire mais aussi en raison d'un manque de connaissances et d'un sous-échantillonnage de ces organismes dans de nombreux biotopes de la planète (Hawksworth and Rossman 1997; Hawksworth 2001). Les champignons représentent un groupe d'organismes ayant un rôle majeur dans le fonctionnement des écosystèmes ainsi que dans des applications biotechnologiques (Lutzoni et al. 2004). Leur capacité d'exploration via l'extension des hyphes, couplée à leur capacité de sécrétion d'enzymes hydrolytiques, a permis une colonisation d'une grande variété de substrats. Les champignons saprophytes participent au cycle de l'azote et du carbone par la dégradation de la matière organique (Lutzoni et al. 2004). Ils ont la capacité de consommer la cellulose ainsi que la lignine et ils sont considérés comme les principaux recycleurs de la matière organique issue des plantes (Lutzoni et al. 2004). Dans leur activité de décomposition les champignons peuvent attaquer les produits commerciaux (tissus, cuir, peintures, papier) et changer la valeur nutritive des aliments (viande et légumes) (Susan-E Eichhorn et al. 2000). L'importance des champignons comme ravageurs est amplifiée par le fait qu'ils sont capables de se développer dans des conditions diverses (Susan-E Eichhorn et al. 2000). De nombreux champignons pathogènes et parasitaires s'attaquent aux organismes vivants, comme par exemple les plantes, les insectes et les animaux (Lutzoni et al. 2004). D'autres champignons fonctionnent comme des symbiotes avec les plantes (endophytes, endomycorhizes, ectomycorhizes) et les algues (lichens). Ces associations ont permis aux autres organismes d'exploiter de nouvelles ressources et de coloniser de nouveaux habitats (Lutzoni et al. 2004). Certains champignons sont capables d'affecter l'homme dont le système immunitaire est affaibli, par exemple *Pneumocystis carinii*, *Aspergillus fumigatus* et *Cryptococcus neoformans* qui s'attaquent aux malades du SIDA (Phair et al. 1990; Marr et al. 2002; Bose et al. 2003).

Les caractères qui font des champignons des ennemis aussi importants sont aussi à l'origine de leurs valeurs commerciales. On peut par exemple citer les processus fermentaires de certaines levures utilisées dans de nombreux domaines en agro-alimentaire, telle que la « levure de boulanger » *Saccharomyces cerevisiae*, utilisée pour la panification mais également en brasserie (Susan-E Eichhorn et al. 2000). Cette levure est aussi un modèle d'études

important en génétique et biologie moléculaire et cellulaire. D'autres champignons donnent les saveurs et les arômes qui caractérisent différents types de fromages (Susan-E Eichhorn et al. 2000). Les champignons sont aussi utilisés dans l'élimination des déchets toxiques, par exemple *Phanerochaete chrysosporium* (John A. et al. 1987). La découverte en 1928 de la pénicilline produite par une souche fongique *Penicillium notatum*, a ouvert la voie à la biosynthèse d'antibiotiques utilisés depuis 1940. On recensait en 2006 plus de 8600 métabolites bioactifs produits par des souches fongiques, dont 4900 auraient une activité antibiotique (Berdy 2005). Certaines substances dérivées des champignons sont utilisées pour supprimer les réactions immunitaires lors de la transplantation d'organes, comme la cyclosporine sécrétée par *Tolypocladium inflatum* (Abdel-Fattah et al. 2007; El Enshasy et al. 2008). Des dizaines d'espèces différentes de champignons sont régulièrement consommées par l'homme et certaines sont cultivées à grandes échelles. Cependant, en considérant le grand nombre d'espèces fongiques non cultivées, de nombreuses souches productrices de molécules bioactives restent vraisemblablement à découvrir.

Les champignons supérieurs, basidiomycètes et les ascomycètes, regroupent 98% des espèces connues. Les Ascomycota sont le sujet d'études intensives. On compte 45000 espèces décrites à ce jour (Hawksworth 2001; Taylor et al. 2004), ils comprennent un certain nombre de champignons familiers et économiquement importants (levures, morilles et truffes). Les ascomycètes sont également responsables de plusieurs maladies graves des plantes comme l'oïdium qui s'attaque aux feuilles (Susan-E Eichhorn et al. 2000). D'autres ascomycètes sont utilisés comme modèles d'étude au laboratoire notamment *Neurospora crassa*, *Aspergillus nidulans* et *P. anserina*. Ces espèces ont leur génome séquencé et annoté (Galagan et al. 2003; Galagan et al. 2005; Espagne et al. 2008). Actuellement, 200 à 300 génomes de champignons sont séquencés et le projet F1000 (1000 Fungal Genomes Project) vise le séquençage de 1000 génomes de champignons, avec 2 espèces séquencées par famille. Ce projet est destiné à promouvoir les études scientifiques et économiques sur les champignons voir le site: <http://1000.fungalgenomes.org/home/>.

Les champignons ont des modes de reproduction variant suivant les espèces, d'une reproduction exclusivement asexuée à une reproduction combinant le mode asexué et le mode sexué. Il faut noter que la plupart des champignons utilisés pour la synthèse et la semi-synthèse apparaissent asexués. Au cours des dernières années, la reproduction sexuée des champignons a fait l'objet d'études approfondies afin de développer de nouvelles techniques d'amélioration des champignons utilisés dans l'industrie, en utilisant la reproduction sexuée. Sur le plan médical, ces études visent à élaborer des stratégies de contrôle de la croissance chez les champignons pathogènes et de prédire et maîtriser la transmission des gènes de

résistance aux principes actifs utilisés contre les champignons pathogènes des plantes et des animaux.

2 Les gènes des types sexuels

2.1 Introduction

L'étude sur la sexualité des champignons a commencé il y a plus d'un siècle (Blakeslee 1904). La vaste majorité des champignons sont capables de se reproduire par voie sexuée et asexuée. Parmi les champignons se reproduisant par voie sexuée, on distingue des espèces homothalliques et hétérothalliques. En 1904, Blakeslee (Blakeslee 1904) a montré que certaines espèces de *Mucorineae* (Mucoromycotina, anciennement zygomycètes) sont divisées en deux groupes, définis par les signes + et -, sexuellement compatibles et morphologiquement identiques. Il a ainsi introduit les notions d'hétérothallisme et de types sexuels. L'hétérothallisme a été découvert ensuite chez plusieurs champignons, en particulier les champignons filamenteux ascomycètes (Edgerton 1914). En 1920, des études réalisées sur plusieurs espèces de *Neurospora* et *Ascobolus*, ont montré que les types sexuels sont déterminés par des allèles présents au locus *mat* (Dodge 1920; Shear and Dodge 1927; Shiu and Glass 2000). Les ascomycètes filamenteux possèdent un seul locus *MAT* avec deux allèles (Shiu and Glass 2000). L'analyse moléculaire de ces allèles a montré qu'ils correspondaient à des séquences différentes contenant plusieurs gènes. Ce système bipolaire est aussi observé chez les levures *S. cerevisiae* et *Schizosaccharomyces pombe*. Les types sexuels de ces deux levures ont été largement étudiés. Ils codent des facteurs de transcription qui régulent l'expression des gènes spécifiques, en particulier les phéromones et les récepteurs nécessaires à la conjugaison. Ces facteurs de transcription contrôlent aussi, mais d'une façon indirecte, la caryogamie et la méiose. Malgré la similarité présente entre *S. cerevisiae* et *S. pombe*, le nombre et la nature des facteurs de transcription et les voies de régulations de la reproduction sexuée sont différents entre ces deux organismes (Coppin et al. 1997).

Chez les ascomycètes filamenteux la reproduction sexuée est beaucoup plus complexe que celle des levures. Chez les levures, la fusion des deux partenaires est immédiatement suivie par la caryogamie. En revanche, chez les ascomycètes filamenteux la fécondation entre un organe reproducteur femelle et un gamète mâle de type sexuel opposé n'est pas immédiatement suivie par la caryogamie entre ces deux événements s'intercalent toute une série d'étapes qui vont permettre d'obtenir plusieurs centaines de méioses indépendantes à partir d'un événement de fécondation. Plusieurs études ont montré que l'étape de la fécondation et les étapes du développement de la fructification sont sous le contrôle des gènes des types sexuels (Debuchy et al. 1993). Pour identifier précisément ces étapes et déterminer

leur contrôle et les mécanismes moléculaires mis en jeu, nous avons utilisé comme système modèle *P. anserina*. Ce champignon est le plus étudié parmi les champignons filamenteux au niveaux de la structure et la fonction des protéines codées par le locus *MAT*.

2.2 La reproduction sexuée chez le champignon filamenteux

Podospora anserina

2.2.1 *Podospora anserina*

Le champignon filamenteux *P. anserina* est un ascomycète coprophile qui vit sur les excréments d'herbivores. Il appartient au sous-embranchement des Pezizomycotina, à la classe des Sordariomycetes (Figure 1).

L'appareil végétatif de *P. anserina* est constitué d'hyphes formant un réseau appelé le mycélium (Figure 2 A et B). Ces hyphes sont des alignements d'articles plurinucléés séparés par des parois perforées ou septum typiques des ascomycètes. Les pores permettent l'échange du cytoplasme et des noyaux entre les articles. Deux hyphes différents peuvent aussi entrer en contact et fusionner pour échanger leur cytoplasme et éventuellement des noyaux, c'est ce qu'on appelle le processus d'anastomose. Il est possible de diviser le mycélium en petits fragments chacun, capable de régénérer un nouveau thalle lorsqu'il est placé sur un milieu nutritif. Le mycélium adopte des morphologies et colorations variées en fonction du milieu de culture sur lequel il pousse. Le mycélium en phase stationnaire se pigmente en vert, produit des hyphes aériens et développe des structures qui permettent la reproduction sexuée.

P. anserina est un organisme hétérothallique hermaphrodite. Il possède deux types sexuels *mat+* et *mat-*. Chaque thalle différencie des gamètes mâles (les spermaties) et des organes femelles (les ascogones) (Figure 2 C et D). Les spermaties ne germent pas, et contrairement à d'autres modèles fongiques comme *A. nidulans* et *N. crassa*, *P. anserina* ne produit pas de spores végétatives participant à la reproduction asexuée. *P. anserina* a un cycle haplobiontique (la caryogamie est immédiatement suivie par la méiose) d'environ une semaine (Figure 3). Chez cet organisme haploïde, les mutations sont donc mises en évidence facilement et rapidement. Les asques contiennent 4 ascospores, chacune contenant initialement deux noyaux haploïdes.

Le génome de la souche *mat+* entièrement séquencé par le Genoscope est d'une taille de 36 Mb repartis sur 7 chromosomes (Espagne et al. 2008). A ce jour 10546 CDS (Coding DNA Sequence) ont été annotés. Ces données sont disponibles sur le site: <http://podospora.igmors.u-psud.fr>.

Le modèle d'étude, *N. crassa* est le plus proche de *P. anserina* et leurs lignées auraient divergées depuis 75 millions d'années (Paoletti and Saupe 2008). Malgré leur relative

proximité, les études génétiques et moléculaires sur *P. anserina* montrent que les deux espèces présentent des différences importantes. *P. anserina* donne accès à d'autres phénomènes biologiques que *N. crassa*. En particulier, le phénomène de la sénescence et l'instabilité du génome mitochondrial (Espagne et al. 2008).

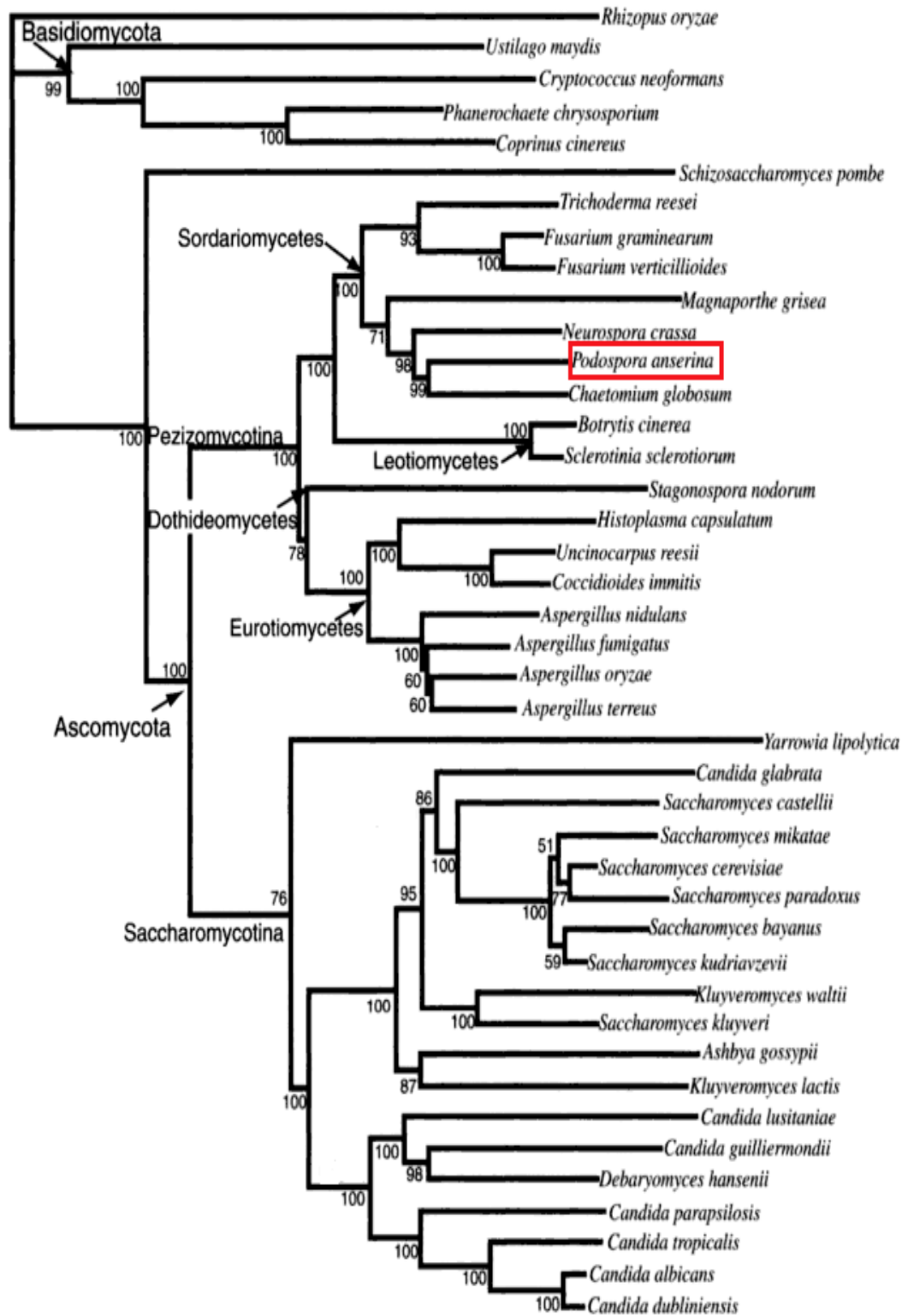


Figure 1: La phylogénie des ascomycètes. L'arbre est construit à partir de 4,805 familles de gènes appartenant à 42 espèces de champignons. *Rizopus oryzae* est utilisée comme groupe externe (Heitman et al. 2007).

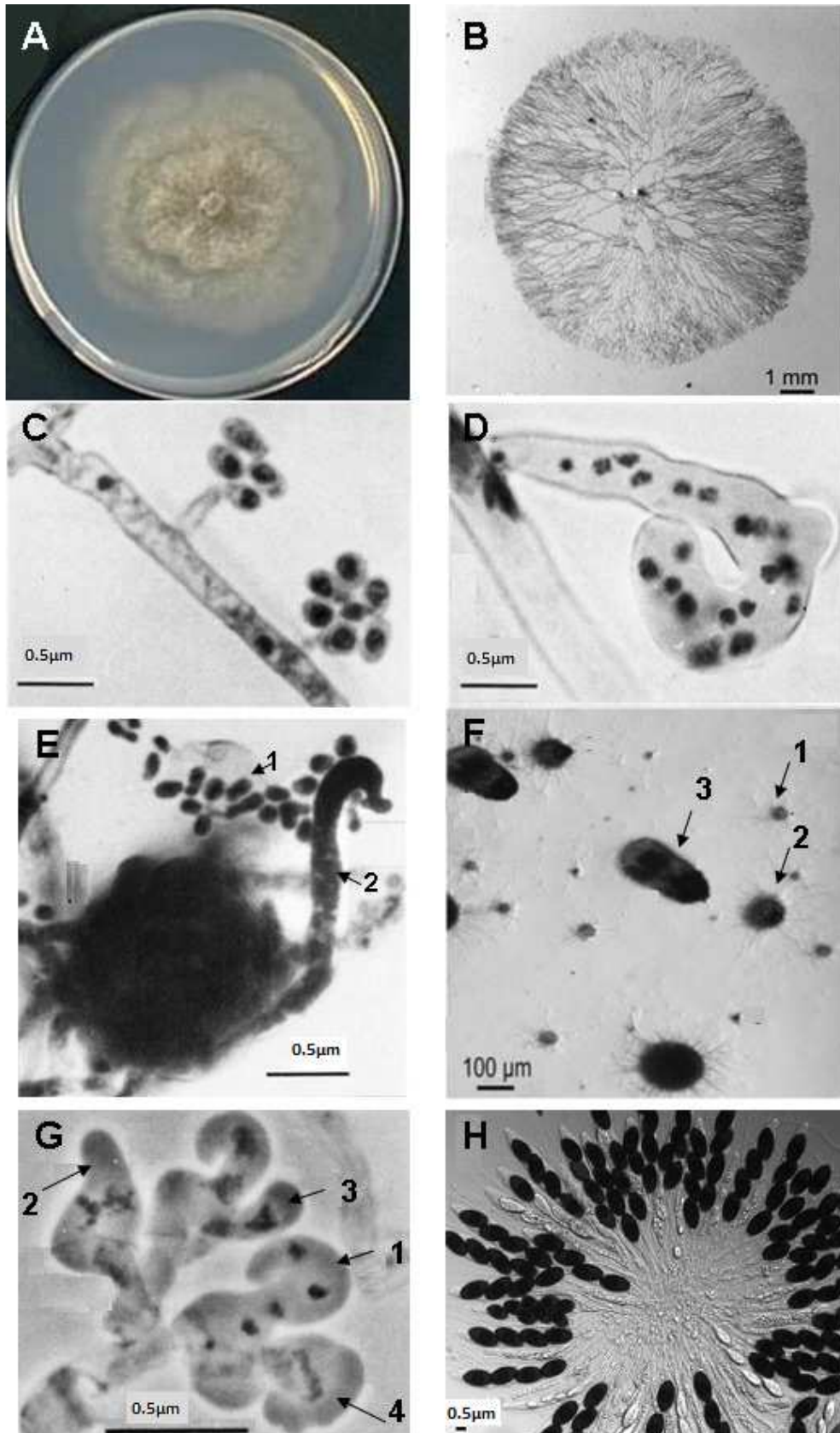


Figure 2: Structures végétatives organes reproducteurs de *P. anserina*.

(A) *P. anserina* cultivé en milieu M2 (d'après <http://podospora.igmors.u-psud.fr/more.php>); (B) Le réseau de mycélium; (C) Différentiation des gamètes mâles (spermaties); (D) Organe femelle (ascogone); (E) Fécondation d'un ascogone par les spermaties : (1) les spermaties, (2) le trichogyne; (F) Différentes étapes de différenciation des périthèces : (1) ascogone non fécondé, (2) périthèce fécondé, (3) périthèce mûr ; (G) Différentes étapes de formation des asques : (1) crochet avec quatre noyaux, (2) la cellule supérieure binucléée s'allonge pour former l'asque, (3) formation d'un nouveau crochet, (4) précaryogamie dans l'asque ; (H) Asques avec quatre spores. B, F, H d'après (Pinan-Lucarre et al. 2007) ; C, D, E, G d'après (Zickler et al. 1995).

2.2.2 Cycle de vie de *Podospora anserina*

Le cycle de reproduction de *P. anserina* (Figure 3) commence par la différenciation des gamètes mâles (spermaties) (Figure 2 C) et des organes femelles (ascogones) (Figure 2 D) qui se forment indépendamment des types sexuels. La reconnaissance intercellulaire entre ces deux structures reproductives de types sexuels opposés fait intervenir le système phéromone/récepteur qui guide la croissance d'un hyphe spécialisé de l'ascogone (trichogyne) vers les cellules mâles (Coppin et al. 2005) (Figure 2 E). La fusion des membranes des deux structures est suivie par la migration du noyau de la spermatie dans l'organe femelle. Chez *P. anserina* l'événement de la fécondation n'est pas immédiatement suivi par la fusion des deux noyaux reproducteurs. La fécondation est suivie par un processus complexe dans une structure particulière : le périthèce (Figure 2 F). Les noyaux d'origine mâle et femelle se divisent et donnent naissance à des cellules plurinucléées (syncytium) d'où vont émerger des cellules dicaryotiques (un noyau de chacun des parents). La transition de l'état syncytial plurinucléé à un état dicaryotique nécessite une étape de reconnaissance internucléaire entre deux noyaux de types sexuels opposés. Ces cellules dicaryotiques vont se diviser pour former les hyphes ascogènes dont les extrémités donnent naissance à des cellules particulières, en forme de crochet dans lesquelles les deux noyaux haploïdes subissent une mitose suivie par la formation de trois cellules différentes: la cellule basale et la cellule latérale, toutes deux uninucléées, et une cellule apicale binucléée. Le noyau de la cellule latérale migre à travers un pore dans la cellule basale, cette cellule binucléée forme un nouveau crochet (Zickler et al. 1995) (Figure 2 G). La cellule apicale s'allonge pour former l'asque binucléé dans lequel aura lieu la caryogamie qui est immédiatement suivie de la méiose, puis de la mitose post-méiotique et de la formation de 4 ascospores (Figure 2 H) binucléées ou 5 ascospores (3 binucléées et 2 uninucléées, Figure 3) avec une fréquence de 1 à 2 %.

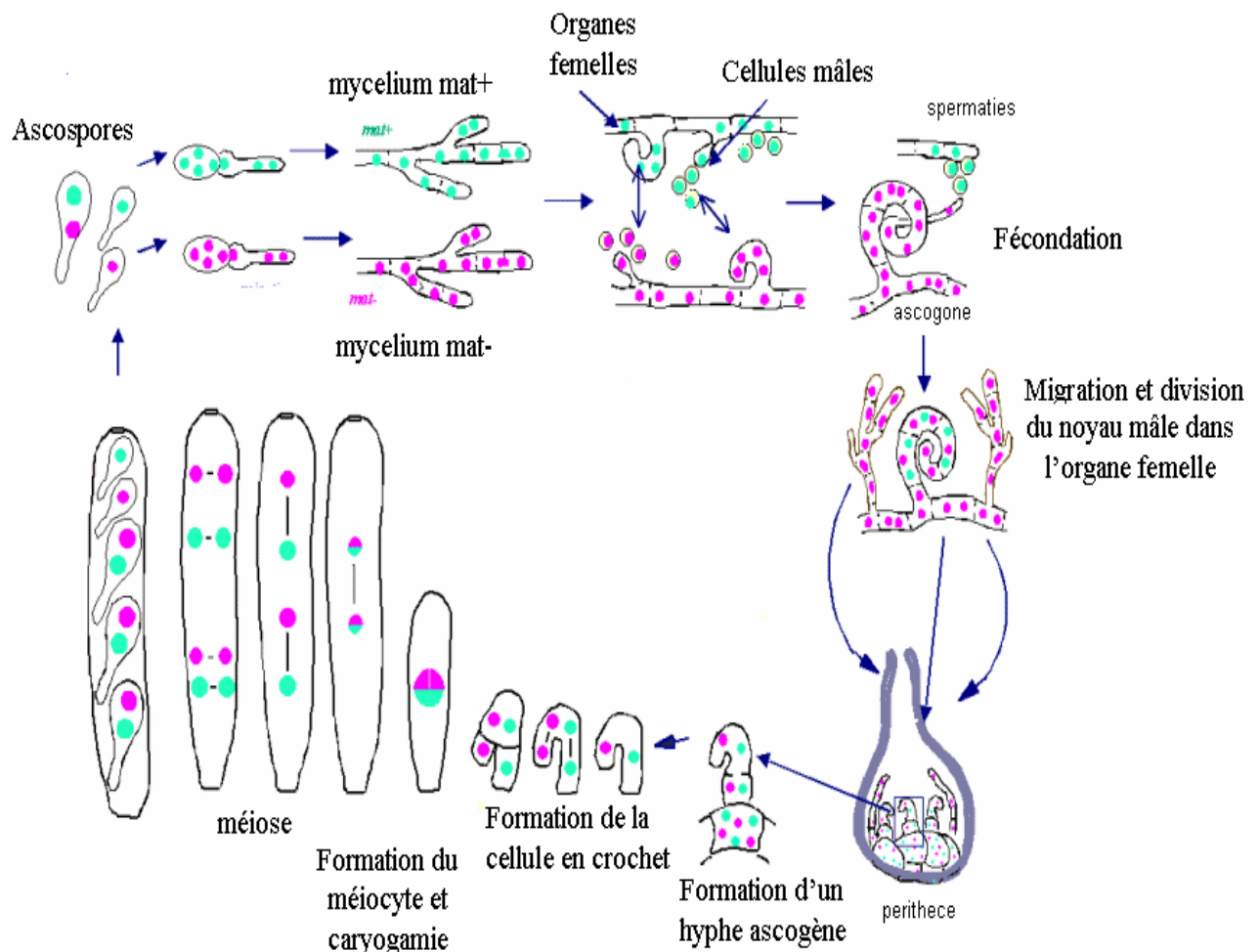


Figure 3: Le cycle sexué de *P. anserina*.
 (D'après <http://podospira.igmors.u-psud.fr/more.php>)

2.2.3 Caractérisation moléculaire et fonctionnelle du locus *mat* chez

Podospira anserina

2.2.3.1 Description du locus *mat*

La reproduction sexuée chez *P. anserina* dépend du type sexuel qui est déterminé par un locus *mat* avec deux idiomorphes: *mat+* et *mat-* (ou *MATI-1* et *MATI-2* selon la nomenclature standard) (Turgeon and Debuchy 2007). Des analyses moléculaires ont montré que ces deux idiomorphes occupent le même endroit sur le chromosome 1.

L'idiomorphe *mat+* est une séquence de 3,8 Kb (Figure 4) contenant un seul gène, *FPR1* (Fertilization Plus Regulator 1) ou *MATI-2-1*. Ce gène code un facteur de transcription à motif MATA_HMG (High Mobility Group).

L'idiomorphe *mat-* est une séquence de 4,7 Kb (Figure 4) contenant trois gènes: *FMR1*, *SMR1* et *SMR2*. *FMR1* (Fertilization Moins Regulator 1) ou *MATI-1-1* code une protéine ayant un domaine $\alpha 1$ classé récemment dans le groupe des MATA α _HMG (Martin et al. 2010). La protéine *SMR2* (Sporeulation Minus Regulator 2) comme la protéine *FPR1*

contient un domaine MATA_HMG (Debuchy and Coppin 1992). La protéine SMR1 (Sporulation Minus Regulator 1) contient un domaine PPF avec deux résidus proline et un résidu phenylalanine conservés chez SMR1 et ses orthologues (Kanematsu et al. 2007). Le rôle de ce domaine est encore inconnu (Turgeon and Debuchy 2007). SMR1 est localisée au niveau du cytoplasme (Turgeon and Debuchy 2007).

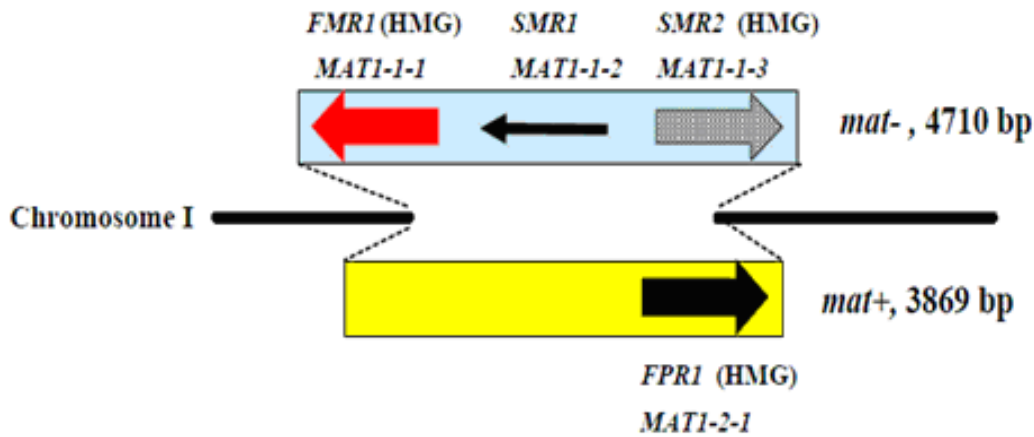


Figure 4: Structure génétique des idiomorphes *mat-* et *mat+* de *P. anserina*.

2.2.3.2 Rôles des gènes *mat* dans le cycle sexué

Le locus *mat* semble avoir un rôle majeur dans la régulation d'au moins trois étapes du cycle sexué: la reconnaissance intercellulaire entre deux partenaires pendant la fécondation, la reconnaissance internucléaire durant la transition d'un syncytium plurinucléé aux cellules dicaryotiques et la formation des hyphes ascogènes après la reconnaissance internucléaire (Arnaise et al. 2001). La délétion du locus *mat* chez *P. anserina* confirme que ce locus contient des gènes essentiels pour la régulation de la reproduction sexuée. La souche Δmat est cependant capable de produire des spermaties et des ascogones, indiquant que les gènes des types sexuels ne contrôlent pas leur formation (Coppin et al. 1993; Turgeon and Debuchy 2007)

2.2.3.2.1 La fonction des gènes des types sexuels pendant la fécondation

La fécondation chez *P. anserina* nécessite la reconnaissance entre la cellule mâle (spermatie) et l'organe femelle pluricellulaire (ascogone) de type sexuel opposé (Figure 3). Les analyses génétiques et moléculaires ont montré que FPR1 chez *mat+* et FMR1 et SMR2 chez *mat-* contrôlent la fécondation. Les parties N-terminales de FMR1 et FPR1 sont nécessaires et suffisantes pour la fécondation (Debuchy and Coppin 1992). Ces régions sont précisément celles qui contiennent les domaines HMG. Les facteurs de transcription FMR1 et FPR1 se comportent respectivement comme des activateurs des gènes *mat-* ou des gènes

mat+ impliqués dans la fécondation (Debuchy and Coppin 1992). Cependant, l'analyse de certains mutants des gènes *FMR1*, *SMR2* et *FPR1* révèle que les modalités du contrôle sont plus complexes qu'une simple relation d'activation entre la protéine régulatrice et ses gènes cibles (Arnaise et al. 2001). Une souche mutante *fmr1⁻* ayant conservé la partie HMG-box de ce gène est capable de se croiser avec une souche *mat+* de référence, mais d'une façon très surprenante elle devient capable de s'autoféconder et de se croiser avec une souche *mat-* avec une faible efficacité. Cette souche *fmr1⁻* exprime donc les deux fonctions *mat-* et *mat+*. L'expression des fonctions *mat+* dans *fmr1⁻* suggère que FMR1 agit non seulement comme un activateur des fonction *mat-* mais aussi comme un répresseur des fonctions *mat+* nécessaires à la fécondation (Arnaise et al. 2001) (Figure 5). Une souche mutante *smr2⁻* est capable de s'autoféconder et de se croiser avec une souche *mat-*. Comme FMR1, SMR2 apparait être un répresseur des fonctions de fécondation *mat+*. Mais contrairement à FMR1 il n'est pas requis pour l'expression des fonctions de fécondation *mat-* (Arnaise et al. 2001) (Figure 5). Des résultats équivalents à ceux de FMR1 ont été obtenus avec FPR1. Une souche mutante *fpr1⁻* ayant conservé la partie HMG-box de ce gène devient capable de s'autoféconder et de se croiser avec une souche *mat+*. FPR1 agit non seulement comme activateur des fonctions *mat+* mais aussi comme répresseur des fonctions *mat-* nécessaires à la fécondation (Figure 5).

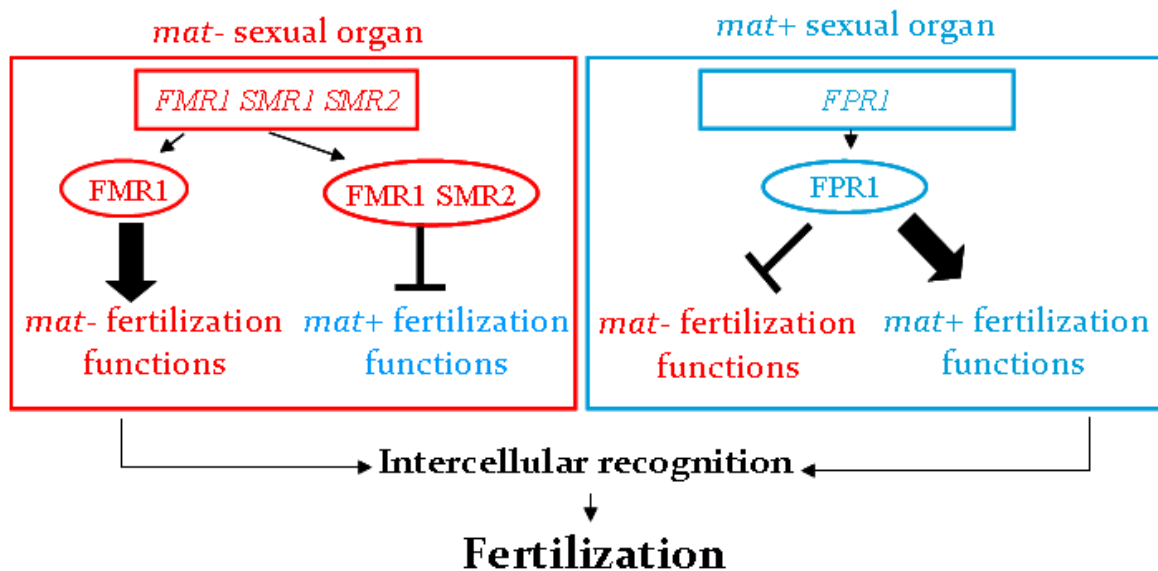


Figure 5 : Régulation de la fécondation par FPR1, FMR1 et SMR2. L'activation est représentée par une flèche, la répression est représentée par une barre (Turgeon and Debuchy 2007).

Un élément essentiel de cette reconnaissance intercellulaire est le système phéromones/récepteurs fonctionnant d'une manière très similaire au système de *S. cerevisiae*. Chez *P. anserina* les gènes qui codent les précurseurs des phéromones étaient les seules cibles des facteurs de transcription MAT identifiées par les études génétiques ce qu'a confirmé l'analyse transcriptomique présentée dans ce manuscrit (Coppin et al. 2005; Bidard et al. 2011).

Le gène *SMR1* n'est pas nécessaire pour la fécondation mais pour des étapes ultérieures (Debuchy et al. 1993).

2.2.3.2.2 La fonction des gènes des types sexuels pendant la reconnaissance internucléaire

Après la fécondation, la transition d'un syncytium plurinucléé aux cellules dicaryotiques nécessite une étape de reconnaissance internucléaire entre un noyau *mat+* et un noyau *mat-* (Figure 3). Cette reconnaissance internucléaire est sous le contrôle des trois gènes *mat* : *FMR1*, *SMR2* et *FPR1* (Zickler et al. 1995), et suit un schéma similaire à celui de la reconnaissance intercellulaire. Le modèle proposé suppose que, dans la cellule plurinucléée, chaque noyau *mat+* et *mat-* porte une identité propre basée sur la présence des protéines spécifiques contrôlées par les gènes *mat* (Figure 6). Dans un noyau *mat-*, *FMR1* et *SMR2* activent des gènes nécessaires à la reconnaissance du noyau *mat-* par le noyau *mat+* et réciproquement, dans un noyau *mat+*, *FPR1* active les gènes nécessaires à la reconnaissance du noyau *mat+* par le noyau *mat-*. Cette activation est accompagnée d'une répression des gènes du système de reconnaissance opposé (Arnaise et al. 2001). Un modèle dit d'identité nucléaire a été présenté à partir de ces données (Figure 6). Ce modèle propose que les hyphes ascogènes ne puissent se développer que lorsqu'elles reçoivent deux noyaux exprimant des identités nucléaires différentes. Dans certains cas, les hyphes ascogènes peuvent se développer si elles reçoivent un seul noyau exprimant les deux identités. C'est ce qu'on observe dans des croisements impliquant des mutants *mat*.

Le croisement des mutants *fmr1⁻*, *smr2⁻* ou *fpr1⁻* chez lesquels le domaine nécessaire à la fécondation n'est pas altéré, avec une souche sauvage de référence, donne des crochets uninucléés et une descendance faible en grande partie uniparentale. La majorité des asques issus de ces croisements sont constitués uniquement de spores du phénotype de la souche mutante (Arnaise et al. 2001).

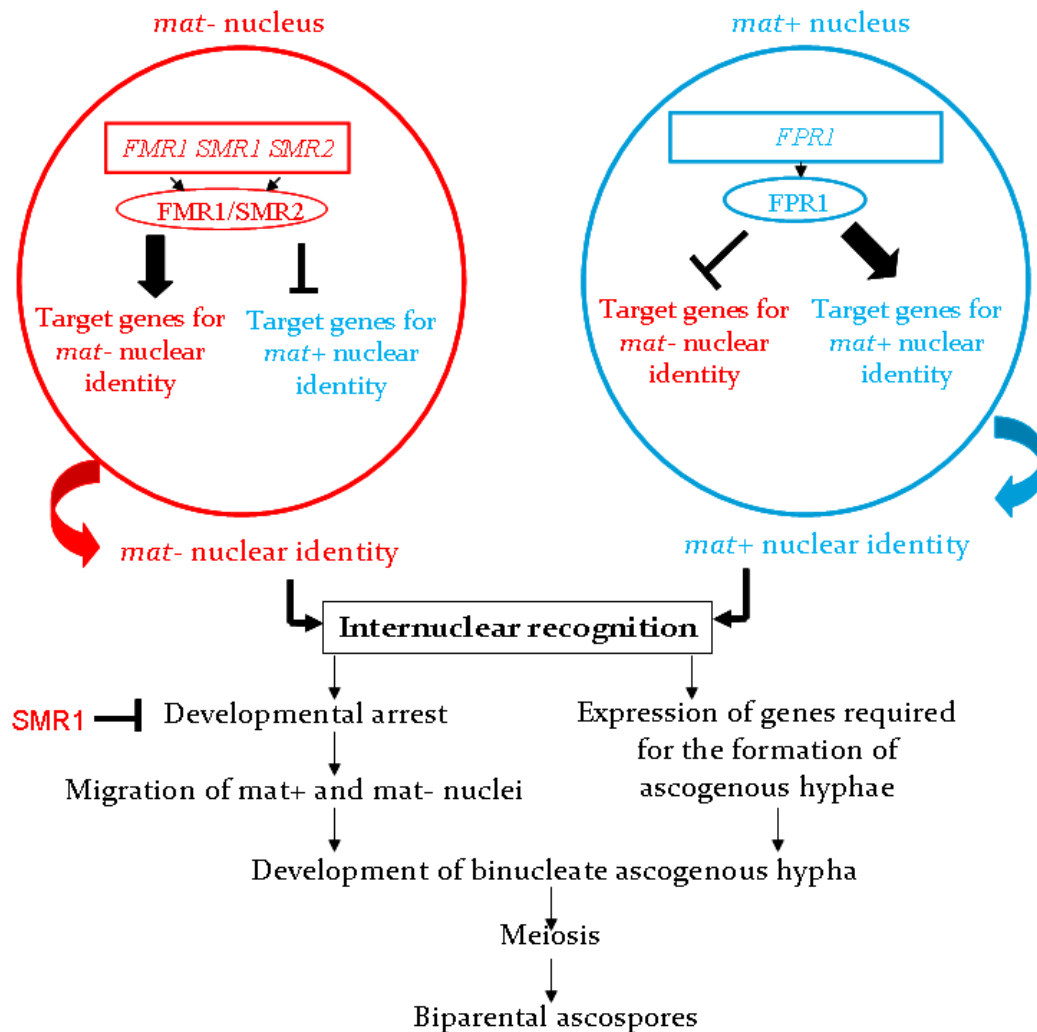


Figure 6: Régulation de la reconnaissance internucléaire par les gènes *mat*. L'activation est représentée par une flèche alors que la répression est représentée par une barre (Turgeon and Debuchy 2007)

Trois types d'asques contenant des spores d'origine uniparentale sont observés : i) des asques contenant quatre grosses ascospores, ou trois grosses ascospores et deux petites spores, ii) des asques avec quatre petites spores et iii) des asques avec des ascospores anormales en taille et en nombre (Zickler et al. 1995). La première catégorie peut s'expliquer par caryogamie entre deux noyaux haploïdes identiques (Figure 7a). Les asques de la deuxième catégorie résultent vraisemblablement d'une méiose haploïde où la première division de méiose est absente (Figure 7b). La ségrégation anormale des chromosomes pendant la méiose haploïde cause la formation des ascospores isolées et anormales de la troisième catégorie (Figure 7c). Ces résultats démontrent que le noyau *mat* mutant se comporte en noyau égoïste, ignorant son partenaire sauvage du type sexuel opposé.

Les mêmes résultats sont observés lors de l'autofécondation des souches mutantes *fmr1*⁻, *smr2*⁻ ou *fpr1*⁻ de même que dans les croisements de ces souches avec des souches de même type sexuel. Ces résultats suggèrent que le noyau mutant exprime les deux identités *mat+* et *mat-*. Cette autorecognition est suffisante pour promouvoir un développement similaire à celui d'un couple de noyaux compatibles *mat+* et *mat-*.

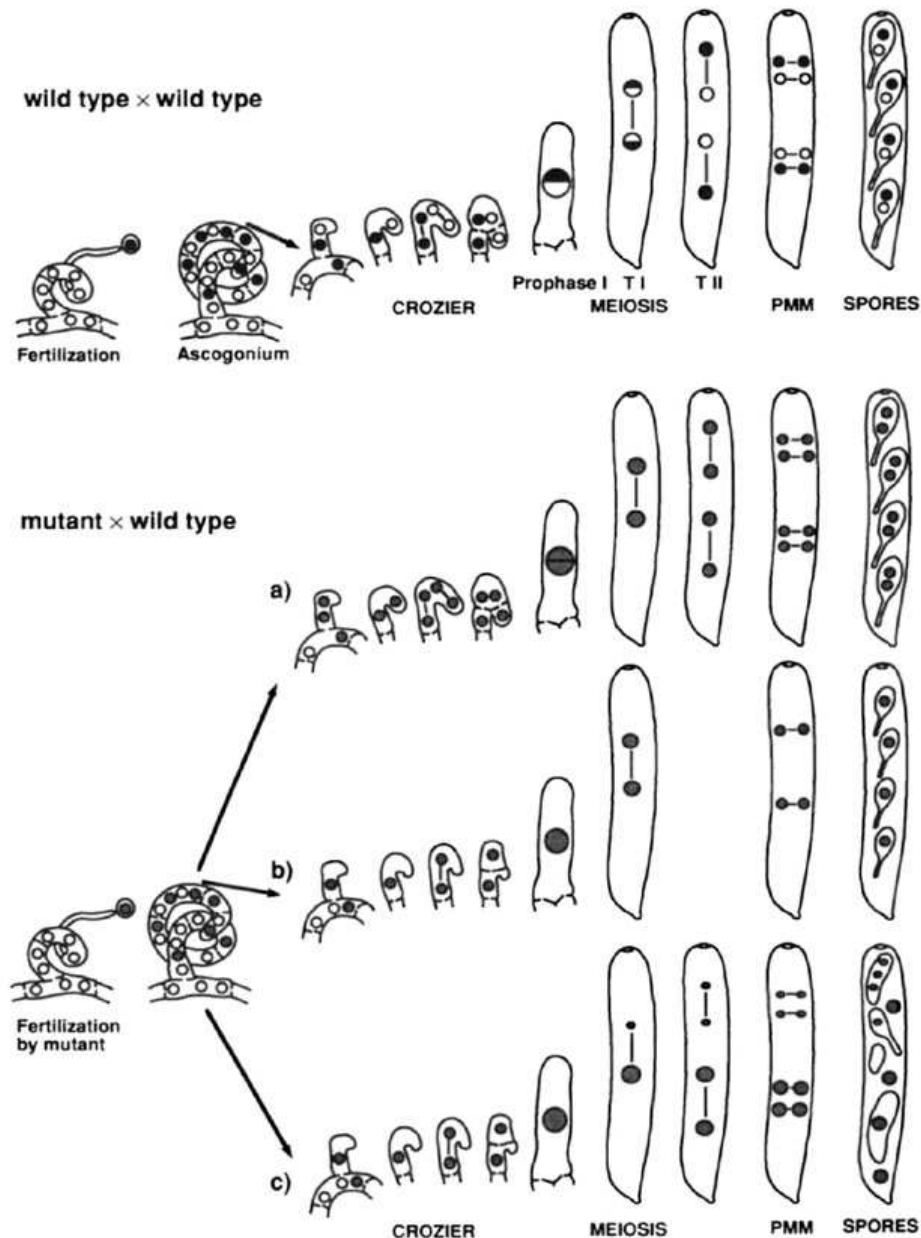


Figure 7: Comparaison entre le croisement sauvage et les croisements mutant \times sauvage à partir du stade de fécondation jusqu'à la formation des spores. Dans le croisement sauvage \times mutant (a), le développement des asques est similaire au sauvage sauf que toutes les quatre ascospores contiennent deux noyaux mutant. Dans le croisement sauvage \times mutant (b), la première division de méiose est absente et le noyau haploïde effectue directement une seconde division de méiose et une division post-méiotique. Dans le croisement sauvage \times mutant (c), la ségrégation anormale des chromosomes pendant la méiose haploïde cause la formation des spores isolées et anormales. Les noyaux des deux types sexuels sont représentés par des couleurs différentes (Zickler et al. 1995).

Le modèle d'identité nucléaire implique que les facteurs de transcription MAT retournent spécifiquement dans le noyau qui contient leur gène, afin de lui conférer son identité. Des analyses de complémentation nucléaire suggèrent effectivement que les produits des gènes *mat* retournent dans le noyau qui les code. Ces expériences comparent le résultat d'un croisement entre un mutant *mat* et une souche testrice sauvage, lorsque le gène complémentant la mutation est présent dans le parent sauvage, et non dans le parent muté. Si un phénotype sauvage est restauré, c'est que le produit du gène diffuse dans le noyau muté ; si

le croisement est identique à celui du mutant avec la souche testrice sauvage, c'est que le produit du gène ne diffuse pas dans le noyau muté. Ainsi ; le croisement d'une souche mutante *smr2⁻* par une souche sauvage *mat⁺* contenant la copie sauvage de *SMR2* donne une descendance identique à celle d'un croisement entre une souche *smr2⁻* par le sauvage (Arnaise et al. 1997). Ces résultats suggèrent que le produit de *SMR2* ne diffuse pas du noyau *mat⁺* au noyau *mat⁻*. Dans le cas de *FMR1* le test de complémentation nucléaire montre que le produit de ce gène peut diffuser vers les noyaux adjacents (Arnaise et al. 1997). Le système de double hybride chez la levure a révélé que les protéines *SMR2* et *FMR1* interagissent physiquement. Ces données suggèrent que *SMR2* séquestre *FMR1* pour empêcher sa diffusion vers les noyaux voisins et l'aider à retourner vers le noyau *mat⁻*. Dans le cas de *FPR1*, la diffusion de la protéine d'un noyau *mat⁺* au noyau *mat⁻* n'a pas été testée (Arnaise et al. 1997).

L'expression noyau-restreinte de *FPR1*, *FMR1* et *SMR2* permet de comprendre comment chaque type de noyau acquiert une identité nucléaire. Mais le mécanisme de la reconnaissance internucléaire ainsi que les gènes cible impliqués dans cette étape restent inconnus. Les mutants *fmr1⁻* et *fpr1⁻* affectent de la même façon la fécondation et la reconnaissance internucléaire. Cette corrélation pouvait suggérer que les deux mécanismes impliquaient des systèmes identiques. L'analyse de mutants délétés des gènes de phéromones a permis de montrer que les systèmes phéromones/récepteurs nécessaires à la reconnaissance intercellulaire pendant la fécondation n'intervenaient pas dans la reconnaissance internucléaire (Coppin et al. 2005).

Metzenberg et Glass ont proposé chez *N. crassa* un autre modèle dit modèle de ségrégation aléatoire (Metzenberg and Glass 1990). Dans ce modèle, la cellule diploïde de *N. crassa* contient un nouveau régulateur produit par la coopération des idiomorphes *a* et *A*. Ce nouvel hétérodimère est responsable de l'expression des gènes requis pour la formation des hyphes ascogènes. L'interaction entre *MAT a-1* (homologue de *FPR1*) et *MAT A-1* (homologue de *FMR1*) a été prouvée par double hybride mais aucune mutation de cette interaction n'a affecté le développement sexué chez *N. crassa*. En ce qui concerne les protéines *MAT* de *P. anserina*, le système double-hybride a révélé l'homodimérisation de *FPR1* mais aucune interaction entre *FPR1* et les autres facteurs de transcription *mat⁻*. Il faut cependant noter que ce système a échoué dans la détection de nombreuses interactions notamment entre *a1* et *a2* chez *S. cerevisiae* (Turgeon and Debuchy 2007). Selon ce modèle de ségrégation aléatoire, l'hétérodimère *FMR1/SMR2* dans le noyau *mat⁻* et l'homodimère *FPR1* dans le noyau *mat⁺* répriment les gènes cibles impliqués dans la formation des hyphes ascogènes. Le couple des noyaux *mat⁺/mat⁻* produit un nouveau régulateur par la coopération des idiomorphes *mat⁺* et *mat⁻*. Ce produit activerait l'expression des gènes impliqués dans la

formation des hyphes ascogènes. Dans ce modèle, la formation des asques uniparentaux *mat-* serait due à la perte de la répression des gènes impliqués dans la formation des hyphes ascogènes en l'absence de l'hétérodimère FMR1/SMR2. L'expression de ces gènes permettrait l'engagement des noyaux mutants dans l'hyphe ascogène ce qui engendre des spores uniparentales. Un raisonnement similaire peut être appliqué au noyau *mat+* mutant (Turgeon and Debuchy 2007).

2.2.3.2.3 La fonction de SMR1

Le gène SMR1 a été trouvé uniquement chez les Sordariomycètes. L'analyse de ce gène chez *P. anserina* fait apparaître une fonction biologique essentielle au cours du cycle sexué. Dans le croisement entre une souche mutante *smr1⁻* et une souche sauvage, on n'observe aucune descendance ; le développement des périthèces est arrêté avant la formation des hyphes ascogènes (Arnaise et al. 1997). Un phénotype similaire a été observé chez *Gibberella zeae* (Marie Dufresne, communication personnelle) et ce rôle essentiel a été confirmé chez *Sordaria macrospora* (Klix et al. 2010). Le gène SMR1 est présent dans l'idiomorphe *mat-*, mais une copie présente dans le noyau *mat+* complémente parfaitement une perte de fonction de SMR1 dans le noyau *mat-*. Ce gène peut même être présent dans les deux parents sans affecter la descendance (Arnaise et al. 1997). Donc ce gène, bien que présent dans un seul des deux idiomorphes *mat*, n'est pas un gène participant à la détermination d'une identité sexuelle ou nucléaire.

Si la fonction moléculaire de SMR1 est encore inconnue, sa fonction biologique a été précisée. Il interviendrait après la reconnaissance internucléaire, pour lever un arrêt du développement. En effet, pour mimer l'étape de reconnaissance internucléaire, les trois gènes impliqués dans cette reconnaissance (*FPRI*, *FMR1* et *SMR2*) ont été regroupés dans un même noyau. Leur expression simultanée entraîne un phénotype léthal. Ce phénotype léthal peut être surmonté par l'expression de SMR1 (Coppin and Debuchy 2000), ce qui suggère que l'expression de *FMR1*, *SMR2* et *FPRI* pendant le développement de la fructification déclenche un arrêt du développement qui est surmonté par l'intervention de SMR1 (Coppin and Debuchy 2000; Turgeon and Debuchy 2007). Cet arrêt pourrait être nécessaire pour la synchronisation des noyaux *mat+* et *mat-* avant l'entrée dans l'hyphe ascogène, où ils subissent des mitoses synchrones (Coppin and Debuchy 2000).

2.2.4 Le système phéromones/récepteurs chez *Podospora anserina*

Chez les champignons, les phéromones sont des substances secrétées par les gamètes mâles pour attirer les organes femelles de type sexuel opposé. Elles se fixent sur des

récepteurs spécifiques existant à la surface du trichogyne ce qui déclenchent une cascade de transduction du signal (Coppin et al. 2005).

Chez *P. anserina*, le génome contient deux gènes codant les précurseurs de phéromones : *mfm* et *mfp*. Le gène *mfp* (mating facteur plus) code un polypeptide de 24 acides aminés qui se termine par un motif CAAX similaire au facteur hydrophobe a de *S. cerevisiae*, alors que le gène *mfm* (mating facteur minus) code un polypeptide de 221 acides aminés qui contient 13 résidus répétés similaires aux facteurs hydrophile α de *S. cerevisiae* (Coppin et al. 2005). Pour être actifs, ces précurseurs des phéromones doivent subir plusieurs étapes de maturation similaires aux étapes observées chez *S. cerevisiae* et *S. pombe* (Davey et al. 1998).

La délétion du gène *mfm* affecte exclusivement la fertilité mâle dans la souche *mat-* sans altérer la fertilité femelle. D'une façon similaire, la délétion du gène *mfp* affecte uniquement la fertilité mâle dans la souche *mat+* (Coppin et al. 2005). Cette stérilité mâle peut être surmontée par l'introduction d'une copie transgénique du gène sauvage *mfm* ou *mfp* dans les souches *mat-* et *mat+* respectivement (Coppin et al. 2005). Par contre, la fertilité mâle n'est pas affectée chez les mutants *mat+ Δmfm* et *mat- Δmfp*. La délétion des phéromones n'affecte pas la croissance, la morphologie du mycélium, la différenciation et la fonction des organes femelles (Coppin et al. 2005). En conclusion, la phéromone *mfp* est exprimée spécifiquement dans la souche *mat+*, tandis que la phéromone *mfm* est exprimée spécifiquement dans la souche *mat-* (Figure 8). L'expression des phéromones est contrôlée par les facteurs de transcription FPR1 (*mat+*) et FMR1 (*mat-*). Chez *mat+*, FPR1 active l'expression de *mfp* et réprime l'expression de *mfm*. Inversement chez *mat-*, FMR1 active l'expression de *mfm* et réprime l'expression de *mfp* (Coppin et al. 2005).

P. anserina possède deux récepteurs de phéromones *PaPRE1* et *PaPRE2* orthologues de *Ncpre-1* et *Ncpre-2* respectivement de *N. crassa*. *NcPRE-1* est le récepteur de la phéromone "NcMFA-1" sécrétée par la souche *mat a*, alors que *NcPRE-2* est le récepteur de la phéromone "NcCCG-4" sécrétée par la souche *mat A* (Poggeler and Kuck 2001). La transcription des gènes *Ncpre-1* et *Ncpre-2* n'est pas spécifique des souches A et a, mais la délétion de *Ncpre1* dans la souche a (ou la délétion de *Ncpre2* dans la souche A) n'ont pas donné d'altérations détectables du phénotype. La délétion de *Ncpre1* cause la stérilité femelle de la souche *mat A* (Kim and Borkovich 2004, 2006).

Chez *P. anserina* on suppose que la régulation des récepteurs est similaire à celle des phéromones. Chez *mat+*, FPR1 activerait l'expression de *PRE2*. Chez *mat-*, FMR1 activerait l'expression de *PRE1*. *PRE1* et *PRE2* sont les récepteurs de *mfp* et *mfm* respectivement (Figure 8).

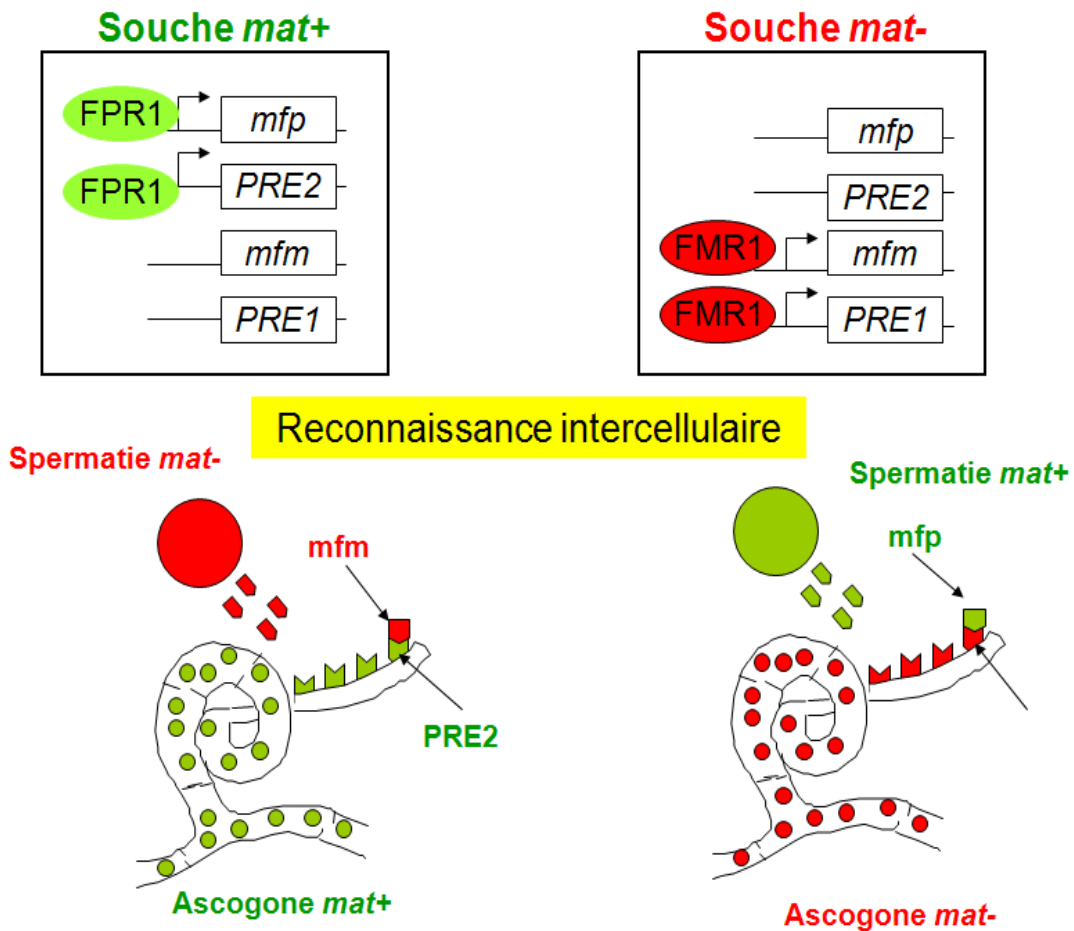


Figure 8 : Régulation de l'expression des phéromones et de leurs récepteurs par les facteurs de transcription MAT. Les gènes *mfm* et *mfp* codent les phéromones, Les gènes *PRE1* et *PRE2* codent les récepteurs.

2.3 La reproduction sexuée chez les levures

2.3.1 Introduction

Les levures sont des champignons qui appartiennent à la famille des ascomycètes. Elles ont la particularité d'être des organismes eucaryotes et unicellulaires, dont les génomes compacts et de petite taille en font des organismes modèles pour l'étude simplifiée des eucaryotes (Forsburg 2005). Il existe de nombreuses espèces de levures différentes, dont les plus connues sont la levure bourgeonnante *S. cerevisiae* (Saccharomycotina) et la levure fissipare *S. pombe* (Taphrinomycotina) (Figure 1). Chacune de ces deux espèces a une évolution bien distincte de l'autre. Elles ont divergé à partir d'un ancêtre commun depuis environ 1000 millions d'années (Forsburg 2005). Ce sont les deux premières levures pour lesquelles les séquences complètes des génomes ont été obtenues, respectivement en 1996 (Goffeau et al. 1996) et 2002 (Wood et al. 2002).

La levure *S. cerevisiae* plus connue sous le nom de « levure du boulanger » est adoptée comme système modèle au laboratoire en 1930 (Hall and Linder 1993). Cette espèce est largement étudiée en biologie cellulaire et moléculaire. En fait, cet organisme représente un système expérimental qui procure de nombreux avantages, entre autre par la facilité à la

manipuler, la possibilité de faire une analyse génétique et ses hautes capacités de recombinaison homologue. La cellule de *S. cerevisiae* est sphéroïde et sa taille dépend de la souche, de son âge (elle se divise par bourgeonnement) et de son degré de ploïdie. Le génome nucléaire de la levure *S. cerevisiae* représente environ 12 méga bases paires (Mbp) réparties sur 16 chromosomes. Le génome contient 6607 gènes codant 4964 ORF vérifiées, 858 ORF non caractérisées et 785 ORF douteuses. Le génome est disponible sur le site : <http://www.yeastgenome.org/>.

La levure *S. pombe* diffère de la levure *S. cerevisiae*, non seulement par la forme des cellules (allongée au lieu de sphériques) mais aussi par de nombreuses autres caractéristiques biologiques. En particulier, elle est appelée levure fissipare car contrairement à *S. cerevisiae*, elle ne se multiplie pas par bourgeonnement mais par fission transversale. Comme *S. cerevisiae* et pour les mêmes raisons, la levure *S. pombe* est devenue un outil scientifique de premier plan. Réparti en seulement trois chromosomes, son génome nucléaire de 12,5 Mbp contient 5122 gènes codant. Le génome est disponible sur le site : <http://old.genedb.org/genedb/pombe/>.

Le cycle de vie de *S. cerevisiae* et *S. pombe*, et en particulier, leurs types sexuels ont fourni un terrain riche pour l'étude des mécanismes moléculaires de régulation des gènes et le fonctionnement des systèmes de signalisation impliqués dans le développement sexué. Comme d'autres eucaryotes, *S. cerevisiae* et *S. pombe* ont une véritable phase sexuelle (Figure 9), les cellules haploïdes de types sexuels, désignées *MATA* ou *MAT α* chez *S. cerevisiae*, et *matP* ou *matM* chez *S. pombe*, sécrètent des phéromones qui se lient spécifiquement aux récepteurs compatibles sur la surface des cellules de types sexuels opposés. La détection réciproque des phéromones induit un phénomène complexe de transduction de signal qui aboutit à la fusion des deux cellules et à la formation d'un zygote. Cette cellule diploïde peut se propager par des divisions mitotiques, ou entrer en méiose en réponse à une carence nutritionnelle pour former quatre spores haploïdes. Ces spores germent et réintègrent le cycle mitotique quand les conditions sont favorables (Casselton 2002).

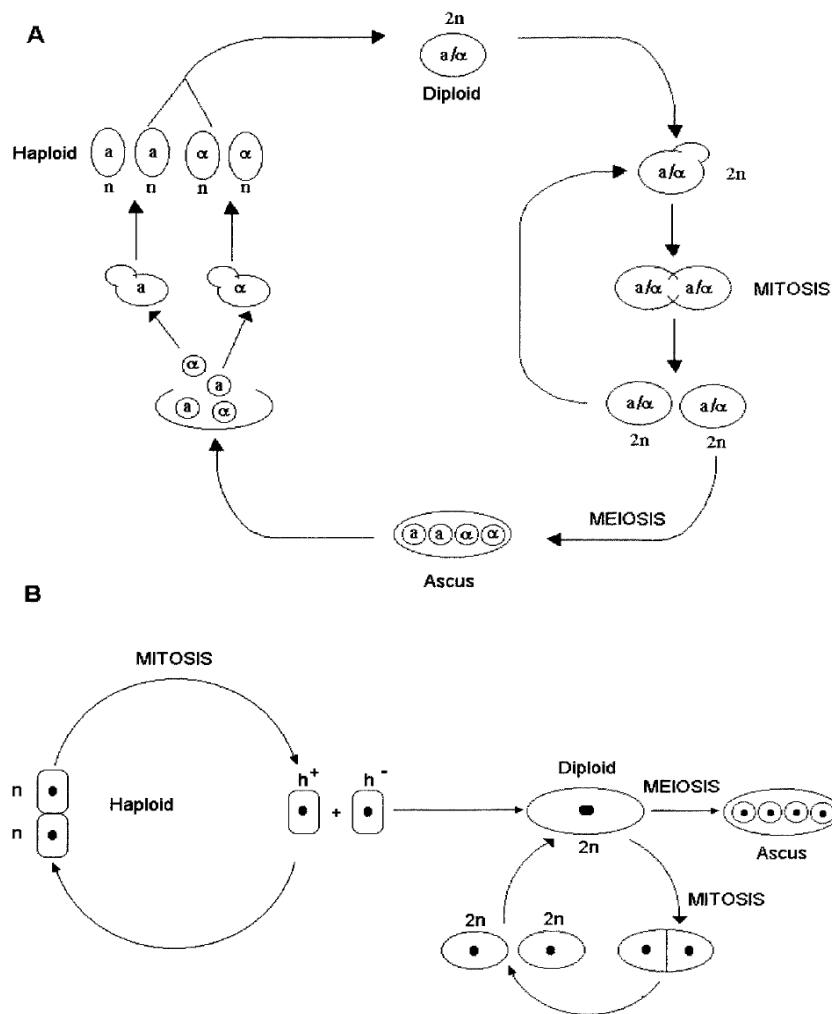


Figure 9: Cycle de vie de (A) *Saccharomyces cerevisiae* et de (B) *Schizosaccharomyces pombe*. (Perego et al. 2000)

2.3.2 La reproduction chez *Saccharomyces cerevisiae*

S. cerevisiae existe en trois types de cellules spécialisées qui jouent un rôle distinct et important dans le cycle de vie (Herskowitz 1988). Les types sexuels de ces cellules sont déterminés par un seul locus appelé *MAT*. Les cellules haploïdes "a" et "α" sont déterminées par *MATa* et *MATα* respectivement, tandis que la cellule diploïde a/α est déterminée par les deux types sexuels *MATa/MATα* (Herskowitz 1988; Casselton 2002; Souza et al. 2003).

Les idiomorphes *MATa* et *MATα* contiennent chacun deux gènes avec des transcrits divergents : *a1*, *a2* et *α1*, *α2* respectivement (Figure 10) ; ces gènes codent des facteurs de transcription qui déterminent l'identité cellulaire (Souza et al. 2003), sauf pour *a2* dont la fonction est inconnue (Herskowitz et al. 1992). Les protéines *a1* et *α1* contiennent des motifs de fixation à l'ADN de type homéodomaine (Shepherd et al. 1984). La protéine *α1* possède un domaine de fixation à l'ADN nommé domaine *α1* caractérisé récemment comme un *MATα_HMG* (Martin et al. 2010), cette protéine est un activateur transcriptionnel (Sprague et al. 1983).

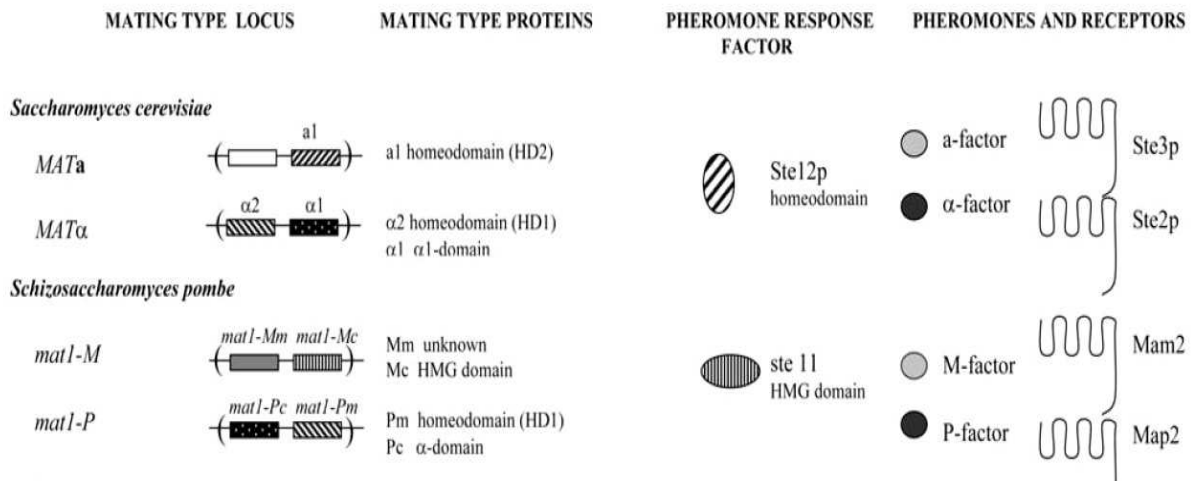


Figure 10: Le locus *mat* et les protéines essentielles dans la conjugaison chez *S. cerevisiae* et *S. pombe*. (Casselton 2002)

Dans des cellules haploïde "a", le facteur de transcription Mcm1p à domaine MADS se fixe sur son site en amont des gènes spécifiques "a" et active leur expression (Ammerer 1990; Casselton 2002) (Figure 11). Ces gènes codent les phéromones a (Mfa1p et Mfa2p), le transporteur de la phéromone a (Ste6p), le récepteur de la phéromone α (Ste2p), la protéase qui dégrade la phéromone α (Bar1p) et une agglutinine (Aga2p) (Galgoczy et al. 2004). Dans les cellules haploïdes "α", Mcm1p s'associe à la protéine α1 pour activer les gènes spécifiques "α", ces gènes codent les phéromones α (Mfa 1p et Mfa 2p), le récepteur de la phéromone a (Ste3p) et une agglutinine (Aga1p) (Galgoczy et al. 2004); Mcm1p s'associe aussi à α2 pour réprimer les gènes spécifiques a (Casselton 2002) (Figure 11).

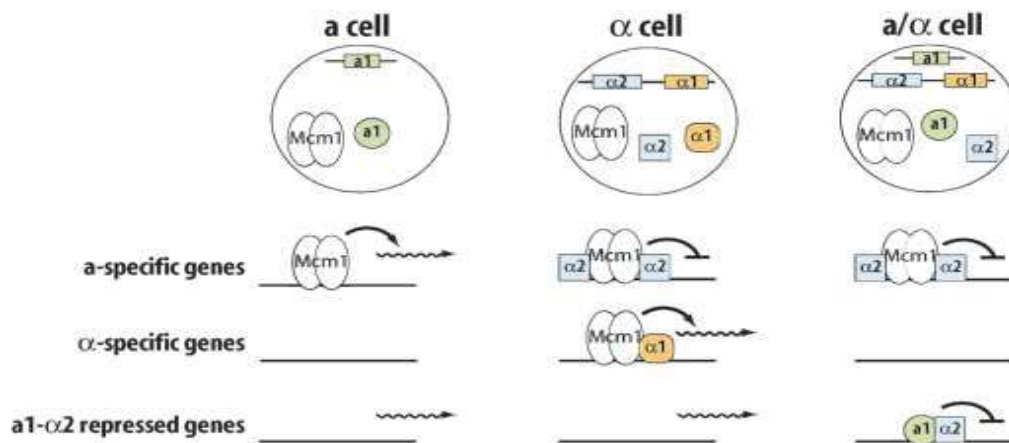


Figure 11: Régulation des types sexuels chez *S. cerevisiae*. (Galgoczy, 2004)

Les phéromones et les récepteurs de phéromones sont spécifiques de chaque type de cellules (Souza et al. 2003). Les cellules "a" produisent le facteur a qui est un peptide de 12 acides aminés avec un motif CAAX dans sa région C-terminal (signal de carboxyméthylation) (Chen et al. 1997). La phéromone mature est sécrétée par Ste6p (ATP-binding cassette transporter) dans le milieu pour se fixer à son récepteur exprimé par les cellules "α". Ce récepteur est codé par le gène *STE3* (Figure 10) (Casselton 2002). Les cellules "α" produisent

le facteur α composé de 13 acides aminés et dérivé de la maturation d'un précurseur à plusieurs copies en tandem de la phéromone α . Le facteur α sécrété dans le milieu par la voie de sécrétion classique reconnaît le récepteur Ste2p exprimé par les cellules "a" (Figure 10) (Casselton 2002).

Une fois reconnues par leurs récepteurs respectifs, les phéromones activent le processus de transduction du signal composé de la protéine G, une cascade de protéines kinase (Ste20p, Ste11p, Ste7p, Fus3p et Kss1p) et un facteur de transcription (Ste12p) (Pi et al. 1997; Casselton 2002). La protéine Ste12p se fixe spécifiquement au motif PRE (pheromone response element) TGAAACA (James R. Broach et al. 1991-1997) présent en amont de nombreux gènes induits lors de la réponse aux phéromones a et α (Errede and Ammerer 1989). Parmi ces gènes, on trouve *FUS1* et *FAR1* impliqués dans la fusion cellulaire et l'arrêt en G1 (Pi et al. 1997). Ste12p se fixe aussi avec Mcm1p aux promoteurs des gènes spécifiques a et des gènes spécifiques α afin d'augmenter l'expression de ces gènes et multiplier le signal phéromone (Pi et al. 1997).

Les cellules diploïdes n'expriment pas les gènes nécessaires à la conjugaison. Elles forment un hétérodimère composé des protéines $\alpha 1$ et $\alpha 2$. Cet hétérodimère $\alpha 1/\alpha 2$ réprime l'expression du gène $\alpha 1$ et les gènes spécifiques des cellules haploïdes (Casselton 2002) (Figure 11), dans les cellules diploïdes, 19 gènes contrôlés par $\alpha 1/\alpha 2$ ont été identifiés; ils sont impliqués dans différents processus biologiques (Galgoczy et al. 2004). Les gènes des types sexuels régulent aussi la méiose par une répression transcriptionnelle. Dans les cellules diploïdes, l'hétérodimère $\alpha 1/\alpha 2$ réprime la transcription de *RME1* qui code le répresseur de la méiose, ce qui permet aux cellules diploïdes d'entrer en méiose (Casselton 2002). En fait, Rme1p réprime la transcription de *IME1* impliqué dans l'activation des gènes nécessaires à la méiose et à la sporulation (James R. Broach et al. 1991-1997).

2.3.3 La reproduction chez *Schizosaccharomyces pombe*

Comme *S. cerevisiae*, les cellules haploïdes de *S. pombe* existent sous forme de deux types sexuels, soit *M* (moins) soit *P* (plus). Ces deux types sexuels sont déterminés par le locus *mat*, *mat1-P* chez les cellules *P* et *mat1-M* chez les cellules *M*. Chaque idiomorphe contient deux gènes, *mat1-Mc* et *mat1-Mm* dans le locus *mat1-M* et *mat1-Pm* et *mat1-Pc* dans le locus *mat1-P*. Les gènes *mat1-Mc* et *mat1-Pc* sont nécessaires pour la conjugaison, alors que *mat1-Mm* et *mat1-Pm* (appelé aussi *mat1-Mi* et *mat1-Pi* respectivement) sont essentiels pour la méiose (Kelly et al. 1988; Bolker and Kahmann 1993). Le locus *mat1-P* ressemble au locus *Mat α* de *S. cerevisiae*, *mat1-Pm* code une protéine à homéodomaine et *mat1-Pc* code une protéine apparentée à la famille HMG, mais dont la structure tridimensionnelles semble différente (Casselton 2002; Martin et al. 2010). Dans le locus *mat1-M*, *mat1-Mc* code une

protéine à domaine HMG, tandis que *mat1-Mm* code une protéine de classe inconnu (Casselton 2002) (Figure 10).

Les carences nutritionnelles et le signal phéromone induit l'expression du facteur de transcription Ste11p qui induit à son tour l'expression des gènes *mat-Mc* et *mat-Pc* (Sugimoto et al. 1991). Dans une cellule *M*, Mat-Mc en collaboration avec Ste11p active l'expression des gènes spécifiques *M* (Kjaerulff et al. 1997). Dans la cellule *P*, Mat-Pc avec Map1p (facteur de transcription de la famille des MADS-box) vont activer les gènes spécifiques *P* (Mata and Bähler 2006). Les autres gènes du locus *mat*, *mat1-Pm* et *mat1-Mm* entrent en action après la caryogamie et activent l'expression des gènes impliqués dans la méiose notamment *mei3* (Willer et al. 1995; Casselton 2002). Dans les cellules diploïdes *P/M* et sous les conditions de carences nutritionnelles, les quatre gènes des types sexuels *mat* sont exprimés, ce qui permet à ces cellules d'entrer en méiose. En revanche, *mat1-Pm* et *mat1-Mm* ne sont pas exprimés chez les cellules diploïdes au stade végétatif (Kelly et al. 1988; Willer et al. 1995). Les mécanismes de régulation de *ste11* et les gènes spécifiques *M* et *P* seront décrits dans le paragraphe 3.2.3.4 de l'introduction des protéines HMG.

Les cellules *P* et *M* produisent des phéromones spécifiques (les facteurs P et M respectivement) ainsi que les récepteurs des phéromones sécrétées par le type sexuel opposé (Mata and Bähler 2006). Le facteur M est issu de la farnesylation et la carboxyméthylation d'un précurseur codé par les gènes *mfm1*, *mfm2* (Davey 1992) et *mfm3* (Mata and Bähler 2006). Ce facteur M reconnaît un récepteur spécifique codé par *map3* dans les cellules *P* (Tanaka et al. 1993). Le facteur P codé par le gène *map2* est un peptides non modifié de 23 acides aminés (Imai and Yamamoto. 1994). Il se fixe sur son récepteur spécifique Mam2p exprimé par les cellules *M* (Kitamura and Shimoda 1991). Les phéromones induisent 163 gènes chez *S. pombe* dont la majorité est induite chez les deux types sexuels *P* et *M*. L'annotation de ces gènes montre qu'ils sont impliqués dans la différenciation sexuelle, l'induction de la conjugaison et la méiose (Xue-Franzen et al. 2006) (Figure 12).

Quand les cellules haploïdes sont exposées aux phéromones, elles répondent par l'arrêt des cellules en phase G1. L'arrêt des cellules *P* et *M* en G1 est suivie par l'élongation du tube de conjugaison vers la cellule de type sexuel opposé. Le contact cellule-cellule conduit à l'agglutination, la caryogamie et la formation du zygote qui entre en méiose et finalement sporule (Xue-Franzen et al. 2006).

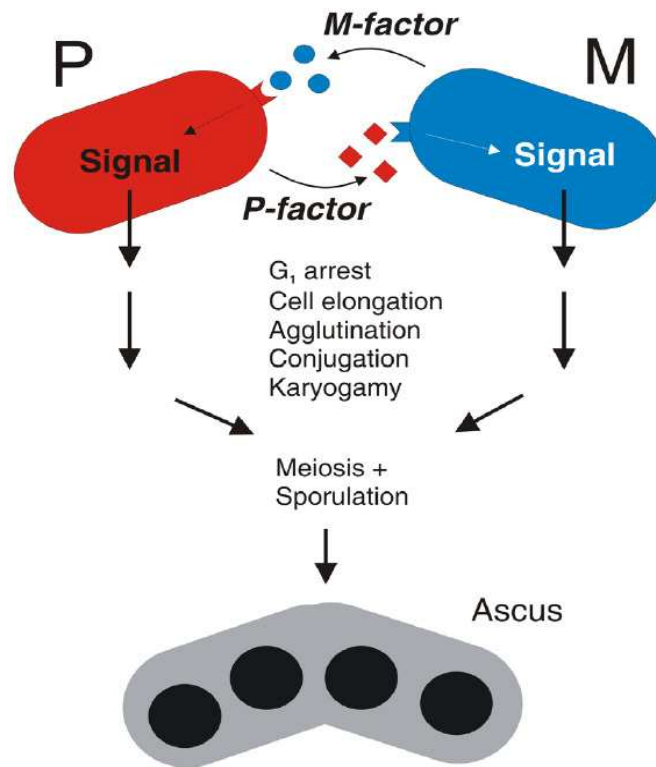


Figure 12 : La transduction du signal phéromone chez *S. pombe*. Les cellules *M* et *P* sécrètent les deux phéromones, M (cercles) et P (carré) respectivement qui se fixent sur leurs récepteurs spécifiques présentés à la surface des cellules de types sexuels opposés. Cela cause l'activation de la voie de signalisation qui prépare la cellule à la conjugaison suivie de la méiose et la sporulation.

3 Les gènes à HMG-box

3.1 Introduction

Dans les cellules eucaryotes, l'ADN nucléaire est étroitement associé à un ensemble de protéines, dont les plus connues sont les histones. La structure qui résulte de cette association est nommée la chromatine. L'unité fondamentale de la chromatine, le nucléosome, composé d'ADN et d'histones, constitue le premier niveau de compaction de l'ADN dans le noyau. Cette structure est ensuite régulièrement répétée pour former le nucléofilament qui peut, lui-même, adopter des niveaux d'organisation plus compacts. La compression de l'ADN en chromosomes est indispensable à l'architecture de la cellule et à la régulation transcriptionnelle des gènes. La chromatine, de par sa structure condensée, représente une barrière à l'accessibilité de l'ADN. Elle est en perpétuel équilibre entre l'état relâché et l'état condensé suivant les effets antagonistes de multiples complexes protéiques et il ne fait guère de doute que l'état condensé est défavorable à la transcription. Le remodelage de la chromatine est donc une étape cruciale dans la régulation de la transcription (Workman and Kingston 1998).

Des travaux initiés en 1973 par Goodwin, Sanders et Johns ont mis en évidence une autre catégorie de protéines chromatiniennes abondantes, distinctes des histones (Goodwin and Johns 1973). Ces protéines sont appelées HMG (high mobility group). La famille des protéines HMG constitue le principal groupe de protéines non-histones associées à la chromatine (Agresti and Bianchi 2003). Les protéines HMG sont ubiquitaires chez les Eucaryotes, mais absentes chez les Eubactéries et les Archae (Marechal, 2003). Officiellement, l'acronyme HMG désigne l'ensemble hétérogène des protéines que l'on peut extraire en condition acide à partir de la chromatine, et qui manifestent une forte mobilité sur gel d'électrophorèse (high mobility group). La nomenclature des protéines HMG nucléaires a été révisée pour définir trois superfamilles classées en fonction de leur motif fonctionnel de liaison à l'ADN (Bustin 2001) (Tableau 1).

- Les protéines HMGA, anciennement appelées HMG-I(Y), contiennent un motif structural AT-hook de neuf acides aminés qui se lie aux régions riches en A+T.

- Les protéines HMGN anciennement HMG-14 et 17 sont caractérisées par un «nucleosomal-binding domain» qui se lie entre les spirales de l'ADN et l'octamère d'histones.

- Les protéines HMGB, anciennement HMG-1 et HMG-2, sont caractérisées par un domaine HMG-box de 80 acides aminés qui lie l'ADN dans le petit sillon.

HMG motif protein	Functional motif	HMG proteins		
		Root symbol	New name (canonical HMGs)	Old name (canonical HMGs)
HMG-box proteins	HMG-box	HMGB	HMGB1,2,..n	HMG-1/HMG-2
NBD proteins	NBD	HMGN	HMGN1,2,..n	HMG-14/HMG-17
ATH proteins	ATH	HMGA	HMGA1,2,..n	HMG-I/HMG-Y/ HMG-C

^aAbbreviations: ATH, AT-hook; HMG, high mobility group; NBD, nucleosome binding domain.

Tableau 1: Nomenclature des protéines HMG.
(Bustin 2001).

Je me suis plus particulièrement intéressée au rôle des protéines HMGB dans le développement sexué des champignons filamenteux et je me limiterai à la présentation de cette famille de protéines HMG. Les protéines HMGB sont caractérisées par la boîte HMG avec un domaine de fixation à l'ADN. Cette superfamille est composée de deux familles distinguées par leur abondance, leurs fonctions et leur spécificité de fixation à l'ADN (Murphy et al. 1999):

- La famille UBF/HMG :

Les protéines de cette famille possèdent plusieurs domaines HMG; elles sont très abondantes dans la cellule et se lient à l'ADN avec une grande affinité pour une variété de structures particulières mais sans spécificité de séquence. Ce groupe de protéines existent dans les trois embranchements des eucaryotes : champignons, plantes et animaux (Soullier et al. 1999). Il comprend entre autres la protéine HMGB1 (Pauken et al. 1994) qui est une protéine ubiquitaire, ainsi que les protéines HMGB2 (Shirakawa and Yoshida 1992) et HMGB3 (Vaccaria et al. 1998). Ces trois protéines possèdent deux domaines HMG notés A et B et un domaine riche en résidus acides dans la partie C-terminale. Le facteur de transcription mitochondrial mtTF-1 (Parisi and Clayton 1991), ainsi que le facteur de transcription nucléaire UBF (Upstream Binding Factor) (Jantzen et al. 1990), appartiennent également à cette classe. Ces deux protéines contiennent respectivement deux et trois domaines HMG.

- La famille MATA/TCF/SOX :

Cette famille rassemble des protéines n'ayant qu'un seul domaine HMG; elles ne sont exprimées que dans certains tissus et se lient à l'ADN en reconnaissant une séquence linéaire spécifique. On peut citer pour exemple des régulateurs transcriptionnels tels que le facteur SRY (Sex determining Region Y) qui est le déterminant du sexe mâle chez les mammifères (Goodfellow and Lovell-Badge 1993), les facteurs activateurs lymphoïdes TCF1 (T Cell Factor-1) et LEF1 (Lymphocyte Enhancer Factor-1) (Giese

et al. 1991). Les sous-familles TCF et SOX ne sont présentes que chez les animaux (Soullier et al. 1999). Les protéines MATA_HMG ne sont présentes que chez les champignons (Soullier et al. 1999). Dans cette sous-famille, on trouve les facteurs de transcription Mat-Mc de *S. pombe* (Kjaerulff et al. 1997) et Mata1p de *S. cerevisiae* (Martin et al. 2010), ainsi que d'autres protéines impliquées dans le développement sexué tel que Ste11p chez *S. pombe* (Mata and Bähler 2006) et FMF-1 chez *N. crassa* (Iyer et al. 2009).

Les facteurs HMGB ont la capacité de courber l'ADN hélicoïdal générant ainsi une conformation qui facilite l'assemblage des complexes nucléoprotéiques, c'est pourquoi ils sont appelés facteurs de transcription architecturaux (Bustin and Reeves 1996; Bustin 1999; Thomas and Travers 2001). En plus de leur rôle dans la transcription, ils sont aujourd'hui très connus comme des éléments essentiels du noyau. Ils participent à de multiples activités cellulaires comme le remodelage de la chromatine et des nucléosomes, les modifications chromosomiques au cours du cycle cellulaire, la réplication de l'ADN, la recombinaison génétique, la réparation de l'ADN et l'apoptose (Agresti and Bianchi 2003; Tetsuya Ueda 2009). Ces protéines servent aussi de chaperons moléculaires. Les protéines HMG sont impliquées dans un certain nombre de pathologies notamment le cancer, l'intégration rétrovirale et la réponse aux dommages tissulaires dans les cas d'inflammation ou nécrose (Stros 2009).

3.2 Structure et fonctions des protéines à domaine HMG

3.2.1 Structure du domaine HMG

Les domaines HMG ont été identifiés premièrement par similarité de séquences entre le facteur de transcription de l'ARN polymérase I (UBF) et d'autres protéines HMG tel que HMG1 et SRY. Depuis, le nombre des protéines à domaine HMG a rapidement augmenté.

Les protéines à domaine HMG sont localisées au niveau du noyau et elles interagissent avec l'ADN. Elles possèdent soit un seul domaine HMG soit deux ou plusieurs. La majorité des protéines HMG particulièrement chez les vertébrés possèdent deux domaines HMG en tandem appelés boîtes A et B. Les plantes, les champignons et les eucaryotes unicellulaires possèdent rarement des protéines HMG avec des boîtes HMG en tandem.

Le domaine HMG est une région d'environ 80 acides aminés avec une forte abondance de résidus aromatiques et basiques. Les études structurales de différents domaines HMG montrent une conservation de la structure tertiaire en forme de L. Le domaine comporte trois hélices α comprenant 75% des acides aminés du domaine. Ces trois hélices α sont stabilisées par un noyau hydrophobe. La partie longue du L est composée de l'extension N-

terminale (résidus 5 à 12) et de l'hélice III (résidus 50 à 74) et la partie courte comprend les hélices I et II. (résidus 13 à 29 et 34 à 48 respectivement) (Figure 13 et 14). Les hélices II et III sont à peu près perpendiculaires, formant entre elle un angle d'environ 80° (Weir et al. 1993).

Certains résidus sont conservés entre les séquences de HMG-box. Ces acides aminés sont impliqués dans le maintien de la structure tridimensionnelle du domaine et dans l'interaction avec l'ADN (Weir et al. 1993) (Figure 13 et 14):

- Phe 14, Phe 17, Trp 45, Lys 53 et Tyr56 se regroupent à la jonction de deux bras du L du domaine HMG afin de maintenir l'angle de la structure tridimensionnelle.
- Les Pro7 et Pro10 de l'extrémité N-terminal non structurée interagissent avec Tyr67 de l'hélice III, tandis que Pro30 est important dans la conformation de la boucle entre les hélices I et II.
- Les acides aminés Arg ou Lys en position 8, 9, 22, 39 et 64 sont conservés entre toutes les HMG-box, la plupart de ces résidus basiques sont sur ou à proximité de la surface concave formée entre les deux bras de la structure L.
- Les résidus à la position 39 et 40 dans le bras court, forment des ponts salins avec Glu34, Asp36 et deux autres résidus acides conservés à la position 20 et 28.

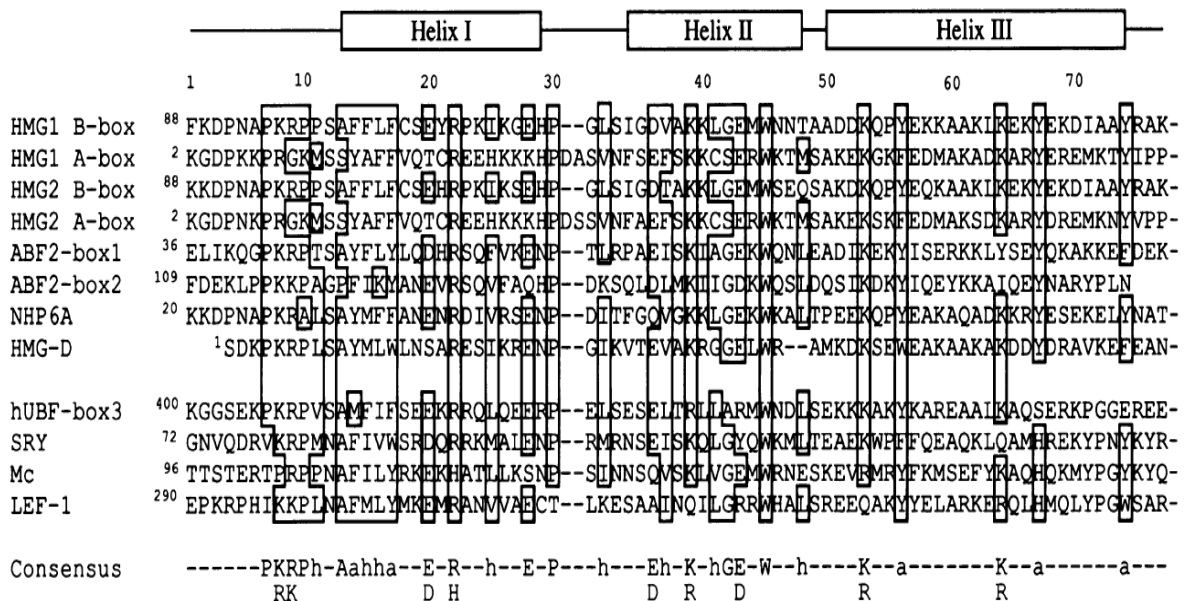


Figure 13 : Alignement des séquences protéiques des HMG-box.

La numérotation en haut se réfère à HMG1 B-box : le résidu 1 de HMG-box correspond au résidu 88 de la protéine HMG1 B (Bianchi et al. 1989). HMG1, high mobility group 1 du rat (Bianchi et al. 1989); HMG2, high mobility group 2 du porc (Shirakawa et al. 1990); ABF2, *S. cerevisiae* ARS binding factor (Diffley and Stillman 1991); NHP6A, *S. cerevisiae* non-histone protein A (Kolodrubetz and Burgum 1990); HMG-D, *Drosophila melanogaster* high mobility group protein (Wagner et al. 1992); hUBF, upstream binding factor for RNA polymerase I de l'homme (Jantzen et al. 1990); SRY, testis-determining factor de l'homme (Sinclair et al. 1990); Mc, mating type protein de *S. pombe* (Kelly et al. 1988); LEF-1, murine lymphoid enhancer factor 1 (Travis et al. 1991). Les boîtes de résidus représentent les résidus conservés entre au moins 8 séquences des 12 séquences analysées. Code du consensus: a, les acides aminés aromatiques (Y, F, W et H); h, les acides aminés hydrophobes (V, I, A, L et F) (Weir et al. 1993).

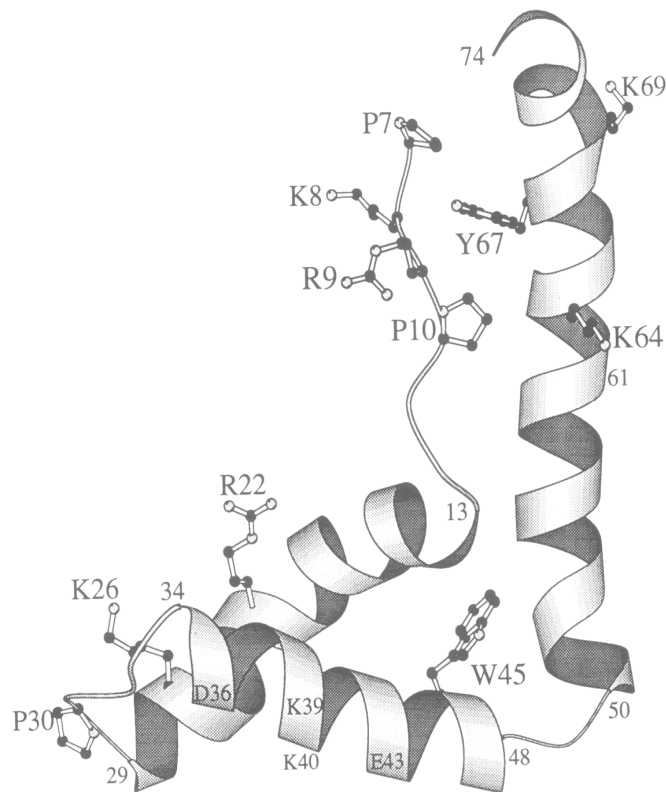


Figure 14: Structure en L d'une boîte HMG. Les résidus représentés sont les acides aminés conservés. (Weir et al. 1993)

Chez les vertébrés et chez les plantes, les protéines HMGB présentent une région basique aux extrémités N et C terminale du domaine HMG ainsi qu'une queue acide en C-terminale de la protéine (Thomas and Travers 2001). Néanmoins, chez les eucaryotes inférieurs, l'extension basique (*D. melanogaster*) ou la queue acide (*S. cerevisiae*) peuvent manquer. L'extension basique semble jouer un rôle dans l'activité de liaison et de courbure de la protéine (Teo et al. 1995). Le rôle de la queue acide est plus vague; elle pourrait interagir avec les charges positives des histones (Thomsen et al. 2004).

Plusieurs différences ont été identifiées entre les protéines HMG qui se fixent spécifiquement à l'ADN et celles qui se fixent non spécifiquement à l'ADN (Murphy et al. 1999). En particulier, dans les domaines HMG se fixant de façon spécifique sur les séquences ADN (SRY, Mc et LEF-1) la position 12 est une asparagine tandis que dans les domaines se fixant de façon non spécifique (HMG1, HMG2, ABF2, NHP6A, HMG-D) on trouve une sérine (Figure 13) (Murphy et al. 1999); cette différence est expliquée par la capacité de la sérine à se lier à n'importe quelle base, soit directement avec la cytosine et la thymine, soit par l'intermédiaire d'une molécule d'eau dans le cas de l'adénine et de la guanine (Murphy et al. 1999). Le résidu 34 est hydrophobe dans les domaines HMG se fixant de façon non spécifique à l'ADN mais hydrophile dans les domaines HMG se fixant de façon spécifique à l'ADN (Balaeff et al. 1998) (Figure 13). Chez LEF-1 par exemple, le résidu Ser34 est

impliqué dans la liaison de HMG avec les bases AA du motif de fixation sur l'ADN (Murphy et al. 1999). La région C-terminale de chacune des deux types de protéines HMG interagit différemment avec l'ADN. Par exemple l'extrémité C-terminale de HMG-D (HMG1 de *D. melanogaster*) n'interagit pas avec l'ADN tandis que celle de LEF-1 se lie à l'ADN pour stabiliser la conformation (Murphy et al. 1999).

Le domaine HMG interagit avec le petit sillon de la double hélice d'ADN. Dans ce cas, des résidus d'acides aminés hydrophobes se positionnent à l'intérieur du sillon et une série de résidus basiques interagit avec le squelette phosphodiester. Certains acides aminés hydrophobes de cette face concave s'intercalent partiellement entre les bases, ce qui provoque l'élargissement du petit sillon, le déroulement de la double hélice, et sa courbure vers le grand sillon (Stott et al. 2006).

HMGB1 contient deux domaines HMG, désignés A et B. Les premières analyses ont clairement mis en évidence des propriétés différentes pour chacune des boîtes. La boîte A a une affinité plus élevée que la boîte B pour les structure d'ADN en croix (Teo et al. 1995; Webb and Thomas 1999). En revanche, les propriétés de la boîte B semblent plus difficiles à caractériser. Pour mieux déterminer la nature de l'interaction de HMGB1 à l'ADN, Stott et ses collaborateurs (Stott et al. 2006) ont construit une protéine chimère, dans laquelle la boîte A de HMGB1 est remplacée par la boîte HMG du facteur de transcription SRY qui se lie à l'ADN de façon spécifique. La protéine ainsi obtenue est capable de former un complexe stable avec l'ADN et l'analyse structurale, par résonance magnétique nucléaire (RMN), des complexes ADN/protéine chimère ainsi formés a permis de montrer, pour la première fois, l'interaction de la boîte B avec de l'ADN (Stott et al. 2006). L'interaction de chacune des deux boîtes HMG entraîne la courbure de l'ADN aux deux sites d'interaction (la somme des deux courbures de l'ADN est 110).

Les analyses RMN du complexe HMG-ADN faites sur SRY et LEF-1 montrent que le domaine HMG reconnaît 10 à 12 paires de bases avec une forte spécificité mais la longueur et la composition de ce site de fixation peut diverger selon les différents domaines HMG (van Beest et al. 2000). Pour LEF-1, le site de fixation est TTCAAAGG (Giese et al. 1992) tandis que SRY se fixe sur le site AACAAAG (Ferrari et al. 1992). D'autres études réalisées sur TCF-1 le facteur de transcription spécifique des cellules T, sur SOX5 et sur Mat-Mc (Kjaerulff et al. 1997) proposent la séquence (A/T)(A/T)CAAAG comme site de fixation du domaine HMG. Dans tous les cas, la fixation du domaine HMG sur son site de fixation est réalisée par l'intercalation des résidus hydrophobes entre deux bases de la séquence (A/T)(A/T)CAAAG (van de Wetering et al. 1991). Ces résidus sont soit Met soit Ile ou bien

Val et ils sont flanqués par deux résidus aromatiques aux positions -1 et +2 (van de Wetering et al. 1991; Martin et al. 2010).

3.2.2 Localisation des protéines HMGB dans la cellule

Plusieurs études se sont intéressées à la localisation des protéines HMG dans la cellule pendant l'état de repos et la division cellulaire. Les protéines de la famille HMGB sont localisées au voisinage de la chromatine dans les cellules en interphase, et sur les chromosomes mitotiques. Cependant, elles circulent très rapidement dans l'espace nucléaire, en interphase et en mitose (Marechal 2003). Les interactions entre ces protéines et la chromatine sont extrêmement brèves et labiles (Pallier et al. 2003).

3.2.3 Fonction des protéines à domaine HMG

3.2.3.1 Assemblage des nucléosomes, remodelage de l'ADN et transcription

Les protéines à domaine HMG, moins abondantes que les histones H1 (environ 10 à 15 fois moins que l'histone H1), sont aussi appelées les protéines chromosomiques non histone (NHP) ; elles sont associées à une chromatine active (Nemeth and Langst 2004) et se lient préférentiellement à l'ADN qui relie les nucléosomes entre eux pour faciliter l'assemblage des complexes nucléoprotéiques (Stros 2009).

L'histone H1 et les protéines HMGB sont impliquées dans les changements dynamiques de conformation de la chromatine. Elles possèdent des rôles antagonistes (Bianchi and Agresti 2005). En effet, H1 diminue l'accessibilité des facteurs de régulation aux nucléosomes, tandis que les protéines à domaine HMG décompactent l'ADN et déstabilisent les nucléosomes pour rendre l'ADN plus accessible au complexe de remodelage de la chromatine et aux facteurs de régulation (Bianchi and Agresti 2005) (Figure 15). HMGB1 a aussi la particularité de protéger les nucléosomes des nucléases (Stros 2009).

Les protéines HMGB favorisent la transcription des gènes par plusieurs mécanismes qui sont décrits ci-dessous.

a- Effet sur la chromatine

Les protéines HMGB entraînent une flexibilité de la structure chromatinienne, en affectant l'espacement et le repliement du nucléo-filament, ce qui permet à la machinerie de transcription d'accéder facilement aux gènes (Bianchi and Agresti 2005) (Figure 15) . Un exemple est le complexe FACT (Facilitates Chromatin Transcription) qui est un complexe hautement conservé chez les eucaryotes. C'est un hétérodimère composé de protéines à boîtes HMG (SPT16/SSRP1 chez l'Homme et Spt16/Cdc68 chez *S. cerevisiae*) qui facilite l'élongation de la transcription par l'ARN polymérase II et permet la restructuration des

nucléosomes en se fixant sur les histones H2A et H2B (Orphanides et al. 1999). Chez la drosophile, FACT (dSPT16 et dSSRP1) a été co-purifié avec la protéine GAF à partir d'extraits embryonnaires. Le complexe GAF-FACT de drosophile est impliqué dans le remodelage de la chromatine nécessaire au maintien de l'expression des gènes homéotiques (Shimajima et al. 2003).

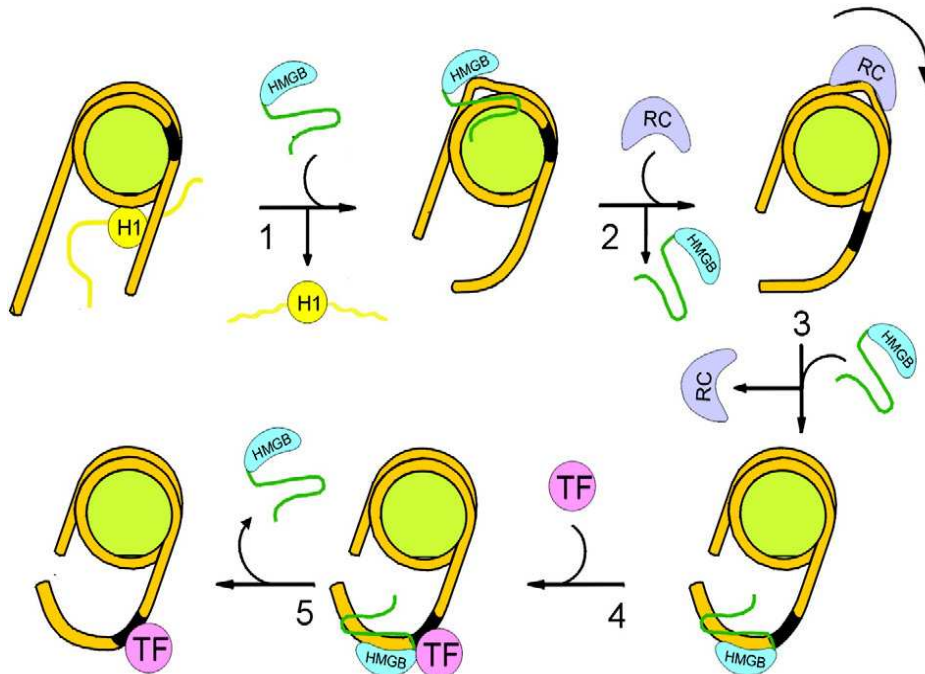


Figure 15 : Modèle de fonctionnement des protéines HMGB.

L'histone H1 se fixe sur le nucléosome (en vert). 1) La protéine HMGB déplace H1 et se fixe sur le bord du nucléosome ce qui permet la formation d'une boucle d'ADN (en jaune). 2) La boucle d'ADN représente une cible d'ancrage du complexe de remodelage (RC) qui déplace la boucle d'ADN autour du nucléosome après le détachement de HMGB. 3) RC se détache et HMGB se fixe à l'ADN pour rendre la séquence spécifique accessible au facteur de transcription (TF). 4) HMGB facilite la fixation de TF qui se fixe sur le site de la séquence spécifique. 5) HMGB se détache de l'ADN et TF régule la transcription de sa cible (Stros 2009).

b- Interaction avec des facteurs de transcription

Les protéines HMGB participent aux interactions entre l'ADN et de multiples facteurs; elles peuvent interagir avec la machinerie de transcription ou avec des facteurs transcriptionnels individuels (les protéines HOX, p53, NF- κ B). En se liant à l'ADN, HMGB induit une courbure qui facilite l'interaction des protéines avec les sites séquences-spécifiques de l'ADN. Par exemple, HMGB1 modifie la conformation de l'ADN pour une meilleure liaison de p53 et augmente ainsi son activité transactivatrice (Jayaraman et al. 1998). Les protéines NHP6 (non-histone proteins) facilitent la fixation de TFIIC et TBP (TATA box Binding Protein) au promoteur de *SNR6* (U6 small nuclear RNA). La délétion de *NHP6A* et *NHP6B* chez la levure provoque un défaut de la transcription de *SNR6* par la polymérase III. Le double mutant *nhp6a* et *b* a une faible croissance et une forte sensibilité à la température (Kruppa et al. 2001; Bianchi and Agresti 2005).

c- Activation de la transcription

L'initiation de la transcription de l'ADN ribosomique (ADNr) requiert l'assemblage d'un complexe multi-protéique spécifique, le complexe de pré-initiation (PIC : Pre-Initiation Complex), qui permet le recrutement de l'ARN Pol I au niveau du promoteur. Chez les mammifères, l'action synergique de deux facteurs de transcription est requise pour l'initiation de la transcription des ADNr: UBF (Upstream Binding Factor), et SL1 (Selectivity factor 1). SL1 est un complexe multiprotéique constitué d'une protéine TBP (TATA box Binding Protein) et de cinq facteurs associés TAFIs (TATA binding protein–Associated Factors) (Gorski et al. 2007). UBF intervient dans la première étape de l'assemblage du complexe d'initiation de la transcription de l'ARN polymérase I. Cette protéine comporte six HMG-box qui vont lui permettre de lier l'ADN. Comme les autres protéines HMG, UBF lie le sillon mineur de l'ADN pour former une boucle de $\approx 360^\circ$ tous les 140 pb ce qui permet de rapprocher les éléments Core et UCE (Upstream Control Element) (Grummt 2003). La mise en place du complexe transcriptionnel sur le promoteur est initiée par la phosphorylation en C-terminal d'UBF. UBF1 phosphorylée stabilise SL1 sur le promoteur et recrute l'ARN Pol I.

La différenciation de l'appareil génital mâle chez les mammifères nécessite la présence du chromosome Y et plus particulièrement du gène *Sry* (Sex-determining Region of Y chromosome). Ce gène code un facteur de transcription avec un domaine HMG qui régule la transcription des gènes impliqués dans la détermination du sexe en se fixant sur un site spécifique : AACAAAG. Les mutations au niveau de ce domaine causent des réversions sexuelles (Goodfellow and Lovell-Badge 1993). L'extrémité C-terminale de SRY interagit avec deux partenaires protéiques, la calmoduline, récepteur de calcium intracellulaire et SIP, partenaire général des protéines à domaine HMG. Divers sites potentiels de fixation pour les protéines SRY ont été localisés dans les régions promotrices du gène *Amh* ou *Hormone anti-müllérienne* (la dégénérescence du canal de Müller), du gène de l'aromatase P450 (conversion de la testostérone en oestradiol) (Haqq et al. 1993) et du gène *Fra-1* ou *Fos-related antigen 1* (régulation de la spermatogenèse) (Cohen et al. 1994).

3.2.3.2 HMGB et les dommages de l'ADN

Les protéines HMGB sont aussi capables de détecter les déformations et les dommages au niveau de l'ADN, de maintenir l'intégrité du génome et de réparer l'ADN (Bianchi and Agresti 2005).

Le rôle joué par les protéines HMGB dans le processus de réparation de l'ADN a été étudié dès 1992, quand Pil et ses collaborateurs (Pil and Lippard 1992) ont montré que HMGB1 se fixait spécifiquement à l'ADN endommagé par le cisplatine, un médicament

anticancéreux. Cette fixation induit le recrutement des protéines du système de réparation NER (Nucleotide Excision Repair) aux sites de lésions de l'ADN. En absence de HMGB1, l'efficacité des protéines du NER est significativement diminuée et cela peut donc entraîner l'accumulation de mutations, la mort cellulaire ou une altération de l'organisation de la chromatine (Reddy et al. 2005). Les cellules fibroblastiques embryonnaires des souris mutantes *Hmgb1*^{-/-} sont hypersensibles aux dommages à l'ADN causés par l'utilisation combinée d'agents chimiques et des UVs (Krynetski et al. 2003; Reeves and Adair 2005)

3.2.3.3 Autres fonctions

Parkkinen et ses collaborateurs (Parkkinen et al. 1993) sont les premiers à avoir mis en évidence la capacité de la protéine HMGB1 à interagir avec l'héparine et les héparanes sulfates, et donc à agir hors de la cellule. Depuis une dizaine d'années la protéine HMGB1 est décrite comme un médiateur soluble quand elle se retrouve dans le milieu extracellulaire.

La protéine HMGB1 est sécrétée au cours de l'activation des monocytes et macrophages et agissent en fait comme des cytokines inflammatoires (Wang et al. 1999). Les modifications par acétylation de multiples lysines délocaliseraient les protéines HMGB1 du noyau vers le cytoplasme pour la sécrétion (Bonaldi et al. 2003). Alors que les macrophages secrètent activement les protéines HMGB hyperacétylées, les cellules en nécrose les libèrent passivement et les cellules en apoptose quant à elles, retiennent les protéines HMGB irréversiblement liées à la chromatine condensée. La libération de la protéine HMGB conduit au recrutement des cellules de l'inflammation et déclenche la production de cytokines TNF α et IL α 1 par les monocytes (Andersson et al. 2000). L'activité cytokinique de la protéine HMGB est liée à un domaine de 20 acides aminés présent dans la boîte B de la protéine HMGB1 humaine (Li et al. 2003). Les protéines HMGB agissent donc comme médiateurs dans la réponse immunitaire et comme facteurs d'activation des monocytes/macrophages (Muller et al. 2001; Dumitriu et al. 2005).

Les protéines HMGB libérées dans le milieu extracellulaire sont capables de se lier à différents récepteurs cellulaires. (Park et al. 2004). A ce jour, trois récepteurs ont été décrits : TLR-2 et 4 (Toll-like récepteur) et le récepteur RAGE (Recepteur for Advanced Glycation end Products).

- Les TLR-2 et 4 sont impliqués dans la reconnaissance de composants microbiens et dans la morphogénèse. Ces récepteurs sont aussi impliqués dans la voie de signalisation qui active la protéine NF-kB, nécessaire au nombreux processus cellulaires comme la prolifération, l'apoptose ou la réponse inflammatoire (Park et al. 2004).

- Le récepteur RAGE appartient à la superfamille des immunoglobulines. L'interaction d'HMGB1 avec RAGE déclenche une cascade de signalisation menant à l'activation des voies MAP kinases et NF- κ B. En fonction du contexte d'activation, RAGE déclenche soit une réponse inflammatoire soit la réparation des tissus endommagés associés à des maladies chroniques. Il reconnaît un large spectre de ligands (Kokkola et al. 2005).

De plus, les protéines HMGB extracellulaires sont un signal lors de la lésion tissulaire : elles induisent la migration et la prolifération des mésoangioblastes et participeraient ainsi à la régénération des tissus musculaires (Palumbo et al. 2004).

3.2.3.4 Rôles des protéines à domaine HMG dans le développement sexué chez les champignons

Les protéines à domaine HMG sont aussi impliquées dans le processus de développement sexué chez les champignons. Les gènes à HMG-box les plus étudiés sont les gènes des types sexuels et le gène *ste11* de *S. pombe*.

3.2.3.4.1 Mating type

Le locus du type sexuel ou mating type (*mat*) des champignons ascomycètes code des protéines avec des domaines de fixation à l'ADN appartenant aux familles des homéodomains et des HMG-box. Les protéines à domaine HMG des types sexuels appartiennent exclusivement au groupe des MATA_HMG (NCBI Conserved Domain Database accession number: cd01389), où l'on peut différencier différents sous-groupes :

- les protéines à domaine MATA_HMG typique (MAT1-2-1 ou FPR1, et MAT1-1-3 ou SMR2 de *P. anserina*). Il faut noter cependant que même dans ce sous-groupe typique, les protéines MAT1-2-1 et MAT1-1-3 ont pu être différenciées phylogénétiquement et fonctionnellement (Martin et al. 2010).
- les protéines à domaine MAT α _HMG (Martin et al. 2010), anciennement connu sous le nom de domaine α 1. Le prototype de cette catégorie est Mata1p de *S. cerevisiae*, et les gènes *MAT1-1-1* que l'on trouve chez quasiment tous les champignons filamenteux ascomycètes (Martin et al. 2010).
- les protéines avec un domaine apparenté aux HMG, de classification douteuse, comme par exemple Pc de *S. pombe* (Martin et al. 2010).

3.2.3.4.2 Ste11p facteur de transcription à domaine HMG chez *S. pombe*

Chez la levure *S. pombe*, le programme de développement sexué est induit en réponse à une carence nutritionnelle, en particulier la carence en azote. Ce signal déclenche la

conjugaison des cellules *P* et *M* et la formation d'un zygote, qui subit la méiose et la sporulation. Le facteur de transcription Ste11p est essentiel au développement sexué; il active plusieurs gènes impliqués dans la conjugaison et la méiose, notamment les gènes *mat* qui contrôlent le système phéromone/récepteur et *mei2* qui déclenche la méiose (Kjaerulff et al. 2005). (Voir chapitre I paragraphe 2.3.3). La mutation de *ste11* provoque une stérilité complète de *S. pombe* et une déficience en méiose (Sugimoto et al. 1991). L'expression hétérochronique de ce gène cause une différenciation sexuelle indépendante des conditions nutritionnelles.

a. La protéine Ste11p

La protéine Ste11p de 468 acides aminés est codée par le gène *ste11* (*aff1* ou *stex*) localisé sur le chromosome 2. Ste11p appartient au groupe des MATA_HMG caractérisé par la présence d'un seul domaine HMG et une fixation spécifique à l'ADN. La localisation de Ste11p dans la cellule régule son activité. Elle est soit nucléaire, soit cytoplasmique, en fonction des mécanismes suivants : la kinase Pat1p/Ranp inhibe l'accumulation de Ste11p dans le noyau par la phosphorylation de la thréonine 173 et la serine 218 localisées sur les deux domaines RKDI et RKDII (Ran Kinase Domain) (Qin et al. 2003) (Figure 16). Cette modification permet la fixation de la protéine Rad24p (Kitamura et al. 2001) qui piège Ste11p dans le cytoplasme (Qin et al. 2003) (Figure 17). Le domaine HMG de Ste11p (du résidu 16 au résidu 50) contient une séquence basique (⁸⁴KKRSTVRRRHKK) (Figure 16) qui est nécessaire à l'accumulation de Ste11p dans le noyau (Qin et al. 2003)



Figure 16: La protéine Ste11p. La boîte bleue représente le domaine HMG. La boîte noire représente la séquence basique BC (basic cluster). Les régions RKDI et RKDII (Ran Kinase Domain) contiennent les sites de phosphorylation (Qin et al. 2003).

b. Régulation de Ste11p

Le signal de carence nutritionnelle induit l'expression de *ste11* par la stimulation de trois voies différentes de signalisation : la voie d'AMP cyclique (AMPC), la voie de réponse au stress (carence en azote) et la voie de signal phéromone (Yamamoto et al. 1997; Davidson et al. 2004).

- la voie AMPC est constituée de plusieurs interactions protéiques: l'adenylyl cyclase codée par *cyr1* qui génère l'AMPC à partir de l'ATP, la protéine trimérique G qui contrôle l'adenylyl cyclase, l'AMPC phosphodiesterase codée par le gène *cgs2* qui régule négativement la voie d'AMPC (Davidson et al. 2004; Jeong et al.

2004). La protéine kinase A (PKA ou cAMP-dependent protéine kinase) est la cible majeure de l'AMPc. La PKA régule la transcription de *ste11* via le facteur de transcription à doigt de zinc Rst2p (Valbuena and Moreno 2010). Lorsque le niveau de l'AMPc intracellulaire est élevé, la PKA phosphoryle Rst2p et le piège dans le cytoplasme, tandis que durant une carence nutritionnelle le niveau d'AMPc et l'activité de la PKA diminuent, ce qui permet la migration de Rst2p dans le noyau pour activer la transcription de *ste11* (Valbuena and Moreno 2010) (Figure 17).

- La voie de réponse au stress, et plus particulièrement la carence en azote déclenche la voie MAPK composée principalement de deux MAPKK Kinase Wis4p et Win1p, de la MAPK Kinase Wis1p et de la MAP kinase Sty1p (Warbrick and Fantès 1991; Kato et al. 1996; Samejima et al. 1997; Samejima et al. 1998; Jeong et al. 2004). En réponse au stress, Sty1p phosphoryle le facteur de transcription Atf1p qui active la transcription de *ste11* et *cgs2*, ce dernier codant l'AMPc phosphodiesterase (Davidson et al. 2004) (Figure 17) .
- La voie de signal phéromone est composée des MAPKK kinase Ras1p et Byr2p, la MAPK kinase Byr1p et la MAP kinase Spk1p (Fukui et al. 1986; Nadin-Davis and Nasim 1988; Toda et al. 1991; Wang et al. 1991; Neiman et al. 1993). La fixation des phéromones sur leur récepteur active la protéine Gpa1p (Obara et al. 1991) qui transmet le signal à Byr2p à l'aide de Ras1p et Ste4p (Barr et al. 1996; Jeong et al. 2004). Byr2p phosphoryle Byr1p qui lui-même phosphoryle Spk1p. La MAP kinase Spk1p active Ste11p via trois voies différentes: i) d'une façon directe via la phosphorylation des thréonines T305 et T317 de Ste11p, et indirectement ii) par la déphosphorylation des sites Pat1p (T173 et S218) et iii) l'inactivation de Pat1p via Mei3p (Kjaerulff et al. 2005) (Figure 17).

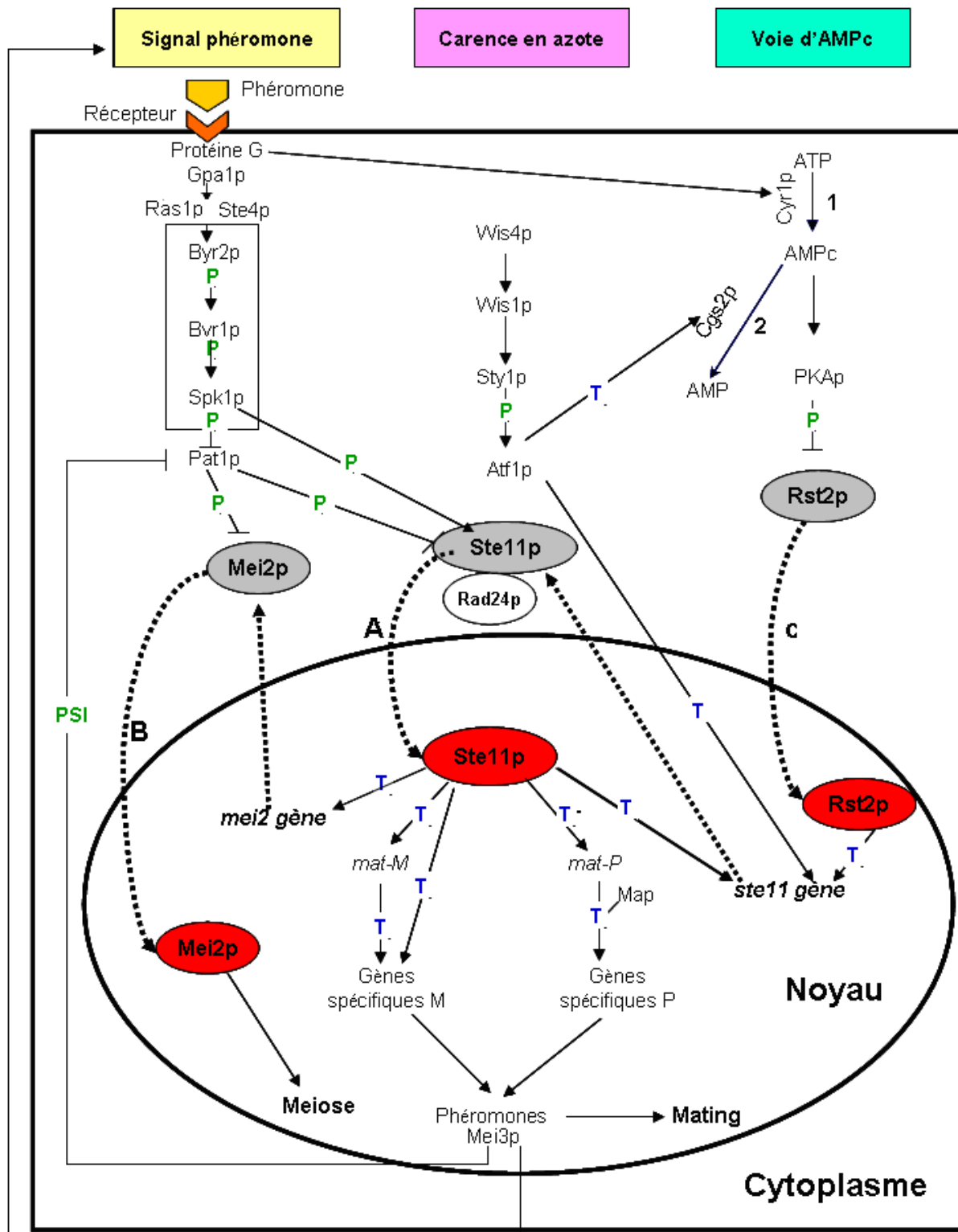


Figure 17 : Régulation de la conjugaison chez *S. pombe*.

La flèche noire représente une activation, la barre représente une répression. (T) indique une régulation transcriptionnelle, (P) indique une régulation traductionnelle par phosphorylation de la protéine et (PSI: Pseudo-Substrate Inhibition) indique une régulation de la protéine par inhibition. En gris Ste11p, Mei2p et Rst2p phosphorylés ; en rouge Ste11p, Mei2p et Rst2p non-phosphorylés. Rad24p se fixe sur Ste11p phosphorylé par Pat1p pour le piéger dans le cytoplasme. (A) Ste11p non-phosphorylé sur les sites reconnus par la kinase Pat1p et phosphorylé dans les sites reconnus par la MAPK Spk1p entre dans le noyau. (B) Mei2p déphosphorylé dans les sites Pat1p entre dans le noyau. (C) Rst2p déphosphorylé entre dans le noyau et active Ste11p. 1) L'adenylyl cyclase (Cyr1p) génère l'AMPc à partir de l'ATP ; 2) La phosphodiesterase (Cgs2p) hydrolyse l'AMPc en AMP. Schéma obtenu par la compilation des données de (Sugimoto et al. 1991; Qin et al. 2003; Davidson et al. 2004; Jeong et al. 2004; Kjaerulff et al. 2005; Mata and Bähler 2006; Valbuena and Moreno 2010).

c. Ste11p régule la différenciation sexuée

Ste11p contrôle l'expression des gènes essentiels pour le développement sexuel dans les cellules *P* et *M*. En effet, dans une cellule *M*, Ste11p active l'expression du locus *matM* (*mat1-Mm* et *mat1-Mc*) (Sugimoto et al. 1991). Mat-Mc (protéine à domaine HMG) et Ste11p activent l'expression des gènes spécifiques *M* (Kjaerulff et al. 1997) (Figure 17). Dans la cellule *P*, Ste11p active l'expression du locus *matP* (Sugimoto et al. 1991), Mat-Pc et Map1p (un orthologue de Mcm1p de *S. cerevisiae*, voir son rôle en 2.3.3) vont activer les gènes spécifiques *P* (Mata and Bähler 2006) (Figure 17). Récemment 78 gènes cibles de Ste11p ont été identifiés durant une analyse transcriptomique des souches *P* et *M* (Mata and Bähler 2006), parmi lesquels 12 gènes spécifiques de *M* et 4 gènes spécifiques de *P* (Tableau 2).

Gènes	Fonctions
Gènes spécifiques des cellules <i>M</i>	
<i>mfm1/ mfm2 /mfm3</i>	Phéromone M
<i>mam1</i>	Transporteur de la phéromone M
<i>cwp1</i>	Geranyl-geranyl transférase I
<i>mam2</i>	Le récepteur de la phéromone P
<i>sax2</i>	La protéase qui dégrade la phéromone P
<i>SPAP11E10.02c</i>	Agglutinine
<i>mam4</i>	S isoprenylcysteine O-methyltransferase
<i>SPAPB1A10.02</i>	Fonction inconnue
<i>SPAC11H11.03c</i>	Fonction inconnue
<i>SPAC11H11.05c</i>	Fonction inconnue
Gènes spécifiques des cellules <i>P</i>	
<i>map2</i>	Phéromone P
<i>map3</i>	Le récepteur de la phéromone M
<i>SPBC21D10.06c</i>	Agglutinine
<i>SPAC1665.03</i>	Fonction inconnue

Tableau 2: Les gènes spécifiques des cellules *M* et les gènes spécifiques des cellules *P* régulés par Ste11p (Mata and Bähler 2006).

d. Ste11p régule la méiose

Mei2p est une RNA-binding protéine qui déclenche la méiose. Cette protéine est présente à un niveau très faible pendant la croissance végétative et reste inactive durant la conjugaison à cause de sa phosphorylation par Pat1p qui la convertit en substrat des enzymes de protéolyse (Watanabe et al. 1997). Par conséquent, la phosphorylation de Mei2p par Pat1p

inhibe l'initiation de la méiose pendant le cycle cellulaire mitotique. Dans un diploïde, *mat1-Pm* et *mat1-Mm* activés par Ste11p, lui-même activé par la voie de signal phéromone, induisent la transcription de *mei3* qui agit comme pseudosubstrat de Pat1p. En conséquence, Pat1p ne peut plus phosphoryler Mei2p qui déclenche la méiose (Davis and Smith 2001; Kjaerulff et al. 2005). De plus, Ste11p se fixe spécifiquement en amont de *mei2* pour activer sa transcription (Sugimoto et al. 1991) (Figure 17).

e. Site de fixation Ste11p

Van Beest et al ont défini le consensus du site de fixation Ste11p $A^1GAACAAAGAAA^{12}$ en amont des gènes cibles de Ste11p (van Beest et al. 2000). Ce site est appelé TR-box en raison de sa richesse en T (T-Rich box). Il est plus long que le motif de fixation (6-8pb) des autres protéines à HMG-box (Laudet et al. 1993). Des analyses RMN ont montré que la région C-terminale du domaine HMG reconnaît A^{11} et A^{12} de la TR-box et que les deux hélices 1 et 2 interagissent avec le cœur du motif $^3ACAAAGA^{10}$ (van Beest et al. 2000) (Figure 18). La séquence 3' adjacente à $A^{11}A^{12}$ est conservée entre les cibles de Ste11p. Deux classes de sites Ste11p peuvent être trouvées en amont des gènes cibles de Ste11p (van Beest et al. 2000) :

- les cibles exprimées d'une façon ubiquitaire possèdent $A^{11}A^{12}$ (Tableau 3).
- les cibles exprimées uniquement chez les cellules M ne possèdent pas $A^{11}A^{12}$ (Tableau 3).

Cette différence nécessite l'intervention de la protéine à domaine HMG Mat-Mc pour l'activation des gènes spécifique M et d'un site appelé M-box. Ce site est adjacent à la TR-box et il est composé de deux séquences ACAAT. La protéine Mat-Mc se fixe spécifiquement au site M-box en amont des gènes spécifiques M. La fixation de Mat1-Mc change la conformation de l'ADN ce qui augmente l'affinité de Ste11p à la TR-box et l'activation des gènes spécifique des cellules M (Kjaerulff et al. 1997). Dans les cellules P, Ste11p ne peut pas se fixer sur son site en amont des gènes spécifique M à cause de l'absence de Mat1-Mc, les gènes spécifiques P sont activés par Mat1-Pc et Map1p.(Kjaerulff et al. 1997).

Gene	DNA motif	Expression
Optimal	AGAACAAAGAAA	
<i>fus1</i>	AGAACAAAGAAA	Ubiquitous
<i>fus1</i>	cGAACAAAGAAA	Ubiquitous
<i>map1</i>	gtAACAAAGAAA	Ubiquitous
<i>mei2</i>	gGAACAAAGAAA	Ubiquitous
<i>mei2</i>	AaAACAAAGAAA	Ubiquitous
<i>mei2</i>	caAACAAAGAAA	Ubiquitous
<i>ste4</i>	AtAACAAAGAAA	Ubiquitous
<i>ste6</i>	taAACAAAGAAA	Ubiquitous
<i>ste11</i>	gcAACAAAGAAA	Ubiquitous
consensus	AACAAAGAAA	Ubiquitous
<i>mam1</i>	ctAACAAAGgcc	MatMc dependent
<i>mam2</i>	taAACAAAGagg	MatMc dependent
<i>mfm1</i>	ccaAACAAAGAAg	MatMc dependent
<i>mfm2</i>	tGAACAAAGAgA	MatMc dependent
<i>mfm3</i>	AcAACAAAGAAg	MatMc dependent
<i>sxa2</i>	gcAACAAAGAcA	MatMc dependent
consensus	AACAAAGA	MatMc dependent

Tableau 3: Comparaison des motifs TR des gènes régulés par Ste11p.
(van Beest et al. 2000)

Sugimoto et ses collaborateurs ont proposé que Ste11p active la transcription de son propre gène. Cette hypothèse est confirmée par la présence d'une TR-box en 5'- UTR de *ste11*. Cette autorégulation permet d'augmenter le signal de réponse en carence (Sugimoto et al. 1991). Deux TR-box ont été aussi trouvées dans les deux locus *matM* et *matP* régulés par Ste11p ; ces motifs sont localisés entre les deux gènes de chaque locus (Sugimoto et al. 1991).

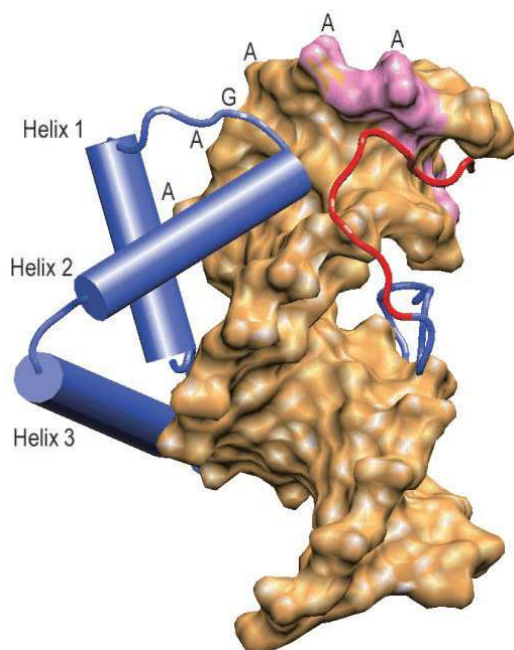


Figure 18 : Modèle de fixation Ste11p-ADN. La protéine Ste11p est en bleue et les cylindres représentent les hélices α . En rouge la partie C-terminale de Ste11p. les paires de base AAGAAA représentent les bases de la position 7 à 12 du motif de fixation de Ste11p (van Beest et al. 2000).

4 Evolution des HMG et du locus *MAT*

Une analyse phylogénique des HMG-box montre que cette superfamille est extrêmement ancienne (Soullier et al. 1999). En fait, les membres de cette famille sont connus dans tous les groupes de métazoaires, chez les eucaryotes unicellulaires tels que *Trypanosoma* (Erondu and Donelson 1992), chez les plantes, et chez les champignons. Il apparait donc clairement que la superfamille des HMG était présente longtemps avant la séparation de ces groupes (Soullier et al. 1999).

La superfamille des protéines HMG-box est composée de deux familles : la famille UBF/HMG présente chez les trois embranchements des eucaryotes plantes, champignons et animaux, et la famille MATA/TCF/SOX (Voir paragraphe 3.1 chapitre I), d'apparition plus récente. La famille MATA/TCF/SOX est composée elle-même de trois sous-familles : la sous-famille MATA regroupant les gènes des types sexuels et Ste11, la sous-famille TCF (T-cell transcription factors) et la sous-famille SOX des gènes apparentés à SRY (Soullier et al. 1999). Les gènes MATA_HMG sont présents exclusivement chez les champignons, tandis que TCF et SOX sont présents chez les animaux. La séquence consensus du domaine HMG dans les trois sous familles affiche une similarité significative (40% d'acide aminés identiques entre MATA_HMG et SOX) (Martin et al. 2010).

Idnurm et ses collaborateurs ont proposé que les gènes à HMG-box représentent l'ancêtre du facteur déterminant la compatibilité sexuelle (qu'ils ont qualifiée de sexe) chez les champignons (Idnurm et al. 2008). Cette hypothèse est basée sur la présence d'un gène à HMG-box déterminant le sexe chez les Zygomycètes et les Microsporidies, qui représentent les premières branches de l'arbre des champignons (Idnurm et al. 2008). Chez les champignons, la détermination du sexe est contrôlée par une petite région spécialisée du génome (locus *MAT*). Le locus *MAT* a été caractérisé chez les Ascomycètes et les Basidiomycètes et présente une structure et un contenu très variable (Figure 19), ce qui constitue une contradiction avec le modèle proposé par Idnurm et al.

En particulier, chez les Ascomycota, l'un des allèles du locus *MAT* est caractérisé par la présence d'une protéine contenant un domaine $\alpha 1$ dont l'origine était inconnue jusqu'à une date récente. La comparaison des séquences, les analyses phylogénique des protéines des types sexuels et la prédiction des structures secondaires et tertiaires, ont permis de classer le domaine $\alpha 1$ dans une nouvelle sous-famille des HMG appelée MAT α _HMG (Martin et al. 2010). Cette découverte a permis de rétablir la continuité de l'histoire évolutive du locus *MAT* entre les Zygomycètes, les Microsporidies et les Ascomycota. En effet, chacun des deux allèles du locus *MAT* des membres de ces groupes contient un gène à HMG-box. La synténie

du locus *MAT* avec certains gènes adjacents à ce locus soutient également cette continuité évolutive (Butler et al. 2004; Martin et al. 2010). Le locus *MAT* a évolué avec des gains et des pertes dans certaines lignées. L'exemple de perte le mieux documenté est celui du gène *MATA_HMG* dans le type sexuel "a" chez *S. cerevisiae* (Butler et al. 2004). Le gain le plus spectaculaire est celui du gène codant une protéine à domaine PPF, qui a un rôle essentiel dans la reproduction sexuée et que l'on trouve chez les Sordariomycètes (Debuchy et al. 2010).

La structure du locus *MAT* des Basidiomycètes est difficile à intégrer dans l'évolution du locus *MAT* entre les Ascomycota et les groupes à la base de l'arbre des champignons. En effet, chez les basidiomycètes le locus *MAT* code des protéines à homéodomaine et des systèmes phéromone/récepteur. Mais pour l'instant, on ne connaît pas chez les Basidiomycètes de locus *MAT* contenant un gène à HMG-box, bien que ceux-ci jouent un rôle essentiel dans la reproduction sexuée, comme cela a été démontré avec *prf1* chez *Ustilago maydis* (Hartmann et al. 1996). Des études supplémentaires seront nécessaires pour comprendre l'origine du locus *MAT* des Basidiomycètes et son éventuelle relation avec les gènes à homeobox du locus *MAT* des Saccharomycotina et des Taphrinomycotina.

		mating type 1					mating type 2		
		$\alpha 1$	PPF	HMG	HD	OTHER	HMG	HD	OTHER
		<i>MAT1-1-1</i>	<i>MAT1-1-2</i>	<i>MAT1-1-3</i>	<i>MAT1-2-1</i>				
Peizizomycotina	Sordariomycetes								
	Sordariales	+	+	+			+		+/-
	Magnaporthales	+	+	+			+		
	Diaporthales	+	+	+			+		
	Hypocreales	+	+	+/-			+		
	Leotiomycetes								
	Helotiales	+		+		+/-			
	Eurotiomycetes								
	Eurotiales	+					+		
	Onygenales	+				+/-	+		
	Dothideomycetes								
	Pleosporales	+					+		
	Dothideales	+					+		
	Ascomycota	<i>S. cerevisiae</i>	<i>MAT $\alpha 1$</i>				<i>MAT $\alpha 2$</i>		
<i>K. lactis</i>		<i>MAT $\alpha 1$</i>				<i>MAT $\alpha 2$</i>	<i>MAT $\alpha 3$</i>	<i>MATa2</i>	<i>MATa1</i>
<i>C. albicans</i>		<i>MTL $\alpha 1$</i>				<i>MTL $\alpha 2$</i>		<i>MTLa2</i>	<i>MTLa1</i>
<i>Y. lipolytica</i>		<i>MAT $\alpha 1$</i>				<i>MAT $\alpha 2$</i>		<i>MATa2</i>	<i>MATa1</i>
		<i>B1</i>				<i>B2</i>		<i>A2</i>	<i>A1</i>
Taphrinomycotina	<i>S. pombe</i>			<i>Pe</i>	<i>Pi</i>		<i>Mc</i>		<i>Mi</i>
Basidiomycota	<i>U. maydis</i>				<i>bW+bE</i>			<i>bW+bE</i>	
Zygomycota	<i>P. blakesleeanus</i>			<i>SexP</i>			<i>SexM</i>		
Microsporidia	<i>E. cuculi</i>						<i>HMG</i>		

Figure 19: La structure des types sexuels chez les champignons. $\alpha 1$, les gènes qui codent les facteurs de transcription à domaine α ; PPF, les gènes qui codent des protéines avec un domaine caractérisé par les résidus conservés proline et phénylalanine; HMG, les gènes qui codent les facteurs de transcription à domaine HMG; HD, les gènes qui codent les facteurs de transcription à homéodomaine. +, gène présent; +/-, gène présent dans certains espèces du groupe (Martin et al. 2010).

5 Objectif de la thèse

L'objectif global de cette thèse est de mieux comprendre la régulation du développement sexuel chez le modèle d'étude des champignons filamenteux *P. anserina*. Les gènes qui régulent les étapes du cycle sexué de *P. anserina* sont maintenant bien déterminés mais les gènes et les mécanismes mis en jeu sont moins connus. Ainsi cette thèse sera divisée en trois principales parties :

- Le premier stade que nous avons analysé par une approche microarray a été celui des souches compétentes pour la fécondation (la reconnaissance intercellulaire). Nous avons recherché les gènes différentiellement exprimés entre une souche sauvage *mat+* et une souche sauvage *mat-*. La comparaison incluait aussi les souches mutantes *fmr1* (*mat-*) et *fpr1* (*mat+*). Ces gènes sont ensuite classés comme étant réprimés et/ou activés par FMR1 et/ou FPR1, les deux facteurs de transcription qui contrôlent l'étape de la fécondation. Les résultats de cette partie seront présentés sous forme d'article qui a été publié récemment dans PLoS ONE sous le titre: «Genome-Wide Gene Expression Profiling of Fertilization Competent Mycelium in Opposite Mating Types in the Heterothallic Fungus *Podospira anserina*».
- La reconnaissance internucléaire entre les noyaux *mat+* et *mat-* est la deuxième étape du cycle sexué étudiée durant ma thèse. Afin d'identifier les gènes cibles des facteurs de transcription MAT : FPR1, FMR1 et SMR2 contrôlant cette étape, nous avons comparé les profils transcriptomiques des deux croisements impliquant pour l'un le mutant du type sexuel *mat+*, pour l'autre le mutant du type sexuel *mat-*. La comparaison incluait aussi les croisements sauvages *mat+* X *mat-*. Les résultats obtenus dans ces comparaisons seront examinés selon deux modèles différents: le modèle d'acquisition de l'identité *mat+* ou *mat-* et le modèle de ségrégation aléatoire.
- La recherche des gènes à HMG-box de *P. anserina* et la détermination de leurs rôles dans le développement sexué. Cette étude a porté spécialement sur l'orthologue de Ste11p de *S. pombe*.

L'identification des gènes acteurs du développement sexuel ne livrera pas nécessairement le mécanisme de ce processus. Pour chaque partie de ma thèse nous avons réalisé des invalidations de gènes, des analyses phénotypiques des mutants et la comparaison de nos résultats avec ceux obtenus chez d'autres ascomycètes.

Chapitre II : Matériels et Méthodes

1. Les souches de *Podospora anserina*, *Escherichia coli* et levure

Les souches de *P. anserina* *mat+* et *mat-* utilisées dans cette étude dérivent de la souche *S* (Rizet 1952). C'est la souche *S mat+* qui a été utilisée pour la séquence du génome (Espagne et al. 2008). Les conditions de culture de *P. anserina* sont décrites dans le site : <http://podospora.igmors.u-psud.fr/> . Les mutants *fprI*⁻, *fmrI*⁻ et *fmrI*⁻*smr2*⁻ utilisés dans les analyses transcriptomiques ont été obtenus par El Khoury et collaborateurs (El-Khoury et al. 2008). Les gènes *FPR1* (dans l'allèle *mat+*), *FMRI* et *SMR2* (dans l'allèle *mat-*) contiennent plusieurs mutations non-sens au début de la phase codante. Les délétions des gènes ont été effectuées dans la souche *S* mutée pour le gène *PaKu70* (El-Khoury et al. 2008) qui est nécessaire à la recombinaison non homologue (Ninomiya et al. 2004).

La souche d'*Escherichia coli* *DH5α* (Hanahan 1983) (de génotype *F*⁻, *supE44*, *ΔlacU169*, *endA1*, *hsdR17*, *thi-1*, *gyrA96*, *λ*⁻) a été utilisée pour l'amplification des plasmides.

La souche de *S. cerevisiae* *W303* (Fan et al. 1996) (de génotype *MATa/MATα leu2-3,112 trp1-1 can1-100 ura3-1 ade2-1 his3-11,15*) a été utilisée pour la construction des cassettes de délétion.

2. Préparation du matériel biologique et extraction des ARN

2.1 Cultures et croisements

P. anserina est cultivé sur des boîtes de Petri contenant du milieu minimum M2 (<http://podospora.igmors.u-psud.fr/methods.php>). Les boîtes sont inoculées avec neuf implants et incubées à la lumière (facteur indispensable pour la reproduction sexuée) dans une chambre humide à 27 C°. Afin de faciliter le prélèvement du mycélium, un support de cellophane (Bio Rad, cat # 1650963) ou de toile à bluter est déposé sur le milieu. Le champignon commence à dégrader la cellophane Bio Rad après 120 heures de croissance. Pour les incubations de plus de 120 heures le support toile à bluter est privilégié. Les conditions de cultures et les souches de *P. anserina* utilisée dans l'étude de la reconnaissance internucléaire et l'expérience de la RT-qPCR de ΔPa_1_13940 sont décrites dans le tableau 4 ci-dessous.

Tout le matériel biologique (mycélium et périthèces) est prélevé par grattage à l'aide d'une lamelle, essoré entre deux feuilles de papier absorbant, pesé et congelé dans l'azote liquide.

$\Delta Pa_6_4110 mat-$	96h de croissance végétative	La souche $\Delta Pa_6_4110 mat-$ est cultivée sur un milieu M2 couvert de cellophane.	5	RT-qPCR
Souches analysées	Heures de prélèvement	Description des conditions de culture	Nombre de réplicats	expérience
$\Delta Pa_7_7190 mat+$	96h de croissance végétative	La souche $\Delta Pa_7_7190 mat+$ est cultivée sur un milieu M2 couvert de cellophane.	5	RT-qPCR
$\text{\textcircled{m}} fpr1^- X \text{\textcircled{f}} WT mat-$	30h après fécondation	La souche sauvage $mat-$ est cultivée 96 h sur un milieu M2 couvert de toile à bluter, puis fécondée par les microconidies de la souche $fpr1^-$ cultivée sur un milieu M2 couvert de cellophane.	4	Reconnaissance internucléaire
$\Delta Pa_7_7190 mat-$	96h de croissance végétative	La souche $\Delta Pa_7_7190 mat-$ est cultivée 6 jours sur milieu M2+hygromycine.	5	RT-qPCR
$\Delta Pa_1_14230 mat+$	96h de croissance 30h après fécondation	La souche $\Delta Pa_1_14230 mat+$ est cultivée sur un milieu M2 couvert de toile à bluter, puis fécondée par les microconidies de la souche $fmr1^- smr2^-$ cultivée 6 jours sur milieu M2+hygromycine.	4	RT-qPCR Reconnaissance internucléaire
$\Delta Pa_1_14230 mat-$	96h de croissance végétative	La souche $\Delta Pa_1_14230 mat-$ est cultivée 6 jours sur milieu M2+hygromycine.	5	RT-qPCR
$\text{\textcircled{m}} WT mat+ X \text{\textcircled{f}} WT mat-$	30h après fécondation	La souche sauvage $mat-$ est cultivée 96 h sur un milieu M2 couvert de toile à bluter, puis fécondée par les microconidies de la souche sauvage $mat+$ cultivée 6 jours sur milieu M2.	3	Reconnaissance internucléaire
$\text{\textcircled{m}} WT mat- X \text{\textcircled{f}} WT mat+$	30h après fécondation	La souche sauvage $mat-$ est cultivée 96 heures sur un milieu M2 couvert de toile à bluter, puis fécondée par les microconidies de la souche sauvage $mat+$ cultivée 6 jours sur milieu M2.	2	Reconnaissance internucléaire
$WT mat+$	96h de croissance végétative	La souche sauvage $mat+$ est cultivée sur un milieu M2 couvert de cellophane.	5	RT-qPCR
$WT mat-$	96h de croissance végétative	La souche sauvage $mat-$ est cultivée sur un milieu M2 couvert de cellophane.	5	RT-qPCR
$\Delta Pa_1_13940 mat+$	96h de croissance végétative	La souche $\Delta Pa_1_13940 mat+$ est cultivée sur un milieu M2 couvert de cellophane.	5	RT-qPCR
$\Delta Pa_1_13940 mat-$	96h de croissance végétative	La souche $\Delta Pa_1_13940 mat-$ est cultivée sur un milieu M2 couvert de cellophane.	5	RT-qPCR
$\Delta Pa_6_4110 mat+$	96h de croissance végétative	La souche $\Delta Pa_6_4110 mat+$ est cultivée sur un milieu M2 couvert de cellophane.	5	RT-qPCR

Tableau 4 : Les conditions de culture des souches de *P. anserina* destinées à l'extraction des ARN.

2.2 Extraction des ARN

Les extractions d'ARN ont été réalisées à l'aide du kit RNeasy Mini Kit de Qiagen (Qiagen, Hilden, Germany) à partir de 60 à 80 mg de matériel biologique. Le protocole du fournisseur a été optimisé pour produire des ARN de qualité adaptée aux expériences de microarray et une digestion à la DNase a été ajoutée systématiquement (Annexe 2).

La quantité, la qualité et la pureté des ARN ont été contrôlées par plusieurs approches complémentaires :

- 1) Vérification de la qualité des ARN totaux par électrophorèse sur gel d'agarose : deux bandes correspondant aux ARNs ribosomiques 18S et 28S doivent être observées avec un rapport 1:2. Une dégradation partielle se traduit par l'apparition de fragments de faible poids moléculaire.
- 2) La quantité de produit initiale est calculée avec le spectrophotomètre NanoDrop ND-1000 (Nanodrop Technologies, Wilmington, USA), permettant par la même occasion de vérifier la pureté des ARN par les rapports d'absorbance A260/A280 et A260/A230. Ces rapports doivent être supérieurs à 1.8.
- 3) L'intégrité des préparations d'ARN est vérifiée avec le Bioanalyzer Agilent 2100 (Agilent, Santa Clara, USA). Les ARN sont déposés sur une puce Agilent RNA 6000 Nano LabChip et analysés avec le Bioanalyzer. Cette technique permet ainsi de détecter une éventuelle dégradation des ARN, mais elle permet également de vérifier qu'il n'y a pas de contamination par de l'ADN génomique.

3. Microarray et analyse des données

Le marquage et l'hybridation se font selon la procédure « Two-Color Microarray-Based Gene expression Analysis » (version 5.0 February 2007, Agilent, Santa Clara, USA) qui comprend notamment les étapes suivantes :

3.1 Etapes de marquage

Les ARNm (environ 1% de la totalité des ARN totaux extraits) permettent d'obtenir des ADNc par transcription inverse en présence d'une reverse transcriptase (MMLV-RT), d'un oligonucléotide d'amorces poly dT et d'un promoteur T7 polymérase. Ces ADNc vont servir par la suite comme matrice pour synthétiser les ARNc en utilisant la T7 RNA polymérase qui va incorporer simultanément le CTP couplé à un fluorochrome soit Cyanine 3 (Cy3) pour les échantillons testés soit le Cyanine 5 (Cy5) pour la référence (Annexe 1).

Les ARNc marqués sont ensuite purifiés sur colonne (Qiagen, Hilden, Germany) afin d'éliminer les colorants non incorporés qui risquent d'augmenter le bruit de fond. Les ARNc

marqués sont mélangés à 350 µl du tampon RLT (RNeasy Lysis Buffer) (Qiagen, Hilden, Germany), la préparation est ensuite déposée sur la colonne pendant 2 min puis centrifugée 30 sec à 13000 rpm à 4 C°. Après l'élimination du filtrat, la colonne est lavée avec 500 µl du tampon de lavage additionné d'éthanol 96 % (RPE) et centrifugée 1 min à 13000 rpm à 4 C°. Cette opération est réalisée 2 à 3 fois afin d'augmenter la qualité des cibles. La colonne est séchée pendant 2 min à 13000 rpm à 4 C° en utilisant un nouveau tube collecteur ; on procède à deux éluions avec 30 µl d'eau (nucléase free) et 2 min de centrifugation à 13000 rpm à 4C°.

3.2 Qualification des cibles

Une fois la synthèse et le marquage des cibles réalisés, une vérification de la qualité des produits d'amplification obtenue est nécessaire avant l'étape d'hybridation sur une lame:

- La quantité des produits d'amplification ainsi que le taux d'incorporation des Cyanines sont calculés à l'aide du spectrophotomètre NanoDrop. Un spectre d'absorption entre 200 nm et 700 nm permet de doser les cibles marquées: 260 nm pour les ARNc, 550 nm pour le Cy3 et 650 nm pour le Cy5
- Le profil électrophorétique des cibles d'ARNc réalisé soit au Bioanalyseur soit sur gel d'agarose permet de vérifier l'homogénéité d'incorporation des fluorochromes et de visualiser les fluorochromes non incorporés.

3.3 Etapes d'hybridation

La puce à ADN de *P. anserina* utilisée lors de ce travail est composée de polymères de 60 nucléotides (60 mers) complémentaires de régions spécifiques aux 10546 CDS identifiées dans le génome. Chaque CDS est représentée quatre fois sur une array et chaque lame présente quatre arrays (format 4X44K, Agilent Technologies). La conception et la fabrication de ces puces a été décrite dans l'article (Bidard et al. 2010).

Après avoir déterminé le volume des ARNc marqués nécessaire pour l'hybridation (825 ng), le couple de cibles Cy3 et Cy5 est mélangé. Le tampon de fragmentation est ajouté et incubé à 60 C° pendant 30 min. La réaction est arrêtée par l'ajout du tampon d'hybridation. L'hybridation est effectuée pendant 17 h à 65 C° dans un four à rotation (10 rpm). Pour chaque lame il faut vérifier la circulation du mélange d'hybridation et l'absence des bulles d'air stationnaire. Les lames sont ensuite lavées successivement dans le bain n°1 [Gene expression Wash buffer1, 1 min à température ambiante], le bain n°2 [Gene expression Wash buffer 2, 1 min à 37 C°] séchées par centrifugation pendant 20 sec et maintenues à l'obscurité avant d'être placées dans le scanner pour la lecture.

3.4 Scan, analyses des images et extraction des données brutes

La première étape d'acquisition des données est l'obtention des images par la détection de la fluorescence émise à la surface de la puce après excitation des fluorochromes. Cette étape réalisée à l'aide du scanner Agilent Microarray s'effectue en deux scans successifs avec des niveaux différents d'amplification des photomultiplicateurs (PMT (Photo Multiplier Tube) 100 et 10 %) permettant d'augmenter la gamme dynamique des intensités.

Afin de vérifier la qualité et la propreté des lames, les images obtenues sont visualisées grâce à un logiciel d'acquisition des images Feature Extraction (FE v9.5.3, Agilent Technologies) qui utilise le protocole GE2-v4_95_Feb07. Ce logiciel permet par la suite la conversion de l'intensité de la fluorescence en données numériques, l'identification et filtrage des spots et l'estimation du bruit de fond. Un contrôle qualité est effectué à partir des données obtenues par la vérification de la dynamique du signal, les contrôles de calibration, l'équilibre entre signaux rouge et vert et la variation des intensités de la référence dans chaque array.

Les images obtenues sont colorées artificiellement, un spot de couleur verte indique un gène dont le niveau d'expression est plus élevé dans l'échantillon marqué avec le Cy3 que celui marqué avec le Cy5, et inversement. Le spot apparaît jaune lorsque le gène est exprimé de manière identique dans les deux échantillons comparés.

3.5 Normalisation des données

La normalisation des données est réalisée par la méthode Lowess (Locally weighted scatter plot smoothing). Cette méthode permet de prendre en compte la non-linéarité existant entre les signaux rouges et verts. Pour chaque valeur A_i (l'abondance relative de transcrits) ($1/2 \log (R*V)$) une correction est calculée en tenant compte d'un ensemble de point situé dans le voisinage de A_i , une courbe de normalisation ajustée est calculé par régression linéaire locale (Dietz et al.).

3.6 Analyse différentielle

Le rapport d'intensité obtenue dans les deux conditions doit nous permettent d'identifier les gènes différentiellement exprimés. Toutefois, une analyse statistique (t-test) robuste est nécessaire pour connaître le niveau de confiance accordé à chacun de ces résultats. Cette étape est effectuée à l'aide du logiciel MAnGO (Marisa et al. 2007). Dans l'analyse, l'approche Bayésienne est employée pour réaliser un t-test pour chaque gène, l'hypothèse nulle "La différence d'expression entre les deux souches est nulle" est testée contre l'hypothèse alternative "La différence d'expression entre les deux souches est non nulle". Le test permet d'obtenir un niveau de confiance (p-value) sur le résultat. Par la suite la méthode

d'ajustement des p-values, FDR (False Discovery Rate) de Benjamini et Hochberg peuvent être utilisées. L'analyse différentielle permet d'obtenir pour chaque gène des informations qui permettront d'évaluer la fiabilité des résultats. Les critères les plus importants sont :

- la p-value ajustée (la significativité statistique du signal par rapport au bruit de fond)
- le niveau de variation d'expression ou Fold Change (FC=intensité du rouge/intensité de vert)
- l'abondance relative des transcrits A ($A = \frac{1}{2} \log(\text{intensité de Rouge} * \text{intensité du vert})$).

4. Annotation des gènes et recherche des orthologues

Les séquences protéiques sont collectées à l'aide de leur numéro "Pa" sur le site <http://podospora.igmors.u-psud.fr/padb/index.php>. La fonction de chaque protéine est ensuite recherchée sur la banque non redondante NCBI avec BLASTp <http://blast.ncbi.nlm.nih.gov/>.

Les gènes à HMG box de *P. anserina* sont annotés par analyse comparative aux gènes HMG box connus chez *N. crassa*

<http://www.broadinstitute.org/annotation/genome/neurospora/Pfam.html>.

La présence des orthologues des gènes étudiés chez d'autres champignons est examinée par FUNGIpath (Grossetete et al. 2010) sur le site <http://embg.igmors.u-psud.fr/fungipath/>.

5. Stratégies de délétion

La construction des cassettes de délétion est réalisée selon deux techniques différentes :

La première méthode consiste à cloner le gène de résistance à l'hygromycine (*hph*) dans un plasmide linéaire contenant les régions 3' et 5' du gène cible. Le plasmide est issu de la banque plasmidique fabriquée dans le projet de séquençage du génome de *P. anserina*. Il porte un insert génomique de 3 kb contenant la CDS à invalider et le gène de résistance à l'ampicilline. Un plasmide linéaire sans le gène cible est généré par PCR en utilisant des amorces complémentaires aux séquences bordant la CDS en 5' et 3'. Le site *AscI* est ajouté en 5' des 2 amorces. La cassette *hph* est amplifiée par PCR à partir du vecteur pBcHygM avec le couple d'amorce 5_CPC1_Mlu (5'- gcgacgcgtccgagatgcgccgctgc-3') et 3_Trpc_Mlu (5'- gtcacgcgtagaggatcctctagcta-3'), puis le produit PCR résultant est digéré par l'enzyme de restriction *MluI* permettant la ligation dans le plasmide linéaire préalablement digéré par *AscI*. Le plasmide pBCHygM est un dérivé de pBC-hygro (Silar 1995) qui a perdu le site *MluI* dans le terminateur *trpC*.

La deuxième technique est fondée sur la méthode de délétion adoptée chez *Neurospora crassa* (Colot et al. 2006) qui utilise la propriété de clonage par recombinaison chez la levure. Les régions 3' et 5' adjacentes au gène cible, la cassette *hph* et le vecteur linéaire sont assemblés dans la levure grâce son système de recombinaison homologue, cette technique comprend plusieurs étapes détaillées dans l'annexe 3:

- 1- Amplification de 1 kb des régions 3'et 5' adjacentes au gène cible avec deux couples d'amorces chimères 5f+5r et 3f+3r: 5r et 3f contiennent 29 nucléotides homologues à la cassette *hph* dans leur régions 5' alors que 5f et 3r possèdent un site *AscI* et 29 nucléotides homologues au vecteur de levure dans la région 5'. La PCR est réalisée avec la polymérase *Pfu* au lieu de la *LA Taq* utilisée dans le protocole décrit par Colot et collaborateurs (Colot et al. 2006) afin de minimiser les erreurs d'amplification.
- 2- Amplification de la cassette *hph* à partir du plasmide pMOcosX grâce au couple d'amorces *hphf* et *hphr*.
- 3- Linéarisation du vecteur de levure (pRS426) par digestion enzymatique.
- 4- Cotransformation de levure (W303) par les régions 3' et 5' adjacentes au gène cible, la cassette *hph* et le plasmide pRS426 linéaire.
- 5- Extraction du vecteur recombinant et validation de la structure par PCR.
- 6- Amplification du vecteur recombinant dans *E. coli* au lieu d'une amplification par PCR utilisée dans le protocole décrit par Colot et collaborateurs (Colot et al. 2006) afin d'éviter les mutations engendrées par la PCR .
- 7- Extraction du vecteur recombinant et linéarisation de la cassette de délétion par *AscI*.
- 8- Transformation de *P. anserina* par la cassette de délétion. Le protocole de transformation est décrit sur le site <http://podospora.igmors.u-psud.fr/methods.php> . Les transformant résistants à l'hygromycine sont sélectionnés et analysés par Southern Blot. Les souches KO sont ensuite purifiées par croisement.

6. Analyses phénotypiques

Plusieurs caractéristiques phénotypiques ont été examinées chez tous les mutants à différents stades de développement

- La fertilité des mutants a été étudiée en croisements hétérozygotes et homozygotes : différenciation des organes femelles, la fécondation, développement des périthèces et production des ascospores. Une estimation de la production de microconidies et de leur pouvoir fécondant a été réalisée pour chaque mutant selon la méthode suivante : les microconidies sont collectées par lavage de la surface d'une culture de 7 jours en M2 gélosé avec 2 ml d'eau stérile ; 1 ml d'une dilution 1/500 est utilisée pour féconder une culture

sauvage de type sexuel complémentaire ; le nombre de périthèces obtenus est comparé au témoin réalisé dans les mêmes conditions avec une culture sauvage. La répartition des périthèces sur un thalle en croissance de chaque mutant a été déterminée en inoculant un broyat de mycelium de souches *mat+* et *mat-* du mutant au centre d'une boîte de Petri contenant du milieu M2.

- La croissance et la morphologie du mycélium de chacun des mutants sont analysées par comparaison avec la souche sauvage. Dans le milieu minimum standard la source de carbone est de la dextrine, un mélange de polysides issus de l'hydrolyse de l'amidon. La croissance et la fertilité sont également testées sur papier Whatmann constitué de cellulose, une source de carbone plus complexe.

- La sensibilité aux inhibiteurs suivants a été analysée: le calcofluor qui agit sur la paroi, le fludioxonil qui agit sur l'osmorégulation, le SDS qui agit sur la membrane plasmidique et la caféine qui est un inhibiteur de phosphodiesterase.

7. Séquençage des souches

7.1 Cultures

Les souches *fmrI⁻* (*mat-*), *fprI⁻* (*mat+*) et *wt mat-* sont cultivées sur rond de cellophane Bio Rad en milieu minimum M2 pendant 3 jours à 27 C°. Le mycélium est prélevé par grattage à l'aide d'une lamelle.

7.2 Extraction de l'ADN

1 à 3 g de mycélium est broyé rapidement dans un mortier contenant de l'azote liquide. Le broyat est repris dans 1 volume de tampon CTAB 2X préchauffé à 65 C° auquel on ajoute un volume de chloroforme. Après centrifugation 5 min à 4000 rpm la phase aqueuse est prélevée et on y ajoute 1/5 volume de tampon CTAB 5X puis 1 volume de chloroforme. Le tout est mélangé délicatement et centrifugé 5 min à 4000 rpm. La phase aqueuse contenant l'ADN est précipité par ajout de 1 volume de tampon CTAB de précipitation. L'ensemble est incubé 1 h à température ambiante, centrifugé 5 min à 3000 rpm, le culot d'ADN est repris dans 2 ml de TE (Tris EDTA) et 1 M NaCl. L'ADN est à nouveau précipité en ajoutant 2 volumes d'éthanol 95 %. La «pelote» formée est lavée dans de l'éthanol 70 % puis séchée et reprise dans 4 ml de TE.

CTAB 2X	2% de cetyltriméthylammonium bromide 100 mM Tris (pH8) 20 mM EDTA (pH8) 1.4 M NaCl 1% Polyvinyl pyrrolidone 40
---------	--

CTAB 5X	5% de CetylTrimethylAmmonium Bromide 0.7 M NaCl
CTAB précipitation	1% de CetylTrimethylAmmonium Bromide 50 mM Tris (pH8) 10 mM EDTA (pH8)

7.3 Purification de l'ADN en gradient de chlorure de césium

Cette technique réalisée à partir de l'ADN extrait précédemment permet de séparer l'ADN génomique de l'ADN mitochondrial.

On rajoute 1,15 g de chlorure de césium (CsCl) et 4 µl de DAPI par ml de solution d'ADN (la densité du mélange doit être $d=1.65$). Le mélange est transféré dans des tubes QuickSeal (Beckman). Les tubes sont scellés puis centrifugés à 50000 rpm pendant 20 h à 20 C°. A l'aide des UV on visualise détecte la bande supérieure, qui correspond à l'ADN mitochondrial, et la bande inférieure, qui correspond à l'ADN nucléaire. Cette dernière est prélevée à l'aide d'une seringue. Le DAPI est éliminé par addition de 1ml de solution d'alcool isoamylique saturé en CsCl à 1,15 g/ml, cette opération est répétée plusieurs fois jusqu'à disparition de la fluorescence due au DAPI. Les échantillons sont ensuite dialysés à 4 C° pendant 72 h pour éliminer le chlorure de césium.

7.4 Construction des banques d'ADN et séquençage

Les trois banques *WT mat-*, *fprI⁺mat+* et *fmrI⁺mat-* ont été construites grâce au protocole Illumina décrit dans l'annexe 4.

L'ADN génomique (10 µg) est fragmenté par nébulisation à la plateforme de séquençage à haut débit à Gif sur Yvette (IMAGIF). Cette technique a permis d'obtenir des fragments inférieurs à 800 pb. Les fragments d'ADN sont réparés et des séquences adaptatrices sont ajoutées à leurs extrémités. Le mélange réactionnel est mis à migrer par électrophorèse et les fragments d'environ 300 pb sont purifiés. L'enrichissement en fragments possédant les bonnes séquences adaptatrices est réalisé par PCR en ajoutant des index qui vont servir à trier les souches pendant l'analyse bioinformatique.

Un aliquot des fragments purifiés est cloné dans le vecteur de clonage PCR4blunt-TOPO du kit de clonage Zero Blunt TOPO d'Invitrogen. Après transformation dans des bactéries compétentes, les clones positifs ayant intégrés le plasmide recombinant ont été sélectionnés par PCR. L'ADN plasmidique des clones positifs a été ensuite isolé par purification grâce au kit QIAprep® Spin Miniprep (QIAGEN) selon le protocole fourni avec le kit. Finalement la séquence du plasmide purifié a été obtenue et vérifiée par séquençage effectué par le service Single Express de Beckman Coulter Genomics.

La banque est validée si au moins 7 clones sur les 10 analysés ont les deux adaptateurs, l'index exact (Annexe 4) et un fragment d'ADN de *P. anserina* d'environ 300 pb. Après cette confirmation de la structure des inserts, le séquençage des trois souches (*fprI*, *fmrI* et *wt mat*-) est réalisé à la plateforme de Séquençage à haut débit IMAGIF. Les alignements du génome de chaque souche mutante avec le génome de la souche de référence ont été effectués par Claire Kuchly (plateforme eBio, plate-forme Bioinformatique de l'Université Paris Sud).

8. Acquisition, analyse et alignement des séquences

Les séquences d'acides aminés des protéines à domaine HMG sont collectées à partir des sites spécifiques de chaque espèce (Annexe 5) ou à partir de NCBI <http://blast.ncbi.nlm.nih.gov/Blast.cgi>, ces séquences sont premièrement analysées par pfam <http://pfam.sanger.ac.uk/> afin d'identifier leur domaine HMG puis alignées par clustalW <http://www.ebi.ac.uk/Tools/msa/clustalw2/> et MEGA4 (Tamura et al. 2007) pour sélectionner précisément les 80 acides aminés du domaine HMG. Les alignements finaux sont visualisés par JALVIEW (Waterhouse et al. 2009). La localisation cellulaire des protéines HMG est réalisée sur <http://wolfsort.org/>.

9. L'analyse phylogénétique

Les analyses phylogénétiques ont été effectuées par la méthode TREEFINDER (Jobb et al. 2004) selon la démarche décrite dans l'article (Martin, 2010).

10. PCR quantitative

10.1 RT-PCR

Les ARN totaux ont été extraits suivant le protocole indiqué en Annexe 2, puis reverse transcrit avec la Superscript III (Invitrogen, Carlsbad, USA) en utilisant des oligo dT suivant les indications du fournisseur.

10.2 Choix des amorces

Les amorces spécifiques pour la PCR quantitative ont été dessinées pour chacun des gènes étudiés à l'aide du logiciel Primer Premier 5 (Premier Biosoft). Des amorces spécifiques du transcrit sont dessinées à cheval sur deux exons consécutifs quand la structure du gène le permet. Pour les deux gènes sans introns *mfm* et *mfp*, nous avons inclus un contrôle avec des ARN non reverse transcrits afin de déterminer le niveau de contamination des préparations d'ARN par de l'ADN génomique.

10.3 Contrôle de l'efficacité et spécificité de la PCR

L'efficacité est calculée par la méthode des dilutions sérielles (1/10, 1/100 et 1/1000) faites à partir d'un mélange des ADNc de toutes les souches dans toutes les conditions. La régression des points sur la droite $Cq = f(\text{conc})$ doit être > 0.99 et l'efficacité $> 80 \%$. La spécificité des amorces est vérifiée par une courbe de fusion réalisée en fin de PCR ; la présence d'un seul pic indique que l'amplification est spécifique.

10.4 RT-qPCR

Le mélange réactionnel est composé d'ADNc (2 μl), de 2X FastStart Universal SYBR Green (5 μl) (Roche Diagnostics), d'amorces sense et antisense (1 μl) 5 pmol/ μl et d'eau (2 μl) « RNase-free ». Les conditions PCR sont celles préconisées par Roche Diagnostics : dénaturation 10 min à 95 C°, 45 cycles (10 sec à 95 C°, 30 sec à 60 C° et lecture des plaques), la réaction est terminée par une courbe de fusion (15 sec à 95 C°, refroidissement rapide jusqu'à 65 C°, puis montée progressive lente de la température jusqu'à 95 C°). Chaque plaque comporte les échantillons d'ADNc examinés pour chaque gène étudié (pour *mfm* et *mfp* ont ajoutés aussi les échantillons d'ARN non reverse transcrits), les trois dilutions sérielles pour vérifier l'efficacité et un contrôle négatif. Toutes les conditions testées sont représentées par au moins cinq replicats biologiques. Les gènes de normalisation sont choisis avec le logiciel geNorm parmi dix gènes de ménage : *AS1* (Pa_1_16650, protéine ribosomal S12), *CIT1* (Pa_3_6780, citrate synthétase), *GPD* (Pa_3_5110, glyceraldéhyde phosphate déshydrogénase), *H2A* (Pa_5_5390, histone H2A), *LEU1* (Pa_7_8770, 3-isopropylmalate dehydrogenase), *PAH1* (Pa_2_6460, facteur de transcription à homéodomaine), *PDF2* (Pa_7_6690, sous unité régulatrice d'une phosphatase), *TIP41* (Pa_7_8490, phosphatase activator), *TBP* (Pa_4_8980, TATA box binding protein), *UBC* (Pa_4_7790, ubiquitine). La normalisation et l'analyse statistique sont effectuées par REST 2009 (Qiagen, Hilden, Germany).

11. Annotation structurale de *Pa_1_13940*

Pour déterminer la position des introns dans le gène *Pa_1_13940*, les amorces 13940_F 5'-gttacgtgctgcttcaatg-3' et 13940_R 5'-ctgtgctgctcaaccgaaag-3' ont été utilisées pour l'amplification des ADNc de *Pa_1_13940* à partir des ARN extraits des souches sauvages *mat+* et *mat-*. Les fragments amplifiés ont été purifiés puis séquencés.

**Chapitre III : Article "Genome-Wide Gene
Expression Profiling of Fertilization Competent
Mycelium in Opposite Mating Types in the
Heterothallic Fungus *Podospora anserina*"**

1. Introduction

Le développement sexué des ascomycètes heterothalliques est initié par l'interaction entre deux cellules ou organes sexuellement compatibles. Cette compatibilité sexuelle est sous le contrôle des types sexuels déterminés par les gènes présents au locus *mat*. La détermination du type sexuel a été largement étudiée chez les levures *S. cerevisiae* et *S. pombe*. Chez ces deux espèces, les gènes *mat* codent des facteurs de transcription qui régulent l'expression des gènes impliqués dans la détermination du type cellulaire, en particulier les systèmes phéromones/récepteurs. Chez ces deux levures, les voies de régulation apparaissent de prime abord très différentes. Un des facteurs contribuant à cette différence est certainement la perte d'un régulateur à domaine HMG chez *S. cerevisiae*, cette perte étant compensée par la mise en place d'un nouveau circuit de régulation (voir chapitre I paragraphe 2.3.2). Pour les champignons filamenteux, les analyses génétiques effectuées chez *P. anserina* indiquent que les systèmes de régulation des types sexuels sont différents de ceux des levures, mais on ne disposait d'aucune vue complète du rôle des gènes *mat* dans la détermination du type sexuel, comme cela est le cas chez les levures. Pour apporter cette vue d'ensemble, une étude globale des gènes régulés par les facteurs de transcription MAT a été réalisée chez le champignon filamenteux *P. anserina* avec une approche microarray.

P. anserina est un champignon hétérothallique, il possède deux types sexuels *mat+* et *mat-* contrôlés chacun par une séquence spécifique. L'idiomorphe *mat+* contient un seul gène : *FPR1*; l'idiomorphe *mat-* contient trois gènes: *FMRI*, *SMR1* et *SMR2* (voir chapitre I paragraphe 2.2.3.2). Pendant la fécondation, *FPR1* (*mat+*) et *FMR1* (*mat-*) contrôlent la reconnaissance entre le gamète mâle et l'organe femelle (Coppin et al. 1993). Dans une souche *mat+*, *FPR1* agit comme activateur des fonctions spécifiques *mat+* et comme répresseur des fonctions spécifiques *mat-*, alors que chez une souche *mat-*, *FMR1* agit comme activateur des fonctions spécifiques *mat-*, et à l'aide de *SMR2*, il réprime les fonctions spécifiques *mat+* (Arnaise et al. 2001). Les gènes *SMR1* et *SMR2* ne sont pas nécessaires pour la fécondation, mais pour des étapes ultérieures (Debuchy et al. 1993). Des études précédentes ont identifié seulement deux gènes cibles de *FPR1* et *FMR1*. Ces gènes cible codent les précurseurs des phéromones MFM et MFP qui sont spécifiquement exprimés dans *mat-* et *mat+* respectivement. Depuis aucun autre gène cible n'a été identifié.

2. Résumé des résultats et additions non publiées

La comparaison du profil transcriptomique des souches *mat+* et *mat-* compétentes pour la fécondation a permis d'identifier 157 gènes. L'analyse des résultats de microarray des souches mutantes *fprI* et *fmrI* a révélé que ces 157 gènes cibles peuvent être soit réprimés,

soit activés par FMR1 et/ou FPR1, ou sous le contrôle du polymorphisme lié au locus *mat* (Figure 20). Ces expériences ont aussi détecté 10 gènes activés ou réprimés au même niveau dans *mat+* et *mat-* (Figure 21).

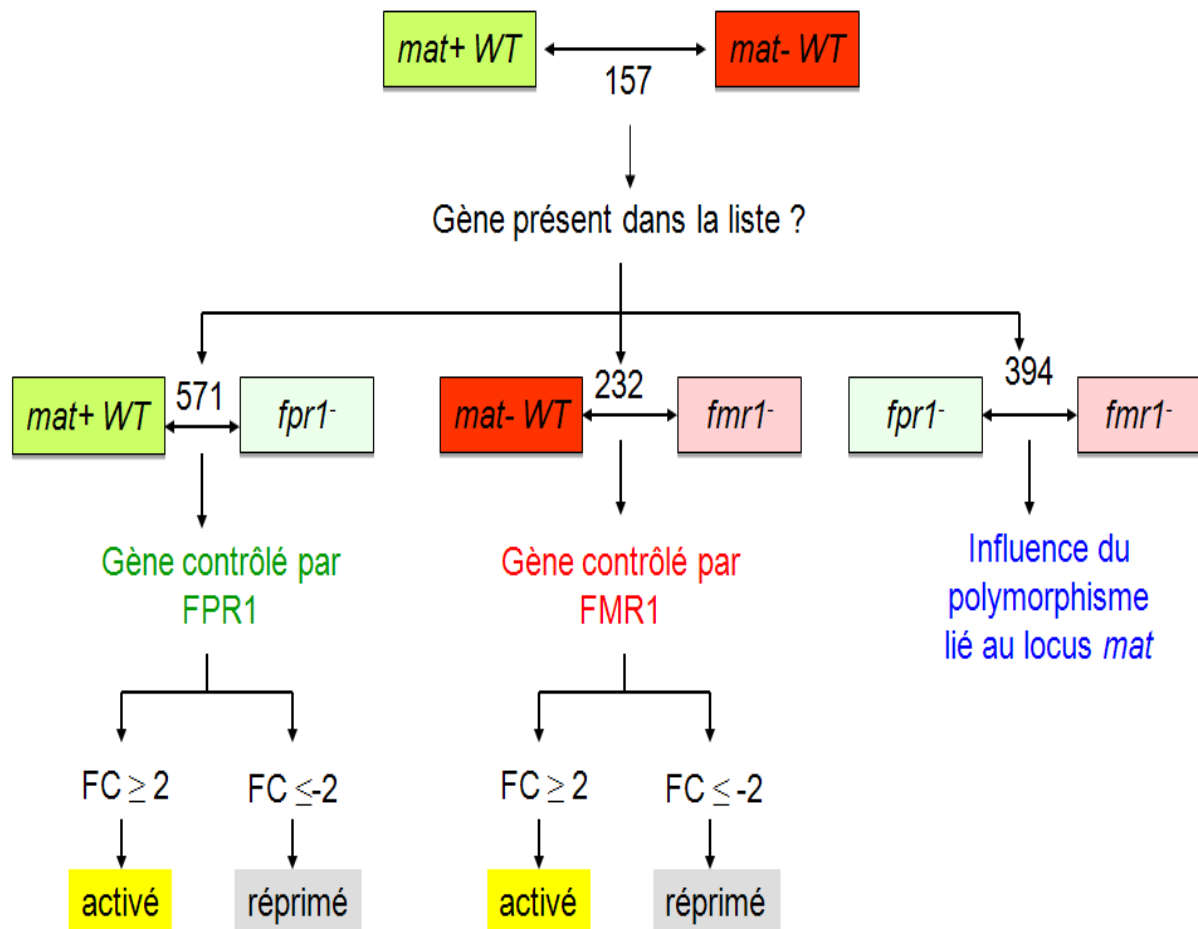


Figure 20: Organigramme de la classification des gènes différentiellement exprimés entre les souches *mat+* et *mat-*. Parmi les 157 gènes différentiellement exprimés entre le sauvage *mat+* et le sauvage *mat-*, les gènes contrôlés par FPR1 ou FMR1 sont les gènes différentiellement exprimés dans la comparaison *WT mat+/fpr1-* ou dans la comparaison *WT mat-/fmr1-* respectivement. Les gènes activés par FPR1 ou FMR1 sont les gènes dont les $FC \geq 2$ dans les comparaisons *WT mat+/fpr1-* ou *WT mat-/fmr1-* respectivement. Les gènes réprimés par FPR1 ou FMR1 sont les gènes dont les $FC \leq -2$ dans les comparaisons *WT mat+/fpr1-* ou *WT mat-/fmr1-* respectivement. La comparaison du mutant *fpr1-* au mutant *fmr1-* révèle l'existence des gènes dont l'expression est sous l'influence du polymorphisme lié au locus *mat*, ces gènes ne sont pas régulés par FPR1 et FMR1.

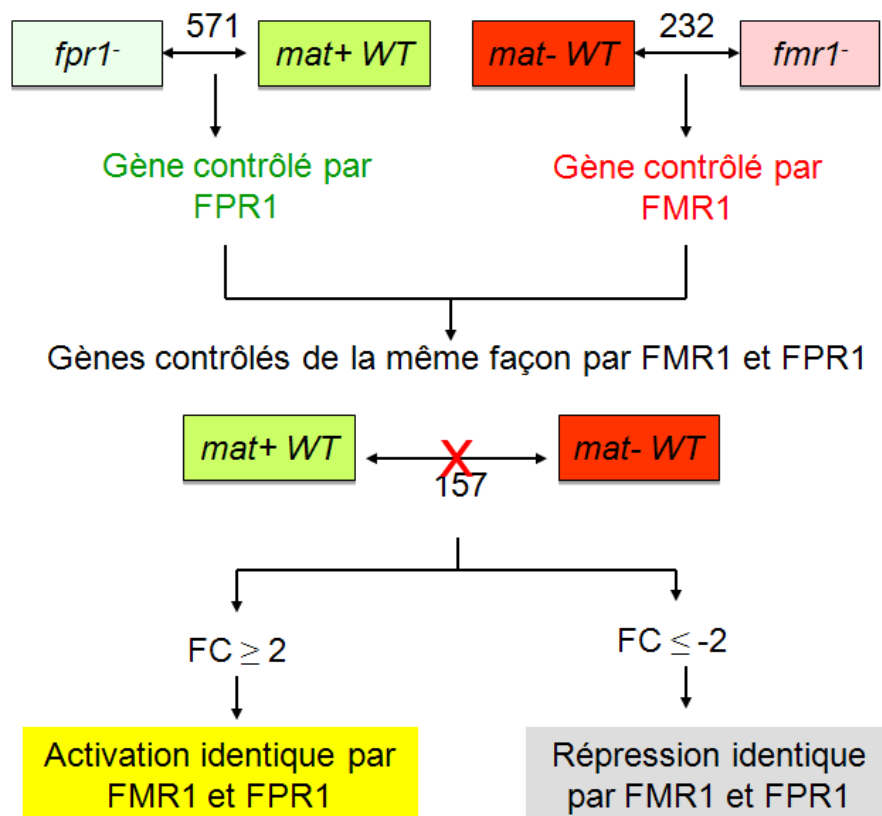


Figure 21: Organigramme de la classification des gènes contrôlés de la même façon par FPR1 et FMR1. Un gène contrôlé au même niveau par FPR1 et FMR1 ne sera pas différenciellement exprimé dans la comparaison du sauvage *mat+* au sauvage *mat-*. Par contre, il est différenciellement exprimé dans les comparaisons du sauvage *mat+* au mutant *fpr1-* et dans la comparaison du sauvage *mat-* au mutant *fmr1-*.

Dans l'ensemble de ces 167 gènes, nous avons déléte 32 gènes, choisis selon trois critères. Nous avons déléte systématiquement les gènes qui avaient une transcription différentielle très élevée entre *mat+* et *mat-*. Nous avons également déléte des gènes qui n'avaient pas un différentiel de transcription très élevé, mais qui étaient intéressants pour leur conservation chez d'autres champignons, ou pour leur fonction (facteurs de transcription, phosphodiesterase). Seuls les mutants des récepteurs de phéromones ont des défauts pendant la fécondation. Nous avons également identifié un gène dont la délétion est létale.

Des recherches partielles de gènes cibles de facteur de transcription MAT ont été publiées chez *Gibberella moniliformis* (Keszthelyi et al. 2007) et *Sordaria macrospora* (Pöggeler et al. 1997). Nous avons utilisé ces résultats pour la recherche d'une éventuelle conservation des gènes cibles *mat* entre ces espèces. Les comparaisons des gènes cibles des facteurs de transcription MAT de *G. moniliformis* et *S. macrospora* avec ceux de *P. anserina* révèlent un nombre significatif de gènes cibles communs entre ces espèces, mais ces gènes ont des profils transcriptomiques différents. Par exemple, le gène *crea* est sur-exprimé dans la souche *mat+* de *P. anserina* et sous-exprimé dans la souche équivalente (*mat1-2*) de *G. moniliformis*. D'autre part, en comparant les listes de gènes cibles chez *P. anserina* et *G.*

moniliformis, j'ai découvert que chez *P. anserina*, bien souvent ce n'est pas l'orthologue du gène cible de *G. moniliformis* qui est lui-même un gène cible, mais un gène voisin (résultats non publiés). J'ai examiné les 160 orthologues des gènes cibles de *G. moniliformis*, et j'ai retrouvé cette situation dans 119 cas (Annexe 6 : Recherche des gènes cibles de *P. anserina* adjacents aux orthologues des gènes cibles de *G. moniliformis*).

Au cours des expériences microarray, on a remarqué que le nombre des gènes différentiellement exprimés dans la comparaison entre la souche sauvage *mat+* et le mutant *fprI*⁻ est beaucoup plus élevé (571 gènes) que celui obtenu dans la comparaison entre les souches sauvages *mat+* et *mat-*. On a fait la même observation avec la comparaison entre la souche sauvage *mat-* et la souche mutante *fmrI*⁻ (232 gènes). Cette situation anormale pourrait s'expliquer s'il y avait beaucoup de gènes contrôlés de la même façon par les types sexuels *mat+* et *mat-*. La comparaison des mutants *fprI*⁻ et *fmrI*⁻ exclut cette hypothèse, puisque là encore, on observe un excès de gènes différentiellement transcrits (394 gènes). Une autre hypothèse serait que le processus de recombinaison homologue en contexte Δ KU lors de la préparation des mutants *fprI*⁻ et *fmrI*⁻ aurait un effet mutateur. Cette hypothèse s'appuie sur les observations allant dans ce sens de Fu et al (Fu et al. 2011). Pour tester cette hypothèse, j'ai resequencé l'ADN génomique des mutants *fprI*⁻ et *fmrI*⁻. Le séquençage de la souche mutante *fprI*⁻ (voir matériels et méthodes) n'a révélé aucune mutation dans les 571 gènes identifiés dans la comparaison *mat+*/*fprI*⁻. Une analyse similaire a été réalisée sur la comparaison *mat-*/*fmrI*⁻, dans les 232 gènes différentiellement exprimés entre *mat-* et *fmrI*⁻ aucune mutation n'a été détectée.

3. Discussion

En utilisant une approche microarray nous avons identifié des cibles potentielles des facteurs de transcription FPR1 et FMR1. Dans cette approche, le profil transcriptomique des souches *mat+* et *mat-* sauvage fait apparaître un nombre très élevé de gènes différentiellement exprimés entre les deux souches par rapport à celui obtenu chez *S. pombe* et *S. cerevisiae*. Chez *S. cerevisiae*, seulement 11 gènes cibles ont été identifiés dont six gènes spécifiques des cellules α et cinq gènes spécifiques des cellules a . Chez *S. pombe*, 16 gènes cibles ont été identifiés dont 12 gènes spécifiques des cellules *M* et quatre gènes spécifiques des cellules *P* (voir tableau 2). Chez *P. anserina* nous avons identifié 157 gènes différentiellement exprimés entre *mat+* et *mat-*. Plusieurs de ces gènes interviennent dans le métabolisme général et n'ont aucune relation avec la fécondation. De nombreux autres gènes ont une fonction inconnue, l'inactivation d'un certain nombre d'entre eux suggère qu'ils n'ont pas de fonction directe dans la fécondation. Deux hypothèses ont été proposées pour expliquer ces résultats. La première hypothèse est basée sur l'observation que *P. anserina* est une espèce pseudo-

homothallique, où les deux types de noyaux cohabitent dans le même mycelium. On peut donc supposer qu'une complémentarité métabolique s'est mise en place entre les noyaux *mat+* et *mat-* pour favoriser leur co-existence. La deuxième hypothèse est basée sur le rôle du mycelium dans l'élaboration de la fructification (Debuchy et al. 2010). Cette hypothèse suppose que le grand nombre de gènes contrôlés par les deux types sexuels pourraient préparer le développement des périthèces en mobilisant les ressources du mycelium.

Dans la liste des gènes différentiellement exprimés entre les types sexuels opposés chez *P. anserina*, *S. cerevisiae* et *S. pombe*, nous avons identifié des orthologues communs (Tableau 5):

- Les gènes codant les précurseurs de phéromones : Chez *P. anserina*, ces gènes ont les différentiels d'expression les plus élevés (144 pour *mfp* et -51 pour *mfm*). De plus, les niveaux des intensités normalisées de ces gènes (Intensité de *mfp* 10795, Intensité de *mfm*: 196734) sont comparables à ceux de gènes connus pour être fortement exprimés dans *P. anserina* (Intensité de GPD 55376, Intensité de EF1A 76733). Cette forte expression des précurseurs de phéromones peut être expliquée par le fait que les phéromones doivent être secrétées en grande quantité dans le milieu afin d'atteindre le partenaire de type sexuel opposé. Toutefois, il est aussi possible que les transcrits subissent des régulations soit traductionnelle (présence du motif conservé avant le codon d'initiation de traduction de gène *mfp*) soit post traductionnelle (au niveau des étapes de maturation).
- Les gènes codant les récepteurs de phéromones spécifiques du type sexuel opposé.
- Le gène codant le transporteur nécessaire pour la sécrétion de la phéromone hydrophobe.

Fonction de la protéine	<i>S. pombe</i>	<i>S. cerevisiae</i>	<i>P. anserina</i>
	Cellules <i>M</i>	Cellules <i>a</i>	<i>mat+</i>
Phéromone	<i>mfm1</i> <i>mfm2</i> <i>mfm3</i>	<i>MFA1</i> <i>MFA2</i>	<i>mfp</i>
Récepteur de phéromone	<i>mam2</i>	<i>STE2</i>	<i>PRE2</i>
Protéase	<i>sxa2</i>	<i>BARI</i>	-
Agglutinine	<i>SPAP11E10.02c</i>	<i>AGA2</i>	-
Farnesyltransferase α	<i>cwp1</i>	-	<i>Pa_2_9270</i>
Farnesyltransferase β	-	-	<i>Pa_4_7760</i>
Méthyltransferase	<i>mam4</i>	-	<i>Pa_7_9690</i>
Transporteur de phéromones	<i>mam1</i>	<i>STE6</i>	<i>Pa_5_11640</i>
	Cellules <i>P</i>	Cellules <i>a</i>	<i>mat-</i>
Phéromone	<i>map2</i>	<i>MFA1</i> <i>MFA2</i>	<i>mfm</i>
Récepteur de phéromone	<i>map3</i>	<i>STE3</i>	<i>PRE1</i>
Protéase	-	-	<i>Pa_6_7350</i>
Agglutinine	<i>SPBC21D10.06c</i>	<i>SAG1</i>	-

Tableau 5: Comparaison entre les gènes spécifiques impliqués dans la fécondation chez *S. pombe*, *S. cerevisiae* et *P. anserina*. Les gènes spécifiques des cellules *a* et α de *S. cerevisiae* sont décrits dans (Galgoczy et al. 2004). Les gènes spécifiques des cellules *P* et *M* de *S. pombe* sont décrits dans (Mata and Bähler 2006).

Les agglutinines sont nécessaires pour l'adhésion des cellules reproductives pendant la fécondation et jouent un rôle important à cette étape chez *S. cerevisiae* et *S. pombe*, où chaque type cellulaire exprime une agglutinine spécifique. Chez *P. anserina*, j'ai identifié une famille multigénique d'agglutinines, mais aucun de ces gènes ne présente une expression différentielle entre *mat+* et *mat-*. Plusieurs hypothèses sont possibles pour expliquer l'absence des gènes agglutinines dans les gènes différentiellement exprimés entre *mat+* et *mat-*:

- Les souches *mat+* et *mat-* utilisent les mêmes agglutinines.
- Les agglutinines sont spécifiques *mat+* ou *mat-* mais l'expression de leurs gènes n'augmente qu'au moment de l'adhésion des cellules reproductives.
- L'expression des gènes qui codent les agglutinines se fait au niveau des cellules de trichogyne où à lieu l'adhésion et on n'arrive pas à la détecter car ce tissu est minoritaire dans l'échantillon analysé.

Les cellules *a* de *S. cerevisiae* et *M* de *S. pombe* utilisent des protéases spécifiques pour dégrader les phéromones hydrophiles du type sexuel opposé. Chez *S. cerevisiae* cette dégradation est utilisée pour limiter la mort cellulaire engendré par la forte concentration des phéromones hydrophiles dans le milieu (Zhang et al. 2006). Chez *P. anserina*, on trouve un gène qui appartient à une famille multigénique codant des protéases (*Pa_6_7350* ; FC=-6.56).

Ce gène est fortement exprimé dans la souche *mat-* équivalente aux cellules α de *S. cerevisiae* et *P* de *S. pombe*. *Pa_6_7350* n'a pas d'orthologues dans *S. cerevisiae* et *S. pombe*. Si la protéase *Pa_6_7350* a la même fonction que *Sax2p* et *BAR1p*, cette protéine serait donc nécessaire pour dégrader la phéromone hydrophobe *mfp*, ce qui constituerait une différence importante par rapport aux systèmes levures. Cette inversion du système de dégradation suggère que la phéromone hydrophobe est présente en grande quantité à proximité des souches *mat-*.

Les gènes *KEX2* et *KEX1* codent des protéases qui sont impliquées dans la maturation des phéromones hydrophiles *MF α 1p* et *2* chez *S. cerevisiae* (Davey et al. 1998). Une analyse de l'expression de leur orthologue chez *P. anserina* montre que ces gènes ne sont pas différentiellement exprimés entre la souche *mat+* et *mat-*. Ces résultats suggèrent que ces gènes ne sont pas exprimés spécifiquement chez les souches *mat-* ou que leur différentiel d'expression est trop faible pour être détecté. Une situation identique est observée chez *S. cerevisiae* et *S. pombe*. J'ai également analysé l'expression des orthologues des gènes de *S. cerevisiae* qui sont nécessaires pour la maturation des phéromones hydrophobes *MFA1p* et *2*. Les gènes qui codent les transférases des chaînes aliphatiques qui sont ajoutées post-traductionnellement à *mfp* sont bien surexprimés chez *mat+* par rapport à *mat-* (*RAM1p*, farnesyltransferase β *Pa_4_7760* ; *STE14p*, protéine-S-isoprenylcysteine O-méthyltransferase *Pa_7_9690*). Le gène codant la farnesyltransferase α (*RAM2*, *Pa_2_9270*) est en dessous du cutoff pour le fold-change, mais avec une p-value très significative. En revanche, comme pour la maturation de la phéromone *mfm*, les gènes codant les protéases (*RCE1p*, *Pa_7_2280*; *AFC1p/STE24p*, *Pa_2_9310*; *STE23p*, *Pa_7_6640*) ont le même niveau d'expression dans les deux souches. Il est possible que les protéases soient déjà exprimées à un niveau suffisamment élevé dans les souches *mat+* et *mat-*, parce qu'elles ont de nombreuses fonctions en dehors de la biogenèse des phéromones. Il ne serait donc pas nécessaire de les surexprimer pour répondre à ce besoin.

La comparaison entre les résultats des études de transcriptomique réalisées chez *G. moniliformis* et *S. macrospora* avec ceux obtenus chez *P. anserina* révèle un nombre significatif des cibles communes entre ces espèces. Cependant ces gènes communs ne sont pas régulés de la même façon chez ces trois espèces. Il y a deux explications possibles à ces observations : 1) un effet métabolique, par exemple la mobilisation des nutriments pour le développement des fructifications. Un même ensemble de gène est mobilisé, mais suivant des modalités différentes d'une espèce à l'autre, pour arriver au même résultat. 2) le point important est la conservation d'un site de fixation pour les facteurs de transcription codés par les gènes *mat*, indépendamment de la fonction du gène et du fait qu'il soit activé ou réprimé.

Cette dernière hypothèse pourrait être en accord avec les résultats indiquant que 119 orthologues des cibles de MAT1-2-1 chez *G. moniliformis* possèdent au moins un gène adjacent régulé par FPR1 chez *P. anserina* (Annexe 6). Une analyse avec MEME serait nécessaire pour identifier la présence d'un éventuel site de fixation commun à tous ces gènes.

Le dernier point non élucidé de cette analyse est le problème de l'excès des gènes différentiellement transcrits chez les mutants *fprI*⁻ et *fmrI*⁻ par rapport à ce que l'on trouve entre les souches *mat+* et *mat-*. J'ai écarté l'hypothèse d'un effet mutateur du contexte génétique Δ KU qui est utilisé lors des constructions de mutants. Il reste encore une hypothèse d'artefact, liée à la sénescence de *P. anserina* (Espagne et al. 2008). En effet, on ne peut pas exclure que les souches *fprI*⁻ et *fmrI*⁻ soient engagées dans la sénescence, à des étapes différentes de ce processus accéléré par leur génotype particulier, et que les différences transcriptomiques observées révèlent les différences physiologiques liées à cet état. Pour tester cette hypothèse, il faudrait recroiser les mutants pour les régénérer et refaire les analyses transcriptomiques. Il convient de noter cependant que cette hypothèse ne remet pas en cause les conclusions essentielles de cet article, puisque il repose sur l'analyse de 157 gènes différentiellement transcrits entre deux souches *WT mat+* et *WT mat-*, qui ont été préparées pour réduire les risques du déclenchement du processus de sénescence au cours des expériences.

**Chapitre IV : Recherche transcriptomique des
gènes cibles des facteurs de transcription
contrôlant la reconnaissance internucléaire chez
*Podospora anserina***

1. Introduction

Chez tous les champignons filamenteux ascomycètes ayant un cycle sexué, la production d'une descendance normale dépend de la reconnaissance internucléaire (RI). Ce processus a été reconnu dès 1956 par Papazian, qui a proposé qu'à l'intérieur de l'organe femelle, les noyaux mâles et femelles doivent avoir des phénotypes différents déterminés par les gènes des types sexuels (Papazian 1956).

Chez *P. anserina* suite à la fécondation les noyaux d'origine mâle et femelle se divisent dans l'organe femelle et donnent naissance à des cellules plurinucléées. Dans ces cellules, les noyaux *mat+* et *mat-* migrent ensemble vers les hyphes ascogènes. Une fois dans les hyphes ascogènes les deux noyaux se divisent par mitose et maintiennent le rapport 1:1 de chaque noyau parental. Finalement les deux noyaux fusionnent et subissent immédiatement la méiose qui engendre une descendance ayant le rapport Mendélien attendu de chaque type sexuel (Debuchy 1999). Des études génétiques et moléculaires (voir chapitre I paragraphe 2.2.3) ont montré que les gènes des types sexuels *FPR1*, *FMR1* et *SMR2* contrôlent la reconnaissance internucléaire, mais les mécanismes mis en jeu au cours de cette étape sont encore inconnus. Deux modèles différents sont proposés pour expliquer l'action des facteurs de transcription codés par ces gènes *mat* au moment de la reconnaissance internucléaire :

Le modèle 1 dit d'identité nucléaire propose que dans la cellule plurinucléée (syncytium), chaque noyau *mat+* et *mat-* porte une identité propre basée sur la présence de protéines spécifiques contrôlées par les facteurs de transcription *FPR1* chez *mat+*, et *FMR1* et *SMR2* dans *mat-* (Figure 22).

- Dans un noyau *mat+*, *FPR1* active les fonctions *mat+* nécessaires à la RI, mais il réprime aussi les fonctions *mat-* impliquées dans la RI. Le noyau *mat+* exprime uniquement et spécifiquement une identité nucléaire *mat+*.
- Dans un noyau *mat-*, *FMR1* active les fonctions *mat-* nécessaires à la RI. La répression des fonctions *mat+* nécessaires pour la RI est assurée par *FMR1* et *SMR2*. Le noyau *mat-* exprime uniquement et spécifiquement une identité nucléaire *mat-*.

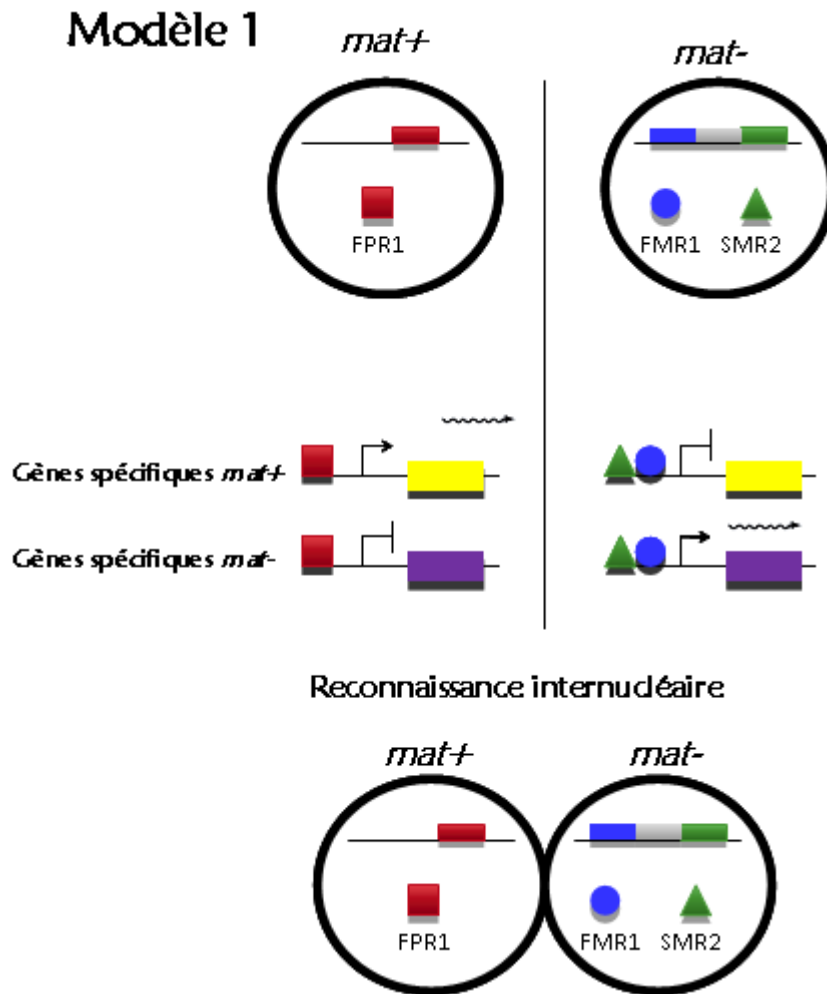


Figure 22: Modèle de l'identité nucléaire. L'activation est représentée par une flèche, la répression est représentée par une barre.

Le modèle 2 dit de ségrégation aléatoire est basé sur la formation d'un complexe de facteurs de transcription FPR1, FMR1 et SMR2. Ce nouveau régulateur retourne dans les noyaux *mat+* et *mat-* afin d'activer les gènes nécessaires à la formation des hyphes ascogènes (Figure 23).

- Dans un noyau *mat+*, FPR1 réprime les gènes nécessaires à la formation des hyphes ascogènes. Au moment de la reconnaissance internucléaire, FPR1 forme un complexe avec FMR1 et SMR2 pour activer l'expression des gènes nécessaires à la formation des hyphes ascogènes.
- Dans un noyau *mat-*, FMR1 et SMR2 répriment les gènes nécessaires à la formation des hyphes ascogènes. Au moment de la reconnaissance internucléaire, FMR1 et SMR2 forment un complexe avec FPR1 pour activer l'expression des gènes nécessaires à la formation des hyphes ascogènes.

Modèle 2

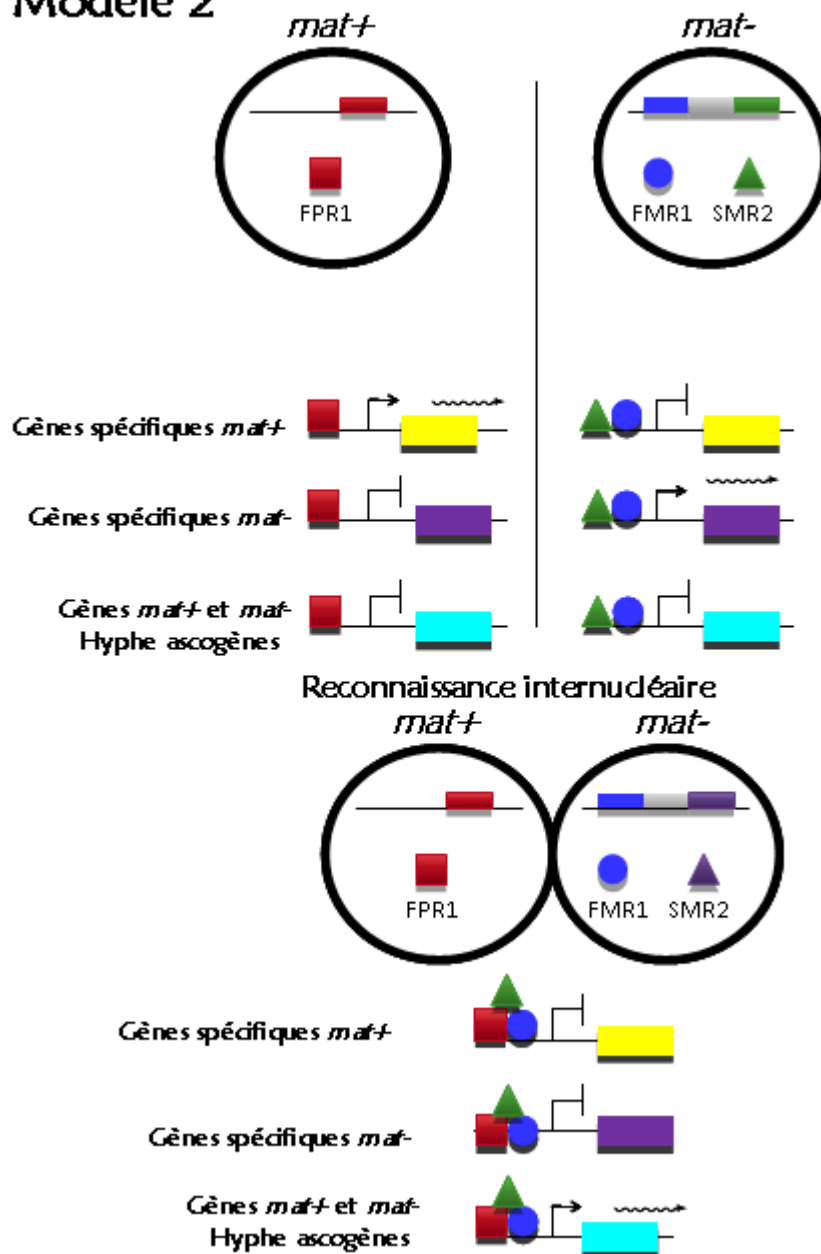


Figure 23: Modèle de ségrégation aléatoire. L'activation est représentée par une flèche, la répression est représentée par une barre.

Les approches gènes candidats, pariant sur le rôle des systèmes phéromones/récepteurs dans le modèle 1, ont abouti à la conclusion que ces gènes cibles n'étaient pas impliqués dans ce processus (Coppin et al. 2005). Les approches génétiques sans a priori se sont soldées par des échecs, probablement en raison de la complexité du processus de reconnaissance internucléaire.

Afin d'identifier les cibles des facteurs de transcription MAT impliqués dans la RI et comprendre les mécanismes de régulation mis en jeu durant cette étape, nous avons utilisé une approche microarray. Dans cette approche, on a comparé le profil transcriptomique de deux croisements, l'un n'exprimant que les gènes cibles *mat+*, l'autre que les gènes cibles

mat-. Le croisement sauvage a été inclus comme témoin dans cette comparaison. Nous avons interprété les résultats de microarray en appliquant l'un ou l'autre modèle pour établir deux listes de gènes cibles potentiels des facteurs de transcription MAT. Nous avons délété certains gènes cibles des deux listes pour déterminer si les mutants nuls obtenus se comportaient en croisement selon les prédictions du modèle.

2. Résultats et discussion

2.1 Choix du temps de prélèvement des croisements pour l'analyse de la reconnaissance internucléaire

La reconnaissance internucléaire entre les noyaux *mat+* et *mat-* se déroule entre l'étape de la fécondation et la formation des crochets. F. Bidard et V. Berteaux-Lecellier, qui ont effectué une analyse cytologique du développement des périthèces sauvages de la fécondation à la maturation des ascospores (résultats non publiés), ont identifié les premiers crochets 30 heures après la fécondation. Ces cellules en crochet sont issues des hyphes ascogènes qui prennent naissance après la RI. En conséquence nous avons étudié la RI 30 h après la fécondation.

2.2 Recherche des gènes impliqués dans la reconnaissance internucléaire selon le modèle dit « d'identité nucléaire »

2.2.1 Principe de comparaison

Le principe de l'identification des gènes d'identité nucléaire (modèle 1) consiste à comparer deux croisements qui n'expriment chacun qu'une seule identité nucléaire (Liste 1 ; Figure 24). Le croisement n'exprimant que l'identité nucléaire *mat+* implique une souche sauvage *mat+* et un partenaire *mat-* portant des mutations *fmr1⁻* et *smr2⁻*, c'est-à-dire un croisement *fmr1⁻smr2⁻ X WT mat+*. Dans ce croisement, les mutations dans *FMR1* et *SMR2* conduisent à une perte de l'expression de l'identité nucléaire *mat-*. Le croisement n'exprimant que les gènes spécifiques *mat-* implique une souche *WT mat-* et un partenaire *mat+* portant une mutation *fpr1⁻*, c'est-à-dire un croisement *fpr1⁻ X WT mat-*. La comparaison des profils transcriptomiques des différents croisements, appuyée par l'analyse d'un croisement WT, doit permettre l'identification des gènes impliqués dans l'identité nucléaire.

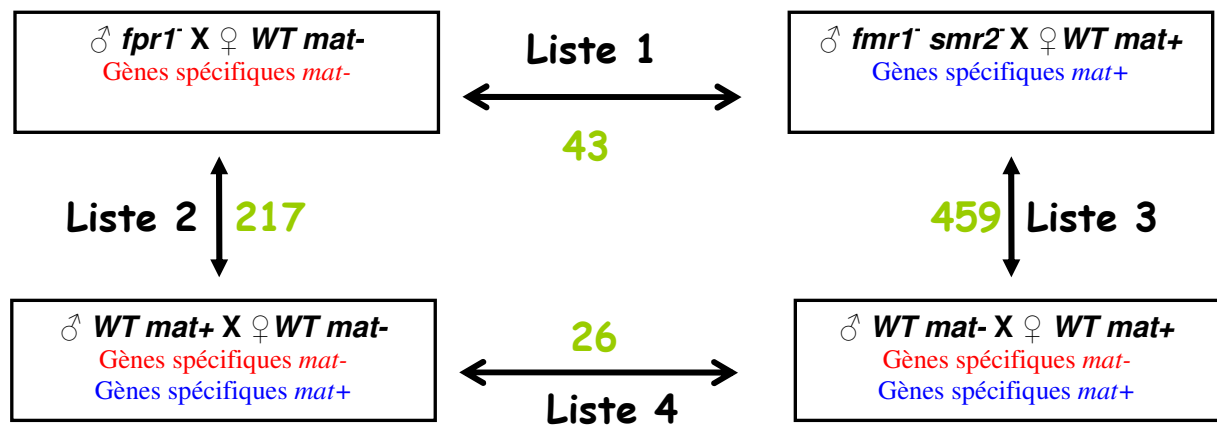


Figure 24: Plan de comparaison des croisements utilisés dans l'étude de la reconnaissance internucléaire selon le modèle 1. Les chiffres en vert représentent le nombre des gènes différentiellement exprimés entre deux croisements.

Les souches mutantes *fpr1⁻* et *fmr1⁻ smr2⁻* contiennent des mutations perte-de-fonction dans les gènes *FPR1* et *FMRI* (voir matériels et méthodes). Donc ces souches sont incapables de féconder ou d'être fécondées. La fécondation est restaurée grâce à l'expression transitoire du gène *FPR1* (type sexuel *mat+*) ou *FMRI* (type sexuel *mat-*) à partir d'un vecteur autorépliquatif pFAC1 (Barreau et al. 1998) présent dans le parent mâle. Ce vecteur est rapidement perdu dans l'organe femelle après la fécondation en l'absence de pression de sélection. En conséquence, nous avons utilisé la souche mutante *fpr1⁻* contenant le vecteur pFAC*FPR1* comme parent mâle pour féconder la souche sauvage *mat-*. D'une façon similaire, la souche *fmr1⁻ smr2⁻* contenant le plasmide pFAC*FMRI* a été utilisée comme mâle pour féconder la souche sauvage *mat+*. Dans un souci de simplification, l'indication des vecteurs pFAC*FPR1* et pFAC*FMRI* est omise dans les croisements.

2.2.2 Analyse transcriptomique et identification des gènes candidats

Les ARN ont été extraits à partir des croisements cultivés sur toile à blutter 30 h après fécondation, puis ils ont été préparés pour l'hybridation sur les puces à ADN de *P. anserina* (voir matériels et méthodes).

Le profil transcriptomique du croisement *fpr1⁻ X WT mat-* est comparé à celui du croisement *fmr1⁻ smr2⁻ X WT mat+* (Figure 24). Les gènes sont définis comme surexprimés dans le croisement *fpr1⁻ X WT mat-* si leur niveau d'expression affiche un fold change ≥ 2 avec une p-value ≤ 0.001 dans la comparaison *fpr1⁻ X WT mat-* par rapport au croisement *fmr1⁻ smr2⁻ X WT mat+* (Liste 1). Réciproquement dans la même comparaison, les gènes sont définis comme surexprimés dans le croisement *fmr1⁻ smr2⁻ X WT mat+* si leur niveau d'expression affiche un fold change ≤ -2 avec une p-value ≤ 0.001 . La comparaison a identifié

une liste de 43 gènes (Liste 1, Figure 24) dont 19 sont surexprimés dans le croisement *fprI*⁻ X *WT mat*⁻ et 24 sont surexprimés dans le croisement *fmrI*⁻ *smr2*⁻ X *WT mat*⁺.

Afin d'affiner nos analyses, les résultats transcriptomiques du croisement sauvage ♂ *WT mat*⁺ X ♀ *WT mat*⁻ et de son croisement réciproque ♂ *WT mat*⁻ X ♀ *WT mat*⁺ ont été inclus dans nos comparaisons. Ces croisements sauvages ont été réalisés au cours d'une étude cinétique du développement du périthèce (F. Bidard et V. Berteaux-Lecellier, résultats non publiés). Le profil transcriptomique du croisement ♂ *fprI*⁻ X ♀ *WT mat*⁻ est comparé avec celui de ♂ *WT mat*⁺ X ♀ *WT mat*⁻ (Liste 2, Figure 24) et le profil transcriptomique du croisement ♂ *fmrI*⁻ *smr2*⁻ X ♀ *WT mat*⁺ est comparé à celui de ♂ *WT mat*⁻ X ♀ *WT mat*⁺ (Liste 3, Figure 24). Le nombre de gènes des listes 2 et 3 est très élevé. La liste 2 compte 217 gènes. Ce nombre important de gènes était attendu dans la mesure où l'on compare un croisement mutant ♂ *fprI*⁻ X ♀ *WT mat*⁻ bloqué au stade RI à un croisement sauvage ♂ *WT mat*⁺ X ♀ *WT mat*⁻ qui progresse au delà de cette étape. Parmi les 217 gènes de la liste 2, 73 sont surexprimés dans le croisement ♂ *fprI*⁻ X ♀ *WT mat*⁻ et 144 gènes sont surexprimés dans le croisement ♂ *WT mat*⁺ X ♀ *WT mat*⁻. La liste 3 compte 459 gènes dont 180 gènes surexprimés dans le croisement ♂ *fmrI*⁻ *smr2*⁻ X ♀ *WT mat*⁺ et 279 gènes surexprimés dans le croisement ♂ *WT mat*⁻ X ♀ *WT mat*⁺. La comparaison des deux croisements sauvages fait apparaître 26 gènes différenciellement exprimés (Liste 4) dont 11 sont retrouvés dans la liste 1.

Selon le modèle 1 les gènes cibles impliqués dans la RI doivent se trouver dans la liste 1 (Figure 24). Nous avons inclus une comparaison avec les croisements sauvages pour affiner l'analyse et préciser si les gènes cibles candidats étaient activés ou réprimés. Nous avons pu ainsi classer les 43 gènes de la liste 1 en 5 classes (Tableau 6) sur la base de leur présence ou absence dans les 3 autres listes établies sur la base des comparaisons de la Figure 24. Par exemple 11 gènes sont retrouvés dans la liste 4 (Tableau 6, colonne de gauche). Ils sont donc différenciellement exprimés dans des croisements sauvages quand le partenaire femelle est de type sexuel différent. Ces gènes « sous effet femelle » listés dans le tableau 7 ont été identifiés comme différenciellement exprimés entre les souches *mat*⁺ et *mat*⁻ au stade de fécondation (Bidard et al. 2011) sauf *Pa_1_19540* et *Pa_1_910*.

Listes	Gènes sous l'effet femelle	Gènes cibles FPR1	Gènes cibles FMR1 et SMR2	Gènes cibles de FPR1, FMR1 et SMR2	Gènes non définis
Liste1	+	+	+	+	+
Liste2	0	+	0	+	0
Liste3	0	0	+	+	0
Liste4	+	0	0	0	0
Nombre de gènes	11 +4*	7 +1*	12 +6*	1	12*

Tableau 6: Classification des 43 gènes différentiellement exprimés entre le croisement ♂ *fpr1*⁻ X ♀ *WT mat*⁻ et le croisement ♂ *fmr1 smr2*⁻ X ♀ *WT mat*⁺.

"0" pas de régulation ; "+" activation ou répression. Liste 1 : ♂ *fpr1*⁻ X ♀ *WT mat*⁻/♂ *fmr1 smr2*⁻ X ♀ *WT mat*⁺ ; liste 2 : ♂ *WT mat*⁺ X ♀ *WT mat*⁻/♂ *fpr1*⁻ X ♀ *WT mat*⁻; liste 3 : ♂ *WT mat*⁻ X ♀ *WT mat*⁺/♂ *fmr1 smr2*⁻ X ♀ *WT mat*⁺ ; liste 4 : ♂ *WT mat*⁺ X ♀ *WT mat*⁻/♂ *WT mat*⁻ X ♀ *WT mat*⁺. +.* reclassement de 11 des 12 gènes de la catégorie «non définis» *(discuté dans le texte).

Gene number	Fonction	FC liste 1	FC liste 4	Classe ^a
Pa_2_2310	mat+ specific pheromone precursor encoded by the mfp gene	-83,9	-65,63	4
Pa_1_24410	Putative SAM-dependent methyltransferases	-11,04	-5,35	4
Pa_1_19170 (Δ)	Unknown function	-3,48	-3,13	1
Pa_4_1380	Putative Pheromone receptor	-3,58	-2,58	4
Pa_0_1270*(Δ)	Unknown function	-3,3	-1,84	3
Pa_1_18670 (Δ)	Unknown function	-3,23	-2,24	1
Pa_7_4100	Unknown function	-3,29	-2,04	4
Pa_4_7760*	Putative farnesyltransferase subunit beta	-2,35	-1,8	4
Pa_1_19540 (Δ)	Putative transporter	2,17	2,09	-
Pa_1_910*	Putative monooxygenase	2,1	1,58	-
Pa_1_6014*(Δ)	Unknown function	2,1	1,9	1
Pa_6_7350	Putative Aspartic protease	2,14	3,52	2
Pa_mito_nad6	Putative NADH-ubiquinone oxidoreductase chain 6	2,21	3,84	1
Pa_mito_nad1	NADH-ubiquinone oxidoreductase chain 1	3,07	4,02	1
Pa_1_8290	mat- specific pheromone precursor encoded by the mfm gene	20,79	32,07	5

Tableau 7: Liste des gènes différentiellement exprimés par un effet femelle dans la comparaison ♂ *fpr1*⁻ X ♀ *WT mat*⁻/♂ *fmr1 smr2*⁻ X ♀ *WT mat*⁺ (liste 1) et dans la comparaison ♂ *WT mat*⁺ X ♀ *WT mat*⁻/♂ *WT mat*⁻ X ♀ *WT mat*⁺ (liste 4).

^a Classe des gènes impliqués dans la fécondation voir article (Bidard et al. 2011) : (1) gènes différentiellement exprimé entre *mat*⁺ et *mat*⁻ sous l'effet de SNP (single nucleotide polymorphism) lié aux types sexuels; (2) gènes différentiellement exprimé entre *mat*⁺ et *mat*⁻ sous l'effet de : SNP lié aux types sexuels, FPR1 et FMR1 ; (4) cible de FPR1 ; (5) cible de FMR1.

(Δ) gènes sélectionnés pour la délétion dans l'analyse de la RI.

* gènes initialement classés dans la catégorie «non définis».

La liste 1 contient aussi bien les gènes cibles activés ou réprimés par FPR1 et FMR1/SMR2. La comparaison du croisement sauvage ♂ *WT mat+* X ♀ *WT mat-* exprimant les gènes spécifiques *mat+* et les gènes spécifiques *mat-*, avec le croisement ♂ *fpr1* X ♀ *WT mat-* n'exprimant que les gènes spécifiques *mat-* (Liste 2) permet d'identifier les gènes cibles de FPR1 parmi les gènes de la liste 1 (Tableau 6; Figure 24). Parmi ces cibles, les gènes activés par FPR1 sont les gènes qui accumulent plus de transcrits dans le croisement sauvage que dans le croisement ♂ *fpr1* X ♀ *WT mat-* (FC ≥ 2 dans la comparaison la liste 2). Inversement, les gènes réprimés par FPR1 ont un FC ≤ -2 dans la même comparaison. Un raisonnement similaire est utilisé dans le cas des gènes cibles FMR1 et SMR2 (Tableau 6). L'ensemble des gènes cibles de FPR1 et FMR1/SMR2 est présenté dans le tableau 8. Le gène *Pa_5_1490* est commun aux listes 1 (FC= 2; P= 0), 2 (FC= 4,17; P= 0) et 3 (FC= 9,88; P= 0). Il semble être sous le contrôle de FPR1, FMR1 et SMR2 (Tableau 6).

Gene number	Function	Fold Change liste1	FPR1 ^a	FMR1 et SMR2 ^b
Pa_5_6970 (Δ)	unknown function	2,81	R	0
Pa_5_6960 (Δ)	unknown function	2,72	R	0
Pa_2_130* (Δ)	unknown function	2,47	0	A
Pa_5_10560 (Δ)	unknown function	2,45	0	A
Pa_1_5940 (Δ)	unknown function	2,43	0	A
Pa_6_6990* (Δ)	Putative Glycoside Hydrolase Family 18	2,41	0	A
Pa_2_7580* (Δ)	Putative protein FAM62B	2,31	0	A
Pa_2_11900	Glucose-repressible protein encoded by the PaGrg1 gene	2,29	0	A
Pa_4_6000 (Δ)	unknown function	2,26	0	A
Pa_2_10300*(Δ)	unknown function	2,07	0	A
Pa_4_977 (Δ)	unknown function	2,06	0	A
Pa_2_7630*(Δ)	unknown function	-2	0	R
Pa_7_1380	Putative oxidoreductase	-2,02	0	R
Pa_2_4890*(Δ)	unknown function	-2,03	A	0
Pa_6_9980 (Δ)	unknown function	-2,08	0	R
Pa_1_13080*(Δ)	Putative transporter	-2,23	0	R
Pa_3_10200 (Δ)	unknown function	-2,23	0	R
Pa_5_6340	mitochondrial medium-chain specific acyl-CoA dehydrogenase precursor	-2,32	0	R
Pa_1_17960 (Δ)	insecticidal toxin complex protein	-2,56	A	0
Pa_5_12020 (Δ)	unknown function	-2,66	A	0
Pa_1_17947 (Δ)	unknown function	-2,67	A	0
Pa_1_17942 (Δ)	unknown function	-2,69	A	0
Pa_5_3115 (Δ)	unknown function	-2,85	0	R
Pa_5_3110 (Δ)	unknown function	-3,03	0	R
Pa_3_8000 (Δ)	unknown function	-3,18	0	R
Pa_6_9560	pisatin demethylase	-4,69	A	0

Tableau 8 : gènes régulés par FPR1 chez *mat+* et par FMR1 et SMR2 chez *mat-*. Le fold change est obtenu dans la comparaison ♂ *fpr1* X ♀ *WT mat-* / ♂ *fmr1 smr2* X ♀ *WT mat+*

^a "A" gènes activés par FPR1 ; "R" gènes réprimés par FPR1.

^b "A" gènes activés par FMR1 et SMR2; "R" gènes réprimés par FMR1 et SMR2.

(Δ) Gènes sélectionnés pour la délétion dans l'analyse de la RI.

* gènes initialement classés dans la catégorie «non définis».

Les 12 gènes de la liste 1 ont été classés comme non définis dans la mesure où dans les autres comparaisons leur FC est > -2 ou < 2 . Six gènes ont un FC proche de la valeur du cut-off $-2 < FC < -1.5$ ou $1.5 < FC < 2$ dans la liste 3 (Tableau 6 et 8), par exemple *Pa_2_7580* (FC= 1.94, P= 0). La p-value associée au FC est très significative ce qui indique qu'ils sont contrôlés par FMR1 et SMR2. De même, le gène *Pa_2_4890* semble être sous le contrôle de FPR1 en se basant sur son FC=1.93 et P= 0 dans la liste 2 (Tableau 6 et 8). Quatre autres gènes possèdent un FC proche de la valeur du cut-off $-2 < FC < -1.5$ ou $1.5 < FC < 2$ et une p-value significative dans la liste 4, exemple *Pa_1_6014* (FC= 1.9 et P= 0,00585). Donc ces quatre gènes peuvent être classés parmi les gènes sous l'effet femelle. En fait, à part le gène *Pa_1_910* qui est différentiellement transcrit au stade périthèce 30 h, les trois autres gènes (*Pa_1_6014*, *Pa_0_1270* et *Pa_4_7760*) sont différentiellement exprimés entre *mat+* et *mat-* au stade de fécondation (Bidard et al. 2011) (Tableau 7). Le gène *Pa_3_3350* a un FC < 1.5 ou FC > -1.5 dans les listes 2, 3 et 4, et s'avère être le seul gène non défini. En conclusion, parmi les 12 gènes initialement considérés comme non définis, six gènes sont contrôlés par FMR1 et SMR2, un gène est contrôlé par FPR1 et quatre gènes résultent de l'effet maternel.

2.2.3 Fonctions des gènes

D'après l'annotation fonctionnelle du génome de *P. anserina*, 50 % des gènes différentiellement exprimés entre les croisements ♂ *fpr1* X ♀ *WT mat-* et ♂ *fmr1 smr2* X ♀ *WT mat+* ont une fonction inconnue et la caractérisation des gènes connus est très succincte.

Afin de déterminer la fonction de ces gènes et leurs rôles dans la RI, nous avons procédé à une analyse comparative de leurs séquences avec celles d'autres ascomycètes modèles et à une annotation fonctionnelle selon la base des données COG (Clusters of Orthologous Groups). L'analyse des gènes cibles spécifiques FPR1 (Tableau 8, Annexe 7) montre l'existence de deux clusters de gènes, l'un est réprimé par FPR1 (*Pa_5_6960* et *Pa_5_6970*) avec une fonction inconnue, et l'autre activé par FPR1 (*Pa_1_17942*, *Pa_1_17947* et *Pa_1_17960*). *Pa_1_17960* possède un domaine RhsA (COG3209) que l'on trouve dans des protéines impliquées dans la biogenèse de l'enveloppe cellulaire. Le gène *Pa_1_17942* possède un paralogue *Pa_5_12020* qui est lui aussi activé par FPR1. La pisatine diméthylase codée par *Pa_6_9560* est activée par FPR1 et impliquée dans plusieurs processus biologiques notamment la biosynthèse des métabolites secondaires. En ce qui concerne les gènes cibles de FMR1 et SMR2 (Tableau 8, Annexe 7) on trouve plusieurs protéines spécifiques *P. anserina* (*Pa_3_8000*, *Pa_6_9980*, *Pa_5_10560* et *Pa_4_977*) dont les fonctions sont inconnues. On trouve aussi un cluster de gènes *Pa_5_3110* et *Pa_5_3115* réprimé par FMR1 et SMR2 et présent uniquement chez les Pezizomycotina. Le *Pa_2_11900*

est l'orthologue de *ccg-1* de *N. crassa*. Le gène *ccg-1* est impliqué dans le contrôle du rythme circadien, la délétion de ce gène chez *N. crassa* n'a aucun effet sur le développement sexué <http://www.broadinstitute.org/annotation/genome/neurospora/>.

Dans la liste 1 les seuls gènes identifiés comme étant impliqués dans le développement sexué de *P. anserina* sont les gènes qui codent les phéromones et le récepteur de phéromone *mfm* qui appartiennent à la classe des gènes différentiellement exprimés sous l'effet femelle (Tableau 7) (Bidard et al. 2011). Dans cette classe on trouve aussi le gène *Pa_4_7760* qui code la farnesyltransferase (Annexe 7).

2.2.4 Construction de mutants de délétion des gènes cibles

Des gènes cibles de FPR1 et des gènes cibles de FMR1 et SMR2 ont été sélectionnés pour la délétion et l'analyse phénotypique (Tableau 8). Les gènes ayant un rôle majeur dans le métabolisme n'ont pas été sélectionnés.

Dans le cas du modèle 1, on attend que la délétion d'un gène cible de FPR1 bloque la reconnaissance internucléaire quand elle est présente dans un parent *mat+*. Cette même délétion dans un parent *mat-* ne devrait avoir aucun effet. Inversement, la délétion d'un gène cible de FMR1 et SMR2 devrait bloquer la reconnaissance internucléaire dans un parent *mat-* et n'avoir aucun effet dans un parent *mat+*.

Au total, 15 souches délétées pour les gènes cibles ont été construites et examinées pour leur fertilité mâle et femelle et pour leur développement sexué. L'analyse phénotypique indique qu'aucune des souches mutantes n'a de défaut dans le développement sexué.

Un certain nombre de gènes « sous effet femelle » communs aux listes 1 et 4 avaient déjà été délétés dans l'étude précédente (voir chapitre III, Annexe 7). Trois gènes de cette catégorie ont été délétés (Tableau 7) ainsi que 10 gènes de la catégorie « gènes non définis » (Tableau 7 et 8). En résumé aucune délétion n'a affecté la reconnaissance internucléaire.

2.3 Recherche des gènes impliqués dans la reconnaissance internucléaire selon le modèle dit de « ségrégation aléatoire »

2.3.1 Principe de comparaison et analyse des gènes

L'échec de la recherche suivant le modèle d'identité nucléaire m'a conduit à appliquer le modèle de ségrégation aléatoire à l'analyse des données (modèle 2).

Les gènes nécessaires à la formation des hyphes ascogènes sont exprimés dans les deux noyaux *mat+* et *mat-*. Afin de les identifier, il faut comparer un croisement où le complexe FPR1/FMR1/SMR2 est absent avec un croisement où il est présent. Le croisement où le complexe FPR1/FMR1/SMR2 est absent implique une souche *fpr1⁻* ou *fmr1⁻smr2⁻* et un

partenaire sauvage de type sexuel opposé. Le croisement formant le complexe FPR1/FMR1/SMR2 implique une souche sauvage *mat-* et un partenaire sauvage *mat+*.

Selon le modèle 2, les gènes nécessaires à la formation des hyphes ascogènes sont les gènes différentiellement exprimés entre le croisement sauvage ♂ *WT mat+* X ♀ *WT mat-* et le croisement ♂ *fpr1* X ♀ *WT mat-* (Liste 2). Ces gènes doivent être également différentiellement exprimés entre le croisement sauvage ♂ *WT mat-* X ♀ *WT mat+* et le croisement ♂ *fmr1 smr2* X ♀ *WT mat+* (Liste 3). Dans la liste 1 et la liste 4 les gènes nécessaires à la formation des hyphes ascogènes ne doivent pas être différentiellement exprimés. En effet, dans la liste 1 on compare deux croisements qui ne forment pas le complexe FPR1/FMR1/SMR2 et aucun gène nécessaire à la formation des hyphes ascogènes n'est exprimé. Dans la liste 4 on compare deux croisements formant le complexe FPR1/FMR1/SMR2 et tous les gènes nécessaires à la formation des hyphes ascogènes sont exprimés (Figure 25).

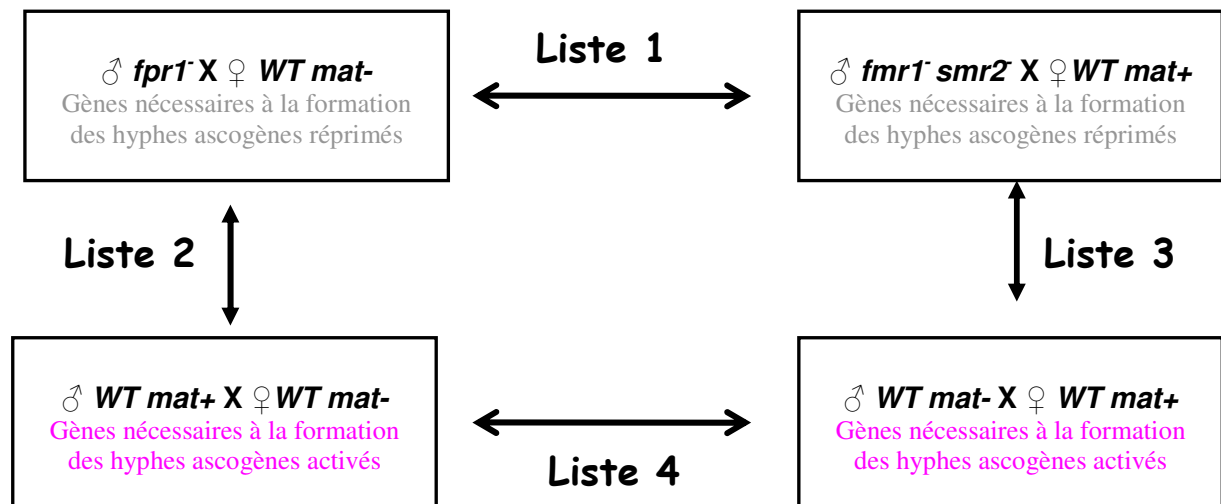


Figure 25 : Plan de comparaison des croisements utilisés dans l'étude de la reconnaissance internucléaire selon le modèle 2.

154 gènes sont communs entre la liste 2 et la liste 3, dont 113 gènes accumulent plus de transcrits dans les croisements sauvages que dans les croisements ♂ *fpr1* X ♀ *WT mat-* et ♂ *fmr1 smr2* X ♀ *WT mat+* et 41 gènes accumulent plus de transcrits dans les croisements ♂ *fpr1 mat+* X ♀ *WT mat-* et ♂ *fmr1 smr2 mat-* X ♀ *WT mat+* que dans les croisements sauvages.

2.3.2 Fonctions des gènes

Comme dans le cas de l'analyse fonctionnelle des gènes impliqués dans la RI selon le modèle 1, la majorité des gènes de la liste des gènes nécessaires à la formation des hyphes ascogènes ont des fonctions inconnues. Le reste des gènes interviennent dans différents processus biologiques (Annexe 8). Le processus de transduction de signal (*Pa_1_18840*), le transport et le trafic cellulaire (*Pa_6_1130*), le métabolisme énergétique (*Pa_5_11750*) et la

transcription (*Pa_1_16130*). Certains de ces gènes ont des orthologues chez *S. cerevisiae* et *N. crassa* qui sont impliqués dans le développement sexué: YDR403W l'orthologue de *Pa_7_190* code une enzyme qui intervient dans la maturation de l'enveloppe des spores. YDR371W l'orthologue de *Pa_5_3290* code une chitinase impliquée dans la sporulation et l'orthologue de *Pa_1_22370* chez *N. crassa* (*NCU07325*) code une protéine spécifique à la conidiation.

2.3.3 Analyse phénotypique et délétion des gènes cibles

12 gènes de la liste des gènes nécessaires à la formation des hyphes ascogènes sont sélectionnés pour la délétion et l'analyse phénotypique (Tableau 9, Annexe 8). Ces gènes ont un FC ≥ 2 dans la comparaison ♂ *WT mat+* X ♀ *WT mat-/♂ fpr1* X ♀ *WT mat-* (Liste 2) et la comparaison ♂ *WT mat-* X ♀ *WT mat+/♂ fmr1 smr2* X ♀ *WT mat+* (Liste 3). Ils possèdent des orthologues dans les champignons ayant une étape de reconnaissance internucléaire dans leur cycle. Les gènes ayant un rôle majeur dans le métabolisme n'ont pas été sélectionnés pour les délétions.

Gene number	Fonction	FC Liste 2	FC Liste 3
Pa_1_290	Unknown function	7,53	16,49
Pa_1_540	Unknown function	7,36	9,57
Pa_1_21670	Unknown function	6,42	12,69
Pa_6_11710	Unknown function	5,98	9,17
Pa_1_18840	Unknown function	5,68	10,37
Pa_5_940	Unknown function	5,59	6,98
Pa_4_2080	Unknown function	5,44	9,81
Pa_1_3290	Unknown function	5,14	6,14
Pa_5_90	Unknown function	5,02	5,78
Pa_1_18880	Unknown function	4,91	7,73
Pa_1_17250	Unknown function	4,85	4,69
Pa_6_1130	Putative Annexin	4,58	5,82

Tableau 9: Les gènes sélectionnés pour la délétion dans la liste des gènes nécessaires à la formation des hyphes ascogènes.

Selon le modèle 2, on attend que la délétion d'un gène nécessaire à la formation des hyphes ascogènes affecte la RI quand elle est localisée dans les deux parents *mat+* et *mat-*. L'analyse de la fertilité mâle et femelle, du développement de périthèces et de la production

d'ascospores dans des croisements homozygotes pour chaque délétion n'a révélé aucun phénotype particulier par comparaison à un croisement sauvage.

3. Conclusion

Les étapes contrôlées par les protéines régulatrices FPR1, FMR1 et SMR2 au cours de la reproduction sexuée sont maintenant bien déterminées mais les mécanismes mis en jeu sont inconnus. Le problème que nous voulons résoudre est le mécanisme de la RI. Les approches génétiques, destinées à rechercher des gènes cibles des facteurs de transcription MAT impliqués dans la RI ont toutes échouées. Devant ces échecs, nous avons choisi de rechercher ces gènes cibles par une autre méthode que les méthodes génétiques en utilisant une approche microarray.

Dans cette étude, nous avons analysé le transcriptome des périthèces issus de deux croisements, l'un entre un mutant *fpr1⁻* et un partenaire sauvage *mat⁻* et l'autre entre un mutant *fmr1⁻smr2⁻* et un partenaire sauvage *mat⁺*. Le profil transcriptomique de chaque croisement mutant est comparé d'une part à celui de l'autre croisement mutant et d'autre part à un croisement témoin sauvage.

Nos résultats montrent une grande complexité du système d'analyse. Les résultats ont été interprétés selon deux modèles. Dans le premier modèle, dit d'identité nucléaire, chaque noyau possède une identité spécifique constituée par des molécules spécifiques des noyaux *mat⁺* ou spécifiques des noyaux *mat⁻*. Selon ce modèle, 8 gènes spécifiques *mat⁺* et 18 gènes spécifiques *mat⁻* ont été identifiés, mais aucun de ces gènes n'a une fonction déjà décrite dans le développement sexué. Le deuxième modèle est basé sur la coopération entre un noyau *mat⁺* et un noyau *mat⁻* pour former un complexe des facteurs de transcription FPR1, FMR1 et SMR2. Ce régulateur contrôlerait les gènes nécessaires à la formation des hyphes ascogènes contenant les deux noyaux *mat⁺* et *mat⁻*. Selon ce modèle, 154 gènes communs entre *mat⁺* et *mat⁻* ont été identifiés dont certains ont des fonctions potentielles dans plusieurs processus biologiques. Nous avons tenté de valider expérimentalement successivement chacun de ces deux modèles. Aucune des 46 souches mutantes construites ne présente un défaut du développement sexué. En conclusion aucun des deux modèles n'a pu être validé et le mécanisme de la RI reste encore à découvrir.

Il existe plusieurs explications possible à cet échec : soit plusieurs mutations sont nécessaires pour affecter la RI, soit nous n'avons pas sélectionné la bonne liste des gènes nécessaires à la RI, soit la régulation de la RI par les facteurs de transcription MAT implique un troisième modèle plus complexe que les deux modèles proposés. Si une approche transcriptomique devait être refaite, il conviendrait d'utiliser un matériel biologique plus

spécifique du stade que nous étudions. En effet, nous avons utilisé un mélange constitué de fragment de mycelium et de périthèces entiers, incluant à la fois l'enveloppe d'origine maternelle et le centrum constitué des tissus sexués où a lieu la RI. Des techniques de microdissection laser sont actuellement développées par le laboratoire de Minou Nowrousian pour isoler des cellules spécifiques des champignons et appliquer les nouvelles méthodes de RNAseq <http://homepage.ruhr-uni-bochum.de/Minou.Nowrousian/projects.html>. Cette technique permettrait d'isoler le centrum et de focaliser l'analyse transcriptomique sans être gêné par l'effet maternel. Je propose également de déterminer le site de fixation de FPR1, FMR1 et SMR2 par des approches moléculaires, ce qui nous aidera à sélectionner facilement leurs cibles dans les listes des gènes obtenus. Enfin il faut étudier la relation de FPR1, FMR1 et SMR2 avec d'autres facteurs de transcription susceptibles de contrôler la RI.

Chapitre V : Etude des gènes à HMG box chez
Podospora anserina

1. Introduction

Le développement sexué chez les champignons est relativement complexe et nécessite l'intervention de plusieurs régulateurs transcriptionnels et de nombreux gènes cibles. Chez *S. cerevisiae* et *S. pombe* un grand nombre de gènes qui affectent le développement sexué ont été identifiés: les gènes qui codent les phéromones et leurs récepteurs, la machinerie de transduction du signal [protéine G, protéines kinases phosphorylant les protéines MAP (Mitogen Activated Proteins), adénylate cyclase, protéine kinase A dépendante de l'AMPc (PKA)] et les gènes qui codent les régulateurs de l'expression des gènes nécessaires pour la conjugaison et les étapes ultérieures. Certaines de ces protéines régulatrices possèdent un domaine de fixation à l'ADN notamment le domaine HMG (High Mobility Group).

Les protéines à domaine HMG appartiennent au groupe des protéines HMGB caractérisés par la présence d'un domaine de fixation à l'ADN très conservé, constitué d'environ 80 acides aminés, et structuré en trois hélices alpha adoptant la forme générale d'un L. Cette structure permet la fixation de la protéine sur le petit sillon de la double hélice d'ADN (voir chapitre I partie 3) et aboutit à la courbure de l'ADN qui est contraint entre la grande et la petite branche du L. Les protéines à domaine HMG sont classées en fonction du nombre de domaines HMG et en fonction des modalités de leur fixation à l'ADN. Par exemple la famille des MATA_HMG, qui nous intéresse particulièrement, est caractérisée par un seul domaine HMG et se fixe spécifiquement sur l'ADN. Elle inclut les protéines des types sexuels et le facteur de transcription Ste11p.

Les gènes à HMG-box ont été identifiés à l'intérieur et à l'extérieur du locus *mat* des champignons filamenteux. A ce jour, la fonction des protéines codées par les gènes à HMG-box situés à l'extérieur du locus *mat* est peu étudiée (Aronstein et al. 2007). Ces protéines peuvent avoir un rôle majeur dans le développement sexué des champignons, comme le montre le cas de Ste11p, la protéine à domaine MATA_HMG de *S. pombe*. Chez notre modèle d'étude *P. anserina*, les seuls gènes à HMG-box dont la fonction a été étudiée sont les gènes des types sexuels *FPR1*, *FMR1* et *SMR2* (Arnaise et al. 2001; Arnaise et al. 2008; Bidard et al. 2011), et *mtHMG1* qui code une protéine mitochondriale (Dequard-Chablat and Alland 2002). *FPR1* et *SMR2* contiennent un domaine MATA_HMG, et *FMR1* contient un domaine MAT α _HMG, qui est un sous-groupe récemment identifié des MATA_HMG (Martin et al. 2010). *FPR1* appartient à l'idiomorphe *mat+* tandis que *FMR1* et *SMR2* sont localisés dans l'idiomorphe *mat-* (voir chapitre I). La délétion du locus *mat* cause la stérilité chez les souches *mat+* et *mat-*. Par contre cette délétion n'a aucun effet sur la croissance végétative et la production des organes reproducteurs (Coppin et al. 1993). Les études

généétiques ont montré que FPR1 et FMR1 contrôlent la reconnaissance intercellulaire pendant la fécondation, tandis que FPR1, FMR1 et SMR2 contrôlent la reconnaissance internucléaire lors du passage d'un syncytium plurinucléé aux hyphes ascogènes dicaryotiques. Ces contrôles sont assurés par la régulation directe ou indirecte de l'expression de gènes cibles. L'analyse transcriptomique effectuée au stade de fécondation sur les souches sauvages *mat+* et *mat-* et les souches mutantes *fpr1* et *fmr1* a permis d'identifier 108 gènes contrôlés par FPR1 dans une souche *mat+* et 49 gènes contrôlés par FMR1 dans une souche *mat-* (Bidard et al. 2011). Parmi ces gènes figurent ceux des systèmes phéromone/récepteur connus chez les champignons pour leur rôle dans la fécondation (Voir chapitre I et chapitre III). La protéine mtHMG1 codée par le gène *Pa_1_13340* (Tableau 10) est localisée dans la mitochondrie en association avec l'ADN mitochondrial. L'étude réalisée par Dequard-Chablat (Dequard-Chablat and Alland 2002) a identifiée deux domaines HMG (HMG1 et HMG2) dans cette protéine qui combine deux types de domaines de fixation à l'ADN : le motif AT-hook et un domaine HMG (Figure 26). La protéine mtHMG1 est impliquée dans le maintien, la stabilité et la transmission du génome mitochondrial (Dequard-Chablat and Alland 2002).

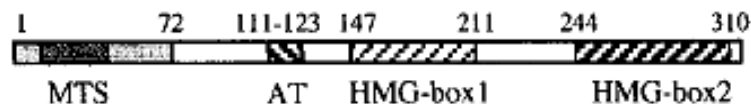


Figure 26: Les caractéristiques de la protéine mtHMG1. MTS, putative mitochondrial targeting sequence. AT, domaine AT-hook. HMG, domaine HMG (Dequard-Chablat and Alland 2002).

Nous avons complété cet inventaire en identifiant et en caractérisant huit autres gènes à HMG-box, dont l'orthologue de *ste11* de *S. pombe* et de *fmf-1* de *N. crassa*. J'ai construit les souches délétées correspondantes afin de déterminer (en collaboration avec E. Coppin) le rôle de ces gènes pendant le développement sexué de *P. anserina*. Je me suis particulièrement intéressée à l'orthologue de *ste11* qui semble avoir un rôle crucial dans le contrôle de l'expression des gènes des types sexuels.

2. Résultats

2.1 Caractérisation des protéines à domaine HMG chez

Podospora anserina

Chez les ascomycètes comme chez les autres groupes de champignons, le nombre de gènes à HMG-box est variable selon les espèces (9 chez *S. pombe*, 8 chez *S. cerevisiae*, 13 chez *N. crassa*). Pour identifier les gènes à HMG-box de *P. anserina*, j'ai recherché les orthologues de protéines HMG de *N. crassa* chez *P. anserina* avec FUNGIpath (Voir matériels et méthodes). Ces deux espèces appartiennent à l'ordre des sordariales; leurs lignées

auraient divergé depuis au moins 75 millions d'années (Saupe et al. 2000). Douze gènes à HMG-box ont été détectés chez *P. anserina* (Tableau 10). Ils codent les protéines des types sexuels (FMR1, SMR2 et FPR1), la protéine mitochondriale mtHMG1 (Pa_1_13340), deux protéines non-histone (Pa_1_9380 et Pa_1_14230) et six protéines de fonction inconnue (Pa_1_13940, Pa_6_4110, Pa_1_7390, Pa_1_11050, Pa_5_8400, Pa_7_7190).

La présence du domaine HMG dans les protéines identifiées chez *P. anserina* a été validée par des recherches effectuées dans *Pfam* (Punta et al. 2012) qui indiquent que les douze protéines de *P. anserina* ne possèdent qu'un seul domaine HMG. L'analyse de la protéine Pa_1_13340 (mtHMG1) nécessite une mention particulière. En effet, Dequard-Chablat et Alland rapportent que mtHMG1 possède deux domaines HMG (Dequard-Chablat and Alland 2002). Le premier de ces domaines n'est pas reconnu par *Pfam*, qui identifie dans cette région un domaine de fonction inconnue (DUF1898), et un domaine HMG en C-terminal. L'analyse de l'orthologue de Pa_1_13340 chez *N. crassa* (NCU02695) par *Pfam* identifie également un domaine DUF1898 et un domaine HMG. L'analyse de Pa_1_13340 (mtHMG1) avec CD-search (Marchler-Bauer et al. 2011) n'identifie qu'un seul domaine HMG, de type HMGB-UBF_HMG dans la région C-terminale, alors que l'analyse de son orthologue NCU02695 identifie deux domaines HMG, le premier ayant un score très faible et recouvrant DUF1898, le second localisé en C-terminal et de type HMGB-UBF_HMG. Il est donc probable que le domaine DUF1898 est apparenté aux domaines HMG. L'analyse des autres protéines HMG avec CD-search m'a permis d'identifier 5 protéines MATA_HMG et 4 protéines HMGB_UBF. Les protéines Pa_5_8400 et Pa_7_7190 n'ont pas été classées dans l'une des sous-familles connues des HMG. Leur alignement avec le consensus du domaine HMG (Martin et al. 2010) indique qu'elles contiennent un résidu chargé au position 9 au lieu d'un résidu aliphatique ou aromatique (Figure 27). Toutefois, en -1 et +2 de cette position 9, on trouve bien les résidus aromatiques caractéristiques des HMG permettant l'ancrage des hélices dans le noyau hydrophobe. Cet alignement montre aussi une conservation de certains résidus entre tous les domaines HMG notamment le résidu tryptophane à la position 57.

J'ai également identifié différents domaines en plus du domaine HMG (Tableau 10). Pa_1_7390 possède un domaine SAM ou Sterile Alpha Motif qui est impliqué dans les interactions protéine-protéine et protéine-ARN. Ce domaine est présent dans différentes protéines notamment les régulateurs de la transcription (par exemple les protéines Polyhomeotic (Ph) et le polymère Sex-comb-on-midleg (Scm) du groupe polycomb, les régulateurs de la traduction (exemple la protéine Smaug qui réprime la traduction des ARNm nanos afin d'établir un gradient morphogène dans l'embryon de la *Drosophile*), les tyrosines kinases, les serine/thréonine kinases et le récepteur Eph de la tyrosine kinase impliqué dans la

différentiation cellulaire (Kim and Bowie 2003). Le domaine SAM est aussi présent dans les protéines Byr2p et Ste4p de *S. pombe* (voir Figure 17, paragraphe 3.3.3.4.2), et Ste11p (MAPKKK) et Ste50p de *S. cerevisiae*. Ces protéines sont impliquées dans la transduction du signal phéromone/récepteur et leur absence induit la stérilité de *S. pombe* et *S. cerevisiae*. Pa_1_9380 contient les domaines SWIRM et amine oxydase, caractéristiques de la famille des NAO (Nuclear Amine Oxydase) qui inclut Splsd1. Le domaine SWIRM est présent dans les protéines Swi3 et Rsc8 de *S. cerevisiae* et Moira de *D. melanogaster*. Ces protéines font partie des complexes de remodelage de la chromatine ATP-dépendant SWI/SNF et RSC chez *S. cerevisiae*, et du complexe Brahma (BRM) chez *D. melanogaster*. Le domaine SWIRM se fixe à l'ADN et au nucléosome afin de faciliter l'interaction des complexes de régulation de la chromatine avec l'ADN (Da et al. 2006). Le domaine amine oxydase est identifié dans les enzymes qui catalysent la réaction d'oxydation dépendante au FAD des substrats amine (Forneris et al. 2009). Ce domaine intervient directement dans la déméthylation de la lysine 9 de l'histone H3 (Forneris et al. 2009). Pa_5_8400 possède un domaine SprT-like caractéristique des métallopeptidases à zinc. La fonction de ce domaine n'a pas été analysée expérimentalement, cependant la présence de métalloprotéases est décelée dans les complexes intervenant dans l'élongation de la transcription, comme par exemple le complexe CP de *S. cerevisiae* et son équivalent FACT chez l'Homme (Ponting 2002). Ces métalloprotéases ont perdu les résidus essentiels du site catalytique mais Ponting suppose qu'elles ont conservé des sites de fixation à l'ADN courbé (Ponting 2002), comme cela a été montré pour la métallopeptidase M24 de *S. pombe* (Yamada et al. 1994).

Les prédictions de localisation cellulaire des protéines HMG indiquent que la protéine Pa_1_13340 (mtHMG1) est mitochondriale en accord avec l'analyse effectuée par (Dequard-Chablat and Alland 2002), FMR1 et SMR2 sont cytoplasmiques et nucléaires tandis que les 9 autres protéines HMG sont toutes nucléaires (Voir matériels et méthodes).

Name	Fonction	Taille de la protéine (Acides aminés)	Domaine HMG	E-value Pfam	Position du domaine HMG	Autres domaines	E-value Pfam
FMR1	Mating type protein	305	MAT α _HMG	7.1e-70	7-193	0	0
SMR2	Mating type protein	288	MATA_HMG	1.2e-20	161-229	0	0
FPR1	Mating type protein	402	MATA_HMG	6.3e-23	169-237	0	0
Pa_1_13940	female and male fertility	559	MATA_HMG	1.2e-10	198-266	0	0
Pa_6_4110	protein of unknown function	945	MATA_HMG	6.5e-18	212-274	0	0
Pa_1_7390	protein of unknown function	478	HMGB_UBF	2.3e-17	119-185	SAM (steril alpha motif)	1.6e-11
Pa_1_9380	Lysine-specific histone demethylase Aof2	1073	HMGB_UBF	1.5e-06	919-994	SWIRM domain	2.2e-11
						Amine oxidase	8.6e-86
Pa_1_11050	protein of unknown function	333	HMGB_UBF	1.1e-05	132-201	0	0
Pa_1_13340	mtHMG1 mitochondrial protein	316	HMGB_UBF	5.3e-07	277-309	DUF 1898/HMG	0
Pa_1_14230	non-histone chromosomal protein	98	HMGB_UBF	1.4e-25	26-94	SprT like family	0
P_5_8400	protein of unknown function	587	HMGB	2.4e-06	511-551	0	1.8e-50
Pa_7_7190	protein of unknown function	353	HMGB	0.018	260-293	0	0

Tableau 10: Caractéristiques des protéines à domaine HMG de *P. anserina*.

0 : Absence d'autres domaines avec un score significatif

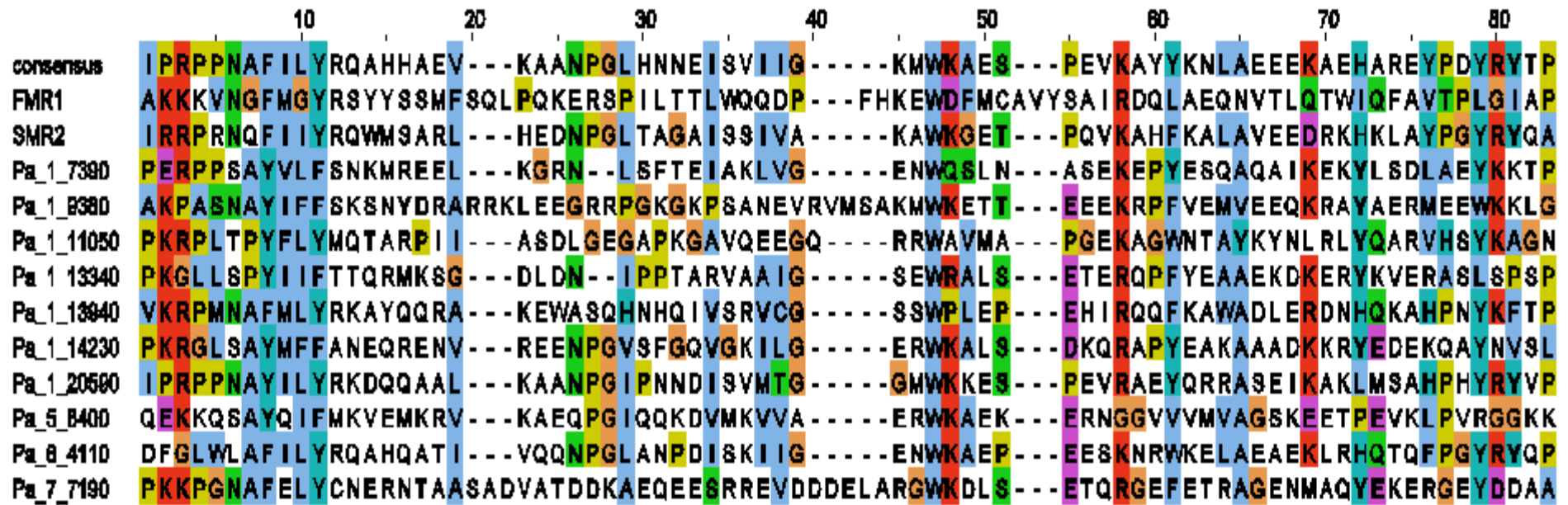


Figure 27: Alignement du consensus HMG avec les protéines à domaine HMG de *P. anserina*. L'alignement est obtenu avec ClustalW2, la coloration est fournie par Jalview (Annexe 9).

2.2 Analyse phylogénique des protéines à domaine HMG chez

Podospora anserina

En collaboration avec Tom Martin et Christina Dixelius (SLU University, Uppsala, Suède), une analyse phylogénique basée sur le maximum de vraisemblance a été effectuée à partir des domaines HMG de protéines d'animaux, de plantes et de champignons (Figure 28 et Tableau 11). La majorité des domaines HMG sont répartis dans l'un des trois groupes : MATA_HMG, HMGB-UBF_HMG et SOX-TCF_HMG. Cependant, quelques domaines HMG des champignons n'ont pas pu être classés dans ces trois groupes. Comme il a été noté précédemment (Soullier et al. 1999), les domaines MATA_HMG sont présents exclusivement chez les champignons, alors que les domaines SOX-TCF_HMG ne se trouvent que chez les animaux. Il y a toutefois une exception, puisque l'une des protéines du type sexuel de *Phycomyces blakesleeanus* (Mucoromycotina, anciennement Zygomycète) est classée dans les SOX-TCF.

Le groupe phylogénique des MATA_HMG contient des protéines typiques codées par les gènes *MAT1-2-1* des types sexuels (FPR1, Podan2a) et 4 clades bien distincts, contenant chacun une protéine HMG de *P. anserina* :

- Un sous-groupe caractérisé par la présence de MAT α p de *S. cerevisiae* et ses orthologues (Figure 28, A). Ce sous-groupe définit la classe des MAT α _HMG identifiée récemment par Martin et al (Martin et al. 2010). Ce sous-groupe contient FMR1 (Podan1).
- Un sous-groupe contenant les protéines codées par le gène *MAT1-1-3* du type sexuel des Pezizomycotina (SMR2, Podan3 de *P. anserina*) (Figure 28, E). Le regroupement de cette classe de protéines avait déjà été noté précédemment (Martin et al. 2010).
- Un sous-groupe monophylétique caractérisé par la présence de Pa_1_13940 (Podan2c) et ses orthologues : Ste11p (Schpo2c) de *S. pombe* et FMF-1 (Neucr2c) de *N. crassa* (Figure 28, G).
- Un sous-groupe contenant Pa_1_9380 (Podan5e) défini comme un HMGB-UBF (Figure 28, H). Ce sous-groupe contient également LSD1 (Schpo7a) de *S. pombe*, dont le classement n'est pas précisé par les logiciels d'analyse de signature. Toutes les protéines de ce sous-groupe (AN3580 de *A. nidulans*, Aspni6; 190976 de *Cochliobolus heterostrophus*, Coche6; NCU9120 de *N. crassa*, Neucr5e) contiennent les domaines SWIRM et amine oxydase, caractéristiques des histones déméthylases. La présence de ces domaines suggère que les protéines de ce sous-groupe sont des orthologues fonctionnels.

On trouve également dans le groupe phylogénique des MATA_HMG la protéine Pa_6_4110 (Podan2b) qui est l'orthologue de ROX1 (Sacce2) chez *S. cerevisiae*. Des études précédentes (Soullier et al. 1999) ont montré que même si ROX1 n'est pas impliquée dans le développement sexué, son domaine HMG appartient au groupe des MATA_HMG.

Le groupe des HMGB-UBF_HMG contient 5 protéines HMG de *P. anserina*. Pa_1_13340 (mtHMG1, Podan5d) est l'orthologue de NCU02695 (Neucr5d). Ces deux protéines possèdent un signal d'adressage à la mitochondrie et appartiennent au groupe HMGB-UBF sans être dans le même clade. Dans deux cas, le phylogramme regroupe des protéines qui ont également en commun d'autres domaines, confirmant ainsi ces regroupements et suggérant qu'on y retrouve des orthologues fonctionnels. Le premier cas est Pa_1_7390 (Podan5c) qui est regroupée avec ses orthologues NCU02819 (Neucr5c) de *N. crassa* et CC14121 (Coche5c) de *C. heterostrophus*. Ces trois protéines possèdent un domaine SAM (Sterile Alpha Domain). Le deuxième cas est Pa_5_8400 (Podan6b) qui est regroupé avec NCU06874 (Neucr6b) de *N. crassa* et NP_595970 (Schpo7b) de *S. pombe*. Ces protéines ont en commun un domaine SprT-like. Dans un cas, on retrouve un regroupement phylogénique qui ne recouvre pas les classements d'orthologie effectués par FUNGIpath. Il s'agit de Pa_1_11050 (Podan5b) qui forme un groupe monophylétique avec ses orthologues NCU03126 (Neucr5b) de *N. crassa*, AN10103 (Aspni5b) de *A. nidulans*, CC33388 (Coche5b) de *C. heterostrophus* et XP_001548220 (Botfu5) de *Botryotinia fuckeliana*. Dans ce groupe monophylétique on trouve NHP6B (Sacce5b) et Hmo1p (Sacce5d) de *S. cerevisiae* qui ne sont pas des orthologues de Pa_1_11050. NHP6B et Hmo1p sont classées par FUNGIpath comme des orthologues de Pa_1_14230 et devraient donc se retrouver dans le même clade que cette protéine, où l'on trouve déjà NHP6A (Sacce5a), un inparalogue de NHP6B.

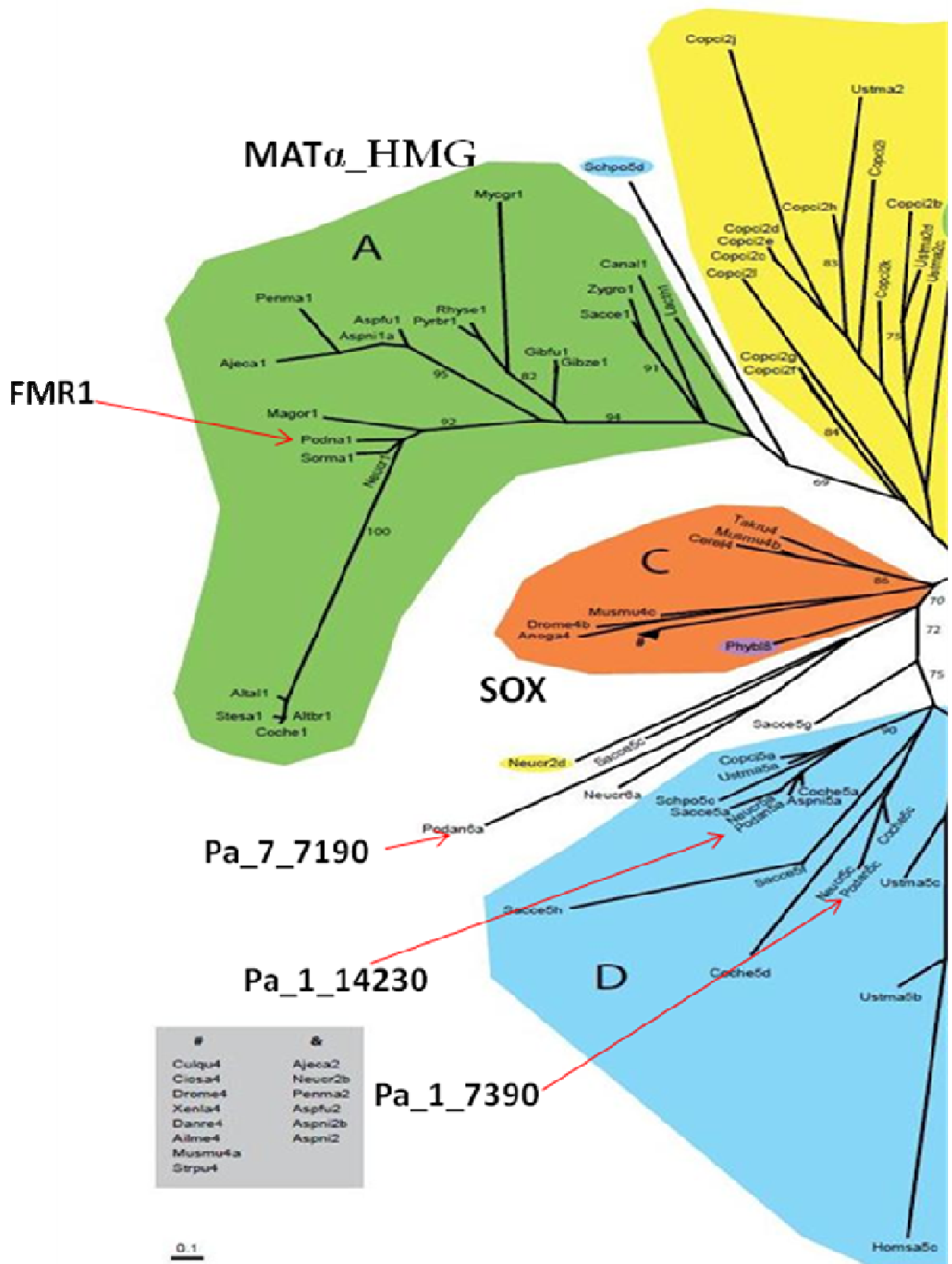
La séquence HMG de Pa_1_7190 (Podan6a) est groupée avec d'autres HMG (NCU07568, Neucr6a) de *N. crassa*, Nhp10p (Sacce5c) et Abf2p (Sacce5g) de *S. cerevisiae* entre les groupes SOX et HMGB-UBF. On trouve également NCU02326 (Neucr2d) dans cette branche située entre SOX et HMG_UBF. FUNGIpath identifie NCU02326 comme un in-paralogue de NCU09387 (Neucr2c), qui est placé dans le sous-groupe G des MATA_HMG avec Pa_1_13940. Selon FUNGIpath, Pa_1_13940 aurait donc un orthologue (NCU02326) et un in-paralogue (NCU09387) chez *N. crassa*.

Groupe	Code	Nom de la Protéine	Domaine	Organisme	nombre d'accession	
Sordariales	Podan1	MAT1-1-1/FMR1	α 1	<i>Podospora anserina</i>	CAA45519	
	Podan2a	MAT1-2-1/FPR1	MATA_HMG	<i>Podospora anserina</i>	CAA45520.1	
	Podan3	MAT1-1-3/SMR2	MATA_HMG	<i>Podospora anserina</i>	CAA52051	
	Podan2b		MATA_HMG	<i>Podospora anserina</i>	Pa_6_4110	
	Podan2c		MATA_HMG	<i>Podospora anserina</i>	Pa_1_13940	
	Podan5a		HMGB	<i>Podospora anserina</i>	Pa_1_14230	
	Podan5b		HMGB	<i>Podospora anserina</i>	Pa_1_11050	
	Podan5c		HMGB	<i>Podospora anserina</i>	Pa_1_7390	
	Podan5d	mtHMGI1	HMGB	<i>Podospora anserina</i>	Pa_1_13340	
	Podan5e		HMGB	<i>Podospora anserina</i>	Pa_1_9380	
	Podan6b		HMG	<i>Podospora anserina</i>	Pa_5_8400	
	Podan6a		HMG	<i>Podospora anserina</i>	Pa_7_7190	
	Neucr1	MAT1-1-1/mat A-1	α 1	<i>Neurospora crassa</i>	AAC37478	
	Neucr2a	MAT1-2-1/mat a-1	MATA_HMG	<i>Neurospora crassa</i>	AAA33598	
	Neucr3	MAT1-1-3/mat A-3	MATA_HMG	<i>Neurospora crassa</i>	AAC37476	
	Neucr2b		MATA_HMG	<i>Neurospora crassa</i>	NCU03481	
	Neucr2c	FMF-1	MATA_HMG	<i>Neurospora crassa</i>	NCU09387	
	Neucr2d		MATA_HMG	<i>Neurospora crassa</i>	NCU02326	
	Neucr5a		HMGB	<i>Neurospora crassa</i>	NCU09995	
	Neucr5b		HMGB	<i>Neurospora crassa</i>	NCU03126	
	Neucr5c		HMGB	<i>Neurospora crassa</i>	NCU02819	
	Neucr5d		HMGB	<i>Neurospora crassa</i>	NCU02695	
	Neucr5e		HMGB	<i>Neurospora crassa</i>	NCU09120	
	Neucr6a		HMG	<i>Neurospora crassa</i>	NCU07568	
	Neucr6b		HMG	<i>Neurospora crassa</i>	NCU06874	
	Sorma1	SMTA-1	α 1	<i>Sordaria macrospora</i>	CAA71623	
	Sorma2	SMTa-1	MATA_HMG	<i>Sordaria macrospora</i>	CAA71624.1	
	Magnaporthales	Magor1	MAT1-1-1	α 1	<i>Magnaporthe oryzae</i>	BAC65083.1
		Magor2	MAT1-2-1	MATA_HMG	<i>Magnaporthe oryzae</i>	BAC65094.1
	Diaporthales	CryPa1	MAT1-1-1	α 1	<i>Cryphonectria Parasitica</i>	AAK83346
		CryPa3	MAT1-1-3	MATA_HMG	<i>Cryphonectria Parasitica</i>	AAK83344.1
	Hypocreales	Fusac3	MAT1-1-3	MATA_HMG	<i>Fusarium acaciae-mearnsii</i>	ABE98373.1
Gibfu1		MAT1-1-1	α 1	<i>Gibberella fujikuroi</i>	AAC71055	
Gibfu3		MAT1-1-3	MATA_HMG	<i>Gibberella fujikuroi</i>	AAC71053.1	
Gibze1		MAT1-1-1	α 1	<i>Gibberella zeae</i>	AAG42809	
Gibze3		MAT1-1-3	MATA_HMG	<i>Gibberella zeae</i>	AAG42812	
Incertae sedis	Verda2	MAT1-2-1	MATA_HMG	<i>Verticillium dahliae</i>	BAG12301.1	
Helotiales	Botfu5	HMG	HMGB	<i>Botryotinia fuckeliana</i>	XP_001548220	
	Pyrbr1	MAT1-1-1/Pad1	α 1	<i>Pyrenopeziza brassicae</i>	CAA06844	
	Pyrbr3	MAT1-1-3/phb1	MATA_HMG	<i>Pyrenopeziza brassicae</i>	CAA06846.1	
	Rhyse1	MAT1-1-1	α 1	<i>Rhynchosporium secalis</i>	CAD71141.1	
Eurotiales	Aspfu1	MAT1-1-1	α 1	<i>Aspergillus fumigatus</i>	AAX83123.1	
	Aspfu2		MATA_HMG	<i>Aspergillus fumigatus</i>	XP_751745.1	
	Aspni1	MAT1-1/MATB	α 1	<i>Aspergillus nidulans</i>	EAA63189.1	
	Aspni2a	MAT1-2/MATA	MATA_HMG	<i>Aspergillus nidulans</i>	AN4734	
	Aspni2b		MATA_HMG	<i>Aspergillus nidulans</i>	AN1962	
	Aspni2c		MATA_HMG	<i>Aspergillus nidulans</i>	AN3667	
	Aspni2d		MATA_HMG	<i>Aspergillus nidulans</i>	AN3549	
	Aspni5a		HMGB	<i>Aspergillus nidulans</i>	AN2885	
	Aspni5b		HMGB	<i>Aspergillus nidulans</i>	AN10103	
	Aspni5c		HMGB	<i>Aspergillus nidulans</i>	AN1267	
	Aspni6		HMG	<i>Aspergillus nidulans</i>	AN3580	
	Penma1	MAT1-1-1/MAT-1	α 1	<i>Penicillium marneffeii</i>	ABC68484.1	
Penma2		MATA_HMG	<i>Penicillium marneffeii</i>	XP_002151220.1		
Onygenales	Ajeca1	MAT1-1-1	α 1	<i>Ajellomyces capsulatus</i>	ABO87596.1	
	Ajeca2	MAT1-2-1	MATA_HMG	<i>Ajellomyces capsulatus</i>	EER39720.1	
Pleosporales	Altal1	MAT1-1-1	α 1	<i>Alternaria alternata</i>	BAA75907.1	
	Altal2	MAT1-2-1	MATA_HMG	<i>Alternaria alternata</i>	BAA75903.1	
	Altbr1	MAT1-1-1	α 1	<i>Alternaria brassicicola</i>	AAK85542.1	
	Bipsa2	MAT1-2-1	MATA_HMG	<i>Bipolaris sacchari</i>	CAA65081.1	
	Coche1	MAT1-1-1	α 1	<i>Cochliobolus heterostrophus</i>	CAA48465	
	Coche2a	MAT1-2-1	MATA_HMG	<i>Cochliobolus heterostrophus</i>	CAA48464.1	
	Cocho2	MAT1-2/1	MATA_HMG	<i>Cochliobolus homomorphus</i>	AAD33441.1	
	Coche2b		MATA_HMG	<i>Cochliobolus heterostrophus</i>	estExt_fgenesH1_p	

					g.C_90036
	Coche2c		MATA_HMG	<i>Cochliobolus heterostrophus</i>	estExt_fgenes1_p g.C_60282
	Coche5e		HMGB	<i>Cochliobolus heterostrophus</i>	estExt_Genewise1 Plus.C_90517
	Coche5c		HMGB	<i>Cochliobolus heterostrophus</i>	Fgenes1_pm.C_s caffold_14000132
	Coche5d		HMGB	<i>Cochliobolus heterostrophus</i>	estExt_fgenes1_p g.C_150217
	Coche5b		HMGB	<i>Cochliobolus heterostrophus</i>	estExt_fgenes1_p g.C_280025
	Coche5a		HMGB	<i>Cochliobolus heterostrophus</i>	e_gw1.5.923.1
	Coche6		HMG	<i>Cochliobolus heterostrophus</i>	estExt_fgenes1_p g.C_30103
	Pyrte2	MAT1-2-1	MATA_HMG	<i>Pyrenophora teres</i>	AAAY35017
	Stesa1	MAT1-1	$\alpha 1$	<i>Stemphylium sarciniforme</i>	AAR04460
Dothideales	Dotpi2	MAT1-2-1	MATA_HMG	<i>Dothistroma pini</i>	ABK91353
	Mycgr1	MAT1-1-1	$\alpha 1$	<i>Mycosphaerella graminicola</i>	AAL30838
	Mycgr2	MAT1-2-1	MATA_HMG	<i>Mycosphaerella graminicola</i>	AAL30836.1
Saccharomycotina	Canal1	MTL $\alpha 1$	$\alpha 1$	<i>Candida albicans</i>	XP_714749
	Canal2	Rfg1p	MATA_HMG	<i>Candida albicans</i>	XP_715804.1
	Lacth1	MAT $\alpha 1$	$\alpha 1$	<i>Lachancea thermotolerans</i>	XP_002554225.1
	Sacce1	Mat $\alpha 1$ p	$\alpha 1$	<i>Saccharomyces cerevisiae</i>	EDN62161.1
	Sacce2	Rox1p	MATA_HMG	<i>Saccharomyces cerevisiae</i>	NP_015390.1
	Sacce5a	NHP6A	HMGB	<i>Saccharomyces cerevisiae</i>	EDN61184
	Sacce5b	NHP6B	HMGB	<i>Saccharomyces cerevisiae</i>	NP_009647
	Sacce5c	Nhp10p	HMGB	<i>Saccharomyces cerevisiae</i>	NP_010282.1
	Sacce5d	Hmo1p	HMGB	<i>Saccharomyces cerevisiae</i>	NP_010459.1
	Sacce5e	Ixr1p	HMGB	<i>Saccharomyces cerevisiae</i>	NP_012893.1
	Sacce5f		HMGB		
	Sacce5g	Abf2p	HMGB	<i>Saccharomyces cerevisiae</i>	NP_013788.1
	Sacce5h		HMGB		
	Zygro1	MAT $\alpha 1$	$\alpha 1$	<i>Zygosaccharomyces rouxii</i>	XP_002497889.1
Taphrinomycotina	Pneca2	STE11	MATA_HMG	<i>Pneumocystis carinii</i>	Q870J1
	Schja2	STE11	MATA_HMG	<i>Schizosaccharomyces japonicus</i>	XP_002175130
	Schpo2c	STE11	MATA_HMG	<i>Schizosaccharomyces pombe</i>	CAA77507.1
	Schpo6	Pc	HMG	<i>Schizosaccharomyces pombe</i>	P10841
	Schpo2a	mat-Mc	MATA_HMG	<i>Schizosaccharomyces pombe</i>	NP_595867.1
	Schpo5a		HMGB	<i>Schizosaccharomyces pombe</i>	NP_595672.1
	Schpo5d	Nht1	HMGB	<i>Schizosaccharomyces pombe</i>	NP_593259.1
	Schpo5c		HMGB	<i>Schizosaccharomyces pombe</i>	NP_593314.1
	Schpo5b	NHP6B	HMGB	<i>Schizosaccharomyces pombe</i>	NP_593693.1
	Schpo7a	SWIRM1	HMG	<i>Schizosaccharomyces pombe</i>	NP_595398.1
	Schpo7b		HMG	<i>Schizosaccharomyces pombe</i>	NP_595970.1
Basidiomycota	Ustma2b	Prf1	MATA_HMG	<i>Ustilago maydis</i>	AAC32736
	Ustma2c	Hmg3 / Rop1	MATA_HMG	<i>Ustilago maydis</i>	UM01457
	Ustma2d		MATA_HMG	<i>Ustilago maydis</i>	UM06025
	Ustma5d		HMGB	<i>Ustilago maydis</i>	UM01390
	Ustma5e		HMGB		
	Ustma6		HMG	<i>Ustilago maydis</i>	UM04007
	Ustma5f		HMGB		
	Ustma5c		HMGB	<i>Ustilago maydis</i>	UM06160
	Ustma5a		HMGB	<i>Ustilago maydis</i>	UM02364
	Ustma5b		HMGB	<i>Ustilago maydis</i>	UM00808
	Copci2b	pcc1	MATA_HMG	<i>Coprinopsis cinerea</i>	CC1G_07392
	Copci5b		HMGB	<i>Coprinopsis cinerea</i>	CC1G_01334
	Copci5a		HMGB	<i>Coprinopsis cinerea</i>	CC1G_11882
	Copci2c		HMG	<i>Coprinopsis cinerea</i>	CC1G_04121
	Copci2d		HMG	<i>Coprinopsis cinerea</i>	CC1G_07470
	Copci2e		HMG	<i>Coprinopsis cinerea</i>	CC1G_11980
	Copci2f		HMG	<i>Coprinopsis cinerea</i>	CC1G_07876
	Copci2g		HMG	<i>Coprinopsis cinerea</i>	CC1G_07859
	Copci2h		HMG	<i>Coprinopsis cinerea</i>	CC1G_03744
	Copci2i		HMG	<i>Coprinopsis cinerea</i>	CC1G_08141
	Copci2j		HMG	<i>Coprinopsis cinerea</i>	CC1G_13994
	Copci2k		HMG	<i>Coprinopsis cinerea</i>	CC1G_00255

	Copci2l		HMG	<i>Coprinopsis cinerea</i>	CC1G_06994	
Zygomycota	Phybl8	SexM	HMG	<i>Phycomyces blakesleeanus</i>	ABX27909.1	
	Phybl9	SexP	HMG	<i>Phycomyces blakesleeanus</i>	ABX27912.1	
Microsporidia	Antlo7	HMG	HMG	<i>Antonospora locustae</i>	ACI87876.1	
	Entbi7	Sex locus	HMG	<i>Enterocytozoon bieneusi</i>	ACI87872	
	Enccu7	Sex locus	HMG	<i>Encephalitozoon cuniculi</i>	NP_585883	
Animalia	Ailme4	hypothetical	SOX	<i>Ailuropoda melanoleuca</i>	EFB23328	
	Anoga4	AGAP003896-PA	SOX	<i>Anopheles gambiae</i>	XP_001230616.1	
	Caeel5	HMG-4	HMGB	<i>Caenorhabditis elegans</i>	NP_498633	
	Cerel4	SRY	SOX	<i>Cervus elaphus yarkandensis</i>	ABK91721	
	Ciosa4	TF	SOX	<i>Ciona savignyi</i>	NP_001071831.1	
	Culqu4	Pangolin	SOX	<i>Culex quinquefasciatus</i>	XP_001864781	
	Danre4	TF-7	SOX	<i>Danio rerio</i>	AAI63927	
	Danre5a	RNA Polymerase 1	HMGB	<i>Danio rerio</i>	CAQ14015	
	Danre5b	TF	HMGB	<i>Danio rerio</i>	NP_957297	
	Danre5c	TF	HMGB	<i>Danio rerio</i>	NP_957297.1	
	Drome4a	Pangolin	SOX	<i>Drosophila melanogaster</i>	NP_001014685.1	
	Drome4b	bobby sox	SOX	<i>Drosophila melanogaster</i>	NP_001027087.1	
	Homsa5a	RNA Polymerase 1	HMGB	<i>Homo sapiens</i>	EAW51616	
	Homsa5b	TF1	HMGB	<i>Homo sapiens</i>	NP_055048.1	
	Homsa5c	HMG	HMGB	<i>Homo sapiens</i>	2CS1_A	
	Musmu4a	lymphoid enhancer	SOX	<i>Mus musculus</i>	EDL12207.1	
	Musmu4b	SOX-1	SOX	<i>Mus musculus</i>	NP_033259.2	
	Musmu4c	HMG	SOX	<i>Mus musculus</i>	NP_694878.2	
	Strpu4	Tcf/Lef	SOX	<i>Strongylocentrotus purpuratus</i>	NP_999640.1	
	Takru4	SOX8b	SOX	<i>Takifugu rubripes</i>	AAQ18506	
	Xenla4	XTCF-3b	SOX	<i>Xenopus laevis</i>	CAA67689	
	Xenla5a	TF1-B	HMGB	<i>Xenopus laevis</i>	NP_001079429	
	Xenla5b	ubtf-b	HMGB	<i>Xenopus laevis</i>	AAH42232	
	Xenla5c	ubtf-b	HMGB	<i>Xenopus laevis</i>	AAH42232.1	
	Planta	Arath5	HMG	HMGB	<i>Arabidopsis thaliana</i>	AAK43965.1

Tableau 11: Les protéines à domaine HMG utilisées dans la construction du phylogramme.



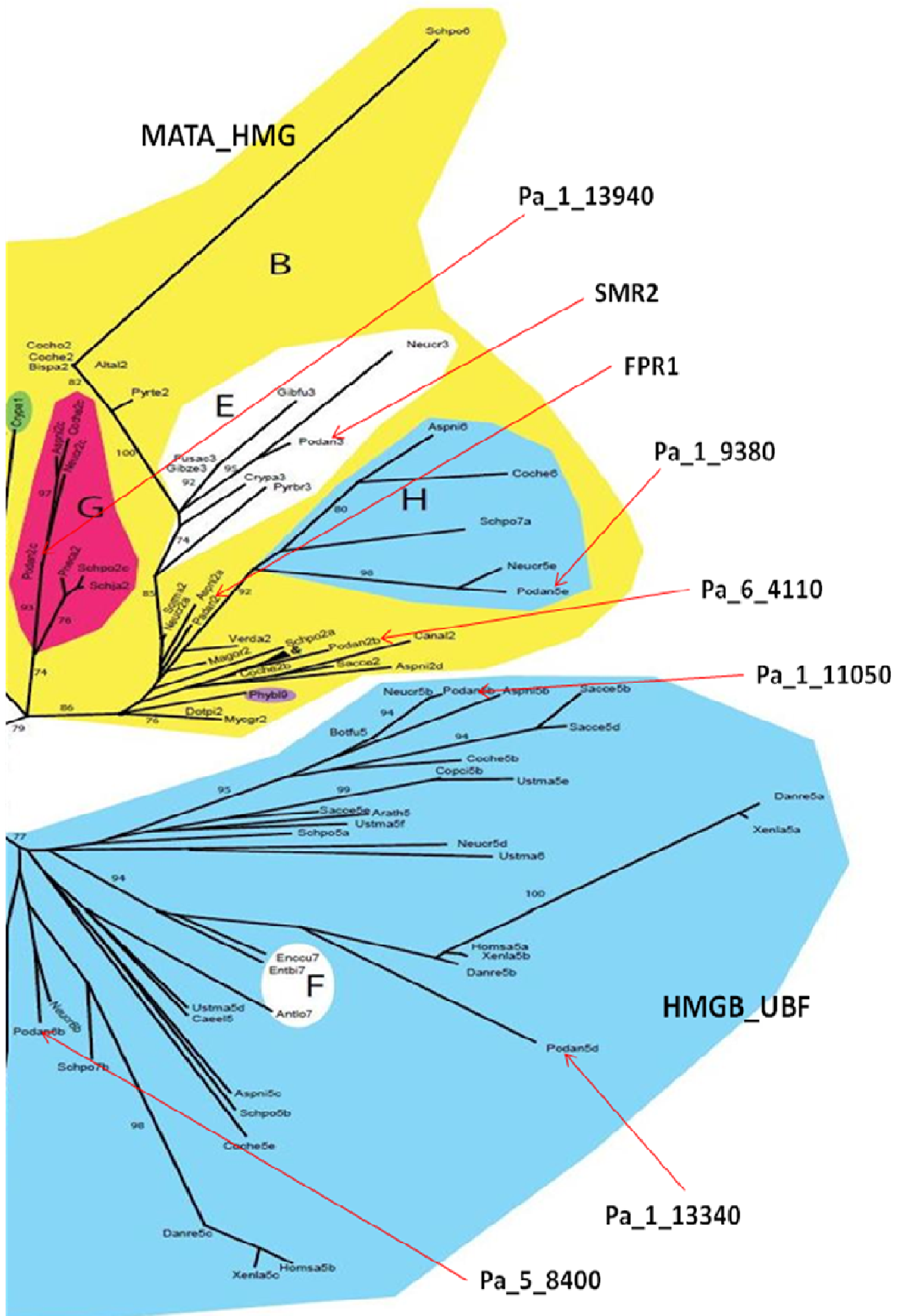


Figure 28: Phylogramme non raciné des domaines HMG. Le clustering du core des sequences HMG a été fait avec la méthode du maximum de vraisemblance et le modèle LG+G.

L'étiquetage est le suivant : MAT α -HMG (A), MATA_HMG (B), SOX (C), HMGB (D), MAT1-1-3 sous groupe de MATA_HMG (E), Microsporidia MAT sex locus HMG (F), STE11 sous groupe de MATA_HMG (G), sous groupe de MATA_HMG non identifié, *Podospora anserina* (Podan) ; *Neurospora crassa* (Neucr) ; *Sordaria macrospora* (Sorma) ; *Magnaporthe oryzae* (Magor) ; *Cryphonectria Parasitica* (CryPa) ; *Fusarium acaciae-mearnsii* (Fusac) ; *Gibberella fujikuroi* (Gibfu) ; *Gibberella zeae* (Gibze) ; *Verticillium dahliae* (Verda) ; *Botryotinia fuckeliana* (Botfu) ; *Pyrenopeziza brassicae* (Pyrbr) ; *Rhynchosporium secalis* (Rhyse) ; *Aspergillus fumigatus* (Aspfu) ; *Aspergillus nidulans* (Aspni) ; *Penicillium marneffeii* (Penma) ; *Ajellomyces capsulatus* (Ajeca) ; *Alternaria alternata* (Altal) ; *Alternaria brassicicola* (Altbr) ; *Bipolaris sacchari* (BisPa) ; *Cochliobolus heterostrophus* (Coche) ; *Pyrenophora teres* (Pyrte) ; *Stemphylium sarciniforme* (Stesa) ; *Dothistroma pini* (Dotpi) ; *Mycosphaerella graminicola* (Mycgr) ; *Candida albicans* (Canal) ; *Lachancea thermotolerans* (Lacth) ; *Saccharomyces cerevisiae* (Sacce) ; *Zygosaccharomyces rouxii* (Zygro) ; *Pneumocystis carinii* (Pneca) ; *Schizosaccharomyces japonicus* (Schja) ; *Schizosaccharomyces pombe* (Schpo) ; *Ustilago maydis* (Ustma) ; *Coprinospora cinerea* (Copci) ; *Phycomyces blakesleeanus* (Phybl) ; *Antonospora locustae* (Antlo) ; *Enterocytozoon bieneusi* (Entbi) ; *Encephalitozoon cuniculi* (Enccu) ; *Ailuropoda melanoleuca* (Alime) ; *Anopheles gambiae* (Anoga) ; *Caenorhabditis elegans* (Caeel) ; *Cervus elaphus yarkandensis* (Cerel) ; *Ciona savignyi* (Ciosa) ; *Culex quinquefasciatus* (Culqu) ; *Danio rerio* (Danre) ; *Drosophila melanogaster* (Drome) ; *Homo sapiens* (Homsa) ; *Strongylocentrotus purpuratus* (Strpu) ; *Takifugu rubripes* (Takru) ; *Xenopus laevis* (Xenla) ; *Arabidopsis thaliana* (Arath). Phylogramme construit par Tom Martin et Christina Dixelius.

2.3 Analyse phénotypique des mutants HMG chez *Podospora anserina*

Les huit nouveaux gènes à HMG-box identifiés dans notre étude ont été invalidés afin d'en préciser la fonction. Pour chacun des mutants, plusieurs caractéristiques phénotypiques ont été analysées : la croissance, la fertilité mâle et femelle, le développement des périthèces et la résistance à certains inhibiteurs spécifiques de différentes fonctions (voir matériels et méthodes). L'ensemble des résultats est résumé dans le tableau 12.

La délétion des gènes *Pa_1_14230*, *Pa_6_4110* et *Pa_7_7190* induit une perte de la fertilité femelle, bien que ces trois souches soient capables de différencier des organes femelles (Figure 29). La quantification de la production de spermaties chez ces mutants montre également une anomalie chez le mutant ΔPa_7_7190 , qui produit 4 à 5 fois plus de spermaties que la souche sauvage. Les mutants ΔPa_1_14230 et ΔPa_7_7190 sont affectés pour la croissance végétative (Figure 29).

Le mutant ΔPa_1_13940 est stérile en tant que mâle et en tant que femelle bien qu'il différencie des spermaties et des ascogones (Tableau 12, Figure 29). Ce phénotype n'a été observé que chez les souches portant une délétion du locus *mat* (Coppin et al. 1993) ou une mutation perte de fonction dans *FMR1* et *FPR1* (El-Khoury et al. 2008). Ce phénotype est aussi observé chez les mutants $\Delta fmf-1$ de *N. crassa* (Johnson 1979) (voir paragraphe suivant).

La souche ΔPa_1_11050 a une croissance végétative et une fertilité femelle comparable au sauvage mais sa production de spermaties est 7 fois plus élevée que celle de la souche sauvage de référence. L'étude de la zone de fructification chez ce mutant (voir matériels et méthodes) montre une distribution anormale des périthèces. Chez la souche sauvage les périthèces forment un anneau à environ 1 cm de l'implant, alors que chez le mutant ΔPa_1_11050 , les périthèces forment un disque autour de l'implant (Tableau 12, Figure 29 et 30).

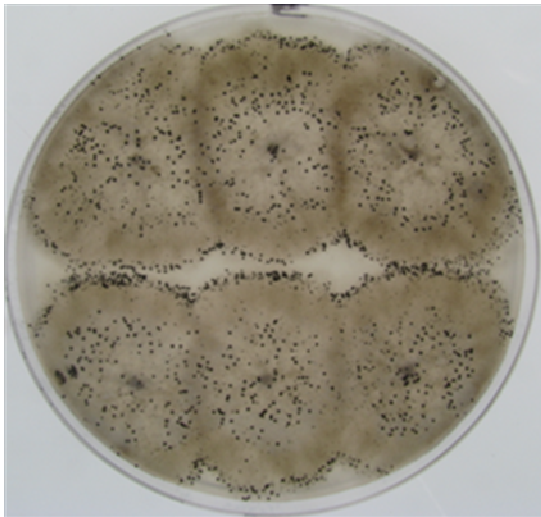
Le stade végétatif du mutant ΔPa_1_7390 est caractérisé par un mycélium pigmenté marbré et sectorisé, alors que son développement sexué est tout à fait normal (Tableau 12, Figure 29).

Enfin les deux souches ΔPa_1_9380 et ΔPa_5_8400 ne présentent aucun défaut ni dans le développement sexué ni dans la croissance végétative (Tableau 12, Figure 29). Le double mutant $\Delta Pa_1_9380 \Delta Pa_5_8400$ ne présente lui non plus aucun défaut par rapport à une souche sauvage.

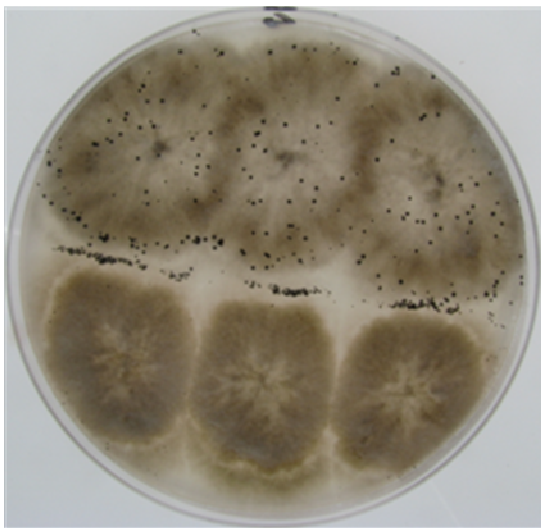
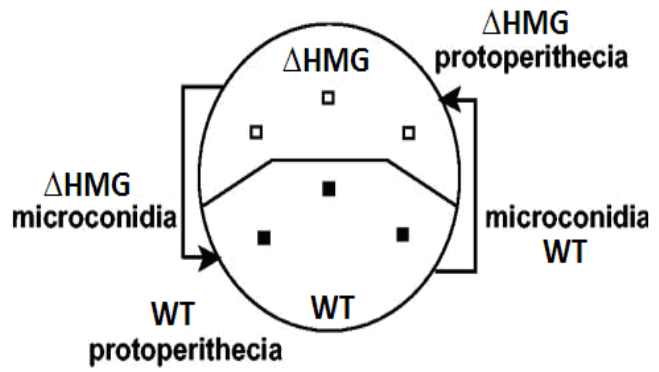
Les souches délétées pour les gènes HMG ont été examinées pour leur sensibilité au calcofluor qui agit sur la paroi, au fludioxonil qui agit sur l'osmorégulation, au SDS qui agit sur la membrane plasmidique et à la caféine qui est un inhibiteur des phosphodiesterases (voir matériels et méthodes). Ces tests n'ont révélé aucune différence entre les mutants HMG et le sauvage.

Souches	Phase végétative	Fertilité	
		Femelle	Mâle
ΔPa_1_14230	croissance en vagues ralentie de 10%	stérile	fertile
ΔPa_7_7190	mycelium ras et clair, croissance 50 à 60% par rapport au WT	stérile	fertile
ΔPa_1_13940	Normale	stérile	stérile
ΔPa_6_4110	Normale	stérile	fertile
ΔPa_1_7390	mycelium pigmenté, marbré et sectorisé	fertile	fertile
ΔPa_1_11050	Normale	fertile	fertile
ΔPa_1_9380	Normale	fertile	fertile
ΔPa_5_8400	Normale	fertile	fertile
WT	Normale	fertile	fertile

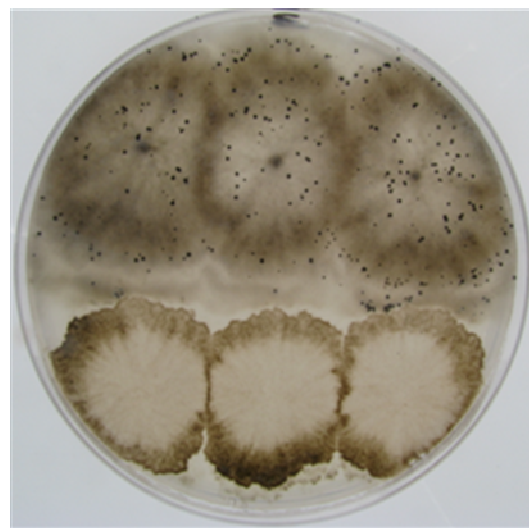
Tableau 12: Phénotypes induits par les délétions des gènes à HMG-box chez *P. anserina*.



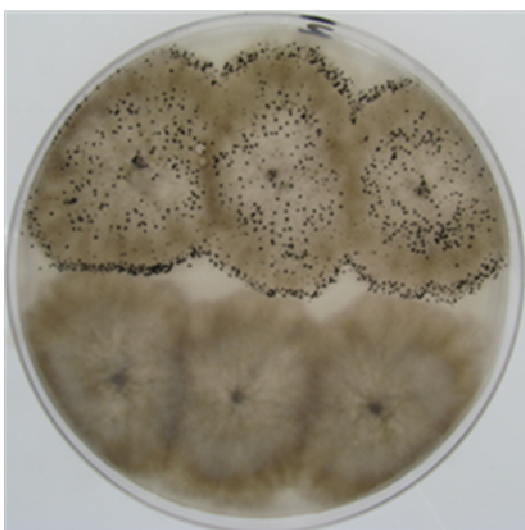
WT X WT



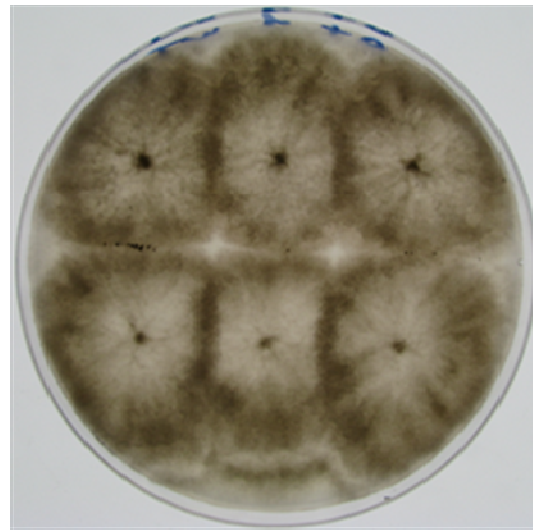
WT X Δ Pa_1_14230



WT X Δ Pa_7_7190



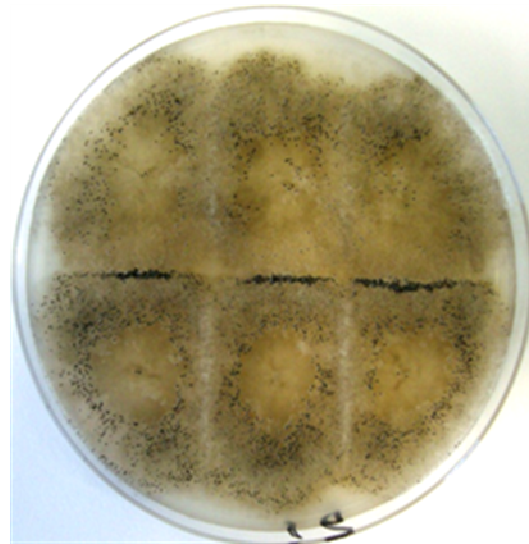
WT X Δ Pa_6_4110



WT X Δ Pa_1_13940



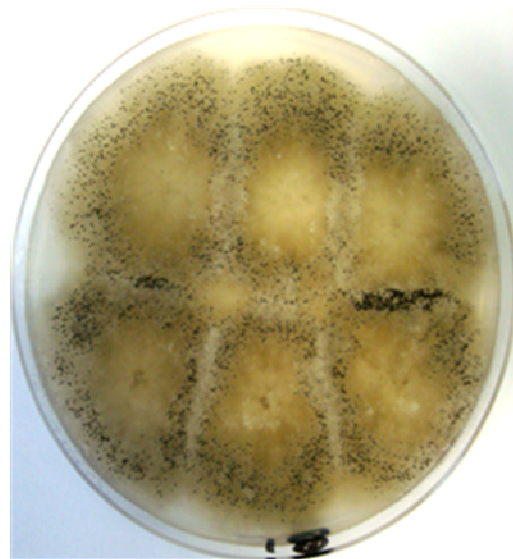
WT X Δ Pa_1_7390



WT X Δ Pa_1_11050

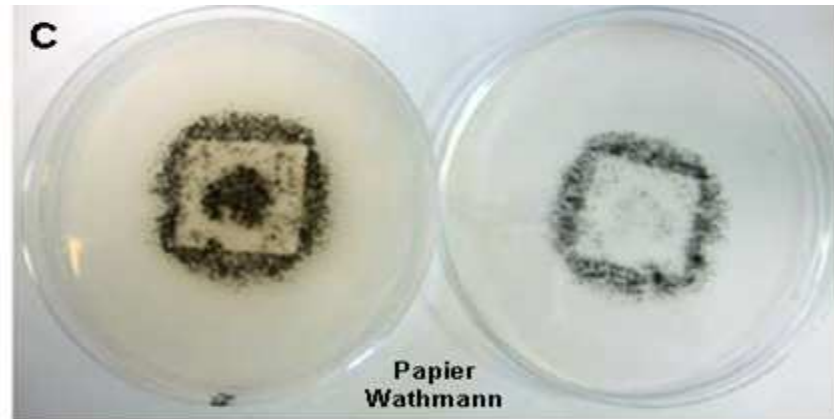
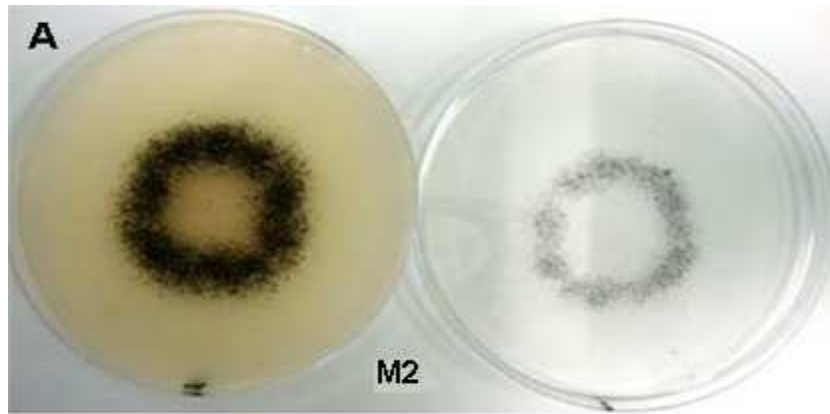


WT X Δ Pa_1_9380

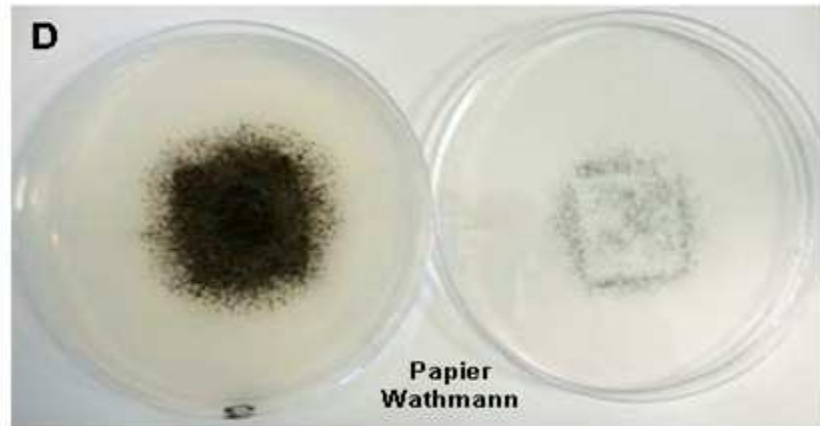
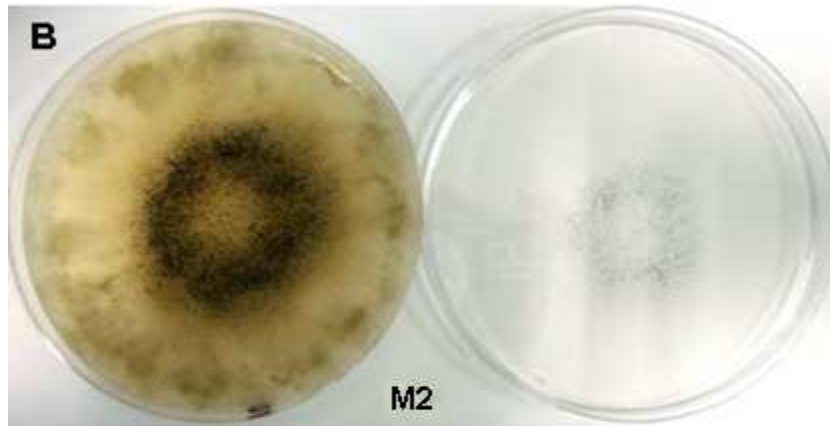


WT X Δ Pa_5_8400

Figure 29: Croisements hétérozygotes entre les souches mutantes *mat+* et la souche sauvage *mat-*. Les croisements réciproques mutant *mat-* X sauvage *at+* sont identiques.



WT mat+ X *WT mat-*



$\Delta pa_1_11050 mat+$ X $\Delta pa_1_11050 mat-$

Figure 30 : Comparaison de la distribution des périthèces entre les croisements sauvages (A et C) et les croisements homozygotes mutants ΔPa_1_11050 (B et D). 5 μ l d'un mixat de mycelium *mat+* et *mat-* de la souche analysée sont déposés au centre d'une boîte contenant du milieu minimum M2 (A et B, boîtes de gauche) ou du milieu minimum avec du papier Wathman comme seule source de carbone (C et D, boîtes de droite). Quand les périthèces arrivent au terme de leur développement ils projettent les ascospores qui peuvent être récoltées sur un couvercle d'agar (boîtes de droite).

2.4 Analyse des mutants HMG affectés dans le développement sexué

2.4.1. Identification des relations de régulation entre les gènes à HMG-box

Les quatre gènes *Pa_6_4110*, *Pa_7_7190*, *Pa_1_14230* et *Pa_1_13940* dont la délétion est associée à un défaut de fécondité femelle et /ou mâle semblent avoir un rôle important dans le développement sexuel de *P. anserina*. Des expériences de RT-qPCR ont été utilisées pour quantifier la transcription de ces quatre gènes dans les mutants ΔPa_6_4110 , ΔPa_7_7190 , ΔPa_1_14230 et ΔPa_1_13940 de type sexuel *mat+* et *mat-* par rapport au sauvage *mat+* et *mat-* (Annexe 10). Les résultats révèlent l'existence d'un lien de régulation entre les gènes à HMG-box *Pa_6_4110*, *Pa_7_7190*, *Pa_1_14230* et *Pa_1_13940*. La transcription de *Pa_1_14230* et de *Pa_7_7190* n'est affectée dans aucune des souches mutantes pour les autres gènes à HMG-box. Cela signifie que *Pa_1_14230* et *Pa_7_7190* sont placés au sommet de la voie de régulation des protéines HMG. La transcription de *Pa_6_4110* varie dans tous les mutants HMG, ce qui nous permet de placer ce gène en bas de cette voie de régulation. *Pa_1_13940* est au centre de la voie de régulation puisqu'il a un phénotype intermédiaire par rapport aux autres gènes HMG (Tableau 13, Figure 31). *Pa_6_4110* et *Pa_1_13940* sont sur-exprimés dans le mutant ΔPa_7_7190 et sous-exprimés dans le mutant ΔPa_1_14230 par rapport au sauvage. Ces résultats signifient que *Pa_7_7190* a un effet négatif sur la transcription de ces deux gènes, alors que *Pa_1_14230* a un effet positif. *Pa_6_4110* est plus faiblement transcrit dans un mutant ΔPa_1_13940 que dans un sauvage, indiquant que *Pa_1_13940* régule positivement la transcription de *Pa_6_4110* (Figure 31).

Gène	Souche ΔPa_1_13940		Souche ΔPa_1_14230		Souche ΔPa_6_4110		Souche ΔPa_7_7190	
	<i>mat+</i>	<i>mat-</i>	<i>mat+</i>	<i>mat-</i>	<i>mat+</i>	<i>mat-</i>	<i>mat+</i>	<i>mat-</i>
<i>Pa_1_13940</i>			down	down	N/D	N/D	up	up
<i>Pa_1_14230</i>	N/D	N/D			N/D	N/D	N/D	N/D
<i>Pa_6_4110</i>	down	down	down	down			up	up
<i>Pa_7_7190</i>	N/D	N/D	N/D	N/D	N/D	N/D		

Tableau 13: Résultats de RT-qPCR des gènes *Pa_6_4110*, *Pa_7_7190*, *Pa_1_14230* et *Pa_1_13940* dans les mutants HMG ΔPa_6_4110 , ΔPa_7_7190 , ΔPa_1_14230 et ΔPa_1_13940 par rapport au sauvage.

Up: gène sur-exprimé dans le mutant par rapport au sauvage.

Down : gène sous-exprimé dans le mutant par rapport au sauvage.

N/D:gène non différentiellement exprimé.

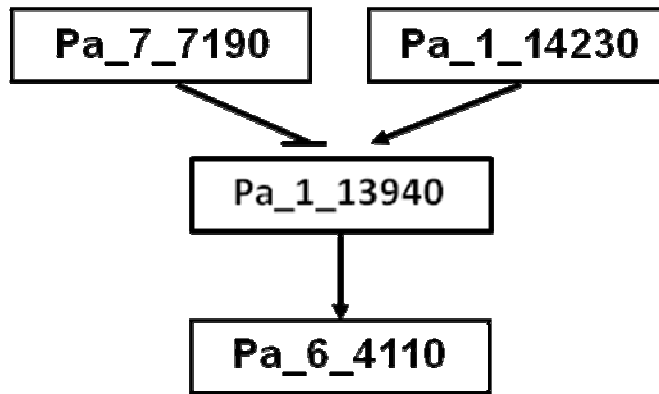


Figure 31: la voie de régulation des gènes à HMG-box. La flèche noire représente une activation, la barre représente une répression.

2.4.2. Profils d'expression des gènes *mat* et de leurs cibles dans les mutants HMG

Le mutant ΔPa_1_13940 est stérile en tant que mâle et comme femelle. Ce phénotype, très rare, est identique à celui des mutants des gènes *mat*. Pour tester l'hypothèse d'un effet de *Pa_1_13940* sur la transcription des gènes *mat*, nous avons quantifié la transcription des gènes des type sexuels (*FPR1* chez *mat+* et *FMRI* chez *mat-*) et de quelques gènes cibles (Bidard et al. 2011) dans les souches sauvages et mutantes ΔPa_1_13940 (Tableau 14 et Annexe 10). La transcription de *FPR1* et de ses cibles a été comparée dans la souche mutante *mat+* ΔPa_1_13940 et la souche sauvage *mat+*, tandis que la transcription de *FMRI* et ses cibles a été mesurée dans la souche mutante *mat-* ΔPa_1_13940 et la souche sauvage *mat-*. Ces comparaisons indiquent que les gènes *FPR1* et *FMRI* ainsi que leurs gènes cibles (à deux exceptions près) sont moins transcrits en contexte ΔPa_1_13940 qu'en contexte sauvage (Tableau 14).

La différence d'expression des gènes *mat* entre un mutant ΔPa_1_13940 et une souche sauvage indique que le produit de *Pa_1_13940* active la transcription des gènes *mat*. En effet, *FPR1* est quatre fois moins transcrit dans une souche *mat+* ΔPa_1_13940 que dans une souche WT, et *FMRI* est dix fois moins transcrit dans une souche *mat-* ΔPa_1_13940 que dans une souche WT. Comme attendu, l'expression des gènes cibles est également diminuée, mais cette diminution peut être soit un simple effet de la diminution de l'expression des gènes régulateurs *FPR1* et *FMRI*, soit un effet plus complexe si *Pa_1_13940* coopère avec les gènes *mat* pour réguler ces cibles.

Les expériences de microarray effectuées sur le mycélium à 96 h et les périthèces à 30 h après fécondation montrent que le gène *Pa_1_13940* est transcrit au même niveau dans les deux souches *mat+* et *mat-* et que la mutation de *fprI* et *fmrI* n'influence pas la transcription de *Pa_1_13940* (résultats de F. Bidard et V. Berteaux-Lecellier non publiés).

Les analyses de RT-qPCR effectuées chez les mutants femelles stériles montrent que la transcription des gènes *mat* et de certaines de leurs cibles est perturbée dans chacun de ces mutants (Tableau 14). Ce résultat peut être expliqué par le fait que les gènes à HMG-box délétés dans ces souches mutantes appartiennent à un réseau de régulation impliquant Pa_1_13940. Cependant, la souche qui porte la délétion du gène *Pa_6_4110* montre une diminution significative des transcrits codant les récepteurs, bien que ce gène *Pa_6_4110* soit placé tout en bas de la voie de régulation. Une observation similaire est faite dans la souche qui porte une délétion dans le gène *Pa_7_7190*, alors que ce gène code un répresseur de la transcription de Pa_1_13940. Ces données signifient sans ambiguïté que Pa_6_4110 et Pa_7_7190 régulent l'expression de certaines des cibles des gènes *mat* sans passer par l'intermédiaire de Pa_1_13940.

	Nom	Fonction	Souche ΔPa_1_13940	Souche ΔPa_1_14230	Souche ΔPa_6_4110	Souche ΔPa_7_7190	La boîte TR
			<i>mat-</i>	<i>mat-</i>	<i>mat-</i>	<i>mat-</i>	
Gènes spécifiques <i>mat-</i>	<i>FMR1</i>	Mating-type regulation	down	down	down	up	+
	<i>mfm</i>	Pheromone precursor	down	down	down	up	-
	<i>PRE1</i>	Pheromone receptor	down	down	down	down	+
	<i>Pa_6_7350</i>	Protéase	down	down	up	up	-
Gènes spécifiques <i>mat+</i>			<i>mat+</i>	<i>mat+</i>	<i>mat+</i>	<i>mat+</i>	
	<i>FPR1</i>	Mating-type regulation	down	down	down	up	+
	<i>mfp</i>	Pheromone precursor	down	down	down	N/S	-
	<i>PRE2</i>	Pheromone receptor	down	down	down	down	-
	<i>Pa_4_3858</i>	unknown	down	down	down	N/S	-
	<i>Pa_1_24410</i>	SAM dependent methyl transferase	down	down	down	up	+
	<i>Pa_5_9770</i>	unknown	down	down	up	up	-
	<i>AOX</i>	alternative mitochondrial oxidase	down	up	N/S	N/S	-
	<i>Pa_4_3160</i>	phosphoenolpyruvate carboxykinase	N/S	down	N/S	up	-
	<i>Pa_4_80</i>	Methyltransferase	N/S	up	N/S	up	-

Tableau 14 : Résultats de RT-qPCR des gènes *mat* et certains de leurs cibles dans les mutants HMG.

N/S : p-value Non significative

Up: gène sur-exprimé dans le mutant par rapport au sauvage.

Down : gène sous-exprimé dans le mutant par rapport au sauvage.

+ Présence de la boîte TR dans la région 5'UTR du gène

- Absence de la boîte TR dans la région 5'UTR du gène

2.5 Caractérisation de *Pa_1_13940*

2.5.1 Structure de *Pa_1_13940*

Selon l'annotation du génome de *P. anserina*, le gène *Pa_1_13940* ne possède qu'un seul exon. Cependant une analyse détaillée de la séquence de ce gène prédit l'existence de deux introns absents dans l'annotation qui était disponible au moment où j'ai commencé l'analyse de ce gène. J'ai confirmé la présence de ces deux introns par RT-PCR (réverse transcription-PCR) et séquençage du ADNc (matériels et méthodes) (Figure 32). Le gène *Pa_1_13940* code une protéine de 559 acides aminés avec un seul domaine HMG entre le résidu 198 et le résidu 266 (Tableau 10).

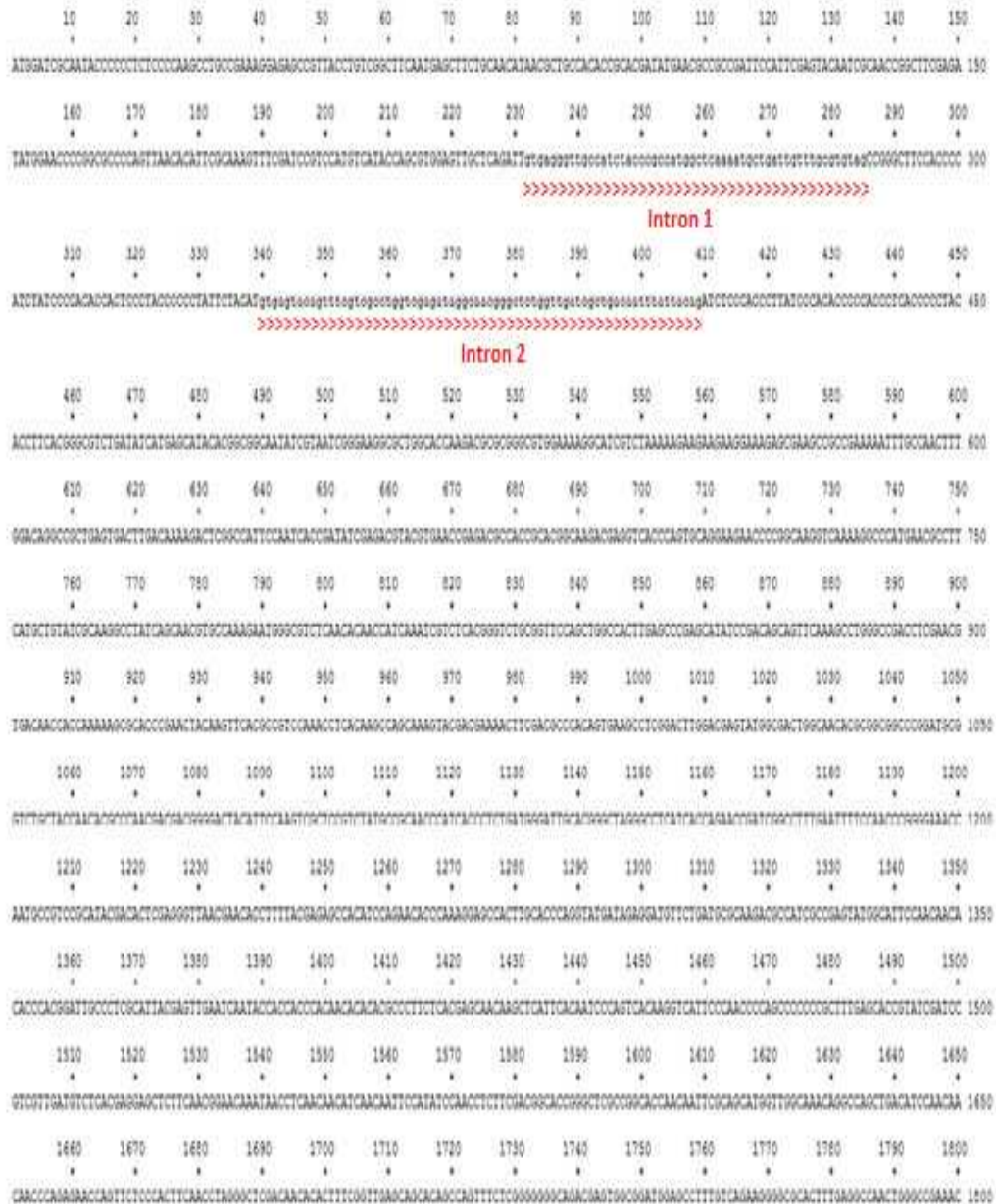
L'analyse phylogénique réalisée sur les domaines HMG de plusieurs espèces (Tableau 11 et Figure 28) a placé le domaine HMG de *Pa_1_13940* dans les MATA_HMG avec Ste11p de *S. pombe* (Voir chapitre I). Un alignement des trois protéines orthologues *Pa_1_13940*, Ste11p de *S. pombe* et FMF-1 de *N. crassa*, indique une forte conservation du domaine HMG (Figure 32), incluant le résidu Asn 6 présent dans les protéines à domaine HMG qui se fixent spécifiquement à l'ADN (Voir chapitre I, paragraphe 3.2.1, Figure 13, position 12). Sur Motif_scan http://myhits.isb-sib.ch/cgi-bin/motif_scan j'ai trouvé que *Pa_1_13940* possède une région N-terminale très riche en asparagine, un signal de localisation nucléaire bipartite (RRAGVEKASSKKKKKERA) et 14 sites de phosphorylation répartis le long de la protéine (Figure 33).

2.5.2 Le mutant ΔPa_1_13940

Dans les deux contextes *mat+* et *mat-* le mutant ΔPa_1_13940 est stérile en tant que mâle et en tant que femelle. Cette souche a une croissance végétative similaire à celle d'une souche sauvage et elle produit un nombre normal de spermaties et d'ascogones. Aucun développement de périthèces n'est observé quand le mutant est croisé à l'état homozygote, alors que dans un croisement hétérozygote $\Delta Pa_1_13940 \times WT$ quelques périthèces se développent à la confrontation entre les deux souches (Figure 31). Ces résultats suggèrent que le produit de *Pa_1_13940* présent dans le sauvage, diffuse dans le mutant ΔPa_1_13940 via des anastomoses et permet le développement sexué, sans que l'on puisse dire à ce stade si cette diffusion restaure la fonction des spermaties et/ou des ascogones mutants. Afin de vérifier que le phénotype de stérilité femelle observé dans le mutant ΔPa_1_13940 n'est pas dû à une anomalie dans la formation de l'enveloppe du périthèce, nous avons construit le tricaryon [Δmat , $\Delta Pa_1_13940 mat+$, $\Delta Pa_1_13940 mat-$] à partir d'un mélange des deux types sexuels *mat+* et *mat-* de la souche mutante ΔPa_1_13940 et la souche Δmat . L'absence totale de périthèces dans ce tricaryon suggère que le défaut de fertilité des souches mutantes

ΔPa_1_13940 n'est pas seulement un problème de formation d'enveloppe, qui serait complété par la souche Δmat . La perte de Pa_1_13940 affecte probablement la formation des tissus sexuels dont le développement est autonome.

A



B

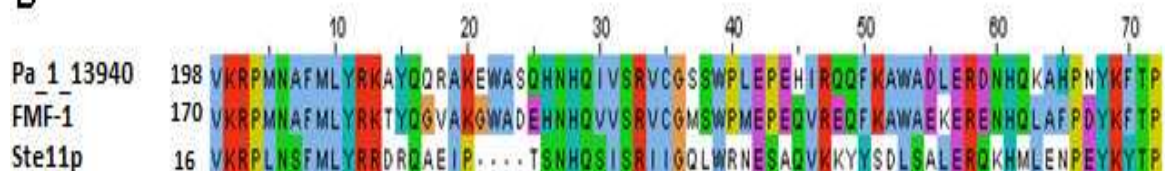
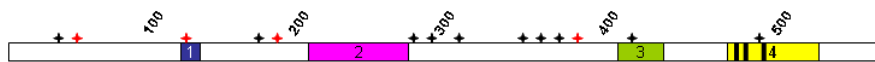


Figure 32: (A) Le gène Pa_1_13940 et la position des introns. (B) alignement des domaines HMG de Pa_1_13940 et ses orthologues FMF-1 de *N. crassa* et Ste11p de *S. pombe*. Le résidu 1 du domaine HMG correspondre aux résidus 198, 170 et 16 de la protéine intacte de Pa_1_13940 , FMF-1 et Ste11p respectivement. L'alignement est obtenu avec ClustalW2, la coloration est fournie par Jalview (Annexe 9).



Sites et domaine	Positions en acides aminés
ASN_GLYCOSYLATION	343 à 346
ASN_GLYCOSYLATION	465 à 468
ASN_GLYCOSYLATION	474 à 477
ASN_GLYCOSYLATION	490 à 493
CK2_PHOSPHO_SITE	31 à 34
CK2_PHOSPHO_SITE	171 à 174
CK2_PHOSPHO_SITE	289 à 292
CK2_PHOSPHO_SITE	310 à 313
CK2_PHOSPHO_SITE	345 à 348
CK2_PHOSPHO_SITE	358 à 361
CK2_PHOSPHO_SITE	368 à 371
CK2_PHOSPHO_SITE	411 à 414
CK2_PHOSPHO_SITE	459 à 462
CK2_PHOSPHO_SITE	555 à 558
PKC_PHOSPHO_SITE	40 à 42
PKC_PHOSPHO_SITE	137 à 139
PKC_PHOSPHO_SITE	182 à 184
PKC_PHOSPHO_SITE	377 à 379
ASN_RICH	465 à 520
HIS_RICH	405 à 433
HMG_BOX	198 à 266
NLS_BP	129 à 143

Figure 33 : Caractéristiques de la protéine Pa_1_13940.

(1) Le signal de localisation nucléaire bipartite (NLS_BP) ; (2) le domaine HMG ; (3) région riche en histidine (HIS_RICH) ; (4) Région riche en asparagine (ASN_RICH). Site de phosphorylation CK2 (étoile noire) ; Site de phosphorylation PKC (étoile rouge). ASN_glycosylation (traits noirs épais) Analyse réalisée par Motif_scan sur la base des données PROSITE http://myhits.isb-sib.ch/cgi-bin/motif_scan.

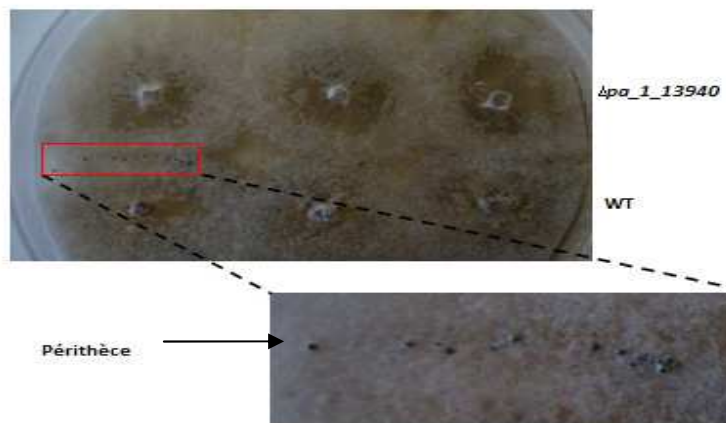


Figure 34 : Développement des périthèces au point de confrontation entre le mutant ΔPa_1_13940 et le sauvage.

2.5.3 Site de fixation potentiel du facteur de transcription codé par

Pa_1_13940

La protéine Stel1p de *S. pombe* se fixe spécifiquement aux TR_box AACAAAGAA présentes dans la région non traduite de la partie 5' (5'UTR) de ses gènes cibles (voir chapitre I) (van Beest et al. 2000). Chez *P. anserina* l'analyse des séquences 5'UTR du gène *Pa_1_13940*, des gènes des types sexuels et leurs cibles (Bidard et al. 2011) montre que le motif TR-box est présent en 5' de la phase codante de *Pa_1_13940*, *FPR1* et *FMRI* (Tableau 15). J'ai également détecté la TR-box en amont de sept gènes cibles de *FPR1* et de quatre gènes cibles de *FMRI* (Tableau 15), notamment *Pa_1_24410* et *Pa_7_9070* qui sont exprimés différemment entre les souches sauvages et les souches mutantes ΔPa_1_13940 (expérience de RT-qPCR Tableau 14). Par contre les gènes *mfm*, *mfp*, *PRE2*, *AOX*, *SAM*, *Pa_4_3858*, *Pa_5_9770* et *Pa_6_7350* ne possèdent pas de TR-box, bien que différemment exprimés entre le mutant ΔPa_1_13940 et la souche sauvage.

Afin de vérifier le rôle de la TR-box pour les gènes *FPR1* et *FMRI*, nous avons réalisé des intégrations ectopiques de *FPR1* et *FMRI* dépourvus du motif TR-box (*fprI*^{TR}, *fmrI*^{TR}) dans les souches mutantes *fprI*⁻ et *fmrI*⁻ (El-Khoury et al. 2008). Si la TR-box a un rôle essentiel dans la transcription de ces gènes, les copies ectopiques de *fprI*^{TR} et de *fmrI*^{TR} ne seront pas transcrites et la fertilité des souches *fprI*⁻ et *fmrI*⁻ ne sera pas restaurée. Nos expériences montrent sans ambiguïté une complémentation de la stérilité des souches *fprI*⁻ et *fmrI*⁻ par les copies ectopiques de *fprI*^{TR} et *fmrI*^{TR} ce qui suggère que la TR-box n'est pas indispensable pour la transcription de ces gènes.

Gène	Séquence	Position
<i>Pa_1_13940</i>	AGAAGCCGGA ACAAAGAA AGGAACAAG	-739
<i>FMRI</i>	AGGATAGCCA ACAAAGAA ATGACTGCC	-153
<i>FPR1</i>	AAGGTGAATA ACAAAGAA ATACACTGA	-371
<u>Gènes cibles de FPR1</u>		
<i>Pa_1_24410</i>	GTAAGTCTTA ACAAAGAA CGCTGATCC	-201
<i>Pa_1_9625</i>	GTGAAGTTCG ACAAAGAA GAGCTGGAG	-116
<i>Pa_6_3770</i>	GTATCCCGAA ACAAAGGA AGGGAGTAG	-865
<i>Pa_2_7170</i>	GAGACATTGA ACAAAGAG GCAAATAGA	-601
<i>Pa_4_860</i>	AAGCCTTGGT GCAAAGAA ATCTTTCAAA	-785
<i>Pa_5_3130</i>	TCGCCTCAAA GCAAAGAA AGTATGAATG	-115

<i>Pa_7_2860</i>	TGGGCCAGTG ACAAAGAA TGCTTTGTCG	-237
Gènes cibles de FMR1		
<i>Pa_7_9070</i>	CCTGGCATCA ACAAAGACCGGCGGGCC	-463
<i>Pa_0_1140</i>	TGACAGAACA ACAAAGTGA ATTTCATT	-812
<i>Pa_5_10940</i>	CCCGAAAAAG ACAAAGCAC GCACGCAT	-17
<i>Pa_6_10330</i>	TAAGACCAAA ACAAAGCCA TCAAAAAGTC	-658

Tableau 15: La TR-box est présente dans les régions 5'UTR (1000pb avant l'ATG) de *Pa_1_13940*, des gènes des types sexuels et des cibles de FMR1 et FPR1. La numérotation des nucléotides commence par -1 au premier nucléotide avant l'ATG du début de la CDS.

3. Discussion

3.1 La majorité des protéines à domaine HMG sont impliquées dans le développement sexué de *Podospira anserina*

Le génome de *P. anserina* code douze protéines à domaines HMG, dont huit identifiées dans cette étude par similarité de séquence avec *N. crassa*. Quatre gènes ont été analysés dans des études précédentes (Debuchy and Coppin 1992; Dequard-Chablat and Alland 2002). Trois de ces gènes ont un rôle direct dans la reproduction sexuée, puisque ce sont les gènes des types sexuels. Le quatrième gène, mtHMG1, est également impliqué dans la reproduction sexuée, mais il s'agit probablement d'un rôle indirect puisqu'il a une localisation mitochondriale. L'analyse du phénotype des délétions des huit gènes HMG identifiés dans cette étude nous a permis de mettre en évidence trois catégories de phénotypes touchant la reproduction sexuée:

- un phénotype de stérilité femelle observée chez ΔPa_6_4110 , ΔPa_7_7190 et ΔPa_1_14230 . Il faut noter que le phénotype de stérilité femelle est observé fréquemment chez *P. anserina*. Dans quelques cas, il peut être attribué à un effet direct, comme par exemple la mutation des récepteurs de phéromones *pre1* et *pre2* (Bidard et al. 2011) ou l'absence d'un ou plusieurs facteurs majeurs de transduction du signal, par exemple les mutants des gènes *PaNOX1*, *PaASK1* et *PaMKK1* (Kicka and Silar 2004; Brun et al. 2009). Dans les autres cas, il s'agit de perturbations du métabolisme général qui affectent de manière indéterminée et indirecte le développement de l'organe femelle (par exemple les mutants *leu1* auxotrophes pour la leucine).
- un phénotype de stérilité mâle et femelle, chez le mutant ΔPa_1_13940 . Ce phénotype n'a été observé que chez les mutants des mating types. Nous examinerons en détail ce mutant dans la section suivante.

- la production des spermaties est aussi l'un des processus affectés par la délétion des gènes à HMG-box. Les mutants ΔPa_7_7190 et ΔPa_1_11050 produisent plus de spermaties que le sauvage.

Finalement, seuls trois gènes sur douze ne semblent pas impliqués dans la reproduction sexuée chez *P. anserina*. Il s'agit de *Pa_1_7390*, *Pa_1_9380* et *Pa_5_8400*. L'analyse de leurs orthologues chez *S. cerevisiae* et *S. pombe* est susceptible de fournir des indications supplémentaires sur un rôle éventuel de ces gènes dans la reproduction sexuée. *Pa_1_7390* ne possède pas d'orthologues chez ces deux levures. *Pa_1_9380* possède un orthologue chez *S. pombe* (*lsd1* ou *SPBC146.09c*) alors qu'aucun orthologue n'est détecté chez *S. cerevisiae* par FUNGIpath. Le gène *Splsdl* code une histone méthylase qui régule l'extension de l'hétérochromatine aux centromères et au locus *MAT* via la déméthylation de la lysine 9 de l'histone H3 (Lan et al. 2007). Les souches de *S. pombe* portant une délétion de *Splsdl* ont une croissance végétative fortement réduite (Nicolas et al. 2006; Lan et al. 2007), alors que la délétion de *Pa_1_9380* n'entraîne aucun phénotype végétatif. Les diploïdes *splsdl*^{+/-} ont des défauts de sporulation, un très petit nombre de diploïdes produisant des spores viables, ce qui suggère que LSD1 a un rôle lors de la méiose (Lan et al. 2007). Là encore, aucune altération de ce type n'est détectable chez les mutants de *P. anserina*. L'ensemble de ces données suggère que LSD1 et *Pa_1_9380* ont des fonctions différentes, ou, plus probablement, que des défauts dans la gestion de l'hétérochromatine ont moins de répercussion chez *P. anserina* que chez *S. pombe*. Il serait intéressant de ce point de vue de confirmer la fonction de *Pa_1_9380* chez *S. pombe*. *Pa_5_8400* est l'orthologue de *SPBC19G7.04* chez *S. pombe*, tandis que chez *S. cerevisiae* aucun orthologue de *Pa_5_8400* n'est détecté par FUNGIpath. *SPBC19G7.04* code un facteur de transcription régulateur du cycle cellulaire. Cette protéine est localisée dans le noyau et au niveau des fuseaux mitotiques et son pic d'activité est observé au début de la phase M (Bushel et al. 2009). Il n'y a aucune indication d'un rôle de cette protéine au cours du cycle sexué de *S. pombe*.

J'ai également recherché et examiné les orthologues des protéines pour lesquelles nous avons détecté un rôle dans la reproduction sexuée. *Pa_7_7190* et *Pa_1_11050* n'ont pas d'orthologues chez *S. cerevisiae* ni chez *S. pombe*. *Pa_1_14230* possède un orthologue, NHP6A, et deux inparalogues, NHP6B et HMO1, chez *S. cerevisiae*. NHP6A et B sont impliqués dans l'organisation de la chromatine et la régulation des gènes appartenant à différents processus cellulaires. En effet, les délétions de *NHP6A* et *NHP6B* affectent l'expression de 197 gènes, et plus particulièrement les gènes impliqués dans le développement sexué de *S. cerevisiae* (exemple *ASG7*, *MFA1* et *AGA1*), les gènes régulés par les phéromones (exemple *FUS2* et *FIG2*) et les gènes de la biosynthèse de la méthionine

(exemple *MET28* et *SUL1*) (Moreira and Holmberg 2000). En dépit de ces nombreuses perturbations, les doubles mutants *nhp6a nhp6b* sont capables de former des projections de conjugaison (shmoo) normales, et de faire un cycle sexué normal en croisement hétérozygote (Costigan et al. 1994). Il n'y a pas d'indication disponible sur un éventuel croisement homozygote des doubles mutants. Enfin, *Pa_6_4110* est l'orthologue de *ROX1* (repressor of hypoxic genes) de *S. cerevisiae*. FUNGIpath n'identifie aucun orthologue de *Pa_6_4110* chez *S. pombe*. *ROX1* est un facteur de transcription répresseur des gènes qui permettent à *S. cerevisiae* d'utiliser l'oxygène en condition partiellement anaérobie (Balasubramanian et al. 1993). La délétion de ce gène chez *S. cerevisiae* cause un retard des cellules en G1 (White et al. 2009). Aucune indication d'un rôle éventuel dans la reproduction sexuée n'est disponible pour *ROX1*.

3.2 ΔPa_1_13940 possède un phénotype particulier de développement sexué

Plusieurs mutants bloqués dans le processus de développement sexué ont été identifiés chez *P. anserina*. Dans la plupart des cas, le développement sexué est bloqué seulement si le mutant est utilisé comme parent femelle. Seules les mutations au niveau du locus *mat* peuvent engendrer une stérilité à la fois en temps que parent mâle et parent femelle.

Pendant cette étude nous avons montré que l'inactivation du gène à HMG-box *Pa_1_13940* provoque aussi une double stérilité mâle et femelle. La mutation ΔPa_1_13940 a une croissance végétative et une production des organes sexués semblables au sauvage mais dans des croisements homozygotes ou hétérozygotes il ne développe pas de périthèces.

En 1979, Thomas E. JOHNSON (Johnson 1979) a décrit un mutant de *N. crassa* *fmf-1* (female and male fertility-1) qui porte une mutation dans un gène qui a été identifié ultérieurement comme l'orthologue de *stell* de *S. pombe* (Iyer et al. 2009). Dans un croisement entre le mutant *fmf-1* et le sauvage (*fmf-1* x *fmf-1*⁺), le développement des périthèces est bloqué précocement. Leur taille atteint 40% de la taille d'un périthèce sauvage mature. Cet arrêt de développement se produit dans tous les cas, que le mutant participe comme parent mâle et le sauvage comme parent femelle ou inversement, En revanche, le croisement *fmf-1* x *fmf-1*⁺ est fertile si le noyau mutant *fmf-1* est issu d'un hétérocaryon *fmf-1* + *fmf-1*⁺, il est donc hétérocaryon récessif. L'utilisation de la mutation *tolerant* (*tol*) pour supprimer l'incompatibilité des types sexuels dans l'hétérocaryon montre que [(*fmf-1*⁺ *tol* A) + (*fmf-1*; *tol* a)] est autofertile tandis que [(*fmf-1* *tol* A) + (*fmf-1*; *tol* a)] est auto stérile. Cela suggère que l'incompatibilité des types sexuels dans l'hétérocaryon empêche la diffusion de la protéine FMF-1 entre les mycéliums *mat* A et *mat* a.

La première différence observée entre *P. anserina* et *N. crassa* est l'absence totale d'initiation du développement des périthèces dans un croisement hétérozygote impliquant ΔPa_1_{13940} , et le début de développement observé dans un croisement impliquant *fmf-1*. Cette différence implique clairement que la fécondation a lieu chez *N. crassa*. Pour *P. anserina*, deux hypothèses sont envisageables. La première suggère que la fécondation n'a pas eu lieu car la souche ΔPa_1_{13940} produit des organes reproducteurs incapables de féconder à cause de la faible expression des phéromones et leurs récepteurs. La deuxième hypothèse propose que la fécondation a eu lieu mais le développement des périthèces n'est pas visible chez *P. anserina* par rapport à *N. crassa*. Afin de choisir entre ces deux hypothèses, on pourrait envisager des expériences de deuxième fécondation avec une souche sauvage. Par exemple, la fécondation d'une souche sauvage par les spermaties d'une souche ΔPa_1_{13940} peut être effective, mais le développement du périthèce rapidement arrêté après la fécondation. Une telle souche femelle ne devrait plus être fécondable par les spermaties d'une souche mâle sauvage. *A contrario*, si les spermaties d'une souche ΔPa_1_{13940} sont incapables de féconder une souche WT, les organes femelles de celle-ci sont disponibles pour une fécondation par les spermaties d'une souche WT de type sexuel complémentaire.

La deuxième différence observée entre *P. anserina* et *N. crassa* concerne la complémentation de la mutation ΔPa_1_{13940} ou *fmf-1* par un noyau sauvage. L'efficacité de cette complémentation a été analysé dans différents procédés expérimentaux chez *P. anserina* et *N. crassa*. Chez *P. anserina*, cette complémentation est beaucoup plus difficile que chez *N. crassa*, voire même impossible dans certains schéma expérimentaux. Nous avons vu (paragraphe 2.4.2 ci-dessus) que cette absence de complémentation suggère que la protéine Pa_1_13940 serait nécessaire pour la formation des tissus sexués du périthèce. A l'inverse, la complémentation observée chez *N. crassa* suggère que la protéine FMF-1 n'est pas nécessaire pour le développement des tissus sexués. Donc les orthologues de SpSte11p auraient des fonctions différentes chez *P. anserina* et *N. crassa*.

3.3 Pa_1_13940 régule le développement sexué chez *Podospora anserina*

Le gène *Pa_1_13940* code une protéine à domaine MATA_HMG, identifiée comme l'orthologue de Ste11p de *S. pombe* (Leupold et al. 1991). Le facteur de transcription Ste11p contrôle le développement sexuel de *S. pombe* via l'activation des gènes des types sexuels et les gènes de la méiose (Sugimoto et al. 1991; Mata and Bähler 2006) (Chapitre I) et, il est possible que Pa_1_13940 et SpSte11p partagent le même rôle.

La faible transcription des gènes *mat* (*FPR1* dans *mat+* et *FMRI* dans *mat-*) et de certaines de leurs cibles dans un mutant ΔPa_1_13940 montre que, comme SpSte11p, Pa_1_13940 régule l'expression des gènes *mat* (*FPR1* et *FMRI*), qui à leur tour régulent directement ou indirectement (en coopération avec Pa_1_13940) les gènes impliqués dans le développement sexué de *P. anserina*. Ainsi, dans un contexte *mat+*, la protéine Pa_1_13940 active l'expression de *FPR1* qui à son tour active l'expression des gènes spécifiques *mat+* et inhibe l'expression des gènes spécifiques *mat-*. Dans le cas d'une souche *mat-*, Pa_1_13940 active l'expression de *FMRI* qui à son tour active l'expression des gènes spécifiques *mat* et inhibe l'expression des gènes spécifiques *mat+*.

La découverte d'une TR-box en amont de *FPR1*, *FMRI*, de certains de leurs gènes cibles et de *Pa_1_13940* suggère que Pa_1_13940 se fixe spécifiquement à l'ADN via le même motif que celui utilisé par SpSte11p. Cependant les copies ectopiques *FPR1* et *FMRI* dépourvues de la TR-box restaurent la stérilité des mutants *fprI* et *fmrI* respectivement et sont donc toujours fonctionnelles. Pour expliquer ce résultat deux hypothèses sont possibles :

- La TR-box est utilisée par Pa_1_13940 mais son absence n'empêche pas l'expression basale de *FPR1* et *FMRI* qui est suffisante pour restaurer la fertilité des mutants *fprI* et *fmrI* respectivement.
- La TR-box n'est pas utilisée par Pa_1_13940 pour réguler les gènes *mat*. La stratégie de régulation adoptée par Pa_1_13940 chez *P. anserina* est peut être plus complexe que celle de SpSte11p chez *S. pombe* ; il est possible que Pa_1_13940 reconnaisse un autre motif spécifique ou régule indirectement *FPR1* et *FMRI*.

3.4 Les protéines à domaine HMG forment un réseau de regulation complexe

Des régulations impliquant des cascades de protéines à domaine HMG ont été identifiées chez plusieurs espèces, notamment la régulation des gènes à HMG-box *mat-Mc* et *mat-Pc* par le facteur de transcription à domaine HMG Ste11p chez *S. pombe* (voir partie 1 paragraphe 3.2.3.4.2). Chez *Ustilago maydis*, le facteur de transcription à domaine HMG Rop1 active l'expression du gène *prf1* qui code également un facteur de transcription à domaine HMG impliqué dans la régulation de l'expression des gènes *mat*, des phéromones et leurs récepteurs (Brefort et al. 2005). Chez les mammifères, le facteur de transcription à domaine HMG SRY active l'expression de *SOX9* à domaine HMG impliqué dans la différenciation de l'appareil génital mâle (Kashimada and Koopman 2010). Chez *P. anserina*, nous avons identifié un réseau de régulation complexe entre plusieurs protéines à domaine HMG. *Pa_1_13940* est régulé positivement par Pa_1_14230 et négativement par Pa_7_7190.

Pa_1_13940 régule la transcription de *Pa_6_4110* et des gènes codant les facteurs de transcription FPR1 et FMR1 (Figure 35). Ces derniers contrôlent à leur tour la transcription des gènes impliqués dans le développement sexué de *P. anserina*. *Pa_7_7190* et *Pa_6_4110* ont aussi un rôle dans la régulation de certains gènes cibles de FPR1 et FMR1, indépendamment de *Pa_1_13940* et des gènes *mat*. Parmi ces gènes on trouve ceux qui codent les récepteurs des phéromones, ce qui explique le phénotype de la stérilité femelle chez les mutants ΔPa_6_4110 et ΔPa_7_7190 .

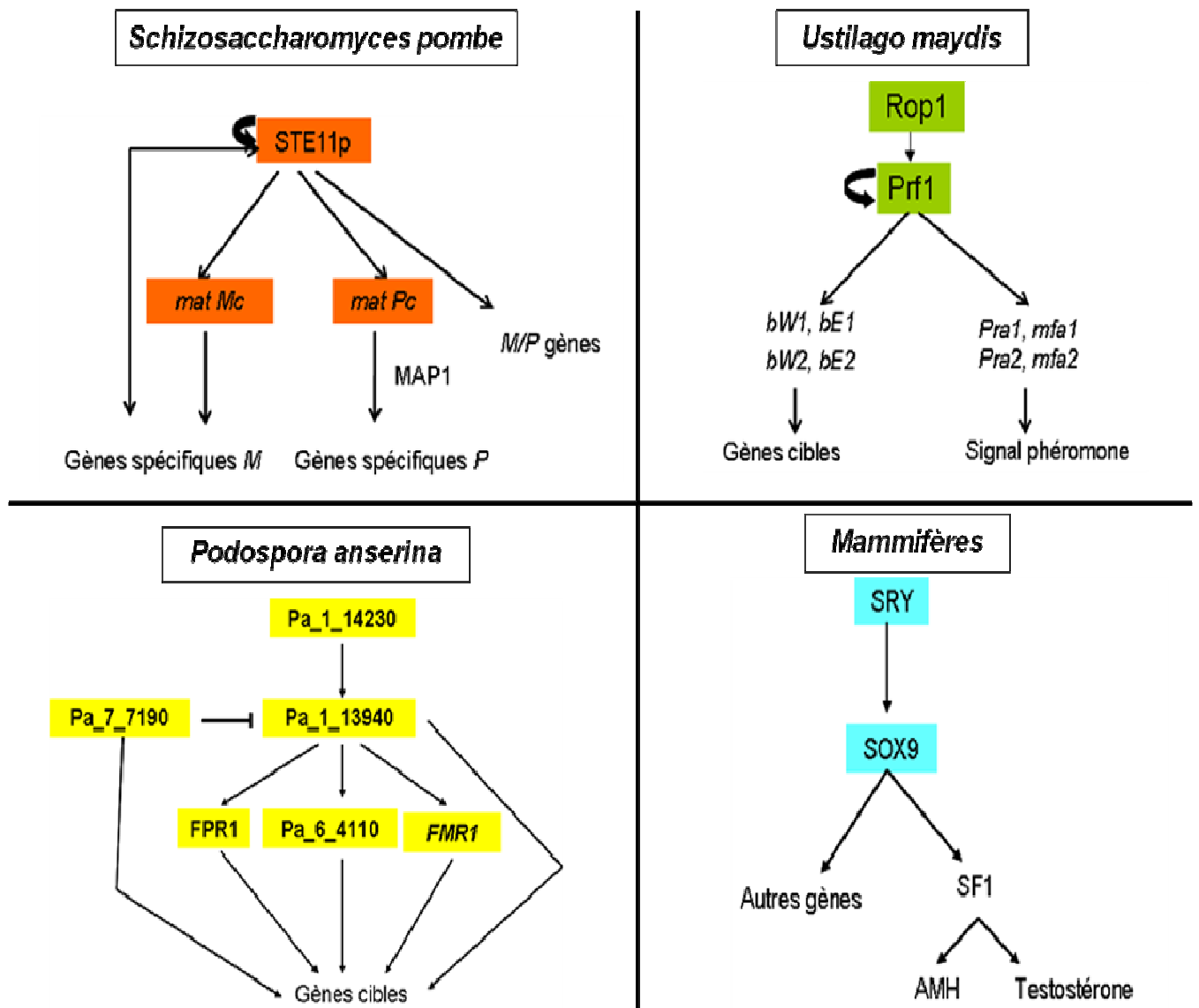


Figure 35: Réseaux de régulation entre les protéines à domaine HMG chez *S. pombe*, *U. maydis*, les mammifères et *P. anserina*.

4. Conclusion

Dans cette étude, nous avons montré que les gènes à HMG-box chez *P. anserina* ont un rôle très important dans le développement sexué de cette espèce, spécialement *Pa_1_13940* le premier acteur de la régulation des gènes nécessaires à la reproduction sexuée. Ce gène code un facteur de transcription qui semble jouer le même rôle que Ste11p de *S. pombe*. Les orthologues de *Pa_1_13940* chez les champignons filamenteux sont très peu étudiés. Nous avons aussi démontré que *Pa_1_13940* a des liens de régulation avec trois autres gènes à HMG-box dont la délétion induit une stérilité femelle chez *P. anserina*. Il est donc nécessaire de poursuivre l'analyse afin de mieux comprendre les mécanismes par lesquels les protéines à domaine HMG régulent le développement sexué chez *P. anserina*. Par exemple l'étude de la cinétique d'expression de *Pa_1_13940* durant le cycle de vie de *P. anserina* permettrait de réaliser des analyses transcriptomiques à des moments précis du cycle et d'identifier les cibles de *Pa_1_13940*, les expériences de retard sur gel permettrait de déterminer le site de fixation de *Pa_1_13940* et trier ses cibles, et l'utilisation de la technique du double-hybride permettrait de vérifier l'interaction entre *Pa_1_13940* et les autres protéines à domaine HMG notamment les facteurs de transcription codés par les gènes *mat*. Enfin, il est aussi nécessaire de préciser si les mutants ΔPa_1_13940 sont bloqués pendant ou après la fécondation.

Conclusion générale

Au cours de ma thèse, j'ai identifié les cibles directes ou indirectes des facteurs de transcription MAT qui sont impliquées dans la reconnaissance intercellulaire et la reconnaissance internucléaire entre *mat+* et *mat-*. Mes résultats montrent une complexité inattendue des circuits de régulation du développement sexué chez *P. anserina*. L'analyse de la fécondation, révèle l'existence de huit possibilités de régulation par les facteurs de transcription FPR1 et FMR1 (activation ou répression). Dans le cas de la reconnaissance internucléaire, le système de régulation est encore plus complexe que celui de la fécondation. Les facteurs de transcription FPR1, FMR1 et SMR2 forment probablement un complexe pour réguler les gènes impliqués dans la formation des hyphes ascogènes. La majorité des gènes cibles identifiés interviennent dans le métabolisme général et n'ont aucune relation avec le développement sexué, contrairement à *S. pombe* et *S. cerevisiae*, chez lesquels toutes les cibles des facteurs de transcription MAT sont impliquées dans la reproduction sexuée.

Il faut compléter les expériences de microarray par des analyses *in silico* afin de chercher le site de fixation des facteurs de transcription MAT et trier leurs cibles directes et indirectes. En plus pour le cas de la reconnaissance internucléaire il faut utiliser un matériel biologique plus spécifique à ce stade, c'est à dire extraire le contenu de périthèces où se déroule la reconnaissance entre un noyau *mat+* et un noyau *mat-*. Je propose également d'étudier la relation entre FPR1, FMR1 et SMR2 pour vérifier la possibilité de la formation d'un complexe hétérotrimérique.

La recherche des gènes impliqués dans le développement sexué parmi les gènes à HMG-box, nous a permis d'identifier pour la première fois chez *P. anserina* le régulateur des gènes *mat*. Pa_1_13940 est un facteur de transcription qui semble jouer le même rôle que son orthologue Ste11p de *S. pombe*. Les orthologues de Pa_1_13940 chez les champignons filamenteux sont très peu étudiés. Il est nécessaire de continuer la recherche des cibles de ce facteur de transcription par des approches microarray ou par des analyse *in silico* en identifiant son site de fixation. Les cibles de Pa_1_13940 peuvent être des acteurs essentiels du développement sexué chez les champignons filamenteux. En outre, l'étude de Pa_1_13940 nous permettra de compléter le schéma de régulation du développement sexué par les gènes *mat*. Enfin, il faut poursuivre la dissection de la fonction des protéines à domaine HMG, plus particulièrement celles qui causent la stérilité femelle.

Annexes

Annexe 1: Composition de la référence utilisée dans les expériences de microarray

- Croisement à 96 h

1. Croissance de 96 h d'une souche *mat+* ou *mat-* sur M2+toile à bluter
2. Arrosage avec des microconidies de la souche de signe opposé récoltés par lavage d'une culture de 120 h de la souche de type sexuel opposé
3. Incubation pendant 96 h après arrosage
4. Prélèvement du matériel biologique

- Croisement à 48 h

- 1 et 2. Comme précédemment
3. Incubation pendant 48h après arrosage
4. Prélèvement du matériel biologique

- Mycélium à 48 h

1. Croissance de 48 h d'une souche *mat+* ou *mat-* sur M2+cellophane
2. Prélèvement du matériel biologique

- Mycélium à 96 h

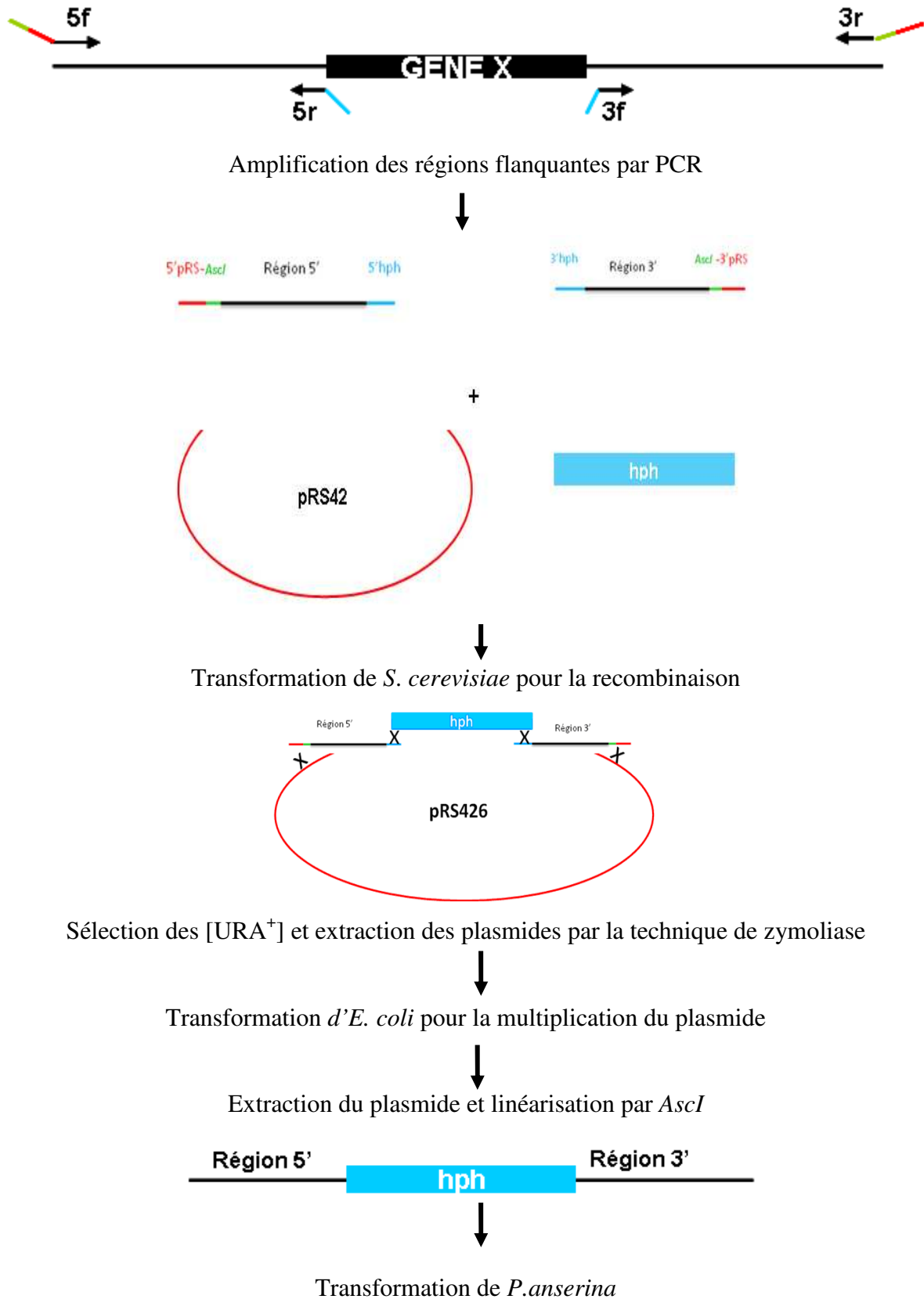
- 1 et 2 Comme précédemment avec une croissance de 96 h

Annexe 2: Protocole de préparation des ARN (Méthode QIAGEN)

1. Placer le mycélium congelé dans le cryotube Nalgene avec la bille d'acier (BBI-8546916, Sartorius) refroidie à l'azote liquide, refermer et placer dans le container à cryotube du Mikro-Dismembrator (Sartorius) également refroidi à l'azote liquide.
2. Broyer 1 min à 2600 rpm.
3. Transférer dans le cryotube Nalgene 700µl de RLT (RNeasy Lysis Buffer) + 7µl de mercaptoethanol et laisser décongeler.
4. Vortexer le mélange tampon + poudre de mycélium puis transférer sur la colonne QIAshredder (permet de casser l'ADN et d'éliminer les fragments de mycélium).
5. La colonne est centrifugée 2 min à 10000 rpm.
6. L'éluât est récupéré (attention car le mycélium résiduel forme un culot qu'il ne faut pas pipeter) et transféré dans un tube RNase-free + 350 µl d'éthanol.
7. Mélanger immédiatement par pipetages.
8. Déposer la moitié du mélange sur une colonne RNeasy et centrifuger 1 min à 10000 rpm.
9. Eliminer l'éluat puis répéter l'étape 7 avec le reste du mélange.
10. Laver la colonne avec 350 µl de tampon RW1 (washing buffer).
11. Centrifuger 1 min à 12000rpm.
12. Traitement à la DNase : dans un tube RNase-free, préparer 10 µl de DNase + 70 ul de tampon RDD, déposer ce mélange sur le centre de la colonne et incubé 20 min à 28°C.
13. Ajouter 350 µl de tampon RW1.
14. Centrifuger 1 minute à 12000 rpm.
15. Eliminer le surnageant.
16. Laver avec 500 ul de tampon RPE.
17. Attendre 2 min avant centrifugation (améliore l'étape de lavage).
18. Centrifuger 2 minutes à 12000 rpm.
19. Répéter le lavage (améliore la qualité des ARNs).
20. Pour finir centrifuger 1 min à 12000 rpm afin d'assécher la colonne.
21. Transférer la colonne dans un tube RNase-free puis éluer avec 2 fois 50 µl d'eau DEPC.
22. Transférer les 100 µl d'éluat dans deux tubes RNase free : 60 µl dans un tube pour congélateur à -80°C (stock) et 40 µl pour -20°C (pour test).

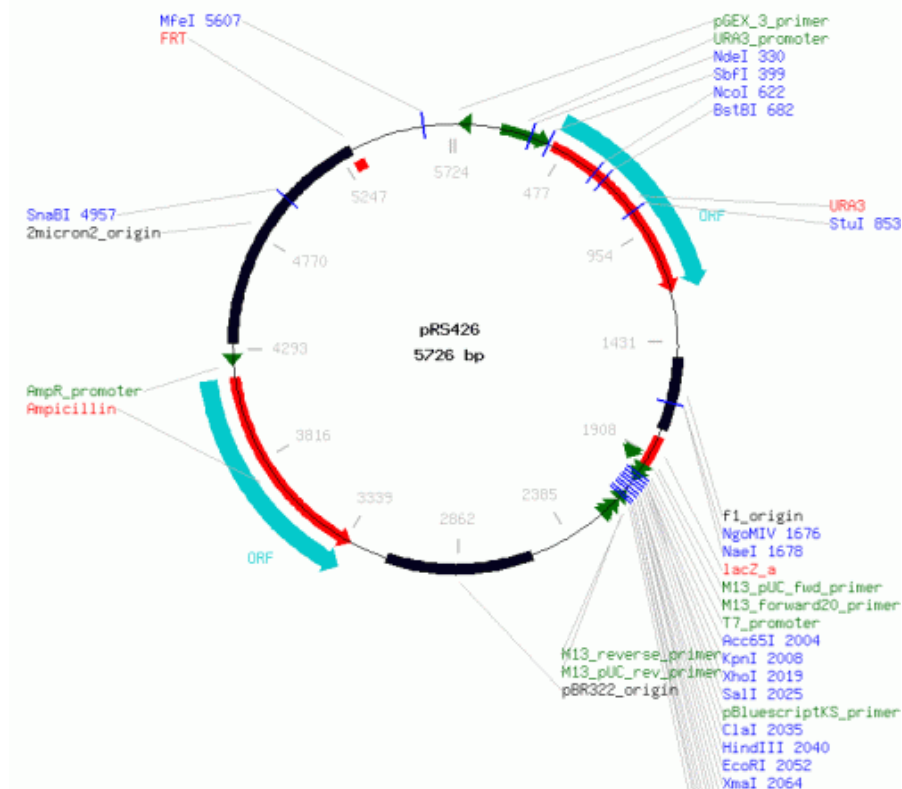
Annexe 3 : La technique de délétion des gènes chez *P. anserina*

1. Principe de la méthode (Colot et al. 2006)



2. Le plasmide pRS426-2 μ

Un vecteur navette *E. coli*/*S. cerevisiae* contenant le marqueur $URA3$, la cassette de résistance à l'ampicilline (Amp^R) et le gène *lacZ* qui code la β -galactosidase.



Carte génétique du plasmide pRS426

3. Dessin des amorces

Dessiner les amorces dans les régions 3' et 5' (20 pb) du gène à déléter de façon à obtenir des fragments de 1 à 1.3kb (voir principe de la méthode).

Pour chaque amorce spécifique d'un gène on ajoute une région 5' commune pour toutes les délétions

5f: **GTAACGCCAGGGTTTTCCAGTCACGACGGCGCGCC**+ séquence 5f de la partie à déléter

5r: **CAGCAGCCGCACGCGGCATCTCGGAAC**+séquence 5r de la partie à déléter

3f: **TCCTTCTTTCTAGCTAGAGGATCCTCTAC**+séquence 3f de la partie à déléter

3r: **GCGGATAACAATTCACACAGGAAACAGCGGGCGCGCC** séquence 3r de la partie à déléter

La séquence en vert présente le site de restriction de l'enzyme *AscI* qui va nous permettre de lineariser la cassette de délétion

Les séquences en rouge sont complémentaires au plasmide pRS426.

Les séquences en bleu sont complémentaires à la cassette *hph*

NB : L'orientation de la phase codante de *hph* est identique à celle du gène à déléter

4. Amplification des régions flanquantes par PCR

Préparer deux tubes pour chaque délétion (un pour la partie 5' et un pour la partie 3')

Produits	Réaction de 50 μ l
Tampon <i>Pfu</i> 10X	5 μ l
H ₂ O	5 μ l
dNTPs 1 mM	10 μ l
Amorce spécifique du gène 5r (ou 3r) 5pmol/ μ l	10 μ l
Amorce spécifique du gène 5f (ou 3f) 5pmol/ μ l	10 μ l
Plasmide qui contient le gène 1 ng/ μ l	10 μ l
ADN polymérase <i>Pfu</i> 1U/ μ l (Promega)	1 μ l

Programme de PCR 94°C 2 min

25 \leq cycles {

- 94°C, 30 sec
- 55°C, 30 sec
- 68°C, x min (2 min par kb)

68°C, 5 min

5. Amplification de la cassette *hph*

Produits	Réaction de 50 μ l
Tampon <i>Pfu</i> 10X	5 μ l
H ₂ O	5 μ l
dNTPs 1 mM	10 μ l
Amorce <i>hphf</i> 5pmol/ μ l	10 μ l
Amorce <i>hphr</i> 5pmol/ μ l	10 μ l
Plasmide pMOcosx 1ng/ μ l	10 μ l
ADN polymérase <i>Pfu</i> 1U/ μ l (Promega)	1 μ l

- *hphf* GTTCCGAGATGCGCCGCGTGC
- *hphr* GTAGAGGATCCTCTAGCTAG

Programme de PCR:

94°C 2 min

{

- 94°C, 30 sec

- Compléter à 100 ml avec de l'eau.
- Filtrer avec un filtre de 0.2 μm .

Mélange 2 :

- Glucose 10g
- Bacto agar pour le milieu solide 7.5g
- Eau 450ml
- Autoclaver 5 min à 120.
- Ajouter 50 ml du Mélange 1.

8. Transformation des levures

8.1 Milieux

- YPG Liquide
- Milieu sélectif SC-Ura solide

8.2 Solutions

- Solution filtrée d'acétate de lithium 1 M (LiAc)
- Solution filtrée d'acétate de lithium 0,1 M (LiAc)
- Solution PEG 3350 50 %
- ADN entraîneur à 5 mg/ml (sheared salmon sperm DNA (2 mg/ml))

8.3 Préparation des cellules compétentes

1. Ensemencer la souche de levure W303 (voir matériels et méthodes) dans 5 ml de milieu YPG et incubé à 30°C avec agitation pendant la nuit.
2. Diluer la culture de la nuit dans 50 ml de milieu YPG et cultiver à 30°C sous agitation jusqu'à une densité cellulaire de 1.
3. Lorsque la $DO_{600}=1$, centrifuger la culture pendant 5 min à 4000 rpm dans un tube Falcon.
4. Éliminer le surnageant
5. Effectuer 1 lavage en resuspendant délicatement le culot à la pipette dans 25 ml d'eau stérile.
6. Centrifuger pendant 5 min à 4000 rpm.
7. Éliminer le surnageant.
8. Resuspendre le culot dans 1 ml de la solution d'acétate de lithium 0.1 M.
9. Transférer dans un tube 1.5 ml.
10. Centrifuger pendant 15 sec à 13000 rpm.
11. Éliminer le surnageant.
12. Resuspendre le culot dans 400 μl de la solution d'acétate de lithium 0.1 M.

8.4 Transformation des levures

1. Pour chaque transformation transférer 50 µl de cellules compétentes dans une tube de 1,5 ml inclure un témoin positif (plasmide réplcatif de levure portant le même marqueur de sélection que celui des transformations à effectuer) et un témoin négatif sans ADN.
2. Centrifuger pendant 15 sec à 13000rpm.
3. Eliminer le surnageant.
4. Dans chaque tube ajouter
 - PEG 3350 50% 240 µl
 - Acétate de lithium 1 M 36 µl
 - Sheared salmon sperm DNA (2 mg/ml) 50 µl
 - H₂O stérile 34-x µl
 - (x) est le volume total de plasmide pRS426 linéaire (2µl), cassette *hph* (2µl) et les fragments 3' (2µl) et 5' (2µl) amplifiés par PCR.

NB : on n'est pas besoin de purifier les produits de PCR pour faire la transformation de levure.
5. Mélanger soigneusement au Vortex.
6. Incuber à 30°C pendant 30 min sans agiter (chambre de culture 30°C)
7. Mélanger en inversant les tubes
8. Effectuer un choc thermique en plaçant les tubes 30 min à 42°C.
9. Centrifuger pendant 15 sec à 13000 rpm.
10. Enlever 50 % de surnageant et resuspendre le culot dans le reste du surnageant.
11. Étaler les cellules avec des billes de verre sur 2 boîtes de milieu sélectif.
12. Incuber les boîtes à 30°C pendant 48 à 72 h.

9. Extraction du plasmide

9.1 Solutions

Solution de zymoliase : pour un volume total de un 1 ml

- Tris Hcl 1 M pH8 50 µl
- EDTA 0.5 M pH8 20 µl
- β-mercaptoethanol 3 µl
- Zymoliase conservé à 4°C 1mg
- H₂O 927µl

Préparer une solution de 10 ml qui peut se conserver un mois à 4°C

Solution d'acétate d'ammonium 7.5M :

- Acétate d'ammonium 28.91 g

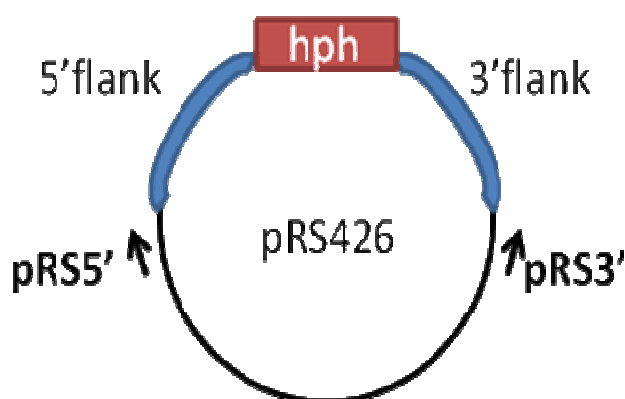
- H₂O

50 ml

Solution SDS 10 %

9.2 Protocoles

1. Mettre en suspension une colonie de levure dans 5 ml de milieu SC-Ura et incuber à 30°C avec agitation pendant toute la nuit (pour chaque transformation analyser 4 colonies)
2. Transférer 2 ml de la culture dans un tube Eppendorf 2 ml.
3. Centrifuger pendant 2 min à 8000 rpm.
4. Eliminer le surnageant puis répéter les étapes 2,3.
5. Resuspendre le culot dans la solution de zymoliase (150 µl).
6. Incuber à 37°C pendant 1 h.
7. Ajouter 20 µl de SDS 10% et 100 µl d'acétate d'ammonium 7.5 M.
8. Vortexer et incuber 15 min à -80°C.
9. Centrifuger 15 min à 15000 rpm.
10. Transférer le surnageant dans un tube 1.5 ml et ajouter 0.6 volume d'isopropanol (150µl).
11. Vortexer et incuber 15 min à -80°C.
12. Centrifuger 15 min à 15000 rpm.
13. Eliminer le surnageant et mettre le culot en solution dans H₂O 100 µl.
14. Traiter à la RNase, puis reprécipiter avec l'éthanol + acétate d'ammonium.
15. Vortexer, centrifuger, éliminer le surnageant et mettre le culot en solution dans H₂O 20 µl.
16. Analyser par PCR pour identifier les transformants positifs (1 µl de la solution).



Produits	Réaction de 25 µl
Tampon de la <i>Dream Taq</i> 10X	2.5 µl
H ₂ O	15µl
dNTPs 1 mM	2.5µl

Amorce pRS3' 5pmol/μl	2μl
Amorce pRS5' 5pmol/μl	2μl
ADN	1μl
<i>Dream Taq</i> 1U/μl (Fermontas)	0.12 μl

- pRS3' CTTTGAGTGAGCTGATACCG
- pRS5' GCTAAGGTAGAGGGTGAACG

Programme de PCR

94°C 2 min

20 cycles {

- 94°C, 30 sec
- 55°C, 30 sec
- 72°C, x min (1min par kb)

72°C, 5 min

17. Transformer *E. coli* avec le reste de l'extraction d'ADN de *S. cerevisiae*.

10. Transformation bactérienne

10.1 Protocole

1. Déposer 10 à 20 μl d'ADN dans un tube en verre placé dans la glace.
2. Laisser décongeler les cellules compétentes dans la glace.
3. Transférer 150 à 200 μl des cellules décongelées dans le tube contenant l'ADN plasmidique.
4. Mélanger doucement (ne pas vortexer).
5. Laisser le mélange ADN/cellules 30 min dans la glace.
6. Soumettre à un choc thermique en plaçant le tube dans un bain marie à 42°C pendant 40 sec.
7. Ajouter 1 ml de milieu LB stérile.
8. Mettre en culture à 37°C sous agitation pendant 1 h.
9. Centrifuger pendant 5 min à 3000 rpm.
10. Éliminer 800 μl du surnageant et resuspendre le culot bactérien dans le volume restant.
11. Étaler la suspension bactérienne avec des billes de verre sur 2 boîtes de LB + Ampicilline (100 μg/ml).
12. Incuber les boîtes 48 h à 37°C

Attention : un essai témoin sera réalisé sur un tube ne contenant pas d'ADN, en suivant la même procédure d'identification des colonies positives par PCR avec les amorces (pRS3' et pRS5')

11. Extraction de l'ADN bactérien

11.1 Solutions

Solution 1 :

- Glucose 50 mM
- Tris Hcl pH8 25 mM
- EDTA 10 mM

Solution 2 :

- NaOH 0.2 M
- SDS 1 %

Solution 3 :

- Potassium acétate 29.43 g
- Glacial acide acétique 11.5 ml
- H₂O 28.5 ml
- Conserver à 4°C

Solution PEG+ NaCl

- PEG 13 %
- NaCl 1.6 M

Chlorure de lithium LiCl 10 M

11.2 Protocole

1. Préparer une préculture des bactéries transformées sur un milieu LB + ampicilline (10 ml) à 37°C sous agitation toute la nuit.
2. Diluer la culture de la nuit dans le milieu LB+ Ampicilline (200 ml) et cultiver à 37°C sous agitation pendant 24 h.
3. Centrifuger les cultures dans des grandes bouteilles pendant 10 min à 4000 rpm.
4. Resuspendre le culot dans la solution 1 (7ml) et le transférer dans un tube Falcon de 50 ml.
5. Ajouter un peu de lysozyme, mélanger et laisser agir 10 min à température ambiante.
6. Ajouter 20 ml de la solution 2, mélanger délicatement par inversion et incubé 10 min à température ambiante. La solution devient très visqueuse.

7. Ajouter 15 ml de la solution 3 froide et mélanger immédiatement mais délicatement par inversion du tube. Observez la précipitation des débris bactériens et de l'ADN chromosomique.
8. Centrifuger 10 min à 4000 rpm.
9. Transférer le surnageant dans un erlenmeyer.
10. Ajouter 0.6 volume d'isopropanol, mélanger et laisser précipiter.
11. Transférer dans des tubes de verre et centrifuger 10 min à 4000 rpm.
12. Eliminer le surnageant et répéter l'étape 11.
13. Ajouter 3 ml d'éthanol à 70 %, vortexer et centrifuger 10 min à 4000 rpm.
14. Eliminer le surnageant et ajouter 3 ml d'eau.
15. Lorsque le culot est totalement dissous, ajouter 1 ml de la solution de chlorure de lithium à 10 M.
16. Laisser dans la glace 10 min.
17. Centrifuger 10 min à 4000 rpm.
18. Transférer le surnageant dans un nouveau tube.
19. Ajouter un volume égale d'isopropanol (4 ml).
20. Centrifuger 10 min à 4000 rpm.
21. Eliminer le surnageant et ajouter 3ml d'éthanol à 70 %.
22. Centrifuger 10 min à 4000 rpm.
23. Eliminer le surnageant puis sécher 1 h au lyophilisateur.
24. Resuspendre le culot dans 300µl d'H₂O et transférer l'ensemble dans un tube Eppendorf 1.5 ml.
25. Laver le premier tube avec 200 µl d'H₂O et ajouter le lavage aux 300 µl.
26. Ajouter 1 µl de RNase 10 mg/ml et incuber à 37°C pendant 1 h.
27. Ajouter 500 µl de NaCl + PEG et laisser dans la glace quelques minutes.
28. Centrifuger 5 min à 11000 rpm.
29. Eliminer le surnageant et ajouter 1ml d'ethanol à 70%.
30. Centrifuger 1min à 11000 rpm
31. Eliminer le surnageant puis sécher le culot sous vide
32. Resuspendre le culot dans 300 µl d'H₂O.
33. Analyser par PCR (amorces pRS3' et pRS5') pour vérifier la présence de la cassette de délétion.

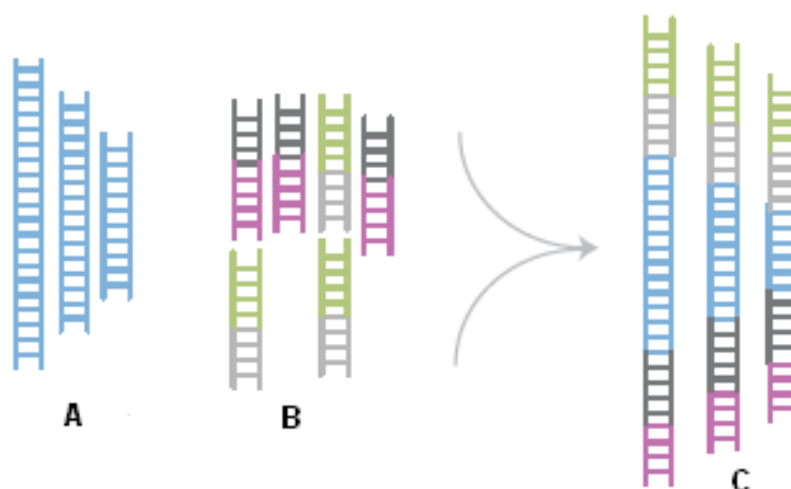
12. Linéarisation de la cassette de délétion

1. Préparer une dilution du plasmide pRS426 avec la cassette de délétion afin d'obtenir entre 5 et 10 µg d'ADN pour transformer *P. anserina*.

2. Dans un volume de 100 μl , ajouter 5 μl de l'enzyme *AscI* et 5 μl de tampon Tango (Fermentas).
3. Incuber 5 h à 37°C.
4. Passer sur gel pour vérifier la taille de la cassette de délétion.

Annexe 4: Protocole de préparation de la banque « paire end » selon Illumina.

1. Principe



"A" l'ADN génomique est fragmenté par nébulisation, "B" des séquences adaptatrices sont liguées aux extrémités des fragments puis le mélange réactionnel est mis à migrer par électrophorèse, "C" les fragments qui possèdent les bonnes séquences adaptatrices sont sélectionnés par PCR.

2. Réparation des extrémités 3' et 5' des fragments d'ADN

Ce protocole permet l'obtention des fragments d'ADN avec des bouts francs, l'activité exonucléase 3' à 5' de la T4 DNA polymérase et l'enzyme Klenow répare la région 3' tandis que l'activité polymérase répare la région 5'.

2.1 Matériels

- Tampon T4 DNA ligase 10 mM ATP (BioLabs)
- dNTP Mix 10 mM
- BSA
- H₂O
- Klenow enzyme 1 U/μl
- T4 PNK (PolyNucléotide Kinase) 1 U/μl
- T4 DNA polymérase 1 U/μl (BioLabs)
- QIAquick PCR Purification Kit (QIGEN)

2.2 Protocols

1. Préparer le Mix suivant:

- ADN fragmentés par nébulisation (1 à 2μg) 25 μl
- Tampon T4 DNA ligase 10 mM ATP 10 μl

- dNTP Mix 10 mM 4 µl
- BSA 1 µl
- H₂O 60 µl
- Klenow enzyme 1 U/µl 2 µl
- T4 PNK (PolyNucléotide Kinase) 1U/µl 2 µl
- T4 DNA polymérase 1 U/µl 2 µl

Le volume total doit être à 100 µl

2. Incuber dans le thermocycleur 30 min à 20 C°.
3. Purifier l'ADN à l'aide du kit PCR purification de QIAquick PCR Purification Kit.
 - Ajouter 5 volumes du tampon PB.
 - Transférer l'ensemble sur la colonne QIAquick.
 - Centrifuger 1 min à 13000 rpm.
 - Eliminer l'éluat et ajouter 750 µl de tampon PE.
 - Centrifuger 1min à 13000 rpm.
 - Eliminer l'éluat et centrifuger 5 min à 13000 rpm.
 - Déposer la colonne sur un tube propre 1.5 ml et éluer l'ADN avec 40 µl de tampon EB.

3. Ajouter la base « A » à la région 3' des fragments d'ADN

L'activité polymérase de l'enzyme Klenow (3' et 5' exo moins) permet d'ajouter la base A à la région 3' afin de fixer les adaptateurs sur l'ADN.

3.1 Matériels

- Tampon NE 10 x (BioLabs)
- BSA
- dATP 1 mM
- Klenow fragment 3' à 5' exo moins (BioLabs)
- Kit MinEluat PCR purification

3.2 Protocole

1. Préparer le mélange suivant :
 - Tampon NE 2 10x 5 µl
 - BSA 1 µl
 - D'ATP 1 mM 10 µl
 - L'échantillon d'ADN 35 µl
 - Klenow fragment 1U/µl 2 µl
2. Incuber le mélange dans le thermocycleur 30min à 37C°.
3. Purifier l'ADN à l'aide du Kit Minieluat PCR purification.

- Ajouter 5 volumes du tampon PB.
- Transférer l'ensemble sur la colonne QIAquick.
- Centrifuger 1min à 13000 rpm.
- Eliminer l'eluât et ajouter 750 µl de tampon PE.
- Centrifuger 1 min à 13000 rpm.
- Eliminer l'eluât et centrifuger 5 min à 13000 rpm.
- Déposer la colonne sur un tube propre 1.5 ml et éluer l'ADN avec 12 µl de tampon EB.

4. Ligation des adaptateurs :

Pendant cette étape nous allons fixer les adaptateurs aux extrémités des fragments d'ADN, ces adaptateurs sont nécessaires durant le séquençage ils permettent la fixation des fragments d'ADN sur les cellules de séquençage.

4.1 Matériels

- Quick DNA Ligation Buffer 2X (BioLabs)
- Tampon DNA ligase (BioLabs)
- DNA ligase (BioLabs)
- Adaptateur 1 : P-GATCGGAAGAGCACACGTCT
- Adaptateur 2 : ACACTCTTTCCCTACACGACGCTCTTCCGATCT

4.2 Protocole

1. Appariement des adaptateurs

- Quick DNA Ligation Buffer 2x 10 µl
- Adaptateur 1 100 pmol/µl 3 µl
- Adaptateur 2 100 pmol/µl 3 µl
- H₂O 4 µl
- Incuber 5 min à 94C° puis laisser refroidir lentement à température ambiante pendant 30 min.

2. Ligation des adaptateurs

- Tampon DNA ligase 10 µl
- Echantillon d'ADN 10 µl
- Adaptateur 1+2 5 µl
- ADN ligase 1 µl

3. Incuber dans un thermocycleur 15 min à 25C°.

5. Purification des produits de ligation

5.1 Matériels

- TBE 5X
- Bromide d'ethidium (BET)
- Bleu de charge
- Marqueur de taille
- Agarose
- Kit QIAquick gel extraction

5.2 Protocole

1. Préparer un gel d'agarose 2 % (pour éviter la contamination entre les souches il faut préparer un gel pour chaque échantillon d'ADN).
2. Déposer 5 µl du marqueur de taille.
3. Déposer 30 µl de l'échantillon d'ADN+ 5 µl de bleu de charge.
4. Lancer l'électrophorèse à 120 V dans du TBE +BET.
5. Visualiser les bandes d'ADN par les UV.
6. Découper la bande à 300 pb.
7. Extraire l'ADN par le Kit QIAquick gel extraction :
 - Ajouter 3 volumes du tampon QG.
 - Incuber 10 min à 50C° jusqu'à dissolution totale du gel.
 - Ajouter 1 volume de l'isopropanol.
 - Transférer le mélange sur la colonne.
 - Centrifuger 1 min à 13000 rpm.
 - Eliminer l'eluat et ajouter 750 µl du tampon PE
 - Attendre 2 min et centrifuger 1min à 13000 rpm
 - Eliminer l'eluat et centrifuger 5 min à 13000 rpm pour sécher la colonne
 - Transférer la colonne sur un tube 1.5 ml puis éluer l'ADN avec 30 µl du tampon EB
 - Centrifuger 1min à 13000 rpm
 - Conserver l'ADN à -20 C°

6. Amplification des fragments d'ADN-adaptateurs par PCR

Cette étape permet non seulement de sélectionner les fragments d'ADN qui contiennent les adaptateurs aux deux extrémités mais aussi d'insérer les index spécifiques à chaque souche. La présence des index sur l'ADN est indispensable lors du séquençage de plusieurs souches à la fois dans une même piste, ils permettent de trier entre les souches pendant l'analyse bioinformatique des banques

6.1 Matériels

- ADN polymérase *Phusion* (FINNZYMES)
- Tampon *Phusion* HF 5 x (FINNZYMES)
- Kit QIAquick PCR Purification
- Gel d'agarose 2 %
- Index 1 (ADN *fprI*) :
CAAGCAGAAGACGGCATAACGAGATCGTGATGTGACTGGAGTTC
- Index 2 (ADN *fmrI*) :
CAAGCAGAAGACGGCATAACGAGATACATCGGTGACTGGAGTTC
- Index 3 (ADN *WT mat*-) :
CAAGCAGAAGACGGCATAACGAGATGCCTAAGTGACTGGAGTTC
- Amorces InpE1 :
AATGATACGGCGACCACCGAGATCTACTCTTTCCCTACACGACGCTCTTC
CGATCT
- Amorces InpE2 : GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT

6.2 Protocole

1. Préparer le mix pour la PCR :

- H ₂ O	30.5 µl
- Tampon <i>Phusion</i> HF 2x	10 µl
- dNTPs	1.5 µl
- Index	1 µl
- Amorces InpE1 25 µM	1 µl
- Amorces InpE2 0.5 µM	1 µl
- L'échantillon d'ADN	5 µl
- ADN polymérase <i>Phusion</i> 1 U/µl	0.5 µl

2. Programme PCR :

	30 s à 98 C°	
18 x	$\left\{ \begin{array}{l} 10 \text{ s à } 98 \text{ C}^\circ \\ 30 \text{ s à } 65 \text{ C}^\circ \\ 30 \text{ s à } 72 \text{ C}^\circ \end{array} \right.$	
	5 min à 72 C°	

3. Purification du produit de PCR par le Kit QIAquick PCR Purification Vérifier sur gel d'agarose 2 % la présence de la bande à 300 pb

4. Clonage des fragments d'ADN

Les produits PCR sont clonés dans le vecteurs pcR4Blunt topo. Une suspension cellulaire d'*E.coli* est transformée (Annexe 3) puis étalée sur un milieu sélectif LB solide

contenant de l'ampicilline (50 $\mu\text{g}\cdot\text{ml}^{-1}$), en présence d'X-Gal. Les boîtes sont incubées une nuit à 37°C. 15 colonies blanches par banque sont repiquées sur milieu LB+ ampicilline et validées par PCR

- Préparer le mix de PCR :
- Tampon green *Gotaq* 5 x 5 μl
- MgCl_2 1.5 μl
- dNTP 1mM 2.5 μl
- Amorces UP 2 μl
- Amorces RP 2 μl
- Polymerase *Gotaq* 1U/ μl 0.1 μl
- H_2O 12 μl

Programme PCR:

2 min à 94°C
30x { 30 s à 94 C°
30 s à 55 C°
1 min à 72 C°
5 min à 72 C°

- Identifier les colonies positives sur gel d'agarose 1 %

Annexe 5 : Sites d'internet utilisé au cours de cette étude.

Nom	Adresse	Fonction
NCBI	http://blast.ncbi.nlm.nih.gov/	Programmes d'alignements Blast
Pfam	http://pfam.sanger.ac.uk/	Programmes de recherche des domaines protéiques
Clustal W	http://www.ebi.ac.uk/Tools/msa/clustalw2/	Programmes d'alignements multiples
Meme	http://meme.nbcr.net/meme4_6_1/cgi-bin/meme.cgi	Programmes de recherche des motifs sur les séquences nucléotidique
<i>Podospora anserina</i> Genome Project	http://podospora.igmors.u-psud.fr/	Site spécifique <i>P. anserina</i>
FUNGIpath	http://embg.igmors.u-psud.fr/fungipath/	Programmes de recherche des orthologues chez les champignons
<i>Neurospora crassa</i> Genome Project	http://www.broadinstitute.org/annotation/genome/neurospora/News.html	Site spécifique <i>N. crassa</i>
SGD	http://www.yeastgenome.org/	Site spécifique <i>Saccharomyces cerevisiae</i>
GDB	http://old.genedb.org/genedb/pombe/	Site spécifique <i>Schizosaccharomyces pombe</i>
<i>Fusarium graminearum</i> Database	http://www.broadinstitute.org/annotation/genome/fusarium_graminearum/MultiHome.html	Site spécifique <i>Fusarium graminearum</i>
<i>Chaetomium globosum</i> Database	http://www.broadinstitute.org/annotation/genome/chaetomium_globosum/Home.html	Site spécifique <i>Chaetomium globosum</i>
<i>Magnaporthe grisea</i> Database	http://www.broadinstitute.org/annotation/fungi/magnaporthe/	Site spécifique <i>Chaetomium globosum</i>

Annexe 6 : Liste des gènes différentiellement exprimés dans la comparaison *mat+/*fprI chez *P. anserina* et adjacentes aux orthologues des gènes différentiellement exprimés dans la comparaison *MAT1-2-1/ΔMAT1-2-1***

best hit accession number <i>G. monoliformis</i>	accession number for <i>P. anserina</i> orthologs	<i>P. anserina</i> gene differentially expressed in <i>mat+</i> vs <i>fprI</i> and adjacent to an ortholog differentially expressed in <i>MAT1-2-1</i> vs <i>ΔMAT1-2-1</i> strains of <i>G. monoliformis</i>
AAX07692	Pa_3_1320	Pa_3_1300
EAA73884	Pa_3_8360	Pa_3_8390
EAL85418	Pa_5_10260	Pa_5_10247
EAA73871	Pa_5_9270	Pa_5_9260 ; Pa_5_9280
EAL90830	Pa_7_2730	Pa_7_2740
EAA76607	Pa_1_17820	Pa_1_17860
CAE76555	Pa_1_11510	Pa_1_11530
EAL95577	Pa_1_17420	Pa_1_17430
EAL89964	Pa_5_10420	Pa_5_10410
EAL88607	Pa_4_6440	Pa_4_6450
EAA74619	Pa_4_5820	Pa_4_5840
EAA74990	Pa_3_5360	Pa_3_5330
EAL90845	Pa_6_5140	Pa_6_5160
CAE76499	Pa_1_13790	Pa_1_13780
XP_711465	Pa_2_13660	Pa_2_13600
EAL88773	Pa_4_6850	Pa_4_6840; Pa_4_6870
EAL86321	Pa_5_8670	Pa_5_8610 ; Pa_5_8620 ; Pa_5_8640 ; Pa_5_8660
EAL93262	Pa_3_1770	Pa_3_1760
EAL91268	Pa_7_7800	Pa_7_7840
EAA67345	Pa_2_1050	Pa_2_1040
CAF05897	Pa_5_9150	Pa_5_9120
ZP_00401899	Pa_1_14320	Pa_1_14390
BAA07817	Pa_3_7480	Pa_3_7460
CAC27836	Pa_1_18080	Pa_1_18090
CAC27836	Pa_2_10220	Pa_2_10240
CAC45043	Pa_2_6020	Pa_2_6010 ; Pa_2_6030
EAL91008	Pa_7_2770	Pa_7_2740 ; Pa_7_2790
P21951	Pa_5_3400	Pa_5_3395
XP_751919	Pa_2_11580	Pa_2_11595
EAA68557	Pa_2_4375	Pa_2_4370
CAA62528	Pa_3_9370	Pa_3_9380 ; Pa_3_9350
AAX07705	Pa_6_10260	Pa_6_10280 ; Pa_6_10220
XP_887169	Pa_1_18830	Pa_1_18840 ; Pa_1_18860 ; Pa_1_18880
CAJ04126	Pa_4_9630	Pa_4_9620 ; Pa_4_9660 ; Pa_4_9690 ; Pa_4_9700
NP_595638	Pa_5_4600	Pa_5_4590 ; Pa_5_4630
EAA78050	Pa_2_11690	Pa_2_11720
NP_497198	Pa_1_16180	Pa_1_16170
AAW69356	Pa_5_870	Pa_5_860
CAD71100	Pa_2_13870	Pa_2_13920

CAA04646	Pa_4_8890	Pa_4_8880 ; Pa_4_8920
EAL92474	Pa_5_5720	Pa_5_5700
EAL92998	Pa_5_7260	Pa_5_7250 ; Pa_5_7290
EAL89069	Pa_5_5500	Pa_5_5490 ; Pa_5_5510
CAE85549	Pa_4_8680	Pa_4_8675 ; Pa_4_8720
NP_011681	Pa_1_21250	Pa_1_21210 ; Pa_1_21280
EAA75069	Pa_2_11630	Pa_2_11595
AAD10252	Pa_2_2150	Pa_2_2160 ; Pa_2_2180 ; Pa_2_2190
AAW44097	Pa_3_360	Pa_3_350
EAL91616	Pa_7_10850	Pa_7_10880
AAH24401	Pa_0_260	Pa_0_290
CAC09327	Pa_1_16140	Pa_1_16130 ; Pa_1_16170 ; Pa_1_16180
XP_324798	Pa_1_14790	Pa_1_14800 ; Pa_1_14820
EAA70869	Pa_1_19310	Pa_1_19290 ; Pa_1_19330
EAL89710	Pa_6_9750	Pa_6_9760
CAD21366	Pa_2_1930	Pa_2_1890 ; Pa_2_1960
AAW69346	Pa_2_13190	Pa_2_13160
CAC28724	Pa_6_10910	Pa_6_10880
XP_646132	Pa_5_4250	Pa_5_4220
EAA73460	Pa_6_10610	Pa_6_10620 ; Pa_6_10590
AAL57485	Pa_5_6340	Pa_5_6330
BAB98356	Pa_1_12430	Pa_1_12450
CAA76330	Pa_1_22930	Pa_1_22920
XP_747610	Pa_2_6540	Pa_2_6590
XP_718105	Pa_3_6370	Pa_3_6360
CAH03680	Pa_2_1240	Pa_2_1200 ; Pa_2_1280 ; Pa_2_1290
ABA49992	Pa_4_7640	Pa_4_7620 ; Pa_4_7600
AAL66352	Pa_6_11240	Pa_6_11260 ; Pa_6_11270
Q00332	Pa_7_10170	Pa_7_10120
EAK96719	Pa_5_5450	Pa_5_5470
EAL85758	Pa_1_19290	Pa_1_19330
EAL84525	Pa_1_22300	Pa_1_22320 ; Pa_1_22330
Q09694	Pa_1_10980	Pa_1_10990
ABA21858	Pa_1_22740	Pa_1_22700
BAA13047	Pa_3_4260	Pa_3_4240
CAE76548	Pa_1_14470	Pa_1_14420
CAI47587	Pa_6_2860	Pa_6_2880
CAE85550	Pa_4_8700	Pa_4_8720
CAE76477	Pa_1_16850	Pa_1_16860
EAL86037	Pa_6_3260	Pa_6_3270 ; Pa_6_3280 ; Pa_6_3290
ZP_00660025	Pa_1_8280	Pa_1_8260
ZP_00660025	Pa_7_860	Pa_7_870 ; Pa_7_880
CAD71035	Pa_2_9800	Pa_2_9790 ; Pa_2_9810
AAW40619	Pa_2_9440	Pa_2_9430 ; Pa_2_9420 ; Pa_2_9410
XP_418353	Pa_3_1720	Pa_3_1710
CAE76242	Pa_1_6390	Pa_1_6350
CAC28812	Pa_2_11870	Pa_2_11900

CAD21326	Pa_2_8400	Pa_2_8430 ; Pa_2_8440
AAB61278	Pa_1_17000	Pa_1_17010 ; Pa_1_17020
EAA68099	Pa_2_2360	Pa_2_2350 ; Pa_2_2340
EAA75931	Pa_7_5250	Pa_7_5260 ; Pa_7_5270
CAD21357	Pa_2_10640	Pa_2_10610
CAD70798	Pa_7_1580	Pa_7_1600
XP_731508	Pa_7_1230	Pa_7_1270 ; Pa_7_1280
EAL90922	Pa_1_15330	Pa_1_15320
EAA68333	Pa_2_980	Pa_2_960
XP_895972	Pa_2_3200	Pa_2_3210
XP_753012	Pa_3_9690	Pa_3_9700 ; Pa_3_9660
EAL90415	Pa_4_8560	Pa_4_8540 ; Pa_4_8590
EAL90415	Pa_5_7760	Pa_5_7750
EAL92505	Pa_5_7800	Pa_5_7830
CAE81941	Pa_1_5740	Pa_1_5710
EAA68216	Pa_5_2510	Pa_5_2520 ; Pa_5_2530 ; Pa_5_2540
BAB72843	Pa_3_8500	Pa_3_8470
XP_322154	Pa_7_9240	Pa_7_9200 ; Pa_7_9270
CAE76096	Pa_2_6460	Pa_2_6484
AAC39498	Pa_4_7730	Pa_4_7740 ; Pa_4_7750 ; Pa_4_7760
CAD21256	Pa_0_1030	Pa_0_1000
EAA72018	Pa_1_17480	Pa_1_17430
EAL93176	Pa_7_3760	Pa_7_3770
EAL93176	Pa_1_5400	Pa_1_5410
EAL93176	Pa_1_7170	Pa_1_7150
EAL93176	Pa_1_23740	Pa_1_23750
EAL93176	Pa_3_5690	Pa_3_5670 ; Pa_3_5710
EAL93176	Pa_7_5110	Pa_7_5140
EAL93176	Pa_6_6450	Pa_6_6440
EAL93176	Pa_6_10440	Pa_6_10430
EAL93176	Pa_5_1430	Pa_5_1420
XP_327310	Pa_1_23930	Pa_1_23910 ; Pa_1_23900
EAA70773	Pa_3_3190	Pa_3_3210

Annexe 7 : Liste des gènes différentiellement exprimés dans la comparaison du croisement ♂ *fpr1*⁻ X ♀ *WT mat*⁻ par rapport au croisement ♂ *fmr1 smr2*⁻ X ♀ *WT mat*⁺.

Δ Gènes délétés. ^a Gènes délétés dans des études précédente (Coppin et al. 2005). ^b Gènes délétés dans l'étude de la fécondation (Bidard et al. 2011). FC fold change. COG Clusters of Orthologous Groups.

ID	Fonction	FC Liste 1	COG	délétion
Pa_2_2310	<i>mat</i> ⁺ specific pheromone precursor encoded by the <i>mfp</i> gene	-83,9	unclassified	Δ ^a
Pa_1_24410	Putative SAM-dependent methyltransferases	-11,04	Cell envelope biogenesis, outer membrane / Coenzyme metabolism	Δ ^b
Pa_6_9560	Putative Pisatin demethylase	-4,69	Secondary metabolites biosynthesis, transport, and catabolism	
Pa_4_1380	Putative Pheromone receptor	-3,58	unclassified	Δ ^b
Pa_1_19170	Unknown function	-3,48	Coenzyme metabolism	Δ
Pa_0_1270	Unknown function	-3,3	Carbohydrate transport and metabolism	Δ
Pa_7_4100	Unknown function	-3,29	Translation, ribosomal structure and biogenesis	Δ ^b
Pa_1_18670	Unknown function	-3,23	unclassified	Δ
Pa_3_8000	Unknown function	-3,18	unclassified	Δ
Pa_5_3110	Unknown function	-3,03	unclassified	Δ
Pa_5_3115	Unknown function	-2,85	unclassified	Δ
Pa_1_17942	Unknown function	-2,69	unclassified	Δ
Pa_1_17947	Unknown function	-2,67	unclassified	Δ
Pa_5_12020	Unknown function	-2,66	unclassified	Δ
Pa_1_17960	Putative insecticidal toxin complex protein	-2,56	Cell envelope biogenesis, outer membrane	Δ
Pa_4_7760	Putative farnesyltransferase subunit beta	-2,35	Posttranslational modification, protein turnover, chaperones	Δ ^b
Pa_5_6340	Putative mitochondrial medium-chain specific acyl-CoA dehydrogenase precursor	-2,32	Lipid metabolism	
Pa_1_13080	Putative transporter	-2,23	Carbohydrate transport and metabolism	Δ
Pa_3_10200	Unknown function	-2,23	unclassified	Δ
Pa_3_3350	Unknown function	-2,23	unclassified	Δ
Pa_6_9980	Unknown function	-2,08	unclassified	Δ
Pa_2_4890	Unknown function	-2,03	unclassified	Δ
Pa_7_1380	Putative oxidoreductase	-2,02	Energy production and conversion	
Pa_2_7630	Unknown function	-2	unclassified	Δ
Pa_5_1490	Putative inositol monophosphatase	2,04	Inorganic ion transport and metabolism / Carbohydrate transport and metabolism	
Pa_4_977	Unknown function	2,06	unclassified	Δ
Pa_2_10300	Unknown function	2,07	unclassified	Δ
Pa_1_6014	Unknown function	2,1	unclassified	Δ
Pa_1_910	Putative monooxygenase	2,1	Inorganic ion transport and metabolism / Secondary metabolites biosynthesis, transport, and catabolism	
Pa_6_7350	Putative Aspartic protease	2,14	unclassified	Δ ^b
Pa_1_19540	Putative transporter	2,17	Carbohydrate transport and metabolism	Δ
Pa_mito_nad6	Putative NADH-ubiquinone oxidoreductase chain 6	2,21	unclassified	
Pa_4_6000	Unknown function	2,26	unclassified	Δ
Pa_2_11900	Glucose-repressible protein encoded by the <i>PaGrg1</i> gene	2,29	unclassified	
Pa_2_7580	Putative protein FAM62B	2,31	unclassified	Δ
Pa_6_6990	Putative Glycoside Hydrolase Family 18	2,41	unclassified	Δ
Pa_1_5940	Unknown function	2,43	unclassified	Δ
Pa_5_10560	Unknown function	2,45	unclassified	Δ
Pa_2_130	Unknown function	2,47	unclassified	Δ
Pa_5_6960	Unknown function	2,72	unclassified	Δ
Pa_5_6970	Unknown function	2,81	unclassified	Δ
Pa_mito_nad1	NADH-ubiquinone oxidoreductase chain 1	3,07	unclassified	
Pa_1_8290	<i>mat</i> ⁻ specific pheromone precursor encoded by the <i>mfm</i> gene	20,79	unclassified	Δ ^a

Annexe 8 : Liste des gènes communs entre les comparaisons ♂ *WT mat+* X ♀ *WT mat-/♂ fpr1* X ♀ *WT mat-* (liste2) et ♂ *WT mat-* X ♀ *WT mat+/♂ fmr1-smr2-* X ♀ *WT mat+* (liste3).

Δ Gènes délétés. FC fold change. COG Clusters of Orthologous Groups.

ID	Fonction	FC Liste 2	FC Liste 3	COG	Délétion
Pa_7_3020	Unknown function	9,25	4,1	unclassified	
Pa_1_290	Unknown function	7,53	16,49	unclassified	Δ
Pa_1_540	Unknown function	7,36	9,57	unclassified	Δ
Pa_1_21670	Unknown function	6,42	12,69	unclassified	Δ
Pa_6_11050	Unknown function	6,1	5,93	unclassified	
Pa_6_11710	Unknown function	5,98	9,17	unclassified	Δ
Pa_7_4430	Unknown function	5,85	4,88	unclassified	
Pa_5_3610	Unknown function	5,69	6,29	unclassified	
Pa_1_18840	Unknown function	5,68	10,37	Signal transduction mechanisms	Δ
Pa_5_4580	Putative pyruvate decarboxylase isozyme 3	5,67	6,92	Carbohydrate transport and metabolism / Coenzyme metabolism / acetolactate synthase, pyruvate dehydrogenase (cytochrome), glyoxylate carboligase, phosphonopyruvate decarboxylase /	
Pa_5_940	Unknown function	5,59	6,98	unclassified	Δ
Pa_4_2080	Unknown function	5,44	9,81	unclassified	Δ
Pa_1_3290	Unknown function	5,14	6,14	unclassified	Δ
Pa_6_6330	Putative NADPH dehydrogenase	5,04	4,93	Energy production and conversion	
Pa_5_90	Unknown function	5,02	5,78	unclassified	Δ
Pa_7_10740	Unknown function	4,93	7,03	unclassified	
Pa_1_18880	Unknown function	4,91	7,73	unclassified	Δ
Pa_1_17250	Unknown function	4,85	4,69	unclassified	Δ
Pa_6_1130	Putative Annexin	4,58	5,82	Intracellular trafficking and secretion	Δ
Pa_4_5980	Unknown function	4,57	4,48	unclassified	
Pa_4_8300	Putative mitochondrial chaperone BCS1	4,32	8,27	Posttranslational modification, protein turnover, chaperones	
Pa_5_315	Unknown function	3,89	4,62	unclassified	
Pa_2_10530	Putative Acyl-CoA dehydrogenase family	3,68	3,74	Lipid metabolism	
Pa_4_9350	peroxisomal 3-keto-acyl-CoA thiolase encoded by pthik	3,63	3,61	Lipid metabolism	
Pa_5_11750	Putative NADP-dependent oxidoreductase	3,61	4,29	Energy production and conversion	
Pa_1_1300	Unknown function	3,57	6,09	unclassified	
Pa_1_10820	Putative bifunctional P-450:NADPH-P450 reductase	3,52	4,22	Inorganic ion transport and metabolism / Secondary metabolites biosynthesis, transport, and catabolism	
Pa_1_6190	Putative Glycerol-3-phosphate dehydrogenase [NAD+]	3,49	4,37	GpsA, Glycerol-3-phosphate dehydrogenase	
Pa_1_24430	Unknown function	3,48	5,19	unclassified	
Pa_1_9760	Putative NADPH dehydrogenase	3,45	2,07	Energy production and conversion	
Pa_6_5270	Putative 26S proteasome subunit RPN4	3,45	5,27	unclassified	
Pa_5_4585	Unknown function	3,39	3,33	unclassified	
Pa_2_11930	Unknown function	3,38	4,79	unclassified	
Pa_5_11490	Putative guanine nucleotide-binding protein alpha-2 subunit	3,37	3,59	unclassified	
Pa_1_18090	Unknown function	3,34	2,75	unclassified	
Pa_2_2160	Putative lipase	3,32	4,64	unclassified	
Pa_6_840	Unknown function	3,32	3,83	unclassified	
Pa_6_90	Unknown function	3,27	2,87	unclassified	
Pa_2_6276	Unknown function	3,24	5,66	unclassified	
Pa_1_21590	Unknown function	3,09	3,76	unclassified	
Pa_4_9800	Putative transporter	3,04	4,65	Carbohydrate transport and metabolism	
Pa_6_9540	Unknown function	3,03	4,24	unclassified	
Pa_3_9660	Putative Carbohydrate Esterase Family 1	3	4,29	Secondary metabolites biosynthesis, transport, and catabolism / Amino acid transport and metabolism	
Pa_7_3090	Unknown function	2,98	2,88	unclassified	

Pa_2_7200	Putative aflatoxin biosynthesis ketoreductase	2,87	2,32	Secondary metabolites biosynthesis, transport, and catabolism
Pa_4_3960	Unknown function	2,84	2,72	unclassified
Pa_1_15530	Unknown function	2,81	3,91	unclassified
Pa_3_7340	Putative Polyamine transporter	2,8	3,51	Carbohydrate transport and metabolism
Pa_0_1410	Unknown function	2,77	3,32	unclassified
Pa_6_4810	Unknown function	2,75	3,56	unclassified
Pa_2_1290	Unknown function	2,74	3,43	unclassified
Pa_5_3130	Unknown function	2,72	4,31	Lipid metabolism
Pa_6_9890	Putative Inositol polyphosphate kinase	2,71	4,01	unclassified
Pa_3_70	Unknown function	2,69	2,17	unclassified
Pa_6_1340	Putative Mitochondrial precursor of Carnitine O-acetyltransferase	2,68	3,16	unclassified
Pa_7_2860	Putative 3',5'-cyclic-nucleotide phosphodiesterase	2,66	2,93	Signal transduction mechanisms
Pa_1_2580	Unknown function	2,65	3,29	unclassified
Pa_1_18410	Unknown function	2,61	3,99	unclassified
Pa_4_5450	Unknown function	2,6	3,29	unclassified
Pa_6_5310	Putative Amino acid permease	2,59	3,3	Amino acid transport and metabolism
Pa_1_6930	Unknown function	2,57	3,46	unclassified
Pa_4_7450	Putative 3-methyl-2-oxobutanoate hydroxymethyltransferase	2,57	3,63	Coenzyme metabolism / Carbohydrate transport and metabolism
Pa_3_9350	Unknown function	2,55	3,52	unclassified
Pa_1_15590	Unknown function	2,54	2,33	unclassified
Pa_5_3290	Putative Glycoside Hydrolase Family 18	2,54	2,55	Carbohydrate transport and metabolism
Pa_5_1500	Unknown function	2,5	3,32	Cell envelope biogenesis, outer membrane
Pa_5_7760	Putative cation-transporting ATPase	2,49	2,93	Inorganic ion transport and metabolism
Pa_3_8550	Unknown function	2,48	3,34	unclassified
Pa_6_5870	Unknown function	2,48	3,08	unclassified
Pa_1_22530	Putative 2-nitropropane dioxygenase	2,47	3	Lipid metabolism
Pa_3_6950	Putative long-chain-fatty-acid--CoA ligase	2,47	2,62	Secondary metabolites biosynthesis, transport, and catabolism / Lipid metabolism
Pa_2_6278	Unknown function	2,44	3,17	unclassified
Pa_4_3400	Putative Nucleoside deaminase	2,4	3,72	Nucleotide transport and metabolism / Translation, ribosomal structure and biogenesis
Pa_1_1460	Putative 3-hydroxyacyl-CoA dehydrogenase	2,38	2,26	Secondary metabolites biosynthesis, transport, and catabolism
Pa_3_720	Unknown function	2,37	2,33	unclassified
Pa_3_9970	Unknown function	2,37	2,84	unclassified
Pa_5_7830	Unknown function	2,37	3,49	unclassified
Pa_1_10340	Unknown function	2,36	2,61	unclassified
Pa_1_5380	Unknown function	2,33	2,86	unclassified
Pa_7_5250	Unknown function	2,32	3,4	unclassified
Pa_4_830	Unknown function	2,29	3,26	unclassified
Pa_4_1960	Unknown function	2,27	2,57	unclassified
Pa_2_4820	Unknown function	2,26	3,57	Cell envelope biogenesis, outer membrane
Pa_5_8390	Unknown function	2,26	3,48	unclassified
Pa_4_1900	Putative Heterokaryon incompatibility protein	2,25	2,07	unclassified
Pa_1_1950	Unknown function	2,24	2,18	unclassified
Pa_2_1550	Unknown function	2,24	2,42	unclassified
Pa_5_12915	Unknown function	2,23	2,48	unclassified
Pa_1_8270	Putative fumarate reductase	2,22	2,54	Energy production and conversion / Coenzyme metabolism
Pa_2_6830	Unknown function	2,22	2,81	unclassified
Pa_1_16550	Unknown function	2,21	2,31	unclassified
Pa_7_10380	Unknown function	2,19	3,35	unclassified
Pa_2_8460	Putative glutathione peroxidase	2,17	2,36	Posttranslational modification, protein turnover, chaperones
Pa_2_8520	Unknown function	2,14	2,74	unclassified
Pa_7_5780	Putative peroxisomal adenine nucleotide transporter ANTI	2,14	2,38	unclassified

Pa_2_8820	Unknown function	2,13	2,59	unclassified	
Pa_4_5510	Unknown function	2,13	2,74	unclassified	
Pa_1_1690	Unknown function	2,11	2,04	unclassified	
Pa_2_4930	Unknown function	2,09	2,31	unclassified	
Pa_1_21400	Unknown function	2,08	3,04	unclassified	
Pa_7_3430	Unknown function	2,08	2,37	unclassified	
Pa_7_190	Putative pyoverdine/dityrosine biosynthesis protein	2,07	2,38	Secondary metabolites biosynthesis, transport, and catabolism	
Pa_1_16130	Unknown function	2,06	3,29	Transcription / RNA processing and modification / Cell division and chromosome partitioning / Chromatin structure and dynamics	
Pa_1_22930	Putative DNA-binding protein cre-1	2,05	3,39	unclassified	
Pa_4_360	Unknown function	2,04	2,27	unclassified	
Pa_3_1760	Putative MFS transporter	2,03	2,27	Carbohydrate transport and metabolism	
Pa_3_2050	Putative AAA superfamily ATPases	2,03	2,28	Posttranslational modification, protein turnover, chaperones	
Pa_4_890	Putative signal transduction histidine-protein kinase	2,03	3,21	Signal transduction mechanisms	
Pa_1_18990	Putative 1-phosphatidylinositol-4,5-bisphosphate phosphodiesterase	2,02	2,6	unclassified	
Pa_1_8110	Unknown function	2,01	2,28	unclassified	
Pa_5_3160	Unknown function	2,01	2,02	Cell envelope biogenesis, outer membrane / Carbohydrate transport and metabolism	
Pa_2_1680	Unknown function	2	2,03	unclassified	
Pa_5_640	Unknown function	2	2,84	unclassified	
Pa_3_2200	Unknown function	-2,01	-2,16	Signal transduction mechanisms	
Pa_6_8770	Putative Aprataxin	-2,01	-2,43	unclassified	
Pa_1_1840	Unknown function	-2,04	-2,12	unclassified	
Pa_5_7720	Unknown function	-2,05	-2,61	unclassified	
Pa_3_925	Unknown function	-2,07	-2,57	unclassified	
Pa_1_22370	Unknown function	-2,13	-3,18	unclassified	
Pa_4_1280	Unknown function	-2,14	-2,2	Coenzyme metabolism	
Pa_4_4550	Putative Aldehyde dehydrogenase	-2,16	-2,49	Energy production and conversion	
Pa_4_1940	Putative dehydrogenase	-2,22	-2,05	Energy production and conversion	
Pa_4_3820	Putative dehydrogenase	-2,32	-4,25	Energy production and conversion	
Pa_1_4050	Putative ATP-dependent permease	-2,36	-2,26	Defense mechanisms	
Pa_4_3890	Putative ATP-binding (ABC) transporters	-2,41	-2,44	Defense mechanisms / Inorganic ion transport and metabolism / Coenzyme metabolism /	
Pa_5_1210	Unknown function	-2,43	-2,18	unclassified	
Pa_1_495	Unknown function	-2,45	-2	unclassified	
Pa_4_6330	Putative protein similar to protein TRI15 of Fusarium sporotrichioides	-2,49	-2,75	unclassified	
Pa_3_210	Putative Isopenicillin N epimerase	-2,5	-3,22	Amino acid transport and metabolism	
Pa_5_3660	Unknown function	-2,51	-4,24	Secondary metabolites biosynthesis, transport, and catabolism	
Pa_7_11030	Unknown function	-2,51	-3,43	unclassified	
Pa_4_3780	Putative oxidoreductase	-2,54	-3,79	Secondary metabolites biosynthesis, transport, and catabolism	
Pa_2_6485	Unknown function	-2,56	-2,43	unclassified	
Pa_1_21160	Unknown function	-2,6	-2,38	unclassified	
Pa_6_850	Unknown function	-2,64	-2,52	unclassified	
Pa_1_205	Unknown function	-2,66	-2,28	unclassified	
Pa_1_8950	Unknown function	-2,78	-4,5	Inorganic ion transport and metabolism	
Pa_4_5090	Unknown function	-2,78	-3,49	unclassified	
Pa_1_16705	Unknown function	-2,91	-2,35	Cell envelope biogenesis, outer membrane	
Pa_5_12470	Putative protein similar to hemerythrin-like protein C869,06c of S.pombe	-2,94	-3,91	unclassified	
Pa_7_4770	Unknown function	-3,03	-2,08	unclassified	
Pa_1_22860	Unknown function	-3,09	-2,97	Inorganic ion transport and metabolism	
Pa_1_55	Unknown function	-3,11	-2,84	unclassified	
Pa_1_4960	Unknown function	-3,15	-4,02	unclassified	
Pa_6_9090	Unknown function	-3,15	-3,67	unclassified	

Pa_7_1060	CATP2 small subunit catalase encoded by the <i>catP2</i> gene	-3,2	-3,46	Inorganic ion transport and metabolism	
Pa_2_7860	Putative vacuolar basic amino acid transporter	-3,37	-3,32	Carbohydrate transport and metabolism	
Pa_4_1610	Putative polyketide synthase / Acyl-coenzyme A synthetases/AMP-(fatty) acid ligases	-3,55	-5,65	Secondary metabolites biosynthesis, transport, and catabolism / Lipid metabolism	
Pa_6_7370	Unknown function	-3,89	-5,22	unclassified	
Pa_1_15495	Unknown function	-4,73	-5,91	unclassified	
Pa_3_10530	Unknown function	-5,89	-7,33	unclassified	
Pa_0_1560	Putative protein similar to YGL110C of <i>Saccharomyces cerevisiae</i>	-6,4	-7,57	unclassified	
Pa_7_4190	Unknown function	-11,31	-9,43	unclassified	
Pa_2_4790	Putative transesterase	-23,64	-31,85	Defense mechanisms / Cell envelope biogenesis, outer membrane	

Annexe 9: Les couleurs des résidus présent dans les alignements des acides aminés

^a si plus d'un résidu est spécifié, les règles sont appliquées à chacun de ces résidus.

^b si les résidus sont connecté ensemble telle que WLVIAMFCHP alors n'importe quelle combinaison de ces résidus doit attendre ou dépasser le pourcentage de la couleur donnée.

Résidu ^a	couleur	Règle de couleur (minimum %, groupe de résidu) ^b
A, I, L, M, F, W, V	BLEU	(+ 60%, WLVIAMFCHP)
R, K	ROUGE	(+ 60%, KR), (+ 80%, K, R, Q)
N	VERT	(+ 50%, N), (+ 85%, N, Y)
C	BLEU	(+ 60%, WLVIAMFCHP)
C	ROSE	(100%, C)
Q	VERT	(+ 60%, KR), (+ 50%, QE), (+ 85%, Q, E, K, R)
E	MAGENTA	(+ 60%, KR), (+ 50%, QE), (+ 85%, E, Q, D)
D	MAGENTA	(+ 60%, KR), (+ 85%, K, R, Q), (+ 50%, ED)
G	ORANGE	(+ 0%, G)
H, Y	CYAN	(+ 60%, WLVIAMFCHP), (+ 85% W, Y, A, C, P, Q, F, H, I, L, M, V)
P	JAUNE	(+ 0%, P)
S, T	VERT	(+ 60%, WLVIAMFCHP), (+ 50%, TS), (+ 85%, S, T)

Annexe 10: Résultats de la quantification de la transcription des gènes des les mutant HMG par rapport au sauvage

N/A: la quantification relative n'est pas disponible.

*: p-value > 0.05

REF: gène de référence

Type de gène	NOM	Souche ΔPa_1_13940		Souche ΔPa_1_14230		Souche ΔPa_6_4110		Souche ΔPa_7_7190	
		<i>mat+</i>	<i>mat-</i>	<i>mat+</i>	<i>mat-</i>	<i>mat+</i>	<i>mat-</i>	<i>mat+</i>	<i>mat-</i>
Gènes HMG-box	<i>Pa_1_13940</i>	N/A	N/A	0.095	0.07	1.3	1.3	3.6	4
	<i>Pa_1_14230</i>	0.9*	1.0*	N/A	N/A	1.0*	1.2	1.0*	1.1*
	<i>Pa_6_4110</i>	0.30	0.36	0.12	0.19	N/A	N/A	1.8	2.1
	<i>Pa_7_7190</i>	1.1*	0.76*	1.1*	1.2*	0.9*	1.4	N/A	N/A
Gène <i>mat-</i>	<i>FMR1</i>	N/A	0.1	N/A	0.12	N/A	0.45	N/A	2.6
Gènes cibles <i>mat-</i>	<i>MFM</i>	N/A	0.004	N/A	0.06	N/A	0.60	N/A	2.6
	<i>PRE1</i>	N/A	0.1	N/A	0.12	N/A	0.1	N/A	0.6
	<i>Pa_6_7350</i>	N/A	0.005	N/A	0.14	N/A	1.5	N/A	4.3
Gène <i>mat+</i>	<i>FPR1</i>	0.25	N/A	0.73	N/A	0.45	N/A	3.3	N/A
Gènes cibles <i>mat+</i>	<i>MFP</i>	0.0004	N/A	0.07	N/A	0.25	N/A	1.1*	N/A
	<i>PRE2</i>	0.25	N/A	0.14	N/A	0.24	N/A	0.58	N/A
	<i>Pa_4_3858</i>	0.030	N/A	0.38	N/A	0.35	N/A	1.2*	N/A
	<i>Pa_1_24410</i>	0.0002	N/A	0.12	N/A	0.72	N/A	3	N/A
	<i>Pa_5_9770</i>	0.3	N/A	0.07	N/A	2	N/A	8.7	N/A
	<i>AOX</i>	0.5	N/A	3.9	N/A	0.79*	N/A	0.88*	N/A
	<i>PEPCK</i>	1.1*	N/A	0.46	N/A	1.3*	N/A	2.7	N/A
Gènes de ménage	<i>Pa_4_80</i>	1.2*	N/A	16	N/A	1.0*	N/A	1.4	N/A
	<i>ASI</i>	REF	REF	REF	REF	REF	REF	REF	REF
	<i>CIT1</i>	N/A	N/A	1.2	0.9	1.0*	0.83	0.87*	0.78
	<i>GPD</i>	1.0*	1.0*	1.8	1.5	REF	REF	1.3	1.3
	<i>H2A</i>	REF	REF	1.6	2	1.1*	1.1	1.0*	1.2
	<i>LEU1</i>	1.2*	1.2	0.76	0.9	0.87	1.2	1.2	1.5
	<i>PAH1</i>	N/A	N/A	0.6	0.4	1.1*	1.2*	0.9*	0.79*
	<i>PDF</i>	0.6	0.7	1.5	1.5	0.95*	1*	1.2	1.3
	<i>TBP</i>	REF	REF	0.78	0.6	1.0*	1.0*	0.85*	0.75
	<i>TIP</i>	REF	REF	1.1*	1.1*	REF	REF	1.1*	1.1
<i>UBC</i>	0.8	1.0*	2	2	0.8	0.9	1.3	1.5	

Article

Genome-Wide Gene Expression Profiling of Fertilization Competent Mycelium in Opposite Mating Types in the Heterothallic Fungus *Podospora anserina*

Frédérique Bidard^{1,2,3a,9}, Jinane Aït Benkhali^{1,2,9}, Evelyne Coppin^{1,2}, Sandrine Imbeaud^{3,4b}, Pierre Grognet^{1,4}, Hervé Delacroix^{3,5}, Robert Debuchy^{1,2*}

1 Univ Paris-Sud, Institut de Génétique et Microbiologie UMR8621, Orsay, France, **2** CNRS, Institut de Génétique et Microbiologie UMR8621, Orsay, France, **3** CNRS, Centre de Génétique Moléculaire FRE3144, GODMAP, Gif sur Yvette, France, **4** UFR des Sciences du Vivant, Université Paris 7-Denis Diderot, Paris, France, **5** Univ Paris-Sud, Orsay, France

Abstract

Background: Mating-type loci in yeasts and ascomycotan filamentous fungi (Pezizomycotina) encode master transcriptional factors that play a critical role in sexual development. Genome-wide analyses of mating-type-specification circuits and mating-type target genes are available in *Saccharomyces cerevisiae* and *Schizosaccharomyces pombe*; however, no such analyses have been performed in heterothallic (self-incompatible) Pezizomycotina. The heterothallic fungus *Podospora anserina* serves as a model for understanding the basic features of mating-type control. Its *mat+* and *mat-* mating types are determined by dissimilar allelic sequences. The *mat-* sequence contains three genes, designated *FMR1*, *SMR1* and *SMR2*, while the *mat+* sequence contains one gene, *FPR1*. *FMR1* and *FPR1* are the major regulators of fertilization, and this study presents a genome-wide view of their target genes and analyzes their target gene regulation.

Methodology/Principal Findings: The transcriptomic profiles of the *mat+* and *mat-* strains revealed 157 differentially transcribed genes, and transcriptomic analysis of *fmr1-* and *fpr1-* mutant strains was used to determine the regulatory actions exerted by *FMR1* and *FPR1* on these differentially transcribed genes. All possible combinations of transcription repression and/or activation by *FMR1* and/or *FPR1* were observed. Furthermore, 10 additional mating-type target genes were identified that were up- or down-regulated to the same level in *mat+* and *mat-* strains. Of the 167 genes identified, 32 genes were selected for deletion, which resulted in the identification of two genes essential for the sexual cycle. Interspecies comparisons of mating-type target genes revealed significant numbers of orthologous pairs, although transcriptional profiles were not conserved between species.

Conclusions/Significance: This study represents the first comprehensive genome-wide analysis of mating-type direct and indirect target genes in a heterothallic filamentous fungus. Mating-type transcription factors have many more target genes than are found in yeasts and exert a much greater diversity of regulatory actions on target genes, most of which are not directly related to mating.

Citation: Bidard F, Aït Benkhali J, Coppin E, Imbeaud S, Grognet P, et al. (2011) Genome-Wide Gene Expression Profiling of Fertilization Competent Mycelium in Opposite Mating Types in the Heterothallic Fungus *Podospora anserina*. PLoS ONE 6(6): e21476. doi:10.1371/journal.pone.0021476

Editor: Jürg Bähler, University College London, United Kingdom

Received: March 2, 2011; **Accepted:** May 29, 2011; **Published:** June 28, 2011

Copyright: © 2011 Bidard et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Funding: This study and the salary of FB were funded by the French National Research Agency (<http://www.agence-nationale-recherche.fr/>) grant number ANR-05-BLAN-0385, project SexDevMycol, coordinator R. Debuchy. The microarray scanner was provided by the Programme Pluri-Formation (PPF) of Univ Paris-Sud 11 "Plate-forme Puces à ADN Gif-Orsay 2005–2008". The <http://podospora.igmors.u-psud.fr/methods.php> address was hosted on a server funded by IFR115. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

* E-mail: robert.debuchy@igmors.u-psud.fr

9 These authors contributed equally to this work.

^a Current address: IBENS-Institut de Biologie de l'École Normale Supérieure, Plate-Forme Transcriptome, Paris, France

^b Current address: INSERM, Génomique Fonctionnelle des tumeurs solides, UMR U-674, IUH, Université Paris-Descartes, Paris, France

Introduction

Sexual reproduction is driven by specific gene expression programs required for identifying a sexually compatible partner, mating, meiosis and generation of progeny. Regulation of the genes involved in the identification of the mating partner and mating is well understood at a genome-wide level in the budding yeast *Saccharomyces cerevisiae* [1] and the fission yeast *Schizosaccharomyces pombe* [2], and these analyses have identified a limited

number of cell specific-genes, most of which have a known function. Filamentous fungi provide an opportunity to study mating-type evolution and function in multicellular organisms, which is not possible in unicellular yeasts; however, no complete description of the regulation exerted by mating-type regulatory genes is available in the filamentous Ascomycetes. Furthermore, our understanding of the functions controlled by the mating-type transcription factors is very incomplete as most of their target genes remain unknown.

To help clarify the role of mating types, sexual development is being studied in an increasing number of filamentous fungi, such as *Podospora anserina* (reviewed in [3]), *Neurospora crassa* [4,5,6], *Sordaria macrospora* [7], *Gibberella zeae* (anamorph *Fusarium graminearum*) [8,9], *G. moniliformis* (anamorph *F. verticillioides*) [10], *Aspergillus nidulans* [11], *Cochliobolus heterostrophus* (reviewed in [3]) and *Tuber melanosporum* [12]. In contrast to the yeasts, Pezizomycotina develop complex female organs and most heterothallic (self-incompatible) fungi also produce male cells, which can fertilize the female organs of the opposite mating type. This fertilization event relies on pheromone receptor systems, which resemble the archetypal system of peptidic pheromones found in the budding yeast [13,14,15,16,17,18]. After fertilization, the female organ undergoes a series of complex differentiation events leading to the formation of several hundred asci, each one resulting from an independent meiotic event [19].

Sexual compatibility in heterothallic filamentous fungi is controlled by a single mating-type locus with two dissimilar allelic sequences, also termed idiomorphs [20]. One idiomorph is characterized by the presence of a gene encoding a transcription factor with a MAT α _HMG domain [21], which was initially identified in the MAT α 1p protein of *S. cerevisiae* [22]. This gene is called *MAT1-1-1* in the standard nomenclature [23] and defines the *MAT1-1* idiomorph. The other idiomorph, *MAT1-2*, is characterized by the presence of a *MAT1-2-1* gene which encodes a transcription factor with a MATA_HMG domain. *MAT1-1-1* and *MAT1-2-1* are essential for fertilization in heterothallic Pezizomycotina [10,24,25,26,27] and development of the fruiting body [28,29,30,31,32].

Various other idiomorphic genes have been described in Pezizomycotina (see [33] for a review), notably *MAT1-1-2* and *MAT1-1-3*. *MAT1-1-2* proteins contain a PPF domain, which is characterized by the conservation of three invariant residues, two prolines and one phenylalanine [34], while *MAT1-1-3* transcription factors are characterized by an HMG-box and form a subgroup within MATA_HMG [21]. Their roles have been investigated in *P. anserina* [29,32] and *N. crassa* [4,35,36], and they are required for post-fertilization development of the fruiting body. Despite numerous genetic analyses of mating-type gene functions, only a few mating-type target genes, essential for mating, have been identified. These genes are involved in the conserved pheromone/receptor system initially found in budding yeast.

The control of cell-type specificity exerted by mating-type genes has been examined in detail through genetic approaches in the heterothallic filamentous ascomycete *P. anserina* (see Figure 1 for the correspondence of *P. anserina* and standard nomenclature for *MAT* genes). These studies revealed that the main regulators of fertilization, *FMR1* and *FPR1*, have activator and repressor activities on the functions required for fertilization (Figure 1) ([13,28] and reviewed in [3]). Two other genes, *SMR1* (*MAT1-1-2*) and *SMR2* (*MAT1-1-3*), are present in the *mat*⁻ idiomorph. These two genes are not involved in the activation of the *mat*⁻ functions required for fertilization, although the MATA_HMG transcription factor, *SMR2*, plays a minor role in the repression of the *mat*⁺ fertilization functions in *mat*⁻ strains. Subsequent studies have identified only two target genes of *FMR1* and *FPR1*. These target genes encode the MFM and MFP pheromone precursors and are transcriptionally activated in a mating type-specific manner (Figure 1) [13]. Coppin *et al.* [13] have suggested that genes encoding pheromone precursor processing proteins are subjected to repression in a mating-type specific manner (Figure 1), but no such genes have been identified. The dual regulatory action of mating-type genes initially reported in *P. anserina* was also reported in *G. moniliformis* for target genes not directly involved in mating [10].

In this study, we present a transcriptomic genome-wide identification of the genes that are differentially transcribed between the *mat*⁺ and *mat*⁻ strains in a vegetative stage competent for fertilization. A total of 157 genes were identified and repression or activation of these genes was determined by analyzing the transcriptomic profile of *fpr1*⁻ and *fmr1*⁻ mutants. The data provided, for the first time, a complete description of the regulation exerted by mating-type genes on their target genes in a heterothallic Pezizomycotina. Many of the target genes are not involved in mating, in agreement with previous observations in *G. moniliformis* [10] and *S. macrospora* [7,37], and deletion of 32 selected genes identified only two genes essential for mating. In addition, the search for common mating-type target genes in *P. anserina* and *G. moniliformis* or *S. macrospora* revealed statistically significant numbers of orthologous pairs; however, these conserved target genes have different transcriptional profiles.

Results

Time-course RT-qPCR analysis of mating-type gene transcript level during vegetative growth

FMR1 and *FPR1* play essential roles in mating-type determination and fertilization (see Introduction). An RT-qPCR experiment was performed to determine the transcription pattern of these two genes and to identify the optimal time to search for differentially expressed genes in *mat*⁻ and *mat*⁺ mycelia. The transcript levels of mating-type genes were investigated on *mat*⁺ and *mat*⁻ mycelia harvested after incubation in Petri dishes for 24 h, 48 h, 72 h, 96 h and 120 h at 27°C under constant light. An additional experiment was performed with *mat*⁻ and *mat*⁺ strains fertilized by spermatization at 96 h and harvested 48 h later. At this time point, perithecia form ascogenous hyphae and nuclei undergo meiosis (Frédérique Bidard and Véronique Bertheaux-Lecellier, unpublished results).

The quantification cycle (C_q) values for *FMR1* and *FPR1* cDNAs were very high at 24 h, indicating that their expression was very low at the beginning of vegetative growth (Table S1). The transcript levels of these two genes increased markedly at 48 h and reached a plateau at 72 h (Figure 2A and B). The nine colonies grown in a Petri dish made contact at 48 h (Figure 2C); thus, maximal *FMR1* and *FPR1* transcript levels occurred 24 h after confluence. A decrease in *FMR1* and *FPR1* transcript levels was observed 48 h after fertilization.

Transcript levels of the *SMR1* and *SMR2* genes were undetectable throughout vegetative development until 120 h, at which time very low levels of transcripts were detectable in the *mat*⁻ mycelium (Table S1). The transcript levels of these two genes increased strongly after fertilization (Figure 2A) in agreement with their essential roles during perithecium development [32,38]. In line with our usual experimental conditions, fertilization was performed typically with 96 h-old mycelia, corresponding to the middle of the plateau phase of maximum transcript accumulation for *FMR1* and *FPR1*. This time-point was thus used in subsequent experiments to determine the genes that were differentially transcribed in *mat*⁻ and *mat*⁺ cultures.

Transcriptomic profiling of *mat*⁻ and *mat*⁺ strains

A transcriptomic comparison of different mating-type strains provided an in-depth analysis of the genes that are differentially transcribed in the *mat*⁻ and *mat*⁺ strains. RNA was extracted from cultures grown under constant light for 96 h and prepared for hybridization (see Materials and Methods) to whole genome gene-expression microarrays developed for *P. anserina* [39].

Genes were defined as up-regulated in the *mat*⁺ strain if their transcript levels showed a ≥ 2 -fold change ($FC \geq 2$) with a p-value

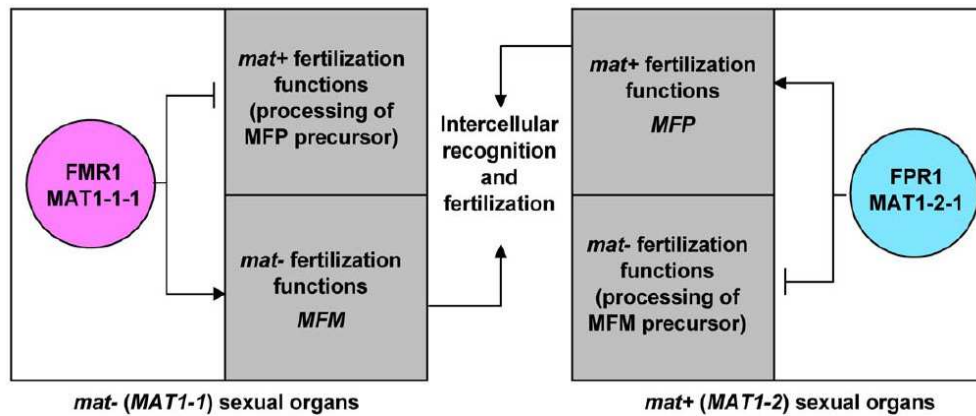


Figure 1. Regulation of fertilization by mating-type genes in *P. anserina*. Mating-type protein names are enclosed in colored circles: magenta, MAT α -HMG protein; cyan, MAT α -HMG proteins. Grey squares represent target genes present in *mat+* and *mat-* strains. The standard nomenclature is indicated below the *P. anserina*-specific gene names. *MFM* and *MFP* encode the pheromone precursors [13]. Arrows with heads and blunt ends indicate target-gene activation and repression, respectively. Data were compiled from [13,28]. doi:10.1371/journal.pone.0021476.g001

of <0.005 in the *mat+* strain compared to the *mat-* strain (*mat+* vs *mat-* comparison). Reciprocally, genes were defined as up-regulated in the *mat-* strain when they showed a ≤ 2 -fold change ($FC \leq -2$) with a p-value of <0.005 in the *mat+* vs *mat-* comparison. A total of 157 genes were differentially transcribed in the *mat+* vs *mat-* comparison (Table S2); 88 genes accumulated more transcripts in *mat+* than in *mat-* strains (Table S3), and 69 genes accumulated more transcripts in *mat-* than in *mat+* strains (Table S4).

Next, we searched for genes expressed in a mating-type specific manner. Assuming that genes not expressed in one mating type should display a signal-to-standard-deviation ratio ($SSR \geq 3$) [39], 13 genes were identified that were transcribed exclusively in the *mat+* strain and six genes were identified that were transcribed exclusively in the *mat-* strain (Table 1). Two of these genes have an ortholog with a known function in *S. cerevisiae*. *STE6*, the ortholog of Pa_5_11640 (ABC transporter), is involved in the transmembrane export of the lipophilic pheromones [40], and *FBP1*, the ortholog of Pa_4_9360 (fructose-1,6-bisphosphatase), which has no reported role in yeast mating. These 19 genes with mating type-specific expression did not include genes encoding the pheromone precursors and the pheromone receptors. This is in agreement with previous genetic analyses, which suggested that, although *mat+* and *mat-* pheromone genes are transcribed at low levels in *mat-* and *mat+* vegetative nuclei, respectively, tight repression of the expression of these genes in inappropriate cells takes place at the post-transcriptional level [13]. Interestingly, a low level of transcription was reported for an *N. crassa* pheromone receptor gene in inappropriate mating-type cells and a role in pheromone gene transcription was evidenced [41].

The consistency of microarray data with RT-qPCR analyses was tested for 14 genes selected to cover a wide range of FC values. Eleven genes displayed a similar FC direction in both experiments (Table S5). As commonly observed in the validation of array data, the FC values were much greater in the RT-qPCR than in the microarray experiments for the genes that were induced strongly [42,43]. The strongest differences were observed for the *MFM* and *MFP* transcripts, which yielded dramatically low Cq values in the *mat-* and *mat+* strains, respectively (Table S6), suggesting that these two genes may be among those with the highest transcription level at this stage of development. Similarly, the pheromone precursor gene of *N. crassa*, *mfa-1*, was identified as the most

abundant clone in starved mycelial cDNA libraries [18]. Of the 14 genes selected for microarray validation, three genes had an FC value close to one with non-significant p-values; microarray experiments gave typically more reliable results than RT-qPCR for genes that displayed low induction levels.

Transcriptomic profiling of *fmr1*⁻ and *fpr1*⁻ mutant strains

Two types of genetic differences can account for differential transcription in *mat+* and *mat-* strains. The first type includes the two mating-type genes that are expressed at 96 h, namely *FPR1* in the *mat+* strain and *FMR1* in the *mat-* strain. Another possible type of genetic difference is the single nucleotide polymorphism (SNP) between *mat+* and *mat-* strains. In these isogenic strains, the polymorphism is restricted to the environment of the mating-type locus and may be associated with differential gene transcription in *mat+* and *mat-* strains. The effect of *FMR1* and *FPR1* on the transcription of the 157 genes, differentially transcribed in *mat+* and *mat-* strains, was determined by transcriptomic profiling of the *fmr1*⁻ and *fpr1*⁻ strains [44]. The *mat-* vs *fmr1*⁻ and *mat+* vs *fpr1*⁻ comparisons (Table S2) enabled the identification of the target genes (direct or indirect) of *FMR1* and *FPR1*, respectively. Genes that were differentially transcribed due to mating-type linked SNPs were differentially transcribed in the *fpr1*⁻ vs *fmr1*⁻ comparison (Table S2). Taken together, these comparisons allowed us to determine the contribution of *FMR1*, *FPR1* and SNPs towards the transcription of each gene differentially expressed in *mat+* and *mat-* strains.

The 157 genes differentially transcribed in the *mat+* vs *mat-* strains were divided into eight classes based on their transcription patterns (Table S7), and the regulatory effect of *FMR1* or *FPR1* (activation or repression) on each target gene was inferred from the FC values. It should be noted that the technology used in this study did not enable us to determine whether *FMR1* and *FPR1* were direct or indirect regulators of their target genes. The number of genes in each class and the control exerted by the mating-type genes are summarized in Figure 3.

Class 1 contained 13 genes differentially transcribed in *mat+* and *mat-* strains as a consequence of the mating-type linked SNPs and independent of any control by mating-type proteins. This class included the *FPR1* (Pa_1_20590) and *FMR1* genes. These two

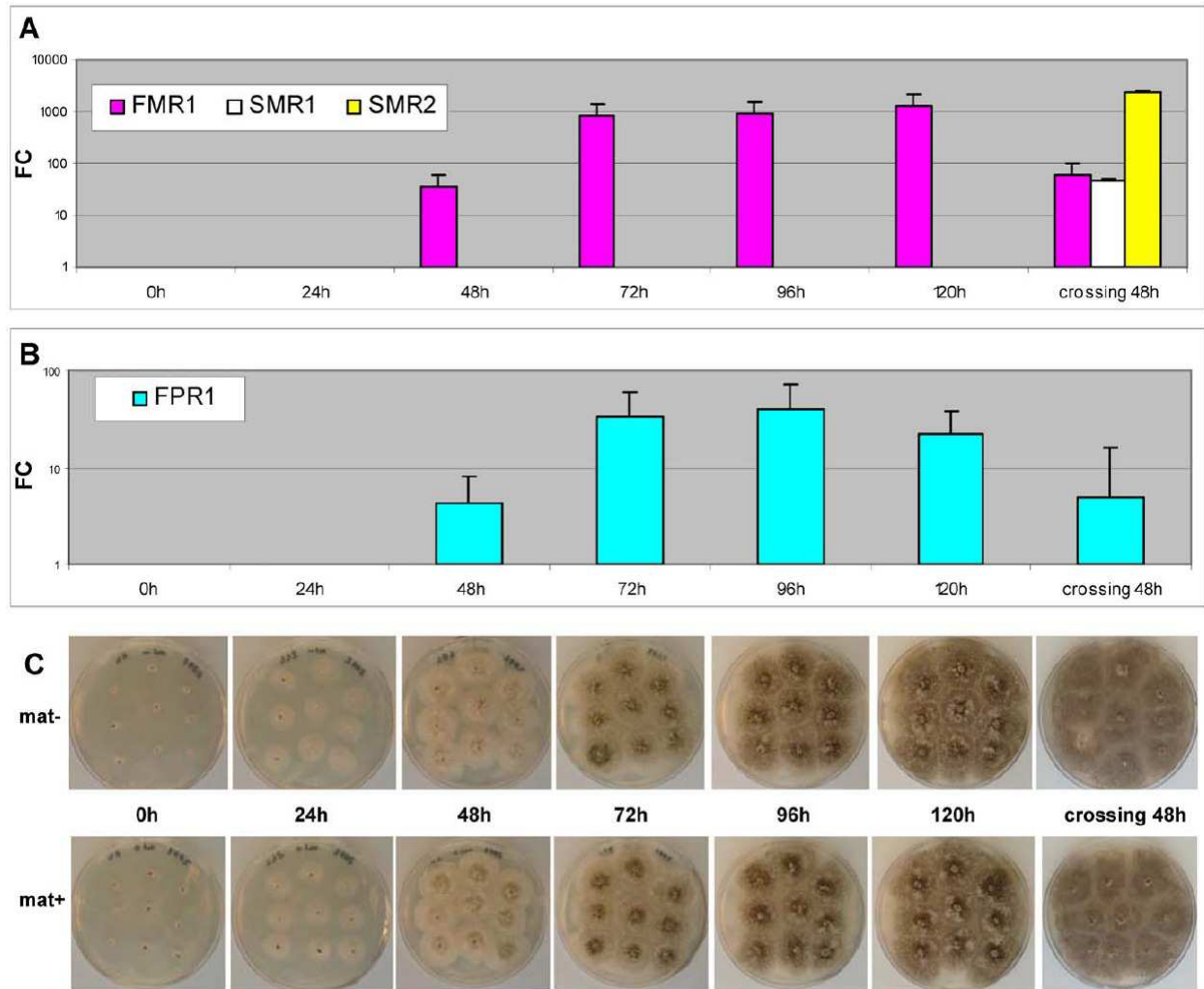


Figure 2. Time-course RT-qPCR analysis of mating-type gene transcript levels during vegetative growth and fruiting-body development. **A:** Relative quantification of *mat-* mating-type gene transcription using the Relative Expression Software Tool (REST) [69]. RNA was extracted from *mat-* colonies harvested during vegetative growth (24 h, 48 h, 72 h, 96 h and 120 h) and from *mat-* colonies grown for 48 h after fertilization with *mat+* microconidia (crossing 48 h). The reference time for *FMR1* fold-change determination was 24 h. The reference time for *SMR1* and *SMR2* fold-change determination was 120 h. Quantification cycle (Cq) values used for relative quantification are in Table S1. **B:** Relative quantification of *mat+* mating-type gene transcription with REST [69]. RNA was extracted from *mat+* colonies harvested during vegetative growth (24 h, 48 h, 72 h, 96 h and 120 h) and from *mat+* colonies grown for 48 h after fertilization with *mat-* microconidia (crossing 48 h). The reference time for *FPR1* fold-change determination was 24 h. Cq values used for relative quantification are shown in Table S1. **C:** Time course development of colonies. Upper row of Petri dishes: *mat+* strain. Bottom row of Petri dishes: *mat-* strain. doi:10.1371/journal.pone.0021476.g002

genes were not differentially transcribed in the *mat+* vs *fpr1*⁻ and *mat-* vs *fmr1*⁻ comparisons, indicating that their transcription was not under the positive or negative control of the proteins that they encode. Classes 2 to 7 displayed all possible combinations of control by FMR1, FPR1 and mating-type linked SNPs. Strikingly, class 2 contained genes, which were differentially transcribed in *mat+* and *mat-* strains due to the effects of the three factors: FPR1 (in *mat+* strain), FMR1 (in *mat-* strain), and mating-type linked SNPs (in either *mat+* or *mat-* strains). Class 8 contained 41 genes differentially transcribed in *mat+* and *mat-* strains, but their control by FMR1, FPR1 or mating-type linked SNPs could not be clearly established because their FC scores or p-values were outside the selected significant values. A total of nine genes had FC scores close to the cut-off values ($-2 < FC < -1.8$ or $1.8 < FC < 2$) in the *mat+* vs *fpr1*⁻ comparison, as exemplified by Pa_4_5280

($FC = -1.97$, $P = 0.00048$). The highly significant p-values associated with these FCs indicated that these genes were controlled by FPR1. Similarly, five genes appeared to be under the control of FMR1, based on their FCs and significant p-values in the *mat-* vs *fmr1*⁻ comparison (e.g. Pa_1_3100, $FC = 1.97$, $P = 0.00056$). The analysis of FC values facilitated the identification of five additional genes from class 8, which were controlled by both mating-type genes. For each of these genes, FCs of wild-type-to-mutant comparisons were below the cut-off value, but their combination resulted in an FC value that corresponded to that observed in the *mat+* vs *mat-* comparison. This was exemplified by Pa_7_9690 (protein-S-isoprenylcysteine O-methyltransferase), activated by FPR1 ($FC = 1.69$ in the *mat+* vs *fpr1*⁻ comparison) and repressed by FMR1 ($FC = -1.75$ in the *mat-* vs *fmr1*⁻ comparison). Activation by FPR1 and repression by FMR1 were combined in

Table 1. Main features of genes specifically transcribed in *mat+* (FC \geq 2) or *mat-* (FC \leq -2) strains.

Gene number ^a	FC	Gene name or function	Class	FPR1 ^b	FMR1 ^c
Pa_4_3858 (Δ)	48.16	unknown function	4	A	0
Pa_1_24410 (Δ)	27.87	SAM-dependent methyltransferase	4	A	0
Pa_5_3435 (Δ)	7.76	unknown function	4	A	0
Pa_5_9770 (Δ)	5.96	Asp protease	4	A	0
Pa_2_13290	3.42	unknown function	4	A	0
Pa_1_16335	3.28	unknown function	5	0	R
Pa_4_9360	2.67	fructose-1,6-bisphosphatase	8	0	0
Pa_2_8800	2.64	unknown function	4	A	0
Pa_6_9980	2.53	unknown function	8	A	0
Pa_5_11640	2.34	ABC transporter	8	0	0
Pa_7_2870	2.31	unknown function	4	A	0
Pa_1_20510	2.05	sugar transporter	2	A	R
Pa_2_7590	2.02	unknown function	8	0	0
Pa_3_10430	-2.08	unknown function	8	0	0
Pa_1_6014	-2.62	unknown function	1	0	0
Pa_4_1433	-2.64	unknown function	8	0	0
Pa_3_10350	-3.61	unknown function	8	0	0
Pa_6_10330 (Δ)	-3.91	polyketide synthase	5	0	A
Pa_4_1290	-5.05	carbohydrate esterase	5	0	A

^aGenes selected for deletion are marked with Δ .

^bA: gene induced by FPR1; R: gene repressed by FPR1; 0: gene not controlled by FPR1.

^cA: gene induced by FMR1; R: gene repressed by FMR1; 0: gene not controlled by FMR1.

doi:10.1371/journal.pone.0021476.t001

the *mat+* vs *mat-* comparison, yielding a 2.95-fold change. Following the same rationale, two genes (Pa_4_4560 and Pa_6_6730) were repressed by FPR1 and activated by FMR1, and two genes (Pa_4_3860 and Pa_2_11120) were activated by FPR1 and induced by SNPs linked to the *mat+* mating type. None of the remaining genes in class 8 displayed differential transcription in the *mat+* vs *mat-* comparison, and FCs close to a value of one with significant p-values in the other comparisons. This precluded the existence of an undetermined regulatory factor that controls class 8 genes in addition to mating-type genes and SNPs. An additional ninth class (Table S7 and Figure 3) contained mating-type target genes that were not differentially transcribed in the *mat+* vs *mat-* comparison (e.g. Pa_5_4645, FC = 1.33) because they were activated or repressed to a similar level in *mat+* and *mat-* strains (e.g. Pa_5_4645, FC = 3.46 in the *mat+* vs *fpr1-* comparison and FC = 3.4 in the *mat-* vs *fmr1-* comparison). This class contained 10 genes with diverse functions (Table 2). Among the 167 genes examined in total, 132 genes were ultimately controlled by FMR1 or FPR1 or by both transcription factors, as shown in Figure 3B. The remaining 35 genes (13 genes from class 1 and 22 genes from class 8) were not included in this clustering. Class 1 consisted of genes exclusively under the control of mating-type linked SNPs, while the remaining genes from class 8 had FC values below the cut-off or not significant p-values and could not be included, with confidence, as genes controlled by FMR1, FPR1 or mating-type linked SNPs.

The transcriptomic profiling of wild-type vs mutant strains revealed many more genes than the 157 that were differentially

transcribed in the *mat+* vs *mat-* comparison (Figure 4 and Table S2). The mutant strains may thus have accumulated adventitious mutations in addition to the *fpr1-* and *fmr1-* mutations. Ten genes were selected from the *fpr1-* and *fmr1-* strains to identify such putative mutations by sequencing (see Materials and Methods). No mutations were found in these ten genes, suggesting that their differential transcription might be the indirect result of adventitious mutations that have yet to be identified. The complete sequencing of the genome for the *fpr1-* and *fmr1-* strains is underway to identify these putative mutations.

Functional annotation of mating-type regulated genes

The published functional annotation of the 157 genes, differentially transcribed in the *mat+* and *mat-* strains (class 1 to 8), and the 10 genes that were up- or down-regulated to similar levels in both strains (class 9) was completed for the proteins with unknown functions. Each of these protein sequences was searched against the non-redundant protein database at NCBI using BLASTP [45] and their putative functions were inferred from conserved domains (Table S8 and Table S9). The Functional Catalogue (FunCat) created by MIPS [46] was used to group the genes, differentially transcribed in *mat+* and *mat-* strains, according to their cellular or molecular functions (Table 3). Chi-squared tests indicated that the FunCat main functional categories were not differently distributed in the *mat+* and *mat-* strains. Approximately 50% of these proteins were of unknown function (FunCat class number 99). Among these proteins, eight did not produce a hit on any other protein and were specific to *P. anserina* (Table S8 and Table S9). Three genes (*nad1*, 2, 3) belonged to the mitochondrial genome and encoded the three sub-units of NADH dehydrogenase [47]. They were placed in class 1 (genes under the control of mating-type linked SNPs, Figure 3A and Table S7).

The search for genes involved in the sexual process identified genes encoding the pheromones, their receptors and some genes required for pheromone biogenesis. Five of these genes were up-regulated in the *mat+* strain: the genes encoding the *mat+* lipophilic pheromone (*MFP*, Pa_2_2310) and the *mat-* pheromone receptor (*PRE2*, Pa_4_1380), and the orthologs of *RAM1* (farnesyl transferase β , Pa_4_7760), *STE24* (protein-S-isoprenylcysteine O-methyltransferase, Pa_7_9690) and *STE6* (ABC transporter, Pa_5_11640), which are required for the biogenesis of the lipophilic pheromone in *S. cerevisiae* [48]. The first three genes clustered with the genes activated by FPR1 (class 4, Table S7), while the two last genes were activated by FPR1 and repressed by FMR1 (class 8, Table S7). The p-value for Pa_5_11640 (ABC transporter, *STE6*) was >0.005 but below the 0.05 upper limit of significance (class 8, Table S7). The prenylation of the *S. cerevisiae* lipophilic pheromone requires also the α subunit of farnesyl transferase, which is encoded by *RAM2* [49]. The ortholog of *RAM2* (Pa_2_9270) was not among the 157 genes that were differentially transcribed in *mat+* and *mat-* strains because its up-regulation in the *mat+* strain was below the cut-off value (FC = 1.59). However, the p-value for this gene was highly significant ($P = 0.00001$) (Table S2), suggesting that this gene may be considered as up-regulated in the *mat+* strain. Orthologs of the other genes required for lipophilic pheromone biogenesis were either not differentially regulated in *mat+* vs *mat-* strains (*RCE1*, *AFC1/STE24*, *STE14* and *STE23*) or were not found in the *P. anserina* genome (*AXL1*). Two genes involved in the sexual process were up-regulated in the *mat-* strain and were activated by FMR1: the genes encoding the *mat-* hydrophilic pheromone (*MFM*, Pa_1_8290) and the *mat+* pheromone receptor (*PRE1*, Pa_7_9070) (class 5, Table S7). Orthologs of the *KEX2* and *KEX1* genes, which are required for hydrophilic pheromone biogenesis in

A

Genes controlled by :	Class								
	1	2	3	4	5	6	7	8	9
FPR1	0	+	+	+	0	0	+	0	S
FMR1	0	+	+	0	+	+	0	0	S
MAT-linked SNPs	+	+	0	0	0	+	+	0	0
Class size	13	7	5	71	18	1	1	41	10

B

Genes controlled by	Type of control								
	FPR1	A		R		A		R	
	FMR1	A		R		A		R	
Class 2							3	4	
Class 3						1	1	3	
Class 4	57	14							
Class 5			17	1					
Class 6			1						
Class 7	1								
Class 8	5	6	5				1	2	
Class 9						7	3		
Total	63	20	23	1	8	3	5	9	

Figure 3. Partitioning of *mat+* vs *mat-* differentially transcribed genes. **A:** Partitioning of differentially transcribed genes into classes controlled by mating-type genes or mating-type linked SNPs. Genes with $FC \geq 2$ or $FC \leq -2$ ($P < 0.005$) in *mat+* vs *fpr1-*, *mat-* vs *fmr1-* and *fpr1-* vs *fmr1-* comparisons are considered as being controlled by FPR1, FMR1 and mating-type linked SNPs, respectively. 0: no regulation; +:induction or repression; s: similar control in *mat+* and *mat-* strains. **B:** Partitioning of the mating-type target genes into groups with identical transcriptional patterns. Genes with $FC \geq 2$ in *mat+* vs *fpr1-* and *mat-* vs *fmr1-* comparisons were considered as activated (A) by FPR1 and FMR1, respectively. Genes with $FC \leq -2$ in *mat+* vs *fpr1-* and *mat-* vs *fmr1-* comparisons were considered as repressed (R) by FPR1 and FMR1, respectively. doi:10.1371/journal.pone.0021476.g003

Table 2. Main features of genes up- or down-regulated to a similar level in *mat+* and *mat-* strains.

Gene number	FC ^a	FC ^b	Gene name or function	Class	FPR1 ^c	FMR1 ^d
Pa_2_4890	3.93	3.61	taurine catabolism dioxygenase	9	A	A
Pa_5_4645	3.46	3.4	NUDIX domain-containing protein	9	A	A
Pa_5_11800	3.08	2.85	MFS transporter	9	A	A
Pa_2_610	2.52	2.28	geranylgeranyl diphosphate synthase	9	A	A
Pa_0_495	2.33	2.05	unkown function	9	A	A
Pa_2_4855	2.08	2.39	unkown function	9	A	A
Pa_4_4010	2.07	2.39	nucleoside hydrolase	9	A	A
Pa_5_6710	-2.28	-2.17	unkown function	9	R	R
Pa_7_10775	-2.27	-2.9	unkown function	9	R	R
Pa_6_6690	-2.2	-2.09	acetyl transferase	9	R	R

^aFC values in *mat+* vs *fpr1-* comparison.

^bFC values in *mat-* vs *fmr1-* comparison.

^cA: gene activated by FPR1; R: gene repressed by FPR1.

^dA: gene activated by FMR1; R: gene repressed by FMR1.

doi:10.1371/journal.pone.0021476.t002

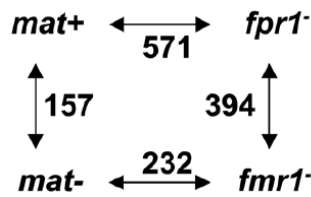


Figure 4. Number of differentially transcribed genes in wild-type and mutant strain comparisons. Cut-off values: $FC \leq -2$ or $FC \geq 2$ and $P < 0.005$.
doi:10.1371/journal.pone.0021476.g004

S. cerevisiae [48] were not differentially transcribed in the *mat+* and *mat-* strains. Pa_6_7350, the only protease-encoding gene to be up-regulated in the *mat-* strain, had no ortholog in *S. cerevisiae* and no ortholog was found in *P. anserina* for the aminopeptidase *STE13*.

Phenotypic analyses of target-gene deletions

The genes that displayed an FC value higher than 5 in the *mat+* vs *mat-* comparison (Table S3) were selected for deletion and phenotype analysis, while two genes, Pa_2_2310 (*MFP*) and Pa_3_1710 (*AOX*), were deleted in previous studies [13,50]. Fourteen additional genes from this list were selected based on (i) putative roles in reproduction (e.g. Pa_4_7760, farnesyltransferase β , RAM1), (ii) conservation in a wide range of fungi (e.g.

Table 3. Distribution of the 157 genes differentially transcribed in *mat+* vs *mat-* in FunCat main functional categories.

FunCat main functional category ID	FunCat main functional category	Number of genes accumulating transcript in <i>mat-</i>	Number of genes accumulating transcript in <i>mat+</i>
Metabolism			
01	Metabolism	17	25
02	Energy	5	8
04	Storage protein		
Information pathways			
10	Cell cycle and DNA processing		
11	Transcription	1	4
12	Protein synthesis		
14	Protein fate (folding, modification, destination)	1	2
16	Protein with binding function	5	5
18	Protein activity regulation		1
Transport			
20	Cellular transport, transport facilitation and transport routes	6	9
Perception and response to stimuli			
30	Cellular communication/signal transduction mechanism		
32	Cell rescue, defense and virulence	6	2
34	Interaction with the environment	3	7
36	Systemic interaction with the environment		
38	Transposable elements, viral and plasmid proteins		
Developmental process			
40	Cell fate		
41	Development (systemic)		
42	Biogenesis of cellular components	1	3
43	Cell type differentiation		
45	Tissue-specific differentiation		
47	Organ differentiation		
Localization			
70	Subcellular localization		1
73	Cell type localization		
75	Tissue localization		
77	Organ localization		
Unclassified			
98	Classification not yet clear-cut		
99	Unclassified proteins	37	45

doi:10.1371/journal.pone.0021476.t003

Pa_2_7180), or (iii) functions of special interest (e.g., the C6 transcription factors, Pa_2_6830 and Pa_6_3770). Nine additional candidates were selected among the genes that were up-regulated in the *mat*⁻ strain (Table S4) based on similar criteria. The phenotype of the Pa_1_8290 (*MFM*) deletion was investigated in a previous study [13].

A total of 32 deletion strains were constructed and examined for male and female fertility and for possible defects in perithecium development (Table 4). The two mutant strains deleted for the *mat*⁺ lipophilic pheromone receptor (*PRE1*, Pa_7_9070) and the *mat*⁻ hydrophilic pheromone receptor (*PRE2*, Pa_4_1380) showed defects in sexual reproduction. The loss of the receptor for the *mat*⁺ lipophilic pheromone resulted in female sterility in *mat*⁻ strains, while Δ *PRE1 mat*⁺ strains displayed wild-type fertility. Identical phenotypes were reported in *N. crassa* strains that were deleted for the lipophilic pheromone receptor, *pre-1* [41]. The *PRE2* deletion resulted in reciprocal phenotypes, namely female sterility in *mat*⁺ strains, while *mat*⁻ strains were not affected. The deletion of Pa_4_7760 (farnesyl transferase β , *RAM1*) was lethal, in contrast to the disruption of its ortholog in *S. cerevisiae* [49]. Farnesylation of Ras proteins is required for essential Ras function, suggesting that *RAM1* is an essential gene. Viable *RAM1* loss-of-function mutations in yeast have been attributed to an alternative prenylation process involving Cdc43p [49]. Similar to *P. anserina*, the deletion of *RAM1* is lethal in *Cryptococcus neoformans* [51], indicating that the farnesylation of Ras proteins cannot be compensated for by alternative pathways in these two fungi, although both contain orthologs of Cdc43p (Pa_5_4440 in *P. anserina*, CNAG_02756 in *C. neoformans*). The 29 remaining deletion strains were not impaired for mating and showed no defects in perithecium development. The mutant strains were also tested for their sensitivity to several inhibitors (see Materials and Methods) without revealing any differences between the mutant and the wild-type strains.

Interspecies comparison of mating-type target gene transcriptional profiles

A cDNA macroarray analysis was conducted in *G. moniliformis* to compare a strain in which the *MAT1-2-1* gene was deleted with the *MAT1-2* wild-type strain [10]. The mutant and wild-type strains used in these experiments were the counterparts of those used in the *mat*⁺ vs *fpr1*⁻ comparison, thus providing the opportunity to search for conserved mating-type target genes in these two heterothallic species. A total of 29 orthologous pairs were found in *G. moniliformis* *MAT1-2-1* and *P. anserina* *FPR1* target genes ($P < 0.0001$) (Table 5 and Table S10), none of which encoded a protein with a known function directly related to mating. The *P. anserina* ortholog of clone #263, PaMpk3 (Pa_1_23930), is involved in osmotic stress resistance, but has no function in sexual development (Hervé Lalucque, Fabienne Malagnac and Philippe Silar, personal communication). The cDNA macroarrays did not provide an exhaustive representation of the genome, which possibly explains the low number of orthologous pairs identified as well as the absence of the expected pheromone and receptor pheromone gene pairs. Of the 29 genes regulated by *MAT1-2-1* and *FPR1* in *G. moniliformis* and *P. anserina*, none was repressed by both transcription factors, only 12 were activated by both transcription factors and a total of 17 genes were activated by *MAT1-2-1* and repressed by *FPR1* or *vice versa*, suggesting low conservation of the transcriptional profiles in *P. anserina* and *G. moniliformis*. Further correlation analysis is precluded because specific FC values were not available for *G. moniliformis* target genes.

S. macrospora is a homothallic species that contains a *MAT α* _HMG gene (*SmtA-1*) and a *MATA*_HMG gene (*Smta-1*) homologous to *FMR1* and *FPR1*, respectively [52]. The *SmtA-1* and *Smta-1* target genes were identified in *S. macrospora*, using *N. crassa* microarrays [7,37] and a search of the common target genes of *FPR1* and *Smta-1* revealed 16 orthologous pairs ($P < 0.0001$) (Table 5 and Table S11). A total of 57 orthologous pairs were also found when searching for common target genes of *FMR1* and *SmtA-1* ($P < 0.0001$) (Table 5 and Table S12). Only one gene (NCU03727, *ham-2*), the ortholog of Pa_2_9440, was reported to be involved in sexual reproduction in *N. crassa* [53]. The *ham-2* mutants are unable to differentiate female organs but male fertility is not affected. The transcriptional profiles of each group of orthologous genes were compared as described in [54]. The calculation of correlation based on the FC values of orthologous pairs reveal no shared transcriptional patterns for *FPR1/Smta-1*, nor for *FMR1/SmtA-1* target genes (Table 5).

Discussion

Mating-type transcription factors *FMR1* and *FPR1* control a large number of target genes not directly involved in mating

The fertilization-competent mycelium analysis described in this study provides the first global view of differential gene expression in strains of opposite mating type in Pezizomycotina. These analyses revealed that the sexual differentiation correlates with the differential transcription of many genes which are not directly involved in sexual reproduction. Genome-wide transcriptomic analyses were conducted in *S. cerevisiae* and *S. pombe* to determine differentially transcribed genes in cells of different mating types. Six α -specific genes and five α -specific genes were identified in *S. cerevisiae* [1] (Figure 5) and with the possible exception of one α -specific gene, all these genes are directly involved in some aspect of mating. In *S. pombe*, Mata and Bähler identified 12 M-specific genes and 4 P-specific genes [2] (Figure 5). Three of these cell type-specific genes were not essential for mating. In *P. anserina*, 157 genes were identified that were differentially transcribed in *mat*⁺ vs *mat*⁻ strains and 10 genes were identified that were up- or down-regulated to a similar level in both strains. Thus, in *P. anserina*, the number of mating-type controlled genes (167) greatly outnumbers that in yeasts. In contrast with the cell-type specific genes in fission and budding yeasts, many *P. anserina* mating-type target genes appeared to be involved in a variety of biological processes, including general metabolism and energy production. A total of 32 genes were selected for deletion, including a few genes expressed in a mating-type specific manner. Only the genes encoding the pheromone receptors *PRE1* (Pa_7_9070) and *PRE2* (Pa_4_1380) were found to play a role in sexual reproduction, and one gene was found to be essential for viability (Pa_4_7760, farnesyl transferase β , *RAM1*). The deletion of the 29 other genes did not reveal any obvious defect in growth and/or mating. These 29 genes included three transcription factors (Pa_2_6830, Pa_4_9520 and Pa_6_3770), which were apparently unrelated to mating. Furthermore, Δ Pa_2_6830 was genetically associated with either Δ Pa_4_9520 or Δ Pa_6_3770, and both double mutants behaved as the wild-type strain (data not shown). Although systematic deletions were not performed, these results suggested that the 167 mating-type target genes identified included very few candidates essential for fertilization or subsequent developmental steps. Currently, only pheromones [13] and pheromone receptor genes have been identified as mating-type target genes essential for mating in *P. anserina*.

Two hypotheses can account for the surprisingly high number of mating type-controlled genes not directly involved in mating,

Table 4. Genes selected for deletion and phenotype of mutant strains.

Gene number	FC ^a	Gene name or function	mat specific expression	mutant phenotype
Pa_4_3858	48.16	unknown function	+	normal mating
Pa_1_24410	27.87	SAM-dependent methyltransferase	+	normal mating
Pa_4_1380	11.23	<i>PRE2</i>		<i>mat+</i> strains female sterile, <i>mat-</i> strains female fertile
Pa_5_3435	7.76	unknown function	+	normal mating
Pa_7_4100	7.5	unknown function		normal mating
Pa_3_3210	7.2	unknown function		normal mating
Pa_5_9770	5.96	Asp protease	+	normal mating
Pa_5_6960	5.94	unknown function		normal mating
Pa_1_9625	5.08	unknown function		normal mating
Pa_1_540	3.6	unknown function		normal mating
Pa_4_7760	3.03	farnesyltransferase subunit beta		essential
Pa_7_2860	2.5	cyclic-nucleotide phosphodiesterase		normal mating
Pa_2_9145	2.46	unknown function		normal mating
Pa_2_7180	2.39	unknown function		normal mating
Pa_2_9500	2.36	unknown function		normal mating
Pa_5_9790	2.35	unknown function		normal mating
Pa_1_5530	2.27	unknown function		normal mating
Pa_2_6830	2.26	C6 transcription factor		normal mating
Pa_2_3690	2.22	unknown function		normal mating
Pa_4_5450	2.17	unknown function		normal mating
Pa_4_9520	2.16	copper fist DNA binding domain protein		normal mating
Pa_2_1170	2.09	unknown function		normal mating
Pa_6_3770	2.06	C6 transcription factor		normal mating
Pa_6_7350	-6.56	protease		normal mating
Pa_7_9070	-5.62	<i>PRE1</i>		<i>mat+</i> strains female fertile, <i>mat-</i> strains female sterile
Pa_5_10930	-4.72	unknown function		normal mating
Pa_5_10935	-4.21	unknown function		normal mating
Pa_6_10330	-3.91	polyketide synthase	+	normal mating
Pa_5_10945	-3.27	unknown function		normal mating
Pa_7_875	-3.11	unknown function		normal mating
Pa_5_10940	-2.61	unknown function		normal mating
Pa_4_110	-2	unknown function		normal mating

^aFC values in *mat+* vs *mat-* comparison.
doi:10.1371/journal.pone.0021476.t004

Table 5. Cross-species searches of mating-type target genes.

Group 1 species	Group 1 transcription database (number of genes)	Group 2 species	Group 2 transcription database (number of genes)	Ortholog pairs	P value ^a	FC correlation (r)	P value ^b
<i>G. moniliformis</i>	<i>MAT1-2-1</i> vs <i>ΔMT1-2-1</i> (160)	<i>P. anserina</i>	<i>FPR1</i> vs <i>fpr1</i> ⁻ (571)	29	<0.0001	N/A ^c	N/A ^c
<i>S. macrospora</i>	<i>Smta-1</i> vs <i>ΔSmta-1</i> (90)	<i>P. anserina</i>	<i>FPR1</i> vs <i>fpr1</i> ⁻ (571)	16	<0.0001	0.27	0.30 ^d
<i>S. macrospora</i>	<i>SmtA-1</i> vs <i>ΔSmtA-1</i> (779)	<i>P. anserina</i>	<i>FMR1</i> vs <i>fmr1</i> ⁻ (232)	57	<0.0001	-0.25	0.06 ^e

^ap-value obtained from Chi-square analysis of interspecies orthologous pair distribution.

^bp-value of interspecies FC correlation analysis.

^cnot applicable because specific FC values were not available.

^dSpearmanRank Correlation: p-value = 0.29; Kendall tau Rank Correlation: p-value = 0.39.

^eSpearmanRank Correlation: p-value = 0.70; Kendall tau Rank Correlation: p-value = 0.65.

doi:10.1371/journal.pone.0021476.t005

The first is based on the observation that *P. anserina* is a pseudo-homothallic (or secondary homothallic) species: it is encountered in nature as a self-fertile mycelium containing *mat+* and *mat-* nuclei. The optimal balance of the two types of nuclei for mycelium self-fertility may require that each type of nucleus controls complementary metabolic pathways. Although no differences in growth rate have been reported between homokaryotic and self-fertile mycelia in *P. anserina* in laboratory conditions, self-fertile mycelia may be favored in nature when less easily digestible nutrients are available. The second, and not exclusive, hypothesis is based on the idea that perithecium development relies on nutrients supplied by the mycelium [33,55,56]. The mating-type genes, which control perithecium development as well as mating, may be involved in the mobilization of nutrients necessary for the developing perithecia. In particular, the differences observed in the transcription profiles of *mat+* and *mat-* strains concern the metabolism of lipids whose role during sexual reproduction has been emphasized (reviewed in [57], [58]); for example, three genes encoding esterase-lipase are up-regulated in the *mat+* strain. Furthermore, six additional *mat+* up-regulated genes involved in lipid metabolism were identified, not including those necessary for processing of the lipophilic pheromone MFP. The lipid reserves may be metabolized during vegetative growth to release fatty acids and other compounds which could be used as nutrients to provide carbon skeletons and metabolic energy during sexual reproduction [56]. Although there are many obvious transcriptional differences between *mat+* and *mat-* strains, no conspicuous differences have ever been noticed during sexual reproductions, besides cell-type specificity. We cannot exclude that some differences do exist between *mat+* and *mat-* female strains during ascogonium formation and initial development after fertilization, although subsequent developmental stages in the maturing fruiting bodies converge in *mat+* and *mat-* female strains due to the complementary information brought by the male nucleus.

Mating-type transcription factors FMR1 and FPR1 have a dual activating and repressing regulatory activity

Our analysis in *P. anserina* revealed a very complex regulatory pattern, which uses all eight possible combinations of regulation (activation and repression) by FMR1 and/or FPR1 (Figure 5). This pattern indicated that each mating-type protein has activator and repressor activity, and that either one or both mating-type proteins may control each target gene. The complex regulation exerted by mating-type genes has been anticipated by previous genetic analyses, which predicted that FPR1 acts as an activator for *mat+* function and as a repressor for *mat-* functions. FMR1 was predicted to have a symmetric but inverse activity to FPR1 (see Introduction and Figure 1). Subsequent genetic studies indicated that the repressor activity of each transcription factor may operate at the level of post-transcriptional pheromone maturation ([13] and reviewed in [3]). In agreement with this prediction, our study indicated that the protein-S-isoprenylcysteine O-methyltransferase (Pa_7_9690), which is required for the maturation of the lipophilic pheromone in yeast [48], is activated by FPR1 and repressed by FMR1 (see Results and Table S7, class 8). Pa_5_11640 (ABC transporter, STE6) could be a good candidate for a similar type of regulation, but this requires confirmation because the p-values were above the threshold. Unfortunately, this study did not succeed in identifying any gene that could be activated by FMR1 and repressed by FPR1 and was required for hydrophilic pheromone maturation.

The dual action (repression and activation) described in this study for the *P. anserina* mating-type genes is also conserved in

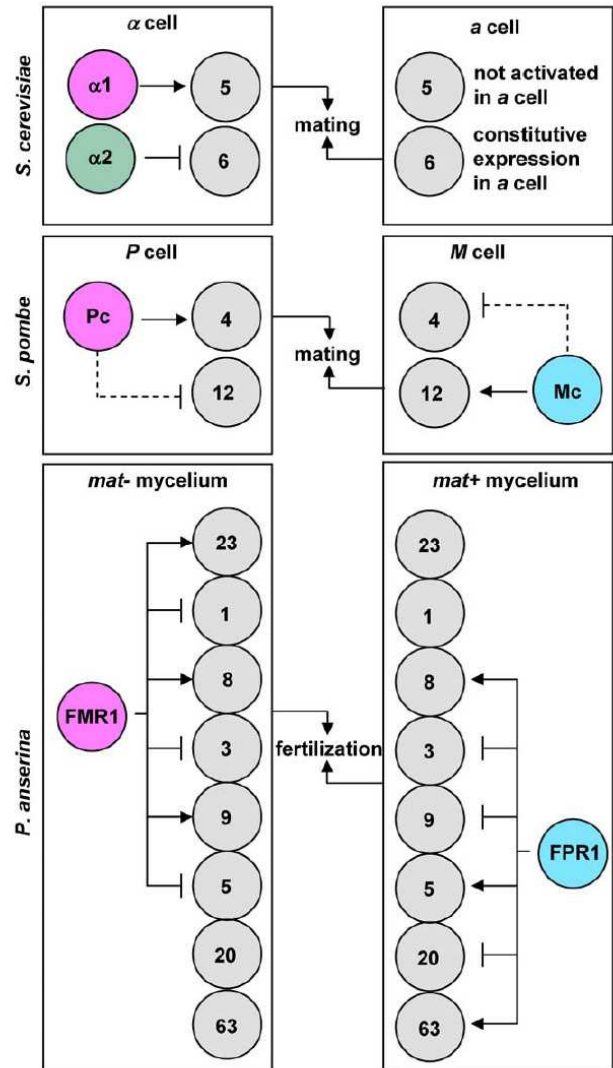


Figure 5. Cell-type regulation and target-gene numbers in *S. cerevisiae*, *S. pombe* and *P. anserina*. Mating-type protein names are enclosed in colored circles: magenta, MAT α -HMG proteins; cyan, MATA-HMG proteins; green: homeodomain protein. Grey circles indicate different types of target genes and enclosed figures indicate their number. Mating-type genes are specific to each cell type, whereas target genes are present in both cell types. Arrows with heads and blunt ends indicate target gene activation and repression, respectively, except for *a* cells from *S. cerevisiae*, which have specific regulation. Stippled arrows indicate hypothetical regulation. *S. cerevisiae* and *S. pombe* data were from compiled from [1,2]. doi:10.1371/journal.pone.0021476.g005

budding yeast (Figure 5) (reviewed in [59]) and in *G. moniliformis* [10]. The fact that the Pc and Mc proteins of *S. pombe* have been reported to use only gene inductions [2] may be due to experimental bias. Mata and Bähler have compared the transcriptome of M and P vegetative cells which overexpress Ste11p relative to corresponding M and P control cells. This enabled the authors to identify activated target genes, but the experimental design could not reveal target genes that were tightly repressed in the control cells since these genes could not be repressed further in cells overexpressing Ste11p. It is therefore possible that M-specific genes are activated in M cells and

repressed in P cells in *S. pombe*, whereas P-specific genes are subject to a symmetrical but reversed action in M cells (Figure 5). The dual action of mating-type genes (activation and repression) may be a conserved feature in Ascomycota, although both types of control are operated by two mating-type proteins in yeast and only one in *P. anserina* and *G. moniliformis*. Further analyses in *S. pombe* and dissection of mating-type regulatory circuits in other heterothallic Ascomycota will help to test this hypothesis.

Interspecies comparisons of mating-type regulated genes revealed conserved genes but no shared transcriptional pattern

The search for common target genes of the mating-type transcription factors revealed a highly significant number of orthologous pairs in the *P. anserina*/*G. moniliformis* and *P. anserina*/*S. macrospora* comparisons. This result was unexpected for SMTA-1 and FMR1, because SMTA-1 is not essential for fertilization and fruiting-body development in *S. macrospora* [7], in contrast to FMR1 in *P. anserina* [24]. However, SMTA-1 is a positive regulator of pheromone gene transcription, although it is not essential for their expression [7]. This observation suggests that SMTA-1 may be involved in the control of the same genes as essential mating-type transcription factors. Orthologous pairs of genes, which are essential for mating, have typically similar expression patterns across species. For instance, the transcription of genes encoding lipophilic pheromones is activated by MATA_HMG transcription factor FPR1 in *P. anserina* and mat a-1 in *N. crassa* [17]. Similarly, the genes encoding hydrophilic pheromones are activated by MAT α _HMG transcription factor FMR1 in *P. anserina* and mat A-1 in *N. crassa* [17]. In contrast, correlation analysis of FC values in orthologous pairs revealed that transcriptional profiles were not maintained among the numerous conserved mating-type target genes that were not involved in mating. One possible explanation to this paradoxical observation is that, for various combinations of genes, different transcriptional profiles might result in the same metabolic effect. This effect may be the mobilization of nutrients during the vegetative phase to fulfil the formation of the fruiting bodies, as proposed above. Another possible explanation is that there is a selection for the conservation of a binding site for a mating-type transcription factor at a conserved chromosomal locus, rather than selection for the function of the target genes. The inter-species variations of transcription patterns would thus reflect random variations in the expression of target genes. Conservation of these binding sites around diverse genes may be required for spatial and temporal sequestration of transcription factors, thereby contributing to the control of the expression of these regulatory proteins.

Conclusions

Our study demonstrated that mating-type genes control many more target genes than those that are directly required for fertilization. Dissection of the regulatory circuits revealed all eight combinations that are possible with two transcription regulators exerting dual control (repression or induction). It must be noted that these experiments were conducted in homokaryotic strains. The control in a heterokaryotic *mat+*/*mat-* self-fertile strain cannot be inferred from our results, as competition of FMR1 and FPR1 for binding sites may impact the transcription of target genes under dual control. Moreover, variation in the nuclear balance during the development of the heterokaryotic mycelium [60] may result in an additional layer of complexity to the temporal expression of target genes. The role of the mating-type target genes in a biological process related to sexual reproduction

is still elusive and its elucidation will require both biochemical analyses of the metabolic pathways involving these genes and ChIP experiments to determine if they are primary or secondary target genes of FMR1 and FPR1. Mating-type regulatory proteins also play essential roles in the development of the ascogenous hyphae in the fruiting body, in association with the other mating-type proteins (reviewed in [33]). Further investigations will be necessary to determine how the switch from vegetative to reproductive tissues modifies the set of target genes that are controlled by FMR1 and FPR1.

Materials and Methods

Strains

The genetic and biological features of *P. anserina* were first described by Rizet and Engelmann [61] and current culture techniques can be found at <http://podospora.igmors.u-psud.fr/methods.php>. The strains used in this study were all derived from the *S* strain [62], which was used to determine the *P. anserina* genome sequence [63]. The *fmr1*⁻ and *fpr1*⁻ mutant strains used for transcriptomic profiling have been described previously [44]. Successive nonsense mutations were introduced at the beginning of *FMR1* and *FPR1* coding sequences, resulting in loss-of-function mutations. Deletion of genes identified in transcriptomic profiling was performed by targeted gene replacement in a strain defective for Pa*KU70*, a gene necessary for the non-homologous end joining recombination [44].

Preparation of biological samples for RNA extraction

Vegetative cultures. Petri dishes containing minimal medium were covered with a cellophane sheet (cat#1650193, Bio Rad, Hercules, USA) and were inoculated with nine implants from *P. anserina mat+* or *mat-* cultures. The dishes were placed at 27°C under constant light (0 h) and were removed from the incubation room at different culture time points (24 h, 48 h, 72 h, 96 h and 120 h). With a growth rate of 7 mm/day, the mycelium reached confluence at 48 h. Mycelia were harvested by scraping with a microscope coverglass and excess water was removed with Precision Wipes (Kimberly-Clark, Reigate, United Kingdom). The mycelium was then stored in aliquots of no more than 100 mg in liquid nitrogen until RNA extraction. Quantitative PCR experiments were performed with five *mat+* plates and five *mat-* plates for each time point. Transcriptomic profiling at 96 h was performed with four biological replicates for each time point, with each plate being considered as a biological replicate.

Fertilized cultures. Plates were prepared as indicated above, except that the medium was covered with cheesecloth (Sefar Nitex 03-48/31, Sefar, Heiden, Swiss) instead of cellophane, which was completely degraded by *P. anserina* after 120 h. Crossing was performed at 96 h by spermatization: each *mat+* plate was washed with 2 ml of water to harvest the *mat+* spermatia, and the suspension was collected and poured onto a *mat-* plate to fertilize the *mat-* female organs. After applying the *mat+* spermatia suspension onto the *mat-* plate, the suspension, which contained the *mat-* spermatia, was collected and poured back onto the *mat+* plate to fertilize the *mat+* female organs. Mycelium and perithecia were harvested by scraping the cheesecloth with a microscope coverglass. Quantitative PCR experiments were performed with five *mat+* plates and five *mat-* plates. Each plate was considered to be a biological replicate.

RNA extraction

A maximum of 100 mg of mycelium was recovered from liquid nitrogen and ground in a Mikro-Dismembrator (Sartorius,

Goettingen, Germany) for 1 min at 2600 rpm in vessels containing 9 mm chromium steel grinding balls frozen in liquid nitrogen. The resulting powder was suspended in RLT buffer (700 μ l) (Qiagen, Hilden, Germany) and the suspension was passed through a Qiasredder column (Qiagen, Hilden, Germany). Total RNA was purified on RNeasy Plant Mini Kit columns (ref 74904, Qiagen, Hilden, Germany) following the manufacturer's protocol, which included DNase treatment. For microarray experiments, washing with RPE buffer (500 μ l and 350 μ l) was performed a further two times to improve RNA quality. The quantity and quality of the total RNA was determined using a NanoDrop ND-1000 spectrophotometer (Nanodrop Technologies, Wilmington, USA) and by electrophoresis on 1% agarose gel. For microarray experiments, a further control was performed on the Bioanalyzer 2100 system (Agilent, Santa Clara, USA) as described previously [64].

RT-qPCR analyses

Total RNA was reverse-transcribed with Superscript III (Invitrogen, Carlsbad, USA) and oligo(dT)₂₀ as recommended by the manufacturer. RT-qPCR amplification mixtures (10 μ l) contained 1 \times FastStart Universal SYBR Green Master (Rox) (Roche Diagnostics, Mannheim, Germany), 2 μ l of a 1/10 dilution of the reverse-transcription reaction and 500 nM of forward and reverse primers. Reactions were run on a CFX96 device (Bio Rad, Hercules, USA). The cycling conditions comprised 10 min polymerase activation at 95°C and 45 cycles at 95°C for 10 sec and 60°C for 30 sec with plate read, followed by a melting curve. Each plate contained cDNA samples representing all of the conditions examined for one gene, a standard curve of three serial dilutions points (1/10 dilutions) and a no-template control. All measures were performed with five biological replicates for each condition. For microarray validation, the biological replicates used for RT-qPCR were different from those used in microarray experiments. All of the PCR efficiencies were over 85%. At least one primer in each primer pair was positioned on two consecutive exons (Table S13), ensuring specific detection of cDNA with minimal interference by genomic DNA. *MFM* and *MFP* genes are intron-less genes and were detected with primers that do not discriminate cDNA from genomic DNA: therefore, the detection of their cDNA was not below the level of genomic DNA contamination. For these two genes, a non-reverse-transcribed control was performed on each biological replicate under the same conditions as the reverse-transcribed samples to determine the level of genomic DNA contamination. Measurements were considered as significant when the genomic contamination was five cycles above the cDNA signal.

For each RT-qPCR analysis (time course and microarray validation), eight housekeeping genes (Table S13) were screened with geNorm [65] to identify the best reference genes for each type of experiment. The minimum number of reference genes was selected that had a pairwise variation (*V*) below 0.15. If two genes were sufficient to reach a value of 0.15 for *V*, three reference genes were used according to MIQE guidelines [66], provided that the addition of the third reference gene did not increase *V* above 0.15. Pa_7_8490 (ubiquitin), Pa_4_8980 (TATA-box binding protein), Pa_3_5110 (glyceraldehyde phosphate dehydrogenase) [67] and Pa_5_5390 (histone H2A) were used for normalization of the time-course RT-qPCR (*V* = 0.15). Pa_7_8490 (ubiquitin), Pa_3_5110 (glyceraldehyde phosphate dehydrogenase) [67] and Pa_1_16650 (ribosomal protein AS1) [68] were used as reference genes for normalization of the microarray RT-qPCR validation (*V* = 0.031).

RT-qPCR normalization, standard error computation and statistical analyses were performed with REST-MCS and REST 2009 (Qiagen, Hilden, Germany) [69].

Labelling of cDNA, hybridization and microarray analyses

The construction and optimization of gene expression microarrays for *P. anserina* has been described previously [39]. Briefly, the microarrays consisted of a 4 \times 44 K platform (AMADID 018343, Agilent, Santa Clara, USA) containing 10,556 probes on each array with each probe present in four replicates.

For the transcriptome microarray experiments, target preparation, hybridization and washing were performed according to the two-color microarray-based gene expression analysis instructions (version 5.0, February 2007) as described by the manufacturer (Agilent, Santa Clara, USA). One-microgram aliquots of total RNA were amplified and Cy-labelled with Agilent's Low RNA input fluorescent linear amplification (LRILAK) PLUS kit and the Two-Color RNA Spike-in Kit (Agilent, Santa Clara, USA). The labelling efficiency and the product integrity were checked as described previously [64].

Four biological replicates labelled with Cy-3, for each of the different experimental conditions were compared with a common reference labelled with Cy-5, in indirect comparisons. The common reference was obtained by mixing RNA extracted from the different conditions as indicated in [39]. A total of 825 ng of each of the Cy3- and Cy5-labeled targets was mixed and incubated on microarray slides for 17 h at 65°C in a rotating oven (10 rpm) using an Agilent *in situ* hybridization kit (Agilent, Santa Clara, USA). The slides were washed and dried by centrifugation at 800 rpm for 1 min.

Microarrays were scanned using the Agilent DNA microarray Scanner (Agilent, Santa Clara, USA) at a resolution of 5 microns using the extended dynamic range (XDR) feature. Spot and background intensities were extracted with the Feature Extraction (FE, v9.5.3) software (Agilent, Santa Clara, USA) using the GE2-v4_95_Feb07 default protocol. Preliminary array quality was assessed through the use of Agilent control features as well as spike-in controls (Agilent 2-Color Spike-in Kit for RNA experiment). Subsequent flagging was done according to the GenePix Pro software (Molecular Devices Sunnyvale, USA) nomenclature, which included four flag levels (good [100], bad [-100], not found [-50], moderate [0]). FE-software normalized data (Lowess normalized, local background subtracted) were processed with MAnGO [70]. A moderated *t* test with adjustment of *p*-values [71] was computed to measure the significance of each difference of expression. Genes were considered as differentially transcribed if their transcript levels showed a ≥ 2 -fold change ($FC \geq 2$) or ≤ 2 -fold change ($FC \leq -2$) with a *p*-value of < 0.005 .

Microarray data accession number

All microarray data is MIAME compliant. The raw data has been deposited in the MIAME compliant Gene Expression Omnibus database [72] and is accessible through the GEO Series accession number GSE27297.

Selection and sequencing of genes from *fpr1*⁻ and *fmr1*⁻ strains

Ten genes were selected in *fpr1*⁻ and *fmr1*⁻ strains to identify putative adventitious mutations that might have occurred during the introduction of site-directed mutations in *FPR1* and *FMR1* [44]. As the two mutant strains were generated independently, it is highly unlikely that they would harbor identical mutations. Therefore, it was assumed that an adventitious mutation in a given gene should affect its expression in either one of the two mutant strains, while its expression should be similar in the three other strains considered, namely the other mutant strain, the *mat+* and the *mat-* strains. Pa_1_23560, Pa_3_11210, Pa_5_4750,

Pa_5_10440 and Pa_7_130 displayed differential transcription with highly significant p-value in the *fpr1⁻ vs fmr1⁻* and *fpr1⁻ vs mat+* comparisons but not in the *mat+ vs mat-* and *fmr1⁻ vs mat-* comparisons (Table S2); thus these genes were selected as the best putative targets for adventitious mutations in the *fpr1⁻* strain. Based on a similar rationale, Pa_1_970, Pa_1_16110, Pa_2_4340, Pa_4_4670 and Pa_6_6260 were selected as possible targets for adventitious mutations in the *fmr1⁻* strain. The selected genes were amplified from the appropriate DNA with Dream Taq (Fermentas, St Leon-Rot, Germany) in accordance with the manufacturer's instructions using primers located 1000 bp upstream and 1000 bp downstream of the translation initiation codon. Each amplification product was sequenced with upstream and downstream primers and primers targeted to regions around the initiation codon. Sequences were assembled with CAP3 [73], which is available on <http://pbil.univ-lyon1.fr/cap3.php>, and were compared with wild-type sequences using BLAST [45].

Gene annotation and the search for orthology

Protein sequences were searched against the NCBI non-redundant protein sequence library using BLASTP [45]. FUNGI-path [74] was used to search for orthologous genes on <http://embg.igmors.u-psud.fr/fungipath/>.

Gene deletion

The construction of deletion cassettes was based either on the *AscI/MluI* method or on the *N. crassa* strategy for high-throughput generation of gene deletions [75].

The *AscI/MluI* method is based on cloning an *MluI* fragment containing the resistance selection gene in an *AscI* linearized plasmid containing the 3' and 5' flanking regions of the gene that was selected for deletion. This plasmid was obtained by the amplification of a vector containing the selected gene with primers targeted to the 5' and 3' borders of the coding sequence to be deleted and with their 3' ends oriented towards the 5' and 3' untranslated regions. These diverging primers were designed with an *AscI* site at their 5' end in order to amplify the vector in a linearized form, ending with an *AscI* site and lacking the coding sequence of the selected gene. The selection gene in an *MluI* fragment was obtained by amplification of the *hph* gene from the pBCHygM vector with the primer pair, 5_CPC1_Mlu (5'-gcgacgcgtccgagatgcccgctgc-3') and 3_Ttrpc_Mlu (5'-gtcacgcgtagagatcctctagcta-3'), followed by digestion with *MluI*. The pBCHygM vector is a derivative of pBC-hygro [76], which has lost the *MluI* site in the trpC terminator.

The *N. crassa* strategy has been previously described [75] and included two modifications aimed at minimizing errors in the 5' and 3' flanking regions and the *hph* cassette obtained from pMOcosX [77]. First, the high-fidelity *Pfu* polymerase (Promega, Madison, USA) was used instead of the low-fidelity *LA Taq* polymerase (TaKaRa BIO Inc, Shiga, Japan) and the cycle conditions were adapted to minimize amplification errors. If amplifications with *Pfu* on 20 cycles failed, 10 additional cycles were performed, and if 30 cycles were not successful, the *LA Taq* polymerase was used according to the manufacturer's instructions. The second modification involved the addition of an *AscI* site at each end of the deletion cassette. These sites were used to linearize and separate the deletion cassette from the vector prepared by amplification in *Escherichia coli* prior to the transformation of *P. anserina*.

Linearized DNA containing the 3' and 5' untranslated regions of the gene to be deleted and the *hph* selection gene [77], was used for the transformation of the *P. anserina* $\Delta PaKU70$ strain as described in <http://podospora.igmors.u-psud.fr/methods.php#>

transfo. In each mutant strain, accurate gene replacement was checked by Southern blot analysis and one transformant containing the correct deletion was crossed with a wild-type strain to generate both *mat+* and *mat-* mutant strains lacking the $\Delta PaKU70$ deletion.

Phenotype analyses

For each deletion, both *mat+* and *mat-* mutant strains were tested for fertility. The number of male cells active in the fertilization of a compatible wild-type tester strain (used as a female) was determined as previously described [13]. Crosses homozygous for each deletion were performed to search for possible defects in the development of the fertilized female organs. The capacity of *mat+/mat-* heterokaryotic deletion strains to grow and reproduce on different carbon sources was tested as described in [78]. Growth and fertility of *mat+*, *mat-* and *mat+/mat-* deletion strains were tested on minimal media containing inhibitors that act on different cell compartments: nikomycin (0.1 mg/ml), which acts on the cell wall by inhibiting chitin synthase; fludioxonil (0.01 mg/ml), which acts on osmoregulation; sodium dodecylsulfate (0.03%) which acts on the cell membrane; and caffeine which induces pleiotropic effects. The concentration of each inhibitor caused an approximately 50% decrease in the wild-type growth rate.

Calculation of interspecies correlation

FUNGIpath [74] was used to identify systemically orthologous groups of genes for pairwise comparisons between *P. anserina* FPR1 and *G. moniliformis* MAT1-2-1 target genes (Table S10), *P. anserina* FPR1 and *S. macrospora* SmtA-1 target genes (Table S11), and *P. anserina* FMR1 and *S. macrospora* SmtA-1 target genes (Table S12). The cut-off for selecting differentially transcribed genes in *P. anserina* comparisons was lowered to ≥ 1.5 -fold change ($FC \geq 1.5$) and a ≤ 1.5 -fold change ($FC \leq -1.5$) with a p-value of < 0.005 . P-values for the overlap of gene sets between *P. anserina* and *G. moniliformis* or *S. macrospora* (Table 5) were determined from a 2x2 contingency table using a Chi-square test with Yates correction run on <http://www.graphpad.com/quickcalcs/contingency1.cfm> (GraphPad Software Inc, La Jolla, USA). The contingency table compared the number of orthologs (group 1, outcome 1) identified in the total number of target genes (group 1, outcome 2) of a given mating-type transcription factor (*G. moniliformis* MAT1-2-1, SmtA-1 or SmtA-1) with the number of targets genes (group 2, outcome 1) identified for the corresponding transcription factor in the 10,556 genes of *P. anserina* (group 2, outcome 2). Calculation of interspecies correlation (Table 5) were performed as described in [54] from log-transformed FC values. Pearson correlations, p-values calculations from T-tests, Spearman rank correlations and Kendall's Tau were performed on <http://www.wessa.net/stat.wasp> [Wessa, P. (2011), Free Statistics Software, Office for Research Development and Education, version 1.1.23-r6].

Supporting Information

Table S1 Time-course RT-qPCR analysis during vegetative growth. Cq values for mating-type gene and reference gene cDNAs are reported during vegetative growth (48 h, 72 h, 96 h and 120 h) and crossing (48 h). N/A: Cq value not available because out-of-scale. (XLS)

Table S2 Whole genome transcriptomic comparisons of wild-type and mutant strains. Fold change (FC), adjusted p-value (p-value) and the binary logarithm of the arithmetic mean of compared intensities (A) are indicated for the 10,556 *P. anserina*

genes and for each comparison (*mat+* vs *mat-*, *fpr1⁻* vs *fmr1⁻*, *mat+* vs *fpr1⁻* and *mat-* vs *fmr1⁻*). Genes with FC values above cut-off and significant p-values are marked by “1” in the “selection” column for each comparison. FC cut-offs: $FC \leq -2$ or $FC \geq 2$. p-value cut-off: $P < 0.005$.
(XLS)

Table S3 Main features of genes accumulating transcripts in the *mat+* strain.
(DOC)

Table S4 Main features of genes accumulating transcripts in the *mat-* strains.
(DOC)

Table S5 RT-qPCR validation of microarray data. Cq values are in Table S6.
(DOC)

Table S6 RT-qPCR validation of microarray. Cq values for mating-type target gene and reference gene cDNAs are reported for *mat+* and *mat-* strains in 96 h-old mycelium.
(XLS)

Table S7 Partitioning of the *mat+* vs *mat-* differentially transcribed genes according to their FC value in wild-type to mutant comparisons. Genes with positive FCs ($FC \geq 2$, $P < 0.005$) in the *mat+* vs *fpr1⁻* and *mat-* vs *fmr1⁻* comparisons were considered as being activated (A) by FPR1 and FMR1, respectively. Genes with negative FCs ($FC \leq -2$, $P < 0.005$) in the *mat+* vs *fpr1⁻* and *mat-* vs *fmr1⁻* comparisons were considered as being repressed (R) by FPR1 and FMR1, respectively.
(XLS)

References

- Galgoczy DJ, Cassidy-Stone A, Llinas M, O'Rourke SM, Herskowitz I, et al. (2004) Genomic dissection of the cell-type-specification circuit in *Saccharomyces cerevisiae*. *Proc Natl Acad Sci U S A* 101: 18069–18074.
- Mata J, Bahler J (2006) Global roles of Ste11p, cell type, and pheromone in the control of gene expression during early sexual differentiation in fission yeast. *Proc Natl Acad Sci U S A* 103: 15517–15522.
- Turgeon BG, Debuchy R (2007) *Cochliobolus* and *Podospora*: mechanism of sex determination and the evolution of reproductive lifestyle. In: Heitman J, Kronstad JW, Taylor JW, Casselton LA, eds. Sex in Fungi, molecular determination and evolutionary implications. Washington, DC: ASM Press. pp 93–121.
- Ferreira AV, An Z, Metzberg RL, Glass NL (1998) Characterization of *mat A-2*, *mat A-3* and *AmatA* mating-type mutants of *Neurospora crassa*. *Genetics* 148: 1069–1079.
- Fleissner A, Diamond S, Glass NL (2009) The *Saccharomyces cerevisiae* *PRM1* homolog in *Neurospora crassa* is involved in vegetative and sexual cell fusion events but also has postfertilization functions. *Genetics* 181: 497–510.
- Wik L, Karlsson M, Johansson H (2008) The evolutionary trajectory of the mating-type (*mat*) genes in *Neurospora* relates to reproductive behavior of taxa. *BMC Evol Biol* 8: 109.
- Klix V, Nowrousian M, Ringelberg C, Loros JJ, Dunlap JC, et al. (2010) Functional characterization of *MAT1-1*-specific mating-type genes in the homothallic ascomycete *Sordaria macrospora* provides new insights into essential and nonessential sexual regulators. *Eukaryot Cell* 9: 894–905.
- Lee SH, Kim YK, Yun SH, Lee YW (2008) Identification of differentially expressed proteins in a *mat1-2*-deleted strain of *Gibberella zeae*, using a comparative proteomics analysis. *Curr Genet* 53: 175–184.
- Qi W, Kwon C, Trail F (2006) Microarray analysis of transcript accumulation during perithecial development in the filamentous fungus *Gibberella zeae* (anamorph *Fusarium graminearum*). *Mol Genet Genomics* 276: 87–100.
- Keszthelyi A, Jeney A, Kerenyi Z, Mendes O, Waalwijk C, et al. (2007) Tagging target genes of the *MAT1-2-1* transcription factor in *Fusarium verticillioides* (*Gibberella fujikuroi* MP-A). *Antonie Van Leeuwenhoek* 91: 373–391.
- Paoletti M, Seymour FA, Alcocer MJ, Kaur N, Calvo AM, et al. (2007) Mating type and the genetic basis of self-fertility in the model fungus *Aspergillus nidulans*. *Curr Biol* 17: 1384–1389.
- Rubini A, Bellioli B, Riccioni C, Tisserant E, Arcioni S, et al. (2010) Isolation and characterization of *MAT* genes in the symbiotic ascomycete *Tuber melanosporum*. *New Phytol* 189: 710–722.
- Coppin E, de Renty C, Debuchy R (2005) The function of the coding sequences for the putative pheromone precursors in *Podospora anserina* is restricted to fertilization. *Eukaryot Cell* 4: 407–420.
- Schmoll M, Seibel C, Tisch D, Dorrer M, Kubicek CP (2010) A novel class of peptide pheromone precursors in ascomycetous fungi. *Mol Microbiol* 77: 1483–1501.
- Shen WC, Bobrowicz P, Ebbole DJ (1999) Isolation of pheromone precursor genes of *Magnaporthe grisea*. *Fungal Genet Biol* 27: 253–263.
- Zhang L, Baasiri RA, Van Alfen NK (1998) Viral repression of fungal pheromone precursor gene expression. *Mol Cell Biol* 18: 953–959.
- Bobrowicz P, Pawlak R, Correa A, Bell-Pedersen D, Ebbole DJ (2002) The *Neurospora crassa* pheromone precursor genes are regulated by the mating type locus and the circadian clock. *Mol Microbiol* 45: 795–804.
- Kim H, Metzberg RL, Nelson MA (2002) Multiple functions of *mfa-1*, a putative pheromone precursor gene of *Neurospora crassa*. *Eukaryot Cell* 1: 987–999.
- Zickler D (1973) La méiose et les mitoses au cours du cycle de quelques Ascomycètes. Orsay: Univ Paris-Sud. 86 p.
- Metzberg RL, Glass NL (1990) Mating type and mating strategies in *Neurospora*. *Bioessays* 12: 53–59.
- Martin T, Lu SW, van Tilbeurgh H, Ripoll DR, Dixelius C, et al. (2010) Tracing the origin of the fungal alpha1 domain places its ancestor in the HMG-box superfamily: implication for fungal mating-type evolution. *PLoS One* 5: e15199.
- Astell CR, Ahlstrom-Jonasson L, Smith M, Tatchell K, Nasmyth KA, et al. (1981) The sequence of the DNAs coding for the mating-type loci of *Saccharomyces cerevisiae*. *Cell* 27: 15–23.
- Turgeon BG, Yoder OC (2000) Proposed nomenclature for mating type genes of filamentous ascomycetes. *Fungal Genet Biol* 31: 1–5.
- Debuchy R, Coppin E (1992) The mating types of *Podospora anserina*: functional analysis and sequence of the fertilization domains. *Mol Gen Genet* 233: 113–121.
- Glass NL, Grotelueschen J, Metzberg RL (1990) *Neurospora crassa* A mating-type region. *Proc Natl Acad Sci U S A* 87: 4912–4916.
- Staben C, Yanofsky C (1990) *Neurospora crassa* a mating-type region. *Proc Natl Acad Sci U S A* 87: 4917–4921.
- Wirsal S, Turgeon BG, Yoder OC (1996) Deletion of the *Cochliobolus heterostrophus* mating-type (*MAT*) locus promotes the function of *MAT* transgenes. *Curr Genet* 29: 241–249.

Table S8 FunCat classification of genes up-regulated in the *mat+* vs *mat-* strain comparison.
(XLS)

Table S9 FunCat classification of genes up-regulated in the *mat-* vs *mat+* strain comparison.
(XLS)

Table S10 Orthologous pairs of *P. anserina* FPR1 and *G. moniliformis* MAT1-2-1 target genes.
(XLS)

Table S11 Orthologous pairs of *P. anserina* FPR1 and *S. macrospora* SMTa-1 target genes.
(XLS)

Table S12 Orthologous pairs in *P. anserina* FMR1 and *S. macrospora* SMTA-1 target genes.
(XLS)

Table S13 Oligonucleotide primers used for RT-qPCR.
(DOC)

Acknowledgments

RD is greatly indebted to Véronique Berteaux-Lecellier for her advice, constant help and support during this study. We thank Philippe Silar for communicating unpublished results, and Stefanie Pöggeler for communicating processed files with the *S. macrospora* mating-type target genes.

Author Contributions

Conceived and designed the experiments: RD FB SI HD. Performed the experiments: FB JAB EC PG. Analyzed the data: FB SI JAB. Wrote the paper: RD.

28. Arnaise S, Zickler D, Le Bilot S, Poisier C, Debuchy R (2001) Mutations in mating-type genes of the heterothallic fungus *Podospora anserina* lead to self-fertility. *Genetics* 159: 545–556.
29. Debuchy R, Arnaise S, Lecellier G (1993) The *mat-1* allele of *Podospora anserina* contains three regulatory genes required for the development of fertilized female organs. *Mol Gen Genet* 241: 667–673.
30. Saupé SJ, Stenberg L, Shiu KT, Griffiths AJ, Glass L, N (1996) The molecular nature of mutations in the *mtA-1* gene of the *Neurospora crassa* A idiomorph and their relation to mating-type function. *Mol Gen Genet* 250: 115–122.
31. Wirsel S, Horwitz B, Yamaguchi K, Yoder OC, Turgeon BG (1998) Single mating type-specific genes and their 3' UTRs control mating and fertility in *Cochliobolus heterostrophus*. *Mol Gen Genet* 259: 272–281.
32. Zickler D, Arnaise S, Coppin E, Debuchy R, Picard M (1995) Altered mating-type identity in the fungus *Podospora anserina* leads to selfish nuclei, uniparental progeny, and haploid meiosis. *Genetics* 140: 493–503.
33. Debuchy R, Berteaux-Lecellier V, Silar P (2010) Mating systems and sexual morphogenesis in Ascomycetes. In: Borkovich KA, Ebbel DJ, eds. *Cellular and Molecular Biology of Filamentous Fungi*. Washington, DC: ASM Press. pp 501–535.
34. Kanematsu S, Adachi Y, Ito T (2007) Mating-type loci of heterothallic *Diaporthe sphaerica*: homologous genes are present in opposite mating-types. *Curr Genet* 52: 11–22.
35. Ferreira AV, Saupé S, Glass NL (1996) Transcriptional analysis of the *mtA* idiomorph of *Neurospora crassa* identifies two genes in addition to *mtA-1*. *Mol Gen Genet* 250: 767–774.
36. Glass NL, Lee L (1992) Isolation of *Neurospora crassa* A mating type mutants by repeat induced point (RIP) mutation. *Genetics* 132: 125–133.
37. Pöggeler S, Nowrousian M, Ringelberg C, Loros JJ, Dunlap JC, et al. (2006) Microarray and real-time PCR analyses reveal mating type-dependent gene expression in a homothallic fungus. *Mol Genet Genomics* 275: 492–503.
38. Coppin E, Debuchy R (2000) Co-expression of the mating-type genes involved in internuclear recognition is lethal in *Podospora anserina*. *Genetics* 155: 657–669.
39. Bidard F, Imbeaud S, Raymond N, Lespinet O, Silar P, et al. (2010) A general framework for optimization of probes for gene expression microarray and its application to the fungus *Podospora anserina*. *BMC Res Notes* 3: 171.
40. Kuchler K, Sterne RE, Thorner J (1989) *Saccharomyces cerevisiae* *STE6* gene product: a novel pathway for protein export in eukaryotic cells. *EMBO J* 8: 3973–3984.
41. Kim H, Borkovich KA (2004) A pheromone receptor gene, *pre-1*, is essential for mating type-specific directional growth and fusion of trichogynes and female fertility in *Neurospora crassa*. *Mol Microbiol* 52: 1781–1798.
42. Wurmbach E (2009) Validation of array data. In: Logan J, Edwards K, Saunders N, eds. *Real-time PCR current technology and applications*. Norfolk, UK: Caister Academic Press. pp 137–148.
43. Wurmbach E (2010) Microarrays and quantitative real-time reverse transcriptase-polymerase chain reaction. In: Bustin SA, ed. *The PCR revolution: basic technologies and applications*. Cambridge, New York, Melbourne, Madrid, Cape Town, Singapore, Sao Paulo, Delhi, Tokyo: Cambridge University Press. pp 262–275.
44. El-Khoury R, Sellem CH, Coppin E, Boivin A, Maas MF, et al. (2008) Gene deletion and allelic replacement in the filamentous fungus *Podospora anserina*. *Curr Genet* 53: 249–258.
45. Altschul SF, Madden TL, Schaffer AA, Zhang J, Zhang Z, et al. (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res* 25: 3389–3402.
46. Ruepp A, Zollner A, Maier D, Albermann K, Hani J, et al. (2004) The FunCat, a functional annotation scheme for systematic classification of proteins from whole genomes. *Nucleic Acids Res* 32: 5539–5545.
47. Cummings DJ, McNally KL, Domenico JM, Matsuura ET (1990) The complete DNA sequence of the mitochondrial genome of *Podospora anserina*. *Curr Genet* 17: 375–402.
48. Davey J, Davis K, Hughes M, Ladds G, Powney D (1998) The processing of yeast pheromones. *Semin Cell Dev Biol* 9: 19–30.
49. He B, Chen P, Chen SY, Vancura KL, Michaelis S, et al. (1991) *RAM2*, an essential gene of yeast, and *RAM1* encode the two polypeptide components of the farnesyltransferase that prenylates a-factor and Ras proteins. *Proc Natl Acad Sci U S A* 88: 11373–11377.
50. Lorin S, Dufour E, Boulay J, Begel O, Marsy S, et al. (2001) Overexpression of the alternative oxidase restores senescence and fertility in a long-lived respiration-deficient mutant of *Podospora anserina*. *Mol Microbiol* 42: 1259–1267.
51. Vallim MA, Fernandes L, Alspaugh JA (2004) The *RAM1* gene encoding a protein-farnesyltransferase beta-subunit homologue is essential in *Cryptococcus neoformans*. *Microbiology* 150: 1925–1935.
52. Pöggeler S, Risch S, Kück U, Osiewacz HD (1997) Mating-type genes from the homothallic fungus *Sordaria macrospora* are functionally expressed in a heterothallic ascomycete. *Genetics* 147: 567–580.
53. Xiang Q, Rasmussen C, Glass NL (2002) The *ham-2* locus, encoding a putative transmembrane protein, is required for hyphal fusion in *Neurospora crassa*. *Genetics* 160: 169–180.
54. McCarroll SA, Murphy CT, Zou S, Pletcher SD, Chin CS, et al. (2004) Comparing genomic expression patterns across species identifies shared transcriptional profile in aging. *Nat Genet* 36: 197–204.
55. Jamet-Vierny C, Debuchy R, Prigent M, Silar P (2007) *IDCI*, a pezizomycotina-specific gene that belongs to the PaMpk1 MAP kinase transduction cascade of the filamentous fungus *Podospora anserina*. *Fungal Genet Biol* 44: 1219–1230.
56. Dyer P, Ingram DS, Johnstone K (1993) Evidence for the involvement of linoleic acid and other endogenous lipid factors in perithecial development of *Nectria haematococca* population. *Mycol Res* 97: 485–496.
57. Dyer P, Ingram DS, Johnstone K (1992) Control of sexual morphogenesis in the ascomycotina. *Biological Reviews*. pp 421–458.
58. Goodrich-Tanrikulu M, Howe K, Stafford A, Nelson MA (1998) Changes in fatty acid composition of *Neurospora crassa* accompany sexual development and ascospore germination. *Microbiology* 144(Pt 7): 1713–1720.
59. Herskowitz I (1989) A regulatory hierarchy for cell specialization in yeast. *Nature* 342: 749–757.
60. Samils N, Johannesson H (2011) Proportion of mating type specific nuclei and their associated gene expression in *Neurospora tetrasperma*. *Fungal Genetics Reports* 58-Supplement: abstract n° 631, 269 p.
61. Rizet G, Engelmann C (1949) Contribution à l'étude génétique d'un ascomycète tetrasperé: *Podospora anserina*. *Rhem Kv Cytol Biol Veg* 11: 201–304.
62. Rizet G (1952) Les phénomènes de barrage chez *Podospora anserina*. I. Analyse génétique des barrages entre souches S and s. *Rev Cytol Biol Veg* 13: 51–92.
63. Espagne E, Lespinet O, Malagnac F, Da Silva C, Jaillon O, et al. (2008) The genome sequence of the model ascomycete fungus *Podospora anserina*. *Genome Biol* 9: R77.
64. Imbeaud S, Graudens E, Boulanger V, Barlet X, Zaborski P, et al. (2005) Towards standardization of RNA quality assessment using user-independent classifiers of microcapillary electrophoresis traces. *Nucleic Acids Res* 33: e56.
65. Vandesompele J, De Preter K, Pattyn F, Poppe B, Van Roy N, et al. (2002) Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. *Genome Biol* 3: RESEARCH0034.
66. Bustin SA, Benes V, Garson JA, Hellemans J, Huggett J, et al. (2009) The MIQE guidelines: minimum information for publication of quantitative real-time PCR experiments. *Clin Chem* 55: 611–622.
67. Ridder R, Osiewacz HD (1992) Sequence analysis of the gene coding for glyceraldehyde-3-phosphate dehydrogenase (gpd) of *Podospora anserina*: use of homologous regulatory sequences to improve transformation efficiency. *Curr Genet* 21: 207–213.
68. Dequard-Chablat M, Sellem CH (1994) The S12 ribosomal protein of *Podospora anserina* belongs to the S19 bacterial family and controls the mitochondrial genome integrity through cytoplasmic translation. *J Biol Chem* 269: 14951–14956.
69. Pfaffl MW, Horgan GW, Dempfle L (2002) Relative expression software tool (REST) for group-wise comparison and statistical analysis of relative expression results in real-time PCR. *Nucleic Acids Res* 30: e36.
70. Marisa L, Ichante JL, Raymond N, Aggerbeck L, Delacroix H, et al. (2007) MANGO: an interactive R-based tool for two-colour microarray analysis. *Bioinformatics* 23: 2339–2341.
71. Benjamini Y, Hochberg Y (1995) Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J Roy Stat Soc Ser B (Statistical Methodology)* 57: 289–300.
72. Edgar R, Domrachev M, Lash AE (2002) Gene Expression Omnibus: NCBI gene expression and hybridization array data repository. *Nucleic Acids Res* 30: 207–210.
73. Huang X, Madan A (1999) CAP3: A DNA sequence assembly program. *Genome Res* 9: 868–877.
74. Grossetete S, Labedan B, Lespinet O (2010) FUNGIpath: a tool to assess fungal metabolic pathways predicted by orthology. *BMC Genomics* 11: 81.
75. Colot HV, Park G, Turner GE, Ringelberg C, Crew CM, et al. (2006) A high-throughput gene knockout procedure for *Neurospora* reveals functions for multiple transcription factors. *Proc Natl Acad Sci U S A* 103: 10352–10357.
76. Silar P (1995) Two new easy to use vectors for transformations. *Fungal Genet Newslett* 42: 73.
77. Orbach MJ (1994) A cosmid with a HyR marker for fungal library construction and screening. *Gene* 150: 159–162.
78. Malagnac F, Bidard F, Lalucque H, Brun S, Lambou K, et al. (2008) Convergent evolution of morphogenetic processes in fungi: Role of tetraspanins and NADPH oxidases 2 in plant pathogens and saprobes. *Commun Integr Biol* 1: 180–181.

Références bibliographiques

- Abdel-Fattah YR, El Enshasy H, Anwar M, Omar H, Abolmagd E et al. (2007) Application of factorial experimental designs for optimization of cyclosporin. A production by *Tolypocladium inflatum* in submerged culture. *Journal of microbiology and biotechnology* 17(12): 1930-1936.
- Agresti A, Bianchi ME (2003) HMGB proteins and gene expression. *Current opinion in genetics & development* 13(2): 170-178.
- Ammerer G (1990) Identification, purification, and cloning of a polypeptide (PRTF/GRM) that binds to mating-specific promoter elements in yeast. *Genes & development* 4(2): 299-312.
- Andersson U, Wang H, Palmblad K, Aveberger AC, Bloom O et al. (2000) High mobility group 1 protein (HMG-1) stimulates proinflammatory cytokine synthesis in human monocytes. *The Journal of experimental medicine* 192(4): 565-570.
- Arnaise S, Debuchy R, Picard M (1997) What is a bona fide mating-type gene? Internuclear complementation of mat mutants in *Podospora anserina*. *Mol Gen Genet* 256(2): 169-178.
- Arnaise S, Zickler D, Le Bilot S, Poisier C, Debuchy R (2001) Mutations in mating-type genes of the heterothallic fungus *Podospora anserina* lead to self-fertility. *Genetics* 159(2): 545-556.
- Arnaise S, Zickler D, Bourdais A, Dequard-Chablat M, Debuchy R (2008) Mutations in mating-type genes greatly decrease repeat-induced point mutation process in the fungus *Podospora anserina*. *Fungal Genet Biol* 45(3): 207-220.
- Aronstein KA, Murray KD, de Leon JH, Qin X, Weinstock GM (2007) High mobility group (HMG-box) genes in the honeybee fungal pathogen *Ascospaera apis*. *Mycologia* 99(4): 553-561.
- Balaeff A, Churchill ME, Schulten K (1998) Structure prediction of a complex between the chromosomal protein HMG-D and DNA. *Proteins* 30(2): 113-135.
- Balasubramanian B, Lowry CV, Zitomer RS (1993) The Rox1 repressor of the *Saccharomyces cerevisiae* hypoxic genes is a specific DNA-binding protein with a high-mobility-group motif. *Molecular and cellular biology* 13(10): 6071-6078.
- Barr MM, Tu H, Van Aelst L, Wigler M (1996) Identification of Ste4 as a potential regulator of Byr2 in the sexual response pathway of *Schizosaccharomyces pombe*. *Molecular and cellular biology* 16(10): 5597-5603.
- Barreau C, Iskandar M, Turcq B, Javerzat JP (1998) Use of a linear plasmid containing telomeres as an efficient vector for direct cloning in the filamentous fungus *Podospora anserina*. *Fungal Genet Biol* 25(1): 22-30.
- Berdy J (2005) Bioactive microbial metabolites. *The Journal of antibiotics* 58(1): 1-26.
- Bianchi ME, Agresti A (2005) HMG proteins: dynamic players in gene regulation and differentiation. *Current opinion in genetics & development* 15(5): 496-506.
- Bianchi ME, Beltrame M, Paonessa G (1989) Specific recognition of cruciform DNA by nuclear protein HMG1. *Science (New York, NY)* 243(4894 Pt 1): 1056-1059.
- Bidard F, Ait Benkhali J, Coppin E, Imbeaud S, Grognet P et al. (2011) Genome-Wide Gene Expression Profiling of Fertilization Competent Mycelium in Opposite Mating Types in the Heterothallic Fungus *Podospora anserina*. *PloS one* 6(6): e21476.
- Bidard F, Imbeaud S, Reymond N, Lespinet O, Silar P et al. (2010) A general framework for optimization of probes for gene expression microarray and its application to the fungus *Podospora anserina*. *BMC research notes* 3: 171-.
- Blakeslee AF (1904) Sexual reproduction in the Mucorineae. 205–319 p.
- Bolker M, Kahmann R (1993) Sexual pheromones and mating responses in fungi. *The Plant cell* 5(10): 1461-1469.

- Bonaldi T, Talamo F, Scaffidi P, Ferrera D, Porto A et al. (2003) Monocytic cells hyperacetylate chromatin protein HMGB1 to redirect it towards secretion. *The EMBO journal* 22(20): 5551-5560.
- Bose I, Reese AJ, Ory JJ, Janbon G, Doering TL (2003) A yeast under cover: the capsule of *Cryptococcus neoformans*. *Eukaryotic cell* 2(4): 655-663.
- Brefort T, Muller P, Kahmann R (2005) The high-mobility-group domain transcription factor Rop1 is a direct regulator of *prf1* in *Ustilago maydis*. *Eukaryotic cell* 4(2): 379-391.
- Brun S, Malagnac F, Bidard F, Lalucque H, Silar P (2009) Functions and regulation of the Nox family in the filamentous fungus *Podospora anserina*: a new role in cellulose degradation. *Molecular microbiology* 74(2): 480-496.
- Bushel PR, Heard NA, Gutman R, Liu L, Peddada SD et al. (2009) Dissecting the fission yeast regulatory network reveals phase-specific control elements of its cell cycle. *BMC systems biology* 3: 93.
- Bustin M (1999) Regulation of DNA-dependent activities by the functional motifs of the high-mobility-group chromosomal proteins. *Molecular and cellular biology* 19(8): 5237-5246.
- Bustin M (2001) Revised nomenclature for high mobility group (HMG) chromosomal proteins. *Trends in biochemical sciences* 26(3): 152-153.
- Bustin M, Reeves R (1996) High-mobility-group chromosomal proteins: architectural components that facilitate chromatin function. *Progress in nucleic acid research and molecular biology* 54: 35-100.
- Butler G, Kenny C, Fagan A, Kurischko C, Gaillardin C et al. (2004) Evolution of the MAT locus and its Ho endonuclease in yeast species. *Proceedings of the National Academy of Sciences of the United States of America* 101(6): 1632-1637.
- Casselton LA (2002) Mate recognition in fungi. *Heredity* 88(2): 142-147.
- Chen P, Sapperstein SK, Choi JD, Michaelis S (1997) Biogenesis of the *Saccharomyces cerevisiae* mating pheromone α -factor. *The Journal of cell biology* 136(2): 251-269.
- Cohen DR, Sinclair AH, McGovern JD (1994) SRY protein enhances transcription of Fos-related antigen 1 promoter constructs. *Proceedings of the National Academy of Sciences of the United States of America* 91(10): 4372-4376.
- Colot HV, Park G, Turner GE, Ringelberg C, Crew CM et al. (2006) A high-throughput gene knockout procedure for *Neurospora* reveals functions for multiple transcription factors. *Proceedings of the National Academy of Sciences of the United States of America* 103(27): 10352-10357.
- Coppin E, Debuchy R (2000) Co-expression of the mating-type genes involved in internuclear recognition is lethal in *Podospora anserina*. *Genetics* 155(2): 657-669.
- Coppin E, de Renty C, Debuchy R (2005) The function of the coding sequences for the putative pheromone precursors in *Podospora anserina* is restricted to fertilization. *Eukaryotic cell* 4(2): 407-420.
- Coppin E, Arnais S, Contamine V, Picard M (1993) Deletion of the mating-type sequences in *Podospora anserina* abolishes mating without affecting vegetative functions and sexual differentiation. *Mol Gen Genet* 241(3-4): 409-414.
- Coppin E, Debuchy R, Arnais S, Picard M (1997) Mating types and sexual development in filamentous ascomycetes. *Microbiol Mol Biol Rev* 61(4): 411-428.
- Costigan C, Kolodrubetz D, Snyder M (1994) NHP6A and NHP6B, which encode HMG1-like proteins, are candidates for downstream components of the yeast SLT2 mitogen-activated protein kinase pathway. *Molecular and cellular biology* 14(4): 2391-2403.
- Da G, Lenkart J, Zhao K, Shiekhattar R, Cairns BR et al. (2006) Structure and function of the SWIRM domain, a conserved protein module found in chromatin regulatory complexes. *Proceedings of the National Academy of Sciences of the United States of America* 103(7): 2057-2062.

- Davey J (1992) Mating pheromones of the fission yeast *Schizosaccharomyces pombe*: purification and structural characterization of M-factor and isolation and analysis of two genes encoding the pheromone. *The EMBO journal* 11(3): 951-960.
- Davey J, Davis K, Hughes M, Ladds G, Powner D (1998) The processing of yeast pheromones. *Seminars in cell & developmental biology* 9(1): 19-30.
- Davidson MK, Shandilya HK, Hirota K, Ohta K, Wahls WP (2004) Atf1-Pcr1-M26 complex links stress-activated MAPK and cAMP-dependent protein kinase pathways via chromatin remodeling of *cgs2+*. *The Journal of biological chemistry* 279(49): 50857-50863.
- Davis L, Smith GR (2001) Meiotic recombination and chromosome segregation in *Schizosaccharomyces pombe*. *Proceedings of the National Academy of Sciences of the United States of America* 98(15): 8395-8402.
- Debuchy R (1999) Internuclear recognition: A possible connection between euascomycetes and homobasidiomycetes. *Fungal Genet Biol* 27(2-3): 218-223.
- Debuchy R, Coppin E (1992) The mating types of *Podospira anserina*: functional analysis and sequence of the fertilization domains. *Mol Gen Genet* 233(1-2): 113-121.
- Debuchy R, Arnaise S, Lecellier G (1993) The *mat-* allele of *Podospira anserina* contains three regulatory genes required for the development of fertilized female organs. *Mol Gen Genet* 241(5-6): 667-673.
- Debuchy R, Berteaux-Lecellier V, Silar P (2010) Mating systems and sexual morphogenesis in Ascomycetes. In: Borkovich KA, Ebbole DJ, editors. *Cellular and Molecular Biology of Filamentous Fungi*. Washington, DC: ASM Press. pp. 501-535.
- Dequard-Chablat M, Alland C (2002) Two copies of *mthmg1*, encoding a novel mitochondrial HMG-like protein, delay accumulation of mitochondrial DNA deletions in *Podospira anserina*. *Eukaryotic cell* 1(4): 503-513.
- Dietz K, Gail M, Krickeberg K, Samet J, Tsiatis A, editors *Design and Analysis of DNA Microarray Investigations* New York Berlin Heidelberg: Springer-Verlag.
- Diffley JF, Stillman B (1991) A close relative of the nuclear, chromosomal high-mobility group protein HMG1 in yeast mitochondria. *Proceedings of the National Academy of Sciences of the United States of America* 88(17): 7864-7868.
- Dodge BO (1920) *The Life History of Ascobolus magnificus: Origin of the Ascocarp from Two Strains*. Mycologia: Mycological Society of America. pp. 115-134.
- Dumitriu IE, Baruah P, Manfredi AA, Bianchi ME, Rovere-Querini P (2005) HMGB1: guiding immunity from within. *Trends in immunology* 26(7): 381-387.
- Edgerton (1914) Plus and minus strain in an ascomycete. *Science (New York, NY)* 35: 151.
- El-Khoury R, Sellem CH, Coppin E, Boivin A, Maas MF et al. (2008) Gene deletion and allelic replacement in the filamentous fungus *Podospira anserina*. *Current genetics* 53(4): 249-258.
- El Enshasy H, Abdel Fattah Y, Atta A, Anwar M, Omar H et al. (2008) Kinetics of cell growth and cyclosporin A production by *Tolypocladium inflatum* when scaling up from shake flask to bioreactor. *Journal of microbiology and biotechnology* 18(1): 128-134.
- Erondu NE, Donelson JE (1992) Differential expression of two mRNAs from a single gene encoding an HMG1-like DNA binding protein of African trypanosomes. *Molecular and biochemical parasitology* 51(1): 111-118.
- Errede B, Ammerer G (1989) STE12, a protein involved in cell-type-specific transcription and signal transduction in yeast, is part of protein-DNA complexes. *Genes & development* 3(9): 1349-1361.
- Espagne E, Lespinet O, Malagnac F, Da Silva C, Jaillon O et al. (2008) The genome sequence of the model ascomycete fungus *Podospira anserina*. *Genome biology* 9(5): R77.
- Fan HY, Cheng KK, Klein HL (1996) Mutations in the RNA Polymerase II Transcription Machinery Suppress the Hyperrecombination Mutant Hpr1 Δ of *Saccharomyces Cerevisiae*. *Genetics*: 749-759.

- Ferrari S, Harley VR, Pontiggia A, Goodfellow PN, Lovell-Badge R et al. (1992) SRY, like HMG1, recognizes sharp angles in DNA. *The EMBO journal* 11(12): 4497-4506.
- Fornieris F, Battaglioli E, Mattevi A, Binda C (2009) New roles of flavoproteins in molecular cell biology: histone demethylase LSD1 and chromatin. *The FEBS journal* 276(16): 4304-4312.
- Forsburg SL (2005) The yeasts *Saccharomyces cerevisiae* and *Schizosaccharomyces pombe*: models for cell biology research. *Gravit Space Biol Bull* 18(2): 3-9.
- Fu C, Iyer P, Herkal A, Abdullah J, Stout A et al. (2011) Identification and characterization of genes required for cell-to-cell fusion in *Neurospora crassa*. *Eukaryotic cell* 10(8): 1100-1109.
- Fukui Y, Kozasa T, Kaziro Y, Takeda T, Yamamoto M (1986) Role of a ras homolog in the life cycle of *Schizosaccharomyces pombe*. *Cell* 44(2): 329-336.
- Galagan JE, Calvo SE, Cuomo C, Ma LJ, Wortman JR et al. (2005) Sequencing of *Aspergillus nidulans* and comparative analysis with *A. fumigatus* and *A. oryzae*. *Nature* 438(7071): 1105-1115.
- Galagan JE, Calvo SE, Borkovich KA, Selker EU, Read ND et al. (2003) The genome sequence of the filamentous fungus *Neurospora crassa*. *Nature* 422(6934): 859-868.
- Galgoczy DJ, Cassidy-Stone A, Llinas M, O'Rourke SM, Herskowitz I et al. (2004) Genomic dissection of the cell-type-specification circuit in *Saccharomyces cerevisiae*. *Proceedings of the National Academy of Sciences of the United States of America* 101(52): 18069-18074.
- Giese K, Amsterdam A, Grosschedl R (1991) DNA-binding properties of the HMG domain of the lymphoid-specific transcriptional regulator LEF-1. *Genes & development* 5(12B): 2567-2578.
- Giese K, Cox J, Grosschedl R (1992) The HMG domain of lymphoid enhancer factor 1 bends DNA and facilitates assembly of functional nucleoprotein structures. *Cell* 69(1): 185-195.
- Goffeau A, Barrell BG, Bussey H, Davis RW, Dujon B et al. (1996) Life with 6000 genes. *Science (New York, NY)* 274(5287): 546, 563-547.
- Goodfellow PN, Lovell-Badge R (1993) SRY and sex determination in mammals. *Annual review of genetics* 27: 71-92.
- Goodwin GH, Johns EW (1973) Isolation and characterisation of two calf-thymus chromatin non-histone proteins with high contents of acidic and basic amino acids. *European journal of biochemistry / FEBS* 40(1): 215-219.
- Gorski JJ, Pathak S, Panov K, Kaschiukovic T, Panova T et al. (2007) A novel TBP-associated factor of SL1 functions in RNA polymerase I transcription. *The EMBO journal* 26(6): 1560-1568.
- Grossetete S, Labedan B, Lespinet O (2010) FUNGIpath: a tool to assess fungal metabolic pathways predicted by orthology. *BMC genomics* 11: 81.
- Grummt I (2003) Life on a planet of its own: regulation of RNA polymerase I transcription in the nucleolus. *Genes & development* 17(14): 1691-1702.
- Hall MN, Linder (1993) *The Early Days of Yeast Genetics*. 477 p.
- Hanahan D (1983) Studies on transformation of *Escherichia coli* with plasmids. *Journal of molecular biology* 166(4): 557-580.
- Haqq CM, King CY, Donahoe PK, Weiss MA (1993) SRY recognizes conserved DNA sites in sex-specific promoters. *Proceedings of the National Academy of Sciences of the United States of America* 90(3): 1097-1101.
- Hartmann HA, Kahmann R, Bolker M (1996) The pheromone response factor coordinates filamentous growth and pathogenicity in *Ustilago maydis*. *The EMBO journal* 15(7): 1632-1641.
- Hawksworth DL, Rossman AY (1997) Where are all the undescribed fungi? *Phytopathology* 87(9): 888-891.

- Hawksworth DL (2001) The magnitude of fungal diversity: the 1.5 million species revisited. *Mycol Res* 105: 1422-1432.
- Herskowitz I (1988) Life cycle of the budding yeast *Saccharomyces cerevisiae*. *Microbiological reviews* 52(4): 536-553.
- Herskowitz I, Rine J, Strathern J (1992) Mating-type Determination and Mating-type Interconversion in *Saccharomyces cerevisiae*. *The Molecular and Cellular Biology of the Yeast* pp. 583-656.
- Idnurm A, Walton FJ, Floyd A, Heitman J (2008) Identification of the sex genes in an early diverged fungus. *Nature* 451(7175): 193-196.
- Imai Y, M., Yamamoto. (1994) The fission yeast mating pheromone P-factor: its molecular structure, gene structure, and ability to induce gene expression and G1 arrest in the mating partner. *Genes & development* 8: 328–338.
- Iyer SV, Ramakrishnan M, Kasbekar DP (2009) *Neurospora crassa* *fmf-1* encodes the homologue of the *Schizosaccharomyces pombe* *Ste11p* regulator of sexual development. *Journal of genetics* 88(1): 33-39.
- James R. Broach, John R. Pringle, Jones EW (1991-1997) *The Molecular and Cellular Biology of the Yeast Saccharomyces: Gene expression*; Press CSHL, editor.
- Jantzen HM, Admon A, Bell SP, Tjian R (1990) Nucleolar transcription factor hUBF contains a DNA-binding motif with homology to HMG proteins. *Nature* 344(6269): 830-836.
- Jayaraman L, Moorthy NC, Murthy KG, Manley JL, Bustin M et al. (1998) High mobility group protein-1 (HMG-1) is a unique activator of p53. *Genes & development* 12(4): 462-472.
- Jeong HT, Ozoe F, Tanaka K, Nakagawa T, Matsuda H et al. (2004) A novel gene, *msa1*, inhibits sexual differentiation in *Schizosaccharomyces pombe*. *Genetics* 167(1): 77-91.
- Jobb G, von Haeseler A, Strimmer K (2004) TREEFINDER: a powerful graphical analysis environment for molecular phylogenetics. *BMC evolutionary biology* 4: 18.
- John A., Bumpus, D. Aust S (1987) Biodegradation of environmental pollutants by the white rot fungus *Phanerochaete chrysosporium*: Involvement of the lignin degrading system. *BioEssays*. pp. 166–170.
- Johnson TE (1979) A *neurospora* mutation that arrests perithecial development as either male or female parent. *Genetics* 92(4): 1107-1120.
- Kanematsu S, Adachi Y, Ito T (2007) Mating-type loci of heterothallic *Diaporthe spp.*: homologous genes are present in opposite mating-types. *Current genetics* 52(1): 11-22.
- Kashimada K, Koopman P (2010) Sry: the master switch in mammalian sex determination. *Development (Cambridge, England)* 137(23): 3921-3930.
- Kato T, Jr., Okazaki K, Murakami H, Stettler S, Fantes PA et al. (1996) Stress signal, mediated by a Hog1-like MAP kinase, controls sexual development in fission yeast. *FEBS letters* 378(3): 207-212.
- Kelly M, Burke J, Smith M, Klar A, Beach D (1988) Four mating-type genes control sexual differentiation in the fission yeast. *The EMBO journal* 7(5): 1537-1547.
- Keszthelyi A, Jeney A, Kerenyi Z, Mendes O, Waalwijk C et al. (2007) Tagging target genes of the MAT1-2-1 transcription factor in *Fusarium verticillioides* (*Gibberella fujikuroi* MP-A). *Antonie Van Leeuwenhoek* 91(4): 373-391.
- Kicka S, Silar P (2004) PaASK1, a mitogen-activated protein kinase kinase kinase that controls cell degeneration and cell differentiation in *Podospora anserina*. *Genetics* 166(3): 1241-1252.
- Kim CA, Bowie JU (2003) SAM domains: uniform structure, diversity of function. *Trends in biochemical sciences* 28(12): 625-628.
- Kim H, Borkovich KA (2004) A pheromone receptor gene, *pre-1*, is essential for mating type-specific directional growth and fusion of trichogynes and female fertility in *Neurospora crassa*. *Molecular microbiology* 52(6): 1781-1798.

- Kim H, Borkovich KA (2006) Pheromones are essential for male fertility and sufficient to direct chemotropic polarized growth of trichogynes during mating in *Neurospora crassa*. *Eukaryotic cell* 5(3): 544-554.
- Kitamura K, Shimoda C (1991) The *Schizosaccharomyces pombe* mam2 gene encodes a putative pheromone receptor which has a significant homology with the *Saccharomyces cerevisiae* Ste2 protein. *The EMBO journal* 10(12): 3743-3751.
- Kitamura K, Katayama S, Dhut S, Sato M, Watanabe Y et al. (2001) Phosphorylation of Mei2 and Ste11 by Pat1 kinase inhibits sexual differentiation via ubiquitin proteolysis and 14-3-3 protein in fission yeast. *Developmental cell* 1(3): 389-399.
- Kjaerulff S, Dooijes D, Clevers H, Nielsen O (1997) Cell differentiation by interaction of two HMG-box proteins: Mat1-Mc activates M cell-specific genes in *S.pombe* by recruiting the ubiquitous transcription factor Ste11 to weak binding sites. *The EMBO journal* 16(13): 4021-4033.
- Kjaerulff S, Lautrup-Larsen I, Truelsen S, Pedersen M, Nielsen O (2005) Constitutive activation of the fission yeast pheromone-responsive pathway induces ectopic meiosis and reveals ste11 as a mitogen-activated protein kinase target. *Molecular and cellular biology* 25(5): 2045-2059.
- Klix V, Nowrousian M, Ringelberg C, Loros JJ, Dunlap JC et al. (2010) Functional characterization of *MAT1-1*-specific mating-type genes in the homothallic ascomycete *Sordaria macrospora* provides new insights into essential and nonessential sexual regulators. *Eukaryotic cell* 9(6): 894-905.
- Kokkola R, Andersson A, Mullins G, Ostberg T, Treutiger CJ et al. (2005) RAGE is the major receptor for the proinflammatory activity of HMGB1 in rodent macrophages. *Scandinavian journal of immunology* 61(1): 1-9.
- Kolodrubetz D, Burgum A (1990) Duplicated NHP6 genes of *Saccharomyces cerevisiae* encode proteins homologous to bovine high mobility group protein 1. *The Journal of biological chemistry* 265(6): 3234-3239.
- Kruppa M, Moir RD, Kolodrubetz D, Willis IM (2001) Nhp6, an HMG1 protein, functions in SNR6 transcription by RNA polymerase III in *S. cerevisiae*. *Molecular cell* 7(2): 309-318.
- Krynetski EY, Krynetskaia NF, Bianchi ME, Evans WE (2003) A nuclear protein complex containing high mobility group proteins B1 and B2, heat shock cognate protein 70, ERp60, and glyceraldehyde-3-phosphate dehydrogenase is involved in the cytotoxic response to DNA modified by incorporation of anticancer nucleoside analogues. *Cancer research* 63(1): 100-106.
- Lan F, Zaratiegui M, Villen J, Vaughn MW, Verdel A et al. (2007) *S. pombe* LSD1 homologs regulate heterochromatin propagation and euchromatic gene transcription. *Molecular cell* 26(1): 89-101.
- Laudet V, Stehelin D, Clevers H (1993) Ancestry and diversity of the HMG box superfamily. *Nucleic acids research* 21(10): 2493-2501.
- Leupold U, Sipiczki M, Egel R (1991) Pheromone production and response in sterile mutants of fission yeast. *Current genetics* 20(1-2): 79-85.
- Li J, Kokkola R, Tabibzadeh S, Yang R, Ochani M et al. (2003) Structural basis for the proinflammatory cytokine activity of high mobility group box 1. *Molecular medicine (Cambridge, Mass)* 9(1-2): 37-45.
- Lutzoni F, Kauff F, Cox CJ, McLaughlin D, Celio G et al. (2004) Assembling the fungal tree of life: progress, classification, and evolution of subcellular traits. *American journal of botany* 91(10): 1446-1480.
- Marchler-Bauer A, Lu S, Anderson JB, Chitsaz F, Derbyshire MK et al. (2011) CDD: a Conserved Domain Database for the functional annotation of proteins. *Nucleic acids research* 39(Database issue): D225-229.
- Marechal V (2003) LA DOUBLE VIE DE LA PROTÉINE HMGB1 :COMPOSANT STRUCTURAL DE LA CHROMATINE ET CYTOKINE PRO-INFLAMMATOIRE

Revue Française des Laboratoires.

- Marisa L, Ichante JL, Reymond N, Aggerbeck L, Delacroix H et al. (2007) MAnGO: an interactive R-based tool for two-colour microarray analysis. *Bioinformatics* (Oxford, England) 23(17): 2339-2341.
- Marr KA, Patterson T, Denning D (2002) Aspergillosis. Pathogenesis, clinical manifestations, and therapy. *Infectious disease clinics of North America* 16(4): 875-894, vi.
- Martin T, Lu SW, van Tilbeurgh H, Ripoll DR, Dixelius C et al. (2010) Tracing the origin of the fungal alpha1 domain places its ancestor in the HMG-box superfamily: implication for fungal mating-type evolution. *PloS one* 5(12): e15199.
- Mata J, Bähler J (2006) Global roles of Ste11p, cell type, and pheromone in the control of gene expression during early sexual differentiation in fission yeast. *Proceedings of the National Academy of Sciences of the United States of America* 103(42): 15517-15522.
- Metzenberg RL, Glass NL (1990) Mating type and mating strategies in *Neurospora*. *Bioessays* 12(2): 53-59.
- Moreira JM, Holmberg S (2000) Chromatin-mediated transcriptional regulation by the yeast architectural factors NHP6A and NHP6B. *The EMBO journal* 19(24): 6804-6813.
- Mueller, Bills (2004) Biodiversity of fungi: inventory and monitoring methods. *Biodiversity of fungi: inventory and monitoring methods*
San Diego: Elsevier Academic Press. pp. 1-4.
- Muller S, Scaffidi P, Degryse B, Bonaldi T, Ronfani L et al. (2001) New EMBO members' review: the double life of HMGB1 chromatin protein: architectural factor and extracellular signal. *The EMBO journal* 20(16): 4337-4340.
- Murphy FVt, Sweet RM, Churchill ME (1999) The structure of a chromosomal high mobility group protein-DNA complex reveals sequence-neutral mechanisms important for non-sequence-specific DNA recognition. *The EMBO journal* 18(23): 6610-6618.
- Nadin-Davis SA, Nasim A (1988) A gene which encodes a predicted protein kinase can restore some functions of the ras gene in fission yeast. *The EMBO journal* 7(4): 985-993.
- Neiman AM, Stevenson BJ, Xu HP, Sprague GF, Jr., Herskowitz I et al. (1993) Functional homology of protein kinases required for sexual differentiation in *Schizosaccharomyces pombe* and *Saccharomyces cerevisiae* suggests a conserved signal transduction module in eukaryotic organisms. *Molecular biology of the cell* 4(1): 107-120.
- Nemeth A, Langst G (2004) Chromatin higher order structure: opening up chromatin for transcription. *Briefings in functional genomics & proteomics* 2(4): 334-343.
- Nicolas E, Lee MG, Hakimi MA, Cam HP, Grewal SI et al. (2006) Fission yeast homologs of human histone H3 lysine 4 demethylase regulate a common set of genes with diverse functions. *The Journal of biological chemistry* 281(47): 35983-35988.
- Ninomiya Y, Suzuki K, Ishii C, Inoue H (2004) Highly efficient gene replacements in *Neurospora* strains deficient for nonhomologous end-joining. *Proceedings of the National Academy of Sciences of the United States of America* 101(33): 12248-12253.
- Obara T, Nakafuku M, Yamamoto M, Kaziro Y (1991) Isolation and characterization of a gene encoding a G-protein alpha subunit from *Schizosaccharomyces pombe*: involvement in mating and sporulation pathways. *Proceedings of the National Academy of Sciences of the United States of America* 88(13): 5877-5881.
- Orphanides G, Wu WH, Lane WS, Hampsey M, Reinberg D (1999) The chromatin-specific transcription elongation factor FACT comprises human SPT16 and SSRP1 proteins. *Nature* 400(6741): 284-288.
- Pallier C, Scaffidi P, Chopineau-Proust S, Agresti A, Nordmann P et al. (2003) Association of chromatin proteins high mobility group box (HMGB) 1 and HMGB2 with mitotic chromosomes. *Molecular biology of the cell* 14(8): 3414-3426.

- Palumbo R, Sampaolesi M, De Marchis F, Tonlorenzi R, Colombetti S et al. (2004) Extracellular HMGB1, a signal of tissue damage, induces mesoangioblast migration and proliferation. *The Journal of cell biology* 164(3): 441-449.
- Paoletti M, Saupe SJ (2008) The genome sequence of *Podospora anserina*, a classic model fungus. *Genome biology* 9(5): 223.
- Papazian HP (1956) Sex and cytoplasm in the fungi. *Transactions of the New York Academy of Sciences* 18(4): 388-397.
- Parisi MA, Clayton DA (1991) Similarity of human mitochondrial transcription factor 1 to high mobility group proteins. *Science (New York, NY)* 252(5008): 965-969.
- Park JS, Svetkauskaite D, He Q, Kim JY, Strassheim D et al. (2004) Involvement of toll-like receptors 2 and 4 in cellular activation by high mobility group box 1 protein. *The Journal of biological chemistry* 279(9): 7370-7377.
- Parkkinen J, Raulo E, Merenmies J, Nolo R, Kajander EO et al. (1993) Amphoterin, the 30-kDa protein in a family of HMG1-type polypeptides. Enhanced expression in transformed cells, leading edge localization, and interactions with plasminogen activation. *The Journal of biological chemistry* 268(26): 19726-19738.
- Pauken CM, Nagle DL, Bucan M, Lo CW (1994) Molecular cloning, expression analysis, and chromosomal localization of mouse Hmg1-containing sequences. *Mamm Genome* 5(2): 91-99.
- Perego P, Jimenez GS, Gatti L, Howell SB, Zunino F (2000) Yeast mutants as a model system for identification of determinants of chemosensitivity. *Pharmacological reviews* 52(4): 477-492.
- Phair J, Munoz A, Detels R, Kaslow R, Rinaldo C et al. (1990) The risk of *Pneumocystis carinii* pneumonia among men infected with human immunodeficiency virus type 1. Multicenter AIDS Cohort Study Group. *The New England journal of medicine* 322(3): 161-165.
- Pi H, Chien CT, Fields S (1997) Transcriptional activation upon pheromone stimulation mediated by a small domain of *Saccharomyces cerevisiae* Ste12p. *Mol Cell Biol* 17(11): 6410-6418.
- Pil PM, Lippard SJ (1992) Specific binding of chromosomal protein HMG1 to DNA damaged by the anticancer drug cisplatin. *Science (New York, NY)* 256(5054): 234-237.
- Pinan-Lucarre B, Paoletti M, Clave C (2007) Cell death by incompatibility in the fungus *Podospora*. *Seminars in cancer biology* 17(2): 101-111.
- Poggeler S, Kuck U (2001) Identification of transcriptionally expressed pheromone receptor genes in filamentous ascomycetes. *Gene* 280(1-2): 9-17.
- Pöggeler S, Risch S, Kück U, Osiewacz HD (1997) Mating-type genes from the homothallic fungus *Sordaria macrospora* are functionally expressed in a heterothallic ascomycete. *Genetics* 147(2): 567-580.
- Ponting CP (2002) Novel domains and orthologues of eukaryotic transcription elongation factors. *Nucleic acids research* 30(17): 3643-3652.
- Punta M, Cogill PC, Eberhardt RY, Mistry J, Tate J et al. (2012) The Pfam protein families database. *Nucleic acids research* 40(Database issue): D290-301.
- Qin J, Kang W, Leung B, McLeod M (2003) Ste11p, a high-mobility-group box DNA-binding protein, undergoes pheromone- and nutrient-regulated nuclear-cytoplasmic shuttling. *Molecular and cellular biology* 23(9): 3253-3264.
- Reddy MC, Christensen J, Vasquez KM (2005) Interplay between human high mobility group protein 1 and replication protein A on psoralen-cross-linked DNA. *Biochemistry* 44(11): 4188-4195.
- Reeves R, Adair JE (2005) Role of high mobility group (HMG) chromatin proteins in DNA repair. *DNA repair* 4(8): 926-938.
- Rizet G (1952) Les phénomènes de barrage chez *Podospora anserina*. I. Analyse génétique des barrages entre souches *S* and *s*. *Rev Cytol Biol Veg* 13: 51-92.

- Samejima I, Mackie S, Fantes PA (1997) Multiple modes of activation of the stress-responsive MAP kinase pathway in fission yeast. *The EMBO journal* 16(20): 6162-6170.
- Samejima I, Mackie S, Warbrick E, Weisman R, Fantes PA (1998) The fission yeast mitotic regulator *win1+* encodes an MAP kinase kinase kinase that phosphorylates and activates *Wis1* MAP kinase kinase in response to high osmolarity. *Molecular biology of the cell* 9(8): 2325-2335.
- Saupe SJ, Clave C, Sabourin M, Begueret J (2000) Characterization of *hch*, the *Podospora anserina* homolog of the *het-c* heterokaryon incompatibility gene of *Neurospora crassa*. *Current genetics* 38(1): 39-47.
- Shear CL, Dodge BO (1927) *Life Histories and Heterothallism of the Red Bread-Mold Fungi of the Monilia sitophila group.* pp. 1019-1042.
- Shepherd JC, McGinnis W, Carrasco AE, De Robertis EM, Gehring WJ (1984) Fly and frog homoeo domains show homologies with yeast mating type regulatory proteins. *Nature* 310(5972): 70-71.
- Shimajima T, Okada M, Nakayama T, Ueda H, Okawa K et al. (2003) *Drosophila* FACT contributes to Hox gene expression through physical and functional interactions with GAGA factor. *Genes & development* 17(13): 1605-1616.
- Shirakawa H, Yoshida M (1992) Structure of a gene coding for human HMG2 protein. *The Journal of biological chemistry* 267(10): 6641-6645.
- Shirakawa H, Tsuda K, Yoshida M (1990) Primary structure of non-histone chromosomal protein HMG2 revealed by the nucleotide sequence. *Biochemistry* 29(18): 4419-4423.
- Shiu PK, Glass NL (2000) Cell and nuclear recognition mechanisms mediated by mating type in filamentous ascomycetes. *Current opinion in microbiology* 3(2): 183-188.
- Silar P (1995) Two new easy to use vectors for transformations *Fungal Genet Newslett* 42: 73.
- Sinclair AH, Berta P, Palmer MS, Hawkins JR, Griffiths BL et al. (1990) A gene from the human sex-determining region encodes a protein with homology to a conserved DNA-binding motif. *Nature* 346(6281): 240-244.
- Soullier S, Jay P, Poulat F, Vanacker JM, Berta P et al. (1999) Diversification pattern of the HMG and SOX family members during evolution. *Journal of molecular evolution* 48(5): 517-527.
- Souza CA, Silva CC, Ferreira AV (2003) Sex in fungi: lessons of gene regulation. *Genet Mol Res* 2(1): 136-147.
- Sprague GF, Jr., Jensen R, Herskowitz I (1983) Control of yeast cell type by the mating type locus: positive regulation of the alpha-specific STE3 gene by the MAT alpha 1 product. *Cell* 32(2): 409-415.
- Stott K, Tang GS, Lee KB, Thomas JO (2006) Structure of a complex of tandem HMG boxes and DNA. *Journal of molecular biology* 360(1): 90-104.
- Stros M (2009) HMGB proteins: interactions with DNA and chromatin. *Biochimica et biophysica acta* 1799(1-2): 101-113.
- Sugimoto A, Iino Y, Maeda T, Watanabe Y, Yamamoto M (1991) *Schizosaccharomyces pombe* *stell1+* encodes a transcription factor with an HMG motif that is a critical regulator of sexual development. *Genes & development* 5(11): 1990-1999.
- Susan-E Eichhorn, Ray-F Evert, Raven P-H (2000) *Biologie végétale.* In: Boeck D, editor. *Biologie végétale.* pp. 944 pages.
- Tamura K, Dudley J, Nei M, Kumar S (2007) MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. *Molecular biology and evolution* 24(8): 1596-1599.
- Tanaka K, Davey J, Imai Y, Yamamoto M (1993) *Schizosaccharomyces pombe* *map3+* encodes the putative M-factor receptor. *Mol Cell Biol* 13(1): 80-88.
- Taylor JW, Spatafora J, O'Donnell K, Lutzoni F, James T et al. (2004) *The Fungi. In Assembling the Tree of Life.* Oxford University Press: 171-194.

- Teo SH, Grasser KD, Thomas JO (1995) Differences in the DNA-binding properties of the HMG-box domains of HMG1 and the sex-determining factor SRY. *European journal of biochemistry / FEBS* 230(3): 943-950.
- Tetsuya Ueda MY (2009) HMGB proteins and transcriptional regulation.
- Thomas JO, Travers AA (2001) HMG1 and 2, and related 'architectural' DNA-binding proteins. *Biochemical Sciences*.
- Thomsen MS, Franssen L, Launholt D, Fojan P, Grasser KD (2004) Interactions of the basic N-terminal and the acidic C-terminal domains of the maize chromosomal HMGB1 protein. *Biochemistry* 43(25): 8029-8037.
- Toda T, Shimanuki M, Yanagida M (1991) Fission yeast genes that confer resistance to staurosporine encode an AP-1-like transcription factor and a protein kinase related to the mammalian ERK1/MAP2 and budding yeast FUS3 and KSS1 kinases. *Genes & development* 5(1): 60-73.
- Travis A, Amsterdam A, Belanger C, Grosschedl R (1991) LEF-1, a gene encoding a lymphoid-specific protein with an HMG domain, regulates T-cell receptor alpha enhancer function [corrected]. *Genes & development* 5(5): 880-894.
- Turgeon BG, Debuchy R (2007) *Cochliobolus* and *Podospora*: mechanism of sex determination and the evolution of reproductive lifestyle. In: Heitman J, Kronstad JW, Taylor JW, Casselton LA, editors. *Sex in Fungi, molecular determination and evolutionary implications*. Washington, DC: ASM Press. pp. 93-121.
- Vaccaria T, Beltramea M, S. Ferrara, Bianchia ME (1998) Hmg4,a New Member of theHmg1/2Gene Family. *Genomics* 49: 247-252
- Valbuena N, Moreno S (2010) TOR and PKA pathways synergize at the level of the Ste11 transcription factor to prevent mating and meiosis in fission yeast. *PloS one* 5(7): e11514.
- van Beest M, Dooijes D, van De Wetering M, Kjaerulff S, Bonvin A et al. (2000) Sequence-specific high mobility group box factors recognize 10-12-base pair minor groove motifs. *The Journal of biological chemistry* 275(35): 27266-27273.
- van de Wetering M, Oosterwegel M, Dooijes D, Clevers H (1991) Identification and cloning of TCF-1, a T lymphocyte-specific transcription factor containing a sequence-specific HMG box. *The EMBO journal* 10(1): 123-132.
- Wagner CR, Hamana K, Elgin SC (1992) A high-mobility-group protein and its cDNAs from *Drosophila melanogaster*. *Molecular and cellular biology* 12(5): 1915-1923.
- Wang H, Bloom O, Zhang M, Vishnubhakat JM, Ombrellino M et al. (1999) HMG-1 as a late mediator of endotoxin lethality in mice. *Science (New York, NY)* 285(5425): 248-251.
- Wang Y, Xu HP, Riggs M, Rodgers L, Wigler M (1991) *byr2*, a *Schizosaccharomyces pombe* gene encoding a protein kinase capable of partial suppression of the *ras1* mutant phenotype. *Molecular and cellular biology* 11(7): 3554-3563.
- Warbrick E, Fantes PA (1991) The *wis1* protein kinase is a dosage-dependent regulator of mitosis in *Schizosaccharomyces pombe*. *The EMBO journal* 10(13): 4291-4299.
- Watanabe Y, Shinozaki-Yabana S, Chikashige Y, Hiraoka Y, Yamamoto M (1997) Phosphorylation of RNA-binding protein controls cell cycle switch from mitotic to meiotic in fission yeast. *Nature* 386(6621): 187-190.
- Waterhouse AM, Procter JB, Martin DM, Clamp M, Barton GJ (2009) Jalview Version 2--a multiple sequence alignment editor and analysis workbench. *Bioinformatics (Oxford, England)* 25(9): 1189-1191.
- Webb M, Thomas JO (1999) Structure-specific binding of the two tandem HMG boxes of HMG1 to four-way junction DNA is mediated by the A domain. *Journal of molecular biology* 294(2): 373-387.
- Weir HM, Kraulis PJ, Hill CS, Raine AR, Laue ED et al. (1993) Structure of the HMG box motif in the B-domain of HMG1. *The EMBO journal* 12(4): 1311-1319.
- White MA, Riles L, Cohen BA (2009) A systematic screen for transcriptional regulators of the yeast cell cycle. *Genetics* 181(2): 435-446.

- Willer M, Hoffmann L, Styrkarsdottir U, Egel R, Davey J et al. (1995) Two-step activation of meiosis by the *mat1* locus in *Schizosaccharomyces pombe*. *Molecular and cellular biology* 15(9): 4964-4970.
- Wood V, Gwilliam R, Rajandream MA, Lyne M, Lyne R et al. (2002) The genome sequence of *Schizosaccharomyces pombe*. *Nature* 415(6874): 871-880.
- Workman JL, Kingston RE (1998) Alteration of nucleosome structure as a mechanism of transcriptional regulation. *Annual review of biochemistry* 67: 545-579.
- Xue-Franzen Y, Kjaerulff S, Holmberg C, Wright A, Nielsen O (2006) Genomewide identification of pheromone-targeted transcription in fission yeast. *BMC genomics* 7: 303.
- Yamada H, Mori H, Momoi H, Nakagawa Y, Ueguchi C et al. (1994) A fission yeast gene encoding a protein that preferentially associates with curved DNA. *Yeast (Chichester, England)* 10(7): 883-894.
- Yamamoto M, Imai Y, Watanabe Y (1997) Mating and Sporulation in *Schizosaccharomyces pombe*. In: J. R. Pringle JR BaEWJ, editor. *The Molecular and Cellular Biology of the Yeast Saccharomyces*. pp. 1037–1106.
- Zhang NN, Dudgeon DD, Paliwal S, Levchenko A, Grote E et al. (2006) Multiple signaling pathways regulate yeast cell death during the response to mating pheromones. *Molecular biology of the cell* 17(8): 3409-3422.
- Zickler D, Arnais S, Coppin E, Debuchy R, Picard M (1995) Altered mating-type identity in the fungus *Podospora anserina* leads to selfish nuclei, uniparental progeny, and haploid meiosis. *Genetics* 140(2): 493-503.