

THÈSE POUR OBTENIR LE GRADE DE DOCTEUR DE MONTPELLIER SUPAGRO

En Sciences agronomique et écologie

École doctorale GAIA – Biodiversité, Agriculture, Alimentation, Environnement, Terre, Eau
Portée par l'Université de Montpellier

Unité de recherche AGAP



Manipulation de la recombinaison chez une plante cultivée, le riz.

Présentée par Delphine MIEULET
Le 27 novembre 2017

Sous la direction de Emmanuel GUIDERDONI

Devant le jury composé de

Anne-Marie CHEVRE, DR2, IGEPP INRA -Rennes
Mathilde GRELON, DR2, IJPB INRA-Versailles
Wyatt PAUL, Chercheur, Biogemma-Clermont
Pierre SOURDILLE, DR2, GEDEC INRA-CLermont
Jacques DAVID, Enseignant-chercheur, AGAP-Montpellier
Raphaël MERCIER, DR1, IJPB INRA-Versailles

Rapporteur
Rapporteur
Examineur
Examineur
Président du jury
Invité



UNIVERSITÉ
DE MONTPELLIER



Remerciements

Avant toute chose, je tiens à remercier les membres du jury qui ont accepté d'évaluer mon travail et de se déplacer jusqu'à Montpellier pour quelques heures. Merci donc aux rapporteurs Anne-Marie Chèvre et Mathilde Grelon et aux examinateurs Wyatt Paul, Pierre Sourdille et Jacques David. Je voudrais également remercier, l'invité de ce jury, Raphaël Mercier, qui m'accompagne sur ce sujet depuis le début et sans lequel il n'y aurait pas eu cette thèse sur la méiose. Je dois également te remercier Raph pour tout le soutien que tu m'as apporté, les coups de fils et les 10 000 mails qui encombrent ta boîte et auxquels tu réponds à vitesse grand V.

Je souhaite évidemment remercier mon chef de toujours, Emmanuel Guiderdoni, qui m'a fait confiance une première fois en 2008 lors de mon recrutement au CIRAD puis, qui a accepté d'être mon directeur de thèse en 2015. Merci aussi à ma direction d'avoir aménagé mon poste de travail afin de me laisser 3 ans pour vivre cette aventure. Avec Manu : Mr Riz, et Raph : Mr Méiose, j'étais super bien partie pour faire de belles choses sur la méiose du Riz. Mille mercis Manu de m'avoir épaulé jusqu'au bout et pour la vision de malade que tu as du sujet. Il me manquait tout de même le petit grain de sel de Brigitte Courtois, ma collègue de bureau et de surcroît sélectionneuse et généticienne. Brigitte m'a appris tout ce que je devais savoir sur la génomique et donné de bons conseils pour conduire ce travail. Merci Bibi aussi pour ton soutien de tous les jours et la relecture du manuscrit.

Merci aussi à Christophe Perin de m'avoir accueilli dans son équipe pour faire ma thèse, ainsi qu'à tous les membres de l'équipe DAR et aux nombreux CDD et thésards que j'ai croisé dans les labos. Merci également à l'équipe de Raphaël qui m'a accueilli plusieurs fois à Versailles et tout appris de la cyto. Merci aussi aux bioinformaticiens de m'avoir aidée à analyser deux trois trucs en un tour de main. Merci surtout aux personnels des plateaux d'Agap, qui chacun à leur tour m'ont guidé. Merci enfin aux serristes sans qui il n'y aurait pas de matériel.

Je voulais aussi remercier tous les chercheurs qui m'ont aidée à financer ce travail, Manu, Christophe, Brigitte, Raph, Nour, Robert, Vincent, Jef ...

Au-delà de l'équipe et de la science, je tiens à remercier tous mes collègues du CIRAD qui m'ont encouragé et surtout mes voisins de bureau qui m'ont fait le café ces derniers temps.

Cette thèse c'est une super aventure ! C'est scientifiquement ce que je voulais faire. C'est psychologiquement ce que je voulais vivre. C'est physiquement aussi dur qu'un marathon mais qu'est-ce que c'est bon quand ça se termine !

Je tiens tout particulièrement à remercier mes sponsors :

- ✓ Mes Amis qui m'amènent régulièrement prendre l'air en montagne ;
- ✓ Mes parents qui ont toujours cru en moi ;
- ✓ Mon petit frère qui sait rester discret mais qui m'a bien remis d'aplomb ;
- ✓ Et merci à mon beauf et ma belle d'être là dans "*la famiglia*" ;
- ✓ Et enfin ma sœur qui depuis le sprint final ne me lâche pas. Merci Coach

Table des matières

Table des matières

Table des figures	1
Table des annexes	3
Abréviations et nom de gènes	4
Contexte et objectifs des travaux de thèse	8
PARTIE I : Synthèse bibliographique	12
1. Le riz, une céréale essentielle	14
1.1. Importance alimentaire.....	14
1.2. Les systèmes de culture rizicoles.....	16
1.3. La diversité et domestication	18
1.3.1. Structuration du genre <i>Oryza</i>	18
1.3.2. Domestication d' <i>Oryza sativa</i>	18
1.3.3. Diversité d' <i>Oryza sativa</i>	19
1.4. Sélection du riz	21
1.4.1. Méthodes de sélection et apport des biotechnologies et de la génomique	21
1.4.2. Hétérosis et hybrides F1	23
1.5. Ressources et outils disponibles pour l'amélioration variétale	24
1.5.1. Diversité naturelle et induite	24
1.5.1.1. Des milliers de variétés de riz.....	26
1.5.1.2. La diversité induite : aléatoire ou ciblée	26
1.5.2. Matériel à haute résolution génétique.....	27
1.5.3. Ressources génomiques et bio-informatiques	27
1.5.4. Les outils d'ingénierie cellulaire et d'édition des génomes	28
1.6. Biologie du riz	30
1.6.1. Phase végétative.....	30
1.6.2. Phase reproductive	32
2. Reproduction sexuée et apomixie	34
2.1. La reproduction sexuée et la double fécondation.....	34
2.1.1. La gamétogenèse mâle et femelle dans la jeune fleur	34
2.1.2. La double fécondation et la formation de grain	37

2.2.	La reproduction asexuée clonale par grain : l'apomixie	39
2.2.1.	Les différents types d'apomixie	39
2.2.1.1.	Apomixie sporophytique	39
2.2.1.2.	Apomixie gamétophytique	39
2.2.2.	Le contrôle génétique et épigénétique de l'apomixie	40
3.	Méiose et recombinaison.....	44
3.1.	Déroulement de la méiose	46
3.1.1.	La prophase I de méiose, le siège de la recombinaison.....	46
3.1.1.1.	Leptotène : Mise en place des axes chromosomiques.....	48
3.1.1.2.	Recherche des homologues via la formation du bouquet.....	50
3.1.1.3.	Zygotène : Mise en place du complexe synaptonémal	51
3.1.1.4.	Pachytène : description et rôle du chiasma ou CO.....	52
3.1.2.	Progression de la méiose et formation des gamètes.....	53
3.2.	Les mécanismes moléculaires de la recombinaison méiotique	54
3.2.1.	Formation de la cassure double-brin de l'ADN : l'initiation.....	54
3.2.2.	Localisation et modèle de formation des cassures double-brin	58
3.2.3.	Maturation des cassures doubles-brins.....	58
3.2.4.	L'invasion par le simple-brin et la formation d'intermédiaires.....	60
3.2.5.	Les voies de réparation des CDB.....	62
3.2.5.1.	La voie ZMM de formation des COs	63
3.2.5.2.	La voie MUS81 de formation des COs	64
3.2.5.3.	Les voies de formation des Non Crossing-Overs (NCO).....	65
3.2.6.	Les voies de régulation des COs : les gènes anti-COs	66
3.2.6.1.	L'hélicase FANCM	66
3.2.6.2.	L'AAA-ATPase FIGETIN-LIKE1.....	67
3.2.6.3.	L'hélicase RECQ	67
3.3.	Comment détecter les événements de recombinaison ?.....	70
3.3.1.	Techniques de cytogénétique.....	70
3.3.2.	Par établissement de carte génétique dans une descendance	70
3.3.3.	Calcul de la fréquence de recombinaison.....	71
3.3.4.	Déséquilibre de liaison et CO.....	73
3.4.	Distribution des évènements de recombinaison.....	75
3.4.1.	Nature et distribution des événements de recombinaison	75
3.4.1.1.	Les phénomènes qui régulent le nombre de CO	75
3.4.1.2.	Nombre moyen de CO par chromosomes selon les espèces	77
3.4.1.3.	Structure du génome et distribution des COs	77

Partie II : Abolir la recombinaison.....	83
4. L'apoméiose chez le riz.....	83
4.1. Introduction.....	83
4.2. Article : "Turning rice meiosis into mitosis"	87
4.3. Discussion et perspectives.....	100
PARTIE III : Augmenter la recombinaison	106
5. Augmenter le taux de recombinaison chez le riz en utilisant les protéines anti-CO <i>FANCM</i> et <i>RECQ14</i>.....	108
5.1. Introduction.....	108
5.2. Matériel et Méthodes	112
5.2.1. Phylogénie : Recherche d'orthologues	112
5.2.2. Matériel génétique	112
5.2.2.1. Recherche de mutants d'insertion et génotypage	112
5.2.2.2. Caractérisation moléculaire des mutants ADN-T	113
5.2.2.3. Développement de populations F2 pour l'étude des gènes anti-CO.....	113
5.2.3. Caractérisation phénotypique des F1.....	116
5.2.3.1. Méiose mâle : cytogénétique.....	116
5.2.3.2. Fertilité paniculaire (que le simple REC sur F1).....	117
5.2.4. Génotypage SNP	117
5.2.5. Construction des cartes génétiques	121
5.2.6. Condition de culture et de croisement de plantes	121
5.3. Résultats.....	124
5.3.1. Recherche des homologues de <i>AtFANCM</i> et <i>AtRECQA/B</i> chez le riz.....	124
5.3.2. Identification de mutant perte de fonction.....	124
5.3.2.1. Détermination de KO dans le gène Os11g07870 (<i>OsFANCM</i>)	124
5.3.2.2. Détermination de KO dans le gène Os04g35420 (<i>OsRECQ14</i>)	126
5.3.3. Cumul des mutations par croisement.....	126
5.3.3.1. Obtention de population F2 simple mutant <i>RECQ4</i>	126
5.3.3.2. Obtention de populations F2 simples mutants <i>FANCM</i>	127
5.3.3.3. Obtention de populations F2 double mutants <i>FANCMxRECQ4</i>	127
5.3.4. Caractérisation phénotypique des hybrides F1	129
5.3.5. Cartographie des événements de recombinaison.....	129
5.3.5.1. Dans une population F2 "simple <i>RECQ14</i> "	129
5.3.5.2. Dans une population F2 "simple <i>FANCM</i> "	136
5.4. Discussion et perspectives.....	137

PARTIE IV : Inactivation de gènes anti-COs par CRISPR/cas9	145
6. Introduction.....	145
7. Création de lignées perte de fonction par CRISPR/Cas9.....	146
7.1. Matériel et méthodes.....	146
7.2. Résultats et discussion	150
8. Transformation d'hybrides F1	151
Discussion générale.....	153
Bibliographie.....	159
Annexes.....	172

Table des figures

Figure 1 : Culture et consommation de riz dans le monde.

Figure 2 : Diversité des écosystèmes rizicoles.

Figure 3 : Diversité du Genre *Oryza* (d'après d'après Zhang & Wing, 2013).

Figure 4 : Classification de 3 000 accessions de riz en 2 sous-espèces et 5 groupes variétaux (adaptation de Li et al. 2014).

Figure 5 : Hérité de la stérilité mâle à contrôle géno-cytoplasmique.

Figure 6 : Stades de développement du riz (d'après l'IRRI).

Figure 7 : Panicule de riz et épillet (d'après CGIAR 2013).

Figure 8 : Développement de la fleur de riz au moment de l'anthèse (d'après Hoshikaw, 1989).

Figure 9 : Gamétogenèse mâle, formation du grain de pollen chez le riz (d'après Hoshikaw, 1989).

Figure 10 : Gamétogenèse femelle, formation du sac embryonnaire chez le riz (d'après Hoshikaw, 1989).

Figure 11 : La double fécondation chez le riz (d'après Hoshikaw, 1989).

Figure 12 : Mécanismes de développement de semences par voie sexuée ou apomictique (Hand et al. 2014).

Figure 13 : Déroulement de la méiose (Mercier et al. 2015).

Figure 14 : Rôle de la cohésine REC8 dans la ségrégation correcte des chromosomes homologues en M1 puis des chromatides-sœurs en MII (Sakuno et Wanatabe 2009).

Figure 15 : Prophase I de méiose et mise en place de l'axe chromosomique (adaptation de Lambing et al. 2017).

Figure 16 : Le bouquet, une organisation caractéristique de la méiose (Zickler and Kleckner, 2015).

Figure 17 : Modèle présentant les mécanismes de la recombinaison méiotique (d'après Mercier et al. 2015).

Figure 18 : Rôle de *Spp1* dans la formation des CDBs chez la levure (Sommermeier et al. 2013).

Figure 19 : Formation et maturation des cassures double-brin de l'ADN (d'après Neale et al. 2005).

Figure 20 : Chargement des recombinaisons et invasion de la molécule homologue (d'après Neale et al. 2005).

Figure 21 : Représentation schématique du déséquilibre (a) et (b) de l'équilibre de liaison entre deux locus (Flint-Garcia et al., 2003 ; Guo et al., 2014).

Figure 22 : Le crossing-over obligatoire et l'interférence (Jones et al. 2006).

- Figure 23** : L'interférence pendant la recombinaison méiotique (Ziolkowski & Handerson 2017).
- Figure 24** : Nombre moyen de Crossing-Overs (COs) par chromosome et par méiose chez différentes espèces eucaryotes et chez le mutant *Atrecq4 figl1* (Mercier et al., 2015).
- Figure 25** : La structure du génome et la distribution des COs diffère entre *A. thaliana*, le riz et le maïs (Lambing et al. 2017).
- Figure 26** : Connection des chromosomes en méiose et attachement des kinétochores (Talbert 2010).
- Figure 27** : Les voies de formation des COs et les protéines anti-CO impliquées (Mercier et al. 2015).
- Figure 28** : Compaction de la chromatine et premières données de recombinaisons chez le riz (d'après Cheng et al. 2001 et Wu et al. 2003).
- Figure 29** : Schéma de croisement pour l'étude d'un simple mutant anti-CO.
- Figure 30** : Schéma de croisement pour l'étude d'un double mutant anti-CO.
- Figure 31** : Position des marqueurs Kasp® sur les chromosomes 1 à 6 du riz.
- Figure 32** : Position des marqueurs Kasp® sur les chromosomes 7 à 12 du riz.
- Figure 33** : Conservation des domaines protéiques entre *A. thaliana* et *Oryza sativa* pour FANCM et RECQI4 (Expasy).
- Figure 34** : Position des mutations dans les gènes.
- Figure 35** : Lignées utilisées dans les croisements en vue de l'obtention des simples et doubles mutants.
- Figure 36** : Fertilité paniculaire des hybrides F1 *Osrecq14* --/-- et des sauvages.
- Figure 37** : Méiose mâle chez *Osrecq14* --/-- (DAPI).
- Figure 38** : Cartes génétiques chez le sauvage et le mutant *Osrecq14* --/--.
- Figure 39** : Distribution des COs sur les 12 chromosomes du riz chez les plantes F2 de *Osrecq14* --/-- et *OsRECQI4* ++/++.
- Figure 40** : Observation de la recombinaison sur différents intervalles en fonction de la position physique de ces intervalles sur le chromosome.
- Figure 41** : Distribution des COs sur 5 chromosomes de plantes F2 de *Osfancm* --/--.
- Figure 42** : Densité en gènes, en éléments transposables et en SNP sur les 12 chromosomes du riz.
- Figure 43** : Relation entre la densité en SNP et la fréquence de recombinaison chez *Osrecq14* --/--.
- Figure 44** : Stratégie de clonage d'un sgRNA dans le vecteur d'expression pC5300-Cas9.
- Figure 45** : Analyse des plantes éditées par CRISPR/cas9.

Tableau 1 : Liste des collections de mutants de riz (Droc et al. 2013).

Table des annexes

Annexe 1 : Liste des mutants identifiés dans les gènes *FANCM* et *RECQ14*.

Annexe 2 : Liste des amorces utilisées pour la Q-RT-PCR.

Annexe 3 : Données d'expression (RiceXPro3.0) pour *OsRECQ14* (Os04g35420), et *OsFANCM* (Os11g07870).

Annexe 4 : Cartes génétiques du sauvage et du mutant *Osrecq14* --/--.

Annexe 5 : Position des 92 marqueurs utilisés pour génotyper la descendance F1 *Osfancm* --/-- et *Osfancm* ++/++.

Annexe 6 : Densité en gènes et en éléments transposables (TE).

Annexe 7 : Amorces correspondant aux crRNA et nom des construits CRISPR/Cas9 réalisées pour produire des plantes ciblant les gènes anti-CO.

Abréviations et nom de gènes

AGO : Argonaute

ASGR : *Apospory Specific Dominant Region*

At : *Arabidopsis thaliana*

BBM : *Baby Boom*

BBMH : *Best Blast Mutual Hits*

BBML : *Baby Boom Like*

BP : *Backbone Primer*

CDB : Cassures double-brin (ou DSB *Double-strand-break*)

CE : *Central element*

CENH3 : *Centromeric Histone 3*

Chr. : chromosome

cM : centiMorgans

CMM : cellule mère de la mégaspore

CO : Crossing-Over

CRC1 : *Central Region Component1*

CRISPR : *Clustered Regularly Interspaced Short Palindromic Repeats* (courtes répétitions palindromiques groupées et régulièrement espacées)

CS : Complexe Synaptonémal

D-loop : *Displacement loop structure*

DEXDc : *DEAD-like helicases superfamily*

DJ : Dongjin

djH : double jonction de Holliday

DL : déséquilibre de liaison

DMC1 : *Disrupted Meiotic cDNA 1*

DMC1 : *Disrupted Meiotic cDNA1*

DSBR : *Double-strand break repair* (réparation des cassures double-brin)

EA = Elément Axial

EC : Element Central

FA : voie *Fanconi Anemia*

FANCM : *Fanconi Anemia Complementation groupM*

FIE : *Fertilization Independent Endosperm*

FIGL1 : *Fidgetin-Like1*

FLIP : Fidgetin-like-1 Interacting Protein

GC : gène cible

GEM : Genome Elimination Induced

HAP : Haploid initiation

HDR : Homology-Directed Repair

He : Hétérozygote

HEI10 : Human Enhancer of Invasion-10

Ho : Homozygote

HRDC : Homologous region RNase D C-terminal

Hv : Hordeum vulgare

HW : Hwayoung

Jas : Jason

KO : Knock Out

LP : Left Primer

M1 : Première division de Méiose

M2 : Seconde division de Méiose

MEICA 1 : Meiotic Chromosome Association 1

MEL1 : Meiosis arrested at Leptotene 1

MiMe : Meiosis Instead of Meiosis

MSI1 : Multicopy of Suppressor of IRA1

MTL : Matrilineal

NB : Nipponbare

NCO : Non Crossing-Over

NGS : Next-Generation Sequencing

NHEJ : Non-Homologous End Joining, (jonction d'extrémité non-homologues)

NLD : Not Like Dad (or Matrilineal)

Os : *Oryza sativa*

OSD1 : Omission of Second Division 1

PAIR1 : Pairing Aberration in Rice meiosis1

PRC2 : Polycomb Repressive Complex 2

PRD : Putative Recombination Initiation Defect

PRD2 : Putative Recombination initiation Defect2

Ps : *Pennisetum squamulatum*

PS1 : Parallele-Spindel 1

QTL : Quantitative Trait Loci

RAD51 : RADiation-sensitive 51
REC8 : "Meiotic recombination and sister chromatid cohesion"
RECQ14 : RecQ like protein 4
 RH : Recombinaison Homologue (*Homologous Recombination* (HR))
RMI1: RecQ-mediated genome instability 1
 RP : *Right Primer*
RQC : RecQ C-terminal
 SC : *Synaptonemal Complex*
 SD : *standard deviation* (Ecart type)
 SDSA : *Synthesis-Dependent Strand Annealing*
 SEC : *Second End Capture"*
 SEI : *Single End Invasion*
 SNP : *Single Nucleotide Polymorphism*
SPO11 : Sporulation11
SWI1/DYAD : Switch1 (or DYAD)
 T0 : transformant primaire
 TAM : *Tardy Aynchronous Meiosis*
TDM1 : Three-Division Mutant 1
 TE : *Eléments transposables*
Tos17 : Transposon of Oryza Sativa 17
UVI4 : UV-B-insensitive 4
 WT : *wild-type*
 Z11 : *Zhonghua 11*
 Zm : *Zea mays*

Contexte et objectifs des travaux de thèse

La recombinaison génétique contribue à la diversité des espèces. Le brassage génétique a lieu lors de la reproduction sexuée. Depuis la domestication des plantes, les cultivateurs puis les sélectionneurs ont exploité ce brassage génétique naturel pour sélectionner les meilleures combinaisons d'allèles et améliorer la qualité et la productivité des espèces cultivées.

La méiose est un mécanisme très conservé chez les eucaryotes. Depuis ces 15 dernières années, les scientifiques ont acquis de nombreuses connaissances sur les mécanismes moléculaires de la méiose chez les plantes (Luo et al. 2014; Mercier et al. 2015). Début 2014, 28 gènes impliqués dans la méiose du riz avaient été caractérisés. Aujourd'hui, la fonction de plus de 44 gènes a été publiée. Ces connaissances moléculaires permettent à présent une manipulation de la recombinaison méiotique. Chez la plante modèle *Arabidopsis thaliana*, les travaux de l'équipe de R. Mercier de l'institut Jean-Pierre Bourgin (UMR1318 INRA-AgroParisTech) ont montré qu'en abolissant la fonction de certains de ces gènes, il était possible d'obtenir des gamètes non réduits possédant le génotype de la plante mère puis, *via* l'induction d'une parthénogenèse, de réaliser une apomixie artificielle (D'Erfurth et al. 2009; Marimuthu et al. 2011). A l'inverse, ces travaux ont également permis d'augmenter considérablement la recombinaison en augmentant le nombre de crossing-overs (CO) par méiose d'un facteur 3 à 10 (Crismani et al. 2012; Fernandes et al. 2017b).

Pour l'amélioration des plantes cultivées et notamment du riz, ces deux voies de manipulation de la recombinaison sont intéressantes. En effet, abolir la recombinaison permettrait de fixer les variétés hybrides F1, très productives à cause de l'hétérosis, mais dont le coût de production des semences est élevé, et l'achat inaccessible aux paysans du Sud. Manipuler les gènes qui augmentent le brassage génétique, permettrait d'aider les sélectionneurs dans la première phase de leur travail qui consiste à générer de la diversité en favorisant la création de nouvelles combinaisons d'allèles.

L'une des thématiques de l'équipe Développement Adaptatif du Riz (DAR, CIRAD, Montpellier), dans laquelle je travaille depuis 2004, est la manipulation de la recombinaison et le ciblage génique. En 2009, R. Mercier, recherchant des collaborateurs pour reproduire le phénotype *MiMe* d'*A. thaliana* chez une plante d'intérêt agronomique, a contacté l'équipe pour travailler sur le riz. L'ANR a accepté de financer ce projet intitulé DipLOSD (*Diploid gametes via the inhibition of the OSD1 gene or its orthologs*) pour 18 mois et E. Guiderdoni m'a confié la partie Riz du projet. C'est dans le cadre de cette collaboration entre l'IJPB et l'équipe DAR qu'en 2014, la direction de l'Unité AGAP m'a autorisée à mettre entre parenthèses mon rôle de "responsable de plateforme" pour faire une thèse sous la direction d'E. Guiderdoni et avec le soutien de R. Mercier.

Cette thèse intitulée "Manipulation de la recombinaison méiotique chez une plante cultivée, le riz" vise à étudier la conservation de certains mécanismes qui régulent la méiose chez *A. thaliana*.

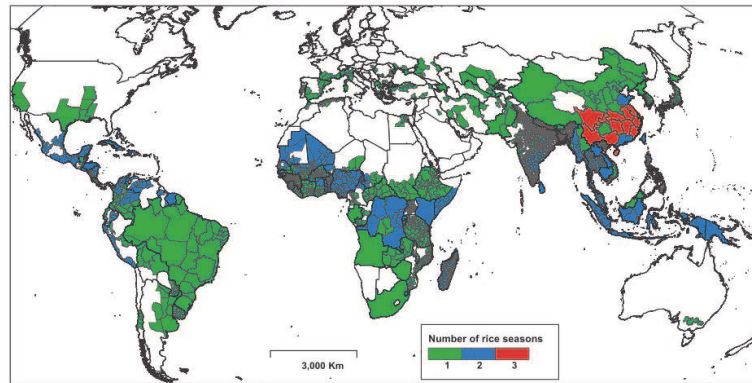
La première partie de cette thèse vise à abolir la recombinaison pour permettre la reproduction clonale par grains (apomixie) afin de fixer l'hétérosis des hybrides F1 cultivés. L'apoméiose est la première des deux composantes de l'apomixie. Elle permet d'obtenir des gamètes non réduits identiques à la plante mère. Ce phénotype a été obtenu en cumulant trois mutations chez *A. thaliana* et a été appelé *MiMe* pour "*Mitosis Instead of Meiosis*" (D'Erfurth et al. 2009). La seconde composante de l'apomixie est la parthénogenèse c'est à dire le développement autonome de l'embryon et de l'albumen. Le croisement d'une lignée *MiMe* avec une lignée capable de déclencher la parthénogenèse, comme la lignée GEN (*Genome elimination induced*), a permis de produire des grains apomictiques chez *A. thaliana* (Marimuthu et al. 2011). Le premier objectif de ma thèse est donc de reproduire le phénotype *MiMe* chez le riz (projet DipLOSD) et ensuite d'étudier les possibilités de déclenchement de la parthénogenèse. Cette étude a fait l'objet d'une publication dans Cell Research en 2016 et est détaillée dans la partie II de ce manuscrit.

S'appuyant toujours sur les travaux conduits à l'IJPB, la seconde partie de ma thèse vise à augmenter globalement la fréquence de recombinaison chez le riz. En effet, chez le riz comme chez 80% des eucaryotes, le nombre de CO est faible (1 à 3) et non homogène le long des chromosomes (Mercier et al. 2015). On observe des régions chaudes et des régions froides de recombinaison ce qui rend très compliqué l'introgression de gènes d'intérêt dans certaines régions froides du génome, en particulier les régions péri-centromériques. Dans le cas de croisement entre deux accessions distantes génétiquement, il est parfois difficile d'obtenir des recombinants à cause du manque d'homologie entre les séquences. Le "linkage drag" dû à ce faible taux de recombinaison est un inconvénient pour les sélectionneurs. Parvenir à casser ces blocs, représente un grand intérêt pour l'amélioration des plantes. Des cribles supprimeurs réalisés chez *A. thaliana* dans une population de mutants constituée dans un fond génétique *zmm* ont permis d'identifier des premiers gènes anti-CO (De Muyt et al. 2009), tels que *AtFANCM* en 2012 puis *AtRECQ4* et *FIGL1* en 2015 (Crismani et al. 2012; Séguéla-Arnaud et al. 2015). L'inactivation d'un ou plusieurs gènes anti-CO permet d'augmenter jusqu'à 780% le nombre de COs en contexte hybride chez le double mutant *Atfigl1* *-/-* *Atrecq4/B* *-/-* par rapport au sauvage (Fernandes et al. 2017b). Tirant avantage de ces découvertes récentes, la seconde partie de cette thèse consiste à évaluer comment l'inactivation des orthologues des gènes anti-CO identifiés chez *A. thaliana*, peut affecter la fréquence globale et la distribution des COs chez le riz. Le travail vise à caractériser des mutants d'insertion dans ces trois gènes, puis à produire des hybrides F1 bi-alléliques mutants dans un fond génétique polymorphe, afin de permettre la cartographie à l'échelle du génome entier des événements de recombinaison. L'étude de la distribution des COs dans la descendance des mutants F1

Osrecql4 -/- et *Osfancm* -/- fait actuellement l'objet de la rédaction d'un article. Les résultats obtenus sur le riz seront associés à ceux obtenus par des collaborateurs de R. Mercier sur le pois (INRA Dijon) et la tomate (INRA Bordeaux) pour une soumission à Nature Communications en fin d'année. Les résultats obtenus sur le riz sont décrits dans la partie III de ce manuscrit. Par ailleurs, nous avons aussi utilisé la technologie CRISPR/Cas9 pour produire de nouveaux allèles mutants. Cette partie du travail est décrite dans la partie résultats complémentaires (Partie IV).

Partie I : SYNTHESE BIBLIOGRAPHIQUE

A



B

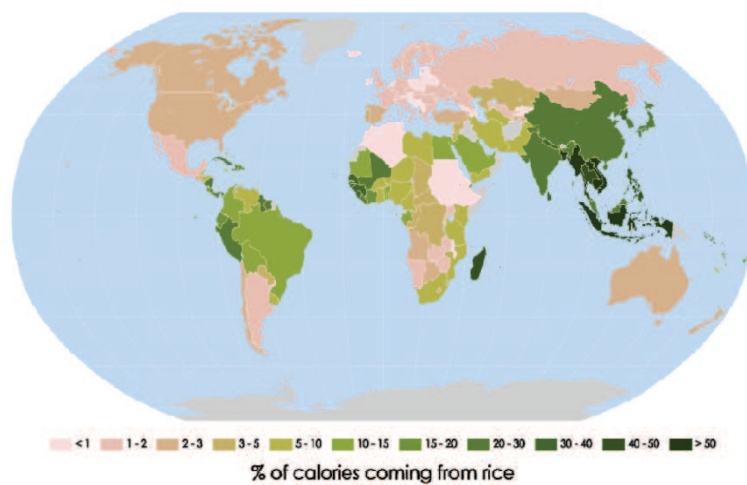


Figure 1 : Culture et consommation de riz dans le monde.

A- Lieux de production. Surfaces cultivées et nombre de cultures par an en fonction des régions (Laborte et al. 2017).

B- Calories apportées par le riz en fonction des pays (Ahmadi 2017)

PARTIE I : Synthèse bibliographique

1. LE RIZ, UNE CEREALE ESSENTIELLE

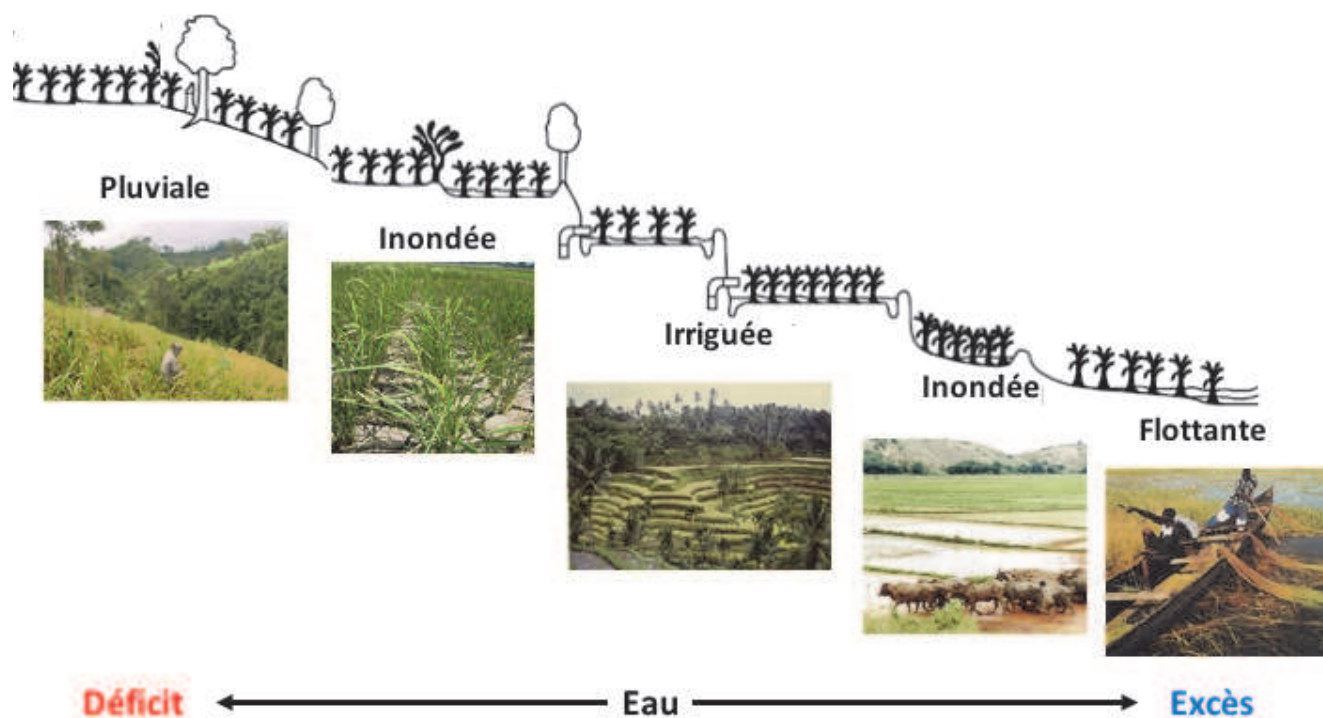
1.1. Importance alimentaire

Le riz, le blé et le maïs, qui sont les 3 principales cultures vivrières du monde, fournissent plus de 42% des calories consommées par la population humaine. En termes de superficie, le blé avec 225 millions d'hectares représente la première céréale suivie par le maïs et le riz qui couvrent chacun 160 millions d'hectares. Le riz est cultivé sur 11% des terres arables de la planète et 90% de ces surfaces se trouvent en Asie du Sud et du Sud-Est (<http://www.fao.org/faosta>). Il est cultivé par plus de 144 millions de paysans dans plus d'une centaine de pays (Laborte et al. 2017) (Figure 1).

Au niveau de la production, le riz est la troisième céréale derrière le maïs et le blé (1 054 millions de tonnes pour le maïs, 740 millions de tonnes pour le blé et 504 Mt de riz blanchi en 2017 (FAOstat). Le riz est la première céréale de consommation humaine (IRRI 2017). 90% du riz est produit et consommé en Asie où il représente entre 35% et 75% des apports caloriques pour plus de 3 milliards d'habitants (Zhang and Wing 2013). En Afrique et en Amérique Latine, où 19 et 25 millions de tonnes sont produites respectivement, le riz devient de plus en plus important puisqu'il représente un tiers des apports caloriques pour 1,5 milliards de personnes (Khush 2005). Il reste l'aliment de base des plus démunis et la consommation annuelle mondiale par habitant demeure stable à 54,1 kg/an (IRRI 2017).

La Chine et l'Inde sont de loin les plus gros producteurs de riz, suivis par l'Indonésie et le Bangladesh, le Vietnam, la Birmanie et la Thaïlande (ricepedia.org). Ces 7 pays représentent à eux seuls 80% de la production mondiale et sont également les principaux consommateurs. Le riz est une production d'autoconsommation, la majorité du riz étant produite dans de petites exploitations familiales de 0,5 à 3 ha maximum (CGIAR 2013) ; de ce fait seuls 6% de la production sont exportés. Les principaux exportateurs de riz sont l'Inde, la Thaïlande, le Vietnam, le Pakistan et les États-Unis (IRRI 2017). Le rendement varie de 1t/ha à plus de 10 t/ha dans les conditions optimales de culture. Le rendement moyen mondial est de 4,3 tonnes par hectare (IRRI 2017). Ce rendement moyen a plus que doublé en 50 ans permettant de ne pas convertir en zones rizicoles près de 230 millions d'hectares (IRRI 2017).

Les Nations Unies prévoient que la population mondiale devrait atteindre 9,7 milliards en 2050 (United Nations 2015). Pour satisfaire les besoins de cette population grandissante, il sera donc nécessaire d'augmenter la production de riz de plus de 42% d'ici à 2050 (production de



	Pluviale	Inondée	Irriguée	Eau profonde	Total (ha)
Afrique (%)	38	34	20	8	9,1
Amérique latine (%)	45		55		6,6
Asie (%)	9	28	57	6	136
USA, UE, Aust. (%)			100		4
Total (%)	10	30	55	5	100
Superficies (Mha)	16	48	86	7	157
Rendement (t/ha)	1 à 5	1 à 5	3 à 10	1 à 3	4,2

Figure 2 : Diversité des écosystèmes rizicoles.

Les quatre grands types de riziculture en fonction de la gestion de l'eau : pluviale, inondée, irriguée, et flottante (illustrations d'Ahmadi, 2017). Le tableau présente les superficies occupées par chaque type de riziculture et les rendements attendus (FAO 2005).

référence année 2008) (Ray et al. 2013). Avec plus de quatre milliards de personnes dans le monde qui se nourrissent de riz chaque jour, le riz est donc un élément essentiel à la sécurité alimentaire et la stabilité politique mondiales (World Rice Statistics, IRRI).

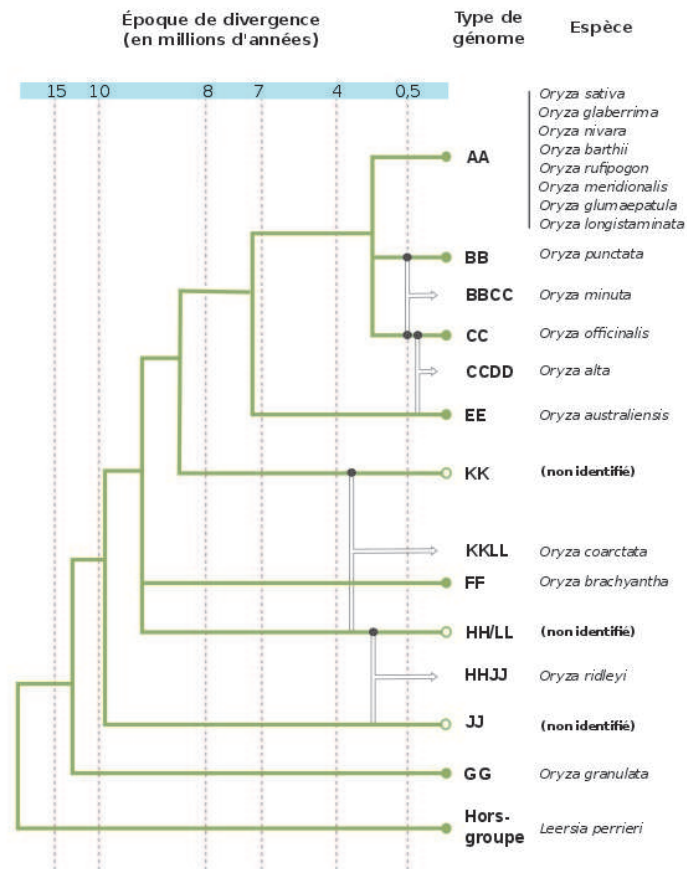
L'enjeu est donc considérable pour les sélectionneurs puisque d'ici 2030 il faudra être capable d'augmenter la production globale de riz de 25% pour supporter la demande croissante (Li et al. 2014). Cette augmentation devra se faire sur des terres arables non extensibles voire en réduction du fait de l'urbanisation ou de la compétition avec d'autres cultures, dans un contexte de raréfaction des ressources (eau, engrais issus de ressources fossiles comme l'azote ou de roches comme le phosphate) et d'instabilité climatique (amenant des extrêmes de température, submersion et salinisation des sols et sécheresse). D'autre part, le riz contribue au réchauffement climatique par la libération de méthane par les rizières. Le riz devra donc devenir "*climate proof*" et "*climate smart*" en réduisant son impact sur l'environnement et en tolérant des variations climatiques.

1.2. Les systèmes de culture rizicoles

Deux espèces de riz sont cultivées : une d'origine Asiatique (*Oryza sativa*) de distribution mondiale et une seconde d'origine Africaine (*Oryza glaberrima*) dont la diffusion est limitée à l'Afrique de l'Ouest. Le riz est une plante annuelle semi aquatique assez exigeante en eau. On estime qu'il faut 4 000 litres d'eau pour produire un Kg de riz contre 1 500 L pour le maïs et 700 L pour le sorgho. Cependant le riz fait preuve d'une grande plasticité vis à vis de ses besoins en eau. La classification des systèmes de riziculture est d'ailleurs fondée sur le type d'alimentation hydrique.

On distingue quatre grands types de riziculture en fonction de la gestion de l'eau dans les cultures (Courtois 1988) : irriguée, inondée, pluviale et flottante. Cette diversité permet de cultiver le riz dans des milieux très variés couvrant une large gamme d'altitudes et de latitudes (Figure 2). La riziculture irriguée, suppose une maîtrise des flux d'eau entrant et sortant. Elle occupe 55 % des surfaces cultivées et représente 75% de la production mondiale. La riziculture inondée est un mode de culture où l'irrigation n'est pas totalement maîtrisée. Elle concerne 30 % des surfaces rizicoles avec des rendements variables du fait de risques importants de submersion et/ou de sécheresse. La riziculture pluviale dont l'alimentation hydrique dépend uniquement de la pluviométrie (riziculture pluviale stricte) ou de la présence d'une nappe éventuelle (riziculture de nappe) représente 10% des surfaces (la majeure partie en Afrique et en Amérique Latine). Enfin, la riziculture flottante, qui suit la crue des grands fleuves, n'occupe que 5 % des surfaces. Les rendements s'améliorent avec la maîtrise de l'eau mais les coûts d'aménagement des rizières augmentent en parallèle (Figure 2). La riziculture irriguée permet une intensification de la culture (jusqu'à 3 cycles de culture par an) et une diminution appréciable des aléas hydrologiques garantissant des rendements élevés (3 t/ha

A



B



Figure 3 : Diversité du Genre *Oryza* (d'après Zhang & Wing, 2013).

A- Arbre phylogénétique du genre *Oryza*.

B- Plantes de 12 espèces représentatives du genre *Oryza*.

en saison des pluies et jusqu'à 10 t/ha en saison sèche). Les variétés très productives et les riz hybrides F1 sont privilégiés dans ce contexte. La culture pluviale, en revanche, ne demande aucun aménagement particulier mais comporte plus de risques, notamment de sécheresse. Il ne peut y avoir qu'un seul cycle de culture et les rendements sont plus faibles et plus variables (de 1 à 5 t/ha) (Figure 2). Afin de pouvoir économiser l'eau dans les systèmes avec maîtrise de l'eau, des itinéraires techniques d'inondation et d'assecs successifs (*alternate water and drying system*) ou totalement aérobies sont en cours de développement. Les variétés utilisées dans ces nouveaux contextes devront faire preuve d'une plasticité morphologique et physiologique comme par exemple un système racinaire réactif allant capturer l'eau et les nutriments en profondeur dans un sol.

1.3. La diversité et domestication

1.3.1. Structuration du genre *Oryza*

Le genre *Oryza* est apparu il y a environ 14 millions d'années (CGIAR 2013). Outre les 2 espèces cultivées *O. sativa* et *O. glaberrima*, le genre *Oryza* comprend 25 espèces sauvages connues à ce jour (CGIAR 2013). La plupart de ces espèces ont été classées en 4 complexes en fonction de leur facilité de croisement permettant de transférer des gènes entre espèces sauvages et espèces cultivées et des observations cytologiques (appariement des chromosomes homéologues chez l'hybride). Mais deux espèces, trop distantes des autres, ont été classées dans une section propre : *O. australiensis* (génome EE) et *O. brachyantha* (génome FF) (Figure 3).

Le complexe *O. sativa* compte les 2 espèces cultivées et 6 espèces sauvages qui ont toutes le génome A d'*O. sativa* ($2n = 24$). Ces espèces sont diploïdes compatibles entre elles et leurs chromosomes montrent un appariement normal des homologues en situation hybride. A l'intérieur de ce complexe *O. sativa*, les croisements entre espèces sont possibles bien que qu'il soit nécessaire dans certain cas de faire du sauvetage d'embryon (exemple de croisement *O. sativa* x *O. longistaminata*). Les hybrides présentent des niveaux très variables de fertilité.

L'ensemble des espèces sauvages constitue un vaste réservoir de gènes, pour la résistance aux stress biotiques et aux contraintes abiotiques, et l'amélioration des variétés de riz de manière plus générale.

1.3.2. Domestication d'*Oryza sativa*

Les traces de la culture d'*Oryza sativa* sont très anciennes et sa domestication remonte à environ 10 000 ans avant J-C. Bien que de nouvelles données génétiques et archéologiques soient disponibles (Gross and Zhao 2014), le processus de domestication de l'espèce *Oryza sativa* n'est toujours pas clair. La première hypothèse est qu'*O. sativa* aurait été domestiquée

deux fois de manière indépendante au sud de la Chine dans la vallée du fleuve Yangtsé et au nord-est de l'Inde dans l'Uttar Pradesh à partir de formes annuelles de l'espèce *O. rufipogon* 8 000 et 5 000 ans avant J-C respectivement. Ces deux domestications indépendantes auraient donné naissance respectivement aux sous espèces japonica et indica. Des données plus récentes issues du séquençage, de l'identification des gènes intervenant dans le syndrome de domestication et de l'analyse des polymorphismes fonctionnels sélectionnés à ces loci conduisent à l'hypothèse d'une première domestication à partir d'une population spécifique d'*O. rufipogon* autour de la zone médiane de la rivière des Perles (ZHU JIANG) dans le sud de la Chine donnant la sous-espèce japonica. La sous-espèce indica aurait ensuite été domestiquée à partir des croisements entre des riz japonica et des riz sauvages locaux alors que les cultivars initiaux se répandaient dans l'Asie du Sud-Est et l'Asie du Sud (Huang et al. 2012). Une troisième théorie prenant en compte les différents groupes variétaux actuels milite pour une triple domestication, la troisième aurait donné naissance au groupe des riz séparé phyllogénétiquement des indica et des japonica tropicaux et tempérés, et présentant des caractères d'adaptation singuliers (Civán et al. 2015).

1.3.3. Diversité d'*Oryza sativa*

La structuration de la diversité de l'espèce *O. sativa* est forte et particulière. C'est une espèce fortement bipolaire avec 2 sous-espèces, les indicas et les japonicas, clairement distingués sur la base de caractéristiques agro-morphologiques, de comportement en croisement, et de marqueurs biochimiques et moléculaires. Si l'obtention de grains hybrides ne pose pas de problème, les plantes F1 issues de ces croisements distants peuvent être sujettes à une stérilité pollinique et paniculaire plus ou moins importante selon les combinaisons. Cette stérilité partielle tend à décroître lors des générations suivantes mais peut s'accompagner d'un retour aux formes parentales et d'un déficit de recombinants. Cette limitation à la recombinaison peut être liée à la méiose (régions d'hétérologies structurales, polymorphisme important) mais également à l'action post-méiotique d'interactions entre gènes de développement sporophytique ou gamétophytique (Ouyang et al. 2010a).

La classification des riz basée sur des critères agro-morphologiques comportait de nombreuses incohérences et ne reflétait pas les observations des sélectionneurs. Par exemple, beaucoup de variétés de type japonica tropicaux avaient été initialement classifiés comme indica du fait du format de leur grain. Les marqueurs biochimiques ont été utilisés pour y mettre de l'ordre. Les premiers travaux conduits avec un échantillon de 1 688 variétés asiatiques à l'aide de 16 marqueurs isoenzymatiques, ont mis en évidence l'existence de six groupes variétaux (Glaszmann 1987). Cette classification n'a pas été démentie avec le développement des marqueurs moléculaires et du séquençage haut débit de 3000 génomes, mais seulement affinée. Le nombre de groupe a été réduit à 5 : les variétés qui, en 1987, constituaient les deux groupes de cette classification les moins nombreux et les plus atypiques

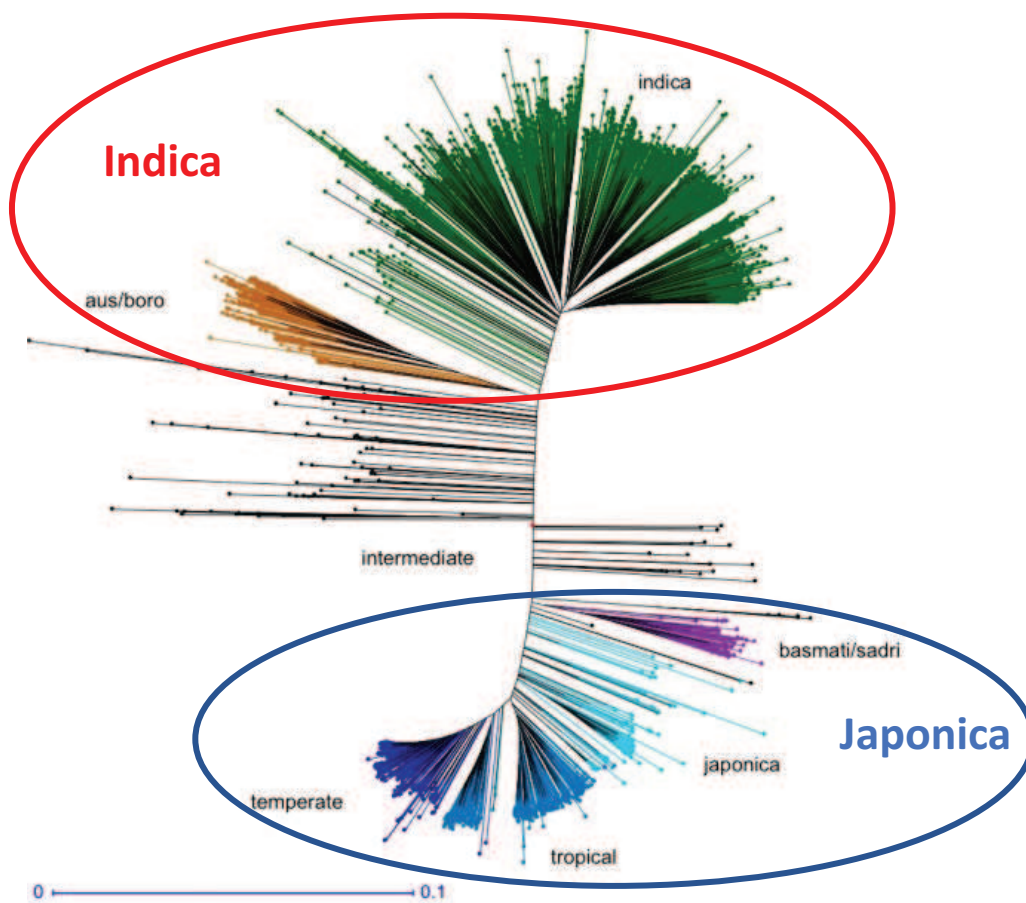


Figure 4 : Classification de 3 000 accessions de riz en 2 sous-espèces et 5 groupes variétaux (adaptation de Li et al. 2014).

(groupe III des Badhoia et groupe IV des Rayadas) ne font plus partie du panel d'accessions utilisé pour les études modernes de diversité et le groupe des japonica a été séparé en deux, les japonica tropicaux et les japonica tempérés. Cinq groupes variétaux – indica, aus/boro, basmati/sadri (ou aromatiques), japonica tropicaux et japonica tempérés – sont reconnus (Li et al. 2014) (Figure 4). Ces 5 groupes variétaux constituent deux sous-espèces, la sous-espèce *Indica* au sens large qui regroupe les indica et les aus/boro et la sous-espèce *Japonica* au sens large qui regroupe les japonica tempérés, les japonica tropicaux et les aromatiques (McCouch et al. 2016).

1.4. Sélection du riz

1.4.1. Méthodes de sélection et apport des biotechnologies et de la génomique

Les efforts des premiers cultivateurs ont permis de transformer les espèces sauvages en cultivars (Sang and Ge 2013) et cette domestication est à présent si poussée que les espèces domestiquées comme *Oryza sativa* ne sont plus capables de survivre à l'état sauvage (CGIAR 2013).

A la fin du 19^e siècle, l'homme réalise les premiers croisements de parents choisis. L'avancée des connaissances et les progrès technologiques ont depuis permis l'évolution des techniques de sélection. Les connaissances en génomique ont ensuite permis d'accélérer le travail d'amélioration marqué par les débuts de la sélection assistée par marqueur. Les premières stations d'amélioration variétale du riz ont vu le jour en Chine, en Inde et au Japon au début du XX^{ème} siècle (Khush 1987).

Les années 60 sont marquées par la révolution verte qui va favoriser la mécanisation, l'irrigation contrôlée, l'utilisation d'engrais et de pesticides et l'utilisation de semences améliorées. La révolution verte a permis une augmentation spectaculaire du rendement, symbolisée par la vulgarisation en 1966 par l'IRRI (*International Rice Research Institute*) d'une variété semi-naine (*semidwarf*) appelée IR8 (Khush 1987). Le rendement moyen mondial du riz en 1960 était d'environ 2 t / ha. Étonnamment, en seulement 40 ans, à mesure que les variétés et les techniques issues de la révolution verte se sont répandues, le rendement moyen a doublé, atteignant 4 t / ha en l'an 2000. En conditions irriguées intensives et en saison sèche, le rendement du riz peut atteindre jusqu'à 10 t/ha. Intensification et culture mono-variétale exposant les variétés à de nombreuses parasites et maladies, et donc d'avoir recours à des pesticides et des fongicides pour protéger les cultures, les années 70-80 ont été consacrées à l'intégration dans les nouvelles variétés de résistances génétiques, pouvant parfois avoir comme source des espèces sauvages apparentées. Les variétés emblématiques issues de ces travaux sont IR36, IR64 et IR72 que l'on a baptisées « méga-variétés » du fait qu'elles ont

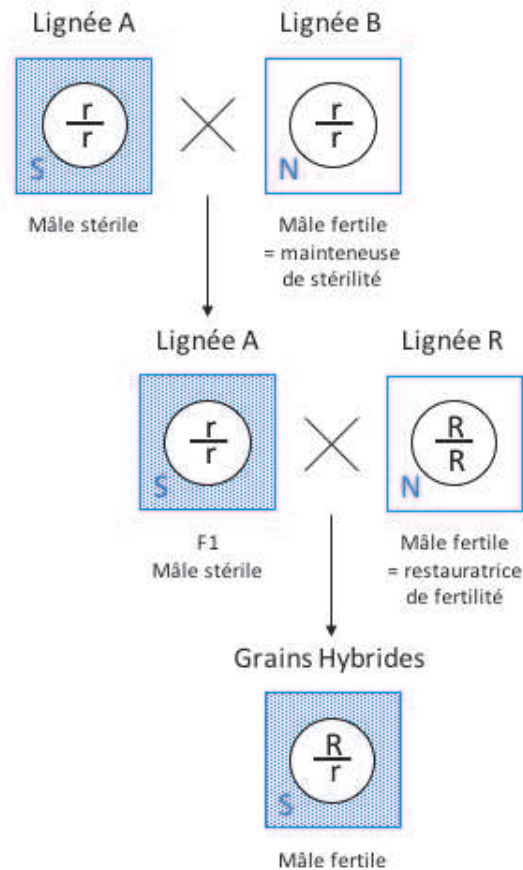


Figure 5 : Hérité de la stérilité mâle à contrôle géno-cytoplasmique.

Utilisation du système CMS (pour *Cytoplasmic Male Sterility*) dans les schémas de production d'hybrides F1. Ce système utilise 3 lignées distinctes : A, B et R. La lignée A est la lignée CMS. Elle porte les gènes de stérilité mâle géno-cytoplasmique c'est à dire qu'au niveau de l'ADN mitochondrial, elle porte des gènes responsables de la stérilité mâle (cytoplasme S pour stérilisant) et au niveau du noyau, elle porte les gènes récessifs de restauration de la fertilité (rr). Les plantes CMS et donc la lignée A ne produisent pas de pollen. La lignée B est aussi appelée lignée mainteneuse. Elle a un cytoplasme normal (N), du pollen fertile et possède les gènes récessifs de restauration de la fertilité pour maintenir le CMS (rr). Enfin la lignée R ou lignée restauratrice est la lignée qui possède les gènes dominant de restauration de la fertilité (RR) et un cytoplasme normal (N). La lignée R, lorsqu'elle est croisée avec la lignée A, permet de produire des grains hybrides (Rr) fertiles qui ont un cytoplasme stérilisant (S) .

couvert à elles-seules plusieurs dizaines de millions d'hectares en Asie. Depuis une quinzaine d'années, l'IRRI réalise un travail d'amélioration génétique plus amont laissant aux programmes nationaux le soin de développer des cultivars mieux adaptés à leurs contextes agro-environnementaux spécifiques.

Plus récemment, les biotechnologies ont été intégrées dans les programmes de sélection. Ce sont des outils supplémentaires à la disposition du sélectionneur pour repousser certaines limites rencontrées lors de l'utilisation des méthodes classiques d'amélioration des plantes.

1.4.2. Hétérosis et hybrides F1

Les deux espèces cultivées étant autogames, les variétés modernes sont en général des lignées pures. La sélection se fait par méthode généalogique après une phase de création de variabilité (hybridation dirigée, rétrocroisements, sélection récurrente ...). Dans les années 70, les sélectionneurs ont cherché d'autres pistes pour augmenter le rendement et se sont intéressés à l'hétérosis qui peut conduire à une augmentation du rendement de 20 à 30% chez le riz. Les Chinois ont été les premiers à développer et vulgariser des riz hybrides F1 après la découverte en 1973 d'un premier système de stérilité mâle géno-cytoplasmique (CMS : *cytoplasmic male sterility*), reposant sur une interaction entre gènes mitochondriaux et gènes nucléaires. Les lignées males stériles peuvent être divisées en 3 groupes : Wild Abortive (WA), Hong Lian (HL) et Boroll (BT) suivant leurs caractéristiques cytologiques et génétiques. Le système CMS WA a été le plus largement utilisé. Des études récentes suggèrent que le gène mitochondrial WA352 confère la stérilité en interagissant avec la protéine mitochondriale COX11 codée au niveau du noyau. La stérilité induite par WA352 peut être supprimée en utilisant deux gènes restaurateurs de fertilité (Luo et al. 2013a). Le système CMS WA reste le plus utilisé car il est stable quelque soient les environnements.

Le système CMS comprend un jeu de trois types de lignées pour multiplier la lignée mâle stérile puis restaurer la fertilité dans l'hybride F1 (Figure 5). La lignée A est la lignée mâle stérile. Elle est donc utilisée comme femelle. C'est elle qui a le cytoplasme qui confère la stérilité mâle ([S]) lorsque le noyau porte le gène de stérilité. La lignée B est la lignée mainteneuse. Elle a un cytoplasme normal ([N]) et un noyau identique à celui de la lignée A. En croisement avec la lignée A, la lignée B permet de multiplier la lignée A et B s'entretient par autofécondation. Enfin, la lignée R est la lignée mâle restauratrice. Elle a un cytoplasme normal ([N]) et son noyau contient un ensemble de gènes qui permettent la restauration de la fertilité. Le plus connu de ces gènes est rf3 mais il y en a d'autres et leur mécanisme d'action reste encore à élucider (Figure 5).

Des systèmes de stérilité mâle géniques induits par la température ou la photopériode ont également été développés : TGMS (for: *temperature genetic male sterility*) ou PGMS (for :

photoperiod genetic male sterility). Ils présentent l'avantage de ne nécessiter que deux parents au lieu de trois mais sont très sensibles aux conditions environnementales.

O. sativa ayant la biologie florale d'une plante autogame, son aptitude à l'allogamie est limitée : il faut donc faciliter la fécondation croisée entre la lignée male stérile et la lignée restauratrice pour produire des semences F1 en quantités suffisantes pour que le système soit économiquement viable. Cette production est assurée au champ en intercalant les deux types de lignées et en procédant à une pollinisation assistée (secouage en Chine). Le coût en main d'oeuvre reste très important, même si des méthodes plus mécanisées pour le semis et la récolte ont été mises au point. Le coût des semences reste donc élevé, et le rachat de semences F1 à chaque récolte rend cette innovation peu accessible aux riziculteurs de subsistance, en dehors des pays où l'achat des semences est subventionné.

D'autre part, le développement des 3 lignées A, B et R prend du temps. Pour accélérer la sortie variétale sur le long terme, l'amélioration de populations parentales pour leur aptitude à la combinaison a été conduit par le Cirad et l'Embrapa (Entreprise brésilienne de recherche pour l'agriculture et l'élevage) en Amérique Latine : L'hybride BRSCIRAD 302 est le premier lancé par une entreprise publique sur ce continent. Cet hybride, fruit d'une longue coopération entre le Cirad et l'Embrapa, résulte du croisement entre deux lignées spécifiquement sélectionnées pour former des hybrides. Ce nouveau riz hybride présente, outre un rendement élevé caractéristique, une qualité de grain égale à celle des meilleures variétés du marché brésilien.

Un système innovant de production des semences, totalement mécanisé pour le semis et la récolte, a aussi été développé. Les semences de BRSCIRAD 302, bien que destinées à la culture irriguée, sont produites en conditions pluviales. Cela permet une diminution des coûts de production des semences et évite toute contamination par les riz rouges adventices très fréquents au Brésil. Des semences de BRSCIRAD 302 sont en phase de commercialisation dans l'état du Rio Grande Do Sul, le plus grand producteur de riz au Brésil.

1.5. Ressources et outils disponibles pour l'amélioration variétale

1.5.1. Diversité naturelle et induite

La diversité est essentielle pour la création de nouvelles variétés. Cette diversité peut provenir soit de la variation allélique naturelle, soit être induite par des agents mutagènes de façon ciblée ou aléatoire. Alors que la diversité naturelle s'est enrichie pas à pas au cours de l'histoire évolutive des espèces puis de la domestication, la mutagenèse par agent chimique ou physique puis les biotechnologies (variation somaclonale, insertion au hasard d'élément transposable ou d'ADN-T pouvant être porteur d'étiquette d'activation) ont ensuite permis de créer de la diversité toujours de façon aléatoire. Depuis 2012, une mutagenèse dirigée par

Library	Institution(s)	Genetic background	Mutagen(s)	Number of lines produced	Number of lines seed increased for distribution	Number of lines with phenotype information
POSTECH Rice Insertion Database (RISD)	Pohang Univ. of Technology and Kyung Hee Univ., Korea	Dong Jin Hwayoung Kitaake	T-DNA GT, AT	105,000	75,000	nd
Rice Mutant Database (RMD)	Huazhong Agricultural Univ., Zhejiang Univ., Biotechnology Research Inst., China	Zhonghua 11 Zhonghua 15 Nipponbare	T-DNA ET <i>Tos17</i>	134,346 100,000	12,000	31,755
Taiwan Rice Insertion Mutant (TRIM)	Academia Sinica, Taiwan	Tainung 67	T-DNA AT	90,000	75,000	75,000
Oryza Tag Line (OTL) Génoplatte	CIRAD-INRA-IRD-CNRS, France	Nipponbare	T-DNA ET (+Ds) <i>Tos17</i>	31,000	25,000	25,000
Shanghai Insertion Population (ShIP)	SIPPE, China	Zhong hua 11	T-DNA ET	10,000	nd	nd
<i>Tos17</i> insertion database	Nat. Inst. of Agrobiological Sciences (NIAS), Japan	Nipponbare	<i>Tos17</i>	50,000	50,000	50,000
CSIRO	CSIRO Plant Industry (Australia)	Nipponbare	Ac/Ds GT/ET	5,000	nd	nd
EU OSTID	European Consortium	Nipponbare	Ac/Ds ET	5,000	5,000	nd
Temasek Ds	Temasek Life Sciences, Singapore	Nipponbare	Ac/Ds GT	20,000	nd	nd
UC Davis Ac/Ds and En/Spm populations	UC Davis, USA	Nipponbare	Ac/Ds GT En/Spm Ac/Ds AT	20,000	nd	nd

Tableau 1 : Liste des collections de mutants de riz (Droc et al. 2013).

l'homologie avec un ADN de réparation devient possible grâce à l'utilisation de la technologie CRISPR/Cas9 qui permet l'induction de coupures double brin à une haute fréquence à un site déterminé du génome (Jinek et al. 2012).

1.5.1.1. Des milliers de variétés de riz

La diversité génétique du riz est considérable avec plus de 150 000 variétés cultivées dans le monde. L'IRRI (International Rice Research Institute) entretient une collection de 130 000 accessions de riz cultivés (variétés modernes, traditionnelles et anciennes) et 5000 accessions d'espèces sauvages apparentées (International-rice-genebank, IRRI) (De Guzman 2016). Cette diversité provient de croisements naturels d'*O. sativa* avec des formes sauvages ou adventices d'*O. rufipogon* ou de croisements intra-sativa combinés à la sélection naturelle et humaine depuis la domestication (Khush 1987).

1.5.1.2. La diversité induite : aléatoire ou ciblée

Conjointement au séquençage de Nipponbare, le génome de référence d'*O. sativa*, un collectif international de chercheurs a développé des collections de mutants à la fin des années 1990 pour relever un autre défi, celui de découvrir la fonction des 39 000 gènes prédits chez Nipponbare (Hirochika et al. 2004; Kawahara et al. 2013). Pour étudier la fonction d'un gène, le moyen le plus direct est d'inactiver son expression et d'observer l'altération du phénotype. Chez le riz, on peut soit inactiver l'expression d'un gène cible par « gene silencing » ou « gene targeting » soit induire des mutations au hasard dans le génome en faisant de la mutagenèse chimique, physique ou insertionnelle. Ainsi, au cours des 20 dernières années, de nombreuses populations de mutants ont été créées (Tableau 1). Les collections de mutants d'insertion permettent aujourd'hui de disposer d'au moins un insert (ADN-T, Ac/Ds, En/Spm ou *Tos17*) dans 72% des gènes et jusqu'à trois inserts pour 22% d'entre eux (Droc et al. 2013). Des collections de RNAi ou de surexpression systématique de cDNA ont également été créées. Des collections de mutants obtenus par traitement chimique de semences à l'EMS ont notamment été produites : Chaque lignée possédant un grand nombre de lésions (une lésion chaque 300 pb), l'identification de mutants possédant une lésion dans le gène recherché se fait par TILLING ou plus récemment par séquençage de la région souhaitée à partir de pools d'ADN (Till et al. 2007). Ces collections de mutants peuvent également être utilisées pour faire de la génétique directe en réalisant des pools de descendances de mutants d'intérêt qui seront séquencés à haut débit. Le polymorphisme responsable de la mutation sera identifié parmi les autres car il sera commun au bulk de descendants présentant le phénotype, permettant l'isolement du gène causatif (Abe et al. 2012). Cette méthode dite MUTMAP a déjà permis d'identifier un nouveau gène de résistance à la pyriculariose et un gène de tolérance à la salinité (Takagi et al. 2013, 2015). Récemment, deux équipes ont rapporté la création de collections de lignées KO CRISPR/Cas9 (*Knock Out*) à l'échelle du génome chez le riz,

ouvrant de nouvelles perspectives pour la découverte de la fonction des gènes (Lu et al. 2017; Meng et al. 2017) .

1.5.2. Matériel à haute résolution génétique

Si la détection de QTL a longtemps utilisé des populations F2 ou des lignées recombinantes fixées (RILs) comme matériel, une façon plus rapide de localiser les QTLs a été de développer à partir de parents bien caractérisés des lignées de substitution ne possédant chacune qu'une petite région du génome d'un parent introgressé dans le fond génétique de l'autre parent (CSSLs ou chromosome segment substitution lines) : de nombreuses populations de CSSLs ont été développées notamment au Japon à partir des parents Nipponbare, Koshihiraki et Kasalath et ont permis d'accélérer la découverte et l'isolement de gènes (Miura et al. 2011). Des populations à haute résolution génétique comme les MAGICs (croisements pyramidaux successifs d'un groupe de parents au génome séquencés maximisant la recombinaison) ou les NAMs (croisements d'un groupe de parents avec un seul parent récurrent, tous génotypés à haut débit ou séquencés) permettent à la fois d'accélérer la cartographie de gènes d'intérêt et de dériver du matériel utilisable en pre-breeding (Fragoso et al. 2017; Raghavan et al. 2017).

1.5.3. Ressources génomiques et bio-informatiques

Le riz possède un génome diploïde simple ($2n=24$) de petite taille (390 Mb ; 0,9 pg ; 37 000 gènes) colinéaire avec celui des autres graminées (Devos and Gale 1997). C'est pourquoi en 1997, les chercheurs ont choisi le riz comme espèce "modèle" pour les monocotylédones (Sasaki and Burr 2000).

En 2002, une première version des séquences génomiques de deux accessions des deux sous espèces japonica (Nipponbare) et indica (93-11) ont été publiées simultanément (Goff et al. 2002; Yu et al. 2002). Par la suite, en 2005, le consortium international pour le séquençage du génome du riz (IRGSP) a publié une séquence de haute qualité du génome du cultivar Nipponbare (Sequencing Project 2005) qui a été annotée par une communauté de spécialistes des familles de gènes ou d'éléments transposables (RAPDB). Cette séquence demeure encore aujourd'hui la plus aboutie au sein de l'espèce *O. sativa* et parmi les plantes cultivées.

Dans le même temps, le « Rice Full-Length cDNA Consortium » a séquencé en 2003 plus de 28000 cDNA de Nipponbare et assigné une fonction à 76% des cDNA par recherche d'homologie de séquence. 64% de ces séquences ont des protéines homologues chez *Arabidopsis* (Kikuchi et al. 2003).

Ces travaux sur le génome ont non seulement permis une avancée considérable en génomique fonctionnelle mais ont aussi fourni un génome de référence de haute qualité indispensable aux nouvelles technologies de re-séquençage NGS (*Next-Generation Sequencing*) qui permettent de séquencer à bas coût des génomes entiers de riz (Guo et al.

2014). Aujourd'hui, le séquençage d'une variété est devenu accessible ce qui représente une nouvelle révolution pour l'amélioration des plantes. En 2014, 3 000 accessions de riz provenant de 89 pays ont été séquencées à une profondeur moyenne de séquençage de 14x (Li et al. 2014) et les SNPs découverts mis à disposition sur le web (<http://snp-seek.irri.org>). Une puce haute densité comportant 700 000 SNP a également été développée à l'Université de Cornell (McCouch et al. 2016). Les travaux de séquençage variétal dans les ressources génétiques du riz ont rapidement montré que se référer à un seul génome était insuffisant pour comprendre la diversité allélique et la variation naturelle. Des gènes d'intérêt agronomique comme les gènes *sub1*, *pstol1* (*pup1*) ou *snorkle* étaient absents de la séquence de Nipponbare. Les génomes d'accessions représentatives des autres groupes variétaux ont donc été séquencés et assemblés de novo. Les séquences de la variété indica IR64 et de la variété aus DJ128 ont ainsi été publiées (Schatz et al. 2014).

Le génome de l'espèce cultivée africaine *O. glaberrima* a été séquencé par plusieurs groupes, ce qui a permis de mettre en évidence que des variations naturelles dans les mêmes gènes de domestication que chez *O. sativa* proviennent de l'action de la sélection humaine.

Dans le cadre du projet I-OMAP, 18 génomes sauvages comprenant les 9 espèces appartenant au génome A et 9 autres espèces représentatives des 9 types de génomes (BB, CC, BBCC, CCDEE, FF, GG, KKLL, et HHJJ) ont été séquencés. Treize de ces séquences ont aujourd'hui été assemblées (Wing communication personnelle) (Figure 3).

1.5.4. Les outils d'ingénierie cellulaire et d'édition des génomes

Le riz a été transformé successivement par traitement chimique ou physique de protoplastes isolés, par suspensions cellulaires (Toriyama et al. 1988) par bombardement de microprojectiles sur des embryons immatures ou des cals dérivés d'embryons de grains matures (Christou et al. 1991), puis par co-culture de cals de grains matures ou d'embryons immatures avec *Agrobacterium tumefaciens* (Hiei et al. 1994). Cette dernière méthode est largement adoptée par les laboratoires car le transfert biologique de l'ADN-T est très efficace et conduit à des intégrations simples par rapport aux méthodes de transfert direct. Le riz de type japonica tempéré demeure la céréale chez laquelle l'efficacité de transformation est la plus élevée (5 événements de transformation par cal co-cultivé pour le génotype modèle Nipponbare). Des progrès significatifs ont été réalisés sur les riz de type indica comme IR64 (20% des embryons co-cultivés forment une plante transgénique). Cette efficacité a permis au riz d'être la seule plante à fleur chez laquelle des événements de ciblage génique ont été obtenus de façon répétée (Terada et al. 2002). Comme nous l'avons vu, cet outil permet le transfert de constructions de surexpression, d'inactivation, de localisation tissulaire ou subcellulaire et aujourd'hui de ciblage par les technologies d'édition du génome (méganucléases, Zinc Finger Nucléases, TALE nucléase, CRISPR/Cas9, CRISPR/Cpf1).

La technologie d'édition du génome CRISPR/Cas9 est aujourd'hui largement adoptée comme en témoigne le nombre exponentiel de publications y faisant référence, y compris sur le riz. Les mécanismes biologiques de CRISPR/Cas9 ont été révélés en 2012 par Jennifer Doudna de l'Université de Californie et Emmanuelle Charpentier de l'Institut Max Planck à Berlin (Jinek et al. 2012). L'équipe de Zhang F. (Broad Institute, MIT) a de son côté montré la possibilité d'utiliser la technologie dans un système eucaryote.

Le mécanisme CRISPR/Cas9 confère à certaines bactéries une « immunité acquise » leur permettant de lutter contre les phages, les virus ou plasmides exogènes. Après une attaque virale, la bactérie va stocker des fragments de cet ADN viral invasif sous forme de "spacers" intercalés entre de petites séquences répétées sur son chromosome. Longtemps considérés comme du « junk DNA », les loci CRISPR ("*Clustered Regularly Interspaced Short Palindromic Repeats*") contiennent donc une alternance de courtes séquences identiques répétées et de séquences "spacers" spécifiques provenant des agresseurs qui sont, en fait, une « bibliothèque » de leurs signatures moléculaires. Si la bactérie est de nouveau attaquée par ce même virus, elle va transcrire cette région qui va être découpée en fragments, les petites séquences répétées étant reconnues par un tracrRNA bactérien transcrit simultanément, qui est va être reconnu par la nucléase Cas9. La partie spécifique (signature moléculaire) ou crRNA va ensuite guider la Cas9 jusqu'au nouvel ADN viral complémentaire, s'y appairer et le couper. Pour que le découpage de l'ADN viral commence, il faut que le site reconnu par le CRISPR possède en amont un motif particulier, le Protospacer Adjacent Motif (PAM) -NGG pour la Cas9. La protéine Cas9 se fixe alors sur l'ADN cible et effectue une coupure double brin de l'ADN. Le découpage de l'ADN viral entraîne sa dégradation. La protéine Cas9 fait partie de la famille des nucléases, et les chercheurs ont découvert que cette protéine peut également couper toute séquence d'ADN à un lieu précis (cible) reposant sur la complémentarité des bases entre l'ARN guide et de l'ADN cible. Il suffit donc d'un ARN guide pour que le système CRISPR/Cas9 coupe l'ADN cible à condition qu'il y ait un site PAM en amont de la cible. La CDB d'ADN est ensuite généralement réparée par ligation des extrémités (NHEJ) mais la Cas9 continue à couper tant que le site de reconnaissance n'est pas muté. La réparation aboutit en général à des insertions de nucléotides A/T ou des délétions de nucléotides (de quelques paires de bases à une dizaine si la réparation procède par recherche de micro-homologies sur la séquence -MMEJ) qui modifient le cadre de lecture. L'expression de la protéine correspondante est abolie. En fournissant une matrice de réparation et en sélectionnant les événements de RH, cette méthode permet de faire également de la réelle édition du génome, allant du changement d'un nucléotide, jusqu'au remplacement de larges régions du génome(Zhou et al. 2014).

La technologie d'édition du génome est à présent basée sur l'utilisation d'une cassette exprimant un ARN guide (gRNA), qui va comporter à la fois la partie spécifique à la cible et la partie constante reconnue par la Cas9. La mise en œuvre est très simple. La seconde cassette

va être conçue pour exprimer une version de l'enzyme Cas9 spécifique de l'espèce. Sur le riz, la mutagenèse ciblée par CRISPR/Cas9 est opérationnelle depuis 2013. Des exemples de multiplexage ont été publiés (Miao et al. 2013b) ainsi que l'utilisation d'une nucléase alternative Cpf1 qui a des propriétés particulières, notamment de reconnaître un PAM différent de la Cas9 riche en T (TTN ou lieu de NGG) (Yin et al. 2017).

Ces nouvelles technologies d'édition du génome ont déjà permis d'améliorer la qualité aromatique du riz (KO de *Osbadh2*) ou le format du grain (KO simultané de 3 gènes).

Des tentatives de remplacement d'allèle sont en cours. Le seul exemple publié reste celui du remplacement de nucléotides sur la séquence du gène d'acétolactate synthase du riz conférant une tolérance à l'herbicide chlorsulfuron, les cellules éditées devenant sélectionnables par l'herbicide. Une autre technologie, basée sur l'utilisation d'une Cas9 nickase (d/nCas9) fusionnée à une cytidine déaminase, permet la correction génique sans induction de CDB (conversion d'une cytidine en uridine permettant le changement de G/C vers A/T). Cette technologie a notamment été utilisée sur le riz pour convertir l'allèle japonica du transporteur de nitrate *NRT1.1b* en allèle indica, qui améliore l'efficacité d'utilisation de l'azote.

La technologie CRISPR/Cas9 qui remplace d'ores et déjà l'utilisation des mutants d'insertion pour les études d'analyse fonctionnelle des gènes sera sans doute utilisée dans le futur dans les programmes d'amélioration des plantes, où elle permettra de remplacer n'importe quel allèle ou combinaison d'allèles par un/une autre plus favorable.

1.6. Biologie du riz

Le riz cultivé est une plante annuelle. Le cycle de biologie du riz, à l'instar de celui des autres céréales s'opère en trois phases essentielles : 1) une phase végétative, allant de la germination à l'initiation des primordiaux floraux ; 2) une phase reproductive, qui va de cette initiation paniculaire à la pollinisation ; et 3) une phase de maturation du grain (Figure 6).

1.6.1. Phase végétative

Il convient de noter que la durée de la phase végétative varie significativement selon les variétés. Les variétés sont d'ailleurs parfois classées selon leur degré de précocité c'est à dire en fonction de la longueur du cycle végétatif car la phase reproductrice, elle, varie très peu (8-10j). On parle alors de variétés très précoces lorsque le cycle complet est compris entre 90 et 100 jours (cas des aus), puis de variétés précoces, semi-précoces, moyennes (120 à 150 jours), tardives et très tardives si le cycle est supérieur à 210 jours (Nona bokra). Ce mode de classement, qui est pratiqué d'un point de vue agronomique, n'a cependant aucune valeur taxonomique.

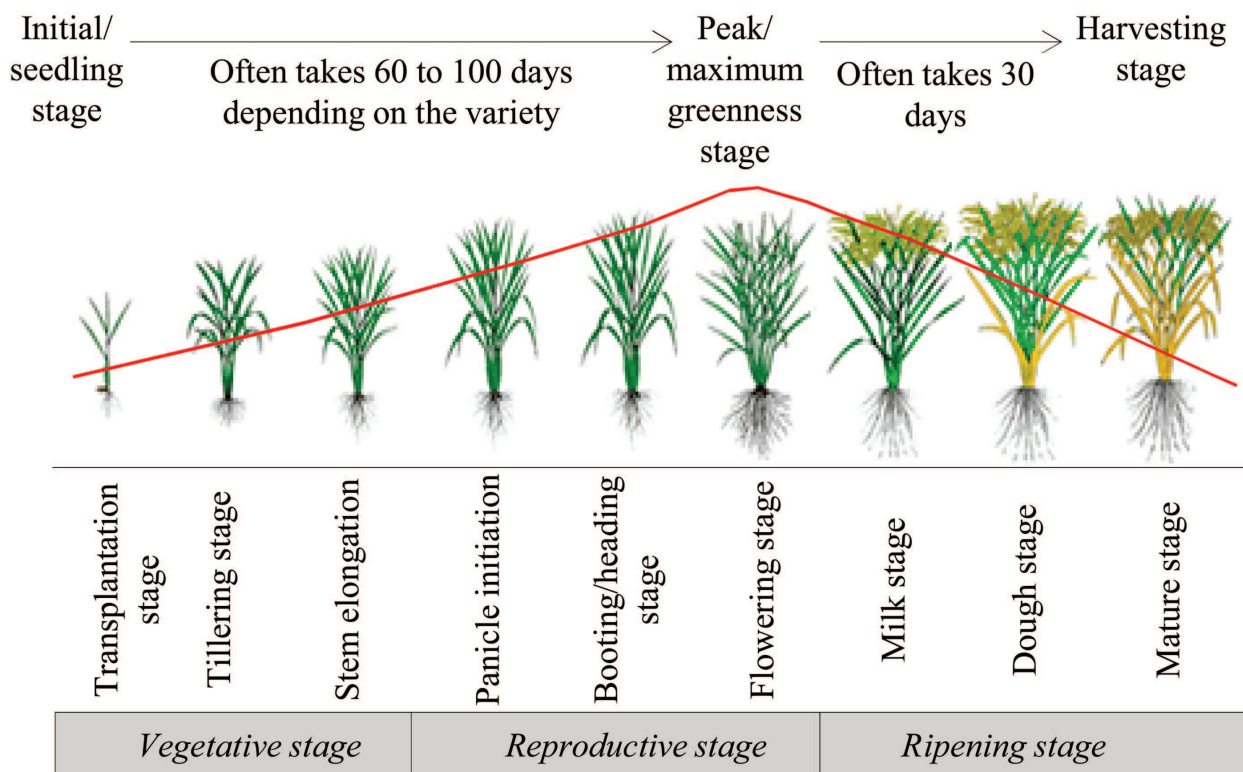


Figure 6 : Stades de développement du riz (d'après l'IRRI).

Le cycle biologique du riz est découpé en trois phases : 1) une phase végétative, allant de la germination à l'initiation des primordiaux floraux ; 2) une phase reproductive, qui va de cette initiation paniculaire à la pollinisation ; et 3) une phase de maturation du grain.

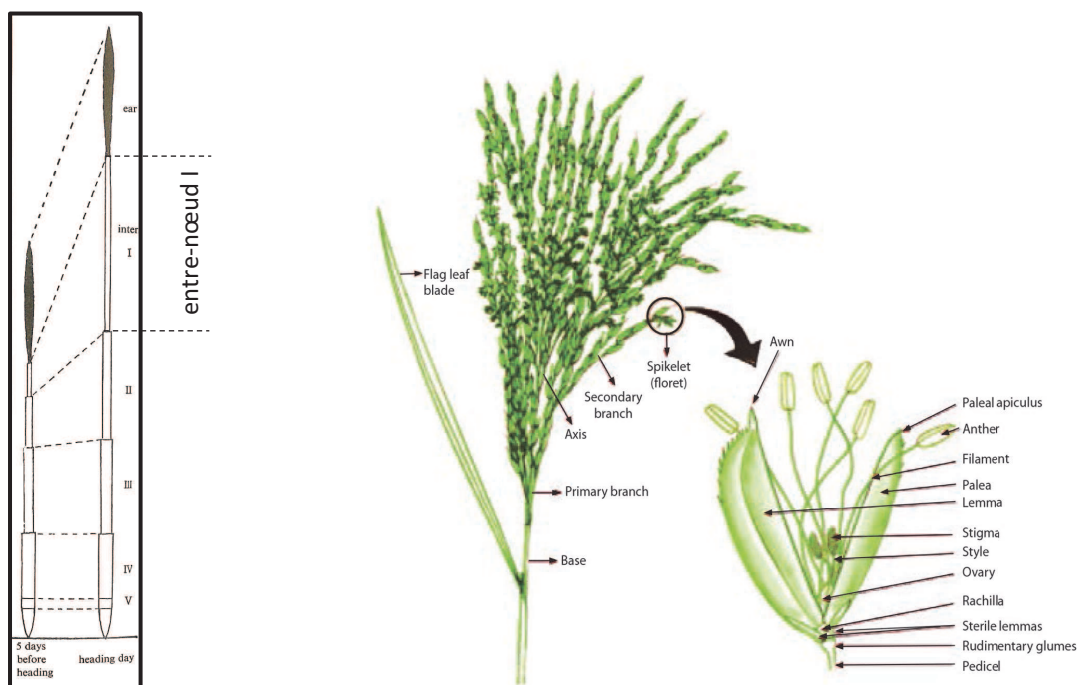


Figure 7 : Panicule de riz et épillet (d'après CGIAR 2013).

Gauche : Lors de la montaison, les entre-nœuds s'allongent jusqu'à l'épiaison. Une taille du premier entre-nœud (I à V) de moins de 1 cm permet de repérer le stade méiocyte. L'inflorescence de riz, ou panicule, est composée de plusieurs épillets. Chaque épillet comprend six anthères et un ovaire surmonté de deux stigmates plumeux.

Le riz germe en 3-4 jours et atteint le stade 7 feuilles en 21 jours. C'est pendant cette phase que la jeune plante peut être repiquée. Puis arrive la phase de tallage, il y a émission de jeunes tiges à partir du plateau de tallage qui se situe à la base de la plante entre la partie aérienne et racinaire. A la fin du stade végétatif, le plante atteint un nombre maximum de talles qui varie de 4 à 30 selon les variétés et les conditions de culture (Figure 6).

1.6.2. Phase reproductive

La phase reproductive débute par l'initiation paniculaire au niveau des primordia floraux, puis la panicule s'élève dans la gaine foliaire (montaison) et sort de la gaine lors de l'émergence de la panicule (ou épiaison). C'est au cours de la montaison qu'a lieu la méiose dans les organes reproducteurs. Le stade méiocyte est corrélé à la distance de l'exertion de la gaine de la feuille paniculaire par rapport à la ligule de la dernière feuille engainante (Figure 7). Lorsque celui-ci mesure entre 0,5 et 1 cm alors les fleurs sont au bon stade pour faire des observations cytologiques de la méiose.

Une panicule est constituée de nombreux épillets (fleurs) et leur floraison s'étale en suivant un gradient basipète (du haut vers le bas de la panicule). La floraison d'une panicule entière dure une dizaine de jours et toutes les talles n'émettent pas leurs panicules en même temps. Ainsi la floraison et la fécondation d'une plante dure plusieurs jours (Figure 7). L'espèce *O. sativa* est autogame. L'épillet est constitué de 6 anthères et d'un ovaire surmonté de deux stigmates. A l'anthèse, les filets des étamines s'allongent et les glumelles s'ouvrent pour laisser sortir les anthères. Cette étape qui correspond à la fécondation ne dure que quelques heures (vers midi heure solaire). Après la pollinisation, les glumelles se referment et viennent pincer les filets laissant les 6 anthères à l'extérieur de l'épillet. Cinq jours après fécondation, le grain atteint le stade laiteux et 35j après fécondation le grain est mature et peut être récolté puis séché. Les étapes la gamétogenèse mâle et femelle (formation du pollen et de l'ovule) seront explicitées dans le paragraphe suivant ainsi que la reproduction (§ 2.1).

Les variétés d'*O sativa*, qui est une espèce annuelle, ont la particularité de pouvoir se multiplier à nouveau en fin de cycle (après récolte) en produisant des bourgeons après recépage (ratooning).

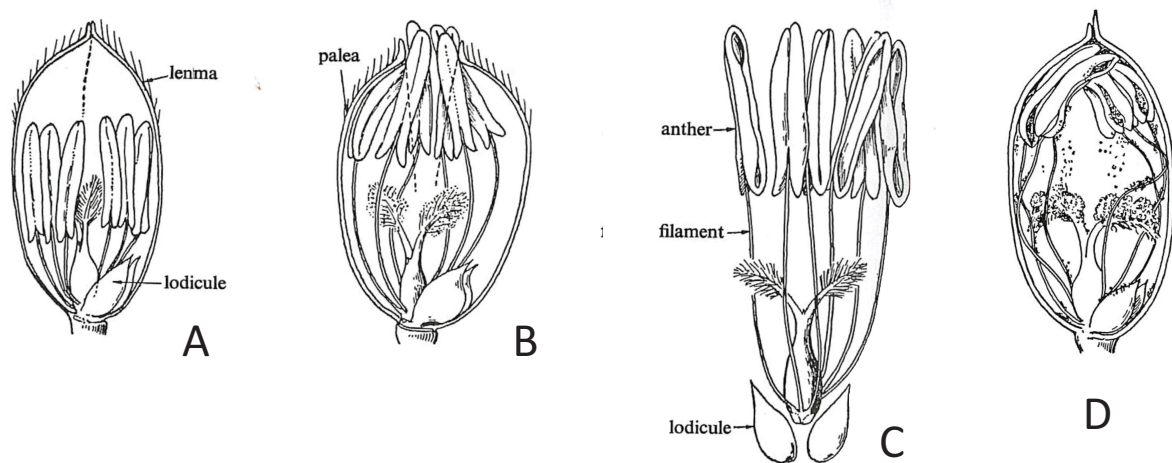


Figure 8 : Développement de la fleur de riz au moment de l'anthèse (d'après Hoshikaw, 1989).

A- Juste avant l'anthèse ; les filaments commencent à s'allonger

B- Début de l'anthèse ; les lodicules se gonflent et l'épillet s'ouvre. Les filaments continuent à s'allonger et les anthères commencent à sortir.

C- Epillet sans les glumelles pour montrer le détail des anthères, des filaments et des lodicules.

D- Illustration de la pollinisation dans un épillet un jour de pluie (cléistogamie).

2. REPRODUCTION SEXUEE ET APOMIXIE

La reproduction se définit comme l'ensemble des processus par lesquels une espèce se perpétue et donne naissance à un nouvel organisme. La reproduction peut être de deux types : sexuée ou asexuée.

La reproduction sexuée est très largement répandue chez les eucaryotes. Pourtant ce mode de reproduction représentant un coût énergétique important pour l'espèce. Il est donc difficile d'en prouver les avantages évolutifs (Hörandl 2009). Moins de 1% des Angiospermes seraient strictement asexués (Whitton et al. 2008). Le terme apomixie a été longtemps utilisé pour décrire l'ensemble des modes de reproduction asexuée, incluant la multiplication végétative. Aujourd'hui, l'apomixie se limite généralement aux modes de reproduction asexuée par graines, sauf dans les cas d'embryonie adventive pour lequel un passage par un stade gamétophytique non réduit s'effectue. Les plantes apomictiques ont la capacité de produire des grains qui sont génétiquement identiques à la plante mère. Pour une plante, l'apomixie représente la capacité à contourner les aspects fondamentaux de la reproduction sexuée que sont la méiose et la fécondation (Hand et al. 2014).

2.1. La reproduction sexuée et la double fécondation

Chez les angiospermes, les organes reproducteurs sont condensés en une fleur. La fleur et le fruit sont caractéristiques de cet embranchement qui comprend les monocotylédones et les dicotylédones. La formation de graines chez les angiospermes nécessite le passage de la phase végétative à la phase reproductive et se manifeste par la formation d'une fleur. Chez le riz, cette fleur est cléistogame et hermaphrodite en renfermant les organes mâles et femelles, c'est-à-dire les anthères et l'ovaire renfermant l'ovule (Figure 8). Le gamétophyte mâle (le grain de pollen) se développe à partir des anthères et le gamétophyte femelle (le sac embryonnaire) se développe à partir de l'ovule.

2.1.1. La gamétogenèse mâle et femelle dans la jeune fleur

C'est dans le tout jeune bouton floral des Angiospermes que se déroule la méiose, étape essentielle de la reproduction sexuée.

Les méiocytes mâles se forment à partir d'un massif de cellules sporogènes dans les étamines. Ces cellules sporogènes ou archéspores évoluent en cellule mère du pollen (méiocytes mâles) et vont subir la méiose à l'issue de laquelle sera formée une tétrade à 4 microspores (cellules haploïdes). La maturation de la tétrade donnera 4 grains de pollen. Un grain de pollen résulte de la mitose d'une microspore et est constitué de trois noyaux haploïdes, un noyau végétatif et deux noyaux reproducteurs (Figure 9).

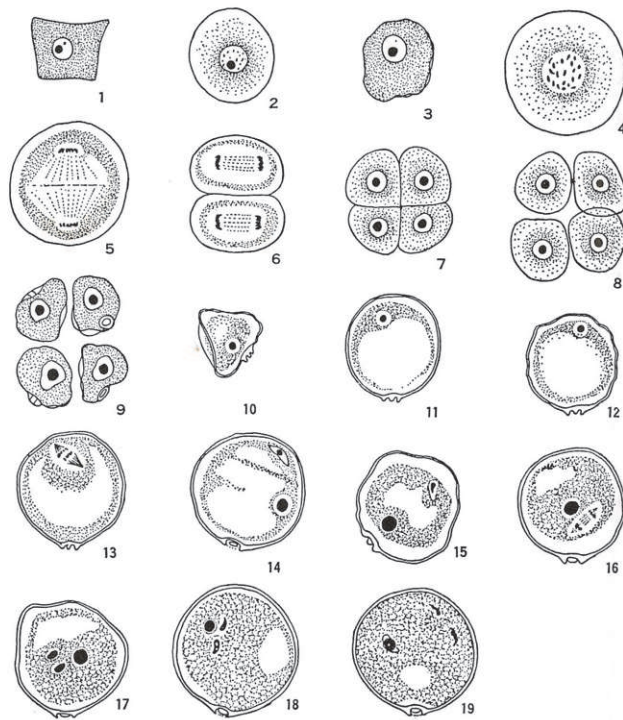


Figure 9 : Gamétogenèse mâle, formation du grain de pollen chez le riz (d'après Hoshikaw, 1989).

1 : Archéspore ; 2. Cellule mère de pollen ; 3. Stade de contraction de la cellule ; 4. Phase S de la première division de méiose ; 5-7. Divisions de méiose ; 8. Tétrade et séparation des 4 microspores ; 9-10. stades de différenciation des microspores ; 11-12. Fin de la différenciation des microspores, formation de l'enveloppe nucléaire et contraction ; 13. première division mitotique ; 14-15. Stade à deux noyaux ; 16. Seconde division mitotique du noyau reproducteur ; 17-19. Maturation du grain de pollen à 3 noyaux (2 reproducteurs et 1 végétatif).

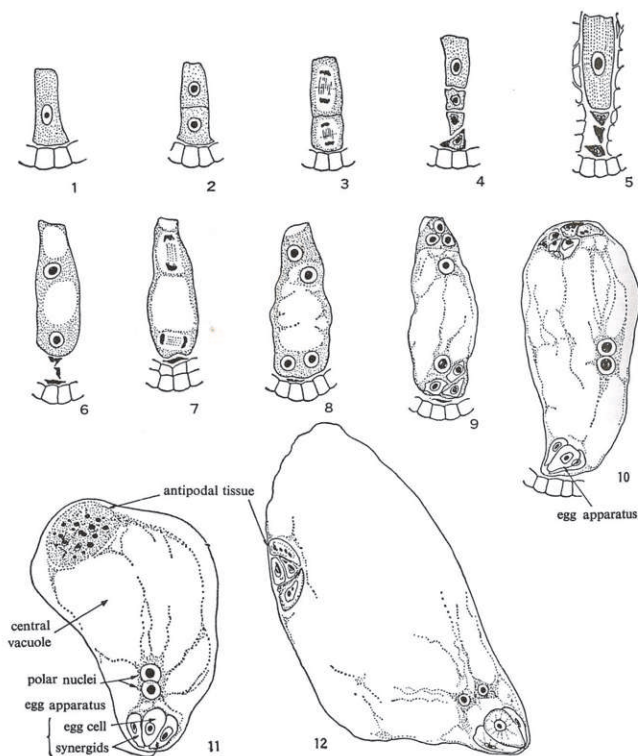


Figure 10 : Gamétogenèse femelle, formation du sac embryonnaire chez le riz (d'après Hoshikaw, 1989).

1. Cellule mère du sac embryonnaire ; 2. Première division de méiose ; 3. Seconde division de méiose ; 4. Stade tétrade du sac embryonnaire ; 5. Trois des cellules dégénèrent et une cellule devient la cellule mère de la mégaspore (CMM) ; 6. Première mitose, stade à 2 noyaux ; 7-8. Seconde mitose, stade à 4 noyaux ; 9. Troisième mitose, cellule à 8 noyaux ; 10-12. Différenciation et croissance du sac embryonnaire.

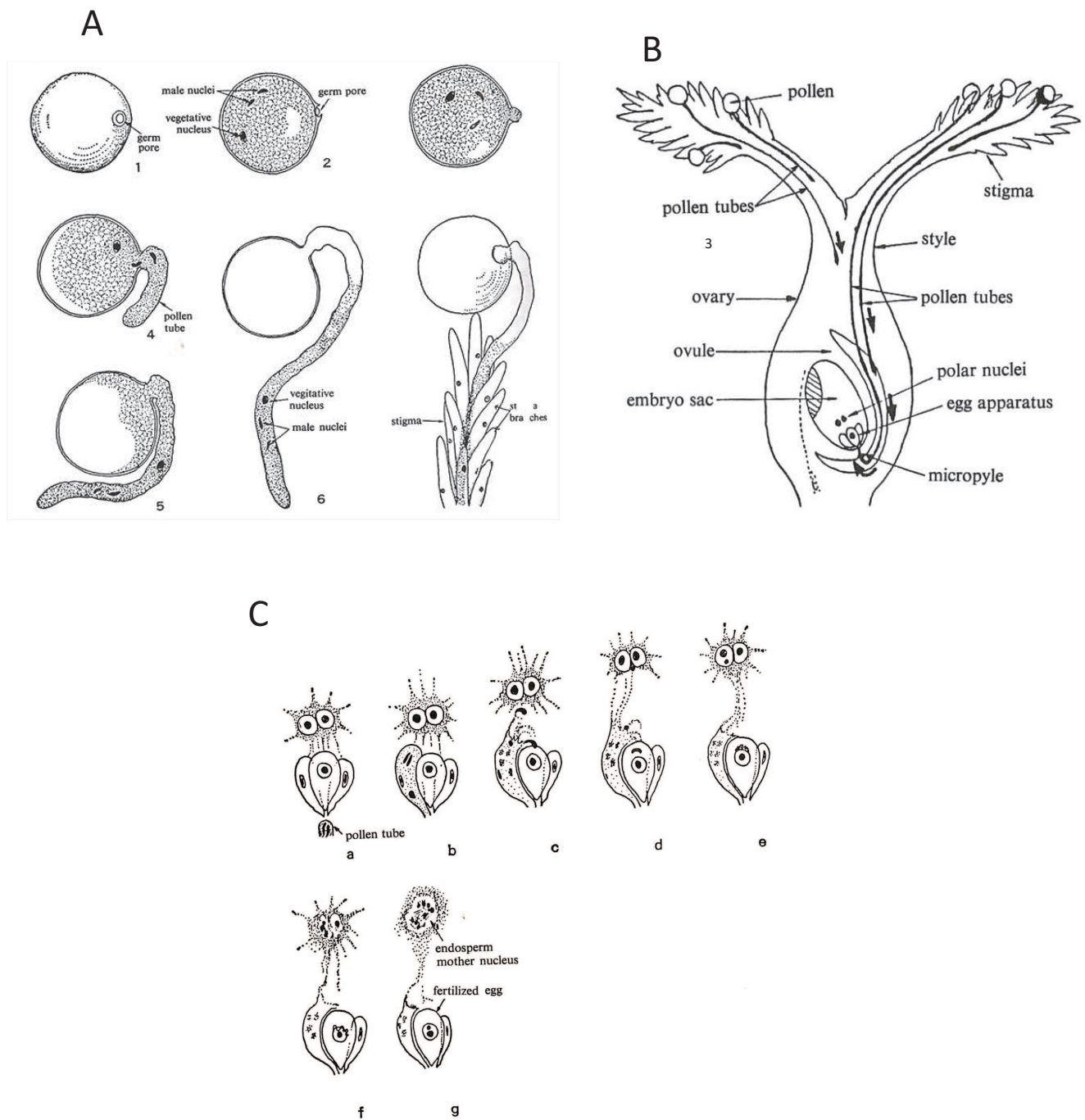


Figure 11 : La double fécondation chez le riz (d'après Hoshikaw, 1989).

A- Germination du grain de pollen et pénétration du tube pollinique dans le stigmate.

B- Elongation du type pollinique dans le style vers l'ovaire.

C- Processus de fécondation

a- Avant la fécondation ; b. Pénétration du tube pollinique ; c. Noyau reproducteur qui se décharge dans la cellule femelle ; d-g. suite de la fécondation.

De même dans l'ovule, une cellule mère du sac embryonnaire (méiocyte femelle) va subir la méiose pour donner 4 macrospores haploïdes, c'est le stade tétrade du sac embryonnaire. Puis 3 de ces 4 cellules avortent et la cellule restante devient la cellule mère de la mégaspore (CMM ou cellule du sac embryonnaire aussi appelée gamétophyte femelle). Cette cellule mère de la mégaspore va subir 3 mitoses successives pour former les 8 noyaux du sac embryonnaire (Figure 10). Chez la plupart des plantes à fleur, le sac embryonnaire est organisé en une structure appelée polygone qui contient 8 cellules de 4 types différents. En partant du côté micropylaire, ces 4 types cellulaires sont : la cellule œuf, qui donnera l'embryon ; les deux synergides, qui sont importantes pour l'attraction du tube pollinique ; deux cellules polaires au centre (ou chez *A. thaliana* une cellule centrale dicaryotique qui résulte de la fusion des 2 noyaux polaires), qui sont importantes pour la formation de l'albumen et trois cellules antipodes à l'opposé au micropyle dont la fonction est inconnue. Soit un total de 7 cellules et de 8 noyaux. (Figure 10) (Sundaresan and Alandete-Saez 2010).

2.1.2. La double fécondation et la formation de grain

Le terme de double fécondation est un abus de langage puisqu'il s'agit en fait d'une fécondation et d'une fusion de noyaux.

La maturation des anthères exerce une pression à l'intérieur de l'épillet qui a pour conséquence d'entrouvrir les deux glumelles. Ce stade est très furtif puisque l'épillet ne sera ouvert qu'une à deux heures au moment précis où va avoir lieu la fécondation. L'ensoleillement, la chaleur et l'humidité sont à ce stade des facteurs essentiels. C'est dans ces conditions que le sac pollinique va exploser et libérer un nuage de grains de pollen. Parmi eux, certains vont venir se fixer sur le stigmate plumeux de l'ovaire. Le grain de pollen germe à une température comprise entre 30 et 35°C et produit un tube pollinique. Le tube pénètre entre les cellules synergides et permet la migration des deux noyaux reproducteurs dans le sac embryonnaire alors que le noyau végétatif reste dans le tube pollinique. L'un des deux noyaux reproducteurs féconde l'oosphère et donne l'embryon à l'origine de la future plante. Le second noyau reproducteur fusionne avec les deux cellules polaires du sac embryonnaire et forme ainsi un tissu triploïde appelé albumen, qui fournira les réserves à l'embryon. Ces deux fusions simultanées constituent la double fécondation (Figure 11).

Le risque d'allo-fécondation est rare, chez les riz cultivés ; il n'est que de 1% en plein champ mais peut parfois atteindre des valeurs plus élevées dans des conditions climatiques particulières. Il existe dans le genre *Oryza* des espèces sauvages allogames (par exemple, *O. longistaminata*) qui ont une biologie florale adaptée à la fécondation croisée. Comme son nom l'indique, les longs filaments des anthères et les longs stigmates leur permettent une exsertion hors des épillets.

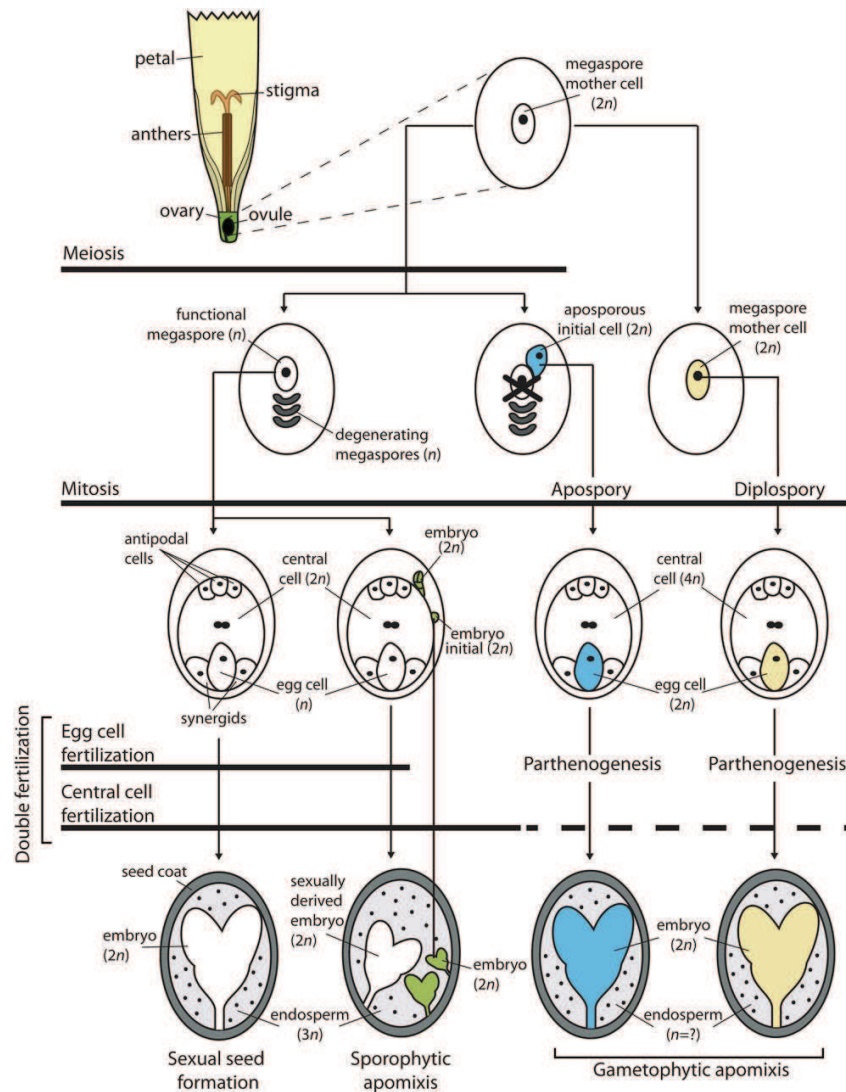


Figure 12 : Mécanismes de développement de semences par voie sexuée ou apomictique (Hand et al. 2014).

Le processus de développement des graines se produit à l'intérieur de l'ovule de la fleur qui chez l'espèce *Hieracium*, représenté dans cette figure, est une fleur unique. Ce schéma illustre les principales différences qui existent entre les voies de développement des graines pour la formation de graines par voie sexuée et les mécanismes apomictiques de l'apomixie sporophytique et gamétophytique. La méiose, la mitose et la double fécondation constituent les principaux composants de la voie de formation des graines. Les flèches passant à travers chacun de ces composants représentent l'implication d'un composant donné dans une voie particulière. Dans le processus d'apomixie gamétophytique, la formation de sac embryonnaire peut se produire soit par apospory soit par diplospory, qui se distinguent par différentes cellules précurseurs du sac embryonnaire. Dans l'apomixie gamétophytique, la formation d'embryon est initiée en l'absence de fécondation (parthénogenèse); cependant, la formation d'endosperme peut se produire avec ou sans fertilisation, ce qui est représenté par une ligne pointillée. Le niveau relatif de ploïdie des cellules (n) est suivi pour divers composants dans chaque voie. Le niveau de ploïdie de l'albumen formé par l'apomixie gamétophytique est variable, dépend d'un certain nombre de facteurs et est donc représenté par un point d'interrogation. Dans la voie de l'aposporie représenté, la voie sexuelle se termine une fois que la cellule initiale de l'apospore subit la mitose. Différentes couleurs suivent les cellules précurseurs qui forment l'embryon pour chaque voie : sexuelle (blanche), apomixie sporophytique (vert), diplospory (jaune), et apospory (bleu).

2.2. La reproduction asexuée clonale par grain : l'apomixie

Cette faculté de produire une descendance génétiquement identique à la plante mère est une valeur ajoutée pour l'agriculture. En effet, l'apomixie pourrait permettre de fixer les génotypes favorables et les hybrides F1. L'apomixie est documentée dans plus de 120 genres d'angiospermes. Cependant aucune grande culture céréalière majeure n'est apomictique et les tentatives visant à introduire le caractère d'apomixie chez les plantes cultivées par croisement classique ont toutes été infructueuses (Savidan 2001).

2.2.1. Les différents types d'apomixie

Contrairement à la formation de graines par reproduction sexuée, l'apomixie peut avoir lieu suivant différents mécanismes qui partagent trois caractéristiques communes à savoir (1) le contournement de la méiose durant la formation du sac embryonnaire (l'apoméiose), (2) le développement de l'embryon indépendamment de la pollinisation (parthénogenèse), (3) la formation d'un albumen viable avec ou sans fécondation par la cellule spermatique (Koltunow & Grossniklaus, 2003). La formation d'un œuf à partir d'une cellule maternelle diploïde et son développement sans fécondation en embryon font que la descendance issue d'un développement apomictique est clonale et donc génétiquement identique au parent maternel.

Les mécanismes de l'apomixie sont historiquement subdivisés en deux catégories et classés soit comme sporophytique soit gamétophytique, selon que l'embryon se développe à partir d'une cellule du gamétophyte (sac embryonnaire) ou directement à partir de cellules somatiques diploïdes (sporophytes) dans l'ovule (Figure 12).

2.2.1.1. Apomixie sporophytique

Au cours de l'apomixie sporophytique, le sac embryonnaire et le gamétophyte vont poursuivre les étapes de la reproduction classique (méiose, mitoses, et double fécondation). Il en résulte une graine qui peut contenir plusieurs embryons dont un résulte de la reproduction sexuée normale (il peut se développer normalement et germer ou se développer partiellement et ne pas germer). Les autres embryons résultent d'apomixie sporophytique. Ce mécanisme est très commun chez le citrus. L'apomixie sporophytique n'a pas été étudiée d'un point de vue moléculaire mais apparaît complexe d'un point de vue génétique (Hand et al. 2014).

2.2.1.2. Apomixie gamétophytique

Dans le cas de l'apomixie gamétophytique, le sac embryonnaire se forme à partir d'une cellule diploïde qui, sans passer par la méiose, va se diviser par mitoses successives et former le sac embryonnaire. En fonction de la provenance de la cellule diploïde on distingue deux mécanismes, la diplosporie et l'aposporie (Figure 12). Dans le cas de la diplosporie, le précurseur est la Cellule Mère de Mégaspore (CMM) ou une cellule avec un programme

modifié qui se différencie en cellule mère de mégaspore. Cette cellule peut entrer dans la méiose et avorter le processus, ou immédiatement commencer la mitose. La diplosporie se retrouve avant tout chez les *Asteraceae* comme le pissenlit. En revanche, l'aposporie implique le développement du sac embryonnaire via la méiose et la mitose. La mitose, au lieu d'avoir lieu à partir d'une CMM, va se faire via une cellule somatique diploïde adjacente à la mégaspore. Cette cellule est nommée "cellule initiale de l'apospore". A l'issue des événements mitotiques de diplosporie et d'aposporie, même si le sac embryonnaire peut ou non être constitué des sept noyaux et de la cellule œuf, on observe, dans tous les cas, la formation d'une cellule œuf diploïde, des synergides et des cellules centrales elles-aussi diploïdes. Dans les deux cas, la cellule œuf diploïde va donner un embryon sans fécondation ; c'est ce que l'on appelle la parthénogenèse. Bien que ce soit rare, le développement de l'albumen peut lui-aussi avoir lieu sans fécondation de la cellule centrale. Ce processus est observé principalement chez la marguerite. Dans tous les cas d'apomixie, l'albumen, selon qu'il résulte ou non de la fusion des cellules centrales, avec ou non une cellule spermatique, va avoir un ratio génome maternel/paternel modifié. Dans un cas normal, le ratio est de 2:1 puisque 2 cellules centrales maternelles haploïdes fusionnent avec 1 cellule spermatique haploïde (albumen $3n$ avec la contribution de deux génomes maternels et un génome paternel). Dans le cas d'apomixie gamétophytique, l'albumen peut être $4n$ ou $5n$ si les deux cellules centrales sont fusionnées avec une cellule spermatique haploïde. Tous ces albumens ne sont pas viables et les plantes apomictiques qui ont besoin de la fécondation des cellules centrales pour que l'albumen se forme ont développé des stratégies diverses pour assurer la viabilité de la graine (Hand et al. 2014).

Nous reviendrons sur l'apomixie gamétophytique car la modification de la méiose, qui permet la formation de diplogamètes (mégaspore diploïde) puis le déclenchement parthénocarpique, permet de former un embryon clonal (Marimuthu et al. 2011).

2.2.2. Le contrôle génétique et épigénétique de l'apomixie

L'apomixie se produit naturellement dans plus de 400 espèces appartenant à plus de 35 familles, dont les *Poaceae*, suggérant ainsi que ce mode de reproduction clonale est apparu de nombreuses fois dans l'histoire de l'évolution des espèces et qu'il n'est sans doute pas lié à un événement génétique très complexe. Le maïs, le mil, le blé et le sorgho ont un ancêtre apomictique alors que le riz et la plupart des plantes cultivées qui ne sont pas des céréales n'en n'ont pas (Rutger 1992; Koltunow and Grossniklaus 2003).

Cependant les bases génétiques qui contrôlent l'apomixie restent difficile à élucider (Hand et al. 2014). Les tentatives d'introgession du caractère d'apomixie chez le maïs, à partir de son cousin sauvage *Tripsacum dactyloides*, ou chez le mil, en utilisant l'espèce sauvage africaine *Pennisetum squamulatum* au XX^{ème} siècle ont toutes échouées. A cette époque, on pensait que l'apomixie était héritée comme un seul locus dominant se trouvant sur un segment de

chromosome non recombinant (Savidan 2001; Barcaccia and Albertini 2013). Plus tard, l'utilisation de la mutagenèse gamma a permis d'obtenir des recombinants de cette région. L'analyse génétique des descendants de ces recombinants a montré que les trois composantes de l'apomixie, que sont l'apoméiose, la parthénogenèse et le développement de l'albumen sans fécondation, sont en fait contrôlées par plusieurs loci indépendants. Chez les espèces apomictiques, la formation de gamétophytes femelles non réduits pourrait s'expliquer par un ou deux loci dominants. Cependant, de multiples facteurs comme la ploïdie, les conditions environnementales ou l'épigénétique compliquent l'analyse de ce caractère. D'autres études montrent que les loci contrôlant l'apomixie pourraient se trouver dans une région génomique très conservée et non recombinante telle que la région hétérochromatique qui est riche en éléments transposables. Toutefois, aucun gène associé à ces loci n'a été décrit, à ce jour, comme contrôlant une quelconque composante de l'apomixie naturelle (Ronceret and Vielle-Calzada 2015).

Les mécanismes sous-jacents à l'hétérosis méritent aussi d'être étudiés car même si on parvient à faire de la reproduction clonale, parvenir à fixer l'hétérosis n'est pas certain. Ceci nécessite une compréhension approfondie des précurseurs cellulaires de la méiose, de la gamétogenèse, du processus de double fécondation, de l'embryogenèse et du développement de l'albumen.

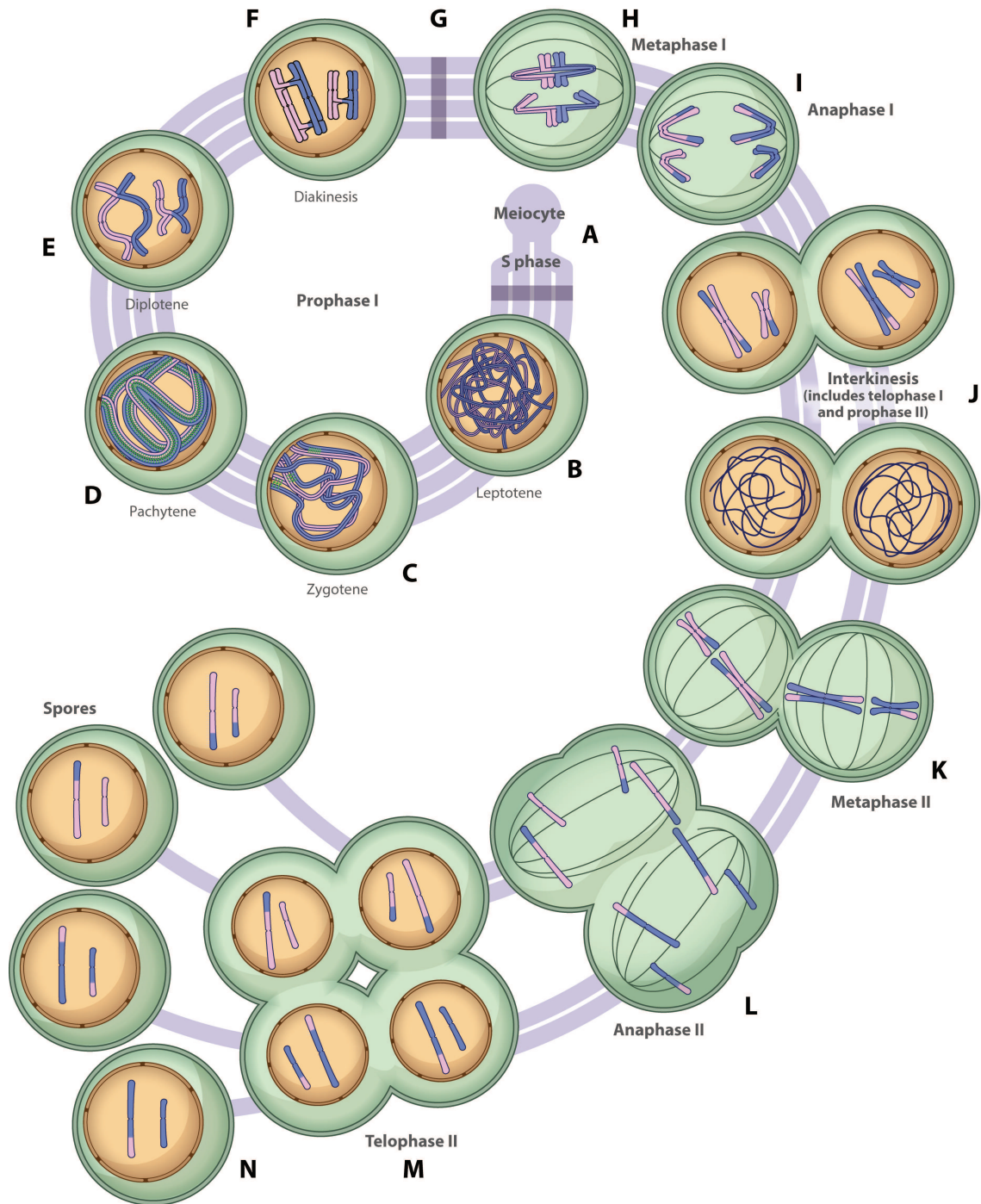


Figure 13. Déroulement de la méiose (Mercier et al. 2015).

A- Réplication préméiotique de l'ADN : les jeux de chromosomes parentaux sont constitués chacun de deux chromatides sœurs parfaitement identiques.

B à N- Deux divisions cellulaires successives composent la méiose, chacune décomposée en quatre phases distinctes : la Prophase, la Métaphase, l'Anaphase et la Télaphase.

B à J- Première division dite "réductionnelle" : les jeux de chromosomes homologues s'apparient, échangent de l'information génétique via la recombinaison méiotique, puis se séparent en deux cellules-filles.

K à N- Seconde division de méiose dite "équationnelle" : les chromatides sœurs de chaque chromosome se séparent donnant ainsi quatre cellules-filles qui formeront les gamètes.

3. MEIOSE ET RECOMBINAISON

La méiose est une division cellulaire très particulière à l'origine de la formation des gamètes. Contrairement à la mitose, qui a pour but de maintenir le niveau de ploïdie entre la cellule mère et les deux cellules filles, la méiose, grâce à deux divisions cellulaires successives, permet de produire 4 cellules filles haploïdes (n) à partir d'une cellule mère diploïde ($2n$). La fusion des gamètes mâle et femelle lors de la fécondation permet de rétablir le niveau de ploïdie de l'espèce ($2n$). Les deux divisions méiotiques sont fondamentalement différentes du point de vue du comportement des chromosomes : la première division permet la séparation des chromosomes homologues alors que la deuxième permet la séparation des chromatides sœurs (Figure 13). Ces deux types d'événements impliquent des modifications importantes de la régulation de la ségrégation des chromosomes.

La méiose assure également un brassage génétique entre chromosomes maternels et paternels grâce à la recombinaison méiotique d'une part et à la ségrégation des chromosomes parentaux dans les cellules filles d'autre part.

La méiose est constituée de deux divisions cellulaires successives nommées méiose I (MI) et méiose II (MII) précédées par un seul événement de réplication de l'ADN appelée réplication pré-méiotique. La première division de méiose est dite réductionnelle car elle permet la ségrégation des chromosomes homologues parentaux. La seconde division est équationnelle et permet la séparation des chromatides sœurs. Chaque division cellulaire est communément découpée en 4 phases identifiables par la structure des chromosomes en cytogénétique : la Prophase, la Métaphase, l'Anaphase et la Télophase (Figure 13) (Mercier et al. 2015).

Le terme recombinaison homologue méiotique (homologous recombination, HR) regroupe les mécanismes de réparation de cassures double-brin de l'ADN (CDB ou *double strand breaks*, DSB) par échange de matériel génétique entre séquences homologues. Il existe deux grands types d'événements : les crossing-overs (COs), qui sont des échanges réciproques d'ADN ; et les NCOs (non-crossing-overs), chez lesquels la cassure double brin est réparée localement, sans échange réciproque d'ADN de part et d'autre de la lésion. S'il existe du polymorphisme génétique, ces COs et NCOs peuvent être associés à des conversions géniques.

Bien que la reproduction représente un coût énergétique élevé d'un point de vu physiologique, elle assure un rôle essentiel pour l'espèce puisqu'elle permet la création de diversité et donc l'adaptation des espèces à différents types d'environnements. Cependant, les anomalies de la méiose sont responsables chez les mammifères, y compris chez les humains, des principales causes de malformations congénitales (Lee 2013).

Dans ce paragraphe, nous allons en premier lieu décrire le déroulement de la méiose en s'appuyant sur les données de cytogénétique qui permettent de distinguer les différentes

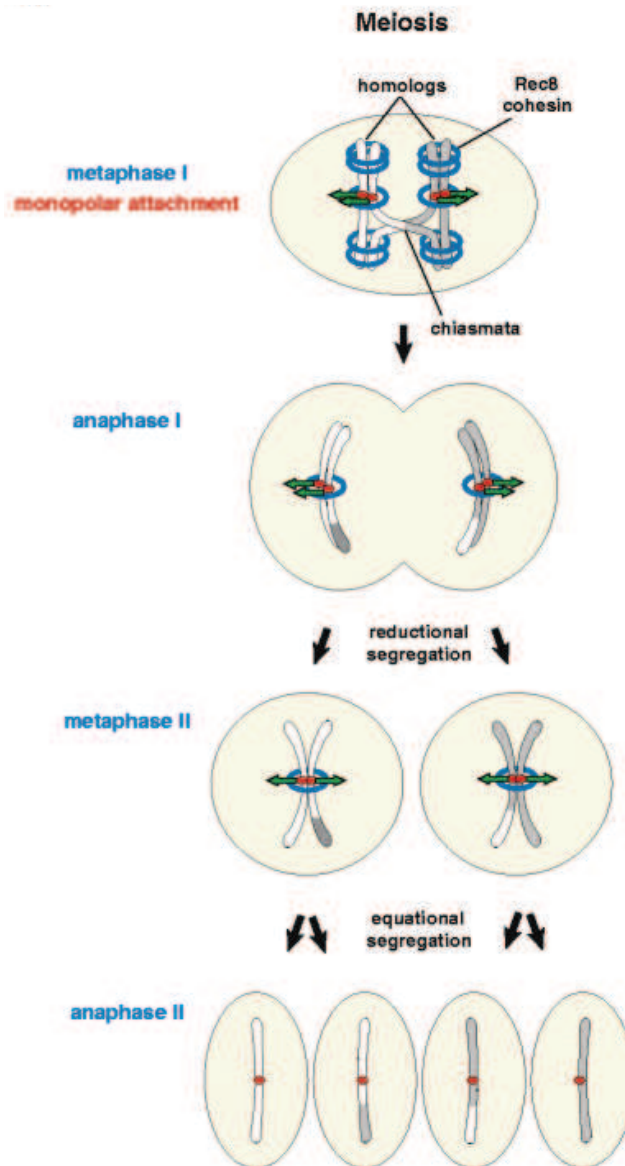


Figure 14. Rôle de la cohésine REC8 dans la ségrégation correcte des chromosomes homologues en M1 puis des chromatides-sœurs en MII (Sakuno et Wanatabe 2009).

étapes de la méiose pas à pas jusqu'à la formation des gamètes recombinaison et équilibrés. En second lieu, nous décrivons les connaissances actuelles sur les acteurs moléculaires intervenant dans la régulation de la méiose.

3.1. Déroutement de la méiose

Dans la majorité des cas, la séparation correcte des chromosomes homologues en méiose ne peut avoir lieu que s'il y a recombinaison entre les homologues. La recombinaison est initiée par la formation de CDB et se poursuit par leur complète réparation.

Le biologiste Allemand Walther Flemming fut le premier en 1882 à observer les chromosomes en mitose (en grec *mitos*, le filament, fait référence à l'aspect des chromosomes). Les techniques de cytogénétique sont des techniques d'étalement des chromosomes qui permettent d'observer le comportement des chromosomes chez de nombreuses espèces. En effet, en fonction de la condensation de la chromatine, les chromosomes prennent différentes formes qui sont caractéristiques des stades du cycle cellulaire. En 1955, la microscopie électronique vient révolutionner la cytologie et Moses observe pour la première fois les chromosomes en prophase de méiose (Moses 2006). Les techniques d'immunofluorescence et la microscopie électronique se complètent et permettent une meilleure compréhension de la dynamique des chromosomes lors de la recombinaison méiotique (Zickler and Kleckner 2015) (Figure 13).

3.1.1. La prophase I de méiose, le siège de la recombinaison

La méiose peut durer de quelques heures, chez la levure, à plus de 40 ans chez la femme (du 7^{ème} mois in utero à la ménopause). Cependant, quel que soit l'organisme considéré, la prophase de la première division de méiose (prophase I) est l'étape la plus longue car elle est le siège de profonds changements de l'état chromatinien. Chez *Arabidopsis thaliana* et en général, la prophase I de méiose représente 90% de la durée totale de la méiose (Armstrong and Jones 2003).

Le début de la méiose est associé à une réorganisation dynamique de la chromatine (phosphorylation et méthylation des histones H3) (Verma and Hong 2008). Ces modifications profondes de la structure des axes chromosomiques et de la compaction de la chromatine sont des marqueurs qui permettent de subdiviser en 5 stades la prophase I de méiose : leptotène, zygotène, pachytène, diplotène et diacinèse (Zickler and Kleckner 1999; Verma and Hong 2008).

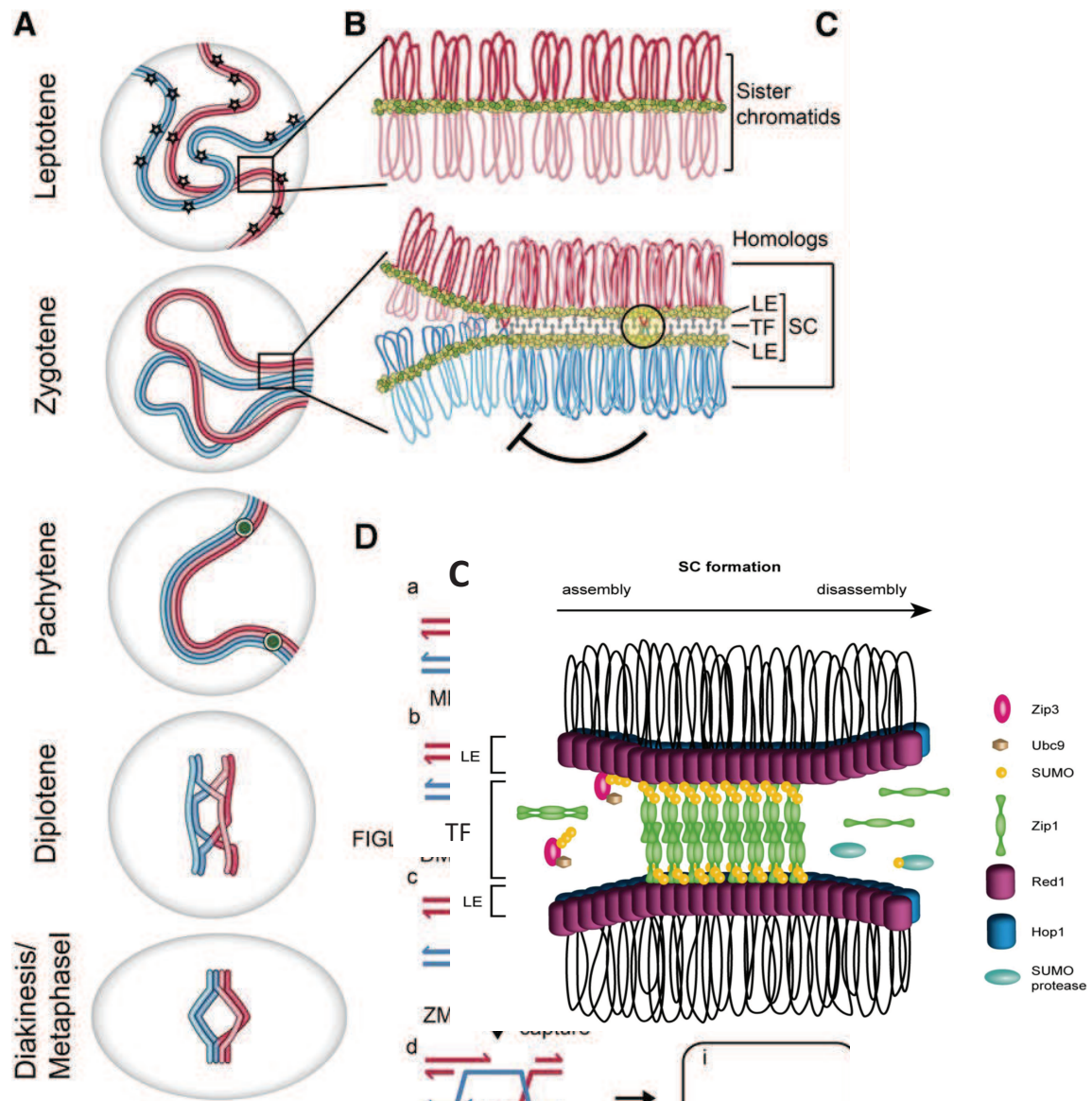


Figure 15 : Prophase I de méiose et mise en place de l'axe chromosomique (adaptation de Lambing et al. 2017).

A- Stades de la prophase I de méiose.

B- Les chromatides méiotiques sont organisées en boucles dont les bases sont ancrées sur des axes chromosomiques (ronds verts) aussi appelés éléments latéraux (LE pour *Lateral Elements*). Les cassures double-brin (CDBs) sont formées dans ce contexte. La structure de la chromatine et les protéines axiales influencent vraisemblablement la formation des CDBs en régulant l'accessibilité et l'activité de la machinerie des CDBs. Au zygotène, les chromosomes homologues synapse grâce à la mise en place des filaments transversaux (TF pour *Transverse Filaments*) qui relient les axes des chromosomes homologues (LE). Les intermédiaires de recombinaison, considérés comme des nodules de recombinaison, sont situés dans la région centrale du complexe synaptonémal (cercle noir). Une fois qu'un CO est désigné (cercle noir), des mécanismes dont on ne connaît pas encore le fonctionnement favorisent la réparation des intermédiaires de recombinaison restants en non-crossing-overs (NCO) (flèche noire).

C- Lors de l'assemblage du complexe synaptonémal (SC pour *Synaptonemal Complex*), Zip3 recrute Ubc9 et des protéines SUMO au sein du siège du synapsis, facilitant ainsi la formation des chaînes SUMO et de la conjugaison de SUMO à d'autres protéines (telles que Red1). Les dimères de Zip1 se polymérisent le long des éléments latéraux (LE) où ils peuvent se lier à des chaînes Red1 et des chaînes SUMO menant à la création de la région centrale (TF). Le démontage du SC pourrait s'effectuer par la dissociation des chaînes SUMO conjuguée à l'action de SUMO protéases ou tout autre mécanisme antagoniste.

3.1.1.1. Leptotène : Mise en place des axes chromosomiques

Suite à la phase de réplication pré-méiotique de l'ADN, les jeux de chromosomes parentaux sont alors constitués chacun de deux chromatides sœurs parfaitement identiques et unies entre elles notamment par un complexe de sous-unités protéiques appelé cohésine. Ces protéines sont extrêmement bien conservées entre les organismes, de la levure aux mammifères. La protéine REC8 est une sous-unité du complexe cohésine spécifique de la méiose. REC8 s'associe aux autres sous-unités protéiques pour former des sortes d'anneaux qui viennent littéralement encercler les chromatides sœurs sur toute leur longueur empêchant ainsi une dissociation précoce des chromatides sœurs avant l'anaphase (Sakuno and Watanabe 2009) (Figure 14). Jusqu'à présent 4 sous-unités de cohésines spécifiques de la méiose ont été identifiées chez les mammifères (Lee 2013). Nous savons également qu'elles forment des complexes avec les cohésines mitotiques. REC8 étant une cohésine spécifique de la méiose (Kudo et al. 2009), c'est un excellent marqueur cytogénétique de l'axe chromosomique (Wang et al. 2010a).

Alors que les cohésines permettent de maintenir deux à deux les chromatides sœurs, les condensines se fixent sur chaque chromatide et condensent l'ADN de sorte qu'il forme des boucles qui viennent s'ancrer sur l'axe formé par les cohésines. En cytologie, on voit apparaître des boucles d'ADN qui s'échappent des axes chromosomiques et s'attachent à l'axe de part et d'autre des condensines (Figure 15). Cette organisation de la chromatine est déterminante pour la poursuite de la méiose et l'initiation du processus de recombinaison chez la plupart des espèces (*S. cerevisiae* ; *S. pombe*; *C. elegans*; *M. musculus*; *A. thaliana*).

Afin de compléter l'axe chromosomique, un troisième type de structure multi-protéique appelé élément axial (EA) s'associe aux cohésines et condensines. L'EA est une structure linéaire qui participe à l'élongation et au maintien de l'axe chromosomique. Bien que très conservées, les protéines qui composent l'élément axial ont des séquences très peu similaires entre les organismes. La protéine HOP1, isolée chez la levure, a pour homologue ASY1 (ASYNAPTIC1) chez *A. thaliana*. Chez le riz, PAIR2 (Pairing 2) est l'orthologue de ASY1 (Nonomura et al. 2004), et a le même profil de localisation que ASY1 et HOP1. PAIR2 est associée à l'EA du leptotène au zygotène puis se retire de l'élément axial le long des bras chromosomiques lorsque le synapsis se forme. En revanche, il reste des protéines PAIR2 au niveau des centromères jusqu'à la diacinèse. Le mutant *pair2* a le même phénotype que *asy1* chez *A. thaliana*, à savoir une diminution de la fertilité de 90% (Nonomura 2006). ASY3 chez *A. thaliana* ou son orthologue PAIR3 chez le riz (Yuan et al. 2009) font aussi partie du complexe de protéines qui forment l'EA. Chez *A. thaliana*, l'utilisation d'anticorps dirigés contre ASY3 (ou d'autre protéines de l'EA) a permis de suivre le développement de l'EA au cours du leptotène (Figure 15).

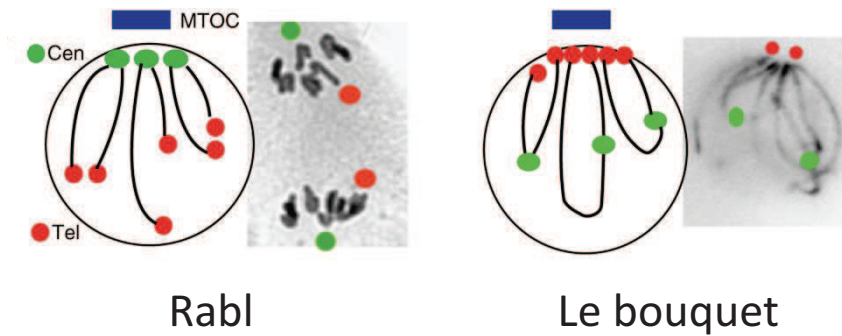


Figure 16 : Le bouquet, une organisation caractéristique de la méiose (Zickler and Kleckner, 2015).

À gauche : Anaphase de division mitotique. Les chromosomes se séparent avec leurs centromères faisant face au pôle du fuseau (et les centres organisateurs des microtubules correspondant (MTOC en bleu). Les télomères se retrouvent à une latitude différente selon la taille des bras chromosomiques et restent dans cette disposition appelée "Rabl".

À droite : Pendant la prophase I de méiose, les télomères se tournent et se regroupent face au MTOC avec les centromères plus ou moins dispersés dans le noyau (images de *Sordaria* de D. Zickler). Cette organisation est appelée bouquet.

Les deux chromatides sœurs, empilées l'une sur l'autre, forment des boucles dans le plan perpendiculaire à cet axe latéral (Figure 15). Cette organisation en boucle est étonnamment conservée bien que la taille des boucles d'ADN varie en fonction des organismes (Kleckner 2006). En effet, dans les oocytes de souris, ces boucles mesurent environ 3-5µm de long, chacune contenant 300-440 bases, alors que la levure présente des boucles de 0,2-0,6µm contenant chacune 20kb d'ADN (Kleckner 2006). L'ADN sortant de l'axe chromosomique est le plus exposé à la recombinaison méiotique. C'est à ce stade qu'apparaissent des nodules précoces de recombinaison (NPRs) visibles en microcopie électronique (Zickler and Kleckner 1999).

Nous verrons dans les paragraphes suivants à quel point cette structuration en boucle le long de l'axe chromosomique est importante pour la régulation des cassures doubles-brins de l'ADN.

3.1.1.2. Recherche des homologues via la formation du bouquet

Les mécanismes qui permettent de faire en sorte que les couples d'homologues se reconnaissent ne sont pas tout à fait élucidés mais une part du mystère se cache dans le bouquet (Zickler and Kleckner 2015) (Figure 16). Le bouquet est une structure dynamique qui se met en place au tout début du zygotène. En effet, les chromosomes migrent vers un même pôle de la cellule et s'attachent à l'enveloppe nucléaire via leurs télomères puis s'amassent en grappe pour former une structure qui ressemble à un bouquet de fleurs. Le bouquet a été observé chez la quasi-totalité des eucaryotes (Harper 2004). Le bouquet peut être très important chez certaines espèces qui ne possèdent pas de complexe synaptonémal. C'est l'exemple de *S.pombe* où le bouquet formé par les chromosomes est très dense et où le noyau est ensuite remué dans la cellule, comme dans une machine à laver (Chikashige et al. 1994). A l'inverse, chez *A. thaliana*, le bouquet est très relâché. Il est proposé que le bouquet pourrait non seulement promouvoir le synapsis, en permettant l'appariement des chromosomes, mais également restreindre la recherche d'homologie à deux dimensions au lieu de trois, en alignant les chromosomes de façon colinéaire. Le bouquet et le mouvement des chromosomes auraient aussi un rôle dans la résolution des nœuds entre chromosomes (interlocks) (Zickler and Kleckner 1999). Les mouvements des chromosomes désavantagent les interactions ectopiques et favorisent les interactions homologues (Zickler and Kleckner 2015). Toutefois, chez certains organismes, tel que *C. elegans* par exemple, l'absence de bouquet n'empêche pas l'alignement des chromosomes homologues (Kleckner 2006). La formation du bouquet et le mouvement rapide des chromosomes en prophase (RPMs for *rapid meiotic prophase chromosome movements*) ne serait qu'une contribution à l'association des chromosomes homologues (Lee et al. 2012).

L'étude du mutant *pam1* (*plural abnormalities of meiosis*) chez le maïs montre, comme son nom l'indique, des anomalies de la méiose. Même si les télomères se polarisent et s'attachent

normalement sur l'enveloppe nucléaire, ils forment plusieurs petit groupes mais ne parviennent pas à se regrouper en un bouquet rigide normal (Harper 2004). On voit aussi des synapses incomplètes et des synapses entre chromosomes non homologues en conséquence de quoi les gamètes males sont totalement stériles alors que coté femelle un peu moins de 1% de gamètes formés sont fertiles. Ceci suggère que la formation du bouquet chez le maïs facilite grandement l'appariement, le synapsis et la résolution des "interlocks" mais le bouquet n'est pas indispensable au bon déroulement de la méiose puisque quelques gamètes peuvent être produits en l'absence de PAM1 (Harper 2004). Comme le souligne Lee, la formation du bouquet ne serait qu'une contribution à l'association des homologues. L'attachement des télomères sur la surface interne de l'enveloppe nucléaire dépend de protéines insérées dans la membrane nucléaire, les protéines SUN et KASH (Kozul and Kleckner 2009). Ces protéines forment des complexes qui permettent l'attachement et le mouvement des télomères sur la membrane nucléaire interne. Le rôle exact de ces mouvements des chromatides reste inconnu même si on suppose qu'il sert à éliminer les associations indésirables entre chromosomes non-homologues et de ce fait à favoriser l'appariement des chromosomes homologues (Kozul and Kleckner 2009). Chez la souris, la délétion *sun1* induit une stérilité des mâles et des femelles due au mauvais attachement des télomères et à un défaut de synapsis (Ding et al. 2007). Chez *C. elegans*, la délétion de *sun1* induit des aberrations de synapsis qui se fait alors entre chromosomes non-homologues (Penkner et al. 2009). Chez *A. thaliana*, l'étude du double mutant *Atsun1/Atsun2* montre des défauts méiotiques sévères, à savoir un retard dans la progression de la méiose, l'absence de synapsis complet, la présence de structures "interlock" non résolues, ainsi qu'une réduction de la fréquence moyenne du nombre de chiasma par cellules (Varas et al. 2015). Chez le riz, la protéine REC8, essentielle à la formation de l'axe chromosomique, joue aussi un rôle lors de la formation du bouquet. L'utilisation de marqueur spécifique des télomères permet d'observer que chez le mutant *Osrec8*, contrairement au sauvage, les télomères restent dispersés dans tout le volume nucléaire au lieu de se regrouper et de s'attacher à la paroi (Shao et al. 2011).

Toutes ces données montrent que, si le bouquet semble présent chez une grande majorité des espèces, son rôle dans le synapsis et la recombinaison est plus ou moins important en fonction des espèces.

3.1.1.3. Zygotène : Mise en place du complexe synaptonémal

C'est au stade zygotène (du grec zygo "paire") que les paires de chromosomes homologues s'apparient (Figure 15). Les protéines qui permettent de rattacher physiquement les chromosomes homologues entre eux sont les protéines du complexe synaptonémal (CS). On sait aujourd'hui que le CS est composé d'une structure tripartite avec un élément central (CE pour *central element*, aussi appelé TF pour *transverse filament*), reliant les axes latéraux des deux chromosomes homologues. Les axes chromosomiques de chaque chromatide sœur

formés au leptotène sont renommés ici LE (pour *lateral element*) et s'assemblent *via* le CE qui est mis en place progressivement au zygotène à partir de plusieurs sites de nucléation, puis qui s'étend de façon bidirectionnelle pour associer les chromosomes homologues entre eux c'est ce qu'on appelle le synapsis (sorte de fermeture éclair qui zippe les 2 homologues).

Une des protéines les plus connues du CE est la protéine ZYP1 chez *A. thaliana* (Higgins et al. 2005) et son homologue ZEP1 chez *O. sativa* (Wang et al. 2010a). L'immunolocalisation de la protéine ZYP1 permet de suivre la progression du synapsis. Au stade pachytène, un marquage continu de Zip1 sur toute la longueur des chromosomes traduit, d'un point de vue cytogénétique, un synapsis complet. Selon les espèces, le

3.1.1.4. Pachytène : description et rôle du chiasma ou CO

Le synapsis est complet au stade pachytène puis se dépolymérise jusqu'au stade diacinèse sauf au niveau des chiasmas. Les chiasmas sont les derniers liens physiques entre les homologues, l'ensemble est appelé bivalents chez les espèces diploïdes.

Le chiasma est une structure cytologique qui traduit la présence d'un événement de crossing-over (CO). Le chiasma résulte, non seulement de l'échange physique de séquences ou morceaux de chromosomes entre 2 chromatides des deux homologues, mais il est aussi dû au fait que les chromatides homologues qui sont engagées dans ce chiasma sont attachées par les cohésines. Les deux chromatides engagées dans un CO subissent un changement de continuité dû au CO (Figure 15) et sont aussi liées par des anneaux de cohésion qui encerclent les chromatides sœurs. Ainsi, les COs permettent non seulement de brasser les allèles parentaux, mais aussi d'assurer un lien entre les chromosomes homologues en méiose I.

C'est à ce stade qu'on voit disparaître les NPRs et apparaître de façon concomitante, mais moins nombreux, les Nodules Tardifs de Recombinaison (NTRs) au niveau des EC (Zickler and Kleckner 1999). Dès 1975, Carpenter a montré chez la *Drosophila* que le NTR est l'endroit où se produit le CO. Elle a simplement montré qu'il y avait le même nombre d'échange d'information génétique que de NTRs dans une cellule de drosophile. Les cytogénéticiens voyaient des chiasmas alors que les généticiens observaient des échanges d'allèles. Carpenter a montré en 1975 que ces deux événements avaient lieu au même endroit (Carpenter 1975).

La formation des COs est étroitement liée à la mise en place du synapsis. Le rôle de Zip1 dans la formation des COs est différent selon les organismes (Wang et al. 2010a). Deux exemples, chez *A. thaliana*, le nombre de CO est réduit de près de 70 à 80% chez le mutant *zyp1a / zyp1b* par rapport au sauvage seuls les CO de classe I seraient maintenus (Higgins et al. 2005). A l'inverse chez le riz, l'absence de synapsis chez le mutant *Oszep1* a pour conséquence d'augmenter le nombre de CO (Wang et al. 2010a).

La présence de chiasmas est une condition nécessaire à la bonne disjonction des chromosomes homologues en première division chez l'immense majorité des espèces. En effet, les chromosomes homologues qui ne sont pas liés par un chiasma ségrègent aléatoirement et indépendamment vers l'un ou l'autre pôle de la cellule. C'est le cas chez de nombreux mutants présentant un défaut de CO, comme le mutant *spo11* ou les mutants de la voie ZMM de formation des COs (Grelon et al. 2001; Stacey et al. 2006). Ces observations ont aussi été faites chez le riz (Luo et al. 2013b).

3.1.2. Progression de la méiose et formation des gamètes

Après la prophase I, la ségrégation des chromosomes homologues débute. Pour rappel, les chromosomes homologues sont maintenus au niveau des chiasmas et stabilisés grâce à la cohésion assurée par les cohésines le long des bras. En métaphase I, les fibres de microtubules émanant des pôles opposés de la cellule s'attachent aux bivalents et alignent les chromosomes sur la plaque métaphasique.

Lorsque les chromosomes sont correctement attachés au fuseau, le complexe APC/C (Anaphase Promoting Complex / Cytosome) est activé et libère la séparase qui à son tour clive les sous-unité REC8 des cohésines présentes le long des bras chromosomiques (Cromer et al. 2013). Ceci permet le relâchement de la cohésion le long des bras, ainsi que la résolution des chiasmas et déclenche l'anaphase I. Les chromosomes homologues s'éloignent l'un de l'autre grâce à la force exercée par le fuseau. La mono-orientation des kinétochores permet aux chromatides sœurs de migrer du même côté (Figure 13) (Hauf and Watanabe 2004). Chaque cellule fille reçoit, en fin de méiose I, un exemplaire de chaque chromosome homologue, c'est ce que l'on appelle la division réductionnelle. La cohésion centromérique est conservée lors de cette première division et permet aux deux chromatides sœurs de rester solidaires et de s'aligner correctement sur la plaque en métaphase. Cette cohésion centromérique est protégée de la dégradation par les protéines SHUGOSHIN (Kitajima et al. 2004) et PATRONUS chez *A. thaliana* (Cromer et al. 2013). En télophase I, le fuseau méiotique disparaît, les chromosomes se décondensent et une nouvelle membrane nucléaire apparaît. A ce stade, la cytodiérèse n'est pas obligatoire. Chez le riz par exemple, une paroi cellulaire est mise en place en fin de méiose I alors que ce n'est pas le cas chez *A. thaliana* (Chen et al. 2005).

Suite à cette première division, une seconde sera initiée sans phase de réplication préalable. En méiose II, les chromosomes vont à nouveau se condenser. La prophase II est très rapide voir quasi inexistantes. Lors de la métaphase II, ces chromosomes à deux chromatides vont s'aligner sur la plaque métaphasique et le fuseau de microtubules se forme à nouveau et capture chaque chromatide sœur. Cette fois les kinétochores des chromatides sœurs seront

3.2. Les mécanismes moléculaires de la recombinaison méiotique

La recombinaison méiotique est basée sur la formation d'une cassure double-brin de l'ADN (CDB) et sa réparation par recombinaison homologue. Plusieurs voies de réparation coexistent, mais ils conduisent à la formation des crossing-over (COs) ou des non-crossing-over (NCOs) (Mercier et al. 2015). La recombinaison homologue (RH) est le mode de réparation principal des CDBs en méiose et il se fait au détriment de la réparation par jonction d'extrémité non-homologues ou NHEJ (*Non-Homologous End Joining*,) (Chapman et al. 2012).

Au cours de ce chapitre nous développerons les mécanismes moléculaires de la RH. Tout commence par la formation des CDBs, suivie de leur maturation. Suite à cette maturation, un des brins lésés envahit l'ADN d'une chromatide homologue - c'est l'invasion simple brin - pour former la D-loop (Displacement loop structure). Cet intermédiaire est un carrefour de la RH qui peut être pris en charge par différentes voies de réparation pour former des CO ou des NCO. La régulation de la formation des CO sera ensuite abordée ainsi que leur distribution dans le génome. Ce modèle de réparation de la cassure double brin appelé "*Double-strand break repair*" (DSBR) a été décrit pour la première fois par Szostak en 1983 puis sans cesse corrigé (Szostak et al. 1983; Allers and Lichten 2001; De Muyt et al. 2012). Pour finir nous verrons comment détecter un CO soit par des approches cytogénétique soit en utilisant la génétique.

Dans cette partie nous allons détailler toutes ces étapes de l'initiation de la CDB, par SPO11 et ses co-facteurs (§3.2.1) ; à la maturation des CDBs (§3.2.3) ; puis l'invasion du simple brin et la formation d'intermédiaire de réparation (§3.2.4) et enfin les voies de formation des COs de classe I ou II et de NCOs (§3.2.5). Nous terminerons par la description des systèmes de régulation entre la formation de CO ou de NCO (§3.2.6).

3.2.1. Formation de la cassure double-brin de l'ADN : l'initiation

En méiose, les cassures double-brin de l'ADN sont catalysées de façon contrôlée par la protéine SPO11 dès la mise en place des axes chromosomiques au leptotène (Zickler and Kleckner 1999, 2015). SPO11 est une protéine hautement conservée chez les eucaryotes, ce qui a permis d'expliquer sa fonction chez *A. thaliana*. SPO11 possède un domaine topoisomérase homologue à celui de la topoisomérase VIA identifiée chez les archées (Bergerat et al. 1997). La Topo VIA est la sous unité catalytique d'une topoisomérase d'ADN de type II (Bergerat et al. 1997). Les topoisomérases sont des enzymes capables de couper et religuer l'ADN. Elles sont impliquées dans la résolution des nœuds formés par l'enchevêtrement des chromosomes lors des étapes de réplication de l'ADN ou de transcription par exemple. Il est supposé que SPO11 catalyse avec l'aide de nombreuses autres protéines la coupure de l'ADN double-brin mais SPO11 n'est pas capable de réaliser la ligation des brins coupés (De Massy 2013).

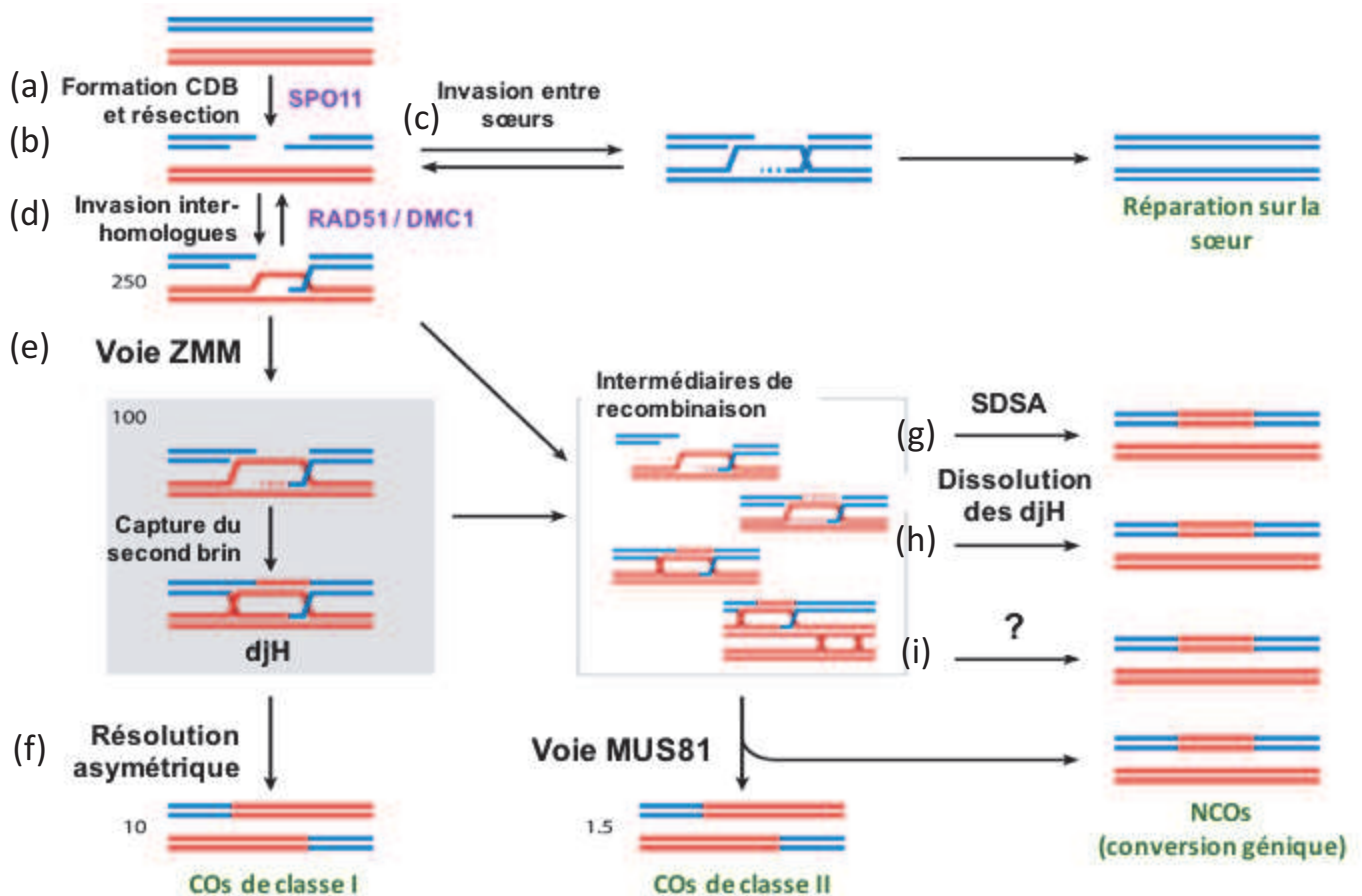


Figure 17 : Modèle présentant les mécanismes de la recombinaison méiotique (d'après Mercier et al. 2015).

La recombinaison méiotique est initiée par la formation d'un grand nombre de cassures double-brin (DSB) (a) qui sont réséquées pour (b) former une extrémité 3'OH simple brin. Cet ADN simple brin peut alors envahir soit la chromatide sœur intacte (c), soit l'une des deux chromatides homologues pour former une D-loop (d). Les intermédiaires inter-homologues peuvent être protégés par des composants de la voie ZMM (e), générant des intermédiaires appelés double-jonction de Holliday (dHJ) pouvant être résolus en crossing-overs de classe I (COs) (f). A partir de la D-loop, une autre alternative permet de maturer les intermédiaires en non-crossing-overs (NCOs) par différents mécanismes, y compris le SDSA (*Synthesis-dependent strand annealing*) (g), ou la dissolution de dHJ (h), et éventuellement d'autres mécanismes encore inconnus (i). De plus, une voie indépendante de ZMM produit des CO de classe II (j). Le nombre moyen de chaque intermédiaire par méiose chez *Arabidopsis thaliana* est indiqué en gris.

On sait que, chez les archées, les Topo VI agissent sous forme d'hétéro-tétramère composés de deux sous-unités A et 2 sous-unités B, ces dernières présentant un site de fixation de l'ATP nécessaire à l'activité catalytique de la sous-unité A (Buhler et al. 2001; Champoux 2001). En 2016, l'équipe de M. Grelon montre qu'il existe, chez *A. thaliana*, une protéine homologue à la sous-unité TopoVIB de l'ADN topoisomérase VI B des archées (MTOPVIB), et montre que cette protéine est capable de former un complexe avec l'hétérodimère SPO11-1/SPO11-2 (Vrielynck et al. 2016). Puis l'équipe de B. De Massy identifie d'autres homologues à la topoisomérase VI B des archées chez *S. cerevisiae*, *S. pombe* et *D. melanogaster*. De Massy en conclut qu'il existe une famille protéique de TOPOVIB-Like et qu'elle est conservée chez les opisthocontes (des archées aux chordés) (Robert et al. 2016). Cette avancée majeure suggère que le complexe catalytique responsable de la formation des CDBs de méiose chez les eucaryotes adopte une structure et un mode de fonctionnement similaire à celle de Topo VI (Vrielynck et al. 2016).

Chaque espèce semble présenter en dehors de SPO11 et probablement de TOPOVIBL des mécanismes qui lui sont spécifiques (Mercier et al. 2015). Chez *S. cerevisiae*, Spo11 a besoin de l'aide d'une dizaine de protéines associées pour faire une CDB (MER11, RAD50, XRS2, SKI8, REC102, REC104, REC114, MEI4 et MER2; appelées aussi complexes MRX et RMM) (Lam and Keeney 2014). Chez les plantes, en plus du complexe SPO11, plusieurs protéines ont déjà été identifiées comme étant essentielles pour la formation des CDBs de méiose. C'est le cas chez *A. thaliana* de PRD1, PRD2, AtPRD3/OsPAIR1 et DFO (Nonomura et al. 2004; Miao et al. 2013a; Lambing et al. 2017). Ces protéines ont probablement, chez les plantes, un rôle similaire à celui des protéines compagnes chez *S. cerevisiae*. Par exemple PRD1 ressemble beaucoup à Mei1 de la souris (De Muyt et al. 2007) et PRD2 est probablement l'orthologue de Mei4 de la souris et la levure (De Muyt et al. 2009; Kumar et al. 2010). Chez le riz, la protéine CRC1 (similaire à AtPCH2) interagirait avec les protéines PAIR1 et P31^{comet} pour former un complexe qui permettrait de réguler la formation des CDBs (Miao et al. 2013a; Ji et al. 2016). Chez le riz, CRC1, PAIR1 et COMET sont, comme SPO11, essentielles à la formation des CDBs. Chez *O. sativa* encore, la cycline SDS serait également impliquée dans la formation des CDBs (Wu et al. 2015) alors que chez *A. thaliana*, SDS est nécessaire lors de la réparation des CDBs avec DMC1. Toutes ces différences suggèrent que les CDBs sont régulées de diverses façon selon les espèces (Lambing et al. 2017).

La façon dont ces protéines interagissent est encore mal connue et, à part SPO11, non seulement ces protéines présentent des séquences peu conservées dans le règne végétal, mais également une divergence fonctionnelle. Il est désormais crucial d'étudier leurs rôles respectifs et d'essayer de comprendre leurs mécanismes de régulation (Mercier et al. 2015).

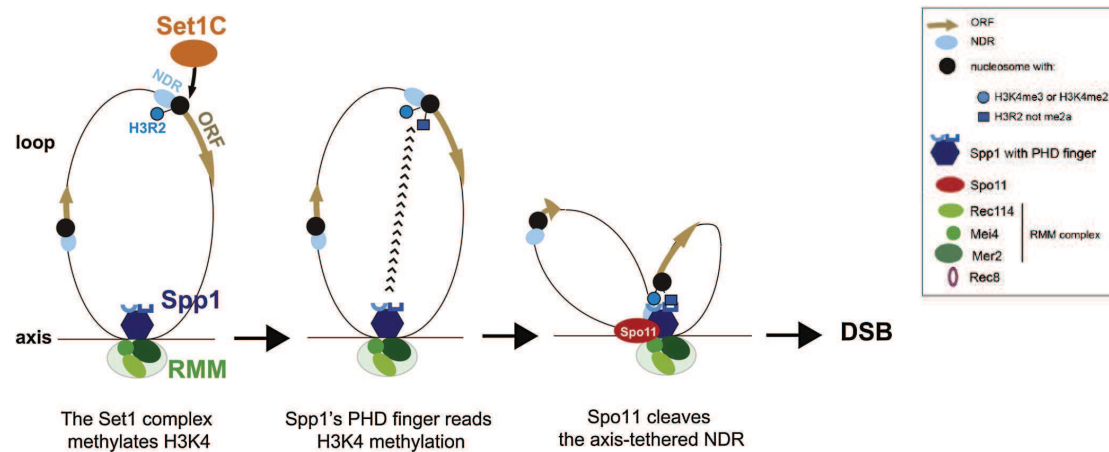


Figure 18 : Rôle de *Spp1* dans la formation des CDBs chez la levure (Sommermeyer et al. 2013).

Les chromosomes méiotiques sont organisés en réseaux de boucles, ancrées à la structure de l'axe des chromosomes. Les cassures double brin d'ADN programmées (DSB) qui initient la recombinaison méiotique sont catalysées par Spo11 et les protéines accessoires. Ces DSBs se forment dans la boucle au niveau des séquences promotrices, tandis que les protéines accessoires (complexe RMM) sont situées sur les axes chromosomiques. Les travaux de Sommermeyer et al. ont permis de montrer que Spp1, un membre conservé du complexe histone H3K4 méthyltransférase, permet de faire le pontage entre ces deux régions chromosomiques (boucles et axe). La présence de Spp1 est nécessaire à la formation d'un nombre normal de DSB et est associée à l'axe chromosomique pendant la méiose. Spp1 interagit physiquement avec la protéine Mer2 (du complexe RMM). Le module PHD (pour *plant homeodomain finger*) de Spp1, détecte la méthylation de H3K4 à proximité des promoteurs et favorise la formation de DSB en attachant ces régions aux axes chromosomiques et en activant le clivage par les protéines DSB (Légende d'après l'article de Sommermeyer et al.).

(NDR : *Nucleosome-Depleted Regions* ; ORF : Open Reading Frame ; DSB : Double-Strand Breaks).

3.2.2. Localisation et modèle de formation des cassures double-brin

L'organisation de la chromatine en boucle autour des axes chromosomiques impose une coordination entre les différents événements et les différents partenaires. Des études ont montré que les cassures sont formées sur les boucles, alors que SPO11 co-localise avec une cohésine de l'axe REC8 (Kugou et al. 2009) et nombre de ses partenaires (Blat et al. 2002; Kumar et al. 2010; Panizza et al. 2011). Le modèle émergent suggère que les séquences situées sur une boucle, et qui subissent la cassure, sont entraînées vers l'axe *via* des interactions protéines-protéines pour procéder à cette cassure (modèle tethered axis-loop, Figure 18). Chez *S.cerevisiae*, il a notamment été montré que SPP1 est essentielle à ce processus car elle interagit spécifiquement avec des marques de l'ADN au niveau des cassures et avec Mer2 sur l'axe chromosomique (Sommermeyer et al. 2013; Acquaviva et al. 2013) (Figure 18).

3.2.3. Maturation des cassures doubles-brins

Chez la levure et les mammifères où le mécanisme est le plus connu, on sait qu'à la suite de la formation de la CDB par SPO11 et les protéines compagnes, SPO11 reste liée de manière covalente aux extrémités 5' des brins d'ADN (Figure 19). Or SPO11 doit être enlevé du site de la cassure pour que la CDB puisse être réparée. La maturation des CDBs s'effectue en trois étapes qui permettront de préparer les extrémités 3' flanquantes à la synthèse d'ADN qui suivra (§ 3.2.5).

La résection des brins d'ADN clivés par SPO11 et ses partenaires se fait en premier lieu grâce à l'action combinée du complexe MRX (MRE11-RAD50-XRS2) et de la protéine SAE2 chez *S. cerevisiae* ou du complexe MRN (MRE11-RAD51-NBS1) et de la protéine COM1 chez les plantes (Keeney and Neale 2006). Chez *S. cerevisiae*, où le mécanisme a été étudié, MRX-SAE2 permet de libérer SPO11 de la chromatine en formant une coupure en 3' du site de liaison de SPO11 à l'ADN, ce qui a pour effet de libérer les extrémités 5' de ce que l'on appelle les oligos SPO11 (SPO11 se détache de la matrice ADN emportant avec lui quelques nucléotides, appelés les oligos spo11) (Neale et al. 2005). Les extrémités 5' libérées sont ensuite dégradées, permettant ainsi l'allongement des extrémités 3' simple-brin sortantes. Ce phénomène appelé résection, nécessite une activité exonucléasique 5'-3' qui pourrait être le fait de EXO1 et de SGS1-DNA2 mais tout ceci reste à démontrer (Zhu et al. 2008; Longhese et al. 2010). Toujours chez *S. cerevisiae*, d'autres études ont montré qu'à la fin de la résection, les deux extrémités 3' simple-brin sortantes sont alors pris en charge par le complexe RPA (pour Replication Protein A), un complexe hétérotrimérique permettant le recrutement des recombinaisons. Les fonctions de RPA sont encore très peu connues mais ce complexe pourrait servir à éliminer les structures secondaires de l'ADN simple-brin afin de permettre le recrutement des recombinaisons (Heyer et al. 2010). Selon des études génétiques et

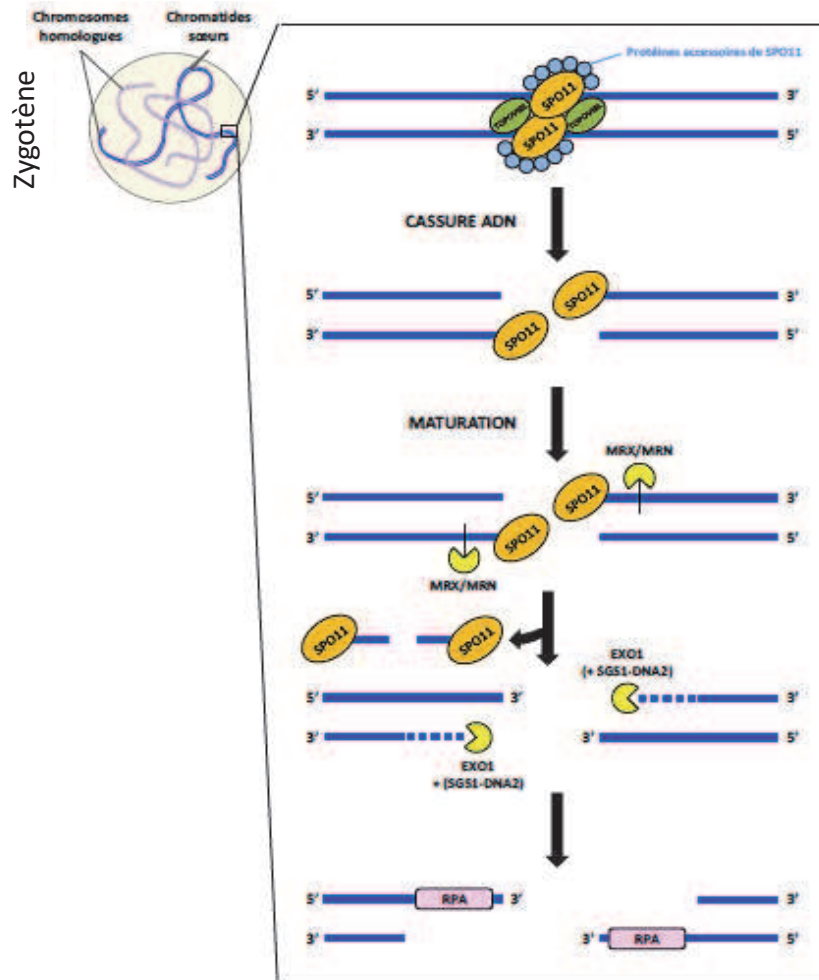


Figure 19 : Formation et maturation des cassures double-brin de l'ADN (d'après Neale et al. 2005).

La recombinaison méiotique est initiée par la formation de cassure double brin (CDB) de l'ADN via notamment l'action de la protéine SPO11. SPO11 et ses protéines accessoires sont recrutées sur la chromatine et procèdent à la cassure grâce à l'activité topoisomérase de SPO11. SPO11 reste accrochée sur la chromatine. Le complexe MRX et SAE2 coupe en 3' de la CDB ce qui mène au relargage des oligos SPO11. EXO1-SGS1 (et peut-être MRX) dégradent l'ADN de l'extrémité 5' vers l'extrémité 3', libérant l'extrémité 3' du brin complémentaire. Ce brin est alors pris en charge par la protéine RPA.

biochimiques réalisées chez le riz et *A. thaliana*, il semblerait qu'il y ait plusieurs complexes RPA ayant des fonctions partiellement redondantes, dont certaines seraient nécessaires à la recombinaison méiotique (Mercier et al. 2015) (Figure 19).

3.2.4. L'invasion par le simple-brin et la formation d'intermédiaires

La recombinaison homologue (RH) procède à la réparation des CDBs en utilisant une séquence homologue comme matrice. Chez les organismes diploïdes, deux matrices ADN sont disponibles pour la réparation : la chromatide sœur et l'une des deux chromatides de l'homologue. Cependant, seule la seconde peut aboutir à la formation de connexions physiques entre les chromosomes homologues indispensable à leur bonne ségrégation en méiose. En mitose, la RH utilise préférentiellement la chromatide sœur comme matrice de réparation (Lao et al. 2013). En méiose, un biais vers une réparation sur une chromatide homologue non sœur assure la formation des crossing-overs : c'est ce que l'on appelle le biais inter-homologue (Schwacha and Kleckner 1997).

La recherche d'homologie et l'invasion du brin d'ADN homologue sont réalisées par deux recombinaisons, RAD51 et DMC1. Elles sont homologues à la protéine RecA d'*Escherichia coli* qui s'assemble en filament nucléo-protéique et s'enroule en spirale autour d'un brin d'ADN (Figure 20). Chez *E.coli*, les agrégats d'ADN et de RecA seraient capables de faire un grand nombre de contacts avec une matrice ADN pour rechercher le brin homologue. Ces contacts seraient établis de façon transitoire, sur un système d'essai-erreur, jusqu'à trouver une séquence homologue dont la complémentarité assure la stabilité de l'interaction (Danilowicz et al. 2014). Après la résection, RAD51 et DMC1 seront chargées sur la chromatine en remplaçant le complexe RPA (Figure 20). Elles sont arrangées en filament, nommé filament présynaptique, et vont permettre l'invasion par le simple-brin d'une des extrémités de la CDB ("*single end invasion*", SEI). Cette invasion mène à la formation de la D-loop ("*Displacement loop structure*") (Figure 20) (Sung, Petukhova, & Stratton, 1998; Hunter et al., 2001).

La stœchiométrie des recombinaisons RAD51 et DMC1 sur les extrémités de la CBD fait débat. Chez *A. thaliana*, il a été montré que RAD51 et DMC1 sont localisées sur les extrémités opposées de la CBD (Kurzbaue et al. 2012). C'est ce que l'on appelle le modèle de chargement asymétrique des recombinaisons. Un autre modèle a été proposé par Brown qui soutient l'hypothèse selon laquelle RAD51 et DMC1 seraient associées et chargées de chaque côté de la CBD (Brown et al. 2015).

Chez la levure comme chez *A. thaliana*, l'activité catalytique de RAD51 est indispensable à la formation de CO en méiose, ce qui indique que RAD51 fonctionne comme un partenaire de DMC1 pendant la formation de CO méiotique (le mutant *Atrad51* est stérile). Chez le mutant *dmc1*, RAD51 permet quand même de réparer efficacement les CDBs méiotiques. Cette voie alternative de réparation, dépendante de RAD51, utilise très probablement la chromatide sœur

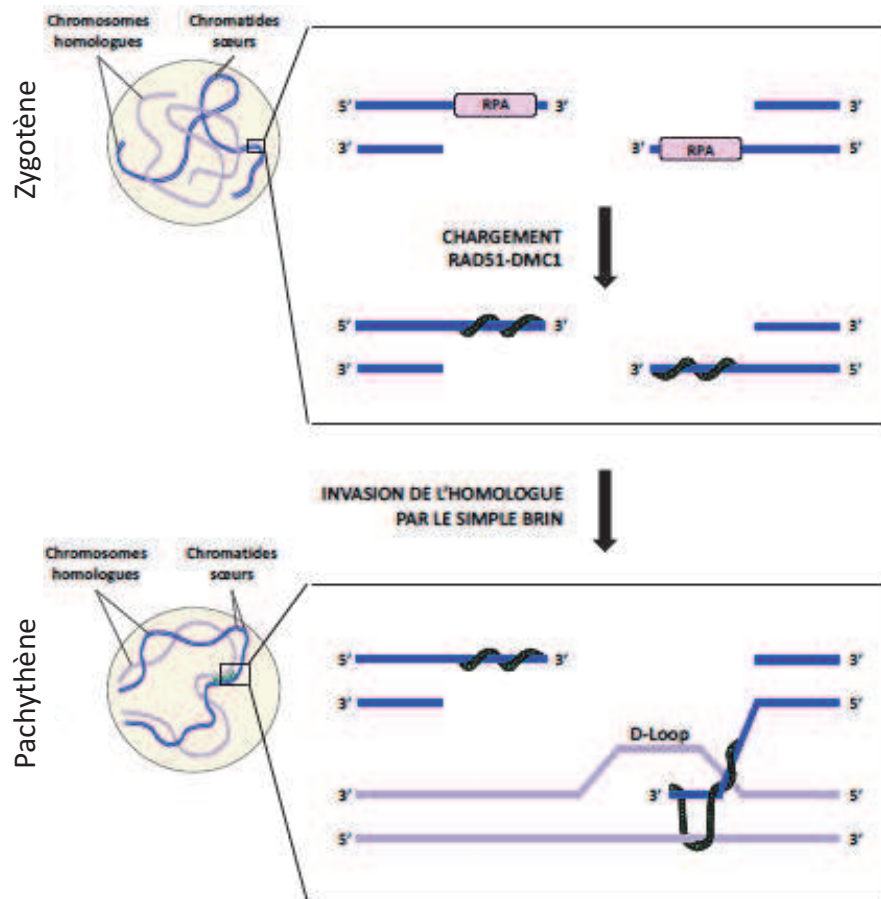


Figure 20 : Chargement des recombinaisons et invasion de la molécule homologue (d'après Neale et al. 2005).

Les recombinaisons RAD51 et DMC1 (représentées en vert) remplacent les complexes RPA sur les extrémités 3' simple-brin sortantes. Celles-ci vont alors favoriser l'invasion de ce simple brin vers la matrice homologue complémentaire entraînant la formation d'une structure appelée D-Loop.

comme matrice de réparation, ce qui entraîne une absence complète de synapsis et la formation de bivalents mais permet de réparer l'ADN. Cependant lorsque DMC1 est présente elle réprime l'activité de réparation de RAD51 favorisant à nouveau la formation de CO (Mercier et al. 2015). L'invasion de la molécule homologue serait donc assurée par la recombinaise DMC1. Quant à RAD51, cette protéine multifonctionnelle catalyserait directement la recombinaison mitotique et indirectement la recombinaison méiotique *via* DMC1 (Cloud et al. 2012; Da Ines et al. 2013).

Jusqu'à présent, peu de différences biochimiques ou structurales ont été identifiées entre RAD51 et DMC1. Cependant un grand nombre de facteurs ont été identifiés comme étant nécessaires au chargement, à la stabilisation et à l'activation de l'une ou l'autre de ces recombinaises (se référer à la liste des protéines identifiées (Mercier et al. 2015)). Les protéines associées pourraient expliquer les rôles divergents de RAD51 et DMC1 pendant la méiose. Par exemple chez *A. thaliana*, la protéine axiale AtASY1, homologue de Hop1 de *S. cerevisiae*, permet de stabiliser DMC1 et non RAD51 pendant la méiose, ce qui donne lieu à une réparation inter-homologue (IH) correcte. De même, SDS, cycline-D spécifique des plantes, est requise au chargement et/ou à la stabilisation de DMC1 chez *A. thaliana* (De Muyt et al. 2009).

De récentes études montrent que HOP2/MND1 est aussi requis pour une activité correcte de DMC1 (Uanschou et al. 2013). Plus récemment encore, l'équipe de R. Mercier a montré que FIGL1 et FLIP limitaient la formation des COs en régulant les interactions dynamiques entre RAD51 et DMC1 (Mercier et al. 2015; Fernandes et al. 2017a) et que cette interaction FIGL1/FLIP était conservée d'*A. thaliana* à l'Humain. De nombreuses autres protéines sont également connues pour être des co-facteurs des recombinaises pendant la méiose. C'est le cas de BRCA2 et des paralogues de RAD51 (XRCC3, RAD51B, RAD51C, RAD51D). Toutefois le recrutement de l'une ou l'autre de ces protéines reste encore flou et varie selon les espèces (se référer à la liste des protéines identifiées (Mercier et al. 2015)). Chez le riz la protéine MEICA 1 vient d'être identifiée et il semblerait qu'elle joue le même rôle que FLIP1 (Hu et al. 2017). Les modalités du contrôle de l'activité de DMC1 et RAD51 restent à préciser et représentent un domaine de recherche fécond (Wu et al. 2015).

3.2.5. Les voies de réparation des CDB

Après avoir envahi le brin homologue, le simple brin d'ADN sert d'amorce (primer) pour une synthèse d'ADN qui va permettre de réparer la cassure. La poursuite de la synthèse d'ADN nécessite le déroulement progressif de l'ADN double-brin de l'homologue. Cela forme une structure en boucle qui se propage le long de la matrice à copier c'est pour cela que cet intermédiaire de recombinaison est appelé D-loop.

La D-Loop est un carrefour dans la réparation des CDBs (Allers and Lichten 2001). En effet, cet intermédiaire peut être pris en charge par différents complexes protéiques, en compétition les uns avec les autres, pouvant former différents intermédiaires de recombinaison qui selon leur la façon dont ils seront stabilisés ou résolus donneront au final des COs ou des NCOs (Figure 17).

Deux classes distinctes de COs coexistent chez la plupart des eucaryotes. La première dépend des protéines ZMM (§3.2.5.1), l'autre des protéines MUS81 (§3.2.5.2). Ces protéines sont fonctionnellement très conservées. Les COs qui résultent de l'une ou l'autre de ces deux voies sont appelés COs de Classe I et COs de Classe II. Ces deux voies ne forment pas des COs à partir des mêmes intermédiaires de recombinaison et présentent des caractéristiques contrastées vis-à-vis du phénomène d'interférence. Ainsi les COs de classe I sont dits "interférents" par opposition aux COs de Classe II appelés "non interférents". L'interférence se définit comme le phénomène qui fait que la formation d'un CO à un endroit donné du chromosome inhibe la formation d'autres COs à proximité (Berchowitz and Copenhaver 2010) (§ 3.3.2).

3.2.5.1. La voie ZMM de formation des COs

La voie ZMM, ou voie I de formation des COs, est la voie principale de formation des CO chez la plupart des eucaryotes et la seule voie chez *C.elegans* (Hillers et al. 2003). Le nom de cette voie provient des sept protéines suivantes : les protéines ZIP1-ZIP2-ZIP3-ZIP4, l'hétérodimère MSH4-MSH5 et l'hélicase MER3. Ces protéines identifiées dans un premier temps chez *S. cerevisiae* ont des homologues chez la plupart des espèces (Lynn et al. 2007). Chez le riz, un certain nombre de ces protéines ont été identifiées ces dix dernières années c'est le cas de ZEP1 (homologue de Zip1), OsHEI10 (homologue de Zip3), OsZIP4, OsMSH5 et OsMER3 (Wang et al. 2009, 2010b, 2012; Shen et al. 2012; Luo et al. 2013c). En cytogénétique l'anticorps HEI10 est très utilisé car il est spécifique des COs de classe I.

Selon le modèle actuellement en vigueur (Szostak et al. 1983; Allers and Lichten 2001; Mercier et al. 2015; Lambing et al. 2017), la formation de CO de classe I dépend de la mise en place d'une structure stable appelée double jonction de Holliday (djH) (Figure 17 e) suivie de la résolution de la djH en CO. La stabilisation complète se fait grâce à la capture du second brin (*second end capture, SEC*) par complémentarité de séquence entre le brin complémentaire et la deuxième extrémité 3' flanquante. Une nouvelle synthèse suivie d'une ligation aboutit à la formation d'une dHj. La dHj est alors résolue de façon asymétrique ce qui conduit à la formation d'un CO de classe I (Figure 17) (Allers and Lichten 2001).

La formation puis la stabilisation de la djH repose sur les protéines ZMMs (Börner et al. 2004). C'est pour cette raison que les CO de classe I sont aussi appelés CO de la voie ZMM. Dans tous les cas l'absence de l'une de ces protéines entraîne la réduction du nombre de COs

formés. Cependant, bien que toutes ces protéines soient requises pour avoir un niveau de COs équivalent aux individus sauvages, la totalité des djH prises en charge par les protéines de la voie ZMM n'aboutit pas à des COs. En effet, par des approches cytologiques, il a été observé chez *A. thaliana* et/ou *O. sativa* que les foyers des protéines ZIP4, MHS4, MSH5, MER3 et HEI10 excèdent largement le nombre de COs en début de Prophase I de méiose ($\geq$ facteur 10) (Figure 17). Le nombre de foci diminue ensuite drastiquement jusqu'à disparaître entre les stades Pachytène et Diacinèse, où seuls quelques foyers HEI10 persistent et co-localisent au niveau des chiasmas qui sont, rappelons-le, la signature cytologique des COs de classe I (Wang et al. 2012; Chelysheva et al. 2012).

Cette résolution des djH est réalisée par les protéines MLH1 et MLH3 qui bien que non identifiées au départ parmi les protéines ZMMs font partie des protéines impliquées dans la voie de formation des CO de classe I (MutL chez *E. coli*) (Jackson et al. 2006). Les protéines MLH1 et MLH3 sont impliquées dans la résolution spécifiquement asymétrique des dHj pour promouvoir la formation des CO de classe I (Zakharyevich et al. 2012). Il semblerait également que les résolvases mitotiques soient impliquées dans la résolution des djH méiotiques bien que leur rôle respectif ne soit pas encore déterminé.

L'ordre dans lequel les protéines ZMM jouent leur rôle pour promouvoir et stabiliser les dHj reste peu connu. Une étude de l'interdépendance entre ces protéines pour leur localisation chez le riz a mis en évidence que MSH5 agirait avant ZIP4 (Luo et al. 2013c) qui lui-même agirait avant MER3. Toujours chez le riz, la localisation de HEI10 ne dépend pas de MER3 jusqu'au stade pachytène, mais la formation des gros foyers spécifiques des CO ne se forment pas en l'absence de MER3 (Wang et al. 2012).

Chez *A. thaliana* et *O. sativa*, la mutation de l'un ou plusieurs gènes de la famille ZMMs aboutit à une forte baisse du nombre de CO mais ce nombre n'est pas réduit à zéro (Börner et al. 2004; Chelysheva et al. 2012; Luo et al. 2013c). A l'exception de *S. pombe* et autres contre exemples, les COs de classe I constituent la classe majoritaire des COs formés. Les COs de classe I représentent 70% des COs chez *S. cerevisiae*, 85-90% chez *A. thaliana* et *O. sativa*, et la totalité des COs chez *C. elegans*.

3.2.5.2. La voie MUS81 de formation des COs

Suite à l'observation de COs résiduels chez les mutants *zmm* d'*A. thaliana*, il a été déduit qu'une seconde classe de COs existait, indépendante cette fois des protéines du groupe ZMM (Mercier et al. 2005). Les COs de Classe II (ou non-ZMM CO) se trouvent être minoritaires chez les différents organismes étudiés à l'exception de *S. pombe* où c'est la seule voie de formation de CO (Smith et al. 2003).

MUS81 est une protéine à domaine nucléase qui fonctionne en tandem avec une protéine régulatrice, EME1 chez les mammifères et MMS4 chez la levure. Chez *A. thaliana* il existe

deux homologues EME1A et EME1B qui ont des fonctions endonucléases différentes (Geuting et al. 2009). Il est aujourd'hui considéré que MUS81 est capable de prendre en charge un grand nombre d'intermédiaires branchés de l'ADN, considérés comme toxiques pour la cellule d'un point de vue intégrité du génome, et qu'elle serait l'endonucléase "à tout faire" pour résoudre les molécules complexes non dissolues et former des CO (Oh et al. 2008; Mercier et al. 2015) (Figure 17).

On sait peu de choses concernant les mécanismes moléculaires régulant la formation des CO de classe II chez les plantes. Seul le gène *MUS81* est caractérisé. Chez les mutant *Atmus81* le nombre de CO est réduit de 10% par rapport au sauvage (Berchowitz et al. 2007) et dans un fond mutant *zmm* on estime que seulement un tiers des COs résiduels est éliminé (Mercier et al. 2015). Cette donnée implique qu'il puisse exister une troisième voie de formation des COs et que MUS81 n'est pas le seul responsable de la formation des COs de classe II. Les facteurs protéiques impliqués dans cette troisième voie putative ne sont pas connus. Cependant les nucléases GEN1 et SLX1, identifiées chez *S. cerevisiae* comme étant impliquées dans une voie II de formation des COs, sont de bons candidats puisqu'elles ont des homologues chez les plantes mais elles n'ont pas été fonctionnellement caractérisées (Mercier et al. 2015). Reste à élucider, dans le futur, comment tout cela fonctionne en détail.

3.2.5.3. Les voies de formation des Non Crossing-Overs (NCO)

Chez la plupart des eucaryotes, il y a beaucoup plus de CDBs que de COs. Par exemple chez *A. thaliana*, 220 à 240 CDBs méiotiques sont réparées en une dizaine de COs soit en moyenne deux par paire de chromosomes et par méiose (Mercier et al. 2015) alors que chez le maïs on estime que la méiose donne lieu à 500 CDBs dont seulement 20 aboutissent à des COs (Franklin et al. 1999).

Les CDBs qui ne sont pas maturées en COs forment des NCOs. Les NCOs constituent le produit principal de réparation des CDBs (Mézard et al. 2007) et se caractérisent par l'absence d'échange réciproque entre les chromosomes homologues. Comme les CO, les NCOs peuvent être générés par plusieurs voies faisant intervenir ou non la djH et impliquent des protéines encore mal connues à ce jour. Dans tous les cas, quelle que soit la voie de formation du NCO, une resynthèse d'ADN prolongeant le simple brin invasif est réalisée via le brin complémentaire homologue. Ainsi, un fragment de quelques centaines à quelques milliers de paires de bases est transféré de manière non réciproque de l'homologue vers les brins soumis à la CDB ; on parle de conversion génique.

Selon le modèle initialement proposé par Szostak (Szostak et al. 1983), les NCOs résulteraient de la résolution des djH par un clivage symétrique et seraient ainsi dépendant des protéines ZMM (Figure 17). Cependant, chez des mutants *zmm* de *S. cerevisiae*, le nombre de NCOs varie peu, indiquant que cette voie ne contribue que dans une moindre mesure à leur formation

(Hunter et al. 2001; Börner et al. 2004). Il est ainsi considéré à ce jour que les NCOs sont majoritairement issus d'une autre voie, intervenant suite à la formation de la D-Loop et avant celle de la djH, appelée SDSA (pour "Synthesis-Dependent Strand Annealing") (Figure 17) (Pâques and Haber 1999). Lors de celle-ci, le simple-brin envahissant la molécule homologue se dissocierait de sa matrice complémentaire, avant la SEC (pour "*second end capture*", capture du second brin), puis se réassocierait à l'autre extrémité de la CDBs.

Chez la plupart des organismes, la voie SDSA serait sous le contrôle de facteurs anti Crossing-Overs tels que SGS1 chez *S. cerevisiae* (*Small growth suppressor 1*), également l'homologue de BLM chez l'humain (*Bloom syndrome*), la protéine RTLE-1 chez *C. elegans* et FANCM chez *A. thaliana* et *S. pombe* (Séguéla-Arnaud et al. 2015). Sur la base d'observations en mitose, il a été montré qu'une partie des NCOs pouvait également provenir de la dissolution des djH (Figure 17) (Youds and Boulton 2011). Dans ce modèle, des activités hélicase et topoisomérase sont requises et semblent portées par le complexe protéique SGS1-TOP3-RMI1 chez *S. cerevisiae* et RECQ4A-TOP3 α -RMI1 chez *A. thaliana* (Hartung et al. 2008).

3.2.6. Les voies de régulation des COs : les gènes anti-COs

Récemment des cribles supprimeurs réalisés chez *A. thaliana* dans une population de mutants constituée dans un fond génétique *zmm* ont permis d'identifier 3 groupes de gènes qui régulent négativement la formation des COs. Ce criblage est basé sur le fait qu'une mutation qui augmenterait la fréquence de CO doit restaurer la fertilité chez un mutant qui ne forme pas assez de CO (i.e. mutant *zmm* d'*A. thaliana* totalement stérile) (Crismani et al. 2012). Les trois "voies" anti-CO identifiées sont premièrement l'hélicase FANCM et ses co-facteurs MHL1 et MHL2 (Crismani et al. 2012; Girard et al. 2014) puis l'AAA-ATPase FIDGETIN-LIKE1 (Girard et al. 2015) et enfin le complexe RTR avec les hélicases RECQ4A, RECQ4B, la TOPOISOMERASE3 α et RMI1 (Bonnet et al. 2013; Séguéla-Arnaud et al. 2015).

3.2.6.1. L'hélicase FANCM

La protéine Fanconi Anemia Complementation Group M (FANCM) est l'un des 17 gènes impliqués dans la voie Fanconi Anemia (FA). La voie FA a été initialement identifiée chez l'homme dans le cadre de la recherche de mutations associées à l'Anémie de Fanconi, qui est une maladie génétique se caractérisant par une instabilité génomique et une prédisposition au cancer. FANCM est une protéine qui interagit directement avec l'ADN et qui est capable de débobiner les fourches de réplication et les djH grâce à son activité hélicase, et qui sert aussi de plateforme d'assemblage au complexe protéique formé par les autres protéines de la voie FA. De plus, la protéine FANCM apparaît conservée, et divers homologues de cette protéine ont été identifiés chez les eucaryotes. Chez *A. thaliana*, on montre que la protéine FANCM et ses cofacteurs MHF1-MHF2 limitent la formation des crossing-overs méiotiques (Girard et al. 2014). Dans ce cas, on pense que l'hélicase FANCM et ses deux co-facteurs MHF1 et MHF2

pourraient déconnecter les intermédiaires post-invasion et promouvoir la formation de NCO à travers la voie SDSA (Crismani et al. 2012; Girard et al. 2014). Chez le mutant *Atfancm*, le nombre de COs est augmenté d'un facteur 3 par rapport au sauvage dans un fond génétique consanguin et 1.1 en fond hybride. Chez le mutant *Atfancm* $-/-$ les bivalents se forment à nouveau normalement et la fertilité est restaurée (base du crible supprimeur dans un fond mutant *zmm*). Ces COs additionnels ne dépendent pas de la voie ZMM mais de la voie de réparation MUS81 car les COs surnuméraires sont des COs de classe II, insensibles à l'interférence.

3.2.6.2. L'AAA-ATPase FIGETIN-LIKE1

La seconde protéine anti-CO isolée à la suite du crible supprimeur chez *A. thaliana* est la voie l'AAA-ATPase FIDGETIN-LIKE1 (FIGL1) (Girard et al. 2015). Cette ATPase est hautement conservée entre les organismes. Chez l'humain il a été montré qu'elle interagit directement avec RAD51 et qu'elle est nécessaire à la réparation efficace de l'ADN dans les cellules U2OS médiée par la RH (Yuan and Chen 2013). En 2015, il a été montré que cette protéine limitait la formation des COs de classe II et agissait indépendamment de FANCM puisque le double mutant *fancm/figl1* augmentait le nombre de COs d'un facteur 6 par rapport au sauvage dans un fond génétique consanguin (lignée pure). Cette augmentation du nombre de CO est supérieure à celle observée chez les simples mutants *figl1* et *fancm* (respectivement 1,8 et 2,8 respectivement dans un contexte lignée pure). La mutation de FANCM est moins efficace en fond hybride que lorsque la méiose a lieu entre deux fonds génétiques proches (lignée pure). FIGL1 n'est pas sensible contrairement à FANCM au contexte génétique ce qui montre que les modes d'action de FANCM et FIGL1 sont différents. Concernant la distribution des COs surnuméraires chez les simples ou doubles mutants *figl1/fancm*, on observe que la recombinaison augmente préférentiellement dans les régions distales des chromosomes alors qu'aux centromères la recombinaison n'augmente pas (Girard et al. 2015). On observe aussi que, chez un mutant *Atfigl1*, le nombre de foci RAD51 augmente et que les foci DMC1 persistent ce qui laisse supposer que FIGL1 contrôle la dynamique de ces deux recombinaisons. FIGL1 entrave l'interaction entre les chromosomes homologues, ce qui suggère que la protéine FIGL1 contrecarre l'invasion du brin inter-homologue médiée par DMC1 / RAD51 pour limiter la formation des COs (Girard et al. 2015). Très récemment et toujours chez *A. thaliana*, une équipe a montré que FIGL1 avait un nouveau partenaire nommé FLIP pour "FIDGETIN-LIKE-1 INTERACTING PROTEIN" qui l'aidait à réguler l'étape d'invasion, étape cruciale de la RH (Fernandes et al. 2017a).

3.2.6.3. L'hélicase RECQ

La troisième voie anti-CO mise en évidence chez *A. thaliana* a été découverte par l'étude des mutants *top3a-R640X* (codon stop à la position 640) et *recq4a recq4b* (Séguéla-Arnaud et al. 2015). La fréquence de CO est augmentée d'un facteur 1,5 chez le simple mutant *top3a-*

R640X et d'un facteur 6 chez le double mutant *recq4a recq4b*. L'étude montre aussi que cette voie est indépendante de celle de l'hélicase FANCM. *RECQ* chez les plantes, est l'homologue du gène *BLM* humain et du gène *SGS1* chez la levure. Chez *A. thaliana* et *Oryza*, il existe 7 gènes *RECQ-like* (Hartung and Puchta 2006). Le gène *BLM* a deux orthologues chez *A. thaliana* qui sont *AtRECQ4A* et *AtRECQ4B* alors qu'*Oryza* n'en a qu'un : *OsRECQ14* (Kwon et al. 2013). La duplication est récente et semble avoir eu lieu uniquement chez les *Brassicaceae* (Séguéla-Arnaud et al. 2015). Kwon a montré qu'*OsRECQ14* jouait un rôle important en mitose dans le maintien de l'intégrité du génome pendant la réplication de l'ADN au niveau des méristèmes (Kwon et al. 2013).

Ces protéines *RECQ4* et *TOP3α* appartiennent au complexe protéique RTR qui comprend, en plus des hélicases *RecQ* et de la Topoisomérase IIIα, la protéine *RMI1* (Mankouri and Hickson 2007; Bonnet et al. 2013). Bonnet montre que la protéine *RMI1* comporte des domaines essentiels au "DNA cross-link repair" lors de la méiose (Bonnet et al. 2013; Hunter 2015). Le complexe RTR est impliqué dans la dissolution d'un éventail de structure d'ADN et a même été nommé "*dissolvasome*" (Mankouri and Hickson 2007). Ce complexe bien conservé correspond aux groupes de protéines appelé STR chez la levure (pour *SGS1-TOP3-RMI1*) ou BTB chez l'Homme (pour *BLM-TOP3-BLAP75*) et s'appelle RTR chez les autres espèces. Le rôle de ces protéines a été majoritairement étudié dans un contexte somatique. Il a été montré que *SGS1* pouvait dissocier les D-loop *in vitro*, de façon analogue à *RTEL1*, probablement pour limiter la RH dans les cellules somatiques (van Brabant et al. 2000). Lors de la méiose, la mutation *sgs1* mène à une augmentation du nombre de CO, et supprime partiellement les défauts d'un mutant *zmm* (Oh et al. 2008), indiquant un rôle anti-CO pour *SGS1*. En revisitant le rôle de cette protéine, il a été montré que *SGS1* serait capable de désassembler des intermédiaires précoces et/ou complexes et de les entraîner vers la voie SDSA pour former des NCOs (De Muyt et al. 2012).

En résumé, ces données montrent que la formation des COs est limitée par au moins 3 voies indépendantes contrôlées, chez *A. thaliana* et très certainement chez les autres espèces de plantes à fleurs, par les protéines FANCM, *RECQ4* et *FIGL1*. De récentes études sur le cumul des 3 mutations montre que le plus gros effet observé est obtenu lors de la double mutation *recq4 fligl1*, en fond hybride, où la carte génétique de l'hybride passe de 389 à 3037 centiMorgans (cM) ce qui correspond à une augmentation jamais observée auparavant de la fréquence de CO d'un facteur 7,8 (Fernandes et al. 2017b). Il est également montré que la recombinaison dans un fond sauvage est plus élevée dans les méiocytes mâles que les méiocytes femelles (490 cM vs 290 cM) alors qu'en fond double mutant *recq4 fligl1*, la recombinaison est plus élevée lors de la gamétogenèse femelle que mâle (3200 cM vs 2720 cM) (Fernandes et al. 2017b). Toujours plus impressionnant et cette fois en modifiant et la voie I et la voie II de formation des COs, une étude montre que la surexpression de la protéine

AtHEI10 dans un fond double mutant *recq4a recq4b* conduit à une augmentation concomittente du nombre de COs interférents et non interférents (Serra et al. 2017).

Ces données montrent que la réparation des cassures double brin n'est pas un processus linéaire bien hiérarchisé, et que si la formation des COs est particulièrement bien régulée (ci-dessus et chapitre suivant), la réparation des CDBs non désignées pour être des COs pourraient être prise en charge par un grand nombre de facteurs protéiques pour donner des NCOs avec des propriétés très différentes suivant le processus de réparation. L'étude de NCOs est plus délicate que celle des COs car il est difficile de les détecter (Mercier et al. 2015).

3.3. Comment détecter les événements de recombinaison ?

Deux types d'approches permettent de détecter un crossing-over l'une directe et l'autre indirecte. Dans les deux cas, différentes méthodes (protocoles) sont possibles. La cytogénétique permet de visualiser directement le CO sur le chromosome alors que la seconde méthode consiste à analyser les produits de la méiose dans une descendance en utilisant des marqueurs moléculaires.

3.3.1. Techniques de cytogénétique

En fin de prophase I de méiose, différentes techniques de cytogénétique permettent de rendre visible les COs. On peut soit faire de l'immuno-localisation (anticorps spécifique des COs de Classe I MLH1 ou HEI10 par exemple), soit observer directement les méiocytes en microscopie électronique et voir directement les chiasmas (revue par Mezard 2015).

Il a été montré que le comptage du nombre de foci HEI10 (anticorps spécifiques des COs de Classe I) est une méthode cytologique robuste pour estimer le nombre de CO de Classe I (Wang et al. 2012). La combinaison de l'immuno-localisation et de la microscopie électronique à transmission permet de différencier les deux classes de CO (Anderson et al. 2014). En effet, les nodules de recombinaisons (RN) qui co-localisent avec les foci MLH1 sont des COs de Classe I alors que les autres RN qui ne fluorescent pas sont considérés comme les COs de Classe II. Cependant, ces approches cytogénétiques sont lourdes à mettre en place et requièrent un grand nombre de cellules à analyser. C'est pourquoi la méiose mâle est plus étudiée par cette approche que la méiose femelle (moins de méiocytes). Toutefois, une étude récente sur *A. thaliana* présente un protocole original ouvrant la porte à l'étude des COs en méiose femelle (Escobar-Guzmán et al. 2015). Un dernier inconvénient de la cytogénétique est lié au fait qu'il peut exister un biais lors de l'étude de la distribution des COs, car les distances physiques correspondent à des mesures en μm qui ne prennent pas en compte les différences de compaction de l'ADN le long des chromosomes.

3.3.2. Par établissement de carte génétique dans une descendance

Chez la plupart des organismes, la détection des COs est principalement réalisée via l'étude de descendance en ségrégation, générées par backcross, haploïdes doublés ou autofécondation, en réalisant leur génotypage via des marqueurs moléculaires. Cependant, chez *A. thaliana*, d'autres techniques comme le typage des gamètes mâles ou "pollen typing" (Drouaud et al. 2013) ou l'utilisation de tétrades fluorescentes (appelée mutant quartet) (Berchowitz and Copenhaver 2008) permettent de quantifier les événements de COs et NCOs (PCR et séquençage ou visualisation directe de la fluorescence pour l'autre technique). Enfin d'autres méthodes comme la chromatine immuno-précipitation (ChIP-Seq) de protéines

dirigées vers des protéines impliquées dans la recombinaison peuvent être employées (Choi and Henderson 2015).

De façon plus classique, on étudie la distribution des crossing-overs en construisant une carte génétique à partir de la descendance d'un hybride F1 obtenu par croisement de deux lignées pures. Le principe étant que deux marqueurs proches physiquement sur un chromosome doivent être co-hérités sauf s'il se produit un événement de CO entre ces marqueurs. Cet événement de CO va casser le lien génétique entre les deux marqueurs et ceux-ci vont alors ségréger de façon indépendante. C'est ainsi qu'historiquement, les événements de CO sont déduits de l'analyse du déséquilibre de liaison (DL). Le DL désigne l'association non aléatoire d'allèles entre paires de marqueurs dans une population (§ 3.3.4).

3.3.3. Calcul de la fréquence de recombinaison

La fréquence de recombinaison ($f_r = r / n$; ou r = nombre de gamètes recombinés et n = nombre total de gamètes) entre deux marqueurs est calculée à partir du nombre de recombinants observé dans une descendance. Sa valeur minimale est 0 et sa valeur maximale est 0.5 ce qui correspond à une situation d'indépendance entre les deux marqueurs. La distance génétique (d) entre les marqueurs est déduite de la fréquence de recombinaison en utilisant la formule de Haldane (qui suppose qu'il n'y a pas d'interférence) ou celle de Kosambi (qui suppose qu'il y a de l'interférence). L'avantage des distances, c'est qu'elles sont additives alors que les fréquences de recombinaison ne le sont pas du fait des crossing-overs multiples.

La définition première développée par Morgan de la distance génétique en centiMorgan (d en cM) entre deux marqueurs, positionnés sur le même chromosome, correspond à la fréquence de recombinants dans une population de type F2 ($r = d$). Ainsi pour deux marqueurs non liés, la probabilité que ces deux marqueurs soient associés dans un gamète est de 50% (une chance sur deux) et la distance génétique associée est infinie (Morgan 1912). Pour des marqueurs liés, la fréquence de recombinaison est inférieure à 0.5. L'hypothèse sous-jacente de Morgan dans cette définition est que l'interférence est complète et qu'aucun double-CO ne se forme sur l'intervalle considéré. Haldane introduit en 1919 la possibilité que ces doubles événements puissent se produire (Haldane 1919). Il infère alors que la probabilité d'avoir un double CO est le carré de celle d'avoir un simple CO (pour 10% de simple CO observés, la fréquence de double CO est de 1%), c'est-à-dire que la présence d'un CO n'empêche pas la formation d'un deuxième, et qu'il n'y a aucune interférence. Là aussi, cette mesure s'éloigne de la réalité. Kosambi propose alors en 1943 de prendre en compte l'interférence dans le calcul de la distance génétique d'un intervalle (Kosambi 1943). Il utilise pour cela la variable c de coïncidence créée par Muller (Muller 1916) : prenons trois marqueurs (appelons les A, B et C), la variable de coïncidence (c) est alors définie comme la fréquence de double-CO observés entre A et C rapportée au nombre attendu, obtenu par le produit des fréquences observées dans les intervalles A-B et B-C. Ce paramètre c donne une mesure de l'interférence (c'est en

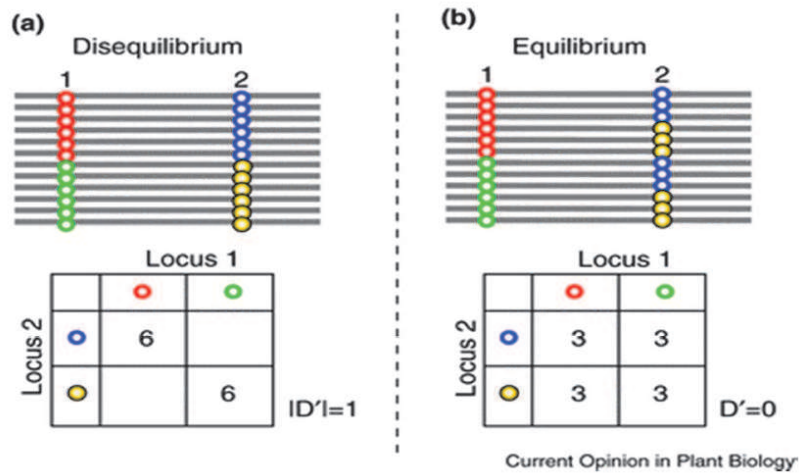


Figure 21 : Représentation schématique du déséquilibre (a) et (b) de l'équilibre de liaison entre deux locus (Flint-Garcia et al., 2003; Guo et al., 2014).

(a) Lorsque le déséquilibre de liaison est total, tous les individus qui possèdent l'allèle rouge au locus 1 ont nécessairement l'allèle bleu au locus 2. (b) Quand il y a équilibre de liaison, l'individu qui possède l'allèle rouge au locus 1 à autant de chance de posséder l'allèle jaune ou l'allèle bleu au locus 2. Les tableaux de contingence respectifs sont fournis et les valeurs du déséquilibre (D') sont données.

fait le coefficient de coïncidence, CoC). Lors du calcul de la distance d'un intervalle (délimité par seulement deux marqueurs), il faut alors estimer la valeur de c , or aucune valeur universelle n'existe. Kosambi a donc estimé la valeur de c chez l'homme et c'est cette valeur qui est utilisée dans son modèle. Pour les autres espèces, des modèles ont été développés pour évaluer la valeur de c d'après les données expérimentales (McPeck et al. 1995). Mais on utilise couramment $c = 2r$ tel que défini par Kosambi. Le modèle de Kosambi est donc identique au modèle de Morgan pour $c = 0$ (aucune interférence), et identique à celui de Haldane pour $c = 1$ (interférence complète). Ces développements théoriques permettent de préciser la notion de centiMorgan. Contrairement à ce qui est souvent conceptualisé, un centiMorgan ne correspond pas à 1% de recombinaison. Même si ces deux quantités sont très proches pour de faibles fréquences, l'égalité n'est pas légitime. En réalité, le centiMorgan pourrait être défini comme la distance séparant deux loci entre lesquels l'espérance du nombre de CO est de un centième (Vienne 1998).

3.3.4. Déséquilibre de liaison et CO

Le déséquilibre de liaison (DL) représente l'association non-aléatoire des allèles à deux marqueurs et résulte d'un déséquilibre gamétique (Flint-Garcia et al. 2003). Le DL est calculé sur une échelle de 0 à 1. La valeur 1 représente un déséquilibre total par rapport à l'association aléatoire, c'est-à-dire une association parfaite entre les allèles. Réciproquement, la valeur 0 représente une ségrégation complètement aléatoire, c'est-à-dire une dissociation complète entre les allèles (Figure 21) (Flint-Garcia et al. 2003; Guo et al. 2014). Différents facteurs peuvent affecter le DL, notamment la recombinaison ou la mutation par la création de nouveaux allèles. Plus le taux de recombinaison sera élevé, plus il abaissera le DL. Au contraire, la sélection favorisera l'association non aléatoire d'allèles et augmentera ainsi le déséquilibre de liaison d'un point de vue local ou à longue distance. De cette liaison entre les allèles découle un autre concept : l'autostop génétique. Si des caractères sont liés physiquement, leur ségrégation a plus de chance de se faire ensemble (liaison génétique). Ainsi, un gène sélectionné sera probablement accompagné d'autres gènes localisés dans son voisinage immédiat à cause (1) du faible nombre de crossing-overs affectant un segment chromosomique de taille réduite ; (2) de la répartition non-aléatoire et non-homogène des crossing-overs. Le déséquilibre de liaison est représenté sous forme de blocs de liaison qu'il y a un grand intérêt à casser (notamment dans les régions centromériques) dans le cas d'un gène cible co-ségrégeant avec un gène délétère ou, à l'inverse, à conserver entier si les gènes qu'ils portent ont un effet favorable. Le contrôle de "activer/désactiver" la recombinaison pour casser ces blocs est donc une voie de recherche étudiée actuellement avec grand intérêt. De plus, l'impact de la recombinaison sur l'association non aléatoire des mutations et l'analyse du polymorphisme génétique actuel permet l'étude de la recombinaison ancestrale. En effet, les crossing-overs vont provoquer des chutes du DL entre polymorphismes liés.

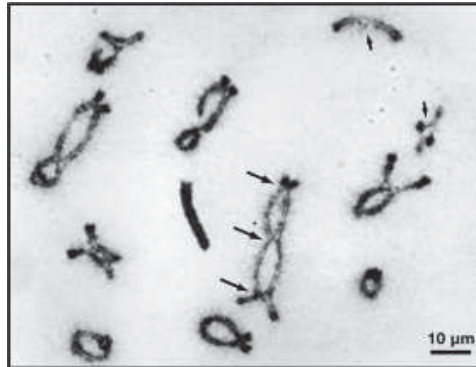


Figure 22 : Le crossing-over obligatoire et l'interférence (Jones et al. 2006).

L'observation de méiocytes mâles de sauterelle (*Schistocerca gregaria* ($2n = 22 + X$)) au stade diplotène permet de voir 11 bivalents autosomiques et un chromosome X (univalent). Les chiasmas (CO) sont souvent multiples et bien espacés sur les bivalents les plus longs (longues flèches noires) alors que les plus petits chromosomes n'ont généralement qu'un seul chiasma obligatoire (flèche courte noire).

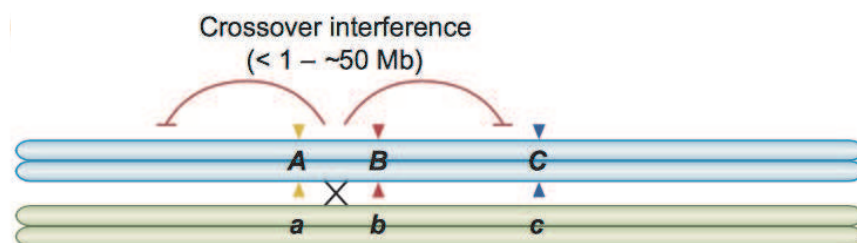


Figure 23 : L'interférence pendant la recombinaison méiotique (Ziolkowski & Handerson 2017).

Le phénomène d'interférence est ici représenté sur des chromosomes homologues répliqués et appariés (bleu et vert). Trois loci sont représentés A, B et C et leurs variants alléliques a, b et c (triangles de couleurs). La croix noire symbolise un crossing-over (CO) entre deux chromatides homologues au niveau des loci A et B. L'occurrence d'un CO dans l'intervalle A-B réduit la probabilité de formation d'un second CO dans les régions adjacentes proches sur une échelle de quelques mégabase (<1-50 Mb).

3.4. Distribution des évènements de recombinaison

La recombinaison méiotique est initiée par la formation programmée d'une CDB de l'ADN. La réparation de cette cassure via la chromatide homologue non-sœur (§3.2.4) aboutit à la formation de COs ou de NCOs. Alors qu'un CO est le résultat d'un échange réciproque d'un long fragment d'ADN entre les chromatides homologues, les NCO résultent de la copie conforme d'une petite portion d'ADN du chromosome homologue sur le chromosome "cassé". En conséquence de quoi un CO modifie les deux chromatides alors qu'un NCO n'en modifie qu'une (Mercier et al. 2015). On parle de brassage génétique dans le cas de la formation des CO et de conversion génique dans le cas de formation de NCO. C'est pourquoi la formation et la réparation des CDBs est soumise à une régulation spatiale et temporelle extrêmement précise.

3.4.1. Nature et distribution des événements de recombinaison

3.4.1.1. Les phénomènes qui régulent le nombre de CO

En dépit de la nature stochastique du CO qui, a priori, peut avoir lieu n'importe où le long du chromosome, les COs sont finement régulés (Wang et al. 2015b). Alors que le nombre de CDBs par noyau est très élevé, le nombre total de CO est très faible, généralement de 1 à 3 par chromosomes. Ainsi le point de vue populaire selon lequel la recombinaison méiotique "brasse" l'information génétique est négligeable puisque le ratio CDBs sur CO montre que la recombinaison méiotique est relativement conservatrice. Jusqu'à la moitié des chromatides peuvent émerger de la méiose avec les haplotypes parentaux (Hunter 2015). Cependant chaque paire de chromosomes homologues subit au moins un CO. C'est un prérequis à la bonne ségrégation des homologues en métaphase et c'est ce que l'on résume par le terme de CO obligatoire (Jones et al. 2006) (Figure 22). En outre, lorsqu'une paire de chromosome subit deux événements de COs ils ont tendance à être largement espacés. Ce phénomène est appelé interférence et a été remarqué pour la première fois en 1913 par Alfred Sturtevant lors de la construction de la première carte génétique (Sturtevant 1915). L'interférence, affecte la formation coïncidente d'un CO de Classe I sur une paire de chromosomes, lorsqu'un CO de cette même classe s'y est déjà produit. Cependant, la formation des COs dit interférents n'affecte pas la position des COs de Classe II qui sont distribués aléatoirement le long des chromosomes (Mézard et al. 2007). Ce phénomène est particulièrement intéressant car il suppose l'existence d'une communication le long des chromosomes sur des dizaines de microns (Figure 23). Bien que l'observation de ce phénomène remonte à une centaine d'années son mécanisme reste largement méconnu (Lambing et al. 2017). La distribution des COs est aussi régulée par un autre phénomène nommé homéostasie. Ce phénomène a été décrit chez le levure et la souris chez qui il a été montré qu'un nombre minimal de CO est maintenu quand bien même le nombre de CDBs par cellule est altéré (Martini et al. 2006).

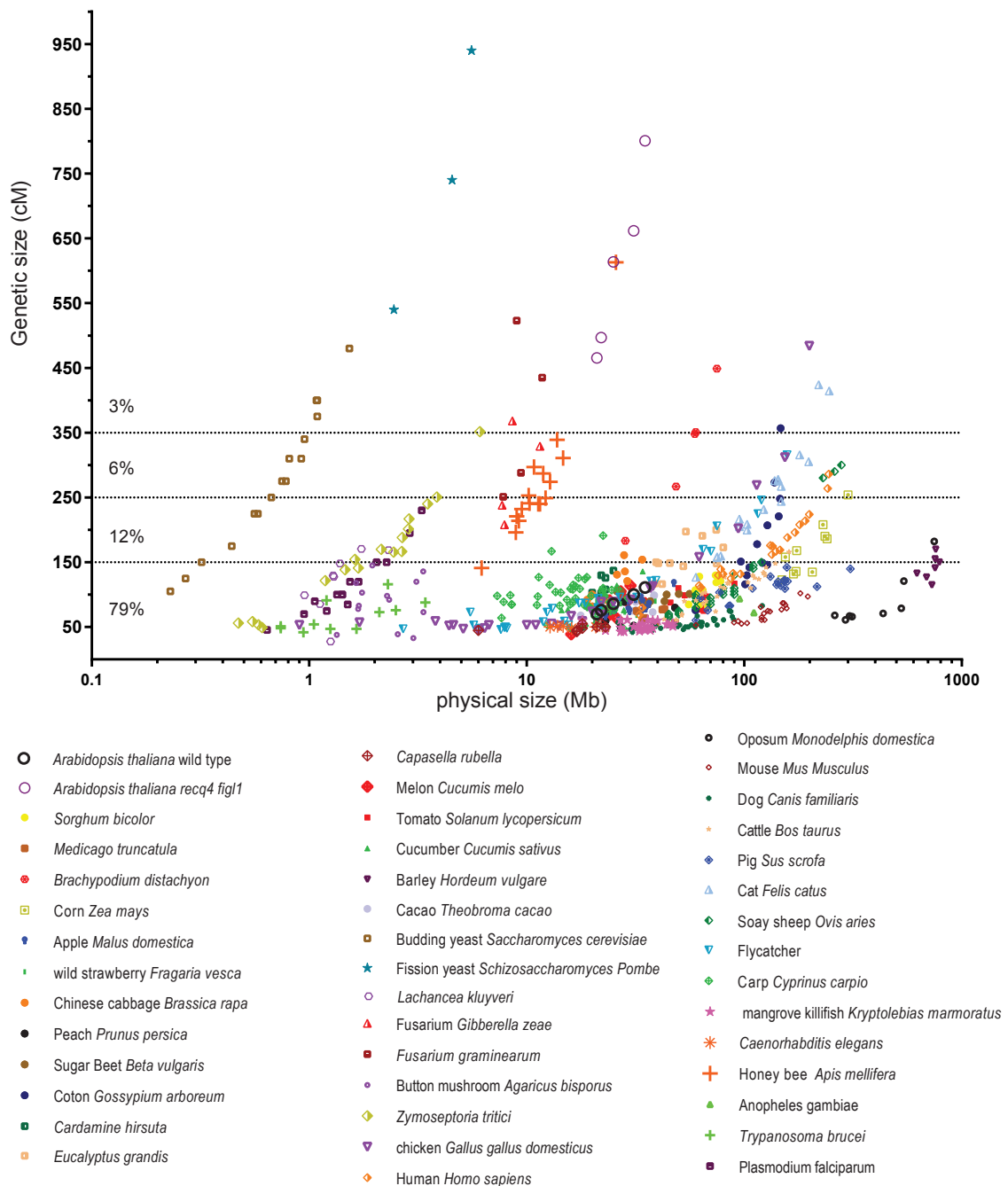


Figure 24 : Nombre moyen de Crossing-Overs (COs) par chromosome et par méiose chez différentes espèces eucaryotes et chez le mutant *Atrecq4 figl1* (Mercier et al., 2015).

Les chromosomes d'une large gamme d'espèces sont positionnés en fonction de leur taille physique (axe X, en Mb, échelle logarithmique) et de leur taille génétique (axe Y, en cM, échelle linéaire). La taille physique repose sur l'assemblage de la séquence génomique de chaque génome et la taille génétique est basée sur une population F2 (ou moyenne homme / femme). Les chromosomes sexuels ont été exclus de l'analyse.

L'homéostasie se réfère aussi au fait qu'il existerait un nombre maximum de CO et que si on augmente le nombre de CO à un endroit, on diminue celui-ci à d'autres. L'interférence peut se mesurer lorsque l'on dispose de données de ségrégation sur deux intervalles mitoyens, comme c'est le cas dans l'expérience de Sturtevant ou dans des systèmes d'analyse de tétrades de levures, ou les tétrades fluorescentes (fluorescent tagged lines, FTL) développé par Copenhaver chez *Arabidopsis* (Berchowitz and Copenhaver 2008). Plusieurs indices mesurant l'interférence ont été développés. Les plus usités sont le coefficient de coïncidence (*Coefficient of coincidence*, CoC) et le ratio d'interférence (*Interference ratio*, IR).

3.4.1.2. Nombre moyen de CO par chromosomes selon les espèces

A l'échelle du chromosome, Mercier montre que 76% des bivalents, chez les eucaryotes, présentent 1 à 3 COs et ce quelle que soit la taille physique du chromosome (Figure 24). Par exemple, le chromosome 3B du blé qui a une taille physique d'environ 1 Gb a en moyenne 3 COs par méiose comme le plus long chromosome d'*A. thaliana* (31 Mb) et le plus petit chromosome de la levure (0,3 Mb). On sait que le nombre de CO n'est en aucun cas limité par le nombre de CDB et on estime que chez les plantes le nombre de CDB est de 10 à 50 fois plus élevé que le nombre de CO (Mercier et al. 2015). Comme nous l'avons vu de nombreux mécanismes moléculaires et autres phénomènes (i.e. Interférence et homéostasie) limitent la réparation des CDBs en CO (§3.2.6). La raison d'un si faible nombre de CO n'est toujours pas connue. L'une des rares hypothèses formulées est qu'un trop grand nombre de CO pourrait conduire à une mauvaise ségrégation des chromosomes homologues pendant la méiose (Louis and Borts 2003). Cette hypothèse est infirmée par le fait que certaines espèces exceptionnelles telles que *S. cerevisiae* ou l'abeille sont capables d'avoir une ségrégation correcte des chromosomes en méiose en dépit d'une densité élevée de CO par chromosome (plus de 10) (Figure 24). De même lorsqu'on augmente artificiellement chez *A. thaliana* le nombre de CO par méiose (3 à 9 fois plus que chez le sauvage) la ségrégation des chromosomes en méiose se fait normalement (Crismani et al. 2013; Séguéla-Arnaud et al. 2015; Girard et al. 2015). Bien que cela ne soit pas expliqué, le nombre optimal de CO par chromosome qui ait permis aux organismes d'évoluer sans que ce soit nuisible est de 1 CO.

3.4.1.3. Structure du génome et distribution des COs

Dans presque toutes les espèces, la distribution des événements de RH n'est pas homogène le long des chromosomes et ce quelle que soit l'échelle considérée (Lambing et al. 2017) (Figure 25). A l'échelle la plus haute, on observe des portions de chromosomes où le taux de CO est inférieur à la moyenne du génome. Ainsi, *A. thaliana*, le blé ou l'Humain, 80% des événements de recombinaison sont distribués sur moins d'un quart du génome (Mercier et al. 2015). A l'échelle la plus fine, les COs et les NCOs se regroupent dans des régions de quelques kilo-bases appelées points chauds de recombinaison ("*hotspots*"). La localisation de ces points chauds reflète la position des CDBs méiotiques. Une partie des mécanismes qui

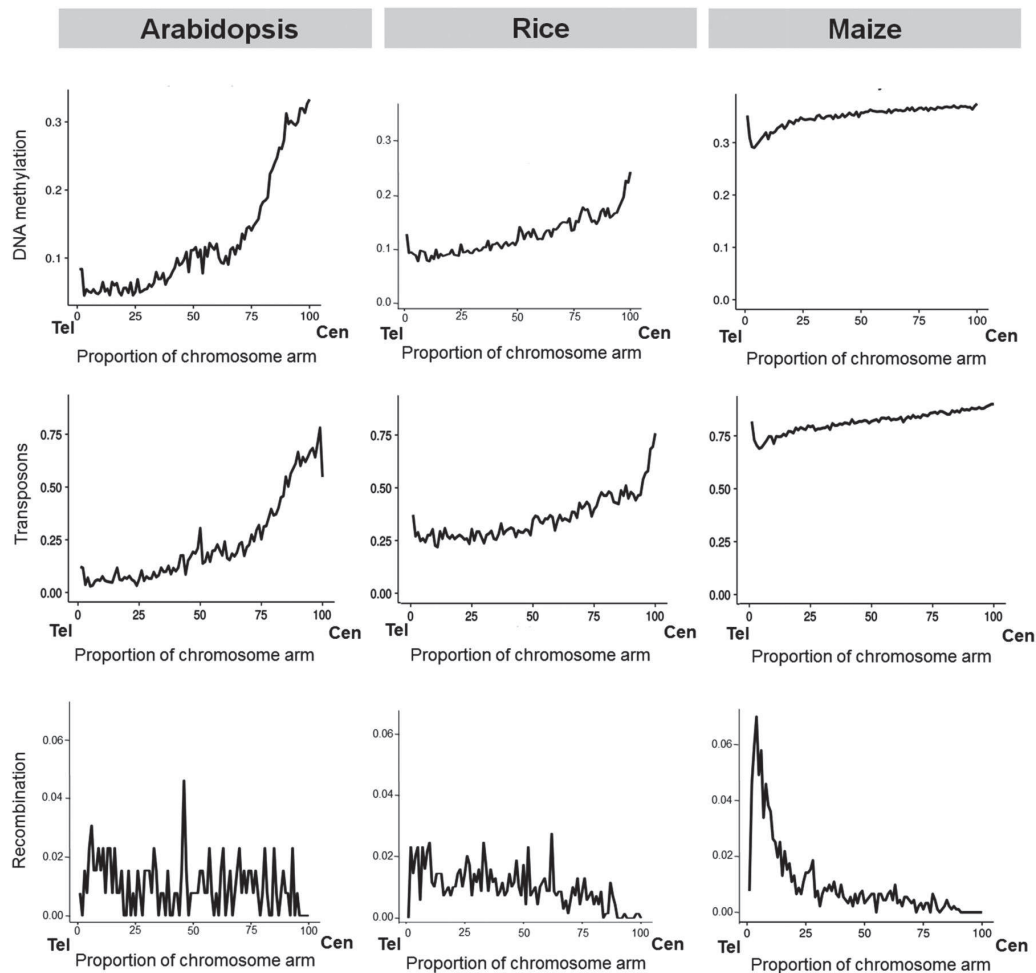


Figure 25 : La structure du génome et la distribution des COs diffère entre *A. thaliana*, le riz et le maïs (Lambing et al. 2017).

Distribution de la méthylation de l'ADN, des éléments transposables et de la recombinaison le long des bras chromosomiques chez 3 espèces, *A. thaliana*, le riz et le maïs.

(Tel : télomères ; Cen : centromères)

gouvernent la localisation des CDBs ont récemment été décryptés chez *S. cerevisiae*, la souris et l'Homme et montrent que certains mécanismes sont communs et d'autre spécifiques aux différents organismes (Mercier et al. 2015). Chez le blé et le maïs, les COs sont plus fréquents sur la partie distale des chromosomes qu'au centre des bras chromosomiques et chez toutes les espèces, les régions centromériques sont dépourvues de CO. En revanche, chez la plante modèle *A. thaliana*, les hotspots de COs se répartissent sur l'ensemble des chromosomes sauf bien sûr aux centromères. On suppose que cette différence est due à la structure du génome (genome organisation). Chez *A. thaliana*, il y a relativement peu de transposons et ils sont souvent dans les régions péri-centromériques et centromériques alors que, chez les céréales où la taille du génome est bien plus importante, les transposons sont nombreux et leur distribution est plus uniforme (Figure 25). On sait aussi que les COs surviennent le plus souvent dans les gènes ce qui suggère un lien entre la structure de la chromatine et la formation des COs. En effet, à l'échelle de la chromatine, l'analyse des hotspots de COs chez le riz, le maïs et le blé montre qu'ils sont plus souvent dans des régions pauvres en transposon et peu méthylées mais riches en gènes (Lambing et al. 2017). Chez *A. thaliana*, on montre que les COs dans les méiocytes mâles sont associés à des marques de chromatine active comme H2A.Z, H3K4me3, une faible méthylation de l'ADN et à une faible densité de nucléosomes (Yelina et al. 2012). Chez la souris et l'Homme, la formation des points chauds de CDBs est associée à la méthyltransférase PRDM9 qui peut grâce, à un domaine en doigt de zinc, se lier à un motif ADN particulier puis le domaine SET de cette protéine permet de triméthyliser la Lysine 4 de l'Histone H3 des nucléosomes présents aux sites de fixation de la protéine (H3K4me3) (Ségurel et al. 2011). Chez les plantes, la sous famille PRDM9 a été perdue indiquant que des mécanismes différents existent en fonction des espèces (Mercier et al. 2015).

Il est également très connu que les COs sont réprimées au niveau des centromères et régions péri-centromériques (Sax 1932). En effet ces régions contiennent des séquences répétées qui permettent aux kinétochores de s'attacher aux chromatides sœurs et assurent la ségrégation correcte des chromosomes (Talbert and Henikoff 2010) (Figure 26). Chez la levure on montre qu'en augmentant la recombinaison avec *recq* dans le génome, les COs qui ont lieu dans les régions péri-centromériques conduisent à la mort des spores (Vincenten et al. 2015).

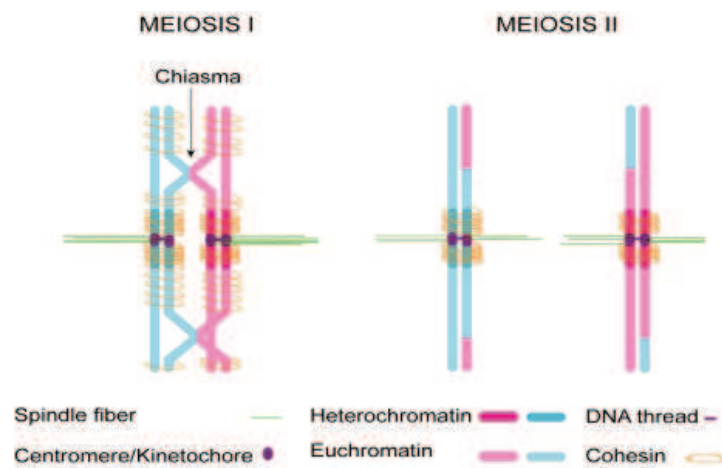


Figure 26 : Connection des chromosomes en méiose et attachement des kinétochores (Talbert 2010).

Les kinétochores s'attachent aux chromosomes homologues au niveau des centromères. Et les fibres du fuseau (spindel fiber) qui viennent de chaque pôle de la cellule s'accrochent aux kinétochores. Les homologues sont maintenus ensemble par des chiasmas, dans lesquels les chromatides recombinantes se croisent. Les chromatides sœurs sont maintenues ensemble par des cohésines. La cohésion est relâchée en deux étapes : sur les bras des chromosomes pour résoudre les chiasmas et séparer les homologues dans la première division méiotique puis autour des centromères pour séparer les sœurs dans la deuxième division méiotique. Ceci est rendu possible par la mono-orientation des kinétochores en première division de méiose puis la bi-orientation des kinétochores en seconde division de méiose.

Partie II : ABOLIR LA RECOMBINAISON

Partie II : Abolir la recombinaison

4. L'APOMEIOSE CHEZ LE RIZ

4.1. Introduction

Comme nous l'avons vu en introduction, l'apomixie gamétophytique résulte de la formation d'un sac embryonnaire à partir d'une cellule diploïde de l'ovule et contourne la méiose. Ce développement du sac embryonnaire par mitose est appelé apoméiose. La formation de l'embryon est indépendante de la fécondation tandis que la formation de l'albumen requiert ou non la fécondation. Il existe deux types de développement d'un sac embryonnaire apoméiotique, la diplosporie et l'aposporie, selon l'origine de la cellule diploïde précurseur qui va donner naissance au sac embryonnaire. Cette cellule précurseur est la cellule mère de la mégaspore dans le cas de la diplosporie, et une cellule diploïde somatique positionnée de façon adjacente à la cellule mère de la mégaspore dans le cas de l'aposporie. Des embryons issus de reproduction sexuée et d'aposporie peuvent également coexister à l'intérieur du même ovule chez les espèces du genre *Brachiaria* (Hand et al. 2014).

L'apomixie est considérée comme une déviation de l'expression spatiale et temporelle de la voie sexuée en recrutant des gènes intervenant dans la reproduction sexuée. On peut en déduire que la conversion d'espèces à reproduction sexuée en espèces apomictiques pourrait être possible. Cette hypothèse est consolidée par l'identification de mutants mimant différents aspects de l'apomixie chez les espèces non apomictiques : apoméiose, parthénogenèse de l'embryon ou développement autonome de l'albumen (Barcaccia and Albertini 2013).

La mutation de certains gènes intervenant dans la méiose peut conduire à la formation de sacs embryonnaires sans méiose qui rappellent l'aposporie. Un des exemples les plus connus est celui du mutant *elongate 1* du maïs qui produit des sacs embryonnaires diploïdes en ne s'engageant pas dans la seconde division de méiose (Barrell and Grossniklaus 2005). De même, le mutant *dyad* identifié chez *A. thaliana* produit des gamètes non réduits suite à une simple division équationnelle (Mercier et al. 2001; Ravi et al. 2008). La mutation de gènes ayant un rôle connu dans la régulation épigénétique de la spécification cellulaire de l'ovule et du développement du sac embryonnaire, comme les gènes de la famille *ARGONAUTE* (*AGO*) chez les plantes sexuées, peuvent conduire à un changement dans le nombre de cellules ayant la capacité d'initier un développement en sac embryonnaire (Olmedo-Monfil et al. 2010; Singh et al. 2011). Chez *A. thaliana*, le mutant *ago9* produit de multiples cellules ayant la capacité de former des sacs embryonnaires à l'intérieur d'un même ovule, un phénotype

rappelant l'aposporie (Olmedo-Monfil et al. 2010). Des phénotypes similaires ont été obtenus avec la mutation de gènes intervenant dans la voie de régulation des petits ARN comme *rdr6* et *sgs3* (Olmedo-Monfil et al. 2010). De même l'inactivation des DNA méthyl-transférases DMT102 et DMT103 conduit aussi à des phénotypes copiant l'apomixie qui comprennent la formation de gamètes diploïdes et de nombreux sacs embryonnaires à l'intérieur d'un même ovule (Garcia-Aguilar et al. 2010).

Chez le maïs, la mutation du gène *AGO104* conduit également à un phénotype rappelant l'apomixie diplosporique avec une seule cellule mère de mégaspore par ovule entrant en mitose plutôt que méiose et aboutissant à la formation de gamètes diploïdes (Singh et al. 2011). Les gènes *AGO* sont également impliqués dans l'identité de la lignée germinale chez le riz, comme le gène *MEL1* (*Meiosis arrested at Leptotene 1*) (Nonomura 2006). Des mutations dans d'autres gènes peuvent conduire au développement parthénogénétique de l'embryon ou de l'albumen sans fécondation. Nombre de ces gènes sont associés avec le complexe de remodelage de la chromatine du groupe Polycomb, notamment le PRC2 (*Polycomb Repressive Complex 2*). L'inactivation des gènes *MEDEA*, *FIS2*, *FIE* (*Fertilization independent Endosperm*) ou *MSI1* (*Multicopy Suppressor of IRA1*) conduit au développement sans fécondation de l'albumen et pour *MSI1* à l'initiation d'un développement parthénogénétique de l'embryon sans toutefois formation de graines viables (Guitton and Berger 2005).

L'ensemble de ces observations converge en faveur d'une différenciation entre espèces apomictiques et non apomictiques ayant un contrôle épigénétique au travers d'altérations de l'état chromatinien. Il a même été suggéré que l'apomixie représenterait un silencing épigénétique réversible de la voie sexuée. La coexistence de la reproduction sexuée et apomictique chez certaines espèces suggère en effet que l'apomixie est une voie superposée sur la voie sexuée (Hand et al. 2014).

L'existence d'espèces apparentées apomictiques chez les céréales, et plus largement chez les graminées, a permis d'entreprendre un criblage chez les espèces sauvages ancêtres du riz cultivé comme source d'apomixie (Rutger 1992; Khush 1994). Afin de mettre en évidence un développement diplosporique du sac embryonnaire, les criblages ont reposé soit sur la détection de grains poly-embryonnés qui auraient dérivé d'une embryogenèse adventive, soit sur l'observation par fluorescence de callose combinée à un éclaircissement du pistil. Les criblages, d'abord réalisés chez les espèces diploïdes du génome AA, ont ensuite été étendus à des accessions d'autres génomes du genre *Oryza*, incluant les espèces polyploïdes connues pour présenter une apomixie plus fréquente. En dépit de ces efforts intenses, aucune source naturelle d'apomixie n'a été identifiée dans le genre *Oryza* à partir de ces cribles. L'alternative est donc de créer l'apomixie chez le riz grâce à la mutation de certains gènes et/ou l'expression de transgènes.

Le premier composant de l'apomixie est l'apoméiose qui est la formation de gamètes fonctionnels diploïdes et clonaux. Des plantes apoméiotiques ont été générées grâce à la mutation de gènes de méiose. En effet, de nombreux mutants de méiose identifiés chez *A. thaliana*, le maïs et le riz, produisent des gamètes non réduits, mais à une faible fréquence pour la plupart d'entre eux, et ces mutations conduisent souvent à une stérilité mâle ou femelle importante. L'apoméiose repose sur la conversion de la méiose en mitose. Nous avons vu que la méiose diffère de la mitose par : 1- l'appariement et la recombinaison des chromosomes homologues ; 2- l'orientation monopolaire du kinétochore conduisant à la migration des chromatides sœurs vers le même pôle à la première division réductionnelle de la méiose ; 3- en une seconde division équationnelle ressemblant à une mitose. Des plantes apoméiotiques fertiles produisant des gamètes diploïdes non recombinaisonnels à une forte fréquence ont été obtenues chez *A. thaliana* grâce à la combinaison de trois mutations empêchant à chacun de ces trois processus de survenir et transformant ainsi la méiose en mitose (D'Erfurth et al. 2009). Les gènes mutés sont *SPO11-1* dont la mutation abolit l'initiation de la recombinaison, *REC8* dont la mutation conduit à la séparation des chromatides sœurs à la première division de méiose et le gène *OSD1* (pour *Omission of Second Division*) dont la mutation bloque la progression de la méiose à la fin de la première division de méiose. De façon remarquable, *osd1* supprime le phénotype de stérilité du double mutant *Atspo11-1 Atrec8*. Les plantes cumulant ces trois mutations sont baptisées *MiMe* (pour *Meiosis instead of Meiosis*). Les descendances de *MiMe* sont tétraploïdes, et peuvent s'autoféconder pour former des plantes octoploïdes à la génération suivante qui, elles, demeurent stériles. L'équipe de R. Mercier a ensuite montré que d'autres combinaisons de mutations pouvaient être utilisées pour recréer le phénotype *MiMe*. Un mutant *MiMe2* produisant également des gamètes génétiquement identiques aux cellules sporophytiques est produit quand une mutation dans *TAM* (*Tardy asynchronous Meiosis*) est utilisée à la place de *OSD1* (D'Erfurth et al. 2010). *TAM1* est une cycline essentielle à la division méiotique (Cromer et al. 2012). De même *Atspo11-1* peut être remplacé par la mutation d'autres gènes essentiels à la formation de CDBs et qui initient la recombinaison tels les gènes *PRD1*, *PRD2* et *PRD3* (*Putative Recombination Initiation Defect*) (Mieulet et al. 2016). Mais pour le moment, aucune autre mutation de gène ne peut remplacer celle de *REC8*. La seule possibilité serait la mutation de *SCC3*, mais ce gène code pour une protéine ayant également un rôle dans la mitose ce qui affecterait le développement de la plante.

L'objectif de la première partie des travaux engagés dans le cadre de cette thèse était de déterminer si le phénotype *MiMe* obtenu chez *A. thaliana* pouvait être reproduit chez une plante cultivée, le riz, afin de réaliser une première étape vers une apomixie synthétique : l'apoméiose.

4.2. Article : "Turning rice meiosis into mitosis"

Turning rice meiosis into mitosis

Delphine Mieulet¹, Sylvie Jolivet², Maud Rivard², Laurence Cromer², Aurore Vernet¹, Pauline Mayonove¹, Lucie Pereira², Gaëtan Droc¹, Brigitte Courtois¹, Emmanuel Guiderdoni¹, Raphael Mercier²

¹CIRAD, UMR AGAP, 34398 Montpellier Cedex 5, France; ²Institut Jean-Pierre Bourgin, INRA, AgroParisTech, CNRS, Université Paris-Saclay, RD10, 78000 Versailles, France

Introduction of clonal reproduction through seeds (apomixis) in crops has the potential to revolutionize agriculture by allowing self-propagation of any elite variety, in particular F1 hybrids. In the sexual model plant *Arabidopsis thaliana* synthetic clonal reproduction through seeds can be artificially implemented by (i) combining three mutations to turn meiosis into mitosis (*MiMe*) and (ii) crossing the obtained clonal gametes with a line expressing modified CENH3 and whose genome is eliminated in the zygote. Here we show that additional combinations of mutations can turn *Arabidopsis* meiosis into mitosis and that a combination of three mutations in rice (*Oryza sativa*) efficiently turns meiosis into mitosis, leading to the production of male and female clonal diploid gametes in this major crop. Successful implementation of the *MiMe* technology in the phylogenetically distant eudicot *Arabidopsis* and monocot rice opens doors for its application to any flowering plant and paves the way for introducing apomixis in crop species.

Keywords: meiosis; mitosis; apomixis; rice; *MiMe*

Cell Research advance online publication 21 October 2016; doi:10.1038/cr.2016.117

Introduction

Asian rice (*Oryza sativa* L.), the first cereal of human consumption, is the staple food for more than half of the mankind. Current rice production (738 million tons of paddy rice over 160 million hectares) has to rise by more than 20% in the next 20 years [1] on the same arable land surface, to meet the demand of an increasing world population mainly residing in rice-eating countries. Enhancing yield potential, limiting inputs and enabling adaptation to unstable biotic and abiotic environments are current breeding targets. Due to the autogamous mode of reproduction, rice varieties are generally pure lines. However, occurrence of heterosis in rice hybrids that translates into a 15%-20% increase in yield potential [2, 3] prompted the development of F1 hybrids. Hybrid rice has been successfully deployed in India, Indonesia, Vietnam, the Philippines, and in China it is grown in more than 50% of the rice-grown areas [4]. However, the seed cost and the necessity to buy new seeds every year prevent

its large adoption by subsistence farmers [5]. Creating hybrids that can breed true through a clonal mode of reproduction would result in major agricultural benefits for many crops, including rice, allowing the full harnessing of hybrid vigor [6].

Apomixis is an asexual reproductive process through seeds that bypasses meiosis and fertilization, to produce offspring genetically identical to the mother [6]. Apomixis occurs in over 400 angiosperm species and is notably common in the *Poaceae* family, but is absent from the most important crops [7]. Genetic analyses of the inheritance of apomixis in several species including the cereal and grass relatives of the genera *Tripsacum*, *Pennisetum*, *Panicum*, *Brachiaria* and *Paspalum* have identified a single-chromosome segment that can induce apomixis (reviewed in [8]). However, attempts to transfer this chromosomal segment to their respective cultivated relatives have been unsuccessful [9, 10]. The existence of apomictic relatives in cereals and grasses has prompted the screening of wild rice relatives but no natural source of apomixis was discovered [11, 12]. One possible alternative to create apomixis in crop plants, and notably in rice, is to *de novo* engineer clonal reproduction by targeted modification of the sexual reproduction mode, which has been shown to be possible in *Arabidopsis* [13] (Supplementary information, Figure S1).

Correspondence: Emmanuel Guiderdoni^a, Raphael Mercier^b

^aE-mail: guiderdoni@cirad.fr

^bE-mail: raphael.mercier@versailles.inra.fr

Received 3 June 2016; revised 9 August 2016; accepted 30 August 2016

The first component of apomixis is apomeiosis, which is the conversion of the meiotic division into a mitotic-like division, leading to the formation of functional, diploid clonal gametes. Several meiotic mutants have been identified in *Arabidopsis*, maize or rice that produce apomeiosis-derived gametes [14]. However, most of these mutations lead to almost complete male and female sterility, and the production of unreduced gametes occurs at a low frequency (e.g., [15]). One exception is the *Arabidopsis* *MiMe* (Mitosis instead of Meiosis) genotype, in which meiosis is turned into a mitotic-like division, associated with both high fertility and production of clonal diploid gametes at a very high frequency (virtually 100%) [16, 17]. *MiMe* is the combination of mutations in three genes (*SPO11-1*, *REC8* and *OSD1*), each mutation impairing one of the three main processes that distinguish meiosis from mitosis. First, *spo11-1* abolishes meiotic recombination [18]. Second, the mutation of *REC8* causes the separation of sister chromatids at first meiotic division, instead of the distribution of homologous chromosomes [19]. Last, *osd1* causes the skipping of the second meiotic division [16]. Therefore, meiosis in *MiMe* occurs without recombination and distributes sister chromatids in a single round of division, mimicking a mitotic division [16]. This results in the production of clonal male and female gametes, but leads to doubling of ploidy at each generation when self-fertilized. Crossing a *MiMe* plant as male or female with a line whose genome is eliminated following fertilization (lines expressing modified CENH3) leads to the production of clonal offspring [13, 20].

The original *MiMe* genotype is a triple mutant *Atspo11-1 Atrec8 osd1* [16]. To facilitate the transfer of the *MiMe* technology to crops, it would be useful to have more potential target genes, notably when the mutant resource is limited or when the genes are represented by several paralogs. For *AtREC8*, there is currently no alternative as the single other gene known to be required for monopolar orientation of kinetochores is *SCC3*, which is also expressed mitotically and thus the corresponding mutation would affect regular development [19]. Two alternatives to *osd1* can be used to create *MiMe*, *tam* and *TDMAD* [17, 21]. Finally, besides *AtSPO11-1*, several genes have been shown to be essential for initiation of meiotic recombination through DNA double-strand break formation (*AtSPO11-2*, *PRD1*, *PRD2*, *PRD3/PAIR1*, *DFO* and *MTOPVIB*) [22–26], but are not yet tested to create the *MiMe* phenotype.

Given the potential interest in developing apomictic plants and establishing the *MiMe* genotype in crops, the objective of this study was three-fold: (i) extend the catalog of genes that can be mutated to create the *MiMe*

phenotype; (ii) identify the ortholog of the key *OSD1* gene in rice; (iii) determine whether one could establish the *MiMe* phenotype in rice.

Results

Several mutant combinations give rise to the MiMe phenotype in Arabidopsis

To test the ability of other mutations than *spo11-1* to generate *MiMe*, we constructed triple mutants *prd1 Atrec8 osd1*, *prd2 Atrec8 osd1* and *prd3 Atrec8 osd1* and analyzed their meiotic behavior (Figure 1). In wild type, five bivalents align at metaphase I (Figure 1A), and segregate into two groups of five chromosomes at anaphase I (Figure 1B). After a short interkinesis (Figure 1C), the five pairs of chromatids align on the two metaphase II plates (Figure 1D). The sister chromatids segregate at anaphase II (Figure 1E), giving rise to four spores each containing five chromatids (Figure 1F). In each of the triple mutants, meiosis was modified toward a mitotic-like division: ten univalents aligned on the metaphase I plate (Figure 1G, 1J and 1M) and 10 pairs of sister chromatids segregated evenly at anaphase I (Figure 1H, 1K and 1N). No second division occurred, resulting in the production of two spores each containing 10 chromatids (Figure 1I, 1L and 1O). Thus in this series of triple mutants, meiosis recapitulated the *MiMe* phenotype. It is likely that mutations in any of the genes required for double-strand break formation, in combination with *osd1* and *Atrec8*, can cause the *MiMe* phenotype, enlarging the list of genes than can be targeted.

Identification of OsOSD1

There is one *OSD1* homologue in the *Arabidopsis* genome, *UVI4*. Mutation of *UVI4* affects the somatic cell cycle but not meiosis [27]. These two genes originated from a whole-genome duplication that is shared by all *Brassicaceae* [28]. Accordingly, two genes representative of this family are typically found in *Brassicaceae* species and sequence similarity analysis clearly distinguishes the *OSD1* proteins from the *UVI4* proteins (Figure 2). In other eudicots, many species have only one representative of the gene family, and when two representatives are present they result from a recent whole-genome duplication [28] and are not shared with the closest diploid species analyzed (see, for example, GM *Glycine max* or ME *Manihot esculenta* in Figure 2). In *Poaceae*, an independent whole-genome duplication [28] led to two subgroups of genes, with each species having at least one representative of both (e.g., OS *Oryza sativa*, HV *Hordeum vulgare*, ZM *Zea mays* in Figure 2). This raised the possibility that these two *Poaceae* gene families could

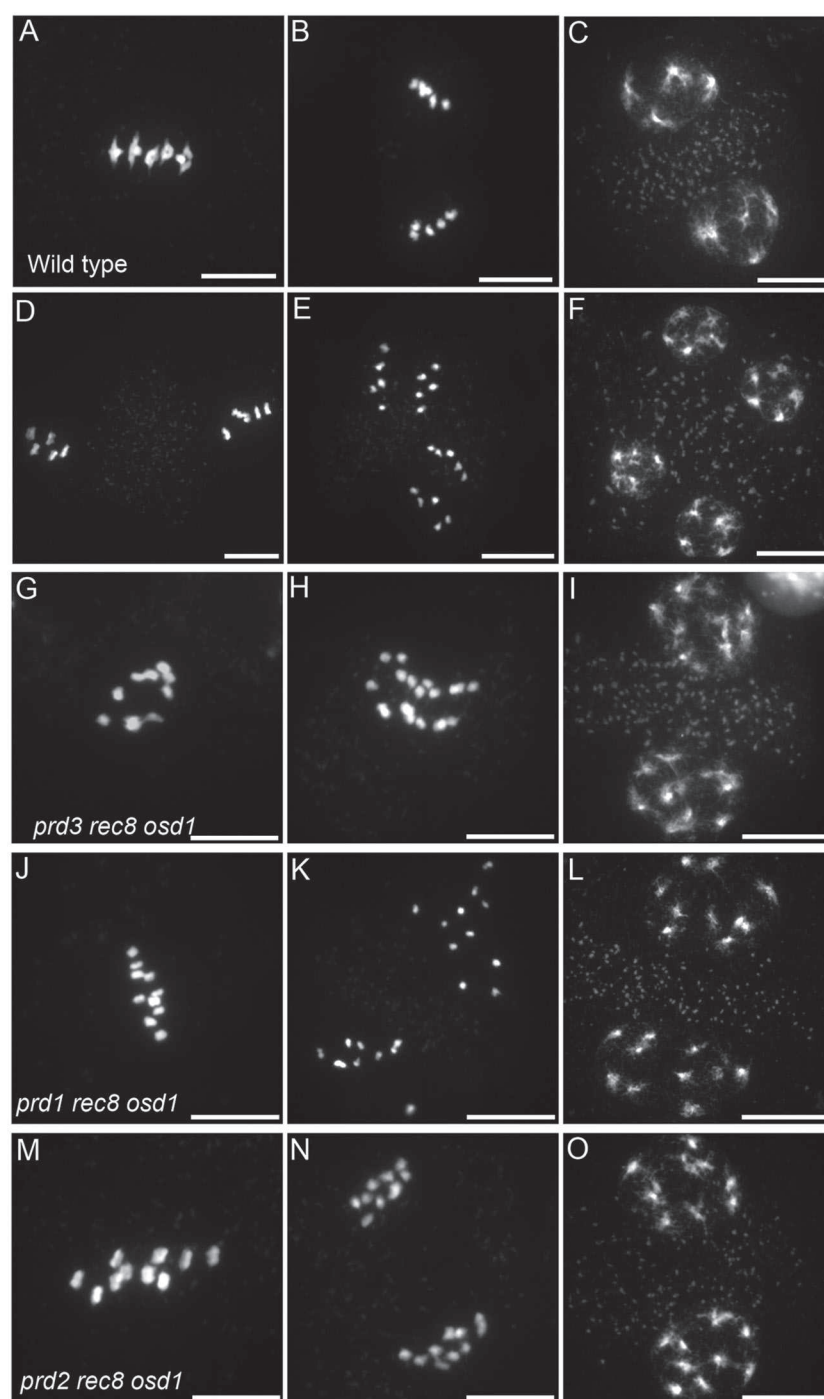


Figure 1 Chromosome spreads of male meiosis in wild-type *Arabidopsis* and *Arabidopsis MiMe* genotypes. **(A-F)** Wild type. **(A)** Metaphase I with five aligned bivalents. **(B)** Anaphase I. **(C)** Telophase I. **(D)** Metaphase II. **(E)** Anaphase II. **(F)** Telophase II. **(G-I)** *prd3 rec8 osd1* triple mutant ($n = 17$). **(J-L)** *prd1 rec8 osd1* triple mutant ($n = 21$). **(M-O)** *prd2 rec8 osd1* triple mutant ($n = 27$). **(G, J, M)** Metaphase I with 10 aligned univalents. **(H, K, N)** Anaphase I with segregation of 10 pairs of chromatids. **(I, L, O)** Telophase. Scale bar = 10 μm .

represent *OSD1* and *UVI4* functional homologues, even though they originated from distinct duplications from

their *Brassicaceae* counterparts. With the aim of exploring the function of these genes in *Poaceae*, we searched

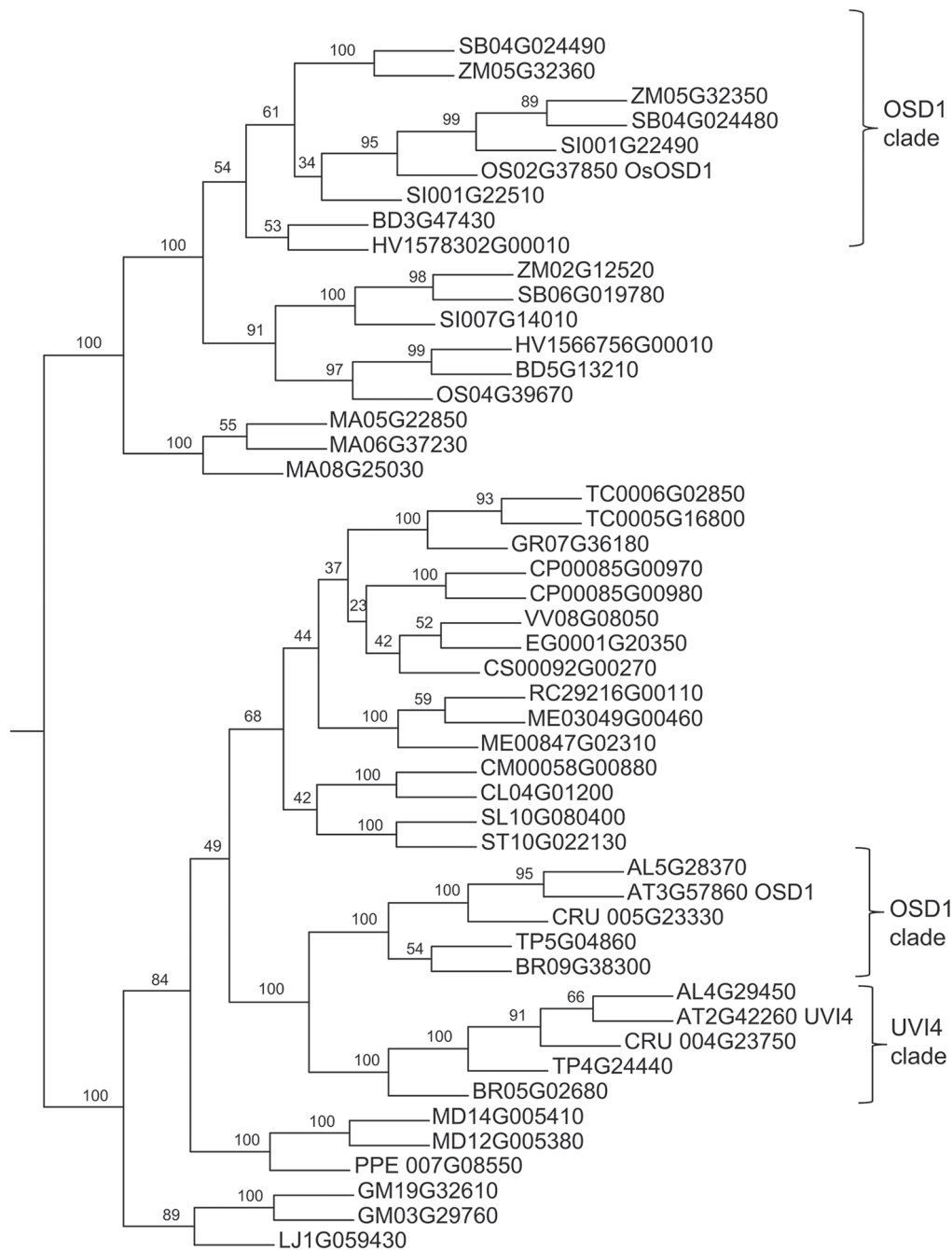


Figure 2 Phylogram of *OSD1* and *UVI4* homologues from flowering plants. Consensus tree of *OSD1* and *UVI4* protein family inferred from a Muscle alignment. Analyses were performed using the MPI Bioinformatic Toolkit [50], Muscle and Phylip-neighbor with default parameters, enabling bootstrapping with 100 replicates. The numbers of forks indicate the number of times the group consisting of the species that are to the right of that fork occurred among the trees, out of 100 trees. AT, *Arabidopsis thaliana*; AL, *Arabidopsis lyrata*; BD, *Brachypodium distachyon*; BR, *Brassica rapa* (turnip mustard); CL, *Citrullus lanatus* (water melon); CM, *Cucumis melo* (muskmelon); CP, *Carica papaya* (papaya); CRU, *Capsella rubella*; CS, *Citrus sinensis* (orange); EG, *Eucalyptus grandis* (eucalyptus); GM, *Glycine max* (soybean); GR, *Gossypium raimondii* (cotton); HV, *Hordeum vulgare* (barley); LJ, *Lotus japonicus*; MA, *Musa acuminata* (banana); MD, *Malus domestica* (apple); ME, *Manihot esculenta* (cassava); OS, *Oryza sativa* (rice); PPE, *Prunus persica* (peach); RC, *Ricinus communis* (castor bean plant); SB, *Sorghum bicolor* (sorghum); SI, *Setaria italica* (foxtail millet); SL, *Solanum lycopersicum* (tomato); ST, *Solanum tuberosum* (potato); TC, *Theobroma cacao* (cacao); VV, *Vitis vinifera* (grape vine); ZM, *Zea mays* (maize).

for mutants in the international rice collections. No insertion was identified in *Os04g39670*, hindering the study of this gene in the present work. However, three insertions were identified in *Os02g37850*. These three insertions, one T-DNA (AMBA12) and two *Tos17* inserts (AMQF10 and ALJA10), are in the second intron (Supplementary information, Figure S2).

Because of omission of the second meiotic division, *Arabidopsis osd1* produces diploid male and female gametes and its self-fertilization generates tetraploid progeny. To address the function of *Os02g37850*, we used flow cytometry to determine the ploidy of the progeny that are homozygous for the AMBA12, AMQF10 and ALJA10 mutations. This revealed the presence of 100% ($n = 82$), 37% ($n = 27$) and 5% ($n = 20$) of tetraploid plants in the respective progeny, the rest being diploid. This suggests that mutations of *Os02g37850* can provoke the production of male and female diploid gametes. AMBA12 appears to be a null mutant, while AMQF10 and ALJA10 appear partially affected. Consistently, RT-PCR experiments showed that the *OSD1* spliced mRNA is barely detectable in AMBA12 homozygotes, and is only reduced in abundance in AMQF10 (Supplementary information, Figure S3). We then compared meiosis in wild type and AMBA12. First, observation of male meiotic products revealed the presence of 100% dyads ($n = 500$) in AMBA12 homozygotes, instead of tetrads in wild type (Figure 3), showing that a single meiotic division occurs in AMBA12. In AMQF10, a mixture of dyads (29%) and tetrads were observed ($n = 130$), in accordance with the mixed ploidy in offspring of this mutant. Chromosome spreads showed that the first meiotic division in AMBA12 was indistinguishable from wild type (Figure 4), suggesting that crossover formation and homologous chromosome segregation were unaffected, but no second division took place. Thus, a single

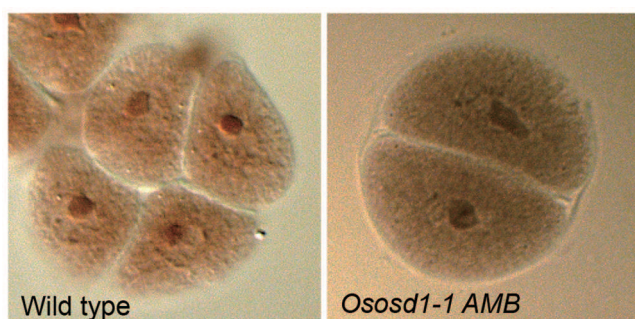


Figure 3 Male meiotic products in wild-type rice and *Ososd1*. Fresh anthers squashed in acetocarmine. (A) Tetrads of spores in wild type. (B) Dyad of spores in *Ososd1-1* ($n = 500$).

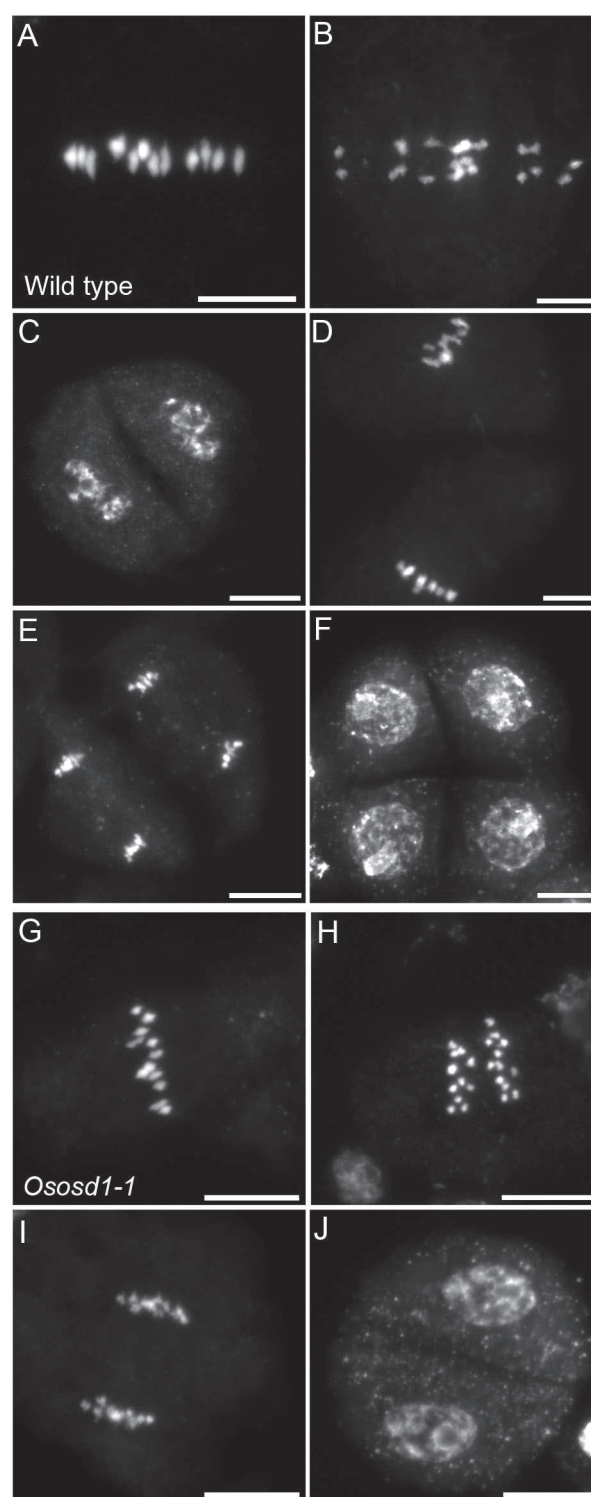


Figure 4 Chromosome spreads of male meiosis in wild-type rice and *Ososd1*. (A-F) Wild type. (A) Metaphase I with 12 aligned bivalents. (B) Anaphase I. (C) Telophase I. (D) Metaphase II. (E) Anaphase II. (F) Telophase II. (G-H) *Ososd1-1* ($n = 62$). (G) Metaphase I with 12 aligned bivalents. (H) Anaphase I. (I) Late anaphase I. (J) Telophase I. No second division was observed. Scale bar = 10 μm .

first meiotic division occurred in AMBA12. In summary, disruption of *Os02g37850* recapitulated the *Arabidopsis osd1* meiotic defect: the omission of the second meiotic division which leads to the production of functional diploid gametes. We therefore named *Os02g37850 OsOSD1* and AMBA12 *Ososd1-1*.

Frequencies of diploid gametes in *Ososd1-1* mutant

To directly determine the ploidy of male gametes in the *Ososd1-1* mutant, we examined the DNA content of nuclei isolated from mature pollen grains with flow cytometry (Figure 5). Pollen nuclei isolated from *OsOSD1^{+/+}* ($n = 5$) and *Ososd1^{-/-}* plants ($n = 6$) exhibited

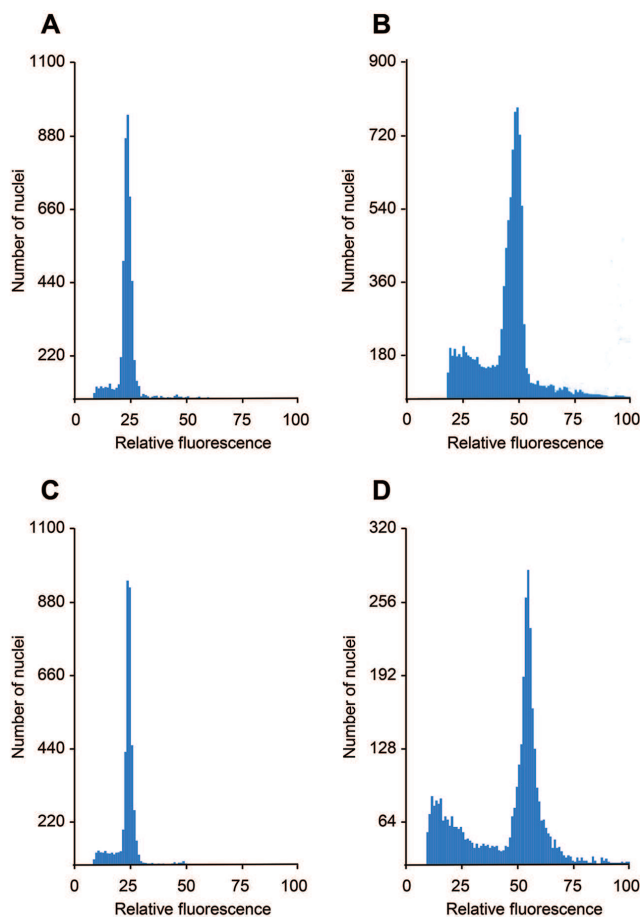


Figure 5 Ploidy of *Ososd1-1* pollen grains. The ploidy was determined using staining of isolated nuclei with propidium iodide followed by flow cytometry. **(A)** Wild-type Nipponbare pollen grains. A single peak is observed, corresponding to haploid nuclei. **(B)** Wild-type Nipponbare leaf. A single peak, corresponding to diploid nuclei. **(C)** *Ososd1^{+/+}* pollen grains. A single peak is observed, corresponding to haploid nuclei. **(D)** *Ososd1^{-/-}* pollen grains. A single peak is observed, corresponding to diploid nuclei.

the expected haploid DNA content (Figure 5A and 5C), which was half of the DNA content in leaf nuclei (Figure 5B). In contrast, pollen nuclei isolated from *Ososd1^{-/-}* plants ($n = 3$) formed a unique peak corresponding to a diploid DNA content. This indicates that all pollen grains produced by *Ososd1* mutants are diploid. We then examined the female meiosis product of *Ososd1^{-/-}* plants by fertilizing them with wild-type pollen. Among 112 seeds obtained, 104 were shrunk and unable to germinate, whereas 8 developed normally and germinated. The obtained eight offspring were diploid and heterozygous for the *Ososd1-1* mutation, indicating they were the result of the fertilization of haploid female gametes produced by *Ososd1-1*. To recover more offspring, we repeated the experiment, by rescuing the embryo at 5 days post fertilization and transferring the resulting plants in soil. Among 47 successfully hybridized flowers, 40 exhibited early signs of abnormal development, including absence of albumen, and developed into triploid plants after embryo rescue. The seven remaining plants, derived from normally developed seeds, were diploid and *Ososd1^{+/+}*. Altogether, these results show that ~91% of female gametes produced by *Ososd1-1* were diploid and ~9% (15/159) were haploid. This is reminiscent of the effect of the *osd1* mutation in *Arabidopsis*, in which ~100% of male gametes and ~85% of female gametes are diploid [16]. However, in contrast to the situation in *Arabidopsis*, seeds resulting from the fertilization of a haploid female gamete by a diploid male gamete in rice are not viable and require embryo rescue, due to abnormal endosperm development [29]. The identification of the *OSD1* ortholog in rice and the high frequency of diploid gametes produced by the corresponding mutant open the way to creating *MiMe* in this species.

Generating *MiMe* in rice

To generate the *MiMe* genotype, *Ososd1* mutation must be combined with a mutation that abolishes recombination, and *Osrec8* that modifies chromatid segregation. The *PAIR1* gene has been previously shown to be required for meiotic recombination initiation in rice [30]. One allele (Supplementary information, Figure S2) was selected in a Nipponbare background (AQUG12/*pair1-4*). As previously described for *pair1* alleles, plants homozygous for this insertion were fully sterile. In *pair1-4* meiocytes, 24 univalents were observed at metaphase I instead of 12 bivalents in wild type, followed by unbalanced distribution of univalents at anaphase I, suggesting that recombination was completely abolished by this allele (Figure 6A and 6B). *OsREC8* has also been functionally characterized [31]. *Osrec8* mutation leads to sister chromatid cohesion defect and chromosome frag-

mentation, similar to *Atrec8* [19, 32, 33]. We identified a novel T-DNA allele in the Hwayoung genetic background (Postech mutant line B01997/*Osrec8-3*) which caused full sterility (Supplementary information, Figure S2). Meiotic chromosome spreads from *Osrec8-3* meiocytes displayed entangled mass of DNA at metaphase I, followed by chromosome fragmentation and missegregation (Figure 6C and 6D), similar to effects of other *Osrec8* alleles. In summary, three single mutants *pair1-4*, *Osrec8-3* and *Ososd1-1* exhibited the same phenotypes as the three mutants used to build the *MiMe* genotype in *Arabidopsis*.

Combination of *pair1*, *Osrec8* and *Ososd1* mutations in rice was performed through sequential crossing of heterozygous plants (Figure 7A). In each generation, the segregation of mutant alleles followed the expected Mendelian segregation pattern. The *pair1 Osrec8* double mutant was sterile. In this double mutant, 24 univalents were observed, which were aligned on the metaphase plate and were segregated into two groups of 24 chromatids (Figure 6E and 6F). This shows that the *pair1* mutation abolishes chromosome fragmentation of *Osrec8*, indicating that *OsREC8* is required for efficient DNA double-strand break repair. It also shows mutating *OsREC8* in *pair1* modifies the mode of segregation of chromosomes: in *pair1* the two sister chromatids segregate as a single unit (monopolar orientation), leading to unbalanced segregation of univalents at anaphase I, while in *pair1 Osrec8* the pairs of chromatids segregate to opposite poles (bipolar orientation), leading to a balanced mitotic-like distribution of sister chromatids at anaphase I. Hence, like in *Arabidopsis* [19] and fission yeast [34], *OsREC8* is required for monopolar orientation of sister kinetochores at meiosis I in rice. In *pair1 Osrec8*, the free chromatids resulting from the mitotic-like segregation at meiosis I did not segregate properly at second division, producing unbalanced spores. The single mutants *pair1*, *Osrec8* and the double mutants *pair1 Osrec8* and *pair1 Ososd1* plants were fully sterile. In contrast, the *pair1 Osrec8 Ososd1* triple mutant showed higher fertility, similar to *Ososd1* single mutant (Figure 7B). Examination of the male meiotic chromosome behavior in *pair1 Osrec8 Ososd1* (Figure 8) revealed the presence of 24 univalents that aligned at metaphase I, and segregated into two groups of 24 chromatids to produce dyads of spores. This behavior resembles a mitotic division, suggesting that, like in *Arabidopsis*, the combination of *pair1*, *rec8* and *osd1* mutations turns meiosis into mitosis. We therefore term this genotype *OsMiMe*.

To gather further support that *OsMiMe* turns meiosis into mitosis, we asked if these plants produce diploid clonal gametes. Ploidy of seedling progeny from

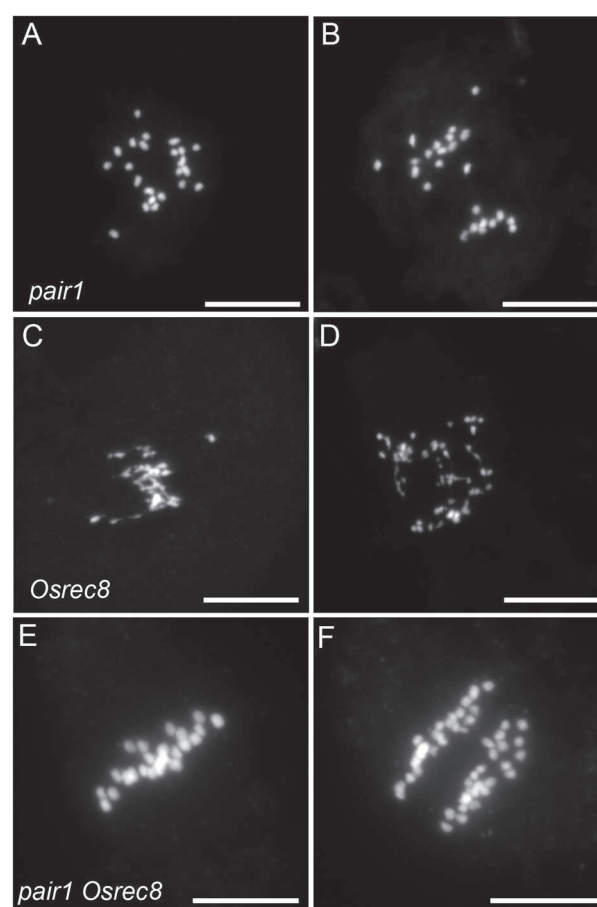


Figure 6 Male meiosis I in rice *pair1*, *Osrec8* and *pair1 Osrec8* mutants. **(A–B)** *pair1* ($n = 154$). **(A)** Metaphase I with 24 unaligned univalents. **(B)** Anaphase I with unbalanced segregation of univalents. **(C, D)** *Osrec8* ($n = 222$). **(C)** Metaphase I with abnormal chromosomes. **(D)** Anaphase I with chromosome fragmentation. **(E, F)** *pair1 Osrec8* ($n = 154$). **(E)** Metaphase I with 24 aligned univalents. **(F)** Anaphase I with segregation of 24 pairs of chromatids. Scale bar = 10 μm .

self-pollination was estimated using flow cytometry, for both *Ososd1* and *OsMiMe* plants. All the progeny (from two and three mother plants, $n = 50$ each) were tetraploid, indicating that both *Ososd1* and *OsMiMe* diploid plants produce diploid male and female gametes. As the chromosome spreads suggested that gametes are derived from a first meiotic division in *Ososd1*, and from a mitotic-like division in *OsMiMe*, their genetic make-up should be different (Supplementary information, Figure S1). *Ososd1* should produce recombined gametes, while *pair1 Osrec8 Ososd1* should produce clonal gametes. More specifically, segregation of a heterozygous marker (Aa) following meiosis I depends on its position along the chromosome [35]. A marker at the centromere would

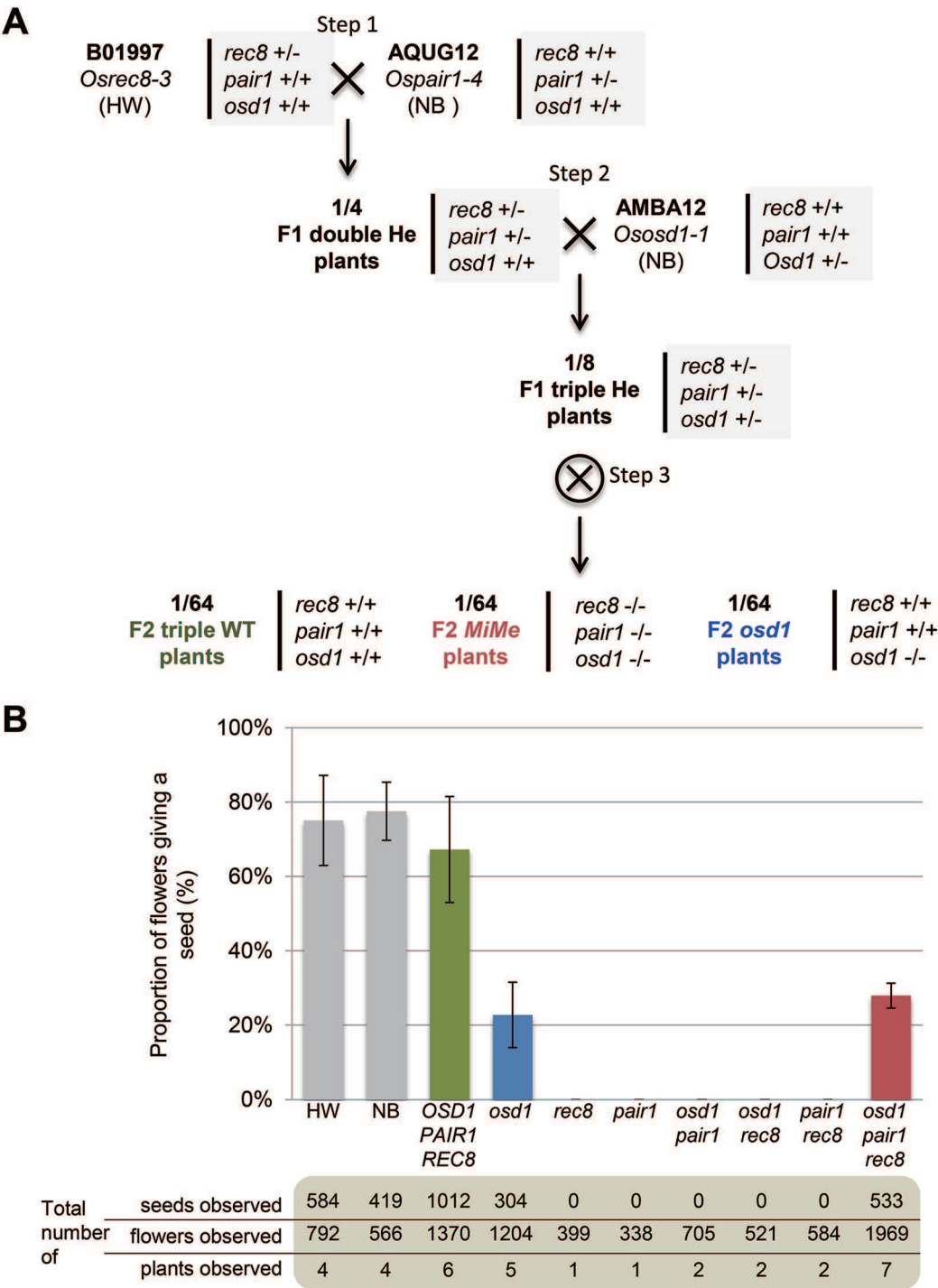


Figure 7 Creating *OsMiMe*. **(A)** Crossing scheme. *OsMiMe* mutant was created first by crossing B01997 plants heterozygous for the *Osrec8* mutation with AQUG12 plants heterozygous for the *pair1* mutation (step 1). 1/4 of F1 were double heterozygous plants and were in turn crossed with AMBA12 plants heterozygous for the *Ososd1* mutation to produce 1/8 of F1 triple heterozygous plants (*Osrec8*^{+/−}; *pair1*^{+/−}; *Ososd1*^{+/−}) (step 2). These plants were selfed (step 3) and the progeny were genotyped to identify 1/64 triple homozygous *OsMiMe* plants and relevant control segregants. HW, Hwayoung; NB, Nipponbare; He, heterozygous. **(B)** Fertility in *MiMe* and related mutants. The histogram represents the percentage of flowers giving a seed by self-fertilization. Numbers of observed flowers and seeds are indicated. Errors bars correspond to the SD for 4–7 plants.

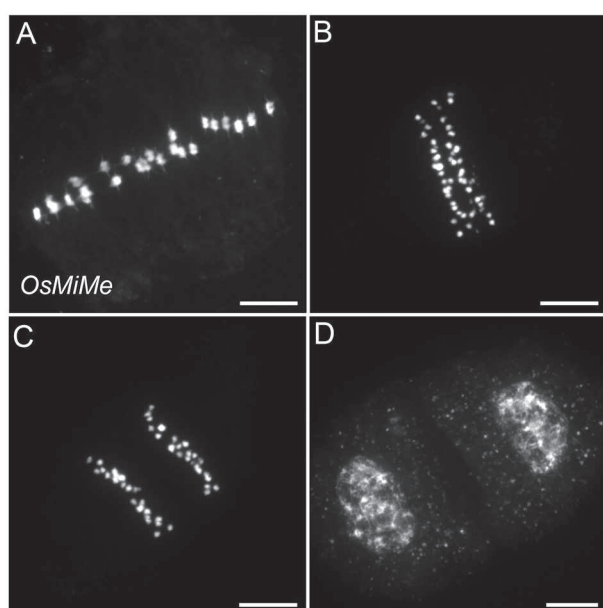


Figure 8 Male meiosis I in *OsMiMe*. **(A)** Metaphase I with 24 aligned univalents. **(B, C)** Anaphase I with segregation of 24 pairs of chromatids. **(D)** Telophase I. $n = 43$. Scale bar = 10 μm .

be homozygous in each diploid gamete (AA or aa), because of co-segregation of sister chromatids. This would render the tetraploid progeny after selfing having AAAA ($\frac{1}{4}$ of plants), aaaa ($\frac{1}{4}$ of plants), and AAaa ($\frac{1}{2}$ of plants). Markers distant from the centromere would segregate with different frequencies because of recombination [35]. The situation following a mitotic-like division is simpler: all heterozygosity would be maintained in each diploid gamete (Aa) because of separation of sister chromatids, and absence of recombination.

To analyze the segregation pattern in *Ososd1* and *pair1 Osrec8 Ososd1*, we took advantage of the fact that the genetic background of the segregating population is derived from a cross between two polymorphic strains Hwayoung (B01997/*Osrec8-3*) and Nipponbare (AUG12/*pair1-4*), which was backcrossed to Nipponbare (AMBA12/*Ososd1-1*) (Figure 7A). We selected markers that were heterozygous in the *Ososd1* and *pair1 Osrec8 Ososd1* diploid plants. The genotypes of tetraploid progeny siblings of *Ososd1* and *pair1 Osrec8 Ososd1* diploid parents at diagnostic markers are shown in Figure 9. In *Ososd1* progeny ($n = 77$), the expected segregation was observed with the loss of heterozygosity at centromeric markers in $\sim 50\%$ of the plants. In sharp contrast, all the diagnostic markers heterozygous in the *OsMiMe* diploid parents were also found heterozygous in all the tetraploid progeny siblings ($n = 73$). This indicates

that all *OsMiMe* gametes fully retained the parental plant genotype.

Discussion

Engineering apomixis in sexual crops would allow the self-propagation of F1 hybrids, which would have immense interest in agriculture by giving access to hybrid vigor in more species and to resource-poor farmers. Apomixis can be separated into two developmental components, the absence of meiosis or its modification into a mitosis-like division and the development of the embryo and its nourishing tissue from the female gametophyte without contribution of a male genome. The most efficient strategy to date to modify meiosis into mitosis has been obtained by combining mutations in three genes *SPO11-1*, *REC8* and *OSD1* (called the *MiMe* genotype) in the model species *Arabidopsis thaliana*. These three mutations affect, respectively, the three pillars of meiosis, homologous recombination, monopolar orientation of sister chromatids and the occurrence of a second division. Here we show that the mutation of *SPO11-1* can be efficiently replaced by mutation in other recombination initiation factors *PRD1*, *PRD2* or *PRD3/PAIR1* to create the *MiMe* phenotype. This suggests that mutation in any gene essential for double-strand break formation, such as *SPO11-2*, *DFO* or *MTOPVIB*, may be used instead of *SPO11-1* [22–26]. This observation may allow us to build *MiMe* in other species, in which mutant collection is limited or in which some genes could be present in more than one copy.

Although *MiMe* technology has been proven to be very efficient in the model species *Arabidopsis*, a major question has been about its transferability to crop species, especially the distant cereal species. *SPO11-1* and *REC8* are widely conserved among eukaryotes, both in sequence and in function, making the identification of their homologs in other plant species easily feasible. Other genes known to be essential for recombination initiation are also well conserved among plants and all these meiosis genes are typically represented by a single copy [28]. Indeed, mutating *REC8* or *PAIR1/PRD3* in rice recapitulates the phenotypes observed in *Arabidopsis*. In contrast, *OSD1* appears to be a plant-specific gene and exhibits a more complex phylogeny (Figure 2). There is one *OSD1* paralog in the *Arabidopsis* genome, *UVI4*, which has a distinct function in regulating somatic cell cycle [27]. *osd1 uvi4* double mutant confers very strong gametophyte and developmental defect, suggesting that these two genes have redundant functions in mitosis [36, 37]. Sequence analysis clearly distinguishes *OSD1* from the *UVI4* protein families in *Brassicaceae* species, but

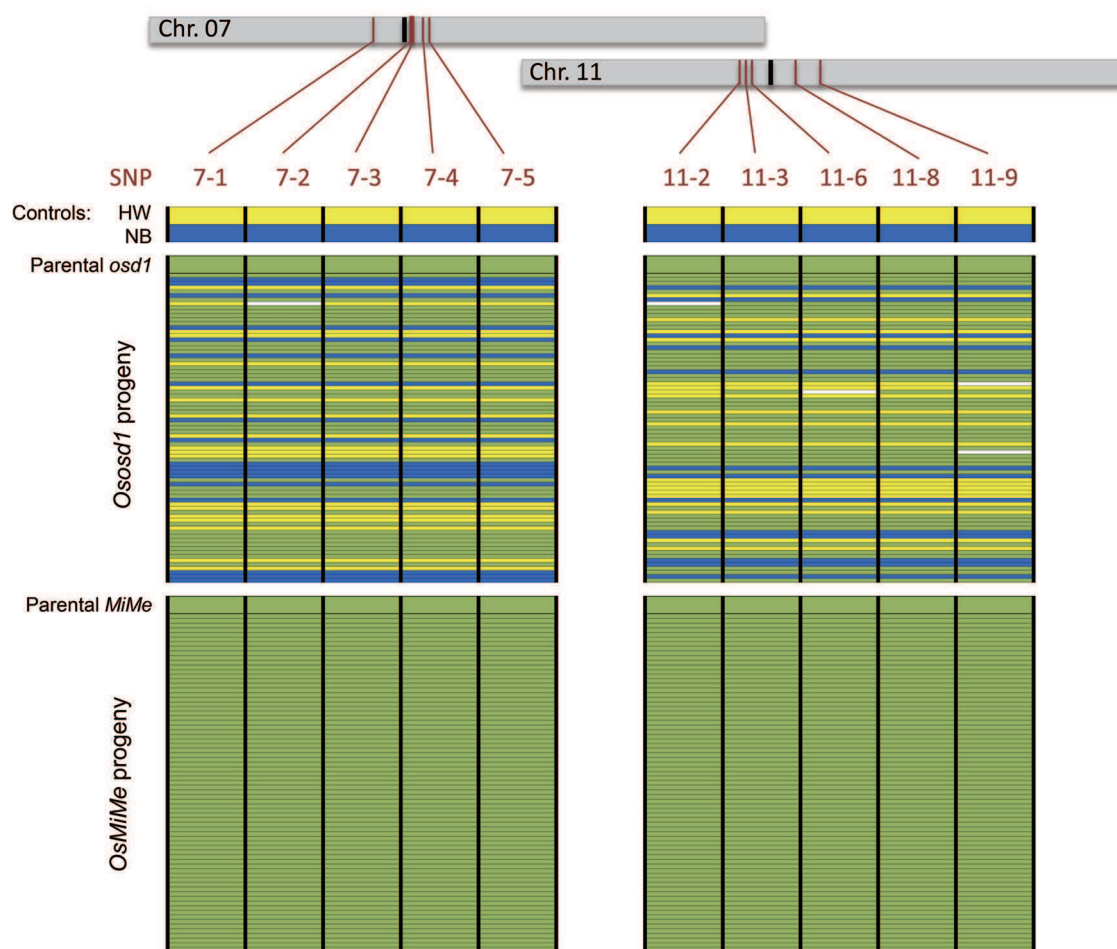


Figure 9 Genotypes of *OsOsd1* and *OsMiMe* diploid parents and their respective tetraploid progeny siblings. Tetraploid offspring of *OsOsd1* and *OsMiMe* plants (refer to Figure 7 for scheme of creation of plant materials) were genotyped using 10 SNP markers residing on rice chromosomes 7 and 11. Positions of markers (brown) and centromeres (black) are indicated along the chromosomes. Each line represents a plant. For each marker, plants carrying only the Hwayoung allele are in yellow, plants carrying only the Nipponbare allele are in blue, while plants with both the Hwayoung and Nipponbare alleles appear in green. *Osd1* and *OsMiMe* parental plants (top line of upper and lower panels, respectively) are heterozygous at the 10 markers. While the offspring of *OsOsd1* exhibit segregation of parental alleles at the 10 markers, all the offspring of *OsMiMe* are heterozygous at the same 10 markers and exhibit the same genotype like the parental plants.

not in more distant flowering plants (Figure 2). In cereals, there are clearly two subgroups of genes, with one or two representatives of both groups in each species. The *Oryza sativa* genome contains two genes, one in each group, and we showed that the single mutation of Os02g37850 is sufficient to give rise to the meiotic defects observed in the *Arabidopsis osd1* mutant. We consequently called this gene *OsOSD1*. Meiosis in *Ososd1* skips the second division, producing diploid spores and gametes. These gametes are functional, leading to increased ploidy in the offspring by selfing. The existence and identification of the functional homolog of *OSD1* in rice suggests that it could be identified in other cereals. Synteny and protein

sequence similarity indicate ortholog of *OSD1* as a single gene in *Hordeum vulgare*, *Brachypodium distachyon* and as a tandem duplication in *Zea mays*, *Sorghum bicolor* and *Setaria italica* [28]. This suggests that diploid gamete production through *osd1* mutations could be obtained in cereals, but may require the use of RNAi or genome editing technologies in the case of tandem duplications.

We identified three meiotic mutations that respectively abolish recombination (*pair1*), allow the separation of sister chromatids (*Osrec8*) and provoke the omission of second division (*Ososd1*). Next, we combined them to construct the *MiMe* genotype. In this triple mutant, meiosis is converted into a mitotic-like division with balanced

segregation of sister chromatids in a single division event. *OsMiMe* plants produce diploid gametes that are genetically identical to their parents. Our study thus establishes that the *MiMe* technology is transferable from the model plant *Arabidopsis* to the distant crop *Oryza sativa*, and suggests that it may be adopted in other important crops. The creation of the rice *OsMiMe* triple mutant was time-consuming and the approach remains so far restricted to genotypes/species with large associated mutant resources. However, recent breakthrough in genome editing by CRISPR/Cas9 may remove many obstacles [38], making apomeiotic conversion a feasible reality in many crops.

Rice *OsMiMe* plants produce functional clonal embryos. However, fertilization still occurs normally, leading to the doubling of ploidy in the next generation. The next challenge is to trigger the development of a *MiMe* gamete into an embryo and then a plant without the contribution of another gamete. One approach is to use genome elimination, where the chromosome set from one parent is removed after fertilization in the zygote. *MiMe* combined with genome elimination indeed leads to clonal offspring, as previously demonstrated in *Arabidopsis* [13]. However, the CENH3 manipulation used in *Arabidopsis* to provoke genome elimination has not been transferred in other species to date [20, 39–41], except in maize but with a lower frequency [42]. The frequency of genome elimination can be genetically controlled in cereals, notably in maize [43–45]. The *gynogenesis inducer 1* (*ggi1*) locus has been shown to induce genome elimination and is widely used in maize breeding to produce haploids and the identification of the underlying gene(s) will open exciting new avenues to clonal reproduction [46]. Alternatively, parthenogenic development of embryo could be induced, and recent progress has been made in the understanding of its genetic control [47, 48].

Materials and Methods

Phylogeny

Protein sequences from the OSD1/UVI4 family were collected using PLAZZA DICOTS3.0 and PLAZZA Monocots 3.0 (<http://bioinformatics.psb.ugent.be/plaza/>) [49] and correspond to family HOM03M004665, HOM03M067707 and HOM03D003883. Some species were removed from the analysis for simplicity. BD5G13210 sequence was replaced by Bradi5g13212.1 (Phytozome 10.3.). Three proteins (SI005G06210, ZM10G18380 and GM11G21435) were excluded because, though they share sequence similarity with the OSD1/UVI4 proteins, they lack most of the conserved domains. Multi-protein alignment (Muscle v3.8) followed by phylogeny inference (PHYLP Neighbor-Joining/UP-GMA method version 3.696) was performed using the MPI toolkit with default parameters [50].

Genetic material

The *Arabidopsis* mutants used were *prd1-2*, *prd2-1*, *prd3-3* [22, 23], *osd1-1* and *atrec8-3* [16]. Rice (*Oryza sativa* L.) mutants used in this study were the following: AMBA12 T-DNA insertion line and two Tos17 insertion mutant AMQF10 and ALJA10 lines for OsOSD1 (Os02g37850); AQUG12 Tos17 insertion line for PAIR1 (Os03g01590); B01997 T-DNA insertion line for OsREC8 (Os05g50410). All insertion lines were generated in the Nipponbare background and are from the Oryza Tag Line insertion line library [51, 52] except B01997, which has a Hwayoung background and is from the POSTECH Rice Insertion Database [53]. Position of the inserts in the respective genes is shown in Supplementary information, Figure S2. The exact junction of the T-DNA insert with the rice genome in the mutants was ascertained by Sanger sequencing.

Genotyping of T-DNA and Tos17 inserts

Genotyping primers are listed in Supplementary information, Table S1. Primers were designed using the “Genotyping Primer Designer” tool of OryGenesDB (<http://www.orygenesdb.cirad.fr>) [54, 55]. In these standard conditions, the expected PCR product is ~800 bp for a mutant allele and ~1 000 bp for a wild-type allele. PCR was performed following the protocol described in [56].

SNP genotyping

Single-nucleotide polymorphism genotyping was performed using Kompetitive Allele Specific PCR (KASP) following the LGC group recommendations for the use of KASP technology on Roche LC480.

Cytological analyses

Arabidopsis chromosome spreads and observation were carried out using the technique described in [18].

Conditions of growth and crossing of the plants

Rice plants were grown under containment greenhouse condition (temperature of 28 °C day and 24 °C night, 60% hygrometry) under natural light that could be assisted by artificial sodium light (light intensity of 700 $\mu\text{mol}/\text{m}^2/\text{s}$). Crossing was accomplished through manual castration of florets and pollination according to the standard procedures, followed by bagging to avoid pollen contamination.

Embryo rescue

Five days after pollination, fertilized flowers were harvested, their lemma and palea removed under a stereo microscope and soaked in 70% ethanol for 5 min. The caryopses were then washed thrice in sterile distilled water. Embryos were isolated following two distinct procedures depending on the presence or absence of endosperm [42, 43]. Immature embryos were plated on MS basal medium [44] supplemented with 50 g/l sucrose and 6 g/l phytagel. Following a 10-day growth period in the dark at 26 °C, the seedlings were transferred to a fresh MS solid medium (2.6 g/l of phytagel; 30 g/l of sucrose) and in a growth chamber (photoperiod 16:8), 28 °C and 60% humidity for 3 weeks.

Flow cytometry determination of DNA content in pollen and leaf cell nuclei

Ploidy was determined by estimating nuclear DNA content

by flow cytometry. Pollen grains from flowers at anthesis were collected by introducing the panicle into a 50 ml tube and gently shaking allowing the shedding of pollen from anthers. One milliliter of LB01 buffer was added to humidify the pollen followed by a mild centrifugation (10 min at 150× g) allowing collection of the pollen pellet. The supernatant was discarded and the pellet was resuspended with 1 ml of a fresh LB01. The nucleus extraction from pollen grains was performed according to Kron *et al.* [57] bursting method. Due to the size of rice pollen grains and nuclei, we used Partec CellTrics at 100 µm for the first filtration and 30 µm for the second one. For leaf nucleus extraction, ~0.5 cm² of leaf blade tissue was chopped in 250 µl of LB01 extraction buffer. The suspension was filtered (30 µm Partec CellTrics filters) to eliminate cell debris. After filtration, 700 µl of LB01 buffer was added. After DNA nucleus extraction from pollen or leaf tissue, DNA was stained by addition of propidium iodide (40 µl of PI at 0.2 mg/ml). The samples were analyzed with UV excitation in a Partec flow cytometer (PA-I; Partec, Munster, Germany).

Acknowledgments

This work has been funded by the ANR Emergence DipLOSD (ANR-09-EBIO-011) to RM and EG. Part of this work has been supported by the RicE FUnctional GENomics (REFUGE) platform funded by Agropolis Fondation. We thank Eve Lorenzini, Christian Chaine and Rémy Michel for providing technical assistance, notably in greenhouse. We thank the guidance of Dr Frédéric Bakry, Cirad, for flow cytometric analyses. We thank Marta Cifuentes, Eric Jenczewski, Wayne Crismani, Rajeev Kumar and Anne Dié-vart for critical reading of the manuscript. The IJPB benefits from the support of the Labex Saclay Plant Sciences-SPS (ANR-10-LABX-0040-SPS).

Author Contributions

EG and RM conceived the project. DM, EG and RM designed experiments. DM, SJ, MR, LC, AV, PM, LP, GD and BD performed experiments and analysed the data. DM, EG and RM wrote the paper with the input from other authors.

Competing Financial Interests

The authors declare no competing financial interests

References

- Rejesus RM, Mohanty S, Balagtas JV. Forecasting global rice consumption. 2012. http://www.agecon.purdue.edu/staff/balagtas/rice_timeseries_v6.pdf. Accessed 01 March 2014
- Virk PS, Khush GS, Virmani SS. Breeding strategies to enhance heterosis in rice. 2003.
- Virmani SS, Mao CX HBE. In: Virmani SS, Mao CX, Hardy B, eds. Hybrid Rice For Food Security, Poverty Alleviation, and Environmental Protection. International Rice Research Institute, Manila, Philippines, 2003.
- Ma GH, Yuan LP. Hybrid rice achievements and development in China. In: Virmani SS, Mao CX, Hardy B, ed. Hybrid Rice For Food Security, Poverty Alleviation, and Environmental Protection. International Rice Research Institute, Manila, Philippines, 2003.
- Virmani SS. Heterosis and Hybrid Rice Breeding. Berlin, Heidelberg: Springer, 1994.
- Spillane C, Curtis MD, Grossniklaus U. Apomixis technology development-virgin births in farmers' fields? *Nat Biotechnol* 2004; **22**:687-691.
- Koltunow AM, Grossniklaus U. Apomixis: a developmental perspective. *Annu Rev Plant Biol* 2003; **54**:547-574.
- Ozias-Akins P, van Dijk PJ. Mendelian genetics of apomixis in plants. *Annu Rev Genet* 2007; **41**:509-537.
- Barcaccia G, Albertini E. Apomixis in plant reproduction: a novel perspective on an old dilemma. *Plant Reprod* 2013; **26**:159-179.
- Savidan Y. Transfer of apomixis through wide crosses. In: Savidan Y, Carman JG, Dresselhaus T, eds. The Flowering of Apomixis: From Mechanisms to Genetic Engineering. Mexico: CIMMYT, IR; 2001:153-167.
- Rutger JN. Searching for apomixis in rice. In: *Proceedings of Apomixis Workshop*. Beltsville, MD, USA: US Department of Agriculture; 1992:36-39.
- Khush GS. Apomixis: Exploiting Hybrid Vigor in Rice. International Rice Research Institute, Manila, Philippines, 1994.
- Marimuthu MP, Jolivet S, Ravi M, *et al.* Synthetic clonal reproduction through seeds. *Science* 2011; **331**:876.
- Ronceret A, Vielle-Calzada JP. Meiosis, unreduced gametes, and parthenogenesis: implications for engineering clonal seed formation in crops. *Plant Reprod* 2015; **28**:91-102.
- Ravi M, Marimuthu MPA, Siddiqi I. Gamete formation without meiosis in *Arabidopsis*. *Nature* 2008; **451**:1121-1124.
- D'Erfurth I, Jolivet S, Froger N, Catrice O, Novatchkova M, Mercier R. Turning meiosis into mitosis. *PLoS Biol* 2009; **7**:e1000124.
- D'Erfurth I, Cromer L, Jolivet S, *et al.* The cyclin-A CYCA1;2/TAM is required for the meiosis I to meiosis II transition and cooperates with OSD1 for the prophase to first meiotic division transition. *PLoS Genet* 2010; **6**:e1000989.
- Grelon M, Vezon D, Gendrot G, Pelletier G. AtSPO11-1 is necessary for efficient meiotic recombination in plants. *EMBO J* 2001; **20**:589-600.
- Chelysheva L, Diallo S, Vezon D, *et al.* AtREC8 and AtSCC3 are essential to the monopolar orientation of the kinetochores during meiosis. *J Cell Sci* 2005; **118**:4621-4632.
- Ravi M, Chan SWL. Haploid plants produced by centromere-mediated genome elimination. *Nature* 2010; **464**:615-618.
- Cifuentes M, Jolivet S, Cromer L, *et al.* TDM1 regulation determines the number of meiotic divisions. *PLoS Genet* 2016; **12**:e1005856.
- De Muyt A, Vezon D, Gendrot G, Gallois JL, Stevens R, Grelon M. AtPRD1 is required for meiotic double strand break formation in *Arabidopsis thaliana*. *EMBO J* 2007; **26**:4126-4137.
- De Muyt A, Pereira L, Vezon D, *et al.* A high throughput genetic screen identifies new early meiotic recombination functions in *Arabidopsis thaliana*. *PLoS Genet* 2009; **5**:e1000654.
- Stacey NJ, Kuromori T, Azumi Y, *et al.* *Arabidopsis* SPO11-2 functions with SPO11-1 in meiotic recombination. *Plant J* 2006; **48**:206-216.
- Zhang C, Song Y, Cheng Z, *et al.* The *Arabidopsis thaliana* DSB formation (*AtDFO*) gene is required for meiotic dou-

- ble-strand break formation. *Plant J* 2012; **72**:271-281.
- 26 Vrielynck N, Chambon A, Vezon D, *et al.* A DNA topoisomerase VI-like complex initiates meiotic recombination. *Science* 2016; **351**:939-943.
 - 27 Hase Y, Trung KH, Matsunaga T, Tanaka A. A mutation in the *uvi4* gene promotes progression of endo-reduplication and confers increased tolerance towards ultraviolet B light. *Plant J* 2006; **46**:317-326.
 - 28 Lloyd AH, Ranoux M, Vautrin S, *et al.* Meiotic gene evolution: can you teach a new dog new tricks? *Mol Biol Evol* 2014; **31**:1724-1727.
 - 29 Sekine D, Ohnishi T, Furuumi H, *et al.* Dissection of two major components of the post-zygotic hybridization barrier in rice endosperm. *Plant J* 2013; **76**:792-799.
 - 30 Nonomura K, Nakano M, Fukuda T, *et al.* The novel gene *HOMOLOGOUS PAIRING ABERRATION IN RICE MEIOSIS1* of rice encodes a putative coiled-coil protein required for homologous chromosome pairing in meiosis. *Plant Cell* 2004; **16**:1008-1020.
 - 31 Shao T, Tang D, Wang K, *et al.* OsREC8 is essential for chromatid cohesion and metaphase I monopolar orientation in rice meiosis. *Plant Physiol* 2011; **156**:1386-1396.
 - 32 Bai X, Peirson BN, Dong F, Xue C, Makaroff CA. Isolation and characterization of SYN1, a *RAD21*-like gene essential for meiosis in *Arabidopsis*. *Plant Cell* 1999; **11**:417-430.
 - 33 Bhatt AM, Lister C, Page T, *et al.* The *DIF1* gene of *Arabidopsis* is required for meiotic chromosome segregation and belongs to the REC8/RAD21 cohesin gene family. *Plant J* 1999; **19**:463-472.
 - 34 Yokobayashi S, Yamamoto M, Watanabe Y. Cohesins determine the attachment manner of kinetochores to spindle microtubules at meiosis I in fission yeast. *Mol Cell Biol* 2003; **23**:3965.
 - 35 Crismani W, Girard C, Mercier R. Tinkering with meiosis. *J Exp Bot* 2013; **64**:55-65.
 - 36 Cromer L, Heyman J, Touati S, *et al.* OSD1 promotes meiotic progression via APC/C inhibition and forms a regulatory network with TDM and CYCA1;2/TAM. *PLoS Genet* 2012; **8**:e1002865.
 - 37 Iwata E, Ikeda S, Matsunaga S, *et al.* GIGAS CELL1, a novel negative regulator of the anaphase-promoting complex/cyclosome, is required for proper mitotic progression and cell fate determination in *Arabidopsis*. *Plant Cell* 2011; **23**:4382-4393.
 - 38 Ma X, Zhang Q, Zhu Q, *et al.* A robust CRISPR/Cas9 system for convenient, high-efficiency multiplex genome editing in monocot and dicot plants. *Mol Plant* 2015; **8**:1274-1284.
 - 39 Karimi-Ashtiyani R, Ishii T, Niessen M, *et al.* Point mutation impairs centromeric CENH3 loading and induces haploid plants. *Proc Natl Acad Sci USA* 2015; **112**:1504333112.
 - 40 Maheshwari S, Tan EH, West A, Franklin FCH, Comai L, Chan SWL. Naturally occurring differences in CENH3 affect chromosome segregation in zygotic mitosis of hybrids. *PLoS Genet* 2015; **11**:e1004970.
 - 41 Kuppu S, Tan EH, Nguyen H, *et al.* Point mutations in centromeric histone induce post-zygotic incompatibility and uniparental inheritance. *PLoS Genet* 2015; **11**:e1005494.
 - 42 Kelliher T, Starr D, Wang W, *et al.* Maternal haploids are preferentially induced by CENH3-tailswap transgenic complementation in maize. *Front Plant Sci* 2016; **7**:1-11.
 - 43 Kermicle J. Androgenesis conditioned by a mutation in maize. *Science* 1969; **217**:446-448.
 - 44 Eder J, Chalyk S. *In vivo* haploid induction in maize. *Theor Appl Genet* 2002; **104**:703-708.
 - 45 Zhang Z, Qiu F, Liu Y, Ma K, Li Z, Xu S. Chromosome elimination and *in vivo* haploid production induced by Stock 6-derived inducer line in maize (*Zea mays* L.). *Plant Cell Rep* 2008; **27**:1851-1860.
 - 46 Barret P, Brinkmann M, Beckert M. A major locus expressed in the male gametophyte with incomplete penetrance is responsible for *in situ* gynogenesis in maize. *Theor Appl Genet* 2008; **117**:581-594.
 - 47 Conner JA, Mookkan M, Huo H, Chae K, Ozias-Akins P. A parthenogenesis gene of apomict origin elicits embryo formation from unfertilized eggs in a sexual plant. *Proc Natl Acad Sci USA* 2015; **112**:11205-11210.
 - 48 Horst NA, Katz A, Pereman I, Decker EL, Ohad N, Reski R. A single homeobox gene triggers phase transition, embryogenesis and asexual reproduction. *Nat Plants* 2016; **2**:15209.
 - 49 Proost S, Van Bel M, Vaneechoutte D, *et al.* PLAZA 3.0: an access point for plant comparative genomics. *Nucleic Acids Res* 2015; **43**:D974-D981.
 - 50 Biegert A, Mayer C, Remmert M, Söding J, Lupas AN. The MPI Bioinformatics Toolkit for protein sequence analysis. *Nucleic Acids Res* 2006; **34**:W335-W339.
 - 51 Larmande P, Gay C, Lorieux M, *et al.* Oryza Tag Line, a phenotypic mutant database for the Génoplante rice insertion line library. *Nucleic Acids Res* 2008; **36**:1022-1027.
 - 52 Sallaud C, Gay C, Larmande P, *et al.* High throughput T-DNA insertion mutagenesis in rice: a first step towards *in silico* reverse genetics. *Plant J* 2004; **39**:450-464.
 - 53 An S, Park S, Jeong DH, *et al.* Generation and analysis of end sequence database for T-DNA tagging lines in rice. *Plant Physiol* 2003; **133**:2040-2047.
 - 54 Droc G, Ruiz M, Larmande P, *et al.* OryGenesDB: a database for rice reverse genetics. *Nucleic Acids Res* 2006; **34**:D736-D740.
 - 55 Droc G, Périn C, Fromentin S, Larmande P. OryGenesDB 2008 update: database interoperability for functional genomics of rice. *Nucleic Acids Res* 2009; **37**:D992.
 - 56 Mieulet D, Diévar A, Droc G, Lanau N, Guiderdoni E. Reverse genetics in rice using Tos17. *Methods Mol Biol* 2013; **1057**:205-221.
 - 57 Kron P, Husband BC. Using flow cytometry to estimate pollen DNA content: improved methodology and applications. *Ann Bot* 2012; **110**:1067-1078.

(Supplementary information is linked to the online version of the paper on the *Cell Research* website.)

4.3. Discussion et perspectives

Ososd1 est bien l'orthologue d'osd1

Chez *A. thaliana*, *UVI4* (*UV-B-insensitive 4*) est le paralogue de *OSD1*. Cependant, *UVI4* a une fonction redondante avec *OSD1* mais n'est pas spécifique de la méiose puisqu'il a également un rôle au niveau de la mitose. Chez les céréales ces deux groupes de protéines existent, et chez le riz, il existe un seul gène pour chaque groupe : Os02g37850 et Os04g39670. Nous avons montré qu'une mutation du gène Os02g37850 est suffisante pour obtenir le phénotype *osd1* et donc nous avons appelé ce gène *OsOSD1*. Au moment de l'étude, nous ne disposions pas d'insertion dans le gène Os04g39670. Toutefois, il serait intéressant de créer des mutations par la technologie CRISPR/Cas9 pour voir si, comme chez *A. thaliana*, ce gène permettrait de produire des gamètes diploïdes. Il serait alors possible de définir si ce gène a une fonction à la fois mitotique et méiotique comme chez *A. thaliana*. En revanche, si la mutation du gène Os04g39670 permet de produire des gamètes non réduits et qu'il présente également une fonction mitotique, il sera moins intéressant que *OsOSD1* pour manipuler la méiose et parvenir à l'apomixie. Nous avons vu qu'il existe d'autres possibilités pour copier le phénotype *MiMe* telles que la mutation du gène de *TAM* nécessaire à la transition de la première à la seconde division de méiose (D'Erfurth et al. 2010) ou l'expression d'un dominant négatif de *TDM1* (Cifuentes et al. 2016).

De nombreuses mutations conduisant à la production de gamètes non réduits ont été identifiées notamment chez l'orge, la pomme de terre, la luzerne et le blé. Cependant, les gènes responsables n'ont pas été clonés (Ronceret and Vielle-Calzada 2015). Chez *A. thaliana*, les gènes *PS1* (*parallel spindle 1*) et *JAS* (*jason*) sont impliqués dans la formation du fuseau durant la seconde division de méiose. Aussi, les mutants *ps1* et *jas* conduisent à la formation de gamètes non réduits, souvent non clonaux, avec une fréquence de 65% et 25% respectivement. La première observation d'une réelle apoméiose a été réalisée chez le mutant *dyad* qui est allélique à *SWI1* et code pour une protéine *coil-coiled* essentielle à l'entrée en méiose. Cependant la mutation *dyad* ne forme que 0,2% de gamètes femelle non réduits viables (Ravi et al. 2008) et ce mutant n'a pas été reproduit chez le riz.

Ososd1-1 et OsMiMe présentent une fertilité incomplète

Nous avons montré que la délétion de *OSOSD1* en combinaison avec la délétion de *PAIR1* et *OSREC8* permet de reproduire le phénotype *MiMe*. La lignée *Ososd1-1 -/-* (AMBA12, ADN-T), présente une fertilité réduite par rapport au sauvage, alors que chez *A. thaliana*, la fertilité du mutant *osd1 -/-* n'est pas affectée. Il est probable que la baisse de fertilité observée chez *Ososd1-1 -/-* soit responsable de la restauration non complète de la fertilité chez le triple mutant. En effet le triple mutant est plus fertile que le simple mutant *Ososd1-1*. Une des hypothèses est que le mutant *Ososd1-1* conserverait un faible niveau d'expression permettant

la formation de gamètes non viables affectant la fertilité. Pour tester cette hypothèse, il serait intéressant de produire par CRISPR/Cas9 de nouveaux allèles mutants dans le gène *OsOSD1* et de mesurer leur fertilité pollinique et paniculaire. Une autre hypothèse est que cette fertilité incomplète résulterait de la formation d'embryons triploïdes à partir de gamètes réduits et non réduits. Chez le mutant *osd1* d'*Arabidopsis*, 100% des gamètes mâles et seulement 85% des gamètes femelles sont diploïdes, ce qui conduit à l'apparition de plantes triploïdes. Si le pollen est réduit (n) alors que la cellule œuf et la cellule centrale sont issues d'une mégaspore non réduite ($2n$), l'albumen du riz sera pentaploïde au lieu de triploïde, ce qui pourrait poser un problème de viabilité. Nos observations semblent conforter cette hypothèse.

Là encore, la reconstitution du triple mutant *OsMiMe* pourrait être reproduite par CRISPR/Cas9. Les nouveaux vecteurs de multiplexages de guides conduisent à une haute fréquence de mutation homozygote dès la première génération (T_0) et ils ciblent simultanément plusieurs gènes avec une fréquence très élevée. Trois sgRNA ciblant simultanément les gènes *PAIR1* (ou *SPO11-1*), *REC8* et *OSD1* pourraient être utilisés pour générer d'autres événements de triple mutation homozygote, ce qui permettrait de approfondir le phénotypage du mutant *MiMe*. Une restauration totale de la fertilité et une absence d'autre phénotype est en effet d'une grande importance pour exploiter, au niveau agronomique, un hybride synthétique apomictique basé sur *MiMe*.

Le mutant *OsMiMe* peut être utilisé pour tester des stratégies d'induction de développement parthénogénétique

A ce jour, les mécanismes contrôlant la parthénogenèse chez les plantes naturellement apomictiques restent largement inconnus (Barcaccia and Albertini 2013). Aucun gène contrôlant le développement apomictique des ovules n'a été identifié et il n'existe pas non plus de lien établi entre les mécanismes causant la formation de gamètes femelle non réduits et ceux de l'embryogenèse. De plus, les progrès réalisés dans l'identification de gènes responsables du développement autonome de l'embryon dans les espèces sexuées restent modestes. Chez l'orge, le mutant *hap* (pour *haploid initiation*) produit des plantes haploïdes représentant 30 à 40% de sa descendance. Chez cette même plante, l'induction d'haploïdes est aussi causée par l'élimination du génome dans le croisement entre *Hordeum bulbosum* (Hb) et *Hordeum vulgare* (Hv). Cette propriété, accompagnée d'un sauvetage d'embryon, est utilisée pour produire des haploïdes doublés depuis plusieurs décennies. Nous avons vu en introduction que chez *A. thaliana*, *msi1* (*multicopy of suppressor of ira1*) est le seul mutant identifié comme causant l'initiation du développement embryonnaire sans fécondation. Cependant, le développement s'arrête après une ou deux divisions (Guitton and Berger 2005). Plus récemment des embryons haploïdes ont été obtenus par élimination d'un génome parental par manipulation de *CENH3* (pour *Centromeric Histone 3*), un variant de l'histone centromérique H3 (Ravi and Chan 2010). Analysant le rôle de *CENH3* durant la ségrégation

des chromosomes, Ravi et Chan ont mis en évidence que le construit de complémentation du mutant *cenh3* qui comprenait un tag GFP remplaçant la partie N-terminale de CENH3 (CENH3-TAIL SWAP) formait une faible fréquence de descendances haploïdes par élimination du génome paternel. La fréquence d'induction d'haploïdes a été améliorée en utilisant deux variants de CENH3 (GFP-CENH3 et GFP TAILSWAP) accumulés en compétition dans le fond mutant *cenh3* pour créer l'élimination du génome (GEM : *Genome Elimination*), induite par un mélange de variants CENH3 (Marimuthu et al. 2011). La pollinisation des mutants apoméiotiques *dyad* et *MiMe* par GEM permet d'obtenir des plantes apomictiques avec une fréquence de 13% et 34% respectivement (Marimuthu et al. 2011). La possibilité d'utiliser l'ingénierie de CENH3 pour l'induction d'haploïdes chez le maïs a été récemment démontrée (Kelliher et al. 2016). Il a l'avantage de produire à la fois des haploïdes paternel et maternel même si la fréquence de formation des haploïdes maternels reste très faible (<1%). Aucun mutant CENH3 n'étant décrit chez le riz, nous avons recherché une insertion dans les collections de lignées d'insertion. Nous avons identifié une lignée (AULB08, *Tos17*, OTL) interrompant un intron du gène. Les plantes homozygotes pour l'insertion étaient viables, même si elles présentaient un retard de développement par rapport au ségrégant nul de la lignée, ce qui laisse supposer que l'insertion ne créait pas de KO. Bien que relativement complexe à mettre en œuvre et encore imparfait, le système GEM introduit chez le riz pourrait être testé en croisement avec le triple mutant *MiMe*. Dans ce croisement, il faudra regarder avec attention le développement de l'albumen, car contrairement à *A. thaliana*, le grain de riz a ses réserves en dehors du cotylédon.

Une alternative à l'utilisation de GEM est le gène récemment cloné dont la mutation est responsable de l'induction d'haploïdes chez le maïs. Des lignées inductrices d'haploïdes qui semblent toutes dérivées d'une ressource commune appelée stock6 ont été identifiées depuis la fin des années 50 chez le maïs. La fréquence d'induction d'haploïdes variant entre 7 et 15% est dépendante de la lignée inductrice mais également de la lignée receveuse. L'étude de la base génétique du caractère a permis d'identifier 8 QTLs dans les lignées inductrices et 2 QTL chez les lignées receveuses pour le caractère (Gilles et al. 2017). Le QTL présent sur le chromosome 1 du maïs représente près de 66% de la variance du caractère a été baptisé gynogenesis inducer 1 (*ggi*) (Barret et al. 2008) et récemment cloné simultanément par deux équipes indépendantes (Gilles et al. 2017; Kelliher et al. 2017). Le gène causatif baptisé *NOT LIKE DAD (NLD) / MATRILINEAL*, code pour une phospholipase de la famille de la patatine exprimée spécifiquement dans le pollen qui serait impliquée dans l'homéostasie des lipides de la membrane des cellules spermatiques, peut-être dans maintien de l'intégrité de leur membrane ou dans une fonction de signalisation. La version inductrice de NLD est une mutation entraînant un décalage du cadre de lecture. L'induction de mutations par TALENs dans le gène convertit une lignée non inductrice en lignée inductrice pouvant induire entre 4 et 12.5% d'embryons haploïdes chez la lignée femelle receveuse (Kelliher et al. 2017). Un

orthologue putatif ayant été identifié chez le riz (Widiez com. Pers.), le KO de ce gène par CRISPR en fond *MiMe* ou en fond sauvage (en multiplexant avec des sgRNA ciblant simultanément *NLD*, *PAIR1*, *REC8* et *OSD1*) serait possible comme preuve de concept, bien que la fréquence de plantes apomictiques attendue soit de l'ordre de 1 à 10%. On peut cependant s'attendre à de nouveaux développements améliorant cette fréquence chez le maïs soit par ingénierie de *NLD* soit par le clonage des gènes sous tendant les QTL responsables du reste de la variation du caractère inducteur, par exemple le QTL *qhir8* qui compte pour 20% de sa variance génétique (Liu et al. 2015).

Toujours récemment, un autre système candidat pour induire le développement parthénogénétique est issu du locus intervenant dans l'apospore chez l'espèce apomictique sauvage) qui ségrège comme un locus unique dominant, *l'apospory specific dominant region* (*PsASGR*). De nombreuses copies du gène *PsASGR-BABY BOOM Like* (*PsASGR-BBML*) résidant au locus, ce gène a été identifié comme un candidat très important dans la fonction apomictique de parthénogenèse car sa perte chez la plante apomictique *Cenchrus* (*Poaceae*) induit une inaptitude à entreprendre la parthénogenèse (Conner et al. 2015).

Un autre faisceau d'évidence provient de la similarité des gènes *ASGR-BBML* avec le gène *BABY BOOM* d'*Arabidopsis* (*BBM*) et des *Brassicaceae*, un facteur de transcription de la famille AP2 semblable à *ANT* (*Aintegumenta*) qui été montré comme directement impliqué dans le développement de la graine. La surexpression de *AtBBM* chez *Arabidopsis* est également connue pour induire la formation d'embryons somatiques sur les racines et les germinations (Boutillier et al. 2002) et celle de *ZmBBM* d'induire la régénération et l'aptitude à la transformation en culture de tissus chez les monocotylédones (Lowe et al. 2016). Le gène *PsASGR-BBML* est exprimé dans la cellule œuf avant la fécondation et peut induire la parthénogenèse et la production de descendances haploïdes chez le mil cultivé transgénique (Conner et al. 2015). L'expression de différents construits plaçant l'ADNc ou le clone génomique de *BBML* sous différents promoteurs (un fragment de 2Kb comprenant le promoteur de *PsASGR-BBML* ou le promoteur du gène d'*Arabidopsis* *DD45* dont l'activité est spécifique de la cellule œuf) a également été montré comme pouvant induire la parthénogenèse chez le maïs et le riz conduisant à la formation de plantes haploïdes (Conner et al. 2017). Cependant les fréquences obtenues restent faibles et de nombreux embryons connaissent un développement avorté indiquant qu'une optimisation du système est nécessaire.

MiMe pourrait être utilisé pour tester l'induction de plantes apomictiques basée sur l'expression de *BBML*. Cette transformation demanderait de sélectionner précocement par génotypage les 1/64^{ème} des grains issus de l'autofécondation du triple hétérozygote pour ne conserver que les cals issus de grains triple homozygotes. Une autre stratégie est d'utiliser un construit ciblant à la fois par CRISPR/Cas9 les 3 gènes causatifs du phénotype *MiMe* et permettant une

expression spécifique de la cellule œuf de *BBML*. Pour que la démonstration soit totalement élégante elle pourrait être effectuée en fond génétique hybride F1 (voir partie IV) pour étudier le maintien de l'hétérosis dans les plantes issues des grains apomictiques synthétiques. Cette exemplification préfigurerait peut-être l'utilisation ultérieure du système en situation agronomique.

Partie III : AUGMENTER LA RECOMBINAISON

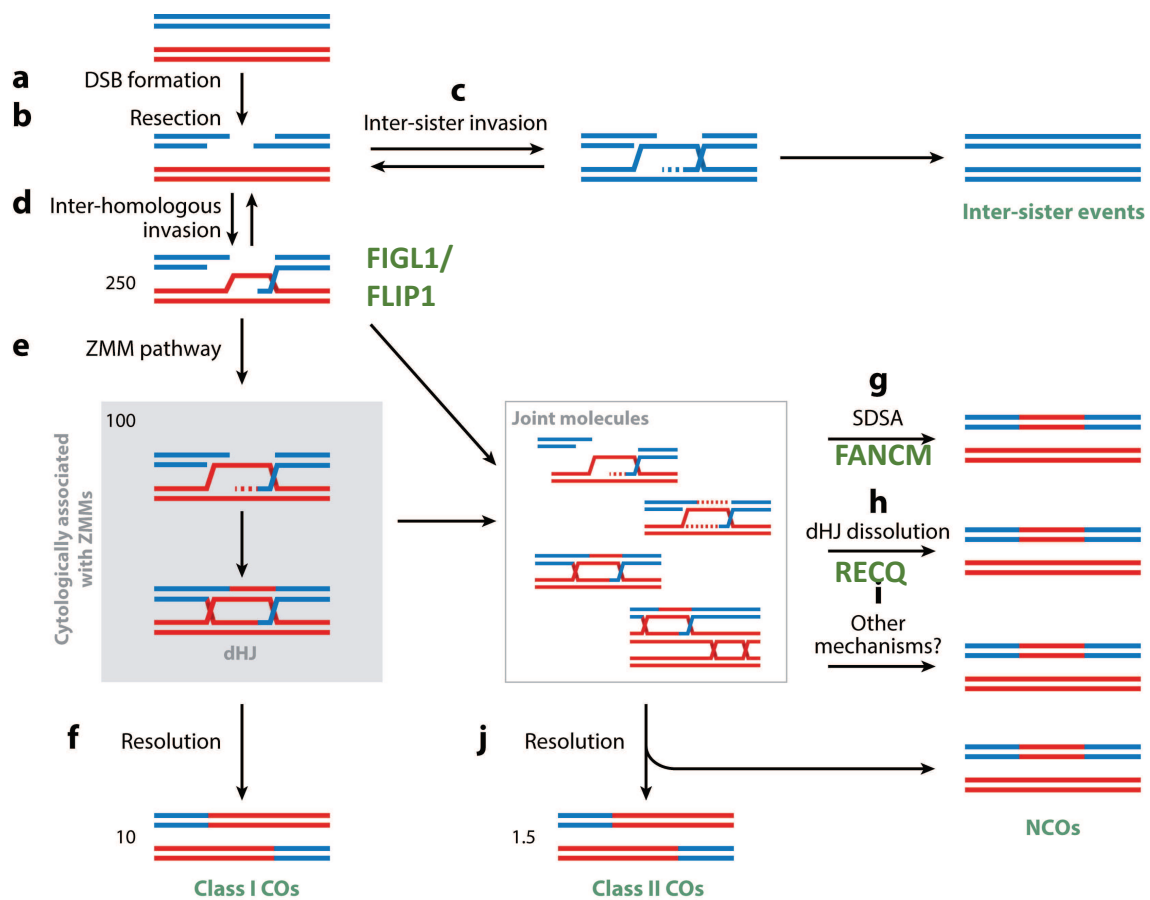


Figure 27 : Les voies de formation des COs et les protéines anti-CO impliquées (Mercier et al. 2015).

L'AAA-ATPase *FIGL1* contrôle négativement la dynamique de *RAD51* et *DMC1* lors de l'étape d'invasion du chromosome homologue. Les hélicases *FANCM* et *RECQ1/4* vont défaire les intermédiaires post invasion pour produire des NCOs (Non crossing-over) dans la voie formant les CO de classe II.

PARTIE III : Augmenter la recombinaison

5. AUGMENTER LE TAUX DE RECOMBINAISON CHEZ LE RIZ EN UTILISANT LES PROTEINES ANTI-CO *FANCM* ET *RECQL4*

5.1. Introduction

Comme nous l'avons vu en introduction, il existe au moins deux voies de formation des COs chez les plantes : les COs de classe I, qui sont sensibles à l'interférence et les COs de classe II, qui, par opposition, sont dit non-interférents. L'interférence désigne le phénomène qui tend à espacer les COs les uns des autres ce qui réduit la probabilité que deux COs se forment l'un à côté de l'autre. Il existe aussi un second phénomène appelé homéostasie qui tend à maintenir un nombre relativement constant de CO par méiocyte quand le nombre de CDBs varie. De ce fait, l'augmentation du nombre de CO dans une région du génome entraînerait une diminution du nombre de CO dans d'autres régions (Berchowitz and Copenhaver 2010) (§ 3.4.1.1). Chez *A. thaliana* comme chez la plupart des organismes, 95 % des CDBs sont réparées sans former de CO. Ainsi, sur 250 CDBs, une dizaine sont réparées en COs (Figure 27). 90% sont des COs de Classe I et 10% sont des COs de classe II (Figure 27).

La voie de formation des COs de classe I est contrôlée par les protéines de la famille ZMM (MSH4, MSH5, MER3, SHOC1, PTD, ZIP4/SPO22, et HEI10) ainsi que par MLH1 et MLH3. Il a été observé chez *A. thaliana* que l'inactivation des gènes de la famille ZMM induisait jusqu'à 85% de réduction de la fréquence des COs (Mercier et al. 2005; Chelysheva et al. 2007; Higgins et al. 2008; Macaisne et al. 2008), cette diminution s'accompagnant d'une forte réduction de la fertilité. Une augmentation du nombre des COs dans un fond mutant *zmm* déficient en COs devrait en toute logique s'accompagner d'une restauration partielle ou totale de fertilité. Partant de cette idée (Crismani et al. 2012) ont mis en place un crible suppresseur chez *A. thaliana* en fond mutant *zmm* et ont ainsi pu identifier 3 groupes de gènes qui régulent négativement la formation des COs. Les trois "voies" anti-CO identifiées impliquent l'hélicase *FANCM* (Crismani et al. 2012; Girard et al. 2014) ; l'AAA-ATPase *FIDGETIN-LIKE1* (Girard et al. 2015) et enfin les hélicases *RECQ4A* et *RECQ4B* (Bonnet et al. 2013; Séguéla-Arnaud et al. 2015). L'inactivation de l'un ou plusieurs de ces gènes permet d'augmenter très significativement la recombinaison méiotique, un facteur maximum d'augmentation de 7,8 fois dans un contexte génétique hybride par rapport au témoin sauvage étant obtenu chez le triple mutant *figl1 recq4a/b* (Fernandes et al. 2017b) (§ 3.2.6).

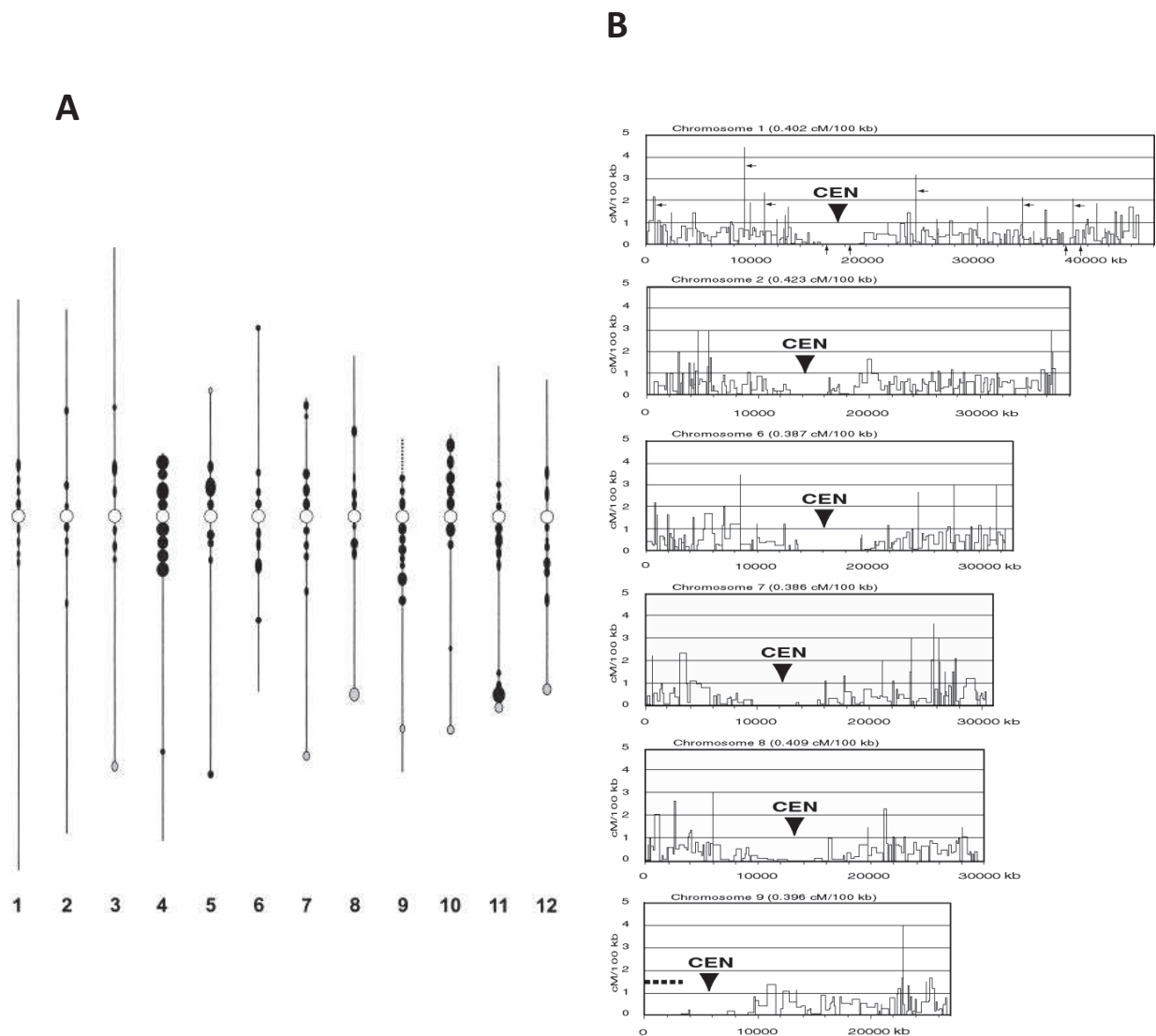


Figure 28 : Compaction de la chromatine et premières données de recombinaisons chez le riz (d'après Cheng et al. 2001 et Wu et al. 2003).

A- Compaction de la chromatine (Cheng et al. 2001).

Les cercles noirs représentent les régions hétéro-chromatiques. La taille du cercle est proportionnelle à l'intensité du marquage au DAPI. Les cercles ombrés représentent les régions marquées au DAPI qui n'existent que chez le riz indica (Zhongxian 3037) et non chez le riz japonica (Nipponbare). Les cercles vides représentent la position des centromères. La ligne en pointillé à l'extrémité du bras court du chromosome 9 représente la section de l'ADNr.

B – Fréquence de la recombinaison en fonction de la distance physique sur six chromosomes de riz (Wu et al. 2003).

Sur chaque graphique, l'axe horizontal représente la distance physique. L'axe vertical représente le rapport relatif de la distance génétique sur la distance physique. La valeur moyenne du taux de recombinaison sur chaque chromosome est donnée entre parenthèses au-dessus de chaque graphique. Les triangles noirs représentent la position des centromères (CEN) sur chaque chromosome. Les petites flèches sur le chromosome 1 indiquent les positions des points chauds (flèches horizontales) et des points froids (flèches verticales). La ligne pointillée sur le chromosome 9 indique la région NOR (pour *Nucleolus Organizer Region*).

La distribution des COs le long des chromosomes ne se fait pas au hasard et n'est pas uniforme. De grandes différences de fréquence de recombinaison peuvent être observées à différentes échelles. Le long des chromosomes on observe une alternance de régions à forte et faible recombinaison et en général les centromères sont dépourvus de CO. Les régions à haute fréquence de CO sont en général associées avec une forte densité en gènes. Ces considérations ne sont cependant pas généralisables à toutes les espèces. Chez le riz, les premières cartes haute densité des chromosomes ont été établies dans les années 90 (Harushima et al. 1998; Wu et al. 2002). L'intégration des cartes physiques et génétiques a permis d'établir les rapports entre cM et kb et leur fluctuation le long des chromosomes qui est en moyenne de 240 kb pour 1 cM (Chen et al. 2002; Wu et al. 2003). Le séquençage des chromosomes 1, 2, 6, 7, 8 et 9 du riz a montré une recombinaison nulle au niveau des centromères occupant entre 2,6 et 5,6 Mb soit 11% de la longueur de ces chromosomes, et à une échelle locale 26 régions chaudes de recombinaison (recombinaison de 5 cM pour 250 kb) et 15 régions froides (0 cM pour une distance allant de 310 à 620 Kb) (Wu et al. 2003). Les bras courts des chromosomes 4 et 10, très hétérochromatiques (Figure 28) (Cheng et al. 2001) ont une recombinaison très réduite (Chen et al 2002, Zhao et al 2002). Plus récemment, deux études utilisant un marquage haute densité (31 000 SNP) de 176 RILs issues du croisement indica x japonica (ex : IR64 x Azucena) (Spindel et al. 2013) et un séquençage 39 x de 41 plantes F2 du croisement PA64 x 93-11 (Si et al 2015) ont confirmé la large variation le long et entre les chromosomes du rapport distance génétique à distance physique. Dans cette dernière étude, un nombre moyen de 33,9 CO et 2,27 non CO a été détecté par plante F2 14% (52 Mb) du génome incluant 11 des 12 régions centromériques étant dépourvu de COs et 0,72% (2,7 Mb) du génome montrant une très haute fréquence de COs (>50cM/Mb). La tendance observée chez le riz comme chez beaucoup de plantes est une augmentation de la fréquence de COs depuis les régions péri-centromériques vers les télomères et concomitante avec une augmentation de la densité en gènes. Même si cette transition est beaucoup moins accentuée que chez des céréales à plus grand génome comme le blé ou le maïs (Figure 25). L'ensemble des COs observés, sur le chromosome 3B du blé (1 Gb), est par exemple regroupé sur seulement 13% de la longueur du chromosome (Choulet et al. 2014). Il a été montré que les conditions environnementales comme la température ou la disponibilité en nutriment peut influencer le nombre et la distribution des COs méiotiques. En effet, une élévation de la température de 22°C à 30°C pour le blé ou de 18°C à 28°C pour *A. thaliana* augmenterait le nombre de CO et modifierait leur distribution. Il semblerait également qu'une meilleure nutrition azotée et phosphatée améliore le taux de recombinaison chez le blé, le seigle et le riz (Lambing et al. 2017). L'utilisation d'agent irradiant provoquant des cassures double-brin de l'ADN comme les UV ou l'utilisation de cisplatine (utilisé en chimiothérapie) permettent aussi d'augmenter le nombre de CO et, chez *A. thaliana*, l'utilisation de cisplatine permet une restauration partielle de la formation des bivalents chez un mutant *spo11*

(Lemmens et al. 2013). On observe aussi que le taux de recombinaison est augmenté chez une plante infectée par des pathogènes. Une étude récente montre que, chez les riz, un stress thermique, osmotique, biotique ou physique (UV) augmente chez certains (30%) des descendants la recombinaison sans que la différence soit significative au niveau d'une population F2 analysée (Si et al. 2015). La manipulation des gènes impliquées dans la méthylation permet aussi de modifier la distribution des COs. C'est ce qui a été observé chez *A. thaliana* lorsque l'on mute des protéines comme MET1 ou DDM1 qui sont impliquées dans la méthylation de l'ADN (Yelina et al. 2012). Une telle étude n'a pas encore été conduite chez le riz. Il a simplement été montré que *Osmet1/2* dérégule l'expression de nombreux gènes en modifiant le profil de méthylation (Yelina et al. 2012; Hu et al. 2014).

Augmenter la fréquence de recombinaison pour révéler de nouvelles combinaisons alléliques ou casser une liaison étroite entre un locus possédant un allèle favorable associé en répulsion avec un autre locus possédant un allèle défavorable seraient d'un grand intérêt chez le riz. Il a par exemple été montré que l'allèle du gène *sd1* conférant le semi nanisme aux variétés de la révolution verte était lié à un allèle défavorable d'un QTL de rendement sous sécheresse situé dans un intervalle de 500 kb à 100 kb de distance de *sd1* (Vikram et al. 2015). Cette liaison était suffisamment forte pour avoir été entraînée dans l'ensemble des variétés de la révolution verte. Le manque de recombinants est également très limitant quand les parents utilisés dans les schémas de croisement sont distants génétiquement. C'est le cas des croisements inter-subspécifiques (*O. sativa japonica* x *O. sativa indica*) et des croisements interspécifiques avec les espèces sauvages apparentées. Ces croisements distants constituent un réservoir considérable d'allèles d'adaptation et parvenir à introgresser des gènes d'un pool génétique à un autre serait un outil formidable pour l'amélioration des variétés. Cependant, divers obstacles à l'obtention de recombinants dans ces types de croisements peuvent être anticipés comme l'existence d'hétérologies structurales entre régions à recombiner empêchant la recherche d'homologie sur le chromosome homologue nécessaire à la formation de CO, ainsi que l'action post zygotique de gènes de stérilité hybride mâle et femelle (Ouyang et al. 2010b). Un exemple d'hétérologie structurale est apporté par la région de 5 Mb d'introgression depuis *O. rufipogon* porteuse du gène de résistance à la pyriculariose *Pi33* présent dans la variété indica IR64 et qui s'est révélée très peu recombinogène (Ballini et al. 2008). Ce qui précède montre que modifier la fréquence globale des COs et leur distribution revêt une importance considérable pour l'amélioration du riz et des plantes d'intérêt agronomique en général.

Dans ce contexte nous avons choisi d'étudier chez le riz, l'effet des homologues des gènes *AtFANCM* et *AtRECQ4A/B*, sur la fréquence et la distribution de la recombinaison méiotique. Comme chez *A. thaliana* nous avons souhaité mesurer l'effet de ces gènes seul ou en combinaison pour voir si d'une part la fonction des gènes était conservée entre *A. thaliana* et *O. sativa* et d'autre part voir si leurs effets étaient additifs. Le double mutant *Atrecq4a/b* étant

celui qui permet d'augmenter le plus la recombinaison parmi les 3 gènes anti-CO découverts, nous avons choisi d'étudier plus particulièrement l'effet d'une mutation de l'unique orthologue chez le riz *OsRECQ14*. Ne sachant pas si les résultats chez le riz allaient être du même ordre que chez *A. thaliana*, nous avons également étudié l'effet de la mutation *fancm* et produit le double mutant *fancm recq14*.

Pour mener à bien cette étude, nous avons identifié des KO dans les trois gènes anti-CO ciblés (GC) et développé des populations hybrides F1 dans un fond génétique polymorphe permettant l'analyse des événements de recombinaisons par le génotypage des descendances F2. L'établissement de cartes génétiques permet ensuite de voir si le taux de recombinaison observé entre les populations sauvage et mutante est différent.

Nous avons essayé au cours de cette thèse de produire l'ensemble des populations F2 nécessaires à l'étude des 3 gènes anti-CO. Les résultats les plus avancés font actuellement l'objet de la rédaction d'un article et sont ceux obtenus par l'étude des simples mutants *OsRECQ14* et *OsFANCM*. Cet article en cours de rédaction inclura les résultats également obtenus chez deux autres espèces cultivées, le pois et la tomate. Il s'intitulera "Manipulating crossover frequency in crop". Les résultats les plus avancés font actuellement l'objet de la rédaction d'un article coordonnée par R. Mercier.

5.2. Matériel et Méthodes

5.2.1. Phylogénie : Recherche d'orthologues

Pour identifier les orthologues des gènes AT1G35530 (*AtFANCM*), AT1G10930 (*AtRECQ4A*) et AT1G60930 (*AtRECQ4B*), nous avons interrogé la plateforme d'analyse comparative des génomes PLAZA3.0 (Proost et al. 2015). L'alignement avec clustalW des séquences protéiques d'*A. thaliana* et d'*Oryza sativa* permet ensuite de valider que les domaines protéiques sont bien conservés.

5.2.2. Matériel génétique

5.2.2.1. Recherche de mutants d'insertion et génotypage

Comme vu dans la première partie (§1.5.1.2), le riz dispose de nombreux outils et en particulier de bases de données de mutants d'insertion ADN-T ou *Tos17* accessibles via différents navigateurs génomiques comme OryGenesDB développé par le CIRAD (Droc et al. 2006). Nous avons développé un outil de génotypage ("Genotyping Primer Designer" sur OryGenesDB) qui permet de générer automatiquement un set de 3 amorces permettant de déterminer si l'intégration du transgène est à l'état homozygotes, hétérozygotes ou sauvages dans la descendance analysée. Le génotypage par PCR des lignées se fait en suivant le

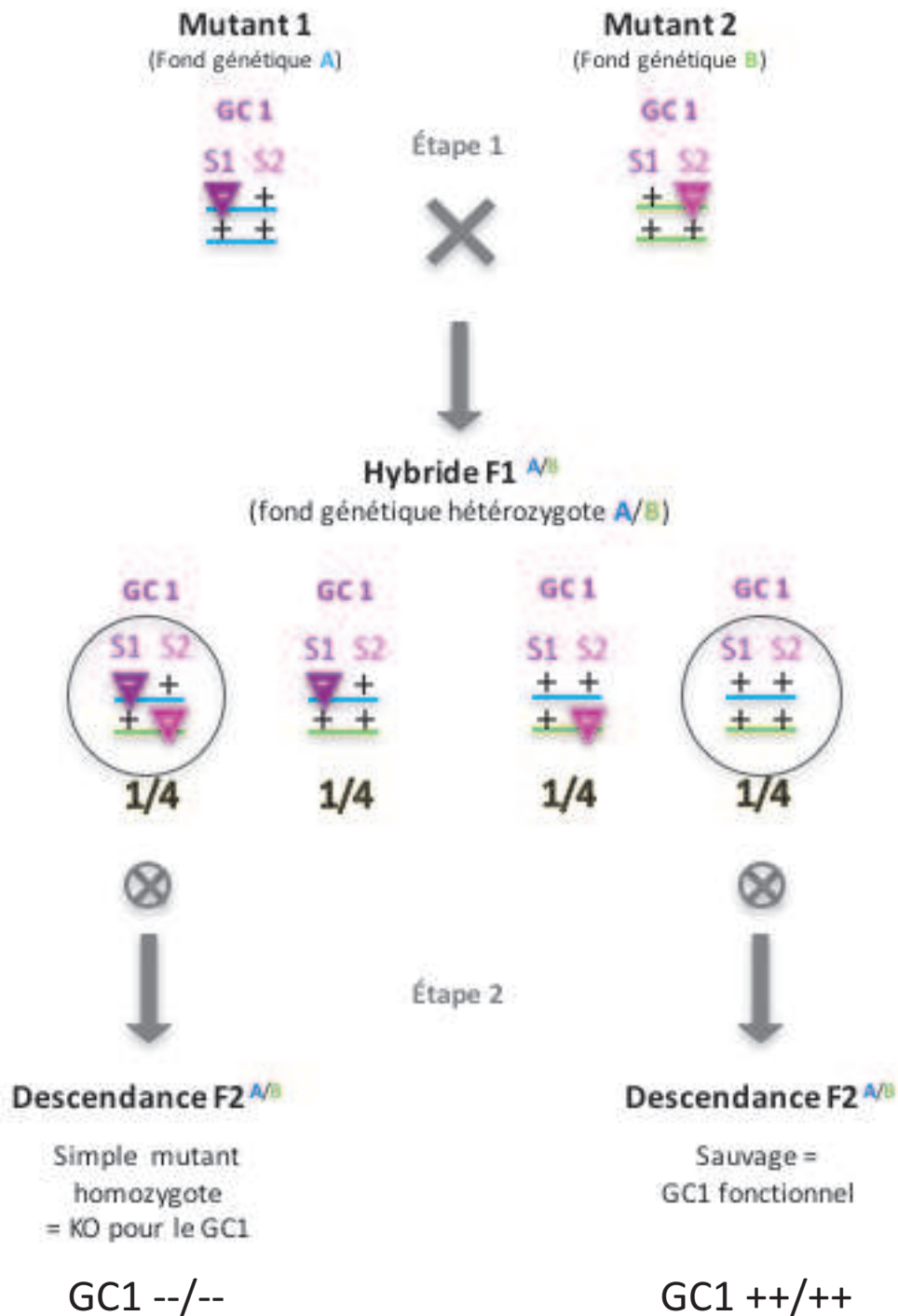


Figure 29 : Schéma de croisement pour l'étude d'un simple mutant anti-CO.

Pour un gène cible (GC) considéré, on utilise deux mutants. Le mutant 1 porte une mutation au site 1 (S1) et le mutant 2 une mutation au site 2 (S2). La mutation peut être à l'état hétérozygote (+/-) ou homozygote (++) pour chaque site considéré.

Chaque triangle représente un site d'insertion ADN-T ou une cible pour la CRISPR/Cas9. Chaque allèle est représenté par un + s'il est sauvage ou un - s'il est muté.

protocole publié (Mieulet et al. 2013). Les amorces utilisées pour génotyper sont listées en annexe (Annexe 1).

Les mutants utilisés dans cette étude sont les lignées d'insertion ADN-T AQSG07 et A46543 pour *OsFANCM* (Os11g07870) ainsi que 3A-03503 et AUFG12 pour *OsRECQ14* (Os04g35420). Les lignées AQSG07 et AUFG12 sont dans un fond génétique Nipponbare (NB), les deux autres sont dans un fond génétique Dongjin (DJ) (Annexe 1).

5.2.2.2. Caractérisation moléculaire des mutants ADN-T

La base de donnée d'expression, RiceXPro (Sato et al. 2013), indique que le gène *Os04g35420* est plus exprimé dans les inflorescences que dans les autres tissus (Annexe 3). Cependant, il n'y a pas de données d'expression pour le gène *Os11g07870*. Les ARN totaux des lignées 3A-3503, AUFG12, AQSG07 et A46543 ont été extraits (RNAzol®RT, RN190, MRC) à partir de jeunes inflorescences (stade méiocyte) mais aussi à partir de plantules de 10 jours. Un lot de 3 plantes homozygotes mutantes (-/-) et 3 plantes sauvages (+/+) par type de mutation soit 6 échantillons par lignée et les témoins variétaux NB et DJ ont été utilisés pour faire l'analyse quantitative. 3µg d'ARN total (dosage BioAnalyser 2100, Agilent) ont été transcrits (Superscript III, Invitrogen). Les ADNc des différents gènes (1µL dilution au 1/10^{ème}) ont été amplifiés en utilisant des amorces dans les exons flanquant le site d'intégration du transgène et en faisant des cycles de PCR de 30 ou 32 cycles (95 °C 20''; 60 °C 20''; 70°C 45'') (KOD Hot Start DNA Polymerase®, Novagen) (Annexe 2). Le gène de ménage *OsEF1α* a été utilisé comme contrôle endogène pour normaliser les échantillons. Les produits PCR ont ensuite été séparés par électrophorèse sur gel d'agarose à 1,5% et l'intensité de chaque bande a été quantifiée avec le logiciel ImageJ. Le niveau relatif d'expression des gènes a été comparé entre les descendants homozygotes mutants et sauvages. Nous avons aussi fait de la PCR quantitative avec des amorces positionnées dans la partie 3'UTR des gènes *FANCM* et *RECQ14* puis avec des amorces situées de part et d'autre du transgène. La chimie utilisée était le SYBR green (SYBR green I-Master 2X®, Roche) sur LightCycler 480 dans un volume réactionnel de 15µL. Pour chaque réaction, 3µL d'ADNc (dilution de 1/10 à 1/100), 7,5µL de 2X SYBR green, 0,75µL de chaque amorce (10µM) et 3µL d'eau ont été utilisés (Annexe 2) (Caldana et al. 2007).

5.2.2.3. Développement de populations F2 pour l'étude des gènes anti-CO

- ✓ Populations F2 permettant d'étudier l'effet d'un gène anti-CO : "simple mutant"

Pour étudier l'impact de la mutation d'un gène anti-CO, nous avons réalisé un croisement entre 2 mutants (mutant 1 et 2) provenant de deux fonds génétiques différents (A ou B) et portant chacun une insertion à l'état hétérozygote dans le gène cible "1" (GC1) (Figure 29). Dans le cas du mutant 1 qui est dans le fond génétique A le GC1 est inactivé par l'intégration d'un ADN-T au site 1 (S1). Dans le cas du mutant 2 le GC1 n'est plus transcrit à cause d'une

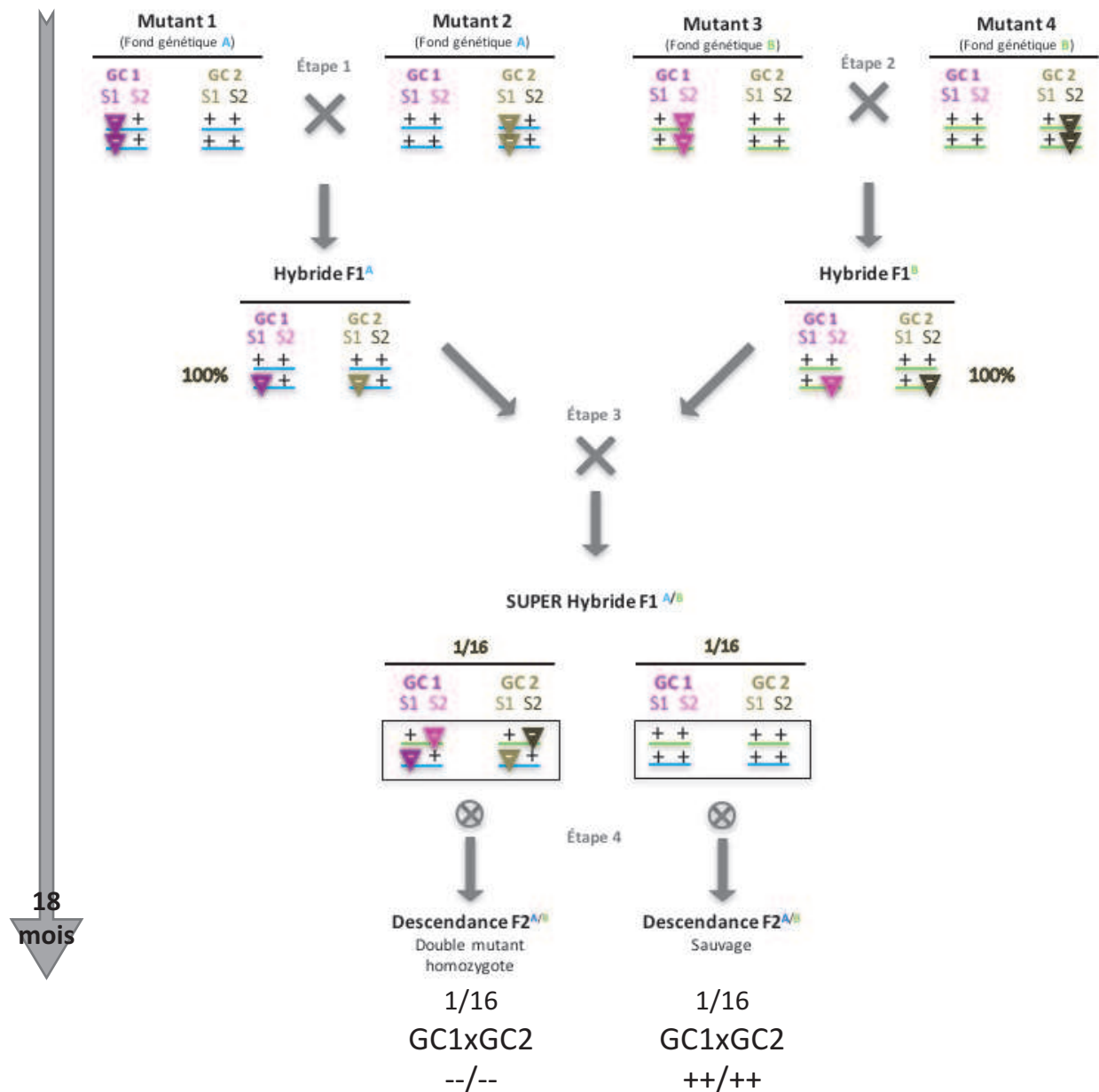


Figure 30 : Schéma de croisement pour l'étude d'un double mutant anti-CO.

Pour deux gènes cibles (GC), on utilise deux mutants par GC. Les mutant 1 et 2 sont mutés au site 1 (S1) du GC1 ou GC2 respectivement. Les mutants 3 et 4 sont mutés au site 2 (S2) du GC1 ou GC2 respectivement. La mutation peut être à l'état hétérozygote (+/-) ou homozygote (++) pour chaque site considéré.

Chaque triangle représente un site d'insertion ADN-T ou une cible pour la CRISPR/Cas9. Chaque allèle est représenté par un + s'il est sauvage ou un - s'il est muté.

mutation (ADN-T ou CRISPR/Cas9) au site 2 (S2). Le croisement ainsi réalisé permet d'obtenir en F1 $\frac{1}{4}$ de sauvages (S1 +/+ ; S2 +/+) et $\frac{1}{4}$ d'individus hétérozygote pour le GC1 au S1 et hétérozygote pour le GC1 au S2. Pour simplifier, on utilisera le terme d'homozygote au GC1, même si à chacun des deux sites de mutation l'insertion est hétérozygote (S1 +/- ; S2 +/-), à l'échelle du gène ciblé (S1+S2), l'insertion est sur les deux brins d'ADNs donc on peut considérer qu'elle est homozygote. L'autofécondation de plantes F1 homozygotes (hétéroalléliques) et des F1 sauvages permet d'obtenir deux populations F2 identiques d'un point de vue génétique à l'exception près du gène anti-CO étudié (GC1) (Figure 29).

✓ Populations F2 permettant d'étudier l'effet de deux gènes anti-CO : "double mutant"

Nous avons établi un schéma de croisement pour développer des population F2 permettant de mesurer l'effet cumulé de deux gènes anti-CO (GC1 et GC2) sur la recombinaison (Figure 30). Les deux premières étapes de ce schéma consistent à croiser dans le même fond génétique "A" ou "B" deux mutants homozygotes l'un pour le GC1 (GC1 S1-/- et S2+/+ ; GC2 S1+/+ et S2+/+) et l'autre pour le GC2 (GC1 S1+/+ et S2+/+ ; GC2 S1-/- et S2+/+). On obtient ainsi dans chacune des F1 100% d'individus hétérozygotes mutants pour les deux gènes dans un fond génétique A (Hybride F1^A) ou dans un fond génétique B (hybride F1^B). L'étape 3 est un croisement entre les deux F1 qui permet d'obtenir les super hybrides F1 dans un fond génétique contrasté (super F1^{A/B}). Dans cette population de super F1^{A/B} nous sélectionnons les 1/16 d'individus double mutants homozygotes (GC1 S1+/- et S2+/- ; GC2 S1+/- et S2+/-) et les sauvages apparentés (GC1 S1+/+ et S2+/+ ; GC2 S1+/+ et S2+/+). Comme précédemment nous considérons les GCs globalement, c'est à dire S1 et S2, c'est pourquoi nous nous permettons de dire que les mutations sont homozygotes hétéroalléliques. Comme pour l'étude du simple mutant, le schéma se poursuit en étape 4 par l'autofécondation de ces deux types de F1. Le génotypage haute densité des populations F2 permet de mesurer l'effet des mutations cumulées des deux gènes anti-CO sur la recombinaison (Figure 30).

5.2.3. Caractérisation phénotypique des F1

5.2.3.1. Méiose mâle : cytogénétique

Les inflorescences immatures sont prélevées puis fixées dans un mélange alcool / acide acétique (3:1) et conservées à 4°C (stade méiocyte §§1.6.2). Les étamines des plus jeunes épillets sont prélevées et triées en fonction de leur taille sous une loupe binoculaire (0,5 à 1mm de long). Les cellules mère de pollen (PMC) sont étalées sur lame en suivant le protocole mise au point sur la plateforme de Cytogénétique de l'INRA de Versailles (L. Cromer, IJPB, Versailles non publié) (Ross et al. 1997; Grelon et al. 2001). Les anthères sont disposées sur une lame sur laquelle une goutte d'acide acétique (45%) est ajoutée. Une cale constituée d'une lamelle en verre de 22x22 est disposée sur le bord de la lame et une lamelle (24x22) est ajoutée sur l'échantillon et la cale. Le dessous de la lame est ensuite chauffé à la flamme

quelques secondes et la préparation est délicatement écrasée avec le dos d'un crayon. La cale est ensuite retirée et la lame est à nouveau chauffée puis mise entre les feuilles d'un papier buvard plié en deux sur lesquelles une forte pression est appliquée. La lame est alors plongée dans l'azote liquide et la lamelle retirée avec un scalpel. La préparation est ensuite mise à sécher et une goutte de DAPI dans du Vectashield (2µg/ml) est ajoutée. L'observation se fait au microscope (ApoTome®, ZEISS) à l'objectif 100.

5.2.3.2. Fertilité paniculaire (que le simple REC sur F1)

La fertilité paniculaire est mesurée à partir de comptages réalisés sur les inflorescences des trois brins-maîtres de la plante de riz. Chaque épillet plein (donnant une graine) et épillet vide (stérile) est compté et on fait le ratio du nombre de grains sur le nombre total d'épillets pour avoir une estimation de la fertilité de la plante. Ce comptage est réalisé chez 3 à 10 plantes pour chaque mutant (GC -/-) et la valeur moyenne obtenue est comparée à celle observée sur un nombre équivalent de plantes sauvages (GC +/+).

5.2.4. Génotypage SNP

✓ Extraction ADN

Les ADNs génomiques des parents F1 et de leurs descendance F2 ont été extraits à partir de feuilles (12 cm du limbe de l'avant-dernière feuille émise sur une plantule de 21 jours coupé en segments de 1 cm qui sont réduits en poudre (Retsch MM311, Qiagen)) avec le robot d'extraction Biomek FX (Beckman Coulter, plateforme GPTR du CIRAD) et un kit d'extraction (NucleoMag® Plant Kit, Macherey-Nagel). Les ADNs sont dosés en utilisant un intercalant de l'ADN (Hoechst) et normalisés à une concentration de 50ng/µL.

✓ Sélection des SNP

Le projet IRIGIN (International Rice Genomic Initiative), financé par France Génomique, a permis, entre autres, le séquençage des lignées mutantes AQS et A4 et du cultivar Dongjin à une profondeur moyenne de 25x. Ces différentes banques des séquences ont suivi un traitement bioinformatique classique en 2 étapes. Dans un premier temps, les séquences ont été alignées sur le génome de référence Nipponbare en utilisant le logiciel BWA (Li and Durbin 2009), puis un appel de variants a été réalisé en utilisant le logiciel GATK (McKenna et al. 2010). Un second traitement a été effectué afin de sélectionner les SNPs pour lesquels on avait le plus de profondeur de séquençage. Les SNPs ont ensuite été choisis en fonction de leur position physique de sorte à couvrir l'ensemble du génome en distribuant en moyenne un marqueur tous les 1,5 Mb. Enfin nous avons retenu préférentiellement les SNPs qui étaient dans des exons et exclus ceux présents dans les éléments transposables (TE).



Figure 31 : Position des marqueurs Kasp® sur les chromosomes 1 à 6 du riz.

La position physique (en Mb) est indiquée à gauche du chromosome et le nom du Kasp® est noté à droite du chromosome.

Seuls les SNPs utilisés pour le génotypage Fluidigm en puce 96.96 sont indiqués (puce 1 en bleu; puce 2 en vert et puce 3 en violet).

✓ Design et validation de marqueurs KASP sur LC480

Nous avons choisi d'utiliser la technologie KASP™ (Kompetitive Allele Specific PCR) pour géotyper les descendance F2 aux locus SNP définis ci-dessus. Le design des amorces a été réalisé par la société LGC. Pour chaque SNP, un mélange de 3 amorces appelés "Mix assay" est fourni par LGC. Il comprend une amorce universelle et 2 amorces spécifiques de chacun des deux allèles du SNP et marquées par un fluorochrome FAM ou VIC. L'ensemble des *Mix assay* a été testé sur un set de 24 ADN comprenant les parents DJ et NB (6 répétitions), 10 hybrides F1 pris au hasard et de l'eau comme contrôle négatif (2 répétitions). Le protocole utilisé est celui recommandé par LGC pour une utilisation des Assay sur LightCycler 480 (Roche). Les ADN sont utilisés à une concentration finale de 10ng/μL et nous travaillons dans un volume réactionnel final de 4 μL en plaque 384 ce qui nous permet de tester 16 sondes et 24 ADN par plaque 384. Pour chaque réaction KASP, on mélange 2 μL de mix (60 μL de *Master Mix Low Rox* (2x), 1 μL de MgCl₂ (2,2mM), et 1,8 μL d'*Assay Mix* ; (LGC Genomic protocol) (Figures 31 et 32).

✓ Puce de géotypage 96.96

Le géotypage SNP a été réalisé en utilisant le système microfluidique développé par Fluidigm sur un appareil Biomark. Nous avons utilisé des puces de géotypage permettant de multiplexer 96 ADN avec 96 Assay (96.96 Genotyping (GT) Dynamic Array Chips, Fluidigm®) et suivi le protocole recommandé par Charles Poncet (Plateforme Gentyane, INRA) pour préparer les puces. Sur chaque puce, 6 puits sont réservés aux témoins (parents de chaque géotype, parent F1 et H₂O). Les ADN sont utilisés à une concentration de 10 ng/μL (4μL / puce).

✓ Analyse des données de géotypage

Les données de géotypage SNP ont été analysées avec le logiciel Fluidigm SNP genotyping Analysis (Biomark). Pour chaque marqueur, trois nuages de points sont attendus. Le nuage le plus collé à l'axe des abscisses contient les individus homozygotes pour l'allèle X (dans notre cas NB car il a toujours été marqué avec le fluorochrome FAM). Le nuage contre l'axe des ordonnées regroupe les individus homozygotes pour l'allèle Y (dans notre cas DJ marqué en VIC). Le troisième nuage, au centre, et dans l'idéal équidistant des deux autres, regroupe les individus qui ont les 2 allèles X et Y (hétérozygotes). Les individus qui n'appartiennent pas clairement à un des 3 nuages sont notés "U" (pour Unknown) ; les individus homozygotes sont notés X ou Y selon l'allèle considéré et les hétérozygotes sont notés H. Pour chaque SNP, si les contrôles ne se placent pas correctement dans les 3 nuages de points ou si les 3 nuages de points ne sont pas suffisamment bien séparés alors le marqueur est écarté de l'analyse ou la réaction est refaite sur une nouvelle puce.



Figure 32 : Position des marqueurs Kasp® sur les chromosomes 7 à 12 du riz.

La position physique (en Mb) est indiquée à gauche du chromosome et le nom du Kasp® est noté à droite du chromosome.

Seuls les SNPs utilisés pour le génotypage Fluidigm en puce 96.96 sont indiqués (puce 1 en bleu; puce 2 en vert et puce 3 en violet).

Après ordonnancement des marqueurs en fonction de leur position physique, les données "U" ont été révisées une par une, et on a parfois pu imputer certaines données quand la position de l'individu dans le graphique le permettait (3 nuages en général avec des points U qui penchent plus vers un nuage que l'autre). Après vérification des données de génotypage, les données brutes ont été triées par type de descendance F2 afin d'avoir un jeu de données pour le simple mutant (--/--) et un jeu de données pour le sauvage (double ségrégant nul).

5.2.5. Construction des cartes génétiques

Les deux jeux de données obtenus ont été analysés avec MapDisto 2.0 (Heffelfinger et al. 2017). Les groupes de liaison ont été établis, à partir des données F2 de la population sauvage, en utilisant un LOD de 5 et une fréquence de recombinaison maximum r de 0,3. Puis, pour chaque groupe de liaison, les marqueurs ont été ordonnés. Les erreurs de génotypage (singletons) ont été filtrées en utilisant la fonction "Iterative error removal" (Number of iteration = 5 ; Start threshold = 0,001; Increase = 0,001).

Du fait de son taux de recombinaison supérieur, l'application des mêmes critères aux données F2 de la population mutante conduit à un nombre de groupes de liaison plus important que le nombre de chromosomes. Les mêmes groupes de liaison et le même ordre de marqueurs ont donc été imposés lors de la construction la carte génétique de la population F2 mutante.

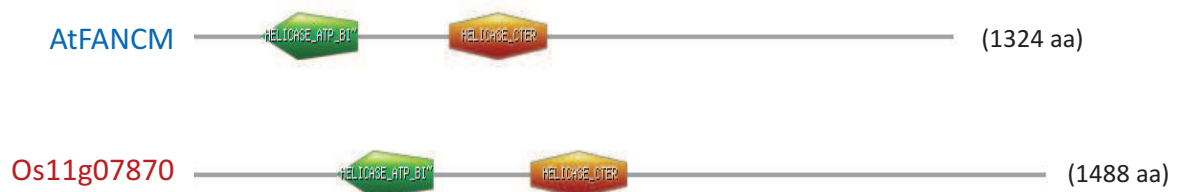
Le calcul des distances entre marqueurs a été fait en utilisant la fonction de Haldane et non Kosambi. Comme nous l'avons vu, la fonction de Kosambi tient compte de l'interférence et les CO surnuméraires que nous observons dans nos populations mutantes sont des CO de classe II non interférents.

Les fréquences de recombinaison par intervalles (r) et l'erreur moyenne de l'estimation de la fréquence (SEM) ont été calculés automatiquement par logiciel MapDisto et un programme R (ggplot2) a permis de réaliser une sortie graphique des résultats et de les mettre en perspective avec différentes données de génomique (nombre de gènes, nombre d'éléments transposables (TE), nombre de SNP par intervalles).

5.2.6. Condition de culture et de croisement de plantes

Les plantes ont été cultivées dans des serres de confinement (température de 28-31°C le jour et 24°C la nuit ; 50%-80% d'hygrométrie ; intensité lumineuse de 400-700 $\mu\text{moles/m}^2\text{s}$). Les croisements ont été réalisés manuellement en utilisant un protocole classique d'émasculatation avec une pompe à vide et de pollinisation par agitation de panicules matures au-dessus des plantes castrées. Les panicules ainsi fécondées ont été ensachées pendant toute la durée de maturation des grains hybrides, c'est à dire 35 jours.

FANCM



RECQ4

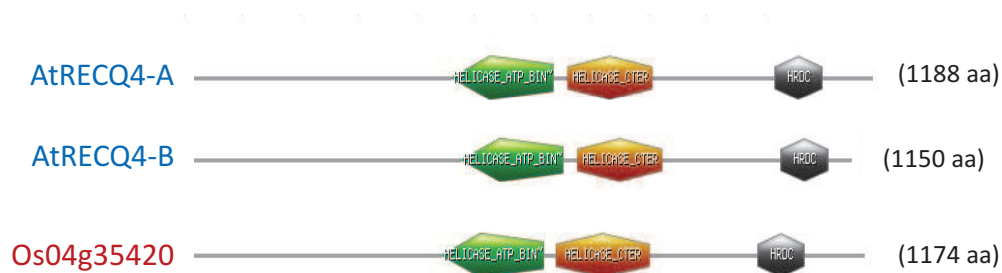


Figure 33 : Conservation des domaines protéiques entre *A. thaliana* et *Oryza sativa* pour FANCM et RECQ4 (Expasy).

A- FANCM chez *A. thaliana* et *Oryza sativa*. Conservation des domaines Hélicase ATP-binding type 1 et Hélicase C-terminal entre AtFANCM et Os11g07870 (OsFANCM).

B- RECQ4 chez *A. thaliana* et *Oryza sativa*. Conservation des domaines hélicase ATP-binding, hélicase_Cter et HRDC entre AtRECQ4-A, AtRECQ4-B et Os04g35420 (OsRECQ4).

5.3. Résultats

5.3.1. Recherche des homologues de *AtFANCM* et *AtRECQA/B* chez le riz

La relation d'homologie entre *A. thaliana* et *Oryza sativa* pour le gène *FANCM* est simple. Le gène Os11g07870 a été identifié par BBMH comme l'unique orthologue de AT1G35530 (*FANCM*) chez *Oryza sativa* ssp japonica. Même si la séquence protéique est peu conservée entre les 2 organismes les deux domaines hélicase ATP-binding et hélicase C-ter sont conservés (Figure 33). Comme nous l'avons vu en introduction, le gène *RECQ* humain (BLM) (§ 3.2.6.3) a deux orthologues chez les *brassicaceae* et un seul chez le riz. L'orthologue de *AtRECQ4A* (AT1G10930) et *AtRECQ4B* (AT1G60930) est *OsRECQI4* (Os04g35420). (Kwon et al. 2013; Séguéla-Arnaud et al. 2015). *AtRECQ-A*, *AtRECQ-B* et *OsRECQI4* ont en commun un hélicase ATP-binding, un domaine hélicase C-ter et un domaine HRDC (*Homologous region RNase D C-terminal*) (Figure 33).

5.3.2. Identification de mutant perte de fonction

Comme expliqué en introduction, pour étudier l'effet d'un gène anti-CO sur la recombinaison il nous faut disposer de 2 mutants perte de fonction (KO) dans ces gènes. Ces deux mutations doivent de plus être présentes dans deux fonds génétiques différents afin de disposer après leur croisement de polymorphismes permettant l'analyse de la recombinaison dans les descendance F2. Le riz dispose de 6 grandes collections internationales de mutants d'insertion (ADN-T ou *Tos17*) qui ont été créées dans des fonds génétiques différents mais appartenant tous au groupe des Japonica tempérés (principalement Dongjin (DJ), Nipponbare (NB), Hwayoung (HW), Tainung 67 (T67) et Zhonghua 11 (Z11) (§ 1.5.1.2). Le niveau d'expression des gènes d'intérêts (GI) dans les lignées obtenues par insertion d'un ADN-T a été mesuré par RT-PCR quantitative et/ou semi-quantitative pour vérifier que l'ARNm du GC n'était plus transcrit.

Cette démarche a été suivie pour caractériser des mutants chez les gènes *FANCM* et *RECQI4* précédemment identifiés par homologie réciproque de séquence avec *A. thaliana*.

5.3.2.1. Détermination de KO dans le gène Os11g07870 (*OsFANCM*)

Les collections de mutants internationales proposent 6 lignées d'insertions ADN-T ou *Tos17* affectant la séquence du gène *OsFANCM*, (Annexe 1). Après génotypage par PCR, 3 lignées présentent une ségrégation normale de l'insertion au site prédit avec 25% des descendants homozygotes mutants (-/-), 25% homozygotes sauvage (+/+) et 50% hétérozygotes (+/-) pour lesquels l'ADN-T n'est présent que sur un des deux allèles et l'autre allèle est sauvage. Les lignées retenues sont AQSG07 (ADN-T, NB, *Osfancm-1*), A46543 (ADN-T, DJ, *Osfancm-2*) et 03Z11CJ39 (ADN-T, Z11, *Osfancm-3*) (Annexe 1). Le séquençage des produits PCR a permis

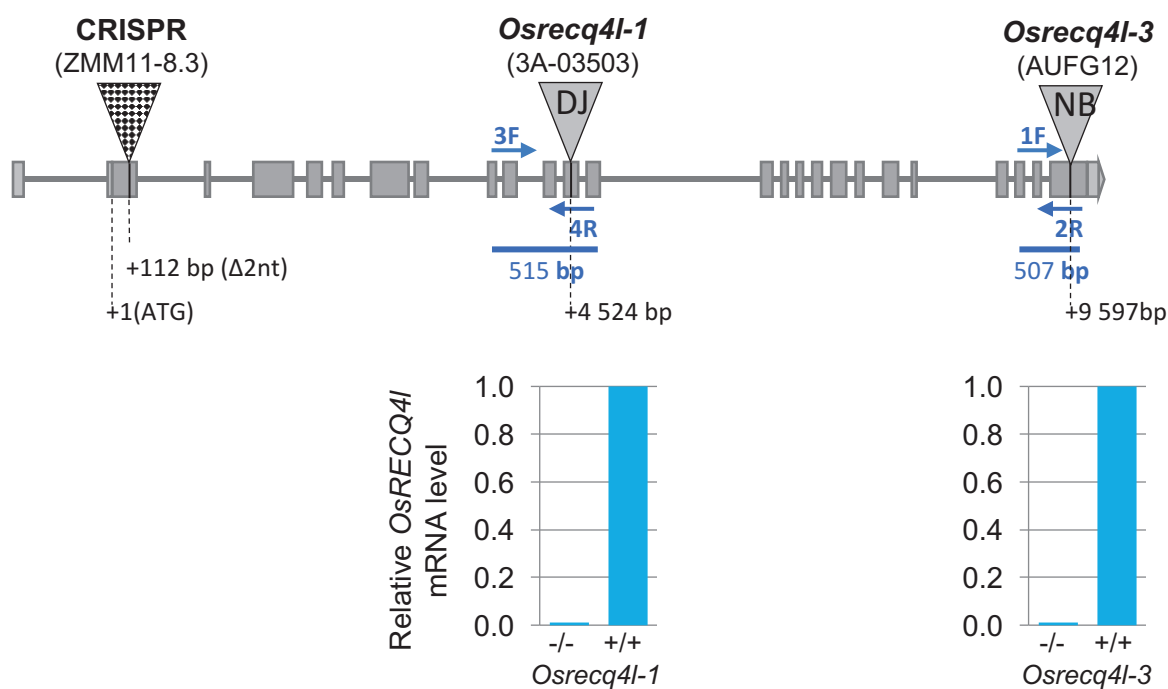
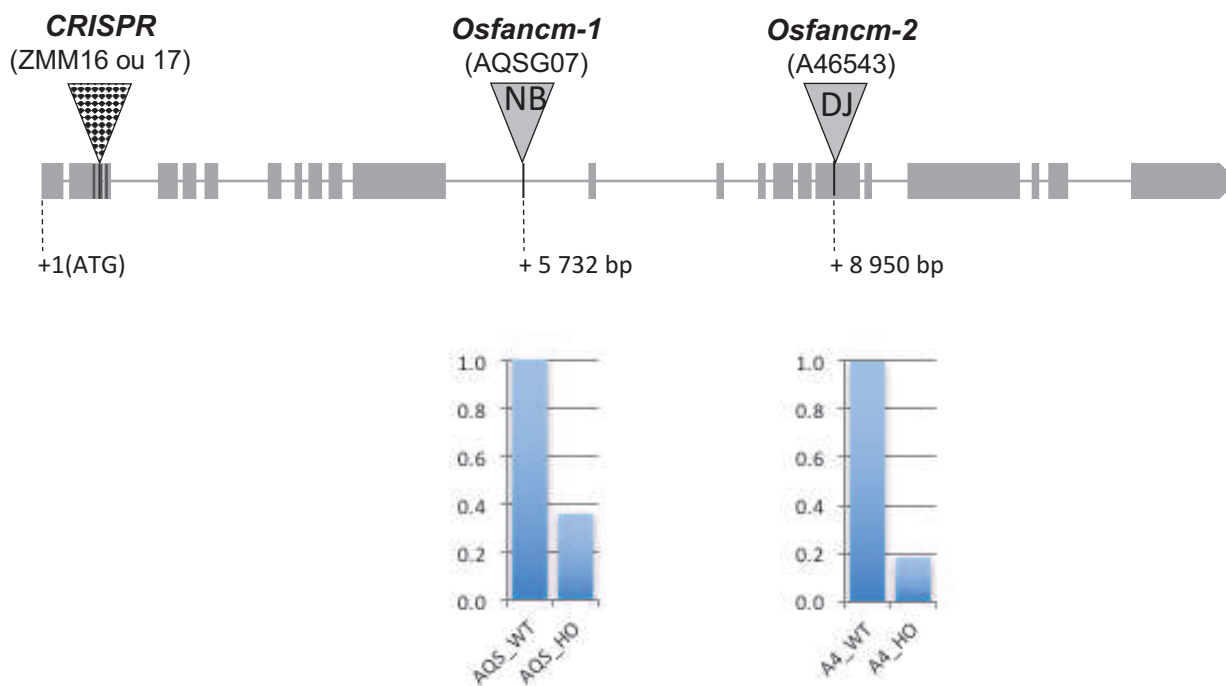


Figure 34 : Position des mutations dans les gènes.

Les triangles représentent la position des insertions ADN-T (gris) ou le site d'édition CRISPR/cas9 (damier noir et blanc). En bleu la position des amorces utilisées pour la PCR semi-quantitative. Les histogrammes montrent les résultats de l'étude d'expression des transcrits soit par Q-PCR pour les allèles *Osfanm-1* et *Osfanm-2* (amorces dans le 3'UTR) soit par PCR semi-quantitative pour les allèles *Osrecq4I-1* et *Osrecq4I-3*.

de valider que l'ADN-T est inséré dans le 12^{ème} intron pour l'allèle *Osfancm-1* et dans qu'il est dans l'exon 19 pour l'allèle *Osfancm-2* (Figure 34).

Le niveau d'expression du gène *OsFANCM* dans les inflorescences est faible et ne permet pas d'avoir des résultats précis par Q-RT-PCR (Nombre de cycles trop élevé). Cependant nous avons pu voir que le niveau de transcription de *FANCM* dans la lignée 03Z11CJ39 est 7 fois plus important que chez le sauvage. Nous avons donc supprimé cet allèle de l'étude. L'accumulation des transcrits du gène *FANCM* est fortement réduite chez les lignées A4 -/- et AQS -/-. Toutefois il est nécessaire de compléter cette analyse (Figure 34). Des essais de PCR semi-quantitative utilisant des amorces de part et d'autre du transgène et cette fois des cDNA obtenus à partir d'ARNm de jeunes plantules de 10 jours ont été réalisés et des oligos fonctionnels ont été définis pour étudier le mutant *Osfancm-2* mais, pour le mutant *Osfancm-1*, plusieurs produits d'amplification ont été observés quelles que soient les températures d'hybridation utilisées (Annexe 2).

5.3.2.2. Détermination de KO dans le gène Os04g35420 (*OsRECQ4*)

Deux allèles, *Osrecq4-1* (3A-03503, ADN-T, DJ) et *Osrecq4-2* (AG023179, *Tos17*, NB) avaient été caractérisés comme étant des mutants perte de fonction (Kwon et al. 2013). Nous avons choisi de reprendre l'allèle *Osrecq4-1* (3503) car la mutation est dans un fond génétique DJ. D'autre part dans la collection OTL nous disposons potentiellement de 4 lignées au locus *OsRECQ4* dans un fond génétique NB (Annexe 1). Le génotypage des F2 n'a permis de retrouver l'insertion de l'ADN-T que dans la lignée AUF12 (38%HO, 54%HE, 8%AZ). Le séquençage des produits PCR a permis de confirmer la présence de l'insertion ADN-T dans le 11^{ème} exon pour l'allèle *Osrecq4-1* et dans le dernier exon pour l'allèle *Osrecq4-3* (Figure 34). Pour l'allèle *Osrecq4-1*, Kwon et al (2013) avaient montré par Northern que les plantes *Osrecq4-1* -/- n'exprimaient plus le transcript dans les pointes racinaires. L'analyse par RT-PCR semi-quantitative des transcrits provenant de jeunes plantes de 10 jours montre que la lignée 3503 -/- n'exprime pas le gène *OsRECQ4* confirmant que l'allèle *Osrecq4-1* est KO (Figure 34). De même, pour la lignée AUF12, aucun transcript à la taille attendue n'est observé. Ces analyses nous permettent de conclure que la lignée AUF-/- est KO. Ce nouvel allèle est appelé *Osrecq4-3*.

5.3.3. Cumul des mutations par croisement

5.3.3.1. Obtention de population F2 simple mutant *RECQ4*

L'analyse des transcrits du gène *OsRECQ4* dans les 2 allèles *Osrecq4-1* -/- (DJ) et *Osrecq4-3* -/- (NB) montre que le gène n'est plus exprimé. Ces deux allèles ont été croisés pour produire l'hybride F1 *Osrecq4* -/- et le sauvage dans un fond génétique hybride NB/DJ. Le niveau d'expression du gène *OsRECQ4* attendu dans l'hybride F1 *Osrecq4* -/- est donc nul. Les plantes *Osrecq4-1* +/- ont été utilisées comme femelles et croisées avec *Osrecq4-3* +/ (Figure

29 et 35). Sur 306 épillets 37% ont donné des grains F1 (114). Le génotypage de 161 F1 par PCR a permis de sélectionner 14 plantes *Osrecql4* --/-- et 31 *OsRECQ4* ++/++. L'autofécondation des hybrides F1 simple mutant homozygote et du sauvage a permis de sélectionner un lot de semences qui constitue la population simple mutant *Osrecql4* --/-- et un lot de semences *OsRECQ4* ++/++. Les ADN de près de 200 plantes F2 (dérivées d'une seule F1) par population constituent le matériel de base pour l'étude génétique de la distribution des COs dans la population mutante *Osrecql4* --/-- et la population témoin dérivant du même fond génétique hybride DJ/NB.

5.3.3.2. Obtention de populations F2 simples mutants *FANCM*

Bien qu'il reste encore à confirmer que les allèles *Osfancm-1* -/- et *Osfancm-2* -/- n'expriment plus *OsFANCM* nous avons quand même procédé au croisement entre les deux mutants car c'est une étape qui prend du temps. Comme décrit dans le schéma de croisement (Figure 29, nous avons croisé *Osfancm-2* +/- (A4; DJ) avec *Osfancm-1* +/- (AQS; NB) (Figure 35). Les hybrides F1 produits (plus de 500) ont été génotypés par PCR (4 points de données par F1) pour sélectionner les plantes *Osfancm* --/-- et les plantes *OsFANCM* ++/++. L'autofécondation de ces F1 permet d'obtenir deux types de descendance F2 qui constituent les populations simples mutant *Osfancm* --/-- et sauvage *OsFANCM* ++/++. Comme pour *RECQ4*, les ADN de 200 plantes F2 par population ont été extraits. Ces 2 lots d'ADNs *Osfancm* --/-- et *OsFANCM* ++/++ servent de matériel génétique pour la construction des cartes génétiques et le calcul de la fréquence de recombinaison.

5.3.3.3. Obtention de populations F2 double mutants *FANCMxRECQ4*

Le schéma de croisement, décrit dans la partie Matériel et méthodes, montre comment en 4 étapes et en un minimum de 18 mois, on peut obtenir du matériel F2 pour étudier l'effet de la mutation de 2 gènes anti-CO sur la recombinaison (Figure 30). Nous avons suivi ce schéma pour produire le double mutant *FANCMxRECQ4* (*FANxREC* pour simplifier). Le détail des lignées croisées à chaque étape est consigné dans la figure 35. Ainsi pour la première étape du schéma, nous avons croisé *Osfancm-1* -/- (AQS ; NB) avec *Osrecql4-3* -/- (AUF ; NB) et avons obtenu 30 grains F1 nommés NB_*FANxREC* (182 épillets castrés soit 14% de réussite). Le croisement a aussi été fait dans l'autre sens et nous avons produit 23 F1 NB_*RECxFAN* (23% de réussite). Le génotypage par PCR des F1 nous a permis de sélectionner les doubles hétérozygotes (GC1 S1+/- ; GC2 S1+/-) et les sauvages dans un fond génétique NB. Pour réaliser la seconde étape nous avons croisé *Osfancm-2* -/- (A4 ; DJ) avec *Osrecql4-1* -/- (3503 ; DJ). La lignée A4 étant peu fertile et naine, nous l'avons utilisée comme femelle exclusivement et avons produit 24 grains F1, appelés DJ_*FANxREC*. Le génotypage par PCR de ces plantes permet de sélectionner les doubles hétérozygotes (GC1 S2+/- ; GC2 S2+/-) et les sauvages respectifs dans un fond génétique DJ. Enfin pour réaliser la troisième étape, nous avons croisé entre elles les plantes F1 issues des étapes 1 et 2. Sur 36 inflorescences castrées et

Gene	<i>A. thaliana</i>	<i>Oryza sativa</i>	Plant Name	FST	Position	Mutagen	Genetic BG	Mutant Id.	Allele name
FANCM	AT1G35530	Os11g07870	AQSG07	SAT7G04	intron	T-DNA	Nipponbare	AQS	<i>Osfancm-1</i>
			A46543	5A-00166R	exon	T-DNA	Dongjin	A4	<i>Osfancm-2</i>
RECQ14	AT1G10930	Os04g35420	AUFG12	GS897606	exon	T-DNA	Nipponbare	AUF	<i>Osrecql4-3</i>
			A12797	3A-03503L	exon	T-DNA	Dongjin	3503	<i>Osrecql4-1</i>

Type de Croisement	Code croisement	Femelle		Mâle		Total grains F1	Produit
		Lignée	Fond génétique	Lignée	Fond génétique		
Simple	FAN	A4	DJ	AQS	NB	> 300	Hybryde F1 ^{DJ/NB}
	REC	3503	DJ	AUF	NB	162	Hybryde F1 ^{DJ/NB}
Double Etape 1	DJ_FANxREC	A4	DJ	3503	DJ	24	Hybryde F1 ^{DJ}
Double Etape 2	NB_FANxREC	AQS	NB	AUF	NB	30	Hybryde F1 ^{NB}
	NB_RECxFAN	AUF	NB	AQS	NB	23	Hybryde F1 ^{NB}
Double Etape 3	SF1_FANxREC	DJ_FANxREC	DJ	NB_FANxREC	NB	> 500	Super Hybryde F1 ^{DJ/NB}
	SF1_FANxREC	DJ_FANxREC	DJ	NB_RECxFAN	NB	> 500	

Figure 35 : Lignées utilisées dans les croisements en vue de l'obtention des simples et doubles mutants.

Le tableau du haut présente le détail des allèles sélectionnées pour réaliser l'étude des gènes *FANCM* et *RECQ14*.

Le tableau du bas illustre les croisements qui ont été faits pour obtenir soit des simples mutants *FANCM* ou *RECQ14*, soit le double mutant *FANCMxRECQ14* dans un contexte hybrides F1 ^{DJ/NB}.

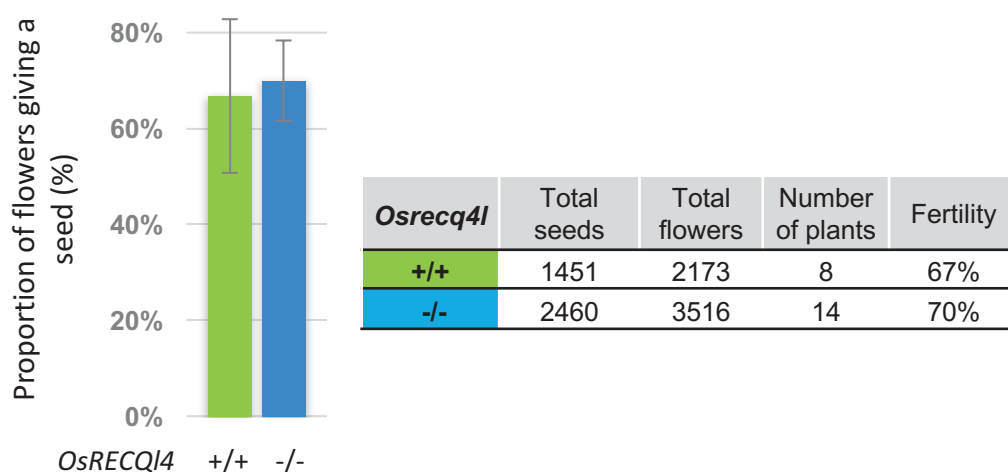


Figure 36 : Fertilité paniculaire des hybrides F1 *Osrecq14* -/- et des sauvages.

pollinisées, nous avons récolté plus 1000 grains super F1. Le taux de réussite sur ce croisement est de plus de 65% (nombre de grains fécondés / nombre de fleurs castrées). On s'attend à ce qu'1/16 de ces plantes super hybrides F1 soient homozygotes pour les deux gènes. Ces super F1 pourront être semés pour une récolte des grains F2 prévu 4 mois plus tard. Ce matériel permettra de réaliser l'étude génome-wide de la population F2 double FANxREC ----/----.

5.3.4. Caractérisation phénotypique des hybrides F1

- ✓ Fertilité paniculaire des Hybrides F1 *Osrecq14* --/--

Le pourcentage moyen de remplissage des épillets des hybrides F1 *Osrecq14* --/-- est de 70% (n=14) tandis que la fertilité paniculaire du témoin *OsRECQ14* ++/++ est de 67% (n=8) (Figure 36). La mutation du gène *Osrecq14* n'affecte pas la fertilité de la plante, la différence mesurée entre les deux populations n'est pas significative.

- ✓ Observation de méiocytes mâles chez *Osrecq14*--/--

L'étalement chromosomique de méiocytes mâles permet de voir que le déroulement de méiose est similaire chez l'homozygote mutant *Osrecq14*--/-- et sauvage (Figure 37). Au stade diplotène les 12 paires de chromosomes homologues sont connectées par le chiasma (n>15). Puis en métaphase I de première division, les chromosomes homologues s'alignent sur la plaque métaphasique (n>30 pour le mutant et le sauvage). En anaphase I la migration de 2 lots de chromosomes homologues vers les pôles opposés de la cellule est équilibrée. Puis la seconde division se déroule elle aussi tout à fait normalement de façon identique au sauvage (n>30).

5.3.5. Cartographie des événements de recombinaison

5.3.5.1. Dans une population F2 "simple *RECQ14*"

Il s'agit ici de comparer la taille d'une carte génétique obtenue dans une population F2 "simple mutant *RECQ14*" de 153 descendant *Osrecq14* --/-- avec celle obtenue à partir d'une population sauvage *Osrecq14* ++/++ de même taille. Cette étude a été faite en utilisant 241 marqueurs distribués sur tout le génome du riz ce qui représente 73 746 points de données de génotypage (12 puces 96.96 GT Fluidigm) (Figures 31 et 32). Quelques distorsions de ségrégation sont observées dans la population sauvage. Le test de χ^2 montre que, sur 241 marqueurs, 26 sont distordus dont 4 assez fortement ($P < 0.005$). Ces marqueurs sont regroupés à l'extrémité du bras long du chr. 7 et ont un déficit d'allèle DJ. Quand on regarde la population F2 *Osrecq14* --/--, trois marqueurs sont assez fortement distordus ($P < 0.005$). Il s'agit de 2 des 4 marqueurs qui sont sur le chr.7. Le dernier marqueur est situé sur le chr. 4. Il possède beaucoup de

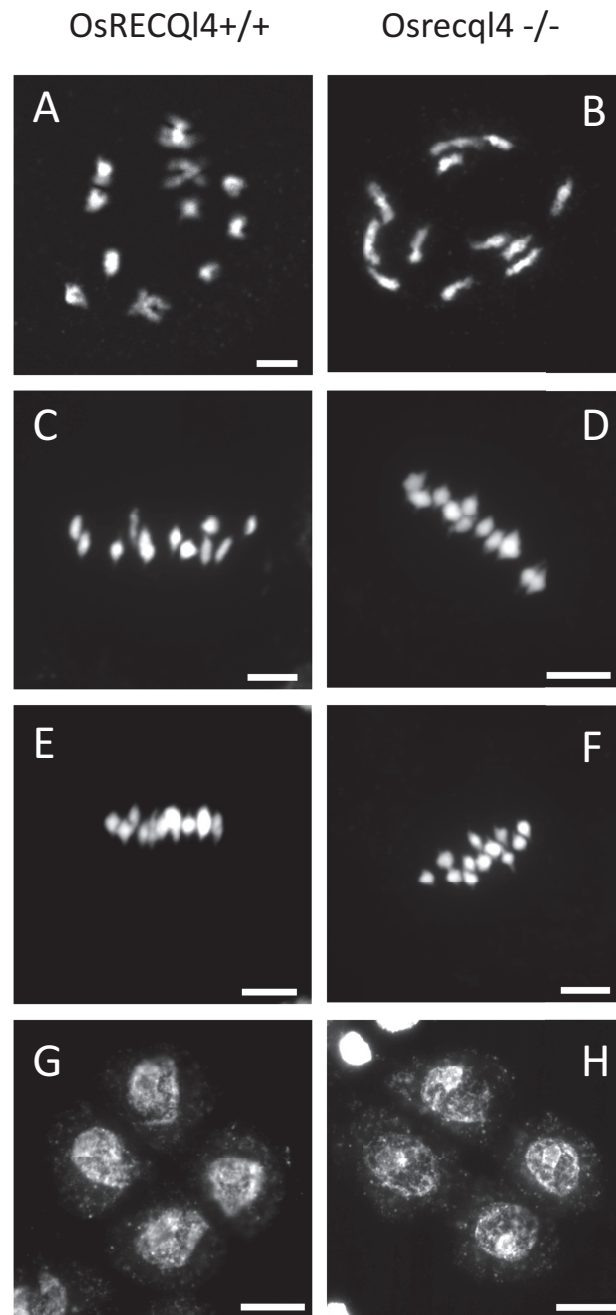
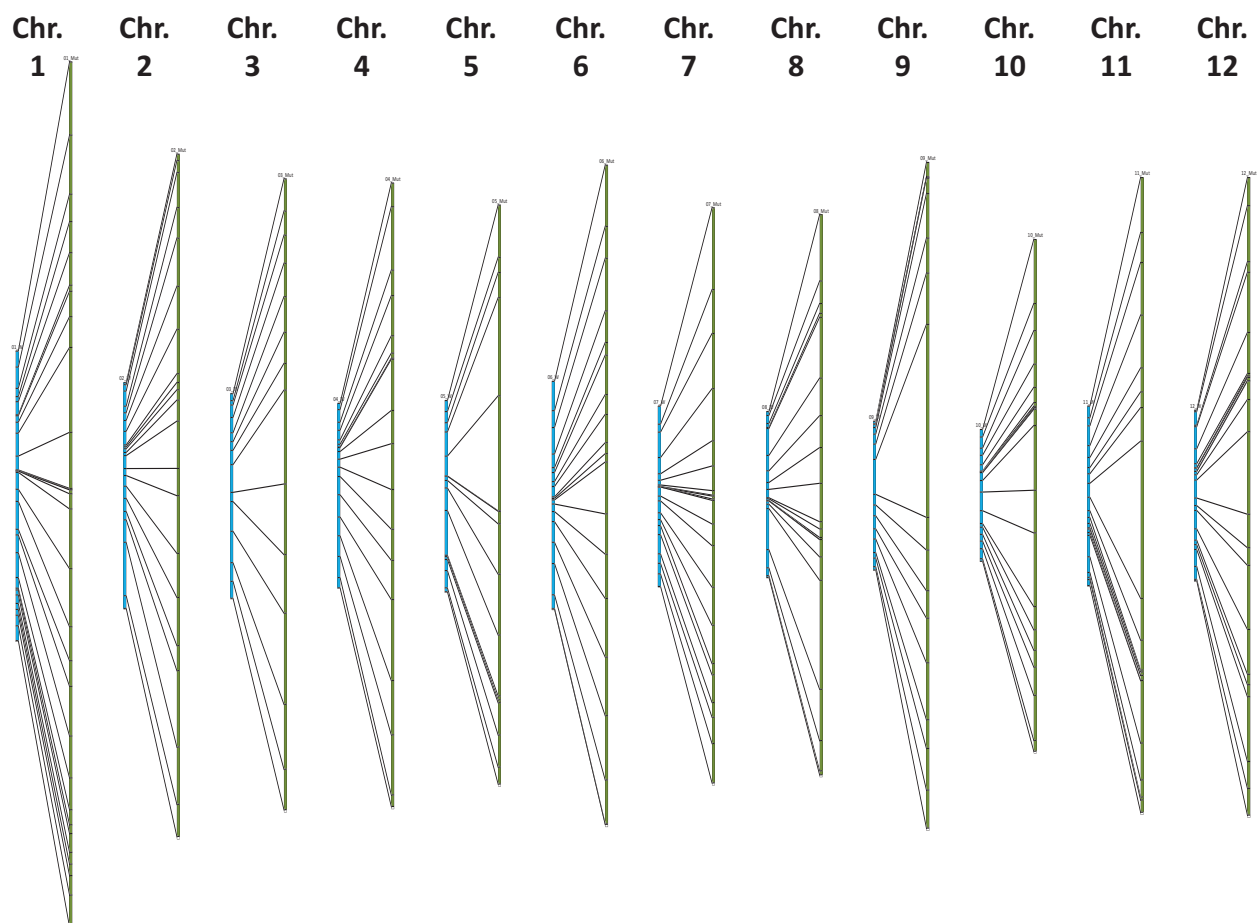


Figure 37 : Meiose mâle chez *Osrecq14* -/- (DAPI).

(A-B) Diplotene : the 12 pairs of chromosome are connected by chiasma. (C-D) Metaphase I with 12 aligned bivalents. (E-F) Metaphase II with 12 pairs of chromatids. (G-H) Tetrad.
Scale bar = 5 μ m.



CHR	<i>OsRECQ14</i> +/+	<i>Osrecq14</i> -/-	Facteur d'augmentation
Chr. 1	210.37	628.23	2.99
Chr. 2	163.89	495.58	3.02
Chr. 3	148.56	458.13	3.08
Chr. 4	134.02	452.84	3.38
Chr. 5	138.86	420.49	3.03
Chr. 6	165.4	478.74	2.89
Chr. 7	130.98	417.66	3.19
Chr. 8	120.4	406.79	3.38
Chr. 9	108.02	483.49	4.48
Chr. 10	95.67	371.62	3.88
Chr. 11	130.35	460.97	3.54
Chr.12	123.84	463.06	3.74
TOTAL	1670.36	5537.6	3.32

Figure 38 : Cartes génétiques chez le sauvage et le mutant *Osrecq14* --/--.

En bleu carte génétique établie chez le sauvage *OsRECQ14*+/+/+

En vert, carte génétique établie chez le mutant *Osrecq14* --/--

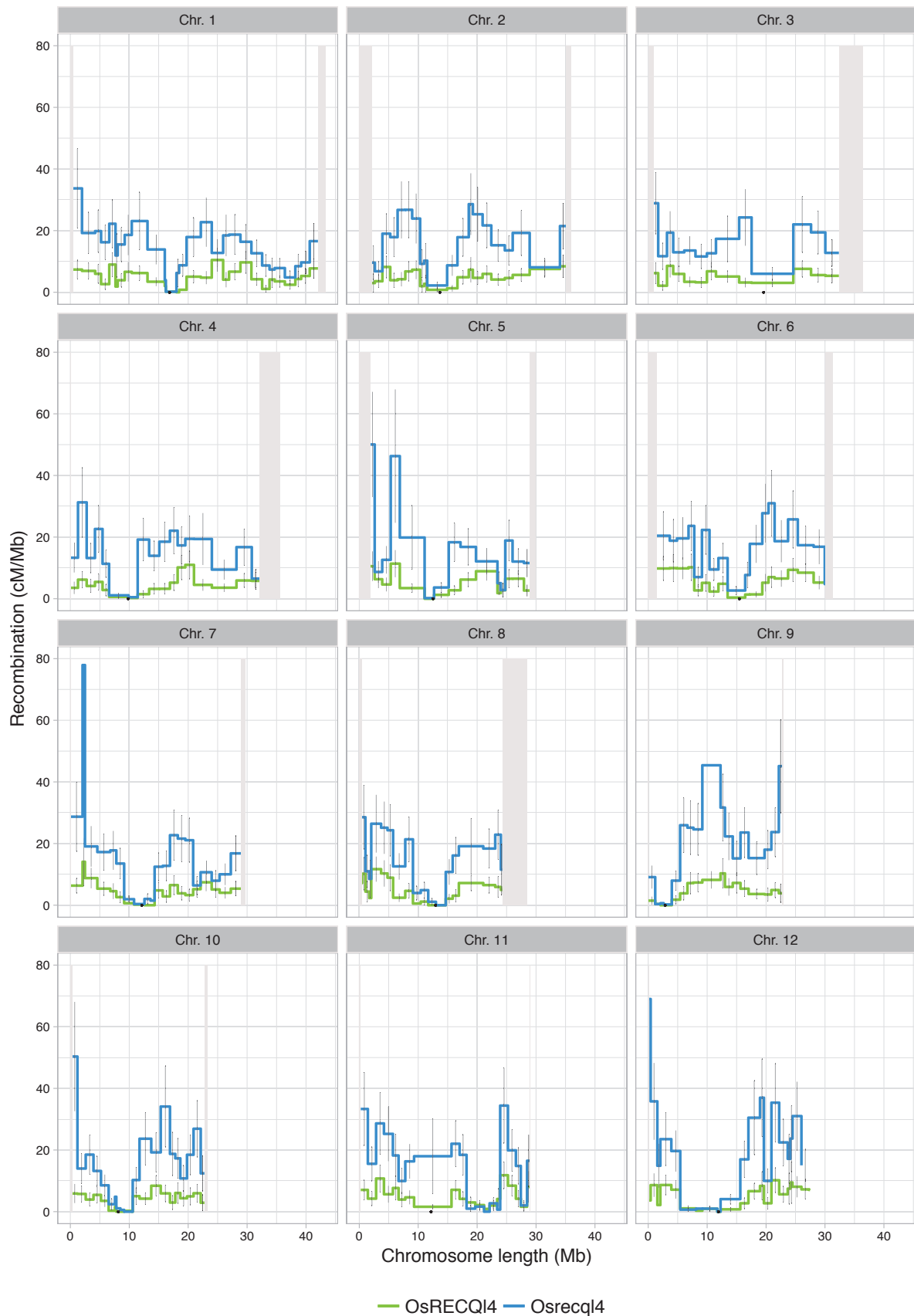


Figure 39 : Distribution des COs sur les 12 chromosomes du riz chez les plantes F2 de *Osrecq14* --/-- et *OsRECQ14* ++/++.

241 SNPs were genotyped using Kasp® technology on Fluidigm Biomark system using 153 plants per F2 population. The recombination ratio (cM/Mb, calculated using Haldane fonction) in each interval was plotted along the 12 rice chromosomes. Black dots represent the centromere position ; arrows represent the telomere positions; The gray area represents area without marker.

données manquantes ce qui peut expliquer que la ségrégation observée soit différente de celle attendue.

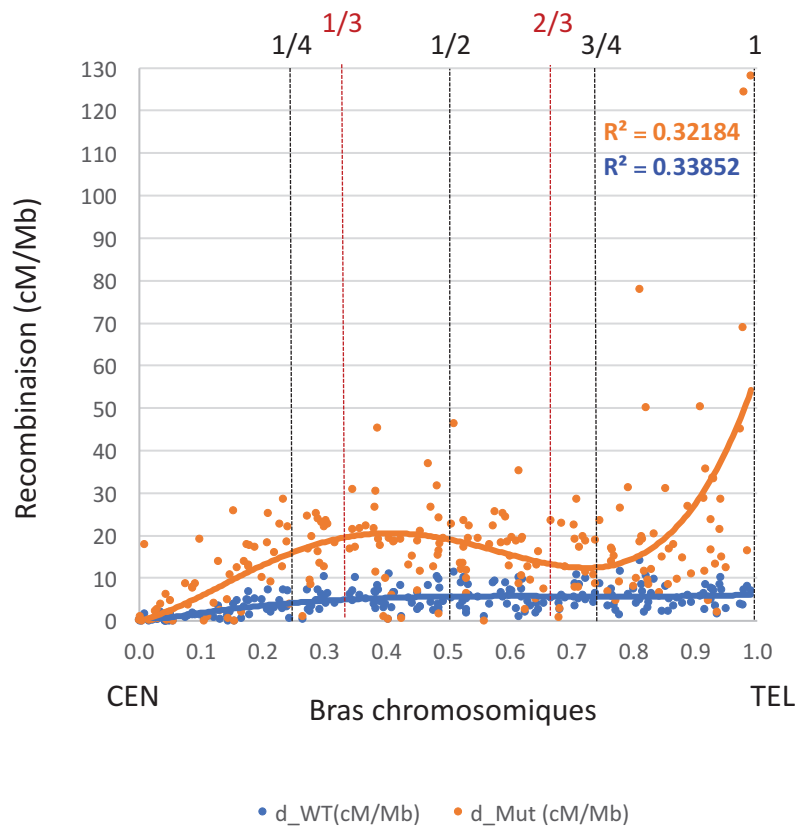
Les groupes de liaison ont été établis à partir des données de la population sauvage. On retrouve bien les 12 groupes de liaisons et l'ordre des marqueurs est bien celui attendu en fonction de la position physique qu'on leur connaît. En revanche, si on essaie de faire la carte de liaison avec les données de génotypage de la population F2 *Osrecql4* --/-- alors, même avec un LOD peu strict de 3 (r max : 0,3), on obtient 28 groupes de liaison et 7 marqueurs sont non liés. Sachant que nous nous attendions à avoir une carte génétique plus grande dans la population mutante, nous avons décidé d'utiliser les groupes de liaison obtenus lors de l'analyse de la population sauvage contrôle pour faire la carte génétique du mutant *Osrecql4* --/--.

En appliquant ces paramètres, et en utilisant l'équation de Haldane pour calculer les distances entre marqueurs, on obtient une carte génétique de 1670 ± 30 cM pour le sauvage et 5538 ± 114 cM pour le mutant ce qui correspond à une augmentation de la fréquence des COs de 3.3 fois lorsque le gène *Osrecql4* est muté. Selon les chromosomes, la taille des cartes peut être agrandie de 2,9 fois sur le chr. 6 à 4,4 fois sur le chr. 9 (Figure 38 et Annexes 4).

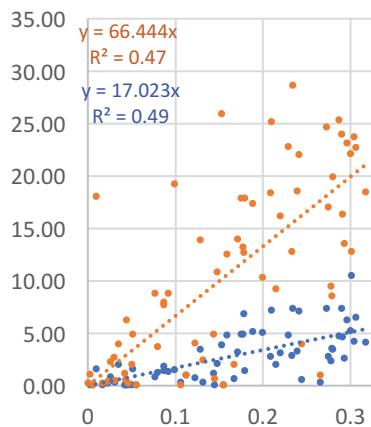
Une représentation graphique du nombre de CO par intervalle et par centiMorgan permet de voir comment sont distribués les COs le long des chromosomes et de voir comment le nombre de recombinants augmente dans la population mutante en fonction de la position de l'intervalle considéré (Figure 39). Toute fois lorsque l'intervalle entre deux marqueurs est très réduit, la recombinaison exprimée ici en cM/Mb (d/taille de l'intervalle) forme un pic artificiel qui ne serait pas observé dans un intervalle plus grand. La solution retenue est de supprimer un marqueur pour lisser le nombre de recombinants sur un plus grand intervalle.

Globalement, on observe que les régions proches du centromère sont dépourvues de CO (Figure 39) et que la recombinaison est augmentée le long de l'ensemble des bras chromosomiques chez le mutant *Osrecql4* --/-- par rapport au sauvage. Lorsqu'on projette tous les bras chromosomiques sur une même figure (distance relative en pourcentage par rapport au centromère) et qu'on regarde la distribution de la recombinaison (Figure 40), on s'aperçoit que chez le sauvage, le taux de recombinaison, part de zéro au centromère puis augmente de façon linéaire pour atteindre 4 cM/Mb ($r^2 = 0,49$) au premier tiers du chromosome où elle atteint un plateau. Concernant le mutant *Osrecql4* --/--, la recombinaison est à zéro au centromère puis augmente de façon linéaire sur le premier tiers du bras chromosomique jusqu'à 20 cM/Mb ($r^2 = 0,47$) puis diminue sur le second tiers jusqu'à 13 cM/Mb et remonte sur le dernier tiers où elle atteint en moyenne 52 cM/Mb (Figure 40).

A



B



C

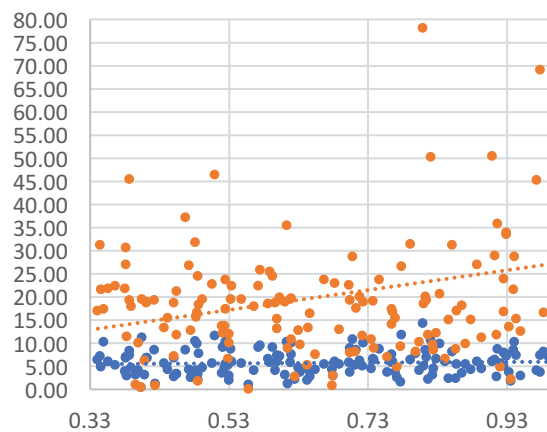


Figure 40 : Observation de la recombinaison sur différents intervalles en fonction de la position physique de ces intervalles sur le chromosome.

A- Rapport entre la distance génétique et physique (cM/Mb) établi chez le sauvage (bleu) et chez le mutant (orange) en fonction de la position de l'intervalle où ont eu lieu les COs sur le bras chromosomique. Les données de tous les bras chromosomiques (sauf les deux bras du chr. 9) sont reportées selon leur distance relative au centromère (la courbe indique la régression polynomiale de degré 4).

B- Zoom sur le premier tiers des bras chromosomiques en partant du centromère (CEN).

C- Zoom les derniers tiers du bras chromosomique coté télomère (TEL).

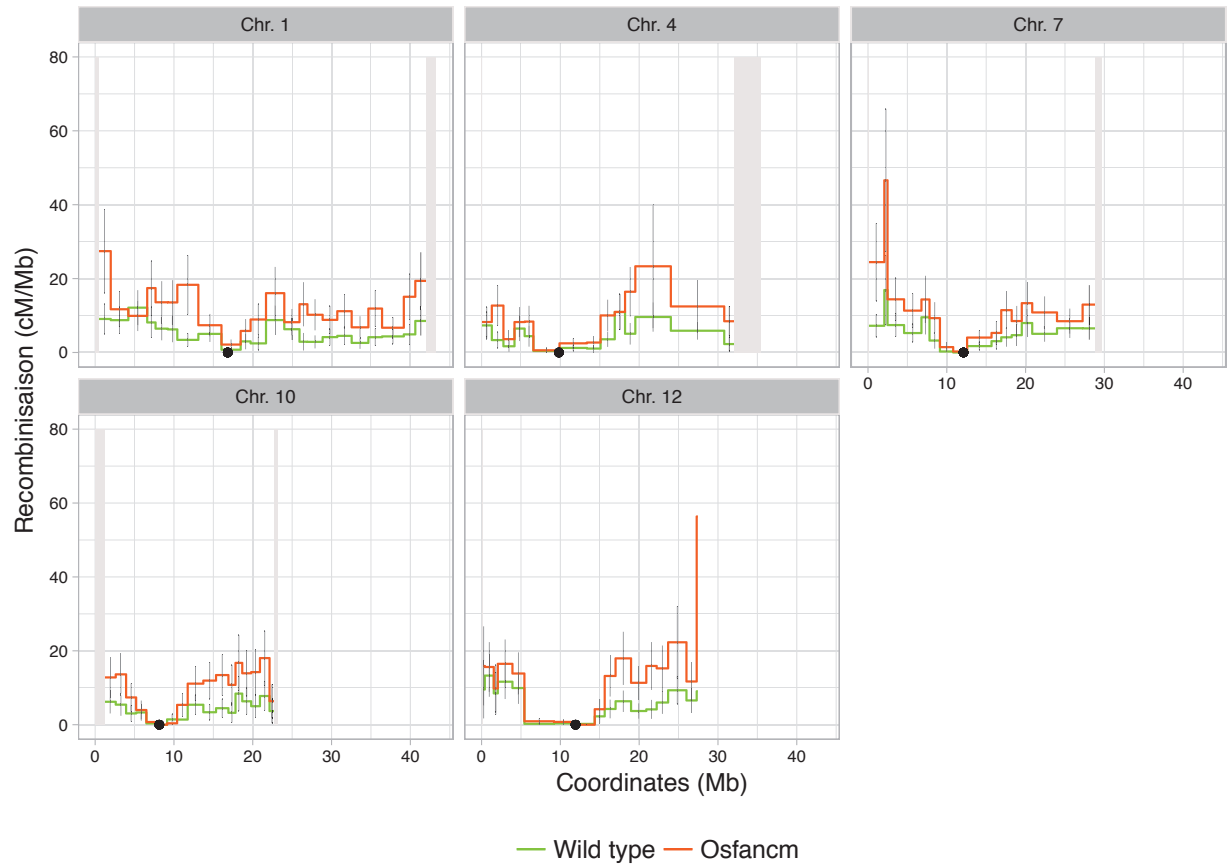


Figure 41 : Distribution des COs sur 5 chromosomes de plantes F2 de *Osfanm* -/- .

92 SNPs were genotyped using Kasp® technology on Fluidigm Biomark system using 108 (Mut.) or 112 (WT) plants per F2 population. The recombination ratio (cM/Mb, calculated using Haldane fonction) in each interval was plotted along the 5 rice chromosomes. *Osfanm* -/- leads to a global 2.15 folds increase in CO frequency (genetic map size: 739 to 1589 cM)

Black dots represent centromere position ; arrows represent telomere positions ; the gray area represents area without marker.

5.3.5.2. Dans une population F2 "simple *FANCM*"

Sachant que chez *A. thaliana*, la mutation de *Atfancm* affecte moins la recombinaison que *Atrecq4A/B* (Girard et al. 2015), nous avons choisi d'étudier l'impact de la mutation de *Osfancm* sur un panel de 5 chromosomes. 96 marqueurs ont été choisis parmi les 557 SNP de notre base de données pour couvrir les 5 chromosomes et avoir une résolution correcte d'un SNP tous les 1,7 Mb (Annexe 6). L'analyse de cette population de 108 individus *Osfancm* --/-- et 112 sauvages avec 92 marqueurs permet de former 5 groupes de liaison comme attendu et les marqueurs sont ordonnés dans l'ordre de leur position physique. Il y a peu de distorsion de ségrégation mais les 3 marqueurs DJ13, 14 et 15 qui sont distordus ($P < 0.001$) sont groupés au milieu du bras long du chromosome 1 et présentent un excès d'allèle DJ par rapport à ce qui est attendu (Annexe 6). Pour la population sauvage, ces trois mêmes marqueurs sont ceux qui ont les χ^2 les plus élevés ($P < 0.005$). En appliquant les mêmes paramètres que pour l'analyse du "simple *RECQ4*", on obtient une carte génétique de 739 ± 25 cM pour le sauvage et 1587 ± 61 cM pour le mutant. La taille de la carte agrandie d'un facteur 2,15 lorsque le gène *Osfancm* est muté. Les distributions de la recombinaison sont parallèles chez le sauvage et chez le mutant (Figure 41).

5.4. Discussion et perspectives

Le mutant *Osrecq14* --/-- présente une progression de méiose et une fertilité normales.

L'analyse cytogénétique des méioses mâles montre que les étapes de la méiose se déroulent normalement chez le mutant comme chez le sauvage. Cependant, au stade diacinèse alors que l'on peut observer des homologues en forme de croix chez le sauvage, on observe plus souvent chez le mutant *Osrecq14* --/-- des barres épaisses comme si les homologues étaient "collés" sur quasi toute la longueur. Cette observation, qui laisse penser que les chromosomes homologues partagent plus de connexions physiques (chiasma), rappelle celle notée chez le double mutant *Atrecq4a/b figl1* d'*A. thaliana* (Fernandes et al. 2017b). En métaphase I, les chromosomes d'*A. thaliana* chez le double mutant *Atrecq4a/b figl1* présentent également une forme différente du sauvage (Fernandes et al. 2017b). Chez le riz, les chromosomes sont deux fois plus petit que chez *A. thaliana* (2,3 μM en moyenne contre 4,0 μM), il faudrait faire de la microscopie électronique pour pouvoir compter les nodules précoces de recombinaison et confirmer cette hypothèse.

La fertilité de la plante *Osrecq14* --/-- ne semble pas affectée par rapport au sauvage. Cependant, le gène *OsRECQ14* a également une fonction démontrée dans les cellules somatiques du riz (Kwon et al. 2013). Le mutant *Osrecq14* présente une sensibilité accrue aux agents mutagènes, une tendance à l'augmentation de la formation de CDBs et Kwon observe des événements de mort cellulaire au niveau des cellules méristématiques racinaires (Kwon et al. 2013). D'autre part, il a été montré que la surexpression de *OsRECQ14* conduit à une augmentation de la fréquence de reconstruction d'un substrat de recombinaison intra-chromosomique pGU_US par SSA. Ce qui semble cohérent avec son rôle méiotique de résolution des intermédiaires de recombinaison post invasion par SDSA (Kwon et al. 2013). Ces résultats, nous conduisent à penser qu'une analyse plus complète de la fertilité pollinique, de la fertilité paniculaire et un phénotype de la plante entière, notamment en situation de stress, pourraient être intéressants.

Nous n'avons pas pour le moment des données cytologiques sur la méiose du mutant *Osfancm* --/--. Ce travail reste à réaliser, comme l'évaluation de la fertilité de la plante. Une analyse préliminaire grossière où seul le nombre total de grains par plante a été compté, semble montrer que *Osfancm* --/-- produit moins de grains que le sauvage. Cependant, la récolte est la résultante de nombreux paramètres morphologiques phénologiques et physiologiques, il faut donc ré-examiner la plante dans son ensemble pour pouvoir se prononcer sur la fertilité paniculaire du mutant *Osfancm* --/--.

La mutation du gène *Osrecql4* --/-- augmente la recombinaison en contexte hybride intra-japonica.

Dans les conditions de cartographie utilisées, la taille de la carte chez le sauvage est de 1670 cM ce qui, compte tenu de l'absence de couverture de certaines régions télomériques, (Figure 39), est tout à fait comparable à ce qui est attendu chez le riz. Plusieurs études de cartographie fine de populations de riz montrent que la taille de la carte génétique est de l'ordre de 1800 cM (Kurata et al. 1994; Chen et al. 2002; Wu et al. 2002). Abolir l'expression du gène *Osrecql4*--/-- augmente la taille de la carte génétique d'un facteur 3,3 fois par rapport au sauvage faisant passer la taille de la carte de 1670 ± 30 à 5537 ± 114 cM.

La distribution de l'augmentation de la recombinaison se fait sur l'ensemble des bras chromosomiques sauf au niveau des régions centromériques comme en témoigne la différence entre le niveau de recombinaison, exprimé en cM/Mb entre le mutant et le sauvage (Figure 39). Comme chez *A. thaliana*, les régions centromériques sont dépourvues de CO (Fernandes et al. 2017b) et l'absence de *Osrecql4* ne permet pas d'augmenter le nombre de recombinants dans ces régions. Il est connu que ces régions du génome sont riches en éléments transposables de type LTR (Long Terminal Repeat) et présentent aussi une faible densité en gènes et une hyper-méthylation de l'ADN ainsi qu'une conformation de la chromatine condensée. Tout cela limite l'accès à l'ADN de la machinerie des protéines nécessaires à la formation des COs comme SPO11. De plus, il a été montré chez la levure, que les kinétochores représentent un mécanisme actif de répression dans les régions péri-centromériques (Vincenten et al. 2015). En effet, la chromatine centromérique est constituée de séquences répétées nécessaires à l'assemblage des kinétochores, une mutation dans cette région n'est pas permise car les conséquences seraient dramatiques. Si les kinétochores ne peuvent pas s'attacher normalement, les chromatides sœurs ne se séparent pas correctement et les gamètes ne sont pas équilibrées (stérilité, mort) (Talbert and Henikoff 2010). Toujours chez la levure, on observe que le placement inadéquat de CO dans le voisinage des centromères influence négativement la ségrégation des chromosomes (séparation précoce des chromatides sœur et mortalité élevée des spores) (Rockmill et al. 2006).

Le niveau de recombinaison des régions immédiatement adjacentes aux télomères est augmenté de façon moins spectaculaire chez le mutant que dans le même contexte chez *A. thaliana* (Figure 40). Cette dépression télomérique est observée dans différentes cartes de recombinaison haute densité chez le riz (Spindle et al 2013). L'augmentation de la recombinaison en fonction de l'éloignement du centromère le long des bras chromosomiques du mutant *Osrecql4* est également beaucoup moins linéaire que chez le double mutant *Atreq4a/b*. Chez le riz, alors que cette augmentation linéaire est observée sur le premier tiers des bras, elle tend après à se stabiliser pour augmenter à nouveau. (Figure 40).

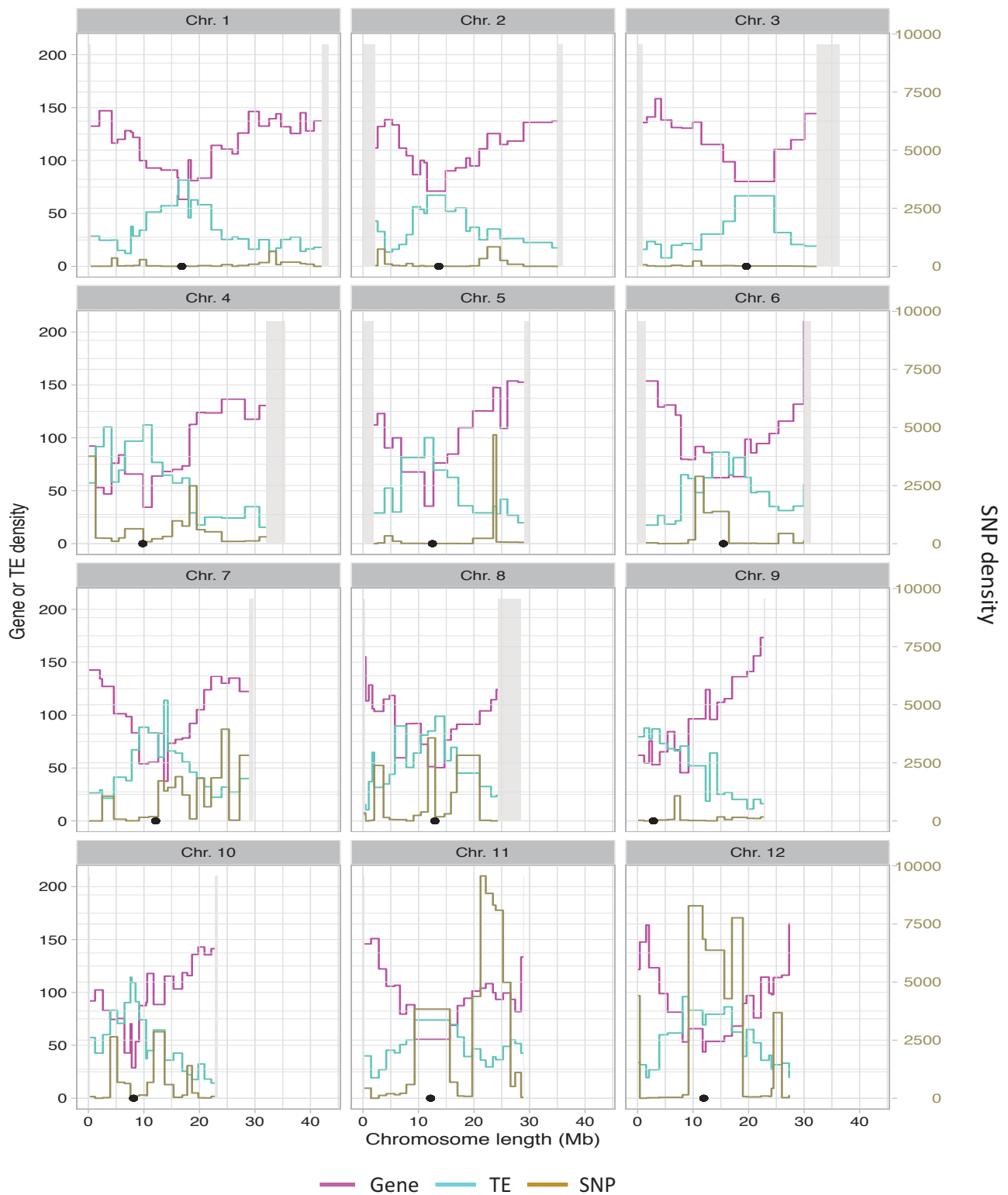


Figure 42 : Densité en gènes, en éléments transposables et en SNP sur les 12 chromosomes du riz. Distribution de la densité en gènes (rose), en éléments transposables (TE; en turquoise), et en SNP (marron) le long des 12 chromosomes (Chr.). (d'après Nipponbare MSU7)

L'effet de *recq4* est comparable à celui observé chez *A. thaliana* en contexte hybride (fois 4). La fonction du gène RECQ4 semble être conservée entre *A. thaliana* et le riz puisqu'on observe aussi une augmentation du taux de recombinaison à l'échelle du génome entier.

La mutation du gène *Osfancm*--/-- augmente la recombinaison en contexte hybride intra-japonica.

Il a été montré, chez *A. thaliana*, que *FANCM* est un anti-CO moins puissant que *RECQ* et que son effet mis en évidence en contexte "lignée" disparaît totalement en contexte hybride (Girard et al. 2015). Nous avons analysé ici des F2 produites dans un contexte hybride (entre deux japonica tempéré). Le taux de polymorphisme estimé entre ces 2 cultivars NB et DJ est de 3,6 % (estimation à partir de 20 000 points de données issus du génotypage GbS de 250 accessions représentant la diversité de l'espèce *O. sativa*). Ce taux de polymorphisme entre les espèces utilisées pour notre étude est 10 fois inférieur, à ce que l'on peut observer dans un croisement intergroupe, mais clairement différent du contexte "lignée" ou l'effet de la mutation *Atfancm* a été observé chez *Arabidopsis*.

L'analyse de cartographie de la population *Osfancm*--/-- permet de montrer que la taille de la carte génétique est augmentée qu'un facteur 2,15 par rapport au sauvage (739 ± 25 cM à 1587 ± 61 cM). La chute drastique de l'effet de la mutation du gène *FANCM* lors de l'augmentation du polymorphisme, entre chromosomes homologues observée chez *A. thaliana*, ne semble pas être observée chez le riz. L'augmentation de la recombinaison liée à la mutation de *FANCM* chez *A. thaliana* est de 3 fois en contexte "lignée" et devient non significative en contexte hybride (Fernandes et al. 2017b). Il serait très intéressant de voir quel est l'effet de *Osfancm* dans des croisements distants tels que IR64 x Azucena (*indica* x *japonica*) ou KalingaIII x Kitaake (aus x japonica) où le polymorphisme intervariétal est de 52% (ici 3,6%).

Comme dans le cas de la mutation *Osrecq4*--/--, l'augmentation de la recombinaison a lieu, sur l'ensemble des bras chromosomiques (Figure 41). La taille de la population F2 étant petite (108), le nombre de recombinants observés au niveau des centromères ne permet pas de conclure que *FANCM* a un effet centromérique. Il faudrait augmenter la taille de la population pour valider cette hypothèse.

Perspectives

Ces résultats montrent que comme chez *A. thaliana*, *FANCM* et *RECQ* ont, chez le riz, un effet anti-CO permettant d'augmenter de façon très significative le nombre de recombinants en méiose F1. L'augmentation obtenue avec *Osrecq4*--/-- est la plus importante jamais enregistrée chez le riz. La mutation partielle du gène *Oszep1*, l'orthologue du gène d'*A. thaliana* *ZIP1*, qui code pour une protéine importante de l'élément central du complexe



Figure 43 : Relation entre la densité en SNP et la fréquence de recombinaison chez *Osrecql4* --/--. Densité en gènes (rose) en éléments transposables (TE; en turquoise) et en SNP (marron).

synaptonémal, a récemment conduit à une augmentation d'un facteur 1,8 fois au détriment du retard de la progression de la méiose et d'une réduction de la fertilité (Wang et al. 2015a).

Le cumul des mutations *Osrecq4* *Osfancm* est en cours, puisque des grains F1 dont 1/16 devraient présenter des mutations homozygotes hétéroalléliques aux deux gènes ont été produites. Chez *Arabidopsis*, le cumul des deux mutations permet d'atteindre, en contexte "lignée pure", un effet cumulatif alors qu'en contexte hybride les augmentations de la recombinaison observées chez le mutant *Atrecq4a/b* et le triple mutant *Atrecq4a/b Atfancm* ne sont pas significativement différentes. Le comportement du riz semblant différent, un effet cumulatif pourrait être observé. Il sera également intéressant de tester cette double mutation en contexte hybride intergroupe (plus distant). Nous verrons que les récents développements de la technologie CRISPR/Cas9 et la transformation de semences hybrides F1 permet de tester à présent rapidement ces hypothèses (§5.5).

Reste à déterminer si l'effet de la mutation du gène *FIGL1* sur la recombinaison est également conservé entre *A. thaliana* et *O. sativa*. La mutation de ce gène et le cumul avec la mutation du gène *OsRECQ4* pourrait permettre d'atteindre un niveau maximum d'augmentation de la recombinaison à l'instar de ce qui est observé chez *A. thaliana* où une augmentation d'un facteur 7,8 fois est obtenue chez le triple mutant *Atrecq4 a/b Atfigl1*. Il sera pertinent de vérifier également si une telle augmentation du nombre de CO ne perturbe pas trop la fertilité pollinique mais surtout la fécondation et la formation des grains. En effet si une étamine contient plusieurs milliers de grains de pollens chez le riz l'ovaire de la fleur ne contient qu'un seul ovule, à la différence d'*Arabidopsis*.

PARTIE IV : Inactivation de gènes anti-COs par CRISPR/cas9

PARTIE IV : Inactivation de gènes anti-COs par CRISPR/cas9

6. INTRODUCTION

Une des limites du travail précédent a été la disponibilité d'allèles KO dans les gènes ciblés pour tenter d'augmenter la recombinaison et d'autre part la contrainte du génotype dans lesquels les ressources ont été constituées. Le développement récent des technologies d'édition du génome chez les plantes, notamment par CRISPR/Cas9 et la haute efficacité de cette dernière chez le riz pour la création de lésions dans des gènes cibles permet à présent de générer dans un panel de fonds génétiques accessibles à la transformation des allèles dans différentes régions d'un gène cible (Miao et al. 2013b). L'efficacité d'édition biallélique homozygote est également très élevée permettant l'obtention de mutant dès la génération T0. Cette propriété est très intéressante pour tester l'effet de l'inactivation d'un gène anti-CO directement en contexte hybride. L'obtention de grains T1 d'une plante KO pour le gène cible (GC) prend environ 12 mois (avec le clonage).

Une autre contrainte, liée à l'utilisation des mutants d'insertion ADN-T, concerne le temps de production du matériel génétique pour étudier l'effet cumulé des mutations chez deux gènes anti-COs. En effet, la stratégie envisagée, jusqu'ici, est très longue à mettre en œuvre puisqu'il faut plus de 18 mois rien que pour produire les populations F2 GC1xGC2 --/-- et GC1xGC2 ++/++ (Figure 30). Un autre avantage de la technologie CRISPR/Cas9, par rapport notamment aux autres technologies de mutagenèse ciblée par nucléase, est de facilement permettre le multiplexage des cibles conduisant à l'obtention simultanée, et avec une certaine fréquence à l'état homozygote, de double voir triple (Yin et al. 2017) mutants dès la génération T0 (Ma et al. 2015). Il est donc là aussi également envisageable de réaliser des KO simultané de plusieurs gènes anti-COs en contexte hybride F1 avec un construit CRISPR/Cas9 visant directement 2 ou 3 gènes cibles. Outre le gain de temps et la simplification, ceci permet de tester l'effet de la distance génétique i.e. du polymorphisme entre chromosomes homologues sur l'augmentation de la fréquence de recombinaison permise par la mutation des gènes anti-COs.

Le fait de pouvoir dériver des séries d'allèles pour un même gène dans un même fond génétique est d'autre part un avantage considérable par rapport à l'utilisation de lignées d'insertion chez le riz. Ceci permet notamment d'identifier plus facilement les effets phénotypiques liés à la mutation de ceux résultant de la variation somaclonale et d'éliminer

les artéfacts liés à la présence d'un ADN-T au site de disruption du gène (nous l'avons vu dans le cas de la lignée RMD 03Z11CJ39). Ainsi la phénocopie de mutants d'insertion ADN-T par la technologie CRISPR/Cas9 peut amener à remettre en cause certains résultats considérés comme établis (Gao et al. 2015). Au niveau de l'application, cette technologie devrait rendre plus opérationnelle l'intégration de la mutation des gènes anti-COs dans les générations de brassage génétique des schémas de sélection.

Dans cette partie nous allons donc voir 1) Comment la technologie CRISPR-Cas9 peut être utilisée pour générer de nouveaux allèles *Osfancm* et *Osrecql4* puis 2) Nous parlerons des premiers résultats obtenus avec la technique directe de transformation d'hybrides F1.

7. CREATION DE LIGNEES PERTE DE FONCTION PAR CRISPR/CAS9

7.1. Matériel et méthodes

✓ Clonage CRISPR/cas9

L'utilisation de la technologie CRISPR/Cas9 (Jinek et al. 2012) pour l'édition du génome nécessite d'exprimer dans la cellule d'une part la protéine Cas9 et d'autre part, la crRNA (séquence de 20 bases complémentaires de l'ADN génomique ciblé) et le tracrRNA qui est une séquence de 84 nt (ou 48) permettant l'adressage du crRNA sur sa cible (§1.5.4) (Figure 44).

Le fond de vecteur pCAMBIA5300 (Cambia®, Australie), très performant pour la transgénèse via *Agrobacterium tumefaciens* chez le riz, a été utilisé comme base. Une cassette contenant la protéine Cas9, optimisée pour le riz, y a été insérée au site unique de clonage KpnI. La cassette Cas9 provient du vecteur pJIT163-2NLS-OsCas9 (Dr Kunling Chen, Beijing), et a été sortie de son vecteur d'origine par double restriction KpnI/SpeI. Le vecteur ainsi obtenu a été nommé pC5300-Cas9 (Breitler J-C, non publié). La seconde cassette permet d'exprimer le sgRNA et est pour chaque clonage usinée de la façon suivante. Après avoir déterminé les 20nt constituant la séquence cible (protospacer), les amorces sont synthétisées en sens et anti-sens en ajoutant en amont de la séquence sens les 20 derniers nucléotides correspondant à la séquence du promoteur U3 (primer crRNA-F), et l'amorce anti-sens, les 20 premiers nucléotides du tracrRNA en anti-sens (amorce crRNA-R). Deux autres amorces sont nécessaires, il s'agit de l'amorce U3-F et de l'amorce tracrRNA-R (Annexe 7). Par le biais de 3 PCR (overlapping-PCR), la cassette pU3-crRNA-tracrRNA est amplifiée puis purifiée avant clonage dans le vecteur de transformation (Figure 44 ; Breitler, non publiée).



```
>EcoRI-pU3(377pb)::crRNA(20pb)::tracrRNA(88pb)-XbaI
```

GAATTC AAGGATCTTTAAACATACGaacagatcacttaaagttcttctgaagcaactaaagtatcaggcatgcatggatcttggaggaaatcagatgtgcagtcaggaccatagcacaggacaggcg
tcttctactggtgctaccagaaatgctggaagcgggaacactgggtcagttggaaaccacgtgatgtggagtaagataaacctgtaggagaaaaacatttctagtggccatgaagccttcaggacatgtattgc
agtatggcgccgcattacgcaattggacgacacaaagactagtattagtaccactcggtatccacatagatcaaaagctggtttaaaagattgtgcagatgatccgtggcNNNNNNNNNNNNNNNNNN
NNNNNGtttagagctagaatatagcaagttaaataaagctagtcggttatcaactgaaaaagtgCACCGAGTCGGTGCTTTTTTCTAGATCTAGA

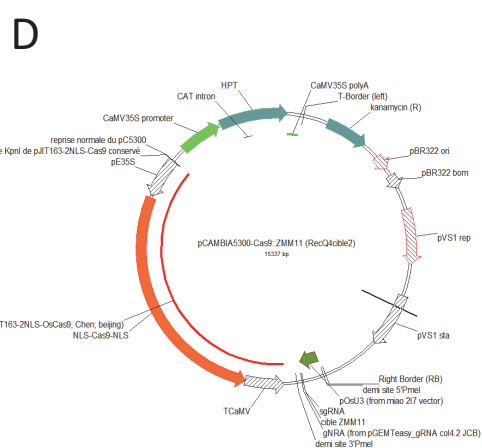


Figure 44 : Stratégie de clonage d'un sgRNA dans le vecteur d'expression pC5300-Cas9.

A- Principe de la technologie d'édition du génome par CRISPR-Cas9.

B- Amplification de la cassette pU3-crRNA-tracrRNA par overlapping PCR.

C- ADN-T du vecteur pC5300-Cas9 optimisé pour le riz.

D- Vecteur binaire pC5300cas9::RECQI4-sgRNA2 (ZMM11).

Les réactions de PCR1 A et B sont faites dans un volume final de 20 µL avec 10 ng de matrice ADN et la polymérase Phusion HF (M0530, NEB®) en suivant le protocole recommandé par NEB (98°C 5' ; 2 cycles (98°C 10"; 50°C 15"; 72°C 15") ; 30 cycles (98°C 10"; 60°C 15"; 72°C 15") ; 72°C 5') (Annexe 7). Pour la PCR2, 1µL de chaque produit PCR 1A et B sont mélangés dans 198µL d'eau puis 1µL du mélange est utilisé pour faire la PCR2 dans 50µL final avec l'enzyme Taq Phusion HF (M0530, NEB®) et les amorces externes U3-F et tracrRNA-R. Les conditions de cycles utilisés pour la PCR2 sont les mêmes que ceux utilisés en PCR 1. Le produit PCR2 est ensuite purifié sur colonne (PCR purification Kit, Promega) et 19µL de ce produit PCR sont utilisés pour transformer les extrémités 5' et 3' sortantes en extrémités cohésives avec l'enzyme T4 DNA polymérase (Quick Blunting kit, E1201S, NEB). Cette matrice ainsi obtenue constitue la cassette d'expression du sgRNA constituée du promoteur U3 du riz (OsU3), du crRNA spécifique de la cible et du tracrRNA : pU3-crRAN-tracrRNA (Figure 44).

Cette cassette d'expression du sgRNA (spécifique de chaque crRNA ciblé) est alors insérée par ligation dans le vecteur pC5300-Cas9 linéarisé par la digestion au site unique PmeI et déphosphorilé. 500ng de vecteur pC5300-Cas9 sont digérés avec 20U de PmeI (R0560S, NEB) à 37°C pendant une nuit puis déphosphorilé en suivant le protocole de NEB (Antarctic Phosphatase, M0289, NEB). La ligation de 50ng de vecteur pC5300-cas9 et 3 ng d'insert se fait avec l'enzyme quick ligase (M0202S, NEB), 5min à 24°C. Le vecteur final ainsi obtenu est cloné dans TOP10 (TOP10 Chemically Competent E. coli, C404010, Invitrogen®) puis séquencé. Le vecteur une fois vérifié est appelée pC5300Cas9::"nom du GENE cible"-gRNA1à3 puis transféré dans la souche d'*Agrobacterium tumefaciens* EHA105. Nous avons ciblé les 3 anti-COs et designer 3 crRNA pour chacun des gènes ciblées (Annexe 7).

✓ Transformation via *Agrobacterium tumefaciens*

Les vecteurs pC5300-Cas9:crRNA d'expression de l'ADN-T "édition du génome" ont ensuite servi à transformer des cals de riz suivant le protocole décrit par Sallaud (Sallaud et al. 2004). Et nous avons produit environ 24 plantes T0 par construits soit des Nipponbare soit dans Dongjin.

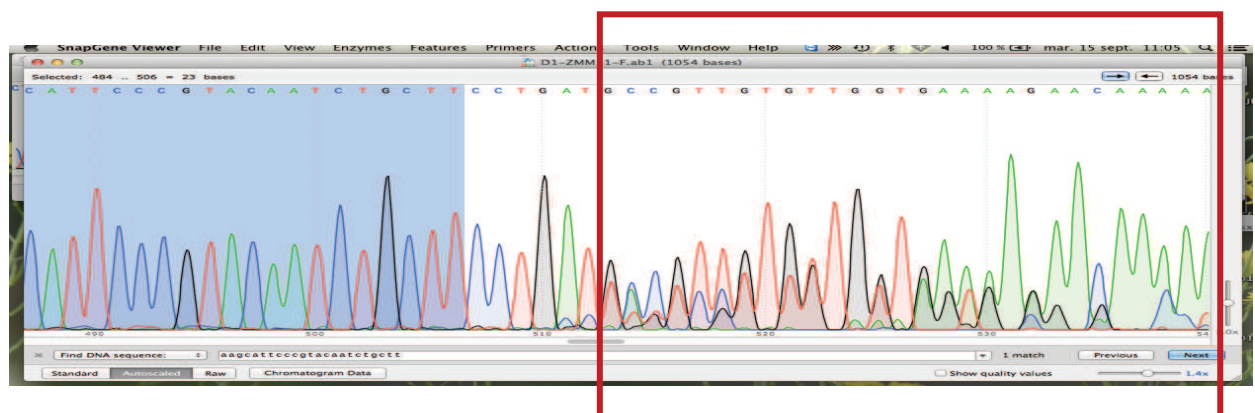
✓ Analyse des lésions par PCR et séquençage

L'ADN des plantes T0 est extrait en suivant le protocole MATAB (Mieulet et al. 2013). Les locus cibles sont amplifiés en utilisant 25 ng d'ADN (dans 25µL final) et des amorces encadrant le site de clivage (Figure 45) (KOD Hot Start DNA polymerase®, Novagen). Les conditions PCR après optimisation sont les suivantes : 95°C 2' et 35 cycles (95°C 20" ; 58°C 15" ; 70°C 35"). Le séquençage Sanger a été fait avec les amorces utilisées pour l'amplification du locus ciblé (GATC-Biotech). L'analyse des allèles éditées se fait par la lecture manuelle

A

Gène cible	Construit	A-(N)19-NGG	Primer	Séquence	Taille (pb)
<i>FANCM</i> (Os11g07870)	ZMM-16	TCCTACCGAGTGGCGCCAG-GGG	Z16-2F	GTAAGCGCCTCTTCTCCTC	595
			Z16-2R	CTTGCATGCCAATGAAAATG	
	ZMM-17	CCACGAGGCAGCGCGAACT-TGG	Z17_2F	GCCGATCCTACCGAGTGG	509
			Z17_2R	TGTTTCGCTTATATTGGAATCGT	
<i>RECQ14</i> (Os04g35420)	ZMM-11	ACCAACACAAAGGCCTAATC-AGG	ZMM11_F	GTTTTGGCTTGACATCATCCTT	1056
			ZMM11_R	TGCTGTCTAACTGAAAGGTGGTTA	
	ZMM-12	ATGCATCAGACGACTGAAAG-TGG	ZMM12_F	AGATTTGGATGACCTCTGG	885
			ZMM12_R	GGTTCCTCACTCCAATC	

B



C

Construits	Plantes TO	type	séquence (en rev/r au sgRNA)	fréquence
pC5300Cas9::RecQ4-gRNA2	ZMM11 (n=5)	Allèle-WT	CCTGATTAGGCCTTTGTGTTGGT	
		Allèle Mut-	CCTGAT- AGGCCTTTGTGTTGGT	1
		Allèle Mut-	CCTGAT - - GGCCTTTGTGTTGGT	2
		Allèle Mut-	CCTGAT - - - GCCTTTGTGTTGGT	3

Figure 45 : Analyse des plantes éditées par CRISPR/cas9.

A- Liste des amorces utilisées pour le génotypage des plantes TO CRISPR/cas9.

Amplification de la région encadrant la cible CRISP/cas9.

B- Chromatogramme double dans le cas d'une édition hétérozygote (1 allèle sauvage et allèle édité).

C- Type de lésions obtenu avec le CRISPR/Cas9 ciblant *OsRECQ14* (ZMM11).

des chromatogrammes (SnapGene® Viewer 4.0.4) (Figure 45). Le type d'édition observé peut être mono-allélique ou bi-allélique homozygote ou hétérozygote (Figure 45).

7.2. Résultats et discussion

✓ Clonage des sgRNA pour les différents gènes cibles

Le tableau de l'annexe 2 montre les différents gènes cibles et les différents guides qui ont été sélectionnés via des logiciels de prédiction des off target (crispor.tefor.net ou CRISPR-PLANT0.1) pour tenter le clonage de la cassette d'expression du gRNA dans le pC5300-Cas9. Dans la première série de clonage, 40% des cibles ont été rapidement clonées dans le vecteur pC5300-Cas9 (7/18) puis nous avons changé de stratégie de clonage (vecteur Miao) et obtenu 100% de réussite (2/2). Cependant l'avantage du premier vecteur était de pouvoir générer des éditions hétérozygotes (1 allèle édité et un allèle sauvage).

✓ Caractérisation des plantes T0 CRISPR/cas9

- De nouveaux allèles KO via CRISPR/Cas9 pour *OsFANCM*

Nous avons produit de nouveaux allèles au locus *OsFANCM*, en utilisant la technologie CRISPR/Cas9 d'édition du génome. Les construits pC5300Cas9::*FancM*-gRNA1 (ZMM16) et pC5300Cas9::*FancM*-gRNA2 (ZMM17) ont chacun permis de produire 18 et 30 plantes T0 respectivement (Annexe 7). L'amplification et/ou le séquençage Sanger de la région ciblée n'est pas d'une qualité suffisante pour pouvoir lire les chromatogrammes et déterminer s'il y a eu édition du génome ou non dans le gène *FANCM*. A l'heure actuelle, nous ne pouvons pas conclure sur la fonctionnalité de ces guides. Comme nous avons déjà deux allèles *Osfancm-1* et *2* dans deux fonds génétiques différents nous n'avons pas poursuivi l'analyse de ces transformants primaires (T0) mais l'analyse plus approfondie de ces plantes devra être réalisée si on veut utiliser les sgRNA pour transformer des hybrides F1, par exemple.

- De nouveaux allèles KO via CRISPR/Cas9 pour *OsRECQ14*

Deux séries de 23 plantes T0 ciblant l'exon 1 (ZMM11) ou 3 (ZMM12) du gène *OsRECQ14* (annexe 1) ont été produites dans un fond génétique NB par CRISPR-Cas9. L'analyse des lésions par PCR et séquençage Sanger montre que l'une des deux cibles sgRNA ne fonctionne pas puisqu'aucune des plantes transformées avec ce vecteur (ZMM12) n'est éditée alors que 22% des plantes T0 obtenues avec l'autre guide (ZMM11) sont éditées (5/23). Parmi ces plantes, 4 sont hétérozygotes mono-allélique (1 allèle édité + 1 allèle WT) et 1 plante est homozygote bi-allélique (2 allèles édités). Il s'agit de délétions de 1 à 3 nucléotides (Figure 45). Le guide ZMM11 qui fonctionne a été utilisé pour produire, cette fois dans un fond génétique DJ, 46 plantes T0 dont les semences F1 ont été récoltées. Ces plantes devront être caractérisées.

8. TRANSFORMATION D'HYBRIDES F1

Ayant validé certains sgRNA lors de la production de plantes T0 avec le premier vecteur de clonage pC5300-Cas9 (Annexe 2), nous avons utilisé le même sgRNA à savoir ZMM11 ciblant le gène *OsRECQ14* dans le vecteur Miao2.0 (Annexe 2 et 3). Ce nouveau construit cloné par I. Fayos a permis de transformer en mai 2017 des hybrides F1 IR64 x Azucena (D. Meynard et stage M. Cao-Van).

Les grains hybrides F1 utilisés ont été obtenus par pollinisation manuelle de la variété indica semi naine IR64 avec la variété japonica tropicale pluviale Azucena. Nous avons vérifié en préalable que la séquence cible des sgRNA définis sur la séquence de référence japonica Nipponbare ne présentait pas de variation chez IR64 et Azucena, qui pourrait diminuer ou abolir l'affinité du sgRNA. Les variétés Indica sont connues pour être plus récalcitrantes en culture in vitro, du fait d'une différenciation très rapide des cellules induites en callogenèse. S'agissant d'hybrides indica x japonica deux milieux de base de culture pouvaient être utilisés : un milieu de base MS adapté aux Indica (Hiei et al 2005) ou un milieu de base N6 plus adapté aux Japonica (Sallaud et al. 2004). Un milieu MS a finalement été choisi, mais s'est avéré moins discriminant lors des étapes de sélection.

Sur 60 grains hybrides mis en induction, 70% ont formé des cals secondaires qui ont été transformés avec le vecteur simple *RECQ14* (Miao2.0_sgRNAZMM11). Une seule plante transformée (ayant au moins une copie du gène de résistance à l'hygromycine, hpt) a été obtenue parmi une forte fréquence de plantes non transformées qui sont des échappées de la sélection. Ceci montre qu'une optimisation de la procédure de sélection doit être conduite. Les grains F2s de ces plantes seront récoltés en novembre. La plante transformée présente la même insertion d'un acide nucléique (adénine) sur les deux allèles. Cette plante T0 est homozygote mono-allélique et nous vérifierons par Q-RT-PCR le niveau d'expression du gène *OsRECQ14* par rapport aux plantes non transformées (contrôles) pour confirmer que la plante est KO. Enfin, le génotypage SNP des descendances simple mutant *Osrecq14*^{-/-} et du sauvage pourra être réalisé en fin d'année 2017.

Compte tenu des avantages présentés en introduction de ce chapitre, le CRISPR/Cas9 est une méthode de choix pour cibler les gènes anti-COs chez le riz. Le fait de pouvoir générer une série d'allèles est intéressant pour la caractérisation fonctionnelle des gènes. La possibilité de transformer des grains matures hybrides F1 est un autre avantage puisqu'il nous permettra de tester l'effet de l'accroissement du polymorphisme inter-parental sur l'augmentation de la recombinaison. En étudiant cette nouvelle descendance F2 de *Osrecq14*^{-/-} dans ce contexte hybride IR64 x Acunena, il sera très intéressant de comparer les cartes génétiques obtenues avec celle de la population NBxDJ (Partie 3) et celle-ci. Les séquences assemblées de IR64 et Azuneca devraient être prochainement disponibles.

Discussion générale

Discussion générale

Nous avons un double objectif à l'initiation de cette thèse : abolir la recombinaison en essayant de reproduire le phénotype *MiMe* et augmenter la recombinaison en utilisant les gènes anti-CO découverts chez *A. thaliana*. Au-delà de l'acquisition de connaissances, l'objectif finalisé était de développer des outils utilisables par les sélectionneurs dans les étapes amont de brassage génétique et aval de diffusion du progrès génétique. Ces objectifs ambitieux reposaient sur les avancées spectaculaires, réalisées ces dernières années dans le domaine de la méiose chez *A. thaliana*.

Conservation de la fonction des gènes méiotiques étudiés entre *Arabidopsis* et le riz

Une question qui se posait était de savoir si la mutation des gènes homologues chez le riz allait produire les mêmes effets que ceux observés chez *A. thaliana*. En effet si la méiose est globalement bien conservée chez les eucaryotes, ses mécanismes peuvent varier notamment en fonction de la taille du génome, comme par exemple la formation du bouquet chromosomique, ou entre dicotylédones et monocotylédones, comme la présence de cytokinèse à l'issue de chaque division méiotique (Hamant et al. 2006). Aussi, différents allèles mutants dans un même gène peuvent avoir des effets plus ou moins pénétrants sur la méiose mâle et femelle ou sur l'altération du processus en lui-même. C'est le cas de *swi1* d'*Arabidopsis* ou *am1* du riz et du maïs (Pawlowski et al. 2009). Les gènes dupliqués ne sont pas toujours fonctionnellement redondants, comme les gènes *SWI1* et *SWI2* d'*Arabidopsis* (Mercier et al. 2001, 2003). C'est aussi le cas de *ZIP1* et *ZIP2* (duplication en tandem), où il faut abolir chez *A. thaliana*, l'expression de *ZIP1* et *ZIP2* par RNAi pour inhiber la formation du complexe synaptonémal. Chez le riz et l'orge, la mutation de *ZEP1* (orthologue de *ZIP1*) est suffisant pour obtenir un phénotype méiotique. De plus le phénotype méiotique des mutants d'insertion *Oszep1* est très différent du phénotype RNAi *zip1/2* d'*Arabidopsis* et *zep1* de l'orge puisque *Oszep1* présente une augmentation de la recombinaison alors qu'une diminution est observée chez les autres plantes (Wang et al. 2010a; Barakate et al. 2014). Un autre exemple, est celui du gène *RAD51D* qui, chez le riz, apparaît essentiel à la méiose (Byun and Kim 2014) et ne l'est pas chez *A. thaliana* (Serra et al. 2013).

Toutes ces informations pouvaient laisser présager des difficultés en caractérisant chez le riz des mutations dans les gènes orthologues putatifs de ceux d'*Arabidopsis*. Or au cours de cette thèse, tous les gènes utilisés ont une fonction conservée et leurs mutations ou la combinaison de leurs mutations ont produit quasiment les mêmes effets phénotypiques. Lors de la reproduction de *MiMe* chez le riz, la seule différence observée est au niveau de la fertilité. Chez *OsMiMe*, la fertilité du triple mutant est restaurée par rapport à *Ososd1* mais n'est pas aussi importante que celle mesurée chez le sauvage. Une explication pourrait venir de la

différence de biologie du développement du grain entre *Oryza* et *A. thaliana*. Il se peut que la formation d'un albumen 5n non viable, empêche le développement de l'embryon triploïde. Dans le cas de l'étude du gène *FANCM*, une différence notable réside dans le fait qu'en contexte hybride (3,6% de polymorphisme entre NB et DJ) la mutation du gène *OsFANCM* a un effet sur la recombinaison alors que chez *A. thaliana* l'hélicase *FANCM* perd son effet anti-CO quand la recombinaison ne se fait plus entre deux lignées pures. Il sera donc très intéressant, de réaliser cette mutation par CRISPR/Cas9 dans des contextes hybrides intergroupes mettant en jeu un polymorphisme plus important pour voir si l'effet de *FANCM* sur la recombinaison se manifeste toujours. Une autre différence, est liée à la distribution des COs surnuméraires sur les bras chromosomiques (Figure 40). Alors que chez *A. thaliana* on observe que la recombinaison augmente de façon linéaire du centromère vers le télomère, chez le riz, celle-ci est beaucoup plus irrégulière (Figure 40).

Enfin, un article très récent fait état de la stérilité du mutant *Osfignl1* chez le riz (Zhang et al. 2017), alors que chez *Arabidopsis*, le mutant *Atfigl1* n'entraîne que rarement des défauts mineurs de méiose (Fernandes et al. 2017b). Si ce constat est avéré, alors il s'agit d'une différence majeure entre les deux plantes et cette stérilité empêcherait, évidemment, l'utilisation de *Osfignl1*. Nous envisageons de vérifier les données de ce papier en reproduisant le mutant *Osfignl1* dans un fond génétique Nipponbare et de vérifier s'il existe une diversité allélique naturelle dans ce gène chez les 3000 génomes de riz.

Intérêt et limites de l'exploitation de l'augmentation de la recombinaison dans les étapes de brassage génétique des schémas de sélection.

Le développement de nouvelles variétés repose sur la création de nouvelles combinaisons d'allèles. Ces nouvelles combinaisons sont obtenues en mixant les allèles existant dans les collections élités (pool de variétés donneuses) ou en introduisant de la variabilité à partir des collections de ressources génétiques (allèles sauvages). Dans tous les cas, l'obtention de nouvelles combinaisons d'allèles dépend de la recombinaison génétique. Augmenter globalement la recombinaison pour accroître la résolution génétique lors de la création de populations de type MAGIC ou NAM (§1.5.2) pourrait avoir un intérêt pour la dissection des caractères polygéniques, l'accélération du clonage positionnel des gènes, et fournir des associations d'allèles originales dans du matériel de pré-breeding. L'introduction de ce type de mutation pourrait également être utile dans des populations de sélection récurrente qui alternent des cycles de brassage génétique par inter-croisements, pouvant être facilités par un gène de stérilité mâle génique, et de sélection, distinctes de la sortie variétale.

Nous avons vu que chez *Arabidopsis*, l'augmentation de la recombinaison liée aux mutations *fancm* et *recq4* survient sur les bras chromosomiques, de façon croissante et linéaire depuis les régions péri-centromériques, totalement dépourvues de recombinaison, vers les télomères. Chez le riz cette variation du taux de recombinaison est nettement moins linéaire mais comme

chez *A. thaliana* la région péri-centromérique ne recombine pas. Ceci semble indiquer, que les régions péri-centromériques, et les gènes qui y résident resteront relativement inaccessibles. La taille des centromères est bien plus importante chez des plantes à grands génomes, comme le maïs et le blé, et l'étude de ce phénomène apportera sans doute plus d'informations sur ce plan que chez le riz où la régions péri-centromérique sont relativement restreintes.

Il sera aussi intéressant d'étudier en détail avec un marquage plus fin ou un séquençage de descendances l'influence de la mutation sur l'augmentation de l'activité recombinogène des régions considérées comme des cold spots ou des hot spots de recombinaison et possédant un polymorphisme nucléotidique contrasté. Chez *Arabidopsis*, les régions très polymorphes, comme les cluster de gènes de résistance de type NBS LRR en évolution rapide semblent plus réfractaires à l'augmentation de la recombinaison causée par la mutation des gènes anti-COs (Fernandes et al. 2017b).

Au-delà de l'augmentation de la recombinaison dans les régions péri-centromériques, les sélectionneurs souhaitent également augmenter la recombinaison dans les zones d'hétérologie structurale telles qu'on les rencontre fréquemment dans le cas de croisement inter-subspécifiques ou interspécifiques. L'intérêt d'augmenter la recombinaison de manière globale est de casser des associations défavorables. Toutefois, quand on s'intéresse à des variétés élites déjà très performantes, le risque symétrique est de casser des associations favorables qui ont été progressivement sélectionnées. Les modèles de génétique quantitative démontrent effectivement que l'augmentation de la recombinaison n'a pas toujours un effet positif sur le progrès génétique mais que cela dépend des caractères et des associations alléliques chez les parents. Il est donc probable que l'augmentation de la recombinaison sera plus intéressante dans un contexte de variétés peu sélectionnées, comme les accessions traditionnelles.

Augmentation globale et / ou ciblage de la recombinaison ?

Augmenter ou cibler la recombinaison, dans les programmes de sélection des espèces cultivées, devient progressivement envisageable. Deux articles récents de généticiens quantitatifs discutent des gains génétiques possible en utilisant deux approches (McClosky and Tanksley 2013; Bernardo 2017).

Le ciblage d'un grand nombre de régions d'intérêt permettrait d'espérer un doublement du gain génétique pour les caractères quantitatifs chez le maïs. Des stratégies de ciblage de la recombinaison chez les plantes sont en cours de développement, basées sur des travaux pionniers réalisés chez la levure. L'induction de COs somatiques a été montré comme étant possible après CDBs induite par CRISPR/Ca9, permettant même la cartographie de gènes sans croisement (Sadhu et al. 2016). L'induction ciblée de COs méiotiques par CRISPR/Cas9 n'a cependant pas encore été rapportée et sera sans doute, sauf surprise, limitée par la

régulation très fine de leur programmation et l'intervention spécifique de nombreux acteurs protéiques devant interagir entre eux pour la résolution de la CDB en CO méiotique. Une autre stratégie en cours pour augmenter la recombinaison de façon ciblée est d'exprimer une fusion entre SPO11 et un domaine de reconnaissance de l'ADN permettant de diriger la protéine SPO11 (technologie dead-cas9) (Robine et al. 2007; Sarno et al. 2017) et de former une CDB au site reconnu donc de donner plus de chance à la formation d'un CO à ce locus. Les fréquences de recombinaison au site cible sont augmentées d'un facteur 2 à 6 en contexte *spo11* mutant, cette fréquence variant en fonction d'une interaction entre la nature de la cible et les systèmes de ciblage (Sarno et al. 2017). Le multiplexage de cibles, qui fait partie de la simulation du gain génétique de Bernardo (Bernardo 2017) chez le maïs, n'a pour le moment pas encore été tenté. La compagnie MeioGenix (liée à l'Institut Curie) explore actuellement cette voie sur le riz, le maïs et le blé en collaboration avec le CIRAD et l'INRA. Chez les plantes le système est complexifié par l'existence de plusieurs protéines SPO11, dont au moins deux, SPO11-1 et SPO11-2 ont été montrées comme étant indispensables à l'initiation de la méiose. Il pourrait être intéressant de voir si la combinaison de la technologie de ciblage avec celle d'augmentation globale de la recombinaison apportée par la mutation des gènes anti-COs pourrait augmenter l'efficacité de ce système.

Implémentation dans les programmes de sélection

Pour pouvoir utiliser un outil qui permette d'augmenter de façon globale la recombinaison les sélectionneurs seraient intéressés par un système génétique dominant, et non récessif comme les mutations, et qui puisse être éliminé par ségrégation quand le sélectionneur le désire. Les semenciers favoriseraient une solution du type pulvérisation d'un produit chimique à effet temporaire lors de la méiose dans la phase de brassage génétique mais aucun produit n'est encore en développement et sa mise au point nécessitera une connaissance très approfondie des mécanismes. Une voie à explorer rapidement serait de réaliser un knock-down le plus strict possible des gènes anti-COs par CRISPR interférence sous le contrôle d'un promoteur spécifique de la méiose. Les exemples de CRISPR interférence restent encore limités chez les plantes et il faut souvent définir plusieurs sgRNA dans la région promotrice ou 5'UTR du gène pour parvenir à la répression de l'expression du gène. Une stratégie alternative à l'interférence est l'utilisation de mutations dominantes négatives comme illustré dans le cas de TDM.

Exploitation de l'apomixie synthétique pour la diffusion clonale par grain de l'hétérosis

Les sélectionneurs ont mis au point des variétés hybrides de riz qui produisent 20% de plus que les lignées pures. Mais du fait du coût de production et de vente élevé des semences, ces hybrides F1 sont assez peu cultivés sauf en Chine. L'ouverture du marché des hybrides F1 pourrait, lui-aussi, aider à satisfaire la demande en riz. Cependant, le riz est une plante autogame et la production d'hybrides F1 nécessite l'utilisation d'un système de stérilité mâle

géo-cytoplasmique. L'apomixie a donc été envisagée comme une solution permettant de faire bénéficier les agriculteurs des pays du Sud des avantages des hybrides F1 sans leurs inconvénients actuels. En permettant la diffusion de formules hybrides F1 clonale par grain. Ce système permettrait de bénéficier de l'hétérosis dans un contexte génétique immortalisé.

Une première étape sera d'améliorer la fertilité des combinaisons *Ososd1*^{-/-} (ou *MiMe*) qui n'est actuellement pas de 100%. Dans un contexte de production et non plus d'analyse génétique, la F1 doit en effet être pleinement fertile. La seconde étape sera de démontrer que l'on peut déclencher la parthénogénèse en utilisant un des inducteurs connus d'haploïdes qui incluent les stratégies transgéniques CENH3 et GFP-tailswap (GEM) ou BBML (§4.3). Pour le moment, chez *A. thaliana*, le croisement de *MiMe* avec une lignée GEM permet d'obtenir 30% de production de graines apomictiques (Marimuthu et al. 2011). La surexpression de BBML (Conner et al. 2015; Conner and Ozias-Akins 2017) chez *MiMe* pourrait permettre également de déclencher la parthénogénèse. Ces stratégies devraient pouvoir être améliorées en optimisant les construits, notamment au niveau des régions régulatrices, et devraient permettre d'aboutir à un pourcentage de parthénogénèse supérieur à celui observé actuellement chez *A. thaliana*.

L'association de *MiMe* avec une mutation déclenchant l'induction du développement parthénogénétique serait donc la plus prometteuse. Le récent clonage du gène de phospholipase Matriléale not Like Dad (*NLD*) (Gilles et al. 2017; Kelliher et al. 2017) ouvre la voie à l'utilisation de cette mutation chez le riz, chez lequel un orthologue putatif est clairement identifié. La création de mutations par CRISPR/Cas9 dans le gène *NLD* chez le maïs a permis d'aboutir à une fréquence moyenne d'induction de 6,5% d'induction chez un génotype non inducteur (Kelliher et al. 2017). Cette stratégie suppose de transformer des grains hybrides F1, simultanément avec des sgRNA ciblant les trois gènes *OSD1*, *PAIR1* et *REC8* et le gène *NLD*. Il est possible que la compréhension du mécanisme d'action de *NLD* permette d'accéder à de nouvelles cibles dont la manipulation pourrait améliorer cette fréquence.

En conclusion, ces travaux sur le riz nous permettent de croire qu'on s'approche de l'objectif convoité depuis plus de 30 ans qui est d'introduire l'apomixie chez les céréales. De plus on est aussi capable d'augmenter le taux de recombinaison d'un facteur 3,3 en moyenne sur l'ensemble du génome. Ces résultats sont une première chez les céréales et ouvrent des perspectives d'utilisation de ces technologies pour l'amélioration des plantes. Il ne fait aucun doute que les travaux de recherche sur la méiose, sur la plante modèle *Arabidopsis*, leurs transpositions sur les espèces cultivées ainsi que les projets visant au développement des nouvelles technologies, qu'ils soient financés par le secteur public ou privé ont grandement facilité ce travail.

Bibliographie

Bibliographie

- Abe A, Kosugi S, Yoshida K, et al (2012) Genome sequencing reveals agronomically important loci in rice using MutMap. *Nat Biotechnol* 30:174–178. doi: 10.1038/nbt.2095
- Acquaviva L, Szekvolgyi L, Dichtl B, et al (2013) The COMPASS Subunit Spp1 Links Histone Methylation to Initiation of Meiotic Recombination. *Science* (80-) 339:215–218. doi: 10.1126/science.1225739
- Allers T, Lichten M (2001) Differential timing and control of noncrossover and crossover recombination during meiosis. *Cell* 106:47–57.
- Anderson LK, Lohmiller LD, Tang X, et al (2014) Combined fluorescent and electron microscopic imaging unveils the specific properties of two classes of meiotic crossovers. *Proc Natl Acad Sci U S A* 111:13415–20. doi: 10.1073/pnas.1406846111
- Armstrong SJ, Jones GH (2003) Meiotic cytology and chromosome behaviour in wild-type *Arabidopsis thaliana*. *J Exp Bot* 54:1–10.
- Ballini E, Morel J-B, Droc G, et al (2008) A Genome-Wide Meta-Analysis of Rice Blast Resistance Genes and Quantitative Trait Loci Provides New Insights into Partial and Complete Resistance. *Mol Plant-Microbe Interact* 21:859–868. doi: 10.1094/MPMI-21-7-0859
- Barakate A, Higgins JD, Vivera S, et al (2014) The synaptonemal complex protein ZYP1 is required for imposition of meiotic crossovers in barley. *Plant Cell* 26:729–40. doi: 10.1105/tpc.113.121269
- Barcaccia G, Albertini E (2013) Apomixis in plant reproduction: a novel perspective on an old dilemma. *Plant Reprod* 26:159–179. doi: 10.1007/s00497-013-0222-y
- Barrell PJ, Grossniklaus U (2005) Confocal microscopy of whole ovules for analysis of reproductive development: the elongate1 mutant affects meiosis II. *Plant J* 43:309–320. doi: 10.1111/j.1365-313X.2005.02456.x
- Barret P, Brinkmann M, Beckert M (2008) A major locus expressed in the male gametophyte with incomplete penetrance is responsible for in situ gynogenesis in maize. *Theor Appl Genet* 117:581–94. doi: 10.1007/s00122-008-0803-6
- Berchowitz LE, Copenhaver GP (2010) Genetic interference: don't stand so close to me. *Curr Genomics* 11:91–102. doi: 10.2174/138920210790886835
- Berchowitz LE, Copenhaver GP (2008) Fluorescent *Arabidopsis* tetrads: a visual assay for quickly developing large crossover and crossover interference data sets. *Nat Protoc* 3:41–50. doi: 10.1038/nprot.2007.491
- Berchowitz LE, Francis KE, Bey AL, et al (2007) The Role of AtMUS81 in Interference-Insensitive Crossovers in *A. thaliana*. *PLoS Genet* 3:e132. doi: 10.1371/journal.pgen.0030132
- Bergerat A, de Massy B, Gadelle D, et al (1997) An atypical topoisomerase II from archaea with implications for meiotic recombination. *Nature* 386:414–417. doi: 10.1038/386414a0
- Bernardo R (2017) Prospective Targeted Recombination and Genetic Gains for Quantitative Traits in Maize. doi: 10.3835/plantgenome2016.11.0118
- Blat Y, Protacio RU, Hunter N, et al (2002) Physical and functional interactions among basic chromosome organizational features govern early steps of meiotic chiasma formation. *Cell* 111:791–802. doi: 10.1016/S0092-8674(02)01167-4
- Bonnet S, Knoll A, Hartung F, Puchta H (2013) Different functions for the domains of the *Arabidopsis thaliana* RMI1 protein in DNA cross-link repair, somatic and meiotic

- recombination. *Nucleic Acids Res* 41:9349–60. doi: 10.1093/nar/gkt730
- Börner GV, Kleckner N, Hunter N (2004) Crossover/noncrossover differentiation, synaptonemal complex formation, and regulatory surveillance at the leptotene/zygotene transition of meiosis. *Cell* 117:29–45.
- Boutillier K, Offringa R, Sharma VK, et al (2002) Ectopic expression of BABY BOOM triggers a conversion from vegetative to embryonic growth. *Plant Cell* 14:1737–49.
- Brown MS, Grubb J, Zhang A, et al (2015) Small Rad51 and Dmc1 Complexes Often Co-occupy Both Ends of a Meiotic DNA Double Strand Break. *PLoS Genet.* doi: 10.1371/journal.pgen.1005653
- Buhler C, Lebbink JHG, Bocs C, et al (2001) DNA Topoisomerase VI Generates ATP-dependent Double-strand Breaks with Two-nucleotide Overhangs*. *Publ JBC Pap Press.* doi: 10.1074/jbc.M101823200
- Byun MY, Kim WT (2014) Suppression of *OsRAD51D* results in defects in reproductive development in rice (*Oryza sativa* L.). *Plant J* 79:256–269. doi: 10.1111/tpj.12558
- Caldana C, Scheible W-R, Mueller-Roeber B, Ruzicic S (2007) A quantitative RT-PCR platform for high-throughput expression profiling of 2500 rice transcription factors. *Plant Methods* 3:7. doi: 10.1186/1746-4811-3-7
- Carpenter ATC (1975) Electron microscopy of meiosis in *Drosophila melanogaster* females: II: The recombination nodule-a recombination-associated structure at pachytene? (crossing-over/serial reconstruction). *Genetics* 72:3186–3189.
- CGIAR (2013) Rice Almanac. IRRI
- Champoux JJ (2001) DNA Topoisomerases: Structure, Function, and Mechanism. *Annu Rev Biochem* 70:369–413. doi: 10.1146/annurev.biochem.70.1.369
- Chapman JR, Taylor MRG, Boulton SJ (2012) Playing the End Game: DNA Double-Strand Break Repair Pathway Choice. *MOLCEL* 47:497–510. doi: 10.1016/j.molcel.2012.07.029
- Chelysheva L, Gendrot G, Vezon D, et al (2007) Zip4/Spo22 Is Required for Class I CO Formation but Not for Synapsis Completion in *Arabidopsis thaliana*. *PLoS Genet* 3:e83. doi: 10.1371/journal.pgen.0030083
- Chelysheva L, Vezon D, Chambon A, et al (2012) The *Arabidopsis* HEI10 Is a New ZMM Protein Related to Zip3. *PLoS Genet* 8:e1002799. doi: 10.1371/journal.pgen.1002799
- Chen C, Zhang W, Timofejeva L, et al (2005) The *Arabidopsis* ROCK-N-ROLLERS gene encodes a homolog of the yeast ATP-dependent DNA helicase MER3 and is required for normal meiotic crossover formation. *Plant J* 43:321–334. doi: 10.1111/j.1365-3113X.2005.02461.x
- Chen M, Presting G, Barbazuk WB, et al (2002) An integrated physical and genetic map of the rice genome. *Plant Cell* 14:537–45. doi: 10.1105/TPC.010485
- Cheng Z, Buell CR, Wing RA, et al (2001) Toward a Cytological Characterization of the Rice Genome. *Genome Res* 11:2133–2141. doi: 10.1101/gr.194601
- Chikashige Y, Ding DQ, Funabiki H, et al (1994) Telomere-led premeiotic chromosome movement in fission yeast. *Science* 264:270–3.
- Choi K, Henderson IR (2015) Meiotic recombination hotspots - a comparative view. *Plant J* 83:52–61. doi: 10.1111/tpj.12870
- Choulet F, Alberti A, Theil S, et al (2014) Structural and functional partitioning of bread wheat chromosome 3B. *Science* (80-) 345:1249721–1249721. doi: 10.1126/science.1249721
- Christou P, Ford TL, Kofron M (1991) Production of Transgenic Rice (*Oryza Sativa* L.) Plants from Agronomically Important Indica and Japonica Varieties via Electric Discharge Particle Acceleration of Exogenous DNA into Immature Zygotic Embryos. *Bio/Technology* 9:957–962. doi: 10.1038/nbt1091-957

- Cifuentes M, Jolivet S, Cromer L, et al (2016) TDM1 Regulation Determines the Number of Meiotic Divisions. *PLoS Genet* 12:e1005856. doi: 10.1371/journal.pgen.1005856
- Civáň P, Craig H, Cox CJ, Brown TA (2015) Three geographically separate domestications of Asian rice. *Nat Plants* 1:15164. doi: 10.1038/nplants.2015.164
- Cloud V, Chan Y-L, Grubb J, et al (2012) Rad51 Is an Accessory Factor for Dmc1-Mediated Joint Molecule Formation During Meiosis. *Science* (80-) 337:1222–1225. doi: 10.1126/science.1219379
- Conner JA, Mookkan M, Huo H, et al (2015) A parthenogenesis gene of apomict origin elicits embryo formation from unfertilized eggs in a sexual plant. *Proc Natl Acad Sci U S A* 112:11205–10. doi: 10.1073/pnas.1505856112
- Conner JA, Ozias-Akins P (2017) *Apomixis: Engineering the Ability to Harness Hybrid Vigor in Crop Plants*. Humana Press, New York, NY, pp 17–34
- Courtois B (1988) Les systèmes de culture du riz pluvial. Mémoire Trav l'IRAT 16 96p.
- Crismani W, Girard C, Froger N, et al (2012) FANCM Limits Meiotic Crossovers. Sup. material. *Science* (80-) 336:1588–1590. doi: 10.1126/science.1220381
- Crismani W, Girard C, Mercier R (2013) Tinkering with meiosis. *J Exp Bot* 64:55–65. doi: 10.1093/jxb/ers314
- Cromer L, Heyman J, Touati S, et al (2012) OSD1 pro1. Cromer L, Heyman J, Touati S, et al. OSD1 promotes meiotic progression via APC/C inhibition and forms a regulatory network with TDM and CYCA1;2/TAM. *PLoS Genet* 2012; 8:e1002865. notes meiotic progression via APC/C inhibition and forms a regulato. *PLoS Genet* 8:e1002865. doi: 10.1371/journal.pgen.1002865
- Cromer L, Jolivet S, Horlow C, et al (2013) Centromeric cohesion is protected twice at meiosis, by SHUGOSHINs at anaphase I and by PATRONUS at interkinesis. *Curr Biol* 23:2090–2099. doi: 10.1016/j.cub.2013.08.036
- D'Erfurth I, Cromer L, Jolivet S, et al (2010) The cyclin-A CYCA1;2/TAM is required for the meiosis I to meiosis II transition and cooperates with OSD1 for the prophase to first meiotic division transition. *PLoS Genet* 6:e1000989. doi: 10.1371/journal.pgen.1000989
- D'Erfurth I, Jolivet S, Froger N, et al (2009) Turning meiosis into mitosis. *PLoS Biol* 7:e1000124. doi: 10.1371/journal.pbio.1000124
- Da Ines O, Degroote F, Goubely C, et al (2013) Meiotic Recombination in Arabidopsis Is Catalysed by DMC1, with RAD51 Playing a Supporting Role. *PLoS Genet* 9:e1003787. doi: 10.1371/journal.pgen.1003787
- Danilowicz C, Peacock-Villada A, Vlassakis J, et al (2014) The differential extension in dsDNA bound to Rad51 filaments may play important roles in homology recognition and strand exchange. *Nucleic Acids Res* 42:526–533. doi: 10.1093/nar/gkt867
- De Muyt A, Jessop L, Kolar E, et al (2012) BLM Helicase Ortholog Sgs1 Is a Central Regulator of Meiotic Recombination Intermediate Metabolism. *Mol Cell* 46:43–53. doi: 10.1016/j.molcel.2012.02.020
- De Guzman F (2016) The International Rice Genebank. International Rice Research Institute (IRRI)
- De Massy B (2013) Spp1 Links Sites of Meiotic DNA Double-Strand Breaks to Chromosome Axes. *MOLCEL* 49:3–5. doi: 10.1016/j.molcel.2012.12.011
- De Muyt A, Pereira L, Vezon D, et al (2009) A high throughput genetic screen identifies new early meiotic recombination functions in Arabidopsis thaliana. *PLoS Genet* 5:e1000654. doi: 10.1371/journal.pgen.1000654
- De Muyt A, Vezon D, Gendrot G, et al (2007) AtPRD1 is required for meiotic double strand break formation in Arabidopsis thaliana. *EMBO J* 26:4126–37. doi: 10.1038/sj.emboj.7601815

- Devos KM, Gale MD (1997) Comparative genetics in the grasses. *Plant Mol Biol* 35:3–15.
- Ding X, Xu R, Yu J, et al (2007) SUN1 Is Required for Telomere Attachment to Nuclear Envelope and Gametogenesis in Mice. *Dev Cell* 12:863–872. doi: 10.1016/j.devcel.2007.03.018
- Droc G, An G, Wu C, et al (2013) Mutant Resources for Functional Analysis of the Rice Genome. In: *Genetics and Genomics of Rice*. Springer New York, New York, NY, pp 81–115
- Droc G, Ruiz M, Larmande P, et al (2006) OryGenesDB: a database for rice reverse genetics. *Nucleic Acids Res* 34:D736–D740. doi: 10.1093/nar/gkj012
- Drouaud J, Khademian H, Giraut L, et al (2013) Contrasted patterns of crossover and non-crossover at *Arabidopsis thaliana* meiotic recombination hotspots. *PLoS Genet* 9:e1003922. doi: 10.1371/journal.pgen.1003922
- Escobar-Guzmán R, Rodríguez-Leal D, Vielle-Calzada J-P, Ronceret A (2015) Whole-mount immunolocalization to study female meiosis in *Arabidopsis*. *Nat Protoc*. doi: 10.1038/nprot.2015.098
- Fernandes J, Duhamel M, Seguela-Arnaud M, et al (2017a) FIGL1 and its novel partner FLIP form a conserved complex that regulates homologous recombination. *bioRxiv*. doi: 10.1101/159657
- Fernandes JB, Seguela-Arnaud M, Larcheveque C, et al (2017b) Unleashing meiotic crossovers in hybrid plants. *bioRxiv*. doi: <http://dx.doi.org/10.1101/159640>.
- Flint-Garcia SA, Thornsberry JM, Buckler ES (2003) Structure of Linkage Disequilibrium in Plants. *Annu Rev Plant Biol* 54:357–374. doi: 10.1146/annurev.arplant.54.031902.134907
- Fragoso CA, Moreno M, Wang Z, et al (2017) Genetic Architecture of a Rice Nested Association Mapping Population. *G3* 7:117. doi: 10.1534/g3.117.041608
- Franklin AE, McElver J, Sunjevaric I, et al (1999) Three-Dimensional Microscopy of the Rad51 Recombination Protein during Meiotic Prophase.
- Gao Y, Zhang Y, Zhang D, et al (2015) Auxin binding protein 1 (ABP1) is not required for either auxin signaling or *Arabidopsis* development. *Proc Natl Acad Sci U S A* 112:2275–80. doi: 10.1073/pnas.1500365112
- Garcia-Aguilar M, Michaud C, Leblanc O, Grimanelli D (2010) Inactivation of a DNA Methylation Pathway in Maize Reproductive Organs Results in Apomixis-Like Phenotypes. *Plant Cell* 22:3249–3267. doi: 10.1105/tpc.109.072181
- Geuting V, Kobbe D, Hartung F, et al (2009) Two distinct MUS81-EME1 complexes from *Arabidopsis* process Holliday junctions. *Plant Physiol* 150:1062–71. doi: 10.1104/pp.109.136846
- Gilles LM, Khaled A, Laffaire J-B, et al (2017) Loss of pollen-specific phospholipase NOT LIKE DAD triggers gynogenesis in maize. *EMBO J* 36:707–717. doi: 10.15252/embj
- Girard C, Chelysheva L, Choinard S, et al (2015) AAA-ATPase FIDGETIN-LIKE 1 and Helicase FANCM Antagonize Meiotic Crossovers by Distinct Mechanisms. *PLOS Genet* 11:e1005369. doi: 10.1371/journal.pgen.1005369
- Girard C, Crismani W, Froger N, et al (2014) FANCM-associated proteins MHF1 and MHF2, but not the other Fanconi anemia factors, limit meiotic crossovers. *Nucleic Acids Res* 42:9087–9095. doi: 10.1093/nar/gku614
- Glaszmann JC (1987) Isozymes and classification of Asian rice varieties. *Theor Appl Genet* 74:21–30. doi: 10.1007/BF00290078
- Goff SA, Ricke D, Lan T-H, et al (2002) A Draft Sequence of the Rice Genome (*Oryza sativa* L. ssp. *japonica*).

- Grelon M, Vezon D, Gendrot G, Pelletier G (2001) AtSPO11-1 is necessary for efficient meiotic recombination in plants. *EMBO J* 20:589–600. doi: 10.1093/emboj/20.3.589
- Gross BL, Zhao Z (2014) Archaeological and genetic insights into the origins of domesticated rice. *Proc Natl Acad Sci U S A* 111:6190–7. doi: 10.1073/pnas.1308942110
- Guillon A-E, Berger F (2005) Loss of function of MULTICOPY SUPPRESSOR OF IRA 1 produces nonviable parthenogenetic embryos in *Arabidopsis*. *Curr Biol* 15:750–4. doi: 10.1016/j.cub.2005.02.066
- Guo L, Gao Z, Qian Q (2014) Application of resequencing to rice genomics, functional genomics and evolutionary analysis. *Rice (N Y)* 7:4. doi: 10.1186/s12284-014-0004-7
- Hamant O, Ma H, Cande WZ (2006) GENETICS OF MEIOTIC PROPHASE I IN PLANTS. *Annu Rev Plant Biol* 57:267–302. doi: 10.1146/annurev.arplant.57.032905.105255
- Hand ML, Koltunow AMG, Akiyama Y, et al (2014) The genetic control of apomixis: asexual seed formation. *Genetics* 197:441–50. doi: 10.1534/genetics.114.163105
- Harper L (2004) A bouquet of chromosomes. *J Cell Sci* 117:4025–4032. doi: 10.1242/jcs.01363
- Hartung F, Puchta H (2006) The RecQ gene family in plants. *J Plant Physiol* 163:287–296. doi: 10.1016/j.jplph.2005.10.013
- Hartung F, Suer S, Knoll A, et al (2008) Topoisomerase 3alpha and RMI1 suppress somatic crossovers and are essential for resolution of meiotic recombination intermediates in *Arabidopsis thaliana*. *PLoS Genet* 4:e1000285. doi: 10.1371/journal.pgen.1000285
- Harushima Y, Yano M, Shomura A, et al (1998) A High-Density Rice Genetic Linkage Map with 2275 Markers Using a Single F₂ Population. *Genetics*, 148, 479–494. Retrieved from <https://www.ncbi.nlm.nih.gov>. *Genetics* 148:479–494.
- Hauf S, Watanabe Y (2004) Kinetochore orientation in mitosis and meiosis. *Cell* 119:317–327.
- Heffelfinger C, Fragoso CA, Lorieux M (2017) Constructing linkage maps in the genomics era with MapDisto 2.0. *Bioinformatics* 33:2224–2225. doi: 10.1093/bioinformatics/btx177
- Heyer W-D, Ehmsen KT, Liu J (2010) Regulation of homologous recombination in eukaryotes. *Annu Rev Genet* 44:113–39. doi: 10.1146/annurev-genet-051710-150955
- Hiei Y, Ohta S, Komari T, Kumashiro T (1994) Efficient transformation of rice (*Oryza sativa* L.) mediated by *Agrobacterium* and sequence analysis of the boundaries of the T-DNA. *Plant J* 6:271–282. doi: 10.1046/j.1365-313X.1994.6020271.x
- Higgins JD, Sanchez-Moran E, Armstrong SJ, et al (2005) The *Arabidopsis* synaptonemal complex protein ZYP1 is required for chromosome synapsis and normal fidelity of crossing over. *Genes Dev* 19:2488–500. doi: 10.1101/gad.354705
- Higgins JD, Vignard J, Mercier R, et al (2008) AtMSH5 partners AtMSH4 in the class I meiotic crossover pathway in *Arabidopsis thaliana*, but is not required for synapsis. *Plant J* 55:28–39. doi: 10.1111/j.1365-313X.2008.03470.x
- Hillers KJ, Villeneuve AM, Duffy JB, et al (2003) Chromosome-wide control of meiotic crossing over in *C. elegans*. *Curr Biol* 13:1641–7. doi: 10.1016/J.CUB.2003.08.026
- Hirochika H, Guiderdoni E, An G, et al (2004) Rice mutant resources for gene discovery. *Plant Mol Biol* 54:325–334. doi: 10.1023/B:PLAN.0000036368.74758.66
- Hörandl E (2009) A combinational theory for maintenance of sex. *Heredity (Edinb)* 103:445–457. doi: 10.1038/hdy.2009.85
- Hu L, Li N, Xu C, et al (2014) Mutation of a major CG methylase in rice causes genome-wide hypomethylation, dysregulated genome expression, and seedling lethality. *Proc Natl Acad Sci* 111:10642–10647. doi: 10.1073/pnas.1410761111
- Hu Q, Li Y, Wang H, et al (2017) MEICA 1 (meiotic chromosome association 1) interacts with

- TOP3 α and regulates meiotic recombination in rice. *Plant Cell*. doi: 10.1105/tpc.17.00241
- Huang X, Kurata N, Wei X, et al (2012) A map of rice genome variation reveals the origin of cultivated rice. doi: 10.1038/nature11532
- Hunter N (2015) Meiotic Recombination: The Essence of Heredity. *Cold Spring Harb Perspect Biol* a016618. doi: 10.1101/cshperspect.a016618
- Hunter N, Kleckner N, Huynh AD, et al (2001) The single-end invasion: an asymmetric intermediate at the double-strand break to double-holliday junction transition of meiotic recombination. *Cell* 106:59–70. doi: 10.1016/S0092-8674(01)00430-5
- IRRI (2017) FAO - News Article: IRRI and FAO step up joint efforts to globally bolster sustainable rice production.
- Jackson N, Sanchez-Moran E, Buckling E, et al (2006) Reduced meiotic crossovers and delayed prophase I progression in AtMLH3-deficient Arabidopsis. *EMBO J* 25:1315–23. doi: 10.1038/sj.emboj.7600992
- Ji J, Tang D, Shen Y, et al (2016) P31comet, a member of the synaptonemal complex, participates in meiotic DSB formation in rice. *Proc Natl Acad Sci U S A* 113:10577–82. doi: 10.1073/pnas.1607334113
- Jinek M, Chylinski K, Fonfara I, et al (2012) A Programmable Dual-RNA-Guided DNA Endonuclease in Adaptive Bacterial Immunity. *Science* (80-) 337:816–821. doi: 10.1126/science.1225829
- Jones GH, Franklin FCH, Whitehorn MAF, et al (2006) Meiotic crossing-over: obligation and interference. *Cell* 126:246–8. doi: 10.1016/j.cell.2006.07.010
- Kawahara Y, de la Bastide M, Hamilton JP, et al (2013) Improvement of the *Oryza sativa* Nipponbare reference genome using next generation sequence and optical map data. *Rice* 6:4. doi: 10.1186/1939-8433-6-4
- Keeney S, Neale MJ (2006) Initiation of meiotic recombination by formation of DNA double-strand breaks: mechanism and regulation.
- Kelliher T, Starr D, Richbourg L, et al (2017) MATRILINEAL, a sperm-specific phospholipase, triggers maize haploid induction. *Nature* 542:105–109. doi: 10.1038/nature20827
- Kelliher T, Starr D, Wang W, et al (2016) Maternal Haploids Are Preferentially Induced by CENH3-tailswap Transgenic Complementation in Maize. *Front Plant Sci* 7:1–11. doi: 10.3389/fpls.2016.00414
- Khush GS (1994) Apomixis: exploiting hybrid vigor in rice.
- Khush GS (2005) What it will take to Feed 5.0 Billion Rice consumers in 2030. *Plant Mol Biol* 59:1–6. doi: 10.1007/s11103-005-2159-5
- Khush GS (1987) Rice breeding: Past, present and future. *J Genet* 66:195–216. doi: 10.1007/BF02927713
- Kikuchi S, Satoh K, Nagata T, et al (2003) Collection, Mapping, and Annotation of Over 28,000 cDNA Clones from japonica Rice.
- Kitajima TS, Kawashima SA, Watanabe Y (2004) The conserved kinetochore protein shugoshin protects centromeric cohesion during meiosis. *Nature* 427:510–517. doi: 10.1038/nature02312
- Kleckner N (2006) Chiasma formation: Chromatin/axis interplay and the role(s) of the synaptonemal complex. *Chromosoma* 115:175–194.
- Koltunow AM, Grossniklaus U (2003) Apomixis: a developmental perspective. *Annu Rev Plant Biol* 54:547–574. doi: 10.1146/annurev.arplant.54.110901.160842
- Koszul R, Kleckner N (2009) Dynamic chromosome movements during meiosis: a way to eliminate unwanted connections? *Trends Cell Biol*. 19:716–724.

- Kudo NR, Anger M, Peters AHFM, et al (2009) Role of cleavage by separase of the Rec8 kleisin subunit of cohesin during mammalian meiosis I. *J Cell Sci* 122:2686–2698. doi: 10.1242/jcs.035287
- Kugou K, Fukuda T, Yamada S, et al (2009) Rec8 guides canonical Spo11 distribution along yeast meiotic chromosomes. *Mol Biol Cell* 20:3064–76. doi: 10.1091/mbc.E08-12-1223
- Kumar R, Bourbon H-M, de Massy B (2010) Functional conservation of Mei4 for meiotic DNA double-strand break formation from yeasts to mice. *Genes Dev* 24:1266–1280. doi: 10.1101/gad.571710
- Kurata N, Nagamura Y, Yamamoto K, et al (1994) A 300 kilobase interval genetic map of rice including 883 expressed sequences. *Nat Genet* 8:365–372. doi: 10.1038/ng1294-365
- Kurzbauer M-T, Uanschou C, Chen D, Schlögelhofer P (2012) The Recombinases DMC1 and RAD51 Are Functionally and Spatially Separated during Meiosis in Arabidopsis.
- Kwon Y-I, Abe K, Endo M, et al (2013) DNA replication arrest leads to enhanced homologous recombination and cell death in meristems of rice *OsRecQ14* mutants. *BMC Plant Biol* 13:62. doi: 10.1186/1471-2229-13-62
- Laborte AG, Gutierrez MA, Balanza JG, et al (2017) RiceAtlas, a spatial database of global rice calendars and production. *Sci Data* 4:170074. doi: 10.1038/sdata.2017.74
- Lam I, Keeney S (2014) Mechanism and regulation of meiotic recombination initiation. *Cold Spring Harb Perspect Biol* 7:a016634. doi: 10.1101/cshperspect.a016634
- Lambing C, Franklin FCH, Wang C-JR (2017) Understanding and manipulating meiotic recombination in plants. *Plant Physiol* 173:pp.01530.2016. doi: 10.1104/pp.16.01530
- Lao JP, Cloud V, Huang C-C, et al (2013) Meiotic crossover control by concerted action of Rad51-Dmc1 in homolog template bias and robust homeostatic regulation. *PLoS Genet* 9:e1003978. doi: 10.1371/journal.pgen.1003978
- Lee CY, Conrad MN, Dresser ME (2012) Meiotic chromosome pairing is promoted by telomere-led chromosome movements independent of bouquet formation. *PLoS Genet* 8:e1002730. doi: 10.1371/journal.pgen.1002730
- Lee J (2013) Roles of cohesin and condensin in chromosome dynamics during mammalian meiosis. *J Reprod Dev* 59:431–6. doi: 10.1262/jrd.2013-068
- Lemmens BBLG, Johnson NM, Tijsterman M (2013) COM-1 Promotes Homologous Recombination during *Caenorhabditis elegans* Meiosis by Antagonizing Ku-Mediated Non-Homologous End Joining. *PLoS Genet* 9:e1003276. doi: 10.1371/journal.pgen.1003276
- Li H, Durbin R (2009) Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* 25:1754–1760. doi: 10.1093/bioinformatics/btp324
- Li J-Y, Wang J, Zeigler RS, et al (2014) The 3,000 rice genomes project: new opportunities and challenges for future rice research. *Gigascience* 3:8. doi: 10.1186/2047-217X-3-8
- Liu C, Li W, Zhong Y, et al (2015) Fine mapping of *qhir8* affecting in vivo haploid induction in maize. *Theor Appl Genet* 128:2507–2515. doi: 10.1007/s00122-015-2605-y
- Longhese MP, Bonetti D, Manfrini N, Clerici M (2010) Mechanisms and regulation of DNA end resection. *EMBO J* 29:2864–74. doi: 10.1038/emboj.2010.165
- Louis EJ, Borts RH (2003) Meiotic Recombination: Too Much of a Good Thing? *Curr Biol* 13:R953–R955. doi: 10.1016/j.cub.2003.11.040
- Lowe K, Wu E, Wang N, et al (2016) Morphogenic Regulators Baby boom and Wuschel Improve Monocot Transformation 4 5. *Plant Cell Adv Publ Publ*. doi: 10.1105/tpc.16.00124
- Lu Y, Ye X, Guo R, et al (2017) Genome-wide Targeted Mutagenesis in Rice Using the CRISPR/Cas9 System. *Mol Plant* 10:1242–1245. doi: 10.1016/j.molp.2017.06.007

- Luo D, Xu H, Liu Z, et al (2013a) A detrimental mitochondrial-nuclear interaction causes cytoplasmic male sterility in rice. *Nat Genet* 45:573–577. doi: 10.1038/ng.2570
- Luo Q, Li Y, Shen Y, Cheng Z (2014) Ten Years of Gene Discovery for Meiotic Event Control in Rice. *J Genet Genomics* 41:125–137. doi: 10.1016/j.jgg.2014.02.002
- Luo Q, Tang D, Wang M, et al (2013b) The role of OsMSH5 in crossover formation during rice meiosis. *Mol Plant* 6:729–42. doi: 10.1093/mp/sss145
- Luo Q, Tang D, Wang M, et al (2013c) The Role of OsMSH5 in Crossover Formation during Rice Meiosis. *Mol Plant* 6:729–742. doi: 10.1093/mp/sss145
- Lynn A, Soucek R, Börner GV (2007) ZMM proteins during meiosis: Crossover artists at work. *Chromosom Res* 15:591–605. doi: 10.1007/s10577-007-1150-1
- Ma X, Zhang Q, Zhu Q, et al (2015) A Robust CRISPR/Cas9 system for convenient, high-efficiency multiplex genome editing in monocot and dicot plants. *Mol Plant* 8:1274–1284. doi: 10.1016/j.molp.2015.04.007
- Macaisne N, Novatchkova M, Peirera L, et al (2008) SHOC1, an XPF Endonuclease-Related Protein, Is Essential for the Formation of Class I Meiotic Crossovers. *Curr Biol* 18:1432–1437. doi: 10.1016/j.cub.2008.08.041
- Mankouri HW, Hickson ID (2007) The RecQ helicase-topoisomerase III-Rmi1 complex: a DNA structure-specific “dissolvasome”? *Trends Biochem Sci* 32:538–46. doi: 10.1016/j.tibs.2007.09.009
- Marimuthu MP a, Jolivet S, Ravi M, et al (2011) Synthetic clonal reproduction through seeds. *Science* 331:876. doi: 10.1126/science.1199682
- Marston AL, Tham W-H, Shah H, Amon A (2004) A genome-wide screen identifies genes required for centromeric cohesion. *Science* 303:1367–70. doi: 10.1126/science.1094220
- Martini E, Diaz RL, Hunter N, Keeney S (2006) Crossover Homeostasis in Yeast Meiosis. *Cell* 126:285–295. doi: 10.1016/j.cell.2006.05.044
- McClosky B, Tanksley SD (2013) The impact of recombination on short-term selection gain in plant breeding experiments. *Theor Appl Genet* 126:2299–2312. doi: 10.1007/s00122-013-2136-3
- McCouch S, Wright MH, Tung C-W, et al (2016) Open access resources for genome-wide association mapping in rice. *Nat Commun* 7:10532. doi: 10.1038/ncomms10532
- McKenna A, Hanna M, Banks E, et al (2010) The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res* 20:1297–303. doi: 10.1101/gr.107524.110
- Meng X, Yu H, Zhang Y, et al (2017) Construction of a Genome-Wide Mutant Library in Rice Using CRISPR/Cas9. *Mol Plant* 10:1238–1241. doi: 10.1016/j.molp.2017.06.006
- Mercier R, Armstrong SJ, Horlow C, et al (2003) The meiotic protein SWI1 is required for axial element formation and recombination initiation in Arabidopsis. *Development* 130:3309–18.
- Mercier R, Jolivet S, Vezon D, et al (2005) Two Meiotic Crossover Classes Cohabit in Arabidopsis: One Is Dependent on MER3, whereas the Other One Is Not. *Curr Biol* 15:692–701. doi: 10.1016/j.cub.2005.02.056
- Mercier R, Mézard C, Jenczewski E, et al (2015) The Molecular Biology of Meiosis in Plants. *Annu Rev Plant Biol* 66:297–327. doi: 10.1146/annurev-arplant-050213-035923
- Mercier R, Vezon D, Bullier E, et al (2001) SWITCH1 (SWI1): a novel protein required for the establishment of sister chromatid cohesion and for bivalent formation at meiosis. *Genes Dev* 15:1859–71. doi: 10.1101/gad.203201
- Mézard C, Vignard J, Drouaud J, Mercier R (2007) The road to crossovers: plants have their say. *Trends Genet* 23:91–99. doi: 10.1016/j.tig.2006.12.007

- Miao C, Tang D, Zhang H, et al (2013a) Central region component1, a novel synaptonemal complex component, is essential for meiotic recombination initiation in rice. *Plant Cell* 25:2998–3009. doi: 10.1105/tpc.113.113175
- Miao J, Guo D, Zhang J, et al (2013b) Targeted mutagenesis in rice using CRISPR-Cas system. *Cell Res* 23:1233–1236. doi: 10.1038/cr.2013.123
- Mieulet D, Diévert A, Droc G, et al (2013) Reverse Genetics in Rice Using Tos17. In: *Methods in molecular biology* (Clifton, N.J.). pp 205–221
- Mieulet D, Jolivet S, Rivard M, et al (2016) Turning rice meiosis into mitosis. *Nat Publ Gr* 1–13. doi: 10.1038/cr.2016.117
- Miura K, Ashikari M, Matsuoka M (2011) The role of QTLs in the breeding of high-yielding rice. *Trends Plant Sci* 16:319–26. doi: 10.1016/j.tplants.2011.02.009
- Moses MJ (2006) A new chromosome structure is revealed. *Chromosoma* 115:152–154. doi: 10.1007/s00412-006-0051-y
- Neale MJ, Pan J, Keeney S (2005) Endonucleolytic processing of covalent protein-linked DNA double-strand breaks. *Nature* 436:1053–7. doi: 10.1038/nature03872
- Nonomura K-I (2006) PAIR2 is essential for homologous chromosome synapsis in rice meiosis I. *J Cell Sci*. doi: 10.1242/jcs.02736
- Nonomura K-I, Nakano M, Fukuda T, et al (2004) The novel gene HOMOLOGOUS PAIRING ABERRATION IN RICE MEIOSIS1 of rice encodes a putative coiled-coil protein required for homologous chromosome pairing in meiosis. *Plant Cell* 16:1008–20. doi: 10.1105/tpc.020701
- Oh SD, Lao JP, Taylor AF, et al (2008) RecQ helicase, Sgs1, and XPF family endonuclease, Mus81-Mms4, resolve aberrant joint molecules during meiotic recombination. *Mol Cell* 31:324–36. doi: 10.1016/j.molcel.2008.07.006
- Olmedo-Monfil V, Durán-Figueroa N, Arteaga-Vázquez M, et al (2010) Control of female gamete formation by a small RNA pathway in Arabidopsis. *Nature* 464:628–632. doi: 10.1038/nature08828
- Ouyang Y, Liu Y-G, Zhang Q (2010a) Hybrid sterility in plant: stories from rice. *Curr Opin Plant Biol* 13:186–192. doi: 10.1016/j.pbi.2010.01.002
- Ouyang Y, Liu YG, Zhang Q (2010b) Hybrid sterility in plant: stories from rice. *Curr. Opin. Plant Biol.*
- Panizza S, Mendoza MA, Berlinger M, et al (2011) Spo11-Accessory Proteins Link Double-Strand Break Sites to the Chromosome Axis in Early Meiotic Recombination. *Cell* 146:372–383. doi: 10.1016/j.cell.2011.07.003
- Pâques F, Haber JE (1999) Multiple pathways of recombination induced by double-strand breaks in *Saccharomyces cerevisiae*. *Microbiol Mol Biol Rev* 63:349–404.
- Pawlowski WP, Wang C-JR, Golubovskaya IN, et al (2009) Maize AME1OTIC1 is essential for multiple early meiotic processes and likely required for the initiation of meiosis. *Proc Natl Acad Sci* 106:3603–3608. doi: 10.1073/pnas.0810115106
- Penkner AM, Fridkin A, Gloggnitzer J, et al (2009) Meiotic chromosome homology search involves modifications of the nuclear envelope protein Matefin/SUN-1. *Cell* 139:920–33. doi: 10.1016/j.cell.2009.10.045
- Proost S, Van Bel M, Vaneechoutte D, et al (2015) PLAZA 3.0: an access point for plant comparative genomics. *Nucleic Acids Res* 43:D974–81. doi: 10.1093/nar/gku986
- Raghavan C, Mauleon R, Lacorte V, et al (2017) Approaches in Characterizing Genetic Structure and Mapping in a Rice Multiparental Population. *G3 (Bethesda)* 7:1721–1730. doi: 10.1534/g3.117.042101
- Ravi M, Chan SWL (2010) Haploid plants produced by centromere-mediated genome

- elimination. *Nature* 464:615–8. doi: 10.1038/nature08842
- Ravi M, Marimuthu MPA, Siddiqi I (2008) Gamete formation without meiosis in *Arabidopsis*. *Nature* 451:1121–1124. doi: 10.1038/nature06557
- Ray DK, Mueller ND, West PC, et al (2013) Yield Trends Are Insufficient to Double Global Crop Production by 2050. *PLoS One* 8:e66428. doi: 10.1371/journal.pone.0066428
- Robert T, Nore A, Brun C, et al (2016) The TopoVIB-Like protein family is required for meiotic DNA double-strand break formation. *Science* (80-) 351:943–949. doi: 10.1126/science.aad5309
- Robine N, Uematsu N, Amiot F, et al (2007) Genome-wide redistribution of meiotic double-strand breaks in *Saccharomyces cerevisiae*. *Mol Cell Biol* 27:1868–80. doi: 10.1128/MCB.02063-06
- Rockmill B, Voelkel-Meiman K, Roeder GS (2006) Centromere-proximal crossovers are associated with precocious separation of sister chromatids during meiosis in *Saccharomyces cerevisiae*. *Genetics* 174:1745–54. doi: 10.1534/genetics.106.058933
- Ronceret A, Vielle-Calzada J-P (2015) Meiosis, unreduced gametes, and parthenogenesis: implications for engineering clonal seed formation in crops. *Plant Reprod* 28:91–102. doi: 10.1007/s00497-015-0262-6
- Ross KJ, Fransz P, Armstrong SJ, et al (1997) Cytological characterization of four meiotic mutants of *Arabidopsis* isolated from T-DNA-transformed lines. *Chromosom Res* 5:551–559. doi: 10.1023/A:1018497804129
- Rutger JN (1992) Searching for apomixis in rice. In: *Proceedings of Apomixis Workshop*. U.S. Department of Agriculture, Beltsville, MD, USA. USA, pp 36–39
- Sadhu MJ, Bloom JS, Day L, Kruglyak L (2016) CRISPR-directed mitotic recombination enables genetic mapping without crosses. *Science* (80-) 352:1113–1116. doi: 10.1126/science.aaf5124
- Sakuno T, Watanabe Y (2009) Studies of meiosis disclose distinct roles of cohesion in the core centromere and pericentromeric regions. *Chromosom. Res.* 17:239–249.
- Sallaud C, Gay C, Larmande P, et al (2004) High throughput T-DNA insertion mutagenesis in rice: A first step towards in silico reverse genetics. *Plant J* 39:450–464. doi: 10.1111/j.1365-313X.2004.02145.x
- Sang T, Ge S (2013) Understanding rice domestication and implications for cultivar improvement. *Curr Opin Plant Biol* 16:139–146. doi: 10.1016/j.pbi.2013.03.003
- Sarno R, Vicq Y, Uematsu N, et al (2017) Programming sites of meiotic crossovers using Spo11 fusion proteins. *Nucleic Acids Res.* doi: 10.1093/nar/gkx739
- Sasaki T, Burr B (2000) International Rice Genome Sequencing Project: the effort to completely sequence the rice genome. *Curr Opin Plant Biol* 3:138–142. doi: 10.1016/S1369-5266(99)00047-3
- Sato Y, Takehisa H, Kamatsuki K, et al (2013) RiceXPro Version 3.0: expanding the informatics resource for rice transcriptome. *Nucleic Acids Res* 41:D1206–D1213. doi: 10.1093/nar/gks1125
- Savidan Y (2001) Transfer of apomixis through wide crosses. In: Savidan Y, J C, T D (eds) *The flowering of apomixis: From mechanisms to genetic engineering.*, CIMMYT, IR. Mexico, pp 153–167
- Sax K (1932) The Cytological Mechanism for Crossin Over. *Proc. Sixth Int. Congr. Genet.* 256–273.
- Schatz MC, Maron LG, Stein JC, et al (2014) Whole genome de novo assemblies of three divergent strains of rice, *Oryza sativa*, document novel gene space of aus and indica. *Genome Biol* 15:506. doi: 10.1186/PREACCEPT-2784872521277375

- Schwacha A, Kleckner N (1997) Interhomolog bias during meiotic recombination: meiotic functions promote a highly differentiated interhomolog-only pathway. *Cell* 90:1123–35.
- Séguéla-Arnaud M, Crismani W, Larchevêque C, et al (2015) Multiple mechanisms limit meiotic crossovers: TOP3 α and two BLM homologs antagonize crossovers in parallel to FANCM. *Proc Natl Acad Sci* 112:4713–4718. doi: 10.1073/pnas.1423107112
- Ségurel L, Leffler EM, Przeworski M (2011) The Case of the Fickle Fingers: How the PRDM9 Zinc Finger Protein Specifies Meiotic Recombination Hotspots in Humans. *PLoS Biol* 9:e1001211. doi: 10.1371/journal.pbio.1001211
- Sequencing Project IRG (2005) The map-based sequence of the rice genome. *Nature* 436:793–800. doi: 10.1038/nature03895
- Serra H, Da Ines O, Degroote F, et al (2013) Roles of XRCC2, RAD51B and RAD51D in RAD51-Independent SSA Recombination. *PLoS Genet* 9:e1003971. doi: 10.1371/journal.pgen.1003971
- Serra H, Lambing C, Griffin CH, et al (2017) Massive crossover elevation via combination of HEI10 and recq4a recq4b during Arabidopsis meiosis. *bioRxiv*. doi: 10.1101/159764
- Shao T, Tang D, Wang K, et al (2011) OsREC8 is essential for chromatid cohesion and metaphase I monopolar orientation in rice meiosis. *Plant Physiol* 156:1386–96. doi: 10.1104/pp.111.177428
- Shen Y, Tang D, Wang K, et al (2012) ZIP4 in homologous chromosome synapsis and crossover formation in rice meiosis. *J Cell Sci* 125:2581–2591. doi: 10.1242/jcs.090993
- Si W, Yuan Y, Huang J, et al (2015) Widely distributed hot and cold spots in meiotic recombination as shown by the sequencing of rice F₂ plants. *New Phytol* 206:1491–1502. doi: 10.1111/nph.13319
- Singh M, Goel S, Meeley RB, et al (2011) Production of Viable Gametes without Meiosis in Maize Deficient for an ARGONAUTE Protein. *Plant Cell* 23:443–458. doi: 10.1105/tpc.110.079020
- Smith GR, Boddy MN, Shanahan P, Russell P (2003) Fission Yeast Mus81·Eme1 Holliday Junction Resolvase Is Required for Meiotic Crossing over but Not for Gene Conversion. *Genetics* 165:2289–2293.
- Sommermeier V, Béneut C, Chaplais E, et al (2013) Spp1, a Member of the Set1 Complex, Promotes Meiotic DSB Formation in Promoters by Tethering Histone H3K4 Methylation Sites to Chromosome Axes. *Mol Cell* 49:43–54. doi: 10.1016/j.molcel.2012.11.008
- Spindel J, Wright M, Chen C, et al (2013) Bridging the genotyping gap: using genotyping by sequencing (GBS) to add high-density SNP markers and new value to traditional bi-parental mapping and breeding populations. *Theor Appl Genet* 126:2699–2716. doi: 10.1007/s00122-013-2166-x
- Stacey NJ, Kuromori T, Azumi Y, et al (2006) Arabidopsis SPO11-2 functions with SPO11-1 in meiotic recombination. *Plant J* 48:206–16. doi: 10.1111/j.1365-313X.2006.02867.x
- Sturtevant AH (1915) The behavior of the chromosomes as studied through linkage. *Z Indukt Abstamm Vererbungsl* 13:234–287. doi: 10.1007/BF01792906
- Sundaresan V, Alandete-Saez M (2010) Pattern formation in miniature: the female gametophyte of flowering plants. *Development* 137:179–189. doi: 10.1242/dev.030346
- Sung P, Petukhova G, Stratton S (1998) Catalysis of homologous DNA pairing by yeast Rad51 and Rad54 proteins. *Nature* 393:91–94. doi: 10.1038/30037
- Szostak JW, Orr-Weaver TL, Rothstein RJ, Stahl FW (1983) The double-strand-break repair model for recombination. *Cell* 33:25–35.
- Takagi H, Tamiru M, Abe A, et al (2015) MutMap accelerates breeding of a salt-tolerant rice cultivar. *Nat Biotechnol* 33:445. doi: 10.1038/nbt.3188

- Takagi H, Uemura A, Yaegashi H, et al (2013) MutMap-Gap: whole-genome resequencing of mutant F2 progeny bulk combined with *de novo* assembly of gap regions identifies the rice blast resistance gene *Pii*. *New Phytol* 200:276–283. doi: 10.1111/nph.12369
- Talbert PB, Henikoff S (2010) Centromeres convert but don't cross. *PLoS Biol* 8:e1000326. doi: 10.1371/journal.pbio.1000326
- Terada R, Urawa H, Inagaki Y, et al (2002) Efficient gene targeting by homologous recombination in rice. doi: 10.1038/nbt737
- Till BJ, Cooper J, Tai TH, et al (2007) Discovery of chemically induced mutations in rice by TILLING. *BMC Plant Biol* 7:19. doi: 10.1186/1471-2229-7-19
- Toriyama K, Arimoto Y, Uchimiya H, Hinata K (1988) Transgenic Rice Plants After Direct Gene Transfer into Protoplasts. *Nat Biotechnol* 6:1072–1074. doi: 10.1038/nbt0988-1072
- Uanschou C, Ronceret A, Von Harder M, et al (2013) Sufficient Amounts of Functional HOP2/MND1 Complex Promote Interhomolog DNA Repair but Are Dispensable for Intersister DNA Repair during Meiosis in Arabidopsis. *Plant Cell* 25:4924–4940. doi: 10.1105/tpc.113.118521
- United Nations (2015) World Population Prospects. New York
- van Brabant AJ, Ye T, Sanz M, et al (2000) Binding and melting of D-loops by the Bloom syndrome helicase. *Biochemistry* 39:14617–25.
- Varas J, Graumann K, Osman K, et al (2015) Absence of SUN1 and SUN2 proteins in Arabidopsis thaliana leads to a delay in meiotic progression and defects in synapsis and recombination. *Plant J* 81:329–346. doi: 10.1111/tjp.12730
- Verma DPS (Desh PS., Hong Z (2008) Cell division control in plants. Springer
- Vikram P, Swamy BPM, Dixit S, et al (2015) Drought susceptibility of modern rice varieties: an effect of linkage of drought tolerance with undesirable traits. *Sci Rep* 5:14799. doi: 10.1038/srep14799
- Vincenten N, Kuhl L-M, Lam I, et al (2015) The kinetochore prevents centromere-proximal crossover recombination during meiosis. *Elife*. doi: 10.7554/eLife.10850
- Vrielynck N, Chambon A, Vezon D, et al (2016) A DNA topoisomerase VI-like complex initiates meiotic recombination. *Science* (80-) 351:939–943. doi: 10.1126/science.aad5196
- Wang K, Tang D, Wang M, et al (2009) MER3 is required for normal meiotic crossover formation, but not for presynaptic alignment in rice. *J Cell Sci* 122:2055–2063. doi: 10.1242/jcs.049080
- Wang K, Wang C, Liu Q, et al (2015a) Increasing the Genetic Recombination Frequency by Partial Loss of Function of the Synaptonemal Complex in Rice. *Mol Plant* 8:1295–1298. doi: 10.1016/j.molp.2015.04.011
- Wang K, Wang M, Tang D, et al (2012) The role of rice HEI10 in the formation of meiotic crossovers. *PLoS Genet* 8:e1002809. doi: 10.1371/journal.pgen.1002809
- Wang M, Wang K, Tang D, et al (2010a) The central element protein ZEP1 of the synaptonemal complex regulates the number of crossovers during meiosis in rice. *Plant Cell* 22:417–30. doi: 10.1105/tpc.109.070789
- Wang M, Wang K, Tang D, et al (2010b) The central element protein ZEP1 of the synaptonemal complex regulates the number of crossovers during meiosis in rice. *Plant Cell* 22:417–30. doi: 10.1105/tpc.109.070789
- Wang S, Zickler D, Kleckner N, Zhang L (2015b) Meiotic crossover patterns: obligatory crossover, interference and homeostasis in a single process. *Cell Cycle* 14:305–14. doi: 10.4161/15384101.2014.991185
- Whitton J, Sears C, Baack E, Otto S (2008) The Dynamic Nature of Apomixis in the Angiosperms. *Int J Plant Sci* 169:169–182. doi: 10.1086/523369

- Wu J, Maehara T, Shimokawa T, et al (2002) A comprehensive rice transcript map containing 6591 expressed sequence tag sites. *Plant Cell* 14:525–35. doi: 10.1105/TPC.010274
- Wu J, Mizuno H, Hayashi-Tsugane M, et al (2003) Physical maps and recombination frequency of six rice chromosomes. *Plant J* 36:720–730. doi: 10.1046/j.1365-313X.2003.01903.x
- Wu Z, Ji J, Tang D, et al (2015) OsSDS is essential for DSB formation in rice meiosis. *Front Plant Sci* 6:21. doi: 10.3389/fpls.2015.00021
- Yelina NE, Choi K, Chelysheva L, et al (2012) Epigenetic remodeling of meiotic crossover frequency in *Arabidopsis thaliana* DNA methyltransferase mutants. *PLoS Genet* 8:e1002844. doi: 10.1371/journal.pgen.1002844
- Yin X, Biswal AK, Dionora J, et al (2017) CRISPR-Cas9 and CRISPR-Cpf1 mediated targeting of a stomatal developmental gene EPFL9 in rice. *Plant Cell Rep* 36:745–757. doi: 10.1007/s00299-017-2118-z
- Youds JL, Boulton SJ (2011) The choice in meiosis - defining the factors that influence crossover or non-crossover formation. *J Cell Sci*. doi: 10.1242/jcs.074427
- Yu J, Hu S, Wang J, et al (2002) A Draft Sequence of the Rice Genome (*Oryza sativa* L. ssp. *indica*).
- Yuan J, Chen J (2013) FIGNL1-containing protein complex is required for efficient homologous recombination repair. *Proc Natl Acad Sci U S A* 110:10640–5. doi: 10.1073/pnas.1220662110
- Yuan W, Li X, Chang Y, et al (2009) Mutation of the rice gene PAIR3 results in lack of bivalent formation in meiosis. *Plant J* 59:303–315. doi: 10.1111/j.1365-313X.2009.03870.x
- Zakharyevich K, Tang S, Ma Y, Hunter N (2012) Delineation of joint molecule resolution pathways in meiosis identifies a crossover-specific resolvase. *Cell* 149:334–47. doi: 10.1016/j.cell.2012.03.023
- Zhang P, Zhang Y, Sun L, et al (2017) The Rice AAA-ATPase OsFIGNL1 Is Essential for Male Meiosis. *Front Plant Sci*. doi: 10.3389/fpls.2017.01639
- Zhang Q, Wing RA (eds) (2013) *Genetics and Genomics of Rice*. Springer New York, New York, NY
- Zhou H, Liu B, Weeks DP, et al (2014) Large chromosomal deletions and heritable small genetic changes induced by CRISPR/Cas9 in rice. *Nucleic Acids Res* 42:10903–10914. doi: 10.1093/nar/gku806
- Zhu Z, Chung W-H, Shim EY, et al (2008) Sgs1 Helicase and Two Nucleases Dna2 and Exo1 Resect DNA Double-Strand Break Ends. *Cell* 134:981–994. doi: 10.1016/j.cell.2008.08.037
- Zickler D, Kleckner N (1999) Meiotic Chromosomes: Integrating Structure and Function. *Annu Rev Genet* 33:603–754. doi: 10.1146/annurev.genet.33.1.603
- Zickler D, Kleckner N (2015) Recombination, Pairing, and Synapsis of Homologs during Meiosis. *Cold Spring Harb Perspect Biol*. doi: 10.1101/cshperspect.a016626

ANNEXES

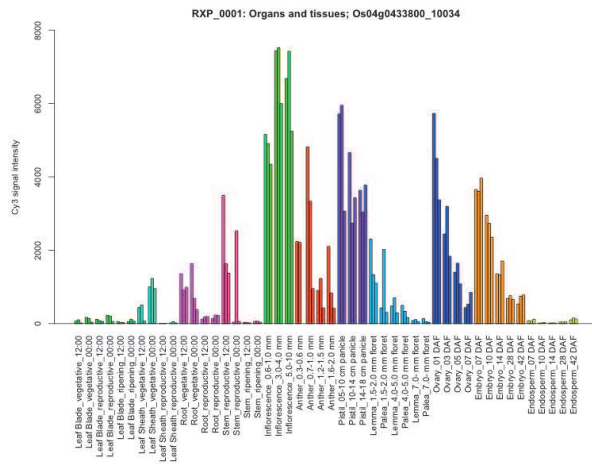
Gene chez A. thaliana	Gene	Locus ID oryza sativa	Plant Name	FST	Position	Collection	Mutagen	Fond génétique	Disponibilité	Code interne	BP- Name	BP- Sequence	LP- Name	LP- Sequence	RP- Name	RP- Sequence	PCR WT (pb)	PCR Mut (pb)	Présence de l'insertion	Nom de l'allèle
AT1G35530 <i>FANCM</i>			AKGE05	GS901065	intron	OTL	Tos17	Nipponbare		AKG	AKG-BP	AGTCGCTGATTTCTTCACCAAGG	AKG-LP	CTTGTGCAGGAGTGGACAA	AKG-RP	CAACTGTGGGCTCTGAAATA			non confirmé	
			APCC05	CU317967	exon	OTL	Tos17	Nipponbare	indisponible	APC	APC-BP	AGTCGCTGATTTCTTCACCAAGG	APC-LP	ATCCATCAGCACGGAATAGG	APC-RP	CCAAACTGGTTTGTGGCTTGT			x	
			A46543	5A-00166R	exon	Postech	T-DNA	Dongjin		A4	A46543-BP	TCGTTAAACTGCTCGCACAGC	A46543-LP	CAACGTA TGGGAAGGACTGG	A46543-RP	GGTGGGAAGAAGACCAACCA	1094	758	confirmé	Oslancm-2
				03211CJ39	intron			Zhonghua11		03211	Hygè	GTCTGGACCGATGGCTGTGTAGAAG	032-LP3	caactgacattggggctct	032-RP3	gatttggcctcaagaatca			confirmé	
				05211HO51	exon	RMD_T-DNA	T-DNA	Zhonghua11		05211	Hygè	GTCTGGACCGATGGCTGTGTAGAAG	052-LP	CCTTTGCTCCCTCACTGTGT	052-RP	TGGAAGGACCTGCTGTATGT			PCR not OK	
			AQS607	SA77G04	intron	OTL	T-DNA	Nipponbare		AQS	AQS-BP	GTCTGGACCGATGGCTGTGTAGAAG	AQS-LP1	aggatcctggggagtg	AQS-RP	TGTAAACATCCGATCAGTGTGC	588	798	confirmé	Oslancm-1
				NF6822_0_7	exon	NIAS	Tos17													
				NE8036_0_5	intron	NIAS	Tos17													
			T06792T	AG023179	exon	NIAS	Tos17		indisponible	x										
			AFUB07	CL519229	promoter	OTL	T-DNA	Nipponbare		AFU	Re-O4-BP1	GTCTGGACCGATGGCTGTGTAGAAG	Re-O4-LP1	ACCTAAGGGGGTGAAGCAAG	Re-O4-RP1	ATCTAAACAGGGTCCGAGCA			PCR not OK	Osnecq4-2
AT1G10930 <i>RECQ4</i>			AUFG12	GS997606	exon	OTL	T-DNA	Nipponbare		AUF	Re-O4-BP2	TCGTTAAACTGCTCGCACAGC	Re-O4-LP2	TCGATGAATCTGTCAGTTCCA	Re-O4-RP2	ACATGCGCTACGGGAATAT	1013	890	confirmé	Osnecq4-3
			AJYH02	GS901180	intron	OTL	Tos17	Nipponbare	indisponible	x										
			ASCG10	CU316652	exon	OTL	Tos17	Nipponbare		ASC	Re-O4-BP3	AGTCGCTGATTTCTTCACCAAGG	Re-O4-LP3	AGTCTGGAAGCTGCAGAGT	Re-O4-RP3	GCAGAGTTAACCCAGATGA			non confirmé	
			A22224	3A-10177L	exon	Postech	T-DNA													
			A26155	3A-12928L	exon	Postech	T-DNA													
			A09780	3A-01037R	intron	Postech	T-DNA													
			A12797	3A-03503L	exon	Postech	T-DNA													
			A18914	3A-07836R	intron	Postech	T-DNA													
			A32312	3A-17083R	exon	Postech	T-DNA													
				03211EH07	intron			Zhonghua11												
				03211EH74	intron	RMD_T-DNA	T-DNA													
				04211DO63	promoter	RMD_T-DNA	T-DNA													
				04211OZ62	exon	RMD_T-DNA	T-DNA													
				05211AZ93	promoter	RMD_T-DNA	T-DNA													
			M0015299	E1010630	exon	TRIM	T-DNA													
			M0015298	E1010629	exon	TRIM	T-DNA													

Annexe 1 : Liste des mutants identifiés dans les gènes *FANCM* et *RECQ4*.

GENE	Insertion	Primer	séq	Tm	Taille
RECQI4	AUF	E-REC-1F	GCTTTGCAATGCCTTAGGAC	60	
		E-REC-2F	AGGAAATCAGCAACCGAATG	60	
		E-REC-1R	TAGGCTATTCTGGCGGACTG	60	
		E-REC-2R	CGCAGGTAAAACCTACCAG	60	
		Q-REC-7F	CTTCTCCGTCACCAACACCT	60	226bp (Q-RT)
		E-REC-2R	CGCAGGTAAAACCTACCAG	60	
	3503	E-REC-3F	CCAAGACCAGATCATGCA'TTA	60	
		E-REC-4F	CGCCTGAAAAGATAGCCAAG	60	
		E-REC-3R	GTTGATCCTATCCTTGCTCCAC	60	
		E-REC-4R	TGATGCCGCATTCTTTATTATG	60	
		Q-REC-8F	AAAACAGAAGTTCCCGCAGA	60	224bp (Q-RT)
		E-REC-4R	TGATGCCGCATTCTTTATTATG	60	
	ZMM11	E-REC-5F	CATTCAACAGGTTTCGCATGT	60	
		E-REC-6F	TAAAGCCAAGGGTCAACTGG	60	
		E-REC-5R	CCTGGGTGGGATTCTATTT	60	
		E-REC-6R	TGTCCATAAGTATGAGCATGG	60	
	3'UTR du gène (primer Q-PCR)	Q-Recq4-F	CTTCTCCGTCACCAACACCT	60	
		Q-Recq4-R	TGCCTTGACGTGCTTTGC	60	
FANCM	AQS ou et A4	E-FAN-1F	GGTATCAAGGCAGCACATCA	60	
		E-FAN-2F	GCACATCAGTCAATCGAAGC	60	
		E-FAN-3F	CAACGTATGGGAAGGACTGG	60	
		E-FAN-4F	GTGATCGTCGCAGATTGAA	60	
		E-FAN-1R	ATCCATCAGCACCGAATAGG	60	
		E-FAN-2R	ATCAGCAGGTCCTTCCAAAG	60	
		E-FAN-3R	CCTTAATGCGGAGATGTTGG	60	
		E-FAN-4R	GGCATCAAAACAAACCACAA	60	
	3'UTR du gène (primer Q-PCR)	fancm-QF	AAGGGATCCATCAACCACAC	60	
		fancm-QR	TGGTGATGACTCACAATGG	60	
OsEXP	Os07g02340	Q-Exp-F	AGGAACATGGAGAACAAGG	60	
		Q-Exp-R	CAGAGGTGGTGCAGATGAAA	60	
		Q-Exp-F	AGGAACATGGAGAACAAGG	60	
		Q-Exp-R	CAGAGGTGGTGCAGATGAAA	60	
		ExpVerif-F	ATGGGCAGAAGTCGAAGATG	60	
		ExpVerif-R	TCTTGAGGTGAGGGAAGCAC	60	

	MIX 1	MIX 2	MIX 3	MIX 4	MIX 5	MIX 6	MIX 7	MIX 8	MIX 9	MIX 10	MIX 11	MIX 12
E-FAN-1F	X	X	X	X								
E-FAN-2F					X	X	X	X				
E-FAN-3F									X	X		
E-FAN-4F											X	X
E-FAN-1R	X				X				X		X	
E-FAN-2R		X				X				X		X
E-FAN-3R			X				X					
E-FAN-4R				X				X				
TAILLE AMPLI	1039	987	477	459	1027	975	465	447	556	504	441	389
Taille si AQN									1255			
Caractérisation AQS seul												
Caractérisation A4 seul												
Caractérisation AQS et A4												
	MIX 1	MIX 2	MIX 3	MIX 4	MIX 5	MIX 6	MIX 7	MIX 8	MIX 9	MIX 10	MIX 11	MIX 12
E-REC-1F	X	X										
E-REC-2F			X	X								
E-REC-1R	X		X									
E-REC-2R		X		X								
E-REC-3F					X	X						
E-REC-4F							X	X				
E-REC-3R					X		X					
E-REC-4R						X		X				
E-REC-5F									X	X		
E-REC-6F											X	X
E-REC-5R									X		X	
E-REC-6R										X		X
TAILLE AMPLI	530	507	448	425	658	515	521	378	459	399	375	315
Taille si AQN	704	681	544	521	2822	1029	2612	819	1575	1517		
Caractérisation AUF seul												
Caractérisation 3503 seul												
Caractérisation Z11 seul												

Annexe 2 : Liste des amorces utilisées pour la Q-RT-PCR (en haut) et tailles des fragments attendues en fonction des couples d'amorces (ci-contre).

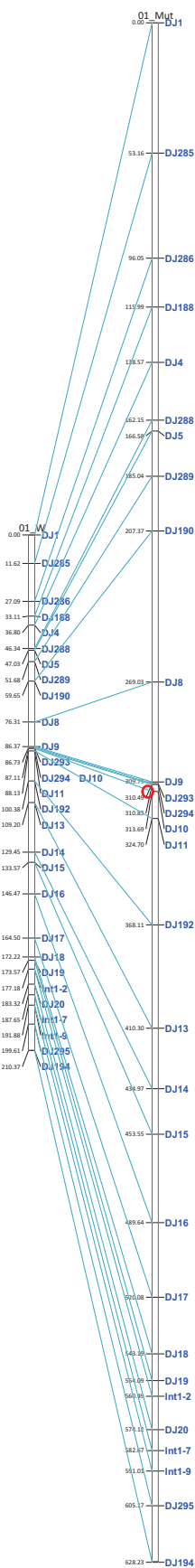


***OsRECQL4* :**

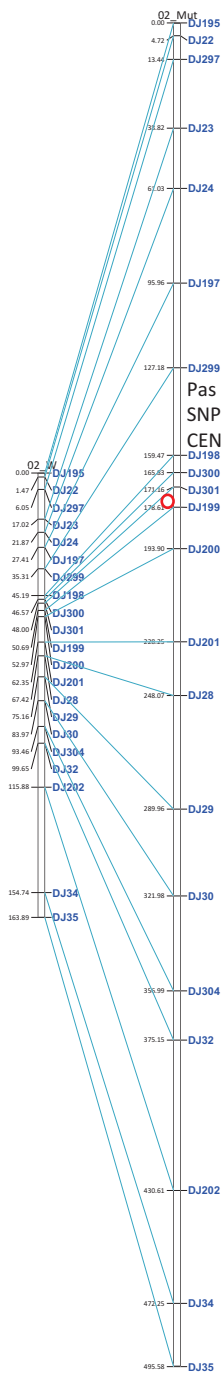
~~Pas de données~~

***OsFANCM* :**

Annexe 3 : Données d'expression (RiceXPro3.0) pour *OsRECQ14* (Os04g35420) et *OsFANCM* (Os11g07870).

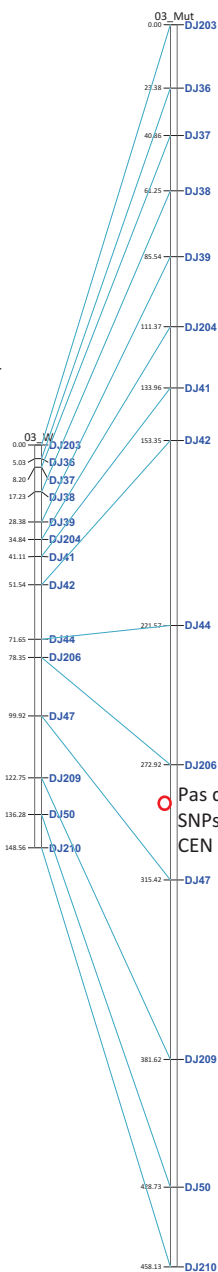


Mut
x2,98
Chr.1



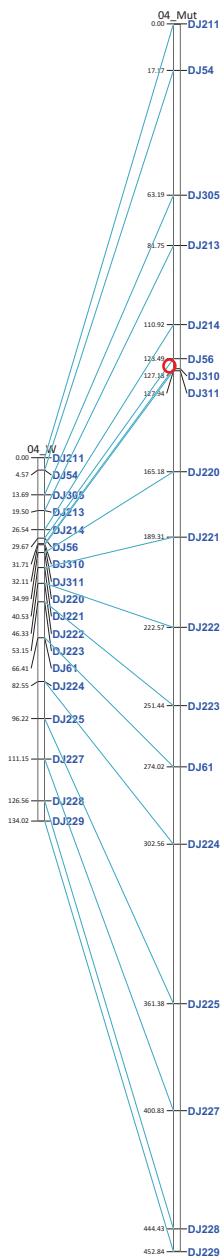
Pas de
SNPs péri-
CEN

Mut
x3.02
Chr.2



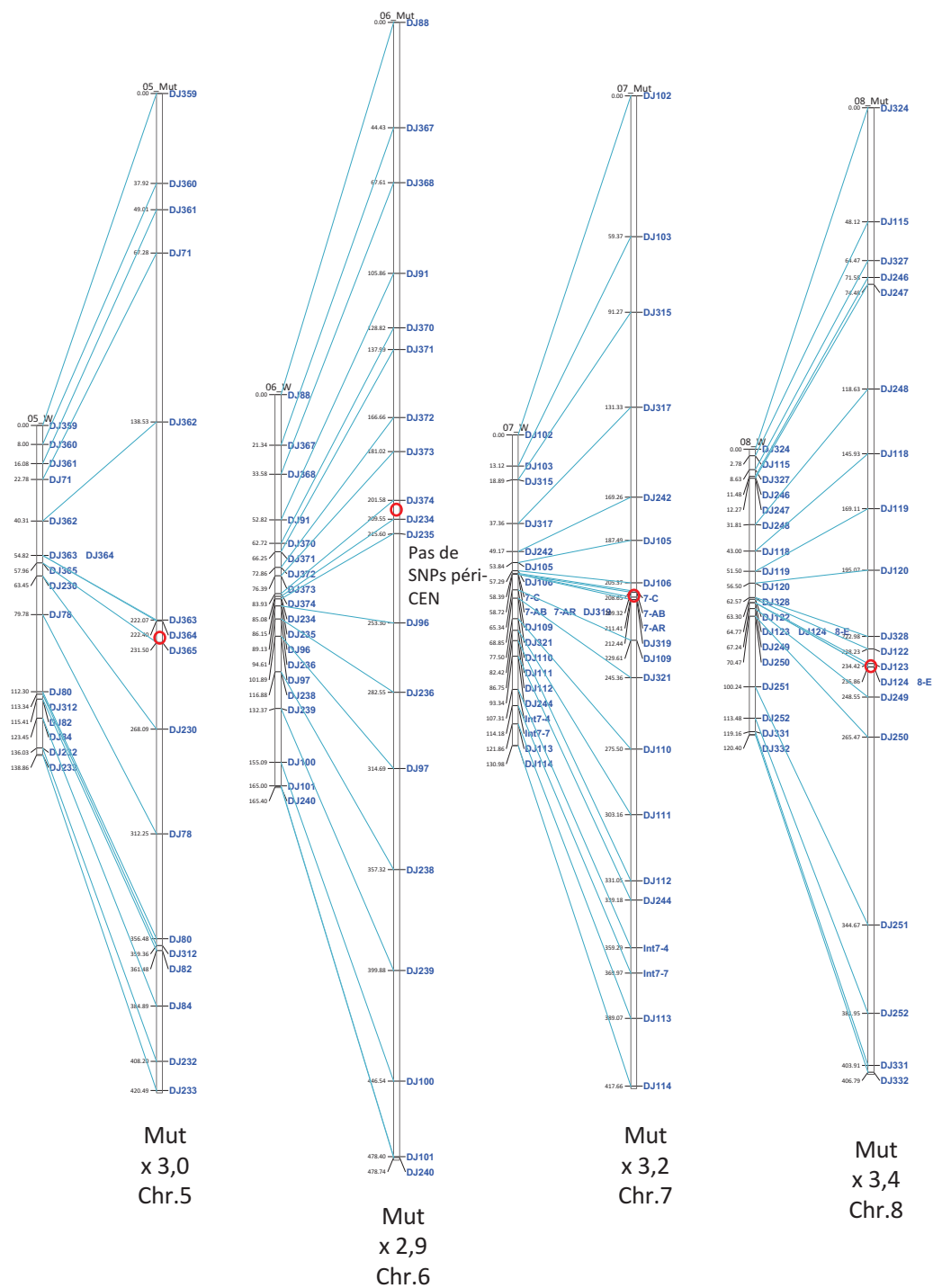
Pas de
SNPs péri-
CEN

Mut
x 3,08
Chr.3

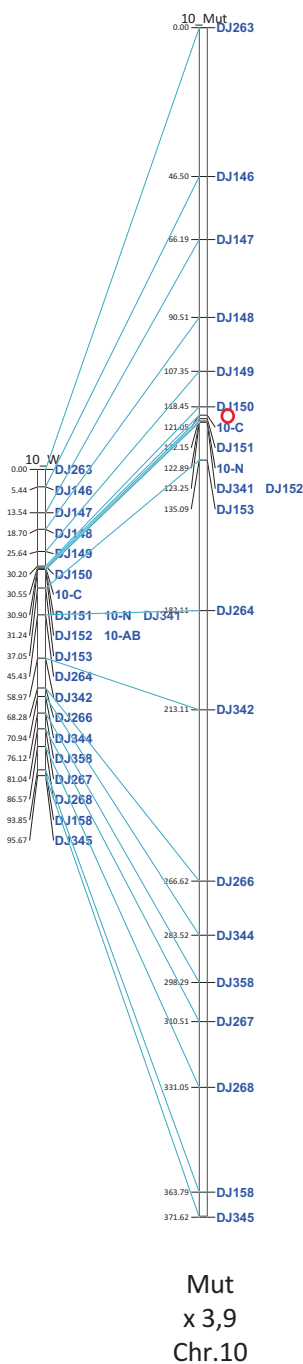
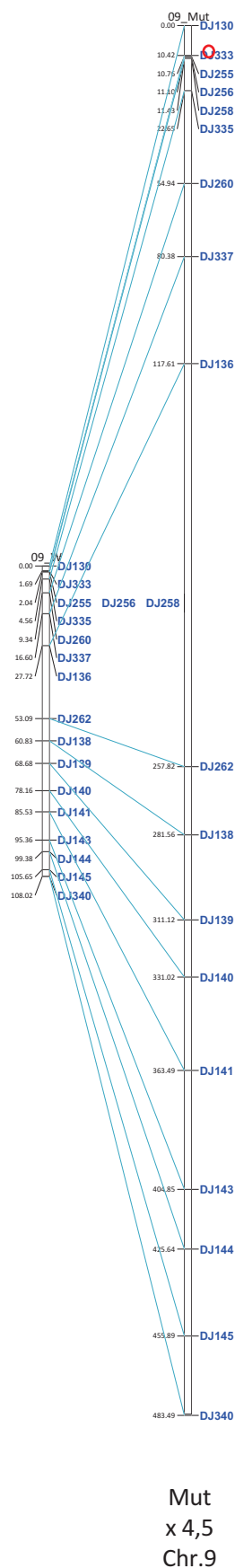


Mut
x3.37
Chr.4

Annexe 4 (1/3) : Cartes génétiques du sauvage et du mutant *Osrecq14* --/-- .



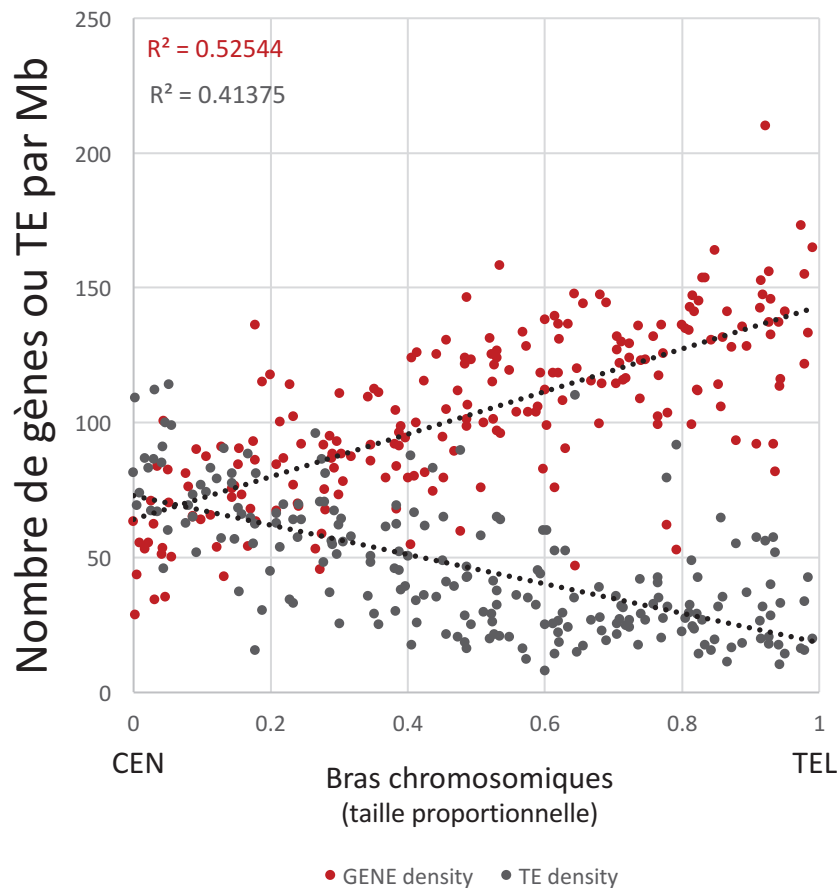
Annexe 4 (2/3) : Cartes génétiques du sauvage et du mutant *Osrecql4* --/-- .





Annexe 5 : Position des 92 marqueurs utilisés pour génotyper la descendance F1 *Osfancm*--/-- et *Osfancm* ++/++.

La position physique (en Mb) est indiqué à gauche du chromosome et le nom du Kasp[®] est noté à droite du chromosome.



Annexe 6 : Densité en gènes et en éléments transposables (TE).

Le nombre de gènes et de TE a été estimé pour chaque intervalle considéré pour la mesure de la recombinaison via les données du génome d'*Oryza sativa* cv Nipponbare (MSU7)

Les valeurs ont ensuite été reportées en fonction de la distance relative au centromère. (Régression linéaire).

Gène cible	Vecteur	Clonage Name	Target génomique loci (crRNA) + PAM (NGG)	crRNA-F	séquence = crRNA en Forward + 20 premiers nt du tracrRNA	crRNA-R	séquence = crRNA en revers + f20 derniers nt du promoteur U3	Nom du construit	code plantes T0
FANCM	PC53000-Cas9	ZMM-16	ATCCTACCGAGTGGCGCCAG-GGG	Fan-F1	ATCCTACCGAGTGGCGCCAGgttttagagctagaaatagc	Fan-R1	CTGGCGCCACTCGGTAGGATgccacggatcatctgcac	pC5300Cas9::FancM-gRNA1	ZMM16
		ZMM-17	ACCACGAGGCAGCGCAACT-TGG	Fan-F2	ACCACGAGGCAGCGCAACTgttttagagctagaaatagc	Fan-R2	AGTTCGCGCTGCCTCTGTGGTgccacggatcatctgcac	pC5300Cas9::FancM-gRNA2	ZMM17
		ZMM-18	AGGCAGCTCGTGTGCCGTTA-AGG	Fan-F3	AGGCAGCTCGTGTGCCGTTAgtttagagctagaaatagc	Fan-R3	TAACGGCACACGAGCTGCCTgccacggatcatctgcac		
RECQJ4		ZMM-10	AAAGCCAAGGGTCAACTGGT-CGG	Q4-F1	AAAGCCAAGGGTCAACTGGTgttttagagctagaaatagc	Q4-R1	ACCAAGTTGACCCCTTGGCTTgccacggatcatctgcac		
		ZMM-11	ACCAACACAAAGGCCTAATC-AGG	Q4-F2	ACCAACACAAAGGCCTAATCgttttagagctagaaatagc	Q4-R2	GATTAGGCCTTTGTGTTGGTgccacggatcatctgcac	pC5300Cas9::RecQ4-gRNA2	ZMM11
		ZMM-12	ATGCATCAGACGACTGAAAG-TGG	Q4-F3	ATGCATCAGACGACTGAAAGgttttagagctagaaatagc	Q4-R3	CTTTCAGTCGTCTGATGCATgccacggatcatctgcac	pC5300Cas9::RecQ4-gRNA3	ZMM12
PAIR1		ZMM-4	AGCAGCGCGCGTCGGCTTC-CGG	Pair-F1	AGCAGCGCGCGTCGGCTTCgttttagagctagaaatagc	Pair-R1	GAAGCCGACGCGCCGCTGCTgccacggatcatctgcac		
		ZMM-5	AACCGAGCACCAGTAGCAGC-TGG	Pair-F2	AACCGAGCACCAGTAGCAGCgttttagagctagaaatagc	Pair-R2	GCTGCTACTGGTGTCTGGTgccacggatcatctgcac		
		ZMM-6	AAGAGGGCCTGGAGATCCTG-CGG	Pair-F3	AAGAGGGCCTGGAGATCCTGgttttagagctagaaatagc	Pair-R3	CAGGATCTCCAGGCCCTCTTgccacggatcatctgcac		
OSD1		ZMM-1	ATGCCTGAAGTGAGAAATC-CGG	Osd-F1	ATGCCTGAAGTGAGAAATCgttttagagctagaaatagc	Osd-R1	GAATTTCTCACTTCAGGCATgccacggatcatctgcac		
		ZMM-2	AGCCGTGGCGGTCAAGCCGC-TGG	Osd-F2	AGCCGTGGCGGTCAAGCCGCgttttagagctagaaatagc	Osd-R2	GCGGCTTGACCGCCACGGCTgccacggatcatctgcac		
		ZMM-3	ACCACCAACTCAAAAGGCC-TGG	Osd-F3	ACCACCAACTCAAAAGGCCCgttttagagctagaaatagc	Osd-R3	GGGCTTTTGAGTTGGTGGTgccacggatcatctgcac		
REC8		ZMM-7	ACTCGCACCAGCTCCTCGCG-CGG	Rec-F1	ACTCGCACCAGCTCCTCGCGgttttagagctagaaatagc	Rec-R1	CGCGAGGAGCTGGTGCAGTgccacggatcatctgcac	pC5300Cas9::Rec8-gRNA1	ZMM7
		ZMM-8	AGGCTCCGCTCGGCCAGATA-TGG	Rec-F2	AGGCTCCGCTCGGCCAGATAgtttagagctagaaatagc	Rec-R2	TATCTGCCGAGCGGAGCCTTgccacggatcatctgcac		
		ZMM-9	AAGGCTCTCCGGAATTCTCA-TGG	Rec-F3	AAGGCTCTCCGGAATTCTCAgttttagagctagaaatagc	Rec-R3	TGAGAATTCGGGAGAGCCTTgccacggatcatctgcac	pC5300Cas9::Rec8-gRNA3	ZMM9
		communs		U3-F (EcoRI)	GAATTCAGGGATCTTTAAACATACG	tracr-R (XbaI)	TCTAGAAAAAGCACCGACTCGGTG		
OSD1	Miao	x		OSD1A-S	5'-GGCAAGTGAGAAATTCCGGCGGT-3'	OSD1A-AS	5'-AAACACCGCCGGAATTTCTCACTT-3'	crOSD1(cible1)-Miao2.0	OSDA
		x		OSD1B-S	5'-GGCAATAAGAACCACCCACCCGA-3'	OSD1B-AS	5'-AAACTCGGGTGGTGGGTTCTTTAT-3'	crOSD1(cible2)-Miao2.0	OSDB
RECQJ4	Miao2.0	ZMM11	ACCAACACAAAGGCCTAATC-AGG	Z11BTZ-F	TGTTACCAACACAAAGGCCTAATC	Z11BTZ-R	GATTAGGCCTTTGTGTTGGTAAA	Miao2.0-sgRNA-Z11	T141

Annexe 7 : Amorces correspondant aux crRNA et nom des construits CRISPR/Cas9 réalisées pour produire des plantes ciblant les gènes anti-CO.

Liste des gènes ciblés et des crRNA identifiés pour parvenir à éteindre l'expression des gènes cibles en utilisant soit le vecteur de transformation pC5300-Cas9 et un clonage par restriction soit les vecteurs de types Miao qui permettent une intégration du sgRNA par Gateway.

RESUME : MANIPULATION DE LA RECOMBINAISON CHEZ UNE PLANTE CULTIVEE, LE RIZ.

L'accroissement prévisible de la population mondiale ainsi que les conséquences du changement climatique obligent les sélectionneurs à créer de nouvelles variétés plus productives et plus résilientes. Les nouvelles combinaisons d'allèles favorables sont issues de la recombinaison génétique entre chromosomes homologues dont le siège est la prophase de première division de méiose. De récentes avancées chez la plante modèle *Arabidopsis* ont montré que l'inactivation de certains gènes permet de manipuler la méiose pour abolir ou au contraire augmenter très significativement la recombinaison. Les mécanismes de la méiose étant relativement bien conservés chez les eucaryotes, l'objectif de cette thèse était de transposer ces avancées chez une plante cultivée importante, le riz. Abolir la recombinaison méiotique permettrait de propager de façon clonale par grain des formules variétales hybrides F1 dont le rendement est de 20% supérieur à celui des lignées pures chez le riz mais dont les semences restent peu utilisées par les riziculteurs de subsistance. Les travaux réalisés dans une première partie de la thèse ont montré que le cumul de trois mutations *Ososd1*, *pair1* et *Osrec8*, permettait d'obtenir des gamètes clonaux diploïdes mâles et femelles. Le phénotype apoméiotique obtenu, appelé *MiMe* (*Mitosis instead of meiosis*) chez *Arabidopsis*, peut être utilisé pour tester différentes stratégies d'induction de la parthénogenèse afin de produire des grains formant des plantes diploïdes clonales apomictiques. Une optimisation du mécanisme permettrait d'envisager l'utilisation de l'apomixie pour fixer l'hétérosis dans les semences hybrides F1. Par ailleurs, une augmentation globale ou locale de la recombinaison méiotique est recherchée car elle permettrait de diminuer la taille des populations de sélection et de réduire la taille des segments chromosomiques introduits dans les variétés élite de riz. Nous avons montré dans une seconde partie, que la mutation du gène *OsRECQ4* codant pour une hélicase permet d'augmenter le taux de recombinaison d'un facteur de 3,3 fois faisant passer la taille de la carte génétique de 1670 ± 30 cM à 5537 ± 114 cM sans affecter la fertilité de la plante ni le déroulement de la méiose. Chez les plantes affectées dans la fonction d'une autre hélicase, *OsFANCM*, le taux de recombinaison a été également augmenté mais dans une moindre mesure (x 2,2). L'augmentation de la recombinaison s'opère sur l'ensemble des bras chromosomiques sauf au niveau des centromères. Ces résultats confirment ceux obtenus chez *A. thaliana* qui ont montré le rôle de régulateur négatif des crossing-overs (CO) des protéines RECQ4 et FANCM. Ces résultats ouvrent la voie à l'utilisation des gènes anti-COs pour augmenter de façon globale le nombre de recombinants dans les croisements chez le riz et sans doute chez les autres céréales.

ABSTRACT : MANIPULATION OF RECOMBINATION IN A CROP, RICE.

The forecasted increase of world population as well as the consequences of global climate change oblige plant breeders to develop new varieties that are both more productive and resilient. Novel combinations of favourable alleles are generated through genetic recombination between homologous chromosomes, which occurs during the prophase of the first division of meiosis. Recent advances in the model plant *Arabidopsis* have demonstrated that the inactivation of some genes allows meiosis manipulation resulting in either an abolishment or in contrast, a significant enhancement of meiotic recombination. The meiosis mechanisms being relatively conserved across eucaryotes, the overall objective of this thesis was to transfer these advances to a crop of crucial importance, rice. To abolish meiotic recombination would allow the clonal propagation by seeds of F hybrids, which exhibit a 20% yield enhancement compared to that of pure lines in rice but remain rarely used in subsistence farming. In a first part, we showed that rice plants cumulating 3 mutations inactivating *Ososd1*, *pair1* and *Osrec8*, formed clonal diploid male and female gametes. This apomeiotic phenotype, called *MiMe* (*Mitosis instead of meiosis*) in *Arabidopsis*, can serve as material to assay several strategies of parthenogenetic induction that would result in seed forming diploid clonal plants. Further optimization of the mechanisms would allow the use of apomixis to fix heterosis in hybrid seeds. Global and local enhancement of recombination is another desirable goal since it would allow a reduction in breeding population size and a downsizing of the introgressed chromosomal segments in elite plant materials. In a second part, we showed that mutation in the DNA helicase gene *OsRECQ4* conducted to a 3.3 fold increase of recombination and inflated the genetic map size from 1670 ± 30 to 5537 ± 114 cM, without altering plant fertility nor meiosis progression. Plants altered in a second DNA helicase, *OsFANCM*, exhibited a more modest 2.2 fold recombination enhancement. Recombination increase operated along the whole chromosome arms except at the centromere level. These results confirms the negative regulator role of RECQ4 and FANCM on crossing overs (CO), previously reported in *Arabidopsis*. These results pave the way to the use of anti-CO genes to enhance recombinant recovery in crosses of rice and possibly of other cereals.