AGROCAMPUS OUEST

CFR Angers

2 rue André Le Nôtre

49045 Angers

INSTITUT NATIONAL DE LA RECHERCHE AGRONOMIQUE

INRA Le Rheu

Domaine de la Motte au Vicomte BP 35327

35653 LE RHEUX CEDEX

Mémoire de Fin d'Études

Diplôme d'Ingénieur de l'Institut Supérieur des Sciences Agronomiques, Agroalimentaires, Horticoles et du Paysage

Année universitaire : 2012-2013

Spécialisation ou option : Amélioration des plantes

La distribution des évènements de Crossing-over, un paramètre modulable chez les Brassicas

Par : Alexandre PELE

*Volet à renseigner par l'enseignant responsable de l'option/spécialisation** Ou son représentant	Bon pour dépôt (version définitive) ❑
Date : …./…/… Signature	Autorisation de diffusion : Oui ❑ Non❑

Devant le jury *:* Soutenu à Rennes, le 12 Septembre 2013

Sous la présidence de : Mme Maria MANZANARES-DAULEUX

Maître de stage : Mme Anne-Marie CHEVRE

Enseignant référent : Mme Maria MANZANARES-DAULEUX

Autres membres du jury : Mr Frédéric LECERF et Mme Mélanie JUBAULT

"*Les analyses et les conclusions de ce travail d'étudiant n'engagent que la responsabilité de son auteur et non celle d'AGROCAMPUS OUEST".*

Fiche de diffusion du mémoire

A remplir par l'auteur[1] avec le maître de stage.

Aucune confidentialité ne sera prise en compte si la durée n'en est pas précisée.

Préciser les limites de la confidentialité [2] **:**

➢ Confidentialité absolue : ☐ oui ☒ non

(ni consultation, ni prêt)

↳ *Si oui* ☐1 an ☐5 ans ☐10 ans

➢ A l'issue de la période de confidentialité **ou** si le mémoire n'est pas confidentiel, merci de renseigner les éléments suivants :

Référence bibliographique diffusable[3] : ☐ oui ☐ non

Résumé diffusable : ☐ oui ☐ non

Mémoire consultable sur place : ☐ oui ☐ non

Reproduction autorisée du mémoire : ☐ oui ☐ non

Prêt autorisé du mémoire : ☐ oui ☐ non

.....

Diffusion de la version numérique : ☒ oui ☐ non

↳ *Si* oui, l'auteur[1] complète l'autorisation suivante :

Je soussigné(e) Alexandre Pelé *, propriétaire des droits de reproduction dudit résumé, autorise toutes les sources bibliographiques à le signaler et le publier.*

Date : *Signature :*

Rennes/Angers, le

Le maître de stage[4]**,** **L'enseignant référent,** **L'auteur**[1]**,** Alexandre Pelé

(1) auteur = étudiant qui réalise son mémoire de fin d'études

(2) L'administration, les enseignants et les différents services de documentation d'AGROCAMPUS OUEST s'engagent à respecter cette confidentialité.

(3) La référence bibliographique (= Nom de l'auteur, titre du mémoire, année de soutenance, diplôme, spécialité et spécialisation/Option)) sera signalée dans les bases de données documentaires sans le résumé.

(4) Signature et cachet de l'organisme.

REMERCIEMENTS

Je tiens tout d'abord à remercier Madame Anne-Marie CHEVRE sans qui, bien évidemment, ce stage n'aurait été possible. Mais je te remercie surtout de m'avoir transmis ta passion ainsi que ta vision de la recherche qui m'attirait déjà avant mon arrivée au Rheu. Ta bonne humeur, tes idées, ton charisme, ta patience et surtout ton soutien, m'ont fait plus qu'apprécier ces 6 derniers mois. Au vu de tout ce qui a été accompli pendant ce stage et des hypothèses amenées, je ne regrette pas de l'avoir réalisé et remercie, par la même, Madame Maria MANZANARES-DAULEUX qui m'a conduit à toi Anne-Marie.

J'ai bénéficié d'un encadrement d'une grande qualité de ta part mais aussi des différents membres de ton équipe. Je remercie donc toutes ces personnes intervenues à des moments clefs de ce stage. En premier lieu, Maryse LODE qui m'a formée en Biologie Moléculaire et à qui je laisse un cahier de laboratoire bien rempli pour la suite. En second lieu Madame Sylvie NEGRE ainsi que Messieurs Gilles LASSALLE et Cyril FALENTIN qui mon eux initiés au domaine bioinformatique. Merci à tous les quatre pour votre patience et votre aide précieuse ainsi qu'à toi Gilles pour tous ces cours particuliers et fous rires. En troisième lieu, je tiens à remercier toutes les personnes qui ont assisté inlassablement aux répétitions de mon oral du concours de l'école doctorale VAS. Merci à Gwenn TROTOUX, Régine DELOURME et Sophie PAILLARD, c'est aussi grâce à vous. Merci également pour le temps que vous m'avez accordé, pour répondre à mes questions et pour le rapport de stage, ainsi que pour vos remarques, qui ont contribuées à l'avancé de cette étude. Je tiens à finir ces remerciements avec Messieurs Eric JENCZEWSKI, Matthieu FALQUE et Franck GAUTHIER intervenus pour apporter des réponses concernant les résultats ainsi que pour m'inculquer la notion d'interférence.

L'étude de la distribution des évènements de recombinaison à été une très belle aventure de 6 mois que j'espère pouvoir poursuivre pour les trois prochaines années, avec vous tous, pour des résultats encore plus surprenants espérons le. Quoi qu'il advienne, je remercie encore une fois toute l'équipe ainsi qu'Anne-Marie de m'avoir permis de participer à ce projet qui m'a beaucoup apporté tant sur le plan personnel et professionnel et m'a définitivement convaincu que les OGM ne sont pas la (seule) solution pour améliorer les programmes de sélection variétale.

ABREVIATIONS

ADN = Acide Désoxyribose Nucléique

ARN = Acide Ribose Nucléique

BAC = Bacterial Artificial Chromosomes

Bases A, C, G, T = Adénine, Cytosine, Guanine, Tyrosine

BLASTn = Nucléotide Blast

CDB = Cassure Double Brin

cM = Centimorgan

CO = Crossing-over

Gb = Gigabase

Kb = Kilobase

LOD = Logarithm of the odds

Mb = Mégabase

NCO = Non crossing-over

Pb = Paire de base

PCR = Polymerase Chain Reaction

QTL = Quantitative Trait Loci

SSR = Simple Sequence Repeat

TEs = Eléments Transposables

ZMM = Zip, Msh, Mer

TABLE DES GRAPHIQUES

Figure 1. Représentation schématique de la ségrégation des chromosomes en méiose *2*

Figure 2. Prophase I de méiose *2*

Figure 3. Représentation schématique des voies de résolution de la cassure double-brin *3*

Figure 4. Immunolocalisation de la protéine MLH1 au stade pachytène *3*

Figure 5. Variation de la fréquence de recombinaison le long du chromosome 4 d'*A. thaliana* *5*

Figure 6. Distribution des évènements de recombinaison et de la teneur en GC chez *A. thaliana* *6*

Figure 7. Eléments Transposables fréquemment observés dans le génome des plantes *7*

Figure 8. Idéogramme de *B. rapa* *7*

Figure 9. Position des centromères entre les blocs de Schranz le long des chromosomes de *B. rapa* ...*7*

Figure 10. Distribution des évènements de recombinaison et de la méthylation sur les cinq chromosomes d'*A. thaliana* *8*

Figure 11. Evènements de duplication des génomes *9*

Figure 12. Relation entre les génomes des Brassicas *10*

Figure 13. Cartographie de populations issues d'un hybride diploïde AA, triploïde AAC et tétraploïde AACC *10*

Figure 14. Représentation schématique de la production des hybrides *12*

Figure 15. Structure du chromosome A07 de *B. rapa* *17*

Figure 16. Ratios des fréquences de recombinaison le long du chromosome A07 par rapport au témoin AA sans chromosomes C en addition pour les populations $1C_9$ et 9C *18*

Figure 17. Cartographie du groupe de liaison A07 pour les populations 0C, $1C_9$ et 9C *19*

Figure 18. Fréquence de CO par chromatide et par individu pour les populations 0C, $1C_9$ et 9C *19*

Figure 19. Représentation de la distribution des évènements de recombinaison le long du chromosome A07 de *B. rapa* chez les populations 0C, $1C_9$ et 9C *20*

Figure 20. Distribution des fréquences en gènes et en TEs ainsi que du pourcentage en bases GC le long du chromosome A07 .....*22*

Figure 21. Schématisation des résultats obtenus pour l'étude de la distribution des évènements de recombinaison sur le chromosome A07 chez les populations 0C, $1C_9$ et 9C *27*

TABLE DES TABLEAUX

Tableau 1. Ratios CDB/CO chez différents modèles d'étude de la recombinaison homologue *4*

Tableau 2. Fréquence de recombinaison en cM/Mb pour quelques espèces de plantes *5*

Tableau 3. Séquences Forward et Reverse des marqueurs utilisés par Suay *et al.* (2013) *13*

Tableau 4. Caractéristiques des motifs SSR *13*

Tableau 5. Caractéristiques des couples d'amorces générés par primer3 .....*13*

Tableau 6. Blastn entre la séquence centromérique CentrBr1 et le chromosome A07 de *B. rapa* *17*

Tableau 7. Blastn entre la séquence péricentromérique pBct et le chromosome A07 de *B. rapa* *17*

Tableau 8. Position des marqueurs des Blocs de Schranz sur le chromosome A07 de *B. rapa* *17*

Tableau 9. Positions physiques des SSR utilisés par Suay *et al.* (2013) sur le génome de *B. rapa* *18*

Tableau 10. Corrélations entre la teneur en GC, gènes et TEs avec la fréquence de recombinaison dans les régions densifiées en SSR du chromosome A07 .....*21*

Tableau 11. Mesure de l'interférence, via les données de génotypage, avec les modèles à 1 voie et à 2 voies de COs chez les populations 0C, $1C_9$ et 9C, sur le chromosome A07 .....*22*

TABLE DES MATIERES

INTRODUCTION GENERALE 1

CHAPITRE I. SYNTHESE BIBLIOGRAPHIQUE 2

I. LA RECOMBINAISON HOMOLOGUE MEIOTIQUE 2

I.1. Un phénomène complexe intervenant au cours de la méiose 2

I.2. Les produits de la recombinaison homologue méiotique 3

I.2.1. Les Crossing-overs (COs) 3

I.2.1.1. Les COs de voie I: Type Interférents *3*

I.2.1.2. Les COs de voie II: Type Non Interférents *4*

I.2.2. La conversion génique ou Non Crossing-Over (NCO) 4

II. REGULATION ET DISTRIBUTION DES EVENEMENTS DE RECOMBINAISON 4

II.1. Des tendances et des lois régissent le nombre et la distribution des COs 4

II.1.1. Deux règles pour réguler et orienter les COs 4

II.1.1.1. L'homéostasie *4*

II.1.1.2. L'interférence *5*

II.2.1. Des variations et des tendances 5

II.2. Des évènements contrôlés 6

II.2.1. Par des facteurs génétiques 6

II.2.2. Par des facteurs génomiques 6

II.2.2.1. La teneur en GC *6*

II.2.2.2. La teneur génique *7*

II.2.2.3. Les éléments transposables (TEs) *7*

II.2.3. Par des facteurs épigénétiques 8

II.2.3.1. La méthylation de l'ADN *8*

II.2.3.2. La modification de la structure de l'ADN *8*

II.2.4. Par des facteurs externes 9

III. LE COLZA, UN MODELE D'ETUDE DE LA RECOMBINAISON 9

III.1. Généralités sur le colza 9

III.2. Origine et sélection 9

III.3. La recombinaison chez le colza 10

III.3.1. La recombinaison homéologue 10

III.3.2. La recombinaison homologue 10

IV. OBJECTIFS ET STRATEGIES DE L'ETUDE 11

CHAPITRE II. MATERIEL ET METHODES 12

I. MATERIEL VEGETAL 12

II. ANALYSES BIOINFORMATIQUES 12

II.1. Etude de la structure du chromosome A07 de *B. rapa* 12

II.2. Localisation physique du centromère du chromosome A07 de *B. rapa* 13
II.3. Localisation physique des SSR établis sur le chromosome A07 de *B. rapa* 13
II.4. Design de SSR 13
III. ANALYSES BIOMOLECULAIRES 14
III.1. Extraction ADN 14
III.2. Amplification 14
III.3. Lecture SSR 14
IV. ANALYSES STATISTIQUES 14
IV.1. Etude de la ségrégation allélique des marqueurs 14
IV.2. Cartographie génétique 15
IV.3. Etude de l'hétérogénéité des fréquences de CO 15
IV.4. Etude des corrélations liées à la structure génomique 15
IV.5. Mesure de l'interférence 16
CHAPITRE III. RESULTATS 17
I. ETUDE DE LA STRUCTURE DU CHROMOSOME A07 DE *B. RAPA* 17
I.1. Comparaison des chromosomes A07 de *B. rapa* et de *B. napus* 17
I.2. Localisation du centromère d'A07 17
II. ETUDE DE LA DE LA DISTRIBUTION DES COs LE LONG DU CHROMOSOME A07 18
II.1. Définition des zones à densifier 18
II.2. Design des marqueurs SSR 18
II.3. Cartographie génétique 19
II.4. Distribution des COs le long du chromosome A07 20
II.4.1. Distribution chez le témoin AA sans chromosome C en addition 20
II.4.2. Comparaison avec les populations $1C_9$ et 9C 20
II.5. Analyse des corrélations entre la fréquence de recombinaison et la structure génomique d'A07 ...21
II.5.1. Etude des corrélations chez la population AA sans chromosome C en addition 21
II.5.2. Etude des corrélations avec des chromosomes C en addition 21
II.5.3. Qu'en est-il sur le chromosome A07 entier ? 22
III. ANALYSE DE L'INTERFERENCE 22
CHAPITRE IV. DISCUSSION 23
I. STRUCTURE DU CHROMOSOME A07 DE *B. RAPA* 23
II. ETUDE DE LA REGULATION ET DE LA DISTRIBUTION DES COs LE LONG D'A07 24
II.1. Régulation et distribution des évènements de COs chez des diploïdes AA 24
II.2. Effets de l'addition des chromosomes C sur la régulation et la distribution des COs 25
CONCLUSION GENERALE 28
BIBLIOGRAPHIE 29
ANNEXES 37

TABLE DES ANNEXES

Annexe I. Protocole d'extraction ADN *38*

Annexe II. Mix et programmes PCR *41*

Annexe III. Résultats du polymorphisme des SSR designés *42*

Annexe IV. position, ségrégation et amorces des différents SSR polymorphes *45*

Annexe V. Cartographie des populations 0C issue de AACxAA ou de AAxAA *46*

Annexe VI. Analyses statistiques *47*

Annexe VII. Corrélations *49*

Annexe VIII. Résultats de l'étude de la distributions des COs le long d'A02 et A04 .....*56*

INTRODUCTION GENERALE

La méiose est le processus biologique, présent chez toutes les espèces à reproduction sexuée, qui conditionne la formation des gamètes et assure la stabilité des génomes. Cet évènement très conservé a également pour rôle de générer de la diversité par brassage génétique via la recombinaison méiotique qui contrôle la formation de crossing-over (COs) lors de la première division. Les gamètes qui en résultent présentent donc de nouvelles combinaisons alléliques utilisées en sélection variétale. Cependant, le nombre de COs est strictement régulé. Il n'existe en effet qu'un CO obligatoire entre chromosomes homologues, garants de leur ségrégation équilibrée dans les gamètes, mais rarement plus. L'augmentation de leur nombre fait actuellement l'objet d'études afin d'optimiser le brassage génétique. Ainsi chez le colza (*Brassica napus*), produit de l'hybridation interspécifique naturelle de *B. rapa* (AA, 2n=2x=20) et de *B. oleracea* (CC, 2n=2x=18), il a récemment été observé que la composition chromosomique influence le taux de recombinaison. En effet, chez des hybrides triploïdes AAC, il a été observé une augmentation du taux de recombinaison homologue entre les chromosome A d'un facteur 4 à 6 par rapport à des diploïdes AA. De plus, il a été constaté qu'une forte partie de cette augmentation est due à la présence du seul chromosome C_9. Cependant, le lien entre la fréquence de recombinaison et la distribution physique des COs le long des chromosomes n'a pas encore été étudié chez cette espèce alors qu'il a été établi chez d'autres que les COs ne sont pas répartis de manière homogène. Les séquençages récents des génomes de *B. rapa* et de *B. napus* nous permettent maintenant de faire ce lien. Nous nous sommes donc fixés pour objectif d'étudier la distribution des COs le long d'un chromosome du génome A chez des diploïdes AA, d'observer l'effet de l'ajout des 9 chromosomes C ou du chromosome C_9 seul sur cette répartition et de déterminer le rôle de la structure génomique du chromosome A dans cet effet. Lors de l'introduction bibliographique, nous présenterons les données acquises sur la méiose, la formation de COs, les facteurs qui les contrôlent et l'intérêt du modèle choisi. Après la description des matériels et méthodes utilisés, les résultats seront décrits pour la structure du chromosome A07, pour la distribution des COs dans des populations contrastées pour la fréquence de recombinaison et pour l'effet de la structure génomique sur cette répartition. Nous discuterons enfin de la régulation de la distribution des COs observée, notamment dans les régions péricentromériques au regard des données bibliographiques afin de dégager des perspectives sur les travaux à conduire et sur l'application de nos résultats dans les programmes d'amélioration des plantes.

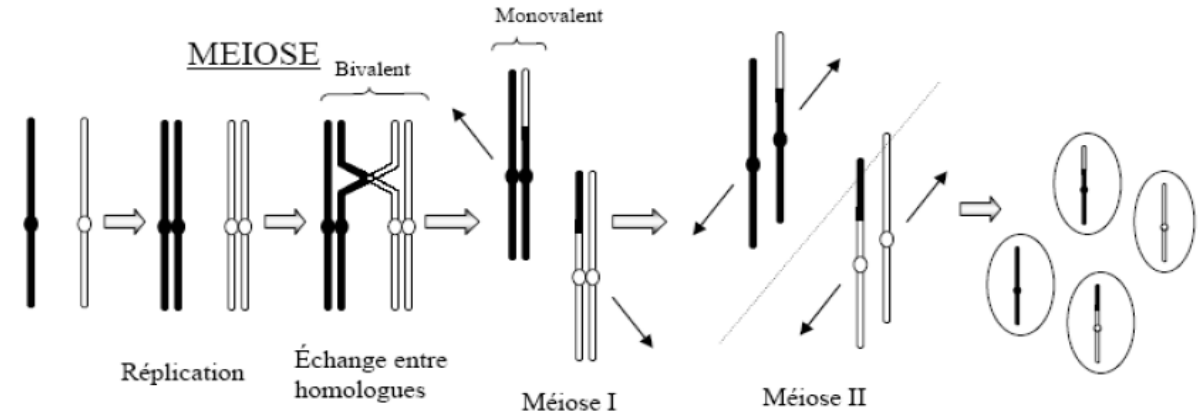


Figure 1. Représentation schématique de la ségrégation des chromosomes en méiose par Smith et Nicolas (1998)

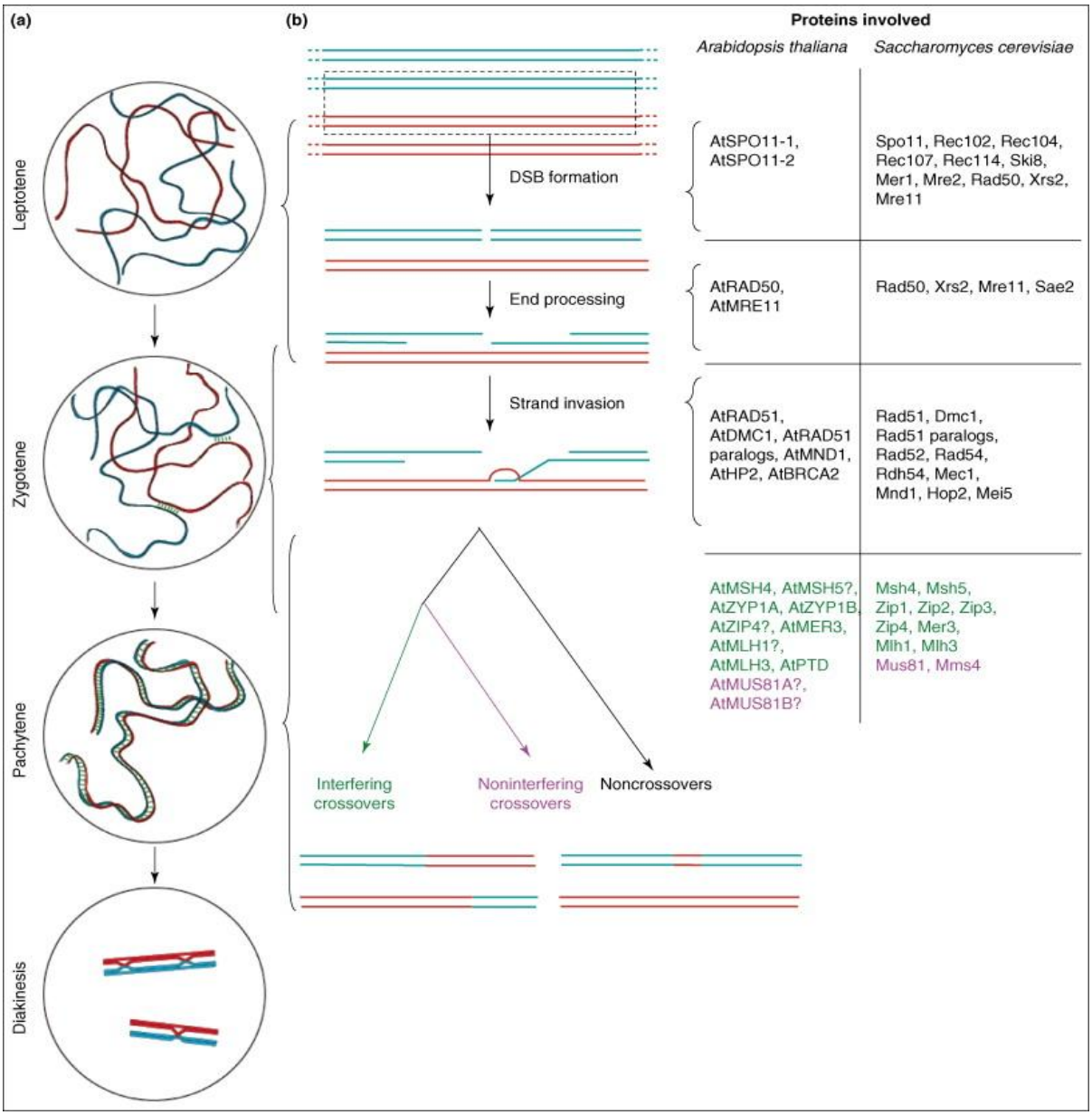


Figure 2. Prophase I de méiose par Mézard *et al.* (2007). **a/** Représentation schématique des étapes de la prophase I de méiose, **b/** Modèle des voies de recombinaison méiotique chez *Saccharomyces cerevisiae* et *Arabidopsis thaliana* et protéines impliquées.

CHAPITRE I. SYNTHESE BIBLIOGRAPHIQUE

I. LA RECOMBINAISON HOMOLOGUE MEIOTIQUE

I.1. Un phénomène complexe intervenant au cours de la méiose

La méiose est le processus biologique à l'origine de la formation de gamètes haploïdes (n) contenant un jeu unique de chromosomes issus du brassage entre les chromosomes homologues de l'organisme diploïde (2n). Retrouvé chez la majorité des eucaryotes, cet évènement très conservé de la reproduction sexuée est contrôlé de manière génétique (Bhatt *et al.* 2001 ; Hamant *et al.* 2006). Précédée par une phase de réplication de l'ADN, appelée réplication pré-méiotique, la méiose se compose de deux divisions cellulaires (Figure 1), chacune découpée en quatre phases repérables par des observations cytologiques : la Prophase, la Métaphase, l'Anaphase et la Télophase. Lors de la première division, appelée division réductionnelle, la cellule mère se divise en deux cellules filles haploïdes bichromatidiennes puis en quatre cellules filles haploïdes mono-chromatidiennes après la seconde division nommée division équationnelle (Zickler et Kleckner, 1999 ; Petronczki *et al.* 2003 ; Hamant *et al.* 2006).

La recombinaison méiotique est également un processus biologique très conservé, sous contrôle génétique et se produit au cours de la division réductionnelle de la méiose (Anderson et Stack, 2002). Celle-ci garantit la ségrégation équilibrée des chromosomes homologues en assurant une connexion physique entre eux par l'intermédiaire de crossing-overs. Elle évite ainsi une composition chromosomique anormale dans les gamètes pouvant générer des plantes aneuploïdes. De plus, son intérêt en sélection est établi : la recombinaison méiotique est le principal mécanisme permettant de brasser la diversité lors de la production de descendance dans les programmes de sélection variétale (Wijnker et De Jong, 2008). Elle correspond à un échange de matériel génétique entre chromosomes homologues (i.e. provenant du même génome) et permet ainsi de créer de nouvelles combinaisons alléliques (Cromie et Smith, 2007). Ce processus complexe, divisé en plusieurs étapes, intervient tout au long de la Prophase I et repose sur un principe de cassure et de réparation des brins d'ADN (Mézard *et al.* 2007) (Figure 2). C'est au cours du stade Leptotène et jusqu'au Zygotène que la recombinaison méiotique est initiée par la formation programmée de cassures double-brin (CDB) de l'ADN sur l'une des chromatides via la protéine Spo11 isolée pour la première fois chez *Saccharomyces cerevisiae* (Atcheson, 1987). La maturation, catalysée par les endonucléases Rad50 et Mre11, intervient ensuite et permet la résection des extrémités 5' des CDB générées, produisant des extrémités 3' simple-brin (Keeney et Kleckner, 1995 ; Keeney et Neale, 2006 ; Borde, 2007). A la fin du stade Zygotène les chromosomes homologues commencent à s'associer, entrainant l'étape dite « d'invasion » de la recombinaison. Lors de cette dernière, les extrémités 3' simple-brin, générées en maturation, envahissent la molécule d'ADN homologue intacte via les protéines recombinases Rad51 et Dmc1, entraînant sur celle-ci la formation d'une boucle appelée D-loop (Baumann *et al*. 1996; Zickler et Kleckner, 1999 ; Sehorn *et al.* 2004 ; Gerton et Hawley 2005).

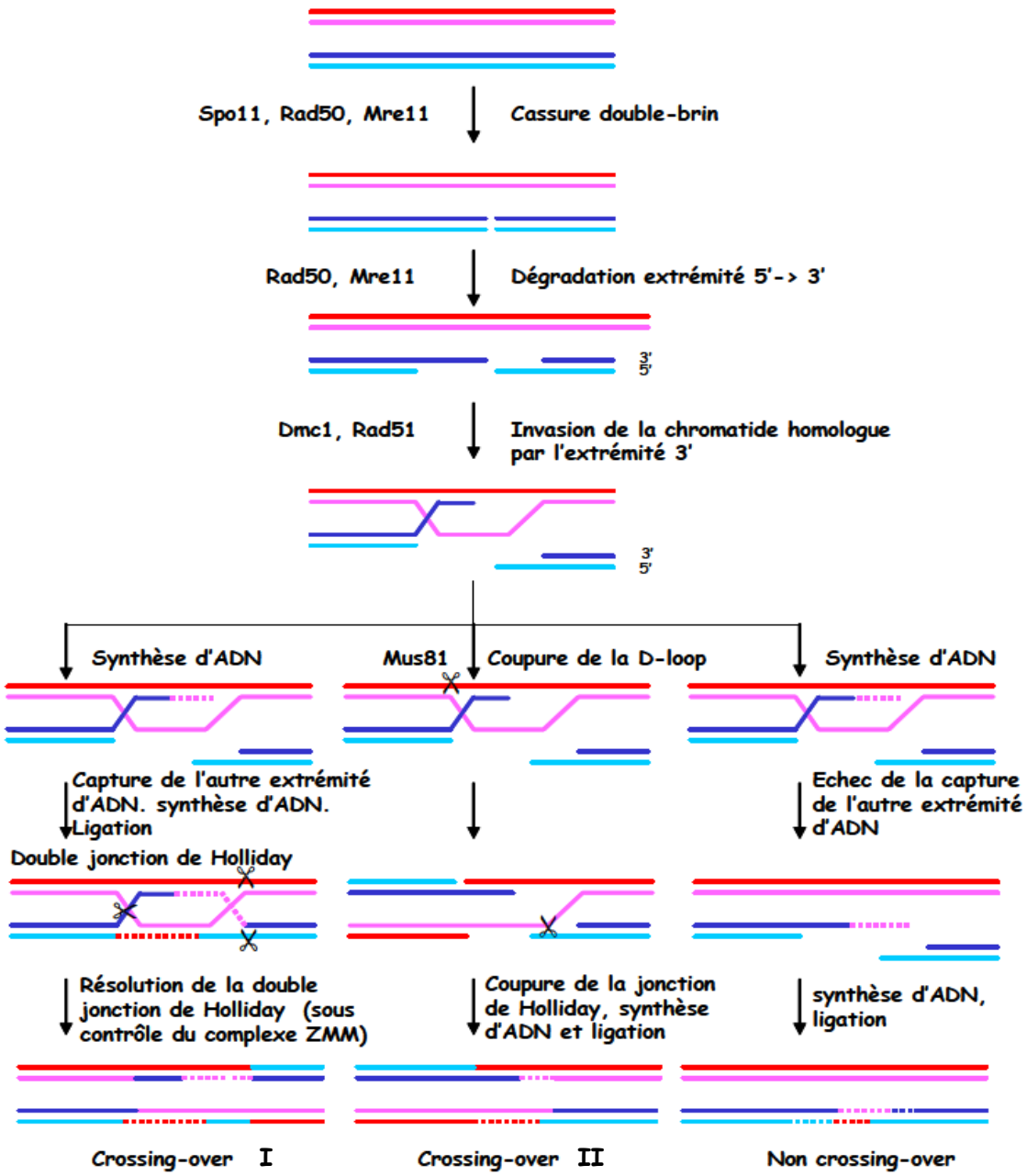


Figure 3. Représentation schématique des voies de résolution de la cassure double-brin par Leflon *et al.* (2007) d'après Hollingsworth et Brill (2004)

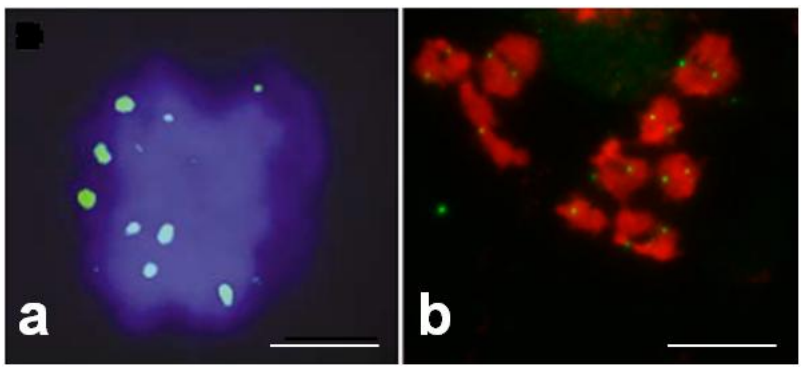


Figure 4. Immunolocalisation de la protéine MLH1 au stade pachytène (en vert) chez les chromosomes **a/** d'*Arabidopsis thaliana* (Jackson *et al.* 2006) et **b/** de *Brassica rapa* (Leflon *et al.* 2010). Barres=10mm.

I.2. Les produits de la recombinaison homologue méiotique

Au stade Pachytène, les chromosomes homologues sont entièrement synapsés et la réparation des CDB peut avoir lieu (Zickler et Kleckner, 1999). Celle-ci constitue l'ultime étape de la recombinaison homologue méiotique et, selon les cas, aboutit à un crossing-over ou à un non-crossing-over (Osman *et al.* 2011 ; Youds et Boulton, 2011).

I.2.1. Les Crossing-overs (COs)

Les COs, produits de la recombinaison, constituent la clef du brassage génétique car ils permettent un échange réciproque de matériel génétique entre des régions homologues. Ils assurent également la ségrégation des chromosomes pendant la méiose, qui ne sont plus maintenus entre eux que par les chiasmas (signature cytologique de la recombinaison) au cours des stades Diplotène et Diacinèse et ce, jusqu'à leur séparation en fin de Métaphase I. Deux grandes approches permettent leurs observations. La première, directe, consiste à visualiser les sites des COs par cytogénétique en localisant les protéines spécifiques des CDB et COs (Anderson et Stack, 2002). Cette technique permet de dénombrer les COs, mais pose cependant problème quand à l'étude de leur distribution du fait d'une évaluation difficile sur chromosome. Elle correspond ainsi à des estimations des distances en µm ne prenant pas en compte les différences de compaction de l'ADN le long des chromosomes. La seconde, indirecte, consiste à analyser les gamètes formés par génotypage des descendants qu'ils ont permis de produire ; les distances entre marqueurs des cartes génétiques traduisant la fréquence de recombinaison qu'il existe dans cet intervalle et mesurée en unité de distance génétique (cM). Avec l'arrivée des données de séquençage des génomes, la localisation physique des marqueurs sur les chromosomes permet de mieux situer la position des COs. Cette méthode a ainsi été mise en place chez la levure, l'humain, la souris et *Arabidopsis thaliana* dans ce but (Mancera *et al*. 2008 ; Myers *et al.* 2005 ; Paigen *et al.* 2008 ; Drouaud *et al.* 2006, 2007). Toutefois, sa résolution est fortement corrélée au nombre de marqueurs et d'individus utilisés.

I.2.1.1. Les COs de voie I : Type Interférents

Dans le cas où l'extrémité 3' simple brin, générée lors de la maturation, est capturée par la D-loop, formée chez le chromosome homologue, le brin est alors réparé en utilisant comme matrice la séquence 5'3' homologue. La double jonction de Holliday se forme et sa résolution entraîne la formation d'un crossing-over Interférent (Figure 3). Cette résolution est catalysée par le complexe multiprotéique ZMM (Lynn *et al.* 2007) et dépend de la protéine MLH1 (Lhuissier *et al.* 2007). Son immunolocalisation au pachytène a été utilisée chez de nombreuses espèces comme *Arabidopsis thaliana*, *Brassica rapa* et *Lycopersicon esculentum* pour révéler la présence de COs Interférents (Jackson *et al.* 2006 ; Leflon *et al.* 2010 ; Lhuissier *et al.* 2007) (Figure 4). Ces derniers représentent d'une manière générale, 70% des COs générés au cours de la méiose (85% des CO identifiés chez *A. thaliana* et *S. cerevisiae* sont de type I et 100% des CO formés chez *C. elegans*) (Mezard *et al.* 2007).

Tableau 1. Ratios CDB/CO chez différents modèles d'étude de la recombinaison homologue méiotique d'après Mézard *et al.* (2007)

Espèces	Nb. Estimé de sites de CDB	Nb. Estimé de CO	Ratio
A. thaliana	250	9,2	27
L. esculentum	290	22	13
L. longiflorum	2000	55	36
S. cerevisiae	130	60	2
C. elegans	12	6	2
M. musculus	130	23	13
H. sapiens	150	50	3

I.2.1.2. Les COs de voie II : Type Non Interférents

Dans le cas où la D-loop est coupée par des endonucléases comme Mus81, chaque extrémité est associée à une extrémité différente de la chromatide invasive et la structure est résolue en CO Non interférents par une synthèse d'ADN suivie d'une ligation (Gaskell *et al.* 2007) (Figure 3). L'immunolocalisation des protéines spécifiques de cette voie reste cependant difficile et n'a pu être réalisée que chez *A. thaliana* (Knoll *et al.* 2012). Ces COs sont minoritaires par rapport aux COs I chez l'ensemble des espèces, hormis chez *Schizosaccharomyces pombe* où la totalité des COs est de type II (Mezard *et al.* 2007).

I.2.2. La conversion génique ou Non Crossing-Over (NCO)

Dans le cas où l'extrémité 3' simple brin n'est pas capturée par la D-loop, l'association entre brins homologues est alors rompue. Le brin d'ADN présentant la CDB est réparé en utilisant comme matrice le brin complémentaire (Börner *et al*. 2004) (Figure 3). Cette réparation n'entraîne pas d'échange ni de modification des séquences à l'exception d'un fragment d'ADN resynthétisé à partir de la molécule homologue. Ce fragment, de l'ordre du pb au kb, identifié chez différentes espèces comme le maïs et *A. thaliana* donne lieu à une conversion génique appelée aussi non-crossing-over (NCO) (Dooner et Martinez-Ferez, 1997 ; Plagnol *et al.* 2006). Ce phénomène reste encore peu étudié car les NCOs sont difficiles à visualiser. Néanmoins, comme le nombre de CDB estimé est largement supérieur au nombre de CO chez plusieurs espèces dont *A. thaliana*, il en a été déduit par Mézard *et al.* (2007) que les NCOs sont le principal produit de la réparation des CDB.

II. REGULATION ET DISTRIBUTION DES EVENEMENTS DE RECOMBINAISON

II.1. Des tendances et des lois régissent le nombre et la distribution des COs

II.1.1. Deux règles pour réguler et orienter les COs

II.1.1.1. L'homéostasie

Jones (1984) a observé qu'il existe toujours un CO obligatoire par chromosome et par méiose, indispensable pour assurer la bonne ségrégation des chromosomes homologues à chacun des pôles de la cellule. C'est pour cela qu'il existe un mécanisme garant du maintien d'un nombre minimal de CO : l'homéostasie (Martini *et al.* 2006). Ainsi, il semble que le nombre de CO par chromosome soit fixé et que la cellule puisse maintenir ce niveau malgré certaines perturbations. Elle est pour cela capable, lorsqu'un CO n'arrive pas à se former sur un chromosome sans homologue, de mettre en place un mécanisme compensatoire conduisant à une augmentation des COs sur d'autres chromosomes recombinants (Carlton *et al.* 2006). Le nombre de CO est donc indépendant du nombre de NCO, plus nombreux, chez la plupart des espèces, tel qu'*A. thaliana, L. esculentum, L. longiflorum, S. cerevisiae, C. elegans, M. musculus* et *H. sapiens* (Mezard *et al*. 2007 ; Chen *et al.* 2008 ; Roig et Keeney, 2008 ; Hillers et Villeneuve, 2003 ; Baudat et De Massy, 2007) (Tableau 1).

Tableau 2. Fréquence de recombinaison en cM/Mb pour quelques espèces de plantes d'après Saintenac (2009)

Espèces	Caryotype	Mb	cM	cM/Mb	Références
A. thaliana	2n=2x=10	157	753,6	4,8	(Gaut *et al.* 2007)
B. napus	2n=4x=38	1127	1605	1,42	(Sun *et al.* 2007)
B. oleracea	2n=2x=18	760	703	0,93	(Gao *et al.* 2007)
B. rapa	2n=2x=20	784	1378	1,75	(Choi *et al.* 2007)
C. papaya	2n=2x=18	368	1069	2,91	(Chen *et al.* 2007)
L. sativa	2n=2x=18	2205	1346	0,61	(Truco *et al.* 2007)
T. cacao	2n=2x=20	417	782	1,88	(Pugh *et al.* 2004)
V. vinifera	2n=2x=38	417	1460	3,50	(Troggio *et al.* 2007)

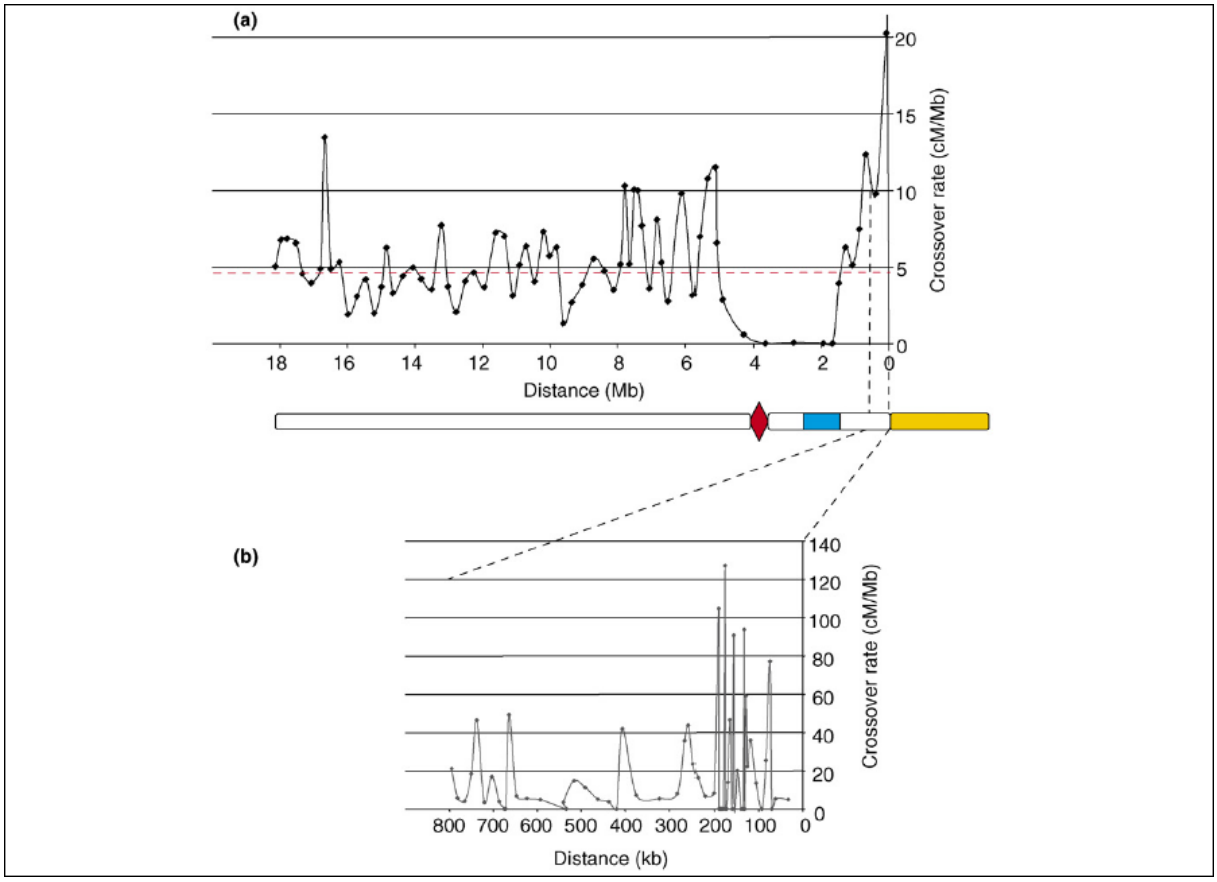


Figure 5. Variation de la fréquence de recombinaison le long du chromosome 4 d'*A. thaliana* par Mézard (2006) **a/** Représentation de la fréquence de recombinaison le long du chromosome 4 d'*A. thaliana* schématisé en abscisse avec le centromère (rouge), le bouton hétérochromatique (bleu) et la région organisatrice nucléolaire (jaune). **b/** Représentation de la fréquence de recombinaison le long des 800 premiers kb du bras court du chromosome 4 d'*A. thaliana*.

II.1.1.2. L'interférence

Si un CO est obligatoire par chromosome et par méiose, un phénomène, nommé Interférence, en limite toutefois le nombre à 1 ou 2 par bivalent chez la plupart des espèces. Celui-ci empêche la formation coïncidente d'un CO I sur une paire de chromosomes lorsqu'un CO I s'y est déjà produit. C'est pour cela que la distance entre COs est plus importante qu'attendu en cas d'indépendance. Cependant, la formation des COs Interférents n'influe pas sur la position des COs II distribués aléatoirement le long des chromosomes (Jones, 1984 ; Muller, 1916 ; Mezard *et al*. 2007). Les mécanismes responsables de ce phénomène sont encore inconnus, mais plusieurs modèles statistiques ont été mis en place pour mesurer l'intensité de l'interférence sur les chromosomes avec des données cytogénétiques (localisation de MLH1) ou issues du génotypage (Foss *et al*. 1993 ; McPeek et Speed, 1995 ; Mézard *et al*. 2007 ; Falque *et al.* 2007 ; Falque *et al.* 2009). Ces modèles permettant de tester la distribution aléatoire ou non des COs ont permis d'identifier le phénomène d'Interférence positive chez la quasi-totalité des espèces testées hormis *S. pombe* et *A. nidulans* (Broman et Weber 2000 ; Broman *et al*. 2002 ; Cromie et Smith 2007 ; Lhuissier *et al*. 2007 ; Copenhaver *et al.* 2002 ; Shaw et Moore 1998).

II.1.2. Des variations et des tendances

Nous avons pu voir précédemment que le nombre et la distribution des COs sont strictement régulés chez l'ensemble des espèces. Toutefois, des variations et des tendances ont été mises en évidence. **A une échelle inter-espèce**, celles-ci semblent trouver leur principale explication dans le nombre de chromosomes (Tableau 2) (Kaback, 1996 ; Mancera *et al*. 2008 ; Shifman *et al.* 2006 ; Anderson et Stack 2002). En effet, à taille de génome équivalente, l'espèce ayant le plus grand nombre de chromosome présente la plus forte fréquence de recombinaison. Par exemple, *Theobroma cacao* (2n=20) et *Vitis vinifera* (2n=38) ont un génome avoisinant les 417 Mb mais la vigne, avec 18 chromosomes de plus que le cacao, présente la fréquence de recombinaison la plus élevée (Tableau 2). La polyploïdisation, qui correspond soit au doublement du nombre de chromosomes (autoploïdisation) soit à la réunion de deux génomes (alloploïdisation), joue également un rôle. Il a notamment été observé chez des autotétraploïde et allotétraploïde d'*A. thaliana* ainsi que chez un allotétraploïde de colza, une augmentation d'un facteur deux du taux de recombinaison par rapport à un diploïde d'*A. thaliana* et à *B. rapa*, l'un des progéniteurs du colza (Pécinka *et al*. 2011 ; Leflon *et al*. 2010). **A une échelle intra-espèce**, des variations de la fréquence de recombinaison ont également été observées chez l'orge et *A. thaliana* fonction des cultivars (Säll *et al.* 1990 ; Sanchez-Moran *et al.* 2002). Celles-ci dépendent principalement de la structure des chromosomes ainsi que de leur configuration définie par la position du centromère (acrocentrique, métacentrique ou télocentrique) (Dumas et Britton-Davidian, 2002). En effet, il est établi que la répartition des COs le long des chromosomes n'est pas homogène, révélant ainsi la présence de régions chaudes (i.e. fréquence de CO supérieure à la moyenne de la fréquence du chromosome ou du génome) et de régions froides pour la recombinaison (Zickler et Kleckner, 1999 ; Mézard, 2006 ; Mézard *et al.* 2007 ; Drouaud *et al.* 2007) (Figure 5). Malgré l'alternance de ces régions le long des chromosomes, il a été observé, chez la plupart des espèces animales et végétales, que la fréquence de

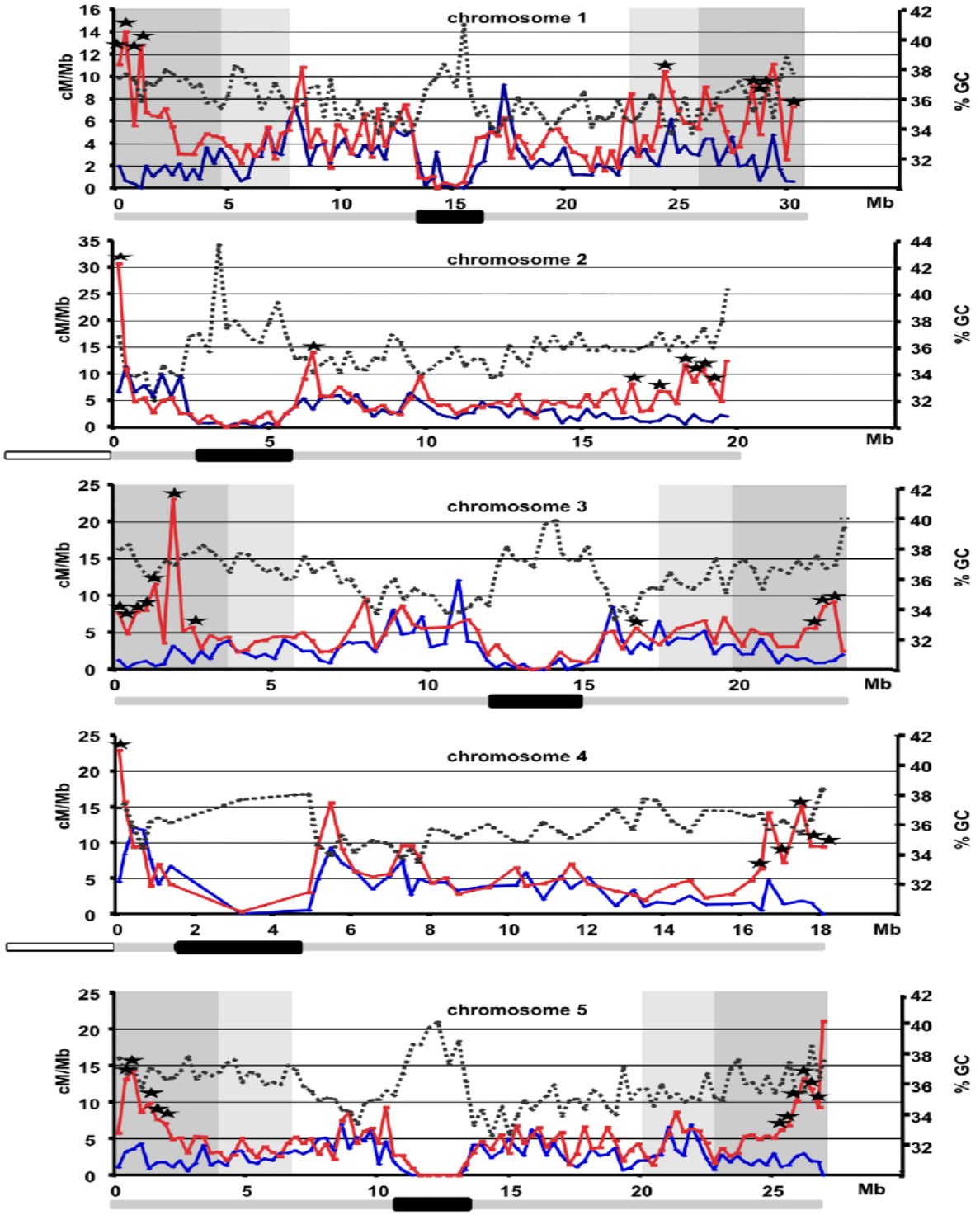


Figure 6. Distribution des évènements de recombinaison et de la teneur en GC sur les cinq chromosomes d'*Arabidopsis thaliana* par Giraut *et al.* (2011). En bleu est représentée la distribution des COs en méiose femelle, en rouge celle des mâles et en pointillés, la teneur en GC. Le chromosome est représenté en gris en bas de chaque graphique avec les régions péricentromériques en foncées.

recombinaison suit un gradient croissant du centromère aux télomères (Kong *et al.* 2002 ; Shifman *et al.* 2006 ; Lukaszewski et Curtis 1993 ; Kunzel *et al*. 2000 ; Anderson *et al.* 2003 ; Chen *et al*. 2002 ; Sherman et Stack 1995). Il semble toutefois que chez certaines espèces, une diminution de la fréquence des COs aux extrémités des télomères soit observée et que le taux dans ces régions varie en fonction du sexe de l'individu (Figure 6) (Kong *et al*. 2002; Shifman *et al*. 2006 ; Drouaud *et al.* 2007). En revanche, il est communément admis que les régions centromériques et péricentromériques sont froides pour la recombinaison, probablement pour empêcher une mauvaise ségrégation des chromosomes pendant l'Anaphase I (Choo, 1998 ; Lamb *et al.* 2005). La présence de régions froides pour la recombinaison dans cette zone ainsi que le long des chromosomes pose cependant la question de l'existence de contrôles régulant fréquence et distribution des évènements des COs (Mézard *et al.* 2007).

II.2. Des événements contrôlés

II.2.1. Par des facteurs génétiques

Les variations intra-espèce observées du taux de recombinaison ont mis en avant la possibilité d'un contrôle génétique de la fréquence moyenne des COs. Cette régulation peut s'effectue via les gènes identifiés dans la formation des COs et des CDB (RAD51, DMC1, MUS81, MLH1, SPO1...) ou via d'autres gènes spécifiques du contrôle de la fréquence des COs. En effet, des QTLs impliqués dans cet effet ont été identifiés chez la souris, *A. thaliana*, le maïs et le blé par Esch *et al*. (2007) en prenant en compte le nombre de CO comme phénotype étudié. D'autres gènes pourraient ainsi avoir un rôle dans la régulation et la distribution des évènements de CO, mais leur connaissance et les mécanismes qu'ils mettent en place ne sont pas encore connus.

II.2.2. Par des facteurs génomiques

Le développement des technologies haute densité de séquençage depuis ces 20 dernières années a permis d'étudier la relation entre la distribution des COs et les caractéristiques de la séquence chez de nombreuses espèces (Mézard *et al*. 2007 ; Wijnker et De Jong, 2008). Certaines corrélations ont ainsi été établies, permettant d'expliquer la distribution des évènements de recombinaison le long des chromosomes.

II.2.2.1. La teneur en GC

La composition en bases A, C, G, T a été le premier élément génomique étudié pour comprendre la distribution des COs le long du génome humain par Kong *et al.* (2002). Ces derniers ont ainsi mis en évidence une corrélation positive entre le pourcentage en GC et la présence de COs, qui a par la suite été étendue au règne animal (Jensen-Seaman *et al.* 2004). Toutefois, au sein du règne végétal, les résultats varient fonction de l'espèce, Giraut *et al*. (2011) identifiant notamment une corrélation négative chez les femelles et une absence de corrélation chez les mâles d'*A. thaliana* (Figure 6). Les variations des résultats semblent toutefois trouver leur explication dans l'échelle de densification. En effet, des études menées à plus petite échelle, c'est-à-dire avec une densification en marqueurs plus importante (de l'ordre du Kb contre 100 Kb à 1 Mb), ont révélé une nette corrélation positive (Myers et

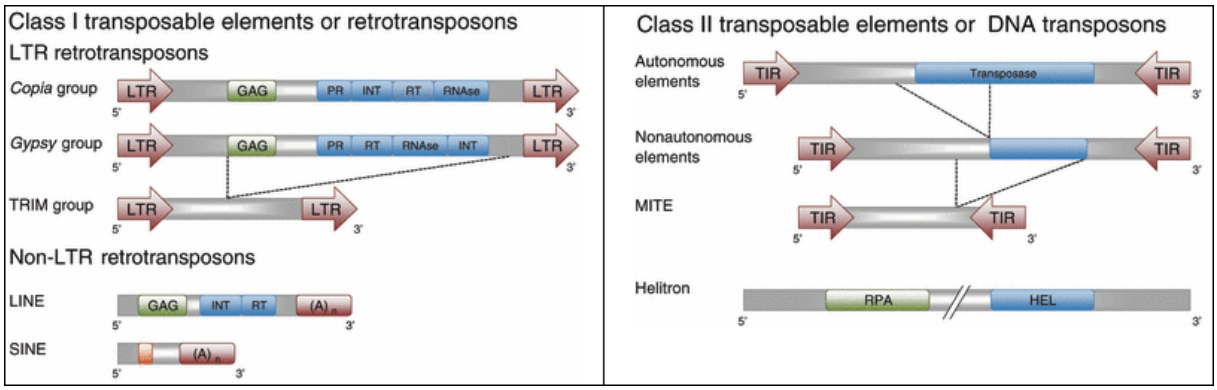


Figure 7. Eléments Transposables fréquemment observés dans le génome des plantes par Parisod *et al.* (2009)

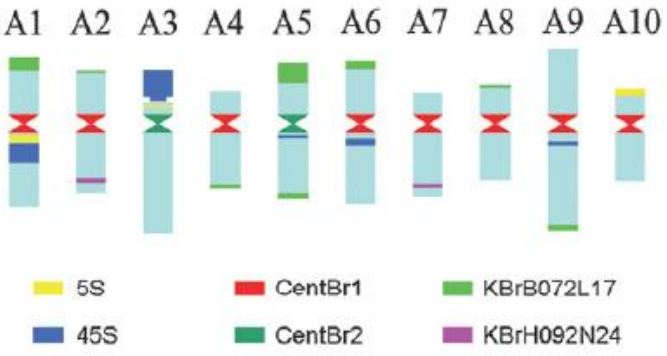


Figure 8. Idéogramme de *B. rapa* par Xiong et Pires (2011)

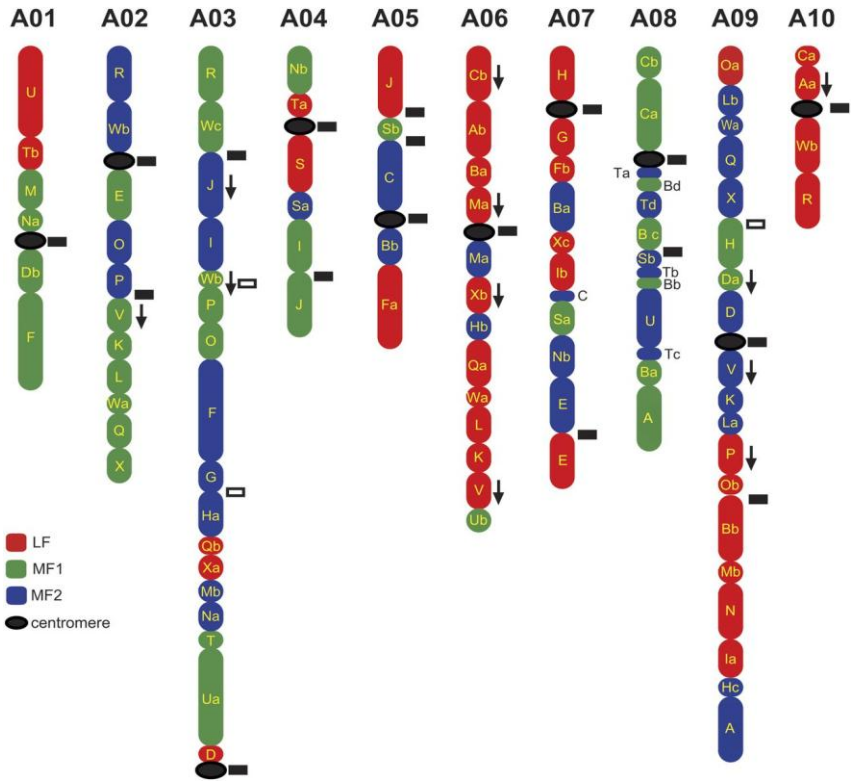


Figure 9. Position des centromères entre les blocs de Schranz le long des chromosomes de *B. rapa* par Cheng *et al.* (2013)

Spencer, 2006). Son existence semble donc universelle mais son explication fait encore l'objet de débats. La seule hypothèse formulée à ce jour propose la recombinaison comme responsable de l'enrichissement en GC de certaines régions. En effet, la formation de CDB en méiose peut induire une réparation en NCO responsable de la conversion génique (vue en I.3). Celle-ci favoriserait la réparation des mésappariements en G:C plutôt qu'en A:T (Birdsell, 2002), enrichissant ainsi en GC les régions où la fréquence en NCO est importante.

II.2.2.2. La teneur génique

Des corrélations positives plus ou moins fortes entre la teneur en gènes et la distribution des COs ont été identifiées chez plusieurs espèces dont l'humain, la souris, le riz, le maïs et le blé (Kong *et al.* 2002 ; Paigen *et al.* 2008 ; Wu *et al.* 2003 ; Anderson *et al.* 2006 ; Erayman *et al.* 2004). Les COs semblent se produire préférentiellement dans les régions riches en gènes même s'il semble difficile de généraliser car chez *A. thaliana*, une corrélation négative est détectée (Giraut *et al.* 2011). A une échelle plus dense en marqueurs, ces corrélations semblent s'établir avec les parties non codantes des gènes, suggérant un contrôle de la distribution des COs pour éviter l'inactivation de gènes lors du brassage génétique (Myers *et al.* 2005 ; de Massy, 2003). Le type de gènes pourrait également influencer la position des COs comme l'ont observé Mancera *et al.* (2008) chez la levure.

II.2.2.3. Les éléments transposables (TEs)

Les éléments transposables (TEs) sont les constituants les plus importants des génomes Eucaryotes. Ces séquences d'ADN répétées et mobiles sont considérées comme des moteurs puissants de l'évolution des génomes et de la biodiversité (Kumar et Bennetzen, 1999). Deux classes principales de TEs, subdivisées en plusieurs sous classes, sont présentes chez l'ensemble des espèces (Figure 7). La première, appelée rétrotransposon, agit via le principe du « copier-coller » par un intermédiaire ARN alors que la seconde, appelée transposon ADN, se déplace via un intermédiaire ADN par un mécanisme de « couper-coller ». Hormis chez *A. thaliana*, des corrélations négatives ont pu être établies entre la présence de points chauds de recombinaison et celle de TEs chez de nombreuses espèces (Rizzon *et al.* 2002 ; Wu *et al.* 2003 ; Jensen-Seaman *et al.* 2004 ; Dooner et He 2008 ; Giraut *et al.* 2011). Ces derniers empêchent donc d'une manière générale la présence de COs. L'accumulation des TEs, et principalement ceux de classe I, dans les régions centromériques et péricentromériques constituent l'une des explications à l'absence de recombinaison dans ces régions. Leur localisation est ainsi couramment utilisée afin d'identifier la position des centromères (Harrison et Heslop-Harrison, 1995 ; Round *et al.* 1997 ; Schmidt et Heslop-Harrison, 1998 ; Hall *et al.* 2006 ; Luce *et al.* 2006 ; Lim *et al.* 2005 ; 2007 ; Pouilly *et al.* 2008). Chez *Brassica rapa*, par exemple, deux types de rétrotransposons centromériques répétées ont été mis en évidence et caractérisées : CentBr1 et CentBr2 (Lim *et al*, 2005 ; 2007; Xiong et Pires, 2011) (Figure 8), ainsi que deux fragments de rétrotransposons spécifiques des régions péricentromériques : pBct et pBgt (Koo *et al.* 2011). Ces séquences ont récemment permis de positionner les centromères chez *B. rapa* (Figure 9) (Cheng *et al.* 2013).

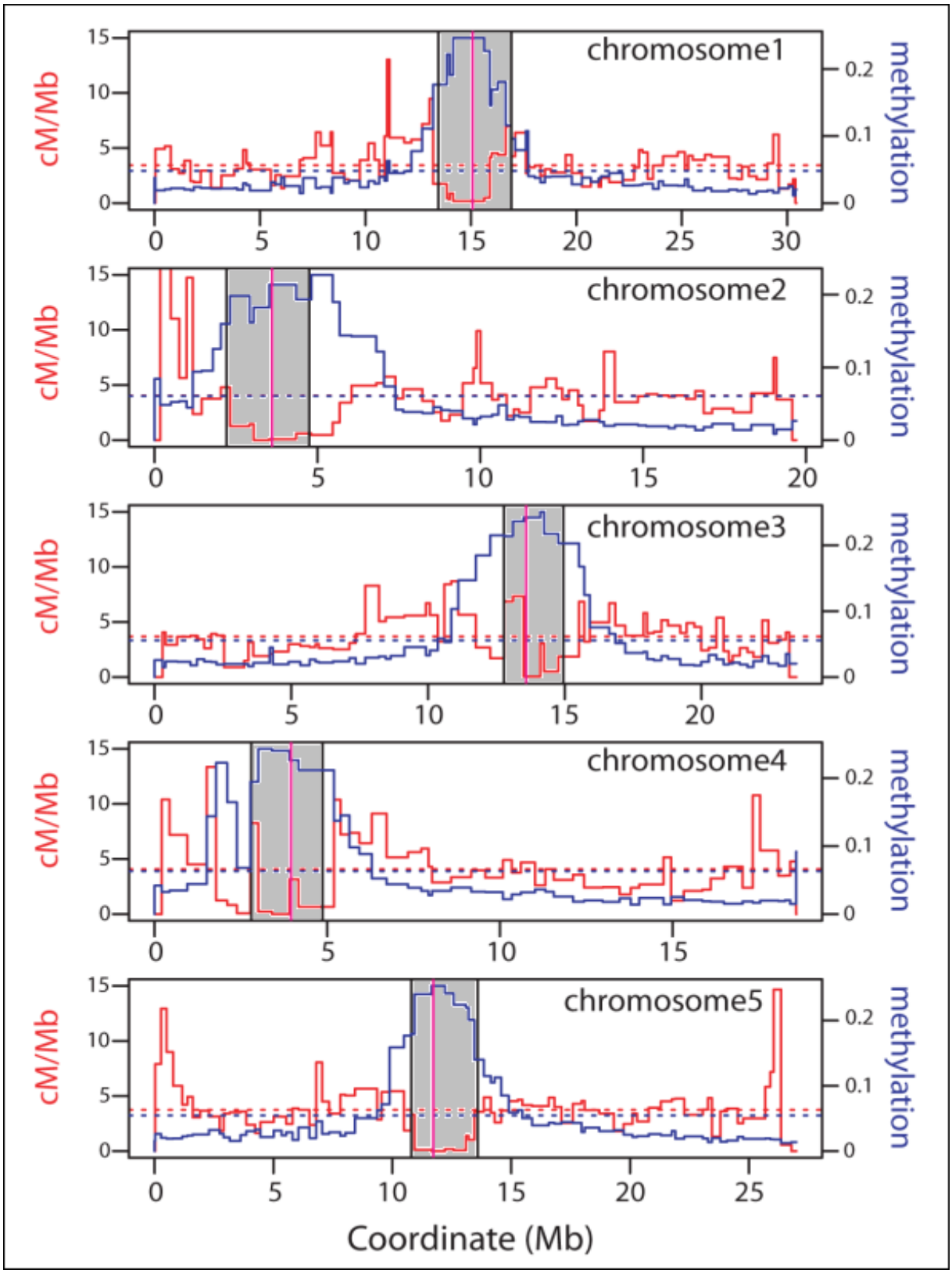


Figure 10. Distribution des évènements de recombinaison (rouge) et de la teneur en 5'-méhtlcytosine (bleu) sur les cinq chromosomes d'*Arabidopsis thaliana* par Yelina *et al.* (2012). Les régions centromériques et péricentromériques apparaissent en grisées.

II.2.3. Par des facteurs épigénétiques

La comparaison de la distribution des évènements de recombinaison entre l'homme et le chimpanzé par Winckler *et al.* (2005) a mis en avant des différences entre ces deux espèces alors que ces dernières présentent près de 99% d'identité de séquence. Ces résultats impliquent l'existence de facteurs autres que génomique. Les modifications épigénétiques, telles que la méthylation ou la modification de la structure de l'ADN jouant un rôle dans la distribution des COs, peuvent expliquer ces différences (Kniewel et Keeney 2009; Petes, 2001).

II.2.3.1. La méthylation de l'ADN

La méthylation de l'ADN, catalysée par des méthyltransférases dans trois sites génomiques avec des motifs symétriques particuliers (CG, CHG et CHH), correspond à la fixation d'un groupement méthyle (CH_3) au niveau du C_5 des bases cytosines (Gruenbaum *et al.* 1981 ; Chan *et al.* 2005). La formation de la 5'-méthylcytosine est ensuite maintenue par différentes protéines telles que DDM1 (Brzeski et Jerzmanowski, 2003). Cette marque épigénétique a été associée à de nombreux processus biologiques dont l'empreinte génomique, la régulation transcriptionnelle des gènes et des TEs, le gène silencing ainsi que dans la distribution des COs (Jost et Saluz 1993; Finnegan *et al.* 2000; Hafiz *et al*. 2001; Martienssen et Colot 2001; Paszkowski et Whitham 2001 ; Petes 2001). Une étude a ainsi démontré que la présence du motif CG méthylé est positivement corrélée à la présence de CO chez l'homme, la souris et le rat (Jensen-Seaman *et al.* 2004). Ces observations indiquent donc que la méthylation favorise la recombinaison ; pourtant cette corrélation est très faiblement retrouvée sur les bras des chromosomes d'*A. thaliana* (R^2=0.021). De plus, celle-ci devient fortement négative dans les régions péricentromériques d'*A. thaliana* (Figure 10) et l'hyperméthylation de ces régions apparait comme une explication supplémentaire à l'absence de recombinaison au centromère (Yelina *et al.* 2012). *A. thaliana* n'est cette fois pas une exception puisque la méthylation inhibe toute présence de CO chez le champignon *Ascobolus immersus* (Maloisel et Rossignol, 1998). Toutefois, il a récemment été observé que l'inactivation de DDM1, provoquant une déficience en méthylation, entraine l'augmentation du taux de recombinaison dans les régions euchromatiques mais pas dans les régions hétérochromatiques (Melamed-Bessudo et Levy, 2012). Ce résultat impliquerait donc l'existence d'une relation entre méthylation et structure de l'ADN dans le contrôle de la régulation et de la distribution des évènements de recombinaison.

II.2.3.2. La modification de la structure de l'ADN

La chromatine est la forme sous laquelle se présente l'ADN dans le noyau. Elle peut être décondensée c'est-à-dire à l'état euchromatique ou condensée, à l'état hétérochromatique. La condensation de la chromatine dépend du nucléosome, complexe protéique composé de quatre histones en doublon : H2A, H2B, H3 et H4, formant un octamère protéique. Celle-ci permet l'enroulement de 147 pb d'ADN conduisant à sa condensation, empêchant ainsi la transcription des gènes présents, bloquant l'accès aux facteurs de transcription et inhibant la formation de COs. La différence de fréquence de recombinaison entre les régions euchromatique et hétérochromatique a pu être observée chez toutes les espèces analysées

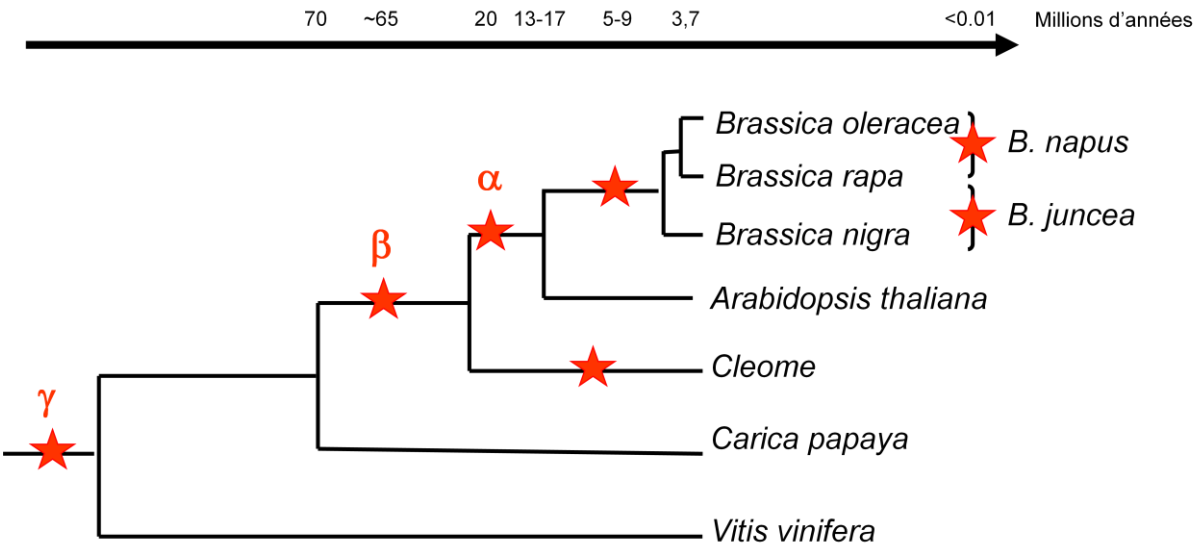


Figure 11. Evènements de duplication des génomes par Jenczewski *et al.* (2013). Les étoiles rouges représentes les évènements de duplication.

comme l'homme, la tomate et *A. thaliana* (Kong *et al.* 2002 ; Gaut *et al.* 2007 ; Melamed-Bessudo et Levy, 2012). Ainsi, la condensation de l'ADN est un facteur clef dans le contrôle de la distribution des évènements de recombinaison. Régulièrement observée dans les régions péricentromériques et télomériques, cette conformation de la chromatine constitue l'une des principales hypothèses avancées pour expliquer l'absence de recombinaison au centromère et la variation de la fréquence de recombinaison aux télomères (Choo, 1998). L'état de la chromatine dépend toutefois de la modification post-traductionnelle des histones liés à l'ADN. La méthylation de ces derniers entraine la condensation de l'ADN ainsi que la diminution drastique de la formation des CDB (Borde *et al.* 2008). Au contraire, leur acétylation décondense l'ADN et entraine une augmentation de la fréquence de recombinaison (Merker *et al.* 2008).

II.2.4. Par des facteurs externes

Borde *et al.* (1999) ont observé que des individus d'une même espèce, identiques pour leur génome, avec pour seule variable l'environnement, présentaient des fréquences de recombinaison très différentes. Cette prise de conscience a conduit à l'identification d'autres facteurs externes à la séquence et à l'épigénétique, affectant la fréquence et la distribution des évènements de recombinaison. L'effet de la température a ainsi été mis en évidence sur diverses espèces (Borner *et al.* 2004 ; Francis *et al.* 2007); l'augmentation de celle-ci entrainant toujours une synthèse supplémentaire de COs. Certains composés chimiques ont également fait l'objet d'étude quant à leur effet sur la recombinaison dans le but de l'améliorer. La cisplatine, le bisphénol A et le RS-1 sont ceux ressortis de ces travaux (Hunt, 2006 ; Jayathilaka *et al*. 2008).

III. LE COLZA, UN MODELE D'ETUDE DE LA RECOMBINAISON

III.1. Généralités sur le colza

Le colza (*Brassica napus* L.) est une espèce annuelle à fleurs jaunes organisées en grappes, dont la pollinisation est principalement autogame avec 10 à 30% d'allogamie fonction des variétés. Appartenant à la classe des dicotylédones et à la famille des Brassicacées, cette crucifère est essentiellement utilisée pour la production d'huile dans l'alimentation ou l'industrie (Biocarburant, lubrifiant, solvant) et de tourteau dans l'alimentation animale après extraction de ses 40 à 42% d'huile (PROLEA, 2013). Le colza est à ce jour la première espèce oléagineuse en Europe et la quatrième au niveau mondial. Elle est cultivée sur plus de 32 millions d'hectares pour une production mondiale s'élevant à 64,8 millions de tonnes dont 8,43% rien que par la France (FAOSTAT, 2012).

III.2. Origine et sélection

Les *Brassica* semblent tous dériver d'un même génome ancestral, à n=6 chromosomes ayant connu un phénomène de triplication de son génome (Figure 11) (Lagercrantz et Lydiate, 1996 ; Parkin *et al*. 2003 ; Lysak *et al*. 2005 ; Wang *et al.* 2011). *Brassica napus* (AACC, 2n=4x=38) est lui le produit de l'hybridation interspécifique naturelle de *B. rapa* (AA, 2n=2x=20) et de *B. oleracea* (CC, 2n=2x=18) mais son origine fait encore

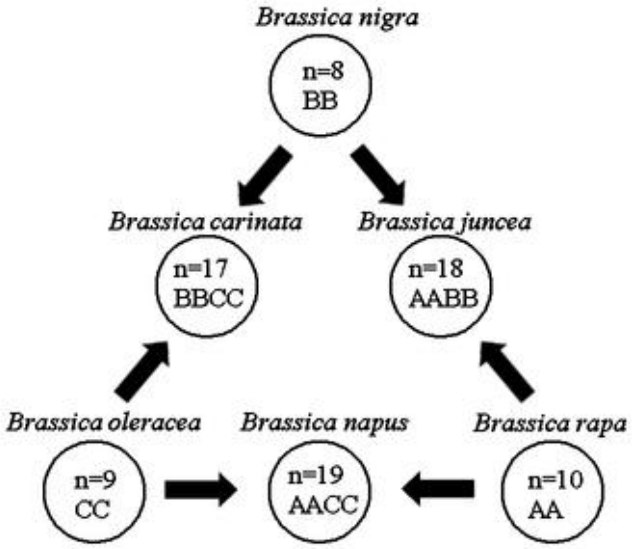


Figure 12. Relation entre les génomes des Brassicas (U, 1935)

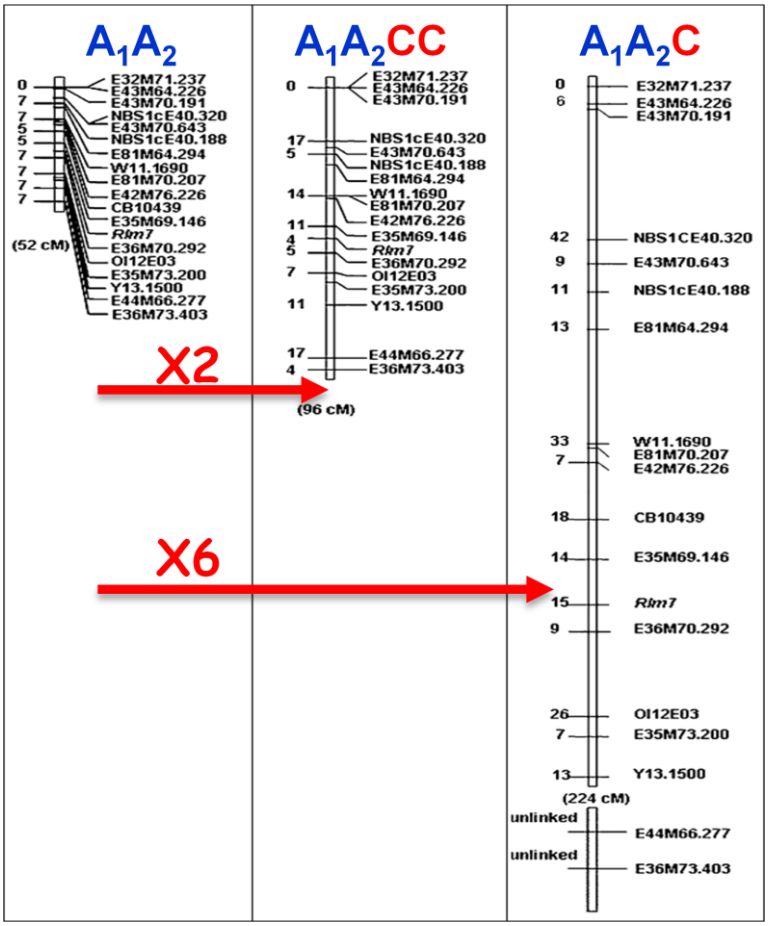


Figure 13. Cartographie de populations issues d'un hybride diploïde AA, triploïde AAC et tétraploïde AACC par Leflon *et al.* (2010)

débat (Figure 12) (U, 1935). Le colza se serait formé il y a moins de 10.000 ans soit dans le pourtour méditerranéen, zone de coexistence de ses progéniteurs (Rana *et al.* 2004), soit dans des jardins au moyen âge (Gomez-Campo et Prakash, 1999). Il n'existe que des variétés cultivées de colza et peu de variabilité au sein de cette espèce (Delourme *et al*. 2013). Celui-ci a en effet été sélectionné principalement pour sa qualité (baisse teneur en glucosinolates et acide érucique, qualité de l'huile) et a ainsi subi un fort goulot d'étranglement. Partant de ce constat, la recherche actuelle s'oriente vers des croisements interspécifiques entre le colza et ses progéniteurs afin de réintroduire de la diversité (Doré et Varoquaux, 2006).

III.3. La recombinaison chez le colza

III.3.1. La recombinaison homéologue

Résultant de l'hybridation interspécifique de deux espèces, *B. napus*, est un allotétraploïde. Chez cette forme de polyploïde, les recombinaisons peuvent se produire non plus uniquement entre chromosomes homologues mais également entre homéologues (i.e. provenant de génomes différents). Lors de la resynthèse d'hybrides de colza, il a récemment été observé que malgré la divergence des génomes A et C, ces derniers sont suffisamment proches pour recombiner entre eux (Gaeta, 2010 ; Szadkowski *et al.* 2010 ; 2011). En revanche chez les colzas naturels, les appariements homéologues sont sous contrôle génétique. Ainsi un QTL majeur, nommé *PrBn*, sur le chromosome C_9, trois QTL mineurs sur les chromosomes C_1, C_6 et A_3 ainsi que des relations épistatiques ont été identifiés comme régulant l'appariement et les recombinaisons entre chromosomes homéologues chez des haploïdes de colza (Jenczewski *et al.* 2003, Liu *et al.* 2006 ; Nicolas *et al*. 2009).

III.3.2. La recombinaison homologue

Le contrôle génétique de la recombinaison homéologue pourrait également influencer le nombre de COs entre chromosomes homologues. En effet, des hybrides AAC produits par croisement entre une navette et la variété de colza Darmor (à fort taux d'appariement au stade haploïde) présentent un plus fort taux de recombinaison homologue entre les chromosomes A que ceux obtenus à partir de la variété Yudal (à faible taux d'appariement au stade haploïde) (Nicolas *et al.* 2009). Par la suite, Leflon *et al*. (2010) ont observé qu'à même fond génomique, le taux de recombinaison entre les chromosomes A07 augmente d'un facteur 6 chez des hybrides triploïdes (AAC) par rapport aux diploïdes (AA) (Figure 13). L'immunolocalisation des protéines MLH1 sur les bivalents a permis de confirmer un nombre plus élevé de CO I chez les triploïdes par rapport aux diploïdes sans toutefois expliquer la totalité des COs observés à partir des cartes génétiques impliquant une augmentation des CO I et II. Il a donc été conclu que le génome C sous la forme haploïde régule la recombinaison en favorisant la formation de COs entre les chromosomes A. La capacité des triploïdes à transmettre leurs chromosomes C à la descendance selon une loi binomiale (Leflon *et al*. 2006) a permis de générer diverses combinaison de diploïdes AA avec des chromosomes C en addition dont le nombre et la nature varient. Les résultats obtenus ont révélé que la nature des chromosomes C en addition, et surtout de C_9, est responsable de l'augmentation de la recombinaison entre les chromosomes A07, A04 et A02 et que celle-ci entraine également une forte diminution de l'intensité de l'interférence par rapport aux diploïdes AA (Suay *et al.*

2013). Cependant, le chromosome C_9 seul ne suffit pas à expliquer le taux de recombinaison observé lorsque les 9 chromosomes C sont présents, impliquant un effet des autres chromosomes C.

IV. OBJECTIFS ET STRATEGIES DE L'ETUDE

Il a donc été établi que l'addition du génome C sous forme haploïde ou du chromosome C_9 chez des diploïdes AA régule favorablement la recombinaison homologue entre chromosomes A. Cependant, le lien entre la fréquence de recombinaison et la distribution des COs le long des chromosomes reste à déterminer. Les séquençages récents des génomes de *B. rapa* et de *B. napus* permettent maintenant de réaliser ce lien (Wang *et al.* 2011 ; Chalhoub *et al.* en prep). En nous appuyant sur celles-ci et sur les populations obtenues par Suay *et al*. (2013), nous nous sommes fixés pour objectif d'étudier la distribution des évènements de recombinaison entre les chromosomes A07 chez des diploïdes AA de *B. rapa* et de déterminer l'effet de l'addition du chromosome C_9 et des 9 chromosomes C. Nous avons ainsi formulé deux hypothèses : (1) la structure génomique du chromosome influence la répartition des COs le long d'A07 et (2) les 9 chromosomes C et C_9 en particulier, affectent cette distribution.

Pour y répondre, nous nous attacherons à (1) étudier la structure du chromosome A07 en le comparant chez *B. rapa* et *B. napus* et en localisant son centromère, (2) densifier des régions présentant des variations de fréquence de recombinaison entre nos populations AA, AA+1C_9 et AA+9C, (3) étudier la distribution des évènements de recombinaison chez ces populations, (4) déterminer les facteurs génomiques comme la teneur en gènes, TEs, %GC, affectant la distribution de ces évènements, et (5) Mesurer l'interférence le long du chromosome A07 chez nos populations afin de déterminer quelle voie de synthèse des COs est affectée.

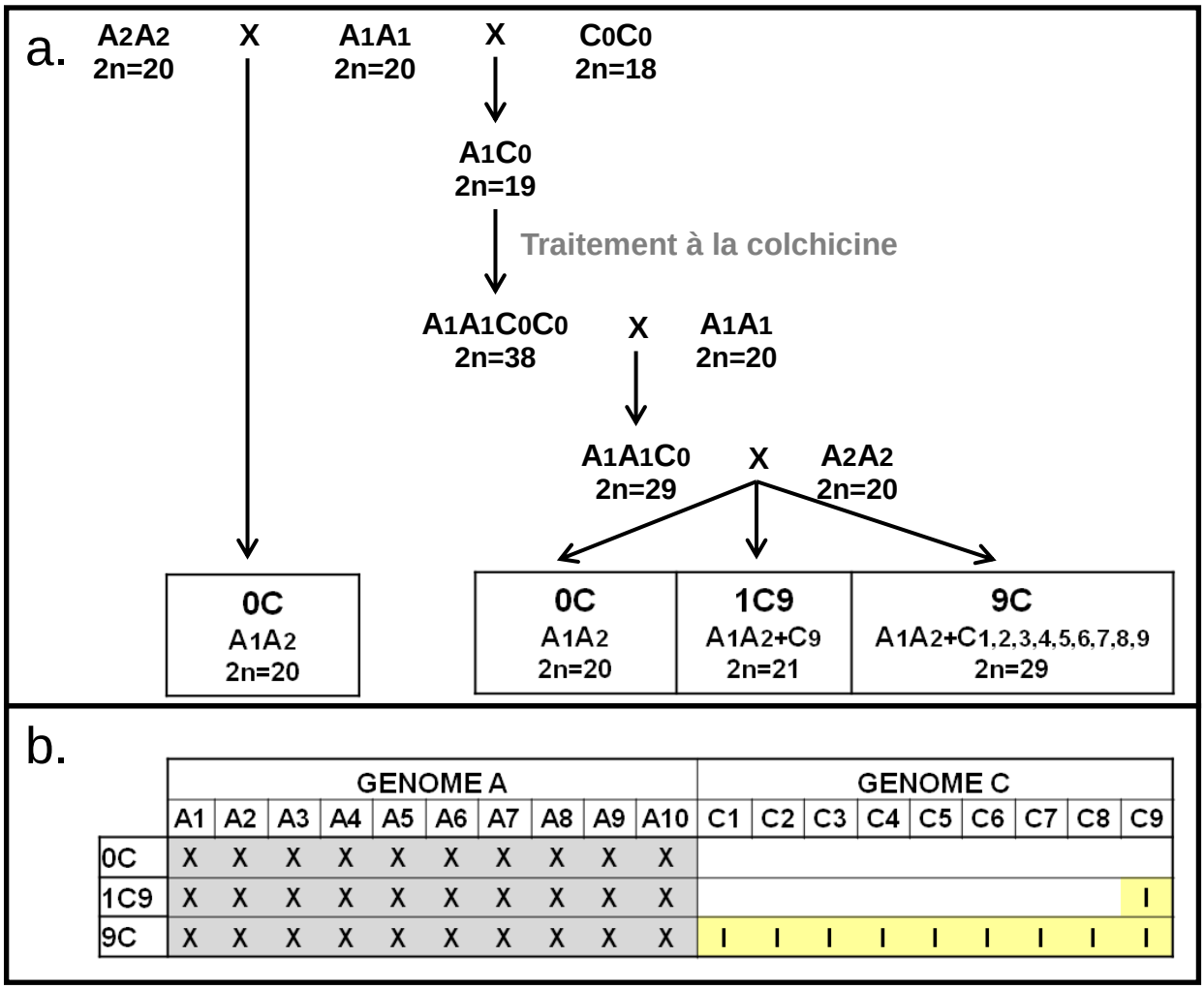


Figure 14. a/ Représentation schématique de la production des hybrides. **b/** Représentation schématique de la composition chromosomique des hybrides. A_1A_1 et A_2A_2 sont des cultivars de *Brassica rapa* (C1.3 et Chiifu) et C_0C_0 est un cultivar de *Brassica oleracea* (RC34). Les **X** représentent les bivalents et les **I** les univalents.

CHAPITRE II : MATERIEL ET METHODES

I. MATERIEL VEGETAL

Les populations utilisées au cours de ce stage ont été obtenues avant mon arrivée pour réaliser l'article de Suay *et al.* (2013). La stratégie développée pour produire les hybrides est détaillée sur la figure 14. Dans un premier temps, une plante de *Brassica oleracea* var. *alboglabra* (RC34), une homozygote haploïde doublée (C*0*C*0*, 2n=18), a été croisée comme femelle avec une plante de *Brassica rapa* var. *chicon* (C1.3), une variété française hétérozygote de navette fourragère (A_1A_1, 2n=20). Le résultat de ce croisement, un hybride allohaploïde A_1C_0, a été traité avec de la colchicine pour resynthétiser un allotétraploïde de colza comme réalisé par Chèvre *et al*. (1989). Un backcross a ensuite été opéré entre cet allotétraploïde utilisé comme femelle et la navette diploïde C1.3 pour synthétiser cette fois des hybrides triploïdes $A_1A_1C_0$ (2n=29). Lorsque les chromosomes C sont sous forme haploïde chez un hybride, leur transmission à la descendance présente la spécificité de suivre une loi binomiale (Leflon *et al.* 2006). Ainsi, de par cette caractéristique, le croisement du triploïdes $A_1A_1C_0$ avec un diploïde A_2A_2 a permis d'obtenir des diploïdes avec des chromosomes C en addition dont le nombre et la nature varient. Le diploïde A_2A_2 utilisé comme plante mâle dans ce croisement est une variété de choux chinois, *Brassica rapa* var. *chiifu*. Son utilisation a été choisie afin de réaliser un lien physique avec la fréquence de recombinaison car son génome a récemment été séquencé (Wang *et al.* 2011). Les 202 plantes produites ont été analysée par cytométrie de flux pour déterminer le nombre de chromosomes C en addition et leur nature a été identifiée ensuite à l'aide de marqueurs spécifiques. Trois hybrides issus de ce croisement ont été sélectionnés A_1A_2 + 0C (0C) (Témoin), A_1A_2 + $1C_9$ ($1C_9$) et A_1A_2 + 9C (9C). Afin d'établir des cartes génétiques à partir de ces individus, ces derniers ont été croisés comme plantes femelles avec la variété Darmor de colza générant ainsi 64 plantes pour 0C et 91 plantes pour $1C_9$ et 9C. Dans le but d'augmenter la taille de la population du témoin 0C, un croisement direct a été opéré en parallèle entre C1.3 et Chiifu. Le produit de ce croisement, un diploïde A_1A_2 (2n=20), a également été croisé par Darmor, permettant ainsi d'obtenir 110 plantes supplémentaires.

II. ANALYSES BIOINFORMATIQUES

II.1. Etude de la structure du chromosome A07 de *B. rapa*

Les séquences du chromosome A07 de *Brassica rapa* var. Chiifu v1.5 disponible sur Brassica Database (http://brassicadb.org/brad/) et de *Brassica. napus* var. Darmor v1.2 disponible en interne (Chalhoub *et al*. en prep.) ont été alignées entre elles à l'aide du package NUCmer de la suite de programme MUMmer v3.0 sous Linux (Kurtz *et al*. 2004). Avant de générer le résultat de ce blast sous forme de Dot-plot, un filtre a été appliqué avec les paramètres -i 95 -l 1000, afin de n'aligner que des séquences d'un minimum de 1000 pb avec au moins 95% d'homologie. Les scaffolds (i.e. assemblage de contigs) de *Brassica rapa* v1.5 ont également été alignés sur le chromosome A07 de *B. rapa* v1.5 via NUCmer avec les mêmes paramètres pour déterminer les régions bien séquencées et assemblées lors du séquençage du génome.

Tableau 3. Séquences Forward et Reverse des marqueurs utilisés par Suay *et al.* (2013) sur le chromosome A07

Marqueurs	Forward	Reverse
NMS1168	TGAGCACATAAAGTGGTTCAGC	CTGAACAAGGACCCGAAGAA
sR4047	AGAGAGGGTGTTGTTCCCAG	ATTCTTTGAATGGCCAGCAC
cnu_m062a	ATCGGCGCTGGTTATGTCA	CTAGGCTGCCCTTTCCGATT
sR0282R	AGGAAGCCCAACAGGACTTT	AATTCGATTCTCCATCGTGC
NMS0956	AACCATCGCGGTAATTAAACC	ACGACTCCAGCTCAGCATCT
Bras023	CCATTGCAAATCCCTTTACTC	AGATGGAATGCGATCAAAGA
nia_m063a	GAAGAAACTCGGTGGGGAGT	AAAGAGTTCCGAAAATGGGC
cnu_m273a	ATAAGGGCATCGCCTCAACA	TGCACGCATCCACATAAACA
sNRA59	CAGATTCGGATTTGGGAAGA	GGCGGAAGAATCAAAGGAGT
cnu_m295a	GCTGCCTAATAGGGTGCTTG	AGAGCGCATTCAAGTCTGGT
nia_m030a	GATCATAAGCCGAAAAAGGTTG	TGCTCTCCTCAAGTGAATCAAA
sR12387	GGGTCTGGGTTTTTCTGTGA	GATTGGGCCGTGTAATATCG

Tableau 4. Caractéristiques des motifs SSR

Caractéristiques	
Taille minimale du motif	2
Taille minimale du motif	4
Nombre de répétition minimale du motif	3
Longueur minimale du microsatellite	12
Poly N	Non

Tableau 5. Caractéristiques des couples d'amorces générés par primer3

Caractéristiques	
Taille minimale du produit amplifié	150 pb
Taille optimale des primers	20 pb
Taille minimale des primers	18 pb
Taille maximale des primers	25 pb
Température optimale d'hybridation	57°C
Température minimale d'hybridation	54°C
Température maximale d'hybridation	60°C
Pourcentage optimale en base G et C	44 %
Taille minimal de l'amplicon	150 pb
Taille minimal de l'amplicon	300 pb
Logiciel Design	Primer3

II.2. Localisation physique du centromère du chromosome A07 de *B. rapa*

Les séquences centromérique, CentrBr1 (Genbank accession n° EU294378.1), et péricentromérique, pBct (Genbank accession n° BZ613967), spécifiques du chromosome A07 de *B. rapa* ont été identifiées dans des revues antérieures (Lim *et al*. 2005 ; 2007 ; Koo *et al.* 2011). Celles-ci ont été alignées sur la séquence du chromosome A07 de *B. rapa* v1.5 par un BLASTn (pour Basic Local Alignement Search Tool) et nous n'avons conservé que les alignements avec e-value < 10^{-10} et dont le pourcentage d'alignement global est supérieur à 80%. Pour confirmer les résultats obtenus par ces deux blastn, nous avons localisé la position des blocs H et G de Schranz qui se situent en amont et en aval du centromère du chromosome A07 de *B. rapa* (Cheng *et al.* 2013). Ces blocs définis à partir d'*A. thaliana* correspondent à des régions hautement conservées regroupant un grand nombre de gènes. Les orthologues de ces gènes ont été identifiés par Wang *et al.* (2011) et positionnés le long du génome de *B. rapa*, disponible sur Brassica Database (2010). Connaissant les gènes constituant les blocs H et G chez *A. thaliana* (Schranz *et al.* 2006), nous avons pu déterminer leurs orthologues chez *B. rapa* ainsi que leur position.

II.3. Localisation physique des SSR établis sur le chromosome A07 de *B. rapa*

Les marqueurs microsatellites cartographiés par Suay *et al*. (2013) sont tous issus de publications antérieures. Les marqueurs présentant les préfixes cnu, nia, CB et Bras sont issus de la publication de Piquemal *et al.* (2005) et ceux avec comme préfixes sN et sR proviennent de Agriculture and Agri-Food Canada (http://www.agr.gc.ca/)(Tableau 3). Les séquences des amorces des marqueurs cartographiés par Suay *et al.* (2013) ont été alignées avec un BLASTn et des paramètres stringents sur la séquence du chromosome A07 de *B. rapa* v1.5. En recoupant les informations acquises lors de l'étude de la structure du chromosome A07 et de la localisation des SSR, nous avons défini la position physique des zones à densifier.

II.4. Design de SSR

Nous avons choisi de réaliser une densification avec des marqueurs SSR (Simple Sequence Repeat), correspondant à des motifs de 2 à 4 nucléotides répétés en tandem, car ceux-ci sont localisables physiquement, codominants, très polymorphes et bien répartis sur l'ensemble du génome. Nous avons ainsi recherché des motifs dont les caractéristiques sont détaillées dans le tableau 4, dans les régions définies pour être densifiées via les données de séquence de *B. rapa* v1.5. Leur détection a été permise via le workflow « Microsat » développé par Gille Lassalle au sein de l'équipe bioinformatique de l'INRA de Rennes. Ce dernier exploite le logiciel "Primer3" permettant de générer des couples d'amorces autour de ces motifs dont les caractéristiques choisies sont détaillés dans le tableau 5 (Rozen et Skalestsky, 2000). Les différents couples d'amorces obtenus sont ensuite testés par PCR « *in silico* » sur les données de séquence de *B. rapa* v1.5 et de *B. napus* v1.2. Celle-ci permet d'aligner les amorces par un BLASTn stringent afin d'éliminer (1) les profils multibandes et (2) les marqueurs non polymorphes entre Chiifu et Darmor (i.e. avec un amplicon de même taille). Suite à ce criblage, les marqueurs potentiels ont été sélectionnés sur le nombre de répétition du motif (>3) et sur leur positionnement dans nos zones afin d'avoir au moins un SSR tous les 250 kb.

III. ANALYSES BIOMOLECULAIRES

III.1. Extraction ADN

Trois à quatre jeunes feuilles ont été prélevées, lyophilisées puis broyées pour chaque individu des populations 0C, $1C_9$ et 9C et les parents Darmor, C1.3, Chiifu et RC34. L'ADN de chaque échantillon a ensuite été extrait à l'aide du kit d'extraction NucleoSpin® Plant II Maxi, dont le protocole est décrit par Lombard et Delourme (2001) et détaillé en Annexe I. Enfin, les solutions d'ADN ont été dosées puis diluées à une concentration finale de 5 ng/µL.

III.2. Amplification

Des amplifications PCR ont été conduites pour chacun des couples d'amorces sélectionnés par bioinformatique, afin de tester le polymorphisme entre les parents Darmor, C1.3, Chiifu et RC34 (Annexe II). Les SSR polymorphes chez les parents ont été de nouveau amplifiés par PCR chez les individus des populations 0C, $1C_9$ et 9C.

III.3. Lecture SSR

La mesure de l'intensité des fragments amplifiés a été réalisée par l'analyseur de fragments 16 capillaires ABI PRISM® 3130xl décrit par Esselink *et al*. (2004). Cette analyse a nécessité la fixation préalable d'une queue M13 sur les amorces forward pour permettre une révélation par immunofluorescence des amplicons (Schuelke, 2000). Les résultats obtenus ont été traités à l'aide du logiciel GeneMapper™ Software v3.7 (Applied Biosystems). Ce dernier nous a d'abord permis de révéler les marqueurs polymorphes qui, compte tenu du schéma de croisement, doivent présenter des amplicons de tailles différentes entre C1.3 et Chiifu mais être absents ou différents chez Darmor et RCC. Puis, les marqueurs présentant un bon profil de polymorphisme ont permis le génotypage des populations 0C, $1C_9$ et 9C afin d'identifier, pour chaque individu, la version C1.3 ou Chiifu présentée par le marqueur.

IV. ANALYSES STATISTIQUES

IV.1. Etude de la ségrégation allélique des marqueurs

La ségrégation allélique de chacun des marqueurs pour les populations 0C, $1C_9$ et 9C à été déterminée en réalisant un test de χ^2 sous l'hypothèse H_0, la ségrégation est de type mendélienne 1:1 entre les allèles des parents C1.3 et Chiifu:

$$\chi^2 = \frac{\sum(\text{Fréquences Observées} - \text{Fréquences Théoriques})^2}{\text{Fréquences Théoriques}}$$

Nous avons fixé notre seuil α à 5% correspondant à une valeur de 3,84 d'après la table de χ^2 pour un degré de liberté (ddl) de 1. Ainsi l'hypothèse H_0 est vérifiée uniquement lorsque la valeur du χ^2 obtenue pour un marqueur au sein d'une de nos populations est strictement inférieure à 3,84. Les marqueurs présentant un $\chi^2>3,84$ sont considérés comme étant en distorsion et ne sont utilisés pour la cartographie qu'après ceux présentant une ségrégation 1:1.

IV.2. Cartographie génétique

Parmi les marqueurs présentant une ségrégation mendélienne, seuls ceux éloignés d'un minimum de 100 Kb ont été retenus pour ne pas fausser les résultats lors de l'étude des corrélations liées à la structure génomique. Les données issues du génotypage de ces marqueurs ont ensuite été utilisées pour établir les cartes génétiques des populations 0C, $1C_9$ et 9C à l'aide du logiciel Mapmaker v3.0 qui accepte des données de type backcross (Lander *et al.* 1987). La liaison génétique entre deux locus est testée par un test de maximum de vraisemblance en fonction du LOD score minimum (Log of the Odds Ratio, décrit par Morton en 1955) et du taux maximum de recombinaison : $\text{LOD} = log10\left(\frac{e^L(\hat{\theta})}{e^L(\theta_0)}\right)$, où $e^L(\hat{\theta})$ représentant la vraisemblance maximale évaluée à $\hat{\theta}$ et $e^L(\theta_0)$ la vraisemblance évalué avec la fonction "d" de Kosambi (1944) prenant en compte l'interférence : $\text{d} = \frac{1}{4}log10\left(\frac{1+2\theta}{1-2\theta}\right)$. Les groupes de liaison ont été obtenus à l'aide de la fonction "group" en utilisant une valeur de LOD de 4 et une distance maximum de 25 cM. La fonction "order" a ensuite permis d'identifier l'ordre de ces marqueurs et les autres ont été ajoutés avec la fonction "try" quand ceux-ci présentaient un LOD>3. Les derniers marqueurs dont le LOD est inférieur à 3 ont été positionnés en nous appuyant sur les données physiques de l'ordonnancement des marqueurs obtenus lors du design des SSR le long du chromosome A07.

IV.3. Etude de l'hétérogénéité des fréquences de CO

Dans le but d'étudier l'hétérogénéité des fréquences de recombinaison entre nos populations, nous avons réalisé des tests de χ^2 pour chaque intervalle entre les marqueurs cartographiés et pour le groupe de liaison A07 sous l'hypothèse H0, absence de différence entre les taux de COs. Entre chaque intervalle le ddl est de 1, et sur le groupe de liaison il est fixé au nombre de marqueurs -1. Enfin, via la fonction LOI.KHIDEUX(x,d) sous Excel, avec x les résultats du test χ^2 et d le ddl, nous avons obtenu les p-valeurs permettant ainsi de comparer les taux de recombinaison pour le groupe de liaison A07 et pour chaque intervalle entre nos populations.

IV.4. Etude des corrélations liées à la structure génomique

Afin d'établir des corrélations avec la structure génomique du chromosome A07, nous avons dans un premier temps calculé pour chaque intervalle entre nos différents marqueurs la fréquence de recombinaison en cM/Mb ainsi que les teneurs en GC (%GC), gènes (Nb gène/Mb) et TEs (Nb TEs/Mb). Nous avons divisé la taille génétique (cM), le nombre de gène et le nombre de TEs entre chaque marqueur par la taille physique de l'intervalle (Mb). Si les tailles génétiques proviennent de la cartographie, le nombre de gènes et de TEs sont eux issus respectivement des banques Brassica database et Repbase (http://www.girinst.org/repbase/). Les séquences des TEs, provenant exclusivement des banques d'*A.thaliana* et des plantes dicotylédones, ont été préalablement alignées sur la v1.5 de *B. rapa* par un BLASTn en ne conservant que les alignements avec une e-value < 10^{-10} et en éliminant les doublons (correspondant aux fragments d'un même TEs identifiés à la même position physique). La teneur en GC résulte, elle, du simple pourcentage de base GC compris entre chaque marqueur par rapport au nombre de base totale. Dans un second temps, nous avons établi des

corrélations entre la fréquence de recombinaison et les différentes teneurs mesurées. Le coefficient de détermination (R^2) a ainsi été obtenu à partir des droites de régression. Celui-ci correspond à une mesure de la précision de l'ajustement de la droite de régression. Il s'agit du rapport entre la variation de la variable dépendante (mesurée par sa variance) expliquée par le modèle de régression et sa variation totale (mesurée par sa variance). Le R^2 correspond au carré du coefficient de corrélation (r) de Pearson mesuré par : $r = \frac{\sum(x-\bar{x})(y-\bar{y})}{\sqrt{\sum(x-\bar{x})^2 \sum(y-\bar{y})^2}}$

Où $\bar{x}$ est la moyenne de x et $\bar{y}$ est la moyenne de y.

Enfin, un test de corrélation de Pearson a permis de déterminer les p-valeurs de ces corrélations sous R.

Nous avons également cherché à établir différents gradients le long du bras long du chromosome A07. Pour ce faire nous avons pour chaque intervalle, du centromère aux télomère, fixé sa position sur le chromosome A07 à celle du milieu de ses 2 marqueurs adjacents. Puis avec le même procédé d'analyse que les corrélations, nous avons ensuite déterminé l'existence de gradient entre teneur en GC, TEs, gènes et fréquence de recombinaison avec leur position le long du bras long du chromosome A07.

IV.5. Mesure de l'interférence

Afin de mesurer l'interférence sur le chromosome A07 à partir des données issues du génotypage, nous avons utilisé le logiciel CODA (Crossover Distribution Analyzer) (http://moulon.inra.fr/index.php/fr/production-scientifique/logiciels/coda) (Gauthier *et al*. 2011). Celui-ci s'appuie sur le modèle GS (Two Pathway Gamma-Sprinkling Model) tenant compte de l'existence de deux types de COs ayant des distributions différentes (Falques *et al*. 2009). Il est ainsi possible d'étudier l'interférence en simulant des COs de voie I ou des COs de voie I et II le long d'un chromosome pour les comparer aux données expérimentales. Pour le modèle à 1 voie, les distances entre COs I successifs (i.e. COs interférents dont la distribution n'est pas aléatoire), sont des variables aléatoires, indépendantes et identiquement distribuées suivant la loi Gamma de paramètre v correspondant à la force de l'interférence (McPeek *et al.* 1995) : $\rho_v(d) = (2v)^v d^{v-1} \frac{e^{-2dv}}{\Gamma(v)}$. Pour le modèle à 2 voies de CO, les COs I sont simulés selon une loi gamma et une fraction *p* du nombre total de CO est estimée comme étant des COs de voie II (i.e. COs non interférents dont la distribution est aléatoire). Les distances entre COs II successifs sont distribuées selon une loi de poisson de paramètre λ (Copenhaver *et al.* 2002) : $p(k) = e^{-\lambda} \frac{\lambda^k}{k!}$

Nous avons ainsi estimé la force de l'interférence pour les modèles à 1 voie de CO (nu_{eff}) et à 2 voies de CO (nu) ainsi que *p*, la fraction de CO II pour le modèle à 2 voies, pour les populations 0C, $1C_9$ et 9C. Les intervalles de confiance inférieure et supérieure ont été obtenus en utilisant une approche re-simulation comme décrit dans (Falque *et al.* 2009) en utilisant 1000 re-simulations. L'interprétation de l'intensité d'interférence à quand à elle été établie grâce à la valeur du paramètre v. Quand celui-ci est égal à 0, il y a absence totale d'interférence et l'intensité de cette dernière augmente à mesure que la valeur de v croit.

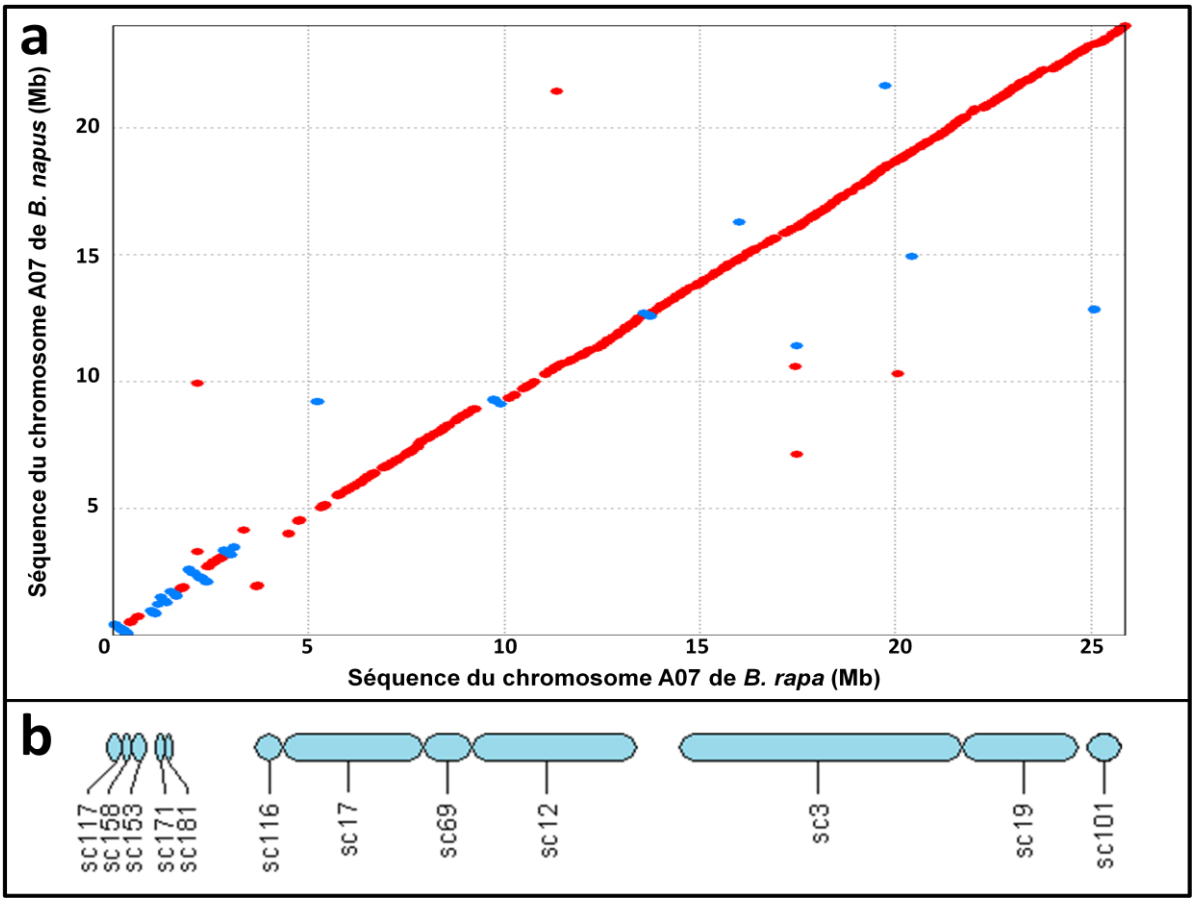


Figure 15. Structure du chromosome A07 de *B. rapa*. **a/** Dot-plot des séquences des chromosomes A07 de *B. rapa* v1.5 (abscisse) et de *B. napus* v1.2 (ordonnée). **b/** Ordonnancement physique des scaffolds du chromosome A07 de *B. rapa* v1.5

Tableau 6. Blastn entre la séquence centromérique CentrBr1 et le chromosome A07 de *B. rapa*

Séquence	Chr.	% identité globale	Alignement (pb)	Position début (pb)	Position fin (pb)	e-value
CentBr1	A07	93.18	173	6493060	6493232	6.00E-72
CentBr1	A07	93.18	176	3591426	3591251	1.00E-66
CentBr1	A07	92.61	173	3591602	3591430	1.00E-69
CentBr1	A07	92.05	176	3591074	3590899	8.00E-62
CentBr1	A07	92.05	176	3591250	3591075	8.00E-62
CentBr1	A07	92.04	173	4852570	4852398	3.00E-67
CentBr1	A07	84.66	166	4863928	4863763	1.00E-48
CentBr1	A07	84.09	166	4861817	4861652	3.00E-46
CentBr1	A07	84.09	166	4862345	4862180	3.00E-46
CentBr1	A07	84.09	166	4862521	4862356	3.00E-46
CentBr1	A07	84.09	166	4862873	4862708	3.00E-46
CentBr1	A07	84.09	157	6493236	6493392	2.00E-62
CentBr1	A07	83.52	166	4863049	4862884	6.00E-44
CentBr1	A07	83.52	166	4863576	4863412	2.00E-41
CentBr1	A07	82.95	166	4861641	4861476	2.00E-41
CentBr1	A07	82.95	166	4862169	4862004	2.00E-41
CentBr1	A07	82.95	166	4863752	4863587	2.00E-41
CentBr1	A07	82.39	166	4861993	4861828	4.00E-39
CentBr1	A07	82.39	166	4863401	4863236	4.00E-39
CentBr1	A07	81.82	166	4862697	4862532	9.00E-37
CentBr1	A07	81.82	166	4863225	4863060	9.00E-37

Tableau 7. Blastn entre la séquence péricentromérique pBct et le chromosome A07 de *B. rapa*

Séquence	Chr.	% identité globale	Alignement (pb)	Position début (pb)	Position fin (pb)	e-value
pBct	A07	90.06	471	3711677	3711207	0
pBct	A07	84.79	469	5127275	5127743	0

CHAPITRE III. RESULTATS

I. ETUDE DE LA STRUCTURE DU CHROMOSOME A07 DE *B. RAPA*

I.1. Comparaison des chromosomes A07 de *B. rapa* et de *B. napus*

Les résultats de l'alignement des séquences des chromosomes A07 de *B. rapa* et de *B. napus* sont présentés sous forme de Dot-plot sur la figure 15.a/. Sur cette représentation, chaque point schématisé en rouge ou bleu représente un alignement d'au moins 1000 pb avec 95% d'homologie. La couleur indique elle, le sens de l'alignement qui peut être conservé (en rouge) ou inversé (en bleu). Nous avons ainsi pu observer sur cette figure d'importantes similitudes de séquence entre ces deux espèces. Quelques points de rupture ont cependant été identifiés, notamment au début du chromosome, pouvant indiquer des différences de séquences entre les deux Brassicas ou des difficultés d'assemblage lors du séquençage des génomes. Afin de connaître la qualité de l'assemblage de ce chromosome, nous avons réalisé un alignement contre lui des séquences de l'ensemble des scaffolds de *B. rapa* (i.e. assemblage de contigs). Les résultats, schématisés sur la figure 15.b/, positionnent les 12 scaffolds identifiés le long du chromosome. Nous avons ainsi pu observer que le début de ce chromosome, dont la taille totale est estimée à 25,88 Mb, n'est actuellement pas « bien séquencé » et pourrait expliquer les points de ruptures identifiés entre *B. rapa* et *B. napus*.

I.2. Localisation du centromère d'A07

Les résultats des alignements des séquences centromérique et péricentromérique sur le chromosome A07 de *B. rapa* sont présentés dans les tableaux 6 et 7. Ces BLASTn nous ont permis de localiser CentBr1 en trois positions à 3,59 Mb, 4,86 Mb et 6,49 Mb et pBct en deux positions à 3,71 et 5,13 Mb. Cette dernière séquence, bordant de part et d'autre le centromère, nous a permis de délimiter l'intervalle des régions péricentromériques entre 3,71 et 5,13 Mb. La présence de CentBr1 dans cet intervalle à 4,86 Mb semble donc indiquer l'emplacement réel du centromère. Dans le but de conforter ces observations, nous avons situé les blocs de Schranz H et G identifiés autour du centromère du chromosome A07 par Cheng *et al.* (2013) (tableau 8). Nous avons ainsi observé que le bloc H s'étend sur A07 de 0,01 à 4,04 Mb et que le bloc G est compris entre 4,78 et 5,15 Mb. L'intervalle entre ces blocs défini entre 4,04 et 4,78 Mb confirme les résultats obtenus précédemment avec pBct mais ne nous permet de garantir la position exacte du centromère. Nous avons donc choisi de délimiter les régions péricentromériques du chromosome A07 de *B. rapa* dans leur intervalle le plus large, c'est-à-dire entre 3,71 et 5,13 Mb et conclu que ce chromosome est acrocentrique.

Tableau 8. Position des marqueurs des Blocs H et G de Schranz sur le chromosome A07 de *B. rapa*

Blocs de Schranz	Marqueur *B. rapa* début	Marqueur *B. rapa* fin	Position début (pb)	Position fin (pb)
H	Bra036508	Bra036452	13 172	377 792
H	Bra039035	Bra038813	439 805	1 166 722
H	Bra028091	Bra028099	1 186 345	1 295 248
H	Bra039603	Bra039893	1 323 148	1 823 775
H	Bra002100	Bra001953	1 835 204	3 168 349
H	Bra015252	Bra015189	3 176 735	4 041 577
G	Bra015184	Bra015160	4 777 186	5 145 236

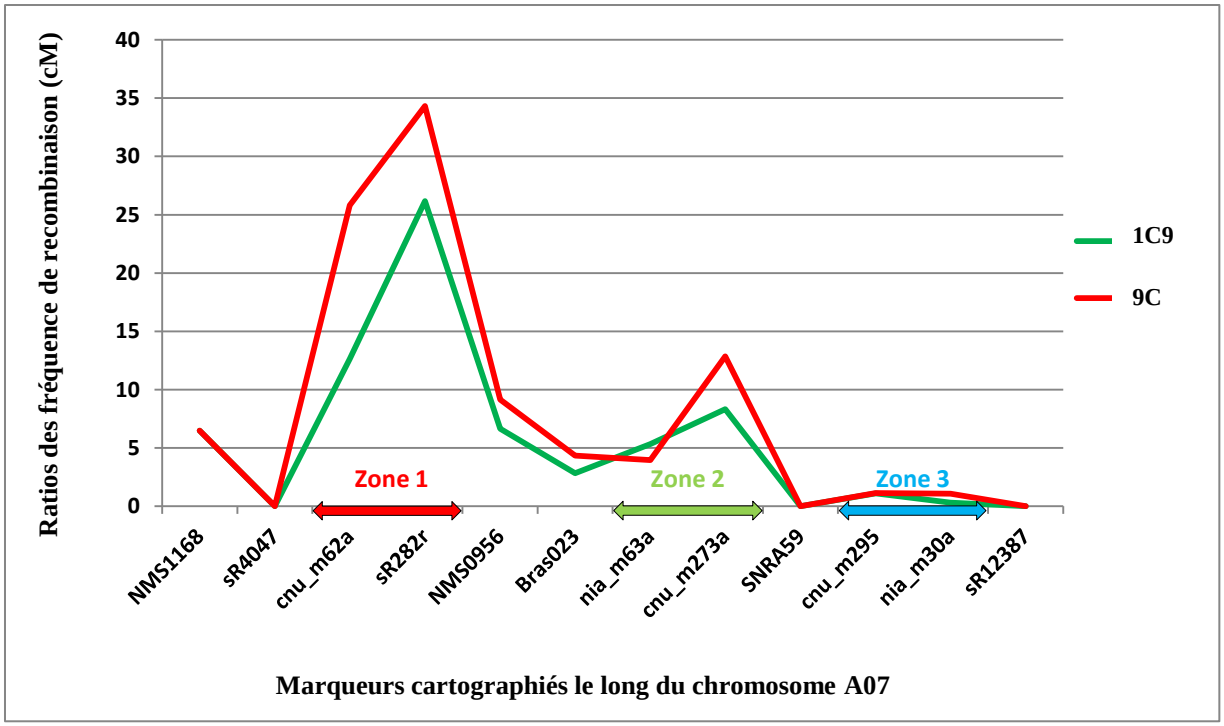


Figure 16. Ratios des fréquences de recombinaison le long du chromosome A07 par rapport au témoin AA sans chromosomes C en addition pour les populations $1C_9$ et 9C d'après Suay *et al.* (2013). Chaque point correspond au ratio de fréquence de recombinaison mesuré dans un intervalle entre deux marqueurs. Les zones définies sont schématisées par les flèches rouge, verte et bleue.

Tableau 9. Positions physiques des marqueurs utilisés par Suay *et al.* (2013) sur le génome de *B. rapa*

Marker	Hit	Debut (pb)	Fin (pb)
NMS1168	A07	11995	12628
sR4047	A07	660840	660571
cnu_m062a	A07	658916	658713
sR0282R	A07	1539077	1539096
NMS0956	A07	6470981	6471137
Bras023	A07	10910411	10910215
nia_m063a	A07	15288691	15288962
cnu_m273a	A07	15823972	15823714
sNRA59	A07	16337165	16337287
cnu_m295a	A07	16445601	16445795
nia_m030a	A07	18890204	18889754
sR12387	A07	20701201	20701497

II. ETUDE DE LA DISTRIBUTION DES COs LE LONG DU CHROMOSOME A07

II.1. Définition des zones à densifier

Dans le but d'étudier la distribution de la fréquence de recombinaison ainsi que les facteurs génomiques pouvant l'affecter, nous avons concentré notre travail sur des régions du chromosome A07 où des variations de fréquence de recombinaison ont été observées par Suay *et al.* (2013) chez les populations 0C, $1C_9$ et 9C (Figure 16). Trois zones ont été choisies pour être densifiées en marqueurs microsatellites (SSR) par la suite. La première (appelée Zone 1), située entre les marqueurs cnu_m62a et NMS956, présente un ratio de recombinaison élevé par rapport au témoin 0C (25-35 cM/Mb), la seconde (Zone 2), localisée entre nia_m63a et sNRA59, un ratio intermédiaire (8-13 cM/Mb) et la troisième (Zone 3) entre cnu_295 et nia_m30a, un ratio faible (<2 cM/Mb). Pour identifier la position de ces régions, nous avons aligné les couples d'amorces de ces SSR sur le génome de *B. rapa*. Les résultats de ce Blastn, présentés dans le tableau 9, nous ont permis de conclure que la Zone 1 est localisée entre 0,66 et 6,47 Mb, la Zone 2 entre 15,29 et 16,34 Mb et la Zone 3 entre 16,45 et 18,89 Mb. Bénéficiant des résultats de l'analyse structurale du chromosome A07, nous avons constaté que les Zones 2 et 3 sont localisées sur un Scaffold unique, le Scaffold 3, alors que la Zone 1 se situe dans la région identifiée comme présentant un assemblage de mauvaise qualité (Figure 1.b/). Afin d'établir un lien physique et d'étudier l'effet de la structure génomique sur la fréquence de recombinaison, il nous est essentiel de travailler sur des régions génomiques dont la qualité du séquençage est assurée. Nous avons donc décidé de redéfinir nos zones en fonction des scaffolds. Seule la partie de la Zone 1 identifiée au niveau du Scaffold 17 a été conservée et rallongée mesurant ainsi 2,5 Mb entre 5 et 7,5 Mb sur le chromosome A07. Sur le Scaffolds 3, les deux zones présentes ont également été réajustées, la Zones 2 mesurant ainsi 2 Mb entre 14,4 et 16,4 Mb et la Zone 3, 2,5 Mb entre 16,4 et 18,9 Mb.

II.2. Design de marqueurs SSR

Pour densifier les régions définies précédemment, nous avons dans un premier temps détecté par bioinformatique, via le workflow « microsat », 623 couples d'amorces entre des motifs microsatellites sur le Scaffold 17 et 1173 sur le Scaffold 3. Suite à cela, une première PCR « *in silico* » réalisée sur le génome de *B. rapa*, suivie d'une seconde sur le génome de *B. napus*, nous ont permis d'éliminer les couples d'amorces présentant une amplification multilocus et ceux non polymorphes entre les deux espèces. Ainsi 191 couples d'amorces ont été retenus sur le Scaffold 17 et 371 sur le Scaffold 3. Parmi eux, seulement 38 couples sur le Scaffold 17 et 31 sur le Scaffold 3 ont été sélectionnés pour tester le polymorphisme. Les résultats du polymorphisme obtenus après PCR sur les parents Darmor, RCC34, Chiifu et C1.3 suivis d'une révélation à l'analyseur de fragments 16 capillaires, sont présentés dans l'Annexe III. Ces derniers ont révélé 13 couples d'amorces présentant le polymorphisme attendu sur le scaffold 17 et 11 sur le Scaffold 3, soit seulement 35% des amorces testées.

Les résultats du génotypage des populations 0C, $1C_9$ et 9C à l'aide des SRR définis et des 12 microsatellites utilisés par Suay *et al.* (2013), sont représentés dans l'Annexe IV. Quatre de nos marqueurs sont absents de ces données car ces derniers ne se sont pas révélés

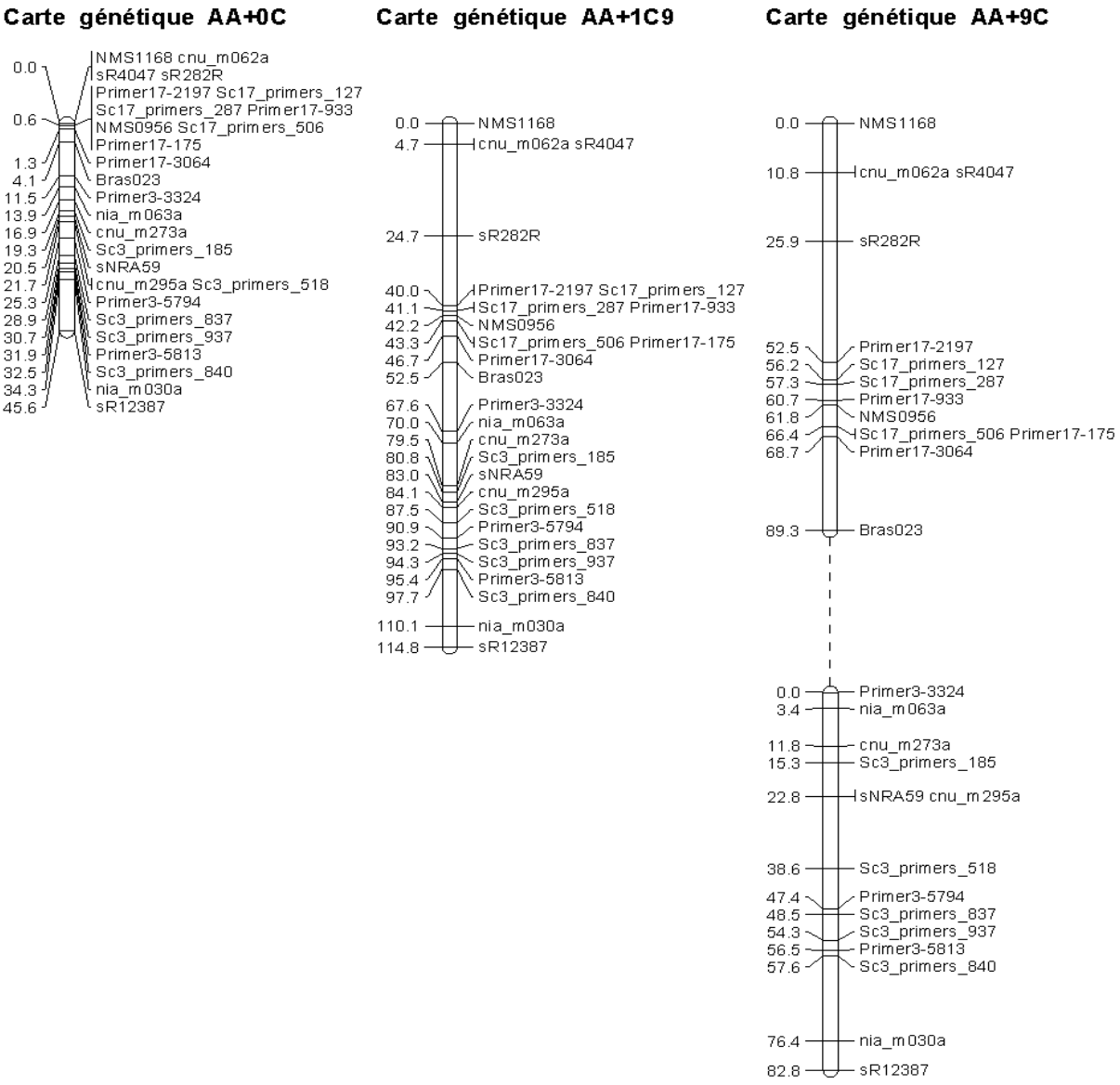


Figure 17. Cartographie du groupe de liaison A07 pour les populations 0C, $1C_9$ et 9C

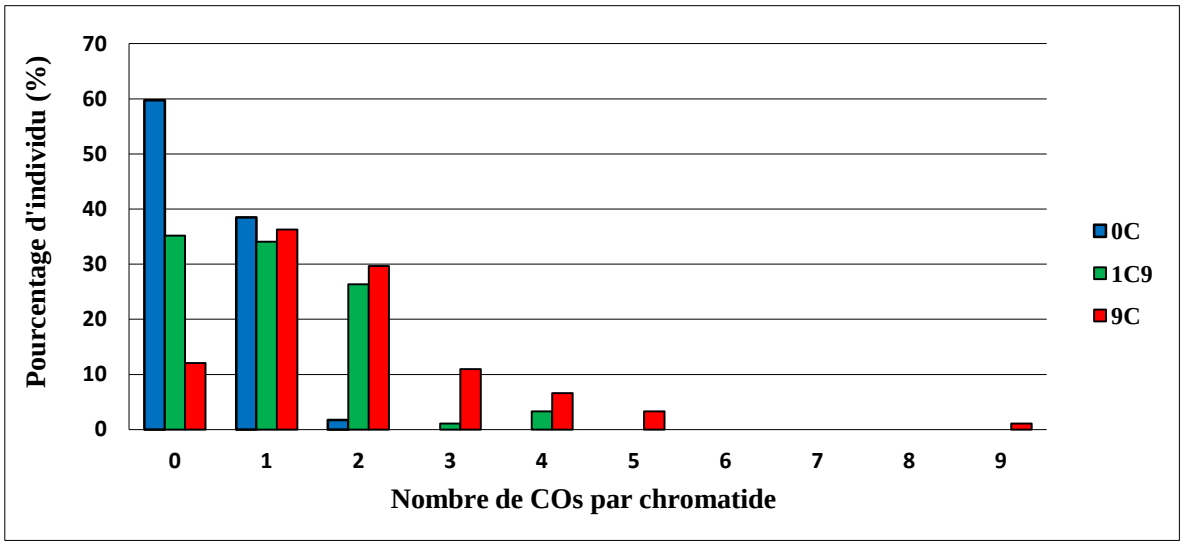


Figure 18. Fréquence de CO par chromatide et par individu pour les populations 0C, $1C_9$ et 9C

polymorphes lors de cette étape et n'ont donc pu être conservés pour la cartographie. Nous avons également pu observer que parmi les SSR définis et ceux génotypés par Suay *et al.* (2013), quatre sont en distorsion (ddl=1, χ^2>3,84) et uniquement chez la population $1C_9$. Ces SSR n'ont donc été cartographiés qu'après les autres microsatellites en ségrégation mendélienne. Enfin, afin de conserver un écartement d'un même ordre de grandeur entre marqueurs dans nos régions, nous avons également éliminé de notre analyse cinq SSR trop proches (<60Kb). Au final, entre nos marqueurs nouvellement définis et ceux de Suay *et al.* (2013), nos Zones 1, 2 et 3 comportent respectivement 8, 5 et 8 SSR avec un espacement entre marqueurs minimum de 108 Kb, maximum de 926 Kb et moyen de 298 Kb.

II.3. Cartographie génétique

Avant de commenter les résultats pour les populations 0C, $1C_9$ et 9C, nous rappelons que la population 0C est issue d'un croisement entre deux diploïdes (AA x AA) et d'un croisement entre un triploïde et un diploïde (AAC x AA) (Figure 14). Nous avons donc commencé par réaliser la cartographie du groupe de liaison A07 pour ces deux populations issues de croisements différents. Les résultats présentés dans l'Annexe V ont révélé l'absence de différence significative entre les populations pour la taille du groupe de liaison (Test χ^2, ddl = 26, p-valeur = 0,889). De plus, seul un intervalle est significativement différent sur les 26 (Test χ^2, ddl = 1, p-valeur = 0,004). Nous avons donc conclu que ces populations sont similaires pour la recombinaison et considéré par la suite, une unique population 0C regroupant les individus issus de deux diploïdes et d'un triploïde croisé avec un diploïde.

Les résultats obtenus lors de la cartographie génétique du groupe de liaison A07 chez les populations 0C, $1C_9$ et 9C sont représentés sur la figure 17. Nous avons tout d'abord observé que l'ordonnancement des marqueurs sur les cartes génétiques respecte l'ordre réel physique ainsi que l'ordre des cartes génétiques établies par Suay *et al.* (2013) avant l'ajout de nos SSR. Nous avons ensuite constaté que la taille du groupe de liaison est significativement différente entre les populations 0C et $1C_9$ (Test χ^2, ddl = 26, p-valeur = 2,9E-09), 0C et 9C (Test χ^2, ddl = 26, p-valeur =7,1E-28) ainsi que $1C_9$ et 9C (Test χ^2, ddl = 26, p-valeur =0,0005). En affinant notre analyse nous avons déterminé qu'entre les populations 0C et $1C_9$, 5 intervalles sur 26 sont significativement différents (Test χ^2, ddl = 1, p-valeur < 0.05), contre 11 entre les populations 0C et 9C et 4 entre $1C_9$ et 9C (Annexe VI). De plus, en nous intéressant au nombre de COs par chromatide, nous avons constaté que les populations $1C_9$ et 9C présentent plus de COs multiple que 0C (Figure 18). Toutefois nous avons identifié que chez la population témoin, le nombre de COs par chromatide observé est significativement différent de celui attendu (Test χ^2, ddl = 1, p-valeur =0,01). A partir de la taille des cartes génétiques, nous avons estimé que la fréquence de recombinaison augmente d'un facteur 2,5 entre les populations 0C et $1C_9$ et qu'elle augmente, sans tenir compte de l'intervalle entre les SSR Bras023 et Primer3-3324 qui présente une rupture, d'un facteur supérieur à 4,5 entre 0C et 9C ainsi que supérieur à 1,7 entre $1C_9$ et 9C. L'ensemble de ces résultats permet de conclure, comme Suay *et al.* (2013), que l'addition des 9 chromosomes C sous forme haploïde ou du chromosome C_9 seul chez un diploïde AA, favorise la recombinaison entre les chromosomes A07. Cependant, l'ajout du chromosome C_9 seul ne suffit pas à expliquer la taille du groupe de liaison A07 retrouvée lorsque les 9 chromosomes C sont présents impliquant un effet d'un

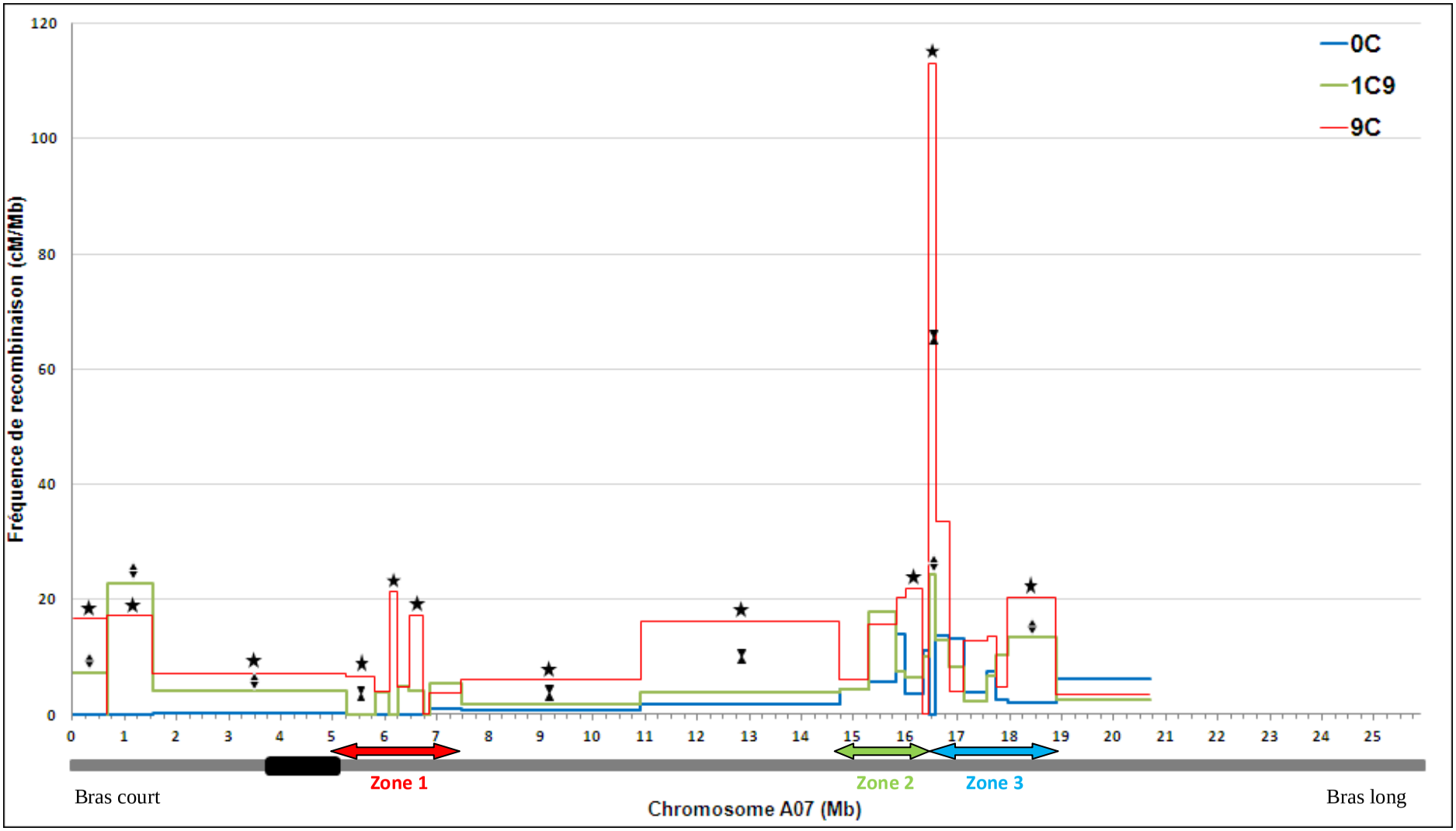


Figure 19. Représentation de la distribution des évènements de recombinaison le long du chromosome A07 de *B. rapa* chez les populations 0C, $1C_9$ et 9C. Le chromosome est schématisé en abscisse (en gris) avec ses régions péricentromériques (en noire). Les ♦ représentent les intervalles significativement différents entre 0C et $1C_9$, les ★ ceux entre 0C et 9C et les ⧗ ceux entre $1C_9$ et 9C. Les zones densifiées sont schématisées par les flèches rouge, verte et bleue.

ou de plusieurs autres chromosomes C dans la régulation de la recombinaison homologue. Nous avons donc conservé, par la suite, l'ensemble des marqueurs définis pour étudier la distribution des évènements de recombinaison.

II.4. Distribution des COs le long du chromosome A07

L'ensemble des marqueurs positionnés a permis d'étudier la distribution des COs sur 80% du chromosome A07, hors région télomérique de son bras long. Les résultats obtenus chez les populations 0C, $1C_9$ et 9C sont présentés sur la figure 19 avec conservation de la valeur de 61,8 cM produite par Mapmaker pour l'intervalle présentant une rupture chez 9C.

II.4.1. Distribution chez le témoin AA sans chromosome C en addition

La distribution des COs chez la population témoin 0C, en bleu sur la figure 19, montre une absence de recombinaison entre 0 et 6,9 Mb, dont l'intervalle correspond au bras court, au centromère et à ses régions péricentromériques ainsi qu'au début du bras long du chromosome A07. Nous avons également pu observer l'existence d'une corrélation positive entre la fréquence de recombinaison et la position physique des intervalles sur le bras long (R^2=0,3976 ; p-valeur=0.0013) traduisant la présence d'un gradient. Ce dernier implique donc l'augmentation progressive de la fréquence de recombinaison du centromère jusqu'au dernier intervalle. Toutefois, la distribution des COs n'est pas homogène le long du chromosome de part l'absence de recombinaison sur les 7 premiers Mb et l'existence de régions froides sur le bras long entre des régions chaudes pour la recombinaison, notamment de 16,4 à 16,6 Mb.

II.4.2. Comparaison avec les populations $1C_9$ et 9C

Entre les populations 0C et $1C_9$, nous avons tout d'abord observé que cinq intervalles sur 26 sont significativement différents pour la fréquence de recombinaison et quatre d'entre eux sont localisés dans des régions identifiées comme froides sur le chromosome A07 chez 0C. Plus étonnant, si l'un de ces intervalles se situe sur le bras long (de 16,4 à 16,6 Mb) et que deux sont sur le bras court (de 0 à 1,54 Mb), le dernier intervalle est celui qui comprend les régions péricentromériques (de 1,5 à 5,3 Mb). Ainsi, il apparait que la présence du chromosome C_9 chez des diploïdes AA induise la formation de nouvelles régions recombinantes et ce, même dans les régions péricentromériques. Nous avons ensuite observé chez $1C_9$, la conservation du gradient positif de la fréquence de recombinaison sur le bras long (R^2=0,2816 ; p-valeur=0.0092), ainsi que l'absence de différence significative pour les 21 autres intervalles par rapport à 0C. Nous pouvons donc conclure que lorsqu'il est en addition chez un diploïde AA, le chromosome C_9 affecte la distribution des COs sur le chromosome A07 en générant de nouvelles régions chaudes pour la recombinaison, localisées principalement sur le bras court et dans les régions péricentromériques.

Chez la populations 9C, nous avons observé que des COs se sont formés dans la quasi totalité des intervalles (23/26). Par rapport à la population 0C, 11 intervalles sont significativement différents pour la fréquence de recombinaison, dont sept sont localisés dans des régions définies comme froides pour la recombinaison chez la population 0C. Parmi ces 7 nouvelles régions recombinantes, nous avons retrouvé les quatre intervalles observés lors de

Tableau 10. Corrélations entre la teneur en GC, gènes et TEs avec la fréquence de recombinaison dans les régions densifiées en SSR du chromosome A07. L'indice (a) correspond au sens de la corrélation observée et le R^2 indique son coefficient de détermination. Les p-valeurs sont indiquées par les couleurs des cases, roses <0,01, jaunes <0,05, bleues <0,1 et blanc >0,1.

	AA+0C	AA+1C9	AA+9C
Teneur en GC			
(a)	(+)	(+)	(+)
R^2	0,4757	0,0387	0,0689
Teneur en gènes			
(a)	(+)	(+)	(+)
R^2	0,5819	0,2636	0,061
Teneur en TEs			
(a)	(-)	(-)	(-)
R^2	0,1725	0,256	0,0721
Classe I			
(a)	(-)	(-)	(-)
R^2	0,1493	0,2552	0,0477
Classe II			
(a)	(-)	(-)	(-)
R^2	0,0246	0,0002	0,0694

la comparaison entre 0C et $1C_9$, indiquant que l'effet identifié par le chromosome C_9 est conservé lorsque les 9 chromosomes C sont en addition chez un diploïde AA. De plus, nous avons constaté que les trois nouvelles régions recombinantes, formées par l'addition des 9 chromosomes C, se situent au début du bras long du chromosome A07 (de 5,25 à 6,75 Mb). L'un de ces intervalles est même significativement différent entre les populations $1C_9$ et 9C (de 5,3 à 5,8 Mb), impliquant un effet des 8 autres chromosomes C dans la formation de nouvelles régions chaudes pour la recombinaison au début du bras long du chromosome A07. Dans les zones présentant de la recombinaison chez 9C, une augmentation de la fréquence de recombinaison est observée pour 4 intervalles par rapport à 0C (de 7,5 à 14,7 Mb, de 16 à 16,3 Mb et de 18 à 18,9 Mb) et 3 par rapport à $1C_9$ (de 7,5 à 14,7 Mb et de 16,4 à 16,6 Mb). De plus, chez 9C, plus aucun gradient significatif pour la fréquence de recombinaison n'est identifié sur le bras long du chromosome A07 (R^2=0.0573 ; p-valeur=0.2715). L'augmentation de la taille de la carte génétique du groupe de liaison A07 semble donc, chez cette population, expliquée par une régulation de la fréquence de recombinaison des régions déjà recombinantes chez 0C et $1C_9$. L'addition des 8 autres chromosomes C pourrait donc affecter la distribution des COs le long du chromosome A07 en générant de nouvelles régions chaudes pour la recombinaison, localisées principalement sur le début du bras long et en régulant la fréquence de recombinaison dans des régions déjà aptes à recombiner.

II.5. Analyse des corrélations entre la fréquence de recombinaison et la structure génomique du chromosome A07

II.5.1. Etude des corrélations chez la population AA sans chromosome C en addition

Afin de comprendre l'impact de la structure génomique des chromosomes A sur la distribution des COs, nous avons déterminé les corrélations existantes entre la fréquence de recombinaison et les teneurs en bases GC, gènes et TEs, uniquement dans les zones densifiées en SSR du chromosome A07 chez les populations 0C, $1C_9$ et 9C (Annexe VII). Ces résultats, présentés dans le tableau 10, nous ont permis d'observer chez la population 0C, des corrélations positives et significatives pour les teneurs en GC et gènes. Une corrélation négative, moindre mais significative à un seuil de 10%, a également été observée avec la teneur en TEs sans pouvoir identifier la classe de TEs responsable. Nous avons conclu de ces résultats que la structure génomique du chromosome A07 explique, dans les zones densifiées, la distribution des COs. Ces derniers sont favorisés dans des régions riches en bases GC et en gènes alors qu'ils semblent moins fréquents dans les régions riches en TEs.

II.5.2. Etude des corrélations avec des chromosomes C en addition

Nous nous sommes ensuite attachés à déterminer si ces relations structure génomique-fréquence de recombinaison sont conservées lorsque le chromosome C_9 ou les 9 chromosomes C sont en addition chez le diploïde AA. Nous avons ainsi observé **chez la population $1C_9$**, l'absence de corrélation avec la teneur en GC et une faible corrélation positive avec la teneur en gènes. Il semble donc que, dans les zones densifiées, ces composantes de la structure génomique du chromosome A07 n'expliquent plus, ou moins, la distribution des COs chez $1C_9$. Cependant la forte corrélation négative établie avec la teneur en TEs et plus particulièrement la classe I chez $1C_9$, indique que comme pour la population

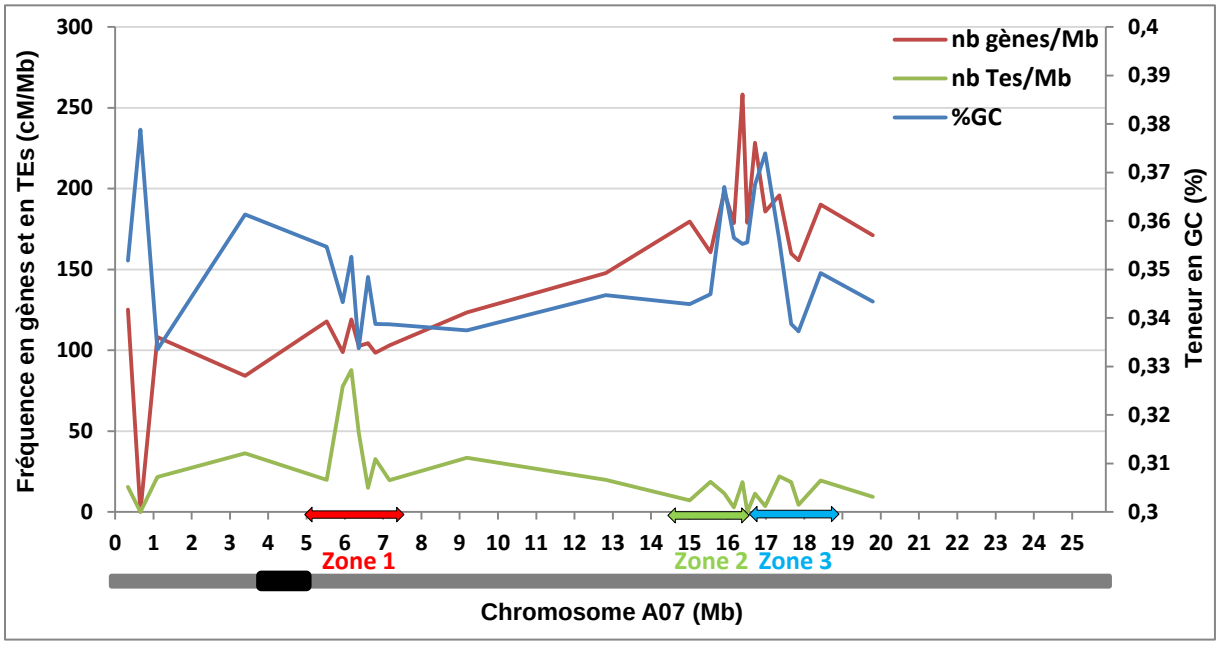


Figure 20. Distribution des fréquences en gènes et en TEs ainsi que du pourcentage en bases GC le long du chromosome A07. Le chromosome est schématisé en abscisse (en gris) avec sa région centromérique (en noire). Les zones densifiées sont schématisées par les flèches rouge, verte et bleue.

Tableau 11. Mesure de l'interférence, via les données de génotypage, avec les modèles à 1 voie et à 2 voies de COs chez les populations 0C, $1C_9$ et 9C, sur le chromosome A07. nu_{eff} représente l'estimation de l'intensité de l'interférence pour le modèle à 1 CO et *nu* pour le modèle à 2 COs. *p* représente la proportion de COs formés par la voie II et nu_{eff}*_inf*, nu_{eff}*_sup*, *nu_inf*, *nu_sup* , *p_inf*, *p_sup* représentent les intervalles de confiance pour 1000 re-simulations.

Populations	nu_{eff}	nu_{eff}*_inf*	nu_{eff}*_sup*	*nu*	*nu_inf*	*nu_sup*	*p*	*p_inf*	*p_sup*
AA+0C	13.009	1.321	24.696	13.009	1.321	24.696	0.000	0.000	0.112
AA+1C9	1.000	1.000	1.355	1.000	1.000	1.621	0.499	0.000	1.130
AA+9C	1.119	1.000	1.413	1.119	1.000	1.413	0.000	0.000	2.787

0C, les COs se forment moins fréquemment dans les régions riches en TEs. Nous avons enfin observé **chez la population 9C**, l'absence de toute corrélation avec les teneurs en GC, gènes et TEs dans les zones densifiées. Ces résultats indiquent donc que lorsque les 9 chromosomes C sont en addition chez un diploïde AA, les COs se produisent entre les chromosomes A07 indépendamment de leur structure génomique. De plus avec les résultats observés chez $1C_9$, il apparait que un ou plusieurs des 8 autres chromosomes C permettent aux COs de se former dans des régions riches en TEs.

II.5.3. Qu'en est-il sur le chromosome A07 entier ?

Avec l'ensemble des SSR positionnés de façon moins homogène, la répartition des fréquences en gènes et TEs ainsi que la teneur en GC a été établie le long du chromosome A07 (figure 20). Nous avons ainsi déterminé des corrélations avec leur position physique sur le bras long, de l'intervalle comprenant les régions péricentromérique à celui le plus éloigné. Celles-ci n'ont toutefois pu être établies sur le bras court faute d'un nombre suffisant d'intervalles (3). Deux gradients ont pu être identifiés pour la teneur en gènes et TEs, respectivement positif (R^2=0,6873 ; p-valeur=$1,86^E$-06) et négatifs (R^2=0,4299 ; p-valeur=$6,83^E$-04). Ainsi, il apparait d'après ces résultats que les régions péricentromériques sont pauvres en gènes et riches en TEs. Etant donné l'absence de recombinaison chez la population 0C, ces résultats étaient attendus. Toutefois, nous avons constaté que le nombre de gènes reste élevé dans les régions péricentromériques (84,29 gènes/Mb). De plus, cette région est beaucoup moins riche en TEs que le début du bras long (37 contre 87 TEs/Mb). Cette répartition des TEs ajoutée à la corrélation négative établie dans les régions densifiées pourrait expliquer l'absence de recombinaison chez 0C et $1C_9$ au début du bras long du chromosome A07 et pourrait confirmer l'effet des 8 autres chromosomes C sur la formation de nouvelles régions recombinantes dans les régions riches en TEs. Concernant la teneur en GC, aucun gradient n'a pu être établi (R^2=0,0497 ; p-valeur=0,1169) indiquant une répartition non homogène. Toutefois, des régions riches en GC au début du chromosome A07 sont observées (de 1,5 à 5,8 Mb) alors que chez la population 0C aucun CO n'est identifié malgré la corrélation positive précédemment établie. Ces résultats indiquent donc que le début du chromosome ne suit pas les mêmes règles de corrélation que le bras long d'A07 et impliquent l'existence de facteurs contrôlant spécifiquement la recombinaison dans ces régions.

III. ANALYSE DE L'INTERFERENCE

Nous nous sommes attachés dans cette partie à étudier l'interférence sur le chromosome A07 chez les populations 0C, $1C_9$ et 9C (Tableau 11). Chez la population témoin sans chromosome C en addition, nous avons d'abord observé une interférence élevée sur le chromosome entier quel que soit le modèle utilisé (nu_{eff}=nu=13,009). En revanche, cette observation diffère chez les populations $1C_9$ et 9C où une très faible interférence est détectée (nu_{eff}=nu=1,000 ; nu_{eff}=nu=1,119). Ainsi, nous avons pu conclure de ces résultats que l'ajout du chromosome C_9 ou des 9 chromosomes C chez un diploïde AA affecte l'interférence initialement forte sur le chromosome A07 comme l'a observé Suay *et al*. (2013).

CHAPITRE IV. DISCUSSION

I. STRUCTURE DU CHROMOSOME A07 DE *B. RAPA*

Nous avons montré dans un premier temps que les structures du chromosome A07 de *B. rapa* et *B. napus* présentent d'importantes similitudes malgré quelques points de rupture (Figure 15.a/). Si l'évolution des Brassicas est marquée par plusieurs évènements de duplication, il semble que la formation récente du colza (<10.000 ans), par hybridation interspécifique d'une navette (*B. rapa*) et d'un chou (*B. oleracea*), explique la conservation des génomes de ses progéniteurs en son sein (Parkin *et al.* 2005 ; Adam et Wendel, 2005 ; Yang *et al.* 2006). En effet, des études de cartographie génétique et de comparaison de séquence de BAC (Bacterial Artificial Chromosomes) ont confirmé la forte similitude entre les génomes A et C de *B. napus* et de ses progéniteurs mais ont toutefois identifiées l'existence d'évènements de translocation chez des colzas naturels (Parkin *et al.* 1995 ; Cheung *et al*. 2009 ; Xiong et Pires, 2011). Ces remaniements pourraient ainsi expliquer les différents points de ruptures que nous avons observés au début du chromosome A07. Toutefois, nous ne sommes qu'aux toutes premières versions d'assemblages des génomes de Brassicas comme en témoigne les résultats obtenus lors de l'alignement des Scaffolds (Figure 15.b/). Les séquençages à venir, nous permettront par la suite d'améliorer la qualité des séquences et de confirmer ces conclusions.

Nous avons dans un second temps identifié la position des régions péricentromériques du chromosome A07 à l'aide de séquences connues et des blocs de Schranz dans un intervalle compris entre 3,71 et 5,13 Mb. Ces résultats sont confortés par l'ensemble de la bibliographie qui identifie la configuration du chromosome A07 comme acrocentrique (Koo *et al.* 2011 ; Cheng *et al.* 2013 ; Pouilly *et al*. 2008 ; Xiong et Pires, 2011). Toutefois, il ne nous a pas été possible de définir la position exacte du centromère malgré 15 séquences centromériques identifiées à 4,86 Mb sur les 21 observées. En effet, aucune des trois localités, présentant un alignement, ne correspondent à l'intervalle défini par les blocs de Schranz. La difficulté à assembler les séquences répétées lors du séquençage peut expliquer pourquoi nous n'identifions pas plus de séquences répétées qui composent le centromère (Harrison et Heslop-Harrison, 1995 ; Round *et al.* 1997 ; Luce *et al*. 2006 ; Lim *et al*. 2005 ; 2007). Cependant elle n'explique pas pourquoi nous localisons la séquence centromérique en différents emplacements. Ces derniers pourraient correspondre à des paléocentromères, c'est-à-dire à des vestiges d'une position plus ancienne du centromère (Cheng *et al.* 2013). Pour confirmer la position des régions péricentromériques et centromérique, il semble par la suite nécessaire de valider physiquement en hybridant des BAC spécifiques de ces régions (contenant des séquences centromériques et péricentromériques) sur le chromosome A07 et d'y ancrer physiquement des marqueurs génétiques cartographiés comme l'on réalisé Cheng *et al.* (2001) sur le riz.

II. ETUDE DE LA REGULATION ET DE LA DISTRIBUTION DES COs LE LONG DU CHROMOSOME A07

II.1. Régulation et distribution des évènements de COs chez des diploïdes sans chromosomes C en addition

Nous avons tout d'abord observé que la population témoin 0C présente moins d'un CO par chromosome et une forte intensité d'interférence. Ce phénomène d'interférence positive, identifié chez la quasi-totalité des modèles d'étude, limite le nombre de CO par chromosome à 1 ou 2 et peut expliquer pourquoi nous identifions peu de CO chez cette population (Jones, 1984 ; Muller, 1916 ; Mezard *et al*. 2007). Toutefois il n'explique pas pourquoi nous n'observons pas un minimum d'un CO par chromosome assuré par l'homéostasie qui permet de maintenir la bonne ségrégation des chromosomes au cours de la méiose. L'absence de marqueurs SSR sur 20% du chromosome A07 nous empêche sans doute de visualiser certains évènements de recombinaison au bout du bras long et peut ainsi expliquer que ce minimum n'est pas atteint.

Concernant la distribution des évènements de recombinaison, nous avons observé que, chez la population 0C, aucun CO n'est présent sur les 7 premiers Mb du chromosome A07 comprenant le bras court, les régions péricentromériques et centromérique ainsi que le début du bras long. Ceci peut s'expliquer par l'existence de facteurs génomiques et surtout épigénétiques autour du centromère inhibant la formation des COs (Choo, 1998). Il a même été observé que ces facteurs peuvent être répandus sur plusieurs Mb et ainsi empêcher les évènements de recombinaison de se produire sur le bras court des chromosomes acrocentriques et au début du bras long (Dumas et Britton-Davidian, 2002). En s'éloignant du centromère, nous avons observé que la fréquence de recombinaison suit un gradient positif malgré une alternance de points chauds et froids sur le bras long du chromosome A07. Cette augmentation du nombre de CO à mesure que l'on s'éloigne du centromère a été identifiée chez la plupart des espèces animales et végétales tout comme l'alternance de point chauds et froids (Kong *et al.* 2002 ; Shifman *et al.* 2006 ; Lukaszewski et Curtis 1993 ; Kunzel *et al*. 2000 ; Anderson *et al.* 2003 ; Chen *et al*. 2002 ; Sherman et Stack 1995).

Cette distribution non homogène, traduite par l'alternance de régions chaudes et froides pour la recombinaison, semble en partie pouvoir être expliquée par la structure génomique du chromosome A07. Nous avons en effet identifié des corrélations chez la population 0C indiquant que les COs se forment préférentiellement dans les régions riches en GC et gènes mais pauvres en TEs. Bien que fonction des espèces le sens de ces corrélations varie (Kong *et al.* 2002 ; Jensen-Seaman *et al.* 2004 ; Paigen *et al.* 2008 ; Wu *et al.* 2003 ; Anderson *et al.* 2006 ; Erayman *et al.* 2004 ; Rizzon *et al.* 2002 ; Dooner et He 2008 ; Giraut *et al.* 2011), les études menées à une échelle plus forte en densification (de l'ordre du Kb) vont dans le même sens que nos résultats (de Massy, 2003 ; Myers *et al*. 2005 ; Myers et Spencer, 2006). Toutefois, même s'il semble que les sens de ces corrélations soient universels, leur explication fait encore débat. En effet, pour la teneur en GC, il semble que la formation de NCOs soit à l'origine d'un enrichissement en GC via la conversion génique conduisant préférentiellement à la réparation des mésappariements en G:C plutôt qu'en A:T (Birdsell, 2002). Pour la teneur

en TEs, aucune explication n'a pu être apportée même s'il est envisagé par différents auteurs que ces éléments mobiles se déplacent dans des régions afin d'y limiter la recombinaison (Rizzon *et al.* 2002 ; Wu *et al.* 2003). La teneur en gènes ne trouve également pas d'explication, mais la distribution des COs intervenant préférentiellement dans les régions non codantes pourrait indiquer l'existence d'un contrôle établi pour limiter la perte de leur fonction lors de leur brassage (Myers *et al*. 2005 ; de Massy, 2003).

A l'échelle du chromosome entier, nous avons constaté certaines irrégularités par rapport aux corrélations établies (Figure 23). En effet, nous avons observé que les régions péricentromériques sont riches en GC bien qu'elles soient dépourvues de recombinaison. Cette constatation a également été faite chez *A. thaliana* et corrobore ainsi l'existence de facteurs, autres que génomiques, dans ces régions inhibant la présence de COs (Giraut *et al.* 2011 ; Choo, 1998). Pour les TEs, le gradient négatif établi du centromère au bout du bras long du chromosome A07 semble indiquer qu'ils sont majoritairement répartis au niveau du centromère. Pourtant, ces derniers semblent plus fortement concentrés au début du bras long du chromosome A07 et pourraient expliquer l'absence de recombinaison dans cette région. Leur accumulation dans les régions proches du centromère a de nombreuse fois été observée et semble être un élément supplémentaire pour confirmer la position des régions péricentromériques que nous avons définies (Harrison et Heslop-Harrison, 1995 ; Round *et al.* 1997 ; Schmidt et Heslop-Harrison, 1998 ; Hall *et al.* 2006 ; Luce *et al*. 2006 ; Lim *et al*. 2005 ; 2007 ; Pouilly *et al*. 2008). Enfin, pour la teneur en gènes, il ne semble pas y avoir d'irrégularité. Nous avons simplement constaté un nombre élevé de gènes dans les régions froides pour la recombinaison, y compris dans les régions péricentromériques (84,29 gènes/Mb) contrairement aux autres espèces étudiées (Kong *et al.* 2002 ; Paigen *et al.* 2008 ; Wu *et al.* 2003 ; Erayman *et al.* 2004). Il apparait ainsi la difficulté de brasser ces différents locus en sélection variétale et la nécessité d'induire des COs dans ces régions chez *B. rapa*.

II.2. Effets de l'addition des chromosomes C chez un diploïde sur la régulation et la distribution des COs

Nous avons conclu que **l'addition du chromosome C_9** chez des diploïdes AA entraine une augmentation du taux de recombinaison sur le chromosome A07, par rapport à des diploïdes sans chromosome C en addition, en générant de nouvelles régions recombinantes sur le bras court et dans les régions péricentromériques. Aucune étude n'a, à ce jour, identifiée pareil phénomène. Seule l'augmentation du taux de recombinaison a pu être observée en cas de polyploïdisation ou de modification de la structure génomique mais la distribution des COs n'a pas été étudiée (Carlton *et al*. 2006 ; Leflon *et al*. 2010 ; Pécinka *et al.* 2011). Avec les résultats dont nous disposons, nous savons que la séquence du chromosome A07 est la même chez les populations 0C et $1C_9$ et que les conditions environnementales qui ont permis leur obtention sont identiques. Ainsi, seuls des facteurs génétiques ou épigénétiques induits par C_9 pourraient expliquer cette modification de la distribution des COs. Le chromosome C_9 pourrait ainsi réguler les gènes liés à la formation des CDB tel que Spo11 ou des COs comme MLH1 ou Mus81 en augmentant leur nombre. Toutefois, l'absence de corrélation entre le nombre de CDB et de COs chez l'ensemble des espèces modèles, nous pousse plutôt à penser que C_9 pourrait favoriser la réparation des CDB en COs de type I ou II et non plus en NCO

(Mezard *et al*. 2007). La formation de CO semble même appuyée par les résultats obtenus après mesure de l'interférence qui montre une forte diminution de son intensité chez la population $1C_9$ par rapport à 0C sur le chromosome A07. Ce résultat pourrait ainsi indiquer une augmentation des COs II selon Leflon *et al*. (2010) pour expliquer la présence des nouvelles régions recombinantes mais doit toutefois être confirmé par immunolocalisation des protéines MLH1(COs I) et MUS81 (COs II). L'étude actuellement menée par Eric Jenczewski sur le QTL majeur *PrBn*, impliqué dans la régulation des recombinaisons homéologues et présent sur le chromosome C_9, a révélé la présence de plusieurs gènes de méiose qui pourraient être impliqués dans la formation de ces COs et pose la question de l'implication de *PrBn* dans la régulation de la recombinaison homologue.

Cependant, la présence de COs dans les régions péricentromériques du chromosome A07 attire notre attention. Celle-ci n'a jamais été observée (Choo, 1998) et constitue donc un résultat majeur, de cette étude, bien que la distance entre les marqueurs de ces régions soit élevée (3,71 Mb), nécessitant à l'avenir une densification. Cette observation a également été validée sur les chromosomes A02 et A04 dont l'étude a été possible grâce aux données de cartographie de Suay *et al*. (2013) sur des populations 0C et 6C (C_1+C_5+C_6+C_7+C_8+C_9) obtenues de la même manière que nos trois populations (Données non présentées voir Annexe VIII). Pour expliquer l'absence de CO dans ces régions, Lamb *et al.* (2005) supposent que des mécanismes empêcheraient leur formation pour éviter une mauvaise ségrégation des chromosomes pendant l'Anaphase I. Certains de ces mécanismes ont été mis en évidence laissant supposer des régulations épigénétiques (Choo, 1998). La méthylation a ainsi été identifiée comme inhibant la présence de COs (Yelina *et al.* 2012 ; Maloisel et Rossignol, 1998 ; Melamed-Bessudo et Levy, 2012). Toutefois, l'inactivation du gène DDM1, impliqué dans le maintien du groupement méthyle sur les cytosines (Brzeski et Jerzmanowski, 2003), a conduit à une augmentation de la fréquence de recombinaison dans les régions euchromatiques mais pas dans les régions hétérochromatiques retrouvées généralement dans les régions péricentromériques (Melamed-Bessudo et Levy, 2012). La condensation de l'ADN semble donc impliquée dans l'inhibition de la formation de COs et pourrait, combinée à la méthylation, expliquer l'absence de recombinaison au centromère (Choo, 1998). La présence de C_9 favoriserait ainsi, soit la décondensation de l'ADN, soit la déméthylation, soit les deux et expliquerait la présence de recombinaison dans les régions péricentromériques ainsi que dans les régions proches comme sur le bras court d'A07 (Dumas et Britton-Davidian, 2002). L'identification des cytosines méthylées et des régions condensées chez les populations 0C et $1C_9$ permettrait de confirmer ces hypothèses.

Lorsqu'il est en addition chez un diploïde AA, le chromosome C_9 semble donc induire des facteurs génétiques voir épigénétiques permettant ainsi de produire des évènements de COs sur la plupart des intervalles entre nos marqueurs sur le chromosome A07 (22/26) et ce même dans des régions identifiées comme froides pour la recombinaison chez 0C (bras court et régions péricentromériques). Cependant, une corrélation négative avec la teneur en TEs, et plus spécifiquement ceux de classe I, dans les zones densifiées est encore identifiée chez cette population. Celle-ci permettrait ainsi d'expliquer l'absence de CO sur le début du bras long du chromosome où nous avons observé la concentration la plus élevée de TEs.

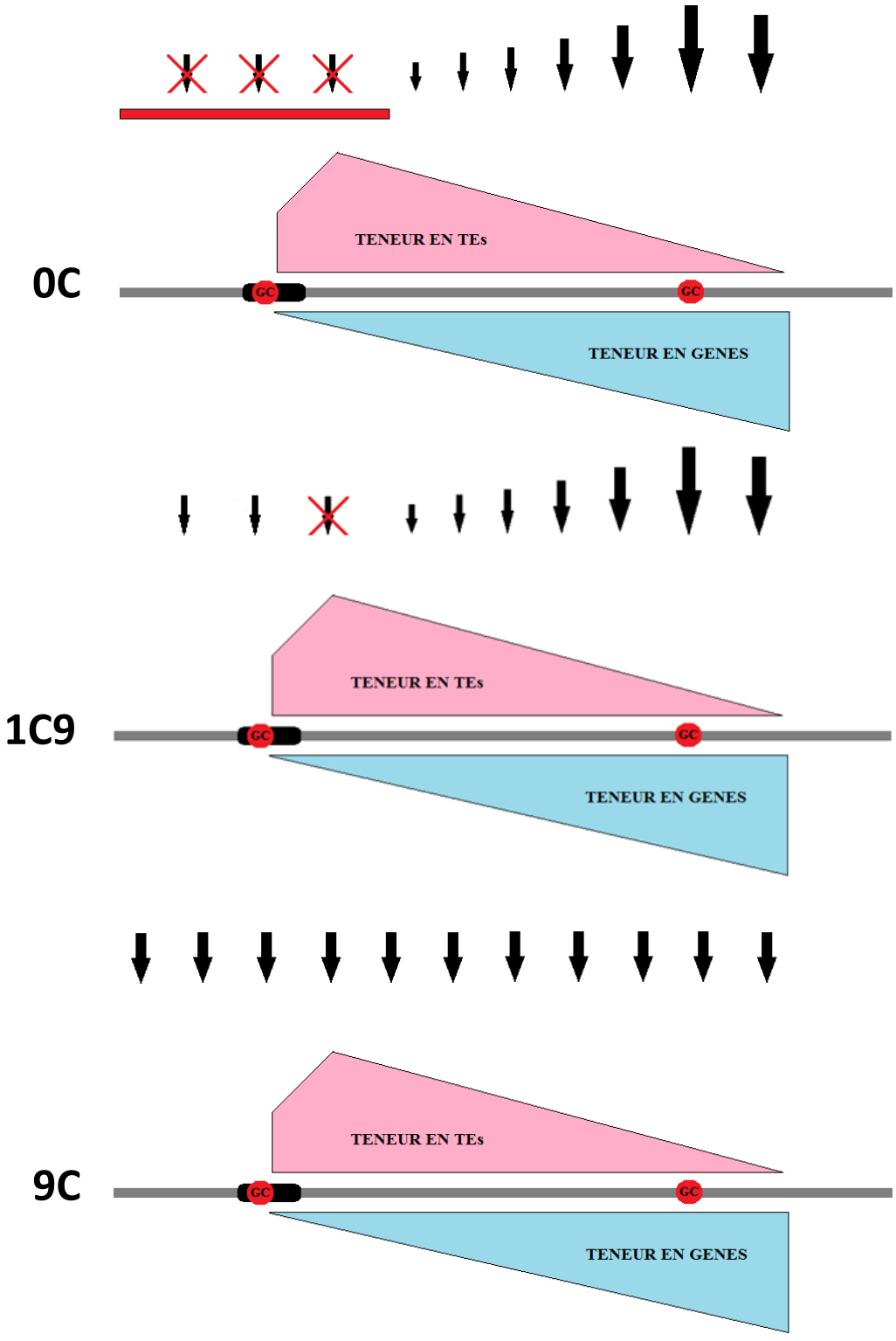


Figure 21. Schématisation des résultats obtenus pour l'étude de la distribution des évènements de recombinaison sur le chromosome A07 chez les populations 0C, $1C_9$ et 9C. Le chromosome A07 est schématisé en gris avec ses régions péricentromériques et centromérique en noire. La proportion des TEs le long du chromosome est représentée en rose et celle des gènes, en bleu. Les régions riches en GC sont schématisées par un rond rouge. Les COs sont eux schématisés par des flèches dont la taille est proportionnelle à la fréquence.

Nous avons conclu que **la présence des 9 chromosomes C** sous la forme haploïdes chez un diploïde AA entraine une augmentation du taux de recombinaison sur le chromosome A07 et la formation de nouvelles régions chaudes pour la recombinaison par rapport aux populations 0C et $1C_9$. La conservation de l'effet produit par $1C_9$ chez la population 9C nous permet de poser les mêmes hypothèses génétiques et/ou épigénétiques. Cependant l'augmentation significative de la fréquence de recombinaison sur plusieurs intervalles où la présence de recombinaison est déjà observée chez 0C et $1C_9$ pourrait impliquer un effet des 8 autres chromosomes C dans la régulation de la recombinaison. Si le QTL *PrBn* est impliqué, il pourrait s'agir des QTLs mineurs identifiés par Liu *et al.* (2006) sur les chromosomes C_1, C_6 et A_3 et des différentes relations épistatiques définies. La faible intensité d'interférence mesurée chez cette population pourrait indiquer, comme pour $1C_9$, la formation de CO II. Chez cette population, la plupart des intervalles entre nos marqueurs présente de la recombinaison (23/26) et plus aucune corrélation avec la structure génomique du chromosome A07 n'est identifiée, indiquant une distribution indépendante des COs le long du chromosome A07. L'absence de corrélation entre la fréquence de recombinaison et la teneur en TEs, dans les régions densifiées, associée à la présence de nouvelles régions chaudes pour la recombinaison par rapport à $1C_9$ sur le début du bras long semble indiquer un effet des 9 chromosomes C dans la formation de COs dans les régions riches en TEs. Ainsi, il semble qu'un ou plusieurs autres chromosomes C, agissant indépendamment ou en combinaison, pourraient permettre la formation de COs dans les régions riches en TEs. Seul chez *A. thaliana*, une corrélation positive a été détectée entre fréquence de recombinaison et teneur en TEs sans pouvoir expliquer ce résultat (Giraut *et al*. 2011). Il semblerait donc, d'après les observations faites chez 0C, $1C_9$ et 9C, qu'un contrôle inhibant la formation des COs dans les régions riches en TEs existe mais que la présence de chromosomes C sous forme haploïde puisse le modifier. Dans le but d'identifier le ou les chromosomes C responsables de cet effet, il parait intéressant de générer des hybrides AA avec des combinaisons de chromosome C en addition dont le nombre et la nature varient puis d'étudier la distribution des COs comme nous l'avons réalisés chez 0C, $1C_9$ et 9C.

Pour conclure, sur l'effet de l'addition du chromosome C_9 ou des 9 chromosomes C chez un diploïde AA, nous avons schématisé nos résultats sur la figure 21. Chez des diploïdes sans chromosome C en addition, la structure génomique explique la distribution des COs peu nombreux. Ceux-ci sont répartis sur le bras long du chromosome A07 et leur fréquence augmente dans les régions riches en gènes et en GC mais diminue dans les régions riches en TEs. De plus, la présence de facteurs empêche la formation de COs dans les régions centromérique et péricentromérique ainsi que sur plusieurs Mb à proximité. L'ajout du chromosome C_9 semble inhiber ces facteurs permettant la formation de COs sur le bras court ainsi que dans les régions péricentromériques du chromosome tout en conservant la distribution observée chez 0C. Cependant, la présence de TEs continue d'inhiber la formation de COs. Enfin, l'ajout des 9 chromosomes C produit le même effet que celui de C_9, mais permet de réguler la recombinaison dans les régions déjà recombinantes tout en rendant possible la recombinaison dans les régions riches en TEs.

Nos résultats doivent cependant être validés. En effet, il nous faut rappeler que nous avons densifié 3 régions qui ne représentent que 29% de la taille du chromosome A07. Il semble donc nécessaire de densifier l'ensemble du chromosome pour nous assurer que nos corrélations soient toujours bien identifiées. Cette densification nous permettra par la même d'observer la distribution des COs sur le télomère du bras long, impossible dans cette étude de par l'absence de marqueur au bout du chromosome. Ensuite, nous rappelons que les intervalles entre nos marqueurs présentent de fortes variations (108 Kb à 926 Kb), due à la difficulté de designer des SSR avec le polymorphisme recherché. Les variations de ces intervalles peuvent ainsi affecter nos résultats et il semble donc indispensable de définir à l'avenir des marqueurs mieux répartis à intervalles réguliers (150 à 300 Kb) comme les autres études menées (Giraut *et al.* 2011). Cette densification uniforme sur tout le chromosome apportera également plus de données augmentant la puissance de nos tests statistiques et affinant la précision des régions chaudes pour la recombinaison. Chez nos différentes populations, nous avons étudié la distribution avec relativement peu de données. En effet, les populations 0C, $1C_9$ et 9C ne nous ont permis d'observer respectivement que 70, 59 et 80 COs. Il semble donc nécessaire de régénérer ces populations afin d'obtenir un nombre plus élevé de COs pour étudier leur distribution mais également l'interférence très sensible à la moindre erreur de génotypage.

CONCLUSION GENERALE

Nous avons montré au cours de cette étude que la distribution des évènements de COs le long du chromosome A07 est expliquée par sa structure génomique chez des diploïdes AA et que cette distribution varie fonction de la composition chromosomique de l'hybride. Ainsi, des régions caractérisées comme étant froides pour la recombinaison telles que les régions péricentromériques peuvent recombiner par le seul ajout du chromosome C_9. Les huit autres chromosomes C semblent eux principalement impliqués dans la régulation de la fréquence de recombinaison mais favorisent tout de même la formation des COs dans des régions riches en TEs. Bien que ces résultats doivent être généralisés avec une densification plus homogène aux autres chromosomes A, la compréhension des mécanismes génétiques et/ou épigénétiques mis en place par le chromosome C_9 et les 8 autres chromosomes C semblent indispensables pour pouvoir être utilisés en sélection variétale. Leurs effets pourront ainsi trouver des applications directes en amélioration des plantes. En effet, dans le clonage de gènes ou la détection de QTL, l'augmentation de la fréquence de recombinaison, associée à une densité en marqueur élevée, peut permettre de réduire l'intervalle de confiance de la position d'un gène à cloner ou d'un QTL tout en réduisant la taille des populations à observer. En sélection variétale, la possibilité d'induire des COs dans des régions froides pour la recombinaison peut elle permettre de briser certains déséquilibres de liaison forts et ainsi favoriser le brassage des caractères très conservés. Toutefois, il semble par la suite nécessaire de déterminer quels types de gènes vont ainsi pouvoir être brassés et leur intérêt pour la sélection de nouvelles variétés (gènes de résistance aux maladies par exemple). L'augmentation de la fréquence de recombinaison associée à la modification de la distribution des COs permettront dans tous les cas de réduire le nombre de plantes générées en sélection variétale pour obtenir les combinaisons alléliques désirées chez un individu.

BIBLIOGRAPHIE

Articles périodiques :

Adams KL, Wendel JF. 2005. Polyploidy and genome evolution in plants. *Current Opinion in Plant Biology*, 8(2): 135-141.

Anderson LK, Stack SM. 2002. Meiotic recombination in plants. *Current Genomics*, 3(6): 507-525.

Anderson LK, Doyle GG, Brigham B, Carter J, Hooker KD, Lai A, Rice M, Stack SM. 2003. High-resolution crossover maps for each bivalent of Zea mays using recombination nodules. *Genetics*, 165: 849-865.

Anderson LK, Lai A, Stack SM, Rizzon C, Gaut BS. 2006. Uneven distribution of expressed sequence tag loci on maize pachytene chromosomes. *Genome Resources*, 16: 115-122.

Atcheson CL, DiDomenico B, Frackman S, Esposito RE, Elder RT. 1987. Isolation, DNA sequence, and regulation of a meiosis-specific eukaryotic recombination gene. *Proceeding of the National Academy of Sciences*, 84(22): 8035-8039.

Baudat F, De Massy M. 2007. Regulating double-stranded DNA break repair towards crossover or non-crossover during mammalian meiosis. *Chromosome Research*, **15:** 565-577.

Baumann P, Benson FE, West SC. 1996. Human Rad51 protein promotes ATP dependent homologous pairing and strand transfer reactions in vitro. *Cell*, 87: 757-766.

Bhatt AM, Canales C, Dickinson HG. 2001. Plant meiosis: the means to 1N. *Plant Science*, 6: 114-121.

Birdsell JA. 2002. Integrating genomics, bioinformatics, and classical genetics to study the effects of recombination on genome evolution. *Molecular Biology and Evolution*, 19(7): 1181-97.

Borde, V., T.C. Wu, and M. Lichten. 1999. Use of a recombination reporter insert to define meiotic recombination domains on chromosome III of Saccharomyces cerevisiae. *Molecular Biology and Evolution*, 19: 4832-4842

Borde V. 2007. The multiple roles of the Mre11 complex for meiotic recombination. *Chromosome Research*, 15: 551-563.

Borde, V., N. Robine, W. Lin, S. Bonfils, V. Geli, and A. Nicolas. 2008. Histone H3 lysine 4 trimethylation marks meiotic recombination initiation sites. *The EMBO Journal*, 28: 99-111.

Börner GV, Kleckner N, Hunter N. 2004. Crossover/Non-crossover Differentiation, Synaptonemal Complex Formation, and Regulatory Surveillance at the Leptotene/Zygotene Transition of Meiosis. *Cell*, 117(1): 29-45.

Broman KW, Weber JL. 2000. Characterization of human crossover interference. *The American Journal of Human Genetics*, 66: 1911-1926.

Broman KW, Rowe LB, Churchill GA, Paigen K. 2002. Crossover interference in the mouse. *Genetics*, 160: 1123-1131.

Brzeski J, Jerzmanowski B. 2003. Deficient in DNA Methylation 1 (DDM1) Defines a Novel Family of Chromatin-remodeling Factors. *The Journal of Biological Chemistry*, 278: 823-828.

Carlton PM, Farruggio AP, Dernburg AF. 2006. A link between meiotic prophase progression and crossover control. *PLoS Genetics*, 2**:** e12.

Chan SWL, Henderson IR, Jacobsen SE. 2005. Gardening the genome: DNA methylation in *Arabidopsis thaliana*. *Nature Reviews Genetics* **6**: 351-360.

Chen M, Presting G, Barbazuk WB, Goicoechea JL, Blackmon B, Fang G, Kim H, Frisch D, Yu Y, Sun S, Higingbottom S, Phimphilai J, Phimphilai D, Thurmond S, Gaudette B, Li P, Liu J, Hatfield J, Main D, Farrar K, Henderson C, Barnett L, Costa R, Williams B, Walser S, Atkins M, Hall C, Budiman MA, Tomkins JP, Luo M, Bancroft I, Salse J, Regad F, Mohapatra T, Singh NK, Tyagi AK, Soderlund C, Dean RA, Wing RA. 2002. An integrated physical and genetic map of the rice genome. *The Plant Cell*, 14**:** 537-545.

Chen C, Yu Q, Hou S, Li Y, Eustice M, Skelton RL, Veatch O, Herdes RE, Diebold L, Saw J, Feng Y, Qian W, Bynum L, Wang L, Moore PH, Paull RE, Alam M, Ming R. 2007. Construction of a sequence-tagged high-density genetic map of papaya for comparative structural and evolutionary genomics in brassicales. *Genetics*, 177**:** 2481-2491.

Chen SY, Tsubouchi T, Rockmill B, Sandler JS, Richards DR, Vader G, Hochwagen A, Roeder GS, Fung JC. 2008. Global analysis of the meiotic crossover landscape. *Developmental Cell*, 15(3): 401-415.

Cheng Z, Presting GG, Buell CR, Wing RA, Jiang J. 2001. High-resolution pachytene chromosome mapping of bacterial artificial chromosomes anchored by genetic markers reveals the centromere location and the distribution of genetic recombination along chromosome 10 of rice. *Genetics*, 157: 1749-1757.

Cheng F, Mandáková T, Wu J, Xie O, Lysak MA, Wang X. 2013. Deciphering the diploid ancestral genome of the mesohexaploid *Brassica rapa*. *The Plant Cell*, tpc.113.110486.

Cheung F, Trick M, Drou N, Lim TP, Park JY, Kwon SJ, Kim JA, Scott R, Pires JC, Paterson AH, Town C, Bancroft I. 2009. Comparative analysis between homoeologous genome segments of *Brassica napus* and its progenitor species reveals extensive sequence-level divergence. *The Plant Cell*, 21(7): 1912-1928.

Chevre AM, Eber F, Thomas G, Baron F. 1989. Cytological studies of tetraploid kale (*Brassica oleracea* L. ssp Acephala) obtained from diploid lines after colchicine treatment. *Agronomie*, 9: 521-525.

Choi SR, Teakle GR, PlahaP, Kim JH, Allender CJ, Beynon E, Piao ZY, Soengas P, Han TH, King GJ, Barker GC, Hand P, Lydiate DJ, Batley J, Edwards D, Koo DH, Bang JW, Park BS, and Lim YP. 2007. The reference genetic linkage map for the multinational *Brassica rapa* genome sequencing project. *Theoretical and Applied Genetics*, 115: 777-792.

Choo KH. 1998. Why is the centromere so cold? *Genome Resources*, 8, 81-82.

Copenhaver GP, Housworth EA, Stahl FW. 2002. Crossover interference in Arabidopsis. *Genetics*, 160: 1631-1639

Cromie GA, Smith GR. 2007. Branching out: meiotic recombination and its regulation. *Review of Cell Biology*, 17(9): 448-455.

De Massy B. 2003. Distribution of meiotic recombination sites. *Genetics*, 19: 514-522.

Delourme R, Falentin C, Fomeju B, Boillot M, Lassalle G, Andre I, Duarte J, Gauthier V, Lucante N, Marty A, Pauchon M. 2013. High-density SNP-based genetic map development and linkage disequilibrium assessment in *Brassica napus* L. *BMC Genomics*, 14: 120.

Dooner HK, Martinez-Ferez IM. 1997. Recombination occurs uniformly within the bronze gene, a meiotic recombination hotspot in the maize genome. *Plant Cell*, 9: 1633-1646.

Dooner HK, He L. 2008. Maize genome structure variation: interplay between retrotransposon polymorphisms and genic recombination. *The Plant Cell*, 20: 249-258.

Doré C, Varoquaux F. 2006. Histoire et amélioration de cinquante plantes cultivées. Quae (ed.), pp. 817.

Drouaud J, Camilleri C, Bourguignon PY, Canaguier A, Berard A, Vezon D, Giancola S, Brunel D, Colot V, Prum B, Quesneville H, Mezard C. 2006. Variation in crossing-over rates across chromosome 4 of *Arabidopsis thaliana* reveals the presence of meiotic recombination "hot spots". *Genome Resources*, 16: 106-114.

Drouaud J, Mercier R, Chelysheva L, Berard A, Falque M, Martin O, Zanni V, Brunel D, Mezard C. 2007. Sex-specific crossover distributions and variations in interference level along *Arabidopsis thaliana* chromosome 4. *PLoS Genetics*, 3: e106.

Dumas D, Britton-Davidian J. 2002. Chromosomal rearrangements and evolution of recombination: comparison of chiasma distribution patterns in standard and robertsonian populations of the house mouse. *Genetics*, 162: 1355-1366

Erayman M, Sandhu D, Sidhu D, Dilbirligi M, Baenziger PS, Gill KS. 2004. Demarcating the gene-rich regions of the wheat genome. *Nucleic Acids Research*, 32: 3546- 3565.

Esch E, Szymaniak JM, Yates H, Pawlowski WP, Buckler ES. 2007. Using crossover breakpoints in recombinant inbred lines to identify quantitative trait loci controlling the global recombination frequency. *Genetics*, 177: 1851-1858.

Esselink GD, Nybom H, Vosman B. 2004. Assignment of allelic configuration in polyploids using the MAC-PR (microsatellite DNA allele counting-peak ratios) method. *Theoretical and Applied Genetics*, 109: 402-408.

Falque M, Mercier R, Mézard C, de Vienne D, Martin OC. 2007. Patterns of recombination and mlh1 foci density along mouse chromosomes: modeling effects of interference and obligate chiasma. *Genetics*. 176: 1453-1467.

Falque M, Anderson LK, Stack SM, Gauthier F, Martin OC. 2009. Two types of meiotic crossovers coexist in maize. *The Plant Cell*, 21: 3915-3925.

Finnegan EJ, Kovac KA. 2000. Plant DNA methyltransferases. Plant Molecilar Biology, 43: 189-201.

Foss E, Lande R, Stahl FW, Steinberg CM. 1993. Chiasma interference as a function of genetic distance. *Genetics*, 133(3): 681-691.

Francis KE, Lam SY, Harrison BD, Bey AL, Berchowitz LE, Copenhaver GP. 2007. Pollen tetrad-based visual assay for meiotic recombination in Arabidopsis. *Proceeding of the National Academy of Sciences*, 104: 3913-3918.

Gaeta RT, Pires JC. 2010. Homoeologous recombination in allopolyploids: the polyploid ratchet. *New Phytologist*, 186: 18-28.

Gao M, Li G, Yang B, Qiu D, Farnham M, Quiros C. 2007. High-density *Brassica oleracea* linkage map: identification of useful new linkages. *Theoretical and Applied Genetics*, 115: 277-287.

Gaskell LJ, Osman F, Gilbert RJC, Whitby MC. 2007. Mus81 cleavage of Holliday junctions: a failsafe for processing meiotic recombination intermediates? *The EMBO Journal***,** 26: 1891 - 1901.

Gaut BS, Wright SI, Rizzon C, Dvorak J, Anderson LK. 2007. Recombination: an underappreciated factor in the evolution of plant genomes. *Nature Reviews Genetics*, 8: 77-84.

Gauthier F, Martin OC, Falque M. 2011. CODA (crossover distribution analyzer): quantitative characterization of crossover position patterns along chromosomes. *BMC Bioinformatics*, 12: 27.

Gerton JL, Hawley RS. 2005. Homologous chromosome interactions in meiosis : diversity amidst conservation. *Nature Reviews Genetics*, 6: 477-487.

Giraut L, Falque M, Drouaud J, Pereira L, Martin OC, Mézard C. 2011. Genome-wide crossover distribution in *Arabidopsis thaliana* meiosis reveals sex-specific patterns along chromosomes. *PLoS Genetics*, 7(11): e1002354.

Gomez-Campo C, Prakash S. 1999. Origin and domestication. Gomez-Campo C. (ed.), *Biology of Brassica Coenospecies*, 33-58. Elsevier, Amsterdam.

Gruenbaum Y, Naveh-Many T, Cedar H, Razin A. 1981. Sequence specificity of methylation in higher plant DNA. *Nature*, 292: 860-62.

Hafiz IA, Anjum MA, Grewal AG, Chaudhary GA. 2001. DNA methylation, an essential mechanism in plant molecular biology. Acta Physiologiae Plantarum, 23: 491-499.

Hall AE, Kettler GC, Preuss D. 2006. Dynamic evolution at pericentromeres. *Genome Resources*, 16: 355-364.

Hamant O, Ma H, Cande WZ. 2006. Genetics of meiotic prophase I in plants. *Annual Review in Plant Biology*, 57: 267-302.

Harrison GE, Heslop-Harrison GS. 1995. Centromeric repetitive DNA sequences in the genus *Brassica*. *Theoretical and Applied Genetics*, 90(2): 157-165.

Hillers KJ, Villeneuve AM. 2003. Chromosome-wide control of meiotic crossing over in C. elegans. *Current Biology*, 13: 1641-1647.

Hollingsworth NM, Brill SJ. 2004. The Mus81 solution to resolution: generating meiotic crossovers without Holliday junctions. *Genes & Development*, 18(2): 117-125.

Hunt PA. 2006. Meiosis in mammals: recombination, non-disjunction and the environment. *Biochemical Society Transactions*, 34: 574-577.

Jackson N, Sanchez-Moran E, Buckling E, Armstrong SJ, Jones GH, Franklin FCH. 2006. Reduced meiotic crossovers and delayed prophase I progression in AtMLH3-deficient Arabidopsis. *The EMBO Journal*, 25: 1315-1323

Jayathilaka K, Sheridan SD, Bold TD, Bochenska K, Logan HL, Weichselbaum RR, Bishop DK, Connell PP. 2008. A chemical compound that stimulates the human homologous recombination protein RAD51. *Proceeding of the National Academy of Sciences*, 105(4): 15848-15853.

Jensen-Seaman MI, Furey TS, Payseur BA, Lu Y, Roskin KM, Chen CF, Thomas MA, Haussler D, Jacob HJ. 2004. Comparative recombination rates in the rat, mouse, and human genomes. *Genome Resources*, 14: 528-538.

Jenczewski E, Eber F, Grimaud A, Huet S, Lucas MO, Monod H, Chevre AM. 2003. PrBn, a Major Gene Controlling Homeologous Pairing in Oilseed Rape (*Brassica napus*) Haploids. *Genetics*, 164: 645-653.

Jenczewski E, Chèvre AM, Alix K. 2013. Chromosomal and Gene Expression Changes in *Brassica* Allopolyploids. *Polyploid and hybrid genomics*. (eds), John Wiley & Sons, Inc., Oxford, UK.

Jones GH. 1984. The control of chiasma distribution. *Symposia of the Society for Experimental Biology*, 38: 293-320.

Jost JP, Saluz HP. 1993. DNA methylation: molecular biology and biological significance. Springer-Verlag (ed.), Basel, pp. 572.

Kaback DB. 1996. Chromosome-size dependent control of meiotic recombination in humans. *Nature Genetics*, 13: 20-21

Keeney S, Kleckner N. 1995. Covalent protein-DNA complexes at the 5' strand termini of meiosis-specific double-strand breaks in yeast. *Proceeding of the National Academy of Sciences*, 92(24): 11274-11278.

Keeney S, Neale MJ. 2006. Initiation of meiotic recombination by formation of DNA double-strand breaks: mechanism and regulation. *Biochemical Society Transactions*, 34(4): 523-525.

Kniewel R, Keeney S. 2009. Histone methylation sets the stage for meiotic DNA breaks. *The EMBO Journal*, 28: 81-83.

Knoll A, Higgins JD, Seeliger K, Reha SL, Dangel NJ, Bauknecht M, Schröpfer M, Franklin FCH, Puchta H. 2012. The Fanconi Anemia Ortholog FANCM Ensures Ordered Homologous Recombination in Both Somatic and Meiotic Cells in Arabidopsis. *The Plant Cell*, 24(4): 1448-1464.

Kong A, Gudbjartsson DF, Sainz J, Jonsdottir GM, Gudjonsson SA, Richardsson B, Sigurdardottir S, Barnard J, Hallbeck B, Masson G, Shlien A, Palsson ST, Frigge ML, Thorgeirsson TE, Gulcher JR, Stefansson K. 2002. A high-resolution recombination map of the human genome. *Nature Genetics*, 31: 241-247.

Kosambi DD. 1944. The estimation of map distances from recombination values. *Annual* Eugen, 12: 172-175.

Koo DH, Hong CP, Batley J, Chung YS, Edwards D, Bang JW, Hur Y, Lim YP. 2011. Rapid divergence of repetitive DNAs in Brassica relatives. *Genomics*, 97(3): 173-185.

Kumar A, Bennetzen JL. 1999. Plant retrotransposons. Annual Review of Genetics, 33: 479-532.

Kunzel G, Korzun L, Meister A. 2000. Cytologically integrated physical restriction fragment length polymorphism maps for the barley genome based on translocation breakpoints. *Genetics*, 154: 397-412.

Kurtz S, Phillippy A, Delcher AL, Smoot M, Shumway M, Antonescu C, Salzberg SL. 2004. Versatile and open software for comparing large genomes. *Genome Biology*, 5: R12.

Lagercrantz U., Lydiate D. J. (1996). Comparative genome mapping in *Brassica*. *Genetics*, 144: 1903-1910.

Lamb NE, Sherman SL, Hassold TJ. 2005. Effect of meiotic recombination on the production of aneuploid gametes in humans. *Cytogenetic and Genome Research*, 111: 250-255.

Lander ES, Green P, Abrahamson J, Barlow A, Daly MJ, Lincoln SE, Newburg L. 1987. MAPMAKER: an interactive computer package for constructing primary genetic linkage maps of experimental and natural populations. *Genomics*, 1: 174-181.

Leflon M, Eber F, Letanneur J, Chelysheva L, Coriton O, Huteau V, Ryder C, Barker G, Jenczewski E, Chèvre A. 2006. Pairing and recombination at meiosis of *Brassica rapa* (AA) × *Brassica napus* (AACC) hybrids. *Theoretical and Applied Genetics*, 113(8): 1467-1480.

Leflon M. 2007. Effets de la structure de l'hybride et du fonds génétique sur le taux de recombinaison entre les génomes A de *Brassica rapa* (AA, 2n=20) et de *B. napus* (AACC, 2n=38). Application à l'introgression de gènes de résistance à *Leptosphaeria maculans* dans le colza. Thèse de l'Ecole Normale Supérieure Agronomique de Rennes, p. 112.

Leflon M, Grandont L, Eber F, Huteau V, Coriton O, Chelysheva L, Jenczewski E, Chèvre AM. 2010. Crossovers get a boost in *Brassica* allotriploid and allotetraploid hybrid. *The Plant Cell*, 22: 2253-2264.

Lhuissier FGP, Offenberg HH, Wittich PE, Vischer NOE, Heyting C. 2007. The mismatch repair protein MLH1 marks a subset of strongly interfering crossovers in tomato. *The Plant Cell*, 19: 862-876.

Lim KB, De Jong H, Yang TJ, Park JY, Kwon SJ, Lim MH, Kim JA, Jin M, Jin YM, Kim SH, Lim YP, Bang JW, Kim HI, Park BS. 2005. Characterization of rDNAs and tandem repeats in heterochromatin of *Brassica rapa*. *Molecular Cell*, 19: 436-444.

Lim KB, Yang TJ, Hwang YJ, Park JY, Kwon SJ, Lim MH, Kim JA, Jin M, Jin YM, Kim SH, Lim YP, Bang JW, Kim HI, Park BS. 2007. Characterization of the centromere and peri centromere retrotransposons in *Brassica rapa* and their distribution in related Brassica species. *The Plant Journal*, 49: 173-183.

Liu Z, Adamczyk K, Manzanares-Dauleux M, Eber F, Lucas M-O, Delourme R, Chevre AM, Jenczewski E. 2006. Mapping PrBn and Other Quantitative Trait Loci Responsible for the Control of Homeologous Chromosome Pairing in Oilseed Rape (*Brassica napus* L.) Haploids. *Genetics*, 174: 1583-1596.

Lombard V, Delourme R. 2001. A consensus linkage map for rapeseed (*Brassica napus* L.): construction and integration of three individual maps from DH populations. *Theoretical and Applied Genetics*, 103: 491-507.

Luce AC, Sharma A, Mollere OS. 2006. Precise Centromere mapping using a combination of repeat junction markers and chromatin immunoprecipitation-polymerase chain reaction. *Genetics*, 174: 1057-1061.

Lukaszewski J, Curtis CA. 1993. Physical distribution of recombination in B-genome chromosomes of tetraploid wheat. *Theoretical and Applied Genetics*, 86: 121-127.

Lynn A, Soucek R, Borner GV. 2007. ZMM proteins during meiosis: crossover artists at work. *Chromosome Research*, 15(5): 591-605.

Lysak MA, Koch MA, Pecinka A, Schubert I. 2005. Chromosome triplication found across the tribe *Brassiceae*. *Genome Research*, 15: 516-525.

Maloisel L, Rossignol JL. 1998. Suppression of crossing-over by DNA methylation in *Ascobolus*. *Genes & Development*, 12: 1381-1389.

Mancera E, Bourgon R, Brozzi A, Huber W, Steinmetz LM. 2008. High-resolution mapping of meiotic crossovers and non-crossovers in yeast. *Nature*, 454: 479-485.

Martienssen RA, Colot V. 2001. DNA methylation and epigenetic inheritance in plants and filamentous fungi. *Plant Physiology*, 129: 733-746.

Martini E, Diaz RL, Hunter N, Keeney S. 2006. Crossover homeostasis in yeast meiosis. *Cell*, 126: 285-295.

McPeek MS, Speed TP. 1995. Modeling Interference in Genetic Recombination. *Genetics*, 139: 1031-1044.

Melamed-Bessudo C, Levy AA. 2012. Deficiency in DNA methylation increases meiotic crossover rates in euchromatic but not in heterochromatic regions in Arabidopsis. *Proceeding of the National Academy of Sciences*, 109(16): 981-988.

Merker JD, Dominska M, Greenwell PW, Rinella E, Bouck DC, Shibata Y, Strahl BD, Mieczkowski P, Petes TD. 2008. The histone methylase Set2p and the histone deacetylase Rpd3p repress meiotic recombination at the HIS4 meiotic recombination hotspot in Saccharomyces cerevisiae. *DNA Repair*, 7: 1298-1308.

Mézard C. 2006. Meiotic recombination hotspots in plants. *Biochemical Society Transactions*, 34: 531-534.

Mézard C, Vignard J, Drouaud J, Mercier R. 2007. The road to crossovers: plants have their say. *Review of Genetics*, 23(2): 91-99.

Morton NE. 1956. The detection and estimation of linkage between genes for elliptocytosis and the Rh blood type. *The American Journal of Human Genetics*, 8: 80-96.

Muller HJ. 1916. The mechanisms of crossing-over. *American Naturalist*, 50: 193-221.

Myers S, Bottolo L, Freeman C, McVean G, Donnelly P. 2005. A fine-scale map of recombination rates and hotspots across the human genome. *Science*, 310: 321-324.

Myers S, Spencer SS. 2006. The distribution and causes of meiotic recombination in the human genome. *Biochemical Society Transactions*, 34(4): 526-30.

Nicolas SD, Leflon M, Monod H, Eber F, Coriton O, Huteau V, Chèvre AM, Jenczewski E. 2009. Genetic regulation of meiotic cross-overs between related genomes in *Brassica napus* haploids and hybrids. *The Plant Cell*, 21(2): 373-385.

Osman K, Higgins JD, Sanchez-Moran E, Armstrong SJ, Franklin FC. 2011. Pathways to meiotic recombination in *Arabidopsis thaliana*. *New Phytologist*, 190(3): 523-544.

Paigen K, Szatkiewicz JP, Sawyer K, Leahy N, Parvanov ED, Ng SH, Graber JH, Broman KW, Petkov PM. 2008. The recombinational anatomy of a mouse chromosome. *PLoS Genetics*, 4: e1000119.

Parkin IAP, Sharpe AG, Keith DJ, Lydiate DJ. 1995. Identification of the A and C genomes of amphidiploid *Brassica napus* (oilseed rape). *Genome*, 38(6): 1122-1131.

Parkin IAP, Sharpe AG, Lydiate DJ. 2003. Patterns of genome duplication within the *Brassica napus* genome. *Genome*, 46: 291-303.

Parkin IAP, Gulden SM, Sharpe AG, Lukens L, Trick M, Osborn TC, Lydiate DJ. 2005. Segmental structure of the *Brassica napus* genome based on comparative analysis with *Arabidopsis thaliana*. *Genetics*, 171: 765-78.

Paszkowski J, Whitham S. 2001. Gene silencing and DNA methylation processes. *Current Opinion in Plant Biology*, 4: 123-129.

Pecinka A, Fang W, Rehmsmeier M, Levy A, Scheid OM. 2011. Polyploidization increases meiotic recombination frequency in Arabidopsis. *BMC Biology*, Vol. 9: 24.

Petes, T.D. 2001. Meiotic recombination hot spots and cold spots. *Nature Reviews Genetics*, 2: 360-369.

Petronczki M, Siomos MF, Nasmyth K. 2003. Un ménage à quatre: the molecular biology of chromosome segregation in meiosis. *Cell*, 112(4): 423-40.

Piquemal J, Cinquin E, Couton F, Rondeau C, Seignoret E, Doucet I, Perret D, Villeger MJ, Vincourt P, Blanchard P. 2005. Construction of an oilseed rape (*Brassica napus* L.) genetic map with SSR markers. *Theoretical and Applied Genetics*, 111: 1514-1523.

Plagnol V, Padhukasahasram B, Wall JD, Marjoram P, Nordborg M. 2006. Relative influences of crossing over and gene conversion on the pattern of linkage disequilibrium in *Arabidopsis thaliana*. *Genetics*, 172: 2441-2448.

Pouilly N, Delourme R, Alix K, Jenczewski E. 2008. Repetitive sequence-derived markers tag centromeres and telomeres and provide insights into chromosome evolution in *Brassica napus*. *Chromosome Research*, 16: 683-700.

Pugh T, Fouet O, Risterucci AM, Brottier P, Abouladze M, Deletrez C, Courtois B, Clement D, Larmande P, N'Goran JA, Lanaud C. 2004. A new cacao linkage map based on codominant markers: development and integration of 201 new microsatellite markers. *Theoretical and Applied Genetics*, 108: 1151-1161.

Rana D, Van Den Boogaart T, O'Neill CM, Hynes L, Bent E, Macpherson L, Park JY, Lim YP, Bancroft I. 2004. Conservation of the microstructure of genome segments in *Brassica napus* and its diploid relatives. *The plant journal*, 40(5): 725-733.

Rizzon C, Marais G, Gouy M, Biémont C. 2002. Recombination Rate and the Distribution of Transposable Elements in the *Drosophila melanogaster* Genome. *Genome Research*, 12: 400–407.

Roig I, Keeney S. 2008. Probing meiotic recombination decisions. *Development Cell*, 15: 331- 332.

Round EK, Flowers SK, Richards EJ. 1997. *Arabidopsis thaliana* centromere regions: genetic map positions and repetitive DNA structure. *Genome Research*, 7: 1045-1053.

Rozen S, Skalestsky H. 2000. Primer3 on the WWW for general users and for biologist programmers. *Methods in Molecular Biology*, 132: 365-386.

Saintenac C. 2009. Analyse de la distribution des crossing-overs sur le chromosome 3B du blé tendre (Triticum aestivum) et des facteurs influençant cette distribution. Thèse de l'Ecole Doctorale des Sciences de la Vie et de la Santé de Clermont-Ferrant II Blaise Pascale. 383 p.

Säll T, Flink F, Bengtsson BO. 1990. Genetic control of recombination in barley: Variation in recombination frequency measured with inversion heterozygotes. *Herditas*, 112(2): 157-170.

Sanchez-Moran E, Armstrong SJ, Santos JL, Franklin FC, Jones GH. 2002. Variation in chiasma frequency among eight accessions of *Arabidopsis thaliana*. *Genetics*, 162: 1415-1422.

Schmidt T, Heslop-Harrison JS. 1998. Genomes, genes and junk: the large scale organization of plant chromosomes. *Plant Science*, 3: 195-199.

Schranz E**,** Lysak MA, Mitchell-Olds T. 2006. The ABC's of comparative genomics in the Brassicaceae: building blocks of crucifer genomes. *Plant Science*, 11(11): 535-542.

Schuelke M. 2000. An economic method for the fluorescent labeling of PCR fragments. *Nature Biotechnology*, 18: 233-234.

Sehorn MG, Sigurdsson S, Bussen W, Unger VM, Sung P. 2004. Human meiotic recombinase Dmc1 promotes ATP-dependent homologous DNA strand exchange. *Nature*, 429: 433-437.

Shaw P, Moore G. 1998. Meiosis: vive la difference! *Current Opinion in Plant Biology*, 1: 458-462.

Sherman JD, Stack SM. 1995. Two-dimensional spreads of synaptonemal complexes from solanaceous plants. VI. High-resolution recombination nodule map for tomato (*Lycopersicon esculentum*). *Genetics*, 141: 683-708.

Shifman S, Bell JT, Copley RR, Taylor MS, Williams RW, Mott R, Flint J. 2006. A high-resolution single nucleotide polymorphism genetic map of the mouse genome. *PLoS Biology*, 4: e395.

Smith KN, Nicolas A.1998. Recombination at work for meiosis. *Current Opinion in Genetic & Development*, 8*:* 200-211.

Suay L, Zhang D, Eber F, Joury H, Lodé M, Huteau V, Coriton O, Szadkowski E, Leflon M, Martin O, Falque M, Jenczewski E, Paillard S, Chèvre AM. 2013. Crossover rate between homologous chromosomes and interference are regulated by the addition of specific unpaired chromosomes in Brassica. *New Physiologist*, in press.

Sun Z, Wang Z, Tu J, Zhang J, Yu F, McVetty PB, Li G. 2007. An ultradense genetic recombination map for *Brassica napus*, consisting of 13551 SRAP markers. *Theoretical and Applied Genetics*, 114: 1305-1317.

Szadkowski E, Eber F, Huteau V, Lodé M, Huneau C, Belcram H, Coriton O, Manzanares-Dauleux MJ, Delourme R, King GJ, Chalhoub B, Jenczewski E, Chèvre AM. 2010. The first meiosis of resynthesized *Brassica napus*, a genome blender. *New Phytologist*. 186 (1): 102-112.

Szadkowski E, Eber F, Huteau V, Lodé M, Coriton O, Jenczewski E, Chèvre AM. 2011. Polyploid formation pathways have an impact on genetic rearrangements in resynthesized *Brassica napus. New Phytologist*, 191(3): 884-894.

Troggio M, Malacarne G, Coppola G, Segala C, Cartwright DA, Pindo M, Stefanini M, Mank R, Moroldo M, Morgante M, Grando MS, Velasco R. 2007. A dense single-nucleotide polymorphism-based genetic linkage map of grapevine (*Vitis vinifera* L) anchoring Pinot Noir bacterial artificial chromosome contigs. *Genetics*, 176: 2637-2650.

Truco MJ, Antonise R, Lavelle D, Ochoa O, Kozik A, Witsenboer H, Fort SB, Jeuken MJ, Kesseli RV, Lindhout P, Michelmore RW, Peleman J. 2007. A highdensity, integrated genetic linkage map of lettuce (*Lactuca* spp). *Theoretical and Applied Genetics*, 115: 735-746.

U N. 1935. genome analysis in Brassica with special reference to the experimental formation of *B. napus* and peculiar mode of fertilization. *Journal of Botany*, 7: 389-452.

Wang X, Wang H, Wang J, Sun R, Wu J, Liu S, Bai Y, Mun J,H, Bancroft I, Cheng F, Huang S, Li X, Hua W, Wang J, Wang X, Freeling M, Pires J,C, Paterson A,H, Chalhoub B, Wang B, Hayward A, Sharpe A,G, Park B,S, Weisshaar B, Liu B, Li B, Liu B, Tong C, Song C, Duran C, Peng C, Geng C, Koh C, Lin C, Edwards D, Mu D, Shen D, Soumpourou E, Li F, Fraser F, Conant G, Lassalle G, King G,J, Bonnema G, Tang H, Wang H, Belcram H, Zhou H, Hirakawa H, Abe H, Guo H, Wang H, Jin H, Parkin I,A, Batley J, Kim J,S, Just J, Li J, Xu J, Deng J, Kim J,A, Li J, Yu J, Meng J, Wang J, Min J, Poulain J, Wang J, Hatakeyama K, Wu K, Wang L, Fang L, Trick M, Links M,G, Zhao M, Jin M, Ramchiary N, Drou N, Berkman P,J, Cai Q, Huang Q, Li R, Tabata S, Cheng S, Zhang S, Zhang S, Huang S, Sato S, Sun S, Kwon S,J, Choi S,R, Lee T,H, Fan W, Zhao X, Tan X, Xu X, Wang Y, Qiu Y, Yin Y, Li Y, Du Y, Liao Y, Lim Y, Narusaka Y, Wang Y, Wang Z, Li Z, Wang Z, Xiong Z, Zhang Z. 2011. The genome of the mesopolyploid crop species *Brassica rapa*. *Nature Genetics*, 43: 1035-1039.

Wijnker E, De Jong H. 2008. Managing meiotic recombination in plant breeding. *Review of Plant Cell*, 13(12): 640-646.

Winckler W, Myers SR, Richter DJ, Onofrio RC, McDonald GJ, Bontrop RE, McVean GAT, Gabriel SB, Reich D, Donnelly P, Altshuler D. 2005. Comparison of Fine-Scale Recombination Rates in Humans and Chimpanzees. *Science*, 308(5718): 107-111.

Wu J, Mizuno H, Hayashi-Tsugane M, Ito Y, Chiden Y, Fujisawa M, Katagiri S, Saji S, Yoshiki S, Karasawa W, Yoshihara R, Hayashi A, Kobayashi H, Ito K, Hamada M, Okamoto M, Ikeno M, Ichikawa Y, Katayose Y, Yano M, Matsumoto T, Sasaki T. 2003. Physical maps and recombination frequency of six rice chromosomes. *Plant Journal*, 36: 720-730.

Xiong Z, Pires JC. 2011. Karyotype and Identification of All Homoeologous Chromosomes of Allopolyploid *Brassica napus* and Its Diploid Progenitors. *Genetics*, 187: 37-49.

Yang TJ, Kim JS, Kwon SJ, Lim KB, Choi BS, Kim JA, Jin M, Park JY, Lim MH, Kim HI, Lim YP, Kang JJ, Hong JH, Kim CB, Bhak J, Bancroft I, Park BS. 2006. Sequence-level analysis of the diploidization process in the triplicated FLC region of *Brassica rapa*. *The Plant Cell*, 18: 1339-1347.

Yelina NE, Choi K, Chelysheva L, Macaulay M, de Snoo B, Wijnker E, Miller N, Drouaud J,Grelon M, Copenhaver GP, Mezard C, Kelly K, Henderson IR. 2012. Epigenetic Remodeling of Meiotic Crossover Frequency in *Arabidopsis thaliana* DNA Methyltransferase Mutants. *PLoS Genetics*, 8(8): e1002844

Youds JL, Boulton SJ. 2011. The choice in meiosis, defining the factors that influence crossover or non-crossover formation. *Journal of Cell Science*, 124: 501-513.

Zickler D, Kleckner N. 1999. Meiotic chromosomes: integrating structure and function. *Annual Review of Genetics*, 33: 603-754.

Sitographie :

FAOSTAT (2012). FAOSTAT Browse - Data 2010. http://faostat3.fao.org/ (consulté le 09/06/2013)

PROLEA (2013). La filière française des huiles et protéines végétales - Colza. http://prolea.com/index.php?id=1477 (consulté le 09/06/2013)

AGROCAMPUS OUEST

CFR Angers

2 rue André Le Nôtre

49045 Angers

INSTITUT NATIONAL DE LA RECHERCHE AGRONOMIQUE

INRA Le Rheu

Domaine de la Motte au Vicomte BP 35327

35653 LE RHEUX CEDEX

Mémoire de Fin d'Études

Diplôme d'Ingénieur de l'Institut Supérieur des Sciences Agronomiques, Agroalimentaires, Horticoles et du Paysage

Année universitaire : 2012-2013

Spécialisation ou option : Amélioration des plantes

Annexes

Par : Alexandre PELE

*Volet à renseigner par l'enseignant responsable de l'option/spécialisation** Ou son représentant	Bon pour dépôt (version définitive) ❑
Date : …./…/… Signature	Autorisation de diffusion : Oui ❑ Non❑

Devant le jury *:* Soutenu à Rennes, le 12 Septembre 2013

Sous la présidence de : Mme Maria MANZANARES-DAULEUX

Maître de stage : Mme Anne-Marie CHEVRE

Enseignant référent : Mme Maria MANZANARES-DAULEUX

Autres membres du jury : Mr Frédéric LECERF et Mme Mélanie JUBAULT

Annexe I : Protocole d'extraction ADN

Unité : UMR APBV Centre de Rennes	Maxi préparation d'ADN génomique de plante (crucifères - pois)		Réf. : EXT10 Version 7
	Imprimé le 12/09/13		Créé le : 06/2004 Modifié le : Archivé le :
	Rédacteur	Vérificateur	Approbateur
Nom :	E. Jenczewski	C. Falentin	

Objectif : Extraire l'ADN de plantes en vue de tests PCR et/ou de digestion par des enzymes de restriction (ex : SSR, RFLP, AFLP). Il permet d'extraire de l'ADN de plantes lyophilisées. Ce protocole comprend une 2nde purification optionnelle au phénol et les précipitations ce font à l'isopropanol (100% et 70%).

Matériel :

A réserver : Salle d'extraction (sur calendrier).

A vérifier : Tubes 2 ml à vis autoclavés, tubes de 50 ml PP (Sarstedt), cônes jaunes et bleus autoclavés, pipettes plastique 3.5 ml jetables.

Préparer 3 séries de tubes 50ml identiques (1 série pour le prélèvement et la lyophilisation des feuilles et 2 séries pour la reprise des surnageants).

Matériel végétal :

Le matériel végétal a été préalablement lyophilisé pendant 66 heures puis broyé à l'aide de billes de verre (5 ml par tube) grâce au mélangeur à peinture.

Réactifs :

-Tampon CTAB (MON1) contenant 1% de β-mercaptoéthanol: Noter le n° de lot et le nom du préparateur. Ajouter le β-mercaptoéthanol au dernier moment.

-*Série de 36 tubes : 1 litre*

-Série de 72 tubes : 2 litres

-Chloroforme-octanol 24:1 v/v (MON 12): Noter le n° de lot et le fournisseur de chaque produit. Si du mélange est préparé à l'avance, il faut le conserver dans une bouteille hermétiquement close.

-Tampon TE 10-0.1 stérile (MON2) et TE 10-1 (MON6): Noter le n° de lot et le nom du préparateur.

-*Série de 36 tubes : 400 ml*

-Série de 72 tubes : 800 ml

-RNase 10 mg/ml (MON3): Noter le n° de lot et le nom du préparateur

-Chlorure de sodium (NaCl) 5M stérile (MON17): Noter le n° de lot et le nom du préparateur

-*Série de 36 tubes : 250 ml*

-Série de 72 tubes : 500 ml

-Isopropanol 100% et 70%

Timing :

J-7: Réserver la salle

J-2 : Vérifier s'il reste assez de tampon CTAB, de chloroforme, d'isopropanol, de tubes de 2 ml autoclavés, de cônes bleus autoclavés, de tubes de 50 ml (stériles d'origine).

Méthodologie:

Préparation des feuilles

Récolter environ 3-4 de préférence jeunes et de bon aspect dans des tubes de 50 ml PP (Sarsted) contenant 5 ml de billes de verre, correctement identifiés. Il est également possible de récolter les feuilles dans des sachets d'autofécondation correctement identifiés et de les transférer, une fois congelées, dans les tubes, juste avant de les lyophiliser.

Extraction

1- En arrivant: Mettre en route le bain-marie à 65°C. Régler la centrifugeuse KR4-22 sur 15°C **(***Sorbonne + gants).*

2- Préparer 25 ml par échantillon de tampon CTAB additionné de 1 % (v/v) de β-mercaptoéthanol dans une bouteille. Il est OBLIGATOIRE de réaliser la manipulation sous une sorbonne (le β-mercaptoéthanol est toxique). Faire chauffer le mélange dans le bain-marie sous la sorbonne pendant une dizaine de minutes (le CTAB froid précipite et peut faire précipiter l'ADN). Ajouter environ 25 ml de tampon CTAB dans chaque tube et bien agiter. Placer une série de 36 tubes au moins 2 heure à 65°C (détruit les membranes et libère l'ADN). Agiter régulièrement pour décoller la poudre et la laisser s'imbiber.

3- Après avoir laissé les tubes refroidir quelques minutes sur la paillasse, compléter à 50 ml avec du chloroforme-octanol 24-1. Bien homogénéiser en agitant jusqu'à formation d'une soupe (élimine les débris cellulaires et précipite les protéines restantes). Mélanger à nouveau en inversant les tubes par série de 24 une vingtaine de fois.

4- Centrifuger à 3800 t/mn et 4°C (centri KR4-22) pendant 45 mn minimum. Pendant la centrifugation, préparer une nouvelle série de tubes de 50 ml identifiés contenant 16 ml d'isopropanol 100%.

5- Récupérer la phase aqueuse (environ 21 ml) à l'aide d'une transferpette stérile. Transférer dans le tube correspondant identifié. Mélanger avec l'isopropanol en inversant les tubes. L'ADN précipite. Laisser décanter l'ADN au moins une heure à 4°C (dans la glace, le réfrigérateur ou la centrifugeuse KR4-22 réglée à 4°C pour une série de 36 tubes).

6- Centrifuger 3800 t/mn à 4°C (centri KR4-22) pendant 10 mn. Eliminer le surnageant et laisser sécher le tube à l'envers sur du papier absorbant pendant ¼ heure. Attention : vérifier que les culots ne se décrochent pas ! Mettre en route le bain-marie à 37°C.

7- Dissoudre le culot dans 10 ml de TE 10-1 stérile contenant 0.5 µg/ml de RNAse (0.5 µl de solution stock à 10 mg/ml par ml de TE). Placer à 37°c pendant 1 heure (agiter de temps en temps).

Remarque : vous pouvez stopper les manips à ce stade en conservant les tubes à 4°C jusqu'au lendemain.

Optionnel : Si le matériel végétal de départ était en mauvais état ou si le culot est anormalement important, ajouter 10 ml de phénol-chloroforme-octanol, agiter et centrifuger 3800 t/mn (centri KR4-22) pendant 10 mn à 15°C. Récupérer le surnageant dans un tube neuf.

8- Ajouter 10 ml de chloroforme-octanol, agiter et centrifuger 3800 t/mn (centri KR4-22) pendant 10 mn à 4°C. Récupérer le surnageant dans un tube neuf.
9- Ajouter 5 ml de NaCl 5M, agiter, puis 10 ml d'isopropanol 100%. Agiter à nouveau : l'ADN précipite. Laisser précipiter au minimum ½ heure à 4°C (dans la glace, le réfrigérateur ou la centrifugeuse KR4-22 réglée à 4°C pour une série de 36 tubes).
10- Centrifuger 3800 t/mn (centri KR4-22) pendant 5 mn à 4°C. Eliminer le surnageant. Ajouter 5 ml d'isopropanol 70%. Bien rincer le tube. Centrifuger 3800 t/mn (centri KR4-22) pendant 5 mn à 4°C. Eliminer le surnageant et laisser sécher le tube à l'envers sur du papier absorbant pendant ¼ heure. Coucher le tube sur le côté et laisser sécher quelques heures (en fonction de la taille du culot).
11- Resuspendre dans un volume minimum d'eau autoclavée (commencer par 100 µl). Doser l'ADN (au spectrophotomètre) et vérifier sa qualité en passant quelques échantillons sur minigel (voir AGA2).

Résultat attendu : De l'ADN de haut poids moléculaire non dégradé doit être obtenu (au même niveau que de l'ADN du phage lambda de 40 kbp). 20 à 100 µgrammes d'ADN sont attendus.

Organisation : Une ou deux (2 personnes) séries de 36 tubes doivent pouvoir être réalisées par jour, et cela au moins jusqu'à l'étape 7.

Annexe II : Mix et programmes PCR

MIX		
Réactifs	**Concentration finale**	**Vol. 1 réaction (µL)**
H2O up	-	1.808
Tp 5X	1X	1.4
DMSO	5%	0.35
MgCl2 25mM	2mM	0.56
dNTP 25mM	75µM	0.028
Amorce M13 10µM	0,36µM	0.252
Taq 5U/µL	0,03U/µL (0,15U)	0.042
Amorce U 1µM	0,04µM	0.28
Amorce L 10µM	0,40µM	0.28
Vol. MIX	-	5
Vol. ADN	-	2
Vol. final	-	7

Programme PCR	CB-Bras-MR-MD-NMS-m	
94°C	4min	
94°C	30sec	
TD 65°C à 54°C	1min	x 12 cycles
72°C	30sec	
94°C	30sec	
53°C	1min	x 25 cycles
72°C	30sec	
72°C	10min	
10°C	10min	

Programme PCR	SSR designés	
94°C	5min	
94°C	30sec	
60°C	1min	x 35 cycles
72°C	1min30	
72°C	10min	
10°C	10min	

Programme PCR	sR-Sn	
94°C	10min	
94°C	15s	
50°C	15s	x 8 cycles
72°C	30s	
89°C	15s	
50°C	15s	x 27 cycles
72°C	30s	
72°C	10min	
10°C	10min	

Programme PCR	cnu-nia	
95°C	15min	
94°C	1min	
55°C	1min	x 35 cycles
72°C	1min	
72°C	10min	
10°C	10min	

Annexe III : Résultats du polymorphisme des SSR designés

Marqueurs	Scaffolds	Amplification (pb)	C1.3	Chiifu	Darmor	RC	Polymorphe
P17A1299	17	0	0	0	0	0	NON
P17A1684	17	329	0	1	0	0	NON
		390	1	0	1	1	
P17A175	17	417	1	0	0	0	OUI
		418	0	0	1	0	
		419	0	1	0	0	
P17A1937	17	0	0	0	0	0	NON
P17A203	17	371	1	1	1	0	NON
		356	0	0	1	1	
P17A2197	17	359	1	0	0	0	OUI
		356	0	0	1	0	
P17A2812	17	0	0	0	0	0	NON
P17A303	17	403	0	1	0	0	OUI
		409	0	0	1	0	
P17A3064	17	414	1	0	0	0	OUI
		408	0	1	0	0	
		446	0	1	0	0	
		419	0	0	1	0	
P17A933	17	202	0	1	0	0	OUI
		189	0	0	1	0	
		198	0	0	0	1	
P17B120	17	264	0	0	1	0	NON
		268	0	0	0	1	
P17B127	17	286	1	0	0	0	OUI
		303	0	1	0	0	
		305	0	1	0	0	
		308	0	0	1	0	
		310	0	0	1	0	
P17B158	17	300	0	1	0	0	NON
		303	1	0	1	0	
P17B169	17	281	0	0	1	0	OUI
		282	1	0	0	0	
		287	0	1	0	0	
		289	0	0	1	0	
		322	0	0	0	1	
P17B184	17	317	0	0	1	0	NON
		320	0	0	0	1	
P17B186	17	325	1	0	0	0	NON
		411	1	1	1	1	
P17B198	17	187	0	0	1	0	NON
		197	1	0	1	0	
		202	1	0	0	0	
		213	0	1	1	0	
P17B204	17	196	1	0	0	0	OUI
		202	0	0	1	0	
P17B214	17	0	0	0	0	0	NON
P17B228	17	277	1	0	0	0	NON
	17	277	?	1	0	0	NON
		282	0	0	1	0	
P17B247	17	293	0	1	0	0	NON
P17B249	17	263	1	0	1	0	NON
		268	0	1	0	0	
		270	0	0	0	1	
P17B287	17	214	1	0	0	0	OUI
		217	0	0	1	1	
		219	0	1	0	0	
		222	0	0	1	0	
P17B296	17	233	0	1	0	0	NON
		238	0	0	0	0	
		248	1	0	1	0	
P17B428	17	0	0	0	0	0	NON
P17B450	17	311	0	0	0	1	OUI
		316	0	1	0	0	
		318	0	0	1	0	
P17B470	17	142	0	1	0	0	NON
		209	0	1	0	0	
P17B472	17	0	0	0	0	0	NON

Marqueurs	Scaffolds	Amplification (pb)	C1.3	Chiifu	Darmor	RC	Polymorphe
P17B478	17	244	0	1	1	0	NON
		265	0	0	1	0	
		304	0	1	0	0	
		315	0	1	0	0	
P17B49	17	225	0	0	1	0	OUI
		238	1	0	0	0	
		242	0	1	0	0	
		260	0	0	1	0	
P17B496	17	314	1	0	1	0	NON
		316	0	1	0	0	
P17B506	17	264	0	0	1	0	OUI
		269	0	1	0	0	
		278	1	0	0	0	
P17B516	17	228	0	0	1	1	OUI
		250	0	1	0	0	
		256	0	0	1	0	
		264	0	0	1	0	
P17B548	17	229	1	0	1	1	NON
		236	0	1	0	0	
P17B574	17	195	0	0	0	0	NON
		210	0	1	0	0	
		215	1	0	0	0	
		223	0	1	1	0	
		229	1	0	0	0	
P17B581	17	283	0	1	0	0	NON
P17B82	17	0	0	0	0	0	NON
P3A1097	3	212	0	0	0	0	NON
P3A185	3	277	0	0	1	0	OUI
		280	0	1	0	0	
P3A274	3	313	0	0	0	1	NON
		318	0	1	1	0	
		328	1	0	1	0	
P3A401	3	262	1	1	0	0	NON
		267	0	0	1	1	
P3A490	3	280	0	1	0	0	NON
		296	0	1	0	0	
		304	1	0	1	0	
P3A506	3	0	0	0	0	0	NON
P3A518	3	237	0	0	1	0	OUI
		244	1	0	0	0	
		248	0	1	0	0	
		250	1	0	0	0	
P3A581	3	300	0	1	0	0	NON
		307	1	0	1	0	
		317	0	1	0	0	
		323	1	0	0	0	
		324	0	0	1	0	
P3A730	3	0	0	0	0	0	NON
P3A925	3	220	0	0	0	1	NON
		223	1	0	1	0	
		228	0	1	0	0	
P3A1001	3	202	0	1	0	0	NON
P3A1042	3	254	1	1	0	0	NON
		318	0	0	1	0	
P3A129	3	304	0	0	0	1	NON
		306	1	1	0	0	
		312	0	0	1	0	
P3A130	3	289	1	0	1	0	NON
		293	0	1	0	0	
P3A282	3	217	0	0	1	0	NON
		242	1	0	0	0	
		246	1	1	0	0	
		250	0	0	1	0	
P3A287	3	193	0	1	0	0	OUI
		195	1	0	0	0	
		199	0	0	1	0	

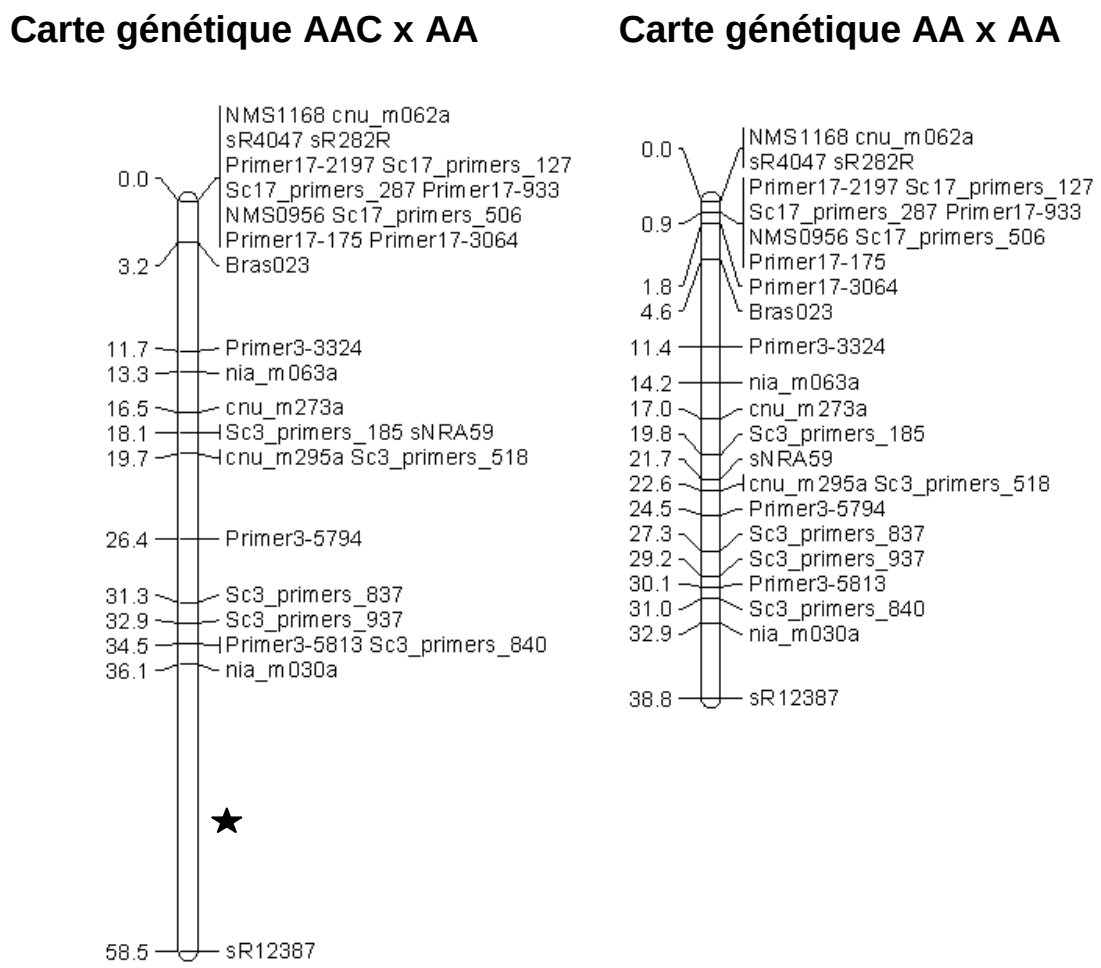
Marqueurs	Scaffolds	Amplification (pb)	C1.3	Chiifu	Darmor	RC	Polymorphe
		200	1	0	1	0	
P3A575	3	221	0	1	0	0	NON
		305	1	1	0	0	
P3A633	3	315	1	0	1	0	NON
P3A712	3	290	0	1	0	0	NON
		298	1	0	0	0	
		306	0	0	1	0	
P3A837	3	310	?	1	0	0	OUI
		214	1	0	0	0	
		259	0	1	0	0	
P3A840	3	261	1	0	0	0	OUI
		269	1	0	1	0	
P3A874	3	280	0	1	0	0	NON
		187	0	0	1	0	
		188	1	0	0	0	
P3A916	3	192	0	1	0	0	OUI
		221	0	1	0	1	
		223	1	0	0	0	
P3A937	3	225	1	0	1	0	OUI
		313	0	1	0	0	
		316	1	0	1	0	
P3A945	3	320	1	0	0	0	NON
		302	1	1	0	0	
P3A954	3	308	0	0	1	0	NON
		245	1	0	?	0	
		248	0	1	?	0	
P3A987	3	249	1	0	?	0	NON
		308	0	0	1	0	
		310	0	1	0	0	
m3324	3	326	1	0	0	0	OUI
		270	1	0	0	0	
m5794	3	281	0	1	0	0	OUI
		268	0	1	0	0	
		313	1	0	0	0	
m3571	3	323	1	0	0	0	OUI
		259	0	0	1	1	
		287	1	0	1	0	
m5813	3	306	0	1	0	0	OUI

Annexe IV : Position, ségrégation et amorces des différents SSR polymorphes

Marqueurs testés	Séquence des amorces		Position (pb)		χ^2		
	Forward	Reverse	Début	Fin	0C	1C9	9C
NMS1168	TGAGCACATAAAGTGGTTCAGC	CTGAACAAGGACCCGAAGAA	11995	12628	0.725	2.528	3.322
cnu_m062a	ATCGGCGCTGGTTATGTCA	CTAGGCTGCCCTTTCCGATT	658916	658713	0.474	1.899	3.048
sR4047	AGAGAGGGTGTTGTTCCCAG	ATTCTTTGAATGGCCAGCAC	660840	660571	0.159	1.636	2.036
sR0282R	AGGAAGCCCAACAGGACTTT	AATTCGATTCTCCATCGTGC	1539077	1539096	0.474	0.099	0.275
Primer17-2197	AAGGTTATGGCTCTGTCAAA	ATGTGCAATTCAGACAAACA	5253232	5252841	0.376	1.6	0.101
Sc17_primers_49	TCCTATTGGACACGGAGTTT	TGAGCTGGGTAAGTTTAGGG	5270921	5271147	0.301	3.322	0.301
Sc17_primers_127	TATATACGACGCGGCTAGTG	CTGCTCGAGGATCCATTTAT	5804323	5804604	0.092	1.6	0.178
Sc17_primers_287	TCTAATCGAAAACGGAAAGG	ATCGAAAAGAACTGACGGAG	6087451	6087652	0.092	2.473	0.178
Sc17_primers_204	TCTGTGGTGCATAGACTCGT	TATGTAGCCAAATGGCAGAA	6088379	6088565	0.093	2.227	0.178
Primer17-933	GGAAATTGCACTCACTAACC	GGATGAAGCTTGATTTGTTT	6246971	6246786	0.006	3.682	0.78
NMS0956	AACCATCGCGGTAATTAAACC	ACGACTCCAGCTCAGCATCT	6470981	6471137	0.053	1.674	0.178
Sc17_primers_506	CTGTGATCAAACTGTCCCAA	GGCCATTAGATAGGGCATTA	6739338	6739588	0.052	2.844	0.4
Primer17-303	TGGAAGCACATACTGTGAGA	TTAAAGCACCGAGAAAAGAG	6752779	6752392	0.006	3.682	0.538
Primer17-175	TCTTGTGTGTTTTCAAATCG	ATACAGCTCCACCATGAAAC	6861182	6860784	0.153	1.899	0.711
Primer17-3064	GCTACTTCGAAGGAAAGGAT	TCTTCTTCCTCATTTTTCCA	7473059	7472671	0.006	2.844	0.711
Bras023	CCATTGCAAATCCCTTTACTC	AGATGGAATGCGATCAAAGA	10910411	10910215	0.023	5.813	0.309
Primer3-3324	TTATATTGGACAACAAATAAGTTCA	AATATGAAACGGACTAATGTTAGAG	14732105	14732399	1.21	1.391	0.178
nia_m063a	GAAGAAACTCGGTGGGGAGT	AAAGAGTTCCGAAAATGGGC	15288691	15288962	0.699	0.538	0.538
cnu_m273a	ATAAGGGCATCGCCTCAACA	TGCACGCATCCACATAAACA	15823972	15823714	0.581	0.103	0.563
Sc3_primers_185	TTTCACCTGAAAACACGCTA	TTTTGCCTGAAGGGTAAAAG	15995621	15995884	0.217	0.047	0.78
sNRA59	CAGATTCGGATTTGGGAAGA	GGCGGAAGAATCAAAGGAGT	16337165	16337287	0.093	0.044	0.045
cnu_m295a	GCTGCCTAATAGGGTGCTTG	AGAGCGCATTCAAGTCTGGT	16445601	16445795	0	0.045	0.045
Sc3_primers_518	CCGGAGATCGGTTAAGTAAA	TCACCACCTCTCATAACCCT	16585365	16585594	0.023	0	0.101
Primer3-5794	GATCAGTAGTAAGCGAGAGAAGAT	CAGCACAGTCGATATAAGATAAAT	16848081	16848330	0	0.045	0.099
Primer3-3571	CAACAAGTTATACAATTAAAATCCA	ATGAGGAGAACGTGTATATAAGACT	16906639	16906859	0.523	11.951	1.756
Sc3_primers_837	GGAACTAGGAAGCATGGAAA	TCAGTATCTGAACTGCTCGTG	17122717	17123007	0.381	0	0.178
Sc3_primers_916	TTAGAATAAACCCCCACCAA	GCATCTGGCTCATTTAGGTT	17130225	17130398	0.094	0.011	0.275
Sc3_primers_937	TCTTGCACTTTCTCACCAAA	TGTGAGACTCAGCCGAATAA	17576926	17577128	1.316	0.101	0.538
Primer3-5813	AAAGTAGAATCGAGGAACTAAGAAT	GGGAAATAAAATGTATTATTAAGGA	17739483	17739767	0.837	0.551	0.538
Sc3_primers_840	TTTTGTCTTGGCTGACTTTG	AAAAGTTCTCAAGTTTGCGG	17964158	17964397	0.977	0.178	0.275
nia_m030a	GATCATAAGCCGAAAAAGGTTG	TGCTCTCCTCAAGTGAATCAAA	18890204	18889754	1.316	4.846	1.22
sR12387	GGGTCTGGGTTTTTCTGTGA	GATTGGGCCGTGTAATATCG	20701201	20701497	0.595	6.4	0.409

Remarque : En rouge, les SSR éliminés à cause de leur trop forte proximité avec d'autres SSR

Annexe V : Cartographie des populations 0C issues de AACxAA ou de AAxAA



Remarque : L'étoile indique le seul intervalle significativement différent observé entre ces 2 populations pour le groupe de liaison A07

Annexe VI : Analyses statistiques

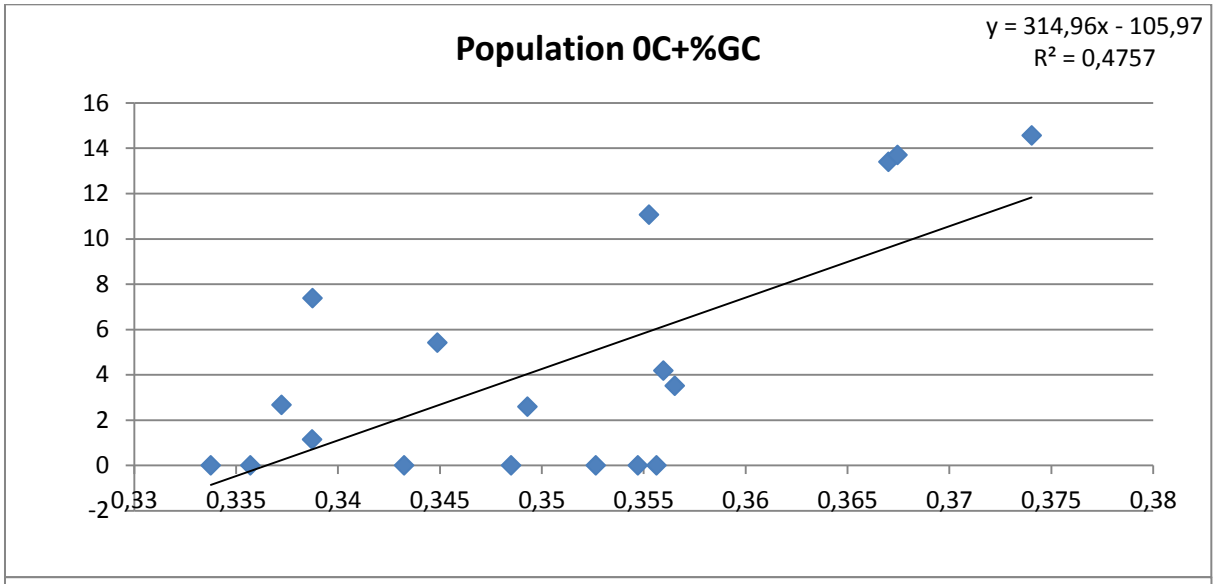
Intervalles	0C		1C9		%attendu		somme(obs-att)²/att	ddl=1	p-value
	%CO	%absCO	%CO	%absCO	%CO	%abs CO			
A-B	0,00	100,00	4,40	95,60	1,51	98,49	7,14	>3,84	0,0076
B-C	0,00	100,00	0,00	100,00	0,00	100,00	0,00	<3,84	1
C-D	0,00	100,00	16,48	83,52	5,66	94,34	27,94	>3,84	1E-07
D-E	0,57	99,43	13,19	86,81	4,91	95,09	18,72	>3,84	2E-05
E-F	0,00	100,00	0,00	100,00	0,00	100,00	0,00	<3,84	1
F-G	0,00	100,00	1,10	98,90	0,38	99,62	1,76	<3,84	0,1842
G-H	0,00	100,00	0,00	100,00	0,00	100,00	0,00	<3,84	1
H-I	0,00	100,00	1,10	98,90	0,38	99,62	1,76	<3,84	0,1842
I-J	0,00	100,00	1,10	98,90	0,38	99,62	1,76	<3,84	0,1842
J-K	0,19	99,81	0,00	100,00	0,13	99,87	0,16	<3,84	0,6888
K-L	0,77	99,23	3,30	96,70	1,64	98,36	2,19	<3,84	0,1393
L-M	2,49	97,51	5,49	94,51	3,52	96,48	1,46	<3,84	0,2272
M-N	6,90	93,10	12,64	87,36	8,87	91,13	2,24	<3,84	0,1346
N-O	2,30	97,70	2,75	97,25	2,45	97,55	0,05	<3,84	0,8299
O-P	2,87	97,13	8,24	91,76	4,72	95,28	3,52	<3,84	0,0606
P-Q	2,30	97,70	1,65	98,35	2,08	97,92	0,11	<3,84	0,7353
Q-R	1,15	98,85	2,20	97,80	1,51	98,49	0,41	<3,84	0,524
R-S	1,15	98,85	1,10	98,90	1,13	98,87	0,00	<3,84	0,9718
S-T	0,00	100,00	3,30	96,70	1,13	98,87	5,33	>3,84	0,0209
T-U	3,45	96,55	3,30	96,70	3,40	96,60	0,00	<3,84	0,9506
U-V	3,45	96,55	2,20	97,80	3,02	96,98	0,29	<3,84	0,5881
V-W	1,72	98,28	1,10	98,90	1,51	98,49	0,14	<3,84	0,704
W-X	1,15	98,85	1,10	98,90	1,13	98,87	0,00	<3,84	0,9718
X-Y	0,57	99,43	2,20	97,80	1,13	98,87	1,29	<3,84	0,2556
Y-Z	1,72	98,28	10,99	89,01	4,91	95,09	10,10	>3,84	0,0015
Z-AA	9,20	90,80	4,40	95,60	7,55	92,45	1,81	<3,84	0,1782

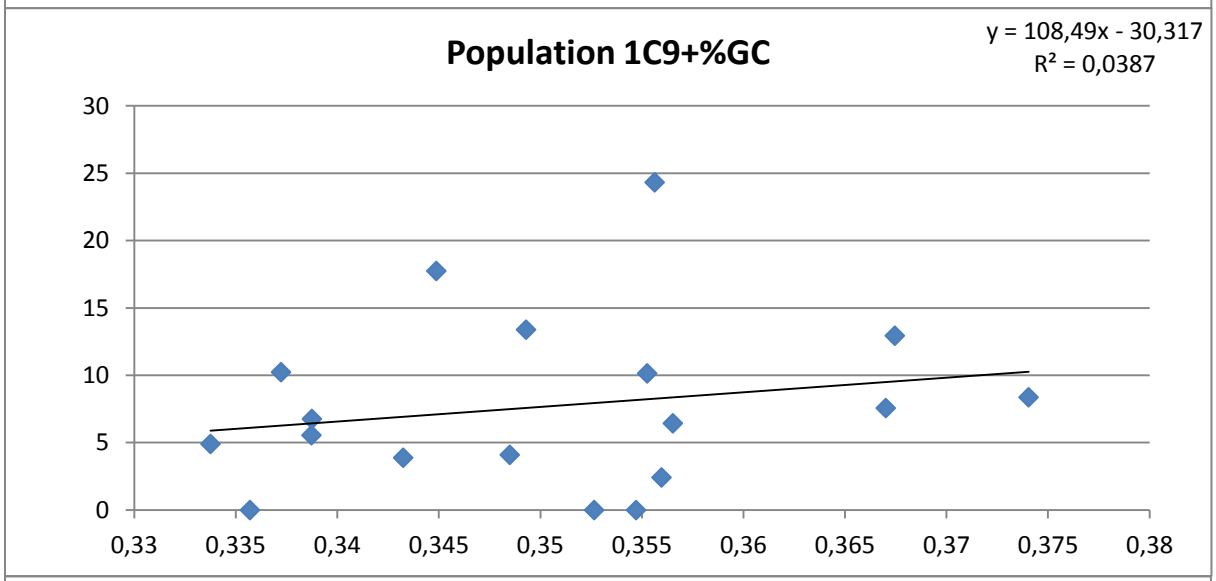
Intervalles	0C		9C		%attendu		somme(obs-att)²/att	ddl=1	p-value
	%CO	%absCO	%CO	%absCO	%CO	%abs CO			
A-B	0,00	100,00	9,16	90,84	3,14	96,86	15,12	>3,84	0,0001
B-C	0,00	100,00	1,47	98,53	0,50	99,50	2,35	<3,84	0,1249
C-D	0,00	100,00	11,36	88,64	3,90	96,10	18,89	>3,84	1E-05
D-E	0,57	99,43	19,78	80,22	7,17	92,83	30,43	>3,84	3E-08
E-F	0,00	100,00	4,40	95,60	1,51	98,49	7,14	>3,84	0,0076
F-G	0,00	100,00	0,00	100,00	0,00	100,00	0,00	<3,84	1
G-H	0,00	100,00	3,30	96,70	1,13	98,87	5,33	>3,84	0,0209
H-I	0,00	100,00	1,10	98,90	0,38	99,62	1,76	<3,84	0,1842
I-J	0,00	100,00	4,40	95,60	1,51	98,49	7,14	>3,84	0,0076
J-K	0,19	99,81	0,00	100,00	0,13	99,87	0,16	<3,84	0,6888
K-L	0,77	99,23	2,20	97,80	1,26	98,74	0,91	<3,84	0,3412
L-M	2,49	97,51	16,48	83,52	7,30	92,70	15,90	>3,84	7E-05
M-N	6,90	93,10	34,07	65,93	16,23	83,77	29,82	>3,84	5E-08
N-O	2,30	97,70	3,30	96,70	2,64	97,36	0,21	<3,84	0,6448
O-P	2,87	97,13	7,69	92,31	4,53	95,47	2,95	<3,84	0,0859
P-Q	2,30	97,70	3,30	96,70	2,64	97,36	0,21	<3,84	0,6448
Q-R	1,15	98,85	6,96	93,04	3,14	96,86	6,09	>3,84	0,0136
R-S	1,15	98,85	0,37	99,63	0,88	99,12	0,39	<3,84	0,5345
S-T	0,00	100,00	13,00	87,00	4,47	95,53	21,76	>3,84	3E-06
T-U	3,45	96,55	8,24	91,76	5,09	94,91	2,61	<3,84	0,1062
U-V	3,45	96,55	1,10	98,90	2,64	97,36	1,18	<3,84	0,2777
V-W	1,72	98,28	5,49	94,51	3,02	96,98	2,67	<3,84	0,1025
W-X	1,15	98,85	2,20	97,80	1,51	98,49	0,41	<3,84	0,524
X-Y	0,57	99,43	1,10	98,90	0,75	99,25	0,20	<3,84	0,6536
Y-Z	1,72	98,28	14,84	85,16	6,23	93,77	16,16	>3,84	6E-05
Z-AA	9,20	90,80	6,04	93,96	8,11	91,89	0,73	<3,84	0,3924

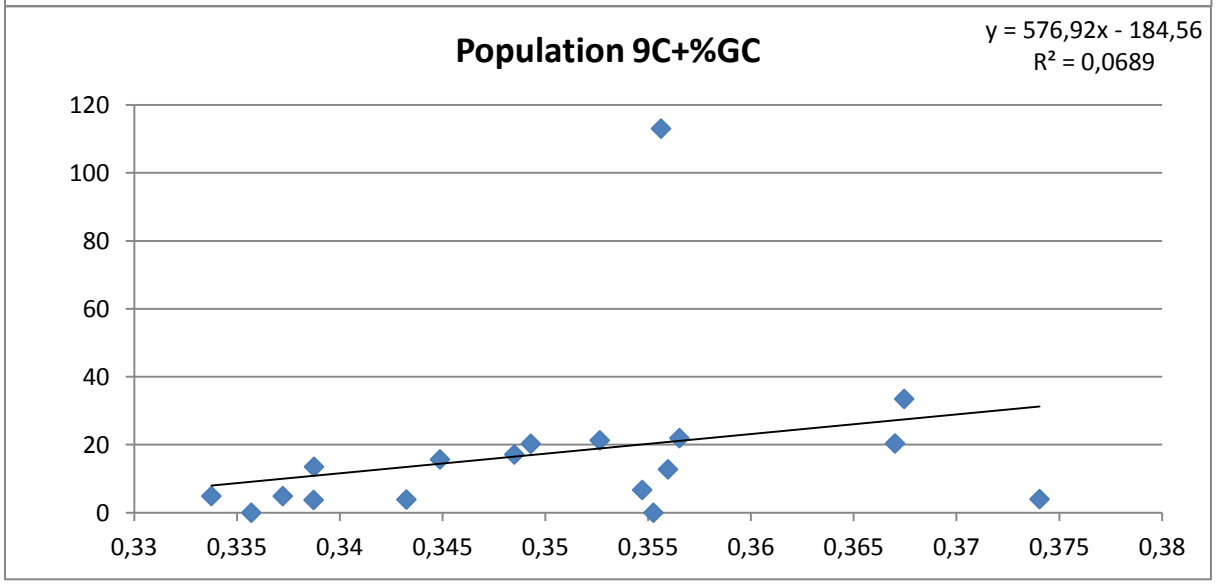
Intervalles	1C9		9C		%attendu		somme(obs-att)²/att	ddl=1	p-value
	%CO	%absCO	%CO	%absCO	%CO	%abs CO			
A-B	4,40	95,60	9,16	90,84	6,78	93,22	1,79	<3,84	0,1804
B-C	0,00	100,00	1,47	98,53	0,73	99,27	1,48	<3,84	0,2244
C-D	16,48	83,52	11,36	88,64	13,92	86,08	1,10	<3,84	0,2948
D-E	13,19	86,81	19,78	80,22	16,48	83,52	1,58	<3,84	0,2089
E-F	0,00	100,00	4,40	95,60	2,20	97,80	4,49	>3,84	0,034
F-G	1,10	98,90	0,00	100,00	0,55	99,45	1,10	<3,84	0,2932
G-H	0,00	100,00	3,30	96,70	1,65	98,35	3,35	<3,84	0,0671
H-I	1,10	98,90	1,10	98,90	1,10	98,90	0,00	<3,84	1
I-J	1,10	98,90	4,40	95,60	2,75	97,25	2,03	<3,84	0,1538
J-K	0,00	100,00	0,00	100,00	0,00	100,00	0,00	<3,84	1
K-L	3,30	96,70	2,20	97,80	2,75	97,25	0,23	<3,84	0,6345
L-M	5,49	94,51	16,48	83,52	10,99	89,01	6,17	>3,84	0,013
M-N	12,64	87,36	34,07	65,93	23,35	76,65	12,83	>3,84	0,0003
N-O	2,75	97,25	3,30	96,70	3,02	96,98	0,05	<3,84	0,8205
O-P	8,24	91,76	7,69	92,31	7,97	92,03	0,02	<3,84	0,8859
P-Q	1,65	98,35	3,30	96,70	2,47	97,53	0,56	<3,84	0,4529
Q-R	2,20	97,80	6,96	93,04	4,58	95,42	2,60	<3,84	0,1072
R-S	1,10	98,90	0,37	99,63	0,73	99,27	0,37	<3,84	0,5435
S-T	3,30	96,70	13,00	87,00	8,15	91,85	6,29	>3,84	0,0121
T-U	3,30	96,70	8,24	91,76	5,77	94,23	2,25	<3,84	0,1337
U-V	2,20	97,80	1,10	98,90	1,65	98,35	0,37	<3,84	0,5417
V-W	1,10	98,90	5,49	94,51	3,30	96,70	3,03	<3,84	0,0817
W-X	1,10	98,90	2,20	97,80	1,65	98,35	0,37	<3,84	0,5417
X-Y	2,20	97,80	1,10	98,90	1,65	98,35	0,37	<3,84	0,5417
Y-Z	10,99	89,01	14,84	85,16	12,91	87,09	0,66	<3,84	0,4174
Z-AA	4,40	95,60	6,04	93,96	5,22	94,78	0,27	<3,84	0,6003

Annexe VII : Corrélations

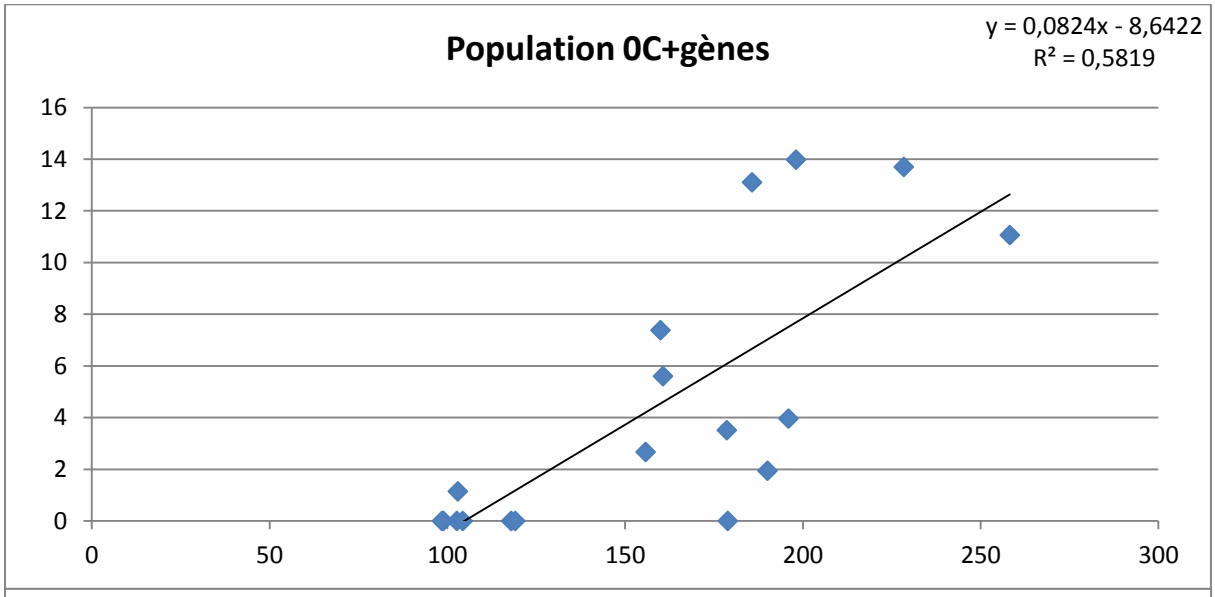
– Fréquence de recombinaison & Teneur en GC

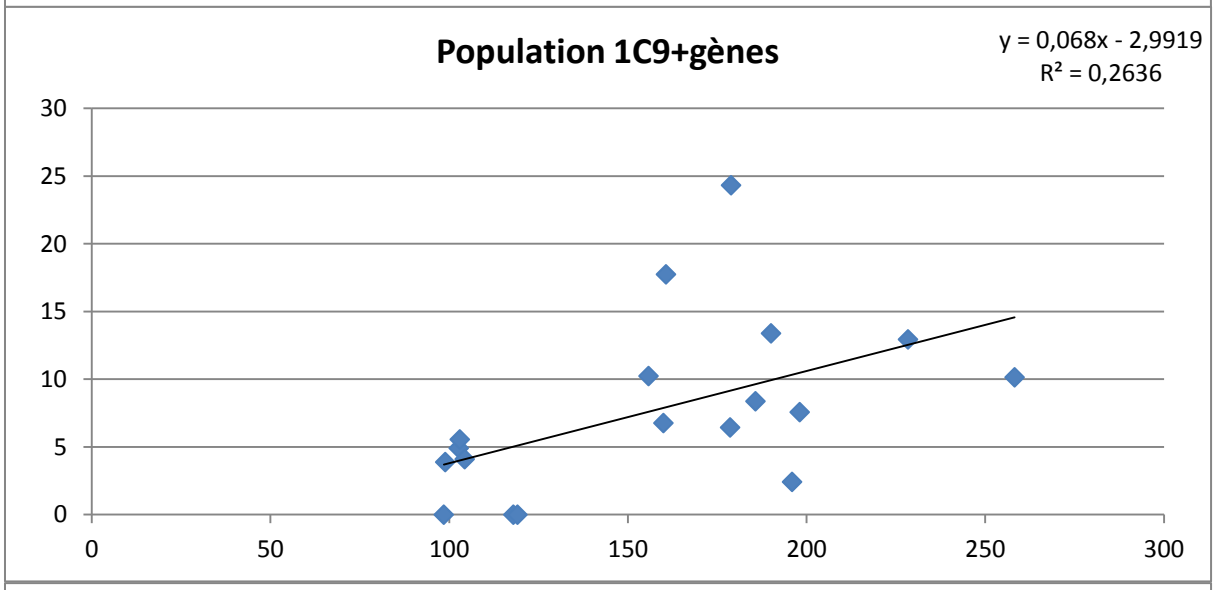


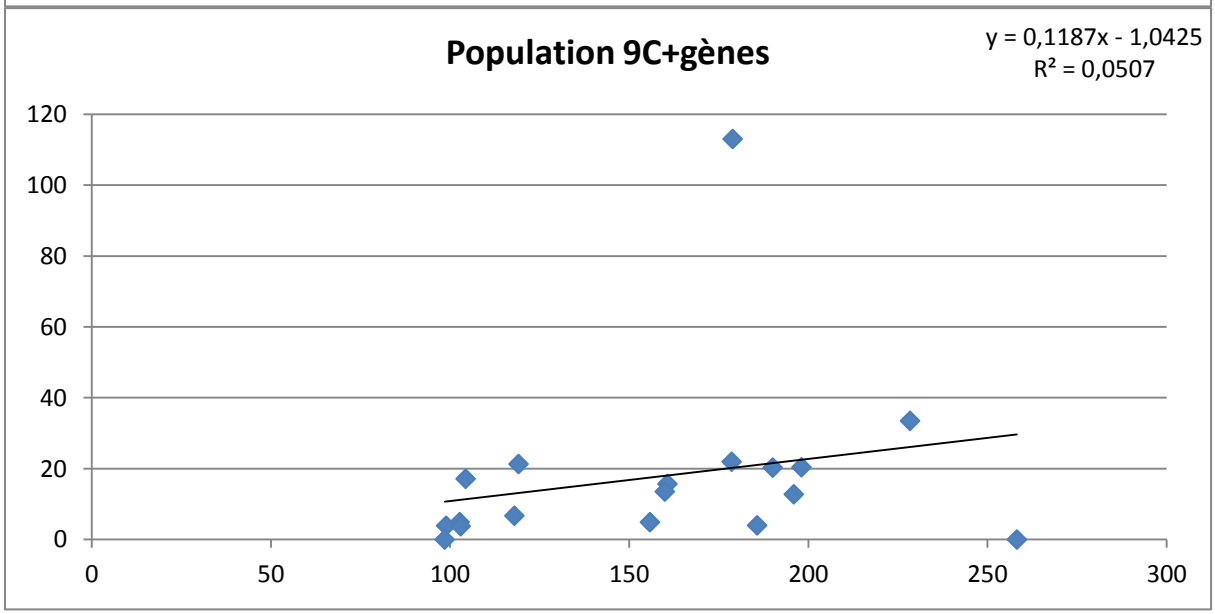




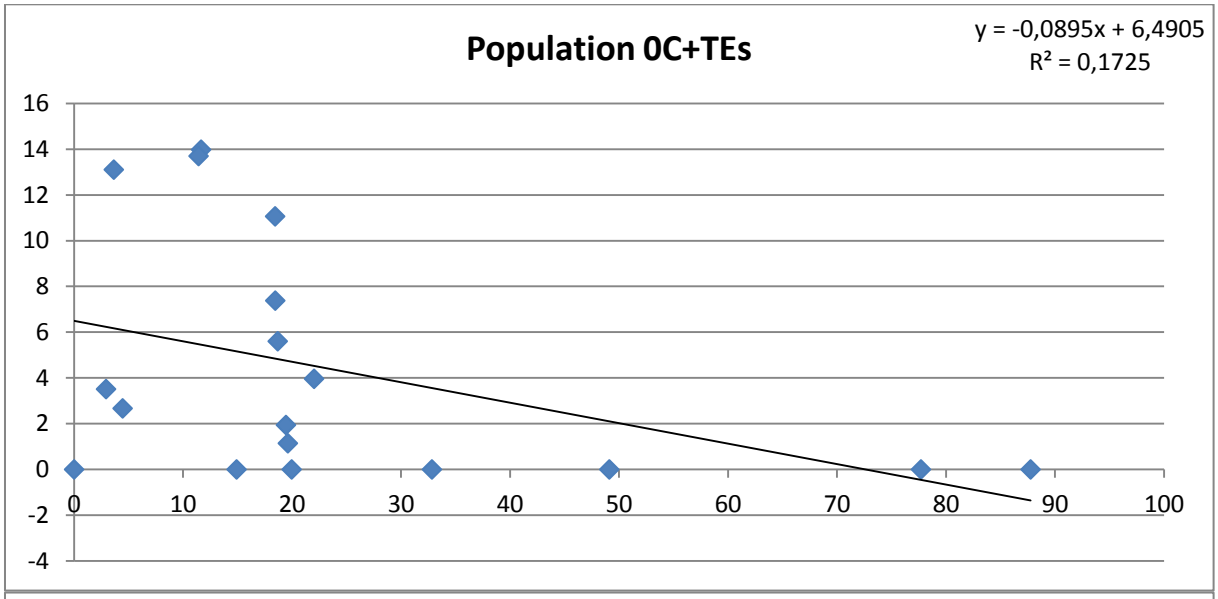
– Fréquence de recombinaison & Teneur en gènes

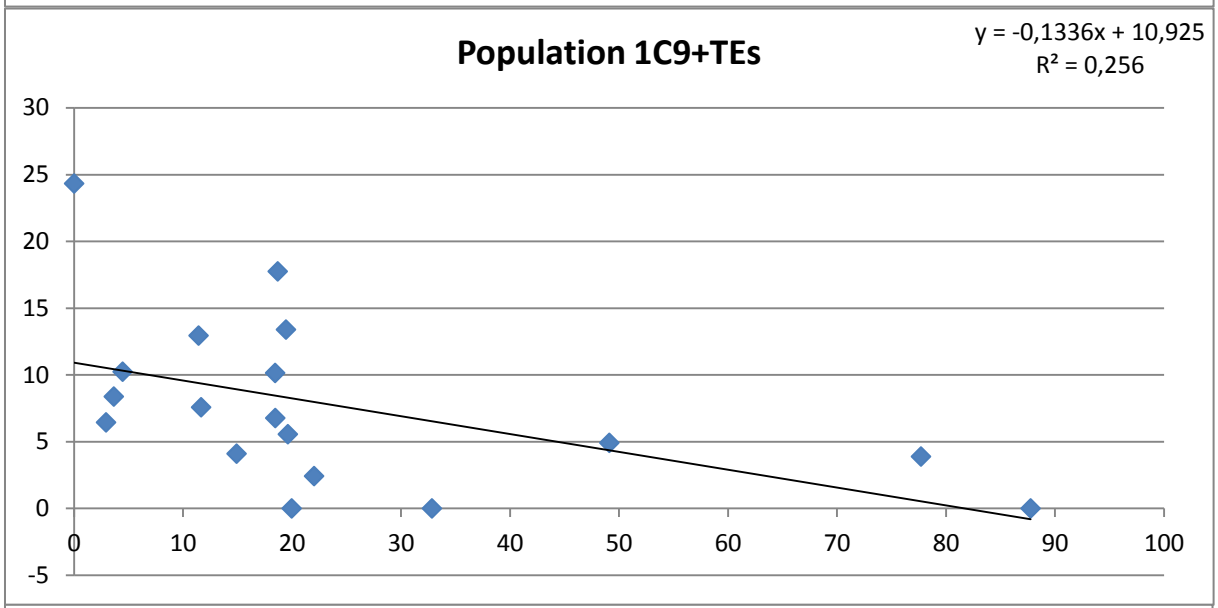


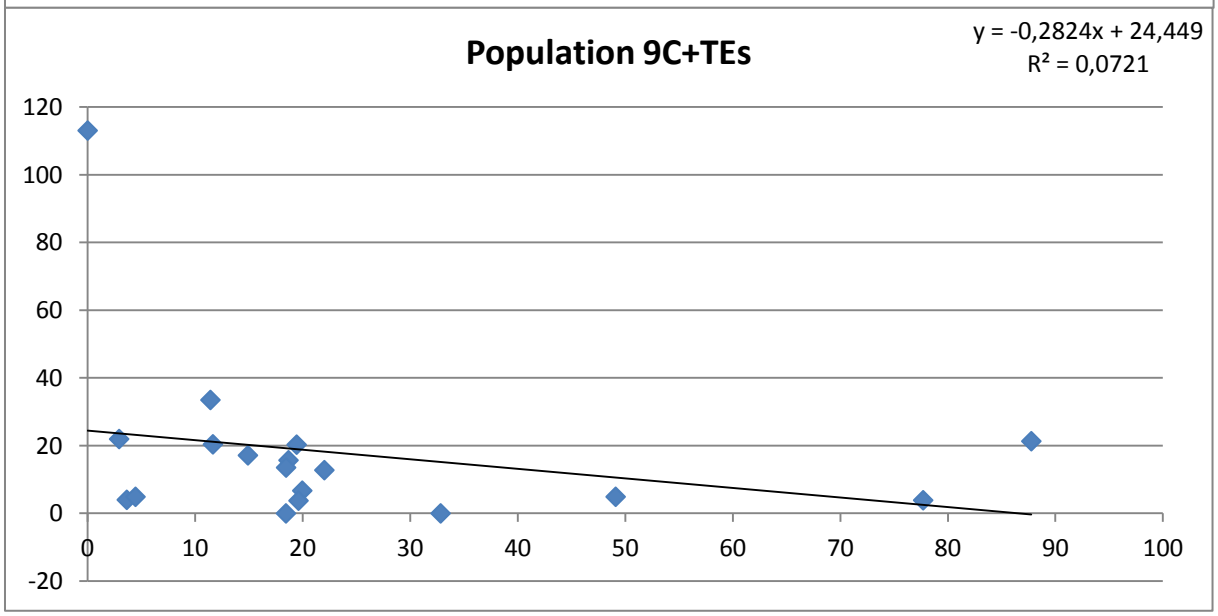




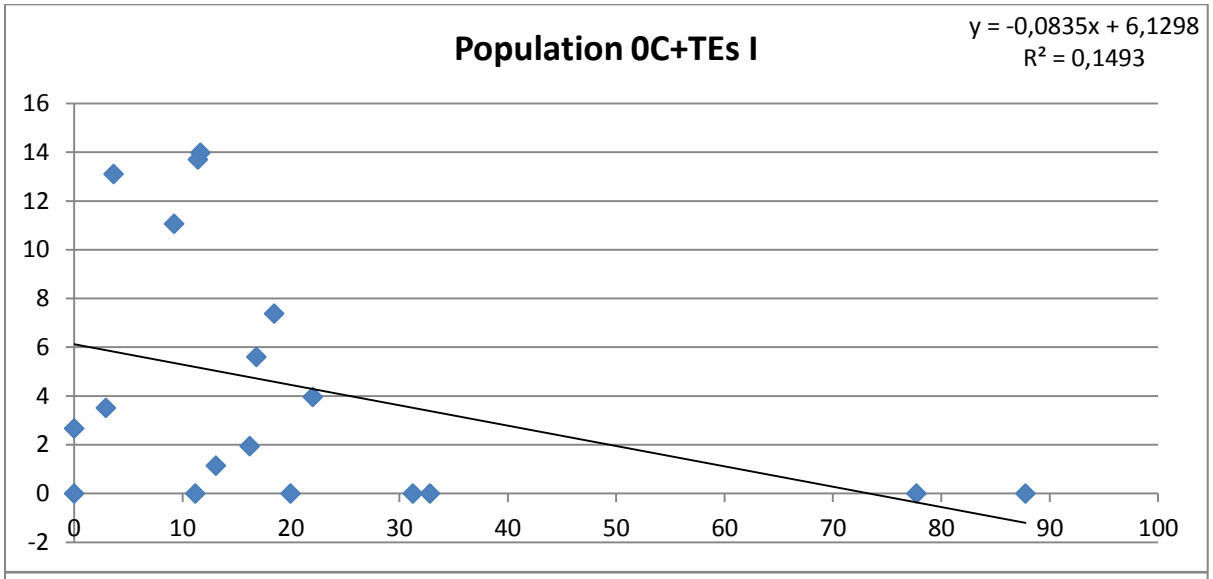
– Fréquence de recombinaison & Teneur en TEs Classes I et II

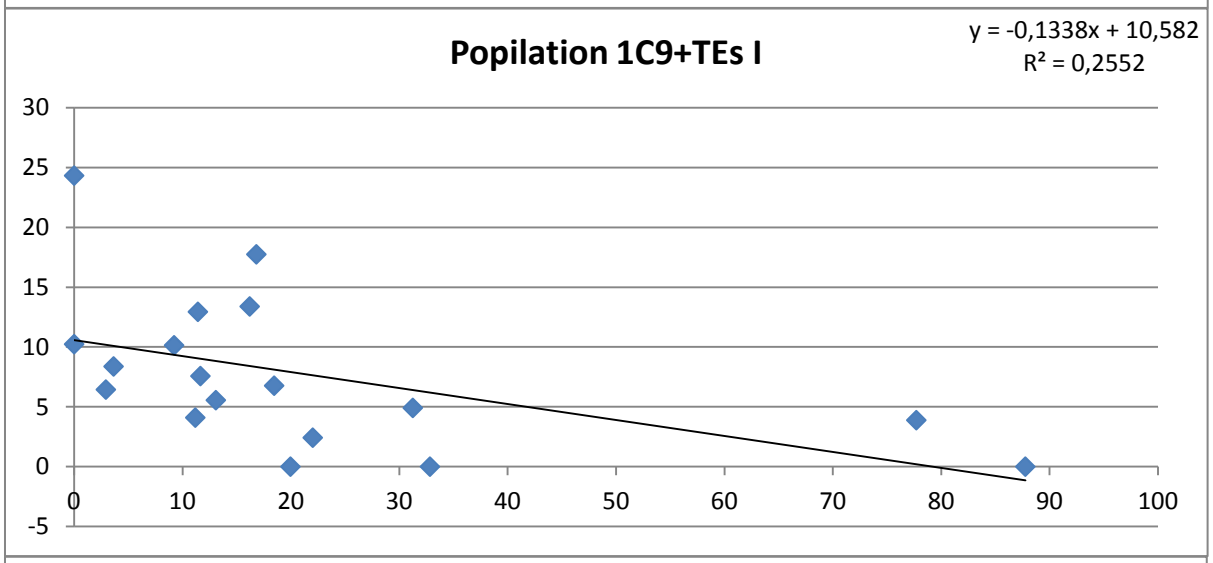


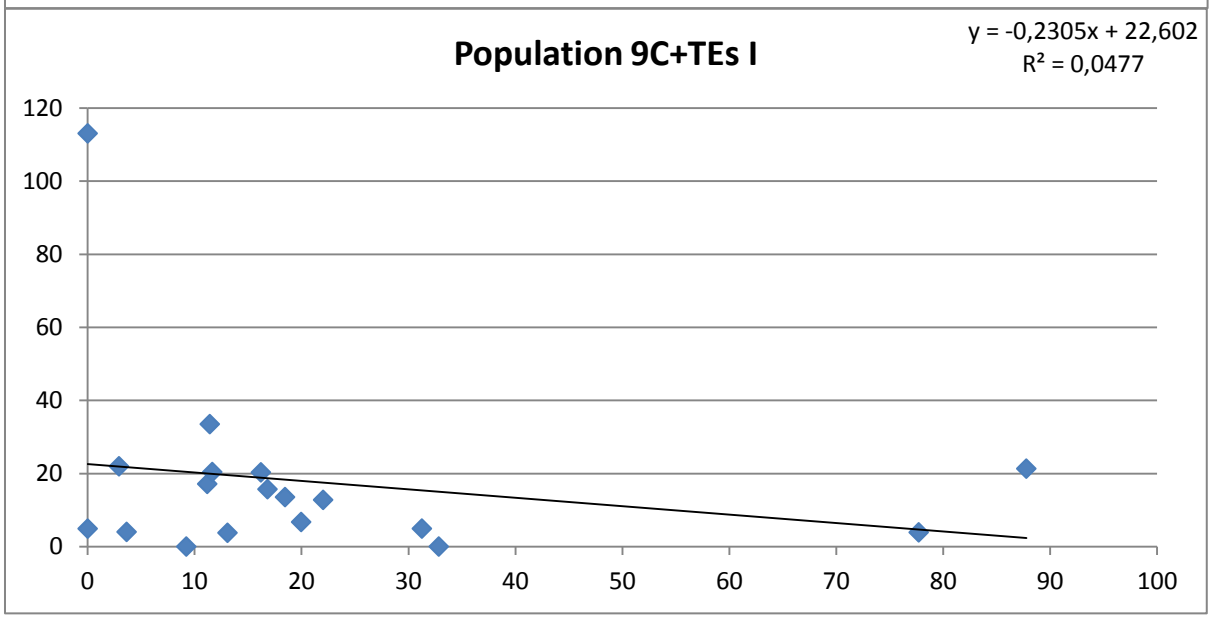




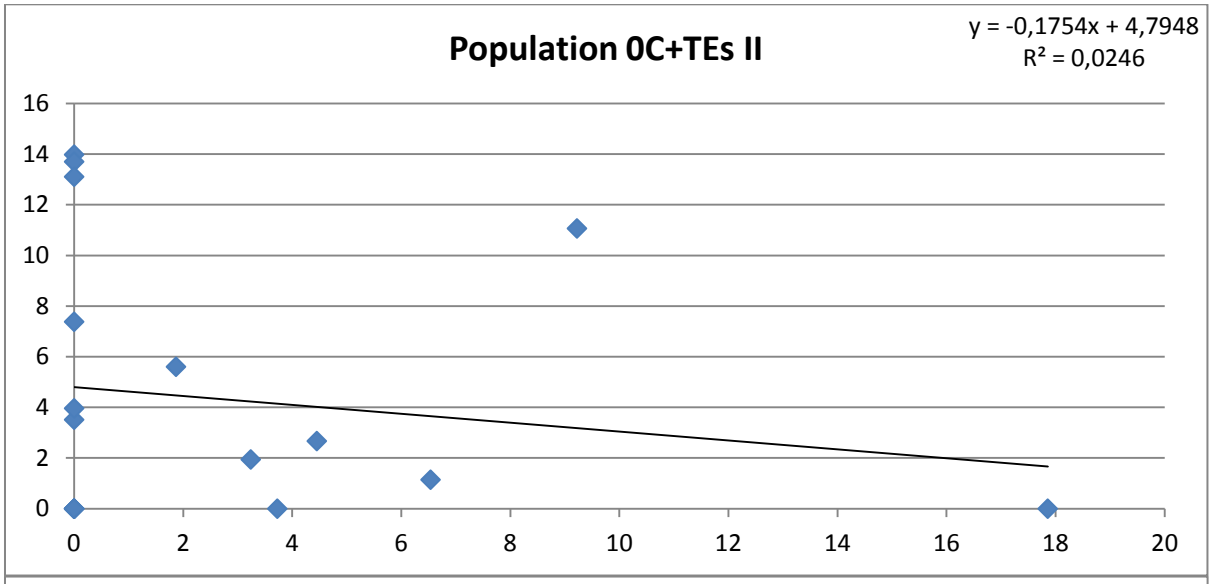
– Fréquence de recombinaison & Teneur en TEs Classes I

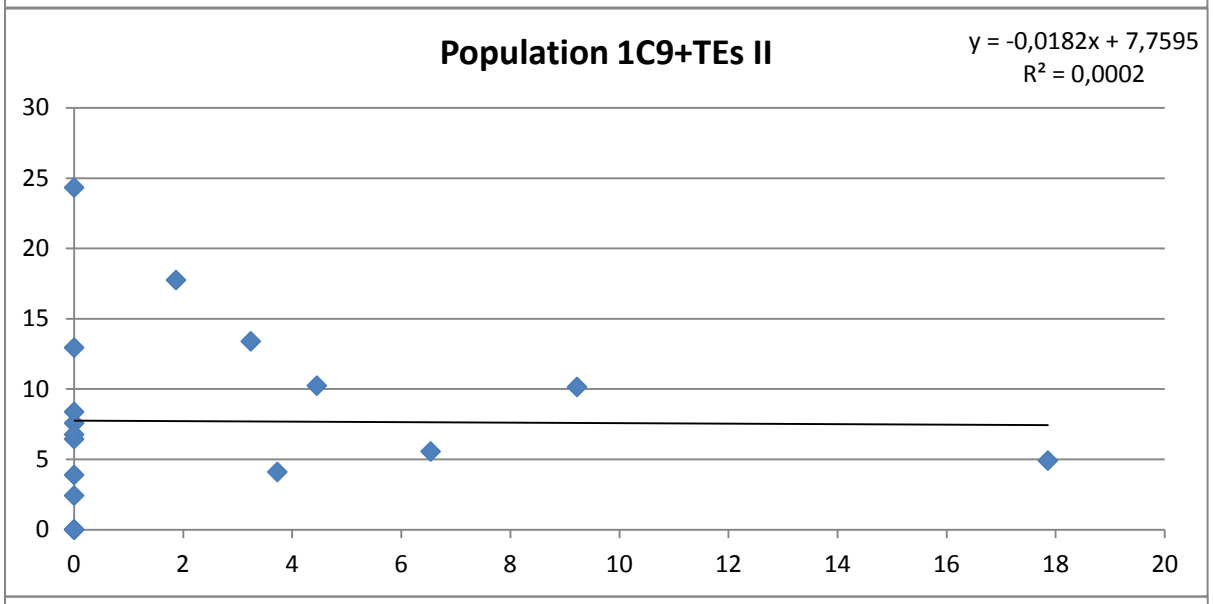


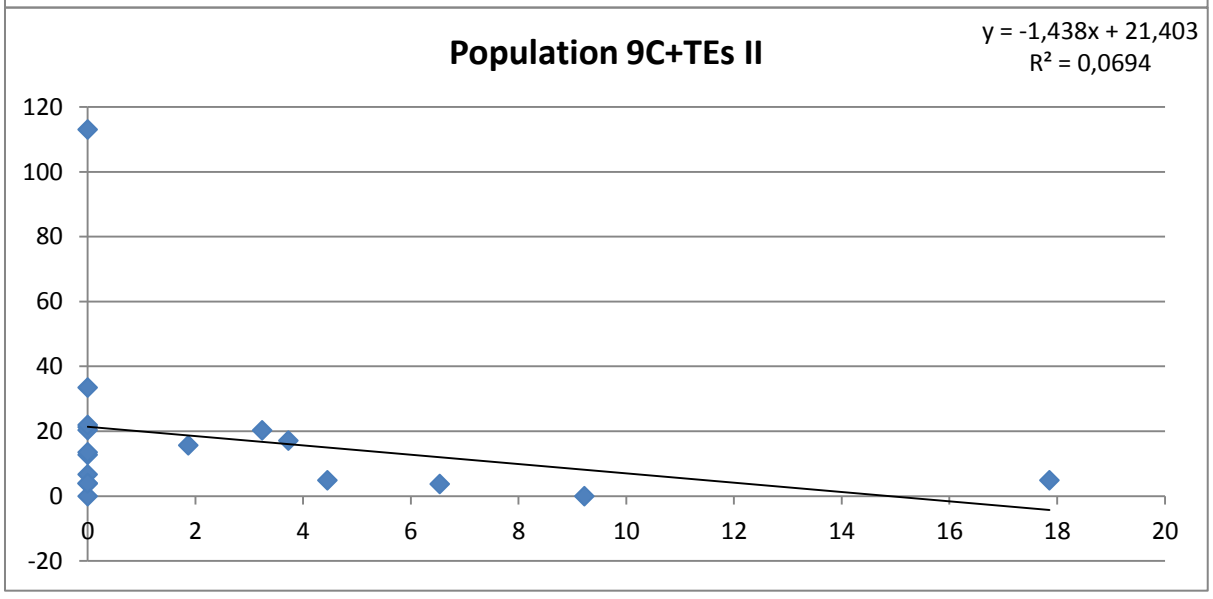




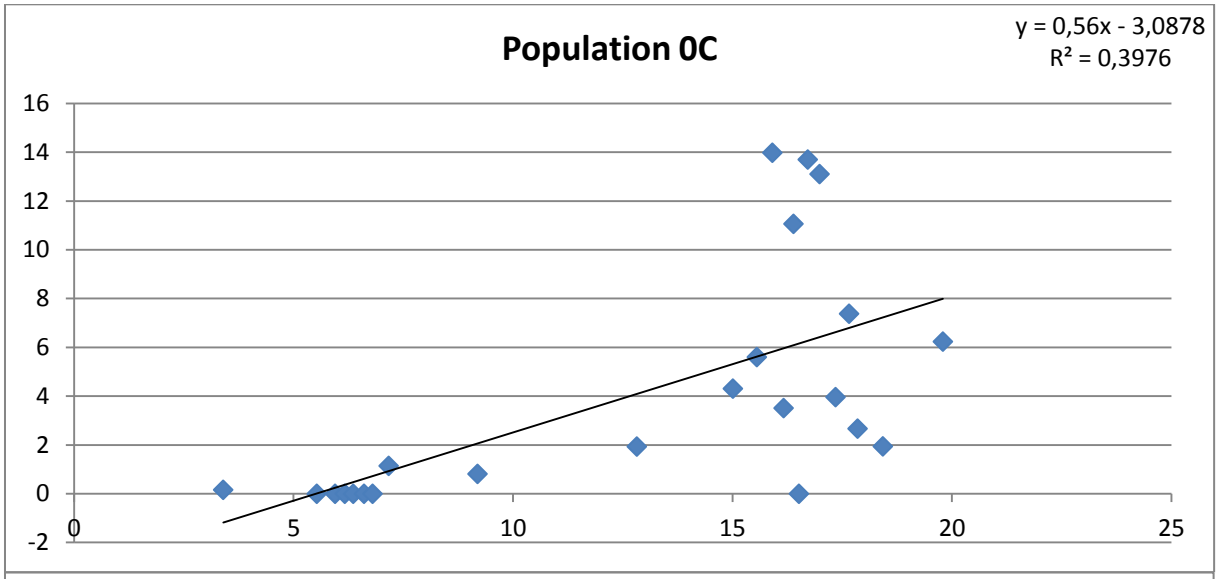
– Fréquence de recombinaison & Teneur en TEs Classes II

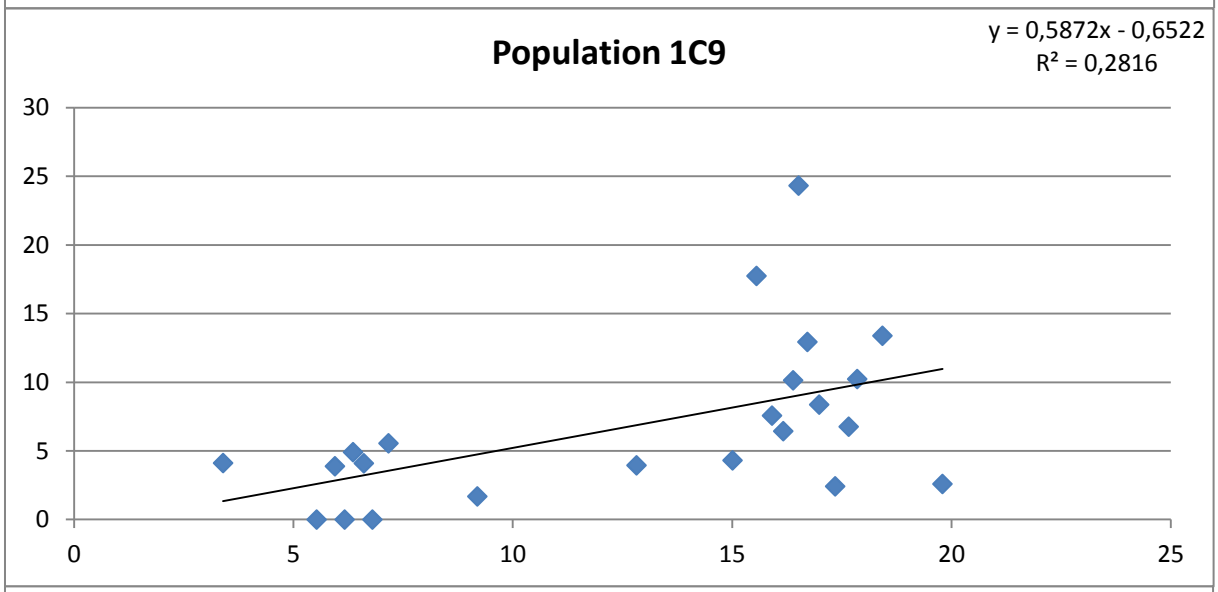


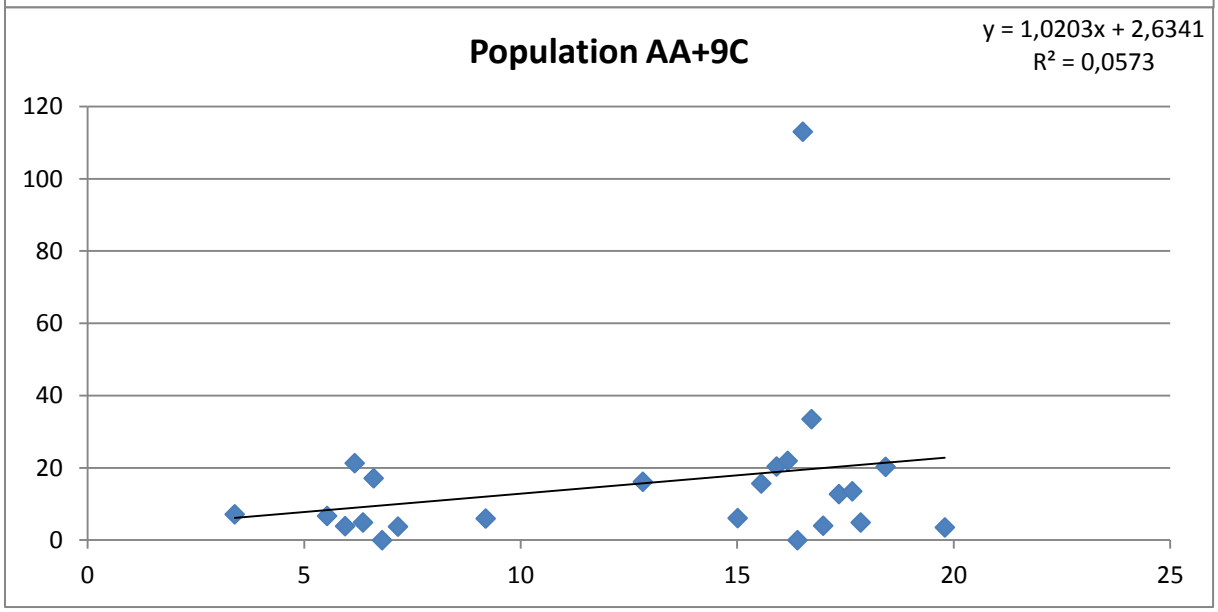




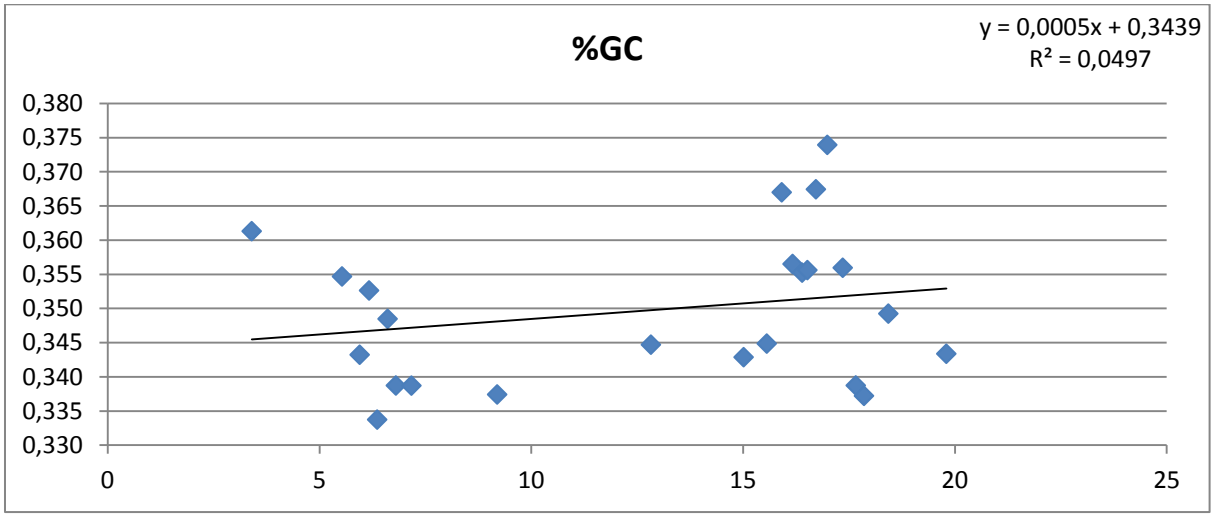
– Fréquence de recombinaison & Distribution



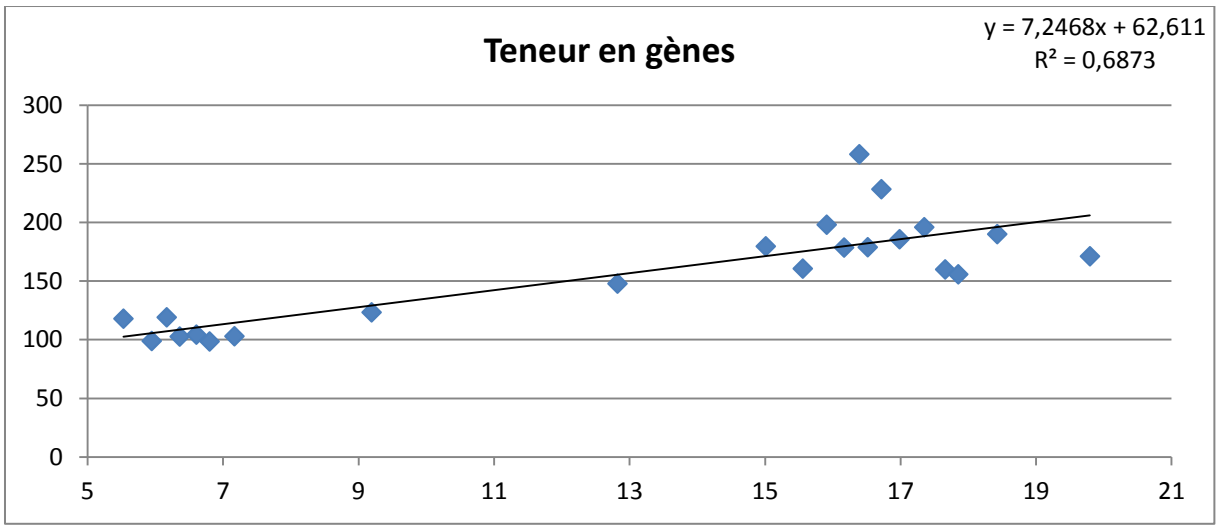




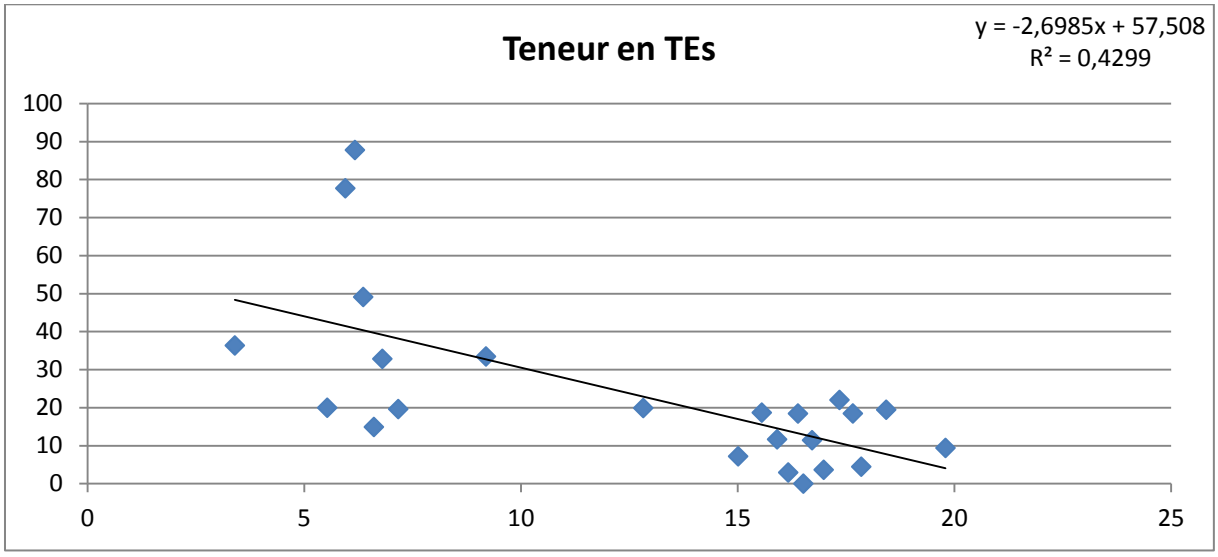
– Teneur en GC & Distribution



– Teneur en gènes & Distribution



– Teneur en TEs & Distribution



Annexe VIII : Résultats de l'étude de la distribution des COs le long des chromosomes A02 et A04

Tableau 1. Positions physiques des marqueurs utilisés par Suay *et al.* (2013) sur le génome de *B. rapa* v1.5

Marker	Hit	Debut (pb)	Fin (pb)
sR12095	A02	214757	214434
cnu_m616a	A02	666883	667153
NMS0515	A02	1660765	1660675
nia_m121a	A02	2606431	2606727
nia_m105a	A02	3251787	3251586
Bras037	A02	5307524	5311625
sR6293	A02	6150306	6150429
cnu_m46a	A02	9249371	9249555
Na12H09	A02	10630998	10630871
NMS0920	A02	13660057	13659908
cnu_m254a	A02	23467286	23467038
NMS0730	A02	25861334	25861538
cnu_m570a	A04	91041	90859
cnu_m225a	A04	3824921	3824806
cnu_m439a	A04	4392312	4392012
CB10347	A04	4611846	4611663
NMS1096	A04	9015854	9016059
sN2025	A04	10752015	10751362
BN54345	A04	10696430	10696300
BN51144X	A04	11341231	11340885
cnu_m246a	A04	12551342	12551649
PFM512	A04	17426886	17427600

Tableau 2. Position des marqueurs des Blocs W et E de Schranz sur le chromosome A02 de *B. rapa*

Blocs de Schranz	Marqueur *B. rapa* début	Marqueur *B. rapa* fin	Position début (Pb)	Position fin (Pb)
W	Bra020251	Bra020463	4 844 256	5 892 850
W	Bra035650	Bra035540	5 917 829	6 614 006
W	Bra022745	Bra022504	6 630 365	8 370 344
E	Bra022498	Bra022470	8 402 633	8 589 701
E	Bra039733	Bra039762	8 609 994	8 830 203
E	Bra040632	Bra040624	8 879 243	8 934 977
E	Bra034043	Bra033923	8 950 841	9 920 637
E	Bra038367	Bra038305	9 950 557	10 369 056
E	Bra007864	Bra008502	10 384 450	15 011 253

Tableau 3. Position des marqueurs des Blocs T et S de Schranz sur le chromosome A04 de *B. rapa*

Blocs de Schranz	Marqueur *B. rapa* début	Marqueur *B. rapa* fin	Position début (Pb)	Position fin (Pb)
T	Bra033482	Bra033487	4 873 477	4 918 407
T	Bra032782	Bra032672	4 962 259	6 073 518
S	Bra028101	Bra028250	6 101 687	7 460 841
S	Bra025665	Bra025427	7 511 591	9 188 107
S	Bra030361	Bra030327	9 212 795	9 465 650

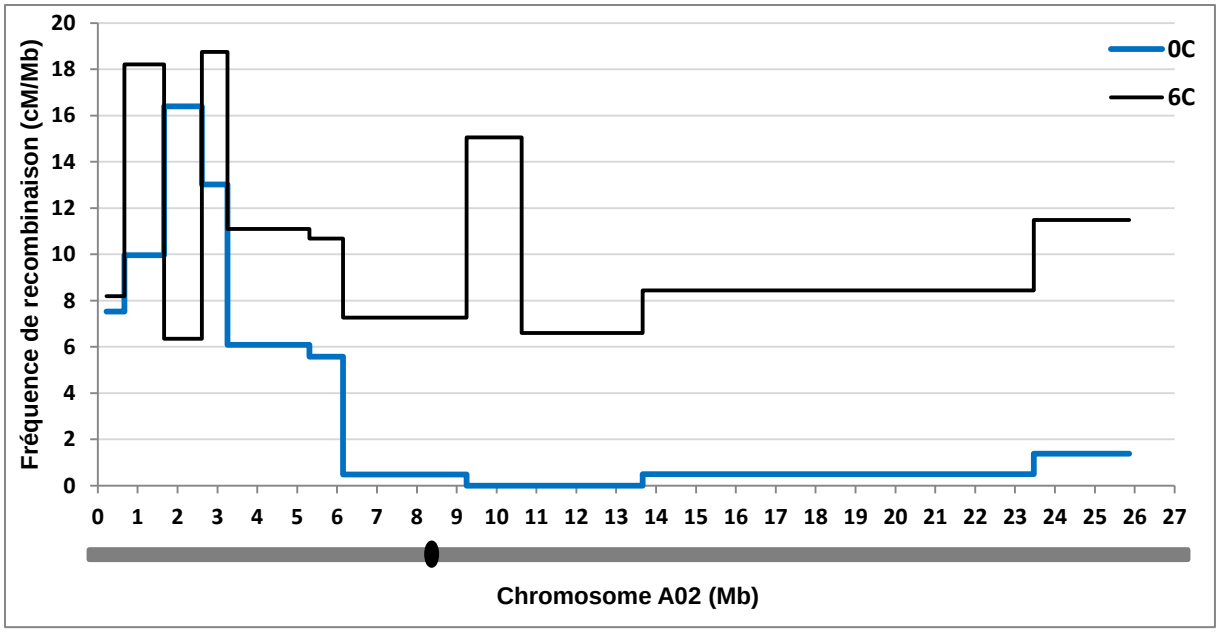


Figure 1. Représentation de la distribution des évènements de recombinaison le long du chromosome A02 de *B. rapa* chez les populations 0C et 6C. Le chromosome identifié par les blocs de Schranz est schématisé en abscisse (en gris) avec sa région centromérique (en noire).

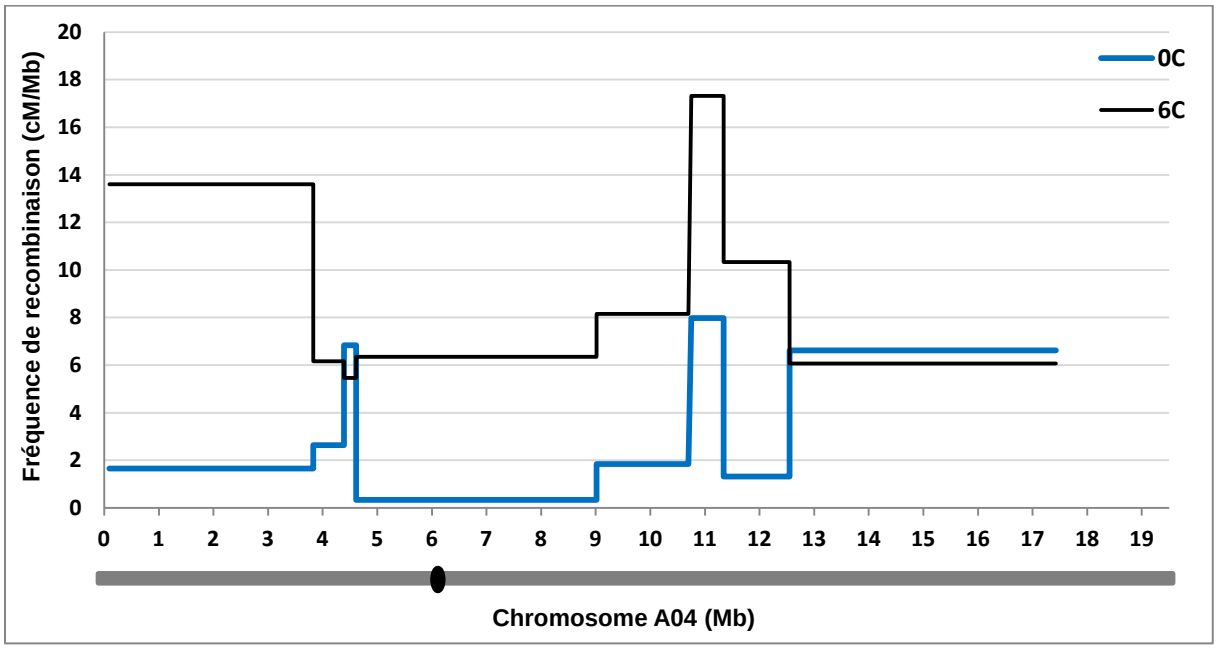


Figure 2. Représentation de la distribution des évènements de recombinaison le long du chromosome A04 de *B. rapa* chez les populations 0C et 6C. Le chromosome identifié par les blocs de Schranz est schématisé en abscisse (en gris) avec sa région centromérique (en noire).

Diplôme: Diplôme d'Ingénieur de l'Institut des Sciences Agronomiques, Agroalimentaires, Horticoles et du Paysage

Spécialité: Horticulture

Spécialisation / option: Amélioration des plantes

Responsable d'option: Mme Maria MANZANARES-DAULEUX

Auteurs(s): Alexandre PELE

Date de naissance: 08 Juin 1989

Nb pages: 28 Annexe(s): 21

Année de soutenance: 2013

Organisme d'accueil: INRA Le Rheu

Adresse: INRA Le Rheu,
Domaine de la Motte au Vicomte BP 35327
35653 LE RHEU CEDEX

Maître de stage: Mme Anne-Marie CHEVRE

Titre français: La distribution des évènements de crossing-over, un paramètre modulable chez les Brassicas

Titre anglais: The distribution of crossover events, an adjustable parameter in Brassicas

Résumé: La recombinaison méiotique, par l'intermédiaire de Crossing-over (CO) est le principal mécanisme permettant le brassage génétique de la diversité lors de la production de descendance dans les programmes de sélection variétale. Cependant le nombre et la répartition des COs sont strictement régulés. En effet, il existe toujours un CO par chromosome mais très rarement plus et leur distribution n'est pas homogène le long des chromosomes. Pourtant, chez les Brassicas, il est possible d'augmenter d'un facteur 6 le taux de recombinaison entre les génomes A chez les hybrides triploïdes AAC (2n=29) par rapport aux diploïdes AA (2n=20), grâce à la nature des chromosomes C en addition et notamment du chromosome C_9. Les séquençages récents des génomes de *B. rapa* et de *B. napus* permettent désormais de faire le lien physique avec ces variations de la fréquence de recombinaison. L'objectif de cette étude est donc de comprendre comment la structure génomique du chromosome A07 conditionne la formation des COs et de déterminer l'effet de l'addition du chromosome C_9 et des 9 chromosomes C chez un diploïde AA sur leur distribution. Nous avons pour ce faire (1) choisi 3 zones contrastées pour la fréquence de recombinaison sur le chromosome A07 chez les populations AA, AA+1C_9 et AAC et (2) densifié puis cartographié des marqueurs microsatellites physiquement ancrés tous les 300 Kb pour caractériser ces zones. Nous avons ainsi montré que la distribution des COs le long du chromosome A07 est expliquée par sa structure génomique chez des diploïdes AA. De nouveaux points de recombinaison apparaissent en présence du chromosome C_9 et des 9 chromosomes C dans des régions froides chez le témoin AA et notamment dans les régions péricentromériques. Ces résultats offrent ainsi la perspective de pouvoir briser certains déséquilibres de liaison et de brasser ces locus conservés en sélection variétale.

Abstract: Meiotic recombination by Crossover (CO) is the primary mechanism for the genetic mixing of diversity in the production of offspring in breeding programs. However, the number and distribution of COs are strictly regulated. Indeed, there is still a CO per chromosome, but rarely more and their distribution is not uniform along the chromosome. Yet in Brassicas, it is possible to increase by a factor of 6 the rate of recombination between the genomes A in AAC triploid hybrid (2n = 29) compared with the AA diploid (2n = 20), due to the nature of C chromosomes in addition, particularly to the C_9. It is now possible to make a physical connection with these changes in the frequency of recombination with the recent sequencing of the genomes of *B. rapa* and *B. napus*. The purpose of this study is to understand how the genomic structure of chromosome A07 determines the formation of COs and to determine the effect of the addition of C_9 chromosome and the 9 chromosomes C in a diploid AA on COs distribution. To make this, we have (1) selected three areas contrasted for the frequency of recombination on chromosome A07 on AA, AA+1C_9 and AAC populations, and (2) densified and mapped microsatellite markers physically anchored every 300 Kb to characterize these areas. We have shown that the distribution of COs along chromosome A07 is explained by the genomic structure in AA diploid. New points of recombination occur in the presence of C_9 chromosome and 9 chromosome C in cold regions in the AA control and especially in pericentromeric regions. These findings offer the prospect of breaking some linkage disequilibrium and stir retained loci in plant breeding.

Mots-clés: Brassicas, recombinaison homologue, génotypage, carte physique, régions péricentromériques

Key Words: Brassicas, homologous recombination, genotyping, physical map, pericentromeric regions