Analyse de la distribution des crossing-overs sur le chromosome 3B du blé tendre (Triticum aestivum) et des facteurs influençant cette distribution

Cyrille Saintenac

▶ **To cite this version:**

Cyrille Saintenac. Analyse de la distribution des crossing-overs sur le chromosome 3B du blé tendre (Triticum aestivum) et des facteurs influençant cette distribution. Génétique des plantes. Université Blaise Pascal - Clermont-Ferrand II, 2009. Français. NNT : 2009CLF21917 . tel-00725777

HAL Id: tel-00725777

https://theses.hal.science/tel-00725777

Submitted on 27 Aug 2012

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

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

Université Clermont-Ferrand II Blaise Pascal *Année 2009*
N° DU : 1917

ECOLE DOCTORALE DES SCIENCES DE LA VIE ET DE LA SANTE
N°ED : 498

DOCTORAT

Présenté à l'Université Blaise Pascal
Pour l'obtention du grade de Docteur d'Université
Spécialité : Génétique

Cyrille Saintenac

Analyse de la distribution des crossing-overs sur le chromosome 3B du blé tendre (*Triticum aestivum*) et des facteurs influençant cette distribution

Soutenu le 30 mars 2009

Président :	S. Mouzeyar (Professeur d'Université)
Rapporteurs :	B. de Massy (Directeur de Recherche, CNRS)
	E. Guiderdoni (Directeur de Recherche, CIRAD)
Examinateurs :	C. Mezard (Directrice de Recherche, INRA)
	C. Feuillet (Directrice de thèse)
	P. Sourdille (Directeur de thèse)

Institut National de la Recherche Agronomique Site de Crouël UMR 1095 Génétique, Diversité et Ecophysiologie des Céréales

Résumé

Les crossing-overs (CO) sont indispensables dans la création variétale pour permettre l'introgression de régions d'intérêt dans les variétés agronomiques d'espèces cultivées telles que le blé tendre (*Triticum aestivum* L.). Afin d'évaluer l'impact des facteurs qui influencent leur formation, nous avons entrepris une caractérisation fine de leur distribution sur le plus grand chromosome (chromosome 3B, 995 Mb) du blé tendre en s'appuyant sur la carte physique récemment développée et le séquençage de quelques régions de plusieurs mégabases.

La comparaison entre une carte génétique dense (102 marqueurs) et une carte physique de délétion montre que 77 % des CO sont présents dans les régions distales couvrant seulement 25 % du chromosome. La comparaison de différentes cartes génétiques montre de plus que cette distribution est conservée entre populations avec cependant des différences de taux de CO locaux entre populations mais également entre méiose mâle et femelle. Cette distribution est influencée par une interférence positive forte à des distances inférieures à 10 cM. Cependant, les faibles fréquences de CO observées au sein des régions proximales restent inexpliquées. En effet, la faible augmentation du taux de CO observée au sein des régions proximales placées en position distale suggère que la position proximale de ces régions sur le chromosome ne semble pas responsable de leur faible fréquence de CO. De plus, nous avons montré que ces faibles fréquences ne seraient pas non plus dues à une divergence de séquence entre chromosomes homologues au sein des régions proximales, la fréquence de CO étant toujours aussi faible au sein de celles-ci entre deux chromosomes homozygotes.

En revanche, l'analyse à l'échelle d'une région séquencée de 3.1 Mb indique que les fréquences de CO importantes sont fortement corrélées avec la présence de gènes. L'inhibition de la formation des CO au sein des régions proximales pourrait ainsi s'expliquer par la présence de gènes en quantité moins importante dans ces régions comparées aux régions distales.

Mots clés : blé, crossing-over, cartes génétiques, cartes physiques, interférence, localisation des gènes.

Abstract

Crossovers (CO) are necessary in breeding to introgress regions of interest in agronomic varieties from cultivated species such as bread wheat (*Triticum aestivum* L.). In order to evaluate the impact of factors influencing CO formation, we started the characterisation of their distribution on the largest wheat chromosome (chromosome 3B, 995 Mb) using the recently developed physical map as well as some sequenced regions of few megabases.

Comparison between a dense genetic map (102 markers) and a physical deletion map shows that 77 % of CO are present in the distal regions covering 25 % of the chromosome. Comparison between different genetic maps indicates that this distribution is conserved despite local differences of CO ratio between different populations and between male and female meiosis. Positive interference impacts significantly CO distribution at distances less than 10 cM on chromosome 3B. However, factors involved in the inhibition of CO formation in the proximal regions remain unknown. Positioning proximal regions in distal region of the chromosome increase only weakly the frequency of CO, suggesting that the low frequency observed in the proximal regions is not due to their proximal position. Moreover, sequence divergence between the two homologous chromosomes is not involved in the low frequency observed in the proximal regions since homozygous regions do not exhibit more CO than heterozygous ones.

On the contrary, analyses at the level of a 3.1 Mb sequenced region show that high CO frequencies correlate well with gene occurrence. Low CO frequency observed in the proximal regions could thus be explained by a low gene density.

Key words: wheat, crossing-over, genetic map, physical map, interference, gene localisation.

Remerciements

Je voudrais tout d'abord remercier l'ensemble des personnes que j'ai côtoyé au cours de ma thèse. Ils ont grandement participé à l'ensemble de ce travail et ont fait de ces trois années passées au sein de l'UMR Génétique, Diversité et Ecophysiologie des Céréales une période très agréable. Je ne les remercierai jamais assez pour tout ce qu'ils ont fait.

Merci à Gilles Charmet pour m'avoir accueilli au sein de l'unité GDEC.

Merci à Catherine Feuillet de m'avoir accordé sa confiance en me proposant ce sujet de thèse. Merci de m'avoir donné l'opportunité de participer à des congrés internationaux. Merci pour sa disponibilité, sa simplicité, son enthousiasme, ses connaissances. La liste serait encore longue… Tout simplement merci à cette femme en or.

Merci à Pierre Sourdille pour son encadrement de tous les jours, ses conseils, sa présence, sa sympathie, son aide précieuse dans l'élaboration de ce mémoire. Merci pour toutes les discussions intéressantes que nous avons eu ensemble. En espérant que cette thématique, récemment développée au sein de l'équipe, continue de progresser.

Merci à Etienne Paux pour sa gentillesse et pour avoir été à l'écoute des milliards de questions que j'ai pu lui poser.

Merci à Delphine Boyer sans qui ces trois années n'auraient pas été les mêmes. Merci pour son aide technique, sa disponibilité, les innnombrables services qu'elle a pu me rendre mais aussi merci pour ces soirées animées et innoubliables au bout du monde (je crois que je ne connais toujours pas la bonne route !). En tout cas MERCI pour tout.

Un énorme merci à Jacqueline Philippon pour toute son aide technique, son application, sa patience et sa persévérance notamment dans la recherche des marqueurs. Merci aussi pour les soins précieux apportés aux plantes et les milliers de sacs posés sur chaque épi (je revois l'image de la chambre de culture).

Merci à Sébastien Faure pour m'avoir grandement aidé au sein de cette thématique recombinaison et pour les échanges intéressants que nous avons pu avoir. Merci pour sa sympathie, sa maîtrise de la langue anglaise et la relecture de rédactions. Merci aussi pour les soirées sympathiques foot-fléchettes (Allez Lyon).

Merci à Fredéric Choulet pour l'énorme travail réalisé sur les contigs (facile il va me dire). Merci pour sa disponibilité et ses connaissances bio-informatiques très précieuses. Merci aussi pour son enthousiasme et sa folie du jeu qui ont souvent mis de l'ambiance.

Merci à Matthieu Falque de l'INRA du Moulon et à Olivier Martin (Université Paris Sud) pour le travail formidable qu'ils ont réalisé concernant l'analyse de l'interférence mais aussi pour l'étude des comparaisons entre taux de CO.

Merci à Georges Gay pour la réalisation des expériences de C-banding. Merci pour toute son application lors des nombreux croisements effectués.

Merci à Mathieu Chicard pour l'ensemble du travail effectué sur le contig *Rph7*.

Merci aux membres de mon comité de thèse (Jacques David, Christine Mezard, Alain Murigneux et Charles White) pour m'avoir guidé, conseillé et pour avoir partagé leurs connaissances avec moi.

Merci à Nicolas Guihot pour toute son aide lors de mes nombreux problèmes informatiques (word n'a plus de secret pour moi maintenant). Merci aussi à Jeff pour son aide informatique.

Comment ne pourrais-je pas remercier mon préparateur physique Bouzid Charef qui m'a permis d'être au meilleur de ma forme pour réaliser cette thèse. Merci aussi pour son aide technique.

Merci aussi à Jerome Salse, Michel Bernard, Jacques Legouis pour les discussions interessantes que nous avons pu avoir.

Merci à l'équipe des serres (Christophe, Richard et Raphaël) pour le travail d'entretien des plantes. Merci aussi à Alain pour les croisements effectués, les récoltes et l'entretien des plantes.

Merci à Monique, Patricia et Valérie pour leur gentilesse, leur aide dans les démarches administratives et dans l'organisation des dépalcements.

Un merci tout particulier à Ludo, Manu, Vincent et Jean-Baptiste avec qui j'ai discuté des heures entières sur l'évolution de nos thèses. Merci pour tous les moments inoubliables (et il y en a eu) passés avec vous. Merci aussi à Walid, Umar, Camille et Mathieu pour leur aide au quotidien.

Merci aux anciens membres du bureau ASICT (Guillaume, Anne-Laure, Véronique, Julien et Jeff) pour cette année sympathique que nous avons passée ensemble.

Merci à Charles pour sa sympathie son enthousiasme et sa joie de vivre.

Merci à Jean et à toute l'équipe de foot de Crouël pour les bons moments passés lors des coupes INRA. Merci aussi à tous les joueurs rencontrés lors des matchs hebdomadaires et en particulier merci à Laurent pour maintenir cette dynamique.

Merci à l'ensemble des personnes que j'ai côtoyé qui ont rendu ce moment agréable et qui m'ont donné un coup de main (Stéphane, Florence, Philippe, Karine, Aurélie, Séverine, Mireille, Catherine, Marie-Reine, Lifeng, Chenyang, Karine, Alicia, Pierre, Nathalie, Gilles, Audrey, Aniko, Jacques, Faha, Bruno…).

Merci aussi à mes amis de Saint-Paulien et d'ailleurs (Christophe, Jérôme, Laurent, Ludo, Yohan, Fred, Adeline, Emilie, Jean-Marie, Vincent, Pierre, Fred, Franck, Cédric…) qui ont tenté de comprendre ces travaux et qui m'ont donné un coup de main dans les moments difficiles (merci Jérôme pour les cadeaux de noël !).

Merci aussi à Aurélie pour la relecture de ce rapport et toute son attention pendant ces derniers mois un peu difficile.

Enfin, je n'ai pas souvent eu l'occasion de remercier mes parents donc je voulais leur dire un énorme MERCI pour tout ce qu'ils font depuis des années pour me permettre d'arriver à concrétiser toutes ces années d'études (et j'espère les voir dans le Kansas).

Abréviations

ADN : Acide Désoxyribonucléique
Are : Arche x Récital
AS-PCR : Allele Specifique Polymerase Chain Reaction
BAC : Bacterial Artificial Chromosome
BGC : Biaised Gene Conversion
BIR : Break-Induced Replication
BLAST : Basic Local Alignment Search Tool
BPA : Bisphénol A
bz : bronze
CDK : Cycline dependent kinase
cM : centiMorgan
CNG : Centre National de Génotypage
CO : Crossing-over
Cs : Chinese spring
CsCtA : Chinese spring x Courtot Androgenèse
CsCtG : Chinese spring x Courtot Gynogenèse
Ct : Courtot
DAPI : 4',6' Di Amidino-2-Phényl Indole
DArT : Diversity Array Technology
DSBR : Double Strand Break Repair
EDTA : Ethylene Diamine Tetra Acetic acid
EST : Expressed Sequence Tag
FAO : Food and Agricultural Organisation
FISH : Fluorescent *in situ* Hybridization
Gb : Gigabase
GISH : Genomic *in situ* Hybridization
HD : Haploïdes Doublés
INRA : Institut National de la Recherche Agronomique
ISBP : Insertion Site Based Polymorphism
ITMI : International Triticeae Mapping Initiative
IWGSC : International Wheat Genome Sequencing Consortium
kb : kilobase
LOD : Logarithm of the Odd ratio
LTR : Long Terminal Repeat
MAT : Mating-Type
Mb : Mégabase
NCO : Non Crossing-Over
NHEJ : Non Homologous End-Joining
NR : Nodule de Recombinaison
Op : Opata
ORF : Open Reading Frame
PCR : Polymerase Chain Reaction
pb : paire de bases
QTL : Quantitative Trait Loci
Re : Renan
RNAi : Ribonucleic Acid Interference
SDSA : Synthesis-Dependent Strand Annealing
Sh : Shrunken
SNP : Single Nucleotide Polymorphism
SSA : Single Strand Annealing
SSR : Simple Sequence Repeat
Sy : Synthétique ou W7984
TREP : Triticeae Repeat sequence database
VDJ : Variable/Diversity/Joining genes
Wx : Waxy
ZMM : Zip, Msh, Mer

Table des Matières

I Introduction générale 1

II Bibliographie 7

1 La recombinaison : definitions et roles 9

1.1 Définitions 9

1.2 Rôle de la recombinaison 9

2 Meiose et recombinaison meiotique 11

2.1 Généralités sur la méiose 11

2.2 Différentes phases de la méiose 11

2.3 La prophase I : étape clé pour la formation des événements de recombinaison 19

2.4 Gènes et protéines impliqués dans la recombinaison homologue 37

3 Distribution et regulation des CO sur les chromosomes 51

3.1 Techniques d'étude de la distribution des événements de recombinaison 51

3.2 Fréquence moyenne de recombinaison chez différentes espèces 59

3.3 Distribution des CO le long des chromosomes 59

3.4 Facteurs influençant les variations de fréquence et de distribution des CO 65

3.5 Synthèse 105

4 La recombinaison meiotique chez le ble 107

4.1 Le blé tendre 107

4.2 Analyse détaillée de la méiose 109

4.3 Distribution des CO sur les chromosomes 113

4.4 Facteurs génétiques impliqués dans la recombinaison chez le blé 117

5 Objectifs de la these 121

III Matériels et Méthodes 125

1 Materiel vegetal 127

1.1 Lignées parentales 127

1.2 Lignées aneuploïdes 129

1.3 Populations F2 129

1.4 Populations haploïdes doublés 131

1.5 Population F3 Chinese spring x Renan 3BS4 133

2 Extraction d'ADN 135

3 Marqueurs moleculaires et conditions d'amplification 135

3.1 Marqueurs moléculaires 135

3.2 Conditions de génotypage des marqueurs SSR et ISBP 139

3.3 Conditions de génotypage des marqueurs SNP 141

4 Cartographie genetique 143

5 CALCUL DES FREQUENCES DE CO 143
6 COMPARAISON DES TAUX DE CO 145
6.1 Calcul du taux de CO 145
6.2 Calcul du taux de CO pour les régions homozygotes Cs sur le bivalent 3B délété 145
6.3 Comparaison des taux de CO 147
7 ETUDE DE L'INTERFERENCE 147
7.1 Test de l'hypothèse H_0 de non-interférence 147
7.2 Le modèle gamma 149
7.3 Intervalle de confiance du paramètre m 149
7.4 Différence entre le paramètre m issu du mâle et celui issu de la femelle 149
7.5 Le coefficient de coïncidence 151
8 ANALYSE DES BANDES C DU CHROMOSOME 3B 151
9 ANALYSE DES SEQUENCES 153
9.1 Annotation des séquences 153
9.2 Calcul des valeurs des différents paramètres de séquences et des coefficients de détermination R^2 153

IV Résultats 157

1 DISTRIBUTION DES CO A L'ECHELLE DU CHROMOSOME ENTIER 159
1.1 Développement et analyse de marqueurs moléculaires 159
1.2 Différentes cartes génétiques développées pour l'étude de la distribution des CO 165
1.3 Distribution des CO sur le chromosome 3B : Comparaison de la carte génétique CsRe avec la carte physique de délétion 167
1.4 Comparaison des cartes génétiques CsCtG et CsCtA : Différence de taux de CO en fonction du sexe 171
1.5 Comparaison des cartes génétiques CsCt et CsRe : Variations du taux de CO sur le chromosome 3B en fonction du croisement étudié 171
1.6 L'interférence entre CO 173
2 DISTRIBUTION DES CO SUR UN CHROMOSOME 3B DELETE 175
2.1 Caractérisation fine du bivalent 3B de l'individu F2CsRe3BS4 n°369 177
2.2 Distribution des CO sur un bivalent 3B composé d'un chromosome délété et d'un chromosome entier 179
2.3 Distribution des CO sur le bivalent 3B composé des deux chromosomes délétés 181
3 DISTRIBUTION DES CO A L'ECHELLE DE REGIONS SEQUENCEES 185
3.1 Marqueurs développés sur les différents contigs étudiés 187
3.2 Distribution des CO sur les contigs séquencés 191

V Discussion et perspectives 203

1 DISTRIBUTION DES CO SUR LE CHROMOSOME 3B 205
2 CHUTE DE LA FREQUENCE DE CO A L'EXTREMITE DU BRAS COURT 211
3 DIFFERENCE DU TAUX DE CO ENTRE LA MEIOSE MALE ET LA MEIOSE FEMELLE 215

4 PRESENCE D'INTERFERENCE POSITIVE SUR LE CHROMOSOME 3B 219
5 EFFET DE L'ETAT HETEROZYGOTE CHINESE SPRING/RENAN SUR LA DISTRIBUTION DES CO A L'ECHELLE DU CHROMOSOME 221
6 EFFET DE LA POSITION DISTALE SUR LA DISTRIBUTION DES CO 225
7 DEVELOPPEMENT D'OUTILS DE GENOMIQUE, CARTOGRAPHIE GENETIQUE ET MODIFICATION DE LA FREQUENCE DE CO 229
7.1 Développement d'outils de génomique 229
7.2 Cartographie génétique 233
7.3 Modification de la fréquence de CO dans les régions proximales 237
8 IMMUNO-LOCALISATION DES PROTEINES INTERVENANT DANS LA RESOLUTION DES JONCTIONS DE HOLLYDAY 239
9 LA TAILLE PHYSIQUE CUMULEE DES GENES GOUVERNE LA DISTRIBUTION DES CO 241
10 LES CARACTERISTIQUES DE LA SEQUENCE NE SONT PAS FORTEMENT CORRELEES AVEC LA DISTRIBUTION DES CO 251
11 L'HETEROZYGOTIE CS/RE IMPACTE LA DISTRIBUTION DES CO SUR LE CONTIG *RPH7* 253

VI Conclusion générale 257

VII Références 267

VIII Annexes 311

IX Publication 351

Tables des Figures

Figure 1 *Différentes phases de la méiose* ... 12
Figure 2 *Modèle de recombinaison homologue proposé par R. Hollyday en 1964.* ... 22
Figure 3 *Visualisation de la double jonction de Holliday par microscopie électronique.* ... 22
Figure 4 *Mécanisme de formation des événements de recombinaison au cours de la méiose.* ... 24
Figure 5 *Détails de la prophase de première division de méiose.* ... 34
Figure 6 *Différentes techniques de cytogénétique utilisées pour étudier la distribution des CO* ... 50
Figure 7 *Variation de la fréquence de CO le long des chromosomes de différentes espèces* ... 58
Figure 8 *Relation entre la taille des génomes et les fréquences moyennes de recombinaison* ... 72
Figure 9 *Schéma d'obtention du blé tendre (Triticum aestivum L.).* ... 106
Figure 10 *Comportement des centromères d'un hybride blé-seigle pendant le début de la prophase I de méiose au cours de la formation du bouquet télomérique (Prieto et al. 2004).* ... 108
Figure 11 *Comportement des régions distales des chromosomes de blé au cours des premiers stades de la prophase I de méiose.* ... 110
Figure 12 *Représentation schématique des chromosomes de blé tendre en début de prophase I.* ... 112
Figure 13 *Lignées aneuploïdes du blé tendre, localisation des marqueurs sur les chromosomes et création des cartes physiques de délétion.* ... 114
Figure 14 *Pose de sachets d'autofécondation sur les épis des plants de blé présents en chambre de culture pour éviter la pollinisation croisée entre différents plants.* ... 130
Figure 15 *Description du protocole d'amplification PCR de type « M13 ».* ... 138
Figure 16 *Génotypage des SNP par la technique AS-PCR* ... 140
Figure 17 *Génotypage des SNP par la technique SNaPshot.* ... 142
Figure 18 *Phase des marqueurs encadrant les régions homozygotes présentes sur le bivalent délété.* ... 144
Figure 19 *Cartes génétiques du chromosome 3B obtenues à partir de différentes populations.* ... 164
Figure 20 *Distribution des CO sur le chromosome 3B.* ... 166
Figure 21 *Différence locale du taux de CO entre la méiose mâle et femelle sur le chromosome 3B.* ... 170
Figure 22 *Comparaison des cartes génétiques F2CsRe et HDCsCt.* ... 172
Figure 23 *Valeurs du coefficient de coïncidence C en fonction des distances entre les deux intervalles étudiés.* ... 174
Figure 24 *Schéma de croisement utilisé pour générer un chromosome 3B délété* ... 176
Figure 25 *Profil des bandes C (C-banding) sur le chromosome 3B* ... 178
Figure 26 *Comparaison de la distribution des CO observés sur un bivalent 3B entier, un bivalent 3B comportant un chromosome entier et un chromosome délété et un bivalent délété.* ... 180
Figure 27 *Variation de la fréquence de CO sur le contig Rph7.* ... 194
Figure 28 *Distribution des CO sur les régions 8 et 9 entre les marqueurs cfb6072 et cfb6123* ... 198
Figure 29 *Distribution des CO sur la région de 23 kb entourant le gène HGA3 entre les marqueurs cfp5046_Aspcr40 et cfb6074.* ... 200
Figure 30 *Différentes caractéristiques du chromosome 3B* ... 210
Figure 31 *Modèle de recombinaison homologue SSA.* ... 313
Figure 32 *Modèle de recombinaison homologue BIR.* ... 315

Table des Tableaux

Tableau 1 *Principaux gènes impliqués dans la recombinaison méiotique.* ... 36

Tableau 2 *Fréquence de recombinaison en cM/Mb pour différentes espèces de plantes.* ... 57

Tableau 3 *Principaux points chauds de CO étudiés chez les eucaryotes supérieurs.* ... 62

Tableau 4 *Différents types de microsatellites testés* ... 134

Tableau 5 *Assignation et polymorphisme de 217 marqueurs SSR provenant de données publiques et privées et présents sur les cartes génétiques du chromosome 3B.* ... 160

Tableau 6 *Assignation et polymorphisme de 70 marqueurs SSR issus des séquences d'extrémités de BAC de la banque du chromosome 3B.* ... 162

Tableau 7 *Caractéristiques principales des trois contig étudiés.* ... 186

Tableau 8 *Polymorphisme des SSR développés sur les contigs Sr2 et Rph7.* ... 188

Tableau 9 *Taux de polymorphisme des SSR entre Chinese spring et Renan sur le contig Rph7 en fonction du type de motif des SSR.* ... 188

Tableau 10 *Valeurs des coefficients de détermination R^2 testant la relation entre la distribution des CO et les caractéristiques de la séquence sur les différents intervalles du contig Rph7.* ... 196

Tableau 11 *Fréquence moyenne de recombinaison chez différentes espèces.* ... 319

Tableau 12 *Différences de fréquences de recombinaison entre la méiose mâle et la méiose femelle chez différentes espèces animales et végétales.* ... 325

Tableau 13 *Composition du milieu solide de type Muller utilisé pour le dévelopement des embryons.* ... 327

Tableau 14 *Marqueurs et conditions d'amplification utilisés pour l'étude chromosome entier.* ... 335

Tableau 15 *Séquences des amorces développées pour révéler le polymorphisme de type SNP ou insertion/délétion au niveau du locus Rph7* ... 337

Tableau 16 *Table de Bravais-Pearson.* ... 339

Tableau 17 *Analyse détaillée de la comparaison entre la carte physique de délétion du chromosome 3B et la carte génétique issue de la population F2CsRe.* ... 347

Tableau 18 *Différences du taux de CO local entre la méiose mâle et la méiose femelle.* ... 349

I Introduction générale

Jusqu'à maintenant, la production mondiale de céréales a couvert les besoins humains, dépassant même la croissance de la population ce qui a conduit à des excédents sur les trente dernières années. Cependant, ce n'est plus le cas et la croissance démographique qui est estimée pour les 50 prochaines années à 9 Milliards d'ici 2050 va induire simultanément un accroissement de la demande mondiale à la fois pour l'alimentation mais également pour des besoins industriels. L'organisation mondiale de l'alimentation et de l'agriculture (FAO pour Food and Agricultural Organisation) estime que la production de céréales devra augmenter de 15 % pour suivre les demandes croissantes en viande dans les pays en dévelopement ainsi que les demandes pour la fabrication de biocarburants (Edgerton 2009). Si la réponse traditionnelle à une demande accrue a jusqu'à présent été une augmentation des surfaces cultivées, cela n'est plus possible du fait d'une réduction des surfaces de terres arables en particulier dans les pays en voie de développement. De plus, cette pratique devient non envisageable compte tenu de son impact néfaste sur les gaz à effet de serre et la biodivisersité. Cet état de fait va conduire inéluctablement à une diminution des stocks avec potentiellement des conséquences désastreuses pour la population ainsi que des risques de famine.

De plus, l'agriculture moderne doit faire face à des contraintes écologiques. En effet, l'agriculture intensive a induit des modifications importantes de l'environnement avec des déforestations et des suppressions de haies et de fossés. Il en a résulté une dégradation des sols, ainsi qu'une augmentation des phénomènes d'inondation. De même, l'utilisation intensive des engrais, des herbicides et des pesticides a conduit à réduire la diversité faunistique et floristique. Enfin, les dérèglements climatiques de ces dernières années et notamment les périodes importantes de sécheresse dans certains pays ont conduit à une chute catastrophique de la production mondiale.

L'amélioration des rendements des céréales dans le contexte d'une agriculture durable est donc un enjeu majeur pour répondre aux besoins de l'humanité d'ici 2050.
Une façon de relever ce challenge est d'une part d'augmenter la connaissance de ces espèces ainsi que les outils nécessaires à ce dévelopement et d'autre part de mieux gérer et exploiter les ressources génétiques disponibles chez les céréales (Able et al. 2007; Feuillet et al. 2008). Si les ressources génétiques sont extrêmement étendues (environ 10 000 espèces connues avec de nombreuses accessions pour chacune), elles restent largement sous-exploitées.

Ceci est dû : (1) à une faible connaissance des capacités de ces ressources en terme de résistance aux stress biotiques ou abiotiques ainsi que de leur potentiel pour la panification ou les rendements ; (2) à la présence de nombreux caractères indésirables et de faible intérêt agronomique comme la hauteur de plante, la sensibilité à la verse, l'égrainage ou les grains vêtus. Cependant, il y a un regain d'intérêt dans l'exploration de la diversité génétique en raison de la présence de gènes de résistance aux champignons mais aussi du fait de la présence de résistances aux stress abiotiques comme la sécheresse. Sous l'impulsion des outils de génomique tels que le marquage moléculaire, la sélection assistée par marqueurs pour les caractères d'intérêt agronomique serait un outil puissant pour exploiter ces gènes.

Une façon d'utiliser cette diversité génétique dans des programmes de sélection est d'effectuer des croisements entre des espèces sauvages et des variétés élites pour introduire dans le patrimoine génétique de ces dernières des fragments provenant des génomes sauvages et portant les gènes d'intérêt. L'introgression de ces régions génomiques est obtenue par recombinaison méiotique entre les chromosomes du donneur et ceux du receveur. La recombinaison méiotique est cependant un processus complexe qui reste encore mal connu chez les céréales. Elle implique de nombreux gènes et dépend d'un nombre considérable de facteurs.

Afin de mieux maîtriser l'exploitation des ressources génétiques et leurs utilisations dans l'amélioration des variétés élites, il apparaît essentiel de mieux connaître la recombinaison méiotique chez les céréales pour faciliter l'introduction d'allèles nouveaux originaux permettant une agriculture respectant les contraintes environnementales. Cette thèse a donc été initiée dans le but de mieux connaître la recombinaison méiotique chez le blé tendre.

II Bibliographie

1 La recombinaison : définitions et rôles

1.1 Définitions

Il est très difficile de donner une définition générale au terme recombinaison. Il peut prendre plusieurs sens selon la personne qui l'emploie ; il peut ainsi correspondre au mécanisme moléculaire en lui-même mais également à un mécanisme d'évolution des espèces et de brassage du matériel génétique ou encore à un mécanisme de réparation de l'ADN. La recombinaison implique aussi un nombre important de gènes et de protéines et entraîne une modification particulière de la structure de la molécule d'ADN.

La recombinaison est ainsi un terme très général et de ce fait, il est souvent associé à un adjectif le qualifiant (Lichten 2001a). La recombinaison homologue se déroule entre deux séquences identiques quelles que soient leur localisation sur les chromosomes. Elle s'oppose à la recombinaison non-homologue ou illégitime qui joint des séquences non identiques. La recombinaison allélique concerne la formation d'un événement entre deux séquences situées au même locus sur les chromosomes alors que la recombinaison ectopique se forme à partir de deux séquences situées en position non allélique. Les adjectifs mitotique et méiotique sont également employés pour qualifier un type particulier de recombinaison qui se déroule dans les cellules respectivement en mitose ou en méiose.

1.2 Rôle de la recombinaison

Le rôle majeur de la recombinaison est de permettre une réparation des lésions de l'ADN et en particulier assurer une protection contre les cassures double-brin. Ces cassures peuvent être volontaires et donc déclenchées par la cellule ou au contraire provoquées par des agents extérieurs comme les agents ionisants ou encore les radicaux libres. La recombinaison assure ainsi l'intégrité du génome tout au long du cycle cellulaire pour réparer les dommages irréversibles causés par ces agents extérieurs. En réparant les cassures induites par la cellule, elle assure des rôles essentiels pour le dévelopement et la survie de l'organisme. Quelques exemples sont maintenant bien connus. Chez l'humain et d'autres vertébrés, la recombinaison non-homologue appelée recombinaison VDJ (Variable/Diversity/Joining genes) permet de générer une diversité considérable d'anticorps qui est essentielle pour combattre l'immense diversité d'antigènes présent dans la nature (Bassing et al. 2002).

Chez la levure, la recombinaison homologue entre les loci MATa et MATα entraîne le changement de sexe de celle-ci (Haber 1998). La recombinaison homologue permet également le maintien de la longueur des télomères en absence de télomérase (Sung et Klein 2006). Enfin, la recombinaison homologue méiotique est essentielle à la reproduction de la plupart des organismes supérieurs car elle permet une ségrégation correcte des chromosomes et entraîne également une augmentation de la diversité dans la descendance (Cromie et Smith 2007). Celle-ci va faire l'objet d'une étude plus détaillée dans la suite de cette partie introductive.

2 Méiose et recombinaison méiotique

2.1 Généralités sur la méiose

La méiose (signifie en grec, réduction) est un phénomène universel. Retrouvée chez pratiquement tous les eucaryotes, elle joue un rôle central dans le cycle vivant. Elle permet le passage de l'état diploïde à l'état haploïde, générant ainsi les gamètes mâles et femelles essentiels à la reproduction sexuée. La méiose est précédée d'une phase de réplication de l'ADN, appelée réplication pré-méiotique et caractérisée par une phase G, une phase S rallongée mais pas de phase G2. La méiose en-elle même se scinde en deux grandes divisions : une division réductionnelle ou première division de méiose et une division équationnelle ou deuxième division de méiose. Chacune de ces divisions se découpe en quatre phases : la prophase, la métaphase, l'anaphase et la télophase.

L'entrée en méiose et donc le passage d'un cycle mitotique à un cycle méiotique semble être contrôlée de manière génétique (Hamant et al. 2006). Chez le maïs, la mutation du gène *ameiotic1-1* entraîne la formation de divisions mitotiques à la place de divisions méiotiques. Le choix du passage en cycle méiotique est conditionné dès la phase de réplication pré-méiotique car une forme spéciale de cohésine responsable de l'association des chromatides sœurs est retrouvée uniquement pendant la méiose et est mise en place dès le stade pré-méiotique (Petronczki et al. 2003).

2.2 Différentes phases de la méiose

Paradoxalement, le processus de méiose a été étudié bien avant la découverte de la structure de l'ADN. La méiose a été découverte et décrite pour la première fois chez le tabac par Kölreuter en 1766 puis chez l'oursin en 1876 par O Hertwig (Petronczki et al. 2003).

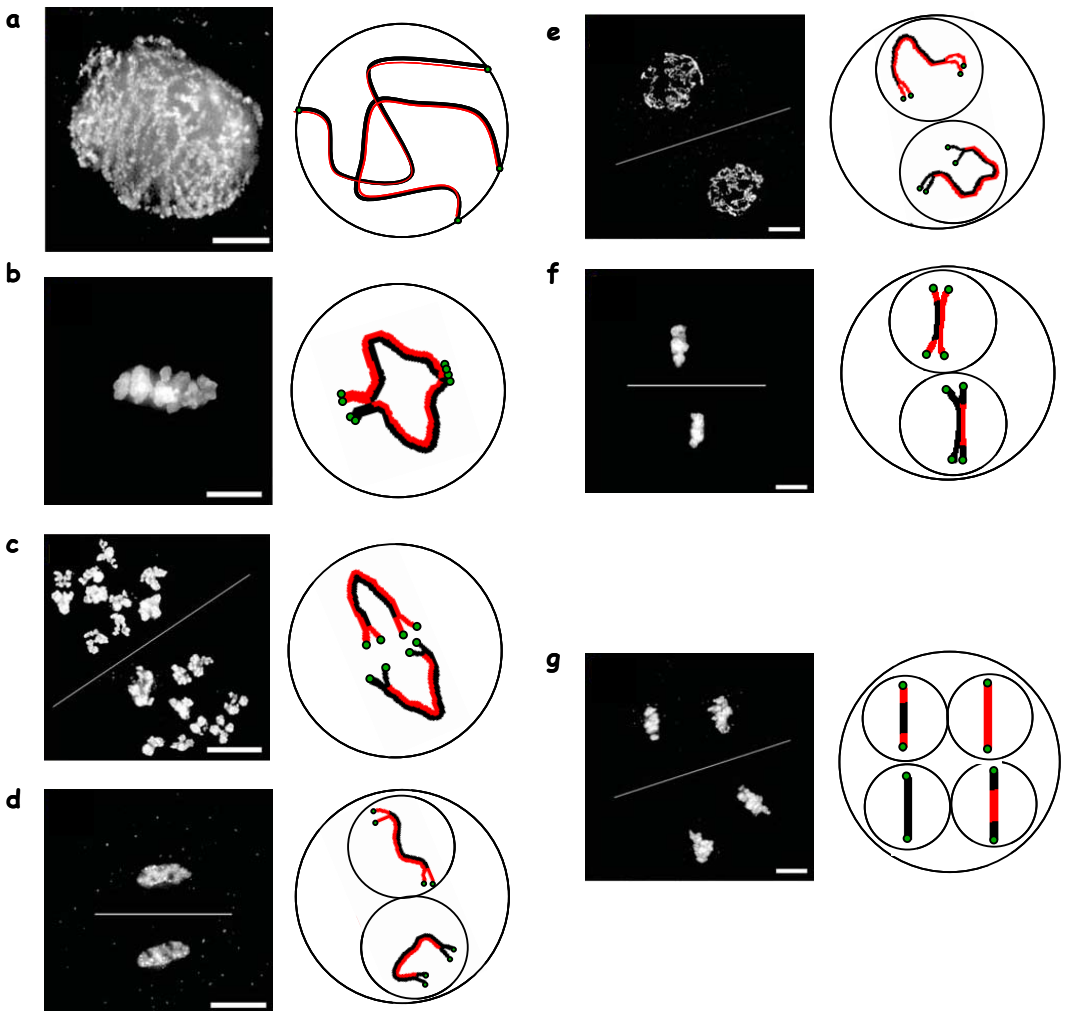


Figure 1 Différentes phases de la méiose.

Les différentes phases de la méiose sont représentées soit par une image montrant les chromosomes de maïs marqués au DAPI (Hamant et al. 2006) soit par un schéma d'une paire de chromosomes homologues où les deux chromatides sont représentées l'une en rouge et l'autre en noire. **a-b-c-d**/ Première division de méiose. **a**/ Prophase I : les chromosomes sont visibles sous forme de longs filaments et ils commencent à se condenser. **b**/ Métaphase I : les chromosomes sont très condensés et ils s'alignent sur la plaque métaphasique. Les chromosomes homologues sont liés au niveau des régions ayant subi un crossing-over qui ont entraîné un échange de chromatide entre chromosomes homologues. Les centromères tirent les chromosomes vers un des pôles de la cellule. **c**/ Anaphase I : les chromosomes homologues se séparent et migrent chacun vers un pôle de la cellule. Les chromatides sœurs restent liées sur les régions proximales aux crossing-over. **d**/ Télophase I : la migration des chromosomes est terminée, deux nouvelles cellules haploïdes bichromatidiennes se forment. **e-f-g**/ Deuxième division de méiose. **e**/ Prophase II. **f**/ Métaphase II : les chromosomes s'alignent sur la plaque équatoriale. **g**/ Anaphase et telophase II : les chromatides de chaque chromosome se séparent et chacune migre vers un des pôles de la cellule. Deux nouvelles cellules haploïdes mono-chromatidiennes se forment qui représentent les gamètes. Les barres sur les images représentent 5 µm. Les points verts représentent les télomères.

Son aspect chromosomique a été découvert quelques années plus tard (1883) par E Van Beneden chez l'ascaris et son implication dans la reproduction a été mise en évidence en 1890 par A.Weismann. La méiose est divisée en deux étapes (première et deuxième division) de quatre phases principales (Figure 1).

2.2.1 La première division de méiose

Lors de la première division ou division réductionnelle, les chromosomes homologues se séparent, entraînant la formation de deux cellules haploïdes bi-chromatidiennes.

2.2.1.1 Prophase I

Au cours de cette phase, il y a condensation des chromosomes qui s'apparient puis formation des événements de recombinaison, les crossing-overs (CO) et les non crossing-overs (NCO). Les nucléoles se dissipent et il y a migration des centrioles de part et d'autre de la cellule. Les fibres du fuseau de microtubules commencent à se former. Cette étape étant prépondérante pour la recombinaison, elle sera reprise plus en détail par la suite.

2.2.1.2 Métaphase I

Lors de la métaphase, les chromosomes s'attachent aux fibres des fuseaux de microtubule au niveau des kinétochores, structure protéique située au niveau des centromères. Ils migrent ainsi sous forme de paires, les deux homologues étant attachés par les CO, vers la plaque métaphasique. A la fin de ce processus, l'ensemble des chromosomes est aligné sur la plaque métaphasique ou équatoriale.

2.2.1.3 Anaphase I

L'anaphase voit le raccourcissement des fibres du fuseau de microtubules, les chromosomes homologues sont tirés de part et d'autre de la cellule par contraction des fibres et il y a déplacement des deux lots de chromosomes jusqu'aux deux pôles de la cellule.

2.2.1.4 Télophase I

Au cours de cette dernière étape, il y a reformation de l'enveloppe nucléaire et dispersion des chromosomes. Les fibres du fuseau sont dissoutes et la cellule est coupée en deux par un anneau de protéines (actine) : c'est la cytokinèse. Ceci permet la formation de deux nouvelles cellules haploïdes bi-chromatidiennes.

2.2.2 La deuxième division de méiose

Lors de la deuxième division ou division équationnelle, celle-ci étant identique à la mitose, il y a séparation des chromatides sœurs. Chaque cellule donne alors naissance à deux cellules haploïdes mono-chromatidiennes. En prophase II, les chromosomes encore constitués des deux chromatides se condensent. En métaphase II, les centromères se positionnent sur la plaque métaphasique. En début d'anaphase II, la cohésine encore présente au niveau des centromères est clivée ce qui permet la séparation des deux chromatides sœurs. Les chromosomes migrent vers les pôles opposés de la cellule. Enfin, en télophase II, les chromosomes se décondensent et l'enveloppe nucléaire se forme. Le cytoplasme se partage entre les deux cellules filles (cytodiérèse) ce qui entraîne la formation de quatre cellules filles haploïdes mono-chromatidiennes qui représentent ainsi les gamètes.

2.2.3 Différents points de contrôle des étapes de la méiose

Au cours de la méiose, la progression des différents événements nécessite une coordination temporelle précise pour éviter des perturbations dans le déroulement de celle-ci et donc dans la transmission du matériel génétique. La cellule possède ainsi tout au long de la méiose des points de contrôle pour veiller à un enchaînement synchronisé des différentes étapes. Deux points de contrôle sont principalement étudiés : le point de couplage entre la réplication pré-méiotique de l'ADN et la recombinaison et le point de contrôle de l'état de l'ADN (Hochwagen et Amon 2006).

2.2.3.1 Réplication pré-méiotique

La méiose est précédée d'une phase de réplication de l'ADN appelé réplication pré-méiotique qui est indispensable pour la recombinaison. Le couplage de ces deux événements se fait de manière très locale. Borde et al. (2000) ont montré qu'un retard de réplication sur un segment précis d'un chromosome retarde la formation des cassures double-brin spécifiquement sur ce segment et non ailleurs. Certaines protéines impliquées dans ce contrôle sont à présent connues. Par exemple chez *S. pombe*, lorsque le gène codant la protéine Rad3 est muté, ceci entraîne une progression de la recombinaison même si la réplication est arrêtée (Hochwagen et Amon 2006).

2.2.3.2 Etat de l'ADN

Indispensable à la méiose, les cassures double-brin sont cependant néfastes pour l'intégrité du génome. Si de telles cassures persistent dans le noyau de la cellule en cours de méiose, le cycle cellulaire est rallongé pour permettre une réparation de l'ADN : c'est le point de contrôle de l'état de l'ADN (Hochwagen et Amon 2006). Ce phénomène a été observé pour un nombre important de mutants incapable de réparer les cassures double-brin chez plusieurs espèces telles que *S. pombe*, la drosophile, la souris et *Arabidopsis thaliana*. Ainsi, un arsenal de protéines vérifie et signale l'état de l'ADN à la cellule pour déclencher ou réprimer la division cellulaire. La plupart des protéines qui participe à ce contrôle est impliquée dans la machinerie de recombinaison (Hochwagen et Amon 2006). Par exemple, la protéine Mre11, en plus de son rôle dans la maturation des cassures double-brin participe à la recherche des dommages retrouvés sur la molécule d'ADN (Borde 2007). De plus, certains composants du complexe synaptonémal (définition p 31) comme les protéines Zip1, Red1 et Hop1 pourraient inhiber la progression du cycle cellulaire jusqu'à la résolution des cassures double-brin (Borner 2006; Hochwagen et Amon 2006).

2.2.4 Conclusion

Au cours de la méiose, les CO, l'attachement particulier des fuseaux de microtubules sur les kinétochores (attachement synthélique à deux orientations, Petronczki et al. 2003) et le maintien des deux chromatides sœurs par la cohésine sont essentiels à la première division de méiose. Les deux forces opposées produites par les CO d'une part et par la tension des kinétochores vers chacun des pôles de la cellule d'autre part entraînent un mouvement des chromosomes homologues permettant leur alignement sur la plaque métaphasique. La libération de la cohésine présente entre les chromatides sœurs et située au-delà des CO ainsi que la résolution des CO conduisent à la migration des chromosomes homologues vers chacun des pôles de la cellule. Une perturbation dans la formation des CO entraîne une aneuploïdie dans la descendance et conduit à la formation de maladie génétique telle que la trisomie 21 ou le syndrome de Turner et de Klinefelter (Hassold et al. 2007).

Ainsi, comprendre comment se forme les CO et comment ils sont régulés est à l'heure actuelle une intense aire de recherche chez toutes les espèces.

2.3 La prophase I : étape clé pour la formation des événements de recombinaison

2.3.1 Deux types d'événements de recombinaison

La recombinaison méiotique est l'un des processus majeurs qui se produit pendant la prophase I. Cette dernière conduit à la formation de deux types d'événements; les crossing-overs (CO) et les non-crossing-overs (NCO) ou événements de conversion génique.

2.3.1.1 Les crossing-overs

Les CO représentent un échange réciproque du brin d'ADN entre deux molécules d'ADN. Ils sont visualisables en microscopie optique sous forme d'une structure appelée chiasma et sont responsables de la ségrégation correcte des chromosomes. Ils sont l'objet de multiples études car ils contribuent aussi à l'évolution des espèces par le biais du brassage génétique et sont simples à identifier. Il existe deux voies de formation des CO. Une voie qui conduit à la formation des CO de classe I qui sont soumis à de l'interférence (définition paragraphe 3.4.1.1b p 65) et une voie qui conduit à la formation des CO de classe II ne subissant pas d'interférence. L'importance entre les deux voies (CO de classe I et CO de classe II) dépend des organismes. Chez la levure *S. pombe*, l'ensemble des CO se forme à partir de la voie de classe II alors que chez *S. cerevisiae* seulement 30 % en dépendent (Mezard et al. 2007). Chez *Arabidopsis thaliana*, la voie II explique 15 % des CO chez un sauvage (Lynn et al. 2007) alors que chez la souris, les CO de classe II représentent 5 à 10 % des CO totaux (Baudat et de Massy 2007).

2.3.1.2 Les non-crossing-overs

Les NCO représentent un transfert unidirectionnel d'un fragment d'ADN d'une chromatide donneuse à une chromatide receveuse. Ils restent encore peu étudiés car ils sont difficiles à visualiser.

En effet, la longueur du fragment transféré varie en fonction des organismes étudiés mais reste très petit. L'observation sur quelques espèces montre que la taille du fragment transféré peut être de l'ordre du kb (kilobase) comme chez la levure (Arnheim et al. 2007) où elle varie de 50 à 1200 pb ou encore chez le maïs où Dooner et Martinez-Ferez (1997) ont observé une taille de 1 à 1.5 kb pour des événements situés au niveau du locus *bronze*.

Cependant, chez d'autres espèces la taille est plus faible comme chez la souris où Paigen et al. (2008) ont observé des tailles variant de 9 à 279 pb ou encore chez *Arabidopsis thaliana* où la moyenne est d'environ 100 pb (Plagnol et al. 2006).

2.3.1.3 Importance des CO et des NCO

L'analyse récente de 52 000 marqueurs sur 204 spores issus d'un hybride de levure est la première étude qui a permis de localiser très finement l'ensemble des CO et des NCO sur un gènome entier (Mancera et al. 2008). Les auteurs déduisent la présence de 66.1 NCO par méiose contre 90.5 CO.

A l'inverse, chez les mammifères ou les plantes supérieures, la présence de NCO semble être beaucoup plus importante que les CO. Le nombre de foyers contenant la protéine Rad51, qui marquent les sites de réparation des cassures double-brin est beaucoup plus important que le nombre de CO observé sur les chromosomes. Le rapport entre le nombre de foyers RAD51 et le nombre de CO est de l'ordre de 3 ou 4 chez l'humain (Baudat et de Massy 2007) et de l'ordre de 25 chez *Arabidopsis thaliana* (Mezard et al. 2007). De plus, chez la plupart des espèces étudiées, le nombre de nodules de recombinaison précoces qui marquent les sites de cassures double-brin excède largement le nombre de CO. Ces observations suggérent que les sites de cassures double-brin non réparés par des CO sont réparés par des NCO et ainsi que ces derniers sont beaucoup plus importants que les CO (Mezard et al. 2007).
La déduction faite par des modèles statistiques à partir des données obtenues sur de larges populations humaines prédit que le nombre de NCO serait 4 à 25 fois plus important que le nombre de CO chez l'humain (Arnheim et al. 2007).

L'étude fine réalisée au niveau de certaines régions où la fréquence de CO est importante montre que la fréquence de NCO l'est aussi mais que le rapport entre les deux varie. Au niveau du locus *β-globine* chez l'humain, le rapport entre la fréquence de CO et la fréquence de NCO est inférieur à 0.08 alors qu'au niveau du locus *NID1* ou *SHOX*, le rapport est de 4 (Arnheim et al. 2007).
L'importance des CO par rapport au NCO dépend de la voie utilisée pour réparer la cassure double-brin induite par la cellule.

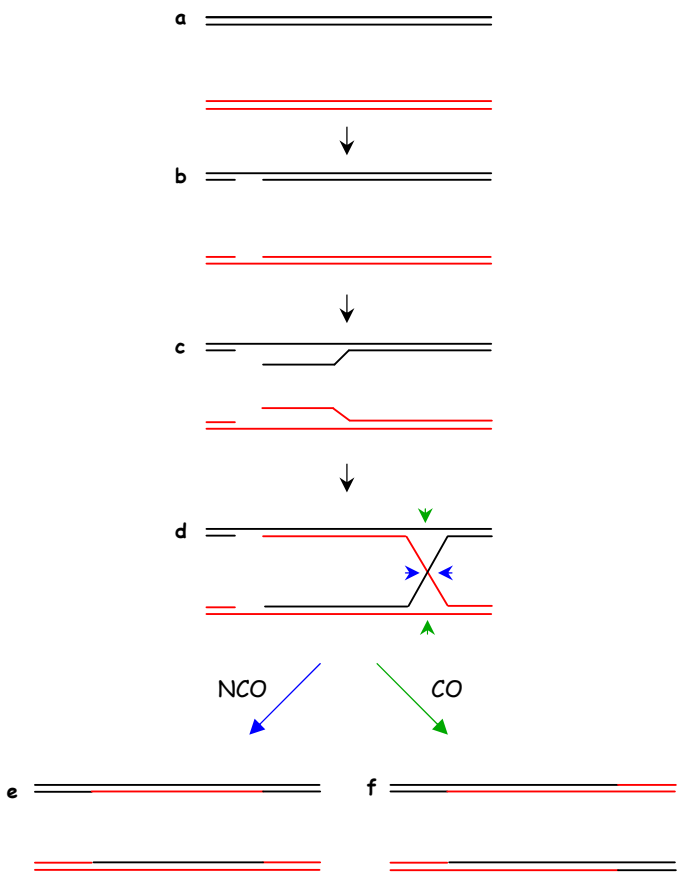


Figure 2 Modèle de recombinaison homologue proposé par R. Hollyday en 1964.

a/ Deux molécules d'ADN homologues double-brin. **b**/ Formation d'une cassure simple-brin au niveau d'un locus sur les deux molécules homologues. **c-d**/ Migration d'une des extrémités de chaque molécule d'ADN vers la molécule homologue et formation de la jonction de Hollyday. Cette jonction peut être résolue soit par la formation d'un NCO (flèches bleues, **e**) soit par la formation d'un CO (flèches vertes, **f**).

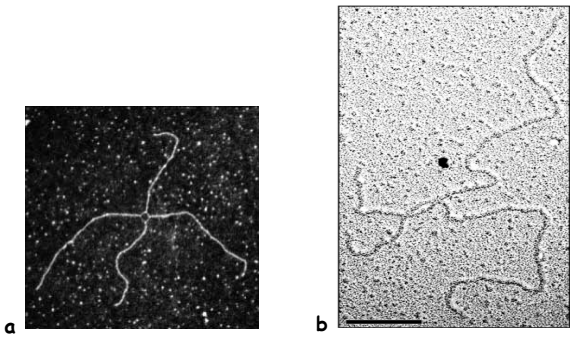


Figure 3 Visualisation de la double jonction de Holliday par microscopie électronique.

a/ Deux plasmides liés par une simple jonction de Hollyday au sein de la bactérie *Escherichia Coli* (Liu et West 2004). **b**/ Deux molécules d'ADN liées par une simple jonction de Hollyday isolées à partir d'une souche de levure de *S. cerevisiae* 4,5 heures après l'induction de la méiose (Cromie et al. 2006). La barre représente 0,2 µm.

2.3.2 Mécanismes moléculaires impliqués dans la formation des événements de recombinaison

2.3.2.1 Introduction

En 1964, Robin Holliday proposa le premier modèle mécanistique pour décrire la recombinaison homologue (Liu et West 2004, Figure 2). Ce modèle établi il y a plus de 40 ans maintenant n'a jamais été remplacé mais a seulement subi certaines modifications suite aux nouvelles découvertes prouvant l'importance de celui-ci.

En 1983, Szostak réactualisa le modèle appelé par la suite : modèle de réparation des cassures double-brin (DSBR pour Double Strand Break Repair) où une double jonction de Holliday se formait (Szostak et al. 1983). Ces jonctions de Holliday furent visualisées par microscopie électronique chez les bactéries (Liu et West 2004) mais aussi chez les levures (Schwacha et Kleckner 1995 ; Cromie et al. 2006 ; Figure 3). De plus, la découverte de protéines capables de catalyser les réactions prédites par ce modèle confirmait clairement l'existence de celui-ci (Haber 1997). En parallèle à ce modèle, de nouveaux virent le jour pour tenter d'expliquer l'ensemble des événements de recombinaison présents au sein des cellules mitotiques et qui n'étaient pas compatibles avec le modèle DSBR. Ainsi, le modèle synthèse dépendante de l'hybridation de brin (SDSA pour Synthesis-Dependant Strand Annealing), le modèle hybridation simple-brin (SSA pour Single Strand Annealing) et le modèle réplication induite par la recombinaison (BIR pour Break Induced Replication) furent élaborés (Pâques et Haber (1999), Figure 31 et Figure 32 en Annexes p 313 et 315).

2.3.2.2 Mécanisme principal de réparation des cassures double-brin lors de la méiose

A l'origine, le modèle DSBR expliquait parfaitement les événements de recombinaison observés à la méiose car en fonction de la résolution des deux jonctions de Holliday un CO ou un NCO pouvait avoir lieu. Cependant, par la suite, il fut montré que ce modèle seul ne pouvait être à l'origine de ces deux types d'événements. En effet, certaines mutations réalisées au sein de gènes impliqués dans le modèle DSBR n'affectaient pas la présence des NCO (Borner et al. 2004).

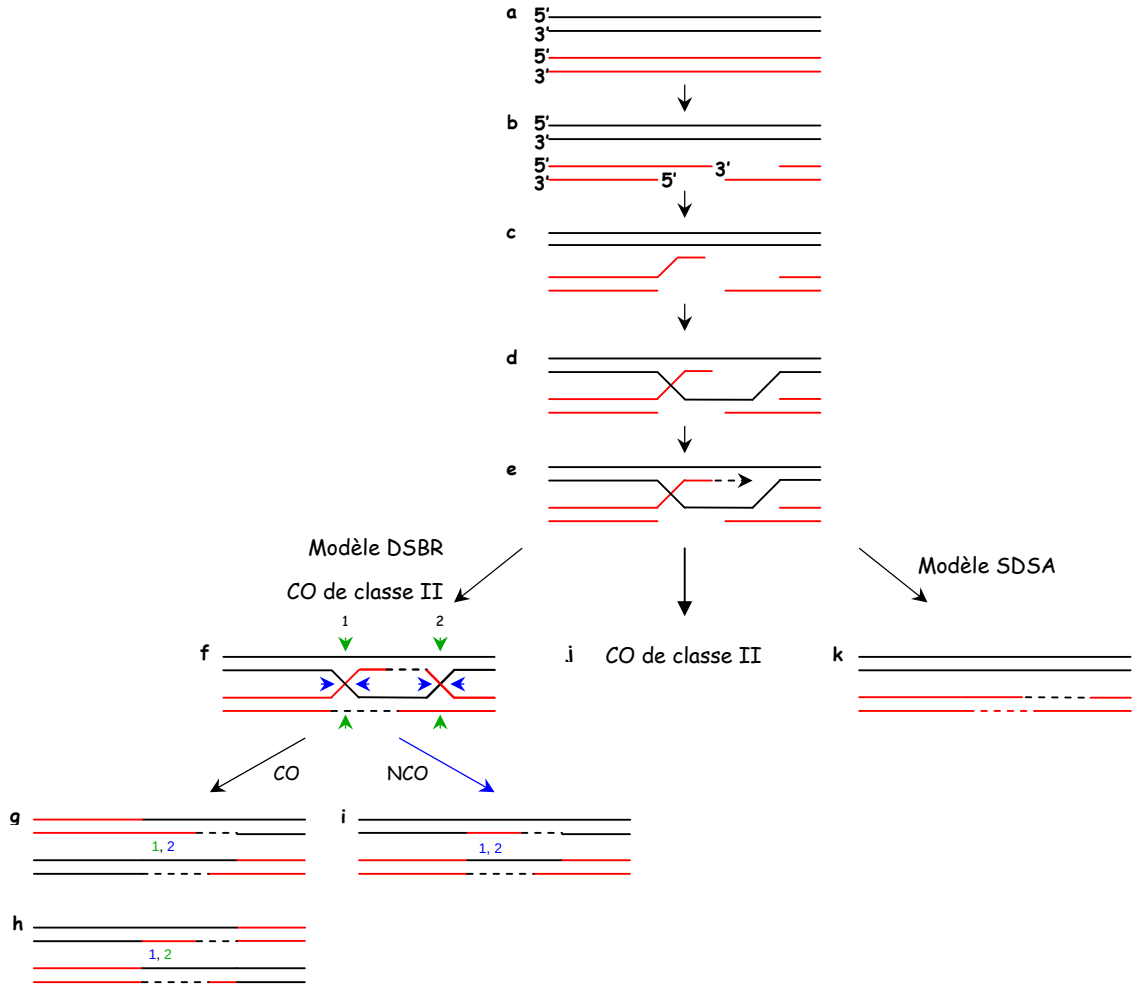


Figure 4 Mécanisme de formation des événements de recombinaison au cours de la méiose.

a/ Deux chromatides homologues, une en rouge et l'autre en noire sont représentées. **b**/ la recombinaison méiotique est initiée par la formation d'une cassure double-brin sur l'une des chromatides. Les extrémités 5' issues de la cassure sont digérées générant ainsi des extrémités 3' saillantes. **c-d**/ L'extrémité 3' saillante envahit la chromatide homologue formant une structure en boucle appelée D-loop. **e**/ La synthèse d'ADN est initiée à partir de l'extrémité 3' envahissante qui sert d'amorce. A ce stade, plusieurs voies peuvent être possibles. **f-g-h-i**/ la voie DSBR. **f**/ La deuxième extrémité 3' saillante issue de la cassure envahit la chromatide homologue et se ligue avec l'ancien brin envahissant, ce qui entraîne la formation d'une double jonction de Holliday (Figure 3 p 22). Suivant le sens de résolution des jonctions, cette voie conduit à la formation de CO de classe I présentant de l'interférence (résolution de la jonction 1 suivant les flèches vertes et de la jonction 2 suivant les flèches bleues (**g**) ou résolution de la jonction 1 suivant les flèches bleues et de la jonction 2 suivant les flèches vertes, **h**) ou à la formation de NCO (dissolution des deux jonctions suivant les flèches bleues, **i**). **j**/ La structure en boucle peut aussi mener à la formation des CO de classe II. **k**/ La voie SDSA. L'extrémité envahissante se retire de la chromatide homologue et retourne s'apparier avec son brin complémentaire. Il y a ensuite synthèse d'ADN pour compléter les trous générés au niveau de la cassure double-brin.

Ainsi, aujourd'hui, la combinaison entre le modèle DSBR et le modèle SDSA semble expliquer parfaitement les événements de recombinaison observés lors de la méiose (Bishop et Zickler 2004 ; McMahill et al. 2007 ; Figure 4). La formation d'une cassure double-brin catalysée par l'enzyme Spo11 est à l'origine du mécanisme de formation des événements de recombinaison (Figure 4). Suite à cela, les deux extrémités 5' générées sont digérées par des exonucléases donnant ainsi naissance à des extrémités 3' saillantes sur leurs brins complémentaires. L'une de ces extrémités saillantes (appelée filament présynaptique) envahit ensuite la deuxième chromatide homologue participant à la recombinaison entraînant ainsi la formation d'une boucle appelée D-loop et d'une jonction de Holliday (Gerton et Hawley 2005; Page et Hawley 2004). Suite à ces premières étapes, il y a migration de la boucle et synthèse d'ADN à partir de l'extrémité 3' envahissante qui sert d'amorce. Cette étape est ensuite le point de divergence entre la formation des CO de classe I, les CO de classe II et le modèle SDSA.

Concernant la formation des CO de classe I, il y a invasion du brin 5' de la deuxième extrémité de la cassure dans la D-loop et ligation de ce brin avec l'ancien brin envahissant. Ceci entraîne la formation de la double jonction de Holliday (Figure 3 et Figure 4). La résolution de cette double jonction conduit soit à la formation d'un CO de classe I soit à la formation d'un NCO (Baudat et de Massy 2007; Cromie et Smith 2007; San Filippo et al. 2008). Les CO de classe II se forment à partir d'une voie différente de la voie des CO de classe I sans formation de la double jonction de Holliday.

Concernant la voie SDSA, la boucle se dissocie et l'extrémité envahissante retourne s'hybrider avec son brin complémentaire d'origine. Il y a ensuite synthèse d'ADN au niveau de la cassure d'origine et ligation. La réparation de la chromatide portant la cassure double-brin par la voie SDSA conduit ainsi à la formation d'un NCO.

L'analyse de gels en deux dimensions a montré que le blocage de la formation des CO a lieu entre la formation de la cassure double-brin et l'invasion de la molécule simple-brin au sein de son homologue chez les mutants des gènes impliqués dans la voie des CO de classe I. Borner et al. (2004) concluent ainsi que la décision du choix entre l'une des voies (CO ou NCO) se fait à ce stade précoce de la méiose. Cependant, comment s'effectue le choix qui détermine le devenir d'une cassure double-brin en CO ou en NCO reste encore sans réponse.

2.3.2.3 Autres voies de réparation des cassures double-brin

Mancera et al. (2008) ont montré que le nombre total de cassures double-brin observées correspondait au nombre de CO et de NCO observés sur les tétrades. Ainsi, chez *S. cerevisiae*, la réparation semble être effectuée selon la voie décrite précédemment (paragraphe 2.3.2.2) en utilisant le chromosome homologue comme matrice de réparation.

À l'inverse des cellules mitotiques où la chromatide sœur est utilisée majoritairement pour la réparation des cassures double-brin, les cellules méiotiques ont mis en place des systèmes pour permettre une réparation en utilisant majoritairement le chromosome homologue. Il existe cependant des exceptions comme la levure *Schyzosaccharomyces pombe* où la majorité des cassures double-brin est résolue en utilisant la chromatide sœur (Cromie et al. 2006). Chez les mutants *dmc1*, un gène codant une protéine responsable de l'invasion du brin saillant, une diminution du nombre de CO est notée chez les deux espèces de levure. Cependant, les cassures double-brin ne sont pas résolues chez *S. cerevisiae* alors qu'elles le sont chez *S. pombe* probablement en utilisant la chromatide sœur (Cromie et Smith 2007). Il existe donc chez *S. cerevisiae* un mécanisme actif qui bloque la formation des réparations via la chromatide sœur.

Les protéines constituantes de l'axe des chromosomes semblent jouer un rôle important dans ce mécanisme de blocage. Chez *S. cerevisiae*, un double mutant *red1 dmc1* est capable de réparer les cassures double-brin grâce à la recombinase Rad51 en utilisant la voie des chromatides sœurs. Ainsi, l'interaction entre la protéine Red1, composant de l'élément axial du complexe synaptonémal et la protéine Dmc1 semble être indispensable pour obtenir le biais en faveur du chromosome homologue (Neale et Keeney 2006). De plus, la protéine Hop1, un autre composant de l'élément axial du complexe synaptonémal (Niu et al. 2005) ainsi que la protéine Hed1, retrouvée à ce jour uniquement chez *S. cerevisiae* (Cromie et Smith 2007), permettraient de bloquer la voie utilisant la chromatide sœur chez cette levure.
Chez les mammifères ou les plantes, la réparation des cassures double-brin au sein des cellules méiotiques se fait aussi en utilisant majoritairement le chromosome homologue comme modèle et rarement des loci non alléliques ou encore la voie de jonction des extrémités non-homologues (NHEJ pour Non-homologous end joining).

Chez *Arabidopsis thaliana* (Couteau et al. 1999 ; Siaud et al. 2004), la mutation du gène *dmc1* entraîne une disparition des CO mais aucune fragmentation des chromosomes comme chez *S. pombe* suggérant une réparation des cassures via la chromatide sœur malgré la présence de la protéine ASY1, l'homologue de Hop1. Très peu de données existent concernant la voie NHEJ opérant au cours de la méiose. Cette voie semble cependant inhibée au sein des cellules méiotiques à l'inverse des cellules mitotiques probablement à cause de la faible expression des protéines Ku intervenant dans cette voie au sein des cellules méiotiques (Goedecke et al. 1999).

2.3.3 Eléments et processus clés en relation avec la formation des événements de recombinaison

2.3.3.1 Un ensemble de processus liés

a Introduction

La recombinaison méiotique est réalisée en association avec plusieurs autres processus cruciaux se déroulant pendant la prophase I : la formation du bouquet télomérique, la mise en place de structures protéiques appelées nodules de recombinaison qui constituent les « usines » responsables de la réparation des cassures, l'appariement des chromosomes homologues et la mise en place du complexe synaptonémal, élément tripartite qui permet une cohésion entre les deux homologues. Les relations entre ces différents processus sont très complexes et varient entre les organismes qui ont fait l'objet d'études approfondies notamment en ce qui concerne l'ordre dans lequel ils se produisent. Par exemple, chez la plupart des organismes, les cassures double-brin et leur maturation sont indispensables pour établir l'appariement et la synapse des chromosomes homologues (Jordan 2006 ; Lichten 2001b). Cependant, chez d'autres comme *C. elegans*, la formation des cassures double-brin a lieu après la formation du complexe synaptonémal. D'autres processus sont alors impliqués dans l'appariement des homologues comme l'existence d'un centre d'appariement situé à une des extrémités de chacun des chromosomes (McKim 2007). De plus, certaines espèces comme la drosophile mâle ne réalisent pas de CO.

Chez la plupart des organismes, tous ces processus fonctionnent cependant main dans la main pour assurer le déroulement correct de la prophase I.

b Le complexe synaptonémal

Le complexe synaptonémal est une structure tripartite schématisée classiquement par une échelle. Il est composé d'un élément axial/latéral qui représente les deux montants et d'éléments transversaux qui jouent le rôle des barreaux (Figure 5 p 34).

Chez la levure, les protéines Red1 et Hop1 sont les deux protéines structurales qui constituent l'élément axial (Page et Hawley 2004). Une mutation dans l'un de ces gènes entraîne une absence de formation du complexe synaptonémal ainsi qu'une diminution de la recombinaison. Les gènes orthologues à *HOP1* ont été isolés chez différentes espèces comme *ASY1* chez *Arabidopsis thaliana* (Caryl et al. 2000), *PAIR2* chez le riz (Nonomura et al. 2004) et *TaASY1* chez le blé (Boden et al. 2007).

Le gène codant la protéine qui constitue les éléments transversaux a également été isolé chez beaucoup d'espèces : *ZIP1* chez la levure, *SCP1* chez les mammifères (Page et Hawley 2004) et *ZYP1a* et *ZYP1b* chez *Arabidopsis thaliana* (Higgins et al. 2005). De manière générale, une mutation dans l'un de ces gènes provoque une diminution du nombre de CO et une perte d'interférence entre CO (De Boer et Heyting 2006).

Le rôle exact du complexe synaptonémal reste encore flou mais il pourrait être important pour la formation des CO interférents (détails paragraphe 3.4.1.1biii, page 69).

c Les nodules de recombinaison

Les nodules de recombinaison (NR) sont des structures protéiques au sein desquelles se déroulent les événements de recombinaison homologue au cours de la méiose (Figure 6 p 50 ; Anderson et Stack 2005). Ils sont en étroite association avec le complexe synaptonémal et ne peuvent être visualisés qu'en microscopie électronique du fait de leur taille, en moyenne 100 nm. Deux types de nodules apparaissent lors de la méiose, ils sont appelés NR précoces et NR tardifs. Ils diffèrent par leur taille, leur nombre, leur composition en protéines et le moment où ils se forment au cours de la méiose.

Les nodules de recombinaison précoces sont plus gros et en moyenne vingt fois plus nombreux que les NR tardifs. Les nodules de recombinaison tardifs représentent les sites de maturation des CO et sont composés des protéines qui interviennent dans la maturation des CO comme les protéines Zip1, Zip2, Zip3, Msh4 et Msh5 chez la levure et la protéine MLH1 chez les mammifères (Page et Hawley 2004). La relation entre les nodules précoces et tardifs reste encore inconnue (Anderson et Stack 2005). Cependant, Moens et al. (2007) ont montré, chez la souris, que certains nodules précoces évoluaient en nodules tardifs.

d Appariement des chromosomes homologues

La recherche du chromosome homologue reste l'un des points les plus intriguant au cours de la méiose. Trois processus majeurs ont été avancés comme étant responsables de l'initiation de l'appariement des chromosomes homologues : des associations pré-méiotiques, des interactions spécifiques ou des contacts aléatoires (Dawe 1998). Des associations pré-méiotiques non-homologues sont retrouvées par exemple au niveau des centromères chez certains organismes comme le blé, *Arabidopsis thaliana* ou encore la levure et seraient importantes pour l'appariement des chromosomes homologues (Jordan 2006).

Le bouquet télomérique (agglomération de l'ensemble des telomères au niveau d'un pôle de la cellule, Figure 5 et Figure 11 p 110) est décrit depuis longtemps comme un mécanisme favorisant l'appariement des chromosomes homologues (Scherthan 2001). Chez le maïs, le mutant *pam1* entraîne une inhibition de la formation du bouquet télomérique, ce qui s'accompagne par une réduction drastique de l'appariement des chromosomes homologues et de la synapse (Hamant et al. 2006). Le rôle prépondérant qui lui est attribué est ainsi de faciliter la proximité entre chromosomes homologues, de les confiner pour promouvoir la synapse ainsi que de les stabiliser quand ils possèdent des cassures double-brin.

Le mécanisme exact de reconnaissance entre chromosomes homologues reste encore une énigme mais au cours de ce celui-ci, les protéines Rad51 et Dmc1 semblent être les acteurs principaux de la recherche d'homologie. Très récemment, Sinha et Peterson (2008) ont montré que le filament présynaptique Rad51 seul était suffisant pour permettre la recherche de l'homologue lors d'une cassure double-brin présente pendant la mitose.

La recherche d'homologie s'effectue naturellement au niveau de la séquence mais aussi au niveau de la morphologie des chromosomes. Des zones d'hétérochromatine ou des zones répétées acquièrent des structures particulières qui aident la recherche de l'homologue (Schwarzacher 2003). De plus, le profil de transcription de chaque chromosome est unique et il détermine donc une structure chromosomique aussi unique qui pourrait aider à la recherche des homologues (Hernandez-Hernandez et al. 2008; Pardo-Manuel de Villena et al. 2000).

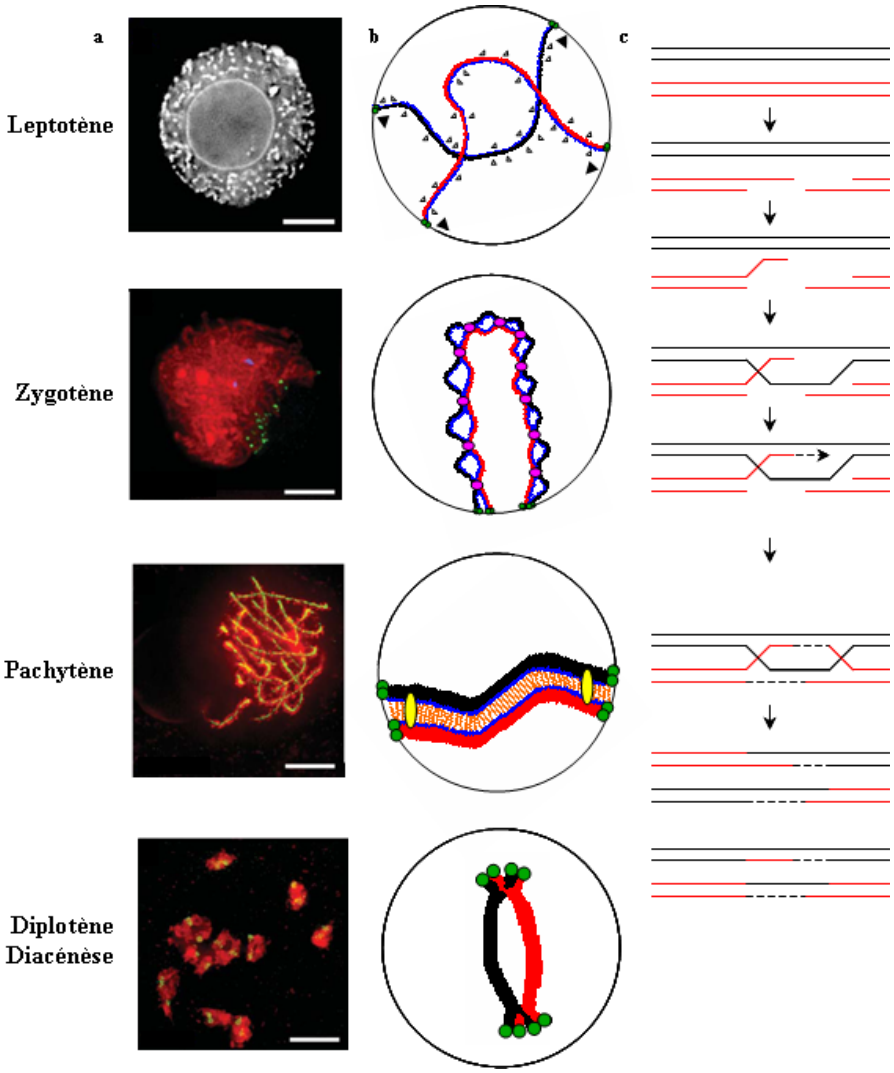


Figure 5 Détails de la prophase de première division de méiose.

a/ Images des chromosomes de maïs pendant les différents stades de prophase I (Hamant et al. 2006). La chromatine est marquée en rouge (DAPI). Au stade zygotène, les télomères sont marqués en vert. Au stade pachytène, le marquage en vert (AFD1) permet de visualiser le complexe synaptonémal présent entre les chromosomes homologues et au stade diplotène, le marquage en vert indique les centromères. Les barres représentent 5 µm. **b**/ Schéma représentant une paire de chromosomes homologues, l'un en rouge l'autre en noir, pendant les différentes étapes de la prophase I. Les télomères sont représentés par les points verts. Au stade leptotène, les télomères migrent (flèches) vers un des pôles de la cellule et les cassures double-brin sont réalisées par la protéine Spo11 (petits triangles). L'élément axial apparaît le long des chromosomes (structure en bleu). Au stade zygotène, les télomères sont agglomérés, c'est le bouquet télomérique. Les nodules de recombinaison précoces (points roses) apparaissent au niveau des cassures double-brin et des sites d'appariement entre chromosomes homologues. Au stade pachytène, le complexe synaptonémal est mis en place (les éléments transversaux sont représentés par les traits orange), les nodules de recombinaison tardifs (ovales jaunes) apparaîssent sur les sites de formation des CO. Au stade diplotène, le complexe synaptonémal disparaît, les CO sont résolus et les télomères se décrochent de l'enveloppe nucléaire. **c**/ Schéma représentant l'évolution de la réparation de la cassure double-brin au cours des différentes phases.

2.3.3.2 Déroulement de la prophase I

La prophase I peut se découper en cinq stades définis selon la morphologie des chromosomes au cours de laquelle l'ensemble des processus décrits précédemment se déroule : le leptotène, le zygotène, le pachytène, le diplotène et la diacinèse (Figure 5 ; Anderson et Stack 2005 ; Hamant et al. 2006 ; Jordan 2006 ; Lynn et al. 2007 ; Mezard et al. 2007 ; Page et Hawley 2004 ; Petronczki et al. 2003 ; Sanchez-Moran et al. 2008 ; Shinohara et al. 2008).

a Leptotène

A ce stade, les chromosomes commencent à se condenser. Ils sont visibles dans le noyau sous forme de longs filaments fins et l'élément axial apparaît le long de chacun d'entre eux. Les télomères s'attachent à l'enveloppe nucléaire et commencent leur migration vers l'un des pôles du noyau. De plus, les premiers stades de recombinaison méiotique s'initient. Ainsi, la protéine Spo11 induit la création des cassures double-brin et dès la fin du leptotène, l'extrémité 3' sortante résultante de la maturation des cassures double-brin commence à interagir avec le chromosome homologue et forme des ponts entre chromosomes homologues (Gerton et Hawley 2005; Page et Hawley 2004).

b Zygotène

Les chromosomes continuent à se condenser et commencent à s'aligner (appelés pour cela bivalents). Certains des ponts entre homologues évoluent en association axiale (AA) et permettent une connexion entre les éléments axiaux des deux homologues. La migration des télomères vers un des pôles de la cellule est terminée : c'est le bouquet télomérique. Les sites de cassure double-brin sont le lieu d'une intense activité protéique qui les mature en CO et/ou en NCO. De plus, à ce stade les nodules de recombinaison précoces apparaissent au niveau des sites de formation du complexe synaptonémal.

c Pachytène

Le complexe synaptonémal est entièrement installé et permet un contact rapproché entre les chromosomes homologues sur toute leur longueur. Il y a dissolution du bouquet et les télomères se dispersent à nouveau le long de l'enveloppe nucléaire. Au niveau des sites de recombinaison, la double jonction de Hollyday se forme et à la fin du stade pachytène la plupart de ces doubles jonctions est résolue. Ces sites sont marqués par l'apparition des nodules de recombinaison tardifs au sein du complexe synaptonémal.

Gènes	Fonctions
SPO11	Formation de la cassure double-brin
PRD1	Initiation de la recombinaison
MRE11	Maturation des cassures double-brin
RAD50	Maturation des cassures double-brin
NBS1/XRS2	Maturation des cassures double-brin
RAD51	Invasion de la molécule d'ADN simple-brin
DMC1	Invasion de la molécule d'ADN simple-brin
MSH4	Formation des CO de classe I
MSH5	Formation des CO de classe I
MER3	Formation des CO de classe I
MLH1	Formation des CO de classe I
MLH3	Formation des CO de classe I
SHOC1	Formation des CO de classe I
MUS81	Formation des CO de classe II
MMS4	Formation des CO de classe II

Tableau 1 Principaux gènes impliqués dans la recombinaison méiotique.

d Diplotène

La condensation des chromosomes devient de plus en plus importante. Le complexe synaptonémal se désagrège et les chromosomes homologues ne sont plus liés entre eux que par les chiasmas, qui sont la manifestation visuelle des CO. Ceux-ci restent présents jusqu'à la séparation en anaphase I.

e Diacinèse

Il y a dissociation de l'enveloppe nucléaire qui est dégradée sous forme de vésicules et le fuseau de microtubules commence à se former. Les chromosomes se condensent davantage et les sites de CO s'enchevêtrent, se superposent, rendant les chiasmas bien visibles. C'est le premier stade de la méiose où les quatre parties des tétrades commencent à être visibles.

2.4 Gènes et protéines impliqués dans la recombinaison homologue

2.4.1 Introduction

Les premières études concernant les gènes impliqués dans la recombinaison homologue ont été réalisées chez la levure *Saccharomyces cerevisiae* et concernent la recherche de mutants sensibles aux agents responsables de cassures de l'ADN. Les gènes impliqués dans ces phénotypes ont été isolés et appelés gènes RAD (Symington 2002). L'ensemble de ces gènes constituent la famille épistatique RAD52 qui est constituée à ce jour de *RAD50*, *RAD51*, *RAD52*, *RAD54*, *RAD55*, *RAD57*, *RAD59*, *MRE11*, *XRS2* et *RDH54/TID1*.

Les protéines codées par ces gènes interviennent aussi bien dans la recombinaison mitotique que dans la recombinaison méiotique. Bizarrement, aucun mutant de la voie SDSA n'a encore été découvert (Baudat et de Massy 2007). La plupart de ces gènes sont aussi retrouvés chez d'autres eucaryotes soulignant le caractère universel du mécanisme de recombinaison homologue (Mercier et Grelon 2008). Une caractérisation fine du phénotype des différents mutants, des propriétés biochimiques *in vitro* de chacune des protéines ainsi que la caractérisation des interactions existantes entre elles ont permis d'avoir une vision globale et détaillée du rôle de chacune dans la recombinaison homologue (Tableau 1).

2.4.2 Formation des cassures double-brin

Chez la levure, la cassure double-brin induite dès le début de la prophase est réalisée par la protéine Spo11 (Keeney 2001; Keeney et Neale 2006). Les mutants *spo11* ne présentent pas de cassure double-brin, pas ou peu d'événements de recombinaison, une perturbation voire une absence complète de formation du complexe synaptonémal (dépendant des organismes) et une disparition des foci Rad51 (Keeney 2001; Lichten 2001b). De plus, chez la levure, Bulher et al. (2007) ont montré que chez les mutants *spo11*, aucune molécule d'ADN simple-brin résultant de la digestion d'un des brins au niveau de la cassure n'était retrouvée démontrant ainsi l'action unique de la protéine Spo11.

Le gène codant la protéine Spo11 a été retrouvé chez tous les autres organismes étudiés. Chez la souris, l'étude des protéines H2AX phosphorylées, un type d'histone particulier qui marque les sites de cassure de l'ADN (Mahadevaiah et al. 2001) et l'étude des foci RAD51 et DMC1 (Romanienko et Camerini-Otero 2000) montrent un nombre important de foyers pendant le leptotène. Cependant, ces foyers disparaissent chez le mutant *spo11* (Lichten 2001b) indiquant que cette protéine semble aussi responsable des cassures double-brin évoluant en CO et NCO chez les organismes supérieurs. Cependant, Chicheportiche et al. (2007) ont observé la présence de deux types de foyers H2AX chez la souris dont l'un d'entre eux ne disparaîssant pas chez les mutants *spo11*. Ils ont suggéraient ainsi que certaines cassures pourraient être indépendantes de la protéine SPO11 et qu'il pourrait donc exister une deuxième voie pour initier la recombinaison méiotique.
Chez *Arabidopsis thaliana*, trois homologues de *SPO11* sont présents : *SPO11-1*, *SPO11-2* et *SPO11-3*. La mutation de *spo11-1* entraîne une chute de 90 % du taux de CO (Grelon et al. 2001). Ainsi, *SPO11-1* est impliqué dans la formation des événements de recombinaison (Hartung et al. 2007) au même titre que *SPO11-2* alors que *SPO11-3* participe à l'endoréduplication somatique (Hartung et al. 2002).
Robine et al. (2007) ont montré, chez *S. cerevisiae*, que malgré la fixation de la protéine Spo11 au niveau de certaines régions, celles-ci restaient dépourvues de cassures double-brin. Ainsi, au moins neuf autres protéines sont nécessaires, en plus de la protéine Spo11, pour induire la cassure double-brin chez la levure : Mer2, Mei4, Ski8/Rec103, Rec102, Rec104, Rec114, Rad50, Mre11 et Xrs2 (Keeney et Neale 2006). La relation qui existe entre ces différentes protéines reste encore floue (Keeney et Neale 2006).

De plus, l'ensemble des gènes codant ces protéines ne semble pas être conservé entre espèces comme l'est le gène *SPO11*. Cependant, d'autres gènes ont été découverts chez les espèces autres que la levure comme récemment, le gène codant une protéine nommée AtPRD1 qui a été isolée chez *Arabidopsis* et qui serait un des partenaires de AtSPO11 pour induire la formation des cassures double-brin (De Muyt et al. 2007).

Ainsi, la protéine Spo11 est la protéine universelle responsable des cassures double-brin. Suite à la création de la cassure, la protéine reste attachée de manière covalente à l'extrémité 5' générée. Elle sera ensuite décrochée lors de la maturation des cassures double-brin.

2.4.3 Maturation des cassures double-brin

Les protéines Mre11, Rad50 et Xrs2/Nbs1 forment un complexe appelé complexe MRX (Xrs2) chez la levure ou complexe MRN (Nbs1) chez les vertébrés et les plantes. Le gène *MRE11* code pour une protéine possédant des motifs d'accrochage à l'ADN, une activité endonucléase simple-brin et une activité exonucléase double-brin 3'/5' (Borde 2007).

Chez la levure, la mutation de l'un des gènes codant une protéine du complexe provoque un ralentissement de la croissance, une sensibilité accrue aux infrarouges et une méiose perturbée (Borde 2007; Paques et Haber 1999). Chez les mammifères, ces gènes sont indispensables pour leur viabilité (D'amours et Jackson 2002) et chez *Arabidopsis*, les mutants *mre11* montrent une fragmentation des chromosomes à la méiose suite à la non résolution des cassures double-brin induites par la protéine SPO11 (Borde 2007).

Chez l'ensemble des organismes où son étude a pu être réalisée, ce complexe joue un rôle majeur dans la libération de la protéine Spo11 qui est liée de manière covalente aux extrémités de la cassure double-brin et permet ainsi une progression de la réparation de la cassure. Ce complexe intervient aussi dans la digestion des extrémités générées lors de cette cassure. De plus, il possède un rôle dans l'activation des points de contrôle qui subsistent tout au long de la méiose (Borde 2007).

Une étude récente chez la levure suggère que la digestion des extrémités de la cassure est réalisée en deux étapes (Mimitou et Symington 2008). La première étape impliquerait le complexe MRX/MRN et la protéine Sae2 qui entraîneraient la digestion de quelques nucléotides. La seconde ferait intervenir les protéines Sgs1 et Exo1 pour permettre une extension plus importante des extrémités simple-brin.

2.4.4 Invasion de l'extrémité simple-brin

2.4.4.1 Les protéines Rad51 et Dmc1

Les protéines Rad51 et Dmc1 sont les homologues de RecA, la protéine responsable de l'invasion du brin homologue chez les bactéries. Elles sont ainsi appelées « recombinases » et permettent l'invasion de la molécule d'ADN simple-brin dans son homologue (Baumann et al. 1996; Sehorn et al. 2004). Ces protéines et notamment Rad51 ont été retrouvées chez tous les organismes dont le gènome est séquencé ou présentant d'importantes bases de données d'EST (Expressed Sequence Tag) et par conséquent ont été trés conservées au cours de l'évolution. Cinq paralogues sont présents chez les vertébrés : *XRCC2*, *XRCC3*, *RAD51B*, *RAD51C* et *RAD51D* et une nouvelle variante, *RAD51-Δex9*, dépourvue de l'exon 9, vient d'être récemment isolée chez l'humain (Park et al. 2008). Ces protéines s'associent sous forme de filaments hélicoïdaux entourant la molécule d'ADN (Sheridan et al. 2008), ce qui forme des foci visibles pendant les premiers stades méiotiques.

La protéine Rad51 est impliquée dans la recombinaison mitotique ainsi que méiotique et la mutation de ce gène entraîne des conséquences différentes en fonction des organismes. Une mutation de *rad51* chez la levure est viable mais elle entraîne une sensibilité accrue aux infrarouges, un défaut dans la formation des spores (Shinohara et al. 1997; Symington 2002). Chez les vertébrés, les mutants *rad51* meurent dès le stade embryon (Symington 2002) alors que chez *Arabidopsis*, les mutants sont stériles et montrent une inhibition de l'appariement et de la synapse entre chromosomes homologues ainsi qu'une fragmentation des chromosomes (Li et al. 2004a). Chez le maïs, les mutants *rad51* montrent un phénotype moins affecté (Li et al. 2007). Les mâles sont stériles et montrent une réduction dans l'appariement des chromosomes homologues de 44 % par rapport au sauvage, une augmentation dans l'appariement et la formation de synapses entre chromosomes non homologues. Chez la femelle, la méiose est moins perturbée, le taux de CO analysés ne diffère pas du taux retrouvé chez le sauvage.

Les mutants *dmc1* diffèrent de ceux de *rad51*. Chez la levure, la recombinaison méiotique est complètement bloquée chez les mutants *dmc1* et une augmentation de la recombinaison ectopique est également observée (Neale et Keeney 2006). Une mutation du gène *dmc1* entraîne la formation d'univalents à la méiose chez *Arabidopsis* (Anderson et Stack 2002). De plus, aucune cassure double-brin n'est observée au sein de l'ADN, suggèrant que les cassures induites par la protéine Spo11 sont réparées en utilisant la chromatide soeur.

Pour expliquer les différences de phénotype que génèrent *rad51* et *dmc1* lorsqu'ils sont mutés, Neale et Keeney (2006) ont émis l'hypothèse que chacune des protéines correspondantes pourrait se fixer spécifiquement sur une des deux extrémités saillantes de la cassure double-brin, ce qui expliquerait aussi le comportement différent des deux extrémités, l'une envahissant la molécule homologue dés la début du mécanisme alors que l'invasion de la seconde s'effectue dans un deuxième temps.

L'interaction entre Rad51 et/ou Dmc1 avec le filament présynaptique, filament qui se charge en recombinases et qui va envahir la molécule d'ADN homologue, ne peut se faire de manière isolée. D'autres facteurs, appelés médiateurs, sont importants pour le chargement et la régulation de ces protéines sur le filament présynaptique.

2.4.4.2 Les médiateurs de Rad51 et Dmc1

La protéine Rpa est une protéine qui se fixe sur les molécules d'ADN simple-brin. Sa présence sur l'extrémité saillante de la cassure double-brin favorise l'entrée des recombinases en dénouant les structures secondaires. Cependant, elle peut aussi empêcher leur accrochage si elle est présente en quantité trop importante (San Filippo et al. 2008).

L'ensemble des protéines du groupe épistatique Rad52 est aussi impliqué dans le chargement de la protéine Rad51 sur le filament présynaptique. Les foci de Rad51 disparaîssent chez les mutants *rad52*, *rad55* et *rad57* (San Filippo et al. 2008). Chez la levure, la protéine Rad52 est l'un des médiateurs les plus importants alors que chez l'humain, elle ne possède pas d'activité médiatrice (San Filippo et al. 2008).

Chez les organismes autres que la levure, d'autres protéines ou complexes protéiques ont été identifiés en tant que médiateurs comme le gène *TEX15* chez la souris (Yang et al. 2008), le gène *PHS1* chez le maïs (Pawlowski et al. 2004), le gène *BRCA2* chez l'humain (San Filippo et al. 2008) et les complexes formés entre les différents paralogues de *RAD51* chez les vertébrés (San Filippo et al. 2008).

Le gène *BRCA2*, absent chez les levures, est un gène très étudié chez l'homme car sa mutation peut conduire à la formation de cancers du sein et des ovaires mais aussi au syndrome de Fanconi, maladie génétique rare (San Filippo et al. 2008). Ce mutant *brca2* montre une sensibilité accrue aux agents ionisants et une disparition des foci RAD51. Chez *Arabidopsis*, un mutant RNAi *brca2* montre une fragmentation des chromosomes due à l'absence de réparation des cassures induites par la protéine SPO11 et entraîne la présence de nombreux univalents lors de la méiose (Siaud et al. 2004). Cette protéine BRCA2 qui interagit physiquement avec les protéines RAD51 et DMC1 est donc impliquée chez l'humain et chez *Arabidopsis* dans le chargement des recombinases sur le filament présynaptique (San Filippo et al. 2008).

2.4.5 Maturation des jonctions de Holliday

Les protéines qui interviennent dans les dernières étapes de la maturation des CO sont encore mal connues. Cependant, certaines protéines comme Rad54 ou les protéines ZMM (Zip, Msh, Mer) ont des rôles multivalents et sont impliquées dans ces dernières étapes.

2.4.5.1 La protéine Rad54

La protéine Rad54 est un membre de la famille des protéines de type Swi2/Snf2 qui interviennent dans le « remodelage » de la chromatine en hydrolysant l'ADN (Heyer et al. 2006). Chez la levure *S. cerevisiae* et chez *Arabidopsis*, les mutants *rad54* sont sensibles aux agents ionisants et engendrent la production d'une descendance viable avec cependant une fréquence moins importante que chez le sauvage. Il existe aussi des paralogues de *RAD54* chez la levure (*RDH4/TID1*) et chez les vertébrés (*RAD54B*). Une mutation au sein de ces paralogues entraîne également une réduction de la fertilité chez la levure (Heyer et al. 2006).
La protéine Rad54 a un rôle multifactoriel. Elle est impliquée en association avec les protéines Rad51 et Dmc1 dans la recherche d'homologie, dans la modification de la structure de la chromatine et dans la formation de la boucle au cours du mécanisme de recombinaison. Elle permet aussi le décrochage de Rad51 du filament présynaptique et la migration de la jonction de Holliday (Heyer et al. 2006).

2.4.5.2 Protéines de la voie interférente

La voie qui conduit à la formation des CO de classe I présentant de l'interférence est sous la dépendance des protéines ZMM (Zip, Msh, Mer). Elles peuvent être subdivisées en trois sous-groupes selon leur fonction (Lynn et al. 2007).

Le groupe I est formé par les protéines Mer3, Msh4 et Msh5. La protéine Mer3 est une hélicase de type 5'_ 3'. Les protéines Msh4 et Msh5 sont des homologues des protéines MutS de réparation des appariements incorrectes chez les bactéries. Ces protéines sont impliquées dans la stabilisation du brin saillant issu de la cassure double-brin envahissant l'homologue. Le groupe II est composé des protéines Zip2, Zip3 et Zip4/Spo22. Le groupe III est composé uniquement de la protéine Zip1 qui est le composant majeur de l'élément transversal du complexe synaptonémal.

En général, une mutation dans l'un de ces gènes entraîne un problème dans la formation du complexe synaptonémal et une diminution voire une absence de CO (Borner et al. 2004). La diminution du taux de CO varie en fonction du gène muté, de l'espèce, de la souche de levure mais aussi des conditions d'études. L'analyse de différentes régions, dans des souches de levure et dans des conditions différentes, montre qu'en moyenne le nombre de CO est réduit de 80 % lors d'une mutation dans un des gènes codant une protéine ZMM alors que le nombre de NCO est conservé (Lynn et al. 2007). Les CO résiduels chez les mutants ne présentent pas d'interférence.

Cependant, des CO résiduels présentant de l'interférence sont retrouvés chez les mutants *spo16* et *spo22* (Shinohara et al. 2008). Le comportement méiotique observé chez ces mutants suggère que les deux protéines Spo16 et Spo22 ne seraient pas nécessaires pour décider du devenir d'une cassure double-brin en CO et pour la mise en place du phénomène d'interférence. Toutefois, elles seraient indispensables pour assurer la formation des CO.

Tout récemment, une nouvelle protéine SHOC1 a été isolée chez *Arabidopsis* (Macaisne et al. 2008). Les mutants *shoc1* présentent le même phénotype que les mutants des gènes codant les protéines ZMM. Cette protéine serait ainsi impliquée dans la maturation des jonctions de Holliday (Macaisne et al. 2008).

Les protéines homologues aux protéines MutL retrouvées chez les bactéries sont aussi impliquées dans la voie interférente. Les protéines MLH1 et MLH3 (pour MutL Homologue) marquent très clairement les CO et sont ainsi utilisées pour l'étude de la distribution des CO sur les chromosomes (voir paragraphe 3.1.1, p 53). Une mutation dans l'un de ces gènes entraîne une diminution importante du nombre de CO chez *Arabidopsis* et chez la souris (Baudat et de Massy 2007; Jackson et al. 2006).

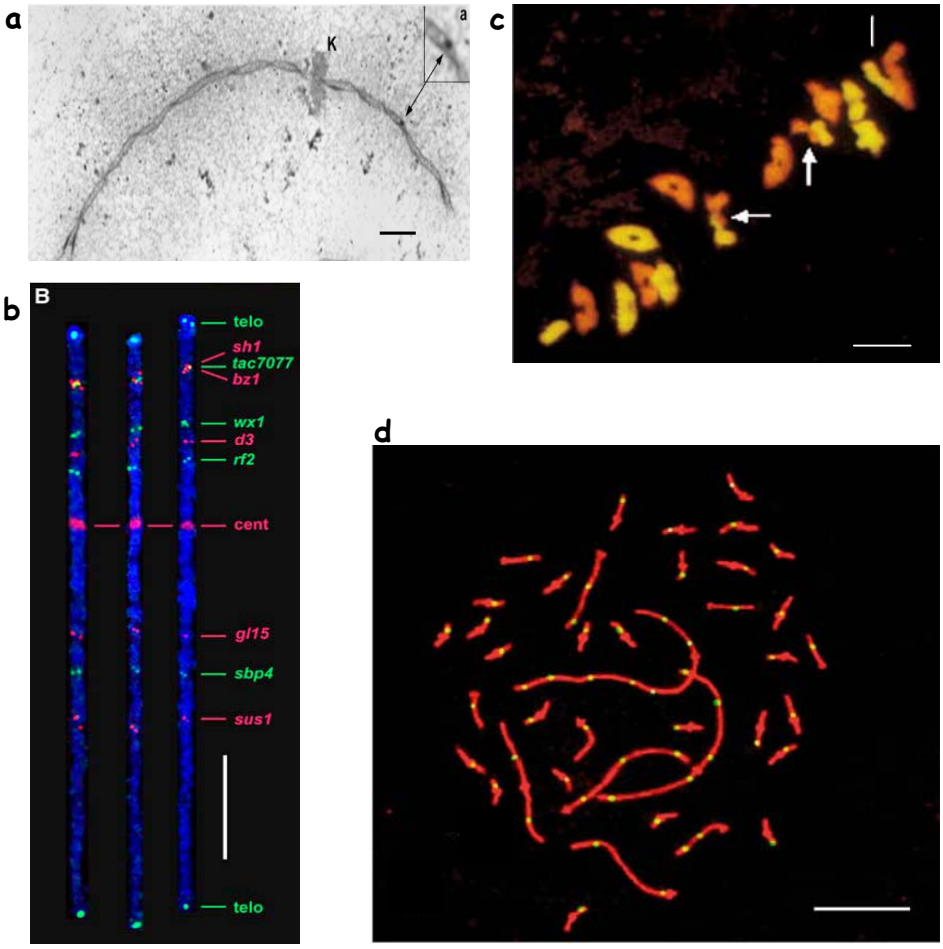


Figure 6 Différentes techniques de cytogénétique utilisées pour étudier la distribution des CO.

a/ Observation en microscopie électronique des nodules de recombinaison tardifs sur le complexe synaptonémal du chromosome 9 de maïs en début de stade diplotène (Anderson et al. 2003). La fenêtre dans le coin en haut à gauche représente un grossissement du nodule. La barre représente 2 µm. **b**/ Visualisation de l'hybridation de neuf sondes FISH sur le chromosome 9 du maïs au stade pachytène pour déterminer la distance physique entre chacune d'entre elles (Wang et al. 2006). La barre représente 10 µm. **c**/ Métaphase I d'un individu F1 issu d'un croisement entre la fétuque et le ray-grass (Zwierzykowski et al. 2008). Les chromosomes sont marqués par la technique GISH en utilisant le génome de la fétuque (jaune) et le génome du ray-grass (orange) comme sonde. Les flèches marquent les sites de formation des CO entre un chromosome de la fétuque et un chromosome du ray-grass. La barre représente 10 µm. **d**/ Foci MLH1 (points jaunes) observés sur le complexe synaptonémal (protéine SCP3, rouge) des chromosomes de caille du japon au stade pachytène (Calderon et Pigozzi 2006). La barre représente 10 µm.

2.4.5.3 Protéines de la voie non-interférente

La formation des CO de classe II nécessite la présence des protéines Mus81 et Mms4 (de los Santos et al. 2003; Lynn et al. 2007). Chez la levure, la mutation du gène *mms4* diminue fortement la viabilité des spores mais entraîne également un fort taux d'échange de matériel entre génotypes parentaux (Mancera et al. 2008). Par exemple, sur une région de 63 kb, Mancera et al. (2008) ont observé 31 événements d'échange entre les deux génotypes parentaux.

La mutation du gène *mus81* a des conséquences moins importantes puisque le phénotype est très voisin de celui du sauvage chez *Arabidopsis* et chez la souris (Higgins et al. 2008; Holloway et al. 2008). De plus, chez les mutants *mus81* ou *mms4*, les CO restants présentent de l'interférence (Berchowitz et al. 2007; Lynn et al. 2007). Ces protéines Mus81 et Mms4 catalysent la résolution des jonctions de Holliday *in vivo* chez la levure *S. cerevisiae* (Jessop et Lichten 2008).

3 Distribution et régulation des CO sur les chromosomes

3.1 Techniques d'étude de la distribution des événements de recombinaison

A l'heure actuelle, la disponibilité des outils de génétique et ou de génomique conditionne l'approche utilisée pour l'étude de la distribution des événements de recombinaison mais aussi la résolution obtenue. L'étude de la distribution de ces événements sur les chromosomes peut se résumer à étudier le rapport entre un taux d'événements (CO ou NCO) et une distance physique. Ainsi, deux grands angles d'attaque ont été utilisés : (1) les techniques de cytogénétique qui permettent une comparaison directe entre distances physiques et distances génétiques car elles sont observables au cours de la même expérience ; (2) la comparaison de distances génétiques issues de l'analyse de données de génotypage obtenues à partir de populations avec des distances physiques.

3.1.1 Techniques de cytogénétique

Les techniques de cytogénétique ont permis de déduire la position des CO à travers leur observation directe sur les chromosomes ou l'observation d'intermédiaires intervenant dans la formation des CO.

L'utilisation du marquage des bandes C (ou C-banding) mais aussi la localisation des chiasmas sur les chromosomes ont permis des études grossières de la distribution des CO chez un grand nombre d'espèces (Anderson et Stack 2002; Lukaszewski et Curtis 1993).

L'hybridation génomique *in situ* (GISH pour Genomic *in situ* Hybridization) a également été utilisée, notamment pour caractériser l'appariement et la formation de CO entre les chromosomes de deux espèces éloignées (Fernandez-Calvin et al. 1995; Ji et Chetelat 2007; Wang et al. 2005b; Zwierzykowski et al. 2008 ; Figure 6).

Plus récemment, les techniques de microscopie électronique ont permis la cartographie des nodules tardifs de recombinaison chez la tomate et le maïs (Anderson et al. 2003 ; Sherman et Stack 1995; Figure 6). De même, les protéines intervenant dans le mécanisme de recombinaison méiotique ont pu être cartographiées par le biais d'anticorps spécifiques. La protéine MLH1 (Figure 6) a été la plus utilisée notamment chez la souris (Froenicke et al. 2002), chez l'oiseau (Calderon et Pigozzi 2006), chez le poisson (Kochakpour et Moens 2008; Moens 2006), chez l'humain (Sun et al. 2006), chez la tomate (Lhuissier et al. 2007) et chez le chien (Basheva et al. 2008).

La distribution des nodules de recombinaison tardifs et des foyers définis par la protéine MLH1 montre une très bonne corrélation avec la distribution des chiasmas (Anderson et al. 2003; Pigozzi et Solari 1999; Sherman et Stack 1995). Ces techniques sont ainsi très intéressantes pour visualiser et estimer le nombre de CO sur les chromosomes.

Cependant, l'ensemble de ces techniques reste très lourd à mettre en place et n'est pas approprié pour l'étude de plusieurs centaines d'individus. De plus, comme l'estimation de la distance physique se fait à partir de l'observation de lames, elle correspond ainsi à des distances en µm. Celles-ci ne prennent pas en compte les différences de compaction de l'ADN le long des chromosomes et en plus elles restent des estimations grossières. Ces techniques ont principalement été utilisées chez des espèces où la visualisation des chromosomes en microscopie est facile et où très peu d'outils de génomique existent. Néanmoins, elles ont permis de comprendre où et comment se localisent les protéines qui interviennent au cours de la recombinaison méiotique et également elles ont permis de déterminer l'effet de certaines mutations sur la formation des CO.

3.1.2 Comparaison entre cartes génétiques et physiques

Les cartes génétiques réalisées depuis plusieurs décennies reflètent la distribution des CO entre les différents marqueurs utilisés. La résolution de ces cartes dépend du nombre de marqueurs et du nombre d'individus utilisés. Chez les plantes, le développement de larges populations, comprenant plusieurs milliers d'individus issus des deux mêmes parents, est largement utilisé à l'inverse de chez les mammifères où ceci n'est pas réalisable.
Ainsi, chez ces derniers, des populations regroupant des milliers individus provenant de différentes familles ont été utilisées. Ces individus très différents les uns des autres ont accumulé chacun un nombre important de CO comparé au génome de leur ancêtre commun, ce qui permet une augmentation considérable de la résolution (Arnheim et al. 2007).
Les marqueurs présents sur ces cartes génétiques peuvent être placés sur les cartes physiques pour ainsi localiser les CO sur les chromosomes. La carte physique idéale est la séquence du génome car la distance physique entre chaque marqueur est connue à la base près permettant une localisation fine des CO. Des études de cette précision ont ainsi pu être réalisées chez la levure (Mancera et al. 2008), l'humain (Myers et al. 2005), la souris (Paigen et al. 2008) et *Arabidopsis* (Drouaud et al. 2006).

Cependant, chez beaucoup d'espèces la séquence du génome n'est pas encore disponible. Des approches préliminaires ont néanmoins été initiées pour déterminer la position des marqueurs sur les chromosomes à l'aide de matériel particulier comme des lignées d'introgression (ray grass et fétuque, King et al. 2002), des lignées transloquées (orge, Kunzel et al. 2000), des hybrides de radiation à partir de lignées d'addition maïs-avoine (Riera-Lizarazu et al. 2000) ou encore des lignées aneuploïdes (blé, Endo et Gill 1996, détails paragraphe 4.3.1 page 113).
Les techniques de cytogénétique ont également été utilisées pour cartographier les marqueurs sur les chromosomes. Par exemple, l'hybridation *in situ* fluorescente (FISH pour Fluorescent *in situ* Hybridization) est employée pour localiser les marqueurs des cartes génétiques sur les chromosomes et estimer les distances physiques les séparant grâce à l'utilisation de ces marqueurs en tant que sonde FISH (Figure 6). Cette technique a notamment été utilisée chez le sorgho (Kim et al. 2005), *Medicago truncatula* (Kulikova et al. 2001), le maïs (Wang et al. 2006), la tomate (Koo et al. 2008), la pomme de terre (Iovene et al. 2008) et le riz (Cheng et al. 2001; Kao et al. 2006).

Espèces	cM	Références	Mb	cM/Mb	Caryotype
Arabidopsis thaliana		(Gaut et al. 2007)[a]	157	4.8	2n=2x=10
Arachis sp	1230	(Moretzsohn et al. 2005)	1243-2989	0.58	2n=2x=20
Allium cepa	1907	(Martin et al. 2005)	16415	0.12	2n=2x=16
Avena sativa	2932	(O'Donoughue et al. 1995)	12961	0.23	2n=6x=42
Beta vulgaris	526, 815[b]	(Laurent et al. 2007)	758	0.73	2n=2x=18
Brassica napus	1605	(Sun et al. 2007)	1127	1.42	2n=38
Brassica oleracea	703	(Gao et al. 2007)	760	0.93	2n=2x=18
Brassica rapa	1378	(Choi et al. 2007)	784	1.75	2n=2x=20
Capsicum annuum	1857	(Barchi et al. 2007)	3920	0.47	2n=2x=24
Carica papaya	1069	(Chen et al. 2007)	368	2.91	2n=2x=18
Cocos nucifera	1971	(Lebrun et al. 2001)	3479	0.57	2n=2x=32
Coffea arabica	1802	(Pearl et al. 2004)	1176	1.53	2n=4x=44
Cucumis melo	1197	(Oliver et al. 2001)	931	1.29	2n=24
Cucumis sativus	706-1164[b]	(Fazio et al. 2003; Wang et al. 2005a)	882	1.06	2n=2x=14
Elaeis guineensis	1743	(Billotte et al. 2005)	980	1.78	2n=2x=32
Eucalyptus sp	1568	(Brondani et al. 2006)	564	2.78	2n=2x=22
Glycine max	3080	(Xia et al. 2007)	1103	2.80	2n=2x=40
Gossypium hirsutum	3425	(Guo et al. 2007)	3161	1.08	2n=4x=52
Helianthus annus	1368	(Tang et al. 2002)	2377	0.58	2n=2x=34
Hordeum vulgare	1118	(Stein et al. 2007)	4900	0.23	2n=2x=14
Humulus lupulus	662	(Cerenak et al. 2006)	2842	0.23	2n=2x=20
Lactucca sativa	1346	(Truco et al. 2007)	2205	0.61	2n=2x=18
Lens culinaris	751	(Hamwieh et al. 2005)	4116	0.18	2n=2x=14
Lycopersicum sp	1469	(Zhang et al. 2002)	950	1.55	2n=2x=24
Lupinus albus	2951	(Phan et al. 2007)	588	5.02	2n=2x=50
Malus domestica	1140-1450[b]	(Liebhard et al. 2003)	2205	0.59	2n=2x=34
Medicago truncatula	513-1225[b]	(Choi et al. 2004; Thoquet et al. 2002)	466	1.86	2n=2x=16
Nicatiana tabacum	1903	(Bindler et al. 2007)	4500	0.42	2n=2x=48
Olea europea	759-798[b]	(Wu et al. 2004)	1911	0.40	2n=2x=46
Oryza sativa		(Gaut et al. 2007)[a]	430	4.12	2n=2x=24
Phaseolus vulgaris	1358	(Grisi et al. 2007)	588	2.31	2n=2x=22
Pinus sp	1700	(Yin et al. 2003)	27244	0.06	2n=2x=24
Pisum sativum	1350	(Loridon et al. 2005)	4778	0.28	2n=2x=14
Populus sp	2100	(Woolbright et al. 2008)	485	4.33	2n=2x=38
Secale cerale	1386	(Bednarek et al. 2003)	7595	0.18	2n=2x=14
Solanum tuberosum	735	(van Os et al. 2006)	858	0.86	2n=4x=48
Sorghum bicolor	1713	(Menz et al. 2002)	1188	1.44	2n=2x=20
Theobroma cacao	782	(Pugh et al. 2004)	417	1.88	2n=2x=20
Triticum aestivum	3552	(Quarrie et al. 2005)	17101	0.21	2n=6x=42
Vitis vinifera	1245-1676[b]	(Di Gaspero et al. 2007; Troggio et al. 2007)	417	3.50	2n=2x=38
Zea mays	1825-1862[b]	(Falque et al. 2005)	2670	0.69	2n=2x=20

Tableau 2 Fréquence de recombinaison en cM/Mb pour différentes espèces de plantes.

Pour chaque espèce, la taille d'une carte génétique (cM), la taille physique (Mb), la fréquence de recombinaison (cM/Mb) et le caryotype sont indiqués. [a]/ correspond à la référence dont sont issues les fréquences de recombinaison ; les autres références faisant lien aux cartes génétiques. [b]/ moyenne utilisée pour calculer la fréquence de recombinaison. Les données de taille des génomes en Mb proviennent de la référence suivante (Bennett et Leitch 2005) sauf pour *Beta vulgaris* (Arumuganathan et Earle 1991), *Populus* sp (Kelleher et al. 2007) et *Nicotiana tabacum* (Bindler et al. 2007).

3.1.3 Autres techniques employées

Une technique utilisée uniquement chez les levures *Saccharomyces cerevisiae* (Baudat et Nicolas 1997; Gerton et al. 2000) et *Schizosaccharomyces pombe* (Cromie et al. 2007) consiste à localiser les cassures double-brin créées par la protéine Spo11 en début de méiose. Chez les mutants *rad50S*, cette protéine reste fixée de manière covalente à la séquence d'ADN qu'elle vient de couper. Ainsi son immuno-précipitation permet de récupérer les molécules d'ADN ayant subies une cassure double-brin. L'hybridation de ces molécules immunoprécipitées sur une puce à ADN permet de déterminer les sites de cassure double-brin. De même, une variante toujours utilisée chez la levure consiste à isoler les molécules d'ADN simple-brin qui résultent de la cassure réalisée par la protéine Spo11 et de la digestion des extrémités saillantes réalisée par les autres protéines. Cette méthode est appliquée sur des mutants *dmc1* où la réparation des cassures ne pouvant avoir lieu, les fragments d'ADN simple-brin persistent dans la cellule. Les molécules simple-brin sont ensuite marquées par des fluorochromes et hybridées sur des puces pour déterminer leurs localisations chromosomiques (Blitzblau et al. 2007; Buhler et al. 2007).

Le marquage des sites de cassure double-brin peut enfin être visualisé grâce à des anticorps dirigés contre l'histone H2AX phosphorylée qui marque spécifiquement les sites endommagés de l'ADN (Mahadevaiah et al. 2001).

Pour étudier la distribution fine des CO sur de petites régions, une méthode consistant à génotyper quelques marqueurs sur de l'ADN issu de spermatozoïdes ou de grains de pollen a été utilisée (Carrington et Cullen 2004) car cette technique permet le génotypage d'un nombre très important de gamètes. Le typage de spermatozoïdes a été particulièrement employé chez l'humain (Arnheim et al. 2007) pour étudier la distribution des CO sur des petites régions du génome.

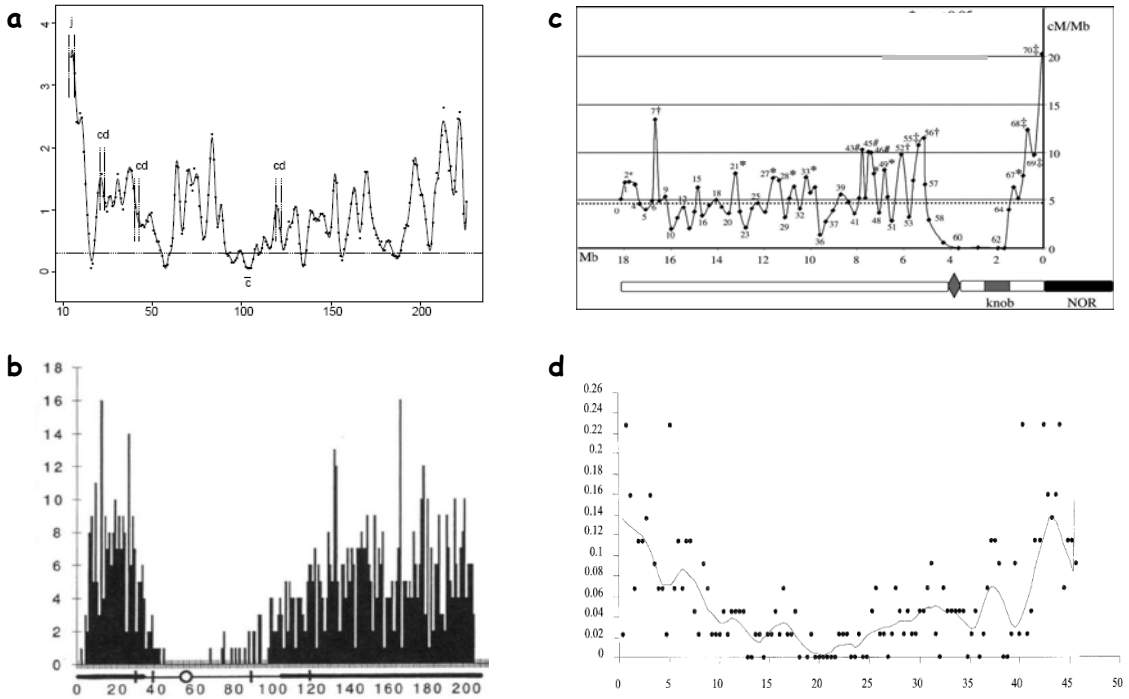


Figure 7 Variation de la fréquence de CO le long des chromosomes de différentes espèces.

L'axe des abscisses représente les distances physiques le long du chromosome soit en Mb (**a** et **c**) soit en µm (**b** et **d**) et l'axe des ordonnées représente les valeurs de fréquence de recombinaison en cM/Mb (**a** et **c**) ou en nombre de nodules de recombinaision tardifs par 0,1 µm (**b**) ou en nombre de nodules de recombinaison tardifs par µm (**d**). Variation de la fréquence de recombinaison le long du chromosome 3 de l'homme (**a**, Kong et al. 2002), le long du chromosome 4 de la tomate (**b**, Sherman et Stack 1995), le long du chromosome 4 d'*Arabidopsis thaliana* (**c**, Drouaud et al. 2006) et le long du chromosome 1 du maïs (**d**, Tenaillon et al. 2002).

3.2 Fréquence moyenne de recombinaison chez différentes espèces

La réalisation de cartes génétiques et la mise en place de base de données concernant la taille physique des génomes (http://www.genomesize.com/ et http://data.kew.org/cvalues/) chez un nombre important d'espèces ont permis de déterminer les fréquences moyennes de recombinaison par le calcul du ratio entre distance génétique et distance physique (en cM/Mb). Ainsi, la fréquence de recombinaison varie énormément de 0.3 cM/Mb chez certains insectes à 75 cM/Mb chez le champignon *Cryptococcus neoformans* (Beye et al. 2006; Wilfert et al. 2007 ; Tableau 11 en Annexes p 319). De plus, ces auteurs ont montré que les fréquences de recombinaison les plus importantes étaient retrouvées chez les protozoaires et chez les champignons.

Chez les eucaryotes supérieurs, les insectes vivant en colonies montrent le plus fort taux de recombinaison comme par exemple l'abeille ou la fourmi qui ont une fréquence de 16 et 14 cM/Mb respectivement (Wilfert et al. 2007). Ces fréquences se distinguent de la moyenne observée chez les eucaryotes supérieures comme celle observée chez l'humain (1.1 cM/Mb), chez la souris (0.4 cM/Mb), chez le chien (0.76 cM/Mb) ou encore chez le pigeon (2.6 cM/Mb) (Basheva et al. 2008; Pigozzi et Solari 1999; Wilfert et al. 2007).

Chez les végétaux, de telles données n'ont jamais été déterminées pour un nombre important d'espèces. Pour cela, nous avons utilisé l'ensemble des cartes génétiques disponibles ainsi que les tailles des génomes en Mb pour déterminer la fréquence de recombinaison en cM/Mb pour 42 espèces de plantes (Tableau 2). Chez celles-ci, la variation de fréquence de recombinaison entre différentes espèces est beaucoup moins importante que chez les animaux, elle varie entre 0.06 cM/Mb chez le pin et 5.02 cM/Mb chez le lupin (Tableau 2).

3.3 Distribution des CO le long des chromosomes

3.3.1 A l'échelle du chromosome

Indépendamment de la technique utilisée et de l'espèce étudiée, les CO se répartissent de manière non homogène le long des chromosomes révélant ainsi la présence de régions chaudes (fréquence de CO supérieure à la moyenne de la fréquence du chromosome ou du génome) et de régions froides (Figure 7). Chez certaines espèces comme chez le blé (Lukaszewski et Curtis 1993), l'orge (Kunzel et al. 2000), le maïs (Anderson et al. 2003; Tenaillon et al. 2002), ou le sorgho (Kim et al. 2005), la fréquence de recombinaison augmente fortement des centromères jusqu'aux télomères.

Ce gradient de recombinaison apparaît aussi mais de manière beaucoup moins importante chez d'autres espèces comme la tomate (Sherman et Stack 1995), le riz (Chen et al. 2002; Wu et al. 2003), le chien (Basheva et al. 2008) ou encore l'humain et la souris (Kong et al. 2002; Shifman et al. 2006) alors que chez d'autres espèces comme *Arabidopsis thaliana*, aucun gradient n'est observé (Drouaud et al. 2006) (Figure 7). Ces études montrent que la formation des CO au sein des régions entourant le centromère est inhibée. Chez les organismes où de telles études n'ont pas encore été réalisées, la concentration de marqueurs sur les cartes génétiques au niveau des centromères suggère que cette inhibition existe également chez ces espèces (Billotte et al. 2005; Hamwieh et al. 2005; Xia et al. 2007).

Une des hypothèses pour expliquer cette inhibition serait que la cellule empêche la formation de CO dans ces régions entourant le centromère car ils entraîneraient une mauvaise ségrégation des chromosomes (Lamb et al. 2005; Rockmill et al. 2006; Ross et al. 1996). Cette régulation pourrait être sous le contrôle des gènes codant les protéines du complexe synaptonémal. En effet, une mutation du gène codant la protéine Zip1 entraîne une augmentation de la fréquence de CO au voisinage des centromères (Chen et al. 2008). De plus, la faible densité de la protéine Red1, protéine qui biaise la réparation de la cassure double-brin en faveur des chromatides homologues, au niveau des répétitions centromèriques CEN3, pourrait ainsi expliquer la réparation des cassures double-brin via la chromatide sœur au sein des régions centromériques et ainsi l'absence de CO.

Une autre hypothèse serait que l'association des centromères en paires non-homologues au moment de l'initiation de la recombinaison empêcherait la formation des événements de recombinaison dans ces régions (Stewart et Dawson 2008).

Chez certaines espèces, les CO sont également inhibés au niveau des régions télomériques. Par exemple, chez la levure une diminution d'un facteur 2 à 3 est observée par rapport à la moyenne du génome (Barton et al. 2008). Cependant, dans ces mêmes régions, le nombre de NCO n'est pas inhibé contrairement aux centromères où la fréquence de NCO reste faible (Chen et al. 2008). L'inhibition des CO au niveau des centromères et des télomères semble donc provenir de deux mécanismes différents, le premier résultant plutôt d'une régulation du nombre des cassures double-brin ou de la réparation des cassures via la chromatide sœur tandis que le deuxième intervient plus au niveau de la régulation du ratio CO/NCO (Chen et al. 2008).

Points chaud de CO	Espèces	Taille en kb	Fréquence de CO en cM/Mb	Références
Ea	Souris	0.4	2250	1
Eb	Souris	4.0	25	1
Pb	Souris	5.3	283	1
Psmb9	Souris	1.8	555	1
B-globine	Homme	1.2	200	2
DMB1	Homme	1.8	5	2
DMB2	Homme	1.2	45	2
DNA2	Homme	1.3	8	2
DNA3	Homme	1.2	100	2
DPA1	Homme	1.6	27	2
MS32	Homme	1.5	40	2
MSTM1a	Homme	1.2	9	2
MSTM1b	Homme	2.1	16	2
NID1	Homme	1.2	38	2
NID2a	Homme	1.4	10	2
NID2b	Homme	1.1	4	2
NID3	Homme	2.0	70	2
PCP-1a	Homme	2.2	21	2
PCP-1b	Homme	0.5	25	2
PCP-2	Homme	1.2	77	2
SHOX	Homme	2.0	300	2
TAP2	Homme	1.0	8	2
BZ	Maïs	1.5	230[a]	3

Tableau 3 Principaux points chauds de CO étudiés chez les eucaryotes supérieurs.

[a] valeur la plus importante sur les différents intervalles étudiés sur la région de 1.5 kb. Références : 1 = (de Massy 2003), 2 = (Arnheim et al. 2007), 3 = (Dooner et Martinez-Ferez 1997)

3.3.2 A l'échelle de régions séquencées : les points chauds de CO

Les études fines qui ont pu être réalisées sur les génomes séquencés soit par la technique de typage de spermatozoïdes soit par des méthodes plus classiques chez l'humain, la souris ou la levure, ont montré que les CO se localisent dans de très petites régions de 1 à 2 kb appelées points chauds de CO. Ces derniers s'opposent ainsi aux points froids de CO qui sont des régions où la fréquence de CO est très faible (inférieure à la moyenne du génome ou aux fréquences des régions voisines).

Les points chauds peuvent être divisés en trois classes (Petes 2001) :

- Les points chauds de type α : leur activité est dépendante de la fixation de facteurs de transcription ;

- Les points chauds de type β : ils sont constitués de régions dépourvues de nucléosomes et donc sensibles aux enzymes de restriction de type nucléase ;

- Les points chauds de type γ : ils sont associés à des régions riches en GC.

Récemment, le génotypage de 52 000 marqueurs sur 204 spores issus d'un hybride de levure a permis de localiser très finement 179 points chauds de recombinaison (Mancera et al. 2008). Chez l'humain, les données de la population HapMap ont permis de déduire la présence de 21 000 points chauds historiques sur le génome, c'est-à-dire représentant la distribution des CO sur de très nombreuses générations d'individus (Arnheim et al. 2007). Cependant, la résolution limitante de certaines régions suggère que le nombre de points chauds serait plutôt autour de 50 000 (Buard et de Massy 2007). Une base de données appelée HumHot a été construite pour référencer l'ensemble de ces points chauds chez l'humain (Nishant et al. 2006). Quelques-uns ont fait l'objet d'études détaillées (Tableau 3) chez la levure, l'humain et la souris (Arnheim et al. 2007; de Massy 2003; Jeffreys et al. 2003; Mezard 2006; Petes 2001). Chez les plantes, seule la région fortement « recombinogène » couvrant le gène *Bronze* (*Bz*) chez le maïs a fait l'objet d'études approfondies. Ce gène code pour une protéine responsable de la pigmentation du grain et il se localise entre deux autres gènes (*Shrunken (Sh)* et *Waxy (Wx)*) influant la forme et la texture du grain. L'utilisation de différentes combinaisons d'allèles pour chacun de ces gènes (*Sh* et *Wx*) a permis d'établir une méthode de criblage des recombinants rapide, basée sur le phénotypage des grains, permettant l'isolement d'un nombre important de recombinants. Grâce à cette méthode, Dooner et Martinez-Ferez (1997) ont étudié la distribution de 130 CO intragéniques au niveau du gène *Bz* et ont montré que les CO se distribuaient de manière homogène le long du gène.

3.4 Facteurs influençant les variations de fréquence et de distribution des CO

Chez la plupart des organismes, les CO sont indispensables pour une ségrégation correcte des chromosomes. Cependant, la formation de CO ectopique peut avoir des effets néfastes sur la descendance.

Ainsi, le nombre et la distribution des CO le long des chromosomes sont extrêmement contrôlés (Jones et Franklin 2006).

3.4.1 Contrôle global des fréquences et de la distribution des CO

3.4.1.1 Deux régles gouvernent la distribution des CO sur les chromosomes

a L'homéostasie

Martini et al. (2006) ont été les premiers à introduire le terme d'homéostasie chez la levure en constatant que l'utilisation de variants alléliques du gène *SPO11*, produisant différents taux de cassures double-brin, ne modifiait pas la fréquence des CO malgré la diminution du nombre de cassures double-brin. L'homéostasie est donc définie comme un mécanisme permettant de tamponner des fluctuations comme la diminution du nombre de cassures double-brin et de maintenir un nombre de base de CO. Ce phénomène d'homéostasie a ensuite été confirmé à l'état sauvage chez la levure où les auteurs ont observé que le nombre de CO était indépendant du nombre de NCO (Chen et al. 2008; Roig et Keeney 2008).

De même, chez *C. elegans*, la fusion de deux chromosomes ayant chacun une taille génétique de 50 cM entraîne la formation d'un chromosome géant mais dont la taille génétique reste inchangée (50 cM ; Hillers et Villeneuve 2003) suggérant que même si le nombre de cassures double-brin est probablement doublé dans ce cas, le nombre de CO sur ce nouveau chromosome reste le même que sur chacun des deux chromosomes de départ.

Il apparaît ainsi que le nombre de CO par chromosome semble être fixé, déterminé et que la cellule a développé un contrôle génétique pour maintenir ce niveau malgré certaines perturbations.

b L'interférence

Le nombre de CO par chromosome reste faible dans la plupart des espèces avec une moyenne de 1 à 2 CO par bivalent.

Ce faible nombre de CO sur chaque chromosome résulte en partie du phénomène d'interférence entre CO, terme défini par Muller en 1916 comme suit : la formation d'un CO dans une région donnée impacte la formation d'autres CO dans le voisinage de cette région. L'interférence peut être positive, il y a alors inhibition de la formation des CO dans le voisinage du premier. Elle peut également être négative et alors la formation du premier CO favorise la formation d'autres CO dans son voisinage. Tout récemment, Mancera et al. (2008) ont montré qu'il existe aussi chez la levure *S. cerevisiae* de l'interférence entre les CO et les NCO.

i Méthodes d'étude de l'interférence

Plusieurs méthodes sont utilisées pour caractériser l'interférence. Le coefficient de coïncidence *C* permet d'estimer l'intensité et l'étendue de l'interférence sur une région précise. Si l'on considère deux intervalles, *C* est le rapport entre la fréquence observée de doubles CO (un dans chaque intervalle) et la fréquence théorique de doubles CO en absence d'interférence. Le coefficient de coïncidence varie entre 0 (interférence positive totale) et 1 (interférence nulle) mais il peut également prendre des valeurs supérieures à 1 lors d'interférence négative.

L'interférence peut également être mesurée directement par l'utilisation de différents modèles. Ainsi, le modèle de comptage permet d'estimer un paramètre *m* qui reflète le nombre de NCO présents entre deux CO. Le modèle gamma aboutit lui aussi au calcul d'un paramètre *v* qui correspond au nombre de NCO + 1 entre chaque CO. Les valeurs de *m* ou de *v* varient de 1 à l'infini et indiquent une interférence positive de plus en plus forte pour des valeurs du paramètre élevées. Ces deux modèles ont été utilisés jusqu'à présent pour décrire l'interférence sur différentes espèces car leur estimation était en accord avec les données expérimentales (Mezard et al. 2007).

Le modèle FIC (Forced Initial CO) a été développé récemment sur la base du modèle de comptage en intégrant la présence d'un CO obligatoire par bivalent pour se rapprocher encore plus des données expérimentales. Ainsi, ce modèle permet une meilleure estimation de la distribution des CO sur les bivalents que le modèle de comptage (Falque et al. 2007).

Enfin, un modèle un peu différent reposant sur des stress mécaniques a été établi par Kleckner et al. (2004). La formation du premier CO est déterminée par un stress important en un point donné du chromosome. La formation de celui-ci déclenche un stress mécanique qui se propage le long de la chromatide dans les deux sens et qui inhibe la formation des autres CO. La simulation de données avec ce modèle montre aussi une bonne corrélation avec des données expérimentales.

ii L'interférence chez différentes espèces

Le phénomène d'interférence positive est retrouvé chez quasiment toutes les espèces où il a été étudié.

Chez la tomate (Lhuissier et al. 2007), l'observation des nodules de recombinaison tardifs a permis d'estimer des valeurs de *m* de 1.3 et 1.9 pour les chromosomes 1 et 2 respectivement. Des valeurs voisines ont été observées chez l'humain (*m* = 2.3 ; Broman et Weber 2000), chez *Arabidopsis* (*m* compris entre 3 et 9 ; Copenhaver et al. 2002). Certaines espèces ont une interférence très forte comme la souris (*m* compris entre 5 et l'infini ; Broman et al. 2002) ou *C. elegans* (Cromie et Smith 2007) et ne possèdent ainsi en général qu'un seul CO par bivalent. A l'opposé, certaines espèces comme la levure *Schizosaccharomyces pombe* ou le champignon *Aspergillus nidulans* ne montrent pas d'interférence (Shaw et Moore 1998).

Des études conduites chez la souris à partir de l'analyse des foci MLH1, MSH4 et RPA suggèrent qu'il existe deux niveaux d'interférence positive (de Boer et al. 2006). Ce phénomène semble aussi exister chez la tomate où il a été observé un faible niveau d'interférence entre les nodules de recombinaison précoces et un fort niveau entre les foci MLH1 (Mezard et al. 2007).

L'interférence négative a été reportée chez quelques espèces telles que l'orge (Esch et Weber 2002), le blé (Peng et al. 2000) et *Arabidopsis* (Drouaud et al. 2007). Cependant, Mezard et al. (2007) suggèrent que l'interférence négative agissant à une distance x serait due au fait que le second CO serait localisé prioritairement à cette distance x du premier CO car c'est la distance sur laquelle l'interférence positive du premier CO agit. Ainsi, l'observation d'interférence négative résulterait uniquement de la méthode d'analyse.

iii Facteurs influençant l'interférence

A l'heure actuelle, le mécanisme responsable de l'interférence entre CO reste encore inconnu. Cependant, une mutation dans l'un des gènes codant les protéines ZMM, protéines responsables de la formation des CO interférents, entraîne une perte de l'interférence entre crossing-overs (Lynn et al. 2007).

Ces protéines ayant également un rôle important dans la formation du complexe synaptonémal, celui-ci pourrait servir de canal pour véhiculer l'information d'interférence au voisinage du premier CO. *S. pombe* et *A. nidulans*, deux organismes ne présentant ni interférence entre CO ni formation de complexe synaptonémal supportent cette hypothèse (de Boer et Heyting 2006; Shaw et Moore 1998). De plus, l'observation d'une corrélation entre l'intensité de l'interférence et la longueur du complexe synaptonémal chez la souris vient également soutenir celle-ci (Petkov et al. 2007). Enfin, chez l'humain, Lian et al. (2008) ont montré qu'une perte de continuité du complexe synaptonémal entraînait une perturbation de l'interférence. Cependant, l'implication du complexe dans la mise en place de l'interférence est remise en cause par certaines observations. En effet, il a été montré que l'interférence serait mise en place avant la formation du complexe. Les foci des protéines Zip2 et Zip3, protéines qui marquent les sites d'initiation de la mise en place du complexe synaptonémal, montrent déjà de l'interférence (Fung et al. 2004). De plus, la mutation du gène *zip1* ne perturbe pas l'interférence des complexes d'initiation, structures initiant la formation du complexe synaptonémal. Ces résultats suggérent que le complexe synaptonémal pourrait intervenir lors d'un deuxième niveau d'interférence (Page et Hawley 2004).

c Mécanismes impliqués dans l'homéostasie et l'interférence

Ces deux points de contrôle, homéostasie et interférence semblent faire partie d'un même mécanisme de régulation. En effet, une mutation au niveau des gènes *zip2* et *zip4* entraîne une diminution de l'homéostasie mais aussi une diminution de l'interférence (Chen et al. 2008). Cependant, ces mutants montrent une totale inhibition de l'interférence alors que la diminution d'homéostasie reste faible. La relation entre ces deux points de contrôle semble donc être complexe.

3.4.1.2 Mesures physiques

a Taille des génomes

Ross-Ibarra (2007) a montré en utilisant le nombre de chiasmas par bivalent chez 279 espèces d'angiospermes différentes que la taille des génomes expliquait 20 % des différences de fréquence de CO observées entre ces espèces.

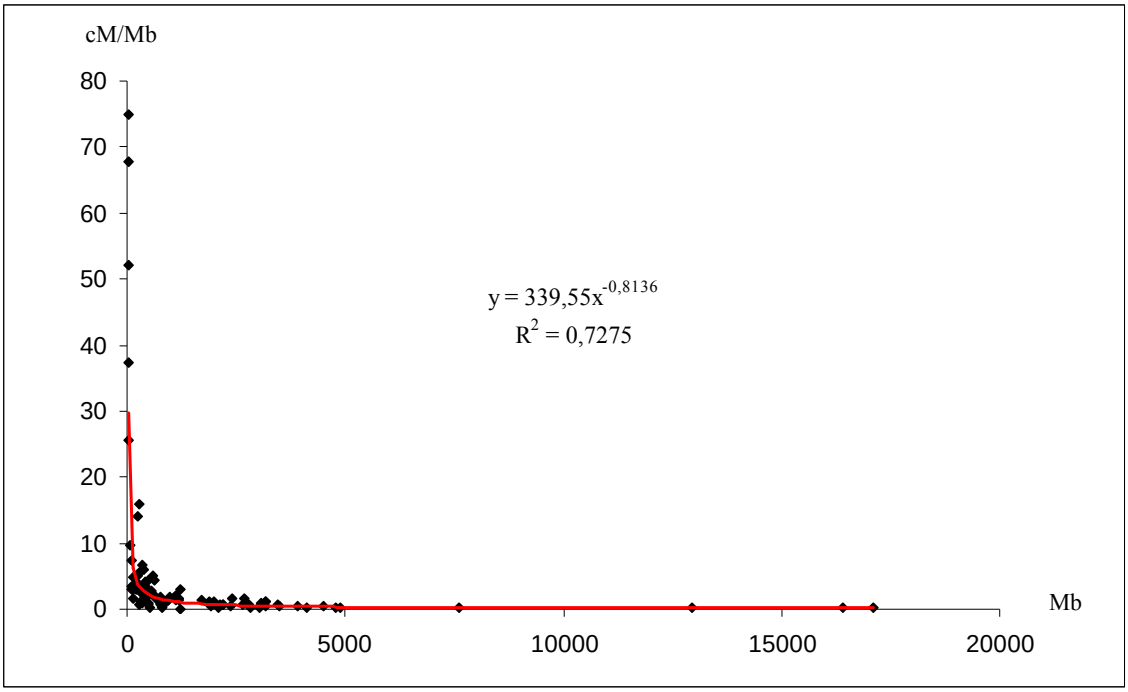


Figure 8 Relation entre la taille des génomes et les fréquences moyennes de recombinaison.

L'axe des ordonnés représente la fréquence moyenne de recombinaison en cM/Mb et l'axe des abscisses la taille des génomes en Mb. Un total de 93 paires de données a été utilisé dont 42 provenant de données issues de plantes (Tableau 2 p 57) et 51 provenant de données issues d'animaux et de champignons (Beye et al. 2006; Wilfert et al. 2007, Tableau 11 en Annexes p 319). L'équation de la courbe ainsi que la valeur du R^2 ont été déterminées grâce au logiciel excel.

En cumulant les données que nous avons obtenues chez les plantes (Tableau 2 p 57) avec celles présentes dans la littérature (Beye et al. 2006; Wilfert et al. 2007), nous avons estimé qu'il existe une relation de puissance entre la taille des génomes en Mb et la fréquence de recombinaison en cM/Mb (R^2 = 0.73, données internes non publiées). L'équation de cette courbe $y = 339{,}55\ x^{(-0.8136)}$ où x représente la taille du génome en Mb et y la fréquence moyenne de recombinaison en cM/Mb permet ainsi de déterminer grossièrement la valeur moyenne de la fréquence de recombinaison d'une espèce en fonction de la taille de son génome (Figure 8).

b Nombre et taille des chromosomes

Pour la plupart des espèces végétales, le nombre de CO observé par gènome est égal ou supérieur au nombre de chromosomes. La notion de un CO par chromosome semble ainsi respectée. Cependant, chez un petit nombre d'espèces comme la pomme de terre ou encore la vigne, le nombre de CO respectivement de 7.3 et 14.6 est inférieur au nombre de chromosomes étant de 24 et 19. De plus, il n'existe pas de corrélation entre le nombre de chromosomes et la fréquence de CO en cM/Mb. Malgré cela, pour des espèces ayant une taille de génome voisine, l'espèce ayant le plus grand nombre de chromosomes montre la plus forte fréquence de recombinaison. Par exemple, le lupin blanc (2n = 50), le haricot (2n = 22) et l'eucalyptus (2n = 22) ont un génome avoisinant les 570 Mb mais le lupin blanc montre une fréquence de recombinaison deux fois plus importante que les deux autres espèces (Tableau 2 p 57). Enfin, la fréquence moyenne de recombinaison sur un chromosome est liée à la taille de celui-ci. Kaback et al. (1992), Fung et al. (2004) ainsi que Mancera et al. (2008) ont montré que chez la levure la fréquence moyenne de recombinaison était négativement corrélée avec la taille des chromosomes. Cette relation a aussi été observée chez l'humain (Kaback 1996), chez la souris (Shifman et al. 2006) et chez *Arabidopsis* où le nombre de chiasmas est corrélé avec la longueur des chromosomes (Anderson et Stack 2002). Cependant, cette relation souffre de quelques exceptions comme chez *Hyacinthus amethystinus* où la taille des chromosomes varie d'un facteur 20 alors que le nombre de CO varie peu (Dawe 1998).

c Longueur du complexe synaptonémal

La longueur du complexe synaptonémal semble être parfaitement corrélée avec la fréquence de recombinaison chez l'ensemble des espèces étudiées à ce jour que ce soit à l'échelle du génome entier lors de comparaison entre recombinaison mâle et femelle (Drouaud et al. 2007; Kochakpour et Moens 2008) ou de comparaison entre différentes espèces (Lopez et al. 2008)

ou à l'échelle de chromosomes individualisés (Kleckner et al. 2003; Lynn et al. 2002; Pigozzi et Solari 1999; Sherman et Stack 1995). Ainsi, deux chromosomes ayant une taille en Mb constante mais un complexe synaptonémal de longueur différente montrent une fréquence de recombinaison différente. A l'inverse, deux chromosomes ayant une taille physique différente mais un complexe synaptonémal de longueur identique montrent un taux de recombinaison identique (Kleckner et al. 2003).

d Autres caractéristiques

D'autres caractéristiques telles que la configuration des chromosomes (acrocentrique, métacentrique ou télocentrique ; Dumas et Britton-Davidian 2002) ou le mode de reproduction (autogame ou allogame ; Gaut et al. 2007; Ross-Ibarra 2007) pourraient également expliquer les différences de fréquence de recombinaison observées entre différents chromosomes ou différents génomes.

3.4.1.3 Contrôle génétique

a Différences de fréquence de recombinaison entre différents génotypes

Des variations de fréquence de recombinaison ont été mentionnées entre différents génotypes de quelques espèces comme la souris (Koehler et al. 2002), *Arabidopsis* (Sanchez-Moran et al. 2002) ou le maïs (Timmermans et al. 1997; Williams et al. 1995). Ces observations ont mis en avant que des facteurs génétiques pouvaient ainsi contrôler le taux moyen de recombinaison sur les chromosomes.

b Les RecQ hélicases

Outre le rôle indispensable des protéines impliquées dans le mécanisme en lui-même, certaines protéines comme les RecQ hélicases jouent un rôle dans la modulation du nombre de CO le long des chromosomes.

Chez *Saccharomyces cerevisiae*, la protéine Srs2 est capable de démanteler le filament présynaptique empêchant la formation des CO (Sung et Klein 2006). Une mutation dans ce gène entraîne ainsi une augmentation du nombre de CO. De même, une mutation au sein du gène *sgs1* codant une autre hélicase entraîne aussi chez *Saccharomyces cerevisiae* une augmentation du nombre de CO (Jessop et al. 2006). Cette protéine serait impliquée dans la résolution des doubles jonctions de Holliday sans formation de CO. Le rôle des ces protéines est cependant différent en fonction de l'espèce étudiée puisque l'orthologue de *SGS1* chez *Schyzosaccharomyces pombe* (*RQH1*) favorise la formation des CO (Cromie et al. 2008).

c Autres facteurs génétiques

Dans le but de découvrir les facteurs responsables des différences de fréquence de recombinaison observées entre différents génotypes de souris, Hunt (2006) a réalisé une étude de QTL classique en prenant en compte le nombre de foci MLH1 comme caractère phénotypique. Il a ainsi déterminé le nombre de foci présents sur des individus F2 issus du croisement entre une souche présentant une faible fréquence de CO (Castaneus, 21.5 foci MLH1 en moyenne sur le gènome) et une souche avec une forte fréquence de CO (C57BL/6, 24.3 foci MLH1 en moyenne sur le gènome). Ces analyses indiquent que le caractère nombre de foci MLH1 et ainsi le nombre de CO est contrôlé par un faible nombre de gènes (Hunt 2006). Ce même type d'analyse a été conduit par Esch et al. (2007) chez la souris, *Arabidopsis*, le maïs et le blé en prenant le nombre de CO comme phénotype étudié. Ils ont ainsi trouvé plusieurs QTL contrôlant la fréquence moyenne de CO sur les chromosomes.

d L'importance des facteurs génétiques

L'isolement de gènes impliqués dans le contrôle de la distribution des CO sur le génome entier pourrait avoir des implications majeures dans les diagnostics médicaux ou dans les programmes de sélection. Chez l'humain, certains individus possédant un allèle spécifique identifiable pourraient présenter des prédispositions à des événements de recombinaison conduisant à de l'aneuploïdie. Chez les plantes, des allèles spécifiques augmentant la fréquence de recombinaison pourraient être utilisés pour limiter la transmission de génotypes défavorables lors de la création variétale dans les phases de pré-sélection.

3.4.1.4 Les facteurs externes

Deux sources externes sont connues comme affectant la formation, la fréquence et la distribution des CO sur les chromosomes : la température et la présence de drogues.

Chez *Saccharomyces cerevisiae*, la formation des cassures double-brin se fait plutôt à des températures importantes (33°C) qu'à des températures plus basses (23°C). De plus, la différence de phénotype chez des mutants ZMM entre basses et hautes températures indique que la température est un facteur important pour la formation des CO (Borner et al. 2004). Chez *C. elegans*, une étude récente vient de caractériser parfaitement les effets de la température sur la fréquence de recombinaison entre méiose mâle et femelle à l'échelle d'un chromosome ou à l'échelle de différents intervalles (Lim et al. 2008).

Chez cette espèce, la température influence la recombinaison mais de manière différente chez le mâle et la femelle. Chez le mâle, Lim et al. (2008) notent une augmentation de la fréquence globale de la recombinaison (pour chaque intervalle) lors de l'augmentation des températures. Chez la femelle, aucune variation globale n'est observée avec l'augmentation de la température mais une redistribution des CO sur le chromosome apparaît. La température influence également les fréquences de CO chez les plantes. Chez *Arabidopsis thaliana*, le passage de 18°C à 28°C entraîne une augmentation de la taille des intervalles étudiés d'environ 2 cM (Francis et al. 2007). Chez l'orge, l'augmentation de la température entraîne aussi une augmentation du nombre de CO alors que chez le blé la relation inverse est observée (Fu et Sears 1973; Powell et Nilan 1963).

L'effet sur la recombinaison de drogues ou de molécules extérieures peut avoir des conséquences importantes dans la recherche médicale, dans l'amélioration de techniques (ciblage de gène par recombinaison homologue) et dans la recherche fondamentale. Ainsi, plusieurs drogues ayant un impact sur la recombinaison ont été isolées comme la cisplatine ou le bisphénol A (BPA) (Hunt 2006). Récemment, le criblage de 10 000 composants biochimiques a permis d'isoler RS-1, composant qui facilite le chargement de la protéine Rad51 sur le filament présynaptique (Jayathilaka et al. 2008).

3.4.2 Contrôle de la distribution des CO à une échelle locale

3.4.2.1 Contrôle de la distribution des cassures double-brin

Les cassures double-brin étant à l'origine de la formation des CO, l'étude de la distribution de celles-ci semble être la première étape pour comprendre la distribution des CO. La distribution des cassures double-brin le long des chromosomes de la levure *S. cerevisiae* est relativement homogène (Blitzblau et al. 2007; Buhler et al. 2007). Bulher et al. (2007) observent ainsi seulement 28 régions dépourvues de cassures double-brin qui représentent 3 % de la taille physique du génome de la levure. La plupart de ces régions correspond aux régions connues comme montrant des taux de CO plus faibles comme les parties les plus distales (20 kb), les régions entourant le centromère (8 kb) ainsi que les longs cadres de lecture.

La plupart des cassures double-brin a cependant lieu dans les régions promotrices des gènes et Blitzblau et al (2007) ont montré qu'il existait une faible corrélation entre le pourcentage en GC et la fréquence de cassures. La distribution des cassures double-brin observées chez *S. cerevisiae* contraste cependant avec celle obtenue chez *S. pombe* où seulement 6 % du génome contiennent des cassures double-brin (Cromie et al. 2007). Chez cette levure, les cassures double-brin se localisent préférentiellement au sein des boucles formées par la chromatide. Les régions de la chromatine situées en dehors des boucles étant fixées à l'élément axial par la protéine Rec8, la distribution des cassures double-brin est faiblement et négativement corrélée avec les sites de fixation de cette protéine (Cromie et al. 2007).

3.4.2.2 Séquences particulières

La séquence des premiers génomes disponibles (comme chez les levures, la souris, l'humain, l'abeille, *Arabidopsis* et le riz) a permis d'étudier la relation entre la distribution des CO et les caractéristiques de la séquence comme la présence de gènes ou de séquences répétées, le pourcentage en GC ou CpG ou encore la présence de motifs particuliers.

a Séquences géniques

Comme cité dans le paragraphe 3.4.2.1, la plupart des cassures double-brin a lieu préférentiellement dans les régions promotrices des gènes (Blitzblau et al. 2007) ce qui suggère une corrélation entre densité en CO et densité en gènes. Chez la souris, Paigen et al. (2008) montrent qu'il existe une corrélation entre la distribution de la recombinaison et la distribution des gènes à une échelle de 10 Mb (r = 0.557). Cependant, à une échelle plus réduite, la corrélation diminue et ne reste significative que sur une partie du chromosome (entre 100 et 150 Mb) jusqu'à une échelle de 200 kb (r = 0.278).

Chez l'humain, Kong et al. (2002) notent, à une échelle de 3 Mb, une faible corrélation entre la recombinaison et la présence de gènes (r^2 = 0.05). Chez l'abeille, Beye et al. (2006) observent une corrélation positive entre la distance moyenne entre les gènes et la recombinaison (r = 0.13) mais aucune corrélation avec la densité en gènes à une échelle de 125 kb (r = 0.01). Ils notent aussi que les gènes avec une taille d'introns importante entraînent une recombinaison plus faible comme cela a été observé chez la drosophile et chez l'humain (Comeron et Kreitman 2000).

Chez le riz, l'étude de dix régions de 13 à 133 kb avec des fréquences de CO variables montre que les régions chaudes sont plus souvent associées à une densité en gènes plus importante (Wu et al. 2003) même si certaines régions froides possèdent une densité équivalente. A l'échelle de plusieurs Mb, il a été montré que la distribution de la recombinaison est corrélée aussi avec la distribution des gènes chez le maïs (r^2 = 0.47) (Anderson et al. 2006), chez le sorgho (Paterson et al. 2009) et chez le blé (Erayman et al. 2004). A l'inverse, chez *Arabidopsis thaliana*, Drouaud et al. (2006) montrent à une échelle de 200 kb qu'il n'existe pas de corrélation entre la présence des gènes et la fréquence de recombinaison.

Lorsqu'une échelle beaucoup plus fine est abordée, les CO se localisent plutôt en dehors de la partie codante des gènes chez l'humain (Myers et al. 2005). Cette observation est confirmée par la caractérisation de quelques points chauds de CO qui se localisent majoritairement au niveau des régions intergéniques ou au niveau des introns (de Massy 2003). Chez la levure, l'initiation se fait majoritairement dans les promotteurs mais 68 % des points chauds chevauchent la partie codante des gènes (Mancera et al. 2008) dus à l'étendue de la zone de conversion génique qui accompagne la formation des CO. A l'inverse, chez le maïs, la caractérisation fine autour du gène *bronze* montre une répartition homogène des CO le long du gène (Dooner et Martinez-Ferez 1997).

Tous les résultats obtenus jusqu'à maintenant différent entre espèces et il est ainsi difficile d'établir une règle générale concernant la relation entre la distribution des CO et celle des gènes. Cependant, chez la plupart des espèces, la présence de gènes semble avoir un impact positif sur la formation des CO.

De manière surprenante, le type de gènes pourrait avoir une influence sur la fréquence de CO et la fréquence de NCO. Mancera et al. (2008) ont observé une corrélation entre les régions chaudes biaisées, c'est-à-dire les régions présentant une différence importante entre la fréquence de CO et la fréquence de NCO, et la présence de gènes impliqués dans le vieillissement des cellules. Ces résultats suggérent que le type de réparation des cassures double-brin sous forme de CO ou de NCO pourrait être dû à la présence de certaines familles de gènes (Mancera et al. 2008).

b Eléments transposables

Chez l'humain, certains éléments comme les rétrotransposons *THE1A/B* présentant le motif CCTCCCT (Myers et al. 2005) ou les LTR solos (Nishant et Rao 2006) sont plus fréquents au sein des points chauds de CO.

Cependant, Myers et al. (2005) ont montré chez l'humain que les éléments répétés de type L1 étaient fortement sous-représentés au sein des points chauds de recombinaison. Cette relation a aussi été reportée chez le rat et la souris où les régions chaudes montrent un faible taux d'éléments répétés (Jensen-Seaman et al. 2004). Chez le riz, la présence d'éléments transposables et leur organisation expliquent en partie la présence de régions froides (Wu et al. 2003). De même, chez le maïs, Dooner et He (2008) montrent que la présence d'un bloc de retrotransposons diminue de moitié le nombre de CO au sein de la région. Chez *Arabidopsis*, aucune corrélation n'a été trouvée entre la présence d'éléments transposables et la présence de CO (Drouaud et al. 2006). Ainsi, chez la plupart des espèces, il existe une relation négative entre la présence d'éléments transposables et la distribution des CO. Le manque de corrélation observée chez *Arabidopsis* provenant peut-être de l'échelle utilisée pour l'étude ou alors du faible taux d'éléments transposables (10 %) présents chez cette espèce (2000, Analysis of the genome sequence of the flowering plant Arabidopsis thaliana).

c Les microsatellites

Les microsatellites sont des motifs de 1 à 6 paires de bases présentant un nombre variable de répétitions en tandem. Ils sont présents chez tous les organismes avec des fréquences différentes (Sharma et al. 2007). La présence des microsatellites est corrélée avec la distribution des CO à différentes échelles chez le rat (r = 0.376) (Jensen-Seaman et al. 2004), la souris (Jensen-Seaman et al. 2004), l'humain (Jensen-Seaman et al. 2004; Myers et al. 2005), *Arabidopsis* (r^2 = 0.13) (Drouaud et al. 2006), l'abeille (r = 0.19) (Beye et al. 2006), ou encore la levure (Bagshaw et al. 2008). De plus, chez cette dernière espèce, la corrélation est plus importante pour des motifs microsatellites courts et hautement répétés (Bagshaw et al. 2008). Ainsi, la présence de certains points chauds s'expliquerait par la présence de ces séquences répétées particulières.

d Pourcentage en GC et îlots CpG

La composition en bases (A, T, C et G) varie le long des chromosomes. Chez l'humain en particulier, le génome est constitué de régions appelées isochores (Costantini et al. 2006) ayant une composition en G + C différente variant de 30 à 60 % de la composition en bases totales. Un second paramètre intéressant est la présence d'îlots riches en di-nucléotides CG et nommés îlots CpG contrairement au reste du génome où la fréquence du di-nucléotide est très faible (Bird 1986).

La méthylation des cytosines sur les di-nucléotides CG conduisant à une réparation en TG, ces îlots riches en CG seraient ainsi des zones particulières où la méthylation des cytosines serait inhibée dans les lignées germinales (Gardiner-Garden et Frommer 1987). Le pourcentage en GC est positivement corrélé avec la distribution des événements de recombinaison chez l'humain (r = 0.449), la souris (r = 0.464), le rat (Jensen-Seaman et al. 2004). Ce paramètre dépend cependant d'un ensemble de facteurs comme la présence de gènes et la présence d'îlots CpG. L'utilisation d'une régression multiple permet de séparer l'ensemble de ces facteurs. Suite à une telle analyse, le pourcentage en GC apparaît négativement corrélé avec la recombinaison chez ces espèces (Jensen-Seaman et al. 2004). Cependant, cette corrélation dépend de la taille des intervalles (Buard et de Massy 2007). Chez *Arabidopsis*, une relation négative est aussi observée (Drouaud et al. 2006).

Chez ces mêmes espèces (rat, souris, homme et *Arabidopsis*), la présence d'îlots CpG est positivement corrélée avec la fréquence de recombinaison (Drouaud et al. 2006; Han et al. 2008; Jensen-Seaman et al. 2004). A l'inverse, chez le riz, aucune relation claire au niveau de dix régions analysées en détail n'a été mentionnée (Wu et al. 2003).

e Présence de motifs particuliers

L'un des motifs les plus étudiés dérive du point chaud *ade6-M26* retrouvé chez la levure *S. pombe*. Cet allèle spécifique (*M26*) résulte d'une mutation G/T qui entraîne la fixation du facteur de transcription Atf1-Pcr1 et induit la présence du point chaud de recombinaison (Kon et al. 1997). Un motif consensus de 18 pb entourant l'heptamère *M26* est important pour déterminer la distribution des points chauds de recombinaison (Steiner et Smith 2005a; Steiner et Smith 2005b). Sur quinze sites étudiés et possédant ce motif, dix se sont révélés être des points chauds. Cependant, pour les quatre restants, la présence de ce motif n'est pas suffisante pour induire la formation d'un point chaud (Cromie et al. 2007; Steiner et Smith 2005a). Chez la levure *S. cerevisiae*, Blumental-Perry et al. (2000) ont identifié un motif de 50 pb contenant une région polyA à partir de l'alignement de la séquence entourant six cassures double-brin. Cependant, Haring et al. (2004) ont montré que la délétion de ce motif consensus au niveau du point chaud *HIS2* n'entraînait pas d'effet sur la présence du point chaud. Ainsi, il semblerait que la présence de ce motif commun n'expliquerait pas systématiquement la présence des points chauds de recombinaison chez *S. cerevisiae*.

Chez l'humain, le motif CCTCCCT est présent dans 11 % des points chauds (Myers et al. 2005). De plus, la modification de celui-ci par un changement d'un nucléotide entraîne une modification de la « chaleur » des points chauds *DNA2* (Jeffreys et Neumann 2002) et *NID1* (Jeffreys et Neumann 2005).

En 2008, Myers et al. ont étendu ce motif à un motif consensus de 13 pb CCNCCNTNNCCNC qui explique 40 % des points chauds dans différents fonds génétiques. De plus, un second motif composé de 9 pb (CCCCACCCC) est associé à 3 % des points chauds du génome chez l'humain (Myers et al. 2005). La présence de ces deux motifs CCTCCCT et CCCCACCCC a aussi été retrouvée chez la souris associée à des régions où la fréquence de recombinaison est importante (Shifman et al. 2006).

Une fraction des points chauds de recombinaison semble ainsi pouvoir être expliquée par la présence de motifs particuliers. En quoi ces motifs guident la formation des événements de recombinaison reste encore une énigme. Il se pourrait qu'ils soient des sites de fixation des protéines intervenant dans la machinerie de recombinaison. Le motif identifié chez l'humain montre une périodicité de 3 pb et il pourrait être identifié par certaines protéines telles que RAD51 qui reconnait des motifs de 3 pb (Myers et al. 2008).

f Présence de séquences codant des ARN messagers non codants

Tout récemment, chez la levure, Wahls et al. (2008) ont comparé la distribution des cassures double-brin avec la distribution des gènes codant des ARN polyadénylés non codants. Ils trouvent une très bonne corrélation (*p-value* = 5.5 x 10^{-8}) entre la distribution des cassures double-brin et les loci codant des ARN messagers non codants de la classe des *prl*. Ces loci sont exprimés pendant la méiose. Les auteurs suggèrent ainsi que les ARN non codants interagiraient avec des régions cibles de l'ADN servant ainsi de guide à la protéine Spo11 pour générer les cassures double-brin.

g Conclusion

La distribution des événements de recombinaison est ainsi corrélée chez certaines espèces avec les différentes caractéristiques de la séquence. Malgré le faible nombre d'espèces étudiées, la plupart du temps et pour la majorité d'entre elles, la corrélation va dans le même sens.

Deux facteurs sont néanmoins critiques pour permettre de telles comparaisons. Premièrement, chez l'humain, les régions chaudes ont été déduites à partir d'une population d'individus complètement différents et représentent ainsi les régions chaudes historiques alors que chez les plantes la distribution des événements de recombinaison est issue d'un croisement entre deux parents et représente ainsi les régions chaudes actuelles. Deuxièmement, la taille des fenêtres d'études utilisée pour étudier ces corrélations est déterminante comme l'ont montré Paigen et al. (2008) chez la souris mais aussi Buard et de Massy (2007) et Myers et al. (2006) chez l'humain pour permettre une comparaison avec d'autres espèces.

La relation de causalité qu'il existe entre les caractéristiques de la séquence et la distribution des événements de recombinaison reste encore peu connue. Meunier et Duret (2004) mais aussi Chen et al. (2006) ont montré que la recombinaison était responsable du fort pourcentage en GC qui lui est associé et ceci à travers le mécanisme appelé biais de conversion génique (Galtier et al. 2001). La formation d'un NCO sur une région donnée entre deux molécules d'ADN présentant des différences ponctuelles de séquence entraîne la formation d'appariements incorrects sur les molécules d'ADN résultant de l'événement de NCO. En effet, celles-ci possédent un brin provenant de chaque molécule d'ADN hétérozygote de départ. La réparation biaisée de ces appariements incorrects va conduire principalement à la formation d'appariements CG plutôt que AT. Ainsi, les regions où la fréquence de NCO est importante subissent un enrichissement en nucléotides C et G.

3.4.2.3 La séquence intrinsèque n'est pas le facteur majeur qui gouverne la distribution des CO

Chez la levure, Jiang et al. (2007) ont mis au point un logiciel de détection des points chauds et des points froids de recombinaison, appelé RF-DYMHC. Il permet d'avoir une estimation correcte de la position des cassures double-brin ou points chauds avec une précision de 82 % basée ainsi uniquement sur les caractéritiques intrinsèques de la séquence.

Cependant, plusieurs données montrent que la séquence en elle-même n'est pas le seul facteur intervenant sur la distribution des événements de recombinaison sur les chromosomes car il a été montré qu'une même séquence pouvait avoir des fréquences de recombinaison réellement différentes lorsqu'elle n'était plus dans le même environnement. Borde et al. (1999) ont ainsi montré que l'insertion d'une cassette en différents points du génome entraîne des variations de fréquence de recombinaison au sein de cette cassette en fonction de sa zone d'insertion. Cette observation n'est cependant pas systématique et par exemple, le déplacement d'une séquence subtélomérique de levure n'entraîne pas de modification de sa fréquence de recombinaison (Barton et al. 2003).

Une séquence identique présente chez des individus différents d'une même espèce ou chez des individus d'espèces différentes peut également avoir des fréquences de recombinaison totalement différentes. Chez l'humain, l'utilisation d'une puce ADN sur une population européenne a confirmé à l'échelle du génome que différents individus présentaient des variations importantes du nombre de CO (Coop et al. 2008).

Winckler et al. (2005) ont aussi remarqué que la distribution des événements de recombinaison entre l'homme et le chimpanzé, qui possèdent 99 % de séquences identiques, était réellement différente. Cette même constatation a été faite au niveau des régions les mieux conservées entre le rat et la souris (Jensen-Seaman et al. 2004).

A l'échelle de certains points chauds comme par exemple *MSTM1b*, il existe une différence de fréquence de recombinaison de l'ordre de 75 fois entre différents individus possédant le même haplotype (Neumann et Jeffreys 2006). Enfin, l'exemple le plus marquant est la différence de fréquence de recombinaison observée entre la méiose mâle et la méiose femelle où les séquences sont identiques.

3.4.2.4 Comparaison mâle/femelle

a L'hétérochiasmie

Chez les eucaryotes, l'hétérochiasmie (différence de fréquence de recombinaison entre le mâle et la femelle) est un phénomène courant et variable (pour revue voir Lenormand et Dutheil 2005b; Lorch 2005). Chez la plupart des espèces, la fréquence moyenne de recombinaison est plus importante chez la femelle que chez le mâle (Tableau 12 en Annexes p 325). Chez d'autres espèces, l'inverse se produit alors que d'autres encore ne montrent pas de différence de fréquence moyenne de recombinaison en fonction du sexe.

Des différences existent aussi dans la distribution des CO sur les chromosomes. Chez l'humain et la souris, les mâles montrent une fréquence de recombinaison beaucoup plus importante que les femelles au niveau des télomères (Kong et al. 2002; Shifman et al. 2006).

b Hypothèses expliquant l'hétérochiasmie

Une des premières hypothèses émises pour expliquer la recombinaison spécifique à chaque sexe fut la règle d'Haldane-Huxley qui disait que chez les espèces hétérogamétiques, c'est-à-dire les espèces possédant des chromosomes sexuels différents, la fréquence de CO était plus faible chez le sexe hétérogamétique (Lenormand et Dutheil 2005a). Ceci serait dû à une sélection empêchant la formation de CO entre les chromosomes sexuels hétérogamétiques en diminuant ou en inhibant la formation des CO sur le génome entier entraînant ainsi une diminution de la fréquence sur les autosomes. Cette règle explique parfaitement les différences observées chez les espèces achiasmates, où aucun CO n'est observé chez le sexe hétérogamétique comme chez la drosophile ou chez le bombyx. Cependant, elle n'explique pas les différences observées chez toutes les espèces hétérogamétiques où parfois le sexe hétérogamétique montre un plus fort taux de CO (Lenormand et Dutheil 2005a).

Les différences peuvent aussi être expliquées par une plus forte activité métabolique observée chez les femelles (Lorch 2005). Ceci provoquerait une formation importante de radicaux libres qui génèreraient de nombreuses cassures de la molécule d'ADN réparées par recombinaison. Cependant, cette hypothèse expliquerait les cas où la femelle montre une fréquence de recombinaison plus importante mais pas l'inverse.

Une troisième hypothèse suggère des différences de compaction des génomes mâles et femelles, mesurées par la longueur des complexes synaptonémaux. Ceci est le cas chez l'humain (Tease et Hulten 2004), chez la souris (Petkov et al. 2007), chez *A. thaliana* (Drouaud et al. 2007) et chez le poisson zébre (Singer et al. 2002). Ces différences sont positivement corrélées avec la fréquence de CO. Par exemple, chez *A. thaliana*, le mâle présente une fréquence de CO plus importante (x 1.8) et la longueur du complexe synaptonémal est aussi plus importante (x 1.7) (Drouaud et al. 2007). De plus, Petkov et al. (2007) ont montré que chez la souris la longueur des bivalents, différente entre le mâle et la femelle, est corrélée négativement avec l'intensité de l'interférence, ceci expliquant les différences de fréquence de CO entre mâle et femelle. Ainsi, la fréquence de CO, l'interférence et la compaction des chromosomes sont trois facteurs intimement liés tout au moins chez la souris mais dont les relations de cause à effet restent encore mal connues.

La quatrième hypothèse est la sélection haploïde post-méiotique qui pourrait entraîner une sélection au niveau des gamètes choisis pour l'élaboration des descendants. Lenormand et Dutheil (2005b) montrent que la sélection haploïde est corrélée avec les différences de fréquence observées entre mâle et femelle chez les plantes. Lorsque la sélection est intense chez l'un des sexes la recombinaison est faible.

Récemment, Kong et al. (2002) ont mentionné qu'un changement d'haplotype au niveau du gène *RNF212* orthologue du gène *ZIP3* chez la levure et du gène *ZHP-3* chez *C. elegans* qui sont impliqués dans la formation des chiasmas, entraîne le changement du sexe qui présente le plus fort taux de recombinaison (Xue et al. 2008) sans entraîner de fortes variations de fréquence de CO moyenne entre les deux sexes. Ainsi, trois changements pontuels d'un nucléotide au sein de ce gène ont un effet sur la fréquence de recombinaison sexe-spécifique et l'un d'entre eux augmente la taille génétique de la carte issue de la méiose femelle de 88.2 cM.

A une échelle plus locale, l'hétérochiasmie peut être expliquée par des patterns d'empreinte génétiques qui résultent de régions présentant des différences de méthylation de l'ADN entre la séquence mâle et la séquence femelle. Chez l'humain, Lercher et Hurst (2003) notent que sur 16 régions analysées qui subissent une empreinte, treize montrent une différence de recombinaison entre la voie mâle et la voie femelle avec dans la plupart des cas une fréquence de recombinaison plus importante chez la femelle.

c Conclusion

Ainsi, une même séquence intrinsèque peut avoir des fréquences de recombinaison très éloignées en fonction de sa position sur les chromosomes et probablement des modifications qu'elle subit. Ce qui implique la présence d'autres facteurs importants pour la distribution des CO sur les chromosomes.

3.4.2.5 Modifications épigénétiques et structure de la chromatine et des chromosomes

a Modifications épigénétiques

Les modifications épigénétiques comme la méthylation ou l'acétylation des histones et de l'ADN jouent un rôle important dans la distribution des événements de recombinaison (Kniewel et Keeney 2009; Petes 2001). Tout récemment, chez *S. cerevisiae*, Borde et al. (2008) ont montré que les régions présentant une importante quantité d'histones de type 3 méthylées sur la lysine 4 étaient les mêmes que les régions de formation des cassures double-brin. De plus, une mutation dans le gène codant la protéine Set1, une méthylase responsable de la méthylation des histones de type 3, entraîne une réduction de la formation des cassures double-brin dans 84 % des sites de cassures observés chez le sauvage (Borde et al. 2008). Chez la levure *S. cerevisiae*, Mieczkowski et al. (2007) ont montré que la mutation d'un seul gène (*sir2*) codant une histone déacétylase entraîne un changement de la fréquence des cassures double-brin pour 12 % des gènes sur les 6000 qu'ils ont étudiés. Le même résultat a été observé par Merker et al. (2008) qui observent une augmentation de la fréquence de recombinaison au niveau du point chaud *HIS4* chez des mutants *rpd3* et *set2* qui codent respectivement pour une protéine qui déacétyle et méthyle les histones. De plus, chez *S. pombe*, une mutation dans le gène codant une histone acétylase *SpGcn5* entraîne une diminution de la recombinaison au sein du point chaud *ade-M26* (Yamada et al. 2004).
Le marquage épigénétique des histones apparaît ainsi comme déterminant pour la distribution des cassures double-brin.

Chez l'homme, les régions qui montrent une empreinte génétique ont une fréquence de recombinaison plus importante que les autres régions (Lercher et Hurst 2003; Sandovici et al. 2006). De plus, tout récemment, Sigurdsson et al. (2009) ont montré que le pattern de méthylation de l'ADN était corrélé avec le taux de recombinaison chez l'humain sans comprendre cependant la relation de cause à effet entre ces deux facteurs.

b Structure de la chromatine et des chromosomes

Dans les années 1920, Heitz avait observé en microscopie que certaines régions du génome avaient une coloration sombre et d'autres apparaîssaient plus claires (Grewal et Jia 2007). L'hétérochromatine a été ainsi identifiée comme une région montrant une coloration intense. Elle constitue une zone particulièrement condensée, composée en majorité de séquences satellites et d'éléments transposables avec une très faible densité en gènes (Grewal et Jia 2007). De plus, elle occupe des parties plus ou moins variables des génomes de 12 % chez *Arabidopsis* jusqu'à plus de 75 % chez la tomate (Gaut et al. 2007).

L'hétérochromatine inhibe la formation des événements de recombinaison. Chez l'humain, Kong et al. (2002) ont montré qu'il existait une relation négative entre la distribution de la recombinaison et la distribution des bandes G correspondant à l'hétérochromatine. Cette relation a été observée chez d'autres espèces. Par exemple, chez le sorgho, la recombinaison est diminuée d'un facteur 34 dans les régions hétérochromatiques par rapport aux régions euchromatiques alors que la diminution peut être 1000 fois inferieure chez la tomate (Gaut et al. 2007). Cette chute de recombinaison dans les parties hétérochromatiques vient d'être également clairement établie chez la pomme de terre (Iovene et al. 2008). Elle intervient de plus dans l'exclusion des CO au niveau des parties télomériques chez la levure (Bisht et al. 2008). Cette structure particulière est ainsi impliquée dans la régulation de la distribution des CO.

La modification de cette structure hétérochromatique et plus généralement de la structure chromatinienne entraîne ainsi des modifications de la distribution de la recombinaison. Chez la levure, Wu et Lichten (1994) ont noté que les sites de formation des cassures double-brin étaient des sites hypersensibles aux nucléases.

De plus, ils ont montré que la fixation de facteurs de transcription au niveau du promoteur du gène *PHO5* entraînait un déplacement des nucléosomes permettant ainsi une meilleure accessibilité pour les facteurs de la machinerie de recombinaison. Ceci a été clairement décrit aussi chez *S. pombe*, où la présence de l'allèle *M26* qui entraîne la formation du point chaud de recombinaison au sein du gène *ADE6* manifeste aussi une sensibilité plus importante aux nucléases (Wu et al. 2003).
La classification qui a été établie par la suite pour décrire les différents points chauds de recombinaison existant suggère que deux types de points chauds sur trois sont dus à une chromatine libre ou à une modification de sa structure (Petes 2001, paragraphe 3.3.2 p 63).

Chez *C. elegans*, la protéine HIM-17, qui est associée à la chromatine, est importante pour la modification des histones et notamment leur méthylation dans la lignée germinale. Si le gène codant cette protéine est muté, très peu de CO se forment (Reddy et Villeneuve 2004). Cette protéine serait donc nécessaire pour permettre le déclenchement des cassures double-brin impliquant ainsi l'état de la chromatine dans l'induction de la recombinaison. De plus, tout récemment, Tsai et al. (2008) ont observé que la mutation du gène codant la protéine Dpy-28 entraînaît une augmentation du nombre de CO par chromosome. Cette protéine est une sous-unité de la condensine intervenant dans la compaction des chromosomes. Elle participerait ainsi au maintien du nombre de CO par bivalent à travers la régulation de la structure des chromosomes.

Chez *Arabidopsis*, le complexe trimérique CAF-1 est impliqué dans l'assemblage des nucléosomes sur la molécule d'ADN. Le mutant *fas1-4* code pour une des sous-unités non fonctionnelle du complexe (Kirik et al. 2006). Cette mutation entraîne une réduction importante de la présence d'hétérochromatine. De plus, la fréquence de recombinaison homologue somatique est augmentée d'un facteur 96 par rapport à celle observée chez le sauvage. Kirik et al. (2006) concluent ainsi que la structure de la chromatine est le facteur majeur qui inhibe la recombinaison homologue.

Ainsi, l'ensemble de ces données montre que la structure de la chromatine et des chromosomes à l'état local mais aussi de manière plus régionale est un élément essentiel pour déterminer la distribution des CO sur les chromosomes.

3.4.2.6 Polymorphisme de séquences

a A l'échelle du chromosome

Mancera et al. (2008) ont étudié la distribution des événements de recombinaison au sein d'une souche de levure hybride S96/YJM789 leur permettant d'obtenir une très bonne résolution au sein de leur étude grâce à l'hétérozygotie entre les souches ayant servi pour réaliser l'hybride. Lorsqu'ils comparent leur distribution avec celle obtenue sur une souche pure, très peu de différences en termes de fréquence et de distribution des CO existent. Le même résultat a été observé chez *C. elegans* où Hammarlund et al. (2005) ne trouvent pas de différence de fréquence de recombinaison entre un hybride Bristol/Hawaiian présentant une différence ponctuelle d'un nucléotide (SNP pour Single nucleotide polymorphism) tous les trois kb en moyenne et une lignée pure Bristol. Cependant, lorsque le niveau de dissimilarité entre les deux homologues est trop important la recombinaison diminue fortement. Par exemple, chez les plantes, les cartes génétiques réalisées à partir de croisement inter-espèces sont plus courtes que celles réalisées à partir de croisement intra-espèces (Chetelat et al. 2000; Luo et al. 2000; Truco et al. 2007; Williams et al. 1995). Chez la tomate, l'utilisation de lignées de pre-breeding a montré que seulement 0 à 10 % des CO attendus se produisaient au niveau des régions d'introgression de regions homéologues (Ji et Chetelat 2007).

b A l'échelle de régions séquencées

A une échelle plus locale, la distribution des CO est très perturbée par la présence de SNP entre les deux molécules d'ADN homologues (Borts et Haber 1987; Chen et Jinks-Robertson 1999; Rocco et Nicolas 1996).

Chez l'humain, les points chauds *DNA2* et *NID1* montrent des différences de fréquence de recombinaison suite à la présence de polymorphisme au niveau de la séquence les chevauchant (Arnheim et al. 2007). Le changement d'un seul nucléotide fait varier la fréquence de recombinaison d'un facteur 20 au sein du premier point chaud et d'un facteur de 3 à 5 au niveau du deuxième en fonction du nucléotide retrouvé sur les deux homologues.

Chez la souris, l'intensité du point chaud *Psmb9* résulte de facteurs trans et de facteurs cis. Baudat et de Massy (2007) observent une réduction de la fréquence des événements de conversion génique au sein de ce point chaud lorsqu'un des homologues provient de la souche SGR.

Comparée à la séquence provenant des autres souches de souris, la présence de deux SNP sur la séquence de l'homologue provenant de la souche SGR laisse penser qu'ils seraient à l'origine de cette chute de la fréquence de CO. De plus, ils montrent que cette réduction pourrait provenir d'une diminution de l'initiation de la recombinaison sur l'homologue issue de la souche SGR.

3.5 Synthèse

La recombinaison est un processus essentiel à la survie de chacune des espèces sur terre et a fait ainsi l'objet de très nombreuses études. Le mécanisme de recombinaison méiotique, de même que les protéines qui sont impliquées, sont maintenant bien connus chez la levure. Malgré la présence de quelques différences, ce mécanisme semble universel et conservé entre la plupart des organismes. La recherche à très grande échelle chez *Arabidopsis*, grâce aux collections importantes de mutants, va permettre dans un futur proche d'isoler la plupart des gènes et protéines importantes chez les plantes et ainsi d'étudier précisément les différences avec la levure.

L'étude de la distribution des événements de recombinaison sur les chromosomes se développe chez la plupart des espèces. Elle a abouti tout récemment chez l'organisme modèle à une résolution encore jamais atteinte où l'ensemble des CO et des NCO a été détecté à l'échelle du génome entier grâce à l'utilisation d'une puce à ADN. De manière inattendue, ceci a permis de montrer que les événements de conversion génique représentaient 1 % en distance physique du génome à chaque méiose.

L'ensemble de ces études a aussi permis de dégager les règles majeures qui contrôlent la distribution des CO sur les chromosomes : le CO obligatoire, l'homéostasie et l'interférence qui entraînent la formation d'un minimum de un CO par chromosome même si le nombre de cassures double-brin diminue et l'inhibition de la formation d'un CO au voisinage d'un second. Les mécanismes qui gouvernent ces deux règles au niveau moléculaire restent encore obscurs, cependant, ils semblent liés puisque l'inhibition de l'interférence entraîne aussi l'inhibition de la formation du CO obligatoire.

Les premiers génomes séquencés ont entraîné la comparaison logique entre les caractéristiques intrinsèques de la séquence et la distribution des CO. De manière générale, la présence de régions codantes est positivement corrélée avec la formation des CO alors que la présence d'éléments transposables l'est de façon négative. Cependant, les différences de fréquences de CO observées entre mâle et femelle suggèrent que d'autres facteurs sont impliqués dans la distribution des événements de CO.

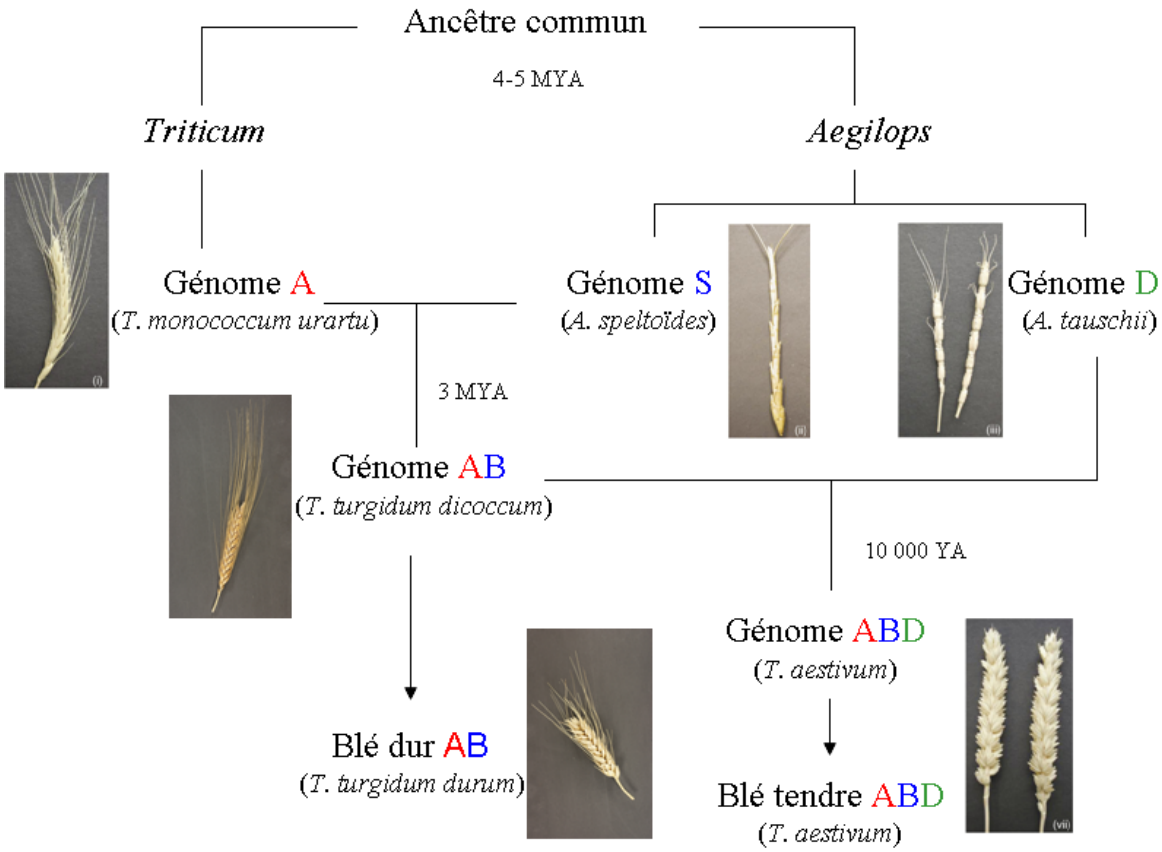


Figure 9 Schéma d'obtention du blé tendre (Triticum aestivum L.).

L'hybridation de *Triticum monococcum urartu* (A) avec *Aegilops speltoïdes* (S) a conduit à la formation de *Triticum turgidum dicoccum* (AB) il y a environ 3 Millons d'années (MYA). Ce dernier a ensuite été domestiqué et selectionné pour donner le blé dur (AB). *Triticum turgidum dicoccum* (AB) s'est hybridé avec *Aegilops tauschii* (D) pour former *Triticum aestivum* (ABD) qui représente le blé tendre.

En effet, les modifications épigénétiques et la compaction des chromosomes à des échelles locales ou plus régionales ont un impact considérable sur le positionnement des CO sur les chromosomes. Par exemple, la longueur différente du complexe synaptonémal entre mâle et femelle explique la variation de recombinaison chez la souris (Petkov et al. 2007). De plus, la mutation de gènes impliqués dans la compaction de la chromatide entraîne une redistribution et/ou une augmentation de la fréquence des CO chez différentes espèces.

La présence de facteurs génétiques en trans contrôlant la fréquence des CO sur les chromosomes a été démontrée mais les gènes gouvernant ces caractéristiques restent jusqu'à maintenant inconnus. La recherche de gènes impliqués sous les QTL majeurs qui contrôlent la distribution des CO va devenir une étape clé.

Le développement considérable des outils de génomique tels que les techniques de génotypage et de séquençage, les collections de mutants et les plateformes de transformation, vont permettre dans les années qui viennent d'élargir ces études à d'autres espèces et ainsi de mieux comprendre quels sont les principaux facteurs contrôlant la distribution des CO sur les chromosomes.

4 La recombinaison méiotique chez le blé

4.1 Le blé tendre

Le blé tendre (*Triticum aestivum* L.) est une monocotylédone de la famille des poacées. C'est une espèce hexaploïde (2n = 6x = 42, AABBDD) résultant de deux croisements interspécifiques naturels successifs (Figure 9) entre trois espèces diploïdes : *Triticum monococcum urartu* (AA), *Aegilops speltoïdes* (BB) et *Aegilops tauschii* (DD). L'hybridation entre *T. urartu* et *Ae. speltoïdes* a conduit dans un premier temps à la création des ancêtres du blé dur *Triticum turgidum* (AABB). L'hybridation de ce dernier avec *Ae. tauschii* a entraîné la formation du blé tendre. Son génome (17 000 Mb) est majoritairement constitué de séquences répétées (environ 90 %) alors que la fraction codante constitue seulement 1 à 2.5 % du génome (Li et al. 2004b; Paux et al. 2006).

Cultivé sur environ 217 millions d'hectares dans le monde en 2007 soit plus de trois fois la superficie de la France, le blé se place au premier rang mondial des plantes de grande culture (données FAO). De plus, avec plus de 600 millions de tonnes produites en moyenne par an, le blé représente 20 % des calories et des protéines de l'alimentation humaine (Gupta et al. 2008).

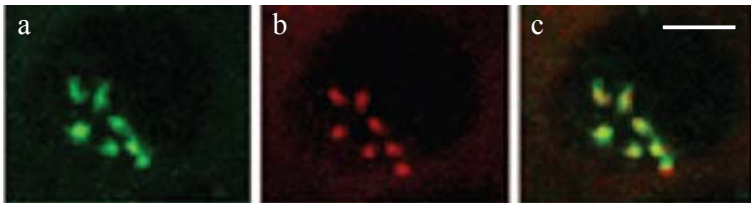


Figure 10 Comportement des centromères d'un hybride blé-seigle pendant le début de la prophase I de méiose au cours de la formation du bouquet télomérique (Prieto et al. 2004).

a/ Les centromères des 21 chromosomes de blé sont marqués en vert et les centromères des 7 chromosomes de seigle sont marqués en rouge (**b**). **c** représente la superposition de **a** et **b**. Les centromères forment sept groupes distincts. La barre représente 10 µm.

4.2 Analyse détaillée de la méiose

4.2.1 Analyse de la méiose chez le blé

Chez le blé, la présence de chromosomes de grande taille (en moyenne 808 Mb ; Paux et al. 2008 : Table S1) a fortement facilité les études cytologiques et l'étude de leur comportement méiotique. Cependant, la présence de trois génomes homéologues, de duplications ancestrales (Salse et al. 2008) mais également d'un fort taux de séquences répétées rend difficile le développement de techniques telles que l'hybridation *in situ* (Jiang et Gill 2006) qui permettent une analyse fine du comportement des chromosomes. L'utilisation du seigle dans des hybrides blé-seigle, dans des lignées de substitution de blé comportant un chromosome de seigle ou encore dans des lignées de blé comportant un fragment de chromosome de seigle, a permis de contourner partiellement ce problème en utilisant des sondes spécifiques au seigle pour visualiser et étudier spécifiquement le comportement des fragments de seigle. Ainsi, pour étudier le comportement des centromères, Corredor et al. (2007) ont utilisé les lignées de blé pour lesquelles les chromosomes 1B et 2B possèdent la région centromérique des chromosomes 1R et 2R du seigle (nommées 1Brc et 2Brc) à l'état homozygote (1Brc-1Brc et 2Brc-2Brc) ou hétérozygote (1B-1Brc et 2B-2Brc). De même, pour étudier spécifiquement le comportement des parties distales des chromosomes, Prieto et al. (2004) ont utilisé une lignée de blé chez laquelle les 15 % distaux du chromosome 1D sont remplacés par la région orthologue du seigle.

4.2.1.1 Comportement des centromères

Pendant le stade pré-méiotique, les centromères sont associés sous forme de paires de manière non-homologue et ils sont regroupés au niveau d'un des pôles du noyau (Maestra et al. 2002; Stewart et Dawson 2008). En observant par microscopie confocale en trois dimensions les chromosomes de la variété Chinese spring provenant de sections d'anthères et marqués par la technique FISH, Martinez-Perez et al. (2003) ont observé que les centromères formaient sept groupes après leur association par paires. En 2004, Prieto et al. ont observé aussi la présence de ces sept groupes sur un hybride blé-seigle (Figure 10). Ces données suggèrent une association entre les centromères des chromosomes homéologues. De plus, Martinez-Perez et al. (2003) montrent que cette association en sept groupes s'éffectue juste avant la formation du bouquet télomérique, suggérant que ce processus serait nécessaire pour l'appariement des chromosomes homologues.

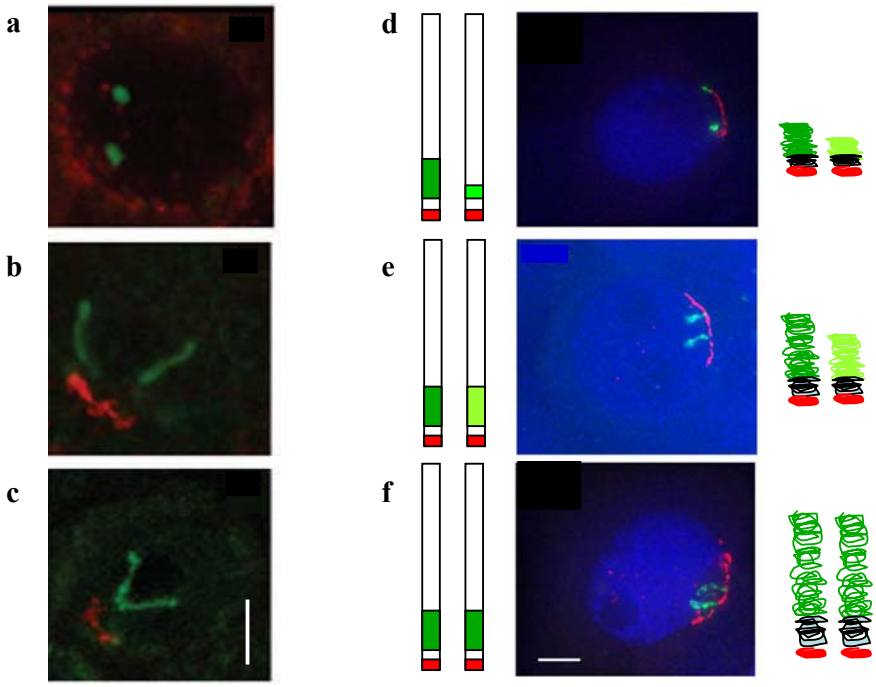


Figure 11 Comportement des régions distales des chromosomes de blé au cours des premiers stades de la prophase I de méiose.

a-b-c/ Comportement méiotique des régions distales du chromosome 1D issues du seigle (régions en vert) (Prieto et al. 2004). Les télomères sont marqués en rouge. **a**/ Stade méiotique très précoce (leptotène). **b-c**/ Formation du bouquet télomérique, élongation des régions distales et début d'appariement entre homologues (zygotène). **d-e-f**/ Comportement méiotique des régions distales de seigle dans différentes configurations d'identité entre les deux homologues (Colas et al. 2008). A gauche, représentation schématique des différentes configurations de la région de seigle présente (structures vertes) sur le bivalent 1D : très différentes (**d**), légérement différentes (**e**) et identiques (**f**). Au milieu est représentée l'observation en microscopie du comportement méiotique des régions issues du seigle (signaux verts). Les télomères sont visualisés en rouge. A droite est représenté un schéma de l'observation en microscopie. Plus la différence de séquences est importante au niveau des régions distales issues du seigle plus l'élongation au niveau de ces régions est réduite et différente entre les deux homologues.

Cependant, Corredor et al. (2007) (Revue dans Naranjo et Corredor 2008) ont observé des résultats différents en utilisant les lignées de blé 1Brc-1Brc, 2Brc-2Brc, 1B-1Brc et 2B-2Brc. Ils ont noté un fort taux de noyaux présentant plus de sept groupes de centromères en utilisant des écrasements d'anthères. Ils ont montré aussi, grâce à l'utilisation de sondes spécifiques des centromères de seigle, que chez les lignées 1Brc-1Brc et 2Brc-2Brc les centromères homologues de seigle étaient rarement associés avant la formation du bouquet télomérique suggérant que l'appariement au niveau des régions distales entraînerait plutôt l'appariement homologue des centromères. De plus, ils ont montré que la présence de centromères homéologues seigle-blé sur des chromosomes homologues n'avait aucune influence sur l'association des deux homologues, suggérant que l'homologie au niveau des centromères n'est pas responsable de l'appariement des chromosomes.

Ainsi, l'association des centromères homologues avant ou après la formation du bouquet télomérique et son rôle dans l'appariement des chromosomes homologues reste encore en discussion. Ces résultats différents pourraient s'expliquer par le fait que ces deux groupes ont utilisé des techniques d'étalement différentes, soit des coupures d'épillets (Martinez-Perez et al. 2003), soit des écrasements d'anthères (Corredor et al. 2007).

4.2.1.2 Comportement des régions distales des chromosomes

Ces études de méiose ont également permis de montrer que les chromosomes subissaient un changement de conformation dès le début de la prophase avant la formation du bouquet télomérique. Ainsi, un chromosome de blé où les 15 % distaux sont remplacés par un segment de seigle montre une décondensation qui se manifeste par une élongation de ce segment pendant le début de la prophase I (Prieto et al. 2004; Figure 11). Cette élongation avait déjà été montrée par Mikhailova et al. en 1998 qui avaient observé une augmentation de la longueur des chromosomes de cinq fois en début de leptotène par rapport à leur longueur au stade pré-méiotique. De plus, Colas et al. (2008) ont montré que cette élongation était corrélée avec le degré d'identité entre les deux homologues au niveau de cette région provenant du seigle. Ainsi, l'élongation diminue au fur et à mesure que le pourcentage d'homologie décroit et devient différente lorsque la divergence entre les deux homolgues est importante (Figure 11). A l'aide du même matériel, ils ont également montré que l'absence de cette élongation entraînait une inhibition de la formation des CO (Colas et al. 2008). Cette structure allongée de la chromatine relative aux régions distales avant la formation du bouquet télomérique est ainsi essentielle pour permettre la formation des CO chez le blé.

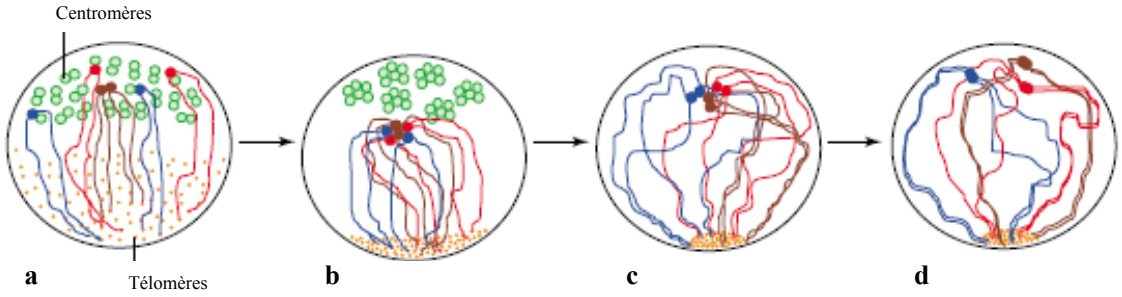


Figure 12 Représentation schématique des chromosomes de blé tendre en début de prophase I.

a/ Seulement trois paires de chromosomes homéologues sont représentées en bleu, rouge et marron, les chromosomes homologues étant de couleur identique. Pendant l'interphase pré-méiotique, les centromères des 42 chromosomes (cercles de grande taille colorés soit en vert soit en bleu, vert ou marron) sont appariés de manière non-homologue. **b**/ Formation du bouquet télomérique et association des centromères des chromosomes homéologues. **c**/ La synapse entre homologue débute au niveau des régions distales. (**d**) La synapse progresse le long des bras des chromosomes et les centromères s'associent sous forme de paires homologues. Schéma issu de Naranjo et Corredor (2004).

4.2.2 Déroulement des premiers stades de méiose

L'ensemble de ces analyses a permis de dégager le comportement des chromosomes de blé au cours des premiers stades de la prophase I qui se découpe en différentes étapes : association des centromères sous forme de paires non-homologues, élongation des parties distales, association des centromères homéologues et homologues au cours de la formation du bouquet télomérique, appariement et synapse des chromosomes homologues au niveau des régions distales qui se propagent vers le centromère et association des centromères sous forme de paires homologues (Figure 12).

4.3 Distribution des CO sur les chromosomes

4.3.1 Méthodes d'étude de la distribution de la recombinaison sur les chromosomes de blé

L'un des avantages de ce génome polyploïde est sa capacité à subir des délétions chromosomiques sans perturber le développement de la plante. La présence de chromosomes homéologues compensant l'absence de certains fragments. Ainsi, plus de 400 lignées aneuploïdes pour lesquelles une partie du génome est absente ont été développées (Figure 13) : des lignées nulli-tétrasomiques (ex : la lignée N3BT3A est une lignée où la paire de chromosomes 3B est remplacée par une paire de chromosomes 3A), des lignées ditélosomiques ou ditélocentriques (ex : la lignée Dt3BS est une lignée ne possédant que le bras court du chromosome 3B) et enfin des lignées de délétion (ex : la lignée 3BS3-0.87 est une lignée où 13 % de la partie distale du bras court du chromosome 3B sont absents) (Endo et Gill 1996; Sears 1954; Sears 1966; Sears et Sears 1978). Ainsi, les différentes lignées de délétion ont permis de découper le génome en fragments de plusieurs Mb appelés bins de délétion, permettant la production de cartes physiques de délétion (Figure 13).

4.3.2 Distribution de la recombinaison sur les chromosomes de blé

Werner et al. (1992) ont été les premiers à utiliser ces cartes physiques de délétion pour réaliser des comparaisons avec des cartes génétiques. Grâce à une dizaine de marqueurs répartis sur les trois chromosomes du groupe 7, ils ont observé que la recombinaison était faible dans les régions proximales et forte dans les régions distales.

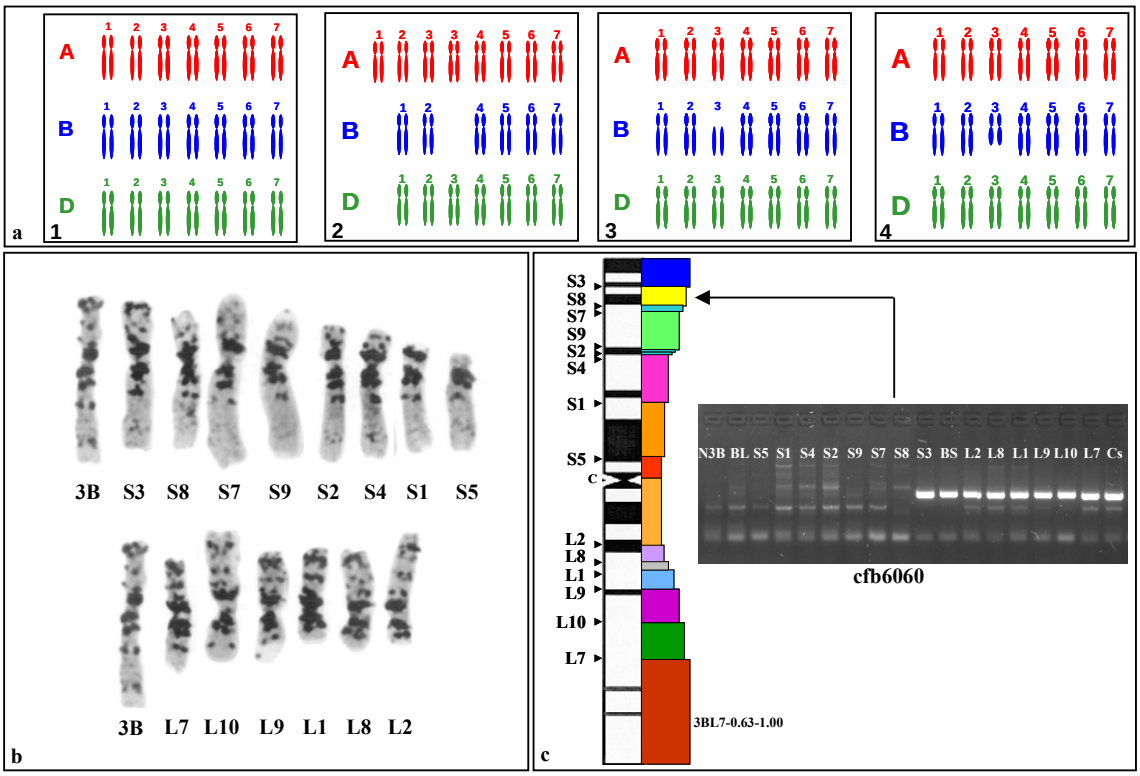


Figure 13 Lignées aneuploïdes du blé tendre, localisation des marqueurs sur les chromosomes et création des cartes physiques de délétion.

a/ Caryotypes du blé tendre et des lignées aneuploïdes. **1**/ Caryotype du blé tendre. Le génome est composé de trois sous-génomes homéologues (A, B, D) contenant chacun sept chromosomes numérotés de 1 à 7. **2**/ Caryotype de la lignée nulli-tétrasomique N3BT3A. Au sein de cette lignée, le chromosome 3B est remplacé par un chromosome 3A. **3**/ Caryotype de la lignée ditélosomique ou ditélocentrique 3BL. Au sein de cette lignée, le bras court du chromosome 3B est absent. **4**/ Caryotype d'une lignée de délétion du chromosome 3B. Au sein de cette lignée, un fragment du chromosome 3B est absent. **b**/ C-banding du chromosome 3B et des différents chromosomes 3B délétés issus des 14 lignées de délétion (Endo et Gill 1996). S3 = 3BS3-0.87, S8 = 3BS8-0.78, S7 = 3BS7-0.75, S9 = 3BS9-0.57, S2 = 3BS2-0.56, S4 = 3BS4-0.55, S1 = 3BS1-0.33, S5 = 3BS5-0.07, L2 = 3BL2-0.22, L8 = 3BL8-0.28, L1 = 3BL1-0.31, L9 = 3BL9-0.38, L10 = 3BL10-0.50, L7 = 3BL7-0.63. Chaque valeur associée à une lignée de délétion représente le pourcentage de bras restant. Par exemple la lignée 3BL7-0.63 possède un chromosome 3B où il ne reste que 63 % du bras long. **c**/ Localisation d'un marqueur sur le chromosome. Représentation schématique du chromosome 3B avec le pattern de C-banding et le lieu des cassures présentes dans les différentes lignées de délétion (gauche). Chaque région définie entre deux cassures ou entre une cassure et une extrémité de bras représente un bin de délétion (case en couleur) comme par exemple le bin de délétion 3BL7-0.63-1.00 (case marron) qui couvre les 37 % des régions distales du bras long. C = centromère. Le gel montre le produit d'amplification PCR du marqueur cfb6060 sur les lignées de délétion du chromosome 3B. Le profil d'amplification montre que ce marqueur est assigné dans le bin de délétion 3BS8-0.78-0.87. Cs = ADN de la variété Chinese spring qui a été utilisé pour réaliser les lignées de délétion. L'assignation des marqueurs au bin de délétion permet ainsi de les localiser sur les chromosomes et de créer des cartes physiques de délétion.

En 1993, Lukaszewski et Curtis ont étudié la distribution des CO grâce au polymorphisme des bandes C des parents utilisés pour réaliser des populations en ségrégation. Ils ont noté que la fréquence de recombinaison augmentait de manière exponentielle des centromères vers les télomères. Dans les années suivantes, de nombreux auteurs ont analysé les relations entre cartes génétiques et cartes physiques de délétion confirmant le gradient des centromères vers les télomères (Akhunov et al. 2003a; Akhunov et al. 2003b; Erayman et al. 2004; Gill et al. 1996a; Gill et al. 1996b; See et al. 2006). Sur le groupe d'homéologie 1, Gill et al. (1996b) ont réalisé une carte physique de délétion consensus à l'aide de 50 marqueurs qu'ils ont comparé à une carte génétique consensus comportant 288 marqueurs et une carte génétique du chromosome 1D de *Triticum tauschii* composée de 68 marqueurs. Grâce à 14 marqueurs en commun, ils ont montré que les régions entourant le centromère et couvrant 40 % du chromosome étaient dépourvues de recombinaison. Les études suivantes (Akhunov et al. 2003a; Akhunov et al. 2003b; Erayman et al. 2004) ont été réalisées en compilant les données de cartographie génétique déjà disponibles dans les bases de données et la littérature pour obtenir des estimations moyennes des distances génétiques et les comparer aux distances physiques issues des bins de délétion pour confirmer la présence du gradient de recombinaison.

4.3.3 Distribution de la recombinaison, des gènes et des régions d'évolution des chromosomes

Chez le blé, la distribution de la recombinaison sur les chromosomes semble corrélée avec la présence des gènes (Erayman et al. 2004; Gill et al. 1996a; Gill et al. 1996b). En effet, en cartographiant physiquement 3025 loci correspondant à des gènes, Erayman et al. (2004) ont estimé que le génome était composé de 48 régions riches en gènes représentant seulement 29 % de la taille physique du génome. De plus, en utilisant une carte génétique consensus, ils suggèrent que 95 % de la recombinaison se localisent dans ces régions riches en gènes.

La distribution de la recombinaison est également positivement corrélée avec la présence de duplications et/ou de délétions. Ceci résulte ainsi d'une perte de synténie entre chromosomes homéologues et entre chromosomes de blé et de riz au sein des régions où la fréquence de recombinaison est forte (Akhunov et al. 2003a; Akhunov et al. 2003b; See et al. 2006).

4.4 Facteurs génétiques impliqués dans la recombinaison chez le blé

4.4.1 Introduction

Les facteurs génétiques qui contrôlent la formation des CO et leur distribution chez le blé sont encore mal connus. En 1976, Sears présentait un premier bilan des gènes potentiellement impliqués dans le contrôle de la recombinaison. Il parlait ainsi de gènes « promoteurs d'appariement » des chromosomes homologues, localisés sur les chromosomes 2AS, 3A, 3B, 3D, 5A, 5B et 5D ainsi que de gènes « suppresseurs d'appariement » des homologues, présents sur les chromosomes 3A, 3D, 4D et 5B.

4.4.2 Le locus *Ph1*

Connu depuis les années 1950, le locus *Ph1* localisé sur le bras long du chromosome 5B a fait l'objet d'une attention particulière (Feldman 1993). En effet, il est responsable du comportement diploïde en méiose du blé en empêchant la formation de bivalents entre chromosomes homéologues. La délétion de ce locus entraîne la formation de multivalents entre chromosomes homéologues pendant la première division de méiose et l'appariement de chromosomes au sein d'hybrides de blé (Feldman 1993). Le clonage récent de *Ph1* (Griffiths et al. 2006) a montré qu'il se localisait dans une région comprenant plusieurs gènes *CDK* (Cycline dependent kinase) (Al-Kaff et al. 2008; Griffiths et al. 2006). Cette région diffère entre les chromosomes homéologues 5A, 5B et 5D par la présence d'un bloc d'hétérochromatine, par la présence de 7 gènes *Cdk* et par une expression plus importante de ces gènes sur la région du chromosome 5B (Al-Kaff et al. 2008; Martinez-Perez et Moore 2008). Le rôle de *Ph1* serait d'induire une modification de la structure des chromosomes de manière synchronisée entre les deux homologues pour permettre leur reconnaissance, leur appariement, la formation des CO et ainsi leur ségrégation correcte (Martinez-Perez et Moore 2008).

4.4.3 Les autres facteurs génétiques

Le locus *Ph2*, situé sur le bras court du chromosome 3D, est un deuxième locus dont la mutation entraîne la présence d'appariements entre chromosomes non homologues chez des hybrides de blé (Sears 1982).

Plusieurs gènes candidats sont localisés dans la délétion de 80 Mb incluant le locus *ph2*, notamment le gène *TaMSH7* ainsi qu'un gène codant une protéine présentant 78 % d'identité avec le gène *SSRP1* qui est impliqué dans la recombinaison chez d'autres espèces (Sutton et al. 2003). La mutation du gène *msh7* chez l'orge entraîne une diminution de la fertilité des plantes et ainsi une diminution du nombre de graines (Lloyd et al. 2007). D'autres gènes de réparation des mauvais appariements comme *MSH2*, *MSH3* et *MSH6* ont été cartographiés chez le blé sur les chromosomes des groupes 1, 2 et 3/5 respectivement (Korzun et al. 1999).

Le gène orthologue à *ASY1* identifié chez *Arabidopsis* ou *PAIR2* chez le riz a récemment été isolé chez le blé. Ce gène nommé *TaASY1* se localise dans la région péri-centromérique du bras long du chromosome 5A (Boden et al. 2007). La protéine TaASY1 est exprimée uniquement dans les tissus méiotiques et elle se localise au niveau de l'élément axial du complexe synaptonémal. L'inhibition de ce gène par RNAi chez le blé entraîne une augmentation de l'appariement entre chromosomes homéologues (Boden et al. 2008). De plus, il serait sous le contrôle du locus *Ph1* car la mutation de ce dernier entraîne une modification de l'expression et de la localisation de la protéine TaASY1.

Le gène *MRE11* a également été caractérisé chez les Triticées où il montre une expression différentielle entre les copies présentes sur les différents homéologues (de Bustos et al. 2007).

Enfin, des études d'expression à plus large échelle ont été entreprises. Ainsi, Crismani et al. (2006) ont observé une expression différentielle pour 1350 gènes durant les premiers stades de la méiose grâce à des analyses d'expression sur puces à ADN. De plus, tout récemment, Bovill et al. (2008) ont mis en évidence chez le blé la présence de 33 gènes impliqués dans toutes les étapes du mécanisme de recombinaison homologue par comparaison de séquences (BLAST) avec les gènes de levure et d'*Arabidopsis* suggérant un mécanisme conservé. Ils ont également réalisé une analyse QTL pour détecter la présence de loci impliqués dans le contrôle de la fréquence des CO et ont mis en évidence la présence d'un QTL majeur présent sur le chromosome 2 et conservé entre l'orge et le blé.

5 Objectifs de la thèse

Jusqu'à très récemment et malgré l'importance socio-économique du blé, très peu d'outils de génomique existaient du fait de la complexité de son génome. Ceci a longtemps ralenti la compréhension de la structure, de l'expression et de l'évolution du génome de cette espèce et par la même, l'amélioration des variétés de blé. Dans ce contexte, le consortium international de séquençage du blé (IWGSC pour International Wheat Genetics Sequencing Consortium) s'est mis en place en 2005 (www.wheatgenome.org). Le projet envisagé va engendrer la création d'un ensemble d'outils de génomique et tout particulièrement le développement de cartes physiques essentielles pour préparer au séquençage du génome. Celles-ci seront réalisées à partir de chromosomes uniques ou de bras de chromosomes pour diminuer la complexité et partager les efforts entre différentes équipes internationales. L'objectif final étant de séquencer le génome du blé tendre et ainsi de contribuer à accélérer la création de nouvelles variétés.

Dans le cadre d'un projet pilote réalisé dans l'équipe d'accueil de mon doctorat, le chromosome 3B du blé tendre (*T. aestivum* cv Chinese spring) a été choisi comme modèle pour établir la première carte physique spécifique d'un chromosome. Ce chromosome a été choisi pour des raisons techniques car il est le plus grand parmi les 21 chromosomes (995 Mb) et ainsi le plus facile à isoler par cytométrie de flux (Safar et al. 2004). De plus, il porte un grand nombre de gènes d'intérêt agronomique (pour revue voir Paux et al. 2008). Enfin, les chromosomes du groupe d'homéologie 3 sont ceux qui montrent la meilleure relation de colinéarité avec le génome du riz (chromosome 1, La Rota et Sorrells 2004), espèce modèle séquencée chez les graminées. Ceci a conduit à la réalisation d'une banque BAC spécifique du chromosome 3B (Safar et al. 2004) et au développement de la première carte physique d'un chromosome de blé (Paux et al. 2008). Cet outil, à travers le développement d'un nombre considérable de marqueurs et du séquençage de régions couvrant plusieurs Mb, offrait ainsi une opportunité unique pour étudier finement la distribution des CO sur ce chromosome.

L'objectif de cette thèse était donc de mieux comprendre la distribution des CO le long des chromosomes du blé tendre et les facteurs qui l'influencent, à travers une étude fine sur le chromosome 3B.

Dans une première étape, nous avons comparé une carte génétique dense du chromosome 3B, représentant la distribution des CO, avec une carte physique de délétion élaborée à l'aide des 14 lignées de délétion disponibles pour le chromosome 3B dans le but d'étudier de manière fine la distribution des CO le long du chromosome 3B entier.

Au cours d'une deuxième étape, nous avons étudié quels sont les facteurs qui influencent la distribution des CO. De part l'ensemble des techniques disponibles au laboratoire mais aussi l'importance des différents facteurs, nous avons choisi de nous focaliser sur trois d'entre eux : la position d'une région sur l'axe centromère/télomère; l'influence du sexe et la présence d'interférence. Pour le premier point, nous avons utilisé une des lignées de délétion du chromosome 3B pour développer un matériel végétal original afin de voir si le fait de déplacer une région initialement située en position intermédiaire sur le chromosome vers une position distale sur le chromosome délété entraînait une augmentation du taux de CO au sein de cette région. Pour le deuxième point, nous avons utilisé deux populations en ségrégation issues du même croisement mais réalisées l'une par androgenèse et représentant ainsi la méiose mâle et l'autre par gynogenèse et représentant ainsi la méiose femelle pour comparer l'influence du sexe sur la distribution et le taux des CO. Enfin, l'étude de l'interférence a été réalisée en collaboration avec l'INRA du Moulon et en particulier avec M Falque et O Martin. Nous avons utilisé les données générées pour l'étude de l'influence du sexe sur la distribution des CO pour déterminer la présence ou non d'interférence entre CO sur le chromosome 3B et regarder si celle-ci varie en fonction du sexe. Le modèle gamma et le coefficient de coïncidence ont été utilisés pour analyser l'intensité et l'étendue de l'interférence.

Finalement, nous nous sommes focalisés sur la première région séquencée de plusieurs Mb (locus *Rph7*, 3.1 Mb) située en partie distale du bras court du chromosome 3B, possédant une densité en gènes importante et une fréquence de CO élevée. Nous avons ainsi utilisé des populations de grande taille pour obtenir un nombre important de CO à étudier dans cette région. Nous avons ensuite regardé la distribution des CO sur cette région et comparé les données avec les caractéristiques intrinsèques de la séquence (contenu en gènes, éléments répétés…) pour comprendre quels pouvaient être les facteurs qui gouvernent la distribution des CO sur ce locus. De plus, nous avons affiné l'analyse de ce locus au niveau de certaines régions chaudes afin de déterminer si, comme chez d'autres espèces, il existe des points chauds et des points froids de CO. Enfin, l'utilisation de deux populations différentes nous a permis de déterminer si la distribution des CO est conservée entre différents croisements.

III Matériels et Méthodes

1 Matériel végétal

1.1 Lignées parentales

Tout le matériel végétal utilisé pour la suite des analyses a été développé à partir de trois lignées principales : Chinese Spring, Renan et Courtot.

Chinese Spring constitue la variété de référence internationale pour les études de génétique chez le blé tendre. Elle doit en grande partie sa notoriété à sa facilité à accepter des croisements interspécifiques. C'est en étudiant les descendances de tels croisements que Sears (1939 ; 1944) a abouti à la production d'aneuploïdes, de lignées monosomiques (1954), de lignées nullisomiques-tétrasomiques (1966) et de lignées ditélosomiques (1978). Plus récemment, Endo et Gill (1996) ont également créé un jeu de plus de 400 lignées de délétion. C'est ce matériel aneuploïde qui est à la base de l'assignation des marqueurs, aussi bien moléculaires que phénotypiques. Chinese Spring est un blé de printemps mutique de taille haute et donc sensible à la verse mais aussi à de très nombreuses maladies ou attaques d'insectes. C'est un blé peu productif et inapte à la panification.

Renan est une variété INRA inscrite au catalogue en 1989 qui se caractérise par une bonne résistance au froid, une bonne résistance à de nombreuses maladies et par une bonne valeur boulangère due à une teneur en protéine élevée. Ses qualités de rusticité en font une variété très appréciée en Europe de l'Est et en particulier en Autriche pour des systèmes de type culture biologique malgré des rendements inférieurs aux variétés plus récentes. Du fait de la présence du géniteur VPM (issus d'un croisement entre les espèces *Ae. Ventricosa*, *T. persicum* et *T. aestivum* cv Moisson) dans son pedigree, elle est régulièrement utilisée dans les schémas de sélection, en particulier comme source de résistance au piétin-verse (présence du gène *Pch1* sur le chromosome 7D).

Courtot est une variété INRA inscrite au catalogue en 1974. C'est un blé tendre d'hiver demi-nain, aristé, assez précoce, résistant à l'attaque de nombreuses rouilles mais assez sensible à l'oïdium ainsi qu'à d'autres maladies telles que la septoriose, la fusariose ou le piétin-verse. Du fait d'une grande résistance à la verse et d'une forte capacité de tallage, Courtot possède une assez bonne productivité. Enfin, c'est également un blé de bonne qualité boulangère. Bien qu'ancien, il reste encore cultivé dans certaines régions, le plus souvent sous contrat.

Les variétés Opata, W7984 (blé synthétique, issu du croisement entre un blé dur (Altar 84) et une accession d'*Ae. Tauschii*), Arche et Récital ont également été utilisées pour des tests de polymorphisme car ils servent de parents à des populations de cartographie génétique travaillées au laboratoire (population ITMI (pour International Triticeae Mapping Initiative) servant de référence internationale et population Arche x Récital (ARe)).

1.2 Lignées aneuploïdes

Pour cartographier physiquement les marqueurs moléculaires, nous avons utilisé les lignées aneuploïdes développées à partir de la variété Chinese spring par Sears (1954 ; 1966 ; 1978) et Endo (1996), (Figure 13 p 114). Plus spécifiquement, nous avons utilisé une lignée nullisomique 3B tétrasomique 3A (N3BT3A), deux lignés ditélosomiques Dt3BS (bras long absent) et Dt3BL (bras court absent) et 14 lignées de délétion du chromosome 3B. A chaque lignée de délétion est associée une valeur qui détermine le pourcentage physique de bras restant. Par exemple, la lignée 3BS4-0.55 conserve 55 % du bras court du chromosome 3B. L'ensemble des 14 lignées de délétion permet de définir 16 bins de délétion sur le chromosome 3B : 3BL7-0.63-1.00, 3BL10-0.50-0.63, 3BL9-0.38-0.50, 3BL1-0.31-0.38, 3BL8-0.28-0.31, 3BL2-0.22-0.28, C-3BL2-0.22, C-3BS5-0.07, 3BS5-0.07-0.33, 3BS1-0.33-0.55, 3BS4-0.55-0.56, 3BS2-0.56-0.57, 3BS9-0.57-0.75, 3BS7-0.75-0.78, 3BS8-0.78-0.87 et 3BS3-0.87-1.00. Chaque bin correspond à la région définie par les points de cassures de deux lignées de délétion. Par exemple, le bin 3BS4-0.55-0.56 correspond à la région comprise entre 55 et 56 % du bras court et couvre donc 1 % de ce bras, les lignées 3BS4-0.55 et 3BS2-0.56 permettant de définir ce bin. La taille physique de chaque bin en Mb (Figure 20 p 166 et Tableau 17 en Annexes p 347) a été estimée à partir du pourcentage de couverture de chaque bin et de la taille en Mb du bras long (562) et du bras court (433) du chromosome 3B.

1.3 Populations F2

Les populations F2 ont été développées à partir de croisements entre les variétés Chinese Spring et Renan ou Courtot. En serre, tous les épis d'une plante Chinese Spring ont été émasculés et la moitié a été pollinisée avec Renan, l'autre avec Courtot. Dix-sept jours après pollinisation, les épis sont récoltés pour réaliser une génération F1 accélérée.

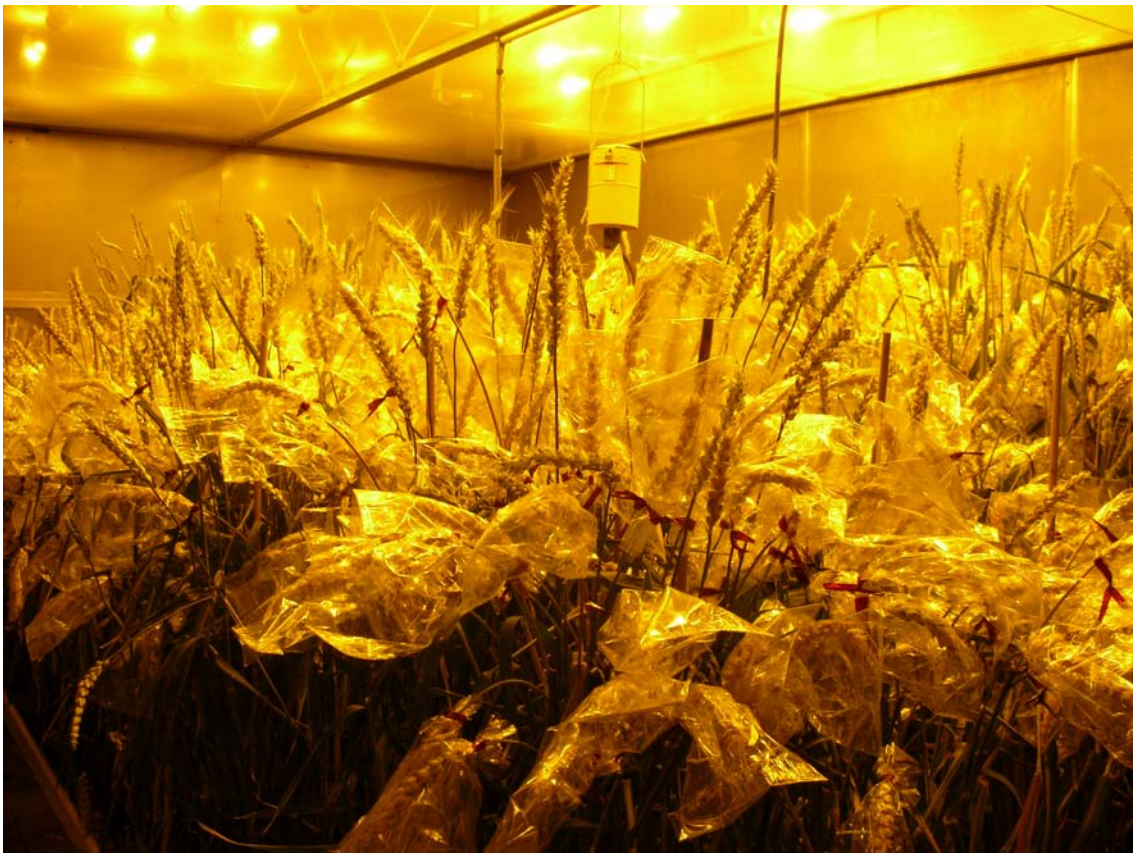


Figure 14 Pose de sachets d'autofécondation sur les épis des plants de blé présents en chambre de culture pour éviter la pollinisation croisée entre différents plants.

En conditions stériles, les embryons de chaque grain sont disséqués sous une loupe binoculaire et mis en culture dans un pot contenant un milieu solide enrichi de type Muller (Composition détaillée Tableau 13 en Annexes p 327). Les pots sont conservés 24 heures à l'obscurité puis mis en chambre de culture sous néons universels jusqu'au développement de plantules au stade deux feuilles (environ quinze jours). Ces individus sont ensuite repiqués en pots 5x5 dans du terreau et ré-acclimatés deux semaines en serre. Ils sont ensuite mis en chambre de vernalisation à 4°C pendant 8 semaines. Les douze plus belles plantes pour les deux croissements sont repiquées en pots de 3 litres contenant un mélange de terreau et de pouzzolane et maintenues dans une chambre de culture. Les plantes sont cultivées pendant 20 jours avec une alternance de température de 17°C (12 h de jour) et de 12°C (12 heures de nuit) puis un mois et demi avec une alternance de température de 23°C (16 h de jour) et de 18°C (8 h de nuit), quatre jours avec une alternance de température de 24°C et de 20°C (mêmes horaires jour/nuit) et enfin sept jours à 24°C jour et nuit (mêmes horaires jour/nuit). Des sachets d'autofécondation sont posés sur tous les épis dès leur apparition (Figure 14). Après récolte, les épis sont battus mécaniquement et les grains sont conditionnés en sachet et mis en chambre de stockage à 4°C. Entre six et huit mille grains ont été obtenus pour chaque croisement. Dans un premier temps, 376 individus F2 Chinese spring x Renan (CsRe) ont été utilisés pour générer la population F2CsRe pour étudier la distribution des CO à l'échelle du chromosome. Les grains ont été semés dans des pots 5x5 contenant du terreau et mis en serre. Environ quinze jours après le stade 2 feuilles, des prélèvements de feuille en plaques 96 sur glace sont effectués. Pour l'étude fine au niveau des régions séquencées, respectivement 16 et 24 plaques de 96 individus pour les populations F2CsRe et F2CsCt ont été semées et récoltées en plaque 96 sur glace.

1.4 Populations haploïdes doublés

Deux populations haploïdes doublés (HD) issues d'un croissement entre les cultivars Chinese spring et Courtot ont été analysées. La première population a été réalisée en 1991 et en 1992 par androgenèse selon la méthode décrite par Martinez et al. (1994) et représente la méiose mâle. Brièvement, les individus F1 issues du croisement entre Cs et Ct sont cultivés en serre jusqu'à épiaison. Les épis récoltés sont stockés 4 à 8 jours à 4°C dans le noir.

Les anthères sont ensuite disséquées, aseptisées, déposées sur un milieu solide Potato-2 (Martinez et al. 1994) et incubées à 27°C pendant 4 à 7 semaines. Les embryons formés sont ensuite déposés sur milieu solide de régénération pour permettre la formation de plantules. Les plantes sont ensuite doublées à la colchicine. Une population de 182 lignées HD a été utilisée pour réaliser notre étude.

La deuxième population a été réalisée en 2001 par gynogenèse selon la méthode décrite par Laurie et Bennett (1988) et représente la méiose femelle. Les individus F1 issus du croisement entre Cs et Ct sont cultivés en serre. Les épis sont émasculés et pollinisés entre un et trois jours plus tard, avec du pollen de maïs. Les grains haploïdes générés sont disséqués et l'embryon est déposé sur un milieu de culture riche solide de type Difco-orchid contenant 8 g/l de sucrose (Laurie et Bennett 1988). Les plantes sont finalement doublées à la colchicine. Une population de 185 lignées HD a été utilisée pour réaliser notre étude.

Une population haploïde doublé de 239 individus issue du croisement entre les variétés Arche et Récital (HDARe) et réalisée par androgenèse (Laperche et al. 2006) a également été utilisée.

1.5 Population F3 Chinese spring x Renan 3BS4

En serre, les épis d'un individu Renan sont émasculés et pollinisés avec du pollen de la lignée de délétion 3BS4-0.55. Deux individus F1 sont conduits en génération accélérée dans les mêmes conditions que décrites dans le paragraphe 1.3 p 129. Un total de 380 individus F2 est semé en barquette de 96 et un prélèvement de feuilles est effectué pour chaque individu au stade deux feuilles. Après sélection moléculaire (paragraphe 2.1 p 177) l'individu F2 n°369 est repiqué en pot de 5x5 contenant du terreau et mis en chambre de vernalisation pendant 8 semaines à 4°C. Il est ensuite repiqué en pot de 3 litres contenant un mélange de terreau et de pouzzolane et mis en chambre de culture dans les mêmes conditions que les individus F1CsRe. Un sachet d'autofécondation est posé sur chaque épi. Un total de 379 individus F3 est ensuite utilisé pour constituer la population délétée F3 Chinese spring x Renan 3BS4 (F3CsRe3BS4). Au stade deux feuilles, des prélèvements de feuilles sont effectués. Ces individus sont mis 8 semaines en chambre de vernalisation à 4°C puis sont repiqués sous serre. Un sachet d'autofécondation est finalement posé sur chaque épi. Un total de 96 grains F4CsRe3BS4 provenant de l'autofécondation de la plante F3CsRe3BS4 n°357 sélectionnée à partir des données moléculaires est semé en barquette 96 et des prélèvements de feuilles sont réalisés au stade deux feuilles.

Désignation des SSR	Nombre de SSR testés	Nom des SSR	Sources
AC	1	Répétition AC	GrainGenes
barc	33	USDA-ARS Beltsville Agricultural Research Center	GrainGenes
cfa	5	Clermont-Ferrand A Genome	Clermont-Ferrand
cfb	224	Clermont-Ferrand B Genome	Clermont-Ferrand
cfd	7	Clermont-Ferrand D Genome	Clermont-Ferrand
cfe	91	Clermont-Ferrand EST	Clermont-Ferrand
cfr	93	Clermont-Ferrand Riz	Clermont-Ferrand
cs-ssr	2	Chinese Spring-SSR	(Shen et al. 2005)
dw	18	NC	Syngenta
gpw	59	Genoplante Wheat	Clermont-Ferrand
gdm	1	Gatersleben D-Genome Microsatellite	GrainGenes
gwm ou wms	63	Gatersleben Wheat Microsatellite	GrainGenes
nw	44	NC	Syngenta
swm	1	Swiss Wheat Microsatellite	B. Keller
wmc	50	Wheat Microsatellite Consortium	GrainGenes
wmj	3	Wheat Monsanto JIC	RAGT
wmm	58	Wheat Monsanto Microsatellites	RAGT

Tableau 4 Différents types de microsatellites testés.

NC = non connu
Sources : GrainGenes = http://wheat.pw.usda.gov/; B. Keller (Université de Zurich, Suisse) ; RAGT : Rouergue Auvergne Gévaudan Tarnais.

2 Extraction d'ADN

Les prélèvements de feuilles et les extractions d'ADN sont effectués en plaque 96 (1.2 ml). L'extraction d'ADN est réalisée à partir d'un protocole utilisant du méta-bisulfite. Environ 100 mg de feuilles sont broyés en poudre fine à l'aide d'un mélangeur à peinture en maintenant la congélation par immersion régulière dans l'azote liquide. 600 µl de tampon d'extraction (0.1 M Tris, 0.05 M EDTA, 0.5 M NaCl, 0.02 M sodium metabisulfite et 0.2 M RNAse) sont ajoutés dans chaque puit (Biomek FX2000) puis les plaques sont mélangées et incubées au bain-marie (95°C) pendant 45 min. Après centrifugation (4000 g, 10 min), 430 µl de surnageant sont prélevés au robot et déposés en plaque 96 contenant 425 µl d'un mélange d'isopropanol/acétate d'ammonium à 7.5 M (400/25). Les plaques sont mélangées et après précipitation, l'ADN est centrifugé 30 min à 4000 g. Le surnageant est éliminé et les culots sont lavés à l'alcool à 70 %. Après 30 min de centrifugation à 4000 g, le surnageant est de nouveau éliminé et les culots sont séchés une nuit à température ambiante puis repris dans 100 µl de Tris-EDTA 0.1 M. L'ADN d'une dizaine d'individus choisis aléatoirement est alors dosé à l'aide d'un nanodrop de façon à estimer une concentration moyenne pour les échantillons de la plaque.

3 Marqueurs moléculaires et conditions d'amplification

3.1 Marqueurs moléculaires

3.1.1 Microsatellites

L'établissement de la carte physique du chromosome 3B (Paux et al. 2008) a permis la collection de 751 microsatellites ou SSR (Simple Sequence Repeat) issus des bases de données publiques et privées, de la banque BAC du chromosome 3B et des EST de blé (Tableau 4). Parmi ceux-ci, 90 répartis de façon homogène sur l'ensemble du chromosome 3B ont été sélectionnés pour l'analyse à l'échelle du chromosome entier sur les populations F2CsRe, HDCsCt et HDARe.

De plus, à partir des données de séquence des contigs *Rph7* et *Sr2*, d'autres marqueurs SSR ont été identifiés à l'aide du logiciel SSRdesign (Frédéric Choulet, communication personnelle) qui combine le logiciel SSRsearch (F. Choulet) pour la recherche des microsatellites et le logiciel primer3 (http://frodo.wi.mit.edu/) pour la réalisation des amorces. Cette recherche a été effectuée sur des séquences masquées, c'est-à-dire où les éléments transposables sont éliminés. Un total de 178 SSR a été sélectionné premièrement en fonction de leur position sur les contigs et deuxièmement en fonction de leurs motifs, en choisissant en priorité les motifs di-nucléotides avec le maximum de répétition pour maximiser la probabilité d'observer du polymorphisme. Ces nouveaux SSR ont été nommés cfb5000, cfb5001… pour ceux issus du locus *Sr2* et cfb6000, cfb6001… pour ceux venant du locus *Rph7*.

3.1.2 ISBP

Les données de séquençage d'extrémité de BAC a permis récemment le développement d'un nouveau type de marqueur issu des sites d'insertion des éléments transposables et nommé ISBP (pour Insertion Site Based Polymorphism ; Paux et al. 2006). Sur le chromosome 3B, 711 marqueurs ISBP ont été isolés dans le cadre du développement de la carte physique (Paux et al. 2008) et 26 ont été choisis pour l'analyse de la recombinaison sur les différentes populations à l'échelle du chromosome entier. Au niveau des régions séquencées, nous avons utilisé 5 marqueurs ISBP développés au cours d'autres programmes de recherche. Nous avons développé 14 nouveaux marqueurs ISBP au niveau de régions spécifiques (régions 8 et 9) du contig *Rph7* pour l'étude fine de la distribution des CO. La recherche d'ISBP a été faite à l'aide du logiciel ISBPfinder (F. Choulet, communication personnelle). Ce logiciel utilise le logiciel RepeatMasker (http://www.repeatmasker.org/) et la base de données de séquences répétées des Triticées (TREP pour Triticeae Repeat sequence database) pour la recherche et l'annotation des éléments transposables. Il localise ensuite les jonctions présentes et utilise primer3 (http://frodo.wi.mit.edu/) pour le dessin des amorces autour des jonctions trouvées.

3.1.3 SNP

La recherche de polymorphisme dû à une différence d'un nucléotide (SNP pour Single Nucleotide Polymorphism) a été réalisée au niveau des marqueurs SSR (37) et ISBP (17) non polymorphes.

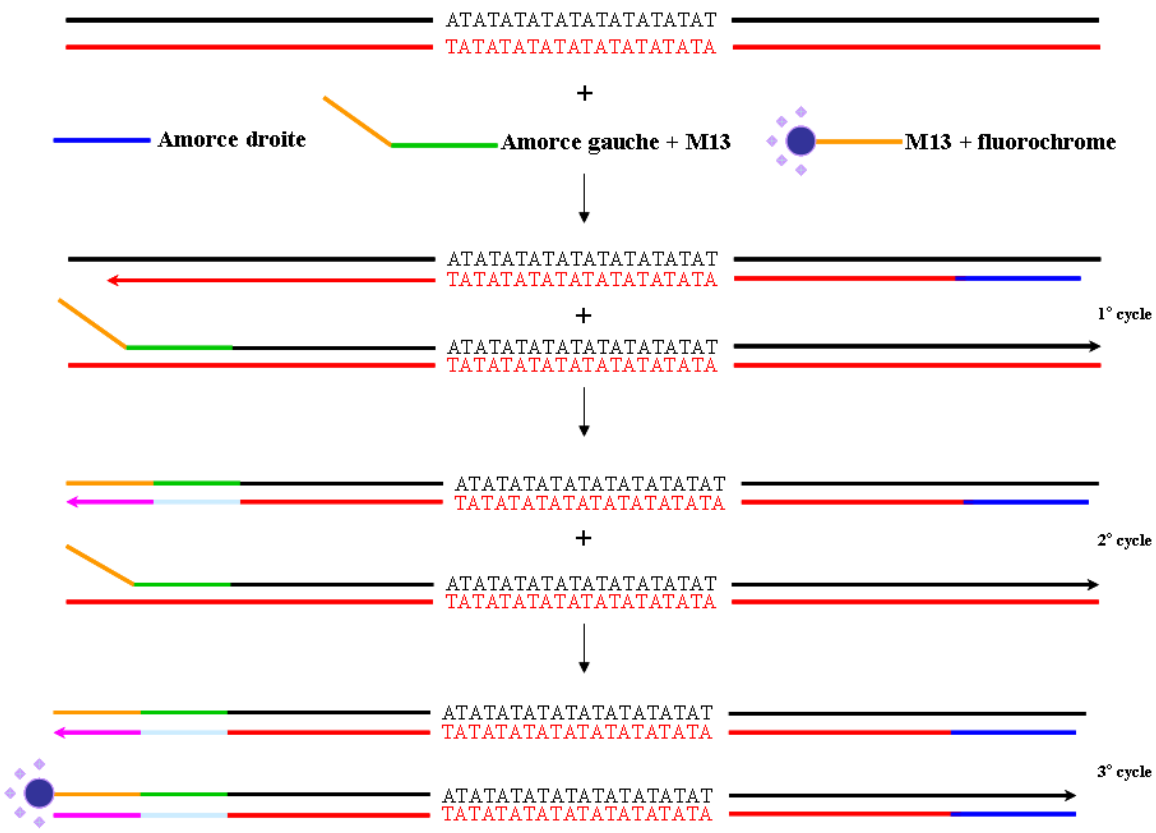


Figure 15 Description du protocole d'amplification PCR de type « M13 ».

La réaction PCR se déroule en présence de la molécule d'ADN contenant le motif microsatellite à amplifier, d'une amorce droite, d'une amorce gauche couplée à une amorce M13 et d'une amorce M13 couplée à un fluorochrome. Lors du 1^{e} cycle de la réaction, l'amorce droite et l'amorce gauche s'hybrident aux brins de la molécule d'ADN sauf au niveau de la partie M13 qui n'est pas complémentaire et il y a synthèse des brins complémentaires. Lors du 2^{e} cycle, les mêmes événements se produisent avec en plus synthèse du brin complémentaire à l'amorce M13. Lors du 3^{e} cycle, l'amorce M13 couplée au fluorochrome peut venir se fixer sur sa région complémentaire pour initier la synthèse d'ADN, la molécule d'ADN synthétisée est ainsi marquée par un fluorochrome. Elle pourra être visualisée par électrophorèse capillaire.

Les amplicons ont été obtenus dans les mêmes conditions que le protocole utilisant le rouge de crésol (paragraphe 3.2.2 p 141) en remplaçant ce dernier par de l'eau et en ajustant les réactifs pour un volume final de 25 µl. Ces produits PCR ont été envoyés au CNG (Centre National de Génotypage) pour séquençage.

La comparaison des séquences des variétés Cs, Ct et Re a été réalisée à l'aide du logiciel Genalys (CNG), afin d'identifier les SNP et les indel (insertion/délétion) présents. Les séquences des gènes *HGA1*, *HGA2* et *HGA3* présents au locus *Rph7* ont été fourni par C. Ravel.

3.2 Conditions de génotypage des marqueurs SSR et ISBP

Trois protocoles différents d'amplification PCR et de visualisation des produits PCR ont été utilisés en fonction du type de marqueurs : le protocole « M13 » (Figure 15), le protocole « rouge de crésol » et le protocole « SYBR Green ». Le test de départ sur l'ensemble des 751 microsatellites a été réalisé par la méthode « M13 ».

Pour la réalisation des cartes génétiques et l'étude de la distribution des CO, les conditions spécifiques utilisées pour chaque marqueur sont résumés dans le Tableau 14 en Annexes p 335.

3.2.1 Protocole « M13 »

Les amplifications sont réalisées dans un volume final de 6,5 µl (Taq : 0.2 U, amorce gauche avec extension M13 : 50 nM, amorce droite : 500 nM, amorce M13 marquée : 3.4 pmoles, dNTP : 5.4 nmoles) à l'aide d'un programme « Touch down » (95°C pendant 5 min, 7 cycles (95°C pendant 30 s, 62°C (-1°C/cycle) pendant 30 s et 72°C pendant 30 s), 20 cycles (95°C pendant 30 s, 55°C pendant 30 s et 72°C pendant 30 s), 8 cycles (95°C pendant 30 s, 56°C pendant 30 s et 72°C pendant 30 s) et 72°C pendant 5 min.

Les produits d'amplification sont ensuite visualisés par électrophorèse capillaire (ABI PRISM®3100, Applied Biosystems) et analysés grâce aux logiciels GeneScan version 3.7.1 (Applied Biosystems) et GenoTyper version 3.7 (Applied Biosystems) pour les tests de polymorphisme et au logiciel GeneMapper version 3.7 (Applied Biosystems) pour les données servant à la cartographie.

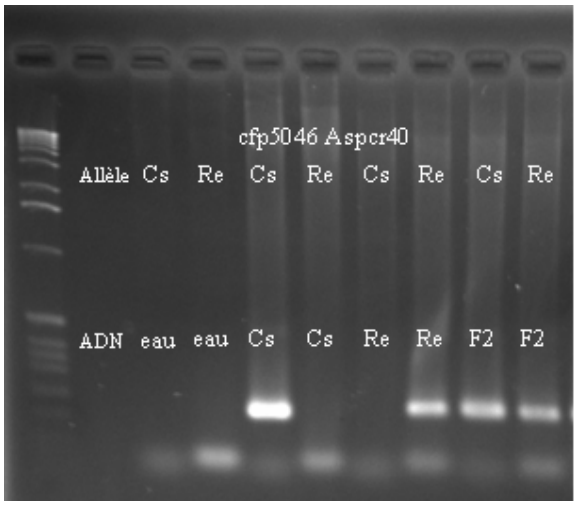


Figure 16 Génotypage des SNP par la technique AS-PCR

L'utilisation de l'amorce cfp5046Aspcr (Cs) spécifique de l'allèle présent chez la variété Chinese spring (Cs) permet une amplification que sur la variété Cs. A l'inverse, l'utilisation de l'amorce cfp5046Aspcr (Re) spécifique de l'allèle présent chez la variété Renan (Re) permet une amplification que sur celle-ci. L'utilisation des deux amorces dans deux réactions différentes sur un individu F2CsRe hétérozygote permet un génotypage co-dominant.

3.2.2 Protocole « rouge de crésol »

Les amplifications sont réalisées dans un volume final de 10 µl (Taq : 0.25 U, amorce droite et gauche : 0.5 µM, rouge de crésol : 1 µM) selon le protocole suivant : 94°C pendant 5 min, 30 cycles (94°C pendant 30 s, 60 °C pendant 30 s et 72°C pendant 30 s) et 72°C pendant 5 min. les produits d'amplifications sont ensuite visualisés sur gel d'agarose 4 % à l'aide du bromure d'éthidium.

3.2.3 Protocole « SYBR Green »

Les amplifications sont réalisées dans un volume final de 10 µl (Taq : 0.25 U, amorce droite et gauche : 0.5 µM, SYBR Green : 1X, chaque dNTP : 0.4 mM et betaïne : 1M) selon la même procédure d'amplification que pour le protocole M13. Les produits PCR sont ensuite visualisés par courbe de dissociation sur un appareil de PCR quantitative (ABI7900HT, Applied Biosystems ; 95°C pendant 15 s, 60°C pendant 15 s et 90°C pendant 15 s) et analysés à l'aide du logiciel SDS v. 2.2.1 (Applied Biosystems).

3.3 Conditions de génotypage des marqueurs SNP

Un total de 9 SNP et 3 indel a été utilisé pour étudier la distribution des CO sur le contig *Rph7*. Deux protocoles ont servi pour réaliser le génotypage des SNP : le protocole de réaction en chaîne allèle-spécifique (AS-PCR pour Allele Specific Polymerase Chain Reaction) et le protocole SNaPshot (Applied Biosystems).

Pour la technique AS-PCR (Figure 16), deux amorces spécifiques, une pour chacun des deux parents, possédant le nucléotide 3' positionné sur le SNP ont été développées (liste des amorces, Tableau 15 en Annexes p 337) afin d'avoir un génotypage de type co-dominant. Ces deux amorces sont ensuite utilisées pour réaliser une amplification PCR avec une troisième amorce classique suivant le protocole rouge de crésol défini précédemment. L'ensemble des marqueurs a d'abord été testé à une température d'hybridation de 60°C puis, suivant les résultats, nous avons fait varier la température d'hybridation pour en obtenir une où chaque amorce est spécifique du parent pour lequel elle porte le nucléotide en position 3'.

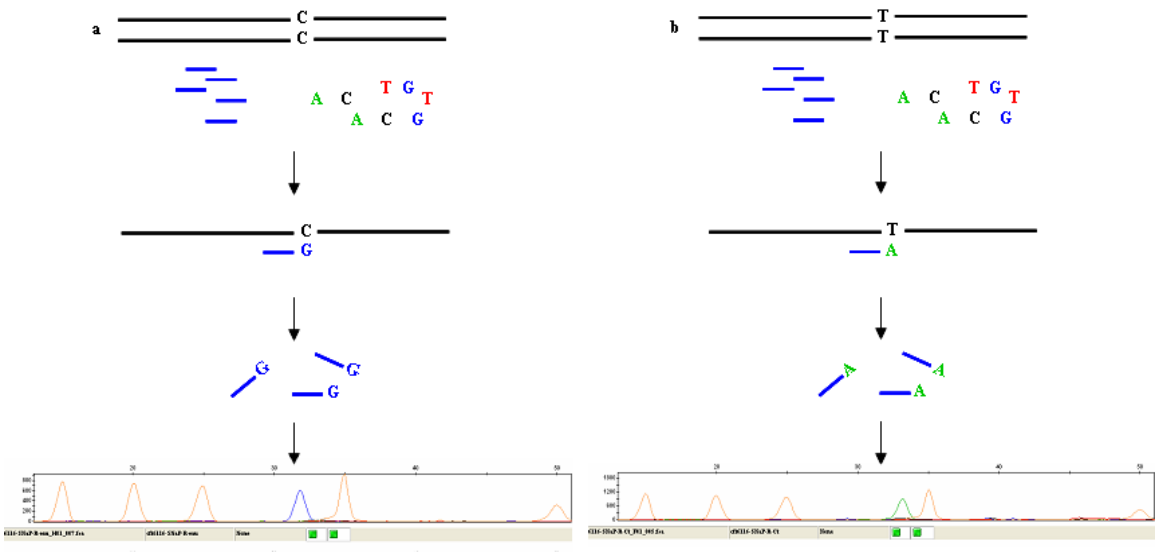


Figure 17 Génotypage des SNP par la technique SNaPshot.

La réaction SNaPshot est réalisée en présence du produit PCR du loci contenant le SNP à génotyper, de l'amorce dont le nucléotide 3' se localise un nucléotide en amont du SNP et des didésoxy-nucléotides, chacun marqué par un fluorochrome de couleur différente. Suite à la fixation de l'amorce, la synthèse d'ADN débute mais elle se limite au premier nucléotide ajouté en raison de la présence des didésoxy-nucléotides. Le premier nucléotide ajouté est celui qui est complémentaire au SNP à génotyper et ainsi pour chaque réaction les amorces sont marquées avec une couleur spécifique de l'allèle de l'individu testé, en bleu par exemple pour l'allèle présent chez la variété Chinese spring (**a**) et en vert pour l'allèle présent chez la variété Renan (**b**). Les amorces sont ensuite purifiées, migrées sur électrophorèse capillaire et analysées grâce au logiciel GeneMapper version 3.7 (Applied Biosystems). Un pic bleu est visualisé pour la variété Chinese spring et un pic vert pour la variété Renan. Deux pics, un vert et un bleu sont visualisés dans le cas d'un individu hétérozygote (donnée non montrée).

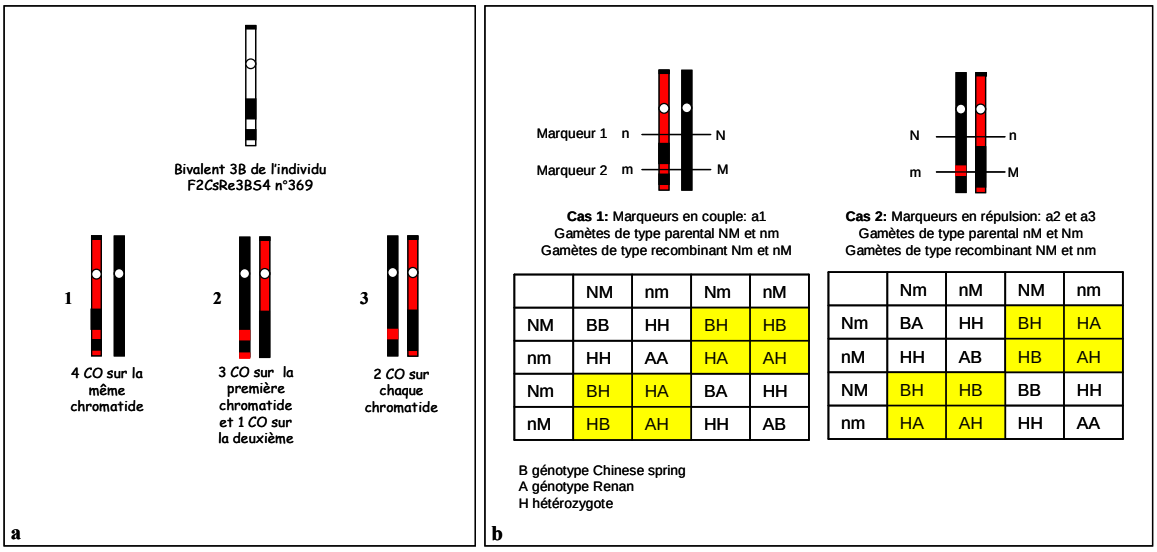
Pour la technique SNaPshot (Figure 17), une amorce dont le nucléotide 3' se localise un nucléotide en amont du SNP à génotyper a été défini. Dans un premier temps, le fragment contenant le SNP est amplifié par PCR selon le protocole rouge de crésol en remplaçant ce dernier par de l'eau. Les produits PCR sont ensuite purifiés (5 µl de produit PCR, 0.2 unités d'enzyme ExoI, 1 unité d'enzyme SAP, 0.9 µl de tampon 10X SAP) et repris dans un volume final de 9 µl, mis à 37°C pendant 1 h puis 80°C pendant 20 min. La réaction SNaPshot est réalisée à l'aide du Kit SNaPshot (Applied Biosystems) dans un volume final de 10 µl (3 µl de produit PCR purifié, 0.75 µl de mix SNaPshot, 1 µl des amorces SNaPshot à 2 µM) suivant le programme : 25 cycles, 96°C pendant 10 s, 50°C pendant 5 s et 60°C pendant 30 s. Les produits d'amplification SNaPshot sont ensuite purifiés par ajout d'une unité d'enzyme SAP et incubés pendant 1 h a 37°C puis pendant 15 min à 65°C. Les produits de purification sont visualisés par électrophorèse capillaire (ABI PRISM®3100, Applied Biosystems) et analysés grâce au logiciel GeneMapper.
Les indel ont été génotypés en suivant le protocole M13 décrit précédemment.

4 Cartographie génétique

Les cartes génétiques ont été construites à l'aide du logiciel Mapmaker (Lander et al. 1987). Les groupes de liaison sont obtenus à l'aide de la fonction « group » en utilisant une valeur de LOD de 3 et une distance maximun de 37.25 cM. La commande « order » est ensuite utilisée pour déterminer l'ordre des marqueurs en utilisant la fonction de cartographie de Kosambi (1944). Lorsque l'ordre des marqueurs n'a pas pu été défini à des valeurs de LOD<2, la fonction « compare » est utilisée en prenant en compte l'ordre le plus vraisemblable. Les fonctions « error-detection » et « genotypes » sont également employées pour déterminer la présence d'erreurs potentielles de génotypage. Les individus responsables de ces erreurs potentielles sont alors vérifiés et éventuellement génotypés de nouveau afin de générer un jeu de données propres et sans ambigüité. Pour les données issues des populations haploïdes doublés, l'ensemble des singletons a été génotypé une deuxième fois.

5 Calcul des fréquences de CO

Les fréquences de CO pour chaque bin de délétion en cM/Mb ont été calculées en divisant la taille génétique de chaque bin en cM par leur taille physique en Mb.



	NM	nm	Nm	nM
NM	BB	HH	BH	HB
nm	HH	AA	HA	AH
Nm	BH	HA	BA	HH
nM	HB	AH	HH	AB

	Nm	nM	NM	nm
Nm	BA	HH	BH	HA
nM	HH	AB	HB	AH
NM	BH	HB	BB	HH
nm	HA	AH	HH	AA

Figure 18 Phase des marqueurs encadrant les régions homozygotes présentes sur le bivalent délété.

a/ Deux régions homozygotes Cs (boîtes noires) sont présentes sur le bras long du bivalent 3B de l'individu F2CsRe3BS4 n°369. Quatre CO ont eu lieu sur les deux chromatides formant les deux chromosomes du bivalent 3B pour expliquer cette alternance de régions homozygotes (noires) et hétérozygotes (blanches). Les deux chromatides de départ formant ce bivalent peuvent ainsi avoir trois configurations différentes (1, 2 et 3). Les régions issues de Renan sont représentées en rouge et celles issues de Cs en noires. Soit les 4 CO ont eu lieu sur la même chromatide (1) soit 3 CO ont eu lieu sur une chromatide et le troisième sur la deuxième (2) soit 2 CO ont eu lieu sur chaque chromatide (3). **b**/ Les marqueurs 1 et 2 portant respectivement les allèles N et M pour le génotype Cs et n et m pour le génotype Re et encadrant une région homozygote peuvent ainsi être en couple (cas 1) si les 4 CO ont eu lieu sur la même chromatide ou alors en répulsion (cas 2) dans les situations a 2 et a 3. Dans le cas 1, les gamètes parentaux sont de type NM et nm et les gamètes recombinants de type Nm et nM et inversement dans le cas 2. Ainsi, dans le tableau de gamètes, les individus parentaux dans le cas 1 correspondent aux individus possédant deux CO dans le cas 2. Seuls les individus possédant un seul CO et de type BH, HB, AH et HA ont des fréquences identiques dans les deux cas (cases en jaune).

La taille génétique du bin a été calculée en prenant la distance génétique entre les deux marqueurs appartenant au bin les plus éloignés sur la carte génétique. Comme la jonction entre les bins de délétion sur la carte génétique est difficile à déterminer car les deux marqueurs situés au niveau de cette jonction et appartenant à deux bins différents sont séparés par de la distance génétique. Pour prendre en compte ces distances génétiques au niveau des jonctions, la distance présente entre les deux marqueurs appartenant à deux bins différents a été divisée en deux et ainsi la moitié a été additionnée à la distance génétique de chacun des deux bins voisins. Pour le calcul des fréquences de CO des différents intervalles du contig *Rph7*, la taille génétique entre deux marqueurs délimitant un intervalle a été calculée par la fonction de Morgan en divisant le nombre de recombinants par le nombre de gamètes analysés. La taille génétique est ensuite divisée par la taille physique de l'intervalle connue à la base près.

6 Comparaison des taux de CO

6.1 Calcul du taux de CO

Les taux de CO *r* utilisés pour les comparaisons entre méioses mâle et femelle, entre les populations CsRe et CsCt et entre le bivalent 3B homozygote délété et le bivalent comportant un chromosome entier et un délété ont été calculés en utilisant la fonction inverse de Kosambi (1944) à partir des distances génétiques obtenues grâce au logiciel Mapmaker. Seul le calcul des valeurs de *r* au sein des régions homozygotes a été réalisé de manière différente.

6.2 Calcul du taux de CO pour les régions homozygotes Cs sur le bivalent 3B délété

Le calcul des valeurs de *r* pour les régions homozygotes a été réalisé en utilisant deux marqueurs situés de part et d'autre de cette région mais dans les régions hétérozygotes flanquantes. Cependant, la phase (répulsion ou couple) de ces deux marqueurs n'est pas connue (Figure 18). Les valeurs de *r* ont ainsi été déterminées en utilisant uniquement les individus possédant un seul CO entre les deux marqueurs étudiés car les valeurs de *r* ne sont pas affectées dans ce cas précis (Figure 18).

De cette façon, les valeurs de *r* ont été calculées à partir du tableau de gamètes par l'équation ci-dessous :

$(1-r/2)(r/2)*8 = n/N$

$(r-r^2)*2 = n/N$

$r^2 - r + (n/2N) = 0$

où *r* représente le taux de CO entre les marqueurs 1 et 2 et (1-r/2) la fréquence des gamètes d'un type parental et (r/2) la fréquence d'un type de gamètes recombinants. n le nombre d'individus possédant un seul CO entre les deux marqueurs (Génotype de type AH, HA, BH, HB, Figure 18) et N le nombre total d'individus de la population utilisée. La résolution de cette équation du second degré conduit à la valeur de *r*.

6.3 Comparaison des taux de CO

Nous avons utilisé la méthode des *p*-value pour déterminer la significativité des différences locales de taux de CO observées entre le mâle et la femelle et entre la population réalisée avec le chromosome entier et celle obtenue avec le chromosome délété. Sous l'hypothèse H_0 (pas de différence entre les taux de CO), des données simulées peuvent être obtenues pour le mâle et la femelle mais aussi pour le chromosome entier et le chromosome délété. La significativité de la différence observée entre les données théoriques et les données expérimentales est ainsi mesurée par le calcul d'une *p*-value. Le test de comparaison multiple a été effectué par la méthode de Bonferroni (Hocheberg 1988).

7 Etude de l'interférence

Pour l'étude de l'interférence, nous avons collaboré avec Matthieu Falque (INRA du Moulon) et Olivier Martin (Université Paris-sud) qui ont réalisé l'ensemble des analyses. Le détail des analyses statistiques est présent dans la partie supplémentaire (SUPPLEMENTARY MATERIAL 1, SM1, p 375) de l'article (Saintenac et al. 2009; Partie IX p 351).

7.1 Test de l'hypothèse H_0 de non-interférence

L'hypothèse H_0 (aucune différence entre la distribution des CO observée et leur distribution théorique) a été testée en utilisant les distances entre deux CO. Les distances obtenues expérimentalement ont été comparées avec celles obtenues théoriquement sous une distribution de Poisson nous permettant d'obtenir les *p*-value associées.

7.2 Le modèle gamma

Le modèle gamma (Broman et Weber 2000; Mcpeek et Speed 1995) est une extension du modèle de comptage (Stahl et al. 2004). Le paramètre *m* est associé au modèle gamma qui indique l'intensité de l'interférence. Quand *m*=0, l'interférence est nulle alors que quand *m* tend vers l'infini, l'interférence devient très forte. Pour différentes valeurs de *m* (augmentant par palier de 0.1), 10^7 gamètes sont simulés par le modèle. Les gamètes possédant deux événements de recombinaison sont conservés et l'ensemble des distances obtenues entre les deux événements est groupés en sous-ensemble. La fréquence de chaque sous-ensemble est estimée pour chaque valeur de *m* et comparée aux valeurs expérimentales. La valeur de *m* est ainsi obtenue lorsque les valeurs théoriques et expérimentales sont les plus proches.

7.3 Intervalle de confiance du paramètre m

Nous avons simulé 10^5 populations de taille identique aux populations CsCt (182 HD pour la population mâle et 185 HD pour la population femelle) avec le modèle gamma. Pour chacune d'entre elles, le modèle gamma a été appliqué pour déterminer la valeur du paramètre *m* comme décrite ci-dessus. Ainsi, 10^5 valeurs de *m* sont obtenues pour chacune des populations permettant une estimation de l'intervalle de confiance de *m*.

7.4 Différence entre le paramètre *m* issu du mâle et celui issu de la femelle

Pour tester si la différence observée entre les valeurs de *m* obtenue pour le mâle (m^*_m) et la femelle (m^*_f) est significative, nous avons utilisé la méthode dite des bootstraps. Les gamètes issus de la voie mâle et de la voie femelle sont groupées et 10^5 nouvelles populations mâle et femelle de tailles identiques aux expérimentales ont été générées en mélangeant gamètes mâles et femelles aléatoirement. Pour chacune des populations, les valeurs m_f associées à la population femelle et m_m associées à la population mâle sont calculées. La méthode des bootstraps permet ainsi de calculer la *p*-value associée à l'hypothèse H_0 (pas de différence significative d'interférence entre le mâle et la femelle) correspondant ainsi à la fraction de paires (m_f et m_m) pour lesquelles $|m_m - m_f| > |m^*_m - m^*_f|$.

7.5 Le coefficient de coïncidence

L'étendue de l'interférence a été mesurée en utilisant le coefficient de coïncidence (*C*) qui est le rapport entre la fréquence de doubles CO observés sur la fréquence de doubles CO théoriques en absence d'interférence. Si l'on considère deux intervalles non chevauchant AB et CD encadrés respectivement par les loci A, B et C, D, la valeur du coefficient de coïncidence est :

$$C(AB, CD) = f(AB,CD) / r_{AB} \times r_{CD}$$

où r_{AB} et r_{CD} représentent les taux de CO pour les intervalles AB et CD et *f(AB,CD)* la fréquence théorique de doubles CO, un dans chaque intervalle. Lorsque l'interférence est totale *C* est égal à 0 alors qu'en absence d'interférence, *C* est égal à 1. Pour étudier l'étendue de l'interférence, nous avons représenté graphiquement la distance *d* entre le milieu des deux intervalles AB et CD en fonction des valeurs du coefficient de coïncidence. Pour éviter de biaiser le coefficient de coïncidence, nous avons utilisé uniquement des intervalles ayant une valeur de *r* supérieure à 0.05. Pour restreindre le nombre de valeurs de *C*, nous avons groupé les distances *d* en 30 groupes et réalisé une moyenne de la valeur de *C* pour chacun des groupes. Les barres d'erreurs ont été calculées en divisant la déviation standard par la racine carrée du nombre de valeurs de *C* utilisée dans chaque groupe.

8 Analyse des bandes C du chromosome 3B

Le protocole détaillé de la coloration des bandes C est fourni en Annexes (Protocole 1 p 341). Les graines des lignées Cs, Re et de la descendance de l'individu F2CsRe3BS4 n°369 sont mises à germer sur boite de pétri. Après plusieurs bains successifs dans de l'acide acétique (45 %) puis dans de l'acide chlorhydrique (HCl) à 0.2 M, les racines sont mises en contact de cellulysine à 45 U/ml (CALBIOCHEM, ref 219466) pendant une nuit à 4°C. Les méristèmes sont ensuite hachés finement à l'aide d'une lame de rasoir sur une lame de verre en présence d'acide acétique 45 %. Après mise en place de la lamelle sur les échantillons, la lame de verre est trempée dans de l'azote liquide, débarrassée de la lamelle et mis dans l'alcool 100 % pendant une nuit. La lame est ensuite trempée dans une solution de baryte saturée (Merck ref 1.01735.1000) à environ 220 µM pendant 6 min puis dans du HCl à 0.2 M pendant 15 s. La lame est ensuite trempée dans une solution de citrate saline de sodium (SSC, pour Sodium Saline Citrate contenant 300 µM de NaCl et 30 µM de tri sodium de citrate) pendant 60 min.

La coloration des bandes est finalement réalisée par trempage dans une solution de Giemsa (TrisHCl à 0.125 M pH=6.8 et 3 % (vol/vol) de colorant Giemsa, Merck ref 1.09204.0500). La lame est ensuite visualisée en microscopie photonique.

9 Analyse des séquences

9.1 Annotation des séquences

Sur les trois contigs analysés, les séquences du contig *Rph7*, du contig *Sr2* et du contig *L* sont issues de la variété Cs. Les contigs *Rph7* et *L* ont été assemblés et annotés au sein de notre équipe (F. Choulet) alors que le contig *Sr2* a été assemblé par une autre équipe (Thomas Wicker, Université de Zurich, Suisse).

Brièvement, les séquences contenant des éléments transposables ont tout d'abord été masquées grâce au logiciel RepeatMasker qui réalise un BLAST à partir des bases de données d'éléments transposables. Les séquences non masquées sont ensuite analysées à l'aide de logiciels de recherche de gènes et les gènes prédits sur les séquences sont vérifiés par BLAST avec l'ensemble des EST présentes dans les bases de données des poacées. Enfin, l'annotation manuelle permet l'annotation fine des différents éléments présents sur les contigs.

9.2 Calcul des valeurs des différents paramètres de séquences et des coefficients de détermination R^2

La densité en gènes pour chaque intervalle a été calculée en nombre de gènes pour 150 kb. Le pourcentage de la taille physique cumulée des gènes et des pseudogènes pour un intervalle donnée a été calculé en additionnant la taille en pb de l'ensemble des gènes et des pseudogènes présents dans l'intervalle en la divisant par la taille en pb de l'intervalle entier et en la multipliant par 100. Ces valeurs ont été déterminées pour la taille physique cumulée des exons + introns des gènes et des pseudogènes, pour la taille physique cumulée des exons + introns des gènes seulement, pour la taille physique cumulée des parties codantes (exons) des gènes et pseudogènes et pour la taille physique cumulée des parties codantes (exons) des gènes seulement. La taille physique en pb des exons et introns a été définie à la base près grâce à l'annotation fine du contig.

Le pourcentage en GC de chaque intervalle a été calculé en divisant le nombre de nucléotides G + C au sein de l'intervalle par le nombre total de nucléotides de l'intervalle et en multipliant par 100. Le ratio CpG a été calculé en divisant la fréquence de di-nucléotides CG (nombre de di-nucléotides CG dans l'intervalle divisé par nombre de di-nucléotides possibles dans l'intervalle) par le produit de la fréquence de nucléotides C multipliée par la fréquence de nucléotide G (nombre total de nucléotides G dans l'intervalle divisé par le nombre total de nucléotides dans l'intervalle).

Le pourcentage d'éléments transposables a été calculé en divisant la taille physique en pb couverte par l'ensemble des éléments transposables dans l'intervalle par la taille en pb de l'intervalle entier et multiplié par 100. Le pourcentage de séquences répétées en tandem a été calculé en divisant la taille physique en pb de l'ensemble des séquences répétées en tandem dans l'intervalle par la taille physique de l'intervalle et multiplié par 100.

Les différentes valeurs du coefficient de détermination R^2 ainsi que le coefficient de corrélation R ont été calculées grâce au logiciel Excel (2003) et leur significativité a été évaluée grâce à la table de Bravais-Pearson (Tableau 16 en Annexes p 339).

IV Résultats

1 Distribution des CO à l'échelle du chromosome entier

La distribution des CO sur le chromosome 3B (995 Mb) a été étudiée en comparant une carte génétique dense avec une carte physique de délétion mais également en comparant différentes cartes génétiques pour déterminer si la distribution observée à l'échelle du chromosome entier est conservée entre différentes populations.

1.1 Développement et analyse de marqueurs moléculaires

Afin de disposer d'un nombre suffisant de marqueurs pour étudier finement la distribution des CO sur le chromosome 3B, nous avons répertorié tous les marqueurs disponibles issus des bases de données publiques et privées, de la banque BAC du chromosome 3B (Paux et al. 2006; Paux et al. 2008) et des EST de blé (Zhang 2006). Nous avons sélectionné 751 marqueurs microsatellites ou SSR (Simple Sequence Repeat) pouvant potentiellement être cartographiables sur le chromosome 3B.

Ces marqueurs ont tout d'abord été testés sur les 14 lignées de délétion du chromosome 3B pour déterminer leur localisation sur le chromosome. Ces 14 lignées de délétion définissent 16 bins de délétion, 9 sur le bras court et 7 sur le bras long dont la taille physique varie de 4 Mb (3BS4-0.55-0.56 et 3BS2-0.56-0.57) à 208 Mb (3BL7-0.63-1.00) avec une taille moyenne de 62 Mb (Tableau 17 en Annexes p 347, Figure 20 p 166). Le découpage du bras court est relativement homogène (taille moyenne des bins 48 Mb) contrairement à celui du bras long qui est moins régulier à cause de la présence du bin 3BL7-0.63-1.00 représentant à lui seul plus du tiers de ce bras (37 %) et plus de 20 % du chromosome entier.

Simultanément, le polymorphisme de ces marqueurs SSR a été testé systématiquement entre les variétés Chinese spring (Cs), Renan (Re) et Courtot (Ct), variétés qui ont été utilisées pour réaliser les populations de cartographie CsRe et CsCt utilisées par la suite. Les marqueurs issus de la banque BAC 3B ont de plus été évalués sur les lignées Opata (Op) et W7984 (Sy), parents de la population de référence internationale OpSy (encore appelée population ITMI pour International Triticeae Mapping Initiative).

		CsRe				CsCt			
Bin de délétion	**Ass**	**Pol**	**N pol**	**DM**	**% pol**	**Pol**	**N pol**	**DM**	**% pol**
3BS3-0.87-1.00	3	3	0	0	100	3	0	0	100
3BS8-0.78-0.87	36	23	7	6	77	24	6	6	80
3BS7-0.75-0.78	6	2	3	1	40	3	2	1	60
3BS9-0.57-0.75	13	6	4	3	60	3	7	3	30
3BS2-0.56-0.57	4	1	1	2	50	1	1	2	50
3BS4-0.55-0.56	4	2	1	1	67	1	2	1	33
3BS1-0.33-0.55	12	8	1	3	89	3	5	4	37
3BS5-0.07-0.33	8	4	2	2	67	3	3	2	50
Cent-3BS5-0.07	6	5	1	0	83	1	5	0	17
Cent-3BL2-0.22	4	4	0	0	100	2	1	1	67
3BL2-0.22-0.28	5	1	2	2	33	2	1	2	67
3BL8-0.28-0.31	3	1	2	0	33	0	3	0	0
3BL1-0.31-0.38	8	4	2	2	67	2	4	2	33
3BL9-0.38-0.50	3	1	2	0	33	3	0	0	100
3BL10-0.50-0.63	13	7	1	5	88	7	4	2	63
3BL7-0.63-1.00	**89**	39	37	13	51	46	29	14	61
Total	217	111	66	40	**63**	104	73	40	**59**

Tableau 5 Assignation et polymorphisme de 217 marqueurs SSR provenant de données publiques et privées et présents sur les cartes génétiques du chromosome 3B.

Ass = assignable ; Pol = polymorphe ; N pol = non polymorphe ; DM = données manquantes ; % pol = pourcentage de marqueurs polymorphes parmi les marqueurs assignables (données manquantes prises en compte) ; CsRe = Chinese spring x Renan ; CsCt = Chinese spring x Courtot.

1.1.1 Microsatellites issus des bases de données publiques et privées

Sur les 343 marqueurs SSR issus des bases de données publiques et privées dont un locus avait déjà été cartographié sur le chromosome 3B, 36 n'ont donné aucun produit d'amplification. Sur les 307 restants, 217 ont pu être assignés sans ambiguïté au chromosome 3B. Les 90 marqueurs SSR restants présentaient une amplification sur toutes les lignées de délétion du chromosome 3B suggérant la présence d'autres loci sur les chromosomes homéologues 3A et 3D ou sur une autre région du génome. Entre 3 et 89 marqueurs ont été assignés à chaque bin de délétion, la majorité des marqueurs (89/217) ayant été assignée au bin de délétion 3BL7-0.63-1.00 qui est le plus grand (Tableau 5).

Les tests de polymorphisme indiquent qu'environ 60 % des marqueurs assignables sont polymorphes entre les variétés Cs et Re et entre Cs et Ct. De plus, nous avons remarqué que les taux de polymorphisme entre Cs et Re ou Cs et Ct diffèrent en fonction des régions du chromosome 3B mais également pour une même région en fonction de la population CsRe ou CsCt. Ainsi, parmi les marqueurs assignables aux bins de délétion, le taux de polymorphisme entre les variétés Cs et Re est de 51 % pour le bin 3BL7-0.63-1.00 ou 60 % pour le bin 3BS9-0.57-0.75 alors qu'il atteint 77 % pour le bin 3BS8-0.78-0.87. Cependant, entre les variétés Cs et Ct, le taux de polymorphisme est de 61 % pour les marqueurs assignables au bin 3BL7-0.63-1.00.

1.1.2 Microsatellites issus de la banque BAC du chromosome 3B

Nous avons également testé 224 marqueurs SSR issus des 11 Mb de séquence obtenus suite au séquençage de 20 000 extrémités de BAC de la banque du chromosome 3B (Paux et al. 2006). Un total de 55 marqueurs n'a donné aucun produit d'amplification et parmi les 169 autres, 41 % soit 70 ont été assignés au chromosome 3B (Tableau 6). Le pourcentage de marqueurs polymorphes parmi les marqueurs assignables est respectivement de 42, 41 et 30 % pour les populations ITMI, CsRe et CsCt. Le nombre de marqueurs assignés à chaque bin étant faible, nous n'avons pas pu étudier les taux de polymorphisme en fonction des 16 bins de délétion.

		ITMI			CsRe			CsCt		
Bin de délétion	**Ass**	**Pol**	**N pol**	**DM**	**Pol**	**N pol**	**DM**	**Pol**	**N pol**	**DM**
3BS3-0.87-1.00	4	3	1	0	2	2	0	2	2	0
3BS8-0.78-0.87	8	4	2	2	4	4	0	2	6	0
3BS9-0.57-0.75	3	3	0	0	3	0	0	1	2	0
3BS1-0.33-0.55	9	1	8	0	3	6	0	2	7	0
3BS5-0.07-0.33	1	0	1	0	0	1	0	0	1	0
Cent-3BS5-0.07	4	1	3	0	0	4	0	0	4	0
Cent-3BL2-0.22	1	0	1	0	1	0	0	0	1	0
3BL1-0.31-0.38	4	2	2	0	1	3	0	2	2	0
3BL9-0.38-0.50	3	1	1	1	2	1	0	1	1	1
3BL10-0.50-0.63	9	6	3	0	5	3	1	2	7	0
3BL7-0.63-1.00	24	7	16	1	7	17	0	9	15	0
Total	70	28	38	4	28	41	1	21	48	1

Tableau 6 Assignation et polymorphisme de 70 marqueurs SSR issus des séquences d'extrémités de BAC de la banque du chromosome 3B.

Ass = assignable ; Pol = polymorphe ; N pol = non polymorphe ; DM = données manquantes ; ITMI = population de référence internationale ; CsRe = Chinese spring x Renan ; CsCt = Chinese spring x Courtot.

Cependant, le découpage du chromosome en trois grandes régions de l'extrémité distale du bras court à la jonction 3BS9-0.57, de la jonction 3BS9-0.57 à la jonction 3BL10-0.50 et de la jonction 3BL10-0.50 à l'extrémité du bras long montre que le taux de polymorphisme des marqueurs assignables est respectivement de 67, 23 et 40 %. Ainsi, le taux de polymorphisme apparaît plus important dans les régions distales que dans les régions proximales.

Sur l'ensemble des marqueurs testés, le taux de marqueurs assignables au bin de délétion du chromosome 3B et polymorphes est de 12.5 % pour les populations ITMI et CsRe et de 9 % pour la population CsCt. Cela signifie que pour obtenir un marqueur SSR assignable et utilisable en cartographie, il faut développer huit marqueurs en moyenne. De plus, ces données montrent que le taux de polymorphisme observé sur le chromosome 3B au sein de la population de référence internationale (ITMI) est le même que celui observé au sein de la population CsRe.

1.1.3 Microsatellites issus d'EST

Finalement, nous avons testé 184 marqueurs SSR dérivés d'EST dont 93 étaient issus du chromosome 1 de riz orthologue du chromosome 3B et 91 de blé (Zhang 2006). Parmi les 93 SSR issus d'EST de riz, 39 ne présentent pas de produit d'amplification et seulement 3 ont pu être assignés au chromosome 3B dont un polymorphe sur la population CsRe et un sur la population CsCt. Parmi les 91 issus d'EST de blé, 23 ne présentent pas de produit d'amplification et 27 sont assignables au chromosome 3B. De plus, parmi ces marqueurs assignables seulement 4 soit 15 % sont polymorphes sur la population CsRe et 4 autres sont aussi polymorphes sur la population CsCt. Les microsatellites dérivés d'EST montrent ainsi des taux de polymorphisme faibles comparés aux autres microsatellites.

1.1.4 Conclusion

L'ensemble de ces tests a permis d'obtenir un total de 144 marqueurs SSR assignables aux bins de délétion du chromosome 3B et cartographiables sur la population CsRe et 130 SSR du même type sur la population CsCt pour permettre une analyse de la distribution des CO sur le chromosome 3B.

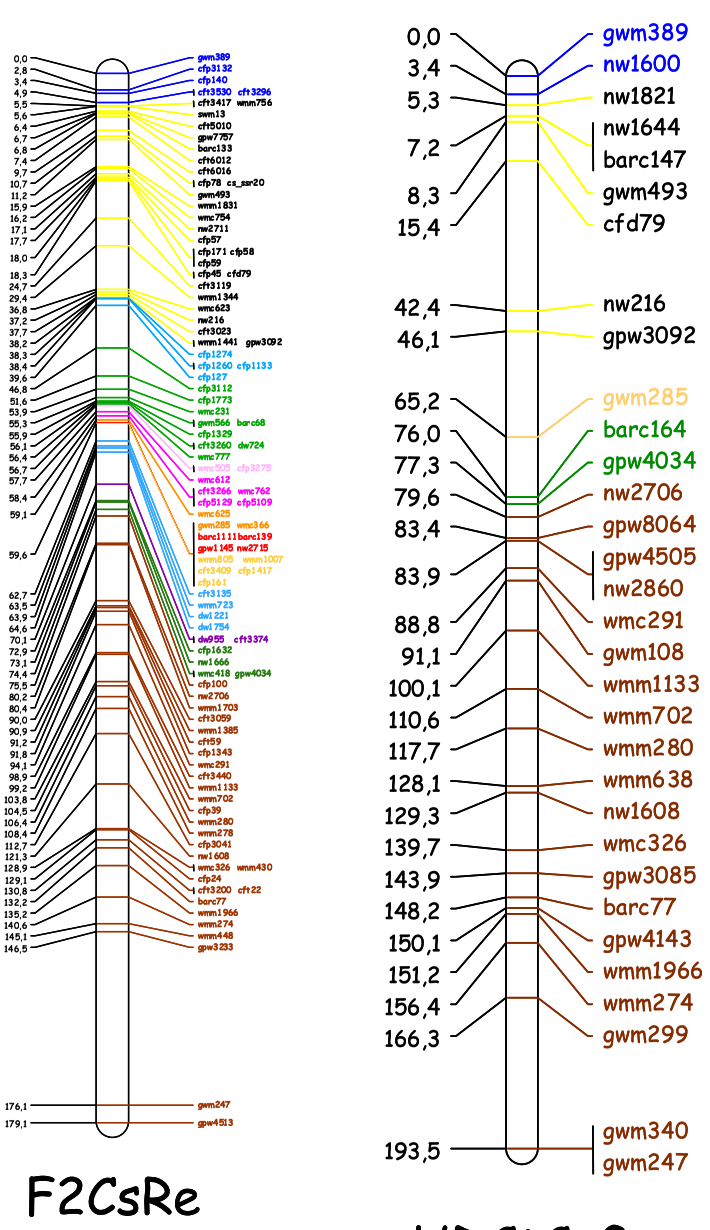


Figure 19 Cartes génétiques du chromosome 3B obtenues à partir de différentes populations.

De gauche à droite : 376 F2 Chinese spring x Renan (F2CsRe), 185 haploïdes doublés Courtot x Chinese spring obtenus par gynogenèse (HDCtCsG), 182 haploïdes doublés Courtot x Chinese spring obtenus par androgenèse (HDCtCsA) et 239 haploïdes doublés obtenus par androgenèse (HDARe). La couleur de chaque marqueur indique son assignation à un bin de délétion du chromosome 3B (code couleur Figure 20).

1.2 Différentes cartes génétiques développées pour l'étude de la distribution des CO

1.2.1 La carte Chinese spring x Renan (CsRe)

Une carte génétique CsRe a été réalisée à partir d'une population de 376 individus F2 issue du croisement entre la variété de référence internationale Chinese Spring et la variété française Renan. Cette carte ayant été développée pour l'étude fine de la distribution des CO par comparaison avec la carte physique de délétion (paragraphe 1.3 p 167), nous avons maximisé la couverture du chromosome 3B en utilisant un total de 102 marqueurs moléculaires dont 76 microsatellites et 26 ISBP issus du développement de la carte physique du chromosome 3B (Insertion Site Based Polymorphism, Paux et al. 2006). Aucune distorsion de ségrégation n'a été observée pour les 102 marqueurs utilisés (khi2 $\alpha = 0.05$). La longueur totale de la carte est de 179.1 cM (Figure 19) en utilisant la fonction de Kosambi (1944), ce qui représente une densité de un marqueur tous les 1.75 cM. Une forte concentration de marqueurs est observée au niveau de la région centromérique avec 11 marqueurs co-ségrégeants (d = 0 cM). Au niveau des parties distales, les distances entre marqueurs sont beaucoup plus importantes (Figure 19). Nous observons, notamment, une région de 30 cM entre les marqueurs gpw3233 et gwm247 situés en partie distale du bras long.

1.2.2 Les cartes Chinese spring x Courtot (CsCt)

Nous avons également développé des cartes génétiques à partir de deux populations haploïdes doublés (HD) issues du croisement entre les variétés Chinese Spring et Courtot. Celle réalisée par androgenèse (CsCtA) représente la méiose mâle et celle réalisée par gynogenèse (CsCtG) la méiose femelle. Ces cartes ont ainsi été utilisées pour étudier les différences de taux de CO entre méiose mâle et femelle (paragraphe 1.4 p 171) et l'interférence (paragraphe 1.6 p 173). Un total de 32 marqueurs SSR communs entre les deux populations HD a été cartographié. Aucune distorsion de ségrégation n'a été observée pour les 32 marqueurs utilisés (khi2 $\alpha = 0.05$). Les cartes mâle et femelle ont des tailles respectives très similaires de 190.1 et 193.5 cM et des densités de un marqueur tous les 6 cM pour les deux cartes.

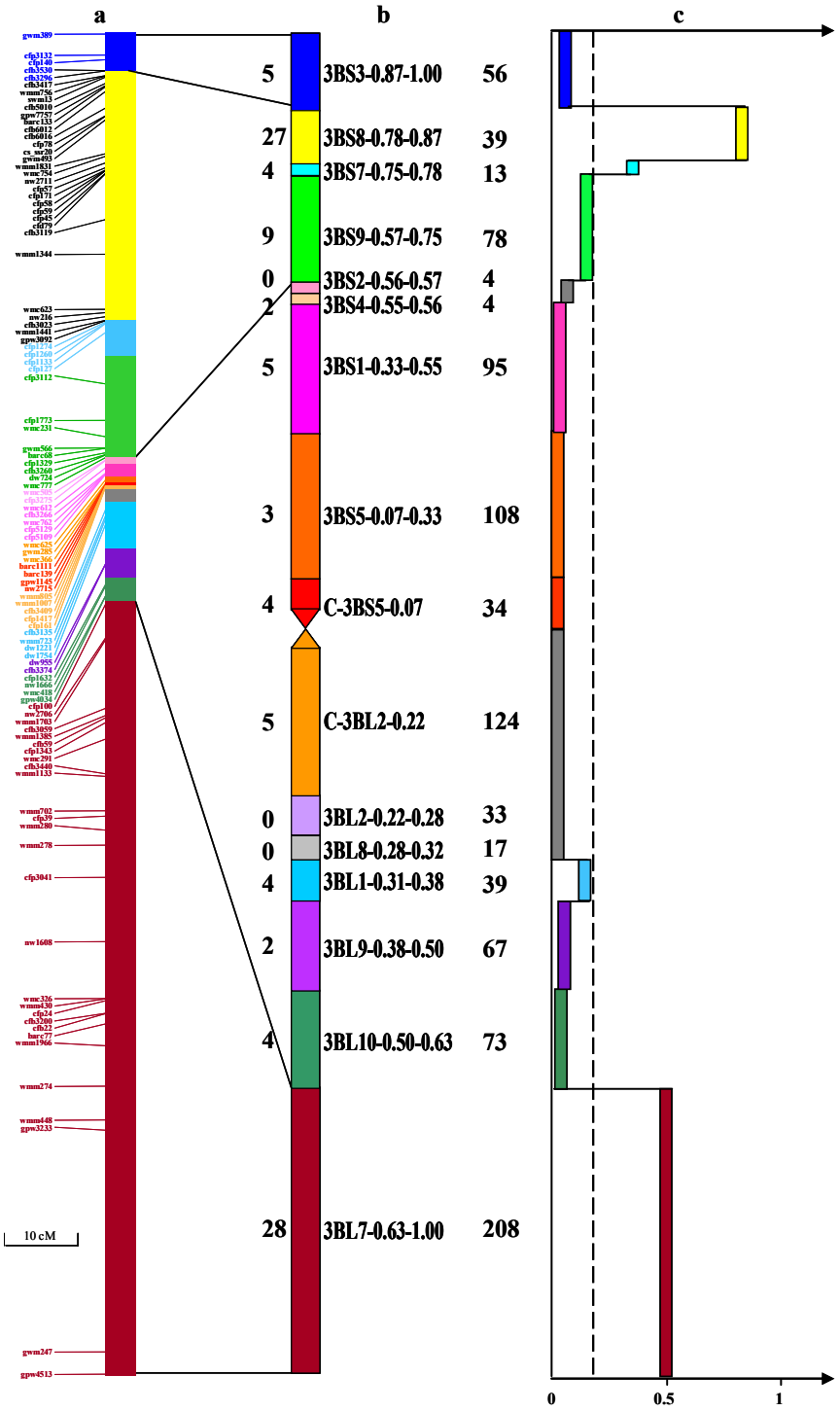


Figure 20 Distribution des CO sur le chromosome 3B.

a/ Carte génétique obtenue à partir d'une population de 376 individus F2CsRe. La couleur des marqueurs représente leur assignation à un bin de délétion. **b/** Carte physique de délétion avec le nom de chacun des bins de délétion, leur taille en Mb à droite et une couleur associée. A gauche est indiqué le nombre de marqueurs utilisés et présents dans chaque bin. **c/** Fréquence de CO estimée en cM/Mb. La ligne en pointillés représente la fréquence moyenne de CO sur le chromosome entier.

1.3 Distribution des CO sur le chromosome 3B : Comparaison de la carte génétique CsRe avec la carte physique de délétion

1.3.1 La carte physique de délétion

L'assignation des 102 marqueurs présents sur la carte F2CsRe a permis la réalisation de la carte physique de délétion. Entre 0 et 28 marqueurs sont présents dans chaque bin. Le mieux couvert étant le bin 3BL7-0.63-1.00. La densité des marqueurs par bin varie de un marqueur tous les 1.4 Mb pour le bin 3BS8-0.78-0.87 jusqu'à un marqueur tous les 36 Mb au sein du bin 3BS5-0.07-0.33 ; la moyenne sur le chromosome étant de un marqueur tous les 9.7 Mb (Tableau 17 en Annexes p 347). La comparaison de cette carte physique de délétion avec la carte génétique CsRe montre une très bonne corrélation, c'est-à-dire que les marqueurs appartenant au même bin de délétion sont liés sur la carte génétique et l'ordre des bins de délétion est respecté sur la carte génétique (Figure 20).

1.3.2 Distribution des CO

Nous avons mesuré la fréquence de CO en cM/Mb pour chaque bin de délétion en divisant la distance génétique en cM, couverte par les marqueurs assignés à ce bin, par sa taille physique en Mb (Figure 20, Tableau 17 en Annexes p 347). Les fréquences de CO varient de 0 cM/Mb dans les régions centromériques (bins : C-3BS5-0.20, C-3BL2-0.22) à 0.85 cM/Mb dans le bin 3BS8-0.78-0.87, avec une moyenne de 0.14, 0.21 et 0.18 cM/Mb pour le bras court, le bras long et le chromosome entier respectivement.

Les CO se localisent majoritairement dans les régions physiques distales qui couvrent 40 % de chaque bras du chromosome. Notamment, 77 % des CO du chromosome se localisent dans les bins 3BS8-0.78-0.87 et 3BL7-0.63-1.00 alors qu'ils ne couvrent physiquement que 25 % du chromosome. Aucun CO n'est présent dans les bins les plus proximaux (C-3BL2-0.22 et C-3BS5-0.07) et seulement trois sont observés dans le bin 3BS5-0.07-0.30.

Sur le bras court, nous observons un gradient de la fréquence de CO du centromère jusqu'au télomère avec un pic dans le bin 3BS8-0.78-0.87. Cependant, la fréquence de CO chute brusquement dans le bin de délétion le plus distal (3BS3-0.87-1.00) pour atteindre une valeur de 0.1 cM/Mb inférieure à la moyenne du bras et du chromosome.

Sur le bras long, 87 % des CO (58 % pour le chromosome entier) se localisent dans le bin le plus distal (3BL7-0.63-1.00) qui couvre 208 Mb. La taille importante de ce bin limite les analyses plus fines au niveau de la partie distale du bras long. Nous observons également, une augmentation de la fréquence de CO au milieu du bras long au niveau du bin 3BL1-0.31-0.38.

1.3.3 Ce pattern de distribution des CO est conservé sur les autres populations

Pour les populations haploïdes doublés CsCtA et CsCtG, nous observons le même pattern de distribution que celui mentionné pour la population F2CsRe. Ainsi, 74 % des CO pour la population CsCtA (haploïdes mâles) et 81 % des CO pour la population CsCtG (haploïdes femelles) se localisent dans les deux bins 3BS8-0.78-0.87 et 3BL7-0.63-1.00 (Figure 19). Nous observons également sur ces deux populations la chute de la fréquence de CO au niveau de la région la plus distale du bras court (bin 3BS3-0.87-1.00).

De plus, nous avons exploité les données acquises dans le cadre d'un autre programme à partir d'un croisement entre les variétés Arche et Récital (ARe) faisant intervenir deux parents différents des trois utilisés pour les populations précédentes pour écarter un effet potentiel de Chinese Spring. Le manque de polymorphisme sur la partie distale du bras long nous a cependant contraints à étudier seulement la distribution des CO sur le bras court. Nous observons le même pattern que celui obtenu sur les autres populations avec une grande majorité des CO localisé dans le bin 3BS8-0.78-0.87 et une chute de la fréquence dans le bin le plus distal (Figure 19).

1.3.4 Conclusion

Pour conclure, nous pouvons dire que quelle que soit la population utilisée CsRe, CsCt ou ARe, à l'échelle du chromosome entier, la distribution des CO est conservée sur le chromosome 3B. La grande majorité des CO est localisée dans les bins 3BS8-0.78-0.87 et 3BL7-0.63-1.00 couvrant 247 Mb. De plus, la région couvrant les 27 % du chromosome qui entoure le centromère est quasi dépourvue de CO. La forte chute de fréquence de CO observée en partie distale du bras court et l'augmentation observée en partie intermédiaire du bras long montrent que la fréquence de CO n'augmente pas de manière uniforme des centromères jusqu'aux télomères.

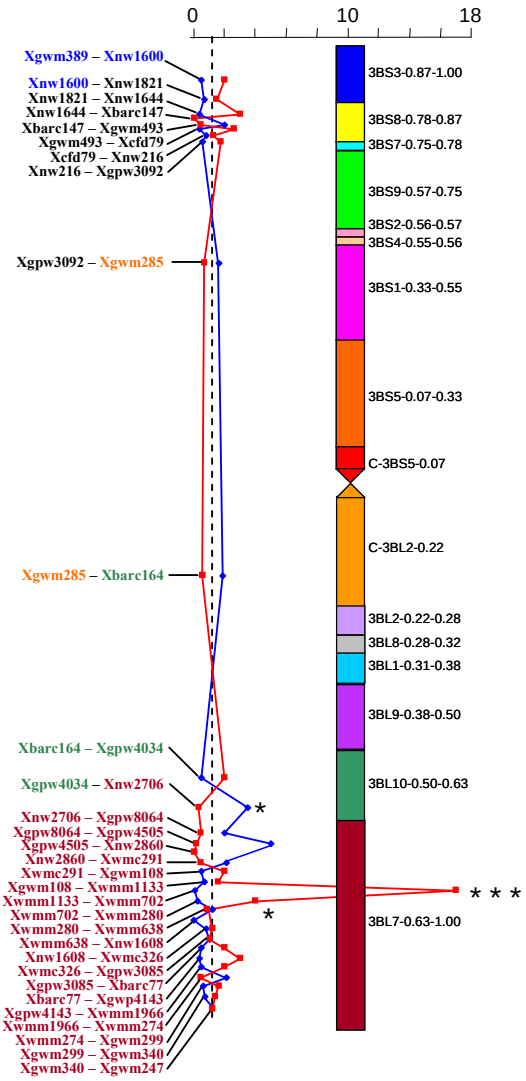


Figure 21 Différence locale du taux de CO entre la méiose mâle et femelle sur le chromosome 3B.

Les 31 intervalles étudiés sont représentés à gauche avec le code couleur (Figure 20) pour l'assignation des marqueurs aux bins de délétion. La position des 31 intervalles sur la carte physique de délétion est arbitraire et a été calculée en divisant le nombre de marqueurs dans chaque bin par la taille physique du bin. La courbe en rouge indique le rapport entre le nombre de CO issu de la méiose femelle sur le nombre de CO issu de la méiose mâle alors que la courbe bleue indique le rapport inverse. Les astérisques indiquent des différences significatives (* $\alpha > 0.05$; *** $\alpha > 0.001$). La ligne en pointillés indique un rpport de 1.

1.4 Comparaison des cartes génétiques CsCtG et CsCtA : Différence de taux de CO en fonction du sexe

A l'échelle du chromosome entier, la taille des cartes génétiques CsCtG et CsCtA et ainsi le nombre total de CO est très proche entre la méiose mâle (334 CO pour 182 individus) et la méiose femelle (337 CO pour 185 individus). Aucune différence significative du taux de CO n'est observée entre les deux sexes à cette échelle.

A l'échelle des 31 intervalles communs entre les deux populations, trois intervalles présentent des différences significatives du taux de CO entre le mâle et la femelle (Figure 21, Tableau 18 en Annexes p 349). Ces trois intervalles sont localisés dans la partie distale du bras long. Le premier (*Xgpw4034-Xnw2706, p*-value = 0.04) montre un taux de CO plus important chez le mâle alors qu'au contraire, les deux autres intervalles (*Xwmm1133-Xwmm702 : p*-value=10^{-4} et *Xwmm702-Xwmm280 : p*-value=0.05) ont un taux de CO plus important chez la femelle, l'intervalle *Xwmm1133-Xwmm702* ayant un taux de CO 17 fois supérieur chez la femelle (Figure 21). Après correction par un test multiple, seul ce dernier intervalle reste significativement différent entre le mâle et la femelle.

En conclusion, nous pouvons donc dire que si le sexe n'a aucune influence sur le nombre de CO à l'échelle du chromosome 3B entier lorsque que l'on étudie le croisement CsCt, les différences de taux de CO locaux entre mâle et femelle peuvent parfois être très importantes.

1.5 Comparaison des cartes génétiques CsCt et CsRe : Variations du taux de CO sur le chromosome 3B en fonction du croisement étudié

Malgré les difficultés pour avoir des marqueurs en commun à cause des contraintes de polymorphisme entre populations, nous avons pu comparer les cartes génétiques CsRe et CsCt pour étudier la variation du taux de CO entre ces populations. Les fréquences locales entre mâles et femelles étant différentes sur la population CsCt, nous avons regroupé l'ensemble des individus issus des deux populations HDCsCt pour réaliser une carte génétique CsCt incluant les CO issus des deux sexes pour ne pas biaiser la comparaison avec la carte F2CsRe qui a été produite à partir de données provenant des deux sexes. De plus, nous avons utilisé seulement 184 F2CsRe pour conserver le même nombre de gamètes analysés entre les deux populations.

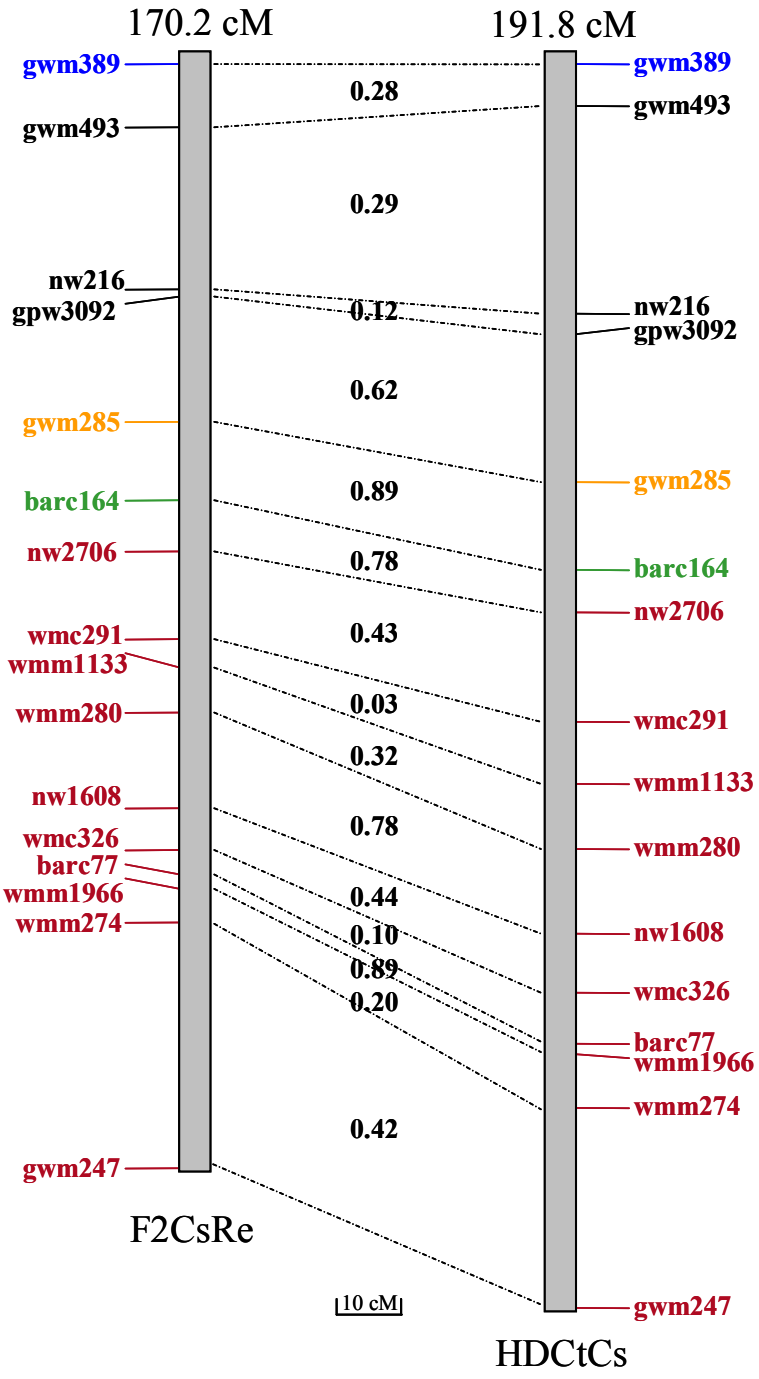


Figure 22 Comparaison des cartes génétiques F2CsRe et HDCsCt.

Ces cartes sont obtenues à partir de 184 F2 Chinese spring x Renan (F2CsRe) et 367 haploïdes doublés Chinese spring x Courtot (HDCsCt).
La population HDCsCt résulte du regroupement des données issues de la population obtenue par androgenèse et celle obtenue par gynogenèse pour avoir des distances génétiques représentant une moyenne des deux sexes. La couleur de chaque marqueur indique son assignation à un bin de délétion du chromosome 3B (code couleur Figure 20). Les valeurs indiquées entre les deux cartes représentent les *p*-values testant la significativité des différences de taux de CO entre les deux populations pour chaque intervalle.

La carte issue de la population F2CsRe et comportant 16 marqueurs (170.2 cM) présente un déficit de 21 cM par rapport à celle obtenue à partir de la population HDCsCt et des 16 mêmes marqueurs (191.8 cM) (Figure 22). Sur les quinze intervalles communs entre les deux populations, dix présentent un taux de CO plus important chez la population HDCsCt. Cependant, seul l'intervalle compris entre les marqueurs wmc291 et wmm1133 montre une *p*-value significative (0.03, Figure 22).

Ainsi, malgré la proximité génétique des deux populations, due au fait qu'elles possèdent un parent commun (Chinese Spring) et que Renan est un descendant de Courtot, la taille génétique du chromosome 3B varie de manière importante entre les populations CsRe et CsCt, ce qui suggère soit un contrôle génétique de la recombinaison différent entre Renan et Courtot, soit une structure du chromosome 3B différente entre ces deux variétés.

1.6 L'interférence entre CO

Pour étudier l'importance de l'interférence sur le chromosome 3B, nous avons utilisé les populations de lignées haploïdes doublés CsCt car le génotype de chaque individu représente celui d'un seul gamète, ce qui facilite la détection de deux CO sur la même chromatide.

Dans un premier temps, nous avons testé l'hypothèse H_0 : la distribution des CO observée sur ces populations est aléatoire et ne dépend pas de l'interférence. Nous avons comparé les distances entre deux CO notées expérimentalement et celles théoriques (Matériels et Méthodes paragraphe 7.1 p 147). Les *p*-values associées à cette hypothèse sont respectivement de 0.084 pour le mâle et 3.6×10^{-4} pour la femelle. Ces résultats suggèrent donc la présence d'interférence à la fois chez le mâle et chez la femelle.

Nous avons ensuite quantifié l'intensité de l'interférence sur le chromosome 3B chez le mâle et chez la femelle en utilisant le modèle gamma. Nous avons utilisé comme base expérimentale les valeurs des fréquences de distance entre deux CO afin de déterminer les valeurs du paramètre *m* associé au modèle gamma reflétant l'intensité de l'interférence. Chez le mâle, la valeur du paramètre *m* est de 1.2 avec une fluctuation de 0 à 6.1 pour un intervalle de confiance à 95 %. Chez la femelle, cette valeur de *m* est de 3.5 et varie de 0.95 à 12.75 pour un intervalle de confiance de 95 %. Nous observons ainsi une interférence faible à modérée sur le chromosome 3B, qui apparaît plus importante chez la femelle même si cette différence n'est pas statistiquement significative (*p*-value= 0.25).

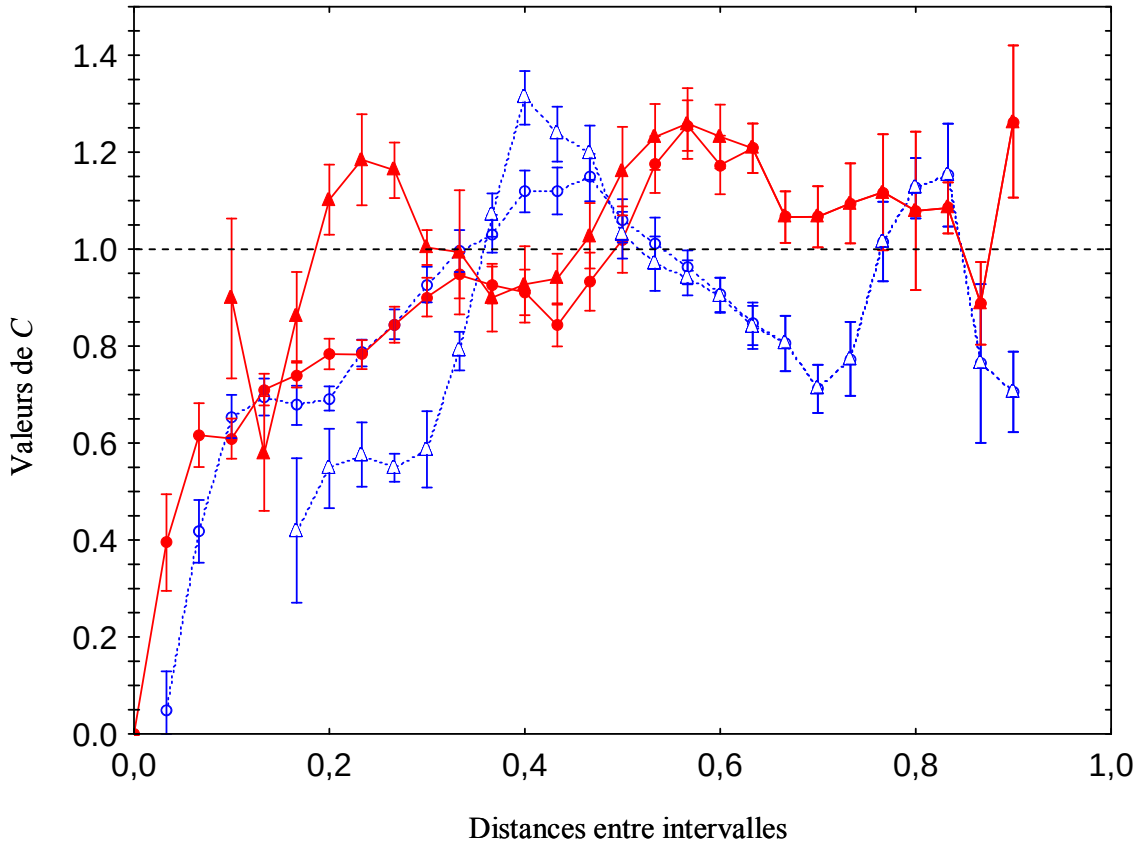


Figure 23 Valeurs du coefficient de coïncidence C en fonction des distances entre les deux intervalles étudiés.

La distance entre intervalles indiquée en abscisse est une distance relative à la distance en cM des cartes génétiques obtenues à partir des populations haploïdes doublés Chinese spring x Courtot issues de la voie femelle (194 cM) et de la voie mâle (190 cM). Les données indiquées en bleu avec des motifs vides et en rouge avec des motifs pleins représentent les données issues de la voie mâle et la voie femelle respectivement. Les données indiquées par des cercles représentent l'ensemble des paires d'intervalles alors que les données indiquées par des triangles représentent les paires d'intervalles qui enjambent le centromère.

Finalement, nous avons évalué l'étendue de l'interférence, c'est-à-dire la distance à laquelle l'interférence agit suite à la formation d'un premier CO. Pour cela, nous avons utilisé le coefficient de coïncidence *C* en fixant la position d'un des deux intervalles et en faisant varier la distance du second par rapport au premier. Le coefficient de coïncidence est égal à 0 lorsque l'interférence est positive et totale, à 1 lorsque l'interférence est nulle et *C* est supérieur à 1 lorsque l'interférence est négative. Sur la Figure 23, nous observons une augmentation des valeurs du coefficient de coïncidence lorsque les distances génétiques entre les intervalles considérés s'accroissent. Pour des distances inférieures à environ 10 cM, les valeurs de *C* sont très faibles (inférieures à 0.5) indiquant une forte interférence. Pour des distances supérieures à environ 45 cM, les valeurs fluctuent autour de 0.8 et l'interférence devient ainsi très faible. Nous pouvons également noter que dans notre cas, l'interférence agit à travers le centromère chez le mâle puisque des valeurs de *C* de 0.4 sont observées pour des intervalles situés de part et d'autre du centromère à des distances de 30 cM. Cependant, une telle action de l'interférence ne se manifeste pas chez la femelle où les valeurs de *C* sont de 0.6 pour des intervalles de part et d'autre du centromère.
Ces résultats indiquent donc la présence d'interférence positive modérée sur l'ensemble du chromosome 3B.

L'ensemble de ces résultats sur la distribution des CO à l'échelle du chromosome 3B a fait l'objet d'une publication dans la revue Genetics (partie IX p 351).

2 Distribution des CO sur un chromosome 3B délété

L'une des hypothèses émises pour expliquer l'importance de la fréquence des CO au niveau des régions distales serait leur position distale sur les chromosomes. Pour tester cette hypothèse sur le chromosome 3B, nous avons étudié la distribution des CO sur un chromosome 3B délété pour déterminer si le passage d'une région située en position intermédiaire sur le chromosome entier à une position distale sur le chromosome délété provoquait une augmentation de la fréquence de CO au niveau de cette région. Pour cette étude, nous avons utilisé la méthode proposée par Jones et al. (2002) et Qi et al. (2002) utilisant des croisements avec des lignées de délétion. Nous avons ainsi croisé la lignée de délétion 3BS4-0.55 issue de la variété Chinese Spring et chez laquelle les 45 % distaux du bras court du chromosome 3B sont absents avec la variété française Renan (Matériels et Méthodes paragraphe 1.5 p 133, Figure 24).

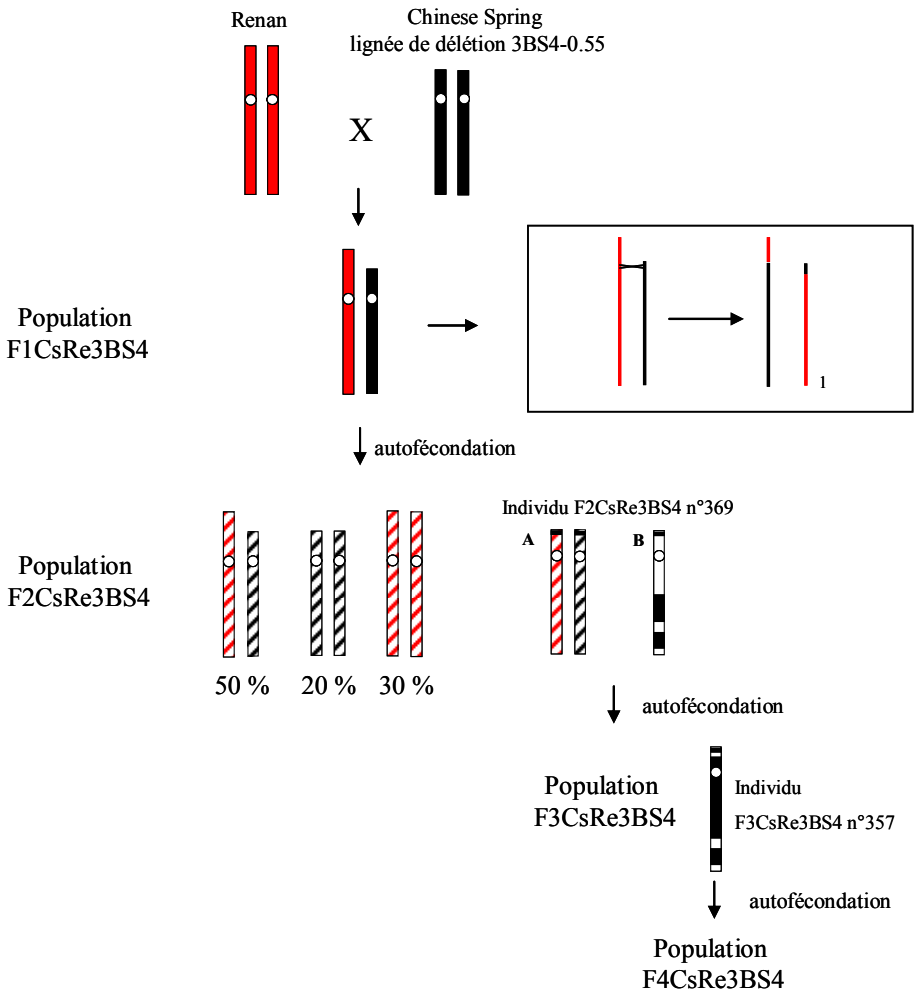


Figure 24 Schéma de croisement utilisé pour générer un chromosome 3B délété.

La lignée de délétion Chinese spring 3BS4-0.55 où les 45 % distaux du bras court du chromosome 3B sont absents a été croisée avec la variété Renan. Les individus F1 portent ainsi un chromosome 3B entier provenant de Renan (rouge) et un chromosome délété provenant de Chinese spring (noir). A la méiose de ces individus (encadré), la formation de CO au niveau de la partie distale du bras court délété entraîne la formation d'une chromatide délétée et de génotype Renan en-dessous de la région de formation du CO et Chinese spring en dessus. L'association de cette chromatide avec une chromatide délétée de type Cs entraîne la formation d'un individu F2 présentant un bivalent 3B hétérozygote CsRe en-dessous de la région de formation du CO d'origine et délété sur les deux chromosomes 3B (**A**, individu F2CsRe3BS4 n°369). Les principaux types de bivalents retrouvés chez les individus F2CsRe3BS4 sont représentés, les hachures indiquant la présence de changement potentiel de génotypes dû aux CO. Une analyse plus fine du bivalent de l'individu n°369 (**B**) montre la présence d'une mosaïque de régions homozygotes Cs (bloc noirs) et hétérozygotes CsRe (bloc blancs) due à la formation de CO sur les chromatides à l'origine de ces chromosomes. La population F3CsRe3BS4 issue de l'autofécondation de l'individu F2CSRe3BS4 n°369 a permis l'étude de la distribution des CO sur un bivalent 3B délété grâce au polymorphisme CsRe. Les régions proximales du bivalent 3B de l'individu n°357 étant à l'état homozygote Cs (blocs noirs), son autofécondation a conduit à la formation d'une population F4CsRe3BS4 utilisée pour déterminer l'impact de l'homozygotie Cs comparé à l'hétérozygotie CsRe sur la fréquence de CO dans les régions proximales.

Lors de la méiose des individus F1 issus de ce croisement, la formation d'un CO entre le chromosome entier provenant de Renan et celui délété provenant de Chinese Spring va entraîner, dans certains cas, la formation d'un chromosome délété provenant de Renan. Ainsi, lors de l'autofécondation des individus F1, l'association de ce nouveau chromosome 3B délété provenant de Renan et du chromosome délété initial provenant de la lignée de délétion Chinese Spring va générer un individu F2 hétérozygote Cs/Re double délété. L'autofécondation de cet individu F2 va ainsi nous permettre de développer une population F3 permettant d'étudier la distribution des CO au sein d'un bivalent délété pour les deux homologues par l'utilisation de marqueurs moléculaires.

2.1 Caractérisation fine du bivalent 3B de l'individu F2CsRe3BS4 n°369

Dans un premier temps, nous avons criblé les individus F2CsRe3BS4 issus du croisement entre la lignée de délétion 3BS4-0.55 et la variété Renan avec deux marqueurs moléculaires situés de part et d'autre de la zone de coupure 3BS4-0.55. Sur 356 individus F2 testés, seulement un individu (n°369) présentait une absence d'amplification pour le premier marqueur et de l'hétérozygotie Chinese Spring/Renan pour le second marqueur suggérant la présence de deux chromosomes homologues délétés au niveau de la jonction 3BS4-0.55 à l'état hétérozygote (présence d'un chromosome provenant de la variété Renan et l'autre de la variété Chinese Spring). De plus, les deux marqueurs utilisés pour réaliser le criblage des individus F2CsRe3BS4 nous ont permis d'observer un biais de ségrégation en faveur des bivalents 3B entiers provenant de Renan (30 %) par rapport aux bivalents délétés provenant de Chinese Spring (20 %).

Pour valider l'hypothèse d'une paire de chromosomes délétés à l'état hétérozygote CsRe, nous avons caractérisé finement le bivalent 3B de l'individu 369 par C-banding car les deux variétés possèdent des patterns différents mais également par le génotypage de plusieurs marqueurs distribués le long du chromosome. L'analyse par C-banding indique clairement la présence de la délétion sur chacun des deux homologues ainsi que l'origine biparentale des deux chromosomes délétés (l'un issu de Renan, l'autre de Chinese Spring, Figure 25). L'analyse de l'individu 369 à l'aide de 51 marqueurs moléculaires répartis sur le chromosome 3B montre une alternance de régions hétérozygotes et de régions homozygotes. Nous détectons trois régions homozygotes, une sur le bras court et deux sur le bras long, toutes de type Chinese Spring (Figure 24).

Chinese spring Renan Individu F2CsRe3BS4 n°369

Figure 25 Profil des bandes C (C-banding) sur le chromosome 3B.

Les patterns de C-banding des chromosomes 3B issus de Chinese spring, de Renan et de l'individu F2CsRe3BS4 n°369 sont représentés. Les observations ont été réalisées sous microscope optique à une magnitude de 1200 (photo G. Gay).

L'assignation des marqueurs aux bins de délétion nous a permis de localiser grossièrement ces zones d'homozygotie. La première localisée dans la partie la plus distale du bras court résulte de l'événement de recombinaison qui a permis la formation du chromosome 3B délété issu de la variété Renan (Qi et al. 2002). La deuxième région homozygote recouvre l'ensemble du bin 3BL10-0.50-0.63 mais également une partie des bins 3BL9-0.38-0.50 et 3BL7-0.63-1.00. La troisième région est localisée dans la partie distale du bin 3BL7-0.63-1.00.

2.2 Distribution des CO sur un bivalent 3B composé d'un chromosome délété et d'un chromosome entier

Nous avons ensuite utilisé la population F2CsRe3BS4 pour étudier la distribution des CO sur le bivalent 3B lorsqu'un des chromosomes est délété et le deuxième est entier (individu F1; bivalent dimorphique) et nous l'avons comparée avec celle observée sur le bivalent 3B où les deux homologues sont entiers. Sur le bras court, le taux de CO entre les loci *Xwmc612* et *Xnw2715* diminue légèrement d'un facteur 3.2 passant de 0.019 à 0.006 mais de manière non significative (*p*-value = 0.08) lorsque l'un des homologues est délété sur la partie distale du bras court. A l'opposé, aucune différence du taux de CO n'est observée entre les deux types de bivalents sur les différents intervalles étudiés sur le bras long (Figure 26).

A l'exception de deux loci situés dans le bin 3BL7-0.63-1.00, l'ensemble des marqueurs testés présente le même biais de ségrégation que celui observé pour les marqueurs ayant servi à faire le criblage, c'est-à-dire un excès des homologues 3B entiers provenant de Renan par rapport aux homologues délétés provenant de Chinese Spring.

Ainsi, au moment de la méiose, la présence d'un bivalent 3B dimorphique composé d'un chromosome avec un bras court délété et un chromosome avec un bras court entier entraîne une réduction du taux de CO sur le bras court dans les régions à l'état diploïde. Au niveau du bras long, le taux de CO reste identique à celui observé sur un croisement impliquant un bivalent 3B entier.

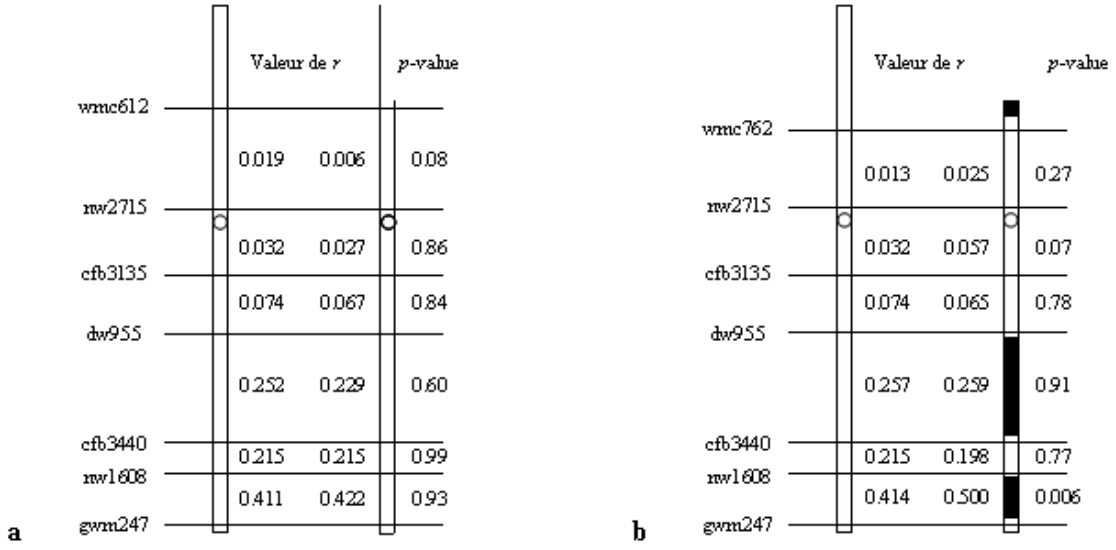


Figure 26 Comparaison de la distribution des CO observés sur un bivalent 3B entier, un bivalent 3B comportant un chromosome entier et un chromosome délété et un bivalent délété.

Les différents intervalles étudiés sont délimités par les marqueurs situés à gauche des chromosomes. Les valeurs du milieu représentent les valeurs du taux de recombinaison *r* obtenues grâce au logiciel Mapmaker pour le bivalent entier (à gauche) et le bivalent comportant un chromosome entier et un chromosome délété (**a**) ou le bivalent délété (**b**) (à droite). La significativité des différences de taux de recombinaison observées entre les deux types de bivalent pour chaque intervalle est indiquée à droite des chromosomes par des *p*-value. **a/** Comparaison de la distribution des CO entre le bivalent entier (Population F2CsRe) et un bivalent 3B comportant un chromosome entier et un chromosome délété pour 45 % de l'extrémité de son bras court (Population F2CsRe3BS4). **b/** Comparaison de la distribution des CO entre le bivalent entier (Population F2CsRe) et le bivalent délété (Population F3CsRe3BS4). Les zones en noires sur le bivalent délété représentent les régions homozygotes Cs. Les *p*-values ont été calculées à partir des individus ne possédant qu'un seul CO au sein de l'intervalle étudié (Matériels et Méthodes paragraphe 6.2 p 145).

2.3 Distribution des CO sur le bivalent 3B composé des deux chromosomes délétés

Nous avons génotypé 379 individus F3CsRe3BS4 issus de l'autofécondation de l'individu F2CsRe3BS4 n°369 pour d'étudier la distribution des CO sur un bivalent dont les deux chromosomes sont délétés sur le bras court et déterminer l'effet de la position distale sur le taux de CO. De plus, au niveau du bras long, nous avons voulu déterminer si la présence de régions homozygotes Chinese Spring avait un impact sur le taux de CO au sein de ces régions comparativement aux états hétérozygotes (Chinese Spring/Renan) et ainsi déterminer si l'état hétérozygote pouvait être responsable des faibles fréquences de CO observées au niveau des régions proximales (Figure 20 p 166).

2.3.1 Influence de la délétion sur la distribution des CO

Nous avons comparé les taux de CO *r* obtenus sur les régions hétérozygotes du bivalent 3B délété à partir de la population F3CsRe3BS4 avec les taux de CO issus des mêmes régions du bivalent 3B entier et obtenus à partir de la population F2CsRe.

Au niveau de la partie hétérozygote du bras court située entre les loci *Xwmc762* et *Xnw2715*, le taux de CO augmente d'un facteur 1.9 chez le bivalent délété par rapport au bivalent entier mais de manière non significative (*p*-value = 0.27). L'étude plus fine de six intervalles compris entre les loci *Xwmc762* et *Xnw2715* montre que l'augmentation du nombre de CO se produit sur l'ensemble des intervalles et qu'elle ne se localise pas uniquement au niveau de l'intervalle le plus distal. Le taux de CO au sein de la partie hétérozygote du bras court du bivalent délété est de 0.05 CO par bivalent à chaque méiose et reste donc très faible.

Au niveau de la région hétérozygote la plus proximale sur le bras long (entre les loci *Xnw2715* et *Xcfb3135*, Figure 26), le taux de CO augmente légèrement mais de manière non significative (*p*-value = 0.07) passant de 0.032 sur le bivalent entier à 0.057 sur le bivalent délété alors que sur les deux autres régions hétérozygotes (entre les loci *Xcfb3135* et *Xdw955* et entre les loci *Xcfb3440* et *Xnw1608,* Figure 26), aucune augmentation du taux de CO sur le chromosome délété n'est observée (*p*-value respectivement de 0.78 et 0.77).

De façon inattendue, les marqueurs présents à l'intérieur de la région hétérozygote comprise entre les loci *Xcfb3440* et *Xnw1608* présentent un fort biais de ségrégation en faveur des allèles Renan au détriment des allèles Chinese Spring sur la population F3CsRe3BS4.

En conclusion, nous pouvons dire que la délétion d'une partie distale du bras court du chromosome 3B n'entraîne qu'une très faible augmentation du taux de CO d'un facteur de 1.85 sur la partie hétérozygote du bras court restant ainsi que sur la partie du bras long la plus proximale couvrant approximativement 30 % de ce bras.

2.3.2 Influence de l'hétérozygotie CsRe sur la distribution des CO

2.3.2.1 Etude des taux de CO sur la descendance de l'individu F2CsRe3BS4 n°369

Le bras long du bivalent 3B de l'individu F2CsRe3BS4 n° 369 étant composé d'une mosaïque de régions homozygotes Cs et hétérozygotes CsRe (Figure 24), nous avons pu comparer les taux de CO au niveau de ces différentes régions avec ceux obtenus sur un bivalent entièrement hétérozygote CsRe de façon à évaluer l'impact des états hétérozygote et homozygote sur les taux de CO. Les taux de CO des deux régions homozygotes du bras long du chromosome 3B ont été estimés à l'aide de marqueurs situés dans les régions hétérozygotes flanquant les régions homozygotes. Cependant, les problèmes de phase rencontrés pour cette étude et notamment pour le calcul des valeurs du taux de CO *r* sur les régions homozygotes Cs nous ont contraints à calculer les valeurs de *r* à partir des individus avec un seul CO dans l'intervalle, individus dont le génotype ne varie pas que les marqueurs soient en couple ou en répulsion (Matériel et Méthodes paragraphe 6.2 p 145).

Au niveau de la région comprise entre les loci *Xdw955* et *Xcfb3440* (Figure 26) sur le chromosome délété, nous n'observons aucune différence du taux de CO par rapport à la même région à l'état hétérozygote sur le chromosome entier (*p*-value = 0.91, Figure 26). Ainsi, l'hétérozygotie CsRe au sein de cette région n'a aucun impact sur le taux de CO comparé à l'état homozygote Cs.

En revanche, au niveau de la région homozygote Cs comprise entre les loci *Xnw1608* et *Xgwm247*, nous observons une augmentation significative du taux de CO (*p*-value = 0.006) par rapport à la même région à l'état hétérozygote sur le chromosome entier (Figure 26). Cependant, nous ne pouvons pas conclure sur la cause de cette augmentation et en particulier si elle est due à la délétion sur le bras court ou à l'homozygotie de la région.

En conclusion, nous pouvons dire que l'état hétérozygote Cs/Re n'influence pas la fréquence de CO au sein de la région proximale étudiée comparé à l'état homozygote Cs de cette même région.

2.3.2.2 Etude sur la descendance de l'individu F3CsRe3BS4 n°357

Pour confirmer que l'état hétérozygote vs homozygote n'entraîne pas de différences de taux de CO dans les régions proximales et ainsi qu'il n'est pas responsable des très faibles fréquences de CO observées au niveau de ces régions à l'échelle du chromosome entier (Figure 20 p 166), nous avons étudié la distribution des CO sur des individus issus de l'autofécondation de l'individu F3CsRe3BS4 n°357. Cet individu porte un bivalent 3B homozygote sur la région entre les loci *Xcfb3266* (bin 3BS4-0.33-0.55) et *Xcfb3440* (bin 3BL7-0.63-1.00), ces deux loci étant à l'état hétérozygote. Cette région homozygote couvre ainsi toute la partie proximale (Figure 24) qui présente une faible fréquence de CO à l'échelle du chromosome entier. Nous avons ainsi comparé le taux de CO au sein de cette région sur le bivalent entier à l'état hétérozygote (population F2CsRe), sur le bivalent délété une petite partie étant homozygote et la majorité étant hétérozygote (descendance de l'individu F2CsRe3BS4 n°369) et sur le bivalent délété à l'état homozygote (descendance de l'individu F3CsRe3BS4 n°357). Aucune différence significative du taux de CO n'est observée au sein de cette région entre l'état homozygote sur le bivalent délété et l'état hétérozygote sur le chromosome entier (p-value = 0.87) mais également entre l'état homozygote sur le bivalent délété et l'état hétérozygote/homozygote sur le bivalent délété (p-value = 0.99). Ainsi, à des échelles importantes de l'ordre de plusieurs Mb, l'état hétérozygote Chinese Spring/Renan comparé à l'état homozygote Chinese Spring n'a aucune influence sur la fréquence des CO au niveau des 600 Mb entourant le centromère.

3 Distribution des CO à l'échelle de régions séquencées

L'impact de la séquence sur la fréquence des CO n'a jamais été évalué en détail chez le blé tendre. Ainsi, nous avons comparé la distribution des CO sur les premières régions séquencées de quelques Mb avec les caractéristiques intrinsèques de la séquence telles que la distribution et la densité des gènes, la présence d'éléments répétés.

Nom du contig	Bin de délétion	Nombre de BAC	Taille en pb	Nombre de gènes	Nombre de pseudogènes	Densité en gènes (nombre pour 150 kb)
Rph7	3BS8-0.78-0.87	26	3109948	41	11	2.0
Sr2	3BS8-0.78-0.87	ND	1300000	ND	ND	ND
L	3BS1-0.33-0.55	6	619476	0	0	0
Nom du contig	**% ET**	**% GC**	**Position sur la carte CsRe[a]**	**SSR développés**	**SSR assignables**	**Autres marqueurs utilisés [b]**
Rph7	63	45.6	7.4	122	89	12
Sr2	ND	ND	5.6	56	42	0
L	98	48.8	58.4	0	0	2

Tableau 7 Caractéristiques principales des trois contig étudiés.

ND = non déterminé. % ET = pourcentage d'éléments transposables. [a] : position du marqueur le plus distal sur le contig par rapport au marqueur le plus distal du bras court sur la carte CsRe en cM. [b] : marqueurs ISBP et SNP.

Nous avons focalisé notre étude sur trois régions du bras court du chromosome 3B dont deux (le contig *Sr2* et le contig *Rph7*) localisées dans le bin 3BS8-0.78-0.87 qui montre une fréquence de CO élevée et l'autre (le contig *L*) dans le bin 3BS1-0.33-0.55 situé en région intermédiaire du bras court avec une faible fréquence de CO. Le contig *Sr2* est situé en position légèrement plus distale que le contig *Rph7* sur la carte génétique CsRe ; le contig *L* est localisé très proche du centromère génétiquement. Le contig *Sr2* (1.3 Mb) est une région qui possède un gène de résistance à la rouille brune chez le blé alors que le contig *Rph7* (3.1 Mb) est une région qui possède des résistances à la rouille chez l'orge et à la fusariose chez le blé. De plus, le contig *Rph7* est une région particulièrement riche en gènes avec une densité de 2 gènes pour 150 kb alors qu'aucun gène n'est présent sur le contig *L* (620 kb) (Tableau 7). La séquence du contig *Sr2* n'étant pas encore optimum, les caractéristiques de ce contig ne sont pas encore connues.

3.1 Marqueurs développés sur les différents contigs étudiés

3.1.1 Microsatellites et assignation aux bins de délétion

A partir des données de séquence, nous avons développé un total de 56 microsatellites sur le contig *Sr2* et 122 sur le contig *Rph7* (Tableau 7). Le contig *L* étant constitué de 98 % de séquences répétées, nous avons préféré utiliser les marqueurs ISBP développés sur ce contig au cours d'un autre programme.

Au niveau du contig *Sr2*, sur 56 marqueurs SSR testés, six n'ont donné aucun produit d'amplification chez l'ensemble des individus testés. Parmi les 50 autres, 84 % ont pu être assignés au bin 3BS8-0.78-0.87 confirmant ainsi la localisation de ce contig au sein de ce bin et la présence de ces marqueurs uniquement sur le chromosome 3B. Les 16 % restants n'ont pas pu être assignés car ils donnent un produit d'amplification sur l'ensemble des lignées de délétion. De plus, la mise en relation de la position physique des marqueurs sur le contig avec leur assignation au bin de délétion a mis en évidence la présence d'un bloc important d'environ 400 kb sur le contig *Sr2* au sein duquel l'ensemble des marqueurs (15) est assignable au bin 3BS8-0.78-0.87 suggérant l'absence ou une évolution importante de ce bloc par rapport aux chromosomes homéologues.

Populations	CsRe				CsCt			
Contigs	***Sr2***		***Rph7***		***Sr2***		***Rph7***	
SSR polymorphes	29		37		9		37	
	% (ass)	**% (tot)**	**% (ass)**	**% (tot)**	**% (ass)**	**% (tot)**	**% (ass)**	**% (tot)**
SSR polymorphes dominants	40	30	19	14	12	9	19	14
SSR polymorphes co-dominants	29	21	22	16	10	7	22	16

Tableau 8 Polymorphisme des SSR développés sur les contigs Sr2 et Rph7.

CsRe : Chinese spring x Renan ; CsCt : Chinese spring x Courtot. % (ass) représente le pourcentage de marqueurs polymorphes par rapport à l'ensemble des marqueurs assignés au bin de délétion 3BS8-0.78-0.87. % (tot) représente le pourcentage de marqueurs polymorphes par rapport à l'ensemble des marqueurs utilisés.

Nombre de nucléotides du motif SSR	**0**	**1**	**2**	**3**	**4**	**Tous**
SSR total	9	7	33	48	25	122
SSR assignables	5	6	24	35	19	89
SSR polymorphes	4	4	13	11	5	37
% polymorphe (ass)	80	67	54	31	26	41
% polymorphe (tot)	44	57	39	23	20	30

Tableau 9 Taux de polymorphisme des SSR entre Chinese spring et Renan sur le contig Rph7 en fonction du type de motif des SSR.

0 représente un marqueur contenant plusieurs motifs SSR. % polymorphe (ass) représente le pourcentage de marqueurs polymorphes CsRe par rapport à l'ensemble des marqueurs ayant le même type de motif et assignés au bin de délétion. % polymorphe (tot) représente le pourcentage de marqueurs polymorphes CsRe par rapport à l'ensemble des marqueurs utilisés ayant le même type de motif.

Au sein du locus *Rph7*, sur les 122 marqueurs testés, 11 ne présentent pas d'amplification chez les individus testés. Parmi les 111 autres, 80 % sont assignés au bin 3BS8-0.78-0.87. Comme pour le contig précédent, les 20 % de marqueurs non assignables sont dus à la présence d'un produit d'amplification sur l'ensemble des lignées de délétion. Sur le contig *Rph7*, de manière identique au contig *Sr2*, sur 51 marqueurs issus des 800 kb les plus proximaux sur le chromosome 3B, 48 sont assignables au bin de délétion.
En conclusion, l'assignation des marqueurs SSR développés sur ces deux contigs au bin de délétion est importante de l'ordre de 80 % et dépend de leur localisation sur ceux-ci.

3.1.2 Microsatellites et polymorphisme au niveau des régions séquencées

Le polymorphisme des microsatellites développés au sein des deux contigs *Sr2* et *Rph7* a été évalué par comparaison des profils de trois lignées utilisées comme parents de populations de cartographie : Chinese Spring (Cs), Renan (Re), Courtot (Ct).

Sur les 42 marqueurs du contig *Sr2* assignables au bin de délétion 3BS8-0.78-0.87, respectivement 69 % et 22 % sont polymorphes entre Cs et Re, Cs et Ct (Tableau 8). De plus, concernant les données de polymorphisme entre les variétés Cs et Re, nous avons observé que 59 % des marqueurs assignables et polymorphes sont dominants à cause d'une absence d'amplification chez la variété Renan.
Parmi les 89 marqueurs du contig *Rph7* assignables au bin de délétion 3BS8-0.78-0.87, 41 % des marqueurs sont polymorphes entre les parents Cs et Re et Cs et Ct, les marqueurs polymorphes étant les mêmes entre les deux couples (Tableau 8). Contrairement aux données du contig *Sr2*, le nombre de marqueurs polymorphes dominants est égal au nombre de marqueurs polymorphes co-dominants. Parmi les microsatellites mono-nucléotides et di-nucléotides assignables, 67 et 54 % sont polymorphes respectivement comparés aux 31 et 26 % de polymorphes pour les tris et tétra-nucléotides (Tableau 9), cette analyse n'étant pas possible sur le contig *Sr2* car la majorité des marqueurs SSR utilisée était de type di-nucléotides. De manière intéressante, la répartition des marqueurs polymorphes indique que la région terminale distale représentant environ 315 kb montre une accumulation importante de marqueurs polymorphes (11) entre Cs et Re et Cs et Ct concernant les marqueurs assignables comparée au reste du locus et présente ainsi un taux de polymorphisme de 78 %.

Pour conclure, sur l'ensemble des 178 marqueurs SSR analysés sur les contigs *Sr2* et *Rph7*, 74 % d'entre eux sont assignables au bin 3BS8-0.78-0.87 et parmi ceux-ci 50 % présentent du polymorphisme entre les variétés Cs et Re soit 37 % de marqueurs assignables et polymorphes entre Cs et Re sur l'ensemble des marqueurs testés. Le taux observé entre les variétés Cs et Ct est plus faible avec 26 % de marqueurs assignables et polymorphes. Cependant, le taux de polymorphisme varie en fonction de la position sur les contigs mais également en fonction du nombre de nucléotides composant le motif SSR.

Afin d'augmenter le nombre de marqueurs polymorphes à notre disposition sur ces deux régions, nous avons séquencé 37 marqueurs SSR monomorphes issus des contigs *Sr2* (5) ou *Rph7* (32) sur les variétés Cs, Re et Ct pour rechercher d'éventuelles différences ponctuelles de nucléotides entre les séquences ou SNP (Single Nucleotide Polymorphism). Des données de séquences inexploitables ont été obtenues pour sept marqueurs SSR et seulement 19 % des 30 autres séquencés soit 7 montrent la présence d'au moins un SNP entre Cs et Re et Cs et Ct. Le séquençage a ainsi permis de trouver du polymorphisme sur les deux populations pour environ un cinquième des marqueurs monomorphes et ainsi d'enrichir des zones non couvertes jusqu'à présent.

3.2 Distribution des CO sur les contigs séquencés

3.2.1 Fréquence de CO au sein du contig *L*

Le contig *L* se situe en position intermédiaire sur le bras court du chromosome 3B, dans une région où la fréquence de CO est très faible (bin 3BS1-0.33-0.55 : 0.02 cM/Mb). Il est constitué de 98 % de séquences répétées insérées les unes dans des autres.

Nous avons utilisé deux marqueurs ISBP (cfp5109 et cfp5129) développés au cours d'un autre programme et situés aux deux extrémités du contig pour identifier les individus portant une ou plusieurs chromatides où un CO a eu lieu dans la génération précédente au niveau de cette région, ils seront appelés individus recombinants ou recombinants par la suite. Sur les 1872 individus F2 de la population CsRe criblés, nous n'avons observé aucun individu recombinant. De plus, le criblage des 379 individus F3CsRe3BS4 où le taux de CO est légèrement augmenté sur le bras court montre également une absence totale d'individu recombinant au sein de ce contig. En conclusion, nous pouvons dire que la formation de CO au sein de ce contig, contenant uniquement des séquences répétées, semble être fortement inhibée sans pouvoir déterminer si cela est dû à son fort taux de séquences répétées ou à sa position dans une région du chromosome faiblement recombinogène.

3.2.2 Distribution des CO sur les contigs situés dans le bin 3BS8-0.78-0.87

3.2.2.1 Criblage d'individus recombinants au sein du locus R

En parallèle au développement de marqueurs sur les contigs *Sr2* et *Rph7*, nous avons criblé 1872 individus F2CsRe avec les marqueurs microsatellites gwm389 et gwm493 qui définissent la région nommée locus *R* (car elle possède un nombre important de gènes de résistance aux maladies) pour isoler les individus possédant un CO au sein de cette région. Ces deux marqueurs sont distants d'environ 12 cM et ils encadrent les marqueurs présents sur les contigs *Sr2* et *Rph7* sur la carte génétique CsRe (Figure 20 p 166, *Sr2* : marqueurs cfb5010 et swm13 ; *Rph7* : marqueurs cfb6012 et cfb6016). De plus, nous avons utilisé un marqueur supplémentaire (gpw7757) situé à égale distance génétique de gwm389 et de gwm493 afin de prendre en compte les individus F2 qui pouvaient posséder une chromatide avec un double CO mais également ceux qui pouvaient avoir deux chromatides recombinées de manière réciproques entre les deux marqueurs. Dans un premier temps, nous avons isolé 390 individus F2 recombinants portant un ou deux CO au sein de cette région. Un re-génotypage de ces individus pour vérification et validation a confirmé la validité de 380 individus recombinants, les 10 autres étant des erreurs de génotypage.

Pour les 1734 individus F2 pour lesquels le génotypage des trois marqueurs testés (gwm389, gpw7757 et gwm493) est disponible, nous avons recherché les individus montrant la présence d'un double CO sur la même chromatide en observant les différentes combinaisons génotypiques de ces individus. Si A représente le génotype Renan, B le génotype Chinese spring et H le génotype hétérozygote, seul les individus dont la combinaison génotypique pour les trois marqueurs testés est AHA, BHB, ABH, HBA, HAB, BAH, ABA et BAB montrent un double CO sur la même chromatide. Aucun individu de l'une ou l'autre de ces combinaisons n'est présent parmi les individus testés indiquant qu'aucun double CO n'est observé au sein de l'intervalle compris entre les marqueurs gwm389 et gwm493 sur cette population et pour cet effectif. Ce résultat confirme ceux observés sur les deux populations HDCsCt où nous avions montré précédemment que l'interférence était très importante pour des distances égales ou inférieures à 10 cM.

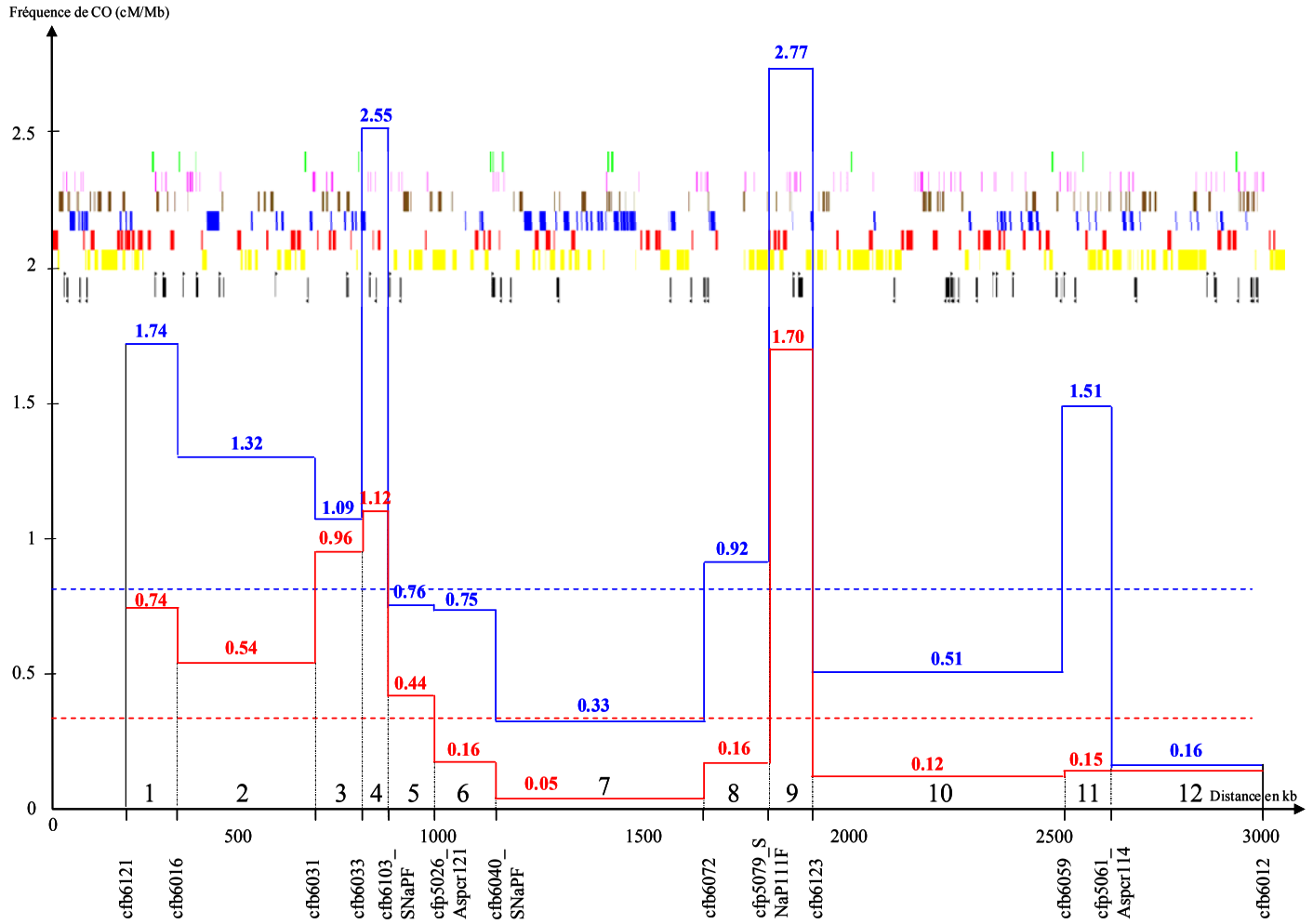


Figure 27 Variation de la fréquence de CO sur le contig Rph7.

Les fréquences de CO ont été obtenues à partir des données issues de la population F2CsRe (bleu) et de la population F2CsCt (rouge) et sont indiqués sur chacun des plateaux. Le contig est divisé en 12 régions numérotées de 1 à 12 et définies par les marqueurs situés en bas de la figure. Les barres en couleur représentent les éléments répétés (vert : LINE, rose : MITE; marron : non classé; bleu : CACTA; rouge : copia; jaune : gypsy) et les barres en noir en-dessous des éléments répétés représentent les gènes et les pseudogènes. Les lignes en pointillés représentent la fréquence moyenne de CO sur le contig *Rph7* pour la population F2CsRe (bleu) et la population F2CsCt (rouge).

3.2.2.2 Fréquence de CO au sein du contig *Sr2*

Nous avons criblé les 380 individus recombinants du locus *R* avec les deux microsatellites cfb5010 et swm13 couvrant un maximum de longueur du contig *Sr2* et à la fois polymorphes et co-dominants entre Cs et Re. Nous avons isolé 16 individus recombinants entre ces deux marqueurs. Ceux-ci montrent un biais de ségrégation en faveur de l'allèle Renan puisque 12 sont de type BH alors que seulement 4 sont de type AH. Cependant, nous n'avons pas pu étudier les relations entre fréquences de CO et caractéristiques de la séquence, le séquençage de cette région n'étant pas optimum. Ce type d'analyse sera donc fait uniquement sur le locus *Rph7*.

3.2.2.3 Localisation fine des CO sur le contig *Rph7*

a Etude à partir de la population CsRe

Nous avons criblé les 380 individus recombinants du locus *R* avec les marqueurs microsatellites cfb6121 et cfb6012, distants d'environ 2.8 Mb chez la variété Chinese Spring, ce qui nous a permis d'isoler 77 individus recombinants dont un présentait deux CO, un sur chaque homologue. La fréquence moyenne de CO du contig est ainsi de 0.82 cM/Mb ce qui est très proche de celle estimée pour le bin dans lequel se trouve ce contig (3BS8-0,78-0,87 : 0.85 cM/Mb).

i Etude à l'échelle du contig entier

Nous avons ensuite utilisé l'ensemble des marqueurs SSR, ISBP et SNP présent sur ce contig et montrant un polymorphisme co-dominant entre Cs et Re pour sélectionner les marqueurs définissant des régions ayant une taille physique la plus homogène possible pour étudier la distribution des CO. Ainsi, douze sous-régions ayant une taille moyenne de 230 kb et variant de 68 kb à 625 kb ont été définies. La fréquence de CO varie de 0.16 cM/Mb au sein de la région 12 à 2.77 cM/Mb au sein de la région 9 (Figure 27) soulignant une distribution non homogène à l'intérieur du contig sur des régions de quelques centaines de kb. Deux régions (Figure 27, régions 7 et 12) présentent une fréquence de CO faible avoisinant la fréquence moyenne du chromosome pour la région 12 alors que deux autres (Figure 27, régions 4 et 9) présentent une fréquence environ 14 fois plus forte que la moyenne du chromosome.

	Densité en gènes	% ex/in g/psd	% ex g/psd	% ex/in g	% ex g	% GC	Ratio CpG	% ET	% SSR
12 intervalles *Rph7* (CsRe)	0.18	0.05	**0.36**	0.32	**0.39**	0.04	0.2	0.3	0.02
11 intervalles *Rph7* (CsCt)	0.17	0.32	**0.82**	**0.68**	**0.6**	0.03	0.03	0.31	0.08
7 intervalles *HGA3* (CsRe)	ND	ND	ND	ND	ND	0.01	0.27	0.01	0.28
7 intervalles *HGA3* (CsCt)	ND	ND	ND	ND	ND	**0.66**	0.04	0.24	0.24

Tableau 10 Valeurs des coefficients de détermination R^2 testant la relation entre la distribution des CO et les caractéristiques de la séquence sur les différents intervalles du contig Rph7.

Les différences de fréquence de CO observées au niveau des différents intervalles définis sur le contig *Rph7* mais aussi les différences de fréquence au niveau des différents intervalles étudiés sur la région encadrant le gène *HGA3* de 23 kb ont été mis en relation avec la distribution des différentes caractéristiques des intervalles. Le calcul des valeurs de chaque caractéristique de la séquence est présenté en détail dans la partie Matériels et Méthodes (paragraphe 9.2 p 153). % ex/in représente le pourcentage de la taille physique cumulée des exons et des introns par rapport à la taille physique de l'intervalle. % ex représente seulement le pourcentage de la taille physique cumulée des exons. g/psd = pourcentage pour l'ensemble des gènes et pseudogènes. g = pourcentage pour l'ensemble des gènes. Les valeurs en gras indiquent les valeurs significatives au risque $\alpha = 0.05$ (Table de Bravais-Pearson, en Annexes p 339). ND = non déterminé.

Nous avons mis en relation la distribution des CO avec différents paramètres de la séquence. Les valeurs des coefficients de détermination montrent que la distribution des fréquences de CO n'est pas corrélée avec la densité en gènes, le pourcentage en GC, le ratio CpG, le pourcentage d'éléments transposables et le pourcentage de séquences répétées en tandem (Tableau 10). Cependant, le pourcentage de la taille physique cumulée des exons de l'ensemble des gènes présents dans chaque intervalle explique de manière significative 39 % de la variation de fréquence des CO au sein des intervalles. De plus, la fréquence de CO est également corrélée de façon significative avec le pourcentage de la taille physique cumulée des exons de l'ensemble des gènes et des pseudogènes ($R^2 = 0.36$).

Ainsi, à cette échelle de quelques centaines de kb, la taille physique cumulée de l'ensemble des gènes sur un intervalle donné joue un rôle important dans la fréquence de CO observée au sein de cette région.

ii Distribution des CO à l'échelle de dizaines de kilobases

Nous nous sommes ensuite focalisés sur deux régions précises (régions 8 et 9 : Figure 27) pour regarder à une échelle plus fine comment se répartissaient les CO. La région 8 (entre cfb6072 et cfp5079_SNaP111F, 157 kb) comporte un seul gène (*HGA1*) et présente une fréquence de CO voisine de celle du contig entier alors que la région 9 (entre cfp5079_SNaP111F et cfb6123, 105 kb) comporte deux gènes (*HGA3* et *SOS1*) et présente la plus forte fréquence de CO du contig. Les marqueurs présents au sein de ces deux régions nous ont permis de les découper en six intervalles et d'encadrer chacun des trois gènes présents dans ces deux régions à des échelles de quelques kb (6.8 kb pour *HGA1*, 23 kb pour *HGA3* et 38 kb pour *SOS1*). Sur les quinze CO présents dans les deux intervalles de départ (région 8 et 9), douze sont localisés dans les trois régions encadrant les gènes. De plus, l'intensité de la fréquence de CO au sein des intervalles contenant un gène varie; la région encadrant le gène *HGA3* montrant la plus forte fréquence avec 8.82 cM/Mb (Figure 28).

Ainsi, au sein de la région située entre les marqueurs cfb6072 et cfb6123, la majorité des CO se localise dans des régions de quelques kb entourant les gènes.

iii Distribution des CO autour du gène *HGA3*

Nous nous sommes finalement focalisés sur la région de 23 kb encadrant le gène *HGA3* où sept CO ont été détectés. Nous avons découpé cet intervalle de 23 kb en sept intervalles et nous avons observé que les CO sont présents dans la région en amont du gène (défini comme la partie transcrite du gène), au niveau de la partie codante du gène mais également au niveau de la région en aval du gène (Figure 29).

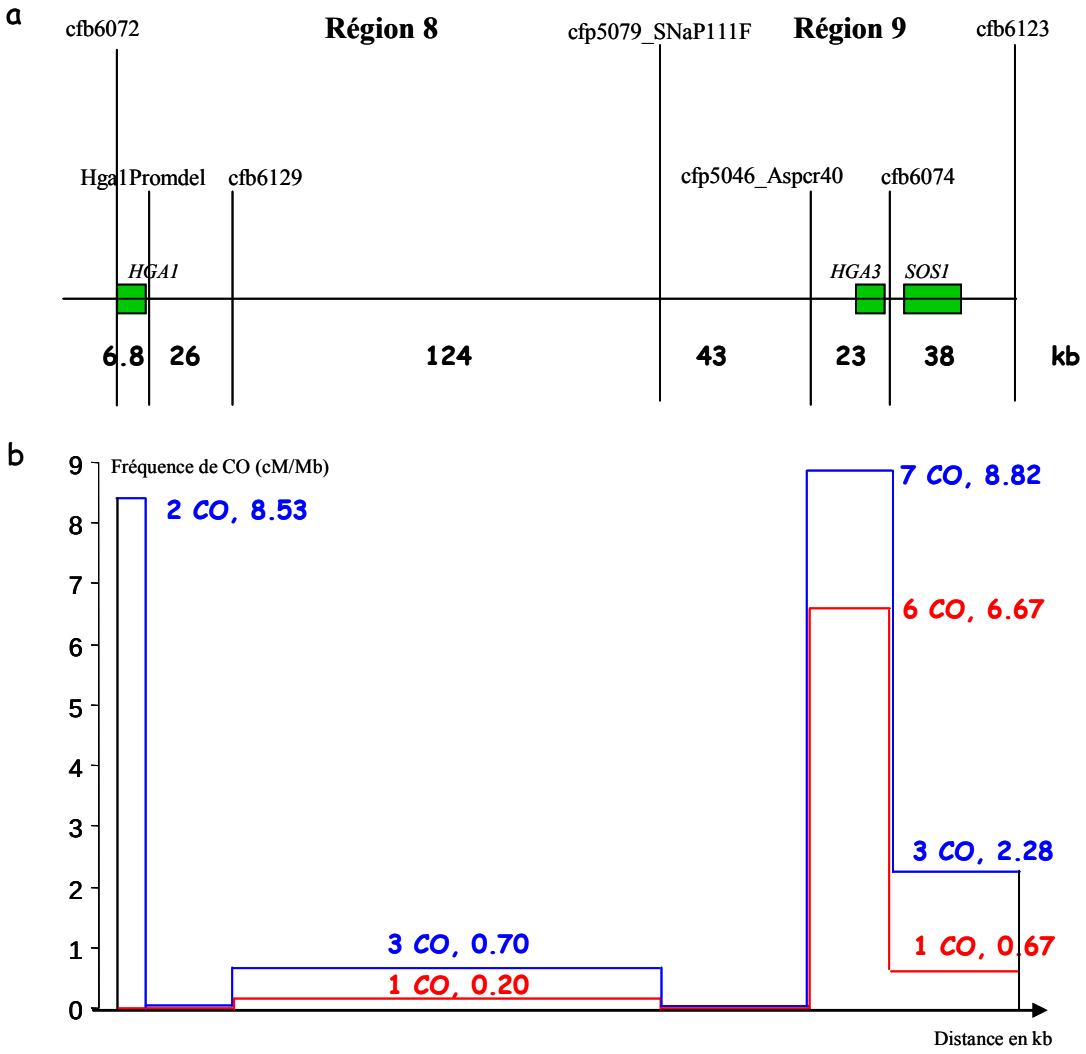


Figure 28 Distribution des CO sur les régions 8 et 9 entre les marqueurs cfb6072 et cfb6123.

a/ Détail des six intervalles étudiés au sein des régions 8 et 9 (Figure 27) du contig *Rph7*. Les trois gènes présents sont indiqués par les boîtes vertes. La taille des six intervalles étudiés est indiquée en kb sous chacun d'entre eux. **b/** Variation de la fréquence de CO sur les régions 8 et 9 à partir des données de la population F2CsRe (bleu, 1719 individus) et de la population F2CsCt (rouge, 1957 individus). Sur chaque plateau sont indiqués le nombre de CO observés et la fréquence de CO en cM/Mb pour chaque intervalle et chaque population.

Ainsi les CO présents dans la région entourant le gène *HGA3* s'étendent au-delà de la partie codante et ne se localisent pas uniquement sur une région particulière du gène. De plus, aucune des caractéristiques de la séquence comme le pourcentage en GC, le ratio CpG et le pourcentage de séquences répétées n'est corrélée avec la distribution des CO à cette échelle (Tableau 10).

b Etude à partir de la population CsCt

Afin de déterminer si la distribution des CO obtenue sur le contig *Rph7* à partir de la population F2CsRe était conservée sur un autre croisement, nous avons criblé 2230 individus F2CsCt avec les mêmes marqueurs cfb6121 et cfb6012 bordant le contig. Parmi les 1957 individus présentant des données de génotypage exploitables, nous avons identifié 35 individus recombinants dans cet intervalle, ce qui correspond à une fréquence de CO sur cette population de 0.32 cM/Mb soit 2.6 fois moins que celle observée sur la population F2CsRe.

i Distribution des CO à l'échelle du contig entier

Nous avons ensuite utilisé les mêmes marqueurs que ceux utilisés pour la population CsRe pour délimiter des intervalles communs entre les deux populations sur ce contig. Le marqueur cfp5061-Aspcr114 délimitant la séparation entre les régions 11 et 12 sur la population CsRe n'étant pas polymorphe entre Cs et Ct, nous n'avons pu définir que onze intervalles. La fréquence de CO varie de 0.05 cM/Mb pour la région 7 à 1.7 cM/Mb pour la région 9 (Figure 27). Malgré une fréquence moyenne plus faible sur la population CsCt, le pattern de distribution des CO sur le contig est globalement conservé avec celui de la population CsRe (R^2 = 0.84). En effet, les deux intervalles ayant la fréquence de CO la plus importante (cfb6033-cfb6103_SNaPF; cfp5079_SNaP111F-cfb6123) et les deux intervalles ayant la fréquence la plus faible (cfb6040_SNaPF-cfb6072 ; cfb6059-cfb6012) sont communs aux deux populations. Cependant, le rapport fréquence de CO obtenue à partir de la population CsRe et fréquence de CO obtenue à partir de la population CsCt montre la présence de différences ponctuelles entre les deux populations. Par exemple, au sein de la région 3, la fréquence de CO est grossièrement la même pour les deux populations alors qu'au niveau de la région 7 la fréquence de CO obtenue à partir de la population CsCt est 6.8 fois plus faible (Figure 27).

La mise en relation des caractéristiques de la séquence et de la distribution des CO obtenue à partir de la population CsCt montre que celle–ci n'est pas corrélée avec la densité en gènes, le pourcentage en GC, le ratio CpG, le pourcentage d'éléments transposables et le pourcentage de séquences répétées en tandem (Tableau 10).

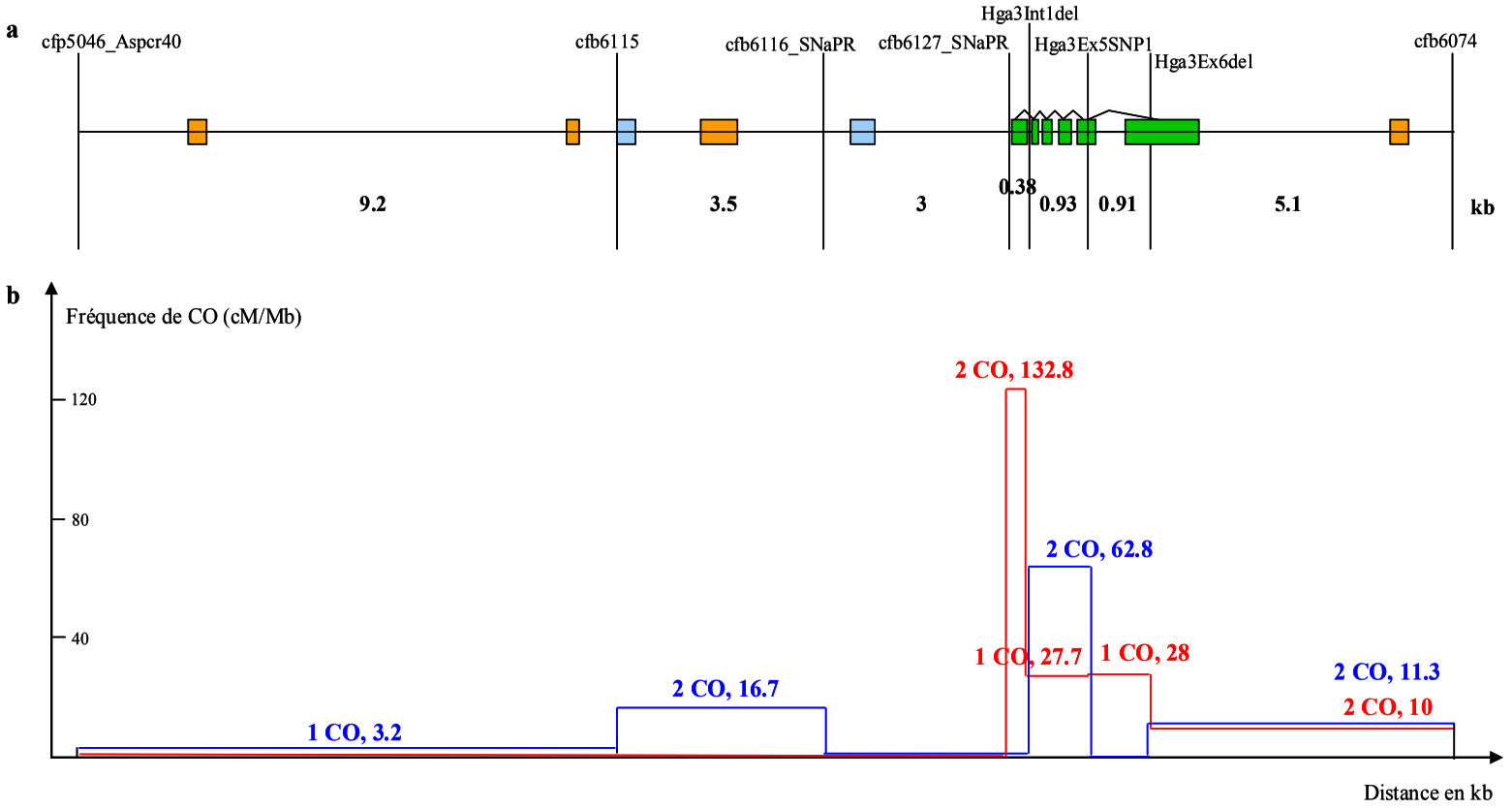


Figure 29 Distribution des CO sur la région de 23 kb entourant le gène HGA3 entre les marqueurs cfp5046_Aspcr40 et cfb6074.

a/ Détail des intervalles étudiés au niveau de la région encadrant le gène *HGA3*. Les exons du gène *HGA3* sont indiqués par les boîtes vertes. Les boîtes orange et bleues représentent respectivement des éléments transposables de type MITE et des éléments répétés non connus. La taille de chaque intervalle est indiquée en kb sous chacun d'entre eux. **b/** Variation de la fréquence de CO sur la région entourant le gène *HGA3* à partir des données de la population F2CsRe (bleu, 1719 individus) et de la population F2CsCt (rouge, 1957 individus). Sur chaque plateau sont indiqués le nombre de CO observés et la fréquence de CO pour chaque intervalle et chaque population.

Cependant, le pourcentage de la taille physique cumulée des exons de l'ensemble des gènes, le pourcentage de la taille physique cumulée des exons/introns de l'ensemble des gènes ainsi que le pourcentage de la taille physique cumulée des exons de l'ensemble des gènes et pseudogènes montrent une corrélation importante avec la fréquence de CO (Tableau 10). En effet, ce dernier facteur explique 82 % de la variation de la fréquence de CO observée entre les onze intervalles.

Ainsi, la taille des parties codantes des gènes et pseudogènes est le facteur majeur expliquant la distribution des CO sur le contig *Rph7*.

ii Distribution des CO à l'échelle de dizaines de kilobases et autour du gène *HGA3*

L'étude détaillée des régions 8 et 9 du contig *Rph7* (Figure 27 p 194) montre que sur les huit CO présents dans ces deux régions, sept sont localisés dans l'un des six intervalles définis qui encadrent un gène (Figure 28). En effet, six sont situés dans l'intervalle contenant le gène *HGA3* et un dans l'intervalle contenant le gène *SOS1*. Le dernier CO est placé dans l'intervalle entre les gènes *HGA1* et *HGA3* sans pouvoir le rattacher à une séquence codante quelconque.

Sur l'intervalle de 23 kb entourant le gène *HGA3* (Figure 29), la distribution des CO est différente de celle observée chez la population CsRe. Aucun CO n'est situé dans la partie en amont du gène, la majorité se localisant sur le gène lui-même. Deux CO sont également présents dans la partie en aval du gène sans que l'on puisse déterminer s'ils sont localisés dans le dernier exon ou dans la partie extérieure au gène. De plus, la distribution des CO est corrélée de manière significative avec le taux de GC ($R^2 = 0.66$, Tableau 10).

Ainsi, au sein de la population CsCt, l'analyse fine montre que les CO se localisent sur des régions encadrant les gènes et notamment au niveau du gène *HGA3* où les CO se localisent sur le gène.

c Conclusion

En conclusion, nous pouvons dire que l'analyse détaillée d'une séquence de plusieurs Mb chez le blé tendre tend à montrer que les CO sont plus fréquemment localisés à proximité ou dans les gènes, suggérant que la distribution des gènes pourrait expliquer la distribution des CO.

V Discussion et perspectives

1 Distribution des CO sur le chromosome 3B

La comparaison de la carte génétique avec la carte physique de délétion a montré que la fréquence de CO augmente des centromères jusqu'aux télomères. Ainsi, 77 % des CO sont présents dans les régions distales couvrant seulement 25 % du chromosome et 27 % des régions entourant le centromère sont quasiment dépourvues de CO. Ce même pattern de distribution des CO a été observé pour la première fois sur les chromosomes du groupe 7 par Werner et al. (1992) puis sur l'ensemble des chromosomes du blé tendre par Erayman et al. (2004). Ainsi, bien que ce chromosome soit le plus grand (995 Mb), son pattern de distribution des CO est identique aux autres. Il peut donc servir de modèle pour comprendre la distribution des CO sur les chromosomes du blé tendre.

Ce type de distribution des CO est aussi observé chez les espèces à grand génome comme le maïs (Anderson et al. 2003), l'orge (Kunzel et al. 2000), le sorgho (Kim et al. 2005) alors que chez les espèces possédant un génome plus petit comme le riz ou la tomate, ce gradient du centromère aux télomères est beaucoup moins important voire inexistant (Chen et al. 2002; Sherman et Stack 1995). La raison de la formation de ce gradient reste pour l'heure inexpliquée mais plusieurs hypothèses ont été émises telles que l'influence de la position d'une région sur l'axe centromère-télomère signifiant que la position distale expliquerait les fortes fréquences de CO au niveau des régions distales ou la présence de différents types de séquences. Tout récemment, Lukaszewski (2008) a montré que sur un chromosome de seigle portant un bras long inversé, les chiasmas se forment au niveau des régions proximales et non plus dans les régions distales sur le bras inversé montrant que seul le type de séquences est responsable de la formation des CO. Ainsi, certaines régions comme celles présentes dans les régions distales des chromosomes de blé et présentant un type particulier de séquence seraient « aptes » à la formation de CO et d'autres non comme celles présentes dans les régions proximales et présentant un autre type particulier de séquence. La différence entre ces deux types de séquences reste encore inconnue ainsi que leur relation précise avec le mécanisme de recombinaison.

Chez les espèces possédant un grand génome, Dawe et al. (1998) et Prieto et al. (2004) ont observé que la chromatine se décondense et s'étire dans les régions terminales et sub-terminales des chromosomes en début de méiose. De plus, Colas et al. (2008) ont récemment observé que si la décondensation des chromosomes est inhibée sur un chromosome de blé, la formation des CO est inhibée simultanément suggérant que cette décondensation est importante pour la formation des CO. Elle pourrait favoriser l'accès à la machinerie de recombinaison, permettre ainsi la formation des CO et expliquer les fortes fréquences de CO dans ces régions.

Cette décondensation, sous le contrôle du locus *Ph1*, serait nécessaire pour permettre la reconnaissance entre chromosomes homologues en entraînant leur décondensation de façon synchronisée et ainsi éviter l'appariement entre homéologues (Martinez-Perez et Moore 2008). Retrouvée uniquement au niveau des régions distales, la restriction à ces régions pourrait être liée avec le type de séquence particulier présent au sein de celles-ci. Ainsi, le type de séquence présent au sein des régions distales pourrait subir la décondensation induite par le mécanisme de reconnaissance des homologues alors que le type de séquence présent dans les régions proximales ne permettrait pas la décondensation des régions proximales. Ce mécanisme de reconnaissance des chromosomes homologues présent au sein du blé hexaploïde combiné aux différents types de séquence entre régions distales et proximales pourraient être la cause des fortes fréquences de CO observées dans les parties distales.

En absence du locus *Ph1*, la décondensation se produit même si le deuxième homologue n'est pas présent suggérant que la décondensation dans ces régions serait donc réalisée par un mécanisme encore inconnu et contrôlé par la présence du locus *Ph1*. Ainsi, la plupart des espèces étant des polyploïdes ou des anciens polyploïdes, la présence de ce gradient chez d'autres espèces comme le maïs pourrait indiquer la présence d'un mécanisme identique de reconnaissance des homologues.

Pour valider ces hypothèses, il serait donc intéressant de regarder sur le chromosome 3B si les régions qui subissent une décondensation de la chromatine dans les régions distales au cours des premiers stades de méiose sont les mêmes que celles qui présentent des fréquences de CO importantes. L'analyse menée par Prieto et al. (2004) et Colas et al. (2008) a été effectuée sur le chromosome 1D comportant un fragment de seigle couvrant les 15 % distaux, ce qui a limité l'analyse de la décondensation de la chromatine à ce fragment de seigle sans pouvoir conclure si l'élongation observée était la même sur les différentes régions du fragment.

La carte physique du chromosome 3B récemment développée pourrait nous permettre de mieux appréhender la décondensation de la chromatine en début de méiose et sa relation avec la formation des CO.

Nous pourrions utiliser les contigs de la carte physique du 3B dont une petite vingtaine est maintenant séquencée pour développer des sondes utilisables en FISH et réparties tout le long des bras du chromosome 3B. Ces sondes permettraient dans un premier temps de déterminer l'étendue de la décondensation de la chromatine en comparant les distances observées entre chacune pendant les premiers stades de la méiose avec celles mesurées pendant des stades pré-méiotiques servant ainsi de témoin. Ces données seraient comparées avec la position génétique des sondes déterminée par les marqueurs moléculaires présents sur les mêmes contigs que celles-ci. Ces analyses permettraient ainsi dans un deuxième temps de déterminer si la décondensation est réellement inhibée au niveau des régions proximales faiblement « recombinogènes » et si les régions distales présentant de fortes fréquence de CO sont les mêmes que celles qui subissent la décondensation. Cette étude serait réalisée sur des individus F1CsRe qui sont les individus à partir desquels nous avons effectué l'étude de la distribution des fréquences de CO mais aussi sur les lignées Cs et Re servant de témoin.

La structure de la chromatine a un impact considérable sur la distribution des CO chez des espèces autres que le blé (Kirik et al. 2006; Lopez et al. 2008; Tsai et al. 2008). Ainsi, nous pourrions également regarder s'il existe, en plus de la décompaction des régions distales qui s'opère sous le contrôle du locus *Ph1*, d'autres modifications particulières de la structure de la chromatine. Cela pourrait être réalisé en regardant la structure en 3D du chromosome 3B au cours des premiers stades de méiose. Tout récemment, une nouvelle technique de microscopie nommée 3dSIM (three-dimensional structured illumination microscopy) a été mise au point (Carlton 2008). Elle permet une amélioration de la visualisation des chromosomes d'un facteur 2 par rapport aux anciennes techniques. Cet auteur montre que celle-ci pourrait être utilisée pour observer l'organisation structurelle des chromosomes à la méiose. Une visualisation du chromosome 3B nous permettrait donc de déterminer s'il existe des différences de structure entre régions proximales et régions distales.

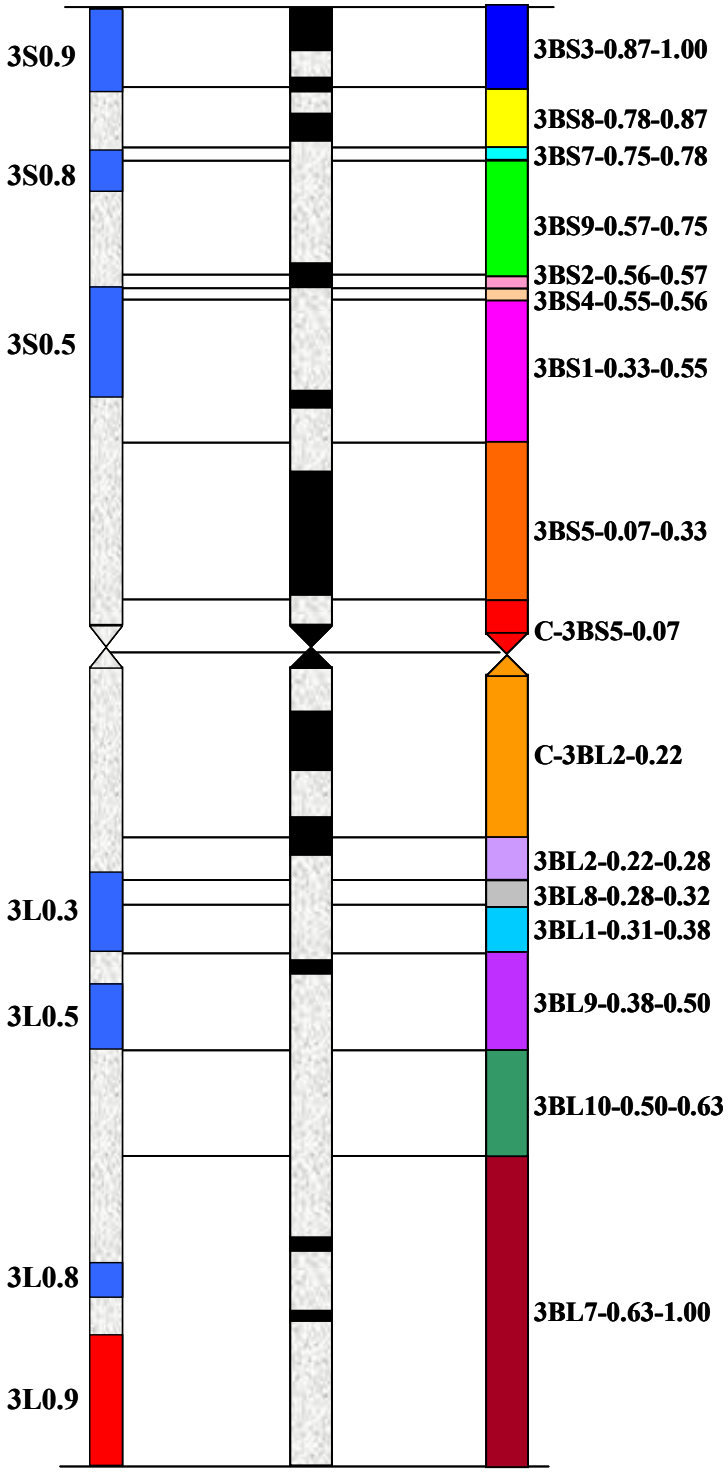


Figure 30 Différentes caractéristiques du chromosome 3B.

Les sept régions riches en gènes décrites par Erayman et al. (2004) sont représentées sur le chromosome de gauche par des segments bleus pour des régions moyennement riche en gènes (20 à 40 %) et par des segments rouges pour des régions très riches en gènes (50 à 60 %). Le nom de chacune est indiqué sur la gauche. Le chromosome du milieu représente les différentes bandes C du chromosome 3B (Gill et al. 1991). Ces deux caractéristiques sont mises en relation avec les 16 bins de délétion du chromosome 3B (représentation de droite).

L'étude de la distribution des CO sur le chromosome 3B réalisée à partir de la carte physique de délétion reste grossière car la taille physique des bins de délétion a été estimée cytologiquement. Le manque de distance physique précise est un frein considérable pour l'étude de la distribution des CO sur le chromosome 3B. La mise en place de la séquence dans les années à venir sera ainsi une révolution pour ce type d'analyse comme nous l'avons constaté à travers l'étude sur la région *Rph7* totalement séquencée. Dans cette perspective lointaine, il serait intéressant de densifier la carte CsRe notamment dans les régions distales où il reste encore des régions pauvres en marqueurs pour pouvoir la comparer avec la séquence du chromosome 3B dès que celle-ci sera disponible.

2 Chute de la fréquence de CO à l'extrémité du bras court

La résolution de notre étude a permis d'observer une chute de la fréquence de CO à l'extrémité du bras court du chromosome 3B plus particulièrement dans le bin de délétion le plus distal (3BS3-0.87-1.00).

Cette observation pourrait s'expliquer par le fait que nous n'avons pas pris en compte l'ensemble des CO qui se produit sur l'extrémité du bras court en raison d'un manque de marqueurs télomériques ou proches du télomère. Cependant, nous avons plusieurs arguments qui laissent penser que la totalité des CO présents dans ce bin a été pris en compte dans notre étude. En particulier, l'ancrage des marqueurs de la carte génétique F2CsRe sur les contigs de la carte physique auxquels ont également été assignés des EST de blé fourni des relations de synténie précieuse. Ainsi, les marqueurs gwm1034 (absent sur les cartes du manuscrit mais co-ségrége avec wmm756) et cfb6012 sont ancrés sur des contigs portant respectivement les EST *BE404656* et *BJ302829*. Ces deux EST correspondent aux gènes de riz Os01g01960 et Os01g03100 localisés en position 500 kb et 1 Mb sur le chromosome orthologue du riz (chromosome 1). Les marqueurs gwm1034 et cfb6012 sont respectivement cartographiés à 5,5 et 7,4 cM du marqueur gwm389 le plus distal du bras court ce qui suggère que la région synténique comprenant ce marqueur chez le riz serait située à une position très inférieure à 500 kb et donc très proche du télomère sur le chromosome 1 de riz.

L'hétérochromatine étant associée à des faibles fréquences de CO chez la plupart des espèces étudiées comme chez la pomme de terre, la tomate, le sorgho, le maïs (Gaut et al. 2007; Iovene et al. 2008), la présence du bloc important d'hétérochromatine retrouvée au sein de ce bin de délétion pourrait expliquer la chute de la fréquence de CO dans cette partie distale (Figure 30).

Nos résultats sont cependant contradictoires avec ceux présentés par Erayman et al. (2004) qui montrent, par comparaison d'une carte génétique et d'une carte physique consensus des chromosomes du groupe 3, que la fréquence de CO est maximum dans le bin de délétion le plus distal (3BS3-0.87-1.00). Ces différences avec nos résultats s'expliquent vraisemblablement par le fait qu'ils ont utilisé des cartes consensus dont la précision est moindre mais également qu'un très faible nombre de leurs marqueurs a été assigné aux bins de délétion.

Cette chute de fréquence de CO observée sur le bras court n'est pas retrouvée au niveau de l'extrémité du bras long dans notre étude. Cependant, une chute similaire au niveau du bras long est mentionnée par Lukaszewski et Curtis (1993) lorsqu'ils utilisent le pattern de C-banding comme marqueur. Le fait que nous n'observions pas de décroissance de la fréquence des CO en position distale sur le bras long est vraisemblablement dû à la trop grande taille du bin distal 3BL7-0.63-1.00 (208 Mb). Cette hypothèse pourrait être vérifiée en utilisant d'autres lignées dont la délétion serait localisée en position plus télomérique. Une telle lignée existe (3BL11, Endo et Gill 1996) mais nous ne sommes pas parvenus à isoler cette délétion à l'état homozygote et nous n'avons donc pas pu exploiter ce matériel. Une façon de pallier ce manque serait de développer d'autres lignées de délétion ou d'autres types de matériel tels que des hybrides d'irradiation (Riera-Lizarazu et al. 2008). De telles lignées sont en cours de production pour faciliter l'ancrage des contigs physiques sur les cartes génétiques (S Kianian, communication personnelle).

Une chute de la fréquence de CO aux extrémités distales des chromosomes a déjà été observée au sein d'un croisement entre le seigle et le blé (Lukaszewski et al. 2004) mais également au sein d'un croisement entre le ray-grass et la fétuque dans les régions terminales du chromosome 3, l'orthologue du chromosome 3B (King et al. 2002; King et al. 2007). Le développement de cartes à haute résolution pour l'ensemble des chromosomes du blé tendre (www.wheatgenome.org) permettra de confirmer ou non si cette chute de la fréquence de CO est un phénomène général ou une spécificité du chromosome 3B. Le fait que See et al. (2006) observent que 70 % des CO se produisent dans les 23 Mb distaux du chromosome 4B suggère que le pattern de distribution des CO dans les régions les plus distales semble spécifique à chaque chromosome.

3 Différence du taux de CO entre la méiose mâle et la méiose femelle

L'utilisation de deux populations haploïdes doublés CsCt, l'une réalisée par gynogenèse et l'autre par androgenèse, a permis de comparer les taux de CO entre méiose mâle et méiose femelle sur le chromosome 3B.

Alors que nous n'avons observé aucune différence en termes de nombre de CO à l'échelle du chromosome entier, des différences apparaîssent à des échelles plus fines. Notamment, l'étude de 31 régions réparties sur le chromosome a montré la présence d'une région située dans le bin de délétion le plus distal du bras long (3BL7-0.63-1.00) où le taux de CO est dix sept fois plus important chez la femelle. Ces résultats diffèrent de ceux obtenus par Wang et al. (1995) et Guzy-Wrobelska et al. (2007). Wang et al. (1995) ont montré, en utilisant 16 marqueurs répartis sur les chromosomes des groupes 6 et 7, que le taux de CO à l'échelle du chromosome est plus important chez un des deux sexes et ceci de manière chromosome spécifique alors que nous n'avons observé aucune différence au niveau du chromosome 3B entier entre les deux sexes. Guzy-Wrobelska et al. (2007) ont montré sur le chromosome 3B, en utilisant cinq marqueurs et deux populations haploïdes doublés (76 lignées issues de gynogenèse et 82 issues d'androgenèse) issues d'un croisement entre les variétés Eta et Darkhan15, que le taux de CO est plus important chez le mâle que chez la femelle sur un intervalle du bras long. Cette différence se localise sur la partie du bras long où au contraire nous observons simultanément un intervalle pour lequel le taux est plus important chez le mâle (*Xgpw4034-Xnw2706*) et celui où le taux est plus important chez la femelle (*Xwmm1133-Xwmm702*). Ces différences de résultats sont vraisemblablement dues soit aux différentes variétés utilisées soit aux faibles effectifs tant en nombre de marqueurs que d'individus utilisés par ces deux équipes comparativement à nos données.

L'hétérochiasmie est un phénomène présent chez la plupart des organismes étudiés (Lenormand et Dutheil 2005b; Lorch 2005). Celle-ci peut se manifester par une différence à l'échelle du génome entier où la femelle présente un taux de CO plus important que le mâle comme chez l'humain, la souris, le cochon ou encore le chien ou l'inverse comme chez le pin ou *Arabidopsis thaliana* (Tableau 12 en Annexes p 325, (Lenormand et Dutheil 2005b; Lorch 2005).

L'hétérochiasmie que nous observons sur le chromosome 3B ne présente pas de différence à l'échelle du chromosome entier mais des différences plus locales comme cela a déjà été observé chez certaines espèces de céréales telles que l'orge ou le millet (Busso et al. 1995; Lenormand et Dutheil 2005b). Ces différences locales et notamment celle au niveau de la région où le taux de CO est dix sept fois plus important chez la femelle restent difficiles à expliquer. De tels phénomènes locaux ont également été mentionnés chez l'humain où une région de 8 Mb présente un taux de CO seize fois plus important chez la femelle par rapport au mâle (Matise et al. 2002).
Plusieurs facteurs sont évoqués pour expliquer ces différences locales tels que l'accessibilité de la chromatine ou l'expression sexe-spécifique de certains gènes (Matise et al. 2002). En effet, les régions montrant des phénomènes d'empreinte génétique, c'est-à-dire des régions où l'expression des gènes est différente entre les allèles provenant du mâle et ceux provenant de la femelle (Efstratiadis 1994), présentent des différences de taux de CO entre les deux sexes chez l'humain (Lercher et Hurst 2003). Ce facteur pourrait donc expliquer les différences observées sur le chromosome 3B.

La connaissance et la prise en compte de l'impact du sexe sur le taux de CO global ou local sont des phénomènes importants en particulier pour déterminer le type de population à développer pour un projet de clonage positionnel dans le but de minimiser le nombre d'individus à analyser et maximiser le nombre de recombinants dans la région d'intérêt.

L'étude fine de la région présentant une différence importante de taux de CO entre la méiose mâle et la méiose femelle pourrait être intéressante pour comprendre quels sont les facteurs autres que la séquence qui gouvernent la formation des CO. Pour cela, dans un premier temps, il serait nécessaire de confirmer cette différence. Deux populations de backcross (une mâle et une femelle) issues du croisement entre Cs et Re en cours de développement nous permettrons de de valider ce résultat sur ce nouveau croisement et de disposer d'effectifs de population plus importants pour l'étude fine de cette région. Si la différence entre le taux de CO mâle et le taux de CO femelle est confirmée dans cette région, il serait ensuite nécessaire d'affiner la position exacte de la zone montrant une différence de taux de CO entre les deux sexes et, une fois la zone identifiée avec plus de précision, d'en établir la carte physique. Celle-ci pourrait être séquencée afin de tenter d'identifier la (ou les) cause (s) de l'écart de recombinaison. En effet, cette séquence permettrait dans un premier temps de déterminer les gènes présents dans cette région et de tester leur expression spécifiquement chez l'un ou l'autre des deux sexes.

De plus, grâce à la séquence nous pourrions développer des sondes FISH et regarder si l'élongation de la chromatine est différente entre les deux sexes par le même type d'analyse que celle décrite auparavant (paragraphe 1 p 205).

4 Présence d'interférence positive sur le chromosome 3B

Nous avons étudié l'interférence entre CO sur le chromosome 3B en utilisant pour la première fois chez le blé le modèle gamma et son paramètre *m* associé ainsi que le calcul des coefficients de coïncidence. Les valeurs de *m* (1.2 chez le mâle et 3.5 chez la femelle) montrent la présence d'une interférence modérée sur le chromosome 3B. Ces valeurs sont proches des valeurs observées chez d'autres espèces comme chez la tomate (1.3 et 1.9 pour les chromosomes 1 et 2 respectivement, Lhuissier et al. 2007), chez l'humain (2.3, Broman et Weber 2000) ou chez *Arabidopsis thaliana* (de 3 à 9, Copenhaver et al. 2002).

L'étude du coefficient de coïncidence montre que pour des distances inférieures à environ 10 cM, l'interférence positive est très importante. Ce résultat a été confirmé sur l'ensemble des 1734 individus de la population F2CsRe lorsque nous avons criblé les individus recombinants entre les locus *Xgwm389* et *Xgwm493* distants d'environ 12 cM. En effet, aucun individu ne présente de double CO sur la même chromatide. Chez le blé et l'orge, Peng et al. (2000) et Esch et Weber (2002) ont montré, à l'inverse de nos résultats, la présence d'interférence négative agissant sur de courtes distances génétiques. Peng et al. (2000) ont utilisé une population F2 pour réaliser cette étude. L'utilisation d'une telle population doit se faire avec précaution car les valeurs du coefficient peuvent être biaisées du fait que deux gamètes indépendants sont présents au sein du même individu. Deux CO ayant eu lieu dans la même région mais de manière indépendante provenant ainsi de deux gamètes différents se retrouvent chez le même individu et risquent d'être comptabilisés comme un double CO sur la même chromatide. De plus, ces deux groupes ont utilisé dans leur étude des intervalles de très petite taille avec moins de 10 individus recombinants pour certains, ce qui introduit un autre biais dans la mesure du coefficient de coïncidence et notamment une surestimation de sa valeur comme observée lors de l'utilisation de valeurs basées sur des rapports (Martin et Hospital 2006).

L'interférence négative observée par ces auteurs pourrait donc être due à des biais dans les méthodes d'analyse utilisées. Ainsi, l'impact de l'interférence serait positif à des faibles distances chez le blé comme observé chez la plupart des organismes étudiés (Mezard et al. 2007).

Nous observons également que l'interférence agit sur des CO situés de part et d'autre du centromère chez le mâle. Ce phénomène est commun et est observé chez la plupart des organismes étudiés jusqu'à présent (Drouaud et al. 2007; Lian et al. 2008).

Contrairement aux autres régions que nous avons analysées, nos résultats suggèrent qu'au niveau des régions proximales, l'interférence agit, chez le mâle, sur des distances génétiques importantes de l'ordre de 30 cM alors qu'elle est restreinte à des distances d'environ 10 cM sur le reste du chromosome. Une hypothèse pour expliquer ce phénomène pourrait être la variation de compaction de l'ADN le long du chromosome. Chez l'humain et la souris, l'interférence semble corrélée avec la longueur du complexe synaptonémal ou encore la longueur du bivalent mesuré en µm (Codina-Pascual et al. 2006; Petkov et al. 2007). Ce résultat suggèrerait que l'intensité de l'interférence serait constante sur un chromosome par unité de taille physique des chromosomes (en µm). Cependant, les distances estimées en µm ne prennent pas en compte la variation de compaction de la chromatine et un µm correspond à une distance physique estimée en Mb plus importante dans les régions fortement compactées par rapport aux régions plus relâchées. Les régions entourant le centromère du chromosome 3B étant riches en hétérochromatine (Figure 30), celles-ci ont vraisemblablement un degré de compaction important. Un µm dans ces régions représente ainsi une distance physique en Mb beaucoup plus importante que dans les régions plus distales. Si l'interférence est la même sur un µm dans les régions proximales et dans les régions distales, elle va agir sur de plus grandes distances physiques en Mb au niveau des régions proximales. L'importante taille physique en Mb de ces régions au niveau du centromère permettrait ainsi la formation d'un nombre plus important de CO, ce qui entraînerait une distance génétique plus importante.

5 Effet de l'état hétérozygote Chinese Spring/Renan sur la distribution des CO à l'échelle du chromosome

La présence d'une mosaïque de régions hétérozygotes Cs/Re et homozygotes Cs sur le bras long du bivalent délété 3B de l'individu F2CsRe3BS4 n°369 nous a permis de déterminer l'impact de l'hétérozygotie CsRe sur la distribution des CO.

Nous avons observé une augmentation significative du taux de CO dans la région homozygote la plus distale du bras long sur le bivalent délété comparée à cette même région à l'état hétérozygote sur le bivalent entier.

Deux facteurs peuvent contribuer à cette augmentation : la délétion du bivalent sur le bras court et l'état homozygote Cs.

En effet, la délétion du bras court entraîne une réduction de la formation des CO sur ce bras qui pourrait être compensée par une augmentation du nombre de CO sur le bras long (Wijnker et de Jong 2008). Cependant, cette hypothèse semble peu probable car nous n'observons pas d'augmentation similaire au niveau de la partie distale du bras long à l'état hétérozygote sur le bivalent 3B composé d'un chromosome délété et d'un chromosome entier et où la fréquence de CO est également inhibée sur la partie diploïde du bras court.

Il semble donc que cette augmentation serait plutôt due à l'état homozygote Cs de cette région sur le bivalent délété. Cependant, le passage à l'état homozygote dans les régions intermédiaires et proximales (etude de la descendance de l'individu F3CsRe3BS4 n°357) ne provoque pas d'augmentation similaire du taux de CO. Cette différence entre le comportement des régions distales et des régions proximales pourrait s'expliquer par une divergence de séquence entre les variétés Cs et Re plus importante au niveau des régions distales qu'au niveau des régions proximales. Le passage à l'état homozygote provoquerait ainsi une augmentation plus importante du taux de CO dans les régions distales. Cette hypothèse est en accord avec un ensemble de résultats présents dans la littérature. En effet, des différences de séquence comme des insertions/délétions entre chromosomes homéologues du blé où des différences de synténie entre les chromosomes orthologues du blé et du riz sont plus importantes dans les régions distales fortement recombinogènes (Akhunov et al. 2003a; Akhunov et al. 2003b; See et al. 2006). De plus, les données de polymorphisme obtenues à partir des marqueurs SSR issus des extrémités de BAC de la banque 3B montre aussi un polymorphisme plus important dans les régions distales du chromosome.

6 Effet de la position distale sur la distribution des CO

Nous avons étudié le taux de CO sur un bivalent 3B délété pour lequel 45 % des régions distales du bras court sont absentes afin de déterminer si le fait de positionner des régions situées initialement en position intermédiaire où la fréquence de CO est faible sur le chromosome entier en position distale sur le chromosome délété entraînait une augmentation de la formation des CO dans ces régions. Ceci dans le but de déterminer l'influence de la position d'une région sur l'axe centromère-télomère et de déterminer ainsi si la position distale d'une région est responsable des forts taux de CO observés dans les régions distales. Le bras court du bivalent délété de l'individu F2CsRe3BS4 n°369 étant composé d'une région homozygote Chinese Spring distale et d'une région hétérozygote Chinese Spring/Renan, l'utilisation de marqueurs moléculaires nous a contraints à étudier le taux de CO uniquement sur la partie hétérozygote. Nous avons noté une très légère augmentation d'un facteur 1.9 du taux de CO au niveau de cette région. Cette augmentation n'est cependant pas significative. Ceci montre que l'impact de la délétion sur la recombinaison au niveau de cette région est très faible.

Lors d'une expérience similaire, Qi et al. (2002) avaient également observé une augmentation du taux de CO d'un facteur 1.6 au niveau de la région hétérozygote proximale à la délétion 5BL11-0.59. Cependant, en étudiant le nombre de chiasmas par cytogénétique, ils ont noté que la distance génétique de la partie homozygote distale qui couvre une région d'environ 25 Mb (4 % du bras long) subissait en revanche une augmentation de 21.9 cM comparée à sa taille génétique lorsque la région était en position intermédiaire sur le chromosome 5B entier. Les auteurs ont ainsi conclu que le positionnement distal des régions entraînait une augmentation importante du taux de CO et était un des facteurs majeurs expliquant la distribution des CO sur les chromosomes de blé. La même conclusion a été obtenue par Jones et al. (2002) lors de l'étude d'un chromosome 1B délété.

L'étude du taux de CO au niveau de la partie la plus distale à l'état homozygote n'ayant pas été réalisée sur le chromosome 3B, nous ne pouvons pas conclure quant à l'effet de ce facteur sur le taux de CO dans la région distale du bin 3BS1-0.33-0.55. Les résultats obtenus par Qi et al. (2002) et par Jones et al. (2002) sont cependant contradictoires avec ceux récemment publiés par Lukaszewski (2008) qui montre que la fréquence de CO est déterminée par le type de séquences et non par la position sur les chromosomes.

Dans l'étude de Qi et al. (2002), la région de 25 Mb, où l'augmentation importante du taux de CO est observée sur le chromosome délété, présente un taux de CO très important sur le chromosome entier. Le type de séquences présent dans cette région serait ainsi « aptes » à la formation de CO et le passage de cette région en position distale et à l'état homozygote pourrait expliquer l'augmentation du taux de CO observée au sein de cette région. Cependant, il est impossible de dire si l'accroissement du taux de CO observé par ces auteurs est dû à l'état homozygote ou à la position distale. L'effet de la position distale reste donc encore inconnu.

Afin de répondre plus précisément à l'impact de ces différents facteurs, homozygotie des régions, position des régions sur l'axe centromère-télomère et type de séquence, nous envisageons deux types d'analyses :

(1) Dans un premier temps, il serait intéressant de poursuivre l'étude sur le bivalent 3B délété en étudiant le nombre de chiasmas présents sur la partie homozygote du bras court délété du bivalent 3B pour déterminer si le type de séquences proximales a réellement la capacité d'inhiber la formation des CO. En effet, la région située en dessous de la fraction 3BS4-0.55 est une région faiblement recombinogène sur le chromosome entier (fréquence de CO de 0.02 cM/Mb) et semble donc faire partie des types de séquences qui inhibent la formation des CO.

Pour étudier cytologiquement les chiasmas sur le bras court délété, nous avons développé une sonde spécifique du bras long du chromosome 3B pour pouvoir identifier et suivre le comportement de ce chromosome au sein des cellules méiotiques sur des descendants de l'individu F2CsRe3BS4 n°369. Si la région homozygote du bin 3BS4-0.55 conserve la propriété d'inhiber la formation des CO même à l'état homozygote et placée en position distale, elle ne devrait permettre la formation que d'un nombre très limité de CO. Ainsi, si le nombre de chiasmas ou CO reste identique à celui observé sur les 55 % proximaux du bras court sur le chromosome entier, les résultats seront en accord avec ceux obtenus par Lukaszewski (2008) sur un bras de chromosome de seigle, ce qui confirmerait que le type de séquence est le facteur prépondérant pour la formation des CO sur les chromosomes de blé sans conclure cependant sur l'impact des facteurs homozygotie et position distale. En revanche, si le nombre de chiasmas augmente, cela pourra être dû soit au passage en position distale de la séquence soit à l'homozygotie de la région, ce matériel particulier n'étant pas le mieux adapté car la région distale reste obligatoirement partiellement homozygote ce qui ne permet pas de conclure quant à l'effet de la position distale seule (Jones et al. 2002; Qi et al. 2002).

(2) Pour déterminer alors réellement l'impact de la position distale, il serait intéressant d'étudier le nombre de chiasmas sur les régions distales des différentes lignées de délétion du bras court du chromosome 3B. Ces analyses nous permettraient de séparer les facteurs homozygotie et position distale.

En effet, le facteur homozygotie ne serait pas pris en compte puisque dans ce cas nous ne travaillerions que sur des lignées totalement homozygotes Cs. De plus, nous pourrions écarter l'effet type de séquences en étudiant uniquement les délétions pour lesquelles les régions situées en position proximale à la délétion présentent des fréquences de CO élevées sur le chromosome entier comme la lignée 3BS3-0.87 ou encore 3BS7-0.75.

7 Développement d'outils de génomique, cartographie génétique et modification de la fréquence de CO

7.1 Développement d'outils de génomique

La complexité du génome du blé tendre (*Triticum aestivum*, taille de 17 Gb, Zonneveld et al. 2005, 80-85% de séquences répétées et hexaploïdie) a longtemps limité le développement d'outils de génomique chez cette espèce ralentissant ainsi les recherches sur la compréhension du fonctionnement et de la structure de son génome. Le développement de la carte génétique Chinese Spring x Renan (CsRe) comprenant 102 marqueurs sur un seul chromosome et dont les données de ségrégation sont issues d'une même population représente la première carte génétique d'une telle densité sur un chromosome de blé. Son développement a été possible grâce à la mise en place de la carte physique du chromosome 3B (Paux et al. 2008) qui a entraîné le développement d'un nombre important de marqueurs. Notamment, les marqueurs ISBP récemment mis au point (Paux et al. 2006) ont permis de combler certains vides présents sur la carte génétique et ainsi, malgré leur dominance, ils se sont avérés être très intéressants pour couvrir des « déserts » de marqueurs et ainsi densifier la carte génétique.

La comparaison de cette carte génétique avec la carte physique de délétion a montré que l'ordre des locus était très conservé entre ces deux cartes. Ceci résulte sûrement de l'utilisation commune de la variété Chinese Spring en tant que parent pour la population de cartographie génétique et pour l'élaboration des lignées de délétion (Endo et Gill 1996).

La carte physique du chromosome 3B étant également développée à partir de la variété Cs, la carte génétique CsRe a été utilisée pour ordonner les contigs de la carte physique du chromosome 3B (Paux et al. 2008). Ainsi, 27 contigs portant 49 EST et couvrant 25 Mb localisés dans la partie distale du bras court ont pu être ordonnés et comparés avec la séquence du chromosome 1 de riz. Cela a permis d'identifier quatre inversions entre le blé et le riz ainsi que des gènes localisés en position non colinéaire. L'utilisation de cette carte génétique associée à la carte physique confirme ainsi mais avec une meilleure résolution, les réarrangements observés par ailleurs (La Rota et Sorrells 2004) et suggère que ces réarrangements locaux sont plus fréquents qu'attendus.

Jusqu'à présent, la carte de référence internationale était basée sur la population issue du croisement entre W7984 et Opata. Cette population avait été choisie car les marqueurs développés montraient un taux de polymorphisme élevé entre les deux variétés (environ 80 %, Roder et al. 1998) grâce à une procédure optimisée pour la recherche de microsatellites polymorphes, W7984 étant un blé synthétique (issu du croisement entre un blé dur, Altar 84 et une accession d'*Ae tauschii* apportant le génome D). Dans notre étude, nous avons montré que les microsatellites dérivés des séquences d'extrémités des BAC de la banque du chromosome 3B mais aussi ceux étudiés au niveau du locus *Sr2*, montrent le même taux de polymorphisme entre les variétés Chinese Spring et Renan qu'entre les variétés W7984 et Opata. Ainsi, en plus de l'excellente corrélation entre l'ordre des marqueurs sur la carte physique et sur la carte génétique et de la disponibilité de banques BAC de grands fragments pour les deux variétés Cs et Re, les parents de cette population présentent un taux de polymorphisme important. Il semble donc que cette population CsRe soit un excellent candidat pour ancrer les cartes physiques en développement pour les vingt autres chromosomes dans le cadre du consortium international de séquençage du génome du blé tendre (www.wheatgenome.org). Une population d'environ 2000 lignées recombinantes issues de ce croisement et d'ailleurs en cours de développement dans le cadre du projet Européen TriticeaeGenome (C Feuillet, communication personnelle).

Avec plus de 100 marqueurs sur ce chromosome 3B, cette carte génétique représente de plus un potentiel de marqueurs très important pour les personnes impliquées dans des projets de clonage positionnel de gènes sur ce chromosome.

7.2 Cartographie génétique

Cette carte CsRe couvre environ 180 cM ce qui a permis de déduire la présence de 3.6 CO par bivalent 3B à chaque méiose. Ces données sont légèrement supérieures aux 2.97 chiasmas observés sur le bivalent 3B à chaque méiose par Sallee et Kimber (1978). Cette différence pourrait être due à la difficulté d'estimer le nombre exact de chiasmas sur un chromosome. En particulier, lorsque deux CO se produisent à une distance proche, ils pourraient n'être visualisés que par un seul chiasma. Cette configuration est d'autant plus importante du fait de la grande taille du chromosome 3B qui facilite la formation de double CO sur le même bras. La comparaison entre le nombre de chiasmas observé génétiquement dans notre étude et la valeur estimée par cytogénétique montre que nous avons vraisemblablement pris en compte l'ensemble des CO présents sur le bivalent 3B et que l'étude de la distribution des CO sur le chromosome entier est de bonne qualité.

Avec une moyenne de 48.3 chiasmas sur l'ensemble des 21 chromosomes (Sallee et Kimber 1978), la taille estimée des cartes génétiques du génome du blé tendre devrait avoisiner les 2400 cM. Compte tenu de la légère sous-estimation du nombre de chiasmas que nous avons observé sur le chromosome 3B, les tailles des cartes génétiques devraient plutôt être comprises entre 2400 et 3000 cM. La taille actuelle des cartes qui couvrent le génome entier, le nombre de groupes de liaison étant égal au nombre de chromosomes, est compris entre 2500 cM et 4000 cM, excédant ainsi pour certaines les valeurs attendues (Akbari et al. 2006; Quarrie et al. 2005; Sourdille et al. 2003; Suenaga et al. 2005; Torada et al. 2006; Xue et al. 2008).

Notre analyse du coefficient de coïncidence montre de l'interférence positive importante pour des intervalles inférieurs à 10 cM et ainsi l'absence de doubles CO dans de tels intervalles. L'ajout de marqueurs dans ces intervalles ne devrait donc pas entraîner d'augmentation de leur taille génétique. Aux 102 marqueurs de départ présents sur la carte F2CsRe, nous avons tout récemment ajouté 233 marqueurs DArT (Diversity Array Technology, http://www.diversityarrays.com/) sans modifier significativement la taille génétique du chromosome (180 cM par rapport à 178.9 cM) confirmant l'absence de doubles CO sur de petits intervalles. La surestimation de la taille des cartes génétiques avec une densité de marqueurs importante résulte ainsi plus probablement d'erreurs de génotypage que de la détection de doubles CO. Ceci souligne la vigilance nécessaire lors de l'exploitation des distances génétiques retrouvées sur les cartes génétiques présentes dans la bibliographie.

La présence de cartes génétiques de taille importante pourrait aussi provenir de facteurs génétiques, présents chez les parents de ces populations, augmentant le taux global de CO sur les chromosomes. La différence de taille des cartes génétiques CsRe et CsCt où chaque donnée de génotypage au voisinage d'un CO a été vérifiée soutient cette hypothèse.

De tels facteurs semblent exister chez les céréales. Un dispositif de type diallèle où quatre parents de blé dur ont été systématiquement inter-croisés indique que les populations, présentant spécifiquement un des parents, montrent des taux de CO plus importants (P. Roumet, communication personnelle). De plus, des études de QTL classiques utilisant le nombre de CO comme caractère quantitatif afin d'identifier les régions du génome impliquées dans la variation du taux de CO ont été menées par Esh et al. (2007) et Bovill et al. (2008). Ils ont identifiés des régions sur plusieurs chromosomes notamment le chromosome 2D du blé et le chromosome 6H de l'orge. Des gènes candidats potentiels impliqués dans le mécanisme de recombinaison comme le gène *RAD54* ont d'ailleurs été cartographiés près de certains de ces QTL (Bovill et al. 2008) et leur analyse détaillée est en cours (J. Able, communication personnelle).

L'isolement de variétés voire de régions entraînant la présence de taux de CO important serait très utile au cours des programmes de sélection pour favoriser la formation de CO entre espèces éloignées mais aussi pour diminuer la taille des régions échangées entre les deux espèces (Feuillet et al. 2008; Martinez-Perez et Moore 2008).

Dans le but d'isoler de telles régions, il serait intéressant de regarder s'il existe une variabilité importante au sein de la diversité des Triticées en termes de taux de CO.

Ainsi, nous pourrions intégrer lors de l'analyse de l'élongation des parties distales du chromosome 3B en utilisant les sondes FISH (décrites dans le paragraphe 1 p 205), un ensemble de lignées sélectionnées sur la base de carte génétique de taille importante. Si la décondensation de la chromatine dans les parties distales est responsable des fortes fréquences de CO observées dans ces régions alors l'analyse de la collection d'individus nous permettrait de voir s'il existe de la variabilité en termes de taille de décompaction de la chromatine entre variétés afin d'isoler celles présentant la décondensation la plus importante et vraisemblablement la plus forte fréquence de CO.

Une recherche des déterminants génétiques permettant des taux de CO plus importants pourrait ensuite être conduite par une recherche de QTL en étudiant la descendance du croisement entre la variété présentant la plus longue élongation avec celle présentant la plus faible.

7.3 Modification de la fréquence de CO dans les régions proximales

L'étude de la distribution des CO sur le chromosome 3B montre que la fréquence de CO peut être très faible dans les régions les plus distales comme à l'extrémité du bras court ou au contraire être modérée dans les parties intermédiaires voire proximales comme dans le bin 3BL1-0,31-0,38. Ainsi, bien que la majorité des gènes clonés chez le blé se localise dans les parties distales, le clonage de gènes pourrait se révéler difficile dans ces régions du fait d'une variation locale de la fréquence de CO alors qu'il pourrait être facilité dans des régions intermédiaires où cette fréquence est ponctuellement accrue.
Cependant, les très faibles fréquences de CO observées dans les régions proximales restent un frein considérable au clonage positionnel de gènes dans le voisinage des centromères de même qu'elles sont un handicap pour ordonner les contigs des cartes physiques au niveau de ces régions.

Pour surmonter ce problème des faibles fréquences de CO dans les régions proximales, plusieurs solutions pourraient être envisagées :

(1) Une première façon de résoudre ce problème serait d'augmenter le nombre de plantes analysées de façon à accroître la probabilité d'identifier des individus montrant des événements de CO dans ces régions. Cependant, des estimations basées sur les données déjà acquises (fréquences de CO dans les régions proximales entre 0 et 0.1 cM/Mb) suggèrent qu'il faudrait développer des populations supérieures à 15 000 individus pour obtenir un CO tous les 500 kb au niveau des régions où la fréquence est de 0.1 cM/Mb et beaucoup plus pour les autres régions. Si cela est techniquement faisable, cette solution serait longue et laborieuse.

(2) Une deuxième façon consiste à travailler non plus sur une liaison génétique entre les marqueurs mais sur une liaison physique qui ne dépend pas de la distribution des CO. Ceci peut être obtenu soit en développant un panel d'échantillons pour réaliser du « happy mapping » (Riera-Lizarazu et al. 2008) soit en travaillant à l'aide d'hybrides d'irradiation qui sont en cours de développement pour le chromosome 3B (S. Kianian, communication personnelle).

(3) Une troisième façon pourrait être d'induire une augmentation de la fréquence des CO dans ces régions proximales. Deux articles récents proposent des solutions pour augmenter la fréquence de CO sur les chromosomes (Martinez-Perez et Moore 2008; Wijnker et de Jong 2008). Wijnker et de Jong (2008) parlent ainsi de modifications structurales des chromosomes pour augmenter la fréquence des CO sur certaines parties. Cependant, les modifications que nous avons entraînées (passage en position distale et passage à l'état homozygote Cs) lors de l'étude du chromosome 3B ne provoquent pas d'augmentation de la fréquence des CO et semblent donc ne pas être adaptées à ce génome. Martinez-Perez et Moore (2008) proposent plusieurs autres solutions, comme inhiber l'interférence entre CO, augmenter le temps des premiers stades de méiose où ont lieu la formation des CO ou encore utiliser des mutants qui provoquent un déplacement de la distribution des CO sur les chromosomes.

Cette dernière méthode semble particulièrement intéressante pour induire la formation de CO dans des régions inhabituelles et pourrait être utilisée chez le blé pour provoquer la formation de CO dans les régions proximales. En effet, Martinez-Perez et Moore (2008) ainsi que Kirik et al. (2006), Mieczkowski et al. (2007) et Tsai et al. (2008) montrent qu'un nombre important de mutants chez les levures, chez *C. elegans* et chez *Arabidopsis thaliana* provoquent un déplacement ou une augmentation de la formation des CO sur les chromosomes. Ces gènes codent pour des protéines qui entraînent une modification des marques épigénétiques de la chromatine ou qui induisent des perturbations de la compaction de la chromatine. L'utilisation de tels mutants pourrait s'avérer efficace pour augmenter les fréquences de CO dans les régions proximales où la chromatine particulièrement compactée semble entraîner les faibles fréquences de CO. A partir de l'ensemble de ces gènes présents chez les autres espèces, nous pourrions faire une recherche dans les bases de données de blé pour déterminer leur existence au sein du génome de blé. Après l'isolement de ces gènes et leur caractérisation, nous pourrions muter les plus intéressants par RNAi et déterminer les conséquences sur la distribution des CO et notamment si de tels mutants provoquent une augmentation de la fréquence de CO dans les régions proximales.

8 Immuno-localisation des protéines intervenant dans la résolution des jonctions de Hollyday

L'utilisation d'anticorps dirigés contre la protéine MLH1 est une technique qui se développe considérablement.

Elle ne s'avère pas extrêmement précise pour la localisation des CO sur les chromosomes mais elle permet d'avoir une très bonne estimation du nombre de CO présents sur chaque chromosome comparée à la visualisation des chiasmas qui reste difficile. Cette protéine marquant vraisemblablement les CO de classe I présentant de l'interférence, l'utilisation d'anticorps dirigés contre les protéines MUS81 ou MMS4 qui marquent les CO de classe II permettrait une analyse exhaustive du nombre de CO présents sur les chromosomes.
Bien que cette technique ne soit pas encore mise au point chez le blé, la visualisation des foci MLH1, MUS81 ou MMS4 permettrait d'observer l'ensemble des CO présents sur les chromosomes sans ambigüité et pourrait s'avérer être la meilleure technique pour étudier la diversité du taux de CO intra-espèces et utiliser l'observation de ces foci comme caractère phénotypique dans les analyses de QTL de taux de CO (2) pour déterminer le nombre de CO présents sur les lignées de délétion du bras court pour l'étude de l'impact de la position distale.
De plus, cette technique couplée au marquage du complexe synaptonémal nous permettrait de déterminer si la longueur du complexe synaptonémal est corrélée avec le nombre de CO présents sur les chromosomes comme cela a été observé chez la souris (Petkov et al. 2007).

Le développement d'une telle technique chez le blé semble ainsi nécessaire pour mieux comprendre la distribution des CO sur les chromosomes.

9 La taille physique cumulée des gènes gouverne la distribution des CO

Chez les espèces dont le génome est séquencé comme l'humain, la souris, *Arabidopsis thaliana* ou encore le riz, les caractéristiques de la séquence telles que la présence de gènes, le pourcentage en GC ou encore la présence d'éléments transposables jouent un rôle dans la distribution des CO sur les chromosomes (Drouaud et al. 2006; Kong et al. 2002; Paigen et al. 2008; Wu et al. 2003). Afin d'analyser pour la première fois chez le blé l'influence de ces facteurs sur la distribution des CO, nous avons étudié la distribution des CO dans deux populations différentes, Chinese Spring x Renan et Chinese Spring x Courtot, au niveau d'une région séquencée de 3,1 Mb située en position distale sur le bras court du chromosome 3B, le contig *Rph7*.

Au sein de cette région, l'étude d'une dizaine de sous-intervalles délimités par des couples de marqueurs montre sans ambiguïté que la distribution des CO est corrélée de manière significative avec la taille cumulée de l'ensemble des gènes au sein de chaque intervalle, facteur qui peut expliquer jusqu'à 82 % de la variation de la fréquence de CO sur le contig. En d'autres termes plus l'intervalle est riche en gènes de grande taille plus il contient de CO. De plus, le découpage fin des sous-régions 8 et 9 montre que la majorité des CO est localisée dans des régions de quelques kb entourant les gènes. Ainsi, sur le contig *Rph7*, les gènes jouent un rôle déterminant dans la distribution des CO.

La distribution des CO observée à partir de la population CsRe montre cependant une corrélation plus faible avec la taille cumulée des gènes comparée aux données issues de la population CsCt. Les CO observés à partir de la population CsCt sont présents uniquement sur le gène (du codon ATG au codon stop) alors que ceux observés sur la population CsRe sont présents aussi en amont dans la région promotrice, ce qui pourrait expliquer cette différence. Cependant celle-ci reste difficile à comprendre du fait du lien de pedigree entre les variétés Renan et Courtot, les séquences identiques retrouvées au niveau des gènes *HGA3*, *HGA2*, *HGA1* et *GAD1*, et seulement deux marqueurs microsatellites polymorphes entre les variétés Ct et Re. La variété Renan est connue pour porter un nombre important de gènes de résistance à diverses maladies du blé et le locus *Rph7* contient justement des gènes de résistance comme le gène *FHB1* qui confère une résistance à la fusariose. Il se peut donc que l'expression de ces gènes soit différente entre la variété sensible Courtot et la variété Renan ce qui pourrait impliquer une variation dans la distribution des CO dans cette région.

L'importance des gènes dans la localisation des CO a déjà été mentionnée auparavant chez le blé. Lors d'expériences de clonage positionnel, la relation entre la présence de gènes et la présence de CO a été observée à travers l'étude des relations entre distances physiques et distances génétiques sur quelques centaines de kb autour des gènes d'intérêt. En particulier, Stein et al. (2000) ont montré que sur un intervalle de 30 kb situé proche du gène *LR10* et contenant trois ORF, la fréquence de CO était importante (1.6 cM/Mb) comparée aux régions voisines. En plus de son importance à l'échelle de quelques Mb, la présence de gènes est vraisemblablement un facteur important pour expliquer la distribution des CO à l'échelle des chromosomes entiers. Ainsi, Erayman et al. (2004) ont montré que 95 % de la recombinaison se localisaient dans les zones riches en gènes.

Sur le chromosome 3B, ils ont identifié 7 régions riches en gènes, quatre sur le bras long et trois sur le bras court (Figure 30 p 210). Sur le bras long, nous avons trouvé que la majorité des CO est localisée dans le bin 3BL7-0.63-1.00 qui correspond à la plus importante région riche en gènes du bras (3L0.9) identifiée par Erayman et al. (2004). De plus, nous observons une augmentation de la fréquence de CO dans le bin 3BL1-0.31-0.38 qui correspond à la région riche en gènes majeure du milieu du bras long (3L0.3). À l'inverse, au niveau du bras court, nous n'observons pas de relation entre la distribution des CO au sein de notre étude et la localisation des régions riches en gènes décrite par Erayman et al. (2004). Nous avons observé que la majorité des CO est localisée dans le bin de délétion 3BS8-0.78-0.87 alors que la région riche en gènes majeure se localise dans le bin le plus distal 3BS3-0.87-1.00. Comme nous l'avons décrit auparavant, cette différence s'explique vraisemblablement par l'assignation non systématique des marqueurs aux bins de délétion par Erayman et al. (2004) car nous observons que certains marqueurs qu'ils localisent dans le bin le plus distal 3BS3-0.87-1.00 par déduction avec l'assignation d'autres marqueurs voisins sont en fait assignés au sein de notre étude dans le bin 3BS8-0.78-0.87. Ainsi, la région riche en gènes qu'ils décrivent dans le bin 3BS3-0.87-1.00 semble plutôt se localiser dans le bin 3BS-0.78-0.87 qui présente le plus fort taux de CO.

En accord avec les données d'Erayman et al. (2004), le fait que nous observons une corrélation importante entre la présence de gènes et les fortes fréquences de CO ; Celles-ci étant plus faibles dans les régions proximales, cela suggère que la densité et/ou la taille des gènes pourraient être plus faibles dans les régions proximales comparées aux régions plus distales où la fréquence de CO est importante. Une telle distribution des gènes vient d'être clairement montrée chez le sorgho (Patterson et al. 2009) où le même pattern de distribution des CO que chez le blé est observé (Kim et al. 2005). Seule la séquence du chromosome entier permettra une analyse exhaustive de l'organisation des séquences géniques. Cependant, l'absence de gènes au sein du contig *L* (620 kb) présent dans les régions proximales, la densité importante de un gène tous les 60 kb au sein du contig *Rph7* mais également la densité importante d'EST dans les régions distales observée après leur assignation dans les bins de délétion semblent être en faveur de cette hypothèse.

La présence de gènes ainsi que leurs tailles semblent être ainsi un des facteurs majeurs contrôlant la distribution des CO sur une région donnée chez le blé. Plusieurs hypothèses peuvent expliquer l'influence des gènes sur la distribution des CO.

Premièrement, les gènes sont des régions présentant des marques épigénétiques différentes du reste du génome en particulier des faibles taux de méthylation de l'ADN (Zhang 2008) pouvant entraîner une compaction plus faible de la chromatine comparée aux régions riches en éléments transposables qui présentent un fort taux de méthylation et une compaction plus importante. Cette compaction moins importante pourrait favoriser l'accès aux facteurs impliqués dans la recombinaison et donc la formation des CO. Deuxièmement, comme pour les points chauds de type α définis par Petes (2001) chez la levure et l'humain, la fixation de facteurs de transcription sur les gènes même s'ils n'induisent pas forcément une transcription, pourrait entraîner également une ouverture de la chromatine. Enfin, le pourcentage d'identité de séquences entre les deux homologues étant un facteur important pour permettre la formation des CO, les gènes constituent les régions les mieux conservées entre les génomes à cause des fortes pressions de sélection. Ils seraient donc les sites préférentiels de reconnaissance entre chromosomes homologues et de formation des CO au sein d'un génome comme celui du blé qui contient environ 85 % de séquences répétées qui elles, évoluent beaucoup plus rapidement et montrent donc moins d'identité entre différentes variétés.

Afin de confirmer la relation entre CO et présence de gènes sur le locus *Rph7*, nous avons commencé un re-découpage de ce locus en encadrant précisément les îlots de gènes qu'il comporte comme nous l'avons fait pour les sous-régions 8 et 9 afin de déterminer si, comme sur ces deux sous-régions, la majorité des CO est localisée dans les quelques kb entourant les gènes. De plus, cela nous permettrait également de déterminer où et combien de CO se localisent dans des zones sans gènes et de comprendre quels sont ces autres régions qui subissent des CO. Au niveau de la sous-région 8, trois CO se trouvent dans une zone de 124 kb dépourvu de gène. La présence de plusieurs satellites dans cette région pourrait en être la raison comme cela a déjà été observé chez l'humain (Jeffreys et al. 2003). Le découpage fin de cette région est en cours ce qui nous permettra de conclure quant à cette hypothèse. L'étude d'autres régions du chromosome semble aussi indispensable pour déterminer si cette relation observée est caractéristique du locus *Rph7* ou si elle est conservée sur le reste du chromosome. L'étude sur d'autres contigs séquencés présents sur la partie distale du bras long est également en cours de réalisation.

Il serait également intéressant de travailler de manière plus approfondie sur le gène *HGA3*. Nous avons montré qu'il présente la plus forte fréquence de CO au sein des sous-régions 8 et 9 et que les CO se répartissent sur l'ensemble du gène sur la population F2CsRe. Cependant, les effectifs de CO obtenus au sein de cette région restent trop faibles pour pouvoir donner des conclusions significatives sur la distribution des CO sur ce gène. Pour obtenir un nombre de CO plus important et étudier finement leur distribution sur ce gène, l'utilisation de la technique du typage de pollen pourrait être envisagée comme cela a déjà été fait sur d'autres espèces (Llaurens et al. 2008). De plus, grâce aux effectifs développés, nous pourrions également étudier la fréquence de conversion génique (NCO) et la comparer avec la fréquence de CO. En effet, nous disposons maintenant de cinq SNP entre les génotypes Cs et Re ou Ct sur la séquence du gène entre le codon ATG et le codon stop couvrant 2954 pb dont trois sont séparés par 200 pb au niveau des exons 4 et 5 où la fréquence de CO est la plus importante. Chez la levure, les points de conversion génique représentent une part importante en terme de taille de séquence totale du génome à chaque méiose (Mancera et al. 2008). Ainsi, il serait intéressant d'étendre l'analyse à plusieurs régions chaudes du locus *Rph7* pour déterminer si les résultats observés chez la levure sont similaires chez le blé et ainsi déterminer si les NCO peuvent jouer un rôle important dans l'évolution du génome du blé.

Enfin, il serait intéressant de comprendre pourquoi la distribution des CO est différente entre les populations CsRe et CsCt et notamment au sein de la région de 23 kb entourant le gène *HGA3*. La comparaison de la séquence du gène *HGA3* entre Ct et Re montre qu'elles sont similaires alors que la distribution des CO est différente au sein du gène. Ce paramètre ne semble donc pas être impliqué dans la distribution des CO dans ce cas. Cependant, il serait intéressant d'établir la séquence de Re et de Ct au niveau de la région promotrice du gène pour voir s'il existe des différences qui pourraient expliquer l'absence de CO sur la population CsCt. Nous voudrions de plus évaluer l'expression de ce gène pendant les premiers stades méiotiques pour déterminer si une des deux variétés montrent une expression différente qui pourrait être à l'origine des différences de distribution des CO. Cette analyse pourrait également être étendue à d'autres sous-régions du contig qui présentent une différence de fréquence de CO mais également une fréquence de CO identique entre CsCt et CsRe afin de voir si l'expression des gènes a un impact sur la distribution des CO.

10 Les caractéristiques de la séquence ne sont pas fortement corrélées avec la distribution des CO

Chez la plupart des organismes comme chez l'humain, la souris ou encore l'abeille, la distribution des CO est corrélée avec le pourcentage de GC (Beye et al. 2006; Kong et al. 2002; Paigen et al. 2008).

Au niveau de la région entourant le gène *HGA3*, nous observons également une corrélation ($R^2 = 0.66$) entre le pourcentage de GC et la distribution des CO avec les données issues de la population CsCt. Chez les organismes où une telle corrélation est observée, celle-ci semble attribuée au mécanisme appelé biais de conversion génique (BGC pour biased gene conversion) qui entraîne une réparation biaisée vers le G ou C lors des appariements non corrects de type A-C, A-G ou T-G, T-C qui se produisent pendant les événements de conversion génique (NCO) à la méiose (Galtier et al. 2001). Ce mécanisme a été également observé chez les Triticées (Haudry et al. 2008), ce qui pourrait ainsi expliquer la corrélation que nous avons identifiée à partir des données issues de la population CsCt sur la région encadrant le gène *HGA3*.

Cependant, l'analyse de la deuxième population (CsRe) pour les mêmes sous-intervalles encadrant le gène *HGA3* ainsi que l'étude à l'échelle du contig *Rph7* entier montre que la distribution des CO n'est plus corrélée avec le pourcentage de GC ni avec le ratio CpG, le pourcentage d'éléments transposables et le pourcentage de SSR. Les corrélations définies ont été réalisées à partir des données de séquence issues de la variété Chinese Spring. A la vue des résultats de polymorphisme des marqueurs moléculaires entre Cs et Re que ce soit à l'échelle du chromosome entier ou à l'échelle du contig *Rph7*, la séquence de la variété française Renan présente vraisemblablement de nombreuses différences avec la séquence de Cs. Ceci pourrait éventuellement entraîner une différence de localisation des CO sur le contig *Rph7* entre des individus portant un bivalent 3B composé de deux chromosomes provenant de la variété Cs et des individus portant un bivalent 3B avec un chromosome provenant de la variété Cs et l'autre de la variété Re. Des résultats confirmant cette hypothèse ont été obtenus chez le maïs au niveau du locus *bronze* (Dooner et He 2008) où des variations importantes des fréquences de CO ont été observées en fonction des génotypes utilisés lors des croisements et notamment lors de la présence ou de l'absence de blocs d'éléments transposables entre les deux génotypes.

Ainsi, si l'on prend l'exemple du pourcentage de GC, la distribution des CO sur un bivalent Cs doit entraîner la distribution du pourcentage de GC observé sur la séquence Cs que nous avons analysée. Cependant, la distribution des CO sur un bivalent hétérozygote Cs/Re probablement différente de celle observée sur un bivalent homozygote Cs entraîne une distribution du pourcentage de GC différente de celle observée sur un bivalent homozygote Cs et donc différente de celle observée sur la séquence issue de Cs. Ce décalage entre la séquence issue d'une variété et la distribution des CO obtenue à partir de deux génotypes pourrait expliquer l'absence de corrélation.

11 L'hétérozygotie Cs/Re impacte la distribution des CO sur le contig *Rph7*

Bien que la fréquence globale de CO soit 2.6 fois plus faible sur la population CsCt par rapport à la population CsRe, la variation de fréquence de CO le long du contig *Rph7* reste conservée entre les deux populations. Au sein des régions 7 et 11 (ou 11 et 12 pour CsRe), la fréquence de CO est très faible sur les deux populations.

Au niveau de la région 7, l'ensemble des marqueurs spécifiques au chromosome 3B est absent chez les variétés Courtot et Renan suggérant la présence d'une délétion d'environ 500 kb chez ces deux variétés ou d'une insertion chez la variété Chinese Spring. L'utilisation de certains de ces marqueurs sur une core-collection de 372 individus (Balfourier et al. 2007) et sur un ensemble d'accessions chinoises et européennes montre une diversité importante dans la présence ou non de cette délétion sur l'ensemble des génotypes étudiés (F. Balfourier, communication personnelle). Cette délétion est vraisemblablement à l'origine de la faible fréquence de CO retrouvée au sein de cet intervalle comme cela a été mentionné chez le maïs suite à la présence d'un bloc de rétrotransposons sur l'un des homologues au niveau du locus *bronze* (Dooner et He 2008) ou encore chez *C. elegans* (Hammarlund et al. 2005). Chez cette espèce, la présence d'une large insertion à l'état hétérozygote entraîne une chute locale de la fréquence de CO mais également une redistribution de la recombinaison de part et d'autre de l'insertion avec une augmentation de la fréquence de CO sur la région voisine située du coté du centre d'appariement. Les auteurs ont suggéré que comme l'appariement est initié au centre d'appariement et s'étend ensuite sur le chromosome, la présence de cette insertion empêcherait la progression de la synapse au sein de certaines cellules entraînant ainsi une augmentation des CO dans la région située entre le centre d'appariement et la délétion.

Dans notre étude, de fortes fréquences de CO sont également observées de part et d'autre de la région présentant la délétion. Bien que nous n'ayons pas étudié la distribution des CO dans cette zone avec cette délétion/insertion à l'état homozygote, ces données suggèrent que comme chez *C. elegans*, l'augmentation de la fréquence de CO dans les régions voisines à cette délétion serait le résultat de cette délétion. De fortes fréquences de CO étant présentes de part et d'autre de la délétion, ceci suggère également que la progression de la synapse pourrait se produire dans les deux sens (proximal-distal et distal-proximal).

Au sein des régions 11 ou 12, nous avons observé une fréquence de CO très faible (0.16 cM/Mb) pour les deux populations CsCt et CsRe. Nous observons aussi que 78 % des marqueurs de cette région montrent du polymorphisme entre CsRe et CsCt. Les séquences de Cs d'une part et de Re et Ct d'autre part sont donc très divergentes dans cette zone. La présence de polymorphisme entre homologues réduisant considérablement la formation des CO comme cela a été observé lors de croisements entre espèces éloignées (Ji et Chetelat 2007), ce facteur pourrait expliquer la faible fréquence de CO notée dans cet intervalle. Seul un séquençage comparatif de la même région chez les variétés Renan et Chinese spring permettrait de déterminer réellement le pourcentage de divergence de séquences entre les deux génotypes et son impact potentiel sur la distribution des CO.

VI Conclusion générale

L'ensemble de nos résultats apporte une nouvelle vision concernant la distribution des CO sur les chromosomes de blé et les facteurs qui impactent cette distribution. Nous avons observé des faibles fréquences de CO, inférieures à 0.1 cM/Mb, pour les régions entourant le centromère et couvrant environ 600 Mb soit les 2/3 du chromosome 3B. De plus parmi celles-ci, 27 % entourant le centromère sont quasiment dépourvus de CO. Ces analyses restent cependant limitées en raison de l'utilisation des lignées de délétion pour le calcul des distances physiques. La disponibilité de la séquence du chromosome 3B dans les années à venir va permettre une révolution dans l'étude de la distribution des CO sur les chromosomes de blé comme nous avons pu le constater sur la région séquencée du *Rph7* et chez les espèces dont le génome est séquencé et où cette étude a déjà été effectuée.

La présence des faibles fréquences de CO observées dans les régions proximales pose un sérieux problème pour ordonner les cartes physiques au sein de ces régions mais aussi pour le clonage de gènes. Une des structures chromosomiques connue comme inhibant la formation des CO est la présence d'hétérochromatine. Le pattern de C-banding (Figure 30 p 210) montre un fort taux d'hétérochromatine dans les régions proximales suggérant son rôle dans les faibles fréquences observées dans ces régions. Cependant, au niveau de certains bins de délétion comme le bin 3BL10-0.50-0.63 l'absence d'hétérochromatine associée aux faibles fréquences de CO relevées (0.05 cM/Mb) montre que d'autres facteurs sont importants pour expliquer la distribution des CO à l'échelle du chromosome entier.

L'influence de la position d'une région sur l'axe centromère-télomère semble vraisemblablement avoir un effet limité sur la formation des CO au sein de cette région mais l'impact réel reste encore à définir. Ces résultats montrent ainsi que la position proximale des régions n'est pas responsable des faibles fréquences de CO observées au sein de ces régions. De plus, la divergence de séquence entre les variétés Chinese spring et Renan n'est pas non plus à l'origine des faibles fréquences de CO retrouvées dans les régions proximales. L'absence de CO dans ces régions reste ainsi encore inexpliquée.

Au cours de notre étude, nous avons également relevé que la distribution des CO était fortement corrélée avec la présence des gènes au niveau du locus *Rph7* et notamment avec la taille cumulée des gènes sur un intervalle donné. Les données publiées récemment montrent que la structure de la chromatine est essentielle dans la distribution des CO (Kirik et al. 2006; Lopez et al. 2008; Tsai et al. 2008).

L'explication à la relation que nous observons pourrait ainsi provenir de la faible compaction de la chromatine au niveau des gènes. Celle-ci faciliterait en effet l'accès aux facteurs de la recombinaison et ainsi la formation des CO. La faible densité en gènes présents au niveau des régions proximales suggérée par Erayman et al. (2004) sucite que ces régions seraient fortement compactées, ce qui provoquerait un accès restreint aux facteurs impliqués dans la recombinaison et l'inhibition observée de la formation des CO. Cette compaction augmenterait de plus l'étendue de l'interférence sur de plus grandes distances physiques mesurées en Mb limitant davantage la formation de CO.

La structure de la chromatine qui vraisemblablement pourrait être naturellement différente en fonction du type de régions (riches ou pauvres en gènes) subit aussi des remaniements durant les premiers stades de la méiose. En effet, le mécanisme responsable de la reconnaissance entre chromosomes homologues, gouverné par le locus *Ph1*, entraîne une décondensation de la chromatine au niveau des régions subterminales des chromosomes. L'étendue de cette modification sur les bras des chromosomes n'a pas encore été déterminée précisement. Cependant, cette modification pourrait être fortement facilitée au niveau des régions où la densité en gènes est importante et au contraire inhibée au sein de régions pauvres en gènes. La combinaison entre le mécanisme de reconnaissance entre chromosomes homologues et le type de séquence présent sur le chromosome pourrait entraîner la décompaction et ainsi les fortes fréquences de CO observées au niveau des régions distales.
L'analyse entre la décondensation observée sur les chromosomes pendant les premiers stades de la prophase I et la distribution des CO nous permettra de vérifier si cette modification de la structure de la chromatine est le facteur majeur impliqué dans la distribution des CO. Dans l'avenir, la compréhension des mécanismes qui influencent cette décondensation pourrait avoir un impact important pour augmenter la fréquence de CO dans les régions proximales.

Lors de la comparaison entre cartes génétiques, nous avons observé des différences de taux de CO entre les croisements CsCt et CsRe suggérant de la diversité entre variétés en termes de taux de CO sur les chromosomes. Cette observation suscite un intérêt particulier car il implique la présence de facteurs pouvant augmenter le taux de CO sur les chromosomes. Chez le blé, Bovill et al. (2008) et Esch et al. (2007) ont déjà observé, par une analyse QTL, la présence de telles régions en utilisant les croisements disponibles dans les bases de données ou au sein de leur laboratoire. Cependant, il serait intéressant dans un premier temps de quantifier cette variabilité intra-espèce en utilisant par exemple la longueur d'élongation de la chromatine dans les premiers stades de méiose comme marqueur phénotypique. La visualisation des foci MLH1, MUS81 ou MMS4, bien que cette technique ne soit pas encore mise au point chez le blé, permettrait d'observer l'ensemble des CO présents sur les chromosomes et pourrait s'avérer être aussi une technique de criblage très intéressante pour étudier la diversité intra-espèces. L'isolement de tels facteurs favorisant la recombinaison serait un atout majeur pour développer un matériel végétal adapté pour augmenter le taux de CO au cours des programmes de sélection et ainsi accélérer la création variétale.

Enfin, pour améliorer voire contrôler la recombinaison dans les schémas de sélection, il apparaît ainsi indispensable de connaître les gènes qui interviennent dans ce processus de formation des CO. Si ceux-ci commencent à être bien connus chez les espèces modèles en particulier chez la levure et chez *Arabidopsis thaliana* (Mercier et Grelon 2008), chez le blé, les études se sont concentrées majoritairement sur le locus *Ph1* (Al-Kaff et al. 2008). Cependant, des gènes impliqués directement dans le processus de recombinaison tels que *MRE11* et *TaASY1* commencent a être étudiés plus en détail (Boden et al. 2008; Boden et al. 2007; Bovill et al. 2008; Crismani et al. 2006; de Bustos et al. 2007; Jenkins et al. 2008). Il semble donc pertinent de conduire des études à plus large échelle afin d'identifier et d'isoler les gènes impliqués dans le mécanisme de recombinaison chez le blé, de les localiser sur les chromosomes et d'étudier leur profil d'expression en particulier en regard des différentes copies homéologues (fonctionnalité). Pour cela, il semble tout à fait opportun de s'appuyer sur les données obtenues sur les espèces modèles et tout particulièrement celles issues d'*Arabidopsis*. Une fois les différentes copies de ces gènes isolées, il sera également intéressant d'aller mesurer la variabilité existante au sein de collections ou de rechercher des allèles nouveaux plus efficaces dans des populations de mutants. Ces derniers pourront de plus être exploités pour identifier de nouveaux gènes impliqués dans le mécanisme et qui pourraient être spécifiques à un organisme polyploïde ou au blé.

Le développement de nouvelles variétés élites répondant aux contraintes environnementales (moins gourmandes en eau, en engrais, moins sensibles aux ravageurs…) passe vraisemblablement par une meilleure exploitation des ressources génétiques disponibles et l'introgression d'allèles nouveaux performants ne pourra se faire sans un contrôle précis de leur intégration dans les génomes élites via une recombinaison efficace. Seule une étude exhaustive des gènes qui sont impliqués dans son processus ainsi qu'une compréhension détaillée des facteurs qui influencent la distribution des CO pourront amener à une meilleure maîtrise de ce phénomène chez le blé tendre.

VII Références

2000. Analysis of the genome sequence of the flowering plant Arabidopsis thaliana. *Nature* **408:** 796-815.

Able, J.A., P. Langridge, and A.S. Milligan. 2007. Capturing diversity in the cereals: many options but little promiscuity. *Trends Plant Sci* **12:** 71-79.

Akbari, M., P. Wenzl, V. Caig, J. Carling, L. Xia, S. Yang, G. Uszynski, V. Mohler, A. Lehmensiek, H. Kuchel, M.J. Hayden, N. Howes, P. Sharp, P. Vaughan, B. Rathmell, E. Huttner, and A. Kilian. 2006. Diversity arrays technology (DArT) for high-throughput profiling of the hexaploid wheat genome. *Theor Appl Genet* **113:** 1409-1420.

Akhunov, E.D., A.R. Akhunova, A.M. Linkiewicz, J. Dubcovsky, D. Hummel, G. Lazo, S. Chao, O.D. Anderson, J. David, L. Qi, B. Echalier, B.S. Gill, Miftahudin, J.P. Gustafson, M. La Rota, M.E. Sorrells, D. Zhang, H.T. Nguyen, V. Kalavacharla, K. Hossain, S.F. Kianian, J. Peng, N.L. Lapitan, E.J. Wennerlind, V. Nduati, J.A. Anderson, D. Sidhu, K.S. Gill, P.E. McGuire, C.O. Qualset, and J. Dvorak. 2003a. Synteny perturbations between wheat homoeologous chromosomes caused by locus duplications and deletions correlate with recombination rates. *Proc Natl Acad Sci U S A* **100:** 10836-10841.

Akhunov, E.D., A.W. Goodyear, S. Geng, L.L. Qi, B. Echalier, B.S. Gill, Miftahudin, J.P. Gustafson, G. Lazo, S. Chao, O.D. Anderson, A.M. Linkiewicz, J. Dubcovsky, M. La Rota, M.E. Sorrells, D. Zhang, H.T. Nguyen, V. Kalavacharla, K. Hossain, S.F. Kianian, J. Peng, N.L. Lapitan, J.L. Gonzalez-Hernandez, J.A. Anderson, D.W. Choi, T.J. Close, M. Dilbirligi, K.S. Gill, M.K. Walker-Simmons, C. Steber, P.E. McGuire, C.O. Qualset, and J. Dvorak. 2003b. The organization and rate of evolution of wheat genomes are correlated with recombination rates along chromosome arms. *Genome Res* **13:** 753-763.

Al-Kaff, N., E. Knight, I. Bertin, T. Foote, N. Hart, S. Griffiths, and G. Moore. 2008. Detailed dissection of the chromosomal region containing the Ph1 locus in wheat Triticum aestivum: with deletion mutants and expression profiling. *Ann Bot (Lond)* **101:** 863-872.

Anderson, K.A. and S.M. Stack. 2002. Meiotic Recombination in Plants. *Current Genomics* **3:** 507-525.

Anderson, L.K., G.G. Doyle, B. Brigham, J. Carter, K.D. Hooker, A. Lai, M. Rice, and S.M. Stack. 2003. High-resolution crossover maps for each bivalent of Zea mays using recombination nodules. *Genetics* **165:** 849-865.

Anderson, L.K., A. Lai, S.M. Stack, C. Rizzon, and B.S. Gaut. 2006. Uneven distribution of expressed sequence tag loci on maize pachytene chromosomes. *Genome Res* **16:** 115-122.

Anderson, L.K. and S.M. Stack. 2005. Recombination nodules in plants. *Cytogenet Genome Res* **109:** 198-204.

Arnheim, N., P. Calabrese, and I. Tiemann-Boege. 2007. Mammalian meiotic recombination hot spots. *Annu Rev Genet* **41:** 369-399.

Arumuganathan, K. and E.D. Earle. 1991. Nuclear DNA content of some important plant species. *Plant Mol Biol Rep* **9:** 208-218.

Bagshaw, A.T., J.P. Pitt, and N.J. Gemmell. 2008. High frequency of microsatellites in S. cerevisiae meiotic recombination hotspots. *BMC Genomics* **9:** 49.

Balfourier, F., V. Roussel, P. Strelchenko, F. Exbrayat-Vinson, P. Sourdille, G. Boutet, J. Koenig, C. Ravel, O. Mitrofanova, M. Beckert, and G. Charmet. 2007. A worldwide bread wheat core collection arrayed in a 384-well plate. *Theor Appl Genet* **114:** 1265-1275.

Barchi, L., J. Bonnet, C. Boudet, P. Signoret, I. Nagy, S. Lanteri, A. Palloix, and V. Lefebvre. 2007. A high-resolution, intraspecific linkage map of pepper (Capsicum annuum L.) and selection of reduced recombinant inbred line subsets for fast mapping. *Genome* **50:** 51-60.

Barton, A.B., M.R. Pekosz, R.S. Kurvathi, and D.B. Kaback. 2008. Meiotic recombination at the ends of chromosomes in Saccharomyces cerevisiae. *Genetics* **179:** 1221-1235.

Barton, A.B., Y. Su, J. Lamb, D. Barber, and D.B. Kaback. 2003. A function for subtelomeric DNA in Saccharomyces cerevisiae. *Genetics* **165:** 929-934.

Basheva, E.A., C.J. Bidau, and P.M. Borodin. 2008. General pattern of meiotic recombination in male dogs estimated by MLH1 and RAD51 immunolocalization. *Chromosome Res* **16:** 709-719.

Bassing, C.H., W. Swat, and F.W. Alt. 2002. The mechanism and regulation of chromosomal V(D)J recombination. *Cell* **109 Suppl:** S45-55.

Baudat, F. and B. de Massy. 2007. Regulating double-stranded DNA break repair towards crossover or non-crossover during mammalian meiosis. *Chromosome Res* **15:** 565-577.

Baudat, F. and A. Nicolas. 1997. Clustering of meiotic double-strand breaks on yeast chromosome III. *Proc Natl Acad Sci U S A* **94:** 5213-5218.

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

Bednarek, P.T., P. Masojc, R. Lewandowska, and B. Myskow. 2003. Saturating rye genetic map with amplified fragment length polymorphism (AFLP) and random amplified polymorphic DNA (RAPD) markers. *J Appl Genet* **44:** 21-33.

Bennett, M.D. and I.J. Leitch. 2005. Nuclear DNA amounts in angiosperms: progress, problems and prospects. *Ann Bot (Lond)* **95:** 45-90.

Berchowitz, L.E., K.E. Francis, A.L. Bey, and G.P. Copenhaver. 2007. The role of AtMUS81 in interference-insensitive crossovers in A. thaliana. *PLoS Genet* **3:** e132.

Beye, M., I. Gattermeier, M. Hasselmann, T. Gempe, M. Schioett, J.F. Baines, D. Schlipalius, F. Mougel, C. Emore, O. Rueppell, A. Sirvio, E. Guzman-Novoa, G. Hunt, M. Solignac, and R.E. Page, Jr. 2006. Exceptionally high levels of recombination across the honey bee genome. *Genome Res* **16:** 1339-1344.

Billotte, N., N. Marseillac, A.M. Risterucci, B. Adon, P. Brottier, F.C. Baurens, R. Singh, A. Herran, H. Asmady, C. Billot, P. Amblard, T. Durand-Gasselin, B. Courtois, D. Asmono, S.C. Cheah, W. Rohde, E. Ritter, and A. Charrier. 2005. Microsatellite-based high density linkage map in oil palm (Elaeis guineensis Jacq.). *Theor Appl Genet* **110:** 754-765.

Bindler, G., R. van der Hoeven, I. Gunduz, J. Plieske, M. Ganal, L. Rossi, F. Gadani, and P. Donini. 2007. A microsatellite marker based linkage map of tobacco. *Theor Appl Genet* **114:** 341-349.

Bird, A.P. 1986. CpG-rich islands and the function of DNA methylation. *Nature* **321:** 209-213.

Bishop, D.K. and D. Zickler. 2004. Early decision; meiotic crossover interference prior to stable strand exchange and synapsis. *Cell* **117:** 9-15.

Bisht, K.K., S. Arora, S. Ahmed, and J. Singh. 2008. Role of heterochromatin in suppressing subtelomeric recombination in fission yeast. *Yeast* **25:** 537-548.

Blitzblau, H.G., G.W. Bell, J. Rodriguez, S.P. Bell, and A. Hochwagen. 2007. Mapping of meiotic single-stranded DNA reveals double-stranded-break hotspots near centromeres and telomeres. *Curr Biol* **17:** 2003-2012.

Blumental-Perry, A., D. Zenvirth, S. Klein, I. Onn, and G. Simchen. 2000. DNA motif associated with meiotic double-strand break regions in Saccharomyces cerevisiae. *EMBO Rep* **1:** 232-238.

Boden, S.A., P. Langridge, G. Spangenberg, and J.A. Able. 2008. TaASY1 promotes homologous chromosome interactions and is affected by deletion of Ph1. *Plant J.*

Boden, S.A., N. Shadiac, E.J. Tucker, P. Langridge, and J.A. Able. 2007. Expression and functional analysis of TaASY1 during meiosis of bread wheat (Triticum aestivum). *BMC Mol Biol* **8:** 65.

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

Borde, V., A.S. Goldman, and M. Lichten. 2000. Direct coupling between meiotic DNA replication and recombination initiation. *Science* **290:** 806-809.

Borde, V., N. Robine, W. Lin, S. Bonfils, V. Geli, and A. Nicolas. 2008. Histone H3 lysine 4 trimethylation marks meiotic recombination initiation sites. *Embo J.*

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. *Mol Cell Biol* **19:** 4832-4842.

Borner, G.V. 2006. Balancing the checks: surveillance of chromosomal exchange during meiosis. *Biochem Soc Trans* **34:** 554-556.

Borner, G.V., N. Kleckner, and N. Hunter. 2004. Crossover/noncrossover differentiation, synaptonemal complex formation, and regulatory surveillance at the leptotene/zygotene transition of meiosis. *Cell* **117:** 29-45.

Borts, R.H. and J.E. Haber. 1987. Meiotic recombination in yeast: alteration by multiple heterozygosities. *Science* **237:** 1459-1465.

Bovill, W.D., P. Deveshwar, S. Kapoor, and J.A. Able. 2008. Whole genome approaches to identify early meiotic gene candidates in cereals. *Funct Integr Genomics*.

Broman, K.W., L.B. Rowe, G.A. Churchill, and K. Paigen. 2002. Crossover interference in the mouse. *Genetics* **160:** 1123-1131.

Broman, K.W. and J.L. Weber. 2000. Characterization of human crossover interference. *Am J Hum Genet* **66:** 1911-1926.

Brondani, R.P., E.R. Williams, C. Brondani, and D. Grattapaglia. 2006. A microsatellite-based consensus linkage map for species of Eucalyptus and a novel set of 230 microsatellite markers for the genus. *BMC Plant Biol* **6:** 20.

Buard, J. and B. de Massy. 2007. Playing hide and seek with mammalian meiotic crossover hotspots. *Trends Genet* **23:** 301-309.

Buhler, C., V. Borde, and M. Lichten. 2007. Mapping meiotic single-strand DNA reveals a new landscape of DNA double-strand breaks in Saccharomyces cerevisiae. *PLoS Biol* **5:** e324.

Busso, C.S., C.J. Liu, C.T. Hash, J.R. Witcombe, K.M. Devos, J.M.J. Wet, and M.D. Gale. 1995. Analysis of recombination rate in female and male gametogenesis in pearl millet (Pennisetum glaucum) using RFLP markers. *Theor Appl Genet* **90:** 242-246.

Calderon, P.L. and M.I. Pigozzi. 2006. MLH1-focus mapping in birds shows equal recombination between sexes and diversity of crossover patterns. *Chromosome Res* **14:** 605-612.

Carlton, P.M. 2008. Three-dimensional structured illumination microscopy and its application to chromosome structure. *Chromosome Res* **16:** 351-365.

Carrington, M. and M. Cullen. 2004. Justified chauvinism: advances in defining meiotic recombination through sperm typing. *Trends Genet* **20:** 196-205.

Caryl, A.P., S.J. Armstrong, G.H. Jones, and F.C. Franklin. 2000. A homologue of the yeast HOP1 gene is inactivated in the Arabidopsis meiotic mutant asy1. *Chromosoma* **109:** 62-71.

Cerenak, A., Z. Satovic, and B. Javornik. 2006. Genetic mapping of hop (Humulus lupulus L.) applied to the detection of QTLs for alpha-acid content. *Genome* **49:** 485-494.

Chen, C., Q. Yu, S. Hou, Y. Li, M. Eustice, R.L. Skelton, O. Veatch, R.E. Herdes, L. Diebold, J. Saw, Y. Feng, W. Qian, L. Bynum, L. Wang, P.H. Moore, R.E. Paull, M. Alam, and R. Ming. 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, J.F., F. Lu, S.S. Chen, and S.H. Tao. 2006. Significant positive correlation between the recombination rate and GC content in the human pseudoautosomal region. *Genome* **49:** 413-419.

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

Chen, S.Y., T. Tsubouchi, B. Rockmill, J.S. Sandler, D.R. Richards, G. Vader, A. Hochwagen, G.S. Roeder, and J.C. Fung. 2008. Global Analysis of the Meiotic Crossover Landscape. *Dev Cell*.

Chen, W. and S. Jinks-Robertson. 1999. The role of the mismatch repair machinery in regulating mitotic and meiotic recombination between diverged sequences in yeast. *Genetics* **151:** 1299-1313.

Cheng, Z., G.G. Presting, C.R. Buell, R.A. Wing, and J. Jiang. 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.

Chetelat, R.T., V. Meglic, and P. Cisneros. 2000. A genetic map of tomato based on BC(1) Lycopersicon esculentum x Solanum lycopersicoides reveals overall synteny but suppressed recombination between these homeologous genomes. *Genetics* **154:** 857-867.

Chicheportiche, A., J. Bernardino-Sgherri, B. de Massy, and B. Dutrillaux. 2007. Characterization of Spo11-dependent and independent phospho-H2AX foci during meiotic prophase I in the male mouse. *J Cell Sci* **120:** 1733-1742.

Choi, H.K., D. Kim, T. Uhm, E. Limpens, H. Lim, J.H. Mun, P. Kalo, R.V. Penmetsa, A. Seres, O. Kulikova, B.A. Roe, T. Bisseling, G.B. Kiss, and D.R. Cook. 2004. A sequence-based genetic map of Medicago truncatula and comparison of marker colinearity with M. sativa. *Genetics* **166:** 1463-1502.

Choi, S.R., G.R. Teakle, P. Plaha, J.H. Kim, C.J. Allender, E. Beynon, Z.Y. Piao, P. Soengas, T.H. Han, G.J. King, G.C. Barker, P. Hand, D.J. Lydiate, J. Batley, D. Edwards, D.H. Koo, J.W. Bang, B.S. Park, and Y.P. Lim. 2007. The reference genetic linkage map for the multinational Brassica rapa genome sequencing project. *Theor Appl Genet* **115:** 777-792.

Codina-Pascual, M., M. Campillo, J. Kraus, M.R. Speicher, J. Egozcue, J. Navarro, and J. Benet. 2006. Crossover frequency and synaptonemal complex length: their variability and effects on human male meiosis. *Mol Hum Reprod* **12:** 123-133.

Colas, I., P. Shaw, P. Prieto, M. Wanous, W. Spielmeyer, R. Mago, and G. Moore. 2008. Effective chromosome pairing requires chromatin remodeling at the onset of meiosis. *Proc Natl Acad Sci U S A* **105:** 6075-6080.

Comeron, J.M. and M. Kreitman. 2000. The correlation between intron length and recombination in drosophila. Dynamic equilibrium between mutational and selective forces. *Genetics* **156:** 1175-1190.

Coop, G., X. Wen, C. Ober, J.K. Pritchard and M. Przeworski. 2008. High-resolution mapping of crossovers reveals extensive variations in fine-scale recombination patterns among humans. **319**: 1395-1398.

Copenhaver, G.P., E.A. Housworth, and F.W. Stahl. 2002. Crossover interference in Arabidopsis. *Genetics* **160:** 1631-1639.

Corredor, E., A.J. Lukaszewski, P. Pachon, D.C. Allen, and T. Naranjo. 2007. Terminal regions of wheat chromosomes select their pairing partners in meiosis. *Genetics* **177:** 699-706.

Costantini, M., O. Clay, F. Auletta, and G. Bernardi. 2006. An isochore map of human chromosomes. *Genome Res* **16:** 536-541.

Couteau, F., F. Belzile, C. Horlow, O. Grandjean, D. Vezon and Droutriaux MP. 1999. Random chromosome segregation without meiotic arrest in both male and female meiocytes of a dmc1 mutant of Arabidopsis. *Plant cell* **11**:1623-1634.

Crismani, W., U. Baumann, T. Sutton, N. Shirley, T. Webster, G. Spangenberg, P. Langridge, and J.A. Able. 2006. Microarray expression analysis of meiosis and microsporogenesis in hexaploid bread wheat. *BMC Genomics* **7:** 267.

Cromie, G.A., R.W. Hyppa, H.P. Cam, J.A. Farah, S.I. Grewal, and G.R. Smith. 2007. A discrete class of intergenic DNA dictates meiotic DNA break hotspots in fission yeast. *PLoS Genet* **3:** e141.

Cromie, G.A., R.W. Hyppa, and G.R. Smith. 2008. The fission yeast BLM homolog Rqh1 promotes meiotic recombination. *Genetics* **179:** 1157-1167.

Cromie, G.A., R.W. Hyppa, A.F. Taylor, K. Zakharyevich, N. Hunter, and G.R. Smith. 2006. Single Holliday junctions are intermediates of meiotic recombination. *Cell* **127:** 1167-1178.

Cromie, G.A. and G.R. Smith. 2007. Branching out: meiotic recombination and its regulation. *Trends Cell Biol* **17:** 448-455.

Dawe, R.K. 1998. Meiotic Chromosome Organization and Segregation in Plants. *Annu Rev Plant Physiol Plant Mol Biol* **49:** 371-395.

de Boer, E. and C. Heyting. 2006. The diverse roles of transverse filaments of synaptonemal complexes in meiosis. *Chromosoma* **115:** 220-234.

de Boer, E., P. Stam, A.J. Dietrich, A. Pastink, and C. Heyting. 2006. Two levels of interference in mouse meiotic recombination. *Proc Natl Acad Sci U S A* **103:** 9607-9612.

de Bustos, A., R. Perez, and N. Jouve. 2007. Characterization of the gene Mre11 and evidence of silencing after polyploidization in Triticum. *Theor Appl Genet* **114:** 985-999.

de los Santos, T., N. Hunter, C. Lee, B. Larkin, J. Loidl, and N.M. Hollingsworth. 2003. The Mus81/Mms4 endonuclease acts independently of double-Holliday junction resolution to promote a distinct subset of crossovers during meiosis in budding yeast. *Genetics* **164:** 81-94.

de Massy, B. 2003. Distribution of meiotic recombination sites. *Trends Genet* **19:** 514-522.

De Muyt, A., D. Vezon, G. Gendrot, J.L. Gallois, R. Stevens, and M. Grelon. 2007. AtPRD1 is required for meiotic double strand break formation in Arabidopsis thaliana. *Embo J* **26:** 4126-4137.

Di Gaspero, G., G. Cipriani, A.F. Adam-Blondon, and R. Testolin. 2007. Linkage maps of grapevine displaying the chromosomal locations of 420 microsatellite markers and 82 markers for R-gene candidates. *Theor Appl Genet* **114:** 1249-1263.

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

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

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

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

Dumas, D. and J. Britton-Davidian. 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.

Edgerton, M.D. 2009. Increasing crop productivity to meet global needs for feed, food, and fuel. *Plant Physiol* **149:** 7-13.

Efstratiadis, A. 1994. Parental imprinting of autosomal mammalian genes. *Curr Opin Genet Dev* **4:** 265-280.

Endo, T.R. and B.S. Gill. 1996. The deletion Stocks of Common Wheat. *Journal of Heredity* **87:** 295-307.

Erayman, M., D. Sandhu, D. Sidhu, M. Dilbirligi, P.S. Baenziger, and K.S. Gill. 2004. Demarcating the gene-rich regions of the wheat genome. *Nucleic Acids Res* **32:** 3546-3565.

Esch, E., J.M. Szymaniak, H. Yates, W.P. Pawlowski, and E.S. Buckler. 2007. Using crossover breakpoints in recombinant inbred lines to identify quantitative trait loci controlling the global recombination frequency. *Genetics* **177:** 1851-1858.

Esch, E. and E. Weber. 2002. Investigation of crossover interference in barley (Hordeum vulgare L.) using the coefficient of coincidence. *Theor Appl Genet* **104:** 786-796.

Falque, M., L. Decousset, D. Dervins, A.M. Jacob, J. Joets, J.P. Martinant, X. Raffoux, N. Ribiere, C. Ridel, D. Samson, A. Charcosset, and A. Murigneux. 2005. Linkage mapping of 1454 new maize candidate gene Loci. *Genetics* **170:** 1957-1966.

Falque, M., R. Mercier, C. Mezard, D. de Vienne, and O.C. Martin. 2007. Patterns of recombination and MLH1 foci density along mouse chromosomes: modeling effects of interference and obligate chiasma. *Genetics* **176:** 1453-1467.

Fazio, G., J.E. Staub, and M.R. Stevens. 2003. Genetic mapping and QTL analysis of horticultural traits in cucumber (Cucumis sativus L.) using recombinant inbred lines. *Theor Appl Genet* **107:** 864-874.

Feldman, M. 1993. Cytogenetic activity and mode of action of the pairing homoeologous (Ph1) gene of wheat. *Crop Science* **33:** 894-897.

Fernandez-Calvin, B., E. Benavente, and J. Orellana. 1995. Meiotic pairing in wheat-rye derivatives detected by genomic in situ hybridization and C-banding--A comparative analysis. *Chromosoma* **103:** 554-558.

Feuillet, C., P. Langridge, and R. Waugh. 2008. Cereal breeding takes a walk on the wild side. *Trends Genet* **24:** 24-32.

Francis, K.E., S.Y. Lam, B.D. Harrison, A.L. Bey, L.E. Berchowitz, and G.P. Copenhaver. 2007. Pollen tetrad-based visual assay for meiotic recombination in Arabidopsis. *Proc Natl Acad Sci U S A* **104:** 3913-3918.

Froenicke, L., L.K. Anderson, J. Wienberg, and T. Ashley. 2002. Male mouse recombination maps for each autosome identified by chromosome painting. *Am J Hum Genet* **71:** 1353-1368.

Fu, T.K. and E.R. Sears. 1973. The Relationship between Chiasmata and Crossing over in TRITICUM AESTIVUM. *Genetics* **75:** 231-246.

Fung, J.C., B. Rockmill, M. Odell, and G.S. Roeder. 2004. Imposition of crossover interference through the nonrandom distribution of synapsis initiation complexes. *Cell* **116:** 795-802.

Galtier, N., G. Piganeau, D. Mouchiroud, and L. Duret. 2001. GC-content evolution in mammalian genomes: the biased gene conversion hypothesis. *Genetics* **159:** 907-911.

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

Gardiner-Garden, M. and M. Frommer. 1987. CpG islands in vertebrate genomes. *J Mol Biol* **196:** 261-282.

Gaut, B.S., S.I. Wright, C. Rizzon, J. Dvorak, and L.K. Anderson. 2007. Recombination: an underappreciated factor in the evolution of plant genomes. *Nat Rev Genet* **8:** 77-84.

Gerton, J.L., J. DeRisi, R. Shroff, M. Lichten, P.O. Brown, and T.D. Petes. 2000. Inaugural article: global mapping of meiotic recombination hotspots and coldspots in the yeast Saccharomyces cerevisiae. *Proc Natl Acad Sci U S A* **97:** 11383-11390.

Gerton, J.L. and R.S. Hawley. 2005. Homologous chromosome interactions in meiosis: diversity amidst conservation. *Nat Rev Genet* **6:** 477-487.

Gill, B.S., B. Friebe and T.R. Endo. 1991. Standard karyotype and nomenclature system for description of chromosome bands and structural aberrations in wheat (*Triticum aestivum*). *Genome* **34**: 830-839.

Gill, K.S., B.S. Gill, T.R. Endo, and E.V. Boyko. 1996a. Identification and high-density mapping of gene-rich regions in chromosome group 5 of wheat. *Genetics* **143:** 1001-1012.

Gill, K.S., B.S. Gill, T.R. Endo, and T. Taylor. 1996b. Identification and high-density mapping of gene-rich regions in chromosome group 1 of wheat. *Genetics* **144:** 1883-1891.

Goedecke, W., M. Eijpe, H.H. Offenberg, M. van Aalderen, and C. Heyting. 1999. Mre11 and Ku70 interact in somatic cells, but are differentially expressed in early meiosis. *Nat Genet* **23:** 194-198.

Grelon, M., D. Vezon, G. Gendrot, and G. Pelletier. 2001. AtSPO11-1 is necessary for efficient meiotic recombination in plants. *Embo J* **20:** 589-600.

Grewal, S.I. and S. Jia. 2007. Heterochromatin revisited. *Nat Rev Genet* **8:** 35-46.

Griffiths, S., R. Sharp, T.N. Foote, I. Bertin, M. Wanous, S. Reader, I. Colas, and G. Moore. 2006. Molecular characterization of Ph1 as a major chromosome pairing locus in polyploid wheat. *Nature* **439:** 749-752.

Grisi, M.C., M.W. Blair, P. Gepts, C. Brondani, P.A. Pereira, and R.P. Brondani. 2007. Genetic mapping of a new set of microsatellite markers in a reference common bean (Phaseolus vulgaris) population BAT93 x Jalo EEP558. *Genet Mol Res* **6:** 691-706.

Guo, W., C. Cai, C. Wang, Z. Han, X. Song, K. Wang, X. Niu, C. Wang, K. Lu, B. Shi, and T. Zhang. 2007. A microsatellite-based, gene-rich linkage map reveals genome structure, function and evolution in Gossypium. *Genetics* **176:** 527-541.

Gupta, P.K., R.R. Mir, A. Mohan, and J. Kumar. 2008. Wheat genomics: present status and future prospects. *Int J Plant Genomics* **2008:** 896451.

Guzy-Wrobelska, J., A. Labocha-Pawlowska, M. Kwasniewski, and I. Szarejko. 2007. Different recombination frequencies in wheat doubled haploid populations obtained through maize pollination and anther culture. *Euphytica* **156:** 173-183.

Haber, J.E. 1997. A super new twist on the initiation of meiotic recombination. *Cell* **89:** 163-166.

Haber, J.E. 1998. Mating-type gene switching in Saccharomyces cerevisiae. *Annu Rev Genet* **32:** 561-599.

Hamant, O., H. Ma, and W.Z. Cande. 2006. Genetics of meiotic prophase I in plants. *Annu Rev Plant Biol* **57:** 267-302.

Hammarlund, M., M.W. Davis, H. Nguyen, D. Dayton, and E.M. Jorgensen. 2005. Heterozygous insertions alter crossover distribution but allow crossover interference in Caenorhabditis elegans. *Genetics* **171:** 1047-1056.

Hamwieh, A., S.M. Udupa, W. Choumane, A. Sarker, F. Dreyer, C. Jung, and M. Baum. 2005. A genetic linkage map of Lens sp. based on microsatellite and AFLP markers and the localization of fusarium vascular wilt resistance. *Theor Appl Genet* **110:** 669-677.

Han, L., B. Su, W.H. Li, and Z. Zhao. 2008. CpG island density and its correlations with genomic features in mammalian genomes. *Genome Biol* **9:** R79.

Haring, S.J., L.J. Lautner, J.M. Comeron, and R.E. Malone. 2004. A test of the CoHR motif associated with meiotic double-strand breaks in Saccharomyces cerevisiae. *EMBO Rep* **5:** 41-46.

Hartung, F., K.J. Angelis, A. Meister, I. Schubert, M. Melzer, and H. Puchta. 2002. An archaebacterial topoisomerase homolog not present in other eukaryotes is indispensable for cell proliferation of plants. *Curr Biol* **12:** 1787-1791.

Hartung, F., R. Wurz-Wildersinn, J. Fuchs, I. Schubert, S. Suer, and H. Puchta. 2007. The catalytically active tyrosine residues of both SPO11-1 and SPO11-2 are required for meiotic double-strand break induction in Arabidopsis. *Plant Cell* **19:** 3090-3099.

Hassold, T., H. Hall, and P. Hunt. 2007. The origin of human aneuploidy: where we have been, where we are going. *Hum Mol Genet* **16 Spec No. 2:** R203-208.

Haudry, A., A. Cenci, C. Guilhaumon, E. Paux, S. Poirier, S. Santoni, J. David, and S. Glemin. 2008. Mating system and recombination affect molecular evolution in four Triticeae species. *Genet Res* **90:** 97-109.

Hernandez-Hernandez, A., G.H. Vazquez-Nin, O.M. Echeverria, and F. Recillas-Targa. 2008. Chromatin structure contribution to the synaptonemal complex formation. *Cell. Mol. Life Sci.*

Heyer, W.D., X. Li, M. Rolfsmeier, and X.P. Zhang. 2006. Rad54: the Swiss Army knife of homologous recombination? *Nucleic Acids Res* **34:** 4115-4125.

Higgins, J.D., E.F. Buckling, F.C. Franklin, and G.H. Jones. 2008. Expression and functional analysis of AtMUS81 in Arabidopsis meiosis reveals a role in the second pathway of crossing-over. *Plant J* **54:** 152-162.

Higgins, J.D., E. Sanchez-Moran, S.J. Armstrong, G.H. Jones, and F.C. Franklin. 2005. The Arabidopsis synaptonemal complex protein ZYP1 is required for chromosome synapsis and normal fidelity of crossing over. *Genes Dev* **19:** 2488-2500.

Hillers, K.J. and A.M. Villeneuve. 2003. Chromosome-wide control of meiotic crossing over in C. elegans. *Curr Biol* **13:** 1641-1647.

Hocheberg, Y. 1988. A sharper Bonferonni procedure for multiple tests of significance. *Biometrika* **75:** 800-803.

Hochwagen, A. and A. Amon. 2006. Checking your breaks: surveillance mechanisms of meiotic recombination. *Curr Biol* **16:** R217-228.

Holloway, J.K., J. Booth, W. Edelmann, C.H. McGowan, and P.E. Cohen. 2008. MUS81 generates a subset of MLH1-MLH3-independent crossovers in mammalian meiosis. *PLoS Genet* **4:** e1000186.

Hunt, P.A. 2006. Meiosis in mammals: recombination, non-disjunction and the environment. *Biochem Soc Trans* **34:** 574-577.

Iovene, M., S.M. Wielgus, P.W. Simon, C.R. Buell, and J. Jiang. 2008. Chromatin structure and physical mapping of chromosome 6 of potato and comparative analyses with tomato. *Genetics* **180:** 1307-1317.

Jackson, N., E. Sanchez-Moran, E. Buckling, S.J. Armstrong, G.H. Jones, and F.C. Franklin. 2006. Reduced meiotic crossovers and delayed prophase I progression in AtMLH3-deficient Arabidopsis. *Embo J* **25**.

Jayathilaka, K., S.D. Sheridan, T.D. Bold, K. Bochenska, H.L. Logan, R.R. Weichselbaum, D.K. Bishop, and P.P. Connell. 2008. A chemical compound that stimulates the human homologous recombination protein RAD51. *Proc Natl Acad Sci U S A*.

Jeffreys, A.J., J.K. Holloway, L. Kauppi, C.A. May, R. Neumann, M.T. Slingsby, and A.J. Webb. 2003. Meiotic recombination hot spots and human DNA diversity. *Phil Trans R Soc Lond* **359:** 141-152.

Jeffreys, A.J. and R. Neumann. 2002. Reciprocal crossover asymmetry and meiotic drive in a human recombination hot spot. *Nat Genet* **31:** 267-271.

Jeffreys, A.J. and R. Neumann. 2005. Factors influencing recombination frequency and distribution in a human meiotic crossover hotspot. *Hum Mol Genet* **14:** 2277-2287.

Jenkins, G., D. Phillips, E.I. Mikhailova, L. Timofejeva, and R.N. Jones. 2008. Meiotic genes and proteins in cereals. *Cytogenet Genome Res* **120:** 291-301.

Jensen-Seaman, M.I., T.S. Furey, B.A. Payseur, Y. Lu, K.M. Roskin, C.F. Chen, M.A. Thomas, D. Haussler, and H.J. Jacob. 2004. Comparative recombination rates in the rat, mouse, and human genomes. *Genome Res* **14:** 528-538.

Jessop, L. and M. Lichten. 2008. Mus81/Mms4 endonuclease and Sgs1 helicase collaborate to ensure proper recombination intermediate metabolism during meiosis. *Mol Cell* **31:** 313-323.

Jessop, L., B. Rockmill, G.S. Roeder, and M. Lichten. 2006. Meiotic chromosome synapsis-promoting proteins antagonize the anti-crossover activity of sgs1. *PLoS Genet* **2:** e155.

Ji, Y. and R.T. Chetelat. 2007. GISH analysis of meiotic chromosome pairing in Solanum lycopersicoides introgression lines of cultivated tomato. *Genome* **50:** 825-833.

Jiang, J. and B.S. Gill. 2006. Current status and the future of fluorescence in situ hybridization (FISH) in plant genome research. *Genome* **49:** 1057-1068.

Jiang, P., H. Wu, J. Wei, F. Sang, X. Sun, and Z. Lu. 2007. RF-DYMHC: detecting the yeast meiotic recombination hotspots and coldspots by random forest model using gapped dinucleotide composition features. *Nucleic Acids Res* **35:** W47-51.

Jones, E., K. Rybka, and J. Lukaszewski. 2002. The effect of a deficiency and a deletion on recombination in chromosome 1BL in wheat. *Theor Appl Genet* **104:** 1204-1208.

Jones, G.H. and F.C. Franklin. 2006. Meiotic crossing-over: obligation and interference. *Cell* **126:** 246-248.

Jordan, P. 2006. Initiation of homologous chromosome pairing during meiosis. *Biochem Soc Trans* **34:** 545-549.

Kaback, D.B. 1996. Chromosome-size dependent control of meiotic recombination in humans. *Nat Genet* **13:** 20-21.

Kaback, D.B., V. Guacci, D. Barber, and J.W. Mahon. 1992. Chromosome size-dependent control of meiotic recombination. *Science* **256:** 228-232.

Kao, F.I., Y.Y. Cheng, T.Y. Chow, H.H. Chen, S.M. Liu, C.H. Cheng, and M.C. Chung. 2006. An integrated map of Oryza sativa L. chromosome 5. *Theor Appl Genet* **112:** 891-902.

Keeney, S. 2001. Mechanism and control of meiotic recombination initiation. *Curr Top Dev Biol* **52:** 1-53.

Keeney, S. and M.J. Neale. 2006. Initiation of meiotic recombination by formation of DNA double-strand breaks: mechanism and regulation. *Biochem Soc Trans* **34:** 523-525.

Kelleher, C.T., R. Chiu, H. Shin, I.E. Bosdet, M.I. Krzywinski, C.D. Fjell, J. Wilkin, T. Yin, S.P. DiFazio, J. Ali, J.K. Asano, S. Chan, A. Cloutier, N. Girn, S. Leach, D. Lee, C.A. Mathewson, T. Olson, K. O'Connor, A.L. Prabhu, D.E. Smailus, J.M. Stott, M. Tsai, N.H. Wye, G.S. Yang, J. Zhuang, R.A. Holt, N.H. Putnam, J. Vrebalov, J.J. Giovannoni, J. Grimwood, J. Schmutz, D. Rokhsar, S.J. Jones, M.A. Marra, G.A. Tuskan, J. Bohlmann, B.E. Ellis, K. Ritland, C.J. Douglas, and J.E. Schein. 2007. A physical map of the highly heterozygous Populus genome: integration with the genome sequence and genetic map and analysis of haplotype variation. *Plant J* **50:** 1063-1078.

Kim, J.S., M.N. Islam-Faridi, P.E. Klein, D.M. Stelly, H.J. Price, R.R. Klein, and J.E. Mullet. 2005. Comprehensive molecular cytogenetic analysis of sorghum genome architecture: distribution of euchromatin, heterochromatin, genes and recombination in comparison to rice. *Genetics* **171:** 1963-1976.

King, J., I.P. Armstead, I.S. Donnison, H.M. Thomas, R.N. Jones, M.J. Kearsey, L.A. Roberts, A. Thomas, W.G. Morgan, and I.P. King. 2002. Physical and genetic mapping in the grasses Lolium perenne and Festuca pratensis. *Genetics* **161:** 315-324.

King, J., I.P. Armstead, S.I. Donnison, L.A. Roberts, J.A. Harper, K. Skot, K. Elborough, and I.P. King. 2007. Comparative analyses between lolium/festuca introgression lines and rice reveal the major fraction of functionally annotated gene models is located in recombination-poor/very recombination-poor regions of the genome. *Genetics* **177:** 597-606.

Kirik, A., A. Pecinka, E. Wendeler, and B. Reiss. 2006. The chromatin assembly factor subunit FASCIATA1 is involved in homologous recombination in plants. *Plant Cell* **18:** 2431-2442.

Kleckner, N., A. Storlazzi, and D. Zickler. 2003. Coordinate variation in meiotic pachytene SC length and total crossover/chiasma frequency under conditions of constant DNA length. *Trends Genet* **19:** 623-628.

Kleckner, N., D. Zickler, G.H. Jones, J. Dekker, R. Padmore, J. Henle, and J. Hutchinson. 2004. A mechanical basis for chromosome function. *Proc Natl Acad Sci U S A* **101:** 12592-12597.

Kniewel, R. and S. Keeney. 2009. Histone methylation sets the stage for meiotic DNA breaks. *Embo J* **28:** 81-83.

Kochakpour, N. and P.B. Moens. 2008. Sex-specific crossover patterns in Zebrafish (Danio rerio). *Heredity* **100:** 489-495.

Koehler, K.E., J.P. Cherry, A. Lynn, P.A. Hunt, and T.J. Hassold. 2002. Genetic control of mammalian meiotic recombination. I. Variation in exchange frequencies among males from inbred mouse strains. *Genetics* **162:** 297-306.

Kon, N., M.D. Krawchuk, B.G. Warren, G.R. Smith, and W.P. Wahls. 1997. Transcription factor Mts1/Mts2 (Atf1/Pcr1, Gad7/Pcr1) activates the M26 meiotic recombination hotspot in Schizosaccharomyces pombe. *Proc Natl Acad Sci U S A* **94:** 13765-13770.

Kong, A., D.F. Gudbjartsson, J. Sainz, G.M. Jonsdottir, S.A. Gudjonsson, B. Richardsson, S. Sigurdardottir, J. Barnard, B. Hallbeck, G. Masson, A. Shlien, S.T. Palsson, M.L. Frigge, T.E. Thorgeirsson, J.R. Gulcher, and K. Stefansson. 2002. A high-resolution recombination map of the human genome. *Nat Genet* **31:** 241-247.

Koo, D.H., S.H. Jo, J.W. Bang, H.M. Park, S. Lee, and D. Choi. 2008. Integration of cytogenetic and genetic linkage maps unveils the physical architecture of tomato chromosome 2. *Genetics* **179:** 1211-1220.

Korzun, V., A. Börner, R. Siebert, S. Malyshev, M. Hilpert, R. Kunze, and H. Puchta. 1999. Chromosomal location and genetic mapping of the mismatch repair gene homologs *MSH2*, *MSH3*, and *MSH6* in rye and wheat. *Genome* **42:** 1255-1257.

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

Kulikova, O., G. Gualtieri, R. Geurts, D.J. Kim, D. Cook, T. Huguet, J.H. de Jong, P.F. Fransz, and T. Bisseling. 2001. Integration of the FISH pachytene and genetic maps of Medicago truncatula. *Plant J* **27:** 49-58.

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

La Rota, M. and M.E. Sorrells. 2004. Comparative DNA sequence analysis of mapped wheat ESTs reveals the complexity of genome relationships between rice and wheat. *Funct Integr Genomics* **4:** 34-46.

Lamb, N.E., S.L. Sherman, and T.J. Hassold. 2005. Effect of meiotic recombination on the production of aneuploid gametes in humans. *Cytogenet Genome Res* **111:** 250-255.

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

Laperche, A., M. Brancourt-Hulmel, E. Heumez, O. Gardet, and J. Le Gouis. 2006. Estimation of genetic parameters of a DH wheat population grown at different N stress levels characterized by probe genotypes. *Theor Appl Genet* **112:** 797-807.

Laurent, V., P. Devaux, T. Thiel, F. Viard, S. Mielordt, P. Touzet, and M.C. Quillet. 2007. Comparative effectiveness of sugar beet microsatellite markers isolated from genomic libraries and GenBank ESTs to map the sugar beet genome. *Theor Appl Genet* **115:** 793-805.

Laurie, D.A. and M.D. Bennett. 1988. The production of haploid wheat plants from wheat x maize crosses *Theor Appl Genet* **3:** 393-397.

Lebrun, P., L. Baudouin, R. Bourdeix, J.L. Konan, J.H. Barker, C. Aldam, A. Herran, and E. Ritter. 2001. Construction of a linkage map of the Rennell Island Tall coconut type (Cocos nucifera L.) and QTL analysis for yield characters. *Genome* **44:** 962-970.

Lenormand, T. and J. Dutheil. 2005a. Recombination difference between sexes: a role for haploid selection. *PLoS Biol* **3**.

Lenormand, T. and J. Dutheil. 2005b. Recombination difference between sexes: a role for haploid selection. *PLoS Biol* **3:** e63.

Lercher, M.J. and L.D. Hurst. 2003. Imprinted chromosomal regions of the human genome have unusually high recombination rates. *Genetics* **165:** 1629-1632.

Lhuissier, F.G., H.H. Offenberg, P.E. Wittich, N.O. Vischer, and C. Heyting. 2007. The mismatch repair protein MLH1 marks a subset of strongly interfering crossovers in tomato. *Plant Cell* **19:** 862-876.

Li, J., L.C. Harper, I. Golubovskaya, C.R. Wang, D. Weber, R.B. Meeley, J. McElver, B. Bowen, W.Z. Cande, and P.S. Schnable. 2007. Functional analysis of maize RAD51 in meiosis and double-strand break repair. *Genetics* **176:** 1469-1482.

Li, W., C. Chen, U. Markmann-Mulisch, L. Timofejeva, E. Schmelzer, H. Ma, and B. Reiss. 2004a. The Arabidopsis AtRAD51 gene is dispensable for vegetative development but required for meiosis. *Proc Natl Acad Sci U S A* **101:** 10596-10601.

Li, W., P. Zhang, J.P. Fellers, B. Friebe, and B.S. Gill. 2004b. Sequence composition, organization, and evolution of the core Triticeae genome. *Plant J* **40:** 500-511.

Lian, J., Y. Yin, M. Oliver-Bonet, T. Liehr, E. Ko, P. Turek, F. Sun, and R.H. Martin. 2008. Variation in crossover interference levels on individual chromosomes from human males. *Hum Mol Genet* **17:** 2583-2594.

Lichten, M. 2001a. Describing recombination. *Trends Genet* **17:** 135.

Lichten, M. 2001b. Meiotic recombination: breaking the genome to save it. *Curr Biol* **11:** R253-256.

Liebhard, R., B. Koller, L. Gianfranceschi, and C. Gessler. 2003. Creating a saturated reference map for the apple (Malus x domestica Borkh.) genome. *Theor Appl Genet* **106:** 1497-1508.

Lim, J.G., R.R. Stine, and J.L. Yanowitz. 2008. Domain-specific regulation of recombination in C. elegans in response to temperature, age, and sex. *Genetics*.

Liu, Y. and S.C. West. 2004. Happy Hollidays: 40th anniversary of the Holliday junction. *Nat Rev Mol Cell Biol* **5:** 937-944.

Llaurens, V., S. Billiard, J.B. Leducq, V. Castric, E.K. Klein, and X. Vekemans. 2008. Does frequency-dependent selection with complex dominance interactions accurately predict allelic frequencies at the self-incompatibility locus in Arabidopsis halleri? *Evolution* **62:** 2545-2557.

Lloyd, A.H., A.S. Milligan, P. Langridge, and J.A. Able. 2007. TaMSH7: a cereal mismatch repair gene that affects fertility in transgenic barley (Hordeum vulgare L.). *BMC Plant Biol* **7:** 67.

Lopez, E., M. Pradillo, C. Romero, J.L. Santos, and N. Cunado. 2008. Pairing and synapsis in wild type Arabidopsis thaliana. *Chromosome Res* **16:** 701-708.

Lorch, P.D. 2005. Sex differences in recombination and mapping adaptations. *Genetica* **123:** 39-47.

Loridon, K., K. McPhee, J. Morin, P. Dubreuil, M.L. Pilet-Nayel, G. Aubert, C. Rameau, A. Baranger, C. Coyne, I. Lejeune-Henaut, and J. Burstin. 2005. Microsatellite marker polymorphism and mapping in pea (Pisum sativum L.). *Theor Appl Genet* **111:** 1022-1031.

Lukaszewski, A.J. 2008. Unexpected behavior of an inverted rye chromosome arm in wheat. *Chromosoma* **117:** 569-578.

Lukaszewski, A.J., K. Rybka, L. Korzun, S. Malyshev, B. Lapinski, and R. Whitkus. 2004. Genetic and physical mapping of homoelogous recombination points involving wheat chromosome 2B and rye chromosome 2R. *Genome* **47:** 36-45.

Lukaszewski, J. and C.A. Curtis. 1993. Physical distribution of recombination in B-genome chromosomes of tetraploid wheat. *Theor Appl Genet* **86:** 121-127.

Luo, M.C., Z.L. Yang, R.S. Kota, and J. Dvorak. 2000. Recombination of chromosomes 3A(m) and 5A(m) of Triticum monococcum with homeologous chromosomes 3A and 5A of wheat: the distribution of recombination across chromosomes. *Genetics* **154:** 1301-1308.

Lynn, A., K.E. Koehler, L. Judis, E.R. Chan, J.P. Cherry, S. Schwartz, A. Seftel, P.A. Hunt, and T.J. Hassold. 2002. Covariation of synaptonemal complex length and mammalian meiotic exchange rates. *Science* **296:** 2222-2225.

Lynn, A., R. Soucek, and G.V. Borner. 2007. ZMM proteins during meiosis: crossover artists at work. *Chromosome Res* **15:** 591-605.

Macaisne, N., M. Novatchkova, L. Peirera, D. Vezon, S. Jolivet, N. Froger, L. Chelysheva, M. Grelon, and R. Mercier. 2008. SHOC1, an XPF endonuclease-related protein, is essential for the formation of class I meiotic crossovers. *Curr Biol* **18:** 1432-1437.

Maestra, B., J. Hans de Jong, K. Shepherd, and T. Naranjo. 2002. Chromosome arrangement and behaviour of two rye homologous telosomes at the onset of meiosis in disomic wheat-5RL addition lines with and without the Ph1 locus. *Chromosome Res* **10:** 655-667.

Mahadevaiah, S.K., J.M. Turner, F. Baudat, E.P. Rogakou, P. de Boer, J. Blanco-Rodriguez, M. Jasin, S. Keeney, W.M. Bonner, and P.S. Burgoyne. 2001. Recombinational DNA double-strand breaks in mice precede synapsis. *Nat Genet* **27:** 271-276.

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

Martin, O.C. and F. Hospital. 2006. Two- and three-locus tests for linkage analysis using recombinant inbred lines. *Genetics* **173:** 451-459.

Martin, W.J., J. McCallum, M. Shigyo, J. Jakse, J.C. Kuhl, N. Yamane, M. Pither-Joyce, A.F. Gokce, K.C. Sink, C.D. Town, and M.J. Havey. 2005. Genetic mapping of expressed sequences in onion and in silico comparisons with rice show scant colinearity. *Mol Genet Genomics* **274:** 197-204.

Martinez-Perez, E. and G. Moore. 2008. To check or not to check? The application of meiotic studies to plant breeding. *Curr Opin Plant Biol* **11:** 222-227.

Martinez-Perez, E., P. Shaw, L. Aragon-Alcaide, and G. Moore. 2003. Chromosomes form into seven groups in hexaploid and tetraploid wheat as a prelude to meiosis. *Plant J* **36:** 21-29.

Martinez, I., M. Bernard, P. Nicolas, and S. Bernard. 1994. Study of androgenetic performance and molecular characterisation of a set of wheat-rye addition lines. *Theor Appl Genet* **89:** 982-990.

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

Matise, T.C., C.J. Porter, S. Buyske, A.J. Cuttichia, E.P. Sulman, and P.S. White. 2002. Systematic evaluation of map quality: human chromosome 22. *Am J Hum Genet* **70:** 1398-1410.

McKim, K.S. 2007. Meiotic pairing: a place to hook up. *current Biology* **17**.

McMahill, M.S., C.W. Sham, and D.K. Bishop. 2007. Synthesis-dependent strand annealing in meiosis. *PLoS Biol* **5:** e299.

Mcpeek, M.S. and T.P. Speed. 1995. Modeling interference in genetic recombination. *Genetics* **139:** 1031-1044.

Menz, M.A., R.R. Klein, J.E. Mullet, J.A. Obert, N.C. Unruh, and P.E. Klein. 2002. A high-density genetic map of Sorghum bicolor (L.) Moench based on 2926 AFLP, RFLP and SSR markers. *Plant Mol Biol* **48:** 483-499.

Mercier, R. and M. Grelon. 2008. Meiosis in plants: ten years of gene discovery. *Cytogenet Genome Res* **120:** 281-290.

Merker, J.D., M. Dominska, P.W. Greenwell, E. Rinella, D.C. Bouck, Y. Shibata, B.D. Strahl, P. Mieczkowski, and T.D. Petes. 2008. The histone methylase Set2p and the histone deacetylase Rpd3p repress meiotic recombination at the HIS4 meiotic recombination hotspot in Saccharomyces cerevisiae. *DNA Repair (Amst)* **7:** 1298-1308.

Meunier, J. and L. Duret. 2004. Recombination drives the evolution of GC-content in the human genome. *Mol Biol Evol* **21:** 984-990.

Mezard, C. 2006. Meiotic recombination hotspots in plants. *Biochem Soc Trans* **34:** 531-534.

Mezard, C., J. Vignard, J. Drouaud, and R. Mercier. 2007. The road to crossovers: plants have their say. *Trends Genet* **23:** 91-99.

Mieczkowski, P.A., M. Dominska, M.J. Buck, J.D. Lieb, and T.D. Petes. 2007. Loss of a histone deacetylase dramatically alters the genomic distribution of Spo11p-catalyzed DNA breaks in Saccharomyces cerevisiae. *Proc Natl Acad Sci U S A* **104:** 3955-3960.

Mikhailova, E.I., T. Naranjo, K. Shepherd, J. Wennekes-van Eden, C. Heyting, and J.H. de Jong. 1998. The effect of the wheat Ph1 locus on chromatin organisation and meiotic chromosome pairing analysed by genome painting. *Chromosoma* **107:** 339-350.

Mimitou, E.P. and L.S. Symington. 2008. Sae2, Exo1 and Sgs1 collaborate in DNA double-strand break processing. *Nature* **455:** 770-774.

Moens, P.B. 2006. Zebrafish: chiasmata and interference. *Genome* **49:** 205-208.

Moens, PB., E. Marcon, JS. Shore, N. Kochakpour, B. Spyropoulos B. 2007. Initiation and resolution of interhomolog connections: crossover and non-crossover sites along mouse synaptonemal complexes. *J Cell Sci* **120**: 1017-1027.

Moretzsohn, M.C., L. Leoi, K. Proite, P.M. Guimaraes, S.C. Leal-Bertioli, M.A. Gimenes, W.S. Martins, J.F. Valls, D. Grattapaglia, and D.J. Bertioli. 2005. A microsatellite-based, gene-rich linkage map for the AA genome of Arachis (Fabaceae). *Theor Appl Genet* **111:** 1060-1071.

Muller, H.J. 1916. The mechanisms of crossing-over. *Am Nat* **50:** 193-221.

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

Myers, S., C. Freeman, A. Auton, P. Donnelly, and G. McVean. 2008. A common sequence motif associated with recombination hot spots and genome instability in humans. *Nat Genet*.

Myers, S., C.C. Spencer, A. Auton, L. Bottolo, C. Freeman, P. Donnelly, and G. McVean. 2006. The distribution and causes of meiotic recombination in the human genome. *Biochem Soc Trans* **34:** 526-530.

Naranjo, T. and E. Corredor. 2004. Clustering of centromeres precedes bivalent chomosome pairing of polyploid wheats. *Trends Plant Science* **9**: 214-217.

Naranjo, T. and E. Corredor. 2008. Nuclear architecture and chromosome dynamics in the search of the pairing partner in meiosis in plants. *Cytogenet Genome Res* **120:** 320-330.

Neale, M.J. and S. Keeney. 2006. Clarifying the mechanics of DNA strand exchange in meiotic recombination. *Nature* **442:** 153-158.

Neumann, R. and A.J. Jeffreys. 2006. Polymorphism in the activity of human crossover hotspots independent of local DNA sequence variation. *Hum Mol Genet* **15:** 1401-1411.

Nishant, K.T., C. Kumar, and M.R. Rao. 2006. HUMHOT: a database of human meiotic recombination hot spots. *Nucleic Acids Res* **34:** D25-28.

Nishant, K.T. and M.R. Rao. 2006. Molecular features of meiotic recombination hot spots. *Bioessays* **28:** 45-56.

Niu, H., L. Wan, B. Baumgartner, D. Schaefer, J. Loidl, and N.M. Hollingsworth. 2005. Partner choice during meiosis is regulated by Hop1-promoted dimerization of Mek1. *Mol Biol Cell* **16:** 5804-5818.

Nonomura, K.I., M. Nakano, K. Murata, K. Miyoshi, M. Eiguchi, A. Miyao, H. Hirochika, and N. Kurata. 2004. An insertional mutation in the rice PAIR2 gene, the ortholog of Arabidopsis ASY1, results in a defect in homologous chromosome pairing during meiosis. *Mol Genet Genomics* **271:** 121-129.

O'Donoughue, L.S., M.E. Sorrells, S.D. Tanksley, E. Autrique, A.V. Deynze, S.F. Kianian, R.L. Phillips, B. Wu, H.W. Rines, P.J. Rayapati, M. Lee, G.A. Penner, G. Fedak, S.J. Molnar, D. Hoffman, and C.A. Salas. 1995. A molecular linkage map of cultivated oat. *Genome* **38:** 368-380.

Oliver, M., J. Garcia-Mas, M. Cardus, N. Pueyo, A.L. Lopez-Sese, M. Arroyo, H. Gomez-Paniagua, P. Arus, and M.C. de Vicente. 2001. Construction of a reference linkage map for melon. *Genome* **44:** 836-845.

Page, S.L. and R.S. Hawley. 2004. The genetics and molecular biology of the synaptonemal complex. *Annu Rev Cell Dev Biol* **20:** 525-558.

Paigen, K., J.P. Szatkiewicz, K. Sawyer, N. Leahy, E.D. Parvanov, S.H. Ng, J.H. Graber, K.W. Broman, and P.M. Petkov. 2008. The recombinational anatomy of a mouse chromosome. *PLoS Genet* **4:** e1000119.

Paques, F. and J.E. Haber. 1999. Multiple pathways of recombination induced by double-strand breaks in Saccharomyces cerevisiae. *Microbiol Mol Biol Rev* **63:** 349-404.

Pardo-Manuel de Villena, F., E. de la Casa-Esperon, and C. Sapienza. 2000. Natural selection and the function of genome imprinting: beyond the silenced minority. *Trends Genet* **16:** 573-579.

Park, J.Y., H.W. Yoo, B.R. Kim, R. Park, S.Y. Choi, and Y. Kim. 2008. Identification of a novel human Rad51 variant that promotes DNA strand exchange. *Nucleic Acids Res* **36:** 3226-3234.

Patterson, A.H., J.E. Bowers, R. Bruggmann, I. Dubchak, J. Grimwood, H. Gundlach, G. Haberer, U. Hellsten, T. Mitros, A. Poliakov, J. Schmutz, M. Spannagl, H. Tang, X. Wang, T. Wicker, A. K. Bharti, J. Chapman, F. A. Feltus, U. Gowik, I. V. Grigoriev, E. Lyons, C. A. Maher, M. Martis, A. Narechania, R. P. Otillar, B. W. Penning, A. A. Salamov, Y. Wang, L. Zhang, N. C. Carpita, M. Freeling, A. R. Gingle, C. T. Hash, B. Keller, P. Klein, S. Kresovich, M. C. McCann, R. Ming, D. G. Peterson, M. Rahman, D. Ware, P. Westhoff, K. F. X. Mayer, J. Messing and D. S. Rokhsar. The sorghum bicolor genome and the diversification of grasses. *Nature* **457**: 551-556.

Paux, E., D. Roger, E. Badaeva, G. Gay, M. Bernard, P. Sourdille, and C. Feuillet. 2006. Characterizing the composition and evolution of homoeologous genomes in hexaploid wheat through BAC-end sequencing on chromosome 3B. *Plant J* **48:** 463-474.

Paux, E., P. Sourdille, J. Salse, C. Saintenac, F. Choulet, P. Leroy, A. Korol, M. Michalak, S. Kianian, W. Spielmeyer, E. Lagudah, D. Somers, A. Kilian, M. Alaux, S. Vautrin, H. Berges, K. Eversole, R. Appels, J. Safar, H. Simkova, J. Dolezel, M. Bernard, and C. Feuillet. 2008. A physical map of the 1-gigabase bread wheat chromosome 3B. *Science* **322:** 101-104.

Pawlowski, W.P., I.N. Golubovskaya, L. Timofejeva, R.B. Meeley, W.F. Sheridan, and W.Z. Cande. 2004. Coordination of meiotic recombination, pairing, and synapsis by PHS1. *Science* **303:** 89-92.

Pearl, H.M., C. Nagai, P.H. Moore, D.L. Steiger, R.V. Osgood, and R. Ming. 2004. Construction of a genetic map for arabica coffee. *Theor Appl Genet* **108:** 829-835.

Peng, J., A.B. Korol, T. Fahima, M.S. Roder, Y.I. Ronin, Y.C. Li, and E. Nevo. 2000. Molecular genetic maps in wild emmer wheat, Triticum dicoccoides: genome-wide coverage, massive negative interference, and putative quasi-linkage. *Genome Res* **10:** 1509-1531.

Petes, T.D. 2001. Meiotic recombination hot spots and cold spots. *Nat Rev Genet* **2:** 360-369.

Petkov, P.M., K.W. Broman, J.P. Szatkiewicz, and K. Paigen. 2007. Crossover interference underlies sex differences in recombination rates. *Trends Genet* **23:** 539-542.

Petronczki, M., M.F. Siomos, and K. Nasmyth. 2003. Un menage a quatre: the molecular biology of chromosome segregation in meiosis. *Cell* **112:** 423-440.

Phan, H.T., S.R. Ellwood, K. Adhikari, M.N. Nelson, and R.P. Oliver. 2007. The first genetic and comparative map of white lupin (Lupinus albus L.): identification of QTLs for anthracnose resistance and flowering time, and a locus for alkaloid content. *DNA Res* **14:** 59-70.

Pigozzi, M.I. and A.J. Solari. 1999. Recombination nodule mapping and chiasma distribution in spermatocytes of the pigeon, Columba livia. *Genome* **42:** 308-314.

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

Powell, J.B. and R.A. Nilan. 1963. Influence of temperature on crossing over in an inversion heterozygote in barley. *Crop Science*.

Prieto, P., P. Shaw, and G. Moore. 2004. Homologue recognition during meiosis is associated with a change in chromatin conformation. *Nat Cell Biol* **6:** 906-908.

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

Qi, L.L., B. Friebe, and B.S. Gill. 2002. A strategy for enhancing recombination in proximal regions of chromosomes. *Chromosome Res* **10:** 645-654.

Quarrie, S.A., A. Steed, C. Calestani, A. Semikhodskii, C. Lebreton, C. Chinoy, N. Steele, D. Pljevljakusic, E. Waterman, J. Weyen, J. Schondelmaier, D.Z. Habash, P. Farmer, L. Saker, D.T. Clarkson, A. Abugalieva, M. Yessimbekova, Y. Turuspekov, S. Abugalieva, R. Tuberosa, M.C. Sanguineti, P.A. Hollington, R. Aragues, A. Royo, and D. Dodig. 2005. A high-density genetic map of hexaploid wheat (Triticum aestivum L.) from the cross Chinese Spring x SQ1 and its use to compare QTLs for grain yield across a range of environments. *Theor Appl Genet* **110:** 865-880.

Reddy, K.C. and A.M. Villeneuve. 2004. C. elegans HIM-17 links chromatin modification and competence for initiation of meiotic recombination. *Cell* **118:** 439-452.

Riera-Lizarazu, O., M.I. Vales, E.V. Ananiev, H.W. Rines, and R.L. Phillips. 2000. Production and characterization of maize chromosome 9 radiation hybrids derived from an oat-maize addition line. *Genetics* **156:** 327-339.

Riera-Lizarazu, O., M.I. Vales, and S.F. Kianian. 2008. Radiation hybrid (RH) and HAPPY mapping in plants. *Cytogenet Genome Res* **120:** 233-240.

Robine, N., N. Uematsu, F. Amiot, X. Gidrol, E. Barillot, A. Nicolas, and V. Borde. 2007. Genome-wide redistribution of meiotic double-strand breaks in Saccharomyces cerevisiae. *Mol Cell Biol* **27:** 1868-1880.

Rocco, V. and A. Nicolas. 1996. Sensing of DNA non-homology lowers the initiation of meiotic recombination in yeast. *Genes Cells* **1:** 645-661.

Rockmill, B., K. Voelkel-Meiman, and G.S. Roeder. 2006. Centromere-proximal crossovers are associated with precocious separation of sister chromatids during meiosis in Saccharomyces cerevisiae. *Genetics* **174:** 1745-1754.

Roder, M.S., V. Korzun, K. Wendehake, J. Plaschke, M.H. Tixier, P. Leroy, and M.W. Ganal. 1998. A microsatellite map of wheat. *Genetics* **149:** 2007-2023.

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

Romanienko, P.J. and R.D. Camerini-Otero. 2000. The mouse Spo11 gene is required for meiotic chromosome synapsis. *Mol Cell* **6:** 975-987.

Ross-Ibarra, J. 2007. Genome size and recombination in angiosperms: a second look. *J Evol Biol* **20:** 800-806.

Ross, L.O., R. Maxfield, and D. Dawson. 1996. Exchanges are not equally able to enhance meiotic chromosome segregation in yeast. *Proc Natl Acad Sci U S A* **93:** 4979-4983.

Safar, J., J. Bartos, J. Janda, A. Bellec, M. Kubalakova, M. Valarik, S. Pateyron, J. Weiserova, R. Tuskova, J. Cihalikova, J. Vrana, H. Simkova, P. Faivre-Rampant, P. Sourdille, M. Caboche, M. Bernard, J. Dolezel, and B. Chalhoub. 2004. Dissecting large and complex genomes: flow sorting and BAC cloning of individual chromosomes from bread wheat. *Plant J* **39:** 960-968.

Saintenac, C., M. Falque, O.C. Martin, E. Paux, C. Feuillet, and P. Sourdille. 2009. Detailed Recombination Studies along Chromosome 3B Provide New Insights on Crossover Distribution in Wheat (Triticum aestivum L.). *Genetics*.

Sallee, P.J. and G. Kimber. 1978. An analysis of the pairing of wheat telocentric chromosomes. *Proc 5th Int Wheat Genetics Symposium***:** 408-419.

Salse, J., S. Bolot, M. Throude, V. Jouffe, B. Piegu, U.M. Quraishi, T. Calcagno, R. Cooke, M. Delseny, and C. Feuillet. 2008. Identification and characterization of shared duplications between rice and wheat provide new insight into grass genome evolution. *Plant Cell* **20:** 11-24.

San Filippo, J., P. Sung, and H. Klein. 2008. Mechanism of Eukaryotic Homologous Recombination. *Annu Rev Biochem* **77**: 9.1-9.29.

Sanchez-Moran, E., S.J. Armstrong, J.L. Santos, F.C. Franklin, and G.H. Jones. 2002. Variation in chiasma frequency among eight accessions of Arabidopsis thaliana. *Genetics* **162:** 1415-1422.

Sanchez-Moran, E., K. Osman, J.D. Higgins, M. Pradillo, N. Cunado, G.H. Jones, and F.C. Franklin. 2008. *ASY1 coordinates early events in the plant meiotic recombination pathway*.

Sanchez-Moran, E., J.L. Santos, G.H. Jones, and F.C. Franklin. 2007. ASY1 mediates AtDMC1-dependent interhomolog recombination during meiosis in Arabidopsis. *Genes Dev* **21:** 2220-2233.

Sandovici, I., S. Kassovska-Bratinova, J.E. Vaughan, R. Stewart, M. Leppert, and C. Sapienza. 2006. Human imprinted chromosomal regions are historical hot-spots of recombination. *PLoS Genet* **2:** e101.

Scherthan, H. A bouquet makes ends meet. *Nat Rev Mol Cell Biol*. **2**: 621-627.

Schwacha, A. and N. Kleckner. 1995. Identification of double Holliday junctions as intermediates in meiotic recombination. *Cell* **83:** 783-791.

Schwarzacher, T. 2003. Meiosis, recombination and chromosomes: a review of gene isolation and fluorescent in situ hybridization data in plants. *J Exp Bot* **54:** 11-23.

Sears, E.R. 1939. Cytogenetic studies with polyploid species of wheat. I. Chromosomal aberrations in the progeny of a haploid of *Triticum vulgare*. *Genetics* **24:** 509-523.

Sears, E.R. 1944. Cytogenetic studies with polyploid species of wheat. II. Additional chromosomal aberrations in *Triticum vulgare*. *Genetics* **29:** 232-246.

Sears, E.R. 1954. The Aneuploids of common Wheat. *University of Missouri Agricultural Experiment Station, Columbia* **572:** 1-58.

Sears, E.R. 1966. *Nullisomic-tetrasomic combinations in hexaploid wheat*.

Sears, E.R. 1976. Genetic control of chromosome pairing in wheat. *Ann. Rev. Genet.* **10**.

Sears, E.R. 1982. A wheat mutation conditionning an intermediate level of homoeologous chromosome pairing. *Canadian Journal of Genetics and Cytology* **24:** 715-719.

Sears, E.R. and L.M.S. Sears. 1978. *The telocentric chromosomes of common wheat*, New Delhy.

See, D.R., S. Brooks, J.C. Nelson, G. Brown-Guedira, B. Friebe, and B.S. Gill. 2006. Gene evolution at the ends of wheat chromosomes. *Proc Natl Acad Sci U S A* **103:** 4162-4167.

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

Sharma, P.C., A. Grover, and G. Kahl. 2007. Mining microsatellites in eukaryotic genomes. *Trends Biotechnol* **25:** 490-498.

Shaw, P. and G. Moore. 1998. Meiosis: vive la difference! *Curr Opin Plant Biol* **1:** 458-462.

Shen, B., D.M. Wang, C.L. McIntyre, and C.J. Liu. 2005. A 'Chinese Spring' wheat (Triticum aestivum L.) bacterial artificial chromosome library and its use in the isolation of SSR markers for targeted genome regions. *Theor Appl Genet* **111:** 1489-1494.

Sheridan, S.D., X. Yu, R. Roth, J.E. Heuser, M.G. Sehorn, P. Sung, E.H. Egelman, and D.K. Bishop. 2008. A comparative analysis of Dmc1 and Rad51 nucleoprotein filaments. *Nucleic Acids Res* **36:** 4057-4066.

Sherman, J.D. and S.M. Stack. 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., J.T. Bell, R.R. Copley, M.S. Taylor, R.W. Williams, R. Mott, and J. Flint. 2006. A high-resolution single nucleotide polymorphism genetic map of the mouse genome. *PLoS Biol* **4:** e395.

Shinohara, A., S. Gasior, T. Ogawa, N. Kleckner, and D.K. Bishop. 1997. Saccharomyces cerevisiae recA homologues RAD51 and DMC1 have both distinct and overlapping roles in meiotic recombination. *Genes Cells* **2:** 615-629.

Shinohara, M., S.D. Oh, N. Hunter, and A. Shinohara. 2008. Crossover assurance and crossover interference are distinctly regulated by the ZMM proteins during yeast meiosis. *Nat Genet* **40:** 299-309.

Siaud, N., E. Dray, I. Gy, E. Gerard, N. Takvorian, and M.P. Doutriaux. 2004. Brca2 is involved in meiosis in Arabidopsis thaliana as suggested by its interaction with Dmc1. *Embo J* **23:** 1392-1401.

Sigurdsson, M.I., A.V. Smith, H.T. Bjornsson, and J.J. Jonsson. 2009. HapMap methylation-associated SNPs, markers of germline DNA methylation, positively correlate with regional levels of human meiotic recombination. *Genome Res* in press.

Singer, A., H. Perlman, Y. Yan, C. Walker, G. Corley-Smith, B. Brandhorst, and J. Postlethwait. 2002. Sex-specific recombination rates in zebrafish (Danio rerio). *Genetics* **160:** 649-657.

Sinha, M. and C.L. Peterson. 2008. A Rad51 presynaptic filament is sufficient to capture nucleosomal homology during recombinational repair of a DNA double-strand break. *Mol Cell* **30:** 803-810.

Solignac, M., F. Mougel, D. Vautrin, M. Monnerot, and J.M. Cornuet. 2007. A third-generation microsatellite-based linkage map of the honey bee, Apis mellifera, and its comparison with the sequence-based physical map. *Genome Biol* **8:** R66.

Sourdille, P., T. Cadalen, H. Guyomarc'h, J.W. Snape, M.R. Perretant, G. Charmet, C. Boeuf, S. Bernard, and M. Bernard. 2003. An update of the Courtot x Chinese Spring intervarietal molecular marker linkage map for the QTL detection of agronomic traits in wheat. *Theor Appl Genet* **106:** 530-538.

Stahl, F.W., H.M. Foss, L.S. Young, R.H. Borts, M.F. Abdullah, and G.P. Copenhaver. 2004. Does crossover interference count in Saccharomyces cerevisiae? *Genetics* **168:** 35-48.

Stein, N., C. Feuillet, T. Wicker, E. Schlagenhauf, and B. Keller. 2000. Subgenome chromosome walking in wheat: a 450-kb physical contig in Triticum monococcum L. spans the Lr10 resistance locus in hexaploid wheat (Triticum aestivum L.). *Proc Natl Acad Sci U S A* **97:** 13436-13441.

Stein, N., M. Prasad, U. Scholz, T. Thiel, H. Zhang, M. Wolf, R. Kota, R.K. Varshney, D. Perovic, I. Grosse, and A. Graner. 2007. A 1,000-loci transcript map of the barley genome: new anchoring points for integrative grass genomics. *Theor Appl Genet* **114:** 823-839.

Steiner, W.W. and G.R. Smith. 2005a. Natural meiotic recombination hot spots in the Schizosaccharomyces pombe genome successfully predicted from the simple sequence motif M26. *Mol Cell Biol* **25:** 9054-9062.

Steiner, W.W. and G.R. Smith. 2005b. Optimizing the nucleotide sequence of a meiotic recombination hotspot in Schizosaccharomyces pombe. *Genetics* **169:** 1973-1983.

Stewart, M.N. and D.S. Dawson. 2008. Changing partners: moving from non-homologous to homologous centromere pairing in meiosis. *Trends Genet*.

Suenaga, K., M. Khairallah, H.M. William, and D.A. Hoisington. 2005. A new intervarietal linkage map and its application for quantitative trait locus analysis of "gigas" features in bread wheat. *Genome* **48:** 65-75.

Sun, F., M. Oliver-Bonet, T. Liehr, H. Starke, P. Turek, E. Ko, A. Rademaker, and R.H. Martin. 2006. Variation in MLH1 distribution in recombination maps for individual chromosomes from human males. *Hum Mol Genet* **15:** 2376-2391.

Sun, Z., Z. Wang, J. Tu, J. Zhang, F. Yu, P.B. McVetty, and G. Li. 2007. An ultradense genetic recombination map for Brassica napus, consisting of 13551 SRAP markers. *Theor Appl Genet* **114:** 1305-1317.

Sung, P. and H. Klein. 2006. Mechanism of homologous recombination: mediators and helicases take on regulatory functions. *Nat Rev Mol Cell Biol* **7:** 739-750.

Sutton, T., R. Whitford, U. Baumann, C. Dong, J.A. Able, and P. Langridge. 2003. The Ph2 pairing homoeologous locus of wheat (Triticum aestivum): identification of candidate meiotic genes using a comparative genetics approach. *Plant J* **36:** 443-456.

Symington, L.S. 2002. Role of RAD52 epistasis group genes in homologous recombination and double-strand break repair. *Microbiol Mol Biol Rev* **66:** 630-670, table of contents.

Szostak, J.W., T.L. Orr-Weaver, R.J. Rothstein, and F.W. Stahl. 1983. The double-strand-break repair model for recombination. *Cell* **33:** 25-35.

Tang, S., J.K. Yu, B. Slabaugh, K. Shintani, and J. Knapp. 2002. Simple sequence repeat map of the sunflower genome. *Theor Appl Genet* **105:** 1124-1136.

Tease, C. and M.A. Hulten. 2004. Inter-sex variation in synaptonemal complex lengths largely determine the different recombination rates in male and female germ cells. *Cytogenet Genome Res* **107:** 208-215.

Tenaillon, M.I., M.C. Sawkins, L.K. Anderson, S.M. Stack, J. Doebley, and B.S. Gaut. 2002. Patterns of diversity and recombination along chromosome 1 of maize (Zea mays ssp. mays L.). *Genetics* **162:** 1401-1413.

Thoquet, P., M. Gherardi, E.P. Journet, A. Kereszt, J.M. Ane, J.M. Prosperi, and T. Huguet. 2002. The molecular genetic linkage map of the model legume Medicago truncatula: an essential tool for comparative legume genomics and the isolation of agronomically important genes. *BMC Plant Biol* **2:** 1.

Timmermans, M.C., O.P. Das, J.M. Bradeen, and J. Messing. 1997. Region-specific cis- and trans-acting factors contribute to genetic variability in meiotic recombination in maize. *Genetics* **146:** 1101-1113.

Torada, A., M. Koike, K. Mochida, and Y. Ogihara. 2006. SSR-based linkage map with new markers using an intraspecific population of common wheat. *Theor Appl Genet* **112:** 1042-1051.

Troggio, M., G. Malacarne, G. Coppola, C. Segala, D.A. Cartwright, M. Pindo, M. Stefanini, R. Mank, M. Moroldo, M. Morgante, M.S. Grando, and R. Velasco. 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, M.J., R. Antonise, D. Lavelle, O. Ochoa, A. Kozik, H. Witsenboer, S.B. Fort, M.J. Jeuken, R.V. Kesseli, P. Lindhout, R.W. Michelmore, and J. Peleman. 2007. A high-density, integrated genetic linkage map of lettuce (Lactuca spp.). *Theor Appl Genet* **115:** 735-746.

Tsai, C.J., D.G. Mets, M.R. Albrecht, P. Nix, A. Chan, and B.J. Meyer. 2008. Meiotic crossover number and distribution are regulated by a dosage compensation protein that resembles a condensin subunit. *Genes Dev* **22:** 194-211.

van Os, H., S. Andrzejewski, E. Bakker, I. Barrena, G.J. Bryan, B. Caromel, B. Ghareeb, E. Isidore, W. de Jong, P. van Koert, V. Lefebvre, D. Milbourne, E. Ritter, J.N. van der Voort, F. Rousselle-Bourgeois, J. van Vliet, R. Waugh, R.G. Visser, J. Bakker, and H.J. van Eck. 2006. Construction of a 10,000-marker ultradense genetic recombination map of potato: providing a framework for accelerated gene isolation and a genomewide physical map. *Genetics* **173:** 1075-1087.

Wahls, W.P., E.R. Siegel, and M.K. Davidson. 2008. Meiotic recombination hotspots of fission yeast are directed to loci that express non-coding RNA. *PLoS ONE* **3:** e2887.

Wang, C.J., L. Harper, and W.Z. Cande. 2006. High-resolution single-copy gene fluorescence in situ hybridization and its use in the construction of a cytogenetic map of maize chromosome 9. *Plant Cell* **18:** 529-544.

Wang, G., V. Hyne, S. Chao, Y. Henry, J. De Buyser, M.D. Gale, and J.W. Snape. 1995. A comparison of male and female recombination frequency in wheat using RFLP maps of homoeologous group 6 et 7 chromosomes. *Theor Appl Genet* **91:** 744-746.

Wang, G., J. Pan, X. Li, H. He, A. Wu, and R. Cai. 2005a. Construction of a cucumber genetic linkage map with SRAP markers and location of the genes for lateral branch traits. *Sci China C Life Sci* **48:** 213-220.

Wang, Y.P., X.X. Zhao, K. Sonntag, P. Welhing, and R.J. Snowdon. 2005b. Behaviour of Sinapis alba chromosomes in a Brassica napus background revealed by genomic in-situ hybridization. *CHROMOSOME RES* **13:** 819-826.

Werner, J.E., T.R. Endo, and B.S. Gill. 1992. Toward a cytogenetically based physical map of the wheat genome. *Proc Natl Acad Sci U S A* **89:** 11307-11311.

Wijnker, E. and H. de Jong. 2008. Managing meiotic recombination in plant breeding. *Trends Plant Sci* **13:** 640-646.

Wilfert, L., J. Gadau, and P. Schmid-Hempel. 2007. Variation in genomic recombination rates among animal taxa and the case of social insects. *Heredity* **98:** 189-197.

Williams, C.G., M.M. Goodman, and C.W. Stuber. 1995. Comparative recombination distances among Zea mays L. inbreds, wide crosses and interspecific hybrids. *Genetics* **141:** 1573-1581.

Winckler, W., S.R. Myers, D.J. Richter, R.C. Onofrio, G.J. McDonald, R.E. Bontrop, G.A. McVean, S.B. Gabriel, D. Reich, P. Donnelly, and D. Altshuler. 2005. Comparison of fine-scale recombination rates in humans and chimpanzees. *Science* **308:** 107-111.

Woolbright, S.A., S.P. Difazio, T. Yin, G.D. Martinsen, X. Zhang, G.J. Allan, T.G. Whitham, and P. Keim. 2008. A dense linkage map of hybrid cottonwood (Populus fremontii x P. angustifolia) contributes to long-term ecological research and comparison mapping in a model forest tree. *Heredity* **100:** 59-70.

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

Wu, S.B., G. Collins, and M. Sedgley. 2004. A molecular linkage map of olive (Olea europaea L) based on RAPD, microsatellite, and SCAR markers. *Genome* **47:** 26-35.

Wu, T.C. and M. Lichten. 1994. Meiosis-induced double-strand break sites determined by yeast chromatin structure. *Science* **263:** 515-518.

Xia, Z., Y. Tsubokura, M. Hoshi, M. Hanawa, C. Yano, K. Okamura, T.A. Ahmed, T. Anai, S. Watanabe, M. Hayashi, T. Kawai, K.G. Hossain, H. Masaki, K. Asai, N. Yamanaka, N. Kubo, K. Kadowaki, Y. Nagamura, M. Yano, T. Sasaki, and K. Harada. 2007. An integrated high-density linkage map of soybean with RFLP, SSR, STS, and AFLP markers using A single F2 population. *DNA Res* **14:** 257-269.

Xue, S., Z. Zhang, F. Lin, Z. Kong, Y. Cao, C. Li, H. Yi, M. Mei, H. Zhu, J. Wu, H. Xu, D. Zhao, D. Tian, C. Zhang, and Z. Ma. 2008. A high-density intervarietal map of the wheat genome enriched with markers derived from expressed sequence tags. *Theor Appl Genet* **117:** 181-189.

Yamada, T., K. Mizuno, K. Hirota, N. Kon, W.P. Wahls, E. Hartsuiker, H. Murofushi, T. Shibata, and K. Ohta. 2004. Roles of histone acetylation and chromatin remodeling factor in a meiotic recombination hotspot. *Embo J* **23:** 1792-1803.

Yang, F., S. Eckardt, N.A. Leu, K.J. McLaughlin, and P.J. Wang. 2008. Mouse TEX15 is essential for DNA double-strand break repair and chromosomal synapsis during male meiosis. *J Cell Biol* **180:** 673-679.

Yin, T.M., X.R. Wang, B. Andersson, and E. Lerceteau-Kohler. 2003. Nearly complete genetic maps of Pinus sylvestris L. (Scots pine) constructed by AFLP marker analysis in a full-sib family. *Theor Appl Genet* **106:** 1075-1083.

Zhang, L. 2006. Etude de la portabilité de marqueurs microsatellites issus d'EST de blé tendre (*T. aestivum*) ou de riz (*O. sativa*) vers des espèces apparentées et évaluation de leur intérêt pour la structuration des ressources génétiques. Blaise Pascal, Clermont-Ferrand.

Zhang, L.P., A. Khan, D. Nino-Liu, and M.R. Foolad. 2002. A molecular linkage map of tomato displaying chromosomal locations of resistance gene analogs based on a Lycopersicon esculentum x Lycopersicon hirsutum cross. *Genome* **45:** 133-146.

Zhang, X. 2008. The epigenetic landscape of plants. *Science* **320:** 489-492.

Zhao, H., M.S. McPeek and T.P. Speed. 1995. Statistical analysis of chromatid interference. *Genetics* **139**: 1057-1065.

Zonneveld, B.J., I.J. Leitch, and M.D. Bennett. 2005. First nuclear DNA amounts in more than 300 angiosperms. *Ann Bot (Lond)* **96:** 229-244.

Zwierzykowski, Z., E. Zwierzykowska, M. Taciak, N. Jones, A. Kosmala, and P. Krajewski. 2008. Chromosome pairing in allotetraploid hybrids of Festuca pratensis x Lolium perenne revealed by genomic in situ hybridization (GISH). *Chromosome Res* **16:** 575-585.

VIII Annexes

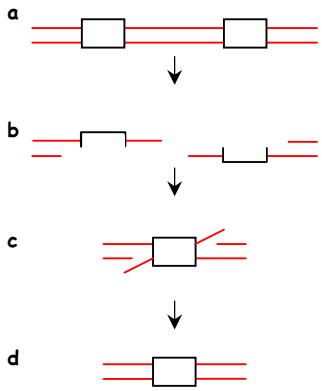


Figure 31 Modèle de recombinaison homologue SSA.

a-b/ Formation d'une cassure double-brin au sein d'une molécule d'ADN entre deux répétitions (boîtes). La digestion des extrémités 5' jusqu'au niveau de ces répétitions entraîne la formation de deux séquences complémentaires, entre les deux répétitions, qui s'apparient (**c**). L'excision des extrémités 3', la synthèse d'ADN et la ligation permettent la restauration des deux brins avec la perte du fragment situé entres les deux répétitions (**d**).

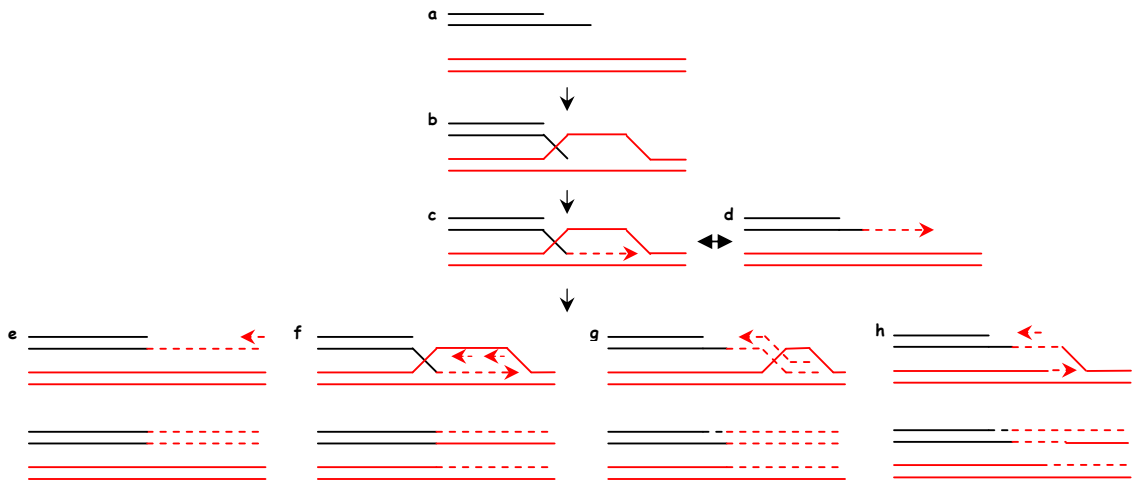


Figure 32 Modèle de recombinaison homologue BIR.

a/ Formation d'une cassure double-brin qui entraîne la présence d'une seule extrémité. **b-c**/ Le bout saillant envahit la chromatide homologue, il y a formation d'une boucle, synthèse d'ADN et déplacement de la boucle. Le bout envahissant peut se dissocier (**d**) et se réassocier. **e**/ Le bout envahissant se dissocie et il y a synthèse du brin complémentaire. **f**/ Pendant la migration de la boucle, la synthèse du brin complémentaire au brin envahissant s'effectue. Il y a formation d'une fourche de réplication unidirectionnelle. La résolution de la structure entraîne la présence de deux molécules semi-réplicatives. **g**/ La migration de la boucle entraîne le déplacement des deux brins nouvellement synthétisés. **h**/ La boucle peut aussi être résolue de manière précoce entraînant la formation de deux molécules semi-réplicatives.

Taxons / Espèces	Caryotypes (N)	Taille génétique en cM	Taille physique en Mb	Fréquence de recombinaison en cM/Mb	Références
Arthropodes					
Ixodes scapularis	14	616	2100	0.29	2
Acyrthosiphon pisum	4	360	303	1.19	2
Rhyzopertha dominica	9	390	476	0.82	2
Leptinotarsa	18	1032	450	2.29	2
Champignons					
Cryptococcus neoformans	14	1500	20	75	1
Coprinus cinereus	13	346	36	37.4	1
Cochliobolus sativus	15	849	33	25.7	1
Hymenoptères sociaux					
Acromyrmex echinatior	18	2236	335	6.7	1
Apis melifera	16	4115	262	16	1
Bombus terrestris	18	2760	625	4.4	1
Pogonomyrmex rugosus	16	3558	255	14	1
Hymenoptères parasites					
Bracon Hebetor	10	800	165	4.8	1
Bracon sp. Near Hebetor	10	536	165	3.2	1
Nasonia vitripennis x giraulti	5	765	312	2.5	1
Trichogramma brassica	5	1330	246	5.4	1
Insectes					
Acrythosiphon pisum pisum	5	294	360	1	1
Aedes aegypti	3	205	793	0.3	1
Anopheles gambiae	3	215	278	0.8	1
Armigeres sulbalbatus	3	182	1215	0.1	1
Bombyx mori	28	1305	510	2.6	1
Culex pipiens	3	166	528	0.3	1
Drosophila mauretania	3	463	147	3.1	1
Drosophila melanogaster	3	259	157	1.6	1
Heliconius erato	21	2400	396	6.1	1
Helioconius melpomene	21	1616	292	5.5	1
Laupala sp.	8	2330	1900	1.2	1
Leptinotarsa decemlineata	18	1302	450	2.3	1
Rhyzopertha dominica	9	390	476	0.8	1

Tribolium castaneum	10	571	199	2.9	1
Tribolium confusum	10	969	245	4.0	1
Nématodes					
Caenorhabdtis elegans	6	298	100	2.94	2
Protostomes					
Crassostrea gigas	10	1096	890	1.2	1
Crassostrea virginica	10	1077	675	1.6	1
Dapnia pulex	12	1206	235	5.1	1
Heterodera Glycines	9	687	93	7.4	1
Penaeus monodon	44	2292	2000	1.1	1
Penaeus vannamei	44	4015	2393	1.7	1
Pristionchus pacificus	6	339	100	3.4	1
Vertébrés					
Ambystoma tigrinum	14	5251	26950	0.2	1
Danio rerio	25	1700	2445	1.35	2
Equus callallus	32	2772	3087	0.9	1
Gallus gallus	39	3800	1225	3.1	1
Homo sapiens	23	3615	3191	1.1	1
Macaca mulatta	21	2048	3077	0.7	1
Mus musculus	20	1361	3179	0.4	1
Takifugu rubripes	22	1135	365	3.1	1
Oncorhynchus mykiss	30	4590	2680	1.71	2
Papio hamadryas	20	2375	3450	0.69	2
Plasmodium falciparum	14	1556	23	67.7	1
Rattus norvegicus	21	1503	2750	0.60	2
Salmo salar	30	901/103	3050	0.16	2
Toxoplasma gondii	14	592	65	9.6	1
Trypanosoma brucei	11	1358	26	52.2	1

Tableau 11 Fréquence moyenne de recombinaison chez différentes espèces.

Références : 1 = (Wilfert et al. 2007), 2 = (Beye et al. 2006).

Espèces	Mâle[a]	Femelle[a]	Rapport mâle / femelle
Animaux			
Bettongia penicillata	28	27.9	1.00
Bos Taurus	3.567	3.765	0.95
Canis familiaris	1290	1822	0.71
Chorthippus brunneus	13.6	13.1	1.04
Chorthippus jucundus	12.66	12.65	1.00
Chorthippus parallelus	13.38	11.81	1.13
Chorthippus vagans	11.25	10.56	1.07
Chortoicetes terminifera	13.1	11.6	1.13
Chrysochraon dispar	12.6	12.1	1.04
Danio rerio	999.9	2852.7	0.35
Dendrocoelum lacteum	11.8	20.4	0.58
Equus caballus	62.4	79.9	0.78
Euchorthippus chopardi	11.62	10.48	1.11
Euchorthippus pulvinatus	11.81	11.06	1.07
Eyprepocnemis plorans	14.1	12	1.18
Gallus domesticus	3062.1	3026.8	1.02
Gyratrix hermaphtoditus	5.2	4.5	1.16
Homo sapiens	2730	4435	0.62
Macaca mulatta	39.6	31.7	1.25
Macropus eugenii	-	-	1.28
Melanoplus femur-rubrum	13.5	14	0.96
Mus musculus	20.9	28.9	0.72
Myrmeleotettix maculatus	14.4	13.2	1.09
Notoplana igiliensis	12.5	18.6	0.67
Omocestus panteli	11.8	11.26	1.05
Oncorhyncus mykiss	467.5	1034.9	0.45
Ovis aries	3875.8	3253.4	1.19
Papio hamadryas	24.3	79.2	0.31
Papio papio	41.5	39.6	1.05
Paradistomoides orientalis	32.3	32.3	1.00
Parapleurus allaceus	12.3	12.9	0.95
Rana esculenta	25.2	45.7	0.55
Salamandra salamandra	24	36.8	0.65
Schimdtea polychroa	0.07	0.23	0.30
Sminthopsis crassicaudata	13.6	10.2	1.33
Stethophyma grossum	11.3	13.7	0.82
Sus scrofa	5452.1	7614.9	0.72
Triturus alpestris	32.3	24.5	1.32
Triturus cristatus	36.5	24	1.52
Triturus helveticus	22	25	0.88
Triturus marmoratus	25.7	29	0.89
Végétaux			
Acacia mangium	1561	1537	1.02
Actinidia species	1104.1	1758.5	0.63
Allium cepa	22.4	17.9	1.25
Allium consanguineum	21.9	17.5	1.25

Allium flavum	14.9	18.8	0.79
Allium macranthum	42.3	58.7	0.72
Allium nigrum	21.9	16.9	1.30
Allium pallens	15	19.4	0.77
Allium paniculatum	14.6	16	0.91
Allium ursinum	13.8	14.1	0.98
Arabidopsis thaliana	417.29	216.23	1.93
Brassica napus	1544	1577	0.98
Brassica nigra	418	401	1.04
Brassica oleracea	1050.8	1749.4	0.60
Castanea sativa	1054	947	1.11
Coffea canephora	211	217	0.97
Cypripedium cordigerum	16.4	19.7	0.83
Dioscorea alata	-	-	1.00
Dioscorea rotundata	852	891	0.96
Dioscorea tokoro	570.9	489.4	1.17
Endymion nonscriptus	17.7	18.2	0.97
Epipactis consimilis	25.8	27.1	0.95
Epipactis latifolia	30.7	29.1	0.95
Eucalyptus grandis	1415	1551	0.91
Eucalyptus urophylla	1101	1331	0.83
Fritillaria meleagris	24.8	37.8	0.66
Hevea species	-	-	0.83
Hordeum bulbosum	1203.7	1016.9	1.18
Hordeum vulgare	13.9	13.7	1.01
Humulus lupulus	227.4	346.7	0.66
Lilium hansonii	40	49	0.82
Lilium henryl	41.2	44.4	0.93
Lilium longiflorum	27.3	31.5	0.87
Lilium martagon	36.3	41	0.89
Lilium pardalinum	31.2	36.9	0.85
Lilium regale	41.8	45	0.93
Lilium sargentiae	31.2	42	0.74
Lilium speciosum	26.4	33.9	0.78
Listera ovata	26.9	30.3	0.89
Malus pumila	559	447	1.25
Manhihot esculenta	49.1	40.8	1.20
Neottia listeroides	29.3	31.1	0.94
Passiflora edulis	783.5	727.7	1.08
Pennisetum glaucum	267	234	1.14
Picea abies	1557	1381	1.13
Pinus pinaster	1538.8	1169.4	1.32
Pinus sylvestris	2437	1885	1.29
Pinus taeda	1983.7	1339.5	1.48
Populus species	1063.6	1071.7	0.99
Quercus robur	921.7	893.2	1.03
Rhododendron species	164	171	0.96
Rhoeo disolor	10.2	11.4	0.89

Rosa species	287.3	238.4	1.21
Secale cereale	10.7	10.6	1.01
Solanum peruvianum	-	-	0.72
Solanum species	1097	1299	0.84
Solanum chacoense	514	709	0.72
Solanum tuberosum	382.9	525.1	0.73
Trigonella foenum	21.3	21.1	1.01
Triticum aestivum	378	328	1.15
Tulbaghia acutiloba	14.4	15.8	0.91
Tulbaghia leucantha	12.4	15.5	0.80
Tulbaghia pulchella	12.2	13.7	0.89
Tulbaghia violacea	11	14.3	0.77
Vicia faba	20.6	16	1.29
Vitis vinifera	816	767	1.06

Tableau 12 Différences de fréquences de recombinaison entre la méiose mâle et la méiose femelle chez différentes espèces animales et végétales.

[a]/ indique des valeurs de nombre de chiasma ou de longueurs de cartes génétiques. – données non déterminées. Les données sont issues de la référence (Lenormand et Dutheil 2005b).

Produits	**Concentration finale en mg/l**
$KH_2\ PO_4$	300
KNO_3	1000
$MgSO_4, 7H_2O$	71.50
$Ca\ (NO_3)_2, 4H_2O$	500
KCl	65
NH_4NO_3	1000
$MnSO_4, H_2O$	4.9
$ZnSO_4, 7H_2O$	2.7
H_3BO_3	1.6
IK	0.75
Glycine	2
Acide Nicotinique	5
Pyridoxine HCl	5
Thiamine HCl	1
Myo-Inositol	100
Glutamine	500
Na_2, EDTA	18.62
$FeSO_4, 7H_2O$	13.94
Acide indolacétique	1
Acide naphtalène acétique	0.2
Agarose	2500
Gelrite	3000
Saccharose	20000

Tableau 13 Composition du milieu solide de type Muller utilisé pour le dévelopement des embryons.

Marqueurs	Assignation	Types	Conditions d'amplification [a]
gwm389	3BS3-0.87-1.00	SSR	M13
cfp3132	3BS3-0.87-1.00	ISBP	SYBR green
cfp140	3BS3-0.87-1.00	ISBP	SYBR green
cfb3530	3BS3-0.87-1.00	SSR	M13
cfb3296	3BS3-0.87-1.00	SSR	M13
nw1600	3BS3-0.87-1.00	SSR	M13
cfb3417	3BS8-0.78-0.87	SSR	M13
wmm756	3BS8-0.78-0.87	SSR	rouge de crésol
swm13	3BS8-0.78-0.87	SSR	M13
cfb5010	3BS8-0.78-0.87	SSR	M13
gwm533	3BS8-0.78-0.87	SSR	M13
gpw7757	3BS8-0.78-0.87	SSR	M13
barc133	3BS8-0.78-0.87	SSR	rouge de crésol
cfb6012	3BS8-0.78-0.87	SSR	M13
cfb6016	3BS8-0.78-0.87	SSR	M13
cfp78	3BS8-0.78-0.87	ISBP	SYBR green
cs_ssr20	3BS8-0.78-0.87	SSR	rouge de crésol
gwm493	3BS8-0.78-0.87	SSR	M13
wmm1831	3BS8-0.78-0.87	SSR	rouge de crésol
wmc754	3BS8-0.78-0.87	SSR	rouge de crésol
nw2711	3BS8-0.78-0.87	SSR	rouge de crésol
cfp57	3BS8-0.78-0.87	ISBP	rouge de crésol
cfp171	3BS8-0.78-0.87	ISBP	rouge de crésol
cfp58	3BS8-0.78-0.87	ISBP	rouge de crésol
cfp59	3BS8-0.78-0.87	ISBP	rouge de crésol
cfp45	3BS8-0.78-0.87	ISBP	rouge de crésol
cfd79	3BS8-0.78-0.87	SSR	M13
cfb3119	3BS8-0.78-0.87	SSR	rouge de crésol
wmm1344	3BS8-0.78-0.87	SSR	M13
wmc623	3BS8-0.78-0.87	SSR	rouge de crésol
nw216	3BS8-0.78-0.87	SSR	M13
cfb3023	3BS8-0.78-0.87	SSR	M13
wmm1441	3BS8-0.78-0.87	SSR	M13
gpw3092	3BS8-0.78-0.87	SSR	M13
nw1821	3BS8-0.78-0.87	SSR	M13

nw1644	3BS8-0.78-0.87	SSR	M13
barc147	3BS8-0.78-0.87	SSR	M13
cfp1274	3BS7-0.75-0.78	ISBP	rouge de crésol
cfp1260	3BS7-0.75-0.78	ISBP	rouge de crésol
cfp1133	3BS7-0.75-0.78	ISBP	rouge de crésol
cfp127	3BS7-0.75-0.78	ISBP	rouge de crésol
cfp3112	3BS9-0.57-0.75	ISBP	rouge de crésol
cfp1773	3BS9-0.57-0.75	ISBP	rouge de crésol
wmc231	3BS9-0.57-0.75	SSR	M13
gwm566	3BS9-0.57-0.75	SSR	M13
barc68	3BS9-0.57-0.75	SSR	M13
cfp1329	3BS9-0.57-0.75	ISBP	rouge de crésol
cfb3260	3BS9-0.57-0.75	SSR	M13
dw724	3BS9-0.57-0.75	SSR	M13
wmc777	3BS9-0.57-0.75	SSR	M13
wmc505	3BS2-0.56-0.57	SSR	M13
cfp3275	3BS2-0.56-0.57	ISBP	SYBR green
wmc612	3BS1-0.33-0.55	SSR	M13
cfb3266	3BS1-0.33-0.55	SSR	M13
wmc762	3BS1-0.33-0.55	SSR	M13
cfp5129	3BS1-0.33-0.55	ISBP	rouge de crésol
cfp5109	3BS1-0.33-0.55	ISBP	rouge de crésol
wmc625	3BS5-0.07-0.33	SSR	M13
gwm285	3BS5-0.07-0.33	SSR	M13
wmc366	3BS5-0.07-0.33	SSR	M13
barc1111	C-3BS5-0.07	SSR	M13
barc139	C-3BS5-0.07	SSR	M13
gpw1145	C-3BS5-0.07	SSR	M13
nw2715	C-3BS5-0.07	SSR	M13
wmm805	C-3BL2-0.22	SSR	rouge de crésol
wmm1007	C-3BL2-0.22	SSR	M13
cfb3409	C-3BL2-0.22	SSR	M13
cfp1417	C-3BL2-0.22	ISBP	rouge de crésol
cfp161	C-3BL2-0.22	ISBP	rouge de crésol
cfb3135	3BL1-0.31-0.38	SSR	M13
wmm723	3BL1-0.31-0.38	SSR	M13
dw1221	3BL1-0.31-0.38	SSR	M13
dw1754	3BL1-0.31-0.38	SSR	M13

dw955	3BL9-0.38-0.50	SSR	M13
cfb3374	3BL9-0.38-0.50	SSR	M13
cfp1632	3BL10-0.50-0.63	ISBP	SYBR green
nw1666	3BL10-0.50-0.63	SSR	rouge de crésol
wmc418	3BL10-0.50-0.63	SSR	M13
gpw4034	3BL10-0.50-0.63	SSR	M13
cfp100	3BL7-0.63-1.00	ISBP	SYBR green
nw2706	3BL7-0.63-1.00	SSR	M13
wmm1703	3BL7-0.63-1.00	SSR	M13
cfb3059	3BL7-0.63-1.00	SSR	M13
wmm1385	3BL7-0.63-1.00	SSR	M13
cfb59	3BL7-0.63-1.00	SSR	rouge de crésol
cfp1343	3BL7-0.63-1.00	ISBP	rouge de crésol
wmc291	3BL7-0.63-1.00	SSR	M13
cfb3440	3BL7-0.63-1.00	SSR	M13
wmm1133	3BL7-0.63-1.00	SSR	rouge de crésol
wmm702	3BL7-0.63-1.00	SSR	M13
cfp39	3BL7-0.63-1.00	ISBP	rouge de crésol
wmm280	3BL7-0.63-1.00	SSR	M13
wmm278	3BL7-0.63-1.00	SSR	M13
cfp3041	3BL7-0.63-1.00	ISBP	rouge de crésol
nw1608	3BL7-0.63-1.00	SSR	M13
wmc326	3BL7-0.63-1.00	SSR	rouge de crésol
wmm430	3BL7-0.63-1.00	SSR	M13
cfp24	3BL7-0.63-1.00	ISBP	rouge de crésol
cfb3200	3BL7-0.63-1.00	SSR	M13
cfb22	3BL7-0.63-1.00	SSR	rouge de crésol
barc77	3BL7-0.63-1.00	SSR	M13
wmm1966	3BL7-0.63-1.00	SSR	M13
wmm274	3BL7-0.63-1.00	SSR	M13
wmm448	3BL7-0.63-1.00	SSR	M13
gpw3233	3BL7-0.63-1.00	SSR	rouge de crésol
gwm247	3BL7-0.63-1.00	SSR	M13
gpw4513	3BL7-0.63-1.00	SSR	M13
gpw8064	3BL7-0.63-1.00	SSR	M13
gpw4505	3BL7-0.63-1.00	SSR	M13
nw2860	3BL7-0.63-1.00	SSR	M13
gwm108	3BL7-0.63-1.00	SSR	M13

wmm638	3BL7-0.63-1.00	SSR	M13
gpw3085	3BL7-0.63-1.00	SSR	M13
gpw4143	3BL7-0.63-1.00	SSR	M13
gwm299	3BL7-0.63-1.00	SSR	M13
gwm340	3BL7-0.63-1.00	SSR	M13

Tableau 14 Marqueurs et conditions d'amplification utilisés pour l'étude chromosome entier.

[a]/ Le protocole détaillé de chaque condition d'amplification est décrit dans la partie Matériels et Méthodes paragraphe 3.2 p 139.

Amorces utilisées pour la technique AS-PCR	**Séquences**
cfp5026_Aspcr121Cs	CCTTCAAGCTAACCCGATCG
cfp5026_Aspcr121Re	CCTTCAAGCTAACCCGATCA
cfp5046_Aspcr40Cs	TAGGTTGCTCACTTAGCTGC
cfp5046_Aspcr40Re	TAGGTTGCCCACTTAGCTGA
cfp5061_Aspcr114Re	GCCACTCTTAGAATTAACTACTG
cfp5061_Aspcr114Cs	GCCACTCTTAGAATTAACTACTC
Amorces utilisées pour la technique SNaPshot	**Séquences**
cfb6103_SNaPF	GACCTTGTAATTGATCAATAAATTGGAG
cfb6040_SNaPF	TCGTCGTCGTCGGGCGCGTG
cfb6116_SNaPR	ACGCGCGTGAATGGTCCGTC
cfb6127_SNaPR	CAGTACACGGAGCAAAAAAT
cfp5079_SNaP111F	ATCGTTTATTATCCATGCTATAATTGTATT
Hga3Ex5SNP1-SNaP-F	CTACGATTCCTGCTCCTGTG
Amorces utilisées pour le génotypage des insertions/délétions	**Séquences**
Hga1PromdelF1	GCAAAAGGTGCCACCTGG
Hga1PromdelR1	CGAAGCCGAGGTTGAGGTA
Hga3int1delR1	GCTGACTGCACATATACACGCATC
Hga3int1delF1	AGACAGAGCAAGGGGGGC
Hga3Ex6delF1	GTGGCGCATCAGGCCGTA
Hga3Ex6delR1	GTAGACGCCGCGCATGTGT

Tableau 15 Séquences des amorces développées pour révéler le polymorphisme de type SNP ou insertion/délétion au niveau du locus Rph7.

ddl	α = 0.05	α = 0.01		ddl	α = 0.05	α = 0.01
1	0.997	1.000		26	0.374	0.479
2	0.950	0.990		27	0.367	0.471
3	0.878	0.959		28	0.361	0.463
4	0.811	0.917		29	0.355	0.456
5	0.755	0.875		30	0.349	0.449
6	0.707	0.834		31	0.344	0.442
7	0.666	0.798		32	0.339	0.436
8	0.632	0.765		33	0.334	0.430
9	0.602	0.735		34	0.329	0.424
10	0.576	0.708		35	0.325	0.418
11	0.553	0.684		36	0.320	0.413
12	0.532	0.611		37	0.316	0.408
13	0.514	0.641		38	0.312	0.403
14	0.497	0.623		39	0.308	0.398
15	0.482	0.606		40	0.304	0.393
16	0.468	0.590		41	0.301	0.389
17	0.456	0.575		42	0.297	0.384
18	0.444	0.561		43	0.294	0.380
19	0.433	0.549		44	0.291	0.376
20	0.423	0.537		45	0.288	0.372
21	0.413	0.526		46	0.285	0.368
22	0.404	0.515		47	0.282	0.365
23	0.396	0.506		48	0.279	0.361
24	0.388	0.496		49	0.276	0.358
25	0.381	0.487		50	0.273	0.354

Tableau 16 Table de Bravais-Pearson.

ddl = degré de liberté.

Protocole 1 Marquage des bandes C.

Matériels nécessaires

Azote liquide, Bain Marie à 60°C, Boîtes de Pétri de diamètre 55 mm ou 90 mm selon le nombre de grains, Bouchons en liège pour tubes 20 ml, Cuves en verre pour 8 lames ou Cuve Hellendahl 16 lames, Filtres Whatman de 90 mm (Whatman réf n°1001090) ou 55 mm (Whatman réf n°1001055) pour le fond des boîtes de Pétri, Glace pilée, Carroussel de 6 tubes 1,5 ml, Minitubes en verre de 1 ml, Seringues , Huile à immersion, Lamelles 18x18 mm, Lames de rasoir (Gilette super Silver réf 356 www.coop-labo.com), Lames porte objet 76x26 mm Superfrost lavées dégraissées à bords rodés (VWR réf 80100505), Logiciel Adobe Photoshop, Microscope avec contraste de phase équipé éventuellement d'une caméra et logiciel pour prise d'images, Minuteur, Portoirs 14 ou 16 tubes en bois ou plastique (fabrication maison) adaptés aux tubes pour manipulations et stockage au frigo, Pinces Dumont modèle suisse n°5 (Moria), Réfrigérateur, Support pour séchage de 40 lames et Tubes verre de 20ml.

Réactifs (chimiques et biologiques)

Acide acétique 45 % (eau acétique)
Acide chlorhydrique 1 N et 0,2 N
Alcool 100 %, 95 %, 70 %
Baryte ($Ba(OH)_2$, 8 H_2O Merck réf 1.01735.1000) : environ 70 g de baryte dans 1000 ml d'eau ultra pure. Laisser pendant 30 min sous agitation et laisser reposer avant utilisation.
Eau ultra pure (H2Oup)
Entellan (Merck réf 1.07960.0500)
Giemsa (Merck réf 1.09204.0500)
2×SSC : 17.53 g NaCl + 8.82 g Na Citrate (2 H_2O), ajuster pH à 7.0 avec HCl, qsp 1000 ml eau ultra pure
TrisHCl 0.125 M pH=6.8: 14.6 g de Trizma Base qsp 1000ml eau ultra pure, ajuster au pH=6.8 avec HCl
Cellulysine (CALBIOCHEM réf 219466, 11200 U/g) : 0.04 g de cellulysine + 10 ml H2Oup. Ajuster le pH à 4

Préparation des racines

Le jeudi matin, disposer les grains sur une épaisseur de deux filtres humidifiés avec de l'eau ultra pure dans une boite de Pétri, les laisser s'imbiber 24 h à température ambiante, (si les grains sont très infectés : les tremper 10 min dans une solution de $KMnO_4$ et rincer à l'eau du robinet).
Le vendredi, la germination étant amorcée, les boîtes sont mises au frigo à 5°C jusqu'au lundi.
Dès le lundi matin, les boîtes sont remises à température ambiante, le mardi ou le mercredi quand les racines atteignent 2 à 3 cm, les couper et les mettre dans l'eau glacée pendant 24 h (tubes dans de la glace pilée à l'intérieur d'un frigo).
Tremper le carroussel dans l'eau acétique, mettre dans les tubes les racines et étiquettes d'identification et laisser les pendant 15 min.
Changez l'eau acétique et mettre le tout à 4°C pendant 3-4 heures.

Rincer 3 fois dans H2Oup à température ambiante puis à nouveau 6 fois 10 min. Entre temps, préchauffer HCl 0.2 N à 60°C et prévoir également HCl 0.2 N à 4°C.
Mettre le carroussel contenant les racines dans HCl 0.2 N à 4°C pendant 15 min, puis dans du HCl 0.2 N à 60°C pendant 5 min. Entre temps, chercher de l'eau glacée et transférer le carroussel dans l'eau glacée pendant 5 min.
Rincer à nouveau 3 fois 5 min dans H2Oup.
Préparer la cellulysine (bien vérifier le pH=4).
Remplir les minis-tubes en verre de cellulysine, couper les méristèmes et les inclure dans les minis-tubes. Laisser les la nuit à température ambiante.

Montage des lames

Vider la cellulysine à l'aide de la seringue et rincer les méristèmes pendant 3 fois 5 min dans H2Oup.
Récupérer les méristèmes (3 maxi) à l'aide de la seringue et les disposer sur la lame. Enlever délicatement par aspiration l'eau sur la lame avec la seringue.
Mettre une goutte d'eau acétique et laisser 1 min 30 s précisément en contact avec les méristèmes.
Aspirer l'eau acétique et remettre une goutte à nouveau.
« Hacher » finement les méristèmes à l'aide d'une lame de rasoir.
Poser délicatement une lamelle 18x18 mm et tapoter légèrement pour éliminer les bulles d'air.
Faire une pression bien verticale à l'aide d'un buvard sans faire bouger la lamelle et observer au microscope en contraste de phase pour apprécier la qualité d'étalement.
Plonger quelques secondes la préparation dans l'azote liquide à l'aide d'une grosse pince en bois puis faire sauter la lamelle avec une lame de rasoir.
Mettre la lame dans l'alcool 100 % pendant au moins une nuit.

Dénaturation et hydrolyse

Mettre la lame 6 min dans une solution saturée de baryte à température ambiante puis 15 s dans HCl 0,2 N à température ambiante.
Rincer sous l'eau du robinet très brièvement puis laisser dans H2Oup.

Renaturation

Mettre la lame 60 min dans 2xSSC à 60°C et rincer sous un filet d'eau de robinet pendant 20 min puis tremper 3 fois 5 min dans H2Oup.

Coloration

Utiliser une cuve à coloration à 8 places.
Préparer pendant le rinçage la solution de coloration sous la hotte : pour une cuve de 50 ml : TrisHCl 0.125 M pH=6.8 (bien vérifier le pH) + 1.5 ml de colorant Giemsa.
Immerger les lames.

Contrôler au bout de quelques minutes l'évolution de la coloration sous un microscope et arrêter quand le contraste coloration des bandes et corps des chromosomes est optimal.
Arrêter au plus tard à 45 mn.
Laver rapide dans de l'eau ultra pure (quelques secondes).
Sécher à l'air verticalement.

Montage permanent : dans une goutte d'Entellan, sécher pendant 3 heures minimum (ou 1 heure sous la hotte), conserver à l'obscurité dans boîtes à lames.

Examen au microscope : avec objectif x10 dans un premier temps puis x100 à immersion avec ou sans contraste de phase et réalisation de photos éventuellement.

Bins de délétion	3BS3-0.87-1.00	3BS8-0.78-0.87	3BS7-0.75-0.78	3BS9-0.57-0.75	3BS2-0.56-0.57	3BS4-0.55-0.56	3BS1-0.33-0.55	3BS5-0.07-0.33	C-3BS5-0.07	C-3BL2-0.022	3BL2-0.22-0.28	3BL8-0.28-0.31	3BL1-0.31-0.38	3BL9-0.38-0.50	3BL10-0.50-0.63	3BL7-0.63-1.00	Chr.*
Nombre de marqueurs	5	27	4	9	0	2	5	3	4	5	0	0	4	2	4	28	102
Taille des bins (Mb)	56	39	13	78	4	4	95	108	34	124	33	17	39	67	73	208	992
Pourcentage de la taille physique (bras)	12,93	9,01	3,00	18,01	0,92	0,92	21,94	24,94	7,85	22,06	5,87	3,02	6,94	11,92	12,99	37,01	
Pourcentage de la taille physique (chromosome)	5,65	3,93	1,31	7,86	0,40	0,40	9,58	10,89	3,43	12,50	3,33	1,71	3,93	6,75	7,36	20,97	
Distances génétiques (F2CsRe en cM)	5,20	33,05	4,85	13,35	-	0,65	1,55	0,85	0,00	1,55			6,20	4,15	3,45	104,15	179
Pourcentage de recombinaison (bras)	8,74	55,55	8,15	22,44	-	1,09	2,61	1,43	0,00	1,30			5,19	3,47	2,89	87,15	
Pourcentage de recombinaison (chromosome)	2,91	18,46	2,71	7,46	-	0,36	0,87	0,47	0,00	0,87			3,46	2,32	1,93	58,18	
Marqueurs/Mb/bin	11,30	1,44	3,25	8,67	-	2,00	19,00	36,00	8,50	34,80			9,75	33,50	18,25	7,43	9,73
Fréquence de recombinaison (cM/Mb)	0,09	0,85	0,37	0,17	-	0,08	0,02	0,01	0,00	0,01			0,16	0,06	0,05	0,50	0,18

Tableau 17 Analyse détaillée de la comparaison entre la carte physique de délétion du chromosome 3B et la carte génétique issue de la population F2CsRe.

* Chr. Chromosome entier.

Intervalles	***r* obtenus sur CsCtA**	***r* obtenus sur CsCtG**	***p*-value**
gwm389-nw1600	0.019	0.035	0.69
nw1600-nw1821	0.012	0.018	0.88
nw1821-nw1644	0.006	0.017	0.65
nw1644-barc147	0.019	0	0.14
barc147-gwm493	0.026	0.013	0.68
gwm493-cfd79	0.030	0.073	0.21
cfd79-nw216	0.208	0.247	0.69
nw216-gpw3092	0.026	0.044	0.68
gpw3092-gwm285	0.253	0.186	0.38
gwm285-barc164	0.165	0.097	0.19
barc164-gpw4034	0.006	0.012	0.80
gpw4034-nw2706	0.082	0.023	**0.04**
nw2706-gpw8064	0.081	0.040	0.27
gpw8064-gpw4505	0.032	0.006	0.24
gpw4505-nw2860	0.019	0	0.17
nw2860-wmc291	0.114	0.053	0.12
wmc291-gwm108	0.012	0.022	0.80
gwm108-wmm1133	0.055	0.084	0.56
wmm1133-wmm702	0.006	0.103	$\mathbf{10^{-4}}$
wmm702-wmm280	0.018	0.073	**0.05**
wmm280-wmm638	0.138	0.107	0.70
wmm638-nw1608	0	0.012	0.46
nw1608-wmc326	0.076	0.097	0.79
wmc326-gpw3085	0.038	0.038	0.97
gpw3085-barc77	0.019	0.037	0.67
barc77-wmm1966	0.006	0.019	0.56
wmm1966-gpw4143	0.006	0.012	0.89
gpw4143-wmm274	0.123	0.055	0.10
wmm274-gwm299	0.062	0.102	0.45
gwm299-gwm247	0.189	0.241	0.51
gwm247-gwm340	0	0	-

Tableau 18 Différences du taux de CO local entre la méiose mâle et la méiose femelle.

r représente les valeurs du taux de CO obtenues à partir des populations haploïdes doublés Chinese spring x Courtot androgenèse (A) et gynogenèse (G). Ces valeurs ont été calculées en utilisant les distances en cM obtenues par le logiciel Mapmaker et en utilisant la fonction inverse de Kosambi (1944). Les *p*-values ont été calculées comme indiqué dans la partie Matériels et Méthodes (paragraphe 6.3 p 147). Les *p*-values significatives sont indiquées en gras.

IX Publication

Copyright © 2009 by the Genetics Society of America
DOI: 10.1534/genetics.108.097469

Detailed Recombination Studies Along Chromosome 3B Provide New Insights on Crossover Distribution in Wheat (*Triticum aestivum* L.)

Cyrille Saintenac,* Matthieu Falque,† Olivier C. Martin,‡,§ Etienne Paux,* Catherine Feuillet* and Pierre Sourdille*,[1]

*INRA-UBP, UMR 1095, Genetic, Diversity and Ecophysiology of Cereals, 63100 Clermont-Ferrand, France, †INRA and ‡Université Paris-Sud, UMR 0320/UMR 8120 Génétique Végétale, 91190 Gif-sur-Yvette, France and §Université Paris-Sud, UMR 8626 LPTMS, 91400 Orsay, France

Manuscript received October 13, 2008
Accepted for publication December 8, 2008

ABSTRACT

In wheat (*Triticum aestivum* L.), the crossover (CO) frequency increases gradually from the centromeres to the telomeres. However, little is known about the factors affecting both the distribution and the intensity of recombination along this gradient. To investigate this, we studied in detail the pattern of CO along chromosome 3B of bread wheat. A dense reference genetic map comprising 102 markers homogeneously distributed along the chromosome was compared to a physical deletion map. Most of the COs (90%) occurred in the distal subtelomeric regions that represent 40% of the chromosome. About 27% of the proximal regions surrounding the centromere showed a very weak CO frequency with only three COs found in the 752 gametes studied. Moreover, we observed a clear decrease of CO frequency on the distal region of the short arm. Finally, the intensity of interference was assessed for the first time in wheat using a Gamma model. The results showed m values of 1.2 for male recombination and 3.5 for female recombination, suggesting positive interference along wheat chromosome 3B.

IN most eukaryotes, meiotic recombination plays a key role in ensuring correct segregation of the chromosomes during gamete formation and for generating diversity in the progeny. Comparative analyses between a wide range of species including human, mouse, yeast, and Arabidopsis suggested a conserved process for recombination starting with the formation of a double-strand break (DSB) and proceeding through either crossover (CO) or non-CO pathways (for review see Anderson and Stack 2002; Hamant *et al.* 2006; Baudat and De Massy 2007; Cromie and Smith 2007). At least one CO per bivalent is required because the obligate chiasma constraint ensures a good segregation of chromosomes at meiosis. Additional COs are subjected to interference that inhibits the formation of a CO in the vicinity of another one (Jones and Franklin 2006). The number of COs on a chromosome is thus highly controlled, presumably because of the deleterious effect of having too many or too few COs (Jones and Franklin 2006). Their distribution can be analyzed by comparing physical and genetic maps or by studying linkage disequilibrium patterns. Studies performed in human, mouse, Arabidopsis, and rice (Wu *et al.* 2003; Drouaud *et al.* 2006; Arnheim *et al.* 2007) provided insight into some of the major features of recombination distribution with the identification of hot and cold regions of recombination along the chromosomes as well as a general suppression of COs in centromeric and pericentromeric regions (for a review see Jones 1987). In some plants with large genomes such as maize (2.5 Gb), barley (5 Gb), and wheat (17 Gb), recombination was shown to follow particular patterns with a gradual increase from the centromeres to the telomeres (Lukaszewski and Curtis 1993; Kunzel *et al.* 2000; Tenaillon *et al.* 2001; Anderson *et al.* 2003).

To date, no consistent relationship has been established between recombination rate and sequence composition, *i.e.*, gene or repetitive DNA content. This question was addressed in humans (Kong *et al.* 2002), where recombination rates were found to be significantly correlated with both cytogenetic structures (staining intensity of G bands) and sequence composition [GC content, CpG motifs, and poly(A)/poly(T) stretches]. In rice, hotspots are generally associated with high gene density (Wu *et al.* 2003) while in maize, Anderson *et al.* (2006) have shown that recombination is localized to the distal part of the chromosomes and is correlated strongly with the distribution of expressed sequenced tags (ESTs). Further, a recombination hotspot has been identified in a highly gene-rich region at the maize *bz* locus and characterized at a fine scale (Fu *et al.* 2001). In contrast, in Arabidopsis (Drouaud *et al.* 2006) fine-scale analysis failed to establish a correlation between recombination and gene density.

Previous studies in wheat indicated that recombination occurs mainly in the telomeric regions of the

[1] *Corresponding author:* UMR 1095, Genetic, Diversity and Ecophysiology of Cereals, Domaine de Crouël, 234 Ave. du Brézet, 63100 Clermont-Ferrand, France. E-mail: pierre.sourdille@clermont.inra.fr

chromosomes with a gradient from the centromeres to the telomeres (Lukaszewski and Curtis 1993; Erayman *et al.* 2004; Sidhu and Gill 2004; See *et al.* 2006). The recombination gradient was proposed to be correlated with gene content. For example, Erayman *et al.* (2004) suggested that 95% of the recombination occurs in 48 gene-rich regions covering 29% of the physical size of the wheat genome. A striking example was observed for a gene-rich region encompassing only 4% of the long arm of chromosome 5B yet explaining at least 30% of the recombination of the entire chromosome (Faris *et al.* 2000). In addition, high recombination rates were found to be related to the decrease of synteny either between homeologous chromosomes (Akhunov *et al.* 2003a) or between wheat and rice orthologous chromosomes (Akhunov *et al.* 2003b). Chromosomal rearrangements such as deletions and duplications were also found to be more frequently located in highly recombinant regions (Dvorak *et al.* 2004; Dvorak and Akhunov 2005). To date, very little is known about the factors underlying the gradient of recombination along the wheat chromosomes as most of the studies were performed at low resolution with a limited number of markers and in the absence of physical maps (fingerprinted BAC libraries organized into genetically anchored contigs) and of a genome sequence for wheat.

In this work, we have used the largest hexaploid wheat chromosome (3B, 995 Mb) as a model to perform detailed recombination studies and identify factors that govern recombination in wheat. By comparing genetic and physical deletion maps, we established a fine-scale profile of CO distribution along the chromosome and confirmed that recombination is restricted mainly to distal regions with a decrease in the very telomeric regions. Comparison of male and female CO rates indicated an effect of both male and female meiosis on local CO variations. Finally, using the Gamma model, we provide the first evidence and characterization of positive interference on a wheat chromosome and discuss the implications of these finding for genetic mapping in wheat.

MATERIALS AND METHODS

Plant material and molecular analyses: *Genetic mapping populations:* The wheat cultivars Chinese Spring (Cs) and Renan (Re) were crossed and grains of the F_1 individuals were dissected 20 days after pollination. Each embryo was cultivated on a rich medium (Muller) and grown in a growth chamber until the first two leaves emerged. Plants were then potted and grown for 2 weeks in cool conditions (17° day and 12° night, 12 hr of day) and then for 2 months at 24° day and 18° night with 16 hr of day. F_1 plants were subsequently selfed and 376 seeds were sown to obtain the CsRe F_2 population. This population was then used to study CO distribution along chromosome 3B.

Interference and sex-specific recombination were studied using two different doubled-haploid (DH) populations originating from the cross between Cs and Courtot (Ct). The androgenetic (CsCt-A) population is described in Felix *et al.* (1996) and a subset of 182 lines was used in the present study. For the gynogenetic population (CsCt-G), Chinese Spring and Courtot were crossed and F_1 individuals were grown in a greenhouse. The spikes were emasculated and pollinated 1–3 days after with maize pollen (Laurie and Bennett 1988). Haploid seeds were then dissected and embryos were rescued and grown on rich medium until doubling with colchicine treatment. A set of 185 DH lines was used for the study.

Aneuploïd lines: To assign markers to chromosome 3B, we used the nullisomic-tetrasomic line N3BT3A, two ditelocentric lines (Dt3BL and Dt3BS), and 14 chromosome 3B deletion lines described earlier (Sears 1954, 1966; Sears and Sears 1978; Endo and Gill 1996). Each deletion line is defined by the fraction length of the arm that remains present on the deleted chromosome (for example, line 3BS4-0.55 conserves the proximal 55% and lacks the distal 45% of the short arm of chromosome 3B). These deletion lines define 16 deletion bins: 3BL7-0.63-1.00, 3BL10-0.50-0.63, 3BL9-0.38-0.50, 3BL1-0.31-0.38, 3BL8-0.28-0.31, 3BL2-0.22-0.28, C-3BL2-0.022, C-3BS5-0.07, 3BS5-0.07-0.33, 3BS1-0.33-0.55, 3BS4-0.55-0.56, 3BS2-0.56-0.57, 3BS9-0.57-0.75, 3BS7-0.75-0.78, 3BS8-0.78-0.87, and 3BS3-0.87-1.00. Each deletion bin is defined by the fraction lengths of two successive deletion lines (for example, the deletion bin 3BS4-0.55-0.56 is the chromosome fragment covering the region between 55 and 56% of the short arm of chromosome 3B and delimited by deletion lines 3BS4-0.55 and 3BS2-0.56). The estimated sizes of the bins (in megabases) are indicated in Figure 1.

DNA extraction: DNA was extracted using a sodium metabisulfite protocol. Briefly, 100 mg of fresh tissue were ground in fine powder in liquid nitrogen. Then 600 μl of extraction buffer (0.1 M Tris, 0.05 M EDTA, 0.5 M NaCl, 0.02 M sodium metabisulfite, 0.2 M RNAse) were added, mixed, and incubated for 45 min at 95°. After centrifugation, 430 μl of the supernatant were collected and added to 425 μl of 7.5 M isopropanol/ammonium acetate (16/1). After centrifugation DNA pellets were washed in 70% ethanol and resuspended in 100 μl of 0.1 M Tris-EDTA buffer.

Markers and PCR conditions: A total of 102 molecular markers (76 SSRs and 26 insertion site-based polymorphisms, ISBPs) for the CsRe population and 32 SSR markers for the two CsCt populations were selected for linkage analysis on the basis of their distribution in the different deletion bins (Table 1).

Ninety SSRs originating from different sources were amplified using either the M13 protocol (Nicot *et al.* 2004) or the cresol red protocol (Hodges *et al.* 1997). The 26 ISBP (Paux *et al.* 2006) markers were amplified using either the SYBR GREEN protocol or the cresol red protocol (supplemental Table S1).

For the M13 protocol, PCR reactions were carried out in 6.5 μl of 0.2 units Taq, 50 nM forward-M13-tailed primer, 500 nM reverse primer, 3.4 pmol labeled M13 primer, and 5.4 nmol dNTP. Amplification was done following a touchdown procedure (Paux *et al.* 2006). Amplification products were visualized using an ABI PRISM 3100 Genetic Analyzer (Applied Biosystems, Foster City, CA).

For the SYBR GREEN protocol, PCR reactions were carried out in 10 μl of 0.25 units Taq, 0.5 μM for each primer, 1× SYBR Green, 0.4 mM for each dNTP, and 1 M betaine. Amplification was done with the same touchdown program as for condition 1 (Paux *et al.* 2006). Visualization of PCR products was done using a quantitative PCR machine (ABI7900HT, Applied Biosystems) by analyzing dissociation curves with SDS 2.2.1 software.

For the Cresol red protocol, PCR reactions were carried out in 10 μl of 0.25 units of Taq, 0.5 μM of each primer, 0.4 mM of each dNTP, and 1 μM cresol red. Amplification was done using a classical PCR protocol (94° for 5 min; 30 cycles of 94° for

30 sec, 60° for 30 sec, and 72° for 30 sec; and a final step of 72° for 5 min). PCR products were directly visualized on a 4% agarose gel stained with Ethidium Bromide.

Statistical analyses: Algorithmic procedures for statistical analyses are presented in supplemental material 1.

Linkage analysis: Genetic maps were constructed with the Mapmaker software (Lander *et al.* 1987). Linkage groups were obtained using default parameters. The "order" command was used to order the loci and the "map" function was further used to generate the genetic map, using the Kosambi (1944) mapping function. When markers were not placed on the map at LOD < 2, the function "compare" was used and the best order was retained. The "error detection" and "genotypes" commands permitted us to identify potential genotyping errors. The individuals concerned were scored again to check and eliminate these potential genotyping errors.

Comparison of CO frequency: The significance of the differences interval-by-interval between male and female genetic maps was tested using a *P*-value approach. We determined the probability of having recombination rate differences at least as large as the values we found in our data sets under the hypothesis H_0 that the recombination rate is sex independent (supplemental material 1). For multiple testing, we used Bonferroni correction (Hochberg 1988).

Test of the no-interference hypothesis: The no-interference hypothesis H_0 was tested via the statistics of distances between recombination events. In this procedure, the (theoretically derived) histogram of these distances was established and tested to determine whether it is compatible with the experimental histogram, from which it is possible to derive a *P*-value (supplemental material 1).

Coefficient of coincidence: Interference strength was estimated using the coefficient of coincidence C_4 that is the frequency of observed double COs divided by their theoretical frequency in the absence of interference. For this measure, we consider two nonoverlapping intervals associated with locus pairs (A, B) and (C, D). If r_{AB} and r_{CD} are the recombination rates for intervals (A, B) and (C, D), the coefficient of coincidence is

$$C_4(AB, CD) = \frac{f(AB, CD)}{r_{AB} r_{CD}},$$

where $f(AB, CD)$ is the frequency of double COs. By construction and in absence of interference, $C_4 = 1$ while interference effects are strong when C_4 is close to 0. To display the putative distance dependence of C_4, we defined the distance d_4, as the distance between the two midpoints of the intervals. Since small intervals may result in large statistical errors in the estimations of r_{AB} or r_{CD}, leading to erroneous C_4 values, we did not consider intervals for which $r < 0.05$. From the list of C_4 values, we performed a binning on d_4 and computed the mean C_4 in each of the 30 bins. The error bars on these mean values are given by the standard deviation divided by the square root of the number of C_4 values contributing to each bin.

Gamma model: The Gamma model (McPeek and Speed 1995; Broman and Weber 2000) can be considered as an extension of the counting model (Stahl *et al.* 2004), allowing for noninteger values of m (supplemental material 1). The advantage over the counting model is that m is continuous, allowing for a broader range of interference effects. When $m = 0$, the model reduces to Haldane's (1919) model of no interference. For each m value (taken in practice in steps of 0.1 or 0.2), we simulated the statistical properties of recombination events and compared the simulated values to the experimental ones. The simulations produced 10^7 gametes using the Gamma model. We constructed the histogram of the distances between recombinant intervals, restricted to gametes where there are exactly two recombinant intervals (A, B) and (C, D), where A and B are adjacent, just as C and D are. Letting this histogram of binned distances have the frequencies $f_1, f_2, f_3, \ldots$, we can then associate the likelihood f_j to a DH plant that has two recombinant events at a distance falling in the bin j. The LOD score for a given m is then the sum of the logarithm of each plant's likelihood. The best fit is obtained for the value of m that maximizes this score (see supplemental material 1).

Confidence interval for the estimate of m: Let m be the interference parameter obtained by fitting to the experimental data and N be the number of experimental DH plants (182 plants for male meiosis and 185 for female meiosis). We generated 10^5 simulated data sets, each one producing N gametes via the Gamma model of parameter m. The fitting was done for each data set, *i.e.*, for the current set of N simulated crosses. The best adjusted interference parameter m_e was found through scanning. By simulations, 10^5 realizations of this random variable m were generated. The confidence interval on m was then given by the two values m_- and m_+ such that the probability of finding $m_e < m_-$ or $m_e > m_+$ is 0.025. Significance of male–female differences in interference strength was tested through the H_0 hypothesis (the same interference parameter describes the male and female data sets) by the bootstrap method based on data reshuffling (supplemental material 1).

RESULTS

CO distribution along chromosome 3B: To study the distribution of COs in wheat, comparisons were made between genetic and physical maps of the largest (995 Mb) hexaploid wheat chromosome (3B). Since no particular CO distribution has been observed previously for chromosome 3B compared to other wheat chromosomes (Lukaszewski and Curtis 1993; Akhunov *et al.* 2003b), this chromosome provides a good model to study CO distribution in wheat.

Linkage analysis using an F_2 Chinese Spring × Renan population with 102 selected markers resulted in a genetic map of 179.1 cM in length (Figure 1), which is in the range of the genetic map length reported for chromosome 3B in the literature (Quarrie *et al.* 2005; Suenaga *et al.* 2005; Akbari *et al.* 2006). This represented an average of 1 marker for every 1.75 cM. On this genetic map, a strong concentration of markers was found in the proximal part of the chromosome and a few gaps remained at the distal end of the long arm, in particular between loci *Xgpw3233* and *Xgwm247*, where a gap of 30 cM was observed (Figure 1). No segregation distortion was observed among the 102 loci.

Physical mapping was performed using aneuploid stocks of chromosome 3B developed from cv. Chinese Spring by Sears (1954, 1966; Sears and Sears 1978) and Endo and Gill (1996). The chromosome 3B deletion map consisted of 16 deletion bins, ranging from 4 Mb (3BS4-0.55–0.56 and 3BS2-0.56–0.57) to 208 Mb (3BL7-0.63–1.00) with an average size of 62 Mb (Figure 1, Table 1). The density of markers per bin ranged from one marker per 1.4 Mb in bin 3BS8-0.78–0.87 to one marker per 36 Mb in bin 3BS5-0.07-0.33, with an average of one

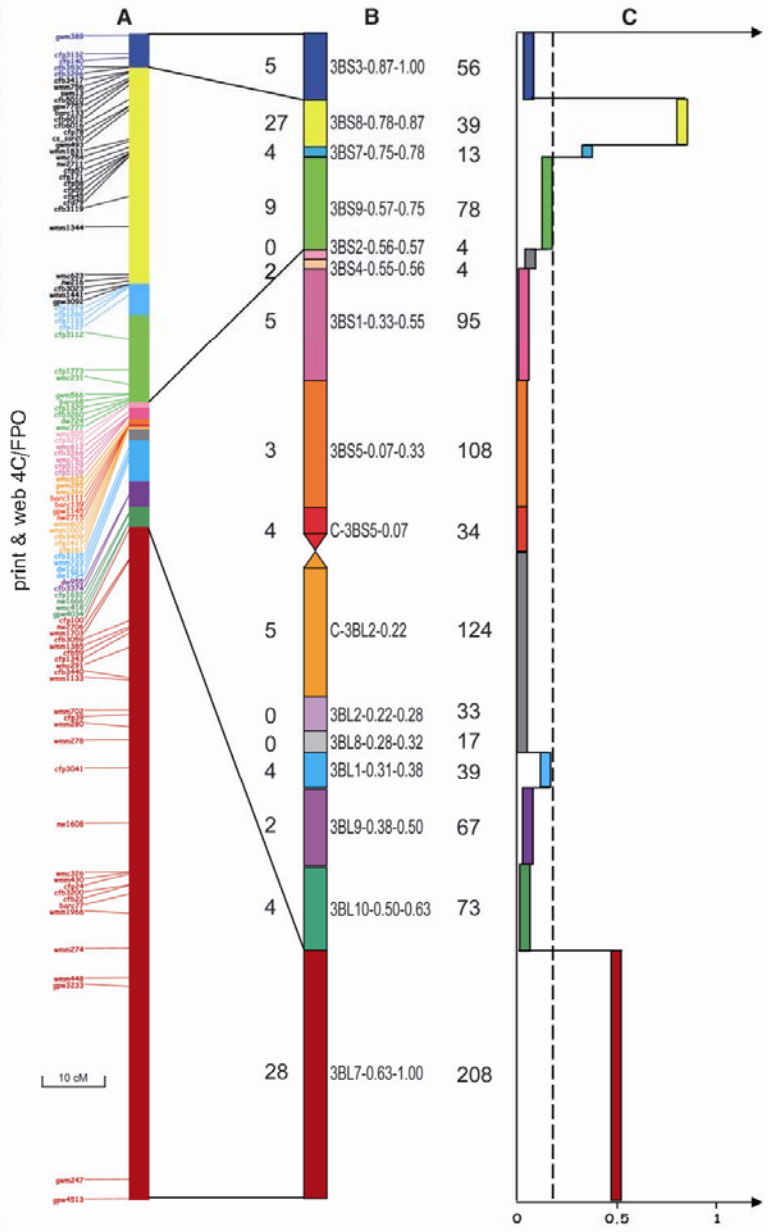


FIGURE 1.—Distribution of COs along chromosome 3B. (A) Genetic map of chromosome 3B. Markers are indicated with the same color code as for the deletion bins. (B) Deletion map of chromosome 3B. The deletion bins used to assign markers are represented with the number of markers in each bin (left of the chromosome) and the size of the bin in megabases (right). (C) Schematic representation of the CO frequency along chromosome 3B. CO frequency calculated in centimorgans per megabase is represented by a colored bar for each bin. The two gray bars indicate bins with no markers that were grouped together. The plotted lines represent the average CO frequency for the chromosome.

marker per 9.7 Mb. Very good correlation was found between the physical and genetic maps as all markers assigned to the same deletion bin were linked on the genetic map and the bin order was conserved on the genetic map.

CO frequency (centimorgans per megabase) was calculated for each bin (Figure 1, Table 1) and ranged from 0 in the centromeric region (C-3BS5-0.20, C-3BL2-0.22) to 0.85 in the terminal bin of the short arm (3BS8-0.78–0.87) with an average of 0.18 cM/Mb for the entire chromosome 3B. Crossovers occurred mostly in the distal 40% of the chromosome arms, with the two distal bins 3BS8-0.78–0.87 and 3BL7-0.63–1.00 corresponding to 25% of the physical size of the chromosome accounting for 77% of the COs. There were no COs in the most proximal centromeric regions (bins C-3BL2-0.22 and C-3BS5-0.07) and only three were observed in bin 3BS5-0.07–0.30 (Table 1, Figure 1). Interestingly, slight differ-

TABLE 1

Detailed analysis of the comparison between the chromosome 3B cytogenetic map and the F_2 CsRe genetic map

Deletion bins	3BS3-0.87-1.00	3BS8-0.78-0.87	3BS7-0.75-0.78	3BS9-0.57-0.75	3BS2-0.56-0.57	3BS4-0.55-0.56	3BS1-0.33-0.55	3BS5-0.07-0.33	C-3BS5-0.07	C-3BL2-0.022	3BL2-0.22-0.28	3BL8-0.28-0.31	3BL1-0.31-0.38	3BL9-0.38-0.50	3BL10-0.50-0.63	3BL7-0.63-1.00	Chr[a]
Marker no.	5	27	4	9	0	2	5	3	4	5	0	0	4	2	4	28	102
Size of the bin (Mb)	56	39	13	78	4	4	95	108	34	124	33	17	39	67	73	208	992
% of physical size per arm	12.93	9.01	3.00	18.01	0.92	0.92	21.94	24.94	7.85	22.06	5.87	3.02	6.94	11.92	12.99	37.01	
% of physical size per chromosome	5.65	3.93	1.31	7.86	0.40	0.40	9.58	10.89	3.43	12.50	3.33	1.71	3.93	6.75	7.36	20.97	
Recombination values on the F_2 CsRe map (cM)	5.20	33.05	4.85	13.35	—	0.65	1.55	0.85	0.00	1.55	1.55	1.55	6.20	4.15	3.45	104.15	179
% of recombination per arm	8.74	55.55	8.15	22.44	—	1.09	2.61	1.43	0.00	1.30	1.30	1.30	5.19	3.47	2.89	87.15	
% of recombination per chromosome	2.91	18.46	2.71	7.46	—	0.36	0.87	0.47	0.00	0.87	0.87	0.87	3.46	2.32	1.93	58.18	
Marker/Mb/bin	11.30	1.44	3.25	8.67	—	2.00	19.00	36.00	8.50	34.80	34.80	34.80	9.75	33.50	18.25	7.43	9.73
Crossover frequency (cM/Mb)	0.09	0.85	0.37	0.17	—	0.08	0.02	0.01	0.00	0.01	0.01	0.01	0.16	0.06	0.05	0.50	0.18

[a] Chr, whole chromosome.

ences were found between the two arms in terms of both average CO frequency and CO distribution. The short arm displayed an average CO frequency of 0.14 cM/Mb and a clear gradient of CO from the centromere to the telomere with a peak (0.85 cM/Mb) in bin 3BS8-0.78–0.87 (Figure 1). In addition, CO frequency decreased dramatically in the most telomeric bin 3BS3-0.87–1.00 with a value of 0.1 cM/Mb only. In the long arm, the average CO frequency was higher (0.21 cM/Mb) with 87% of the crossovers occurring in bin 3BL7-0.63–1.00 (58% of the whole chromosome CO). In contrast to the short arm, no decrease in the CO frequency was observed in the telomeric part of the long arm. However, the large size of the terminal deletion bin (208 Mb) precluded the fine-scale study of the CO gradient in the telomeric region and possibly prevented the identification of such a decrease. Finally, bin 3BL1-0.31–0.38 showed a higher CO rate than bins 3BL9-0.38–0.50 and 3BL10-0.50–0.63 (Figure 1), though it was difficult to draw any conclusion due to the paucity of markers in these regions.

Similar results were observed using androgenetic and gynogenetic CsCt DH populations with a distribution of COs in bins 3BS8-0.78–0.87 and 3BL7-0.63–1.00 explaining 74% of the COs in the male and 81% in the female maps (Figure 1 in supplemental material 2). Interestingly, the same decrease in CO number was observed in both maps at the end of the short arm of chromosome 3B.

These results clearly demonstrate that COs on chromosome 3B are localized mainly in two distal regions (bins 3BS8-0.78–0.87 and 3BL7-0.63–1.00) that represent 247 Mb and 25% of the chromosome and that 27% of the chromosome in the centromeric region is devoid of crossovers. In addition, the decrease of CO frequency observed in bin 3BS3-0.87–1.00 and the average rate observed in bin 3BL1-0.31–0.38 strongly suggest that the density of COs is not increasing uniformly from the centromere to the telomere.

Positive interference along chromosome 3B: To study the interference between COs and potential differences between male and female meiosis, we used two mapping populations produced by andro- or gynogenesis from the cultivars Chinese Spring and Courtot.

No significant difference between the two populations was noted in terms of map size (190.3 and 193.8 cM; Figure 1 in supplemental material 2) and average number of COs per gamete (male, 1.82; female, 1.87). However, among the 31 intervals studied, one from the distal region of the long arm (*Xwmm1133-Xwmm702*, *P*-value = 10^{-4}; Figure 2 in supplemental material 2) showed a 17-fold increase in female CO rate, suggesting a local impact of sex on CO distribution.

These 31 intervals (mean size 6.2 cM; range 0.5–27.2 cM) were then used for interference analysis. In the absence of interference, COs produced during meiosis are independent. The noninterference hypothesis H_0

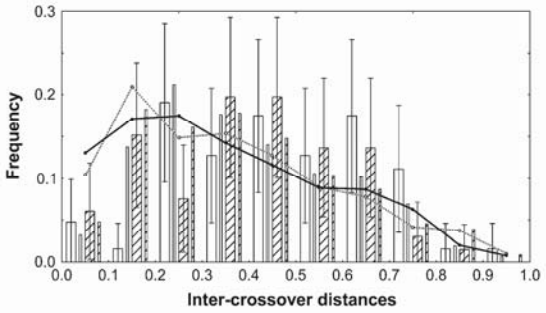


FIGURE 2.—Frequency distribution of distances between two COs referred to as intercrossover distances estimated from recombinant intervals. Distances are relative to chromosome genetic map length, which is 190 cM in the male map and 194 cM in the female map. Open bars, female meiosis; hatched bars, male meiosis; wide bars, experimental data; narrow bars, distribution of simulated gametes generated using the Gamma model (see MATERIALS AND METHODS). Curves: theoretical expectations under the hypothesis of no crossover interference for male (hatched line and open symbols) and female (solid line and symbols) meiosis.

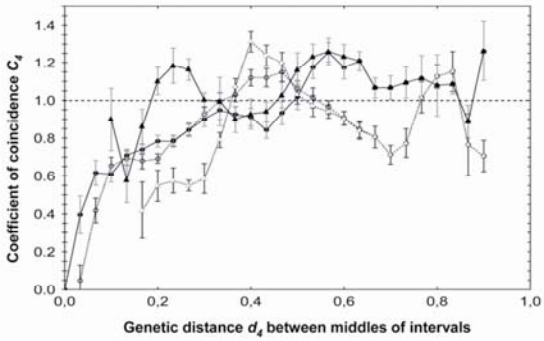


FIGURE 3.—Variation of coefficient of coincidence C_4 as a function of the distance d_4 between intervals. Distances are relative to chromosome genetic map length, which is 190 cM in the male map and 194 cM in the female map. Each point represents the mean value of C_4 for all interval pairs with d_4 falling into the bin. Solid lines and symbols, female meiosis; dashed lines and open symbols, male meiosis. Circles, all interval pairs; triangles, only interval pairs spanning the centromere (one interval on the short arm and one on the long arm). The dashed horizontal line indicates the expected value of C_4 in the absence of interference.

(expected CO distribution without interference equals observed CO distribution) was tested using the frequency of the distance between two crossovers, referred as intercrossover distances on the CtCs andro- and gynogenetic maps (Figure 2). H_0 *P*-values of 0.084 for male and 3.6×10^{-4} for female recombination suggested interference in both male and female meioses with a larger effect in female than in male meiosis.

The intensity of interference was computed using the Gamma model in which the parameter m provides a measure of the degree of interference on the basis of the genetic distances between successive COs. Positive interference, *i.e.*, the inhibition of crossovers by nearby crossovers, is indicated by m values >0. We used the frequency of the genetic distance between two crossovers for male and female meiosis (Figure 2) and compared it with the theoretical distribution obtained with the Gamma model. For male meiosis, the fitted value for m was 1.2, ranging from 0 to 6.1 with a 95% confidence interval. For female meiosis, m varied from 0.95 to 12.75 with a fitted value of 3.5. While these results confirmed the previously suggested greater interference in female meiosis than in male meiosis, the difference was too weak to be statistically significant when tested by bootstrap analysis (*P*-value = 0.25). Interference was further investigated using the coefficient of coincidence (C), which corresponds to the ratio between the frequency of observed double recombinants and the expected frequency in the absence of interference. Here, we used the C_4 parameter (see MATERIALS AND METHODS) that allows evaluating the distance range in which interference is effective. The results showed that positive interference is strong within short distances (<10 cM, Figure 3) and decreases with distance to the point where it becomes very weak at distances >45 cM ($C_4 > 0.75$).

The results also indicated small C_4 values <1 for intervals spanning the centromere in male meiosis (Figure 3), suggesting that interference extends through the centromere and that CO formation in the two arms of chromosome 3B is not independent. Moreover, and in contrast to the other regions of the chromosome, interference in the centromeric regions acts on long genetic distances of at least 31 cM (Figure 3). No such pattern was observed for female meiosis.

DISCUSSION

CO distribution along chromosome 3B: The construction of a dense genetic map of wheat chromosome 3B comprising 102 markers well distributed across the genetic and deletion maps was challenging in a species with a large genome of 17 Gb (ZONNEVELD *et al.* 2005), consisting of ~90% of repeated sequence (FLAVELL and SMITH 1977; LI *et al.* 2004; PAUX *et al.* 2006), for which no genome sequence is available and showing a low level of polymorphism (CHAO *et al.* 1989; KAM-MORGAN *et al.* 1989; CADALEN *et al.* 1997). The development of the CsRe genetic map of chromosome 3B was facilitated greatly by the construction of a BAC-based physical map (PAUX *et al.* 2008) that provided a large source of markers well distributed along the minimal tiling path (MTP) of this chromosome. ISBP markers (PAUX *et al.* 2006) derived from the BAC end sequences of the MTP were particularly useful for filling gaps on the genetic map and ensuring a homogeneous marker distribution along the chromosome.

Comparison of the CsRe genetic map with the physical deletion map showed that the order of the loci is highly conserved between the two maps. This is probably due to the fact that Chinese Spring was used as one of the parents in our mapping population and for the construction of the physical map. Thus, the CsRe mapping population seems very well adapted for ordering contigs of the physical maps from the other chromosomes of cv. Chinese Spring that are currently under construction in the framework of the International Wheat Genome Sequencing Initiative (see http://www.wheatgenome.org for further information).

Analysis of the CO distribution along chromosome 3B showed a recombination gradient from the centromere to the telomeres on both arms and a suppression of COs in 27% of the chromosome in the proximal regions. This latter phenomenon was observed in the first-generation deletion bin maps of group-7 chromosomes (Werner *et al.* 1992) and the other homeologous group chromosomes of wheat (for review see Erayman *et al.* 2004). The absence of CO at the centromere is a general feature of most of the species studied so far independently of their genome size and complexity (Anderson *et al.* 2003; Wu *et al.* 2003; Jensen-Seaman *et al.* 2004; Drouaud *et al.* 2006). Thus, our results on chromosome 3B confirm the general trend and demonstrate that even if it is the largest chromosome (995 Mb), the CO pattern is representative of all other wheat chromosomes.

Interestingly, the high resolution of the maps allowed us to observe a dramatic decrease in the COs' gradient at the very distal end of the short arm, suggesting suppressed recombination in this region. This decrease was also observed in the cross between Courtot and Chinese Spring (Figure 1 in supplemental material 2) as well as in an unrelated population developed between the cultivars Arche and Recital (data not shown), suggesting that the pattern observed in the CsRe map reflects a general behavior of CO distribution on chromosome 3B. A possible explanation could be the lack of markers close to the telomere that would lead to an underestimation of the total CO number and of the CO frequency for the terminal bin. However, synteny with rice suggests that the most distal marker (*Xgwm389*) is very close to the telomere. Indeed, although the synteny with *Xgwm389* could not be determined because it is a microsatellite, two BAC contigs located proximal to X*gwm389* and carrying the SSR markers *Xgwm1034* and *cfb6012* and also ESTs BE404656 and BJ302829 correspond to sequences containing the orthologous genes OS01G01960 and OS01G03100 located on positions 500 kb–1 Mb, respectively, on rice chromosome 1 (data not shown). Thus, these results suggest that the whole deletion bin is covered and that most of the COs occurring in the telomeric bin are included in our genetic map. Another possibility to explain the reduced or suppressed recombination in the terminal region is the high heterochromatin content of the telomeric end of wheat chromosome 3B (Gill *et al.* 1991), with heterochromatin being well known to inhibit crossover formation (Gaut *et al.* 2007). Our findings, observed on four different populations, conflict with the previous study of Erayman *et al.* (2004), who observed high recombination frequency in the most distal bin (3BS3-0.87–1.00) and no recombination in the subterminal bin 3BS8-0.78–0.87. The discrepancy probably originates from differences in the relative position of markers on the genetic and deletion maps between the two studies that likely result from the use of a consensus deletion map from the *A*, *B*, and *D* group-3 chromosomes and a non-systematic assignation of markers in deletion bins by Erayman *et al.* (2004).

In contrast, no decrease in CO number was observed at the telomere of the long arm of chromosome 3B. Lukaszewski and Curtis (1993) showed that the CO frequency was higher in the subtelomeric region of the long arm of chromosome 3B, suggesting a pattern similar to the one we observed on the short arm. Probably the size of terminal deletion bin 3BL7-0.63–1.00 (208 Mb) precluded the detection of such a decrease in our study. Decreasing CO number at the end of chromosomes was also reported in a cross between rye and wheat (Lukaszewski *et al.* 2004) and in the telomeric parts of chromosome 3, in the orthologous chromosome of wheat chromosome 3B, in a cross between Lolium and Festuca (King *et al.* 2002, 2007). Additional mapping data on chromosome 3B and other wheat chromosomes for which high-resolution genetic and physical maps are underway at the international level (http://www.wheatgenome.org) will help to confirm the CO pattern in the telomeric ends of the chromosomes. The fact that 70% of the recombination occurs in the most distal part (23 Mb) of the long arm of chromosome 4B (See *et al.* 2006) may indicate that each chromosome has a specific pattern of recombination distribution in the telomeric regions.

In addition to the decrease of CO frequency in the distal part of the short arm, we observed an increase of CO frequency in the middle part of the long arm (bin 3BL1-0.31–0.38). Thus, it appears that at least along chromosome 3B, CO frequency does not follow a simple gradient from the centromeres to the telomeres. This implies that cloning a gene in a distal position on a wheat chromosome may sometimes be more difficult than expected while cloning a gene in the middle part would be possible without developing very large segregating populations.

CO distribution was shown to be correlated with genes in human (Kong *et al.* 2002), mouse (Paigen *et al.* 2008), rice (Wu *et al.* 2003), and maize (Anderson *et al.* 2006). In wheat, Erayman *et al.* (2004) showed that seven gene-rich regions (GRRs) are present on a consensus chromosome 3 map (three on the short arm and four on the long arm) and that most of the recombination occurs within the GRRs. In our study, we also observed a

correlation between the CO frequency and the presence of the GRRs on the long arm. Especially, the increase of CO frequency that was found in bin 3BL1-0.31–0.38 corresponded exactly to the major GRR (3L0.3) mentioned by Erayman *et al.* (2004) in the middle of this arm. In contrast, we did not find the same correlation on the short arm. Maximal recombination was observed in bin 3BS8-0.78–0.87, which does not correspond to one of the GRR regions that was defined on the short arm of chromosome 3B. As discussed previously, this is likely due to the nonsystematic assignment by Erayman *et al.* (2004) of markers in deletion bins and the identification of the GRR on the short arm might need to be reassessed with more markers and on individual chromosomes.

If a large recombination gradient can be observed in large genome species such as wheat, barley (Kunzel *et al.* 2000), and maize (Anderson *et al.* 2003), recombination is more evenly distributed in species with smaller genomes such as rice (Wu *et al.* 2003). In the small genome species *Arabidopsis thaliana*, no gradient has been observed (Drouaud *et al.* 2006), suggesting that the pattern of CO distribution is correlated with genome size. In species displaying a CO gradient, the high CO rates observed in subtelomeric regions are explained generally by the distal position of these regions. This hypothesis is supported by the fact that chromosomes with terminal deletions show increased CO number in regions that were formerly proximal (Jones *et al.* 2002; Qi *et al.* 2002). Higher CO rates in distal subtelomeric regions of chromosomes have been related to the onset of meiosis with the bouquet formation occurring in the vicinity of these regions thereby favoring CO formation (Corredor *et al.* 2007). This feature, however, is a common feature of meiosis and thus also appears in species with small genomes that do not show such contrast in CO rate between distal and proximal regions. Moreover, very recently, Lukaszewski (2008) showed that a rye chromosome with an inverted long arm where the telomeric regions are located at the centromere exhibited simultaneously an inverted pattern of chiasma distribution; *i.e.*, recombination was higher close to the neocentromeric regions. He concluded that CO frequency along a chromosome is not position dependent but sequence dependent. In contrast to species with small genomes, those with large genomes show extensive chromosome remodeling in their subterminal chromosomal regions (Dawe *et al.* 1994; Prieto *et al.* 2004), thereby suggesting that high CO rates are due to remodeling in these regions. A recent study (Colas *et al.* 2008) of the meiotic behavior of translocated rye chromosomes showed that when chromosome remodeling is inhibited in the heterozygous subterminal regions of rye chromosome 1R in a 1BL.1RS wheat–rye translocation line, COs are also inhibited. To confirm this hypothesis, it would be interesting to study whether remodeling is observed in new distal regions of deleted chromosomes or chromosomes with inverted arms.

Positive interference along chromosome 3B: In addition to the frequency of individual CO formation, recombination is influenced by interference, *i.e.*, the effect of one CO on the occurrence of other COs in its vicinity. Here, we analyzed for the first time in wheat the intensity of the interference on a chromosome, using a Gamma model. Estimates of the interference parameter m of 1.2 for male meiosis and 3.5 for female meiosis were found. In tomato, late recombination nodules observations indicated m values of 1.3 and 1.9 for chromosomes 1 and 2, respectively (Lhuissier *et al.* 2007). Similar values were obtained from genetic mapping data, in human ($m = 2.3$) (Broman and Weber 2000) and Arabidopsis ($m = 3$–9) (Copenhaver *et al.* 2002), whereas m values between 5.0 and infinity were observed in mouse (Broman *et al.* 2002). Understanding the molecular basis of interference intensity difference is a very challenging task that is under progress in model species such as yeast and Arabidopsis.

On the basis of the coefficient of coincidence, our results indicated strong positive interference at distances <10 cM. We confirmed this interference in a different cross where the screening of >1400 F_2 CsRe using two markers separated by 12 cM and localized in the distal part of the short arm showed an absence of double COs (P-value = 0.015, data not shown). In wheat and barley, Peng *et al.* (2000) and Esch and Weber (2002) reported negative interference within short distances. However, in contrast to our study that is based on doubled-haploid populations, Peng *et al.* (2000) used F_2 plants. Estimation of the coefficient of coincidence with this type of population is biased because two independent gametes are considered as one. To deal with dependency, an expectation-maximization (EM) procedure needs to be used on F_2 populations for distinguishing independent meiotic events present in an F_2 individual. In any case, using an F_2 population results in substantial loss of statistical power to detect interference compared to DH populations in which individual gametes are considered. Moreover, in their studies, Peng *et al.* (2000) and Esch and Weber (2002) considered intervals with very small numbers of recombinants (<10), thereby introducing another bias in the measurement of the coefficient of coincidence and leading to potentially large overestimations (Martin and Hospital 2006). Thus using an F_2 population and analyzing small intervals result in an overestimation of double recombination events, in high values for the coefficient of coincidence ($C > 1$), and, finally, in what appears to be negative interference.

The coefficient of coincidence analysis also shows that interference acts across the centromere in male meiosis, a phenomenon that has been seen already in other species (Colombo and Jones 1997; Drouaud *et al.* 2007; Lian *et al.* 2008). Our results suggest that, in wheat, the interference range differs between proximal and distal regions by acting on longer genetic distances

around the centromere (30 cM) compared to the other regions (10 cM). In humans and mice (Codina-Pascual et al. 2006; Petkov et al. 2007), interference was shown to be correlated with the synaptonemal complex length and therefore was constant on physical distances measured in micrometers along the entire chromosome. However, distances estimated in micrometers depend on chromatin compaction, implying that 1 μm represents a longer physical distance measured in megabases in condensed regions compared to less condensed regions. As proximal regions of chromosome 3B are rich in heterochromatin (Gill et al. 1991), they are probably highly condensed and 1 μm in these regions represents a much larger physical size in megabases compared to 1 μm in the distal regions. Consequently, more COs can occur in the large physical (megabases) proximal regions, which could explain the longer genetic interval observed for the interference range around the centromere.

Genetic mapping in wheat: The CsRe mapping data presented in this article showed that chromosome 3B exhibits on average 3.60 COs per bivalent at each meiosis. This is slightly higher than the 2.97 chiasmata observed by Sallee and Kimber (1978) and is probably due to the difficulty to accurately estimate chiasma number on this very large chromosome since the frequency of two close COs on the same arm is more frequent than on short chromosomes. However, it confirms that the coverage of the genetic map was complete and took into account all the possible CO events. With 48.3 chiasma suggested for the entire genome (Sallee and Kimber 1978), the size of a wheat genetic map should be ~2400 cM. Considering the slight underestimation of chiasma number, the size should rather be between 2400 and 3000 cM. Wheat genetic maps that were developed initially were smaller than expected (1700–1800 cM; for a review see Varshney et al. 2004). However, these maps covered only poorly the whole genome with sometimes chromosomes that were not represented (Cadalen et al. 1997). The actual sizes of wheat genetic maps range from 2500 to 4000 cM (Sourdille et al. 2003; Quarrie et al. 2005; Suenaga et al. 2005; Akbari et al. 2006; Torada et al. 2006; Xue et al. 2008), indicating that some of them exceed the expected value. This may suggest genotyping errors, the use of not well-adapted mapping functions, or differences in CO frequencies between different wheat varieties and provides a note of caution on the quality of the genetic maps of wheat. Analysis of the coefficient of coincidence showed that interference is strong at short distances, inhibiting the formation of double COs. This was proved by the size of our genetic map that was similar to previously published maps even though a larger number of markers (102) were used in our study. Thus, these results confirm that the size increase of genetic maps that results from the addition of markers reflects genotyping errors rather than double COs.

Conclusions: In this work, we analyzed the recombination gradient and the interference, two factors that influence recombination on wheat chromosome 3B. At least one additional factor affects recombination distribution and rate: sequence similarity between the homologous chromosomes. This was demonstrated in yeast (Datta et al. 1997), in mammalian cells (Lukacsovich and Waldman 1999), and in Arabidopsis (Opperman et al. 2004). In wheat, genetic maps obtained from interspecific crosses are shorter than those issued from intraspecific crosses (Luo et al. 2000), suggesting lower recombination rates in the former case. Moreover, recombination is known to be reduced at the level of introgressed segments (Ji and Chetelat 2007). The direct effect of sequence similarity has not been studied so far in wheat because large stretches of sequence from different varieties are not available yet. Crossover distribution is best studied in species for which whole genome sequences are available. To better study this essential phenomenon in wheat, we are currently sequencing and comparing large stretches of DNA originating from recombinant and nonrecombinant regions in different wheat varieties.

We thank J. Philippon and D. Boyer for their technical help. We also thank G. Gay, M. R. Perretant, S. Bernard, and A. Madeore for the development of populations. We are grateful to C. Mezard, B. S. Gill, and S. Faure for critical reading of the manuscript and to K. Eversole for editing. Génoplante, RAGT, and Syngenta are greatly acknowledged for providing, respectively, the gpw, wmm, and nw/dw SSRs. We thank also the members of the GENTYANE genotyping platform of the UMR GDEC 1095 for their technical support during genotyping. This work has been supported by grants from the Agence Nationale de la Recherche (ANR-05-BLANC-0258-01 and ANR-07-BLANC-COPATH) and from the Institut National de la Recherche Agronomique. C.S. is funded by a grant from the French Ministry of Research.

LITERATURE CITED

Akbari, M., P. Wenzl, V. Caig, J. Carling, L. Xia *et al.*, 2006 Diversity arrays technology (DArT) for high-throughput profiling of the hexaploid wheat genome. Theor. Appl. Genet. **113:** 1409–1420.

Akhunov, E. D., A. R. Akhunova, A. M. Linkiewicz, J. Dubcovsky, D. Hummel *et al.*, 2003a Synteny perturbations between wheat homoeologous chromosomes caused by locus duplications and deletions correlate with recombination rates. Proc. Natl. Acad. Sci. USA **100:** 10836–10841.

Akhunov, E. D., A. W. Goodyear, S. Geng, L. L. Qi, B. Echalier *et al.*, 2003b The organization and rate of evolution of wheat genomes are correlated with recombination rates along chromosome arms. Genome Res. **13:** 753–763.

Anderson, K. A., and S. M. Stack, 2002 Meiotic recombination in plants. Curr. Genomics **3:** 507–525.

Anderson, L. K., G. G. Doyle, B. Brigham, J. Carter, K. D. Hooker *et al.*, 2003 High-resolution crossover maps for each bivalent of *Zea mays* using recombination nodules. Genetics **165:** 849–865.

Anderson, L. K., A. Lai, S. M. Stack, C. Rizzon and B. S. Gaut, 2006 Uneven distribution of expressed sequence tag loci on maize pachytene chromosomes. Genome Res. **16:** 115–122.

Arnheim, N., P. Calabrese and I. Tiemann-Boege, 2007 Mammalian meiotic recombination hot spots. Annu. Rev. Genet. **41:** 369–399.

Baudat, F., and B. de Massy, 2007 Regulating double-stranded DNA break repair towards crossover or non-crossover during mammalian meiosis. Chromosome Res. **15:** 565–577.

Broman, K. W., and J. L. Weber, 2000 Characterization of human crossover interference. Am. J. Hum. Genet. **66:** 1911–1926.

Broman, K. W., L. B. Rowe, G. A. Churchill and K. Paigen, 2002 Crossover interference in the mouse. Genetics **160:** 1123–1131.

Cadalen, T., C. Boeuf, S. Bernard and M. Bernard, 1997 An intervarietal molecular marker map in *Triticum aestivum* L. em. Thell. and comparison with a map from a wide cross. Theor. Appl. Genet. **94:** 367–377.

Chao, S., P. J. Sharp, A. J. Worland, E. J. Warham, R. M. D. Koebner *et al.*, 1989 RFLP-based genetic maps of wheat homoeologous group 7 chromosomes. Theor. Appl. Genet. **78:** 495–504.

Codina-Pascual, M., M. Campillo, J. Kraus, M. R. Speicher, J. Egozcue *et al.*, 2006 Crossover frequency and synaptonemal complex length: their variability and effects on human male meiosis. Mol. Hum. Reprod. **12:** 123–133.

Colas, I., P. Shaw, P. Prieto, M. Wanous, W. Spielmeyer *et al.*, 2008 Effective chromosome pairing requires chromatin remodeling at the onset of meiosis. Proc. Natl. Acad. Sci. USA **105:** 6075–6080.

Colombo, P. C., and G. H. Jones, 1997 Chiasma interference is blind to centromeres. Heredity **79:** 214–227.

Copenhaver, G. P., E. A. Housworth and F. W. Stahl, 2002 Crossover interference in Arabidopsis. Genetics **160:** 1631–1639.

Corredor, E., A. J. Lukaszewski, P. Pachon, D. C. Allen and T. Naranjo, 2007 Terminal regions of wheat chromosomes select their pairing partners in meiosis. Genetics **177:** 699–706.

Cromie, G. A., and G. R. Smith, 2007 Branching out: meiotic recombination and its regulation. Trends Cell Biol. **17:** 448–455.

Datta, A., M. Hendrix, M. Lipsitch and S. Jinks-Robertson, 1997 Dual roles for DNA sequence identity and the mismatch repair system in the regulation of mitotic crossing-over in yeast. Proc. Natl. Acad. Sci. USA **94:** 9757–9762.

Dawe, R. K., J. W. Sedat, D. A. Agard and W. Z. Cande, 1994 Meiotic chromosome pairing in maize is associated with a novel chromatin organization. Cell **76:** 901–912.

Drouaud, J., C. Camilleri, P. Y. Bourguignon, A. Canaguier, A. Berard *et al.*, 2006 Variation in crossing-over rates across chromosome 4 of Arabidopsis thaliana reveals the presence of meiotic recombination "hot spots". Genome Res. **16:** 106–114.

Drouaud, J., R. Mercier, L. Chelysheva, A. Berard, M. Falque *et al.*, 2007 Sex-specific crossover distributions and variations in interference level along Arabidopsis thaliana chromosome 4. PLoS Genet. **3:** e106.

Dvorak, J., and E. D. Akhunov, 2005 Tempos of gene locus deletions and duplications and their relationship to recombination rate during diploid and polyploid evolution in the Aegilops-Triticum alliance. Genetics **171:** 323–332.

Dvorak, J., Z. L. Yang, F. M. You and M. C. Luo, 2004 Deletion polymorphism in wheat chromosome regions with contrasting recombination rates. Genetics **168:** 1665–1675.

Endo, T. R., and B. S. Gill, 1996 The deletion stocks of common wheat. J. Hered. **87:** 295–307.

Erayman, M., D. Sandhu, D. Sidhu, M. Dilbirligi, P. S. Baenziger *et al.*, 2004 Demarcating the gene-rich regions of the wheat genome. Nucleic Acids Res. **32:** 3546–3565.

Esch, E., and E. Weber, 2002 Investigation of crossover interference in barley (Hordeum vulgare L.) using the coefficient of coincidence. Theor. Appl. Genet. **104:** 786–796.

Faris, J. D., K. M. Haen and B. S. Gill, 2000 Saturation mapping of a gene-rich recombination hot spot region in wheat. Genetics **154:** 823–835.

Felix, I., J. P. Martinant, M. Bernard, S. Bernard and G. Branlard, 1996 Genetic characterization of storage proteins in a set of F1-derived haploid lines in bread wheat. Theor. Appl. Genet. **92:** 340–346.

Flavell, R. B., and D. B. Smith, 1977 Hyperpolymer formation during renaturation of DNA from genomes with different sequence organisation. Nucleic Acids Res. **4:** 2429–2444.

Fu, H., W. Park, X. Yan, Z. Zheng, B. Shen *et al.*, 2001 The highly recombinogenic bz locus lies in an unusually gene-rich region of the maize genome. Proc. Natl. Acad. Sci. USA **98:** 8903–8908.

Gaut, B. S., S. I. Wright, C. Rizzon, J. Dvorak and L. K. Anderson, 2007 Recombination: an underappreciated factor in the evolution of plant genomes. Nat. Rev. Genet. **8:** 77–84.

Gill, B. S., B. Friebe and T. R. Endo, 1991 Standard karyotype and nomenclature system for description of chromosome bands and structural aberrations in wheat (*Triticum aestivum*). Genome **34:** 830–839.

Hamant, O., H. Ma and W. Z. Cande, 2006 Genetics of meiotic prophase I in plants. Annu. Rev. Plant Biol. **57:** 267–302.

Hochberg, Y., 1988 A sharper Bonferonni procedure for multiple tests of significance. Biometrika **75:** 800–803.

Hodges, E., S. M. Boddy, S. Thomas and J. L. Smith, 1997 Modification of IgH PCR clonal analysis by the addition of sucrose and cresol red directly to PCR reaction mixes. Mol. Pathol. **50:** 164–166.

Jensen-Seaman, M. I., T. S. Furey, B. A. Payseur, Y. Lu, K. M. Roskin *et al.*, 2004 Comparative recombination rates in the rat, mouse, and human genomes. Genome Res. **14:** 528–538.

Ji, Y., and R. T. Chetelat, 2007 GISH analysis of meiotic chromosome pairing in Solanum lycopersicoides introgression lines of cultivated tomato. Genome **50:** 825–833.

Jones, E., K. Rybka and J. Lukaszewski, 2002 The effect of a deficiency and a deletion on recombination in chromosome 1BL in wheat. Theor. Appl. Genet. **104:** 1204–1208.

Jones, G. H., 1987 *Chiasmata.* Academic Press, New York.

Jones, G. H., and F. C. Franklin, 2006 Meiotic crossing-over: obligation and interference. Cell **126:** 246–248.

Kam-Morgan, L. N. W., B. S. Gill and S. Muthukrishnan, 1989 DNA restriction fragment length polymorphisms: a strategy for genetic mapping of D genome of wheat. Genome **32:** 724–732.

King, J., I. P. Armstead, I. S. Donnison, H. M. Thomas, R. N. Jones *et al.*, 2002 Physical and genetic mapping in the grasses *Lolium perenne* and *Festuca pratensis*. Genetics **161:** 315–324.

King, J., I. P. Armstead, I. S. Donnison, L. A. Roberts, J. A. Harper *et al.*, 2007 Comparative analyses between Lolium/Festuca introgression lines and rice reveal the major fraction of functionally annotated gene models is located in recombination-poor/very recombination-poor regions of the genome. Genetics **177:** 597–606.

Kong, A., D. F. Gudbjartsson, J. Sainz, G. M. Jonsdottir, S. A. Gudjonsson *et al.*, 2002 A high-resolution recombination map of the human genome. Nat. Genet. **31:** 241–247.

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

Lander, E. S., P. Green, J. Abrahamson, A. Barlow, M. J. Daly *et al.*, 1987 MAPMAKER: an interactive computer package for constructing primary genetic linkage maps of experimental and natural populations. Genomics **1:** 174–181.

Laurie, D. A., and M. D. Bennett, 1988 The production of haploid wheat plants from wheat x maize crosses. Theor. Appl. Genet. **3:** 393–397.

Lhuissier, F. G., H. H. Offenberg, P. E. Wittich, N. O. Vischer and C. Heyting, 2007 The mismatch repair protein MLH1 marks a subset of strongly interfering crossovers in tomato. Plant Cell **19:** 862–876.

Li, W., P. Zhang, J. P. Fellers, B. Friebe and B. S. Gill, 2004 Sequence composition, organization, and evolution of the core Triticeae genome. Plant J. **40:** 500–511.

Lian, J., Y. Yin, M. Oliver-Bonet, T. Liehr, E. Ko *et al.*, 2008 Variation in crossover interference levels on individual chromosomes from human males. Hum. Mol. Genet. **17:** 2583–2594.

Lukacsovich, T., and A. S. Waldman, 1999 Suppression of intrachromosomal gene conversion in mammalian cells by small degrees of sequence divergence. Genetics **151:** 1559–1568.

Lukaszewski, A., and C. A. Curtis, 1993 Physical distribution of recombination in B-genome chromosome of tetraploid wheat. Theor. Appl. Genet. **86:** 121–127.

Lukaszewski, A. J., 2008 Unexpected behavior of an inverted rye chromosome arm in wheat. Chromosoma(in press).

Lukaszewski, A. J., K. Rybka, V. Korzun, S. V. Malyshev, B. Lapinski *et al.*, 2004 Genetic and physical mapping of homoeologous recombination points involving wheat chromosome 2B and rye chromosome 2R. Genome **47:** 36–45.

Luo, M. C., Z. L. Yang, R. S. Kota and J. Dvorak, 2000 Recombination of chromosomes 3A(m) and 5A(m) of *Triticum monococcum* with homeologous chromosomes 3A and 5A of wheat: the distribution of recombination across chromosomes. Genetics **154:** 1301–1308.

Martin, O. C., and F. Hospital, 2006 Two- and three-locus tests for linkage analysis using recombinant inbred lines. Genetics **173:** 451–459.

McPeek, M. S., and T. P. Speed, 1995 Modeling interference in genetic recombination. Genetics **139:** 1031–1044.

Nicot, N., V. Chiquet, B. Gandon, L. Amilhat, F. Legeai *et al.*, 2004 Study of simple sequence repeat (SSR) markers from wheat expressed sequence tags (ESTs). Theor. Appl. Genet. **109:** 800–805.

Opperman, R., E. Emmanuel and A. A. Levy, 2004 The effect of sequence divergence on recombination between direct repeats in Arabidopsis. Genetics **168:** 2207–2215.

Paigen, K., J. P. Szatkiewicz, K. Sawyer, N. Leahy, E. D. Parvanov *et al.*, 2008 The recombinational anatomy of a mouse chromosome. PLoS Genet. **4:** e1000119.

Paux, E., D. Roger, E. Badaeva, G. Gay, M. Bernard *et al.*, 2006 Characterizing the composition and evolution of homoeologous genomes in hexaploid wheat through BAC-end sequencing on chromosome 3B. Plant J. **48:** 463–474.

Paux, E., P. Sourdille, J. Salse, C. Saintenac, F. Choulet *et al.*, 2008 A physical map of the 1Gb bread wheat chromosome 3B. Science **322:** 101–104.

Peng, J., A. B. Korol, T. Fahima, M. S. Roder, Y. I. Ronin *et al.*, 2000 Molecular genetic maps in wild emmer wheat, Triticum dicoccoides: genome-wide coverage, massive negative interference, and putative quasi-linkage. Genome Res. **10:** 1509–1531.

Petkov, P. M., K.W. Broman, J. P. Szatkiewicz and K. Paigen, 2007 Crossover interference underlies sex differences in recombination rates. Trends Genet. **23:** 539–542.

Prieto, P., P. Shaw and G. Moore, 2004 Homologue recognition during meiosis is associated with a change in chromatin conformation. Nat. Cell Biol. **6:** 906–908.

Qi, L. L., B. Friebe and B. S. Gill, 2002 A strategy for enhancing recombination in proximal regions of chromosomes. Chromosome Res. **10:** 645–654.

Quarrie, S. A., A. Steed, C. Calestani, A. Semikhodskii, C. Lebreton *et al.*, 2005 A high-density genetic map of hexaploid wheat (Triticum aestivum L.) from the cross Chinese Spring x SQ1 and its use to compare QTLs for grain yield across a range of environments. Theor. Appl. Genet. **110:** 865–880.

Sallee, P. J., and G. Kimber, 1978 An analysis of the pairing of wheat telocentric chromosomes. Proceedings of the 5th International Wheat Genetics Symposium, pp. 408–419.

Sears, E. R., 1954 The aneuploids of common wheat. University of Missouri Agricultural Experiment Station, Columbia **572:** 1–58.

Sears, E. R., 1966 *Nullisomic-Tetrasomic Combinations in Hexaploid Wheat.*

Sears, E. R., and L. M. S. Sears, 1978 *The Telocentric Chromosomes of Common Wheat.* New Delhi.

See, D. R., S. Brooks, J. C. Nelson, G. Brown-Guedira, B. Friebe *et al.*, 2006 Gene evolution at the ends of wheat chromosomes. Proc. Natl. Acad. Sci. USA **103:** 4162–4167.

Sidhu, D., and K. S. Gill, 2004 Distribution of genes and recombination in wheat and other eucaryotes. Plant Cell Tissue Organ Cult. **79:** 257–270.

Sourdille, P., T. Cadalen, H. Guyomarc'h, J. W. Snape, M. R. Perretant *et al.*, 2003 An update of the Courtot x Chinese Spring intervarietal molecular marker linkage map for the QTL detection of agronomic traits in wheat. Theor. Appl. Genet. **106:** 530–538.

Stahl, F. W., H. M. Foss, L. S. Young, R. H. Borts, M. F. Abdullah *et al.*, 2004 Does crossover interference count in *Saccharomyces cerevisiae*? Genetics **168:** 35–48.

Suenaga, K., M. Khairallah, H. M. William and D. A. Hoisington, 2005 A new intervarietal linkage map and its application for quantitative trait locus analysis of "gigas" features in bread wheat. Genome **48:** 65–75.

Tenaillon, M. I., M. C. Sawkins, A. D. Long, R. L. Gaut, J. F. Doebley *et al.*, 2001 Patterns of DNA sequence polymorphism along chromosome 1 of maize (Zea mays ssp. mays L.). Proc. Natl. Acad. Sci. USA **98:** 9161–9166.

Torada, A., M. Koike, K. Mochida and Y. Ogihara, 2006 SSR-based linkage map with new markers using an intraspecific population of common wheat. Theor. Appl. Genet. **112:** 1042–1051.

Varshney, R. K., V. Korzun and A. Börner, 2004 Molecular maps in cereals: methodology and progress, pp. 35–82 in *Cereal Genomics*, edited by P. K. Gupta and R. K. Varshney. Kluwer Academic Publishers,

Werner, J. E., T. R. Endo and B. S. Gill, 1992 Toward a cytogenetically based physical map of the wheat genome. Proc. Natl. Acad. Sci. USA **89:** 11307–11311.

Wu, J., H. Mizuno, M. Hayashi-Tsugane, Y. Ito, Y. Chiden *et al.*, 2003 Physical maps and recombination frequency of six rice chromosomes. Plant J. **36:** 720–730.

Xue, S., Z. Zhang, F. Lin, Z. Kong, Y. Cao *et al.*, 2008 A high-density intervarietal map of the wheat genome enriched with markers derived from expressed sequence tags. Theor. Appl. Genet. **117:** 181–189.

Zonneveld, B. J., I. J. Leitch and M. D. Bennett, 2005 First nuclear DNA amounts in more than 300 angiosperms. Ann. Bot. **96:** 229–244.

Communicating editor: G. P. Copenhaver

Comparison of CO frequency: In many organisms, the male and female meioses lead to different recombination rates. For the wheat chromosome of interest here, the total genetic lengths are very close, but the local rates seem to differ. We have tested for the significance of such differences using *p*-values. Given two markers, the experimental genotyping results provide a number of informative gametes N_m and N_f for male and female respectively. Among these, n^*_m and n^*_f are recombinant. We want to test the hypothesis H_0 that the recombination rate is sex independent. Applying the maximum likelihood approach to the pooled data gives the estimate $r_{pool} = (n^*_m + n^*_f)/(N_m + N_f)$ for the associated recombination rate. Then within H_0, the probability that the data set produces n_m and n_f recombinants is

$$P(n_m, n_f) = r^{(n_m+n_f)}(1-r)^{(N_m-n_m+N_f-n_f)}\binom{N_m}{n_m}\binom{N_f}{n_f}$$

where $\binom{N}{p} = \frac{N!}{(N-p)!\,p!}$ is the binomial coefficient. Given the experimental values (n^*_m, n^*_f), the *p*-value associated with H_0 is

$$P(H_0) = \sum_{n_m=0}^{N_m} \sum_{n_f=0}^{N_f} P(n_m, n_f)\,\theta\left(P(n^*_m, n^*_f) - P(n_m, n_f)\right)$$

where $\theta(x) = 0$ for $x \leq 0$ and 1 for $x > 0$. These sums can be computed straightforwardly by computer.

Test of the no-interference hypothesis: Interference arises in many organisms where it has been tested for; in wheat we are aware of no previous study, so the most basic question is whether the data suggest the presence of interference or not. We have considered this by using the distance between recombination events. To begin, we consider the frequencies of gametes with zero, one or two COs. Under the no interference hypothesis H_0, crossover events arise independently with uniform density (in genetic position) equal to 1. If the chromosome is of length *L*, the frequency of gametes with k crossovers is $e^{-L}(L^k/k!)$ (The Poisson distribution). However, we cannot resolve crossovers, we only detect recombination events and these can be due to any odd number of crossovers. For an interval of genetic length *l*, there is a recombination event with probability $r = \frac{1}{2}(1-e^{-2l})$. Let $r_1, r_2, \ldots r_k$ be the *k* recombinant rates for the *k* adjacent intervals (there are *k*+1 markers). Since the intervals are independent under H_0, the probability to have 0 recombinant gametes is

$$P(0) = \prod_{i=1}^{k} (1 - r_i)$$

Similarly, one has

$$P(1) = P(0) \sum_{i=1}^{k} \frac{r_i}{(1 - r_i)}$$

$$P(2) = P(0) \sum_{i=1}^{k-1} \sum_{j=i+1}^{k} \frac{r_i}{(1 - r_i)} \frac{r_j}{(1 - r_j)} \ldots \text{ etc.}$$

The test of H_0 is based on the statistics of distances between recombination events. Again, we do not have crossover positions, we only know whether an interval is recombinant or not. We focus on gametes for which there are exactly two recombinant intervals. The probability that intervals i and j are recombinant under the no-interference hypothesis is

$$P(i, j) = \frac{P(0)}{P(2)} \frac{r_i r_j}{(1 - r_i)(1 - r_j)}$$

where $P(0)$ and $P(2)$ were previously computed. We can construct the histogram of the corresponding distances (we take the distance of the midpoints of the intervals *i* and *j* for simplicity) and test whether this histogram is compatible with the experimental one. The theoretical frequencies determine the likelihood of any realization of the histogram. If L^* is the likelihood of the experimental histogram, to compute the *p*-value associated with the no-interference hypothesis H_0, we must sum the likelihoods of all possible realizations of this histogram whose likelihood is less than L^*. As long as this *p*-value is not too small, it can be computed reliably by sampling methods. This was done using a Mathematica code.

The counting model: Several models have been used to describe the positioning of crossover events during meiosis. Most are based on producing crossovers in succession, allowing for interference by preventing crossovers from being too close to one another. Probably the best known example is the counting model (Stahl et al. 2004) which operates on points randomly placed on a bivalent and takes each $(m+1)^{th}$ point to be a crossover.

First a succession of points are generated on the infinite line, say going from left to right; each point is generated from its predecessor by advancing a random (genetic) distance *d*. The model is parameterized by an integer *m*, $m \geq 0$. Second, one of these points is considered to be a crossover, but the *m* following ones are thought of as gene conversions, i.e., do not produce crossovers. The point after these *m* gene conversions is again a crossover, and the process of counting is repeated so that all points are treated. Having each distance *d* drawn from the distribution

$$\rho(d) = 2(m+1)e^{-2(m+1)d}$$

ensures that the density of crossovers on the bivalent is 2 (when *d* is measured in Morgan). Of course only those crossover points falling onto the bivalent are relevant. Finally, to go from the bivalent to a gamete, we assume as is almost always the case (Zhao et al. 1995) that there is no chromatid interference, so each crossover has a 50% chance of being transmitted to the considered gamete; the density of crossovers on the gametes after this "thinning" is thus 1.

The Gamma model: The Gamma model (Broman and Weber 2000; McPeek and Speed 1995) can be considered as an extension of the counting model, allowing for non integer values of *m*. One advantage over the counting model is that *m* is continuous, allowing for a broader range of interference effects. For chromosomes that are long >1.5 Morgan, the Gamma model is also nearly identical to the FIC framework (Falque et al. 2007) but has the advantage of being simpler. Before the thinning (on the bivalent), distances between crossovers in the counting model are distributed as

$$G_m(d) = 2(m+1)^{(m+1)} d^m \frac{e^{-2(m+1)d}}{m!}$$

as this is the law obtained when adding (m+1) variables with the distribution ρ(d). The Gamma model simply extends this to arbitrary *m*, not necessarily integer; the distance between two successive crossovers is then distributed as

$$G_m(d) = 2\nu^{\nu} d^{(\nu-1)} \frac{e^{-2\nu d}}{\Gamma(\nu)}$$

where Γ Gamma is Euler's Gamma function and ν plays the role of (*m*+1).

For integer ν, $\Gamma(\nu) = (\nu-1)\ !$

Recombination and consequences of missing data: When we use the Gamma model to simulate gamete production, we do so on a continuous chromosome; given the simulated crossover positions, we determine the alleles at the markers, thereby creating virtual genotypes in the same format as the experimental results. In both situations, the genotyping allows one to determine which intervals have recombined. However in some cases the marker allele is not determined in the experimental data set; we have missing data. This feature has to be taken into account in the analysis.

For the estimation of a recombinant frequency r_{AB}, we ignore gametes for which data is missing at either locus A or B. Similarly, when estimating the frequency of double recombinants for intervals (A,B) and (C,D), we ignore gametes for which data was missing at any of these 4 loci.

Consider now the frequency $P(k)$ of DH plants having k recombinant intervals, each such interval being delimited by adjacent loci. However, if there are missing data, we cannot be sure of the value of k, and so a cautious approach would be to ignore plants with any missing data. But given the many loci, we would loose a large amount of plants and thus have poor precision. So instead we estimate the number of recombination events by restricting attention to those loci that are informative; for instance if one has four adjacent loci, A, B, C, and D such that the interval (A,D) is recombinant but the data at B and C are missing, we say we have one recombination (even though it cannot be localized precisely). This procedure gives a lower bound on the actual number of recombinations, but in fact it is expected to almost always give the correct value if few of the loci have missing data. Indeed double recombinants in adjacent intervals are rare, first because the intervals are small and second because interference strongly suppresses them.

Consider now the construction of the histogram of distances between recombinant intervals. We worked only with gametes having two recombination events; we decided whether a given gamete had two recombinant events by applying the same criteria as for estimating $P(2)$. The gametes passing that test could have missing data, but we required the markers enclosing the two intervals under consideration to be fully informative. From these we produced the experimental histogram of distances between recombinant intervals by using the distance between the mid-points of the two intervals.

Finally, we compare the experimental data to what arises in simulated gametes, including if possible the effects of missing data. For each of the N experimental gametes, there may be specific loci for which the genotyping was not conclusive; we use these missing data as *templates* for the simulated genotypes. Specifically, in our procedure we first use the

Gamma model to produce 10^7 gametes with identified markers at each locus. Then for each simulated gamete G_5, we choose at random one of the experimental gametes G_e and transform the haplotype of G_5 so that the pattern of missing data is the same as in G_e. Thus if G_e has a missing data on the 5^{th} locus, we label as missing the data on that locus in G_5. In that way, the missing data in the experiments and in the simulations have the same statistics, making the comparison of theory and experiment more direct.

Significance of male-female differences in interference strength: The hypothesis H_0 to be tested is that the *same* interference parameter describes the male and female data sets. We test H_0 by the bootstrap method based on data reshuffling as follows.

The experimental data consists of N_m gametes produced during male meiosis and N_f during female meiosis. Analysis of these two data sets leads to the fitted values m^*_m and m^*_f for the interference parameter under the Gamma model. Now we shuffle these two data sets 10^5 times. For each shuffle, elements are exchanged between the two sets at random, maintaining a first data set of N_m gametes and a second of N_f gametes, but with no correlation with the male or female nature of the gametes. For each (mixed) set, we perform a fit to the Gamma model, leading to say the values m_1 and m_2. The difficulty is that the fitting requires handling gametes with different maps; for each map and each value of the putative interference parameter m, there is a theoretical distribution of distances between recombinant intervals, and thus we can assign a likelihood or score to any gamete. Given the list of 10^5 pairs (m_1, m_2), the bootstrap method specifies the p-value associated with H_0 as the fraction of these pairs for which $|m_1-m_2| > |m^*_m-m^*_f|$.

Résumé

Les crossing-overs (CO) sont indispensables dans la création variétale pour permettre l'introgression de régions d'intérêt dans les variétés agronomiques d'espèces cultivées telles que le blé tendre (*Triticum aestivum* L.). Afin d'évaluer l'impact des facteurs qui influencent leur formation, nous avons entrepris une caractérisation fine de leur distribution sur le plus grand chromosome (chromosome 3B, 995 Mb) du blé tendre en s'appuyant sur la carte physique récemment développée et le séquençage de quelques régions de plusieurs mégabases.

La comparaison entre une carte génétique dense (102 marqueurs) et une carte physique de délétion montre que 77 % des CO sont présents dans les régions distales couvrant seulement 25 % du chromosome. La comparaison de différentes cartes génétiques montre de plus que cette distribution est conservée entre populations avec cependant des différences de taux de CO locaux entre populations mais également entre méiose mâle et femelle. Cette distribution est influencée par une interférence positive forte à des distances inférieures à 10 cM. Cependant, les faibles fréquences de CO observées au sein des régions proximales restent inexpliquées. En effet, la faible augmentation du taux de CO observée au sein des régions proximales placées en position distale suggère que la position proximale de ces régions sur le chromosome ne semble pas responsable de leur faible fréquence de CO. De plus, nous avons montré que ces faibles fréquences ne seraient pas non plus dues à une divergence de séquence entre chromosomes homologues au sein des régions proximales, la fréquence de CO étant toujours aussi faible au sein de celles-ci entre deux chromosomes homozygotes.

En revanche, l'analyse à l'échelle d'une région séquencée de 3.1 Mb indique que les fréquences de CO importantes sont fortement corrélées avec la présence de gènes. L'inhibition de la formation des CO au sein des régions proximales pourrait ainsi s'expliquer par la présence de gènes en quantité moins importante dans ces régions comparées aux régions distales.

Abstract

Crossovers (CO) are necessary in breeding to introgress regions of interest in agronomic varieties from cultivated species such as bread wheat (*Triticum aestivum* L.). In order to evaluate the impact of factors influencing CO formation, we started the characterisation of their distribution on the largest wheat chromosome (chromosome 3B, 995 Mb) using the recently developed physical map as well as some sequenced regions of few megabases.

Comparison between a dense genetic map (102 markers) and a physical deletion map shows that 77 % of CO are present in the distal regions covering 25 % of the chromosome. Comparison between different genetic maps indicates that this distribution is conserved despite local differences of CO ratio between different populations and between male and female meiosis. Positive interference impacts significantly CO distribution at distances less than 10 cM on chromosome 3B. However, factors involved in the inhibition of CO formation in the proximal regions remain unknown. Positioning proximal regions in distal region of the chromosome increase only weakly the frequency of CO, suggesting that the low frequency observed in the proximal regions is not due to their proximal position. Moreover, sequence divergence between the two homologous chromosomes is not involved in the low frequency observed in the proximal regions since homozygous regions do not exhibit more CO than heterozygous ones.

On the contrary, analyses at the level of a 3.1 Mb sequenced region show that high CO frequencies correlate well with gene occurrence. Low CO frequency observed in the proximal regions could thus be explained by a low gene density.