
Abstract: The evolution of recombination in self-fertilising species: theoretical approach and genomic study of linkage disequilibrium between deleterious mutations in several Angiosperm species

Genetic recombination is a building block of the sexual cycle of eukaryotes, and is often considered as one of the main benefits of sexual reproduction. It is the consequence of crossovers occurring during meiosis, enabling the formation of recombinant gametes. However, the number and position of these crossovers along the genome varies greatly between species, as well as between individuals within the same species. Despite the fact that mechanical constraints related to the segregation of chromosomes during meiosis can impose a minimum and maximum number of crossovers, several empirical results show that recombination rates can evolve rapidly. Interestingly, higher recombination rates are often observed in self-fertilising hermaphroditic species compared with outcrossing species, suggesting that recombination is more strongly favoured in selfing species. The first part of the thesis involved developing theoretical models to better understand the effect of selfing on the evolution of recombination. A first model considered the evolution of a 'modifier' gene affecting the average number of crossovers per chromosome, in the presence of deleterious mutations occurring along the chromosomes. Analytical approximations and simulation results show that, under realistic parameter values, selection for recombination is generally stronger in selfing species; moreover, this selection is mainly due to the negative linkage disequilibrium between deleterious mutations generated by the Hill-Robertson effect (a stochastic effect related to finite population size). A second simulation model explored more realistic scenarios concerning the genetic architecture of recombination rate variation, as well as the distribution of crossovers along the chromosomes. The results show, in particular, that higher recombination rates can be maintained when this distribution is not uniform, or when the effect of recombination modifiers is restricted to a portion of the chromosome. The second part of the thesis focused on an empirical estimation of the linkage disequilibrium between deleterious mutations (an important component of selection for recombination), using genomic data from natural populations of the outcrossing plant *Capsella grandiflora*, as well as the highly selfing plants *Arabidopsis thaliana* and *Capsella orientalis*. This study highlights several methodological biases that can generate positive linkage disequilibrium between deleterious mutations. The first source of bias arises when the analysis is restricted to low-frequency mutations. Indeed, mutations with similar frequencies tend to be in positive linkage disequilibrium. A second source of bias may be generated by the presence of duplications present in the polymorphic state in the sample and absent from the reference genome. The positive linkage disequilibrium between deleterious mutations observed in previous studies (in particular in *C. grandiflora*) is probably largely due to these biases. In the selfing species *A. thaliana* (in which duplications present in the polymorphic state are easier to detect) we nevertheless observed positive linkage disequilibrium between deleterious mutations, which may be caused by epistatic effects or by the strong spatial structure observed in this species. Finally, the implications of these various theoretical and empirical results are discussed, in order to identifying prospects for future works.

Résumé : L'évolution de la recombinaison chez les espèces autogames : approche théorique et étude génomique du déséquilibre de liaison entre mutations délétères chez plusieurs espèces d'Angiospermes.

La recombinaison génétique est un élément constitutif du cycle sexué des Eucaryotes, et est souvent considérée comme l'un des principaux avantages de la reproduction sexuée. Elle est la conséquence de crossing-over se produisant pendant la méiose, permettant la formation de gamètes recombinants. Or, le nombre et la position de ces crossing-over le long du génome est très variable entre espèces, ainsi qu'entre individus au sein d'une même espèce. Malgré le fait que des contraintes mécaniques liées à la ségrégation des chromosomes pendant la méiose peuvent imposer un nombre minimal et maximal de crossing-over, plusieurs résultats empiriques montrent que les taux de recombinaison peuvent évoluer rapidement. De façon intéressante, on observe souvent des taux de recombinaison plus élevés chez les espèces hermaphrodites autogames par rapport aux espèces allogames, suggérant que la recombinaison est plus fortement favorisée chez les autogames. Une première partie de la thèse a consisté à développer des modèles théoriques permettant de mieux comprendre cet effet de l'autofécondation sur l'évolution de la recombinaison. Un premier modèle a considéré l'évolution d'un gène "modificateur" affectant le nombre moyen de crossing-over par chromosome, en présence de mutations délétères se produisant le long des chromosomes. Les approximations analytiques ainsi que les résultats de simulation montrent que, sous des valeurs de paramètres réalistes, la sélection pour la recombinaison est généralement plus forte chez les espèces autogames ; par ailleurs, cette sélection est principalement due au déséquilibre de liaison négatif entre mutations délétères généré par l'effet Hill-Robertson (un effet stochastique lié à la taille finie des populations). Un deuxième modèle de simulation a exploré des scénarios plus réalistes concernant l'architecture génétique de la variation des taux de recombinaison, ainsi que la distribution des crossing-over le long des chromosomes. Les résultats montrent notamment que des taux de recombinaison plus élevés peuvent être maintenus lorsque cette distribution n'est pas uniforme, ou lorsque l'effet des modificateurs de recombinaison est restreint à une portion de chromosome. La deuxième partie de la thèse a porté sur une estimation empirique du déséquilibre de liaison entre mutations délétères (une composante importante de la sélection pour la recombinaison), en utilisant des données génomiques issues de populations naturelles de la plante allogame *Capsella grandiflora*, ainsi que des plantes fortement autogames *Arabidopsis thaliana* et *Capsella orientalis*. Cette étude met en avant plusieurs biais méthodologiques pouvant générer du déséquilibre de liaison positif entre mutations délétères. Une première source de biais se produit lorsque l'analyse est restreinte aux mutations présentes en faible fréquence. En effet, des mutations en fréquences similaires ont tendance à être en déséquilibre de liaison positif. Une deuxième source de biais peut être générée par la présence de duplications présentes à l'état polymorphe dans l'échantillon, et absentes du génome de référence. Ainsi, le déséquilibre de liaison positif entre mutations délétères observé lors de précédents travaux (notamment chez *C. grandiflora*) résulte probablement en grande partie de ces biais. Chez l'espèce autogame *A. thaliana* (chez qui les duplications présentes à l'état polymorphe sont plus faciles à détecter) on observe néanmoins du déséquilibre de liaison positif entre mutations délétères, pouvant être causé par des effets épistatiques ou par la forte structure spatiale observée chez cette espèce. Les implications de ces différents résultats théoriques et empiriques sont enfin discutées afin d'en dégager des perspectives pour des travaux futurs.