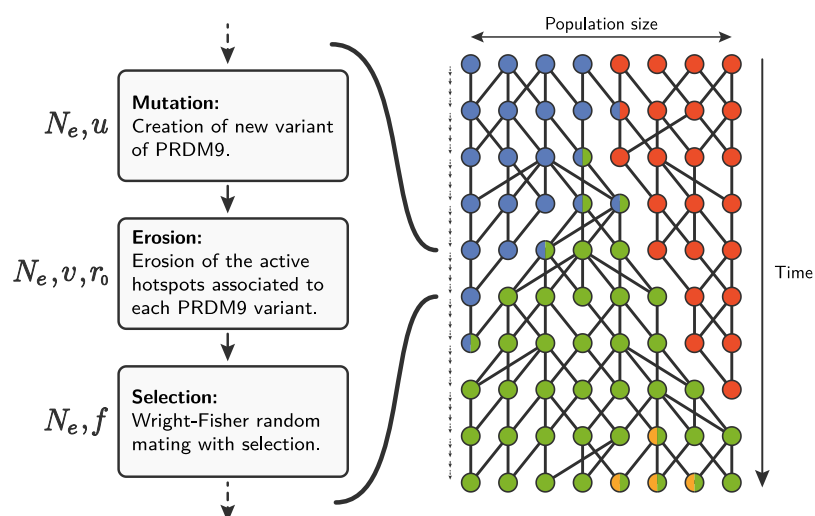


Intitulé du sujet : Modélisation de la dynamique évolutive de la recombinaison chez les mammifères.

Encadrants: Nicolas LARTILLOT (nicolas.lartillot@univ-lyon1.fr), Laurent DURET (laurent.duret@univ-lyon1.fr), LBBE - UMR 5558.

Lieu du stage: Laboratoire de Biométrie et Biologie Evolutive (LBBE), 43 boulevard du 11 novembre 1918, Villeurbanne



Résumé du projet:

Par bien des aspects, l'évolution de la recombinaison méiotique reste encore une des grandes énigmes de la biologie évolutive. Pour résoudre cette énigme, il est important de mieux connaître les mécanismes moléculaires de la recombinaison. Sur ce point, les recherches récentes effectuées chez l'humain et la souris ont mis en évidence plusieurs faits marquants. D'une part, la recombinaison est concentrée dans des points chauds, dont la durée de vie est très courte à l'échelle évolutive -- les paysages de recombinaison évoluent donc très rapidement [1]. D'autre part, les points chauds de recombinaison coïncident avec les sites de fixation d'une protéine à doigts de zinc, codée par le gène **PRDM9**, qui est sous forte sélection positive et évolue lui-aussi très rapidement [2]. Un modèle de type Reine Rouge a été proposé pour expliquer cette dynamique évolutive très particulière, faisant intervenir deux forces évolutives opposées: un **biais de transmission méiotique** en défaveur des sites de fixation de PRDM9, faisant disparaître les points chauds de recombinaison; auquel s'oppose une **sélection positive** en faveur de nouveaux allèles de PRDM9, reconnaissant de nouveaux sites de fixation et restaurant ainsi la recombinaison à travers le génome [3].

Ce modèle de Reine Rouge a été récemment exploré par l'équipe d'accueil [4], au moyen de **simulations informatiques sous un modèle de Wright-Fisher** [5]. Ce travail, effectué lors d'un stage de master, a permis de mieux comprendre et de quantifier le fonctionnement de la Reine Rouge, en fonction de plusieurs paramètres clés: taux de mutation, taille de la population, et force de la sélection pour maintenir la recombinaison. De telles approches par modélisation sont cruciales pour mieux comprendre à la fois la régulation moléculaire et la dynamique évolutive de la recombinaison. Toutefois, ce travail est encore préliminaire. En particulier, le modèle ne formalise pas correctement la fonction moléculaire de la recombinaison, alors même que cette fonction détermine la sélection agissant sur la recombinaison, et donc aussi sur PRDM9. Les connaissances actuelles suggèrent que la recombinaison possède en réalité deux fonctions distinctes: en plus de son rôle dans le brassage de l'information génétique, elle est aussi requise pour l'appariement et la ségrégation des chromosomes

homologues au cours de la méiose. Or, des résultats récents suggèrent que PRDM9 joue un rôle central dans le processus d'appariement des chromosomes [6]. Ces résultats ont abouti à un modèle moléculaire quantitatif [7], qui fournit tous les ingrédients nécessaires pour estimer la fécondité d'un individu en fonction de son génotype PRDM9 et de la séquence de son génome.

Objectifs:

L'objectif de ce stage de Master sera de poursuivre le travail de simulation entamé dans l'équipe d'accueil autour du modèle de la Reine Rouge, et d'utiliser les résultats empiriques récents en vue de modéliser plus finement le rôle de PRDM9 dans la méiose. Plus précisément, on intégrera le modèle mécaniste proposé en [7] avec le simulateur de Wright Fisher déjà développé dans l'équipe d'accueil [5], afin de tester si la fonction de PRDM9 dans l'appariement des chromosomes suffit à expliquer la dynamique évolutive de la recombinaison telle qu'observée chez l'humain et la souris.

Résultats escomptés:

En partant du programme de simulation de la Reine Rouge de la recombinaison est déjà développé dans l'équipe d'accueil, la démarche de travail pourra se dérouler selon les étapes suivantes:

- adapter le modèle de simulation existant; en particulier, formaliser et programmer la fonction de fitness suggérée par [7], permettant de calculer la fécondité d'un individu en fonction de ses allèles PRDM9 et de la distribution des sites de fixation de ces allèles de PRDM9 dans son génome.
- faire tourner le simulateur à travers une large gamme de valeurs de paramètres;
- observer le comportement des populations simulées (en termes de diversité génétique, de fitness, de distribution de la recombinaison dans leurs génomes, etc), en fonction des paramètres de la simulation (taux de mutation, taille de population, etc);
- dégager des tendances, mettre à jour des lois d'échelle entre quantités mesurables et paramètres du modèle (par exemple, est ce que la fitness de la population, ou le taux moyen de recombinaison à travers les génomes, augmentent ou diminuent en fonction de la taille de la population?);
- proposer des arguments simples pour expliquer de façon qualitative les relations ainsi observées (par ex., pourquoi une telle relation entre taux de recombinaison et taille de population?)
- ajuster le modèle aux données empiriques disponibles chez l'homme et la souris;
- en tirer des conclusions sur l'adéquation du modèle ou son rejet par les données empiriques.

Profile et compétences recherchées

L'étudiant(e) doit montrer un intérêt particulier pour les questions de génétique évolutive, aussi bien sur ses aspects empiriques que théoriques. Une connaissance de base en programmation (en python ou en C++) sera par ailleurs utile (mais sans forcément avoir des compétences poussées en maths/info). Le stage sera alors l'occasion d'acquérir et d'améliorer des compétences en bioinformatique et d'approfondir sa compréhension des concepts et des théories de la génétique évolutive et de la génétique des populations.

Modeling the evolutionary dynamics of recombination landscapes in mammals.

Context: The evolution of meiotic recombination is still one of the big enigmas in evolutionary biology. To solve this mystery, it is important to first acquire a better knowledge of the molecular regulation of recombination. In this direction, recent research conducted in human and in the mouse has provided important new insights. First, recombination is typically concentrated in hot spots, which are short-lived over evolutionary timescales (~ 1 Myr or less) -- thus, recombination landscapes evolve very rapidly [1]. Second, recombination hot spots coincide with the binding sites of a zinc-finger protein, encoded by the **PRDM9** gene, which is under strong positive selection and thus also evolves very rapidly [2]. A **Red Queen model** has been proposed to explain this very peculiar evolutionary dynamics. The model involves two opposing evolutionary forces: a **meiotic transmission bias** against PRDM9 binding sites, leading to the extinction of recombination across the

genome; which is then counteracted by **strong positive selection** in favor of new PRDM9 alleles, recognizing new binding sites and thus recombination genome wide [3].

The Red Queen model of recombination evolution has been recently explored by the host team [4], using computer simulations under a Wright-Fisher model [5]. This work, conducted by a master student, has led to a better quantitative understanding of the evolutionary arms race between PRDM9 and the recombination hot spots, as a function of several key parameters: mutation rate, population size and the strength of selection for maintaining recombination. Such simulation approaches are essential in order to get a better understanding of the regulation and of the evolutionary dynamics of recombination. However, the work accomplished thus far is still preliminary. In particular, the model does not correctly formalize the molecular function of recombination, yet this function is what determines the selection acting on recombination, and thus also on PRDM9. Current knowledge suggests that recombination has in fact two functions: it dissipates linkage between neighboring loci, but it is also essential for **chromosome pairing and segregation** during meiosis. Importantly, recent results suggest that PRDM9 plays a key role in chromosome pairing [6]. These results have led to a detailed and quantitative model [7], giving all the necessary ingredients to estimate the fertility of an individual, as a function of its PRDM9 genotype and its genomic background.

Objectives: The aim of this internship is to continue the work done in the host team on the Red Queen of recombination, by incorporating in the simulator a detailed and mechanistic model of the role of PRDM9 in chromosome pairing, based on the recently published results on the subject.

Methods and expected results: A simulation program of the Red Queen model of recombination is already available in the host team. From there, the following research strategy is suggested:

- use this program as a starting point, and incorporate the mechanistic model of the role of PRDM9 in chromosome pairing and synapsis proposed in [7];
- running the simulation program across a broad range of parameter regimes, and observing the behavior of simulated populations (in terms of its genetic diversity, fitness, mean recombination activity, etc), as a function of the parameters of the model (mutation rate, effective population size, etc) should reveal the broad trends and the scaling behavior of the model (for instance: does the mean fitness or the mean recombination activity over the genome increase or decrease as a function of effective population size?, etc)
- simple arguments can then be derived to explain these scaling relations;
- finally, a qualitative fit of the model to empirical data available in human and in the mouse will be attempted, in order to test the model for its empirical adequacy.

Expected profile and skills of the candidate: The candidate should have a particular interest for questions of evolutionary genomics, both from an empirical and a theoretical perspective. Basic programming skills (in python or C++) will also be useful. The internship will be an occasion to acquire and improve one's skills in bioinformatics, and one's understanding of the concepts and the theory in evolutionary and population genetics.

References

1. Myers S, Bottolo L, Freeman C, McVean G, Donnelly P. A fine-scale map of recombination rates and hotspots across the human genome. *Science*. 2005 Oct 14;310(5746):321-4.
2. Baudat F, Buard J, Grey C, Fledel-Alon A, Ober C, Przeworski M, Coop G, de Massy B. PRDM9 is a major determinant of meiotic recombination hotspots in humans and mice. *Science*. 2010 Feb 12;327(5967):836-40.
3. Myers, S., Bowden, R., Tumian, A., Bontrop, R. E., Freeman, C., MacFie, T. S., et al. (2010). Drive against hotspot motifs in primates implicates the PRDM9 gene in meiotic recombination. *Science*, 327(5967), 876–879.
4. Lesecque Y, Glémin S, Lartillot N, Mouchiroud D, Duret L. The red queen model of recombination hotspots evolution in the light of archaic and modern human genomes. *PLoS Genet*. 2014 Nov 13;10(11):e1004790.

5. Latrille, T., Duret, L., & Lartillot, N. (2017). The Red Queen model of recombination hot-spot evolution: a theoretical investigation. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 372(1736).
6. Davies, B., Hatton, E., Altemose, N., Hussin, J. G., Pratto, F., Zhang, G., et al. (2016). Re-engineering the zinc fingers of PRDM9 reverses hybrid sterility in mice. *Nature*, 530(7589), 171–176.
7. Gregorová, S., Gergelits, V., Chvatalova, I., Bhattacharyya, T., Valiskova, B., Fotopulosova, V., et al. (2018). Modulation of Prdm9-controlled meiotic chromosome asynapsis overrides hybrid sterility in mice. *eLife*, 7.