



STAGE DE RECHERCHE M2 ECOLOGIE EVOLUTION GENOMIQUE

Rentrée 2018

Identification et étude des gènes sur le chromosome Y des plantes

Université Claude Bernard Lyon 1

Laboratoire de Biométrie et Biologie Evolutive UMR CNRS 5558 Villeurbanne

Encadrants : Hélène Badouin (helene.badouin@univ-lyon1.fr), Jos Käfer (jos.kafer@univ-lyon1.fr)

Contexte

Parmi les 350000 plantes à fleurs, une minorité (15600 espèces) ont des sexes séparés, ce qu'on appelle « dioécie » (Renner, 2014). Parmi celles-ci, des chromosomes sexuels ont été identifiés chez quelques dizaines d'espèces (Ming et al., 2011). Les chromosomes sexuels des plantes évoluent dans l'ensemble de la même façon que ceux des animaux : l'arrêt de recombinaison fait que la sélection est moins efficace pour le chromosome Y, ce qui conduit à la dégénérescence de ce chromosome, accompagnée d'une perte de gènes. Ce processus peut même conduire à la perte totale du Y. Les chromosomes sexuels des plantes sont en général beaucoup plus jeunes que ceux des animaux (ce qui est lié à l'origine récente de la dioécie chez la plupart des plantes). Cela donne l'opportunité d'étudier les phases précoces de l'évolution des chromosomes sexuels. Notamment, une des observations faites chez les plantes est l'agrandissement initial du chromosome Y par accumulation d'éléments transposables. Ainsi, un débat scientifique a lieu actuellement sur la question de savoir si cet agrandissement du chromosome Y constitue une phase précoce commune aux plantes et aux animaux (plus facile à observer chez les plantes où les chromosomes sexuels sont plus jeunes), ou si cela constitue une spécificité des plantes liée à leur cycle de vie. En effet, l'alternance des phases diploïdes et haploïdes pourrait empêcher ou freiner la dégénérescence du Y, du fait de la sélection opérant dans la phase haploïde.

Pour répondre à cette question il est nécessaire de disposer de plus de données sur les chromosomes sexuels des plantes. L'ADN de ces chromosomes est difficile à assembler, à cause de leur plus faible couverture dans des projets de séquençage des génomes (la couverture est réduite de moitié pour les chromosomes Y), et à cause d'un plus grand contenu en ADN non codant (répétitions, éléments transposables) sur le Y. Pour ces raisons, nous avons développé dans l'équipe une stratégie basée sur le séquençage des transcriptomes (RNAseq), qui nous donne directement accès aux gènes, et permet d'identifier les paires de gènes XY et les gènes absents du chromosome Y en analysant la répartition des allèles et génotypes chez les individus des deux sexes (Muyle et al. 2016).

Le but de ce stage est d'identifier et d'étudier les gènes spécifiques du chromosome Y, qui n'entrent pas encore dans le champ de nos méthodes de détection. Il s'agit de gènes présents uniquement sur le chromosome Y, ou dont l'allèle Y est trop divergent de l'allèle X pour être identifié par comparaison des séquences d'ARN. Ces gènes pourraient avoir des rôles importants dans le déterminisme du sexe mâle, comme c'est le cas du gène SRY chez les mammifères. La méthode à mettre en place repose sur la comparaison des transcriptomes des mâles et des femelles. Les gènes spécifiques du transcriptome mâle sont soit des gènes exprimés uniquement chez les mâles, soit des gènes spécifiques du Y.

Pour distinguer ces deux cas, il faut s'appuyer sur la comparaison avec les séquences les plus proches dans le transcriptome femelle (en particulier sur la comparaison séquences d'acides aminés prédites, plus conservées que les séquences nucléiques), et sur le fait que les séquences du Y ne peuvent pas être hétérozygotes.

Pour cela, l'étudiant ou étudiante aura deux jeux de données à sa disposition : les transcriptomes femelle et mâle de *Coccinia grandis* (Cucurbitaceae), une espèce avec un chromosome Y beaucoup plus grand que le X, et des données RNAseq de *Cannabis sativa* (Cannabaceae), une espèce avec les chromosomes sexuels qui sont parmi les plus vieux connus des plantes. Pour *Coccinia*, les transcriptomes sont déjà assemblés, mais personne n'a encore séquencé le génome, tandis que pour *Cannabis*, nous disposons du génome et du transcriptome d'une femelle (Van Bakel et al., 2011), mais pas d'un mâle. Pour les deux espèces, un certain nombre de gènes XY ont déjà été identifiés par des travaux récents de l'équipe ; ces gènes serviront dans un premier temps à tester la méthodologie mise en place dans ce stage. Après cette validation, la détection de nouveaux gènes permettra d'étudier les questions encore ouvertes sur les gènes Y-spécifiques.

Profil Requis

L'étudiant(e) a déjà des compétences de base dans les domaines suivants, et est prêt(e) à approfondir ses connaissances : évolution moléculaire et génétique des populations, traitement de grands jeux de données, manipulation des outils bioinformatiques sous linux (ligne de commande, utilisation d'un cluster de calcul intensif)

Bibliographie associée

Ming, R., A. Bendahmane et S. S. Renner (2011). "Sex chromosomes in land plants."

In : Annu. Rev. Plant Biol. 62, p. 485-514.

Muyle, A., J. Käfer, N. Zemp, S. Mousset, F. Picard et G. A. Marais (2016).

"SEX-DETECTOR: A Probabilistic Approach to Study Sex Chromosomes in Non-Model Organisms."

In : Genome Biol. Evol. 8.8, p. 2530-2543.

Renner, S. S. (2014). "The relative and absolute frequencies of angiosperm sexual systems: dioecy, monoecy, gynodioecy, and an updated online database."

In : Am. J. Bot. 101.10, p. 1588-1596.

Van Bakel, H., J. M. Stout, A. G. Cote, C. M. Tallon, A. G. Sharpe, T. R.

Hughes et J. E. Page (2011). "The draft genome and transcriptome of Cannabis sativa."

In : Genome Biol. 12.10, R102.