



STAGE DE RECHERCHE M2 ECOLOGIE EVOLUTION GENOMIQUE Rentrée 2018

Un scan de sélection pour identifier des gènes sous sélection équilibrante chez le tournesol

Université Claude Bernard Lyon 1

Laboratoire de Biométrie et Biologie Evolutive UMR CNRS 5558 Villeurbanne

Encadrante : Hélène Badouin (helene.badouin@univ-lyon1.fr)

Contexte

Comprendre comment et pourquoi les espèces évitent l'autofécondation est une question importante en biologie. Chez de nombreuses plantes à fleurs, l'autofécondation est empêchée par des mécanismes d'auto-incompatibilité génétique (AI), où l'interaction entre les protéines spécifiques du pistil et du pollen entraîne le rejet de l'autopollen¹. L'AI est généralement contrôlé par un locus unique, multi-allélique (locus S) qui évolue dans le cadre d'une forte sélection équilibrante, c'est-à-dire une sélection pour le maintien de la diversité génétique². L'acquisition et la perte multiples d'AI ont eu lieu au cours de l'évolution des Angiospermes, mais les gènes sous-jacents (gènes S) sont inconnus dans la plupart des familles, limitant grandement notre compréhension des forces évolutives et des mécanismes proximaux qui régissent ces transitions³. Identifier les gènes S est également d'un intérêt majeur pour les programmes de sélection dans les plantes domestiquées. Les Astéracées constituent la plus grande famille d'Angiospermes (> 25 000 espèces) et comprennent des espèces d'intérêt agricole telles que la laitue, la carotte, l'artichaut ou le tournesol. Plus de la moitié des Astéracées sont auto-incompatibles, mais les gènes S n'ont encore été identifiés chez aucune espèce d'Asteraceae. Le tournesol commun *Helianthus annuus* est une plante oléagineuse majeure (~ 10% de la production mondiale). Alors que la plupart des lignées domestiquées sont auto-compatibles, les tournesols sauvages sont auto-incompatibles, ce qui constitue un défi pour les programmes de sélection visant à exploiter le réservoir de la diversité sauvage pour faire face au changement climatique. Le tournesol est actuellement l'un des très rares modèles de plantes AI dans lequel les ressources et les outils nécessaires pour une étude intégrative de la découverte des gènes S sont disponibles. Un locus de caractères quantitatifs pour l'IS a été identifié⁴, et un génome de référence de haute qualité a été récemment obtenu (Badouin et al., 2017, Nature⁵). Le projet GENESIS (Exploration génomique des gènes d'auto-incompatibilité chez le tournesol) utilise une combinaison d'approches génomiques et génétiques pour identifier et valider les gènes causant l'auto-incompatibilité chez le tournesol.

Objectifs du stage

Le but de cette formation est d'effectuer un scanner du génome⁶ pour équilibrer la sélection dans un ensemble de données sur la génomique des populations de tournesols sauvages. Nous avons reséquéncé les génomes complets de 36 individus sauvages et d'un groupe externe en utilisant la technologie Illumina paired end. La cartographie des lectures courtes sur un génome de référence de haute qualité (mapping) et la détection des variants (SNP calling) ont été réalisés. L'étudiant ou étudiante utilisera l'ensemble des données de mapping et SNP calling pour: 1) Identifier les gènes montrant une couverture suffisante pour être analysé; 2) Calculer des statistiques descriptives de génomique de population; 3) Effectuer des tests de sélection équilibrante (tests Tajima D, Fay et Wu H). Cela sera fait avec les bibliothèques scientifiques existantes pour la génomique des populations (bio ++, libsequence ou egglib⁷). En fonction de son intérêt et de ses compétences, l'étudiant pourrait également participer à l'analyse d'un jeu de données en transcriptomique à lectures longues pour lequel un pipeline bioinformatique *ad hoc* devra être développé.

Compétences

L'étudiant ou étudiante acquerra ou améliorera des compétences en bioinformatique (utilisation d'un environnement Unix, d'un cluster informatique partagé, langage script Python ou Perl, R), analyse de données NGS et de génomique des populations.

Bibliographie associée

1. Nasrallah, J. B. Recognition and Rejection of Self in Plant Reproduction. *Science* **296**, 305–308 (2002).
2. Charlesworth, D., Vekemans, X., Castric, V. & Glémin, S. Plant self-incompatibility systems: a molecular evolutionary perspective. *New Phytol.* **168**, 61–69 (2005).
3. Allen, A. M. & Hiscock, S. J. Evolution and Phylogeny of Self-Incompatibility Systems in Angiosperms. in *Self-Incompatibility in Flowering Plants* 73–101 (Springer Berlin Heidelberg, 2008). doi:10.1007/978-3-540-68486-2_4
4. Gandhi, S. D. *et al.* The self-incompatibility locus (S) and quantitative trait loci for self-pollination and seed dormancy in sunflower. *Theor. Appl. Genet.* **111**, 619–629 (2005).
5. **Badouin**, H. *et al.* The sunflower genome provides insights into oil metabolism, flowering and Asterid evolution. *Nature* **546**, 148–152 (2017).
6. **Badouin**, H. *et al.* Widespread selective sweeps throughout the genome of model plant pathogenic fungi and identification of effector candidates. *Mol. Ecol.* **26**, 2041–2062 (2017).
7. Mita, S. D. & Siol, M. EggLib: processing, analysis and simulation tools for population genetics and genomics. *BMC Genet.* **13**, 27 (2012).