



LBBE



Sujet de stage de M2

Annotation et analyse évolutive des gènes responsables du parfum de la lavande laineuse

Ce sujet de stage de M2 s'intègre dans le projet LaLaGéVoTé, qui vise à séquencer, assembler et annoter le génome de la lavande laineuse et à étudier l'évolution des voies de biosynthèse des terpènes volatils. Ce projet est le fruit d'une collaboration, au sein de la FR BioEnviS, entre le Laboratoire de Biotechnologies Végétales Appliquées aux Plantes Aromatiques et Médicinales (LBVpam, Saint Etienne), la plateforme de bio-informatique PRABI-AMSB (UCBL, Lyon), et le Laboratoire de Biométrie et Biologie Évolutive (LBBE, Lyon).

Les travaux de recherche du LBVpam visent à développer des connaissances fondamentales sur la biologie, la biochimie, l'écologie et l'évolution des composés organiques volatils (COV) responsables du parfum des plantes et constituants des huiles essentielles. Les COV sont produits par des enzymes du métabolisme spécialisé. Parmi les COV, les terpénoïdes volatils sont largement distribués dans l'arbre du vivant. Chez les plantes, ils sont des médiateurs de la communication chimique intra- et inter-organismes et participent en particulier aux mécanismes de défense ou de pollinisation. Chez certaines plantes, des regroupements (ou «clusters») de gènes du métabolisme des terpénoïdes, comprenant des gènes de terpène synthases, ont été révélés dans les génomes. En particulier, le génome de la lavande vraie *Lavandula angustifolia*, endémique des montagnes sèches du sud de la France, comporte de nombreuses duplications de gènes impliqués dans le métabolisme des terpénoïdes. Plus de cent séquences de terpène synthases ont été identifiées dans ce génome [1].

Le projet LaLaGéVoTé vise à effectuer une étude évolutive fine de la biosynthèse des terpénoïdes chez des espèces proches de la lavande vraie. Une étape cruciale de ce projet est de séquencer, assembler et annoter le génome de la lavande laineuse (*Lavandula lanata*), une espèce endémique des montagnes du sud de l'Espagne mais proche phylogénétiquement de *Lavandula angustifolia*. Un premier assemblage du génome, réalisé grâce à des données de séquençage Nanopore, est à ce jour disponible. Des données de séquençage du transcriptome des fleurs et des feuilles de la lavande laineuse ont également été produites, dans le but d'améliorer l'annotation des gènes et d'évaluer leurs patrons d'expression.

Dans le cadre de ce stage, nous proposons d'annoter et d'analyser les gènes impliqués dans le métabolisme des terpénoïdes dans le génome de *Lavandula lanata*. Pour réaliser cette annotation, nous nous baserons sur les annotations existantes des séquences de ces gènes dans le génome de *Lavandula angustifolia*, ainsi que sur des données de séquençage ciblé produites au LBVpam. Une première étape sera d'effectuer une annotation par homologie, en utilisant des outils bio-informatiques existants [2]. Dans un deuxième temps, nous proposons d'évaluer les patrons d'expression des gènes identifiés, en utilisant les données de séquençage de transcriptome (RNA-seq) disponibles dans l'équipe. Enfin, nous proposons de réaliser une

étude évolutive des gènes impliqués dans le métabolisme des terpénoïdes. Pour ce faire, des homologues seront prédits dans les génomes des espèces de la famille des *Lamiaceae* [3], leurs séquences seront alignées et des arbres phylogénétiques seront reconstruits.

Ce projet permettra à l'étudiant.e recruté.e de se familiariser avec l'utilisation d'outils bioinformatiques couramment utilisés dans le domaine (BLAST, DIAMOND, GeMoMa, kallisto, MUSCLE, PhyML,...). L'étudiant.e pourra également renforcer ses connaissances en programmation en Python, analyses statistiques et représentations graphiques avec R.

Il ou elle collaborera sur le projet avec Sandrine Moja et Aurélie Bony du LBVpam, Christine Oger du PRABI-AMSB et Anamaria Necsulea du LBBE et sera impliqué.e dans les trois étapes du travail. Il ou elle sera chargée de réaliser en autonomie les phylogénies des gènes qui auront été sélectionnés d'après leur niveau d'expression.

Les candidat.e.s devront justifier d'un bon niveau théorique et pratique en biologie et évolution et avoir des compétences en analyses et fouille de données sous R. Des compétences en bio-informatique seront un plus.

Pour candidater : envoyer un CV et une lettre de motivation à sandrine.moja@univ-st-etienne.fr

Bibliographie

- [1] Li, J. et al. The chromosome-based lavender genome provides new insights into Lamiaceae evolution and terpenoid biosynthesis. *Hortic Res* 8, 53 (2021).
- [2] Keilwagen, J., Hartung, F. & Grau, J. GeMoMa: Homology-based gene prediction utilizing intron position conservation and RNA-seq data. *Methods Mol Biol* 1962, 161–177 (2019).
- [3] Bornowski, N. et al. Genome sequencing of four culinary herbs reveals terpenoid genes underlying chemodiversity in the Nepetoideae. *DNA Research*, 27, dsaa016 (2020).