

Proposition de stage de Master 2 Bioinformatique

Sujet : Annotation du génome de la centaurée de la Clape

Résumé

L'annotation d'un génome permet de définir des unités fonctionnelles à partir de séquences d'ADN. L'annotation permet notamment de reconnaître les unités génomiques essentielles dans la physiologie de l'organisme. L'annotation est une étape dans l'étude d'un éventuel polymorphisme de ces unités génomiques, par exemple diversité synonyme/non synonyme entre unités.

L'objectif du stage proposé est d'annoter le génome de *Centaurea corymbosa*, la centaurée de la Clape. Cette plante, de la famille des Astéracées, est une espèce endémique du Massif de la Clape (Narbonne) qui présente des enjeux de conservation forts. L'espèce est étudiée depuis 1994 dans l'équipe et parmi les travaux actuels, nous nous intéressons à sa capacité de résistance au stress hydrique, enjeu majeur face au changement climatique.

Un génome de référence a déjà été construit à partir de bibliothèques HiFi sur PacBio. La taille du génome a été estimée à 750Mb par haplotype (avec $2n=18$). Une carte optique est en cours de construction par un prestataire (plate-forme Gentyan INRAE Clermont Ferrand). Récemment, un transcriptome pour 3 espèces proches (*Centaurea jacea*, *Centaurea solstitialis* et *Cyanus segetum*) a été publié (<https://doi.org/10.1111/jipb.13078>).

Le/la stagiaire devra dans un premier temps produire une synthèse bibliographique sur les outils les plus pertinents pour l'assemblage, trouver le génome annoté le plus proche phylogénétiquement dans les bases de données existantes et enfin annoter le génome de référence. Ce travail permettra de produire des statistiques descriptives sur le génome annoté.

Par ailleurs, des séquences génomiques sur 96 individus sont en cours de production dans le cadre d'un autre projet. Ces séquences pourront servir de base pour travailler sur la recherche d'un éventuel polymorphisme de quelques gènes d'importance dans la résistance au stress hydrique.

Compétences attendues : production de synthèse bibliographique, pipe-line d'annotation

Encadrement : il sera assuré conjointement par des membres de la plate-forme Bioinformatique de l'ISEM et par des membres de l'équipe EVODEMO.

Jos Käfer, évolution des génomes et diversité des plantes à fleur

Khalid Belkhir, bioinformatique

Benjamin Penaud, bioinformatique

Eric Imbert, biologie et écologie de la Centaurée de la Clape