

OFFRE DE STAGE

Période du stage : 1^{er} semestre 2026

Durée du stage : 5-6 mois

Nom du tuteur responsable du stage : G. Hernandez Raquet et S. Déjean

INTEGRATION DE DONNEES METAPROTEOMIQUES ET METAGENOMIQUES POUR LA COMPREHENSION DU FONCTIONNEMENT DE CONSORTIA MICROBIENS LIGNOCELLULOLYTIQUES

La biomasse lignocellulosique représente une ressource renouvelable très intéressante pour la production d'énergie et de synthons d'intérêt industriel. La transformation de la lignocellulose peut être effectuée par des consortia microbiens sélectionnés. Outre leur diversité enzymatique, les consortia microbiens présentent une diversité métabolique élevée ainsi qu'une forte tolérance à la variabilité propre de la biomasse lignocellulosique. Cependant, leur fonctionnement interne est encore mal caractérisé.

Dans notre équipe, des consortia issus de rumen bovin, très actifs pour la dégradation de lignocellulose et la production de molécules d'intérêts, ont été stabilisés lors de précédents travaux. Pour mieux comprendre leur fonctionnement, une approche multi-échelle a été réalisée. Les consortia ont été étudiés lors de cinétiques de dégradation de substrats prétraités, en analysant : (1) les paramètres macro-cinétiques (dégradation du substrat, accumulation des produits et activités enzymatiques clés), (2) la diversité microbienne (métagénomique amplicon des ADNr et ARNr 16S), et (3) les protéines exprimées (métaprotéomique 'shotgun'). L'intégration et l'analyse conjointe de ces données omiques est donc extrêmement prometteuse en termes de compréhension du fonctionnement des consortia lignocellulolytiques étudiés.

L'objectif de ce travail est de mener une analyse statistique intégrative des différents jeux de données -omiques cités précédemment. Dans un premier temps, pour se familiariser avec le contexte, le stagiaire pourra revenir sur l'analyse séparée des différents jeux de données (contrôle qualité, analyse différentielle, analyse exploratoire...). Dans un second temps, les méthodes statistiques à mettre en œuvre pour l'analyse intégrative pourront être l'analyse des corrélations canoniques, la régression PLS et son extension à l'analyse discriminante (PLS-DA) ainsi que des méthodes d'apprentissage. Les développements se feront avec le logiciel R à l'aide de plusieurs packages dédiés à l'analyse de données -omiques (limma, mixOmics et d'autres disponibles sur Bioconductor).

Compétences requises : le stagiaire devra posséder une expérience significative en bioinformatique et en programmation R et une bonne connaissance des méthodes statistiques standards (tests statistiques, ANOVA, régression, analyse descriptive...). M2 en bioinformatique, ingénierie mathématique ou statistique.

Lieu : Ce stage s'effectue au Toulouse Biotechnology Institute – Bio & Chemical Engineering (TBI) et l'Institut de Mathématiques de Toulouse (IMT). TBI est un environnement pluridisciplinaire qui permettra à l'étudiant de se sensibiliser à des thématiques autour des procédés de traitement des déchets, de la biotechnologie, du génie microbien et de la biologie moléculaire. Le stagiaire pourra s'appuyer sur les compétences en statistiques de l'IMT.

Gratification de stage prévue : ~650€/mois

Contact : Guillermina HERNANDEZ-RAQUET

e-mail : hernandg@insa-toulouse.fr

Sébastien DEJEAN

e-mail : sebastien.dejean@math.univ-toulouse.fr