

Analyse de données de séquençage en « noyau unique » pour comprendre l'effet de signaux bactériens sur le fonctionnement de la racine d'une plante légumineuse modèle

Contexte :

Les plantes sont des organismes sessiles dont le système racinaire, important pour leur ancrage et leur nutrition, présente un développement très plastique qui leur permet de s'adapter à un environnement changeant. Par exemple, les racines de la plante modèle *Medicago truncatula* peuvent modifier leur développement en réponse à des signaux produits par des bactéries et des champignons du sol. L'équipe de S. Bensmihen (LIPME) cherche à comprendre comment ces signaux microbiens peuvent influencer le développement racinaire de la plante et, particulièrement, comment ces signaux perçus à la surface des racines peuvent réguler l'activité et le devenir de différents tissus de la plante. Pour cela, l'équipe a récemment acquis un grand jeu de données transcriptomiques au niveau « noyau unique » (4 traitements, 2 fonds génétiques de plante et 2 réplicats biologiques, plus de 100000 noyaux).

Le but du stage vise à analyser ces données d'expression de grande dimension dans le cadre d'un design complexe afin de répondre à la question biologique. Ces résultats pourront être comparés à ceux obtenus par l'équipe de S. Bensmihen lors de l'analyse de données de bulk RNA-seq.

Objectifs du stage :

- Il faudra tout d'abord se familiariser avec les données d'expression disponibles et reprendre les premières analyses faites par un binôme d'étudiantes en projet de Mathématiques Appliquées de l'INSA Toulouse à l'aide du package Seurat.
- Il faudra consolider le pipeline d'analyse pour chaque condition séparément.
- Il faudra ensuite approfondir l'analyse différentielle pour pouvoir comparer l'effet des traitements ou des différents fonds génétiques (dont l'un ne peut plus percevoir les signaux microbiens) sur la réponse génique. Ce point nécessitera de faire un point sur l'état de l'art sur cette question et de comparer les résultats obtenus par les différentes méthodes. Un des challenges est de tenir compte du design (traitements et fond génétiques) dans l'analyse.
- Durant ce stage, les analyses seront effectuées avec le logiciel R. Des scripts en Rmarkdown et des outils de visualisation/exploitation des résultats seront produits pour la reproductibilité des résultats et aider l'équipe de S. Bensmihen.

Encadrantes :

Sandra Bensmihen
DR CNRS, LIPME, Auzeville-Tolosane
sandra.bensmihen@inrae.fr

Cathy Maugis-Rabusseau
INSA Toulouse / IMT
cathy.maugis@insa-toulouse.fr

L'étudiant-e bénéficiera d'un environnement pluridisciplinaire entre biologie et biostatistique au sein des laboratoires du LIPME et de l'IMT.

Compétences attendues :

- De bonnes connaissances en statistiques
- De bonnes compétences en programmation R
- Une aptitude pour les applications à la biologie

Conditions du stage :

- Durée du stage : 4 à 6 mois
- Gratification de stage selon les règles en vigueur
- Localisation : L'étudiant-e partagera son temps de travail entre le LIPME (campus INRAE, Auzeville) et l'Institut de Mathématiques de Toulouse (campus INSA Toulouse, GMM)

Candidature :

Veuillez envoyer un email contenant votre CV et une lettre de motivation aux deux encadrantes citées ci-dessus, en indiquant dans l'objet « Candidature Stage scRNAseq LIPME-IMT ».