

Titre de la proposition du projet :

Développement d'un outil pour la comparaison et intégration de méthodes bioinformatiques/statistiques dédiées à l'analyse de l'expression différentielle et co-expression.

Contexte de la proposition du projet :

Les récents progrès des nouvelles technologies de séquençage ouvrent de nouvelles voies pour l'étude du vivant grâce à une augmentation sans précédent du nombre de données et de leur disparité. Cet amas d'information est essentiel pour nous permettre d'améliorer notre compréhension des relations entre génotype et phénotype. Pour répondre à ces besoins et proposer des signatures pour un phénotype donné, des approches bioinformatiques et biostatistiques ont été développées pour analyser les expressions différentielles et co-expression entre gènes. Le choix d'une méthode est difficile, chaque approche ayant ses forces et ses faiblesses, chacune d'entre elles visant à identifier des groupes de gènes exprimés de manière différentielle dans différentes conditions expérimentales, ainsi que des groupes de gènes co-régulés (Gadgil et al., 2008, Våremo et al. 2013). L'exploitation simultanée de plusieurs méthodes d'analyse est plus que pertinente pour améliorer la fiabilité des prédictions. De même, la comparaison des résultats d'expression différentielle (différence d'expression d'un gène entre deux conditions) et de co-expression (corrélation de l'expression de deux gènes toutes conditions confondues) est une information complémentaire dans l'analyse des données (Farahbod et al. 2019).

Description du travail :

Plus concrètement, ce projet vise à mettre en oeuvre un outil (scripts R) basé sur les méthodes actuelles pour le calcul de différence d'expression et de co-expression:

- charger, normaliser et filtrer les données d'expression,
- intégrer des méthodes de calcul pour l'expression différentielle et co-expression
- visualiser les résultats de comparaison avec des représentations basées sur les heatmaps, diagrammes de Venn, courbe ROC...

Nom(s) du (des) responsables du projet:

Elodie Darbo: elodie.darbo@u-bordeaux.fr

Patricia Thébault : patricia.thebault@u-bordeaux.fr

Gadgil M. A Population Proportion approach for ranking differentially expressed genes. *BMC Bioinformatics*. 2008;9:380. Published 2008 Sep 18. doi:10.1186/1471-2105-9-380

Marjan Farahbod, Paul Pavlidis, Differential coexpression in human tissues and the confounding effect of mean expression levels, *Bioinformatics*, Volume 35, Issue 1, 01 January 2019, Pages 55–61, <https://doi.org/10.1093/bioinformatics/bty538>

Leif Våremo, Jens Nielsen, Intawat Nookaew, Enriching the gene set analysis of genome-wide data by incorporating directionality of gene expression and combining statistical hypotheses and methods, *Nucleic Acids Research*, Volume 41, Issue 8, 1 April 2013, Pages 4378–4391, <https://doi.org/10.1093/nar/gkt111>