



Proposition de stage de master 2

Lieu : Hospices Civils de Lyon, Groupement Hospitalier Est, Centre de Biologie et Pathologie Est

Sujet : Mise en place d'outils/pipelines pour l'analyse de données de pharmacogénétique

La pharmacogénétique consiste à étudier des variations génétiques influençant la manière dont un individu répond à un médicament. Par exemple, le 5-fluorouracile est une chimiothérapie qui est très souvent prescrite pour le traitement des tumeurs solides. Or, certains patients possèdent des mutations dans le gène DPYD qui induisent un déficit en DPD, une enzyme qui métabolise le 5-fluorouracile. Chez ces patients, le 5-fluorouracile va s'accumuler et induire une toxicité pouvant être fatale.

Au sein des Hospices Civils de Lyon, l'identification de variations potentielles dans des gènes impliqués dans la réponse à des médicaments se fait en routine diagnostique par le séquençage d'un panel de gènes. Un pipeline d'analyse de ces données tourne actuellement, néanmoins il n'est pas optimal et de nouveaux outils sont à présent disponibles.

En parallèle, les HCL testent aussi la possibilité d'utiliser le séquençage de 3^e génération. Ces technologies de séquençage long reads sont particulièrement adaptées à l'analyse de données de pharmacogénétique car elles nécessitent une étape d'haplotypage qui est grandement facilitée lorsqu'on dispose de longs fragments.

Le but de ce stage est donc double.

- sur les données de séquençage de seconde génération (short read, technologie illumina), tester les différents outils d'analyse de données de pharmacogénétique disponibles, les comparer (entre eux ainsi qu'aux résultats obtenus en diagnostic) et élaborer un nouveau pipeline d'analyse de données de pharmacogénétique
- sur les données de séquençage de 3^e génération, nous nous intéressons plus particulièrement à la mise en évidence de l'existence d'hybrides entre 2 gènes (CYP2D6/CYP2D7). L'objectif du stage serait de développer une méthode permettant d'identifier et de visualiser ces hybrides.

Les données nécessaires pour le stage ont été séquencées en 2024 et sont donc déjà disponibles.

Pré-requis : bonne maîtrise du système d'exploitation linux (le stage implique de travailler sur des serveurs linux distant et sur un cluster de calcul distant), maîtrise d'un ou plusieurs langages de scripts (bash/perl/python), utilisation d'outils de développement collaboratif (git), connaissance éventuelle d'un langage de workflow (nextflow idéalement), connaissances de base en génétique et concernant les données de séquençage à haut débit, intérêt pour le travail en milieu hospitalier.

Ce travail sera réalisé au sein de la cellule bioinformatique de la plateforme de séquençage à haut débit des Hospices Civils de Lyon (4 ingénieurs bioinformaticiens expérimentés + 2 enseignants-chercheur), en collaboration avec des pharmacogénéticiens (Pr. L. Payen, Dr. D. Barthelemy).

Contact : Claire Bardel (claire.bardel@univ-lyon1.fr)