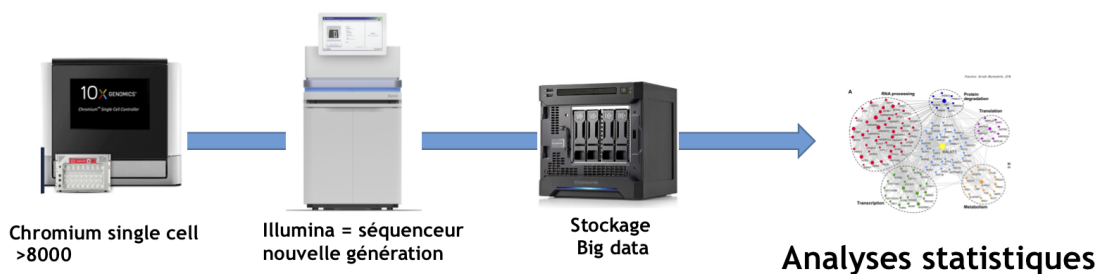
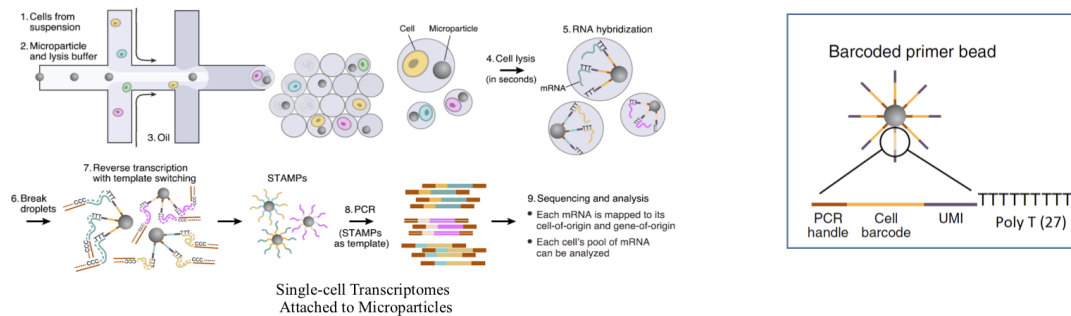


Étude des populations sous-clonales du myélome multiple Clustering de données de transcriptome sur cellule unique

Le myélome multiple

Le myélome multiple (MM) est caractérisé par une profonde diversité génétique et épigénétique des sous-populations de cellules malignes. La maladie évolue à travers un processus darwinien d'expansion clonale d'un mélange de sous-clones concurrents, sous l'influence du microenvironnement tumoral. Chez la plupart des patients, la pression du traitement entraîne une réorganisation et une diversification profonde des populations sous-clonales avec une dynamique complexe d'évolution. Les processus permettant aux sous-clones de survivre et de proliférer au fil du temps sont pour l'instant mal connus.



Objectifs scientifiques

- Le projet que nous proposons consiste à identifier les différents sous-clones qui composent la maladie et à analyser les composantes de son microenvironnement tumoral à partir de données du transcriptome sur cellule unique (scRNA-seq), pour des échantillons de moelle de patient.
- Pour identifier les sous-clones, des méthodes de classification non supervisée peuvent ici être mises en oeuvre. Dans ce contexte de grande dimension, il s'agit aussi de sélectionner les variables pertinentes pour appuyer le clustering sur les données les plus pertinentes. Dans un premier temps, les méthodes de clustering classiquement utilisées pour les données de cellule unique seront étudiées et évaluées.
- Dans un deuxième temps, il s'agira de comparer des profils de patients en fonction de leur population de sous clones. On cherche ici à comparer des clusterings d'un même patient à plusieurs instants, ou de comparer les profils de plusieurs patients

distincts. Il n'existe pas de réponse satisfaisante à ce problème dans la littérature pour le clustering de cellule unique ; il s'agit ici d'un problème de recherche assez ouvert.

- Le.e stagiaire travaillera sur une traduction mathématique satisfaisante du problème. Il.Elle contribuera aussi à l'élaboration d'un package R, ceci en collaboration étroite avec les biologistes et les statisticiens impliqués dans le projet.

Environnement du stage

Le stage sera encadré par Aurélien Serandour (MCU Centrale Nantes/ Chaire Inserm, Centre de Recherche en Cancérologie et Immunologie de Nantes et Angers) et Bertrand Michel (Ecole Centrale et Nantes, Laboratoire de Mathématiques Jean Leray). Le stagiaire sera accueilli dans les deux laboratoires à Nantes afin de pouvoir interagir plus facilement avec les deux encadrants. Début du stage : avril 2018. Ces travaux pourraient éventuellement être poursuivis par une thèse.

Contact : bertrand.michel@ec-nantes.fr

Profil attendu

Le profil attendu est celui d'un.e étudiant.e en Master 2 de statistiques et sciences des données, ou d'un.e étudiant.e ingénieur.e avec une spécialisation dans ce domaine. De solides connaissances en apprentissage statistique et en statistique mathématique sont nécessaires pour réussir ce stage, de même qu'une bonne maîtrise de R ou de Python.