

MÉTHODES POUR L'ANALYSE DE DONNÉES TRANSCRIPTOMIQUE

ECTS	Cours (h)	T.D. (h)	T.P. (h)	Stage (semaines)
6	30		30	

Mention du master transmettant la fiche UE : **BIOINFORMATIQUE**

Composante de gestion de l'UE : **Département de BIOLOGIE**

Responsable de l'UE : **LACROIX Vincent**

Statut du responsable : **MC**

PROGRAMME DE L'UNITE D'ENSEIGNEMENT :

Cette UE a pour but de former les étudiants à l'utilisation avancée et/ou au développement de méthodes d'analyse de données transcriptomiques.

Les protocoles d'acquisition des données (e.g. puces, RNA-seq) et les biais correspondants seront détaillés.

Les deux grandes familles de traitement bioinformatique des données NGS seront rappelées (mapping ou assemblage de novo). Les traitements spécifiques aux données transcriptomiques seront approfondis: identification et quantification d'exons, de jonctions, de gènes. Le problème de la reconstruction de tous les transcrits alternatifs d'un gène sera étudié.

L'analyse comparative des niveaux d'expression entre deux conditions expérimentales (traité / non traité, tissu1 / tissu2) sera l'occasion d'introduire les méthodes statistiques pour décider si un gène est différentiellement exprimé ou épissé. À cette occasion une brève introduction aux modèles linéaires généralisés sera faite. Les étudiants seront formés à bioconductor.

Les mécanismes de régulation de l'expression des gènes seront abordés par l'analyse de données ChIP-seq.

Le protocole ChIP-seq sera donc présenté, avec ses biais. Les méthodes bioinformatiques spécifiques de l'analyse de données ChIP-seq seront détaillées (peak calling, normalisation). Les analyses de séquences promotrices seront détaillées, notamment les aspects combinatoires et statistiques de la recherche et l'inférence de motifs dans les séquences.

Enfin, l'impact fonctionnel de différentiel d'expression ou d'épissage sera abordé via l'analyse intégrative de diverses sources de données. Pour cela, les étudiants seront familiarisés avec l'analyse GO, les méthodes de clustering de gènes co-exprimés (analyse multivariée, machine learning), la visualisation de gènes différentiellement exprimés dans un réseau de gènes connus. Les techniques récentes (Hi-C, ChIP-seq polII, ribosome profiling) seront abordées brièvement.

MUTUALISATION :

Master Santé Publique en tant qu'option.