

Environnement scientifique  
et technique de la formation



Institut de génétique moléculaire de  
Montpellier  
<http://www.igmm.cnrs.fr>

### RESPONSABLES

**Jean-Christophe ANDRAU**

Directeur de recherche  
UMR 5535

**Amal MAKRINI**

Ingénieure d'études  
UMR 5535

### LIEU

MONTPELLIER (34)

### ORGANISATION

4 jours

De 8 à 12 stagiaires

TD encadrés par 1 intervenant pour 4  
stagiaires maximum

### COÛT PÉDAGOGIQUE

1700 Euros

### À L'ISSUE DE LA FORMATION

Evaluation de la formation par les  
stagiaires

Envoi d'une attestation de formation

### DATE DU STAGE

**Réf. 19 023** : du lundi 20/05/19 à 09:00  
au jeudi 23/05/19 à 17:00

Janvier	Février	Mars	Avril
Mai 19 023	Juin	Juillet	Août
Sept.	Oct.	Nov.	Déc.

## ChIP-seq, RNA-seq et Hi-C : traitement, analyse et visualisation de données

### OBJECTIFS

- Savoir planifier une expérience simple de type ChIP-seq ou RNA-seq
- Savoir évaluer la qualité des données obtenues dans le cadre de la régulation des gènes (ChIP-seq) et de l'analyse du transcriptome (RNA-seq)
- Maîtriser les principales méthodes et outils d'analyse des données ChIP-seq et RNA-seq
- Savoir les visualiser dans un navigateur de génome et en extraire les coordonnées des pics enrichis
- Savoir manipuler et annoter les fichiers d'enrichissement (bedtools)
- Maîtriser les principales méthodes d'analyse des données Hi-C et de génération des cartes de contact chromosomique

### PUBLIC

Ingénieurs et chercheurs biologistes / bioinformaticiens

**Prérequis** : avoir des bases d'utilisation de lignes de commande sous R

### PROGRAMME

#### 1<sup>er</sup> jour : pre-processing des données ChIP-seq et RNA-seq

- Notions de bases du ChIP-seq, RNA-seq et Hi-C et de leurs pipelines d'analyse, principe de normalisation avec spike-in
- Prétraitement des données (qualité des données brutes, alignements, artefacts de séquençage)
- Exploration des données alignées (estimation de la taille des fragments, qualité de l'alignement)

#### 2<sup>ème</sup> et 3<sup>ème</sup> jours : analyse ChIP-seq

- Génération des fichiers d'enrichissement (.wig) avec le package R-PASHA
- Visualisation des fichiers .wig et initiation à la sélection de régions enrichies, aux paramètres de détection (MACS2 ou directement sur genome browser)
- Initiation à la mise en place de métaprofiles autour des régions d'intérêt (gène, TSS, TES, enhancers)
- Analyse d'enrichissement d'annotations fonctionnelles ("gene ontology")
- Recherche de motifs dans les régions enrichies

#### Analyse RNA-seq

- Quantification de l'expression des gènes dans les données RNA-seq à travers le RPKM
- Principe de normalisation utilisant des spike-in
- Analyse de l'expression différentielle des gènes et des exons

#### 4<sup>ème</sup> jour : analyse Hi-C

- Traitement et alignement des séquences et construction de la carte de contact chromosomique
- Visualisation et interprétation de la carte de contact
- Présentation de quelques bases de données et d'outils de visualisation de Hi-C
- Discussion (3 h) sur les problématiques des participants et sur les pistes de solutions à y apporter

#### Alternance de cours (4 h) et de TD (24 h)

### EQUIPEMENT

Mise à disposition d'un ordinateur par stagiaire sous linux avec logiciels nécessaires à la formation

### INTERVENANTS

J.-C. Andrau, L. Pioger (chercheurs) et A. Markini (ingénieure d'études)