

Cycle « Analyse de données de séquençage à haut-débit »

Module 1/5 : Analyses ADN

Le cycle de formation « Analyses de données de séquençage haut-débit » est constitué de cinq modules. :

- 1) **Analyses ADN**
- 2) Analyses RNA-seq
- 3) Analyses ChIP-seq
- 4) Analyses de variants
- 5) Métagénomique

Ces modules peuvent être suivis indépendamment. Toutefois maîtriser le contenu du module 1 (nettoyage et qualité des données, mapping, assemblage) est nécessaire à la bonne compréhension des modules suivants. De manière générale, le cycle de formation s'adresse à des participants qui ont des connaissances en biologie moléculaire et génomique.

OBJECTIFS	<ul style="list-style-type: none"> - Apprendre à manipuler des données de séquençage d'ADN - Présenter les méthodes et outils d'alignement et d'assemblage - Réaliser des contrôles de qualité et des alignements sur une référence - Utiliser la plateforme Galaxy pour ces analyses
PUBLIC CONCERNE	Biologistes et médecins : enseignants-chercheurs, chercheurs, ingénieurs, techniciens, doctorants contractuels, post-doctorants
PRE-REQUIS	<ul style="list-style-type: none"> - Etre familier avec la plate-forme web Galaxy (idéalement avoir suivi la formation bilille « Initiation à Galaxy »)
CONTENU	<ul style="list-style-type: none"> - Présentation d'un protocole classique pour des analyses de données de séquençage d'ADN - Introduction aux format des données NGS (<i>fastq pour les séquences, sam/bam pour les alignements</i>) et aux différents scores de qualité - Rappel sur l'utilisation de la plateforme Galaxy - Contrôle qualité des lectures et nettoyage (ex : FastQC, Cutadapt, Prinseq) - Présentation des méthodes et outils d'alignement et de l'influence des différents paramètres - Alignement sur un génome de référence (ex : BWA, Bowtie2) - Contrôle de qualité et filtrage de l'alignement (ex : samtools) - Présentation des méthodes et outils d'assemblage génomique et de l'influence des différents paramètres; aperçu des techniques pour l'assemblage transcriptomique <p>Les parties pratiques sur Galaxy aborderont entre autre</p> <ul style="list-style-type: none"> - L'importation et l'exportation de données - Les bonnes pratiques pour la gestion des dataset et des historiques - La création et l'utilisation d'un workflow Galaxy

Cycle « Analyse de données de séquençage à haut-débit »

Module 1/5 : Analyses ADN

METHODE	Alternance entre parties théoriques et parties pratiques, essentiellement à l'aide d'outils logiciels libres accessibles sous Galaxy
DATES	Session 1 : 6 et 7 février 2017 Eventuelle Session 2 : 6 et 7 mars 2017
INTERVENANTS	Sophie Gallina, IR CNRS Rayan Chikhi, CR CNRS Laurent Noé, MCF, Université de Lille, Sciences et Technologies En collaboration avec <i>bilille</i> , la plateforme de bioinformatique de Lille
EFFECTIF	Un groupe de 12 personnes
LIEU	Université de Lille
INSCRIPTIONS	Avant le 10 janvier 2017

	Anne-Sophie LECLERCQ Université des Sciences et Technologies de Lille, Cité Scientifique-Bâtiment A3, Service du Personnel-Formation Continue, 59655 VILLENEUVE D'ASCQ Cedex Tél.: 03 20 33 70 71. Fax : 03 20 33 59 72. Courriel : Anne-Sophie.Grare@univ-lille1.fr
	Dorothee TERRYN Inserm ADR Lille, 6 rue du Pr. Laguesse, CS 50027, 59045 LILLE Cedex Tél.: 03 20 29 86 78. Fax : 03 20 49 01 38. Courriel : dorothee.terryn@inserm.fr
	Pierre SILVEIRA CNRS - Délégation Nord - Pas de Calais et Picardie, 2 rue des Canoniers, 59046 LILLE Cedex Tél.: 03 20 12 36 88. Fax : 03 20 12 36 90. Courriel : pierre.silveira@dr18.cnrs.fr
	Nadia BENYAHIA Université Lille 2 - Université du Droit et de la Santé, 42 rue Paul Duez, 59800 LILLE Tél.: 03 20 96 43 37. Fax : 03 20 88 24 32. Courriel : nadia.benyahia@univ-lille2.fr