

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

Module 4/5 : Analyses de variants

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"> - Comprendre les grands principes de la détection de variants - Réaliser les différentes étapes du post-traitement des données d'alignement à la détection de variants - Adapter l'analyse en fonction du type de données NGS générées - Comprendre la structure des données de variants - Savoir annoter des variants - Etre capable d'interpréter une liste de variants grâce aux outils libres disponibles
PUBLIC CONCERNE	Biologistes : 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 ») - Avoir suivi le module 1/5 « Analyses ADN » de ce cycle ou toute autre formation permettant de justifier de connaissances sur les données de séquençage haut débit et leur alignement.
CONTENU	<ul style="list-style-type: none"> - Présentation des algorithmes de détection de variants - Format de données : alignement, variants, régions génomiques - Présentation et application des bonnes pratiques GATK - Réalisation des différentes étapes d'analyse post-alignement : « du bam au vcf annoté » (post-traitement des fichiers bam, détection et post-traitement des variants...) - Annotation des variants à l'aide de l'outil VeP (région exonique, HGVS...) - Découverte de différentes bases de données d'annotation publiques (1000 Genomes, HapMap, ExAc, dbSNP, ClinVar, COSMIC,...) - Interprétation des données obtenues
METHODE	Alternance entre parties théoriques et parties pratiques, essentiellement à l'aide d'outils logiciels libres accessibles sous Galaxy ou sur Internet

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

Module 4/5 : Analyses de variants

DATES	Du 2 au 4 octobre 2017 (dates prévisionnelles)
INTERVENANTS	<p>Emilie Ait Yahya et Christophe Demay, Bioinformaticiens, Cellule bio-informatique du Plateau Commun de Biologie Moléculaire, CHRU Lille</p> <p>Martin Figeac, Bioinformaticien, Plate-forme de Génomique Fonctionnelle et Structurale, Université de Lille 2</p> <p>Benjamin Grenier-Boley, Bioinformaticien, Institut Pasteur de Lille</p> <p>En collaboration avec <i>bilille</i>, la plateforme de bioinformatique de Lille</p>
EFFECTIF	Un groupe de 12 personnes
LIEU	Université de Lille
INSCRIPTIONS	Avant le 10 janvier 2017

	<p>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</p>
	<p>Dorothee TERRY 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</p>
	<p>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</p>
	<p>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</p>