

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

Module 5/6 : Analyses RNA-seq – partie 1 (bioinformatique)

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

- 1) Analyses ADN
- 2) Analyses de variants
- 3) Métagénomique
- 4) Analyses ChIP-seq
- 5) Analyses RNA-seq - partie 1 (bioinformatique)**
- 6) Analyses RNA-seq - partie 2 (biostatistique)

Ces modules peuvent être suivis indépendamment. Toutefois, maîtriser le contenu du module 1 (nettoyage et qualité des données, mapping) 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.

Cette formation est proposée par la plate-forme de bioinformatique et biostatistique de Lille, bilille : <https://wikis.univ-lille.fr/bilille>

OBJECTIFS	<ul style="list-style-type: none"> - Savoir réaliser une analyse transcriptomique par RNA-seq avec ou sans (de novo) génome de référence à l'aide du portail Galaxy - Avoir un regard critique sur la qualité des lectures obtenues par le séquenceur - Connaître et savoir paramétrer les outils nécessaires à l'analyse
PUBLIC CONCERNÉ	Biologistes et médecins : enseignants-chercheurs, chercheurs, ingénieurs, techniciens, doctorants contractuels, post-doctorants
PRÉ-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 1/6 « 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	<p><i>Connaissances générales en RNA-seq</i></p> <ul style="list-style-type: none"> - Spécificités du séquençage - État de l'art des outils d'analyse - Contrôle qualité des séquences - Introduction à la troisième génération de séquençage (PacBio/Nanopore) <p><i>Pour RNA-seq avec génome de référence spécifiquement</i></p> <ul style="list-style-type: none"> - Performances comparées des logiciels - Assemblage des transcrits (ex : HiSAT2+StringTie) - Alignement complet des lectures (ex : Hisat2, STAR) - Visualisation des lectures alignées (ex: IGV) - Quantification de l'expression des transcrits (ex : Salmon, featureCounts)

Cycle « Analyse de données de séquençage à haut-débit » Module 5/6 : Analyses RNA-seq – partie 1 (bioinformatique)

	<p><i>Pour RNA-seq de novo spécifiquement</i></p> <ul style="list-style-type: none"> - Assemblage de novo des lectures (ex : Trinity) - Visualisation des assemblages (ex: BLAT de l'UCSC) - Quantification / expression différentielle (ex: Trinity workflow) - Variant (SNP, alternative splicing) calling (ex: KisSplice)
MÉTHODE	Alternance entre parties théoriques et parties pratiques, essentiellement à l'aide d'outils logiciels libres accessibles sous Galaxy
DATES	16 et 17 septembre 2021 (2 jours)
INTERVENANTS	Camille Marchet, CR CNRS, BONSAI, UMR 9189, CRISTAL Pierre Pericard, IR Université de Lille, bilille, UMS PLBS
EFFECTIF	Un groupe de 10 personnes
LIEU	Université de Lille (Campus Cité Scientifique)
INSCRIPTIONS	Avant le 6 janvier 2021

	<p>Jennifer Chouchaoui Service formation des personnels Domaine universitaire du Pont de Bois - BP 60149 59653 Villeneuve d'Ascq Cédex Tél. 03.62.26.95.42. Courriel : jennifer.chouchaoui@univ-lille.fr</p>
	<p>Dorothee TERRYN Inserm DR 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 : dr18.liste.fp@dr18.cnrs.fr</p>