

Traitements de masse



Composante
École
Supérieure
d'Ingénieurs en
Technologies
Innovantes

Présentation

Description

Technologies du Vivant - 4e année - Semestre 2
UE4 - Bio-informatique et stratégie expérimentale

Cette matière contient deux parties :

- Traitement de masse de données en biologie (8 hC, 16 hTD)
 - Séquençage nouvelle génération (4h sous forme de conférences)
-

Pré-requis obligatoires

Connaissance de base en informatique et sur la structure d'un gène

Séquençage nouvelle génération

Connaissances de base en biologie moléculaire

Contrôle des connaissances

Contrôle continu

Syllabus

Traitement de masse

- Manipulation de fichiers simples (« plat » de type texte) présentant une structuration des informations
- extraire des informations ciblées, les importer et les analyser dans une feuille de calcul d'un simple tableur.)
- construire et utiliser une base de données
- Conception et la création de tables de stockage des données
- Construction de requêtes sur ces données stockées pour les analyser.

Séquençage nouvelle génération

- Historique : De la découverte de l'ADN au séquençage du premier génome humain.
- Sanger vs NGS
- Préparation des librairies
- Stratégies de séquençage : WGS, WES, capture, amplicon...
- Exemple du séquençage par synthèse d'Illumina
- Workflow bioinformatique
- Le service commun de génomique de Rouen : équipements, applications, prestations.
- Les nouvelles générations de séquenceurs (Pacific Biosciences, Oxford Nanopore) : la révolution du séquençage en temps réel de reads longs
- Les applications du NGS : DNA-Seq, RNA-Seq, ChiP-Seq, Single-Cell, Methyl-Seq...
- Pipelines bioinformatiques : détection de SNV, CNV. Stratégies d'interprétation. Infrastructure informatique.

L'ensemble est vu sous Linux

Liste des TP :

- Fichiers plats (Utilisation des commandes Linux) : redirection et chaînage de commandes, recherche des lignes contenant un motif, sélection de champs.
- Utilisation d'un tableur (LibreOffice) : Importation des données extraites, Définition de formules, Tracé de graphes, Limitations à l'utilisation d'un tableur.
- Conception et utilisation d'une base de données : organisation, interdépendance, clés primaires et clés étrangères, organisation physique des tables, Conception des tables et résolution des relations n-n, Requêtes d'interrogation, Introduction à l'interfaçage web.

Séquençage nouvelle génération

Formation théorique aux bases du NGS, à ses applications et à ses évolutions. Introduction au traitement bioinformatique des données issues du NGS

Informations complémentaires

Mutualisé avec la licence de biologie

Compétences visées

- Savoir extraire et importer automatiquement des informations à partir d'un fichier texte vers un tableur, utiliser une feuille de calcul pour analyser les données. Lorsque les données sont massives, savoir basculer vers la conception et l'utilisation d'une base de données.

- Connaître les outils permettant de manipuler des données biologiques
- Savoir créer et utiliser une base de données