

Intitulé de l'UE	Service Bioinformatique
Section(s)	- (4 ECTS) Master en Sciences de l'Ingénieur industriel orientation Life data technologies / Cycle 2 Bloc 1

Responsable(s)	Heures	Période
Vincent BRANDERS	50	Quad 1

Activités d'apprentissage	Heures	Enseignant(s)
Compléments de techniques bioinformatiques 2	30h	Axel THIEFFRY
Programmation R	20h	Vincent BRANDERS

Prérequis	Corequis

Répartition des heures
Compléments de techniques bioinformatiques 2 : 10h de théorie, 20h d'exercices/laboratoires
Programmation R : 20h d'exercices/laboratoires

Langue d'enseignement
Compléments de techniques bioinformatiques 2 : Français, Anglais
Programmation R : Français, Anglais

Connaissances et compétences préalables
- Principes de bases en biologie moléculaire et technologies de séquençage haut-débit (Illumina) - Utiliser la ligne de commande (bash) dans un système LINUX/UNIX

Objectifs par rapport au référentiel de compétences ARES
Cette UE contribue au développement des compétences suivantes
- Master en Sciences de l'ingénieur industriel :
<ul style="list-style-type: none"> • Identifier, conceptualiser et résoudre des problèmes complexes <ul style="list-style-type: none"> ◦ Intégrer les savoirs scientifiques et technologiques afin de faire face à la diversité et à la complexité des problèmes rencontrés ◦ Sélectionner et exploiter les logiciels et outils conceptuels les plus appropriés pour résoudre une tâche spécifique

- Concevoir et gérer des projets de recherche appliquée
 - Exploiter les résultats de recherche
- Communiquer face à un public de spécialistes ou de non-spécialistes, dans des contextes nationaux et internationaux
 - Maîtriser les méthodes et les moyens de communication en les adaptant aux contextes et aux publics
- S'engager dans une démarche de développement professionnel
 - Assumer la responsabilité de ses décisions et de ses choix

- Master en Sciences de l'ingénieur industriel en Life Data Technologies :

- Comprendre l'origine des données biologiques, les méthodes d'acquisition, de transmission, de stockage et de traitement
 - Savoir dialoguer avec les acteurs du domaine de la médecine/ recherche scientifique/ imagerie
 - Comprendre l'origine biologique des données à traiter
 - Connaître et utiliser les méthodes d'acquisition des données biologiques
 - Utiliser, adapter et/ou créer des outils bioinformatiques en réponse aux problèmes biologiques posés par les acteurs du domaine
 - Développer des outils informatiques et statistiques destinés à la gestion et à l'intégration des données
- S'adapter aux nouvelles technologies d'avenir dans un domaine en plein essor
 - Être capable d'apprentissage, d'adaptabilité et créativité pour répondre à des besoins spécifiques
 - S'adapter aux nouvelles technologies tant dans les domaines médicaux et scientifiques qu'informatiques

Objectifs de développement durable

Aucun

Acquis d'apprentissage spécifiques

- Produire des figures scientifiques pertinentes et attractives en terme de visualisation mais également en terme de communication effective de résultats
- Familiarisation au langage de programmation R, le Tidyverse, et Bioconductor, et leurs utilisation dans le cadre d'analyses de données de séquençage
- Développer une capacité critique vis-à-vis de résultats de l'analyse de données CAGE et de leur distinction par rapport aux méthodes classiques (mRNA-seq)

Contenu de l'AA Compléments de techniques bioinformatiques 2

Au travers d'une analyse complète d'un jeu de données de type CAGE (Cap Analysis of Gene Expression):

- 1- Nettoyage de données et contrôle qualité (LINUX/Unix, command line, bash/zsh)
- 2- Analyses bioinformatiques des données nettoyées: mapping, quantification, annotation
- 3- Quantification et expression différentielle (R, Tidyverse, Bioconductor)
- 4- Présentation des résultats sous la forme de figures scientifiques

Contenu de l'AA Programmation R

Au terme de l'AA *Programmation R*, l'étudiant sera capable :

- de comprendre la syntaxe de base de R et d'utiliser des concepts tels que les boucles, fonctions et apply(s)
- de connaître les types de données courants et leur manipulation en R
- d'importer, lire et écrire des données dans différents formats (CSV, TSV, Excel, ...)
- d'effectuer des opérations de base sur les données, telles que la sélection, le filtrage et le tri
- de créer des graphiques de base pour explorer et représenter des données et d'utiliser des packages tels que ggplot2 pour créer des visualisations plus avancées
- de réaliser de la manipulation avancée de données
- de respecter les conventions de codage R et maintenir un code propre et bien documenté
- de collaborer efficacement en utilisant des outils de gestion de versions comme Git

Méthodes d'enseignement

Compléments de techniques bioinformatiques 2 : approche interactive, approche par situation problème, approche déductive, étude de cas, utilisation de logiciels

Programmation R : travaux de groupes, approche par projets, approche par situation problème, utilisation de logiciels

Supports

Compléments de techniques bioinformatiques 2 : copies des présentations, notes d'exercices, Ressources et scripts disponible via un répertoire GitHub public

Programmation R : copies des présentations, notes d'exercices

Ressources bibliographiques de l'AA Compléments de techniques bioinformatiques 2

- Collection de "cheat sheets" pour la programmation R, la manipulation de données, et la production de figures scientifiques: <https://www.rstudio.com/resources/cheatsheets/>
- Guide pour la création et validation de figures scientifiques: <https://www.data-to-viz.com>
- Répertoire GitHub public créé spécifiquement pour le cours

Évaluations et pondérations

Évaluation	Note globale à l'UE
Langue(s) d'évaluation	Français, Anglais
Méthode d'évaluation	La note finale est définie comme suit : <ul style="list-style-type: none">• 60 % pour l'AA <i>Compléments de techniques bioinformatiques 2</i><ul style="list-style-type: none">◦ Une évaluation de la participation tout au long du cours (présence, attention, motivation, questions, réponses, etc.)◦ Complétion d'un examen dont le format (individuel/groupe ; interne ou <i>via service tiers</i>) sera défini en fonction de l'avancement par rapport au programme prévu• 40 % pour l'AA <i>Programmation R</i> dont la note est répartie comme suit :<ul style="list-style-type: none">◦ 50 % pour le travail continu (certifications DataCamp)◦ 25 % pour la réalisation d'un travail avec Git et RShiny◦ 25% pour la réalisation d'un examen écrit (sur ordinateur)

Report de note d'une année à l'autre pour l'AA réussie en cas d'échec à l'UE

Compléments de techniques bioinformatiques 2 : **oui**
Programmation R : **oui**

Année académique : **2023 - 2024**