



LMD5 — Fiche UE du département Biologie-Écologie

Niveau : M2 S9

Intitulé : Bioinformatique pour l'Évolution et l'Écologie (BioInfo)

Code : HAB907B

Responsable(s) : Vincent RANWEZ

Co-responsable(s) : Emmanuel DOUZERY

Contact(s) : vincent.ranwez [at] supagro.fr, emmanuel.douzery [at] umontpellier.fr

Moodle : <https://moodle.umontpellier.fr/course/view.php?id=19107>

Nombre ECTS : 2

Nombre d'heures : CM = 0 — TD = 9 — TP = 0 — Terrain = 0 — Séminaires = 6 — SPS = 0

Statut par parcours : B2I (—), BioGET1 (—), BioGET2 (—), CEPAGE (—), DARWIN (Choix), EcoSystèmes (—), EGEN (-), MédiACCES (—), MEME (—), PAL (), Agreg (—).

Description de l'UE : L'UE vise trois objectifs principaux.

- Premièrement donner aux étudiants un socle de connaissance et de compétences informatiques et ainsi leur fournir des bases solides pour apprendre et utiliser les outils bioinformatiques utilisées plus spécifiquement en évolution et écologie.
- Deuxièmement les sensibiliser à la nécessité de produire des résultats reproductibles et leur présenter les concepts et outils clés pour cela.
- Troisièmement faire travailler les étudiants sur des exemples concrets remobilisables lors de leur stage de master et de leur future vie professionnelle.

Compétences visées (Savoir, Savoir-Faire, Savoir-Être) : "Savoirs :

- Commandes clés de linux/bash en bioinfo (ls, cp, mv, grep, cut, sort, uniq, sed, ...) ;
- Notions clés de programmation impératives (variables, boucles, fonctions, structures de données basiques) ;
- Notions de flux de données (pipes en bash et R ggplot2) ;
- Notions clés liés aux clusters de calculs (organisation d'un cluster, utilisation d'un cluster) ;
- Bonnes pratiques en bioinformatique (commentaires, indentation, nommage des variables/fonctions, données FAIR, utilisation basique de git et de md, "test first approach") ;
- Notions clés pour tester et valider des programmes et des modèles ("training/validation/ test sets", mesurer l'apport d'un modèle/programme par rapport à l'existant) ;
- Limites inhérentes aux outils bioinfo (fonction objectif, heuristique).

Savoir-Faire :

- Cluster de calculs : manipuler et chercher du contenu dans de nombreux fichiers/dossiers (bash/R) ;
- Cluster de calculs : se connecter à un cluster, transférer des données de/vers le cluster, lancer des "jobs" ;
- Programmation : lancer des tâches répétitives (boucle for en bash), définir une fonction et l'appeler sur plusieurs jeux de données (R, bash) ;
- Programmation : enchaîner des commandes (bash/R-ggplot2) ;
- Reproductibilité : savoir créer un dépôt git et l'utiliser pour gérer plusieurs version d'un script R/bash ;
- Reproductibilité : utiliser markdown pour documenter les scripts."

Prérequis (compétences et/ou UE) : -

Modalités de contrôle des connaissances : Contrôle continu intégral : 100%

Informations additionnelles : Cette UE peut avoir lieu à n'importe quelle période d'enseignement du semestre, et les compétences bioinformatiques acquises peuvent être facilement (re)mobilisées dans d'autres UE (par exemple, l'UE de Phylogénie Approfondie).

Les principes de bases des méthodes NGS (kmers, assemblage, mapping) seront traitées dans le module de Phylogénie Approfondie (HAB946B).

Les notions de bases de données (SIG, créations de base de données, ...) ne seront pas abordées dans cette UE.

Cette UE est soutenue par DataCamp [<https://www.datacamp.com/>], la plateforme d'apprentissage la plus intuitive pour apprendre la science des données et de leur analyse. La méthodologie d'apprentissage par la pratique de DataCamp combine de courtes vidéos d'experts et des exercices pratiques au clavier pour aider les étudiant(e)s à acquérir leurs compétences.