Syllabus de la spécialisation

Biologie Computationnelle pour les Biotechnologies

Elie Maza (INP-ENSAT, elie.maza@toulouse-inp.fr)

César Aceves (INSA, cesar.aceves@insa-toulouse.fr)

2026-2027

BioComp

Table des matières

Tableau récapitulatif du volume horaire et des ECTS par UF 4
Prologue 5
UF1 Outils numériques et concepts fondamentaux 6
Module 1 Extraction et mise en forme des données 6
Module 2 Stat & R 7
UF2 Bio-informatique pour la génomique 8
Module 1 Assemblage *de novo* et annotation des génomes 8
Module 2 Alignement de séquences et recherche de polymorphisme 8
Module 3 Bio-informatique pour l'épigénome 9
UF3 Post-Génomique 10
Module 1 Transcriptomique (Microarray) 10
Module 2 Protéomique 10
Module 3 Métabolome et Fluxome 11
UF4 Biologie intégrative 12
Module 1 Intégration statistique des données (avec *mixOmics*) 12
Module 2 Modélisation dynamique d'un système biologique avec régulation 12
Module 3 Introduction à l'IA pour la génomique 13
UF5 Biologie structurale computationnelle 14
Module 1 Bases théoriques et méthodologiques 14
Module 2 Mise en pratique des méthodes 14
UF6 Relations Humaines, Professionnelles et Ethiques 16
Module 1 Management d'équipe et Psychologie sociale (étudiants INSA) 16
Module 2 Sport (étudiants INSA et ENSAT) 16
Module 3 Parcours Professionnel Individualisé (étudiants INSA et ENSAT) 17
UF7 Projet « Défi en bio-informatique » & Anglais 18
Module 1 Organisation d'une table ronde au forum carrière (étudiants ENSAT) 18
Module 2 Définition du « Défi en bio-informatique » (étudiants ENSAT) 19
Module 3 Réalisation du « Défi en bio-informatique » (étudiants ENSAT et INSA) 19
UF8 Projet « RNA-Seq » 20
UF9 Stage de fin d'études 21
UF10 Entreprise 22

Cas de l'alternant en contrat de professionnalisation 22
Cas de l'alternant en contrat d'apprentissage 23
Codes et domaines INSA et codes ENSAT des UF de BioComp 24

Tableau récapitulatif du volume horaire et des ECTS par UF

Numéro et nom de l'UF	Référents	ECTS	Vol. horaire en présentiel	Vol. horaire en autonomie
UF1* – Outils numériques et concepts fondamentaux	E. Maza	5	40h	20h
UF2* – Bio-informatique pour la génomique	M. Zouine	4	41h20	21h20
UF3* – Post-Génomique	B. Enjalbert	3	22h30	5h
UF4 – Biologie intégrative	C. Aceves	4	36h40	18h20
UF5* – Biologie structurale computationnelle	I. André	3	35h	13h45
UF6 – Relations Humaines, Professionnelles et Ethiques	H. Herring F. Rey (Sport) Q. Marechal (PPI)	INSA : 6 ENSAT : 2	INSA : 65h ENSAT : 30h	INSA : 0h ENSAT : 0h
UF7 – Projet Défi bioinfo & Anglais**	C. Laplanche A. Alibert (Anglais)	INSA : 2 ENSAT : 6	INSA : 5h20 ENSAT : 9h20	INSA : 33h20 ENSAT : 66h40
UF8* – Projet RNA-Seq	F. Regad	3	30h40	9h20
UF9 – Stage de fin d'études	C. Aceves E. Maza	30		
UF10* Entreprise**	V. Barraud-Didier	10		
TOTAL INSA		60	276h40	121h20
TOTAL ENSAT		60	245h40	154h20

* UF suivie par les alternants ENSAT

** UF suivie en partie par les alternants ENSAT

*** UF suivie exclusivement par les alternants ENSAT

Prologue

A la croisée de la biologie, de l'informatique, des mathématiques, de la statistique, des sciences de l'information, de la physique et de la chimie, la biologie computationnelle se positionne aujourd'hui comme une discipline essentielle, aussi bien du point de vue de la gestion et de l'interprétation des données biologiques affluant massivement dans les banques de données, que du développement d'outils d'analyse et de prédiction *in silico*.

Cette spécialisation de 3ème année, par nature interdisciplinaire, est commune aux deux écoles d'ingénieurs toulousaines que sont l'AgroToulouse et l'INSA. Cette spécialisation se veut une formation intégrée, à la programmation et à l'analyse statistique, au traitement et à l'analyse de données génomiques et post-génomiques, ainsi qu'à leur intégration, et à la biologie structurale computationnelle.

L'aisance à analyser des documents et à communiquer en langue anglaise sera évaluée dans une unité de formation projet avec une mise en situation lors d'un « project review meeting ». Un enseignant d'anglais de l'AgroToulouse accompagnera les étudiants dans la préparation de l'exposé, aussi bien à l'écrit qu'à l'oral.

UF1 Outils numériques et concepts fondamentaux

Référent : Elie Maza (INP-ENSAT)

Intervenants : Anis Djari (INP-ENSAT), Christophe Klopp (INRAE), Elie Maza (INP-ENSAT)

ECTS : 5

Volume horaire global : 40h en présentiel, 20h en autonomie = 60h étudiant

Finalités. La finalité principale de cette UF est de fournir à tous les étudiants l'ensemble des prérequis nécessaires pour pouvoir poursuivre correctement la formation proposée. Il s'agit, plus précisément, de rappeler (ou d'introduire) des notions d'informatique, de statistique et/ou de génomique de base, au travers d'exemples multiples issus, entre autres, des technologies de séquençage à haut-débit.

Objectifs. L'étudiant devra être capable de travailler sous l'environnement *Unix*, de manipuler des fichiers à l'aide des langages *AWK* et *Python*, d'automatiser des traitements de données et de lancer des commandes sur un cluster de calcul. Il devra aussi être capable d'effectuer des analyses statistiques classiques à l'aide du logiciel R. L'étudiant devra connaître les techniques de séquençage de 2nde et 3ème génération, et savoir réaliser des annotations de séquences à l'aide d'outils d'alignement local lancés en ligne de commande.

Place dans la formation. Le suivi de cette UF est un prérequis à la poursuite de la formation.

Module 1 Extraction et mise en forme des données

Intervenant : Christophe Klopp, Anis Djari

Volume horaire : 26h40 en présentiel, 13h20 en autonomie = 40h étudiant

Coefficient : 1

Description. Ce module a pour objectif de familiariser les étudiants avec l'environnement *Unix*, l'extraction d'information et la manipulation de fichiers de grande taille, l'automatisation de ces traitements. Les fichiers utilisés, issus de bases de données publiques, proviendront de différents appareils à haut-débit dit "omiques" couramment employés sur les projets de biologie. Ce module comprend les 4 parties/techniques décrites ci-dessous :

1. L'environnement ligne de commande sous Unix (4h+2h). Comment naviguer dans cet environnement ? Comment construire une ligne de commande et en enchaîner plusieurs ? Comment filtrer des fichiers de grande taille ?
2. La modification des lignes d'un fichier à l'aide des commandes sed et awk (3h+1h). Comment corriger des erreurs ? Comment ajouter des valeurs manquantes ? Comment modifier les plages de valeurs d'une variable ?
3. La manipulation de fichiers à l'aide des langages de programmation Perl et Python (6h+4h). Comment construire de nouvelles variables à partir des données figurant dans les fichiers ? Comment traiter des blocs de lignes d'un fichier ? Quelles structures de données utiliser (liste, table de hashage, objet, ...) pour réaliser des traitements complexes ?
4. L'automatisation des traitements sur le poste de travail et sur un cluster de calcul (3h+1h). Comment lancer simultanément plusieurs commandes ? Comment suivre l'avancement des traitements ? comment arrêter un traitement en cours ? Comment allouer des ressources à un traitement ?

Evaluation. Epreuve de 2 heures sur machine avec rapport à rendre.

Modalités de rattrapage. Les mêmes que précédemment.

Module 2 Stat & R

Intervenant : Elie Maza

Volume horaire : 13h20 en présentiel, 6h40 en autonomie = 20h étudiant

Coefficient : 1

Description. Ce module présente les principaux concepts statistiques (de la statistique descriptive, de la en théorie de l'échantillonnage, de l'estimation statistique, des tests statistiques) au travers de l'utilisation du logiciel et langage de programmation R (http://www.r-project.org/). Les jeux de données traités seront issus, entre autres, des technologies dites à haut-débit. Plus précisément, les étudiants travailleront les outils statistiques suivants :

1. Introduction à R. Manipulation de tableaux de données. Traitement de données. Description/manipulation des outils classiques de l'analyse descriptive uni et bivariée. (2h40 + 1h20 de TA)
2. Description/manipulation avec R des outils classiques de la théorie de l'échantillonnage, de l'estimation statistique, des tests statistiques usuels. (2h40 + 1h20 de TA)
3. Description/manipulation avec R de méthodes statistiques descriptives multivariées classiques. Analyse en Composantes Principales (ACP). Analyse Factorielle des Correspondances (AFC). Analyse des Correspondances Multiples (ACM). Classification non supervisée (algorithme *k-means*, classification hiérarchique). Analyse Factorielle Discriminante (AFD). Multidimensional scaling (MDS). (2h40 + 1h20 de TA)
4. Description/manipulation avec R de méthodes statistiques prédictives multivariées. Modèle linéaire gaussien. Régression linéaire simple, multiple. Analyse de la Variance (ANOVA) à 2 facteurs ou plus. ANCOVA. (2h40 + 1h20 de TA)
5. Possibles compléments. Introduction aux méthodes de sélection de variables. Introduction au modèle linéaire généralisé (modèle log-linéaire, régression logistique, …). Test non paramétrique de Kruskal-Wallis pour données non gaussiennes. (2h40 + 1h20 de TA)

Evaluation. Epreuve de 2 heures sur machine avec rapport à rendre.

Modalités de rattrapage. Les mêmes que précédemment.

UF2 Bio-informatique pour la génomique

Référent : Mohamed Zouine (INP-ENSAT)

Intervenants : Philippe Bardou (INRAE), Cédric Cabau (INRAE), Christophe Klopp (INRAE), Benjamin Linard (INRAE), Mohamed Zouine (INP-ENSAT)

ECTS : 4

Volume horaire global : 41h20 en présentiel, 21h20 en autonomie = 62h40 étudiant

Finalités. Cette UF permettra d'acquérir les connaissances et les compétences nécessaires pour manipuler les données issues des approches expérimentales faisant appel au séquençage haut débit dit de 2ème et 3ème génération, pour :

1. Générer un génome ou un transcriptome de référence.
2. Annoter ces séquences en cherchant les régions géniques et en prédisant leur fonction.
3. Aligner des séquences haut débit sur ces génomes de référence pour chercher des variants alléliques (SNP calling).
4. Identifier l'épigénome par séquençage bisulfite et par ChiP-Seq.

Objectifs. L'étudiant devra être capable d'utiliser les outils bio-informatiques installés sur un cluster de calcul pour générer, annoter et exploiter un génome ou un transcriptome de référence.

Place dans la formation. Cette UF, située en amont des autres UF, permettra d'apprendre comment générer les ressources nécessaires pour la post-génomique (UF3) et l'intégration des données "omiques" (UF4).

Module 1 Assemblage *de novo* et annotation des génomes

Intervenants : Christophe Klopp (assemblage) et Benjamin Linard (annotation)

Volume horaire : 13h20 en présentiel, 10h40 en autonomie

Coefficient : 1

Description. Ce module a pour objectif de présenter certaines techniques d'assemblage des génomes et des transcriptomes et leur annotation structurale et fonctionnelle pour générer une séquence de référence pour les espèces d'intérêt. Cette référence est souvent utilisée comme une ressource de départ pour aborder d'autres approches "omiques" comme le RNA-Seq, ChiP-seq et autres. Certaines de ces approches seront abordées au cours de cette spécialisation. Il abordera ensuite les aspects d'annotation fonctionnelle des contigs et scaffolds produits par l'assemblage. Nous introduisons les techniques et algorithmes utilisés pour l'annotation des régions codantes, les éléments répétés, et les ARN non codants, ainsi que les principales différences distinguant l'annotation fonctionnelle des procaryotes et eucaryotes.

Évaluation assemblage : Présentation orale par groupe de résultats d'assemblage.

Évaluation annotation : Présentation orale par groupe de résultats d'annotation.

Modalités de rattrapage. Oral individuel commun avec les modules 1, 2 et 3.

Module 2 Alignement de séquences et recherche de polymorphisme

Intervenants : Philippe Bardou, Cédric Cabau

Volume horaire : 10h40 en présentiel, 10h40 en autonomie

Coefficient : 1

Description. Cette formation a pour objectif d'apprendre à traiter les séquences issues des NGS pour la recherche de polymorphisme. Seront abordés les formats de séquences, d'alignements et de variants ainsi que les logiciels dédiés à l'alignement de séquences sur un génome de référence, à la recherche et à l'annotation de polymorphismes.

Evaluation. Présentation orale par groupe de résultats de recherche de variants + QCM.

Modalités de rattrapage. Oral commun aux modules 1, 2 et 3.

Module 3 Bio-informatique pour l'épigénome

Intervenant : Mohamed Zouine

Volume horaire : 20h00 en présentiel

Coefficient : 1

Description. Les nouvelles technologies de génomique offrent l'opportunité de pouvoir analyser l'épigénome à différents niveaux : la méthylation du génome et les modifications post-traduction des histones. Ce module vise (i) à présenter les approches bio-informatiques permettant de caractériser ces modifications (TDs sur machine) et (ii) à discuter d'applications concrètes de ces méthodes en recherche (articles scientifiques).

Evaluation. Examen sur table de 1h.

Modalités de rattrapage. Oral commun aux modules 1, 2 et 3.

UF3 Post-Génomique

Référent : Brice Enjalbert (INSA)

Intervenants : David Bouysset (CNRS), Brice Enjalbert (INSA), Pierre Millard (INSA)

ECTS : 3

Volume horaire global : 22h30 en présentiel, 5h00 en autonomie = 27h30 étudiant

Finalités. Connaître et savoir utiliser les principales approches "omiques" (gène, ARN, protéines, métabolites et flux). Apprendre à manipuler les jeux de données et à en extraire l'information essentielle.

Objectifs. Être capable de choisir la ou les meilleures approches en fonction de la question biologique posée. Savoir poser un plan expérimental satisfaisant les objectifs et respectant les contraintes. Connaître les principaux outils et stratégies de manipulation et de traitement de données "omiques".

Place dans la formation. Cette UF explore les avancées permises par les efforts de la génomique (UF2) et détaille les techniques dont sont issues les jeux de données analysés dans l'UF4 (Intégration).

Module 1 Transcriptomique (Microarray)

Intervenants : Brice Enjalbert

Volume horaire : 7h30 en présentiel, 2h30 en autonomie

Coefficient : 1

Description. Ce module a comme objectif de présenter l'approche de transcriptomique par microarray. La partie RNA-Seq est vue dans le projet RNA-Seq (UF8).

- Cours (2h30, BE). Les ARNs, la transcription, technique de Northern, biopuces, stabilome, overview RNA-Seq.

- TD (5h00, BE ; 2h30 en autonomie). Traitement des données : puces à ADN, E. coli, raw jusqu'au GO et visualisation.

Évaluation. Épreuve sur machine de 2h30 commune aux modules 1, 2 et 3.

Modalités de rattrapage. Épreuve sur machine ou oral.

Module 2 Protéomique

Intervenant(s) : David Bouysset, Brice Enjalbert

Volume horaire : 7h30 en présentiel, 1h15 en autonomie

Coefficient : 1

Description. Ce module a comme objectif de présenter les approches visant à identifier et à quantifier les protéines à l'échelle génomique, du gel2D à la protéomique comparative.

- Cours (1h15, BE)

  a. Les protéines, la traduction, traductome, Gels 2D, MS dégradome, notion quantitative (1h15, BE).

  b. Protéomique comparative, séquençage, modif post-trad (1h15 ; DB).

- TD (5h00, DB, BE ; 1h15 en autonomie)

  a. Traitement de données ITRAQ (1h15, BE ; 1h15 en autonomie).

  b. Séquençage de protéines et modification post-traductionnelle (3h45, DB).

Évaluation. Épreuve sur machine de 2h30 commune aux modules 1, 2 et 3.
Modalités de rattrapage. Épreuve sur machine.

Module 3 Métabolome et Fluxome

Intervenant(s) : Brice Enjalbert, Pierre Millard
Volume horaire : 7h30 en présentiel, 1h15 en autonomie
Coefficient : 1
Description. Ce module a comme objectif de présenter les approches visant à identifier et à quantifier les pools de métabolites et les distributions de flux dans les voies métaboliques.

- Cours (2H30, BE, PM).
  a. Les métabolites, organisation des voies, RMN, MS, Notions de flux, analyse des flux (1h15, BE).
  b. Cartographie des flux (1h15, PM)
- TD (5h00, BE, PM ; 1h15 en autonomie).
  a. Analyse de métabolome (1h15, BE ; 1h15 en autonomie).
  b. Détermination d'un fluxome (3h45, PM).

Évaluation. Épreuve sur machine de 2h30 commune aux modules 1, 2 et 3.
Modalités de rattrapage. Épreuve sur machine.

UF4 Biologie intégrative

Référents : C. Aceves (TBI, INSA)

Intervenants : C. Aceves (TBI, INSA), S. Déjean (IMT, UT3), R. Mourad (MCD, CBI, UT3)

ECTS : 4

Volume horaire global : 36h40 en présentiel, 18h20 en autonomie = 55h étudiant

Finalités. L'étudiant devra être capable d'analyser et de modéliser des données de grandes dimensions et/ou provenant de différents niveaux d'organisation du vivant. Pour cela, l'étudiant devra savoir identifier et utiliser à bon escient les domaines et les méthodes adéquats présentés dans cette unité de formation.

Objectifs. L'étudiant devra connaître et savoir utiliser des méthodes d'intégration de données statistiques (Module 1) ou de données issues de systèmes dynamiques (Module 2).

Place dans la formation. Cette UF est consacrée à la valorisation, entre autres, des données collectées lors de l'UF3 Post-génomique. Cette UF est donc "naturellement" précédée de cette UF3. Cela dit, cette UF peut être traitée de manière indépendante puisque les méthodes étudiées peuvent aussi être appliquées à d'autres sources de données.

Module 1 Intégration statistique des données (avec *mixOmics*)

Intervenant : S. Déjean

Volume horaire : 10h40 en présentiel, 9h20 en autonomie = 20h étudiant

Coefficient : 1

Description. Ce module a comme objectif de présenter un certain nombre de techniques statistiques d'intégration de données pour analyser des jeux de données de grandes dimensions. Les données illustrant ces méthodes sont ici issues de techniques dites à haut débit (données "omiques") telles que les données de transcriptomique, de protéomique, etc. Cela dit, les méthodes statistiques décrites peuvent être employées dans des domaines bien plus larges que ceux étudiés ici, pour intégrer des données de grandes dimensions. Ce module comprend les 3 parties/techniques décrites ci-dessous :

1. Méthodes pour un seul jeu de données (Analyse en Composantes Principales et Analyse Discriminante) : (i) Principe : ACP, AFD, PLS-DA, extensions sparse, (ii) Exemples, (iii) Cas d'étude avec mixOmics.
2. Méthodes multi-blocs pour l'intégration horizontale (mêmes individus, plusieurs ensembles de variables) : (i) Principe : PLS, PLS multi-blocs, méthodes supervisées et non supervisées, extensions sparse, (ii) Exemples, (iii) Cas d'étude avec mixOmics.
3. Méthodes multi-groupes pour l'intégration verticale (plusieurs groupes d'individus, même ensemble de variables) : (i) Principe : méthodes multi-groupes supervisées et non supervisées, (ii) Exemples, (iii) Cas d'étude avec mixOmics.

Evaluation. Projet avec rapport à rendre.

Modalités de rattrapage. Oral

Module 2 Modélisation dynamique d'un système biologique avec régulation

Intervenants : C. Aceves

Volume horaire : 10h en présentiel, 5h en autonomie = 15h étudiant

Coefficient : 1

Description. Ce module a comme objectif de montrer les principales méthodologies utilisées pour l'analyse des réseaux métaboliques. Les méthodes analysées sont divisées entre celles qui prennent l'hypothèse d'un état pseudo-stationnaire et celles qui sont dynamiques. Le module est composé de cinq parties :

1. Introduction aux systèmes biologiques (1h15 présentiel)
2. L'analyse d'ordres de magnitude pour les vitesses des réactions (1h15 présentiel)
3. Analyse d'états pseudo-stationnaires (2h30 présentiel). Les propriétés de la matrice de l'état pseudo-stationnaire à droite et à gauche.
4. Analyse métabolique pour proposer des knockouts avec CellNet sur *MatLab* (1h15 présentiel, 1h15 autonomie)
5. Les bilans matières avec ODEs, Modèles hybrides et Capteurs logiciels avec Machine Learning pour les bioprocédés (2h30 présentiel, 8h45 autonomie). Principes basiques. Analyse de phase. Hybridation.

Evaluation. Projet avec rapport à rendre.

Modalités de rattrapage. Oral

Module 3 Introduction à l'IA pour la génomique

Intervenants : R. Mourad

Volume horaire : 16h en présentiel, 4h en autonomie = 20h étudiant

Coefficient : 1

Description. Ce module a pour objectif d'introduire l'apprentissage des réseaux de neurones profonds ainsi que la mise en œuvre pratique, avec R (librairie Keras), d'approches d'apprentissage classique et profond pour résoudre des problèmes rencontrés en biologie moléculaire et génétique humaine. Le module sera organisé de la manière suivante :

1. Cours d'introduction aux réseaux de neurones, et application pour la génétique humaine (4h).
2. TP 1 : Prédire les éléments régulateurs avec des approches d'apprentissage automatique (SVM, K-mers, GLM, forêts aléatoires) (4h) et avec des approches d'apprentissage profond (CNN, interprétation de modèles) (4h).
3. TP 2 : Application à la génétique humaine : SNP issus des études d'association à l'échelle du génome et locus de traits quantitatifs d'expression (eQTL ou eSNP). Prédiction de l'effet du SNP par apprentissage profond pour la priorisation des variants et la compréhension des mécanismes biologiques (4h).
4. Travail en autonomie sur les TP (4h).

Evaluation. Projet avec rapport à rendre.

Modalités de rattrapage. Oral

UF5 Biologie structurale computationnelle

Référents : Isabelle André

Intervenants : Isabelle André (CNRS-INSA), Jérémy Esque (INRAE-INSA), Nicolas Panel (CNRS-INSA)

ECTS : 3

Volume horaire global : 35h en présentiel, 13h45 en autonomie = 48h45 étudiant

Finalités. Cette unité de formation permettra d'acquérir les connaissances et compétences nécessaires en modélisation moléculaire pour l'étude des relations séquence-structure-dynamique-fonction des protéines.

Objectifs. Être capable de visualiser, manipuler, analyser des structures tridimensionnelles de (macro)molécules. Connaitre les principales méthodes de calcul utilisées en modélisation 3D ainsi que leurs limites pour prédire la structure de protéines et assemblage moléculaires ainsi que pour étudier la réactivité, la conformation, et la dynamique des molécules. Être capable de choisir les approches de modélisation moléculaire les plus adaptées en fonction de la question biologique posée.

Place dans la formation. Ce module nécessite exclusivement l'UF1 comme prérequis.

Module 1 Bases théoriques et méthodologiques

Intervenants : Isabelle André, Jérémy Esque, Nicolas Panel

Volume horaire : 7h30 en présentiel

Coefficient : 1

Description. Ce module a pour objectif de fournir une introduction générale à la modélisation moléculaire, aux méthodes de représentation et analyse des structures, et bases de données structurales, et de présenter les approches de modélisation moléculaire multi-échelles :

a. Introduction générale à la biologie structurale (2h30, JE).
b. Méthodes de calculs théoriques (2h30, IA).
c. Prédiction de structures tridimensionnelles des protéines et dynamique moléculaire (2h30, JE).

Module 2 Mise en pratique des méthodes

Intervenants : Isabelle André, Jérémy Esque, Nicolas Panel

Volume horaire : 27h30 en présentiel, 13h45 en autonomie = 41h15 étudiant

Coefficient : 1

Description. Ce module a comme objectif de mettre en application quelques méthodes :

a. Analyse & Modélisation de structures 3D de protéines
b. Analyse & Construction de modèles 3D de complexes protéine-ligand
c. Champs de force de mécanique moléculaire et minimisation d'énergie
d. Dynamique moléculaire

La rédaction de comptes rendus est prévue dans chacun des TPs.

Evaluation. Compte-rendu de chaque TP.

Modalités de rattrapage. Présentation orale.

UF6 Relations Humaines, Professionnelles et Ethiques

Référent : H. Herring (INSA)

Intervenants : H. Herring (INSA), F. Rey (INSA, Sport), Q. Marechal (INSA, PPI) et Enseignants CSH INSA.

ECTS : 6 (INSA) et 2 (ENSAT)

Volume horaire global : 65h en présentiel (INSA) et 30h en présentiel (ENSAT)

Finalités. A la fin de ce module, l'étudiant devra avoir compris et pourra expliquer les objectifs et l'organisation de la fonction RH, l'analyse des emplois, le processus de recrutement, la gestion prévisionnelle des emplois et des compétences, la motivation au travail, les rémunérations, le processus d'appréciation des salariés, la formation, la gestion des carrières, la gestion des conflits, les contrats de travail. Il devra aussi avoir compris ce qu'est un groupe, ce qui l'influence et le dynamise. Enfin, les différences de fonctionnement des principes de l'expression écrite et orale en anglais devront être acquises. L'aisance à analyser des documents et à communiquer en langue anglaise sera évaluée aussi bien dans cette unité de formation que dans les unités scientifiques. Un enseignant référent accompagnera l'étudiant dans la préparation de certains exposés, aussi bien à l'écrit qu'à l'oral.

Objectifs. L'étudiant devra être capable de (i) décoder les problèmes de Gestion des Ressources Humaines (GRH), les situer dans leur contexte et proposer des solutions pertinentes pour les résoudre, (ii) évaluer l'efficacité des diverses pratiques de GRH et les interrelations qui existent entre elles, (iii) faire l'analyse d'une situation de groupe.

Place dans la formation. Aucun prérequis.

Module 1 Management d'équipe et Psychologie sociale (étudiants INSA)

Intervenants : Enseignants INSA

Volume horaire : 35h en présentiel

Coefficient : 1

Description. Ce module comprend :

- une description de ce qu'est une notion de groupe, un processus de décision, la gestion des conflits,
- les notions d'autorité, de minorités actives, des influences qui peuvent s'exercer,
- une analyse des emplois, du recrutement, des motivations au travail, la rémunération, l'appréciation des salariés, la formation, gestion des carrières, relations professionnelles, flexibilité des RH et contrats de travail.

L'enseignement sera effectué sous forme de cours, d'études de cas, d'exposés et de jeux de rôles. Une mise en situation de 3 jours autour d'activités sportives est au programme.

Evaluation. Contrôle continu avec travail sur des études de cas à toutes les séances, exposés en groupe et un examen final sur table.

Modalités de rattrapage. Examen écrit.

Module 2 Sport (étudiants INSA et ENSAT)

Intervenants : Enseignants INSA

Volume horaire : 20h en présentiel

Coefficient : 1
Description. Voir module 1.
Evaluation. Evaluation du comportement en groupe pendant les 3 jours de stage sportif.

Module 3 Parcours Professionnel Individualisé (étudiants INSA et ENSAT)

Intervenants : Enseignants INSA
Volume horaire : 10h en présentiel
Coefficient : 1
Description. Le Parcours Professionnel Individualisé (PPI) a pour vocation d'accompagner les élèves-ingénieurs dans la construction de leur projet professionnel, de développer leurs compétences et d'accroître leur employabilité dans une perspective humaine durable et globale. Des ingénieurs et professionnels d'entreprises accompagnent les étudiants dans une dynamique individuelle et collective : entretiens individuels, simulations de recrutement, mises en situations, sessions de groupes, équipes-projets, rencontres, conférences-métiers et/ou tables rondes.
Evaluation. Un rendu écrit et une simulation d'entretien.

UF7 Projet « Défi en bio-informatique » & Anglais

Référent : C. Laplanche (INP-ENSAT)

Intervenant : C. Laplanche (INP-ENSAT), A. Alibert (Anglais, INP-ENSAT)

ECTS : 2 (INSA) et 6 (ENSAT)

Volume horaire INSA : 5h20 en présentiel et 33h20 en autonomie = 38h40 étudiant

Volume horaire ENSAT : 9h20 en présentiel et 66h40 en autonomie = 76h étudiant

Finalités. L'étudiant devra être capable de réaliser une tâche complexe dans le domaine de la biologie computationnelle, sous la forme d'un projet collaboratif.

Objectifs. L'étudiant sera capable de résoudre une tâche d'ingénierie complexe dans le domaine de la biologie computationnelle pour les biotechnologies ; de travailler efficacement sous la forme d'un projet collaboratif ; de structurer, synthétiser et présenter ses résultats, ainsi que d'animer une discussion à la pointe dans le domaine, en français et en anglais.

Liens avec le référentiel de compétences ENSAT. Cette UF permet aux étudiants d'acquérir les compétences et apprentissages critiques suivants :

- Diagnostiquer, via l'Apprentissage Critique DIAG 3.1 : « Sélectionner et mettre en œuvre des méthodes d'analyse et de traitement pertinentes en fonction de la demande du prescripteur »,
- Valider, via l'Apprentissage Critique VAL 3.1 : « Manipuler des données complexes par leur hétérogénéité ou leur taille »,
- Communiquer, via le Jalon COM 3 : « Adapter sa communication à un public extérieur et à la situation en représentant une entité (soi-même, un groupe de travail, un organisme, une entreprise) ».

Place dans la formation. La réalisation du défi fera appel aux connaissances techniques développées dans les UFs 1 à 5 et 8, et aux connaissances relatives à la gestion d'un projet en entreprise acquises dans l'UF6. L'expérience acquise dans cette UF, en termes de mise en œuvre d'un projet, sera remobilisée pendant le stage de fin d'études (UF9).

Module 1 Organisation d'une table ronde au forum carrière (étudiants ENSAT)

Intervenant : C. Laplanche

Volume horaire : 2h40 en présentiel, 20h en autonomie = 22h40 étudiant

Coefficient : 1

Description. Les étudiants organiseront et animeront une table ronde au forum carrière ENSAT, à laquelle ils inviteront à participer des professionnels du domaine de la biologie computationnelle.

Evaluation. La qualité des échanges qui auront lieu lors de la table ronde permettra d'évaluer l'acquisition de la compétence Communiquer, via l'Apprentissage Critique COM 3.2 « Sensibiliser un public à des questions d'avenir (durabilité, choix politiques, préoccupations environnementales, trajectoires de changement, …) ».

Modalités de rattrapage. Proposition de pistes d'amélioration en réponse aux problèmes rencontrés.

Module 2 Définition du « Défi en bio-informatique » (étudiants ENSAT)

Intervenant : C. Laplanche

Volume horaire : 1h20 en présentiel, 13h20 en autonomie = 14h40 étudiant

Coefficient : 1

Description. Les étudiants contacteront des entreprises et définiront, avec l'une d'entre elles, un projet lié à un besoin en biologie computationnelle. Les étudiants matérialiseront le projet sous la forme d'une lettre de mission, qui inclura un cahier des charges formalisant la demande et une description de la démarche en terme de gestion de projet qui sera suivie pour réaliser le projet (PBS, matrice de risques, GANTT). Le projet sera réalisé dans le module 3 de cette UF.

Evaluation. L'obtention d'un projet à l'issu du démarchage auprès des entreprises qui correspond aux objectifs de la formation permettra d'évaluer la compétence Communiquer via l'Apprentissage Critique COM 3.3 « Convaincre et négocier pour arriver à mener à bien un projet (obtenir un job, un financement, un partenariat, des données, …) ».

Modalités de rattrapage. Révision de la lettre de mission.

Module 3 Réalisation du « Défi en bio-informatique » (étudiants ENSAT et INSA)

Intervenants : C. Laplanche, A. Alibert

Volume horaire : 5h20 en présentiel, 33h20 en autonomie = 38h40 étudiant

Coefficient : 1

Description. Les étudiants réaliseront le projet défini dans le module 2 de cette UF. Le projet permettra de résoudre, sous la forme d'un projet collaboratif, une tâche complexe d'ingénierie répondant à un besoin d'une entreprise, d'un laboratoire, etc. dans le domaine de la biologie computationnelle.

Evaluation. Le travail réalisé en analyse de données permettra d'évaluer l'acquisition des Apprentissages Critiques DIAG 3.1 « Sélectionner et mettre en œuvre des méthodes d'analyse et de traitement pertinentes en fonction de la demande du prescripteur » et VAL 3.1 « Manipuler des données complexes par leur hétérogénéité ou leur taille ». La présentation du projet lors d'une soutenance à mi-parcours sous la forme d'un « project review meeting » (en anglais) ainsi que lors de la soutenance finale suivie d'une discussion permettront d'évaluer la compétence Communiquer via l'Apprentissage Critique COM 3.1 « Informer un public extérieur des résultats obtenus dans le cadre d'un projet ».

Modalités de rattrapage. Nouvelle soutenance orale.

UF8 Projet « RNA-Seq »

Référent : Farid Regad (INP-ENSAT)

Intervenants : Guillaume Madignier (INP-ENSAT), Elie Maza (INP-ENSAT), Farid Regad (INP-ENSAT)

ECTS : 3

Volume horaire : 30h40mn en présentiel + 9h20mn en autonomie = 40h étudiant

Finalités. La finalité de cette UF projet est de permettre aux étudiants de répondre à une question biologique, portant sur l'expression génique d'un modèle vivant, au travers de la réalisation d'une expérience de transcriptomique. Cette UF fournira à tous les étudiants l'ensemble des prérequis nécessaires pour mener à bien une telle expérience depuis sa conception jusqu'à l'interprétation des résultats. Plus précisément, cette UF permettra de mobiliser conjointement les notions de planification et de réalisation d'expériences, et les notions d'informatique, de statistique et de génomique nécessaires au traitement et à l'exploitation des données RNA-Seq.

Objectifs. A la fin de cette UF, l'étudiant sera capable de concevoir et de préparer un plan expérimental associé à une question biologique posée, de réaliser les extractions d'ARN en vue de leur séquençage à haut débit en plateforme, de mettre en place un control qualité des données issues du séquençage, de travailler sous l'environnement *Unix* pour la manipulation des fichiers issus du séquençage et pour le lancement de leurs traitements et analyses sur un cluster de calcul, d'automatiser ces traitements et analyses informatiques, d'effectuer les analyses statistiques sur R (différentiel d'expression, …), et de répondre à la question posée initialement au vu des résultats obtenus.

Liens avec le référentiel de compétences ENSAT. Cette UF permet aux étudiants de mettre en pratique l'ensemble des 3 apprentissages critiques du jalon 2 de la compétence Concevoir. De plus, cette UF permet également aux étudiants d'acquérir les compétences et apprentissages critiques suivants :

- Diagnostiquer, via l'Apprentissage Critique DIAG 3.1 : Sélectionner et mettre en œuvre des méthodes d'analyse et de traitement pertinentes en fonction de la demande du prescripteur
- Valider, via l'Apprentissage Critique VAL 3.1 : Manipuler des données complexes par leur hétérogénéité ou leur taille
- Valider, via l'Apprentissage Critique VAL 3.2 : Evaluer les enjeux et les conséquences de la décision découlant des résultats
- Valider, via l'Apprentissage Critique VAL 3.3 : Proposer des améliorations méthodologiques en vue de consolider l'évaluation

Place dans la formation. La réalisation de cette UF projet fait appel aux connaissances techniques développées dans les UFs 1 à 3. Les compétences acquises pourront être remobilisées pour l'UF7.

Evaluation. Restitution écrite, par groupe, et présentation orale.

Modalités de rattrapage. Oral sur la base de la restitution écrite.

UF9 Stage de fin d'études

ECTS : 30

Finalités. Les élèves-ingénieurs doivent réaliser un stage de 5 à 7 mois, donnant lieu à un projet de fin d'études (PFE) présenté à la fois dans un rapport écrit (appelé mémoire de fin d'études) et lors d'une soutenance orale devant un jury composé d'enseignants et du maître de stage.

UF10 Entreprise

Lors du **semestre 9**, les élèves, en contrat de professionnalisation ou en contrat d'apprentissage, alternent des **périodes à l'école** et **des périodes en entreprise**. Quelle que soit la nature du contrat, les élèves ont le même rythme d'alternance pour leur dernière année de formation d'ingénieur agronome.

Les **périodes en entreprise** sont validées par l'**UF Entreprise**. Cette UF permet à l'alternant de valider **10 crédits ECTS** ; les UF rattachées à la formation académique comptant pour 20 crédits ECTS.

Les **modalités d'obtention** de ces 10 crédits ECTS diffèrent selon que l'alternant est lié à l'entreprise par un contrat de professionnalisation ou par un contrat d'apprentissage.

Cas de l'alternant en contrat de professionnalisation

L'attribution des 10 crédits ECTS repose sur deux évaluations :

- Une **évaluation des missions** réalisées par l'alternant en entreprise : 6 crédits ECTS. Cette évaluation est réalisée par le tuteur professionnel.
- Une évaluation d'un **rapport d'étonnement et sa soutenance** : 4 crédits ECTS. Cette évaluation est réalisée par le tuteur professionnel et le tuteur pédagogique.

Ces 2 évaluations reposent sur des **grilles** qui sont à compléter par les 2 tuteurs de l'élève sur le Livret Electronique de l'Alternant (LEA) de MidiSup. L'élève ainsi que les deux tuteurs reçoivent en début d'année universitaire un **guide** pour les aider à se connecter au LEA et à compléter les grilles d'évaluation.

Le **rapport d'étonnement**, d'une dizaine de pages, conforme aux normes de rédaction habituelles, devra présenter l'entreprise et mettre en avant son fonctionnement avec des points forts et des points possibles d'amélioration.

Habituellement, un rapport d'étonnement permet de répondre aux questions suivantes :

- Qu'est-ce qui vous a le plus étonné dans l'entreprise ?
- Quel est le point fort qui vous a le plus surpris ?
- Quel a été pour vous le point faible le plus inattendu ?
- Qu'est-ce qui devrait être amélioré, modifié ou abandonné prioritairement selon vous ?
- Si vous aviez une baguette magique, quelle est la chose que vous changeriez dans votre entreprise ?
- Qu'est-ce qui vous a étonné dans la manière dont l'entreprise sert ses clients ?
- Quelle est la force des produits et/ou services que vous ne soupçonniez pas avant de travailler dans l'entreprise ?
- Quel est la faiblesse ou le manquement qui vous inquiète le plus dans les produits et les services ?
- Dans les relations interpersonnelles, qu'est-ce qui vous a étonné ?
- Quelles sont les améliorations concrètes que vous suggéreriez ?

La **soutenance** a lieu en entreprise ou en visio et dure 15 minutes. Elle est suivie par une discussion.

Cas de l'alternant en contrat d'apprentissage

L'attribution des 10 crédits ECTS repose sur l'**évaluation des compétences** de l'apprenti sur la base des missions confiées par l'entreprise. Les compétences évaluées sont celles du **Référentiel de Compétences** de l'élève-ingénieur de l'AgroToulouse.

Contrairement aux semestres précédents (semestres 5 à 8), le semestre 9 de l'apprenti est évalué **seulement** sur la base des missions qui lui sont confiées. Autrement dit, l'UF Entreprise de ce semestre ne comporte pas de projet semestriel.

Cette évaluation des compétences est réalisée **conjointement** par le maître d'apprentissage et le tuteur pédagogique en présence de l'apprenti au cours d'une rencontre entre les différentes parties.

Le tuteur pédagogique complète l'onglet « Evaluation 3A (S9) » du fichier Excel de suivi de l'apprenti (fichier disponible sur Moodle). En parallèle, au cours d'un entretien mené à la fin de chaque période en entreprise, l'apprenti et le maître d'apprentissage remplissent de façon concertée la fiche de liaison, sur le LEA de MidiSup ; fiche détaillant la nature des missions conduites sur la période et celles envisagées sur la période suivante, l'avis du Maitre d'apprentissage sur le comportement en entreprise et les points à améliorer.

Codes et domaines INSA et codes ENSAT des UF de BioComp

Numéro UF	Code INSA	Domaine INSA	Code ENSAT
UF 1	I5GBBC01	Outils et méthodes en bioinformatique	
UF 2	I5GBBC02	Bioinformatique pour la génomique, la post-génomique et la structure des macromolécules	
UF 3	I5GBBC03	Bioinformatique pour la génomique, la post-génomique et la structure des macromolécules	
UF 4	I5GBBC04	Outils et méthodes en bioinformatique	
UF 5	I5GBBC05	Bioinformatique pour la génomique, la post-génomique et la structure des macromolécules	
UF 6	I5CCGE21	CSH	
UF 7	I5GBBC07	Projets	
UF 8	I5GBBC08	Projets	
UF 9	PFE	PFE	