

Syllabus de la spécialisation

Biologie Computationnelle
pour les Biotechnologies

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

Brice Enjalbert (INSA, brice.enjalbert@insa-toulouse.fr)

2022-2023

Bio Comp

Sommaire

UF1 Outils numériques et concepts fondamentaux	5
Module 1 Extraction et mise en forme des données	5
Module 2 Stat & R	6
UF2 Bio-informatique pour la génomique	7
Module 1 Assemblage <i>de novo</i> et annotation des génomes	7
Module 2 Alignement de séquences et recherche de polymorphisme	7
Module 3 Bio-informatique pour l'épigénome	8
UF3 Post-Génomique	9
Module 1 Transcriptomique	9
Module 2 Protéomique	9
Module 3 Métabolome et Fluxome	10
UF4 Biologie intégrative	11
Module 1 Intégration statistique des données (avec <i>mixOmics</i>)	11
Module 2 Modélisation dynamique d'un système biologique avec régulation	11
UF5 Biologie structurale computationnelle	13
Module 1 Bases théoriques et méthodologiques	13
Module 2 Mise en pratique des méthodes	13
UF6 Gestion des Ressources Humaines et Vie dans les organisations, Anglais et Sport	14
Module 1 Gestion des Ressources Humaines et Vie dans les organisations (étudiants INSA)	14
Module 2 Sport (étudiants ENSAT et INSA)	15
Module 3 Parcours professionnel individualisé (étudiants ENSAT et INSA)	15
UF7 Projet	16
Module 1 Organisation d'une table ronde au forum carrière ENSAT (étudiants ENSAT)	16
Module 2 Définition du « Défi en bio-informatique » (étudiants ENSAT)	16
Module 3 Réalisation du « Défi en bio-informatique » (étudiants ENSAT et INSA)	17
UF8 Stage de fin d'études	18

Tableau récapitulatif du volume horaire et des ECTS par UF

UF	Référent(s)	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	5	30h	40h
UF3 - Post-Génomique	B. Enjalbert	4	33h	9h
UF4 – Biologie intégrative	C. Aceves	4	23h20	11h40
UF5 – Biologie structurale computationnelle	I. André J. Esque	5	35h	21h
UF6– Gestion des Ressources Humaines et Vie dans les organisations, Anglais et Sport	A. Alibert (Anglais ENSAT) F. Rey (Sport INSA) Q. Marechal (Humanité INSA)	INSA : 5 ENSAT : 2	INSA : 67.5h ENSAT : 32.5h	INSA : 0h ENSAT : 0h
UF7 – Projets	C. Laplanche	INSA : 2 ENSAT : 5	INSA : 5h20 ENSAT : 9h20	INSA : 33h20 ENSAT : 66h40
UF8 – Stage de fin d'études		30		
TOTAUX INSA		60	235h	135h
TOTAUX ENSAT		60	203h	169h

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 3^{ème} année, par nature interdisciplinaire, est commune aux deux écoles d'ingénieurs toulousaines que sont l'INP-ENSAT 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 plusieurs des unités de formation scientifiques décrites ci-dessous (UF2 et UF7). Un enseignant d'anglais de l'ENSAT accompagnera les étudiants dans la préparation de certains exposés, 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 (INRA), 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 *Perl* 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 3^{ème} 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 indispensable 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 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. Description/manipulation avec R des outils classiques de l'analyse descriptive uni et bivariable, de la théorie de l'échantillonnage, de l'estimation statistique, des tests statistiques usuels. (3h+2h)
2. 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). (3h+2h)
3. 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. (3h+2h)
4. Introduction aux méthodes de sélection de variables. Introduction au modèle linéaire généralisé (modèle log-linéaire, régression logistique, etc.). Test non paramétrique de Kruskal-Wallis pour données non gaussiennes. (3h+2h)

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

Modalités de rattrapage. Les mêmes que précédemment.

UF2 Bio-informatique pour la génomique

Réfèrent : Mohamed Zouine (INP-ENSAT)

Intervenant : Mohamed Zouine (INP-ENSAT)

ECTS : 5

Volume horaire global : 30h en présentiel, 40h en autonomie

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 2^{ème} et 3^{ème} 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 bissulfite 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

Intervenant : Mohamed Zouine

Volume horaire : 10h en présentiel, 15h 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.

Evaluation. Examen sur machine de 2h commun aux modules 1, 2 et 3.

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

Module 2 Alignement de séquences et recherche de polymorphisme

Intervenant : Mohamed Zouine

Volume horaire : 10h en présentiel, 10h 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. Examen sur machine de 2h commun aux modules 1, 2 et 3.

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

Module 3 Bio-informatique pour l'épigénome

Intervenant : Mohamed Zouine

Volume horaire : 10h en présentiel, 15h en autonomie

Coefficient : 1

Description. Les nouvelles technologies de génomique offrent l'opportunité de pouvoir analyser l'épigénome à différents niveaux : la méthylation du génome et l'acétylation des histones. Les approches bio-informatique permettant de caractériser ces trois niveaux seront abordées sous forme de TD sur machine.

Evaluation. Examen sur machine de 2h commun aux modules 1, 2 et 3.

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), Elie Maza (INP-ENSAT), Pierre Millard (INSA), Mohamed Zouine (INP-ENSAT)

ECTS : 4

Volume horaire global : 33h en présentiel, 9h en autonomie

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

Intervenants : Brice Enjalbert, Elie Maza, Mohamed Zouine

Volume horaire : 17.5h en présentiel, 7.5h en autonomie

Coefficient : 1

Description. Ce module a comme objectif de présenter les approches visant à quantifier les molécules d'ARN, des premières approches aux RNA-Seq.

- Cours (5H15; BE, EM, MZ)

- a. Les ARNs, la transcription, technique de Northern, biopuces, stabilome, overview RNA-Seq (1H15 ; BE).
- b. RNA-Seq : techniques, interprétation, applications (2H00; MZ).
- c. Traitement statistique des données (2H00 ; EM).

- TD (9H45 ; BE, EM, MZ. 11h autonomie étudiants)

- a. Traitement des données : puces à ADN, E. coli, raw jusqu'au GO et visualisation (3H45 ; BE. 6H autonomie).
- b. Traitement des données : RNA-Seq, tomate, raw jusqu'au GO et visualisation (3H00 ; MZ. 3H autonomie).
- c. Traitement statistique des données : DE Analysis (3H00 ; EM. 2H autonomie).

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 : 7.5h 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 (2H30 ; DB, 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 (3H45 ; DB, BE. 2H30 autonomie étudiants)

a. Traitement de données ITRAQ (1H15 ; BE. 1H45 autonomie).

b. Séquençage de protéines et modification post-traductionnelle (2H30 ; DB. 1H45 ; autonomie).

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 : 7.5h en présentiel, 0h 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 (3H45 ; BE, PM. 5H autonomie étudiants).

a. Analyse de métabolome (1H15 ; BE. 1H45 autonomie étudiants).

b. Détermination d'un fluxome (2H30 ; PM. 1H45 autonomie étudiants).

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 (INSA)

Intervenants : C. Aceves (INSA), S. Déjean (IMT)

ECTS : 4

Volume horaire global : 23h20 en présentiel, 11h40 en autonomie = 35h 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 : ???

Volume horaire : 13h20 en présentiel, 6h40 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. L'analyse Canonique des Corrélations (4h présentiel, 2h autonomie). (i) Quand utiliser l'ACC ? (ii) Principes. ACC classique. ACC régularisée. Réglage des paramètres. (iii) Usage avec *mixOmics*, cas d'étude.
2. La régression PLS (4h présentiel, 2h autonomie). (i) Quand utiliser la régression PLS ? (ii) Principes. La régression PLS classique. La régression PLS *sparse*. Réglage des paramètres. (iii) Usage avec *mixOmics*, cas d'étude.
3. L'Analyse Discriminante PLS (4h présentiel, 2h autonomie). (i) Quand utiliser la PLS-DA ? (ii) Principes. PLS-DA classique. PLS-DA *sparse*. Réglage des paramètres. (iii) Usage avec *mixOmics*, cas d'étude.

Evaluation. Examen en salle informatique de 2h commun aux modules 1, 2 et 3.

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 et les ODEs pour un réseau métabolique avec Copasi (2h30 présentiel, 8h45 autonomie). Principes basiques. Analyse de phase. Analyse de sensibilité.

Evaluation. Examen en salle informatique de 2h commun aux modules 1, 2 et 3.

Modalités de rattrapage. Oral

UF5 Biologie structurale computationnelle

Référents : Isabelle André

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

ECTS : 5

Volume horaire global : 33h45 en présentiel, 26h15 en autonomie = 60h 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. Connaître 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

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

Volume horaire : 26h15 en présentiel, 26h15 en autonomie = 52h30 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 Gestion des Ressources Humaines et Vie dans les organisations, Anglais et Sport

Référent : A. Alibert (Anglais ENSAT), F. Rey (Sport INSA), Q. Marechal (Humanité INSA)

Intervenants : Enseignants INSA (GRH, VO) et INP-ENSAT (Anglais).

ECTS : 5 (INSA) et 2 (ENSAT)

Volume horaire global (INSA) : 67.5h en présentiel

Volume horaire global (ENSAT) : 32.5h en présentiel

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 Gestion des Ressources Humaines et Vie dans les organisations (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 ENSAT et INSA)

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 ENSAT et INSA)

Intervenants : Enseignants INSA

Volume horaire : 12.5h 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

Référent : C. Laplanche (INP-ENSAT)

ECTS : 2 (INSA) et 5 (ENSAT)

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

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

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

Objectifs. L'étudiant sera capable de démarcher une entreprise, d'entretenir un contact professionnel, de travailler efficacement en groupe, de définir un cahier des charges, de planifier et réaliser une tâche complexe, de mobiliser des connaissances techniques diverses, de planifier la réalisation d'un évènement, de structurer et présenter ses résultats.

Place dans la formation. La réalisation du défi fera appel aux connaissances techniques développées dans les UF1 à 5 et aux connaissances relatives à la gestion d'un projet en entreprise (UF6). L'expérience acquise dans cette UF en termes de mise-en-œuvre de projet sera remobilisée pendant le stage de fin d'études (UF8).

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

Intervenant : C. Laplanche (INP-ENSAT)

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

Coefficient : 1

Description. Les étudiants inviteront des professionnels du domaine de la biologie computationnelle à participer au forum carrière ENSAT en y présentant leur parcours professionnel. Cette présentation prendra la forme d'une table ronde que les étudiants organiseront et animeront.

Objectifs. L'étudiant sera capable de démarcher une entreprise, d'entretenir un contact professionnel, de travailler efficacement en mode projet en appliquant les outils associés (SMART, SWOT, WBS, GANTT), de planifier la réalisation d'un évènement, d'animer des échanges dans le domaine de la biologie computationnelle.

Evaluation. Qualité du déroulement de la table ronde. Evaluation de la bonne utilisation des outils de gestion de projet.

Modalités de rattrapage. Evaluation réflexive concernant le déroulement de la table ronde au forum carrière et de sa préparation et propositions de pistes d'amélioration.

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

Intervenant : C. Laplanche (INP-ENSAT)

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. Outre la demande de l'entreprise, le cahier des charges élaboré devra prendre en considération les ressources dont disposeront

les étudiants (ressources internes/externes, moyens matériels, calendrier). Le projet sera mis en œuvre dans le module 3 de cette UF.

Objectifs. L'étudiant sera capable de démarcher une entreprise, d'entretenir un contact professionnel, de travailler efficacement en mode projet en appliquant les outils associés (SMART, SWOT, WBS, GANTT), de définir un cahier des charges, de planifier une tâche complexe.

Evaluation. Evaluation de la lettre de mission présentant le cahier des charges et la structure du projet. Evaluation de la bonne utilisation des outils de gestion de projet.

Modalités de rattrapage. Révision de la lettre de mission, révision de l'utilisation des outils de gestion de projet.

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

Intervenants : C. Laplanche

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

Coefficient : 1

Description. Les étudiants réaliseront sous la forme d'un projet collaboratif une tâche complexe d'ingénierie répondant à un besoin d'une entreprise dans le domaine de la biologie computationnelle.

Objectifs. L'étudiant sera capable d'entretenir un contact professionnel, de travailler efficacement en mode projet en appliquant les outils associés, de planifier et réaliser une tâche complexe, de mobiliser des connaissances techniques diverses, de structurer et présenter ses résultats.

Evaluation. Evaluation du livrable par l'entreprise. Evaluation de la qualité de la soutenance orale. Evaluation de la bonne utilisation des outils de gestion de projet.

Modalités de rattrapage. Nouvelle soutenance orale.

UF8 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.