

UFR SCIENCES du VIVANT

ANNÉE 2015 / 2016

L3 Mention : Sciences de la Vie

Parcours Biologie Biomolécules Bio-informatique
(Bio3)

Dominante Bio-informatique



UFR Sciences du Vivant : Bâtiment Lamarck - Case 7044
35 rue Hélène Brion - 75205 Paris Cedex 13

Responsable scolarité : Virginie Bruère
e-mail : virginie.bruere@univ-paris-diderot.fr

☎ 01 57 27 82 33

Responsable pédagogique : Anne Badel
e-mail : anne.badel@univ-paris-diderot.fr

☎ 01 57 27 83 78

Site internet de la formation : <http://biteach.sdv.univ-paris-diderot.fr/lbi/>

Organigramme de la formation

La Licence de **Biologie Biomolécules (Bio³)** comporte 6 semestres. Le niveau L3 correspond aux semestres S5 et S6. Le **semestre L3 S5** est partiellement commun aux 2 parcours Biologie Biomolécules et Bio-Informatique pour la partie Biologie/Biochimie et avec le M1 Biologie-Informatique.

Semestre S5			Semestre S6		
		crédits			crédits
UE 1	Biologie moléculaire	6	UE 1	Bio-informatique	3
UE 2	Génétique	3	UE 2	Projet Professionnel 3	3
UE 3	Métabolisme	3	UE 3	Génomique fonctionnelle et génétique	3
UE 4	Programmation I	6	UE 4	Travaux Pratiques	3
UE 5	Algorithmique	3	UE 5	Programmation II	6
UE 6	Base de données I	3	UE 6	Bio-statistiques avancées	3
UE 7	biostatistique	3	UE 7	Modélisation de systèmes biologiques	3
UE 8	Anglais scientifique	3	UE 8	UE de spécialisation	6

Validation du diplôme de la licence

Au sein d'un semestre, les ECUEs et les UEs se compensent entre elles. Il peut cependant être préférable de refuser une compensation pour améliorer sa note globale.

Les modalités de contrôle des connaissances de chaque UE sont décrites dans le tableau suivant.

Quelques rappels :

- Les **TD** et les **TP** sont **obligatoires**.
- Toute absence lors d'un TP ou toute absence prolongée devront être justifiées (au secrétariat de Licence). Une absence non justifiée sera sanctionnée par un zéro.

BIOLOGIE MOLECULAIRE

Responsable pédagogique : P. Nadal

Résumé du programme :

- 1 Techniques de Génie génétique : l'ADN, caractérisation et détection. Isolement et caractérisation d'un fragment d'ADN particulier, PCR, clonage et séquençage, banques génomiques et ADNc.
- 2 Structure de l'ADN : superstructure de l'ADN, compaction, courbure, super-enroulement, topologie de l'ADN circulaire. Interactions ADN - protéines : exemples et méthodes d'étude.
- 3 Réplication de l'ADN : l'ADN polymérase III de E. coli : mécanisme des étapes du démarrage et de l'élongation; problèmes topologiques. Réplication des plasmides : compatibilité et maintenance.
- 4 Généralités sur la transcription : les ARN polymérases : Exemples chez les phages, les procaryotes, les eucaryotes. Les promoteurs et la régulation de la transcription. Transcription des ARN ribosomaux et des ARN de transfert. Méthodes d'étude des transcrits.
- 5 Régulation de la transcription : L'opéron lactose, régulation par l'inducteur et répression catabolique. Régulation de la transcription chez les phages : Exemple des phages lambda et T4. Transcription eucaryote : Généralités ; régulations ; épissage. Les virus eucaryotes : Adénovirus. Notion de vecteur pour la thérapie génique.
- 6 Traduction et sa régulation : Code génétique : Les ARNt assurent la correspondance triplet / acide aminé. Structure des ARNt et notions « d'identité » des ARNt, de codons synonymes et de Wobble. Mutants suppresseurs de non sens : Le code génétique est modifié. Le ribosome bactérien : Les sites A, P, E. La dynamique de la traduction : facteurs d'initiation, d'élongation et de terminaison. Régulation de la synthèse des protéines et des ARNr qui composent les ribosomes. Rôle des structures secondaires de l'ARNm dans la régulation de la traduction, accessibilité des sites de démarrage.
- 7 Techniques d'analyse de l'expression : Techniques mettant en jeu l'hybridation ARN-sonde spécifique. Gènes rapporteurs. Vecteurs d'expression.

Compétences visées :

Acquérir les connaissances de biologie moléculaire permettant l'accès aux enseignements de Masters

Horaires : 38h CM + 22 TD

Modalités d'évaluation : le contrôle continu consiste en deux mini-projets, le premier comme initiation, le second commun aux matières d'algorithmique et de programmation I. Les modalités sont :

- 1ère session : 55% CT + 45% CC
- 2nde session : 55% CT + 45% CC ou 100% CT

GENETIQUE

Responsable pédagogique : P. Silar

Résumé du programme :

Histoire de la génétique
Génétique mendélienne : caractères, gènes et génomes
Mitose - Méiose
Relation génotype / phénotype
Du gène à l'ADN
La démarche génétique : crible de sélection de mutants
Notions de Génétique Humaine

Compétences visées :

Connaître les règles de bases de la génétique

Horaires : 8h CM + 22h TD

Modalités d'évaluation :

- 1ère session : 100% CT
- 2nde session : 100% CT

METABOLISME

Responsable pédagogique : A. Méjean

Résumé du programme :

- 1 MÉTABOLISME des SUCRES : Glycolyse, néoglucogenèse, voie des pentoses. Glycogénolyse, glycogénogenèse. Métabolisme chez les microorganismes, plantes, vertébrés. Insuline, glucagon, diabète.
- 2 CYCLE de KREBS : Présentation. Aspect énergétique. Aspects anabolique et catabolique. Voies primitives du cycle de Krebs.
- 3 RESPIRATION, FERMENTATION : Chaîne respiratoire. Force proton motrice, potentiel de membrane, Δ pH. Synthèse d'ATP par l'ATP synthase. Respiration chez les différents organismes. Fermentations et synthèse d'ATP au niveau du substrat. Maladies et mutants des chaînes respiratoires.
- 4 MÉTABOLISME des LIPIDES : Acétyl ~ CoA, métabolite central. Dégradation des lipides. Synthèse des lipides. Anabolisme à partir des lipides ou de l'acétate : le shunt glyoxylique. Transport des lipides dans le sang. Régulations hormonales.
- 5 MÉTABOLISME des ACIDES AMINÉS et de l'AZOTE : Assimilation de N₂ (nitrogénase), NH₄⁺, acides aminés. Glutamate déshydrogénase, glutamate synthase, glutamine synthétase, transaminases. Biosynthèse des acides aminés (acides aminés essentiels chez l'homme). Dégradation des acides aminés. Excrétion de l'azote sous forme de NH₄⁺, urée, acide urique, cycle de l'urée. Régulations hormonales du métabolisme des acides aminés.
- 6 PHOTOSYNTHÈSE : Absorption de la lumière par les photo-systèmes. Chaînes de transport d'électrons de O₂ vers NADPH et synthèse d'ATP. Assimilation du CO₂, cycle de Calvin et cycle des plantes C₄.

Compétences visées :

Connaître les règles de bases de la génétique

Horaires : 18h CM + 12h TD

Modalités d'évaluation :

- 1ère session : 60% CT + 40% CC
- 2nde session : 60% CT + 40% CC ou 100% CT

PROGRAMMATION I

Responsable pédagogique : V. Padovani

Résumé du programme :

Le cours couvre tous les aspects fondamentaux de la programmation, à l'exception de la compilation séparée et des pointeurs de fonctions.

- types, variables, expressions, types de base,
- structures de contrôle,
- définitions de fonctions, passage par valeur,
- tableaux et chaînes de caractères, tableaux multidimensionnels,
- pointeurs, arithmétique des pointeurs,
- allocation dynamique,
- types de structures et d'unions,
- listes chaînées, arbres binaires et programmation récursive.

Les exemples sont réalisés à l'aide du langage C ou d'un autre langage non orienté objet.

Bibliographie : "Le Langage C", B.W. Kernighan et D.M. Ritchie, MASSON

Compétences visées :

Acquérir les bases de la programmation impérative.

ECUE 1 : Initiation, 7h cours + 15h TP

ECUE 2 : Base de la programmation, 15h cours + 20h TP

Modalités d'évaluation : le contrôle continu consiste en deux mini-projets, le premier comme initiation, le second commun aux matières d'algorithmique et de programmation I. Les modalités sont :

- 1ère session : 70% examen final + 30% de la note de contrôle continu
- 2nde session : examen session 2

Algorithmique

Responsable pédagogique : R. Mantaci

Résumé du programme :

Le cours introduira la notion de complexité algorithmique et présentera des méthodes générales pour la conception d'algorithmes efficaces. Dans la première partie, on s'intéressera au problème fondamental du tri d'un ensemble de valeurs/objets, en présentant différentes méthodes (tris élémentaires, le tri fusion, le tri rapide...) dont les complexités seront évaluées et comparées. Dans la seconde partie, on s'intéressera aux structures arborescentes qui permettent entre autre d'obtenir des algorithmes efficaces pour la recherche d'une donnée dans un ensemble. En particulier, on présentera la structure de tas, les arbres binaires de recherche, les fonctions de hachage.

Compétences visées : Ce cours est complémentaire du cours de programmation. Alors que le cours de programmation vise l'apprentissage de la syntaxe du langage de programmation choisi par son responsable, le cours d'algorithmique vise les compétences suivantes :

- l'évaluation des performances (essentiellement en termes de temps de calcul et de mémoire utilisée) des algorithmes qui seront ensuite traduits en programmes ;
- la conception d'algorithmes performants à l'aide de stratégies algorithmiques telles que le « diviser pour régner », ou de structures de données efficaces comme les structures arborescentes.

Horaires : 18h cours + 18 h TD

Modalités d'évaluation :

session 1 : 70% CT + 30% CC
session 2 100% CT

BASE DE DONNEES I

Responsable pédagogique : W. Zielonka

Résumé du programme :

- 1 Notions d'architecture d'un système de gestion de bases de données.
 - 2 Modélisation de base de données.
 - 3 Transformation de modèle en définition de données SQL.
 - 4 Insertion, suppression et modification de données.
 - 5 Interrogation des bases de données relationnelles : le langage de requêtes SQL.
- Un projet de bases de données sera à réaliser dans le cadre de cet enseignement.

Compétences visées :

Apporter aux étudiants les connaissances théoriques, techniques et pratiques pour leur permettre d'étudier, de concevoir et d'implémenter une base de données relationnelle dans le domaine de la biologie ou de la médecine.

Horaires : 10h cours + 10h TD + 10h TP

Modalités d'évaluation :

- session 1 : 30% CC + 70% CT
- session 2 : 100% CT

BIostatistique

Responsable pédagogique : A. Badel

Résumé du programme :

Après une rapide reprise de la théorie des tests et des tests basés sur la loi normale, les tests non paramétriques seront abordés. L'accent sera ensuite mis sur l'analyse de la variance et les plans factoriels.

Compétences visées :

L'enseignement dispensé dans le cadre de ce module a pour but de donner une formation de base aux étudiants dans le domaine des statistiques appliquées à la Biologie. Le programme de statistiques doit permettre aux étudiants de comprendre, d'appliquer et d'interpréter correctement les méthodes fondamentales du domaine.

Horaires : 12h cours + 10h TD + 10h TP

Modalités d'évaluation :

Session 1 : 50% CC (projet et contrôle continu) + 50% examen final

Session 2 : 100% CT

ANGLAIS SCIENTIFIQUE

Responsable pédagogique : C. Champanet

Résumé du programme :

Les étudiants sont répartis en groupes de niveaux. L'enseignement, adapté à chaque niveau, consiste essentiellement à étudier et à commenter des textes scientifiques et à faire des exposés.

Compétences visées :

Les objectifs du cours d'anglais destinés aux étudiants se destinant à un mastère sont d'acquérir les compétences suivantes :

- faire un exposé en anglais sur un sujet scientifique
- rédiger une étude de texte dans un anglais scientifique
- soutenir une conversation supposant un bon niveau de compréhension orale
- lire couramment un texte scientifique
- maîtriser suffisamment la grammaire anglaise pour exprimer clairement ses idées

Horaires : 22h CM/TD

Modalités d'évaluation :

Session 1 : 50% CC + 50% CT

Session 2 : 50% CC + 50% CT

BIOINFORMATIQUE

Responsable pédagogique : S. Pasquali

Résumé du programme :

- 1 La bioinformatique : Pour quoi faire ?
- 2 Fondements : matrices, scores, alignements
- 3 Bases de données biologiques : Swiss-Prot, Genbank, prosite, Prodom, PDB, PFAM, Bases de génomes
- 4 Recherche d'homologie et fonctions biologiques

Compétences visées :

L'objectif de cet enseignement est

- d'initier l'étudiant à quelques principes de la bio-informatique, notamment la recherche de propriétés biologiques sur les serveurs WEB.
- d'être capable d'étudier la structure d'une macromolécule biologique pour en comprendre et mettre en évidence les éléments qui peuvent expliquer sa fonction/activité biologique.

Horaires : 16h CM + 6 h TD + 8 h TP

Modalités d'évaluation :

Session 1 : 40% CC + 60% CT

Session 2 : 100% CT

PROJET PROFESSIONNEL, PP3

Responsable pédagogique : F. Deshayes

Résumé du programme :

Les enseignements de PP3 ont pour but d'engager une réflexion des étudiants sur leur avenir professionnel : Quel métier veulent-ils faire ? Comment accéder à ce métier concrètement ?

La méthodologie employée au cours de ces enseignements consiste d'une part en une discussion constructive autour de leur projet professionnel et d'autre part à rencontrer étudiant dans le cursus visé. L'étudiant devra également constituer un dossier de candidature à un emploi désiré ou à l'entrée dans le M1 souhaité avec une simulation d'entretien pour un laboratoire.

Compétences visées :

Après une interrogation sur ses motivations et son projet professionnel, l'étudiant doit contacter des étudiants dans leur futur cursus ainsi que remplir des dossiers de candidature dans la profession ou le cursus désiré. Il apprendra comment mener un entretien et rédiger un CV.

Modalités d'évaluation : la présence est obligatoire, même avec certificat médical, les étudiants ne peuvent manquer plus de 2 séances.

1^{ère} session : 30% CC + 70% final

2^{ème} session : 100% final

GENOMIQUE FONCTIONNELLE

Responsable pédagogique : N. Janel

Résumé du programme :

- 1 Cartographie et structure des génomes
- 2 Analyse du transcriptome, du protéome, de l'interactome. Exemples d'application en pathologie.

Compétences visées :

Aborder l'analyse globale des génomes et leur fonction par l'appréhension des techniques combinées de données.

Horaires : 14h CM + 12 h TD

Modalités d'évaluation :

- 1^{ère} session : 30% CC + 70% final
2^{ème} session : 100% final

TRAVAUX PRATIQUES

Responsable pédagogique : A. Badel & G. Moroy

Résumé du programme :

TP BM

Mutagenèse par insertion d'un transposon chez *Escherichia coli*.
Sélection de diverses mutations dans l'opéron lactose.
Construction de méro-diploïdes par conjugaison et test de dominance et récessivité pour diverses mutations de *lac I*, gène du répresseur de l'opéron lactose.
Localisation des modifications provoquées sur la structure tridimensionnelle de la protéine.
Importance du dosage de gène pour une régulation.

TP pymol

TP de Bioinformatique structurale : Etude de la relation structure/fonction/activité des macromolécules biologiques par des méthodes bioinformatiques à l'aide du programme de graphisme moléculaire PyMOL : structure des protéines (structure primaire à quaternaire), interaction protéine-ligand et protéine-acide nucléique, notion de flexibilité des protéines, structure de biomolécules non protéique.

Horaires : 30h

Modalités d'évaluation :

- 1^{ère} session : 100% TP
2^{ème} session : 100% TP

PROGRAMMATION AVANCEE

Responsable pédagogique : F. de Montgolfier

Résumé du programme :

L'objectif principal de cet enseignement est d'amener les étudiants à concevoir, réaliser et tester un projet informatique ambitieux en rapport direct avec la biologie à partir d'un cahier des charges pré-défini. En pratique cet enseignement complète ceux du premier semestre par (i) des cours en programmation qui font suite à ceux du premier semestre et (ii) de la programmation d'algorithmes spécifiques. Ces connaissances et exercices prépareront la réalisation du projet lui-même.

En somme, l'enseignement comporte aussi bien un apprentissage théorique qu'un apprentissage par la pratique.

Compétences visées :

- structuration et modélisation de l'information biologique
- écriture et programmation en C d'un algorithme
- conception d'un logiciel

Horaires :

EC Programmation avancée : 16h C + 20h TP

EC Projet de programmation : 16h C / TP

Modalités d'évaluation (CC, examen final...)

Session 1 : 50% CC + 50% CT

Session 2 : $\max(25\%CC+75\%CT, 100\%CT)$

BIostatistiques Avancées

Responsable pédagogique : L. Regad

Résumé du programme :

En biostatistique, deux problèmes principaux ressortent et seront approfondis dans ce cours :

1. le problème de l'estimation : estimation et estimateur, maximum de vraisemblance, rapport de vraisemblance
2. les modèles linéaires et log-linéaires : les modèles de régression linéaires cherchent à expliquer ou à prédire une variable d'intérêt Y à partir d'un ensemble de variables explicatives (X_1, X_2, \dots, X_p) . Ils sont largement appliqués en biologie. Ces modèles, très simples, apportent souvent une description adéquate et interprétable de l'influence des variables explicatives sur la variable d'intérêt Y . Finalement, on peut appliquer des transformations de variables qui permettent d'étudier largement l'impact des variables explicatives sur la variable Y . La validation des approches de prédiction sera présentée.

De plus les méthodes de re-échantillonnages, largement utilisées en bioinformatique seront étudiées.

L'ensemble de ces connaissances sera appliqué lors d'un projet sur des données à l'aide d'un logiciel de statistique (R), très utilisé dans le domaine de la bioinformatique.

Horaires : 12h CM + 4h TD + 8h TP

Modalités d'évaluation :

1^{ère} session : 40% CC + 60% CT

2^{ème} session : 100% CT

Modélisation des systèmes biologiques

Responsable pédagogique : S. Eskiizmirli

Résumé du programme :

L'objectif de ce module est de donner des méthodes et outils pour analyser ou prédire des phénomènes issus de la biologie (ce qui fera entre autre appel aux outils mathématiques). Pour cela, on va

- 1/ montrer comment formaliser ou modéliser des connaissances de systèmes biologiques,
- 2/ analyser et étudier la modélisation des problèmes biologiques étudiés afin de montrer quel est l'apport de cette modélisation pour la compréhension du système biologique,
- 3/ mettre en évidence des propriétés nouvelles.

Compétences visées :

Modélisation, analyse de modèles. Prise en main d'un logiciel de calcul numérique.

Horaires : 15h cours + 20h TD/TP

Modalités d'évaluation : 35% CC + 65% examen final

SPECIALISATION

Responsable pédagogique : A. Badel

2 ECUE au choix parmi les suivantes :

Bases de données II (I. Fagnot)

Pré-requis : une première approche des bases de données telle qu'elle est enseignée dans le cours de Bases de Données du L3 de Bioinformatique.

Résumé du programme :

- 1 Programmation en langage procédural (PL/PGSQL). Programmation de Triggers.
- 2 Compléments de modélisation. Notions de normalisation.
- 3 Algèbre relationnelle.
- 4 Index. Notions d'optimisation des requêtes.
- 5 Accès concurrents à une base de données, transactions.

Un projet de bases de données sera à réaliser dans le cadre de cet enseignement.

Compétences visées :

Approfondir les connaissances acquises au 1er semestre dans le cours de bases de données et les compléter sur le plan théorique et surtout pratique.

Horaires: 15h cours + 15h TD

Modalités d'évaluation :

- Session 1 : 30% projet + 70% examen
- Session 2 : max (30% projet + 70% examen, 100% examen)

Automates et langages formels (K. Pakdaman)

Résumé du programme :

Notions d'alphabet, de mot, de langage, de langage rationnel et de langage reconnaissable. Définition d'automates. Détermination. Théorème de Kleene. Construction d'un automate reconnaissant un langage rationnel dont on connaît l'expression rationnelle. Construction de l'expression rationnelle d'un langage reconnaissable étant donné un automate qui le reconnaît. Lemme d'itération pour les langages rationnels et preuves de non-rationalité. Automate minimal et minimisation d'automates. Applications à la recherche de motifs.

Langages libre de contexte (context-free languages): définition par les grammaires algébriques et par les automates à pile. Version déterministe et non déterministe des automates à pile. On ne fera pas la théorie générale on se contentera de montrer sur des exemples l'équivalence entre spécification par automates à pile et par grammaires algébriques.

Compétences visées :

Le cours se veut une introduction à la théorie de la reconnaissance automatique des langages par moyen de machines dites « automates finis ». On apprendra à identifier quelles propriétés des mots peuvent être « testées » par le moyen des automates. Les applications à la biologie moléculaire sont évidentes puisqu'une

séquence de bases azotées sur les nucléotides de l'ADN ou de l'ARN n'est rien d'autre qu'un mot sur l'alphabet {A, C, G, T} (ou {A, C, G, U}). Les automates permettent donc une analyse automatique efficace du génome ainsi que la recherche rapide de certaines sous-séquences, notamment de gènes. Les langages libres de contexte sont une façon de modéliser la structure secondaire du génome.

Horaires : 15h cours + 15h TD

Modalités d'évaluation :

Session 1 : 70% Examen Final + 30% CC (projet + partiel)

Session 2 : 100% Examen Final

Neurosciences (S. Eskiizmirliler)

Résumé du programme :

Donner une formation de base dans le domaine des Neurosciences et une introduction au domaine des Neurosciences Computationnelles. Pendant les cours l'étude du système nerveux central est effectuée de bas en haut, c.à.d., commençant par la description du fonctionnement d'un neurone, de la formation du potentiel d'action, de sa transmission allant jusqu'à l'étude des structures anatomophysiologiques de certains aspects principaux du contrôle sensori-moteur tel que l'apprentissage (plasticité synaptique), LTP, LTD, Vision, Sommeil. Les dernières avancées dans le domaine de l'Interface-Cerveau-Machine seront également exposées et discutées. Pendant les TPs les simulateurs "Neuron" et "Brian" seront utilisés afin de compléter ce contenu théorique par des expérimentations en simulation.

Compétences visées :

Double compétence en neurobiologie cellulaire et computationnelle.

Horaires : 20 h cours + 12 h TP

Modalités d'évaluation :

Session 1 : 60% examen final + 40% compte-rendu des TP

Session 2 : 100% examen final