

## Cours IFT3295/BIN6000/IFT6291 Bio-Informatique

### 1 Description

Ce cours est destiné aux étudiants désireux de découvrir le domaine de la biologie computationnelle. Aucun prérequis en biologie n'est nécessaire: le cours commencera par une introduction au code génétique. Des connaissances biologiques supplémentaires seront introduites, au besoin, tout au long de la session. Cependant, des bases suffisantes en algorithmique et programmation sont souhaitables.

Les thèmes biologiques abordés concernent l'analyse des séquences biologiques (ADN, ARN, protéines), et l'étude de leur évolution. Le cours sera orienté vers les méthodes algorithmiques utilisées pour aborder les problématiques biologiques soulevées.

### 2 Objectifs

L'objectif de ce cours est d'initier l'étudiant au domaine de la "biologie computationnelle", qui consiste à formaliser un problème biologique, et à développer les outils algorithmiques et mathématiques permettant de le résoudre. À l'issue de ce cours, l'étudiant sera familiarisé avec les problématiques bio-informatiques classiques liées à l'analyse et au traitement des séquences biologiques, ainsi qu'avec les méthodes algorithmiques les plus utilisées pour les résoudre.

Dans le cadre du cours BIN6000/IFT6295, l'objectif est également d'initier l'étudiant à la recherche en biologie computationnelle. Dans ce cadre l'étudiant devra réaliser un projet permettant d'approfondir un sujet du cours, basé sur une revue de littérature. Ce projet donnera lieu à un rapport et à une présentation orale.

### 3 Programme

Les heures indiquées sont des estimations d'heures de cours. Les heures manquantes correspondent aux heures de démonstration, et aux examens.

1. Introduction au cours et à la Bio-Informatique (**2h**)
2. Alignement de séquences biologiques; méthodes exactes (**6h**)
3. Prédiction de structures secondaires d'ARN (**3h**)

4. Recherche de motifs: méthodes exactes et heuristiques; recherche de répétitions; arbres des suffixes (**9h**)
5. Phylogénie: Modèles évolutifs, inférence (**7h**)
6. Génomique comparative (**6h**)

## 4 Mode d'évaluation

### 4.1 IFT3295

1. Examen Intra 20%
2. Examen Final 40%
3. Quatre TP 40%

Pour réussir il faut une note finale supérieure à 50 et une moyenne pondérée des examens supérieure à 50, c'est-à-dire  $(0.20 \times \text{Intra} + 0.40 \times \text{Final}) / 0.60 \geq 50$ .

### 4.2 BIN6000/IFT6291

1. Examen Intra 20%
2. Examen Final 40%
3. Présentation d'article 20%
4. Projet 20% (Inclu 5% pour le résumé à envoyer le 9 octobre)

Pour réussir il faut une note finale supérieure à 50 et une moyenne pondérée des examens supérieure à 50.

## 5 Quelques manuels de référence

- *Algorithms on Strings, Trees and Sequences - Computer Science and Computational Biology*, Dan Gusfield, Cambridge University Press, 1997.
- *An Introduction to Bioinformatics Algorithms*, Neil C. Jones and Pavel A. Pevzner, The MIT Press, 2004.
- *Bioinformatics Algorithms - Techniques and Applications*, Ion I. Mandoiu and Alexander Zelikovsky eds., Wiley, 2008.
- *Algorithms in Computational Molecular Biology - Techniques, Approaches and Applications*, Mourad Elloumi and Albert Y. Zomaya eds., Wiley Series in Bioinformatics, 2011.

- *Introduction to Computational Molecular Biology*, J. Setubal and J. Meidanis, PWS Publishing Company, 1997.
- *Biological sequence analysis - Probabilistic models of proteins and nucleic acids*, R. Durbin, S. Eddy, A. Krogh and G. Mitchison, 1998.