

Stages courts qualifiants en bio-informatique

plate-forme bio-informatique genouest (IRISA)

Introduction aux systèmes dynamiques

Durée : 2 jours - intervenante : Anne SIEGEL

Objectifs, compétences développées

Diverses approches expérimentales et/ou bio-informatiques permettent d'obtenir des données (quantitatives et /ou qualitatives) relatives aux protéines et à l'expression des gènes. La modélisation de réseaux biologiques fournit un cadre où ces observations expérimentales peuvent être étudiées et analysées, en fonction du type d'observation et de connaissances disponibles.

L'objectif de cette formation est d'introduire les différentes notions nécessaires pour débiter dans le domaine de la modélisation de systèmes en s'appuyant sur des exemples pratiques. On se concentrera sur une application à la modélisation de réseaux transcriptionnels.

Les différentes approches existantes seront discutées et mises en perspective.

Programme

- Définition d'un réseau : un même concept pour de multiples significations :
 - réseau de régulation génique ;
 - réseau métabolique ;
 - réseau protéine-protéine ;
 - corrélations bibliographiques.
- Informations concernant les interactions disponibles dans différentes bases de données :
 - KEGG, Ingenuity, Ecocyc, YGD, MINT, BIND, Transpath...
- Visualisation des réseaux :
 - formats classiques SBML et BioPax ;
 - logiciels : Cytoscape et plugins associés.
- Confrontation et annotation des réseaux et données.
- Exploration des réseaux :
 - annotations Gene Ontology ;
 - différentes corrélations.

Introduction à Cytoscape

Durée : 1 jour - intervenant : François MOREEWS

Objectifs, compétences développées

Cytoscape est un logiciel permettant de visualiser les réseaux d'interactions moléculaires et les voies métaboliques, avec leurs annotations, leurs profils d'expression ou d'autres attributs associés.

L'objectif de cette formation est de proposer une première expérience pratique du logiciel. À l'issue de cette formation, les stagiaires posséderont les bases suffisantes pour une utilisation autonome de cet outil.

Programme

- Présentation de Cytoscape et des sources de données d'interaction.
- Construction d'un premier réseau d'interaction :
 - mise en place de cytoscape ;
 - importation d'un réseau ;
 - manipulations.
- Les plugins de Cytoscape :
 - installation d'un plugin ;
 - présentation de quelques plugins.

Introduction à la suite logicielle EMBOSS

Durée : 1 jour - intervenant : Olivier COLLIN

Objectifs, compétences développées

Acquérir l'autonomie nécessaire pour l'utilisation de cette suite logicielle pour l'analyse de données de séquences.

Programme

- Tour d'horizon et catégorisation des programmes Emboss.
- Modalités d'utilisation des programmes Emboss en ligne de commande.
- Emboss et les banques de séquences.
- Enchaînement de programmes et automatisation.

Recherche et découverte de motifs

Durée : 1 jour - intervenant : François COSTE

Objectifs, compétences développées

Acquérir et maîtriser les différentes méthodes utilisées en recherche et découverte de motifs dans les séquences biologiques.

Programme

- La découverte de motifs :
 - définition et représentation des motifs ;
 - les banques de motifs ;
 - présentation d'outils dédiés ;
 - mise en pratique.
- La recherche de motifs :
 - recherche de motifs simples ;
 - recherche de motifs complexes ;
 - recherche de sites de fixation de facteurs de transcription dans les zones promotrices ;
 - mise en pratique.

Introduction aux méthodes de phylogénie

Durée : 1 jour - intervenant : Erwan CORRE

Objectifs, compétences développées

L'exploitation des outils de phylogénie moléculaire repose sur la bonne compréhension des concepts sous-jacents.

L'objectif de cette formation est de proposer une approche générale des outils d'étude et de reconstruction phylogénétique. À l'issue de cette formation, les stagiaires auront acquis une connaissance des divers outils et méthodes utilisables et seront en mesure de mieux interpréter les résultats fournis par ces mêmes outils.

Programme

- Session théorique :
 - préambule historique et lexical ;
 - la phylogénie moléculaire ;
 - les méthodes d'analyse phylogénétiques :
 - méthodes de maximum de parcimonie,
 - les distances évolutives,
 - méthodes de distances,
 - méthodes de maximum de vraisemblance ;

• divers :

- robustesse des arbres phylogénétiques,
- attraction des longues branches

Session pratique :

- utilisation des outils d'analyse phylogénétique : phylowin, Mega, etc.

Introduction au langage PERL

Durée : 3 jours - intervenant : Fabrice LEGEAI

Objectifs, compétences développées

Acquérir les notions de base de PERL pour développer des outils nécessaires à l'analyse et à l'automatisation de l'analyse des données biologiques.

Programme

- Les bases de PERL.
- Les expressions régulières.
- Les références.
- Les fonctions.
- Bases de données et CGI.
- Les modules Perl.

Perl avancé: Perl Objet et BioPerl

Durée : 3 jours - intervenant : Fabrice LEGEAI

Objectifs, compétences développées

Maîtriser les notions avancées de PERL pour le développement d'applications en bioinformatique.

Programme

- Rappels de quelques notions indispensables: référence, hachage, modules.
- Notions de programmation orientée objet.
- PERL objet.
- BioPERL.

Introduction à R et Bioconductor

Durée : 1 jour - intervenante : Nolwenn LE MEUR

Objectifs, compétences développées

Acquérir l'autonomie nécessaire pour l'utilisation de cette suite logicielle pour l'analyse de données de séquences.

Programme

- Présentation du langage R et de ses fonctionnalités :
 - structures de données ;
 - manipulation des données.
- Présentation du projet Bioconductor :
 - packages pour les données à haut-débit ;
 - méta-données ou annotation des données génomiques ;
 - traitement des données génomiques (application pratique).

Introduction à Unix et SGE (Sun Grid Engine)

Durée : 1 jour - intervenant : Olivier COLLIN

Objectifs, compétences développées

Acquérir l'autonomie nécessaire pour évoluer au sein de l'environnement Unix et pouvoir utiliser les outils bio-informatiques disponible sous cet environnement. Pouvoir utiliser l'environnement de soumission Sun Grid Engine et développer des scripts de soumission.

Programme

- Linux :
 - présentation de l'environnement et de sa philosophie ;
 - quelques commandes de base.
- SGE :
 - fonctionnement de SGE ;
 - élaboration de scripts de soumission.

Introduction à Python en sciences biologiques

Durée : 2 jours - intervenante : Solenn STOECKEL

Objectifs, compétences développées

L'objectif de cette formation est d'introduire aux concepts de programmation de haut/bas niveau, orienté objet ou non, interprétés / compilés. Il s'agit également d'amener les stagiaires à comprendre quand, comment et pourquoi utiliser tel ou tel type de langages par notamment la comparaison caractéristique de certains d'entre eux ainsi que par la mise en pratique sur la base de cas concrets.

Programme

- Histoire du langage python, philosophie, portage, de son utilisation, concurrent (Ruby et consorts)
- Variables, types de données
- Boucles
- Conditions
- Fonctions
- Librairies
- Objets/classes : avantages, exemple de créations d'interfaces graphiques conviviales
- Gestion dynamique des erreurs (essais/exceptions)
- Gestion de données externalisées : fichiers, (bases de données ?)
- BioPython
- L'utilisation de C, R ou Matlab dans Python
- Analyse de code génétique et expressions régulières en python
- Programmation de calculs matriciels : utilisation de Numpy, Maths, Matplotlib
- Création de simulations stochastiques
- Création d'une interface graphique pour la gestion de l'un des problèmes rencontrés.

Perfectionnement R et Bioconductor (cours avancé)

Durée : 1 jour - intervenante : Nolwenn LE MEUR

Objectifs, compétences développées

Acquérir des méthodes pour intégrer au mieux méta-données et données expérimentales afin d'extraire du sens des données génomiques. Ce deuxième niveau de formation (avancée) s'adresse aux chercheurs et ingénieurs, biologistes et bio-informaticiens souhaitant avoir une approche intégrative et aller plus loin dans l'analyse de leurs données génomiques avec R.

Programme

- Gene set enrichment analysis (GSEA)
- Les graphes
- Méthodes d'apprentissages (clustering et autres méthodes classifications)

INFORMATIONS / CONTACT

DATES & TARIFS : nous consulter

LIEU DE LA FORMATION : institut de formation supérieure en informatique et communication (IFSIC),
campus de Beaulieu - 35000 Rennes

CHARGÉE DE MISSION : Anne BRIAND
02 23 23 39 50
sfc-irisa@univ-rennes1.fr

COORDINATION PÉDAGOGIQUE : Olivier COLLIN - ingénieur de recherche
CNRS / IRISA