

FICHE DESCRIPTIVE D'UN MODULE DE FORMATION

33-E223-03

OBI3 - INFORMATIQUE POUR LA BIOLOGIE

Bases Algorithmiques et Statistiques de l'Analyse des Séquences
Applications avec les logiciels disponibles sur le WEB

Atelier Approfondissement		Organisation : ED 223 LdV
Durée : 4,0 jours		Contact email : sec_ldv@snv.jussieu.fr
2 sessions/ an	12 places/session	Recommandé en <input checked="" type="checkbox"/> 1 ^{re} <input checked="" type="checkbox"/> 2 ^e <input checked="" type="checkbox"/> 3 ^e année

Objectifs

Montrer les méthodes et algorithmes utilisés en Génomique et analyse de séquence. Ces algorithmes seront étudiés de manière à mettre en évidence la pertinence du paramétrage des logiciels utilisés. Les stagiaires seront capables de mettre en œuvre de manière adéquate les méthodes bioinformatiques disponibles sur le WEB.

Contenu et méthode

- Les algorithmes de programmation dynamique : alignements 2 à 2 (NWS [align], Smith&Waterman [align], BestFit)
- Les algorithmes d'alignements multiples : alignement globaux (clustal-multalin-malign), alignements par blocs (Macaw - Dialign)
- Les scores d'alignements - Les matrice de similarités / matrice de scores, scores d'alignement pour les acides nucléiques et les protéines, pénalités de gap et d'extension de gap
- La validation statistique : distributions limites, p-value, statistiques de BLAST, statistiques sur les séquences, z-score, randomisation.
- Les banques de données de séquences - Principaux logiciels de requête : banques nucléiques EMBL, PDB,...banques protéiques SWISSPROT, PIR, PDB,...autres banques (de domaines PRODOM, banques de banques,...)
- Les algorithmes de recherche de similarités dans les banques (algorithmes de FASTA, BLAST, SSEARCH)
- Détermination des motifs, blocs, "profiles", matrice score/position, modèles de Markov cachés - recherche de motifs sur les séquences ou sur les banques - Application à la localisation de zones codantes.
- Stratégies d'analyse : recherche de similarité dans les banques/alignements multiples/retour sur les profiles

Responsable pédagogique / Formateurs

B. Billoud, E. Ollivier, J. Pothier , E. Rocha – Atelier de Bio-Informatique
M-O. Delorme, UMR 8116
C. Gerbaud, Laboratoire de Biologie et Génétique moléculaire
I. Gonçalves, Institut Jacques Monod
I. Lafontaine, Institut Pasteur

Pré-requis éventuels

Ce module est destiné à un public de biologistes (au sens large).
Les prérequis pour ce module sont donc par ordre d'importance:

- 1) être intéressé
- 2) avoir déjà pianoté sur un clavier d'ordinateur (même avec un doigt, mais c'est mieux avec 2 ! :-).

Pour en savoir plus (calendrier, places disponibles, etc.) : <http://www.snv.jussieu.fr/edv>

Date de mise à jour de la fiche (mois/année) : novembre 2006