

Séminaires de l'IDRIS

Les besoins de la communauté des sciences de la vie
en termes d'infrastructure informatique

Jeudi 14 novembre 2013 (10 h 30 – 12 h)

Jean-François Gibrat (INRA) et Christophe Blanchet (CNRS)

La communauté des sciences de la vie produit des masses toujours croissantes de données brutes hétérogènes. En particulier, les nouvelles technologies de séquençage (Next Generation Sequencing technologies) sont désormais capables de générer plusieurs téraoctets de données par expérience. Le rôle de la bioinformatique est de gérer ces grandes masses de données mais surtout de les analyser en utilisant une large palette de méthodes informatiques afin d'aider les biologistes à en extraire des connaissances biologiques. Le séminaire décrira brièvement l'organisation de la communauté bioinformatique française, en particulier l'Institut Français de Bioinformatique (IFB) qui est une nouvelle infrastructure de service en bioinformatique issue du programme « Investissements d'Avenir ». Une des missions de l'IFB est de mettre à disposition une infrastructure informatique dédiée à la gestion et l'analyse des données produites par les sciences du vivant et d'y déployer un Cloud académique. Le séminaire présentera les particularités des traitements bioinformatiques de ces données et les solutions qui sont envisagées dans le cadre de l'IFB pour y répondre.

Jean-François Gibrat est DR1 INRA dans l'unité Mathématique, Informatique et Génome du centre de recherche de Jouy-en-Josas. Il est également directeur de l'Institut Français de Bioinformatique et est chargé de la mise en place de cette nouvelle infrastructure nationale. Ses domaines de recherche portent sur l'analyse in silico des relations séquence — structure tridimensionnelle (3D) — fonction des protéines et sur l'analyse des données génomiques microbiennes. Il a ainsi été amené à utiliser différentes infrastructures informatiques, du Cray 1 du Centre de Calcul Vectoriel pour la Recherche fin des années 80 aux fermes de calcul déployées par les plates-formes de bioinformatique.

Christophe Blanchet est IR1 CNRS, responsable de la plateforme Infrastructure Distribuée pour la Biologie (IDB, idee-b.ibcp.fr) à l'Institut de Biologie et Chimie des Protéines (IBCP CNRS-FR3302). Il exerce une activité de recherche et veille technologique sur le calcul scientifique intensif, particulièrement sur les solutions de *cloud computing* pour la Biologie. Il a coordonné le projet RENABI GRISBI « Grille, Support pour la Bioinformatique » (www.grisbio.fr) de 2008 à 2011. Il est également en charge depuis 2012 de la cellule Infrastructure de l'Institut Français de Bioinformatique et à ce titre du déploiement d'un *cloud académique* pour la communauté nationale des sciences du vivant.

L'accès à ce séminaire est libre mais l'enregistrement est obligatoire à l'adresse www.idris.fr/seminaires