

A panel of European Public Galaxy Instances

Hans-Rudolf Hotz, Friedrich Miescher Institute for Biomedical Research, Maulbeerstrasse 66, 4058 Basel, Switzerland

Björn Grüning, Albert-Ludwigs-Universität Freiburg, Georges-Köhler-Allee 106, 79110 Freiburg, Germany

Alban Lermine, Institut Curie, INSERM U900, Mines ParisTech, Bioinformatics, Biostatistics, Epidemiology and computational Systems Biology of Cancer, 26 rue d'Ulm, 75248 Paris cedex 05, France

Valentina Boeva, Institut Curie, INSERM U900, Mines ParisTech, Bioinformatics, Biostatistics, Epidemiology and computational Systems Biology of Cancer, 26 rue d'Ulm, 75248 Paris cedex 05, France

Olivier Inizan, Institut National de la recherche agronomique, Unité de recherche en génomique info. RD 10 Rte de Saint Cyr, 78026 Versailles Cedex.

Christophe Blanchet, French Institute of Bioinformatics, IFB-core CNRS UMS 3601, Avenue de la Terrasse - Bât 21, 91190 Gif-sur-Yvette

Abstract

Galaxy is a web-based platform for data intensive biomedical research. Galaxy is available as a free public web server (<http://UseGalaxy.org>), and also as open source software that can be downloaded and customized for local needs. In this demonstration, we present three publicly accessible Galaxy servers: "deepTools", "Nebula" and "IFB's cloud".

Description

Galaxy is an open, web-based platform for accessible, reproducible, and transparent computational biomedical research. Users without programming experience can easily specify parameters and run tools and workflows. Galaxy captures information so that any user can repeat and understand a complete computational analysis. Users can share and publish analyses via the web and create interactive, web-based documents that describe a complete analysis. This is all available on Galaxy's public service web site usegalaxy.org.

In addition to using the public Galaxy server (<http://UseGalaxy.org>), you can also install your own instance of Galaxy or create a cloud-based instance of Galaxy. Another option is to use one of the ever-increasing number of public Galaxies hosted by other organizations. This presentation includes a demonstration of three specialised public Galaxies: "deepTools", "Nebula" and "URGI/IFB." An introduction to the Galaxy Project and deepTools will be covered in the first half hour, and Nebula and IFB's cloud in the second half hour.

deepTools (<http://deeptools.ie-freiburg.mpg.de/>) is a suite of user-friendly tools for the visualization, quality control and normalization of data from high-throughput DNA sequencing experiments. deepTools offers multiple methods for highly-customizable data visualization that immensely aid hypothesis generation and data interpretation. It also offers all the tools needed to create coverage files in standard bedGraph and bigWig file formats allowing various normalization procedures and comparisons between two files (for example, treatment and control).

Nebula (<http://nebula.curie.fr/>) is a web service which allows inexperienced users to perform a complete bioinformatics analysis of ChIP-seq data, including data for transcription factors and histone marks. The Nebula workflow includes tools for (1) read mapping, (2) peak calling and a module for immunoprecipitation quality control, (3) *de novo* motif discovery, (4) calculation of the density and the cumulative distribution of peak locations relative to gene transcription start sites, and (5) peak-to-gene and gene-to-peak annotation. Nebula generates the graphs and the enrichment statistics at each step of the process. During steps 4 and 5, Nebula optionally repeats the analysis on a control dataset and compares these results with those from the main dataset. Nebula can also incorporate gene expression (or gene modulation) data during these steps. In summary, Nebula is an innovative web service that provides an advanced ChIP-seq analysis pipeline providing ready-to-publish results.

The academic cloud infrastructure of the French Institute of Bioinformatics (IFB, <http://www.france-bioinformatique.fr>) is a community cloud devoted to life science. Its purpose is to provide scientists with a bioinformatics infrastructure to run the required tools close to public reference data sources. Biologists and bioinformaticians have to face the deluge of data and to perform nowadays intensive analyses with many bioinformatics tools among the arsenal of thousands ones available from the international community. Moreover, most of the time they need to combine multiple software packages to study their data with public pipelines or their own workflows. The Galaxy environment can provide scientists with such an integrated environment to perform their analyses. Hence, among the different bioinformatics cloud services provided by the IFB, one is a Galaxy appliance containing a predefined Galaxy portal ready to run by life science researchers on the IFB's cloud infrastructure.