Babelomics

Web Application Environment Introduction

Buenos Aires, October 2011

Javier Santoyo-Lopez
jsantoyo@cipf.es
http://bioinfo.cipf.es
Genomics Department
Centro de Investigacion Principe Felipe (CIPF)
(Valencia, Spain)
Motivation
pre-genomic analysis and data

Extract as much information as possible for one single data

Very easy data analysis of the results

Western blot
Southern blot
Northern blot

Linkage studies

FISH
Motivation
Huge amount of data

DNA sequencer

DNA microarray: *expression*, *genotyping*, exons, ChIP-on-chip, ...

2063 Microbial Genomes selected:
Complete - 774, Assembly - 596, Unfinished - 693

501 Eukaryotic Genome Sequencing Projects Selected:
Complete - 23, Assembly - 243, In Progress - 235

*Illumina HiSeq 2000* is the first commercially available sequencer to enable researchers to obtain ~30x coverage of two human genomes in a single run for under $10,000
Motivation

Extracting information from data

- Many new statistical considerations:
  - Data normalization
  - Multiple test correction
- New analysis are possible
  - Predictors
  - Clustering
- Computational problems:
  - Many gigas of data
  - How to process all this data
- Biological Databases
  - More than 1200 thousand

Data ≠ Information
Babelomics
The Web


- It's a merge and rewrite of Babelomics-3 and GEPAS-4:
  - Integration
  - Speed up
  - Development platform

- More than 10 publications in NAR in the last 5 years

- Babelomics was chosen by NAR editors as Featured Article in 2010
Why a Web application?

- Software paradigm change, from local desktop to web applications, some advantages:
  - **Multi-platform**: Windows, GNU/Linux, Mac
  - **Web storage**: 1GB of disk space for free
  - **Automatic updates**: nothing to install locally is required
  - **High CPU/RAM requirements**: many of the algorithms will take ages to run in a typical PC

- Requirements:
  - **Internet connection**: the best you can get
  - **Web browser (HTML5)**: Chrome, Firefox, Safari, Opera, Explorer?
Babelomics
Web usage

- Live demo:
  - Anonymous, Registration and Login
  - Reset and Change password
  - Status bar
  - Jobs and Data widgets
  - Data and Project management
  - Upload data and Tags
  - Online examples and help links
  - Contact us
Documentation

- Bioinformatics Group software site:
  - http://docs.bioinfo.cipf.es/

- Project site:
  - http://docs.bioinfo.cipf.es/projects/show/babelomics

- Tutorial site:
  - http://bioinfo.cipf.es/babelomicstutorial

- Google Groups:
  - http://groups.google.com/group/bioinfo-cipf