ALLBIO training workshop – Marseille, June 23-25, 2014

Analysing thousands of bacterial genomes:
gene annotation, metabolism, regulation

Practical info

Registration, schedule, details   http://www.allbioinformatics.eu/doku.php?id=public:valid-ulb

Dates	June 23-25, 2014
Location	Site de Luminy, Aix-Marseille Université (AMU), Marseille, France
Organizer	Jacques van Helden (Jacques.van-Helden@univ-amu.fr)
Participants	20
Cost	0€
Registration	before June 11, 2014

Funding

This training workshop is funded by the ALLBIO project (http://www.allbioinformatics.eu/), whose goal is to establish the link between bioinformatics and biologist communities who express uncovered needs for other fields than human genomics (microbial, plant, livestock).

Motivation

Sequencing technologies allow microbiologists to sequence, for a moderate cost, not only the genome of their favourite species, but a few dozens of related species, a collection of strains, and even to characterize inter-individual variations. The Ensembl Genomes resource contains over 10,000 completely sequenced genomes. The availability of such data requires new bioinformatics resources suited for annotating, querying and analysing multitudes of sequenced bacterial genomes in parallel.

This training workshop will show how to combine several specialized bioinformatics resources to extract information about bacterial genes and their annotations (Ensembl Genomes), metabolism (MICROSCOPE), and regulation (RSAT). The course will be oriented towards comparative genomics, and give a perspective about advances that can be expected from the massive reduction of sequencing costs resulting from Next Generation Sequencing.

The course will mainly be based on user-friendly Web interfaces. It will also include a basic introduction about the programmatic access to bioinformatics resources, showing by simple and well-documented examples how to automatize the collection of data, and extract information from multiple data sets.

Target audience

This course is addressed to microbiologists confronted to the need to use bioinformatics resources to analyse multiple genomes.

Prerequisites

The course does not require any computing skills. A basic knowledge of the Linux environment is welcome, but not required.

Contents

This training course will include

- Short talks illustrating the functionalities offered by bioinformatics resources suited for the analysis of multiple bacterial genomes.
- Practicals based on concrete study cases;
  - collecting multiple genomes from bacteria;
  - comparing annotations between different species;
  - extracting metabolic information for selected genomes;
  - metabolic projection: coverage of reference pathways by enzymes found in a genome;
  - phylogenetic profiling: detecting genes co-occurrences across genomes;
  - phylogenetic footprints: detection of conserved cis-regulatory elements.
- First steps in the programmatic use of some tools:
  - extracting data for multiple genomes;
  - combining data from diverse sources.

Computer environment

The training room is equipped with Linux computers, but participants are welcome to bring their own laptop if they wish to work in their familiar environment.

Schedule

Date	From	To	Duration	Topics	Teachers
Mon 23/06	08:30	09:00	00:30	Welcome and computer configuration	
	09:00	10:00	01:00	Introduction: comparative genomics, data, concepts and perspectives	Jacques van Helden, Claudine Médigue
	10:00	*10:30*	*00:30*	*Coffee break*	
	11:00	12:30	01:30	Genome annotations (EnsemblGenomes)	Dan Staines
	12:30	*14:00*	*01:30*	*Lunch*	
	14:00	15:30	01:30	Genome annotations (EnsemblGenomes)	Dan Staines
	15:30	*16:00*	*00:30*	*Tea break*	
	16:00	17:30	01:30	Metabolism (MICROSCOPE)	Claudine Médigue, François Le Fevre, Eugeni Belda
	17:30			*End of first day*	
Tue 24/06	09:00	10:30	01:30	Metabolism (MICROSCOPE)	Claudine Médigue, François Le Fevre, Eugeni Belda
	10:30	*11:00*	*00:30*	*Coffee break*	
	11:00	12:30	01:30	Regulation (RSAT)	Jacques van Helden, Denis Puthier
	12:30	*14:00*	*01:30*	*Lunch*	
	14:00	15:30	01:30	Regulation (RSAT)	Jacques van Helden, Denis Puthier
	15:30	*16:00*	*00:30*	*Tea break*	
	16:00	17:30	01:30	Programmatic access to the Web resources	All teachers
	19:30	*22:00*	*02:30*	*Dinner*	
Wed 25/06	09:00	10:30	01:30	Combining resources to analyse multiple genomes	All teachers
	10:30	*11:00*	*00:30*	*Coffee break*	
	11:00	12:30	01:30	Combining resources to analyse multiple genomes	All teachers
	12:30	*14:00*	*01:30*	*Lunch*	
	14:00			*End of the training*	

Teachers

People	Institution	Resources
Dan Staines	European Bioinformatics Institute (EBI), UK	Ensembl Genomes http://ensemblgenomes.org/
François LeFevre, Eugeni Belda, Claudine Médigue	Genoscope (CEA), France	MICROSCOPE http://www.genoscope.cns.fr/agc/microscope/
Jacques van Helden Denis Puthier	Aix-Marseille Université (AMU), France	Regulatory Sequence Analysis Tools (RSAT) http://www.rsat.eu/