



Short Course

Hands-on Bioinformatics and Functional Genomic Data Mining

University of Palermo

Environmental & hEredity Evolution Consulting (e³C) srls

Held by:

Federico Martinelli – Assistant Professor in Plant Genetics, Department of Agricultural and Forest Sciences, University of Palermo

Bipin Balan – Bioinformatician, PhD Student, University of Palermo

Course Language: English/Italian

Focus: Bioinformatics and functional genomic data mining

Target audience: PhD, graduate students, post-docs and all other staff researchers willing to perform functional genomic studies

Educational certification/accreditation status: Certificate of participation provided by Relators

Needed equipment: Audience should bring their own laptops

Key features and objective of the course:

The course aims at providing knowledge and practical experience on genomic data analysis and functional mining. The course is addressed for students or researchers who are willing to use and produce NGS datasets, but have scarce experience in their functional analysis. The goal of the course is to teach audience how to exploit RNA-seq, proteomic and metabolomic data for research purposes. At the end of the course, the audience is expected to be able to design, develop and produce high-impact factor publications dealing with functional genomic studies.

The course is organized in three days.

First day: In the morning, students will be introduced to bioinformatics and they will learn basic informations on linux commands. In afternoon sessions they will learn perl programming and how to use files formats for data analysis (Fastq, BAM & SAM, bed file format).

Second day: In the morning, RNA-Seq analysis will be introduced. Audience will learn how:

- 1) download reference genome and index creation
- 2) perform important operations such as fastq quality check, base and adapter trimming and contamination removal.

In the afternoon, some data-mining tools frequently used in functional genomics will be described. Students will learn how:

- visualize the data using Mapman software
- perform gene set and pathway enrichment analysis (Pageman, David software).

Third day: Methods for proteomic and metabolomic data analysis will be explained, qRT-PCR validation of transcriptomic data will be described. Instructions and advices for the production of high impact functional genomic publications will be provided.

You will learn how:

- Perform quality control of raw reads
- Align reads to a reference genome
- Convert and process file formats

- Perform quantitative expression analysis
- Analyze functional genomic data (transcriptomics, proteomics and metabolomics)
- Extract important biological informations from large-scale datasets

Registration costs: 150 Euro/person (included slide materials of course lessons)

Program

Day 1

Date: 15/06/2017

Time: 9 AM to 6PM

Session1: 9AM – 11AM (Linux Commands)

- ✓ Basic Linux Command
- ✓ Vim editor

Tea Break: 15 min

Session2: 11.15AM – 1.15PM (Perl Programming for NGS data analysis)

- ✓ Basic Perl programming

Lunch Break: 1 hour

Session3: 2.15PM – 4.15PM (Perl Programming for NGS data analysis)

- ✓ Basic Perl programming

Tea Break: 15 min

Session4: 4.30PM – 6.30PM (NGS File formats)

- ✓ Fastq file format

- ✓ BAM & SAM
- ✓ Bed file format
- ✓ Question Answer session

Day 2

Date: 16/06/2017

Time: 9 AM to 6PM

Session1: 9AM – 11AM (RNA-Seq Prerequisite Steps)

- ✓ Introduction to RNA-Seq Analysis
- ✓ Downloading reference genome & index creation
- ✓ Fastq Quality Check, Base Trim, Adapter Trim
- ✓ Contamination removal

Tea Break: 15 min

Session2: 11.15AM – 1.15PM (Reference RNA-Seq Analysis)

- ✓ Alignment with reference genome
- ✓ Expression estimation & Comparison
- ✓ Cumberbund R package.
- ✓ Question Answer session

Lunch Break: 1 hour

Session3: 2.15PM – 4.15PM (Functional Analysis and Data Mining)

- ✓ Principal Component Analysis
- ✓ Visualization of “omic” data
- ✓ Gene Set Enrichment Analysis
- ✓ *Tea Break: 15 min*

Session3: 4.30PM – 6.30PM (Functional Analysis and Data Mining)

- ✓ Pathway Enrichment Analysis
- ✓ Networks Analysis
- ✓ Question-Answer session

Day 3

Date: 17/06/2017

Time: 9 AM to 12 AM

Session1: 9AM – 12AM (Proteomic and metabolomic data mining and validation)

- ✓ Proteomic and metabolomic data analysis
- ✓ Validation of transcriptomic data through qRT-PCR
- ✓ Structure and write publications on functional genomics

-----**END OF COURSE**-----