

BIOINFORMATICS FOR FOREST PATHOLOGISTS

DIAROD Training School. 21 to 24th July 2015
Palencia (Spain)

Tuesday 21 July

9:00-11:00. Issue 1: **NCBI: Introduction to BLAST (I)** - theory

Trainer: Irene Barnes

- What is BLAST and why use it?
- Basics of using NCBI BLAST
 - Local alignment vs global alignment
 - How to search for a record
 - How to read a GenBank flatline file
 - Using BLASTn to search GenBank
 - Different input formats
 - Interpreting the output
 - Refining searches
 - Exporting data in different file formats
- Primer BLAST
- MOLE-BLAST

11:00-11:30 Coffee Break

11:30-13:30. Issue 1: **NCBI: Introduction to BLAST (2)** – practical

Trainer: Irene Barnes

- To know how to search for records at NCBI, both using search terms or identifiers and GQuery, or using a nucleotide sequence and BLAST.
- To be able to use nucleotide BLAST (Blastn) to search GenBank, and be able to interpret the output.
- To be able to use primer BLAST to design primers
- To be able construct simple phylogenetic trees using MOLE-BLAST

13:30-14:30 Lunch

14:30-17:00. Issue 2: **NCBI: Advanced BLAST (3)** (15:30-16:00 Coffee Break)

Trainer: Jarkko Hantula

- To understand the general concept underlying substitution matrices used for scoring protein similarity.
- To know which type of matrix to use to identify more distantly related sequences or those that are more closely related.
- To be able to interpret a dot matrix alignment.
- To be familiar with the theory behind the working of the BLAST algorithm.
- To know which flavour of BLAST (blastn, blastp, tblastx etc.) to use when – the key is to know whether your query sequence is nucleotide or protein, and whether your database is nucleotide or protein.
- To be able to use the appropriate GenBank database.

Wednesday 22 July

9:00-10:00. Issue 3: **Multiple Sequence Alignment** (MAFFT (online), MEGA)

Trainer: Irene Barnes

- To know the different alignment options available: MAFFT, CLUSTAL X, MUSCLE
- To know how to use the online version of MAFFT for sequence alignment
- Know how to import your data into MEGA
- Know how to align sequences in MEGA using Clustal W and Muscle

10:00 -17:00. Issue 4: **Phylogenetic Analyses** (Model-test, MEGA)

(11:00-11:30 & 15:30-16:00 Coffee Break; 13:30-14:30 Lunch)

Trainer: Jarkko Hantula

- To know the terminology associated with phylogenetic trees and be able to identify the most recent common ancestor of any two terminal nodes (taxa) on the tree.
- To know the fundamental elements and the terminology of phylogenetics.
- To be able to identify the root of a tree and to know the difference between rooted and unrooted trees.
- To understand distance and character-based phylogenetic methods.
- To be acquainted with the various substitution models;
- To be familiar with bootstrapping and with a bootstrap score on a node
- To be able practically to construct neighbour-joining and maximum likelihood trees using MEGA.

Trainer: Piotr Boroń

- To be able to practically use jModeltest
- To know the principles of Bayesian analysis
- To be able to practically set up the phylogenetic analysis in MrBayes
- Exporting of MrBayes results to MEGA tree drawing tool

Thursday 23 July

9:00-17:00. Issue 5: **Next Generation Sequencing applications: METAGENOMICS**

(11:00-11:30 & 15:30-16:00 Coffee Break; 13:30-14:30 Lunch)

Trainer: Juan Imperial & Beatriz Jorrín

- To understand some of the current technologies for performing next generation sequencing.
- To be acquainted with the concept of metagenomes and understand how these can be used to assess species diversity of a community.
- To be familiar with how NGS assemblies are generated.
- To know how to explore RNA-Seq data on a Genome Browser and understand what an RNA-Seq track can be used for in terms of ascertaining the expression level.

Friday 24 July

9:00-13:30. Issue 6: **Protein Analyses**

(11:00-11:30 Coffee Break)

Trainer: Marta Vasconcelos

- To be familiar with the use of Protein and Specialized Sequence Databases.
- To be able to generate a protein sequences using a nucleotide sequence and choosing the correct open reading frame (Expasy)
- To know how to work with a Single Protein Sequence and with Protein 3-D Structures.
- To be familiar with databases that can be used to predict certain properties about a protein and which can be useful for its empirical investigation
- How to use ligand information and bioinformatics to identify molecules with potential fungicidal properties