

GRAINGER BIOINFORMATICS CENTER WORKSHOP: ANALYSES OF PHYLOGENETIC AND PHYLOGENOMIC DATA SETS OF LICHENS

7 – 9 June 2022

The Workshop will take place at the Faculty of Pharmacy, Complutense University of Madrid,
Spain

Phylogenetic analyses are increasingly based upon massive genomic data rather than single marker genes. While PCR and Sanger sequencing allowed to study the relationship of organisms based on one to multiple marker regions, new technologies such as high-throughput sequencing provides opportunities to sequence and analyze whole genomes. Therefore, computational work becomes increasingly important for modern biologist to analyze massive data sets generated by high-throughput sequencing techniques. Many biologists are now challenged to work with gigabytes of sequencing data generated in only hours by parallel sequencing machines. These sequencing data are usually produced in the form of large text files, which makes the Unix operating system (including Linux) particularly suited to processing such files - especially when operated from the command line. Hence, the aim of this workshop is to introduce phylogenetic and phylogenomic concepts and how to apply these analyses by using the Linux command line.

During the first sessions of the course, you'll learn the history of phylogenetic methods and about traditional molecular marker regions which were used for decades to estimate the relationships of organisms. Then you will learn basic but powerful Linux commands to manage your folder structure, handle large files, and install and execute programs. After you feel comfortable in the Linux command line environment, you will apply your new knowledge to: build alignments, run complex phylogenetic methods, assemble high throughput sequencing data into continuous genomes, verify the integrity of these genomes by sequence mapping, use search methods to identify gene regions, and use these regions for phylogenetic reconstruction – all on the Linux operating system. We'll finish this course with discussing advantages and problems of this new era of sequencing technologies.

You'll learn in this workshop how to:

- ⇒ Operate Linux from the command line
- ⇒ Install and execute Linux programs
- ⇒ Work with phylogenetic data set (build alignments with clustalW, learn how to convert files with Mesquite)
- ⇒ Use aligned sequences for phylogenetic reconstruction (using RAxML, Phylip, and MrBayes)

- ⇒ Work with the structure of large high-throughput sequencing data formats (e.g. fastq, sam, vcf)
- ⇒ Assembly and analyze genome sequencing (SPAdes, blast)
- ⇒ Quality control genome assemblies by read mapping (Bowtie)
- ⇒ Reduced genome representation methods for phylogenomic such as target enrichment analysis (HybPiper and UnFATE)

This workshop is open to master's and Doctoral students, post-Doctoral fellows, research scientists and lecturers, both directly related or not to research in lichens. **OPEN to ALL: NO REGISTRATION FEES. The number of participants is limited to 15**, so that personal attention can be provided throughout the training. **REGISTRATION** would be appreciated no later than [31st May 2022](#) to the Secretary of our department at: dp426@ucm.es

The Faculty of Pharmacy, UCM will provide the venue and informatics facilities. The Field Museum Chicago will provide the literature and bioinformatics' scripts and programs during the workshop.

Venue: Informatics room 219 at the Faculty of Pharmacy, Complutense University of Madrid, Plaza Ramon y Cajal, s/n, 28040, MADRID, SPAIN.

Organizers:

PRADEEP K. DIVAKAR, MARIA JOSEFA HERNAIZ GOMEZ-DEGANO, AND THE SYSTEMOL TEAM

Department of Pharmacology, Pharmacognosy and Botany, & Faculty of Pharmacy, UCM.

WORKSHOP SCHEDULE

Grainger Bioinformatics Center Workshop: Analyses of Phylogenetic and Phylogenomic Data Sets of Lichens

7 – 9 June 2022

Venue: Informatics room 219 at the Faculty of Pharmacy, Complutense University of Madrid, Plaza Ramon y Cajal, s/n, 28040, MADRID, SPAIN.

Day 1, June 7th, 2022 (Tuesday) [chair: Pradeep K. Divakar]	
9:15 – 9:30 h	Opening Remarks
9:30 – 10:15 h: Felix Grewe	History of Sequencing
10:15 – 11:00 h: Thorsten Lumbsch	Molecular Markers
11:00 – 11:30 h	Coffee Break
11:30 – 13:00 h: Felix Grewe	Linux Command Line 1
13:00 – 14:00 h	Lunch Break
14:00 – 15:30 h: Felix Grewe	Linux Command Line 2
15:30 – 16:00 h	Coffee Break
16:00 – 16:45 h: Thorsten Lumbsch	Alignments
Day 2, June 8th, 2022 (Wednesday) [chair: David Pizarro]	
9:30 – 11:00 h: Felix Grewe	Check on homework 1 and 2
11:00 – 11:30 h	Coffee Break
11:30 – 12:15 h: Thorsten Lumbsch	Maximum Parsimony
12:15 – 13:00 h: Thorsten Lumbsch	Maximum Likelihood
13:00 – 14:00 h	Lunch Break
14:00 – 15:30 h: Felix Grewe	Linux Command Line 3
15:30 – 16:00 h	Coffee Break
16:00 – 16:45 h: Felix Grewe	Assemblies
Day 3, June 9th, 2022 (Thursday) [chair: Leopoldo G. Sancho]	
9:30 – 11:00 h: Felix Grewe	Check on homework 3 and 4
11:00 – 11:30 h	Coffee Break
11:30 – 13:00 h: Felix Grewe	Target Enrichment Analysis
13:00 – 14:00 h	Lunch Break
14:00 – 14:45 h: Thorsten Lumbsch	Issues with Phylogenies
14:45 – 15:30 h: Thorsten Lumbsch	Species Delimitation
15:30 – 16:00 h	Coffee Break
16:00	Closing Remarks