

Computational Biology at the Field Museum Workshop - CBFM

Computational work becomes increasingly important for modern biologists to analyze massive data sets generated by high-throughput sequencing techniques. Biologists who never thought they would need computational skills are now finding that using an Excel spreadsheet is simply not enough to work with their gigabytes of sequencing data. These sequencing data are usually produced in the form of large text files, which make 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 computational biology work using Linux to anyone who has never previously worked with the Linux command line.

During the first weeks of the course, 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: 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 a basic introduction to concepts of the programming language python which will allow you to script your personal bioinformatics routines.

You'll learn how to:

- Operate Linux from the command line
- Install and execute Linux programs
- The concepts of Git version control
- Work with the structure of large high-throughput sequencing data formats (e.g. fastq, sam, vcf)
- Assemble and analyze a genome sequence (using SPAdes and/or velvet)
- Verify genome integrity by sequence mapping (using bowtie and samtools)
- Identify gene regions (using blast)
- Use the assembled sequences for phylogenetic reconstruction (using RAxML and/or PhyML)
- Basic concepts of programming with python

What you need to know about Linux:

- Nothing, absolute beginners are more than welcome (If you already know your way around the Linux command line or the python programming language, there's no need to attend this class)

What you need to bring and prepare:

- Please bring a personal laptop with the *CBFM Workshop Appliance* installed, since this class will be more workshop than lecture
- Installation instruction of the *CBFM Workshop Appliance* and course material will be made available at <http://cbfm.github.io/website/>.

The workshop will be instructed by Felix Grewe, Lyndon Coghill, and Rick Ree. If you're interested in joining this workshop or have any questions, email fgrewe@fieldmuseum.org or stop by in my office.

The workshop is scheduled for April 27, 29, May 1, 4, 6, 13, 15, 18, 20 from 1-2pm in the Zoology classroom.