Genomics and Clinical Microbiology 2024 Software Installation

Table of Contents

Introduction to Software Installation

This guide provides extensive instructions for software installation on Ubuntu/Windows/MacOS systems, applicable to both virtual machine (VM) and host machine setups, as taught and utilized in the Genomics and Clinical Microbiology 2024 course. For VM installation specifics, please consult the supplementary VM installation guide attached separately. Additionally, valuable troubleshooting links are provided below for additional assistance.

Note: The Virtual Machine file (.vdi) contains all the software installed for the course.

Software list according to course modules

Module	Software	Summary	Version	Website
	Inkscape	Inkscape is a Free and open source vector graphics editor for GNU/Linux, Windows and macOS.	0.92.4	https://inkscape.org/
Constructing phylogenetic trees	MEGA	Sophisticated and user-friendly software suite for analyzing DNA and protein sequence data from species and populations.	11.0.3	https://www.megasoftware.net/

Constructing phylogenetic trees	Splits Tree	SplitsTree is a widely used application for computing unrooted phylogenetic networks from molecular sequence data.	4.18.3	**SplitsTree | University of Tübingen**
Introduction to the Linux command line and BLAST	BLAST+	The BLAST+ applications can write the query, database, and command-line options for a BLAST search into a "strategy" file.	2.15.0	https://blast.ncbi.nlm.nih.gov/doc/blast-help/downloadblastdata.html
Genome assembly	SPAdes	SPAdes Genome Assembler is an open source tool for de novo sequencing.	3.15.4	https://github.com/ablab/spades
Genome assembly	QUAST	QUAST can evaluate assemblies both with a reference genome, as well as without a reference	5.2.0	https://github.com/ablab/quast
Genome assembly	FastQC	FastQC is used to quality control checks on raw sequence data coming from high throughput sequencing pipelines.	0.12.1	https://github.com/s-andrews/FastQC

Note: The versions might differ as new releases are enrolled by the software company.

Installation of Inkscape

Download Inkscape:

Visit the Inkscape website (https://inkscape.org/release/inkscape-1.3.2/) and download the Windows/MacOS/Ubuntu version of Inkscape.

In Windows and MacOS save the downloaded file and use on-screen instructions to install Inkscape.

Installing Inkscape on Ubuntu:

1. Open the Terminal on your Ubuntu system.
2. Add Inkscape Repository: Run the following commands to add the Inkscape repository and update the package lists:

```bash
sudo add-apt-repository ppa:inkscape.dev/stable
sudo apt-get update
```

3. Use the package manager to install Inkscape:

```bash
sudo apt-get install inkscape
```

4. Once the installation is complete, you can launch Inkscape from the application menu or by running Inkscape in the Terminal.

Installation of MEGA

Download MEGA:

Visit the official MEGA website (https://mega.io/download) and download the Windows/MacOS version of MEGA Sync Client.

In Windows and MacOS save the downloaded file and use on-screen instructions to install MEGA.

Installing MEGA on Ubuntu:

1. Open the Terminal on your Ubuntu system.
2. Add MEGA Repository: Run the following commands to add the MEGA repository and update the package lists:

```bash
sudo add-apt-repository ppa:megasync/megasync
sudo apt-get update
```

3. Use the package manager to install MEGA:

```bash
sudo apt-get install megasync
```

4. Once the installation is complete, you can launch MEGA from the application menu or by running 'megasync' in the Terminal.

Installation of SplitsTree

Download SplitsTree:

Visit the SplitsTree website (https://www.splitstree.org/) and download the Windows/MacOS/Ubuntu version of SplitsTree.

In Windows and MacOS save the downloaded file and use on-screen instructions to install SplitsTree.

Installing SplitsTree on Ubuntu:

1. Open the Terminal on your Ubuntu system.
2. Download SplitsTree:

```bash
sudo apt-get update
sudo apt-get install splitstree
```

3. Once the installation is complete, you can run SplitsTree by entering the following command in the Terminal:

```bash
splitstree
```

Installation of BLAST+, SPAdes, QUAST and FASTQC (with CONDA)

Conda Overview:

Conda is an open-source package management and environment management system that runs on Windows, macOS, and Linux. It simplifies the installation and management of software packages, ensuring dependencies are correctly handled.

Installation Script:

Ensure you have Conda installed on your system before running the script.

1. Create a Conda Environment (Optional): If you prefer to create a separate environment for these tools, you can do so by running:

```bash
conda create -n bioinformatics
conda activate bioinformatics
```

2. Add Conda Channels: Add necessary channels for accessing bioinformatics packages:

```bash
conda config --add channels conda-forge
conda config --add channels bioconda
```

3. Installations of software and verify:

```
# Install BLAST+, SPAdes, QUAST, and FASTQC
conda install blast spades quast fastqc

# Verify installations
echo "BLAST+ version: $(blastn -version)"
echo "SPAdes version: $(spades.py --version)"
echo "QUAST version: $(quast.py --version)"
echo "FASTQC version: $(fastqc --version)"
```

Note: In order to install individual software, use the command line command “conda install softwareName”. For example: conda install spades.

4. If you created a separate environment, deactivate it:

```bash
conda deactivate
```

Note:

- Conda environments provide isolation for different sets of tools, enhancing reproducibility.
- You can customize the script based on your preferences, such as creating a dedicated environment or modifying the channels. Activate the Conda base environment (conda activate base) if you choose not to create a separate environment.
- By using Conda, this script streamlines the installation process, ensuring that the specified bioinformatics tools and their dependencies are correctly configured on your system.
- Save this script to a file (e.g., install_bioinformatics.sh), make it executable (chmod +x install_bioinformatics.sh), and run it in your terminal (./install_bioinformatics.sh). Adjust the script as needed for your specific requirements.

Additional Resources and Troubleshooting

- Ubuntu Documentation: https://help.ubuntu.com/
- Ubuntu Community Support: https://ubuntu.com/support/community-support
- Conda Installation: https://conda.io/projects/conda/en/latest/user-guide/install/index.html
- Conda Documentation: https://conda.io/projects/conda/en/latest/user-guide/getting-started.html#managing-python
- Stack overflow: https://stackoverflow.com/ (Public Q&A platform for debugging)
- Bioinformatics (BioStars) Forum: https://www.biostars.org/t/Forum/ (General Bioinformatics queries)