

How to install programs in Linux



Most bioinformatics programs have been designed to run in Unix machines.



Or similar

Or connect to a server that has unix



<https://www.putty.org/>



With over 6 million users, the open source [Anaconda Distribution](#) is the fastest and easiest way to do Python and R data science and machine learning on Linux, Windows, and Mac OS X. It's the industry standard for developing, testing, and training on a single machine.

Beyond the propaganda, ANACONDA contains many of the bioinformatics programs and libraries.

Advantages:

- Works in unix, MacOS and windows: though it is not able to install everything in all the platforms.
- Generally it installs all needed dependencies
- It can create different environments for different programs that you can activate only when you need a given program
- Does not need superuser permissions

Disadvantages:

- If you're not careful it can mess up your normal computer configuration

Anaconda 2018.12 For Linux Installer

Python 3.7 version *

 Download

[64-Bit \(x86\) Installer \(652.5 MB\)](#) ⓘ
[64-Bit \(Power8 and Power9\) Installer \(313.6 MB\)](#)
[32-Bit Installer \(542.7 MB\)](#)

Python 2.7 version *

 Download

[64-Bit \(x86\) Installer \(628.2 MB\)](#) ⓘ
[64-Bit \(Power8 and Power9\) Installer \(289.7 MB\)](#)
[32-Bit Installer \(518.6 MB\)](#)

After downloading the distribution you want to, you just need to execute the file and it will do the installation process.

```
bash Anaconda-latest-Linux-x86_64.sh
```

In order to use anaconda, you need to activate it by exporting the folder where anaconda was installed into the path.

Warning: during the last steps of the installation it will ask you if you want to add the export path to your .bashrc

Once anaconda has been installed in your computer, you can simply install programs by printing:

```
conda install PACKAGENAME
```

By default, conda will install the latest version, but if, after a while you want to update it, you will need to run:

```
conda update PACKAGENAME
```

You may also want to install a specific version of a program:

```
conda install scipy=0.15.0
```

You also will very likely install programs that are not directly part of Conda but from other, associated channels. To install those packages you will need to tell conda from which channel it can get the package:

```
conda install -c bioconda raxml
```

To search which packages can be installed you can use:

bioconda



conda search scipy

Or just search in google if your package is available.

Organization created on Sep 11, 2015

	biobb_md	public	Biobb_md is the Biobb module collection to perform molecular dynamics simulations.	2019-01-27
	epa-ng	public	Massively parallel phylogenetic placement of genetic sequences	2019-01-27
	ataqv	public	ataqv is a toolkit for measuring and comparing ATAC-seq results. It was written to help understand how well ATAC-seq assays have worked, and to make it easier to spot differences that might be caused by library prep or sequencing.	2019-01-27
	bioconductor-rhdf5	public	This package provides an interface between HDF5 and R. HDF5's main features are the ability to store and access very large and/or complex datasets and a wide variety of metadata on mass storage (disk) through a completely portable file format. The rhdf5 package is thus suited for the exchange of large and/or complex datasets between R and other software package, and for letting R applications work on datasets that are larger than the available RAM.	2019-01-27
	r-msm	public	Functions for fitting continuous-time Markov and hidden Markov multi-state models to longitudinal data. Designed for processes observed at arbitrary times in continuous time (panel data) but some other observation schemes are supported. Both Markov transition rates and the hidden Markov output process can be modelled in terms of covariates, which may be constant or piecewise-constant in time.	2019-01-27
	r-phylobase	public	Provides a base S4 class for comparative methods, incorporating one or more trees and trait data.	2019-01-27
	r-polysat	public	A collection of tools to handle microsatellite data of any ploidy (and samples of mixed ploidy) where allele copy number is not known in partially heterozygous genotypes.	2019-01-27
	r-rtfbs	public	Identifies and scores possible Transcription Factor Binding Sites and allows for FDR analysis and pruning. It supports splitting of sequences based on size or a specified GFF, grouping by G+C content, and specification of Markov model order. The heavy lifting is done in C while all results are made available via R.	2019-01-27

What if it's not in conda?

Check out your unix repositories just in case it may be there.

Advantage: it's very easy to install

Disadvantage: it may be an old version. You need superuser permissions to install.

What if it's not in conda nor in my unix repository?

You will have to search for the website where it is published and try and follow the instructions of the authors to install the software.

Advantages: you'll always have the latest improvements

Disadvantages: a lot of times it can take some effort to install all the dependencies and some programs may not explain exactly how to install. Also you may not have superuser permissions which will limit your ability to install things.

Just a few words on github:

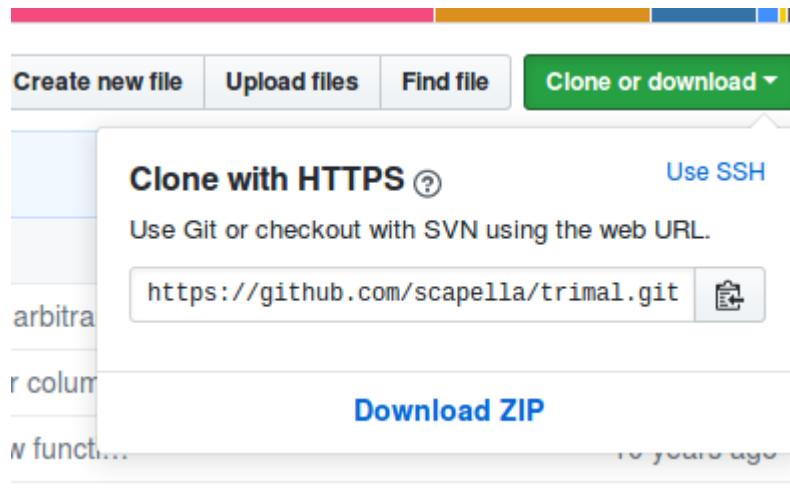
It has become one of the main program repositories because it allows for control version. So, a lot of programs not found in conda may be found there.

The screenshot shows the GitHub repository page for 'scapella / trimAI'. At the top, there are buttons for 'Watch' (10), 'Star' (53), 'Fork' (18), and navigation links for 'Code', 'Issues 29', 'Pull requests 0', 'Projects 0', and 'Insights'. Below this, a description states: 'A tool for automated alignment trimming in large-scale phylogenetic analyses. Development version: 1.4 <http://trimal.cgenomics.org>'. The repository has 124 commits, 1 branch, 1 release, and 2 contributors. A green 'Clone or download' button is visible. The commit history lists several commits from 'scapella', including adding a Python script for random alignment, fixing bugs, and updating the CHANGELOG and LICENSE files. The README section contains instructions for basic installation, which involve running 'make' in the 'source' directory. The page also includes a 'multiple-sequence-alignment', 'trimming', and 'bioinformatics-tool' tag cloud.

This is for instance the github page for trimAl.

In order to download it and install it you can press on the green button.

It will give you the option to either download the repository as a zipped folder or give you an address to use the git application.



If you have git installed you can simply:

```
git clone https://github.com/scapella/trimal.git
```

```
Cloning into 'trimal'...
remote: Enumerating objects: 836, done.
remote: Total 836 (delta 0), reused 0 (delta 0), pack-reused 836
Receiving objects: 100% (836/836), 8.48 MiB | 1.15 MiB/s, done.
Resolving deltas: 100% (566/566), done.
Checking connectivity... done.
mmarcet@phobos:~/software$ ls
trimal
```

Once you have the folder downloaded you will have to follow the instructions of the author to install the program

