

# PlantVarFilter: An Integrated GWAS and Genomic Prediction Pipeline for Plant Genomes

## Documentation:

How to run it on Linux with MiniFrog and Mamba

Developed by

Ahmed Yassin, and Dr Falak Sher Khan - postdoc scientist

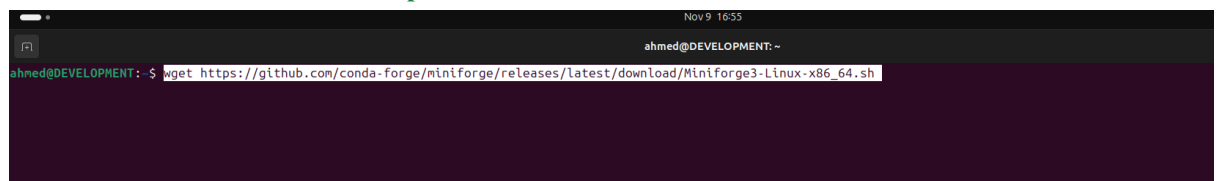
## Installing steps in Ubuntu:

### Uses the Commands:

Open the Ubuntu terminal and use the commands one by one:

1. `sudo apt update && sudo apt upgrade -y`
2. Install the minifrog version from conda by these commands: pull the conda from the GitHub repository “`wget` [https://github.com/conda-forge/miniforge/releases/latest/download/Miniforge3-Linux-x86\\_64.sh](https://github.com/conda-forge/miniforge/releases/latest/download/Miniforge3-Linux-x86_64.sh)”

Put the link in the terminal and press Enter, and wait for the installation.

A screenshot of a terminal window with a dark background. The title bar at the top shows "Nov 9 16:35" and "ahmed@DEVELOPMENT: ~". The terminal prompt is "ahmed@DEVELOPMENT:~\$". The command being entered is "wget https://github.com/conda-forge/miniforge/releases/latest/download/Miniforge3-Linux-x86\_64.sh". The command is highlighted with a light blue selection box.

Now you save a minifrog on your device

[illegible]

**Now we are installing the minifro from this command:**

## “bash Miniforge3-Linux-x86\_64.sh”

**Then press “Enter,” then Yes to create licenses and Enter to install**

the GNU General Public License and can be found in `pkgs/<pkg-name>/info/licenses` folders.

Miniforge install comes with a bootstrapping executable that is used when installing miniforge and is deleted after miniforge is installed. The bootstrapping executable uses micromamba, cll11, cp-file-system, curl, c-ares, krb5, libarchive, libev, lz4, nghttp2, openssl, libz, nlohmann-json, repro and zstd which are licensed under BSD-3-Clause, MIT and OpenSSL licenses. Licenses and copyright notices of these projects can be found at the following URL.

<https://github.com/conda-forge/micromamba-feedstock/tree/master/recipe>.

Copyright (c) 2019-2022, conda-forge  
All rights reserved.

Redistribution and use in source and binary forms, with or without modification, are permitted provided that the following conditions are met:

1. Redistributions of source code must retain the above copyright notice, this list of conditions and the following disclaimer.

2. Redistributions in binary form must reproduce the above copyright notice, this list of conditions and the following disclaimer in the documentation and/or other materials provided with the distribution.

3. Neither the name of the copyright holder nor the names of its contributors may be used to endorse or promote products derived from this software without specific prior written permission.

THIS SOFTWARE IS PROVIDED BY THE COPYRIGHT HOLDERS AND CONTRIBUTORS "AS IS" AND ANY EXPRESS OR IMPLIED WARRANTIES, INCLUDING, BUT NOT LIMITED TO, THE IMPLIED WARRANTIES OF MERCHANTABILITY AND FITNESS FOR A PARTICULAR PURPOSE ARE DISCLAIMED. IN NO EVENT SHALL THE COPYRIGHT HOLDERS OR CONTRIBUTORS BE LIABLE FOR ANY DIRECT, INDIRECT, INCIDENTAL, SPECIAL, EXEMPLARY, OR CONSEQUENTIAL DAMAGES (INCLUDING, BUT NOT LIMITED TO, PROCUREMENT OF SUBSTITUTE GOODS OR SERVICES; LOSS OF USE, DATA, OR PROFITS; OR BUSINESS INTERRUPTION) HOWEVER CAUSED AND ON ANY THEORY OF LIABILITY, WHETHER IN CONTRACT, STRICT LIABILITY, OR TORT (INCLUDING NEGLIGENCE OR OTHERWISE) ARISING IN ANY WAY OUT OF THE USE OF THIS SOFTWARE, EVEN IF ADVISED OF THE POSSIBILITY OF SUCH DAMAGE.

```
Do you accept the license terms? [yes|no]
```

```
>>> yes
```

**When you finish the minifrog install, put the command in the terminal**

```
ahmed@DEVELOPMENT:~$ source ~/.bashrc
ahmed@DEVELOPMENT:~$
```

### 3- Create the plantvarfilter environment to install:

**Use this command:**

**“mamba create -n pvf -c conda-forge -c bioconda python=3.11 samtools bcftools bowtie2 minimap2 plink”**

```
ahmed@DEVELOPMENT:~$ mamba create -n pvf -c conda-forge -c bioconda python=3.11 santools bcftools bowtie2 mininap2 plink
Found conda-prefix at '/home/ahmed/miniforge3/envs/pvf'. Overwrite?: [y/N] y
warning libmamba 'repo.anaconda.com', a commercial channel hosted by Anaconda.com, is used.

warning libmamba Please make sure you understand Anaconda Terms of Services.
warning libmamba See: https://legal.anaconda.com/policies/en/
[+] 4.6s
conda-forge/linux-64 _____ 4.0MB / 48.6MB @ 883.8kB/s 4.6s
conda-forge/noarch _____ 2.9MB / 23.2MB @ 629.8kB/s 4.6s
bioconda/linux-64 _____ 2.6MB / 5.3MB @ 573.2kB/s 4.6s
bioconda/noarch _____ 2.8MB / 5.8MB @ 688.3kB/s 4.6s
pkgs/main/linux-64 _____ 2.2MB / 9.8MB @ 475.1kB/s 4.6s
```

#### 4- When finished, activate mamba env

“mamba activate pvf”

Note:

The “PVF”, our env name, you can put any name.

```
ahmed@DEVELOPMENT:~$ mamba activate pvf
(pvf) ahmed@DEVELOPMENT:~$
```

#### 5- Now Install the Plantvarfilter package, Command:

“pip install plantvarfilter”

```
(pvf) ahmed@DEVELOPMENT:~$ pip install plantvarfilter
Collecting plantvarfilter
  Using cached plantvarfilter-0.2.4-py3-none-any.whl.metadata (13 kB)
Collecting numpy>=1.26 (from plantvarfilter)
  Using cached numpy-2.3.4-cp311-cp311-manylinux_2_27_x86_64.manylinux_2_28_x86_64.whl.metadata (62 kB)
Collecting pandas>=2.0 (from plantvarfilter)
  Using cached pandas-2.3.3-cp311-cp311-manylinux_2_24_x86_64.manylinux_2_28_x86_64.whl.metadata (91 kB)
Collecting scipy>=1.11 (from plantvarfilter)
  Using cached scipy-1.16.3-cp311-cp311-manylinux2014_x86_64.manylinux_2_17_x86_64.whl.metadata (62 kB)
Collecting matplotlib>=3.7 (from plantvarfilter)
  Using cached matplotlib-3.10.7-cp311-cp311-manylinux2014_x86_64.manylinux_2_17_x86_64.whl.metadata (11 kB)
Collecting scikit-learn>=1.3 (from plantvarfilter)
```

Then install fastlmm, geneview, xgboost

“pip install fastlmm”

“pip install geneview”

“pip install xgboost”

From this command, you can turn on the package:

“plantvarfilter”

