



# PIMGAVir Installation Manual

The current document describes the necessary steps to install and configure PIMGAVir pipeline on your machine. The PIMGAVir pipeline has been tested on Linux Ubuntu 20.04

## ADD NEW USER

```
sudo useradd -m -d /home/pimgavir pimgavir -s /bin/bash
sudo passwd pimgavir
sudo usermod -aG sudo pimgavir
```

## INSTALL PKGS

### TrimGalore

```
sudo apt install trim-galore
```

### SortmeRNA

```
sudo mkdir -p /usr/share/NGS-PKGs/Sortmerna
sudo chown -Rf pimgavir.pimgavir /usr/share/NGS-PKGs/
cd /usr/share/NGS-PKGs/Sortmerna
wget https://github.com/biocore/sortmerna/releases/download/v4.3.4/sortmerna-4.3.4-Linux.sh
./sortmerna-4.3.4-Linux
```

### Add the following line to .bashrc

```
/usr/share/NGS-PKGs/Sortmerna/sortmerna-4.3.4-Linux/bin/
```

### Kraken2

```
sudo apt install kraken2
```

### Kaiju

```
cd /usr/share/NGS-PKGs/
git clone https://github.com/bioinformatics-centre/kaiju.git
cd kanji/src
make
cd .. ; mkdir kaijudb ; cd kaijudb
../bin/kaiju-makedb -s viruses
```

### Add the following line to .bashrc

```
/usr/share/NGS-PKGs/kaiju/bin
```

### Krona-tools

```
cd Krona-master/KronaTools
sudo ./install.pl
mkdir taxonomy
```

```
./updateTaxonomy.sh
./updateAccessions.sh
mv Krona-master/KronaTools /usr/share/NGS-PKGs/
```

### Add the following line in .bashrc

```
/usr/share/NGS-PKGs/Krona-master/KronaTools/scripts
```

## Anaconda

```
wget -c https://repo.anaconda.com/archive/Anaconda3-2020.02-Linux-x86_64.sh
sudo mkdir /opt/Anaconda3
sudo bash Anaconda3-2020.02-Linux-x86_64.sh -u
PATH --> /opt/Anaconda3 [multiusers]
Do you wish the installer to initialize Anaconda3 by running conda init? YES
sudo chmod -Rf a+w Anaconda3/
```

### Add the following PATH to the .bashrc

```
/opt/Anaconda3/bin
```

## MEGAHIT

```
conda install -c bioconda megahit
```

## SeqKit

```
conda install -c bioconda seqkit
```

## SPADES

```
cd /usr/share/NGS-PKGs/
wget -c https://cab.spbu.ru/files/release3.15.3/SPAdes-3.15.3-Linux.tar.gz
tar -xzf SPAdes-3.15.3-Linux.tar.gz
```

### Add the following line to .bashrc

```
/usr/share/NGS-PKGs/SPAdes-3.15.3-Linux/bin
```

## QUAST

```
sudo apt-get install python-setuptools
cd /usr/share/NGS-PKGs/
wget -c https://jaist.dl.sourceforge.net/project/quast/quast-5.0.2.tar.gz
tar -xzf quast-5.0.2.tar.gz
cd quast-5.0.2/
sudo ./setup.py install_full

sudo vi /usr/local/lib/python3.8/dist-packages/quast-5.0.2-
py3.8.egg/quast_libs/site_packages/jsontemplate/jsontemplate.py
```

**NB: All occurrences of cgi must be replaced with html (specifically on line 583)**

## Bowtie2

```
sudo apt-get install bowtie2
```

## Samtools

```
sudo apt-get install autoconf automake make gcc perl zlib1g-dev libbz2-dev liblzma-dev libcurl4-gnutls-dev libssl-dev libncurses5-dev
wget https://github.com/samtools/samtools/releases/download/1.15.1/samtools-1.15.1.tar.bz2
bunzip2 samtools-1.15.1.tar.bz2
tar -xvf samtools-1.15.1.tar
cd samtools-1.15.1
./configure
make
sudo make install
```

## Pilon

```
sudo apt install pilon
```

## Prokka

```
conda install -c conda-forge -c bioconda -c defaults prokka
prokka --setupdb
```

## NCBI-Blast+

```
sudo apt install ncbi-blast+
```

## Vsearch

```
sudo apt install vsearch
```

## Diamond

```
cd /usr/share/NGS-PKGs
wget http://github.com/bbuchfink/diamond/releases/download/v2.0.13/diamond-linux64.tar.gz
tar -xzvf diamond-linux64.tar.gz
```

**Add the following line to the .bashrc**

```
/usr/share/NGS-PKGs
```

## Taxonkit

```
cd /usr/share/NGS-PKGs
wget https://github.com/shenwei356/taxonkit/releases/download/v0.10.1/taxonkit_linux_amd64.tar.gz
tar -xzvf taxonkit_linux_amd64.tar.gz
```

**Add the following line to the .bashrc**

```
/usr/share/NGS-PKGs
```

## Seqtk

```
sudo apt install seqtk
```

## INSTALL DBS

It is supposed the following databases are already configured and working on a local machine

```
mkdir /usr/share/DBs
cp -Rf /mnt/Diamond-RefSeqProt/ /usr/share/DBs [check for permission]
cp -Rf /mnt/Refseq/ /usr/share/DBs
cp -Rf /mnt/DBs/blastdb/ /usr/share/DBs/
```

## VARIABLES TO SET

```
pimgavir.sh:
DiamondDB="/DBs/Diamond-RefSeqProt/refseq_protein_nonredund_diamond.dmnd"
PathToRefSeq="/DBs/RefSeq"

pre-process.sh:
refSLR138="/DBs/SILVA/slr138.fasta"
refSSR138="/DBs/SILVA/ssr138.fasta"
SilvaDB_idx="/DBs/SILVA/idx/"

NB: ln -s $SilvaDB_idx sortmeRNA_wd/idx

reads-filtering.sh:
TMP="/tmp"

taxonomy.sh:
KrakenViralDB="/DBs/KrakenViral"
kaijuNodes="/usr/share/NANOPORE-PKGs/kaiju/kaijudb/nodes.dmp"
kaijuNames="/usr/share/NANOPORE-PKGs/kaiju/kaijudb/names.dmp"
kaijuDB="/usr/share/NANOPORE-PKGs/kaiju/kaijudb/viruses/kaiju_db_viruses.fmi"

krona-blast.sh
ref_viruses_rep_genomes="/DBs/blastdb/ref_viruses_rep_genomes"
krona="/usr/local/bin/ktImportTaxonomy"

clustering.sh
ConcScript="/usr/share/NGS-PKGs/scripts/concatenate_reads.py"
```

## COPY SCRIPTS

```
scp * pimgavir@192.168.1.103:/usr/share/NGS-PKGs/scripts

pimgavir@frodo:/usr/share/NGS-PKGs/scripts$ ll
total 48
drwxrwxr-x 2 pimgavir pimgavir 4096 4月 19 16:03 ./
drwxr-xr-x 7 pimgavir pimgavir 4096 4月 15 15:34 ../
-rwxr-xr-x 1 pimgavir pimgavir 3938 4月 19 16:03 clustering.sh*
-rwxrwxr-x 1 pimgavir pimgavir 974 4月 19 16:03 concatenate_reads.py*
-rwxr-xr-x 1 pimgavir pimgavir 2226 4月 19 16:03 krona-blast.sh*
-rwxr-xr-x 1 pimgavir pimgavir 3021 4月 19 16:03 Misaele_Filter_Param.sh*
-rwxr-xr-x 1 pimgavir pimgavir 8583 4月 19 16:03 pimgavir.sh*
-rwxr-xr-x 1 pimgavir pimgavir 3273 4月 19 16:03 pre-process.sh*
-rwxr-xr-x 1 pimgavir pimgavir 1987 4月 19 16:03 reads-filtering.sh*
-rwxr-xr-x 1 pimgavir pimgavir 3947 4月 19 16:03 taxonomy.sh*
```

**Add the following line to .bashrc**

```
/usr/share/NGS-PKGs/scripts
```

**At the end of the .bashrc file of pigmavir user, the following line have to be present:**

```
#PIMGAVir-Pipeline
export PATH=$PATH:/usr/share/NANOPORE-PKGs/kaiju/bin:/usr/share/NANOPORE-PKGs/Krona-
master/KronaTools/scripts:/opt/Anaconda3/bin:/usr/share/NGS-PKGs/SPAdes-3.15.3-Linux/bin:/usr/share/NGS-
PKG:/usr/share/NGS-PKGs/scripts:/usr/share/NGS-PKGs/Sortmerna/sortmerna-4.3.4-
Linux/bin:/usr/share/NGS-PKGs/sratoolkit.2.11.2-ubuntu64/bin:/opt/Anaconda3/bin/
```

**Load the environment by**

source .bashrc command