Installation Guide

MutaNET

Version 2.0

MutaNET comes with a next generation sequencing (NGS) pipeline that calls
mutations based on paired–end NGS reads, an automated mutation analysis tool and
various file converters and mergers. The mutation analysis feature considers the
coding region, protein domains, regulation and transcription factor binding site
information, and can be used to analyse the potential impact of mutations on genes
of interest.

This guide gives instructions for installing MutaNET as well as required software on

Linux, Mac OS X and Windows. These instructions were tested on Ubuntu 16.10,

Mac OS El Capitan, Windows 8 and Windows 10.

Contents

1 Quickstart

1.1 Starting MutaNET

MutaNET comes as **Python 3 source code** as well as an **executable** for Windows.

Executable (Windows) To start MutaNET, double–click on the executable **MutaNET32.exe** or **MutaNET64.exe**, depending on whether you have a 32–bit or 64–bit Windows installation. If you are not sure, choose the 32–bit executable. Make sure that the executable remains in the same directory as the **config.yaml** file. Otherwise the user interface will not start.

From Source Open a command prompt or terminal and execute the following command:

```
python3 source_folder_path/mutaNET.py
```

or on **Windows** depending on your Python installation:

```
python source_folder_path/mutaNET.py
```

Source_folder_path is the path to the folder containing the source code of MutaNET. This requires Python 3 to be installed.

The following chapters explain how to install Python 3, as well as programs required for the NGS pipeline of MutaNET on **Windows**, **Linux** and **Mac OS X**.

1.2 Example Data Sets

When starting MutaNET for the first time, the file paths for small example data sets for the NGS pipeline, mutation analysis and file converters are already loaded to allow quick testing. Keep in mind that for the NGS pipeline extra programs need to be installed.

The settings for the example data can be restored any time by clicking on **Settings → Restore default settings**. The data sets can be found in the **example_data** folder, if you want to have a look at the file formats.

2   Installing MutaNET on Linux

For Linux, MutaNET can be downloaded as Python 3 source code, which requires Python 3 to be installed. The NGS pipeline, mutation analysis and file conversion function independently. **Just running the mutation analysis or converting/merging files does not require any additional programs!** For the NGS pipeline, however, Burrows Wheeler Aligner (BWA) [1], SAMTools [2], VarScan [3] and a Java Runtime Environment (JRE) are necessary.

In order to install the required programs, you can either use the provided installation script (Section 2.1, on Debian or Ubuntu) **or manually install** BWA, SAMTools, VarScan and JRE (Section 2.2) as well as Python 3 (Section 2.3).

2.1   Installation Script for Ubuntu or Debian

The installation script can be used to install BWA, SAMTools, VarScan and/or Python 3. Open a terminal and execute the following command:

```
bash MutaNET_path/install_ubuntu_debian.sh
```

MutaNET_path is the path to the folder containing the file **install_ubuntu_debian.sh**.

2.2   Installing BWA, SAMTools, VarScan and JRE

These programs are required for the NGS pipeline. **You do not need to install them if you only wish to run the mutation analysis or use the file converter/merger!**

Open a terminal and execute the following commands:

```
sudo apt-get install bwa
sudo apt-get install samtools
sudo apt-get install varscan
sudo apt-get install default-jre
```

If you are not using Ubuntu or Debian, replace **apt–get** with the name of the package manager of your Linux installation.

2.3   Installing Python 3

Python 3 is required to run MutaNET from source code. To install Python 3, open a terminal and execute the following commands:

```
sudo apt-get install python3
sudo apt-get install python3-pip
```

If you are not using Ubuntu or Debian, replace **apt–get** with the name of the package manager of your Linux installation. To install the required Python 3 packages, open a terminal and execute the following commands:

```
pip3 install tkinter
pip3 install numpy
pip3 install matplotlib
pip3 install scipy
pip3 install fpdf
pip3 install pyyaml
```

MutaNET can now be started by executing the following command:

```
python3 source_folder_path/mutaNET.py
```

source_folder_path is the path to the source folder of MutaNET.

3 Installing MutaNET on Mac OS

MutaNET can be downloaded as Python source code, which requires Python 3 to be installed. The NGS pipeline, mutation analysis and file conversion function independently. **Just running the mutation analysis or converting/merging files does not require any additional programs!**

For the NGS pipeline, however, Burrows Wheeler Aligner (BWA) [1], SAMTools [2], VarScan [3] and a Java Runtime Environment (JRE) are necessary.

In order to install the required programs, you can either use the provided installation script (Section 3.1) **or manually install** BWA, SAMTools, VarScan and JRE (Section 3.3) as well as Python 3 (Section 3.4).

3.1 Installation Script

The installation script can be used to install BWA, SAMTools, VarScan and/or Python 3.

Open a terminal, e.g. via your launchpad, and execute the following command:

```
bash MutaNET_path/install_mac_os.sh
```

MutaNET_path is the path to the folder containing the file **install_mac_os.sh**.

3.2 Installing Homebrew

Homebrew is a package manager for Mac OS that is used for easily installing various programs such as Python 3, BWA, SAMTools, VarScan and JRE.

Open a terminal, e.g. via your launchpad, and execute the following command:

```
/usr/bin/ruby -e "$(curl -fsSL https://raw.githubusercontent.com/Homebrew/install/
master/install)"
```

3.3 Installing BWA, SAMTools, VarScan and JRE

These programs are required for the NGS pipeline. **You do not need to install them if you only wish to run the mutation analysis or use the file converters/mergers!**

After installing Homebrew, open a terminal, e.g. via your launchpad, and execute the following commands:

```
brew update
brew cask install java
brew install bwa
brew install samtools
brew install brewsci/bio/varscan
```

3.4 Installing Python 3

Python 3 is required to run MutaNET from source code. First install Homebrew. Open a terminal, e.g. via your launchpad, and execute the following commands:

```
brew install python3
pip3 install numpy
pip3 install matplotlib
pip3 install scipy
pip3 install fpdf
pip3 install pyyaml
```

MutaNET can now be started by executing the following command:

```
python3 source_folder_path/mutaNET.py
```

Source_folder_path is the path to the folder containing the source code of MutaNET.

On some Mac OS versions, some default Python packages might cause errors while starting MutaNET. If that happens, try the following work–around:

```
brew install homebrew/dupes/tcl-tk
brew uninstall python3
brew install python3 --with-brewed-tk
```

4   Installing MutaNET on Windows

MutaNET can be downloaded as an executable, which can be started via double–click, or as Python source code. The source code requires Python 3 to be installed, whereas the executable does not. See Section 4.6 for instructions on how to install Python 3, if you wish to run the source code.

The NGS pipeline, mutation analysis and file conversion function independently. **Just running the mutation analysis or converting/merging files does not require any additional programs!**

For the NGS pipeline, however, Burrows Wheeler Aligner (BWA, Section 4.4) [1], SAMTools (Section 4.4) [2], VarScan (Section 4.5) [3] and a Java Runtime Environment (JRE, Section 4.2) are necessary. Since BWA and SAMTools are UNIX based programs, Cygwin needs to be installed first (Section 4.1).

4.1   Installing Cygwin

Cygwin is required to install and run BWA, SAMTools and VarScan, which in turn are required for the NGS pipeline. **You do not need to install it if you only wish to use the mutation analysis or the file converter/merger!**

1. Go to the download page (Cygwin) and either download the **setup.exe** for 32–bit or 64–bit depending on your Windows installation.

2. Navigate to the download folder.

3. Install Cygwin.

   1. Double-click on the **setup.exe** you just downloaded.
   2. Click on **Next** in the installation wizard.
   3. Select **Install from Internet** and click on **Next**.
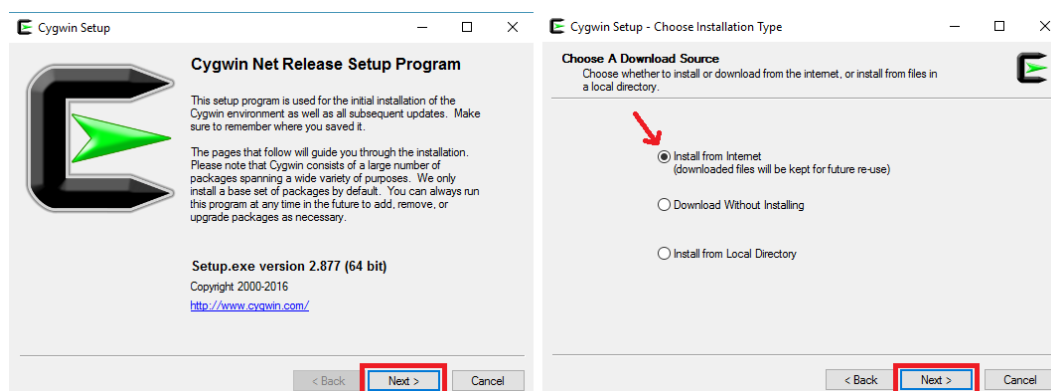
   4. **Optional:** Choose a custom installation folder or select **Just Me** (not recommended). Continue with **Next**.

5. **Optional:** Choose a custom download and intermediary folder and continue by clicking on **Next**.
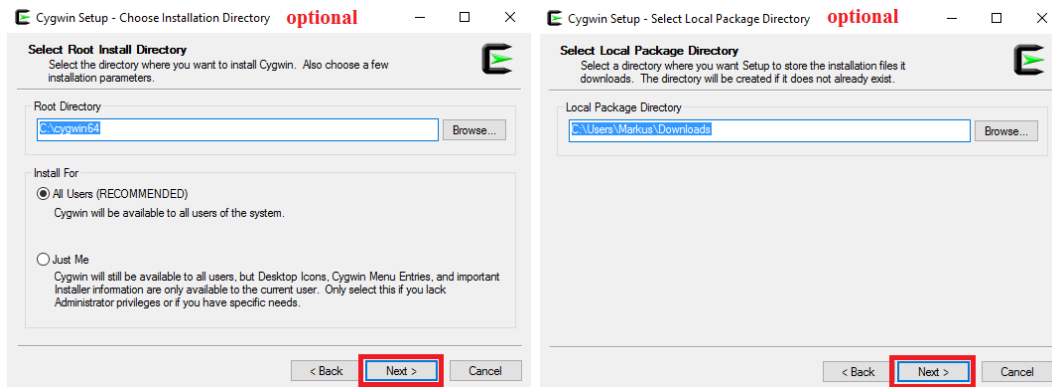
6. Select **Direct Connection** and continue by clicking on **Next**.

7. From the list of download sites, select a **http** site (**ftp** is going to take a lot longer), preferably near you. Continue with **Next**.
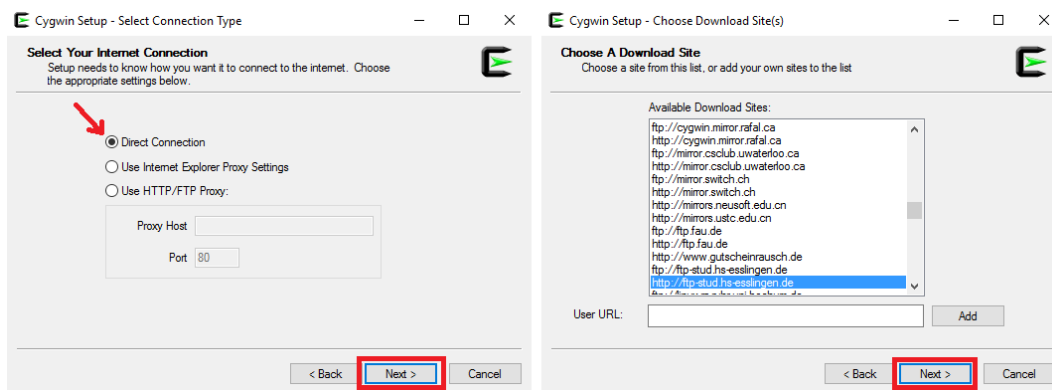

8. Type **gdb** into the search bar. Click on **Default** next to the sub–categories to change them to **Install**. **Do NOT click on Default next to All!**
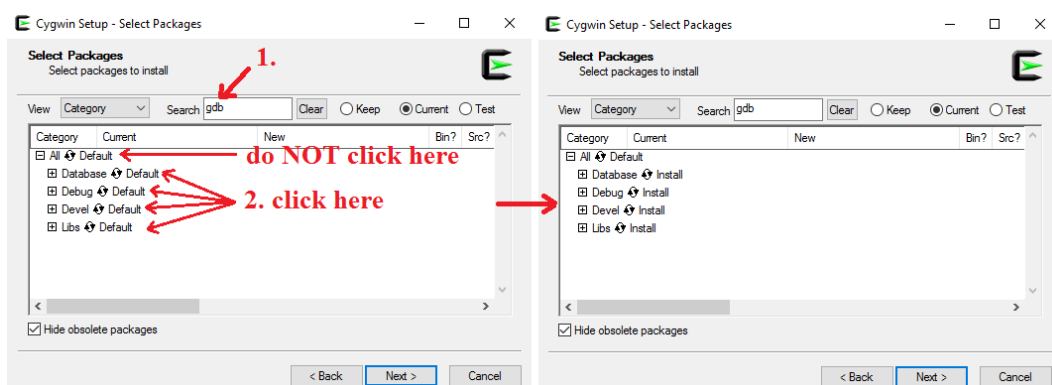

9. Repeat the previous step for the packages **gcc-g++**, **make**, **libncurses**, **zlib**, **zlib-devel** and **wget**.

10. Continue by clicking on **Next**.

11. Select **Select required packages (RECOMMENDED)** and click on **Next**. Wait for Cygwin to download and install the packages. This can take a while.
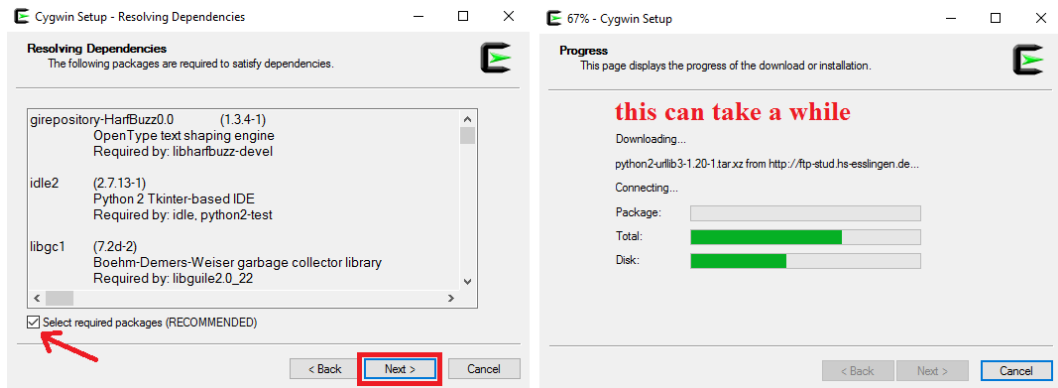
12. **Optional:** Create a Desktop or Start Menu icon for the Cygwin command prompt.

13. Complete the installation by clicking on **Finish**.
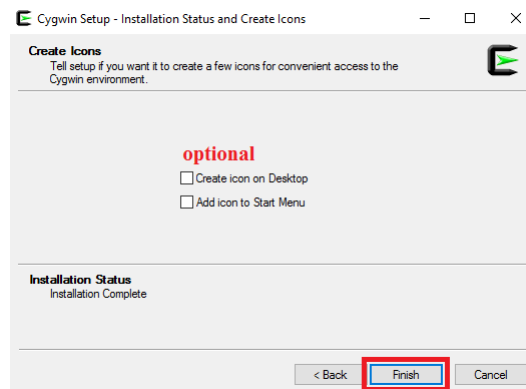
4. Continue with Section 4.3 to add Cygwin to the Windows path.

4.2 Installing a Java Runtime Environment

A Java Run Time Environment (JRE) is required for VarScan, which in turn is required for the NGS pipeline. **You do not need to install it if you only wish to use the mutation analysis or the file converter/merger!**

1. Download a Java Runtime Environment installer.

    1. Go to main download page (Oracle).
    2. Scroll down until you see **JRE download** in the middle, and follow the link to a page with installers for different operating systems.

3. Select **Accept License Agreement**.
4. Download the offline installer for windows (**jre-something-windows.exe**). Choose the **x86** version if you have a 32-bit and **x64** if you have a 64-bit Windows installation.
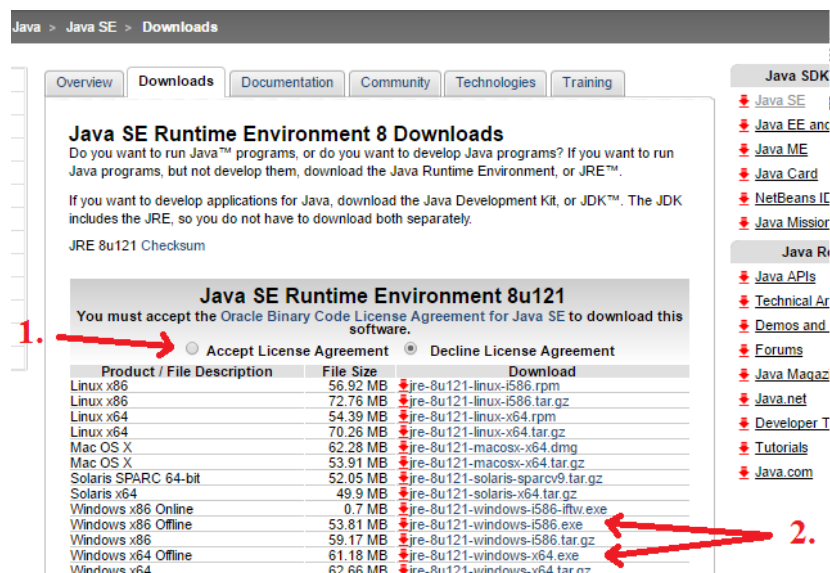
2. Navigate to the downloads folder.

3. Install the Java Runtime Environment.

    1. Double click on the executable you just downloaded.

    2. **Optional:** Select **Change destination folder** if you want to customise where the Java Runtime Environment is going to be installed.

    3. Click on **Install** and wait for the installation to proceed.
       Optional: If you have performed step 2, you now have the opportunity to select the installation folder. Click on **Next** to continue with the installation.

    4. Click on **Close** to complete the installation.
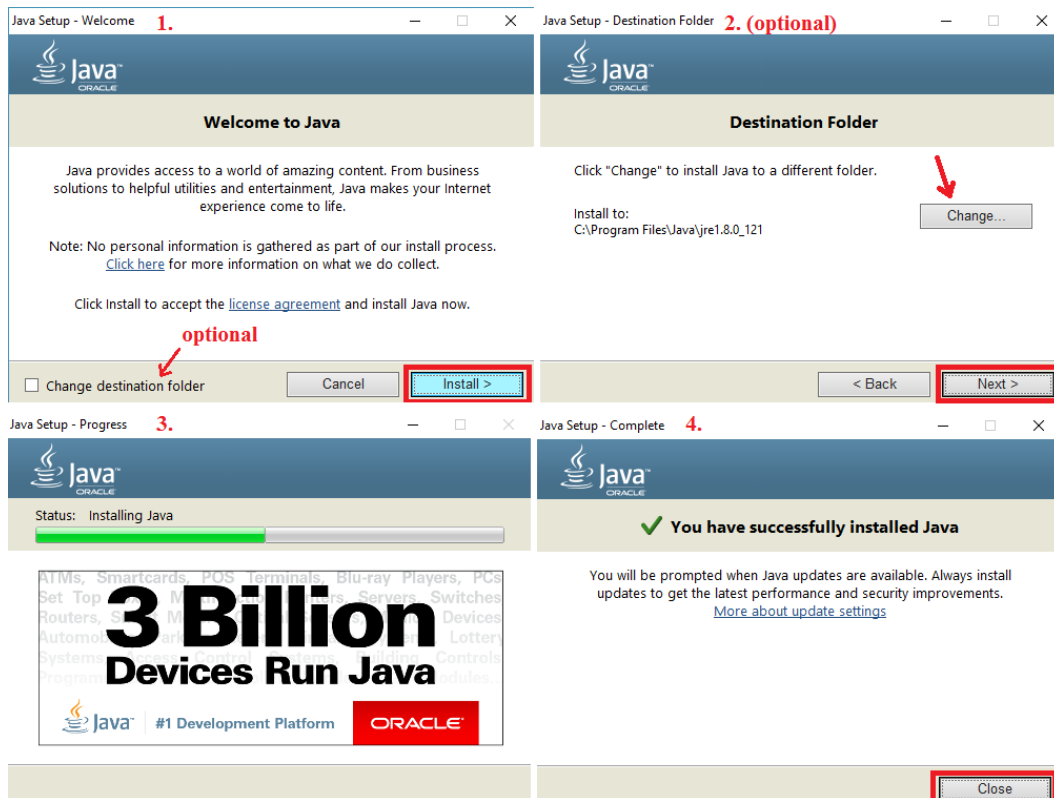
4. Continue with Section 4.3 to add Java to the Windows path.

4.3 Adding Java and Cygwin to the Windows Path

This step is required for installing BWA and SAMTools, as well as running VarScan and the NGS pipeline. **You do not need to this if you only wish to use the mutation analysis or the file converter/merger!**

1. Open the **System Properties** menu.

   1. Press **Windows key + R**. The **Run** window opens.
   2. Type **SystemPropertiesAdvanced** into the entry field and click on **OK**.

   Alternatively, type **environment variables** into the Windows search bar and click on **Edit the system environment variables**.
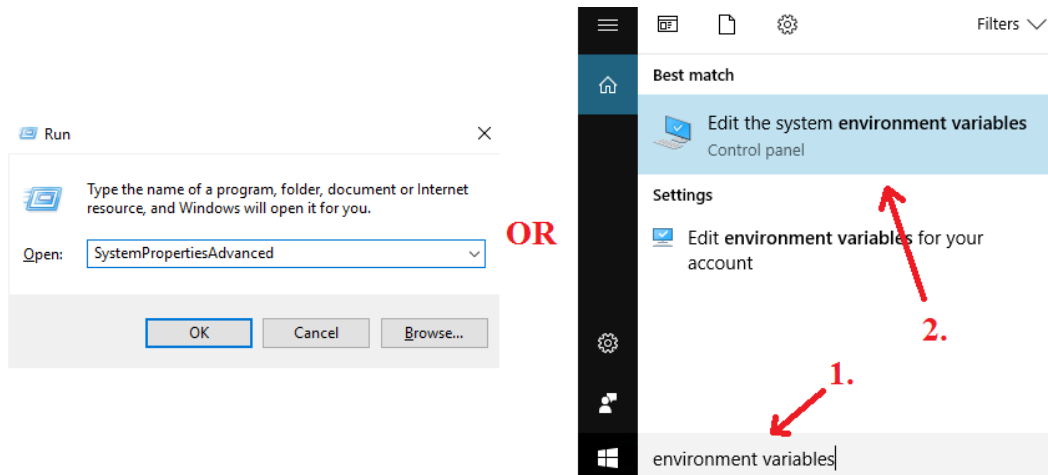
2. Click on **Environment Variables...** in the **System Properties** menu.
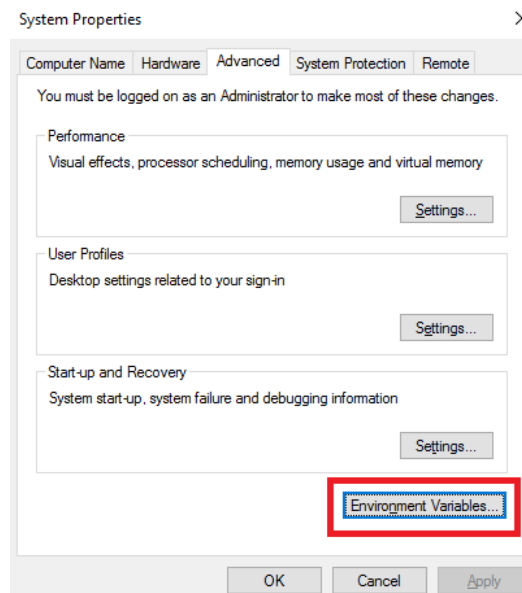
3. Add Cygwin and Java to the **System Path**.

   1. Select **Path** and click on **Edit....**
   2. Click on **New** and type **%JAVA_HOME%\bin**.
   3. Click on **New** and enter the path to Cygwin installation folder, the default is **C:\cygwin64\bin** for 64–bit and **C:\cygwin\bin** for 32–bit.
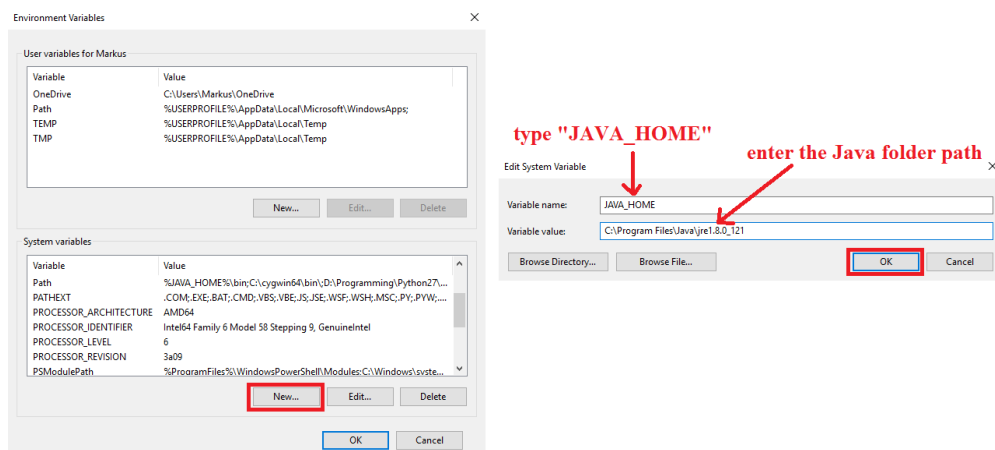   4. Click on **OK** to close the window.
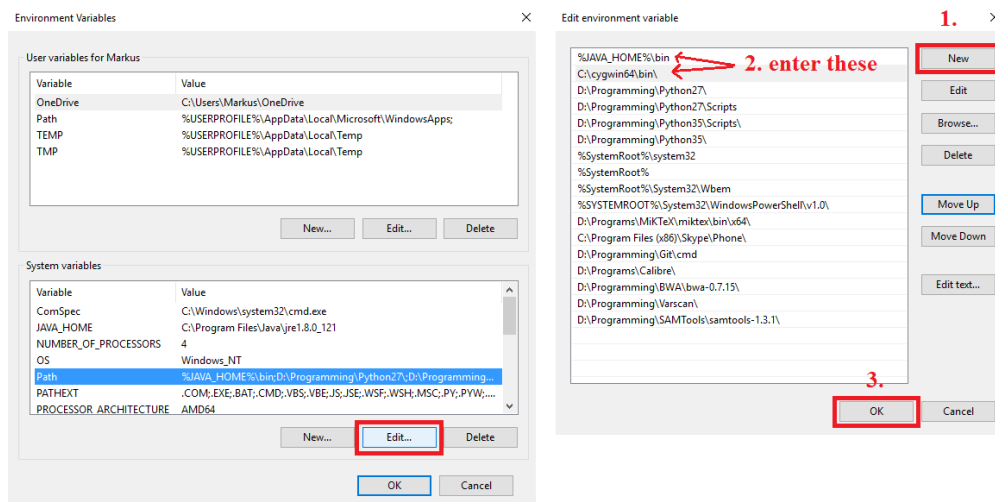
4. Add Java **.jar** files to recognised executables
   1. Select **PATHTEXT** and click on **Edit...**.
   2. Add **;.JAR** to the **Variable Values** if it is not already in the list.
   3. Click on **OK** to close the window.
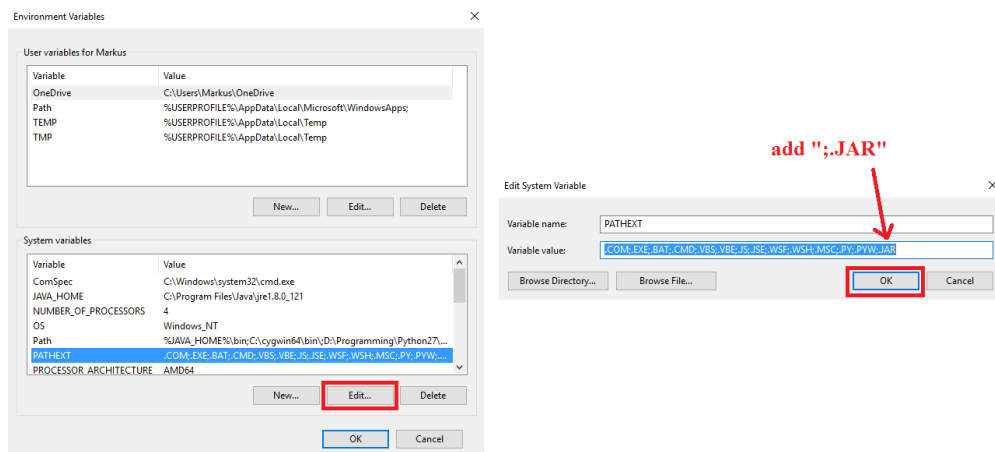
5. Set the Java home folder.
   1. If **JAVA_HOME** is not already in the **System Variable** list, click on **New**, otherwise click on **Edit...**.
   2. Type **JAVA_HOME** into the **Variable name** field.
   3. Enter the path to the Java installation folder into the **Variable value** field. The default path is **C:\Program Files\Java \jre_version**, where **version** is the version number you downloaded.

6. Click on **OK** to close all windows.

## 4.4 Installing Burrows Wheeler Aligner and SAMTools

Burrows Wheeler Aligner and SAMTools are required for the NGS pipeline. **You do not need to install them if you only wish to use the mutation analysis or the file converter/merger!** Their installation requires Cygwin to be installed on the Windows path (see Section 4.1).

1. Download BWA (SourceForge) and SAMTools **1.3** (SourceForge). **Newer versions of SAMTools currently do not work on Cygwin due to a missing library!**
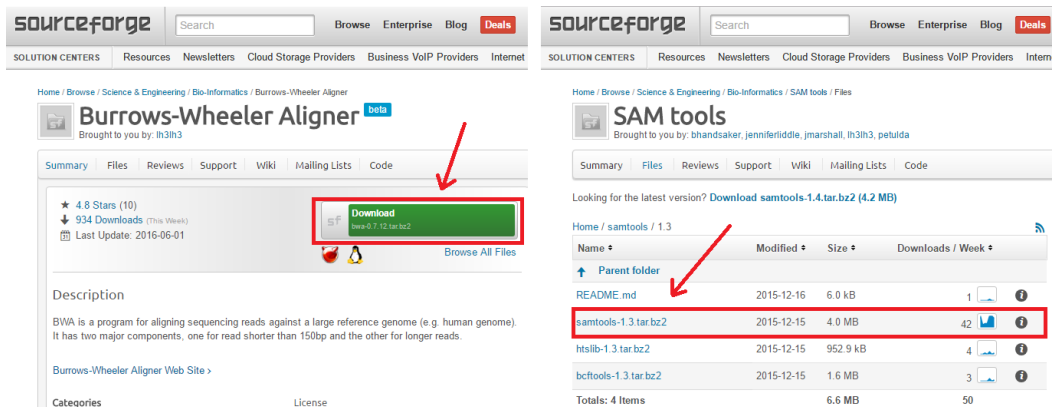

2. Navigate to the download folder.

3. Right click on the compressed downloaded **.tar.bz2** file(s) and extract the contents using a program such as 7zip or WinRAR (here shown with WinRAR).

   If you use 7zip, you need to perform to extraction steps: the first one from **.tar.bz2** to *.tar* and the second one to extract the folders from the **.tar** file.
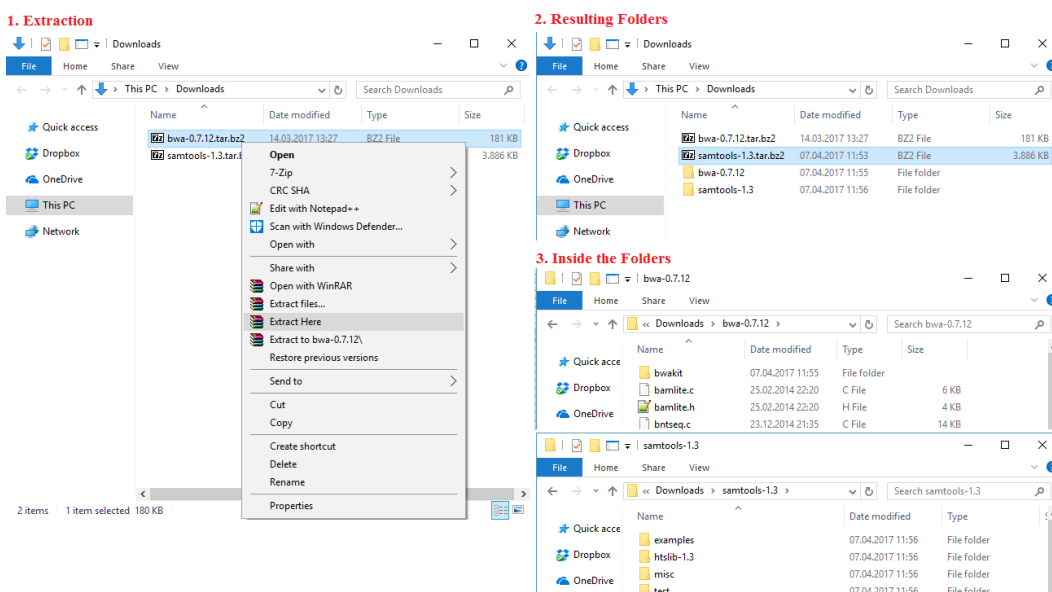

4. Open the windows **Command Prompt**.

   1. Press **Windows key + R**. The **Run** window opens.
   2. Type **cmd** into the entry field and click on **OK**.

   Alternatively, type **cmd** into the Windows search bar and click on **Command Prompt**.

OR
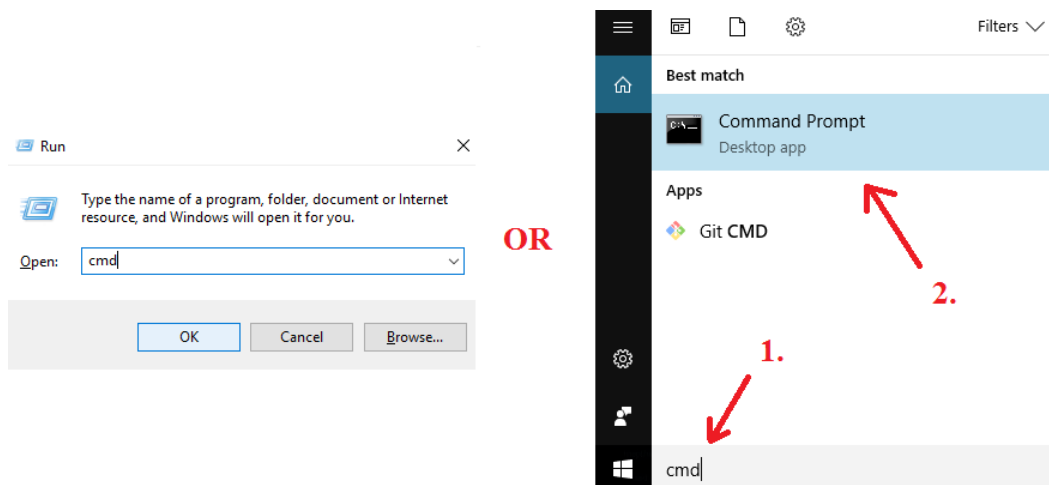
5. Compile BWA and/or SAMTools. This requires Cygwin to be installed on the Windows path (see Section 4.1).

   1. Navigate to the BWA (or SAMTools) folder inside the command prompt by using the **cd** command followed by the path to the folder.
   2. Compile with the **make** command. A wall of text appears in the command window that you can ignore.
   3. Check if there is a file called **bwa.exe** (or **samtools.exe**) in the BWA (or SAMTools) folder.
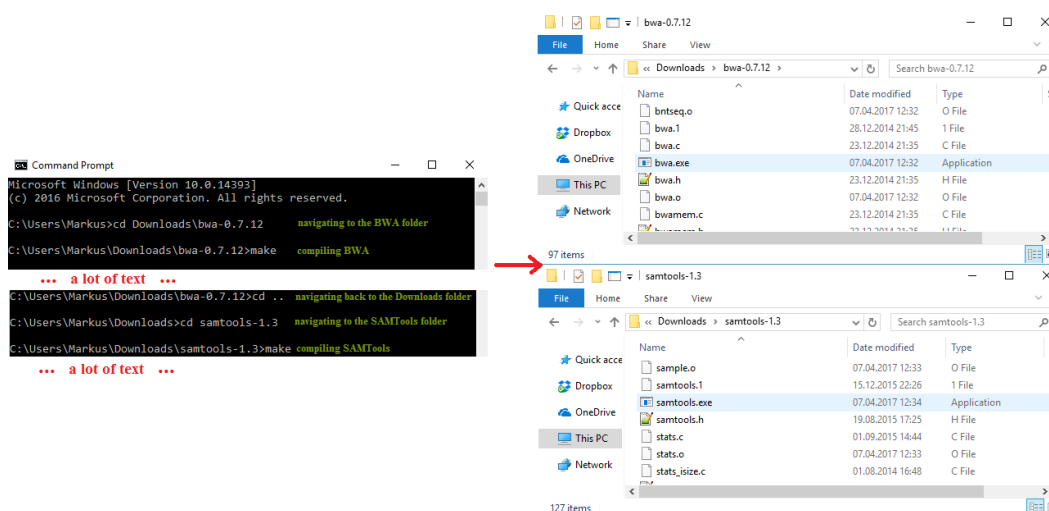
6. You can now move the folders to any location you like.

4.5   Installing VarScan

VarScan is required for the NGS pipeline. **You do not need to install it if you only wish to use the mutation analysis or the file converter/merger!** VarScan requires a Java Runtime Environment to run (see Section 4.2).

1. Download VarScan (SourceForge).

2. Navigate to the download folder.

3. Move the **VarScan.jar** executable you just downloaded to any location you like.

4. If you have not done so already, install a Java Runtime Environment (Section 4.2).

4.6 Installing Python 3

Python 3 is required for running MutaNET from source. **If you are using the executable, you do not need to install it!**

1. Download Python 3 and the required packages.

   1. Go to the Python download page (Python).

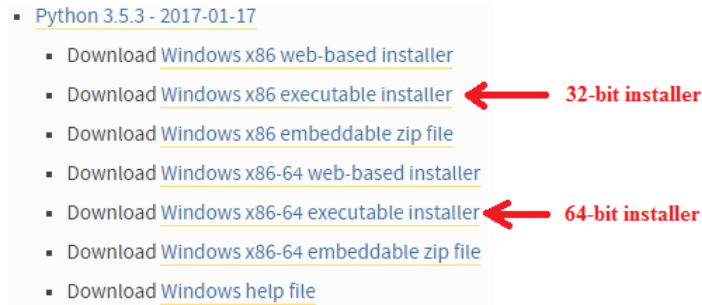   2. Download an **executable installer** of Python 3 version, e.g. Python 3.5. Select the **x86** installer for 32–bit and the **x86–x64** installer for 64–bit Windows.

   - Python 3.5.3 - 2017-01-17
     - Download Windows x86 web-based installer
     - Download Windows x86 executable installer ← **32-bit installer**
     - Download Windows x86 embeddable zip file
     - Download Windows x86-64 web-based installer
     - Download Windows x86-64 executable installer ← **64-bit installer**
     - Download Windows x86-64 embeddable zip file
     - Download Windows help file

   3. Go to the download page for **numpy** for Windows (University of California).

   4. Download the **numpy+mlk.whl** for your Python and Windows version. For example, if you installed Python 3.5 and have a 64–bit Windows installation, download **numpy+mlk–cp35–cp35m–win_amd64.whl**

   NumPy, a fundamental package needed for scientific computing with Python.
   Numpy+MKL is linked to the Intel® Math Kernel Library and includes required DLLs in the numpy.core directory.
   numpy-1.11.3+mkl-cp27-cp27m-win32.whl
   numpy-1.11.3+mkl-cp27-cp27m-win_amd64.whl
   numpy-1.11.3+mkl-cp34-cp34m-win32.whl
   numpy-1.11.3+mkl-cp34-cp34m-win_amd64.whl
   numpy-1.11.3+mkl-cp35-cp35m-win32.whl
   numpy-1.11.3+mkl-cp35-cp35m-win_amd64.whl
   numpy-1.11.3+mkl-cp36-cp36m-win32.whl
   numpy-1.11.3+mkl-cp36-cp36m-win_amd64.whl
   numpy-1.12.1+mkl-cp27-cp27m-win32.whl
   numpy-1.12.1+mkl-cp27-cp27m-win_amd64.whl
   numpy-1.12.1+mkl-cp34-cp34m-win32.whl
   numpy-1.12.1+mkl-cp34-cp34m-win_amd64.whl
   numpy-1.12.1+mkl-cp35-cp35m-win32.whl ← **Python 3.5 numpy version for**
   numpy-1.12.1+mkl-cp35-cp35m-win_amd64.whl ← **32-bit or 64-bit Windows**
   numpy-1.12.1+mkl-cp36-cp36m-win32.whl
   numpy-1.12.1+mkl-cp36-cp36m-win_amd64.whl

   5. Go to the download page for **scipy** for windows (University of California).

   6. Download the **scipy.whl** for your Python and Windows version. For example, if you installed Python 3.5 and have a 64–bit Windows installation, download **scipy–cp35–cp35m–win_amd64.whl**

   7. Go to the download page for **matplotlib** for windows (University of California).

   8. Download the newer version of **matplotlib.whl** for your Python and Windows version. For example, if you installed Python 3.5 and have a 64–bit Windows installation, download **matplotlib–cp35–cp35m–win_amd64.whl**

2. Navigate to the download folder.

3. Install Python 3.

   1. Double click on the Python installer you just downloaded.

   2. Select **Add Python 3 to PATH**.

   3. Click on **Customize installation**.

   4. Select at least **pip**, **tcl/tk and IDLE** and **py launcher** before cclicking on **Next**.
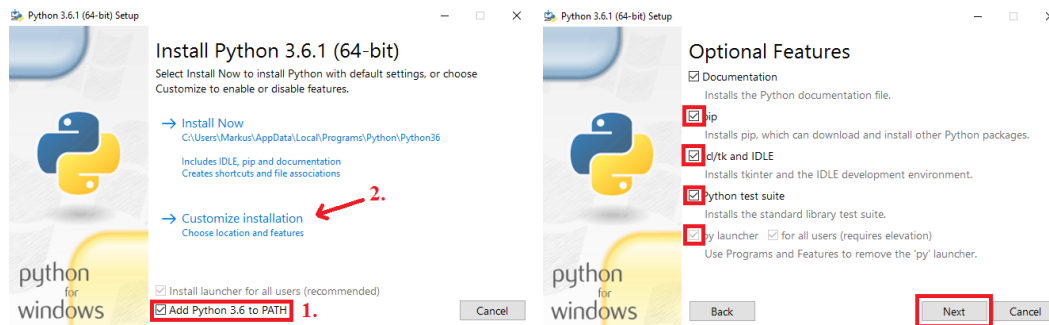
5. Select at least **Associate files with Python** and **Add Python to environment variables**. Choose where to install Python and continue by clicking on **Install**.

6. Wait for the installation to complete and then click on **Close**.
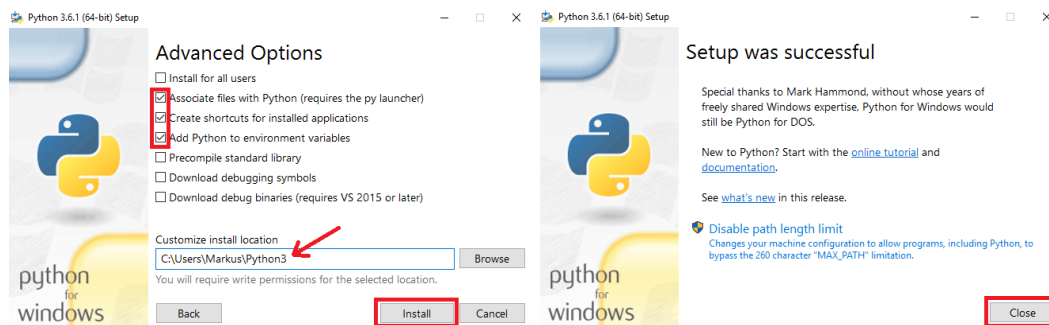

4. Add Python 3 to the Windows path.

    1. Open the **System Properties** menu.

        1. Press **Windows key + R**. The **Run** window opens.

        2. Type **SystemPropertiesAdvanced** into the entry field and click on **OK**.

       Alternatively, type **environment variables** into the Windows search bar and click on **Edit the system environment variables**.
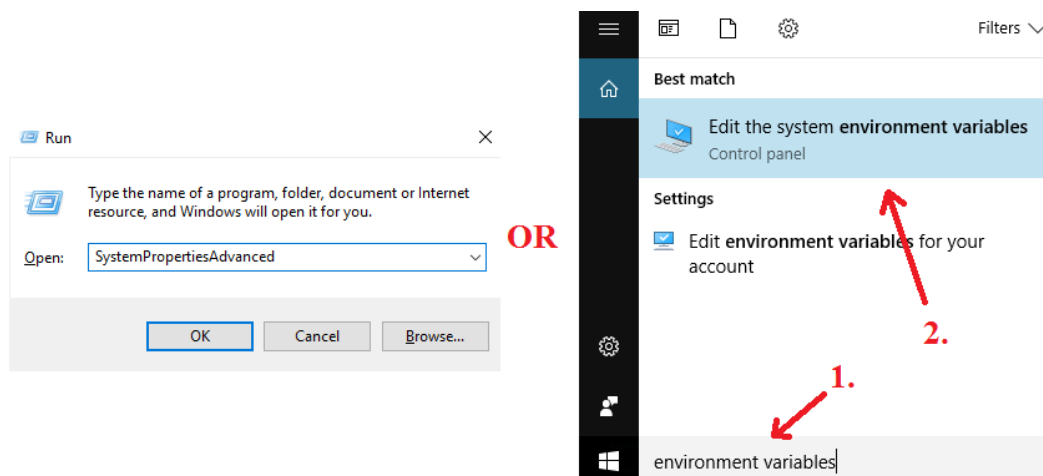

    2. Click on **Environment Variables...** in the **System Properties** menu.
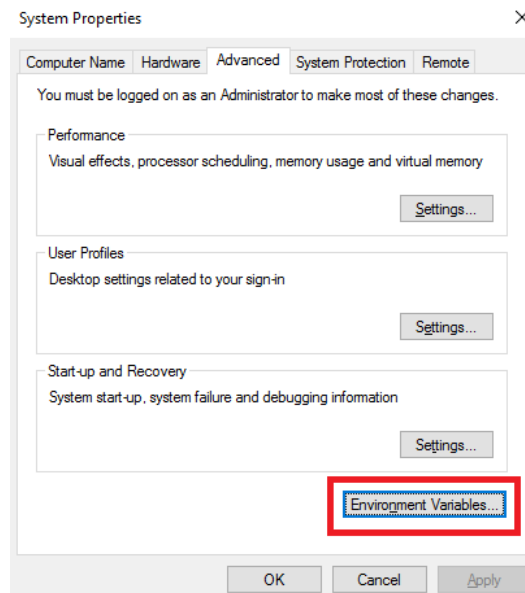
3. Select **Path** and click on **Edit...**.

4. If your Python 3 installation folder is not in the list, click on **New** and type **python_path**, where **python_path** is the path to your Python installation folder, in this example **C:\Users\Markus\Python3**.

5. Do the same for **python_path\Scripts**.

6. Click on **OK** to close all windows.
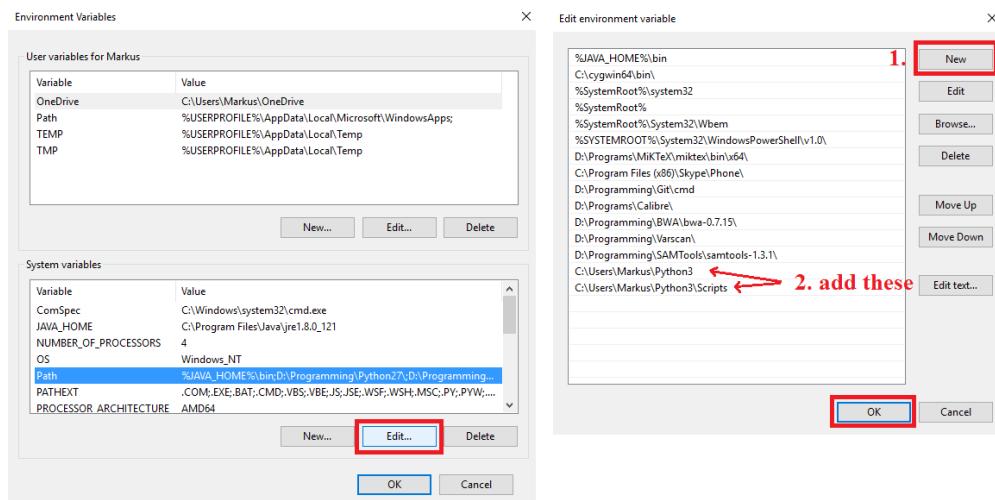
5. Install the required Python 3 packages.

   1. Open the Windows **Command Prompt**.
      1. Press **Windows key + R**. The **Run** window opens.
      2. Type **cmd** into the entry field and click on **OK**.

      Alternatively, type **cmd** into the Windows search bar and click on **Command Prompt**.
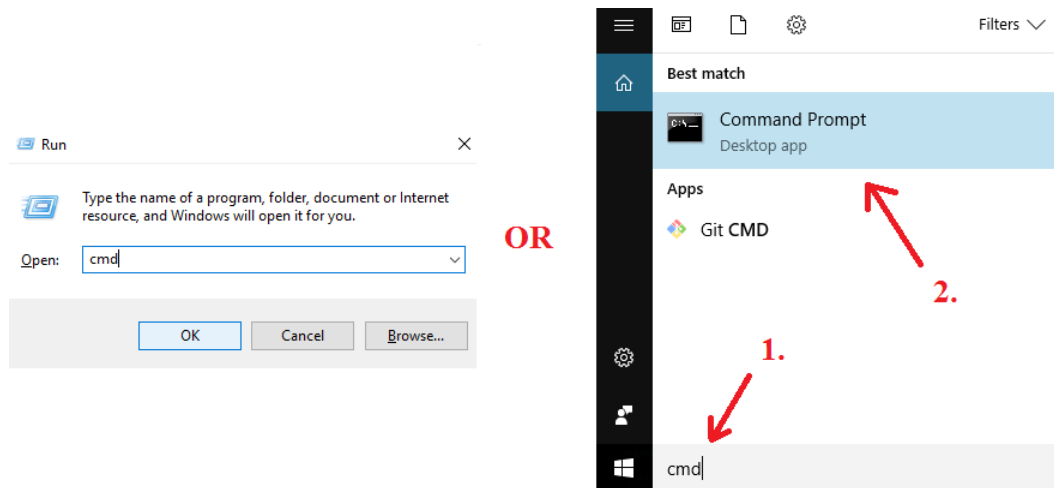
2. Navigate to the Python package download folder using the **cd** command followed by the folder path. For example,

```
cd C:\Users\username\Downloads
```

3. Install **numpy** using the **numpy+mlk.whl** you downloaded earlier using the **pip.exe install package.whl** command.

For example, if you downloaded **numpy-1.12.1+mkl-cp36-cp36m-win_amd64.whl**, the command is as follows:

```
pip3.exe install numpy-1.12.1+mkl-cp36-cp36m-win_amd64.whl
```

4. Repeat the previous step first for the **scipy** package and then for the **matplotlib** package.

5. Install **pyyaml** and **fpdf** with the following commands:

```
pip3.exe install fpdf
pip3.exe install pyyaml
```

6. MutaNET can now be started by opening the **Command Prompt** (see above), and then executing the following command:

```
python source_folder_path\gui.py
```

Source_folder_path is the path to the folder containing the source code of MutaNET.

References

[1] H. Li and R. Durbin. "Fast and accurate long-read alignment with Burrows-Wheeler transform". In: *Bioinformatics* 26.5 (Mar. 2010), pp. 589–595. URL: http://bio-bwa.sourceforge.net/.

[2] H. Li et al. "The Sequence Alignment/Map format and SAMtools". In: *Bioinformatics* 25.16 (Aug. 2009), pp. 2078–2079. URL: http://samtools.sourceforge.net/.

[3] D. C. Koboldt et al. "VarScan 2: somatic mutation and copy number alteration discovery in cancer by exome sequencing". In: *Genome Res.* 22.3 (Mar. 2012), pp. 568–576. URL: http://varscan.sourceforge.net.