

IQMMA: installation guide

IQMMA is a python program which combines several other programs while working. Among them are three feature detections (Dinosaur, Biosaur2, OpenMS FeatureFinderCentroided) and Diffacto for quantitation analysis. It is required to install all of them to fully use all the advantages of IQMMA. This installation guide is separated into three parts.

- IQMMA basic package (IQMMA, Diffacto, Biosaur2)
- Dinosaur
- OpenMS FeatureFinderCentroided

Linux (ubuntu/Debian)

It is highly recommended (however, not necessary) to use some kind of virtual environment to install all needed packages with needed versions.

Here python virtualenv package is used:

Package installation:

```
$ pip install virtualenv
```

Virtual environment creation:

The command creates a new folder named <virtual-environment-name> in the directory where it is executed.

```
$ python<version> -m venv <virtual-environment-name>
```

Then to activate virtual environment:

```
$ source ./<virtual-environment-name>/bin/activate
```

To deactivate virtual environment:

```
$ deactivate
```

IQMMA basic package

Using pip:

This will install IQMMA and Biosaur2, Diffacto as requirements.

```
$ pip install iqmma
```

After successful installation the help command should work:

```
$ iqmma --help
```

Dinosaur

To run Dinosaur Java 1.8 or greater is needed (For Dinosaur versions up to 1.1.4, Java 1.6 or higher is needed). It has a good installation guide on its github page (<https://github.com/fickludd/dinosaur>).

Firstly, download Dinosaur released .jar file from its github:

<https://github.com/fickludd/dinosaur/releases>.

Next, installing java:

For ubuntu/Debian the command is (8 is replaceable with any other version higher than 8):

```
$ sudo apt install openjdk-8-jre-headless
```

After successful installation the help command should work:

```
$ java -jar /PATH/TO/Dinosaur-1.2.0.free_1.jar --help
```

OpenMS FeatureFinderCentroided

To install OpenMS FeatureFinderCentroided the easiest option is to install the pre-made OpenMS command-line tools package.

There are other options and paths that are described in the OpenMS installation guide (<https://openms.readthedocs.io/en/latest/openms-applications-and-tools/installation.html>).

Going through the 3-rd section of that guide for Linux the deb-package is downloaded from the OpenMS archive

(<https://abibuilder.cs.uni-tuebingen.de/archive/openms/OpenMSInstaller/release/latest/>).

Then the gdebi package is installed (if it's not already installed).

```
$ sudo apt-get install gdebi
```

And finally gdebi installs OpenMS tools (including FeatureFinderCentroided into the environment)

```
$ sudo gdebi /PATH/TO/OpenMS.deb
```

After successful installation the help command should work:

```
$ FeatureFinderCentroided -help
```

Usage Tooltips

After installation to use all installed tools it is needed to specify paths to its run files.

On Linux:

- Diffacto
- dif /PATH/TO/VIRTUALENV/<virtual-environment-name>/bin/diffacto
- Biosaur2
- bio2 /PATH/TO/VIRTUALENV/<virtual-environment-name>/bin/biosaur2
- Dinosaur
- dino /PATH/TO/Dinosaur-1.2.0.free_1.jar
- path to the downloaded .jar file
- OpenMS
- openms /PATH/TO/VIRTUALENV/<virtual-environment-name>/bin/FeatureFinderCentroided

Windows (Anaconda)

First of all, it is needed to activate your working environment:

```
(base) PS C:\Users\User> conda activate install_iqmma
(install_iqmma) PS C:\Users\User> █
```

Creation and useful commands to manage your environments properly are here (<https://conda.io/projects/conda/en/latest/user-guide/tasks/manage-environments.html>).

IQMMA basic package

Using pip:

This will install IQMMA and Biosaur2, Diffacto as requirements.

\$ pip install iqmma

```
(install_iqmma) PS C:\Users\User> pip install iqmma
Collecting iqmma
  Using cached iqmma-0.1.3.2-py3-none-any.whl (29 kB)
Collecting lxml
  Using cached lxml-4.9.2-cp310-cp310-win_amd64.whl (3.8 MB)
Collecting biosaur2
  Using cached biosaur2-0.2.11-cp310-cp310-win_amd64.whl (119 kB)
Collecting numpy==1.23.1
  Using cached numpy-1.23.1-cp310-cp310-win_amd64.whl (14.6 MB)
Collecting pyteomics
  Downloading pyteomics-4.6-py2.py3-none-any.whl (235 kB)
  235.1/235.1 kB 2.9 MB/s eta 0:00:00
```

After collecting all needed packages it will successfully installed:

```
Successfully installed biosaur2-0.2.11 contourpy-1.0.7 cycler-0.11.0 cython-0.29.34 diffacto-1.0.6 fonttools-4.39.4 iqmma-0.1.3.2 joblib-1.2.0 kiwisolver-1.4.4 lxml-4.9.2 matplotlib-3.7.1 networkx-3.1 numpy-1.23.1 packaging-23.1 pandas-2.0.1 pillow-9.5.0 pyparsing-3.0.9 pyteomics-4.6 python-dateutil-2.8.2 pytz-2023.3 scikit-learn-1.2.2 scipy-1.10.1 six-1.16.0 threadpoolctl-3.1.0 tzdata-2023.3 venn-0.1.3
(install_iqmma) PS C:\Users\User> █
```

After successful installation the help command should work:

\$ iqmma --help

Dinosaur

To run Dinosaur Java 1.8 or greater is needed (For Dinosaur versions up to 1.1.4, Java 1.6 or higher is needed). It has a good installation guide on its github page (<https://github.com/fickludd/dinosaur>).

Firstly, download Dinosaur released .jar file from its github:

<https://github.com/fickludd/dinosaur/releases>.

Next, installing java. For Anaconda this command is enough:

\$ conda install -c conda-forge openjdk

```
(install_iqmma) PS C:\Users\User> conda install -c conda-forge openjdk
```

After successful installation the help command should work:

\$ java -jar /PATH/TO/Dinosaur-1.2.0.free_1.jar --help

```
(install_iqmma) PS C:\Users\User> java -jar C:\Users\User\Desktop\env_openms_installation\env_openms_installation\Dinosaur-1.2.0.free_1.jar --help
usage:
> java -jar Dinosaur-1.2.0.jar [OPTIONS] mzML
OPTIONS:
    PARAMETER  DEFAULT  DESCRIPTION
    advHelp    false   set to output adv param file help and quit
    advParams   path to adv param file
    concurrency 2       the number of assays to analyze in parallel
    force      false   ignore missing mzML params
    maxCharge  6       max searched ion charge
    minCharge  1       min searched ion charge
    mode       global  analysis mode: global or target. Global mode reports all isotope patterns, targeted only those matching targets.
    mzML       -       The shotgun MzML file to analyze
    nReport    10      number of random assay to export control figure for
    outDir     output directory (by default same as input mzML)
    outName    basename for output files (by default same as input mzML)
    profiling  false   set to enable CPU profiling
    reportDeisoMzHeight 15.0  mz range in deisotoper reports
    reportHighRes false   generate high-resolution plot trail when supported (for print)
    reportSeed -1       seed to use for report assay selection (<0 means random)
    reportTargets false  set to create a special report figure for each target
    seed       -1       seed to use for bootstrapping of mass calibration (<0 means random)
    targetPreference rt   if multiple isotope patterns fit target, take the closest rt apex (rt) or the most intense (intensity)
    targets    path to isotope patterns target file (not used by default)
    verbose    false   increase details in output
    writeBinary false   set to output binary MSFeatureProtocol file
    writeHills false   set to output csv file with all hills assigned to isotope patterns
    writeMsInspect false  set to output MsInspect feature csv file
    writeQuantML false   set to output m2QuantML file
    zipQcFolder false   set to zip the entire qc folder on algorithm completion

Error parsing 'help'. Option does not exist.
Not enough arguments!
(install_iqmma) PS C:\Users\User>
```

OpenMS FeatureFinderCentroided

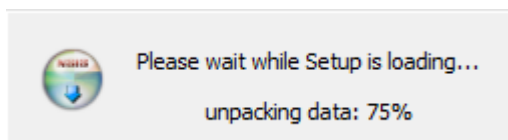
To install OpenMS FeatureFinderCentroided the easiest option is to install the pre-made OpenMS command-line tools package.

There are other options and paths that are described in the OpenMS installation guide (<https://openms.readthedocs.io/en/latest/openms-applications-and-tools/installation.html>).

Going through the 1-st section of that guide for Windows the installer is downloaded from the OpenMS archive

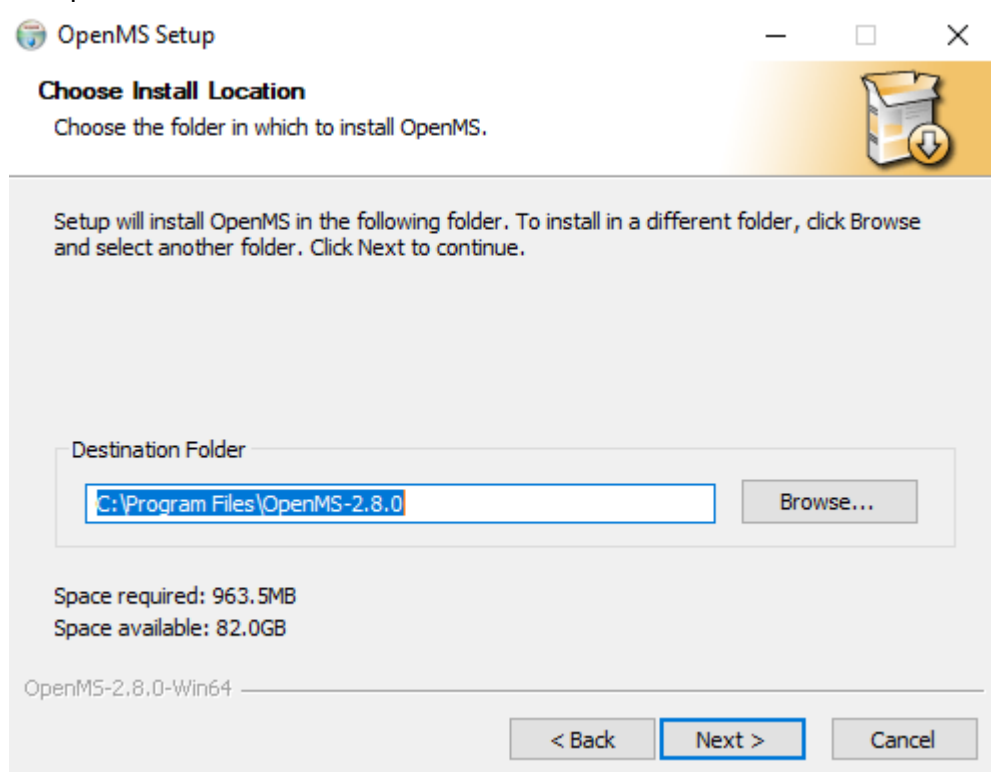
(<https://abibuilder.cs.uni-tuebingen.de/archive/openms/OpenMSInstaller/release/latest/>).

Then run the installer and accept administrator run.

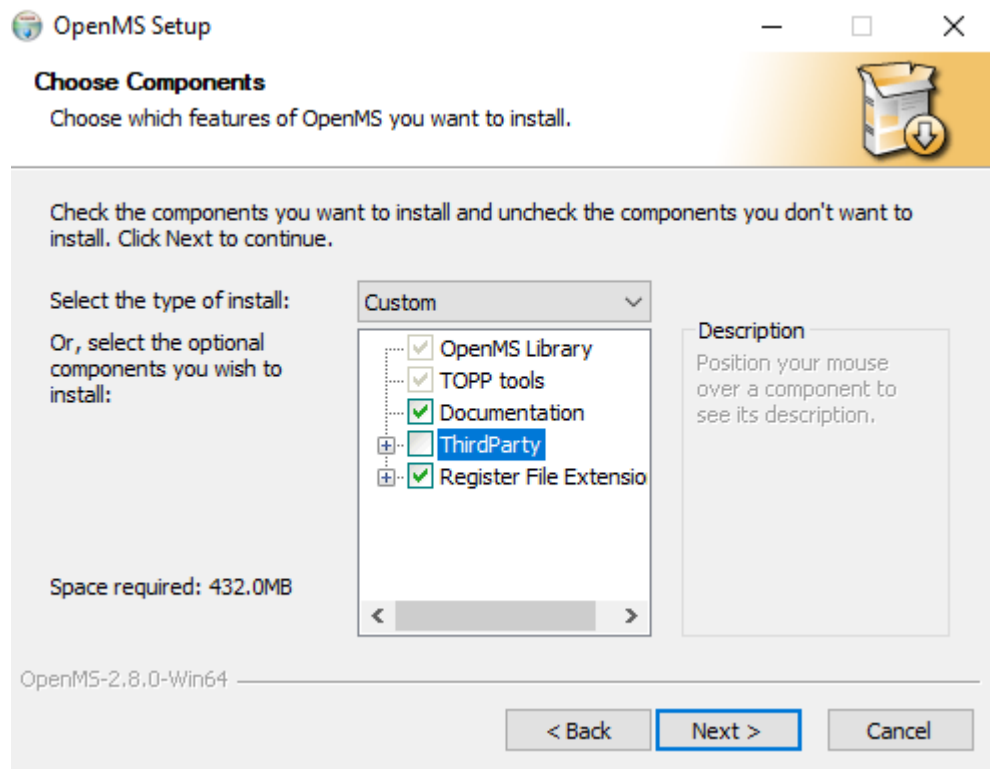




Accept the license and choose the location:



Uncheck ThirdParty since those tools are not used in IQMMA. (However, a lot of them are truly useful in proteomics data analysis.



Create or not the start menu shortcuts and finally install.

After successful installation the help command should work (version may change):

\$ /PATH/TO/OpenMS-2.8.0/bin/FeatureFinderCentroided --help

```
(install_iqmma) PS C:\Users\User> D:\Lerost\Lab_data\OpenMS-2.8.0\bin\FeatureFinderCentroided
--help

FeatureFinderCentroided -- Detects two-dimensional features in LC-MS data.
Full documentation: http://www.openms.de/doxygen/release/2.8.0/html/TOPP\_FeatureFinderCentroided.html
Version: 2.8.0 Feb 22 2022, 12:20:27, Revision: d203985
To cite OpenMS:
  Rost HL, Sachsenberg T, Aiche S, Bielow C et al.. OpenMS: a flexible open-source software p
  latform for mass spectrometry data analysis. Nat Meth. 2016; 13, 9: 741-748. doi:10.1038/nmet
  h.3959.
To cite FeatureFinderCentroided:
  Sturm M. A novel feature detection algorithm for centroided data. Dissertation, 2010-09-15,
  p.37 ff. doi:https://publikationen.uni-tuebingen.de/xmlui/bitstream/handle/10900/49453/pdf/Dissertation\_Marc\_Sturm.pdf.

Usage:
  FeatureFinderCentroided <options>

This tool has algorithm parameters that are not shown here! Please check the
ini file for a detailed description or use the --helphelp option.

Options (mandatory options marked with '*'):
  -in <file>*      Input file (valid formats: 'mzML')
  -out <file>*     Output file (valid formats: 'featureXML')
  -seeds <file>    User specified seed list (valid formats: 'featureXML')

Common TOPP options:
  -ini <file>      Use the given TOPP INI file
  -threads <n>     Sets the number of threads allowed to be used by the TOPP
                  tool (default: '1')
  -write_ini <file> Writes the default configuration file
  --help          Shows options
  --helphelp      Shows all options (including advanced)

The following configuration subsections are valid:
  - algorithm     Algorithm section

You can write an example INI file using the '-write_ini' option.
Documentation of subsection parameters can be found in the doxygen documentatio
n or the INIFileEditor.
For more information, please consult the online documentation for this tool:
  - http://www.openms.de/doxygen/release/2.8.0/html/TOPP\_FeatureFinderCentroided.html
```

Usage Tooltips

After installation to use all installed tools it is needed to specify paths to its run files.

In Anaconda on Windows:

- Diffacto

```
-dif %PATH%\TO\Anaconda3\envs\<env_name>\Scripts\diffacto
```

where <env_name> is working environment name install_iqmma in the installation or

```
-dif %PATH%\TO\Anaconda3\Scripts\diffacto
```

if the installation was done in the base environment.

- Biosaur2

```
-bio2 %PATH%\TO\Anaconda3\envs\<env_name>\Scripts\biosaur2
```

where <env_name> is working environment name install_iqmma in the installation or

```
-bio2 %PATH%\TO\Anaconda3\Scripts\biosaur2
```

if the installation was done in the base environment.

- Dinosaur

-dino /PATH/TO/Dinosaur-1.2.0.free_1.jar

path to the downloaded .jar file

- OpenMS

-openms /PATH/TO/OpenMS-2.8.0/bin/FeatureFinderCentroided

path to a certain tool in the OpenMS installation folder (:C\Program Files\ in the installation example).

Troubleshooting

Updating pip

Sometimes (for example on clean ubuntu subsystem for windows) it is needed to install pip and update it (<python> - is python version that is used):

```
$ sudo apt install <python>-pip
```

```
$ pip install --upgrade pip
```

Updating conda

Sometimes with old versions of conda incompatibility issues may occur. In that case, creating the environment with up-to-date python and conda versions may help. Here is the official guide on how to do it:

(<https://conda.io/projects/conda/en/latest/user-guide/tasks/manage-environments.html>).

However, in some cases “Solving environment” problem occurs. Then updating or even reinstalling conda (or at least) is a good way to solve it.

Some useful links:

<https://www.anaconda.com/blog/a-faster-conda-for-a-growing-community>

<https://stackoverflow.com/questions/63734508/stuck-at-solving-environment-on-anaconda>

<https://github.com/conda/conda/issues/11919>

Using IQMMA from the jupyter-notebook

In some cases, when IQMMA was installed in the virtual environment, and then is called from the notebook outside of the environment (which is not the proper way of using) errors like “ModuleNotFoundError: No module named 'ipython'” (or matplotlib.inline) may occur while plotting Venn diagram.

To solve this simply install needed modules in your environment. For example with pip:

```
$ pip install ipython
```