



MZmine 2 manual

About MZmine 2

Copyright (c) 2005-2015 MZmine Development Team

MZmine 2 is an open-source framework for processing, visualization and analysis of mass spectrometry based molecular profile data. It is based on the original MZmine toolbox described in 2006 *Bioinformatics* publication.

References

- Pluskal, T., Castillo, S., Villar-Briones, A., Orešič, M. (2010). MZmine 2: Modular framework for processing, visualizing, and analyzing mass spectrometry-based molecular profile data, *BMC Bioinformatics* **11**:395
- Katajamaa, M., Miettinen, J., and Orešič, M. (2006). MZmine: Toolbox for processing and visualization of mass spectrometry based molecular profile data, *Bioinformatics* **22**, 634-636

License

This program is free software; you can redistribute it and/or modify it under the terms of the GNU General Public License as published by the Free Software Foundation; either version 2 of the License, or (at your option) any later version.

This program is distributed in the hope that it will be useful, but WITHOUT ANY WARRANTY; without even the implied warranty of MERCHANTABILITY or FITNESS FOR A PARTICULAR PURPOSE. See the GNU General Public License for more details.

You should have received a copy of the GNU General Public License along with this program; if not, write to the Free Software Foundation, Inc., 51 Franklin Street, Fifth Floor, Boston, MA 02110-1301, USA.

Table of contents

[About MZmine 2](#)

[License](#)

[Table of contents](#)

[1. How to report problems with MZmine 2](#)

[2. Installation](#)

[2.1. System requirements](#)

[2.2. Installation](#)

[3. Starting MZmine 2](#)

[3.1. Starting in GUI mode](#)

[3.2. Starting in batch mode](#)

[4. MZmine 2 desktop](#)

[4.1. Number formatting](#)

[5. Data processing](#)

[5.1. Raw data import](#)

[5.2. Setting sample parameters](#)

[5.3. Data processing workflow](#)

[5.4. Batch mode](#)

[6. Documentation for individual modules](#)

1. How to report problems with MZmine 2

We appreciate any feedback from the users of MZmine 2. Please do let us know your suggestions regarding the design of the software. Especially, please inform us about any problems or bugs you might encounter while using the software.

In order to fix the problems reported by users, we need to obtain sufficient information to reproduce each problem. Therefore, please include the following details with your reports:

1. Your MZmine version.
2. Your OS and its version (32-bit or 64-bit?).
3. The exact output of the following two commands:

```
java -version  
R --version
```
4. Let us know if you changed any settings in your startMZmine script file.
5. The exact steps you took until you encountered the problem. Please list all the relevant parameter values you set for each data processing step. Preferably, you can send us your raw data files or your MZmine project file so we can reproduce your steps (do not send large files by email, use a service such as [DropBox](#)).
6. Screenshots of all error messages or unexpected output. Also, please include the log messages reported in the MZmine console window. If you feel that the results provided by MZmine were wrong, please describe exactly why you think they were wrong and how they differ from the results you expected.

Please file your problem report at <https://github.com/mzmine/mzmine2/issues>. We shall respond promptly.

2. Installation

2.1. System requirements

Minimum requirements

- 2 GHz CPU
- 1 GB RAM

Recommended requirements

- 64-bit CPU + 64-bit OS + 64-bit Java runtime
- 4 or more CPU cores
- 4 GB or more RAM

2.2. Installation

Prerequisites

1. Install Java SE runtime (JRE) version 1.7 or later (available from <http://java.oracle.com>)
2. *(Optional)* Install R (<http://http://www.r-project.org>). R is required for certain modules such as baseline correction or heat maps. Several packages must be installed in your R distribution. MZmine will try to install them automatically, but this process may fail. To install them manually, run the following commands in R:

```
install.packages(c("Rserve", "ptw", "gplots", "baseline", "hyperSpec",  
"ggplot2"))  
source("http://bioconductor.org/biocLite.R")  
biocLite("xcms")  
biocLite("CAMERA")  
biocLite("PROcess")
```

MZmine 2 installation

1. Download the MZmine distribution (ZIP file) and unpack it into a folder of your choice.
2. That's all.

3. Starting MZmine 2

3.1. Starting in GUI mode

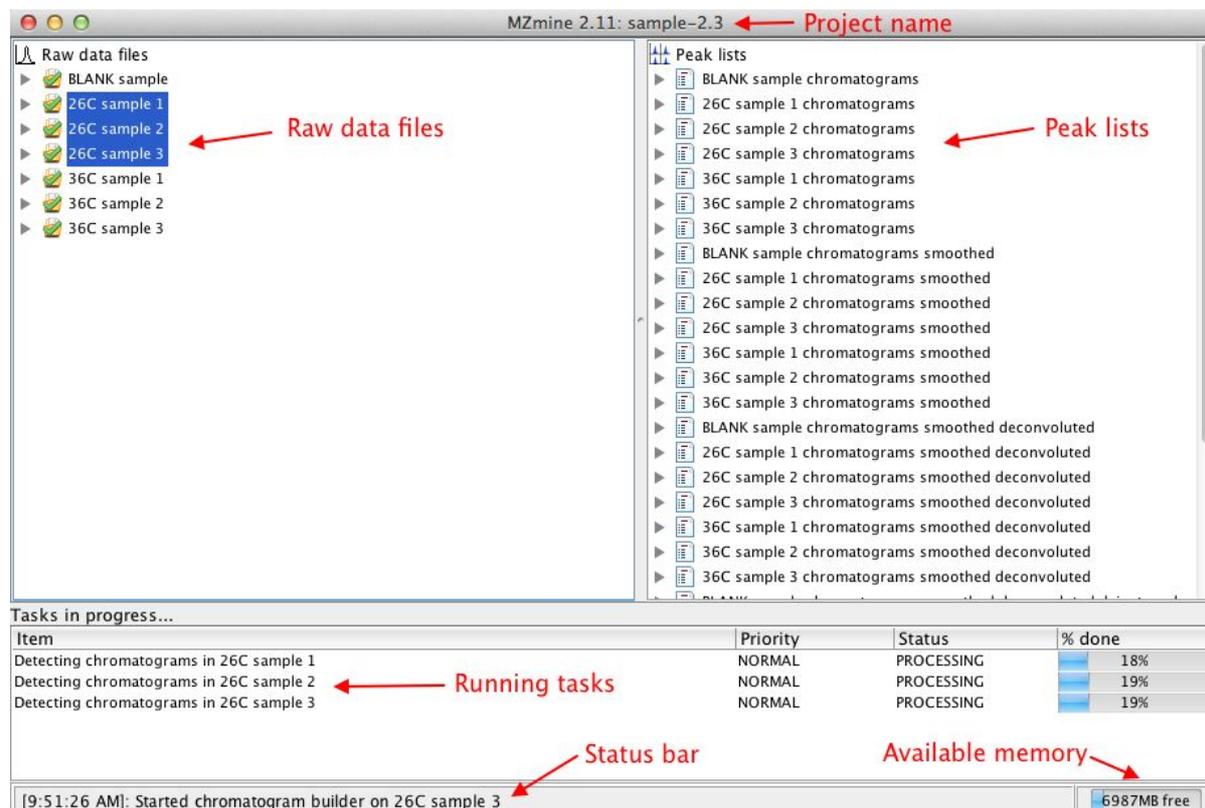
A startup script is provided for MS Windows (*startMZmine_Windows.bat*), Mac OS X (*startMZmine_MacOSX.command*) and Linux (*startMZmine_Linux.sh*) environments. Start MZmine in the GUI mode by running this script. Please note that the amount of memory used by the program is limited by the value of the `HEAP_SIZE` parameter in the startup script. The default value is set to half of your physical RAM. If you wish to use more RAM for MZmine 2, please edit the startup script.

3.2. Starting in batch mode

If the *startMZmine...* script is supplied with a command line parameter, this parameter is treated as a batch queue file (see [Batch mode](#)). In such case MZmine is started as a command-line application and the batch file is loaded into the batch mode module and executed. Upon finishing the whole batch, MZmine will automatically quit. This allows to integrate MZmine into scriptable work flows. Note that the first step of the batch should be always either the “Raw data import” module or the “Project load” module, otherwise MZmine will have no data to process.

4. MZmine 2 desktop

The following screenshot describes the components of the main MZmine 2 window:

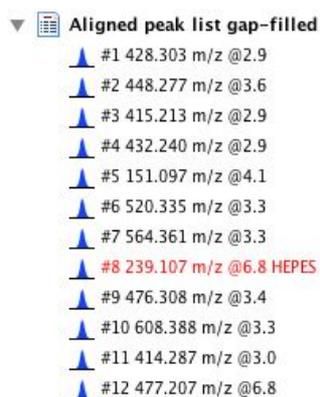


The order of the raw data files and peak lists in the project tree may be changed using mouse (drag & drop).

If you expand the raw data files in the left section of the main window, you will see a list of scans in the file. MS scans are shown in blue, while MS/MS scans are shown in red. Each scan is annotated with its sequential number (#), retention time in minutes (@), MS level (for MS/MS scans the precursor m/z is also shown), type of spectrum (p = profile, c = centroid, t = thresholded), and polarity of ionization (+ = positive, - = negative, ? = unknown).

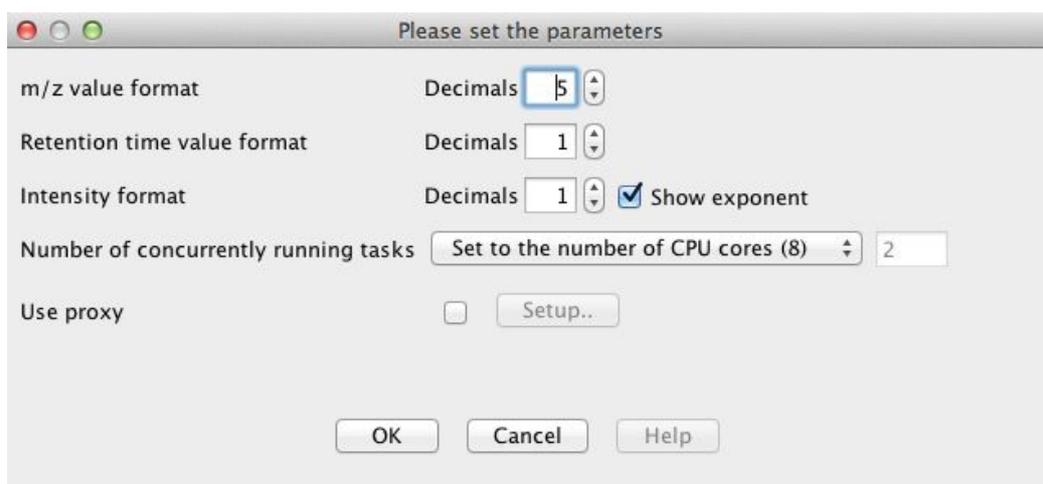
```
▼ 090112_B_VE_26_neg_06.mzXML
  #1 @0.0 MS1 p -
  #2 @0.0 MS1 p -
  #3 @0.1 MS2 (154.948) p -
  #4 @0.1 MS1 p -
  #5 @0.1 MS2 (258.921) p -
  #6 @0.1 MS1 p -
  #7 @0.1 MS2 (219.175) p -
  #8 @0.2 MS1 p -
  #9 @0.2 MS2 (228.932) p -
```

Similarly, expanding the peak list in the right section of the project window will show its individual rows with their ID (#), m/z value and retention time (@). Rows that have identity information attached are displayed in red and their name is also shown.



4.1. Number formatting

Formatting of the m/z, retention time and intensity values may be modified in the dialog activated by the Project / Set preferences menu item. Note that the retention time is always represented in minutes.



5. Data processing

5.1. Raw data import

MZmine 2 can read and process both low-resolution (unit mass) and high-resolution (exact mass) data in both continuous and centroided modes, including fragmentation (MS^n) scans. Supported data formats are:

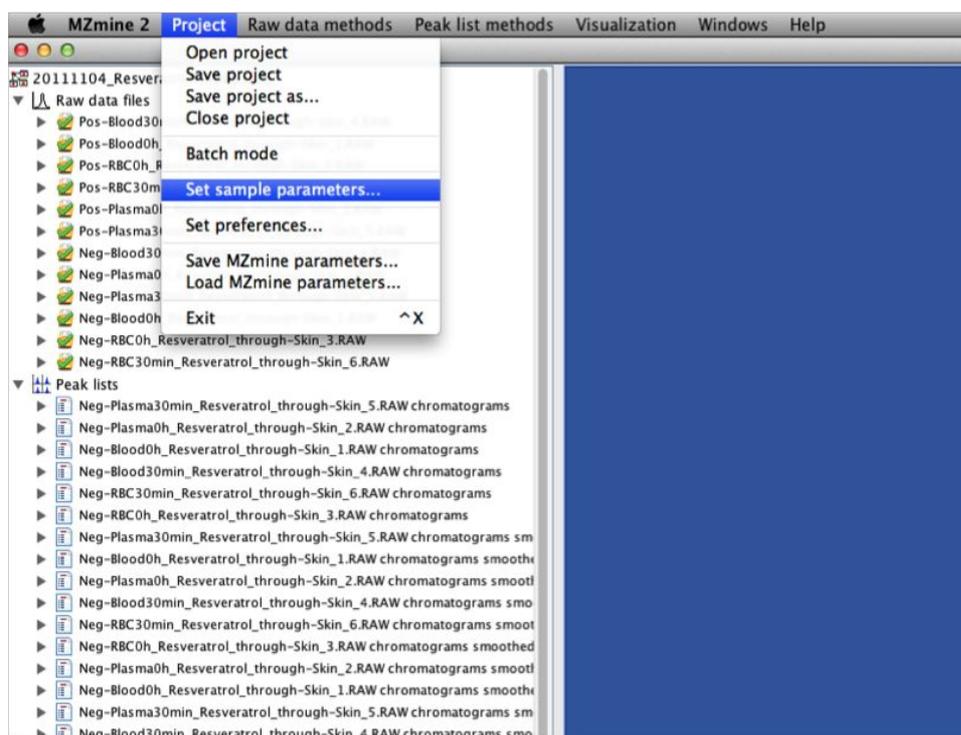
- **mzML** (mzML version 1.0 and 1.1)
- **mzXML** (mzXML versions 2.0, 2.1 and 3.0)
- **mzData** (mzData versions 1.04 and 1.05)
- **NetCDF** (no MS^n data support)
- **Thermo RAW** (native data format)
- **Waters RAW** (native data format)

The support for native data formats is implemented using a small external program, which runs only on Windows. On Linux or other platforms, MZmine will attempt to execute the external program through the Windows emulator Wine (<http://www.winehq.org>), so you can try to install it.

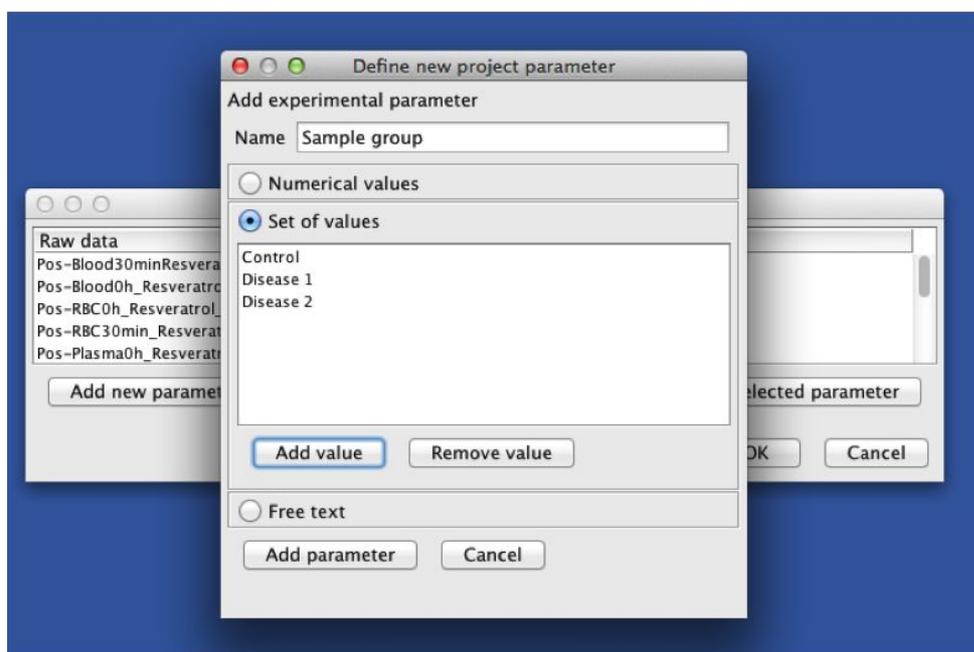
5.2. Setting sample parameters

Parameters can be added to the MZmine project and their values can be set for each sample individually. These parameters can later be used in further stages of data processing, for example in the Intensity plot. The addition of the parameters to the project is done in the following steps:

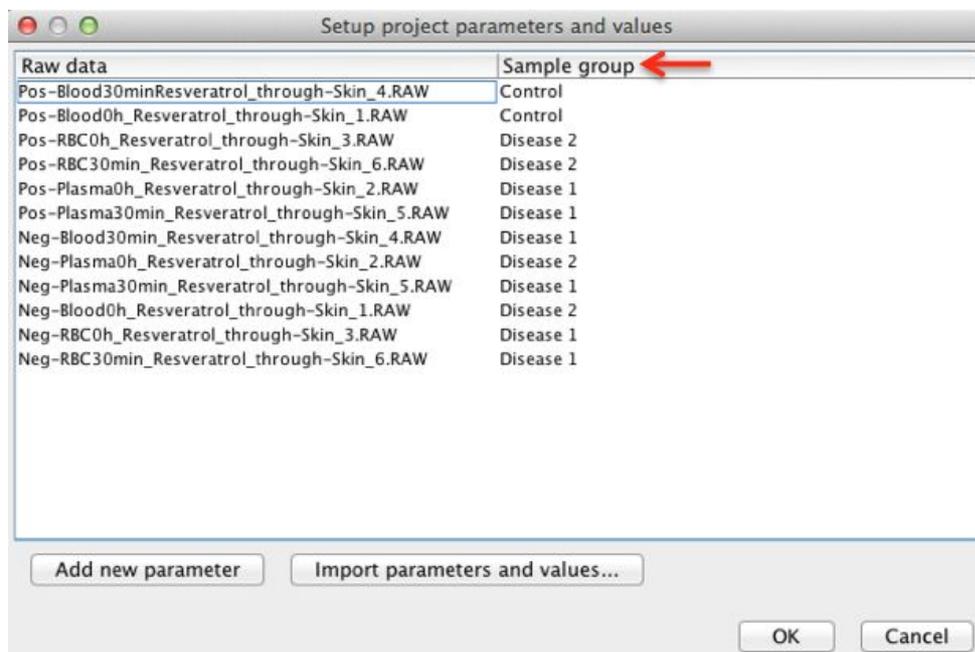
1. Open the Set sample parameters dialog



2. Add a new parameter. Each parameter can have numerical values, text values, or a set of defined values (options).



3. Set the parameter value to each sample



The parameters and values can also be imported from a comma-separated (CSV) file. The format of the input CSV file is as described below.

First column of the comma-separated file must contain file names matching to names of raw data files opened in MZmine. Each other column corresponds to one project parameter.

First row in the file must contain column headers. Header for the first column (filename) is ignored but must exist. Rest of the column headers are used as names for project parameters. All column names in the file must be unique. If main dialog already contains a parameter with the same name, a warning is shown to the user before overwriting previous parameter. The type of the imported parameter is decided according to the following rules:

1. If all values in a column (except column header on the first row) are numeric, then the parameter type is set to numerical.
2. If there are at least some duplicate strings among values for a parameter, then the parameter type is "Set of values" and possible parameter values are all unique strings.
3. Otherwise it is a free text parameter.

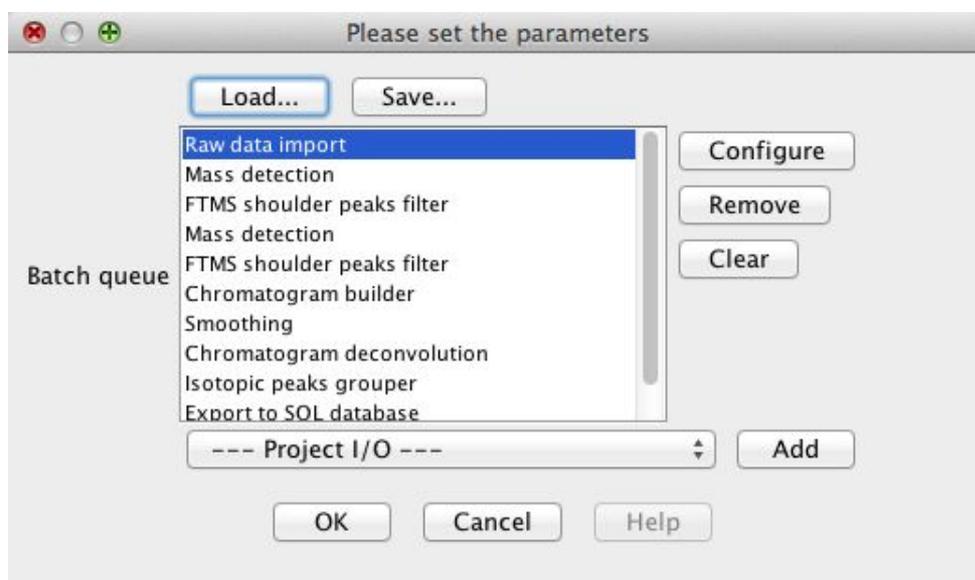
5.3. Data processing workflow

The typical workflow for processing mass spectrometry data using MZmine 2 consists of the following steps (note that some of these steps are optional and may be skipped). The bold font indicates the name of the corresponding item in the MZmine application menu.

1. **Raw data import**
2. *(Optional)* **Raw data methods / Filtering**
3. **Peak detection**, using one of the available methods
 - a. Generation of mass lists (detected ions) for each scan using **Mass detection**, followed by **FTMS shoulder peaks filter** *(optional)*, followed by detection of chromatograms using the **Chromatogram builder**, followed by **Smoothing** *(optional)*, followed by separation of individual peaks in the chromatograms using the **Chromatogram deconvolution** module
 - b. 2D peak detection using the **GridMass** module
 - c. Peak list generation from precursor ions of MS/MS scans using the **MS/MS peaklist builder** module, followed by the **Peak extender** module
 - d. **Targeted peak detection** from a given list of target ions
4. Removing of isotopes using the **Isotopic peak grouper** module
5. *(Optional)* **Identification** of fragments, adducts and peak complexes
6. *(Optional)* **Normalization** of retention time using the **Retention time normalizer**
7. **Alignment** using the **Join aligner** or **RANSAC aligner**
8. *(Optional)* **Gap filling** using the **Peak finder** or **Same RT and m/z range gap filler**
9. *(Optional)* **Normalization** of peak heights/areas using the **Linear normalizer** or **Standard compound normalizer**
10. *(Optional)* **Identification** using **Custom database search**, **Online database search**, **Formula prediction**, etc.
11. **Data analysis, Export, Visualization** etc..

5.4. Batch mode

MZmine 2 has the option to run multiple data processing methods in batch mode. The following figure shows the batch mode setup dialog, where the user can setup the individual batch steps. At the moment of adding a new step, the corresponding parameter setup dialog appears. It is also possible to change the steps or the parameters of each step with "Configure" and "Remove" buttons. The contents of the batch queue can be saved to a XML file and later restored by clicking the "Load.." button.



6. Documentation for individual modules

Please refer to MZmine's internal Help for the documentation of individual modules and algorithms. Each parameter setup dialog has a Help button as shown below:

