

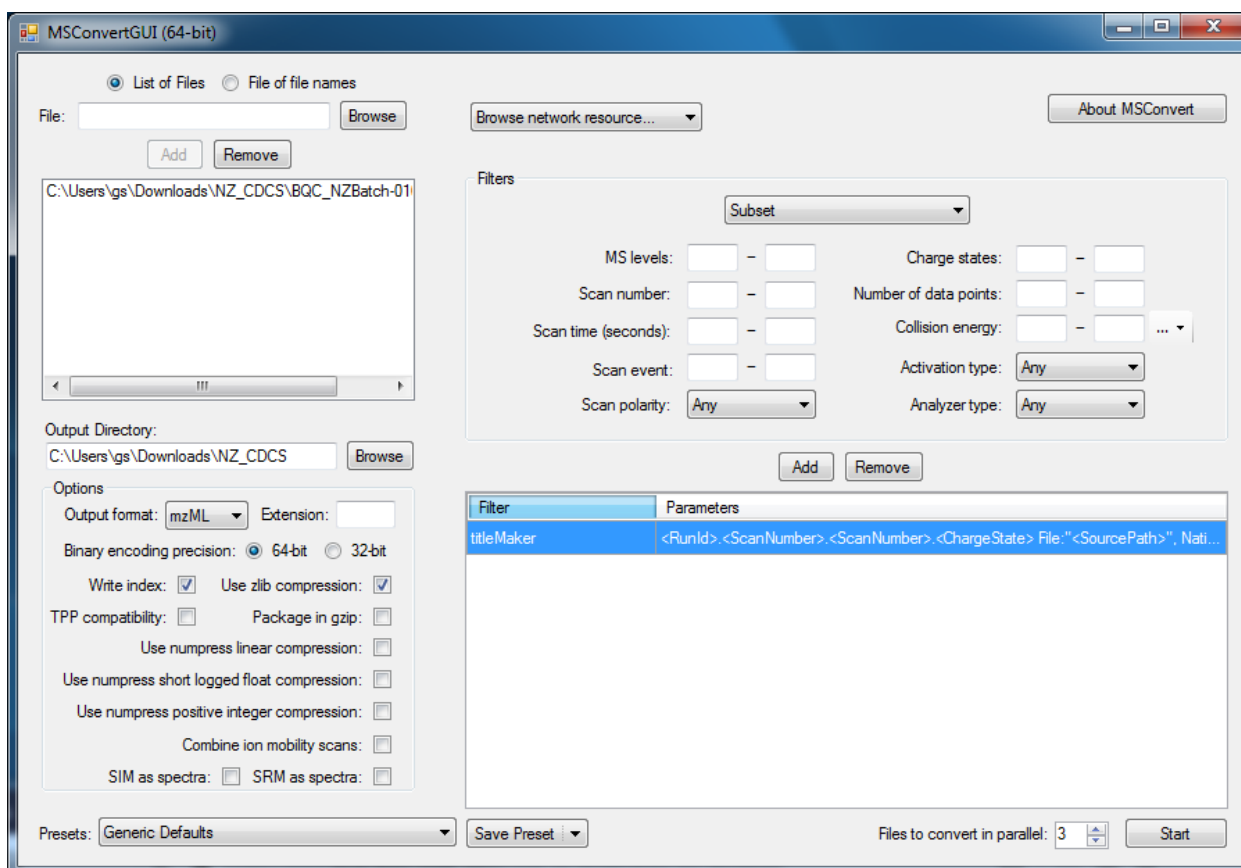
# MRMkit

July 9, 2020

## 1 Raw Data Conversion into mzML

The input files need to be in mzML format for MRMkit. MSConvert, provided through the ProteoWizard software suite, enables conversion of proprietary raw data files (.wiff for SCIEX, .raw files for Thermo Fisher, .d folders for Agilent, etc) into the mzML format.

1. Download and install a recent version ProteoWizard on a Windows computer (the versions for other OS platforms do not work).
2. Start the software by opening the Start Menu, type “MSConvert” in the search field and click on “MSConvert”.
3. In the MSConvert window, use the “Browse” button to select the raw files to be converted (the .wiff and .wiff.scan files need to be located in the same folder). Select “mzML” for “Output format”, “64-bit” for “Binary encoding precision”, check “Write index” and “Use zlib compression” checkbox (see figure below).



## 2 Input Parameters

Input parameters must be specified according to the user's preference in order to control various aspects of the data processing. Input parameters are to be specified in the "param.txt" file. MRMkit requires that the user provide the following options for data processing:

- **mzML\_files**: Names of the mzML files. Use "\*.mzML" to select all mzML files in the current working directory.
- **batch\_info**: analytical sequence of samples with indication of metabolite extraction / sample preparation batches. An example shown in Figure 1.
- **length\_of\_ion\_chromatogram**: range of ion chromatogram to be detected in each MRM window (in seconds)
- **transition\_list**: assay information with compound name, Q1, Q3, retention time and the name of the designated internal standards (ISTD). An example shown in Figure 2.
- **ISTD\_trace\_all**: Set "1" to use the peak shape of the corresponding ISTD for peak integration for all transitions, "0" to directly take the area under peak for transitions

with unimodal chromatograms. Peak areas will still be extracted based on the peak shape of ISTD for multi-modal ion chromatograms

- **batch\_correction**: Set “1” to enable time trend and batch effect correction, “0” to disable.

### 3 Analysis Output Table

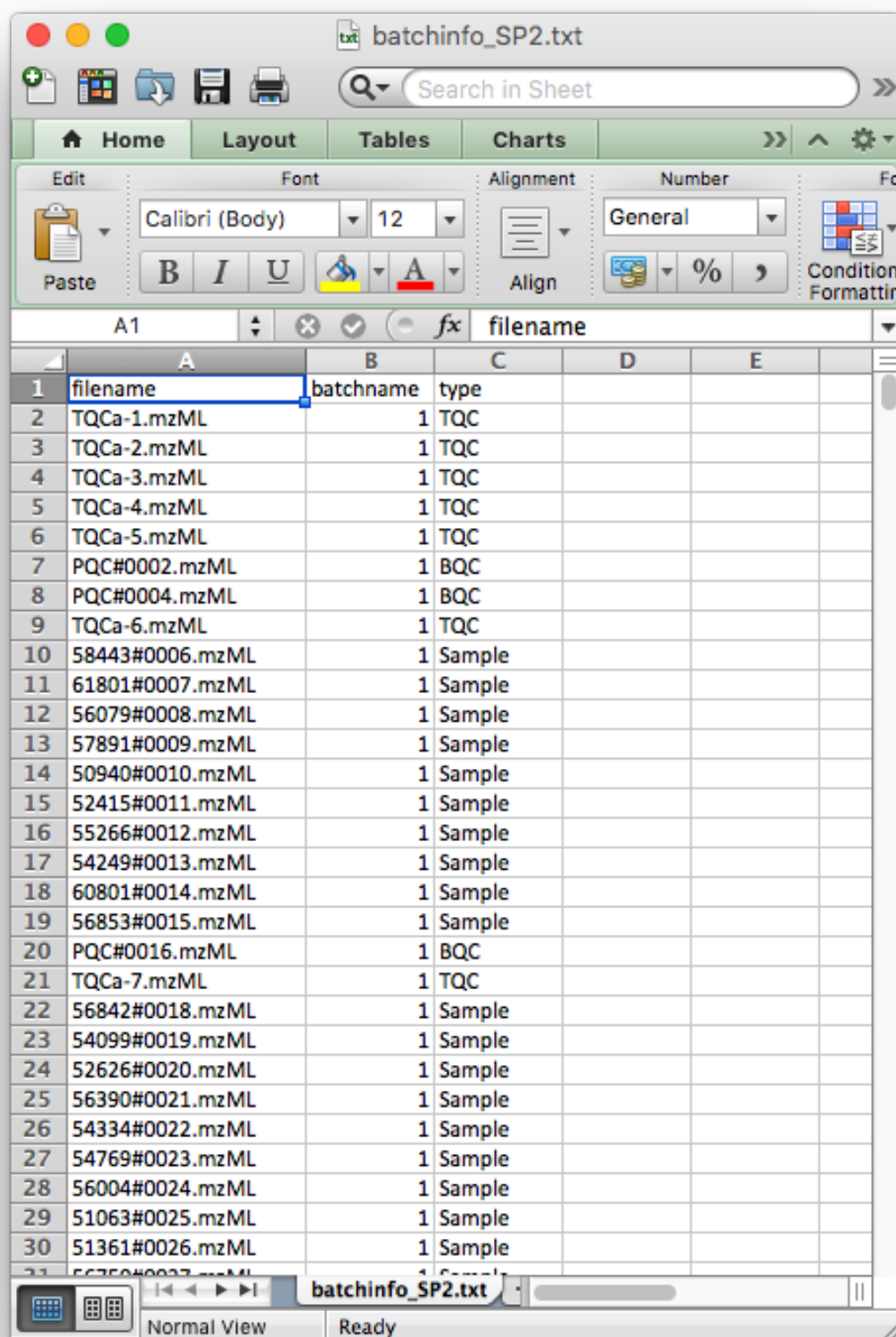
- **batch\_adjusted.txt**: Quantitative data (compound peak area / ISTD) after correction of time trends and batch effects
- **quant\_auc\_rt.txt**: Integrated peak areas for target analytes and ISTDs
- **BQCtable.txt**: Batch specific and overall coefficient of variation using batch pooled QC samples (a pooled sample with metabolite extraction performed along with the samples from a particular batch)
- **TQCtable.txt**: Batch specific and overall coefficient of variation using technical QC samples (a pooled sample formed at the beginning of the entire analytical sequence, intended to capture instrument variation only)
- **run\_seq.pdf**: Visualization of concentration data before and after time trend and batch effect correction, per compound basis (produced from R script)
- **IonC\_{transition\_name}.pdf**: Ion Chromatogram images with integrated peaks across all samples, per transition
- **users\_RT.txt**: The RT selection table, refer to next section.
- **quant\_table.txt**: The final quantification table after the user indicates the desired peak in “users\_RT.txt”.

### 4 User’s RT selection step

Upon running the analysis, MRMkit produces the aforementioned output files, but it also reports a RT selection table. This step is necessary: for transitions with a multi-modal ion chromatogram, the user must specify which peak he/she wishes to use to quantify the target compound. To facilitate the process, the template file shows the TQC/BQC-based coefficients of variation and their respective RTs. In addition, the user can refer to the ion chromatogram image files to see which peak corresponds to the “real” peak, based on the knowledge of the assay. See the template in Figure 3:

Once the user fills up this form, run the following to produce the final quantification table.

Figure 1: batch\_info file



The screenshot shows a spreadsheet application window titled "batchinfo\_SP2.txt". The spreadsheet has five columns labeled A, B, C, D, and E. The data is organized as follows:

|    | A               | B         | C      | D | E |
|----|-----------------|-----------|--------|---|---|
| 1  | filename        | batchname | type   |   |   |
| 2  | TQCa-1.mzML     | 1         | TQC    |   |   |
| 3  | TQCa-2.mzML     | 1         | TQC    |   |   |
| 4  | TQCa-3.mzML     | 1         | TQC    |   |   |
| 5  | TQCa-4.mzML     | 1         | TQC    |   |   |
| 6  | TQCa-5.mzML     | 1         | TQC    |   |   |
| 7  | PQC#0002.mzML   | 1         | BQC    |   |   |
| 8  | PQC#0004.mzML   | 1         | BQC    |   |   |
| 9  | TQCa-6.mzML     | 1         | TQC    |   |   |
| 10 | 58443#0006.mzML | 1         | Sample |   |   |
| 11 | 61801#0007.mzML | 1         | Sample |   |   |
| 12 | 56079#0008.mzML | 1         | Sample |   |   |
| 13 | 57891#0009.mzML | 1         | Sample |   |   |
| 14 | 50940#0010.mzML | 1         | Sample |   |   |
| 15 | 52415#0011.mzML | 1         | Sample |   |   |
| 16 | 55266#0012.mzML | 1         | Sample |   |   |
| 17 | 54249#0013.mzML | 1         | Sample |   |   |
| 18 | 60801#0014.mzML | 1         | Sample |   |   |
| 19 | 56853#0015.mzML | 1         | Sample |   |   |
| 20 | PQC#0016.mzML   | 1         | BQC    |   |   |
| 21 | TQCa-7.mzML     | 1         | TQC    |   |   |
| 22 | 56842#0018.mzML | 1         | Sample |   |   |
| 23 | 54099#0019.mzML | 1         | Sample |   |   |
| 24 | 52626#0020.mzML | 1         | Sample |   |   |
| 25 | 56390#0021.mzML | 1         | Sample |   |   |
| 26 | 54334#0022.mzML | 1         | Sample |   |   |
| 27 | 54769#0023.mzML | 1         | Sample |   |   |
| 28 | 56004#0024.mzML | 1         | Sample |   |   |
| 29 | 51063#0025.mzML | 1         | Sample |   |   |
| 30 | 51361#0026.mzML | 1         | Sample |   |   |
| 31 | 56350#0027.mzML | 1         | Sample |   |   |

Figure 2: transition\_list file

|    | A     | B                           | C                           | D             | E           | F      | G |
|----|-------|-----------------------------|-----------------------------|---------------|-------------|--------|---|
| 1  | class | Compound Name               | ISTD                        | Precursor Ion | Product Ion | RT     |   |
| 2  | Cer   | Cer d18:1/C8:0 (ISTD)       |                             | 426.4         | 264.2       | 196.74 |   |
| 3  | Cer   | Cer d18:1/C8:0 (ISTD)(-H2O) |                             | 408.4         | 264.2       | 196.74 |   |
| 4  | Cer   | Cer d16:1/C16:0             | Cer d18:1/C8:0 (ISTD)       | 510.5         | 236.2       | 266.22 |   |
| 5  | Cer   | Cer d16:1/C16:0(-H2O)       | Cer d18:1/C8:0 (ISTD)(-H2O) | 492.5         | 236.2       | 266.22 |   |
| 6  | Cer   | Cer d16:1/C18:0             | Cer d18:1/C8:0 (ISTD)       | 538.5         | 236.2       | 289.68 |   |
| 7  | Cer   | Cer d16:1/C18:0(-H2O)       | Cer d18:1/C8:0 (ISTD)(-H2O) | 520.5         | 236.2       | 289.68 |   |
| 8  | Cer   | Cer d16:1/C20:0             | Cer d18:1/C8:0 (ISTD)       | 566.5         | 236.2       | 310.14 |   |
| 9  | Cer   | Cer d16:1/C20:0(-H2O)       | Cer d18:1/C8:0 (ISTD)(-H2O) | 548.5         | 236.2       | 310.14 |   |
| 10 | Cer   | Cer d16:1/C22:0             | Cer d18:1/C8:0 (ISTD)       | 594.5         | 236.2       | 330.48 |   |
| 11 | Cer   | Cer d16:1/C22:0(-H2O)       | Cer d18:1/C8:0 (ISTD)(-H2O) | 576.5         | 236.2       | 330.48 |   |
| 12 | Cer   | Cer d16:1/C22:1             | Cer d18:1/C8:0 (ISTD)       | 592.5         | 236.2       | 317.4  |   |
| 13 | Cer   | Cer d16:1/C22:1(-H2O)       | Cer d18:1/C8:0 (ISTD)(-H2O) | 574.5         | 236.2       | 317.4  |   |
| 14 | Cer   | Cer d16:1/C23:0             | Cer d18:1/C8:0 (ISTD)       | 608.6         | 236.2       | 345.6  |   |
| 15 | Cer   | Cer d16:1/C23:0(-H2O)       | Cer d18:1/C8:0 (ISTD)(-H2O) | 590.6         | 236.2       | 345.6  |   |
| 16 | Cer   | Cer d16:1/C24:0             | Cer d18:1/C8:0 (ISTD)       | 622.6         | 236.2       | 352.8  |   |
| 17 | Cer   | Cer d16:1/C24:0(-H2O)       | Cer d18:1/C8:0 (ISTD)(-H2O) | 604.6         | 236.2       | 352.8  |   |
| 18 | Cer   | Cer d16:1/C24:1             | Cer d18:1/C8:0 (ISTD)       | 620.6         | 236.2       | 334.2  |   |
| 19 | Cer   | Cer d16:1/C24:1(-H2O)       | Cer d18:1/C8:0 (ISTD)(-H2O) | 602.6         | 236.2       | 334.2  |   |
| 20 | Cer   | Cer d16:1/C25:1             | Cer d18:1/C8:0 (ISTD)       | 634.6         | 236.2       | 345.42 |   |
| 21 | Cer   | Cer d16:1/C25:1(-H2O)       | Cer d18:1/C8:0 (ISTD)(-H2O) | 616.6         | 236.2       | 345.42 |   |
| 22 | Cer   | Cer d16:1/C26:0             | Cer d18:1/C8:0 (ISTD)       | 650.6         | 236.2       | 366.6  |   |
| 23 | Cer   | Cer d16:1/C26:0(-H2O)       | Cer d18:1/C8:0 (ISTD)(-H2O) | 632.6         | 236.2       | 366.9  |   |
| 24 | Cer   | Cer d16:1/C26:1             | Cer d18:1/C8:0 (ISTD)       | 648.6         | 236.2       | 351.6  |   |
| 25 | Cer   | Cer d16:1/C26:1(-H2O)       | Cer d18:1/C8:0 (ISTD)(-H2O) | 630.6         | 236.2       | 351.6  |   |
| 26 | Cer   | Cer d18:0/C16:0             | Cer d18:1/C8:0 (ISTD)       | 540.5         | 266.2       | 288.36 |   |
| 27 | Cer   | Cer d18:0/C16:0(-H2O)       | Cer d18:1/C8:0 (ISTD)(-H2O) | 522.5         | 266.2       | 288.36 |   |
| 28 | Cer   | Cer d18:0/C18:0             | Cer d18:1/C8:0 (ISTD)       | 568.5         | 266.2       | 316.32 |   |
| 29 | Cer   | Cer d18:0/C18:0(-H2O)       | Cer d18:1/C8:0 (ISTD)(-H2O) | 550.5         | 266.2       | 318.96 |   |
| 30 | Cer   | Cer d18:0/C20:0             | Cer d18:1/C8:0 (ISTD)       | 596.5         | 266.2       | 336    |   |
| 31 | Cer   | Cer d18:0/C20:0(-H2O)       | Cer d18:1/C8:0 (ISTD)(-H2O) | 578.5         | 266.2       | 336    |   |
| 32 | Cer   | Cer d18:0/C22:0             | Cer d18:1/C8:0 (ISTD)       | 624.6         | 266.2       | 352.86 |   |

Figure 3: **users\_RT** file, the user will fill in RT of true peak where multiple peaks are detected (highlighted in yellow)

|    | A                     | B     | C     | D                 | E                 | F          | G             | H       | I |
|----|-----------------------|-------|-------|-------------------|-------------------|------------|---------------|---------|---|
|    | name                  | Q1    | Q3    | %CoV(TQC)         | %CoV(BQC)         | expectedRT | detectedRT    | usersRT |   |
| 2  | Cer d16:1/C16:0       | 510.5 | 236.2 | 26.4, 56.7        | 28.6, 56.5        | 4.437      | 234, 270      | 270     |   |
| 3  | Cer d16:1/C16:0(-H2O) | 492.5 | 236.2 | 13.9, 15.9        | 15.2, 16.4        | 4.437      | 234, 271      | 271     |   |
| 4  | Cer d16:1/C18:0       | 538.5 | 236.2 |                   | 58                | 4.828      |               | 294     |   |
| 5  | Cer d16:1/C18:0(-H2O) | 520.5 | 236.2 | 15.6              | 15.3              | 4.828      |               | 294     |   |
| 6  | Cer d16:1/C20:0       | 566.5 | 236.2 | 38.6              | 40.1              | 5.169      |               | 315     |   |
| 7  | Cer d16:1/C20:0(-H2O) | 548.5 | 236.2 | 15.5              | 17.5              | 5.169      |               | 315     |   |
| 8  | Cer d16:1/C22:0       | 594.5 | 236.2 | 25.9              | 19.8              | 5.508      |               | 335     |   |
| 9  | Cer d16:1/C22:0(-H2O) | 576.5 | 236.2 | 17.4              | 15.3              | 5.508      |               | 335     |   |
| 10 | Cer d16:1/C22:1       | 592.5 | 236.2 | 260.2, 255.9      | 521.3, 557.4      | 5.29       | 305, 325      | 325     |   |
| 11 | Cer d16:1/C22:1(-H2O) | 574.5 | 236.2 | 284.5, 169.5      | 470.9, 68.3       | 5.29       | 303, 317      | 317     |   |
| 12 | Cer d16:1/C23:0       | 608.6 | 236.2 |                   | 24.2              | 5.76       |               | 344     |   |
| 13 | Cer d16:1/C23:0(-H2O) | 590.6 | 236.2 | 13.9              | 14.2              | 5.76       |               | 344     |   |
| 14 | Cer d16:1/C24:0       | 622.6 | 236.2 | 18.6              | 14.9              | 5.88       |               | 352     |   |
| 15 | Cer d16:1/C24:0(-H2O) | 604.6 | 236.2 | 11.7              | 13                | 5.88       |               | 353     |   |
| 16 | Cer d16:1/C24:1       | 620.6 | 236.2 | 32.8, 32.9        | 27.2, 36.8        | 5.57       | 335, 344      | 335     |   |
| 17 | Cer d16:1/C24:1(-H2O) | 602.6 | 236.2 | 16.9              | 15.2              | 5.57       |               | 335     |   |
| 18 | Cer d16:1/C25:1       | 634.6 | 236.2 | 332.6, 269.0      | 411.8, 248.1      | 5.757      | 343, 352      | 343     |   |
| 19 | Cer d16:1/C25:1(-H2O) | 616.6 | 236.2 | 26.5              | 27.3              | 5.757      |               | 344     |   |
| 20 | Cer d16:1/C26:0       | 650.6 | 236.2 | 550.8, 45.5, 95.8 | 677.2, 51.4, 85.2 | 6.11       | 354, 368, 375 | 368     |   |
| 21 | Cer d16:1/C26:0(-H2O) | 632.6 | 236.2 | 14.1              | 26.6              | 6.115      |               | 368     |   |
| 22 | Cer d16:1/C26:1       | 648.6 | 236.2 | 260.7, 33.4       | 421.3, 34.4       | 5.86       | 354, 364      | 352     |   |
| 23 | Cer d16:1/C26:1(-H2O) | 630.6 | 236.2 | 25.6              | 27.6              | 5.86       |               | 352     |   |
| 24 | Cer d18:0/C16:0       | 540.5 | 266.2 | 50.7, 43.1, 54.6  | 51.4, 45.8, 53.5  | 4.806      | 264, 301, 306 | 301     |   |
| 25 | Cer d18:0/C16:0(-H2O) | 522.5 | 266.2 | 38.5, 21.7        | 37.8, 21.9        | 4.806      | 268, 294      | 294     |   |
| 26 | Cer d18:0/C18:0       | 568.5 | 266.2 | 288.4, 57.6       | 568.3, 53.4       | 5.272      | 312, 325      | 325     |   |
| 27 | Cer d18:0/C18:0(-H2O) | 550.5 | 266.2 | 19.7              | 18.8              | 5.316      |               | 317     |   |
| 28 | Cer d18:0/C20:0       | 596.5 | 266.2 | 44.8, 42.5        | 43.1, 42.0        | 5.6        | 332, 340      | 332     |   |
| 29 | Cer d18:0/C20:0(-H2O) | 578.5 | 266.2 | 21.5              | 20.1              | 5.6        |               | 333     |   |

The user does not have to do this for every assay – as long as the MRM method stays the same and the LC-MS instrumentation stays the same, this template file can be applied to all future experiments.

## 5 Example “param.txt”

```
mzML_files
*.mzML

batch_info
batchinfo_SP2.txt

length_of_ion_chromatogram
4 40

transition_list
SP2 MRM list.txt

ISTD_trace_all
1

batch_correction
1
```

## 6 Command line for unix/macOS users

Installation of python3 ( $\geq 3.7$ ) and R ( $\geq 4.0.0$ ) is required.

```
cd {path_to_folder_with_mzML_files,_param.txt,_assay_info_and_batch_info}

# internal standards (ISTDs) peak extraction
python3 {path_to_MRMkit}/MRMistd.py

# peak detection
python3 {path_to_MRMkit}/MRMgetpeak.py

# draw ion chromatograms
Rscript {path_to_MRMkit}/MRMionc.r

# batch correction
```

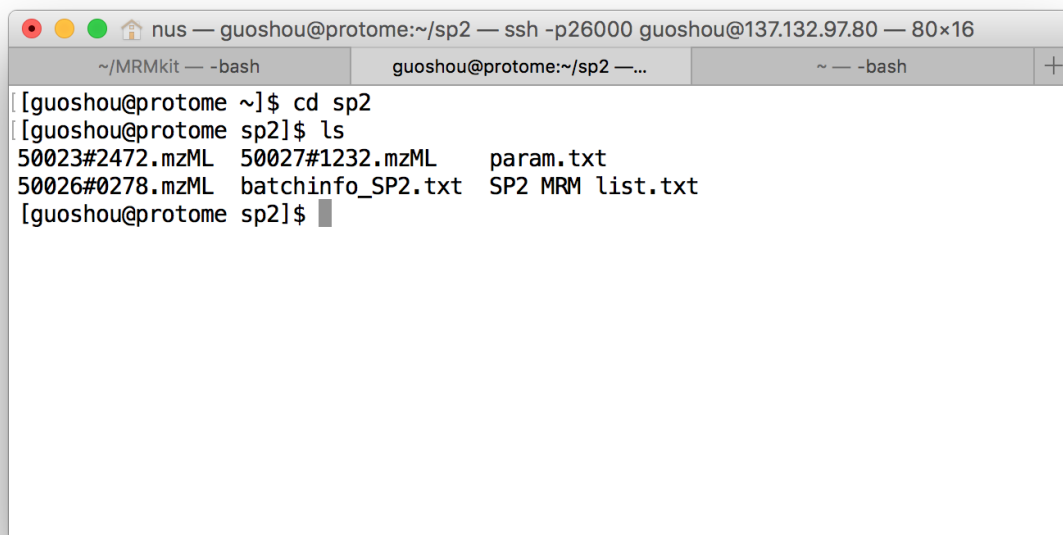
```
python3 {path_to_MRMkit}/MRMcorrect.py
```

```
# RT selection (optional)
```

```
# users to input RT of true peaks in user_RT.txt
```

```
python3 {path_to_MRMkit}/MRM_RT.py
```

1. Change directory to the folder (e.g. sp2) containing mzML files (a truncated list of mzML files shown here), param.txt, assay info and batch info



A terminal window with a title bar showing 'nus — guoshou@protome:~/sp2 — ssh -p26000 guoshou@137.132.97.80 — 80x16'. The window has three tabs: '~ /MRMkit — -bash', 'guoshou@protome:~/sp2 —...', and '~ — -bash'. The active tab is the second one. The terminal content shows the user navigating to the 'sp2' directory and listing its contents. The files listed are: 50023#2472.mzML, 50027#1232.mzML, param.txt, 50026#0278.mzML, batchinfo\_SP2.txt, and SP2 MRM list.txt.

```
[guoshou@protome ~]$ cd sp2
[guoshou@protome sp2]$ ls
50023#2472.mzML  50027#1232.mzML  param.txt
50026#0278.mzML  batchinfo_SP2.txt  SP2 MRM list.txt
[guoshou@protome sp2]$
```

2. internal standards (ISTDs) peak extraction, MRMkit's path in this example was “~/02srm”.





A terminal window titled "nus — guoshou@protome:~/sp2 — ssh -p26000 guoshou@137.132.97.80 — 80x16". The window has three tabs: "~ /MRMkit — -bash", "guoshou@protome:~/sp2 —...", and "~ — -bash". The active tab shows the command prompt "[guoshou@protome sp2]\$ python3 ~/02srm/MRMistd.py" with a cursor at the end.

### 3. peak detection



A terminal window titled "nus — guoshou@protome:~/sp2 — ssh -p26000 guoshou@137.132.97.80 — 80x16". The window has three tabs: "~ /MRMkit — -bash", "guoshou@protome:~/sp2 —...", and "~ — -bash". The active tab shows the command prompt "[guoshou@protome sp2]\$ python3 ~/02srm/MRMgetpeak.py" with a cursor at the end.

### 4. draw ion chromatograms

## 5. batch correction



## 7 Command line for unix/macOS users

## 8 Copyrights

BSD 3-Clause License

Copyright (c) 2020, Guoshou Teo 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 HOLDER 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.