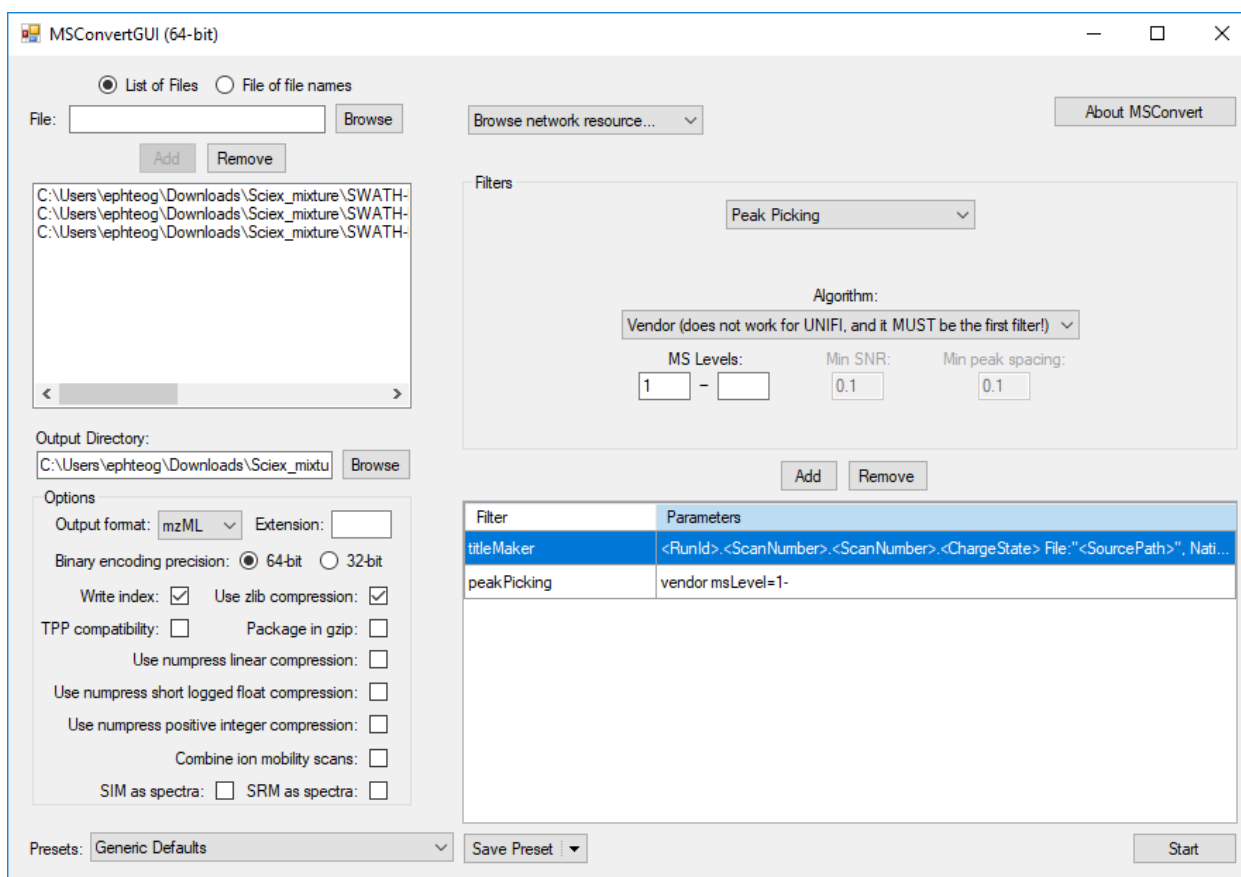# MetaboKit

July 7, 2020

1 Raw Data Conversion into mzML

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

1. Download and install a recent version ProteoWizard on a Windows computer.

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).

MSConvertGUI (64-bit)   — □ ×

◉ List of Files   ○ File of file names

File: [   ] Browse

Add   Remove

C:\Users\ephteog\Downloads\Sciex_mixture\SWATH-
C:\Users\ephteog\Downloads\Sciex_mixture\SWATH-
C:\Users\ephteog\Downloads\Sciex_mixture\SWATH-

Output Directory:
C:\Users\ephteog\Downloads\Sciex_mixtu Browse

Options
Output format: [mzML ∨] Extension: [   ]
Binary encoding precision: ◉ 64-bit ○ 32-bit
Write index: ☑   Use zlib compression: ☑
TPP compatibility: ☐   Package in gzip: ☐
Use numpress linear compression: ☐
Use numpress short logged float compression: ☐
Use numpress positive integer compression: ☐
Combine ion mobility scans: ☐
SIM as spectra: ☐   SRM as spectra: ☐

Presets: [Generic Defaults   ∨]

Browse network resource... ∨   About MSConvert

Filters
Peak Picking ∨

Algorithm:
Vendor (does not work for UNIFI, and it MUST be the first filter!) ∨
MS Levels: Min SNR: Min peak spacing:
[1] – [   ]   [0.1]   [0.1]

Add   Remove

Filter	Parameters
titleMaker	<RunId>.<ScanNumber>.<ScanNumber>.<ChargeState> File:"<SourcePath>", Nati...
peakPicking	vendor msLevel=1-

Save Preset ▼   Start

2   Input Parameters

Input parameters must be specified according to the user's preference for various aspects of the data processing. Input parameters are to be specified in the "param.txt" file.

MetaboKit requires that the user provide the following options for data processing:

- **mzML_files**: Names of the centroided mzML files. Use "*.mzML" to select all mzML files in the folder.

- **batch_info**: range of ion chromatogram of feature in seconds

- **length_of_ion_chromatogram**: range of ion chromatogram of feature in seconds

- **transition_list**: assay information with compound name, Q1, Q3, retention time and designated internal standards (ISTD).

- **ISTD_trace_all**: Set "1" to use the XIC shape of associated ISTD for peak integration for all transitions, "0" to directly take area under peak in the case for transitions with unimodal chromatogram.

- **batch_correction**: Set "1" to enable batch effect correction, "0" to disable.

3   Analysis Output Table

- **batch_adjusted.txt** A table of quantification.

4   Example "param.txt"

```
mzML_files
*.mzML

batch_info
batchinfo.txt

length_of_ion_chromatogram
4 40

transition_list
SP2 MRM list.txt

ISTD_trace_all
1

batch_correction
1
```

5   Copyrights

ING, 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.