LipidMatch Flow Instructions

Video tutorials are available at:

<https://www.youtube.com/playlist?list=PLZtU6nmcTb5kAOHAPjtpWXwyjfpDnaB2M>

Yang Li

yliweb@ufl.edu

Jeremy Koelmel

jeremkoelmel@gmail.com

1) Dependencies: (most windows computers will already have the two dependencies below installed, but if you do not you will need to install them)

Windows (built for Windows 10)

Microsoft .NET (Developed using Ver 4.7.2)

Java 64-bit version

2) Acquire data for the following data types:

You will need 3-4 data types: Blanks, Target, Samples, and MS/MS

Blanks (at least 3, recommended 3-4): Data with full scan (can also have MS/MS) of blank injections with the same chromatography used for samples. Ideally, extraction blanks would be used, since features found in the blanks at similar intensities to those in samples are removed. Therefore extraction blanks would remove ions/molecules contributed from the extraction procedure not the biological sample. Solvent blanks will work as well, but won't remove any features introduced during extraction.

Target (at least 1, recommended 3-5): these can be randomly picked samples, pooled samples, or any samples which are representative of all other samples. We recommend using 3-5 target samples, the less you use the faster the program will go (they are used for doing untargeted feature finding, which is used to generate a targeted list of peaks to search across all samples). The more targets you use the more features will be detected, and hence more lipid identifications will be obtained. They are also treated as normal samples and will be aligned with all you other samples in the final dataset. So for instance if you have 3 diseased and 3 healthy samples, you could run 1 diseased and 1 healthy as targeted samples, and the other 2 as normal samples (for each polarity). Or you could pool all 3 diseased and all 3 healthy and run the pools as targeted samples.

Samples (0 needed, 100s to 1000s possible depending on computer specs): These are data of samples for final feature/lipid molecule intensities to be determined for statistics or determining the relative amount of lipids in a certain sample type. Ideally both targeted and sample data is in full scan mode only, because this maximizes the number of scans across a peaks. More scans across a peak means more accurate calculation of peak area, etc. But data files containing full scan and MS/MS will run fine, and often provide an adequate number of full scans across a peak. More samples generally don't increase data-processing time, since a rapid peak integration method using a targeted list from the "target" file type is employed.

MS/MS (ddMS2): this can be auto-MS/MS (otherwise know as data-dependent MS/MS) or data of MS/MS data acquired using a targeted inclusion list. We recommend using between 1-6 MS/MS files for each polarity. The more MS/MS files used, generally the more identifications will be obtained, especially if intelligent data-acquisition methods such as iterative exclusion are used. But, the more data-files are used, the longer data-processing will take. All-ion-fragmentation data using thermo instruments is also accepted, and takes at least twice as long to process as data-dependent files.

2) Change .raw (or other vendor format) file names:

You can change your vendor formatted files to the right naming convention manually (or during data-acquisition) or using the tool that comes with LipidMatch Flow (see link for tutorial):

<https://youtu.be/eIpdhDnhh9I>

All file names should end in \_Neg.raw or \_Pos.raw, depending on the polarity.

All ion fragmentation data should have AIF somewhere in the name

Data-dependent are targeted MS/MS files should have "ddMS2" somewhere in the name

Blanks, for blank filtration (for example extraction blanks) should have "blank" somewhere in the file name

Files which are representative of all samples (eg. pooled samples) and can be processed to determine which features to target in all samples, should have "target" somewhere in the file name.

Length of the file names of MS/MS (ddMS2 or AIF) should not be longer than 23 characters. It is OK if full scan data used for feature detection is over 23 characters.

3) Input Files:

Double click LipidMatch\_Batch.exe to open the GUI interface.

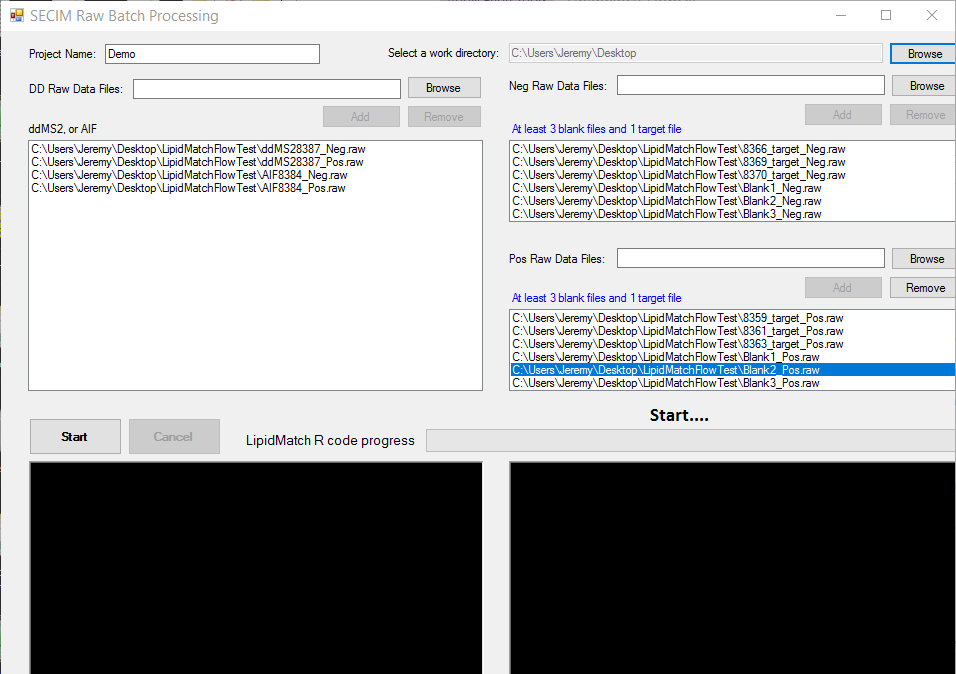
You should have at least one MS/MS file per polarity (AIF or ddMS2)

Atleast 3 blanks and 1 file with "target" in their name for each polarity

You can run only a single polarity if desired

After uploading files and choosing project name and working directory the start button should no longer be grayed out

Select start.



4) Interpreting results:

You will get multiple files outputted in the working directory you selected in the previous step; these include the converted .raw files to .mzXML and the MZmine files in the Temp\_Work folder which is generated, and the LipidMatch output files in the LipidMatch\_Run directory. Your final data with each feature's *m/z*, retention time, intensities across samples, and lipid annotation(s) will be in the LipidMatch\_Run directory saved as Pos\_IDed.csv and Neg\_IDed.csv. You will also have Neg\_MolecularSpecies.csv, Pos\_MolecularSpecies.csv, and PosNeg\_MolecularSpecies.csv. The latter contains one feature for each lipid molecular species and combines positive and negative mode data, this is the file you will most likely want for further statistics, etc.

Appended in the column labeled ID of PosNeg\_MolecularSpecies.csv contains lipid annotations or multiple lipid annotations per feature separated by “I”.

* There are currently 4 identity markers:
  + 1\_[ID] represents confirmation by ddMS2
  + 2\_[ID] represents confirmation by AIF
  + 3\_[ID] represent ddMS2 by class
  + 4\_[ID] represents confirmation by exact mass.
* The order for multiple confirmations is based on summed fragment intensity.
* The data files in the folder “…Neg/Confirmed/”, “…Pos/Confirmed/”, and “…PosByClass/Confirmed/” can be used to better determine confidence and the most abundant lipids defining a specific feature (these tables contain fragment intensities, retention times at max intensity, experimental and actual mass of fragments, number of scans containing fragments, etc.).

MZmine project files can be found in the Temp\_Work folder (use the most recently saved files for negative and positive mode). These files can be open in MZmine and the gap-filled table can be double clicked to manually investigate the quality of peak picking.