User Manual for *Lipid Wizard*

1. Installation

1.1 Environment

Users must have *Python 3.9.0* or later installed before downloading the *Lipid Wizard* software. *Python 3.9.0* can be downloaded from https://www.python.org/downloads/release/python-390/

1.2 Software Download

Lipid Wizard can be downloaded from https://github.com/RaoboXu/Lipidwizard/releases/

After downloading, the users need to unzip the package to a desired folder.

1.3 Operating Systems

1.3.1 Windows

For Windows users, double click the *LipidWizard.exe*

1.3.2 Linux

For Ubuntu Users, double click the *LipidWizard*

1.3.3 Apple OSX

For Apple OSX Users, you must open the *Terminal* and change the directory to the unzipped folder, e.g., if the unzipped folder is stored on the Desktop, run the command:

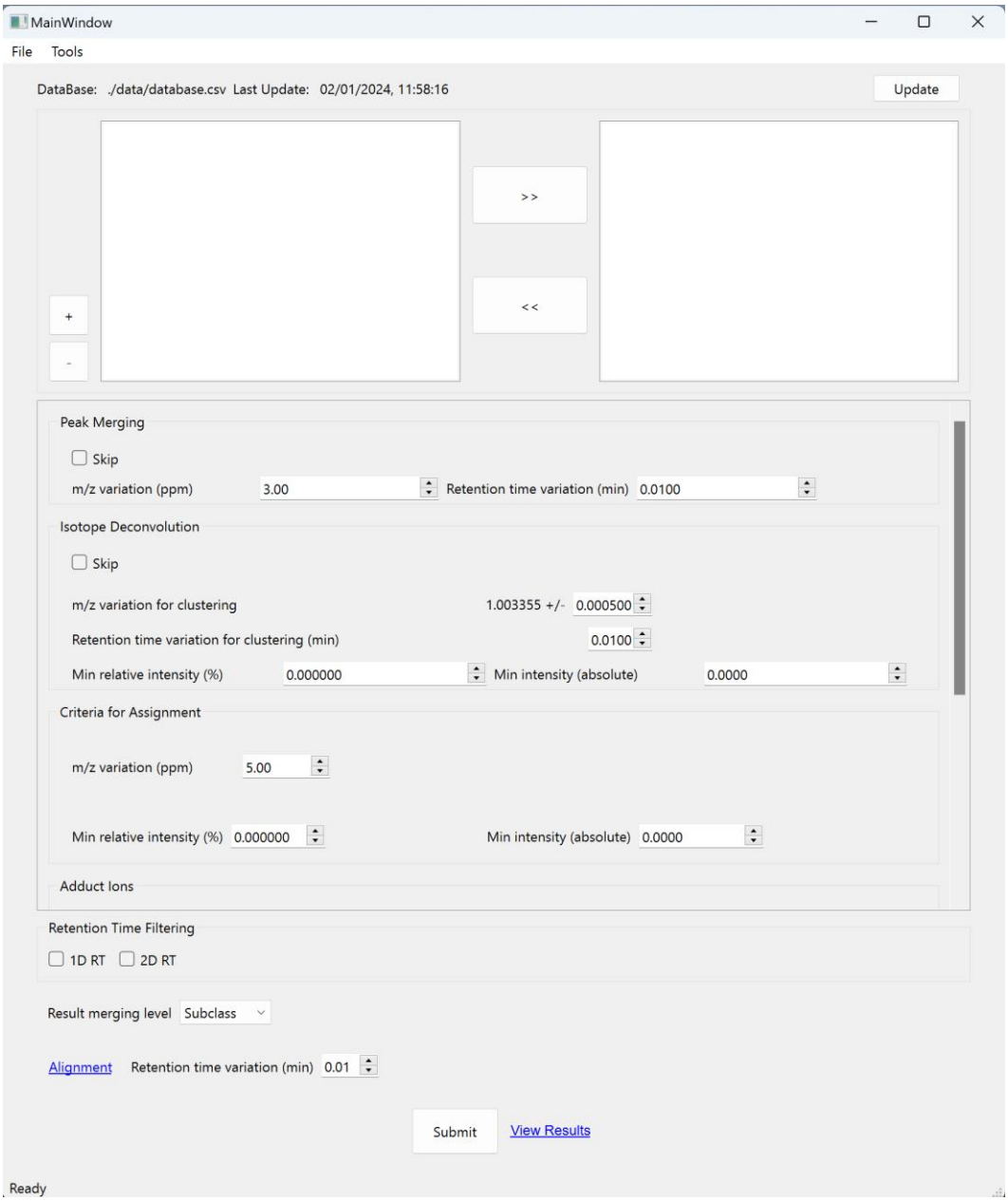
cd desktop/LipidWizard-v0.9.9-OSX-Intel/

Then, run the command:

/LipidWizard

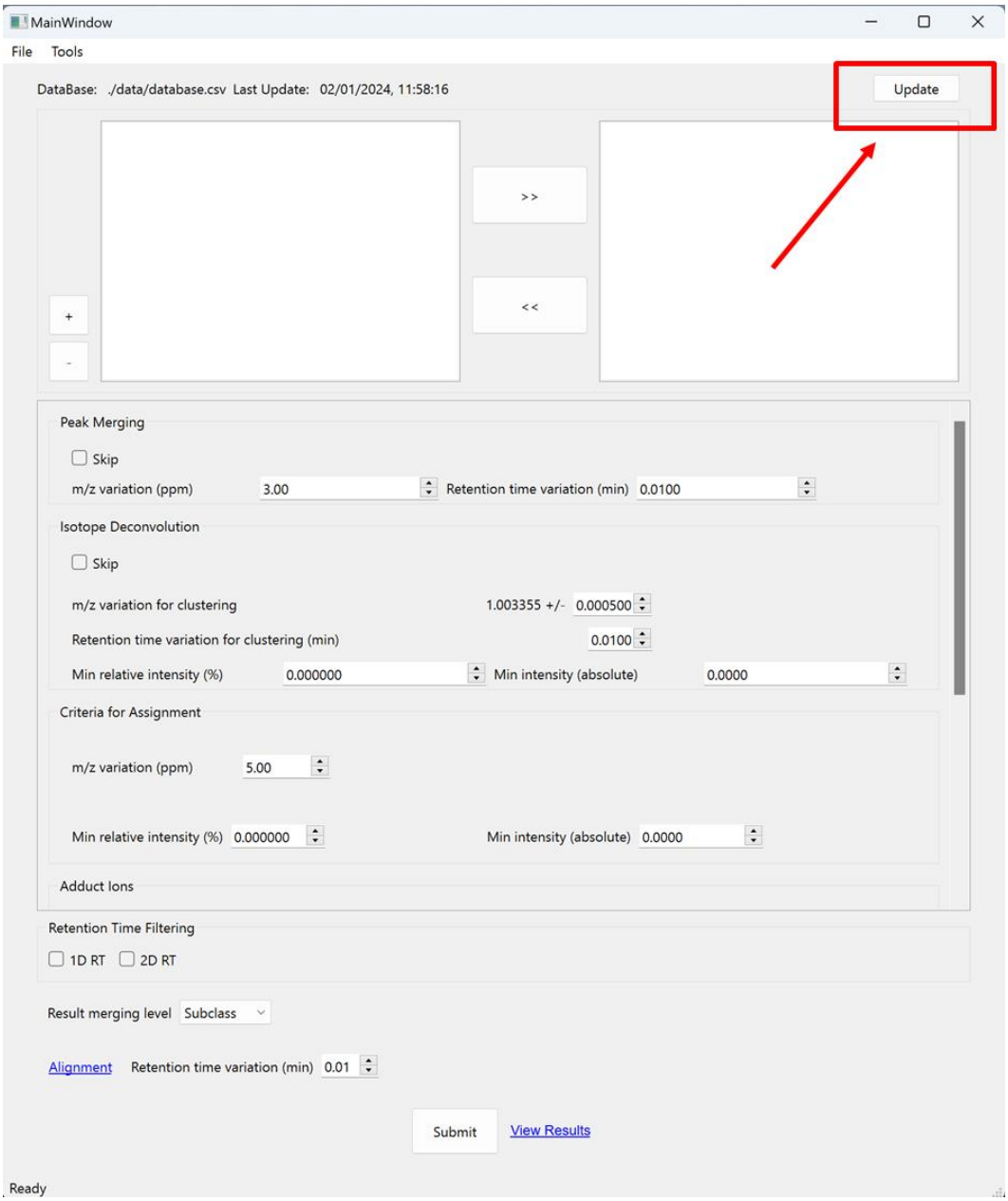
to launch the software.

2. Main User Interface



The main user interface of *Lipid Wizard* has three panels, an input panel, a parameter panel, and an analysis panel. The input panel enables users to update *LIPID MAPS*® database and input the peak list(s) generated by *XCMS* or other peak-picking software. The parameter panel lets users define the parameters to be used for peak merging, isotopic peak deconvolution, adduct ion selection, lipid class and/or subclass selection, and lipid assignment. The users can perform retention time filtering, merging assigned lipids, and cross-sample peak list alignment in the analysis panel.

3. Updating Lipid Database



Lipid Wizard saves a copy of all lipids in the *LIPID MAPS*® and uses them as its local database. The users can update the local lipid database by clicking the 'Update' button in the input panel. After clicking that button, the software will automatically download all lipids of the latest version of the *LIPID MAPS*®. The local lipid database in *Lipid Wizard* will be updated using the newly downloaded lipids.

4. Input Files

The input files of peak lists **MUST** have the exact format as provided in the test files. This file format is identical to that generated by *XCMS* software.
For more details about *XCMS*, please visit
https://xcmsonline.scripps.edu/landing_page.php?pgcontent=mainPage
If other software packages are used for peak picking, the users must reformat the peak lists.

4.1 Format of *XCMS* deconvoluted peak list

If multiple LC-MS data are analyzed together, *XCMS* software generates one alignment table with peaks detected in each LC-MS file listed in the same column. These peak lists deconvoluted from *XCMS* need to be split before *Lipid Wizard* performs lipid assignment. To split the peak lists from the *XCMS* output file into individual sample-based lists, users should ensure that the file format is .xlsx, and then apply peak list split using "*Split Peak Lists"*, a built-in component of *Lipid Wizard* (see Section 5. Splitting Peak List).

An example peak list generated by *XCMS* can be found in the folder:
Test_Files/XCMS_Peak_List_for_Splitting.

4.2 Format of split peak lists

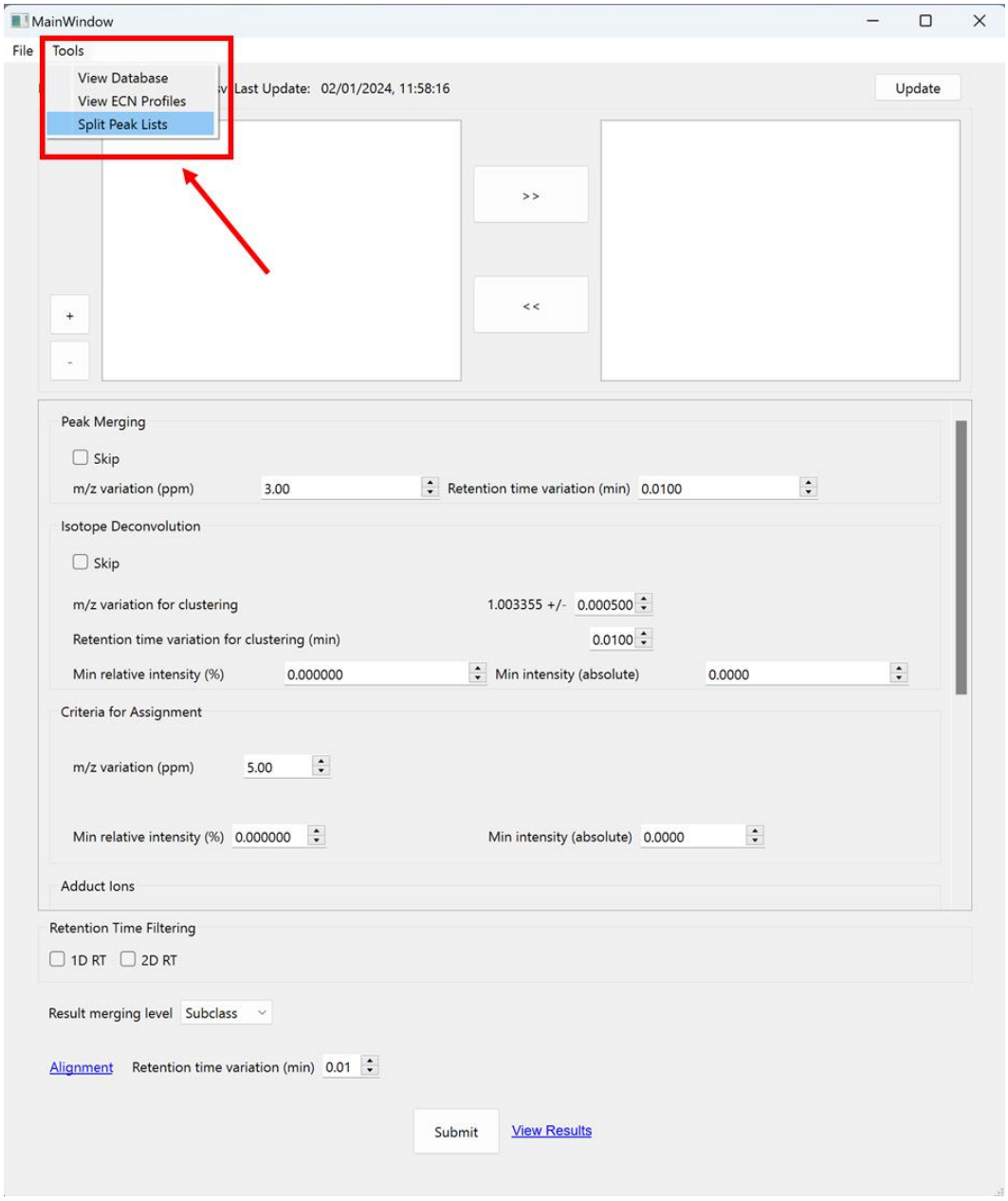
Each split peak list is saved in a .csv file. The users need to remove the illegal characters in the file name located in the first row of each .csv file. Users can only use the underscore mark '_' to separate words if needed. Each split peak list contains eight columns with the following column heads.

m/z	Intensity	Retention time	Relative intensity	mzmin	mzmax	rtmin	rtmax

Multiple examples of the split peak lists are provided in the downloaded *Lipid Wizard* installation package. Users can find those files in the folder:
Test_Files/Split_Peak_Lists

Users can view them as references of input file format and perform the assignment using *Lipid Wizard* as test files.

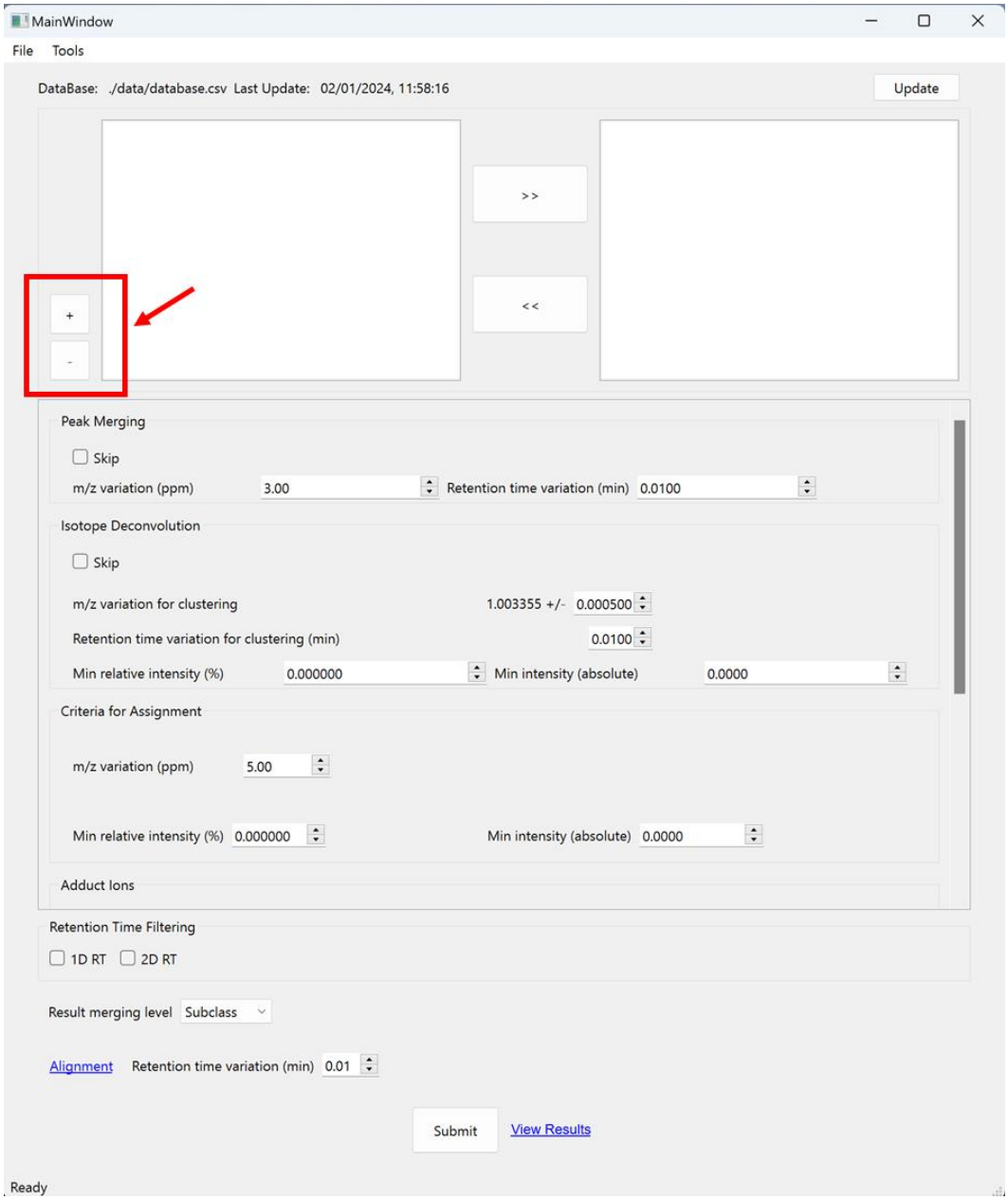
5. Splitting Peak List



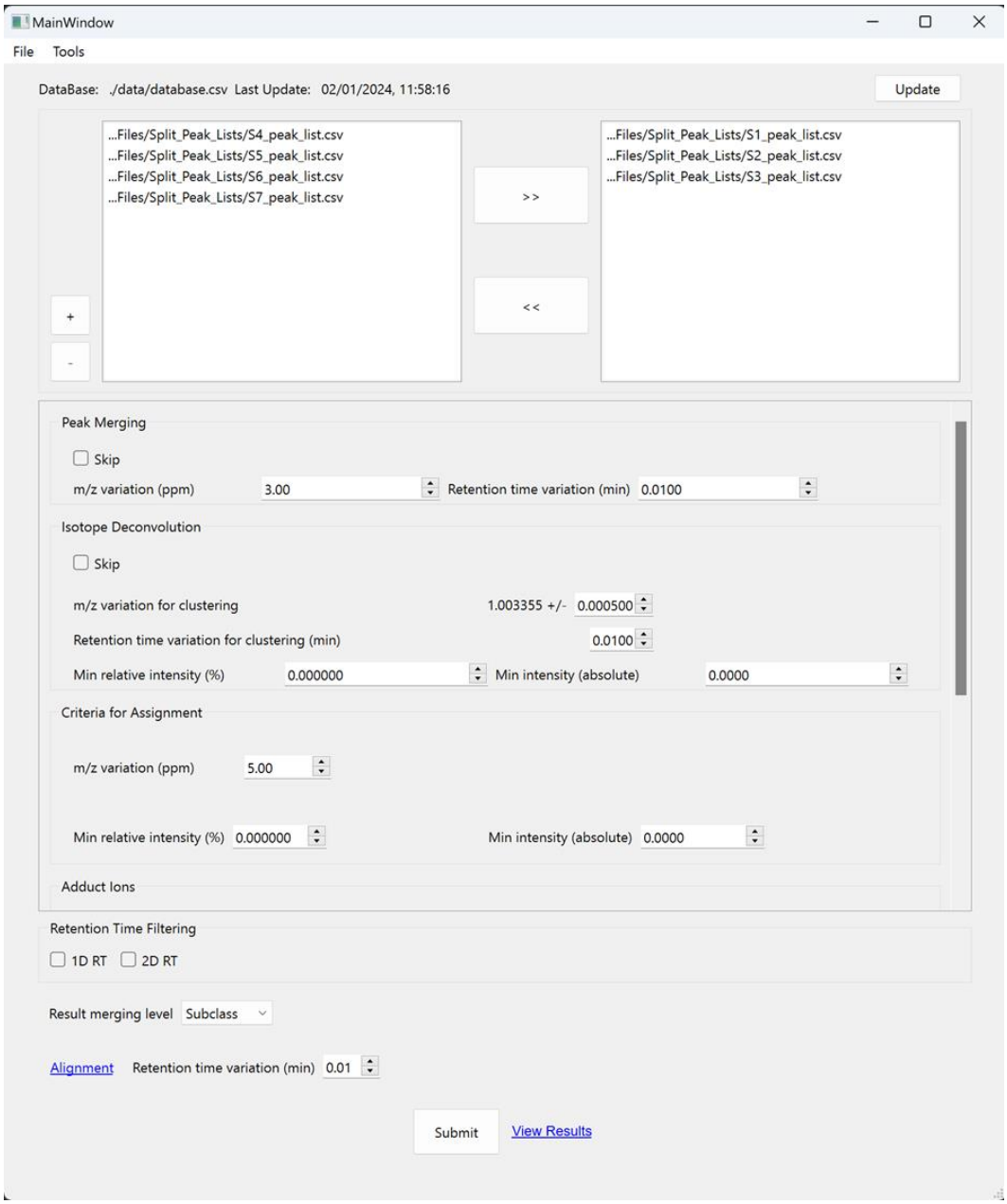
If multiple HILIC×RPC-MS data files are analyzed by *XCMS*, the output file of *XCMS* is a single .xlsx file, i.e., an aligned peak list table. Each column contains the peaks detected in an input 2DLC-MS data file, and each row contains all peaks generated by the same compounds but detected in different samples. *Lipid Wizard* processes the peak list of each sample, one at a time. Therefore, the users need to split the peak lists in the *XCMS* output file into individual peak lists, by clicking the "*Split Peak Lists*" to select the *XCMS* file

from the pop-up window. After clicking OK, *Lipid Wizard* splits the combined peak list file into individual sample-based peak lists and saves them in the user-defined location.

6. Loading Sample-based Peak Lists



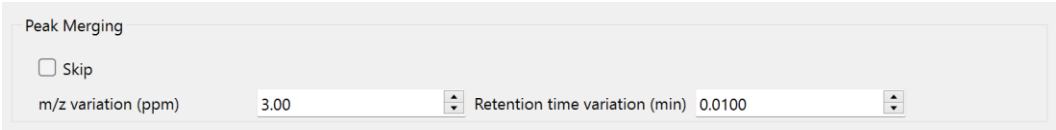
After peak list splitting, the users can upload the split peak lists into *Lipid Wizard*. The users can use the "+" and "-" buttons to select and unselect the individual sample-based peak lists that users want to analyze. After selecting and highlighting the peak lists, click the arrow button in the middle of the input panel to confirm.


MainWindow
File
Tools
DataBase: ./data/database.csv Last Update: 02/01/2024, 11:58:16
Update
...Files/Split_Peak_Lists/S4_peak_list.csv
...Files/Split_Peak_Lists/S5_peak_list.csv
...Files/Split_Peak_Lists/S6_peak_list.csv
...Files/Split_Peak_Lists/S7_peak_list.csv
>>
<<
+
-
...Files/Split_Peak_Lists/S1_peak_list.csv
...Files/Split_Peak_Lists/S2_peak_list.csv
...Files/Split_Peak_Lists/S3_peak_list.csv
Peak Merging
Skip
m/z variation (ppm)
3.00
Retention time variation (min)
0.0100
Isotope Deconvolution
Skip
m/z variation for clustering
1.003355 +/-
0.000500
Retention time variation for clustering (min)
0.0100
Min relative intensity (%)
0.000000
Min intensity (absolute)
0.0000
Criteria for Assignment
m/z variation (ppm)
5.00
Min relative intensity (%)
0.000000
Min intensity (absolute)
0.0000
Adduct Ions
Retention Time Filtering
1D RT
2D RT
Result merging level
Subclass
Alignment
Retention time variation (min)
0.01
Submit
View Results

7. Parameter Selection for Analysis

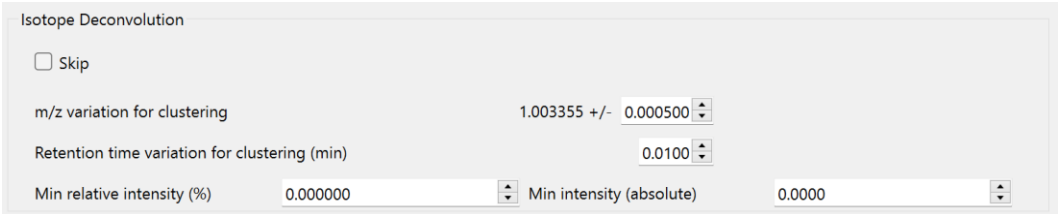
7.1 Peak Merging

The users can define m/z variation and retention time variation for peak merging to get rid of the potential peak-picking problems generated by *XCMS*.



7.2 Isotopic Peak Deconvolution

The users can define parameters for isotopic peak clustering and stripping to quantitatively deconvolute the overlapping peaks.

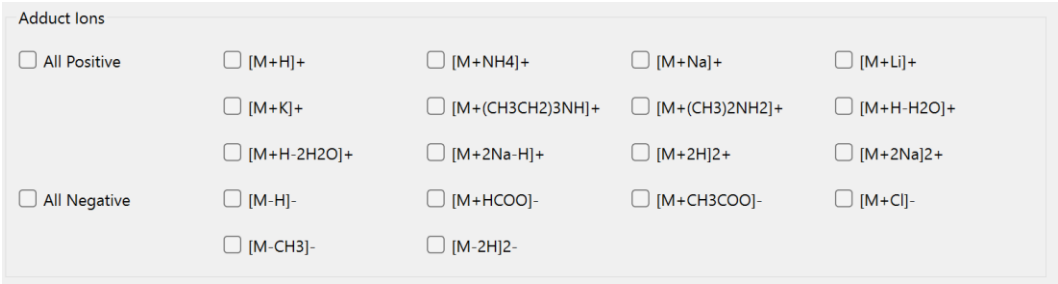


7.3 Criteria for Assignment

The users can define parameters for lipid assignment, including the m/z variation in ppm, and minimum relative or absolute intensity.

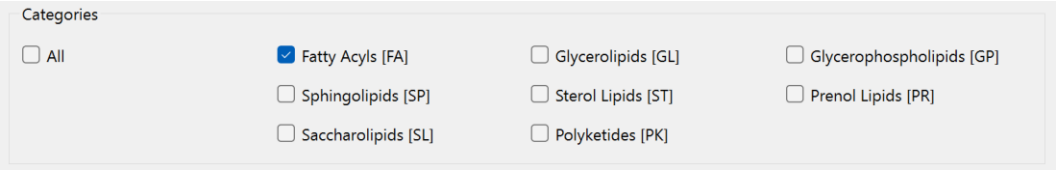
7.4 Adduct Ions

The users can select possible adduct ions.



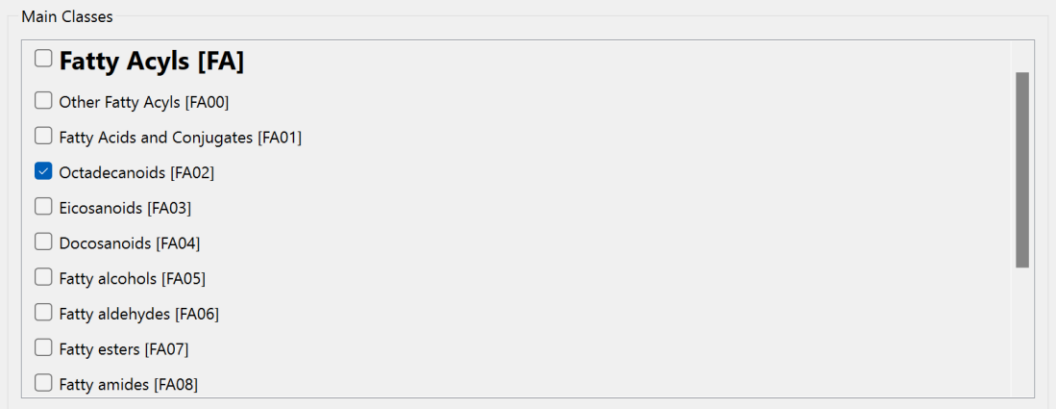
7.5 Lipid Categories

The users can define parameters for interested lipid categories that need to be assigned, e.g., checking Fatty Acyl [FA] if the users are only interested in fatty acyls.



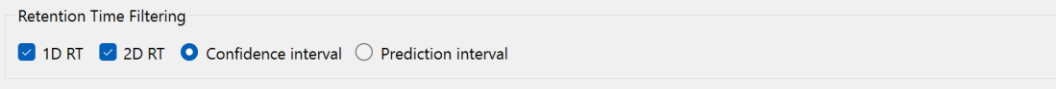
7.6 Main Classes

The users can define parameters for the lipid main classes of interest, e.g., checking Octadecanoids [FA02] if users only focus on Octadecanoids.



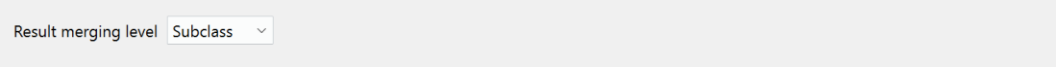
7.7 Retention Time Filtering

Lipid Wizard can filter the m/z matched lipids by the ^{1}D and ^{2}D retention time, respectively. Users can select either confidence interval or prediction interval for ^{2}D RT filtering, by checking and unchecking the corresponding boxes. For details, please refer to Section 8. Configuration Files for ^{1}D and ^{2}D RT Filtering.



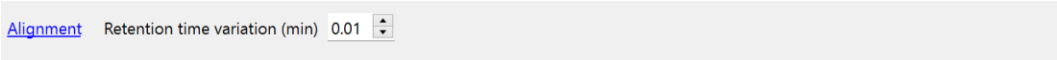
7.8 Result Merging

Lipid Wizard can merge the assigned lipids at different levels. For instance, if the users are only interested in lipids at the subclass level, the users can select the subclass. By doing so, all assigned lipids in the same subclass will be merged, and the output file will contain assigned lipids only with different subclasses in one row.



7.9 Alignment

After performing an assignment for each submitted sample peak list, the users can define retention time variation for cross-sample peak list alignment. By default, *Lipid Wizard* aligns all peak lists together. However, the users have the option to align a subset of assigned peak lists by clicking the "Alignment" button (blue color) to select the samples of interest.



8. Configuration Files for ^{1}D and ^{2}D RT Filtering

8.1 ^{1}D RT filtering

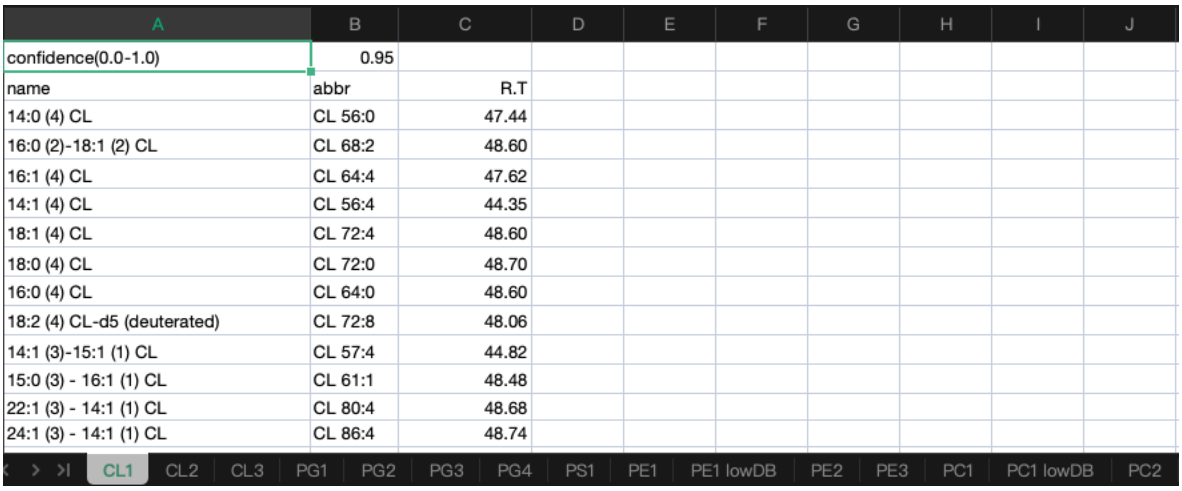
ExpectedRTRangle.xlsx (File path: data/ERT_DATA) is used for ^{1}D RT filtering. The users must determine ^{1}D expected RT range for each subclass based on the retention time of the lipid standards of that subclass. Column A is the lipid class name with a format of [categories.main class.subclass], and the users must define Categories. For the main class and subclass, the users may skip if the information is unavailable for the lipid standards they used. Columns B and C are user-defined RT time ranges. Column D is a space for users to write comments as a reminder, and *Lipid Wizard* does not read the information in this column.

A	B	C	D
class	r.t.start	r.t.end	comment
Fatty Acyls [FA]	1	15	LCFA
Glycerolipids [GL]	1	15	DG/MG/TG
Sphingolipids [SP].Ceramides [SP02]	1	15	Cer
Glycerophospholipids [GP].Glycerophosphoglycerols [GP04].Diacylglycerophosphoglycerols [GP0401]	15	73	PG
Glycerophospholipids [GP].Glycerophosphoglycerols [GP04].Monoacylglycerophosphoglycerols [GP0405]	32	90	LPG
Glycerophospholipids [GP].Glycerophosphoinositols [GP06].Monoacylglycerophosphoinositols [GP0605]	32	90	LPI
Glycerophospholipids [GP].Glycerophosphoethanolamines [GP02].Monoacylglycerophosphoethanolamines [GP0205]	55	90	LPE
Glycerophospholipids [GP].Glycerophosphocholines [GP01].Monoacylglycerophosphocholines [GP0105]	72	90	LPC
Glycerophospholipids [GP].Glycerophosphates [GP10].Monoacylglycerophosphates [GP1005]	72	90	LPA
Sphingolipids [SP].Sphingoid bases [SP01].Lysosphingomyelins and lysoglycosphingolipids [SP0106]	72	90	LSM
Sphingolipids [SP].Phosphosphingolipids [SP03].Ceramide phosphocholines (sphingomyelins) [SP0301]	72	90	SM
Glycerophospholipids [GP].Glycerophosphoserines [GP03].Diacylglycerophosphoserines [GP0301]	73	90	PS
Glycerophospholipids [GP].Glycerophosphoinositols [GP06].Diacylglycerophosphoinositols [GP0601]	32	90	PI
Glycerophospholipids [GP].Glycerophosphoethanolamines [GP02].Diacylglycerophosphoethanolamines [GP0201]	32	90	PE
Glycerophospholipids [GP].Glycerophosphocholines [GP01].Diacylglycerophosphocholines [GP0101]	55	90	PC
Glycerophospholipids [GP].Glycerophosphates [GP10].Diacylglycerophosphates [GP1001]	72	90	PA
Glycerophospholipids [GP].Glycerophosphoglycerophosphoglycerols [GP12].Monoacylglycerophosphoglycerophosphomonoradylglycerols [GP1207]	35	90	CL
Glycerophospholipids [GP].Glycerophosphoglycerophosphoglycerols [GP12].Diacylglycerophosphoglycerophosphomonoradylglycerols [GP1202]	35	90	CL
Glycerophospholipids [GP].Glycerophosphoglycerophosphoglycerols [GP12].Diacylglycerophosphoglycerophosphodiradylglycerols [GP1201]	35	90	CL

8.2 ^{2}D RT filtering

"Lipid category".xlsx (File path: data/ECN_DATA) is used for the ^{2}D RT filtering. It may contain up to eight *.xlsx* files because *LIPID MAPS*® has a total of eight lipid categories. Each file name must be the same as the lipid category given by *LIPID MAPS*®. In each file, the users can generate a new worksheet for a different lipid class under the same categories. The worksheet name has no format, e.g., in file Glycerophospholipids GP.xlsx, *Lipid Wizard* provides Cardiolipin (CL), Phosphatidylglycerol (PG), Phosphatidylethanolamine (PE), Phosphatidylcholine (PC) in four worksheets. In each worksheet, the users can determine the confidence level or prediction level (0% to 100%) for the ^{2}D RT filtering, and update the name of lipid

standards in column A, lipid name of the standards at species level in column B, and ^{2}D experimental RT in column C.

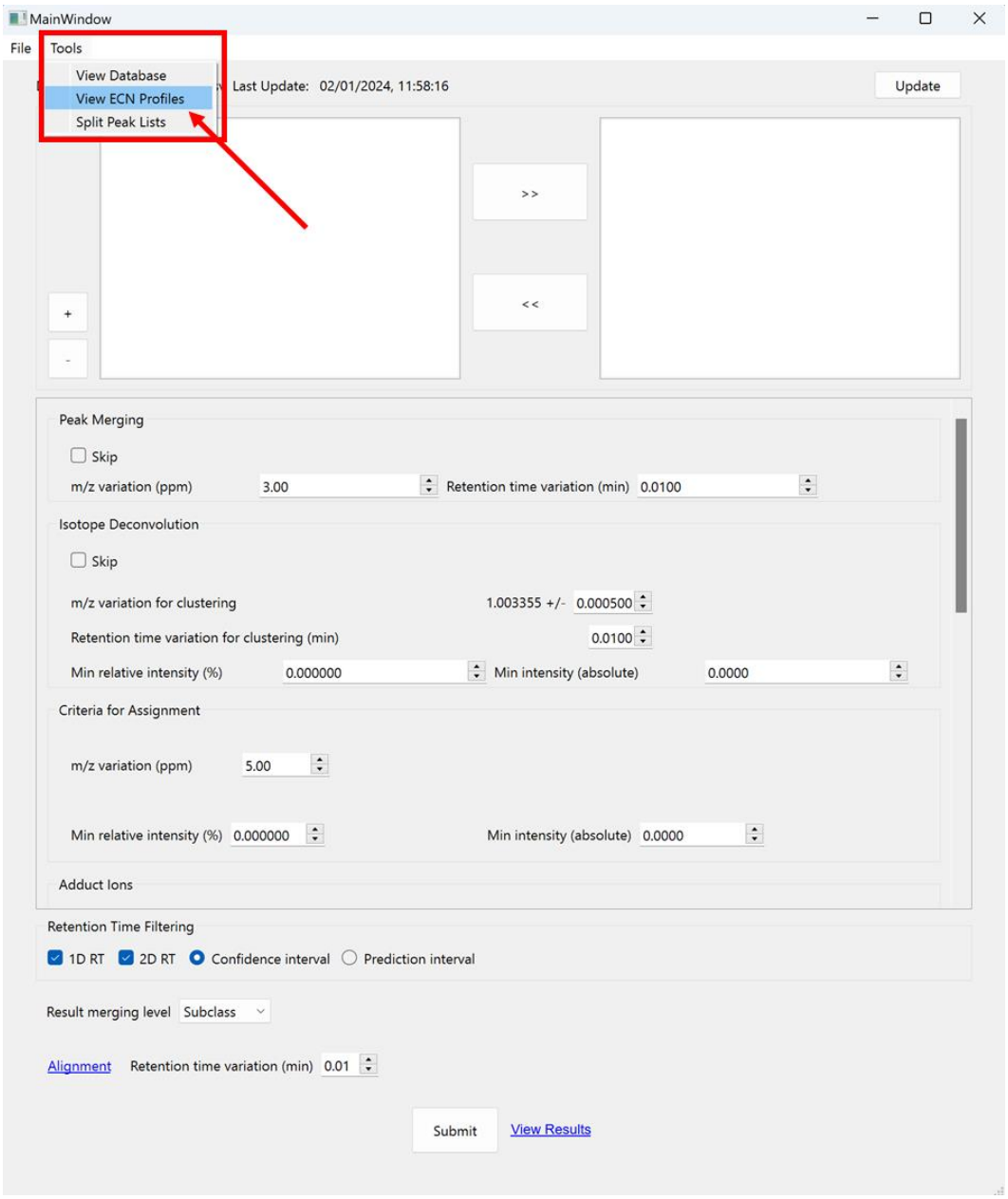


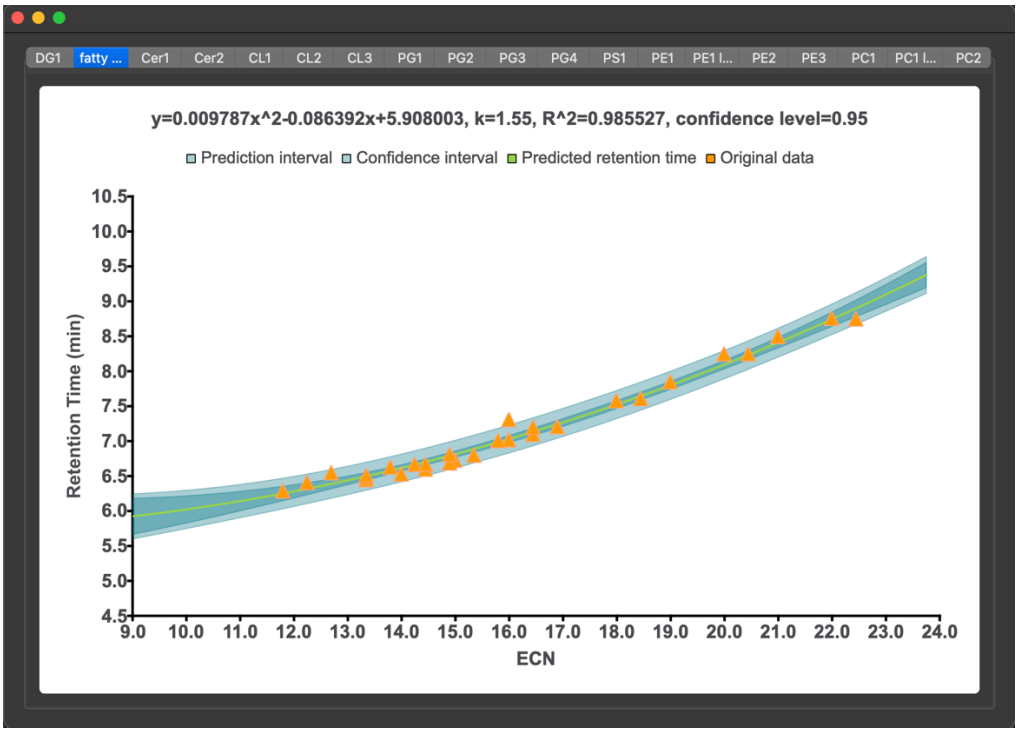
	A	B	C	D	E	F	G	H	I	J
1	confidence(0.0-1.0)	0.95								
2	name	abbr	R.T							
3	14:0 (4) CL	CL 56:0	47.44							
4	16:0 (2)-18:1 (2) CL	CL 68:2	48.60							
5	16:1 (4) CL	CL 64:4	47.62							
6	14:1 (4) CL	CL 56:4	44.35							
7	18:1 (4) CL	CL 72:4	48.60							
8	18:0 (4) CL	CL 72:0	48.70							
9	16:0 (4) CL	CL 64:0	48.60							
10	18:2 (4) CL-d5 (deuterated)	CL 72:8	48.06							
11	14:1 (3)-15:1 (1) CL	CL 57:4	44.82							
12	15:0 (3) - 16:1 (1) CL	CL 61:1	48.48							
13	22:1 (3) - 14:1 (1) CL	CL 80:4	48.68							
14	24:1 (3) - 14:1 (1) CL	CL 86:4	48.74							

CL1 | CL2 | CL3 | PG1 | PG2 | PG3 | PG4 | PS1 | PE1 | PE1 lowDB | PE2 | PE3 | PC1 | PC1 lowDB | PC2

8.3 Visualizing the ECN Models

After updating the experimental RT of lipid standards in the targeted files, the users can visualize the results of each ECN model by clicking the menu Tools → View ECN Profiles. A pop-up window will display the results of ECN models of all lipid standards. The menu of the popup window lists all ECN models for lipid subclasses. The users can click an item on the menu to visualize the ECN model of that subclass. The second-order polynomial regression function and the optimal *k* value are also provided for each ECN model.


MainWindow
File
Tools
View Database
View ECN Profiles
Split Peak Lists
Last Update: 02/01/2024, 11:58:16
Update
>>
<<
+
-
Peak Merging
Skip
m/z variation (ppm)
3.00
Retention time variation (min)
0.0100
Isotope Deconvolution
Skip
m/z variation for clustering
1.003355 +/-
0.000500
Retention time variation for clustering (min)
0.0100
Min relative intensity (%)
0.000000
Min intensity (absolute)
0.0000
Criteria for Assignment
m/z variation (ppm)
5.00
Min relative intensity (%)
0.000000
Min intensity (absolute)
0.0000
Adduct Ions
Retention Time Filtering
1D RT
2D RT
Confidence interval
Prediction interval
Result merging level
Subclass
Alignment
Retention time variation (min)
0.01
Submit
View Results



9. Submit Task

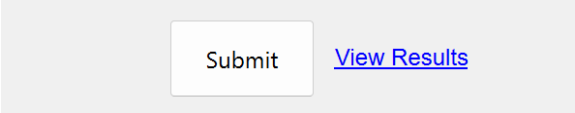
After visualizing the ECN models, the users can click the "Submit" button in the main panel to submit the task. A popup window will allow the users to save the analysis result, i.e., an alignment table to a specific location. The analysis is done when the command:

Writing data C:/ "USER DEFINED DESTINATION"/stat.csv SUCCESS!

shows in the command prompt or terminal.

10. Visualizing the Results Online

After the analyses, the users can view the results online by clicking 'View Results'. A popup window will show up. The users can view the lipid assignment results by further clicking the lipid category, main class, or subclass on the left panel of the popup window.



The assignment results are also saved in the user-defined folder named 'MatchResults-year-month-day-time'. Inside the folder, 'aligned_result.csv' file contains the assignment results, and 'stat.csv' file contains the peak area summation result at the user-defined

level. The folder ‘sample name’ contains several .csv files which are the results of each step of the assignment including peak merging, peak clustering, assignment results before ^{1}D RT filtering, assignment results after ^{1}D filtering, etc.

In the file of ‘aligned_result.csv’, *Lipid Wizard* keeps all assigned lipids, which include lipids passed ^{1}D and ^{2}D RT filters and the lipids are assigned only based on m/z matching. Assigning lipids only by parent ion m/z matching will have low confidence, and would not be at the same level as the other assigned lipids that passed ^{1}D and ^{2}D RT filtering. For those assigned lipids, the users can decide on whether to use these lipids in downstream analyses. If not, the users can sort the alignment table by “class” and “subclass” columns in the alignment file, and then, remove those lipids that do not belong to the classes of the lipid standards.