Lipid annotation with MS2Analyzer

Yan Ma

10/24/2013

Checklist before you start…

You need to have:

1.A computer with Java environment and Office(2003 or higher)

2.MS/MS spectra in MGF files

3.Latest version of MS2Analyzer jar file

4.Query text files

For additional validation with lipidblast, download the library and NIST search software in the following link(download the 87 MB full version):

http://fiehnlab.ucdavis.edu/projects/LipidBlast

Instruction outline

Step1: Using MS2Analyzer to search the spectra feature

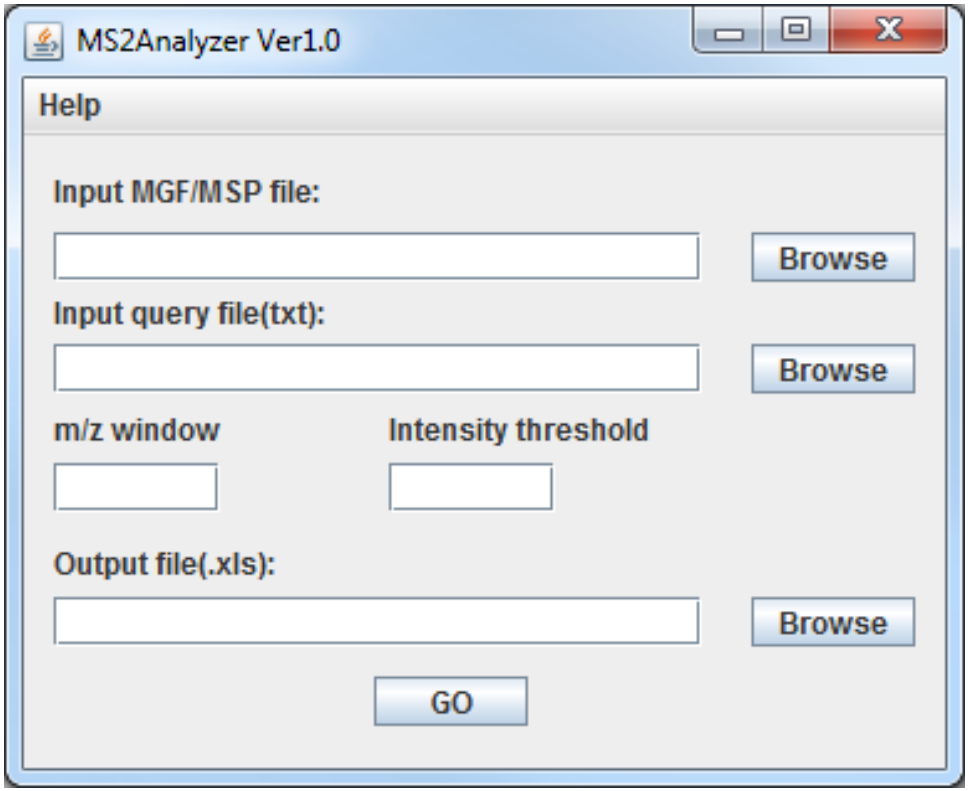
Step2: Using excel filters to annotate lipids

Step3: Confirmation by checking retention time (optional)

Step 4:Using NIST search for library validation(optional)

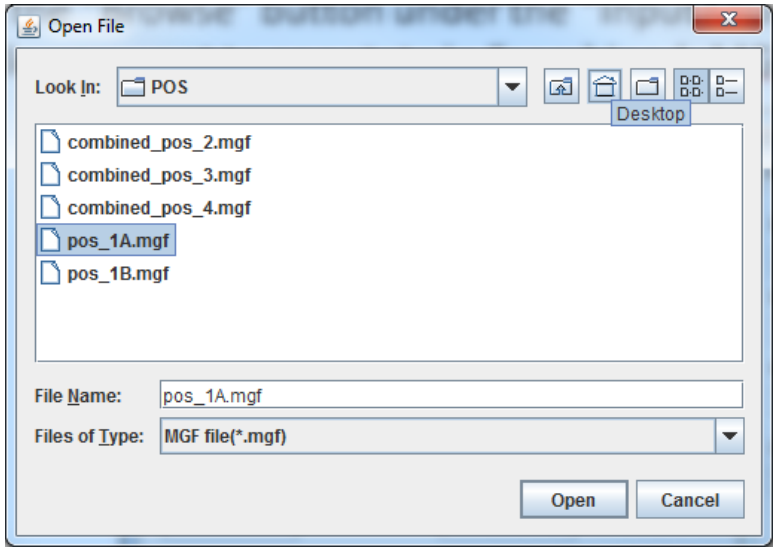
Step 1: Using MS2Analyzer to search the spectra feature

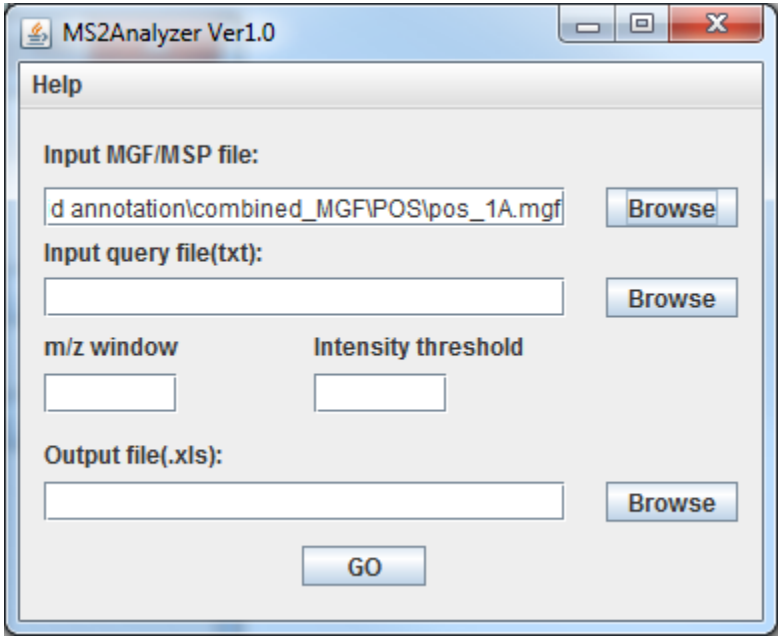
1.Double click on MS2Analyzer-ver1.jar to start the program



Step 1: Using MS2Analyzer to search the spectra feature

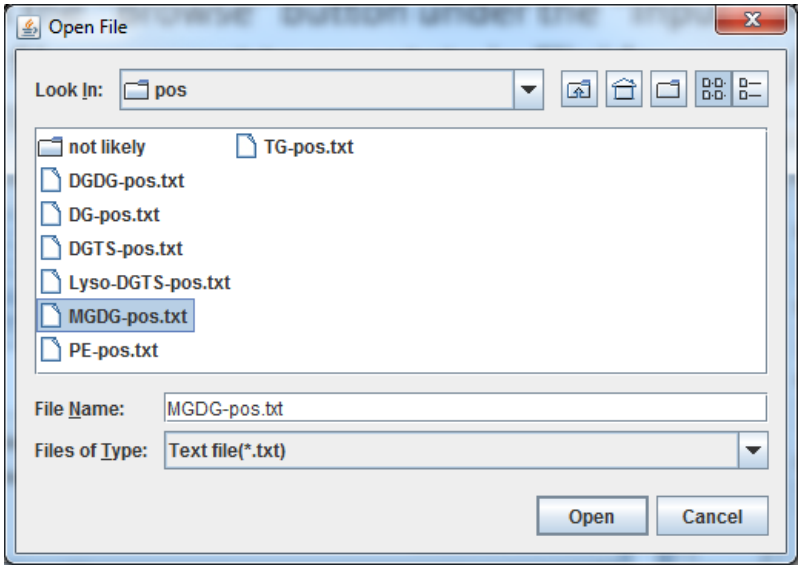
2.Click the “Browse” button under the “Input MGF/MSP file” and select the MGF file you want to annotate.

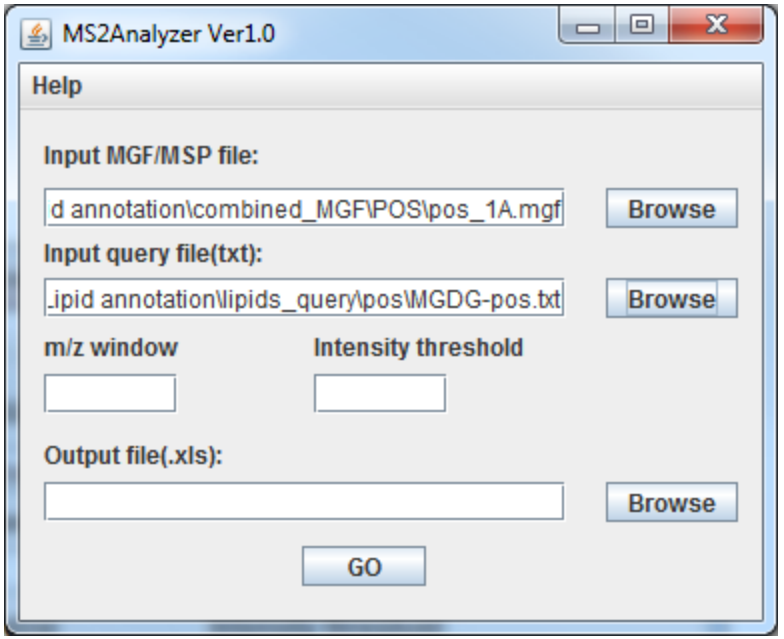




Step 1: Using MS2Analyzer to search the spectra feature

3.Click the “Browse” button under the “Input query file(txt)” and select the query file you want.

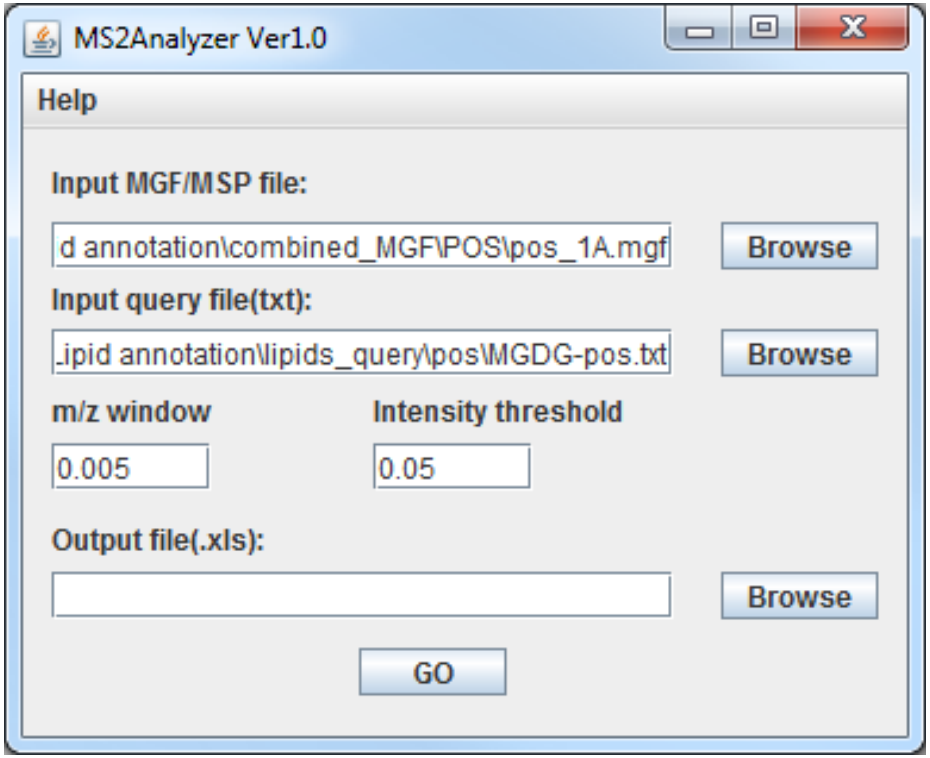




Step 1: Using MS2Analyzer to search the spectra feature

4.Type your desired parameters into “m/z window” and “intensity threshold”.

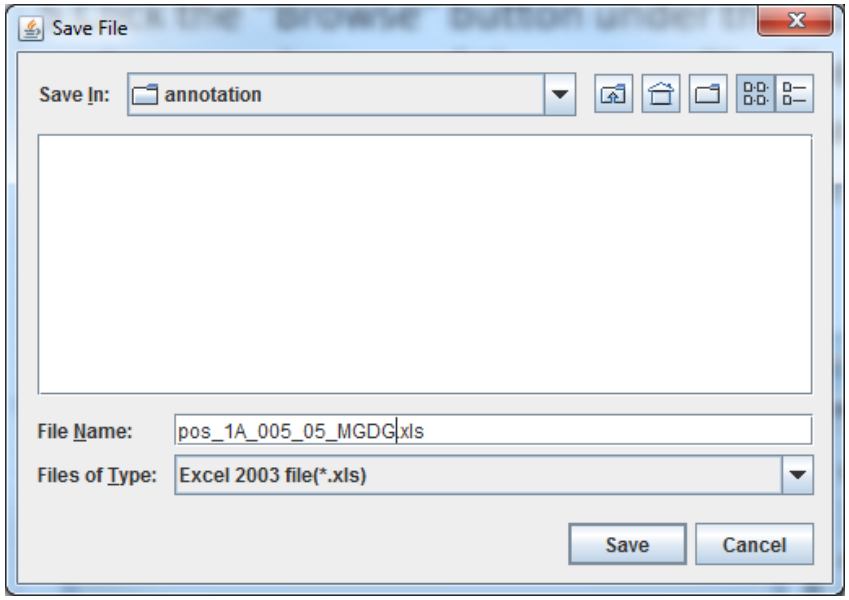
The first one depends on the mass accuracy of the instrument and the second one depends how noisy the spectra are. Intensity threshold are relative intensity to the base peak, from 0 to 1.0.

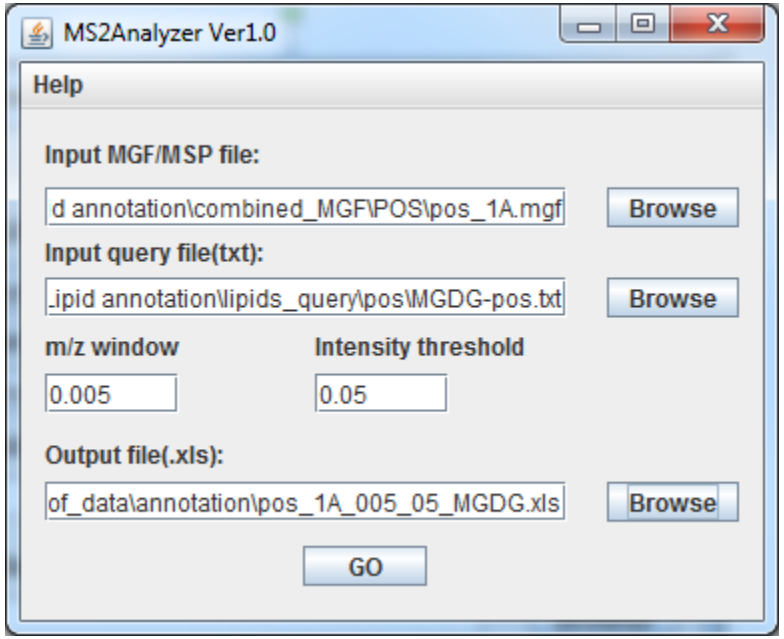


Step 1: Using MS2Analyzer to search the spectra feature

5.Click the “Browse” button under the “Output files (.xls)” and choose the pathway and name of the output file.

The default name of the output file is same as the name of the MGF files, but you may want to add more information about the lipid class and parameters. Here’s an example:

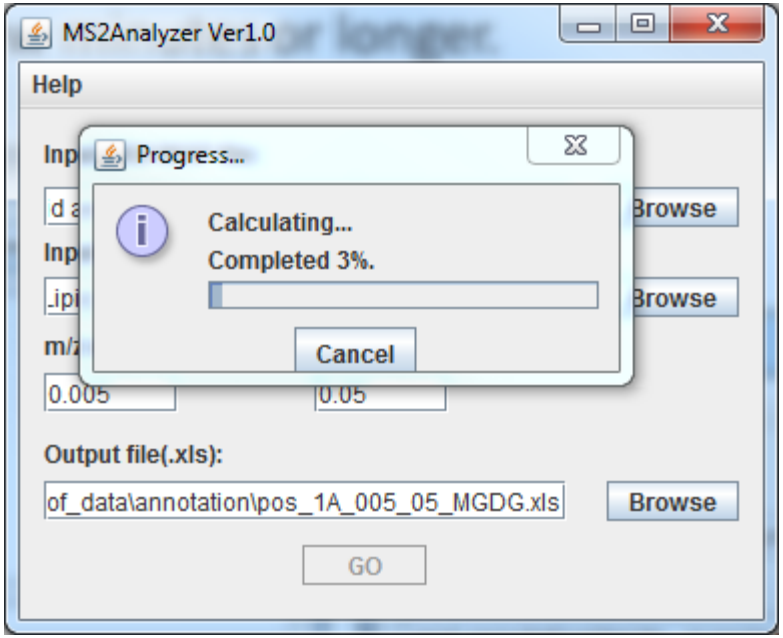


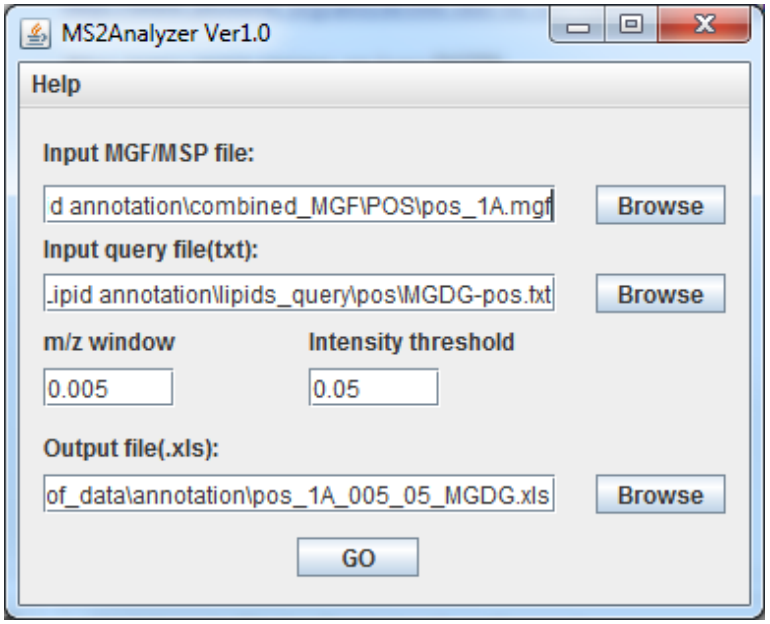


Step 1: Using MS2Analyzer to search the spectra feature

6.Click on “GO” button. The program will start searching mass spectra features in the MGF file and generate a Excel 2003 file.

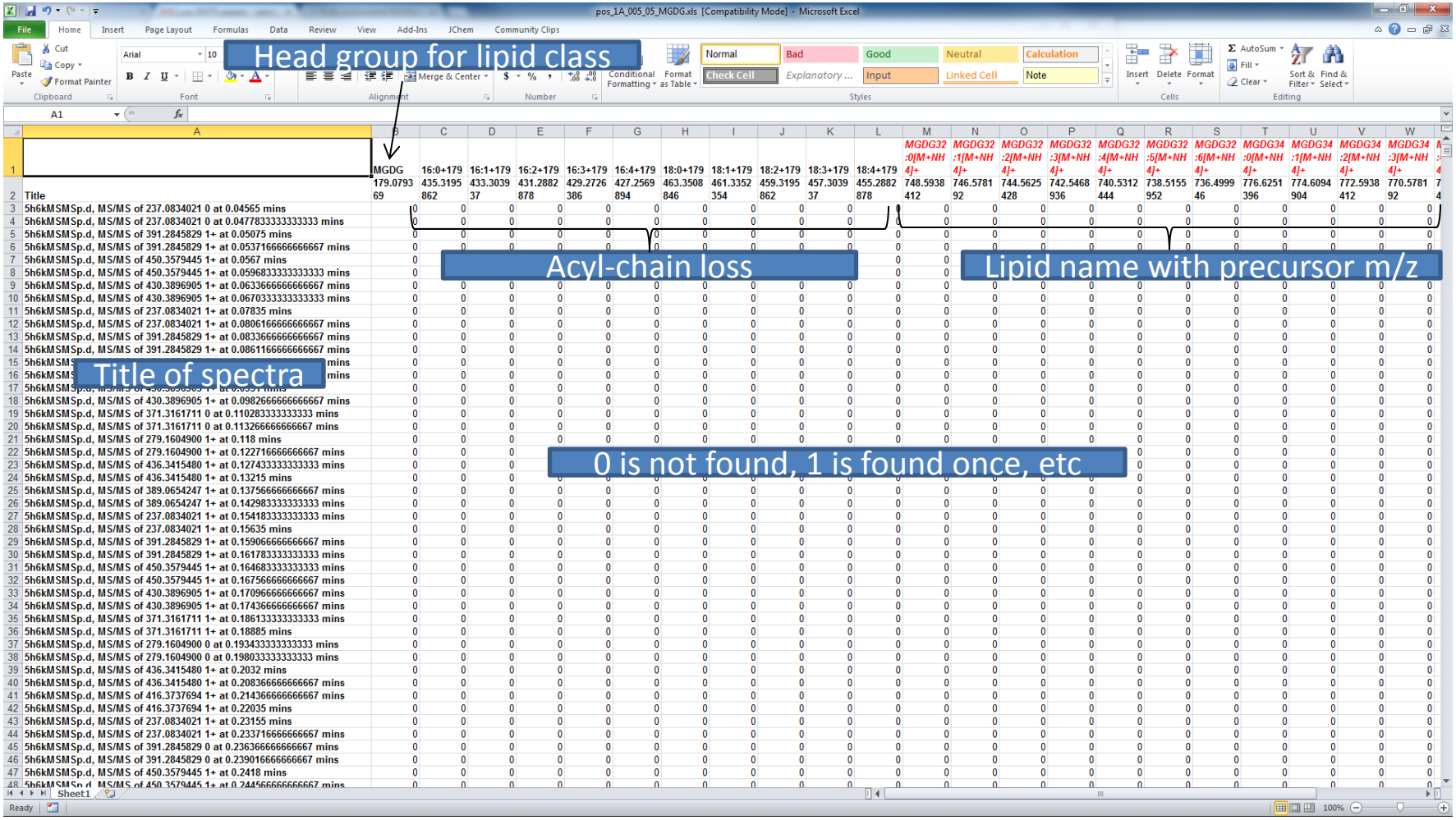
Depending on the speed of the computer, it will take a few minutes or longer. A progress window will show up. When it’s done, the progress window will disappear, and you are ready to move to the next Step!





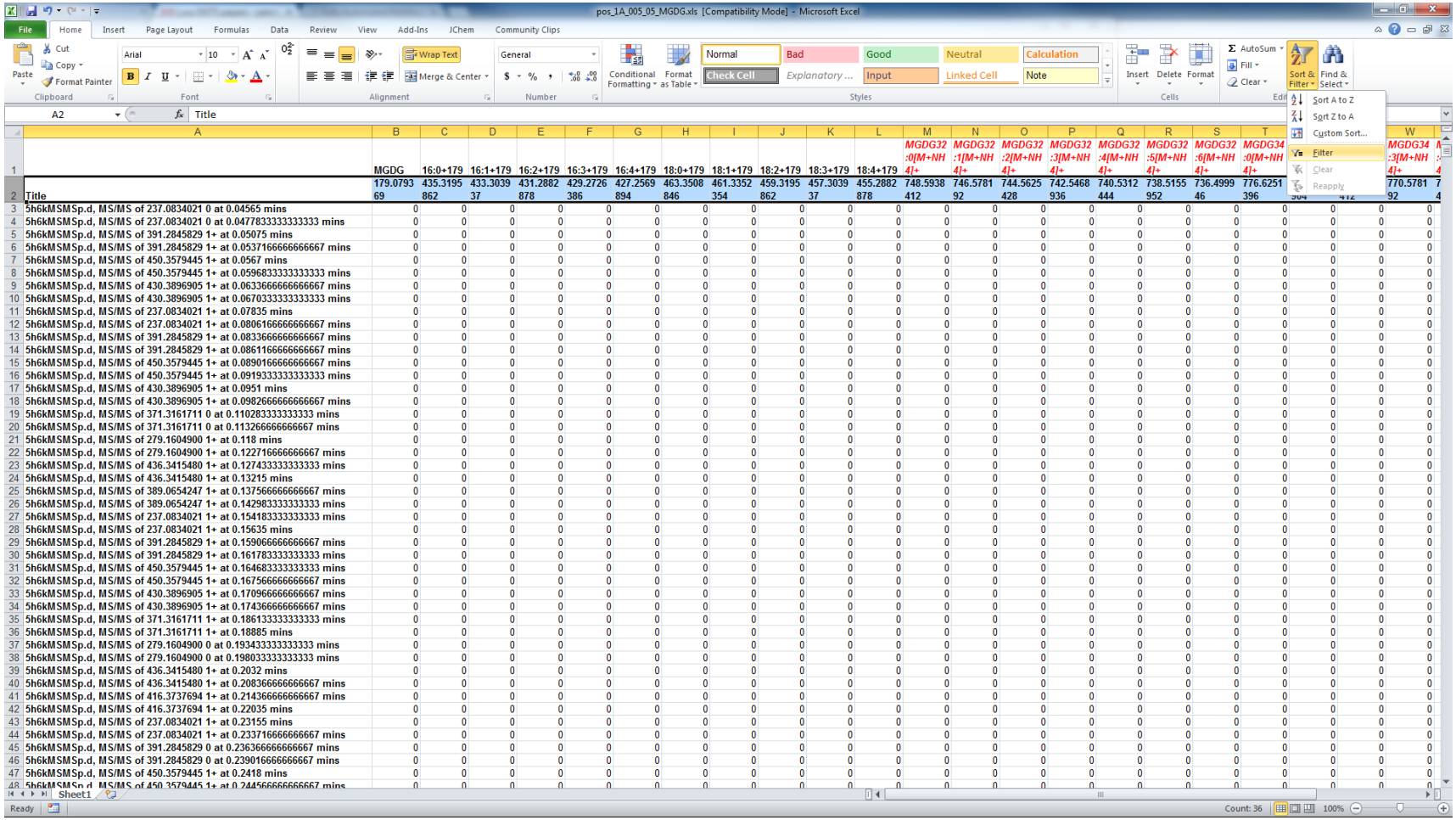
Step 2: Using excel filters to annotate lipids

1.Open the Excel 2003 files you generated from the last Step.



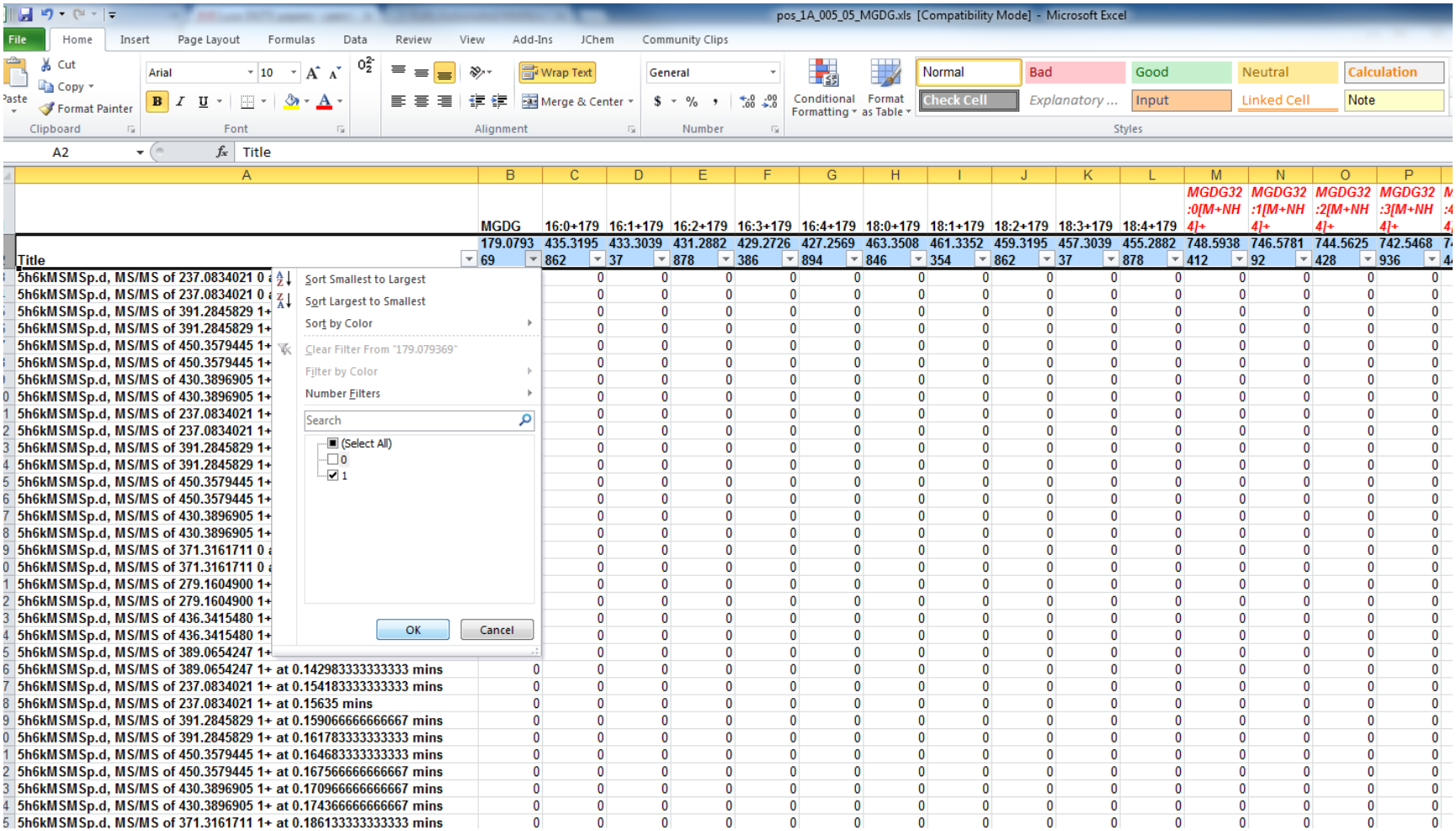
Step 2: Using excel filters to annotate lipids

2.Set filters on the second row with Home-Sort&Filter-Filter (here's a example of excel 2010)



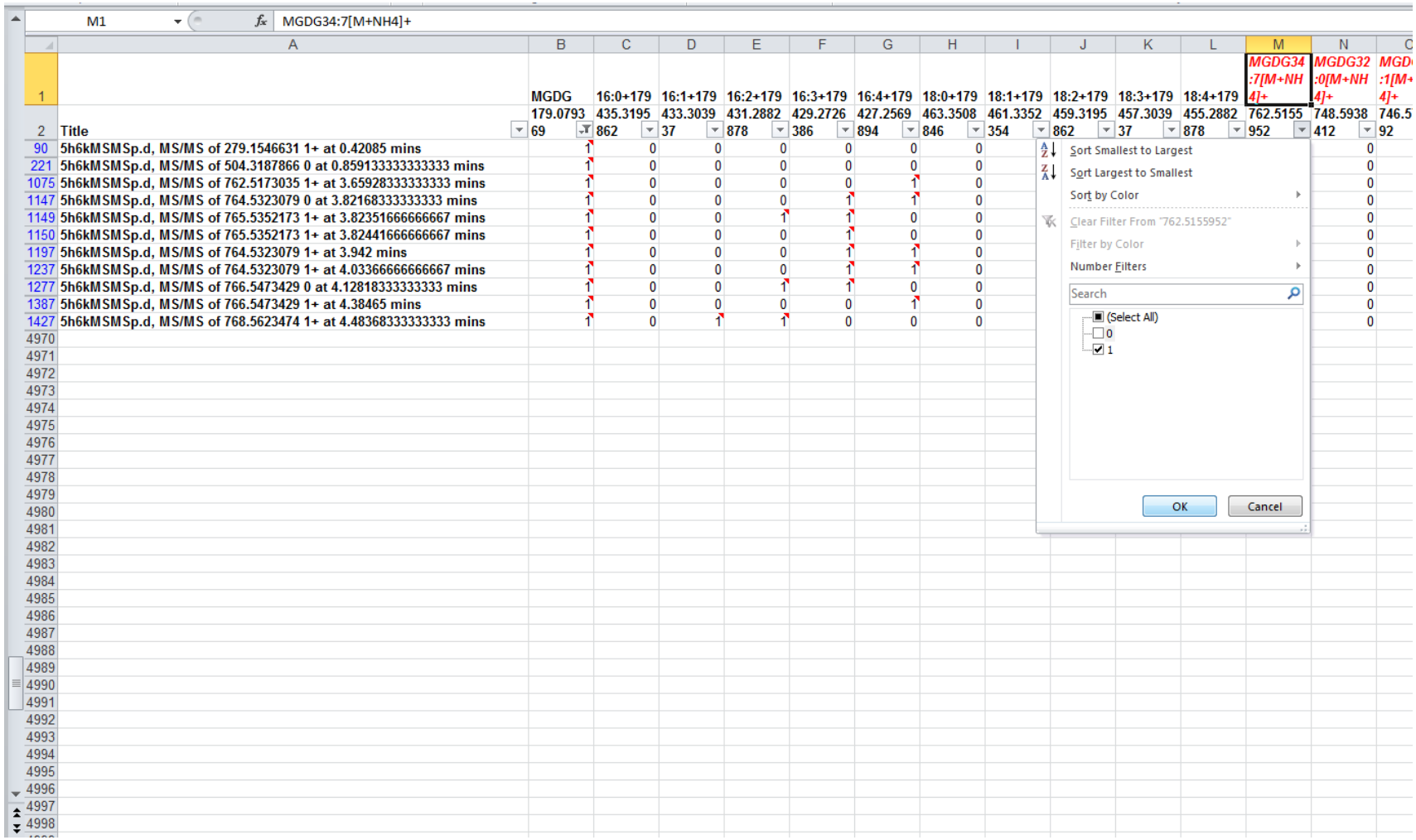
Step 2: Using excel filters to annotate lipids

3. Filter out the 0 values in the head group(has the same name as lipid class, in this case, MGDG)column, only leave the 1 value there.



Step 2: Using excel filters to annotate lipids

4.Click on each filter under the name of lipids(in red), and check if there's 1 value there. If yes, filter out the 0 and show only the 1 value.



	A	B	C	D	E	F	G	H	I	J	K	L	M	N	C
1		MGDG	16:0+179	16:1+179	16:2+179	16:3+179	16:4+179	18:0+179	18:1+179	18:2+179	18:3+179	18:4+179	MGDG34 :7[M+NH 4]+	MGDG32 :0[M+NH 4]+	MGD :1[M+ 4]+
2	Title	179.0793 69	435.3195 862	433.3039 37	431.2882 878	429.2726 386	427.2569 894	463.3508 846	461.3352 354	459.3195 862	457.3039 37	455.2882 878	762.5155 952	748.5938 412	746.5 92
90	5h6kMSMSp.d, MS/MS of 279.1546631 1+ at 0.42085 mins	1	0	0	0	0	0	0						0	
221	5h6kMSMSp.d, MS/MS of 504.3187866 0 at 0.859133333333333 mins	1	0	0	0	0	0	0						0	
1075	5h6kMSMSp.d, MS/MS of 762.5173035 1+ at 3.65928333333333 mins	1	0	0	0	0	1	0						0	
1147	5h6kMSMSp.d, MS/MS of 764.5323079 0 at 3.82168333333333 mins	1	0	0	0	1	1	0						0	
1149	5h6kMSMSp.d, MS/MS of 765.5352173 1+ at 3.82351666666667 mins	1	0	0	1	1	0	0						0	
1150	5h6kMSMSp.d, MS/MS of 765.5352173 1+ at 3.82441666666667 mins	1	0	0	0	1	0	0						0	
1197	5h6kMSMSp.d, MS/MS of 764.5323079 1+ at 3.942 mins	1	0	0	0	1	1	0						0	
1237	5h6kMSMSp.d, MS/MS of 764.5323079 1+ at 4.03366666666667 mins	1	0	0	0	1	1	0						0	
1277	5h6kMSMSp.d, MS/MS of 766.5473429 0 at 4.12818333333333 mins	1	0	0	1	1	0	0						0	
1387	5h6kMSMSp.d, MS/MS of 766.5473429 1+ at 4.38465 mins	1	0	0	0	0	1	0						0	
1427	5h6kMSMSp.d, MS/MS of 768.5623474 1+ at 4.48368333333333 mins	1	0	1	1	0	0	0						0	

Step 2: Using excel filters to annotate lipids

5. Check the 1 values in the acyl chain loss columns(16:0 to 18:4 in this case) and see if the sum of them matches the number in lipids name. If the sum matches, this is a correct hit!

Example 1: 16:4+18:3=34:7 which means this lipid is

MGDG 34:7(16:4/18:3) [M+NH4]+

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
1		MGDG	16:0+179	16:1+179	16:2+179	16:3+179	16:4+179	18:0+179	18:1+179	18:2+179	18:3+179	18:4+179	MGDG34:7[M+NH4]+	MGDG32:0[M+NH4]+	MGDG32:1[M+NH4]+
2	Title	179.0793	435.3195	433.3039	431.2882	429.2726	427.2569	463.3508	461.3352	459.3195	457.3039	455.2882	762.5155	748.5938	746.5781
1075	5h6kMSMSp.d, MS/MS of 762.5173035 1+ at 3.65928333333333 mins	69	862	37	878	386	894	846	354	862	37	878	952	412	92
4970		1	0	0	0	0	1	0	0	0	1	0	1	0	0

Example 2: Sometimes there's only one acyl chain loss found, but maybe that's also a potential hit. In this case, the lipid name is written as

MGDG 34:7(16:4/18:3) [M+NH4]+

While 16:4 is found by program, and 18:3 is back calculated: 34:7-16:4=18:3

	A	B	C	D	E	F	G	H	I	J	K	L	M	N
1		MGDG	16:0+179	16:1+179	16:2+179	16:3+179	16:4+179	18:0+179	18:1+179	18:2+179	18:3+179	18:4+179	MGDG34:7[M+NH4]+	MGDG32:0[M+NH4]+
2	Title	179.0793	435.3195	433.3039	431.2882	429.2726	427.2569	463.3508	461.3352	459.3195	457.3039	455.2882	762.5155	748.5938
1075	5h6kMSMSp.d, MS/MS of 762.5173035 1+ at 3.65928333333333 mins	69	862	37	878	386	894	846	354	862	37	878	952	412
4970		1	0	0	0	0	1	0	0	0		0	1	0

Step 2: Using excel filters to annotate lipids

5. Example 3:Sometimes,there are multiple possibilities…

First spectra:

16:2+18:3=34:5 MGDG 34:5(16:2/18:3) [M+NH4]+

Second spectra:

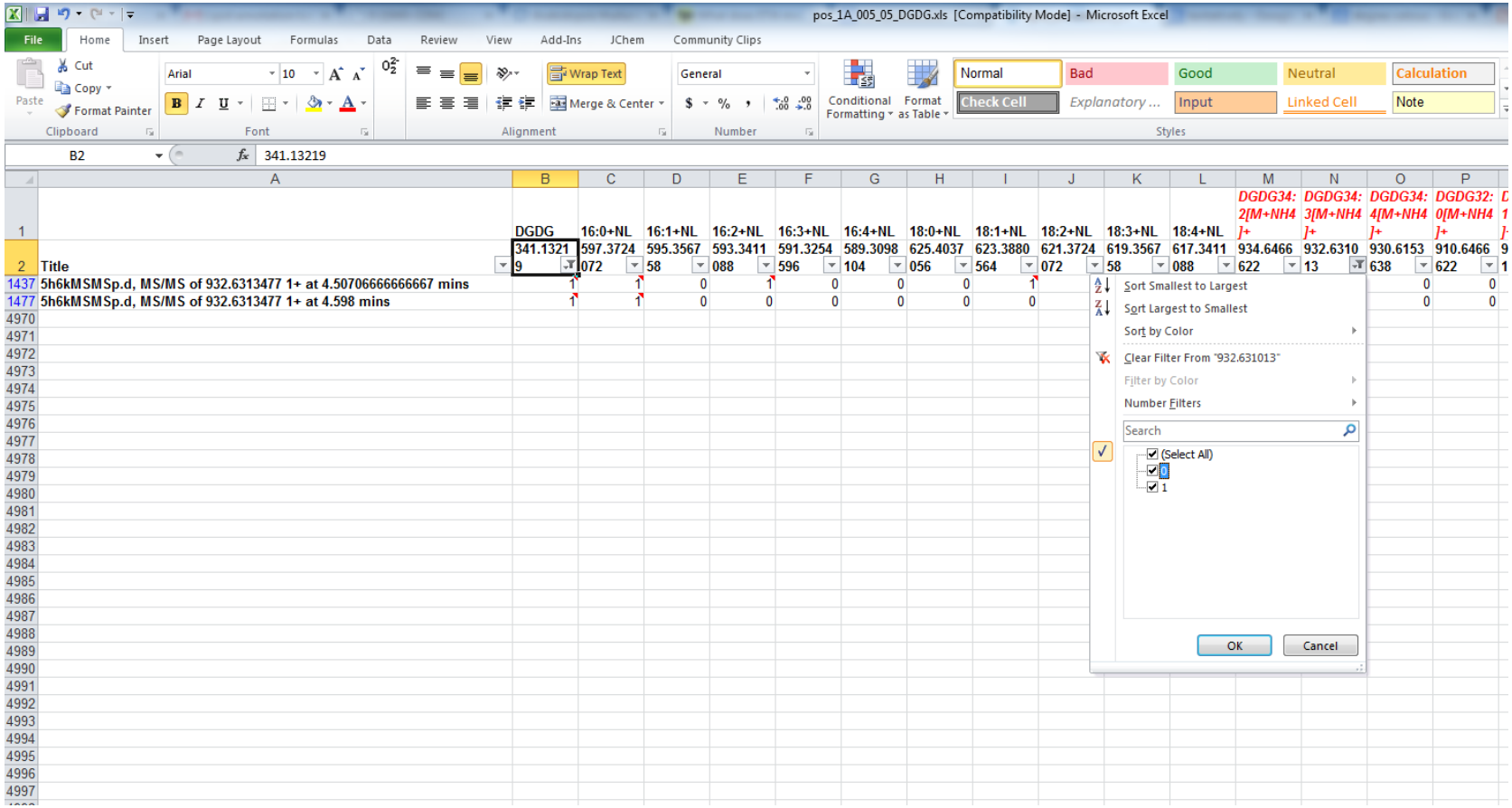
16:4+18:1=34:5 MGDG 34:5(16:4/18:1) [M+NH4]+

	A	B	C	D	E	F	G	H	I	J	K	L	M	N
1		MGDG	16:0+179	16:1+179	16:2+179	16:3+179	16:4+179	18:0+179	18:1+179	18:2+179	18:3+179	18:4+179	MGDG34:5[M+NH4]+	MGDG34:7[M+NH4]+
2	Title	179.0793 69	435.3195 862	433.3039 37	431.2882 878	429.2726 386	427.2569 894	463.3508 846	461.3352 354	459.3195 862	457.3039 37	455.2882 878	766.5468 936	762.5155 952
1277	5h6kMSMSp.d, MS/MS of 766.5473429 0 at 4.12818333333333 mins	1	0	0	1	1	0	0	0	0	1	1	1	0
1387	5h6kMSMSp.d, MS/MS of 766.5473429 1+ at 4.38465 mins	1	0	0	0	0	1	0	1	0	0	0	1	0
4970														
4971														

Both are correct and should be reported separately.

Step 2: Using excel filters to annotate lipids

5. (Continue) After you are done with one lipid precursor m/z , clear the filter by checking the 0 value under its filter again, and move to the next precursor m/z.

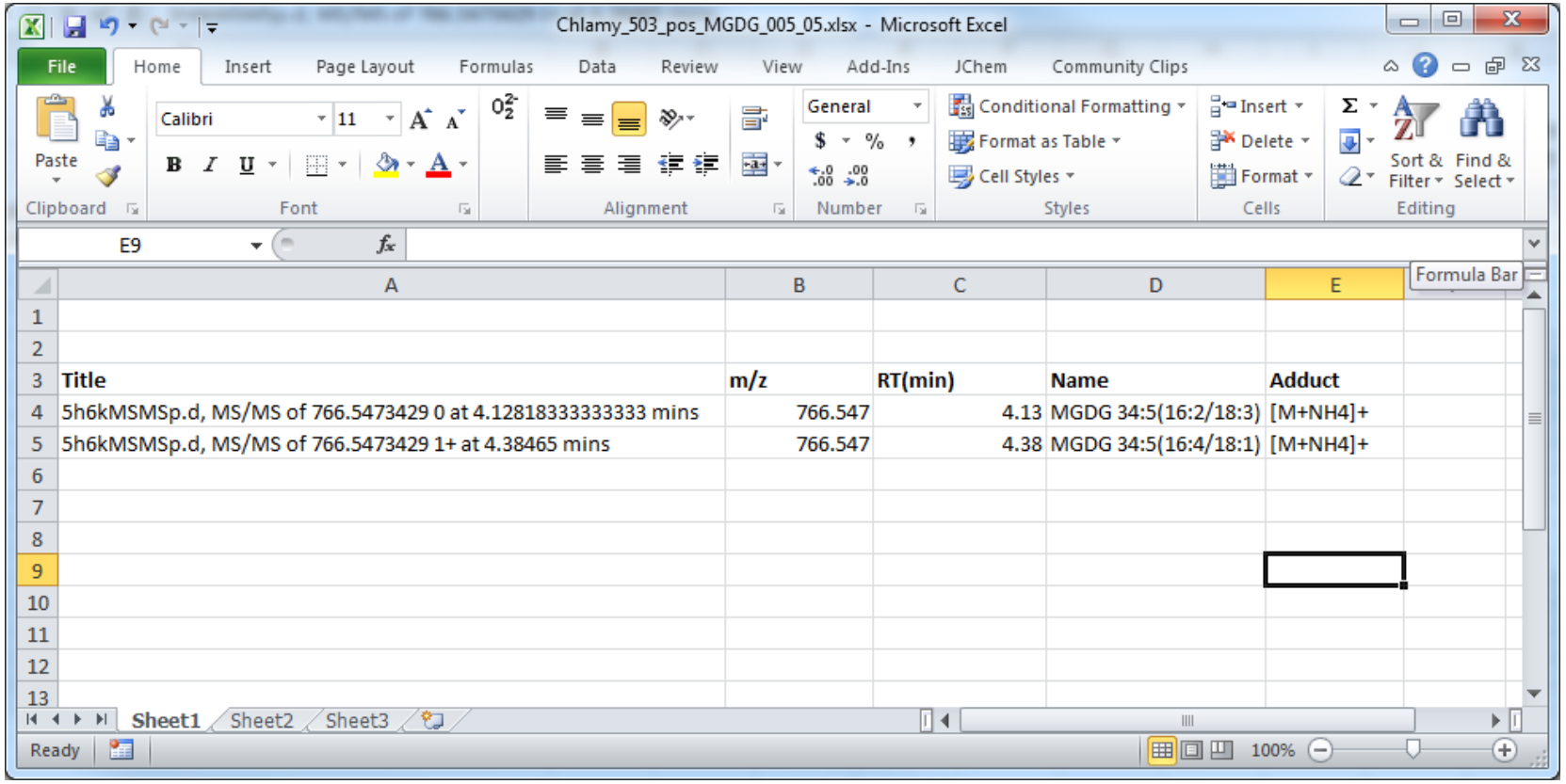


Step 2: Using excel filters to annotate lipids

6. Write the result into a Excel file.

The following is an example of the report:

Report: Title, m/z(with 3 decimals, e.g. 766.547), RT(retention time, with 2 decimals, e.g. 4.13), Name of lipid, and Adduct



	A	B	C	D	E
1					
2					
3	**Title**	**m/z**	**RT(min)**	**Name**	**Adduct**
4	5h6kMSMSp.d, MS/MS of 766.5473429 0 at 4.12818333333333 mins	766.547	4.13	MGDG 34:5(16:2/18:3)	[M+NH4]+
5	5h6kMSMSp.d, MS/MS of 766.5473429 1+ at 4.38465 mins	766.547	4.38	MGDG 34:5(16:4/18:1)	[M+NH4]+
6					
7					
8					
9					
10					
11					
12					
13					

Step 2: Using excel filters to annotate lipids

TIPS:

MGDG, DGDG, DGTS,PE all have head groups and two acyl chains(same as the example).

Lyso-DGTS only has head group and one acyl chain, in this case, the acyl chain loss must be found in order to report.

DG and TG don’t have head groups. Therefore,

DG: only look for the two acyl chain loss(both must be found)

TG: only look for the three acyl chain loss(at least two of them must be found)

Step3: Confirmation by checking retention time

If a RP column is used for the separation, among lipids in each lipid class, the retention times usually follow the following rule:

Retention times increase with the number of carbon, while decrease with the number of double bond.

This can be used as a further confirmation of the annotation.

Step3: Confirmation by checking retention time

For example

RT MGDG 34:0< RT MGDG 36:0

RT DGDG 32:2 >RT DGDG 32:3

Quiz:

RT PE 34:1 ? RT PE 32:4

RT TG 50:4? RT DG 36:4

Step 4:Using NIST search for library validation

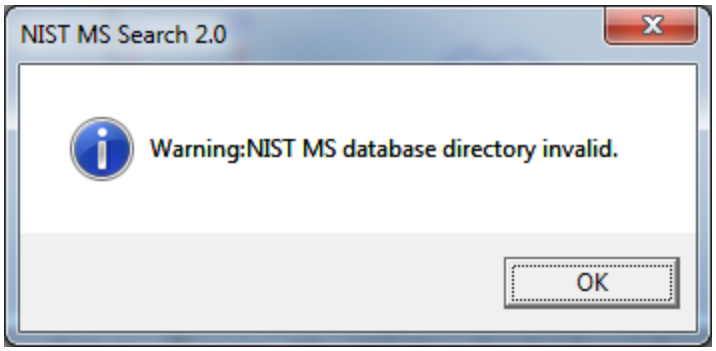
Now you have a list of potential lipid with its title and annotation, you can visualize the spectra in NIST search and search against available libraries to see if the annotation results match.

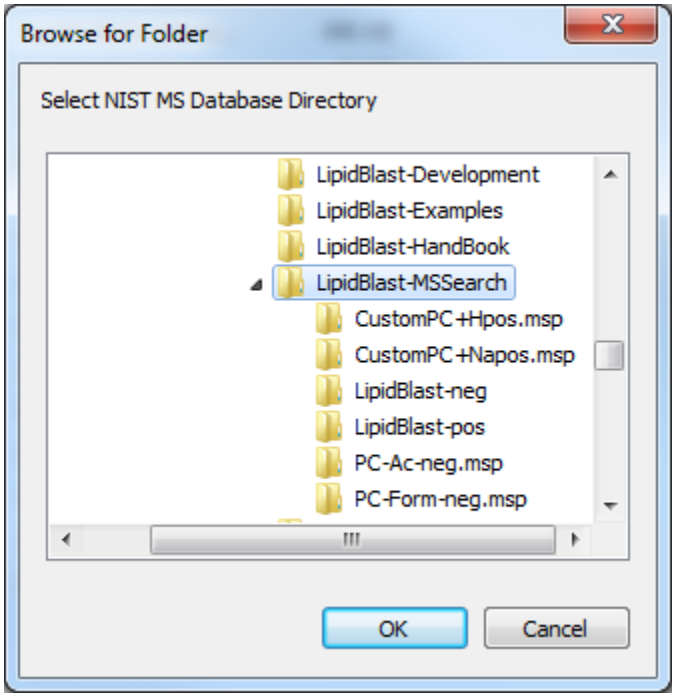
Here we use lipidblast library as an example.

Step 4:Using NIST search for library validation

1. Extract the LipidBlast-Full.zip file, and double click on the nistms.exe file in the LipidBlast-MSSearch folder.

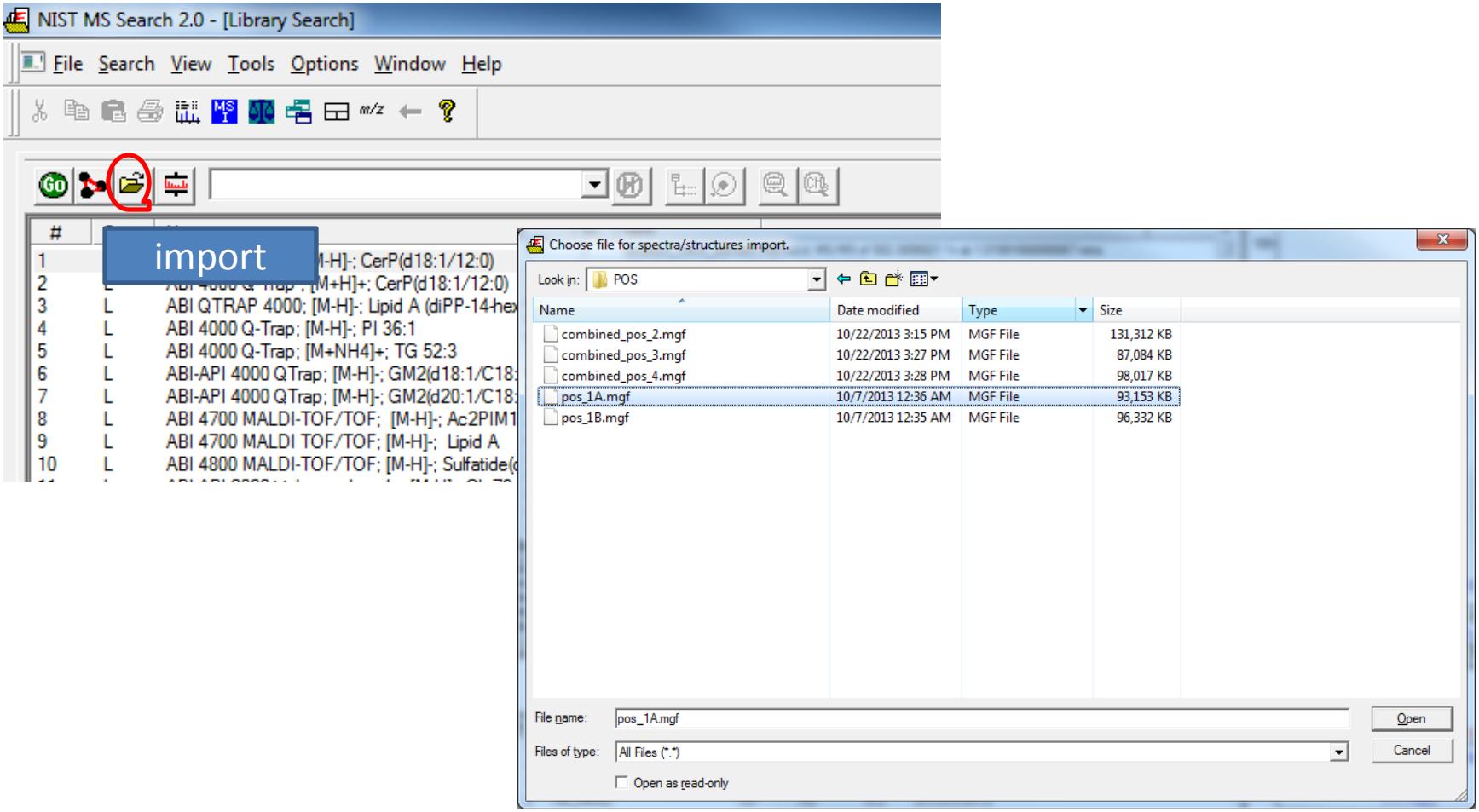
When you open Nist search for the first time, it will give you a warning message. After clicking OK, a window pops up to ask you to select library directory. Here we just keep it highlighting LipidBlast-MSSearch and click OK, then the software will open.





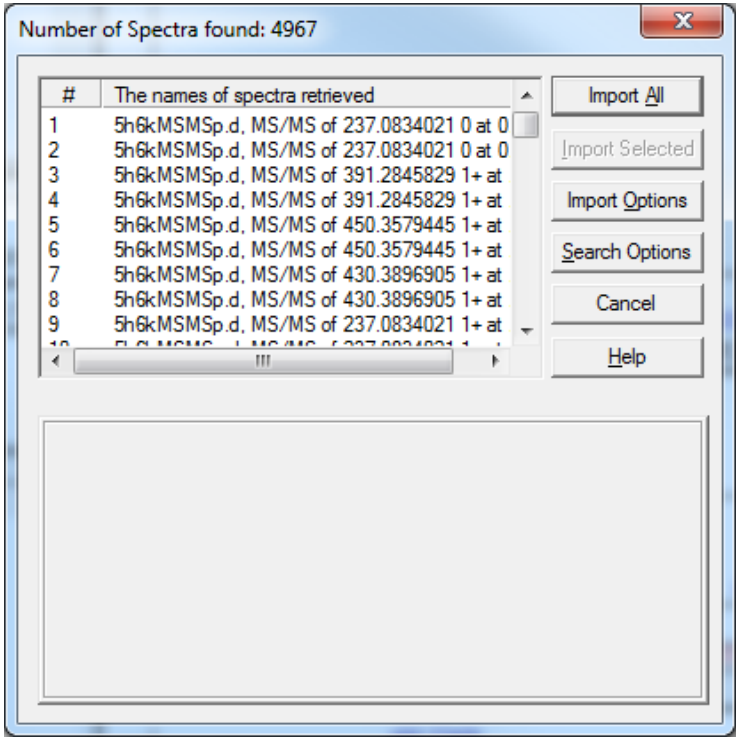
Step 4:Using NIST search for library validation

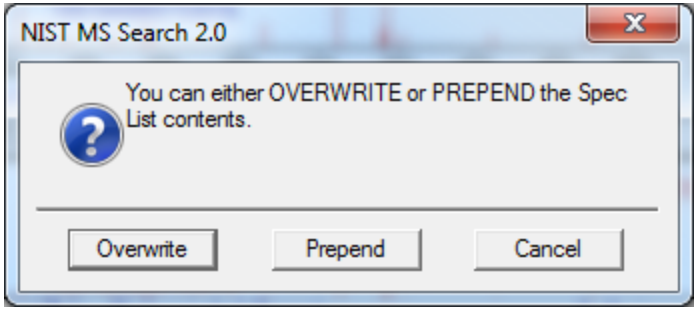
2. Go to the import button, and select the MGF file where you have your lipids in.

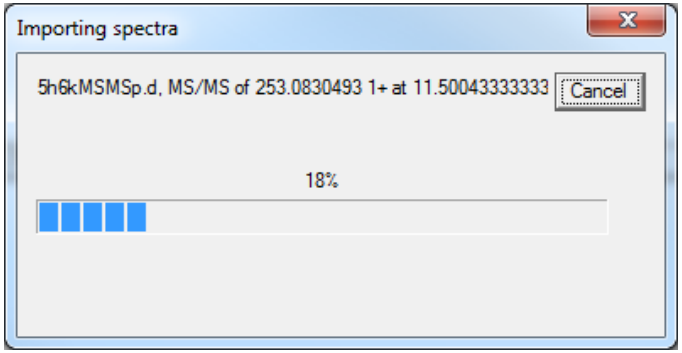


Step 4:Using NIST search for library validation

Click on “Import All” and in the next window, choose Overwrite or Prepend. It’ll take some time to load all the spectra.



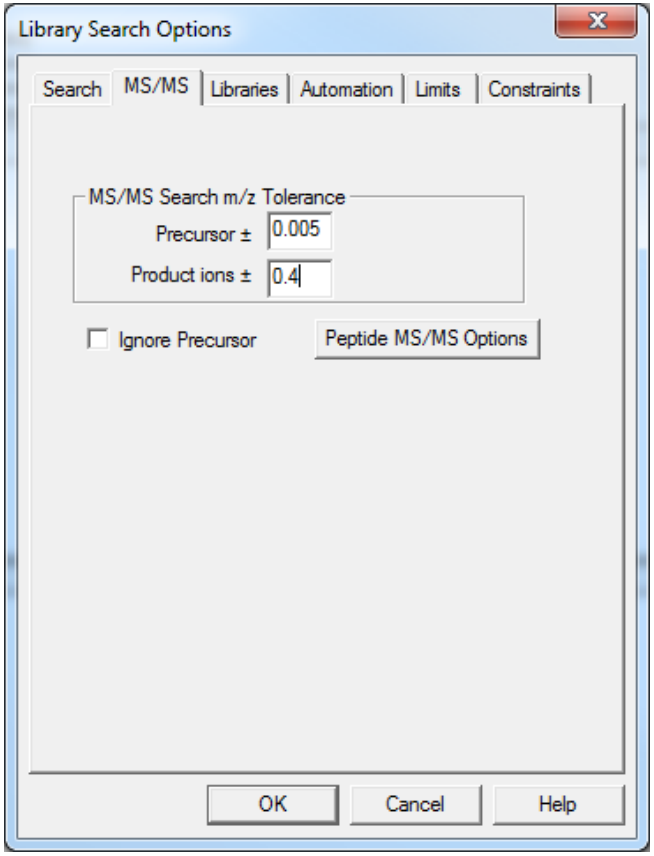


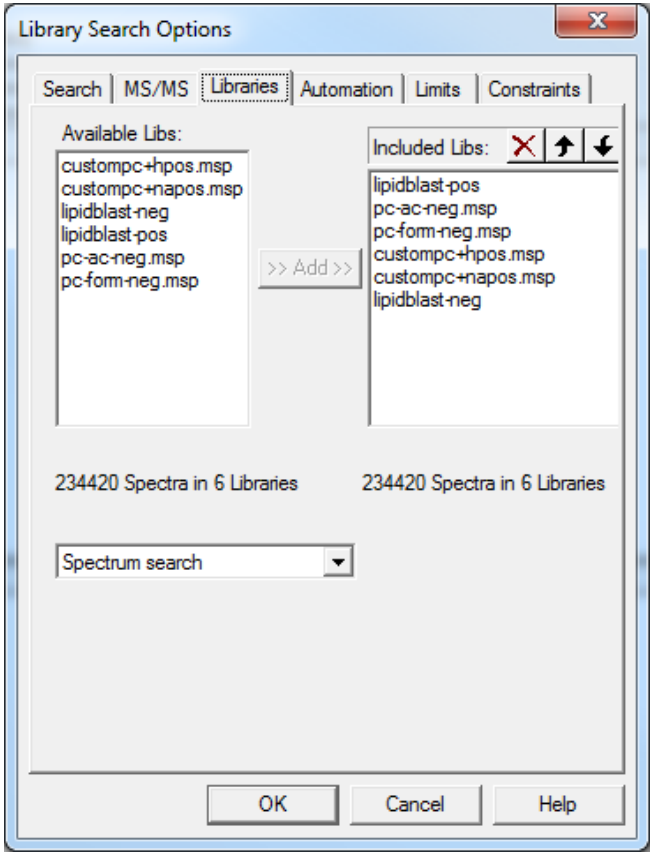


Step 4:Using NIST search for library validation

3. When the importing is done, go to the Library Search Options button, and set appropriate parameters.

For details, please refer to the user manual of lipidblast.



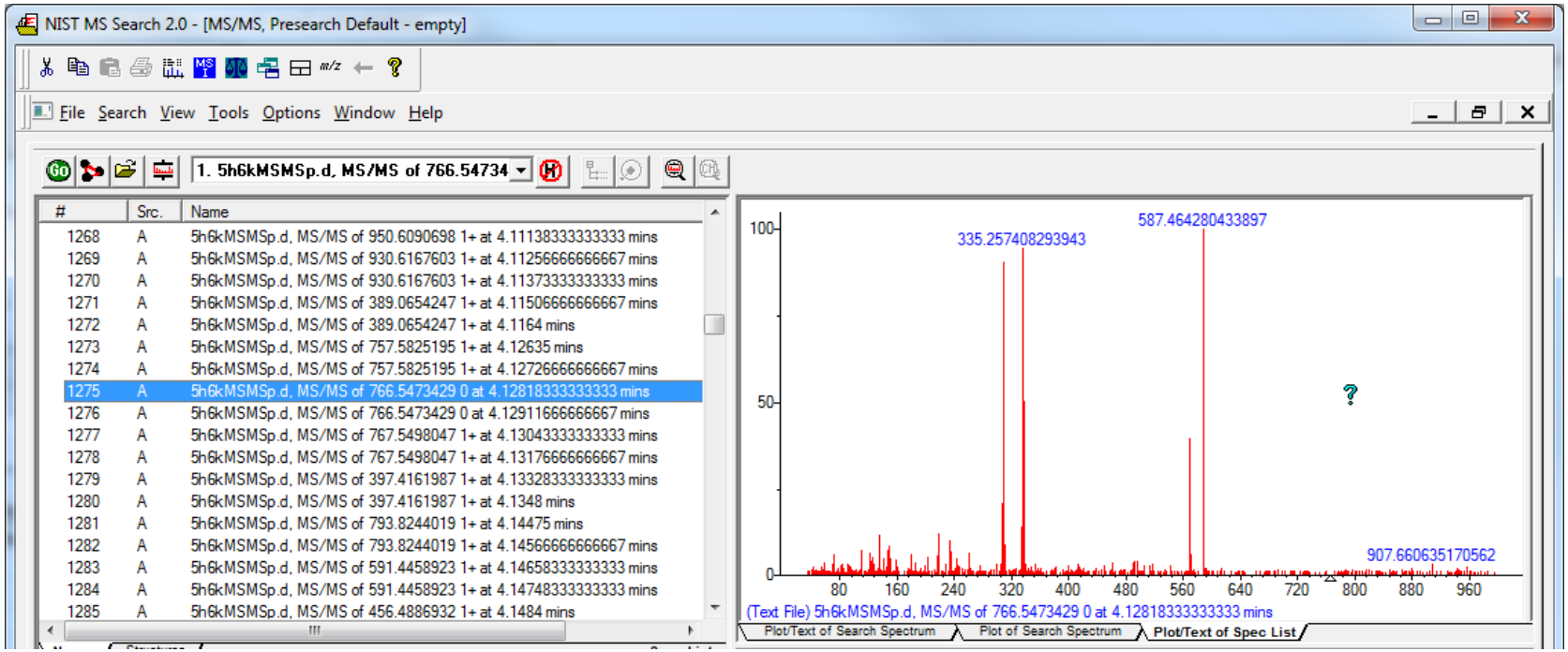


Step 4:Using NIST search for library validation

4. Scroll down in the upper-left window and find your potential lipids by its title(mainly RT, since RT is the increasing order).

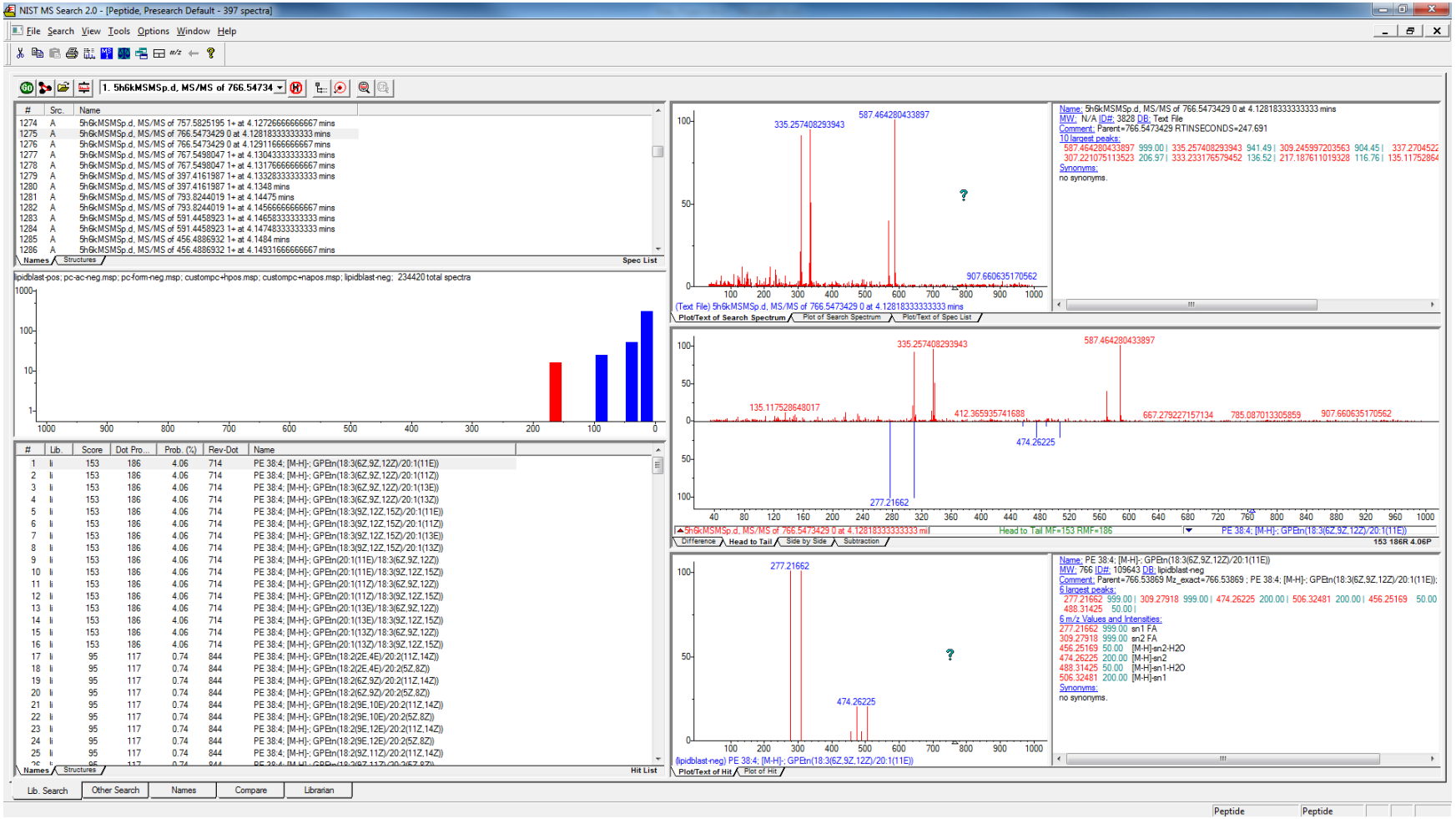
Click on its title to show it spectra in the upper-right window.

Title	m/z	RT(min)	Name	Adduct
5h6kMSMSp.d, MS/MS of 766.5473429 0 at 4.12818333333333 mins	766.547	4.13	MGDG 34:5(16:2/18:3)	[M+NH4]+



Step 4:Using NIST search for library validation

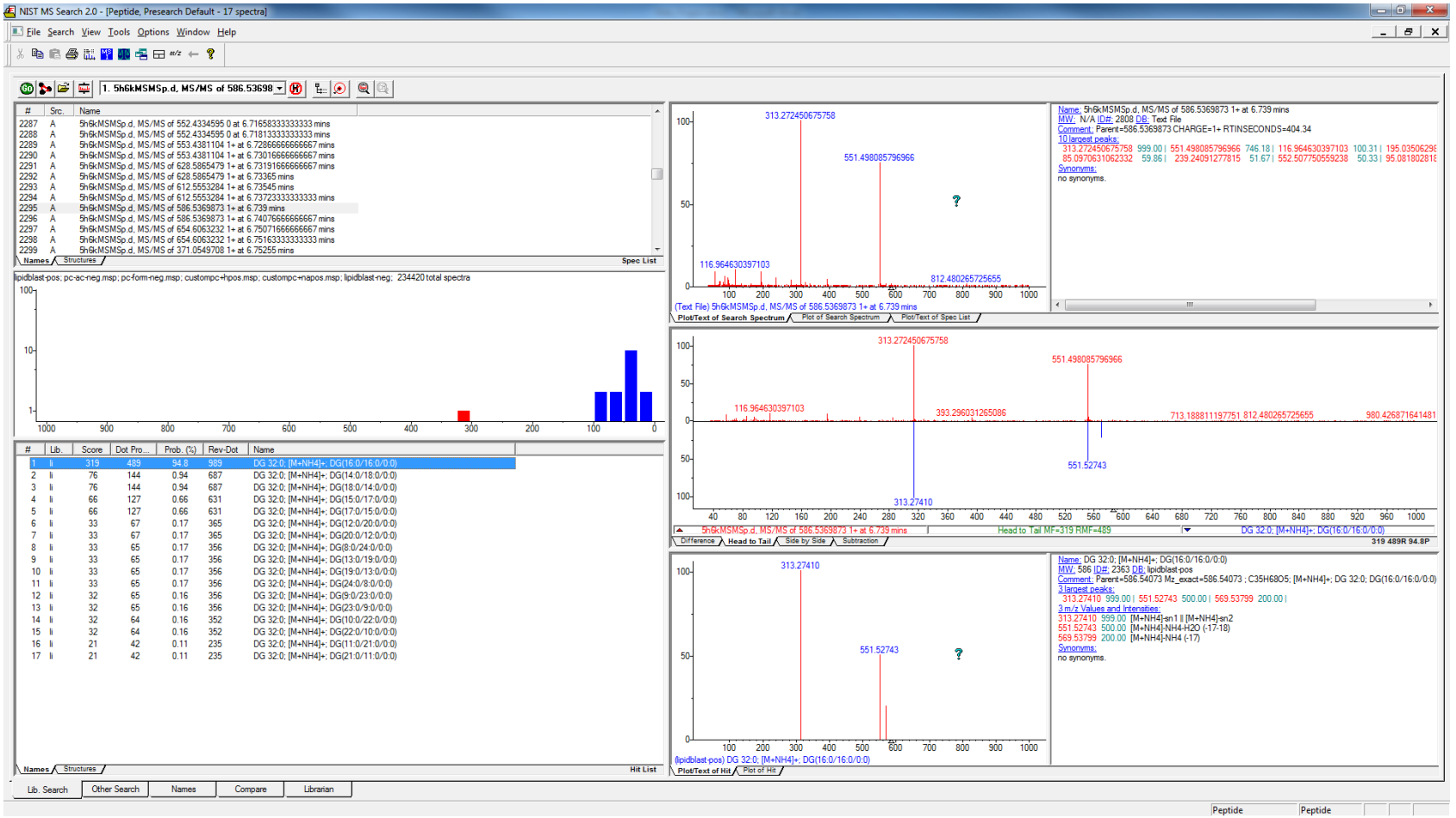
Double click on the title to search against libraries. Sometimes there's no good match in lipidblast because the specific lipid is not covered in lipidblast.



Step 4:Using NIST search for library validation

If the library search result match the MS2Analyzer result, then the annotation is further confirmed. Here's an example of good match.

Title	m/z	RT(min)	Name	Adduct
5h6kMSMSp.d, MS/MS of 586.5369873 1+ at 6.739 mins	586.537	6.74	DG 32:0(16:0/16:0)	[M+NH4]+



That’s it!

Thanks for using MS2Analyzer.

Feel free to email yanma@ucdavis.edu if you have any question and comment.