Application of SimLipid® Software in Lipid Profiling of Secreted Lung Lipids by MSMS^ALL Shotgun Lipidomics

Ningombam Sanjib Meitei[1], Himani Gupta[1], Rajesh Pujari[1], Arun Apte[2], Hong Yin[3], Vladimir Capka[3], David Rowlands[3], Sejal Patel[3], Sandeep Daya[3], Katherine Choy[3]
[1]PREMIER Biosoft, Indore, India; [2]PREMIER Biosoft, CA, US; [3]Novartis Institutes for BioMedical Research, Inc., Cambridge, MA, USA
Corresponding author e-mail: sanjib@premierbiosoft.com

PREMIER Biosoft
Accelerating Research in Life Sciences

NOVARTIS
NOVARTIS INSTITUTES FOR BIOMEDICAL RESEARCH

1

INTRODUCTION

There has been rapid advancement in the field of analytical technology, particularly mass spectrometry that fuels the lipidomics research [1]. One of such MS based techniques was the sequential precursor ion fragmentation, MSMS^ALL, that delivers both qualitative and quantitative information for hundreds of lipid species in a short period of time without compromising data quality[2-3].

We present a comprehensive lipidomics workflow using automated large volume loop injection (FIA) coupled to a TripleTOF® 5600+ mass spectrometer (SCIEX, Concord, ON) with SimLipid software. The lipid species in the sample were identified using data from MSMS^ALL experiment and subsequently quantified using data generated by TOF-MS experiments. SimLipid software has been re-developed to streamline the data analysis from this workflow.

METHODS

Mouse bronchoalveolar lavage fluid samples from lipopolysaccharide treated and control mice were extracted using a modified Bligh-Dyer method. Extracts were dried and reconstituted in 10mM ammonium acetate in a methylene chloride/isopropanol/methanol (2:1:1, v/v/v). Samples were introduced into a TripleTOF 5600+ MS by automated flow injection at 17uL/min for 5.2 min. MSMS^ALL acquisition was one cycle of 801 precursor fragmentation experiments stepping from 400 to 1200 Da with 1 Da selection window with positive electro-spray ionization. For quantification, 252 cycles of TOF-MS scan was performed from 200 to 1200 Da.

DATA ANALYSIS

Data Analysis was done using SimLipid Software v. 6.04 (PREMIER Biosoft, USA) customized for MSMS^ALL workflow, Microsoft (MS) Excel, and Genedata software [4]. Figure 1 shows the schematic representation of the lipidomics workflow using SCIEX's tripleTOF5600+ MS with SimLipid software, and Genedata software.

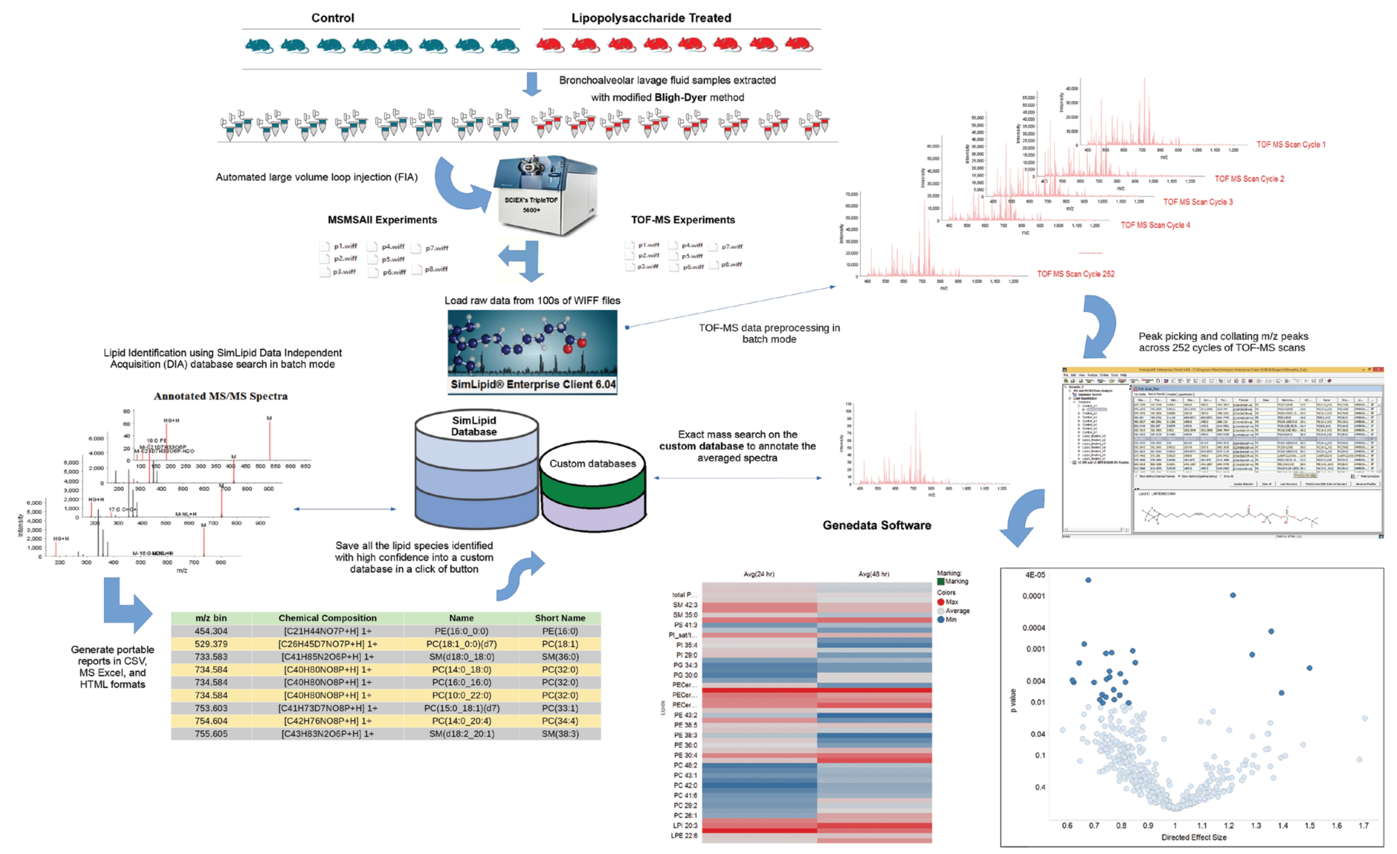

Control
Lipopolysaccharide Treated
Bronchoalveolar lavage fluid samples extracted with modified Bligh-Dyer method
Automated large volume loop injection (FIA)
MSMSAll Experiments
TOF-MS Experiments
Load raw data from 100s of WIFF files
SimLipid® Enterprise Client 6.04
TOF-MS data preprocessing in batch mode
Peak picking and collating m/z peaks across 252 cycles of TOF-MS scans
Lipid Identification using SimLipid Data Independent Acquisition (DIA) database search in batch mode
Annotated MS/MS Spectra
SimLipid Database
Custom databases
Exact mass search on the custom database to annotate the averaged spectra
Save all the lipid species identified with high confidence into a custom database in a click of button
Generate portable reports in CSV, MS Excel, and HTML formats
Genedata Software

Figure 1: Schematic representation of the MSMS^ALL shotgun lipidomics workflow with SimLipid software.

1. RAW DATA LOADING AND DATA PROCESSING

SimLipid software directly read raw data of the MSMS^ALL and TOF-MS experiments from SCIEX's native format file ".wiff". The program can read data from hundreds of .wiff files in batch mode and process the data to generate peaklists.

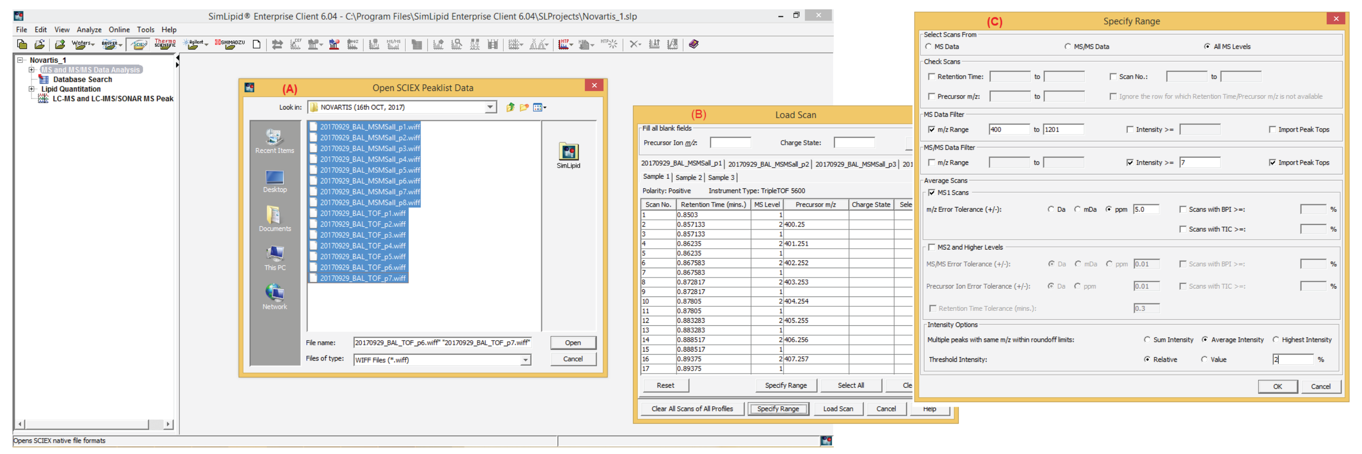


Figure 2: Typical SimLipid software GUIs showing the raw data loading steps: (A) Select the .wiff files, (B) Select the samples you want to read the data from, and (C) Specify filters and data processing parameters.

SimLipid software enables its users to visualize the raw data and results in six different but inter-related windows (Figure 3): (A) Project management panel; (B) Table View of the Data; (C) Chromatogram; (D) Spectrum; (E) Result; (F) Lipid Information. The inter-relationship between the windows are explained as follows. The project management panel displays all the raw data files and peaklists. Upon selecting a data file in (A), windows (B)-(F) are updated. Selection of a particular row of the table (A), or a time point in the Chromatogram, (B) will update the content of (D)-(F). Selection of a particular row of the table (E) will update the content of (F).

The program facilitates raw data, peaklists, and results to be visualized, reviewed, and re-processed using different parameters without having to re-load the data files, all in a single workbench.

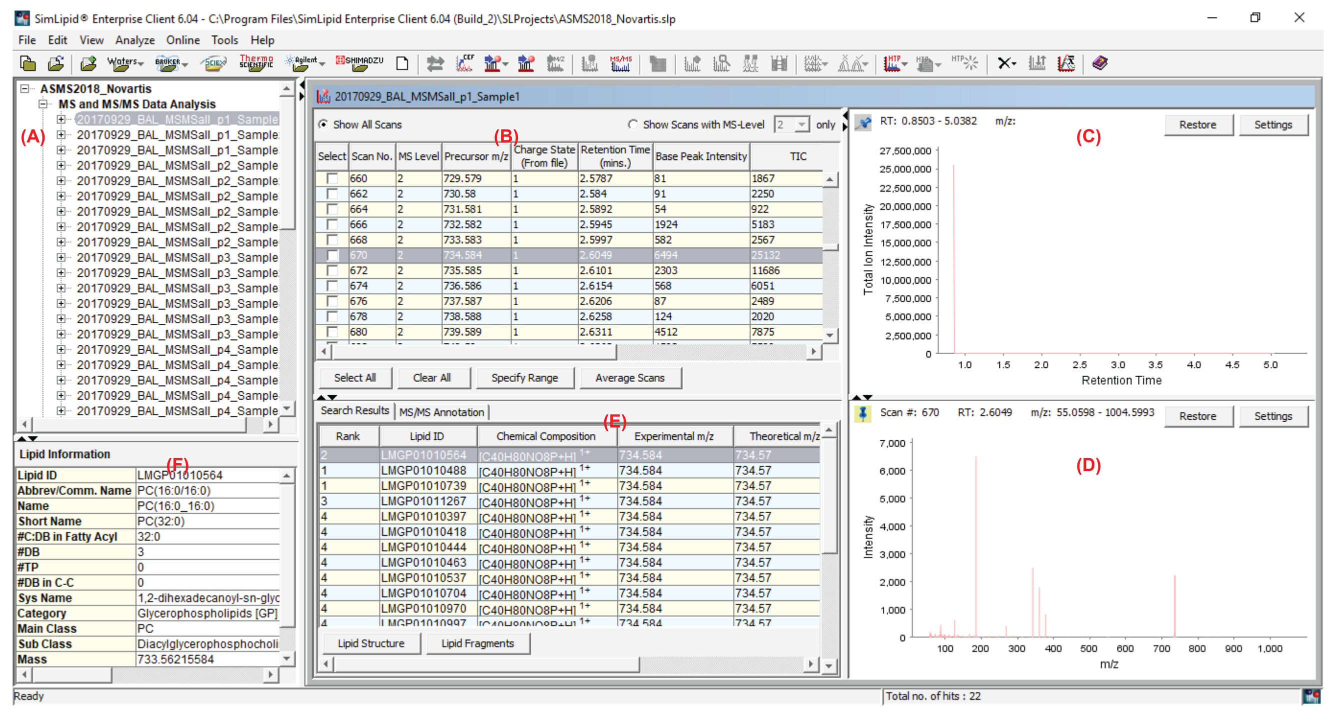


Figure 3: Typical SimLipid software GUI providing a single workbench solution to visualize raw data, peaklists, and results.

2. LIPID IDENTIFICATION USING MSMS^ALL DATA

In this workflow, the Q1 quadrupole transmits a wide mass range, 1 amu in our case, containing multiple analytes through to the collision cell. The transmitted ions are fragmented and the resulting fragments are analyzed in the TOF-MS Analyzer at high resolution. This produces a complex MS/MS spectrum which is a composite of all the analytes within that Q1 m/z window.

The MSMS^ALL data was subjected into SimLipid search for automated identification of lipids using product ion data. The database search was performed for PC, SM, PE, PS, PG, and PI classes (±0.5 Da for precursors, ±0.01 Da for product ions) to identify probable lipids using MS/MS spectra from sequential precursor ion fragmentation (unit m/z bin). [M+H+] was designated to PC, PE, SM, PS, and PI; [M+NH4+] was designated to PG, and PI. A total of 452 unique lipids were identified: 186 PC, 109 PE, 47 PG, 43 PI, 40 PS, and 27 SM.

Figure 4 shows the schematic representation of SimLipid identifying multiple lipid species belonging to different sub-classes using product ions data from MS/MS spectrum at m/z bin 734.584. SimLipid identified 5 isobaric compounds (Table 1) with different 37 lipid species and 13 distinguishable structures from the MS/MS spectrum.

Short Name	# Isobaric Lipids	Exact Mass	Parent Ion Species
SM(35:1)	2	716.5832	NH4+
PC(32:0)	14	733.56216	H+
PC(31:4)	2	711.48391	Na+
PE(35:0)	11	733.56216	H+
PA(37:1)	8	716.535607	NH4+

Table 1: List of 5 isobaric lipids identified using product ions data from the MS/MS spectrum at m/z bin 734.584.

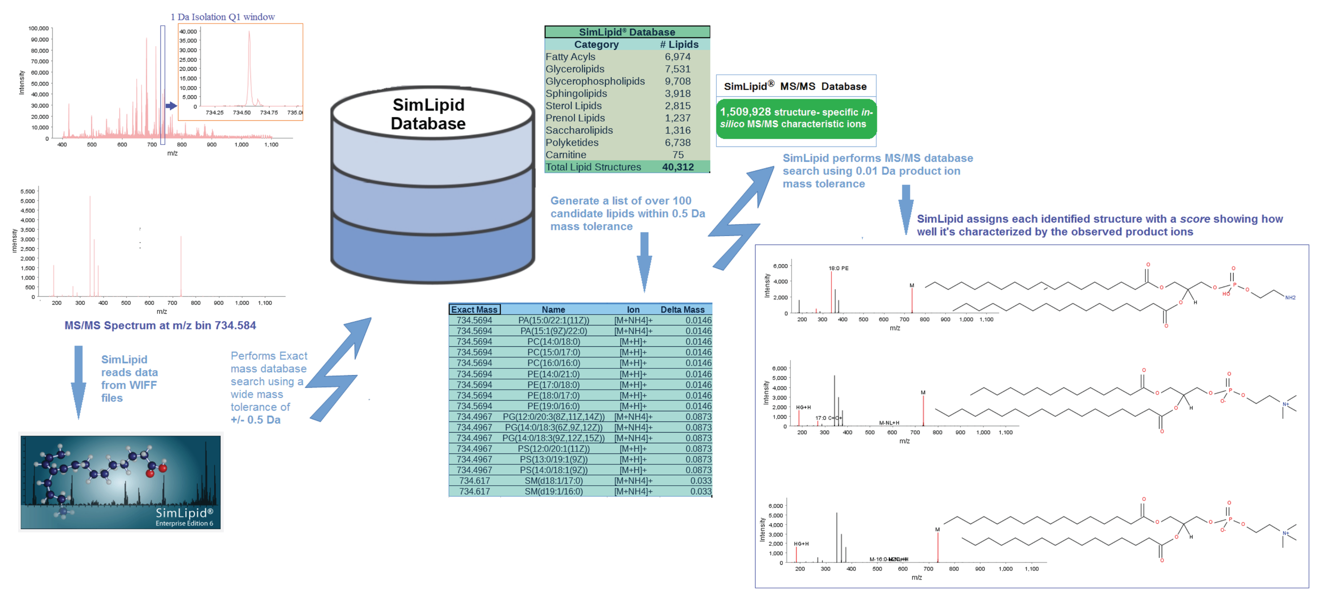

1 Da Isolation Q1 window
MS/MS Spectrum at m/z bin 734.584
SimLipid reads data from WIFF files
Performs Exact mass database search using a wide mass tolerance of +/- 0.5 Da
SimLipid Database
SimLipid® Database
Category
Lipids
Fatty Acyls
6,974
Glycerolipids
7,531
Glycerophospholipids
9,708
Sphingolipids
3,918
Sterol Lipids
2,815
Prenol Lipids
1,237
Saccharolipids
1,316
Polyketides
6,738
Carnitine
75
Total Lipid Structures
40,312
SimLipid® MS/MS Database
1,509,928 structure- specific in-silico MS/MS characteristic ions
Generate a list of over 100 candidate lipids within 0.5 Da mass tolerance
SimLipid performs MS/MS database search using 0.01 Da product ion mass tolerance
SimLipid assigns each identified structure with a score showing how well it's characterized by the observed product ions

Figure 4: Schematic representation of SimLipid software interpreting a complex MS/MS spectrum containing product ions from multiple lipids from a MSMS^ALL shotgun lipidomics workflow with Q1 m/z window of 1 amu.

2

Figure 5 (A), and (B) show the MS/MS database search parameters, and the annotated MS/MS spectra with the top 5 distinguishable lipid species.

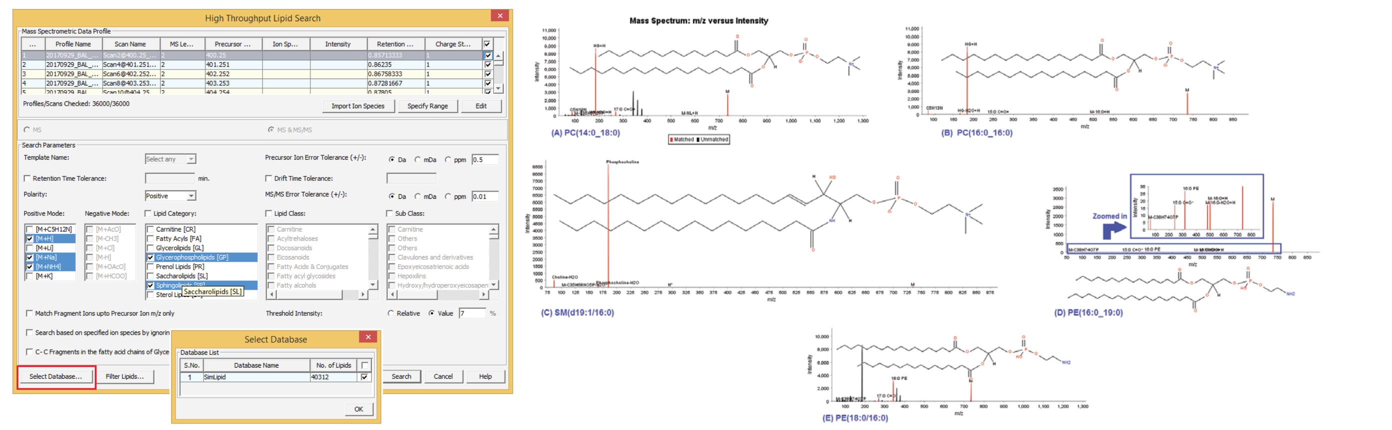


Figure 5: Typical SimLipid software GUIs showing (A) MS/MS database search parameters; and (B) Five identified lipid species using product ions of MS/MS spectrum at m/z bin: 734.584.

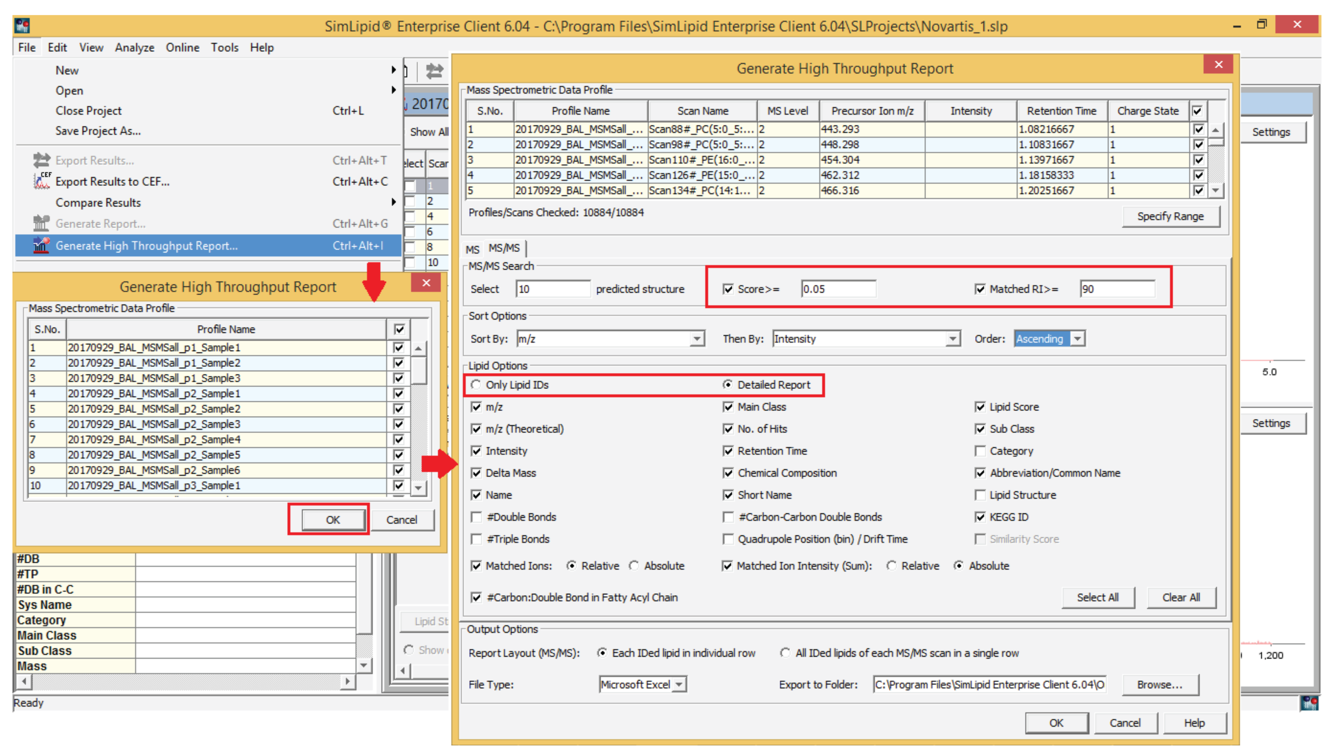


Figure 6: Typical SimLipid software GUIs that facilitates unique lipid IDs that satisfy the filter criteria based on Score, and Matched RI.

3. RESULTS FILTERS AND CREATE LIPID SUB-DATABASE

A custom database of these lipids was created in the SimLipid database by exporting the unique lipid IDs that satisfy the Score>=0.05 and Matched RI filter>=90% (Figure 6 and Figure 7).

4. LIPID QUANTITATION

TOF-MS data of two biological groups, each consisting of 7 replicates, and each replicate containing 252 TOF-MS scans were imported into the program. For each replicate, data from all TOF-MS scans were averaged using 5 ppm tolerance. Averaged spectrum from each replicate was annotated by performing exact mass search within 5 ppm error on the custom database containing 454 unique lipid species identified using MSMS^ALL data. Overlapping isotopic peaks were corrected using algorithm explained in [5]. Lipids annotated across the biological replicates were mapped based on their short names, e.g., PC(32:1) for PC(14:0/18:1(9Z)).

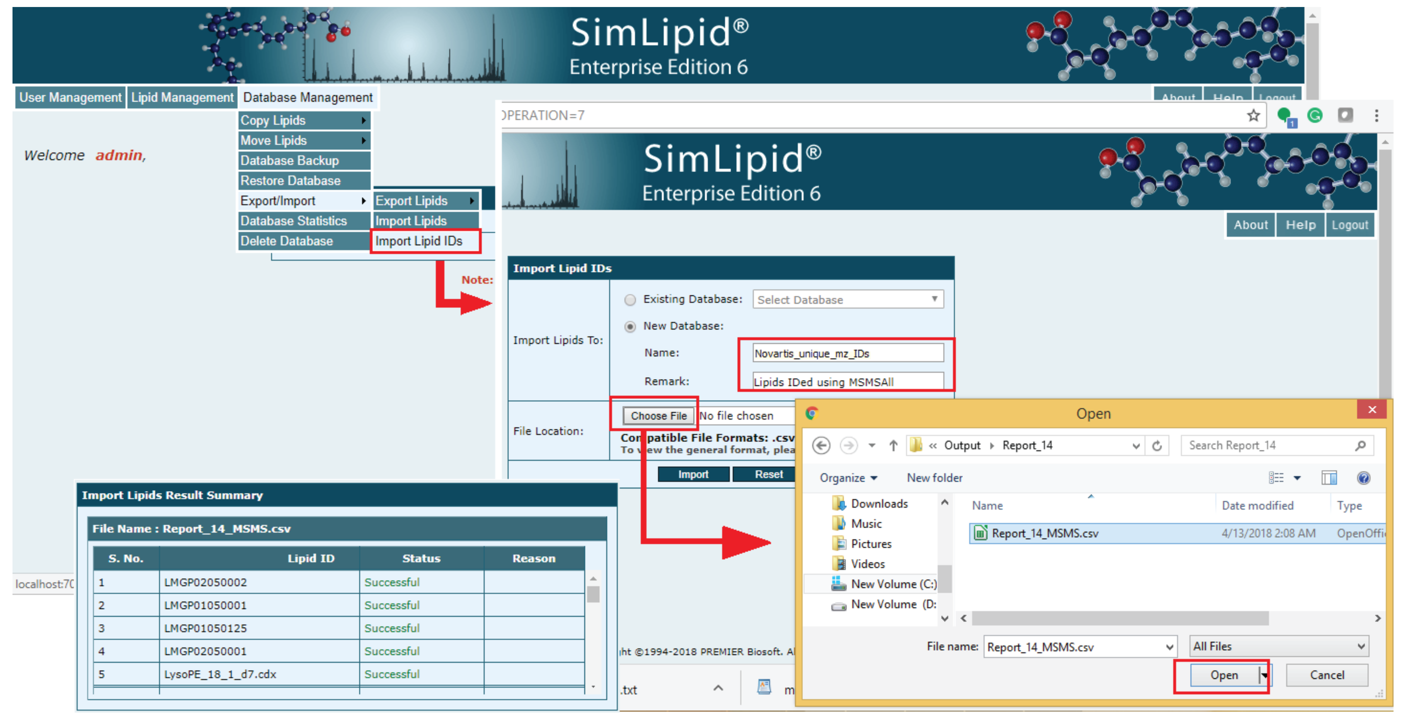


Figure 7: Typical SimLipid software GUIs that facilitates users to directly create a custom database of unique lipid IDs that satisfy the filter criteria based on Score, and Matched RI.

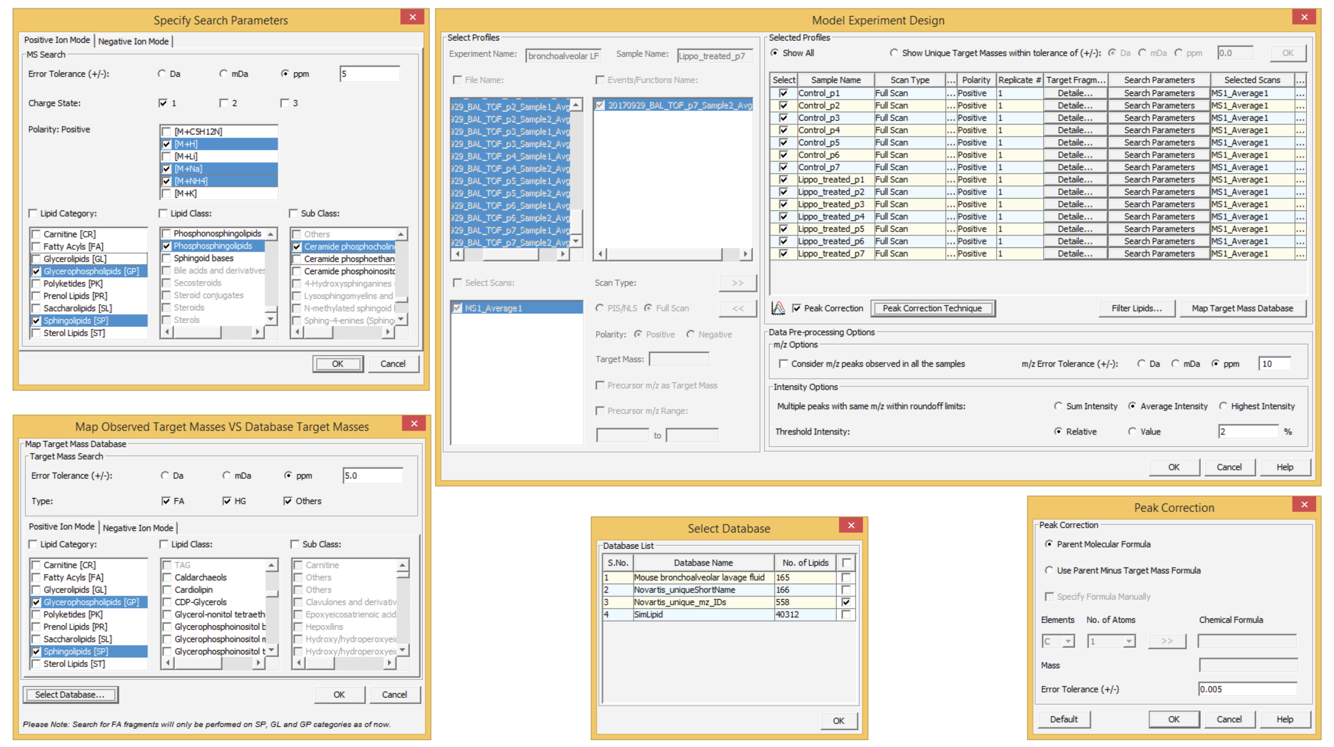


Figure 8: Using the TOF-MS data - Typical SimLipid software GUIs that facilitates users to model experimental design, map peaklists across biological replicates, perform exact mass search on custom database, and instruct the program to perform isotopic peak correction.

5. STATISTICAL ANALYSIS

Results were exported into MS excel format file that can be imported into Genedata software [4] for further statistical analysis. Using this approach we were able to identify lipid biomarkers in mouse lung in response to lipopolysaccharides treatment.

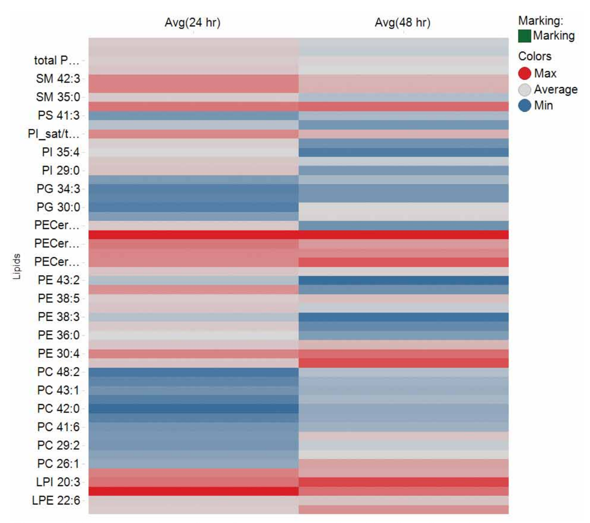


Figure 9: Heat map from GeneData (as visualized by Spotfire) showing treatment-affected lipids across treatment groups.

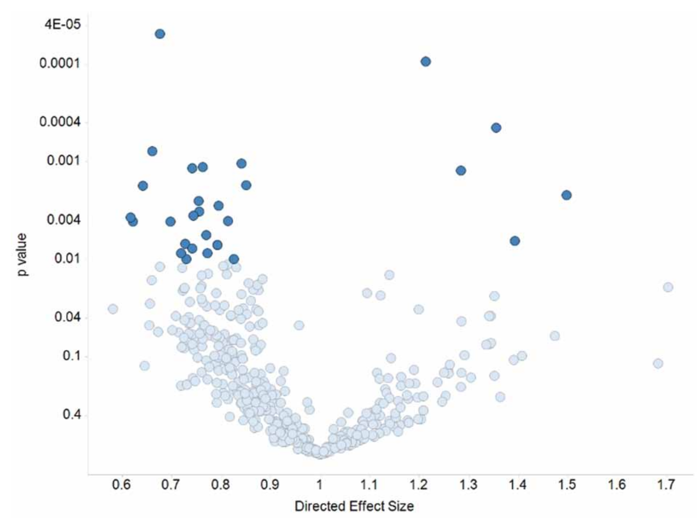


Figure 10: Volcano plot showing statistical significance of the effect size of each lipid in 24hr treatment group vs control. Highlighted data are potential lipid biomarkers showing a significant change between the control and the 24hr treatment group at p<0.05.

CONCLUSION

Using the Infusion MSMS^ALL workflow combined with SimLipid software, we were able to identify potential secreted lung lipid biomarkers upon stimulation by lipopolysaccharide treatment. This workflow can be subsequently used to design further targeted experiments to validate potential biological hypotheses relevant to a disease model or a treatment intervention.

SimLipid software alleviates key shortcoming of MSMS^ALL workflow: Rapid and accurate identification of lipid species for MS/MS spectra containing product ions from multiple precursors. However, users need to realize this is a fundamental feature of Infusion MSMS^ALL data and further interrogate results based on the MS/MS score and sum of relative intensities of the characteristic ions of the identified lipid species assigned by the program as appropriate to their needs.

SimLipid software provides the following fey features that support MSMS^ALL shotgun analysis:

1. Directly read raw data from hundreds of .wiff files in batch mode.
2. Collate data from multiple scans and perform peak picking for hundreds of WIFF files in batch mode.
3. Identify lipid species using shotgun-MS/MS spectra containing product ions from multiple analytes acquired using data independent acquisition method.
4. Create custom database of lipid species that were identified with high confidence using MSMS^ALL data.
5. Model experimental design facilitating differential lipidomics analysis by annotating lipids from custom database and map them across biological replicates.
6. Export results into HTML, MS Excel, spreadsheet compatible format files facilitating easy downstream analysis using third party software tools.

REFERENCES

1. Köfeler et al, (2012). Metabolites, 2, 19-38.
2. Simons et al, (2012). Metabolites, 2(1), pp.195-213.
3. Jung et al, (2011). Biochimica et Biophysica Acta (BBA)-Molecular and Cell Biology of Lipids, 1811(11), pp.925-934.
4. Lempiäinen et al, (2018). Scientific Reports, 8(1):3434.
5. Liebisch et al, (2004). Biochim. Biophys. Acta. 1734:86–89.