Lipid Profiling of Malaria Samples Using Orbitrap Velos Pro Mass Spectrometer with SimLipid Software

Ningombam Sanjib Meitei[1,2]; Himani Gupta[2]; Fatima Rahlouni[3]; David J. Sullivan[4]; Vladimir Shulaev[3]
[1]PREMIER Biosoft, Palo Alto, CA; [2]PREMIER Biosoft, Indore, India; [3]University of North Texas, Denton, Texas; [4]Bloomberg School of Public Health, Johns Hopkins University, Baltimore, MD

1

INTRODUCTION

Rapid advancement in the technologies related to liquid chromatography (LC) and mass spectrometry (MS) leads to development of new lipidomics methods. We have developed LC-MS lipidomics method based on UHPLC coupled with high resolution Orbitrap Velos Pro MS instrument for lipid profiling of *Plasmodium berghei* samples. Lipid identification was achieved using MS/MS data from both positive and negative ion modes and quantitation of lipid species was performed using ion abundance corresponding to its monoisotopic peak of the isotope cluster and statistical analysis. We performed profiling of *P. berghei* wild type and hemolysin III knockout strains to identify significant lipid species that discriminate these strains.

METHODS

Lipids were extracted from *Plasmodium berghei* wild type strains or hemolysin III knockout ANKA (PbHlyIII KO) isolate using Folch method. Two independent experiments using different KO isolates were performed with 3 biological replicates per wild type or KO strain each resulting in total of 12 samples.

LC-MS: LC-MS analysis was carried out on Accela 1250 quaternary pump with Accela Open autosampler on-line coupled to an LTQ Orbitrap Velos Pro Hybrid MS (Thermo Fisher Scientific, USA). *Plasmodium* lipid extracts were separated on an Accucore C18 2.1x150 mm 2.6 µm column using 30 min gradient [1].

Data Analysis: SimLipid® software (PREMIER Biosoft) was used for lipid identification using MS/MS data, and SIMCA-P software (Umetrics) was used for multivariate statistical analysis. Charts were created using Microsoft Excel in-built functions.

The lipidomics analysis constitutes the following steps: (1) raw data processing, (2) lipid identification using MS/MS data, and (3) statistical analysis. The statistical analysis was performed on the following experimental design: group 1: KO1+KO1 vs group 2: WT1+WT2. Multivariate techniques namely, Principal Component Analysis(PCA), Partial Least Square-Discriminant Analysis (PLS-DA) , and Orthogonal Projections to Latent Structures Discriminant Analysis (OPLS-DA) were also conducted to identify the lipid species that significantly contributed in classifying the groups. The raw data were processed and then subjected to MS/MS database searches with a 5 ppm error tolerance for both the precursor, and product ions.

Stringent filter criteria were applied in SimLipid software to remove false positives in the results. Figure 1 (A) shows typical graphical user interface (GUI) of SimLipid software with optimized parameter settings for peak detection from LC-MS data generated by LTQ Orbitrap Pro Velos system. The MS/MS database search parameters shown in Figure 1(B) were used for lipid identification by SimLipid software.

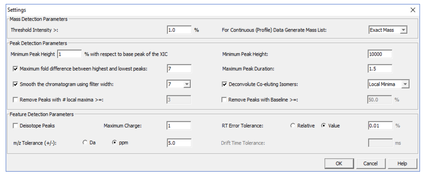

Figure 1 (A): Typical GUI of SimLipid software showing the LC-MS data processing parameters for LTQ Orbitrap Velos Pro Hybrid MS.

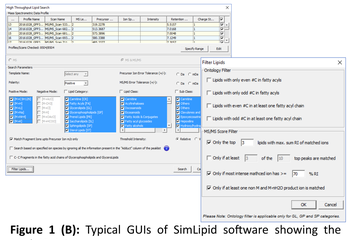

Figure 1 (B): Typical GUIs of SimLipid software showing the MS/MS database search parameters for structural identification of lipids using product ions data.

RESULTS AND DISCUSSION

The following numbers of unique lipids were identified by SimLipid MS/MS database searches between the samples.

Table 1 shows the parent molecular ions of the lipid species belonging to the top 11 lipid classes from which maximum number of lipids were identified by SimLipid from the +ve, and -ve mode data.

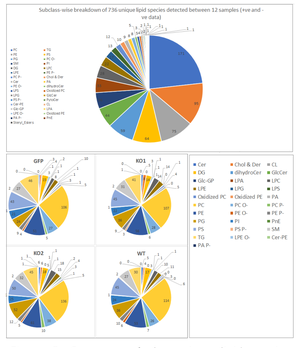

Figure 2: Summary of the unique lipid species identified from both the positive and negative modes data by SimLipid software tool using MS/MS data.

Main Class	-ve Data	+ve Data
PC	$[M+HCOO]^-$, $[M-H]^-$, $[M+CH_3]^-$, $[M+Cl]^-$	$[M+H]^+$
PE	$[M-H]^-$	$[M+H]^+$
PS	$[M-H]^-$, $[M+CH_3]^-$	$[M+H]^+$, $[M+Na]^+$
SM	$[M+HCOO]^-$	$[M+H]^+$, $[M+Na]^+$
TAG	-	$[M+NH_4]^+$, $[M+H]^+$
PG	$[M-H]^-$, $[M+OAcO]^-$	$[M+H]^+$
Cer	$[M-H]^-$	$[M+H]^+$
DAG	-	$[M+H]^+$, $[M+NH_4]^+$
PI	$[M-H]^-$	$[M+H]^+$
PA	$[M-H]^-$	$[M+H]^+$
CL	$[M-H]^-$	$M+H]^+$

Table 1: Parent Ion Molecules of the lipids identified.

2

Overview of Subclass-Specific MS/MS Characteristic Ions

Example 1: **PC(16:0_18:2) -** PC(16:0_18:2) at around 19 minutes was identified by the software tool with the following characteristic ions (Table 2). SimLipid identification is based on the nine characteristic ions (Figure 2). Description of the characteristic ions of the lipid is described in Table 3:

	Parent Ion Species: [M-HCOO]
Data ID	808
Retention Time:	18.9134
Precursor m/z	802.5599
Delta mass	0.5385 (in ppm)
Best Score/Grade	0.2202 (very good)
Group (replicates)	WT1 (replicate 2)
Characteristic Ions (SimLipid fragment name (observed m/z))	1. 15:0 COO-(255.2327), 2. 17:2 COO-(279.2326), 3. M-18:2-H2O(462.2989), 4. M-18:2(480.3089), 5. M-16:0-H2O(486.2975), 6. M-16:0(504.3083), 7. M-C5H13N(671.4639), 8. M-CH3(742.5374), 9. M(802.558)

Table 2: Table showing the characteristic ions based on which SimLipid identify the lipid species using MS/MS data

Fragment Type	Description	SimLipid Nomenclature
A	Sn1 chain	15:0 COO -
B	Sn2 chain	17:0 COO -
A*	M- Sn1 chain + Adduct	M-16:0 + HCOO-
A**	A* - H2O	M-16:0 -H2O + HCOO-
B	M- Sn2 chain + Adduct	M-18:2+HCOO-
B*	B* - H2O	M-18:2-H2O+HCOO-
C	M- Choline	M-C5H13N
D	M-CH3	M-CH3
E	M+Adduct	M

Table 3: Description of the characteristic ions of the lipid PC(16:0/18:2)

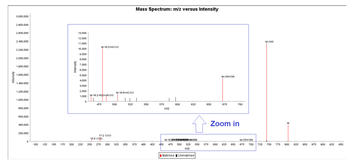

Figure 2: Typical GUI of SimLipid software showing the MS/MS spectrum annotated with fragment ions of the identified lipid species PC(16:0_18:2). The zoom in spectrum has been manually inserted for clear visualization of the fragment ions.

Example 2: **PE(16:0_18:2) -** For this lipid, the characteristic ions are the peaks at m/z values 313.274, 337.274, 575.5042 (Table 4). In addition to these common characteristic ions, SimLipid software make use of the characteristic ions 124.0159 (observed in the reported standard spectrum: (http://www.hmdb.ca/spectra/ms_ms/29673), and the characteristic ions corresponding to the fatty acids' acylium ions 239.2372, and 263.2372 (Figure 3).

	$[M+H]^+$
Data ID	364
Retention Time:	19.26
Precursor m/z	716.5235
Delta mass	2.1007 (in ppm)
Best Score/Grade	0.1586 (good)
Group (replicates)	WT2 (replicate 2)
Characteristic Ions (SimLipid fragment name (observed m/z))	i. HG-H2O (124.0159), ii. 15:0C=O+(239.2372), iii. 17:2C=O+(263.2372), iv. 16:0 PE (313.274), v. 18:2 PE (337.274), vi. M- NL+H (575.5042)

Table 4: Description of the characteristic ions of the lipid PE(16:0_18:2)

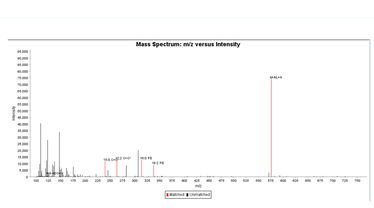

Figure 3: Typical GUI of SimLipid software showing the MS/MS spectrum annotated with fragment ions of the identified lipid species PE(16:0_18:2). The MS/MS spectra from positive mode data feature a noisy region (m/z values < 200).

Example 3: **CL(18:2_18:2_18:2_18:2) -** This lipid species was detected at around 24 minutes in the sample: WT2 Technical Replicate: 3

	SimLipid
Data ID	2544
Retention Time:	23.55
Precursor m/z	1447.9648
Lipid (Short Name)	CL (72:4)
Lipid (Common Name)	CL(18:2_18:2_18:2_18:2)
Delta mass	0.0745 (in ppm)
Score/Grade	0.1821 (good)

Table 5: Description of the characteristic ions of the lipid CL(18:2_18:2_18:2_18:2)

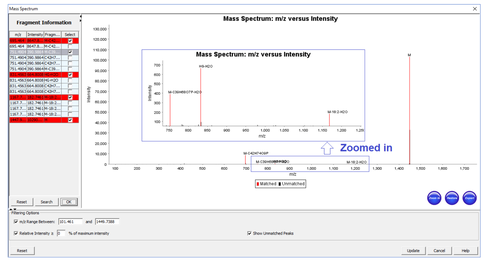

Figure 4: Typical GUI of SL software showing the MS/MS spectrum annotated with fragment ions of the identified lipid species CL(18:2_18:2_18:2_18:2). The zoom in spectrum has been manually inserted for clear visualization of the fragment ions.

Example 4: **PI(18:0_22:6) -** The MS/MS spectra of PI lipid species in –ve mode feature m/z peaks corresponding to both the fatty acids as well as neutral loss of the fatty acids. Figure 5 shows the annotated MS/MS spectrum of the lipid species PI(18:0/22:6) by SimLipid software.

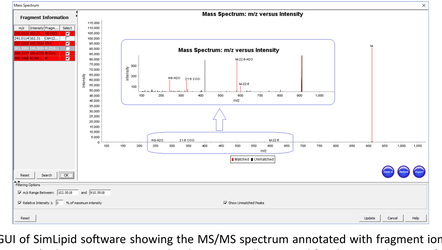

Figure 5: Typical GUI of SimLipid software showing the MS/MS spectrum annotated with fragment ions of the identified lipid species PI(18:0_22:6). The zoom in spectrum has been manually inserted for clear visualization of the fragment ions.

3

Downstream Quantitative and Statistical Analysis

Unique lipids identified from all the biological groups, and their technical replicates were exported into MS Excel file along with their peak areas. One of the challenges we face in quantitative measurement of detected lipid species from LC-MS data is the convoluted LC-MS peaks of isomeric species. For example, Figure 6 displays the convoluted peak belonging to isomeric species of TG 54:3 at m/z 902.8177. Hence, we create isomeric groups, each isomeric group representing all the individual isomeric lipid species based on number of carbons and double bonds in the fatty acid chains e.g., PC(32:0), TG(54:3), etc. (Figure 7).

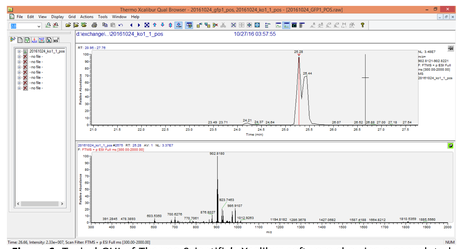

Figure 6: Typical GUI of Thermo Scientific's Xcalibur software showing a convoluted peak belonging to the extracted ion chromatogram of m/z 902.8177 with 0.005 mass tolerance.

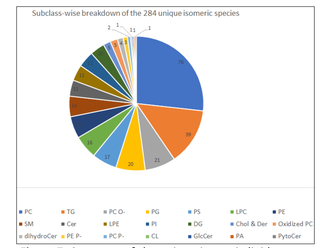

Figure 7: Summary of the unique isomeric lipid groups identified from both the positive and negative modes.

Data Normalization

This is an important step to ensure that we get correct, and reliable results from the statistical analyses that we may subject our data into for further hypothesis testing. For a biological group, the exported peak area of every ID lipid was normalized with respect to the sum of peak areas of all the identified ID lipids from all the technical replicates of the biological group. (Figure 8) This data normalization is done to fulfil the assumption that if a lipid is identified, its abundance is observed consistently throughout the biological replicates.

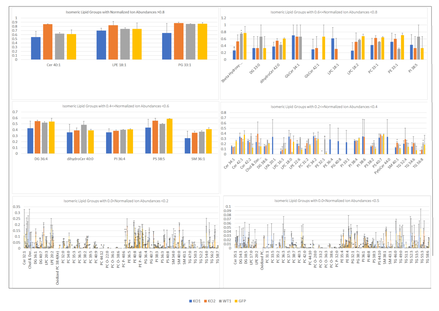

Figure 8: Bar charts showing isomeric lipid groups with their normalized ion abundances; ** denotes the isomeric group with significantly varying normalized ion abundances between the groups (p-value<0.01); * denotes the isomeric group with significantly varying normalized ion abundances between the groups (p-value<0.05).

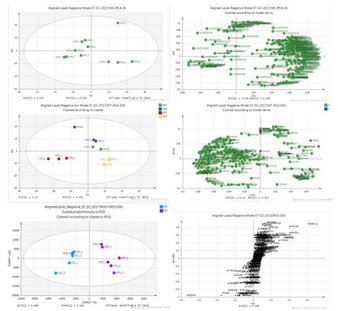

Figure 9: Typical GUIs of the SIMCA-P software showing Scores plots, and their corresponding loadings plots from PCA, PLS-DA, and OPLS-DA. observed in the score plot (only data from negative ion mode were used).

The normalized data was subjected into the following multivariate statistical techniques namely, PCA, PLS-DA, and OPLS-DA. From every technique, we display the (a) Score plot: showing the classification of the biological groups while their corresponding (b) Correlation loadings plot: showing the lipid species responsible for classification of the biological groups.

CONCLUSIONS

1. We have developed a lipidomics analysis method –LC coupled with high resolution Thermo's Orbitrap Velos Pro MS instrument methods; lipid identification using database search of MS/MS data from both the positive, and negative modes.
2. SimLipid software identified 736 unique lipid species, and 284 isomeric lipid groups across the 12 samples
3. The MS/MS spectra in positive mode feature a noisy region (peaks with m/z <200) which lead to false identification of fatty ester lipids by SimLipid software instead of TAG lipids. This could be controlled by removing fatty esters class in the SimLipid software database search filters.

REFERENCES

1. Hu, C.,et al. (2008) *Proteome Res. 7: 4982–4991*.