

LipidMS application workflow

Maribel Alcoriza Balaguer

2021-12-13

LipidMS Overview

LipidMS v3 is an R-package aimed to confidently identify lipid species in untargeted LC-MS for DIA or DDA data analysis. It combines a set of fragmentation and intensity rules and a parent-to-fragment co-elution score (PFCS, only applied for DIA analysis) calculated in predefined retention time windows. Depending on the MS evidence reached by the annotations, lipids can be identified at three different levels: i) subclasslevel, e.g., PG(34:1); ii) fatty acyl level, e.g., PG(16:0_18:1); and iii) fatty acyl position level, e.g., PG(16:0/18:1). As a general rule, parent ions will be found when no collision energy is applied, while fragment ions will be found when it is. Each lipid class has characteristic ionization and fragmentation properties that allow to filter informative fragments among all fragment ions to reconstruct the parent's structure. Next figure summarizes the basis of LipidMS annotation:

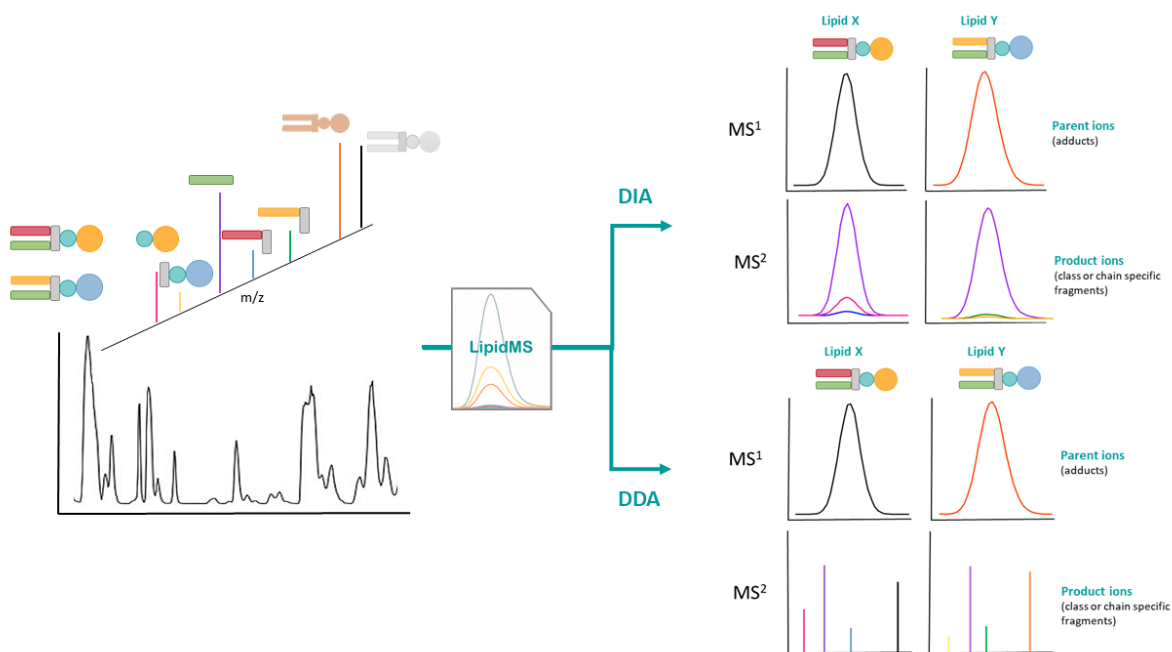


Figure 1: LipidMS abstract

And the following figure shows the overall workflow of LipidMS:

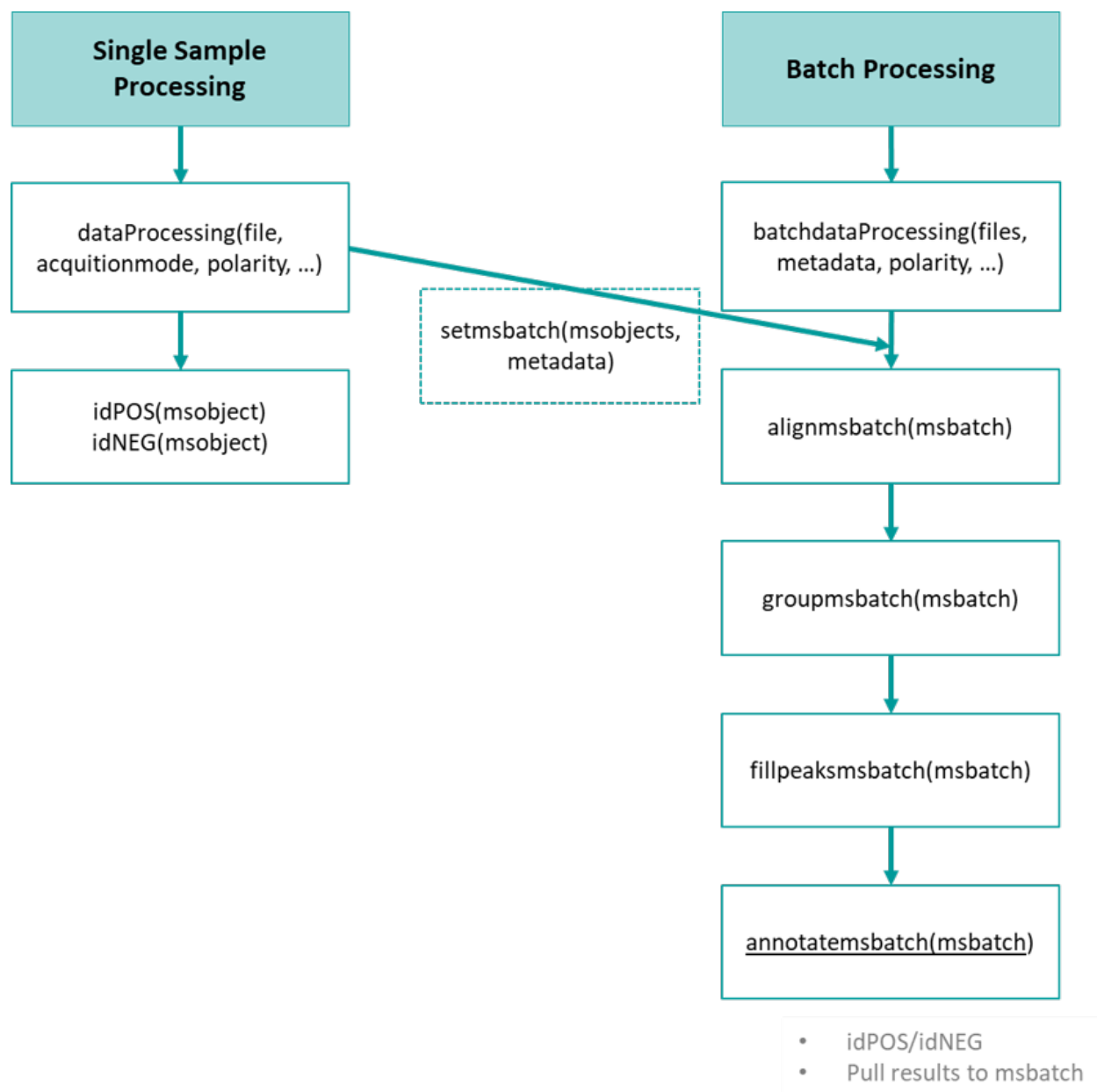


Figure 2: LipidMS workflow

Installation

The LipidMS application can be accessed locally from R or through our website at <https://www.lipidms.com>. In case you want to run LipidMS locally you need to install the package, otherwise skip this step.

```
# Install LipidMS  
install.packages("LipidMS", dependencies = c("Depends", "Imports"))
```

Files conversion

To start the MS analysis raw files need to be converted into mzXML format (you can use any software such as MSConvert from proteowizard) and then, LipidMS can be run.

Example data files

Some example files and scripts can be downloaded here.

LipidMS application

How to run it locally

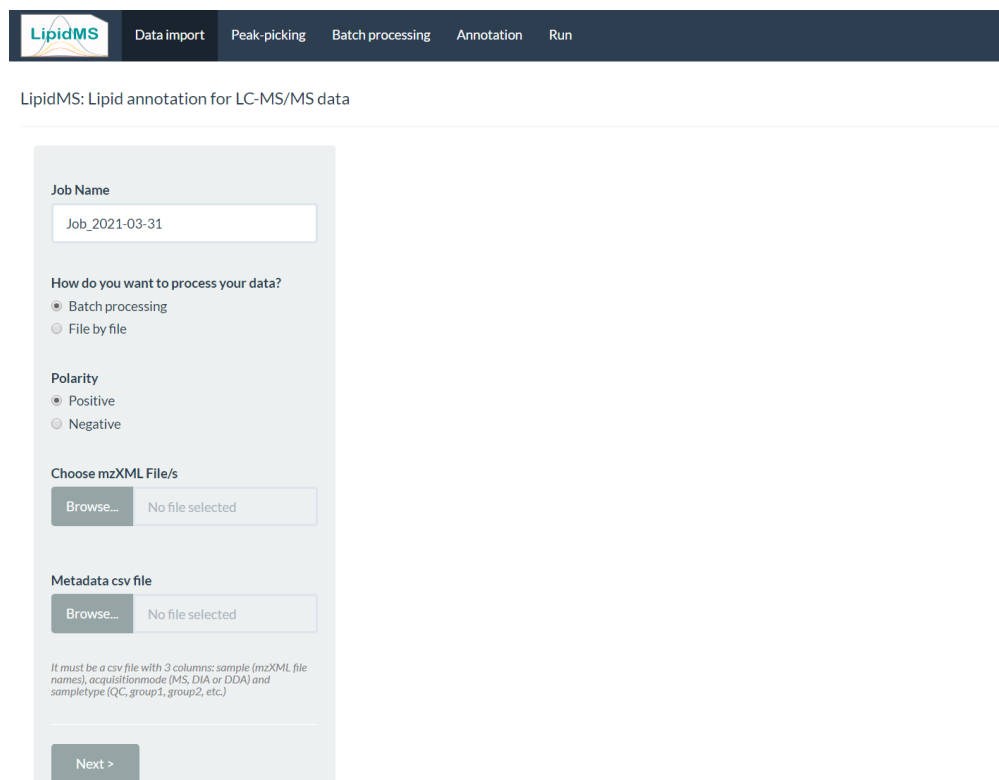
In case you are working locally on your computer execute the following lines and the app will open on a new window:

```
# load LipidMS library  
library(LipidMS)  
  
# Example data files can be downloaded from:  
# https://drive.google.com/drive/folders/1hSYrQBkh-rAA-oiaKqGkrNL7uWQraV75?usp=sharing  
  
# Lunch the app  
LipidMSapp()
```

In case you would prefer to use the web application, it can be accessed from <https://www.lipidms.com>.

Data Import

Once all your files have been converted to mzXML and the app is running, upload your files and tune the parameters. At the first tab, a metadata csv file is required. It must have 3 columns: sample (mzXML file names), acquisitionmode (MS, DIA or DDA) and sampletype (QC, group1, group2, etc.)



The screenshot shows the 'Data import' tab of the LipidMS application. The interface includes a top navigation bar with the LipidMS logo and tabs for 'Data import', 'Peak-picking', 'Batch processing', 'Annotation', and 'Run'. Below the navigation bar, the title 'LipidMS: Lipid annotation for LC-MS/MS data' is displayed. The main form area contains the following sections:


- Job Name:** A text input field containing 'Job_2021-03-31'.
- How do you want to process your data?:** Two radio button options: 'Batch processing' (selected) and 'File by file'.
- Polarity:** Two radio button options: 'Positive' (selected) and 'Negative'.
- Choose mzXML File/s:** A 'Browse...' button and a text field showing 'No file selected'.
- Metadata csv file:** A 'Browse...' button and a text field showing 'No file selected'.
- Instructions:** A note stating: 'It must be a csv file with 3 columns: sample (mzXML file names), acquisitionmode (MS, DIA or DDA) and sampletype (QC, group1, group2, etc.)'.
- Next >** A button at the bottom of the form.

Figure 3: Data import tab of LipidMS app

If you have samples acquired with different polarities they must be analyzed in different batches.

Peak-picking

At the second tab we find all parameters required for peak-picking. More detailed information about all these parameters can be found at the documentation page of the dataProcessing function.



Data import

Peak-picking

Batch processing

Annotation

Run


LipidMS: Lipid annotation for LC-MS/MS data

dmzagglom (in ppm) <small>m/z tolerance used for partitioning and clustering. 5 by default.</small>	MS1 <input type="text" value="5"/>	MS2 <input type="text" value="5"/>
drtagglom (in seconds) <small>rt window used for partitioning (in seconds). 25 by default.</small>	MS1 <input type="text" value="25"/>	MS2 <input type="text" value="25"/>
drtclost (in seconds) <small>rt window used for clustering (in seconds). 25 by default.</small>	MS1 <input type="text" value="25"/>	MS2 <input type="text" value="25"/>
minpeak <small>minimum number of measurements required for a peak. By default, 5 for MS1 and 4 for MS2.</small>	MS1 <input type="text" value="5"/>	MS2 <input type="text" value="4"/>
minint <small>minimum intensity of a peak. By default, 1000 for MS1 and 100 for MS2.</small>	MS1 <input type="text" value="1000"/>	MS2 <input type="text" value="100"/>
grtgap (in seconds) <small>maximum rt gap length to be filled. 5 by default.</small>	MS1 <input type="text" value="5"/>	MS2 <input type="text" value="5"/>
drtmminpeak (in seconds) <small>minimum rt width of a peak. 15 by default. At least minpeak within the drtmminpeak window are required to define a peak.</small>	MS1 <input type="text" value="15"/>	MS2 <input type="text" value="15"/>
drtmmaxpeak (in seconds) <small>maximum rt width of a single peak. 100 by default.</small>	MS1 <input type="text" value="100"/>	MS2 <input type="text" value="100"/>
maxeicpeaks <small>maximum number of peaks within one EIC. By default, 5 for MS1 and 3 for MS2.</small>	MS1 <input type="text" value="5"/>	MS2 <input type="text" value="10"/>
weight <small>weight for assigning measurements to a peak. By default, 2 for MS1 and 3 for MS2.</small>	MS1 <input type="text" value="2"/>	MS2 <input type="text" value="3"/>

Figure 4: Peak-picking tab of LipidMS app

Batch processing

The third tab contains parameters required for alignment, grouping and filling missing peaks.

 Data import Peak-picking **Batch processing** Annotation Run


LipidMS: Lipid annotation for LC-MS/MS data

dmzalign <i>mass tolerance between peak groups for alignment (in ppm). 5 by default.</i>	<input type="text" value="5"/>
drtalign <i>maximum rt distance between peaks for alignment (in seconds). 30 by default.</i>	<input type="text" value="30"/>
span <i>span parameter for loess rt smoothing. 0.4 by default.</i>	<input type="text" value="0,4"/>
minsamplesfracalign <i>minimum samples fraction represented in each cluster used for alignment. 0.75 by default.</i>	<input type="text" value="0,75"/>
dmzgroup <i>mass tolerance between peak groups for grouping (in ppm). 5 by default.</i>	<input type="text" value="5"/>
drtagglomgroup <i>maximum rt distance in mz partitions for grouping (in seconds). 30 by default. It shouldn't be smaller than drtgroup.</i>	<input type="text" value="30"/>
drtgroup <i>maximum rt distance between peaks for grouping (in seconds). 30 by default.</i>	<input type="text" value="30"/>
minsamplesfracgroup <i>minimum samples fraction represented in each cluster used for grouping. 0.25 by default.</i>	<input type="text" value="0,25"/>
parallel <i>parallelize processing. FALSE by default.</i>	<input type="text" value="FALSE"/>
ncores <i>number of cores to be used in case parallel is TRUE. 2 by default.</i>	<input type="text" value="2"/>

Figure 5: Batch processing tab of LipidMS app

Annotation

This tab contains parameters related with the annotation step and which lipid classes will be included in the analysis.

Data importPeak-pickingBatch processingAnnotationRun

LipidMS: Lipid annotation for LC-MS/MS data

dmzprecursor
mass tolerance for precursor ions. 5 by default.

5

dmzproducts
mass tolerance for product ions. 10 by default.

10

rttol
total rt window for coelution between precursor and product ions. 5 by default.

5

coelcutoff
coelution score threshold between parent and fragment ions. Only applied if rawData Info is supplied. 0.7 by default.

0.7

Lipid classes to annotate for ESI+ :

- ☒ MG
- ☒ LPC
- ☒ LPE
- ☒ PC
- ☒ PE
- ☒ PG
- ☒ Sphingoid bases
- ☒ Sphingoid bases phosphate
- ☒ Cer
- ☒ SM
- ☒ Carnitines
- ☒ CE
- ☒ DG
- ☒ TG

Lipid classes to annotate for ESI+ :

- ☒ FA
- ☒ FAHFA
- ☒ LPC
- ☒ LPE
- ☒ LPG
- ☒ LPI
- ☒ LPS
- ☒ PC
- ☒ PE
- ☒ PG
- ☒ PI
- ☒ PS
- ☒ Sphingoid bases
- ☒ Sphingoid bases phosphate

Figure 6: Annotation tab of LipidMS app

Run

Finally, run your job. You will obtain two or three types of csv files with the results tables (feature matrix if batch processing has been performed, summary tables and the whole peak tables with annotations) and pdf files with plots of the peaks supporting lipid identifications for each one of your files.

If you are using the web application, an email will be required to send back your results. Otherwise, if you are using LipidMSapp() on your computer, you will find four buttons to download your results. Wait until you can see the results on the main panel to download (you may need to write the extension .zip to save your files properly).

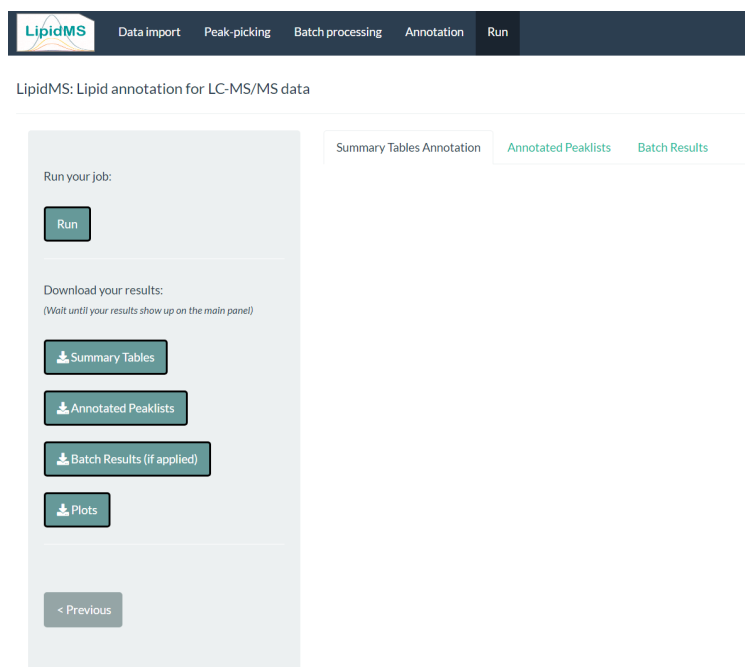


Figure 7: Run tab of LipidMS application

If you have any further questions, please do not hesitate to contact us at: maribel_alcoriza@iislife.es or maribel_alcoriza@hotmail.com