May 30, 2024

BAF_Protocol_011 Metabolomics: Database Search MS-DIAL and Analysis using Metaboanalyst 6.0

DOI
dx.doi.org/10.17504/protocols.io.dm6gpzj1jlzp/v1

Nicholas Sherman[1]

[1]University of Virginia Biomolecular Analysis Facility Core

Biomolecular Analysis Fac…

Dilza Silva
UVA

DOI: dx.doi.org/10.17504/protocols.io.dm6gpzj1jlzp/v1

Protocol Citation: Nicholas Sherman 2024. BAF_Protocol_011 Metabolomics: Database Search MS-DIAL and Analysis using Metaboanalyst 6.0. **protocols.io https://dx.doi.org/10.17504/protocols.io.dm6gpzj1jlzp/v1**

Protocol status: Working
We use this protocol and it's working

Created: May 17, 2024

Last Modified: May 30, 2024

Protocol Integer ID: 100061

Keywords: Metabolomics, MS-DIAL, Metaboanalyst, data analysis

Abstract

These steps represent a starting point for analysis of metabolomics data. Other data analysis and/or packages may be used.

Materials

Raw files.

MS-DIAL Software: https://systemsomicslab.github.io/compms/msdial/main.html
Databases: https://systemsomicslab.github.io/compms/msdial/main.html#MSP

MetaboAnalyst 6.0 online tool: https://www.metaboanalyst.ca/MetaboAnalyst/home.xhtml

MS-DIAL V 4.9.221218 starting

1 Thermo RAW files are organized in two folders - one for all positive mode raw files and another for all negative mode raw files. These files will be loaded into MS-DIAL in two separate searches. Download MSP library files for Positive and Negative mode and save them in a folder in the computer you are using to perform the searches.

2 Open MS-DIAL

3 Click File -> New project

4 Project file path: Browse the folder with Positive or Negative raw files
Ionization type: Soft ionization (LC-MSMS)
Separation type: Chromatography (LC)
MS method type: Conventional LC/MS or data dependent MS/MS
Data type: MS1 Profile, MS/MS Centroid
Ion mode: Positive or Negative
Target omics: Metabolomics
Click "Next"

5 Analysis file path: browse to the folder with raw files, select all the apply for the search. Should include at least 3 blanks and 3 QC runs.
Choose the file type for each sample (Blank, QC, sample)
Class ID - add short group names for each replicate without adding numbers among replicates.

Analysis parameter setting - keep default if not mentioned in this section

6 **Data collection tab**
Mass accuracy: MS1 tolerance 0.01 Da, MS2 tolerance 0.025 Da
- Advanced:
Data collection parameters: retention time begin 0 min, retention time end 15 min, MS1 range begin 50 Da, MS1 range end 1000 Da, MSMS range begin 50 Da, MSMS range end 1000 Da, Isotope recognition: max charge number 2.

7 **Peak Detection tab**
Peak detection parameters: minimum peak height 10000 amplitude, mass slice 0.1 Da
- Advanced:
Smoothing method: linear weighted moving average
Smoothing level: 3 scan
Minimum peak width: 5 scan

8 **MS2Dec tab**

Deconvolution parameters: sigma window value 0.5, MS2 abundance cut off 10 amplitude.
- Advanced:
Check Exclude after precursor ion
Keep the isotope ions until 0.5 Da
Check Keep the isotopic ions w/o MS2Dec

9 **Identification tab**

MSP file and MS/MS identification setting: add MSP file downloaded Positive or Negative depending on the data that you loaded in the step 04.
Retention time tolerance: 100 min, Accurate mass tolerance (MS1) 0.01 Da, Accurate mass tolerance MS/MS 0.05, Identification score cut off 80%.
- Advanced:
add txt file of an in-house library Positive or Negative depending on the data that you loaded in the step 04.
Retention time tolerance: 0.1 min, Accurate mass tolerance 0.01 Da, Identification score cut off 85%.
Check only report the top hit

10 **Adduct tab**

Check [M+H]+
Check [M+H-H2O]+
Check [2M+H]+

11 **Alignment tab**

Alignment parameters settings: Under result name add a name for the alignment file, reference file choose one of the QC runs, retention time tolerance 0.05 min, MS1 tolerance 0.015 Da
- Advanced:
Check remove feature based on blank information
Check keep "reference matched" metabolite features
Check keep "suggested (w/o MS2)" metabolite features
Check keep removable features and assign the tag
Check Gap filling by compulsion

12 **At the bottom**

Check together with Alignment
Click "Finish" --> processing will start.
Once the search is finished, click on the name of the file that is showing at the Alignment Navigator
Click Export - Alignment result, choose a folder to save the file, check remove features from blank, and export.
Data can be loaded into Excel/R or any other software environment for further data analysis.

13 Data Analysis: load data into Excel, filter out: m/z match FALSE and identifcations named w/o MS2. Combine postive and negative identifications removing redundancy and save as txt file. Load the combined non-redundant data into R or the online MetaboAnalyst 6.0 tool.

14 MetaboAnalyst 6.0 --> statistical analysis
Upload .txt data as peak intensities (check file format), no data filtering, Sample normalization: normalization by median, Data transformation: Log transformation (base 10), no Data Scaling.
At this point there are distinct analyses that can be done, the following is provided by our core as a first-step analysis:
Chemometrics analysis: PCA, PLS-DA with importance measures using VIP.
Univariate analysis: fold change and t-test -> volcano plot (two groups comparisons).
Clustering analysis: heatmap.

Protocol references

Tsugawa H, Cajka T, Kind T, Ma Y, Higgins B, Ikeda K, Kanazawa M, VanderGheynst J, Fiehn O, Arita M. MS-DIAL: data-independent MS/MS deconvolution for comprehensive metabolome analysis. Nat Methods. 2015 Jun;12(6):523-6. doi: 10.1038/nmeth.3393. Epub 2015 May 4. PMID: 25938372; PMCID: PMC4449330.

Pang, Z., Lu, Y., Zhou, G., Hui, F., Xu, L., Viau, C., Spigelman, A., MacDonald, P., Wishart, D., Li, S., and Xia, J. (2024) **MetaboAnalyst 6.0: towards a unified platform for metabolomics data processing, analysis and interpretation** Nucleic Acids Research (doi: 10.1093/nar/gkae253)