DOM_Interlab-LCMS study 2021 - Sample preparation and analysis guide

The Shipment

Arriving in the package is 7 vials: QC Mix, A, A45M, A15M, A5M, M, PPL blank. The QC vial is not part of the study, it is simply for you to check/adjust/optimise your method.

If the vials are to be stored for >1 week before analysis, please store them at 4 degrees C, although the contents are expected to be stable at room temperature

The PPL blank should be resuspended in 100 uL 50% MeOH. All other vials should be resupendent with 200µL 50% MeOH (LCMS grade water and MeOH) vortexed/sonicated, transferred to an insert vial and injected at 5 - 10 µL (depending on your instrument sensitivity).
Note. In case some of the material does not completely dissolve after sonication, we recommend to briefly centrifuge the sample and only transfer the supernatant to the micro insert.

Please run all 5 complex samples (A, A45M, A15M, A5M, M) as triplicates, in POS MS1 and POS MS/MS DDA mode as well as in NEG MS1 and NEG MS/MS DDA mode. Ideally, the sample order should be randomized within each mode (MS1 pos, MS/MS pos, MS1 neg, MS/MS neg)
Note: QC Mix and PPL Blank should be run one time in each of the four modes.

Table 1: The QC-Mix contains the compounds stated in the table below.

Compound	MF	[M-H]-	[M+H]+	tr (min)	Log P	LogP Data
Sulfathiazole	$C_9H_9N_3O_2S_2$	254.0063	256.0209	5.76	0.05	Experimental
Guanosine-5'-monophosphate	$C_{10}H_{14}N_5O_8P$	362.0507	364.0653	2.13	-3.5	Computed
Vanillin	$C_8H_8O_3$	151.0401	153.0546	6.49	1.21	Experimental
Hippuric Acid	$C_9H_9NO_3$	178.051	180.0655	5.93	0.31	Experimental
Capsicin	$C_{18}H_{27}NO_3$	304.1918	306.2064	9.32	4.42	Experimental
Dihydrocapsicin	$C_{18}H_{29}NO_3$	306.2075	308.222	9.8	4.4	Computed
glycyrrhizic acid	$C_{42}H_{62}O_{16}$	821.3965	823.4111	7.72	2.8	Experimental
Caffeine	$C_8H_{10}N_4O_2$	193.0731	195.0877	5.68	-0.07	Experimental
Methyl-D-Mandelate	$C_9H_{10}O_3$	165.0557	167.0703	8.16	1.2	Computed
Raffinose	$C_{18}H_{32}O_{16}$	503.1618	505.1763	1.52	-5.8	Computed

Note: Some of these compounds clearly ionise better in one polarity or the other, but they should all give good peaks with a 5µL injection in at least one polarity (pos/neg). If you re-dissolve the vials to 200µL, they should be 10µg/ml.

Table 2: Exact masses of internal standards in DOM samples

Compound	**Molecular Formula**	**Mass [M+H+]+**	**Mass [M+Na+]+**	**Mass [M-H+]-**
Domoic acid	$C_{15}H_{21}NO_6$	312.1441639	334.1261085	310.1296109
Kanic acid	$C_{10}H_{15}NO_4$	214.1073845	236.089329	212.0928315
Isoxaben	$C_{18}H_{24}N_2O_4$	333.1808837	355.162828	331.1663308
Irgarol	$C_{11}H_{19}N_5S$	254.1433929	276.125338	252.1288399
Imazapyr	$C_{13}H_{15}N_3O_3$	262.1186179	284.100563	260.1040649
Heroin	$C_{21}H_{23}NO_5$	370.1648993	392.146844	368.1503464
Methamphetamine	$C_{10}H_{15}N$	150.1277261	172.109671	148.113173
Cocaine	$C_{17}H_{21}NO_4$	304.1543347	326.13628	302.1397817

LC-MS Parameters

The goal of the study is to compare high-resolution LC-MS and LC-MS/MS data from algal extracts and dissolved organic matter from multiple laboratories and different instrument platforms. In order to enable a comparison of different instruments, the chromatographic and MS acquisition settings need to be as unified as possible. While we acknowledge that differences in maximum pressure limits of each LC system will also determine the type of column particle size (HPLC vs. UHPLC) that can be applied, we want to make sure that at least the mobile phase, flow rate, column chemistry and column dimensions are relatively similar.
For the study we recommend the following parameters:

Chromatography

- Stationary phase: C18 with column diameter of 2 mm and length around 150 mm.
- Particle size: As low as you have. Jeff and Daniel are both using 1.7µm particle size Phenomenex Kinetex C18 150x2.1
- Mobile phase: LC-MS grade water (A), acetonitrile (B) with 0.1% formic acid (FA) each.
- Gradient: Start at 5% B, increase to 50% B at 7 min, increase to 99% B at 10 min, 3 min washout phase at 99% B, 4 min equilibration phase at 5% B (method length = 17 min).
- Flow rate: Should be around 0.4 mL/min.
- If available, the column oven should be set to 30 C
- MS data should be acquired over the full LC run (17 min).

- Injection volume 5-10 uL

NOTE: The majority of labs to date (Jun 26th 2021) have injected 5µL, meaning a total of 60µL injected and 140µL remaining, which can be stored at -20degC for future analysis.

General MS Settings

- Mass range: 150-1500 m/z
- ESI settings: ESI pos and neg tuned for HPLC flow rate used (Use your regular infusion rate of calibration/tune mix + makeup flow from the HPLC through a T-mixer) with the vender specific cal mix (eg Thermo cal mix , tuned for m/z 514 in pos mode).

MS/MS Runs

- Isolation width should be as narrow as possible to obtain still good transmission (e.g. 1 m/z for Orbitraps)
- Normalized Collision Energy: If available, use stepped collision energy with 20 to 30 to 40 % with z = 1 as default charge state. If no stepped collision energy is available use 35 % with z = 1 as default charge state
- Data Dependent MS/MS Acquisition: DDA top5 (5 MS/MS events per duty cyle).
- Dynamic exclusion on, with repeat count of 1 and exclusion duration of 5 sec.
- Resolution for MS Survey Scan 70K - 140K (depending on instrument platform, and duty cycle time, generally resolution should be as high as possible while keeping duty cycles < 1sec)
- Resolution for MS/MS scans should be around 17K - 35K (depending on instrument platform, and duty cycle time, generally resolution should be reasonably high while keeping duty cycle < 1sec)
- Maximum Injection time (C-trap fill time, TOF accumulation time): 100 ms for MS1, 100-200 ms for MS/MS
- Automatic Gain Control (AGC): default settings (if available and depending on instrument used)
- MS/MS threshold: Depending on Instrument type we recommend to use an MS/MS threshold of 1-5% of the AGC.
- Apex-Trigger: If your systems support apex triggering (e.g. Q-Extactive or Orbi Fusion), turn this setting on for time spans of 2-15 sec.

MS1 Runs

- MS Resolution should be as high as possible without exceeding the duty cycle time by 1 sec (depending on instrument platform)
- Maximum Injection time (C-trap fill time, TOF accumulation time): 200 ms for MS1
- AGC: default settings (depending on instrument used)

Documentation of Parameters

In order to check and help you set consensus settings for the study we ask everybody to list the settings for their methods in this list before you run the samples.

https://docs.google.com/spreadsheets/d/1LvL8klby8bC2OpP7AYNVPck-lz281-a9MFe4laQTjqg/edit?usp=sharing
Please list your settings under the appropriate instrument section (e.g. Q-Exactive, LTQ-Orbi, qTOF). Once we checked the settings, we will mark it as “checked” and you can proceed running the samples.

Note: Please don't run the sample without our OK on your settings.

Office Hours
If you have any questions regarding your settings, we will offer office hours on Thursdays from 7:00 - 8:00 PST on zoom (https://zoom.us/my/functionalmetabolomicslab). If those days/times don’t work for you feel free to email us and we will set a different meeting with you.

Check your instrument performance
In order to optimize measurement settings apart from these fixed parameters, we recommend to check the detectability of 10 target peaks in the QE mix shipped along with the DOM/algal extract samples before you run the sample (see table above).

Data handling
Once all data has been acquired, data needs to be **peak picked (centroided)** and converted to mzML file format with MS convert (From Proteowizard). And both raw and .mzML files uploaded to the massive repository (massive.ucsd.edu).

We made a short video tutorial on how to convert, upload, and generate the MassIVE data set here: https://www.youtube.com/watch?v=sudY7UtkMQg

Files from the different modes (POS_MSMS_mzML, POS_MS1_mzML, NEG_MSMS_mzML, NEG_MS1_mzML) and (POS_MSMS_raw, POS_MS1_raw, NEG_MSMS_raw, NEG_MS1_raw) should be organized in separate folders. The name of the data files should follow the following convention:
DOM_Interlab-LCMS_LabXXX_A45M_Pos_MS2_rep1 (replace XXX with the number of your lab, if you run samples on two machines, you will get two lab numbers, lab numbers are show here:
https://docs.google.com/spreadsheets/d/1LvL8klby8bC2OpP7AYNVPck-lz281-a9MFe4laQTjqg/edit#gid=1563737078

During the dataset creation it will be helpful to organize the data in raw (.raw) and peak (.mzML) files and eventually also separate them in subfolders for MS1 and MS2 and pos and neg data. (eg. raw_MS1_pos, raw_MS1_neg, raw_MS2_pos, raw_MS2neg, mzml_MS1_pos, mzml_MS1_neg, mzml_MS2_pos, mzml_MS2_neg)

Before you create the massive dataset, I'd recommend using an FTP client (WinSCP) to upload the data to your GNPS/massive user space.

https://ccms-ucsd.github.io/GNPSDocumentation/fileupload/

With the FT client, you can directly organize the files in the desired folder structure. (As suggested above)

Once the data is uploaded and you logged in on the massive page (massive.ucsd.edu) you can create the dataset via this link:

https://massive.ucsd.edu/ProteoSAFe/index.jsp?params={%22workflow%22:%22MASSIVE-COMPLETE%22}

Detailed documentation can be found here:

https://ccms-ucsd.github.io/GNPSDocumentation/datasets/#submitting-gnps-massive-datasets