**Sections:**



# Metabolomics Sample Submission and Guidelines

The first step to submitting samples is a discussion with the JGI Metabolomics leads and program manager regarding experimental design and sample preparation for metabolomics analysis. All samples are different, and it is important to make sure that the correct controls are prepared, enough sample provided for analysis, experimental design and samples are compatible with MS analysis, extraction method / solvents are suitable for detecting metabolites of interest and compatible with the sample / sample container. Measurements for normalization, if necessary, are also discussed (biomass, dry weight, OD, TOC, etc).

## Pilot Project

Prior to submitting the final set of experimental samples, a small subset of samples for a **pilot project** may be recommended to de-risk the experiment. This is typically about 3-20 samples in size, depending. According to the methods and procedures discussed, pilot samples will be extracted, run on LC-MS and undergo a preliminary analysis.

Here we will test, evaluate and finalize:

Amount of sample and sample preparation
- Is there sufficient signal for analysis?
- Are the metabolites of interest detected?
- Is there anything contaminating the MS signal?

Sample container
- Is this compatible with extraction solvents, etc?

Extraction method

Additional measurements (e.g. dry weight, OD)
Number of replicates
- How variable are replicates?

Controls
- Extraction blanks (empty tubes), conditioned (uninoculated) media

LC-MS method (w/ chromatography)
- HILIC, C18, etc

Once the pilot experiment has been completed and everything “goes well,” the final samples can be submitted. .

## Sample Submission Form

Before shipping samples, a “Sample Submission Form” must be filled out and approved. This provides the relevant metadata necessary to catalog and appropriately analyze the submitted samples. Review of this sheet also ensures all the appropriate controls and samples are being included.

An example form can be viewed here (Dec 2024 - V9)

The form has 4 sheets/tabs:

Sheet 1 – DATA SHEETS
Sheet 2 – EXPERIMENT DESCRIPTION
Sheet 3 – MEDIA RECIPES
Sheet 4 - TARGETED COMPOUNDS

**DATA SHEETS.** This is a list of all samples submitted and associated sample information. The requested metadata is important for identifying each sample and the components of each sample. This information is needed for generating high quality LC-MS data for subsequent analysis.

**EXPERIMENT DESCRIPTION.** A brief description of the samples and related experiment, as well as any processing that has been performed on the samples and types of analysis requested.

**MEDIA RECIPES.** Submit detailed media recipes used in the experiment including a list of specific compounds and concentrations. When appropriate, include media preparation steps. For complex media or “environmental” media, include the source or catalog information.

**TARGETED COMPOUNDS.** This information is used to ensure that if specific metabolites of interest are important to an experiment, especially in a “targeted” analysis, these will be considered in the analysis. Metabolic pathways or classes of metabolites may also be listed

(e.g. amino acids) but please provide as much detail as possible and hyperlinks. In some cases, a small number of compound standards may be run inline with the samples themselves for quantitative analysis, more robust identification, or if the standard is available but not yet added to JGI's database.

## Considerations in Experimental Design

**Sample amount.** Usually >1 mg C for soil, media, plant exudates. Other typical amounts – soil (0.5-2 g), media (1 mL), wet biomass (~100 mg, 3 mm pellet – often pellet from 1 mL media), dry biomass (2-10 mg). *Note:* For soil, if high clay content (doesn't extract well) or low C is expected (not rich), then up to 5, 10 or 30 g may be needed. For cell pellets / plant biomass, 100 mg typically results in a very good signal. Lower amounts will have less signal (e.g. 20 mg) with small numbers of metabolites detected, with 5 mg typically being a lower limit. Pilot experiments are used to determine what amount of material for a particular sample will be needed.

**Number of replicates.** Usually 4 replicates (up to 8 if highly variable between sample replicates). *Note:* For each type of MS analysis, a separate replicate sample may need to be prepared (different extractions used for each). Exceptions to this include when the extraction solvent is 100% methanol (or similar) - in this case, we run the same extract on both C18/nonpolar and HILICZ/polar chromatographies with good results. Lipid analysis is typically performed on a chloroform-based extraction, but can also be performed on a 100% MeOH extract with some variation in results based on lipid solubility.

**Sample processing / things to avoid.** Some chemicals are not compatible with sample extraction (e.g. glycerol, DMSO), or have strong signals that suppress and interfere with detecting metabolites of interest. This includes surfactants, PEG, dirty glass with soap residue, some buffers. Other sources of contamination include polymers leaching from plastic, especially if they contact organic solvent (e.g. filters, acrylate, pipette tips with low adhesive polymer coating). For glassware, it is best to "bake" off carbon-based residues prior to use or acid-wash. Washing and re-using plastics is not advised. When using plastic containers, polypropylene is recommended for solvent compatibility and minimal leaching.

**Media.** Defined and minimal media are preferred since competitive ionization and signal suppression is minimized. Also, since metabolites in these can be fully characterized, this makes it easier to detect new peaks / metabolites and to track increases or decreases in metabolites versus control media. Complex media is OK but analysis may also be more complex.

**pH.** For certain chromatographies, pH of the sample and resulting extract will influence the retention time of a metabolite (e.g. our HILICZ chromatography) and interfere with data analysis. For JGI's HILICZ chromatography, near neutral pH is OK. If low (<5) or high (>8) pH is expected / measured, adjusting the pH of the sample or extract may be necessary

prior to running LC-MS. For standard C18 analysis (secondary metabolites, non-polar), pH is typically not an issue.

**Salt.** In general, high salt is not mass spectrometry friendly when using chromatography for primary metabolite analysis (HILICZ). For these samples (e.g. ocean), a large amount of salt is removed during extraction (sometimes multiple salt crashed are performed in sequential extractions) but sometimes the remaining salt may still suppress signal during portions of the LC-MS run. For C18 analysis (secondary metabolites, non-polar), salt is typically not an issue.

## Controls – Do not forget!!

**Conditioned media control.** Incubate media but do not inoculate. This media should be subject to the same processing conditions and environments as the inoculated samples (e.g. filtering, temperature, light, time, etc). This control is necessary to determine what is being made by the organism, or consumed, when compared to the control. Fresh media is not comparable because metabolite degradation or other changes may occur during incubation, especially in the case of high temperature samples.

**Extraction controls, ExCtrl (just the containers).** Please send the same containers that samples were shipped in but with no sample. This control is necessary to determine sample signal from "background" signal that arises from the container itself, extraction solvents, or somewhere else during LC-MS sample preparation or system itself. Same # of replicates as each sample group.

**Fresh media.** This control is used to evaluate how much the experimental procedures have changed the media, and to determine the initial metabolite composition.

**Positive / Negative.** When possible, please include a positive and negative control. For instance, if you expect to see synthesis of a metabolite of interest in one sample and not in another.

**Stable isotope probing (SIP), unlabeled controls (no isotopic label).** When performing an experiment with stable isotopic labeling, include replicate samples that do not have any labeling (no heavy isotope added). This control is required for initial detection / identification of a compound (detecting the M0 mass with zero heavy isotopes), and is also needed to determine the relative isotopic levels of a specific metabolite when unlabeled to calculate isotopic incorporation when labeled.

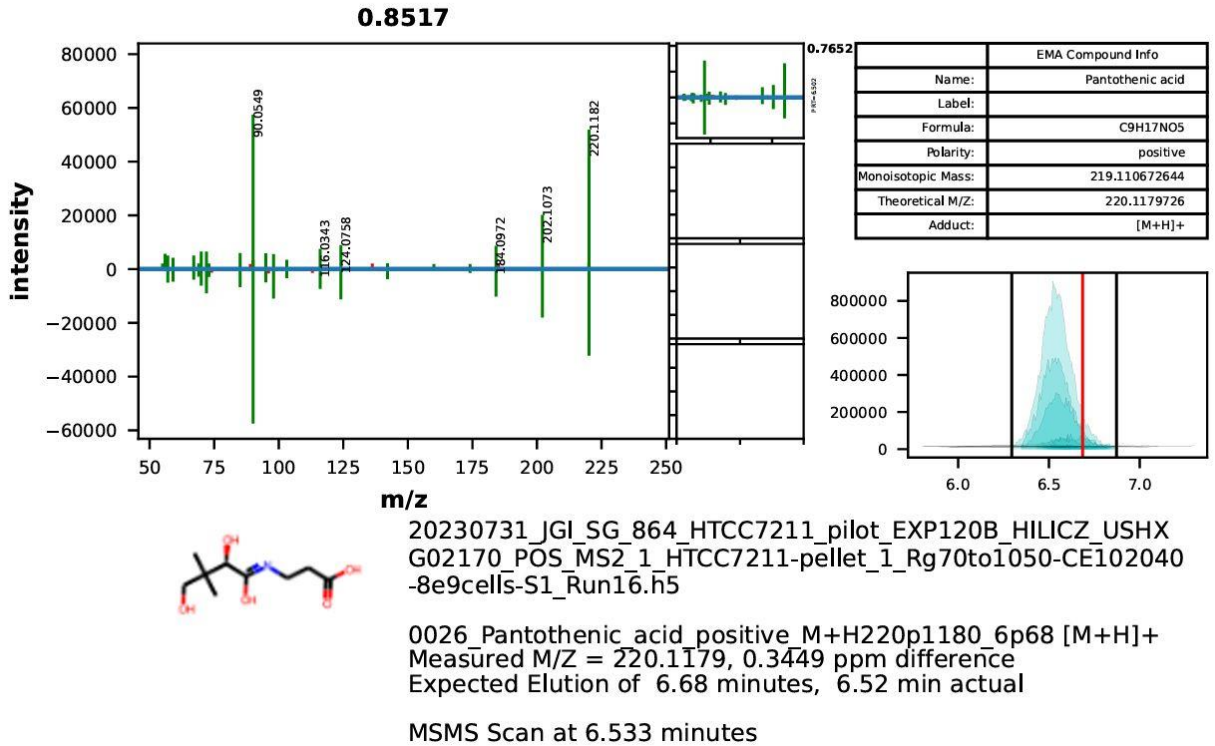
**Other.** Additional controls may be necessary depending on your experimental design.

# Metabolomics Results

For each metabolite (or feature – unique *m/z* & RT pair) identified, ion intensity (peak height or peak area) is reported for each sample and delivered in a spreadsheet format. Semi-quantitative analysis can be performed comparing the relative intensity of a specific metabolite between samples. However, due to differing ionization efficiencies of each metabolite, intensity cannot be compared between different metabolites. Quantitative analysis is possible when a standard is available (preferably an isotopically labeled standard) and a calibration curve has been generated for the experiment. Mass spectrometry data can be challenging to interpret due to a number of factors inherent to the technique, including competitive ionization of metabolites co-eluting in a sample, limits of detection, different metabolites that are isomers or isobars, in-source fragmentation, ionization efficiency and chromatographic resolution. The JGI Metabolomics group is available to discuss and aid with analysis and interpretation.

A metabolite “atlas” is also provided in spreadsheet format. This lists each identified metabolite as well as relevant compound information (e.g. structure; compound ids; description; links to pubchem, chebi, metacyc, etc).

As supplementary / supporting information, pdf images are also provided of (1) chromatograms of each compound in each sample and (2) MS/MS identification figures in which the MS/MS fragmentation spectra of a metabolite in a sample is given on the positive axis and MS/MS of the identified standard (if available) on the negative axis. An example is shown below:



Documentation of methods, workflows and analyses developed and used by JGI Metabolomics is available at protocols.io in the workspace JGI/LBNL Metabolomics Repository. Metabolomics datasets are publicly available at the MassIVEive data repository, and can be easily explored using tools available at GNPS2.

# Metabolite Standards in JGI Library

These are the metabolite standards in our library that we have already run on LC-MS and use for JGI's standard Targeted Environmental Metabolite Atlas (EMA) analysis. Here, we compare retention time, *m/z* and fragmentation spectra between sample and standard to definitively identify a metabolite in a sample.

This is a list of compounds we have run in C18 POS and NEG, and HILICZ POS and NEG.

Reference MSMS spectra for the 1000s of compound standards run at JGI have also been deposited in the GNPS Public Spectral Library, under the library name "BERKELEY-LAB", as a community resource for aiding compound identification.