Metabolomics Workbench
NMDR study submission tutorial

Version 16 (Feb,2024)

Updates from version 15
New requirement for ‘Sample source’ column in study-design section

Eoin Fahy, UCSD

National Metabolomics Data Repository Online Data Submission Tutorial

Before you start:

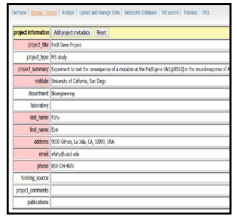
- Have a summary of the study ready (a paragraph describing the goal and design of the experiment for the benefit of the general research community- a publication abstract or equivalent would be ideal).

- Have a study design table ready with sample names and experimental variables in separate columns. Subject information and other measurements may also included.

- Make sure that sample names in submitted results tables exactly match those in the study design table. Otherwise you won't be able to proceed with the submission.

- Collect all relevant protocols and raw data for upload to the WorkBench. Protocol/methods files may be uploaded at appropriate points during the online metadata submission process whereas (large) raw data is uploaded during the final registration step (Step 7).

Online data/metadata submission flowchart



Online GUI's

Project → Study → Subject → Study design

Collection → Treatment → Sample prep.

Sample prep. → Analysis type?

Analysis type? NMR → Load binned NMR data table

Analysis type? MS → # of chromatography methods?

# of chromatography methods? → # of MS methods?

# of MS methods? → # of analyses =#CH * #MS

# of analyses =#CH * #MS → An_1, An_2, An_3, ..., An_n

An_1, An_2, An_3, ..., An_n → Load results and metabolite metadata for each analysis

Chromatography

MS

NMR

mwTab file

Online Data Submission

https://www.metabolomicsworkbench.org/data/DRCCDataDeposit.php

Register/login

Upload/manage studies

Requirements for depositing data via the Metabolomics Workbench:

*****Please read the **NMDR study submission tutorial** first!*****

1. **Register** (you will then be authorized to submit studies)
2. **Log in**.
3. View the list of exemplary studies (below) for examples of best practices for study submissions.
4. The use of the common metabolite names in the **RefMet** database is *strongly encouraged* in order to be able to compare and contrast metabolite data across different experiments and studies. For your convenience an **online tool** is available to map your current metabolite identifications (where possible) to the corresponding RefMet names.
5. Use the **'New Data Upload'** tab to (a) register your study, (b) submit metadata and processed data and (c) upload raw data/supplementary material. Please indicate the date when the study may be made available to the public.
6. **E-mail us for additional assistance** (help@metabolomicsworkbench.org), if needed.

(a): Complete the registration form

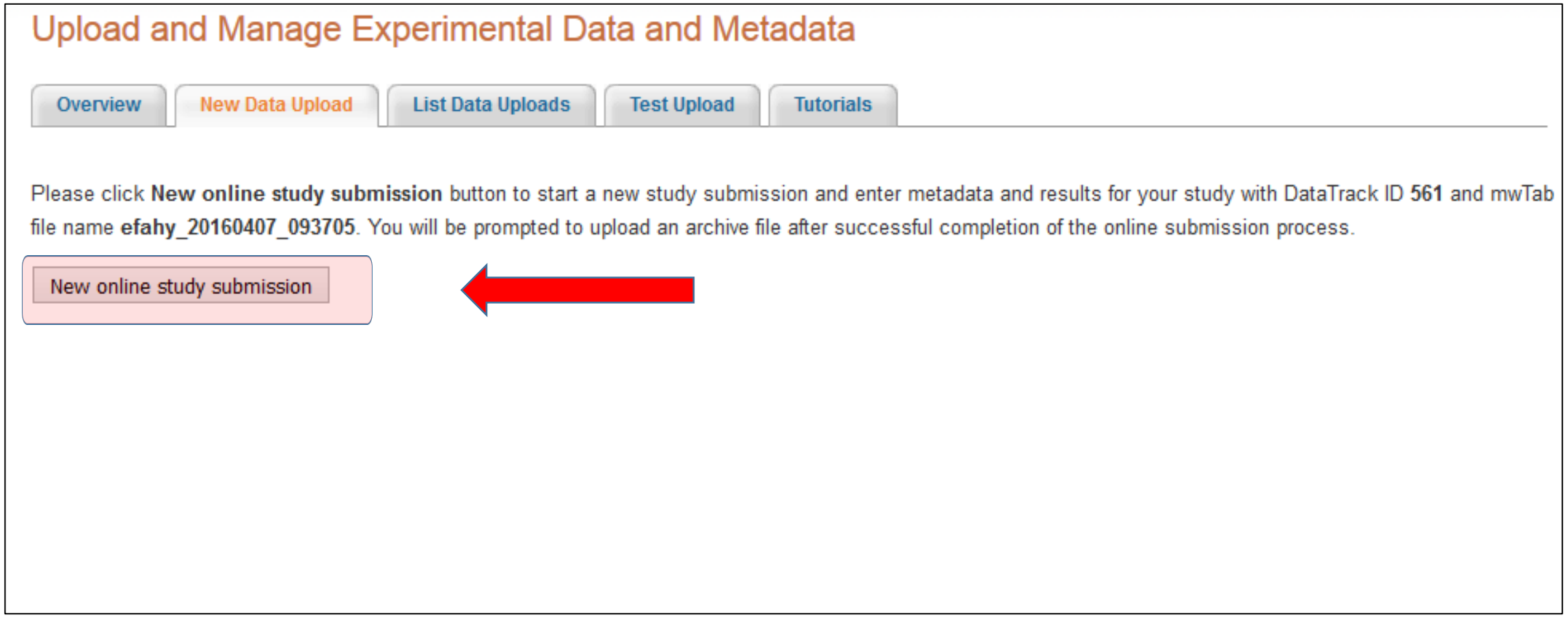
Use separate submissions if your study contains both MS and NMR data

Specify the embargo date if applicable

Please tell us about the data you plan to upload. (* = required)

* mwTab file name	efahy_20210210_135208_mwtab.txt	(Automatically assigned name)
* Name of archive file to be uploaded	EF-_45.zip	(e.g. MyData.zip, MyData.7z or MyData.gz)
* Data type being submitted	MS	(Use separate submissions for studies containing both MS and NMR data)
* Protocol methods filename(s)	PR_SP45.pdf PR_TR45.pdf	
* MS/NMR instrument manufacturer	ABI-SCIEX	
* MS/NMR instrument model	4000-QTRAP	
* Binary data format	.wiff	(e.g. .WIFF (ABI/Sciex), .RAW (Thermo) or .d (Agilent))
* Multi-part study	No	(For multi-part studies, add additional information such as "Study part m of n" in comments field)
* Embargo	Yes	(e.g. If Yes, then please specify date below)
Embargo until	2021-06-12	(e.g. 1 year, 6 months, or YYYY-MM-DD)
Open source text formats	.mzML	

(b): Begin the online submission of metadata and results



(b): Begin the online submission of metadata and results

Start a new study from scratch (most common option, especially for new users)

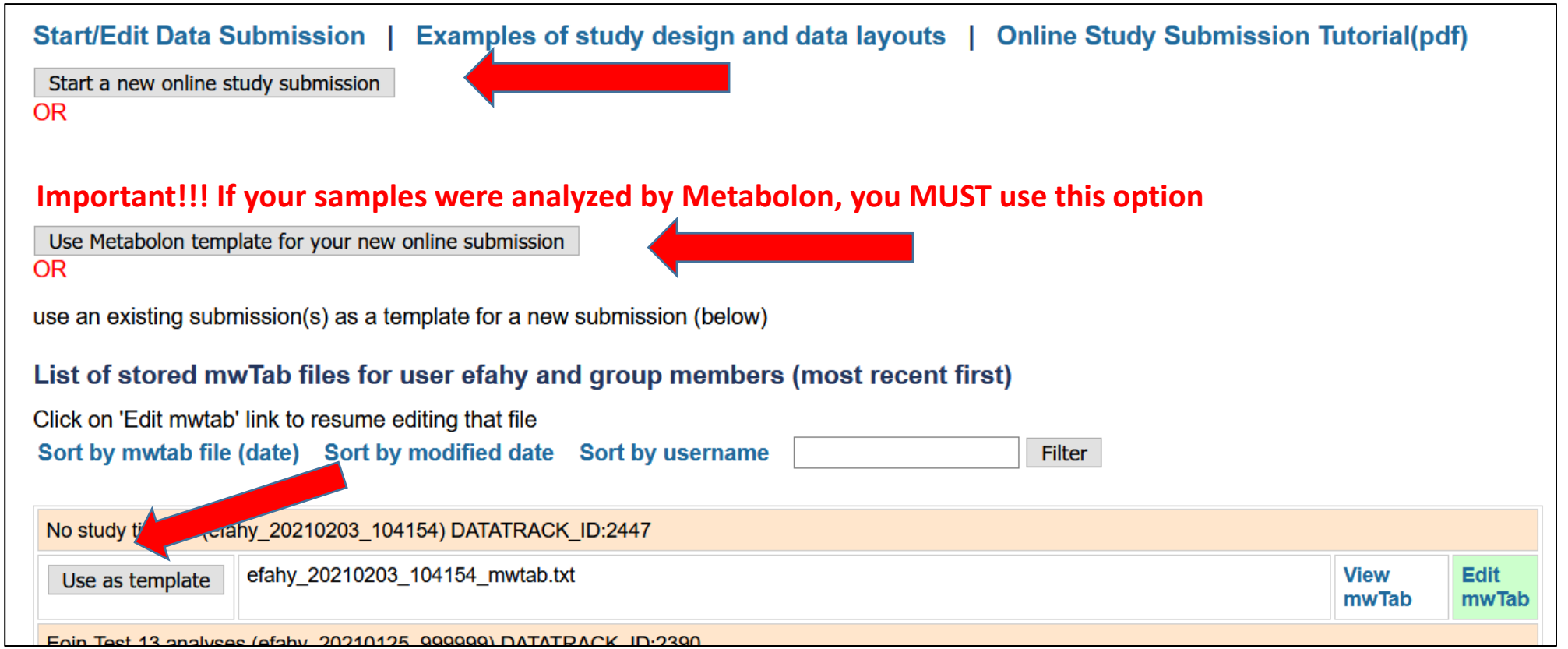
or

use the Metabolon template if the new submission is composed of Metabolon analyses

(If your samples were analyzed by Metabolon, you MUST use this option)

or

use an existing study as a template for a new submission



(b): Begin the online submission of metadata and results
Entering Metabolon data

The Metabolon template on the Metabolomics Workbench has 4 different combined LC/MS methods:
Low pH polar (LC/MS Pos early)
Low pH Lipophilic (LC/MS Pos late)
High pH (LC/MS Neg)
HILIC (LC/MS Polar Neg)
which correspond to 4 different sections (respectively) of the Metabolon results spreadsheet file:
Pos Early
Pos Late
Neg
Polar

Split your Metabolon results (1st column is metabolite names, subsequent columns are sample data) and metabolite metadata (1st column is metabolite names, subsequent columns are Pubchem id, KEGG, SMILES, etc) into these 4 sections based on the "PLATFORM" heading in the spreadsheet. Enter these data in the 4 "Data(Results)" sections of the submission form. Specify units of measurement for the data that you're providing (Unnormalized data, normalized-Imputed Data, log-transformed data, etc.)

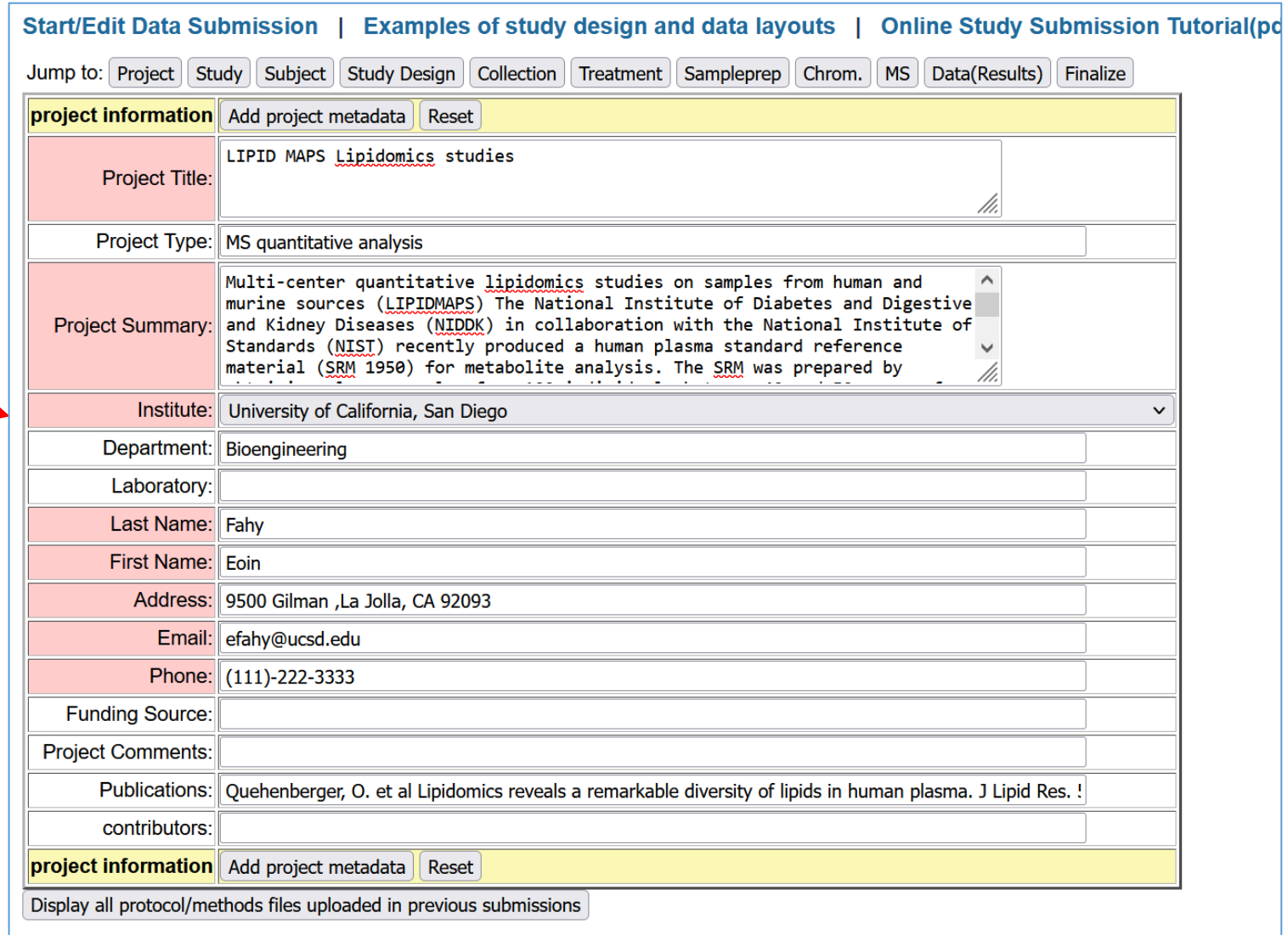
PLATFORM in Metabolon results spreadsheet	Workbench template results section	LC/MS method	MS Polarity
Pos Early	1	Low pH polar (LC/MS Pos early)	POS
Pos Late	2	Low pH Lipophilic (LC/MS Pos late)	POS
Neg	3	High pH (LC/MS Neg)	NEG
Polar	4	HILIC (LC/MS Polar)	NEG

Project information

Items in pink are required fields

Personal information such as name, address, email is autopopulated in the form based on your login credentials (but you may edit these fields if not correct or appropriate)

If the item you want is not available in a pull-down menu, choose "Add new item" from the list and type in your own value



Study information

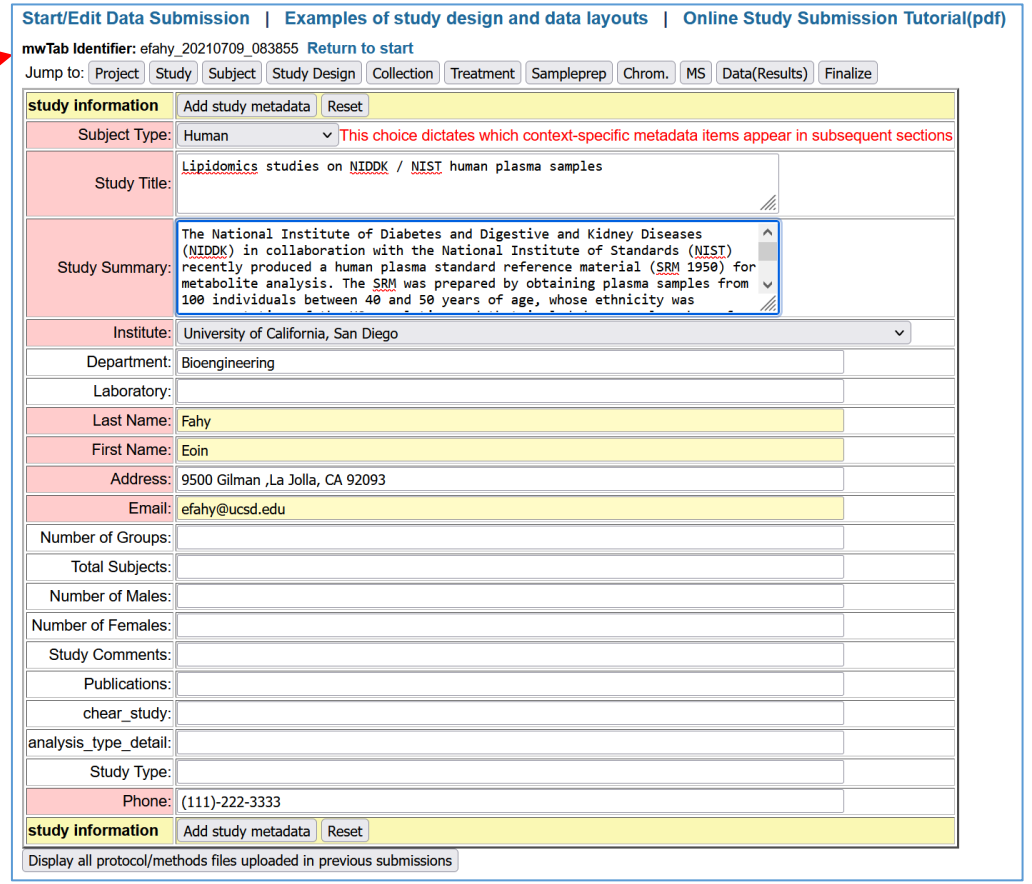
Subject type is mandatory and creates context-specific metadata items in subsequent sections

Study title should be unique (if you're submitting multiple studies)

Study summary is **very important** in order to describe the objectives of the experiment to the general public. Ideally it should be a paragraph similar to an abstract in a publication

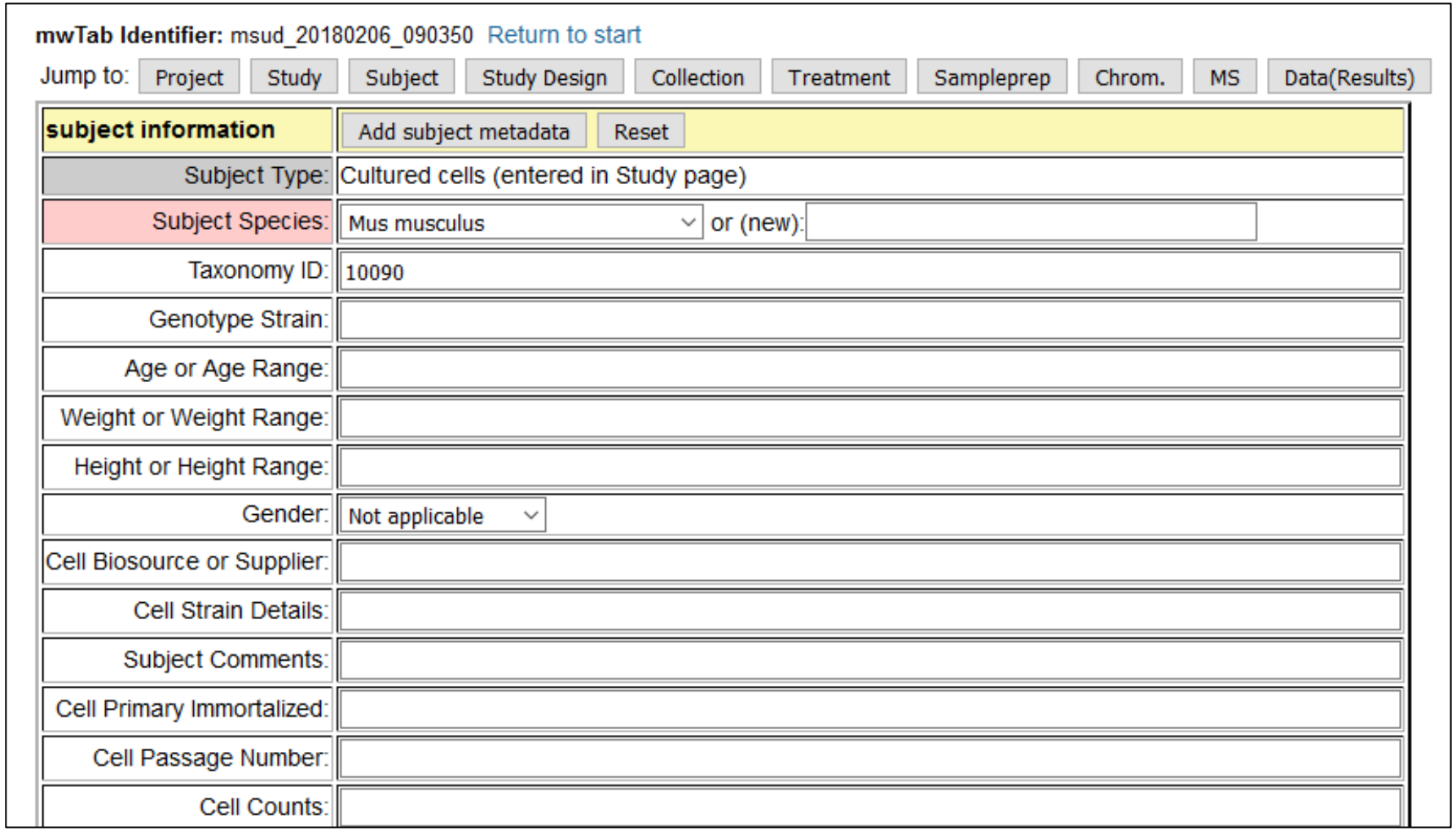
Personal information such as name, address, email is autopopulated in the form based on your login credentials (but you may edit these fields if not correct or appropriate)

This is your unique Submission identifier (contains your login name and date/time)



Subject information

Choose subject species from pulldown menu or enter a new species (Latin name)
In cases where metabolites from multiple species are being assayed in the same experiment, separate the species names with a "/"



Study design information

This section contains essential study design information for the study which must include sample identifiers, sample source and at least one experimental variable(factor) in tabular format. An additional "subject_id" column relating the samples to a particular source (patient, animal, cell etc.) may also be included. Additional information unique to each sample (e.g. height, weight, BMI, age, assay measurement, etc.) may also be included but should NOT be designated as factors (Designate these as "Other" in the next step)

Include a column with raw data file names. Use multiple raw file columns if you have more than 1 raw file per sample (e.g. different LC methods, ion polarity)

Subject_ID	Sample_ID	Sample source	Genotype	Treatment	Batch	RAW_FILE_NAME
CA11	CA11W0	Liver	Wild-type	Control	B1a	SC_CA11W0.mzML
CA12	CA12W0	Liver	Wild-type	Control	B1a	SC_CA12W0.mzML
CA13	CA13W0	Liver	Wild-type	Control	B1a	SC_CA13W0.mzML
CA11	CA11W50	Liver	Wild-type	50uM	B1a	SC_CA11W50.mzML
CA12	CA12W50	Liver	Wild-type	50uM	B1a	SC_CA12W50.mzML
CA13	CA13W50	Liver	Wild-type	50uM	B1a	SC_CA13W50.mzML
CA14	CA14W0	Liver	Mutant	Control	B1a	SC_CA14W0.mzML
CA15	CA15W0	Liver	Mutant	Control	B1a	SC_CA15W0.mzML
CA16	CA16W0	Liver	Mutant	Control	B1a	SC_CA16W0.mzML
CA17	CA17W50	Liver	Mutant	50uM	B1a	SC_CA17W50.mzML
CA18	CA18W50	Liver	Mutant	50uM	B1a	SC_CA18W50.mzML
CA19	CA19W50	Liver	Mutant	50uM	B1a	SC_CA19W50.mzML

Start/Edit Data Submission | Examples of study design and data layouts | Online Study Submission

mwTab Identifier: efahy_20230814_171135 Return to start

IMPORTANT!!

Make sure that sample names in submitted results table(s) or file(s) exactly match those in the study design table.
Otherwise you won't be able to complete the submission.
(One needs to be able to relate experimental conditions in the study-design section via sample names in ALL submitted datasets)

Input **Study Design** information. Sample names, experimental factor(s),sample source and raw data file.
Subject name and additional sample data are optional.
First row must contain headings. Data must be tab-delimited.

Copy/paste as tab-delimited data from Excel or text file (View the "See examples.." link for more help)

Subject_ID	Sample_ID	Sample source	Genotype	Treatment	Batch	RAW_FILE_NAME
CA11	CA11W0	Liver	Wild-type	Control	B1a	SC_CA11W0.mzML
CA12	CA12W0	Liver	Wild-type	Control	B1a	SC_CA12W0.mzML
CA13	CA13W0	Liver	Wild-type	Control	B1a	SC_CA13W0.mzML
CA11	CA11W50	Liver	Wild-type	50uM	B1a	SC_CA11W50.mzML
CA12	CA12W50	Liver	Wild-type	50uM	B1a	SC_CA12W50.mzML
CA13	CA13W50	Liver	Wild-type	50uM	B1a	SC_CA13W50.mzML
CA14	CA14W0	Liver	Mutant	Control	B1a	SC_CA14W0.mzML
CA15	CA15W0	Liver	Mutant	Control	B1a	SC_CA15W0.mzML
CA16	CA16W0	Liver	Mutant	Control	B1a	SC_CA16W0.mzML
CA17	CA17W50	Liver	Mutant	50uM	B1a	SC_CA17W50.mzML

View/check study design See examples of study design layout

Then click on "View/check study design" to view in tabular format

Study design information

Instructions:

Sample names/identifiers in the required **'Sample_ID'** column should be unique and should exactly match those names used in the processed results.
The required **'Sample source'** column (e.g. blood, urine, HEK cells, blank, buffer) must be completed. This may be the same or different for all samples
The required **'Raw file name'** column must be completed when submitting raw data. The sample name to raw file name mapping is essential in order to enable re-analysis of raw data.
The optional **'Subject_ID'** column may be used to designate the submitter's source identifier for a given sample (e.g. subject/patient/animal identifer).
The required **'Factor'** column(s) are used to assign experimental variables (factors) to sample groups (e.g. treatment condition, time, genotype, phenotype, etc.).Use an appropriate name for the factor heading (e.g. Genotype, Time, Drug treatment) - don't use 'Factor' which is too vague.
The optional **'Other'** column(s) may be used to include additional data such as BMI, age, glucose measurements, etc. that are unique to each sample.
These types of measurements should NOT be designated as factors.

Assign every column(below) as 'Subject ID'(optional), 'Sample ID'(required:1 and only 1), 'Sample source'(required:1 and only 1), 'Factor'(required:at least 1) or 'Other'(optional additional sample data). Columns assigned 'Ignore' will be ignored.

Process study design data **Does submission contain raw data?:**

Subject_ID	Sample_ID	Sample source	Genotype	Treatment	Batch	RAW_FILE_NAME
CA11	CA11W0	Liver	Wild-type	Control	B1a	SC_CA11W0.mzML
CA12	CA12W0	Liver	Wild-type	Control	B1a	
CA13	CA13W0	Liver	Wild-type	Control	B1a	
CA11	CA11W50	Liver	Wild-type	50uM	B1a	
CA12	CA12W50	Liver	Wild-type	50uM	B1a	

Assign each column as "Subject_ID", "Sample_ID", "Factor", "Raw file name", "Other" or "Ignore"

Process study design data **Does submission contain raw data?:** Yes

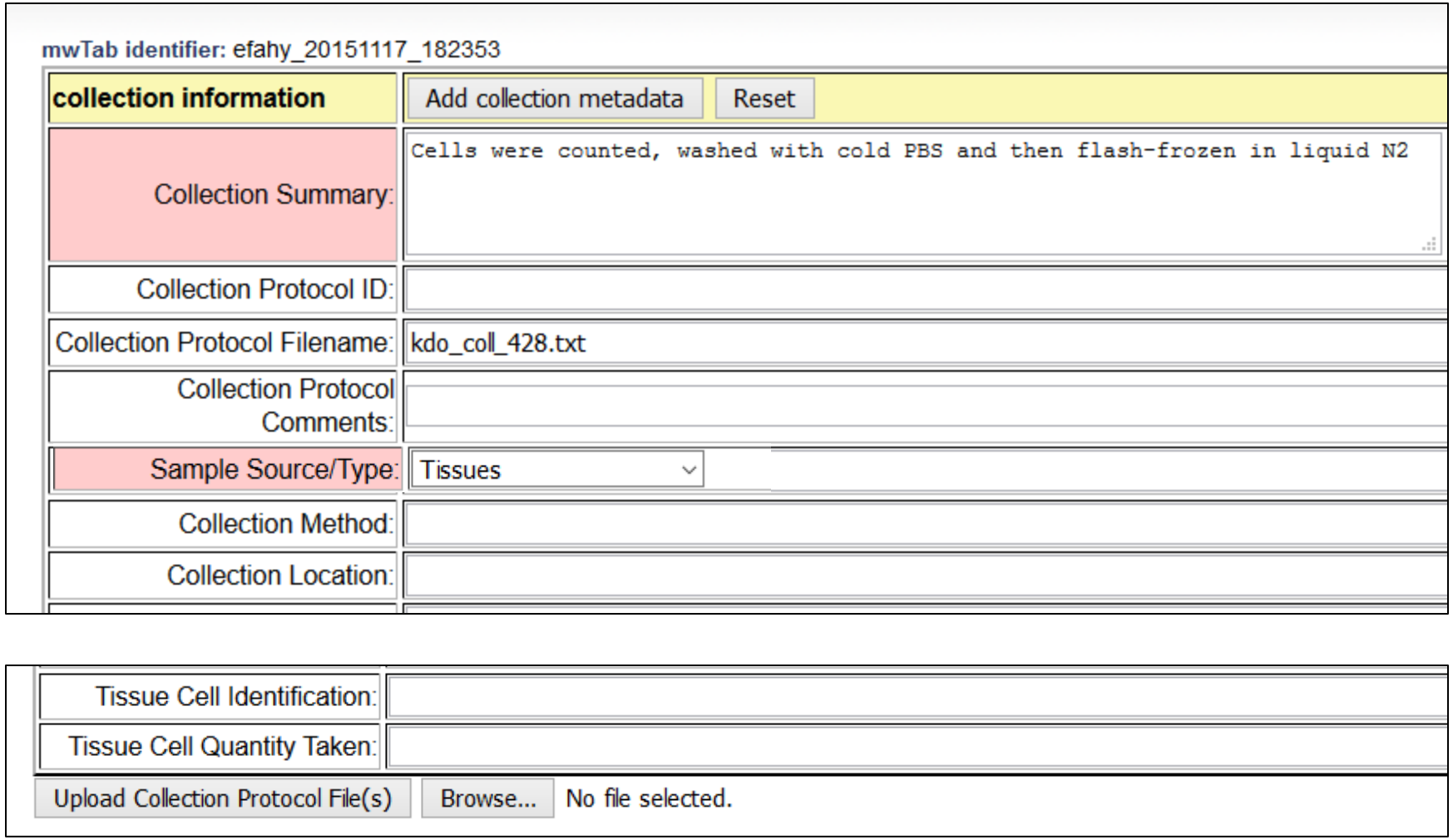
Subject_ID	Sample_ID	Sample source	Factor	Factor	Other	Raw file name
Subject_ID	Sample_ID	Sample source	Genotype	Treatment	Batch	RAW_FILE_NAME
CA11	CA11W0	Liver	Wild-type	Control	B1a	SC_CA11W0.mzML
CA12	CA12W0	Liver	Wild-type	Control	B1a	SC_CA12W0.mzML
CA13	CA13W0	Liver	Wild-type	Control	B1a	SC_CA13W0.mzML
CA11	CA11W50	Liver	Wild-type	50uM	B1a	SC_CA11W50.mzML
CA12	CA12W50	Liver	Wild-type	50uM	B1a	SC_CA12W50.mzML

After editing/assignment, proceed by clicking on "Process study design data"

Collection information

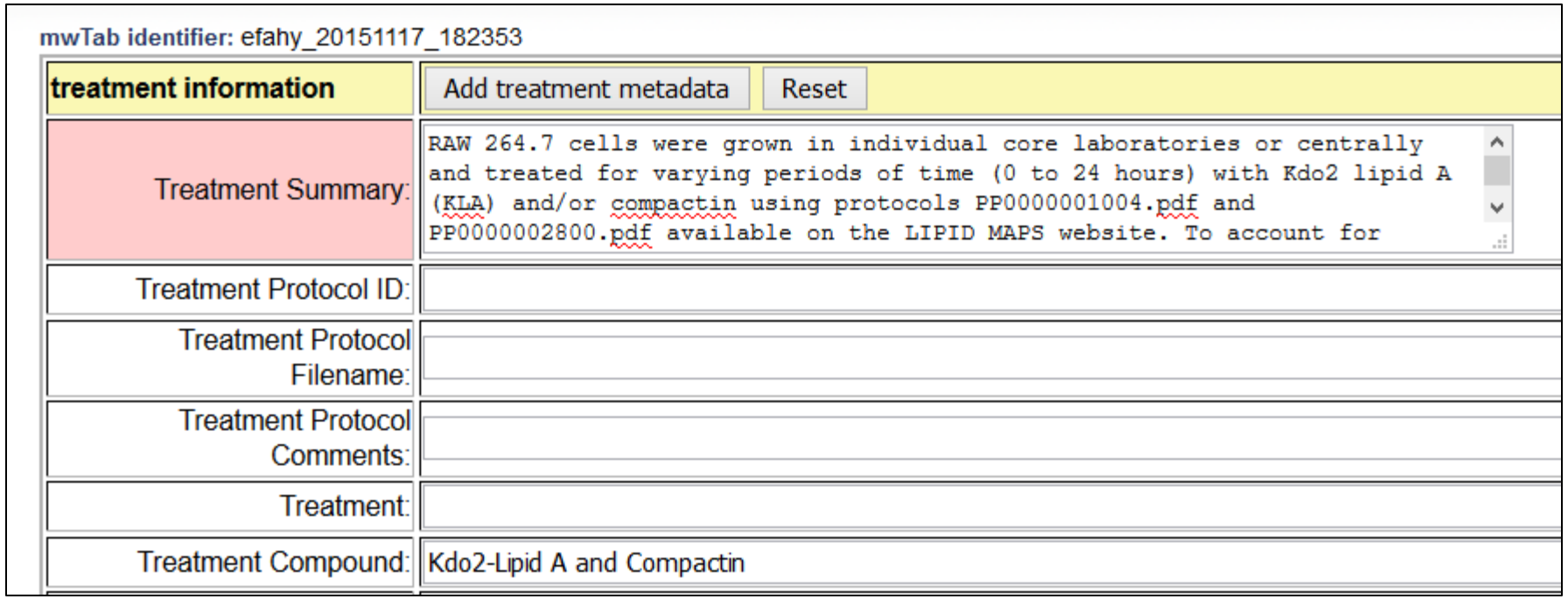
Metadata related to sample collection/ handling/storage

User may upload a methods/protocol file relating to sample collection



Treatment information

Metadata related to treatment protocols



mwTab identifier: efahy_20151117_182353

treatment information	Add treatment metadata Reset
Treatment Summary:	RAW 264.7 cells were grown in individual core laboratories or centrally and treated for varying periods of time (0 to 24 hours) with Kdo2 lipid A (KLA) and/or compactin using protocols PP0000001004.pdf and PP0000002800.pdf available on the LIPID MAPS website. To account for
Treatment Protocol ID:	
Treatment Protocol Filename:	
Treatment Protocol Comments:	
Treatment:	
Treatment Compound:	Kdo2-Lipid A and Compactin

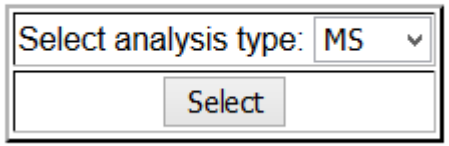
Sample prep. information

Metadata related to sample preparation protocols

mwTab identifier: efahy_20151117_182353

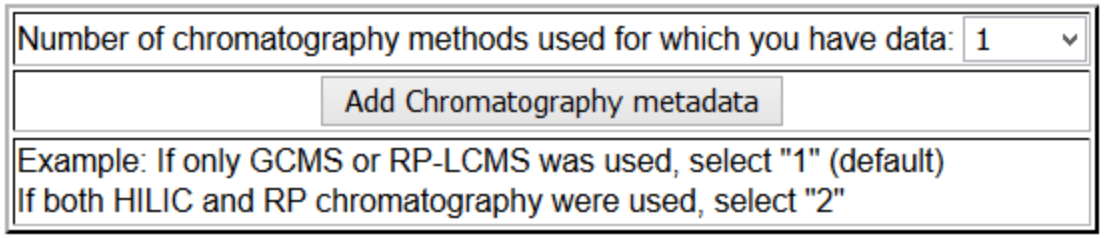
sampleprep information	Add sampleprep metadata Reset
Sampleprep Summary:	Total lipids were extracted from the cell suspension (Bligh/Dyer). Ice-cold methanol (2.5 ml) was added to each 1 ml of DPBS containing the scraped cell suspension. A volume containing 600 pmol of each of the 18 d5-labeled DAG and TAG internal standards in toluene/methanol (1:1) was
Sampleprep Protocol ID:	
Sampleprep Protocol Filename:	
Sampleprep Protocol Comments:	
Processing Method:	
Processing Storage Conditions:	

Decision point: MS or NMR experiment?



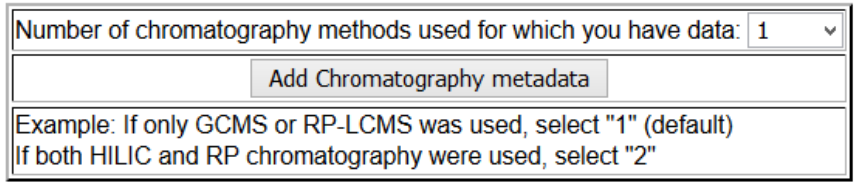
If MS is chosen, the user is prompted to enter chromatography information

Choose number of chromatography methods for which you have data (default=1)

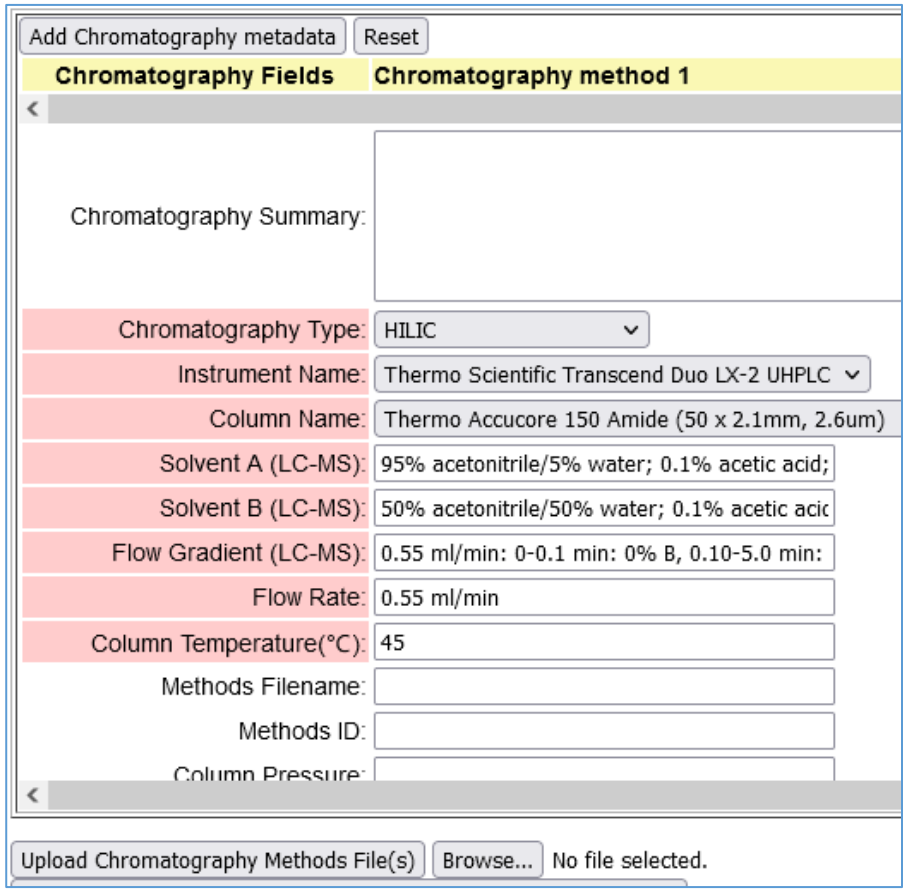


Chromatography information

Metadata related to chromatography (LC/GC) protocols

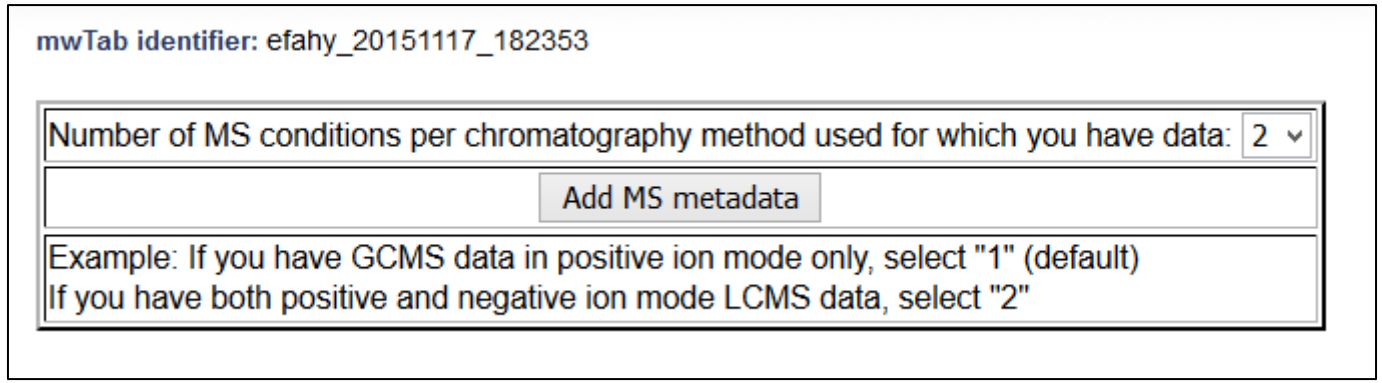


User may upload a methods/protocol file relating to chromatography



Number of MS conditions per chromatography method

In the case of LCMS this is typically 2 (Positive and negative ion mode)

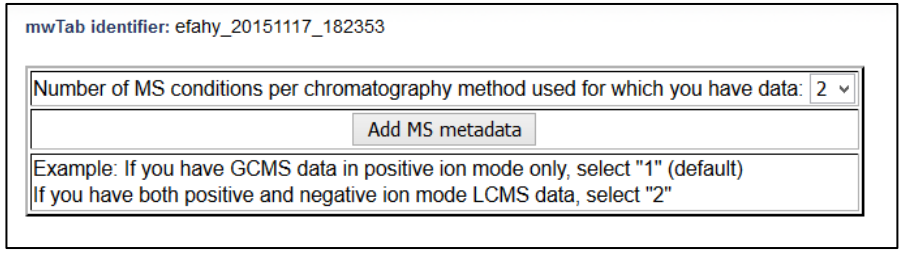


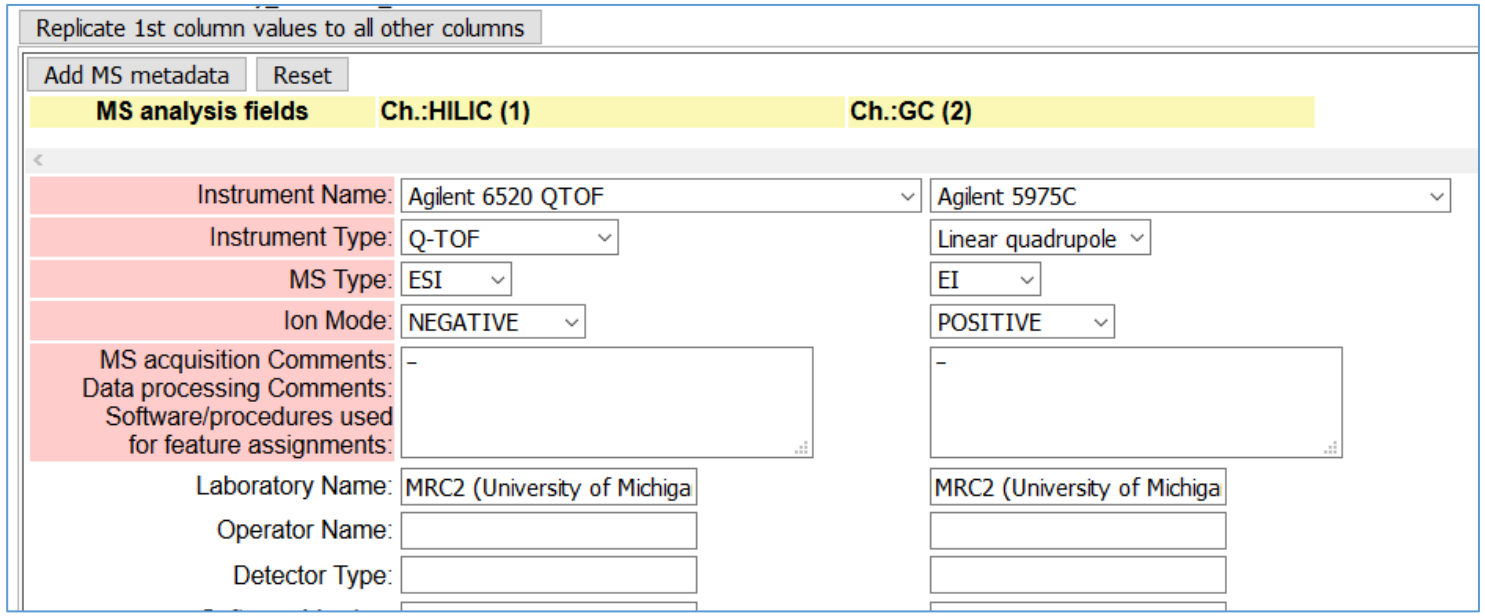
MS information

Metadata related to MS methods

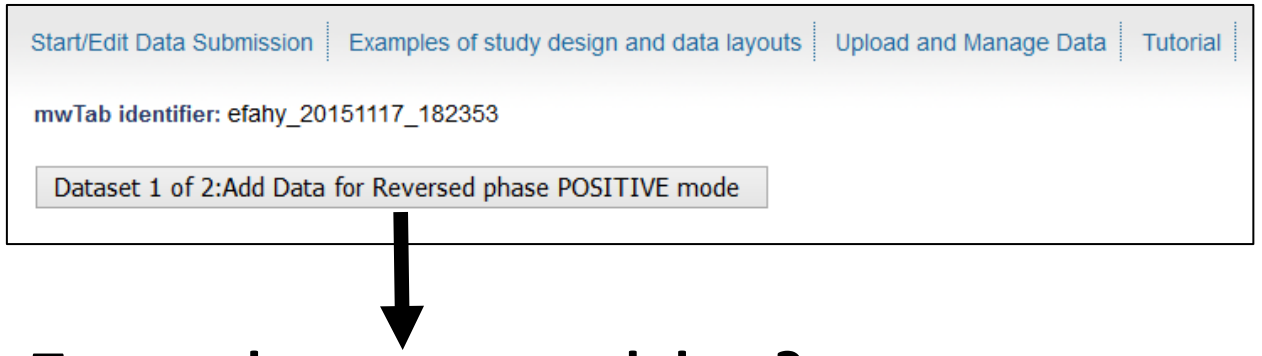
The number of data columns will equal the number of chromatography conditions multiplied by the number of MS conditions. For example, if reversed-phase and HILIC chromatography were specified in conjunction with 2 MS modes (+ and – mode detection), then 2x2=4 columns are displayed

Hint: Fill out the parameters in the column on the left only, click the "Replicate.." button to copy the content to the other columns, then adjust any unique values as appropriate





Enter processed data for each chromatography/MS combination that you have specified



Decision point: Targeted or untargeted data?

Option1: Measurements for **named** metabolites from **targeted** experiments, e.g. GC-MS analyses or LC-MS assays with known standards.

Option2: Measurements from **untargeted** experiments e.g. high-resolution LC-MS analyses. Detected features are typically m/z-retention time values

OPTION 1: Targeted assays containing identified (named) metabolites

Input Data in tab-delimited format in the text area below.
First column must contain metabolite names.
Subsequent columns must contain sample data with identical sample names as in Study Design submission.
First row must contain sample names.

Tabular results (typically tens or hundreds of named metabolites) are pasted into this textarea

Units of measurement (required): xxx

View/check metabolite data | See examples of metabolite data layout | Delete existing metabolite data (this analysis only)

OPTION 2: Untargeted assays not containing identified (named) metabolites

(e.g. datasets with m/z,retention time features from LC-MS experiments, NMR binned data)

A file of tabular results (typically thousands of unidentified features) is uploaded here

Add/replace results as a tab-delimited text file:
Ideally, feature names should be formatted as 'm/z underscore retention time', e.g. 645.5327_
IMPORTANT!:If unidentified featues are listed by neutral mass rather than m/z ratio, this must be specified in the menu below
This will enable analysis of the dataset by a larger number of tools on the Metabolomics Workbench

Units of measurement (required):	Feature names contain m/z values*? (required):	Feature names contain retention time values? (required):	Time units:

*** By "m/z values" we are referring to mass-to-charge ratios and NOT neutral masses.**

Upload tab-delimited datafile | Browse... | No file selected.

The first line in the submitted file should contain sample names exactly matching those that you submitted in the 'Study Design' section.

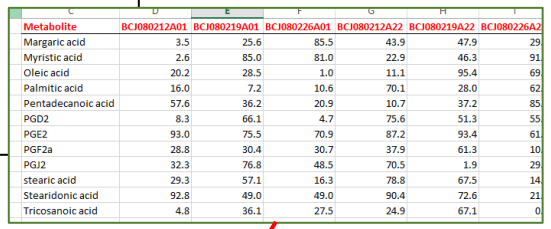
Option1:
Targeted data

Enter processed data (metabolite identifications and measurements) for each chromatography/MS combination that you have specified

Start/Edit Data Submission | Examples of study design and data layouts | Upload and Manage Data | Tutorial

mwTab identifier: efahy_20151117_182353

Dataset 1 of 2:Add Data for Reversed phase POSITIVE mode



Metabolite	BCJ080212A01	BCJ080219A01	BCJ080226A01	BCJ080212A22	BCJ080219A22	BCJ080226A2
Margaric acid	3.5	25.6	85.5	43.9	47.9	29
Myristic acid	2.6	85.0	81.0	22.9	46.3	91
Oleic acid	20.2	28.5	1.0	11.1	95.4	69
Palmitic acid	16.0	7.2	10.6	70.1	28.0	62
Pentadecanoic acid	57.6	36.2	20.9	10.7	37.2	85
PGD2	8.3	66.1	4.7	75.6	51.3	55
PGE2	93.0	75.5	70.9	87.2	93.4	61
PGF2a	28.8	30.4	30.7	37.9	61.3	10
PGJ2	32.3	76.8	48.5	70.5	1.9	29
stearic acid	29.3	57.1	16.3	78.8	67.5	14
Stearidonic acid	92.8	49.0	49.0	90.4	72.6	21
Tricosanoic acid	4.8	36.1	27.5	24.9	67.1	0

mwTab identifier: efahy_20151117_182353

Input Data in tab-delimited format. First column must contain metabolite names.
Subsequent columns must contain sample data with identical sample names as in Study Design submission.
First row must contain sample names.

Copy/paste tabular data here
First row MUST contain sample names identical to those submitted in the "study design" step. First column must contain metabolite names or m/z-retention time identifiers in the case of unidentified ions (e.g. "231.4185_17.68")

Units of measurement:

View/check metabolite data See examples of metabolite data layout

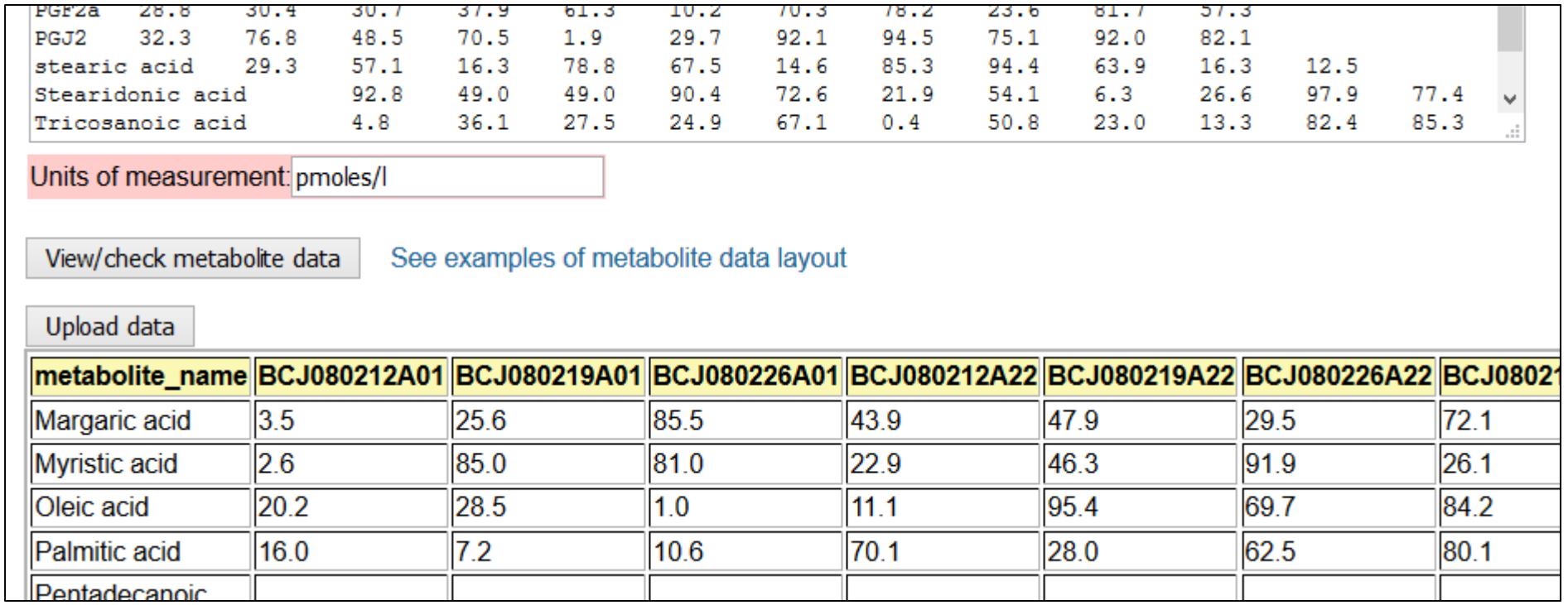
Units is a required field

View the "See examples.." link for examples of datasets

Option1:
Targeted data

Processed data upload: Review in tabular form, then Upload data

Note: sample names must match those submitted in the "study design" section, otherwise a warning will be generated and this must be resolved before proceeding



PGF2a	28.8	30.4	30.7	37.9	61.3	10.2	70.3	78.2	23.6	81.7	57.3	
PGJ2	32.3	76.8	48.5	70.5	1.9	29.7	92.1	94.5	75.1	92.0	82.1	
stearic acid	29.3	57.1	16.3	78.8	67.5	14.6	85.3	94.4	63.9	16.3	12.5	
Stearidonic acid		92.8	49.0	49.0	90.4	72.6	21.9	54.1	6.3	26.6	97.9	77.4
Tricosanoic acid		4.8	36.1	27.5	24.9	67.1	0.4	50.8	23.0	13.3	82.4	85.3

Units of measurement: pmoles/l

View/check metabolite data See examples of metabolite data layout

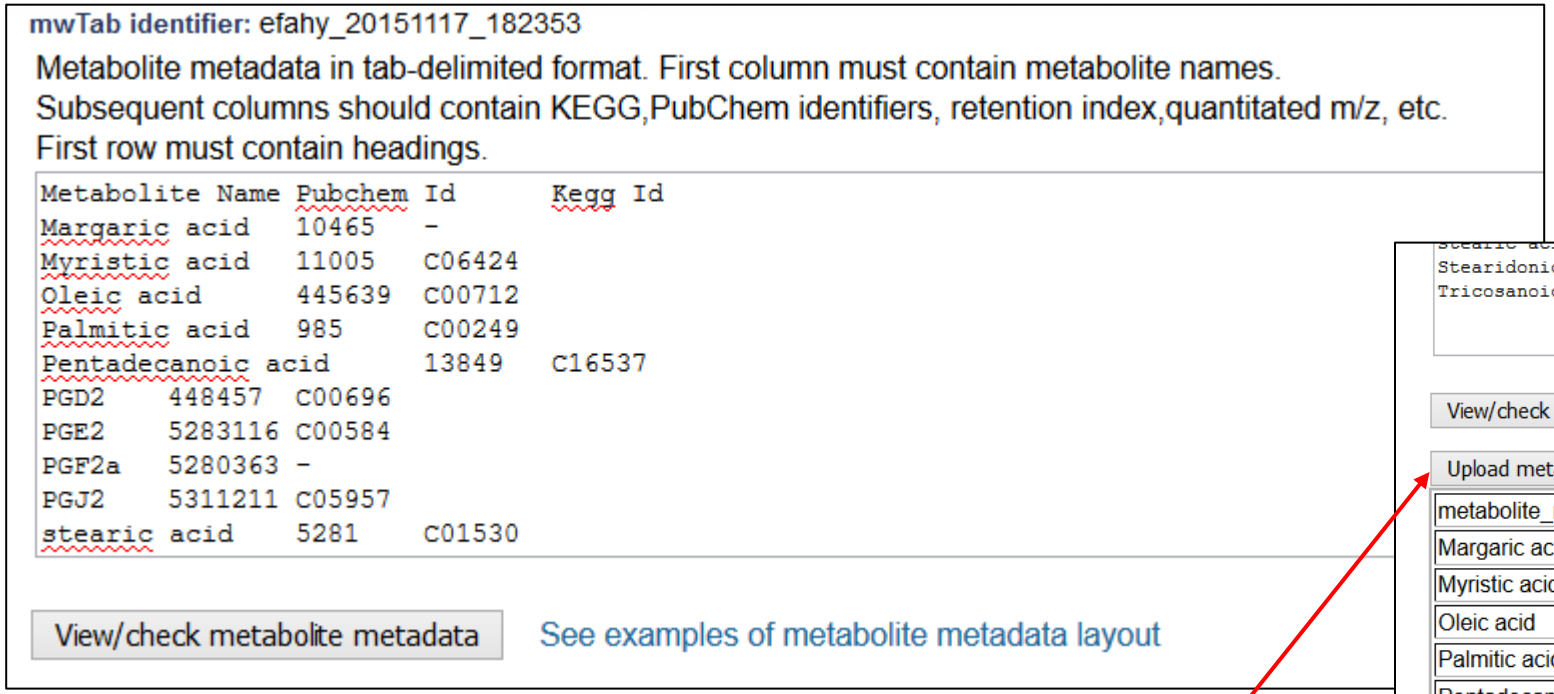
Upload data

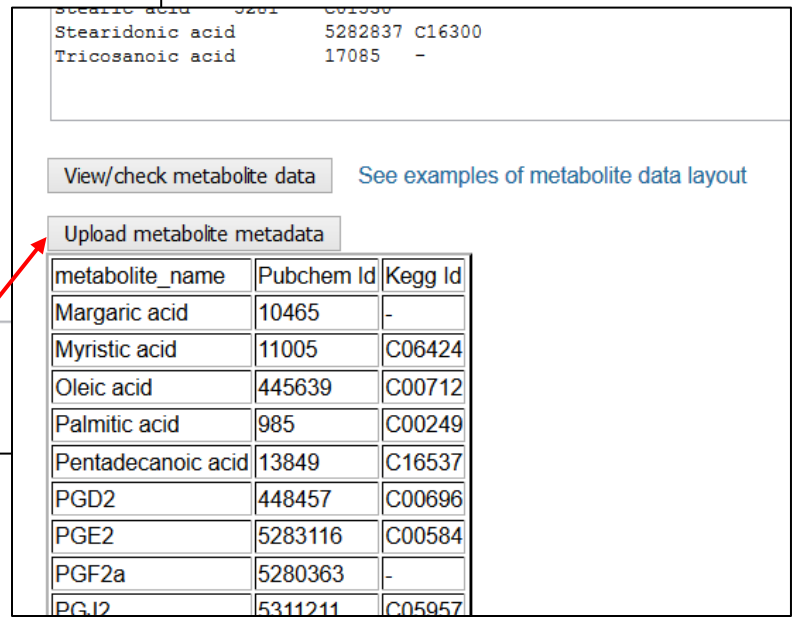
metabolite_name	BCJ080212A01	BCJ080219A01	BCJ080226A01	BCJ080212A22	BCJ080219A22	BCJ080226A22	BCJ0802
Margaric acid	3.5	25.6	85.5	43.9	47.9	29.5	72.1
Myristic acid	2.6	85.0	81.0	22.9	46.3	91.9	26.1
Oleic acid	20.2	28.5	1.0	11.1	95.4	69.7	84.2
Palmitic acid	16.0	7.2	10.6	70.1	28.0	62.5	80.1
Pentadecanoic							

Option1:
Targeted data

Metabolite metadata upload

Copy/paste metabolite annotations in tabular format (PubChem CID, KEGG ID, InCHi Key, LC/GC retention time/index, etc.) Metabolite names MUST match those submitted in the previous data section. If you don't have any metabolite annotations, just submit the column of metabolite names.

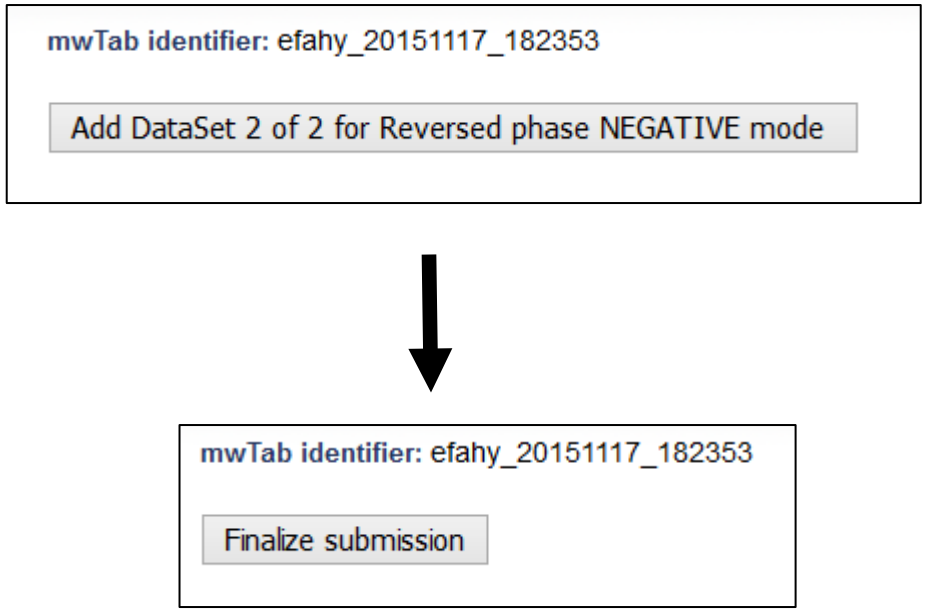




After checking the table of metabolite annotations, click "Upload metabolite metadata"

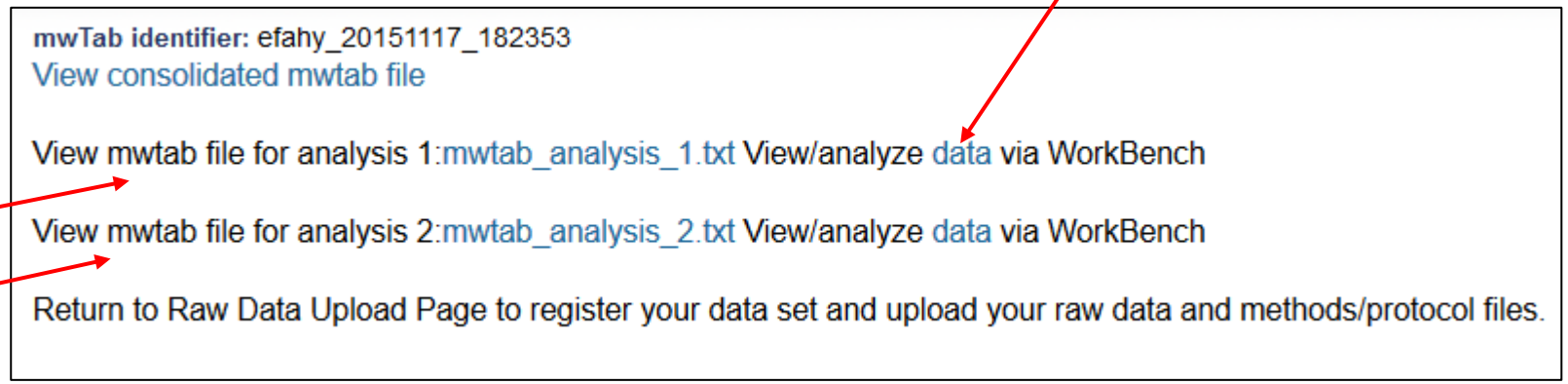
Option1:
Targeted data

Repeat the data/metabolite metadata upload steps for each chromatography/MS analysis combination that you have specified



When upload steps are complete click on "Finalize submission"

Users may view and analyze the study to review the data/metadata. This viewer simulates how the study will appear on the Metabolomics WorkBench after curation and database upload



Separate mwTab files are generated for each analysis combination

Option2:
Untargeted data

Measurements from untargeted experiments e.g. high-resolution LC-MS analyses are uploaded as a tab-delimited text file containing a table of unidentified features (typically m/z-retention time values) and associated measurements.

OPTION 2: Untargeted assays not containing identified (named) metabolites

(e.g. datasets with m/z,retention time features from LC-MS experiments, NMR binned data)

Add/replace results as a tab-delimited text file:
Ideally, feature names should be formatted as 'm/z underscore retention time', e.g. 645.5327_24.91
IMPORTANT!:If unidentified featues are listed by neutral mass rather than m/z ratio, this must be specified in the menu below
This will enable analysis of the dataset by a larger number of tools on the Metabolomics Workbench

Units of measurement (required):	Feature names contain m/z values*? (required):	Feature names contain retention time values? (required):	
Peak area	Yes	Yes	Time units: Minutes

*** By "m/z values" we are referring to mass-to-charge ratios and NOT neutral masses.**

Upload tab-delimited datafile | Browse... | jwalejko_20181204_201054_mwtab.txt

The first line in the submitted file should contain sample names exactly matching those that you submitted in the 'Study Design' section.

Select results file from your file system. Sample names should exactly match those submitted in the "Study Design" section of the metadata submission

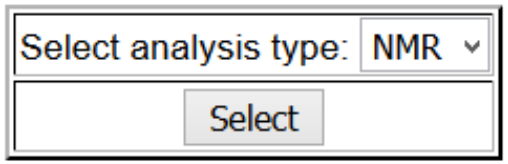
C:\Users\eoinf\Downloads\untargeted_data_table.txt

	Feature(m/z_RT)	samp1	samp2	samp3	samp4	sam
2	100.02005_15.5	8875.5	9273.9	1559.0	1160.0	894
3	100.07742_65.4	2744.3	2152.3	6895.3	9465.8	212
4	101.06952_73.9	6646.6	3736.5	1458.4	9832.6	653
5	102.08992_29.2	4164.2	2195.9	8447.9	1920.1	274
6	102.08983_25.0	8187.6	8647.8	4984.4	9747.3	741
7	103.05251_42.6	2432.0	2431.9	4988.6	3383.4	820
8	103.78777_17.7	5714.7	3217.8	4914.0	8954.6	414
9	104.05150_20.6	9814.3	8541.1	6641.6	2744.3	215
10	104.06962_16.9	1481.1	1368.9	2780.0	2206.6	513
11	104.10595_11.6	5430.2	6389.2	8495.9	9654.2	848
12	104.10562_72.7	2614.9	2431.9	2140.9	9045.2	155
13	104.99081_88.7	6193.2	5506.5	7210.6	5457.4	991
14	106.04841_24.7	5995.0	8896.7	4185.6	2675.6	556
15	106.08454_13.2	2862.1	9659.3	2016.6	1539.5	527
16	108.01019_20.6	5768.7	4539.3	4992.9	1156.6	166
17	109.09961_16.9	4128.3	5113.5	6015.4	8823.3	348
18	110.05838_71.2	9221.6	1079.8	7146.5	8210.4	155
19	110.06358_79.4	5995.0	8896.7	1570.0	2258.1	991
20	110.10667_56.9	602.8	1942.7	4983.4	1102.9	556

Example of a file with untargeted MS data. Note the 1st column contains m/z_retention time features. Subsequent columns contain measurements for each sample.

Decision point: MS or NMR experiment?

NMR option



mwTab identifier: efahy_20151112_141949

nmr information	Add nmr metadata / Reset
Instrument Name:	Bruker Avance III
Instrument Type:	FT-NMR
NMR Experiment Type:	1D-1H
NMR Comments:	
Field Frequency Lock:	Deuterium
Standard Concentration:	0.5 mM
Spectrometer Frequency:	950 MHz
NMR Probe:	cryo, inverse
NMR Solvent:	D2O
NMR Tube Size:	5mm x 7 in
Shimming Method:	Topshim

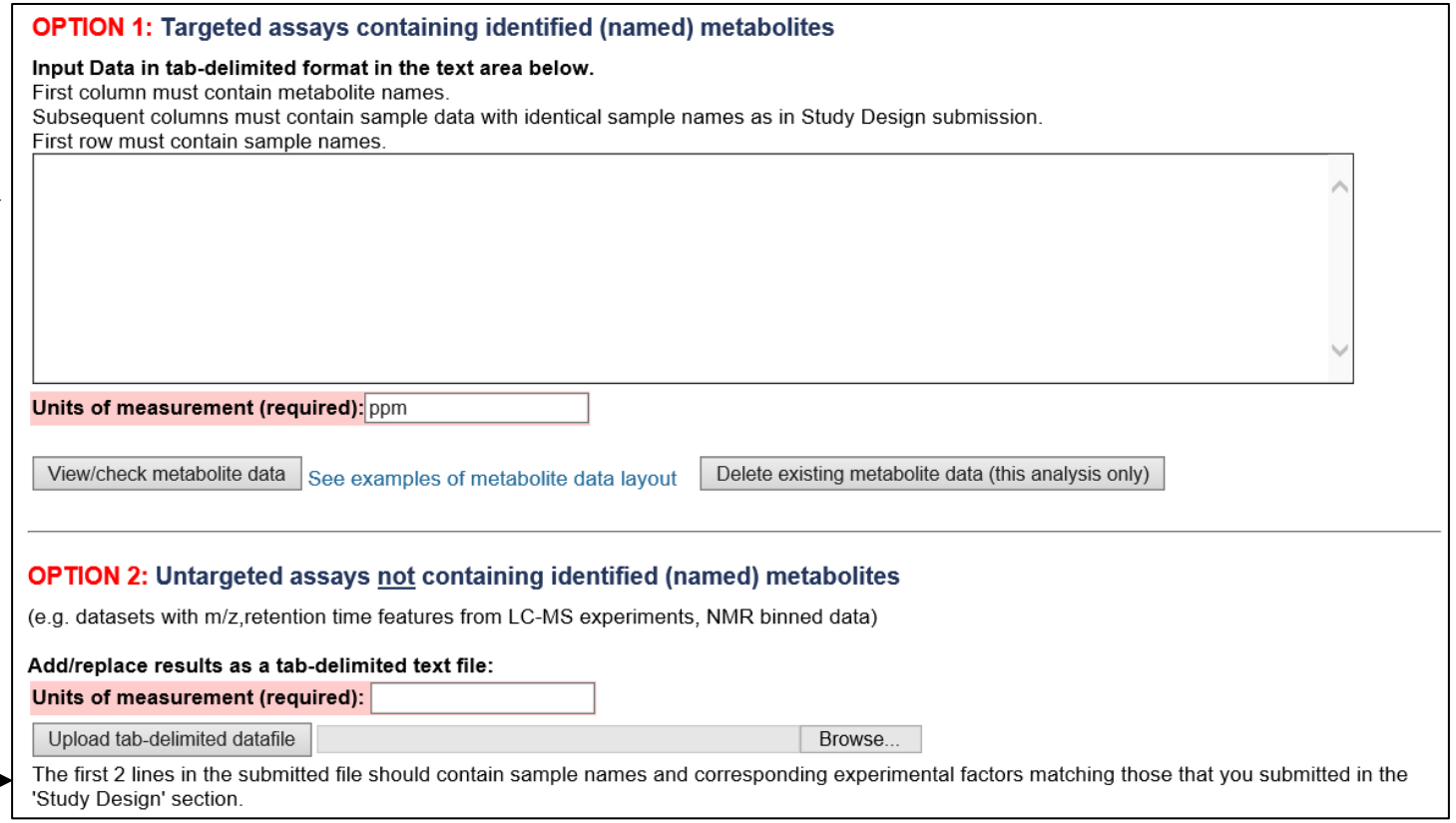
NMR experiment option

Add NMR results data

Decision point: Targeted or untargeted data?

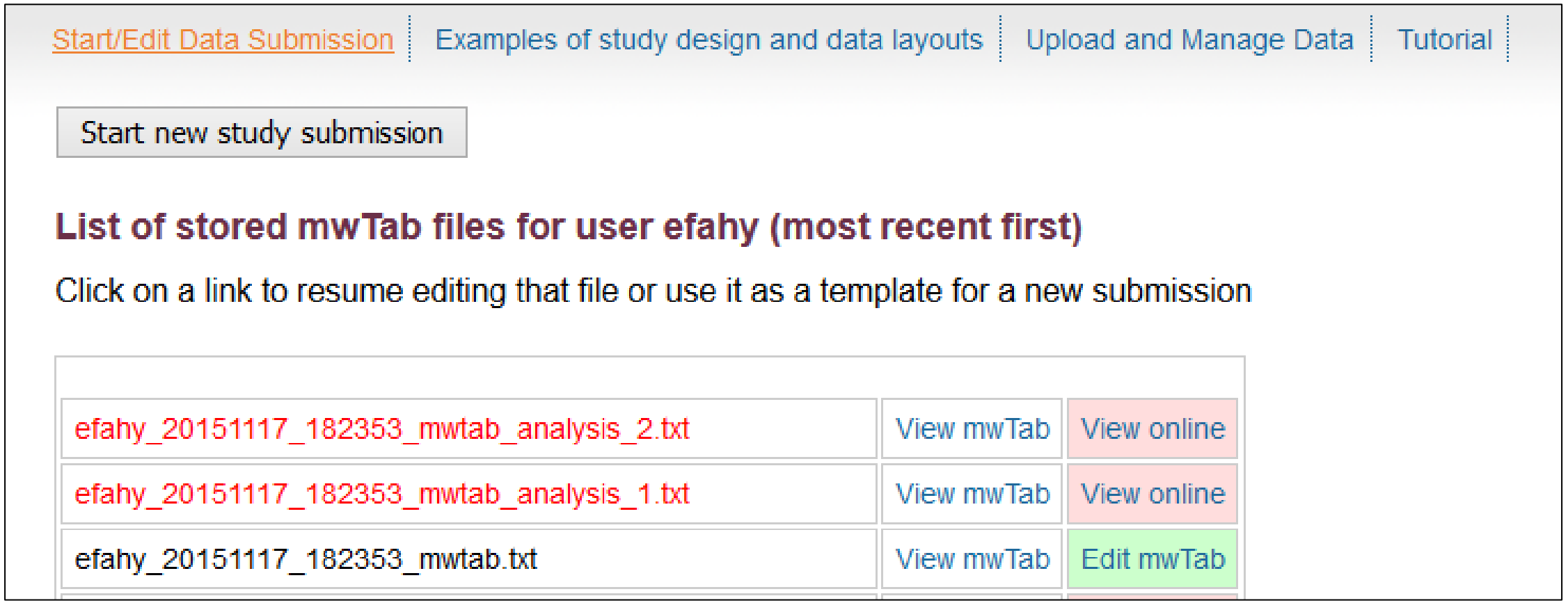
Option1: Measurements for **named** metabolites from **targeted** NMR analyses.

Option2: Measurements from **untargeted** experiments e.g. binned NMR analyses. Detected features are typically binned chemical shift ranges.



Start/Edit Data Submission link

All of a user's submissions are visible on this page

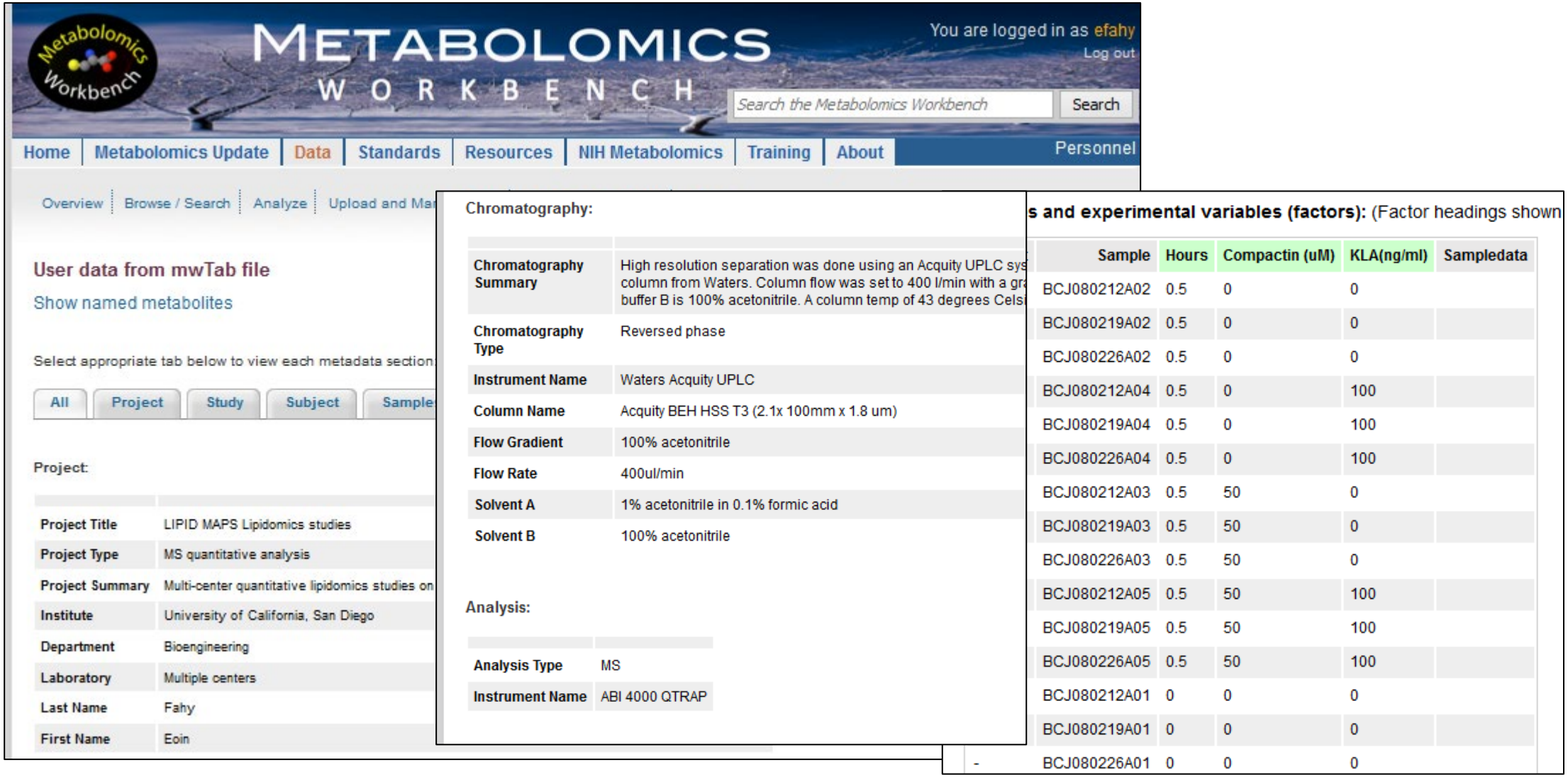


View/download the completed mwTab files

These are saved in the user's login area

```
#METABOLOMICS WORKBENCH efahy_20151117_182353
VERSION                 1
CREATED_ON              November 17, 2015, 6:23 pm
#PROJECT
PR:PROJECT_TITLE                        LIPID MAPS Lipidomics studies
PR:PROJECT_TYPE                         MS quantitative analysis
PR:PROJECT_SUMMARY                      Multi-center quantitative lipidomics s
PR:PROJECT_SUMMARY                      sources (LIPIDMAPS)
PR:INSTITUTE                            University of California, San Diego
PR:DEPARTMENT                           Bioengineering
PR:LABORATORY                           Multiple centers
PR:LAST_NAME                            Fahy
PR:FIRST_NAME                           Eoin
PR:ADDRESS                              9500 Gilman, La Jolla, CA, 92093, USA
PR:EMAIL                                efahy@ucsd.edu
PR:PHONE                                858-534-4076
PR:FUNDING_SOURCE                       NIGMS
#STUDY
ST:STUDY_TITLE                          Timecourse on RAW 264.7 cells treated
ST:STUDY_TYPE                           Timecourse experiment
ST:STUDY_SUMMARY                        Lipidomics studies on macrophages - RA
ST:STUDY_SUMMARY                        and compactin. Experiments were conduc
ST:STUDY_SUMMARY                        serum. 8-timepoint study: Measurements
ST:STUDY_SUMMARY                        24hrs for: (i) compactin, (ii) Kdo2-Li
ST:STUDY_SUMMARY                        (iv) control
ST:INSTITUTE                            University of California, San Diego
ST:DEPARTMENT                           Bioengineering
ST:LABORATORY                           Multiple centers
ST:LAST_NAME                            Fahy
ST:FIRST_NAME                           Eoin
```

The “View Online “ link allows users to view and analyze the study to review the data/metadata. This viewer simulates how the study will appear on the Metabolomics WorkBench after NMDR curation and database upload



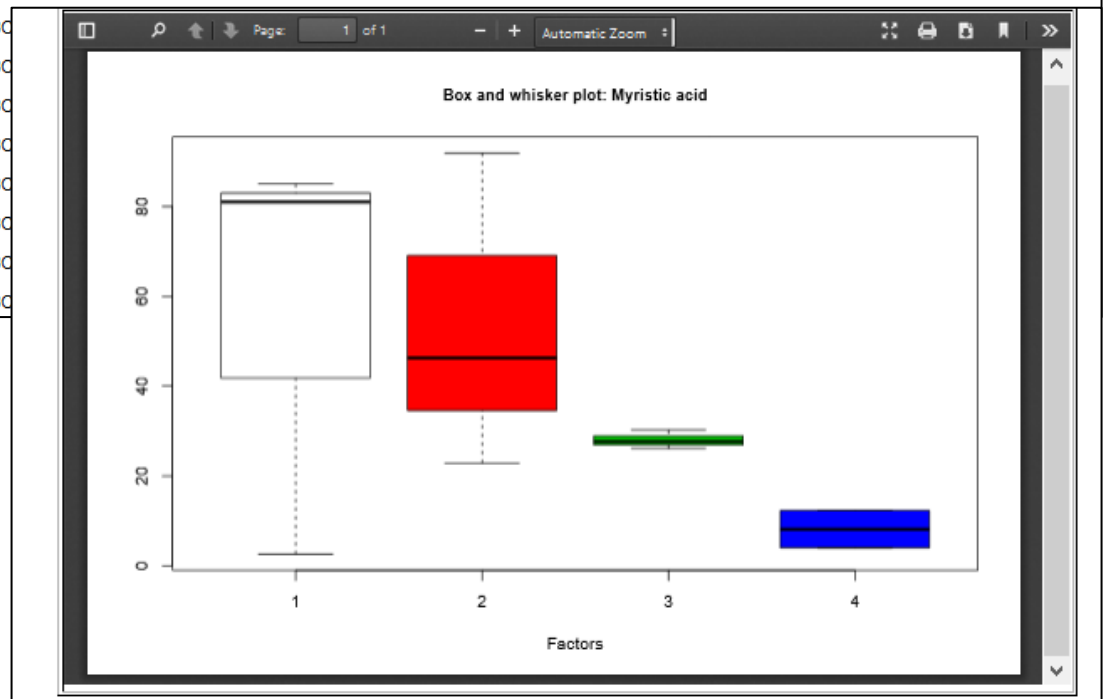
The “View Online “ link allows users to perform analysis on their datasets via the mwTab format prior to NMDR registration and database upload

User data from mwTab file

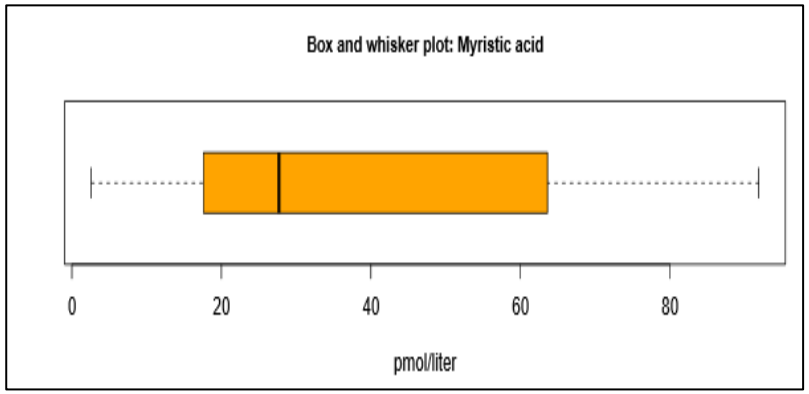
ANOVA analysis | Display t-test grid | Show z-scores | PCA analysis | LDA analysis

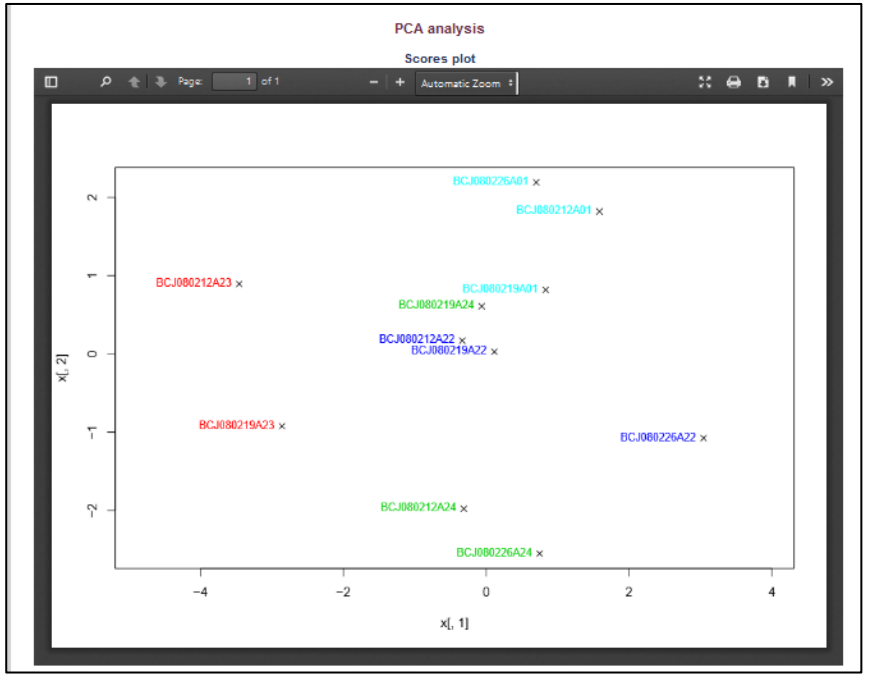
Bar graph by sample	Boxplot	Boxplot	Bar graph by factor level	Select one factor
Bar graph (samples)	All samples	By factor	Bar graph by factor level	Display data for one factor

Sample	Myristic acid	Factors
BCJ080212A01	2.6	Hours:0 | Compactin (uM):0 | KLA(ng/ml):0



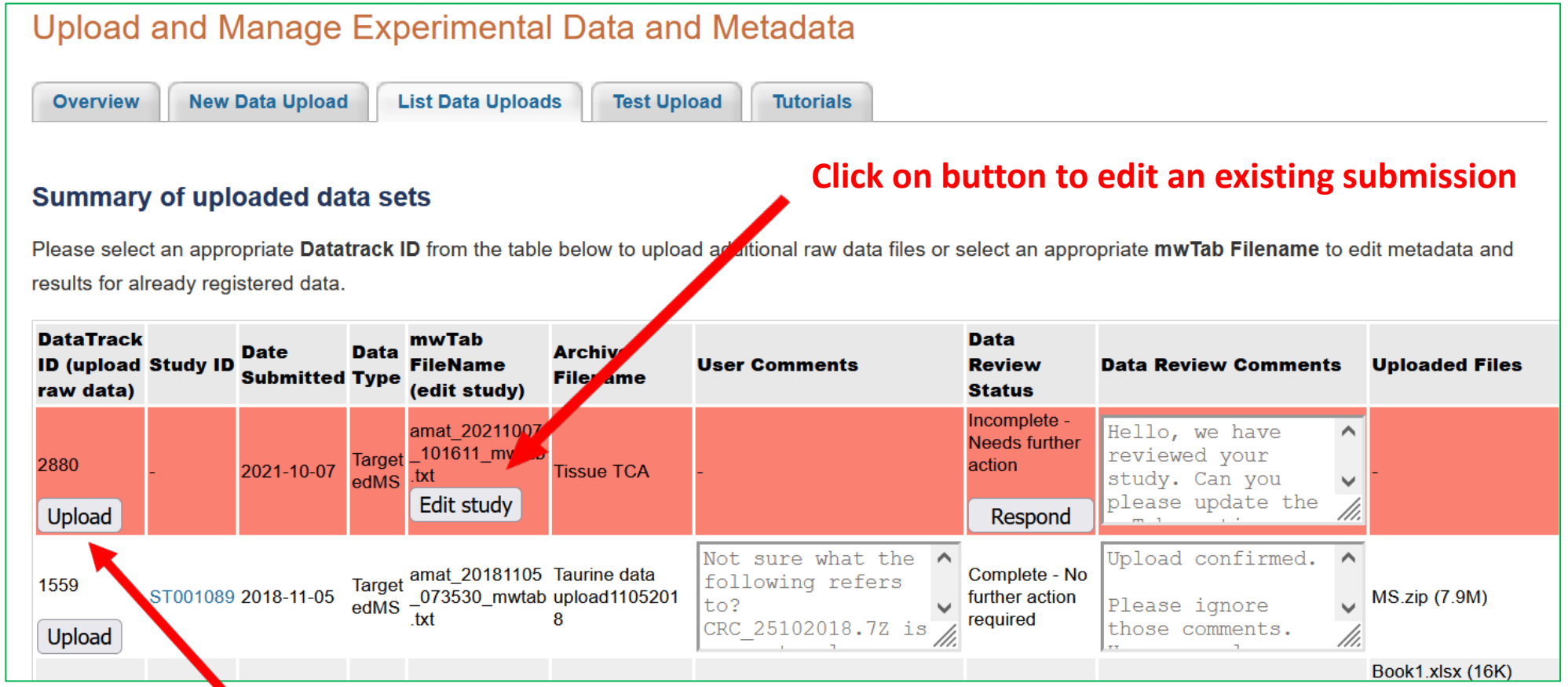
Factor level F1: Hours:0 | Compactin (uM):0 | KLA(ng/ml):0
Factor level F2: Hours:12 | Compactin (uM):0 | KLA(ng/ml):0
Factor level F3: Hours:12 | Compactin (uM):0 | KLA(ng/ml):100
Factor level F4: Hours:12 | Compactin (uM):50 | KLA(ng/ml):0





Edit your Data Submission (DataTrack_ID)

Resume submission or edit an existing submission from the “List Data Uploads” section at https://www.metabolomicsworkbench.org/data/DRCCDataDeposit.php

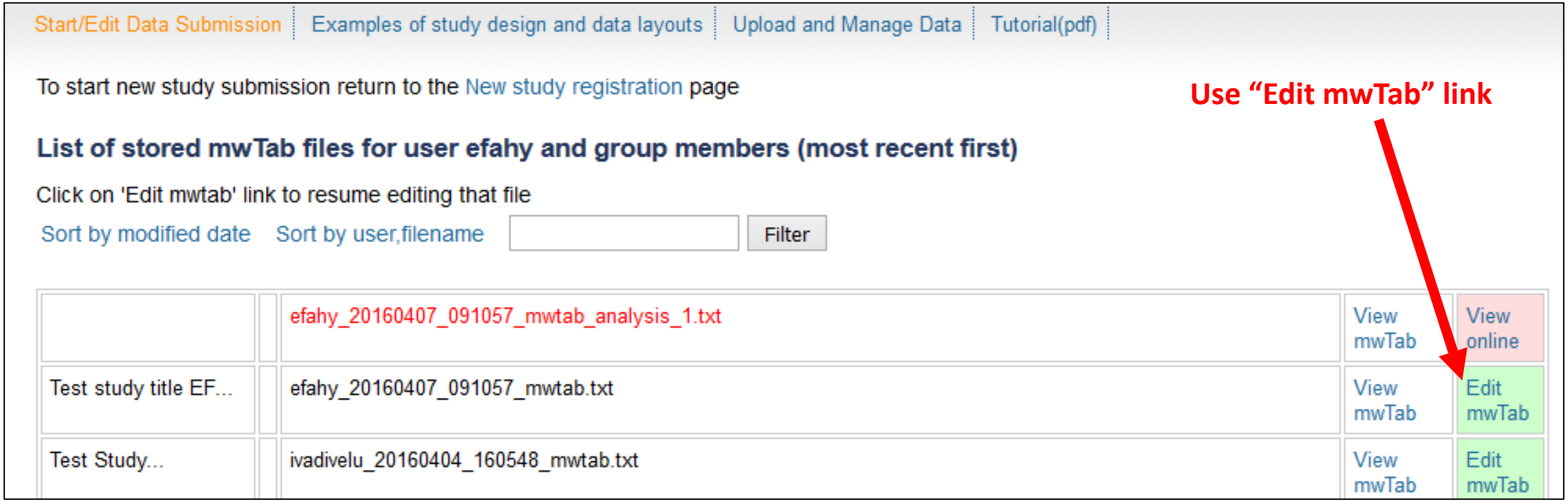


Click on button to edit an existing submission

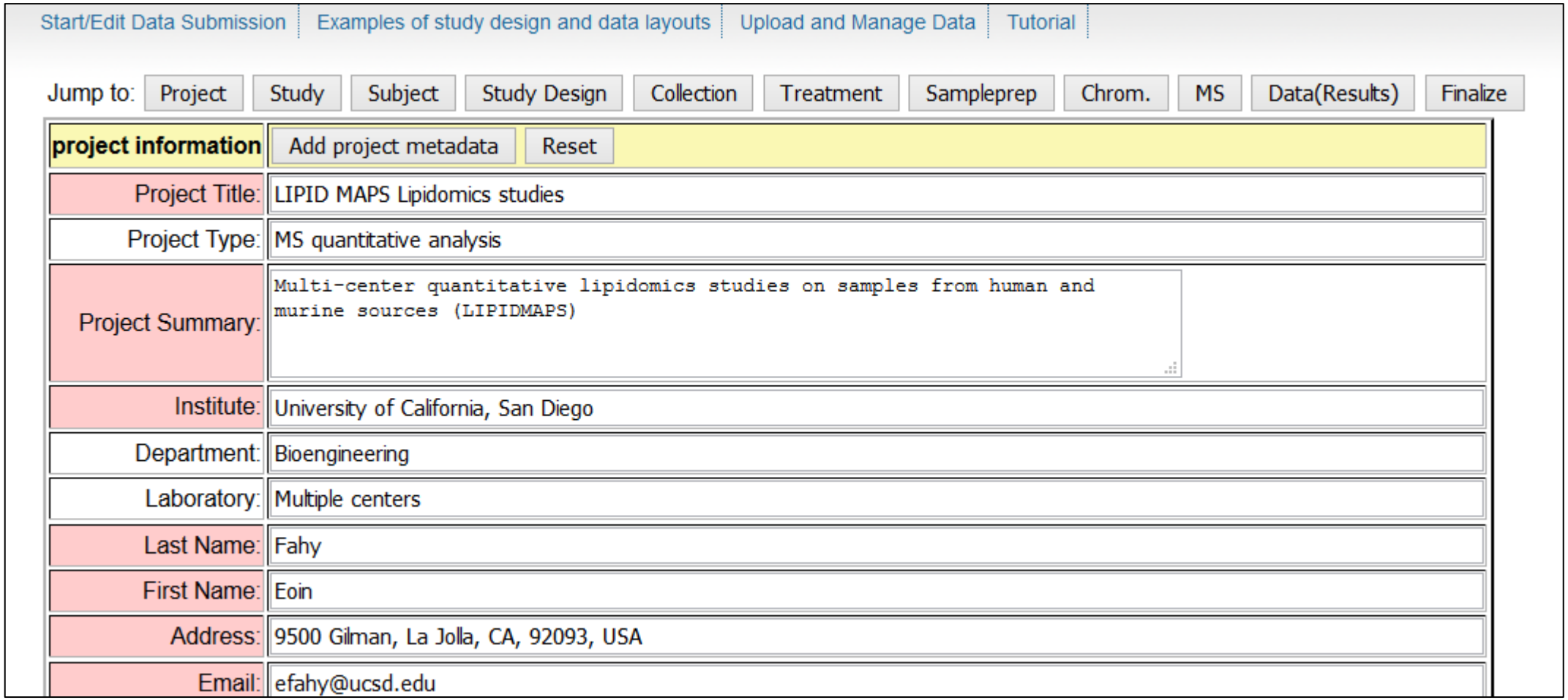
Upload raw data for a submission by clicking on the Upload button

Edit your Data Submission

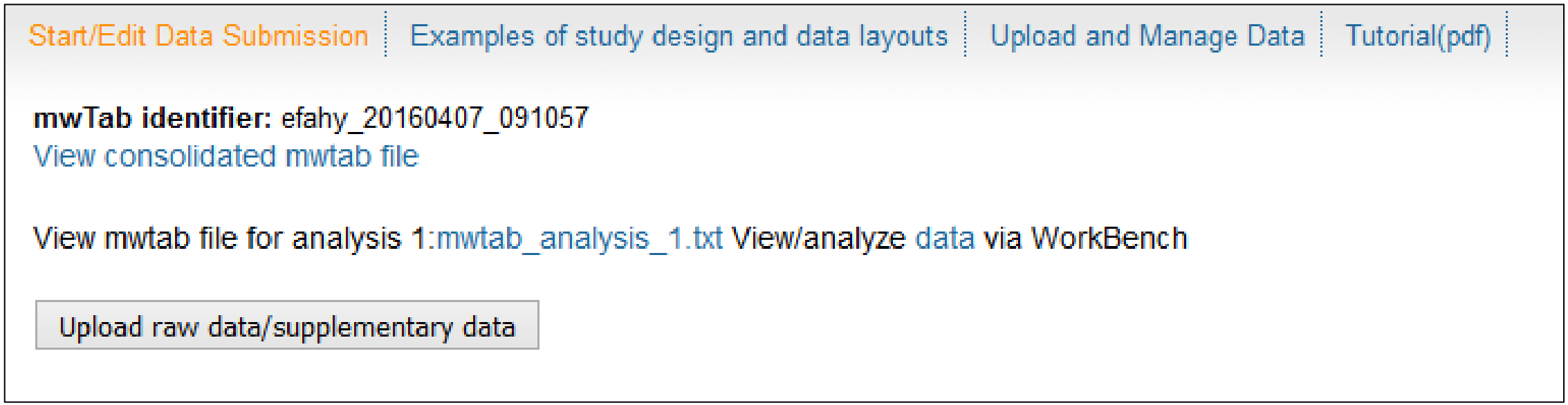
Resume submission of a new study or edit an existing study from the online GUI at http://www.metabolomicsworkbench.org/data/ds_main.php



Study editing interface: Jump to section of interest



After finalizing your online submission, upload your raw and supplementary data



Upload your raw and supplementary data via a standalone FTP client

Your raw data should be submitted as a compressed file (.zip, .7Z, .gz, etc)

IMPORTANT! Please upload raw data in open-source format (e.g. mzML, mzXML, CDF) if at all possible to enable re-use and re-analysis by other researchers

Please do not upload individual raw files- combine them in a single compressed archive (.zip,.7z)

Upload additional files for already registered data with data track ID [redacted]

Please review the following information before you continue to upload data to remote directory [redacted]:

- **Raw data upload**:
  IMPORTANT! Please upload raw data in open-source format (mzML,mzXML,CDF) if at all possible to enable re-use and re-analysis.
  Please do not upload individual raw files- combine them in a single compressed archive (.zip,.7z) first.
  Compressed (zip, 7z) data files are selected and uploaded to the NMDR FTP server through a FTP client. A variety of free and commercial standalone file transfer clients exist to upload large data files directly to the FTP servers: **FileZilla** , **WinSCP** , etc. The usage of **FileZilla** is recommended for uploading data to the NMDR.
- Use the following credentials to upload data to the NMDR:
  - **Server Name:** www.metabolomicsworkbench.org
  - **User Name:** drccupload
  - **Password:** #Vgy7ujmnbv$

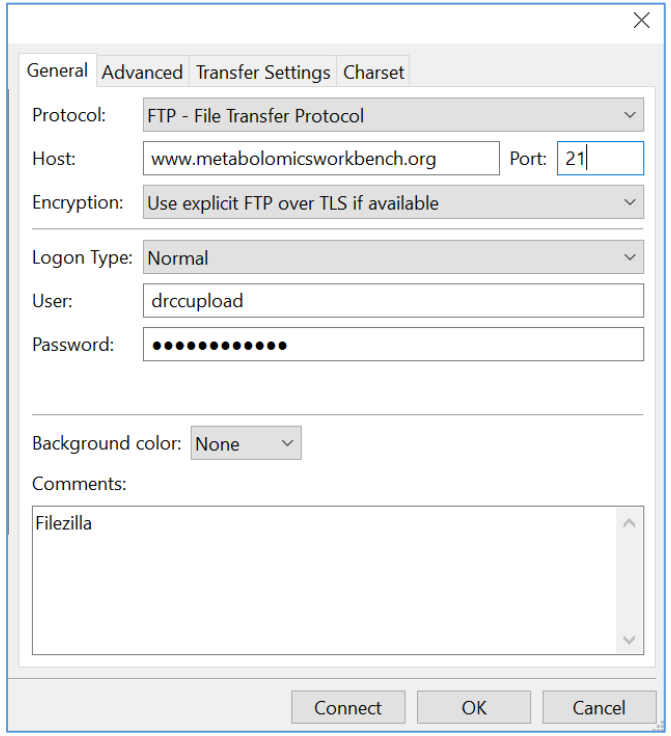
  Password and directory location will be visible when you navigate to the appropriate datatrack_id in your data upload list
- After connecting to the NMDR FTP server, please change remote directory to [redacted] before uploading the compressed (zip, 7z) data file.

Upon completion of registration, your dataset is entered in the NMDR processing queue where it will be curated and uploaded on the Metabolomics WorkBench public website (depending on embargo conditions)

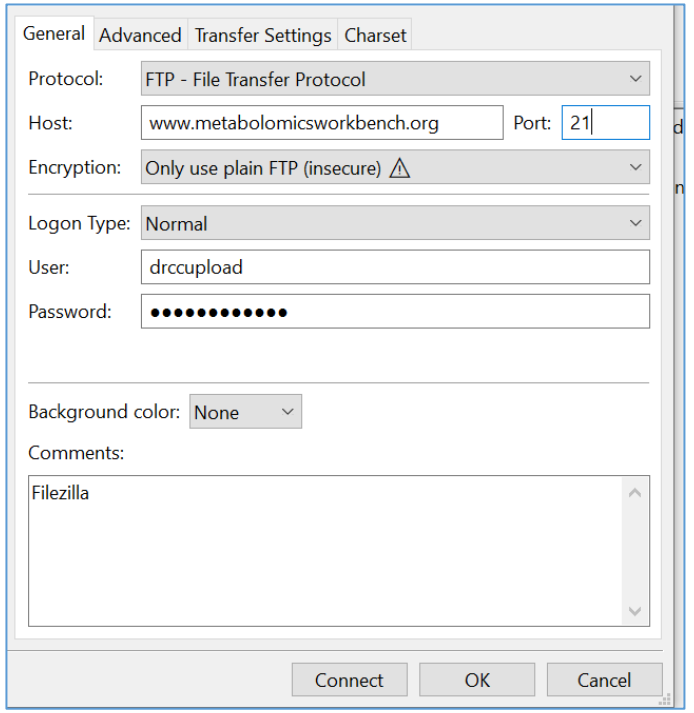
Upload your raw and supplementary data via a standalone FTP client
Fillezilla is the recommended FTP client.
Download the free client at https://filezilla-project.org

Fillezilla settings

(a)



(b)



Note: If your upload fails using the default encryption settings (a), switch to the unencrypted setting (b)
The firewall rules at your institution may not allow FTP over TLS

Then navigate to the remote directory given to you in the instructions before uploading files

Upon completion of registration, your dataset is entered in the NMDR processing queue where it will be curated and uploaded onto the Metabolomics WorkBench website (depending on user-specified embargo conditions)

It typically takes 5-10 working days for a submission to be reviewed and processed. The submitter will then be notified and provided with a DOI and a private link to the study which may be shared with reviewers

Things that slow down submission processing:

The submitter has not provided raw data files (or the files are incomplete/corrupt)
The submitter has not listed which raw files match each sample in the study-design section
The submitter has neglected to provide a required item in one of the metadata fields