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 <u>exactly match</u> 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 Data Submission

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



Register/login

(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	EF45.zip	(e.g. MyData.zip, MyData.7z or MyData.gz)						
* Data type being submitted	MS 📉 (Use sepa	rate submissions for studies containing both MS and NMR data)						
* Protocol methods filename(s)	PR_SP45.pdf							
	PR_TR45.pdi							
* MS/NMR instrument manufacturer	ABI-SCIEX							
* MS/NMR instrument model	4000-QTRAP							
* Binary data format	.wiff	(e.gWIFF (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 t	pelow)						
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

Upload and Manage Experimental Data and Metadata
Overview New Data Upload List Data Uploads Test Upload Tutorials
Please click New online study submission button to start a new study submission and enter metadata and results for your study with DataTrack ID 561 and mwTab file name efahy_20160407_093705. You will be prompted to upload an archive file after successful completion of the online submission process.
New online study submission

(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 pulldown menu, choose "Add new item" from the list and type in your own value

ump to: Project S	tudy Subject Study Design Collection Treatment Sampleprep Chrom. MS Data(Results) Finalize
project information	Add project metadata Reset
Project Title	LIPID MAPS Lipidomics studies
Project Type	MS quantitative analysis
Project Summary	Multi-center quantitative lipidomics studies on samples from human and murine sources (LIPIDMAPS) The National Institute of Diabetes and Digestive and Kidney Diseases (NIDDK) in collaboration with the National Institute of Standards (NIST) recently produced a human plasma standard reference material (SRM 1950) for metabolite analysis. The SRM was prepared by
Institute	University of California, San Diego 🗸
Department	Bioengineering
Laboratory	
Last Name	Fahy
First Name	Eoin
Address	9500 Gilman ,La Jolla, CA 92093
Email	efahy@ucsd.edu
Phone	(111)-222-3333
Funding Source	
Project Comments	
Publications	Quehenberger, O. et al Lipidomics reveals a remarkable diversity of lipids in human plasma. J Lipid Res. !
contributors	
project information	Add project metadata Reset

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 <u>very important</u> 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)

Jump to: Project Stu	dy Subject Study Design Collection Treatment Sampleprep Chrom. MS Data(Results) Finalize	
study information	Add study metadata Reset	
Subject Type:	Human	ions
Study Title:	<u>Lipidomics</u> studies on <u>NIDDK</u> / <u>NIST</u> human plasma samples	
Study Summary:	The National Institute of Diabetes and Digestive and Kidney Diseases (<u>NIDDK</u>) in collaboration with the National Institute of Standards (<u>NIST</u>) recently produced a human plasma standard reference material (<u>SRM</u> 1950) for metabolite analysis. The <u>SRM</u> was prepared by obtaining plasma samples from 100 individuals between 40 and 50 years of age, whose ethnicity was	
Institute:	University of California, San Diego 🗸 🗸	
Department:	Bioengineering	
Laboratory:		
Last Name:	Fahy	
First Name:	Eoin	
Address:	9500 Gilman ,La Jolla, CA 92093	
Email:	efahy@ucsd.edu	
Number of Groups:		
Total Subjects:		
Number of Males:		
Number of Females:		
Study Comments:		
Publications:		
chear_study:		
analysis_type_detail:		
Study Type:		
Disease	(111) 222 2222	

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 "/"

mwTab Identifier: msud_20180206_090350 Return to start							
Jump to: Project Study	Subject Study Design Collection Treatment Sampleprep Chrom. MS Data(Results)						
subject information	Add subject metadata Reset						
Subject Type:	Cultured cells (entered in Study page)						
Subject Species:	Mus musculus v or (new):						
Taxonomy ID:	10090						
Genotype Strain:							
Age or Age Range:							
Weight or Weight Range:							
Height or Height Range:							
Gender:	Not applicable V						
Cell Biosource or Supplier:							
Cell Strain Details:							
Subject Comments:							
Cell Primary Immortalized:							
Cell Passage Number:							
Cell Counts:							

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)

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) **Copy/paste as tab-delimited data from Excel or text** Input Study Design information. Sample names, experimental factor(s), sample source and raw data file Subject name and additional sample data are optional. file (View the "See examples.." link for more help) First row must contain headings. Data must be tab-delimited. Subject_ID Sample_ID RAW FILE NAME Sample source Genotype Treatment Batch Wild-type CA11 CA11W0 Liver Control B1a SC CA11W0.mzML CA12 CA12W0 Liver Wild-type Control B1a SC CA12W0.mzML Wild-type CA13 CA13W0 Liver Control B1a SC CA13W0.mzML CA11 CA11W50 Liver Wild-type 50uM B1a SC CA11W50.mzML CA12 CA12W50 Liver Wild-type 50uM SC CA12W50.mzML B1a CA13 CA13W50 Liver Wild-type 50uM SC_CA13W50.mzML B1a SC CA14W0.mzML CA14 CA14W0 Liver Mutant Control B1a CA15 Liver Control CA15W0 Mutant B1a SC_CA15W0.mzML CA16 CA16W0 Control SC_CA16W0.mzML Liver Mutant B1a CA17 SC CA17W50.mzM CA17W50 Liver Mutant 50uM B1a Then click on "View/check study design" View/check study design See examples of study design layout to view in tabular format

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
1	CA12	CA12W50	Liver	Wild-type	50uM	B1a	SC_CA12W50.mzML
٦	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
1	CA18	CA18W50	Liver	Mutant	50uM	B1a	SC_CA18W50.mzML
	CA19	CA19W50	Liver	Mutant	50uM	B1a	SC_CA19W50.mzML

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 <u>unique to each sample</u>. 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.

CA12

Process study design data Does submission contain raw data?:															
~		~		~			~			~		~			
Subject_ID	Sample_ID	Sample source	Genotype		Treat	me	nt	Bato	h		RAW_FILE_N	AME			
CA11	CA11W0	Liver	Wild-type		Contr	ol		B1a			SC_CA11W0.	mzML]		
CA12	CA12W0	Liver	Wild-type		Contr	ol		B1a		Assign	each colu	mn a	s "Sub	iect ID)″
CA13	CA13W0	Liver	Wild-type		Contr	ol		B1a		"Sam		actor ⁴	" "Rav	y filo	
CA11	CA11W50	Liver	Wild-type		50uM			B1a		sample_id, factor, naw me					
CA12	CA12W50	Liver	Wild-type		50uM			B1a		name	, Other	or ig	nore		
0142	044014/50	Liver	Process study design	n data 🛛	Does su	ıbmi	ssion conta	in raw	/ data?	Yes 🗸					
			Subject_ID v Subject_ID	Samp Samp	le_ID l e_ID	~	Sample sour	rce 🗸	Factor Genot	v) ype	Factor V Treatment	Other Batch	~	Raw file na	me ∽ _NAME
			CA11	CA11V	VO		Liver		Wild-ty	ре	Control	B1a		SC_CA11V	V0.mzML
After editing/assignment, proceed by			CA12	CA12	NO		Liver		Wild-ty	ре	Control	B1a		SC_CA12V	V0.mzML
elieking on "Drocoss study design date"			CA13	CA13	N0		Liver		Wild-ty	ре	Control	B1a		SC_CA13V	V0.mzML
	ocess study de	sign udld	CA11	CA11V	V50		Liver		Wild-ty	ре	50uM	B1a		SC CA11V	V50.mzML

Liver

CA12W50

B1a

SC CA12W50.mzML

50uM

Wild-type

Collection information

Metadata related to sample collection/ handling/storage

	mwTab identifier: efahy_20151117	7_182353
	collection information	Add collection metadata Reset
	Collection Summary:	Cells were counted, washed with cold PBS and then flash-frozen in liquid N2
1	Collection Protocol ID:	
	Collection Protocol Filename:	kdo_col_428.txt
	Collection Protocol Comments:	
	Sample Source/Type:	Tissues
	Collection Method:	
	Collection Location:	

Tissue	Cell Identification:		
Tissue Ce	II Quantity Taken:		
Upload Colle	ction Protocol File(s)	Browse	No file selected.

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

Treatment information

Metadata related to treatment protocols

mwTab identifier: efahy_20151112	7_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	< > II
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

ļ	mwTab identifier: efahy_20151117	7_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		

Decision point: MS or NMR experiment?

Select analysis type: MS				
Select				

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

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

Number of chromatography methods used for which you have data: 1				
Add Chromatography metadata				
Example: If only GCMS or RP-LCMS was used, select "1" (default) If both HILIC and RP chromatography were used, select "2"				

Chromatography information

Metadata related to chromatography (LC/GC) protocols

Number of chromatography methods used for which you have data: 1

Add Chromatography metadata

Example: If only GCMS or RP-LCMS was used, select "1" (default) If both HILIC and RP chromatography were used, select "2"

Add Chromatography metadata Reset				
Chromatography Fields Chromatography method 1				
<				
Chromatography Summary:				
Chromatography Type:	HILIC V			
Instrument Name:	Thermo Scientific Transcend Duo LX-2 UHPLC \checkmark			
Column Name:	Thermo Accucore 150 Amide (50 x 2.1mm, 2.6um)			
Solvent A (LC-MS):	95% acetonitrile/5% water; 0.1% acetic acid;			
Solvent B (LC-MS):	50% acetonitrile/50% water; 0.1% acetic acic			
Flow Gradient (LC-MS):	0.55 ml/min: 0-0.1 min: 0% B, 0.10-5.0 min:			
Flow Rate:	0.55 ml/min			
Column Temperature(°C):	45			
Methods Filename:				
Methods ID:				
Column Pressure:				
<				
Upload Chromatography Methods File(s) Browse No file selected.				

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)

mwTab identifier: efahy_20151117_182353		
Number of MS conditions per chromatography method used for which you	have data: 2 👻	
Add MS metadata		
Example: If you have GCMS data in positive ion mode only, select "1" (default) If you have both positive and negative ion mode LCMS data, select "2"		

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

mwTab identifier: ef	mwTab identifier: efahy_20151117_182353						
Number of MS conditions per chromatography method used for which you have data: 2 v							
	Add MS metadata						
Example: If you h If you have both p	ave GCMS data in positive ion m positive and negative ion mode L	node only, select "1" (defa CMS data, select "2"	ult)				
Replicate 1st column values to all ot	her columns						
Add MS metadata Reset							
MS analysis fields CI	h.:HILIC (1)	Ch.:GC	: (2)				
<							
Instrument Name:	Agilent 6520 QTOF	~	Agilent 5975C	~			
Instrument Type:	Q-TOF ~		Linear quadrupole 🗸				
MS Type:	ESI ~		EI ~				
Ion Mode:	NEGATIVE ~		POSITIVE ~				
MS acquisition Comments: Data processing Comments: Software/procedures used for feature assignments:	-	.::	-				
Laboratory Name:	MRC2 (University of Michiga		MRC2 (University of Michiga				
Operator Name:							
Detector Type:							

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 containin	ng identified (named) r	netabolites	
Input Data in tab-delimited format in the te First column must contain metabolite names. Subsequent columns must contain sample da First row must contain sample names.	xt area below. ta with identical sample na	mes as in Study De	sign submission.
Tabular results (typically are pasted into this text	r tens or hundre area	ds of name	d metabolites)
Units of measurement (required): xxx View/check metabolite data See examples of	of metabolite data layout	Delete existing meta	 abolite data (this analysis only)
OPTION 2: Untargeted assays <u>not</u> co	ntaining identified (na	ned) metabolite	 A file of tabular results (typically thousands of
(e.g. datasets with m/z, retention time features Add/replace results as a tab-delimited text Ideally, feature names should be formatted	from LC-MS experiments, file: as 'm/z underscore reten	tion time', e.g. 645	uploaded here
This will enable analysis of the dataset by a	a larger number of tools o	on the Metabolomic	cs Workbench
Units of measurement (required):	(required):	m/z values"?	(required): Time units:
* By "m/z values" we are referring to mass	-to-charge ratios and NO	r neutral masses.	
Upload tab-delimited datafile Browse No	file selected.		

Option1: Targeted data

Units is a required

field

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

mwTab identifier: efahy_20151117_182353	Metabolite BC/080212A01 BC/080219A01 BC/080226A01 BC/080222A22 BC/080219A01 Margaric acid 3.5 25.6 85.5 43.9 Myristic acid 2.6 85.0 81.0 22.9
Dataset 1 of 2:Add Data for Reversed phase POSITIVE mode	Oleic acid 20.2 28.5 1.0 11.1 Palmitic acid 16.0 7.2 10.6 70.1 Pentadecanoic acid 57.6 36.2 20.9 10.7 Penp 8.3 66.1 4.7 75.6
	PGE2 93.0 75.5 70.9 87.2 PGF2a 28.8 30.4 30.7 37.9 PGF2a 28.3 76.8 48.5 70.5
	stearic acid 29.3 57.1 16.3 78.8 Stearidonic acid 52.8 43.0 49.0 90.4 Tricosanoic acid 4.8 36.1 27.5 24.9
mutah idantifian ofaby 20151117 182252	/
Input Data in tab-delimited format. First column must contain metabolite names. Subsequent columns must contain sample data with identical sample names as in St First row must contain sample names.	tudy Design submission.
Input Data in tab-delimited format. First column must contain metabolite names. Subsequent columns must contain sample data with identical sample names as in Si First row must contain sample names. Copy/paste tabular data here	tudy Design submission.
Input Data in tab-delimited format. First column must contain metabolite names. Subsequent columns must contain sample data with identical sample names as in Si First row must contain sample names. Copy/paste tabular data here First row MUST contain sample names identical to those	tudy Design submission.
Input Data in tab-delimited format. First column must contain metabolite names. Subsequent columns must contain sample data with identical sample names as in Si First row must contain sample names. Copy/paste tabular data here First row MUST contain sample names identical to those in the "study design" step. First column must contain me	tudy Design submission.
Input Data in tab-delimited format. First column must contain metabolite names. Subsequent columns must contain sample data with identical sample names as in Si First row must contain sample names. Copy/paste tabular data here First row MUST contain sample names identical to those in the "study design" step. First column must contain me names or m/z-retention time identifiers in the case of ur	e submitted etabolite nidentified
Input Data in tab-delimited format. First column must contain metabolite names. Subsequent columns must contain sample data with identical sample names as in Si First row must contain sample names. Copy/paste tabular data here First row MUST contain sample names identical to those in the "study design" step. First column must contain me names or m/z-retention time identifiers in the case of ur ions (e.g. "231.4185, 17.68")	e submitted etabolite nidentified
Input Data in tab-delimited format. First column must contain metabolite names. Subsequent columns must contain sample data with identical sample names as in Si First row must contain sample names. Copy/paste tabular data here First row MUST contain sample names identical to those in the "study design" step. First column must contain me names or m/z-retention time identifiers in the case of ur ions (e.g. "231.4185_17.68")	tudy Design submission.

Option1: Processed data upload: Review in tabular form, then Upload data Targeted 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	./ 3/.	9 61.3	10.2 /0.3	/8.2 23.6	81./ 5/.	3	
PGJ2 32.3	76.8 48	.5 70.5	5 1.9	29.7 92.1	94.5 75.1	92.0 82.	1	
stearic acid	29.3 57	.1 16.3	3 78.8	67.5 14.6	85.3 94.4	63.9 16.	3 12.5	
Stearidonic aci	1 92	.8 49.0	0 49.0	90.4 72.6	21.9 54.1	6.3 26.	6 97.9 77	.4 🗸
Tricosanoic aci	d 4.	8 36.1	1 27.5	24.9 67.1	0.4 50.8	23.0 13.	3 82.4 85	i.3 🧮
Units of measurer	n <mark>ent:</mark> pmoles/	/						
View/check metal	oolite data	See exam	ples of meta	abolite data layou	t			
metabolite_nam	e BCJ08021	IZA01 BCJ	J080219A01	BCJ080226A01	BCJ080212A22	BCJ080219A22	BCJ080226A22	BCJ0802
Margaric acid	3.5	25.6	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	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								



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.

mwTab identifier: efahy_20151117_182353						
Metabolite metadata in tab-delimited format. First column must contain metabolite names.						
Subsequent columns should contain KEGG, PubChem identifiers, retention index, quantitated m	/z, e	tc.				
First row must contain headings.						
Metabolite Name Pubchem Id Kegg Id						
Margaric acid 10465 -						
Myristic acid 11005 C06424		Stearidonic acid	5282	837 C16300		
Oleic acid 445639 C00712		Tricosanoic acid	1708	5 -		
Palmitic acid 985 C00249						
Pentadecanoic acid 13849 C16537						
PGD2 448457 C00696		View/check metabo	lite data S	See examples of metabolite dat	a lavout	
PGE2 5283116 C00584		,				
PGF2a 5280363 -		Upload metabolite r	netadata			
PGJ2 5311211 C05957		metabolite_name	Pubchem Ic	d Kegg Id		
stearic acid 5281 C01530		Margaric acid	10465	-		
		Myristic acid	11005	C06424		
View/check metabolite metadata See examples of metabolite metadata layout		Oleic acid	445639	C00712		
view check metabolice metabolice metabolice metabolite metabolite metabolite		Palmitic acid	985	C00249		
		Pentadecanoic acio	13849	C16537		
After checking the table of metabolite constations, alight		PGD2	448457	C00696		
After checking the table of metabolite annotations, click /		PGE2	5283116	C00584		
"Upload metabolite metadata"		PGF2a	5280363			
		PG.I2	5311211	C05957		



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



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*?	Feature names contain retention time values?		
Peak area	(required): Yes	(required): Yes V Time units: Minutes V		
* 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 conta	in sample names exactly matching those that	t you submitted in the 'Study Design' section.		

<i>(</i>						
🚺 C:\	Users\eoinf\Downloads\untargeted_da	ta_table.txt				
1	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	1912 7	4983 4	1102 9	556

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

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 v		
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



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_ 1	182353
CREATED ON	November 17, 20	15. 6:23 pm
#PROJECT		10, 0120 pm
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

Netabolomic Workbench		BOLC KBEN	You are logg MICS C H Search the Metabolomics Workbench	ed in as efahy Log out Search				
Home Metabo	omics Update Data Standards	Resources NI	H Metabolomics Training About	Personnel				
Overview Brow	se / Search Analyze Upload and Mar	Chromatography:		s and experime	ental v	ariables (factor	rs): (Factor	headings shown
llear data fror	n mwTah filo	Chromatography	High resolution separation was done using an Acquity UPLC sys	Sample	Hours	Compactin (uM)	KLA(ng/ml)	Sampledata
Show named m	etabolites	Summary	column from Waters. Column flow was set to 400 l/min with a grabuffer B is 100% acetonitrile. A column temp of 43 degrees Celsion	BCJ080212A02	0.5	0	0	
onow named m	etabolites	Chromatography	Reversed phase	BCJ080219A02	0.5	0	0	
Select appropriate	tab below to view each metadata section.	Туре		BCJ080226A02	0.5	0	0	
		Instrument Name	Waters Acquity UPLC	BCJ080212A04	0.5	0	100	
All Projec	t study subject sample	Column Name	Acquity BEH HSS T3 (2.1x 100mm x 1.8 um)	BCJ080219A04	0.5	0	100	
		Flow Gradient	100% acetonitrile	BCJ080226A04	0.5	0	100	
Project:		Flow Rate	400ul/min	BC-1080212A03	0.5	50	0	
Project Title	LIPID MAPS Linidomics studies	Solvent A	1% acetonitrile in 0.1% formic acid	BC 1080219403	0.5	50	0	
Project Type	MS quantitative analysis	Solvent B	100% acetonitrile	BC3080219A03	0.5	50	0	
Project Summary	Multi-center quantitative lipidomics studies on			BCJ080220A03	0.5	50	400	
Institute	University of California, San Diego	Analysis:		BCJ080212A05	0.5	50	100	
Department	Bioengineering			BCJ080219A05	0.5	50	100	
Laboratory	Multiple centers	Analysis Type	MS	BCJ080226A05	0.5	50	100	
Last Name	Fahy	Instrument Name	Instrument Name ABI 4000 QTRAP		0	0	0	
First Name	Eoin			BCJ080219A01	0	0	0	
and an and a second	And the second se			BCJ080226A01	0	0	0	

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







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

Uploa	Upload and Manage Experimental Data and Metadata									
Overvie	Overview New Data Upload List Data Uploads Test Upload Tutorials									
Summa Please se results for	Summary of uploaded data sets Please select an appropriate Datatrack ID from the table below to upload adaitional raw data files or select an appropriate mwTab Filename to edit metadata and results for already registered data.									
DataTra ID (uplo: raw data	ck ad Study ID a)	Date Submitted	Data Type	mwTab FileName (edit study)	Archiv Filer ame	User Comments	Data Review Status	Data Review Comments	Uploaded Files	
2880 Upload		2021-10-07	Target edMS	amat_20211007 _101611_my .txt Edit study	Tissue TCA	-	Incomplete - Needs further action Respond	Hello, we have reviewed your study. Can you please update the	• //	
1559 Upload	ST001089	2018-11-05	Target edMS	amat_20181105 _073530_mwtab .txt	Taurine data upload1105201 8	Not sure what the following refers to? CRC_25102018.7Z is	Complete - No further action required	Upload confirmed. Please ignore those comments.	MS.zip (7.9M)	
									Book1.xlsx (16K)	

• 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

Start/Edit Data Submiss	on Examples of study design and data layouts Upload and Manage Data Tutorial(pdf)					
To start new study submission return to the New study registration page Use "Edit n						
List of stored mw	Tab files for user efahy and group members (most recent first)					
Click on 'Edit mwtab' lii	nk to resume editing that file					
Sort by modified date	Sort by user, filename Filter					
	efahy_20160407_091057_mwtab_analysis_1.txt	View mwTab	View online			
Test study title EF	efahy_20160407_091057_mwtab.txt	View mwTab	Edit mwTab			
Test Study	ivadivelu_20160404_160548_mwtab.txt	View mwTab	Edit mwTab			

Study editing interface: Jump to section of interest

Start/Edit Data Submission Examples of study design and data layouts Upload and Manage Data Tutorial					
Jump to: Project	Study Subject Study Design Collection Treatment Sampleprep Chrom. MS Data(Results) Finalize				
project informatio	n Add project metadata Reset				
Project Title	e: LIPID MAPS Lipidomics studies				
Project Type	e: MS quantitative analysis				
Project Summar	ulti-center quantitative lipidomics studies on samples from human and urine sources (LIPIDMAPS)				
Institute	Jniversity of California, San Diego				
Departmen	Bioengineering				
Laborator	y: Multiple centers				
Last Name	Fahy				
First Name	Eoin				
Address	s: 9500 Gilman, La Jolla, CA, 92093, USA				
Ema	il: efahy@ucsd.edu				

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

Start/Edit Data Submission Examples of study design and data layouts Upload and Manage Data Tutorial(pdf)
mwTab identifier: efahy_20160407_091057 View consolidated mwtab file
View mwtab file for analysis 1:mwtab_analysis_1.txt View/analyze data via WorkBench
Upload raw data/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)



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

 \times

Port: 21

OK

Cancel

Connect

	(a)
General Adva	nced Transfer Settings Charset
Protocol:	FTP - File Transfer Protocol
Host:	www.metabolomicsworkbench.org
Encryption:	Use explicit FTP over TLS if available
Logon Type:	Normal
User:	drccupload
Password:	•••••
Background d	color: None 🗸

General A Protocol

Comments: Filezilla

Fillezilla settings

(b)

General Adv	anced Transfer Settings Charset	
Protocol:	FTP - File Transfer Protocol ~	
Host:	www.metabolomicsworkbench.org Port: 21	
Encryption:	Only use plain FTP (insecure) \triangle	
Logon Type:	Normal ~	
User:	drccupload	
Password:	•••••	
Background	color: None ~	
Comments:		
Filezilla	^	
	×	
	Connect OK Cancel	

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