



MetaboAnalyst 5.0

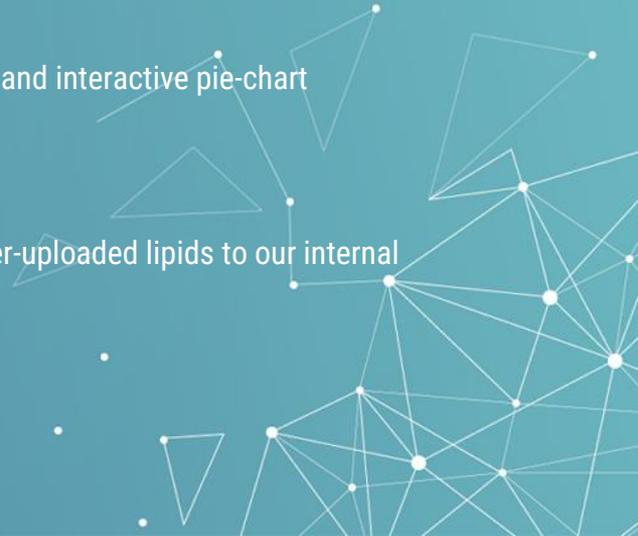
A Web-based Tool for Streamlined
Metabolomics Data Analysis

2022.07.12

5. Enrichment Analysis

Metabolite Enrichment Analysis is used to directly investigate if a group of functionally related metabolites are significantly enriched, eliminating the need to preselect compounds based on some arbitrary cut-off threshold. The Enrichment Analysis module of MetaboAnalyst has added multiple enhancements for Version 5.

Highlights:

- Added 1615 new chemical class metabolite sets
 - Enhanced enrichment analysis visualizations ~ interactive bar-chart, bubble plot, and interactive pie-chart
 - Enhanced support for lipidomics data
 - Expanded our underlying database to cover 197 854 lipids
 - Implemented a smart-matching algorithm to enhance the matching of user-uploaded lipids to our internal compound database.
- 
- A decorative network diagram in the bottom right corner, consisting of white dots connected by thin white lines, forming a complex web of connections against the teal background.

5.1 Start Enrichment Analysis

MetaboAnalyst 5.0 - user-friendly, end-to-end metabolomics data analysis

Module Overview

Input Data Type

Click here to start

on a module to proceed, or scroll down for more details)

Raw Spectra (mzML, mzXML or mzData)			LC-MS Spectral Processing			
MS Peaks (peak list or intensity table)			Functional Analysis	Functional Meta-analysis		
Annotated Features (compound list or table)		Enrichment Analysis	Pathway Analysis	Joint-Pathway Analysis	Network Analysis	
Generic Format (.csv or .txt table files)	Statistical Analysis	Biomarker Analysis	Time-series/Two-factor Analysis	Statistical Meta-analysis	Power Analysis	Other Utilities

Show R command history

5.2.1 Enrichment Analysis (Upload one column compound list)

MetaboAnalyst 5.0 - user-friendly, end-to-end metabolomics data analysis

Users can upload either a list of compounds, a list of compounds with concentration values, or a concentration table.

Choose one of the following options to proceed

Over Representation Analysis | Single Sample Profiling | **Quantitative Enrichment Analysis**

Please enter a one-column compound list:

Creatine
Dimethylglycine
Fumaric acid
Glycine
Homocysteine
L-Cysteine
L-Isolucine
L-Phenylalanine
L-Serine
L-Threonine
L-Tyrosine
L-Valine
Phenylpyruvic acid
Propionic acid
Pyruvic acid
Sarcosine

Download
Exit

Specify the input type.

Input Type: Compound names
Feature Type: -- Specify --
Try Example: None List 1 (metabolites) List 2 (lipids)

Specify here whether the features are lipids or metabolites.

Click **Submit** to upload your data.

Submit

Check the mapping results and then Click **Submit** to process your data.

MetaboAnalyst 5.0 - user-friendly, end-to-end metabolomics data analysis

Name/ID Standardization:

- For enrichment analysis, only well-annotated HMDB compounds (i.e. those in our pathway libraries & metabolite sets) will be mapped. For general purpose name mapping, all Compound IDs (ChEBI, KEGG, PubChem, etc.) are supported.
- Green alphabetics are not recognized; they should be replaced by English names (i.e. alpha, beta).
- Query names in normal alpha-numeric case match - matched by "1" in the downloaded file.
- Query names highlighted include no exact or unique match - matched by "0" in the downloaded file.
- For compound name, you should click the View link to perform approximate search and manually select the correct match if found.
- For KEGG ID, it is possible to have multiple hits, you should click the View link to manually select the correct match if found.

Query	Hit	HMDB	PubChem	KEGG	Details
Aspartic acid	Aspartic acid	156000000	58	C00184	
beta-Alanine	Beta-Alanine	156000000	239	C00089	
Creatine	Creatine	156000000	366	C00266	
Dimethylglycine	Dimethylglycine	156000000	623	C00269	
Fumaric acid	Fumaric acid	156000000	68623	C00122	
Glycine	Glycine	156000000	338	C00007	
Homocysteine	Homocysteine	156000000	238	C00268	
L-Cysteine	L-Cysteine	156000000	3662	C00007	
L-Phenylalanine	L-Phenylalanine	156000000	616	C00073	View
L-Serine	L-Serine	156000000	5851	C00005	
L-Threonine	L-Threonine	156000000	602	C00002	
L-Tyrosine	L-Tyrosine	156000000	5399	C00188	
L-Valine	L-Valine	156000000	6107	C00180	
Phenylpyruvic acid	Phenylpyruvic acid	156000000	607	C00186	
Propionic acid	Propionic acid	156000000	3832	C00163	
Pyruvic acid	Pyruvic acid	156000000	3399	C00022	
Sarcosine	Sarcosine	156000000	3398	C00123	

You can download the result [here](#)

Submit

5.2.2 Enrichment Analysis (Upload two column compound list)

MetaboAnalyst 5.0 - user-friendly, end-to-end

Choose one of the following options to proceed

Over Representation Analysis | Single Sample Profiling | Quantitative Enrichment Analysis

Enter your data below (two-column data):

```

D-Glucose 23.92
Pyroglutamic acid 26.38
Formic acid 26.72
Indoxyl sulfate 34.21
Dimethylamine 38.28
Ethanolamine 39.29
Glycolic acid 41.39
L-Glutamine 52.99
L-Histidine 55.95
Trigonelline 57.4
3-Aminoisobutanoic acid 89.76
Taurine 116
Glycine 123.52
Trimethylamine N-oxide 128.04
Citric acid 225.31
Hippuric acid 278.53
    
```

Input Type: Compound names
 Feature Type: -- Specify --
 Biofluid (unit): Urine (umol/mmol_creatinine)
 Use the example data
 - urine sample (umol/mmol_creatinine)

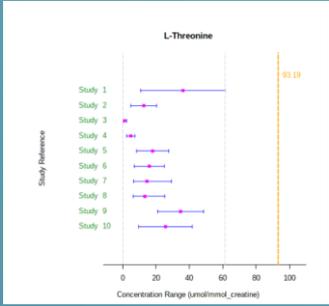
Submit

1. Upload a list of compounds with concentration values

2. Specify the unit of the concentration values you input above.

3. Click "Submit" to upload your data.

4. Check the compound mapping results.



6. Click view to inspect the compound comparison.

NameID Standardization:

- For enrichment analysis, only well-annotated HMDB compounds (i.e. those in our pathway libraries & metabolite lists) will be mapped. For general-purpose name mapping use Compound ID Conversion tool in Other Utilities module.
- Close synonyms are not recognized; they should be replaced by English names (i.e. alpha, beta).
- Query names in normal white indicate exact match - marked by "E" in the download file.
- Query names highlighted indicate no exact or unique match - marked by "N" in the downloaded file.
- For compound names, you should click the View link to perform approximate search and manually select the correct match if found.
- For KEGG ID, if it is possible to have multiple IDs, you should click the View link to manually select the correct match if found.

UnameID	Query	HM	HMDB	PubChem	KEGG	Details
Fumaric acid	Fumaric acid	HMDB0000324	448872	C00322		
Acetone	Acetone	HMDB0000350	358	C00207		
Succinic acid	Succinic acid	HMDB0000254	1111	C00042		
1-Methylhistidine	1-Methylhistidine	HMDB0000001	82109	C02152		
L-Asparagine	L-Asparagine	HMDB0000166	6267	C00152		
3-Methylhistidine	3-Methylhistidine	HMDB0000479	64939	C02152		
L-Threonine	L-Threonine	HMDB0000187	6268	C00189		
Creatine	Creatine	HMDB0000008	509	C00109		
iso-Alanine acid	iso-Alanine acid	HMDB0000022	64235	C00042		
L-Tryptophan	L-Tryptophan	HMDB0000029	6325	C00123		
L-Carnitine	L-Carnitine	HMDB0000062	272489	C00119		
L-Serine	L-Serine	HMDB0000107	5951	C00085		

5. Check the concentration comparison results.

Comparison with Reference Concentration

Note: reference concentrations are in the form of median (range) - exact format. In cases where the ranges were not reported in the original literature, the min and max were calculated using the 95% confidence intervals. In the Comparison column, H, M, L means high, medium (within ranges), lower compared to the reference concentration. Click the image icon below to see a graphical summary for the comparison.

Compound	Concentration	Reference concentrations	Comparison	Detail	Include
Fumaric acid	6.47	6.95 (0.02 - 1.88) 10.2 (2.8 - 43.7) 5.9 (1 - 7.7) 10.7 (0.1 - 28.2) 4.4 (0.2 - 8.6) 8.7 (0.2 - 1.7) 3.542	M	View	
Acetone	0.55	4.2 (0.90 - 19.3) 0.92 (0.2 - 2.8) 3.9 (0.8 - 17.6) 2.24 (0 - 6.75) 3.94 (0 - 7.84)	M	View	
Succinic acid	9.4	7.7 (1.9 - 20) 19.2 (2.4 - 48.6)	M	View	
1-Methylhistidine	9.4	4.4 (1.9 - 10.2) 2.3 (1.4 - 4) 0.9 (0.3 - 3.3) 14.59 (1.919 - 24.24) 10.491 (0 - 30.25)	M	View	
L-Threonine	89.76	8.96 (0.31 - 1.61) 0.211 (0.289 - 1.5) 10.4 (6 - 18.32) 19.52 (8.67 - 14.75) 8.8 (4.6 - 17.7) 9.3 (0 - 20) 18.1	M	View	
Creatine	726	54.3 (24 - 76.6) 87 (14.3 - 160.7) 73.8 (64 - 100.3) 37 (17.3 - 45.3) 25.5 (15 - 35.5)	M	View	
iso-Alanine acid	14.36	15.2 (0.16 - 28.8) 1.25 (1.116 - 1.37) 1.7 (1.2 - 2.4) 1.3 (1.1 - 1.5) 1.8 (1.6 - 2.0)	M	View	
L-Tryptophan	15.78	13.2 (0.16 - 28.8) 1.25 (1.116 - 1.37) 1.7 (1.2 - 2.4) 1.3 (1.1 - 1.5) 1.8 (1.6 - 2.0)	H	View	
L-Carnitine	16.81	22.5 (18.3 - 27.6) 6.4 (5.02 - 15.2) 5 (0.7 - 18.4) 5.7 (0.8 - 16.5) 16.109 (1.737 - 34.601)	M	View	
L-Serine	17.32	28 (1.02 - 44.4) 2.6 (2.37 - 3) 6.9 (4.9 - 13.5) 30.168 (4 - 45.6) 28 (16 - 36) 74.87 (36.9 - 102.41) 27.36	M	View	
L-Tyrosine	67.51	17.1 (10.8 - 27.1) 1.9 (1.6 - 2.2) 1.9 (1.6 - 2.2) 1.9 (1.6 - 2.2) 1.9 (1.6 - 2.2)	M	View	

5.2.3 Enrichment Analysis (Upload concentration table)

1. Upload your data table and specify the format etc.

1. Input the Metabolomic Workbench Study ID to import the data.

Choose one of the following options to proceed

[Over Representation Analysis](#) [Single Sample Profiling](#) [Quantitative Enrichment Analysis](#)

Upload your concentration data (.csv or .txt)

Group Label: Discrete (Classification) Continuous (Regression)

ID Type:

Feature Type:

Data Format:

Data File: No file chosen

Download from Metabolomics Workbench

Study ID:

Try our test data:

Data	ID Type	Group Label	Description
<input checked="" type="radio"/> Data 1	Common name	Discrete	Urinary metabolite concentrations from 77 cancer patients measured by 1H NMR. Phenotype: N - cachexic; Y - control
<input type="radio"/> Data 2	PubChem CID	Continuous	Urinary metabolite concentrations from 97 cancer patients measured by 1H NMR. Phenotype: muscle gain (percentage within 100 days, negative values indicate muscle loss)
<input type="radio"/> Data 3	Common name	Discrete	Lipidomic biomarkers in serum samples from patients with benign and secondary progressive MS obtained from the Metabolomics Workbench (ST000688).

2. Click "**Submit**" to upload your data.

TIP1: MetaAnalyst now allows users to use the study results from Metabolomics Workbench directly by simply providing the STUDY ID.

TIP2: The data pre-processing steps, including data integrity check, ID standardization, and normalization need to be performed step by step.

5.3 Parameter Selection



MetaboAnalyst 5.0 - from raw spectra to patterns and biological insights



Processing

Parameter Setting:

Please select a metabolite set library

Pathway based	<input checked="" type="radio"/> SMPDB	99 metabolite sets based on normal human metabolic pathways.
	<input type="radio"/> KEGG	84 metabolite sets based on KEGG human metabolic pathways (Oct. 2019).
	<input type="radio"/> Drug related	461 metabolite sets based on drug pathways from SMPDB.
Disease signatures	<input type="radio"/> Blood	344 metabolite sets reported in human blood.
	<input type="radio"/> Urine	384 metabolite sets reported in human urine.
	<input type="radio"/> CSF	166 metabolite sets reported in human cerebral spinal fluid (CSF).
	<input type="radio"/> Feces	44 metabolite sets reported in human feces.
Chemical structures	<input type="radio"/> Super-class	35 super chemical class metabolite sets or lipid sets
	<input type="radio"/> Main-class	464 main chemical class metabolite sets or lipid sets
	<input type="radio"/> Sub-class	1072 sub chemical class metabolite sets or lipid sets
Other types	<input type="radio"/> SNPs	4,598 metabolite sets based on their associations with SNPs loci.
	<input type="radio"/> Predicted	912 metabolic sets predicted to change in the case of dysfunctional enzymes.
	<input type="radio"/> Localions	73 metabolite sets based on organ, tissue, and subcellular localizations.
Self defined	Upload here	define your own customized metabolite sets

Only use metabolite sets containing at least 2 entries

Please specify a reference metabolome

- Use all the compounds in the selected library
- [Upload a reference metabolome based on your analytical platform](#)

Submit

R Command History

Keep collapsed

Save

```
1. mSet<-InitDataObjects("conc", "asetor
a", FALSE)
2. cmdP.vec<-c("CerP(d18:1/26:1)", "DG(18:
0/15:0)", "DG(18:2/19:0)", "LysoPC(18:
0)", "LysoPC(17:0)", "LysoPE(22:2)", "PA
(18:1/18:0)", "PA(18:1/21:0)", "PA(20:4/
20:0)", "PA(22:2/24:0)", "PA(22:6/18:
1)", "PC(20:5/18:2)", "PC(P-18:0/18:
1)", "PE(18:1/22:1)", "PE(18:2/16:0)", "P
E(18:2/21:0)", "PE(18:2/22:1)", "PE(20:
2/18:2)", "PE(20:3/20:2)", "PE(20:3/22:
0)", "PE(20:4/18:0)", "PE(20:4/20:0)", "P
E(P-16:0/18:0)", "PE(P-18:0/13:0)", "PE
(P-18:0/17:0)", "PE(P-18:0/20:4)", "PE(P
-18:0/20:5)", "PE(P-18:0/22:1)", "PE(P-2
0:0/22:6)", "PE(18:0/16:0)", "PI(18:1/1
8:0)", "PI(22:6/20:1)", "PI(18:2/18:
1)", "PI(22:2/16:0)", "PS(18:0/21:0)", "P
S(18:1/20:3)", "PS(18:1/22:0)", "PS(18:
1/24:1)", "PS(18:2/22:1)", "PS(20:1/18:
0)", "PS(20:3/21:0)", "PS(22:6/17:2)", "P
S(22:6/18:0)", "SODG(18:0/12:0)")
3. mSet<-Setup_MapData(mSet, cmdP.vec)
4. mSet<-CrossReferencing(mSet, "name", 1
lipid = T)
5. mSet<-CreateMappingResultTable(mSet)
```

Metabolite set library selection. New libraries in V5 include the fecal-disease metabolite set and all chemical structure metabolite sets.

Click "Submit" to perform enrichment analysis.

TIP1: Select the metabolites sets you are going to explore or upload a self-defined sets.

5.4.1 Enrichment Analysis Results – Bar Chart

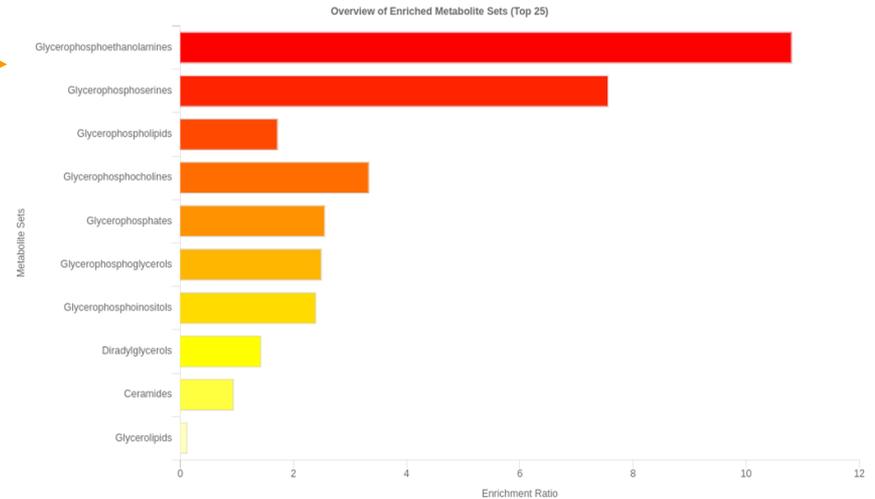


MetaboAnalyst 5.0 - from raw spectra to patterns and biological insights

Results of Enrichment Analysis

- Please scroll down for the detailed result table on the bottom
- **Enrichment Ratio** is computed by Hits / Expected, where hits = observed hits, expected = expected hits (see the Table below)

Bar Chart



Interactive bar-chart of the enrichment analysis results. Hover over the bars to view the Enrichment Ratio per metabolite set.

5.4.2 Enrichment Analysis Results – Network View

MetaboAnalyst 5.0 - user-friendly, end-to-end metabolomics data analysis

Network View

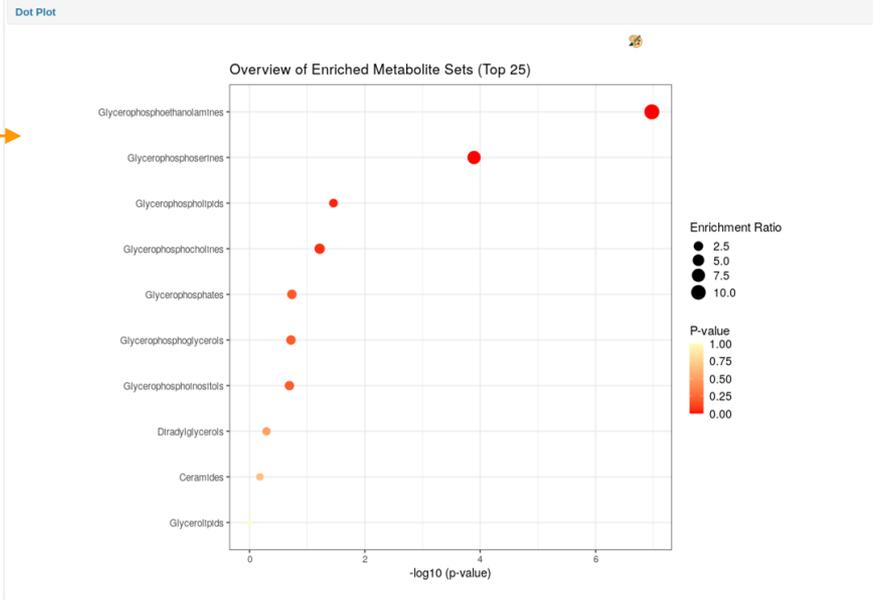
Interactive network of different pathways are displayed here. You can try to drag or zoom in to explore the connection between pathways.

5.4.3 Enrichment Analysis Results – Dot Plot



MetaboAnalyst 5.0 - from raw spectra to patterns and biological insights

Dot plot of the enrichment analysis results. The size of the circles per metabolite set represents the Enrichment Ratio and the color represents the p-value.



Interactive Pie Chart

5.4.4 Enrichment Analysis Results – Summary table



MetaboAnalyst 5.0 - user-friendly, end-to-end metabolomics data analysis

Home

Upload

Processing

Normalization

Enrichment

Set parameter

View result

Download

Exit

Metabolite Set	Total	Hits	Statistic	Expected	P value	Holm P	FDR	Details
Schizophrenia	32	17	12.55	1.3158	3.5583E-4	0.039141	0.003311	View
Rats Treated With Hgd2	13	13	12.814	1.3158	3.6236E-4	0.039497	0.003311	View
Difference Between Normal Han-Wistar (Hw) And Sprague-Dawley (Sd) Rats	10	10	12.763	1.3158	3.633E-4	0.039497	0.003311	View
Propionic Acidemia	20	11	12.843	1.3158	3.8979E-4	0.041708	0.003311	View
Paraquat Poisoning	7	7	13.083	1.3158	3.9458E-4	0.041825	0.003311	View
Tyrosinemia I	8	6	12.992	1.3158	3.9493E-4	0.041825	0.003311	View
Phenylketonuria	12	4	13.466	1.3158	4.0033E-4	0.041825	0.003311	View
Feeding: Amino Acid Formula [Dd]	5	2	14.973	1.3158	4.695E-4	0.048358	0.003311	View
Maple Syrup Urine Disease	9	7	13.855	1.3158	4.9515E-4	0.050506	0.003311	View
Metabolites Affected By Age	12	12	14.547	1.3158	4.9763E-4	0.050506	0.003311	View
Metabolites Affected By Gender	8	5	14.234	1.3158	5.0144E-4	0.050506	0.003311	View
Peripheral Inflammatory Pain In The Rat (Fca Model)	10	7	14.169	1.3158	5.0252E-4	0.050506	0.003311	View
X-Linked Creatine-Transporter Defect	4	4	14.205	1.3158	5.027E-4	0.050506	0.003311	View
Primary Hypomagnesemia	4	2	14.251	1.3158	5.0295E-4	0.050506	0.003311	View
Creatine Deficiency, Guanidinoacetate Methyltransferase Deficiency	4	3	14.933	1.3158	5.115E-4	0.050506	0.003311	View
Asphyxia [Dd]	12	4	14.902	1.3158	5.115E-4	0.050506	0.003311	View
Metabolites Affected By Diurnal Variation	6	5	14.928	1.3158	5.117E-4	0.050506	0.003311	View
Cachexia	3	3	11.205	1.3158	7.428E-4	0.06908	0.0045393	View
Fructose-1,6-Diphosphatase Deficiency	3	1	13.16	1.3158	0.0011854	0.10905	0.0068627	View
Hartnup Disease	12	10	10.35	1.3158	0.001424	0.12958	0.007832	View
Fumarase Deficiency	5	2	11.176	1.3158	0.0018152	0.16337	0.0090793	View
Spastic Ataxia	5	1	12.143	1.3158	0.0018984	0.16896	0.0090793	View
Tryptophanuria	2	1	12.143	1.3158	0.0018984	0.16896	0.0090793	View
Hemodialysis	12	1	11.154	1.3158	0.002993	0.26039	0.01359	View
Lung Cancer	32	27	6.9615	1.3158	0.0030887	0.26563	0.01359	View
Ketosis, Unspecific [Dd]	8	3	7.9804	1.3158	0.0042017	0.35715	0.017776	View
Rats Treated With Hydrazine	9	7	7.0029	1.3158	0.0047579	0.39966	0.019384	View
Multiple Sclerosis	21	1	9.6108	1.3158	0.0060732	0.50408	0.023859	View
Malnutrition	10	2	7.7399	1.3158	0.0066558	0.54577	0.024601	View
Medium Chain Acid Coa Dehydrogenase Deficiency	15	2	7.7731	1.3158	0.0075444	0.64107	0.024601	View

Tabular summary of the enrichment analysis results. Click "Details" to view which features were hits in the metabolite set.

5.5 Result Downloading & New Journey

MetaboAnalyst 5.0 - user-friendly, end-to-end metabolomics data analysis

Click the **"Generate Report"** to download a pdf report summarizing your analysis.

Click the **"Explore Data in Other Modules"** to directly use your data in other modules (e.g. Statistical Analysis Module)

Click the **"Download.zip"** to download a zipped folder containing all data and created images.

Click the **"Logout"** to finish the processing.

Results Download | Start New Journey

General Statistics

Targeted Metabolomics

Untargeted Metabolomics

GO!

Download.zip	oea_0_dpi72.png
Rhistory.R	oea_map.csv
data_processed.csv	oea_dot_0_dpi72.png
snorm_1_dpi72.png	Schizophreniadi72.png
Tryptophanuriadi72.png	data_normalized.csv
mseae_oea_result.csv	data_original.csv
mseae_network.json	norm_1_dpi72.png

Logout

Thanks

*If you have any questions please read through the FAQs or contact us at
[Zhiqiang.pang\[at\]xialab.ca](mailto:Zhiqiang.pang@xialab.ca) or [Jeff.xia\[at\]xialab.ca](mailto:Jeff.xia@xialab.ca)*

