



# MetaboAnalyst 5.0

---

A Web-based Tool for streamlined  
Metabolomics Data Analysis

2022.07.12

### 3. Meta-Analysis (MS Peaks)

The **Meta-Analysis of MS Peaks** module extends the mummichog algorithm (Li et al. 2013) implemented in the MS Peaks to Paths module to perform biological-interpretation based meta-analysis.

#### Highlights:

- Support for meta-analysis of untargeted metabolomics data integration at pathway level;
- Support for meta-analysis of multiple untargeted metabolomics data by pooling their peaks.



## 3.0 Knowledge & Background

### Why meta-analysis?

While individual studies may identify certain results, such results may not be reproducible in other independent studies of the same biological questions due to low sample size, sample heterogeneity, the type of LC-MS platform used, or metrics used for interpreting results. Meta-analysis, which is the combination of findings from independent studies - can be used to overcome such limitations and ultimately increase the power, precision, and generalizability of a study.

The **Meta-Analysis (MS Peaks)** module of MetaboAnalyst now permits users to perform meta-analysis of untargeted metabolomics data. The method implemented in this module also reduces the bias individual studies may carry towards specific sample processing protocols or LC-MS instruments.

A decorative network diagram in the bottom right corner of the slide. It consists of a complex web of white lines connecting various white dots of different sizes. Some dots are larger and act as hubs, while others are smaller and peripheral. The lines form a dense, interconnected structure that resembles a molecular network or a data graph.

## 3.2 Start Meta-Analysis (MS Peaks)

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

### Module Overview

Input Data Type	Available Modules (click 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

Click here to start

Show R command history

# 3.3.1 Meta-Analysis MS Peaks Upload

**Please upload and process your data**

Use the panel below to prepare each individual data. Click the individual cells to activate each process. Click **Add New** to add a new data set. The maximum total number of samples allowed is **1000**. When all data sets have been processed, Click **Proceed** to proceed. Click the **Try Examples** button to try our example datasets.

Data Upload	Integrity Check	Normalization	Analysis	Summary	Include
<input type="button" value="Upload"/>	<input type="button" value="Process"/>	<input type="button" value="Normalize"/>	<input type="button" value="Analyze"/>	<input type="button" value="Detail"/>	<input type="checkbox"/>
			<input type="button" value="Try Examples"/>	<input type="button" value="Add New"/>	<input type="button" value="Proceed"/>

1. Upload each peak table individually.
2. Perform sanity check and view data.
3. Normalize the data.
4. Perform a t-test to identify important features.
5. View summary of statistical test.

Try our example data here

**TIP 1:** The uploading table should include all peaks after processing rather than several interested peaks.

Click here to include (or not) the peak table in the meta-analysis.

**TIP 2:** The mummichog algorithm (Li et al. 2013) works best when there are 10~25% significant peaks. Please try to adjust the percentage of significant features by changing the p value cutoff.

**TIP 3:** After performing the uploading and analyzing. Users could optionally define the potential adducts for more accurate mapping in the coming steps.

## 3.3.2 Meta-Analysis MS Peaks Adducts Customization

Please upload and process your data

Use the panel below to prepare each individual data. Click the individual cells to activate each process. Click **Add New** to add a new data set. The maximum total number of samples allowed is **1000**. When all data sets have been processed, Click **Proceed** to proceed. Click the **Try Examples** button to try our example datasets.

Data Upload	Integrity Check	Normalization	Analysis	Summary	Include
<input checked="" type="checkbox"/> A1_pos	<input checked="" type="checkbox"/> Process	<input checked="" type="checkbox"/> Normalize	<input checked="" type="checkbox"/> Analyze	<input checked="" type="checkbox"/> Detail	<input checked="" type="checkbox"/>
<input checked="" type="checkbox"/> B1_pos				<input checked="" type="checkbox"/> Detail	<input checked="" type="checkbox"/>
<input checked="" type="checkbox"/> C1_pos				<input checked="" type="checkbox"/> Detail	<input checked="" type="checkbox"/>

Try Examples Add New Proceed

### Data Analysis

T-tests will be performed on the uploaded peak table to determine the number of significant peaks. Note that the mummichog algorithm (Li et al. 2013) works best when there are 10–25% significant peaks. The pie chart below shows the percentage of significant peaks in your dataset.

Set p value cutoff:

Customize adducts (Optional): [Edit Adducts](#)

Submit Done

### Adduct Customization

Please customize the adducts based on your experimental conditions (e.g. Mobile Phases). Click the **Confirm** button to finalize these changes.

Available		Include
M+2Na [2+]	>	M [1+]
M+3Na [3+]	>>	M+H [1+]
M+H+2Na [3+]	<	M+2H [2+]
M+2H+Na [3+]	<<	M+3H [3+]

Confirm

Customize the adducts, click 'Done' to finish the setting.

## 3.4 Upload Using Example Data

**TIP:** Users can either upload a single ion mode table as a dataset, or tables from both ion modes together by choosing the appropriate 'Ion Mode'.

**Please upload and process your data**

Use the panel below to prepare each individual data. Click the individual cells to activate each process. Click **Add New** to add a new data set. The maximum total number of samples allowed is **1000**. When all data sets have been processed, Click **Proceed** to proceed. Click the **Try Examples** button to try our example datasets.

Data Upload	Integrity Check	Normalization	Analysis	Summary	Include
✓ A1_pos	✓ Process	✓ Normalize	✓ Analyze	✓ Detail	✓
✓ B1_pos	✓ Process	✓ Normalize	✓ Analyze	✓ Detail	✓
✓ C1_pos	✓ Process	✓ Normalize	✓ Analyze	✓ Detail	✓

> Try Examples   ⊕ Add New   ⌂ Proceed

Click **Proceed** to move forward with the meta-analysis

# 3.5.1 Parameter Setting : Pathway-level Integration

1. Start meta-analysis integration at pathway or peaks

2. Set parameters for pathway integration

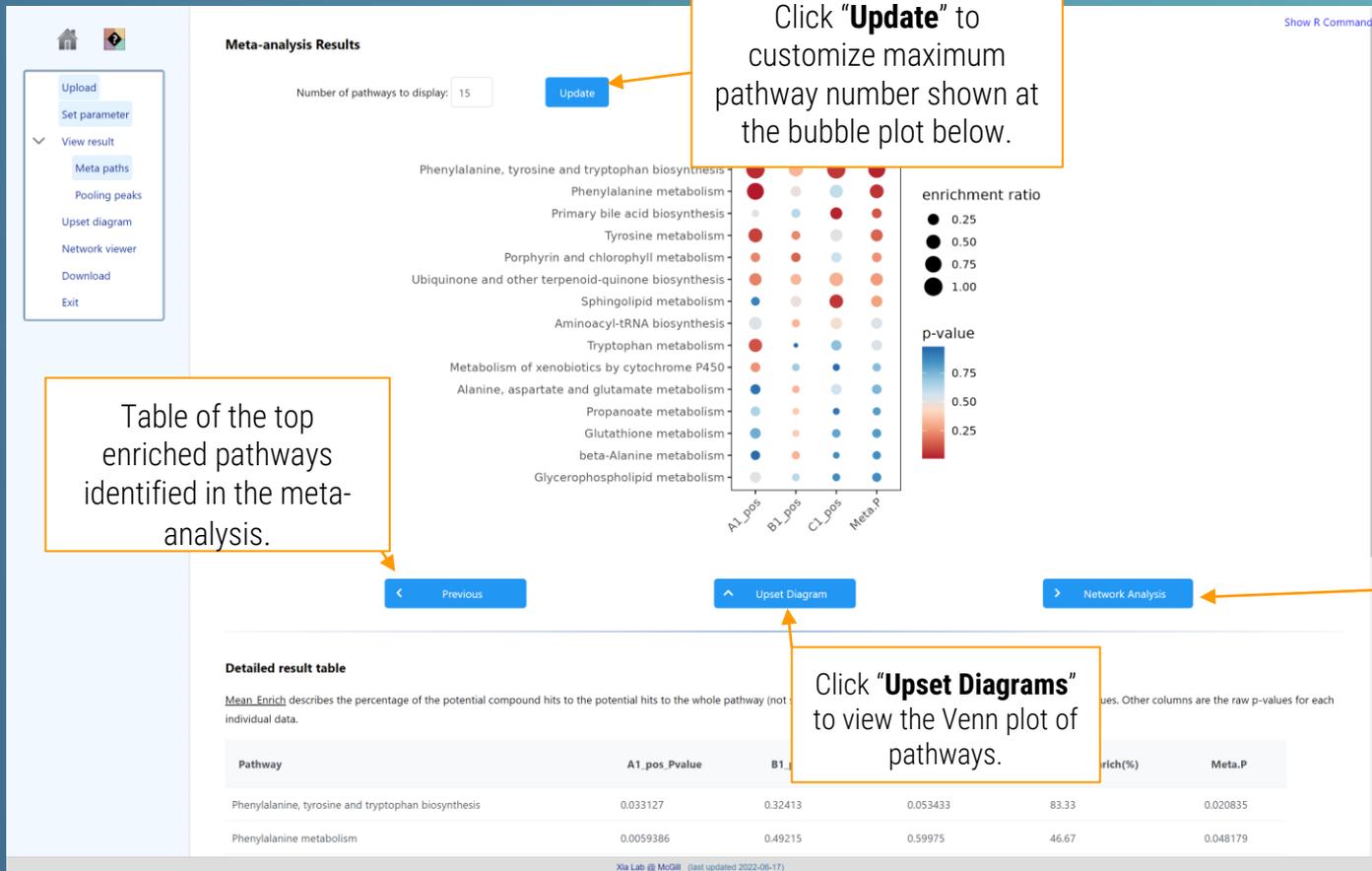
3. Click "Submit" to process the analysis

**TIP1:** Users could select to perform the meta-analysis at pathway level (by integrating p values) or at peaks level from pooling peaks together.

**TIP2:** p value integration provide 6 approaches here. Briefly, **Fisher** is sensitive to small p values and not recommended for 6 or more datasets. **Min** and **Max** are mathematically the minimum or maximum p values. **Stouffer** is used for data following gaussian curve, while **Edgington** bests fits circular data and sensitive to large p value. **Vote-counting** weight all datasets same and give a vote if there is a study meet the confidence interval ( $p < 0.05$ ), at least 5 datasets are required to reach the significant level (meta.p < 0.05) for vote-counting.

[>>Read More<<](#)

## 3.5.2 Results (Pathway Level integration)



Click **“Update”** to customize maximum pathway number shown at the bubble plot below.

Table of the top enriched pathways identified in the meta-analysis.

**TIP:** Pathway integration results are shown as a bubble plot. The node size is positively correlated with the enrichment ratio of the corresponding pathway. Meta.P is the integrated P values.

Click **“Network Analysis”** to explore the data in the context of a KEGG metabolic network.

Click **“Upset Diagrams”** to view the Venn plot of pathways.

# 3.5.3 Results (Upset Diagram)

**TIP:** Venn diagram is providing the intersection results of different datasets.

Update

Adjust the **p-value cutoff** below to control the number of selected pathways.

P-value cutoff: 0.05

Name	Pathways	Include
A1_pos	3	<input checked="" type="checkbox"/>
B1_pos	0	<input checked="" type="checkbox"/>
C1_pos	1	<input checked="" type="checkbox"/>
meta_dat	2	<input checked="" type="checkbox"/>

Manually set the cut-off of p values of pathways to include for Venn diagram.

This UpSet diagram is useful for displaying all possible comparisons between datasets.

- Different nodes and connecting nodes correspond to the possible comparisons between datasets that contain hits
- Different vertical bars represent the number of hits corresponding to select comparison. Horizontal bars represent the number of hits corresponding to select dataset.
- **Click** on nodes and bars to show the corresponding features on the left panel.
- At most **four** datasets can be compared at the same time. Datasets without hits will **NOT** be shown here.

Intersection Size

● (A1\_pos ∩ C1\_pos ∩ meta\_dat): 2

● (A1\_pos ∩ C1\_pos ∩ meta\_dat): 2

Intersection	Intersection Size
A1_pos	13
B1_pos	10
C1_pos	7
meta_dat	7
A1_pos ∩ B1_pos	5
A1_pos ∩ C1_pos	4
A1_pos ∩ meta_dat	4
B1_pos ∩ C1_pos	4
B1_pos ∩ meta_dat	4
C1_pos ∩ meta_dat	2
A1_pos ∩ B1_pos ∩ C1_pos	2
A1_pos ∩ B1_pos ∩ meta_dat	2
A1_pos ∩ C1_pos ∩ meta_dat	2
B1_pos ∩ C1_pos ∩ meta_dat	2
A1_pos ∩ B1_pos ∩ C1_pos ∩ meta_dat	2

7 meta\_dat

7 C1\_pos

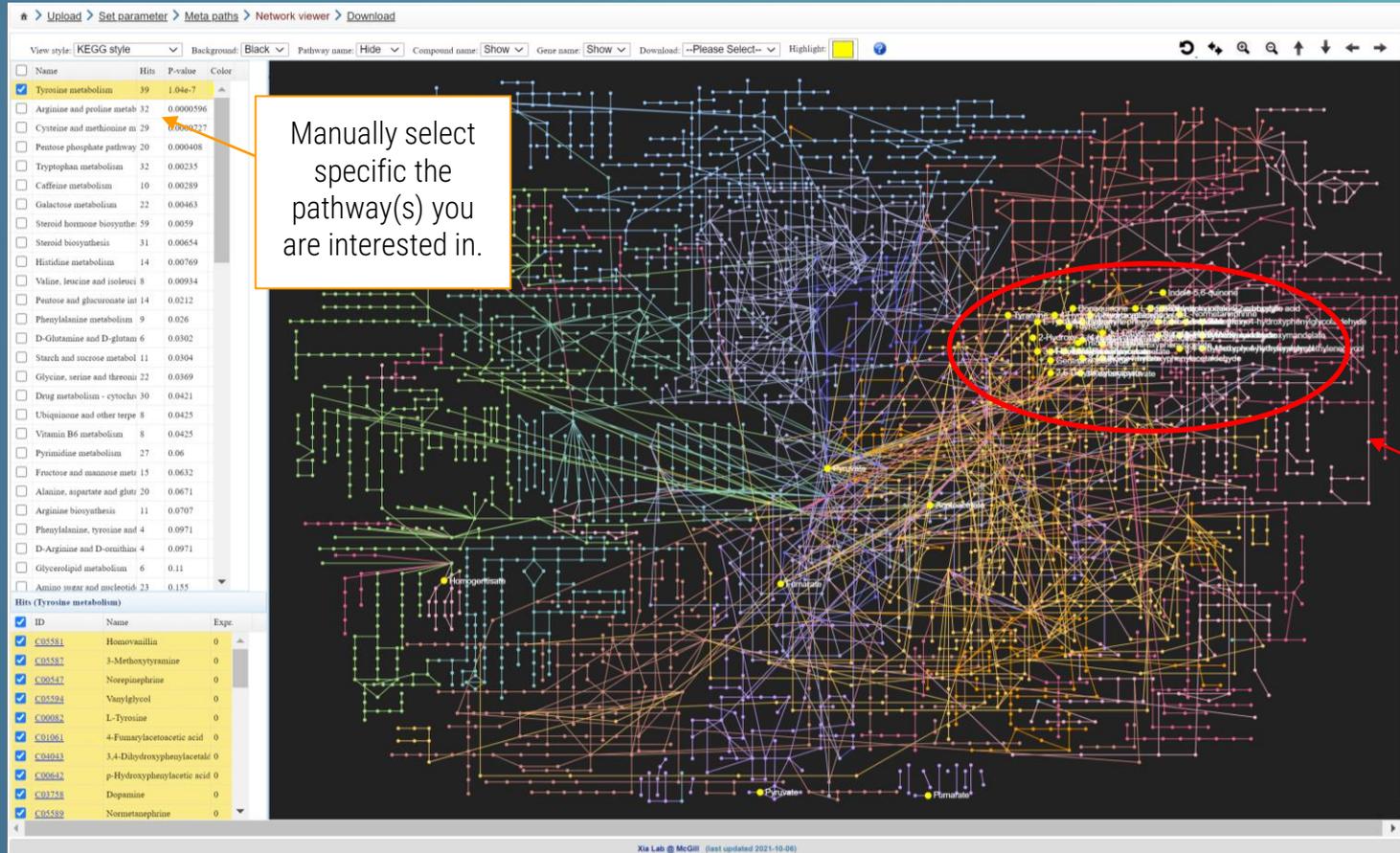
10 B1\_pos

13 A1\_pos

Click certain part to view the intersected pathways

- Total:2
- Phenylalanine, tyrosine and tryptophan biosynthesis
- Ubiquinone and other terpenoid-quinone biosynthesis

# 3.5.4 Results (Network Exploration)



# 3.6.1 Parameter Setting – Pooling peaks

**TIP:** Mummichog version 1 is more compatible for heterogenous data (like different LC-MS platforms/ extraction method etc.), while the version 2 is only applicable for the same LC-MS condition.

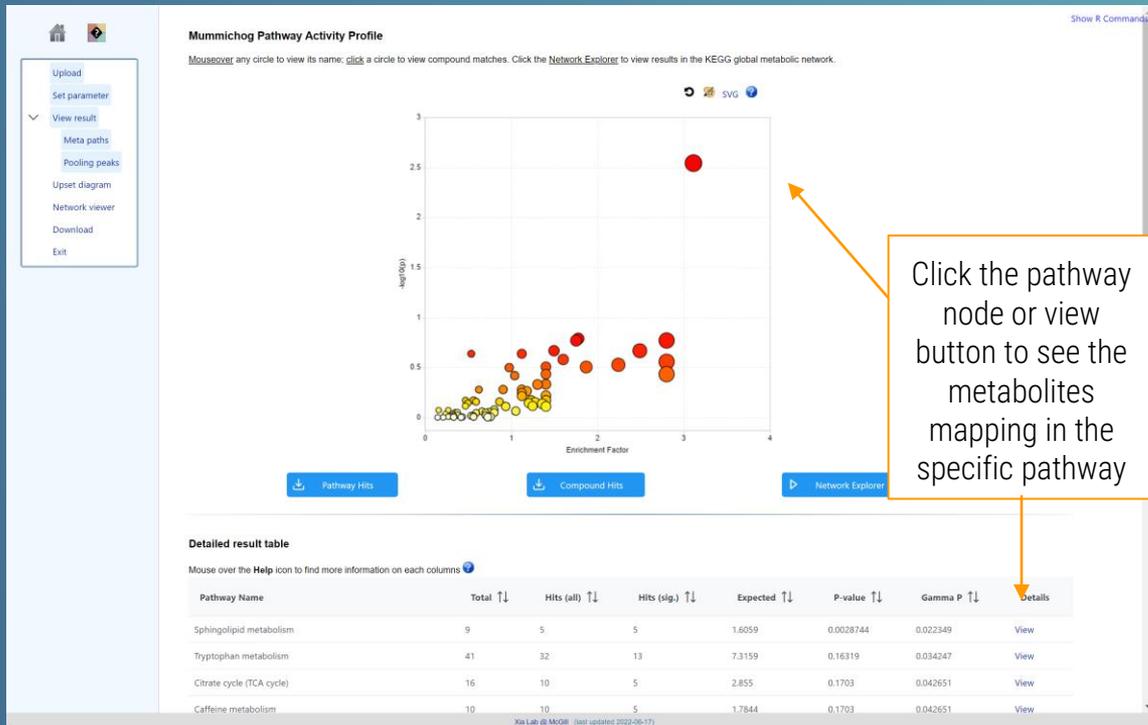
1. Start meta-analysis from pooling peaks

2. Set parameters for functional analysis from pooling peaks.

3. Click "Submit" to process the analysis

The screenshot displays the 'Parameter Settings' page for 'Pooling Peaks'. On the left is a navigation menu with options: Upload, Set parameter, View result, Meta paths, Pooling peaks, Upset diagram, Network viewer, Download, and Exit. The main content area is titled 'Parameter Settings' and contains two sections: 'Pathway-level Integration' and 'Pooling Peaks'. The 'Pooling Peaks' section includes a description of the method, a tip, and a form with the following fields: 'Algorithm' (radio buttons for Mummichog and GSEA), 'Version' (dropdown menu set to 'Version 1'), 'P-value cutoff' (input field set to '0.001'), and 'Pathway library' (dropdown menu set to 'Homo sapiens (human) [KEGG]'). A blue 'Submit' button is located to the right of the form. Three orange callout boxes with arrows point to the 'Pooling Peaks' section, the 'Submit' button, and the 'Submit' button in the 'Pathway-level Integration' section above.

## 3.6.2 Results (Pooling peaks)



**TIP:** Click the network explorer at the bottom of this page to start the exploration of network analysis. You can also download the pathway or compound hits by clicking the corresponding buttons.

The colored compounds/empirical compounds indicate potential matches from the user's input, with red colors indicating significant hits and blue colors indicating non-significant hits.

Pathway	Metabolites
Sphingolipid metabolism	C06124, C01120, <b>C00836</b> , <b>C00319</b> , <b>C00065</b> , C00154, <b>C02934</b> , <b>C12144</b> , C00346

Sphingolipid metabolism C43444, C00346, C00154, C02934, C00836, C00319, C00065, C00154, C02934, C12144, C00346

# 3.7 Result Downloading & New Journey

Click the “**Start New Journey**” to directly use your data in other modules (e.g. Statistical Analysis Module)

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

**Download Results & Start New Journey**

Please download the results (tables and images) from the **Results Download** tab below. The **Download** button. Finally, you can continue to explore other compatible modules using the **Start New Journey** button.

**Results Download**   **Start New Journey**

**Generate Report**

<a href="#">download.zip</a>	<a href="#">A1_nos_norm.csv</a>
<a href="#">Rhistory.R</a>	<a href="#">meta_bubble_0_7_300.png</a>
<a href="#">t_test.csv</a>	<a href="#">mummichog_matched_compound_all.csv</a>
<a href="#">A1_eos_norm_boxdei72.png</a>	<a href="#">C1_eos.csv</a>
<a href="#">C1_eos_nprocessed.csv</a>	<a href="#">B1_eos_nprocessed.csv</a>
<a href="#">eos_norm_boxdei72.png</a>	<a href="#">ms_peaks_meta_anal_cred_matching.json</a>
<a href="#">eos_norm_boxdei72.png</a>	<a href="#">B1_eos.csv</a>
<a href="#">eos_to_paths_0_dei72.png</a>	<a href="#">A1_eos.csv</a>
<a href="#">eos_norm.csv</a>	<a href="#">Result_summary.csv</a>
<a href="#">peaks_meta_anal_all_results.csv</a>	<a href="#">C1_eos_norm.csv</a>
<a href="#">mummichog_query.json</a>	<a href="#">scattermum.json</a>
<a href="#">A1_eos_nprocessed.csv</a>	<a href="#">mummichog_pathway_enrichment.csv</a>

**Logout**



**Results Download**   **Start New Journey**

<b>General Statistics</b>	<input checked="" type="radio"/> Statistical Analysis
	<input type="radio"/> Biomarker Analysis
	<input type="radio"/> Time-series/Two-factor
	<input type="radio"/> Power Analysis
<b>Targeted Metabolomics</b>	<input type="radio"/> Enrichment Analysis
	<input type="radio"/> Pathway Analysis
<b>Untargeted Metabolomics</b>	<input type="radio"/> Functional Analysis

**GO!**

# *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)*

