OmicsNet Tutorial: Omics Network Starting from MS Peaks

Computer Requirement

- Modern browser supporting **WebGL**
- Chrome 50+, Firefox 47+, Safari 10.1+ and Edge 12+
- Please make sure WebGL is enabled in your browser
  - Please consult this web page to verify: https://get.webgl.org/
- If not enabled, please consult our **FAQ** page for instructions
- For best performance and visualization, use:
- Latest version of **Google Chrome**
- A modern computer with at least 4GB of physical RAM
- A 15-inch screen or bigger (larger is better)
- Retina Display is supported

Motivation

1. **Background :** Metabolomics based on Liquid Chromatography coupled with Mass spectrometer (LC-MS) has been extensively used for global detection on whole metabolome of a certain system;
2. **Bottleneck:** Annotation of MS peaks usually laborious and require commercial standards for chemical identification;
3. **Aim:** Comprehensively annotating the whole metabolomics peak set without tedious compound identification;
4. **Approach:** Using biochemical/abiotic transformation network-based global annotation.

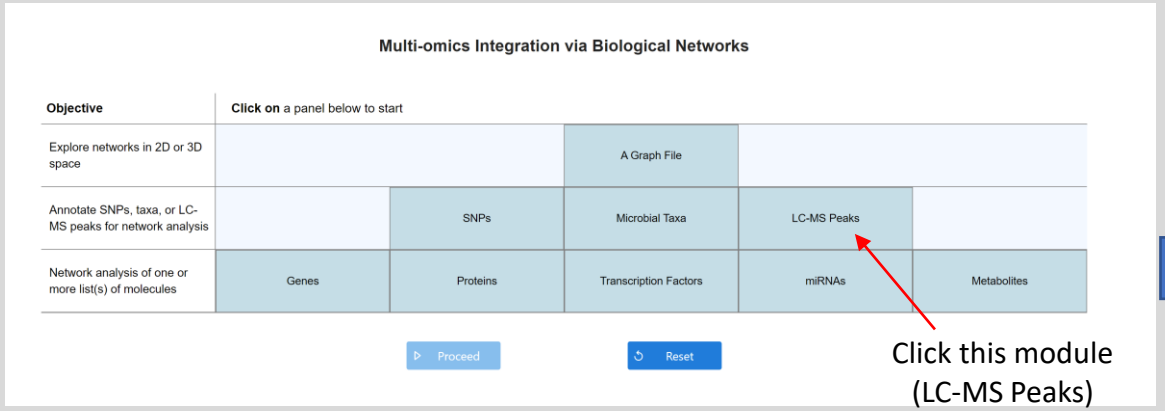
Data format for uploading

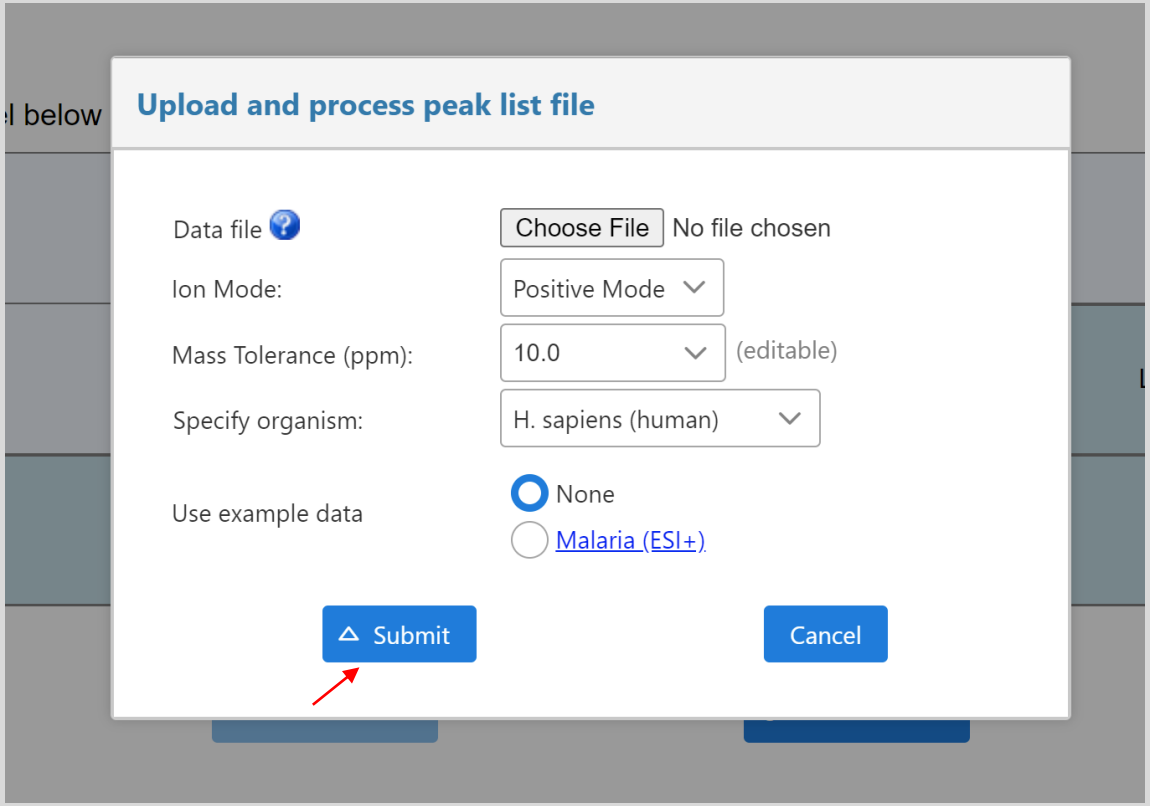
- **Uploading data should be a peak list with four required columns.**
- 1st column: m/z values;
- 2nd column: RT (retention time);
- 3rd column: intensity values of the peaks;
- 4th column: statistical p values of the peak intensities between different groups.

	A	B	C	D
1	Mz	Rt	Int	Pvalue
2	85.065	0.41	109023.1	0.486
3	85.0843	2.79	2002051	0.197
4	86.0602	0.62	623104.3	0.016
5	86.0602	2.15	2304644	0.003
6	86.0603	0.52	878208.6	0
7	86.0966	0.98	1.14E+08	0.22
8	86.0966	1.57	90493483	0.096
9	86.0966	1.38	95475583	0.124
10	86.0966	1.13	2.1E+08	0.053
11	86.0966	1.87	1.92E+08	0
12	86.3224	1.88	61372.2	0.029
13	86.4142	1.66	74522.7	0
14	86.9088	1.3	1365505	0.003
15	86.9089	1.99	298156.1	0.754
16	87.0555	2.09	1422187	0.114
17	87.0807	0.4	16051646	0.092
18	87.1	1.57	6070507	0.167
19	87.1	1.37	5840515	0.092
20	87.1001	0.99	8257790	0
21	87.1001	1.87	11949551	0.063
22	87.1001	1.14	11987004	0.038
23	88.0396	1.97	2594988	0.117
24	88.0634	2.69	11900171	0.087
25	88.084	0.41	1087337	0.413
26	88.5093	1.51	11359624	0.338
27	88.5094	0.41	730502.2	0.83

Example

Parameters for peak preprocessing

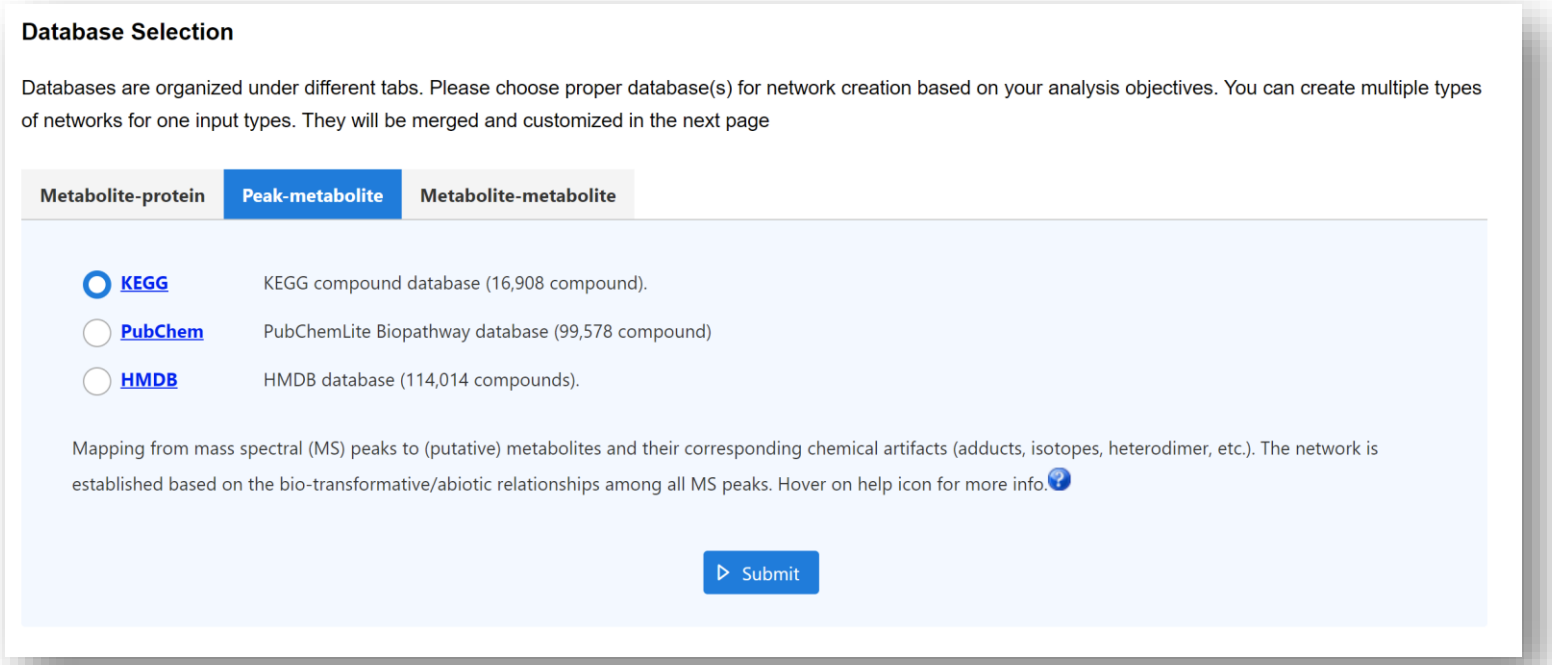




From the uploading panel (as shown in the left side), user need to specify the Ion mode (polarity of ESI), mass tolerance and organism. Click ‘Submit’ to upload your data. Then click “Proceed” button from the home page to go the next.

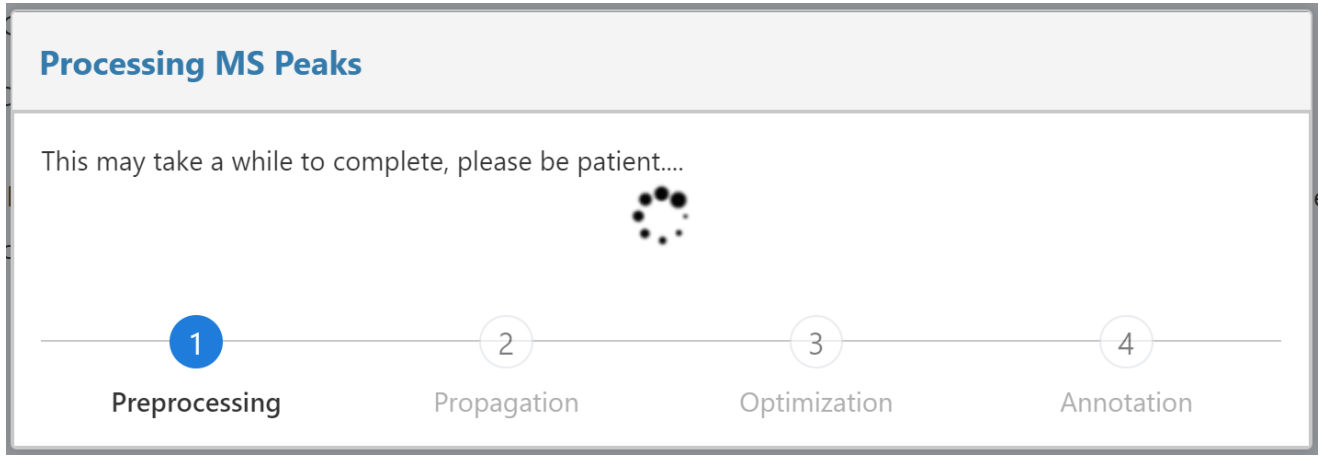
Try to select the example for practice!

Database Selection



There are three different compound databases (KEGG, PubChem Lite BioPathway and HMDB) provided for compound annotation. Different databases have different sizes. The smaller of the size, the faster annotation/processing will be.

Data Processing Steps



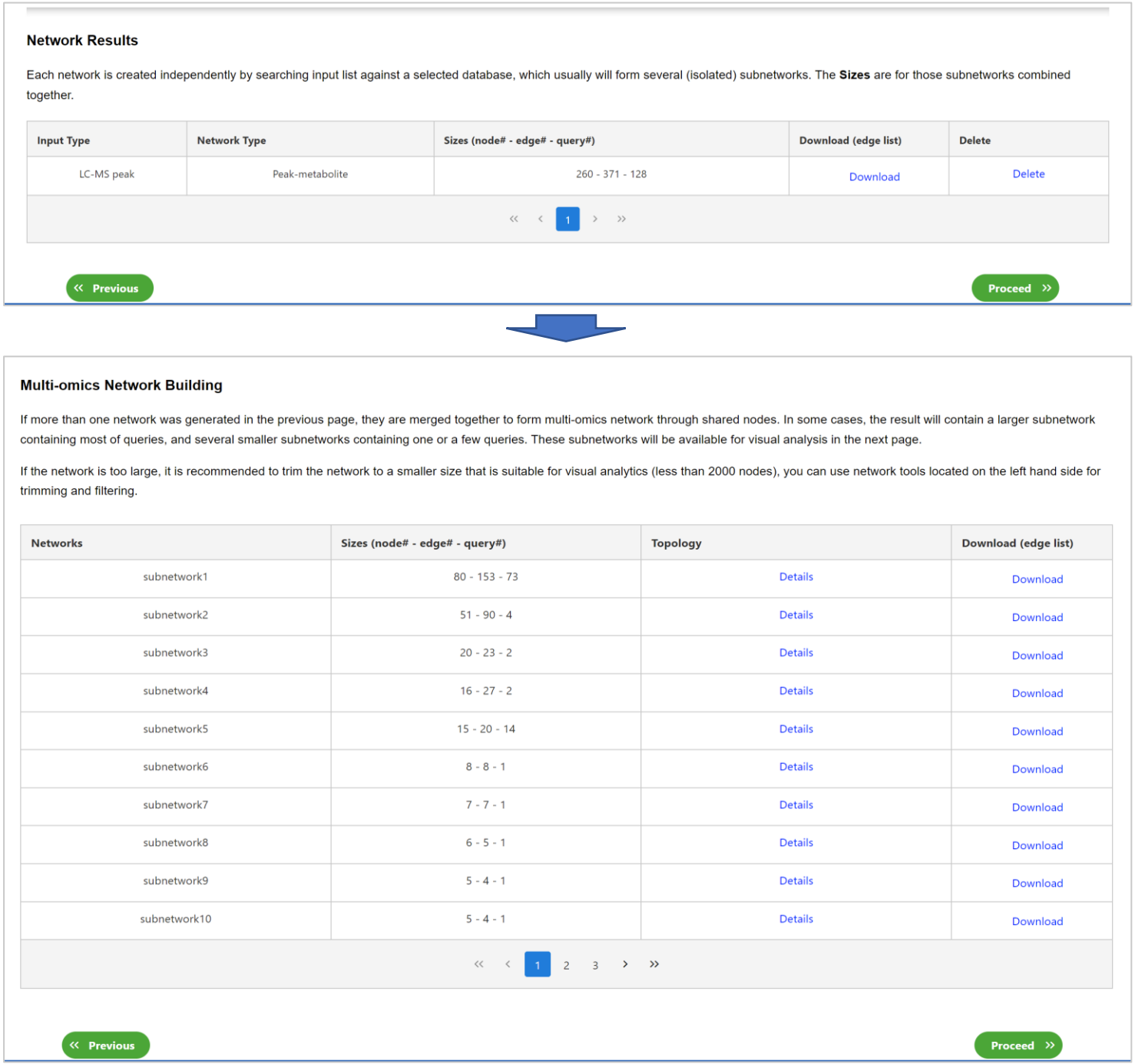
MS peak processing mainly includes 4 steps:

1. Processing: Seed compounds identification and structure base generation;
2. Propagation: Bio-transformative/abiotic propagation to cover the metabolome;
3. Network Construction: Optimal network construction with integer linear programming;
4. Annotation: Data organization and MS peak results output.

Peak to metabolite Network construction

- After uploading the peak list, the annotation will be executed automatically. Once the annotation and processing finished, user will be able to view the summary of the constructed network.

- Click “Proceed” to get the details of all subnetworks and click “Proceed” again to view the networks in 2D or 3D.



Network Results

Each network is created independently by searching input list against a selected database, which usually will form several (isolated) subnetworks. The **Sizes** are for those subnetworks combined together.

Input Type	Network Type	Sizes (node# - edge# - query#)	Download (edge list)	Delete
LC-MS peak	Peak-metabolite	260 - 371 - 128	Download	Delete

« 1 » »

« Previous Proceed »

Multi-omics Network Building

If more than one network was generated in the previous page, they are merged together to form multi-omics network through shared nodes. In some cases, the result will contain a larger subnetwork containing most of queries, and several smaller subnetworks containing one or a few queries. These subnetworks will be available for visual analysis in the next page.

If the network is too large, it is recommended to trim the network to a smaller size that is suitable for visual analytics (less than 2000 nodes), you can use network tools located on the left hand side for trimming and filtering.

Networks	Sizes (node# - edge# - query#)	Topology	Download (edge list)
subnetwork1	80 - 153 - 73	Details	Download
subnetwork2	51 - 90 - 4	Details	Download
subnetwork3	20 - 23 - 2	Details	Download
subnetwork4	16 - 27 - 2	Details	Download
subnetwork5	15 - 20 - 14	Details	Download
subnetwork6	8 - 8 - 1	Details	Download
subnetwork7	7 - 7 - 1	Details	Download
subnetwork8	6 - 5 - 1	Details	Download
subnetwork9	5 - 4 - 1	Details	Download
subnetwork10	5 - 4 - 1	Details	Download

« 1 2 3 » »

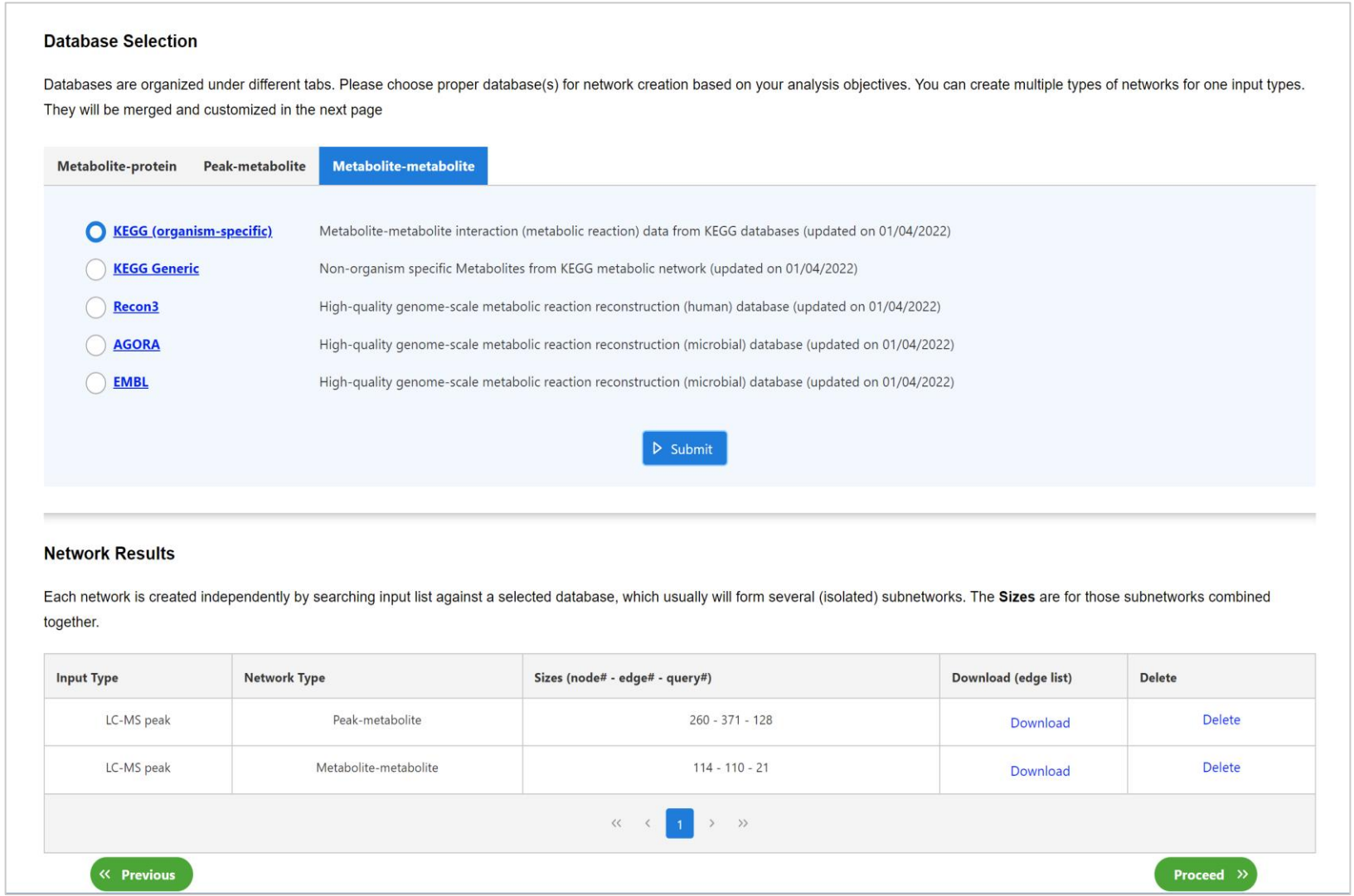
« Previous Proceed »

Metabolic reaction network expansion

- The peak-metabolite network can be extended according to the metabolic reactions from “Metabolite-metabolite” tab.

- Herein, user could select organism specific KEGG database or other metabolic databases.

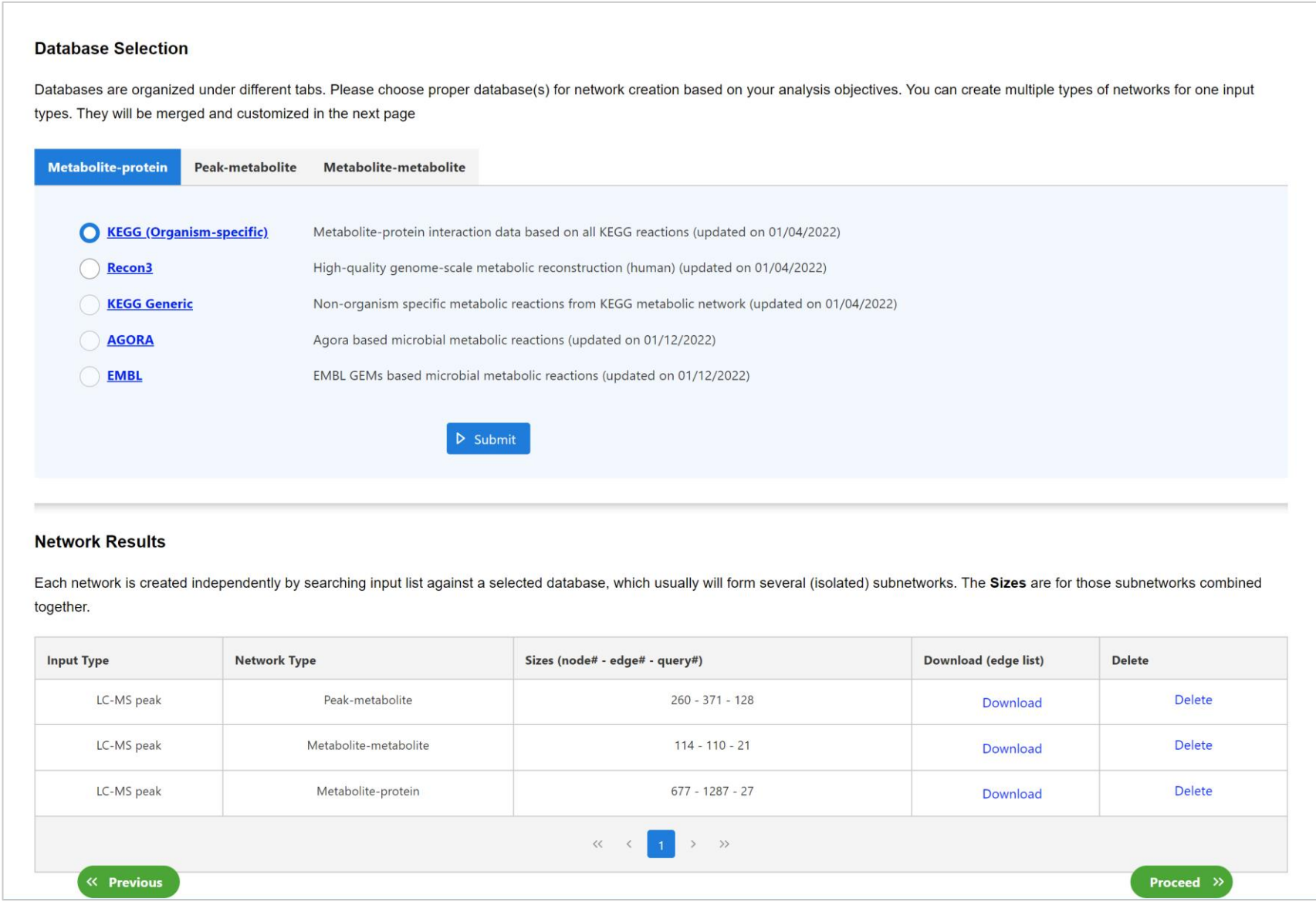
- Click “Submit” to expand the network.



Metabolite-enzyme network construction

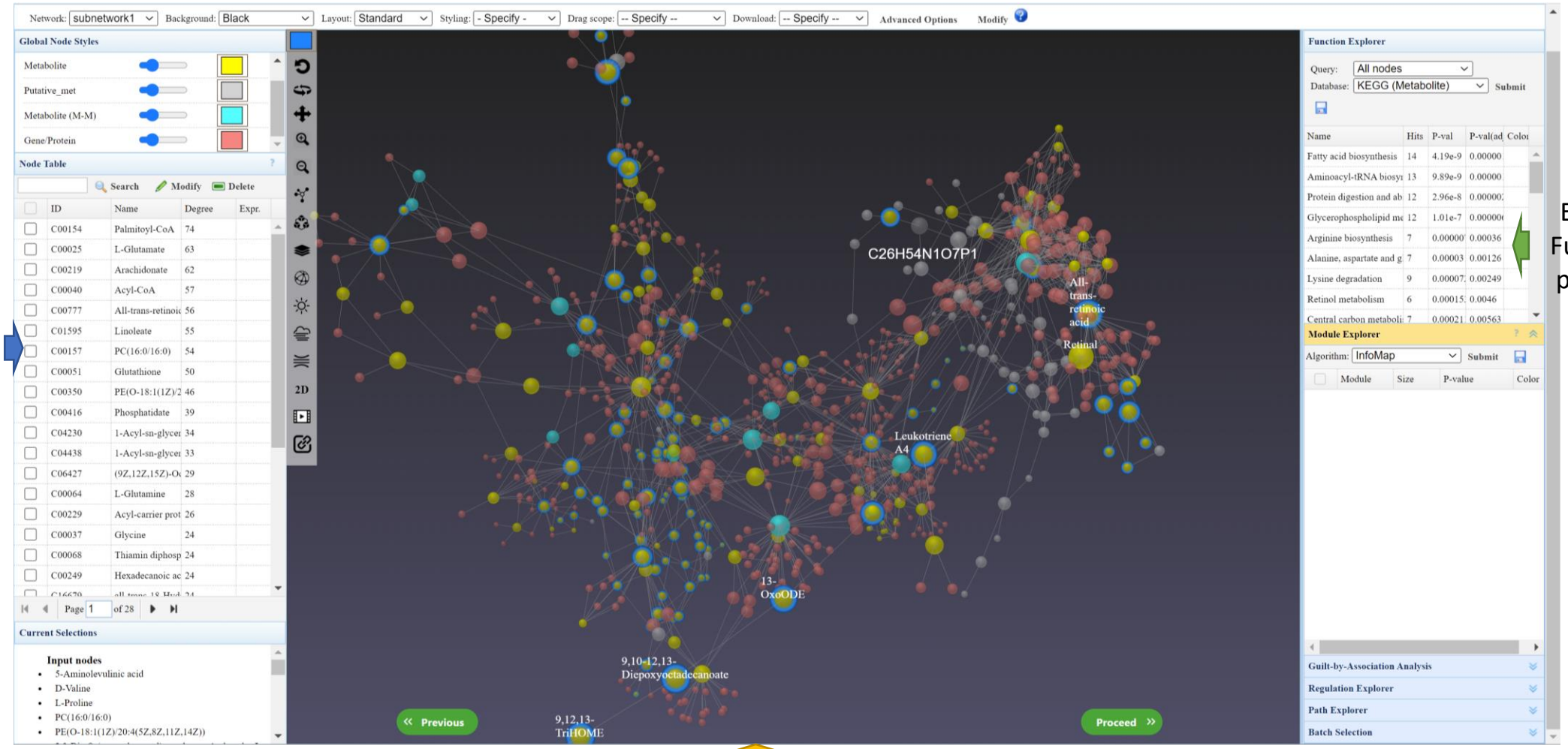
- The peak-metabolite network can be extended according to the metabolite-enzyme relationship from “Metabolite-protein” tab.

- Herein, user could select multiple databases and Click “Submit” to expand the network.



Results of Omics Network

Compound and genes (proteins) of the whole network



Enriched Functional pathways

Constructed network

Result Explanation

- The MS peaks from LC-MS based metabolomics usually consists of many chemical artifacts, and as a result, only a small percentage of the MS matches to compounds;
- “Peak-metabolite” module allows to show the peak annotation network from NetID algorithms, with 20~90% of all peaks annotated as a certain chemical component;
- “Metabolite-metabolite” enhances the connection of the annotated metabolomics data by leveraging the metabolic reaction from knowledgebase (KEGG et al);
- “Metabolite-protein” further adds an extra layer to show the enzyme component of the corresponding metabolic network;
- In brief, the integration results originating from MS peaks intuitively display the (potential) complicated interaction within a metabolism-related system.

The End