Metabolomic Data Analysis with MetaboAnalyst 5.0

Name: guest17886441071731322890

June 9, 2023

1 Background

The Pathway Analysis module combines results from powerful pathway enrichment analysis with pathway topology analysis to help researchers identify the most relevant pathways involved in the conditions under study.

There are many commercial pathway analysis software tools such as Pathway Studio, MetaCore, or Ingenuity Pathway Analysis (IPA), etc. Compared to these commercial tools, the pathway analysis module was specifically developed for metabolomics studies. It uses high-quality KEGG metabolic pathways as the backend knowledgebase. This module integrates many well-established (i.e. univariate analysis, over-representation analysis) methods, as well as novel algorithms and concepts (i.e. Global Test, GlobalAncova, network topology analysis) into pathway analysis. Another feature is a Google-Map style interactive visualization system to deliver the analysis results in an intuitive manner.

2 Data Input

The Pathway Analysis module accepts either a list of compound labels (common names, HMDB IDs or KEGG IDs) with one compound per row, or a compound concentration table with samples in rows and compounds in columns. The second column must be phenotype labels (binary, multi-group, or continuous). The table is uploaded as comma separated values (.csv).

3 Compound Name Matching

The first step is to standardize the compound labels used in user uploaded data. This is a necessary step since these compounds will be subsequently compared with compounds contained in the pathway library. There are three outcomes from the step - exact match, approximate match (for common names only), and no match. Users should click the textbfView button from the approximate matched results to manually select the correct one. Compounds without match will be excluded from the subsequently pathway analysis.

Table 1 shows the conversion results. Note: *1* indicates exact match, *2* indicates approximate match, and *0* indicates no match. A text file contain the result can be found the downloaded file *name_ map.csv*

Table 1: Result from Compound

	Query	Match	HMDB	PubChem	KEGG	SMILES
1	HMDB0000763	5-Hydroxyindoleacetic acid	HMDB0000763	1826	C05635	C1=CC2=C(C=C1O)C(=CN2)CC(=O)O
2	HMDB0000023	(S)-3-Hydroxyisobutyric acid	HMDB0000023	87	C06001	C[C@@H](CO)C(=O)O
3	HMDB0000355	3-Hydroxymethylglutaric acid	HMDB0000355	1662	C03761	CC(CC(=O)O)(CC(=O)O)O
4	HMDB0002064	N-Acetylputrescine	HMDB0002064	122356	C02714	CC(=O)NCCCCN
5	HMDB0000191	L-Aspartic acid	HMDB0000191	5960	C00049	C([C@@H](C(=O)O)N)C(=O)O
6	HMDB0000755	Hydroxyphenyllactic acid	HMDB0000755	9378	C03672	C1=CC(=CC=C1CC(C(=O)O)O)O
7	HMDB0000033	Carnosine	HMDB0000033	439224	C00386	C1=C(NC=N1)C[C@@H](C(=O)O)NC(=O
8	HMDB0000254	Succinic acid	HMDB0000254	1110	C00042	C(CC(=O)O)C(=O)O
9	HMDB0001401	Glucose 6-phosphate	HMDB0001401	5958	C00092	C([C@@H]1[C@H]([C@@H]([C@H](C(O1)O

10	HMDB0000252	Sphingosine	HMDB0000252	5353955	C00319	CCCCCCCCCCCCC/C=C/C(C(CO)N)O
11	HMDB0000099	L-Cystathionine	HMDB0000099	439258	C02291	C(CSC[C@@H](C(=O)O)N)[C@@H](C(=O)
12	HMDB0000269	Sphinganine	HMDB0000269	91486	C00836	CCCCCCCCCCCCCCCC[C@H]([C@H](CO)
13	HMDB0011737	gamma-Glutamylglutamic acid	HMDB0011737	92865	C05282	C(CC(=O)N[C@@H](CCC(=O)O)C(=O)O
14	HMDB0000661	Glutaric acid	HMDB0000661	743	C00489	C(CC(=O)O)CC(=O)O
15	HMDB0000965	Hypotaurine	HMDB0000965	107812	C00519	C(CS(=O)O)N
16	HMDB0000211	myo-Inositol	HMDB0000211		C00137	O[C@H]1[C@H](O)[C@@H](O)[C@H](O)[C@
17	HMDB0000226	Orotic acid	HMDB0000226	967	C00295	C1=C(NC(=O)NC1=O)C(=O)O
18	HMDB0003229	Palmitoleic acid	HMDB0003229	5312427	C08362	CCCCCC/C=C\CCCCCCCC(=O)O
19	HMDB0000162	L-Proline	HMDB0000162	145742	C00148	C1C[C@H](NC1)C(=O)O
20	HMDB0000190	L-Lactic acid	HMDB0000190	61503	C00186	C[C@@H](C(=O)O)O
21	HMDB0000251	Taurine	HMDB0000251	1123	C00245	C(CS(=O)(=O)O)N
22	HMDB0000806	Myristic acid	HMDB0000806	11005	C06424	CCCCCCCCCCCCCC(=O)O
23	HMDB0000296	Uridine	HMDB0000296	6029	C00299	C1=CN(C(=O)NC1=O)[C@H]2[C@@H]([C
24	HMDB0000148	L-Glutamic acid	HMDB0000148	33032	C00025	C(CC(=O)O)[C@@H](C(=O)O)N
25	HMDB0000641	L-Glutamine	HMDB0000641	5961	C00064	C(CC(=O)N)[C@@H](C(=O)O)N
26	HMDB0000122	D-Glucose	HMDB0000122	5793	C00221	C([C@@H]1[C@H]([C@@H]([C@H](C(O1)O
27	HMDB0000574	L-Cysteine	HMDB0000574	5862	C00097	C([C@@H](C(=O)O)N)S
28	HMDB0000169	D-Mannose	HMDB0000169	18950	C00936	C([C@@H]1[C@H]([C@@H]([C@@H](C(O1)
29	HMDB0000691	Malonic acid	HMDB0000691	867	C04025	C(C(=O)O)C(=O)O
30	HMDB0000300	Uracil	HMDB0000300	1174	C00106	C1=CNC(=O)NC1=O
31	HMDB0000235	Thiamine	HMDB0000235	1130	C00378	CC1=C(SC=[N+]1CC2=CN=C(N=C2N)C
32	HMDB0000258	Sucrose	HMDB0000258	5988	C00089	C([C@@H]1[C@H]([C@@H]([C@H]([C@H](C
33	HMDB0000267	Pyroglutamic acid	HMDB0000267	7405	C01879	C1CC(=O)N[C@@H]1C(=O)O
34	HMDB0000224	O-Phosphoethanolamine	HMDB0000224	1015	C00346	C(COP(=O)(O)O)N
35	HMDB0011741	gamma-Glutamyltyrosine	HMDB0011741	94340		C1=CC(=CC=C1C[C@@H](C(=O)O)NC(=
36	HMDB0000766	N-Acetyl-L-alanine	HMDB0000766	88064		C[C@@H](C(=O)O)NC(=O)C
37	HMDB0002166	(S)-b-aminoisobutyric acid	HMDB0002166	439434	C03284	C[C@@H](CN)C(=O)O
38	HMDB0000289	Uric acid	HMDB0000289	1175	C00366	C12=C(NC(=O)N1)NC(=O)NC2=O
39	HMDB0000707	4-Hydroxyphenylpyruvic acid	HMDB0000707	979	C01179	C1=CC(=CC=C1CC(=O)C(=O)O)O
40	HMDB0002817	N-Acetylglucosamine 6-phosphate	HMDB0002817	439219		CC(=O)N[C@@H]1[C@H]([C@@H]([C@H](C
41	HMDB0029151	gamma-Glutamylhistidine	HMDB0029151	7017195		C1=C(NC=N1)C[C@@H](C(=O)O)NC(=O
42	HMDB0001123	2-Aminobenzoic acid	HMDB0001123	227	C00108	C1=CC=C(C(=C1)C(=O)O)N
43	HMDB0011171	gamma-Glutamylleucine	HMDB0011171	4524287		CC(C)C[C@@H](C(=O)O)NC(=O)CC[C@@
44	HMDB0000754	3-Hydroxyisovaleric acid	HMDB0000754	69362	C20827	CC(C)(CC(=O)O)O
45	HMDB0000209	Phenylacetic acid	HMDB0000209	999	C07086	C1=CC=C(C=C1)CC(=O)O
46	HMDB0000714	Hippuric acid	HMDB0000714	464	C01586	C1=CC=C(C=C1)C(=O)NCC(=O)O
47	HMDB0000450	5-Hydroxylysine	HMDB0000450	4433	C16741	C(C[C@@H](C(=O)O)N)[C@H](CN)O
48	HMDB0004041	L-Allothreonine	HMDB0004041	99289	C05519	C[C@@H]([C@@H](C(=O)O)N)O
49	HMDB0003357	N-Acetylornithine	HMDB0003357	439232	C00437	CC(=O)N[C@@H](CCCN)C(=O)O
50	HMDB0000299	Xanthosine	HMDB0000299	64959	C01762	C1=NC2=C(N1[C@H]3[C@@H]([C@@H]([C
51	HMDB0003355	5-Aminopentanoic acid	HMDB0003355	138	C00431	C(CCN)CC(=O)O
52	HMDB0028854	Glycyl-Valine	HMDB0028854	97417		CC(C)C(NC(=O)CN)C(O)=O
53	HMDB0000594	gamma-Glutamylphenylalanine	HMDB0000594	111299		C1=CC=C(C=C1)C[C@@H](C(=O)O)NC(
54	HMDB0004666	2-Arachidonylglycerol	HMDB0004666	5282280	C13856	CCCCC/C=C\C/C=C\C/C=C\C/C=C\C
55	HMDB0009815	PI(18:0/20:4(5Z,8Z,11Z,14Z))	HMDB0009815		C00626	[H][C@@](COC(=O)CCCCCCCCCCCCCCC
56	HMDB0010734	(R)-3-Hydroxy-hexadecanoic acid	HMDB0010734	15569776		CCCCCCCCCCCCC[C@H](CC(=O)O)O
57	HMDB0002203	3-Hydroxycapric acid	HMDB0002203	26612		CCCCCCCC(CC(=O)O)O
58	HMDB0002039	2-Pyrrolidinone	HMDB0002039	12025	C11118	C1CC(=O)NC1
59	HMDB0011172	gamma-Glutamylvaline	HMDB0011172	7015683		CC(C)[C@@H](C(=O)O)NC(=O)CC[C@@H
60	HMDB0000409	(5R)-5-Hydroxyhexanoic acid	HMDB0000409	7131043		C[C@H](CCCC(=O)O)O
61	HMDB0006695	Prolylhydroxyproline	HMDB0006695	11902892		C1C[C@H]([NH2+]C1)C(=O)N2C[C@@H](C
62	HMDB0033724	Undecylenic acid	HMDB0033724	5634	C13910	C=CCCCCCCCCC(=O)O
63	HMDB0002000	Myristoleic acid	HMDB0002000	5281119	C08322	CCCC/C=C\CCCCCCCC(=O)O
64	HMDB0002024	Imidazoleacetic acid	HMDB0002024	96215	C02835	C1=C(NC=N1)CC(=O)O
65	HMDB0003312	Daidzein	HMDB0003312	5281708	C10208	C1=CC(=CC=C1C2=COC3=C(C2=O)C=
66	HMDB0000759	Glycylleucine	HMDB0000759	92843	C02155	CC(C)C[C@@H](C(=O)O)NC(=O)CN
67	HMDB0062557	N-Acetylthreonine	HMDB0062557	152204		C[C@@H](O)[C@H](NC(C)=O)C(O)=O
68	HMDB0029160	gamma-Glutamyltryptophan	HMDB0029160	3989307		C1=CC=C2C(=C1)C(=CN2)C[C@@H](C(=
69	HMDB0000855	Nicotinamide riboside	HMDB0000855	439924	C03150	C1=CC(=C[N+](=C1)[C@H]2[C@@H]([C@
70	HMDB0006809	Nicotinate D-ribonucleoside	HMDB0006809	161234	C05841	C1=CC(=C[N+](=C1)[C@H]2[C@@H]([C@
71	HMDB0094701	N-Acetylproline	HMDB0094701	66141		[H][C@]1(CCCN1C(C)=O)C(O)=O
72	HMDB0013678	4-Hydroxyhippuric acid	HMDB0013678	151012		C1=CC(=CC=C1C(=O)NCC(=O)O)O
73	HMDB0000897	7-Methylguanine	HMDB0000897	11361	C02242	CN1C=NC2=C1C(=O)N=C(N2)N
74	HMDB0240577	NA	NA	NA	NA	NA
75	HMDB0011170	gamma-Glutamylisoleucine	HMDB0011170	22885096		CC[C@H](C)[C@@H](C(=O)O)NC(=O)CC
76	HMDB0000710	4-Hydroxybutyric acid	HMDB0000710	10413	C01991	C(CC(=O)O)CO
77	HMDB0000378	2-Methylbutyroylcarnitine	HMDB0000378	6426901		CCC(C)C(=O)OC(CC(=O)[O-])C[N+](C)(
78	HMDB0000701	Hexanoylglycine	HMDB0000701	99463		CCCCCC(=O)NCC(=O)O
79	HMDB0002031	Ureidoisobutyric acid	HMDB0002031	160663	C05100	CC(CNC(=O)N)C(=O)O
80	HMDB0001068	D-Sedoheptulose 7-phosphate	HMDB0001068	22833559	C05382	C([C@@H]1[C@H]([C@H]([C@@H](C(O1)(C
81	HMDB0002035	4-Hydroxycinnamic acid	HMDB0002035	637542	C00811	C1=CC(=CC=C1/C=C/C(=O)O)O
82	HMDB0000206	N6-Acetyl-L-lysine	HMDB0000206	92832	C02727	CC(=O)NCCCC[C@@H](C(=O)O)N
83	HMDB0002931	N-Acetylserine	HMDB0002931	65249		CC(=O)N[C@@H](CO)C(=O)O
84	HMDB0006111	12-HETE	HMDB0006111	5312983	C14777	CCCCC/C=C\CC(/C=C\C=C/C/C=C/C
85	HMDB0010381	LysoPC(15:0)	HMDB0010381	24779458	C04230	CCCCCCCCCCCCCCC(=O)OC[C@H](CO
86	HMDB0003320	Indole-3-carboxylic acid	HMDB0003320	69867	C19837	C1=CC=C2C(=C1)C(=CN2)C(=O)O
87	HMDB0000522	3-Methylglutaconic acid	HMDB0000522	1551553		C/C(=C\C(=O)O)/CC(=O)O
88	HMDB0001212	Hydantoin-5-propionic acid	HMDB0001212	782	C05565	C(CC(=O)O)C1C(=O)NC(=O)N1
89	HMDB0001344	NA	NA	NA	NA	NA
90	HMDB0003219	Sedoheptulose	HMDB0003219	439645	C08355	C(C(C(C(C(=O)CO)O)O)O)O)O
91	HMDB0029131	Valyl-Leucine	HMDB0029131	6993118		CC(C)CC(NC(=O)C(N)C(C)C)C(O)=O
92	HMDB0002271	Imidazolepropionic acid	HMDB0002271	70630	C20522	C1=C(NC=N1)CCC(=O)O
93	HMDB0012881	Acetylcarnosine	HMDB0012881	9903482		CC(=O)NCCC(=O)N[C@H](CC1=CN=CN
94	HMDB0004983	Dimethyl sulfone	HMDB0004983	6213	C11142	CS(=O)(=O)C
95	HMDB0041646	NA	NA	NA	NA	NA
96	HMDB0000621	D-Ribulose	HMDB0000621	151261	C00309	C([C@H]([C@H](C(=O)CO)O)O)O

97	HMDB0061652	3-hydroxyhexanoic acid	HMDB0061652	11829482		CCC[C@H](O)CC(O)=O
98	HMDB0240296	NA	NA	NA	NA	NA
99	HMDB0060014	NA	NA	NA	NA	NA
100	HMDB10737	(R)-3-Hydroxy-Octadecanoic acid	HMDB0010737	5312838		CCCCCCCCCCCCCCCC[C@H](CC(=O)O)O
101	HMDB0007969	PC(16:0/16:1(9Z))	HMDB0007969	6443788	C00157	CCCCCCCCCCCCCCCC(=O)OC[C@H](C
102	HMDB0008046	PC(18:0/20:3(5Z,8Z,11Z))	HMDB0008046	24778855	C00157	CCCCCCCCCCCCCCCCCC(=O)OC[C@H
103	HMDB0240328	NA	NA	NA	NA	NA
104	HMDB0240644	NA	NA	NA	NA	NA
105	HMDB0013676	2,6-Dihydroxybenzoic acid	HMDB0013676	9338	C21298	C1=CC(=C(C(=C1)O)C(=O)O)O
106	HMDB0240388	NA	NA	NA	NA	NA
107	HMDB0011176	L-phenylalanyl-L-hydroxyproline	HMDB0011176	53480675		C1C(CN(C1C(=O)O)C(=O)C(CC2=CC=C
108	HMDB0028930	Leucyl-Hydroxyproline	HMDB0028930	20847829		CC(C)CC(N)C(=O)N1CC(O)CC1C(O)=O
109	HMDB0028908	Isoleucyl-Hydroxyproline	HMDB0028908	61158802		CCC(C)C(N)C(=O)N1CC(O)CC1C(O)=O
110	HMDB0037115	NA	NA	NA	NA	NA

4 Pathway Analysis

In this step, users are asked to select a pathway library, as well as specify the algorithms for pathway enrichment analysis and pathway topology analysis.

4.1 Pathway Library

There are 15 pathway libraries currently supported, with a total of 1173 pathways :

- Homo sapiens (human) [80]
- Mus musculus (mouse) [82]
- Rattus norvegicus (rat) [81]
- Bos taurus (cow) [81]
- Danio rerio (zebrafish) [81]
- Drosophila melanogaster (fruit fly) [79]
- Caenorhabditis elegans (nematode) [78]
- Saccharomyces cerevisiae (yeast) [65]
- Oryza sativa japonica (Japanese rice) [83]
- Arabidopsis thaliana (thale cress) [87]
- Escherichia coli K-12 MG1655 [87]
- Bacillus subtilis [80]
- Pseudomonas putida KT2440 [89]
- Staphylococcus aureus N315 (MRSA/VSSA)[73]
- Thermotoga maritima [57]

Your selected pathway library code is **mmu** (KEGG organisms abbreviation).

4.2 Over Representation Analysis

Over-representation analysis tests if a particular group of compounds is represented more than expected by chance within the user uploaded compound list. In the context of pathway analysis, we are testing if compounds involved in a particular pathway are enriched compared to random hits. MetPA offers two of the most commonly used methods for over-representation analysis:

- Fishers'Exact test
- Hypergeometric Test

Please note, MetPA uses one-tailed Fisher's exact test which will give essentially the same result as the result calculated by the hypergeometric test.

The selected over-representation analysis method is **Hypergeometric test**.

4.3 Pathway Topology Analysis

The structure of biological pathways represent our knowledge about the complex relationships among molecules within a cell or a living organism. However, most pathway analysis algorithms fail to take structural information into consideration when estimating which pathways are significantly changed under conditions of study. It is well-known that changes in more important positions of a network will trigger a more severe impact on the pathway than changes occurred in marginal or relatively isolated positions.

The pathway topology analysis uses two well-established node centrality measures to estimate node importance - **degree centrality** and **betweenness centrality**. Degree centrality is defined as the number of links occurred upon a node. For a directed graph there are two types of degree: in-degree for links come from other nodes, and out-degree for links initiated from the current node. Metabolic networks are directed graph. Here we only consider the out-degree for node importance measure. It is assumed that nodes upstream will have regulatory roles for the downstream nodes, not vice versa. The betweenness centrality measures the number of shortest paths going through the node. Since the metabolic network is directed, we use the relative betweenness centrality for a metabolite as the importance measure. The degree centrality measure focuses more on local connectivities, while the betweenness centrality measure focuses more on global network topology. For more detailed discussions on various graph-based methods for analyzing biological networks, please refer to the article by Tero Aittokallio, T. et al. [1]

Please note, for comparison among different pathways, the node importance values calculated from centrality measures are further normalized by the sum of the importance of the pathway. Therefore, the total/maximum importance of each pathway is 1; the importance measure of each metabolite node is actually the percentage w.r.t the total pathway importance, and the pathway impact value is the cumulative percentage from the matched metabolite nodes.

Your selected node importance measure for topological analysis is **out degree centrality**.

5 Pathway Analysis Result

The results from pathway analysis are presented graphically as well as in a detailed table.

A Google-map style interactive visualization system was implemented to facilitate data exploration. The graphical output contains three levels of view: **metabolome view**, **pathway view**, and **compound view**. Only the metabolome view is shown below. Pathway views and compound views are generated dynamically based on your interactions with the visualization system. They are available in your downloaded files.

[1] Tero Aittokallio and Benno Schwikowski. *Graph-based methods for analyzing networks in cell biology*, Briefings in Bioinformatics 2006 7(3):243-255

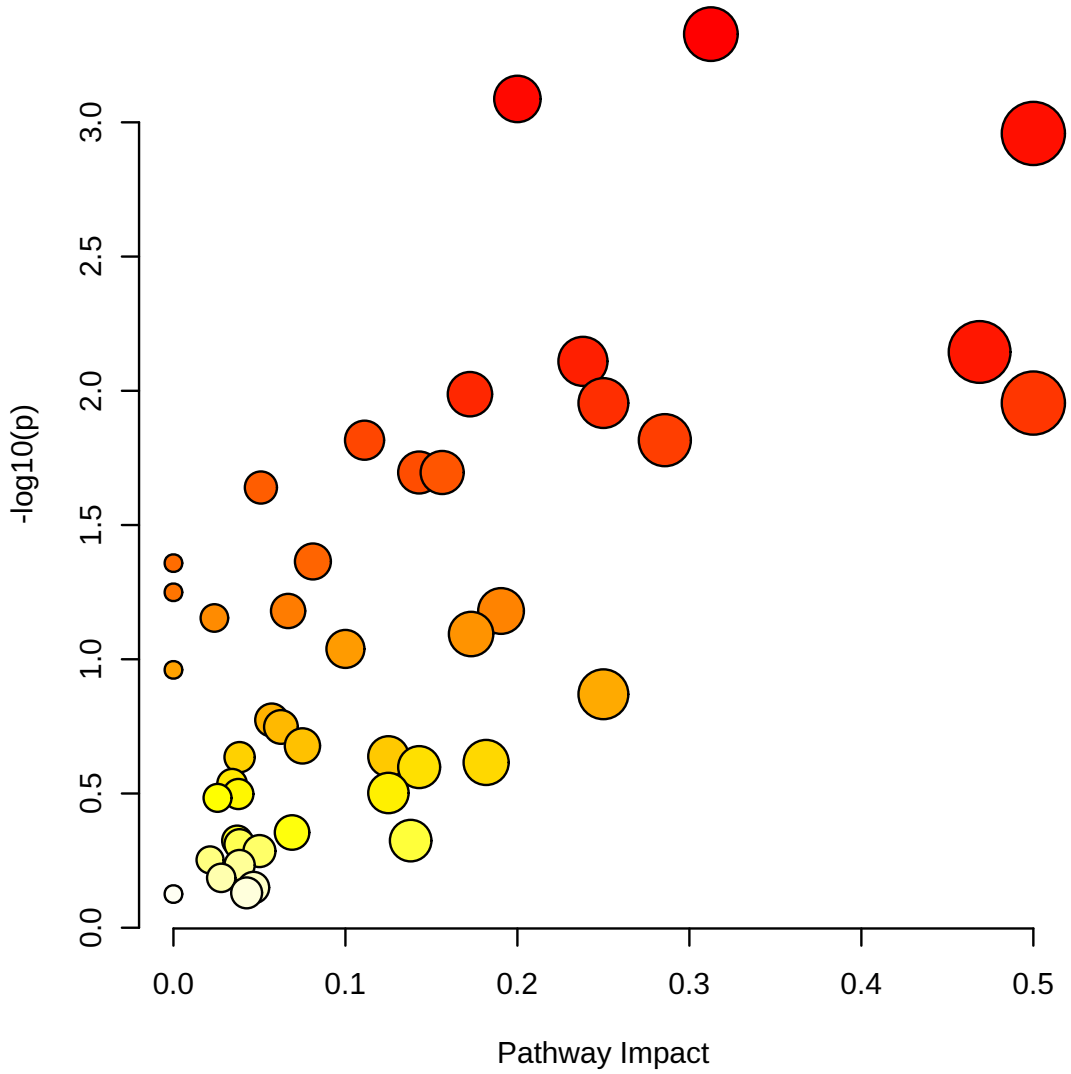

Figure 1: Summary of Pathway Analysis

The table below shows the detailed results from the pathway analysis. Since we are testing many pathways at the same time, the statistical p values from enrichment analysis are further adjusted for multiple testings. In particular, the **Total** is the total number of compounds in the pathway; the **Hits** is the actually matched number from the user uploaded data; the **Raw p** is the original p value calculated from the enrichment analysis; the **Holm p** is the p value adjusted by Holm-Bonferroni method; the **FDR p** is the p value adjusted using False Discovery Rate; the **Impact** is the pathway impact value calculated from pathway topology analysis.

Table 2: Result from Pathway Analysis

	Total	Expected	Hits	Raw p	-log10(p)	Holm adjust	FDR	Impact
Arginine biosynthesis	14	0.40	4	4.69E-04	3.33E+00	3.94E-02	3.08E-02	0.31
Histidine metabolism	16	0.46	4	8.18E-04	3.09E+00	6.79E-02	3.08E-02	0.20
Taurine and hypotaurine metabolism	8	0.23	3	1.10E-03	2.96E+00	9.02E-02	3.08E-02	0.50
Alanine, aspartate and glutamate metabolism	28	0.80	4	7.17E-03	2.14E+00	5.81E-01	1.17E-01	0.47
Nicotinate and nicotinamide metabolism	15	0.43	3	7.77E-03	2.11E+00	6.21E-01	1.17E-01	0.24
Aminoacyl-tRNA biosynthesis	48	1.37	5	1.03E-02	1.99E+00	8.13E-01	1.17E-01	0.17
Nitrogen metabolism	6	0.17	2	1.11E-02	1.95E+00	8.67E-01	1.17E-01	0.25
D-Glutamine and D-glutamate metabolism	6	0.17	2	1.11E-02	1.95E+00	8.67E-01	1.17E-01	0.50
Thiamine metabolism	7	0.20	2	1.53E-02	1.82E+00	1.00E+00	1.28E-01	0.29
Pantothenate and CoA biosynthesis	19	0.54	3	1.53E-02	1.82E+00	1.00E+00	1.28E-01	0.11
beta-Alanine metabolism	21	0.60	3	2.02E-02	1.70E+00	1.00E+00	1.41E-01	0.14
Sphingolipid metabolism	21	0.60	3	2.02E-02	1.70E+00	1.00E+00	1.41E-01	0.16
Pyrimidine metabolism	39	1.11	4	2.29E-02	1.64E+00	1.00E+00	1.48E-01	0.05
Glutathione metabolism	28	0.80	3	4.32E-02	1.36E+00	1.00E+00	2.46E-01	0.08
Phenylalanine metabolism	12	0.34	2	4.38E-02	1.36E+00	1.00E+00	2.46E-01	0.00
Neomycin, kanamycin and gentamicin biosynthesis	2	0.06	1	5.63E-02	1.25E+00	1.00E+00	2.96E-01	0.00
Butanoate metabolism	15	0.43	2	6.61E-02	1.18E+00	1.00E+00	3.08E-01	0.07
Starch and sucrose metabolism	15	0.43	2	6.61E-02	1.18E+00	1.00E+00	3.08E-01	0.19
Glycine, serine and threonine metabolism	34	0.97	3	7.02E-02	1.15E+00	1.00E+00	3.10E-01	0.02
Glycerophospholipid metabolism	36	1.03	3	8.05E-02	1.09E+00	1.00E+00	3.38E-01	0.17
Arginine and proline metabolism	38	1.08	3	9.15E-02	1.04E+00	1.00E+00	3.66E-01	0.10
Phenylalanine, tyrosine and tryptophan biosynthesis	4	0.11	1	1.10E-01	9.61E-01	1.00E+00	4.18E-01	0.00
Linoleic acid metabolism	5	0.14	1	1.35E-01	8.70E-01	1.00E+00	4.93E-01	0.25
Glycolysis / Gluconeogenesis	26	0.74	2	1.68E-01	7.74E-01	1.00E+00	5.89E-01	0.06
Galactose metabolism	27	0.77	2	1.79E-01	7.48E-01	1.00E+00	6.00E-01	0.06
Inositol phosphate metabolism	30	0.86	2	2.10E-01	6.77E-01	1.00E+00	6.79E-01	0.07
Ubiquinone and other terpenoid-quinone biosynthesis	9	0.26	1	2.30E-01	6.38E-01	1.00E+00	6.95E-01	0.12
Glyoxylate and dicarboxylate metabolism	32	0.91	2	2.32E-01	6.35E-01	1.00E+00	6.95E-01	0.04
Cysteine and methionine metabolism	33	0.94	2	2.42E-01	6.16E-01	1.00E+00	7.02E-01	0.18
Ascorbate and aldarate metabolism	10	0.29	1	2.52E-01	5.98E-01	1.00E+00	7.06E-01	0.14
Purine metabolism	66	1.88	3	2.91E-01	5.37E-01	1.00E+00	7.88E-01	0.03
alpha-Linolenic acid metabolism	13	0.37	1	3.15E-01	5.02E-01	1.00E+00	8.09E-01	0.12
Valine, leucine and isoleucine degradation	40	1.14	2	3.18E-01	4.98E-01	1.00E+00	8.09E-01	0.04
Tryptophan metabolism	41	1.17	2	3.28E-01	4.84E-01	1.00E+00	8.11E-01	0.03
Citrate cycle (TCA cycle)	20	0.57	1	4.42E-01	3.55E-01	1.00E+00	1.00E+00	0.07
Pyruvate metabolism	22	0.63	1	4.74E-01	3.25E-01	1.00E+00	1.00E+00	0.04
Pentose phosphate pathway	22	0.63	1	4.74E-01	3.25E-01	1.00E+00	1.00E+00	0.14
Propanoate metabolism	23	0.66	1	4.89E-01	3.11E-01	1.00E+00	1.00E+00	0.04
Lysine degradation	25	0.71	1	5.18E-01	2.86E-01	1.00E+00	1.00E+00	0.05
Phosphatidylinositol signaling system	28	0.80	1	5.59E-01	2.53E-01	1.00E+00	1.00E+00	0.02
Porphyrin and chlorophyll metabolism	30	0.86	1	5.84E-01	2.33E-01	1.00E+00	1.00E+00	0.04
Arachidonic acid metabolism	36	1.03	1	6.52E-01	1.86E-01	1.00E+00	1.00E+00	0.03
Tyrosine metabolism	42	1.20	1	7.09E-01	1.49E-01	1.00E+00	1.00E+00	0.05
Primary bile acid biosynthesis	46	1.31	1	7.42E-01	1.30E-01	1.00E+00	1.00E+00	0.04
Fatty acid biosynthesis	47	1.34	1	7.49E-01	1.25E-01	1.00E+00	1.00E+00	0.00

6 Appendix: R Command History

```
 [1] "mSet<-InitDataObjects(\"conc\", \"pathora\", FALSE)"
 [2] "cmpd.vec<-c(\"HMDB0000763\",\"HMDB0000023\",\"HMDB0000355\",\"HMDB0002064\",\"HMDB0000191\",\"
 [3] "mSet<-Setup.MapData(mSet, cmpd.vec);"
 [4] "mSet<-CrossReferencing(mSet, \"hmdb\");"
 [5] "mSet<-CreateMappingResultTable(mSet)"
 [6] "mSet<-PerformDetailMatch(mSet, \"HMDB0240577\");"
 [7] "mSet<-GetCandidateList(mSet);"
 [8] "mSet<-SetKEGG.PathLib(mSet, \"mmu\", \"current\")"
 [9] "mSet<-SetMetabolomeFilter(mSet, F);"
[10] "mSet<-CalculateOraScore(mSet, \"rbc\", \"hyperg\")"
[11] "mSet<-PlotPathSummary(mSet, F, \"path_view_0_\", \"png\", 72, width=NA, NA, NA )"
[12] "mSet<-SetKEGG.PathLib(mSet, \"mmu\", \"current\")"
[13] "mSet<-SetMetabolomeFilter(mSet, F);"
[14] "mSet<-CalculateOraScore(mSet, \"dgr\", \"hyperg\")"
[15] "mSet<-PlotPathSummary(mSet, F, \"path_view_1_\", \"png\", 72, width=NA, NA, NA )"
[16] "mSet<-SetKEGG.PathLib(mSet, \"mmu\", \"current\")"
[17] "mSet<-SetMetabolomeFilter(mSet, F);"
[18] "mSet<-CalculateOraScore(mSet, \"rbc\", \"hyperg\")"
[19] "mSet<-PlotPathSummary(mSet, F, \"path_view_2_\", \"png\", 72, width=NA, NA, NA )"
[20] "mSet<-SetKEGG.PathLib(mSet, \"mmu\", \"current\")"
[21] "mSet<-SetMetabolomeFilter(mSet, F);"
[22] "mSet<-CalculateOraScore(mSet, \"dgr\", \"hyperg\")"
[23] "mSet<-PlotPathSummary(mSet, F, \"path_view_3_\", \"png\", 72, width=NA, NA, NA )"
[24] "mSet<-SaveTransformedData(mSet)"
[25] "mSet<-PreparePDFReport(mSet, \"guest17886441071731322890\")\n"
```

The report was generated on Fri Jun 9 10:51:41 2023 with R version 4.2.2 (2022-10-31), OS system: Linux, version: -Ubuntu SMP Wed Feb 22 14:14:39 UTC 2023 .