Metabolomic Data Analysis with MetaboAnalyst 4.0

Name: guest4061267642427780 94

July 28, 2020

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*

T

	Query	Match	HMDB	PubChem	KEGG	SMILES
1	Pyruvic acid	Pyruvic acid	HMDB0000243	1060	C00022	CC(=O)C
2	L-Alanine	L-Alanine	HMDB0000161	5950	C00041	C[C@@H]
3	L-Lactic acid	L-Lactic acid	HMDB0000190	61503	C00256	C[C@@H]
4	2-Ketobutyric acid	2-Ketobutyric acid	HMDB0000005	58	C00109	CCC(=O)
5	Acetoacetic acid	Acetoacetic acid	HMDB0000060	96	C00164	CC(=O)C
6	3-Hydroxybutyric acid	3-Hydroxybutyric acid	HMDB0000357	441	C01089	CC(CC(=
7	L-Serine	L-Serine	HMDB0000187	5951	C00065	C([C@@H
8	Glyceric acid	Glyceric acid	HMDB0000139	439194	C00258	C([C@H](
9	Uracil	Uracil	HMDB0000300	1174	C00106	C1=CNC

10	L-Proline	L-Proline	HMDB0000162	145742	C00148	C1C[C@H
11	Maleic acid	Maleic acid	HMDB0000176	444972	C00122	C(=C\C(
12	Fumaric acid	Fumaric acid	HMDB0000134	444972	C00122	C(=C/C(
13	Alpha-ketoisovaleric acid	Alpha-ketoisovaleric acid	HMDB0000019	49	C00141	CC(C)C(=
14	Succinic acid	Succinic acid	HMDB0000254	1110	C00042	C(CC(=C
15	L-Homoserine	L-Homoserine	HMDB0000719	12647	C00263	C(CO)[C@
16	L-Cysteine	L-Cysteine	HMDB0000574	5862	C00097	C([C@@H
17	Nicotinic acid	Nicotinic acid	HMDB0001488	938	C00253	C1=CC(=
18	Taurine	Taurine	HMDB0000251	1123	C00245	C(CS(=O
19	Pyroglutamic acid	Pyroglutamic acid	HMDB0000267	7405	C01879	C1CC(=C
20	Pipecolic acid	Pipecolic acid	HMDB0000070	849	C00408	C1CCNC(
21	Citraconic acid	Citraconic acid	HMDB0000634	643798	C02226	C/C(=C/
22	Ketoleucine	Ketoleucine	HMDB0000695	70	C00233	CC(C)CC
23	Hydroxyproline	4-Hydroxyproline	HMDB0000725	5810	C01157	C1[C@H](
24	L-Asparagine	L-Asparagine	HMDB0000168	6267	C00152	C([C@@H
25	Ornithine	Ornithine	HMDB0000214	6262	C00077	C(C[C@@
26	L-Malic acid	L-Malic acid	HMDB0000156	222656	C00149	C([C@@H
27	Homocysteine	Homocysteine	HMDB0000742	778	C05330	C(CS)C(C
28	2-Aminobenzoic acid	2-Aminobenzoic acid	HMDB0001123	227	C00108	C1=CC=
29	Acetylphosphate	Acetylphosphate	HMDB0001494	186	C00227	CC(=O)C
30	Oxoglutaric acid	Oxoglutaric acid	HMDB0000208	51	C00026	C(CC(=C
31	L-Glutamine	L-Glutamine	HMDB0000641	5961	C00064	C(CC(=C
32	L-Lysine	L-Lysine	HMDB0000182	5962	C00047	C(CCN)C
33	2-Oxo-4-methylthiobutanoic acid	2-Oxo-4-methylthiobutanoic acid	HMDB0001553	473	C01180	CSCCC(=
34	L-Methionine	L-Methionine	HMDB0000696	6137	C00073	CSCC[C@
35	D-Ribose	D-Ribose	HMDB0000283	5779	C00121	C([C@@H
36	Xanthine	Xanthine	HMDB0000292	1188	C00385	C1=NC2=
37	p-Hydroxyphenylacetic acid	p-Hydroxyphenylacetic acid	HMDB0000020	127	C00642	C1=CC(=
38	2,3-Dihydroxybenzoic acid	2-Pyrocatechuic acid	HMDB0000397	19	C00196	C1=CC(=
39	L-Histidine	L-Histidine	HMDB0000177	6274	C00135	C1=C(NC
40	Indole-3-carboxylic acid	Indole-3-carboxylic acid	HMDB0003320	69867	C19837	C1=CC=
41	Phenylpyruvic acid	Phenylpyruvic acid	HMDB0000205	997	C00166	C1=CC=
42	L-Phenylalanine	L-Phenylalanine	HMDB0000159	6140	C00079	C1=CC=
43	Phenyllactic acid	Phenyllactic acid	HMDB0000779	3848	C01479	C1=CC=
44	Quinolinic acid	Quinolinic acid	HMDB0000232	1066	C03722	C1=CC(=
45	Phosphoenolpyruvic acid	Phosphoenolpyruvic acid	HMDB0000263	1005	C00074	C=C(C(=
46	Uric acid	Uric acid	HMDB0000289	1175	C00366	C12=C(N
47	Dihydroxyacetone phosphate	Dihydroxyacetone phosphate	HMDB0001473	668	C00111	C(C(=O)
48	Glyceraldehyde 3-phosphate	NA	NA	NA	NA	NA
49	Glycerol 3-phosphate	Glycerol 3-phosphate	HMDB0000126	439162	C00093	C([C@H](
50	cis-Aconitic acid	cis-Aconitic acid	HMDB0000072	643757	C00417	C(/C(=C
51	N-Acetylornithine	N-Acetylornithine	HMDB0003357	439232	C00437	CC(=O)N
52	L-Arginine	L-Arginine	HMDB0000517	6322	C00062	C(C[C@@
53	N-Acetyl-L-aspartic acid	N-Acetyl-L-aspartic acid	HMDB0000812	65065	C01042	CC(=O)N
54	Citrulline	Citrulline	HMDB0000904	9750	C00327	C(C[C@@
55	Ascorbic acid	Ascorbic acid	HMDB0000044	54670067	C01041	C([C@@H
56	Ureidosuccinic acid	Ureidosuccinic acid	HMDB0000828	93072	C00438	C([C@@H
57	Allantoic acid	Allantoic acid	HMDB0001209	203	C00499	C(C(=O)
58	2-Isopropylmalic acid	2-Isopropylmalic acid	HMDB0000402	5280523	C02504	CC(C)[C@
59	Gluconolactone	Gluconolactone	HMDB0000150	7027	C00198	C([C@@H
60	4-Hydroxyphenylpyruvic acid	4-Hydroxyphenylpyruvic acid	HMDB0000707	979	C01179	C1=CC(=
61	D-Glucose	D-Glucose	HMDB0000122	5793	C00031	C([C@@H
62	4-Pyridoxic acid	4-Pyridoxic acid	HMDB0000017	6723	C00847	CC1=NC
63	DL-O-Phosphoserine	DL-O-Phosphoserine	HMDB0001721	106	C01005	C(C(C(=
64	3-Phosphoglyceric acid	3-Phosphoglyceric acid	HMDB0000807	724	C00597	C(C(C(=
65	N-Alpha-acetyllysine	N-Alpha-acetyllysine	HMDB0000446	192590	C12989	CC(=O)N
66	Kynurenic acid	Kynurenic acid	HMDB0000715	3845	C01717	C1=CC=
67	N-Acetylglutamic acid	N-Acetylglutamic acid	HMDB0001138	185	C00624	CC(=O)N
68	Isocitrate	Isocitric acid	HMDB0000193	1198	C00311	C(C(C(C(
69	2-Keto-L-gluconate	2-Keto-L-gluconate	HMDB0011732	50	C15673	C(C(C(C(
70	D-Erythrose 4-phosphate	D-Erythrose 4-phosphate	HMDB0001321	122357	C00279	C([C@H](
71	L-Tryptophan	L-Tryptophan	HMDB0000929	6305	C00078	C1=CC=
72	Xanthurenic acid	Xanthurenic acid	HMDB0000881	5699	C02470	C1=CC2=
73	Glucaric acid	Glucaric acid	HMDB0000663	33037	C00818	[C@H]([C
74	Deoxyribose 1-phosphate	Deoxyribose 1-phosphate	HMDB0001351	439287	C00672	C1[C@@H
75	Pantothenic acid	Pantothenic acid	HMDB0000210	6613	C00864	CC(C)(C(
76	L-Cystathionine	L-Cystathionine	HMDB0000099	439258	C02291	C(CSC[C
77	Deoxyuridine	Deoxyuridine	HMDB0000012	13712	C00526	C1[C@@H
78	D-Ribulose 5-phosphate	D-Ribulose 5-phosphate	HMDB0000618	439184	C00199	C([C@H](
79	D-Ribose 5-phosphate	D-Ribose 5-phosphate	HMDB0001548	439167	C00117	C([C@@H
80	Uridine	Uridine	HMDB0000296	6029	C00299	C1=CN(C
81	Glucosamine 6-phosphate	Glucosamine 6-phosphate	HMDB0001254	439217	C00352	C([C@@H
82	D-Glucose 1,6-bisphosphate	Alpha-D-Glucose 1,6-bisphosphate	HMDB0003514	82400	C01231	C([C@@H
83	Fructose 6-phosphate	Fructose 6-phosphate	HMDB0000124	69507	C00085	C([C@H](
84	Inosine	Inosine	HMDB0000195	6021	C00294	C1=NC(=
85	6-Phosphogluconic acid	6-Phosphogluconic acid	HMDB0001316	91493	C00345	C([C@H](
86	Pantetheine	Pantetheine	HMDB0003426	479	C00831	CC(C)(C(
87	Xanthosine	Xanthosine	HMDB0000299	64959	C01762	C1=NC2=
88	5'-Methylthioadenosine	5'-Methylthioadenosine	HMDB0001173	439176	C00170	CSC[C@@
89	D-4'-Phosphopantothenate	D-4'-Phosphopantothenate	HMDB0001016	131	C03492	CC(C)(C(
90	dCMP	dCMP	HMDB0001202	13945	C00239	C1[C@@H
91	Glutathione	Glutathione	HMDB0062697	745	C00051	NC(CCC(
92	5-Thymidylic acid	5-Thymidylic acid	HMDB0001227	9700	C00364	CC1=CN
93	Cytidine monophosphate	Cytidine monophosphate	HMDB0000095	8117	C00055	C1=CN(C
94	Uridine 5'-monophosphate	Uridine 5'-monophosphate	HMDB0000288	6030	C00105	C1=CN(C
95	Cyclic AMP	Cyclic AMP	HMDB0000058	6076	C00575	C1[C@@H
96	Deoxyadenosine monophosphate	Deoxyadenosine monophosphate	HMDB0000905	12599	C00360	C1[C@@H

97	Fructose 1,6-bisphosphate	Fructose 1,6-bisphosphate	HMDB0001058	445557	C00354	C([C@@H
98	Sucrose	Sucrose	HMDB0000258	5988	C00089	C([C@@H
99	Adenosine monophosphate	Adenosine monophosphate	HMDB0000045	6083	C00020	C1=NC2=
100	Inosinic acid	Inosinic acid	HMDB0000175	8582	C00130	C1=NC(=
101	Guanosine monophosphate	Guanosine monophosphate	HMDB0001397	6804	C00144	C1=NC2=
102	Xanthosine 5'-phosphate	Xanthylic acid	HMDB0001554	73323	C00655	C1=NC2=
103	Sedoheptulose 1,7-bisphosphate	Sedoheptulose 1,7-bisphosphate	HMDB0060274	164735	C00447	O[C@H](C
104	Riboflavin	Riboflavin	HMDB0000244	493570	C00255	CC1=CC
105	S-Adenosylhomocysteine	S-Adenosylhomocysteine	HMDB0000939	439155	C00021	C1=NC2=
106	CDP	CDP	HMDB0001546	6132	C00112	C1=CN(C
107	Uridine 5'-diphosphate	Uridine 5'-diphosphate	HMDB0000295	6031	C00015	C1=CN(C
108	Trehalose 6-phosphate	Trehalose 6-phosphate	HMDB0001124	122336	C00689	C([C@@H
109	ADP	ADP	HMDB0001341	6022	C00008	C1=NC2=
110	Guanosine diphosphate	Guanosine diphosphate	HMDB0001201	8977	C00035	C1=NC2=
111	CDP-Ethanolamine	CDP-Ethanolamine	HMDB0001564	123727	C00570	C1=CN(C
112	Flavin Mononucleotide	Flavin Mononucleotide	HMDB0001520	643976	C00061	CC1=CC
113	5-Methyltetrahydrofolic acid	5-Methyltetrahydrofolic acid	HMDB0001396	439234	C00440	CN1C(CN
114	Uridine triphosphate	Uridine triphosphate	HMDB0000285	6133	C00075	C1=CN(C
115	Citicoline	Citicoline	HMDB0001413	13804	C00307	C[N+](C)
116	Uridine diphosphate glucose	Uridine diphosphate glucose	HMDB0000286	53477679	C00029	C1=CN(C
117	Uridine diphosphate glucuronic acid	Uridine diphosphate glucuronic acid	HMDB0000935	17473	C00167	C1=CN(C
118	ADP-glucose	ADP-glucose	HMDB0006557	16500	C00498	C1=NC2=
119	Uridine diphosphate-N-acetylglucosamine	Uridine diphosphate-N-acetylglucosamine	HMDB0000290	9547196	C00043	CC(=O)N
120	Glutathione disulfide	Oxidized glutathione	HMDB0003337	975	C00127	C(CC(=C
121	FAD	FAD	HMDB0001248	643975	C00016	CC1=CC

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 Pathway Enrichment Analysis

Pathway enrichment analysis usually refers to quantitative enrichment analysis directly using the compound concentration values, as compared to compound lists used by over-representation analysis. As a result, it is more sensitive and has the potential to identify **subtle but consistent** changes amongst compounds involved in the same biological pathway.

Many procedures have been developed in the last decade for quantitative enrichment analysis, the most famous being the Gene Set Enrichment Analysis. Many new and improved methods have been implemented since. The enrichment analysis is based on GlobalTest and GlobalAncova. Both methods support enrichment analysis with binary, multi-group, as well as continuous phenotypes. The p-values can be approximated based on the asymptotic distribution without using permutations which is computationally very intensive and is not suitable for web applications. Please note, when sample sizes are small, the approximated p values may be slightly less accurate compared to p values obtained by using a permutation-based method (for details, please refer to the paper by Goeman, J.J. et al. [1] and by

[1] Jelle J. Goeman and Peter Buhlmann. *Analyzing gene expression data in terms of gene sets: methodological issues*, Bioinformatics 2007 23(8):980-987

Hummel, M. et al. [2]) However, since our focus is to identify the most relevant pathways within the pathways in the library, we are more interested in the rank of the pathway, not its absolute p-value. Therefore, this disadvantage may be tolerated.

The selected pathway enrichment analysis method is **Globaltest**.

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. [3]

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 **relative betweenness 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.

[2]Manuela Hummel, Reinhard Meister and Ulrich Mansmann. *GlobalANCOVA: exploration and assessment of gene group effects*, Bioinformatics 2008 24(1):78-85

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

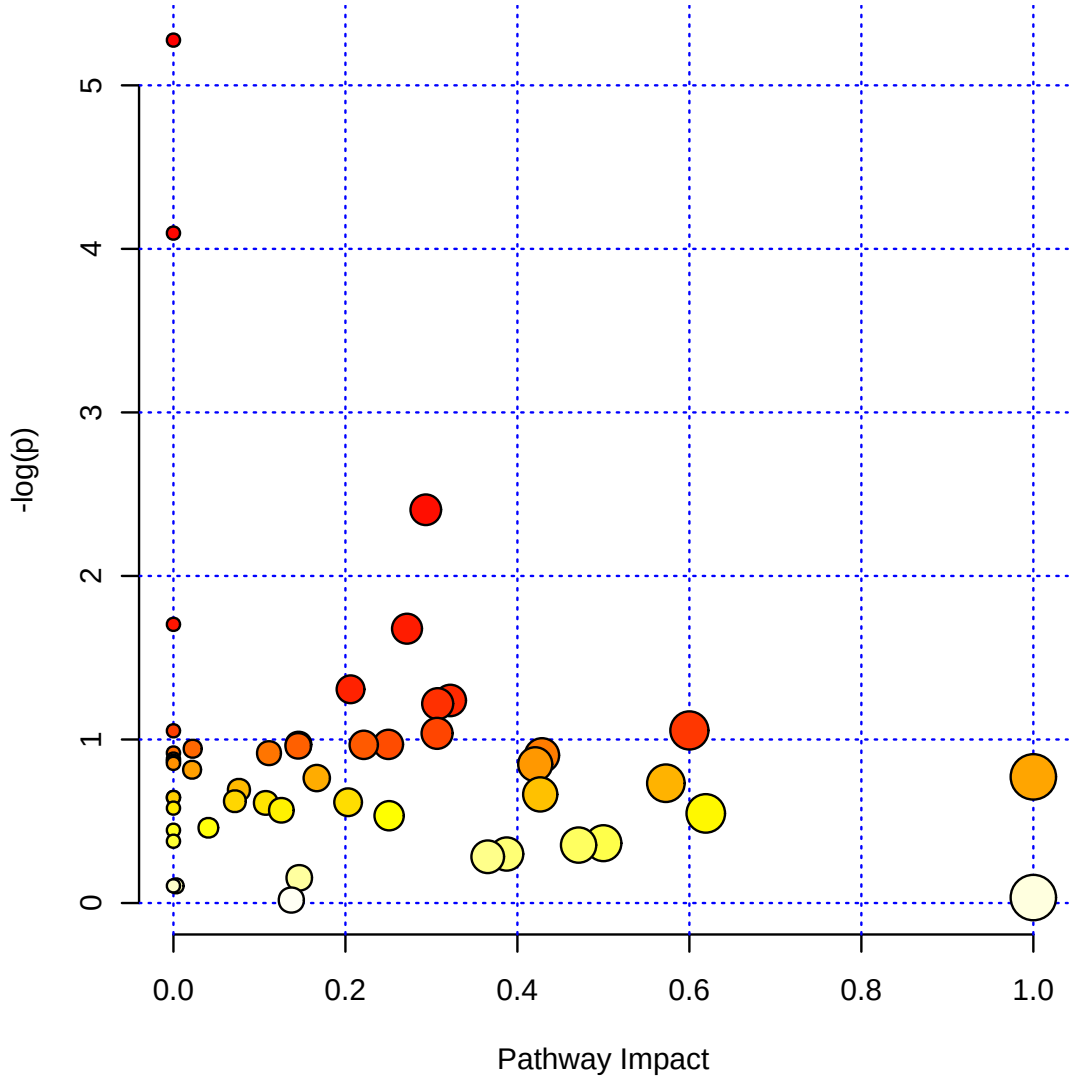

Figure 1: Summary of Pathway Analysis

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 Cmpd	Hits	Raw p	-log(p)	Holm adjust	FDR	Impact
Nicotinate and nicotinamide metabolism	15	2	5.11E-03	5.28E+00	2.56E-01	2.56E-01	0.00
Neomycin, kanamycin and gentamicin biosynthesis	2	1	1.66E-02	4.10E+00	8.15E-01	4.16E-01	0.00
Glutathione metabolism	28	5	9.03E-02	2.40E+00	1.00E+00	7.72E-01	0.29
Lysine degradation	25	2	1.82E-01	1.70E+00	1.00E+00	7.72E-01	0.00
Glycine, serine and threonine metabolism	34	7	1.87E-01	1.68E+00	1.00E+00	7.72E-01	0.27
Glycolysis / Gluconeogenesis	26	3	2.71E-01	1.31E+00	1.00E+00	7.72E-01	0.21
Pyruvate metabolism	22	6	2.90E-01	1.24E+00	1.00E+00	7.72E-01	0.32
Arginine and proline metabolism	38	5	2.96E-01	1.22E+00	1.00E+00	7.72E-01	0.31
Synthesis and degradation of ketone bodies	5	2	3.48E-01	1.06E+00	1.00E+00	7.72E-01	0.60
Thiamine metabolism	7	1	3.49E-01	1.05E+00	1.00E+00	7.72E-01	0.00
Citrate cycle (TCA cycle)	20	8	3.54E-01	1.04E+00	1.00E+00	7.72E-01	0.31
Ascorbate and aldarate metabolism	10	3	3.79E-01	9.70E-01	1.00E+00	7.72E-01	0.25
Glyoxylate and dicarboxylate metabolism	32	7	3.79E-01	9.70E-01	1.00E+00	7.72E-01	0.15
Histidine metabolism	16	1	3.80E-01	9.67E-01	1.00E+00	7.72E-01	0.22
Tryptophan metabolism	41	2	3.82E-01	9.61E-01	1.00E+00	7.72E-01	0.15
Primary bile acid biosynthesis	46	1	3.89E-01	9.43E-01	1.00E+00	7.72E-01	0.02
beta-Alanine metabolism	21	2	4.00E-01	9.17E-01	1.00E+00	7.72E-01	0.00
Butanoate metabolism	15	4	4.00E-01	9.16E-01	1.00E+00	7.72E-01	0.11
Taurine and hypotaurine metabolism	8	2	4.05E-01	9.03E-01	1.00E+00	7.72E-01	0.43
Nitrogen metabolism	6	1	4.15E-01	8.80E-01	1.00E+00	7.72E-01	0.00
Sphingolipid metabolism	21	1	4.20E-01	8.67E-01	1.00E+00	7.72E-01	0.00
Selenocompound metabolism	20	1	4.26E-01	8.54E-01	1.00E+00	7.72E-01	0.00
Starch and sucrose metabolism	15	5	4.28E-01	8.49E-01	1.00E+00	7.72E-01	0.42
Valine, leucine and isoleucine degradation	40	3	4.43E-01	8.15E-01	1.00E+00	7.72E-01	0.02
Riboflavin metabolism	4	3	4.62E-01	7.71E-01	1.00E+00	7.72E-01	1.00
Aminoacyl-tRNA biosynthesis	48	12	4.66E-01	7.64E-01	1.00E+00	7.72E-01	0.17
Cysteine and methionine metabolism	33	10	4.81E-01	7.33E-01	1.00E+00	7.72E-01	0.57
Galactose metabolism	27	3	5.01E-01	6.92E-01	1.00E+00	7.72E-01	0.08
Pentose phosphate pathway	22	8	5.15E-01	6.64E-01	1.00E+00	7.72E-01	0.43
Valine, leucine and isoleucine biosynthesis	8	3	5.24E-01	6.46E-01	1.00E+00	7.72E-01	0.00
Biotin metabolism	10	1	5.25E-01	6.45E-01	1.00E+00	7.72E-01	0.00
Pantothenate and CoA biosynthesis	19	6	5.37E-01	6.23E-01	1.00E+00	7.72E-01	0.07
Pentose and glucuronate interconversions	18	3	5.40E-01	6.16E-01	1.00E+00	7.72E-01	0.20
Tyrosine metabolism	42	5	5.42E-01	6.12E-01	1.00E+00	7.72E-01	0.11
D-Glutamine and D-glutamate metabolism	6	2	5.60E-01	5.80E-01	1.00E+00	7.72E-01	0.00
Amino sugar and nucleotide sugar metabolism	37	5	5.67E-01	5.67E-01	1.00E+00	7.72E-01	0.13
Phenylalanine metabolism	12	2	5.78E-01	5.48E-01	1.00E+00	7.72E-01	0.62
Alanine, aspartate and glutamate metabolism	28	10	5.86E-01	5.34E-01	1.00E+00	7.72E-01	0.25
Propanoate metabolism	23	2	6.31E-01	4.60E-01	1.00E+00	8.01E-01	0.04
Vitamin B6 metabolism	9	1	6.41E-01	4.45E-01	1.00E+00	8.01E-01	0.00
One carbon pool by folate	9	1	6.85E-01	3.78E-01	1.00E+00	8.16E-01	0.00
Phenylalanine, tyrosine and tryptophan biosynthesis	4	3	6.94E-01	3.66E-01	1.00E+00	8.16E-01	0.50
Pyrimidine metabolism	39	13	7.02E-01	3.54E-01	1.00E+00	8.16E-01	0.47
Purine metabolism	66	15	7.41E-01	3.00E-01	1.00E+00	8.37E-01	0.39
Arginine biosynthesis	14	8	7.54E-01	2.83E-01	1.00E+00	8.37E-01	0.37
Glycerophospholipid metabolism	36	4	8.57E-01	1.54E-01	1.00E+00	9.32E-01	0.15
Fructose and mannose metabolism	18	1	9.00E-01	1.05E-01	1.00E+00	9.38E-01	0.00
Inositol phosphate metabolism	30	1	9.00E-01	1.05E-01	1.00E+00	9.38E-01	0.00
Ubiquinone and other terpenoid-quinone biosynthesis	9	1	9.67E-01	3.35E-02	1.00E+00	9.82E-01	1.00
Glycerolipid metabolism	16	3	9.82E-01	1.81E-02	1.00E+00	9.82E-01	0.14

6 Appendix: R Command History

```
 [1] "mSet<-InitDataObjects(\"conc\", \"pathqea\", FALSE)"
 [2] "mSet<-Read.TextData(mSet, \"Replacing_with_your_file_path\", \"rowu\", \"disc\");"
 [3] "mSet<-CrossReferencing(mSet, \"name\");"
 [4] "mSet<-CreateMappingResultTable(mSet)"
 [5] "mSet<-PerformDetailMatch(mSet, \"Glyceraldehyde 3-phosphate\");"
 [6] "mSet<-GetCandidateList(mSet);"
 [7] "mSet<-SetCandidate(mSet, \"Glyceraldehyde 3-phosphate\", \"D-Glyceraldehyde 3-phosphate\");"
 [8] "mSet<-PerformDetailMatch(mSet, \"D-Glucose 1,6-bisphosphate\");"
 [9] "mSet<-GetCandidateList(mSet);"
[10] "mSet<-SetCandidate(mSet, \"D-Glucose 1,6-bisphosphate\", \"Alpha-D-Glucose 1,6-bisphosphate\")
[11] "mSet<-PerformDetailMatch(mSet, \"Glyceraldehyde 3-phosphate\");"
[12] "mSet<-GetCandidateList(mSet);"
[13] "mSet<-SetCandidate(mSet, \"Glyceraldehyde 3-phosphate\", \"D-Glyceraldehyde 3-phosphate\");"
[14] "mSet<-SanityCheckData(mSet)"
[15] "mSet<-ReplaceMin(mSet);"
[16] "mSet<-PreparePrenormData(mSet)"
[17] "mSet<-Normalization(mSet, \"NULL\", \"NULL\", \"NULL\", ratio=FALSE, ratioNum=20)"
[18] "mSet<-PlotNormSummary(mSet, \"norm_0_\", \"png\", 72, width=NA)"
[19] "mSet<-PlotSampleNormSummary(mSet, \"snorm_0_\", \"png\", 72, width=NA)"
[20] "mSet<-SetKEGG.PathLib(mSet, \"rno\", \"current\")"
[21] "mSet<-SetMetabolomeFilter(mSet, F);"
[22] "mSet<-CalculateQeaScore(mSet, \"rbc\", \"gt\")"
[23] "mSet<-PlotPathSummary(mSet, \"path_view_0_\", \"png\", 72, width=NA)"
[24] "mSet<-PlotKEGGPath(mSet, \"Riboflavin metabolism\",528, 480, \"png\", NULL)"
[25] "mSet<-RerenderMetPAGraph(mSet, \"zoom1595945164579.png\",528.0, 480.0, 100.0)"
[26] "mSet<-PlotKEGGPath(mSet, \"Ubiquinone and other terpenoid-quinone biosynthesis\",528, 480, \"pn
[27] "mSet<-PlotKEGGPath(mSet, \"Nicotinate and nicotinamide metabolism\",528, 480, \"png\", NULL)"
[28] "mSet<-SetKEGG.PathLib(mSet, \"mmu\", \"current\")"
[29] "mSet<-SetMetabolomeFilter(mSet, F);"
[30] "mSet<-CalculateQeaScore(mSet, \"rbc\", \"gt\")"
[31] "mSet<-PlotPathSummary(mSet, \"path_view_1_\", \"png\", 72, width=NA)"
[32] "mSet<-PlotKEGGPath(mSet, \"Riboflavin metabolism\",528, 480, \"png\", NULL)"
[33] "mSet<-RerenderMetPAGraph(mSet, \"zoom1595945381246.png\",528.0, 480.0, 100.0)"
[34] "mSet<-PlotKEGGPath(mSet, \"Ubiquinone and other terpenoid-quinone biosynthesis\",528, 480, \"pn
[35] "mSet<-PlotKEGGPath(mSet, \"Glutathione metabolism\",528, 480, \"png\", NULL)"
[36] "mSet<-PlotPathSummary(mSet, \"path_view_1_\", \"png\", 300, width=NA)"
[37] "mSet<-PlotKEGGPath(mSet, \"Riboflavin metabolism\",528, 480, \"png\", NULL)"
[38] "mSet<-PlotKEGGPath(mSet, \"Glutathione metabolism\",528, 480, \"png\", NULL)"
[39] "mSet<-SaveTransformedData(mSet)"
[40] "mSet<-PreparePDFReport(mSet, \"guest406126764242778094\")\n"
```

The report was generated on Tue Jul 28 10:12:16 2020 with R version 3.6.3 (2020-02-29).