Metabolomic Data Analysis with MetaboAnalyst 5.0

Name: guest5645338793786851237

May 21, 2021

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*

	Query	Match	HMDB	PubChem	KEGG	SMILES
1	HMDB0001858	p-Cresol	HMDB0001858	2879	C01468	CC1=CC=C(
2	HMDB0033895	Anisole	HMDB0033895	7519	C01403	COC1=CC=C
3	HMDB0003119	Benzyl alcohol	HMDB0003119	244	C03485	C1=CC=C(C
4	HMDB0002055	o-Cresol	HMDB0002055	335	C01542	CC1=CC=CC
5	HMDB0002048	m-Cresol	HMDB0002048	342	C01467	CC1=CC(=C
6	HMDB0163354	NA	NA	NA	NA	NA
7	HMDB0163353	NA	NA	NA	NA	NA
8	HMDB0163351	NA	NA	NA	NA	NA
9	HMDB0163350	NA	NA	NA	NA	NA

10	HMDB0163352	NA	NA	NA	NA	NA
11	HMDB0163702	NA	NA	NA	NA	NA
12	HMDB0179351	NA	NA	NA	NA	NA
13	HMDB0179346	NA	NA	NA	NA	NA
14	HMDB0179345	NA	NA	NA	NA	NA
15	HMDB0179344	NA	NA	NA	NA	NA
16	HMDB0179343	NA	NA	NA	NA	NA
17	HMDB0179342	NA	NA	NA	NA	NA
18	HMDB0179341	NA	NA	NA	NA	NA
19	HMDB0169695	NA	NA	NA	NA	NA
20	HMDB0169664	NA	NA	NA	NA	NA
21	HMDB0165583	NA	NA	NA	NA	NA
22	HMDB0165582	NA	NA	NA	NA	NA
23	HMDB0165581	NA	NA	NA	NA	NA
24	HMDB0165584	NA	NA	NA	NA	NA
25	HMDB0163703	NA	NA	NA	NA	NA
26	HMDB0000453	Delta-Hexanolactone	HMDB0000453	13204		CC1CCCC(=
27	HMDB0060476	epsilon-Caprolactone	HMDB0060476	10401	C01880	O=C1CCCC
28	HMDB0032490	Prenyl formate	HMDB0032490	110373		CC(=CCOC=
29	HMDB0039581	NA	NA	NA	NA	NA
30	HMDB0031158	2-Methyl-4-pentenoic acid	HMDB0031158	549519		CC(CC=C)C(
31	HMDB0031562	2-Methyl-3-pentenoic acid	HMDB0031562	71587183		C\C=C/C(C)
32	HMDB0031559	4-Methyl-2,3-pentanedione	HMDB0031559	24115		CC(C)C(=O)
33	HMDB0031491	2,3-Hexanedione	HMDB0031491	19707		CCCC(=O)C(
34	HMDB0031560	2-Methyl-2-pentenoic acid	HMDB0031560	6436344		CC/C=C(/C)
35	HMDB0031501	3-Hexenoic acid	HMDB0031501	5355152		CC/C=C\CC
36	HMDB0029487	Allyl propionate	HMDB0029487	61319		CCC(=O)OC
37	HMDB0033568	5-Oxohexanal	HMDB0033568	10953466		CC(=O)CCC
38	HMDB0040328	NA	NA	NA	NA	NA
39	HMDB0031492	3,4-Hexanedione	HMDB0031492	62539		CCC(=O)C(=
40	HMDB0031561	4-Methyl-2-pentenoic acid	HMDB0031561	12216200		CC(C)/C=C\
41	HMDB0010719	trans-Hex-2-enoic acid	HMDB0010719	5282707		CCC/C=C/C
42	HMDB0003843	Gamma-Caprolactone	HMDB0003843	12756		CCC1CCC(=
43	HMDB0240384	NA	NA	NA	NA	NA
44	HMDB0061929	NA	NA	NA	NA	NA
45	HMDB0061965	NA	NA	NA	NA	NA
46	HMDB0062135	NA	NA	NA	NA	NA
47	HMDB0161115	NA	NA	NA	NA	NA
48	HMDB0161116	NA	NA	NA	NA	NA
49	HMDB0163679	NA	NA	NA	NA	NA
50	HMDB0061652	3-hydroxyhexanoic acid	HMDB0061652	11829482		CCC[C@H](O
51	HMDB0171222	NA	NA	NA	NA	NA
52	HMDB0165042	NA	NA	NA	NA	NA
53	HMDB0165043	NA	NA	NA	NA	NA
54	HMDB0165041	NA	NA	NA	NA	NA
55	HMDB0163682	NA	NA	NA	NA	NA
56	HMDB0163680	NA	NA	NA	NA	NA
57	HMDB0163681	NA	NA	NA	NA	NA
58	HMDB0000317	2-Hydroxy-3-methylpentanoic acid	HMDB0000317	10796774		CC[C@@H](C
59	HMDB0000409	(5R)-5-Hydroxyhexanoic acid	HMDB0000409	7131043		C[C@H](CCC
60	HMDB0032158	(+/-)-1-Acetoxy-1-ethoxyethane	HMDB0032158	97895		CCOC(C)OC(
61	HMDB0000624	D-Leucic acid	HMDB0000624	439960	C03264	CC(C)C[C@H
62	HMDB0000665	Leucinic acid	HMDB0000665	92779		CC(C)CC(C(=
63	HMDB0000746	Hydroxyisocaproic acid	HMDB0000746	83697		CC(C)C[C@@
64	HMDB0000525	5-Hydroxyhexanoic acid	HMDB0000525	170748		CC(CCCC(=O
65	HMDB0001624	2-Hydroxycaproic acid	HMDB0001624	99824		CCCCC(C(=O
66	HMDB0001975	2-Ethyl-2-Hydroxybutyric acid	HMDB0001975	77199		CCC(CC)(C(=
67	HMDB0010718	(R)-3-Hydroxyhexanoic acid	HMDB0010718	10197713		CCC[C@H](C
68	HMDB0012843	6-Hydroxyhexanoic acid	HMDB0012843	14490	C06103	C(CCC(=O)O
69	HMDB0033851	4,4-Dimethoxy-2-butanone	HMDB0033851	228548		CC(=O)CC(C
70	HMDB0032456	Paraldehyde	HMDB0032456	31264	C07834	CC1OC(OC(O
71	HMDB0040409	NA	NA	NA	NA	NA
72	HMDB0165040	NA	NA	NA	NA	NA
73	HMDB0059770	NA	NA	NA	NA	NA
74	HMDB0041844	NA	NA	NA	NA	NA
75	HMDB0029166	2-Methyl-3-hydroxyvaleric acid	HMDB0029166	152968		CCC(C(C)C(=
76	HMDB0124942	NA	NA	NA	NA	NA
77	HMDB0167829	NA	NA	NA	NA	NA
78	HMDB0062810	NA	NA	NA	NA	NA
79	HMDB0125535	NA	NA	NA	NA	NA
80	HMDB0183895	NA	NA	NA	NA	NA
81	HMDB0126385	NA	NA	NA	NA	NA
82	HMDB0031459	3-Methoxybenzaldehyde	HMDB0031459	11569		COC1=CC=C
83	HMDB0002340	2-Methylbenzoic acid	HMDB0002340	0	C07215	CC1=C(C=C
84	HMDB0041485	NA	NA	NA	NA	NA
85	HMDB0033968	Methyl benzoate	HMDB0033968	7150		COC(=O)C1=
86	HMDB0032400	3-(5-Methyl-2-furyl)prop-2-enal	HMDB0032400	11434978		CC1=CC=C(
87	HMDB0032603	2-Hydroxy-4-methylbenzaldehyde	HMDB0032603	61200		CC1=CC(=C
88	HMDB0029686	4-Methoxybenzaldehyde	HMDB0029686	31244	C10761	COC1=CC=C
89	HMDB0040733	Phenyl acetate	HMDB0040733	31229	C15583	CC(=O)OC1=
90	HMDB0032568	2'-Hydroxyacetophenone	HMDB0032568	9398	C07189	CC(=O)C1=C
91	HMDB0167830	NA	NA	NA	NA	NA
92	HMDB0038176	NA	NA	NA	NA	NA
93	HMDB0040638	NA	NA	NA	NA	NA
94	HMDB0029635	4-Methylbenzoic acid	HMDB0029635	7470	C01454	CC1=CC=C(
95	HMDB0033128	4-(2-Furanyl)-3-buten-2-one	HMDB0033128	1549522		CC(=O)/C=C
96	HMDB0003767	4-Hydroxyphenylacetaldehyde	HMDB0003767	440113	C03765	C1=CC(=CC

97	HMDB0033766	2-Methoxybenzaldehyde	HMDB0033766	8658		COC1=CC=C
98	HMDB0000209	Phenylacetic acid	HMDB0000209	999	C07086	C1=CC=C(C
99	HMDB0126617	NA	NA	NA	NA	NA
100	HMDB0169080	NA	NA	NA	NA	NA
101	HMDB0059966	NA	NA	NA	NA	NA
102	HMDB0131479	NA	NA	NA	NA	NA
103	HMDB0135785	NA	NA	NA	NA	NA
104	HMDB0146894	NA	NA	NA	NA	NA
105	HMDB0146874	NA	NA	NA	NA	NA
106	HMDB0131231	NA	NA	NA	NA	NA
107	HMDB0131232	NA	NA	NA	NA	NA
108	HMDB0125534	NA	NA	NA	NA	NA
109	HMDB0037724	NA	NA	NA	NA	NA
110	HMDB0005784	Hydroxytyrosol	HMDB0005784	82755		C1=CC(=C(C
111	HMDB0030570	Sylvopinol	HMDB0030570	592986		COC1=CC(=
112	HMDB0032012	4-Hydroxy-3-methoxybenzenemethanol	HMDB0032012	62348	C06317	COC1=C(C=
113	HMDB0037728	NA	NA	NA	NA	NA
114	HMDB0034158	2,6-Dimethoxyphenol	HMDB0034158	7041	C10787	COC1=C(C(=
115	HMDB0172312	NA	NA	NA	NA	NA
116	HMDB0170961	NA	NA	NA	NA	NA
117	HMDB0170960	NA	NA	NA	NA	NA
118	HMDB0038679	NA	NA	NA	NA	NA
119	HMDB0172313	NA	NA	NA	NA	NA
120	HMDB0029359	(R)-1,3-Octanediol	HMDB0029359	90927		CCCCCC(CC
121	HMDB0037635	NA	NA	NA	NA	NA
122	HMDB0163103	NA	NA	NA	NA	NA
123	HMDB0136694	NA	NA	NA	NA	NA
124	HMDB0163105	NA	NA	NA	NA	NA
125	HMDB0163106	NA	NA	NA	NA	NA
126	HMDB0163104	NA	NA	NA	NA	NA
127	HMDB0163100	NA	NA	NA	NA	NA
128	HMDB0163101	NA	NA	NA	NA	NA
129	HMDB0163099	NA	NA	NA	NA	NA
130	HMDB0163102	NA	NA	NA	NA	NA
131	HMDB0182795	NA	NA	NA	NA	NA
132	HMDB0169534	NA	NA	NA	NA	NA
133	HMDB0176281	NA	NA	NA	NA	NA
134	HMDB0240639	NA	NA	NA	NA	NA
135	HMDB0182798	NA	NA	NA	NA	NA
136	HMDB0182799	NA	NA	NA	NA	NA
137	HMDB0182797	NA	NA	NA	NA	NA
138	HMDB0182796	NA	NA	NA	NA	NA
139	HMDB0182685	NA	NA	NA	NA	NA
140	HMDB0182684	NA	NA	NA	NA	NA
141	HMDB0181635	NA	NA	NA	NA	NA
142	HMDB0181634	NA	NA	NA	NA	NA
143	HMDB0181636	NA	NA	NA	NA	NA
144	HMDB0181637	NA	NA	NA	NA	NA
145	HMDB0179485	NA	NA	NA	NA	NA
146	HMDB0179394	NA	NA	NA	NA	NA
147	HMDB0178761	NA	NA	NA	NA	NA
148	HMDB0163117	NA	NA	NA	NA	NA
149	HMDB0163109	NA	NA	NA	NA	NA
150	HMDB0163110	NA	NA	NA	NA	NA
151	HMDB0181638	NA	NA	NA	NA	NA
152	HMDB0181639	NA	NA	NA	NA	NA
153	HMDB0181690	NA	NA	NA	NA	NA
154	HMDB0182680	NA	NA	NA	NA	NA
155	HMDB0182683	NA	NA	NA	NA	NA
156	HMDB0182620	NA	NA	NA	NA	NA
157	HMDB0182621	NA	NA	NA	NA	NA
158	HMDB0182619	NA	NA	NA	NA	NA
159	HMDB0182618	NA	NA	NA	NA	NA
160	HMDB0182617	NA	NA	NA	NA	NA
161	HMDB0181693	NA	NA	NA	NA	NA
162	HMDB0181692	NA	NA	NA	NA	NA
163	HMDB0181691	NA	NA	NA	NA	NA
164	HMDB0163108	NA	NA	NA	NA	NA
165	HMDB0182681	NA	NA	NA	NA	NA
166	HMDB0037821	NA	NA	NA	NA	NA
167	HMDB0036067	NA	NA	NA	NA	NA
168	HMDB0037175	NA	NA	NA	NA	NA
169	HMDB0040639	NA	NA	NA	NA	NA
170	HMDB0035714	Epoxyartemisia ketone	HMDB0035714	11816089		CC1(C(O1)C(
171	HMDB0182682	NA	NA	NA	NA	NA
172	HMDB0031156	Ethyl (4Z)-4,7-octadienoate	HMDB0031156	5352798		CCOC(=O)C
173	HMDB0032313	2,4-Hexadienyl isobutyrate	HMDB0032313	5978897		C/C=C/C=C
174	HMDB0035829	Dihydronepetalactone	HMDB0035829	519465		CC1CCC2C1C
175	HMDB0032312	2,4-Hexadienyl butyrate	HMDB0032312	6537958		CCCC(=O)O
176	HMDB0031273	Methyl octynecarboxylate	HMDB0031273	8137		CCCCCCC#
177	HMDB0033700	(-)-(Z)-Tetrahydro-6-(2-pentenyl)-2H-pyran-2-one	HMDB0033700	6438070		CC/C=C/CC
178	HMDB0034452	5,6-Dihydro-6-pentyl-2H-pyran-2-one	HMDB0034452	39914		CCCCCC1CC
179	HMDB0036103	NA	NA	NA	NA	NA
180	HMDB0035766	Ascaridole	HMDB0035766	10545	C09836	CC(C)C12CC
181	HMDB0013105	trans-4,5-epoxy-2(E)-decenal	HMDB0013105	5352429		CCCCCC1C(
182	HMDB0037014	NA	NA	NA	NA	NA
183	HMDB0034670	6-Hydroxy-2,6-dimethyl-2,7-octadien-4-one	HMDB0034670	71379147		CC(C)=CC(=

184	HMDB0032434	(+/-)-2-(5-Methyl-5-vinyltetrahydrofuran-2-yl)propionaldehyde	HMDB0032434	155007		CC(C=O)C1C
185	HMDB0032319	cis-3-Hexenyl crotonate	HMDB0032319	16220110		CC/C=C\CC
186	HMDB0032219	(+/-)-Dihydromintlactone	HMDB0032219	14344497		CC1CCC2C(C
187	HMDB0180148	NA	NA	NA	NA	NA
188	HMDB0180508	NA	NA	NA	NA	NA
189	HMDB0164896	NA	NA	NA	NA	NA
190	HMDB0164962	NA	NA	NA	NA	NA
191	HMDB0180052	NA	NA	NA	NA	NA
192	HMDB0180055	NA	NA	NA	NA	NA
193	HMDB0180054	NA	NA	NA	NA	NA
194	HMDB0180051	NA	NA	NA	NA	NA
195	HMDB0180506	NA	NA	NA	NA	NA
196	HMDB0180507	NA	NA	NA	NA	NA
197	HMDB0180505	NA	NA	NA	NA	NA
198	HMDB0180509	NA	NA	NA	NA	NA
199	HMDB0180053	NA	NA	NA	NA	NA
200	HMDB0180512	NA	NA	NA	NA	NA
201	HMDB0180510	NA	NA	NA	NA	NA
202	HMDB0180511	NA	NA	NA	NA	NA
203	HMDB0163825	NA	NA	NA	NA	NA
204	HMDB0010724	3-Oxodecanoic acid	HMDB0010724	5282982		CCCCCCCC(
205	HMDB0033010	(1S,2S,4S,5R)-1,8-Epoxy-p-menthane-2,5-diol	HMDB0033010	85195105		CC1(C)OC2(C
206	HMDB0039533	NA	NA	NA	NA	NA
207	HMDB0041376	NA	NA	NA	NA	NA
208	HMDB0030369	3-Hydroxy-5Z-octenyl acetate	HMDB0030369	92034372		CC\C=C/CC
209	HMDB0040204	NA	NA	NA	NA	NA
210	HMDB0040205	NA	NA	NA	NA	NA
211	HMDB0030367	Ethyl (3R,5Z)-3-hydroxy-5-octenoate	HMDB0030367	101998819		CCOC(=O)C
212	HMDB0038978	NA	NA	NA	NA	NA
213	HMDB0041545	NA	NA	NA	NA	NA
214	HMDB0039587	NA	NA	NA	NA	NA
215	HMDB0038692	NA	NA	NA	NA	NA
216	HMDB0039582	NA	NA	NA	NA	NA
217	HMDB0038523	NA	NA	NA	NA	NA
218	HMDB0240780	NA	NA	NA	NA	NA
219	HMDB0006469	Linoleyl carnitine	HMDB0006469	6450015		CCCCC/C=C
220	HMDB0060145	NA	NA	NA	NA	NA
221	HMDB0000631	Deoxycholic acid glycine conjugate	HMDB0000631	3035026	C05464	C[C@H](CCC
222	HMDB0000637	Chenodeoxycholic acid glycine conjugate	HMDB0000637	22833540	C05466	C[C@H](CCC
223	HMDB0161141	NA	NA	NA	NA	NA
224	HMDB0184641	NA	NA	NA	NA	NA
225	HMDB0006898	Chenodeoxyglycocholic acid	HMDB0006898	53477907	C05462	CC(CCC(=O)
226	HMDB0161140	NA	NA	NA	NA	NA
227	HMDB0000708	Glycoursodeoxycholic acid	HMDB0000708	12310288		C[C@H](CCC
228	HMDB0161139	NA	NA	NA	NA	NA
229	HMDB0173224	NA	NA	NA	NA	NA
230	HMDB0173223	NA	NA	NA	NA	NA
231	HMDB0173227	NA	NA	NA	NA	NA
232	HMDB0173228	NA	NA	NA	NA	NA
233	HMDB0173226	NA	NA	NA	NA	NA
234	HMDB0173225	NA	NA	NA	NA	NA
235	HMDB0012516	11'-Carboxy-alpha-tocotrienol	HMDB0012516	53481452		CC1=C(C(=C
236	HMDB0030140	Adlupulone	HMDB0030140			CCC(C)C(=C
237	HMDB0030041	Lupulone	HMDB0030041	51397980	C10706	CC(C)CC(=C
238	HMDB0015354	Aminophylline	HMDB0015354	9433		CN1C2=C(C(
239	HMDB0010366	Hyaluronan	HMDB0010366	24759		CC(=O)N[C@
240	HMDB0061300	NA	NA	NA	NA	NA
241	HMDB0060932	NA	NA	NA	NA	NA
242	HMDB0181876	NA	NA	NA	NA	NA
243	HMDB0181865	NA	NA	NA	NA	NA
244	HMDB0181867	NA	NA	NA	NA	NA
245	HMDB0181870	NA	NA	NA	NA	NA
246	HMDB0181877	NA	NA	NA	NA	NA
247	HMDB0181878	NA	NA	NA	NA	NA
248	HMDB0181844	NA	NA	NA	NA	NA
249	HMDB0181871	NA	NA	NA	NA	NA
250	HMDB0181866	NA	NA	NA	NA	NA
251	HMDB0181860	NA	NA	NA	NA	NA
252	HMDB0178480	NA	NA	NA	NA	NA
253	HMDB0165062	NA	NA	NA	NA	NA
254	HMDB0182986	NA	NA	NA	NA	NA
255	HMDB0181843	NA	NA	NA	NA	NA
256	HMDB0181842	NA	NA	NA	NA	NA
257	HMDB0181848	NA	NA	NA	NA	NA
258	HMDB0181850	NA	NA	NA	NA	NA
259	HMDB0181849	NA	NA	NA	NA	NA
260	HMDB0181854	NA	NA	NA	NA	NA
261	HMDB0181856	NA	NA	NA	NA	NA
262	HMDB0181855	NA	NA	NA	NA	NA
263	HMDB0181859	NA	NA	NA	NA	NA
264	HMDB0181861	NA	NA	NA	NA	NA
265	HMDB0178481	NA	NA	NA	NA	NA
266	HMDB0182976	NA	NA	NA	NA	NA
267	HMDB0182978	NA	NA	NA	NA	NA
268	HMDB0182981	NA	NA	NA	NA	NA
269	HMDB0182968	NA	NA	NA	NA	NA
270	HMDB0182982	NA	NA	NA	NA	NA

271	HMDB0182987	NA	NA	NA	NA	NA
272	HMDB0182985	NA	NA	NA	NA	NA
273	HMDB0182990	NA	NA	NA	NA	NA
274	HMDB0034120	Lusitanicoside	HMDB0034120	4484591	C10474	CC1C(C(C(C(
275	HMDB0182989	NA	NA	NA	NA	NA
276	HMDB0182991	NA	NA	NA	NA	NA
277	HMDB0182977	NA	NA	NA	NA	NA
278	HMDB0182973	NA	NA	NA	NA	NA
279	HMDB0181872	NA	NA	NA	NA	NA
280	HMDB0182895	NA	NA	NA	NA	NA
281	HMDB0182896	NA	NA	NA	NA	NA
282	HMDB0182961	NA	NA	NA	NA	NA
283	HMDB0182962	NA	NA	NA	NA	NA
284	HMDB0182963	NA	NA	NA	NA	NA
285	HMDB0182966	NA	NA	NA	NA	NA
286	HMDB0182967	NA	NA	NA	NA	NA
287	HMDB0182971	NA	NA	NA	NA	NA
288	HMDB0182972	NA	NA	NA	NA	NA
289	HMDB0061351	NA	NA	NA	NA	NA
290	HMDB0036983	NA	NA	NA	NA	NA
291	HMDB0030556	Tovophyllin B	HMDB0030556	509268		CC(=CCC1=
292	HMDB0161562	NA	NA	NA	NA	NA
293	HMDB0161601	NA	NA	NA	NA	NA
294	HMDB0161594	NA	NA	NA	NA	NA
295	HMDB0161592	NA	NA	NA	NA	NA
296	HMDB0161568	NA	NA	NA	NA	NA
297	HMDB0161555	NA	NA	NA	NA	NA
298	HMDB0161551	NA	NA	NA	NA	NA
299	HMDB0161553	NA	NA	NA	NA	NA
300	HMDB0145086	NA	NA	NA	NA	NA
301	HMDB0145075	NA	NA	NA	NA	NA
302	HMDB0040608	NA	NA	NA	NA	NA
303	HMDB0015315	Domperidone	HMDB0015315	3151		C1CN(CCC1N
304	HMDB0145596	NA	NA	NA	NA	NA
305	HMDB0030174	Isotetrandrine	HMDB0030174	5422	C17060	CN1CCC2=C
306	HMDB0165995	NA	NA	NA	NA	NA
307	HMDB0165996	NA	NA	NA	NA	NA
308	HMDB0165974	NA	NA	NA	NA	NA
309	HMDB0165971	NA	NA	NA	NA	NA
310	HMDB0165968	NA	NA	NA	NA	NA
311	HMDB0165970	NA	NA	NA	NA	NA
312	HMDB0165997	NA	NA	NA	NA	NA
313	HMDB0166000	NA	NA	NA	NA	NA
314	HMDB0168061	NA	NA	NA	NA	NA
315	HMDB0168059	NA	NA	NA	NA	NA
316	HMDB0168058	NA	NA	NA	NA	NA
317	HMDB0168062	NA	NA	NA	NA	NA
318	HMDB0168073	NA	NA	NA	NA	NA
319	HMDB0168070	NA	NA	NA	NA	NA
320	HMDB0168074	NA	NA	NA	NA	NA
321	HMDB0165969	NA	NA	NA	NA	NA
322	HMDB0168071	NA	NA	NA	NA	NA
323	HMDB0165994	NA	NA	NA	NA	NA
324	HMDB0160769	NA	NA	NA	NA	NA
325	HMDB0160767	NA	NA	NA	NA	NA
326	HMDB0034104	Neferine	HMDB0034104	398793		CN1CCC2=C
327	HMDB0170752	NA	NA	NA	NA	NA
328	HMDB0170753	NA	NA	NA	NA	NA
329	HMDB0165403	NA	NA	NA	NA	NA
330	HMDB0165402	NA	NA	NA	NA	NA
331	HMDB0165404	NA	NA	NA	NA	NA
332	HMDB0166528	NA	NA	NA	NA	NA
333	HMDB0166529	NA	NA	NA	NA	NA
334	HMDB0173096	NA	NA	NA	NA	NA
335	HMDB0173097	NA	NA	NA	NA	NA
336	HMDB0173095	NA	NA	NA	NA	NA
337	HMDB0171551	NA	NA	NA	NA	NA
338	HMDB0171552	NA	NA	NA	NA	NA
339	HMDB0169612	NA	NA	NA	NA	NA
340	HMDB0169611	NA	NA	NA	NA	NA
341	HMDB0169610	NA	NA	NA	NA	NA
342	HMDB0167774	NA	NA	NA	NA	NA
343	HMDB0174559	NA	NA	NA	NA	NA
344	HMDB0174558	NA	NA	NA	NA	NA
345	HMDB0186975	NA	NA	NA	NA	NA
346	HMDB0186976	NA	NA	NA	NA	NA
347	HMDB0145615	NA	NA	NA	NA	NA
348	HMDB0180065	NA	NA	NA	NA	NA
349	HMDB0180066	NA	NA	NA	NA	NA
350	HMDB0169573	NA	NA	NA	NA	NA
351	HMDB0169572	NA	NA	NA	NA	NA
352	HMDB0169575	NA	NA	NA	NA	NA
353	HMDB0169574	NA	NA	NA	NA	NA
354	HMDB0177912	NA	NA	NA	NA	NA
355	HMDB0177909	NA	NA	NA	NA	NA
356	HMDB0060154	NA	NA	NA	NA	NA
357	HMDB0013063	S-(PGJ2)-glutathione	HMDB0013063	53481603		CCCCC[C@@

358	HMDB0013062	S-(PGA2)-glutathione	HMDB0013062	53481602		CCCCCC(/C=
359	HMDB0013058	S-(9-deoxy-delta9,12-PGD2)-glutathione	HMDB0013058	53481598		CCCCC[C@@
360	HMDB0184823	NA	NA	NA	NA	NA
361	HMDB0030115	Kuwanon V	HMDB0030115	72550349		CC(C)=CCC1
362	HMDB0040384	NA	NA	NA	NA	NA
363	HMDB0040881	NA	NA	NA	NA	NA
364	HMDB0039330	NA	NA	NA	NA	NA
365	HMDB0039414	NA	NA	NA	NA	NA
366	HMDB0035429	Assamsaponin B	HMDB0035429	131751747		C\C=C(\C)C
367	HMDB0038606	NA	NA	NA	NA	NA
368	HMDB0014259	Goserelin	HMDB0014259	5311128		CC(C)C[C@@

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 **hsa** (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 **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.

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

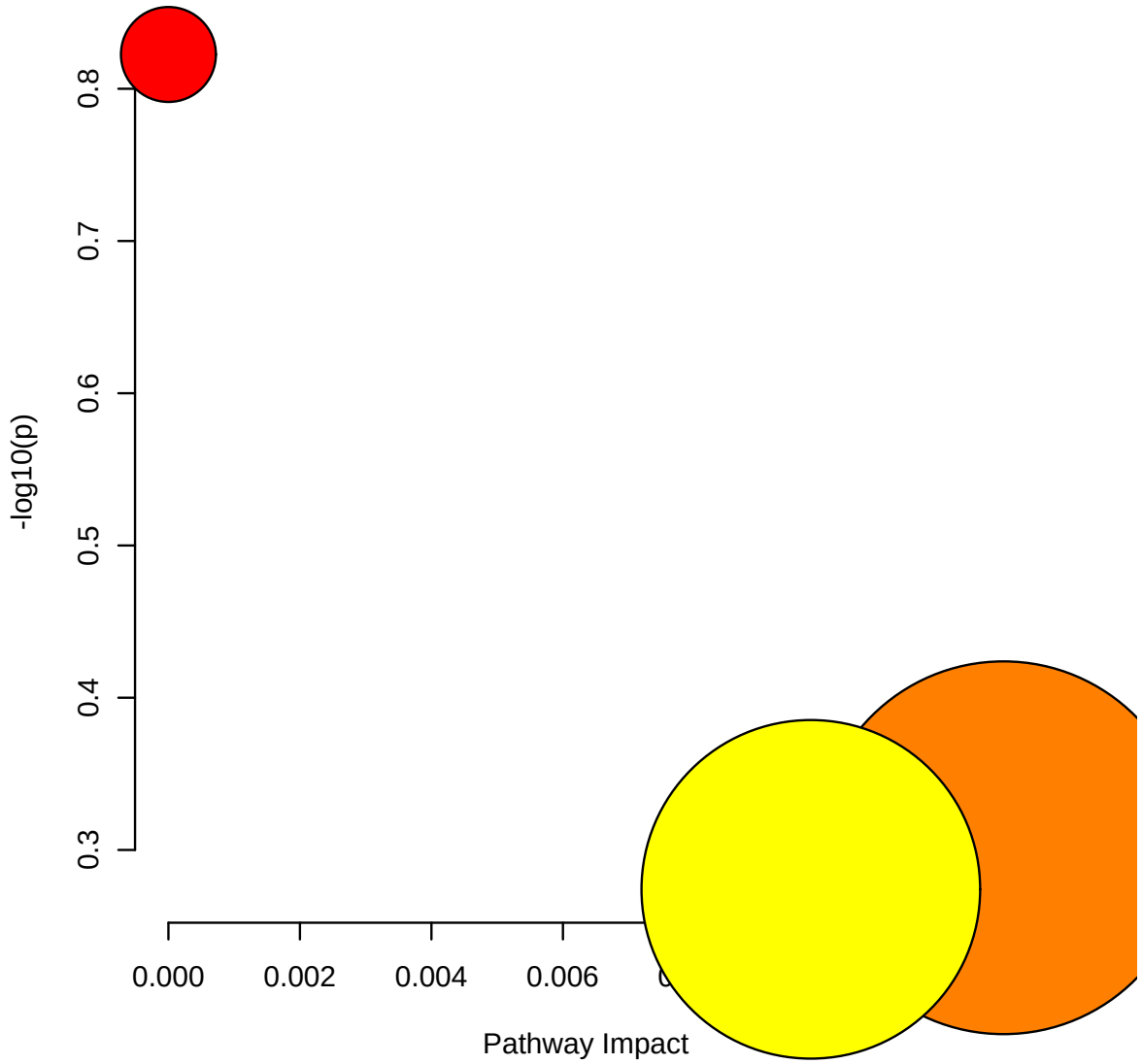

Figure 1: Summary of Pathway Analysis

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
Phenylalanine metabolism	10	0.16	1	1.50E-01	8.23E-01	1.00E+00	1.00E+00	0.00
Tyrosine metabolism	42	0.68	1	5.00E-01	3.01E-01	1.00E+00	1.00E+00	0.01
Primary bile acid biosynthesis	46	0.74	1	5.32E-01	2.74E-01	1.00E+00	1.00E+00	0.01

6 Appendix: R Command History

```
 [1] "mSet<-InitDataObjects(\"conc\", \"pathora\", FALSE)"
 [2] "cmpd.vec<-c(\"HMDB0001858\",\"HMDB0033895\",\"HMDB0003119\",\"HMDB0002055\",\"HMDB0002048\",\"
 [3] "mSet<-Setup.MapData(mSet, cmpd.vec);"
 [4] "mSet<-CrossReferencing(mSet, \"hmdb\");"
 [5] "mSet<-CreateMappingResultTable(mSet)"
 [6] "mSet<-SetKEGG.PathLib(mSet, \"hsa\", \"current\")"
 [7] "mSet<-SetMetabolomeFilter(mSet, F);"
 [8] "mSet<-CalculateOraScore(mSet, \"rbc\", \"hyperg\")"
 [9] "mSet<-PlotPathSummary(mSet, F, \"path_view_0_\", \"png\", 72, width=NA)"
[10] "mSet<-PlotKEGGPath(mSet, \"Primary bile acid biosynthesis\",528, 480, \"png\", NULL)"
[11] "mSet<-RerenderMetPAGraph(mSet, \"zoom1621628358573.png\",528.0, 480.0, 100.0)"
[12] "mSet<-PlotKEGGPath(mSet, \"Tyrosine metabolism\",528, 480, \"png\", NULL)"
[13] "mSet<-PlotKEGGPath(mSet, \"Primary bile acid biosynthesis\",528, 480, \"png\", NULL)"
[14] "mSet<-PlotKEGGPath(mSet, \"Tyrosine metabolism\",528, 480, \"png\", NULL)"
[15] "mSet<-PlotKEGGPath(mSet, \"Primary bile acid biosynthesis\",528, 480, \"png\", NULL)"
[16] "mSet<-PlotKEGGPath(mSet, \"Tyrosine metabolism\",528, 480, \"png\", NULL)"
[17] "mSet<-PlotKEGGPath(mSet, \"Phenylalanine metabolism\",528, 480, \"png\", NULL)"
[18] "mSet<-PlotKEGGPath(mSet, \"Primary bile acid biosynthesis\",528, 480, \"png\", NULL)"
[19] "mSet<-SaveTransformedData(mSet)"
[20] "mSet<-PreparePDFReport(mSet, \"guest5645338793786851237\")\n"
```

The report was generated on Fri May 21 16:19:46 2021 with R version 4.0.2 (2020-06-22).