Metabolomic Data Analysis with MetaboAnalyst 5.0

Name: guest16292574549826598929

March 20, 2021

1   Background

MSEA or Metabolite Set Enrichment Analysis is a way to identify biologically meaningful patterns that are significantly enriched in quantitative metabolomic data. In conventional approaches, metabolites are evaluated individually for their significance under conditions of study. Those compounds that have passed certain significance level are then combined to see if any meaningful patterns can be discerned. In contrast, MSEA directly investigates if a set of functionally related metabolites without the need to preselect compounds based on some arbitrary cut-off threshold. It has the potential to identify subtle but consistent changes among a group of related compounds, which may go undetected with the conventional approaches.

Essentially, MSEA is a metabolomic version of the popular GSEA (Gene Set Enrichment Analysis) software with its own collection of metabolite set libraries as well as an implementation of user-friendly web-interfaces. GSEA is widely used in genomics data analysis and has proven to be a powerful alternative to conventional approaches. For more information, please refer to the original paper by Subramanian A, and a nice review paper by Nam D, Kim SY. [1]. [2]

2   MSEA Overview

Metabolite set enrichment analysis consists of four steps - data input, data processing, data analysis, and results download. Different analysis procedures are performed based on different input types. In addition, users can also browse and search the metabolite set libraries as well as upload their self-defined metabolite sets for enrichment analysis. Users can also perform metabolite name mapping between a variety of compound names, synonyms, and major database identifiers.

3   Data Input

There are three enrichment analysis algorithms offered by MSEA. Accordingly, three different types of data inputs are required by these three approaches:

- A list of important compound names - entered as a one column data (*Over Representation Analysis (ORA)*);

- A single measured biofluid (urine, blood, CSF) sample- entered as tab separated two-column data with the first column for compound name, and the second for concentration values (*Single Sample Profiling (SSP)*);

---

[1] Subramanian *Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles.*, Proc Natl Acad Sci USA. 2005 102(43): 15545-50

[2] Nam D, Kim SY. *Gene-set approach for expression pattern analysis*, Briefings in Bioinformatics. 2008 9(3): 189-197.

- A compound concentration table - entered as a comma separated (.csv) file with the each sample per row and each metabolite concentration per column. The first column is sample names and the second column for sample phenotype labels (*Quantitative Enrichment Analysis (QEA)*)

You selected Over Representation Analysis (ORA) which requires a list of compound names as input.

4   Data Process

The first step is to standardize the compound labels. It is an essential step since the compound labels will be subsequently compared with compounds contained in the metabolite set library. MSEA has a built-in tool to convert between compound common names, synonyms, identifiers used in HMDB ID, PubChem, ChEBI, BiGG, METLIN, KEGG, or Reactome. **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	HMDB0000169	D-Mannose	HMDB0000169	18950	C00936	C([C@@H]1[C@H]([C@@
2	HMDB0000190	L-Lactic acid	HMDB0000190	61503	C00186	C[C@@H](C(=O)O)O
3	HMDB0000122	D-Glucose	HMDB0000122	5793	C00221	C([C@@H]1[C@H]([C@@
4	HMDB0029965	Methyl beta-D-glucopyranoside	HMDB0029965	94214		COC1C(C(C(C(O1)CO)
5	HMDB0000125	NA	NA	NA	NA	NA
6	HMDB0000078	Cysteinylglycine	HMDB0000078	439498	C01419	C([C@@H](C(=O)NCC(
7	HMDB0240253	NA	NA	NA	NA	NA
8	HMDB0038670	NA	NA	NA	NA	NA
9	HMDB0061717	NA	NA	NA	NA	NA
10	HMDB0000725	4-Hydroxyproline	HMDB0000725	5810	C01157	C1[C@H](CN[C@@H]1C
11	HMDB0002712	1,5-Anhydrosorbitol	HMDB0002712	64960	C07326	C1[C@@H]([C@H]([C@@H
12	HMDB0000247	Sorbitol	HMDB0000247	5780	C00794	C([C@H]([C@H]([C@@H
13	HMDB0000660	D-Fructose	HMDB0000660	439709	C02336	C([C@@H]1[C@H]([C@@
14	HMDB0000532	Acetylglycine	HMDB0000532	10972		CC(=O)NCC(=O)O
15	HMDB0003219	Sedoheptulose	HMDB0003219	439645	C08355	C(C(C(C(C(C(=O)CO)C
16	HMDB0000211	myo-Inositol	HMDB0000211		C00137	O[C@H]1[C@H](O)[C@@
17	HMDB0000161	L-Alanine	HMDB0000161	5950	C00041	C[C@@H](C(=O)O)N
18	HMDB0001264	Dehydroascorbic acid	HMDB0001264	7786	C05422	C([C@H]([C@@H]1C(=C
19	HMDB61711	Methylphosphate	HMDB0061711	13130		COP(O)(O)=O
20	HMDB0001401	Glucose 6-phosphate	HMDB0001401	5958	C00092	C([C@@H]1[C@H]([C@@
21	HMDB0000645	Galactose 1-phosphate	HMDB0000645	123912	C00446	C([C@@H]1[C@@H]([C@
22	HMDB0002108	Methylcysteine	HMDB0002108	384585100	C22040	CSC[C@@H](C(=O)O)N
23	HMDB0001078	Mannose 6-phosphate	HMDB0001078	439198	C00275	C([C@@H]1[C@H]([C@@
24	HMDB0001373	Dephospho-CoA	HMDB0001373	439335	C00882	CC(C)(COP(=O)(O)OP
25	HMDB0000210	Pantothenic acid	HMDB0000210	6613	C00864	CC(C)(CO)C(C(=O)NC
26	HMDB0011732	2-Keto-L-gluconate	HMDB0011732	50	C03342	C(C(C(C(C(=O)C(=O)
27	HMDB0000875	Trigonelline	HMDB0000875	5570	C01004	C[N+]1=CC=CC(=C1)C
28	HMDB0001262	Maltotriose	HMDB0001262	439586	C01835	C([C@@H]1[C@H]([C@@
29	HMDB0001423	Coenzyme A	HMDB0001423	87642	C00010	CC(C)(COP(=O)(O)OP
30	HMDB0002092	Itaconic acid	HMDB0002092	811	C00490	C=C(CC(=O)O)C(=O)
31	HMDB0007112	DG(16:0/20:4(5Z,8Z,11Z,14Z)/0:0)	HMDB0007112	9543736	C00165	CCCCCCCCCCCCCCCC
32	HMDB0004827	Proline betaine	HMDB0004827	7016563	C10172	C[N+]1(CCC[C@H]1C(=
33	HMDB0002329	Oxalic acid	HMDB0002329	971	C00209	C(=O)(C(=O)O)O
34	HMDB0001254	Glucosamine 6-phosphate	HMDB0001254	439217	C00352	C([C@@H]1[C@H]([C@@
35	HMDB0000177	L-Histidine	HMDB0000177	6274	C00135	C1=C(NC=N1)C[C@@H
36	HMDB0012141	(R)-2,3-Dihydroxy-isovalerate	HMDB0012141	440279	C04272	CC(C)([C@H](C(=O)O)
37	HMDB0001049	gamma-Glutamylcysteine	HMDB0001049	123938	C00669	C(CC(=O)N[C@@H](CS
38	HMDB0000011	(R)-3-Hydroxybutyric acid	HMDB0000011	92135	C01089	C[C@H](CC(=O)O)O
39	HMDB0240578	NA	NA	NA	NA	NA
40	HMDB0000001	1-Methylhistidine	HMDB0000001	92105	C01152	CN1C=C(N=C1)C[C@@
41	HMDB0000355	3-Hydroxymethylglutaric acid	HMDB0000355	1662	C03761	CC(CC(=O)O)(CC(=O)
42	HMDB0000259	Serotonin	HMDB0000259	5202	C00780	C1=CC2=C(C=C1O)C(
43	HMDB0010395	LysoPC(20:4(5Z,8Z,11Z,14Z))	HMDB0010395	24779476	C04230	CCCCC/C=C\C/C=C\
44	HMDB0000560	5,8-Tetradecadienoic acid	HMDB0000560	5312409		CCCCC/C=C\C/C=C\
45	HMDB0002000	Myristoleic acid	HMDB0002000	5281119	C08322	CCCC/C=C\CCCCCCC
46	HMDB0013133	Methylmalonylcarnitine	HMDB0013133	53481628		CC(C(=O)O)C(=O)O[C
47	HMDB0028844	Glycyl-Isoleucine	HMDB0028844	88079		CCC(C)C(NC(=O)CN)C
48	HMDB0240506	NA	NA	NA	NA	NA
49	HMDB0001348	SM(d18:1/18:0)	HMDB0001348	5283588	C00550	CCCCCCCCCCCCCCCC
50	HMDB0009003	PE(18:0/20:4(5Z,8Z,11Z,14Z))	HMDB0009003	5289133	C00350	CCCCCCCCCCCCCCCC
51	HMDB0013127	Hydroxybutyrylcarnitine	HMDB0013127	53481617		CC(CC(=O)O[C@@H](C
52	HMDB0000126	Glycerol 3-phosphate	HMDB0000126	439162	C00093	C([C@H](COP(=O)(O)O
53	HMDB0240316	NA	NA	NA	NA	NA
54	HMDB0003426	Pantetheine	HMDB0003426	479	C00831	CC(C)(CO)C(C(=O)NC
55	HMDB0000855	Nicotinamide riboside	HMDB0000855	439924	C03150	C1=CC(=C[N+](=C1)[C@
56	HMDB0000917	Ursocholic acid	HMDB0000917	122340	C17644	C[C@H](CCC(=O)O)[C@
57	HMDB0003331	1-Methyladenosine	HMDB0003331	27476	C02494	CN1C=NC2=C(C1=N)N
58	HMDB0000023	(S)-3-Hydroxyisobutyric acid	HMDB0000023	87	C06001	C[C@@H](CO)C(=O)O
59	HMDB0004645	S-Nitrosoglutathione	HMDB0004645	104858		C(CC(=O)N[C@@H](CS

60	HMDB0029200	Ferulic acid 4-sulfate	HMDB0029200	6305574		COC1=C(C=CC(=C1)/...
61	HMDB0000892	Valeric acid	HMDB0000892	7991	C00803	CCCCC(=O)O
62	HMDB0001068	D-Sedoheptulose 7-phosphate	HMDB0001068	22833559	C05382	C([C@@H]1[C@H]([C@H...
63	HMDB0000407	2-Hydroxy-3-methylbutyric acid	HMDB0000407	99823		CC(C)C(C(=O)O)O
64	HMDB0000883	L-Valine	HMDB0000883	6287	C00183	CC(C)[C@@H](C(=O)O...
65	HMDB0000271	Sarcosine	HMDB0000271	1088	C00213	CNCC(=O)O
66	HMDB0000086	Glycerophosphocholine	HMDB0000086	71920	C00670	C[N+](C)(C)CCOP(=O...
67	HMDB0034879	1-[(5-Amino-5-carboxypentyl)amino]-1-deoxyfructose	HMDB0034879	9839580	C16488	NC(CCCCNCC1(O)OC(...
68	HMDB0000860	Phenylpropionylglycine	HMDB0000860	1608347		C1=CC=C(C=C1)CCC(...
69	HMDB0001296	Maltotetraose	HMDB0001296	439639	C02052	C([C@@H]1[C@H]([C@@...
70	HMDB0010575	PG(16:0/18:2(9Z,12Z))	HMDB0010575	52927246		CCCCCCCCCCCCCCCC(O...
71	HMDB0011637	Taurohyocholate	HMDB0011637	0	C15516	[H][C@@]12C[C@H](O)...
72	HMDB0041900	NA	NA	NA	NA	NA
73	HMDB0000722	Lithocholyltaurine	HMDB0000722	53477716	C02592	C[C@H](CCC(=O)NCCS...
74	HMDB0002917	D-Xylitol	HMDB0002917	6912	C00379	C([C@H](C([C@H](CO)O...
75	HMDB0002064	N-Acetylputrescine	HMDB0002064	122356	C02714	CC(=O)NCCCCN
76	HMDB0000317	2-Hydroxy-3-methylpentanoic acid	HMDB0000317	10796774		CC[C@@H](C)[C@H](C(...
77	HMDB0000874	Tauroursodeoxycholic acid	HMDB0000874	12443252		C[C@H](CCC(=O)NCCS...
78	HMDB0000288	Uridine 5'-monophosphate	HMDB0000288	6030	C00105	C1=CN(C(=O)NC1=O)...
79	HMDB0011757	N-Acetylvaline	HMDB0011757	227752		CC(C)C(C(=O)O)NC(=...
80	HMDB0000623	Dodecanedioic acid	HMDB0000623	12736	C02678	C(CCCCCC(=O)O)CC...
81	HMDB0000163	D-Maltose	HMDB0000163	10991489	C00208	C([C@@H]1[C@H]([C@@...
82	HMDB0000665	Leucinic acid	HMDB0000665	92779		CC(C)CC(C(=O)O)O
83	HMDB0010388	LysoPC(18:3(9Z,12Z,15Z))	HMDB0010388	24779469		CC/C=C\C/C=C\C/C=...
84	HMDB0001416	Pantetheine 4'-phosphate	HMDB0001416	987	C01134	CC(C)(COP(=O)(O)O)C...
85	HMDB0000932	Tauro-b-muricholic acid	HMDB0000932	21124703		C[C@H](CCC(=O)NCCS...
86	HMDB0001325	N6,N6,N6-Trimethyl-L-lysine	HMDB0001325	440120	C03793	C[N+](C)(C)CCCC[C@@...
87	HMDB0000752	Methylglutaric acid	HMDB0000752	12284		CC(CC(=O)O)CC(=O)O...
88	HMDB0033433	(S)-Homostachydrine	HMDB0033433	441447	C08283	C[N+]1(CCCC[C@H]1C...
89	HMDB0062174	NA	NA	NA	NA	NA
90	HMDB0000638	Dodecanoic acid	HMDB0000638	3893	C02679	CCCCCCCCCCCC(=O...
91	HMDB0007121	DG(16:0/22:6(4Z,7Z,10Z,13Z,16Z,19Z)/0:0)	HMDB0007121	9543827		CCCCCCCCCCCCCCCC...
92	HMDB0011517	LysoPE(20:4(5Z,8Z,11Z,14Z)/0:0)	HMDB0011517	42607465		CCCCC/C=C\C/C=C\...
93	HMDB0000539	Arabinonic acid	HMDB0000539	122045	C00878	C([C@H]([C@H]([C@@H...
94	HMDB0000729	Alpha-Hydroxyisobutyric acid	HMDB0000729	11671		CC(C)(C(=O)O)O
95	HMDB0013676	2,6-Dihydroxybenzoic acid	HMDB0013676	9338	C21298	C1=CC(=C(C(=C1)O)C(...
96	HMDB0000656	Cysteineglutathione disulfide	HMDB0000656	53477713		C(CC(=O)N[C@@H](CS...
97	HMDB0006210	Heptadecanoyl carnitine	HMDB0006210	53477803		CCCCCCCCCCCCCCCC...
98	HMDB0000848	Stearoylcarnitine	HMDB0000848	52922056		CCCCCCCCCCCCCCCC...
99	HMDB0001494	Acetylphosphate	HMDB0001494	186	C00227	CC(=O)OP(=O)(O)O
100	HMDB0000325	3-Hydroxysuberic acid	HMDB0000325	22328017		C(CCC(=O)O)CC(CC(=...
101	HMDB0000224	O-Phosphoethanolamine	HMDB0000224	1015	C00346	C(COP(=O)(O)O)N
102	HMDB0000477	7Z,10Z-Hexadecadienoic acid	HMDB0000477	13932172		CCCCC/C=C\C/C=C\...
103	HMDB0002453	4-Deoxythreonic acid	HMDB0002453	10964471		C[C@H]([C@@H](C(=O...
104	HMDB0000172	L-Isoleucine	HMDB0000172	6306	C00407	CC[C@H](C)[C@H](C(...
105	HMDB0000872	Tetradecanedioic acid	HMDB0000872	13185		C(CCCCCCC(=O)O)CC...
106	HMDB0012254	Maltopentaose	HMDB0012254	13489094		C([C@@H]1[C@H]([C@@...
107	HMDB0000498	4-Deoxyerythronic acid	HMDB0000498	13120901		C[C@H]([C@H](C(=O)C...
108	HMDB0240577	NA	NA	NA	NA	NA
109	HMDB0000243	Pyruvic acid	HMDB0000243	1060	C00022	CC(=O)C(=O)O
110	HMDB0000893	Suberic acid	HMDB0000893	10457	C08278	C(CCCC(=O)O)CCC(=...
111	HMDB0002271	Imidazolepropionic acid	HMDB0002271	70630	C20522	C1=C(NC=N1)CCC(=...
112	HMDB0000459	3-Methylcrotonylglycine	HMDB0000459	169485		CC(=CC(=O)NCC(=O)...
113	HMDB0000061	Adenosine 3',5'-diphosphate	HMDB0000061	159296	C00054	C1=NC2=C(C(=N1)N)N...
114	HMDB0000267	Pyroglutamic acid	HMDB0000267	7405	C01879	C1CC(=O)N[C@@H]1C...
115	HMDB0000448	Adipic acid	HMDB0000448	196	C06104	C(CCC(=O)O)CC(=O...
116	HMDB0003320	Indole-3-carboxylic acid	HMDB0003320	69867	C19837	C1=CC=C2C(=C1)C(=...
117	HMDB0041623	N6-Carbamoyl-L-threonyladenosine	HMDB0041623	92021849		C[C@@H](O)[C@H](NC(...
118	HMDB41718	NA	NA	NA	NA	NA
119	HMDB0000565	Galactonic acid	HMDB0000565	128869	C00880	C([C@H]([C@@H]([C@@H...
120	HMDB0007248	DG(18:2(9Z,12Z)/18:2(9Z,12Z)/0:0)	HMDB0007248	9543729	C00165	CCCCC/C=C\C/C=C\...
121	HMDB0003334	Symmetric dimethylarginine	HMDB0003334	169148		CNC(=NC)NCCC[C@@...
122	HMDB0011667	gamma-Glutamylglycine	HMDB0011667	165527		C(CC(=O)NCC(=O)O)...
123	HMDB0000511	Capric acid	HMDB0000511	2969	C01571	CCCCCCCCCC(=O)O
124	HMDB0013713	N-acetyltryptophan	HMDB0013713	700653		[H][C@@](CC1=CNC2=...
125	HMDB0000529	5-Dodecenoic acid	HMDB0000529	5312378		CCCCCC/C=C\CCCC(...
126	HMDB0000951	Taurochenodesoxycholic acid	HMDB0000951	387316	C05465	C[C@H](CCC(=O)NCCS...
127	HMDB0000749	Mesaconic acid	HMDB0000749	643798	C01732	C/C(=C\C(=O)O)/C(=...
128	HMDB0002820	Methylimidazoleacetic acid	HMDB0002820	75810	C05828	CN1C=C(N=C1)CC(=O...
129	HMDB0062551	4-ethylphenylsulfate	HMDB0062551	20822574		CCC1=CC=C(OS(O)(=...
130	HMDB0001396	5-Methyltetrahydrofolic acid	HMDB0001396	439234	C00440	CN1C(CNC2=C1C(=O)...
131	HMDB0003312	Daidzein	HMDB0003312	5281708	C10208	C1=CC(=CC=C1C2=C...
132	HMDB0004666	2-Arachidonylglycerol	HMDB0004666	5282280	C13856	CCCCC/C=C\C/C=C\...
133	HMDB0000536	Adenylsuccinic acid	HMDB0000536	440122	C03794	C1=NC2=C(C(=N1)NC...
134	HMDB0002815	LysoPC(18:1(9Z))	HMDB0002815	16081932	C04230	CCCCCCCC/C=C\CC...
135	HMDB0009093	PE(18:2(9Z,12Z)/18:2(9Z,12Z))	HMDB0009093	9546812	C00350	CCCCC/C=C\C/C=C\...
136	HMDB0000630	Cytosine	HMDB0000630	597	C00380	C1=C(NC(=O)N=C1)N...
137	HMDB01448	Sulfate	HMDB0001448	1117	C00059	[O-]S(=O)(=O)[O-]
138	HMDB0000754	3-Hydroxyisovaleric acid	HMDB0000754	69362	C20827	CC(C)(CC(=O)O)O
139	HMDB0000235	Thiamine	HMDB0000235	1130	C00378	CC1=C(SC=[N+]1CC2=...
140	HMDB0001316	6-Phosphogluconic acid	HMDB0001316	91493	C00345	C([C@H]([C@@H]([C@@H...
141	HMDB0000339	2-Methylbutyrylglycine	HMDB0000339	193872		CCC(C)C(=O)NCC(=O...
142	HMDB0000222	L-Palmitoylcarnitine	HMDB0000222	11953816	C02990	CCCCCCCCCCCCCCCC...
143	HMDB0009012	PE(18:0/22:6(4Z,7Z,10Z,13Z,16Z,19Z))	HMDB0009012	9546798	C00350	CCCCCCCCCCCCCCCC...
144	HMDB0002302	Indole-3-propionic acid	HMDB0002302	3744		C1=CC=C2C(=C1)C(=...
145	HMDB0000123	Glycine	HMDB0000123	750	C00037	C(C(=O)O)N
146	HMDB0008147	PC(18:2(9Z,12Z)/20:4(5Z,8Z,11Z,14Z))	HMDB0008147	24778979	C00157	CCCCC/C=C\C/C=C\...

3

147	HMDB0000387	3-Hydroxydodecanoic acid	HMDB0000387	94216		CCCCCCCCCC(CC(=O...
148	HMDB0000884	Ribothymidine	HMDB0000884	445408		CC1=CN(C(=O)NC1=O...
149	HMDB0011537	MG(0:0/18:1(9Z)/0:0)	HMDB0011537	5319879		CCCCCCCCC/C=C\CCC...
150	HMDB0059724	NA	NA	NA	NA	NA
151	HMDB0000792	Sebacic acid	HMDB0000792	5192	C08277	C(CCCCC(=O)O)CCC...
152	HMDB0000672	Hexadecanedioic acid	HMDB0000672	10459	C19615	C(CCCCCCCC(=O)O)C...
153	HMDB0000036	Taurocholic acid	HMDB0000036	6675	C05122	CC(CCC(=O)NCCS(=O...
154	HMDB0007228	DG(18:1(9Z)/20:4(5Z,8Z,11Z,14Z)/0:0)	HMDB0007228	9543786	C00165	CCCCCCCC/C=C\CC...
155	HMDB0001275	Propionyl-CoA	HMDB0001275	439164	C00100	CCC(=O)SCCNC(=O)...
156	HMDB0000114	Glycerylphosphorylethanolamine	HMDB0000114	22833510	C01233	C(CO)N(CC(CO)O)P(=...
157	HMDB0000714	Hippuric acid	HMDB0000714	464	C01586	C1=CC=C(C=C1)C(=O...
158	HMDB0000131	Glycerol	HMDB0000131	753	C00116	C(C(CO)O)O
159	HMDB0000512	N-Acetyl-L-phenylalanine	HMDB0000512	74839	C03519	CC(=O)N[C@@H](CC1=...
160	HMDB10383	LysoPC(16:1(9Z))	HMDB0010383	24779461	C04230	CCCCCC/C=C\CCCC...
161	HMDB0001954	3-Hydroxyoctanoic acid	HMDB0001954	26613	C20793	CCCCCC(CC(=O)O)O...
162	HMDB0000522	3-Methylglutaconic acid	HMDB0000522	1551553		C/C(=C\C(=O)O)/CC(...
163	HMDB0000897	7-Methylguanine	HMDB0000897	11361	C02242	CN1C=NC2=C1C(=O)N...
164	HMDB0011557	MG(0:0/22:6(4Z,7Z,10Z,13Z,16Z,19Z)/0:0)	HMDB0011557	46896303		CC/C=C\C/C=C\C/C= ...
165	HMDB0010734	(R)-3-Hydroxy-hexadecanoic acid	HMDB0010734	15569776		CCCCCCCCCCCCC[C@...
166	HMDB0011550	MG(0:0/20:5(5Z,8Z,11Z,14Z,17Z)/0:0)	HMDB0011550	11660820		CC/C=C\C/C=C\C/C= ...
167	HMDB0007219	DG(18:1(9Z)/18:2(9Z,12Z)/0:0)	HMDB0007219	9543722	C00165	CCCCCCCC/C=C\CC...
168	HMDB0011741	gamma-Glutamyltyrosine	HMDB0011741	94340		C1=CC(=CC=C1C[C@@...
169	HMDB0009789	PI(16:0/20:4(5Z,8Z,11Z,14Z))	HMDB0009789		C00626	[H][C@@](COC(=O)CC...
170	HMDB0001552	2-Keto-glutaramic acid	HMDB0001552	48	C00940	C(CC(=O)N)C(=O)C(=...
171	HMDB0000323	3-Amino-2-piperidone	HMDB0000323	5200225		C1CC(C(=O)NC1)N
172	HMDB0000176	Maleic acid	HMDB0000176	444972	C01384	C(=C\C(=O)O)\C(=O)...
173	HMDB0003072	Quinic acid	HMDB0003072	6508	C00296	OC1C[C@@](O)(C[C@@...
174	HMDB0005065	Oleoylcarnitine	HMDB0005065	46907933		CCCCCCCC/C=C\CC...
175	HMDB0000896	Taurodeoxycholic acid	HMDB0000896	2733768	C05463	C[C@H](CCC(=O)NCCS...
176	HMDB0000736	Isobutyryl-L-carnitine	HMDB0000736	168379		CC(C)C(=O)OC(CC(=O...
177	HMDB0000044	Ascorbic acid	HMDB0000044	54670067	C01041	C([C@@H]([C@@H]1C(=...
178	HMDB000225	NA	NA	NA	NA	NA
179	HMDB0060018	NA	NA	NA	NA	NA
180	HMDB0004949	Ceramide (d18:1/16:0)	HMDB0004949	5283564	C00195	CCCCCCCCCCCCCCC...
181	HMDB0000782	Octadecanedioic acid	HMDB0000782	70095		C(CCCCCCCC(=O)O)...
182	HMDB0000413	3-Hydroxydodecanedioic acid	HMDB0000413	16663321		C(CCCCC(=O)O)CCC...
183	HMDB0001178	Adenosine diphosphate ribose	HMDB0001178	33576	C00301	C1=NC2=C(C(=N1)N)N...
184	HMDB00939	S-Adenosylhomocysteine	HMDB0000939	439155	C00021	C1=NC2=C(C(=N1)N)N...
185	HMDB0000094	Citric acid	HMDB0000094	311	C00158	C(C(=O)O)C(CC(=O)O...
186	HMDB0010383	LysoPC(16:1(9Z))	HMDB0010383	24779461	C04230	CCCCCC/C=C\CCCC...
187	HMDB00653	Cholesterol sulfate	HMDB0000653	65076	C18043	C[C@H](CCCC(C)C)[C@...
188	HMDB0000071	Deoxyinosine	HMDB0000071	65058	C05512	C1[C@@H]([C@H](O[C@...
189	HMDB0062557	N-Acetylthreonine	HMDB0062557	152204		C[C@@H](O)[C@H](NC(...
190	HMDB0000613	Erythronic acid	HMDB0000613	2781043		C([C@H]([C@H](C(=O)...
191	HMDB0061116	NA	NA	NA	NA	NA
192	HMDB0007972	PC(16:0/18:1(9Z))	HMDB0007972	5497103	C00157	CCCCCCCCCCCCCCC...
193	HMDB0002721	1-Methylinosine	HMDB0002721	65095		CN1C=NC2=C(C1=O)N...
194	HMDB0004225	2-Oxoarginine	HMDB0004225	558	C03771	C(CC(=O)C(=O)O)CN=...
195	HMDB0007974	PC(16:0/18:3(6Z,9Z,12Z))	HMDB0007974	24778699	C00157	CCCCCCCCCCCCCCC...
196	HMDB0240365	NA	NA	NA	NA	NA
197	HMDB0028691	Alanyl-Leucine	HMDB0028691	96801		CC(C)CC(NC(=O)C(C)...
198	HMDB0061652	3-hydroxyhexanoic acid	HMDB0061652	11829482		CCC[C@H](O)CC(O)=O...
199	HMDB0000269	Sphinganine	HMDB0000269	91486	C00836	CCCCCCCCCCCCCCC...
200	HMDB0240588	NA	NA	NA	NA	NA
201	HMDB0000089	Cytidine	HMDB0000089	6253	C00475	C1=CN(C(=O)N=C1N)...
202	HMDB0001522	Methylguanidine	HMDB0001522	10111	C02294	CN=C(N)N
203	HMDB0000626	Deoxycholic acid	HMDB0000626	222528	C04483	C[C@H](CCC(=O)O)[C...
204	HMDB0003290	Gulonic acid	HMDB0003290	152304	C00800	C([C@H]([C@@H]([C@H...
205	HMDB0007170	DG(18:0/20:4(5Z,8Z,11Z,14Z)/0:0)	HMDB0007170	6438587	C00165	CCCCCCCCCCCCCCC...
206	HMDB0013010	N-Heptanoylglycine	HMDB0013010	10932172		CCCCCCC(=O)NCC(=...
207	HMDB0002203	3-Hydroxycapric acid	HMDB0002203	26612		CCCCCCC(CC(=O)O)...
208	HMDB0000138	Glycocholic acid	HMDB0000138	23617285	C01921	C[C@H](CCC(=O)NCC...
209	HMDB0000305	Vitamin A	HMDB0000305	445354	C17276	CC1=C(C(CCC1)(C)C)/...
210	HMDB0011635	p-Cresol sulfate	HMDB0011635	4615423		CC1=CC=C(C=C1)OS(...
211	HMDB0000337	(S)-3,4-Dihydroxybutyric acid	HMDB0000337	150929		C(C(CO)O)C(=O)O
212	HMDB0000687	L-Leucine	HMDB0000687	6106	C00123	CC(C)C[C@@H](C(=O)...
213	HMDB0011534	MG(0:0/16:1(9Z)/0:0)	HMDB0011534	53480960		CCCCCC/C=C\CCCC...
214	HMDB0000350	3-Hydroxysebacic acid	HMDB0000350	3017884		C(CCCC(=O)O)CCC(C...
215	HMDB0000277	Sphingosine 1-phosphate	HMDB0000277	5353956	C06124	CCCCCCCCCCCCC/C= ...
216	HMDB0013336	3-Hydroxyhexadecanoylcarnitine	HMDB0013336	53481691		CCCCCCCCCCCCCC(...
217	HMDB0000263	Phosphoenolpyruvic acid	HMDB0000263	1005	C00074	C=C(C(=O)O)OP(=O)(...
218	HMDB0011506	LysoPE(18:1(9Z)/0:0)	HMDB0011506	9547071		CCCCCCCC/C=C\CC...
219	HMDB0002201	N-Carboxyethyl-g-aminobutyric acid	HMDB0002201	2572		C(CC(=O)O)CNCCC(=...
220	HMDB0013287	Ne,Ne dimethyllysine	HMDB0013287	4478779	C05545	CN(C)CCCCC(C(=O)O...
221	HMDB0004980	cis-4-Decenoic acid	HMDB0004980	5312351		CCCCC/C=C\CCC(=O...
222	HMDB0012091	SM(d18:0/22:0)	HMDB0012091	44260132	C00550	CCCCCCCCCCCCCCC...
223	HMDB0035514	xi-2,3-Dihydro-2-oxo-1H-indole-3-acetic acid	HMDB0035514	3080590		OC(=O)CC1C(O)=NC2...
224	HMDB0000881	Xanthurenic acid	HMDB0000881	5699	C02470	C1=CC2=C(C(=C1)O)N...
225	HMDB0003217	Genistein	HMDB0003217	5280961	C06563	C1=CC(=CC=C1C2=C...
226	HMDB0000779	Phenyllactic acid	HMDB0000779	3848	C01479	C1=CC=C(C=C1)CC(C...
227	HMDB0000898	1-Methylhistamine	HMDB0000898	3614	C05127	CN1C=C(N=C1)CCN
228	HMDB0001056	Dihydrofolic acid	HMDB0001056	98792	C00415	C1C(=NC2=C(N1)NC(=...
229	HMDB0000958	trans-Aconitic acid	HMDB0000958	444212	C02341	C(/C(=C\C(=O)O)/C(=...
230	HMDB0000555	3-Methyladipic acid	HMDB0000555	6999745		C[C@@H](CCC(=O)O)C...
231	HMDB0001173	5'-Methylthioadenosine	HMDB0001173	439176	C00170	CSC[C@@H]1[C@H]([C@...
232	HMDB0011621	Cinnamoylglycine	HMDB0011621	709625		C1=CC=C(C=C1)/C=C...
233	HMDB0000056	Beta-Alanine	HMDB0000056	239	C00099	C(CN)C(=O)O

234	HMDB0000929	L-Tryptophan	HMDB0000929	6305	C00078	C1=CC=C2C(=C1)C(=
235	HMDB0028929	Leucyl-Glycine	HMDB0028929	97364		CC(C)CC(N)C(=O)NCC
236	HMDB0001008	Biliverdin	HMDB0001008	5353439	C00500	CC\1=C(/C(=C/C2=C
237	HMDB0061684	N-Acetylisoleucine	HMDB0061684	7036275		CC[C@H](C)[C@H](NC(
238	HMDB0000767	Pseudouridine	HMDB0000767	15047	C02067	C1=C(C(=O)NC(=O)N
239	HMDB0012458	7alpha-Hydroxy-3-oxo-4-cholestenoate	HMDB0012458	3081085	C17337	C[C@H](CCCC(C)C(=O
240	HMDB0000201	L-Acetylcarnitine	HMDB0000201	7045767	C02571	CC(=O)OC(CC(=O)[O-
241	HMDB0000832	Capryloylglycine	HMDB0000832	84290		CCCCCCCC(=O)NCC(
242	HMDB0007132	DG(16:1(9Z)/18:2(9Z,12Z)/0:0)	HMDB0007132	9543699	C00165	CCCCCC/C=C\CCCC
243	HMDB0000699	1-Methylnicotinamide	HMDB0000699	457	C02918	C[N+]1=CC=CC(=C1)C
244	HMDB0005781	Glycitein	HMDB0005781	5317750	C14536	COC1=C(C=C2C(=C1)
245	HMDB0011352	PE(P-16:0/20:4(5Z,8Z,11Z,14Z))	HMDB0011352	52925126		CCCCCCCCCCCCCC/C
246	HMDB000506	NA	NA	NA	NA	NA
247	HMDB0000927	Valerylglycine	HMDB0000927	4737557		CCCCC(=O)NCC(=O)C
248	HMDB0001893	Alpha-Tocopherol	HMDB0001893	14985	C02477	CC1=C(C(=C2CC[C@@
249	HMDB0011738	N2-gamma-Glutamylglutamine	HMDB0011738	150914	C05283	C(CC(=O)N[C@@H](CC
250	HMDB0011152	PE(P-16:0e/0:0)	HMDB0011152	42607469		CCCCCCCCCCCCCC/C
251	HMDB0011538	MG(0:0/18:2(9Z,12Z)/0:0)	HMDB0011538	5365676		CCCCC/C=C\C/C=C\
252	HMDB0000959	Tiglylglycine	HMDB0000959	6441567		C/C=C(\C)/C(=O)NCC
253	HMDB0000912	Succinyladenosine	HMDB0000912	20849086		C1=NC2=C(C(=N1)N[C
254	HMDB0000283	D-Ribose	HMDB0000283	5779	C00121	C([C@@H]1[C@H]([C@H
255	HMDB0000064	Creatine	HMDB0000064	586	C00300	CN(CC(=O)O)C(=N)N
256	HMDB0000562	Creatinine	HMDB0000562	588	C00791	CN1CC(=O)N=C1N
257	HMDB0007883	PC(14:0/20:4(5Z,8Z,11Z,14Z))	HMDB0007883	24778634	C00157	CCCCCCCCCCCCCC(=
258	HMDB0001358	Retinal	HMDB0001358	6436079	C00376	CC1=C(C(CCC1)(C)C),
259	HMDB0013130	Glutarylcarnitine	HMDB0013130	71317118		C[N+](C)(C)C[C@@H](O
260	HMDB0000833	N-Glycolylneuraminic acid	HMDB0000833	16738688	C03410	C1[C@@H]([C@H]([C@@
261	HMDB0002095	Malonylcarnitine	HMDB0002095	22833583		C[N+](C)(C)C[C@H](CO
262	HMDB0000158	L-Tyrosine	HMDB0000158	6057	C00082	C1=CC(=CC=C1C[C@
263	HMDB0000294	Urea	HMDB0000294	2447	C00086	C(=O)(N)N
264	HMDB0002005	Methionine sulfoxide	HMDB0002005	847	C02989	CS(=O)CCC(C(=O)O)N
265	HMDB11580	MG(20:5(5Z,8Z,11Z,14Z,17Z)/0:0/0:0)	HMDB0011580	53480988		CC/C=C\C/C=C\C/C=
266	HMDB0000128	Guanidoacetic acid	HMDB0000128	763	C00581	C(C(=O)O)N=C(N)N
267	HMDB0000299	Xanthosine	HMDB0000299	64959	C01762	C1=NC2=C(N1[C@H]3[
268	HMDB0041724	NA	NA	NA	NA	NA
269	HMDB0000651	Decanoylcarnitine	HMDB0000651	10245190		CCCCCCCCCC(=O)OC
270	HMDB0001176	Cytidine monophosphate N-acetylneuraminic acid	HMDB0001176	448209	C00128	CC(=O)N[C@@H]1[C@H
271	HMDB0010325	Ethyl glucuronide	HMDB0010325	18392195		CCO[C@H]1[C@@H]([C@
272	HMDB0001547	Corticosterone	HMDB0001547	5753	C02140	C[C@]12CCC(=O)C=C1
273	HMDB0000788	Orotidine	HMDB0000788	92751	C01103	C1=C(N(C(=O)NC1=O)
274	HMDB0000462	Allantoin	HMDB0000462	204	C01551	C1(C(=O)NC(=O)N1)N
275	HMDB0000840	Salicyluric acid	HMDB0000840	10253	C07588	C1=CC=C(C(=C1)C(=
276	HMDB0000062	L-Carnitine	HMDB0000062	2724480	C00318	C[N+](C)(C)C[C@@H](CO
277	HMDB0000866	N-Acetyl-L-tyrosine	HMDB0000866	68310	C01657	CC(=O)N[C@@H](CC1=
278	HMDB0014903	Metyrosine	HMDB0014903	441350	C07921	C[C@](CC1=CC=C(C=
279	HMDB0007102	DG(16:0/18:1(9Z)/0:0)	HMDB0007102	5282283	C13861	CCCCCCCCCCCCCCCC
280	HMDB0001414	Putrescine	HMDB0001414	1045	C00134	C(CCN)CN
281	HMDB0000678	Isovalerylglycine	HMDB0000678	546304		CC(C)CC(=O)NCC(=O
282	HMDB0000679	Homocitrulline	HMDB0000679	65072	C02427	C(CCNC(=O)N)C[C@@
283	HMDB0061677	NA	NA	NA	NA	NA
284	HMDB0001397	Guanosine monophosphate	HMDB0001397	6804	C00144	C1=NC2=C(N1[C@H]3[
285	HMDB11760	Cer(d18:0/16:0)	HMDB0011760	5283572		CCCCCCCCCCCCCCCC
286	HMDB0007969	PC(16:0/16:1(9Z))	HMDB0007969	6443788	C00157	CCCCCCCCCCCCCCCC
287	HMDB0006469	Linoleyl carnitine	HMDB0006469	6450015		CCCCC/C=C\C/C=C\
288	HMDB0000943	Threonic acid	HMDB0000943	151152	C01620	C([C@H]([C@@H](C(=O
289	HMDB0029159	gamma-Glutamylthreonine	HMDB0029159	53861142		C[C@H]([C@@H](C(=O)
290	HMDB0001555	Pyridoxamine 5'-phosphate	HMDB0001555	1053	C00647	CC1=NC=C(C(=C1O)C
291	HMDB0001895	Salicylic acid	HMDB0001895	338	C00805	C1=CC=C(C(=C1)C(=
292	HMDB0007103	DG(16:0/18:2(9Z,12Z)/0:0)	HMDB0007103	9543695	C00165	CCCCCCCCCCCCCCCC
293	HMDB0007249	DG(18:2(9Z,12Z)/18:3(6Z,9Z,12Z)/0:0)	HMDB0007249	14275409		CCCCC/C=C\C/C=C\
294	HMDB0000479	3-Methylhistidine	HMDB0000479	64969	C01152	CN1C=NC=C1C[C@@H
295	HMDB13326	trans-2-Dodecenoylcarnitine	HMDB0013326	53481671		CCCCCCCCC/C=C/C(
296	HMDB0000696	L-Methionine	HMDB0000696	6137	C00073	CSCC[C@@H](C(=O)O)
297	HMDB0001206	Acetyl-CoA	HMDB0001206	444493	C00024	CC(=O)SCCNC(=O)CC
298	HMDB0000127	D-Glucuronic acid	HMDB0000127	94715	C00191	[C@@H]1([C@@H]([C@H
299	HMDB0011756	N-Acetylleucine	HMDB0011756	70912	C02710	CC(C)C[C@@H](C(=O)
300	HMDB0000671	Indolelactic acid	HMDB0000671	92904	C02043	C1=CC=C2C(=C1)C(=
301	HMDB0094710	3-[3-(Sulfooxy)phenyl]propanoic acid	HMDB0094710	187488		OC(=O)CCC1=CC(OS(
302	HMDB0004194	N1-Methyl-4-pyridone-3-carboxamide	HMDB0004194	440810	C05843	CN1C=CC(=O)C(=C1)
303	HMDB0001874	D-threo-Isocitric acid	HMDB0001874	5318532	C00451	C([C@@H]([C@H](C(=O
304	HMDB0000206	N6-Acetyl-L-lysine	HMDB0000206	92832	C02727	CC(=O)NCCCC[C@@H
305	HMDB0000014	Deoxycytidine	HMDB0000014	13711	C00881	C1[C@@H]([C@H](O[C@
306	HMDB0003464	4-Guanidinobutanoic acid	HMDB0003464	500	C01035	C(CC(=O)O)CN=C(N)N
307	HMDB0240592	NA	NA	NA	NA	NA
308	HMDB0059586	NA	NA	NA	NA	NA
309	HMDB0013205	9-Decenoylcarnitine	HMDB0013205	53481651		C[N+](C)(C)CC(CC(=O
310	HMDB0000921	Cholestenone	HMDB0000921	91477	C00599	C[C@H](CCCC(C)C)[C@
311	HMDB0000107	Galactitol	HMDB0000107	11850	C01697	C([C@H]([C@@H]([C@@
312	HMDB0007982	PC(16:0/20:4(5Z,8Z,11Z,14Z))	HMDB0007982	10747814	C00157	CCCCCCCCCCCCCCCC
313	HMDB0002817	N-Acetylglucosamine 6-phosphate	HMDB0002817	439219		CC(=O)N[C@@H]1[C@H
314	HMDB0001276	N1-Acetylspermidine	HMDB0001276	496	C00612	CC(=O)NCCCNCCCCN
315	HMDB0000821	Phenylacetylglycine	HMDB0000821	68144	C05598	C1=CC=C(C=C1)CC(=
316	HMDB0008138	PC(18:2(9Z,12Z)/18:2(9Z,12Z))	HMDB0008138	5288075	C00157	CCCCC/C=C\C/C=C\
317	HMDB0094696	N-Methyl-proline	HMDB0094696	643474		CN1CCC[C@H]1C(O)=O
318	HMDB0002100	Palmitoylethanolamide	HMDB0002100	4671	C16512	CCCCCCCCCCCCCCCC
319	HMDB11578	MG(20:4(5Z,8Z,11Z,14Z)/0:0/0:0)	HMDB0011578	16019980		CCCCC/C=C\C/C=C\
320	HMDB0007218	DG(18:1(9Z)/18:1(9Z)/0:0)	HMDB0007218	9543716	C00165	CCCCCCCC/C=C\CC(

5

321	HMDB0012097	SM(d18:1/14:0)	HMDB0012097	11433862		CCCCCCCCCCCCCC/C=
322	HMDB0011691	Cytidine 2',3'-cyclic phosphate	HMDB0011691	53481030	C00127	C1=CN(C(=O)N=C1N)
323	HMDB0005320	NA	NA	NA	NA	NA
324	HMDB0000162	L-Proline	HMDB0000162	145742	C00148	C1C[C@H](NC1)C(=O)(
325	HMDB0001070	Octanoyl-CoA	HMDB0001070	380	C01944	CCCCCCCC(=O)SCCN
326	HMDB0013128	Valerylcarnitine	HMDB0013128	53481619		CCCCC(=O)O[C@@H](
327	HMDB0000669	Ortho-Hydroxyphenylacetic acid	HMDB0000669	11970	C05852	C1=CC=C(C(=C1)CC(
328	HMDB0011686	p-Cresol glucuronide	HMDB0011686	154035		CC1=CC=C(C=C1)O[C
329	HMDB0028995	Phenylalanyl-Glycine	HMDB0028995	98207		NC(CC1=CC=CC=C1)(
330	HMDB0009784	PI(16:0/18:2(9Z,12Z))	HMDB0009784		C00626	[H][C@@](COC(=O)CC
331	HMDB0000982	5-Methylcytidine	HMDB0000982	92918		CC1=CN(C(=O)N=C1N
332	HMDB0001301	1-Pyrroline-5-carboxylic acid	HMDB0001301	6642	C03912	C1CC(N=C1)C(=O)O
333	HMDB04988	Pi-Methylimidazoleacetic acid	HMDB0004988	6451814		CN1C=NC=C1CC(=O)
334	HMDB0008994	PE(18:0/18:2(9Z,12Z))	HMDB0008994	9546749	C00350	CCCCCCCCCCCCCCCC
335	HMDB0000045	Adenosine monophosphate	HMDB0000045	6083	C00020	C1=NC2=C(C(=N1)N)N
336	HMDB0001041	2-Methylbutyryl-CoA	HMDB0001041	439371	C15980	CCC(C)C(=O)SCCNC(=
337	HMDB0010168	SM(d18:0/16:0)	HMDB0010168	5283591	C00550	CCCCCCCCCCCCCCCC
338	HMDB0004620	N-a-Acetyl-L-arginine	HMDB0004620	67427		CC(=O)N[C@@H](CCC
339	HMDB0004950	Ceramide (d18:1/18:0)	HMDB0004950	5283565	C00195	CCCCCCCCCCCCCCCC
340	HMDB0000005	2-Ketobutyric acid	HMDB0000005	58	C00109	CCC(=O)C(=O)O
341	HMDB0011170	gamma-Glutamylisoleucine	HMDB0011170	22885096		CC[C@H](C)[C@@H](C(
342	HMDB0000450	5-Hydroxylysine	HMDB0000450	4433	C16741	C(C[C@@H](C(=O)O)N
343	HMDB0061636	NA	NA	NA	NA	NA
344	HMDB0000133	Guanosine	HMDB0000133	6802	C00387	C1=NC2=C(N1[C@H]3[
345	HMDB0003518	Homocitric acid	HMDB0003518	439459	C01251	C(C[C@@](CC(=O)O)(C
346	HMDB0002024	Imidazoleacetic acid	HMDB0002024	96215	C02835	C1=C(NC=N1)CC(=O)
347	HMDB0032055	N-Acetylhistidine	HMDB0032055	273260	C02997	CC(=O)NC(CC1=CN=C
348	HMDB0005779	PE(O-18:1(1Z)/20:4(5Z,8Z,11Z,14Z))	HMDB0005779	9547058	C00350	CCCCCCCCCCCCCCCC
349	HMDB0002331	Imidazoleacetic acid riboside	HMDB0002331	440569	C05131	C1=C(N=CN1[C@H]2[C
350	HMDB0029151	gamma-Glutamylhistidine	HMDB0029151	7017195		C1=C(NC=N1)C[C@@H
351	HMDB0001257	Spermidine	HMDB0001257	1102	C00315	C(CCNCCCN)CN

The second step is to check concentration values. For SSP analysis, the concentration must be measured in *umol* for blood and CSF samples. The urinary concentrations must be first converted to *umol/mmol_creatinine* in order to compare with reported concentrations in literature. No missing or negative values are allowed in SSP analysis. The concentration data for QEA analysis is more flexible. Users can upload either the original concentration data or normalized data. Missing or negative values are allowed (coded as *NA*) for QEA.

5 Selection of Metabolite Set Library

Before proceeding to enrichment analysis, a metabolite set library has to be chosen. There are seven built-in libraries offered by MSEA:

- Metabolic pathway associated metabolite sets (*currently contains 99 entries*);

- Disease associated metabolite sets (reported in blood) (*currently contains 344 entries*);

- Disease associated metabolite sets (reported in urine) (*currently contains 384 entries*)

- Disease associated metabolite sets (reported in CSF) (*currently contains 166 entries*)

- Metabolite sets associated with SNPs (*currently contains 4598 entries*)

- Predicted metabolite sets based on computational enzyme knockout model (*currently contains 912 entries*)

- Metabolite sets based on locations (*currently contains 73 entries*)

- Drug pathway associated metabolite sets (*currently contains 461 entries*)

In addition, MSEA also allows user-defined metabolite sets to be uploaded to perform enrichment analysis on arbitrary groups of compounds which researchers want to test. The metabolite set library is simply a two-column comma separated text file with the first column for metabolite set names and the second column for its compound names (**must use HMDB compound name**) separated by ”; ”. Please note, the built-in libraries are mainly from human studies. The functional grouping of metabolites may not be valid. Therefore, for data from subjects other than human being, users are suggested to upload their self-defined metabolite set libraries for enrichment analysis.

6 Enrichment Analysis

Over Representation Analysis (ORA) is performed when a list of compound names is provided. The list of compound list can be obtained through conventional feature selection methods, or from a clustering algorithm, or from the compounds with abnormal concentrations detected in SSP, to investigate if some biologically meaningful patterns can be identified.

ORA was implemented using the *hypergeometric test* to evaluate whether a particular metabolite set is represented more than expected by chance within the given compound list. One-tailed p values are provided after adjusting for multiple testing. **Figure 2** below summarizes the result.
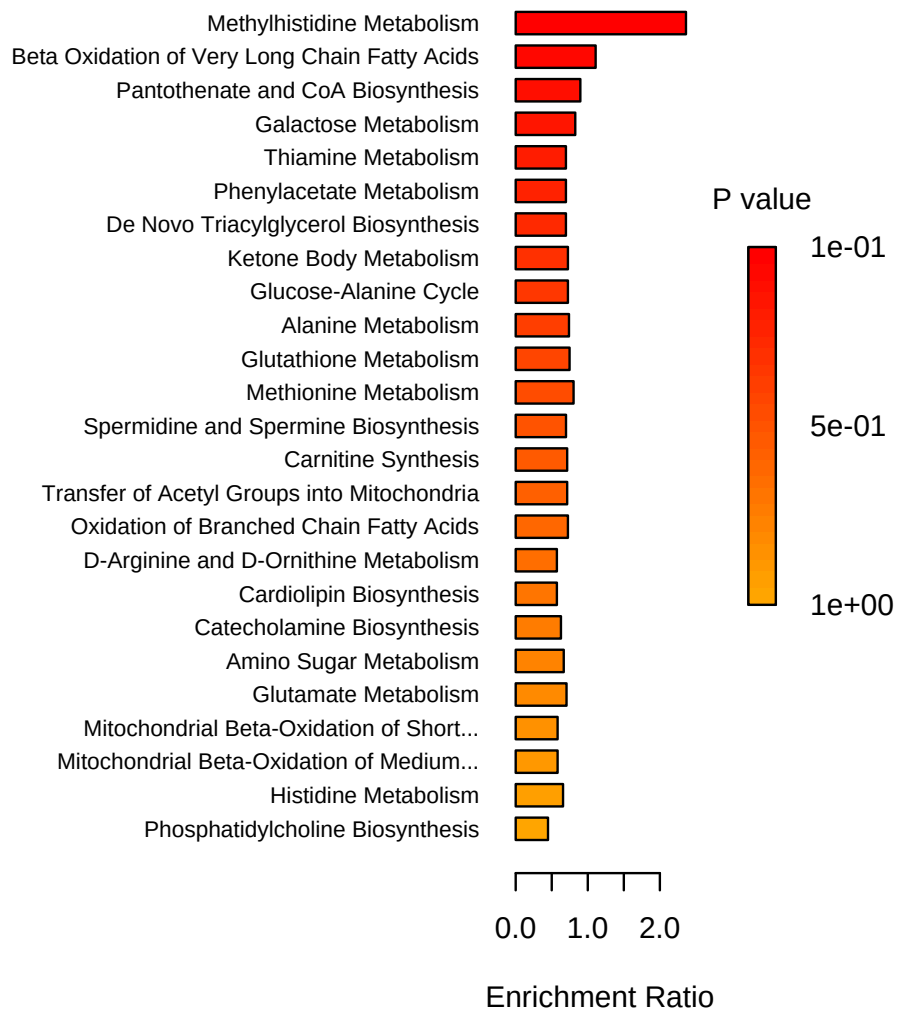
Enrichment Overview (top 25)

Figure 1: Summary Plot for Over Representation Analysis (ORA)

Table 2: Result from Over Representation Analysis

	total	expected	hits	Raw p	Holm p	FDR
Methylhistidine Metabolism	4	1.27	3	9.78E-02	1.00E+00	1.00E+00
Beta Oxidation of Very Long Chain Fatty Acids	17	5.41	6	4.69E-01	1.00E+00	1.00E+00
Pantothenate and CoA Biosynthesis	21	6.69	6	7.05E-01	1.00E+00	1.00E+00
Galactose Metabolism	38	12.10	10	8.21E-01	1.00E+00	1.00E+00
Thiamine Metabolism	9	2.87	2	8.36E-01	1.00E+00	1.00E+00
Phenylacetate Metabolism	9	2.87	2	8.36E-01	1.00E+00	1.00E+00
De Novo Triacylglycerol Biosynthesis	9	2.87	2	8.36E-01	1.00E+00	1.00E+00
Ketone Body Metabolism	13	4.14	3	8.36E-01	1.00E+00	1.00E+00
Glucose-Alanine Cycle	13	4.14	3	8.36E-01	1.00E+00	1.00E+00
Alanine Metabolism	17	5.41	4	8.42E-01	1.00E+00	1.00E+00
Glutathione Metabolism	21	6.69	5	8.50E-01	1.00E+00	1.00E+00
Methionine Metabolism	43	13.70	11	8.58E-01	1.00E+00	1.00E+00
Spermidine and Spermine Biosynthesis	18	5.73	4	8.75E-01	1.00E+00	1.00E+00
Carnitine Synthesis	22	7.00	5	8.79E-01	1.00E+00	1.00E+00
Transfer of Acetyl Groups into Mitochondria	22	7.00	5	8.79E-01	1.00E+00	1.00E+00
Oxidation of Branched Chain Fatty Acids	26	8.28	6	8.84E-01	1.00E+00	1.00E+00
D-Arginine and D-Ornithine Metabolism	11	3.50	2	9.11E-01	1.00E+00	1.00E+00
Cardiolipin Biosynthesis	11	3.50	2	9.11E-01	1.00E+00	1.00E+00
Catecholamine Biosynthesis	20	6.37	4	9.23E-01	1.00E+00	1.00E+00
Amino Sugar Metabolism	33	10.50	7	9.41E-01	1.00E+00	1.00E+00
Glutamate Metabolism	49	15.60	11	9.49E-01	1.00E+00	1.00E+00
Mitochondrial Beta-Oxidation of Short Chain Saturated Fatty Acids	27	8.60	5	9.63E-01	1.00E+00	1.00E+00
Mitochondrial Beta-Oxidation of Medium Chain Saturated Fatty Acids	27	8.60	5	9.63E-01	1.00E+00	1.00E+00
Histidine Metabolism	43	13.70	9	9.63E-01	1.00E+00	1.00E+00
Phosphatidylcholine Biosynthesis	14	4.46	2	9.66E-01	1.00E+00	1.00E+00
Butyrate Metabolism	19	6.05	3	9.69E-01	1.00E+00	1.00E+00
Ethanol Degradation	19	6.05	3	9.69E-01	1.00E+00	1.00E+00
Homocysteine Degradation	9	2.87	1	9.69E-01	1.00E+00	1.00E+00
Lactose Degradation	9	2.87	1	9.69E-01	1.00E+00	1.00E+00
Sphingolipid Metabolism	40	12.70	8	9.69E-01	1.00E+00	1.00E+00
Mitochondrial Beta-Oxidation of Long Chain Saturated Fatty Acids	28	8.91	5	9.71E-01	1.00E+00	1.00E+00
Lactose Synthesis	20	6.37	3	9.77E-01	1.00E+00	1.00E+00
Threonine and 2-Oxobutanoate Degradation	20	6.37	3	9.77E-01	1.00E+00	1.00E+00
Pentose Phosphate Pathway	29	9.23	5	9.78E-01	1.00E+00	1.00E+00
Pyruvaldehyde Degradation	10	3.18	1	9.79E-01	1.00E+00	1.00E+00
Glycolysis	25	7.96	4	9.80E-01	1.00E+00	1.00E+00
Propanoate Metabolism	42	13.40	8	9.80E-01	1.00E+00	1.00E+00
Beta-Alanine Metabolism	34	10.80	6	9.81E-01	1.00E+00	1.00E+00
Betaine Metabolism	21	6.69	3	9.83E-01	1.00E+00	1.00E+00
Phytanic Acid Peroxisomal Oxidation	26	8.28	4	9.85E-01	1.00E+00	1.00E+00
Glycine and Serine Metabolism	59	18.80	12	9.85E-01	1.00E+00	1.00E+00
Glycerol Phosphate Shuttle	11	3.50	1	9.86E-01	1.00E+00	1.00E+00
Phosphatidylinositol Phosphate Metabolism	17	5.41	2	9.87E-01	1.00E+00	1.00E+00
Phosphatidylethanolamine Biosynthesis	12	3.82	1	9.90E-01	1.00E+00	1.00E+00
Nicotinate and Nicotinamide Metabolism	37	11.80	6	9.91E-01	1.00E+00	1.00E+00
Thyroid hormone synthesis	13	4.14	1	9.93E-01	1.00E+00	1.00E+00
Urea Cycle	29	9.23	4	9.94E-01	1.00E+00	1.00E+00
Glycerolipid Metabolism	25	7.96	3	9.95E-01	1.00E+00	1.00E+00
Nucleotide Sugars Metabolism	20	6.37	2	9.95E-01	1.00E+00	1.00E+00
Gluconeogenesis	35	11.10	5	9.96E-01	1.00E+00	1.00E+00
Fatty Acid Biosynthesis	35	11.10	5	9.96E-01	1.00E+00	1.00E+00
Cysteine Metabolism	26	8.28	3	9.96E-01	1.00E+00	1.00E+00
Starch and Sucrose Metabolism	31	9.87	4	9.96E-01	1.00E+00	1.00E+00
Warburg Effect	58	18.50	10	9.97E-01	1.00E+00	1.00E+00
Ammonia Recycling	32	10.20	4	9.97E-01	1.00E+00	1.00E+00
Citric Acid Cycle	32	10.20	4	9.97E-01	1.00E+00	1.00E+00
Fructose and Mannose Degradation	32	10.20	4	9.97E-01	1.00E+00	1.00E+00
Retinol Metabolism	37	11.80	5	9.98E-01	1.00E+00	1.00E+00
Sulfate/Sulfite Metabolism	22	7.00	2	9.98E-01	1.00E+00	1.00E+00
Selenoamino Acid Metabolism	28	8.91	3	9.98E-01	1.00E+00	1.00E+00
Fatty acid Metabolism	43	13.70	6	9.98E-01	1.00E+00	1.00E+00
Pyruvate Metabolism	48	15.30	7	9.99E-01	1.00E+00	1.00E+00
Arginine and Proline Metabolism	53	16.90	8	9.99E-01	1.00E+00	1.00E+00
Caffeine Metabolism	24	7.64	2	9.99E-01	1.00E+00	1.00E+00
Estrone Metabolism	24	7.64	2	9.99E-01	1.00E+00	1.00E+00
Mitochondrial Electron Transport Chain	19	6.05	1	9.99E-01	1.00E+00	1.00E+00
Inositol Phosphate Metabolism	26	8.28	2	9.99E-01	1.00E+00	1.00E+00
Bile Acid Biosynthesis	65	20.70	10	9.99E-01	1.00E+00	1.00E+00
Vitamin B6 Metabolism	20	6.37	1	1.00E+00	1.00E+00	1.00E+00
Ubiquinone Biosynthesis	20	6.37	1	1.00E+00	1.00E+00	1.00E+00
Riboflavin Metabolism	20	6.37	1	1.00E+00	1.00E+00	1.00E+00
Inositol Metabolism	33	10.50	3	1.00E+00	1.00E+00	1.00E+00
Phenylalanine and Tyrosine Metabolism	28	8.91	2	1.00E+00	1.00E+00	1.00E+00
Phospholipid Biosynthesis	29	9.23	2	1.00E+00	1.00E+00	1.00E+00
Folate Metabolism	29	9.23	2	1.00E+00	1.00E+00	1.00E+00
Valine, Leucine and Isoleucine Degradation	60	19.10	8	1.00E+00	1.00E+00	1.00E+00

Fatty Acid Elongation In Mitochondria	35	11.10	3	1.00E+00	1.00E+00	1.00E+00
Aspartate Metabolism	35	11.10	3	1.00E+00	1.00E+00	1.00E+00
Lysine Degradation	30	9.55	2	1.00E+00	1.00E+00	1.00E+00
Tryptophan Metabolism	60	19.10	7	1.00E+00	1.00E+00	1.00E+00
Androgen and Estrogen Metabolism	33	10.50	2	1.00E+00	1.00E+00	1.00E+00
Plasmalogen Synthesis	26	8.28	1	1.00E+00	1.00E+00	1.00E+00
Pterine Biosynthesis	29	9.23	1	1.00E+00	1.00E+00	1.00E+00
Porphyrin Metabolism	40	12.70	2	1.00E+00	1.00E+00	1.00E+00
Purine Metabolism	74	23.60	8	1.00E+00	1.00E+00	1.00E+00
Pyrimidine Metabolism	59	18.80	5	1.00E+00	1.00E+00	1.00E+00
Steroid Biosynthesis	48	15.30	2	1.00E+00	1.00E+00	1.00E+00
Steroidogenesis	43	13.70	1	1.00E+00	1.00E+00	1.00E+00
Tyrosine Metabolism	72	22.90	4	1.00E+00	1.00E+00	1.00E+00
Arachidonic Acid Metabolism	69	22.00	1	1.00E+00	1.00E+00	1.00E+00

7 Appendix: R Command History

```
 [1] "mSet<-InitDataObjects(\"conc\", \"msetora\", FALSE)"
 [2] "cmpd.vec<-c(\"HMDB0000169\",\"HMDB0000190\",\"HMDB0000122\",\"HMDB0029965\",\"HMDB0000125\",\"
 [3] "mSet<-Setup.MapData(mSet, cmpd.vec);"
 [4] "mSet<-CrossReferencing(mSet, \"hmdb\");"
 [5] "mSet<-CreateMappingResultTable(mSet)"
 [6] "mSet<-SetMetabolomeFilter(mSet, F);"
 [7] "mSet<-SetCurrentMsetLib(mSet, \"smpdb_pathway\", 2);"
 [8] "mSet<-CalculateHyperScore(mSet)"
 [9] "mSet<-PlotORA(mSet, \"ora_0_\", \"net\", \"png\", 72, width=NA)"
[10] "mSet<-PlotEnrichDotPlot(mSet, \"ora\", \"ora_dot_0_\", \"png\", 72, width=NA)"
[11] "mSet<-CalculateHyperScore(mSet)"
[12] "mSet<-PlotORA(mSet, \"ora_1_\", \"net\", \"png\", 72, width=NA)"
[13] "mSet<-PlotEnrichDotPlot(mSet, \"ora\", \"ora_dot_1_\", \"png\", 72, width=NA)"
[14] "mSet<-SaveTransformedData(mSet)"
[15] "mSet<-PreparePDFReport(mSet, \"guest16292574549826598929\")\n"
```

The report was generated on Sat Mar 20 21:24:04 2021 with R version 4.0.2 (2020-06-22).