

**ENHANCED DETECTION AND ANNOTATION OF SMALL MOLECULES IN
METABOLOMICS USING MOLECULAR NETWORK-ORIENTED PARAMETER
OPTIMIZATION**

SUPPORTING INFORMATION

Running title: Parameter optimization for molecular network-based compound ID

¹Rui Xu, ¹Jisun Lee, ¹Li Chen, ^{1,2}Jiangjiang Zhu[#]

Author affiliations:

¹Human Nutrition Program, The Ohio State University, Columbus, Ohio 43210

²James Comprehensive Cancer Center, The Ohio State University, Columbus, Ohio 43210

[#]To whom correspondence should be addressed to:

Jiangjiang (Chris) Zhu, Ph.D., Comprehensive Cancer Center and Department of Human Sciences, The Ohio State University, 400 W 12th Ave, Columbus, OH 43210

Tel: 614-685-2226, Email: zhu.2484@osu.edu

Key words: Metabolomics, molecular network, parameter optimization, mass spectrometry

Grant Support: Research reported in this publication was supported by the National Institute of General Medical Sciences of the National Institutes of Health under Award Number R35GM133510.

Figure S1. An example of molecular network and corresponding mass spectrum at m/z 449.101.

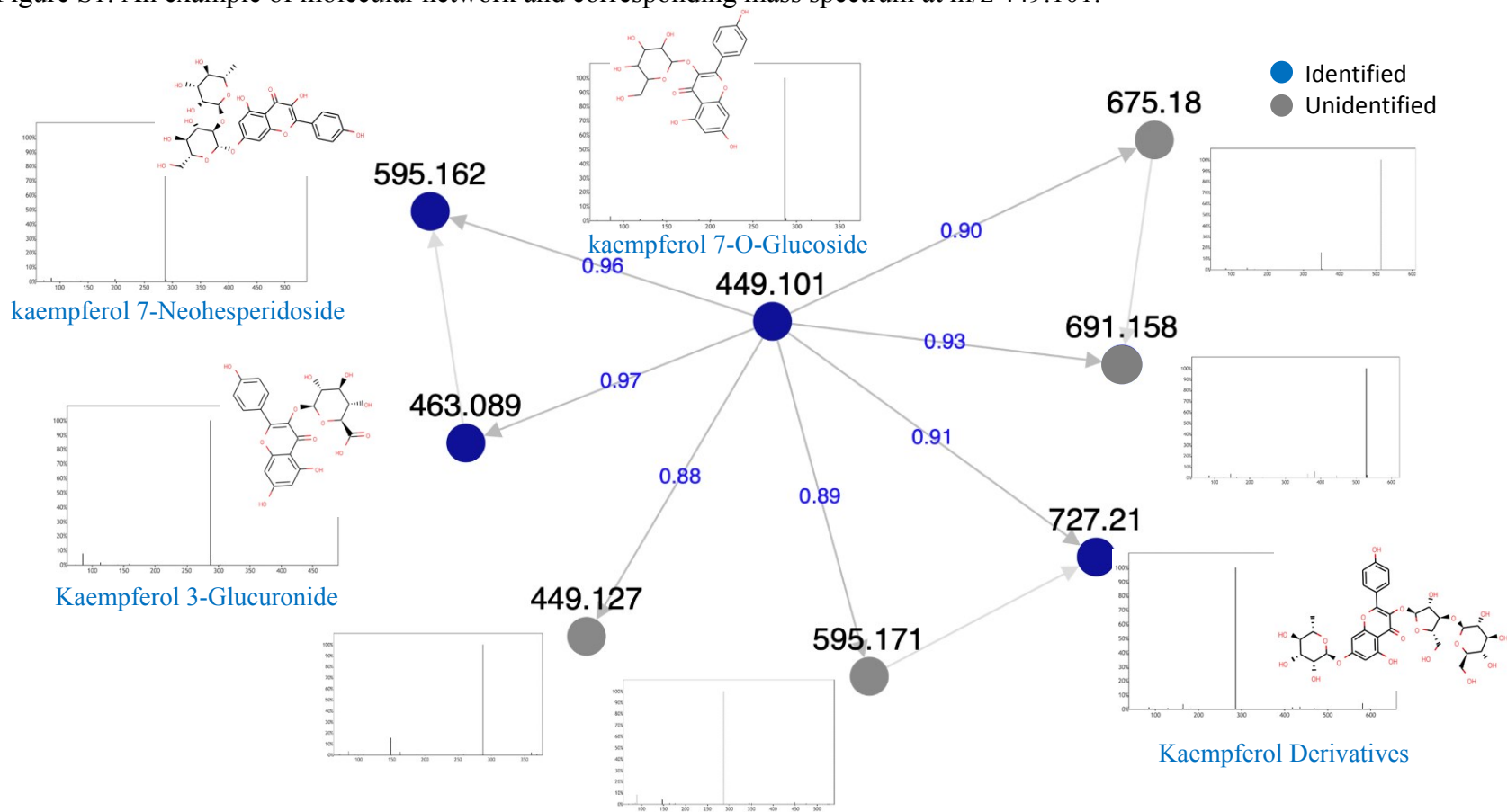


Figure S2. Average statistical values for molecular network with BRB Powder raw data generated under different resolution, NCE, intensity threshold, and exclusion time. TN: total nodes; IN: identified nodes; NN: nodes in network; SN: self-loop nodes; P/E: pairs/edges; CS: cosine scores; NW: networks; CC: cluster coefficients.

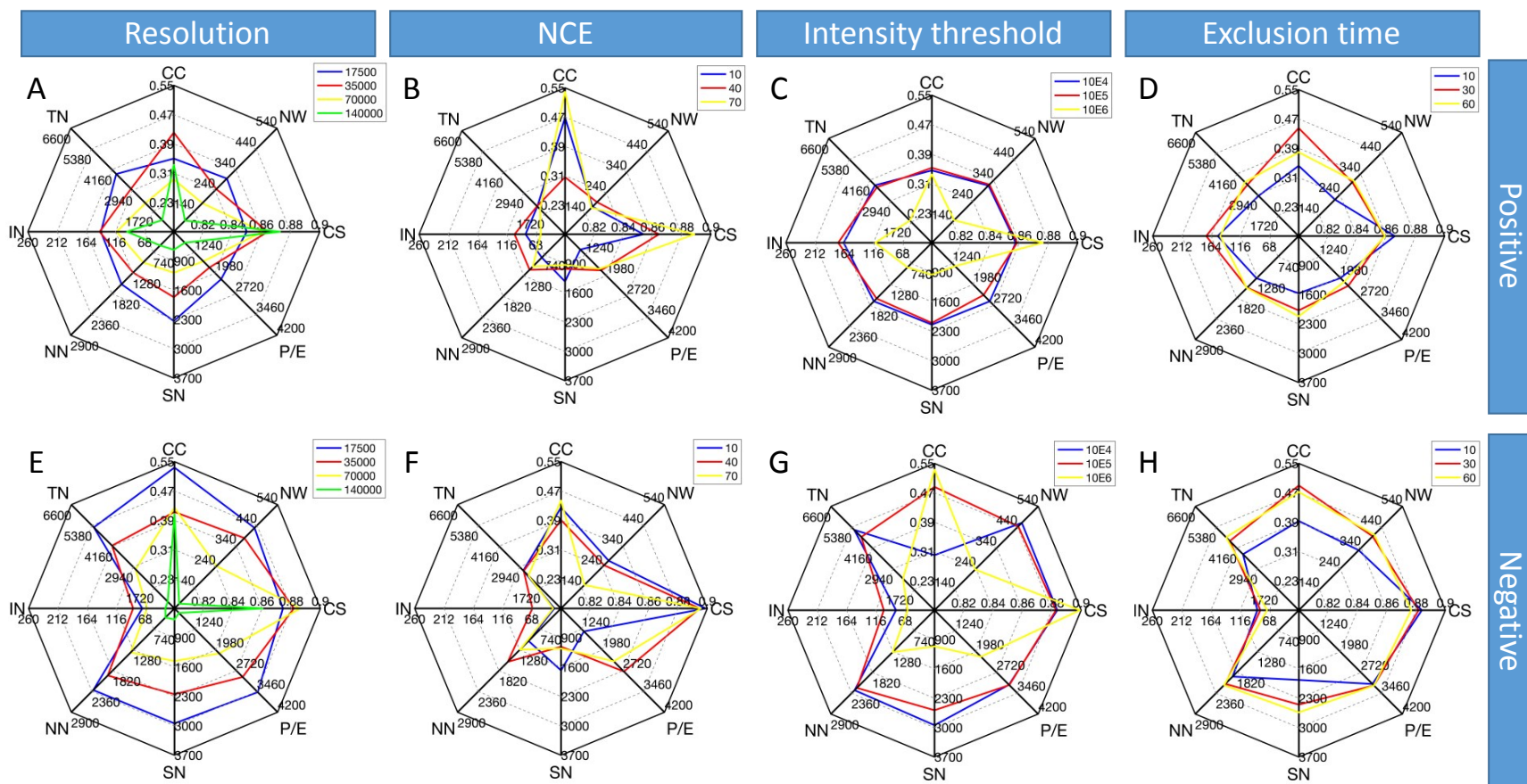


Figure S3. The intensity of the largest cluster of each MN under different parameter settings using powder BRB extract.

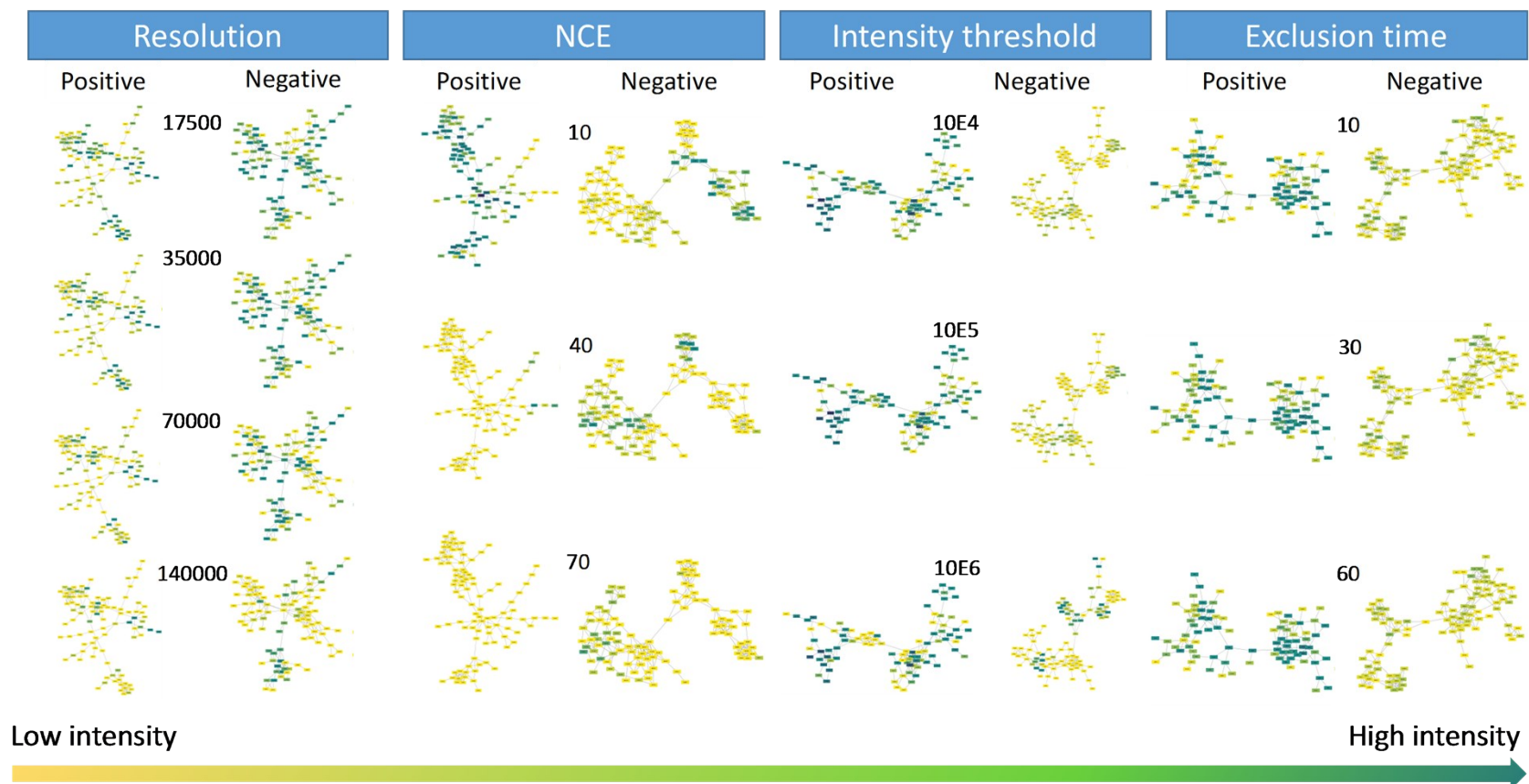


Figure S4. The node origins of the largest cluster of each MN under different parameter settings using powder BRB extract.

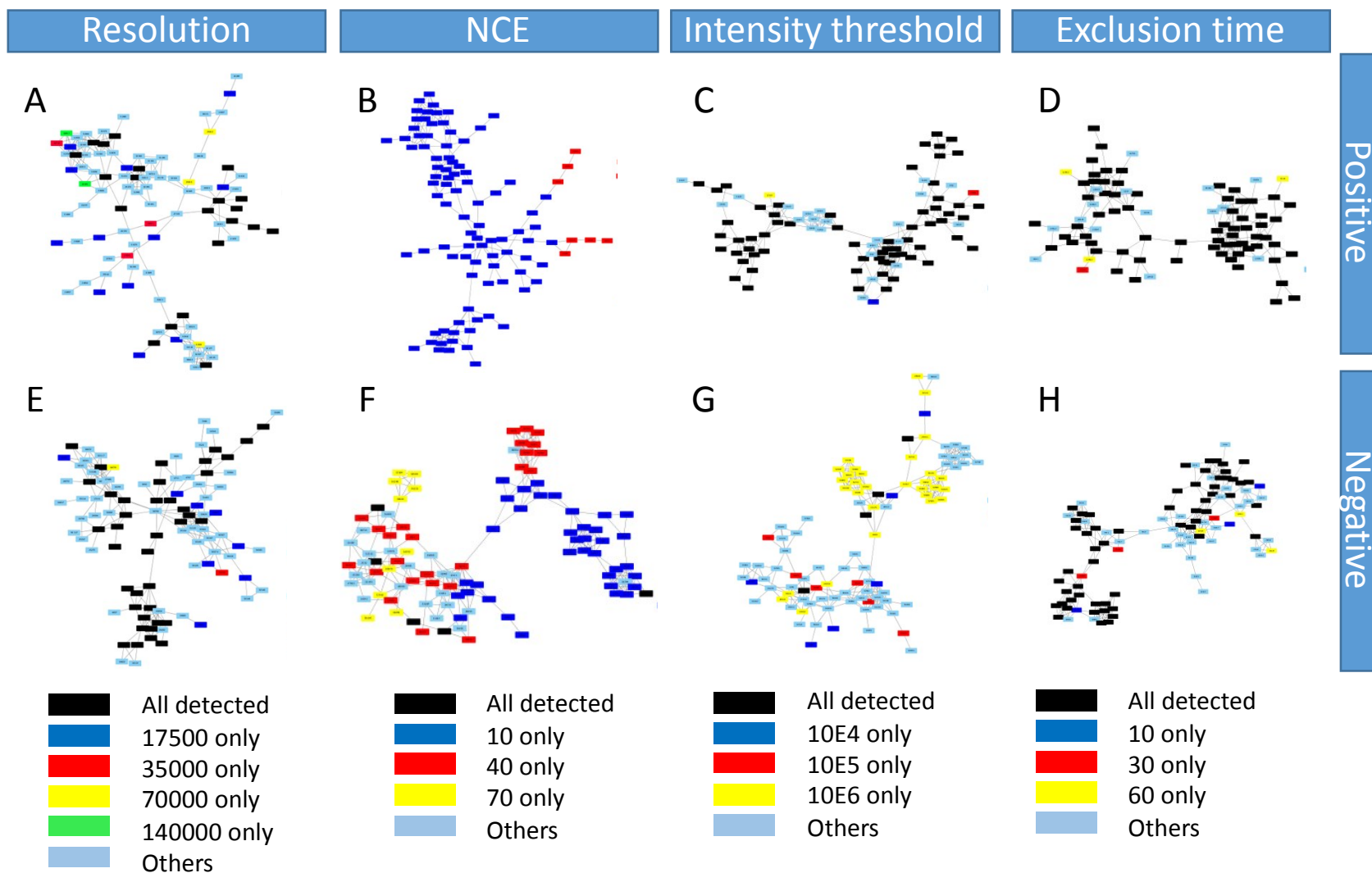


Figure S5. The node origins distribution of each MN under different parameter settings using powder BRB extract.

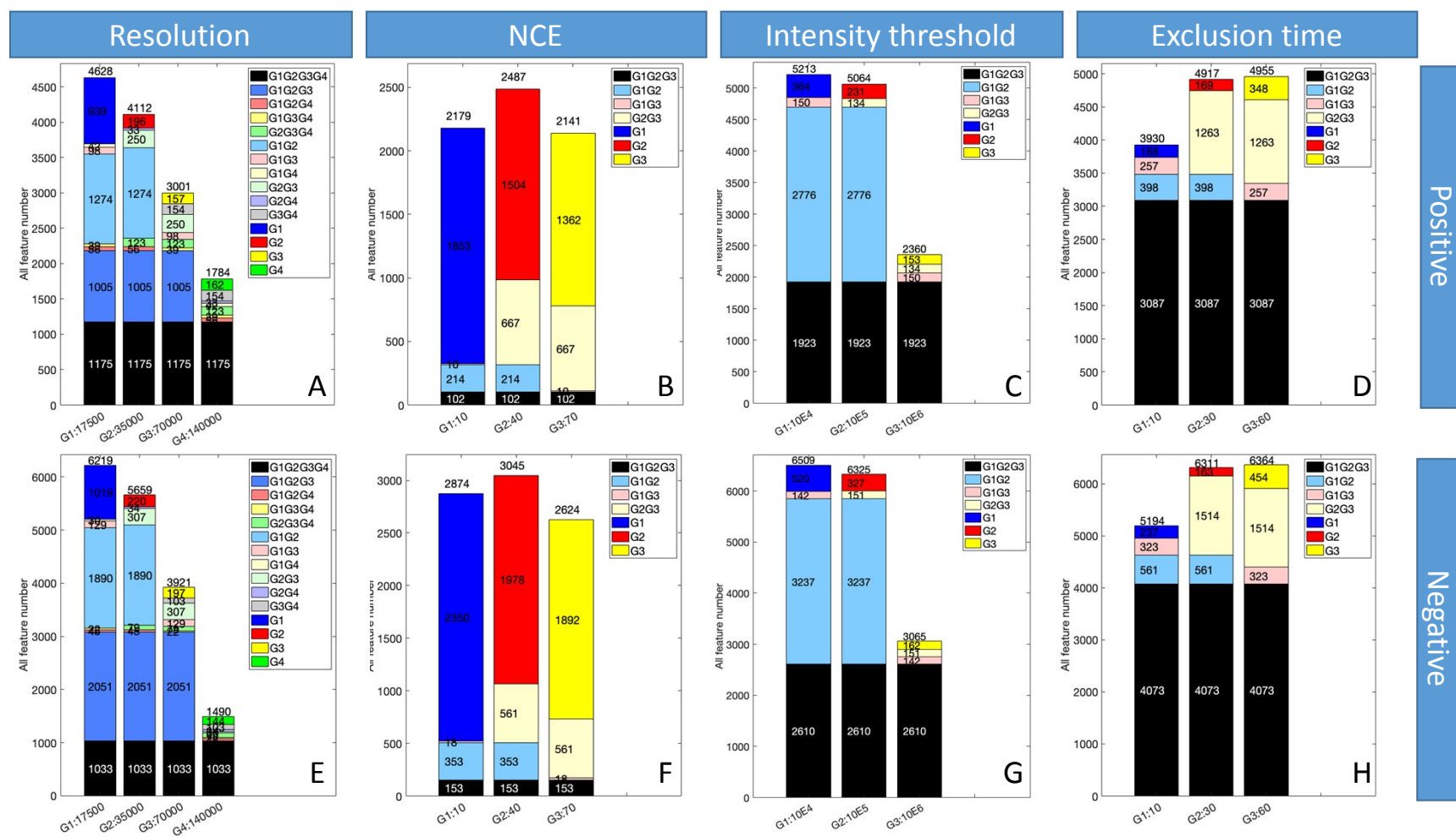


Figure S6. A. The annotated features under both polarities with both BRB samples.

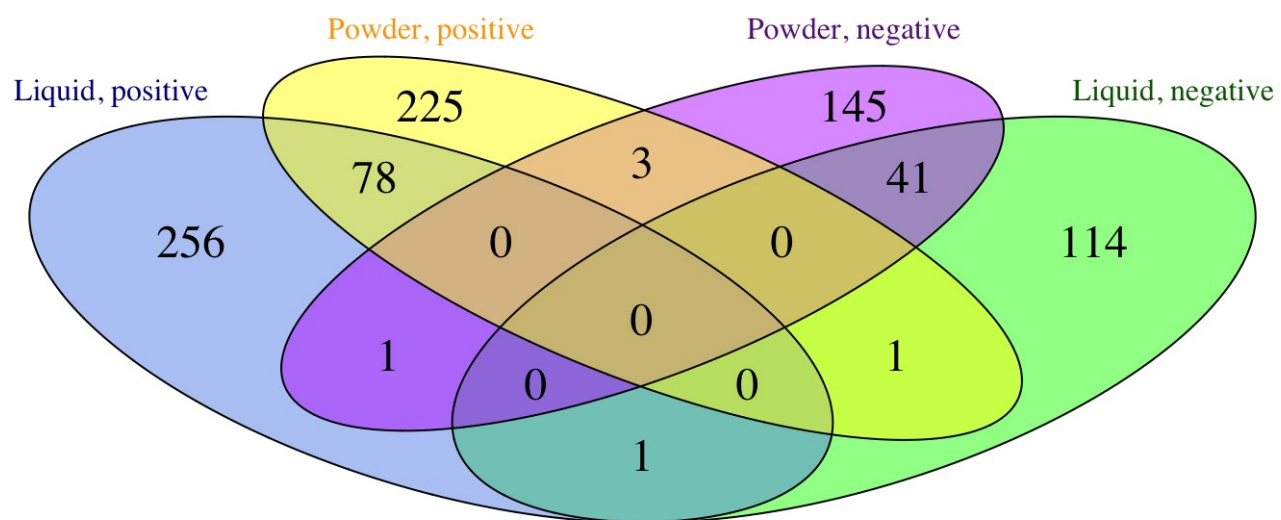


Table S1. A list of the different combinations of parameter settings for LC-MS experiments.

No.	Resolution	NCE	Intensity threshold	Exclusion time
1	17,500	10	1.00E+05	10
2	17,500	10	1.00E+05	30
3	17,500	10	1.00E+05	60
4	17,500	10	1.00E+06	10
5	17,500	10	1.00E+06	30
6	17,500	10	1.00E+06	60
7	17,500	10	1.00E+07	10
8	17,500	10	1.00E+07	30
9	17,500	10	1.00E+07	60
10	17,500	40	1.00E+05	10
11	17,500	40	1.00E+05	30
12	17,500	40	1.00E+05	60
13	17,500	40	1.00E+06	10
14	17,500	40	1.00E+06	30
15	17,500	40	1.00E+06	60
16	17,500	40	1.00E+07	10
17	17,500	40	1.00E+07	30
18	17,500	40	1.00E+07	60
19	17,500	70	1.00E+05	10
20	17,500	70	1.00E+05	30
21	17,500	70	1.00E+05	60
22	17,500	70	1.00E+06	10
23	17,500	70	1.00E+06	30
24	17,500	70	1.00E+06	60
25	17,500	70	1.00E+07	10
26	17,500	70	1.00E+07	30
27	17,500	70	1.00E+07	60

28	35,000	10	1.00E+05	10
29	35,000	10	1.00E+05	30
30	35,000	10	1.00E+05	60
31	35,000	10	1.00E+06	10
32	35,000	10	1.00E+06	30
33	35,000	10	1.00E+06	60
34	35,000	10	1.00E+07	10
35	35,000	10	1.00E+07	30
36	35,000	10	1.00E+07	60
37	35,000	40	1.00E+05	10
38	35,000	40	1.00E+05	30
39	35,000	40	1.00E+05	60
40	35,000	40	1.00E+06	10
41	35,000	40	1.00E+06	30
42	35,000	40	1.00E+06	60
43	35,000	40	1.00E+07	10
44	35,000	40	1.00E+07	30
45	35,000	40	1.00E+07	60
46	35,000	70	1.00E+05	10
47	35,000	70	1.00E+05	30
48	35,000	70	1.00E+05	60
49	35,000	70	1.00E+06	10
50	35,000	70	1.00E+06	30
51	35,000	70	1.00E+06	60
52	35,000	70	1.00E+07	10
53	35,000	70	1.00E+07	30
54	35,000	70	1.00E+07	60
55	70,000	10	1.00E+05	10
56	70,000	10	1.00E+05	30

57	70,000	10	1.00E+05	60
58	70,000	10	1.00E+06	10
59	70,000	10	1.00E+06	30
60	70,000	10	1.00E+06	60
61	70,000	10	1.00E+07	10
62	70,000	10	1.00E+07	30
63	70,000	10	1.00E+07	60
64	70,000	40	1.00E+05	10
65	70,000	40	1.00E+05	30
66	70,000	40	1.00E+05	60
67	70,000	40	1.00E+06	10
68	70,000	40	1.00E+06	30
69	70,000	40	1.00E+06	60
70	70,000	40	1.00E+07	10
71	70,000	40	1.00E+07	30
72	70,000	40	1.00E+07	60
73	70,000	70	1.00E+05	10
74	70,000	70	1.00E+05	30
75	70,000	70	1.00E+05	60
76	70,000	70	1.00E+06	10
77	70,000	70	1.00E+06	30
78	70,000	70	1.00E+06	60
79	70,000	70	1.00E+07	10
80	70,000	70	1.00E+07	30
81	70,000	70	1.00E+07	60
82	140,000	10	1.00E+05	10
83	140,000	10	1.00E+05	30
84	140,000	10	1.00E+05	60
85	140,000	10	1.00E+06	10

86	140,000	10	1.00E+06	30
87	140,000	10	1.00E+06	60
88	140,000	10	1.00E+07	10
89	140,000	10	1.00E+07	30
90	140,000	10	1.00E+07	60
91	140,000	40	1.00E+05	10
92	140,000	40	1.00E+05	30
93	140,000	40	1.00E+05	60
94	140,000	40	1.00E+06	10
95	140,000	40	1.00E+06	30
96	140,000	40	1.00E+06	60
97	140,000	40	1.00E+07	10
98	140,000	40	1.00E+07	30
99	140,000	40	1.00E+07	60
100	140,000	70	1.00E+05	10
101	140,000	70	1.00E+05	30
102	140,000	70	1.00E+05	60
103	140,000	70	1.00E+06	10
104	140,000	70	1.00E+06	30
105	140,000	70	1.00E+06	60
106	140,000	70	1.00E+07	10
107	140,000	70	1.00E+07	30
108	140,000	70	1.00E+07	60

Table S2. The detailed data for radar map of liquid BRB extract analysis in Figure 2.

Modes	Indicators	Resolution				NCE			Intensity threshold			Exclusion time		
		17,500	35,000	70,000	140,000	10	40	70	10E4	10E5	10E6	10	30	60
Positive mode	total nodes	3432	3208	5546	1157	6512	2264	1947	3763	3813	1406	2782	3383	3611
	identified nodes	112	121	223	74	252	112	58	149	154	107	133	140	151
	nodes in networks	1355	1331	2360	490	2875	1165	1022	1583	1533	685	1239	1405	1490
	self-loop nodes	2077	1877	3186	667	3637	1099	925	2180	2280	721	1543	1978	2121
	pairs/edges	1878	1969	3350	732	4194	1901	1799	2241	2231	1119	1919	2113	2157
	cosine score	0.83	0.84	0.86	0.82	0.87	0.85	0.88	0.85	0.85	0.84	0.84	0.84	0.84
	networks	262	236	451	70	530	190	147	298	269	112	191	233	268
	cluster coefficient	0.35	0.26	0.35	0.26	0.34	0.36	0.23	0.26	0.39	0.34	0.32	0.31	0.32
Negative mode	total nodes	3143	2769	1940	1165	5074	1955	1773	3822	2737	541	2525	2844	3015
	identified nodes	66	67	52	49	131	50	46	77	76	31	70	75	73
	nodes in networks	1361	1225	902	505	2361	980	903	1716	1299	318	1187	1282	1360
	self-loop nodes	1782	1544	1038	660	2713	975	870	2106	1438	223	1338	1562	1655
	pairs/edges	2006	1687	1333	682	3405	1565	1732	2352	1957	569	1783	1838	1916
	cosine score	0.86	0.87	0.87	0.87	0.89	0.89	0.89	0.88	0.87	0.86	0.88	0.88	0.87
	networks	227	221	149	100	472	159	110	323	231	48	207	224	252
	cluster coefficient	0.38	0.34	0.24	0.21	0.20	0.38	0.37	0.40	0.36	0.33	0.16	0.22	0.49

Table S3. The detailed data for radar map of powder BRB extract analysis in Figure S2.

Modes	Indicators	Resolution				NCE			Intensity threshold			Exclusion time		
		17,500	35,000	70,000	140,000	10	40	70	10E4	10E5	10E6	10	30	60
Positive mode	total nodes	3903	3031	2182	1189	2137	2154	1975	3860	3747	1796	2870	3525	3676
	identified nodes	140	142	114	97	85	103	60	165	174	113	153	172	149
	nodes in networks	1569	1266	1001	558	808	1127	1042	1716	1645	837	1302	1547	1556
	self-loop nodes	2334	1765	1181	631	1329	1027	933	2144	2102	959	1568	1978	2120
	pairs/edges	2196	1778	1557	893	1046	1790	1736	2577	2355	1331	1999	2279	2136
	cosine score	0.85	0.86	0.86	0.87	0.85	0.86	0.89	0.86	0.86	0.88	0.87	0.86	0.86
	networks	297	246	178	94	173	195	168	314	319	148	216	300	307
	cluster coefficient	0.35	0.42	0.29	0.33	0.47	0.31	0.54	0.35	0.35	0.34	0.34	0.45	0.38
Negative mode	total nodes	5267	4192	2800	900	2725	2702	2427	5223	4806	2353	3789	4566	4786
	identified nodes	73	88	65	35	31	67	33	83	103	64	88	83	72
	nodes in networks	2324	1945	1346	437	1047	1590	1271	2281	2221	1291	1926	2115	2149
	self-loop nodes	2943	2247	1454	463	1678	1112	1156	2942	2585	1062	1863	2451	2637
	pairs/edges	3485	2945	2105	662	1319	2749	2385	3157	3168	2157	3129	3183	3185
	cosine score	0.87	0.88	0.89	0.86	0.90	0.89	0.90	0.88	0.88	0.90	0.88	0.88	0.88
	networks	428	379	244	63	270	248	153	461	444	236	329	396	404
	cluster coefficient	0.53	0.41	0.43	0.41	0.43	0.39	0.44	0.30	0.49	0.53	0.39	0.49	0.47

Table S4. The optimal parameters for both samples under both polarities.

Parameters	Positive Mode		Negative Mode	
	Liquid BRB	Powder BRB	Liquid BRB	Powder BRB
Resolution	70,000	35,000	17,500	17,500
NCE	10	70	10	40
Intensity threshold	10E5	10E5	10E4	10E5
Exclusion time	60	30	60	60

Data deposition

The mass spectra data has been deposited to MassIVE database (<https://massive.ucsd.edu/ProteoSAFe/static/massive.jsp>) with access # MSV000087069.

All GNPS molecular networking jobs can be found at the links below:

BRB_Liquid_pos_resolution_17, 500

<https://gnps.ucsd.edu/ProteoSAFe/status.jsp?task=09eb2a325c2f45c3bbcc55d1d0939227>

BRB_Liquid_pos_resolution_35, 000

<https://gnps.ucsd.edu/ProteoSAFe/status.jsp?task=b4f9c9a7e7f24a0e873c891c86af7c15>

BRB_Liquid_pos_resolution_70, 000

<https://gnps.ucsd.edu/ProteoSAFe/status.jsp?task=32f2af1dc69b4db19e065a8cff06ee4>

BRB_Liquid_pos_resolution_140, 000

<https://gnps.ucsd.edu/ProteoSAFe/status.jsp?task=2e01603972844c71b78444654feb87f>

BRB_Liquid_pos_NCE_10

<https://gnps.ucsd.edu/ProteoSAFe/status.jsp?task=85ac6971a7254fbfa0ff94c4074b8a25>

BRB_Liquid_pos_NCE_40

<https://gnps.ucsd.edu/ProteoSAFe/status.jsp?task=580b67740f6f4a938686c2eeba508b36>

BRB_Liquid_pos_NCE_70

<https://gnps.ucsd.edu/ProteoSAFe/status.jsp?task=438ceda0fc52463cae8e9703f4b78338>

BRB_Liquid_pos_intensity_10e4

<https://gnps.ucsd.edu/ProteoSAFe/status.jsp?task=928a140a34614d81b6d0317c8e8ff9b1>

BRB_Liquid_pos_intensity_10e5

<https://gnps.ucsd.edu/ProteoSAFe/status.jsp?task=f319ff2b6f9b4528871004bd59350263>

BRB_Liquid_pos_intensity_10e6

<https://gnps.ucsd.edu/ProteoSAFe/status.jsp?task=225e518b2aef42d9b8c51124c99395d6>

BRB_Liquid_pos_exclusion_10

<https://gnps.ucsd.edu/ProteoSAFe/status.jsp?task=c3b1bc752e224db5b79166cd74d2507b>

BRB_Liquid_pos_exclusion_30

<https://gnps.ucsd.edu/ProteoSAFe/status.jsp?task=a3a8dc65714440f2ba25a76d9015a868>

BRB_Liquid_pos_exclusion_60

<https://gnps.ucsd.edu/ProteoSAFe/status.jsp?task=02e5ca80d4ab429cbafb50e311dd1f6b>

BRB_Liquid_neg_resolution_17, 500

<https://gnps.ucsd.edu/ProteoSAFe/status.jsp?task=73b2f34cbcb145b9b50627d4b27f8f07>

BRB_Liquid_neg_resolution_35, 000

<https://gnps.ucsd.edu/ProteoSAFe/status.jsp?task=40a4caa9ae274c8398df2b16d47f1110>
BRB_Liquid_neg_resolution_70,000

<https://gnps.ucsd.edu/ProteoSAFe/status.jsp?task=93f5897db76c47c4a7c0d122af90f6b9>
BRB_Liquid_neg_resolution_140,000

<https://gnps.ucsd.edu/ProteoSAFe/status.jsp?task=a9cdc5cdd5964e1ea128eae536791ba3>
BRB_Liquid_neg_NCE_10

<https://gnps.ucsd.edu/ProteoSAFe/status.jsp?task=24a51f26c18146fc95594b51f4c22f88>
BRB_Liquid_neg_NCE_40

<https://gnps.ucsd.edu/ProteoSAFe/status.jsp?task=593cda348d384550b49c3e4ee10843a4>
BRB_Liquid_neg_NCE_70

<https://gnps.ucsd.edu/ProteoSAFe/status.jsp?task=47c62915c7734b5abca9ef212f725a33>
BRB_Liquid_neg_intensity_10e4

<https://gnps.ucsd.edu/ProteoSAFe/status.jsp?task=f6b53e0cbcef4700b222c9bf0a6b924d>
BRB_Liquid_neg_intensity_10e5

<https://gnps.ucsd.edu/ProteoSAFe/status.jsp?task=7ffa52544ef74a68b6ff3e78d6288bcb>
BRB_Liquid_neg_intensity_10e6

<https://gnps.ucsd.edu/ProteoSAFe/status.jsp?task=84192dd5feaf47bc8014abf5949c67a1>
BRB_Liquid_neg_exclusion_10

<https://gnps.ucsd.edu/ProteoSAFe/status.jsp?task=0877fd0d4a4645acb094b5a71dd5d1f7>
BRB_Liquid_neg_exclusion_30

<https://gnps.ucsd.edu/ProteoSAFe/status.jsp?task=da6b188a47ef47dfa77631e6fb16c790>
BRB_Liquid_neg_exclusion_60

<https://gnps.ucsd.edu/ProteoSAFe/status.jsp?task=9b7950bc977048b7988c05073a396314>
BRB_Powder_pos_resolution_17,500

<https://gnps.ucsd.edu/ProteoSAFe/status.jsp?task=5c0890b14c454edcbbb1edb7f0fe2f4b>
BRB_Powder_pos_resolution_35,000

<https://gnps.ucsd.edu/ProteoSAFe/status.jsp?task=7dbcee880bf344a7bd386c8df3653fcb>
BRB_Powder_pos_resolution_70,000

<https://gnps.ucsd.edu/ProteoSAFe/status.jsp?task=b594ddb054f843c6bc93205aafca536c>
BRB_Powder_pos_resolution_140,000

<https://gnps.ucsd.edu/ProteoSAFe/status.jsp?task=241f4a05b7774cab863ff2a8d031626d>
BRB_Powder_pos_NCE_10

<https://gnps.ucsd.edu/ProteoSAFe/status.jsp?task=d836d5a58f3f40e3bf9114f756ae8a59>

BRB_Powder_pos_NCE_40

<https://gnps.ucsd.edu/ProteoSAFe/status.jsp?task=4b509f3f114c40c9be2d781045d1ecaa>

BRB_Powder_pos_NCE_70

<https://gnps.ucsd.edu/ProteoSAFe/status.jsp?task=365e9da936eb46439ff984217eeedd57>

BRB_Powder_pos_intensity_10e4

<https://gnps.ucsd.edu/ProteoSAFe/status.jsp?task=45335fbfcbfb43e3a70e8c262ae36eec>

BRB_Powder_pos_intensity_10e5

<https://gnps.ucsd.edu/ProteoSAFe/status.jsp?task=31ab64efd7fe474cb74def93d0d53b73>

BRB_Powder_pos_intensity_10e6

<https://gnps.ucsd.edu/ProteoSAFe/status.jsp?task=bac46c42c2434a7da328fc66a91cfd0b>

BRB_Powder_pos_exclusion_10

<https://gnps.ucsd.edu/ProteoSAFe/status.jsp?task=54965c0b9e114225bf7b05b0a1be822d>

BRB_Powder_pos_exclusion_30

<https://gnps.ucsd.edu/ProteoSAFe/status.jsp?task=0a5818f92e2e41d3a91ec3b45fcbe2af>

BRB_Powder_pos_exclusion_60

<https://gnps.ucsd.edu/ProteoSAFe/status.jsp?task=a2b51e9bbd2047b8a252417493a02d91>

BRB_Powder_neg_resolution_17, 500

<https://gnps.ucsd.edu/ProteoSAFe/status.jsp?task=3098210f7e5c4f9096b55722c0a193a1>

BRB_Powder_neg_resolution_35, 000

<https://gnps.ucsd.edu/ProteoSAFe/status.jsp?task=540960e0d473451e8bd2f425df984307>

BRB_Powder_neg_resolution_70, 000

<https://gnps.ucsd.edu/ProteoSAFe/status.jsp?task=804d5d8a94194006b12e83a0cb6c440d>

BRB_Powder_neg_resolution_140, 000

<https://gnps.ucsd.edu/ProteoSAFe/status.jsp?task=f32d53a0e4cb4f5b97b06f360fcbdf34>

BRB_Powder_neg_NCE_10

<https://gnps.ucsd.edu/ProteoSAFe/status.jsp?task=8dd36491677f41b8b5443356619eace0>

BRB_Powder_neg_NCE_40

<https://gnps.ucsd.edu/ProteoSAFe/status.jsp?task=085f68f639ea4246935c871c53de499b>

BRB_Powder_neg_NCE_70

<https://gnps.ucsd.edu/ProteoSAFe/status.jsp?task=a965dddb6b394dfdb2f5c48a11e86171>

BRB_Powder_neg_intensity_10e4

<https://gnps.ucsd.edu/ProteoSAFe/status.jsp?task=48659846ce3848c383d08056116fe425>

BRB_Powder_neg_intensity_10e5

<https://gnps.ucsd.edu/ProteoSAFe/status.jsp?task=c30f0ed5951a411786eb4c4a02b10f34>
BRB_Powder_neg_intensity_10e6
<https://gnps.ucsd.edu/ProteoSAFe/status.jsp?task=0626421a984e41db99897c1bf988b519>
BRB_Powder_neg_exclusion_10
<https://gnps.ucsd.edu/ProteoSAFe/status.jsp?task=6a5da4394cc74e54b85fbaa6679e5b15>
BRB_Powder_neg_exclusion_30
<https://gnps.ucsd.edu/ProteoSAFe/status.jsp?task=032a43feae30421897be55401e7bab6f>
BRB_Powder_neg_exclusion_60
<https://gnps.ucsd.edu/ProteoSAFe/status.jsp?task=d21ed4d1252742d2b99ed74597d38ed7>
Liquid 70, 000-10-10e5-60 / Powder 35, 000-70-10e5-30 pos
<https://gnps.ucsd.edu/ProteoSAFe/status.jsp?task=36037e30c6ab420fa6d1c55464515626>
Liquid 17, 500-10-10e4-60 / Powder 17, 500-40-10e5-60 neg
<https://gnps.ucsd.edu/ProteoSAFe/status.jsp?task=b6736cce222243599949639c0811f4c9>