XCMS and Metaboanalyst

Stephen Barnes

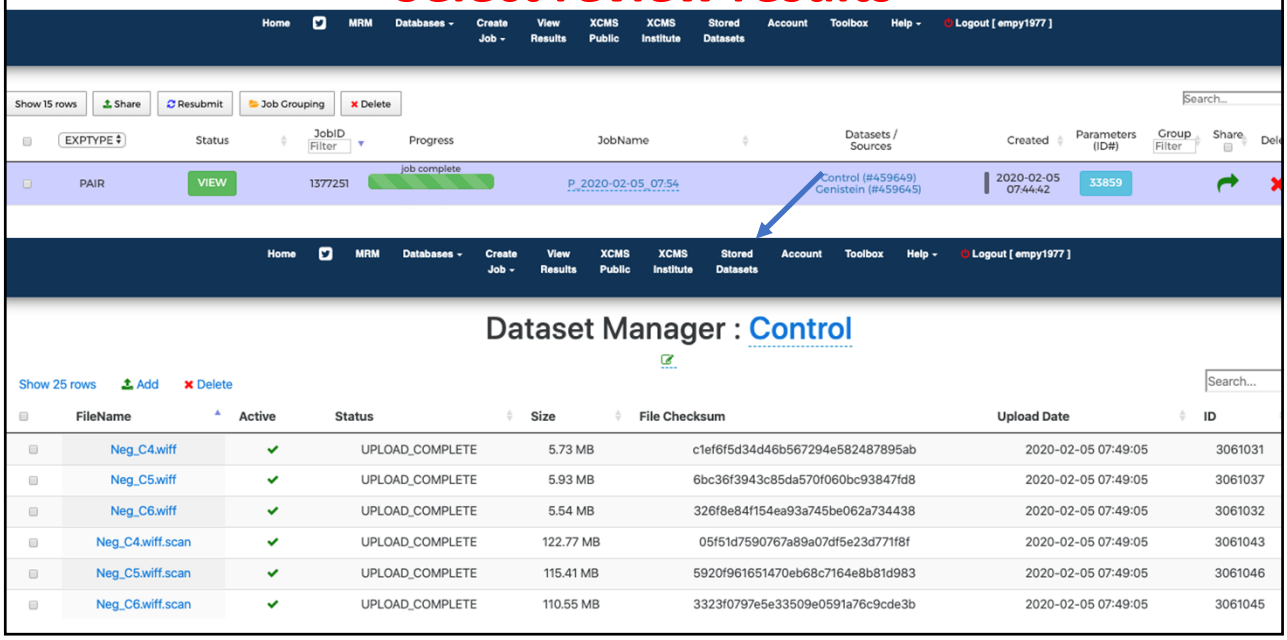
02-07-20

1

Go to https://xcmsonline.scripps.edu and login

2

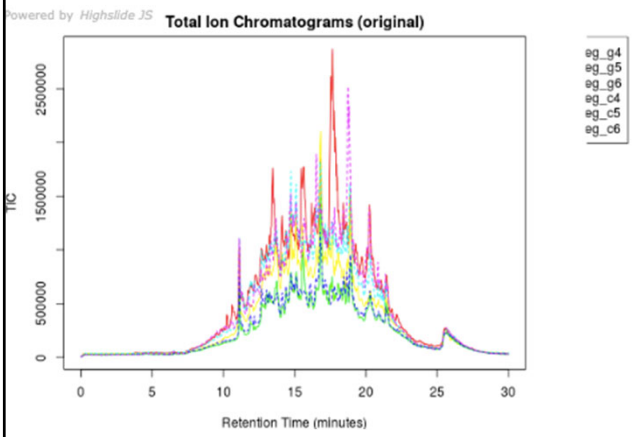
Select review results



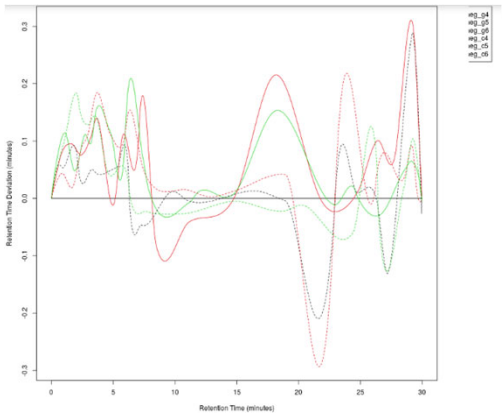
3

Click on view button ->

Overlay of raw data and retention time variation



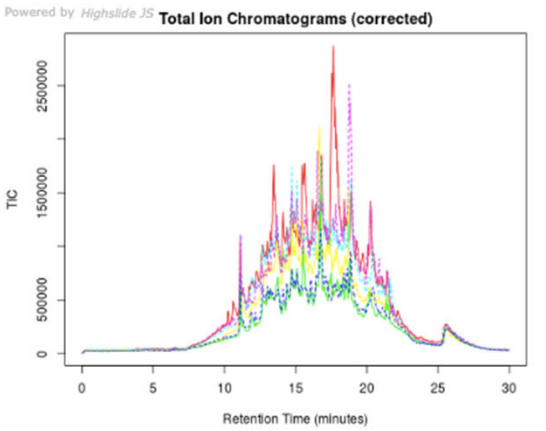
Note that there is a 2-3 fold variation in peak intensities

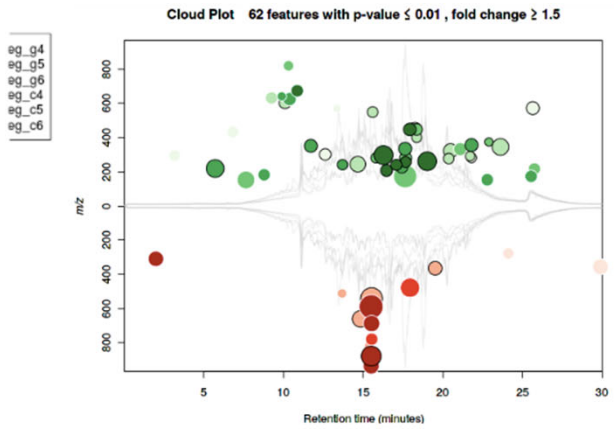


In the region between 10-22 min, there are 12-18 sec variations in retention time compared to the first sample (G4). This sample had very high amounts of Gen-SO_4.

4

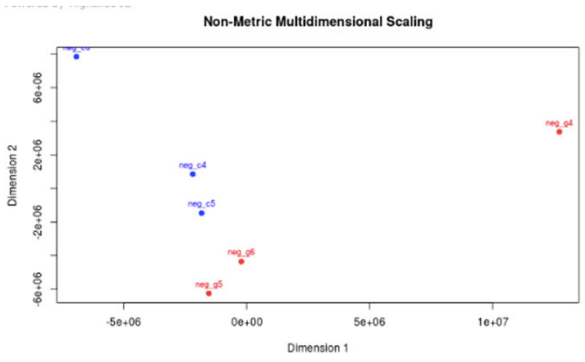
Corrected retention time and Cloud plot

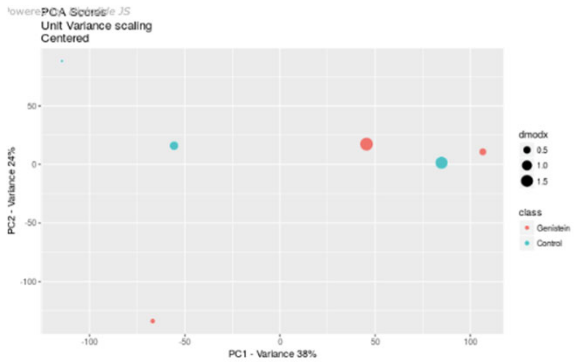




5

Initial statistics provided by XCMS





6

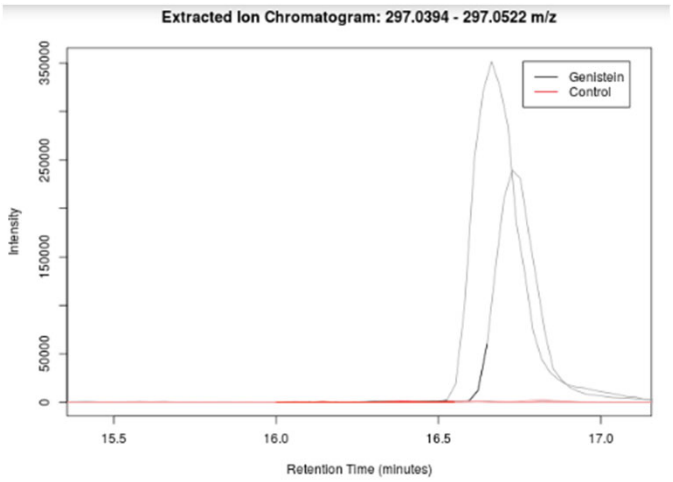
Click on the Results table in sidebar

Columns | Hide isotopic peaks | Page 1 of 220 | 100

featureidx	fold	pvalue	updown	mzmed	rtmed	maxint	dataset1_mean	dataset2_mean	isotopes	adducts	peakgroup
0	0.0	0.00000e+0		0.00	0.00	0.00	0.00	0.00			
0	0.0	0.00000e+0		0.00	0.00	0.00	0.00	0.00			
0	0.0	0.00000e+0		0.00	0.00	0.00	0.00	0.00			
1	8.3	0.00017	UP	297.0479	16.27	836	1,603	13,277		[M-H]- 298.056	59
2	7.8	0.00025	UP	261.0811	19.02	880	987	7,697		[M+Cl]- 226.113	117
2	0.0	0.00000e+0		0.00	0.00	0.00	0.00	0.00			
3	5.0	0.00025	DOWN	688.2208	15.52	1,915	20,880	4,191	[1621][M]-		14
4	2.1	0.00039	UP	240.1254	17.08	746	6,272	13,020		[M-H+Cl]2- 446.	21
5	4.1	0.00041	DOWN	309.8931	1.97	37	97	23			368
6	2.6	0.00062	UP	673.7176	10.85	124	300	773			212
7	2.0	0.00063	UP	257.1041	17.66	5,845	20,970	42,701	[222][M]-	[M-H+HCOOH]-	16
8	1.3	0.00074	DOWN	471.2782	25.36	64	560	437			153
9	10.3	0.00079	DOWN	879.2911	15.50	3,692	22,774	2,221	[2001][M]-	[M-H-CH3]- 895	14
10	2.3	0.00079	UP	207.0704	16.47	456	1,447	3,287		[M-H-COCH2]- 2	6
11	1.3	0.00088	DOWN	392.1401	22.84	158	611	464	[721][M+1]-		177
12	1.4	0.00095	UP	315.9503	1.26	20	31	43			602

7

Right sidebar of EIC of *m/z* 297.0479

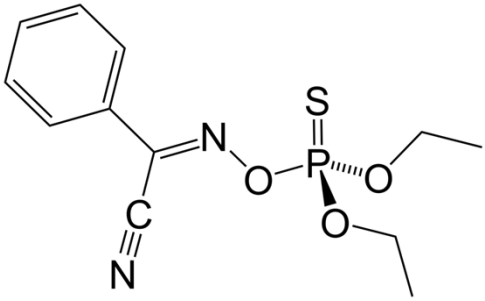


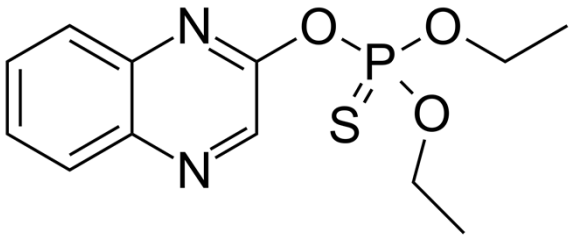
ppr	NAME	ADDUCT	METLIN	TOXIN
1		M-H2O-H	998499	n
4		M-H	1046395	n
4		M-H	1051620	n
4		M-H	1082880	n
4		M-H2O-H	1071195	n
4	Phoxim	M-H	72530	n
4	Quinalphos	M-H	68785	n

The control samples (red line) did not contain this ion – so, is it a metabolite of genistein? And is it in the METLIN database? Phoxim and Quinalphos each have the mass of 298.0541 and *m/z* of 297.0469.

8

Phoxim and Quinalphos structures



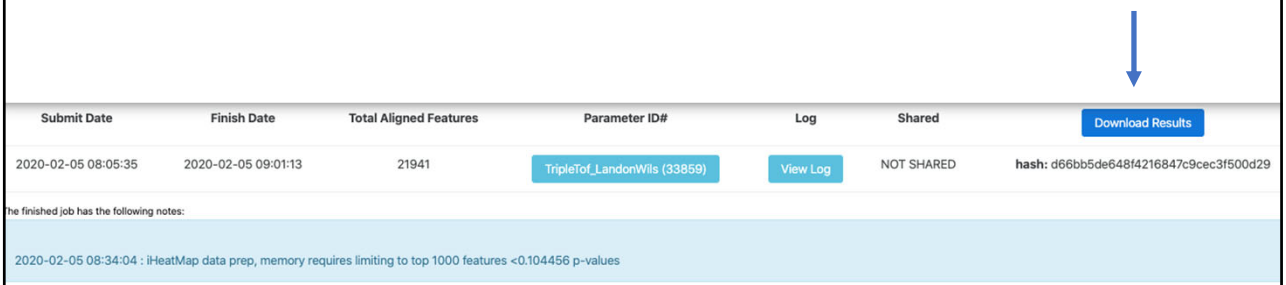


Empirical formula $C_{12}H_{15}N_2O_3PS$

Dead-reckoning the mass defect
14 H @+0.007825 = 0.10955
2 N @+0.003074 = 0.006148
3 O @ -0.005086 = -0.015258
P = -0.026239, S = -0.027929
Total = 0.046272

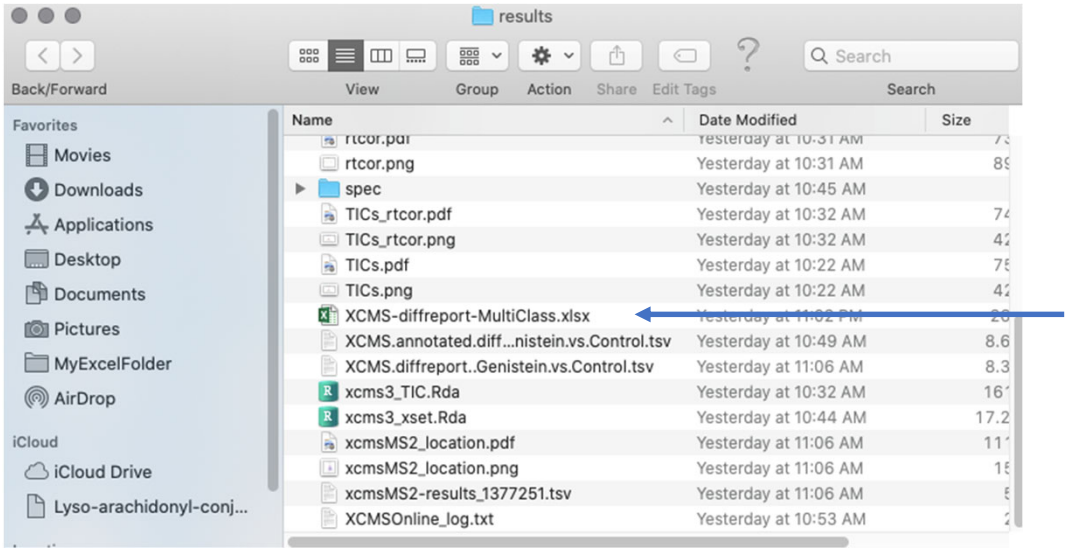
9

Downloading the processed data



Once downloaded, unzip the file (Results)

10



11

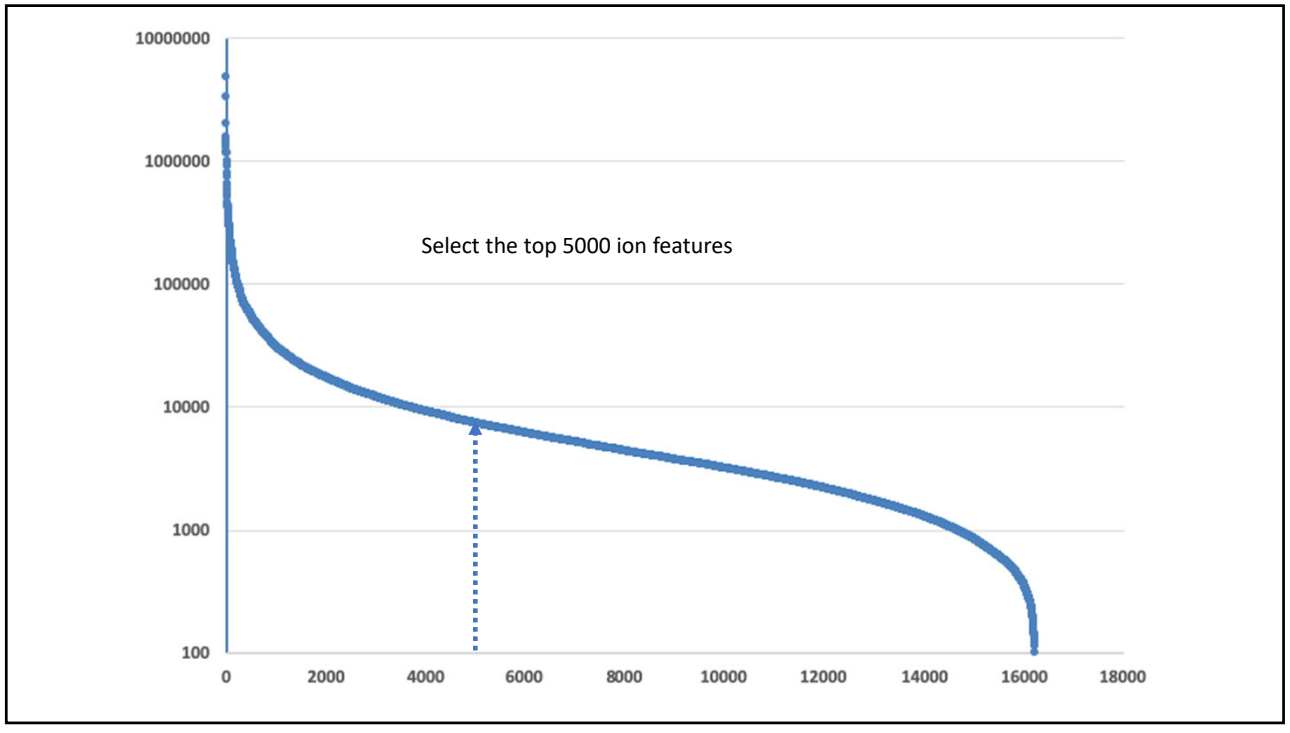
	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U
1	name	fold	log2fold	tstat	pvalue	qvalue	updown	mzmed	mzmin	mzmax	rtmed	rtmin	rtmax	npeaks	Genistein	Control	maxint	mean1	sd1	mean2	sd2
2	M297T16_	8.283522	3.050244	19.0533	0.000174	1	UP	297.0479	297.0413	297.0496	16.27148	16.15827	16.35415	3	3	0	836.1301	1602.76	550.3791	13276.5	907.3281
3	M261T19_	7.795805	2.962698	13.83332	0.000246	1	UP	261.0811	261.0794	261.082	19.01917	19.018	19.09267	3	3	0	879.8413	987.2949	506.3779	7696.759	670.3139
4	M688T16_	4.982727	-2.31694	-13.7956	0.000248	1	DOWN	688.2208	688.2177	688.2246	15.52298	15.3927	15.74043	4	3	0	1915.04	20880.17	1263.487	4190.512	1671.611
5	M240T17_	2.075918	1.053749	15.64491	0.000394	1	UP	240.1254	240.1248	240.1261	17.07933	16.9915	17.08967	3	1	2	746.3985	6271.754	373.5183	13019.65	646.9801
6	M310T2	4.118643	-2.04217	-11.5939	0.000407	1	DOWN	309.8931	309.8925	309.9034	1.967567	1.830133	2.067917	3	0	3	36.78276	96.73505	8.53576	23.48712	6.84717
7	M674T11	2.57281	1.363345	11.21279	0.000616	1	UP	673.7176	673.7168	673.7193	10.85063	10.84668	10.8927	4	1	3	123.8508	300.3898	59.44125	772.8457	42.34292
8	M257T18_	2.03629	1.025943	10.42087	0.000626	1	UP	257.1041	257.103	257.1053	17.66467	17.48383	17.70133	6	3	3	5845.161	20969.87	2237.369	42700.73	2835.467
9	M471T25_	1.280748	-0.35699	-9.44545	0.00074	1	DOWN	471.2782	471.2752	471.2796	25.355	25.28833	25.42767	3	2	1	64.05115	560.3242	15.04764	437.4976	16.75898
10	M879T15	10.25602	-3.3584	-12.642	0.000791	1	DOWN	879.2911	879.2895	879.2927	15.49958	15.49462	15.5012	3	3	0	3691.939	22774.11	2443.614	2220.56	1399.478
11	M207T16_	2.271242	1.183481	12.12411	0.000792	1	UP	207.0704	207.0686	207.071	16.46547	16.36125	16.62183	5	3	1	455.8566	1447.048	225.5648	3286.596	134.8451
12	M392T23	1.31488	-0.39493	-9.02669	0.000885	1	DOWN	392.1401	392.1377	392.143	22.84075	22.81667	22.865	6	3	3	157.9479	610.6244	18.68669	464.3956	20.9306
13	M316T1	1.368723	0.45283	10.35851	0.000954	1	UP	315.9503	315.9503	315.9504	1.263917	1.176217	1.351617	2	2	0	19.5705	31.48005	1.611732	43.08745	1.081356
14	M937T15	4.49709	-2.16899	-8.55057	0.001624	1	DOWN	937.2353	937.2351	937.2355	15.49663	15.47277	15.52048	2	2	0	635.9555	2721.853	350.2295	605.2477	247.3198
15	M447T18_	2.551593	1.351398	7.957769	0.001764	1	UP	447.2589	447.2558	447.2593	17.9245	17.73317	18.04217	7	3	3	4707.9	14908.36	3064.607	38040.08	3994.597
16	M590T15_	29.42193	-4.87882	-22.0333	0.001895	1	DOWN	589.9996	589.9981	590.0002	15.49958	15.42068	15.52048	3	3	0	1587.976	14700.26	1111.749	499.6361	100.8869
17	M227T17	2.328883	1.219638	7.218261	0.002042	1	UP	227.0937	227.0932	227.0947	17.42117	17.29617	17.4475	5	2	3	2098.236	6497.621	1541.983	15132.2	1383.86
18	M219T6	6.424274	2.683533	8.710933	0.00205	1	UP	218.9449	218.9418	218.9477	5.695417	5.605333	5.7692	3	3	0	82.29803	42.22965	24.09778	271.2949	38.64947
19	M479T18	9.003492	-3.17048	-12.1973	0.002075	1	DOWN	478.9563	478.948	478.9621	17.93925	17.73933	18.07167	4	2	2	752.7654	4720.456	550.9348	524.2916	226.9951
20	M780T16	2.50805	-1.32657	-10.3643	0.002309	1	DOWN	780.2218	780.2185	780.2251	15.52661	15.47277	15.58045	2	2	0	512.6193	2150.976	93.426	857.6286	194.9065
21	M641T10	1.671018	0.740727	9.738851	0.002339	1	UP	641.2035	641.2017	641.2045	9.875567	9.867233	9.8926	3	1	2	100.0927	426.1714	23.25036	712.1401	45.23385
22	M153T23	2.521678	1.334384	6.701285	0.002594	1	UP	153.0573	153.0568	153.0578	22.7915	22.77867	22.80433	2	0	2	30.38939	50.49542	14.31532	127.3332	13.76544
23	M241T14_	1.904398	0.929335	8.588919	0.002812	1	UP	241.0734	241.0729	241.0744	13.69837	13.65912	13.71953	6	3	3	866.4532	5397.736	478.8588	10279.44	860.1364
24	M625T10_	2.888792	1.530467	6.663759	0.002904	1	UP	625.189	625.1857	625.1905	10.36641	10.33667	10.42667	6	3	3	1920.918	4147.634	1313.991	11981.65	1555.52
25	M182T9	2.466902	1.302701	9.025981	0.00314	1	UP	182.0479	182.0463	182.0491	8.7701	8.710167	8.8225	3	2	0	69.52372	154.8222	39.01952	381.9313	19.41162
26	M676T16_	2.159368	-1.11061	-6.35758	0.003199	1	DOWN	676.2209	676.2182	676.2237	15.57671	15.54638	15.60703	2	2	0	871.6788	3372.814	335.32	1561.945	361.878
27	M356T22_	2.711366	1.43902	6.797359	0.00324	1	UP	356.1015	356.0981	356.103	21.8015	21.56083	21.86	4	3	1	352.9981	566.2024	147.484	1535.182	198.0201

Copy the whole data and transfer to a new datasheet. Then sort the data by retention time and remove the lines of data with retention times less than 10 min and greater than 22 min.

12

	I	J	K	L	M	N	O	P	Q	R	S	T	U
	mzmin	mzmax	rtmed	rtmin	rtmax	npeaks	Genistein	Control	maxint	mean1	sd1	mean2	sd2
	577.114823	577.118315	10.0002	9.98428333	10.08875	6	3	3	5819.27273	15652.6766	6052.10577	27142.4086	15533.0731
	578.116943	578.12133	10.0002	9.98428333	10.08875	6	3	3	1562.57776	4239.41894	1882.55322	7063.7064	3942.29954
	432.971115	432.974526	10.0002	9.98428333	10.0605	6	3	3	244.840976	878.984048	404.227798	1312.79697	768.520567
	380.060739	380.062717	10.0002	9.98428333	10.0605	6	3	3	4558.22808	16700.9284	7333.06704	24488.7939	14991.1405
	677.320966	677.324475	10.0002	9.98555	10.08875	4	2	2	661.686261	1104.19126	1221.11153	1664.01209	2276.98364
	645.108428	645.111358	10.0002	9.98866667	10.0117333	2	0	2	126.929998	594.563619	265.106878	676.027128	310.213139
	599.097825	599.102954	10.0002	9.98428333	10.0605	6	3	3	198.338784	817.476845	415.713068	869.905794	268.083836
	606.283721	606.289623	10.0002	9.98428333	10.2050167	6	2	3	583.274221	1704.61584	1303.76713	1916.69405	2384.38901
	822.285153	822.291246	10.00185	9.98866667	10.0150333	2	0	2	104.442442	625.61984	461.265372	678.840542	511.19541
	583.233734	583.234215	10.0061667	9.98866667	10.0236667	2	0	2	78.4087572	614.154199	279.263082	494.244356	241.89756
	776.355734	776.359378	10.0073	9.98428333	10.0605	5	2	3	1672.49754	2803.68585	4067.74249	6454.11988	7671.46077
	623.042997	623.044694	10.008575	9.93201667	10.0851333	2	0	2	56.7314982	572.702655	194.966695	356.038638	191.38933
	382.066063	382.069539	10.00985	9.96043333	10.0605	4	2	2	1448.7371	3365.90074	4912.68743	3915.11761	5742.35892
	470.026387	470.031363	10.0104833	9.98428333	10.0605	4	3	1	83.8597532	351.57579	125.023528	502.248826	152.865551
	I	J	K	L	M	N	O	P	Q	R	S	T	U
16214	290.141925	290.144168	21.9761667	21.9746667	21.9776667	2	2	0	380.413931	2743.01527	1396.88845	2873.95165	369.706721
16215	298.201533	298.202728	21.9765	21.9036667	22.0036667	3	1	2	467.662485	4025.60955	1647.77582	2874.57987	2277.49887
16216	298.125517	298.131232	21.9765	21.931	22.1546667	3	0	3	173.181348	823.740349	264.104445	785.246934	108.895088
16217	383.076415	383.079267	21.9776667	21.9776667	21.9776667	2	2	0	343.194075	1332.15603	1591.52138	3000.17134	875.027185
16218	281.081409	281.084089	21.9776667	21.9393333	22.0033333	3	2	1	348.54677	1016.70928	1372.63357	1784.63992	1369.70715
16219	453.136991	453.137177	21.9821667	21.931	22.0333333	2	0	2	288.789065	843.588548	43.8790933	1620.5716	636.764313
16220	301.165912	301.166169	21.9821667	21.931	22.0333333	2	0	2	1418.99342	13536.4671	2130.41675	12866.4302	970.792114
16221	227.147398	227.149332	21.9868333	21.8018333	22.1175	3	1	2	549.547714	5680.56266	2825.71165	3620.77891	1676.36616
16222	556.988138	556.997507	21.9879167	21.8921667	22.1451667	4	2	2	89.788908	579.136534	56.0383595	518.289496	210.141536
16223	240.167367	240.169097	21.9905	21.9485	22.2536667	6	3	3	374.984181	1511.26626	402.67445	2019.47367	626.919151
16224	239.165176	239.166931	21.9905	21.9485	22.2536667	6	3	3	2348.14237	8036.74977	1897.52459	11935.0995	5195.79544
16225	733.419768	733.420209	21.9930833	21.9776667	22.0085	2	2	0	160.758017	885.061338	660.846233	1205.20387	203.357591
16226	585.196154	585.201533	21.9935833	21.8958333	22.0913333	2	2	0	188.855846	619.834623	201.037677	1289.36868	267.376402
16227	163.114031	163.114308	21.9963333	21.9593333	22.0333333	2	0	2	77.1672733	139.610312	1.57265278	360.433053	148.874186
16228	317.0936	317.094326	21.9963333	21.9593333	22.0333333	2	0	2	496.861884	1460.90117	532.556053	2136.57567	752.872931
16229	385.147594	385.14818	21.9963333	21.9593333	22.0333333	2	0	2	573.817642	5496.51844	1181.11939	4484.32925	1761.94527
16230	344.101211	344.107621	21.99675	21.9313333	22.0621667	2	2	0	83.013986	379.683531	167.606789	490.696097	40.5740687

13



14

Select mzmed and rtmed and make new file

Formula Bar

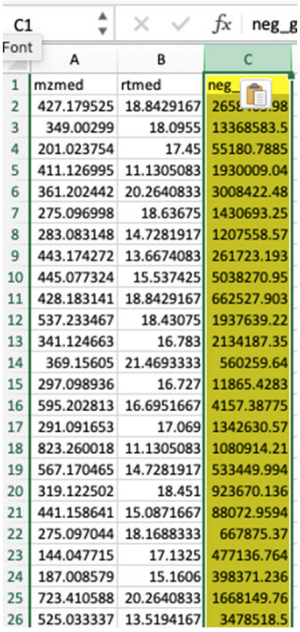
mzmed	mzmin	mzmax	rtmed	rtmin	rtmax	npeaks	Genistein	Control	maxint	mean1	sd1	mean2	sd2
427.179525	427.178088	427.181567	18.8429167	18.6953333	18.9068333	6	3	3	963460.11	3248009.63	1130471.98	6595400.41	5692890.26
349.00299	349.002682	349.003872	18.0955	17.8455	18.1018333	5	2	3	847738.735	6164608.52	6239665.79	578410.88	589944.776
201.023754	201.023303	201.025448	17.45	17.2176667	17.5013333	5	3	2	416393.822	4070005.41	3875251.78	10319.521	7068.81882
411.126995	411.124416	411.127481	11.1305083	11.1033667	11.1555167	6	3	3	278201.378	1305200.93	544816.011	1889834.15	430231.924
361.202442	361.200885	361.2042	20.2640833	20.1931667	20.4188333	6	3	3	309013.074	1394654.83	1413808.06	1632512.33	1118606.54
275.096998	275.095963	275.098057	18.63675	18.4908333	18.6806667	6	3	3	264985.171	1105088.16	325535.005	1735015.06	865324.662
283.083148	283.082521	283.084247	14.7281917	14.7267833	14.7542333	6	3	3	245492.781	1322023.04	317029.717	1427665.06	578933.157
443.174272	443.171242	443.17535	13.6674083	13.5581667	13.68745	6	3	3	194822.368	1158700.44	873617.017	1510345.82	947753.21
445.077324	445.075423	445.07836	15.537425	15.4995833	15.58045	6	3	3	226833.604	2624633.43	2097190.68	14138.1006	10175.335
428.183141	428.181441	428.184911	18.8429167	18.6953333	18.9068333	6	3	3	241333.293	795007.384	279555.516	1626053.63	1398494.55
537.233467	537.230533	537.234612	18.43075	18.2481667	18.4543333	6	3	3	351041.033	917555.761	890682.795	1422306.36	1221595
341.124663	341.123866	341.125697	16.783	16.6111667	16.8075	6	3	3	264216.604	1141138.01	875311.037	1177113.75	908002.23
369.15605	369.154368	369.157498	21.4693333	21.4235	21.7481667	6	3	3	109253.308	786951.812	237344.381	1223892.17	628813.198
297.098936	297.09764	297.099913	16.727	16.59685	16.8725	7	3	3	351464.949	1963294.47	1845002.57	16422.6965	7516.14404
595.202813	595.202128	595.203498	16.6951667	16.6633333	16.727	2	2	0	589148.917	1914099.33	2123022.57	4019.33515	1025.65466
291.091653	291.091107	291.092441	17.069	16.8955	17.1145	6	3	3	137013.856	686118.748	568601.69	1127334.03	45663.4263
823.260018	823.254914	823.261759	11.1305083	11.1033667	11.1555167	6	3	3	288897.422	541628.88	468141.217	1086715.18	525026.058
567.170465	567.168674	567.17118	14.7281917	14.7267833	14.7542333	6	3	3	281225.421	667485.304	339537.256	896691.569	666875.878
319.122502	319.121303	319.123766	18.451	18.2481667	18.7725	9	3	3	102506.583	493530.666	396507.217	1067031.67	142121.3
441.158641	441.156434	441.159753	15.0871667	15.0665	15.0996333	6	3	3	148885.753	725238.711	554770.601	816884.037	344122.877
275.097044	275.095739	275.097578	18.1688333	18.0143333	18.2195	6	3	3	151033.535	566141.233	118828.917	957125.325	458623.507
144.047715	144.047391	144.048428	17.1325	16.975	17.1703333	6	3	3	122580.635	636363.349	163405.092	885975.039	273610.577
187.008579	187.008291	187.009296	15.1606	15.1209667	15.1777167	6	3	3	130111.556	618663.6	317326.92	883883.891	473097.032
723.410588	723.406962	723.413081	20.2640833	20.1931667	20.4188333	6	3	3	258449.578	598118.308	927489.366	709978.5	886810.779
525.033337	525.031149	525.034867	13.5194167	13.45205	13.5338167	6	3	3	329846.341	1266847.86	1915432.54	31025.2247	34465.5586
350.006071	350.005653	350.006852	18.0955	17.8455	18.1018333	5	2	3	154259.415	1104459.26	1147041.55	101910.891	102783.969
349.093312	349.092037	349.094471	14.3770583	14.3566667	14.3953833	6	3	3	125217.836	683956.825	192997.673	489383.835	317184.108
291.091675	291.090876	291.092335	16.4187667	16.3022833	16.64865	9	3	3	93353.5956	592362.456	516016.909	546709.727	195690.082
891.162495	891.155776	891.162504	15.5204833	15.4995833	15.554	3	3	0	191063.401	1126799.22	1426317.65	886.263176	219.088928
429.181154	429.180536	429.184226	18.8561667	18.6953333	18.908	6	3	3	94511.9048	353024.185	126456.612	733119.799	596647.414
377.196754	377.19591	377.197522	16.7504167	16.6111667	16.8006667	6	3	3	75263.1223	441764.474	444926.638	573022.026	433644.809
309.102066	309.101249	309.10272	15.40555	15.3884333	15.4388333	6	3	3	64861.993	376852.532	282673.871	542373.863	174455.209

15

Scroll over to the sample area values and select one at a time

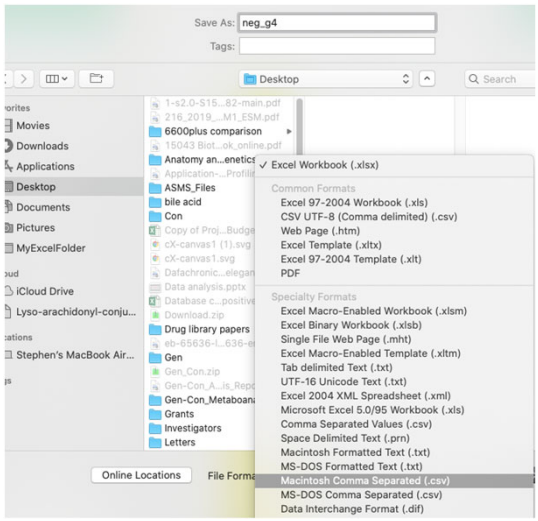
V	W	X	Y	Z	AA	AB	AC
neg_g4	neg_g5	neg_g6	neg_c4	neg_c5	neg_c6	rank #	average
2658463.98	4551392.31	2534172.59	4479791.26	2263200.54	13043209.4	1	4921705.02
13368583.5	2665031.68	2460210.41	1257271.29	287932.159	190029.192	2	3371509.7
55180.7885	7788693.2	4366142.23	11167.9792	16925.8175	2864.76649	3	2040162.46
1930009.04	929270.446	1056323.3	1946190.64	1434201.27	2289110.55	4	1597517.54
3008422.48	374068.62	801473.398	1314457.8	707375.591	2875703.59	5	1513583.58
1430693.25	779623.284	1104947.94	1048263.61	1449847.95	2706933.63	6	1420051.61
1207558.57	1680388.56	1078122	1989305.15	832864.274	1460825.77	7	1374844.05
261723.193	2006914.67	1207463.46	1593956.14	523557.521	2413523.8	8	1334523.13
5038270.95	1247596.61	1588032.74	25491.4107	11081.408	5841.48296	9	1319385.77
662527.903	1116174.91	606319.336	1121804.16	549612.842	3206743.89	10	1210530.51
1937639.22	293988.1	521039.958	912885.645	537866.924	2816166.51	11	1169931.06
2134187.35	481643.502	807583.173	1705600.66	1697085.97	128654.607	12	1159125.88
560259.64	1033678.37	766917.429	1330292.49	548666.881	1792717.15	13	1005421.99
11865.4283	3679264.36	2198753.63	24445.5754	9544.64192	15277.8724	14	989858.585
4157.38775	4199978.33	1538162.28	3687.1879	3200.91595	5169.9016	15	959059.333
1342630.57	350648.088	365077.586	1166597.85	1138180.18	1077224.05	16	906726.388
1080914.21	239819.471	304152.96	1152456.72	531914.429	1575774.38	17	814172.028
533449.994	1053582.02	415423.895	1564526.97	230779.369	894768.371	18	782088.437
923670.136	142619.109	414302.754	949707.435	1026322.1	1225065.47	19	780281.167
88072.9594	1101138.89	986504.287	1183879.95	501454.076	765318.082	20	771061.374
667875.37	435537.501	595010.827	680424.951	704434.339	1486516.68	21	761633.279
477136.764	803648.731	628304.55	1163147.1	616070.756	878707.258	22	761169.194
398371.236	982383.894	475235.671	1323669.99	383345.257	944636.429	23	751273.746
1668149.76	24230.5126	101974.649	332088.844	74706.7827	1723139.87	24	654048.404
3478518.5	177338.993	144686.092	70808.8919	12038.7087	10228.0734	25	648936.544

16



	A	B	C
1	mzmed	rtmed	neg_
2	427.179525	18.8429167	265[illegible]98
3	349.00299	18.0955	13368583.5
4	201.023754	17.45	55180.7885
5	411.126995	11.1305083	1930009.04
6	361.202442	20.2640833	3008422.48
7	275.096998	18.63675	1430693.25
8	283.083148	14.7281917	1207558.57
9	443.174272	13.6674083	261723.193
10	445.077324	15.537425	5038270.95
11	428.183141	18.8429167	662527.903
12	537.233467	18.43075	1937639.22
13	341.124663	16.783	2134187.35
14	369.15605	21.4693333	560259.64
15	297.098936	16.727	11865.4283
16	595.202813	16.6951667	4157.38775
17	291.091653	17.069	1342630.57
18	823.260018	11.1305083	1080914.21
19	567.170465	14.7281917	533449.994
20	319.122502	18.451	923670.136
21	441.158641	15.0871667	88072.9594
22	275.097044	18.1688333	667875.37
23	144.047715	17.1325	477136.764
24	187.008579	15.1606	398371.236
25	723.410588	20.2640833	1668149.76
26	525.033337	13.5194167	3478518.5

Make a .csv file for each sample



17

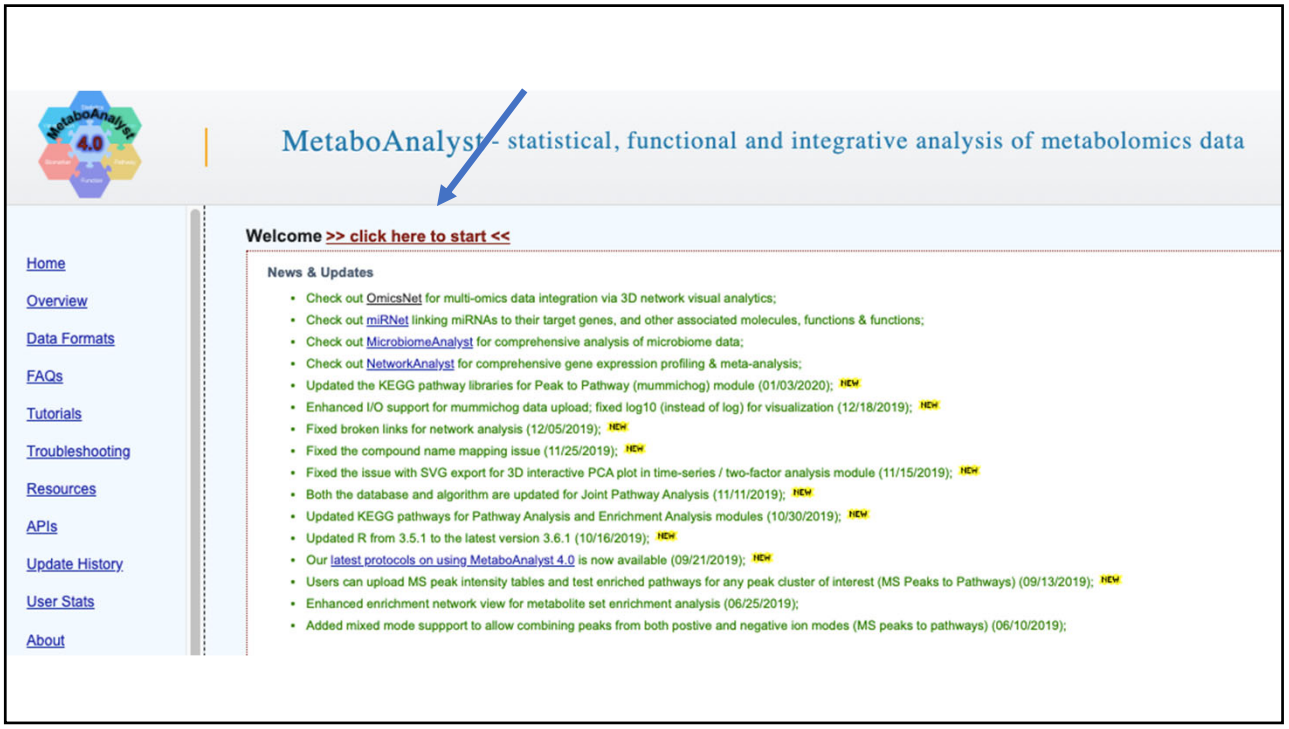
More on .csv files

- **The name of each file and the folders they'll be placed in must not have a space – add an "-" or "_" instead**
- **The mzmed and rtmed values will be the same for each sample, so just copy the next sample data into column C of the first .csv file and then save with the new sample lane**
- **Once all .csv files have been made, create two group folders (Gen and Con) and place the files in them**
- **Then zip up the two folders and rename the zip file as you choose (remember no spaces)**

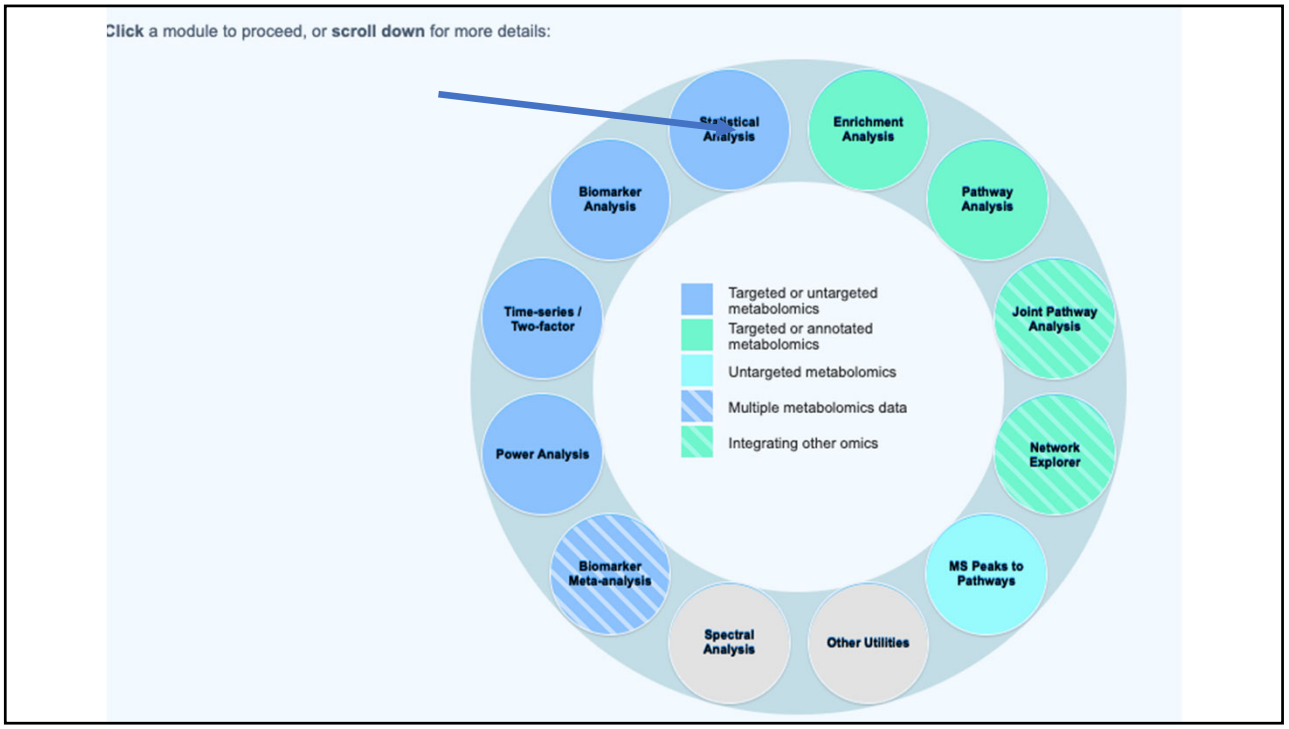
18

Go to http://www.Metaboanalyst.ca

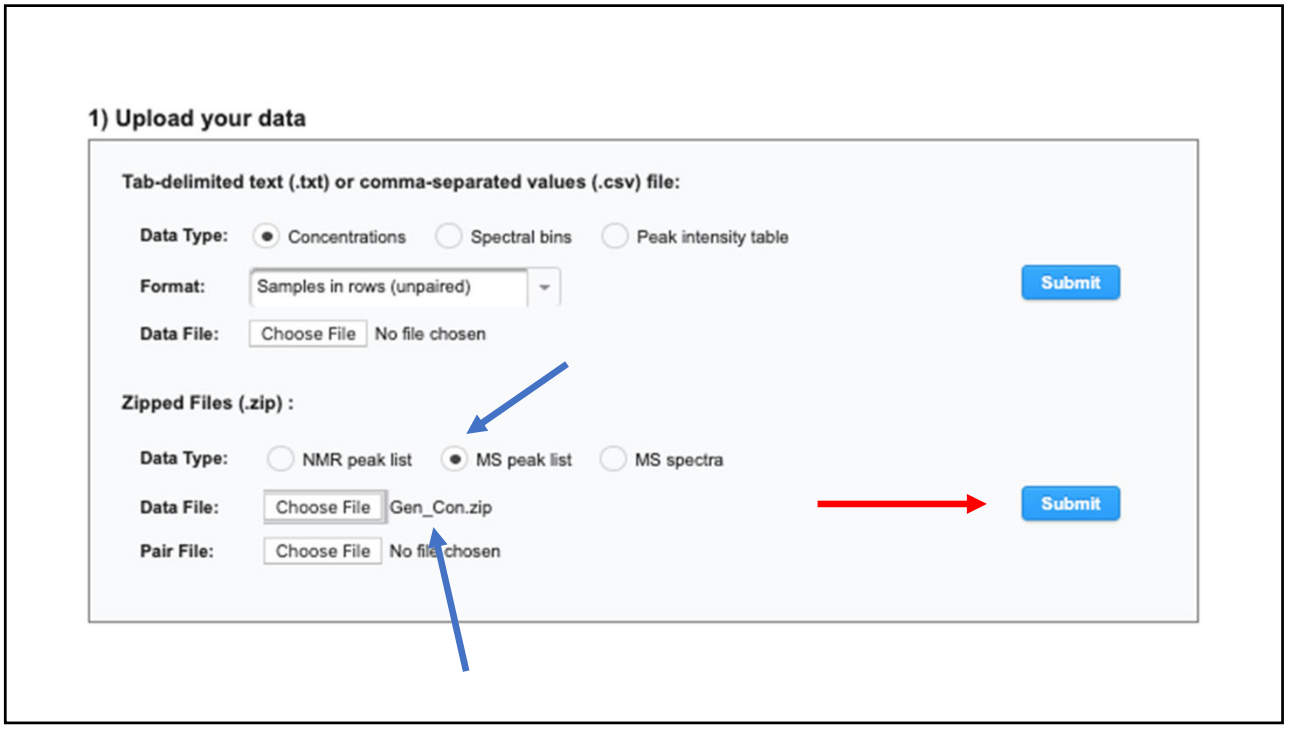
19



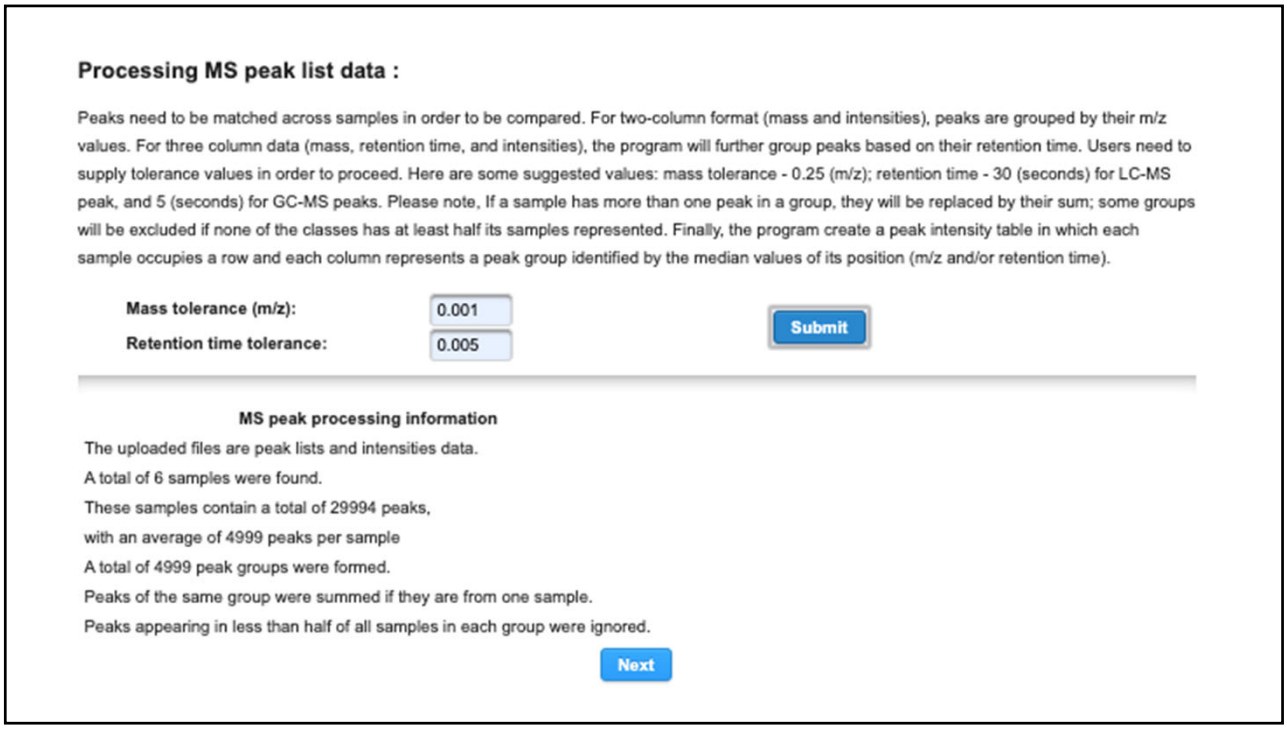
20


Click a module to proceed, or scroll down for more details:
Statistical Analysis
Enrichment Analysis
Pathway Analysis
Joint Pathway Analysis
Network Explorer
MS Peaks to Pathways
Other Utilities
Spectral Analysis
Biomarker Meta-analysis
Power Analysis
Time-series / Two-factor
Biomarker Analysis
Targeted or untargeted metabolomics
Targeted or annotated metabolomics
Untargeted metabolomics
Multiple metabolomics data
Integrating other omics

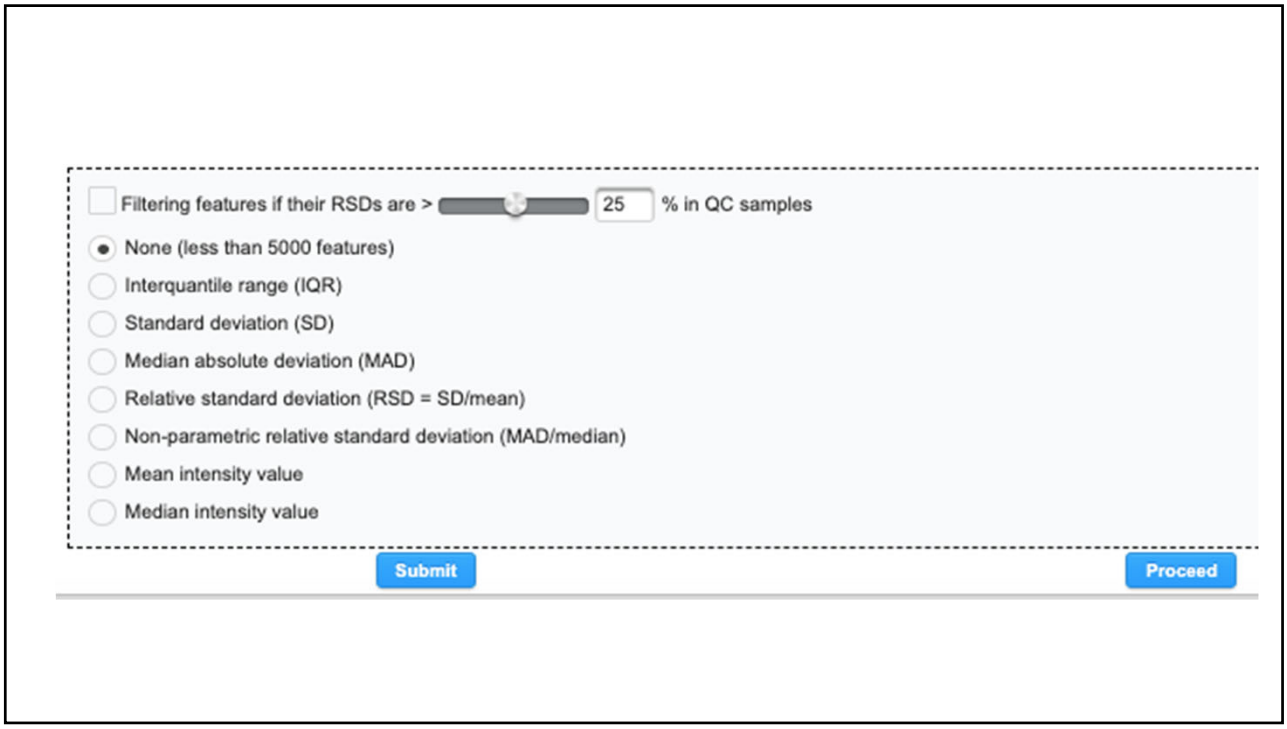
21


1) Upload your data
Tab-delimited text (.txt) or comma-separated values (.csv) file:
Data Type: Concentrations Spectral bins Peak intensity table
Format: Samples in rows (unpaired)
Submit
Data File: Choose File No file chosen
Zipped Files (.zip) :
Data Type: NMR peak list MS peak list MS spectra
Data File: Choose File Gen_Con.zip
Submit
Pair File: Choose File No file chosen

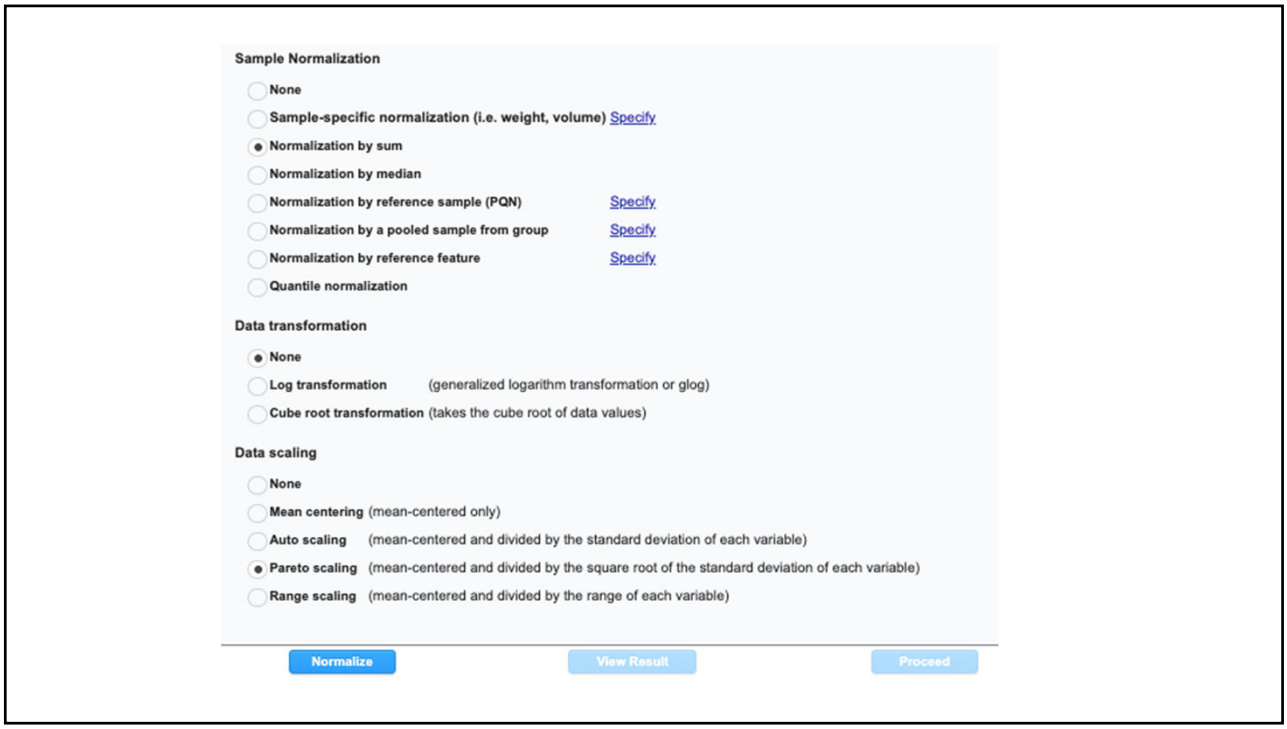
22


Processing MS peak list data :
Peaks need to be matched across samples in order to be compared. For two-column format (mass and intensities), peaks are grouped by their m/z values. For three column data (mass, retention time, and intensities), the program will further group peaks based on their retention time. Users need to supply tolerance values in order to proceed. Here are some suggested values: mass tolerance - 0.25 (m/z); retention time - 30 (seconds) for LC-MS peak, and 5 (seconds) for GC-MS peaks. Please note, If a sample has more than one peak in a group, they will be replaced by their sum; some groups will be excluded if none of the classes has at least half its samples represented. Finally, the program create a peak intensity table in which each sample occupies a row and each column represents a peak group identified by the median values of its position (m/z and/or retention time).
Mass tolerance (m/z):
0.001
Retention time tolerance:
0.005
Submit
MS peak processing information
The uploaded files are peak lists and intensities data.
A total of 6 samples were found.
These samples contain a total of 29994 peaks,
with an average of 4999 peaks per sample
A total of 4999 peak groups were formed.
Peaks of the same group were summed if they are from one sample.
Peaks appearing in less than half of all samples in each group were ignored.
Next

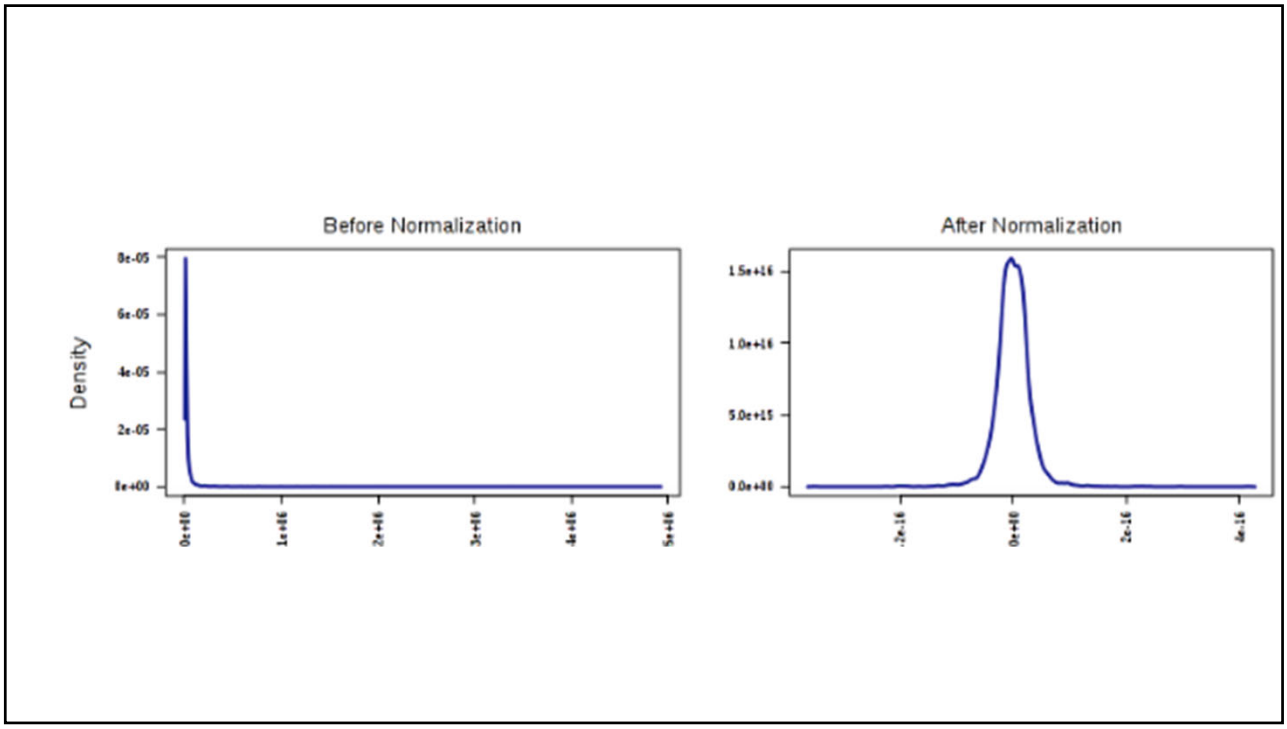
23


Filtering features if their RSDs are >
25
% in QC samples
None (less than 5000 features)
Interquantile range (IQR)
Standard deviation (SD)
Median absolute deviation (MAD)
Relative standard deviation (RSD = SD/mean)
Non-parametric relative standard deviation (MAD/median)
Mean intensity value
Median intensity value
Submit
Proceed

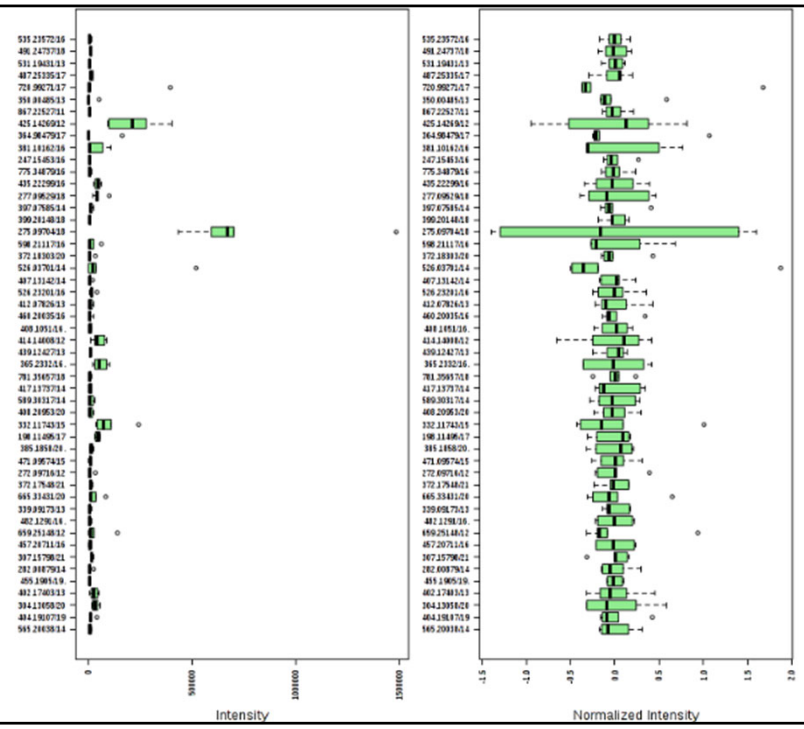
24


Sample Normalization
None
Sample-specific normalization (i.e. weight, volume) Specify
Normalization by sum
Normalization by median
Normalization by reference sample (PQN) Specify
Normalization by a pooled sample from group Specify
Normalization by reference feature Specify
Quantile normalization
Data transformation
None
Log transformation (generalized logarithm transformation or glog)
Cube root transformation (takes the cube root of data values)
Data scaling
None
Mean centering (mean-centered only)
Auto scaling (mean-centered and divided by the standard deviation of each variable)
Pareto scaling (mean-centered and divided by the square root of the standard deviation of each variable)
Range scaling (mean-centered and divided by the range of each variable)
Normalize
View Result
Proceed

25


Before Normalization
After Normalization
Density

26



Close the window and click on the proceed button

27

Select an analysis path to explore :

Univariate Analysis

Fold Change Analysis T-tests Volcano plot

One-way Analysis of Variance (ANOVA)

Correlation Analysis Pattern Searching

Chemometrics Analysis

Principal Component Analysis (PCA)

Partial Least Squares - Discriminant Analysis (PLS-DA)

Sparse Partial Least Squares - Discriminant Analysis (sPLS-DA)

Orthogonal Partial Least Squares - Discriminant Analysis (orthoPLS-DA)

Feature Identification

Significance Analysis of Microarray (and Metabolites) (SAM)

Empirical Bayesian Analysis of Microarray (and Metabolites) (EBAM)

Cluster Analysis

Hierarchical Clustering: Dendrogram Heatmaps

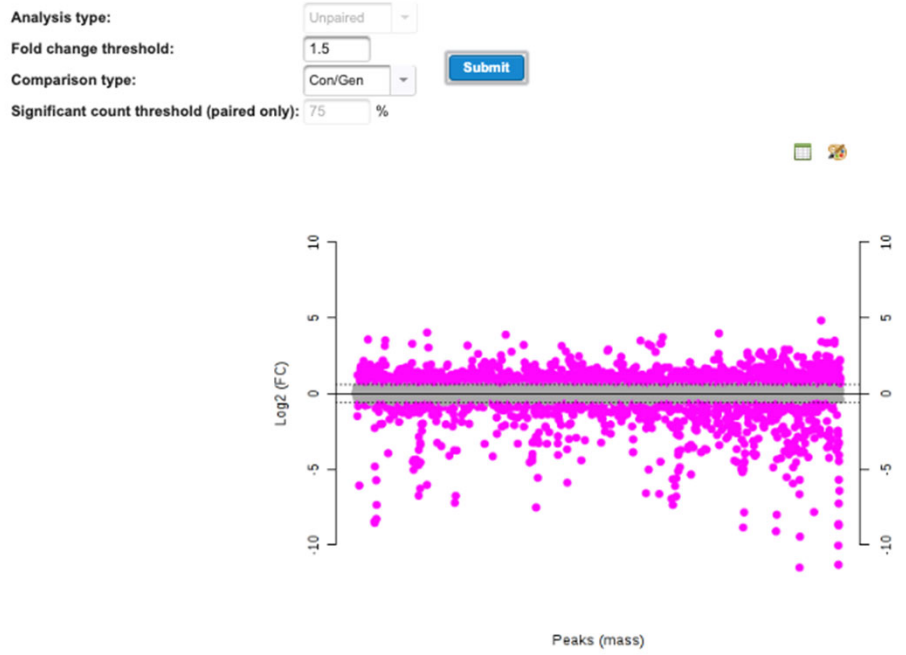
Partitional Clustering: K-means Self Organizing Map (SOM)

Classification & Feature Selection

Random Forest

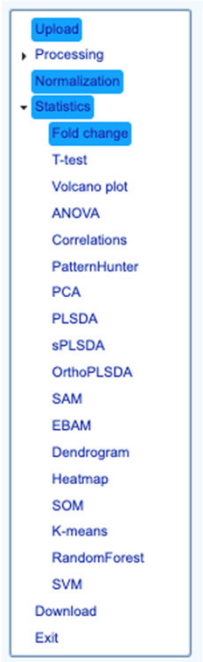
Support Vector Machine (SVM)

28



Note that fold change has been changed to 1.5 from the default value

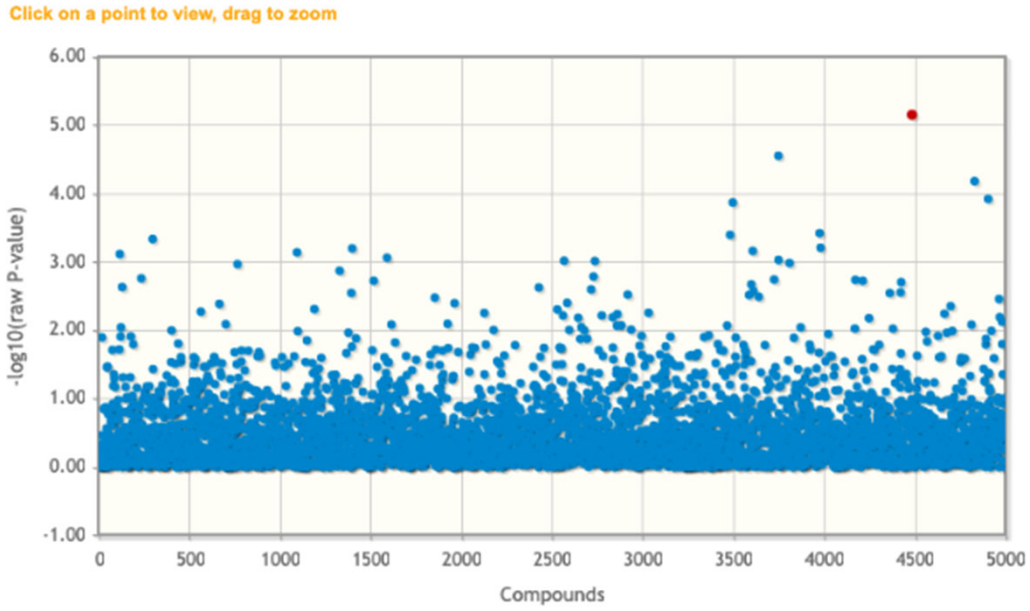
29



This road map to the statistical routines now becomes available

30

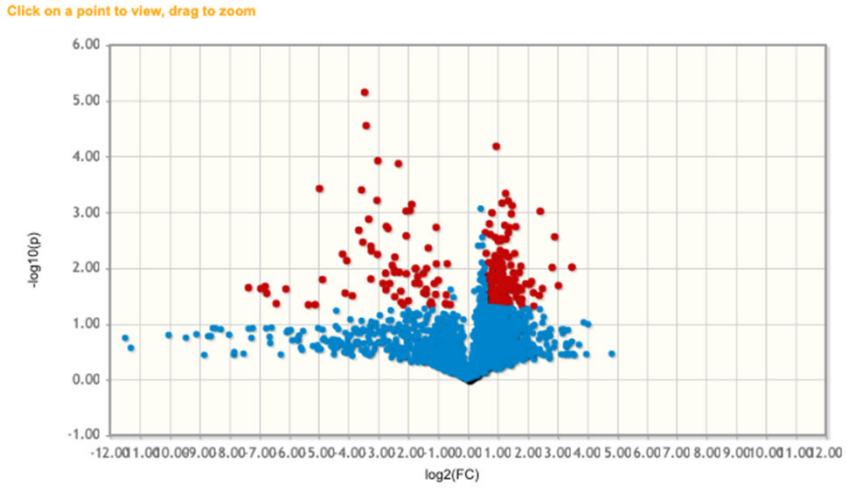
The correction for multiple tests is very strong



31

A volcano plot

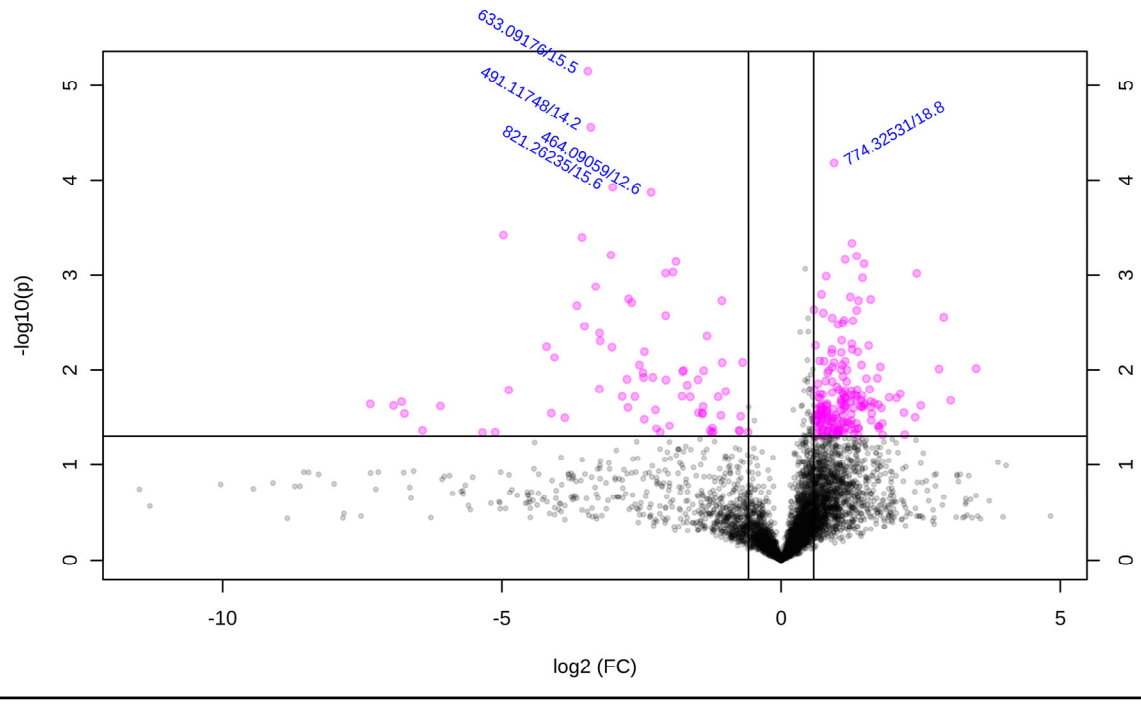
$-\log_{10}$p-value versus $\log_2$Fold change



32

Clicking on the palette enables publication quality pictures

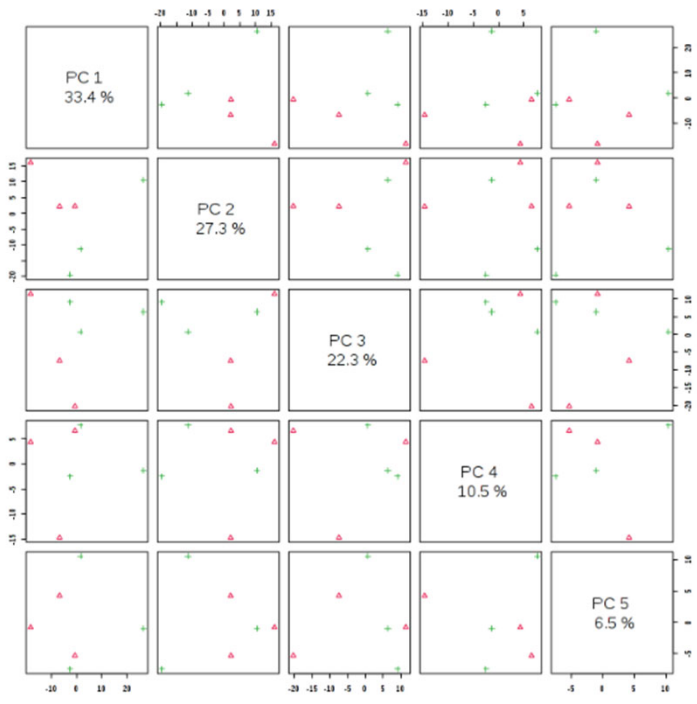
33



34

Multivariate analysis

35

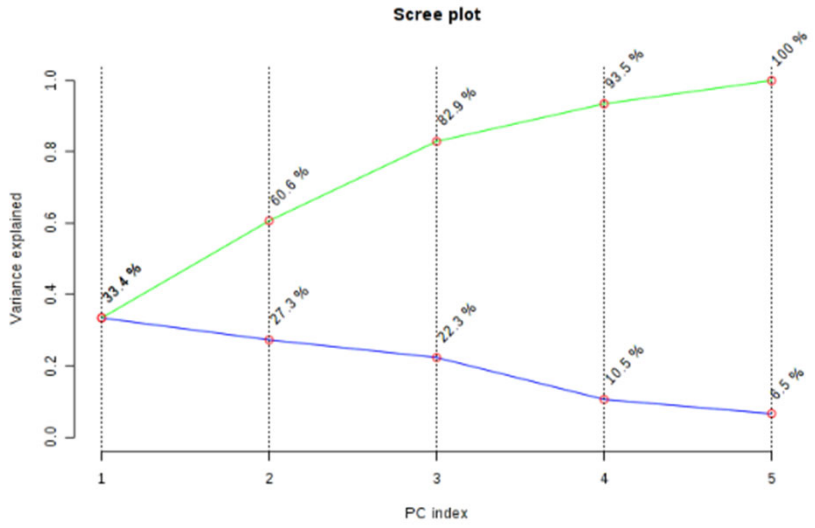


Selecting PCA in the side bar leads to the calculation of up to five (5) principal components. These are plotted against each other.

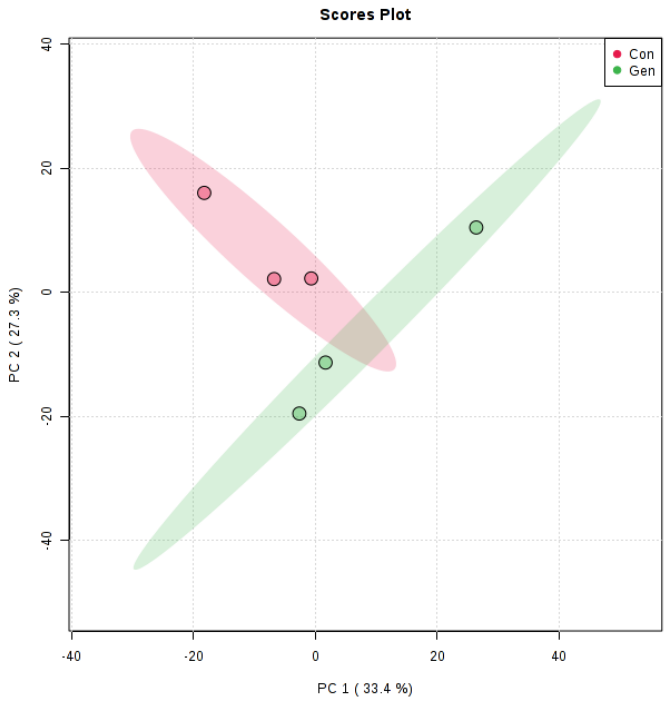
PCA is an unsupervised plot.

36

A scree plot – how much variation is explained by each PC (accumulated)



37

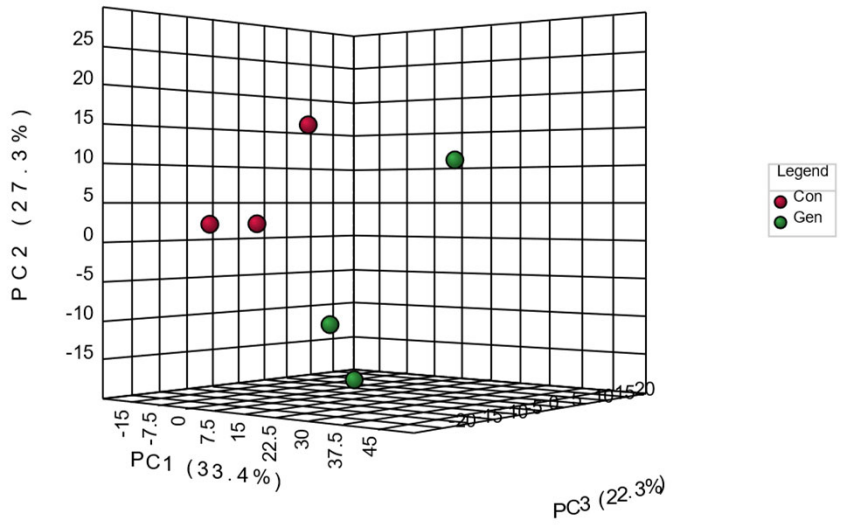


2D-PCA plot

Note that the samples are color-coded after the statistical calculations are made. The ellipses are the 95% confidence limits.

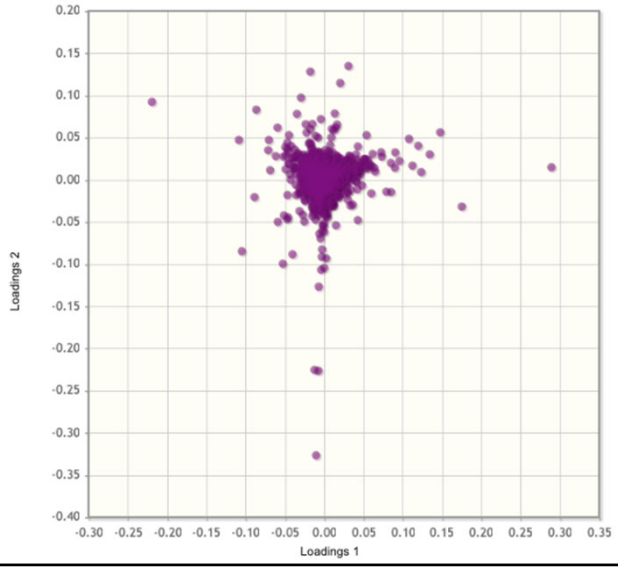
38

3D-PCA analysis

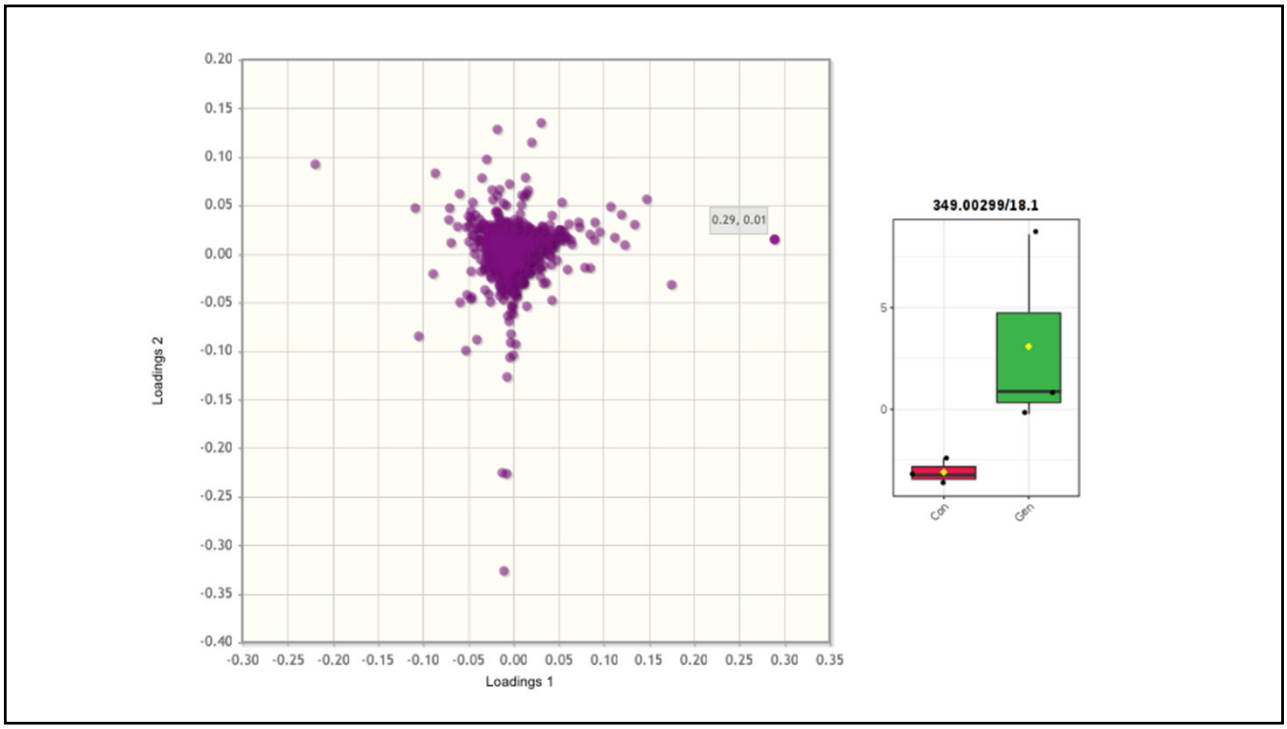


39

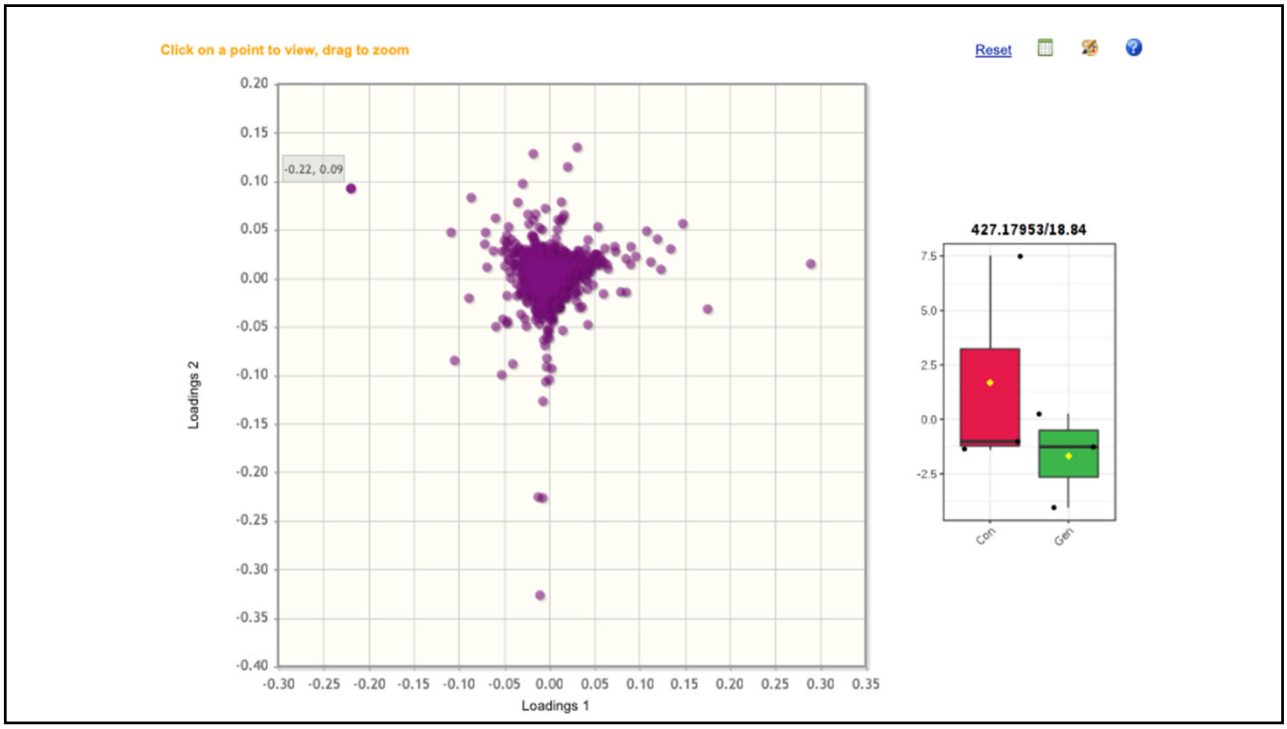
A loadings plot – 0,0 means no contribution



40


0.29, 0.01
Loadings 2
Loadings 1
349.00299/18.1
Con
Gen

41

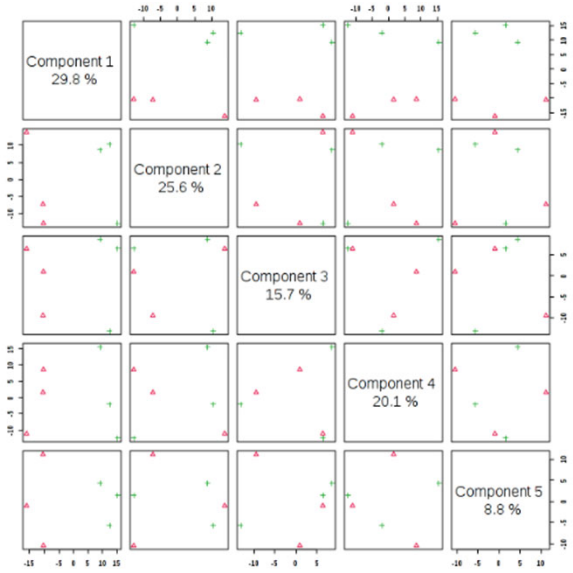

Click on a point to view, drag to zoom
Reset
-0.22, 0.09
Loadings 2
Loadings 1
427.17953/18.84
Con
Gen

42

Partial least squares-discriminant analysis

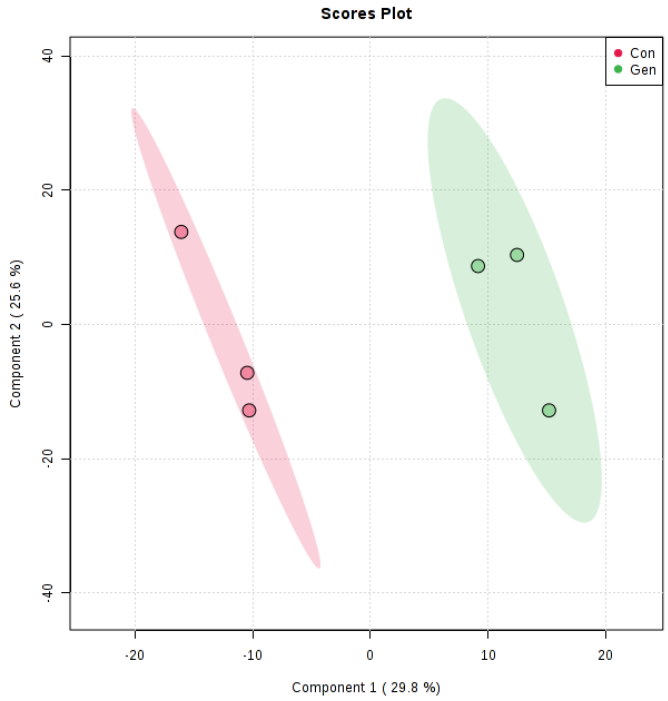
43

A scree plot – how much variation is explained by each PC (accumulated)



This is for PLSDA analysis

44

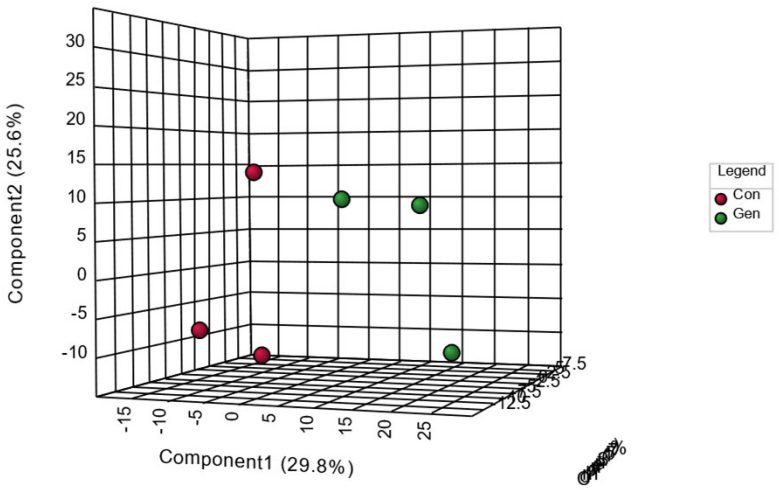


2D-PLS-DA plot

Note that the presence of two groups is included in the statistical analysis. The ellipses are the 95% confidence limits.

45

3D-PLS-DA plot

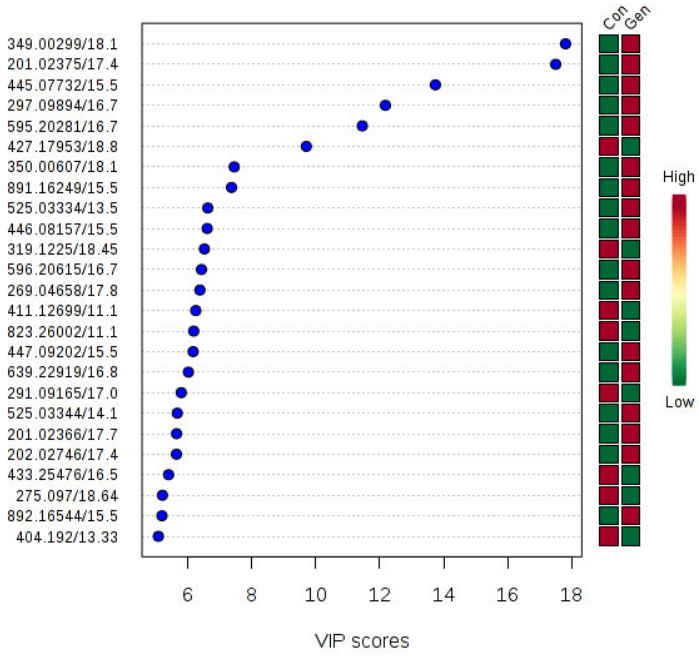


46

47

A loadings plot – 0,0 means no contribution

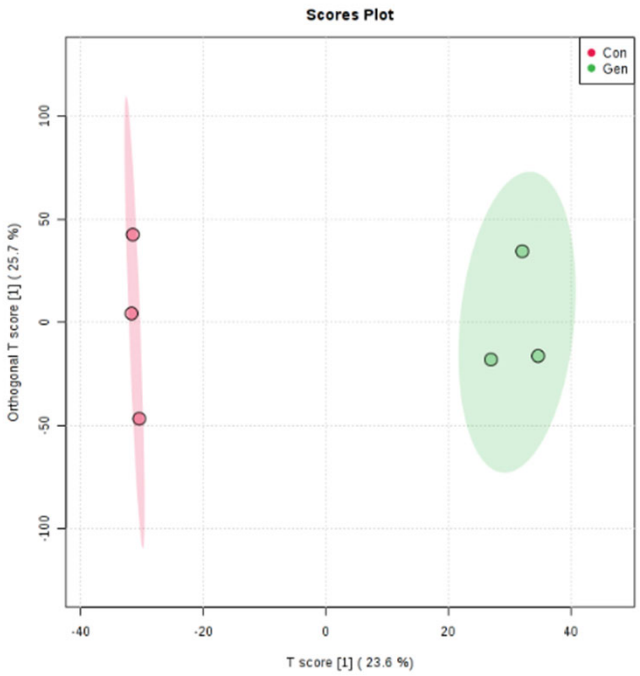
48



The ion features that most contribute to the group separation – VIP, variation in projection

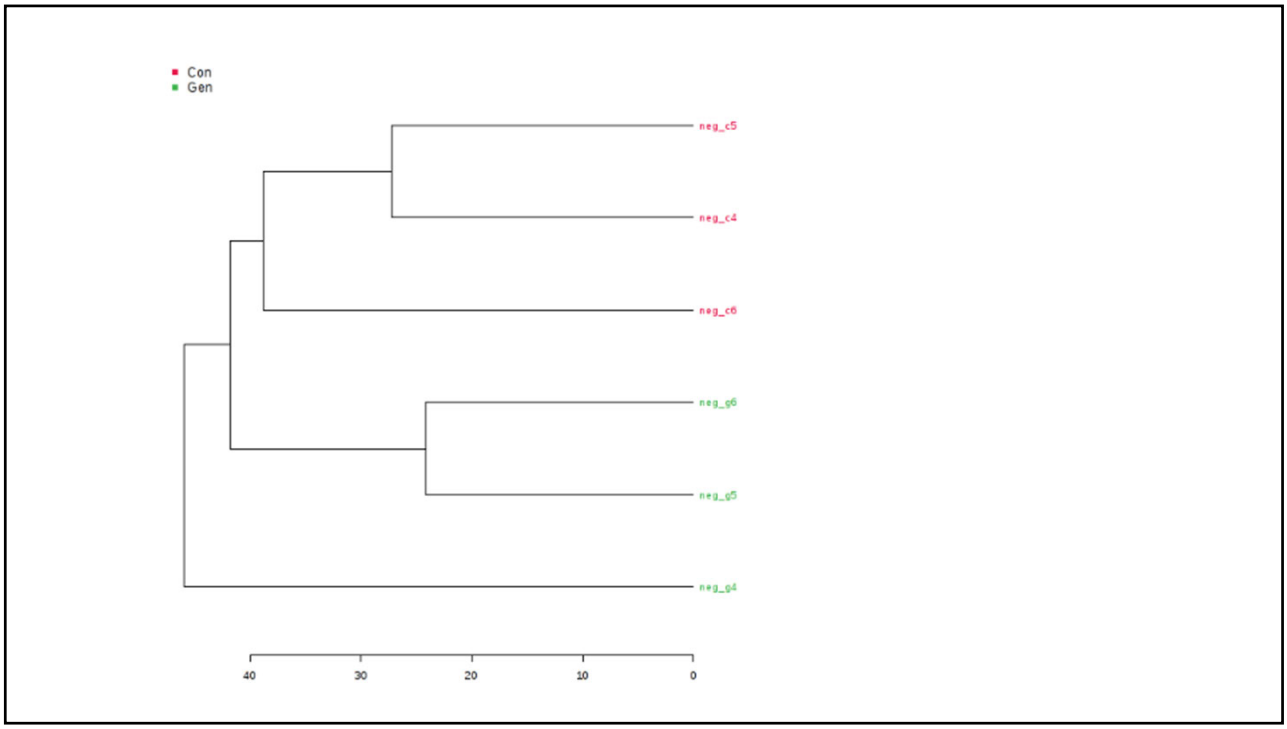
Try to identify the ones you can – ^{13}C-isotopes and multimers

49

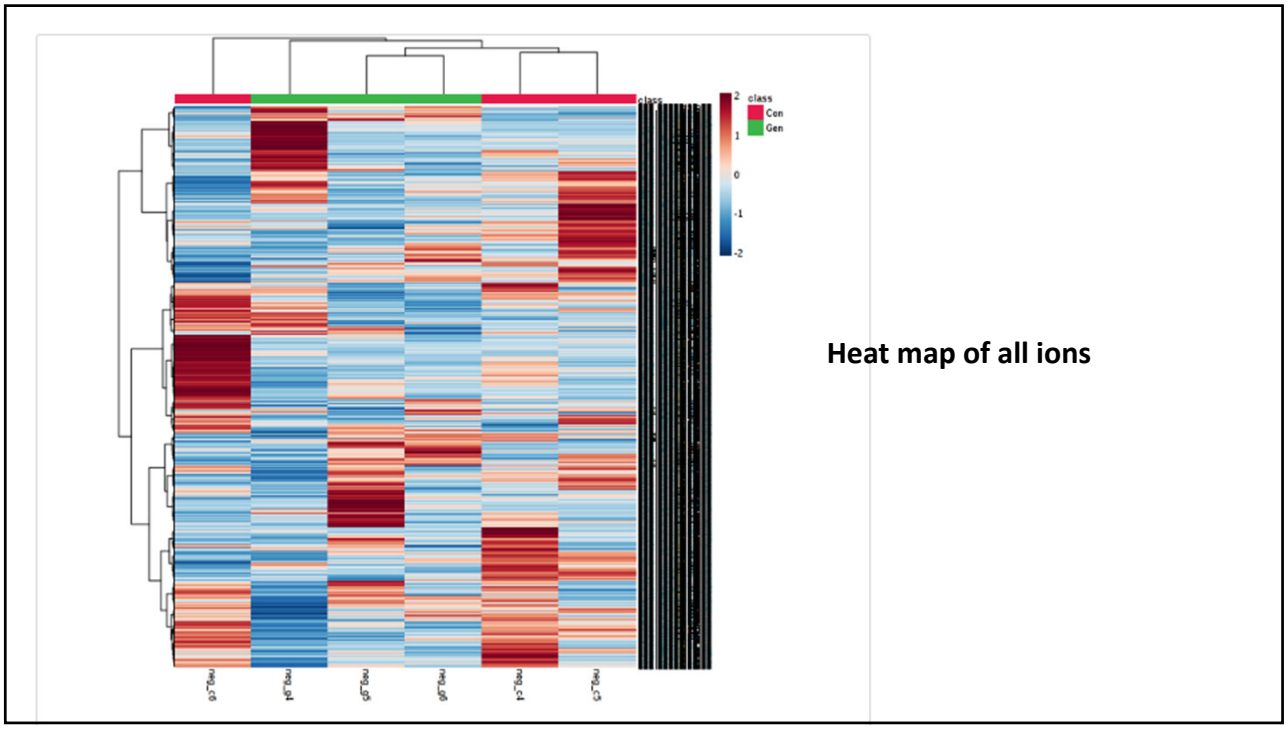


This is an orthogonal PLSDA plot – oPLSDA – you'll see this in published papers

50

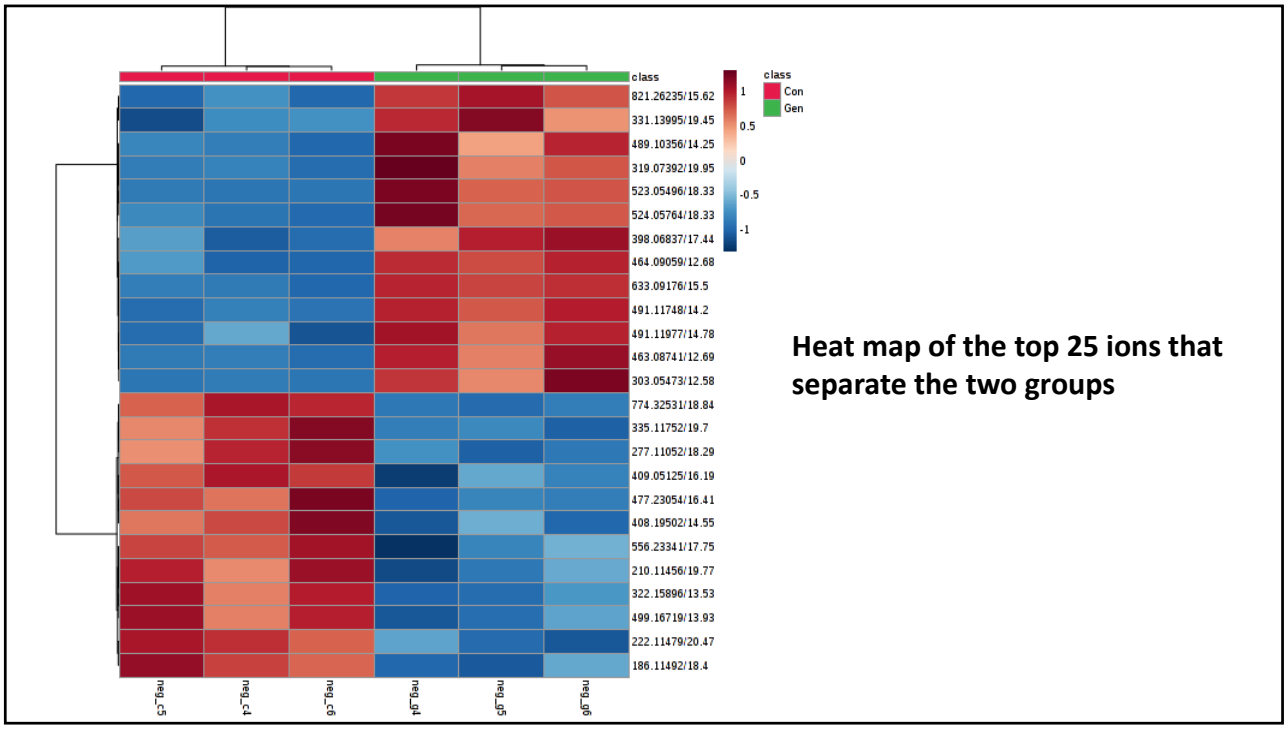


51



Heat map of all ions

52



Heat map of the top 25 ions that separate the two groups

53

After clicking download

54

Click

Generate Report

Analysis Report

Download.zip	opls_splot_0_dpi72.png
Rhistory.R	heatmap_1_.json
pls_score3d_0_.json	fc_0_dpi72.png
plsda_coef.csv	tree_0_dpi72.png
fold_change.csv	pls_imp_0_dpi72.png
oplsda_score.csv	opls_mdl_0_dpi72.png
volcano_0_dpi72.png	349.00299_18.1_dpi72.png
heatmap_1_dpi72.png	fc_1_dpi72.png
oplsda_splot.csv	pls_loading_0_dpi72.png
plsda_vip.csv	pca_score2d_0_dpi72.png

55

Metabolomic Data Analysis with MetaboAnalyst 4.0

Name: guest7521987882463575885

February 7, 2020

1 Data Processing and Normalization

1.1 Reading and Processing the Raw Data

MetaboAnalyst accepts a variety of data types generated in metabolomic studies, including compound concentration data, binned NMR/MS spectra data, NMR/MS peak list data, as well as MS spectra (NetCDF, mzXML, mzDATA). Users need to specify the data types when uploading their data in order for MetaboAnalyst to select the correct algorithm to process them. Table 1 summarizes the result of the data processing steps.

1.1.1 Reading MS Peak List and Intensities Data

MS peak list and intensities data should be uploaded as one zip file. It contains subfoulders with one for each group. Each folder contains peak list files, one per spectrum. The MS peak list format is either a two-column (mass and intensities) or three-column (mass, retention time, and intensities) comma separated values. The first line is assumed to be column labels. The files should be saved in .csv format. For paired analysis, users need to upload separately a text file specifying the paired information. Each pair is indicated by their sample names seperated by a colon ":" with one pair per line.

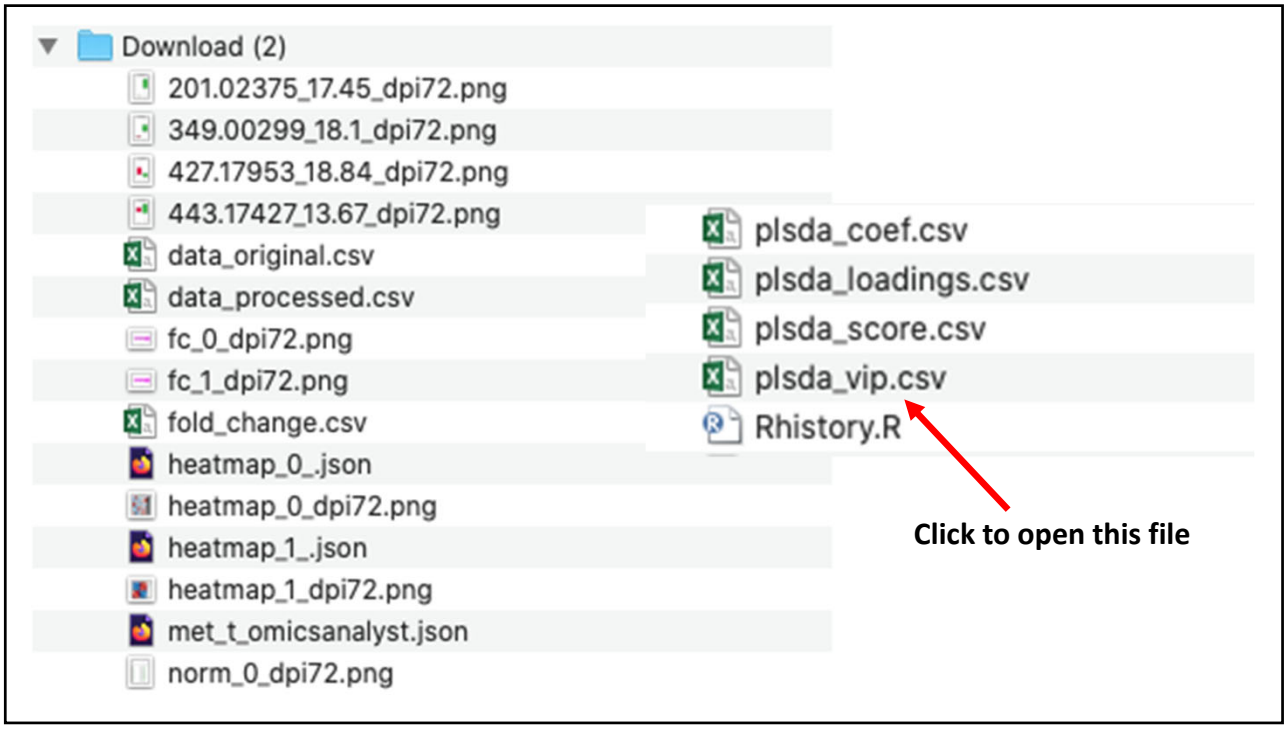
The uploaded files are peak lists and intensities data. A total of 6 samples were found. These samples contain a total of 29994 peaks, with an average of 4999 peaks per sample

Save this PDF file as a permanent record of the stats analysis

56


Generate Report
Click to download
Analysis Report
Download.zip
Rhistory.R
pls_score3d_0_.json
plsda_coef.csv
fold_change.csv
oplsda_score.csv
volcano_0_dpi72.png
heatmap_1_dpi72.png
oplsda_splot.csv
plsda_vip.csv
met_t_omicsanalyst.json
opls_splot_0_dpi72.png
heatmap_1_.json
fc_0_dpi72.png
tree_0_dpi72.png
pls_imp_0_dpi72.png
opls_mdl_0_dpi72.png
349.00299_18.1_dpi72.png
fc_1_dpi72.png
pls_loading_0_dpi72.png
pca_score2d_0_dpi72.png
data_processed.csv

57


Download (2)
201.02375_17.45_dpi72.png
349.00299_18.1_dpi72.png
427.17953_18.84_dpi72.png
443.17427_13.67_dpi72.png
data_original.csv
data_processed.csv
fc_0_dpi72.png
fc_1_dpi72.png
fold_change.csv
heatmap_0_.json
heatmap_0_dpi72.png
heatmap_1_.json
heatmap_1_dpi72.png
met_t_omicsanalyst.json
norm_0_dpi72.png
plsda_coef.csv
plsda_loadings.csv
plsda_score.csv
plsda_vip.csv
Rhistory.R
Click to open this file

58

	A	B	C	D
1		Comp. 1	Comp. 2	Comp. 3
2	61.99289/15	0.090303	0.1383	0.13957
3	61.99239/14	0.10955	0.11619	0.11844
4	61.99283/14	0.12453	0.12386	0.12377
5	61.99277/14	0.028696	0.037486	0.037772
6	61.99271/18	0.23098	0.23773	0.23708
7	61.99234/17	0.47814	0.46984	0.47234
8	61.99267/19	0.13618	0.26042	0.40942
9	61.99287/13	0.50075	0.51901	0.51838
10	61.99333/19	0.050071	0.073464	0.11908
11	96.96241/11	0.049401	0.049917	0.049868
12	107.05247/1	0.0075746	0.063315	0.07338
13	112.98773/1	0.75982	0.75378	0.75586
14	112.98779/1	0.63804	0.62778	0.63235
15	112.9878/14	0.056392	0.075238	0.091758

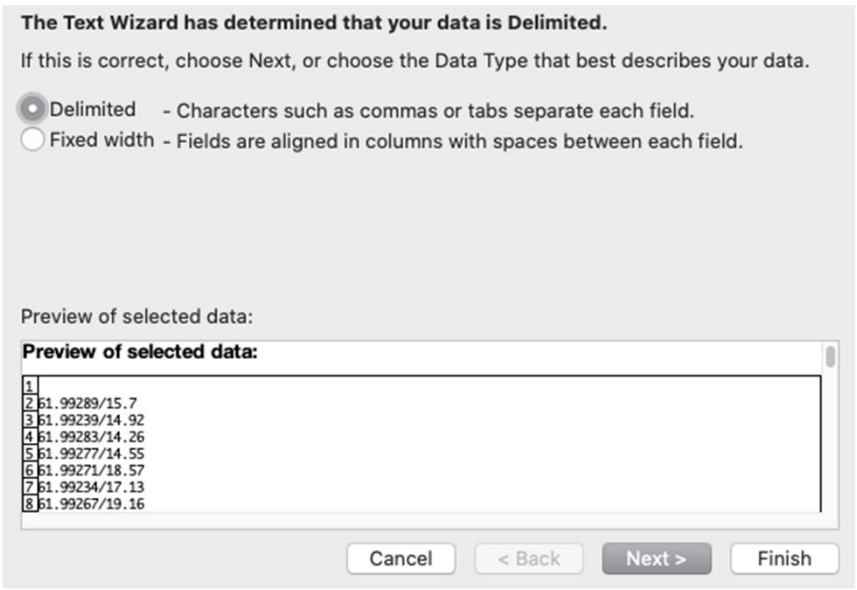
The *m/z* and RT information is concatenated

59

	A	B	C	D	E
1			Comp. 1	Comp. 2	Comp. 3
2	61.99289/15.7		0.090303	0.1383	0.13957
3	61.99239/14.92		0.10955	0.11619	0.11844
4	61.99283/14.26		0.12453	0.12386	0.12377
5	61.99277/14.55		0.028696	0.037486	0.037772
6	61.99271/18.57		0.23098	0.23773	0.23708
7	61.99234/17.13		0.47814	0.46984	0.47234
8	61.99267/19.16		0.13618	0.26042	0.40942
9	61.99287/13.74		0.50075	0.51901	0.51838
10	61.99333/19.6		0.050071	0.073464	0.11908
11	96.96241/11.08		0.049401	0.049917	0.049868
12	107.05247/15.15		0.0075746	0.063315	0.07338
13	112.98773/15.79		0.75982	0.75378	0.75586
14	112.98779/14.53		0.63804	0.62778	0.63235
15	112.9878/14.94		0.056392	0.075238	0.091758
16	112.98773/16.6		0.3355	0.32963	0.32863

Insert a new column (B)

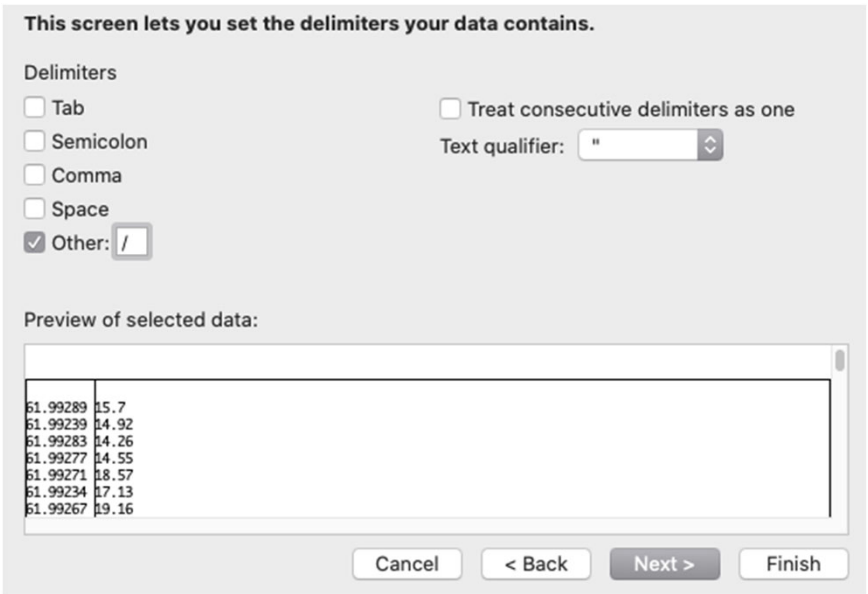
60



Go to data and "Text to column"

Select Next

61



Set the delimiters

Note how the concatenated item has been separated

62

	A	B	C	D	E
1	mass	RT	Comp. 1	Comp. 2	Comp. 3
2	61.99289	15.7	0.090303	0.1383	0.13957
3	61.99239	14.92	0.10955	0.11619	0.11844
4	61.99283	14.26	0.12453	0.12386	0.12377
5	61.99277	14.55	0.028696	0.037486	0.037772
6	61.99271	18.57	0.23098	0.23773	0.23708
7	61.99234	17.13	0.47814	0.46984	0.47234
8	61.99267	19.16	0.13618	0.26042	0.40942
9	61.99287	13.74	0.50075	0.51901	0.51838
10	61.99333	19.6	0.050071	0.073464	0.11908
11	96.96241	11.08	0.049401	0.049917	0.049868
12	107.05247	15.15	0.0075746	0.063315	0.07338
13	112.98773	15.79	0.75982	0.75378	0.75586
14	112.98779	14.53	0.63804	0.62778	0.63235
15	112.9878	14.94	0.056392	0.075238	0.091758
16	112.98773	16.6	0.3355	0.32963	0.32863
17	112.98775	15.31	0.83102	0.82943	0.82709
18	112.98782	13.89	0.11919	0.14435	0.16771
19	112.9877	12.3	2.5067	2.4651	2.457

Add headings to columns A and B

63

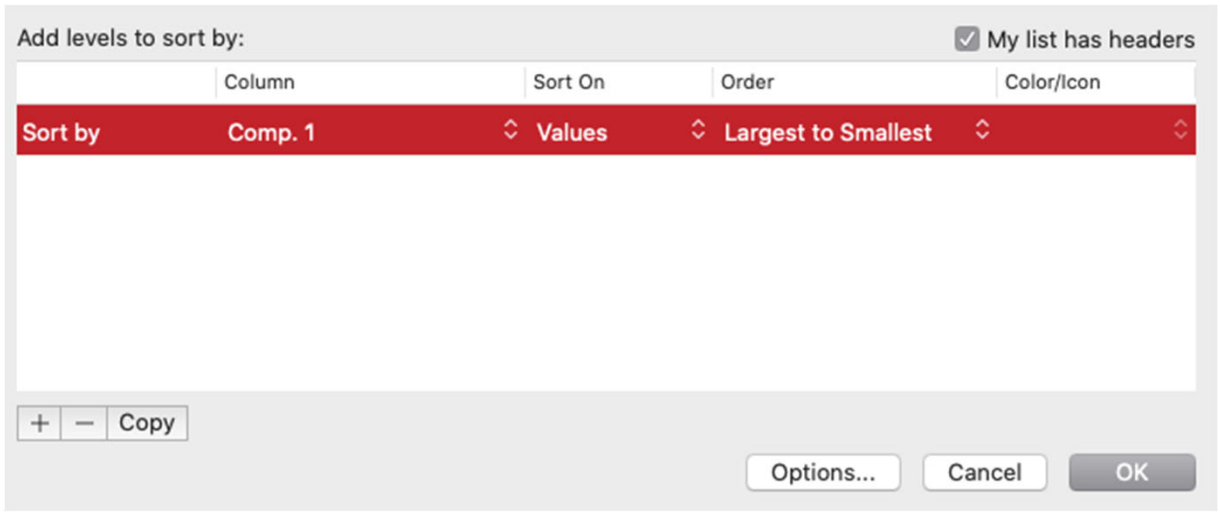
	A	B	C	D	E
1	mass	RT	Comp. 1	Comp. 2	Comp. 3
2	349.00299	18.1	17.817	17.55	17.501
3	201.02375	17.45	17.508	17.411	17.36
4	445.07732	15.54	13.742	13.502	13.474
5	297.09894	16.73	12.176	12.12	12.082
6	595.20281	16.7	11.452	11.344	11.335
7	427.17953	18.84	9.6997	9.9662	9.9341
8	350.00607	18.1	7.4419	7.334	7.3134
9	891.16249	15.52	7.3569	7.2547	7.2389
10	525.03334	13.52	6.6131	6.5834	6.5658
11	446.08157	15.55	6.5943	6.4787	6.4649
12	319.1225	18.45	6.5077	6.4327	6.4229
13	596.20615	16.7	6.415	6.3529	6.3489
14	269.04658	17.87	6.3683	6.2845	6.2667
15	411.12699	11.13	6.2377	6.1476	6.1438
16	823.26002	11.13	6.1736	6.0729	6.0549
61	699.24661	14.48	3.0521	2.9996	2.9916
62	242.1407	14.89	3.043	2.9902	2.9803
63	269.04632	20.54	3.0379	3.0067	2.9987
64	668.11943	15.52	3.0214	2.9803	2.9744

Sort the data according to Comp 1 and largest to smallest and select those with Comp 1 >3

There are 63 ions that fit this.

Copy these lines to two new sheets

64



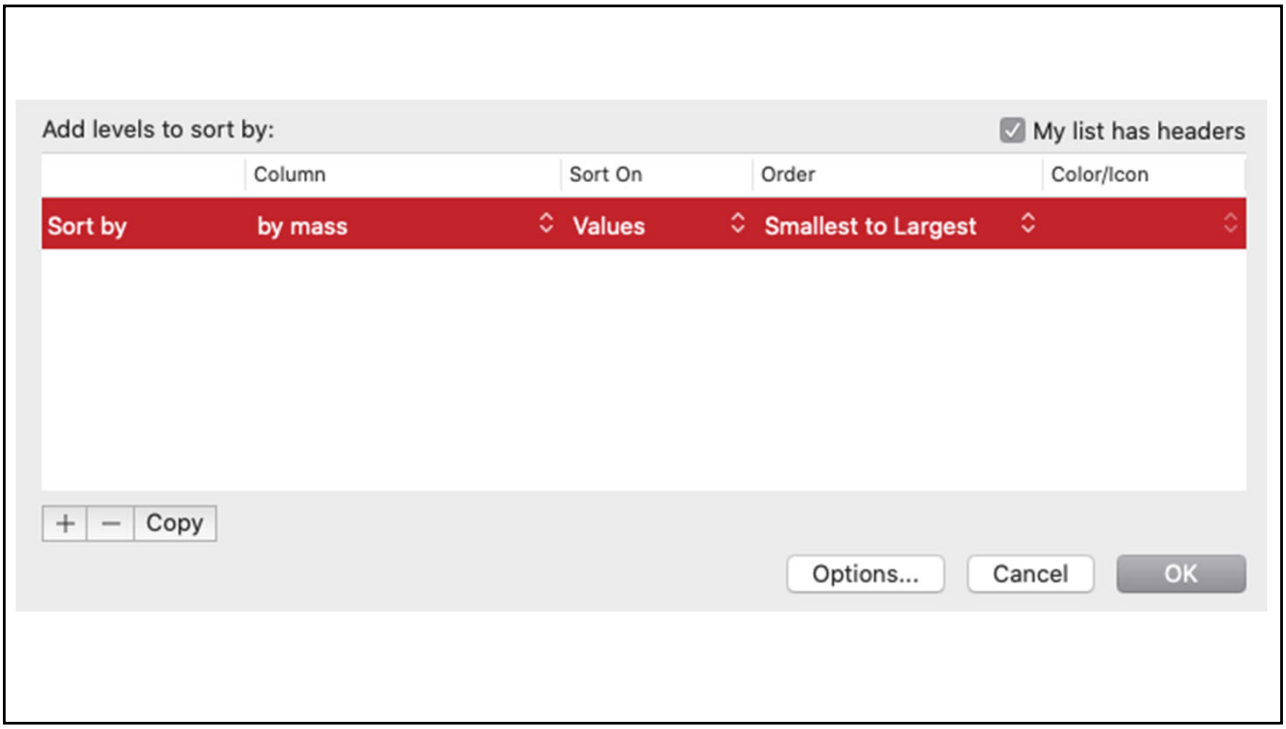
65

1	by mass	RT	Comp. 1	Comp. 2	Comp. 3
2	172.09962	15.71	3.217	3.2033	3.1932
3	186.11528	17.78	4.9117	4.8352	4.8199
4	198.11498	18.67	4.7213	4.6687	4.6558
5	201.02366	17.72	5.6374	5.6844	5.6848
6	201.02375	17.45	17.508	17.411	17.36
7	202.02746	17.43	5.6319	5.5943	5.58
8	203.02021	17.46	4.3899	4.3613	4.3497
9	214.10998	17.38	3.777	3.7226	3.7107
10	216.12551	16.5	4.3984	4.3781	4.3637
11	242.1407	14.89	3.043	2.9902	2.9803
12	269.04632	20.54	3.0379	3.0067	2.9987
13	269.0464	18.1	5.0041	4.9337	4.9244
14	269.04648	15.53	3.6337	3.5701	3.5622
15	269.04658	17.87	6.3683	6.2845	6.2667
16	275.097	18.64	5.1923	5.1321	5.1669
17	275.09704	18.17	4.2569	4.2184	4.2328
18	291.09146	17.9	3.3351	3.2838	3.2797
19	291.09165	17.07	5.783	5.828	5.8109
20	297.09894	16.73	12.176	12.12	12.082
21	298.10317	16.7	4.8879	4.8628	4.8483
22	309.10186	15.65	3.5386	3.5236	3.5141
23	309.10207	15.41	3.3566	3.4641	3.4532
24	317.06893	14.53	3.799	3.7338	3.7381
25	319.1225	18.45	6.5077	6.4327	6.4229

The data are sorted by mass. The m/z of a ^{13}C-isotope of a metabolite will be ~1.003 amu larger. It also has to have the same retention time. Two ^{13}C-isotopes are highlighted in yellow. Note that the comp 1 value will be smaller for the ^{13}C-isotope.

As an exercise, find all the other isotope pairs.

66



67

	A	B	C	D	E
1	mass	RT	Comp. 1	Comp. 2	Comp. 3
2	411.12699	11.13	6.2377	6.1476	6.1438
3	823.26002	11.13	6.1736	6.0729	6.0549
4	824.26237	11.13	3.7999	3.7382	3.7273
5	373.1102	12.77	3.3894	3.3484	3.3558
6	404.192	13.33	5.0639	4.9752	4.9589
7	525.03334	13.52	6.6131	6.5834	6.5658
8	526.03718	13.53	3.2907	3.2759	3.2669
9	525.03344	14.14	5.6579	5.6157	5.6009
10	699.24661	14.48	3.0521	2.9996	2.9916
11	317.06893	14.53	3.799	3.7338	3.7381
12	242.1407	14.89	3.043	2.9902	2.9803
13	309.10207	15.41	3.3566	3.4641	3.4532
14	891.16249	15.52	7.3569	7.2547	7.2389
15	892.16544	15.52	5.1786	5.1087	5.0973
16	893.17404	15.52	4.5894	4.5276	4.5175
17	668.11943	15.52	3.0214	2.9803	2.9744
18	269.04648	15.53	3.6337	3.5701	3.5622
19	445.07732	15.54	13.742	13.502	13.474
20	447.09202	15.54	6.1545	6.0467	6.0316
21	446.08157	15.55	6.5943	6.4787	6.4649

The data are sorted by RTs. The *m/z* of a ^{13}C-isotope of a metabolite will be ~1.003 amu larger. It also has to have the same retention time. Note that the comp 1 value will be smaller for the ^{13}C-isotope.

As an exercise, find all the other isotope pairs and the dimers $[2M-H]^-$ and $[2M-2H]^{2-}$.

68