Analyzing MS-DIAL data

Stephen Barnes, PhD

BBRB 711

sbarnes@uab.edu

1

Open the Area.txt file from MS-DIAL in Excel

A	B	C	D	E	F	G	H	I	J	K	L	M	N
Alignment ID	Average Rt(r	Average Mz	Metabolite n	Adduct type	Post curatior	Fill %	MS/MS assig	Reference R	Reference m	Formula	Ontology	INCHIKEY	SMILES
0	25.967	61.80658	Unknown	[M-H]-		1	FALSE	null	null	null	null	null	null
1	17.134	61.99174	Unknown	[M-H]-		0.667	TRUE	null	null	null	null	null	null
2	16.073	61.99218	Unknown	[M-H]-		0.167	TRUE	null	null	null	null	null	null
3	13.208	61.9929	Unknown	[M-C6H10O5	adduct linkec	0.167	TRUE	null	null	null	null	null	null
4	25.971	61.99305	Unknown	[M-H]-		1	TRUE	null	null	null	null	null	null
5	16.073	68.9986	Unknown	[M-H]-		0.167	TRUE	null	null	null	null	null	null
6	20.489	68.99867	Unknown	[M-H]-		0.167	FALSE	null	null	null	null	null	null
7	26.208	68.99916	Unknown	[M-H]-	found in high	1	TRUE	null	null	null	null	null	null
8	14.321	74.02563	w/o MS2:GL	[M-H]-	similar chror	0.333	FALSE	null	74.0242	C2H5NO2	Alpha amino	DHMQDGOQ	O=C(O)CN
9	19.987	74.02805	w/o MS2:GL	[M-H]-	found in high	0.167	TRUE	null	74.0242	C2H5NO2	Alpha amino	DHMQDGOQ	O=C(O)CN
10	16.531	74.02966	w/o MS2:GL	[M-H]-	found in high	0.167	TRUE	null	74.0242	C2H5NO2	Alpha amino	DHMQDGOQ	O=C(O)CN
11	17.43	74.03057	w/o MS2:GL	[M-H]-		0.167	FALSE	null	74.0242	C2H5NO2	Alpha amino	DHMQDGOQ	O=C(O)CN
12	16.202	76.02376	Unknown	[M-H]-	similar chror	0.167	FALSE	null	null	null	null	null	null
13	12.231	79.95722	Unknown	[M-H]-	found in high	0.167	FALSE	null	null	null	null	null	null
14	18.611	79.95961	Unknown	[M-H]-	found in high	0.333	TRUE	null	null	null	null	null	null
15	15.141	79.96063	Unknown	[M-H]-		0.5	TRUE	null	null	null	null	null	null
16	17.462	79.96455	Unknown	[M-H]-	found in high	0.333	TRUE	null	null	null	null	null	null
17	0.131	88.98872	Unknown	[M-H]-		0.333	TRUE	null	null	null	null	null	null
18	7.351	88.99012	Unknown	[M-H]-		0.167	FALSE	null	null	null	null	null	null
19	6.49	88.99024	Unknown	[M-H]-		0.167	TRUE	null	null	null	null	null	null
20	13.212	92.05163	Unknown	[M-C6H10O5	similar chror	0.667	TRUE	null	null	null	null	null	null
21	16.057	96.96115	w/o MS2:Ph	[M-H]-		0.333	TRUE	null	96.96962	H3O4P	Non-metal p	NBIIXXVUZA	O=P(O)(O)O

2

Delete the top four lines

	A	B	C	D	E	F	G	H	I	J	K	L	M	N
1	Alignment ID	Average Rt(r	Average Mz	Metabolite n	Adduct type	Post curatior	Fill %	MS/MS assig	Reference R	Reference m	Formula	Ontology	INCHIKEY	SMILES
2	0	25.967	61.80658	Unknown	[M-H]-		1	FALSE	null	null	null	null	null	null
3	1	17.134	61.99174	Unknown	[M-H]-		0.667	TRUE	null	null	null	null	null	null
4	2	16.073	61.99218	Unknown	[M-H]-		0.167	TRUE	null	null	null	null	null	null
5	3	13.208	61.9929	Unknown	[M-C6H10O5	adduct linkec	0.167	TRUE	null	null	null	null	null	null
6	4	25.971	61.99305	Unknown	[M-H]-		1	TRUE	null	null	null	null	null	null
7	5	16.073	68.9986	Unknown	[M-H]-		0.167	TRUE	null	null	null	null	null	null
8	6	20.489	68.99867	Unknown	[M-H]-		0.167	FALSE	null	null	null	null	null	null
9	7	26.208	68.99916	Unknown	[M-H]-	found in high	1	TRUE	null	null	null	null	null	null
10	8	14.321	74.02563	w/o MS2:GL	[M-H]-	similar chror	0.333	FALSE	null	74.0242	C2H5NO2	Alpha amino	DHMQDGOQ	O=C(O)CN
11	9	19.987	74.02805	w/o MS2:GL	[M-H]-	found in high	0.167	TRUE	null	74.0242	C2H5NO2	Alpha amino	DHMQDGOQ	O=C(O)CN
12	10	16.531	74.02966	w/o MS2:GL	[M-H]-	found in high	0.167	TRUE	null	74.0242	C2H5NO2	Alpha amino	DHMQDGOQ	O=C(O)CN
13	11	17.43	74.03057	w/o MS2:GL	[M-H]-		0.167	FALSE	null	74.0242	C2H5NO2	Alpha amino	DHMQDGOQ	O=C(O)CN
14	12	16.202	76.02376	Unknown	[M-H]-	similar chror	0.167	FALSE	null	null	null	null	null	null
15	13	12.231	79.95722	Unknown	[M-H]-	found in high	0.167	FALSE	null	null	null	null	null	null
16	14	18.611	79.95961	Unknown	[M-H]-	found in high	0.333	TRUE	null	null	null	null	null	null
17	15	15.141	79.96063	Unknown	[M-H]-		0.5	TRUE	null	null	null	null	null	null
18	16	17.462	79.96455	Unknown	[M-H]-	found in high	0.333	TRUE	null	null	null	null	null	null
19	17	0.131	88.98872	Unknown	[M-H]-		0.333	TRUE	null	null	null	null	null	null
20	18	7.351	88.99012	Unknown	[M-H]-		0.167	FALSE	null	null	null	null	null	null
21	19	6.49	88.99024	Unknown	[M-H]-		0.167	TRUE	null	null	null	null	null	null
22	20	13.212	92.05163	Unknown	[M-C6H10O5	similar chror	0.667	TRUE	null	null	null	null	null	null
23	21	16.057	96.96115	w/o MS2:Phe	[M-H]-		0.333	TRUE	null	96.96962	H3O4P	Non-metal p	NBIIXXVUZA	O=P(O)(O)O
24	22	26.031	96.96125	Phosphoric a	[M-H]-		1	TRUE	null	96.96962	H3O4P	Non-metal p	NBIIXXVUZA	O=P(O)(O)O
25	23	17.669	96.96135	w/o MS2:Phe	[M-H]-	found in high	0.167	TRUE	null	96.96962	H3O4P	Non-metal p	NBIIXXVUZA	O=P(O)(O)O
26	24	18.241	96.96317	Phosphoric a	[M-H]-	found in high	0.167	TRUE	null	96.96962	H3O4P	Non-metal p	NBIIXXVUZA	O=P(O)(O)O

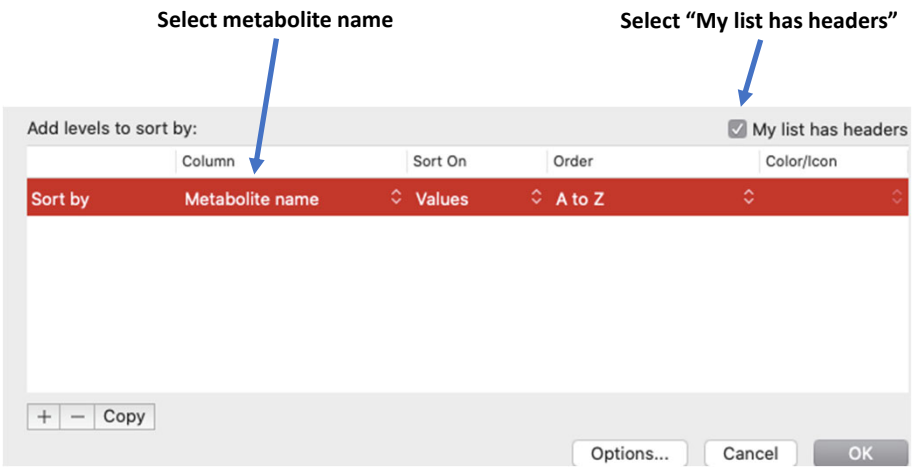
3

Now highlight all the cells

	A	B	C	D	E	F	G	H	I	J	K	L	M	N
1	Alignment ID	Average Rt(r	Average Mz	Metabolite n	Adduct type	Post curatior	Fill %	MS/MS assig	Reference R	Reference m	Formula	Ontology	INCHIKEY	SMILES
2	0	25.967	61.80658	Unknown	[M-H]-		1	FALSE	null	null	null	null	null	null
3	1	17.134	61.99174	Unknown	[M-H]-		0.667	TRUE	null	null	null	null	null	null
4	2	16.073	61.99218	Unknown	[M-H]-		0.167	TRUE	null	null	null	null	null	null
5	3	13.208	61.9929	Unknown	[M-C6H10O5	adduct linkec	0.167	TRUE	null	null	null	null	null	null
6	4	25.971	61.99305	Unknown	[M-H]-		1	TRUE	null	null	null	null	null	null
7	5	16.073	68.9986	Unknown	[M-H]-		0.167	TRUE	null	null	null	null	null	null
8	6	20.489	68.99867	Unknown	[M-H]-		0.167	FALSE	null	null	null	null	null	null
9	7	26.208	68.99916	Unknown	[M-H]-	found in high	1	TRUE	null	null	null	null	null	null
10	8	14.321	74.02563	w/o MS2:GL	[M-H]-	similar chror	0.333	FALSE	null	74.0242	C2H5NO2	Alpha amino	DHMQDGOQ	O=C(O)CN
11	9	19.987	74.02805	w/o MS2:GL	[M-H]-	found in high	0.167	TRUE	null	74.0242	C2H5NO2	Alpha amino	DHMQDGOQ	O=C(O)CN
12	10	16.531	74.02966	w/o MS2:GL	[M-H]-	found in high	0.167	TRUE	null	74.0242	C2H5NO2	Alpha amino	DHMQDGOQ	O=C(O)CN
13	11	17.43	74.03057	w/o MS2:GL	[M-H]-		0.167	FALSE	null	74.0242	C2H5NO2	Alpha amino	DHMQDGOQ	O=C(O)CN
14	12	16.202	76.02376	Unknown	[M-H]-	similar chror	0.167	FALSE	null	null	null	null	null	null
15	13	12.231	79.95722	Unknown	[M-H]-	found in high	0.167	FALSE	null	null	null	null	null	null
16	14	18.611	79.95961	Unknown	[M-H]-	found in high	0.333	TRUE	null	null	null	null	null	null
17	15	15.141	79.96063	Unknown	[M-H]-		0.5	TRUE	null	null	null	null	null	null
18	16	17.462	79.96455	Unknown	[M-H]-	found in high	0.333	TRUE	null	null	null	null	null	null
19	17	0.131	88.98872	Unknown	[M-H]-		0.333	TRUE	null	null	null	null	null	null
20	18	7.351	88.99012	Unknown	[M-H]-		0.167	FALSE	null	null	null	null	null	null
21	19	6.49	88.99024	Unknown	[M-H]-		0.167	TRUE	null	null	null	null	null	null
22	20	13.212	92.05163	Unknown	[M-C6H10O5	similar chror	0.667	TRUE	null	null	null	null	null	null
23	21	16.057	96.96115	w/o MS2:Phe	[M-H]-		0.333	TRUE	null	96.96962	H3O4P	Non-metal p	NBIIXXVUZA	O=P(O)(O)O
24	22	26.031	96.96125	Phosphoric a	[M-H]-		1	TRUE	null	96.96962	H3O4P	Non-metal p	NBIIXXVUZA	O=P(O)(O)O
25	23	17.669	96.96135	w/o MS2:Phe	[M-H]-	found in high	0.167	TRUE	null	96.96962	H3O4P	Non-metal p	NBIIXXVUZA	O=P(O)(O)O
26	24	18.241	96.96317	Phosphoric a	[M-H]-	found in high	0.167	TRUE	null	96.96962	H3O4P	Non-metal p	NBIIXXVUZA	O=P(O)(O)O

4

Click on <data> and <sort>



5

Sorted by metabolite name

	Alignment ID	Average Rt(r	Average Mz	Metabolite name	Adduct type	Post curatior	Fill %	MS/MS assig	Reference R	Reference m
2	322	16.477	197.08145	(-)-Camphanic acid	[M-H]-		1	TRUE	null	197.08192
3	323	17.203	197.08147	(-)-Camphanic acid	[M-H]-		0.833	TRUE	null	197.08192
4	326	16.109	197.08331	(-)-Camphanic acid	[M-H]-	found in high	1	TRUE	null	197.08192
5	2717	12.467	375.13495	(-)-Riboflavin; LC-ESI-QTOF; MS2; CE	[M-H]-		0.167	TRUE	null	375.13101
6	105	25.663	149.01062	(R,R)-TARTARIC ACID	[M-H]-		0.333	TRUE	null	149.009
7	53	14.418	121.03027	2-Hydroxybenzaldehyde	[M-H]-	similar chror	1	TRUE	null	121.0295
8	131	11.452	160.04082	2,8-Quinolinediol	[M-H]-	found in high	1	TRUE	null	160.04041
9	150	15.215	165.05748	3-(3-Hydroxyphenyl)propionic acid	[M-H]-		0.333	TRUE	null	165.05573
10	501	14.526	217.10864	3-Hydroxysebacic acid	[M-H]-	similar chror	0.833	TRUE	null	217.10815
11	438	13.374	212.00157	3-Indoxyl sulfate; LC-ESI-QTOF; MS2; CE	[M-H]-	similar chror	1	TRUE	null	212.0023
12	261	12.12	188.03555	4-HYDROXY-2-QUINOLINECARBOXYLIC ACID	[M-H]-		1	TRUE	null	188.035
13	93	12.12	144.04584	4-Hydroxyquinoline	[M-H]-	similar chror	1	TRUE	2.933	144.04549
14	94	17.145	144.04784	4-Hydroxyquinoline	[M-H]-	similar chror	1	TRUE	2.933	144.04549
15	95	18.645	144.04805	4-Hydroxyquinoline	[M-H]-		0.5	TRUE	2.933	144.04549
16	5121	12.561	621.10785	4'-O-GlcA-7-O-GlcA Apigenin (NMR)	[M-H]-	adduct linkec	0.5	TRUE	null	621.10974
17	3560	18.074	431.2114	5-hydroxy-2,2,6,6-tetramethyl-4-[2-methyl-1-[2,4,6-t	[M-H]-		1	TRUE	null	431.2077
18	3490	23.751	426.96555	6:2 Fluorotelomer sulfonic acid	[M-H]-		0.833	TRUE	null	426.9679
19	121	25.48	157.03886	ALLANTOIN	[M-H]-		0.333	TRUE	null	157.03671
20	122	11.438	157.04041	ALLANTOIN	[M-H]-		0.333	TRUE	null	157.03671
21	1084	20.534	269.04648	Aloe-emodin	[M-H]-	similar chror	0.833	TRUE	5.889	269.04553
22	3734	16.341	445.07587	apigenin-7-O-glucuronide	[M-H]-	adduct linkec	0.5	TRUE	null	445.07761
23	3735	15.558	445.07611	apigenin-7-O-glucuronide	[M-H]-	adduct linkec	0.833	TRUE	null	445.07761
24	1082	17.634	269.04385	Apigenin; LC-ESI-QTOF; MS2; CE	[M-H]-	found in high	0.167	TRUE	null	269.04553
25	1083	18.079	269.04626	Apigenin; LC-ESI-QTOF; MS2; CE	[M-H]-	similar chror	0.5	TRUE	null	269.04553

6

Scroll down and across to “unknowns”

	A	B	C	D	E	F	G	H	I	J	K	L
82	22	26.031	96.96125	Phosphoric acid	[M-H]-		1	TRUE	null	96.96962	H3O4P	Non-metal p
83	24	18.241	96.96317	Phosphoric acid	[M-H]-	found in high	0.167	TRUE	null	96.96962	H3O4P	Non-metal p
84	753	6.508	243.06226	Pseudouridine	[M-H]-		0.833	TRUE	null	243.06226	C9H12N2O6	Nucleoside a
85	754	7.368	243.06299	Pseudouridine	[M-H]-		0.5	TRUE	null	243.06226	C9H12N2O6	Nucleoside a
86	2715	13.035	375.1297	RIBOFLAVIN	[M-H]-	adduct linkec	1	TRUE	null	375.13101	C17H20N4O(	Flavins
87	367	18.475	201.11378	Sebacic acid; LC-ESI-QTOF; MS2; CE	[M-H]-	similar chror	1	TRUE	null	201.11324	C10H18O4	Medium-cha
88	1756	19.114	311.15079	Thymol-beta-D-glucoside	[M-H]-		0.333	TRUE	null	311.14999	C16H24O6	
89	295	12.86	193.05223	trans-Ferulic acid	[M-H]-	found in high	0.833	TRUE	4.228	193.05063	C10H10O4	Hydroxycinn
90	29	20	112.98609	Trifluoroacetic acid	[M-H]-		0.167	TRUE	null	112.98559	C2HF3O2	PFSA
91	30	21.705	112.98636	Trifluoroacetic acid	[M-H]-		0.333	TRUE	null	112.98559	C2HF3O2	PFSA
92	31	25.769	112.98677	Trifluoroacetic acid	[M-H]-		0.5	TRUE	null	112.98559	C2HF3O2	PFSA
93	32	18.316	112.98697	Trifluoroacetic acid	[M-H]-	found in high	0.333	TRUE	null	112.98559	C2HF3O2	PFSA
94	33	26.191	112.98698	Trifluoroacetic acid	[M-H]-	found in high	1	TRUE	null	112.98559	C2HF3O2	PFSA
95	34	15.853	112.98714	Trifluoroacetic acid	[M-H]-	found in high	0.333	TRUE	null	112.98559	C2HF3O2	PFSA
96	35	22.366	112.98718	Trifluoroacetic acid	[M-H]-		0.167	TRUE	null	112.98559	C2HF3O2	PFSA
97	36	19.116	112.98722	Trifluoroacetic acid	[M-H]-		0.167	TRUE	null	112.98559	C2HF3O2	PFSA
98	37	15.28	112.98778	Trifluoroacetic acid	[M-H]-		0.167	TRUE	null	112.98559	C2HF3O2	PFSA
99	38	20.58	112.9887	Trifluoroacetic acid	[M-H]-		0.833	TRUE	null	112.98559	C2HF3O2	PFSA
100	39	17.252	112.98892	Trifluoroacetic acid	[M-H]-		1	TRUE	null	112.98559	C2HF3O2	PFSA
101	40	14.959	112.98996	Trifluoroacetic acid	[M-H]-		0.167	TRUE	null	112.98559	C2HF3O2	PFSA
102	41	17.957	112.9902	Trifluoroacetic acid	[M-H]-		0.333	TRUE	null	112.98559	C2HF3O2	PFSA
103	1757	25.368	311.16684	Triptophenolide	[M-H]-		1	TRUE	null	311.16525	C20H24O3	Oxosteroids
104	1758	22.277	311.16809	Triptophenolide	[M-H]-		0.5	TRUE	null	311.16525	C20H24O3	Oxosteroids
105	1759	23.956	311.16809	Triptophenolide	[M-H]-		0.167	TRUE	null	311.16525	C20H24O3	Oxosteroids
106	1760	24.112	311.16885	Triptophenolide	[M-H]-		1	TRUE	null	311.16525	C20H24O3	Oxosteroids
107	1761	22.875	311.16901	Triptophenolide	[M-H]-		0.667	TRUE	null	311.16525	C20H24O3	Oxosteroids
108	1762	21.719	311.16913	Triptophenolide	[M-H]-		1	TRUE	null	311.16525	C20H24O3	Oxosteroids
109	1763	22.684	311.17126	Triptophenolide	[M-H]-		0.333	TRUE	null	311.16525	C20H24O3	Oxosteroids

7

Right hand side of the selected ions

	AD	AE	AF	AG	AH	AI	AJ	AK	AL	AM	AN	AO	AP
1	Spectrum ref	MS1 isotopic	MS/MS spec	Neg_C4	Neg_C5	Neg_C6	Neg_G4	Neg_G5	Neg_G6	1	2	1	2
2	Neg_C4	197.08209:2(	59.01986:10	344920	83329	603885	116211	323322	242554	344044.51	227362.26	260279.12	104388.105
3	Neg_G5	197.08418:2(	55.05489:5 5	50701	44360	122057	49585	60899	59043	72372.8073	56509.1393	43144.5092	6067.57852
4	Neg_C5	197.08324:9	59.01529:52	550194	103353	1554724	204927	835196	438318	736090.099	492813.474	743328.747	318648.731
5	Neg_G6	375.13495:1(	155.07622:1(	50693	21724	47911	22717	34506	24888	40109.6439	27370.2689	15982.8625	6274.13484
6	Neg_G4	149.01062:4(	57.04531:2 5	46985	39226	20013	139683	29930	26305	35408.1113	65305.929	13885.8022	64437.584
7	Neg_C4	121.03027:7	92.02544:52	123805	58653	44688	33904	40858	75164	75715.5104	49975.2891	42228.0794	22089.1818
8	Neg_C5	160.04216:2(	58.84072:1 6	76374	44816	117040	42273	83257	42133	79409.8997	55887.7526	36207.314	23702.7164
9	Neg_G6	165.05573:2(	83.67742:10	2421	1935	2729	957	52151	24572	2361.28764	25893.3448	400.415093	25622.5063
10	Neg_C6	217.10864:1(	59.01612:17	44439	26070	144260	94534	55775	38390	71589.485	62899.7461	63600.7729	28742.0649
11	Neg_C4	212.00157:1(	77.03726:21	1847406	1126152	1065461	1118220	1648442	1335651	1346339.67	1367437.67	434996.224	266536.048
12	Neg_G5	188.03577:1(	79.96193:8 1	175359	104729	276318	121653	217057	97967	185468.888	145559.219	86240.2086	63041.5183
13	Neg_C6	144.04584:1(	58.35411:7 7	123917	64246	188987	86421	144416	60157	125716.552	96998.1745	62389.516	43113.3671
14	Neg_C5	144.04575:2(	66.00053:4 8	6804898	3596176	4852426	2638733	4603250	3554693	5084500.17	3598892.08	1616901	983004.025
15	Neg_C4	144.04805:7(	102.04022:2	122757	62658	52565	22165	66059	53345	79326.6602	47189.5137	37948.9525	22584.9747
16	Neg_G4	621.10785:4(	57.04045:28	1790	2116	1277	438935	43220	74094	1727.7902	185416.169	422.673767	220095.607
17	Neg_G5	431.2088:17(	59.01629:21	437000	150007	166958	157803	279458	171862	251321.854	203040.906	161025.615	66551.716
18	Neg_G6	426.96823:1(	64.97403:3 7	223847	4835	204270	304212	175943	197054	144317.184	225736.557	121190.734	68776.9159
19	Neg_C6	157.03886:2(	59.05852:5 6	41215	20490	73151	28803	14081	14583	44951.9577	19155.568	26528.7432	8358.32342
20	Neg_C6	157.04041:4(	71.02656:29	73055	46932	137093	51498	26769	41104	85693.2174	39790.2194	46390.5812	12416.76
21	Neg_G6	269.048:154(	59.01894:8 6	55273	21296	11868	2410236	213013	223041	29479.0072	948763.38	22830.126	1265682.56
22	Neg_G4	445.07587:1(	57.44416:10	6787	4324	2663	1767900	149479	209869	4591.28947	709082.995	2075.18383	917459.877
23	Neg_G5	445.07581:2(	57.04179:14	58854	29584	12120	12591124	3142208	4685160	33519.1982	6806164.33	23614.3989	5068973.66
24	Neg_G4	269.04385:4(	63.02591:11(	3452	2219	1348	15000309	1432	4181	2339.69963	5001974.1	1057.53692	8658812.13
25	Neg_G6	269.0466:70(	60.76401:7 6	644085	135367	95863	1231423	1319751	1164510	291771.51	1238561.33	305750.802	77865.9683
26	Neg_G4	297.15125:2(	112.99245:9	270012	300870	311577	709225	279298	310000	294152.854	432840.729	21581.3039	239847.25
27	Neg_C6	297.15259:1(	112.99438:3	67619	116564	320058	313777	85403	431420	168080.273	276866.977	133872.396	175936.617
28	Neg_G4	297.15158:2(	113.02833:5	308234	314936	295827	741284	264364	217827	306332.417	407824.995	9695.71266	289719.932
29	Neg_G6	187.09901:4(	57.03636:21	1035280	635437	1206034	864901	596228	526510	958916.854	662546.333	292863.241	178677.288
30	Neg_G6	891.16492:4(	55.01998:14	328	73	74	3320490	279701	467387	158.333725	1355859.43	146.806651	1704005.8

8

Open a new Excel file and copy the highlighted rows into it

Then go back the original file and scroll down to the end of the unknowns (no entries in this dataset)

Then scroll down to the very end of the file and collect metabolites beginning with X, Y and Z

9

Edit the compounds in the new file

	A	B	C	D	E	F	G	H	I	J	K
88	1756	19.114	311.15079	Thymol-beta	[M-H]-		0.333	TRUE	null	311.14999	C16H24O6
89	295	12.86	193.05223	trans-Ferulic	[M-H]-	found in high	0.833	TRUE	4.228	193.05063	C10H10O4
90	29	20	112.98609	Trifluoroacet	[M-H]-		0.167	TRUE	null	112.98559	C2HF3O2
91	30	21.705	112.98636	Trifluoroacet	[M-H]-		0.333	TRUE	null	112.98559	C2HF3O2
92	31	25.769	112.98677	Trifluoroacet	[M-H]-		0.5	TRUE	null	112.98559	C2HF3O2
93	32	18.316	112.98697	Trifluoroacet	[M-H]-	found in high	0.333	TRUE	null	112.98559	C2HF3O2
94	33	26.191	112.98698	Trifluoroacet	[M-H]-	found in high	1	TRUE	null	112.98559	C2HF3O2
95	34	15.853	112.98714	Trifluoroacet	[M-H]-	found in high	0.333	TRUE	null	112.98559	C2HF3O2
96	35	22.366	112.98718	Trifluoroacet	[M-H]-		0.167	TRUE	null	112.98559	C2HF3O2
97	36	19.116	112.98722	Trifluoroacet	[M-H]-		0.167	TRUE	null	112.98559	C2HF3O2
98	37	15.28	112.98778	Trifluoroacet	[M-H]-		0.167	TRUE	null	112.98559	C2HF3O2
99	38	20.58	112.9887	Trifluoroacet	[M-H]-		0.833	TRUE	null	112.98559	C2HF3O2
100	39	17.252	112.98892	Trifluoroacet	[M-H]-		1	TRUE	null	112.98559	C2HF3O2
101	40	14.959	112.98996	Trifluoroacet	[M-H]-		0.167	TRUE	null	112.98559	C2HF3O2
102	41	17.957	112.9902	Trifluoroacet	[M-H]-		0.333	TRUE	null	112.98559	C2HF3O2

Delete the trifluoroacetic acid

10

Delete more contaminants

40	275	26.155	191.01985	CITRATE	[M-H]-		0.833	TRUE	null	
41	274	15.74	191.01971	Citric acid	[M-H]-		0.167	TRUE	null	
42	276	20.462	191.02	Citric acid	[M-H]-		0.167	TRUE	null	
43	278	14.743	191.02106	Citric acid (Not validated, isomer of 22	[M-H]-		0.333	TRUE		1.64
44	279	6.506	191.02122	Citric acid (Not validated, isomer of 22	[M-H]-		0.667	TRUE		1.64
45	280	7.373	191.02129	Citric acid (Not validated, isomer of 22	[M-H]-		0.167	TRUE		1.64
46	294	21.401	193.03987	D-(+)-Galacturonic acid	[M-H]-	found in high	1	TRUE	null	
47	507	9.964	218.10304	D-PANTOTHENIC ACID	[M-H]-		0.833	TRUE	null	
48	833	17.487	249.03384	Diphenylphosphate	[M-H]-		0.5	TRUE	null	
49	218	11.453	181.05011	DL-3-(4-Hydroxyphenyl)lactic acid; LC-E	[M-H]-		0.833	TRUE	null	
50	1979	19.685	325.17944	Dodecylbenzenesulfonic acid	[M-H]-	adduct linkec	0.167	TRUE	null	
51	1981	20.806	325.18381	Dodecylbenzenesulfonic acid	[M-H]-		0.667	TRUE	null	
52	1982	21.765	325.18414	Dodecylbenzenesulfonic acid	[M-H]-		0.833	TRUE	null	
53	1983	24.591	325.18436	Dodecylbenzenesulfonic acid	[M-H]-		0.5	TRUE	null	
54	1984	25.694	325.18555	Dodecylbenzenesulfonic acid	[M-H]-		1	TRUE	null	
55	1985	22.492	325.18637	Dodecylbenzenesulfonic acid	[M-H]-		0.333	TRUE	null	
56	1986	22.731	325.1864	Dodecylbenzenesulfonic acid	[M-H]-		0.667	TRUE	null	
57	1987	24.423	325.18658	Dodecylbenzenesulfonic acid	[M-H]-		0.5	TRUE	null	
58	873	14.742	253.07675	FA 9:0 + 1O, sulfate; PlaSMA ID-416	[M-H]-		0.833	TRUE		4.24
59	290	14.139	193.03535	Glucuronate	[M-H]-	adduct linkec	0.667	TRUE	null	
60	73	10.066	131.03609	Glutaric acid; LC-ESI-QTOF; MS2; CE	[M-H]-		0.333	TRUE	null	

11

Order by retention time and delete after 20 min

	A	B	C	D	E	F	G	H	I	J	K
65	367	18.475	201.11378	Sebacic acid; LC-ESI-QTOF; MS2; CE	[M-H]-	similar chror	1	TRUE	null	201.11324	C10H18O4
66	95	18.645	144.04805	4-Hydroxyquinoline	[M-H]-		0.5	TRUE	2.933	144.04549	C9H7NO
67	1756	19.114	311.15079	Thymol-beta-D-glucoside	[M-H]-		0.333	TRUE	null	311.14999	C16H24O6
68	3902	19.719	459.08923	oroxindin	[M-H]-	adduct linkec	0.5	TRUE	null	459.09329	C22H20O11
69	276	20.462	191.02	Citric acid	[M-H]-		0.167	TRUE	null	191.01973	C6H8O7
70	1084	20.534	269.04648	Aloe-emodin	[M-H]-	similar chror	0.833	TRUE	5.889	269.04553	C15H10O5
71	294	21.401	193.03987	D-(+)-Galacturonic acid	[M-H]-	found in high	1	TRUE	null	193.03537	C6H10O7
72	1762	21.719	311.16913	Triptophenolide	[M-H]-		1	TRUE	null	311.16525	C20H24O3
73	1758	22.277	311.16809	Triptophenolide	[M-H]-		0.5	TRUE	null	311.16525	C20H24O3
74	592	22.447	227.13028	C12H20O4; PlaSMA ID-334	[M-H]-		0.5	TRUE	6.29	227.12781	C12H20O4
75	1763	22.684	311.17126	Triptophenolide	[M-H]-		0.333	TRUE	null	311.16525	C20H24O3
76	1761	22.875	311.16901	Triptophenolide	[M-H]-		0.667	TRUE	null	311.16525	C20H24O3
77	2550	23.099	362.96942	Perfluoroheptanoic acid; LC-ESI-ITFT; M	[M-H]-		0.833	TRUE	null	362.9696	C7HF13O2
78	3490	23.751	426.96555	6:2 Fluorotelomer sulfonic acid	[M-H]-		0.833	TRUE	null	426.9679	C8H5F13O3S
79	1759	23.956	311.16809	Triptophenolide	[M-H]-		0.167	TRUE	null	311.16525	C20H24O3
80	1034	24.044	265.14731	C12-AS (TENTATIVE)	[M-H]-		1	TRUE	null	265.14792	C12H26O4S
81	1534	24.072	297.15143	aurapten	[M-H]-		0.667	TRUE	null	297.1496	C19H22O3
82	1760	24.112	311.16885	Triptophenolide	[M-H]-		1	TRUE	null	311.16525	C20H24O3
83	1535	24.395	297.15158	aurapten	[M-H]-		0.333	TRUE	null	297.1496	C19H22O3
84	3270	24.617	412.96658	Perfluorooctanoic acid; LC-ESI-ITFT; MS	[M-H]-		1	TRUE	null	412.96643	C8HF15O2
85	1533	25.257	297.15125	aurapten	[M-H]-		1	TRUE	null	297.1496	C19H22O3
86	1757	25.368	311.16684	Triptophenolide	[M-H]-		1	TRUE	null	311.16525	C20H24O3
87	121	25.48	157.03886	ALLANTOIN	[M-H]-		0.333	TRUE	null	157.03671	C4H6N4O3
88	105	25.663	149.01062	(R,R)-TARTARIC ACID	[M-H]-		0.333	TRUE	null	149.009	C4H6O6
89	22	26.031	96.96125	Phosphoric acid	[M-H]-		1	TRUE	null	96.96962	H3O4P
90	275	26.155	191.01985	CITRATE	[M-H]-		0.833	TRUE	null	191.019	C6H8O7

12

Now make the .csv file for Metaboanalyst

	A	B	C	D	E	F	G	H	I	J
1	Average Rt(r	Average Mz			Neg_C4	Neg_C5	Neg_C6	Neg_G4	Neg_G5	Neg_
2	6.506	191.02122			235848	53153	85784	123138	35588	59
3	6.508	243.06226			19692	6511	16178	13515	7707	4292
4	7.368	243.06299			16732	3962	16253	9737	4639	2805
5	7.373	191.02129			92896	25257	44343	113153	5783	33098
6	7.76	129.02118			84099	35847	51680	58043	18988	27532
7	7.761	173.01138			213978	89710	150305	154546	50933	76814
8	8.241	173.01088			127133	26632	128326	42323	16503	19900
9	8.263	129.02121			44624	20650	43276	31488	18854	12860
10	8.809	283.06799			18749	6375	25046	20589	5816	4751
11	8.926	296.10205			34463	14379	51741	35751	13218	12828
12	9.83	310.11398			42950	17037	38915	41848	12340	13893
13	9.964	218.10304			63110	35459	4987	146302	26614	26211
14	10.066	131.03609			13679	5426	65176	41387	7318	5757
15	11.438	157.04041			73055	46932	137093	51498	26769	41104
16	11.452	160.04082			76374	44816	117040	42273	83257	42133
17	11.452	204.03236			349988	222678	540669	198630	414675	238475
18	11.453	181.05011			64747	39361	35054	46219	11776	32574
19	12.12	188.03555			175359	104729	276318	121653	217057	97967
20	12.12	144.04584			123917	64246	188987	86421	144416	60157
21	12.467	375.13495			50693	21724	47911	22717	34506	24888

Make the concatenated string in column C and copy (special) into column D

13

Added the concatenated string

	A	B	C	D	E	F	G	H	I	J
1	Average Rt(r	Average Mz	RT/MZ	RT/M	Neg_C4	Neg_C5	Neg_C6	Neg_G4	Neg_G5	Neg_G6
2	6.506	191.02122	6.51/191.02	6.51/	235848	53153	85784	123138	35588	56659
3	6.508	243.06226	6.51/243.06	6.51/243.06	19692	6511	16178	13515	7707	4292
4	7.368	243.06299	7.37/243.06	7.37/243.06	16732	3962	16253	9737	4639	2805
5	7.373	191.02129	7.37/191.02	7.37/191.02	92896	25257	44343	113153	5783	33098
6	7.76	129.02118	7.76/129.02	7.76/129.02	84099	35847	51680	58043	18988	27532
7	7.761	173.01138	7.76/173.01	7.76/173.01	213978	89710	150305	154546	50933	76814
8	8.241	173.01088	8.24/173.01	8.24/173.01	127133	26632	128326	42323	16503	19900
9	8.263	129.02121	8.26/129.02	8.26/129.02	44624	20650	43276	31488	18854	12860
10	8.809	283.06799	8.81/283.06	8.81/283.06	18749	6375	25046	20589	5816	4751
11	8.926	296.10205	8.93/296.10	8.93/296.10	34463	14379	51741	35751	13218	12828
12	9.83	310.11398	9.83/310.11	9.83/310.11	42950	17037	38915	41848	12340	13893
13	9.964	218.10304	9.96/218.10	9.96/218.10	63110	35459	4987	146302	26614	26211
14	10.066	131.03609	10.07/131.0	10.07/131.0	13679	5426	65176	41387	7318	5757
15	11.438	157.04041	11.44/157.0	11.44/157.0	73055	46932	137093	51498	26769	41104
16	11.452	160.04082	11.45/160.0	11.45/160.0	76374	44816	117040	42273	83257	42133
17	11.452	204.03236	11.45/204.0	11.45/204.0	349988	222678	540669	198630	414675	238475
18	11.453	181.05011	11.45/181.0	11.45/181.0	64747	39361	35054	46219	11776	32574
19	12.12	188.03555	12.12/188.0	12.12/188.0	175359	104729	276318	121653	217057	97967
20	12.12	144.04584	12.12/144.0	12.12/144.0	123917	64246	188987	86421	144416	60157
21	12.467	375.13495	12.47/375.1	12.47/375.1	50693	21724	47911	22717	34506	24888
22	12.561	621.10785	12.56/621.1	12.56/621.1	1790	2116	1277	438935	43220	74094
23	12.586	367.10422	12.59/367.1	12.59/367.1	229088	204082	106350	147689	160699	141379

14

Completed .csv file (added a line for groups)

RT/MZ	Neg_C4	Neg_C5	Neg_C6	Neg_G4	Neg_G5	Neg_G6
Group	1	1	1	2	2	2
6.51/191.021	235848	53153	85784	123138	35588	56659
6.51/243.062	19692	6511	16178	13515	7707	4292
7.37/243.063	16732	3962	16253	9737	4639	2805
7.37/191.021	92896	25257	44343	113153	5783	33098
7.76/129.021	84099	35847	51680	58043	18988	27532
7.76/173.011	213978	89710	150305	154546	50933	76814
8.24/173.011	127133	26632	128326	42323	16503	19900
8.26/129.021	44624	20650	43276	31488	18854	12860
8.81/283.068	18749	6375	25046	20589	5816	4751
8.93/296.102	34463	14379	51741	35751	13218	12828
9.83/310.114	42950	17037	38915	41848	12340	13893
9.96/218.103	63110	35459	4987	146302	26614	26211
10.07/131.03	13679	5426	65176	41387	7318	5757
11.44/157.04	73055	46932	137093	51498	26769	41104
11.45/160.04	76374	44816	117040	42273	83257	42133
11.45/204.03	349988	222678	540669	198630	414675	238475
11.45/181.05	64747	39361	35054	46219	11776	32574
12.12/188.03	175359	104729	276318	121653	217057	97967
12.12/144.04	123917	64246	188987	86421	144416	60157

Now save as a .csv file (no spaces in name!)

15

Submit to Metaboanalyst

16

Data Integrity Check:

1. Checking the class labels - at least three replicates are required in each class.
2. If the samples are paired, the pair labels must conform to the specified format.
3. The data (except class labels) must not contain non-numeric values.
4. The presence of missing values or features with constant values (i.e. all zeros).

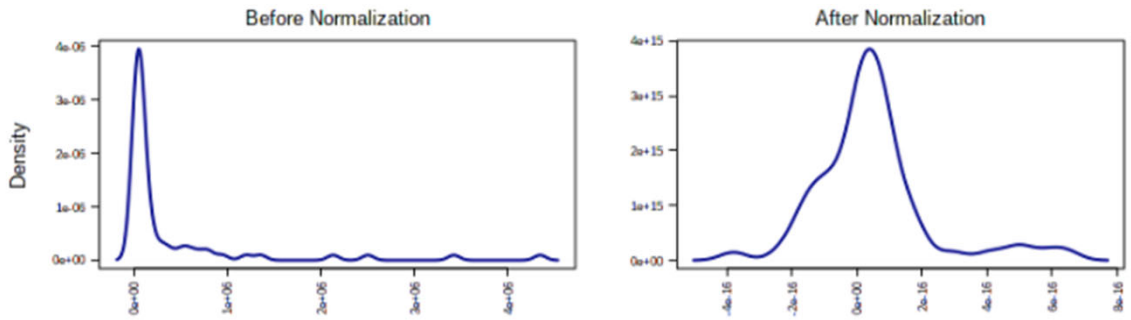
Data processing information:

Checking data content ...passed.
Samples are in columns and features in rows.
The uploaded file is in comma separated values (.csv) format.
The uploaded data file contains 6 (samples) by 67 (peaks(mz/rt)) data matrix.
Samples are not paired.
2 groups were detected in samples.
Only English letters, numbers, underscore, hyphen and forward slash (/) are allowed.
Other special characters or punctuations (if any) will be stripped off.
All data values are numeric.
A total of 0 (0%) missing values were detected.
By default, missing values will be replaced by 1/5 of min positive values of their corresponding variables
Click the **Skip** button if you accept the default practice;
Or click the **Missing value imputation** to use other methods.

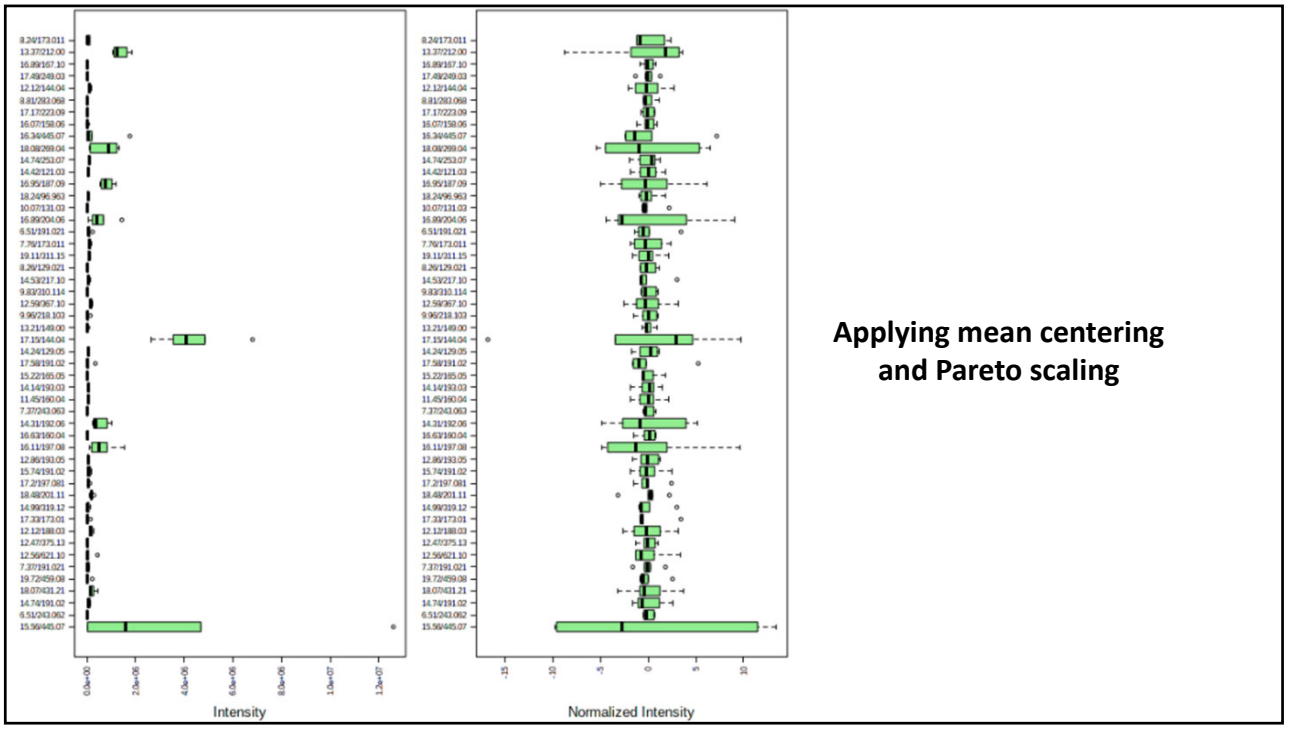
Edit Groups | Missing Values | → Proceed

17

Normalizing the data

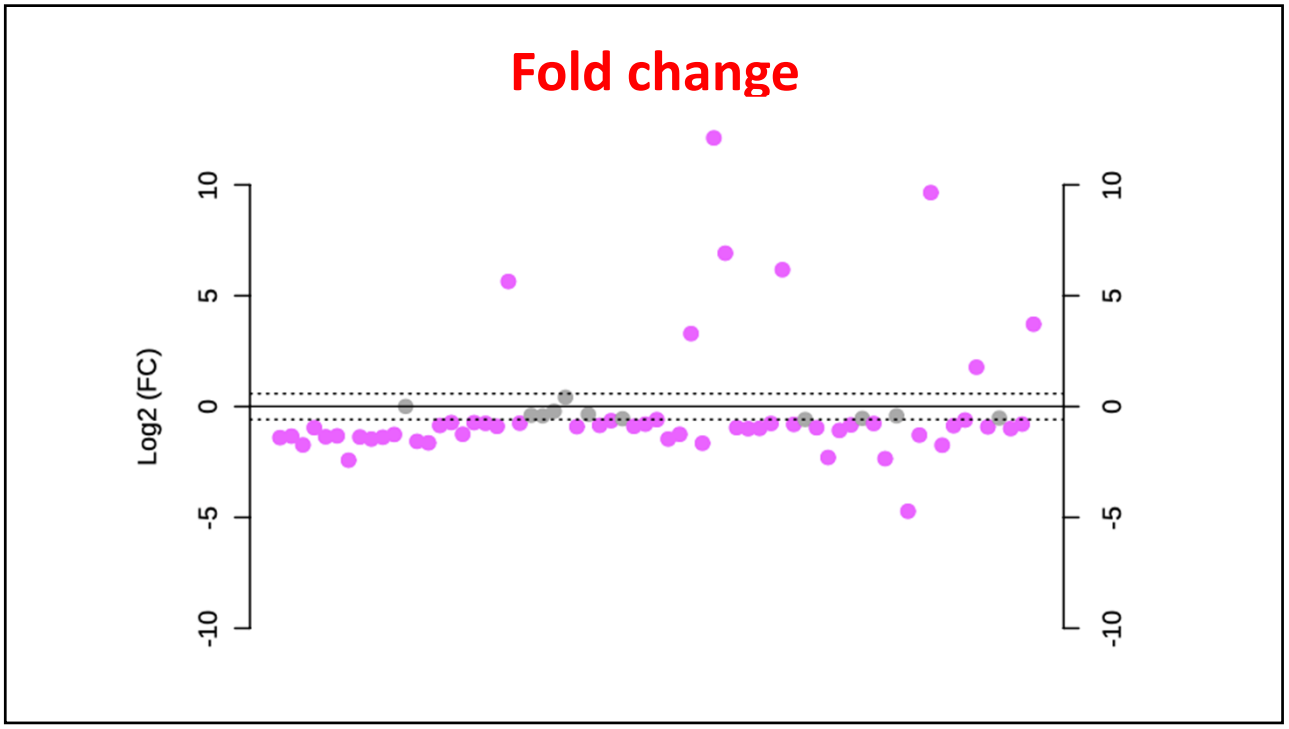


18

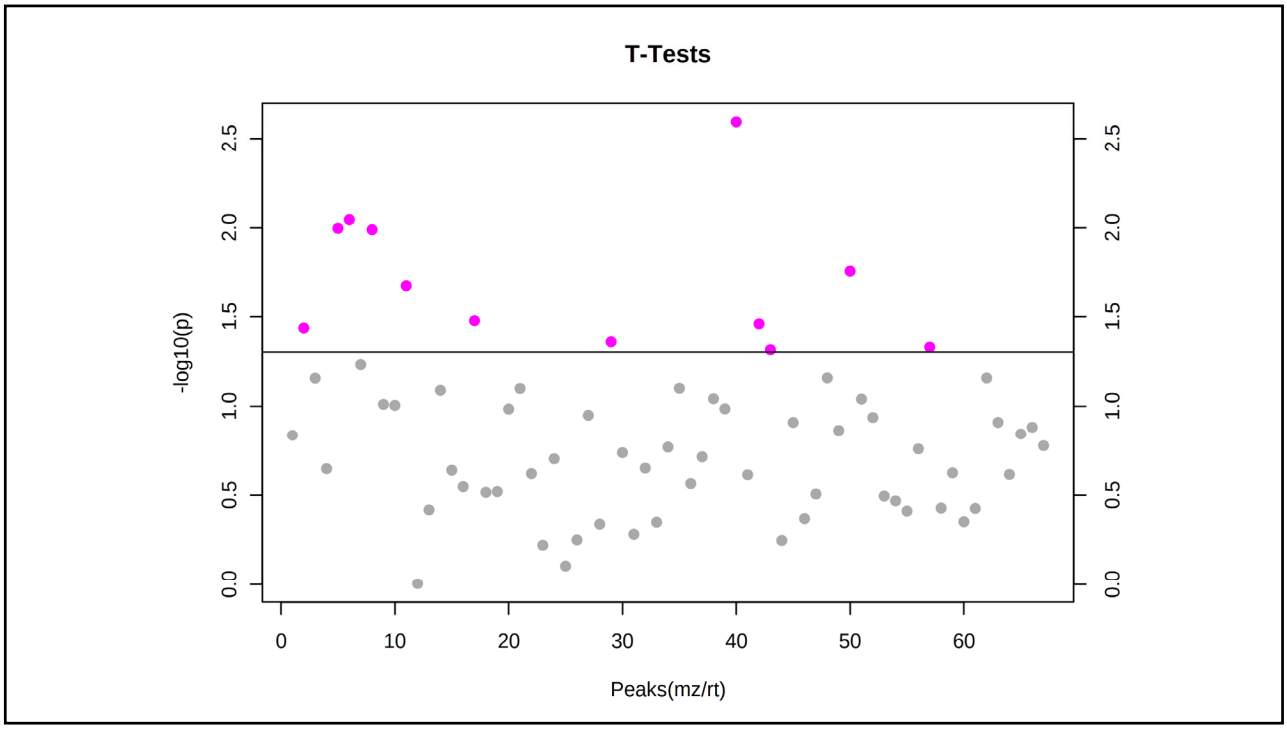


Applying mean centering and Pareto scaling

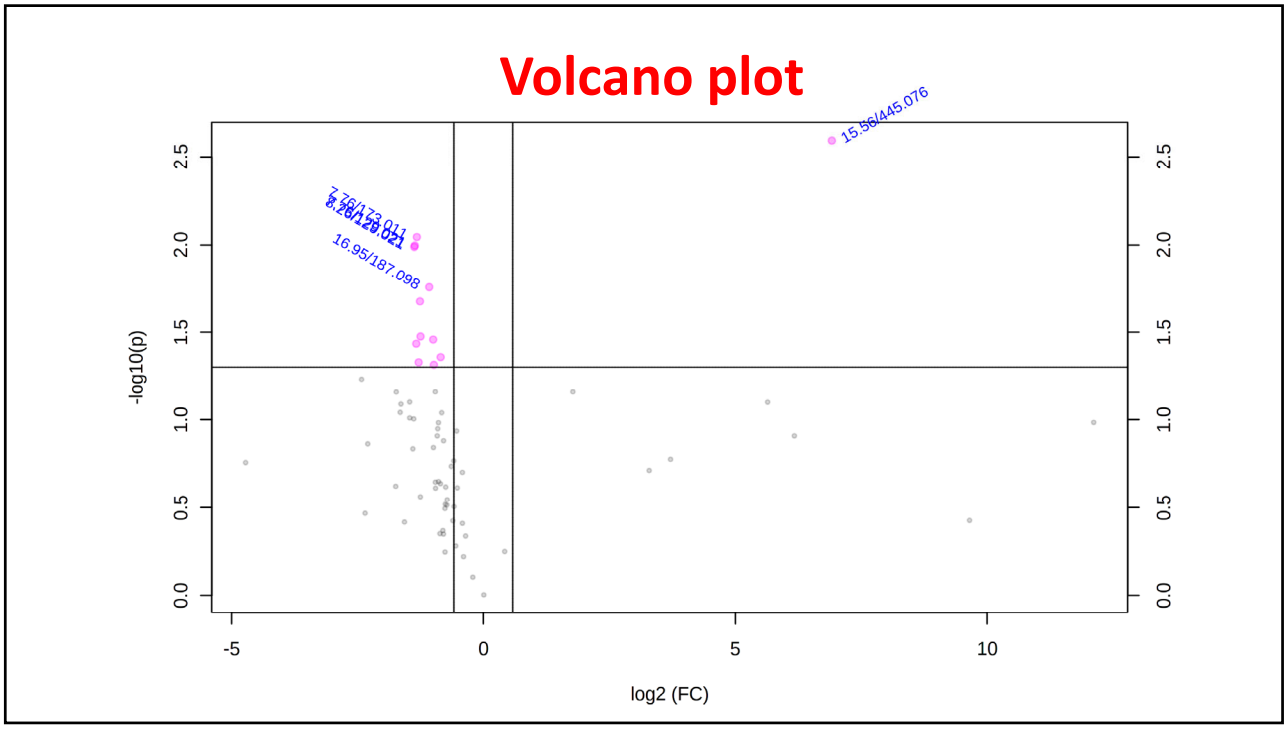
19



20

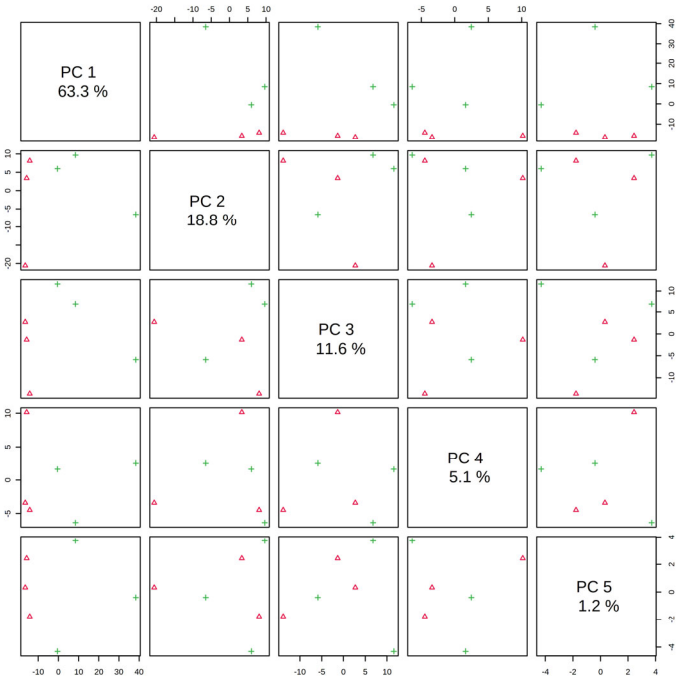

T-Tests
-log10(p)
Peaks(mz/rt)
0.0
0.5
1.0
1.5
2.0
2.5
0
10
20
30
40
50
60

21


Volcano plot
-log10(p)
log2 (FC)
15.56/445.076
7.76/173.011
6.28/128.021
16.95/187.098
0.0
0.5
1.0
1.5
2.0
2.5
-5
0
5
10

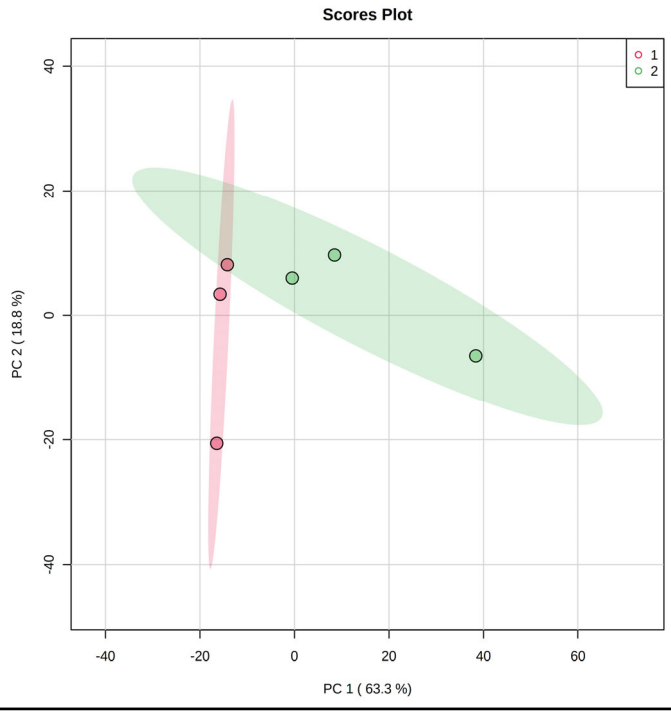
22

PCA plots



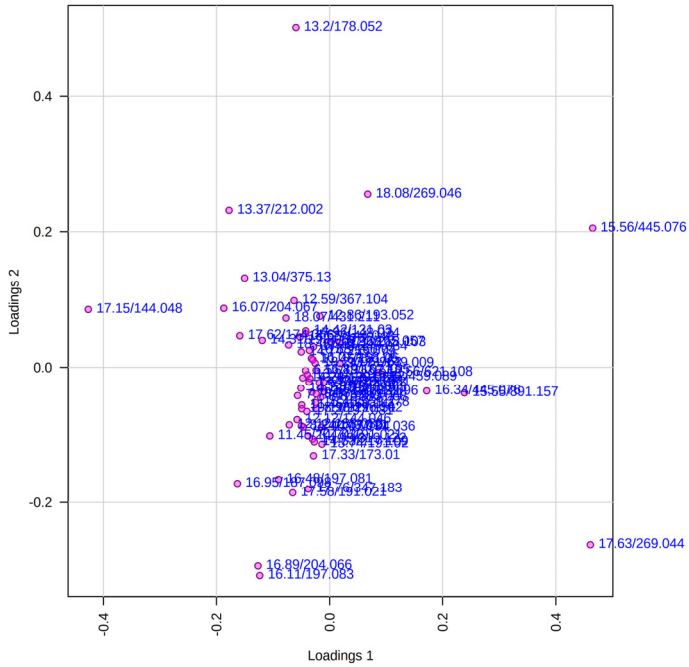
23

2D-PCA plot



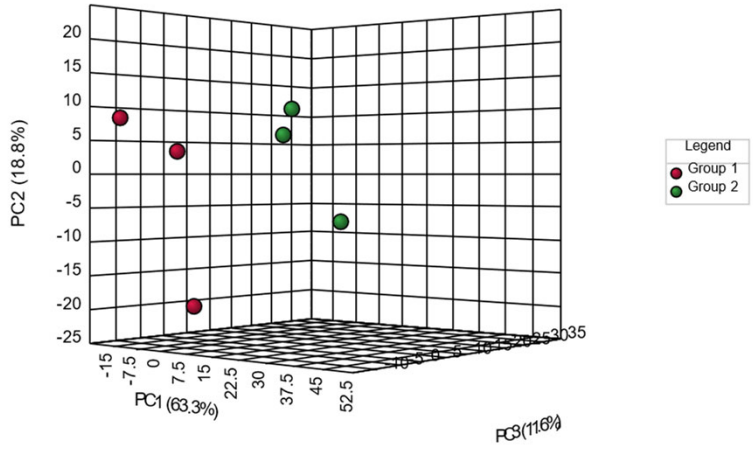
24

Loadings plot for PCA

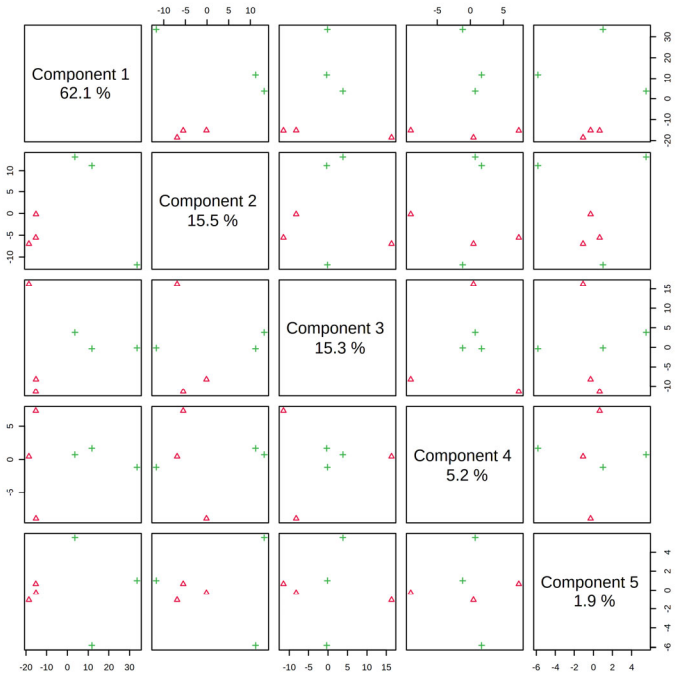


25

3D-PCA score plot

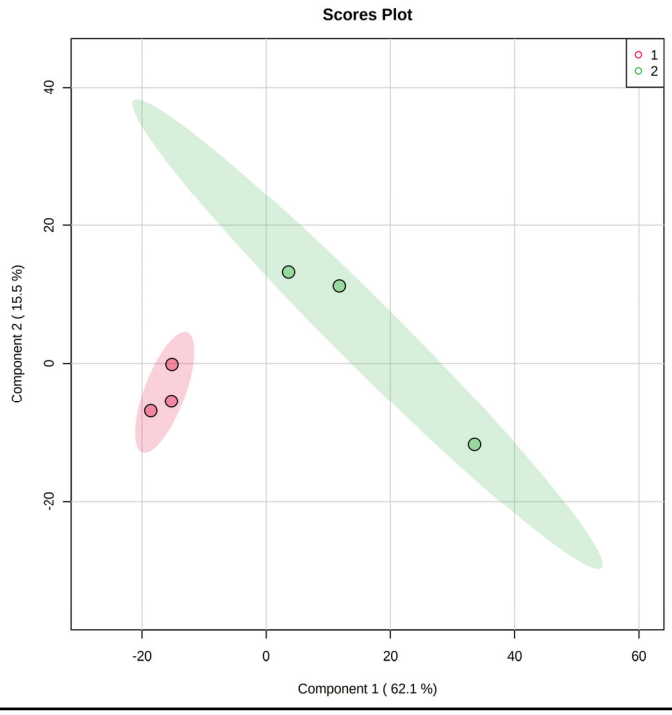


26



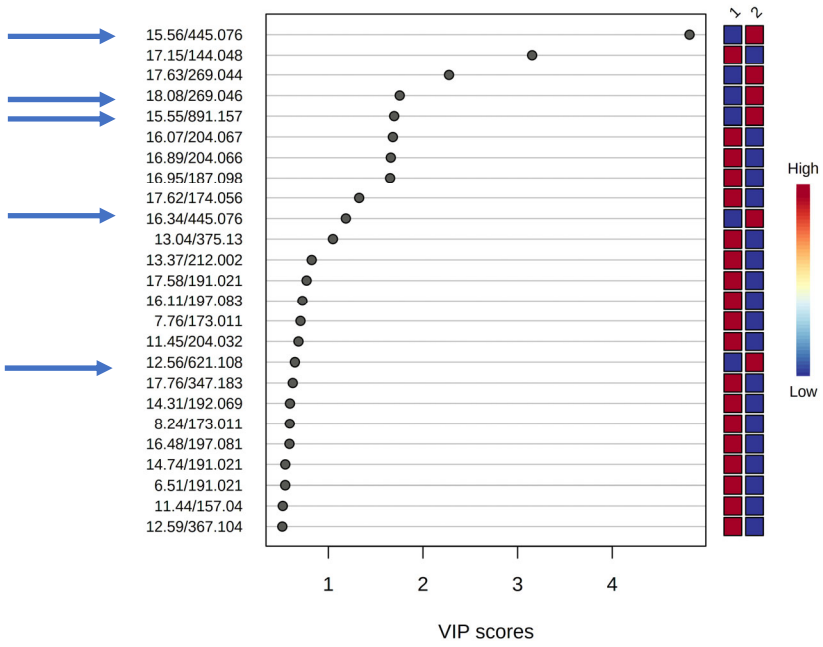
27

2D-PLSDA scores plot



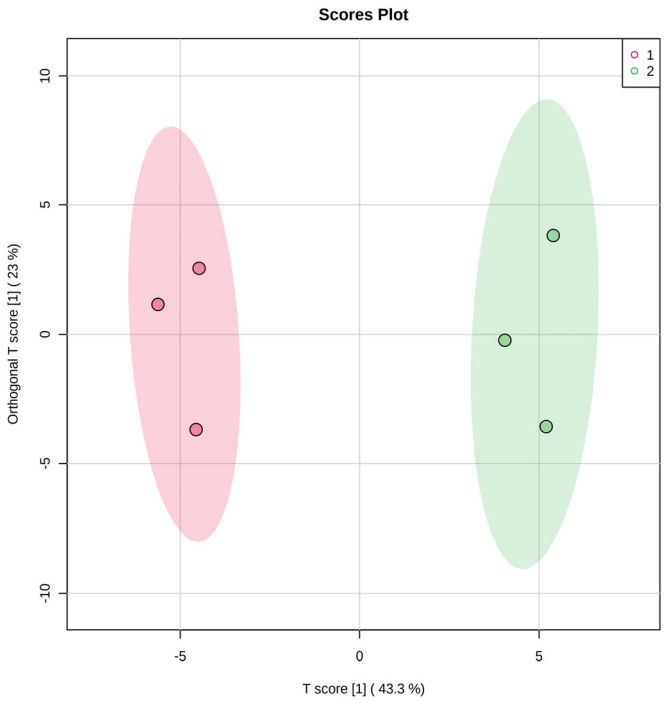
28

PLSDA VIP scores



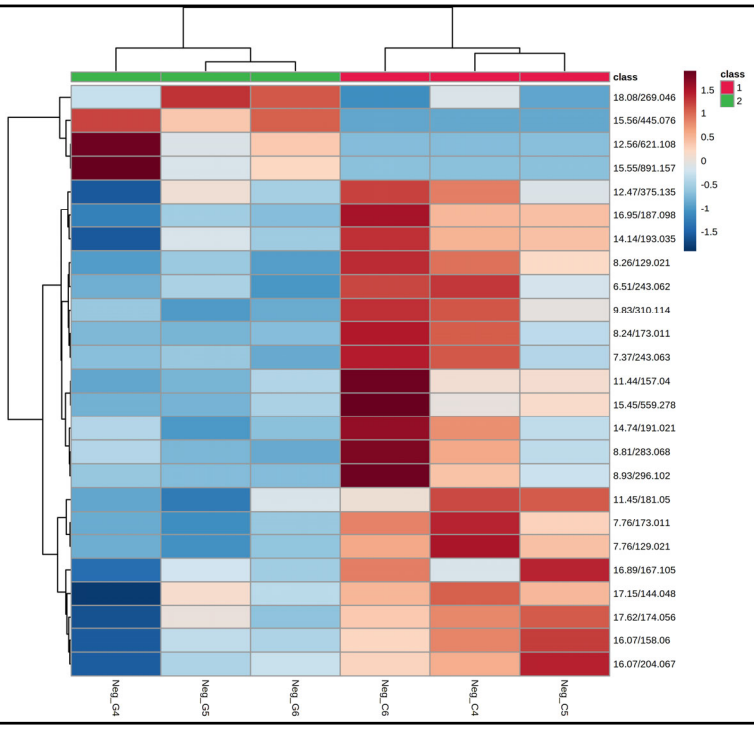
29

oPLSDA plot



30

HeatMap



31

What did we get out of this experiment?

32