THE UNIVERSITY OF ALABAMA AT BIRMINGHAM

Knowledge that will change your world

Working with metabolomics data

Stephen Barnes, PhD

Professor of Pharmacology & Toxicology

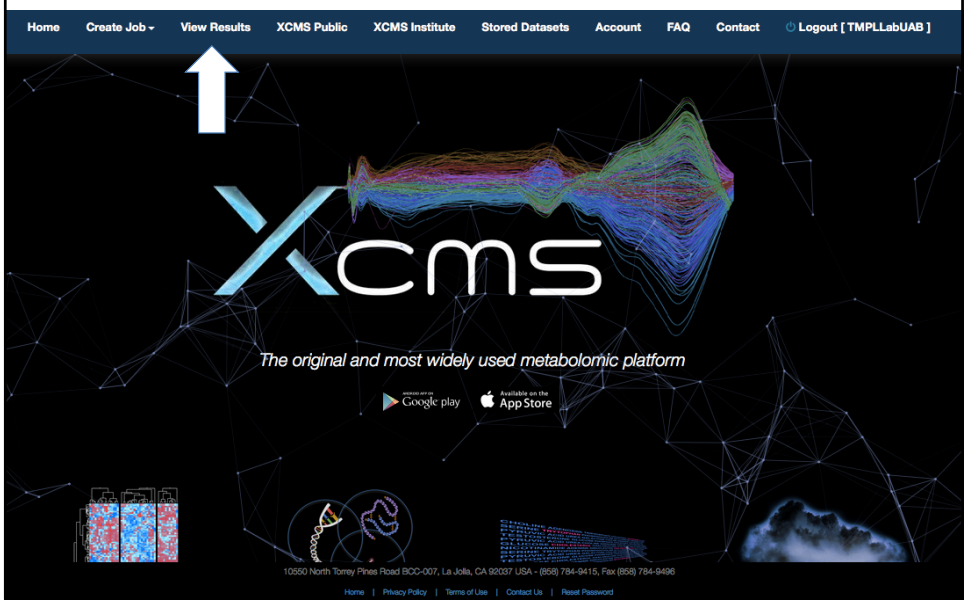
sbarnes@uab.edu

Targeted Metabolomics & Proteomics Laboratory

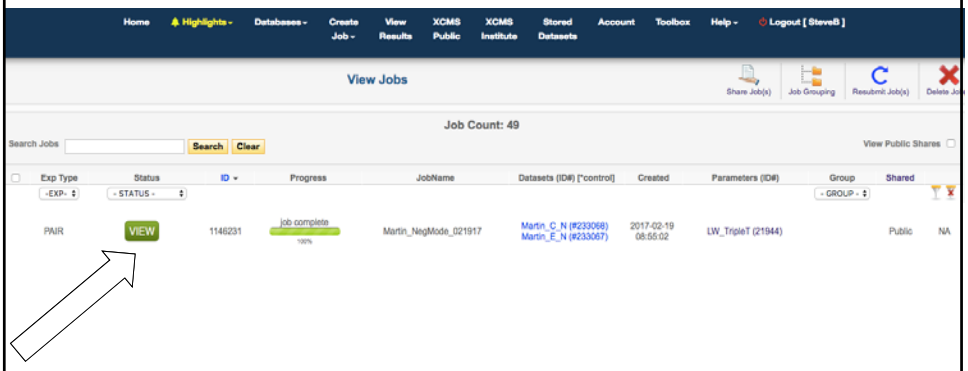
Examining output on XCMS

- We'll logon to XCMS and inspect data from a small study
- Then we'll download the XCMS output file
  - Prepare files for MetaboAnalyst

Logon to XCMS with your name and password

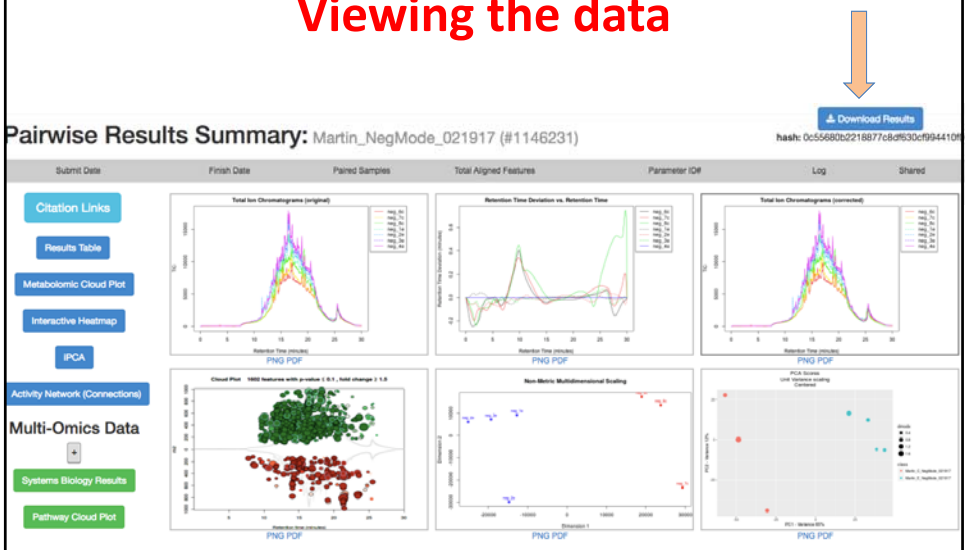


Viewing the datasets



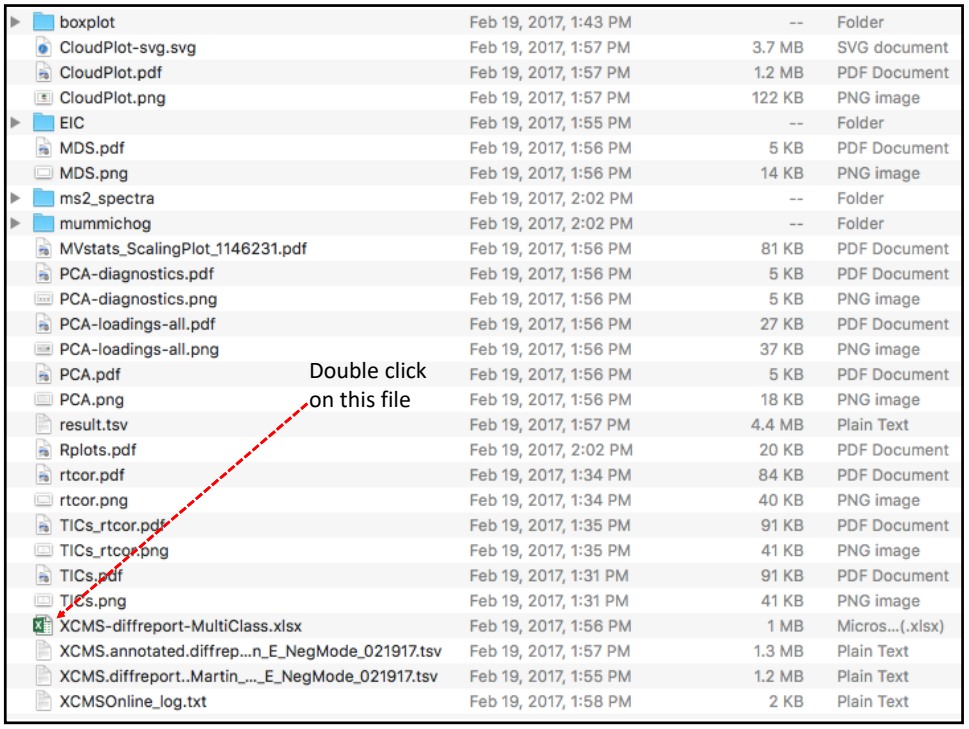
Click on this button

Viewing the data



Next steps

- **Opening the .zip file you've downloaded (or the folder on your thumbdrive)**
- **Open the folder and create the Excel file from the XCMSdiffreport.xlsx file**
- **Preparing for MetaboAnalyst**



boxplot	Feb 19, 2017, 1:43 PM	--	Folder
CloudPlot-svg.svg	Feb 19, 2017, 1:57 PM	3.7 MB	SVG document
CloudPlot.pdf	Feb 19, 2017, 1:57 PM	1.2 MB	PDF Document
CloudPlot.png	Feb 19, 2017, 1:57 PM	122 KB	PNG image
EIC	Feb 19, 2017, 1:55 PM	--	Folder
MDS.pdf	Feb 19, 2017, 1:56 PM	5 KB	PDF Document
MDS.png	Feb 19, 2017, 1:56 PM	14 KB	PNG image
ms2_spectra	Feb 19, 2017, 2:02 PM	--	Folder
mummichog	Feb 19, 2017, 2:02 PM	--	Folder
MVstats_ScalingPlot_1146231.pdf	Feb 19, 2017, 1:56 PM	81 KB	PDF Document
PCA-diagnostics.pdf	Feb 19, 2017, 1:56 PM	5 KB	PDF Document
PCA-diagnostics.png	Feb 19, 2017, 1:56 PM	5 KB	PNG image
PCA-loadings-all.pdf	Feb 19, 2017, 1:56 PM	27 KB	PDF Document
PCA-loadings-all.png	Feb 19, 2017, 1:56 PM	37 KB	PNG image
PCA.pdf	Feb 19, 2017, 1:56 PM	5 KB	PDF Document
PCA.png	Feb 19, 2017, 1:56 PM	18 KB	PNG image
result.tsv	Feb 19, 2017, 1:57 PM	4.4 MB	Plain Text
Rplots.pdf	Feb 19, 2017, 2:02 PM	20 KB	PDF Document
rtcor.pdf	Feb 19, 2017, 1:34 PM	84 KB	PDF Document
rtcor.png	Feb 19, 2017, 1:34 PM	40 KB	PNG image
TICs_rtcor.pdf	Feb 19, 2017, 1:35 PM	91 KB	PDF Document
TICs_rtcor.png	Feb 19, 2017, 1:35 PM	41 KB	PNG image
TICs.pdf	Feb 19, 2017, 1:31 PM	91 KB	PDF Document
TICs.png	Feb 19, 2017, 1:31 PM	41 KB	PNG image
XCMS-diffreport-MultiClass.xlsx	Feb 19, 2017, 1:56 PM	1 MB	Micros...(.xlsx)
XCMS.annotated.diffrep...n_E_NegMode_021917.tsv	Feb 19, 2017, 1:57 PM	1.3 MB	Plain Text
XCMS.diffreport..Martin_..._E_NegMode_021917.tsv	Feb 19, 2017, 1:55 PM	1.2 MB	Plain Text
XCMSOnline_log.txt	Feb 19, 2017, 1:58 PM	2 KB	Plain Text

The Excel DiffReport from XCMS

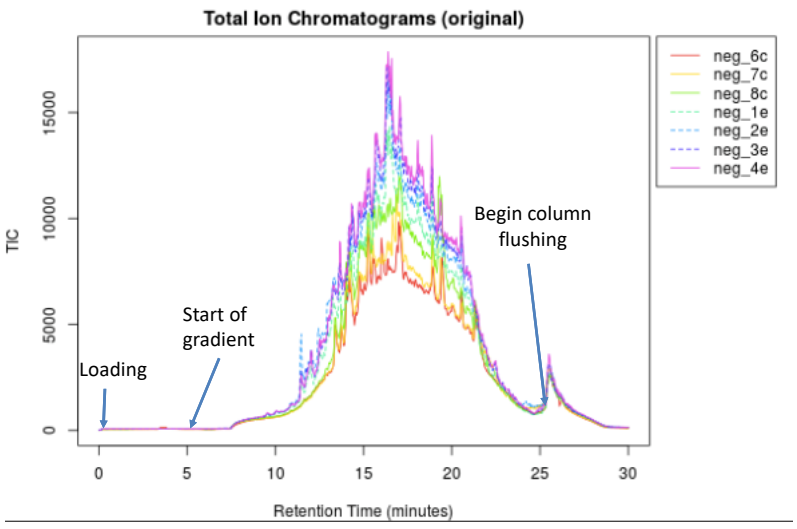
A	B	C	D	E	F	G	H	I	J	K	L	M
name	fold	log2fold	tstat	pvalue	qvalue	updown	mzmed	mzmin	mzmax	rtmed	rtmin	rtmax
M577T14_	6.909129	-2.7885	-88.4435	4.11E-09	2.71E-06	DOWN	577.3835	577.3817	577.3884	14.04068	13.95005	14.05762
M197T14	7.313674	2.870596	38.81764	2.58E-07	6.3E-05	UP	197.0626	197.0618	197.0636	13.64693	13.62275	13.66638
M470T19	5.500509	-2.45957	-36.201	3.03E-07	6.66E-05	DOWN	470.2817	470.2815	470.2878	19.10917	19.09967	19.12683
M408T21	3.854812	-1.94666	-43.3553	5.15E-07	8.49E-05	DOWN	408.2797	408.2789	408.2837	21.3165	21.2685	21.3315
M695T13	4.782005	2.257616	32.42623	9.93E-07	0.000116	UP	695.1871	695.1861	695.1886	12.89197	12.88315	12.91105
M771T13	6.486291	2.697394	28.00855	1.09E-06	0.00012	UP	771.198	771.1969	771.1989	12.91252	12.88432	12.9297
M288T14	4.536327	-2.18152	-25.8751	2.04E-06	0.000153	DOWN	288.1874	288.1859	288.1893	14.04068	13.95005	14.05762
M357T20	6.23272	2.639862	25.70254	2.53E-06	0.000163	UP	357.1324	357.13	357.133	20.2855	20.278	20.32267
M445T20	3.172539	1.665638	22.76875	3.04E-06	0.000171	UP	445.1471	445.1445	445.152	20.2855	20.278	20.32267
M229T15	4.152587	-2.05401	-22.609	3.17E-06	0.000173	DOWN	229.1106	229.1104	229.1107	15.08253	15.04087	15.10412
M364T17	4.9317	2.302085	25.75596	3.25E-06	0.000174	UP	364.0447	364.0404	364.049	16.67074	16.66432	16.67717
M597T14_	8.57949	-3.10089	-28.5356	3.75E-06	0.000179	DOWN	597.363	597.3587	597.367	14.07135	13.98287	14.08795
M167T10	3.284931	1.715863	25.24395	4.33E-06	0.000185	UP	167.0021	167.0018	167.0025	10.04034	10.02168	10.059
M545T26	5.004307	-2.32317	-31.686	4.62E-06	0.000187	DOWN	545.3093	545.3089	545.3186	25.515	25.33967	25.57733
M679T16_	6.050417	2.597034	28.29277	4.8E-06	0.000188	UP	679.1936	679.1889	679.1984	16.34269	16.31365	16.36915
M990T21	4.078085	-2.02789	-59.403	5.19E-06	0.00019	DOWN	989.5686	989.5668	989.5704	21.48192	21.358	21.60583
M575T14	3.706953	-1.89023	-21.1545	5.69E-06	0.000193	DOWN	575.3561	575.3526	575.3596	13.99537	13.95005	14.04068
M532T12_	3.396063	1.763863	43.76261	6.07E-06	0.000195	UP	532.0504	532.048	532.0528	11.87441	11.86025	11.88857
M434T12	8.745729	3.128579	57.31898	6.08E-06	0.000195	UP	434.1217	434.1192	434.1249	11.85023	11.83317	11.86993
M287T14_	4.760581	-2.25114	-21.3931	6.21E-06	0.000195	DOWN	287.2837	287.2781	287.2894	14.00323	13.91852	14.08795
M187T16	3.544341	-1.82552	-23.8039	6.23E-06	0.000195	DOWN	187.0077	187.006	187.0077	15.50122	15.48107	15.54597
M287T14_	7.297991	-2.8675	-33.5488	6.87E-06	0.000201	DOWN	287.1846	287.1842	287.1872	14.04068	13.95005	14.05762
M576T14	2.604531	-1.38102	-19.8191	7.07E-06	0.000202	DOWN	576.3582	576.3495	576.3585	14.04068	13.95005	14.05762
M993T21	3.671652	-1.87643	-20.2252	7.94E-06	0.000213	DOWN	992.5787	992.5766	992.5808	21.36575	21.32883	21.40267
M407T21_	3.451609	-1.78727	-39.727	8.61E-06	0.00022	DOWN	407.2746	407.2746	407.2801	21.3165	21.2685	21.3315

This file has 2,944 lines – we need to apply filters

Preparing for analysis using MetaboAnalyst

http://www.metaboanalyst.ca

The analytical method



Make a copy of the sheet onto a new sheet and sort the data by retention time

	A	B	C	D	E	F	G	H	I	J	K	L	M
	name	fold	log2fold	tstat	pvalue	qvalue	updown	mzmed	mzmin	mzmax	rtmed	rtmin	rtmax
	M151T8	6.18910801	-2.6297315	-15.564112	0.00158982	0.00222516	DOWN	151.025789	151.025546	151.025898	7.76826667	7.74256667	7.79691667
	M125T8	2.65382915	-1.4080755	-6.6658732	0.00251771	0.00271494	DOWN	125.037215	125.036146	125.037785	8.43006667	8.36876667	8.62013333
	M251T8	9.22049744	-3.2048446	-10.334004	0.00895064	0.00541785	DOWN	251.076623	251.076055	251.076797	8.48723333	8.47791667	8.54845
	M267T9	15.8925741	-3.9902809	-36.581905	0.00027564	0.00105245	DOWN	267.074064	267.071506	267.075305	8.64074167	8.63426667	9.13695
	M164T9	2.44780886	-1.2914909	-2.8660081	0.03818147	0.01638319	DOWN	164.072332	164.069495	164.072523	9.39105	9.3395	9.77725
	M178T10	3.51748667	1.81454496	6.93218141	0.00431152	0.00353059	UP	178.052028	178.050841	178.052429	9.6206	9.00206667	9.6693
	M218T10	2.83372396	-1.5026992	-5.1640395	0.03314535	0.01471798	DOWN	218.096668	218.095458	218.097878	9.65415	9.62046667	9.68783333
	M222T10	13.4384764	3.74829768	8.21595805	0.00301695	0.00294217	UP	222.04142	222.041131	222.041887	9.65736667	9.15706667	9.67505
	M408T10	3.78104823	1.91878625	7.32156689	0.00502445	0.00383041	UP	407.992363	407.991374	407.993201	9.65740833	9.64666667	9.67505
	M463T10	5.0317721	2.33106658	7.3283287	0.005237	0.00392159	UP	462.949392	462.94859	462.949691	9.65740833	9.6206	9.67505
	M358T10	3.70214918	1.88836303	6.8948228	0.00586231	0.00418629	UP	358.009146	358.003173	358.01122	9.65740833	9.64666667	9.6693
	M352T10	2.50289154	1.32359578	5.97677354	0.00617498	0.00431853	UP	351.98303	351.97767	351.983726	9.65745	9.65736667	9.6693
	M160T10	10.5606868	3.40063175	9.03421286	0.00248494	0.00269921	UP	160.04163	160.040415	160.042225	9.6644	9.14763333	9.70156667

Keep metabolites eluting between 5.00 and 25.00 minutes

In this example, because of the use of ethyl acetate to extract fecal water, the most hydrophilic metabolites were eliminated and there are no metabolites eluting before 7.7 min

Deleting the data after 25 min

	A	B	C	D	E	F	G	H	I	J	K	L	M
	M616T25	1.74066121	-0.7996354	-3.4556534	0.04632104	0.01904429	DOWN	615.53544	615.535198	615.536317	24.75225	24.7016667	24.8288333
	M617T25	1.70327884	-0.7683146	-2.8883319	0.08728587	0.03074642	DOWN	616.538353	616.534299	616.538873	24.772	24.7216667	24.8288333
	M663T25	1.02330246	0.03323264	0.14826556	0.8914236	0.20442618	UP	662.522891	662.521371	662.523744	24.78575	24.6803333	24.8823333
	M256T25	1.34184796	0.42422121	1.37006888	0.27817946	0.07800847	UP	256.236287	256.234265	256.236744	24.7903333	24.6965	24.823
	M511T25	1.71488977	0.77811585	2.28123449	0.08993258	0.03147224	UP	511.463499	511.457028	511.46997	24.7935	24.7581667	24.8288333
	M662T25	1.06540888	-0.0914072	-0.2503189	0.82206378	0.19156727	DOWN	661.52141	661.520415	661.522412	24.8065	24.7238333	24.8558333
	M255T25	1.25673233	0.3296774	0.98894228	0.39652352	0.10479348	UP	255.232886	255.231687	255.234695	24.8126667	24.7238333	24.8296667
	M630T25	2.70963468	-1.4380984	-1.1860642	0.35613606	0.09600543	DOWN	629.516223	629.515237	629.524589	24.8133333	24.7513333	24.8648333
	M445T25	1.42472325	-0.5106817	-4.0220301	0.01746489	0.00894294	DOWN	444.971171	444.970438	444.976428	24.8528333	24.8065	24.954
	M381T25	1.54774037	-0.6301635	-4.0345626	0.03275954	0.01458561	DOWN	380.975952	380.974802	380.979722	24.8591667	24.8133333	24.9616667
	M549T25_2	6.66652118	2.73693411	3.6083927	0.02756	0.01272993	UP	549.247547	549.247189	549.247906	24.8945833	24.868	24.9211667
	M660T25	4.0842453	-2.0300695	-1.6292071	0.24402551	0.0699418	DOWN	659.565856	659.561129	659.570583	24.9065	24.8141667	24.9988333
	M661T25	4.79371946	-2.2611455	-1.6713356	0.23519904	0.06785819	DOWN	660.569125	660.564609	660.573641	24.9215	24.8591667	24.9838333
	M249T25	1.30097666	0.37959508	1.86418912	0.17901878	0.05392111	UP	248.960443	248.959285	248.962145	24.946	24.7845	25.0308333
	M528T25_2	6.4094141	2.68019248	4.29406879	0.02027959	0.00998888	UP	528.269414	528.268836	528.269715	24.94825	24.868	25.0308333
	M527T25	6.85313497	2.7767641	4.06515266	0.02364163	0.01124932	UP	527.266321	527.265914	527.266706	24.94825	24.868	25.0308333
	M392T25	2.04351954	1.03105604	2.54305433	0.05827333	0.0227366	UP	392.000088	391.99581	392.004326	24.9883333	24.946	25.0425
	M391T25	1.85301432	0.88987403	1.91633661	0.1158983	0.03812132	UP	390.999087	390.993375	390.999758	25.0018333	24.8973333	25.0425
	M405T25	1.47026819	-0.5560793	-1.6996074	0.15905058	0.04897716	DOWN	405.302631	405.301619	405.30729	25.1195	25.0423333	25.2465
	M513T25	1.5658062	0.6469057	1.1925939	0.35377674	0.09548116	DOWN	512.939487	512.929988	512.941509	25.188	25.1616667	25.3361667
	M406T25	2.01392861	-1.0100125	-3.0101833	0.08591016	0.03036538	DOWN	406.305702	406.304516	406.310196	25.2068333	25.1376667	25.2585
	M671T25	1.01785577	0.02553315	0.13242075	0.90281782	0.20649958	UP	671.47149	671.470083	671.474565	25.2209167	25.0753333	25.291
	M672T25	1.0057711	0.00830201	0.03882298	0.97226751	0.2189073	UP	672.47549	672.474947	672.477408	25.2525	25.0813333	25.291
	M703T25	1.09318544	0.12853815	0.64811943	0.54657342	0.13665885	UP	703.461184	703.459721	703.466371	25.2711667	25.233	25.3361667
	M718T25	1.02819514	0.0401141	0.04734225	0.96519357	0.21766127	UP	718.467716	718.466446	718.468775	25.2711667	25.135	25.3183333
	M717T25	1.01161096	-0.0166546	-0.0191981	0.98597389	0.22131035	DOWN	717.46644	717.46572	717.467564	25.2771667	25.1616667	25.3183333

Delete the lines in blue (and those below them in the file)

Now order them according to peak areas

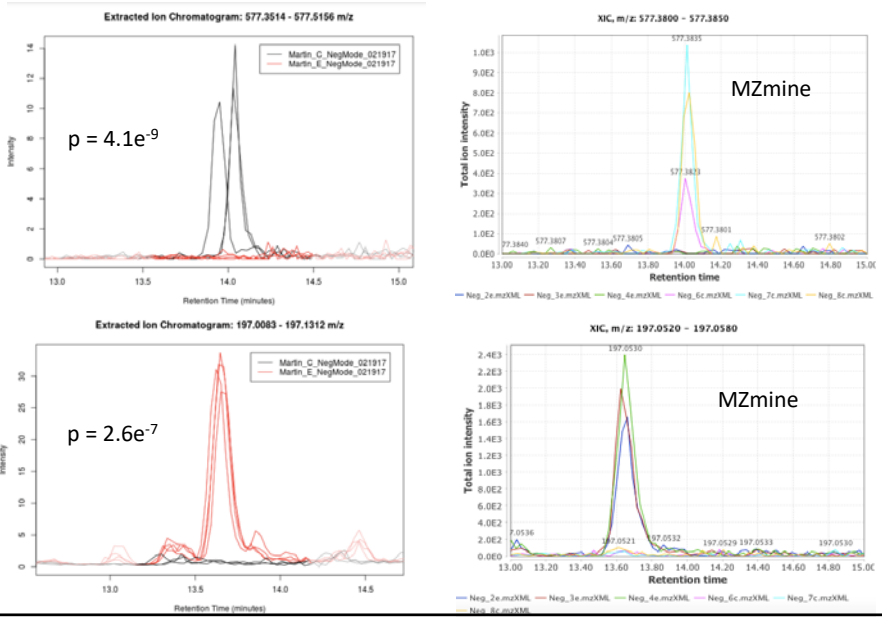
- The goal here is to remove low intensity noisy peaks that contribute little to the overall separation of groups
- Copy the data to a new sheet and order them according to area – use mean 1

Q	R	S	T	U	V	W	X	Y	Z	AA	AB
naxint	mean1	sd1	mean2	sd2	neg_6c	neg_7c	neg_8c	neg_1e	neg_2e	neg_3e	neg_4e
315.359002	15329.1064	3508.69228	19278.9409	978.433407	12315.9496	14490.1388	19181.2308	18238.362	18650.9845	20075.4118	20151.0052
409.926623	14256.9176	3242.63705	17848.0004	870.784185	11427.6214	13547.6523	17795.4789	17275.9454	16941.7691	18716.8129	18457.474
201.433567	10173.8633	2394.50668	12743.6498	640.949504	8181.44133	9509.84732	12830.3011	12052.7061	12384.0435	13076.6991	13461.1504
428.071851	9055.13032	2382.11478	12453.7192	598.449918	7091.01905	8369.50161	11704.8703	11766.1142	12139.8574	12925.7361	12983.1689
499.720104	8822.58615	2382.38491	12746.2851	1011.90362	6742.22552	8303.97442	11421.5585	11908.8996	11845.4367	13770.9273	13459.8767
459.301497	6885.79443	384.311154	330.418743	28.5214215	6520.27575	7286.48067	6850.62687	288.140326	340.949033	341.928682	350.656931
616.169065	6615.67269	131.710042	67.295694	15.9862306	6494.93524	6756.13128	6595.95155	87.1719116	49.6531781	60.8001073	71.5575791
144.768658	6436.40042	1490.47834	8373.68136	392.258809	5283.28233	5906.49452	8119.42442	7919.71783	8259.6772	8456.13676	8859.19366
932.891307	6255.6509	227.872978	857.174388	185.307732	6061.17072	6506.38102	6199.40097	609.563677	995.948822	820.694806	1002.49025
256.140414	6092.36299	1434.56201	8414.56677	607.675978	4956.85316	5615.64137	7704.59444	7779.37807	8012.49927	8981.8627	8884.52704
1344.17536	5693.53013	9659.13669	86.5096666	2.12898827	128.569211	105.088944	16846.9322	88.4501072	88.0919415	83.9986066	85.498011
596.850227	5673.4855	866.812603	4207.19046	280.270228	4750.74909	5799.02306	6470.68436	3910.68437	4543.68309	4317.02706	4057.36733
588.699204	5524.58591	1430.87299	8203.23373	460.577116	4346.77312	5110.01572	7116.9689	7734.1042	7956.46134	8783.84692	8338.52246
210.65354	5392.39078	413.690742	5118.19253	413.986812	5009.95295	5335.72662	5831.49275	4898.17695	4660.64364	5572.84233	5341.1072
707.769812	5360.63269	531.861851	675.897883	181.544359	4783.01895	5468.74976	5830.12937	554.223593	946.131526	599.465599	603.770815
385.747382	5330.12241	710.419913	4007.87707	265.048333	4659.8042	5255.7677	6074.79534	3665.40921	3939.49704	4166.42998	4260.17207

Rationale for data clean up

- It's a good idea to inspect the data with mzmine or xcms to see where the peaks become ragged
- It's important to write down where the cutoff is for a given analysis
- If there are >2,000 features left, then I artificially set a 2,000 features limit
  - Why? Metaboanalyst restricts the number of features it uses to 2,000 or less
  - There are other stats programs that use bigger matrices

Data inspection – top two most significant



Editing to 2000 features

	Q	R	S	T	U	V	W	X	Y	Z	AA	AB
1	maxint	mean1	sd1	mean2	sd2	neg_6c	neg_7c	neg_8c	neg_1e	neg_2e	neg_3e	neg_4e
1989	2.61917341	33.5978788	1.48863178	30.280362	2.19600726	32.0134493	33.8128484	34.9673387	32.1293066	27.2629341	31.6587068	30.0705004
1990	69.1746301	33.5590864	1.22897094	572.669752	103.581142	33.3442481	34.8813113	32.4516999	528.407324	715.308379	573.697587	473.265719
1991	3.0183045	33.4899067	15.9000329	87.6331531	14.8674346	51.8179468	25.2603686	23.3914048	73.899086	76.1213205	96.8430963	103.669109
1992	1.52320965	33.3293358	5.06559254	29.088342	15.580468	37.0130602	35.4223	27.5526472	24.1842869	18.5078723	21.4622669	52.198942
1993	2.96657214	33.2952166	3.10460897	36.2209871	4.92204677	29.8021413	35.7399057	34.3436028	35.0626474	29.7467025	39.6545159	40.4200824
1994	130.011142	33.082448	2.60263108	978.330064	276.734498	32.0529788	31.1520172	36.0423479	732.28634	758.924587	1290.8926	1131.21673
1995	13.9866844	33.0743694	8.2087205	67.690226	32.2561792	26.0087161	31.1354047	42.0789873	50.4878507	46.0073122	115.410777	58.8549644
1996	3.50070544	33.052158	7.08025191	4.5614519	1.73162222	36.4457874	37.7968047	24.9138817	6.9627407	4.13145831	4.31791678	2.83369184
1997	43.2464989	33.0382852	0.50676523	496.753269	53.8125057	33.3444084	32.4533354	33.3171119	484.471518	439.517641	569.223136	493.800781
1998	3.24932925	32.97894	11.2029773	58.6124157	11.6628323	45.2573068	30.366509	23.3130043	59.9965283	44.574603	72.9714184	56.9071129
1999	12.0030845	32.9373465	5.75257267	139.614936	12.491432	27.8650391	31.7592215	39.1877788	129.121977	151.846853	128.575067	148.915846
2000	1.93289122	32.7690126	7.41359532	38.0260091	4.64780409	25.911992	40.6356922	31.7593535	37.5117352	31.7817041	42.5696077	40.2409893
2001	9.45620123	32.7560844	3.81249821	77.6999853	15.7211751	28.5606933	33.6987203	36.0088396	99.6203179	65.7403202	78.5644794	66.8748237
2002	3.51502901	32.729768	1.65823438	39.9822181	2.29922087	31.1026122	32.6692668	34.417425	41.0919997	37.9804628	42.6890456	38.1673641
2003	18.667884	32.7191919	13.0800268	209.71085	81.0399932	23.5116621	26.9545844	47.6913293	134.410226	145.019824	284.630981	274.782369
2004	7.96139122	32.6354196	24.6833137	37.4611607	18.7255178	5.31863585	53.337317	39.2503059	29.7529661	31.6755536	65.0407785	23.3753445
2005	7.58793118	32.6295644	8.36223267	133.540774	17.840258	23.1856837	35.6088708	39.0941387	112.492234	134.676299	131.002684	155.991878
2006	2.29114842	32.5669083	14.8205305	24.269133	3.33705679	30.1046046	19.131741	48.4643794	21.6118007	26.009346	21.34409	28.1112954

Delete features below #2000 in the Excel file

Color code the sample groups

V	W	X	Y	Z	AA	AB
neg_6c	neg_7c	neg_8c	neg_1e	neg_2e	neg_3e	neg_4e
12315.9496	14490.1388	19181.2308	18238.362	18650.9845	20075.4118	20151.0052
11427.6214	13547.6523	17795.4789	17275.9454	16941.7691	18716.8129	18457.474
8181.44133	9509.84732	12830.3011	12052.7061	12384.0435	13076.6991	13461.1504
7091.01905	8369.50161	11704.8703	11766.1142	12139.8574	12925.7361	12983.1689
6742.22552	8303.97442	11421.5585	11908.8996	11845.4367	13770.9273	13459.8767
6520.27575	7286.48067	6850.62687	288.140326	340.949033	341.928682	350.656931
6494.93524	6756.13128	6595.95155	87.1719116	49.6531781	60.8001073	71.5575791
5283.28233	5906.49452	8119.42442	7919.71783	8259.6772	8456.13676	8859.19366
6061.17072	6506.38102	6199.40097	609.563677	995.948822	820.694806	1002.49025
4956.85316	5615.64137	7704.59444	7779.37807	8012.49927	8981.8627	8884.52704
128.569211	105.088944	16846.9322	88.4501072	88.0919415	83.9986066	85.498011
4750.74909	5799.02306	6470.68436	3910.68437	4543.68309	4317.02706	4057.36733
4346.77312	5110.01572	7116.9689	7734.1042	7956.46134	8783.84692	8338.52246
5009.95295	5335.72662	5831.49275	4898.17695	4660.64364	5572.84233	5341.1072
4783.01895	5468.74976	5830.12937	554.223593	946.131526	599.465599	603.770815
4659.8042	5255.7677	6074.79534	3665.40921	3939.49704	4166.42998	4260.17207
1578.85957	12555.4679	1024.96128	1063.54487	5006.70661	1188.14954	1170.94225
4433.02674	4823.36962	5865.32669	6679.6988	7734.74039	7871.59246	7484.43816
5009.01449	4922.87369	4756.25011	1299.67751	1394.45544	1512.37663	1467.41944
4001.06004	4821.06233	5661.30303	507.907159	584.175095	600.471771	673.264794
3839.26431	4481.58449	5571.47329	5197.48765	5185.33452	5532.43879	5484.90238

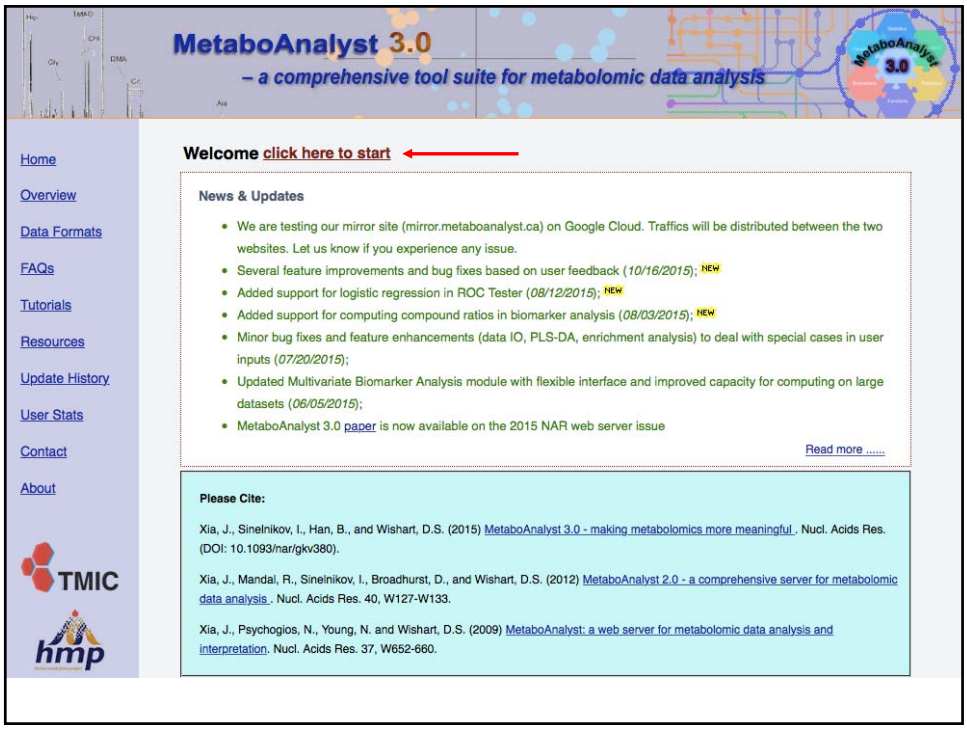
Creating .csv files for each sample

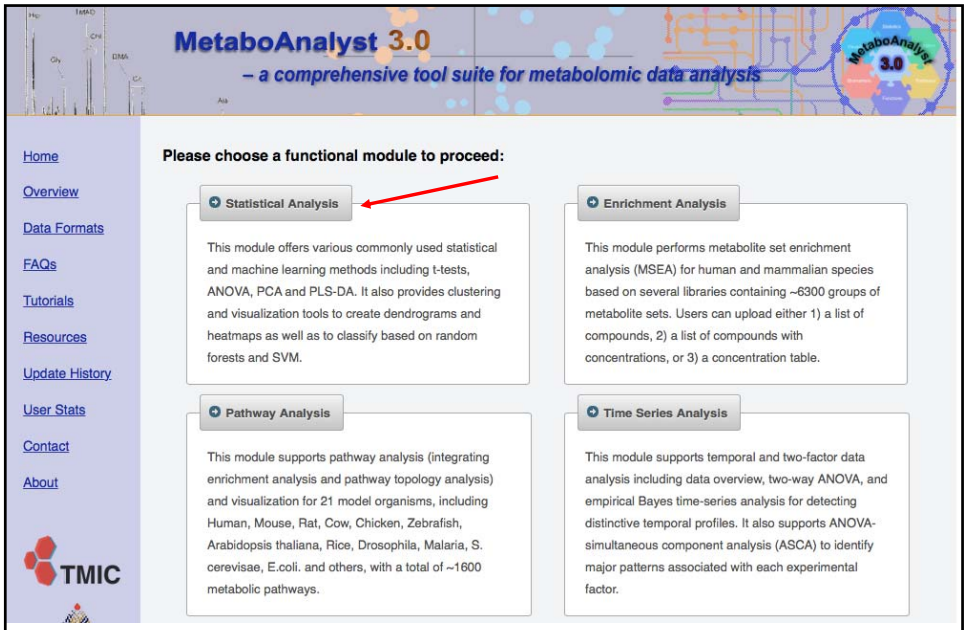
mzmed	rtmed	neg_6c
407.191392	16.7261667	12315.9496
333.155409	15.9005	11427.6214
481.227872	17.3965	8181.44133
419.191434	16.4857833	7091.01905
345.155342	15.6982167	6742.22552
329.102953	16.8775	6520.27575
453.284388	21.3526667	6494.93524
555.264314	17.9361667	5283.28233
287.184577	14.0406833	6061.17072
493.227675	17.1411667	4956.85316
236.097277	19.3865	128.569211
241.109806	14.97655	4750.74909
317.130863	14.1771	4346.77312
305.138376	15.1711	5009.95295
243.124987	14.1763917	4783.01895
315.144717	16.039875	4659.8042
311.168872	22.0998333	1578.85957
389.180407	17.0568333	4433.02674
407.274637	21.3165	5009.01449
301.151733	15.5167167	4001.06004

- **Copy the median *m/z* and median Rt values into a new Excel file. Then copy the column of areas from the first sample in Group_1. Save as an Excel .csv file.**
  - **Note that the file name must not have spaces – use an underscore instead of a space.**
- **Leave the file open and replace the yellow column with the areas from the next Group_1 sample. Save as a second .csv file.**
- **Continue until all Group_1 and Group_2 samples have a corresponding .csv file.**

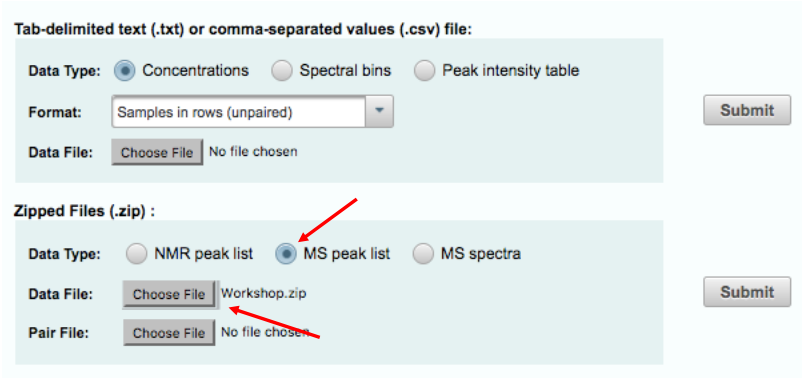
Preparing a .zip file

- Put each of the .csv files for group_1 samples into a folder named “Group_1”.
- Put each of the .csv files for group_2 samples into a folder named “Group_2”.
- Click on Group_1 and Group_2 folders and combine to form a .zip file.
  - Rename the .zip file as [your_name].zip
- You’re now ready to submit it to MetaboAnalyst
  - http://www.metaboanalyst.ca

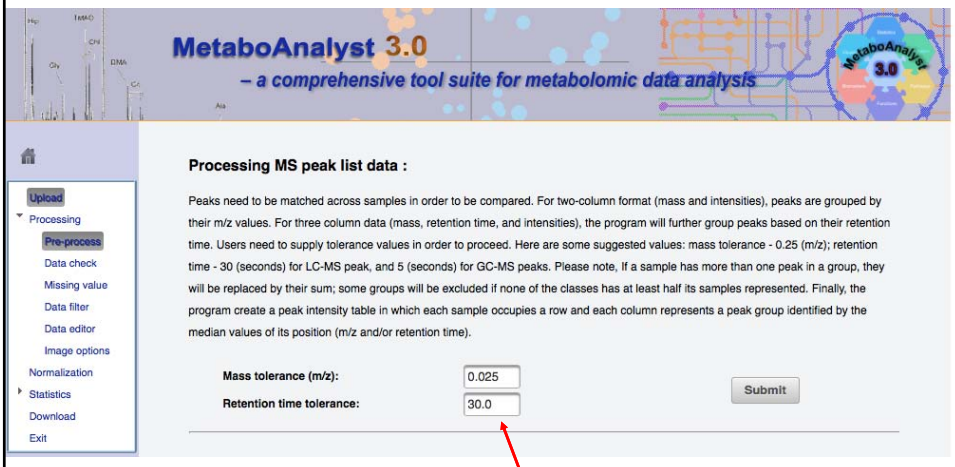




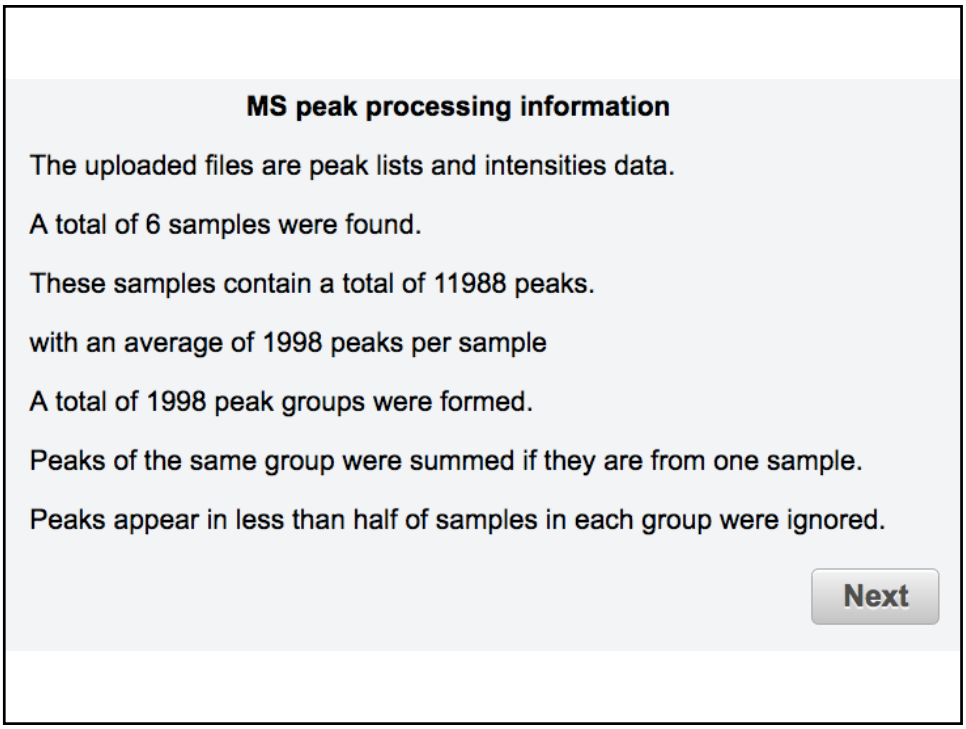
Uploading data to Metaboanalyst

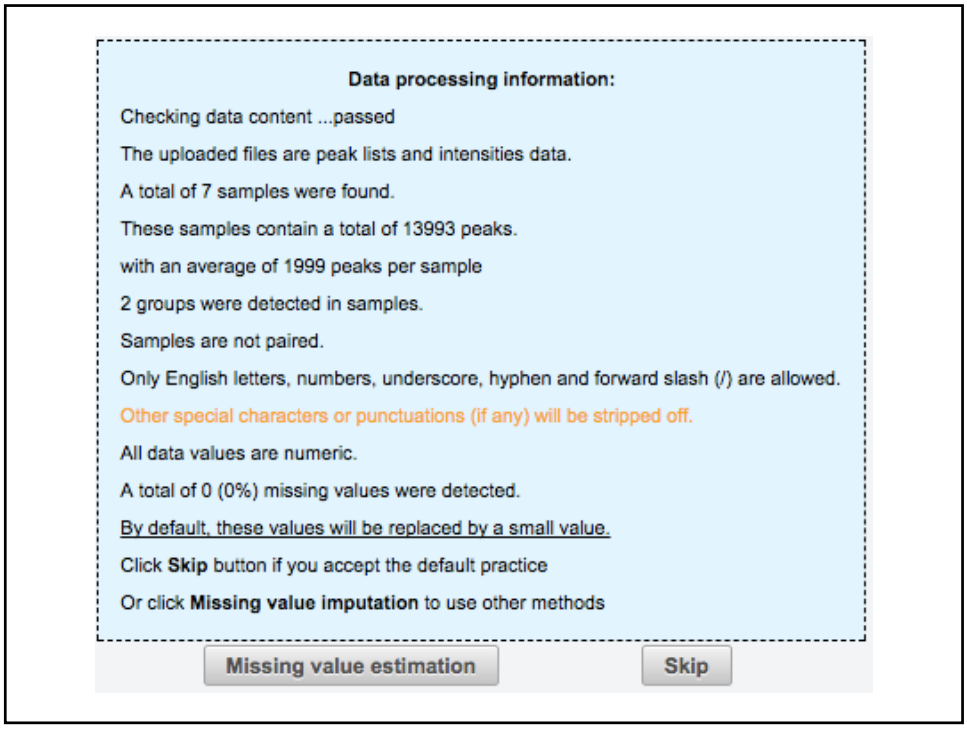


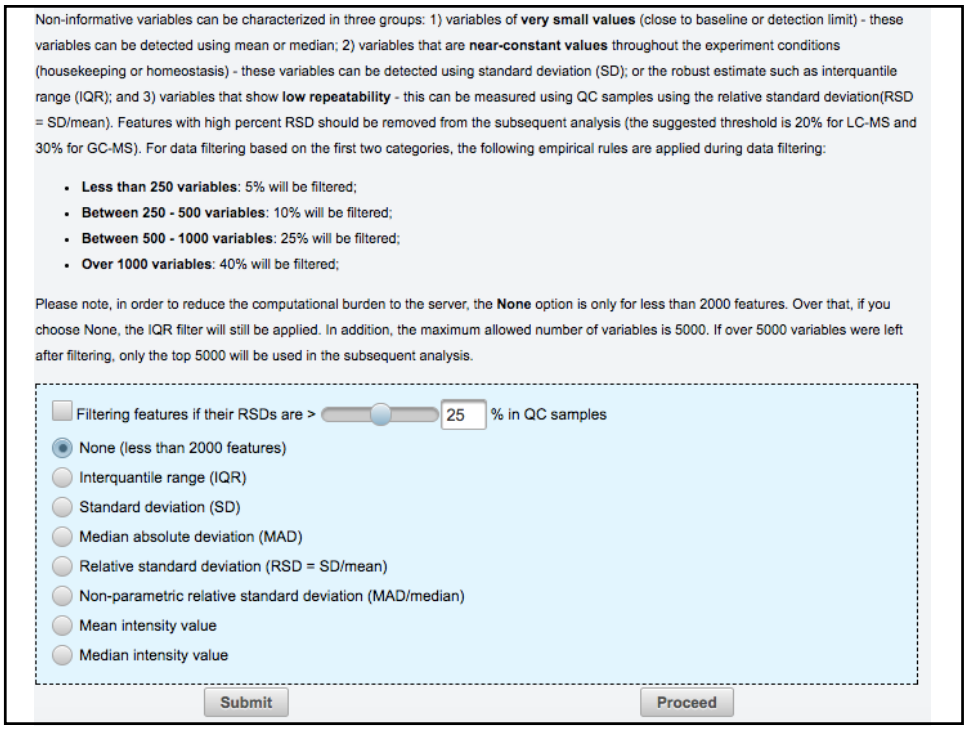
Select MS peak list option and then load the .zip file



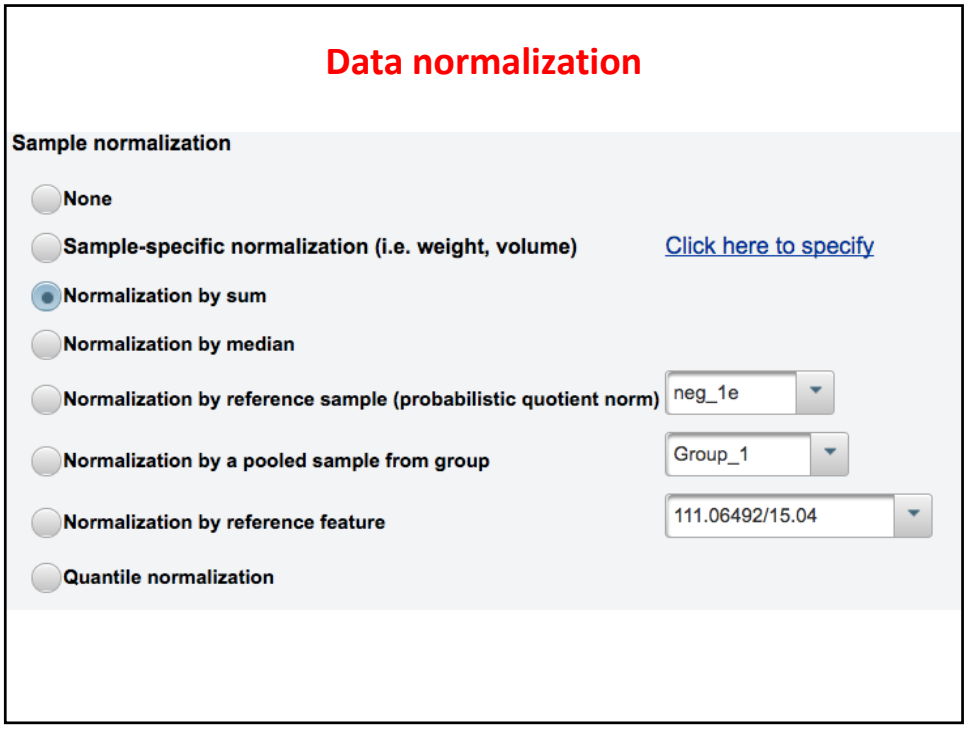
reset these to 0.001 and 0.01, respectively



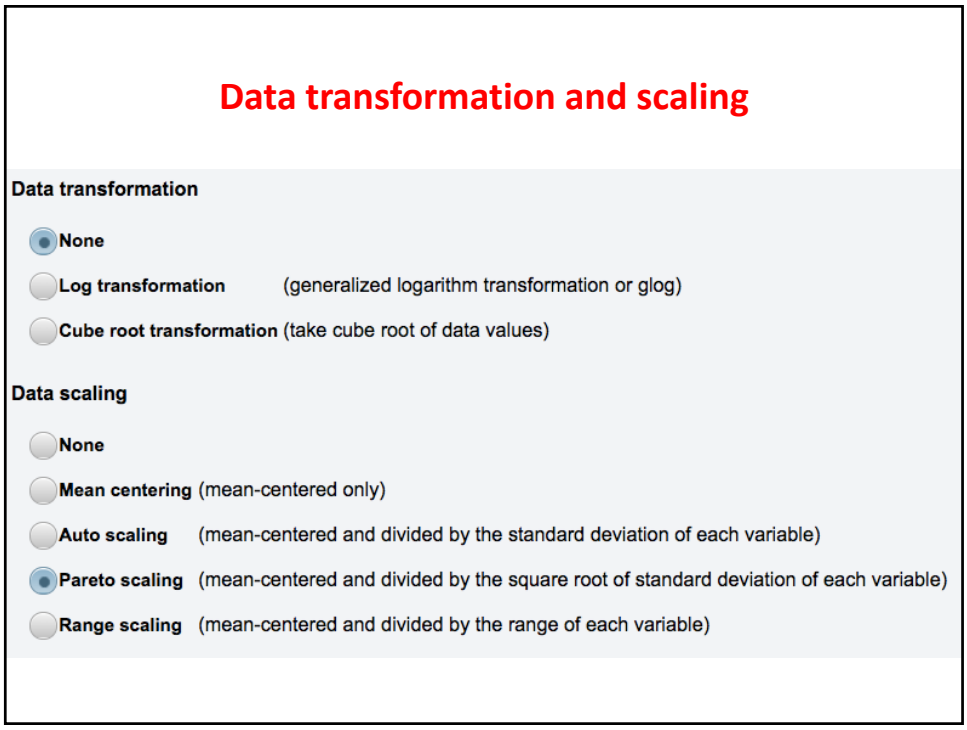

Data processing information:
Checking data content ...passed
The uploaded files are peak lists and intensities data.
A total of 7 samples were found.
These samples contain a total of 13993 peaks.
with an average of 1999 peaks per sample
2 groups were detected in samples.
Samples are not paired.
Only 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, these values will be replaced by a small value.
Click Skip button if you accept the default practice
Or click Missing value imputation to use other methods
Missing value estimation
Skip

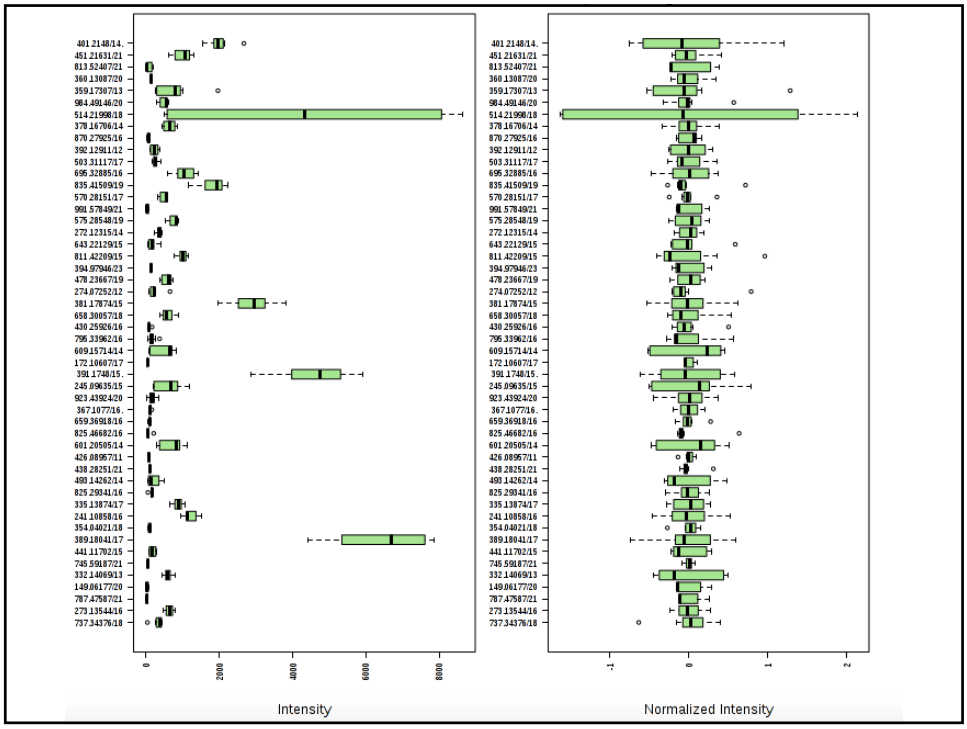

Non-informative variables can be characterized in three groups: 1) variables of very small values (close to baseline or detection limit) - these variables can be detected using mean or median; 2) variables that are near-constant values throughout the experiment conditions (housekeeping or homeostasis) - these variables can be detected using standard deviation (SD); or the robust estimate such as interquantile range (IQR); and 3) variables that show low repeatability - this can be measured using QC samples using the relative standard deviation(RSD = SD/mean). Features with high percent RSD should be removed from the subsequent analysis (the suggested threshold is 20% for LC-MS and 30% for GC-MS). For data filtering based on the first two categories, the following empirical rules are applied during data filtering:
Less than 250 variables: 5% will be filtered;
Between 250 - 500 variables: 10% will be filtered;
Between 500 - 1000 variables: 25% will be filtered;
Over 1000 variables: 40% will be filtered;
Please note, in order to reduce the computational burden to the server, the None option is only for less than 2000 features. Over that, if you choose None, the IQR filter will still be applied. In addition, the maximum allowed number of variables is 5000. If over 5000 variables were left after filtering, only the top 5000 will be used in the subsequent analysis.
Filtering features if their RSDs are > 25 % in QC samples
None (less than 2000 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

Data normalization

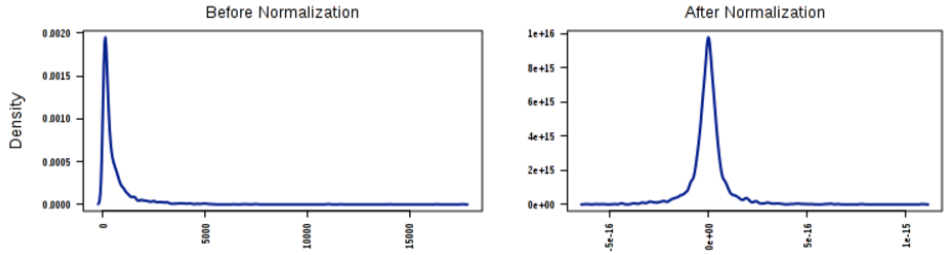


Data transformation and scaling





Effect of normalization and scaling



You're now ready to use the statistical packages of Metaboanalyst