xMWAS

Courtesy of Karan Uppal and Xin Hu

1

Setting up the files

- **Get the normalized file output from Metaboanalyst**
- **Get the files for the pre-study comparing plasma and serum**
- **Convert the ion features back to separate *m/z* and RT columns**
  - Delete the RT column
  - Delete the group line
  - Check for duplicate *m/z* values – delete one (or more) of them
- **Separate the plasma and serum data into two .csv files**
  - Save these files
- **Create a class.csv file**
- **Submit the files to xMWAS**

2

This is how the negative plasma file should look

MZ	Neg_P1	Neg_P2	Neg_P3	Neg_P4	Neg_P5
73.029	-0.400887	0.8170159	-0.5495578	0.2553124	0.3364961
103.039	-1.1159636	2.1335149	-1.4589931	-0.4419286	1.2648773
109.028	-0.4765861	-0.7072879	-0.5862482	2.3215912	-0.4931386
116.072	2.1523037	1.2043164	-0.5918394	3.8922807	0.1175659
117.056	-2.9303821	1.9029645	-0.6293591	-1.1170443	4.6200127
121.029	0.2098905	-0.3829778	-0.3628214	-0.1328817	0.7434454
123.046	-0.005414	-0.4704006	-0.3517024	0.46476488	0.47530019
128.035	0.18609276	0.99519313	1.4527391	0.20830901	0.95441109
129.056	0.19110818	0.13107439	-0.1873413	0.33411859	-0.1015534
130.087	3.7349043	-0.9810414	-3.77168	3.0106388	-0.7349647
130.088	3.1297478	-0.7433436	-2.4126061	3.1771107	-1.8837641
131.071	-0.770047	0.2486026	-0.2005632	0.9523655	0.5854646
131.072	-0.556235	-0.116612	0.5663073	0.5946854	-0.1074998
135.031	5.2295258	0.187076	1.2081436	2.4564859	0.8080518
137.024	-0.9654724	-0.5625457	-0.7164285	-0.9150006	3.567047
144.046	-0.361132	-0.0504132	-0.4325516	0.85813603	1.50700443
144.066	-0.3845491	-0.0042597	0.14862254	-0.0068356	0.41313839
144.104	-0.1008955	-0.0497774	-0.0431978	0.49466212	-0.2272134
145.06	-0.6046738	1.2515007	0.8869745	1.0370456	0.7620225

3

This is how the negative serum file should look

MZ	Neg_S1	Neg_S2	Neg_S3	Neg_S4	Neg_S5
73.029	-0.7066796	0.2701793	-0.6095243	0.2846503	0.3029947
103.039	-1.28746	1.2614952	-1.48564	-0.1884103	1.3185082
109.028	-0.6609991	-0.7722004	-0.5863185	2.4622475	-0.5010598
116.072	-2.8234198	-1.2951277	-2.5600634	1.7398179	-1.8358343
117.056	-3.9525935	0.2232241	-0.9898427	-1.5528633	4.4258835
121.029	0.107048	-0.3653264	-0.3768747	-0.1462416	0.7067397
123.046	-0.188088	-0.5598863	-0.3479292	0.42264337	0.560712
128.035	-1.7571482	-1.6154556	0.02628267	0.06290076	-0.5133248
129.056	-0.1286156	-0.4778058	-0.4308022	0.62346641	0.04635066
130.087	1.6030706	-2.323879	-4.2840137	4.3082191	-0.561254
130.088	0.1891679	-2.5863634	-0.4762054	3.7310177	-2.1247616
131.071	-1.2705645	-0.6433814	-0.346667	1.1896377	0.2551527
131.072	-0.6146509	-0.4036271	0.2828018	0.1691924	0.1856379
135.031	-1.7793796	-2.3391007	-1.4529648	-2.0577609	-2.2600771
137.024	-1.0716456	-0.8201156	-0.7788997	-0.8778128	3.1408739
144.046	-0.7941788	-0.5594279	-0.6581984	-0.4659494	0.9567108
144.066	-0.4991423	-0.2117689	0.09079729	0.10468898	0.34930837
144.104	-0.2053877	-0.1570552	-0.0519183	0.57650822	-0.2357251
145.06	-1.176977	-1.2540316	-0.7165189	1.2911	-1.4764419

4

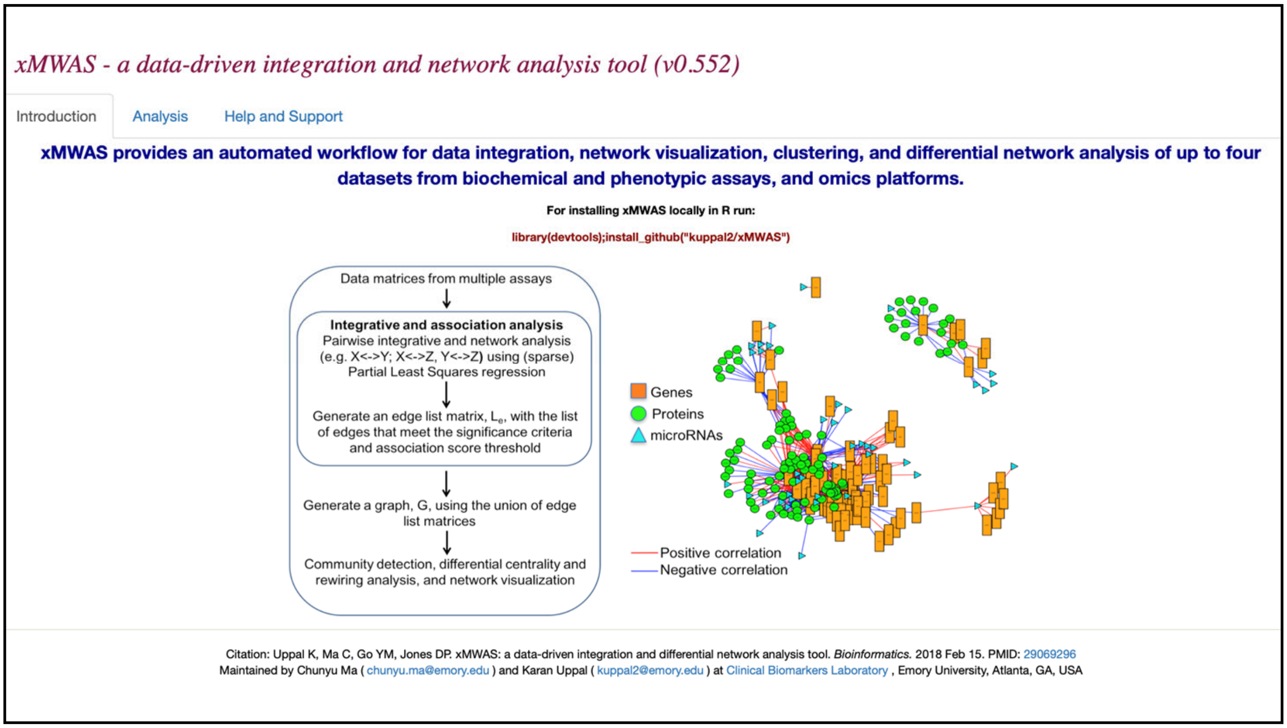
This is how the class file should look

	Class	
Neg_P1	Plasma	
Neg_P2	Plasma	
Neg_P3	Plasma	
Neg_P4	Plasma	
Neg_P5	Plasma	
Neg_S1	Serum	
Neg_S2	Serum	
Neg_S3	Serum	
Neg_S4	Serum	
Neg_S5	Serum	

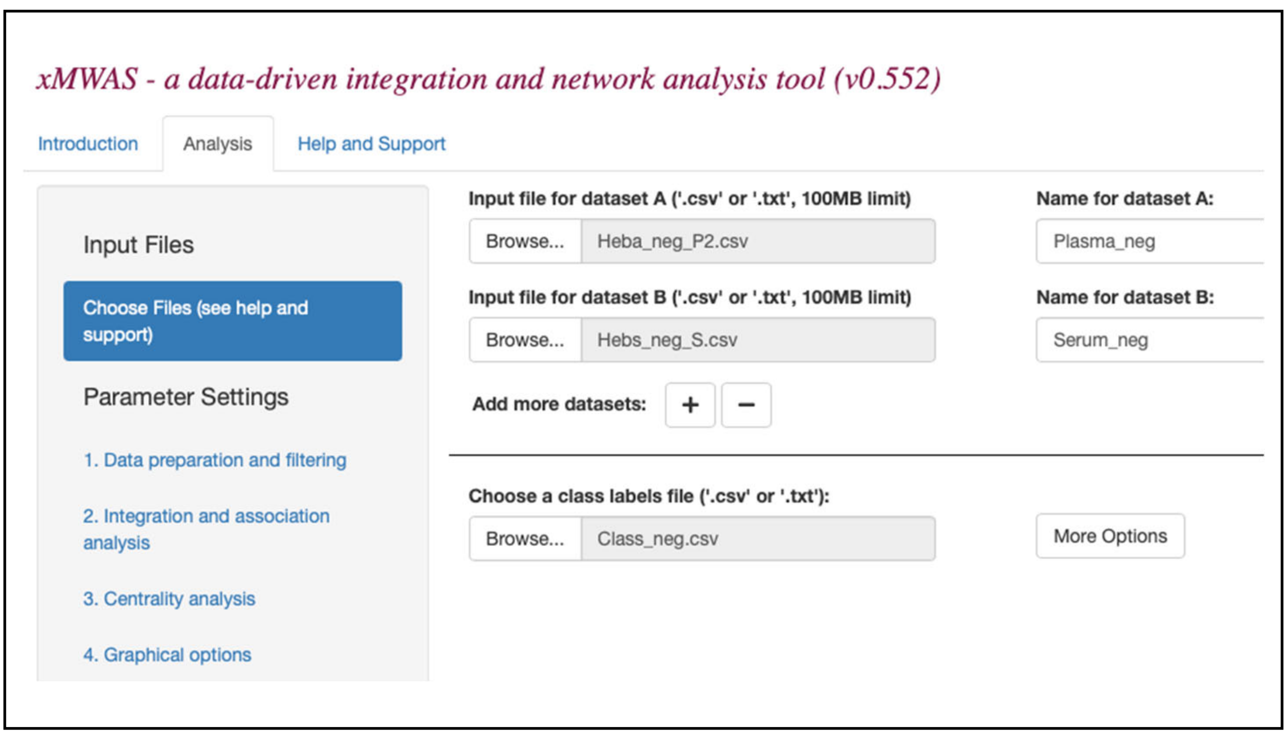
5

Go to https://kuppal.shinyapps.io/xmwas/

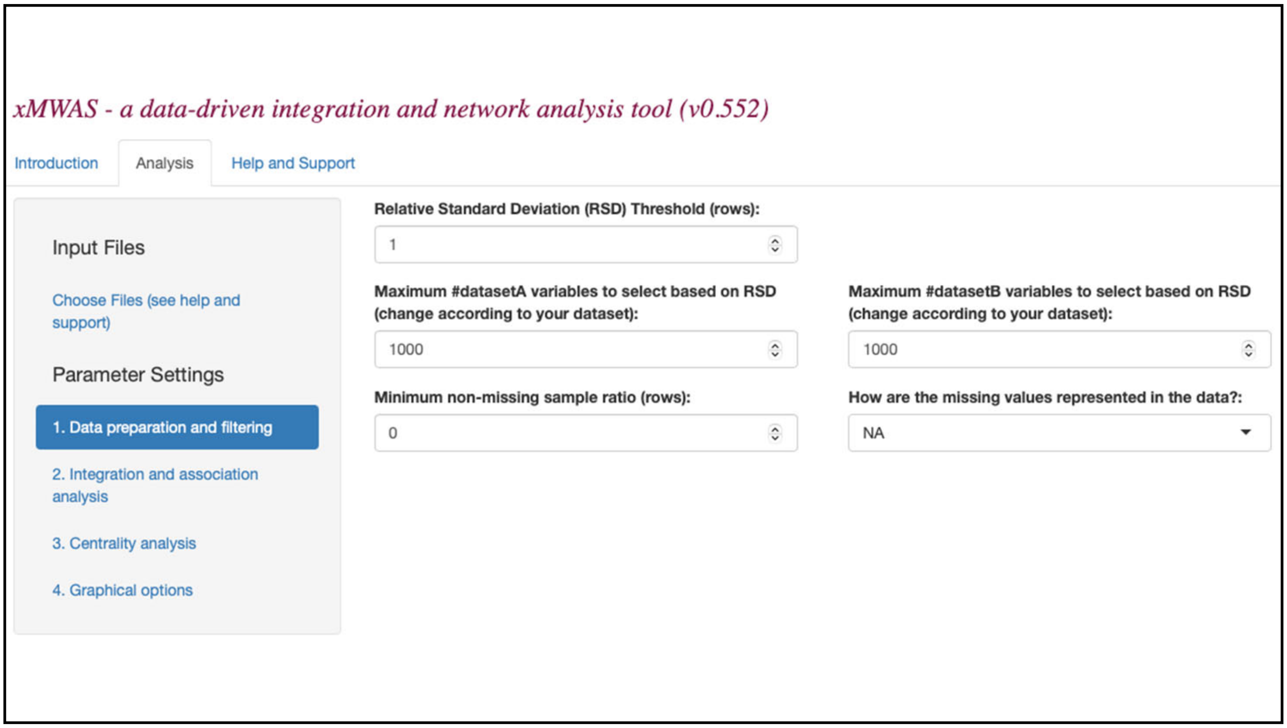
6


xMWAS - a data-driven integration and network analysis tool (v0.552)
Introduction
Analysis
Help and Support
xMWAS provides an automated workflow for data integration, network visualization, clustering, and differential network analysis of up to four datasets from biochemical and phenotypic assays, and omics platforms.
For installing xMWAS locally in R run:
library(devtools);install_github("kuppal2/xMWAS")
Data matrices from multiple assays
Integrative and association analysis
Pairwise integrative and network analysis (e.g. X<->Y; X<->Z, Y<->Z) using (sparse) Partial Least Squares regression
Generate an edge list matrix, Le, with the list of edges that meet the significance criteria and association score threshold
Generate a graph, G, using the union of edge list matrices
Community detection, differential centrality and rewiring analysis, and network visualization
Genes
Proteins
microRNAs
Positive correlation
Negative correlation
Citation: Uppal K, Ma C, Go YM, Jones DP. xMWAS: a data-driven integration and differential network analysis tool. Bioinformatics. 2018 Feb 15. PMID: 29069296
Maintained by Chunyu Ma (chunyu.ma@emory.edu) and Karan Uppal (kuppal2@emory.edu) at Clinical Biomarkers Laboratory , Emory University, Atlanta, GA, USA

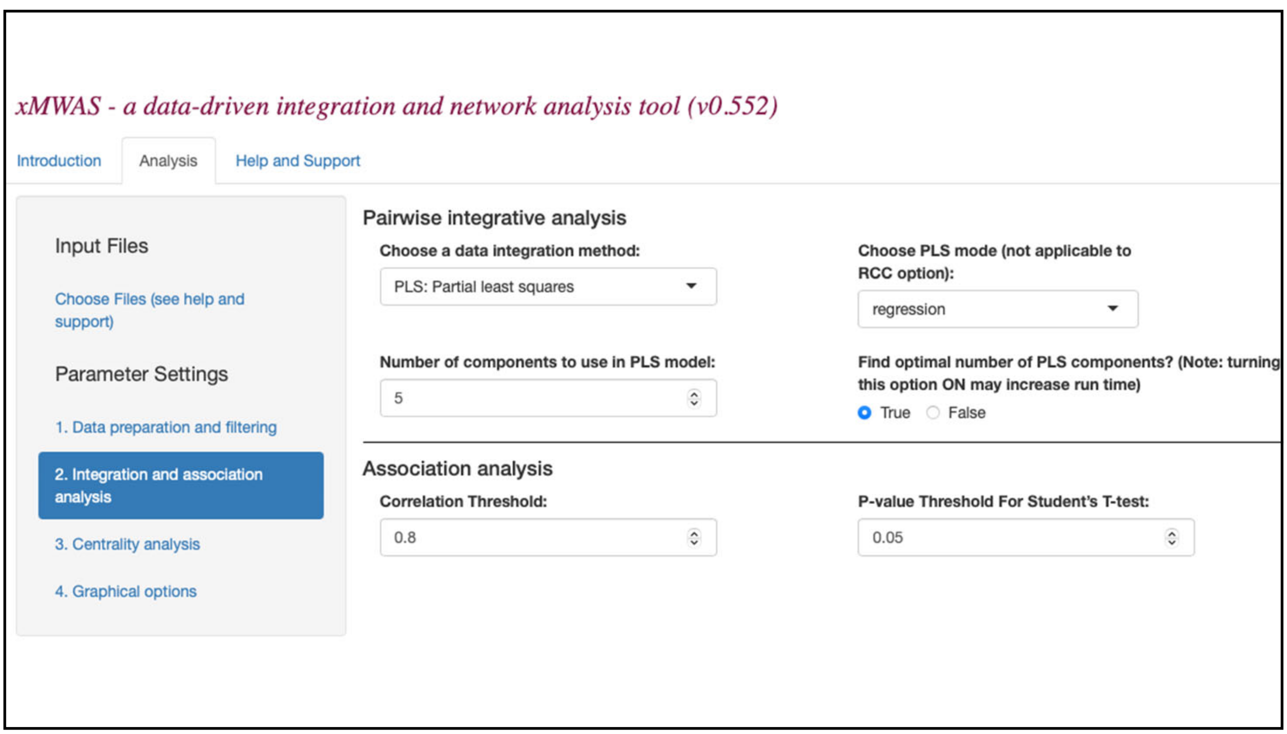
7


xMWAS - a data-driven integration and network analysis tool (v0.552)
Introduction
Analysis
Help and Support
Input Files
Choose Files (see help and support)
Parameter Settings
1. Data preparation and filtering
2. Integration and association analysis
3. Centrality analysis
4. Graphical options
Input file for dataset A ('.csv' or '.txt', 100MB limit)
Browse...
Heba_neg_P2.csv
Name for dataset A:
Plasma_neg
Input file for dataset B ('.csv' or '.txt', 100MB limit)
Browse...
Hebs_neg_S.csv
Name for dataset B:
Serum_neg
Add more datasets:
Choose a class labels file ('.csv' or '.txt'):
Browse...
Class_neg.csv
More Options

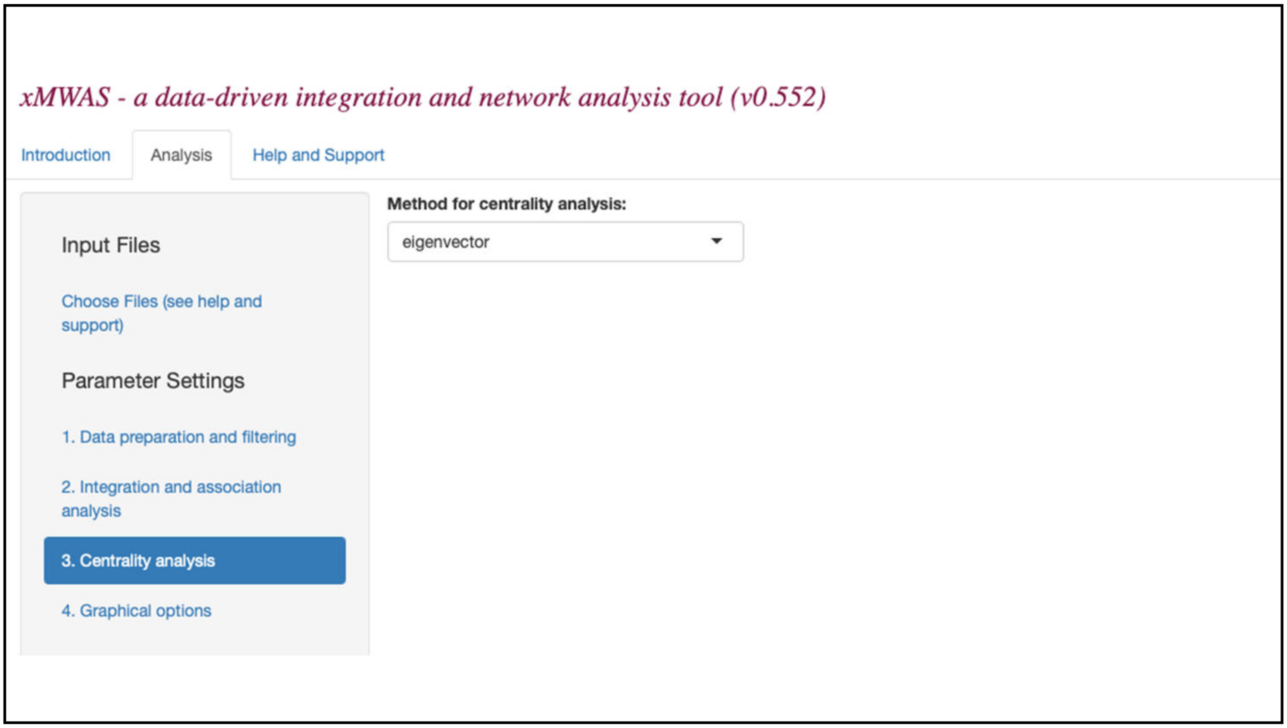
8


xMWAS - a data-driven integration and network analysis tool (v0.552)
Introduction
Analysis
Help and Support
Input Files
Choose Files (see help and support)
Parameter Settings
1. Data preparation and filtering
2. Integration and association analysis
3. Centrality analysis
4. Graphical options
Relative Standard Deviation (RSD) Threshold (rows):
1
Maximum #datasetA variables to select based on RSD (change according to your dataset):
1000
Maximum #datasetB variables to select based on RSD (change according to your dataset):
1000
Minimum non-missing sample ratio (rows):
0
How are the missing values represented in the data?:
NA

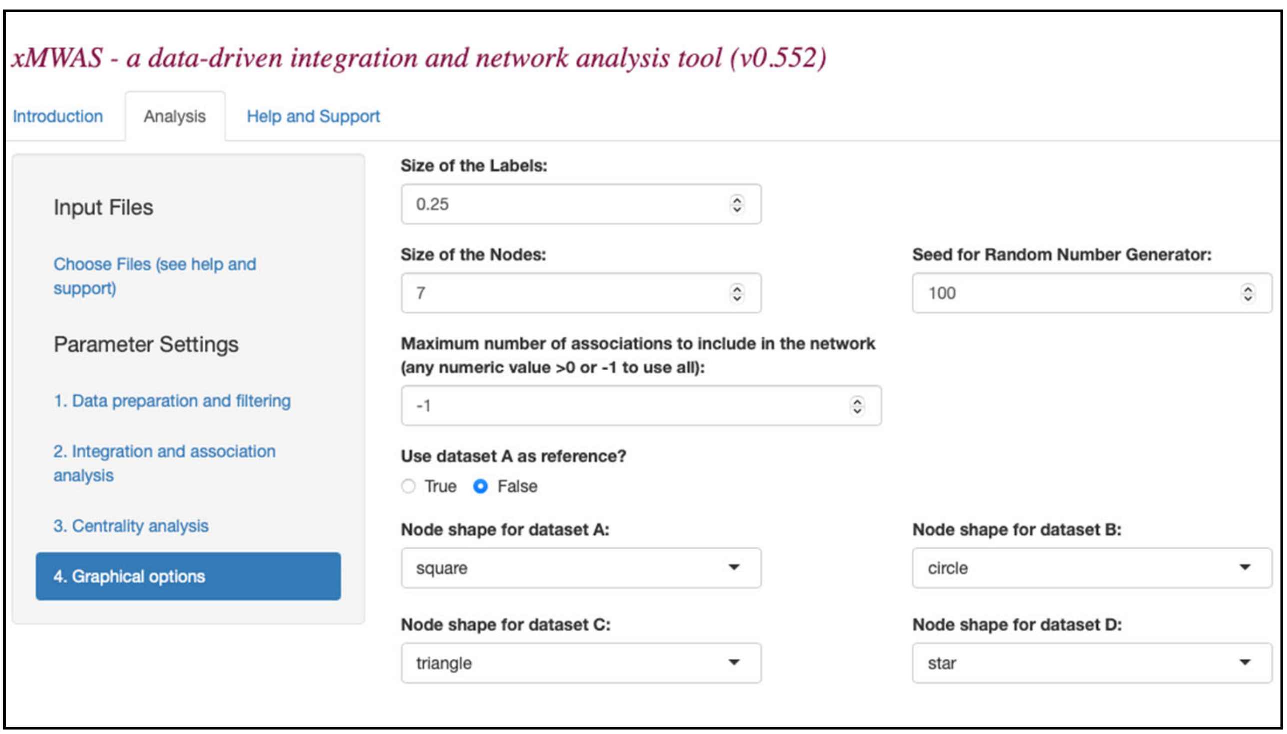
9


xMWAS - a data-driven integration and network analysis tool (v0.552)
Introduction
Analysis
Help and Support
Input Files
Choose Files (see help and support)
Parameter Settings
1. Data preparation and filtering
2. Integration and association analysis
3. Centrality analysis
4. Graphical options
Pairwise integrative analysis
Choose a data integration method:
PLS: Partial least squares
Choose PLS mode (not applicable to RCC option):
regression
Number of components to use in PLS model:
5
Find optimal number of PLS components? (Note: turning this option ON may increase run time)
True
False
Association analysis
Correlation Threshold:
0.8
P-value Threshold For Student's T-test:
0.05

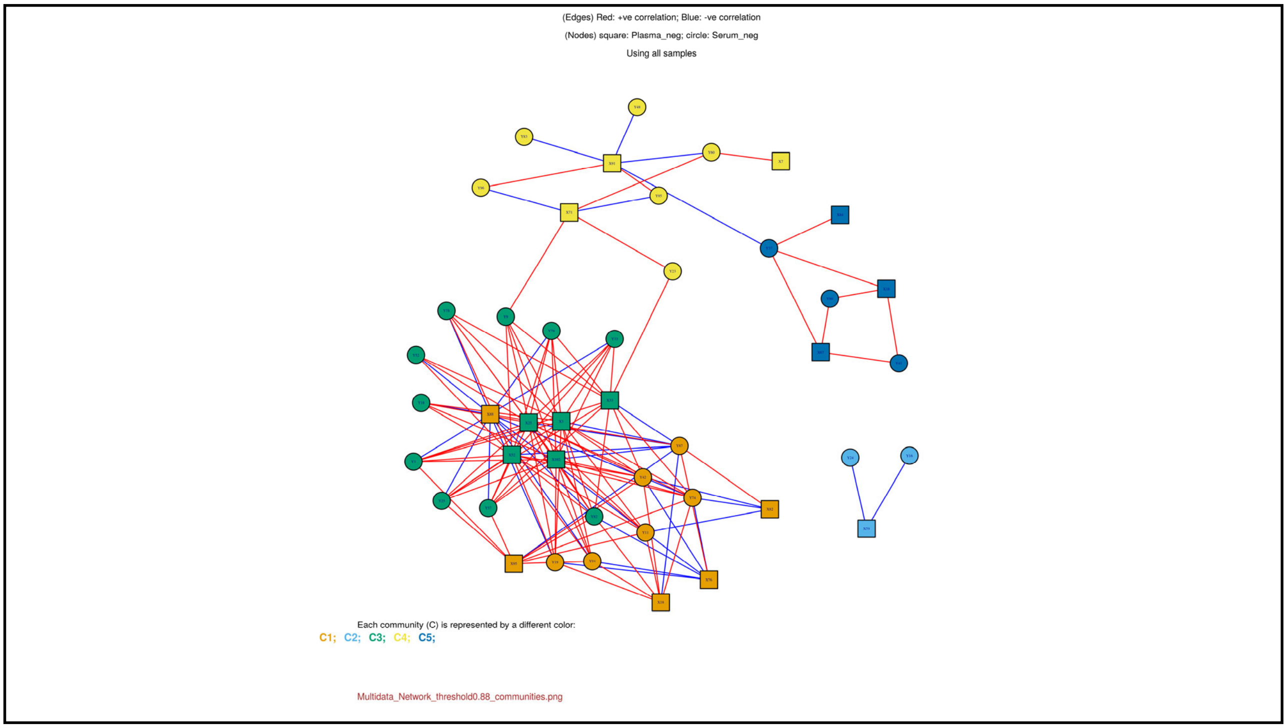
10


xMWAS - a data-driven integration and network analysis tool (v0.552)
Introduction
Analysis
Help and Support
Input Files
Choose Files (see help and support)
Parameter Settings
1. Data preparation and filtering
2. Integration and association analysis
3. Centrality analysis
4. Graphical options
Method for centrality analysis:
eigenvector

11


xMWAS - a data-driven integration and network analysis tool (v0.552)
Introduction
Analysis
Help and Support
Input Files
Choose Files (see help and support)
Parameter Settings
1. Data preparation and filtering
2. Integration and association analysis
3. Centrality analysis
4. Graphical options
Size of the Labels:
0.25
Size of the Nodes:
7
Seed for Random Number Generator:
100
Maximum number of associations to include in the network (any numeric value >0 or -1 to use all):
-1
Use dataset A as reference?
True
False
Node shape for dataset A:
square
Node shape for dataset B:
circle
Node shape for dataset C:
triangle
Node shape for dataset D:
star

12



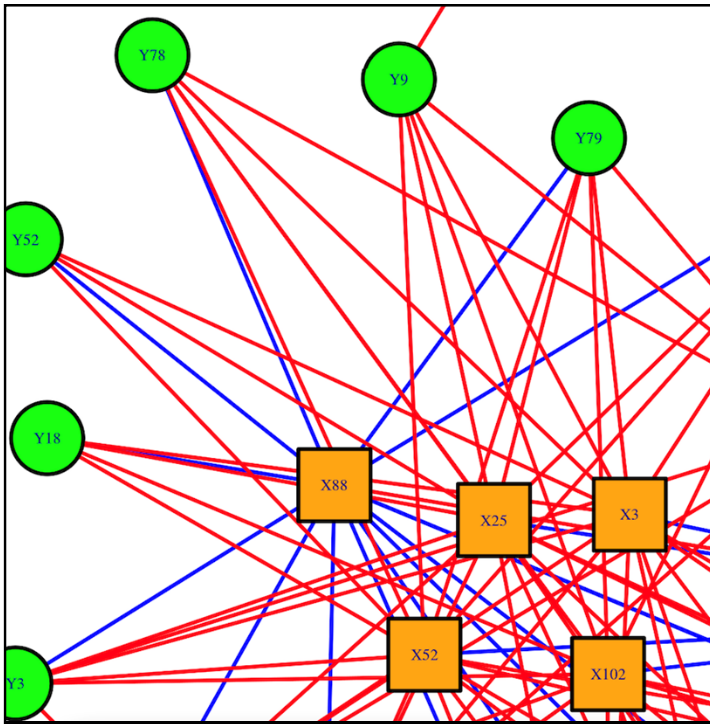
13

Ions that are positively and negatively correlated

"from"	"to"	"weight"
"X88"	"Y74"	-0.974461554765936
"X88"	"Y11"	-0.957519575289324
"X88"	"Y42"	-0.957247786453691
"X76"	"Y11"	-0.949781456148959
"X76"	"Y42"	-0.945544297253841
"X91"	"Y60"	-0.943228376751098
"X76"	"Y74"	-0.940560219642942
"X95"	"Y87"	-0.93987475281858
"X88"	"Y82"	-0.939421156416451
"X3"	"Y87"	-0.937401246222139
"X52"	"Y87"	-0.936641194555845
"X25"	"Y87"	-0.935356776695009
"X88"	"Y3"	-0.935299964205794
"X59"	"Y24"	-0.933581025743903
"X88"	"Y25"	-0.928943972777584
"X88"	"Y19"	-0.928310446095442
"X18"	"Y87"	-0.928106664960895
"X88"	"Y94"	-0.926704447018128
"X102"	"Y87"	-0.925694615764478
"X88"	"Y57"	-0.924681137818432
"X88"	"Y78"	-0.920761033370608
"X88"	"Y18"	-0.912587193789293
"X88"	"Y33"	-0.910162191832505
"X88"	"Y79"	-0.90453034204073
"X59"	"Y16"	-0.903099500203589
"X82"	"Y11"	-0.900057950407366

"X71"	"Y9"	0.926709444453419
"X33"	"Y79"	0.928359633726528
"X102"	"Y3"	0.928672146482609
"X102"	"Y78"	0.93033884763627
"X18"	"Y74"	0.931089645299694
"X25"	"Y33"	0.932349492143085
"X95"	"Y42"	0.933471312419702
"X3"	"Y25"	0.933704176880728
"X95"	"Y11"	0.934913989215837
"X102"	"Y74"	0.938412750009708
"X52"	"Y3"	0.938833450188028
"X52"	"Y78"	0.940132097334762
"X25"	"Y3"	0.942581797262386
"X95"	"Y74"	0.943733154425184
"X3"	"Y33"	0.94497587617022
"X76"	"Y87"	0.945542632719101
"X25"	"Y78"	0.946258580656965
"X25"	"Y74"	0.948875507199841
"X52"	"Y74"	0.949379765386428
"X33"	"Y3"	0.950118076536889
"X3"	"Y74"	0.952168988470733
"X3"	"Y3"	0.95234105622655
"X71"	"Y60"	0.953196814908068
"X33"	"Y9"	0.95688924784227
"X33"	"Y33"	0.95691151701754
"X3"	"Y78"	0.95966444007513
"X88"	"Y87"	0.966636929289798
"X33"	"Y78"	0.974588386326116

14



"X88"	"Y74"	-0.974461554765936
"X88"	"Y11"	-0.957519575289324
"X88"	"Y42"	-0.957247786453691

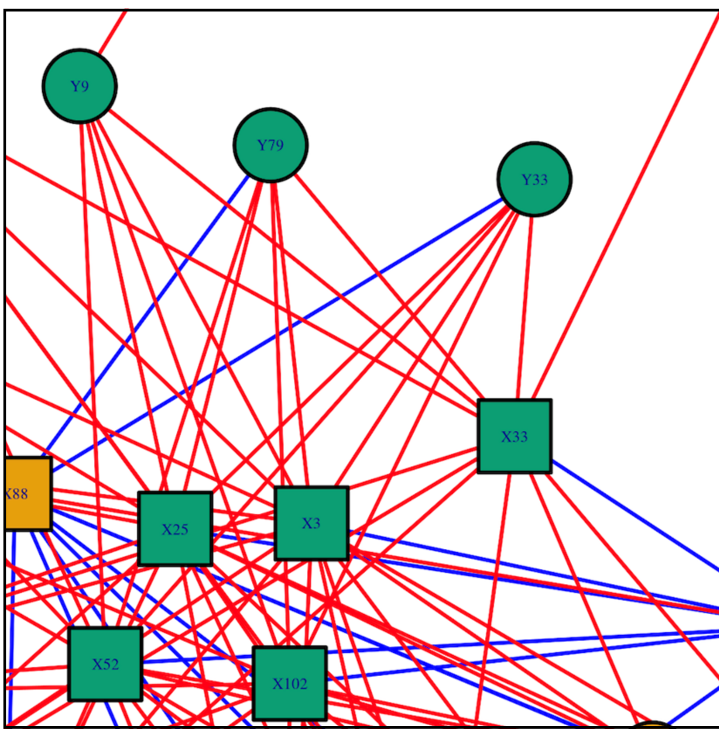
X88 is *m/z* 367.157 – DHEA sulfate

Y11 is *m/z* 130.088 - Leucine

Y42 is *m/z* 180.063 - tyrosine

Y74 is *m/z* 267.073 - Inosine

15



"X33"	"Y9"	0.95688924784227
"X33"	"Y33"	0.95691151701754
"X33"	"Y78"	0.974588386326116

X33 is *m/z* 158.063 – indole3-acetate

Y33 is *m/z* 158.063 – indole-3-acetate

Y78 is *m/z* 227.104 - prolyhydroxyproline

16

In class, we will now process the positive ion data

17