xMWAS installation and tutorial (Sep, 2017)

The main input parameters include file paths or data matrices of input datasets, file with phenotypic labels (e.g. case, control, time points), integration method ("pls", "spls"), threshhold for association score, and statistical significance. The software is designed to work with data from different sources and does not utilize any knowledgebase information for integration. For metabolomics studies, the software is designed to work with both targeted and untargeted datasets and does not require metabolite identification prior to integration. If the classlabels file is not provided, the analysis does not include class or group-wise comparisons and only includes integrative analysis results using all samples. The output includes PDF of integrative network with cluster assignments, text files for association matrices and cluster assignments for each node and their centrality scores, and a GML format file that can be used for visualization with tools such as Cytoscape (Shannon 2003). The instructions for installing the R pckage, preparing the input files, using the online version of the tool, and description of output files is provided below:

A. xMWAS R package installation

a) Install R: https://mirrors.nics.utk.edu/cran/

b) R commands for installing dependencies:

install.packages("devtools",repos="http://cran.r-project.org")

c) Install xMWAS:

From GitHub:

1) install_github("kuppal2/xMWAS")

2) Run "library(xMWAS)" command to make sure the package is successfully installed

B. Input files

I.   Data file: file with expression or intensity values from omics experiment or other assay (at least 2 data matrices are required)

Rows: Variables (genes, metabolites, proteins)
Columns: Intensity or expression values for individual samples

Example:

EntrezGeneID	CT22	CT23	CT24	CT48
1608	8.58	8.31	8.73	8.37
11287	5.21	5.39	5.15	5.16
11298	6.02	5.87	5.59	5.64

II.   Class labels file: file with phenotypic classification for individual samples (optional). The software supports multiple study designs: multiclass (2 or more classes, repeated measures with one factor, and repeated measures with two factors). Example class labels files for different study designs are provided below:

Multiclass (2 or more classes):

FileName	Class
CT22	CNS
CT23	CNS
CT24	CNS
CT48	CNS
CT28	LE
CT29	LE
CT30	LE
CT31	LE

Repeated measures study design (one factor):

FileName	Subject	Factor1
S1	Subject001	TP1
S2	Subject002	TP1
S3	Subject003	TP1
S4	Subject004	TP1
S5	Subject005	TP1
S6	Subject006	TP1
S8	Subject001	TP2
S9	Subject002	TP2
S10	Subject003	TP2
S11	Subject004	TP2
S12	Subject005	TP2
S13	Subject006	TP2
S15	Subject001	TP3
S16	Subject002	TP3
S17	Subject003	TP3
S18	Subject004	TP3
S19	Subject005	TP3
S20	Subject006	TP3

Repeated measures study design (two factors):

FileName	Subject	Factor1	Factor2
S1	Subject001	TP1	Case
S2	Subject002	TP1	Case
S3	Subject003	TP1	Case
S4	Subject004	TP1	Control
S5	Subject005	TP1	Control
S6	Subject006	TP1	Control
S8	Subject001	TP2	Case
S9	Subject002	TP2	Case
S10	Subject003	TP2	Case
S11	Subject004	TP2	Control
S12	Subject005	TP2	Control
S13	Subject006	TP2	Control
S15	Subject001	TP3	Case
S16	Subject002	TP3	Case
S17	Subject003	TP3	Case
S18	Subject004	TP3	Control
S19	Subject005	TP3	Control
S20	Subject006	TP3	Control

C. Using xMWAS in R

Example script:
https://sourceforge.net/projects/xmwas/files/example_xmwas_runscript_v0.4.R

```
#load package
library(xMWAS)

#example dataset that includes metabolome, transcriptome, and cytokine data from the
H1N1 mice study (Chandler 2016)
data(exh1n1)
data(classlabels_casecontrol) #example classlabels file for case vs control design
data(classlabels_repeatmeasures) #example classlabels file for repeat measures design
xMat<-exh1n1$metabolome
yMat<-exh1n1$transcrlptome
zMat<-exh1n1$cytokine
classlabels<-exh1n1$classlabels
```

output<-"/Users/karanuppal/xMWASoutput/" #change for your computer

Please see user manual for description of arguments:
https://sourceforge.net/projects/xmwas/files/xMWAS-manual.pdf/download

```
#call the run_xmwas() function:
xmwas_res<-
run_xmwas(Xome_data=xMat,Yome_data=yMat,Zome_data=zMat,Wome_data=NA,outl
oc=output,
classlabels=classlabels,class_fname=NA,xmwasmethod="spls",plsmode="canonical",m
ax_xvar=5000,max_yvar=5000,
max_zvar=5000,max_wvar=5000,rsd.filt.thresh=1,corthresh=0.3,keepX=100,keepY=100
,keepZ=100,keepW=100,
pairedanalysis=FALSE,optselect=TRUE,rawPthresh=0.05,numcomps=10,net_edge_col
ors=c("blue","red"),
net_node_colors=c("orange",
"green","cyan","gold"),Xname="X",Yname="Y",Zname="Z",Wname="W",
net_node_shape=c("rectangle","circle","triangle","star"),all.missing.thresh=0.3,
seednum=100,label.cex=0.2,vertex.size=6,graphclustering=TRUE,interactive=FALSE,m
ax_connections=10000,
centrality_method="eigenvector",use.X.reference=FALSE,removeRda=TRUE)
suppressWarnings(try(sink(file=NULL),silent=TRUE))
```
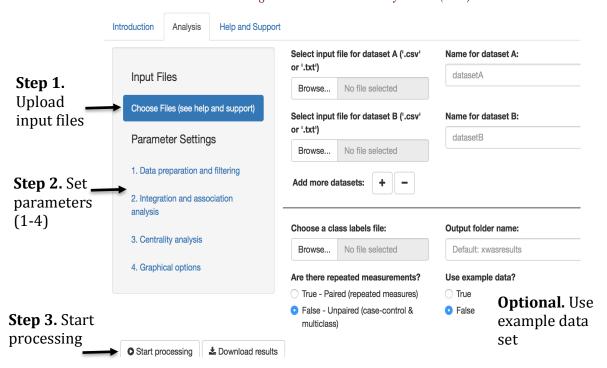
D. Using xMWAS online (URL: https://kuppal.shinyapps.io/xmwas/)

The format of input files is same as described in Section 2: "Input Files". Please see the "Help and Support" page on the website for user manual, description of the arguments, input files and output description:

xMWAS - a data-driven integration and network analysis tool (v0.4)

E. xMWAS output:

	cluster_membership_centrality_table.txt
▶ 📁	Ct
▶ 📁	CtH
	InputParameters.txt
	LogWed_Sep_13_15_08_07_2017.txt
	matrix_centrality.txt
	Multidata_Network_corthresh0.64cytoscapeall.gml
🖼	Multidata_Network_threshold0.64_communities.png
🖼	Multidata_Network_threshold0.64.png
▶ 📁	pairwise_results
	README.txt

Description of output files:

1. cluster_membership_centrality_table.txt: includes the community membership and centrality measures for nodes that meet the association criteria and included in the association networks.

2. InputParameters.txt: includes list of input arguments provided to the software and version numbers of all packages to facilitate reproducibility of results.

3. LogMon_Jul_10_01_52_08_2017.txt: includes runtime output messages

4. matrix_centrality.txt: includes the centrality measures across different conditions for nodes that meet the association criteria and included in the association networks.

5. Multidata_Network_corthreshX_communities.pdf: includes multiome network plot with the communities identified using the multilevel community detection algorithm. Members of each community are assigned colors based on community/module/cluster membership (1: orange; 2: light blue; 3: dark green, and so on). X labels correspond to xome_fname data, Y labels correspond to yome_fname data, Z labels correspond to zome_fname data, W labels correspond to wome_fname data.

6. Multidata_Network_corthreshX.pdf: includes multiome network plot using all significantly associated variables. X labels correspond to xome_fname data, Y labels correspond to yome_fname data, Z labels correspond to zome_fname data, W labels correspond to wome_fname data.

7. Multidata_Network_corthreshXcytoscapeall.gml: includes GML file for all significantly associated variables that can be uploaded to Cytoscape

8. pairwise_results/

   Pairwise integrative analysis results are under pairwise_results folder. The files corresponding to each pairwise comparison (X<->Y, X<->Z, Y<-Z,..) are: XYassociation_matrix_corthresh0.9.txt (correlation matrix with mapping between node labels and original variable names), XYassociation_networkthresholdX.pdf that includes the pairwise network plots, XYBoolean_association_matrix_corthreshX.txt (same as correlation matrix but correlations meeting the threshold are represented 1, and 0 otherwise)

9. README.txt: includes description of output files and additional notes

10. Ct/ and CtH/: sub-folders with network analysis and community detection results for each class

▼ 📁 Ct
    📄 cluster_membership_centrality_table.txt
    📄 Multidata_Network_corthresh0.88cytoscapeall.gml
    🖼 Multidata_Network_threshold0.88_communities.png
    🖼 Multidata_Network_threshold0.88.png
  ▶ 📁 pairwise_results
▼ 📁 CtH
    📄 cluster_membership_centrality_table.txt
    📄 Multidata_Network_corthresh0.88cytoscapeall.gml
    🖼 Multidata_Network_threshold0.88_communities.png
    🖼 Multidata_Network_threshold0.88.png
  ▶ 📁 pairwise_results