

Package ‘xMWAS’

February 26, 2018

Type Package
Title xMWAS: R package for data-integration and differential network analysis
Version 0.52
Date 2018-02-26
Author Karan Uppal
Maintainer Karan Uppal <kuppal2@emory.edu>

Description
xMWAS allows integration of omics data using multivariate methods such as PLS. The package performs community detection and network analysis to allow visualization of positive or negative associations between different datasets generated using samples from the same individuals.

License GPL2.0
LazyLoad no
Depends mixOmics, WGCNA, snow, igraph, graph, RBGL, plyr, plsgenomics

R topics documented:

xMWAS_0.52-package	1
run_xmwas	2
Index	5

xMWAS_0.52-package	<i>xMWAS</i>
--------------------	--------------

Description
xMWAS allows integration of omics data using multivariate methods such as PLS. The package performs community detection and network analysis to allow visualization of positive or negative associations between different datasets generated using samples from same individuals. The run_xmwas() function performs the integrative, community detection, and network analysis. The output files include .gml files that can be used with software like Cytoscape.

Details

Package: xMWAS
 Type: Package
 Version: 0.52
 Date: 2018-02-26
 License: gpl2.0
 LazyLoad: yes

Author(s)

Karan Uppal Maintainer: <kuppal2@emory.edu>

run_xmwas	<i>run_xmwas</i>
-----------	------------------

Description

The function uses sPLS or PLS and network functions in mixOmics package to perform pairwise integrative and correlation analysis. The pairwise graphs are merged using igraph and community detection is performed using the Multilevel clustering algorithm. Association networks can be visualized in R or using Cytoscape.

Usage

```
run_xmwas(xome_fname = NA, yome_fname = NA, zome_fname = NA, wome_fname = NA,
  outloc = NA,
  class_fname = NA, Xome_data = NA, Yome_data = NA, Zome_data = NA, Wome_data = NA,
  classlabels = NA, xmwasmethod = "spls", plsmode = "canonical", max_xvar = 5000,
  max_yvar = 5000, max_zvar = 5000, max_wvar = 5000, rsd.filt.thresh = 1,
  all.missing.thresh = 0.5, missing.val = 0, corthresh = 0.4, keepX = 1000,
  keepY = 1000, keepZ = 1000, keepW = 1000, pairedanalysis = FALSE,
  optselect = TRUE, rawPthresh = 0.05, numcomps = 10,
  net_edge_colors = c("blue", "red"), net_node_colors = c("orange", "green", "blue", "pink"),
  Xname = "X", Yname = "Y", Zname = "Z", Wname = "W",
  net_node_shape = c("circle", "rectangle", "triangle", "star"), seednum = 100,
  label.cex = 0.3, vertex.size = 6, max_connections = 1e+05,
  centrality_method = "eigenvector", use.X.reference = FALSE,
  removeRda = TRUE, compare.classes = TRUE, ...)
```

Arguments

xome_fname	Full path with filename for dataset A. Default: NA; The software uses the value provided for Xome_data when this is set to NA.
yome_fname	Full path with filename for dataset B Default: NA; The software uses the value provided for Yome_data when this is set to NA.
zome_fname	Full path with filename for dataset C Default: NA; The software uses the value provided for Zome_data when this is set to NA.

wome_fname	Full path with filename for dataset D Default: NA; The software uses the value provided for Wome_data when this is set to NA.
Xome_data	Data matrix for dataset A. Run: data(exnci60); head(exnci60\$mrna) to see how to format data matrices.
Yome_data	Data matrix for dataset B
Zome_data	Data matrix for dataset C
Wome_data	Data matrix for dataset D
outloc	Output directory
classlabels	Data matrix with phenotype information. Set to NA if this information is not available. see: data(classlabels_casecontrol) for case vs control design or data(classlabels_repeatmeasures) for repeat measures
class_fname	File with phenotype information. Set to NA if this information is not available. see: data(classlabels_casecontrol) for case vs control design or data(classlabels_repeatmeasures) for repeat measures
xmwasmethod	Method for data integration. eg: "pls": partial least squares regression "spls": sparse partial least squares regression "olpls": orthogonal partial least squares regression
plsmode	"canonical" for bi-directional relationships; "regression" for regression/predictive relationships
max_xvar	Maximum number of X variables to select based on relative standard deviation (RSD). e.g. 10000
max_yvar	Maximum number of Y variables to select based on relative standard deviation (RSD). e.g. 10000
max_zvar	Maximum number of Z variables to select based on relative standard deviation (RSD). e.g. 10000
max_wvar	Maximum number of W variables to select based on relative standard deviation (RSD). e.g. 10000
rsd.filt.thresh	Relative standard deviation (sd/mean) threshold
all.missing.thresh	Remove variables (rows) that do not meet the minimum threshold for presence of non-missing values. e.g. 0.8
missing.val	How are the missing values represented in the input data files? Default: 0
corthresh	Correlation threshold. eg: 0.7
keepX	Maximum number of X variables to select in sPLS. Note: keepX, keepY, keepZ, and keepW are only used when xmwasmethod is set to "spls"
keepY	Maximum number of Y variables to select in sPLS. Note: keepX, keepY, keepZ, and keepW are only used when xmwasmethod is set to "spls"
keepZ	Maximum number of Z variables to select in sPLS. Note: keepX, keepY, keepZ, and keepW are only used when xmwasmethod is set to "spls"
keepW	Maximum number of W variables to select in sPLS. Note: keepX, keepY, keepZ, and keepW are only used when xmwasmethod is set to "spls"
pairedanalysis	Are their repeated measurements? TRUE or FALSE
optselect	Find optimal number of PLS components. TRUE or FALSE
rawPthresh	p-value threshold calculated using Student's t-test. eg: 0.05

numcomps	Number of components to use in PLS model. eg: 3
net_edge_colors	Colors for edges.
net_node_colors	Colors for nodes.
Xname	Name for X dataset. eg: "Genes"
Yname	Name for Y dataset. eg: "Proteins"
Zname	Name for Z dataset. eg: "Metabolites"
Wname	Name for W dataset. eg: "EnvironmentalExposures"
net_node_shape	Shapes for nodes.
seednum	Seed for random number generator used for plotting the network.
label.cex	Size of the labels. eg: 0.8
vertex.size	Size of the nodes.
max_connections	Maximum number of associations to include in the network.
centrality_method	Method for centrality measures. Options: 1) "eigenvector" for eigenvector centrality, which is based on the number and quality of connections - centrality scores range from 0 to 1, where 1 means high centrality and 0 means low or no centrality. Nodes/vertices with high centrality scores are connected to many other nodes, which are in turn connected to many other nodes. Please see <code>igraph::eigen_centrality</code> function for more details. 2) "betweenness" for betweenness centrality, which is based on the number of shortest paths going through a node/vertex - centrality scores are normalized and scaled to 0 to 1 range, where 1 means high betweenness centrality and 0 means low or no betweenness centrality. Please see <code>igraph::betweenness</code> function for more details. 3) "degree.count" based on the number of connections of a node. -centrality scores are normalized and scaled to 0 to 1 range. High centrality means more connections. Please see <code>igraph::degree</code> function for more details. 4) "degree.weight" is based on the sum of absolute weights of connections of a node. -centrality scores are normalized and scaled to 0 to 1 range. High centrality means stronger connections.
use.X.reference	TRUE or FALSE if you want to use <code>Xome_data</code> as reference. If TRUE, only <code>X<->Y</code> , <code>X<->Z</code> , and <code>X<->W</code> pairwise analysis will be performed.
removeRda	TRUE or FALSE; set to TRUE if you want to remove the intermediate files.
compare.classes	TRUE or FALSE; set to TRUE if you want to compare individual classes as provided in class labels file.

Author(s)

Karan Uppal kuppal2@emory.edu

Index

*Topic **xMWAS**

xMWAS_0.52-package, [1](#)

run_xmwas, [2](#)

xMWAS (xMWAS_0.52-package), [1](#)

xMWAS_0.52-package, [1](#)