Package 'fastLiquidAssociation'

November 12, 2026

Type Package

Title functions for genome-wide application of Liquid Association

Version 1.48.0

Date 2017-10-14

Author Tina Gunderson

Maintainer Tina Gunderson <gunderson.tina@gmail.com>

Description

This package extends the function of the LiquidAssociation package for genome-wide application. It integrates a screening method into the LA analysis to reduce the number of triplets to be examined for a high LA value and provides code for use in subsequent significance analyses.

Depends methods, LiquidAssociation, parallel, doParallel, stats, Hmisc, utils

Imports WGCNA, impute, preprocessCore

Suggests GOstats, yeastCC, org.Sc.sgd.db

biocViews Software, GeneExpression, Genetics, Pathways, CellBiology

License GPL-2

LazyLoad yes

git_url https://git.bioconductor.org/packages/fastLiquidAssociation

git_branch RELEASE_3_23

git_last_commit fc71194

git_last_commit_date 2026-04-28

Repository Bioconductor 3.23

Date/Publication 2026-11-11

Contents

fastLiquidAssociation-package 2
fastboots.GLA-methods 3
fastMLA-methods 5
mass.CNM-methods 7
testmat 8

Index **10**

`fastLiquidAssociation-package`
Functions to extend liquid association (LA) analysis for genome-wide use

Description

This package provides three external functions to extend liquid association (LA) analysis for genome-wide use: fastMLA, mass.CNM, and fastboots.GLA. It contains updated versions of several functions available in the LiquidAssociation package, but whereas those functions are written to be used with a single triplet, this package's functions are intended to be applied against a full data set.

Details

Package:	fastLiquidAssociation
Type:	Package
Version:	1.1.7
Date:	2014-10-01
License:	GPL-2

Contains three external functions: fastMLA, mass.CNM, and fastboots.GLA. fastMLA uses an algorithm to approximate liquid association values in order to screen for a subset likely to have high LA values and then evaluates that subset for their LA values. mass.CNM takes the results from fastMLA and attempts to estimate significance based on conditional normal models. In the event that mass.CNM is unable to provide an estimate, fastboots.GLA parallelizes the bootstrapping required to produce a more robust direct estimate for LA.

Author(s)

Author: Tina Gunderson Maintainer: Tina Gunderson <gunderson.tina@gmail.com>

References

[1] Yen-Yi Ho, Giovanni Parmigiani, Thomas A Louis, and Leslie M Cope. Modeling liquid association. Biometrics, 67(1):133-141, 2011.

[2] Ker-Chau Li. Genome-wide coexpression dynamics: theory and application. Proceedings of the National Academy of Sciences, 99(26):16875-16880, 2002.

[3] Paul T Spellman, Gavin Sherlock, Michael Q Zhang, Vishwanath R Iyer, Kirk Anders, Michael B Eisen, Patrick O Brown, David Botstein, and Bruce Futcher. Comprehensive identification of cell cycle regulated genes of the yeast saccharomyces cerevisiae by microarray hybridization. Molecular biology of the cell, 9(12):3273-3297, 1998.

See Also

LiquidAssociation,WGCNA,parallel,yeastCC,GOstats,org.Sc.sgd.db

Examples

```
library(yeastCC)
data(spYCCES)
lae <- spYCCES[,-(1:4)]
### get rid of samples with high % NA elements
lae <- lae[apply(is.na(exprs(lae)),1,sum) < ncol(lae)*0.3,]
data <- t(exprs(lae))
data <- data[,1:50]

##fastMLA
example <- fastMLA(data=data, topn=25, nvec=1:10, rvalue=1.0, cut=4, threads = detectCores())
example[1:5,]

##mass.CNM
CNMcalc <- mass.CNM(data=data, GLA.mat=example, nback=10)
CNMcalc

##fastboots.GLA
clust <- makeCluster(4)
ex <- example[1:5,]
GLAeasy <- fastboots.GLA(ex, data=data, clust=clust, boots=30, perm=100, cut=4)
GLAeasy
stopCluster(clust)
closeAllConnections()
```

---

fastboots.GLA-methods *Function to parallelize bootstrapping for MLA robust estimates.*

Description

Function to increase speed of bootstrapping for MLA robust estimates. It performs the bootstrapping in parallel, so the time decrease will be dependent on the number of cores or CPUs available. Intended for use with results of triplets not processed sensibly by call to either CNM.full or CNM.simple in mass.CNM, though it can also be used with the direct results of fastMLA.

Usage

```
fastboots.GLA(tripmat, data, clust, boots = 30, perm = 100, cut = 4)
```

Arguments

tripmat	Matrix specifying the triplets to be estimated. Intended for use with the results of the mass.CNM or fastMLA functions.
data	Matrix of numeric data, with columns representing genes and rows representing observations.
clust	Cluster of CPU cores created by makeCluster. See parallel.
boots	Number of bootstrap iterations to estimate direct estimate SE.

perm	Number of permutations to use in estimating p-value.
cut	Value passed to the internal clusterGLA function to create buckets (equal to number of buckets+1). Values placing between 15-30 samples per bucket are optimal. Must be a positive integer>1. See GLA.

Details

Choosing the number of bins: For example, assume that our data has 100 observations. Since values between 15-30 observations per bin are optimal, good values to choose would be 5-7.

Value

fastboots.GLA returns a data.frame that specifies the genes in positions X1, X2 and X3; the rhodiff value of the triplet, the GLA value of the triplet, the direct estimate statistic, and the direct estimate p-value. More detailed explanation is available in the package vignette.

Warning

The data matrix must be numeric.

Note

The cluster of CPUs to use for bootstrapping must be created with makeCluster from the parallel package before running the fastboots.GLA function.

Author(s)

Tina Gunderson

References

[1] Yen-Yi Ho, Giovanni Parmigiani, Thomas A Louis, and Leslie M Cope. Modeling liquid association. Biometrics, 67(1):133-141, 2011.

See Also

fastMLA, mass.CNM, LiquidAssociation, parallel

Examples

```
#to view the code for the function
selectMethod("fastboots.GLA", signature=c("ANY","matrix"))

#
library(yeastCC)
data(spYCCES)
lae <- spYCCES[,-(1:4)]
### get rid of samples with high NA elements
lae <- lae[apply(is.na(exprs(lae)),1,sum) < ncol(lae)*0.3,]
data <- t(exprs(lae))
data <- data[,1:50]
```

```
dim(data)

example <- fastMLA(data=data, topn=25, nvec=1:10, rvalue=1.0, cut=4)
clust <- makeCluster(4)
ex <- example[1:5,]
GLAeasy <- fastboots.GLA(ex, data=data, clust=clust, boots=30, perm=100, cut=4)
GLAeasy
stopCluster(clust)
closeAllConnections()
```

`fastMLA-methods` *Function to more efficiently screen for gene triplets for those with a high liquid association value.*

Description

Function reduces the processing power and memory needed to calculate modified liquid association (MLA) values for a genome by using a pre-screening method to reduce the candidate pool to triplets likely to have a high MLA value. It does this using matrix algebra to create an approximation to the direct MLA estimate for all possible pairs of X1X2|X3.

Usage

```
fastMLA(data, topn = 2000, nvec = 1, rvalue = 0.5, cut = 4, threads = detectCores())
```

Arguments

`data`	Matrix of numeric data, with columns representing genes and rows representing observations.
`topn`	Number of results to return, ordered from highest |MLA| value descending.
`nvec`	Numeric vector of the gene(s) to use in the X3 position of the X1X2|X3 screening. This should be a numeric vector representing the column #(s) of the gene.
`rvalue`	Tolerance value for LA approximation. Lower values of rvalue will cause a more thorough search, but take longer.
`cut`	Value passed to the GLA function to create buckets (equal to number of buckets+1). Values placing between 15-30 samples per bucket are optimal. Must be a positive integer>1. See GLA.
`threads`	Number of cores to use for multi-threading in correlation calculation (enableWGCNAThreads argument). See WGCNA.

Details

Choosing the number of bins: For example, assume that our data has 100 observations. Since values between 15-30 observations per bin are optimal, good values to choose for cut would be 5-7.

Value

A data frame with 5 variables: the genes in positions X1, X2 and X3; the rhodiff value of the triplet; and the GLA value of the triplet. A more comprehensive discussion of these values is available in the vignette.

Warning

The data matrix must be numeric.

Note

While this is intended to significantly reduce processing time for identifying high MLA values (and in our estimates did so by >90

Author(s)

Tina Gunderson

References

[1] Yen-Yi Ho, Giovanni Parmigiani, Thomas A Louis, and Leslie M Cope. Modeling liquid association. Biometrics, 67(1):133-141, 2011.

See Also

LiquidAssociation, parallel, WGCNA

Examples

```
#to view function code
selectMethod("fastMLA", "matrix")

#
library(yeastCC)
data(spYCCES)
lae <- spYCCES[,-(1:4)]
### get rid of samples with high % NA elements
lae <- lae[apply(is.na(exprs(lae)),1,sum) < ncol(lae)*0.3,]
data <- t(exprs(lae))
data <- data[,1:50]

example <- fastMLA(data=data, topn=25, nvec=1:10, rvalue=1.0, cut=4)
example[1:5,]
closeAllConnections()
```

`mass.CNM-methods` *Function to obtain CNM significance estimates from an object result of fastMLA.*

Description

Function which takes the results of fastMLA and estimates significance based on the beta5 values returned from the conditional normal model estimate. As there are two conditional normal model functions available, full and simple (see CNM.simple, CNM.full, or the fastLiquidAssociation vignette for further details), it attempts to first process all triplets through the full model and for those that do not appear to be well fit by the full model, processes those triplets through the simple model. It returns the triplet(s), fastMLA correlation difference, MLA estimate, beta5 results from the CNM.full or CNM.simple function, and model type, sorted by most significant results according beta5 p-value. Also returns a list of triplets not sensibly fit by either model.

Usage

```
mass.CNM(data, GLA.mat, nback = 100)
```

Arguments

`data`	Matrix of numeric data, with columns representing genes and rows representing observations.
`GLA.mat`	Matrix of triplets to be processed. Intended for use with results of fastMLA function. The X3 position will be defaulted to the value in the third column. see fastMLA.
`nback`	Number of results to return. Results are sorted from most significant p-value.

Value

`top p-values` A data frame with 10 variables: the genes in positions X1, X2 and X3; the rhodiff value of the triplet; the GLA value of the triplet; the estimates for the beta_5 value, standard error, Wald test statistic, and p-value for the Wald test statistic; and the conditional normal model the estimates were obtained from (either F or S, standing for full and simple respectively). A more detailed discussion of these values is available in the vignette.

`bootstrap triplets` A matrix with triplets that did not appear to be well fit by either the full or simple CNM model and thus are recommended to have significance estimated via fastboots.GLA. Its variables are otherwise as specified in the top p-values results description.

Warning

The data matrix must be numeric. Observing the warning message: "In sqrt(diag(object$valpha)) : NaNs produced" is not uncommon in the mass.CNM function as it can be caused by failures to fit either the full or simple CNM model to a triplet.

Author(s)

Tina Gunderson

References

[1] Yen-Yi Ho, Giovanni Parmigiani, Thomas A Louis, and Leslie M Cope. Modeling liquid association. Biometrics, 67(1):133-141, 2011.

See Also

fastMLA, LiquidAssociation

Examples

```
#to view function code
selectMethod("mass.CNM", signature=c("matrix","data.frame"))

#
library(yeastCC)
data(spYCCES)
lae <- spYCCES[,-(1:4)]
### get rid of samples with high % NA elements
lae <- lae[apply(is.na(exprs(lae)),1,sum) < ncol(lae)*0.3,]
data <- t(exprs(lae))
data <- data[,1:50]
example <- fastMLA(data=data, topn=25, nvec=1:10, rvalue=1.0, cut=4)
CNMcalc <- mass.CNM(data=data, GLA.mat=example, nback=10)
CNMcalc
closeAllConnections()
```

---

testmat *Example results from fastMLA requiring all significance methods*

Description

This data.frame contains example results from fastMLA. It is comprised of example triplets for all three methods of calculating significance: the full CNM, the simple CNM, and the direct estimate method. It is intended for use with examples using mass.CNM or fastboots.GLA

Usage

```
testmat
```

Format

A data frame containing 9 observations of 5 variables (three character("X1 or X2","X2 or X1","X3") and two numeric("est.rho1-rho0"and "GLA.value")

Source

Generated based on results obtained using the Spellman et al. dataset using methods extended from Ho et al.

References

Yen-Yi Ho, Giovanni Parmigiani, Thomas A Louis, and Leslie M Cope. Modeling liquid association. Biometrics, 67(1):133-141, 2011.

Paul T Spellman, Gavin Sherlock, Michael Q Zhang, Vishwanath R Iyer, Kirk Anders, Michael B Eisen, Patrick O Brown, David Botstein, and Bruce Futcher. Comprehensive identification of cell cycle regulated genes of the yeast saccharomyces cerevisiae by microarray hybridization. Molecular biology of the cell, 9(12):3273-3297, 1998.

Index

∗ **datasets**
testmat, 8
∗ **methods**
fastboots.GLA-methods, 3
mass.CNM-methods, 7
∗ **package**
fastLiquidAssociation-package, 2

fastboots.GLA *(*fastboots.GLA-methods*)*, 3
fastboots.GLA,ANY,matrix-method *(*fastboots.GLA-methods*)*, 3
fastboots.GLA-class *(*fastboots.GLA-methods*)*, 3
fastboots.GLA-methods, 3
fastLiquidAssociation *(*fastLiquidAssociation-package*)*, 2
fastLiquidAssociation-package, 2
fastMLA *(*fastMLA-methods*)*, 5
fastMLA,matrix-method *(*fastMLA-methods*)*, 5
fastMLA-class *(*fastMLA-methods*)*, 5
fastMLA-methods, 5

LiquidAssociation, *6*

mass.CNM *(*mass.CNM-methods*)*, 7
mass.CNM,matrix,data.frame-method *(*mass.CNM-methods*)*, 7
mass.CNM-class *(*mass.CNM-methods*)*, 7
mass.CNM-methods, 7

parallel, *6*

testmat, 8

WGCNA, *6*