Package 'ADAM'

November 11, 2026

Type Package

Title ADAM: Activity and Diversity Analysis Module

Version 1.28.0

Date 2023-04-10

Maintainer Jose Luiz Rybarczyk Filho <jose.luiz@unesp.br>

Description ADAM is a GSEA R package created to group a set of genes from comparative samples (control versus experiment) belonging to different species according to their respective functions (Gene Ontology and KEGG pathways as default) and show their significance by calculating p-values referring togene diversity and activity. Each group of genes is called GFAG (Group of Functionally Associated Genes).

Depends R(>= 3.5), stats, utils, methods

Imports Rcpp (>= 0.12.18), GO.db (>= 3.6.0), KEGGREST (>= 1.20.2), knitr, pbapply (>= 1.3-4), dplyr (>= 0.7.6), DT (>= 0.4), stringr (>= 1.3.1), SummarizedExperiment (>= 1.10.1)

Suggests testthat, rmarkdown, BiocStyle

LinkingTo Rcpp

SystemRequirements C++11

Encoding UTF-8

RoxygenNote 6.1.0

License GPL (>= 2)

biocViews GeneSetEnrichment, Pathways, KEGG, GeneExpression, Microarray

VignetteBuilder knitr

LazyLoad yes

NeedsCompilation yes

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

git_branch RELEASE_3_23

git_last_commit 6e95948

git_last_commit_date 2026-04-28

Repository Bioconductor 3.23

Date/Publication 2026-11-11

Author André Luiz Molan [aut],
Giordano Bruno Sanches Seco [ctb],
Agnes Takeda [ctb],
Jose Rybarczyk Filho [ctb, cre, ths]

Contents

ADAnalysis 2
ExpressionAedes 4
GFAGAnalysis 5
KeggPathwaysAedes 7

Index **9**

`ADAnalysis` *Group of Functionally Associated Genes (GFAG) partial analysis*

Description

Analysis of functionally associated gene groups, based on gene diversity and activity, for different species according to existing annotation packages or user's personal annotations. ADAnalysis function allows to run a partial analysis, where is calculated just gene diversity and activity of each GFAG with no signicance by bootrstrap, Wilcoxon or Fisher.

Usage

```
ADAnalysis(ComparisonID, ExpressionData, MinGene, MaxGene,
DBSpecies, AnalysisDomain, GeneIdentifier)
```

Arguments

`ComparisonID`	Sample comparisons identification. It must be a vector in which each element corresponds to 2 sample columns from the expression data. The data's sample columns in each element from the vector are comma separated. The first one refers to the control sample, while the second refers to the experiment. This argument must be informed by the user. There is no default value for it.
`ExpressionData`	Gene expression data (microarray or RNA-seq, for example). It must be a SummarizedExperiment object, a data frame or a path for a text file tab separated containing at least 3 columns. First column mandatory corresponds to the gene identification, according to GeneIdentifier argument. Second, third and the other columns correspond to the gene expression values realated to the genes in the first column and each of these columns correspond to a different sample (control versus experiment). This argument must be informed by the user. There is no default value for it.

MinGene	Minimum number of genes per GFAG. It must be a positive integer value different from zero and lower than MaxGene argument. Default is 3.
MaxGene	Maximum number of genes per GFAG. It must be a positive integer value different from zero and greater than MinGene argument. Default is 2000.
DBSpecies	A string corresponding to the name of an OrgDb species gene annotation package: org.Ag.eg.db (Anopheles gambiae), org.At.tair.db (Arabdopsis thaliana), org.Bt.eg.db (Bos taurus), org.Ce.eg.db (Caenorhabditis elegans), org.Cf.eg.db (Canis familiaris), org.Dm.eg.db (Drosophila melanogaster), org.Dr.eg.db (Danio rerio), org.EcK12.eg.db (Escherichia coli K12), org.EcSakai.eg.db (Escherichia coli Sakai), org.Gg.eg.db (Gallus gallus), org.Hs.eg.db (Homo sapiens), org.Mm.eg.db (Mus musculus), org.Mmu.eg.db (Macaca mulatta), org.Pf.plasmo.db (Plasmodium falciparum), org.Pt.eg.db (Pan troglodytes), org.Rn.eg.db (Rattus norvegicus), org.Sc.sgd.db (Saccharomyces cerevisiae), org.Ss.eg.db (Sus scrofa) and org.Xl.eg.db (Xenopus laevis). If there is no package, it's possible for the user to create a personal gene annotation file, tab separated, containing 3 columns: gene, term annotation code and description of the term annotation. So, istead of a string with an OrgDb name, inform a data frame or a path for the file. This argument must be informed by the user. There is no default value for it.
AnalysisDomain	Analysis domain to be considered for building GFAGs, according: gobp (Gene Ontology - Biological Processes), gocc (Gene Ontology - Celular Components), gomf (Gene Ontology - Molecular Functions), kegg (KEGG Pathways) or own (if there is no annotation package - the annotations were defined in a file by the user). This argument must be informed by the user. There is no default value for it.
GeneIdentifier	Gene nomenclature to be used: entrez or tairID for Arabdopsis thaliana, entrez or orfID for Saccharomyces cerevisiae, symbol or orfID for Plasmodium falciparum and symbol or entrez for the other species. If there is no annotation package, just put the gene nomenclature present in the user's personal annotations. This argument must be informed by the user. There is no default value for it.

Details

The genes present in the expression data are grouped by their respective functions according to the domains described by AnalysisDomain argument. The relationship between genes and functions are made based on the species annotation package. If there is no annotation package, a three column file (gene, function and function description) must be provided. For each GFAG, gene diversity and activity in each sample are calculated. As the package always compare two samples (control versus experiment), relative gene diversity and activity for each GFAG are calculated.

Value

Return a list with two elements. The first one refers to a data frame with the GFAGs and their respective genes. The second one is a a list where each position is a data frame presenting the result of the analysis, according to ComparisonID argument.

Author(s)

André Luiz Molan (andre.molan@unesp.br)

References

CASTRO, M. A., RYBARCZYK-FILHO, J. L., et al. Viacomplex: software for landscape analysis of gene expression networks in genomic context. Bioinformatics, Oxford Univ Press, v. 25, n. 11, p. 1468–1469, 2009.

Examples

```
##
#Partial Analysis with Aedes aegypti through ADAnalysis function
##
data(ExpressionAedes)
data(KeggPathwaysAedes)
ResultAnalysis <- ADAnalysis(ComparisonID = c("control1,experiment1"),
ExpressionData = ExpressionAedes, MinGene = 3L, MaxGene = 20L,
DBSpecies = KeggPathwaysAedes, AnalysisDomain = "own",
GeneIdentifier = "geneStableID")
## Not run:
head(ResultAnalysis[[1]]) #Relation between genes and functions
head(ResultAnalysis[[2]][1]) #Result comparison 1

## End(Not run)
```

ExpressionAedes *Aedes aegypti RNA-seq data expression*

Description

A sample fragment of gene expression from an RNA-seq experiment of Aedes aegypti mosquito.

Format

A data frame with 34 rows and 5 variables

References

AKBARI, O. S. et al. The developmental transcriptome of the mosquito aedes aegypti, an invasive species and major arbovirus vector. G3: Genes— Genomes— Genetics, Genetics Society of America, v. 3, n. 9, p. 1493–1509, 2013.

Examples

```
data(ExpressionAedes)
```

GFAGAnalysis *Group of Functionally Associated Genes (GFAG) complete analysis*

Description

Analysis of functionally associated gene groups, based on gene diversity and activity, for different species according to existing annotation packages or user's personal annotations. GFAGAnalysis function allows to run a complete analysis, using all available arguments.

Usage

```
GFAGAnalysis(ComparisonID, ExpressionData, MinGene, MaxGene,
SeedNumber, BootstrapNumber, PCorrection, DBSpecies, PCorrectionMethod,
WilcoxonTest, FisherTest, AnalysisDomain, GeneIdentifier)
```

Arguments

ComparisonID	Sample comparisons identification. It must be a vector in which each element corresponds to 2 sample columns from the expression data. The data's sample columns in each element from the vector are comma separated. The first one refers to the control sample, while the second refers to the experiment. This argument must be informed by the user. There is no default value for it.
ExpressionData	Gene expression data (microarray or RNA-seq, for example). It must be a SummarizedExperiment object, a data frame or a path for a text file tab separated containing at least 3 columns. First column mandatory corresponds to the gene identification, according to GeneIdentifier argument. Second, third and the other columns correspond to the gene expression values realated to the genes in the first column and each of these columns correspond to a different sample (control versus experiment). This argument must be informed by the user. There is no default value for it.
MinGene	Minimum number of genes per GFAG. It must be a positive integer value different from zero and lower than MaxGene argument. Default is 3.
MaxGene	Maximum number of genes per GFAG. It must be a positive integer value different from zero and greater than MinGene argument. Default is 2000.
SeedNumber	Seed for bootstrap. A numeric positive value used as seed for random number generating by bootstrap function. Default is 1049.
BootstrapNumber	Number of bootstraps. A numeric value greater than zero, used by bootstrap function generates p-values for each GFAG. Default is 1000.
PCorrection	Cutoff for p-value correction. A numeric value between 0 and 1. Default is 0.05.
DBSpecies	A string corresponding to the name of an OrgDb species gene annotation package: org.Ag.eg.db (Anopheles gambiae), org.At.tair.db (Arabdopsis thaliana), org.Bt.eg.db (Bos taurus), org.Ce.eg.db (Caenorhabditis elegans), org.Cf.eg.db (Canis familiaris), org.Dm.eg.db (Drosophila melanogaster), org.Dr.eg.db (Danio rerio), org.EcK12.eg.db (Escherichia coli K12), org.EcSakai.eg.db (Escherichia

coli Sakai), org.Gg.eg.db (Gallus gallus), org.Hs.eg.db (Homo sapiens), org.Mm.eg.db (Mus musculus), org.Mmu.eg.db (Macaca mulatta), org.Pf.plasmo.db (Plasmodium falciparum), org.Pt.eg.db (Pan troglodytes), org.Rn.eg.db (Rattus norvegicus), org.Sc.sgd.db (Saccharomyces cerevisiae), org.Ss.eg.db (Sus scrofa) and org.Xl.eg.db (Xenopus laevis). If there is no package, it's possible for the user to create a personal gene annotation file, tab separated, containing 3 columns: gene, term annotation code and description of the term annotation. So, istead of a string with an OrgDb name, inform a data frame or a path for the file. This argument must be informed by the user. There is no default value for it.

`PCorrectionMethod`
: Method p-value correction: holm, hochberg, hommel, bonferroni, bh, by or fdr. Default is "fdr".

`WilcoxonTest`
: A logical value indicating whether or not to perform Wilcoxon test (TRUE for performing and FALSE for not performing). Default is FALSE.

`FisherTest`
: A logical value indicating whether or not to perform Fisher's exact test (TRUE for performing and FALSE for not performing). Default is FALSE.

`AnalysisDomain`
: Analysis domain to be considered for building GFAGs, according: gobp (Gene Ontology - Biological Processes), gocc (Gene Ontology - Celular Components), gomf (Gene Ontology - Molecular Functions), kegg (KEGG Pathways) or own (if there is no annotation package - the annotations were defined in a file by the user). This argument must be informed by the user. There is no default value for it.

`GeneIdentifier`
: Gene nomenclature to be used: entrez or tairID for Arabdopsis thaliana, entrez or orfID for Saccharomyces cerevisiae, symbol or orfID for Plasmodium falciparum and symbol or entrez for the other species. If there is no annotation package, just put the gene nomenclature present in the user's personal annotations. This argument must be informed by the user. There is no default value for it.

Details

The genes present in the expression data are grouped by their respective functions according to the domains described by AnalysisDomain argument. The relationship between genes and functions are made based on the species annotation package. If there is no annotation package, a three column file (gene, function and function description) must be provided. For each GFAG, gene diversity and activity in each sample are calculated. As the package always compare two samples (control versus experiment), relative gene diversity and activity for each GFAG are calculated. Using bootstrap method, for each GFAG, according to relative gene diversity and activity, two p-values are calculated. The p-values are then corrected, according to the correction method defined by PCorrectionMethod argument, generating a q-value. The significative GFAGs will be those whoose q-value stay under the cutoff set by PCorrection argument. Optionally, it's possible to run Wilcoxon test and/or Fisher's exact test. These tests also provide a corrected p-value, and siginificative groups can be seen through them.

Value

Return a list with two elements. The first one refers to a data frame with the GFAGs and their respective genes. The second one is a a list where each position is a data frame presenting the result

of the analysis, according to ComparisonID argument.

Author(s)

André Luiz Molan (andre.molan@unesp.br)

References

CASTRO, M. A., RYBARCZYK-FILHO, J. L., et al. Viacomplex: software for landscape analysis of gene expression networks in genomic context. Bioinformatics, Oxford Univ Press, v. 25, n. 11, p. 1468–1469, 2009.

Examples

```
##
#Complete Analysis with Aedes aetypti through GFAGAnalysis function
##
data(ExpressionAedes)
data(KeggPathwaysAedes)
ResultAnalysis <- GFAGAnalysis(ComparisonID = c("control1,experiment1",
"control2,experiment2"), ExpressionData = ExpressionAedes, MinGene = 3L,
MaxGene = 20L, SeedNumber = 1049, BootstrapNumber = 1000L,
PCorrection = 0.05, DBSpecies = KeggPathwaysAedes,
PCorrectionMethod = "fdr", WilcoxonTest = TRUE, FisherTest = TRUE,
AnalysisDomain = "own", GeneIdentifier = "gene")
## Not run:
head(ResultAnalysis[[1]]) #Relation between genes and functions
head(ResultAnalysis[[2]][1]) #Result comparison 1
head(ResultAnalysis[[2]][2]) #Result comparison 2

## End(Not run)
```

KeggPathwaysAedes *Relation between Aedes aegypti genes and KEGG pathways as ADAM input*

Description

A relation between the genes in the ExpressionAedes data and their respective KEGG pathways (GFAGs).

Format

A data frame with 200 rows and 2 variables

References

Molan, A. L. 2018. “Construction of a Tool for Multispecies Genic Functional Enrichment Analysis Among Comparative Samples.” Master’s thesis Institute of Biosciences of Botucatu – Univ. Estadual Paulista. http://hdl.handle.net/11449/157105.

Examples

```
data(KeggPathwaysAedes)
```

Index

ADAnalysis, 2

ExpressionAedes, 4

GFAGAnalysis, 5

KeggPathwaysAedes, 7