Package 'agalmar'

November 9, 2016

Title A Collection of Tools for Analyzing Agalma Expression Projects

Version 0.0.0.9000

Date 2015-06-01

Description A Collection of tools for analyzing agalma (https://bitbucket.org/caseywdunn/agalma) expression projects.

Depends R (>= 3.2.2)

License MIT + file LICENSE

LazyData true

Collate 'utility_functions.R'
'agalmar.R'
'data.R'

Imports hutan,
edgeR,
ape (>= 3.3),
ggtree,
DESeq2,
plyr,
digest,
methods,
graphics

Suggests testthat,
roxygen2,
devtools,
knitr

VignetteBuilder knitr

RoxygenNote 5.0.1

R topics documented:

agalmar	*agalmar: A Collection of Tools for Analyzing Agalma Expression Projects*

Description

The agalmar package contains a collection of tools for analyzing expression and phylogenetic data exported from Agalma. See https://bitbucket.org/caseywdunn/agalma

apply_normalizations	*Takes a DGEList object, and returns a matrix of normalized counts*

Description

Takes a DGEList object, and returns a matrix of normalized counts

Usage

```
apply_normalizations(dge)
```

Arguments

dge	edgeR DGEList object, to which normalizations have been applied

Value

matrix of normalized counts

create_DESeq2 *Creates a DESeqDataSet object*

Description

Creates a DESeqDataSet object

Usage

```
create_DESeq2(object, design)
```

Arguments

object	An Expression object
design	a formula that explains the project design

Value

A DESeqDataSet object

create_DESeq2,Expression,formula-method
Creates a DESeqDataSet object

Description

Creates a DESeqDataSet object

Usage

```
## S4 method for signature 'Expression,formula'
create_DESeq2(object, design)
```

Arguments

object	An Expression object
design	a formula that explains the project design

Value

A DESeqDataSet object

---

dataframe_to_node_labels
                    Converts data frame of node annotations to NHX node labels

---

Description

Converts data frame of node annotations to NHX node labels

Usage

```
dataframe_to_node_labels(df)
```

Arguments

df   Data frame return character A vector of node label names

---

decompose_orthologs *Decomposes a gene tree into a list of subtrees that have no duplication events. Assumes notung style node annotations.*

Description

Decomposes a gene tree into a list of subtrees that have no duplication events. Assumes notung style node annotations.

Usage

```
decompose_orthologs(nhx)
```

Arguments

nhx   The tree, as a ggtree nhx object

Value

The subtrees as a list of ape::phylo object

DGEList-class   *"DGEList" class*

Description

"DGEList" class

drop.tip.nhx	*Drops tips from an NHX tree, similar to how ape::drop.tip drops tips from a phylo object*

Description

Drops tips from an NHX tree, similar to how ape::drop.tip drops tips from a phylo object

Usage

```
drop.tip.nhx(nhx, tip, test = FALSE)
```

Arguments

nhx	A ggtree nhx object
tip	A vector of mode numeric or character specifying the tips to delete
test	A logical to designate whether to rotate some internal nodes for more robust testing

Value

A ggtree::nhx object

Expression	*Construct an Expression object from list of experiment data and meta-data provided by Agalma.*

Description

Construct an Expression object from list of experiment data and metadata provided by Agalma.

Usage

```
Expression(data_list)
```

Arguments

data_list	A list containing the expression data

Value

An Expression object

Expression-class   *An S4 class to represent gene expression data for multiple samples*
                            for a given species. Assumes that all expession data are derived from
                            mapping to the same reference. Applies to data for g genes across s
                            samples (ie, sequenced libraries).

Description

The fields that apply to the s samples correspond to those

Slots

species The species

edgeR An edgeR DGEList object holding data for all samples. Expression matrix of dimension g,s.

lengths Summary of length, in bp, of transctipts for each gene. Length g.

individual Factors indicating which individual each sample is from. Length s.

treatment Factors indicating which treatment applies to each sample. Length s.

id Factors indicating the unique sequencing run id of each sample, eg HWI-ST625-75-D0PBDACXX-
    6-ATCACG. Length s.

library_id Factors indicating the unique library id of each sample, eg FEG365. Length s.

sample_prep Sample prep strategy. Length s.

genome_type Character indicating genome type, eg nuclear. See agalma documentation. Length
    g.

molecule_type Character indicating encoded molecule type, eg protein. See agalma documenta-
    tion. Length g.

blast_hit Blast hit. Length g.

rRNA Fraction of reads in sample that are rRNA. Length s.

protein Fraction of reads in sample that are protein coding. Length s.

x The counts matrix

---

get_lane   *Get lane from Illumina fastq sequence header*

Description

Get lane from Illumina fastq sequence header

Usage

```
get_lane(header)
```

Arguments

header   character strings

Value

lane character strings

---

get_run *Get run from Illumina fastq sequence header*

Description

Get run from Illumina fastq sequence header

Usage

```
get_run(header)
```

Arguments

header   character strings

Value

run character strings

---

get_tip_info *Parse species and sequence id from a phy*

Description

Parse species and sequence id from a phy

Usage

```
get_tip_info(phy)
```

Arguments

phy   The tree, as an ape phylo object. Tip names must have 'species@id' format.

Value

Dataframe with one row per tip, a species column, and an id column

---

has_species   *Returns true if all species are in the tree tips*

Description

Returns true if all species are in the tree tips

Usage

```
has_species(phy, species)
```

Arguments

phy   The tree, as an ape phylo object

species   Vector of species names

Value

logical

---

janedoe   *Expression data for different tissues from an animal*

Description

Expression data for different tissues from an animal

Usage

```
janedoe
```

Format

A list parsed from a json that was exported from agalma with https://bitbucket.org/caseywdunn/agalma/src/master/scripts/
export-expression

nhx_label_to_list	*Converts a NHX node label string to a list of named values*

Description

Converts a NHX node label string to a list of named values

Usage

```
nhx_label_to_list(label)
```

Arguments

label	Character string

Value

list A named list of values

Examples

```
nhx_label_to_list("[&&NHX:Ev=S:S=58:ND=0]")
```

node_support	*Parse support from node name*

Description

Parse support from node name

Usage

```
node_support(node_name)
```

Arguments

node_name	A character string of format '83:N', where '83' is node support and N indicates whether the node is a duplication or not

Value

Numeric indicating node support

parse_gene_tree   *Parses text to a gene tree*

Description

Parses text to a gene tree

Usage

```
parse_gene_tree(tree_text)
```

Arguments

tree_text   Text representation of a tree in nhx format with notung or phyldog fields.

Value

phy The tree, as an ape phylo object

plot_matrix   *Plots a matrix*

Description

Plots a matrix

Usage

```
plot_matrix(m, ...)
```

Arguments

m   The matrix

...   additional arguments for image

species   *Get species*

Description

Get species

Usage

```
species(object)
```

Arguments

object   An Expression object

Value

Character string with species

species,Expression-method
                              Get the species

Description

Get the species

Usage

```
## S4 method for signature 'Expression'
species(object)
```

Arguments

object   An Expression object

Value

Species name

---

summarize_edges   *Create a data frame with summary statistics for edges in a phyldog*
                         NHX tree

Description

Create a data frame with summary statistics for edges in a phyldog NHX tree

Usage

```
summarize_edges(nhx)
```

Arguments

nhx   A ggtree nhx object

Value

A data frame of edge summary statistics

---

summarize_libraries   *Summarize expression libraries*

Description

Summarize expression libraries

Usage

```
summarize_libraries(object)
```

Arguments

object   An Expression object

Value

A data frame that summarizes each library with expression data in the Expression object

summarize_libraries,Expression-method

Summarize expression libraries

Description

Summarize expression libraries

Usage

```
## S4 method for signature 'Expression'
summarize_libraries(object)
```

Arguments

object   An Expression object

Value

A data frame that summarizes each library with expression data in the Expression object

---

summarize_nodes   *Create a data frame with summary statistics for nodes in a phyldog NHX tree*

Description

Create a data frame with summary statistics for nodes in a phyldog NHX tree

Usage

```
summarize_nodes(nhx)
```

Arguments

nhx   A ggtree nhx object

Value

A data frame of node summary statistics

summarize_reference   *Summarize reference sequences*

Description

Summarize reference sequences

Usage

```
summarize_reference(object)
```

Arguments

object   An Expression object

Value

A data frame that summarizes the reference sequences in the Expression object

summarize_reference,Expression-method
                           Summarize reference sequences

Description

Summarize reference sequences

Usage

```
## S4 method for signature 'Expression'
summarize_reference(object)
```

Arguments

object   An Expression object

Value

A data frame that summarizes the reference sequences in the Expression object

---

summary_libraries *Create a data frame with summary statistics for expression libraries*

Description

Create a data frame with summary statistics for expression libraries

Usage

```
summary_libraries(e)
```

Arguments

e   A list of Expression objects

Value

A data frame of summary statistics

---

summary_references *Create a data frame with summary statistics for reference sequences*

Description

Create a data frame with summary statistics for reference sequences

Usage

```
summary_references(e)
```

Arguments

e   A list of Expression objects

Value

A data frame of summary statistics

Index