Package 'BACA'

May 27, 2015

Type Package

Title Bubble Chart to Compare Biological Annotations by using DAVID

Version 1.3

Date 2015-05-27

Author Vittorio Fortino and Dario Greco

Maintainer Vittorio Fortino <vittorio.fortino@ttl.fi>

Description R-based graphical tool to concisely visualise and compare biological annotations queried from the DAVID web service. It provides R functions to perform enrichment analysis (via DAVID - http://david.abcc.ncifcrf.gov) on several gene lists at once, and then visualizing all the results in one generated figure that allows R users to compare the annotations found for each list.

License GPL (>= 2)

Depends R (>= 3.0.0)

Imports RDAVIDWebService, ggplot2, rJava

VignetteBuilder knitr

Suggests knitr

NeedsCompilation no

Repository CRAN

Date/Publication 2015-05-27 08:55:17

R topics documented:

BACA–package   *A graphical tool to visualize and compare functional annotations retrieved from DAVID knowledgebase.*

Description

DAVID Bioinformatics Resources (DAVID) at http://david.abcc.ncifcrf.gov is the most popular tool in functional annotation and enrichment analysis. It provides an integrated biological knowledgebase and tools to systematically extract relevant biological terms (e.g., GO terms, KEGG pathways) associated with a given gene list. After submitting a gene list, DAVID annotation tool finds the most enriched annotations and presents them in a table format. This table contains many different types of data, such as text, numbers, bars and hyperlinks, that can be hard tor read and compare when multiple enrichment results are available. The BACA package tries to address these issues by providing a novel R-based graphical tool to concisely visualize DAVID annotations and show how they change across different experimental conditions. This R package uses some functions available in the R package RDAVIDWebService available at http://www.bioconductor.org/packages/release/bioc/html/RDAVIDWebService.html.

Details

Package:	BACA
Type:	Package
Version:	1.0
Date:	2014-06-19
License:	GPL (>=2)

The package BACA provides three different functions: DAVIDsearch, BBplot and Jplot. DAVIDsearch: to call the Functional Annotation Tool of DAVID. BBplot: to build a grid where each row represents an enriched annotation found by DAVID and each column the condition/treatment where that annotation was highlighted. While, each cell reports a bubble indicating the number of genes enriching the correpsonding annotation and the state of these genes in terms of down- and up-regulation (default setting: green = "down" - red = "up"). Jplot: to make a table/matrix with colored boxes. The color of each box indicated the Jaccard index based similarity score calculated between the genes sets associated with two different functional annotations.

Author(s)

Vittorio Fortino <vittorio.fortino@ttl.fi>

Examples

```
data(result.kegg)

bbplot.kegg <- BBplot(result.kegg, max.pval = 0.05, min.ngenes = 10,
                name.com = c("Cond.1_12h","Cond.1_24h","Cond.2_12h","Cond.2_24h"),
```

```
                          labels = c("down", "up"), colors = c("#009E73", "red"),
                          title = "BBplot - KEGG", print.term = "description")
        bbplot.kegg
        ##ggsave("KEGG_terms.tiff", width=6, height=4, scale=2, dpi=200)

        jplot.kegg <- Jplot(result.kegg[[4]], result.kegg[[2]], max.pval = 0.05, min.ngenes = 10)

        jplot.kegg
        ##ggsave("j_plot.tiff", width=6, height=4, scale=2, dpi=200)
```

BBplot	*To compare enriched functional annotations found by using DAVID webservice.*

Description

It uses a list of DAVIDFunctionalAnnotationChart objects to build a chart that shows how the functional annotations found by DAVID have changed across different experimental conditions. BBplot function can work with any kind of gene list and not only with up/down regulated gene lists.

Usage

```
BBplot(list.david.obj, max.pval = 0.01, min.ngenes = 5, max.ngenes = 500,
        adj.method = "Benjamini", title = "BBplot", name.com = "***",
      labels = c("down", "up"), colors = c("#009E73", "red"), print.term = "full")
```

Arguments

list.david.obj	(Required). List of DAVIDFunctionalAnnotationChart objects.
max.pval	(Optional). Numeric with the p-value that must be present in order to select the most significant enriched annotations. Default values is 0.01.
min.ngenes	(Optional). Integer with the minimum number of genes (greater or equal) requested for each enriched annotations. Default values is 5.
max.ngenes	(Optional). Integer with the maximum number of genes (greater or equal) requested for each enriched annotations. Default values is 5.
adj.method	(Optional). Character with the name of the adjustment method. Default valuse is "Benjamini". Methods that are available: "Benjamini", "Bonferroni" or "FDR". It can be empty "".
title	(Optional). Character with the name of the buble plot. Default valuse is "Bubble plot".
name.com	(Optional). Character vector indicating the experimental conditions.
labels	(Optional). Character vector to specify the name used to indicate the down- and up-regulated gene lists. The default value is c("down", "up").
colors	(Optional). Character vector to specify the colors used to distinguish down- and up-regulated gene lists. The default value is c("#009E73", "red").
print.term	(Optional). Character vector to specify the term used to indicate the annotation: 'full', 'name' or 'description'. The default value is "full".

Details

The list of DAVIDFunctionalAnnotationChart objects should be generated by using the function DAVIDsearch.

Value

Return a ggplot2 graph.

Note

For each condition, provide two DAVIDFunctionalAnnotationChart objects: one for the up-regulated genes and other for the down-regulated genes.

Author(s)

Vittorio Fortino

See Also

DAVIDsearch

Examples

```
data(result.kegg)

bbplot.kegg <- BBplot(result.kegg, max.pval = 0.05, min.ngenes = 10,
                name.com = c("Cond.1_12h","Cond.1_24h","Cond.2_12h","Cond.2_24h"),
                labels = c("down", "up"), colors = c("#009E73", "red"),
                title = "BBplot - KEGG", print.term = "full")
bbplot.kegg
##ggsave("KEGG_terms.tiff", width=6, height=4, scale=2, dpi=200)
```

DAVIDsearch	*Get functional annotation charts of gene lists using DAVID.*

Description

Uploads multiple gene lists to DAVID, then perform an automated enrichment analysis based on a given database/resource (i.e., GO terms, KEGG, etc.) for each gene list. Returns DAVID-FunctionalAnnotationChart objects, one for each gene list. Requires a DAVID user id first http://david.abcc.ncifcrf.gov/webservice/register.htm. Among the optional input paramters, can use the easeScore (or P-value) to do thresholding on results. However, we suggest to get DAVID to return all possible annotations despite non-significant P-values (easeScore = 1) and perform your own thresholding when using the plot functions.

Usage

```
DAVIDsearch(gene.lists, david.user, idType = "AFFYMETRIX_3PRIME_IVT_ID",
listType = "Gene", easeScore = 1, annotation = "KEGG_PATHWAY",
species = NA)
```

Arguments

gene.lists   (Required). Lists of character vectors representing the gene lists. To compare
                  the enrichment of differential genes specify for each list the up- and down- reg-
                  ulated genes, separately.

david.user   (Required). Character vector to specify the email of a given registered DAVID;
                  it is necessary to use the DAVID web service.

idType   (Optional). Character with the type of submitted ids. Default value is "AFFYMETRIX_3PRIME_IVT_ID

listType   (Optional). Character with the type of list (Gene, Background). Default value is
                  "Gene".

easeScore   (Optional). Numeric with the EASE score. Default value is 1 (see DAVID Help
                  page).

annotation   (Optional). Character vector with the category name to use in the functional
                  annotation analysis. Default value is "KEGG_PATHWAY".

species   (Optional). Character vector with the species to use. Default value is NA. This
                  means that DAVID will use the specie found for the submitted gene list.

Details

The following information are printed out during the the querying process: For the list "name.list"
you have: - Number of genes loaded = #. - Number of genes mapped/annotated in DAVID <in-
David> = #. - Number of unmapped genes <unmappedIds> = #. - Species involved.

Value

List of DAVIDFunctionalAnnotationChart objects, one for each specified list of genes.

Author(s)

Vittorio Fortino

See Also

BBplot Jplot

Examples

```
#data(gene.lists.ex)
#result.kegg <- DAVIDsearch(gene.lists.ex, david.user = "vittorio.fortino@ttl.fi")
```

`gene.lists.ex`	*Eight gene lists*

Description

This data contains artifical up- and down-regulated gene lists corresponding to two time points of two different experimental conditions.

Usage

```
data(gene.lists.ex)
```

Format

A list of eight character vectors.

Value

Return a list of eight character vectors representing artifical up- and down-regulated gene lists.

Jplot	*To compute and visualize Jaccard similarity score between two enrichment analysis results.*

Description

Comparing two enrichment analysis results, it reveals the relationships existings between annotations found by using two different gene lists. The Jplot() function takes in input two different DAVIDFunctionalAnnotationChart objects (1 and 2) and provides in output a table/matrix with colored boxes. Each box reports the Jaccard index-based similarity score computed between the gene sets enriching two functional annotations.

Usage

```
Jplot(david.obj.1, david.obj.2,  max.pval = 0.01, min.ngenes = 5, title = "Jplot",
        print.term ="full")
```

Arguments

`david.obj.1`	(Required). First DAVIDFunctionalAnnotationChart objects.
`david.obj.2`	(Required). Second DAVIDFunctionalAnnotationChart objects.
`max.pval`	(Optional). Numeric with the p-value that must be present in order to select the most significant enriched annotations. Default values is 0.01.
`min.ngenes`	(Optional). Integer with the minimum number of genes (greater or equal) that must be present in an enriched annotations. Default values is 5.

title	(Optional). Character with the name of the plot. Default valuse is "Bubble plot".
print.term	(Optional). Character vector to specify the term used to indicate the annotation: 'full', 'name' or 'description'. The default value is "full".

Value

Return a ggplot2 graph.

Note

The two DAVIDFunctionalAnnotationChart objects should be associated to two different enrichment analysis results.

Author(s)

Vittorio Fortino

See Also

[DAVIDsearch](#) [BBplot](#)

Examples

```
data(result.kegg)
jplot.kegg <- Jplot(result.kegg[[4]], result.kegg[[2]], max.pval = 0.05, min.ngenes = 10)
jplot.kegg
```

---

result.kegg *David search results*

Description

This data contains an example of list of DAVIDFunctionalAnnotationChart objects computed by the DAVIDsearch function.

Usage

```
data(result.kegg)
```

Format

A list of of DAVIDFunctionalAnnotationChart objects.

Value

Return a list of of DAVIDFunctionalAnnotationChart objects.

Index