# Package 'DEGRE'

November 2, 2022

**Type** Package

**Title** Inferring Differentially Expressed Genes using Generalized
Linear Mixed Models

**Version** 0.2.0

**Description** Genes that are differentially expressed between two or more experimental conditions can be detected in RNA-Seq. A high biological variability may impact the discovery of these genes once it may be divergent between the fixed effects. However, this variability can be covered by the random effects. 'DEGRE' was designed to identify the differentially expressed genes considering fixed and random effects on individuals. These effects are identified earlier in the experimental design matrix. 'DEGRE' has the implementation of preprocessing procedures to clean the near zero gene reads in the count matrix, normalize by 'RLE' published in the 'DESeq2' package, 'Love et al. (2014)' <doi:10.1186/s13059-014-0550-8> and it fits a regression for each gene using the Generalized Linear Mixed Model with the negative binomial distribution, followed by a Wald test to assess the regression coefficients.

**License** Artistic-2.0

**Encoding** UTF-8

**Depends** R (>= 4.0)

**Imports** utils, parglm, glmmTMB, foreach, tibble, ggplot2, ggpubr,
ggrepel, car, dplyr

**RoxygenNote** 7.2.1

**Suggests** testthat (>= 3.0.0)

**Config/testthat/edition** 3

**NeedsCompilation** no

**Author** Douglas Terra Machado [aut, cre]
(<https://orcid.org/0000-0002-6580-7628>),
Otávio José Bernardes Brustolini [aut]
(<https://orcid.org/0000-0001-8132-9753>),
Yasmmin Côrtes Martins [aut] (<https://orcid.org/0000-0002-6830-1948>),
Marco Antonio Grivet Mattoso Maia [aut]
(<https://orcid.org/0000-0001-5130-095X>),
Ana Tereza Ribeiro de Vasconcelos [aut]
(<https://orcid.org/0000-0002-4632-2086>)

## R topics documented:

| BarGraphDEGRE | *Create a bar plot showing the number of downregulated and upregulated genes.* |
|---|---|

### Description

Create a bar plot showing the number of downregulated and upregulated genes.

### Usage

```
BarGraphDEGRE(
  results,
  log2FC_cutoff = 1,
  downregulated_color = "coral2",
  upregulated_color = "cornflowerblue",
  xlab = "Regulation",
  ylab = "Number of genes",
  font.x = 10,
  font.y = 10,
  font.tickslab = 10,
  legend_position = "right",
  legend.title = "Regulation"
)
```

### Arguments

results         a data.frame object. It receives the output of the DEGRE function, filtered or
                not, as input.

log2FC_cutoff   it stores the cutoff of the log2FoldChange. The default is 1.

downregulated_color
                the bar color related to the number of downregulated genes. The default is
                "coral2".

upregulated_color

     the bar color related to the number of upregulated genes. The default is "cornflowerblue".

| | |
|---|---|
| xlab | the x lab text. The default is "Regulation". |
| ylab | the y lab text. The default is "Number of genes". |
| font.x | the font size of the x axis. The default is 10. |
| font.y | the font size of the y axis. The default is 10. |
| font.tickslab | the font size of the ticks lab. The default is 10. |

legend_position

     you need to specify here the position of the legend. The default is "right".

| | |
|---|---|
| legend.title | the title of the legend. The default is "Regulation". |

## Value

No return value, called for side effects

## Examples

```
# Reading a CSV file for an example:
dir <- system.file("extdata", package = "DEGRE")
results_DEGRE_example <- read.csv(file.path(dir,"results_DEGRE_example.csv"))
# Running the BarGraphDEGRE function
BarGraphDEGRE(results = results_DEGRE_example,
          log2FC_cutoff = 1,
          downregulated_color = "coral2",
          upregulated_color = "cornflowerblue",
          xlab = "Regulation",
          ylab = "Number of genes",
          font.x = 10,
          font.y = 10,
          font.tickslab = 10,
          legend_position = "right",
          legend.title = "Regulation")
```

---

DEGRE           *Infer the Differentially Expressed Genes*

---

## Description

The DEGRE function acts inferring the differentially expressed genes fitting them using generalized linear mixed models and then applies the Wald test to the regression coefficients. It offers Benjamini-Hochberg or Bonferroni P-values techniques to adjust the P-values.

## Usage

```
DEGRE(count_matrix, p_value_adjustment = "BH", formula, design_matrix)
```

## Arguments

count_matrix      a data.frame object. It receives the raw matrix as input.

p_value_adjustment

     a data.frame object. It receives the experimental design matrix. The sample names must be identified in the first column. This matrix can also have more columns with information for the fixed and the random effects for the samples.

formula      it receives fixed and random effects descriptions.

design_matrix      All the P-values computed must be corrected and the DEGRE package offers two possibilities: "BH" (Benjamini-Hochberg) correction (default) and "BON" (Bonferroni) correction.

## Value

dataframe object

## Examples

```
# Reading the count matrix and the design matrix for an example:
dir <- system.file("extdata", package = "DEGRE")
tab <- read.csv(file.path(dir,"count_matrix_for_example.csv"))
des <- read.csv(file.path(dir,"design_matrix_for_example.csv"))
# Running DEGRE function:
results <- DEGRE(count_matrix = tab,
                 p_value_adjustment = "BH",
                 design_matrix = des,
                 formula = "condition + (1|sex)")
```

---

DEGRE_package          *DEGRE package*

---

## Description

Genes that are differentially expressed (DEGs) between two or more experimental conditions can be found in investigations using RNA-Seq data. The gene reads matrix is a count matrix that is used to infer DEGs, and fixed effects are often utilized in this inference. A high biological variability may impact the discovery of DEGs once it may be divergent between the fixed effects. This variability can be covered by the random effects. The DEGRE R package was designed to identify DEGs in RNA-Seq experiments considering fixed and the random effects on individuals. These effects are identified earlier in the experimental design matrix. DEGRE has the implementation of preprocessing procedures to clean the count matrix of the gene reads, and it fits a regression for each gene using the Generalized Linear Mixed Model with the negative binomial distribution,followed by a Wald test to test to assess the regression coefficients.

## Author(s)

Douglas Terra Machado <dougterra@gmail.com>, Otávio José Bernardes Brustolini, Yasmmin Côrtes Martins, Marco Antonio Grivet Mattoso Maia, Ana Tereza Ribeiro de Vasconcelos

---

| | |
|---|---|
| VolcanoDEGRE | *Create a volcano plot to visualize the proportion of downregulated and upregulated genes by applying a log2FC cutoff.* |

---

## Description

Create a volcano plot to visualize the proportion of downregulated and upregulated genes by applying a log2FC cutoff.

## Usage

```
VolcanoDEGRE(
  results,
  log2FC_cutoff = 1,
  padj = 0.05,
  font.x = 10,
  font.y = 10,
  font.tickslab = 10,
  downregulated_color = "coral2",
  upregulated_color = "cornflowerblue",
  xlab = "log2Foldchange",
  ylab = "-log10(P-value)",
  legend_position = "right",
  legend.title = "Regulation"
)
```

## Arguments

| | |
|---|---|
| results | a data.frame object. It receives the output of the DEGRE function, filtered or not, as input. |
| log2FC_cutoff | it stores the cutoff of the log2FoldChange. The default is 1. |
| padj | it stores the cutoff of the P-adjusted value (Q-value). The default is 0.05. |
| font.x | the font size of the x axis. The default is 10. |
| font.y | the font size of the y axis. The default is 10. |
| font.tickslab | the font size of the ticks lab. The default is 10. |
| downregulated_color | |
| | the colors of the downregulated genes. The default is "coral2". |
| upregulated_color | |
| | the colors of the upregulated genes. The default is "cornflowerblue". |
| xlab | the x lab text. The default is "log2Foldchange". |

ylab                  the y lab text. The default is "-log10(P-value)".

legend_position

                      you need to specify here the position of the legend. The default is "right".

legend.title     the title of the legend. The default is "Regulation".

## Value

No return value, called for side effects

## Examples

```
dir <- system.file("extdata", package = "DEGRE")
results_DEGRE_example <- read.csv(file.path(dir,"results_DEGRE_example.csv"))
# Running the VolcanoDEGRE function
VolcanoDEGRE(results = results_DEGRE_example,
          log2FC_cutoff = 1,
          padj = 0.05,
          font.x = 10,
          font.y = 10,
          font.tickslab = 10,
          downregulated_color = "coral2",
          upregulated_color = "cornflowerblue",
          xlab = "log2Foldchange",
          ylab = "-log10(P-value)",
          legend_position = "right",
          legend.title = "Regulation")
```

# Index