# Package 'DEsubs'

December 16, 2024

**Version** 1.32.0

**Date** 2017-07-23

**Title** DEsubs: an R package for flexible identification of differentially expressed subpathways using RNA-seq expression experiments

**Author** Aristidis G. Vrahatis and Panos Balomenos

**Maintainer**
Aristidis G. Vrahatis <agvrahatis@upatras.gr>, Panos Balomenos <balomenos@upatras.gr>

**Description** DEsubs is a network-based systems biology package that extracts disease-perturbed subpathways within a pathway network as recorded by RNA-seq experiments. It contains an extensive and customizable framework covering a broad range of operation modes at all stages of the subpathway analysis, enabling a case-specific approach. The operation modes refer to the pathway network construction and processing, the subpathway extraction, visualization and enrichment analysis with regard to various biological and pharmacological features. Its capabilities render it a tool-guide for both the modeler and experimentalist for the identification of more robust systems-level biomarkers for complex diseases.

**Depends** R (>= 3.3), locfit

**License** GPL-3

**NeedsCompilation** no

**LazyLoad** yes

**Imports** graph, igraph, RBGL, circlize, limma, edgeR, EBSeq, NBPSeq, stats, grDevices, graphics, pheatmap, utils, ggplot2, Matrix, jsonlite, tools, DESeq2, methods

**Suggests** RUnit, BiocGenerics, knitr, rmarkdown

**VignetteBuilder** knitr

**biocViews** SystemsBiology, GraphAndNetwork, Pathways, KEGG, GeneExpression, NetworkEnrichment, Network, RNASeq, DifferentialExpression, Normalization, ImmunoOncology

**git_url** https://git.bioconductor.org/packages/DEsubs

**git_branch** RELEASE_3_20

**git_last_commit** 5a48f9f

**git_last_commit_date** 2024-10-29
**Repository** Bioconductor 3.20
**Date/Publication** 2024-12-16

# Contents



---

`DEsubs` *Default run of DEsubs*

---

## Description

Default run of DEsubs

## Usage

```
DEsubs(org, mRNAexpr, mRNAnomenclature, pathways, DEtool, DEpar,
    CORtool, CORpar, subpathwayType, rankedList, verbose)
```

## Arguments

| | |
|---|---|
| `org` | Organism identifier ('hsa', 'mmu', 'rno', 'sce', 'ath', 'dme', 'dre') |
| `mRNAexpr` | RNA-seq expression data in the form of either a matrix or a filename of a text file stored in the 'User' directory. |
| `mRNAnomenclature` | mRNAnomenclature ('entrezgene', 'ensembl_gene_id', 'ensembl_transcript_id', 'ensembl_peptide_id', 'hgnc_id', 'hgnc_symbol', 'hgnc_transcript_name', 'refseq_mrna', 'refseq_peptide') |
| `pathways` | Pathway type ('All', 'Non-Metabolic', 'Metabolic') |
| `DEtool` | DEG analysis tool selection for NodeRule ('edgeR', 'DESeq2', 'EBSeq', 'NBPSeq', 'voom+limma', 'vst2+limma', 'TSPM') |
| `DEpar` | DE analysis tools Q-value threshold of NodeRule (default: DEGpar = 0.05) |
| `CORtool` | Correlation measure selection for EdgeRule ('pearson', 'kendall', 'spearman') |
| `CORpar` | Correlation measure threshold of EdgeRule (default: CORpar = 0.6) |
| `subpathwayType` | Subpathway extraction type selection (get all availiable options from subpathwayTypes) |
| `rankedList` | A named vector of genes and their corresponding significance of differential expression in the form of a Q-value. If the argument is not null, no DEtool is used for differential expression analysis. |
| `verbose` | TRUE to display informative messages, FALSE to hide. |

### Details

- Class vector needed values 1 and 2 for each class (eg. control and disease samples).
- DEpar should be less than 0.05 in order to return statistically significant DEGs.
- Higher CORpar values result in stricter correlation criteria, i.e. less acceptable interactions.

### Value

A list used as input in geneVisualization, subpathwayVisualization, subpathwayToGraph, organismVisualization

### Examples

```
load(system.file('extdata', 'data.RData', package='DEsubs'))

DEsubs.run <- DEsubs(   org='hsa',
                        mRNAexpr=mRNAexpr,
                        mRNAnomenclature='entrezgene',
                        pathways='All',
                        DEtool=NULL, DEpar=0.05,
                        CORtool='pearson', CORpar=0.6,
                        subpathwayType=NULL,
                        rankedList=rankedList)
```

---

## geneVisualization *Gene level visualization*

---

### Description

Visualizes topologically and functionally significant genes using graph theory measures as well as their correlation to pathway, disease, drug, ontology, microRNA and Transcription Factor terms based on external references.

### Usage

```
geneVisualization(DEsubs.out, measures.topological, measures.functional,
    measures.barplot, topGenes, colors.topological, colors.functional,
    colors.barplot, size.topological, size.functional, size.barplot,
    outfile.topological, outfile.functional, outfile.barplot,
    export, verbose)
```

### Arguments

| | |
|---|---|
| DEsubs.out | Return value from DEsubs |
| measures.topological | Functional visualization type(s). |
| measures.functional | Topological visualization type(s). |
| measures.barplot | Gene level Visualization type |
| topGenes | Number of genes with greater Q-values. Default value is 10. |

`colors.topological`
: A custom color mode which overrrides the default settings.

`colors.functional`
: A custom color mode which overrrides the default settings.

`colors.barplot` A custom color mode which overrrides the default settings.

`size.topological`
: A vector storing width and height of the topological measures visualization

`size.functional`
: A vector storing width and height of the functional measures visualization

`size.barplot` A vector storing width and height of the barplot visualization

`outfile.topological`
: Output file name of the topological measures visualization.

`outfile.functional`
: Output file name of the functional measures visualization.

`outfile.barplot`
: Output file name of the barplot visualization.

`export` Export type of visualizations ('plot', 'pdf')

`verbose` TRUE to display informative messages, FALSE to hide.

## Details

- Topological visualization type contains six topological graph theory measures, namely degree, betweeness centrality, closeness centrality, hub_score, eccentricity and page_rank. The availiable options are 'degree', 'betweenness', 'closeness', 'hub_score', 'eccentricity' and 'page_rank'. If no argument is supplied, all availiable options are selected by default. If the argument is NULL, no visualization is exported.
- Functional visualization type contains associations with (i) KEGG's pathway terms, (ii) Gene Ontologies of Molecular function, biological processes and cellular components) (iii) Disease terms from OMIM and GAD databases), (iv) Drug substances from DrugBank) and the influence of (v) microRNA targets from miRecords and (vi) Transcription Factor targets from Transfac and Jaspar. The availiable options are 'KEGG', 'GO_bp', 'GO_cc', 'GO_mf', 'Disease_OMIM' , 'Disease_GAD', 'Drug_DrugBank', 'miRNA' and 'TF'. If no argument is supplied, all availiable options are selected by default. If the argument is NULL, no visualization is exported. Using option 'all' results in the selection of all afforementioned options, along with any other custom gene sets within the 'DEsubs/Data' user specified-directory.

## Value

Individual measure results in matrix form.

## Examples

```
load(system.file('extdata', 'data.RData', package='DEsubs'))

outfile.topological <- tempfile(fileext='.pdf')
outfile.functional <- tempfile(fileext='.pdf')
outfile.barplot <- tempfile(fileext='.pdf')

res <- geneVisualization(
            DEsubs.out=DEsubs.out, top=10,
            measures.topological=c( 'degree', 'betweenness', 'closeness',
```

```
                          'eccentricity', 'page_rank'),
measures.functional=c(  'KEGG',
                          'Disease_OMIM', 'Disease_GAD',
                          'Drug_DrugBank','miRNA', 'TF'),
size.topological=c(5,4),
size.functional=c(7,4),
size.barplot=c(5,6),
export='pdf',
outfile.topological=outfile.topological,
outfile.functional=outfile.functional,
outfile.barplot=outfile.barplot,
verbose=FALSE)
```

---

`organismVisualization` *Organism level visualization*

---

## Description

Organism level measures

## Usage

```
organismVisualization( DEsubs.out, references, topSubs, topTerms, colors,
                       export, width, height, outfiles, verbose )
```

## Arguments

| | |
|---|---|
| `DEsubs.out` | Return value from `DEsubs` |
| `references` | Functional associations with (i) KEGG's pathway terms, (ii) Gene Ontologies of Molecular function, biological processes and cellular components) (iii) Disease terms from OMIM and GAD databases), (iv) Drug substances from DrugBank) and the influence of (v) microRNA targets from miRecords and (vi) Transcription Factor targets from Transfac and Jaspar. The corresponding options are 'KEGG', 'GO_bp', 'GO_cc', 'GO_mf', 'Disease_OMIM' , 'Disease_GAD', 'Drug_DrugBank', 'miRNA' and 'TF'. If no argument is supplied, no option is selected. Using option 'all' results in the selection of all afforementioned options, along with any other custom gene sets within the 'DEsubs/Data' user specified-directory. |
| `topSubs` | Default value is 10 |
| `topTerms` | Default value is 20 |
| `colors` | A custom color mode which overrides the default settings. |
| `export` | Export type of visualizations ('plot', 'pdf') |
| `width` | The width of the printable area (pdf) |
| `height` | The height of the printable area (pdf) |
| `outfiles` | Output filenames of the visualizations. If the argument is not specified, default filenames are used ('DEsubs/Output'). |
| `verbose` | TRUE to display informative messages, FALSE to hide. |

## Value

A list of matrices, containing Subpathway/Term/P-Value results for each reference.

## Examples

```
load(system.file('extdata', 'data.RData', package='DEsubs'))

outfile <- tempfile(fileext='.pdf')

res <- organismVisualization(
                DEsubs.out=DEsubs.out, references='KEGG',
                topSubs=10, topTerms=20,
                width=7, height=6,
                export='pdf',
                outfile=outfile,
                verbose=FALSE)
```

---

# subpathwayToGraph *Subpathway To Graph*

---

## Description

Subpathway plotting as a graph.

## Usage

```
subpathwayToGraph( DEsubs.out, submethod, subname, colors, size, export,
                   width, height, outfile, verbose )
```

## Arguments

| | |
|---|---|
| DEsubs.out | Return value from DEsubs |
| submethod | Subpathway extraction type selection (get all availiable options from subpathwayTypes) |
| subname | Subpathway name as contained in DEsubs return value, i.e. 'sub1'. |
| colors | A custom color mode which overrides the default settings. |
| size | A vector storing width and height of the barplot visualization. |
| export | A set of options for exporting subpathway data. Possible options are 'plot', 'pdf', 'edgelist', 'json', 'gml', 'ncol', 'lgl', 'graphml','dot'. |
| width | The width of the printable area (pdf) |
| height | The height of the printable area (pdf) |
| outfile | Output file name of the visualization. If multiple export types have been selected, the outfile should have an extension '.*' |
| verbose | TRUE to display informative messages, FALSE to hide. |

## Value

No value is returned.

## Examples

```
load(system.file('extdata', 'data.RData', package='DEsubs'))

outfile <- tempfile(fileext='.pdf')

res <- subpathwayToGraph(
                DEsubs.out=DEsubs.out,
                submethod='community.walktrap',
                subname=paste0('sub', 6),
                size=c(10,10),
                export='pdf',
                outfile=outfile )
```

---

`subpathwayTypes` *All subpathway types*

---

## Description

All subpathway types.

## Usage

```
subpathwayTypes(grouping)
```

## Arguments

| | |
|---|---|
| `grouping` | By supplying one of the availiable groupings, specific subsets of availiable subpathway types can be extracted. |

## Details

Apart from the 124 distinct subpathway types, several groupings are availiable:

**'all'** All subpathway options.

**'all.bwd'** Forward propagation.

**'all.fwd'** Backward propagation.

**'all.stream'** All possible genes targeting or are targeted via a path starting from a gene of interest.

**'all.neighbourhood'** All adjacent genes of a gene of interest with incoming or outgoing links.

**'all.cascade'** A finite sequence of interactions connecting a sequence of genes starting or ending from a gene of interest.

**'all.community'** Group of genes sharing common properties.

**'all.component'** Subgraphs on which any two vertices are (strongly) connected to each other by paths.

**'all.functional'** Forward and backward streams starting from genes/nodes with crucial functional role within the network. Individual options include GO_bp, GO_cc, GO_mf, KEGG, Disease_OMIM, Disease_GAD, Drug_DrugBank, miRNA, TF, DEG, which are defined below.

**'all.GO_bp'** Genes acting as a bridge among Gene Ontology (GO) Biological Process terms.

**'all.GO_cc'** Genes acting as a bridge among Gene Ontology (GO) Cellular Component terms.

**'all.GO_mf'** Genes acting as a bridge among Gene Ontology (GO) Molecular Function terms.

**'all.KEGG'** Genes acting as a bridge among KEGG pathway maps.

**'all.Disease_OMIM'** Genes acting as a bridge among OMIM Disease targets.

**'all.Disease_GAD'** Genes acting as a bridge among GAD Disease targets.

**'all.Drug_DrugBank'** Genes acting as a bridge among DrugBank Drug targets.

**'all.miRNA'** Genes acting as a bridge among microRNA-gene targets.

**'all.TF'** Genes acting as a bridge among TF-gene targets.

**'all.DEG'** Genes with highly Differentially expressed by each experimental data.

**'all.topological'** Forward and backward streams starting from genes/nodes with crucial topological role within the network. Individual options include degree, betweenness, closeness, hub_score, eccentricity, page_rank, start_nodes which are defined below.

**'all.degree'** Number of gene's adjacent interactions.

**'all.betweenness'** Number of shortest paths from all vertices to others passing through a node.

**'all.closeness'** Inverse of farness, which is the sum of distances to all other nodes.

**'all.hub_score'** Kleinberg's hub centrality score.

**'all.eccentricity'** Shortest path distance from the farthest other node in the graph.

**'all.page_rank'** Google Page Rank.

**'all.start_nodes'** Gene nodes without incoming links.

An exhaustive list of all 124 subpathway types follows:

STREAM-TOPOLOGICAL

| | |
|---|---|
| 'bwd.stream.topological.degree' | 'fwd.stream.topological.degree' |
| 'bwd.stream.topological.betweenness' | 'fwd.stream.topological.betweenness' |
| 'bwd.stream.topological.closeness' | 'fwd.stream.topological.closeness' |
| 'bwd.stream.topological.hub_score' | 'fwd.stream.topological.hub_score' |
| 'bwd.stream.topological.eccentricity' | 'fwd.stream.topological.eccentricity' |
| 'bwd.stream.topological.page_rank' | 'fwd.stream.topological.page_rank' |
| 'bwd.stream.topological.start_nodes' | 'fwd.stream.topological.start_nodes' |

STREAM-FUNCTIONAL

| | |
|---|---|
| 'bwd.stream.functional.GO_bp' | 'fwd.stream.functional.GO_bp' |
| 'bwd.stream.functional.GO_cc' | 'fwd.stream.functional.GO_cc' |
| 'bwd.stream.functional.GO_mf' | 'fwd.stream.functional.GO_mf' |
| 'bwd.stream.functional.KEGG' | 'fwd.stream.functional.KEGG' |
| 'bwd.stream.functional.Disease_OMIM' | 'fwd.stream.functional.Disease_OMIM' |
| 'bwd.stream.functional.Disease_GAD' | 'fwd.stream.functional.Disease_GAD' |
| 'bwd.stream.functional.Drug_DrugBank' | 'fwd.stream.functional.Drug_DrugBank' |
| 'bwd.stream.functional.miRNA' | 'fwd.stream.functional.miRNA' |
| 'bwd.stream.functional.TF' | 'fwd.stream.functional.TF' |
| 'bwd.stream.functional.DEG' | 'fwd.stream.functional.DEG' |

NEIGHBOURHOOD-TOPOLOGICAL

| | |
|---|---|
| 'bwd.neighbourhood.topological.degree' | 'fwd.neighbourhood.topological.degree' |
| 'bwd.neighbourhood.topological.betweenness' | 'fwd.neighbourhood.topological.betweenness' |
| 'bwd.neighbourhood.topological.closeness' | 'fwd.neighbourhood.topological.closeness' |
| 'bwd.neighbourhood.topological.hub_score' | 'fwd.neighbourhood.topological.hub_score' |
| 'bwd.neighbourhood.topological.eccentricity' | 'fwd.neighbourhood.topological.eccentricity' |
| 'bwd.neighbourhood.topological.page_rank' | 'fwd.neighbourhood.topological.page_rank' |
| 'bwd.neighbourhood.topological.start_nodes' | 'fwd.neighbourhood.topological.start_nodes' |

NEIGHBOURHOOD-FUNCTIONAL

| | |
|---|---|
| 'bwd.neighbourhood.functional.GO_bp' | 'fwd.neighbourhood.functional.GO_bp' |
| 'bwd.neighbourhood.functional.GO_cc' | 'fwd.neighbourhood.functional.GO_cc' |
| 'bwd.neighbourhood.functional.GO_mf' | 'fwd.neighbourhood.functional.GO_mf' |
| 'bwd.neighbourhood.functional.KEGG' | 'fwd.neighbourhood.functional.KEGG' |
| 'bwd.neighbourhood.functional.Disease_OMIM' | 'fwd.neighbourhood.functional.Disease_OMIM' |
| 'bwd.neighbourhood.functional.Disease_GAD' | 'fwd.neighbourhood.functional.Disease_GAD' |
| 'bwd.neighbourhood.functional.Drug_DrugBank' | 'fwd.neighbourhood.functional.Drug_DrugBank' |
| 'bwd.neighbourhood.functional.miRNA' | 'fwd.neighbourhood.functional.miRNA' |
| 'bwd.neighbourhood.functional.TF' | 'fwd.neighbourhood.functional.DEG' |

CASCADE-TOPOLOGICAL

| | |
|---|---|
| 'bwd.cascade.topological.degree' | 'fwd.cascade.topological.degree' |
| 'bwd.cascade.topological.betweenness' | 'fwd.cascade.topological.betweenness' |
| 'bwd.cascade.topological.closeness' | 'fwd.cascade.topological.closeness' |
| 'bwd.cascade.topological.hub_score' | 'fwd.cascade.topological.hub_score' |
| 'bwd.cascade.topological.eccentricity' | 'fwd.cascade.topological.eccentricity' |
| 'bwd.cascade.topological.page_rank' | 'fwd.cascade.topological.page_rank' |
| 'bwd.cascade.topological.start_nodes' | 'fwd.cascade.topological.start_nodes' |

CASCADE-FUNCTIONAL

| | |
|---|---|
| 'bwd.cascade.functional.GO_bp' | 'fwd.cascade.functional.GO_bp' |
| 'bwd.cascade.functional.GO_cc' | 'fwd.cascade.functional.GO_cc' |
| 'bwd.cascade.functional.GO_mf' | 'fwd.cascade.functional.GO_mf' |
| 'bwd.cascade.functional.KEGG' | 'fwd.cascade.functional.KEGG' |
| 'bwd.cascade.functional.Disease_OMIM' | 'fwd.cascade.functional.Disease_OMIM' |
| 'bwd.cascade.functional.Disease_GAD' | 'fwd.cascade.functional.Disease_GAD' |
| 'bwd.cascade.functional.Drug_DrugBank' | 'fwd.cascade.functional.Drug_DrugBank' |
| 'bwd.cascade.functional.miRNA' | 'fwd.cascade.functional.miRNA' |
| 'bwd.cascade.functional.TF' | 'fwd.cascade.functional.TF' |
| 'bwd.cascade.functional.DEG' | 'fwd.cascade.functional.DEG' |

COMMUNITY

| | |
|---|---|
| 'community.walktrap' | 'community.edge_betweenness' |
| 'community.fast_greedy' | 'community.leading_eigen' |
| 'community.infomap' | 'community.louvain' |

COMPONENT-CLIQUES

| | |
|---|---|
| 'component.max_cliques' | 'component.decompose' |
| 'component.3-cliques' | ... |
| ... | 'component.9-cliques' |

COMPONENT-CORENESS

'component.3-coreness'
...
...
'component.9-coreness'

## Value

A vector of all 124 basic subpathway types. See Details section for handly groupings.

## Examples

```
basic.types <- subpathwayTypes()
stream.types <- subpathwayTypes(grouping='all.stream')
```

---

`subpathwayVisualization`
*Subpathway level visualization*

---

## Description

Circular diagrams containing subpathways enrichment in potential key regulators (miRNAs, TFs) and biological, biomedical and pharmacological issues.

## Usage

```
subpathwayVisualization( DEsubs.out, references, submethod, subname, colors,
                         scale, shuffleColors, outfiles, export, verbose )
```

## Arguments

| | |
|---|---|
| `DEsubs.out` | Return value from `DEsubs`. |
| `references` | Topological references include degree, betweeness centrality, closeness centrality, hub_score, eccentricity and page_rank. The corresponding options are 'degree', 'betweenness', 'closeness', 'hub_score', 'eccentricity' and 'page_rank'. Functional references include (i) KEGG's pathway terms, (ii) Gene Ontologies of Molecular function, biological processes and cellular components) (iii) Disease terms from OMIM and GAD databases), (iv) Drug substances from DrugBank) and the influence of (v) microRNA targets from miRecords and (vi) Transcription Factor targets from Transfac and Jaspar. The corresponding options are 'KEGG', 'GO_bp', 'GO_cc', 'GO_mf', 'Disease_OMIM' , 'Disease_GAD', 'Drug_DrugBank', 'miRNA' and 'TF'.Using option 'all' results in the selection of all afforementioned options, along with any other custom gene sets within the 'DEsubs/Data' user specified-directory. |
| `submethod` | Subpathway extraction type selection (see all 124 options along with their R commands in supplementary document) |
| `subname` | Subpathway name as contained in `DEsubs` return value, i.e. 'sub1' |
| `colors` | A custom color mode which overrrides the default settings. |
| `scale` | A value in (0,1] used to scale the visualization. Useful in the case of long labels which are trimmed by default. |
| `shuffleColors` | TRUE to shuffle user defined or default colors. Defaults to FALSE. |

| | |
|---|---|
| outfiles | Output filenames of the visualizations. If the argument is not specified, default filenames are used ('DEsubs/Output'). |
| export | Export type of visualizations ('plot', 'pdf') |
| verbose | TRUE to display informative messages, FALSE to hide. |

## Details

The associations of subpathways with various biological and pharmacologicalfeatures are estimated through a hypergeometric test. The enriched associations of a subpathway to each feature are illustrated through circular diagrams.

## Value

A list of matrices, each containing the P-Value of enrichment between the terms for a specific reference (rows) and each of the subpathway genes (columns). If there is no enrichment, the value is NA.

## Examples

```
load(system.file('extdata', 'data.RData', package='DEsubs'))

outfile <- tempfile(fileext='.pdf')

res <- subpathwayVisualization(
                DEsubs.out=DEsubs.out,
                references=c('TF'),
                submethod='community.walktrap',
                subname='sub1',
                scale=c(1.0),
                export='pdf',
                outfile=outfile )
```

# Index