

# Package ‘FunNet’

July 18, 2009

**Version** 1.00-7

**Date** 2009-06-30

**Title** Integrative Functional Analysis of Transcriptional Networks

**Author** Corneliu Henegar <corneliu@henegar.info>

**Maintainer** Corneliu Henegar <corneliu@henegar.info>

**Depends** R (>= 2.6.0), ade4, cluster, Hmisc, nlme, sna, Cairo

**Description** FunNet is an integrative tool for analyzing gene co-expression networks built from microarray expression data. The analytic model implemented in this library involves two abstraction layers: transcriptional and functional (biological roles). A functional profiling technique using Gene Ontology & KEGG annotations is applied to extract a list of relevant biological themes from microarray expression profiling data. Afterwards multiple-instance representations are built to relate significant themes to their transcriptional instances (i.e. the two layers of the model). An adapted non-linear dynamical system model is used to quantify the proximity of relevant genomic themes based on the similarity of the expression profiles of their gene instances. Eventually an unsupervised multiple-instance clustering procedure, relying on the two abstraction layers, is used to identify the structure of the co-expression network composed from modules of functionally related transcripts. Functional and transcriptional maps of the co-expression network are provided separately together with detailed information on the network centrality of related transcripts and genomic themes.

**License** GPL (>= 2)

**URL** <http://corneliu.henegar.info/FunNet.htm>,  
<http://www.geneontology.org/GO.tools.microarray.shtml#funnet>, <http://www.funnet.info>,  
<http://www.funnet.ws>

**Repository** CRAN

**Date/Publication** 2009-07-18 15:22:15

## R topics documented:

FunNet.R-package	2
Annotations	4
bypass	6
FunNet	7
obese	13
yeast.cycle	14
yeast.dna	15
<b>Index</b>	<b>16</b>

---

FunNet.R-package     *Integrative Functional Analysis of Transcriptional Networks*

---

### Description

FunNet is an integrative tool for analyzing gene co-expression networks built from microarray expression data. The analytic model implemented in this library involves two abstraction layers: transcriptional and functional (biological roles). A functional profiling technique using Gene Ontology & KEGG annotations is applied to extract a list of relevant biological themes from microarray expression profiling data. Afterwards multiple-instance representations are built to relate significant themes to their transcriptional instances (i.e. the two layers of the model). An adapted non-linear dynamical system model is used to quantify the proximity of relevant genomic themes based on the similarity of the expression profiles of their gene instances. Eventually an unsupervised multiple-instance clustering procedure, relying on the two abstraction layers, is used to identify the structure of the co-expression network composed from modules of functionally related transcripts. Functional and transcriptional maps of the co-expression network are provided separately together with detailed information on the network centrality of related transcripts and genomic themes.

### Details

Together with the FunNet algorithm this package provides also:

1. GO and KEGG annotations automatically extracted from their respective web resources and updated on a regular basis
2. The routine for the automated extraction and update of the functional annotations from their respective web resources. The use of this routine is simple: `annotations(date.annot = "")`. Under common circumstances these routine will provide up-to-date annotations, stored into environmental variables, directly formatted for FunNet's use. Some errors may be seen when using this routine related to a lack of availability of the GO annotations for the current month. In case of extraction errors, explained most usually by a delay in updating GO web servers, the release date can be expressly indicated (see [annotations](#)).
3. Four test data sets (see examples below and the dedicated man pages). Two of these datasets are related to adipose tissue expression profiling in obese subjects at baseline and after a bariatric surgery. The other two are yeast datasets related to the cell cycle and DNA repairing processes

induced by irradiation.

The format of the data should be respected in order to perform a successful analysis. The only transcript identification system acceptable for FunNet analysis is EntrezGene GeneID's. The transcript expression data should be organized in dataframes within one row for each transcript. The first column contains the transcript identifiers for each transcript and the rest of them the expression level of that transcript in each of the available microarray samples. See the provided test data for more details.

The results of the FunNet analysis of transcript expression data are stored as HTML, tab separated text or R data files in a "Results" subfolder of the working folder. For each type of available biological annotations and for each list of transcript expression data to be analyzed (one or two), FunNet provides a ranked list with the significantly enriched annotating categories, as well as network structures as text files designed to be imported in Cytoscape for graphical analysis. Detailed findings on the terminological composition and transcript enrichment significance of the resulting functional clusters, as well as various network centrality measures are equally provided.

#### Author(s)

Corneliu Henegar (corneliu@henegar.info)

#### References

1. Prifti E, Zucker JD, Clement K, Henegar C. FunNet: an integrative tool for exploring transcriptional interactions. *Bioinformatics*. 2008 Nov 15;24(22):2636-8.
2. Henegar C, Tordjman J, Achard V, Lacasa D, Cremer I, Guerre-Millo M, Poitou C, Basdevant A, Stich V, Viguerie N, Langin D, Bedossa P, Zucker J-D, Clement K. Adipose tissue transcriptomic signature highlights the pathologic relevance of extracellular matrix in human obesity. *Genome Biology* 2008, 9(1):R14.
3. Henegar C, Clement K, and Zucker JD (2006). Unsupervised multiple-instance learning for functional profiling of genomic data. *Lecture Notes in Computer Science: ECML 2006*. Springer Berlin / Heidelberg, 4212/2006 : 186-197.
4. Henegar C, Canello R, Rome S, Vidal H, Clement K, Zucker JD. Clustering biological annotations and gene expression data to identify putatively co-regulated biological processes. *J Bioinform Comput Biol*. 2006 Aug;4(4):833-52.
5. Canello R, Henegar C, Viguerie N, Taleb S, Poitou C, Rouault C, Coupaye M, Pelloux V, Hugol D, Bouillot JL, Bouloumie A, Barbatelli G, Cinti S, Svensson PA, Barsh GS, Zucker JD, Basdevant A, Langin D, Clement K. Reduction of macrophage infiltration and chemoattractant gene expression changes in white adipose tissue of morbidly obese subjects after surgery-induced weight loss. *Diabetes* 2005; 54(8):2277-86.
6. Zhang B, Horvath S. A general framework for weighted gene co-expression network analysis. *Stat Appl Genet Mol Biol* 4 (2005) Article17.
7. FunNet websites: <http://corneliu.henegar.info/FunNet.htm>, <http://www.funnet.ws>, <http://www.funnet.info>

**See Also**

[cluster](#).

**Examples**

```
## Not run:
## most common use
data(obese)
FunNet(org="HS", two.lists=TRUE, up.frame=up.frame, down.frame=down.frame,
       genes.frame=NULL, restrict=TRUE, ref.list=ref.list, logged=TRUE,
       discriminant=TRUE, go.bp=TRUE, go.cc=TRUE, go.mf=TRUE, kegg=TRUE,
       annot.method="specificity", annot.details=TRUE,
       direct=FALSE, enriched=TRUE, fdr=NA, build.annot.net=TRUE,
       coexp.matrix=NULL, coexp.method="spearman", estimate.th=FALSE,
       hard.th=0.8, soft.th=NA, topological = FALSE, keep.sign=FALSE, level=1,
       annot.clust.method="umilds", annot.prox.measure="dynamical",
       test.recovery=FALSE, test.robust=FALSE, replace.annot=NA,
       build.gene.net=TRUE, gene.clust.method="hclust", gene.net.details=TRUE,
       gene.clusters=NA, alpha=0.05, RV=0.90, sigma=NA, keep.rdata=FALSE, zip=TRUE)

## the following example forces the use of a previous GO release
(e.g. October 2007) for updating annotations
annotations(date.annot = "200710")
## End(Not run)
```

---

Annotations

*Integrative Functional Analysis of Transcriptional Networks*

---

**Description**

This routine is close to a similar one belonging to the package FunCluster. The actual one, provided with the package FunNet, performs a slightly different automated extraction and update of the Gene Ontology & KEGG annotations which are needed for FunNet analysis. The difference relies in the fact that additionally to the gene annotation data, extracted from the NCBI resources, this routine provides also the ontological lattice of GO required by FunNet enrichment computation routine for specificity, terminological or decorrelated annotation.

**Usage**

```
annotations(cust.specs=NULL)
```

**Arguments**

`cust.specs` this parameter is experimental at this moment.

## Details

For details concerning the utilization of the annotations data by the FunNet package please see FunNet help or man page `help(FunNet)`. The "Annotations" routine is allowing the automated extraction and update of the functional annotations from their respective web resources. Under common circumstances this routine will provide up-to-date annotations, stored into environmental variables and directly formatted for FunNet use. Some errors may be seen when using this routine to update annotations provided within the package. These errors are mainly related to the availability of GO annotations for the current month. In case of extraction errors, explained most usually by a delay in updating GO web servers, the date of the GO release to be used can be expressly indicated through the parameter `annot.date` (see example below). The transcript identification system used for FunNet analysis is EntrezGene GeneID's. Also some errors may be related to changes in the format of annotation files provided by the NCBI. Please update your package with the last version available as this may correct such issues.

The annotation data is provided as a unique R data archive which should be manually loaded into R after loading the FunNet package in order to be able to use the updated annotations for your analysis.

Important note for Microsoft Windows users: in order to use this routine you will need additional software for handling TAR and GZIP archives. This software is freely available for Windows under the GNU license.

For TAR packages please go to:

<http://gnuwin32.sourceforge.net/packages/tar.htm>. For GZIP you can go to: <http://gnuwin32.sourceforge.net/packages/gzip.htm>. The TAR and GZIP executable and their dependencies (DLL's) should be placed somewhere into the PATH (like "C:/Windows" for example) in order to be available for R calls.

## References

1. Prifti E, Zucker JD, Clement K, Henegar C. FunNet: an integrative tool for exploring transcriptional interactions. *Bioinformatics*. 2008 Nov 15;24(22):2636-8.
2. Henegar C, Tordjman J, Achard V, Lacasa D, Cremer I, Guerre-Millo M, Poitou C, Basdevant A, Stich V, Viguerie N, Langin D, Bedossa P, Zucker J-D, Clement K. Adipose tissue transcriptomic signature highlights the pathologic relevance of extracellular matrix in human obesity. *Genome Biology* 2008, 9(1):R14.
3. Henegar C, Clement K, and Zucker JD (2006). Unsupervised multiple-instance learning for functional profiling of genomic data. *Lecture Notes in Computer Science: ECML 2006*. Springer Berlin / Heidelberg, 4212/2006 : 186-197.
4. Henegar C, Canello R, Rome S, Vidal H, Clement K, Zucker JD. Clustering biological annotations and gene expression data to identify putatively co-regulated biological processes. *J Bioinform Comput Biol*. 2006 Aug;4(4):833-52.
5. Canello R, Henegar C, Viguerie N, Taleb S, Poitou C, Rouault C, Coupaye M, Pelloux V, Hugol D, Bouillot JL, Bouloumie A, Barbatelli G, Cinti S, Svensson PA, Barsh GS, Zucker JD, Basdevant A, Langin D, Clement K. Reduction of macrophage infiltration and chemoattractant gene expression changes in white adipose tissue of morbidly obese subjects after surgery-induced weight loss. *Diabetes* 2005; 54(8):2277-86.

6. Zhang B, Horvath S. A general framework for weighted gene co-expression network analysis. *Stat Appl Genet Mol Biol* 4 (2005) Article17.

7. FunNet websites: <http://corneliu.henegar.info/FunNet.htm>, <http://www.funnet.ws>, <http://www.funnet.info>

### See Also

[FunNet](#).

### Examples

```
## Not run:
annotations()

## End(Not run)
```

---

bypass

*Gene Expression Profiling of the Human Adipose Tissue of Massively Obese Subjects Before and After Bariatric Surgery*

---

### Description

This data set resulted from a study investigating the dynamics of adipose tissue gene expression profiles induced in massively obese subjects by a form of bariatric surgery (gastric bypass). Please see the references provided below for further experimental details.

### Usage

```
data(bypass)
```

### Format

Three data frames containing significant up-regulated or down-regulated genes and a reference list with all the genes tested for differential expression during this experiment.

### Source

<http://corneliu.henegar.info/FunNet.htm>

### References

1. Canello R, Henegar C, Viguerie N, Taleb S, Poitou C, Rouault C, Coupaye M, Pelloux V, Hugol D, Bouillot JL, Bouloumie A, Barbatelli G, Cinti S, Svensson PA, Barsh GS, Zucker JD, Basdevant A, Langin D, Clement K. Reduction of macrophage infiltration and chemoattractant gene expression changes in white adipose tissue of morbidly obese subjects after surgery-induced weight loss. *Diabetes* 2005; 54(8):2277-86.
2. Henegar C, Tordjman J, Achard V, Lacasa D, Cremer I, Guerre-Millo M, Poitou C, Basdevant A, Stich V, Viguerie N, Langin D, Bedossa P, Zucker J-D, Clement K. Adipose tissue transcriptomic

signature highlights the pathologic relevance of extracellular matrix in human obesity. In positive revision in Genome Biology (October 2007).

3. GEO accession number: GSE9007. <http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE9007>.

### See Also

[FunNet](#).

---

FunNet

*Integrative Functional Analysis of Transcriptional Networks*

---

### Description

FunNet is an integrative tool for analyzing gene co-expression networks built from microarray expression data. The analytic model implemented in this library involves two abstraction layers: transcriptional and functional (biological roles). A functional profiling technique using Gene Ontology & KEGG annotations is applied to extract a list of relevant biological themes from microarray expression profiling data. Afterwards multiple-instance representations are built to relate significant themes to their transcriptional instances (i.e. the two layers of the model). An adapted non-linear dynamical system model is used to quantify the proximity of relevant genomic themes based on the similarity of the expression profiles of their gene instances. Eventually an unsupervised multiple-instance clustering procedure, relying on the two abstraction layers, is used to identify the structure of the co-expression network composed from modules of functionally related transcripts. Functional and transcriptional maps of the co-expression network are provided separately together with detailed information on the network centrality of related transcripts and genomic themes.

### Usage

```
FunNet(wd="", org="hsa", two.lists=TRUE, up.frame=NULL, down.frame=NULL,
genes.frame=NULL, restrict=FALSE, ref.list=NULL, logged=TRUE,
discriminant=FALSE, go.bp=TRUE, go.cc=TRUE, go.mf=TRUE, kegg=TRUE,
annot.method="specificity", annot.details=TRUE,
direct=FALSE, enriched=TRUE, fdr=NA, build.annot.net=TRUE,
coexp.matrix=NULL, coexp.method="spearman", estimate.th=FALSE,
hard.th=NA, soft.th=NA, topological = FALSE, keep.sign=FALSE, level
annot.clust.method="umilds", annot.prox.measure="unilat.pond.norm.m
test.recovery=FALSE, test.robust=FALSE, replace.annot=NA, random.ar
build.gene.net=FALSE, gene.clust.method="hclust", gene.net.details=
gene.clusters=NA, alpha=0.05, RV=0.90, sigma=NA, keep.rdata=FALSE,
```

### Arguments

`wd` sets the working directory where the expression data files are to be found and where results are to be stored.

<code>org</code>	indicates the biological species to which analyzable transcript expression data is related; currently only three possibilities are available with FunNet: "hsa" for human expression data, "mmu" for mouse ( <i>Mus Musculus</i> ) expression data, "rna" for rat ( <i>Rattus Norvegicus</i> ) expression data, and "sce" for yeast ( <i>Saccharomyces Cerevisiae</i> ) expression data. Default value is "hsa".
<code>up.frame</code>	a dataframe containing expression data for up-regulated genes for the case where two lists of genes are analyzed comparatively (i.e. up- vs. down-regulated genes). The first column should contain the GeneID's of the analyzed genes (text data), while the other columns should provide expression measurements (numeric data) in the analyzed microarray experiments (one column for each array). The default value is NULL.
<code>down.frame</code>	a dataframe containing expression data for down-regulated genes for the case where two lists of genes are analyzed comparatively (i.e. up- vs. down-regulated genes). The first column should contain the GeneID's of the analyzed genes (text data), while the other columns should provide expression measurements (numeric data) in the analyzed microarray experiments (one column for each array). The default value is NULL.
<code>genes.frame</code>	a dataframe containing expression data for analyzed genes for the case where one list of genes is analyzed. The first column should contain the GeneID's of the analyzed genes (text data), while the other columns should provide expression measurements (numeric data) in the analyzed microarray experiments (one column for each array). The default value is NULL.
<code>ref.list</code>	a dataframe containing a single column providing GeneID's (text data) to be used as the reference list of analyzed genes (i.e. all genes spotted on the microarrays used for expression profiling). The default value is NULL.
<code>two.lists</code>	possible values are TRUE if a discriminatory functional analysis of two lists of transcripts is required (e.g. significantly up-regulated transcripts versus down-regulated transcripts) or FALSE if only one list of transcripts is to be analyzed. Please see the provided datasets for the required format of the data files. The default value of this parameter is TRUE.
<code>restrict</code>	possible values are TRUE if a reference list of transcripts is provided for the statistical significance calculation of the transcript enrichment of the biological annotations or FALSE if such a restriction is not imposed and the transcript enrichment significance is therefore estimated with regards of the whole genome. The purpose of the reference list is to correct the significance of the enrichment calculations for those situations in which expression data is not available for the whole genome but only for a fraction of it, either because of microarray processing errors which limits the number of transcripts available for analysis, or for the case of dedicated microarrays, designed to scan only a fraction of the genome. The transcripts should be identified only by their EntrezGene ID number. The default value for this parameter is FALSE.
<code>logged</code>	possible values are TRUE or FALSE. This parameter indicates whether the expression measurements are provided as log fold changes or as simple fold changes. This parameter is used exclusively in the filtering and pretreatment of the provided expression data.
<code>discriminant</code>	possible values are TRUE or FALSE. This parameter indicate whether a discriminant annotation of two lists of genes should be performed (if TRUE) or

alternatively only an independent annotation of each list of genes (if FALSE). In the case of a discriminant analysis the enrichment of each list is computed with regards to their union, while otherwise the provided reference list is used for enrichment computation. The default value for this parameter is FALSE.

<code>go.bp</code>	possible values are TRUE or FALSE. This parameter indicates whether GO Biological Process annotations should be used. Default value is TRUE.
<code>go.cc</code>	possible values are TRUE or FALSE. This parameter indicates whether GO Cellular Component annotations should be used. Default value is TRUE.
<code>go.mf</code>	possible values are TRUE or FALSE. This parameter indicates whether GO Molecular Function annotations should be used. Default value is TRUE.
<code>kegg</code>	possible values are TRUE or FALSE. This parameter indicates whether KEGG annotations should be used. Default value is TRUE.
<code>annot.method</code>	indicates the type of ontological annotation to be used (for GO annotations only). Three values are possible: "specificity", "terminological", "decorrelated". "Specificity" will direct the enrichment computation with regards to the precision of the annotation. In this situation "specificity" levels are identified in reference to the most precise annotations available (i.e. direct annotations) which constitute the first specificity level. Further levels of annotation precision are identified by following annotation inheritance in the GO lattice. "Terminological" will indicate that the enrichment computations should be performed in reference to the terminological levels of GO (i.e. levels of conceptual precision within the ontology). "Decorrelated" will use a specific decorrelating technique to reduce informational redundancy within GO. A single list of enriched categories is provided in this last case, while in the previous two a hierarchy of lists, one for each level of annotation specificity or conceptual precision, is provided. The default value for this parameter is "specificity".
<code>annot.details</code>	indicates whether detailed results reflecting enrichment computation should be stored as HTML files. The default value for this parameter is TRUE.
<code>direct</code>	indicates whether the terminological annotation (see <code>annot.method</code> parameter) should consider the precision of annotating categories or not. If FALSE (default) a conventional terminological annotation is performed. Otherwise the enrichment computation considers for each ontological level only categories with a similar annotation precision.
<code>enriched</code>	if FALSE it allows to use all available annotations regardless of their enrichment significance. Provided only for experimental purposes. The default value for this parameter is TRUE.
<code>fdr</code>	if numeric it uses the Storey FDR approach for enrichment p-values correction to a False Discovery Rate equal to the FDR in percentages. The default value for this parameter is NA (no FDR correction is performed in this case).
<code>build.annot.net</code>	indicates whether a functional interaction network should be computed. The default value for this parameter is TRUE.
<code>coexp.matrix</code>	allows to specify another co-expression matrix than the one computed by default. The default value is NULL.

<code>coexp.method</code>	indicates the co-expression measure to be used to compare gene expression profiles. Four values are possible: "spearman", "pearson", "kendall", "euclid". The first three are conventional correlation coefficients while the last one indicates an Euclidean distance between expression profiles.
<code>estimate.th</code>	indicates whether an estimation of the co-expression threshold based on a theoretical scale-free distribution (power law) should be performed. Both a "hard" (i.e. the adjustment quality for various discrete values of the used co-expression measure) and a "soft" (i.e. the adjustment quality for various power law exponents) threshold are computed (for details please see Zhang & Horvath's paper indicated as reference). Based on these calculations the values of the hard or soft threshold must be manually chosen and indicated in the respective parameters. The default value for this parameter is FALSE.
<code>hard.th</code>	indicates the value of the discrete hard threshold for co-expression significance. Either the hard or the soft threshold must be specified except when their estimation is requested. The default value is NA.
<code>soft.th</code>	indicates the value of the discrete soft threshold for co-expression significance. Either the hard or the soft threshold must be specified except when their estimation is requested. The default value is NA.
<code>topological</code>	indicates whether a topological measure of similarity between gene expression profiles should be computed based on the co-expression network build with the co-expression measure indicated by the <code>coexp.method</code> parameter. The default value for this parameter is FALSE.
<code>keep.sign</code>	indicates whether the sign of the correlation coefficient should be considered when computing gene co-expression. The default value for this parameter is FALSE.
<code>level</code>	indicates the level of terminological or specificity annotation to be used for building functional interaction networks. The default value for this parameter is NA which implies that the most specific (i.e. the first) level will be used. It has no impact on KEGG and "decorrelated" GO annotation.
<code>annot.clust.method</code>	indicates the multiple-instance clustering algorithm to be used to explore the structure of functional interaction modules. The two possibilities "umilds" and "ucknn" are detailed in the respective references provided below. The default value is "umilds" which performs a dynamical estimation of the proximity between annotating categories based on the co-expression of their annotated transcripts and then uses a spectral clustering technique to explore the modular structure of the network.
<code>annot.prox.measure</code>	indicates the type of measures to be used for computing the proximity between annotating categories. It has five values: "dynamical", "unilat.pond.norm.mean", "unilat.norm.sum", "norm.sum" and "pond.norm.mean". The first one is to be used with the "umilds" algorithm (see <code>annot.clust.method</code> parameter), while the others are symmetrical or asymmetrical proximity measures to be used with the "ucknn" algorithm. By default "umilds" algorithm uses the dynamical proximity estimation so there is no need to indicate it except for experimental purposes.

<code>test.recovery</code>	this parameter is purely experimental and therefore it should not be modified. The default value is FALSE.
<code>test.robust</code>	this parameter is purely experimental and therefore it should not be modified. The default value is FALSE.
<code>replace.annot</code>	this parameter is purely experimental and therefore it should not be modified. The default value is NA.
<code>random.annot</code>	this parameter is purely experimental and therefore it should not be modified. The default value is FALSE.
<code>build.gene.net</code>	indicates whether a simple gene co-expression network should be computed. If TRUE the results of this computation does not take into account functional assignment of analyzed genes. The default value is FALSE.
<code>gene.clust.method</code>	this parameter is experimental at this stage and should not be modified.
<code>gene.net.details</code>	indicates whether detailed results containing network centrality information in relation to the analyzed transcripts should be provided. The default value is FALSE.
<code>gene.clusters</code>	allows to specify a predefined number of gene clusters to be considered in the analysis of the gene co-expression network. By default the number of clusters (i.e. modules) is computed by optimizing the Silhouette of the resulting gene clustering partitions.
<code>alpha</code>	indicates the threshold of p-values significance (alpha) resulting from statistical calculations concerning transcript enrichment of biological annotations. Default value is 0.05.
<code>RV</code>	allows to control the strength of the co-inertia level used to evaluate the convergence of the dynamical model on which relies the dynamical computation of proximity between annotating categories. It is used in combination with a co-inertia analysis and a Mantel test (see <code>ade4</code> package). The default value is 0.9.
<code>sigma</code>	this parameter is experimental at this stage and therefore it should not be modified.
<code>keep.rdata</code>	this parameter allows to remove the temporary RData files in the end of the computations. The default value is FALSE
<code>zip</code>	this parameter allows to create a ZIP archive for storing the results in the end of the computations. The default value is TRUE

## Details

FunNet can be used with the currently available R distributions (tested with distributions posterior to 2.5.0), either with Microsoft Windows operating environments (tested with Windows XP), Mac OS (tested with OS X Leopard) or, better, with a Linux operating environment (tested with

Debian 4.0 and OpenSuse 10.3). Please be aware that FunNet analysis implies a lot of computations and therefore high processing power and good stability of the operating system are absolute requirements.

Together with the FunNet algorithm this package provide also:

1. GO and KEGG annotations (as of February 2008) automatically extracted from their respective web resources
2. The routine for the automated extraction and update of the functional annotations from their respective web resources. The use of this routine is simple: `annotations(date.annot = " ")`. Under common circumstances these routine will provide up-to-date annotations, stored into environmental variables, directly formatted for FunNet's use. Some errors may be seen when using this routine related to a lack of availability of the GO annotations for the current month. In case of extraction errors, explained most usually by a delay in updating GO web servers, the release date can be expressly indicated (see `annotations`).
3. Four test data sets (see examples below and the dedicated man pages). Two of these datasets are related to adipose tissue expression profiling in obese subjects at baseline and after a bariatric surgery. The other two are yeast datasets related to the cell cycle and DNA repairing processes induced by irradiation.

The format of the data should be respected in order to perform a successful analysis. The only transcript identification system acceptable for FunNet analysis is EntrezGene GeneID's. The transcript expression data should be organized in dataframes within one row for each transcript. The first column contains the transcript identifiers for each transcript and the rest of them the expression level of that transcript in each of the available microarray samples. See the provided test data for more details.

The results of the FunNet analysis of transcript expression data are stored as HTML, tab separated text or R data files in a "Results" subfolder of the working folder. For each type of available biological annotations and for each list of transcript expression data to be analyzed (one or two), FunNet provides a ranked list with the significantly enriched annotating categories, as well as network structures as text files designed to be imported in Cytoscape for graphical analysis. Detailed findings on the terminological composition and transcript enrichment significance of the resulting functional clusters, as well as various network centrality measures are equally provided.

## References

1. Prifti E, Zucker JD, Clement K, Henegar C. FunNet: an integrative tool for exploring transcriptional interactions. *Bioinformatics*. 2008 Nov 15;24(22):2636-8.
2. Henegar C, Tordjman J, Achard V, Lacasa D, Cremer I, Guerre-Millo M, Poitou C, Basdevant A, Stich V, Viguerie N, Langin D, Bedossa P, Zucker J-D, Clement K. Adipose tissue transcriptomic signature highlights the pathologic relevance of extracellular matrix in human obesity. *Genome Biology* 2008, 9(1):R14.
3. Henegar C, Clement K, and Zucker JD (2006). Unsupervised multiple-instance learning for functional profiling of genomic data. *Lecture Notes in Computer Science: ECML 2006*. Springer Berlin / Heidelberg, 4212/2006 : 186-197.

4. Henegar C, Canello R, Rome S, Vidal H, Clement K, Zucker JD. Clustering biological annotations and gene expression data to identify putatively co-regulated biological processes. *J Bioinform Comput Biol*. 2006 Aug;4(4):833-52.
5. Canello R, Henegar C, Viguerie N, Taleb S, Poitou C, Rouault C, Coupaye M, Pelloux V, Hugol D, Bouillot JL, Bouloumie A, Barbatelli G, Cinti S, Svensson PA, Barsh GS, Zucker JD, Basdevant A, Langin D, Clement K. Reduction of macrophage infiltration and chemoattractant gene expression changes in white adipose tissue of morbidly obese subjects after surgery-induced weight loss. *Diabetes* 2005; 54(8):2277-86.
6. Zhang B, Horvath S. A general framework for weighted gene co-expression network analysis. *Stat Appl Genet Mol Biol* 4 (2005) Article17.
7. FunNet websites: <http://corneliu.henegar.info/FunNet.htm>, <http://www.funnet.ws>, <http://www.funnet.info>

### See Also

`cluster`, `cluster`, `link{FunCluster}`.

### Examples

```
## Not run:
## load bypass data (see Diabetes and Genome Biology papers for details)
data(bypass)

## or load adipose tissue expression profiling data (see Genome Biology paper for
data(obese)

## most common use
data(obese)
FunNet(org="HS", two.lists=TRUE, up.frame=up.frame, down.frame=down.frame,
genes.frame=NULL, restrict=TRUE, ref.list=ref.list, logged=TRUE,
discriminant=TRUE, go.bp=TRUE, go.cc=TRUE, go.mf=TRUE, kegg=TRUE,
annot.method="specificity", annot.details=TRUE,
direct=FALSE, enriched=TRUE, fdr=NA, build.annot.net=TRUE,
coexp.matrix=NULL, coexp.method="spearman", estimate.th=FALSE,
hard.th=0.8, soft.th=NA, topological = FALSE, keep.sign=FALSE, level=1,
annot.clust.method="umilds", annot.prox.measure="dynamical",
test.recovery=FALSE, test.robust=FALSE, replace.annot=NA, random.annot=FALSE,
build.gene.net=TRUE, gene.clust.method="hclust", gene.net.details=TRUE,
gene.clusters=NA, alpha=0.05, RV=0.90, sigma=NA, keep.rdata=FALSE, zip=TRUE)
## End(Not run)
```

**Description**

This data set resulted from microarray experiments performed on human white adipose tissue comparing the expression profiles in obese subjects and lean controls. Please see the references provided below for further experimental details.

**Usage**

```
data(obese)
```

**Format**

Three data frames containing significant genes up-regulated or down-regulated in human WAT of obese subjects and a reference list with all the genes tested for differential expression during this experiment.

**Source**

<http://corneliu.henegar.info/FunNet.htm>

**References**

1. Clement K, Viguerie N, Poitou C, Carette C, Pelloux V, Curat CA, Sicard A, Rome S, Benis A, Zucker JD et al: Weight loss regulates inflammation-related genes in white adipose tissue of obese subjects. *FASEB J* 2004, 18(14):1657-1669.
2. Henegar C, Tordjman J, Achard V, Lacasa D, Cremer I, Guerre-Millo M, Poitou C, Basdevant A, Stich V, Viguerie N, Langin D, Bedossa P, Zucker J-D, Clement K. Adipose tissue transcriptomic signature highlights the pathologic relevance of extracellular matrix in human obesity. In positive revision in *Genome Biology* (October 2007).
3. GEO accession number: GSE9153. <http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE9153>.

**See Also**

[FunNet](#).

---

yeast.cycle

*Yeast Cell Cycle Expression Profiling*

---

**Description**

This data set resulted from microarray experiments performed on yeast (*Saccharomyces cerevisiae*) to study the dynamics of cell cycle related processes in yeast. Please see the reference provided below for further experimental details.

**Usage**

```
data(yeast.cycle)
```

**Format**

A data frame containing temporal series gene expression data obtained from yeast (*Saccharomyces cerevisiae*) cultures performed to explore the cell cycle.

**Source**

<http://corneliu.henegar.info/FunNet.htm>

**References**

Spellman P, Sherlock G, Zhang M, Iyer V, Anders K, Eisen M, Brown P, Botstein D, Futcher B. Comprehensive identification of cell cycle-regulated genes of the yeast *Saccharomyces cerevisiae* by microarray hybridization. *Mol Biol Cell* 1998; 9(12):3273-97.

**See Also**

[FunNet](#).

---

yeast.dna

*Expression Profiling of the Yeast after Irradiation*

---

**Description**

This data set resulted from microarray experiments performed on irradiated yeast (*Saccharomyces cerevisiae*) to study the dynamics of DNA repairing processes induced by the damages inflicted by ionizing radiation in yeast. Please see the reference provided below for further experimental details.

**Usage**

```
data(yeast.dna)
```

**Format**

A data frame containing temporal series gene expression data obtained from irradiated yeast (*Saccharomyces cerevisiae*).

**Source**

<http://corneliu.henegar.info/FunNet.htm>

**References**

Gasch A, Huang M, Metzner S, Botstein D, Elledge S, Brown P. Genomic expression responses to DNA-damaging agents and the regulatory role of the yeast ATR homolog Mec1p. *Mol Biol Cell* 2001; 12(10):2987-3003.

**See Also**

[FunNet](#).

# Index

## \*Topic **cluster**

Annotations, [4](#)  
FunNet, [7](#)  
FunNet.R-package, [2](#)

## \*Topic **datasets**

bypass, [6](#)  
obese, [13](#)  
yeast.cycle, [14](#)  
yeast.dna, [15](#)

Annotations, [4](#)  
annotations, [2](#), [12](#)  
annotations (*Annotations*), [4](#)

bypass, [6](#)

cluster, [3](#), [13](#)

FunNet, [5](#), [6](#), [7](#), [14](#), [15](#)  
FunNet.R-package, [2](#)

obese, [13](#)

yeast.cycle, [14](#)  
yeast.dna, [15](#)