Package 'Mulder2012'

September 23, 2012

Type Package

Title Predicting functional networks and modules of chromatin factors
controlling adult stem cell fate from RNA interference screens

Version 0.2.1

Date 2012-05-30

Author Xin Wang <Xin.Wang@cancer.org.uk>, Flo-
rian Markowetz <Florian.Markowetz@cancer.org.uk>

Maintainer Xin Wang <Xin.Wang@cancer.org.uk>

Depends R (>= 2.14), org.Hs.eg.db, KEGG.db

Imports igraph, HTSanalyzeR, PANR, RedeR, pvclust, MASS

Suggests snow

Description This package provides functions to reproduce results and figures in Mulder K. et. al. pub-
lished in Nature Cell Biology 2012 and Wang X. et. al. published in PLoS Computational Biol-
ogy 2012.

License Artistic-2.0

biocViews Bioinformatics

Collate Mulder2012_fun.R Mulder2012.R Arora2010_fun.R Arora2010.R

LazyLoad yes

R topics documented:

1

Arora2010.BMfitting   *Fitting a beta-mixture model to association scores*

Description

The function fits a three-beta mixture model to functional association scores computed from rich
phenotyping screens in Wang X et al., 2012 for RNAi screening data of Ewing's sarcoma in Arora
S et al., 2010.

Usage

```
Arora2010.BMfitting(pheno, model="global", metric="cosine", nPerm=20)
```

Arguments

pheno	a numeric matrix of z-scores with rows and columns corresponding to samples and genes, respectively.
model	a character value specifying whether the original (if global) or extended (if stratified) model is used.
metric	a character value specifying the metric to compute similarity scores. Currently, 'cosine' and 'correlation' are supported (see assoScore for more details).
nPerm	a positive numeric or integer value specifying the number of permutations.

Details

The function implements beta-mixture modelling in the application to Ewing's sarcoma. The func-
tion first initiates an object of class BetaMixture with inputted pheno, model and metric. The
'x' component of the beta-mixture model is then fitted to permuted data for nPerm times, and the
estimated shape parameters of beta distributions are summarized by taking the median values. After
fixing the shape parameters of the 'x' component, the function invokes fitBM to fit the global or
stratified beta-mixture model.

Value

an object of class BetaMixture with beta-mixture modelling results stored at slot result

Author(s)

Xin Wang <xw264@cam.ac.uk>

References

Arora S, Gonzales IM, Hagelstrom RT, et al. (2010). RNAi phenotype profiling of kinases identifies potential therapeutic targets in Ewing's sarcoma. Molecular Cancer, 9(1), 218.

Wang X, Castro MA, Mulder KW and Markowetz F (2012), Posterior association networks and enriched functional gene modules inferred from rich phenotypic perturbation screens, PLoS Computational Biology, doi:10.1371/journal.pcbi.1002566.

See Also

[BetaMixture](), [fitNULL](), [fitBM]()

Examples

```
## Not run:
data(Arora2010, package="Mulder2012")
BMfitting(pheno=Arora2010, model="global", metric="cosine", nPerm=20)

## End(Not run)
```

Arora2010.fig	*A pipeline function to reproduce the main figures of network analyses in Wang et al 2012 for Arora S et al 2010*

Description

This is a pipeline function for the user to reproduce the main figures regarding network analyses in Wang X et al., 2012 for RNAi screening data of Ewing's sarcoma in Arora S et al., 2010.

Usage

```
Arora2010.fig(what="ALL")
```

Arguments

what   a character value specifying what to plot: 'NULLfitting', 'BMfitting', 'sigMod' or 'ALL' (see details).

Details

The user can run this function with a proper argument to reproduce the main figures regarding network analyses in Wang X et al., 2012 for RNAi screening data of Ewing's sarcoma in Arora S et al., 2010.

These are possible options to plot:

'NULLfitting': a figure illustrating the fitting results of a beta distribution to permuted RNAi screening data.

'BMfitting': a figure illustrating the fitting results of a beta-mixture model to the real RNAi screening data.

'sigMod': significant functional modules identified by in `pvclust`.

'pathway': a figure illustrating significantly overrepresented KEGG pathways.

Author(s)

Xin Wang <xw264@cam.ac.uk>

References

Arora S, Gonzales IM, Hagelstrom RT, et al. (2010). RNAi phenotype profiling of kinases identifies potential therapeutic targets in Ewing's sarcoma. Molecular Cancer, 9(1), 218.

Wang X, Castro MA, Mulder KW and Markowetz F (2012), Posterior association networks and enriched functional gene modules inferred from rich phenotypic perturbation screens, PLoS Computational Biology, doi:10.1371/journal.pcbi.1002566.

See Also

`pvclust`, `RedeR`

Examples

```
## Not run:
Arora2010.fig("BMfitting")

## End(Not run)
```

Arora2010.hypergeo　　　　　　*Hypergeometric tests for pathway analysis*

Description

The function is used to perform hypergeometric tests of overrepresentation of KEGG pathways in inferred functional modules in the application to Ewing's sarcoma.

Usage

```
Arora2010.hypergeo(pan, mod.pval.cutoff=0.05, mod.size.cutoff=4,
avg.degree.cutoff=0.5, filter.effects=TRUE)
```

Arguments

pan　　　　　　　　an object of class PAN, which includes the inferred posterior association network and functional modules.

mod.pval.cutoff
　　　　　　　　　　a numeric value specifying the p-value cutoff for selecting significant functional modules.

mod.size.cutoff
　　　　　　　　　　an integer value specifying the minimal size of functional modules.

avg.degree.cutoff
　　　　　　　　　　a numeric value specifying the cutoff of module density, which is the ratio of predicted significant associations to all possible associations.

filter.effects   a logical value specifying whether or not to filter out modules that are not of interest. In this application to Ewing's sarcoma, modules associated with positive loss-of-function can be filtered out because we are only interested in modules related to inhibiting growth and proliferation of cancer cells.

Details

The function is used to do pathway analyses in the application of PANs to Ewing's sarcoma. Hypergeometric tests are performed to test the overrepresentation of KEGG pathways in inferred functional modules. Four arguments can be specified to filter modules. More details about module filtering procedures can be found in the Methods section of Wang X et al., 2012.

Value

a list including hypergeometric test result for each module.

Author(s)

Xin Wang <xw264@cam.ac.uk>

References

Arora S, Gonzales IM, Hagelstrom RT, et al. (2010). RNAi phenotype profiling of kinases identifies potential therapeutic targets in Ewing's sarcoma. Molecular Cancer, 9(1), 218.

Wang X, Castro MA, Mulder KW and Markowetz F (2012), Posterior association networks and enriched functional gene modules inferred from rich phenotypic perturbation screens, PLoS Computational Biology, doi:10.1371/journal.pcbi.1002566.

See Also

PAN, infer

Examples

```
## Not run:
data(pan_Arora2010, package="Mulder2012")
Arora2010.hypergeo(pan_Arora2010, mod.pval.cutoff=0.05, mod.size.cutoff=4,
avg.degree.cutoff=0.5, filter.effects=TRUE)

## End(Not run)
```

Arora2010.InferPAN   *Infer a posterior association network given a fitted BetaMixture object*

Description

The function infers a functional association network given an object of class BetaMixture that has been fitted.

Usage

```
Arora2010.InferPAN(bm, type="SNR", log=TRUE, sign=TRUE, cutoff=log(10),
filter=FALSE)
```

Arguments

bm	an object of S4 class BetaMixture, which models the densities of first-order similarities between genes (see BetaMixture).
type	a character value giving the type of edge weights: signal-to-noise ratio ('SNR'), posterior probability ratio ('PPR') or posterior probability ('PP')
log	a logical value specifying whether or not to compute logrithms for edge weights.
sign	a logical value specifying whether a signed graph should be inferred. It is only used when type='SNR'.
cutoff	a numeric value giving the threshold to tell the significance of an edge.
filter	a logical value specifying whether or not to filter out genes without any significant association with all the other genes.

Details

This function first initiates an object of class PAN with bm, which is an object of class BetaMixture that is supposed to have been fitted. The function infer is then invoked to infer a posterior association network given the inputted arguments type, log, sign, cutoff and filter.

Value

an object of class PAN with the inferred posterior association network stored at slot graph

Author(s)

Xin Wang <xw264@cam.ac.uk>

References

Arora S, Gonzales IM, Hagelstrom RT, et al. (2010). RNAi phenotype profiling of kinases identifies potential therapeutic targets in Ewing's sarcoma. Molecular Cancer, 9(1), 218.

Wang X, Castro MA, Mulder KW and Markowetz F (2012), Posterior association networks and enriched functional gene modules inferred from rich phenotypic perturbation screens, PLoS Computational Biology, doi:10.1371/journal.pcbi.1002566.

See Also

PAN, infer

Examples

```
## Not run:
data(bm_Arora2010, package="Mulder2012")
Arora2010.InferPAN(bm=bm_Arora2010, type="SNR", log=TRUE, sign=TRUE, cutoff=
log(10), filter=FALSE)

## End(Not run)
```

`Arora2010.module.visualize`

Infer a posterior association network given a fitted BetaMixture object

---

Description

The function represent inferred significant functional modules in `RedeR`, which is a bioconductor package for network visualization.

Usage

```
Arora2010.module.visualize(rdp, pan, mod.pval.cutoff=0.05, mod.size.cutoff=4,
avg.degree.cutoff=0.5, filter.effects=TRUE)
```

Arguments

`rdp`	an object of class `RedPort`, which builds the interface between R and RedeR Java server.
`pan`	an object of class `PAN`, which includes the inferred posterior association network and functional modules.
`mod.pval.cutoff`	
	a numeric value specifying the p-value cutoff for selecting significant functional modules.
`mod.size.cutoff`	
	an integer value specifying the minimal size of functional modules.
`avg.degree.cutoff`	
	a numeric value specifying the cutoff of module density, which is the ratio of predicted significant associations to all possible associations.
`filter.effects`	a logical value specifying whether or not to filter out modules that are not of interest. In this application to Ewing's sarcoma, modules associated with positive loss-of-function can be filtered out because we are only interested in modules related to inhibiting growth and proliferation of cancer cells.

Details

The inferred functional modules are represented in a very compact layout in `RedeR`, as many of them are nested to each other due to the nature of hierarchical clustering. Four arguments can be specified to filter modules. More details about module filtering procedures can be found in the Methods section of Wang X et al., 2012.

Author(s)

Xin Wang <xw264@cam.ac.uk>

References

Arora S, Gonzales IM, Hagelstrom RT, et al. (2010). RNAi phenotype profiling of kinases identifies potential therapeutic targets in Ewing's sarcoma. Molecular Cancer, 9(1), 218.

Wang X, Castro MA, Mulder KW and Markowetz F (2012), Posterior association networks and enriched functional gene modules inferred from rich phenotypic perturbation screens, PLoS Computational Biology, doi:10.1371/journal.pcbi.1002566.

See Also

RedeR

Examples

```
## Not run:
library(RedeR)
rdp <- RedPort('MyPort')
calld(rdp)
data(pan_Arora2010, package="Mulder2012")
Arora2010.module.visualize(rdp, pan_Arora2010, mod.pval.cutoff=0.05,
mod.size.cutoff=4, avg.degree.cutoff=0.5)

## End(Not run)
```

Arora2010.ModuleSearchByPvclust

Searching for enriched functional modules by pvclust

Description

This function invokes pvclustModule to search for enriched functional gene modules by hierarchical clustering with multiscale bootstrap resampling (using R package pvclust).

Usage

```
Arora2010.ModuleSearchByPvclust(pan, nboot=1000, metric="cosine2", hclustMethod=
"average", filter=TRUE)
```

Arguments

pan	an object of class PAN, in which a posterior association network has already been inferred.
nboot	a numeric value giving the number of bootstraps for pvclust.
metric	a character value specifying which distance metric to use for the hierarchical clustering: 'cosine' (first-order), 'cosine2' (second-order), 'correlation', 'uncentered', 'abscor' or those allowed by the argument 'method' in dist.
hclustMethod	the agglomerative method used in hierarchical clustering: 'average', 'ward', 'single', 'complete', 'mcquitty', 'median' or 'centroid' (see the argument method in hclust for more details).
filter	a logical value specifying whether or not to filter out screening data of genes without significant associations with all the other genes.

Value

This function will return an object of class PAN with inferred gene modules (modules$clusters) and corresponding p-values (modules$pval) updated in slot 'modules'.

Author(s)

Xin Wang <xw264@cam.ac.uk>

References

Arora S, Gonzales IM, Hagelstrom RT, et al. (2010). RNAi phenotype profiling of kinases identifies potential therapeutic targets in Ewing's sarcoma. Molecular Cancer, 9(1), 218.

Wang X, Castro MA, Mulder KW and Markowetz F (2012), Posterior association networks and enriched functional gene modules inferred from rich phenotypic perturbation screens, PLoS Computational Biology, doi:10.1371/journal.pcbi.1002566.

See Also

`PAN`, `pvclustModule`

Examples

```
## Not run:
data(pan_Arora2010, package="Mulder2012")
pan_Arora2010<-ModuleSearchByPvclust(pan=pan_Arora2010, nboot=10000, metric=
"cosine2", hclustMethod="average", filter=TRUE)
summarize(pan_Arora2010, what="module")

## End(Not run)
```

Arora2010.pipeline	*A pipeline to do network analyses in Wang et al 2012 for Arora S et al 2010*

Description

This pipeline function reproducing results in Wang X et al., 2012 for RNAi screening data of Ewing's sarcoma in Arora S et al., 2010.

Usage

```
Arora2010.pipeline(par4BM=list(model="global", metric="cosine", nPerm=20),
par4PAN=list(type="SNR", log=TRUE, sign=TRUE, cutoff=log(10), filter=FALSE),
par4ModuleSearch=list(nboot=10000, metric="cosine2", hclustMethod="average",
filter=FALSE))
```

Arguments

par4BM	a list of parameters for beta-mixture modelling (see `Arora2010.BMfitting`) for details.
par4PAN	a list of parameters for inferring a posterior association network (see `Arora2010.InferPAN`) for details.
par4ModuleSearch	
	a list of parameters for seaching enriched functional gene modules (see `Arora2010.ModuleSearchByI` for details.

Details

This function reproduces results of network inference in Wang X et al., 2012 for RNAi screening data of Ewing's sarcoma in Arora S et al., 2010.:

(1) Beta-mixture modelling of functional gene associations (see `Arora2010.BMfitting`)) for details.

(2) Inferring a posterior association network (see `Arora2010.InferPAN`)) for details.

(3) Searching enriched gene modules by pvclust (see `Arora2010.ModuleSearchByPvclust`)) for details.

(4) Pathway analyses (see `Arora2010.hypergeo`)) for details.

Value

All results will be saved in the current work directory.

Author(s)

Xin Wang <xw264@cam.ac.uk>

References

Arora S, Gonzales IM, Hagelstrom RT, et al. (2010). RNAi phenotype profiling of kinases identifies potential therapeutic targets in Ewing's sarcoma. Molecular Cancer, 9(1), 218.

Wang X, Castro MA, Mulder KW and Markowetz F (2012), Posterior association networks and enriched functional gene modules inferred from rich phenotypic perturbation screens, PLoS Computational Biology, doi:10.1371/journal.pcbi.1002566.

See Also

`Arora2010.BMfitting`, `Arora2010.InferPAN`, `Arora2010.ModuleSearchByPvclust`, `Arora2010.hypergeo`

Examples

```
## Not run:
Arora2010.pipeline()

## End(Not run)
```

---

data-Arora2010   *Raw and processed data for the application of PANs to Ewing's sar-*
                                        coma

Description

The data used in this application is a matrix (572 x 8) of Z-scores from high-throughput RNAi screens run in duplicates targeting 572 human kinases in four Ewing's sarcoma cell lines: TC-32, TC-71, SK-ES-1 and RD-ES. In these phenotyping screens, viability was assessed using a luminescence-based cell to quantify each gene's function in cancer cell growth and proliferation. The screening data was corrected for plate row variations and normalized using Z-score method as described in Arora S et al., 2010.

`Arora2010` is a matrix of z-scores summarized from the raw screening data in Arora2010.

bm_Arora2010 is an object of `BetaMixture` including fitting results using the global BM model.

pan_Arora2010 is an object of `PAN` including the inferred functional network and modules identified by `pvclust`.

pw.rslt is a list of hypeogeometric test results for overrepresentation of KEGG pathways in identified modules.

Usage

```
##see example for details
```

Author(s)

Xin Wang <xw264@cam.ac.uk>

References

Arora S, Gonzales IM, Hagelstrom RT, et al. (2010). RNAi phenotype profiling of kinases identifies potential therapeutic targets in Ewing's sarcoma. Molecular Cancer, 9(1), 218.

Examples

```
data(Arora2010, package="Mulder2012")
dim(Arora2010)
data(bm_Arora2010, package="Mulder2012")
bm_Arora2010
data(pan_Arora2010, package="Mulder2012")
pan_Arora2010
```

---

data-Mulder2012 *Raw and processed data for the application to epidermal stem cells*

Description

The data set we use here comes from RNA interference screening on 332 chromatin modifiers under five different conditions: vehicle, AG1478, BMP2/7, AG1478 BMP2/7 and serum stimulation (details in Mulder K et al., 2012).

rawScreenData and rawScreenAnnotation are the raw RNAi screening data and annotations, respectively.

Mulder2012 is a matrix of z-scores summarized from the raw data and annotation.

PINdbProteins and PPI are the original PINdb database and extracted protein-protein interactions matrix, respectively. proteinModules is a list of protein complexes in the PINdb database.

PPIEnrich is a list of enrichment analyses results including: gs–a matrix of running scores of the real and permuted data in the enrichment analyses results. pval–p-values derived from enrichment analyses.

bm_Mulder2012 and bm_ext_Mulder2012 is an object of `BetaMixture` including fitting results using the global and stratified BM model, respectively.

pan_Mulder2012 and pan_ext_Mulder2012 is an object of `PAN` including the inferred functional network and modules identified by `pvclust` with and without incorporating prior information, respectively.

Usage

```
##see example for details
```

Author(s)

Xin Wang <xw264@cam.ac.uk>

References

Mulder KW, Wang X, Escriu C, et al. (2012). Diverse epigenetic strategies interact to control epidermal differentiation. Nature Cell Biology. doi:10.1038/ncb2520.

Examples

```
data(Mulder2012, package="Mulder2012")
dim(Mulder2012)
data(bm_Mulder2012, package="Mulder2012")
bm_Mulder2012
data(pan_Mulder2012, package="Mulder2012")
pan_Mulder2012
```

---

dist.pvclust4PAN   *Compute distances between gene phenotypes for pvclust*

Description

This is a function for PAN to compute distances between phenotypes of two genes. This function modifies the function dist.pvclust from R package pvclust by adding in first- and second-order cosine distances.

Usage

```
##see package \code{\link[pvclust:pvclust]{pvclust}} for details
```

Author(s)

Xin Wang

GSEARandomWalkFig   *Plot a figure illustrating enrichment analysis results*

Description

The function invokes functions in R package HTSanalyzeR to plot results of enrichment analysis of functional interactions for protein-protein interactions.

Usage

```
GSEARandomWalkFig(pheno, PPI, bm, what="pos")
```

Arguments

pheno	a numeric matrix of z-scores with rows and columns corresponding to samples and genes, respectively.
PPI	an adjacency matrix of protein-protein interactions extracted from the PINdb (details in `Mulder2012.PPIPre`).
bm	an object of S4 class `BetaMixture`, which models functional interactions between genes (see `BetaMixture`).
what	a character value specifying which component should be plotted: 'pos' (positive association), 'neg' (negative association) or 'none' (lack of association).

Author(s)

Xin Wang <xw264@cam.ac.uk>

References

Mulder KW, Wang X, Escriu C, et al. (2012). Diverse epigenetic strategies interact to control epidermal differentiation. Nature Cell Biology. doi:10.1038/ncb2520.

Wang X, Castro MA, Mulder KW and Markowetz F (2012), Posterior association networks and enriched functional gene modules inferred from rich phenotypic perturbation screens, PLoS Computational Biology, doi:10.1371/journal.pcbi.1002566.

See Also

gseaScores, gseaPlots

Examples

```
## Not run:
data(Mulder2012, package="Mulder2012")
data(Mulder2012.PPI, package="Mulder2012")
data(bm_Mulder2012, package="Mulder2012")
GSEARandomWalkFig(Mulder2012, PPI, bm_Mulder2012, what="pos")

## End(Not run)
```

Mulder2012.BMfitting   *Fitting a beta-mixture model to association scores*

Description

The function fits a three-beta mixture model to functional association scores computed from rich phenotyping screens in Mulder K et al., 2012 and Wang X et al., 2012.

Usage

```
Mulder2012.BMfitting(pheno, model="global", metric="cosine", nPerm=100)
Mulder2012.BMfitting.extended(pheno, model="stratified", metric="cosine",
nPerm=20, partition)
```

Arguments

pheno	a numeric matrix of z-scores with rows and columns corresponding to samples and genes, respectively.
model	a character value specifying whether the original (if global) or extended (if stratified) model is used.
metric	a character value specifying the metric to compute similarity scores. Currently, 'cosine' and 'correlation' are supported (see assoScore for more details).
nPerm	a positive numeric or integer value specifying the number of permutations.
partition	a numeric of gene partition labels (e.g. c(rep(1, 100), rep(2, 20)) is a valid vector of partition labels for a vector of associations falling into two categories of interaction types 1 and 2).

Details

The two functions implement beta-mixture modelling in the application to epidermal stem cells. Both functions first initiate an object of class BetaMixture with inputted pheno, model and metric (and partition for the stratified model). The 'x' component of the beta-mixture model is then fitted to permuted data for nPerm times, and the estimated shape parameters of beta distributions are summarized by taking the median values. After fixing the shape parameters of the 'x' component, the function invokes fitBM to fit the global or stratified beta-mixture model.

Value

an object of class BetaMixture with beta-mixture modelling results stored at slot result

Author(s)

Xin Wang <xw264@cam.ac.uk>

References

Mulder KW, Wang X, Escriu C, et al. (2012). Diverse epigenetic strategies interact to control epidermal differentiation. Nature Cell Biology. doi:10.1038/ncb2520.

Wang X, Castro MA, Mulder KW and Markowetz F (2012), Posterior association networks and enriched functional gene modules inferred from rich phenotypic perturbation screens, PLoS Computational Biology, doi:10.1371/journal.pcbi.1002566.

See Also

BetaMixture, fitNULL, fitBM

Examples

```
## Not run:
data(Mulder2012, package="Mulder2012")
BMfitting(pheno=Mulder2012, model="global", metric="cosine", nPerm=100)

## End(Not run)
```

| Mulder2012.fig | *A pipeline function to reproduce the main figures of network analyses in Mulder K et al 2012 and Wang et al 2012* |

Description

This is a pipeline function for the user to reproduce the main figures regarding network analyses in Mulder K et al., 2012 and Wang X et al., 2012.

Usage

```
Mulder2012.fig(what="ALL")
```

Arguments

what   a character value specifying what to plot: 'NULLfitting', 'BMfitting', 'PPIenrich', 'sigMod', 'selMod' or 'ALL' (see details).

Details

The user can run this function with a proper argument to reproduce the main figures regarding network analyses in Mulder K et al. 2012 and Wang X et al. 2012.

These are the possible options to plot:

'NULLfitting': a figure illustrating the fitting results of a beta distribution to permuted RNAi screening data.

'BMfitting': a figure illustrating the fitting results of a beta-mixture model to the real RNAi screening data.

'PPIenrich': figures describing enrichment results of posterior probabilities in protein-protein interactions.

'sigMod': significant functional modules identified by in `pvclust`.

'selMod': a figure of the module selected for further validation using combinatorial knock-down experiments.

Author(s)

Xin Wang <xw264@cam.ac.uk>

References

Mulder KW, Wang X, Escriu C, et al. (2012). Diverse epigenetic strategies interact to control epidermal differentiation. Nature Cell Biology. doi:10.1038/ncb2520.

Wang X, Castro MA, Mulder KW and Markowetz F (2012), Posterior association networks and enriched functional gene modules inferred from rich phenotypic perturbation screens, PLoS Computational Biology, doi:10.1371/journal.pcbi.1002566.

See Also

`pvclust`, `RedeR`

Examples

```
## Not run:
Mulder2012.fig("BMfitting")

## End(Not run)
```

Mulder2012.InferPAN   *Infer a posterior association network given a fitted BetaMixture object*

Description

The function infers a functional association network given an object of class BetaMixture that has been fitted.

Usage

```
Mulder2012.InferPAN(bm, type="SNR", log=TRUE, sign=TRUE, cutoff=log(10),
filter=FALSE)
```

Arguments

bm	an object of S4 class BetaMixture, which models the densities of first-order similarities between genes (see BetaMixture).
type	a character value giving the type of edge weights: signal-to-noise ratio ('SNR'), posterior probability ratio ('PPR') or posterior probability ('PP')
log	a logical value specifying whether or not to compute logrithms for edge weights.
sign	a logical value specifying whether a signed graph should be inferred. It is only used when type='SNR'.
cutoff	a numeric value giving the threshold to tell the significance of an edge.
filter	a logical value specifying whether or not to filter out genes without any significant association with all the other genes.

Details

This function first initiates an object of class PAN with bm, which is an object of class BetaMixture that is supposed to have been fitted. The function infer is then invoked to infer a posterior association network given the inputted arguments type, log, sign, cutoff and filter.

Value

an object of class PAN with the inferred posterior association network stored at slot graph

Author(s)

Xin Wang <xw264@cam.ac.uk>

References

Mulder KW, Wang X, Escriu C, et al. (2012). Diverse epigenetic strategies interact to control epidermal differentiation. Nature Cell Biology. doi:10.1038/ncb2520.

Wang X, Castro MA, Mulder KW and Markowetz F (2012), Posterior association networks and enriched functional gene modules inferred from rich phenotypic perturbation screens, PLoS Computational Biology, doi:10.1371/journal.pcbi.1002566.

See Also

PAN, infer

Examples

```
## Not run:
data(bm_ext_Mulder2012, package="Mulder2012")
Mulder2012.InferPAN(bm=bm_ext_Mulder2012, type="SNR", log=TRUE, sign=TRUE,
cutoff=log(10), filter=FALSE)

## End(Not run)
```

Mulder2012.module.visualize

Infer a posterior association network given a fitted BetaMixture object

---

Description

The function represent inferred significant functional modules in RedeR, which is a bioconductor package for network visualization.

Usage

```
Mulder2012.module.visualize(rdp, pan, mod.pval.cutoff=0.05, mod.size.cutoff=4,
avg.degree.cutoff=0.5, edgeWidthLeg=TRUE, filter.effects=TRUE)
```

Arguments

rdp an object of class RedPort, which builds the interface between R and RedeR Java server.

pan an object of class PAN, which includes the inferred posterior association network and functional modules.

mod.pval.cutoff
a numeric value specifying the p-value cutoff for selecting significant functional modules.

mod.size.cutoff
an integer value specifying the minimal size of functional modules.

avg.degree.cutoff
a numeric value specifying the cutoff of module density, which is the ratio of predicted significant associations to all possible associations.

edgeWidthLeg a logical value specifying whether or not the legend of edge width should be created.

filter.effects a logical value specifying whether or not to filter out modules that are not of
                interest. In this application to epidermal stem cells, modules associated with
                negative loss-of-function can be filtered out because we are only interested in
                modules related to stem cell self-renewal.

Details

The inferred functional modules are represented in a very compact layout in RedeR, as many of
them are nested to each other due to the nature of hierarchical clustering. Four arguments can be
specified to filter modules. More details about module filtering procedures can be found in the
Methods section of Wang X et al., 2012.

Author(s)

Xin Wang <xw264@cam.ac.uk>

References

Arora S, Gonzales IM, Hagelstrom RT, et al. (2010). RNAi phenotype profiling of kinases identifies
potential therapeutic targets in Ewing's sarcoma. Molecular Cancer, 9(1), 218.

Wang X, Castro MA, Mulder KW and Markowetz F (2012), Posterior association networks and
enriched functional gene modules inferred from rich phenotypic perturbation screens, PLoS Com-
putational Biology, doi:10.1371/journal.pcbi.1002566.

See Also

[RedeR](#)

Examples

```
## Not run:
library(RedeR)
rdp <- RedPort('MyPort')
calld(rdp)
data(pan_ext_Mulder2012, package="Mulder2012")
Mulder2012.module.visualize(rdp, pan_ext_Mulder2012, mod.pval.cutoff=0.05,
mod.size.cutoff=4, avg.degree.cutoff=0.5)

## End(Not run)
```

Mulder2012.ModuleSearchByPvclust

Searching for enriched functional modules by pvclust

Description

This function invokes [pvclustModule](#) to search for enriched functional gene modules by hierarchi-
cal clustering with multiscale bootstrap resampling (using R package pvclust).

Usage

```
Mulder2012.ModuleSearchByPvclust(pan, nboot=10000, metric="cosine2",
hclustMethod="average", filter=TRUE)
```

Arguments

pan	an object of class PAN, in which a posterior association network has already been inferred.
nboot	a numeric value giving the number of bootstraps for pvclust.
metric	a character value specifying which distance metric to use for the hierarchical clustering: 'cosine' (first-order), 'cosine2' (second-order), 'correlation', 'uncentered', 'abscor' or those allowed by the argument 'method' in dist.
hclustMethod	the agglomerative method used in hierarchical clustering: 'average', 'ward', 'single', 'complete', 'mcquitty', 'median' or 'centroid' (see the argument method in hclust for more details).
filter	a logical value specifying whether or not to filter out screening data of genes without significant associations with all the other genes.

Value

This function will return an object of class PAN with inferred gene modules (modules$clusters) and corresponding p-values (modules$pval) updated in slot 'modules'.

Author(s)

Xin Wang <xw264@cam.ac.uk>

References

Mulder KW, Wang X, Escriu C, et al. (2012). Diverse epigenetic strategies interact to control epidermal differentiation. Nature Cell Biology. doi:10.1038/ncb2520.

Wang X, Castro MA, Mulder KW and Markowetz F (2012), Posterior association networks and enriched functional gene modules inferred from rich phenotypic perturbation screens, PLoS Computational Biology, doi:10.1371/journal.pcbi.1002566.

See Also

PAN, pvclustModule

Examples

```
## Not run:
data(pan_ext_Mulder2012, package="Mulder2012")
pan_ext_Mulder2012<-ModuleSearchByPvclust(pan=pan_ext_Mulder2012, nboot=10000,
metric="cosine2", hclustMethod="average", filter=TRUE)
summarize(pan_ext_Mulder2012, what="module")

## End(Not run)
```

Mulder2012.pipeline   *A pipeline to do network analyses for Mulder 2012*

Description

This pipeline function reproducing results in Mulder 2012.

Usage

```
Mulder2012.pipeline(par4BM=list(model="global", metric="cosine", nPerm=20),
par4PAN=list(type="SNR", log=TRUE, sign=TRUE, cutoff=log(10), filter=FALSE),
par4ModuleSearch=list(nboot=10000, metric="cosine2", hclustMethod="average",
filter=FALSE))
```

Arguments

par4BM   a list of parameters for beta-mixture modelling (see `Mulder2012.BMfitting`)
                    for details.

par4PAN   a list of parameters for inferring a posterior association network (see `Mulder2012.InferPAN`)
                    for details.

par4ModuleSearch

                    a list of parameters for seaching enriched functional gene modules (see `Mulder2012.ModuleSearchBy`
                    for details.

Details

This function reproduces results of network inference in Mulder K et al., 2012:

(1) Beta-mixture modelling of functional gene associations (see `Mulder2012.BMfitting`)) for details.

(2) Inferring a functional association network (see `Mulder2012.InferPAN`)) for details.

(3) Searching enriched gene modules by pvclust (see `Mulder2012.ModuleSearchByPvclust`)) for details.

(4) enrichment analyses of posterior probabilities in protein-protein interactions (see `Mulder2012.PPIenrich`)) for details.

Value

All results will be saved in the current work directory.

Author(s)

Xin Wang <xw264@cam.ac.uk>

References

Mulder KW, Wang X, Escriu C, et al. (2012). Diverse epigenetic strategies interact to control epidermal differentiation. Nature Cell Biology. doi:10.1038/ncb2520.

Wang X, Castro MA, Mulder KW and Markowetz F (2012), Posterior association networks and enriched functional gene modules inferred from rich phenotypic perturbation screens, PLoS Computational Biology, doi:10.1371/journal.pcbi.1002566.

See Also

[Mulder2012.BMfitting](), [Mulder2012.InferPAN](), [Mulder2012.ModuleSearchByPvclust](), [Mulder2012.PPIenrich]()

Examples

```
## Not run:
Mulder2012.pipeline()

## End(Not run)
```

Mulder2012.PPIenrich *Enrichment analyses of posterior probabilities in protein-protein interactions*

Description

This function performs enrichment analyses of protein-protein interactions in posterior probabilities of functional associations belonging to '+', 'x', '-' mixture component.

Usage

```
Mulder2012.PPIenrich(pheno, PPI, bm)
```

Arguments

pheno   a numeric matrix of z-scores with rows and columns corresponding to samples and genes, respectively.

PPI   a binary adjacency matrix of protein-protein interactions.

bm   an object of [BetaMixture]().

Details

After fitting a beta-mixture model to the densities of functional gene associations, a posterior probability for each pair of genes belonging to positive, negative or lack of association component is computed. Using the vector of posterior probabilities belonging to each one of the three components, an enrichment analysis is performed based on R package [HTSanalyzeR](), which does gene set enrichment analysis. The 'gene set' we use here is a vector of protein-protein interactions, which is extracted from PINdb (Proteins Interacting in the Nucleus).

Value

an object of class [BetaMixture]() with beta- mixture modelling results stored at slot `result`

Author(s)

Xin Wang <xw264@cam.ac.uk>

References

Luc PV & Tempst P, PINdb (2004): a database of nuclear protein complexes from human and yeast. Bioinformatics 20(9):1413-5.

Mulder KW, Wang X, Escriu C, et al. (2012). Diverse epigenetic strategies interact to control epidermal differentiation. Nature Cell Biology. doi:10.1038/ncb2520.

Wang X, Castro MA, Mulder KW and Markowetz F (2012), Posterior association networks and enriched functional gene modules inferred from rich phenotypic perturbation screens, PLoS Computational Biology, doi:10.1371/journal.pcbi.1002566.

See Also

BetaMixture, HTSanalyzeR

Examples

```
## Not run:
data(Mulder2012, package="Mulder2012")
data(bm_Mulder2012, package="Mulder2012")
data(Mulder2012_PPI, package="Mulder2012")
PPIenrich(pheno=Mulder2012, PPI=PPI, bm=bm_Mulder2012)

## End(Not run)
```

Mulder2012.PPIPre   *Extracting protein-protein interactions from PINdb*

Description

The function extracts protein-protein interaction information to a binary adjacency matrix from PINdb (Protein Interacting in the Nucleus).

Usage

```
Mulder2012.PPIPre()
```

Value

a list consisting of a binary adjacency matrix of protein-protein interactions and a list of protein modules (complexes).

Author(s)

Xin Wang <xw264@cam.ac.uk>

References

Luc PV & Tempst P, PINdb (2004): a database of nuclear protein complexes from human and yeast. Bioinformatics 20(9):1413-5.

Mulder KW, Wang X, Escriu C, et al. (2012). Diverse epigenetic strategies interact to control epidermal differentiation. Nature Cell Biology. doi:10.1038/ncb2520.

Wang X, Castro MA, Mulder KW and Markowetz F (2012), Posterior association networks and enriched functional gene modules inferred from rich phenotypic perturbation screens, PLoS Computational Biology, doi:10.1371/journal.pcbi.1002566.

Examples

```
## Not run:
ppi<-Mulder2012.PPIPre()

## End(Not run)
```

Mulder2012.RNAiPre　　　　*Preprocessing of the raw RNAi screens*

Description

The function performs normalization and summarization of RNAi screening data for Mulder K et al. 2012 and Wang X et al. 2012.

Usage

```
Mulder2012.RNAiPre(rawScreenData, rawScreenAnnotation)
```

Arguments

rawScreenData   a data frame of the raw screening data.

rawScreenAnnotation
　　　　　　　a data frame of annotations.

Value

a numeric matrix of z-scores with rows and columns specifying samples and genes, respectively.

Author(s)

Xin Wang <xw264@cam.ac.uk>

References

Mulder KW, Wang X, Escriu C, et al. (2012). Diverse epigenetic strategies interact to control epidermal differentiation. Nature Cell Biology. doi:10.1038/ncb2520.

Wang X, Castro MA, Mulder KW and Markowetz F (2012), Posterior association networks and enriched functional gene modules inferred from rich phenotypic perturbation screens, PLoS Computational Biology, doi:10.1371/journal.pcbi.1002566.

Examples

```
## Not run:
Mulder2012<-Mulder2012.RNAiPre()

## End(Not run)
```

---

parPvclust4PAN   *Compute distances between gene phenotypes for pvclust*

Description

This is a function for PAN to do parallel computing for `pvclust`. This function modifies the function
parPvclust from R package pvclust to allow clustering on first- and second-order cosine distances.

Usage

```
##see \code{\link[pvclust:parPvclust]{parPvclust}} for details
```

Author(s)

Xin Wang

Index