Package ‘deTS’

October 13, 2022

Type Package

Title Tissue-Specific Enrichment Analysis

Version 1.0

Date 2019-02-06

Author Guangsheng Pei

Maintainer Guangsheng Pei <`peiguangsheng@gmail.com`>

Imports pheatmap, RColorBrewer

Description Tissue-specific enrichment analysis to assess lists of candidate genes or RNA-Seq expression profiles. Pei G., Dai Y., Zhao Z. Jia P. (2019) deTS: Tissue-Specific Enrichment Analysis to decode tissue specificity. Bioinformatics, In submission.

License GPL (>= 2)

NeedsCompilation no

Depends R (>= 2.10)

Repository CRAN

Date/Publication 2019-02-22 13:30:10 UTC

R topics documented:

`deTS-package` *Tissue-Specific Enrichment Analysis Tissue-Specific Enrichment Analysis*

Description

Tissue-specific enrichment analysis to assess lists of candidate genes and tissue-specific expression decode analysis for RNA-seq data to decode RNA expression matrices tissue heterogeneity.

Details

Since disease and physiological condition are often associated with a specific tissue, understanding the tissue-specific genes (TSG) expression patterns will substantially reduce false discoveries in biomedical research. However, due to cell complexity in human system, heterogeneous tissues are frequently collected. Making it difficult to distinguish gene expression variability and mislead result interpretation. Here, we present deTS, an R package that conducts Tissue-Specific Enrichment Analysis (TSEA) using two built-in reference panels: the Genotype-Tissue Expression (GTEx) data and the ENCyclopedia Of DNA Elements (ENCODE) data. We implemented two major functions in TSEA to assess lists of candidate genes or expression matrices.

The DESCRIPTION file:

Package:	deTS
Type:	Package
Title:	Tissue-Specific Enrichment Analysis
Version:	1.0
Date:	2019-02-06
Author:	Guangsheng Pei
Maintainer:	Guangsheng Pei <peiguangsheng@gmail.com>
Imports:	pheatmap, RColorBrewer
Description:	Tissue-specific enrichment analysis to assess lists of candidate genes or RNA-Seq expression profiles. Pei G.,
License:	GPL (>= 2)

Index of help topics:

```
ENCODE_z_score          ENCODE z-score to define tissue-specific genes
GTEx_t_score            GTEx t-score to define tissue-specific genes
GWAS_gene               Gene symbol query data for single sample
GWAS_gene_multiple      Gene symbol query data for multiple samples
correction_factor       Gene average expression level and standard
                        deviation in GTEx data
deTS-package            Tissue-Specific Enrichment Analysis
                        Tissue-Specific Enrichment Analysis
query_ENCODE            ENCODE raw query data
query_GTEx              GTEx raw query data
```

`tsea.analysis`	Tissue-specific enrichment analysis for query gene list
`tsea.analysis.multiple`	Tissue-specific enrichment analysis for multi query gene lists
`tsea.expression.decode`	Tissue-specific enrichment analysis for RNA-Seq expression profiles
`tsea.expression.normalization`	RNA-Seq expression profiles normalization
`tsea.plot`	Tissue-specific enrichment analysis result heatmap plot
`tsea.summary`	Tissue-specific enrichment analysis result summary

Author(s)

Guangsheng Pei

Maintainer: Guangsheng Pei

References

Pei G., Dai Y., Zhao Z., Jia P. (2019) deTS: Tissue-Specific Enrichment Analysis to decode tissue specificity. Bioinformatics, In submission.

See Also

https://github.com/bsml320/deTS

Examples

```
data(GTEx_t_score)
data(ENCODE_z_score)
library(pheatmap)

data(GWAS_gene)
query_gene_list = GWAS_gene
tsea_t = tsea.analysis(query_gene_list, GTEx_t_score, 0.05,
p.adjust.method = "bonferroni")
tsea_t_summary = tsea.summary(tsea_t)

data(GWAS_gene_multiple)
query_gene_list = GWAS_gene_multiple[,1:2]
tsea_t_multi = tsea.analysis.multiple(query_gene_list,
GTEx_t_score, 0.05, p.adjust.method = "BH")

data(query_GTEx)
query_matrix = query_GTEx[,1:2]
data(correction_factor)
query_mat_zscore_nor = tsea.expression.normalization(query_matrix,
correction_factor, normalization = "z-score")
```

```
tseaed_in_ENCODE = tsea.expression.decode(query_mat_zscore_nor,
ENCODE_z_score, 0.05, p.adjust.method = "BH")
tseaed_in_ENCODE_summary = tsea.summary(tseaed_in_ENCODE)
```

`correction_factor` *Gene average expression level and standard deviation in GTEx data*

Description

Gene average expression level and standard deviation in GTEx data

Usage

```
data("correction_factor")
```

Format

A data frame with 14725 observations on the following 2 variables.

`avg.all` a factor with gene average expression level

`sd.all` a factor with gene standard deviation of expression level

Details

nothing

Source

nothing

References

Pei G., Dai Y., Zhao Z., Jia P. (2019) Tissue-Specific Enrichment Analysis deTS: Tissue-Specific Enrichment Analysis to decode tissue specificity. Bioinformatics, In submission.

Examples

```
data(correction_factor)
```

`ENCODE_z_score` *ENCODE z-score to define tissue-specific genes*

Description

ENCODE z-score matrix to define tissue-specific genes

Usage

```
data("ENCODE_z_score")
```

Format

A data frame with z-score of 14031 genes in 44 ENCODE tissues.

Row is genes symbol and column is tissue names.

Adrenal Gland Body of Pancreas Breast Epithelium Camera-type Eye Cerebellum

C1orf112 -0.674 -0.440 -0.246 3.892 1.333

FGR -0.078 -0.345 0.159 -0.354 -0.407

CFH -0.093 -0.365 -0.134 -0.133 -0.160

FUCA2 3.028 1.467 0.040 0.228 -0.601

NFYA -0.637 -0.872 0.053 2.364 0.619

Details

nothing

Source

nothing

References

Pei G., Dai Y., Zhao Z., Jia P. (2019) deTS: Tissue-Specific Enrichment Analysis to decode tissue specificity. Bioinformatics, In submission.

Examples

```
data(ENCODE_z_score)
```

`GTEx_t_score` *GTEx t-score to define tissue-specific genes*

Description

GTEx t-score matrix to define tissue-specific genes

Usage

```
data("GTEx_t_score")
```

Format

A data frame with t-score of 14725 genes in 47 GTEx tissues.

Row is genes symbol and column is tissue names.

Adipose - Subcutaneous Adipose - Visceral (Omentum) Adrenal Gland Artery - Aorta Artery - Coronary

OR4F5 -0.524 -0.597 0.134 -1.109 -0.588

SAMD11 -9.921 -1.734 3.633 3.595 0.017

KLHL17 -6.812 -4.553 -3.084 -0.744 0.494

PLEKHN1 -7.785 -6.882 -3.915 -6.570 -4.892

C1orf170 -7.113 -6.257 -4.465 -5.897 -4.004

Details

nothing

Source

nothing

References

Pei G., Dai Y., Zhao Z., Jia P. (2019) deTS: Tissue-Specific Enrichment Analysis to decode tissue specificity. Bioinformatics, In submission.

Examples

```
data(GTEx_t_score)
```

`GWAS_gene` *Gene symbol query data for single sample*

Description

An example of input gene symbol query data for single sample tissue-specific enrichment analysis

Usage

```
data("GWAS_gene")
```

Format

The format is:

"A1BG" "A1BG-AS1" "A1CF" "A2M" "A2M-AS1" "A2ML1" "A2MP1" "A3GALT2" "A4GALT" "A4GNT" "AA06" "AAAS" "AACS" "AACSP1" "AADAC" ...

Details

nothing

Source

nothing

References

Pei G., Dai Y., Zhao Z., Jia P. (2019) deTS: Tissue-Specific Enrichment Analysis to decode tissue specificity. Bioinformatics, In submission.

Examples

```
data(GWAS_gene)
```

`GWAS_gene_multiple` *Gene symbol query data for multiple samples*

Description

An example of input gene symbol query data for multiple samples tissue-specific enrichment analysis

Usage

```
data("GWAS_gene_multiple")
```

Format

A data frame with 22134 genes if associated with following 5 neuropsychiatric disorders GWAS traits.

Row is genes symbol and column is sample names.

ADHD ASD BD MDD SCZ

A1BG 0 0 0 0 0

A1BG-AS1 0 0 0 0 0

A1CF 0 1 0 0 0

A2M 0 0 0 0 0

A2M-AS1 0 0 0 0 0

Details

nothing

Source

nothing

References

Pei G., Dai Y., Zhao Z., Jia P. (2019) deTS: Tissue-Specific Enrichment Analysis to decode tissue specificity. Bioinformatics, In submission.

Examples

```
data(GWAS_gene_multiple)
```

`query_ENCODE` *ENCODE raw query data*

Description

An example of RNA-Seq query data from ENCODE data for tissue-specific enrichment analysis

Usage

```
data("query_ENCODE")
```

Format

A data frame with average expression level of 18661 genes in 44 ENCODE tissues.

Row is genes symbol and column is sample names.

Adrenal Gland Body of Pancreas Breast Epithelium Camera-type Eye Cerebellum

TSPAN6 11.64 5.390 11.04 24.65 13.238

TNMD 0.01 0.147 2.24 12.43 0.090

DPM1 18.82 9.812 14.21 24.02 10.505

SCYL3 3.81 2.593 5.63 10.50 3.783

C1orf112 1.64 2.308 2.86 14.61 7.345

Details

nothing

Source

nothing

References

Pei G., Dai Y., Zhao Z., Jia P. (2019) deTS: Tissue-Specific Enrichment Analysis to decode tissue specificity. Bioinformatics, In submission.

Examples

```
data(query_ENCODE)
```

`query_GTEx` *GTEx raw query data*

Description

An example of RNA-Seq query data from GTEx data for tissue-specific enrichment analysis

Usage

```
data("query_GTEx")
```

Format

A data frame with average expression level of 18067 gene in 47 GTEx tissues.

Row is genes symbol and column is sample names.

Adipose - Subcutaneous Adipose - Visceral (Omentum) Adrenal Gland Artery - Aorta Artery - Coronary

OR4F5 0.0317 0.0284 0.0469 0.0133 0.0225

SAMD11 0.4451 2.3056 3.8928 3.5822 2.5632

NOC2L 21.9084 21.0439 19.4613 19.4929 19.8367

KLHL17 4.1406 4.4075 4.4227 5.0840 5.3749

PLEKHN1 0.4531 0.3452 1.1795 0.3081 0.3722

Details

nothing

Source

nothing

References

Pei G., Dai Y., Zhao Z., Jia P. (2019) deTS: Tissue-Specific Enrichment Analysis to decode tissue specificity. Bioinformatics, In submission.

Examples

```
data(query_GTEx)
```

`tsea.analysis` *Tissue-specific enrichment analysis for query gene list*

Description

Tissue-specific enrichment analysis by Fisher's Exact Test for given gene list.

Usage

```
tsea.analysis(query_gene_list, score, ratio = 0.05,
p.adjust.method = "BH")
```

Arguments

`query_gene_list`	a gene symbol list object.
`score`	a gene tissue-specific score matrix, c("GTEx_t_score" or "ENCODE_z_score"), can be loaded by data(GTEx) or data(ENCODE), the default value is recommended "GTEx_t_score".
`ratio`	the threshold to define tissue-specific genes (with top t-score or z-score), the default value is 0.05.
`p.adjust.method`	p.adjust.method, c("holm", "hochberg", "hommel", "bonferroni", "BH", "BY", "fdr", "none")

Details

Tissue-specific enrichment analysis by Fisher's Exact Test for given gene list.

Value

A data frame with p-value of tissue-specific enrichment result.

Rows stand for tissue names and columns stand for sample names.

Note

nothing

Author(s)

Guangsheng Pei

References

Pei G., Dai Y., Zhao Z., Jia P. (2019) deTS: Tissue-Specific Enrichment Analysis to decode tissue specificity. Bioinformatics, In submission.

See Also

https://github.com/bsml320/deTS

Examples

```
data(GWAS_gene)
data(GTEx_t_score)
query_gene_list = GWAS_gene
tsea_t = tsea.analysis(query_gene_list, GTEx_t_score, 0.05,
p.adjust.method = "bonferroni")
```

`tsea.analysis.multiple` *Tissue-specific enrichment analysis for multi query gene lists*

Description

Tissue-specific enrichment analysis by Fisher's Exact Test for multiple gene list.

Usage

```
tsea.analysis.multiple(query_gene_list, score, ratio = 0.05,
p.adjust.method = "BH")
```

Arguments

`query_gene_list`	a 0~1 gene~sample table object, row should be gene symbol, column should be sample name. In the table, gene labeled with 1 indicated it is target gene for a given sample, while 0 indicated it is not target in a given sample.
`score`	a gene tissue-specific score matrix, c("GTEx_t_score" or "ENCODE_z_score"), can be loaded by data(GTEx) or data(ENCODE), the default value is recommended "GTEx_t_score".
`ratio`	the threshold to define tissue-specific genes (with top t-score or z-score), the default value is 0.05.
`p.adjust.method`	p.adjust.method, c("holm", "hochberg", "hommel", "bonferroni", "BH", "BY", "fdr", "none")

Details

Tissue-specific enrichment analysis by Fisher's Exact Test for multiple gene list.

Value

A data frame with p-value of tissue-specific enrichment result.

Rows stand for tissue names and columns stand for sample names.

Note

nothing

Author(s)

Guangsheng Pei

References

Pei G., Dai Y., Zhao Z. Jia P. (2019) deTS: Tissue-Specific Enrichment Analysis to decode tissue specificity. Bioinformatics, In submission.

See Also

https://github.com/bsml320/deTS

Examples

```
data(GWAS_gene_multiple)
data(GTEx_t_score)
query_gene_list = GWAS_gene_multiple
tsea_t_multi = tsea.analysis.multiple(query_gene_list,
GTEx_t_score, 0.05, p.adjust.method = "BH")
```

`tsea.expression.decode` *Tissue-specific enrichment analysis for RNA-Seq expression profiles*

Description

Tissue-specific enrichment analysis to decode whether a given RNA-seq sample (RPKM) with potential confounding effects based on expression profiles.

Usage

```
tsea.expression.decode(query_mat_normalized_score, score,
ratio = 0.05, p.adjust.method = "BH")
```

Arguments

`query_mat_normalized_score`	a normalized RNA-seq RPKM object, which produced by "tsea.expression.normalization".
`score`	a gene tissue-specific score matrix, c("GTEx_t_score" or "ENCODE_z_score"), can be loaded by data(GTEx) or data(ENCODE), the default value is recommended "GTEx_t_score".
`ratio`	the threshold to define tissue-specific genes (with top t-score or z-score), the default value is 0.05.
`p.adjust.method`	p.adjust.method, c("holm", "hochberg", "hommel", "bonferroni", "BH", "BY", "fdr", "none")

Details

Tissue-specific enrichment analysis for RNA-Seq expression profiles.

Value

A data frame with p-value of tissue-specific enrichment result for RNA-Seq expression profiles.

Rows stand for tissue names and columns stand for sample names.

Note

nothing

Author(s)

Guangsheng Pei

References

Pei G., Dai Y., Zhao Z., Jia P. (2019) deTS: Tissue-Specific Enrichment Analysis to decode tissue specificity. Bioinformatics, In submission.

See Also

https://github.com/bsml320/deTS

Examples

```
data(query_GTEx)
query_matrix = query_GTEx[,1:2]
data(correction_factor)
data(ENCODE_z_score)
query_mat_zscore_nor = tsea.expression.normalization(query_matrix,
correction_factor, normalization = "z-score")
tseaed_in_ENCODE = tsea.expression.decode(query_mat_zscore_nor,
ENCODE_z_score, 0.05, p.adjust.method = "BH")
```

`tsea.expression.normalization` *RNA-Seq expression profiles normalization*

Description

To avoid the data bias and adapt better data heterogeneity, before tsea.expression.decode() analysis, the raw discrete RPKM value have to normalized to continuous variable meet the normal distribution before t-test.

Usage

```
tsea.expression.normalization(query_mat,
correction_factor, normalization = "abundance")
```

Arguments

`query_mat`	a RNA-seq RPKM object, row name should be gene symbol, and column name should be sample name.
`correction_factor`	correction_factor, a gene table object contain genes average expression level and standard variance in GTEx database, can be loaded by data(correction_factor).
`normalization`	normalization methods, c("z-score", "abundance")

Details

As RNA-Seq samples are often heterogeneous, before in-depth analysis, it is necessary to decode tissue heterogeneity to avoid samples with confounding effects. However, the raw discrete RPKM value have to normalized to continuous variable meet the normal distribution before t-test.

Value

A data frame with normalized RNA-Seq expression profiles.

Rows stand for tissue names and columns stand for sample names.

Note

nothing

Author(s)

Guangsheng Pei

References

Pei G., Dai Y., Zhao Z., Jia P. (2019) deTS: Tissue-Specific Enrichment Analysis to decode tissue specificity. Bioinformatics, In submission.

See Also

https://github.com/bsml320/deTS

Examples

```
data(query_GTEx)
query_matrix = query_GTEx[,1:2]
data(correction_factor)
query_mat_zscore_nor = tsea.expression.normalization(query_matrix,
correction_factor, normalization = "z-score")
```

`tsea.plot` *Tissue-specific enrichment analysis result heatmap plot*

Description

Heat map plot for tissue-specific enrichment analysis result.

Usage

```
tsea.plot(tsea_result, threshold = 0.05)
```

Arguments

`tsea_result`	the result of tissue-specific enrichment analysis, which produced by "tsea.analysis", "tsea.analysis.multiple" or "tsea.expression.decode".
`threshold`	the p-value threshold to define if the gene list or RNA-seq data enriched in a given tissue, p-value greater than threshold will not be labeled in the plot. The default value is 0.05.

Details

Heat map plot for tissue-specific enrichment analysis result

Value

Heatmap plot

Note

nothing

Author(s)

Guangsheng Pei

References

Pei G., Dai Y., Zhao Z., Jia P. (2019) deTS: Tissue-Specific Enrichment Analysis to decode tissue specificity. Bioinformatics, In submission.

See Also

https://github.com/bsml320/deTS

Examples

```
data(GWAS_gene_multiple)
data(GTEx_t_score)
query_gene_list = GWAS_gene_multiple
tsea_t_multi = tsea.analysis.multiple(query_gene_list,
GTEx_t_score, 0.05, p.adjust.method = "BH")
tsea.plot(tsea_t_multi, 0.05)
```

`tsea.summary` *Tissue-specific enrichment analysis result summary*

Description

Tissue-specific enrichment analysis result summary (list the top 3 most enriched tissues) from the given gene list or RNA-seq expression profiles.

Usage

```
tsea.summary(tsea_result)
```

Arguments

`tsea_result`	the result of tissue-specific enrichment analysis, which produced by "tsea.analysis", "tsea.analysis.multiple" or "tsea.expression.decode".

Details

Tissue-specific enrichment analysis result summary

Value

A data frame with summary result of top 3 most enriched tissues.

Rows stand for sample names and columns stand for top 3 most enriched tissues (with p-value).

Note

nothing

Author(s)

Guangsheng Pei

References

Pei G., Dai Y., Zhao Z., Jia P. (2019) deTS: Tissue-Specific Enrichment Analysis to decode tissue specificity. Bioinformatics, In submission.

See Also

https://github.com/bsml320/deTS

Examples

```
data(query_GTEx)
query_matrix = query_GTEx
data(correction_factor)
data(ENCODE_z_score)
query_mat_zscore_nor = tsea.expression.normalization(query_matrix,
correction_factor, normalization = "z-score")
tseaed_in_ENCODE = tsea.expression.decode(query_mat_zscore_nor,
ENCODE_z_score, 0.05, p.adjust.method = "BH")
tseaed_in_ENCODE_summary = tsea.summary(tseaed_in_ENCODE)
```

Index