

Introduction for FunTFPair package

October 24, 2014

Contents

1	Introduction	1
2	Functions in the package	1
3	Example	2

1 Introduction

Summary: Transcription factors (TFs) are fundamental regulators of gene expression and generally function in a complex and cooperative manner. Identifying context-dependent cooperative TFs is essential for understanding how cells respond to environmental change. The huge amount of various omics data currently available, providing genome-wide physical binding and functional effect information about transcription, holds a great opportunity to study TF cooperativity. Here we developed an R package, FunTFPair, which provides an easy and powerful way to identify condition-specific TF pairs by integrating transcription factor binding sites from ENCODE and gene expression profiles from GEO. Users only need to provide GEO ID that they are interested in. FunTFPair will automatically retrieve the expression profiles of the input GEO ID, get TF targets from ENCODE, and identify TF pairs whose common targets show statistically significant differences under changed experimental conditions or exhibit coordinated transcription in a specific condition. The functional TF pairs and their relative importance will be reported in a TF cooperation network. Two datasets from GEO are used as examples to demonstrate the usage and the reliability of the package.

Availability and Implementation: FunTFPair is implemented in R language. The source code is provided at <https://github.com/slzhao/FunTFPair>. A more detailed vignette is also released along the package.

2 Functions in the package

- *prepareGeoData*: A function to download GDS or GSE data and convert ID if needed;
- *differentialAnalysis*: A function to perform differential analysis and then find the enrichment of target genes;

- *correlationAnalysis*: A function to perform differential analysis for expression data, to see if the shared targets of two tf were significantly changed between two conditions;
- *networkVis*: A function to perform functional TF network visualization;

3 Example

Here we used one dataset (GDS2213) profiles gene expression of HepG2 cells treated by 5-aza-2'-deoxycytidine (5-aza-dC), trichostatin A (TSA) or both, as an example of differential analysis-based functional TF pairs identification. The gene expression profiles under three kinds of treatment were compared with those in the control group, respectively. As a result, 24 TF pairs, involving 8 TFs, were identified to be functionally important under the treatment of 5-aza-dC, which inhibit DNA methylation. Consistently, most of TFs has been reported to be influenced by DNA methylation. More interestingly, HDAC2, an important histone deacetylase, joined the functional TF network when TSA, inhibitor of histone decetylation, was added.

```
library(FunTFPair)

## Loading required package: Biobase
## Loading required package: BiocGenerics
## Loading required package: parallel
##
## Attaching package: 'BiocGenerics'
##
## The following objects are masked from 'package:parallel':
##
##   clusterApply, clusterApplyLB, clusterCall, clusterEvalQ,
##   clusterExport, clusterMap, parApply, parCapply, parLapply,
##   parLapplyLB, parRapply, parSapply, parSapplyLB
##
## The following object is masked from 'package:stats':
##
##   xtabs
##
## The following objects are masked from 'package:base':
##
##   anyDuplicated, append, as.data.frame, as.vector, cbind,
##   colnames, do.call, duplicated, eval, evalq, Filter, Find, get,
##   intersect, is.unsorted, lapply, Map, mapply, match, mget,
##   order, paste, pmax, pmax.int, pmin, pmin.int, Position, rank,
##   rbind, Reduce, rep.int, rownames, sapply, setdiff, sort,
##   table, tapply, union, unique, unlist, unsplit
##
## Welcome to Bioconductor
##
## Vignettes contain introductory material; view with
## 'browseVignettes()'. To cite Bioconductor, see
```

```
##      'citation("Biobase")', and for packages 'citation("pkgname")'.
##
## Setting options('download.file.method.GEOquery'='auto')
## Error:  package or namespace load failed for 'FunTFPair'

example(differentialAnalysis, run.dontrun = TRUE)

## Warning in example(differentialAnalysis, run.dontrun = TRUE): no help found for
'differentialAnalysis'
```

Here the size of nodes indicated the importance of the TF in the cooperative network; while the width of the edges between pairs indicated the negative log 10 p value for the enrichment of shared targets in significantly changed genes.

You can also see the example for correlationAnalysis by the following codes:

```
example(correlationAnalysis, run.dontrun = TRUE)
```