Manual of multiOmicsViz

Jing Wang

October 29, 2019

1 Introduction

multiOmicsViz package can calculate the spearman correlation between the source omics data and other target omics data, identify the significant correlations and plot the significant correlations on the heat map in which the x-axis and y-axis are ordered by the chromosomal location.

2 Environment

multiOmicsViz requires R version 3.3 or later, which can be downloaded from the website http://www.r-project.org/. multiOmicsViz package can be installed as follows.

if (!requireNamespace("BiocManager", quietly=TRUE)) install.packages("BiocManager") install.package("multiOmicsViz")

3 multiOmicsViz

After building up the basic environment mentioned above, the users can install the multiOmicsViz package and use it to analyze networks.

```
> library("multiOmicsViz")
> sourceOmics <- system.file("extdata","sourceOmics.txt",package="multiOmicsViz")
> sourceOmics <- read.table(sourceOmics,header=TRUE,sep="\t",stringsAsFactors=FALSE,check.names=FALS
> targetOmics1 <- system.file("extdata","targetOmics.txt",package="multiOmicsViz")
> targetOmics1 <- read.table(targetOmics1,header=TRUE,sep="\t",stringsAsFactors=FALSE,check.names=FAL
> targetOmicsList <- list()
> targetOmicsList[[1]] <- targetOmics1
> outputfile <- paste(tempdir(),"/heatmap",sep="")
> multiOmicsViz(sourceOmics,"CNA","20",targetOmicsList,"mRNA","All",0.001,outputfile)


Identify the significant correlations...
Plot figure...
null device
          1
```

3.1 Input

This section describes the arguments of the multiOmicsViz function:

1. *sourceOmics*: A R matrix, data.frame or SummarizedExperiment object containing the omics data. The data should contain the row names representing the genes and column names representing the samples.

2. *sourceOmicsName*: The name of the source omics data (e.g. CNA).

3. *chrome_sourceOmics*: The *multiOmicsViz* function will extract the genes in the selected chromosome(s) from genes in the source omics data and then identify and visualize the significant correlations based on the

selected genes. *chrome_sourceOmics* can be one character containing the chromosome name (e.g. "1"), a R vector object containing multiple chromosomes (e.g. c("1","2","3")) or "All" representing all 24 chromosomes.

4. *targetOmicsList*: A R list object containing at most 5 target omics data. Each omics data in the list should be a R matrix, data.frame or SummarizedExperiment object and contain the row names representing the genes and column names representing the samples. There should have multiple overlapping genes among all target omics data and at least 6 overlapping sample between source omics data and each target omics data.

5. *targetOmicsName*: A R vector object containing the name of all target omics data stored in the *targetOmicsList*.

6. *chrome_targetOmics*: The *multiOmicsViz* function will extract the genes in the selected chromosome(s) from the overlapping genes among all target omics data and then identify and visualize the significant correlations based on the selected genes. *chrome_sourceOmics* can be one character containing the chromosome name (e.g. "1"), a R vector object containing multiple chromosomes (e.g. c("1","2","3")) or "All" representing all 24 chromosomes.

7. *fdrThr*: The FDR threshold for identifying the significant correlations.

8. *outputfile*: The output file name.

9. *nThreads*: If *targetOmicsList* contains 2 or 3 omics data, multiOmicsViz will use the parallel computing method to calculate the significant correlations between the source omics data and each of target omics data. *nThreads* is the number of cores used for the parallel computing.

10. *legend*: If *legend* is TRUE, the output heat map will contain the legend.

3.2   Output

If the *targetOmicsList* contains one target omics data, the *multiOmicsViz* function will plot a heat map in which x-axis represents the genes in the source omics data, y-axis represents the genes in the target omics data, x-axis and y-axis are ordered by chromosomal location, each point represents a significant correlation, red color represents the significant positive correlation and blue color represents the significant negative correlation. If the *targetOmicsList* contains multiple target omics data, the *multiOmicsViz* function will not only plot multiple heat maps for each target omics data but also plot mutliple bar charts in which blue bars represent the number of specific significant correlations for the target omics data and black bars represents the number of common significant correlations among all target omics data.

The following figure is one example to use heat maps to compare the effect of CNA (copy number alteration) on mRNA and protein abundance.
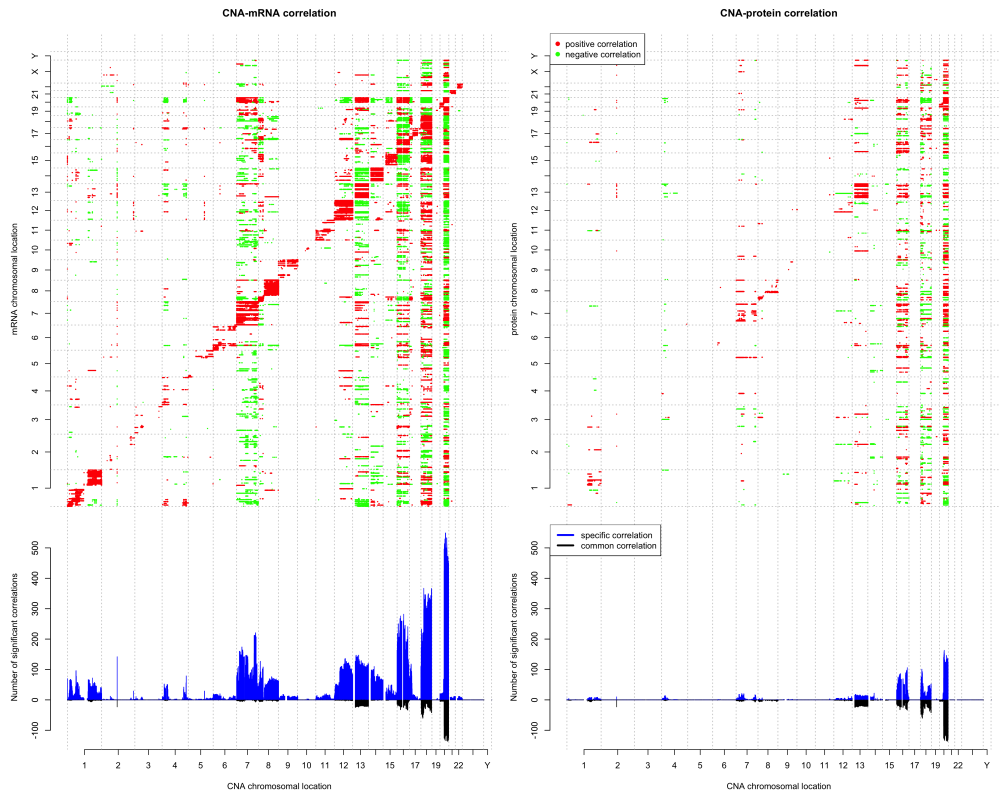
4 calculateCorForTwoMatrices

The calculateCorForTwoMatrices function can Identify the significant correlations between two matrices.

```
> library("multiOmicsViz")
> matrix1 <- system.file("extdata","sourceOmics.txt",package="multiOmicsViz")
> matrix1 <- read.table(matrix1,header=TRUE,sep="\t",stringsAsFactors=FALSE,check.names=FALSE)
> matrix2 <- system.file("extdata","targetOmics.txt",package="multiOmicsViz")
> matrix2 <- read.table(matrix2,header=TRUE,sep="\t",stringsAsFactors=FALSE,check.names=FALSE)
> sig <- calculateCorForTwoMatrices(matrix1=matrix1,matrix2=matrix2,fdr=0.01)
```

4.1 Input

1. *matrix1* and *matrix2*: A R matrix, data.frame or SummarizedExperiment object containing the numeric values. *matrix2* should have at least 6 overlapping samples with *matrix1*.

2. *fdr*: The FDR threshold for identifying significant correlations.

4.2 Output

This function will return a R matrix object containing significant correlations. "1" represents the significant positive correlation, "-1" represents the significant negative correlation and "0" represents no significant correlation.