Package 'MixMAP'

July 2, 2014

Type Package

Title Implements the MixMAP algorithm

Version 1.3.1

Date 2013-10-01

Author Gregory J. Matthews

Maintainer Gregory J. Matthews <gjm112@gmail.com>

Depends R (>= 2.15.1), lme4

Imports

Suggests methods

Description This function implements the MixMAP algorithm, which performs gene level tests of association using data from a previous GWAS or data from a meta-analysis of several GWAS. Conceptually, genes are detected as significant if the collection of p-values within a gene are determined to be collectively smaller than would be observed by chance.

License GPL-3

Collate 'mixmapTest.R' 'mixmapPI.R' 'MixMAPclass.R' 'plot.MixMAP.R' 'summary.MixMAP.R'

NeedsCompilation no

Repository CRAN

Date/Publication 2014-03-10 22:29:59

R topics documented:

`MixMAP-package` *MixMAP: Mixed Modeling of Meta-Analysis P-values*

Description

Mixed Modeling approach based on Meta-analysis p-values for detecting gene level associations from genome wide assocation studies (GWAS) or candidate gene studies.

This package uses raw p-values from previous GWAS, and information about gene-level groupings to search for gene-level assocations between complex disease phenotypes and genetic loci.

Details

Package:	MixMAP
Type:	Package
Version:	1.0
Date:	2012-08-13
License:	GPL-3

This goal of this package is to implement the MixMAP algorithm. The aim of the algorithm is to search for assocations between genes and complex diseases by using individual SNP-level p-values. The user must provides a file with SNP name and SNP p-value as well as gene name, chromosome, and basepair location. The output of the function MixMAP is an object of class 'MixMAP', which contains information on genes that have been detected as being associated with the phenotype of interest. An object of class 'MixMAP' has a plot method associated with it to visually display the result of the MixMAP algorithm in a Manhattan style plot as well as a summary method.

Author(s)

Gregory J. Matthews and Andrea Foulkes

Maintainer: Gregory J. Matthews <gjm112@gmail.com>

References

Foulkes, A.S., Matthews, G.J., Das, U., Ferguson, J., Reilly, M. (2013) "Mixed Modeling of Meta-Analysis P-Values (MixMAP) Suggests Multiple Novel Gene Loci for Low Density Lipoprotein Cholesterol". PLoS ONE 8(2): e54812.

See Also

`lme4-package`

Examples

```
library(MixMAP)
```

```
#Load data
#This data has been prepared to be used as input to the MixMAP function directly
data(MixMAP_example)
str(MixMAP_example)
#Run mixmapTest
MixOut<-mixmapTest(MixMAP_example,pval="GC.Pvalue",snp="MarkerName",
  chr="Chr",coord="Coordinate",gene="Gene")
#Display first ten detected genes
summary(MixOut)
#MixManhattan Plot
plot(MixOut)
```

`MixMAP-class` *MixMAP-class*

Description

Objects of class MixMAP are returned as output from the function MixMAP.

Arguments

`output`	A data.frame containing posterior estimates for all genes, as well as other associated gene level information.
`num.genes.detected`	A numeric vector containing the number of genes detected and the total number of genes.
`detected.genes`	A data.frame containing information on the detected genes. This includes the posterior estimate, variance, and upper bound of the prediction interval fo all genes. Other gene level information that is included is the gene location including chromosome and base pair along with the total number of SNPs within each gene (snpCount), and some summary information regarding the SNP p-values within each gene.
`num.genes.detected`	A numeric vector containing the number of genes detected and the total number of genes.
`lmer.out`	A mer object containing all of the model output information, including parameter estimates, from the lmer function.

Author(s)

Gregory J. Matthews

Examples

```
showClass("MixMAP")
```

mixmapPI *mixmapPI mixmapPI*

Description

This function implements the MixMAP algorithm, which performs gene-level tests of association using data from a previous GWAS or data from a meta-analysis of several GWAS. Conceptually, genes are detected as significant if the collection of p-values within a gene are determined to be collectively smaller than would be observed by chance.

Usage

```
mixmapPI(data.set, pval="pval", snp="snp", gene="gene",
           coord="coord", chr="chr", alpha = 0.05)
```

Arguments

data.set	A data.frame containing the input data. Each observation in this data set is a SNP. This file must contain, at least, the SNP name, p-value for a SNP, group name (usually a gene name) where SNP is located, SNP location coordinate, and chromosome number of the SNP.
pval	A character string with the name of the variable containing the p-values in the data.set. Default is "pval".
snp	A character string with the name of the variable containing the SNP name in the data.set. Default is "snp".
gene	A character string with the name of the variable containing the gene name in the data.set. Default is "gene".
coord	A character string with the name of the variable containing the location coordinate of the SNP in the data.set. Default is "coord".
chr	A character string with the name of the variable containing the chromosome number of the SNP location in the data.set. Default is "chr".
alpha	A numeric scalar indicating the level of significance the user chooses to use for detection. The default is 0.05.

Details

The user must provide a file that includes SNP name, SNP p-value, and a group name (assumed to be gene, but the user can use any grouping they choose), among other inputs.

Note about SNPs in genes: It is possible for individual SNPs to be located in two overlapping genes. In this case, the user can choose, when creating the input file, to list that SNP in both genes, only one of the genes, or simply ignore SNPS that are not in a unique gene. If the user chooses to list a SNP in more than one gene, the SNP must have multiple rows in the input file with a different gene in each row.

If the user only has SNP names and p-values, gene name must be appended to the file. The user can either use their own file to append gene name, base pair, and chromosome. The R package biomaRt located in bioconductor is a good source for linking SNPs to genes.

Value

An object of class 'MixMAP'.

`output` data.frame with a row for each gene containing gene symbol, posterior estimates for all gene level effects, variance used in intervals, upper bound of one sided interval, the number of SNPs in each gene, the chromsome of the SNP, the location coordinate of the SNP, the gene-level p-value, the Bonferroni adjusted gene-level p-value, and the q-value based on Banjamini-Hochberg fasle discovery rate.

`num.genes.detected` A vector containing the number of SNPs detected and the total number of genes

`detected.genes` data.frame with a row for each gene containing gene symbol, posterior estimates for all gene level effects, variance used in intervals, upper bound of one sided interval, the number of SNPs in each gene, the chromsome of the SNP, the location coordinate of the SNP, the gene-level p-value, the adjusted gene-level p-value, the name of the SNP with the smallest p-value, the minimum p-value in the gene, and a 5 number summary of the p-values within each gene

`lmer.out` A mer object containing all of the model output information, including parameter estimates, from the lmer function.

Author(s)

Gregory J. Matthews

References

Foulkes, A.S., Matthews, G.J., Das, U., Ferguson, J., Reilly, M. (2013) "Mixed Modeling of Meta-Analysis P-Values (MixMAP) Suggests Multiple Novel Gene Loci for Low Density Lipoprotein Cholesterol". PLoS ONE 8(2): e54812.

Examples

```
library(MixMAP)
#Load data
#This data has been prepared to be used as input to the MixMAP function
data(MixMAP_example)
#Run MixMAP
MixOut<-mixmapPI(MixMAP_example,pval="GC.Pvalue",snp="MarkerName",
  chr="Chr",coord="Coordinate",gene="Gene")
#Display first ten detected genes
summary(MixOut)
#MixManhattan Plot
plot(MixOut)
```

`mixmapTest` *mixmapTest mixmapTest*

Description

This function implements the MixMAP algorithm, which performs gene-level tests of association using data from a previous GWAS or data from a meta-analysis of several GWAS. Conceptually, genes are detected as significant if the collection of p-values within a gene are determined to be collectively smaller than would be observed by chance.

Usage

```
mixmapTest(data.set, pval="pval", snp="snp", gene="gene",
  coord="coord", chr="chr", alpha = 0.05)
```

Arguments

`data.set`	A data.frame containing the input data. Each observation in this data set is a SNP. This file must contain, at least, the SNP name, p-value for a SNP, group name (usually a gene name) where SNP is located, SNP location coordinate, and chromosome number of the SNP.
`pval`	A character string with the name of the variable containing the p-values in the data.set. Default is "pval".
`snp`	A character string with the name of the variable containing the SNP name in the data.set. Default is "snp".
`gene`	A character string with the name of the variable containing the gene name in the data.set. Default is "gene".
`coord`	A character string with the name of the variable containing the location coordinate of the SNP in the data.set. Default is "coord".
`chr`	A character string with the name of the variable containing the chromosome number of the SNP location in the data.set. Default is "chr".
`alpha`	A numeric scalar indicating the level of significance the user chooses to use for detection. The default is 0.05.

Details

The user must provide a file that includes SNP name, SNP p-value, and a group name (assumed to be gene, but the user can use any grouping they choose), among other inputs.

Note about SNPs in genes: It is possible for individual SNPs to be located in two overlapping genes. In this case, the user can choose, when creating the input file, to list that SNP in both genes, only one of the genes, or simply ignore SNPS that are not in a unique gene. If the user chooses to list a SNP in more than one gene, the SNP must have multiple rows in the input file with a different gene in each row.

If the user only has SNP names and p-values, gene name must be appended to the file. The user can either use their own file to append gene name, base pair, and chromosome. The R package biomaRt located in bioconductor is a good source for linking SNPs to genes.

Value

An object of class 'MixMAP'.

`output` data.frame with a row for each gene containing gene symbol, posterior estimates for all gene level effects, variance used in intervals, upper bound of one sided interval, the number of SNPs in each gene, the chromsome of the SNP, the location coordinate of the SNP, the gene-level p-value, the Bonferroni adjusted gene-level p-value, and the q-value based on Banjamini-Hochberg fasle discovery rate.

`num.genes.detected` A vector containing the number of SNPs detected and the total number of genes

`detected.genes` data.frame with a row for each gene containing gene symbol, posterior estimates for all gene level effects, variance used in intervals, upper bound of one sided interval, the number of SNPs in each gene, the chromsome of the SNP, the location coordinate of the SNP, the gene-level p-value, the adjusted gene-level p-value, the name of the SNP with the smallest p-value, the minimum p-value in the gene, and a 5 number summary of the p-values within each gene

`lmer.out` A mer object containing all of the model output information, including parameter estimates, from the lmer function.

Author(s)

Gregory J. Matthews

References

Foulkes, A.S., Matthews, G.J., Das, U., Ferguson, J., Reilly, M. (2013) "Mixed Modeling of Meta-Analysis P-Values (MixMAP) Suggests Multiple Novel Gene Loci for Low Density Lipoprotein Cholesterol". PLoS ONE 8(2): e54812.

Examples

```
library(MixMAP)
#Load data
#This data has been prepared to be used as input to the MixMAP function
data(MixMAP_example)
#Run MixMAP
MixOut<-mixmapTest(MixMAP_example,pval="GC.Pvalue",snp="MarkerName",
  chr="Chr",coord="Coordinate",gene="Gene")
#Display first ten detected genes
summary(MixOut)
#MixManhattan Plot
plot(MixOut)
```

MixMAP_example *MixMAP_example*

Description

Data containing the results of a meta-analysis consisting of many GWAS whose phenotype of interest is LDL cholesterol. This file contains all of the information necessary to run the MixMAP function directly without any pre-processing.

Usage

```
data(MixMAP_example)
```

Format

A data frame with 31825 observations and the following five variables:

MarkerName: a character with SNP name

GC.Pvalue: a numeric with p-values for each SNP

Coordinate: integer containing the base pair location of the SNP within the chromosome

Chr: integer containing the chromsome number of the SNP

Gene: A character with gene symbol

Details

This file contains all of the information needed to run the MixMAP function directly.

Source

Nature 466, 707–713 (05 August 2010) The p-values in this file are from the paper Teslovich Et Al. (2009) “Biological, clinical and population relevance of 95 loci for blood lipids" Nature 466: 707–713. The phenotype of interest was low-density lipoprotein (LDL) cholesterol. The SNPs chosen are a subset that are found in the IBC array and have a unique gene name associated with them.

References

Foulkes, A.S., Matthews, G.J., Das, U., Ferguson, J., Reilly, M. (2013) “Mixed Modeling of Meta-Analysis P-Values (MixMAP) Suggests Multiple Novel Gene Loci for Low Density Lipoprotein Cholesterol". PLoS ONE 8(2): e54812.

Examples

```
library(MixMAP)
#Load data
#This data has been prepared to be used as input to the MixMAP function directly
data(MixMAP_example)
str(MixMAP_example)
```

```
#Run mixmapTest
MixOut<-mixmapTest(MixMAP_example,pval="GC.Pvalue",snp="MarkerName",
  chr="Chr",coord="Coordinate",gene="Gene")
#Display first ten detected genes
summary(MixOut)
#MixManhattan Plot
plot(MixOut)
```

`Plot` *Plot*

Description

The purpose of plot is to graphically display the results of MixMAP in a similar way as a Manhattan plot that is used for Genome Wide Association Studies.

Usage

```
## S4 method for signature 'MixMAP'
plot(x,col.genes=c("black","gray"),col.detected=c("blue","violet"),
  col.text="black",title="MixMAP Manhattan Plot",display.text=TRUE)
```

Arguments

`x`	An object of class mixmap, which is output from the function MixMAP.
`col.genes`	A character vector containing colors for alternate chromosomes. Default is black and gray.
`col.detected`	A character vector containing with the names of color to display genes detected by MixMAP. Genes detected by MixMAP but not by single SNP analysis will be dispalyed using the first color. Genes deteced in both MixMAP and single SNP analysis are displayed using the second colors. Default is blue and violet.
`col.text`	A character string with the name of a color to display text of the names of detected genes. Default is black.
`title`	A character string used a title for the MixManhattan plot. Default is "MixMAP Manhattan Plot".
`display.text`	Either TRUE or FALSE. FALSE will not label detected genes; TRUE will add text. Default is TRUE.

Details

Since MixMAP searches for genes with small empirical Bayes estimates, the resulting graph plots all genes with a positive empirical Bayes estimate as zero, and genes with a negative Bayes estimate are displayed as their absolute values.

Value

Returns a MixManhattan plot.

Author(s)

Gregory J. Matthews

References

Foulkes, A.S., Matthews, G.J., Das, U., Ferguson, J., Reilly, M. (2013) "Mixed Modeling of Meta-Analysis P-Values (MixMAP) Suggests Multiple Novel Gene Loci for Low Density Lipoprotein Cholesterol". PLoS ONE 8(2): e54812.

Examples

```
data(MixMAP_example)
#Run mixmapTest
MixOut<-mixmapTest(MixMAP_example,pval="GC.Pvalue",snp="MarkerName",
  chr="Chr",coord="Coordinate",gene="Gene")
plot(MixOut)

#Run mixmapPI
MixOutPI<-mixmapTest(MixMAP_example,pval="GC.Pvalue",snp="MarkerName",
  chr="Chr",coord="Coordinate",gene="Gene")
plot(MixOutPI)
```

Summary *Summary*

Description

Displays a summary of the MixMAP object.

Usage

```
## S4 method for signature 'MixMAP'
summary(mixmap.object)
```

Arguments

`mixmap.object`	An object of class MixMAP, which is output from the function MixMAP.

Details

This method returns up to the top ten detected genes as well as information on the total number of genes and the number of genes detected.

Value

Returns a summary of the MixMAP object.

Author(s)

Gregory J. Matthews

References

Foulkes, A.S., Matthews, G.J., Das, U., Ferguson, J., Reilly, M. (2013) "Mixed Modeling of Meta-Analysis P-Values (MixMAP) Suggests Multiple Novel Gene Loci for Low Density Lipoprotein Cholesterol". PLoS ONE 8(2): e54812.

Examples

```
data(MixMAP_example)
#Run mixmapTest
MixOut<-mixmapTest(MixMAP_example,pval="GC.Pvalue",snp="MarkerName",
  chr="Chr",coord="Coordinate",gene="Gene")
summary(MixOut)

#Run mixmapPI
MixOutPI<-mixmapPI(MixMAP_example,pval="GC.Pvalue",snp="MarkerName",
  chr="Chr",coord="Coordinate",gene="Gene")
summary(MixOutPI)
```

Index