

Package ‘GeneticMediation’

May 10, 2020

Type Package

Title An R Package for Causal Mediation Analysis of ROSMAP Data

Version 1.1.0.9000

License GPL-3

URL <https://github.com/tydarnell/GeneticMediation>,
<http://www.tydarnell.com/GeneticMediation>

Author Ty Darnell [aut, cre]

Maintainer Ty Darnell <tydarnell@gmail.com>

Description Provides methods for cleaning and processing data from the ROSMAP study
and preparing it for causal mediation analysis

Encoding UTF-8

LazyData true

RoxygenNote 7.1.0

Depends R (>= 3.6.0)

biocViews

Imports data.table,
IRanges,
mediation,
purrr,
readr,
stats,
utils

Suggests dplyr,
knitr,
magrittr,
rmarkdown,
testthat,
tibble,
tidyselect

VignetteBuilder knitr

R topics documented:

| | |
|-------------------------------|-----------|
| check_data | 2 |
| chr_numeric | 3 |
| clean_data | 3 |
| combine_data_spc | 3 |
| combine_data_spg | 4 |
| last_to_first | 4 |
| lenu | 4 |
| make_folder | 5 |
| match_all | 5 |
| match_snp_peak | 5 |
| med.res_spc | 6 |
| med.res_spg | 6 |
| mediation_spg | 6 |
| med_all_spc | 7 |
| med_chr_spc | 7 |
| med_chr_spg | 8 |
| med_results_spg | 8 |
| med_table_spc | 9 |
| med_table_spg | 9 |
| merge_csv | 9 |
| sens_list | 10 |
| sens_summary | 10 |
| setup_folders | 10 |
| sig_chr_spg | 11 |
| sig_spg | 11 |
| snp_peak_bychr | 11 |
| sum_zero | 12 |
| transpose_readcount | 12 |
| Index | 13 |

| | |
|------------|-------------------|
| check_data | <i>Check Data</i> |
|------------|-------------------|

Description

Checks if the data is in the data folder

Usage

check_data()

| | |
|-------------|---------------------------|
| chr_numeric | <i>Chromosome Numeric</i> |
|-------------|---------------------------|

Description

Convert chromosome names from character to numeric. Useful when sorting a dataframe by chromosome number.

Usage

```
chr_numeric(Chr)
```

Arguments

| | |
|-----|--|
| Chr | a character column or vector of chromosome names |
|-----|--|

| | |
|------------|-------------------|
| clean_data | <i>Clean Data</i> |
|------------|-------------------|

Description

Clean snp.info, peak.info data and save as data/chipseq.RData

Usage

```
clean_data(snp.path, peak.path)
```

Arguments

| | |
|-----------|-----------------------------------|
| snp.path | path to SNP information dataframe |
| peak.path | path to Peak annotation dataframe |

| | |
|------------------|---------------------------------------|
| combine_data_spc | <i>Combine Data SNP-Peak-Clinical</i> |
|------------------|---------------------------------------|

Description

SNP-Peak-Clinical data: combine projid, SNPs, PCs, membership, age, gender, peaks, and outcome data for a chromosome

Usage

```
combine_data_spc(chr)
```

Arguments

| | |
|-----|-----------------|
| chr | chromosome name |
|-----|-----------------|

| | |
|------------------|-----------------------------------|
| combine_data_spg | <i>Combine Data SNP-Peak-Gene</i> |
|------------------|-----------------------------------|

Description

SNP-Peak-Gene Data: combine projid, SNPs, PCs, membership, age, gender, peaks, gene and outcome data for a chromosome.

Usage

```
combine_data_spg(chr)
```

Arguments

| | |
|-----|-----------------|
| chr | chromosome name |
|-----|-----------------|

| | |
|---------------|----------------------|
| last_to_first | <i>Last to First</i> |
|---------------|----------------------|

Description

Make the last column the first column in a dataframe

Usage

```
last_to_first(df)
```

Arguments

| | |
|----|-----------|
| df | dataframe |
|----|-----------|

| | |
|------|----------------------|
| lenu | <i>Length Unique</i> |
|------|----------------------|

Description

Get the length of unique values in a vector

Usage

```
lenu(x)
```

Arguments

| | |
|---|--------|
| x | vector |
|---|--------|

| | |
|-------------|--------------------|
| make_folder | <i>Make Folder</i> |
|-------------|--------------------|

Description

Make a folder only if the folder does not already exist

Usage

```
make_folder(path)
```

Arguments

| | |
|------|-------------|
| path | folder path |
|------|-------------|

| | |
|-----------|---|
| match_all | <i>Match SNPs Peaks All Chromosomes</i> |
|-----------|---|

Description

Match SNPs and Peaks in all chromosomes and return dataframe of matches

Usage

```
match_all(snp.info, peak.info, chrs)
```

Arguments

| | |
|-----------|--------------------------------------|
| snp.info | SNP information |
| peak.info | Peak information |
| chrs | character vector of chromosome names |

| | |
|----------------|-----------------------------|
| match_snp_peak | <i>Match SNPs and Peaks</i> |
|----------------|-----------------------------|

Description

Match SNPs and Peaks in a chromosome

Usage

```
match_snp_peak(snp.info, peak.info)
```

Arguments

| | |
|-----------|-------------------------|
| snp.info | SNP information for chr |
| peak.info | Peak information chr |

| | |
|-------------|--|
| med.res_spc | <i>SNP-Peak-Clinical Mediation Data Prep</i> |
|-------------|--|

Description

Prepare a chromosome of SNP-Peak-Clinical data for mediation analysis and save as "data/spc.res/res_chr.RData"

Usage

```
med.res_spc(chr, all_matches)
```

Arguments

| | |
|-------------|---|
| chr | chromosome name |
| all_matches | dataframe of SNP-Peak matches for all chromosomes |

| | |
|-------------|--|
| med.res_spg | <i>SNP-Peak-Gene Mediation Data Prep</i> |
|-------------|--|

Description

Prepare a chromosome SNP-Peak-Gene data for mediation analysis data and save as "data/spg.res/res_chr.RData"

Usage

```
med.res_spg(chr, med.data, matches)
```

Arguments

| | |
|----------|---------------------------------|
| chr | chromosome name |
| med.data | gene mediation data |
| matches | SNP-Peak-Gene matches dataframe |

| | |
|---------------|--------------------------------|
| mediation_spg | <i>Mediation SNP-Peak-Gene</i> |
|---------------|--------------------------------|

Description

Mediation analysis one SNP-Peak-Gene match

Usage

```
mediation_spg(med.res.data, match_row, simulations = 1000)
```

Arguments

| | |
|--------------|--|
| med.res.data | gene mediation data |
| match_row | row in match dataframe: col1 SNP, col2 Peak, col3 Gene |
| simulations | number of simulations, 1000 by default |

| | |
|-------------|--|
| med_all_spc | <i>SNP-Peak-Clinical Mediation Table</i> |
|-------------|--|

Description

Create a table of SNP-Peak-Clinical mediation results for all chromosomes

Usage

```
med_all_spc(chrs, covar, simulations)
```

Arguments

| | |
|-------------|--|
| chrs | character vector of chromosome names |
| covar | covariates string, each covariate separated by + |
| simulations | number of simulations to run |

| | |
|-------------|---|
| med_chr_spc | <i>SNP-Peak-Clinical Mediation Table Chromosome</i> |
|-------------|---|

Description

Create SNP-Peak-Clinical mediation table for SNP-Peak matches in a chromosome

Usage

```
med_chr_spc(matches, med.res.data, covar, simulations)
```

Arguments

| | |
|--------------|---|
| matches | SNP-Peak matches dataframe for a chromosome |
| med.res.data | mediation data |
| covar | covariates string, separate covariates with + |
| simulations | number of simulations to run |

| | |
|-------------|---|
| med_chr_spg | <i>SNP-Peak-Gene Mediation Table Chromosome</i> |
|-------------|---|

Description

Create a SNP-Peak-Gene mediation table for a chromosome

Usage

```
med_chr_spg(chr, gene_matches, covar, simulations)
```

Arguments

| | |
|--------------|---|
| chr | chromosome name |
| gene_matches | list of peak-SNP matches for each gene |
| covar | covariates string, separate covariates with + |
| simulations | number of simulations to run |

| | |
|-----------------|--|
| med_results_spg | <i>Mediation Results SNP-Peak-Gene</i> |
|-----------------|--|

Description

Obtain Mediation Results for Sensitivity Analysis

Usage

```
med_results_spg(med.res.data, match_row)
```

Arguments

| | |
|--------------|------------------------------|
| med.res.data | mediation data |
| match_row | row of SNP-Peak-Gene matches |

| | |
|---------------|--|
| med_table_spc | <i>SNP-Peak-Clinical Mediation Table</i> |
|---------------|--|

Description

Create SNP-Peak-Clinical mediation table for a SNP-Peak match

Usage

```
med_table_spc(med.res.data, match_row, covar, simulations)
```

Arguments

| | |
|--------------|---|
| med.res.data | mediation data |
| match_row | row in match dataframe, should have 2 columns |
| covar | covariates character variable, separate covariates with + |
| simulations | number of simulations to run |

| | |
|---------------|--------------------------------------|
| med_table_spg | <i>SNP-Peak-Gene Mediation Table</i> |
|---------------|--------------------------------------|

Description

Create a mediation table for one SNP-Peak-Gene match

Usage

```
med_table_spg(med.res.data, match_row, covar, simulations)
```

Arguments

| | |
|--------------|---|
| med.res.data | gene mediation data |
| match_row | row in match dataframe: col1 SNP, col2 Peak, col3 Gene |
| covar | covariates character variable, separate covariates with + |
| simulations | number of simulations to run |

| | |
|-----------|------------------|
| merge_csv | <i>Merge CSV</i> |
|-----------|------------------|

Description

merge multiple csv files into a single dataframe

Usage

```
merge_csv(path)
```

Arguments

| | |
|------|-------------------------------------|
| path | path to folder containing csv files |
|------|-------------------------------------|

| | |
|-----------|-------------------------|
| sens_list | <i>Sensitivity List</i> |
|-----------|-------------------------|

Description

Sensitivity analysis list

Usage

```
sens_list(res)
```

Arguments

res

| | |
|--------------|----------------------------|
| sens_summary | <i>Sensitivity Summary</i> |
|--------------|----------------------------|

Description

Sensitivity Analysis Summary

Usage

```
sens_summary(res, i)
```

Arguments

| | |
|-----|-----------------------------|
| res | mediation results dataframe |
| i | row index of dataframe |

| | |
|---------------|----------------------|
| setup_folders | <i>Setup Folders</i> |
|---------------|----------------------|

Description

Creates data, data/mediator, results folders

Usage

```
setup_folders()
```

| | |
|-------------|----------------------------------|
| sig_chr_spg | <i>Significant SNP-Peak-Gene</i> |
|-------------|----------------------------------|

Description

Returns summary of mediation results for a chr with significant results

Usage

```
sig_chr_spg(chr)
```

Arguments

| | |
|-----|-----------------|
| chr | chromosome name |
|-----|-----------------|

| | |
|---------|----------------------------------|
| sig_spg | <i>Significant SNP-Peak-Gene</i> |
|---------|----------------------------------|

Description

Obtain mediation results for significant SPG match for a chromosome as a list. This is needed for sensitivity analysis

Usage

```
sig_spg(chr, sig_matches)
```

Arguments

| |
|-------------|
| chr |
| sig_matches |

| | |
|----------------|-------------------------------------|
| snp_peak_bychr | <i>SNPs and Peaks by Chromosome</i> |
|----------------|-------------------------------------|

Description

save snp.info, peak.info for each chromosome as "data/chipseq/chipseq_chr.RData"

Usage

```
snp_peak_bychr(chrs, snps, peaks)
```

Arguments

| | |
|-------|---------------------------------------|
| chrs | Character vector of chromosomes names |
| snps | SNP information dataframe |
| peaks | Peak information dataframe |

`sum_zero`*Sum Zero*

Description

Get names of columns that have a sum of zero

Usage

```
sum_zero(df)
```

Arguments

`df` dataframe

`transpose_readcount`*Transpose Readcount*

Description

Read in and transpose ChIP-seq Readcount dataframe and match project id

Usage

```
transpose_readcount(readcount.path)
```

Arguments

`readcount.path` ChIP-seq readcount dataframe file path

Index

*Topic **dataprep**

- [clean_data](#), [3](#)
- [combine_data_spc](#), [3](#)
- [combine_data_spg](#), [4](#)
- [med.res_spc](#), [6](#)
- [med.res_spg](#), [6](#)
- [snp_peak_bychr](#), [11](#)
- [transpose_readcount](#), [12](#)

*Topic **helper**

- [chr_numeric](#), [3](#)
- [last_to_first](#), [4](#)
- [lenu](#), [4](#)
- [make_folder](#), [5](#)
- [merge_csv](#), [9](#)
- [sum_zero](#), [12](#)

*Topic **match**

- [match_all](#), [5](#)
- [match_snp_peak](#), [5](#)

*Topic **mediation**

- [med_all_spc](#), [7](#)
- [med_chr_spc](#), [7](#)
- [med_chr_spg](#), [8](#)
- [med_table_spc](#), [9](#)
- [med_table_spg](#), [9](#)
- [mediation_spg](#), [6](#)
- [sig_chr_spg](#), [11](#)

*Topic **setup**

- [check_data](#), [2](#)
- [setup_folders](#), [10](#)

*Topic **spc**

- [clean_data](#), [3](#)
- [combine_data_spc](#), [3](#)
- [med.res_spc](#), [6](#)
- [med_all_spc](#), [7](#)
- [med_chr_spc](#), [7](#)
- [med_table_spc](#), [9](#)

*Topic **spg**

- [combine_data_spg](#), [4](#)
- [med.res_spg](#), [6](#)
- [med_chr_spg](#), [8](#)
- [med_table_spg](#), [9](#)
- [mediation_spg](#), [6](#)
- [sig_chr_spg](#), [11](#)

- [check_data](#), [2](#)
- [chr_numeric](#), [3](#)
- [clean_data](#), [3](#)
- [combine_data_spc](#), [3](#)
- [combine_data_spg](#), [4](#)

- [last_to_first](#), [4](#)
- [lenu](#), [4](#)

- [make_folder](#), [5](#)
- [match_all](#), [5](#)
- [match_snp_peak](#), [5](#)
- [med.res_spc](#), [6](#)
- [med.res_spg](#), [6](#)
- [med_all_spc](#), [7](#)
- [med_chr_spc](#), [7](#)
- [med_chr_spg](#), [8](#)
- [med_results_spg](#), [8](#)
- [med_table_spc](#), [9](#)
- [med_table_spg](#), [9](#)
- [mediation_spg](#), [6](#)
- [merge_csv](#), [9](#)

- [sens_list](#), [10](#)
- [sens_summary](#), [10](#)
- [setup_folders](#), [10](#)
- [sig_chr_spg](#), [11](#)
- [sig_spg](#), [11](#)
- [snp_peak_bychr](#), [11](#)
- [sum_zero](#), [12](#)

- [transpose_readcount](#), [12](#)