Package 'MetaSubtract'

October 12, 2022

Type Package

Title Subtracting Summary Statistics of One or more Cohorts from Meta-GWAS Results

Version 1.60

Date 2020-03-29

Author Ilja M. Nolte

Maintainer Ilja M. Nolte <i.m.nolte@umcg.nl>

Description If results from a meta-GWAS are used for validation in one of the cohorts that was included in the meta-analysis, this will yield biased (i.e. too optimistic) results. The validation cohort needs to be independent from the meta-Genome-Wide-Association-Study (meta-GWAS) results. 'MetaSubtract' will subtract the results of the respective cohort from the meta-GWAS results analytically without having to redo the meta-GWAS analysis using the leave-one-out methodology. It can handle different meta-analyses methods and takes into account if single or double genomic control correction was applied to the original meta-analysis. It can also handle different meta-analysis methods. It can be used for whole GWAS, but also for a limited set of genetic markers. See for application: Nolte I.M. et al. (2017); <doi:10.1038/ejhg.2017.50>.

License GPL (>= 3)

NeedsCompilation no

Repository CRAN

Date/Publication 2020-03-30 12:10:06 UTC

R topics documented:

MetaSubtract-package	*Subtracting Summary Statistics of One or more Cohorts from Meta-GWAS Results*

Description

If results from a meta-GWAS are used for validation in one of the cohorts that was included in the meta-analysis, this will yield biased (i.e. too optimistic) results. The validation cohort needs to be independent from the meta-GWAS results. MetaSubtract will subtract the results of the respective cohort from the meta-GWAS results analytically without having to redo the meta-GWAS analysis using the leave-one-out methodology. It can handle different meta-analyses methods and takes into account if single or double genomic control correction was applied to the original meta-analysis. It can be used for whole GWAS, but also for a limited set of SNPs or other genetic markers.

Details

The core of this package is the function `meta.subtract`. The merging of cohort results with the meta-GWAS summary statistics is done based on marker identification code (e.g. rs-number or CHR:POSITION:TYPE). The script will check for allele flips and strand mismatches. If after allele flip and strand correction the alleles still don't match, the statistics for those markers will be set to missing in the cohort results (implying that the meta-GWAS results are not corrected) and a list of corresponding marker identification codes will be saved to a file with the name of the cohort's results file extended with '.allele_mismatch.txt'

Author(s)

Ilja M. Nolte

Maintainer: Ilja M. Nolte <i.m.nolte@umcg.nl>

References

Nolte et al. (2017). Missing heritability: is the gap closing? An analysis of 32 complex traits in the Lifelines Cohort Study. Eur J Hum Genet. 2017;25:877-885.

See Also

`meta.subtract`.

---

alternative_headers   *A file with alternative header names in the input files*

Description

A file with alternative header names in the input files, which is used to internally convert the header names from your file(s) to the ones used by the script. The file has two columns. The first one contains the variable names as used in the script, the second one contains the header names as specified in the input files. A default file can be found in the subdirectory 'extdata' of the installation directory of MetaSubtract. You can add header names of your files to this file to make the internal conversion.

The variable names in the script are the following:

MARKER marker name

EFFECTALLELE effect allele

OTHERALLELE other allele

BETA effect size

EAF effect allele frequency

N sample size

P p-value

SE standard error of BETA

Z z-score

NSTUDIES number of studies

DIRECTION direction of effects

LP -log10(p-value)

QHET Q-value for heterogeneity

QHETP p-value of QHET

I2HET I2 value for heterogeneity

---

meta.subtract   *Subtracts GWAS results from one or more cohorts from the meta-GWAS results*

Description

Subtracts GWAS results from one or more cohorts from the meta-GWAS results. It can handle the following meta-analysis methods: fixed effects inverse variance weighted, fixed effects sample size weighted, and fixed effect sample size z-sore method. It also can account for single or double genomic control correction of the meta-GWAS results. It can be used for whole GWAS, but also for a limited set of SNPs or other genetic markers. A column with markername is required in both the meta-GWAS and cohort results files as well as the specific statistics used for the type of meta-analysis. Columns with effect and non-effect alleles are preferably given in the files. If this is not the case, it is assumed that statistics are given for the same alleles in the meta-GWAS and cohort results files.

Usage

```
meta.subtract(metafile, cohortfiles, metamethod = "FIV", lambda.meta = 1,
      lambdas.cohort = 1, gc_meta = TRUE, calculate_lambda.meta = TRUE,
      calculate_lambdas.cohort = TRUE, alternative = "alternative_headers.txt",
   save.as.data.frame = TRUE,
   savefile = "meta.results_corrected.with.MetaSubtract.txt.gz",
      logfile = "MetaSubtract.log", dir=tempdir(), ...)
```

Arguments

metafile	Name of the file containing the results of the meta-GWAS
cohortfiles	Names of the files containing the results of the cohorts
metamethod	What kind of meta-GWAS method was used. Can be: FIV = fixed effect inverse variance meta-analysis (default); FSSW = fixed effect sample size weighted meta-analysis; FSZ = fixed effect sample size weighted z-score meta-analysis. In case of FIV, columns with effect size and its standard error are required in both the meta summary statistics file and the cohort results file. In case of FSZ, columns with z-score and sample size are required. In case of FSSW, columns with effect size and sample size are required.
lambda.meta	Genomic control lambda of the meta-results. If no or single genomic control correction was applied, lambda.meta should be set to 1. Default = 1.
lambdas.cohort	Genomic control lambdas of the cohorts. If no genomic control correction was applied, lambdas.cohort should be set to 1 for each cohort. Default = 1.
gc_meta	Apply genomic control to the corrected meta-results? If FALSE, corrected meta will not be corrected for genomic control. Default = TRUE.
calculate_lambda.meta	
	Calculate genomic control lambda from the corrected meta-results? If FALSE and gc_meta=TRUE, lambda.meta will be used for genomic control correction of the corrected meta-results. If gc_meta=FALSE, this will be ignored. Default = TRUE.
calculate_lambdas.cohort	
	Calculate genomic control lambdas from the cohort results? If TRUE, lambdas.cohort will be ignored. Default = TRUE.
alternative	Name of the file containing alternative codes for the column headers in the GWAS files.
save.as.data.frame	
	Set to FALSE if you don't want to save the corrected meta-GWAS summary statistics as a data frame. In that case the results will only be saved to file. Default = TRUE.
savefile	Name of file to which corrected meta-GWAS results are saved. If NULL or NA, results are not saved to file. Default filename = "meta.results_corrected.with.MetaSubtract.txt".
logfile	Name of the file where information on the run is saved.
dir	Directory where the meta and cohort GWAS files are stored.
...	Other possible arguments for reading in the files, e.g. sep, naStrings, or nrow

Value

A data frame with the corrected meta-GWAS results. This data frame contains the same genetic markers as in the input meta-GWAS file. If a marker was not present in the cohort(s) files, the result of that marker will be the same as in the input meta-GWAS file. If a marker was present in the cohort(s) but not in the meta-GWAS file, it will not be included in the output.

Author(s)

Ilja M. Nolte

References

Nolte et al. (2017). Missing heritability: is the gap closing? An analysis of 32 complex traits in the Lifelines Cohort Study. Eur J Hum Genet. 2017;25:877-885.

See Also

MetaSubtract.

Examples

```
metafile="meta_results.txt"
cohortfiles=c("cohort1_results.txt","cohort2_results.txt","cohort3_results.txt")
# Subtract summary statistics of the two cohorts from the meta summary statistics in metafile.
# Genomic control correction of the imported meta summary statistics will be undone using
# lambda.meta. Genomic control will be applied to the corrected meta summary statistics
# (gc_meta=TRUE by default) based on a lambda derived from the new meta summary statistics
# (calculate_lambda.meta=TRUE by default). The cohorts summary statistics will always be first
# corrected for genomic control before subtracting them from the meta summary statistics. By
# default, the lambda will be calculated from the data (calculate_lambdas.cohort=TRUE by default).
m1<-meta.subtract(metafile=metafile, cohortfiles=cohortfiles, lambda.meta=1.046, dir=tempdir())
head(m1)

# If you would like pre-set genomic control lambdas (e.g. in case you use only part of the data),
# set calculate_lambda.meta=FALSE and calculate_lambdas.cohort=FALSE and define the
# lambda for the meta summary statistics by lambda.meta, and for the cohort(s) by lambda.cohorts.
m2<-meta.subtract(metafile=metafile, cohortfiles=cohortfiles, lambda.meta=1.046,
    lambdas.cohort=c(1.0012,1.03), calculate_lambda.meta=FALSE, calculate_lambdas.cohort=FALSE,
  dir=tempdir())
head(m2)

# If you do not want to apply genomic control, set gc_meta=FALSE,
# calculate_lambda.meta=FALSE and calculate_lambdas.cohort=FALSE.
# lambda.meta=1 and lambda.cohorts=c(1,1) are default values.
m3<-meta.subtract(metafile=metafile, cohortfiles=cohortfiles, gc_meta=FALSE,
      calculate_lambda.meta=FALSE, calculate_lambdas.cohort=FALSE,
  dir=tempdir())
head(m3)

cohortfiles=c("cohort1_results.txt","cohort2_results.txt")
# If you performed a sqrt(smaple sizes) weighted meta-analysis on the z-scores, change the
```

```
# meta-analysis method by 'metamethod=FSZ'.
m4<-meta.subtract(metafile=metafile, cohortfiles=cohortfiles, metamethod="FSZ",
      lambda.meta=1.046, dir=tempdir())
head(m4)
```

Index