Package: MRInstruments (via r-universe)

September 30, 2024

Title Data sources for genetic instruments to be used in MR

Version 0.3.2

Description Datasets of eQTLs, GWAS catalogs, etc.

License CC0

LazyData true

Depends R (>= 2.10)

Suggests dplyr

URL https://github.com/MRCIEU/MRInstruments

BugReports https://github.com/MRCIEU/MRInstruments/issues

RoxygenNote 7.1.0

Encoding UTF-8

Roxygen list(markdown = TRUE)

Repository https://mrcieu.r-universe.dev

RemoteUrl https://github.com/MRCIEU/MRInstruments

RemoteRef HEAD

RemoteSha efa2ca0d7a57cd51067be8e445a0f46bb6060a2e

Contents

`aries_mqtl` *Data frame of mQTLs from ARIES study*

Description

mQTL analysis was performed in five time points in mothers and children. These data are obtained from http://www.gtexportal.org/home/ (GCTA version).

Usage

```
aries_mqtl
```

Format

A data frame with 187318 rows and 19 columns

SNP SNP

timepoint timepoint

cpg cpg

beta beta

pval pval

se se

snp_chr snp_chr

snp_pos snp_pos

effect_allele effect_allele

other_allele other_allele

eaf eaf

sex sex

age age

units units

island_location island_location

cpg_chr cpg_chr

cpg_pos cpg_pos

gene gene

gene_location gene_location

cis_trans cis or trans acting mQTL

Source

http://www.gtexportal.org/home/

blueprint_monocyte_eqtl

Data frame of eQTLs (monocyte cells) from the BLUEPRINT study

Description

Data was downloaded from ftp://ftp.ebi.ac.uk/pub/databases/blueprint/blueprint_Epivar/qtl_as/QTL_RESULTS/mono_gene_nor_combat_peer_10_all_summary.txt.gz. Top hit SNPs filtered for (all associations with P<5e-8).

Usage

```
blueprint_monocyte_eqtl
```

Format

A data frame with 452537 rows and 8 columns

Phenotype Gene name

SNP SNP rs ID

beta Effect size

se Standard error

eaf Effect allele frequency

effect_allele Effect allele

other_allele Non effect allele

pval p-value

Source

https://scmv-ieugit.epi.bris.ac.uk/gh13047/blueprint_monocyte_tophits

drug_interactions *Drug interactions*

Description

List of drug names and the genes that they target, along with information about how they interact and source of information

Usage

```
drug_interactions
```

Format

A data frame with 38116 rows and 6 columns

entrez_gene_symbol entrez gene symbol

gene_long_name gene long name

interaction_claim_source interaction claim source

interaction_types interaction types

drug_name drug name

drug_primary_name drug primary name

Source

The "Interactions TSV" at http://dgidb.genome.wustl.edu/downloads

---

gene_trials *Drug trial information regarding target genes*

Description

The drug trial information comes from PMID 26121088 supplementary table 13. For a number of genes information is provided about diseases for which they are targeted, and the status of the trials in which they were tested.

Usage

```
gene_trials
```

Format

A data frame with 19085 rows and 5 columns

entrez_gene_symbol Entrez Gene symbol

MeSH1 MeSH term 1

MeSH2 Top MeSH term

drug_phase Phase of drug development

approved Approved in EU or US

Source

http://www.ncbi.nlm.nih.gov/pubmed/26121088

gtex_eqtl *Data frame of cis-eQTLs from GTEx study (version 6p)*

Description

Gene expression - SNP associations from 27094 gene names and 44 tissues. Obtained from http://www.gtexportal.org/home/. These data only include associations that GTEx labelled as is_choson_snp.

Usage

```
gtex_eqtl
```

Format

A data frame with 280630 rows and 11 columns

gene_name Gene name

tissue Tissue

gene_start Gene chromosome : gene start position (b37)

SNP SNP rs ID

snp_position SNP chromosome : SNP position (b37)

effect_allele Effect allele

other_allele Non effect allele

beta Effect size

se Standard error

pval p-value

n samplesize

Source

https://scmv-ieugit.epi.bris.ac.uk/gh13047/gtex_data

gwas_catalog *Data frame of GWAS significant SNPs obtained from the EBI GWAS catalog*

Description

Downloaded from EBI on 18/03/2016. Standardised to make units as consistent as possible, trait names, obtain gene annotation. Effect sizes and standard errors are all converted to log odds ratio for binary traits. Ensembl used to identify alleles and effect allele frequency where missing.

Usage

```
gwas_catalog
```

Format

A data frame with 22783 rows and 17 columns

Phenotype_simple Simple phenotype name

Phenotype This phenotype also has units and details if they are necessary to distinguish different studies

Phenotype_info Phenotype_info

PubmedID PubmedID

Author Author

Year Year

SNP SNP

chr chr

bp_ens_GRCh38 bp_ens_GRCh38

Region Region

gene gene

Gene_ens Gene_ens

effect_allele effect_allele

other_allele other_allele

beta beta

se se

pval pval

units units

eaf eaf

date_added_to_MRBASE Date GWAS catalog downloaded and added to MR-Base

Initial_sample_description Sample size and ancestry description for stage 1 of GWAS (summing across multiple Stage 1 populations, if applicable)

Replication_sample_description Sample size and ancestry description for subsequent replication(s) (summing across multiple populations, if applicable)

MAPPED_TRAIT_EFO_URI URI of the EFO trait

MAPPED_TRAIT_EFO Mapped Experimental Factor Ontology trait for this study

STUDY.ACCESSION Accession ID allocated to a GWAS Catalog study

Source

https://scmv-ieugit.epi.bris.ac.uk/gh13047/gwas_catalog_standardisation

`metab_qtls` *Data frame of QTLs of metabolites*

Description

Data provided by Peter Wurtz and Mika Ala-Korpela ahead of publication

Usage

```
metab_qtls
```

Format

A data frame with 1088 rows and 12 columns

phenotype phenotype

chromosome chromosome

position position

SNP SNP rs ID

effect_allele Effect allele

other_allele Non effect allele

eaf effect allele frequency

beta Effect size

se Standard error

pval p-value

n_studies number of studies

n samplesize

Source

unpublished

`mrbase_instruments` *Data frame of GWAS significant SNPs obtained from the EBI GWAS catalog*

Description

Created from the MR-Base database 23/12/2017 The top hits (5e-8) are extracted from these files and then the 1kg European dataset is used to clump the top hits (rsq = 0.001, kb = 10000)

Usage

```
mrbase_instruments
```

Format

A data frame with 29792 rows and 16 columns

exposure This phenotype also has units and details if they are necessary to distinguish different studies

id.exposure id.exposure

samplesize.exposure samplesize.exposure

ncase.exposure ncase.exposure

ncontrol.exposure ncontrol.exposure

mr_keep.exposure mr_keep.exposure

pval_origin.exposure pval_origin.exposure

data_source.exposure data_source.exposure

SNP SNP

effect_allele.exposure effect_allele.exposure

other_allele.exposure other_allele.exposure

beta.exposure beta.exposure

se.exposure se.exposure

pval.exposure pval.exposure

units.exposure units.exposure

eaf.exposure eaf.exposure

Source

http://www.mrbase.org

`proteomic_qtls` *Data frame of proteomic QTLs from Deming et al 2016*

Description

GWAS of immunoassays of 190 protein levels from plasma

Usage

```
proteomic_qtls
```

Format

A data frame with 56 rows and 13 columns

analyte analyte

chr chr

position position

SNP SNP

gene gene

location location

annotation annotation

effect_allele effect_allele

other_allele other_allele

eaf eaf

maf MAF

pval pval

beta beta

se se

Source

https://scmv-ieugit.epi.bris.ac.uk/gh13047/proteomic_qtls

Index