

User guide

BioWinfordMR is a platform dedicated to performing Mendelian randomization causal inference. The platform integrates various Mendelian-related analytical functions and GWAS data. Users can learn how to utilize these functions for their analyses by reading the user guide documentation. Next, we will introduce each module individually.

Clean_format

This module can help users easily process GWAS data in different formats into a standard format. The purpose of doing this is to make it easier for other functional modules to read the same standardized input file to complete the analysis. Generally, GWAS data contains information such as SNP, effect allele, other allele, beta, standard error, p-value, effect allele frequency, trait, ID, etc. However, GWAS data from different sources do not have a standardized naming or arrangement order. Therefore, users can clean up this data using the clean format module.

BioWinford Platform

The screenshot displays the 'Clean Format' module within the BioWinford Platform. The interface includes a sidebar on the left with options to choose a demo file (currently 'exposure.xlsx'), download results, and reset inputs. The main area contains several input fields: 'gwas file*' (with a 'Browse...' button and a 'No file s' error message), 'SNP list' (with a 'Browse...' button and a 'No file s' error message), 'gwas type' (with radio buttons for 'exposure' and 'outcome'), 'SNP*' (with a text input field containing 'SNP'), 'effect_allele*' (with a text input field containing 'effect_allele'), and 'other_allele*' (with a text input field containing 'other_allele'). Red arrows and boxes highlight these fields, pointing to explanatory text boxes on the right. The text boxes provide instructions on file uploads, SNP list filtering, file type selection, and the necessity of variables marked with asterisks.

TwoSampleMR MVMR Multi-response MR instrument Colocalisation MRomics MR circoos MR Forest

chrpos2SNP Mediator_MR Proxy SNP LDSC SMR SMRplot MR_meta MR_CAUSE Clean Format

Choose a demo file: (we provide demo files for user to test)

exposure.xlsx

Download

Download your own results

Get Your Results

Reset inputs

Send output to your email

gwas file*

Browse... No file s

SNP list

Browse... No file s

gwas type

☒ exposure

☐ outcome

SNP*

SNP

effect_allele*

effect_allele

other_allele*

other_allele

Users can upload their own GWAS plain text files (txt, tsv, csv, xlsx) here. The platform also supports compressed files in formats like gz and zip.

If a SNP list is provided, the GWAS data will be filtered accordingly. Only SNP records present in the list are retained.

Users can choose the type of the uploaded file to clean, including exposure or outcome.

Users can enter the column names corresponding to each variable in the GWAS file, allowing the program to standardize and clean up the data format.

All the variables marked with asterisks are necessary for the calculation or figure generation. Other variables like samplesize, and eaf are not necessarily required.

other_allele*

beta/OR*

se

pvalue/LP*

samplesize

trait*

eaf

ID*

fix eaf & samplesize

☒ FALSE

☐ TRUE

Run

Although samplesize and eaf are not necessary for most MR analyses, it's still recommended to provide these two variables in certain situations, such as when conducting coloc analysis. Therefore, users can choose to supplement the samplesize and eaf information of SNPs through reference data from the 1000 Genome Project when these variables are missing. This may not fully restore the true original information, but it can somewhat compensate for incomplete data.

chrpos2snp

In some special cases, users may have obtained GWAS data that is missing SNP information (rsid). For MR analysis, the rsid for each SNP is necessary. In such situations, users can use the "chrpos2snp" module on the Biowinfordmr platform to map the SNP's chromosome and position coordinates to the dbSNP database and convert them to the corresponding rsid.

BioWinford Platform

Choose a demo file: (we provide demo files for user to test)

chrpos.txt

Download

Download your own results

Get Your Results

Reset inputs

Send output to your email

Example demo file

TwoSampleMR MVMR Multi-response MR instrument Colocalisation MRomics MR circoos MR Forest

chrpos2SNP Mediator_MR Proxy SNP LDSC SMR SMRplot MR_meta MR_CAUSE Clean Format

region (8:1-100)

chr pos (chr1 100)

Browse... No file selected

rsid list

Browse... No file selected

Reference

GRCh37

GRCh37

GRCh38

Users upload input files, where the first column is "chr" and the second column is "position". Other columns are not necessary but will be retained in the output file.

Users can choose the version of reference. Then the SNP coordinates in the dbSNP database will be used accordingly.

TwoSampleMR

In this module, we mainly implemented the interactive interface for TwoSampleMR R package. We also supplemented other functions, including mr presso and phenoscanner. Mr presso can call the MRPRESSO R package to evaluate horizontal pleiotropy. The phenoscanner function is used to evaluate if there are confounding factors in the exposure. Other parameters are consistent with the twosamplemr R package.

exposure id (ID1;ID2)

ieu-b-4877;ieu-b-49

outcome id

ebi-a-GCST008043

exposure file

Browse... No file

outcome file

Browse... No file

p value

5e-08

clump_kb

10000

clump_r2

0.001

models

mr_egger_regression

mr_ivw

mr_weighted_mode

mr_weighted_median

mr_simple_median

PhenoScanner

TRUE

MR presso

TRUE

Run

As same as the usage of the TwoSampleMR R package, users can directly enter the GWAS ID and fetch data through the Open GWAS API. In addition, we also provide over 7000 local GWAS files of MRomics module. These data mostly originate from catalogs GWAS, so they were originally not supported for online analysis. However, the BioWinfordMR platform has already processed these data in the local disk, allowing users to fetch and analyze the local files directly by entering GWAS IDs.

If users need to analyze data that cannot be obtained online, such as GWAS data from UK Biobank or FinnGen, we also support users to upload local files. We strongly recommend users to first use the clean format module to clean the local files before uploading them to this module for analysis.

Here are the parameters involved in the other TwoSampleMR R package, including clump_r2, clump_kb, pvalue, and the regression algorithms to be used. The specific meanings of the parameters can be found in the TwoSampleMR document, we won't go into too much detail here.

If the PhenoScanner function is set to TRUE, it will enable the phenoscanner function to conduct a confounding factor analysis on the SNPs. It's important to note that since the phenoscanner website is currently inactive, we have saved the involved data locally for analysis.

If the MR presso function is set to TRUE, the MRPRESSO R package will be activated. MRPRESSO can help users evaluate whether there is effect of horizontal pleiotropy and outliers in the MR results.

MVMR

In some cases, we need to conduct the causal inference between multiple exposures and outcome. These exposures may be genetically correlated. To address this issue, we need to use the `mv_multiple` function in the TwoSampleMR package (MVMR). The parameter interface is similar to TwoSampleMR. Users can analyze in online or local file mode. For MVMR module, we also provide the `mv_lasso` function to eliminate collinearity between multiple exposure factors.

[TwoSampleMR](#)
[MVMR](#)
[Multi-response](#)
[MR instrument](#)
[Colocalisation](#)
[MRomics](#)
[MR circos](#)
[MR Forest](#)

[chrpos2SNP](#)
[Mediator_MR](#)
[Proxy SNP](#)
[LDSC](#)
[SMR](#)
[SMRplot](#)
[MR_meta](#)
[MR_CAUSE](#)
[Clean Format](#)

exposure id (ID1;ID2)

outcome id

exposure files

outcome file

p value

clump_kb

clump_r2

enable mv_lasso

☒ FALSE
 ☐ TRUE

Enter GWAS ID via the Online mode

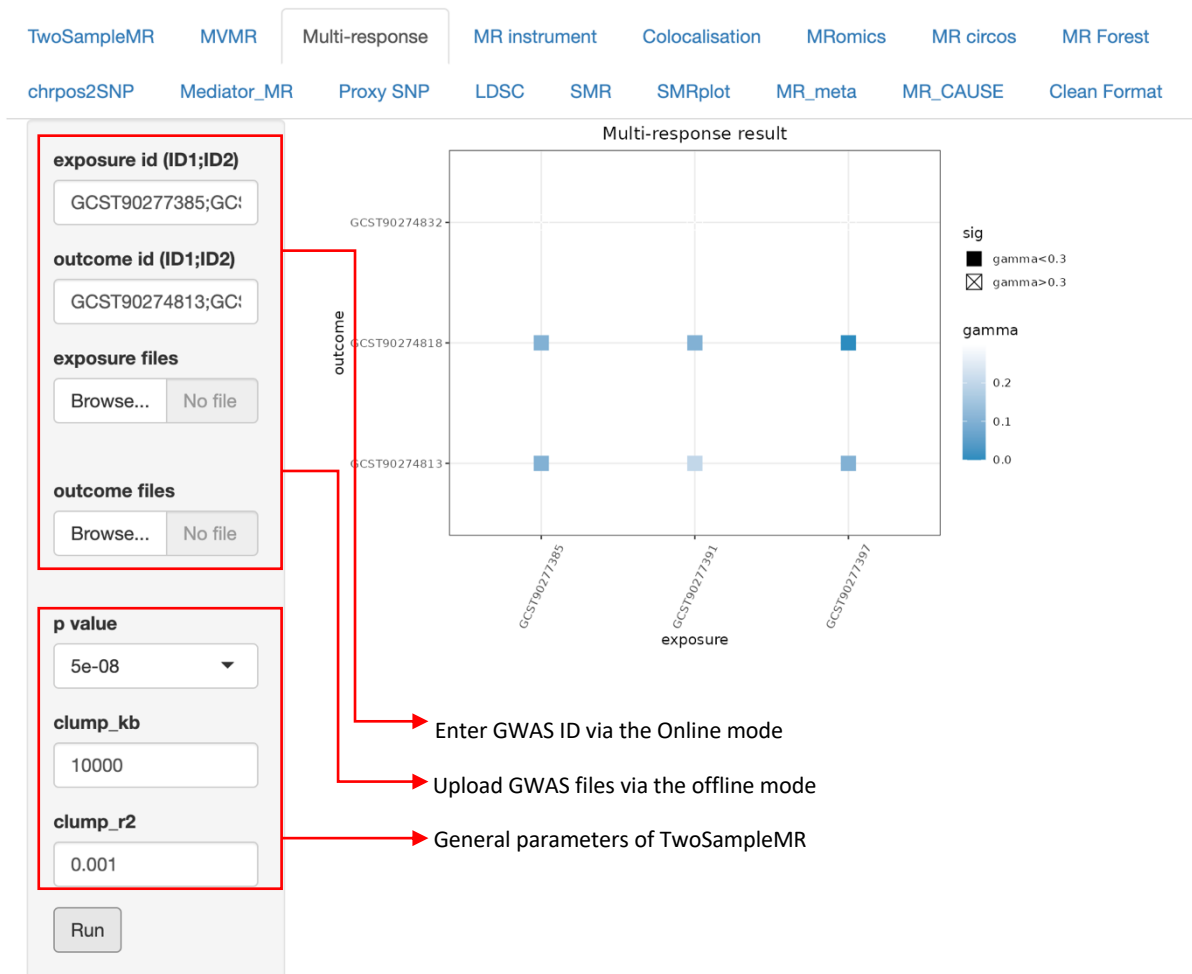
Upload GWAS files via the offline mode

General parameters of TwoSampleMR

If mv_lasso is enabled, the collinearity relationship between multiple exposures is considered.

Multi-response

MVMR can only conduct the causal inference between multiple exposures and a single outcome. If we want to consider the relationship between multiple exposures and multiple outcomes, we can use the Multi-response module. This module invokes the MR2 algorithm (<https://github.com/lb664/MR2/>). The MR2 R package uses a Bayesian model to estimate the correlation between multiple exposures and multiple outcomes. Users can easily use this function via the interface on BiowinfordMR platform.



MR instrument

This is an R package that contains a number of data files from various sources to provide instruments in two sample MR (<https://github.com/MRCIEU/MRInstruments>). We have further supplemented the GWAS data from FinnGen, including the versions from R8 to R10. Users can easily search for matching GWAS data in this module via entering keywords.

TwoSampleMR

MVMR

Multi-response

MR instrument

Colocalisation

MRomics

MR circos

MR Forest

chrpos2SNP

Mediator_MR

Proxy SNP

LDSC

SMR

SMRplot

MR_meta

MR_CAUSE

Clean Format

Source

gwas_catalog

ieu open gwas

gwas_catalog

finngen_R8

finngen_R9

finngen_R10

aries_mqtl

drug_interactions

gene trials

Show 25 entries

Search:

	Phenotype_simple	MAPPED_TRAIT_EFO	MAPPED_TRAIT_EFO_URI
1	Eosinophil percentage of white cells	eosinophil percentage of leukocytes	http://www.ebi.ac.uk/efo/
2	Eosinophil counts	eosinophil count	http://www.ebi.ac.uk/efo/
3	Medication use (agents acting on the renin-angiotensin system)	Agents acting on the renin-angiotensin system use measurement	http://www.ebi.ac.uk/efo/

Colocalization

In this module, users can perform colocalization analysis by entering GWAS ID(s) or uploading local file(s). We provide various types of QTL data for users to choose from, including eQTL, pQTL, and mQTL. To facilitate users to specify the colocalization region, in addition to directly entering an absolute coordinate, users can also specify a genomic region by entering a gene symbol. If a user enters a gene symbol, the region used for colocalization will include the whole gene's region from start to end position, as well as the surrounding 2500kb regions.

TwoSampleMR

MVMR

Multi-response

MR instrument

Colocalisation

MRomics

MR circos

MR Forest

chrpos2SNP

Mediator_MR

Proxy SNP

LDSC

SMR

SMRplot

MR_meta

MR_CAUSE

Clean Format

exposure id (ID1;ID2)

GCST90277380;GCST90277380

gwas files

Browse... No file :

pval.exposure

5e-06

QTL type (option1)

gene name (option2)

genome loc (option3)

1:109724880-109904

Run

Triacylglycerol (46:2) levels

Triacylglycerol (52:4) levels

Position on chromosome 1

Enter GWAS ID or upload local files

eQTL, pQTL, mQTL, etc can be chosen

Specify the coloc region using either a gene name or an absolute genomic coordinate

MRomics

The MRomics module currently contains more than 15 types of omics data involving over 7000 GWAS datasets. Such number keeps growing as the BioWinfordMR platform develops. These omics data consist of gut microbiome, oral microbiome, skin microbiome, cytokines factors, immune cells, lipids, serum metabolites, etc. Users only need to provide one GWAS ID or local file to conduct MR analysis in batches with the aforementioned omics data. To mitigate false positives from multiple testing, the original p-values will be adjusted using the Benjamini-Hochberg correction to generate adjusted p-values.

[TwoSampleMR](#) [MVMR](#) [Multi-response](#) [MR instrument](#) [Colocalisation](#) [MRomics](#) [MR circos](#) [MR Forest](#)
[chrpos2SNP](#) [Mediator_MR](#) [Proxy SNP](#) [LDSC](#) [SMR](#) [SMRplot](#) [MR_meta](#) [MR_CAUSE](#) [MRlap](#)

[Clean Format](#)

ieugwas ID
ieu-b-5088

outcome_file
Browse... No file

p value of SNP
5e-06

Exposure Type
gut211

models (<=5)
mr_ivw
mr_egger_regression
mr_weighted_median
mr_weighted_mode
mr_simple_mode

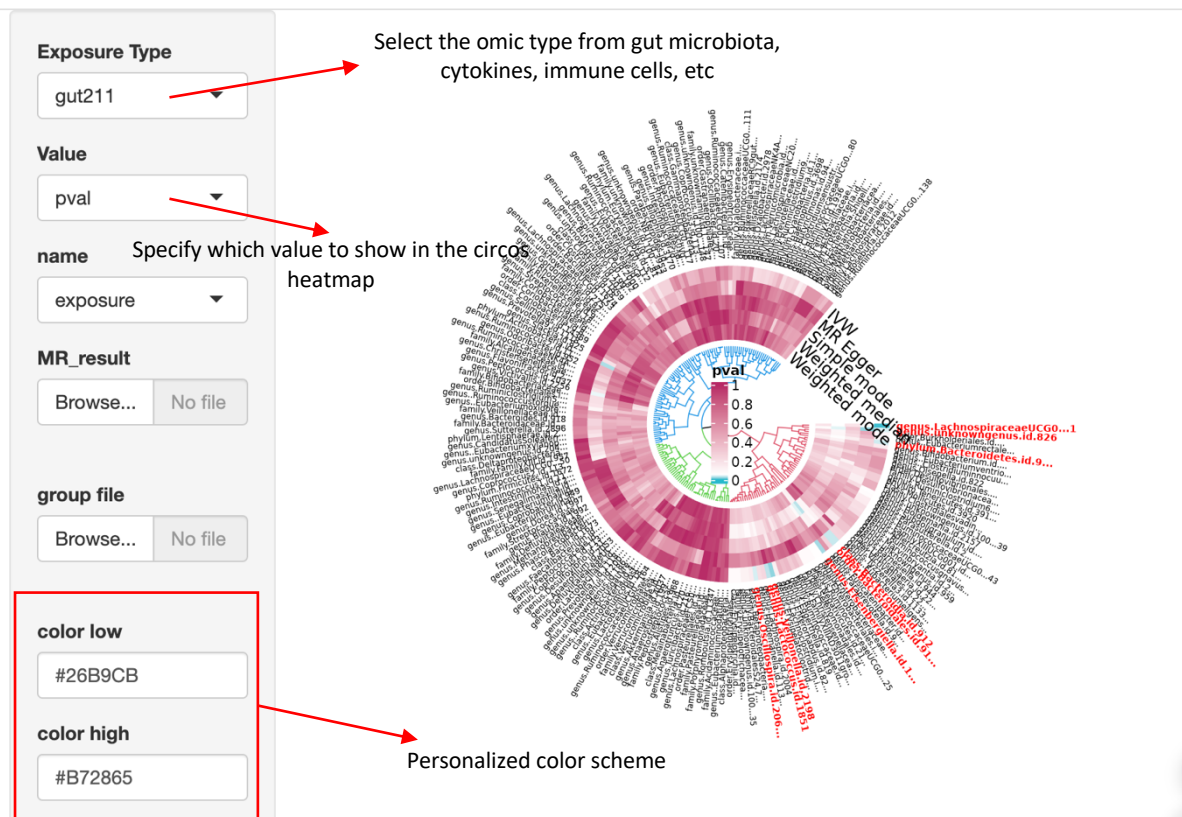
id.exposure **nsnp** **OR (95% CI)** **pval**

GCST90016911	8	1.231 (1.016 to 1.491)	3.4e-02
GCST90016991	6	1.151 (1.018 to 1.301)	2.4e-02
GCST90017026	7	0.763 (0.638 to 0.912)	3.0e-03
GCST90017031	7	1.105 (1.009 to 1.210)	3.1e-02
GCST90017037	4	1.235 (1.012 to 1.508)	3.8e-02
GCST90017086	8	0.844 (0.717 to 0.992)	4.0e-02
GCST90017088	6	0.831 (0.712 to 0.970)	1.9e-02
GCST90017092	8	1.231 (1.016 to 1.491)	3.4e-02
GCST90017111	8	1.215 (1.012 to 1.458)	3.6e-02

0.63 1 1.55

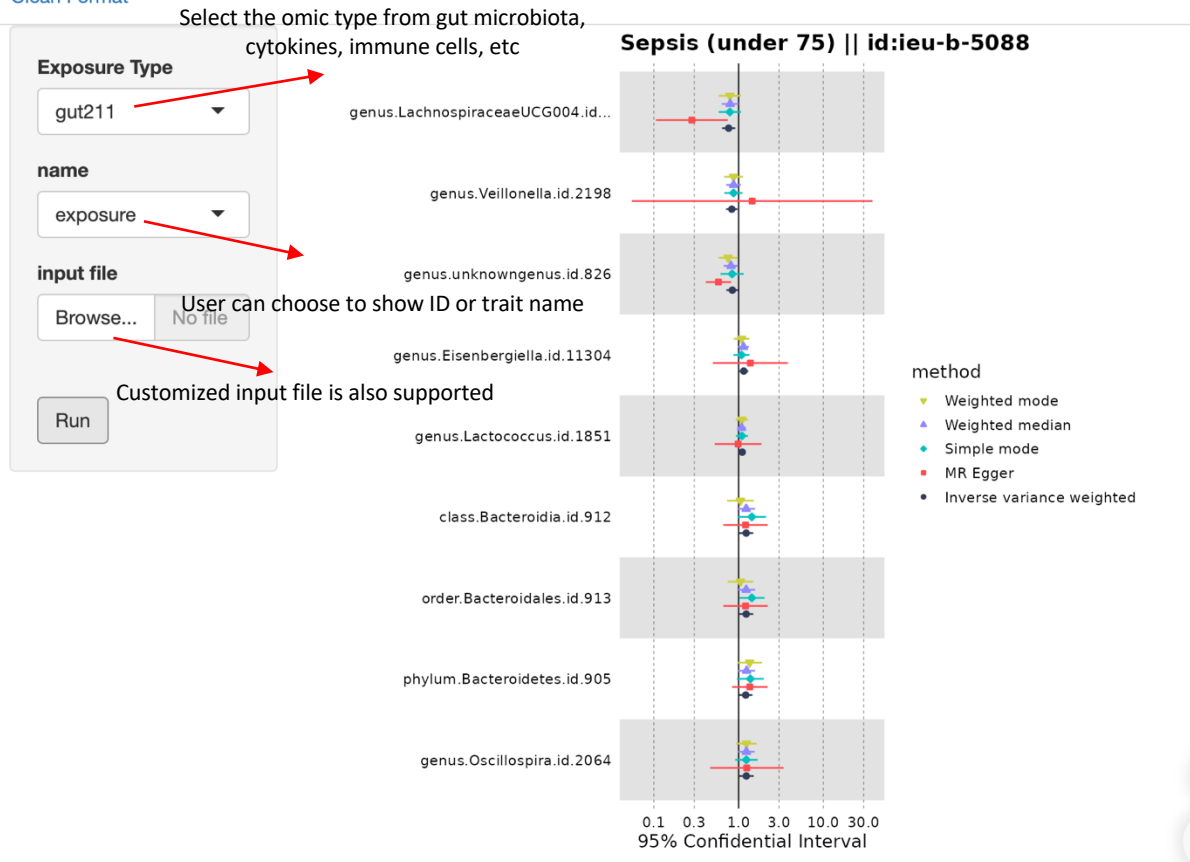
MR circos

This module can directly load the output results from MRomics for visualization analysis. Users can choose the indicators to display (beta, se, OR, p-value) and the type of labels (ID or exposure name). Additionally, this module supports user-defined color schemes.



MR forest

This module can directly load the output results from MRomics for visualization analysis. Users can also upload the customized file to show forest graph. The file content and format should be consistent with the mr_result.txt file exported by MRomics.



Mediator_MR

Although MR can infer causal relationships between two phenotypes, in most cases, the exposure does not directly act on the outcome but rather through certain mediators. The exposure first affects the mediator, which then influences the outcome. This mediator pathway is known as a complete mediator effect. To identify such mediator pathways like exposure-mediator-outcome, we need to utilize the Mediator_MR module. In the Mediator_MR module, we employ the two-step algorithm. The first step involves estimating the effects of exposure-mediator, mediator-outcome, and exposure-outcome. In the second step, after removing the mediator effect, the genetic relationship from exposure to outcome is recalculated.

TwoSampleMR
MVMR
Multi-response
MR instrument
Colocalisation
MRomics
MR circo
MR Forest
chrpos2SNP
Mediator_MR
Proxy SNP
LDSC
SMR
SMRplot
MR_meta
MR_CAUSE
Clean Format

Exposure
GCST90017098
Exposure file
Browse... No file s

Mediator
GCST90001472
Mediator file
Browse... No file s

Outcome
Outcome file
Browse... Sepsis.ç
Upload complete

p value
5e-06
Run

before removing mediator effect

id.exposure	id.outcome	OR (95% CI)	pval
GCST90017098	GCST90001472	0.67 (0.45 to 0.98)	0.038
GCST90001472	GCST90044692	1.11 (1.02 to 1.21)	0.012
GCST90017098	GCST90044692	1.92 (1.04 to 3.53)	0.037

Exposure GWAS ID or local file

after removing mediator effect

id.exposure	id.outcome	OR (95% CI)	pval
GCST90017098	GCST90001472	0.67 (0.45 to 0.98)	0.038
GCST90001472	GCST90044692	1.11 (1.02 to 1.21)	0.012
GCST90017098	GCST90044692	2.00 (1.13 to 3.55)	0.112

Mediator GWAS ID or local file

Outcome GWAS ID or local file

P value cutoff to filter SNPs

Proxy SNP

In the MR analysis process, sometimes we are limited by a small number of SNPs. In such cases, we can consider using proxy SNPs. In the Proxy SNP module, BioWinfordMR allow users to identify proxy SNPs in a specified window range that are in linkage disequilibrium with the root SNP. The code was retrieved from the Github repository of Kamil Slowikowski (https://gist.github.com/slowkow/3d13aa44cf65ca9ad2a0570346ba05?permalink_comment_id=2605197)

[TwoSampleMR](#)
[MVMR](#)
[Multi-response](#)
[MR instrument](#)
[Colocalisation](#)
[MRomics](#)
[MR circos](#)
[MR Forest](#)

[chrpos2SNP](#)
[Mediator_MR](#)
[Proxy SNP](#)
[LDSC](#)
[SMR](#)
[SMRplot](#)
[MR_meta](#)
[MR_CAUSE](#)
[Clean Format](#)

SNP file

Browse...
rsid_list.

Upload complete

win length (kb)

500

Run

Show 25 entries

Search:

rootSNP	seqnames	start	end	R.squared	proxySNP
rs10910048	1	2211849	2211849	0.971771736138365	rs10910047
rs9434469	1	7175858	7175858	0.965223362605805	rs11120852
rs61763314	1	5375524	5375524	0.986206439680886	rs11625671
rs2050815	1	31692242	31692242	0.999157901816867	rs12079944
rs6672915	1	31692242	31692242	0.999157901816867	rs12079944
rs6662203	1	31692242	31692242	0.995791408187299	rs12079944
rs7523841	1	31692242	31692242	0.997473832695384	rs12079944

Upload the root SNP list

Specify the window length to estimate LD between SNPs

SMR

The SMR module allows users to estimate association between the expression level of a gene and a complex trait of interest using summary-level data from GWAS and expression quantitative trait loci (eQTL) studies (Zhu et al. 2016 Nature Genetics). The SMR & HEIDI methodology can be interpreted as an analysis to test if the effect size of a SNP on the phenotype is mediated by gene expression. This tool can therefore be used to prioritize genes underlying GWAS hits for follow-up functional studies. The methods are applicable to all kinds of molecular QTL (xQTL) data, including DNA methylation QTL (mQTL) and protein abundance QTL (pQTL).

TwoSampleMR
MVMR
Multi-response
MR instrument
Colocalisation
MRomics
MR circos
MR Forest
chrpos2SNP
Mediator_MR
Proxy SNP
LDSC
SMR
SMRplot
MR_meta
MR_CAUSE
MRlap

Clean Format

gwas id

gwas file
Browse...
Sepsis.gz
Upload complete

pval.exposure
5e-08

eQTL source
Whole_Blood |
Adipose_Subcutaneous
Adipose_Visceral_Omentum
Adrenal_Gland
Artery_Aorta
Artery_Coronary
Artery_Tibial
Brain_Amygdala
Brain_Anterior cingulate cortex I

Outcome GWAS ID or local file

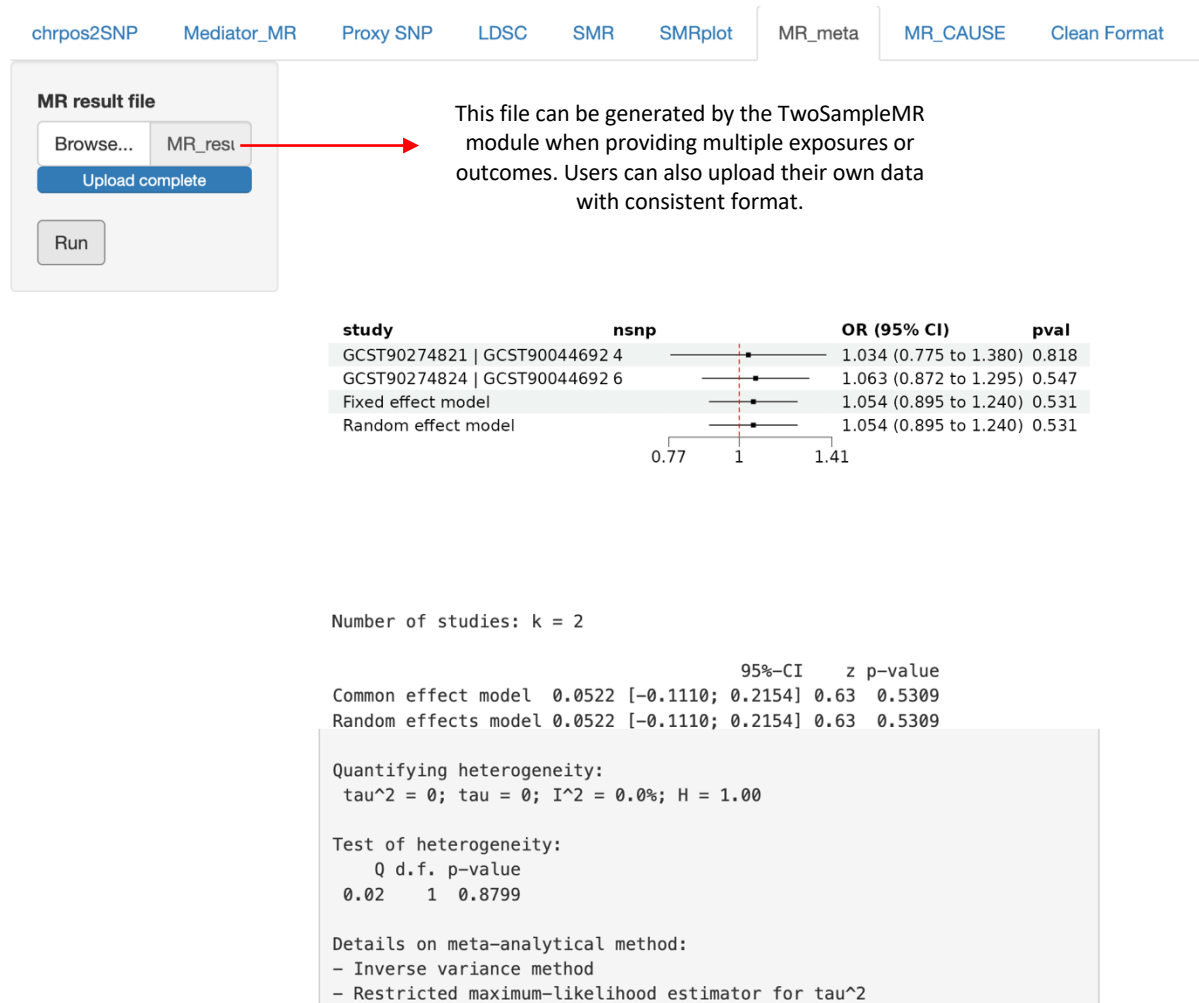
Show 25 entries
Search:

Gene	Probe_bp	topSNP	topSNP_chr	topSNP_bp	A1	
ISG15	1007839	rs1921	1	1014228	A	C
AGRN	1038119	rs4970394	1	1027511	T	C
B3GALT6	1233639	rs60252802	1	1231507	C	1
MXRA8	1357233	rs181929025	1	1355705	T	C
VWA1	1438872	rs115503338	1	1434243	A	C
NADK	1765844	rs4648629	1	1764023	A	C
KCNAB2	6046060	rs806109	1	5992521	A	C
PARK7	7969898	rs226251	1	7964630	C	1
CA6	8960480	rs3765963	1	8974539	G	/
H6PD	9253056	rs7555568	1	9232547	A	

Multiple tissues regarding eQTL GWAS can be selected here.
These data come from the V8 release of the GTEx eQTL/sQTL summary data.

MR_meta

For analyzing multiple exposures or multiple outcomes, in addition to considering the MVMR and multi-response modules, users can also consider conducting meta-analysis on multiple MR results via the meta R package. First, we utilize TwoSampleMR to individually analyze each pair of exposure and outcome. Subsequently, the results from multiple analyses are aggregated for comprehensive evaluation using both fixed-effect and random-effect models.



MR_cause

CAUSE is a Mendelian Randomization method using genome-wide summary statistics. CAUSE models correlated and uncorrelated horizontal pleiotropy in order to avoid false positives that can occur using other methods. You can find the tutorial here (https://jean997.github.io/cause/ldl_cad.html).

[TwoSampleMR](#)[MVMR](#)[Multi-response](#)[MR instrument](#)[Colocalisation](#)[MRomics](#)[MR circoos](#)[MR Forest](#)[chrpos2SNP](#)[Mediator_MR](#)[Proxy SNP](#)[LDSC](#)[SMR](#)[SMRplot](#)[MR_meta](#)[MR_CAUSE](#)[Clean Format](#)**exposure file**

Browse...

diabetes

Upload complete

outcome file

Browse...

Sepsis.g

Upload complete

p value

5e-06

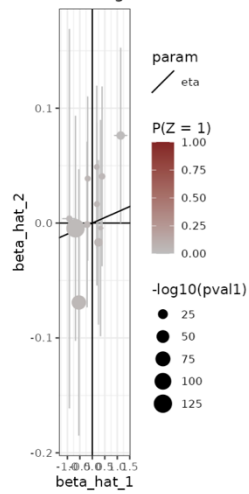
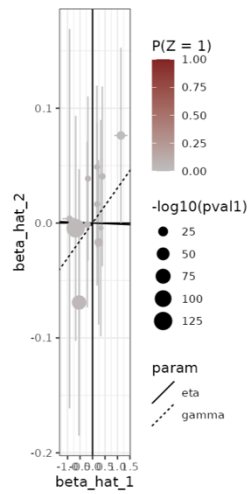
clump_kb

10000

clump_r2

0.001

Run

Sharing Model**Causal Model****ELPD Contribution**