

# Package ‘bGWAS’

October 9, 2023

**Type** Package

**Title** Bayesian Genome-Wide Association Study

**Description** Package regrouping functions to perform Bayesian Genome-Wide Association Studies (bGWAS). See McDaid et al (2017) for more information about the method.

**Version** 1.0.3

**License** GPL-2 | file LICENSE

**Author** Ninon Mounier

**Maintainer** Ninon Mounier <mounier.ninon@gmail.com>

**URL** <https://github.com/n-mounier/bGWAS>

**Encoding** UTF-8

**LazyDataCompression** xz

**RoxygenNote** 7.2.3

**Depends** R (>= 4.3.1),  
dplyr (>= 1.1.3),  
magrittr (>= 2.0.3),

**Suggests** testthat,  
knitr,  
remotes,  
rmarkdown

**Imports** calibrate (>= 1.7.7),  
data.table (>= 1.14.8),  
ggplot2 (>= 3.4.3),  
gplots (>= 3.1.3),  
qqman (>= 0.1.9),  
readr (>= 2.1.4),  
Rcpp (>= 1.0.11),  
rlang (>= 1.1.1),  
R.utils (>= 2.12.2),  
stringr (>= 1.5.0),  
tibble (>= 3.2.1),  
tidyr (>= 1.3.0),  
TwoSampleMR (>= 0.5.7)

**Remotes** MRCIEU/TwoSampleMR,

**VignetteBuilder** knitr

R topics documented:

all.equal.bGWAS . . . . .	2
bGWAS . . . . .	3
coefficients_plot_bGWAS . . . . .	5
extract_MRcoeffs_bGWAS . . . . .	6
extract_results_bGWAS . . . . .	6
get_RSquared_bGWAS . . . . .	8
heatmap_bGWAS . . . . .	8
list_files . . . . .	9
list_priorGWASs . . . . .	9
list_traits . . . . .	10
manhattan_plot_bGWAS . . . . .	10
print.bGWAS . . . . .	11
print_log_bGWAS . . . . .	12
select_priorGWASs . . . . .	12
SmallGWAS_Timmers2019 . . . . .	13
<b>Index</b>	<b>14</b>

---

all.equal.bGWAS	<i>Equality test for bGWAS objects</i>
-----------------	--

---

Description

Equality test for bGWAS objects

Usage

```
## S3 method for class 'equal.bGWAS'  
all(target, current)
```

Arguments

- target            an object of class bGWAS
- current          an object of class bGWAS

Value

all.equal

bGWAS

*bGWAS - main function***Description**

Performs a Bayesian GWAS from Summary Statistics, using publicly available results to calculate the prior effects of the SNPs and compare it to observed z-scores

**Usage**

```
bGWAS(
  name,
  GWAS,
  Z_matrices = "~/ZMatrices/",
  prior_studies = NULL,
  MR_threshold = 1e-06,
  MR_ninstruments = 3,
  MR_pruning_dist = 500,
  MR_pruning_LD = 0,
  MR_shrinkage = 1,
  stepwise_threshold = NULL,
  prior_shrinkage = NULL,
  sign_method = "p",
  sign_thresh = 5e-08,
  use_permutations = FALSE,
  res_pruning_dist = 500,
  res_pruning_LD = 0,
  save_files = FALSE,
  verbose = TRUE
)
```

**Arguments**

name	The name of the analysis (character)
GWAS	The path to the conventional GWAS of interest, the ID of the GWAS from the list of studies available (prior GWASs), or a <code>data.frame</code> (character, numeric or <code>data.frame</code> )
Z_matrices	The path to the folder containing Z-Matrices, default="~/ZMatrices/" (character)
prior_studies	The IDs of prior GWASs to use for the analysis, default=NULL, will include all the prior GWASs available (numeric vector)
MR_threshold	The threshold used to select strong instruments for MR, should be lower than 1e-5, default=1e-6 (numeric)
MR_ninstruments	The minimum number of strong instruments needed to use a prior GWAS, should be between 2 and 8, default=3 (numeric)
MR_pruning_dist	The distance used for pruning MR instruments (in Kb), should be between 10 and 1000, default=500 (numeric)

MR_pruning_LD	The LD threshold used for pruning MR instruments, should be between 0 and 1 (if 0, distance-based pruning is used), default=0 (numeric)
MR_shrinkage	The p-value threshold used for shrinkage before performing MR, should be between MR_threshold and 1 (no shrinkage), default=1 (numeric)
stepwise_threshold	The p-value threshold used for inclusion/exclusion of Prior GWASs during the stepwise selection approach, should be between 0.05 and 0.0005, default=NULL will use 0.05 divided by the number of Prior GWASs tested (numeric)
prior_shrinkage	The p-value threshold used for shrinkage before calculating the prior, should be between MR_threshold and 1, default=NULL will use MR_shrinkage (numeric)
sign_method	The method used to identify significant SNPs, should be "p" for p-value or "fdr" for false discovery rate, default="p" (character)
sign_thresh	The threshold used to identify significant SNPs, default="5e-8" (numeric)
use_permutations	A logical indicating if BF p-values should be estimated using the permutation approach, default=FALSE
res_pruning_dist	The distance used for pruning results (in Kb), should be between 10 and 1000, (if set to NULL, no pruning is done), default=500 (numeric)
res_pruning_LD	The LD threshold used for pruning results, should be between 0 and 1 (if 0, distance-based pruning is used), default=0 (numeric)
save_files	A logical indicating if the results should be saved as files, default=FALSE
verbose	A logical indicating if information on progress should be reported, default=TRUE

## Details

Name and GWAS are required arguments. If GWAS is a path to a file (regular or .gz) or a data.frame, it should contain the following columns :

SNPID (rs numbers) should be : rs, rsid, snp, snpid, rnpid

A1 should be : a1, alt, alts

A2 should be : a2, a0, ref

Z should be : z, Z, zscore

If Z is not present, it can be calculated from BETA and SE.

BETA should be : b, beta, beta1

SE should be : se, std

Note: in order to get rescaled (prior/posterior/corrected) effects, BETA and SE should be provided.

Z-Matrix files, containing Z-scores for all prior GWASs should be downloaded separately and stored in "~/ZMatrices" or in the folder specified with the argument Z\_matrices.

See [here](https://github.com/n-mounier/bGWAS) for more informations.

Use `list_priorGWASs()` to see all the prior GWASs available. Using one of them as your conventional GWAS (argument GWAS = numeric ID) will automatically remove it from the list of prior GWASs used to build the prior.

Use `select_priorGWASs()` to automatically select the prior GWASs to be included/excluded when building the prior (argument prior\_studies).

**Value**

bGWAS() returns an object of class "bGWAS".

Additionally, if save\_files=T, several files are created in the folder ./name/ :

- "PriorGWASs.tsv" - contains information about all prior GWASs (general info + status (used/excluded) + MR coefficients)
- "CoefficientsByChromosome.csv" - contains the MR estimates when masking the focal chromosome (22 coefficients / prior GWASs used for prior estimation)
- "PriorBFp.csv" - contains BF and p-values, prior, posterior and direct effects estimates for all SNPs
- "SignificantSNPs.csv" - contains BF and p-values, prior, posterior and direct effects estimates for a subset of significant SNPs

**Examples**

```
# Permorm bGWAS, using a small conventional GWAS included in the package (data.frame)
# and selecting a subset of studies for the prior
## Not run: top
data("SmallGWAS_Timmers2019")
MyStudies = select_priorGWASs(include_traits=c("Blood Pressure", "Education"),
                              include_files=c("cardiogram_gwas_results.txt",
                                                "All_ancestries_SNP_gwas_mc_merge_nogc.tbl.uniq.gz"))

# 6 Prior GWASs used
list_priorGWASs(MyStudies)

A = bGWAS(name="Test_UsingSmallDataFrame",
          GWAS = SmallGWAS_Timmers2019,
          prior_studies=MyStudies,
          MR_threshold = 1e-6,
          stepwise_threshold=0.05,
          save_files=T)

## End(Not run)

# Permorm bGWAS, using a conventional GWAS from the list of prior GWASs
## Not run: MyGWAS = 3
list_priorGWASs(MyGWAS)
# Coronary Artery Disease GWAS (CARDIoGRAM)
B = bGWAS(name = "Test_UsingGWASfromPriorGWASs",
          GWAS = MyGWAS)

## End(Not run)
```

---

coefficients\_plot\_bGWAS

*Coefficients Plot from bGWAS results*

---

**Description**

Creates a Coefficients Plot (causal effect of each Prior GWASs)

**Usage**

```
coefficients_plot_bGWAS(obj, save_file = F, file_name = NULL)
```

**Arguments**

obj	an object of class bGWAS created using <a href="#">bGWAS()</a>
save_file	A logical indicating if the graphic should be saved, default=FALSE, graphic will be displayed on the on-screen device
file_name	The name of the file saved (is save_file is TRUE) default=NULL, will used NameOfYourAnalysis_CoefficientsPlot.png

**Value**

a Coefficients Plot

---

```
extract_MRcoeffs_bGWAS
```

*Extract MR coefficients from bGWAS results*

---

**Description**

Extracts MR coefficients (multivariable genome-wide and per-chromosome estimates)

**Usage**

```
extract_MRcoeffs_bGWAS(obj)
```

**Arguments**

obj	an object of class bGWAS created using <a href="#">bGWAS()</a>
-----	--

**Value**

a tibble containing the MR coefficients (1 estimate using all chromosomes + 22 estimates with 1 chromosome masked)

---

```
extract_results_bGWAS
```

*Extract SNPs results from bGWAS results*

---

**Description**

Extracts SNPs results from bGWAS results (BFs, p-value, prior, posterior and direct effects, depending on the value of the parameter results)

**Usage**

```
extract_results_bGWAS(obj, SNPs = "significant", results = "BF")
```

## Arguments

obj                    an object of class bGWAS created using `bGWAS()`  
 SNPs,                "all" / "significant", default="significant"  
 results,            "BF" / "posterior" / "direct" / "everything", default="BF"

## Details

For all value of results, basic informations about the SNPs will be returned:

rsid : rs number  
 chr<sub>UK10K</sub> : chromosome (obtained from UK10K data)  
 pos<sub>UK10K</sub> : position (obtained from UK10K data)  
 alt : alternative (effect) allele  
 ref : reference allele  
 beta : observed effect size (if possible)  
 se : observed effect size (if possible)  
 z\_obs : observed Z-score

In addition, if results = "BF" the following information will be returned:

mu\_prior\_estimate : prior effect estimate (z-score scale)  
 mu\_prior\_std\_error : prior effect standard error (z-score scale)  
 beta\_prior\_estimate : prior effect estimate (beta scale, if possible)  
 beta\_prior\_std\_error : prior effect standard error (beta scale, if possible)  
 BF : Bayes Factor  
 BF\_p : Bayes Factor p-value  
 BF\_fdr : Bayes Factor FDR (only if FDR used to identify significant SNPs)

Alternatively, if results = "posterior" the following information will be returned:

mu\_posterior\_estimate : posterior effect estimate (z-score scale)  
 mu\_posterior\_std\_error : posterior effect standard error (z-score scale)  
 beta\_posterior\_estimate : posterior effect estimate (beta scale, if possible)  
 beta\_posterior\_std\_error : posterior effect standard error (beta scale, if possible)  
 z\_posterior : posterior Z-score  
 p\_posterior : posterior effect p-value  
 fdr\_posterior : posterior effect FDR (only if FDR used to identify significant SNPs)

Alternatively, if results = "direct" the following information will be returned:

mu\_direct\_estimate : direct effect estimate (z-score scale)  
 mu\_direct\_std\_error : direct effect standard error (z-score scale)  
 beta\_direct\_estimate : direct effect estimate (beta scale, if possible)  
 beta\_direct\_std\_error : direct effect standard error (beta scale, if possible)  
 z\_direct : direct Z-score  
 p\_direct : direct effect p-value  
 fdr\_direct : direct effect FDR (only if FDR used to identify significant SNPs)  
 CRR : corrected to raw ratio (ratio between direct effect and observed effect)

Alternatively, if results = "everything" all the results described above will be returned (possible only if SNPs = "all").

**Value**

a tibble containing the results for all / significant SNPs

---

get_RSquared_bGWAS	<i>Get squared correlation between observed and prior effects from bGWAS results</i>
--------------------	--

---

**Description**

Returns squared correlation between observed and prior effects, for different subsets of SNPs (all, the ones having at least a moderate effects - p-value < 0.001 -, MR instruments)

**Usage**

```
get_RSquared_bGWAS(obj, SNPs = "all")
```

**Arguments**

obj	an object of class bGWAS created using <a href="#">bGWAS()</a>
SNPs,	"all" / "moderate" / "instruments"

**Value**

a squared correlation

---

heatmap_bGWAS	<i>Heatmap of SNP effects on prior traits from bGWAS results</i>
---------------	--

---

**Description**

Creates a heatmap of SNP effects on prior traits

**Usage**

```
heatmap_bGWAS(obj, SNPs = NULL, save_file = F, file_name = NULL)
```

**Arguments**

obj	an object of class bGWAS created using <a href="#">bGWAS()</a>
SNPs	A data.frame containing the SNPs (rsid) to use in the first column, and optionally the text that should be plotted in addition to rsid in the second column default=NULL.
save_file	A logical indicating if the graphic should be saved, default=FALSE, graphic will be displayed on the on-screen device
file_name	The name of the file saved (is save_file is TRUE) default=NULL, will used NameOfYourAnalysis_Heatmap.png

**Value**

a Heatmap



---

list_files	<i>List prior GWASs files</i>
------------	-------------------------------

---

**Description**

Lists the filenames of the prior GWASs

**Usage**

```
list_files(IDs = NULL, Z_matrices = "~/ZMatrices/")
```

**Arguments**

IDs	the IDs of the studies to print, default="~/ZMatrices/" will list all of them (numeric),
Z_matrices	The path to the folder containing Z-Matrices, default="~/ZMatrices/" (character)

**Value**

List of files

---

list_priorGWASs	<i>List prior GWASs</i>
-----------------	-------------------------

---

**Description**

Lists the studies that can be used as prior GWASs

**Usage**

```
list_priorGWASs(IDs = NULL, Z_matrices = "~/ZMatrices/")
```

**Arguments**

IDs	the IDs of the studies to print, default="~/ZMatrices/" will list all of them (numeric),
Z_matrices	The path to the folder containing Z-Matrices, default="~/ZMatrices/" (character)

**Value**

a tibble containing prior GWASs information

---

list_traits	<i>List prior GWASs traits Lists the traits of the prior GWASs</i>
-------------	--

---

### Description

List prior GWASs traits Lists the traits of the prior GWASs

### Usage

```
list_traits(Z_matrices = "~/ZMatrices/")
```

### Arguments

Z_matrices	The path to the folder containing Z-Matrices, default="~/ZMatrices/" (character)
------------	--

### Value

List of traits

---

manhattan_plot_bGWAS	<i>Manhattan Plot from bGWAS results</i>
----------------------	--

---

### Description

Creates a Manhattan Plot from bGWAS results (for performance, only SNPs with p-value or FDR < 0.05 are plotted)

### Usage

```
manhattan_plot_bGWAS(
  obj,
  save_file = F,
  file_name = NULL,
  annotate = T,
  SNPs = NULL,
  results = "BF"
)
```

### Arguments

obj	an object of class bGWAS created using <a href="#">bGWAS()</a>
save_file	A logical indicating if the graphic should be saved, default=FALSE, graphic will be displayed on the on-screen device
file_name	The name of the file saved (if save_file is TRUE) default=NULL, will used NameOfYourAnalysis_ManhattanPlot.png

annotate	A logical indicating if the significant SNPs identified in the analysis should be annotated on the plot, default=TRUE. If your results are not pruned or if you have a high number of significant SNPs, be aware that annotate=TRUE might decrease readability of the figure. You could define a set of SNPs to annotate using SNPs.
SNPs	A data.frame containing the SNPs (rsid) to annotate in the first column, and optionnally the text that should be plotted in the second column, and the color in the third column, default=NULL, only evaluated if annotate=TRUE.
results,	"BF" / "posterior" / "direct", default="BF"

### Details

If results = "BF", BF p-values / fdr-values will be used.

If results = "direct", direct effect p-values / fdr-values will be used.

If results = "posterior", posterior effect p-values / fdr-values will be used.

### Value

a Manhattan Plot

---

print.bGWAS	<i>Print a bGWAS object</i>
-------------	-----------------------------

---

### Description

Print a bGWAS object

### Usage

```
## S3 method for class 'bGWAS'
print(x, ...)
```

### Arguments

x	an object of class bGWAS
...	further arguments

### Value

print

---

print_log_bGWAS	<i>Print log from bGWAS results</i>
-----------------	-------------------------------------

---

### Description

Prints the log (everything that is printed during a bGWAS analysis) with verbose=TRUE)

### Usage

```
print_log_bGWAS(obj)
```

### Arguments

obj	an object of class bGWAS created using <a href="#">bGWAS()</a>
-----	--

---

select_priorGWASs	<i>Select prior GWASs</i>
-------------------	---------------------------

---

### Description

Allow the quick selection of a subset of prior GWASs based on 2 criteria. First, include all the files specified (if all including parameters are NULL, include all studies), and then remove all the files specified (if all excluding parameters are NULL, keep all studies included at the step before)

### Usage

```
select_priorGWASs(
  include_files = NULL,
  include_traits = NULL,
  exclude_files = NULL,
  exclude_traits = NULL,
  Z_matrices = "~/ZMatrices/",
  verbose = F
)
```

### Arguments

include_files	list of file names (see <a href="#">list_files()</a> ) (character)
include_traits	list of trait (see <a href="#">list_traits()</a> ) (character)
exclude_files	list of file names (see <a href="#">list_files()</a> ) (character)
exclude_traits	list of trait (see <a href="#">list_traits()</a> ) (character)
Z_matrices	The path to the folder containing Z-Matrices, default="~/ZMatrices/" (character)
verbose	boolean, default = FALSE

### Value

IDs (numeric) of studies that meet the criteria

**Examples**

```
## Not run:
AllStudies = list_priorGWASs()
list_traits()
MyStudies = select_priorGWASs(include_traits=c("Heart Rate", "Body Mass Index", "Smoking"))
AllStudies[AllStudies$ID %in% MyStudies, c("ID", "Name", "Trait", "File")]
## End(Not run)
```

---

SmallGWAS\_Timmers2019 *Association results between genotypes and parental lifespan (LifeGen Consortium).*

---

**Description**

Subset of the original dataset containing the estimated effect of SNPs on parental survival

**Usage**

```
SmallGWAS_Timmers2019
```

**Format**

A data frame with 100000 rows and 5 variables:

**rsid** rsid of the SNP

**a1** effect allele for the SNP

**a0** reference allele for the SNP

**beta** estimated effect size for the SNP

**se** standard error of the estimated effect size for the SNP

**Source**

<https://datashare.is.ed.ac.uk/handle/10283/3209>

# Index

## \* datasets

SmallGWAS\_Timmers2019, [13](#)

all.equal.bGWAS, [2](#)

bGWAS, [3](#), [6–8](#), [10](#), [12](#)

coefficients\_plot\_bGWAS, [5](#)

extract\_MRcoeffs\_bGWAS, [6](#)

extract\_results\_bGWAS, [6](#)

get\_RSquared\_bGWAS, [8](#)

heatmap\_bGWAS, [8](#)

list\_files, [9](#), [12](#)

list\_priorGWASs, [4](#), [9](#)

list\_traits, [10](#), [12](#)

manhattan\_plot\_bGWAS, [10](#)

print.bGWAS, [11](#)

print\_log\_bGWAS, [12](#)

select\_priorGWASs, [4](#), [12](#)

SmallGWAS\_Timmers2019, [13](#)