The combGWAS() in our package is designed to combine the univariate GWAS results (or association tests results on a larger number of SNPs) of multiple phenotypes to infer possible pleiotropy effects.

Below we explain the input and output of this function in detail.

Input of combGWAS():

For each phenotype, there should be an association results file with columns being SNPid, beta, SE, coded allele and p-value etc. The user can use any names for these columns, but must specify these column names to corresponding arguments in the function. Specifically, `snpid`, `coded_all` and `AF_coded_all` must be assigned explicitly and the corresponding columns must appear in the input datasets. The alleles are supposed called on positive strand. If not, the user should convert the coded allele to that on a positive strand in the result file.

At least one of `beta`(`SE`) and `Z` should be assigned. In particular, if "beta" method is implemented, `beta` and `SE` must be assigned.

`n_total` and/or `pvalue` can be missing in the input datasets.

Here are two hypothetical univariate GWAS result files. We only give first 10 rows of each file to show the format.

The files can be comma, space, or tab delimited. combGWAS() is able to detect the delimiter used in the file automatically.

Phen1GWAS.csv

SNPID	n_total	AF_coded_all	beta	SE	pval	coded_all
rs11700354	4131	0.204673	0.467787	1.18E-01	7.71E-05	A
rs443192	4131	0.013993	0.630154	1.60E-01	7.79E-05	A
rs7329695	4131	0.064885	0.630473	1.60E-01	7.77E-05	C
rs1015447	4131	0.012561	0.766581	1.94E-01	7.87E-05	C
rs12718223	4131	0.079402	0.7667	1.94E-01	7.74E-05	T
rs691163	4131	0.079467	1.183652	2.99E-01	7.74E-05	T
rs9516869	4131	0.064935	1.372606	3.48E-01	7.88E-05	G
rs443112	4131	0.033272	1.656353	4.20E-01	7.88E-05	A
rs8082605	4131	0.022356	3.821093	9.67E-01	7.73E-05	T

Phen2GWAS.csv

SNPID	n_total	AF_coded_all	beta	SE	pval	coded_all
rs11700354	4131	0.204673	2.819804	7.14E-01	7.88E-05	A
rs443192	4131	0.013993	1.182689	3.00E-01	7.91E-05	A
rs7329695	4131	0.064885	0.595338	1.51E-01	7.94E-05	C

rs1015447	4131	0.012561	2.943738	7.46E-01	7.97E-05	C
rs12718223	4131	0.079402	0.8755	2.22E-01	7.98E-05	T
rs691163	4131	0.079467	0.766317	1.94E-01	7.99E-05	T
rs9516869	4131	0.064935	0.514169	1.30E-01	8.00E-05	G
rs443112	4131	0.033272	8.172358	2.07E+00	8.04E-05	A
rs8082605	4131	0.022356	1.181883	3.00E-01	8.05E-05	T

no change of beta signs before combining

```
combGWAS(project="mv",traitlist=c("phen1","phne2",),traitfile=c("Ph
en1GWAS.csv", "Phen2GWAS.csv",), comb_method=c("z","chisq"),
betasign=c(1,1), snpid=" SNPID ",
beta="beta",SE="SE", coded_all="coded_all ", AF_coded_all="
AF_coded_all ", pvalue="pval")
```

change of beta signs before combining
the beta sign for the 2nd phenotype reversed

```
combGWAS(project="mv",traitlist=c("phen1","phne2",),traitfile=c("Ph
en1GWAS.csv", "Phen2GWAS.csv",), comb_method=c("z","chisq"),
betasign=c(1,-1), snpid=" SNPID ",
beta="beta",SE="SE", coded_all="coded_all ", AF_coded_all="
AF_coded_all ", pvalue="pval")
```

Output from combGWAS():

No value is returned. Instead, results are written to `outfile`(named as "project_traits_method.csv") in the current working directory. In the `outfile`, there are some new variables (listed below) created by the package along with the existing variables in the original datasets. The correlation matrix and the degree of freedom of "sumsq" method (if requested) will be written into "correlation_df.out".

`zi`	Z statistic for the ith phenotype in traitlist. They will appear in "z", "chisq" and "sumsq" methods.
`pi`	p-value for the ith phenotype in traitlist.
`beta`	combined statistic of "z" and "beta" methods.
`SE`	standard error of the combined statistic of "z" and "beta" methods.
`Z.comb`	Z statistic (Z.comb=beta/SE) of "z" and "beta" methods.
`betai`	beta statistic for the ith phenotype in traitlist. They will appear output for in "beta" method.
`chisq.comb`	combined and test statistic of "chisq" and "sumsq" methods.
`pval`	p-value of the combined statistic.
`meanN`	the mean sample size with phenotype and genotype for the genetic marker. N/A if

	`n_total` is not specified.
`minN`	the minimum sample size with phenotype and genotype for the genetic marker. N/A if `n_total` is not specified.
`maxN`	the maximum sample size with phenotype and genotype for the genetic marker. N/A if `n_total` is not specified.
`remark1`	The sign of beta will be flipped if coded alleles different between two datasets.
`remark2`	If the minimum eigen value of the covariance matrix is less than 0.01, we consider it as nearly singular and the analyses will stop.