Overview

About

OSCA (OmicS-data-based Complex trait Analysis) is a software tool written in C/C++ for the analysis of complex trait using multi-omics data. It is developed by Futao Zhang, Zhihong Zhu and Jian Yang at Institute for Molecular Bioscience, the University of Queensland. Bug reports or questions: Jian Yang <jian.yang@uq.edu.au>.

Functions currently supported are:

- Estimating the epigenetic (or transcriptomic) relationships between individuals from genome-wide DNA methylation (or gene expression) data.
- Estimating the proportion of phenotypic variance for a complex trait can be “explained” by all DNA methylation (or gene expression) probes.
- Mixed linear model based analysis to test for associations between DNA methylation (or gene expression) probes and a complex trait.
- Estimating the joint “effects” of all methylation (transcription) probes on a phenotype (e.g. BMI) in a mixed linear model (analogous to BLUP). These estimated effects can be used to

predict the phenotype in a new independent sample.

Note: Although the software is designed for gene expression and DNA methylation data, it can be applied to any other source of omics data including microbiome, proteome and brain connectome.

Credits

Futao Zhang and Jian Yang developed the methods, software and webpage. Zhihong Zhu contributed to the development of the ORM and OREML methods. Zhili Zheng provided the template of the website.

Questions and Help Requests

If you have any bug reports or questions please send an email to Jian Yang at jian.yang@uq.edu.au

Citations

EWAS method:
Zhang, F. et al. (2017) OSCA: a tool for omics-data-based complex trait analysis. In preparation.

Last update: 21 July 2017

Download

Executable Files (version 0.3)

The Linux version is available at: OSCA_Linux.zip.

The MacOS version is available at: OSCA_Mac.zip.

The Windows version is available at: OSCA_Win.zip.

The example files are in example.zip.

The executable files (binary code) are release under MIT lincense.

Update log

Version 0.30 (21 July 2017)

- Beata version released.

Data Management

Data Format

The DNA methylation (or gene expression) data are stored in a binary format for consistency and storage efficiency. We store the data in three separate files .oii (individual information, similar as a PLINK .fam file), .opi (probe information) and .bod (a binary file to store the DNA methylation or gene expression profiles).

myeed.oii

```
F01     I01     0       0       NA
F02     I01     0       0       NA
F03     I02     0       0       NA
...
```

Columns are family ID, individual ID, paternal ID, maternal ID and sex (1=male; 2=female; 0=unknown). Missing data are represented by "NA".

myeed.opi

```
1       probe101        924243  Gene01  +
1       probe102        939564  Gene01  -
1       probe103        1130681 Gene01  -
...
```

Columns are chromosome, probe ID (can be the ID of an exon or a transcript for RNA-seq data), physical position, gene ID and gene orientation.

myeed.bod

DNA methylation (or gene expression) data in binary format. Please do not try to open this file with a text editor.

Note: The text below are for advanced users only. We use the first 12 bytes of the .bod file to store the descriptive information for the data. The first 2 bytes are reserved for further extension of the software. The type of data is indicated by the 3rd byte (0 for gene expression data, 1 for DNA methylation data, and 2 for any other type of data). The type of value is indicated by the 4th byte (0 for DNA methylation beta value, 1 for DNA methylation m value and 2 for any other type of value). The number of individuals and number of probes are stored as integers (4 bytes each). For example, for a DNA methylation data set (0x01) of m values (0x01) on 1,342 (0x53e) individuals and 228,694 (0x37d56) CpG sites, the first 12 bytes of the .bod file are:

```
0101 0000 3e05 0000 567d 0300
```

The bytes afterwards are for the DNA methylation or gene expression data.

Make a binary file

To compile data in binary format

```
osca --efile myprofile.txt --methylation-beta --make-bod --out myprofile
```

- **--efile** reads a DNA methylation (or gene expression) data file in plain text format.
- **--methylation-beta** indicates methylation beta values in the file.
- **--make-bod** saves DNA methylation (or gene expression) data in binary format.
- **--out** saves data (or results) in a file.

myprofile.txt

```
FID     IID cg00000658  cg26036652      cg00489772      ...
F01     I01     0.909   0.845   0.41    ...
F01     I02     0.832   0.732   0.503   ...
...
```

This is a file with a header line that contains family ID, individual ID and names of probes. The column of family ID is optional. Please use the flag "--no-fid" for data without family ID.

```
osca --efile myprofile.txt  --methylation-beta --make-bod --no-fid --out myprofile
```

- **--no-fid** indicates data without family ID.

```
osca --efile myprofile.txt  --methylation-m --make-bod --out myprofile
```

- **--methylation-m** indicates DNA methylation m values in the file.

If the profile is of gene expression value.

```
osca --efile myprofile.txt  --gene-expression --make-bod --out myprofile
```

- **--gene-expression** indicates gene expression profiles in the file.

For any other type of data:

```
osca --efile myprofile.txt --make-bod --out myprofile
```

A binary file also can be made from a transposed file in text format.

```
osca --tefile mytprofile.txt --methylation-beta --make-bod --no-fid --out myprofile
```

mytprofile.txt

```
FID     F01     F01     ...
IID     I01     I02     ...
cg00000658      0.909   0.832   ...
cg26036652      0.845   0.732   ...
cg00489772      0.41    0.503   ...
...
```

The row of family ID is optional. Please use the flag "--no-fid" if there is no family ID in your data.

```
osca --tefile mytprofile.txt --methylation-beta --make-bod --no-fid --out myprofile
```

Update .opi file

Probe information are sometimes not available in the original DNA methylation (or gene expression) data. These information can be updated the command below.

```
osca --befile myprofile --update-opi annotated.opi
```

- **--befile** reads a DNA methylation (or gene expression) data file in binary format.
- **--update-opi** reads a fully annotated .opi file.

Make a text format file

```
osca --befile myprofile --make-efile --out myprofile.txt
```

- **--make-efile** saves the DNA methylation (or gene expression) data in text format.

```
osca --befile myprofile --make-tefile --out mytprofile.txt
```

- **--make-tefile** saves the DNA methylation (or gene expression) data in transposed text format.

Make a subset of profile data

To extract a probe

```
osca --befile myprofile --probe cg00000658 --make-bod --out mysubprofile
```

- **--probe** extracts a specified probe.

To exclude a probe

```
osca --befile myprofile --probe-rm cg00000658 --make-bod --out mysubprofile
```

- **--probe-rm** excludes a specified probe.

To extract a subset of probes

```
osca --befile myprofile --extract-probe probe.list --make-bod --out mysubprofile
```

- **--extract-probe** extracts a subset of probes.

probe.list

```
cg00000658
cg26036652
...
```

To exclude a subset of probes

```
osca --befile myprofile --exclude-probe probe.list --make-bod --out mysubprofile
```

- **--exclude-probe** excludes a subset of probes.

To extract a subset of individuals

```
osca --befile myprofile --keep indi.list --make-bod --out mysubprofile
```

- **--keep** extracts a subset of individuals.

indi.list

```
F01     I01
F01     I02
...
```

To exclude a subset of individuals

```
osca --befile myprofile --remove indi.list --make-bod --out mysubprofile
```

- **--remove** excludes a subset of individuals.

To extract a subset of genes

```
osca --befile myprofile --genes gene.list --make-bod --out mysubprofile
```

- **--genes** extracts a subset of genes.

gene.list

```
MAN1B1
EDEM2
...
```

To extract a gene

```
osca --befile myprofile --gene MAN1B1 --make-bod --out mysubprofile
```

- **--gene** extracts a gene.

To extract a subset of probes based on the order

```
osca --befile myprofile --from-probe probe0 --to-probe probe1 --make-bod --out mysubprofi
le
```

- **--from-probe** specifies the start probe.
- **--to-probe** specifies the end probe.

To extract a subset of probes in a genomic region centred at a specific probe

```
osca --befile myprofile --probe cg00000658 --probe-wind 500 --make-bod --out mysubprofile
```

- **--probe-wind** defines a window <kb> centred on a specified probe.

To extract a subset of probes on a chromosome

```
osca --befile myprofile --chr 1 --make-bod --out mysubprofile
```

- **--chr** specifies a chromosome to select probes.

To extract a subset of probes based on physical positions

```
osca --befile myprofile --from-probe-kb 100 --to-probe-kb 200 --make-bod --out mysubprofi
le
```

- **--from-probe-kb** specifies the start physical position of the probes.
- **--to-probe-kb** specifies the end physical position of the probes.

Merge binary files

```
osca --befile-flist mybod.flist --make-bod --out myprofile
```

- **--befile-flist** reads a file to get the full paths of the binary files.

mybod.flist

```
path1/my_bod1
path2/my_bod2
...
```

Methylation beta value to methylation m value and vice versa

```
osca --befile myprofile --m2beta --make-bod --out newprofile
```

- **--m2beta** calculates the methylation beta value from the methylation m value.

```
osca --befile myprofile --beta2m --make-bod --out newprofile
```

- **--beta2m** calculates the methylation m value from the methylation beta value.

Calculate variance and mean

```
osca --befile myprofile --get-variance --get-mean --out newprofile
```

- **--get-variance** calculates the variance of each probe.
- **--get-mean** calculates the mean of each probe.

newprofile.var.txt

```
cg00000957      0.000812127
cg00001349      0.00560701
...
```

newprofile.mean.txt

```
cg00000957      0.901574
cg00001349      0.860279
...
```

Data trimming

```
osca --befile myprofile --std 0.02 --make-bod --out newprofile
```

- **--std** removes the probes with the standard deviation smaller than a specified threshold.

```
osca --befile myprofile --upper-beta 0.8 --lower-beta 0.2 --make-bod --out newprofile
```

- **--upper-beta** removes the DNA methylation probes with the mean beta value larger than a specified threshold.
- **--lower-beta** removes the DNA methylation probes with the mean beta value smaller than a specified threshold.

Quality control with detection p-value

```
osca --befile myprofile --detection-pval-file dpval.txt --dpval-mth 0  --dpval-thresh 0.0
5 --make-bod --out newprofile
```

- **--detection-pval-file** reads a file that contains DNA methylation detection p-values.
- **--dpval-mth** specifies a method to do quality control with the detection p-values, 0 for removing the probes with one or more detection p-values violating a threshold (default as 0.05), and 1 for dropping the samples violating a proportion threshold (default as 1%) and simultaneously dropping the probes violating a proportion threshold (default as 1%).
- **--dpval-thresh** specifies a threshold of detection p-value.

```
osca --befile myprofile --detection-pval-file dpval.txt --dpval-mth 1  --ratio-probe 0.01
 --ratio-sample 0.01 --dpval-thresh 0.05 --make-bod --out newprofile
```

- **--ratio-probe** specifies a proportion threshold to remove probes.
- **--ratio-sample**specifies a proportion threshold to remove individuals.

the file "dpval.txt" is in the same format with a transposed profile file. Option "--no-fid" is also valid here.

Quality control with missing rate

```
osca --befile myprofile --missing-ratio-probe 0.01 --make-bod --out newprofile
```

- **--missing-ratio-probe** specifies a missing proportion threshold to remove probes.

Probe pruning

To be updated....

Probe clumping

To be updated....

Estimation of probe correlation structure

To be updated....

ORM

```
osca --befile myprofile --make-orm --out myorm
```

```
osca --befile myprofile --make-orm-bin --out myorm
```

--make-orm
or
--make-orm-bin estimates the omics relationship matrix (ORM) between pairs of individuals from a set of probes and save the lower triangle elements of ORM to binary files. This format is compatible with the GRM in the GCTA software.

```
osca --befile myprofile --make-orm-gz --out myorm
```

- **--make-orm-gz** estimates the ORM and save the lower triangle elements of ORM to compressed plain text files.

```
osca --befile myprofile --make-orm --orm-alg 1 --out myorm
```

- **--orm-alg** specifies the algorithm to estimate the ORM. 1 for standardized data of each probe, 2 for centred data of each probe and 3 for standardized data of each individual. The default option is 1.

Note that although we describe the options above using DNA methylation and gene expression data, all the options can be applied to any other source of omics data mentioned

above.

Principal Component Analysis

```
osca --orm myorm --pca 20 --out mypca
```

- **--orm** reads the ORM binary files.
- **--pca** conducts principal component analysis and saves the first n (default as 20) PCs.

mypca.eigenval

```
122.014
95.3064
70.8055
...
```

mypca.eigenvec

```
R06C01  R06C01  0.012637        -0.00328532     0.0263842       -0.00595246
R05C02  R05C02  0.0106692       -0.0184545      -0.0159826      0.013951
R04C02  R04C02  0.0024727       -0.0118185      -0.0256503      -0.0106235
...
```

Users can also manipulate the ORM in the analysis.

```
osca --orm myorm --keep indi.list --pca 20 --out mypca
```

```
osca --orm myorm --remove indi.list --pca 20 --out mypca
```

```
osca --orm myorm --orm-cutoff 0.05 --pca 20 --out mypca
```

- **--orm-cutoff** removes one of a pair of individuals with estimated omics relationships larger than the specified cut-off value.

```
osca --merge-orm myorm.flist --pca 20 --out mypca
```

- **--merge-orm** reads multiple ORMs in binary format.

myorm.flist

```
myorm0
myorm1
myorm2
...
```

OREML

```
osca --reml --pheno my.phen --out myreml
```

- **--reml** performs REML (restricted maximum likelihood) analysis. This option is usually followed by the option --orm (one ORM) or --merge-orm (multiple ORMs) to estimate the variance explained by the probes that were used to estimate the omics relationship matrix.
- **--pheno** reads phenotype data from a plain text file. Missing value should be represented by "NA".

my.phen

```
R06C01  R06C01  32.6332
R05C02  R05C02  23.9411
R04C02  R04C02  29.7441
...
```

```
osca --reml --orm myorm --pheno my.phen --mpheno 1 --out myreml
```

- **--mpheno** reads a list of comma-delimited trait numbers if the phenotype file contains more than one trait, e.g. "1,3" tells OSCA to take the first and the third trait for analysis. OSCA always takes the first trait for analysis unless this option is specified.

  NOTE: current version only supports single trait analysis in one run.

```
osca --reml --orm myorm --pheno my.phen --out myreml
```

```
osca --reml --orm myorm --pheno my.phen --keep indi.list --out myreml
```

```
osca --reml --orm myorm --pheno my.phen --orm-cutoff 0.05 --out myreml
```

With multiple ORMs

```
osca --reml --merge-orm myorm.flist --pheno my.phen --out myreml
```

```
osca --reml --merge-orm myorm.flist --pheno my.phen --reml-alg 0 --out myreml
```

- **--reml-alg** specifies the algorithm to do REML iterations, 0 for average information (AI), 1 for Fisher-scoring and 2 for EM. The default option is 0, i.e. AI-REML, if this option is not specified.

```
osca --reml --merge-orm myorm.flist --pheno my.phen --reml-maxit 100 --out myreml
```

- **--reml-maxit** specifies the maximum number of iterations. The default number is 100 if this option is not specified.

```
osca --reml --orm myorm --pheno my.phen --covar my.covar --out myreml
```

- **--covar** reads discrete covariates from a plain text file.

my.covar

```
R06C01  R06C01  F       0
R05C02  R05C02  F       1
R04C02  R04C02  M       1
...
```

```
osca --reml --orm myorm --pheno my.phen --qcovar my.qcovar --out myreml
```

- **--qcovar** reads quantitative covariates from a plain text file.

my.qcovar

```
R06C01  R06C01  25
R05C02  R05C02  16
R04C02  R04C02  30
...
```

```
osca --reml --orm myorm --pheno my.phen --reml-est-fix --out myreml
```

- **--reml-est-fix** displays the estimates of fixed effects on the screen.

```
osca --reml --orm myorm --pheno my.phen --reml-no-lrt --out myreml
```

- **--reml-no-lrt** turns off the LRT.

EWAS

Mixed Linear Model Association

```
osca --mlma --befile myprofile --pheno my.phen --out my
```

If you have already computed the ORM

```
osca --mlma --befile myprofile --pheno my.phen --orm myorm --out my
```

- **--mlma** initiates an MLM based association analysis including the target probe (the probe to be tested for association) in the ORM. The results will be saved in a plain text file with **.mlma** as the filename extension.

my.mlma

```
Chr     Probe   bp      Gene    Orientation     b       se      p
1       cg00003287      201346149       TNNT2   -       -0.156  0.597   0.794
1       cg00008647      207082900       IL24    +       0.032   0.354   0.926
1       cg00009292      50882082        DMRTA2  -       0.120   1.182   0.919
...
```

This is a text file with headers. Columns are chromosome, probe, probe BP, gene, orientation, effect size, standard error and p-value.

```
osca --mlma-loco --befile myprofile --pheno my.phen --out my
```

- **--mlma-loco** initiates an MLM based association analysis with the chromosome where the target probe is located excluded from the ORM. The results will be saved in a plain text file with **.loco.mlma** as the filename extension.

```
osca --mlma --lxpo 0.1 --befile myprofile --pheno my.phen --out my
```

```
osca --mlma-loco --lxpo 0.1 --befile myprofile --pheno my.phen --out my
```

- **--lxpo** specifies a percentage of probes to exclude from calculating the ORM. This option should accompany **--mlma** or **--mlma-loco**. It will first conduct a linear-regression-based association analysis and then excludes a percentage of top associated probes from the ORM.

Linear Regression

```
osca --befile myprofile --pheno my.phen --linear --out my
```

```
osca --befile myprofile --pheno my.phen --qcovar my.qcovar --covar my.covar --linear --ou
t my
```

- **--linear** saves linear regression statistics to a plain text file.

my.linear

```
probeChr        ProbeID Probe_bp        BETA    SE      P       NMISS
1       cg00003287      201346149       -0.0594 0.531   9.11e-01        1337
1       cg00008647      207082900       0.5003  0.263   5.72e-02        1337
```

```
1       cg00009292      50882082        1.0512  1.026   3.06e-01        1337
...
```

This is a text file with headers. Columns are chromosome, probe, probe BP, effect size, standard error, p-value and number of non-missing individuals.

EWAS simulation

The phenotypes are simulated based on a set of real DNA methylation (or gene expression) data and a simple model $y= sum(x_ib_i)+\varepsilon$, where y is a vector of phenotypes, x_i is a vector of raw DNA methylation (or gene expression) profile or standardized profile of the i-th "causal" probe, b_i is the effect of the i-th causal probe and $\varepsilon \sim N(0,var(sum(x_ib_i))(1/h^2-1))$ is a vector of residual effect.

```
osca --simu-qt --simu-hsq 0.1 --befile myprofile --simu-causal-loci mycausal.list --out m
ypheno
```

- **--simu-qt** smulates a quantitative trait.
- **--simu-hsq** specifies the proportion of variance in phenotype explained by the causal probes. The default value is 0.1.
- **--simu-causal-loci** reads a list of probes as causal probes. If the effect sizes are not specified in the file, they will be generated from a standard normal distribution.

mycausal.list

```
cg04584301      1.55182
cg04839274      -0.106226
cg16648571      0.0257417
...
```

This is a text file with no headers. Columns are probe ID and effect size.

```
osca --simu-qt --simu-hsq 0.1 --simu-eff-mod 0 --befile myprofile --simu-causal-loci myca
usal.list --out mypheno
```

- **--simu-eff-mod** specifies whether or not to standardize the causal probe, 0 for standardized profile and 1 for raw profile. The default value is 0.

```
osca --simu-cc 100 300 --simu-hsq 0.1 --simu-k 0.1 --befile myprofile --simu-causal-loci
 mycausal.list --out mypheno
```

- **--simu-cc** simulates a case-control trait and specifies the number of cases and the number of controls.

- **--simu-k** specifies the disease prevalence. The default value is 0.1 if this option is not specified.

Prediction Analysis

```
osca --reml --orm myorm --pheno my.phen --reml-pred-rand --out myblp
```

- **--reml-pred-rand** predicts the random effects by the BLUP (best linear unbiased prediction) method. This option is to estimate the aggregated effect of all the probes (used to compute the ORM) to the phenotype of an individual. The aggregated omics effects of all the individuals will be saved in a plain text file ***.indi.blp**.

myblp.indi.blp

```
R06C01  R06C0   1.02275 0.07065 1.05692 1.05692
R05C02  R05C02  0.18653 0.27059 0.86650 0.86650
R04C02  R04C02  -0.1982 -0.1673 -0.9209 -0.9209
...
```

This is a text file with no headers. Columns are family ID, individual ID, an intermediate variable, the aggregated omics effect, another intermediate variable and the residual effect.

```
osca --befile myprofile --blup-probe myblp.indi.blp --out myblp
```

- **--blup-probe** calculates the BLUP solutions for the probe effects.

myblp.probe.blp

```
cg04584301      -0.000654646
cg04839274      0.000602484
cg16648571      -4.70356e-05
...
```

This is a text file with no headers. Columns are probe ID and BLUP of the probe effect.

```
osca --befile myprofile --score myblp.probe.blp --out myscore
```

```
osca --befile myprofile --score myblp.probe.blp 1 2 --out myscore
```

- **--score** reads score files for probes and generates predicted omics profiles for individuals. (Note that this option largely follows the --score option in PLINK.) It allows users to specify the column numbers for probe ID and score (the default values are 1 and 2 as shown in the example above).

myscore.profile

```
FID     IID     PHENO   CNT     SCORE
131000028       422572  -9      20000   -7.269198e-07
131000031       243421  -9      20000   9.322096e-06
131000179       338728  -9      20000   -1.250443e-05
...
```

This is a text file with headers. Columns are family ID, individual ID, phenotype, Number of non-missing probes and score.

```
For example
       In the score file:
cg04584301      -0.065
cg04839274      0.060
cg16648571      -1.03

In the DNA methylation data:
FID     IID     cg04584301      cg04839274      cg16648571
R05C02  R05C02  0.18653 0.27059 0.86650

The score should be:
  ( 0.18653*(-0.065) + 0.27059*0.060 + (-1.03)*0.86650 ) / 3 = (-0.8883841) / 3 = -0.296
```

```
osca --befile myprofile --score myblp.probe.blp --score-has-header --out myscore
```

- **--score-has-header** indicates probe score file has headers.

Options Reference

--befile	reads a DNA methylation (or gene expression) data file in binary format
--befile-flist	reads a file to get the full paths of the binary files
--beta2m	calculates the methylation m value from the methylation beta value
--blup-probe	calculates the BLUP solutions for the probe effects
--chr	specifies a chromosome to select probes
--covar	reads discrete covariates from a plain text file
--detection-pval-file	reads a file that contains DNA methylation detection p-values
--dpval-mth	specifies a method to do quality control with the detection p-values
--dpval-thresh	specifies a threshold of detection p-value

--efile	reads a DNA methylation (or gene expression) data file in plain text format
--exclude-probe	excludes a subset of probes
--extract-probe	extracts a subset of probes
--from-probe	specifies the start probe
--from-probe-kb	specifies the start physical position of the probes
--gene	extracts a gene
--genes	extracts a subset of genes
--gene-expression	indicates gene expression profiles in the file
--get-mean	calculates the mean of each probe
--get-variance	calculates the variance of each probe
--keep	extracts a subset of individuals
--linear	saves linear regression statistics to a plain text file
--lower-beta	removes the DNA methylation probes with the mean beta value smaller than a specified threshold
--lxpo	specifies a percentage of probes to exclude from calculating the ORM
--m2beta	calculates the methylation beta value from the methylation m value
--make-bod	saves DNA methylation (or gene expression) data in binary format
--make-efile	saves the DNA methylation (or gene expression) data in text format
--make-orm	estimates the omics relationship matrix (ORM) and save the lower triangle elements of ORM to binary files
--make-orm-bin	estimates the omics relationship matrix (ORM) and save the lower triangle elements of ORM to binary files
--make-orm-gz	estimates the omics relationship matrix (ORM) and save the lower triangle elements of ORM to compressed plain text files
--make-tefile	saves the DNA methylation (or gene expression) data in transposed text format

--merge-orm	reads multiple ORMs in binary format
--methylation-m	indicates methylation m values in the file
--methylation-beta	indicates methylation beta values in the file
--mlma	initiates an MLM based association analysis including the target probe (the probe to be tested for association) in the ORM
--mlma-loco	initiates an MLM based association analysis with the chromosome where the target probe is located excluded from the ORM
--mpheno	reads a list of comma-delimited trait numbers if the phenotype file contains more than one trait
--missing-ratio-probe	specifies a missing proportion threshold to remove probes
--no-fid	indicates data without family ID
--orm	reads the ORM binary files
--orm-alg	specifies the algorithm to estimate the ORM
--orm-cutoff	removes one of a pair of individuals with estimated omics relationships larger than the specified cut-off value
--out	saves data (or results) in a file
--pca	conducts principal component analysis and saves the first n (default as 20) PCs
--pheno	reads phenotype data from a plain text file
--probe	extracts a specified probe
--probe-wind	defines a window centred on a specified probe
--probe-rm	excludes a specified probe
--qcovar	reads quantitative covariates from a plain text file
--ratio-probe	specifies a proportion threshold to remove probes
--ratio-sample	specifies a proportion threshold to remove individuals

--reml	performs REML (restricted maximum likelihood) analysis
--reml-alg	specifies the algorithm to do REML iterations
--reml-est-fix	displays the estimates of fixed effects on the screen
--reml-maxit	specifies the maximum number of iterations
--reml-no-lrt	turns off the LRT
--reml-pred-rand	predicts the random effects by the BLUP (best linear unbiased prediction) method
--remove	excludes a subset of individuals
--score	reads score files for probes and generates predicted omics profiles for individuals
--score-has-header	indicates probe score file has headers
--simu-causal-loci	reads a list of probes as causal probes
--simu-cc	simulates a case-control trait and specifies the number of cases and the number of controls
--simu-eff-mod	specifies whether or not to standardize the causal probe
--simu-hsq	specifies the proportion of variance in phenotype explained by the causal probes
--simu-k	specifies the disease prevalence. The default value is 0.1 if this option is not specified
--simu-qt	simulates a quantitative trait
--std	removes the probes with the standard deviation smaller than a specified threshold
--to-probe	specifies the end probe
--to-probe-kb	specifies the end physical position of the probes
--update-opi	reads a fully annotated .opi file
--upper-beta	removes the DNA methylation probes with the mean beta value larger than a specified threshold