Model-Based Multifactor Dimensionality Reduction
Continuous trait

MBMDR-3.1.0 is a software that is able to detect multiple sets of significant gene-gene and/or gene-environment interactions in relation to a trait of interest, while efficiently controlling type I error rates. The continuous version is dedicated to continuous trait and by default performs a 2order interactions analysis. To see the command line help, type

mbmdr.out --help --continuous

The instruction to perform an MB-MDR analysis is as follows:

```
mbmdr.out --continuous [options] 'mbmdrFile'
```

The 'mbmdrFile' must be represented using the following structure:

Tr S1 S2 ... Sm
X1 Y11 Y12 ... Y1m
... ...
Xk Yk1 Yk2 ... Ykm

The first line is a title line: Tr is the name of the trait and the Sj's are the names of the markers (SNPs or environment variables).

The first column contains the trait values: Xi is a continuous value representing the state of the i^{th} subject. The other columns contain the markers values:

- if Sj is a SNP: Yij is 0 if the i^{th} subject is homozygous for the first allele, 1 if heterozygous and 2 if homozygous for the second allele.
- if Sj is an environment variable: the X different possible values of the environment variables should be coded 0, 1, ..., X-1.

Missingness: a missing Yij value must be coded -9. Missing Xi values are not accepted (the program will ignore the subject and generate a warning)

If your dataset is in PLINK format, you can first use the following command line to create the 'mbmdrFile'

```
mbmdr.out --plink2mbmdr --continuous -ped 'pedFile' -map 'mapFile' -o 'mbmdrFile' -tr 'trFile'
```

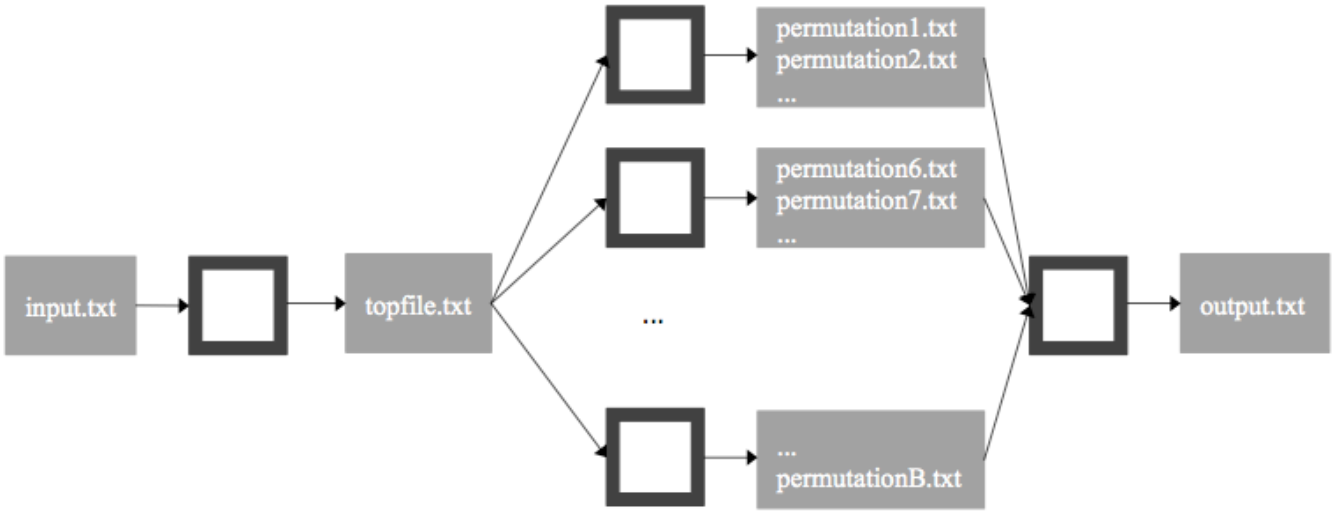
This command will also generate a translation file ‘trFile’, giving the chosen label for each genotype of each SNP.

The different options of the program are: (the options between square brackets are not mandatory)

OPTION	DESCRIPTION	DEFAULT
[-n INT]	Sets the number of p-values to compute	1000
[-p INT]	Sets the permutation amount	999
[-r INT]	Sets the random sequence parameter	Random value generated for you
[-m INT]	Sets the minimum group size to be statistically relevant	10
[-x DOUBLE]	Sets the cutoff value for the chi-square test	0.1
[-o STRING]	Sets the output file name	'infileprefix'_output.txt
[-a STRING]	Sets the adjust method to be used: CODOMINANT, ADDITIVE or NONE	CODOMINANT
[--1D]	Use this flag to do a main effect analysis	2-order
[--3D]	Use this flag to do a three-order interaction analysis	2-order
[-f STRING]	Filter out the pairs that do not contain the given marker (useful for testing one environment variable vs all snps ; can be combined with --3D to test one environment variable vs all pairs of snps)	Compute all pairs
[--no-permut]	Use this flag if you are only interested in the ranking of the markers, not the p-values (1000 times faster!)	Compute p-values
[--no-ranking]	By default, the program does not work with the continuous values themselves, but sorts them and uses the rank instead of the value. Use this flag to avoid this procedure.	Work with ranks
[--verbose]	Verbose the HLO matrices generated for the top pairs (H=high risk L=low risk O=no evidence N=no test performed)	Do not verbose the HLO matrices
[--no-progress]	Do not verbose the kind of progress bar	Verbose progress

Parallel Workflow

We have developed a parallel workflow, composed of three steps, allowing you to run the permutations on different machines. The first step prepares the parallel work, the second step does the parallel work and the last step sums up the results. To see the command line help, type *mbmdr.out --help --parallel*



Step 1 (on one machine)

```
mbmdr.out --continuous --pstep1 [options] 'mbmdrFile'
```

At step 1 you can use the options -n, -m, -a, --1D, --3D, -f, --no-ranking and --verbose described above. In addition, you can use the "-t" option to specify the name of the top file (default: topFile.txt)

STEP 2 (on N machines, let us number them 1, 2,..., B)

```
mbmdr.out --continuous --pstep2 -p INT -o STRING [options] 'mbmdrFile'
```

At step 2, the -p and -o options are mandatory. The name of the output file must be 'xxxi.txt' where xxx is a common prefix and i the machine number. Furthermore, you can also use the options –r, -a, --1D, --3d, -f, --no-ranking and -t described above.

STEP 3 (on one machine)

```
mbmdr.out --continuous --pstep3 -c STRING -q INT [options] 'mbmdrFile'
```

At step 3, we introduce two new mandatory options: "-c" sets the common prefix 'xxx' of the files generated at step 2 and "-q" sets the quantity of files generated at step 2 (= the amount of machines). Furthermore, you can also use the -p and -o options.