The program is coded in R language and is delivered as a main folder (e.g., named "mTDT" here) containing:

 * 1 sub-folder "libs" containing the R script "modules.R", for programs definition
 * 1 R script "mtdt.R", the main program
 * 3 example input files: "MTDT_pedfile.ped", "MTDT_phenfile.phen" and "MTDT_mapfile.map".

## 1) INPUT FILES FORMAT

       * No header in any of the 3 input files
       * Columns in all input files are tab separated

       (i) the PED file: begins with 6 columns following by genotypes columns, respecting the following order
             -       family ID
             -       individual ID
             -       Father ID
             -       Mother ID
             -       sex
             -       affection (this column will not be taken into account, but its presence is required)

             -       genotypes columns (...)

             **!!! ADVICES:** The 'ped' file can contain several markers, but we advise at the running step to include a limited number of markers, e.g. 3 to 5, (using the --markerset option describe below) to avoid both (i) facing a huge number of alternative hypotheses to test and (ii) having sparse count tables for transmitted versus non transmitted alleles combinations across markers. As discussed in the method paper, M-TDT has not been optimized for screen multiple markers effect within a large number of markers, but within a set predefined subgroup of markers of interest, e.g. markers with intermediate marginal effect.

       (ii) the PHENOTYPE file: begins with 2 columns following by phenotypes columns, respecting the following order
             -       family ID
             -       individual ID
             -       phenotypes columns (...)

             NOTA: Only individuals present in the phenotype file will be included in the analysis

       (iii) MAP file: contains one column for genetic marker labels
             -       markers names (keeping the same order as in the 'ped' file)

## 2) SOME RULES

       -       alleles should be coded in integer (1: major allele, 2: minor allele for a bi-allelic marker for instance)
       -       allele separator is SPACE (e.g., "1 2")
       -       missing genotypes should be coded "0 0"

## 3) HOW TO RUN

       * Place yourself in directory containing the main program file (cd to "mTDT" folder here)
       * parameters
             **--pedfile:** name of 'ped' file
             **--mapfile:** name of 'map' file
             **--phenfile:** name of phenotype file
             **--markerset:** list indicating positions of the set of markers to analyze jointly (separated by "," without SPACE); position from the 'ped' file without counting the 6 first columns. If this "--markerset" option is not precise, the software will run like single-marker GWAS, screen all markers in the 'ped' file one-by-one in family-based association. A limited number of top SNPs from a GWAS screening can be subsequently used for multilocus analysis.
             **--phen:** the position of the phenotype (only one) to analyze from the 'phenotype' file without counting the 2 first columns.
             **--nbsim:** number of simulations for the computation of empirical P-values. The default value is 0 and will correspond to the case where only asymptotic P-values are provided. If sample size is limited or if there is LD (linkage disequilibrium) among markers analyzed, asymptotic theory is no more valid and then empirical P-values should be computed using simulations.
             **--nbcores:** number of cores to use for the run, if the user wants to speed up the run by using multiple cores, as the program can run in parallelized at the simulation step if included to obtain empirical P-values. So, this option is useful only if "--nbsim" option is used.

## 4) EXAMPLE OF RUN

       * shell command 1: *cd to "mTDT" folder*
       * shell command 2: *Rscript mtdt.R --nbsim 10 --nbcores 4 --markerset 1,2,3 --pedfile MTDT_pedfile.ped --mapfile MTDT_mapfile.map --phenfile MTDT_phenfile.phen --phen 1*

       NOTA: Results are written in this file "weighted_res_multilocus.csv", located in the main folder.