QTLMap, a software for QTL detection in outbred populations

O. Filangi[1], C. Moreno[2], H. Gilbert[3], A. Legarra[2], P. Le Roy[1], *J.M. Elsen*[2]

Introduction

QTLMap is a software dedicated to the detection of QTLs which control traits in outbred populations. QTLMap has been developed at INRA to analyze genetic variability of traits in various experimental designs with linkage analysis approaches. Such populations comprise mixture of large full sib and half sib families in 2 or 3 generation pedigrees. The underlying methods are interval mapping ones, all based on Linkage Analysis (Elsen et al., 1999, Knott et al., 1996, Gilbert and Le Roy 2007, Moreno et al., 2005) and on Linkage Disequilibrium Linkage Analysis (Legarra and Fernando, 2009).

Methods

The main feature of QTLMap is that it offers the possibility of varying the genetic model depending on:

- the QTLs model (biallelic in crosses between QTL homozygous lines, biallelic without hypothesis on the alleles fixation in grand parental lines, multiallelic, with haplotype identity between families)
- the number of QTLs segregating (one, two linked, several unlinked)
- the interactions between the QTL and fixed effects or between QTLs.

The phenotypic model can also vary depending on:

- the trait distributions (gaussian trait (including gene expression), survival trait or discrete trait)
- the number of traits each QTL influences
- the residual variance structure (homo or heteroskedasticity between half-sib families).

The statistical model may include fixed nuisance effects or covariates, and the polygenic relationships between the animals may be included using an animal model (Henderson, 1973). QTLMap deals with high numbers of markers (SNPs) thanks to an original algorithm for the estimation of transmission probabilities (Elsen et al., 2009).

[1] INRA Agrocampus Ouest, UMR0598, 35042 Rennes cedex, France
[2] INRA, UR0631, 31326 Castanet Tolosan cedex, France
[3] INRA AgroParisTech, UMR1313, 78352 Jouy en Josas cedex, France

The test statistic is an approximated likelihood ratio test. The distribution of the quantitative phenotype is modelized as a mixture of sub-distributions, which correspond to QTL genotypes, with proportions according to the dam phase probabilities. The test statistic can be simplified to its first order and then corresponds to the regression approach (Haley & Knott, 1992).

QTLMap includes a choice of simulation procedures which aim the calculation of rejection thresholds, by simulation or permutation, and the estimation of power for experimental designs optimization.

QTLMap is written in Fortran and uses the NAG or SLATEC librairies. QTLMap is available on the web, with a graphical user interface (http://qgp.jouy.inra.fr/), or in Linux, UNIX or Windows environments (upon request).

Input/Output Files

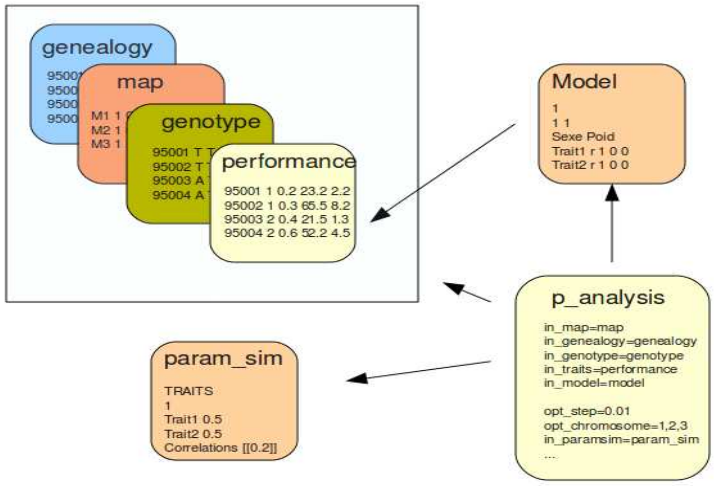

Figure 1 : set of files to analyze a dataset

To carry on an analysis, QTLMap requires 4 data files which give the genetic map, the pedigree, the markers genotypes and the traits values, and 1 model file which describes traits. Moreover, the analysis is fully parametrizable thanks to a parameter file (p_analysis) and runtime environment options:

- pedigree structure (half sib or mixture of full and half sib families with minimal size to take into account dam meioses)

- linkage step analysis (between 10^{-5} and 10^{5} Morgan)
- estimability of the model parameters (Cholesky's decomposition)
- algorithm for the likelihood optimization (quasi Newton (multiple implementations) or Simplex)

Finally, a supplementary file must be given to describe the simulation framework (param_sim). A full description of these files, with examples, is given in the QTLMap documentation.

In output, a general listing is produced. It presents the content of the data files. The most probable phases of the reproducers are built from available marker and pedigree information and then listed. In the third part, results of the genome scan are given. Details are given according to tests and models. Auxiliary files may also be edited, for example with likelihoods of each sire or dam family or transmission probabilities of haplotypes for each progeny.

References

Elsen, J.M., Mangin, B., Goffinet, B. et al. (1999) *Genet. Sel. Evol.* 31 : 213-224.
Elsen J.M., Filangi O., Gilbert H. et al. (2009) *Genet. Sel. Evol.* 41 : 50.
Gilbert, H. & Le Roy, P. (2007) *Genet. Sel. Evol*. 39(2) : 139-58.
Goffinet B., Le Roy P., Boichard D. et al. (1999) *Genet. Sel. Evol*., 31 : 341-350.
Haley, C.S. & Knott, S.A. (1992) *Heredity* 69(4) : 315-24.
Henderson, C.R. (1973) In: *Proceedings of the Animal Breeding and Genetics Symposium in Honour of Dr JL Lush*, American Society of Animal Science and Dairy Science association, Champaign, Il.
Knott, S.A., Elsen, J.M. and Haley, C. S. (1996) *Theor. Appl. Genet.* 93 : 71-80.
Le Roy P., Elsen J.M., Boichard D. et al. (1998). In: *Proc 6th WCGALP*, Volume 26, pp. 257– 260.
Mangin B., Goffinet B., Le Roy P. et al. (1999) *Genet. Sel. Evol*., 31 : 225-237.
Moreno, C.R., Elsen, J.M., Le Roy, P. & Ducrocq, V. (2005) *Genet. Res. Camb.* 85 : 139-149.