BIOINFORMATICS APPLICATIONS NOTE Vol. 18 no. 0 2002
Pages 1–2

QTL Express: mapping quantitative trait loci in simple and complex pedigrees

George Seaton[1], Chris S. Haley[2], Sara A. Knott[1], Mike Kearsey[3] and Peter M. Visscher[1,*]

[1]Institute of Cell, Animal and Population Biology, University of Edinburgh, West Mains Road, Edinburgh EH9 3JT, UK, [2]Roslin Institute (Edinburgh), Roslin, Midlothian EH25 9PS, UK and [3]School of Biosciences, University of Birmingham, Edgbaston, Birmingham B15 2TT, UK

Received on June 27, 2001; revised on August 24, 2001; accepted on September 5, 2001

ABSTRACT

Summary: QTL Express is the first application for Quantitative Trait Locus (QTL) mapping in outbred populations with a web-based user interface. User input of three files containing a marker map, trait data and marker genotypes allows mapping of single or multiple QTL by the regression approach, with the option to perform permutation or bootstrap tests.

Availability: http://qtl.cap.ed.ac.uk/

Contact: george.seaton@ed.ac.uk

INTRODUCTION

The mapping of Quantitative Trait Loci (QTL) is the first step towards the identification of genes and causal polymorphisms for traits of importance in agriculture and human medicine. We have developed fast, efficient and robust linear regression methods to map QTL in simple and complex pedigrees (e.g. Haley and Knott, 1992; Haley *et al.*, 1994; Knott *et al.*, 1996; Visscher *et al.*, 1996; Knott *et al.*, 1998; George *et al.*, 2000). The purpose of the development of QTL Express was to make these QTL mapping tools available to the wider scientific community via a user-friendly web-based user interface.

METHODS

We follow a robust two-step procedure for QTL mapping, by firstly determining the Identity-By-Descent (IBD) probabilities at specific chromosomal locations from multiple marker data, and secondly fitting a statistical model to the observations and IBD coefficients. Populations that are currently suitable for QTL Express are halfsib outbred populations (e.g. several or many sires, each with a number of offspring) and F_2 populations derived from crosses between either inbred or outbred lines.

Linear models are fitted to phenotypic data using a general linear model, i.e. allowing for additional fixed effects and covariates that explain trait variation. For the genetic component in the linear model, a single QTL or a two QTL model is fitted. Additional (known) QTL can be fitted through the use of cofactors. For populations derived from crosses, the model for the QTL can be specified in terms of an additive and a dominance effect, with the option of an interaction between the QTL and a fixed effect (e.g. sex or population). Additionally, for outbred line crosses, the QTL model can include a parent-of-origin (imprinting) effect.

For the F_2 populations, QTL are mapped that explain genetic variation between the founder lines. Implicitly, the founder lines are assumed to be fixed for alternative alleles at the QTL. For outbred lines, this assumption can be tested by including an interaction between the QTL and family (e.g. F_1 or grandparental sires). For the halfsib population structure, QTL are mapped that explain within-family variation, with the evidence for QTL segregation accumulated across the common parents (Knott *et al.*, 1996).

Permutation tests to set chromosome or genome-wide significance thresholds (Churchill and Doerge, 1994) and a bootstrap procedure to estimate the confidence interval of a QTL location (Visscher *et al.*, 1996) are implemented.

The format for each of the three input files is general and these are checked for errors. Segregation distortion and multi-point information content in each linkage group are calculated and graphically displayed. Results from fitting the linear models are displayed both in tables and graphically. Help files and example data files are available. Graphic output and the raw data underlying graphs can be downloaded.

We illustrate the graphical output from a halfsib analysis in Figure 1. The data set consisted of 693 progeny in 20 halfsib families. Progeny and the common parents were genotyped at nine microsatellite marker loci in the

*To whom correspondence should be addressed.

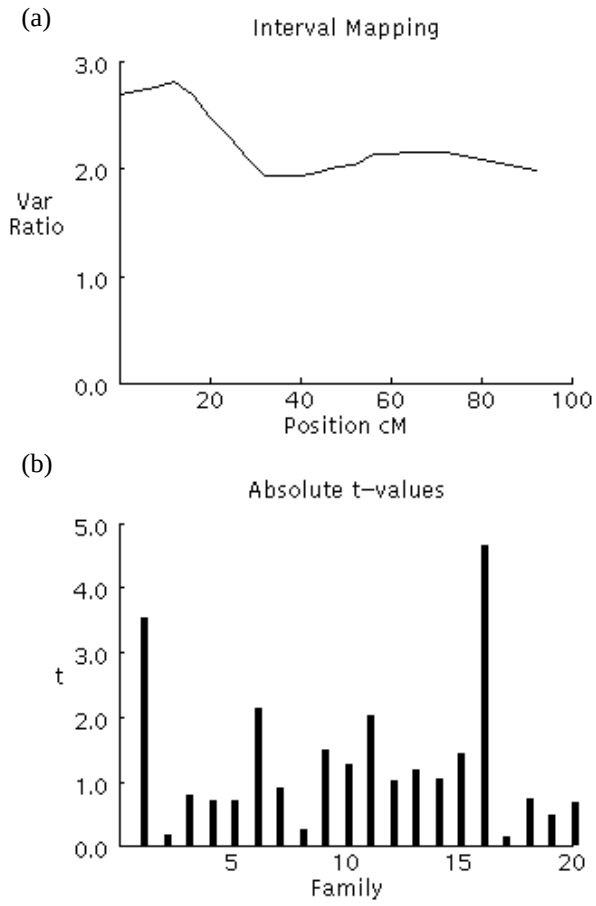

Fig. 1. Example of graphical output from QTL Express. (a) Results from fitting a single QTL model. (b) Individual test statistics per halfsib family.

same linkage group, and a phenotypic measurement was available on progeny. In Figure 1a the test statistic for one QTL at the given location versus no QTL, an *F*-ratio with 20 degrees of freedom in the numerator and 657 degrees of freedom in the denominator, is shown across the linkage group. In Figure 1b, the absolute *t*-values for each family are shown indicating the strength of evidence for a QTL at the location estimated in the across-family analysis. Permutation analyses on these data takes approximately 2.7 and 50.4 min, for 1000 iterations and a single position permutation and chromosome-wide permutation at the default 4 cM stepsize, respectively. A bootstrap analysis for 1000 samples takes approximately 88 min. Corresponding times for a single chromosome analysis on an F_2 population of size 400 and chromosome length of 60 cM are, approximately, 0.2, 2.7 and 2.6 min.

IMPLEMENTATION

The QTL Express site uses Java servlets. All numerical analysis, graphics and HTML output is handled by classes produced and executed by a 1.1.8 Java Development Kit (JDK 1.1.8) compiler and virtual machine respectively. The site's Apache web-server (http://www.apache.org) takes user submissions and passes them to its associated JServ module/java classes for processing by the QTL analysis servlet methods, results are then sent directly to the browser output stream of the user. Netscape Navigator/Communicator is the recommended browser (800 × 600 resolution) for use with QTL Express, as it requires less re-posting of data during analysis, although no specific browser is required. However as all analysis is performed server-side the users machine can be of a modest specification and a Java-enabled browser is not absolutely necessary.

CONCLUSIONS

To our knowledge, this is the first set of web-based QTL mapping tools which deal with a number of outbred population structures and allows for general linear models to be fitted. QTL Express is designed to enable researchers to analyze their data, but is also ideal for teaching purposes.

ACKNOWLEDGEMENTS

This work was funded by the UK Biotechnology and Biological Sciences Research Council. S.A.K. acknowledges support from the Royal Society.

REFERENCES

Churchill,G.A. and Doerge,R.W. (1994) Empirical threshold values for quantitative trait mapping. *Genetics*, **138**, 963–971.

George,A.W., Visscher,P.M. and Haley,C.S. (2000) Mapping quantitative trait loci in complex pedigrees: a two step variance component approach. *Genetics*, **156**, 2081–2092.

Haley,C.S. and Knott,S.A. (1992) A simple regression method for mapping quantitative trait loci in line crosses using flanking markers. *Heredity*, **69**, 315–324.

Haley,C.S., Knott,S.A. and Elsen,J.M. (1994) Mapping quantitative trait loci in crosses between outbred lines using least squares. *Genetics*, **136**, 1195–1207.

Knott,S.A., Elsen,J.M. and Haley,C.S. (1996) Methods for multiple marker mapping of quantitative trait loci in half-sib populations. *Theor. Appl. Genet.*, **93**, 71–80.

Knott,S.A., Marklund,L., Haley,C.S., Andersson,K., Davies,W., Ellegren,H., Fredholm,M., Hoyheim,B., Hannsson,I., Lundstrom,K., Moller,M. and Andersson,L. (1998) Multiple marker mapping of quantitative trait loci in an outbred cross between wild boar and large white pigs. *Genetics*, **149**, 1069–1080.

Visscher,P.M., Thompson,R. and Haley,C.S. (1996) Confidence intervals for QTL locations using bootstrapping. *Genetics*, **143**, 1013–1020.