Tutorial - Multiple-QTL Mapping (MQM) Analysis

Danny Arends, Pjotr Prins, Karl W. Broman and Ritsert C. Jansen

August 18, 2010

1 Introduction

Multiple QTL Mapping (MQM) provides a sensitive approach for mapping quantititive trait loci (QTL) in experimental populations. MQM adds higher statistical power compared to many other methods. The theoretical framework of MQM was introduced and explored by Ritsert Jansen, explained in the 'Handbook of Statistical Genetics' (see references), and used effectively in practical research, with the commercial 'mapqtl' software package. Here we present the first free and open source implementation of MQM, with extra features like high performance parallelization on multi-CPU computers, new plots and significance testing.

MQM is an automatic three-stage procedure in which, in the first stage, missing data is 'augmented'. In other words, rather than guessing one likely genotype, multiple genotypes are modeled with their estimated probabilities. In the second stage important markers are selected by multiple regression and backward elimination. In the third stage a QTL is moved along the chromosomes using these pre-selected markers as cofactors, except for the markers in the window around the interval under study. QTL are (interval) mapped using the most 'informative' model through maximum likelihood. A refined and automated procedure for cases with large numbers of marker cofactors is included. The method internally controls false discovery rates (FDR) and lets users test different QTL models by elimination of non-significant cofactors.

R/qtl-MQM has the following advantages:

- Higher power, as long as the QTL explain a reasonable amount of variation
- Protection against overfitting, because it fixes the residual variance from the full model. For this reason more parameters (cofactors) can be used compared to, for example, CIM
- Prevention of ghost QTL (between two QTL in coupling phase)
- Detection of negating QTL (QTL in repulsion phase)

The current implementation of R/qtl-MQM has the following limitations: (1) MQM is limited to experimental crosses F2, BC, and selfed RIL, (2) MQM does not treat sex chromosomes differently from autosomal chromosomes - though one can introduce sex as a cofactor. Future versions of R/qtl-MQM may improve on these points. Check the website and change log (`http://www.rqtl.org/STATUS.txt`) for updates.

Despite these limitations, *MQM*[1] is a valuable addition to the QTL mapper's toolbox. It is able to deal with QTL in coupling phase and QTL in repulsion phase. *MQM* handles missing data and has higher power to detect QTL (linked and unlinked) than other methods. R/qtl's *MQM* is faster than other implementations and scales on multi-CPU systems and computer clusters. In this tutorial we will show you how to use *MQM* for QTL mapping.

MQM is an integral part of the free *R/qtl* package [**?**, **?**, **?**] for the R statistical language[2].

2 A quick overview of *MQM*

These are the typical steps in an *MQM* QTL analysis:

[1] MQM should not be confused with composite interval mapping (CIM) [**?**, **?**]. The advantage of MQM over CIM is reduction of type I error (a QTL is indicated at a location where there is no QTL present) and type II error (a QTL is not detected) for QTL detection [**?**].

[2] We assume the reader knows how to load his data into R using the R/qtl `read.cross` function; see also the R/qtl tutorials [**?**] and book [**?**].