Implementation of robust methods for locating quantitative trait loci in R

Andreas Baierl and Andreas Futschik

Abstract:

One approach to QTL-mapping is to regress a quantitative trait of interest (e.g. height, yield, tumor count) on the observed genotypes at various positions on the genome (markers) in order to detect locations that influence the trait. We investigate additive and non-additive (epistatic) effects of markers, even if the corresponding main effects are not included in the model. As pointed out by Broman and Speed (2002), the task of selecting the correct markers can be treated as a model selection problem.

Since the distribution of the quantitative trait is often reported to be non-normal, we investigate the application of robust methods (M-estimators, L1-regression) and compare the results to least square regression to estimate the likelihood of the model.

We adapt the modified Bayesian Information Criterion (BIC) proposed by Bogdan et al (2004) that controls the overall type I error of detecting additive effects and pair wise interactions by an additional penalty term. Markers are chosen by an extended forward selection procedure with a backward elimination step.

The performance of the different methods is investigated by an extensive simulation study applying various error-distributions and genetic setups. We have implemented programs carrying out the proposed QTL-mapping approach in R. The robust regression estimators have been calculated using derived the procedures *rlm* from the *MASS*-package and *rq* from the package *quantreg*.

References:

Bogdan, M., J. K. Ghosh and R. W. Doerge, 2004. Modifying the Schwarz Bayesian Information Criterion to Locate Multiple Interacting Quantitative Trait Loci. Genetics, **167**: 989-999.

Broman, K. W. and T. P. Speed, 2002. A model selection approach for the identification of quantitative trait loci in experimental crosses. J Roy Stat Soc B, **64**: 641-656.