

Analyzing Data from a Splice Array Experiment

Yuan-Ming Zhang, Shizhong Xu
University of California, Riverside

Abstract:

Although epistasis is an important phenomenon in genetics and evolution of complex traits, epistatic effects are hard to estimate. The main problem comes from the over parameterized epistatic genetic model. An epistatic genetic model should include all potential pair-wise interaction effects of all loci. However, the model is saturated quickly as the number of loci increases. Therefore, a variable selection technique is usually considered to exclude those interactions with negligible effects from the model. With the variable selection technique, we may take a risk of missing some important interaction effects due to lack of full exploration of the extremely large parameter space of models. We develop a penalized maximum likelihood method to handle such an oversaturated epistatic effects model without resorting to any model selection techniques. In contrast to the Bayesian information criterion (BIC) and other information-based criteria where the penalty is independent of the parameter values, the penalized likelihood method developed here adopts a penalty that depends on the parameters. The penalized likelihood method allows spurious QTL effects to be shrunk towards zero while QTL with large effects are estimated with virtually no shrinkage. A simulation study shows that the penalized likelihood method can handle a model with the number of model effects 15 times larger than the number of observations. Simulation studies also show that results of the penalized maximum likelihood method are comparable to those of the Bayesian shrinkage analysis, but the computational speed of the penalized method is faster than the Bayesian method by orders of magnitude.