Chapter 1

**Introduction**

* 1. QTL mapping/Marker Regression

Numerous methods exist and are being developed to measure and find quantitative trait loci (QTL) effects. These methods can broadly fall into three main categories. These categories are Least-Square methods, maximum likelihood and Bayesian approaches. (Statistical Genetics of Quantitative Traits, Wu et al.) Each method has advantages and considerations that you would need to be aware before conducting analyses to find QTL effects. Brief discussions on a few of the methods are given to highlight some areas of consideration and how the methods proposed can handle such considerations.

Marker Regression would fall in the category of Least Squares approaches. If looking at one marker analysis general t-test and ANOVA procedures can be used to analyze the relationship. It is not recommended however for use in general practice because you do not know how dense the markers are measured. QTL interval mapping would be preferred in such an analysis because the methods take account for missing genotype data that may not have been measured. When estimating a QTL position through maximum likelihood methods, like interval mapping, positions of other possible QTLs could affect the detection of the true position. Neighboring QTLs could possibly flatten the likelihood in instances where there are multiple QTLs on the same chromosome. This would make an effect look less significant at a given location than it actually is. Another possibility is that in the search over the interval you may find an area where the likelihood could reach a peak but could be a “ghost” QTL. This is where an effect is observed because a neighboring QTL is skewing the results at the particular position you are looking in and the result is a false discovery of the position. Marker Regression has been shown to improve interval mapping, which is call Composite Interval Mapping. This is where the QTL position found is also combined in a linear regression where the covariates are the other markers in the dataset. By including the markers as covariates the other position in the chromosome are accounted for in the analysis and false discovery is reduced.

The analysis of interval mapping and single marker analyses has shown to be effective but it limits our inference to one marker at a time as a possible loci that controls a trait. Using Marker Regression however you can incorporate multiple markers in a single analysis to test for possible QTL for a given trait. It is cautioned that running such an analysis is only an approximate test because the null hypothesis is there is no difference between the marker levels and therefore a non-mixture distribution but the alternative is a mixture of distributions. The assumption regression would make of the errors within the marker type to be normally distributed may not be entirely met if the QTL’s fall between the marker regions. However Whittaker et al. (1996) have shown that a direct regression of phenotypes on marker types, provides the same information about location of QTL-effects without having to step to all positions on the interval. With this information using the entire marker set in a regression analysis would provide a nice, computationally efficient way to map out the genetic architecture of a trait. This is stifled by financial cost of fully genotyping subjects to have an adequate sample size to properly create such a map.

* 1. Stepwise forward selection

Forward Selection Algorithm

* 1. Chapter Overview