# QTL Analysis Report for Baseline White Blood Cell Count Study

**The Jackson Laboratory**

## Document modification history

| Revision number | Date | Who | Comments |
|---|---|---|---|
| ***1.0*** | 5/6/2008 | TJL | First version |

## Table of contents

# Chapter 1. Experiment Design

The overall goal for this experiment is to detect and localize QTL that is responsible for variation of baseline white blood cell count between two mouse inbred strains. An intercross experiment (C57BLKS/J x SM/J ) was conducted. White blood cell counts were recorded for 186 F2 progeny that included 91 females and 95 males. Body weight was also recorded for each animal.

# Chapter 2. Statistical Methods

## 2.1 Genome-wide one-dimensional scan

Pseudo-markers were generated at 2-cM spacing for each chromosome, and a whole genome scan was performed using 128 imputations (Sen and Churchill, 2001). One thousand permutations were performed to determine the thresholds for QTL detection (Doerge and Churchill, 1996). Four thresholds 1%, 5%, 10% and 63% were calculated from the permutation results. QTL with LOD score above the 1% threshold were strong QTL, while those above 63% were suggestive QTL (Lander and Kruglyak, 1995). Both full and additive models were analyzed. Full model includes QTL, covariates (sex, body weight) and QTL*covariates interaction effects. Additive model only includes additive QTL and covariates effects.

## 2.2 Genome-wide two--dimensional scan

Pair-wise scans were performed using Pseudo-markers generated from one-dimensional scan process. All possible pairs of QTL locations on each chromosome were tested for association with the phenotype of interest. The likelihood from the full model (pseudo-marker pair and the interaction between them) and the null model (no genetic effect) was compared and LOD scores were calculated. In addition, LOD scores from comparing the likelihood from the full model and the additive model (with only the main effects of pseudo-markers and but no interaction) were also calculated.

## 2.3 Multiple regression

QTL and possible QTL*QTL interactions identified from a one and two-dimensional scans were fit into multiple regression models. By doing so, variations of the phenotype in the models were estimated. Probabilities (p-values) for the significance of terms in the multiple regression models were calculated. Terms were dropped sequentially until all of the terms in the model were significant at 1% level for main effects and 0.1% for the QTL*QTL interaction effects.

## 2.4 Statistical software

R/QTL version 1.08-56 was used (Broman et al, 2003).

# Chapter 3. Distributions of Data and Quality Control Diagnostics

## 3.1 Phenotype distribution

As shown in Figure 1, histogram plots were generated for WBC. The original data were left-skewed, and logarithm transformation of raw data was used to reduce the skewness.

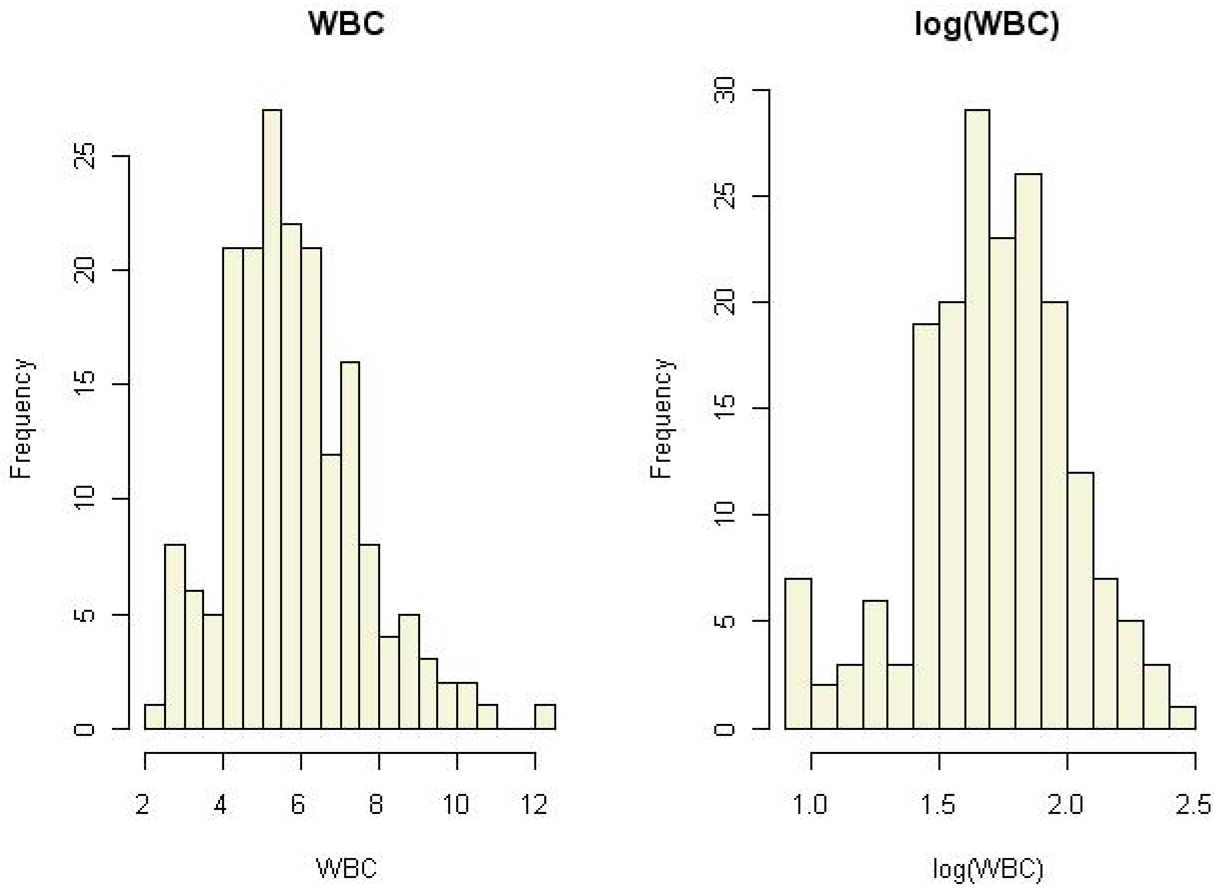


**Figure 1. Histogram plots of WBC and logarithm transformed WBC.**

The Jackson Laboratory

JAX® Services

## 3.2 Genotyping quality assessment

### 3.2.1 Recombination fraction

Recombination fraction (RF) plots are shown in Figure 2. Markers that are physically closer to each other are strongly linked (red color); Markers that are far away from each other on the same chromosome or markers on the different chromosomes are more independent from each other (blue). The left panel in Figure 2 suggests that the last marker on chromosome 16 may have a quality issue. RF plot (right panel) appears in good quality after excluding the last marker on chromosome 16.

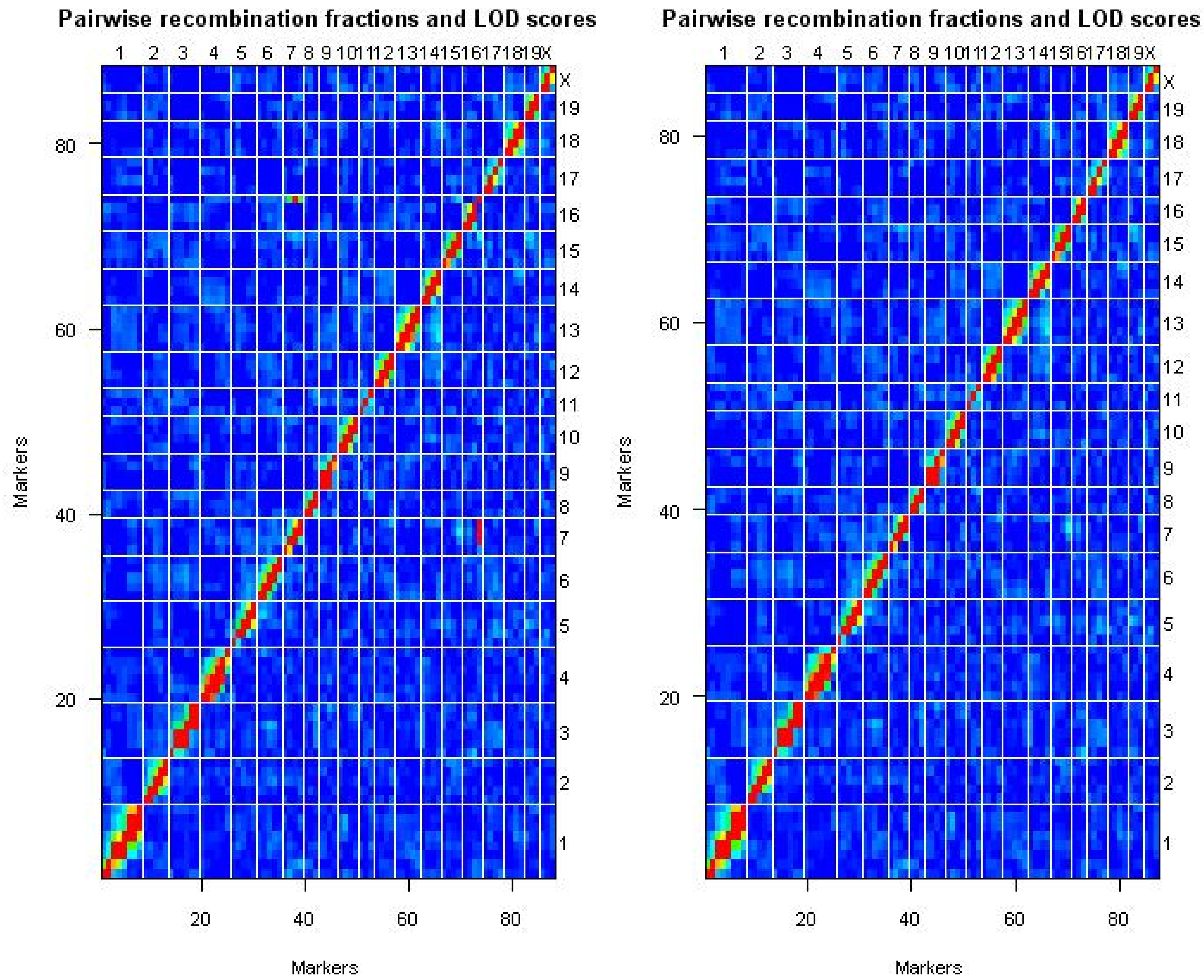


**Figure 2. Recombination plots.**

### 3.2.2 Genetic map estimates

Genetic map was re-estimated using the data from this experiment and presented in Figure 3. Figure 3 shows the length of chromosome 16 was too long. Removing the last marker on chromosome 16 fixed the problem.

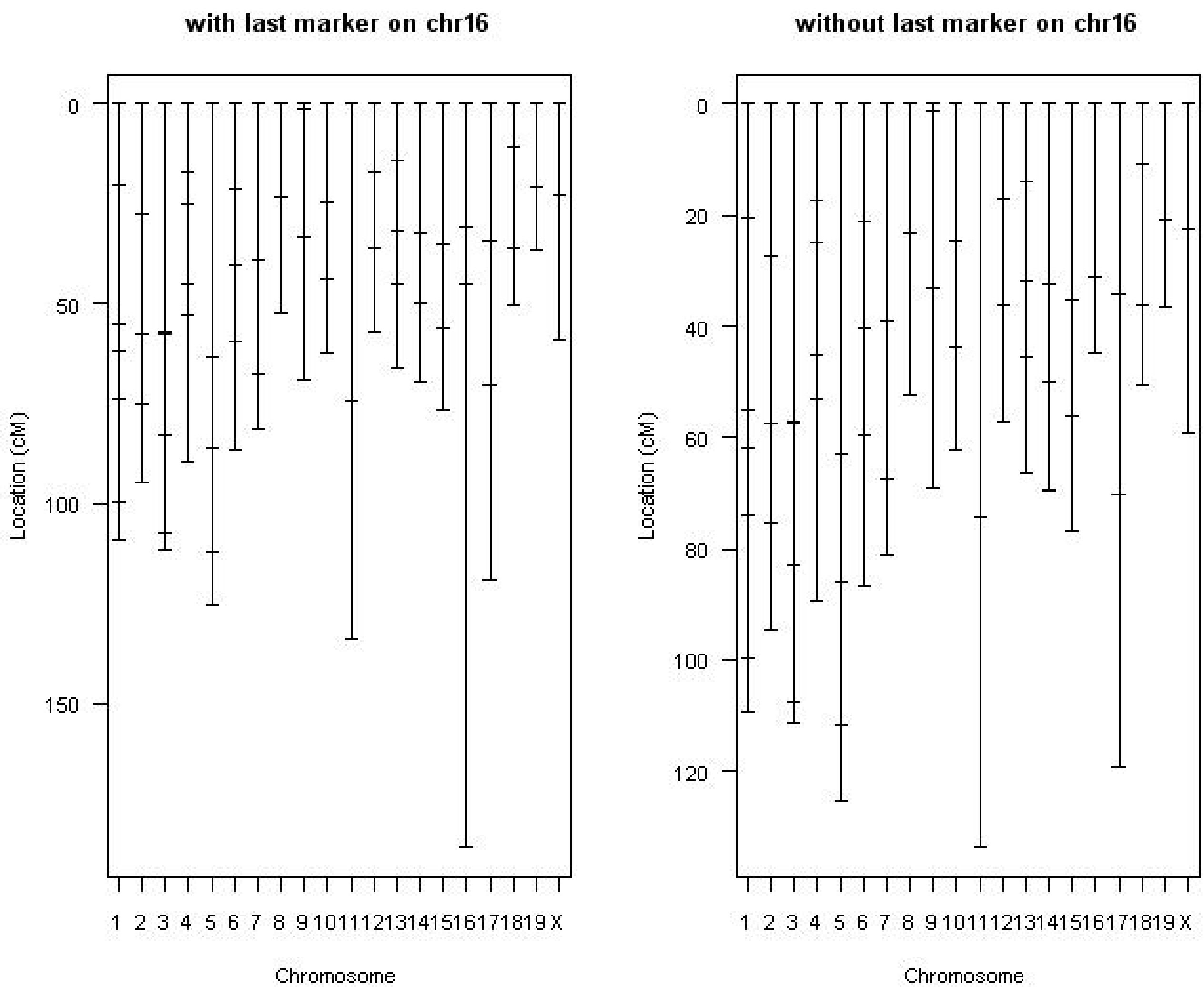


**Figure 3. Re-estimated genetic maps.**

# Chapter 4. Results

## 4.1 Genome-wide one-dimensional scan

Figure 4 presents the results from a genome-wise one-dimensional QTL scan. The lines bottom-up represent significant thresholds of 63%, 10%, 5% and 1%, respectively. The top panel presents the results from a full model analysis; the middle panel shows the results from an additive model analysis; the bottom panel is the difference between the full and additive model, which can be used to detect sex-specific QTL. LOD threshold of 2.0 was used for the sex-specific QTL detection. Candidate QTL on chr1, chr3, chr5 and chr15 were selected for further analysis.

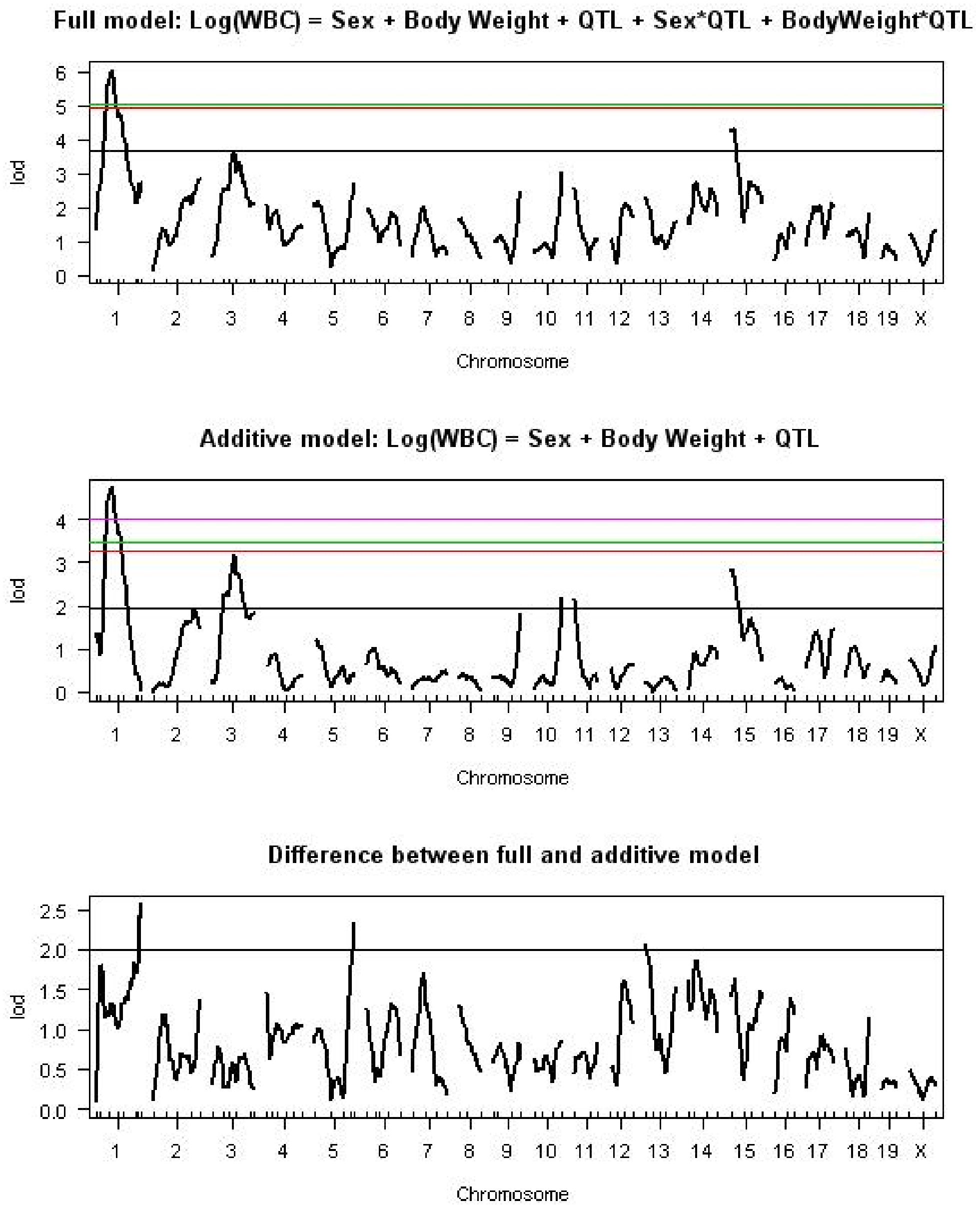


**Figure 4. One-dimensional scan plots.**

### 4.2 Genome-wide two-dimensional scan

Genome-wide two-dimensional QTL scan was performed, and the results are show in Figure 5. The upper triagonal matrix in Figure 5 represents the LOD scores contributed by the interactive QTL only. The lower triagonal matrix represents the LOD scores from a full model considering two QTL and their interaction together. The more reddish, the greater the LOD score is. As shown in Figure 5, no clear evidence of interactive QTL was found.

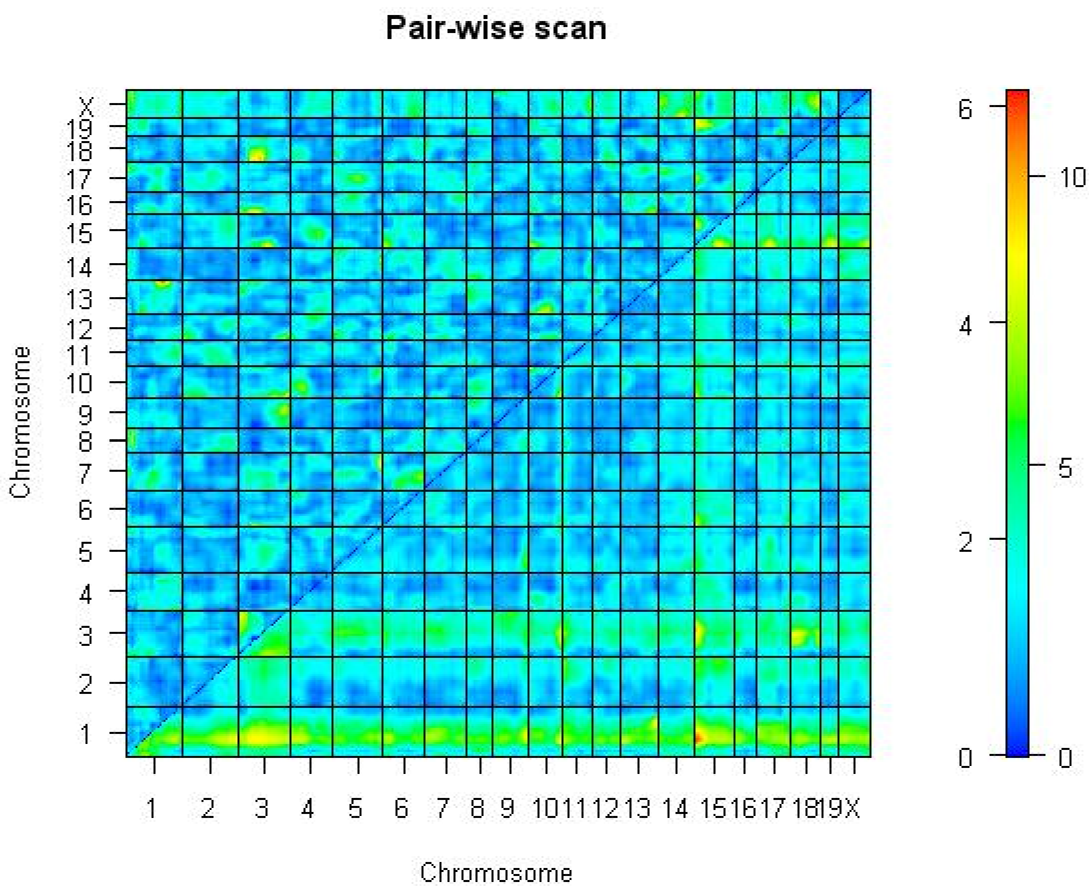


**Figure 5. Two-dimensional QTL scan plot.**

### 4.3 Multiple regression analysis

Candidate QTL from genome-wide one and two-dimensional QTL scans were fit into multiple regression models. Covariates such as sex and body weight were also added as an “adjustment” factor in the multiple regression models. As shown in Table 1, the phenotypic variance explained by each QTL was 10.75%, 5.28% and 7.31% for chr1@44cM, chr3@44.2cM and chr15@0cM,respectively.

| | df | Type III SS | LOD | %var | F value | Pvalue(Chi2) | Pvalue(F) |
|---|---|---|---|---|---|---|---|
| Chr1@44cM | 2 | 1.8371 | 5.463 | 10.7474 | 12.8178 | 0 | 0 |
| Chr3@44.2cM | 2 | 0.9034 | 2.7789 | 5.2852 | 6.3034 | 0.0017 | 0.0023 |
| Chr15@0cM | 2 | 1.2502 | 3.7966 | 7.3137 | 8.7226 | 2e-04 | 2e-04 |
| sex | 1 | 0.0056 | 0.018 | 0.033 | 0.0787 | 0.7737 | 0.7794 |
| Body weight | 1 | 0.1251 | 0.3965 | 0.7321 | 1.7462 | 0.1766 | 0.1881 |

**Table 1. Multiple regression results.**

### 4.4 QTL effect

Figure 6 below shows the 3 QTL effects indentified from multiple regression analysis. Animals with alleles from SM/J for QTL chr3@44.2cM and chr15@0cM are expected to have greater WBC; Animals with alleles from SM/J for QTL chr1@44cM, however, are expected to have smaller number of WBC. The QTL effects are assumed to be additive.

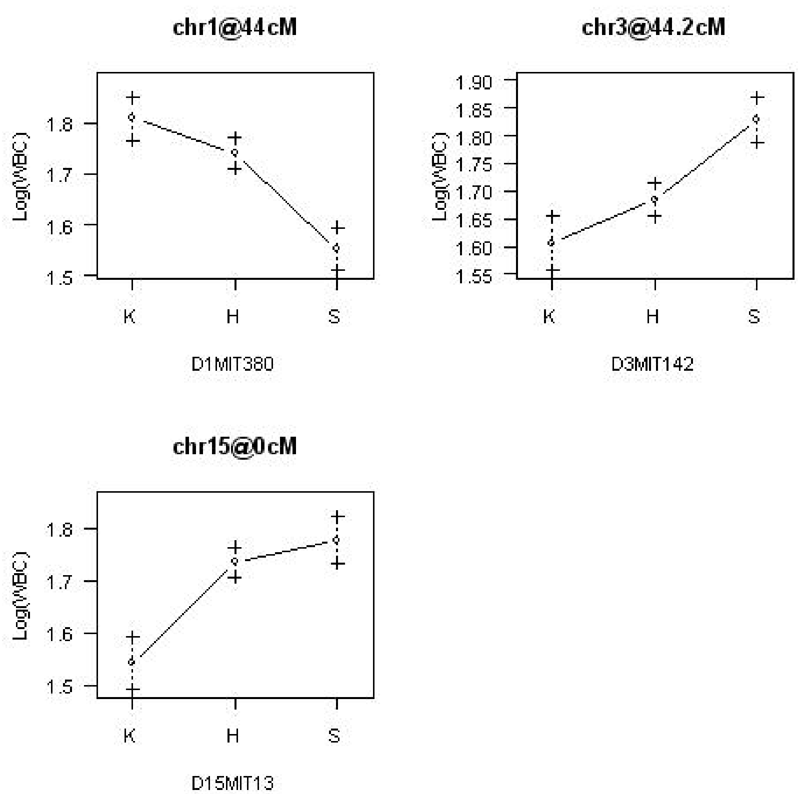


**Figure 6. Effect plots.**

### 4.5 QTL confidence intervals

QTL confidence interval plots are presented in Figure 7. The horizontal bar in each plot represents confidence interval for each QTL detected.

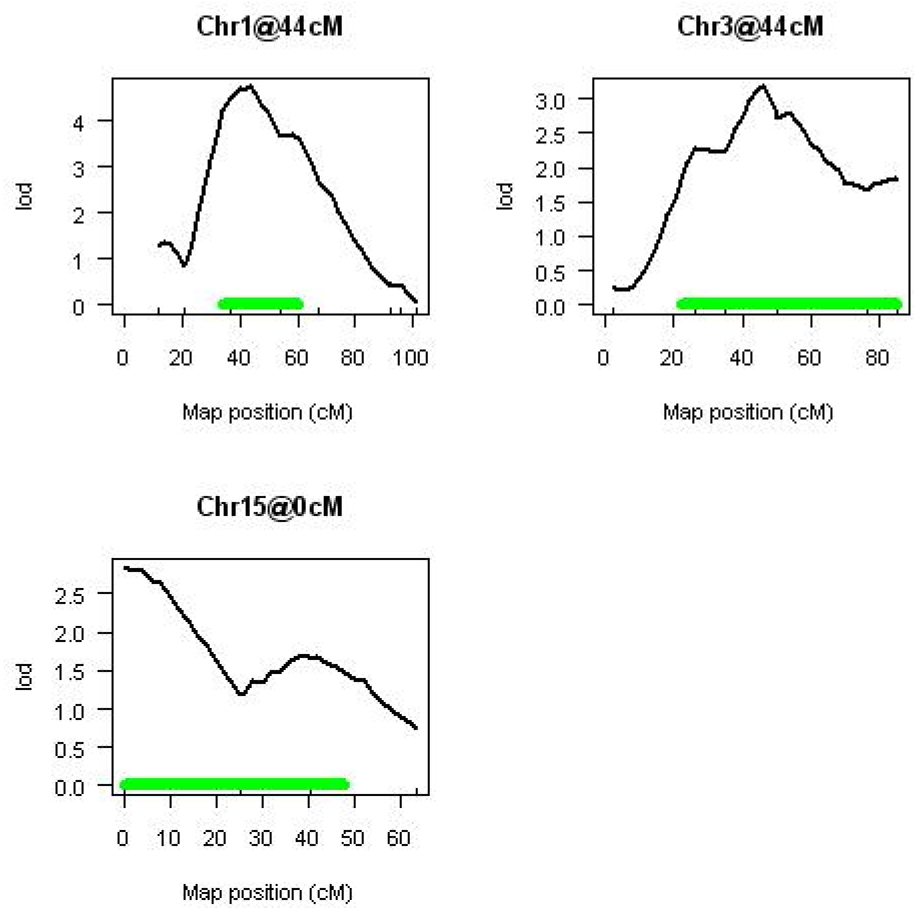


**Figure 7. Confidence interval plots.**

## Chapter 5. Conclusions

Three QTL were detected for baseline white blood cell count. They are chr1@44cM, chr3@44.2cM and chr15@0cM. The LOD scores for the 3 QTL are 5.46, 2.78 and 3.79, respectively. Alleles from SM/J for QTL on chr3 and chr15 are expected to raise WBC, while the effect was opposite for QTL on chr1.

## Chapter 6. References

Broman, K. W., Wu H., Sen S., Churchill, G. A. 2003. R/qtl: QTL mapping in experimental crosses. *Bioinformatics* 19:889-890.
Doerge, R. W. and Churchill, G. A.. 1996. Permutation tests for multiple loci affecting a quantitative character. *Genetics* 142, 285-294.
Lander, E. and Kruglyak, L.. 1995. Genetic dissection of complex traits: guidelines for interpreting and reporting linkage results. *Nature Genetics*.11: 241-247.
Sen, S. and Churchill, G. A.. 2001. A statistical framework for quantitative trait mapping. *Genetics* 159, 371-387. 12