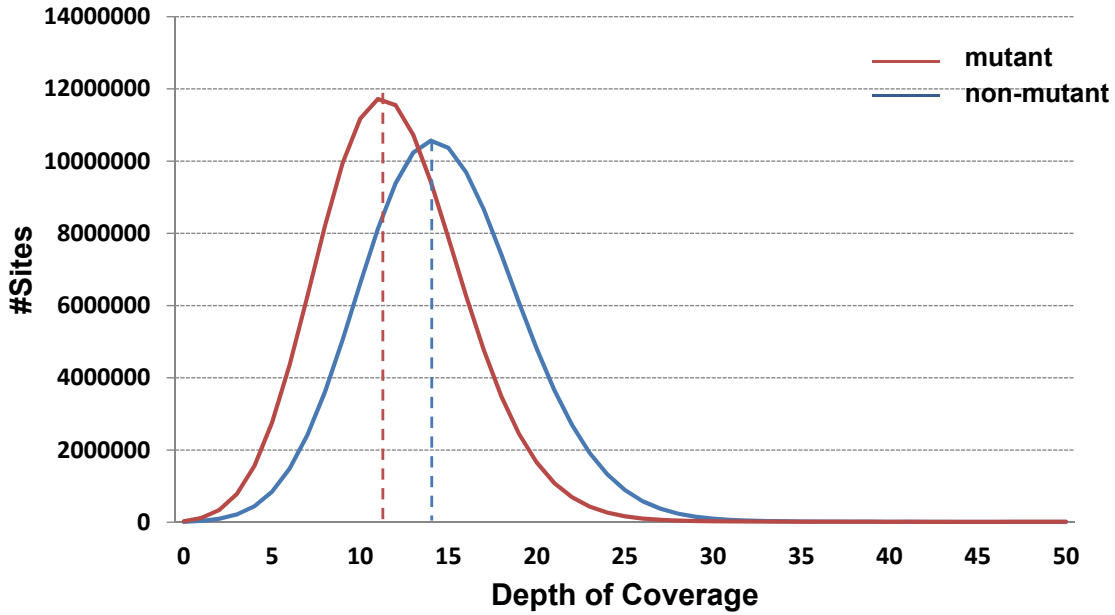

Figure S1 Distribution of the depth of coverage from single-end 51 bp sequencing on Illumina Hi-Seq. The mutant pool has on an average of 11 x coverage and the non mutant pool has on an average of 14 x coverage.

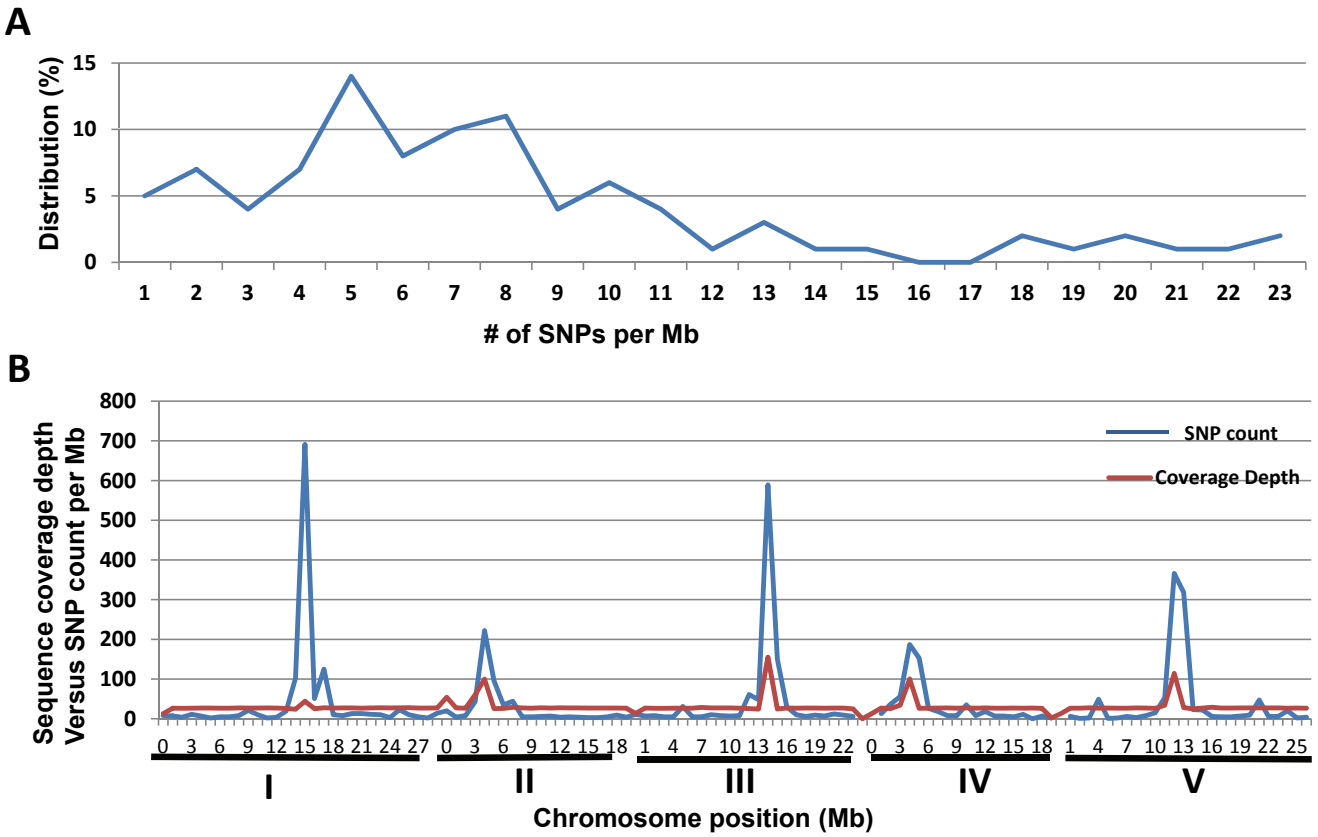

Figure S2 SNP distribution along the chromosomes. (A) Distribution of SNP count per Mb. Shown are the number of SNPs identified by SAMtool per Mb (X Axis) and the percentage of chromosome regions with that number of SNP count (Y Axis). The average SNP count is 8 per Mb across all regions. (B) The red line shows sequence coverage depth (sequence reads per site) and blue line shows the SNP count (number per 1 Mb sliding window). Several regions of chromosomes appear to have a high SNP count and most of them correlate with a high coverage depth, indicating a mis-calling of SNPs at the repetitive regions.

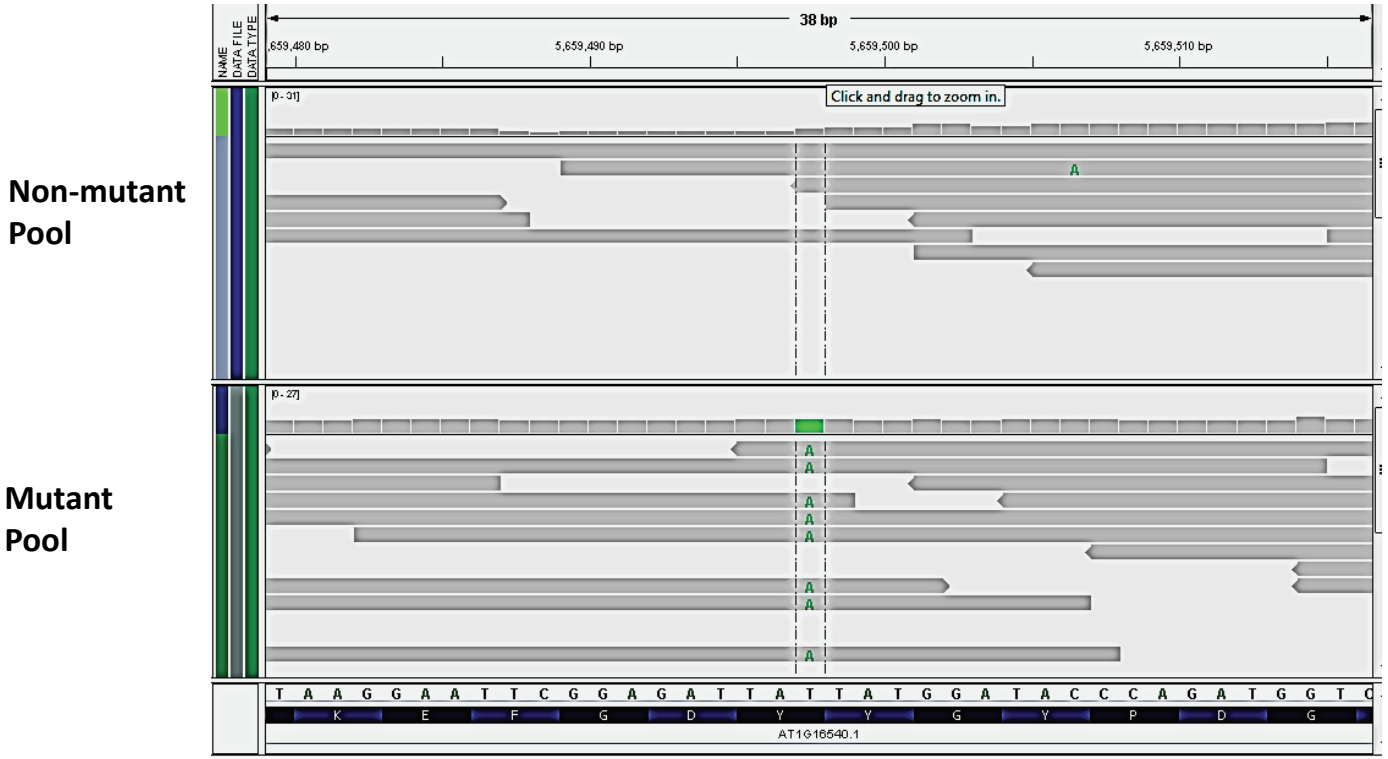

Figure S3 Sequence view at the region with the top candidate SNP. Sequence reads from the mutant pool and the non-mutant pool were viewed along the reference sequences with IGV software (http://www.broadinstitute.org/igv/). Non-reference alleles are indicated by green. All eight reads in the mutant pool has a non-reference base 'A' while all four reads in the non-mutant pool has the reference base of 'T'.