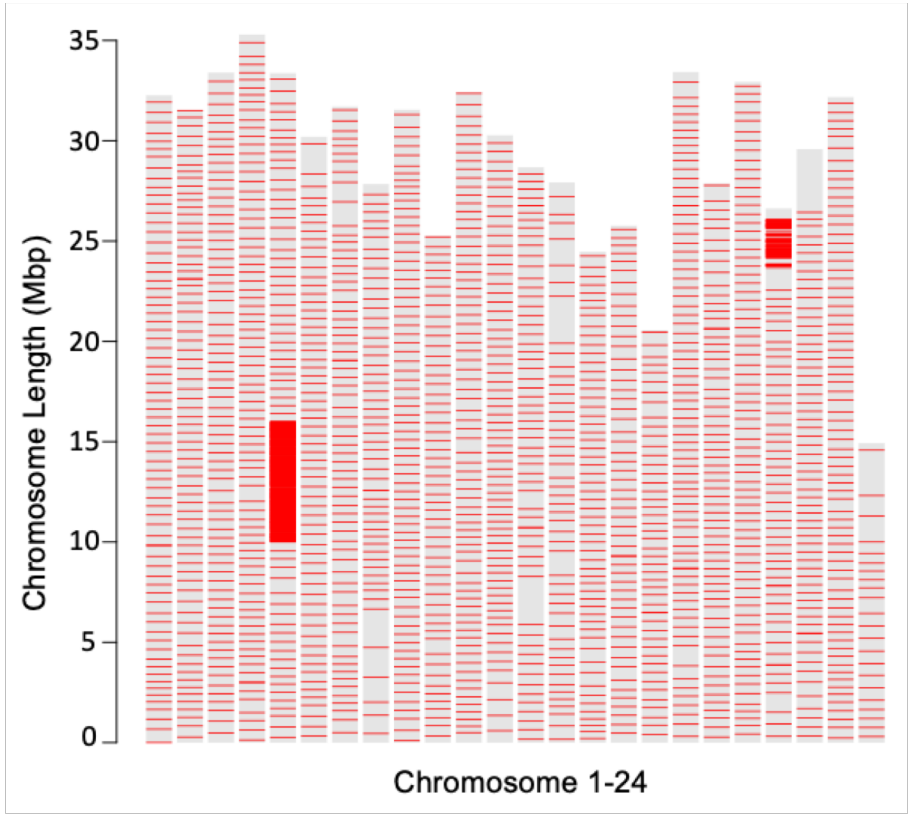

Figure S1. Genomic locations of amplicons

Chromosome plot showing the location of amplicons. Capture probes were designed to sequence the BC progeny genome with different coverage. A total of 2,017 probe sets were produced for amplicon sequencing: 406 sets of probe were designed to reach a resolution of 14.8 kbp within 6 Mbp candidate *R(Diff)* loci on chromosome 5 (chr 5: 10,000,000 – 16,000,000); 101 sets or capture probes were designed for 24.8 kbp definition for sex determining regions (chr 21: 23,750,000 – 26,250,000) that is close to *xmrk*; 1510 sets of probes were designed for genotyping and establish individual backcross progeny haploid map at definition of 459 kbp for the rest of the genome.

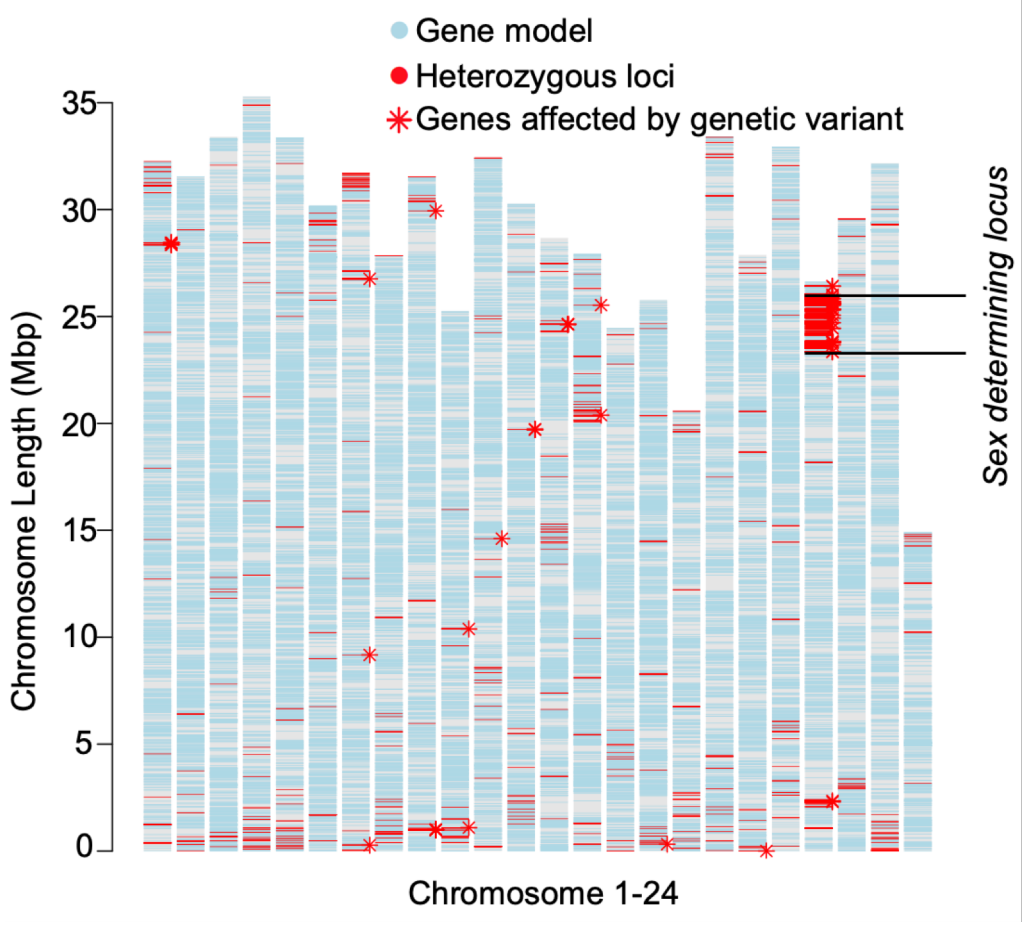

Figure S2. Heterozygous loci in male *X. maculatus* genome

Chromosome plot of heterozygous loci of *X. maculatus* genome. Heterozygous loci were identified using bcftool and varscan. Loci that are heterozygous in at least 75% of tested male *X. maculatus* are highlighted, with the one leading to non-synonymous codon change labeled with asterisks. *X. maculatus* exhibits a XY sex determining locus. This locus is known to be at the end of chromosome 21, close to *xmrk*.

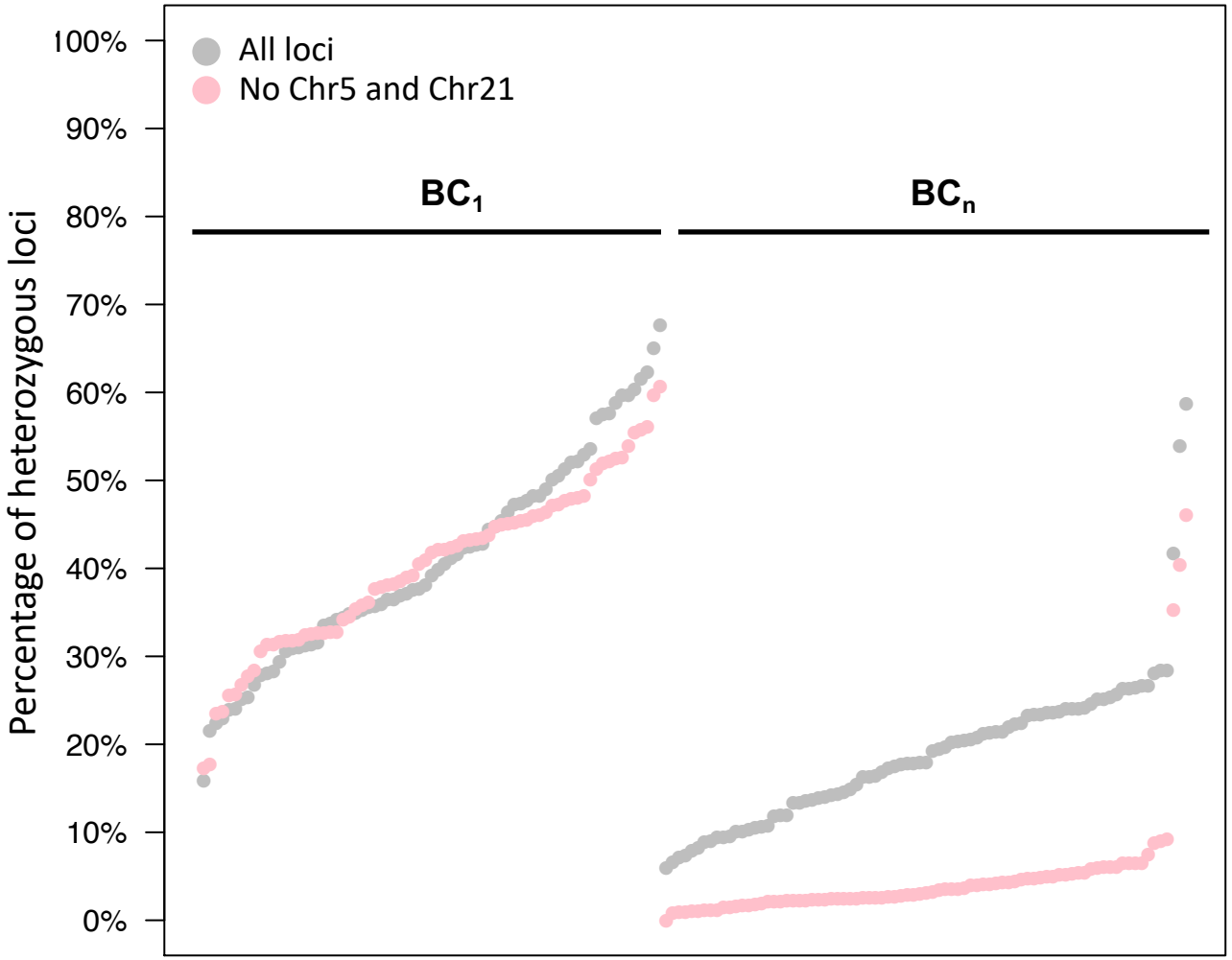

Figure S3. Heterozygous loci in backcross interspecies hybrids

Dot plot showing percentage of heterozygous loci in BC_1 and BC_n (i.e., BC_2-BC_8) hybrids. Heterozygous genotype refers to inheritance of both *X. maculatus* and *X. hellerii* alleles, and homozygous genotype refer to inheritance of only *X. hellerii* alleles. Each spot represents percentage of heterozygosity per BC individual. All percentages of BC_1 and BC_n are ordered numerically. Gray spots represent percentages calculated using all genotype loci, and pink spots represent percentages calculated using only loci out of Chr5 and Chr21.

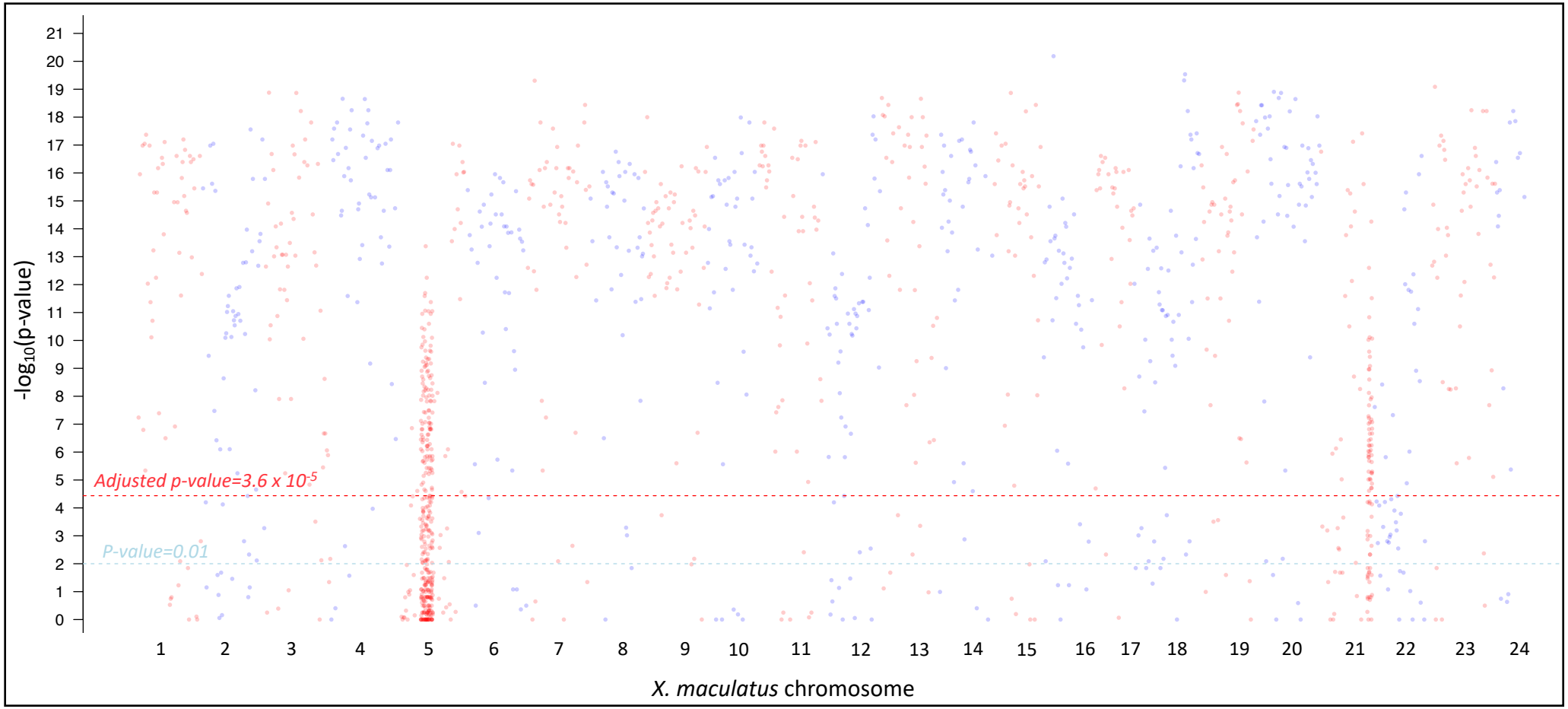

Figure S4. Linkage disequilibrium in advanced backcross hybrids

Manhattan plot showing -$\log_{10}$p-value (Chisq-test) across the genome. Y-axis represents -$\log_{10}$p-value, X-axis represents amplicon chromosomal coordinates, which is labeled as red or blue. Manhattan plot shows statistically significant linkage disequilibrium throughout the genome due to introgression. Light blue dash line represent p-value of 0.01 that is suggestive of statistical significance. Chisq-test p-values were corrected using Bonferroni method across the genome-wide data. Red dashed line represents adjusted p-value of 0.05 that is corresponding to 3.6 x 10^{-5}.

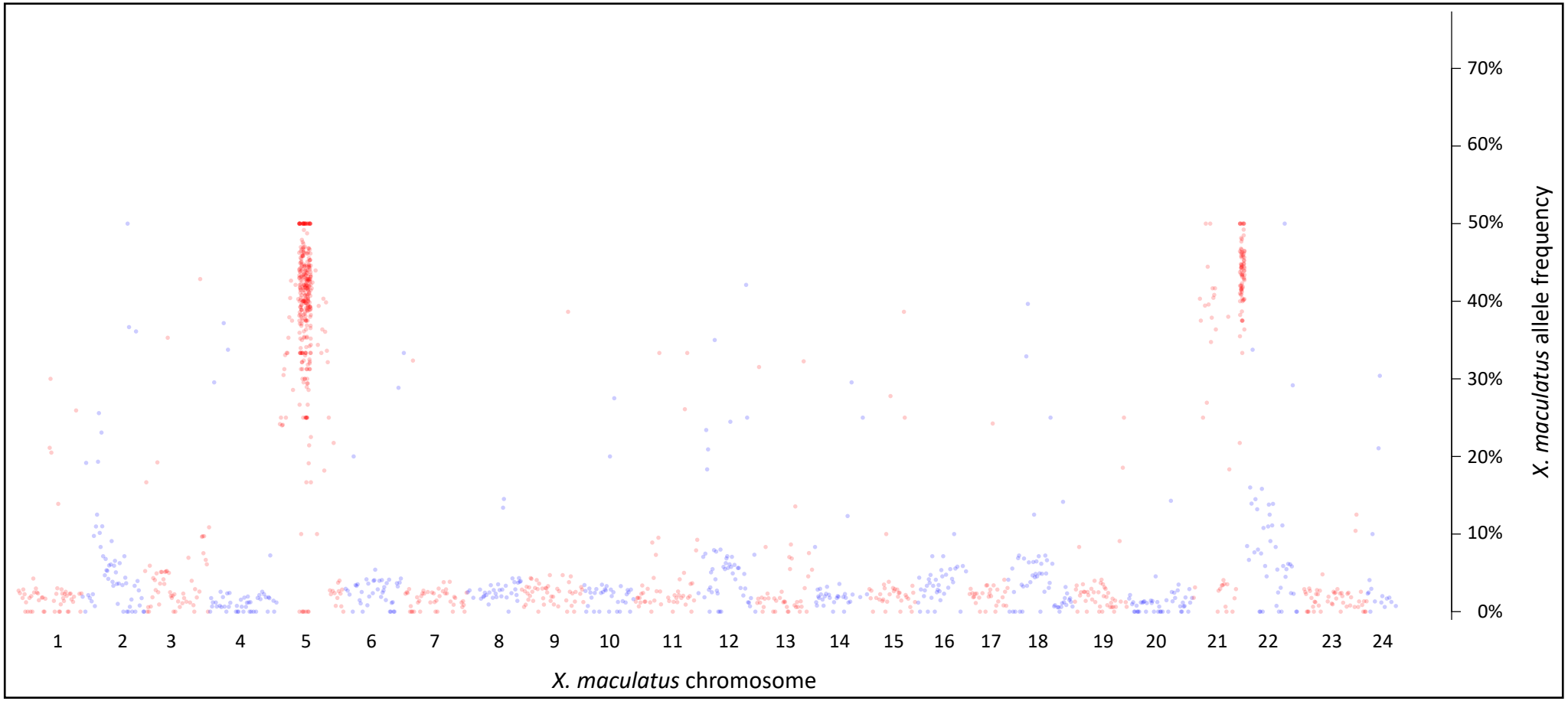

Figure S5. *X. maculatus* allele frequency in advanced backcross hybrids

X. maculatus allele frequencies of all sequenced polymorphic sites in advanced backcross hybrids are shown as dot plot. Red and blue colors represent different chromosomes (red: odd number; blue: even number). The maximum *X. maculatus* allele frequency in backcross progeny is 50% as observed in heterozygous loci, and minimum *X. maculatus* allele frequency is 0% in homozygous loci.

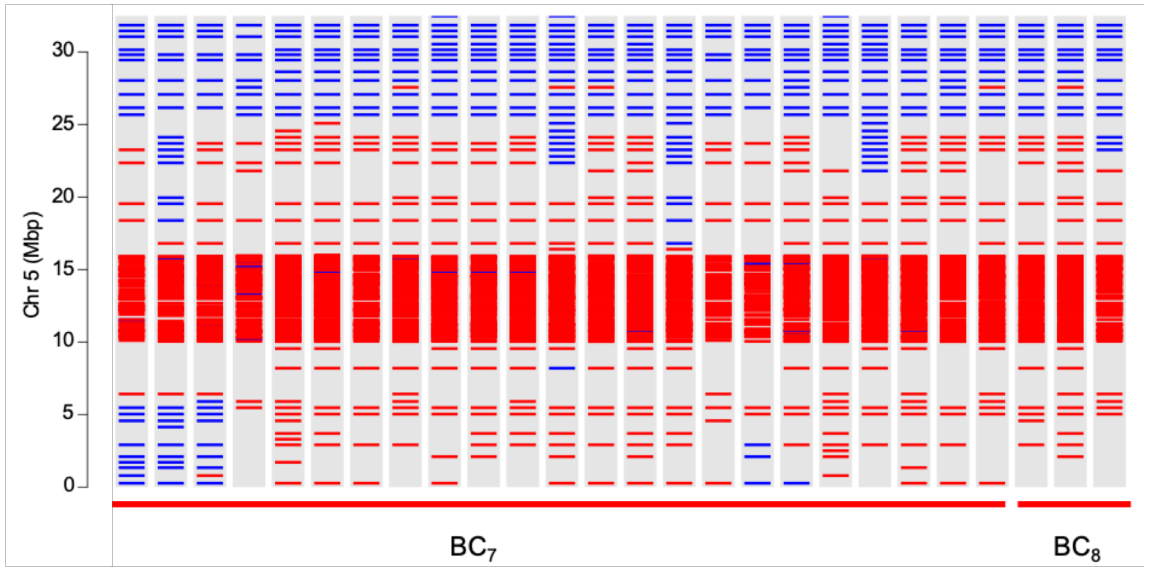

Figure S6. Haploid maps of $BC_{7\text{-}8}$ chromosome 5

Chr5 haploid maps for BC_7 and BC_8 hybrids. Blue bar represents homozygous genotype, and red bar represent heterozygous genotype. Locations of the colored bars represent chromosomal location of amplicons. Heterozygous regions, by calculation, should only account for a total of 5.5 or 2.7 Mbp of a 700 Mbp hybrid genome. These heterozygous loci should include both the locus surrounding *xmrk* and *R(Diff)*. The haploid map of $BC_{7\text{-}8}$ hybrid chromosome 5 show an average of 22 Mbp heterozygous region for BC_7, and 23 Mbp heterozygous region for BC_8. Chromosome 5 heterozygous loci of BC_7 and BC_8 are statistically different from the expected 5.5 and 2.7 Mbp heterozygous region lengths (p-values of 1.18×10^{-16} and 2.4×10^{-3}; t-test, two-tailed).