A simple test for genome-wide evidence of selection on complex traits

Timothy Beissinger [1,2], Aaron Lorenz [3], Jochen Kruppa [4], and Henner Simianer [5]

[1]USDA Agricultural Research Service, Columbia, Missouri, USA
[2]Division of Plant Sciences, University of Missouri, Columbia, Missouri, USA
[3]Department of Agronomy, University of Minnesota, Minneapolis, Minnesota, USA
[4]Inst. for Animal Breeding and Genetics, University of Veterinary Medicine, Hannover, Germany
[5]Department of Animal Breeding and Genetics, Georg-August University, Göttingen, Germany

Introduction

There is increasing evidence that quantitative selection, or selection operating concurrently on many loci, each of small effect, is an important driver of evolution. We developed a simple test, applicable to multi-generation experimental or breeding populations, to identify selection operating on polygenic traits. The test involves the relationship between allele frequency changes over time and estimated allelic effects at every position in the genome. We construct a statistic, Δ_α, which is the sum of the estimated allelic effect multiplied by the change in allele frequency over time. Assuming a multi-generational population has been genotyped at i SNPs, which have effects estimates of α_i and show frequency changes over time of Δ_i, we define Δ_α according to:

$$\Delta_\alpha = \sum_{i=1}^{n} \Delta_i \alpha_i \quad (1)$$

The simple nature of Δ_α allows straight-forward permutation testing to assess significance. This statitic can be used to either 1) confirm that prior phenotypic selection was effective at shaping the genome, or 2) identify traits that may have been under quantitative selection in the past.

Identifying selection by ignoring the underlying genes

▶ Mapping studies rarely have the power to identify QTL of small effect.

▶ Identifying QTL is not necessary to draw genome-wide conclusions.

▶ Trends may emerge from many insignificant loci that are more informative than certainty at only a few loci.
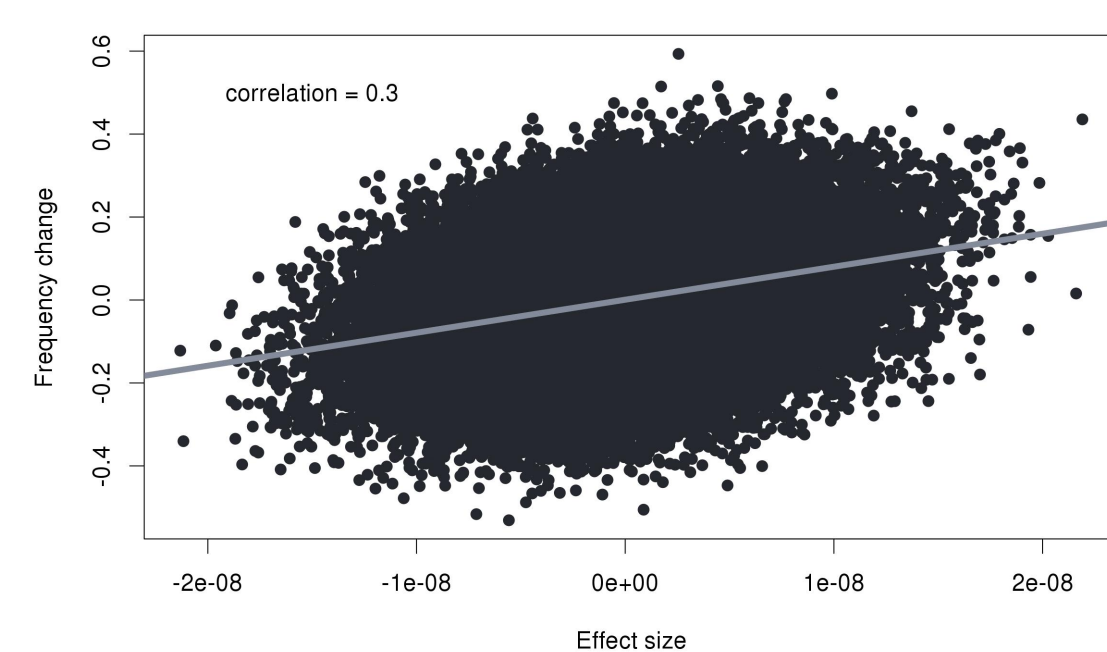

Figure 1: A positive correlation between effect size and allele frequency change is observed in a simulated 20-generation selection experiment with 200 QTL.
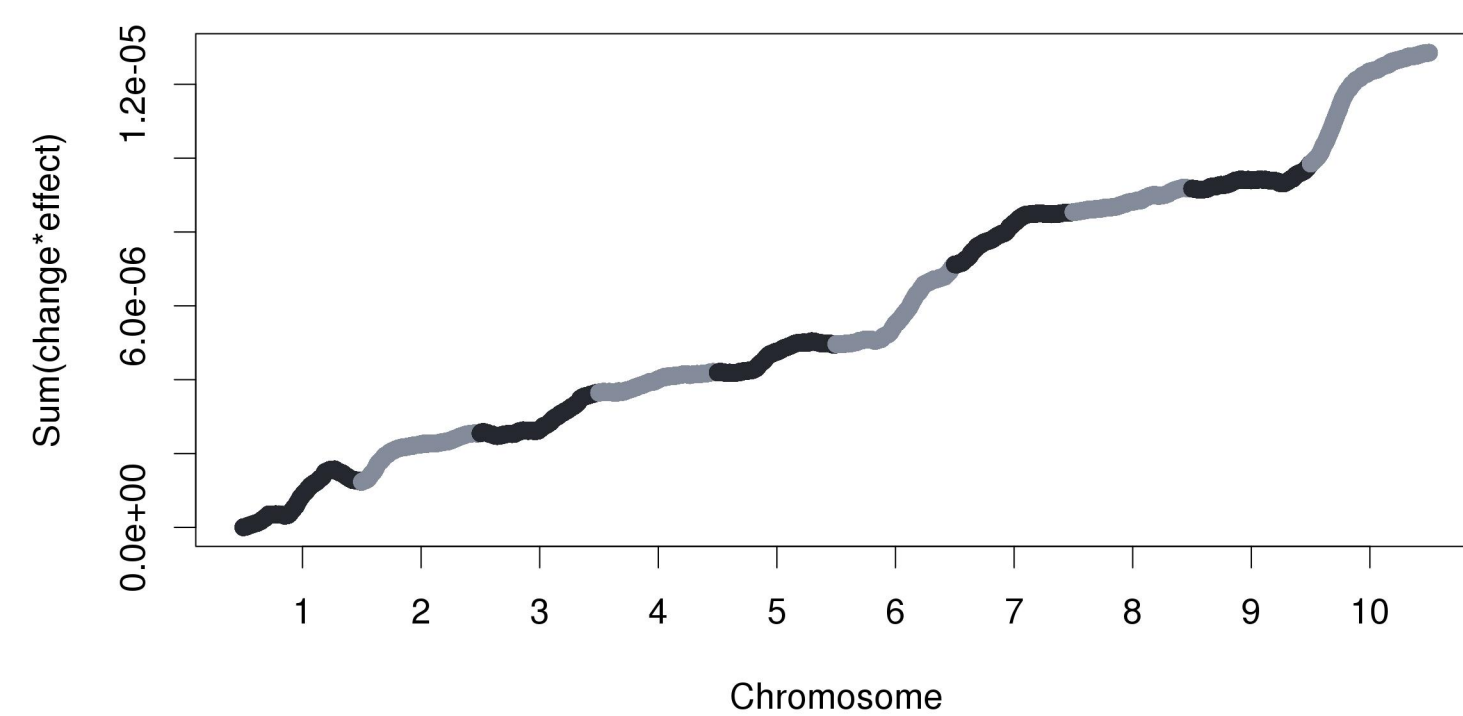

Figure 2: The cumulative sum of effect size multiplied by allele frequency change in a simulated 20-generation selection experiment with 200 QTL

Simulations demonstrate that permutation testing is robust

Figure 3: A highly significant correlation between effect size and frequency change is observed in a simulated selection experiment with 100 QTL.
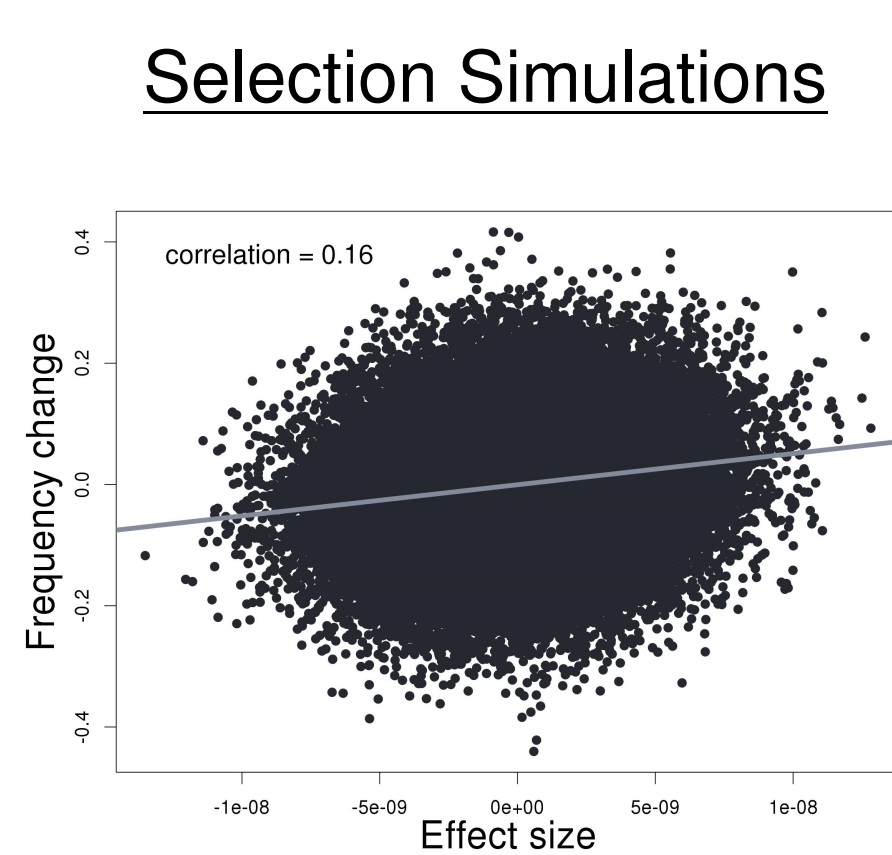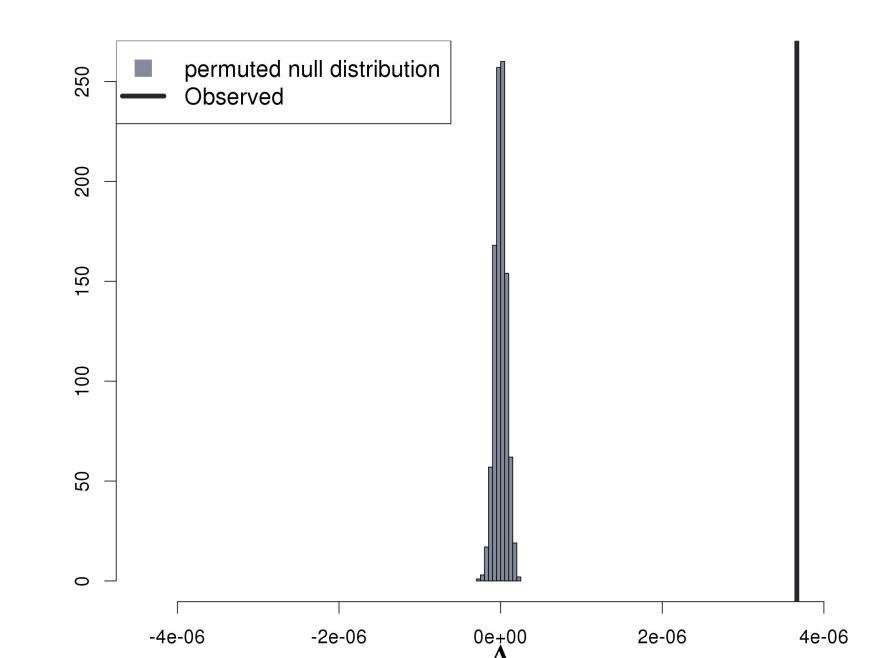
Figure 4: Permutation test of Δ_α in a simulated 20-generation selection experiment with 100 QTL.
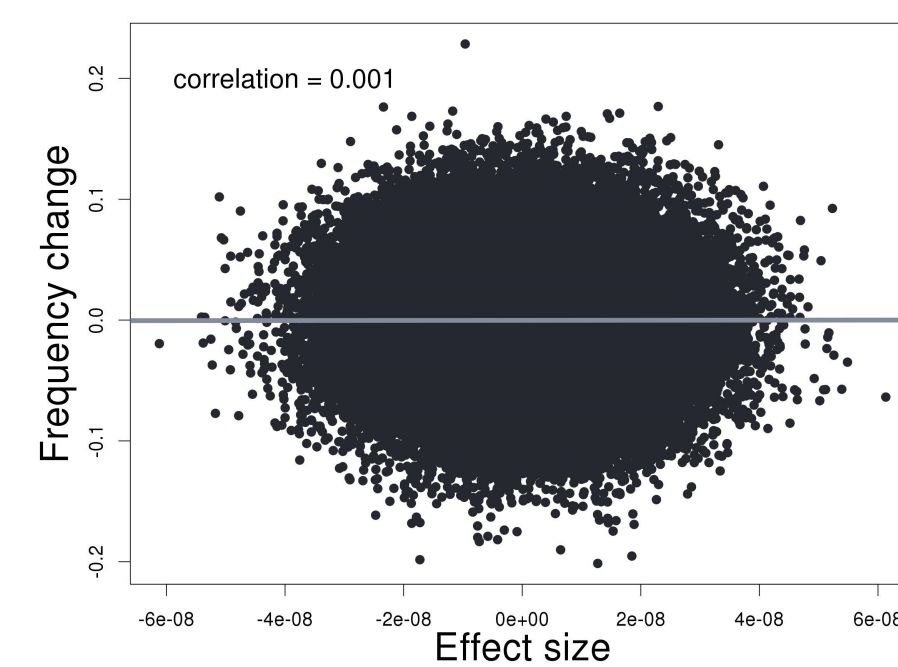

Figure 5: No significant correlation between effect size and frequency change is observed in a simulated drift experiment with 100 QTL.
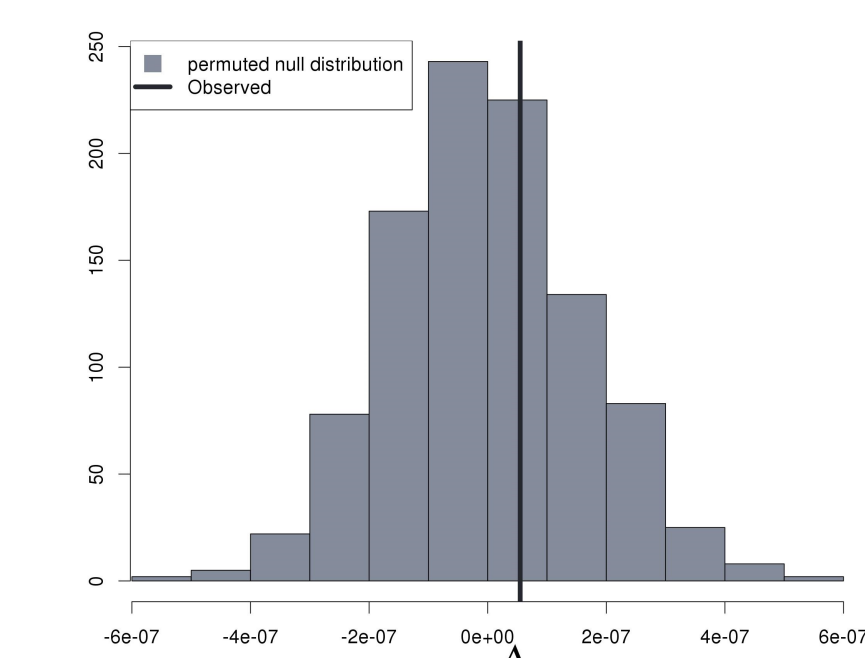
Figure 6: Permutation test of Δ_α in a simulated 20-generation drift experiment with 100 QTL.

Evidence of selection on complex traits in maize

Figure 7: GWAS for several silage-related traits in the Wisconsin Quality Synthetic (WQS) population identifies no significant associations. Adapted from [1].
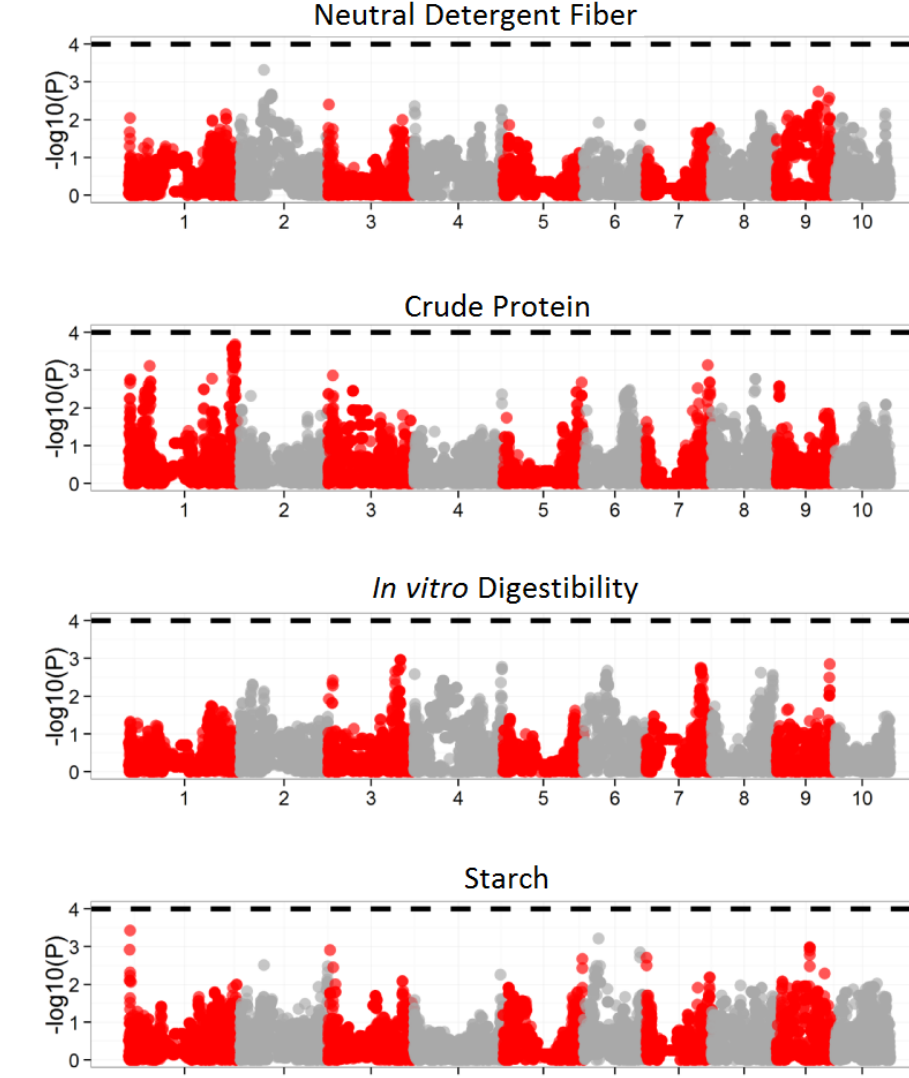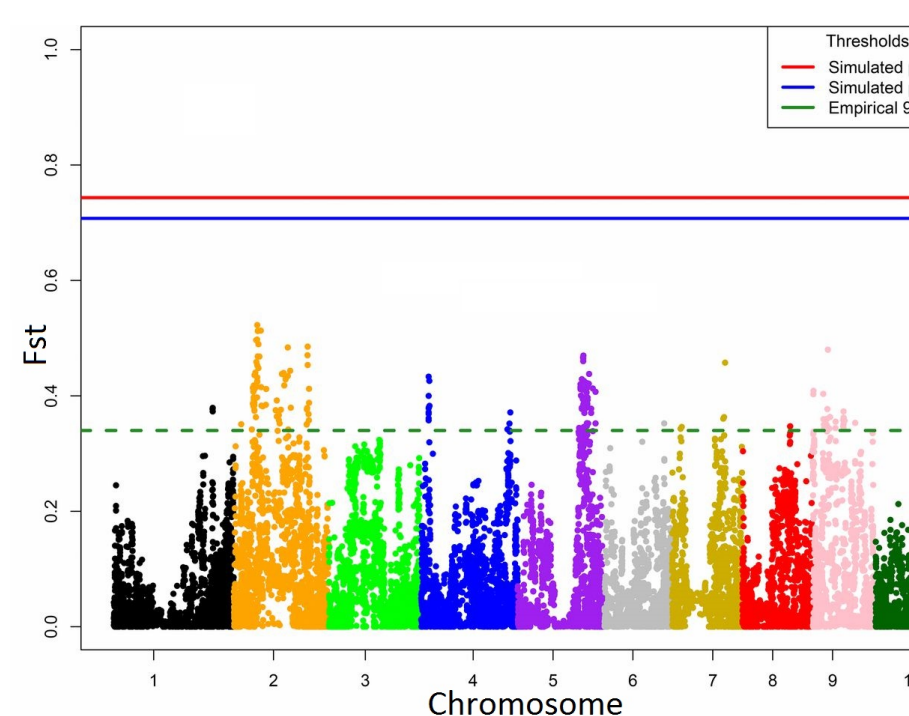
Figure 8: Selection Mapping in the WQS population identifies no significant associations. Adapted from [1].
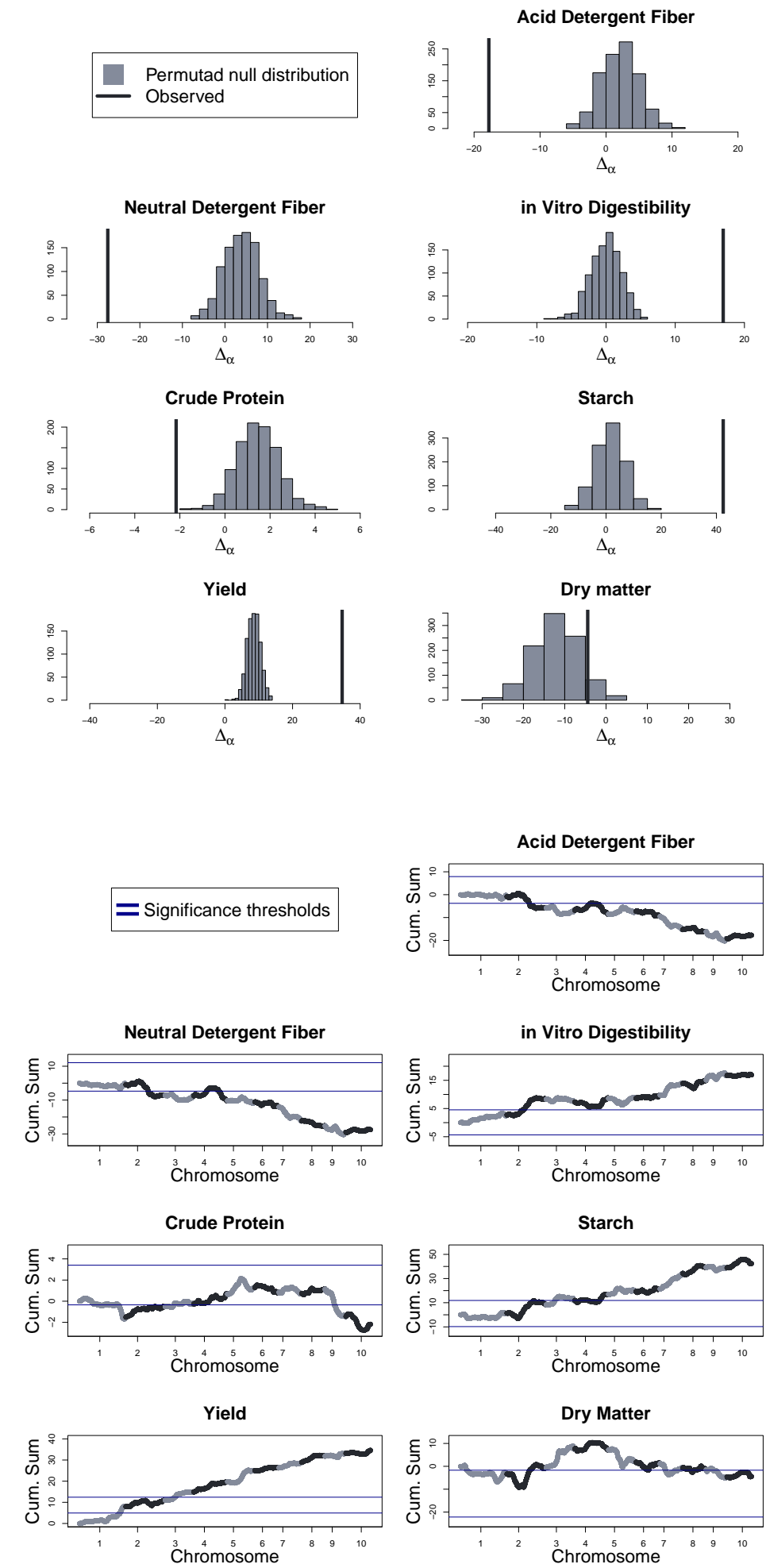
Figure 9: Permutation tests for genome-wide quantitative selection among 7 silage-related traits in the WQS population. Effects were estimated via GWAS.

Figure 10: Cumulative summation of Δ_α among 7 silage-related traits in the WQS population. Effects were estimated via GWAS.

Evidence of selection on complex traits in chickens

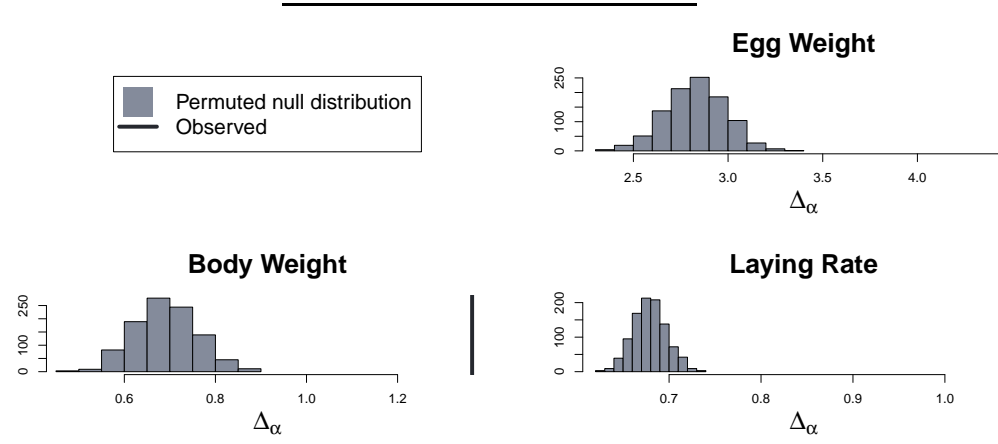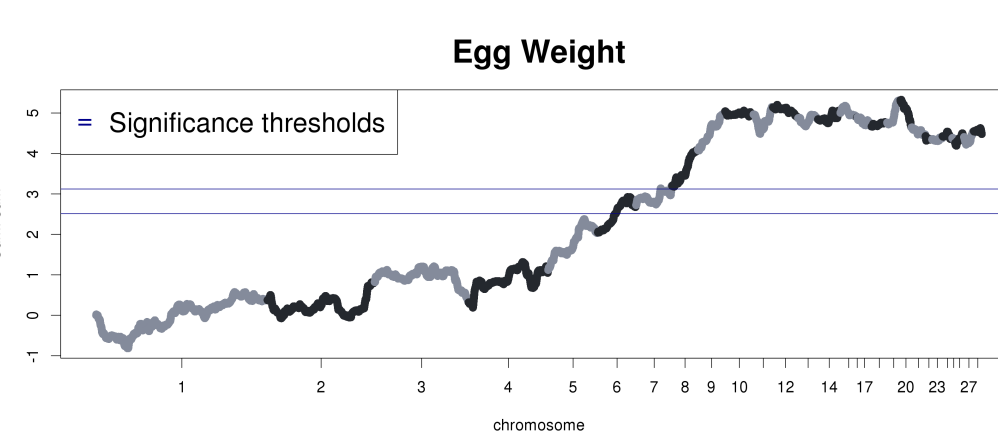Figure 11: Permutation tests for genome-wide quantitative selection among four traits in white-laying chickens.

Figure 12: Cumulative sum of Δ_α for egg weight in white layers.

Figure 13: Permutation tests for genome-wide quantitative selection among 4 traits in brown-laying chickens.
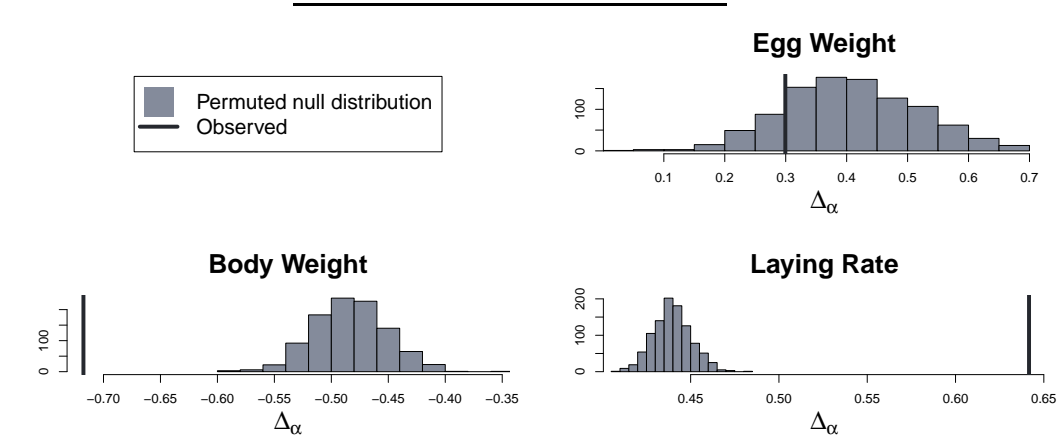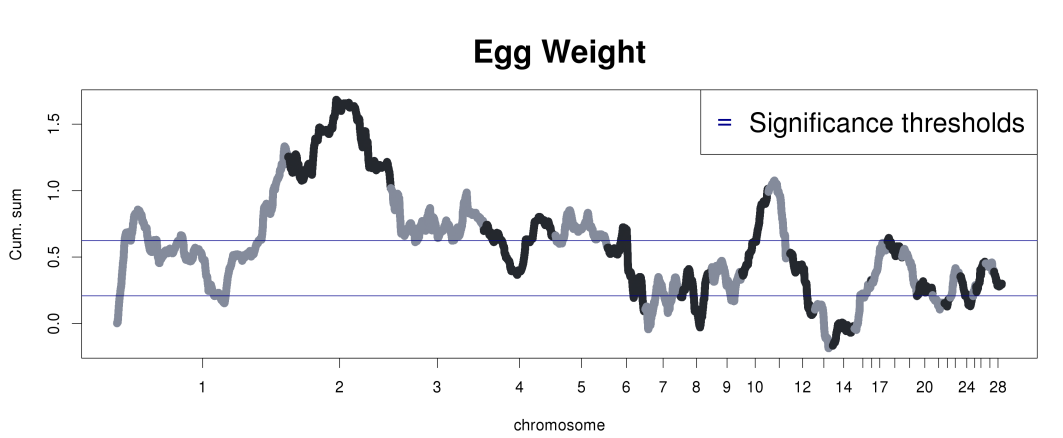
Figure 14: Cumulative sum of $\Delta_i \times \alpha_i$ for egg weight in white layers.

Using $\Delta_i \times \alpha_i$ for mapping

Effect sizes and allele frequency changes can both be informative for mapping, so there exists the possibility that combining them may be more powerful for mapping genes than either is alone. We simulated a population with 10 chromosomes and 100 evenly-distributed QTL to test this hypothesis.
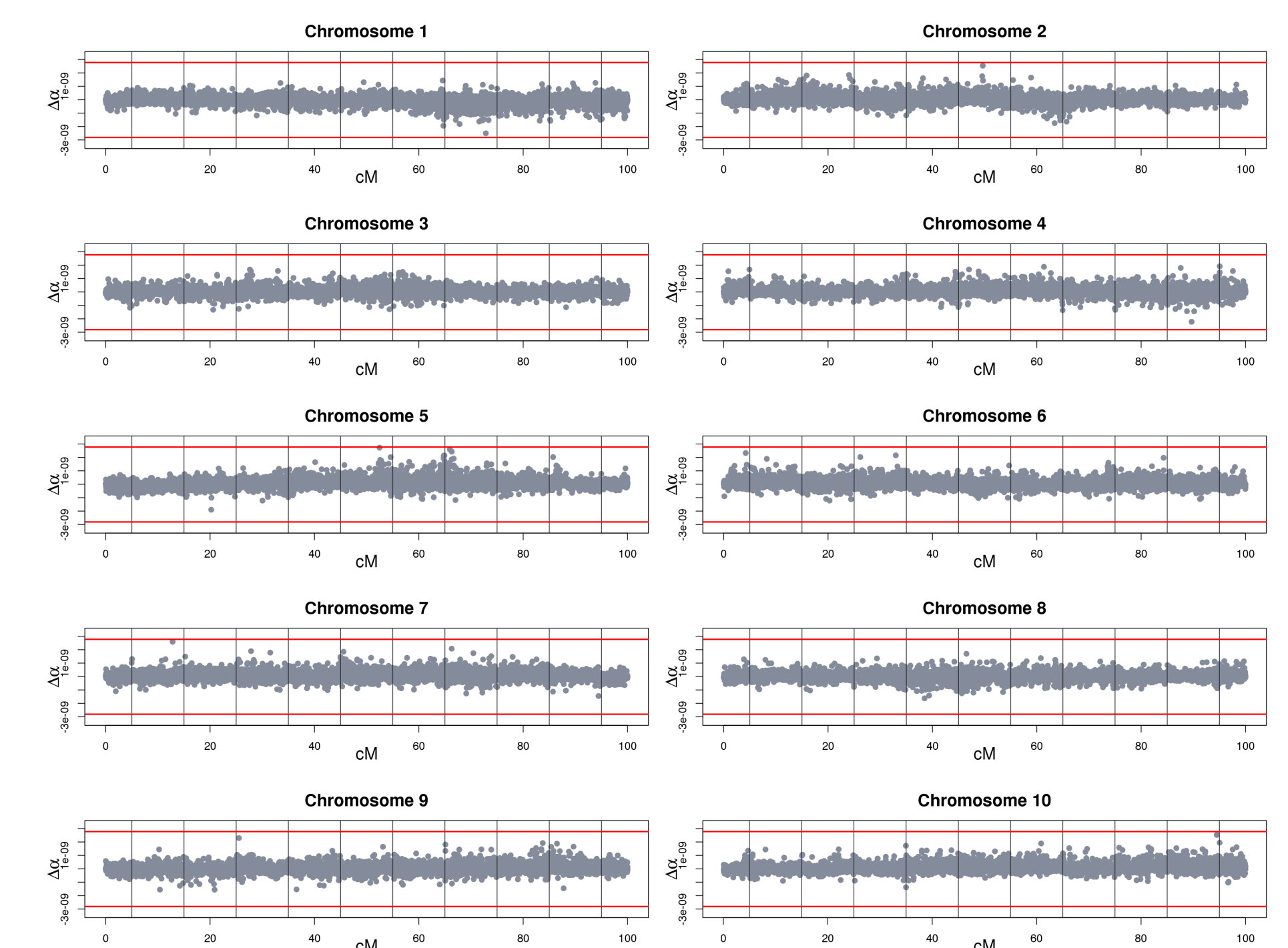

Figure 15: A Manhattan-style plot showing the product of $\Delta_i \times \alpha_i$ for each marker in a simulated selection experiment. Horizontal lines depict simulated significance thresholds while vertical lines depict the location of QTL.

References and acknowledgements

[1] Lorenz, A., et al. 2015. Selection for silage yield and compostion did not affect genomic diversity within the Wisconsin Quality Synthetic maize population. G3. doi: 10.1534/g3.114.015263.
[2] Gengler, N., et al. (2007) A simple method to approximate gene content in large pedigree populations: application to the myostatin gene in dual-purpose belgian blue cattle. Animal 1: 21– 28.

beissingert@missouri.edu