Supplementary Information:
Power analysis of artificial selection experiments

Darren Kessner and John Novembre

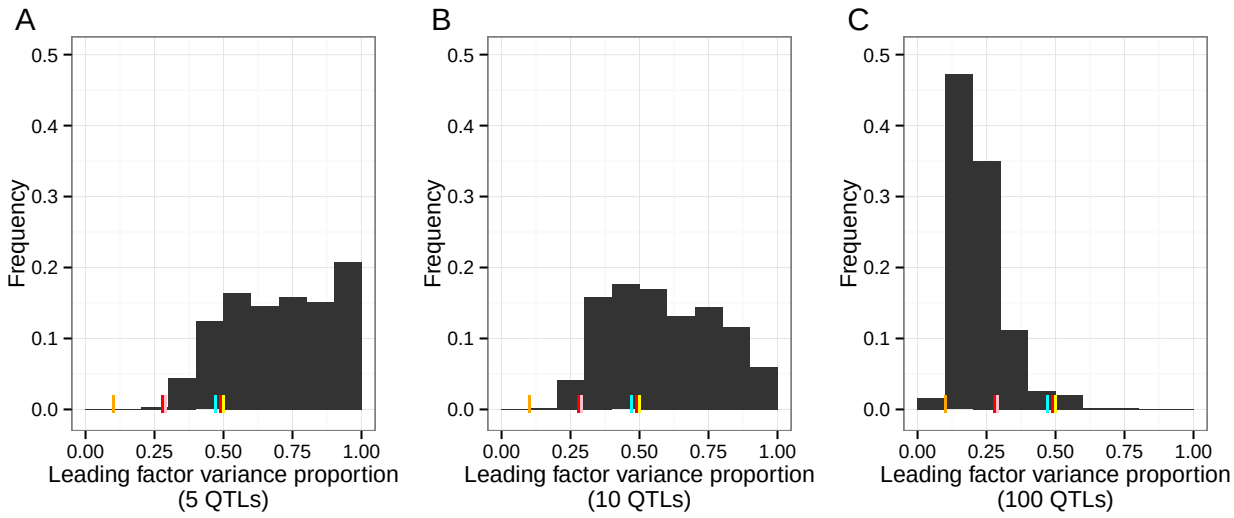

Figure S1. Proportion of variance explained by leading factor. Above are histograms of the proportion of variance explained by the QTL of largest effect, for random traits generated based on the DGRP SNPs and exponentially distributed effect sizes. Shown are histograms for random traits with A) 5 QTLs, B) 10 QTLs, and C) 100 QTLs (500 random trait simulations each). For comparison, we have included colored segments to indicate the proportion of variance explained by the leading factor of several quantitative traits as reported by selected empirical studies. [*Adh* activity, 2 experimental lines (.49, .28, red) (King *et al.*, 2012); susceptibility to viral infection, Drosophila C virus (.47, cyan), Sigma virus (.29, pink) (Magwire *et al.*, 2012); larval nicotine resistance (.50, yellow) (Marriage *et al.*, 2014); bristle number (.1, orange) (Mackay and Lyman, 2005)]. Note that the calculation of variance proportion assumes that QTLs segregate independently in the population (no linkage disequilibrium), which is generally not true for natural populations nor for the simulated populations.

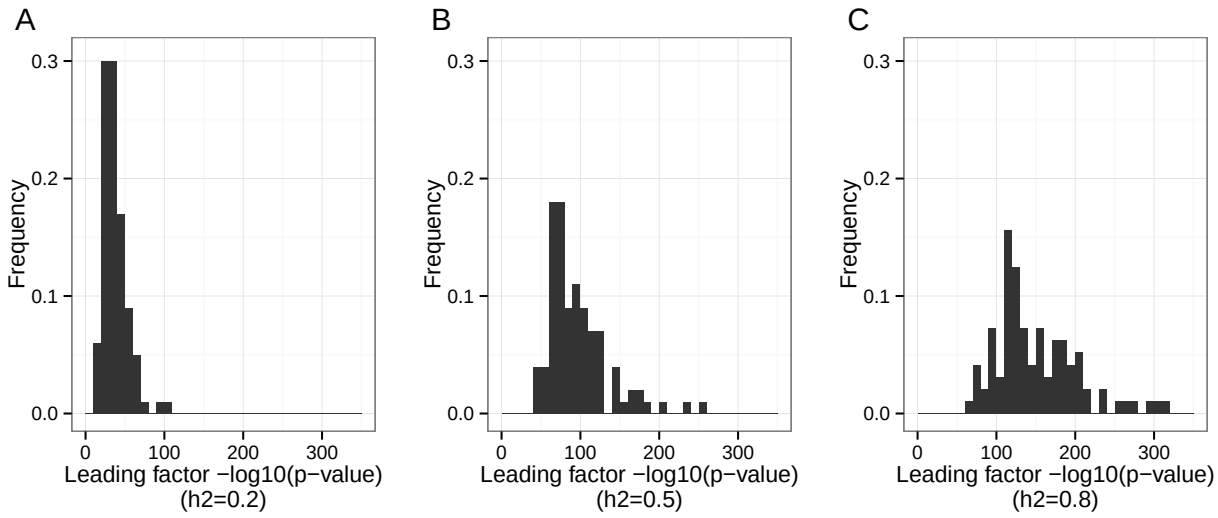

Figure S2. The -log10(p-values) for leading factors in simulated GWASs using the DGRP lines. Above are histograms of the -log10(p-value) for the QTL of largest effect, for random traits generated based on the DGRP SNPs, 100 randomly chosen QTL loci with exponentially distributed effect sizes, and the specified heritability value. We assume a GWAS is conducted using the DGRP founder lines with 20 replicate flies per line and using a standard ANOVA for computing p-values. Shown are histograms for random traits with A) $h^2 = 0.2$, B) $h^2 = 0.5$, and C) $h^2 = 0.8$ (100 random trait simulations each). (For smaller numbers of QTLs [not shown], the leading factor -log10(p-value)'s are larger than those shown here). The results show that under the assumed genetic architecture used in this study the leading factor should be readily detectable by a GWAS unless heritability is low. In practical cases, where a leading effect is not detected in a GWAS for a trait with modest or high heritability, it may be because the leading effect is smaller and/or at lower frequency than the simulated values used here. Together with Supplementary Figure S1 these results highlight that for some traits the results of this simulation study may not be applicable.

References

King, E. G., Merkes, C. M., McNeil, C. L., Hoofer, S. R., Sen, S., Broman, K. W., Long, A. D., and Macdonald, S. J. (2012). Genetic dissection of a model complex trait using the Drosophila Synthetic Population Resource. *Genome Res.*, **22**(8), 1558–1566.

Mackay, T. F. and Lyman, R. F. (2005). Drosophila bristles and the nature of quantitative genetic variation. *Philos. Trans. R. Soc. Lond., B, Biol. Sci.*, **360**(1459), 1513–1527.

Magwire, M. M., Fabian, D. K., Schweyen, H., Cao, C., Longdon, B., Bayer, F., and Jiggins, F. M. (2012). Genome-wide association studies reveal a simple genetic basis of resistance to naturally coevolving viruses in Drosophila melanogaster. *PLoS Genet.*, **8**(11), e1003057.

Marriage, T. N., King, E. G., Long, A. D., and Macdonald, S. J. (2014). Fine-mapping nicotine resistance loci in Drosophila using a multiparent advanced generation inter-cross population. *Genetics*, **198**(1), 45–57.