

Figures

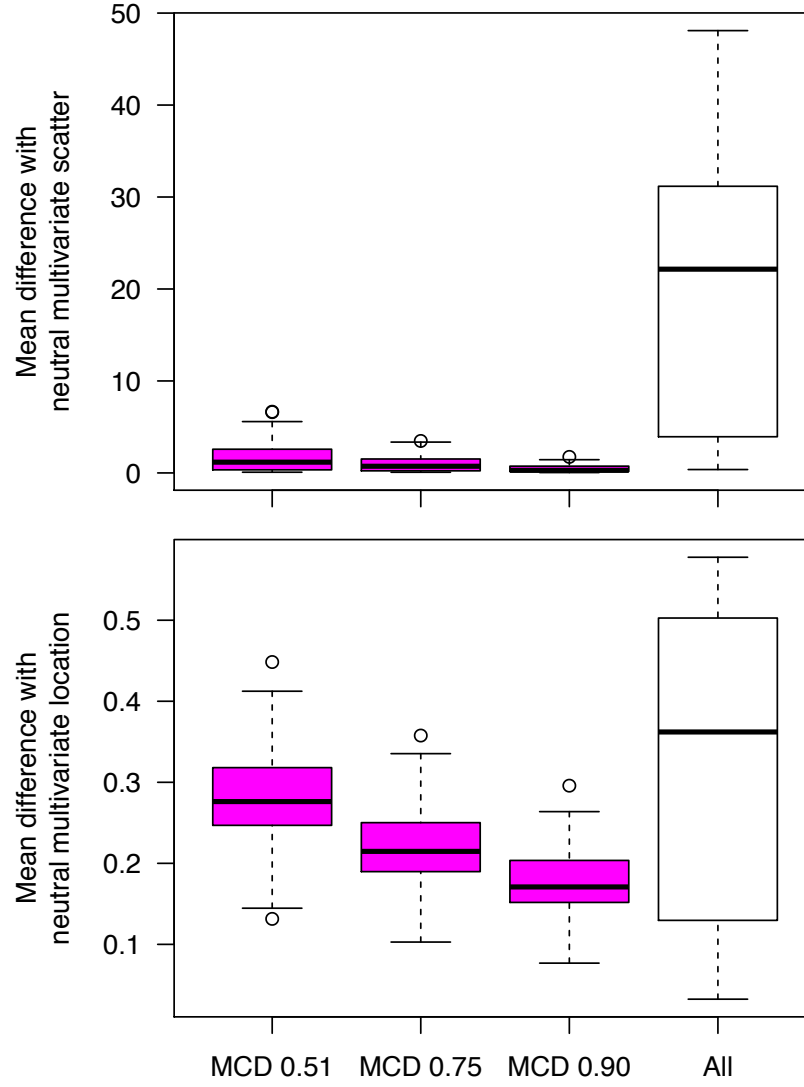


FIGURE 1. Mean difference between the estimate of location and scatter with the true neutral estimate. "All" refers to all the data (including selected loci), and MCD refers to the minimum covariance determinant estimate with the proportion of data used to create the subsets. Boxplots are summarized over all demographics and sampling designs.

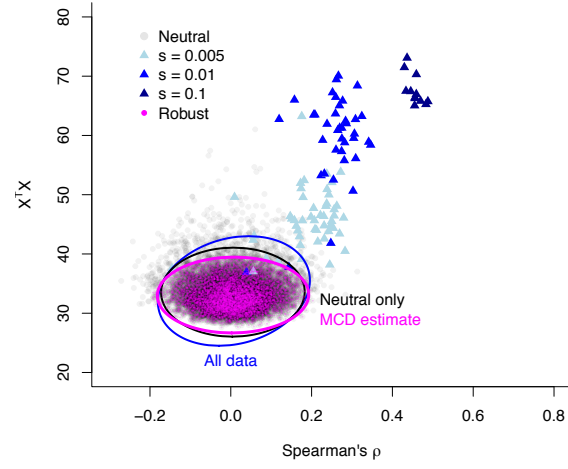


FIGURE 2. Comparison of 95% confidence intervals on the two dimensional ellipse given by the determinant on the covariance matrix between these two variables calculated from: (i) all the data, (ii) neutral loci only, and (iii) the MCD estimate.

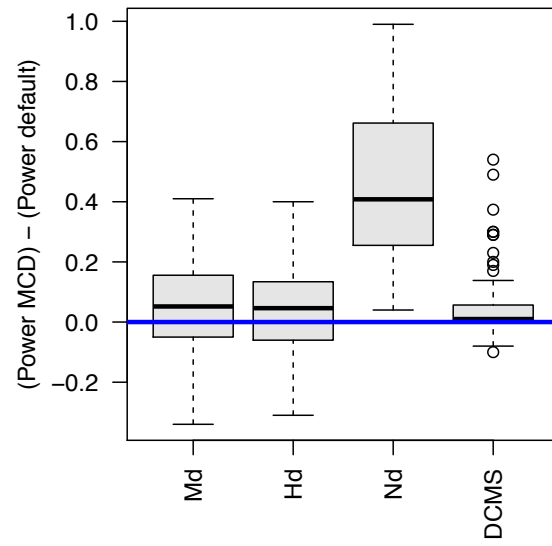


FIGURE 3. The difference in empirical power of the statistic to detect selection with and without the robust points identified by minimum covariance determinant (MCD) used in the calculation. Boxplots are summarized over all demographics.

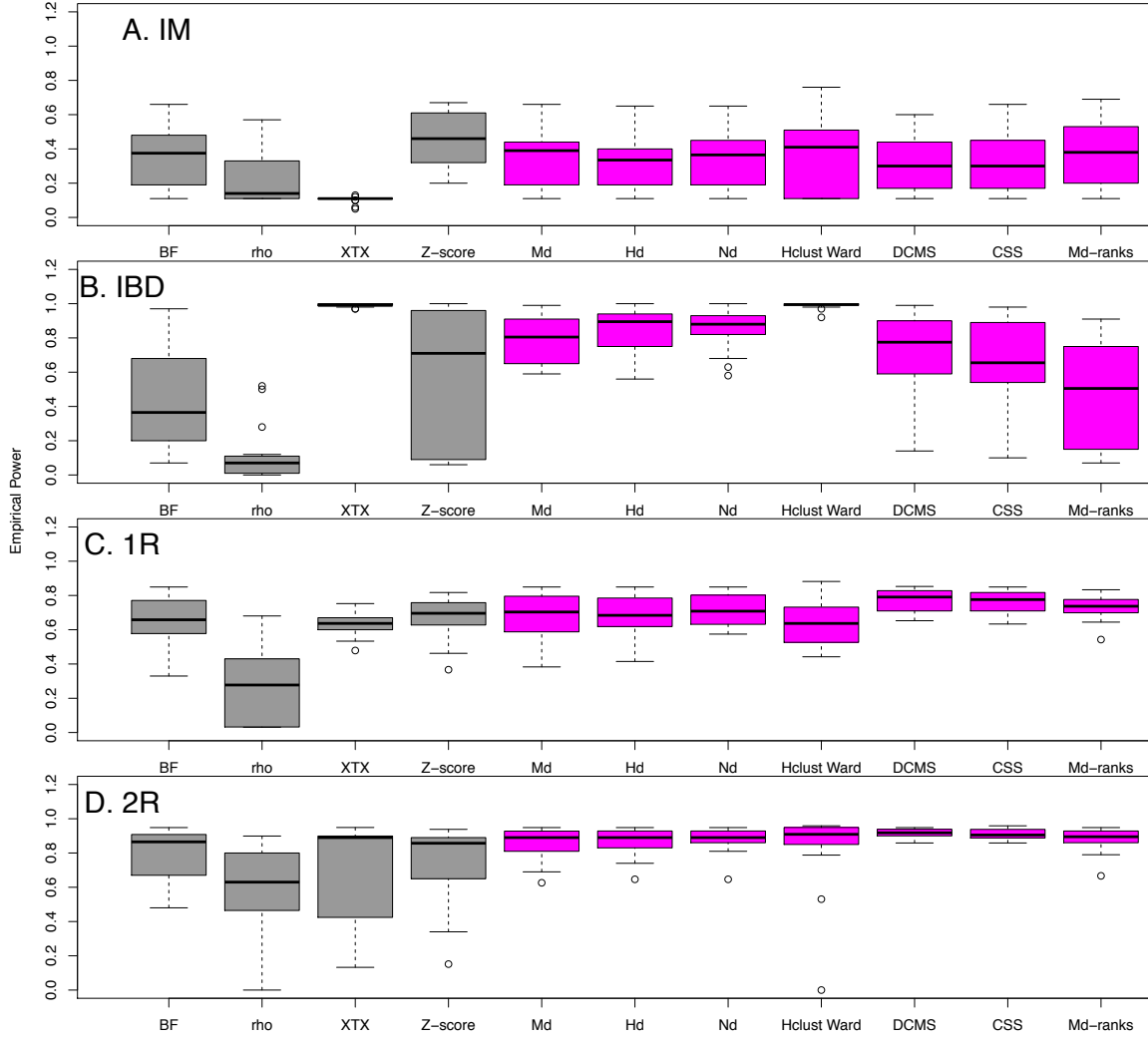


FIGURE 4. Comparison of empirical power of the univariate statistics and compound measures evaluated from the demography simulations. The univariate statistics include Bayes factor (BF) and Spearman's ρ (rho) from a genetic-environment association, a measure of genetic **diffentation** (XTX), Z -score from the program LFMM. The compound measures based on multivariate distances include: Mahalanobis distance (Md), harmonic mean distance (Hd), nearest neighbor distance (Nd), and hierarchical clustering (Hclust.Ward). The compound measures based on combining P -values include: decorrelated composite of multiple signals (DCMS), composite signal of selection (CSS), and Mahalanobis distance based on minus-log fractional P -values (Md-ranks).

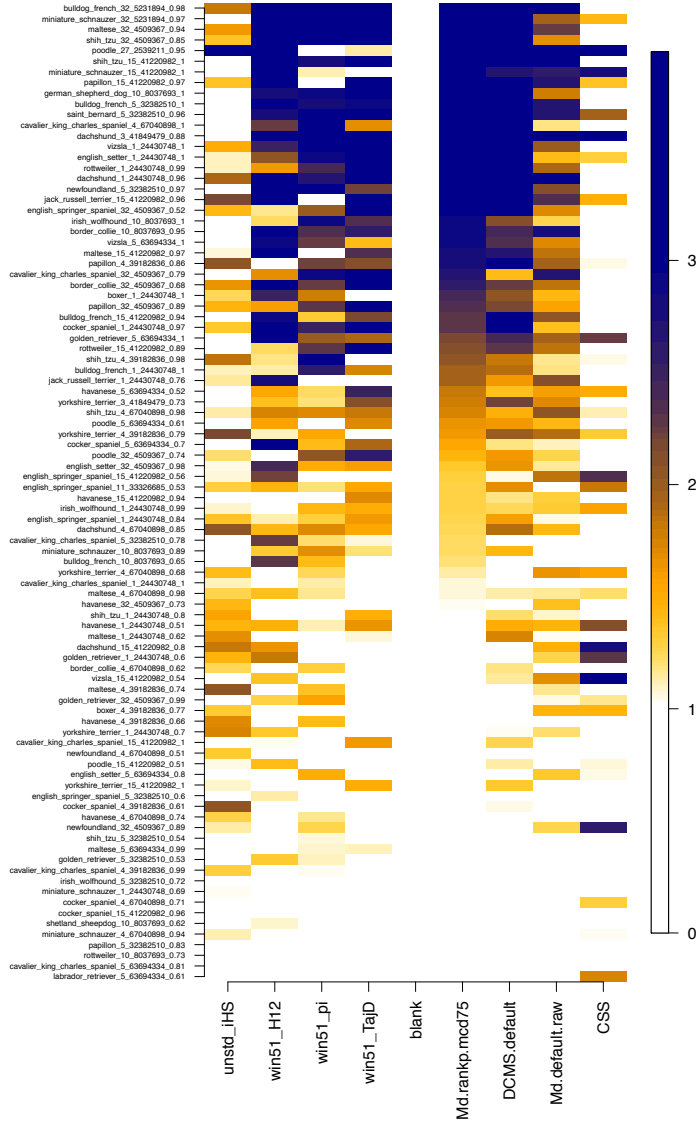


FIGURE 5. Heatmap of statistical quantile based on a minus-log P -value calculated from ranks evaluated for each individual locus in the empirical dog dataset. The univariate statistics include iHS, H12, nucleotide diversity π , and Tajima's D . The compound measures include: the Mahalanobis distance based on minus-log fractional P -values (Md-rank- P), the decorrelated composite of multiple signals (DCMS), the Mahalanobis distance based on the raw statistics, and the composite signal of selection (CSS). The row labels indicate the dog breed, chromosome, base pair of the focal SNP, and the allele frequency of the focal SNP.

TO DO: EDIT X-AXIS LABELS

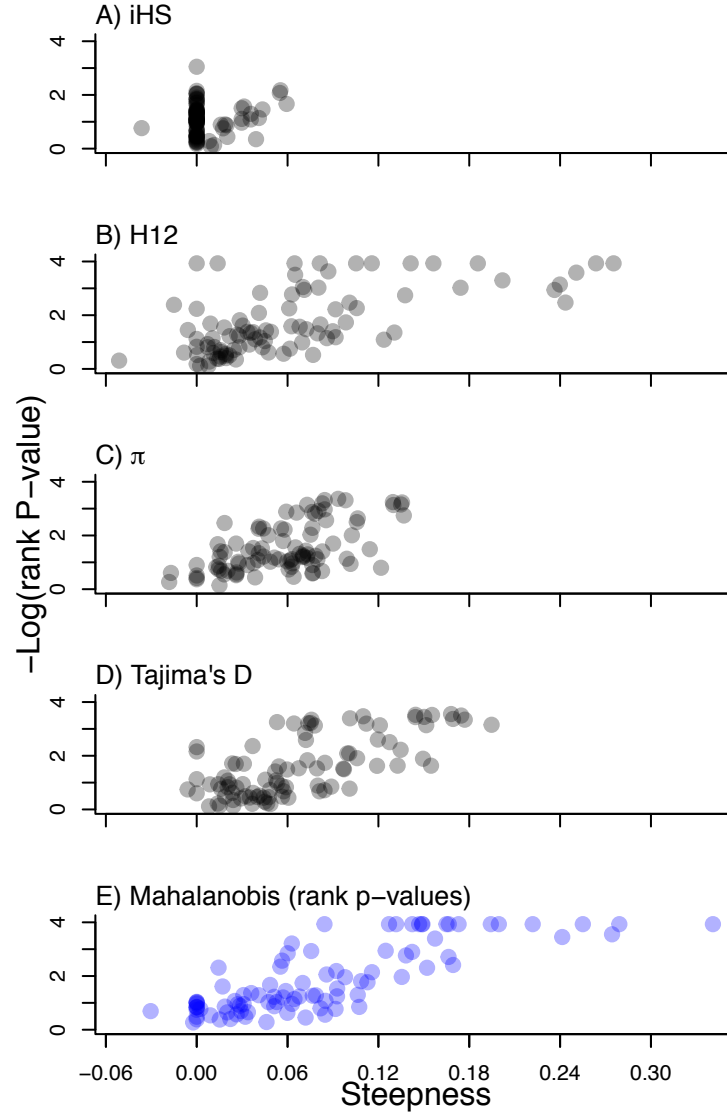


FIGURE 6. Statistical quantile based on a minus-log P -value calculated from ranks is plotted against steepness for each individual locus. Loci in the upper right hand of the graph have a steeper signal of selection and a more significant signal. The univariate statistics are plotted in A-D, compared to the Md-rank- P plotted in E.