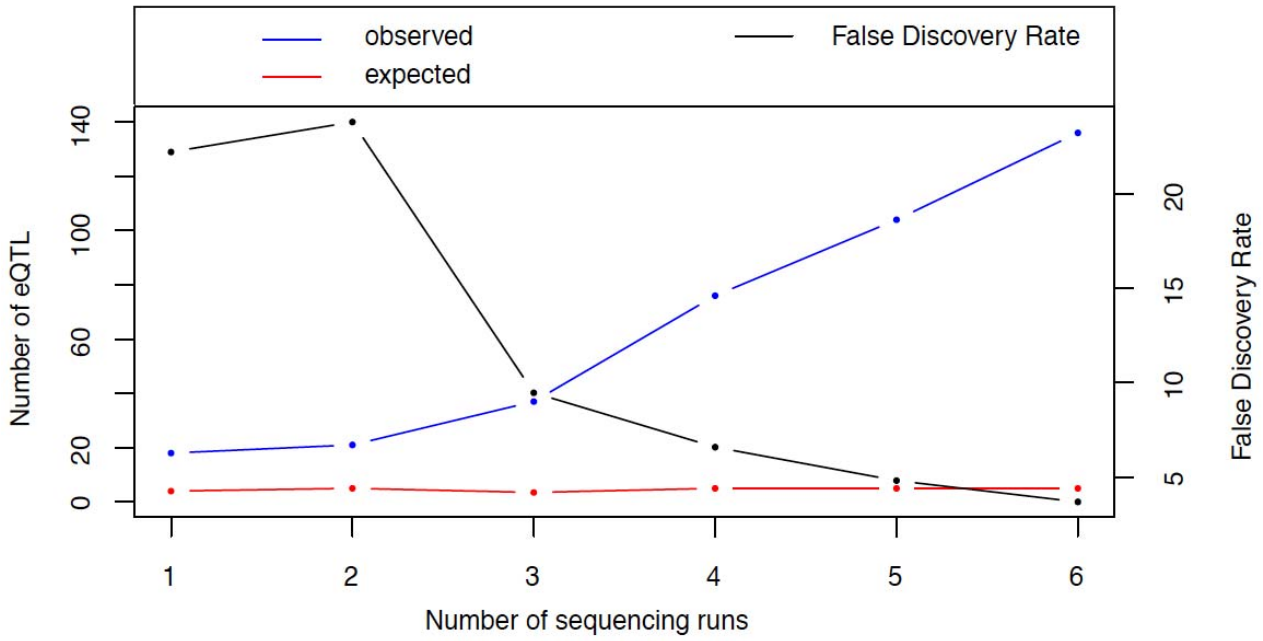

Figure S1. Expected and observed eQTL as a function of sequencing depth. These down-sampling analyses were used to evaluate the power of the study. As a proxy of sequencing depth we used the number of Illumina sequencing runs. Each Illumina sequencing run corresponded to approximately two million additional reads per animal.

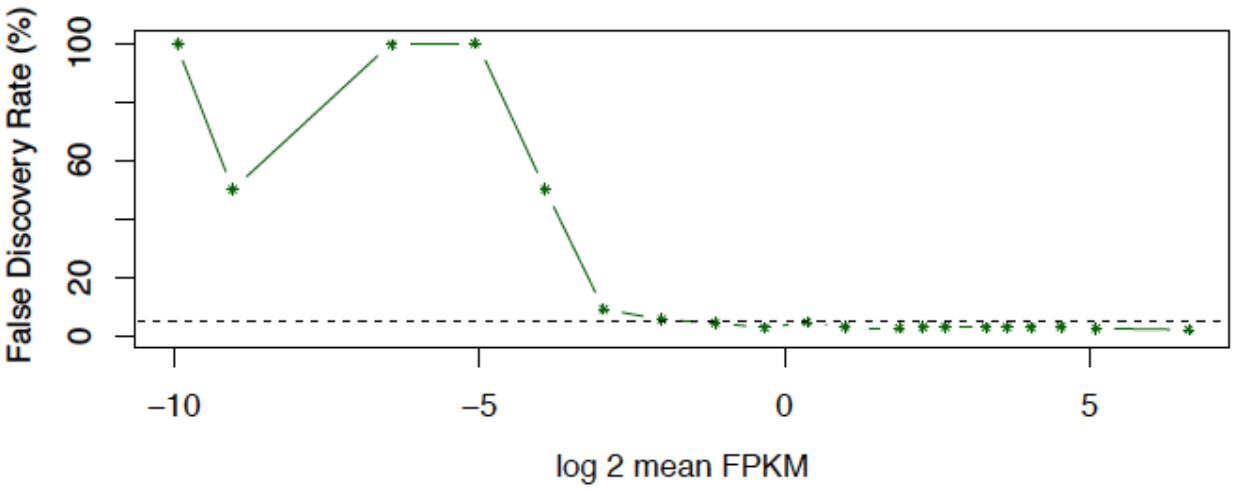

Figure S2. Power to detect eQTL as a function of gene expression level. Genes were grouped into 24 bins of 1,000 genes sorted by expression level. We performed eQTL mapping and permutations within each abundance bin, and plotted the resulting FDR as a function of the mean expression level within the given bin. For clarity, the four bins with the lowest abundance are not shown. The dashed horizontal line indicates the threshold of an FDR of 5%. We used the result shown in this figure to limit eQTL mapping to genes whose expression level was high enough so that eQTL can be reliably detected.

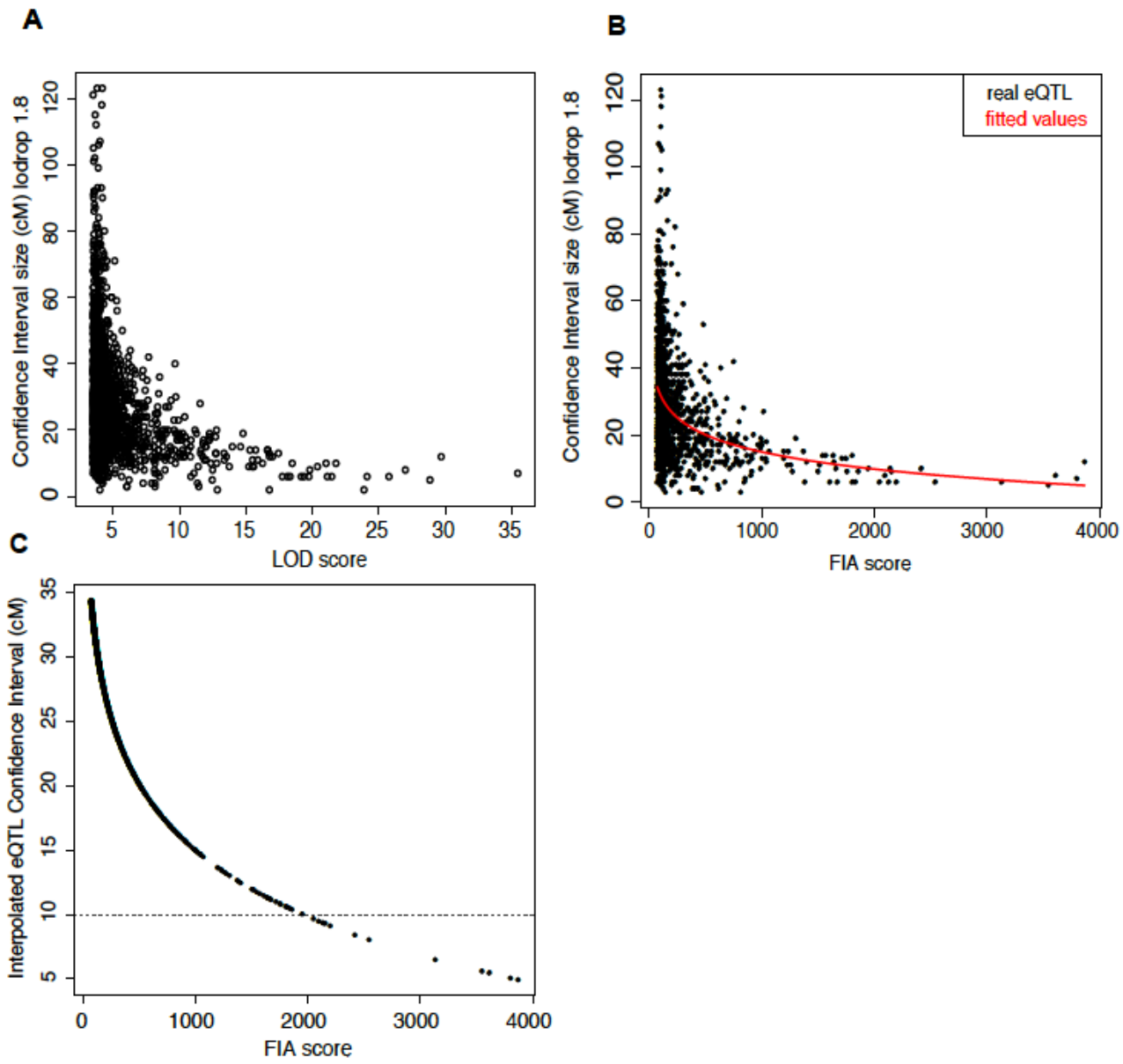

Figure S3. Defining FIA eQTL confidence intervals based on eQTL also detected by HKR. A, LOD scores of HKR eQTL and their 1.8 LOD drop confidence intervals. B, For the same set of eQTL as in A, shown are the FIA scores vs. the HKR 1.8 LOD drop confidence interval. The fitted function is shown as a red line. C, Interpolated confidence intervals for all FIA eQTL versus FIA scores. The confidence interval cut-off at 10 cM is shown as the dashed horizontal line.

D

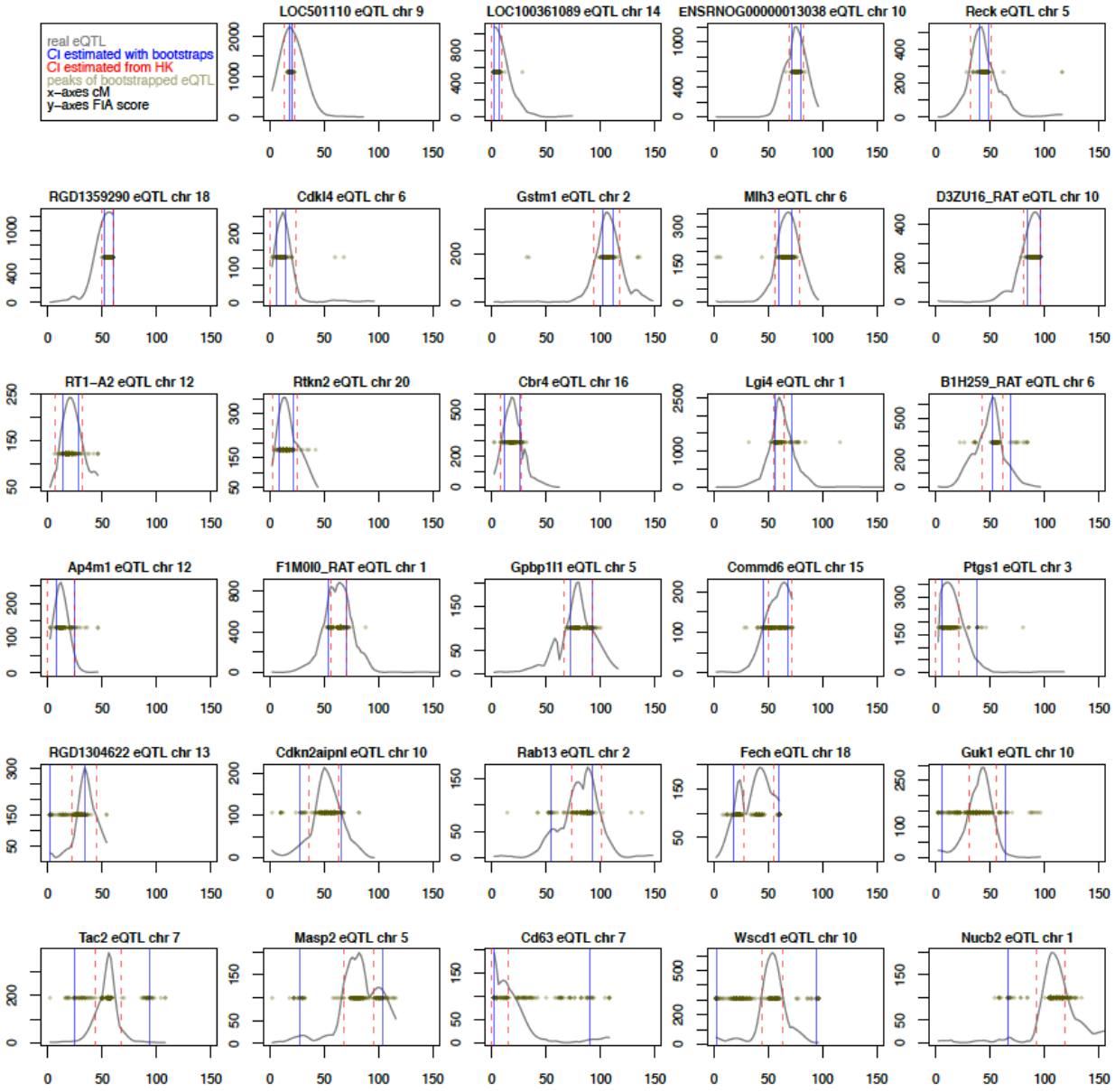

Figure S4 Comparing inferred FIA eQTL confidence intervals to bootstrap results. Shown are the 95% respective confidence intervals for 29 randomly chosen FIA eQTL. The FIA score profiles are shown as grey lines. X-axes show genomic locations in cM. Y-axes show FIA significance scores. The peak location of each of 200 bootstraps are shown as transparent green dots. Bootstrap derived 95% confidence intervals are shown as blue vertical lines and confidence intervals derived via HKR are shown as red dashed vertical lines.

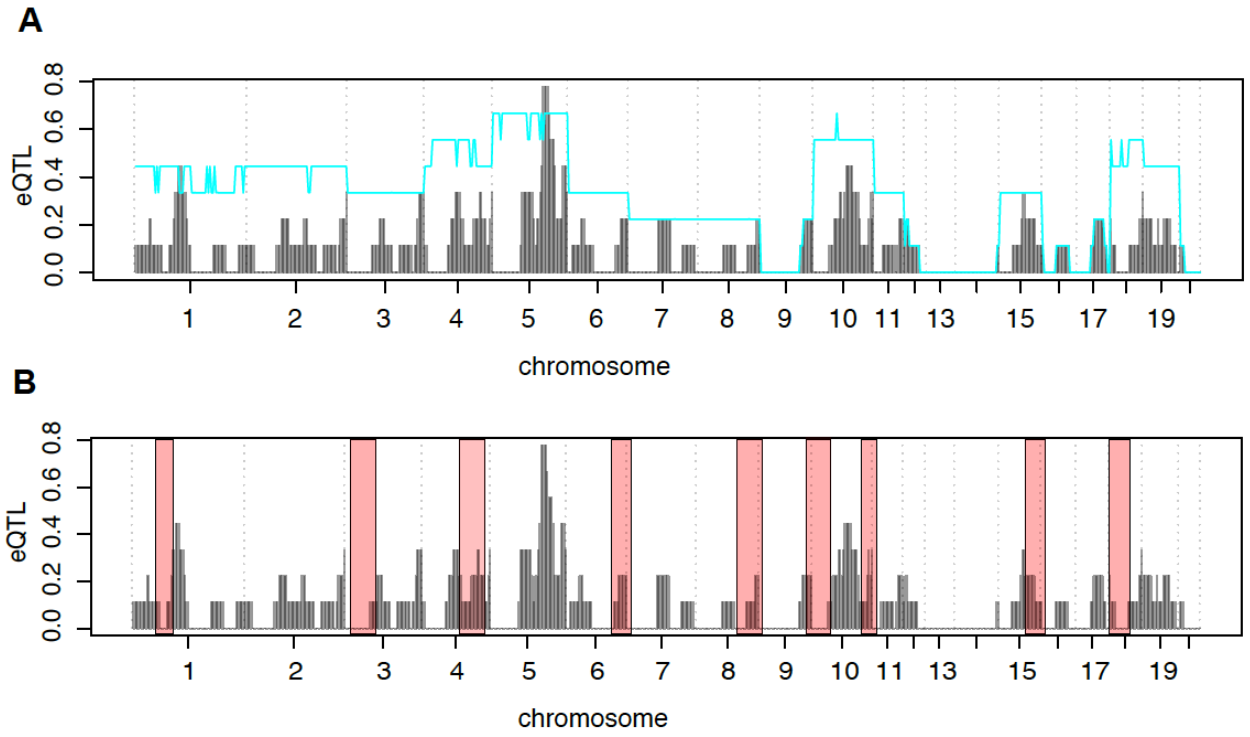

Figure S5. The number of distant eQTL along the genome. The grey bars are identical in A and B and indicate the average number of distant eQTL per 2cM window based on a sliding window analysis. The width of the window was 18 cM, corresponding to the median size of the eQTL confidence intervals. A, The 99th quantile of the expected eQTL density from 1,000 randomizations of eQTL positions per chromosome is shown as a blue line. B, Light red rectangles: Confidence intervals for tameness QTL locations.

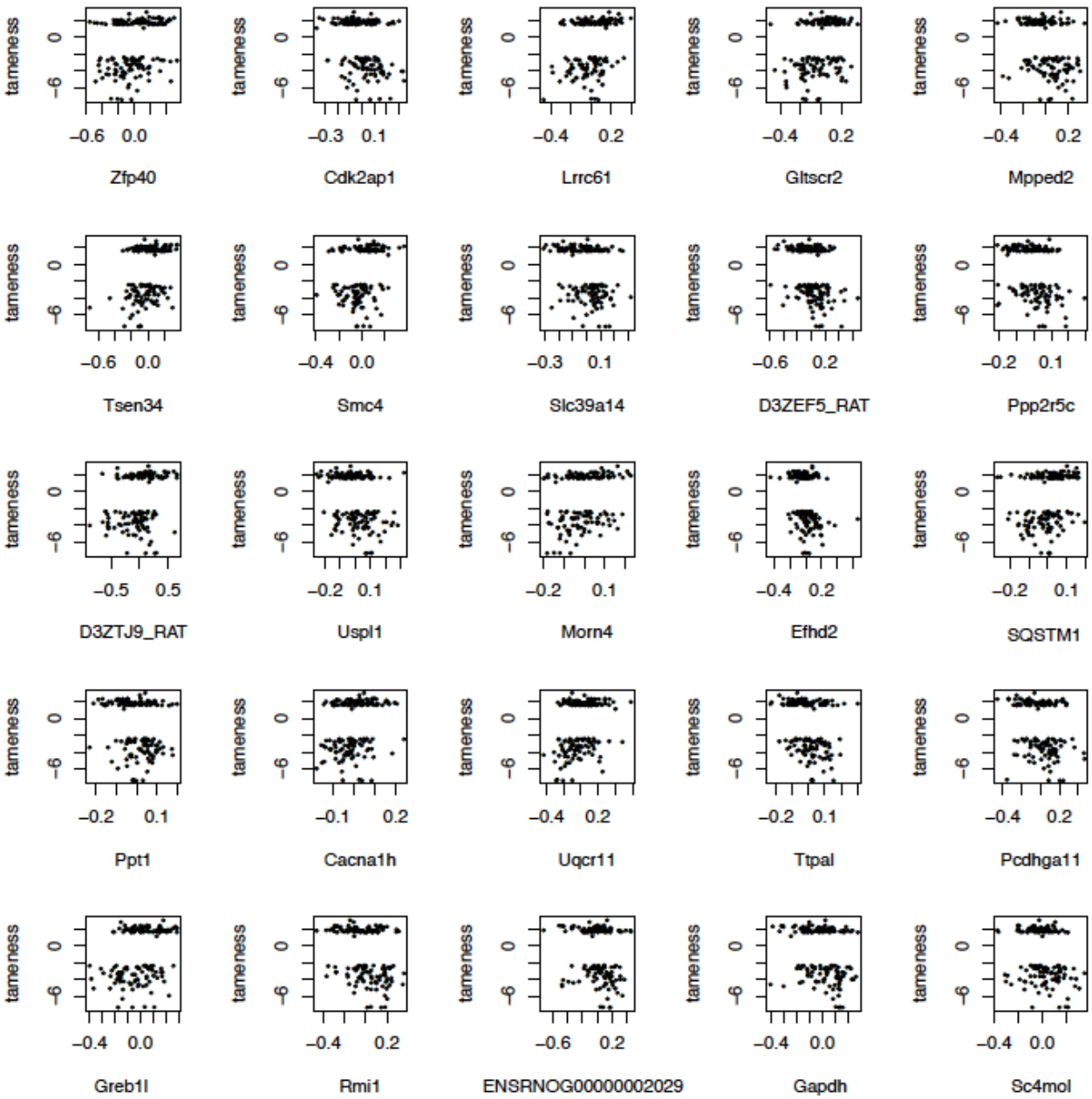

Figure S6. Tameness and gene expression values for the 25 genes with the highest correlations between these two traits. The bi-modality in the tameness score is a result of the selection of the 150 most extreme tame and aggressive individuals of the F2 intercross for inclusion in this study.

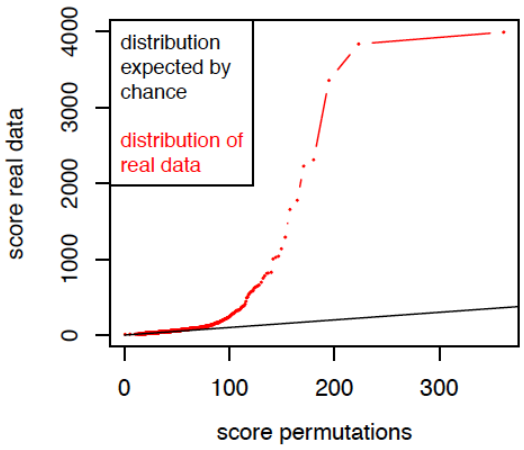

Figure S7. Q-Q plot of expected vs. real FIA eQTL mapping scores.

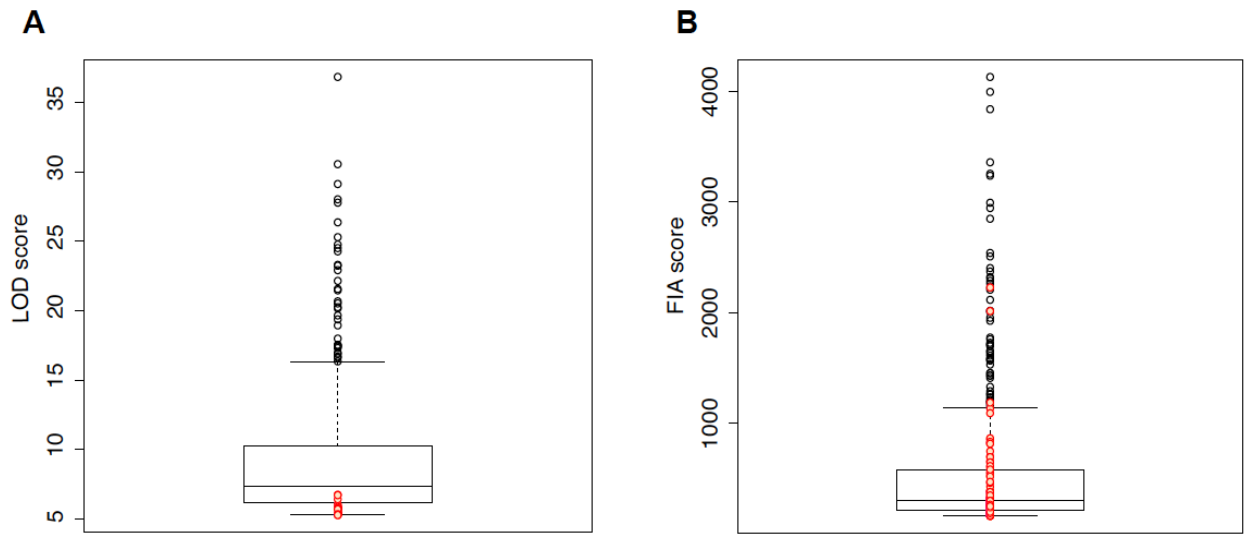

Figure S8. Statistical support for eQTL only found by Haley-Knott Regression (HKR) or Flexible Intercross Analysis (FIA). A. Boxplot of the LOD scores for all (n=443) eQTL identified by HKR. The LOD scores for eQTL (n=32) only detected by HKR (and not by FIA) are shown as red dots. B. Boxplot of the significance scores of all (n=689) eQTL detected by FIA. The scores for eQTL only detected by FIA (n=278) (and not by HKR) are shown as red dots.

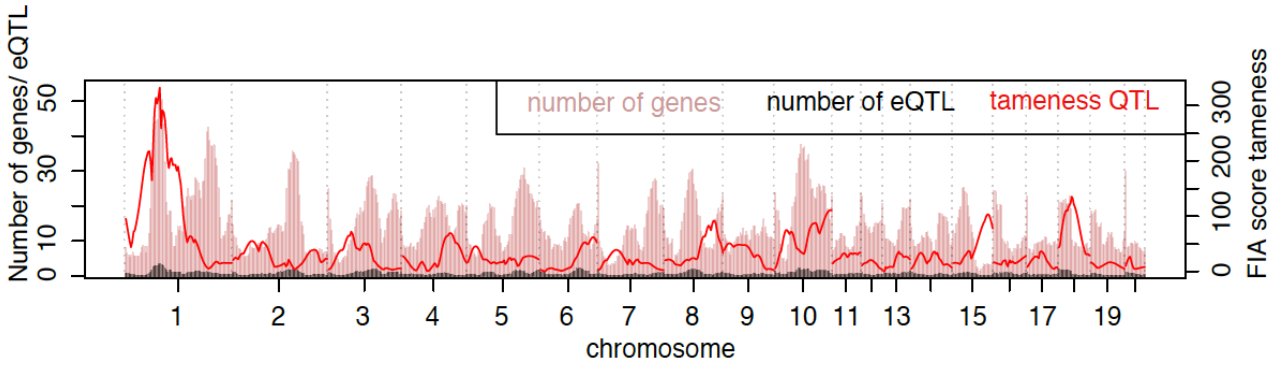

Figure S9. Gene density and statistical support for linkage to tameness along the genome. Light red bars: number of genes. Grey bars: number of local FIA eQTL. Red curve: FIA scores of tameness QTL mapping.

File S1

Detailed Results of Co-expression Network Analysis.

We applied Weighted Gene Co-expression Network Analysis (WGCNA) to investigate whether gene expression traits are clustered in expression networks. We also mapped the "Eigengene" for each cluster (the first principal component of expression level of all genes in a given cluster) to the genome using FIA. Detailed results from the cluster analyses are available in Supplementary Tables S3. The analysis grouped genes into 54 co-expression cluster. Whereas the „Eigengenes" in five of the clusters were correlated with tameness, none of the QTL influencing these clusters overlapped with tameness QTL. There is therefore currently no evidence that these clusters and tameness are influenced by shared genetic loci. Gene Ontology and transcription factor binding site enrichment analysis of the clusters that correlated with tameness showed enrichment for a large variety of biological processes. For example, the „grey60" cluster was enriched for retrotransposable elements (L1 transposase, FDR < 0.002%). Modulation of the transcription of retrotransposable elements occurs in the brain during exposure to stressful environmental conditions, possibly in response to glucocorticoid receptor activation [1]. The blue cluster was enriched in proteasomal ubiquitin-dependent protein catabolic processes (FDR < 2%) and enriched for binding sites of transcription factor Elk-1 (FDR < 10e-8). Elk-1 has been shown to be activated subsequently to glucocorticoid receptor activation [2] comforting a possible role of glucorticoids receptors in the formation of the tame-aggressive phenotype.

References to Supplementary Note

1. Hunter, R.G., B.S. McEwen, and D.W. Pfaff, *Environmental stress and transposon transcription in the mammalian brain.* Mob Genet Elements, 2013. **3**(2): p. e24555.

2. Gutierrez-Mecinas, M., et al., *Long-lasting behavioral responses to stress involve a direct interaction of glucocorticoid receptors with ERK1/2-MSK1-Elk-1 signaling.* Proc Natl Acad Sci U S A, 2011. **108**(33): p. 13806-11.

Supplementary Tables S1-S8

Tables S1-S8 are available for download as Excel sheets at

http://www.genetics.org/lookup/suppl/doi:10.1534/genetics.114.168948/-/DC1