File S1. SUPPLEMENTAL METHODS

Associations with MDD, bipolar disorder and schizophrenia

The summary statistics for the PGC GWAS study on bipolar disorder, the PGC GWAS study on major depressive disorder (MDD) and the PGC GWAS study on schizophrenia were obtained and used to test for putative associations between nine of the ten candidate genes identified (on further investigation the gene *LOC770352* was found to have only a relatively weak human orthologue and hence excluded). The summary statistics contained all the p-values calculated in the original paper, meaning it was not necessary to recalculate these scores. Associations were assessed at the orthologous gene positions +/- 120kb in the case of the bipolar dataset (the total average interval size for each gene was ~248kb, with these 120kb regions up and down-stream also being included to ensure local cis-eQTL control regions would also be analysed). In the case of the schizophrenia and MDD datasets as these were both meta-analyses and the number of individuals were so large (over 20k individuals in each schizophrenia study and 18k in the MDD study) and from so many diverse populations that the search was restricted solely to the orthologous gene position +/- 50kb due to the LD being greatly reduced in these datasets. To find a significance threshold, we randomly selected 1000 sections of the genome each time selecting the lowest p-value from each segment. As nine separate regions were being tested, the total interval size to select was 250kb x 9 = 2.25Mb for the bipolar studies, and 120kb x 9 = 1080kb for the schizophrenia and MDD studies. In this way, a 5% significance (P=0.0001 for the PGC bipolar study, $P=1x10^{-5}$ for the schizophrenia study, P=0.0002 for the MDD study) and a 20% suggestivity threshold (P=0.0004 for the PGC bipolar study, $P=1x10^{-4}$ for the schizophrenia study, P=0.001 for the

MDD study) was found, tailored to each individual study and the size and number of the regions tested.

Associations with Mouse HS cross

The mouse HS cross dataset was used to assess the ten candidate genes. This cross consists of repeated crossing between eight founder strains (with the repeated intercrossing between the different founders giving the short haplotype blocks and increased resolution). SNPs were selected in the orthologous gene regions, using the mus genome build-37, with the closest SNP to each gene selected. Three different open field behaviours – latency to move, total activity and time spent in the centre of the arena were used. Associations were tested using the freely available website (http://mus.well.ox.ac.uk/gscandb/), with the standard additive model option used. . Assessing significance is somewhat problematic as the significance threshold is only provided for a whole genome scan, rather than 10 individual SNPs as assessed here. A LOD score of ~4.8 was considered suggestive at a whole genome level and ~5.6 was significant for the full scan. We attempted a relatively crude randomisation test by randomly selecting 100 genes and recording the LOD score for each trait at the gene. This gave a significance threshold ~LOD 4 for each behavioural trait.