Genetic Liability for Internalizing Versus Externalizing Behavior Manifests in the Developing and Adult Hippocampus: Insight From a Meta-Analysis of Transcriptional Profiling Studies in a Selectively-Bred Rat Model

Supplement 1

Supplemental Methods

Detailed Methods for Animal Husbandry:

Selective Breeding: For the first generation, we chose 120 Sprague–Dawley rats (Charles River, Inc) to begin the bHR and bLR lines that exhibited extreme locomotion scores within a novel environment (the top and bottom 20%, respectively, detailed methods below). Since then, we have maintained 12 breeding families in each line. For generations F1-F41, our breeding strategy amplified the differences between the two lines while minimizing inbreeding. During this time, to propagate the bHR and bLR lines we only mated the males and a females with the most extreme locomotor scores from each family. During the F30 generation, a second colony was begun at the University of Alabama-Birmingham using a selection of bHR/bLR rats from the MBNI colony. Later, at generation F42, the breeding strategy at MBNI changed to target locomotor scores closest to the designated phenotypic score (~1800 beam breaks for bHR and ~150 beam breaks for bLR) to prevent the selectively-bred lines from becoming more extreme.

Animal Husbandry: On postnatal day 1 (PND1), litters were reduced to 12 pups (6 males, 6 females) and raised by their mothers within the breeder room on a 14:10 light:dark schedule (lights on 4 a.m., Standard Time, illumination levels approximately 200-400 lux, target temperature 72°F (71-74°F), target humidity 30%). Following weaning (PND 21), the rats were housed 2-3 per cage in-house and maintained on a 12:12 light:dark schedule (lights on 6 a.m., Standard Time). Access to food and water was *ad libitum*.

Locomotor Testing: For each generation, locomotor response to a novel environment was assessed between P50–75 using methods previously described (1). On the morning of the test (between 7:30 a.m.-12:30 p.m.), the rats were placed in a standard clear acrylic cage (43x 21.5x 24.5 cm - akin to their normal housing) in a different room with housing conditions similar to the breeding room (on shelving, with similar lighting, temperature, and humidity). Locomotor

activity was monitored every 5 min for 1 hr by two panels of photocells, one of which recorded horizontal activity and one of which recorded rearing or vertical activity. Locomotion scores were than calculated as a sum or horizontal and vertical beam breaks. Up to 36 animals were tested simultaneously. Males and females were tested on separate days.

Detailed Methods for the Individual Datasets:

MBNI_AffymetrixRae230_F4

This previously unpublished microarray dataset was generated at MBNI using bHR and bLR male rats from generation F4 (*n*=6 per group). The rats underwent typical locomotor testing to validate phenotype and were adults (between P102-P108) at the time of sacrifice. The rats were sacrificed via rapid decapitation followed by immediate brain extraction. The whole hippocampus, frontal cortex, and hypothalamus samples were rapidly dissected on ice, fast-frozen at –40°C, and stored at –80°C before processing. Only the hippocampal data was included in our current analysis. TRIzol reagent (Invitrogen, Calsbad, CA) was used to extract total RNA, followed by purification using RNeasy RNA purification columns (Qiagen, Valencia, CA). The quality and concentration of the RNA was determined using an Agilent bioanalyzer (Palo Alto, CA) and wavelength absorbance (260/280 nm ratio) by Nanodrop. The samples were transcriptionally profiled using Affymetrix Rat Expression Set 230 (RAE230A) microarray according to standard manufacturer's procedures. The RAE230A arrays contained around 15,900 probe sets that primarily targeted well-annotated full-length genes. Typically, there were eleven pairs (perfect match/mismatch) of 25-mer oligonucleotide probes in each probeset. The samples were run in two batches: the first included the hippocampus and cortex samples and the second included the hypothalamus samples. Each brain region was run on a separate chip.

For our current analyses, the *ReadAffy()* function (from R package *affy* version 1.54.0; (2)) was used to read the data from the hippocampal Affymetrix .CEL files into R studio (version 1.0.153). An expression set was generated using the Robust Multi-Array Average method (RMA: (3)). Using a custom .cdf (4) we summarized the probe signals into probesets for the RAE230A chip ("rae230arnentrezgcdf_19.0.0") which mapped the probe sequences to Entrez Gene IDs (downloaded from http://nmg-r.bioinformatics.nl/NuGO_R.html on Jan 2017, release date Nov 2015). We re-annotated the data according to Official Gene Symbol using the function *org.Rn.egSYMBOL()* from the R package *org.Rn.eg.db* (version 3.4.1; (5)), and then removed rows lacking annotation (leaving 9,959 gene symbols). Quality control identified two samples with low sample-sample correlation values (around 0.66), leaving a final sample size of *n*=5 per group.

MBNI_AffymetrixRgU34A_F6

This dataset was obtained from a published microarray study (6) conducted at MBNI using bHR and bLR rats from the F6 generation (GEO accession #: GSE29552). Three different age groups were included: P7, P14, and P21, with *n*=6 per group (36 total). The rats were all baseline males, having undergone no behavioral testing. The rats were sacrificed via rapid decapitation, and the whole hippocampus and nucleus accumbens were immediately dissected. Only the hippocampal data was included in our current analysis. After extracting mRNA from the targeted tissue, cDNA was synthesized and run on Affymetrix Rat Genome (RG) U34A GeneChips following procedures previously described (6). This chip contained probesets targeting approximately 7,000 full-length sequences and around 1,000 expressed sequence tag clusters. Typically, there are 16 pairs (perfect match/mis match) 25-mer oligonucleotide probes in each probeset. We reanalyzed the original data (.CEL files) using methods similar to those used for the *MBNI_AffymetrixRae230_F4* dataset and a custom .cdf for the rgU34A chip ("rgu34arnentrezgcdr_19.0.0"). Following re-analysis, there were a total of 4,588 unique gene

symbols present in the data. Data underwent standard quality control and no outlier samples were identified.

MBNI_IlluminaRatRef12v1_F15

This dataset was obtained from an unpublished microarray study conducted at MBNI using bHR and bLR rats from generation F15. The study was performed on male P14 baseline animals (*n*=6 per group). Rats were sacrificed via rapid decapitation and tissue was immediately extracted from the whole hippocampus. RNA was extracted and purified as discussed in ((6); *MBNI_AffymetrixRgU34A_F6*) and divided into two aliquots earmarked for transcriptional profiling using different microarray platforms. One aliquot of RNA was analyzed using the Illumina RatRef-12v1 Beadchip microarray following Illumina's standard recommended procedures.

The Illumina RatRef-12v1 contains 23,375 50-mer oligonucleotide probes targeting coding transcripts. To extract the AVG_Signal for all probes for each of the samples, we used the *read.idat()* function (R package *limma* version 3.28.21; (7)). We performed a log2 transformation and then quantile normalized the data using the *normalize.quantiles()* function in the R package *preprocessCore* (version 1.34.0; (8)). The final dataset represented a total of 21,568 unique gene symbols. During quality control, a potential outlier was noted in the PCA. However, due to the strong sample-sample correlations among all samples (>0.98), removing the outlier was deemed unnecessary.

MBNI_AffymetrixRae230_F15

This microarray dataset was produced as part of the same unpublished study as the *MBNI_IlluminaRatRef12v1_ F15* dataset. It used aliquots of the same RNA samples, but was run on a different platform: Affymetrix Rat Expression Set 230 A. This chip was designed with 15,900 probe sets that primarily targeted well-annotated full-length genes and some expressed sequence tags. There are typically 11 pairs (perfect match/mis match) of 25-mer oligonucleotide probes in

each probeset.

We reanalyzed the original data (.CEL files) using methods similar to those used for the *MBNI_AffymetrixRae230_F4* dataset. The final dataset contained 9,958 unique gene symbols. During quality control, we found consistent irregularity associated with one subject (bLR subject X6_RN230_HC_P14_F15_L03.CEL) and it was removed prior to further analyses. Since this dataset is a technical replicate of MBNI_IlluminaRatRef12v1_F15, we considered accounting for covariation between the two sets of biological replicates in our meta-analysis using a multilevel model, but additional analyses revealed an almost perfect lack of correlation between the results produced by the two platforms (**Fig S2**). Therefore, the information that the two datasets was providing in our meta-analysis model appeared for all practical purposes independent.

MBNI_RNASeq_F29

This previously unpublished RNA-Seq dataset was generated at MBNI using bHR and bLR rats from the F29 generation. It included both adult (P60) and P14 rats (*n*=2 per group). These rats were all baseline males and had not undergone any behavioral testing, including standard locomotor testing. Rats were sacrificed through rapid decapitation, followed by immediate brain extraction and whole hippocampus dissection. Extraction, purification, and assessment of the quality and concentration of RNA followed similar procedures to those used for *MBNI_AffymetrixRae230_F4.*

Purified RNA was shipped to the Genomic Services Laboratory at HudsonAlpha (https://gsl.hudsonalpha.org/index) for short-read (50 bp length) single-end sequencing. RNA Qubit was used to assess concentration (range: 720-1120 ng/uL). The dilutions were denatured through heat application at 70˚C for two minutes and then checked for quality with a RNA Nano bioanalysis chip. Indexed non-stranded cDNA libraries were then prepared using standard Illumina procedures and a TruSeq library preparation kit (Illumina, San Diego, CA) after which they were pooled for sequencing with four samples per pool. The pooled libraries were clustered

on a HiSeq flowcell (Illumina, San Diego, CA) and sequenced using a first generation HiSeq 2000 sequencer. Reads were mapped to the rat reference genome (RGSC v3.4) using Tophat2 (9) and Bowtie2 (10) with default parameters and a standard sequence read output. Only reads that aligned with the genome once were maintained for later differential expression analysis. Transcript assembly was collected using Cufflinks and Cuffmerge (11), and bHR vs. bLR differential expression was assessed for each age group (P14 and adult) using CuffDiff (11).

When performing our meta-analysis, we did not have access to original gene level summary data per subject at the time of the analysis, and therefore performed no further quality control or re-analysis procedures besides averaging across *t*-test statistics by gene symbol to eliminate over-representation (final unique gene symbol count in each dataset: 16,237 for adults and 16,342 for P14).

Alabama_NimbleGen_F34

This dataset was obtained from a published microarray study conducted at University of Alabama-Birmingham within the laboratory of a former MBNI investigator ((12); GEO accession #: GSE88874). The tissue in this study originated from the F4 generation of a new colony of selectively bred bHR/bLR rats. This new colony was created by selectively breeding F30 rats shipped from the MBNI colony, therefore we refer to their generation as F34 (F30+F4). This study included bHR and bLR baseline male rats from four different age groups: P7, P14, P21, and adult (*n*=5 per group, or 40 total). Rats were sacrificed by rapid decapitation and brains were flash frozen in isopentane and cooled by dry ice to -30˚C before being stored at -80˚C. Tissue extraction was conducted by first sectioning the brains in a cryostat to alternating sections of 20 and 300 µm, after which the 20 um sections were histologically stained to provide guidance. Tissue punches from the 300 um sections were used to extract samples from the dorsal hippocampus, amygdala, and prefrontal cortex (details in (12)). Only the hippocampal data was used in our meta-analysis.

Transcriptional profiling was conducted using NimbleGen Rat Gene Expression 12x135 microarray targeting 26,419 target genes using five 60mer oligonucleotide probes/target (135,000 probes total). Since Nimblegen uses proprietary software, we were unable to reprocess the raw microarray data and instead depended on the normalized, gene-summary data from the original study (provided on the GEO database). That said, the original data preprocessing (12) followed a pipeline very similar to our own, including RMA and quantile normalization. We performed a log2 transformation and then used the R package *org.Rn.eg.db* (version 3.4.1; (5)) to translate the original annotation (Refseq accession number) to gene symbol. After annotation, the data represented 10,674 unique gene symbols. Quality control procedures were similar to those described earlier. Both PCA and the sample-sample correlations revealed an outlier (a P21 bLR) which was removed from our analysis.

MBNI_RNASeq_F37

This previously unpublished RNA-Seq dataset was generated at MBNI using bHR and bLR male adults from generation F37 (*n*=6 per group). Unlike the other datasets in our study, this one also included animals that showed an intermediate locomotor response to a novel field (Intermediate Responder (bIR) rats), which were obtained by cross-breeding bHR and bLR rats. We did not include the bIR data in the meta-analyses, but incorporated it later in follow-up behavioral analyses. In adulthood, the rats received our standard locomotor testing (discussed above), and then were later tested for anxiety-like behavior (ages: bHR/bLR: P160-P167; bIR: P65-75) using the Elevated Plus Maze (EPM) and methods described previously (1,13). In brief, the animals were placed for 5 min on a black Plexiglas device with elevated arms (70 cm from the floor) joined together in the shape of a cross. Two of the arms were enclosed by 45 cm high walls (“closed arms”), whereas the other two arms (“open arms”) as well as the intersection were left open. The room was dimly lit (approximately 30 lux). The percentage time spent in each section (open arms, closed arms, or central intersection) was monitored using automated

videotracking software (Noldus Ethovision, Leesburg, VA).

Following behavioral testing, the rats were sacrificed (age: bHR/bLR=P160-P167, bIR=P126-134) by rapid decapitation and the whole hippocampus was extracted on ice, rapidly frozen, and stored at -80 degrees C. Nucleotides were extracted using Qiagen AllPrep DNA RNA miRNA Universal Kit 50. Extracted RNA was evaluated for total concentration and quality using a Nanodrop spectrophotometer (concentration range 285-432 ng/ul, 260/280 ratio range 1.61-1.80) and then sent to the University of Michigan DNA Sequencing Core (https://seqcore.brcf.med.umich.edu). Before additional processing, the RNA was re-assessed for quality using the TapeStation automated sample processing system (Agilent, Santa Clara, CA) and only samples with RNA integrity numbers (RINs) of >8 were included in the analysis. The cDNA library was constructed using 0.1-3ug of total RNA and the Illumina TruSeq Stranded mRNA Library Preparation kit (Catalog #s RS-122-2101, RS-122-2102) (Illumina, San Diego, CA).

The final cDNA libraries were checked for quality once again by TapeStation (Agilent) as well as qPCR through the use of Kapa's library quantification kit for Illumina Sequencing platforms (catalog # KK4835, Kapa Biosystems,Wilmington MA). The samples were clustered on a cBot automated cluster generation system (Illumina, San Diego, CA) for clonal amplification. The samples were then hybridized to the slide ("flow cell") of a HiSeq 2000 (Illumina, San Diego, CA) with 6.66 samples per lane and underwent a 100 cycle paired end run in High Output mode using version 3 reagents. Following sequencing and demultiplexing, the RNA-Seq reads were aligned to the rat genome (Rnor_6.0) using the SubRead aligner (14) using default parameters with the exception of indel detection (maximum length of indel that could be detected=0). The featureCounts program (15) then generated the gene-level RNA-Seq count summaries for each sample based on ENSEMBL annotation (Ensembl v.81). This count summary dataset was then filtered to exclude rows of data from genes that did not meet a minimum threshold of 4 samples

with greater than or equal to 10 counts. Rows that lacked official gene symbol annotation were also excluded.

Our current analysis used the log2 fragments per million gene-level summary output for each sample provided by the *voom()* function (R package *limma*; (7)). Quality control included 1) visualization of the overall log (base2) transformed transcript expression across all subjects via boxplot, 2) examination of the overall reads (mean and standard deviation) per subject, 3) visualization of a subject/subject correlation matrix to identify particularly atypical samples (R<.95). No outlier samples were identified.

MBNI_RNASeq_F43

This previously unpublished RNA-Seq dataset was generated at MBNI using adult bHR and bLR male rats from generation F43. These rats were not truly baseline, as the original study compared repeated injections of either antidepressant medication (fluoxetine or desipramine) or vehicle (VEH=1:1 saline and water) solution in bHR and bLR rats (14 days of intraperitoneal injections - P78-P92, 1 per day) under either standard laboratory housing conditions or chronic variable stress. In our meta-analysis, we included only the five male bHR VEH rats and five male bLR VEH rats housed in standard conditions. These rats did not undergo our standard locomotor testing in adulthood, but on the last day of injections they underwent social interaction testing after 15 minutes of exposure to the anxiogenic open arms of the EPM (conditions and device described above). Social interaction testing was conducted using a procedure described previously (16). In brief, under dim light (approximately 30 lux) experimental rats were placed in an open-topped, rectangular, transparent social interaction box (23 5/8" x 16 3/8" x 13 1/4") with a stimulus rat. The stimulus rat was a conspecific (male, outbred Sprague-Dawley) of similar age and weight and had no previous contact with the experimental rat. Rats were placed simultaneously in the box and the amount of time the experimental rat spent initiating social interaction (i.e., grooming, sniffing,

following, and crawling over or under) with the stimulus rat was scored for 5 min by an observer blind to the experimental condition of the animal. In between each test, the box was cleaned with 30% ethanol.

One hour after undergoing exposure to the EPM, the animals were sacrificed by rapid decapitation and the brains were flash frozen in isopentane cooled on dry ice. The whole hippocampus was later dissected from each brain and immediately put in TRIzol™. RNA was extracted using the Zymo RNA isolation Kit and shipped to the University of Michigan DNA Sequencing Core for non-stranded short read (50-bp length) single-end sequencing using a procedure almost identical to that reported above for the *MBNI_RNASeq_F37* study.

The preprocessing procedures for the data closely follow those employed for the *MBNI_RNASeq_F37* dataset, including procedures for alignment to the reference genome (Rnor_6.0), the production of gene-level RNA-seq counts summaries for each sample using ENSEMBL annotation (Ensembl v.85), transformation into log2 fragments per million, and quality control. The read count dataset produced by featureCounts was filtered to exclude rows of data from genes that did not meet a minimum threshold of 5 samples with greater than or equal to 16 counts. After annotation, there were 16,393 unique gene symbols represented in the data. No outlier samples were identified.

Detailed Methods for qPCR Validation:

Animal husbandry and tissue collection: Using selectively-bred rats from later generations (F51, F55), 6 bHR and 6 bLR males were sacrificed at ages P14 and P90 (**Fig S1**). For the P14 collection, the rats were sacrificed within 3-5 min of separation from the dam. The rats designated for the P90 collection were weaned, housed, and loco-tested as previously described. Sacrifice was performed via rapid decapitation without anesthesia. Brains were immediately hemisected and flash frozen by submersion in -30°C 2-methylbutane.

RNA extraction and cDNA synthesis: Brains were stored at -80°C for fewer than 6 months before processing. Hippocampus was dissected from one hemisphere and homogenized using a QIAshredder kit (Qiagen #79654), and RNA was extracted using an RNeasy Mini Kit (Qiagen #74104). cDNA was synthesized using a 20 µL reaction containing 400 ng of RNA template (iScript cDNA Synthesis Kit, Biorad#1708891).

qPCR: The primers were custom-designed to target Bmp4 (ACC# NM_012827.2; forward primer: 5'-CCCTGGTCAACTCCGTTAAT-3', start = 1214; reverse primer: 5'-AACACCACCTTGTCGTACTC-3', start = 1319) and the reference gene Gapdh (ACC# NM_017008.4; forward primer: 5'- GTTTGTGATGGGTGTGAACC-3', start = 459; reverse primer: 5'-TCTTCTGAGTGGCAGTGATG-3', start = 628). Calibration curves for the Bmp4 and Gapdh primers were constructed using a standard dilution analysis of stock cDNA that had been previously synthesized from a mixture of adult and P14 bHR and bLR hippocampi (H20, 0.1 uL, 0.5 uL, 1 uL, 2 uL) in triplicate using iQ™ SYBR® Green Supermix. The bLR/bHR samples were then analyzed using a similar procedure in triplicate, with the samples from each time point processed within a separate batch.

Data analysis: The calibration curves revealed efficiencies close to 1 (Bmp4: R^2=0.98, Gapdh: R^2=0.99, **Fig S1**), therefore the sample data for each time point was analyzed using the traditional Livak method (17,18). After averaging the triplicate quantification cycle (C_q) values for each sample for each probe, the data was normalized by subtracting the C_q for the reference gene (Gapdh) from the target (Bmp4) for each sample (ΔC_q). Group differences in ΔC_q were assessed using Welch's two sample t-test (18). Group differences in the Cq values for the reference gene (Gapdh) were also examined as a control (18). To provide an intuitive comparison to the microarray and RNA-Seq data, the data were plotted as log(2) fold change values ($-\Delta\Delta C_q$), by subtracting the average ΔC_q for the bLRs from all ΔC_q values, and then multiplying by -1.

Additional Methods for Collective Analysis Procedures:

Examining the Relationship Between Gene Expression and Behavior: For the two adult datasets that contained associated behavioral data we used simple bivariate linear models to explore the relationship between gene expression and behavior: total locomotor score, percent time in the open arm of the EPM, and percentage of time spent interacting socially following exposure to a single mild stressor. A general effect of phenotype on these variables was evaluated using ANOVA (Type 3). We also examined the number of fecal boluses excreted during the EPM test served as a measure of anxiety independent from spontaneous exploration (19), but these results are not discussed due to the general redundancy of the findings with more traditional EPM measures.

Effect Size Calculation: We calculated the effect size (Cohen's d and variance of d) for the effect of bHR/bLR phenotype on log(2) gene expression (hybridization signal or FPM) within each age group (R package *compute.es* (20)). The raw data for the small MBNI_RNASeq_F29 dataset were inaccessible at the time of analysis, so we re-derived effect sizes from previous *CuffDiff* t-statistic output (11), averaging the few results representing duplicated gene symbols. To visually compare effect sizes across datasets we created forest plots using *forest.rma()* in the package *metafor* (21)).

Comparison with Genetic Results: The original bHR vs. bLR exome sequencing study targeted 129,237 genetic variants (single nucleotide variants or SNVs, identified by Rnor5 coordinates (22)) in 12 bHR and 12 bLR F37 rats (22). This sample included the six bHR males and bLR males used in the hippocampal transcriptional profiling dataset *MBNI_RNASeq_F37.* Of the variants profiled,1.44% exhibited full allelic segregation (average heterozygosity around 6x10E-4). To identify QTLs for exploratory locomotor behavior, 416 of the SNVs that segregated bHR and bLR rats were then targeted in an F2 bHRxbLR intercross (*n*=314), and then QTL peaks were identified using the Haley-Knott (H-K) regression method with 1 cM resolution (22). To

compare the exome sequencing results with our differential expression findings from the adult meta-analysis, we accessed the coordinates for all genes included in the adult meta-analysis (Rnor_6 annotation from the R package *org.Rn.eg.db* (5)) and used the NCBI Genome Remapping Service (https://www.ncbi.nlm.nih.gov/genome/tools/remap, accessed 8/8/2019 using the first 100 bp of each gene and default parameters).

To determine whether there was an enrichment of differential expression nearby segregating variants, we filtered the exome sequencing results to only include results from genes located on the 20 autosomes and 2 sex chromosomes (Rnor_5). We then examined the percentage of differentially expressed genes (adult meta-analysis FDR<0.05) found nearby segregating variants as defined using either standard (Bonferonni-corrected p<0.05) or more stringent criteria (Bonferonni-corrected p<5.00E-05). The distances examined included within +/- 100 kb, 250 kb, 500 kb, or 1 MB of the segregating variants. Enrichment was defined using Fisher's exact test to evaluate the cross table for differential expression vs. presence of a nearby segregating variant for each of the 8 permutations of stringency & distance.

To determine whether there was an enrichment of differential expression within QTLs for exploratory locomotor behavior, for each gene symbol we extracted the logarithm of the odds (LOD) score from the QTL analysis for the 1 MB bin that encompassed the gene and the two adjacent (+/- 1 MB) bins (Rnor_5), and then calculated the maximum. Enrichment was defined using Fisher's exact test to evaluate the cross table for differential expression (adult meta-analysis FDR<0.05) vs. location within 1 MB of a QTL peak (using the cut-off for genome-wide significance in (22): LOD>4).

The results from these genetic analyses were compared to QTLs identified in the Rat Genome Database ((23): accessed 08/08/2019 with key words "Anxiety", "Stress", and "Despair", and Rnor6 chromosomal coordinates) using Rnor_6 annotation. The top enriched chromosomal loci (p<0.001, FDR<0.01) identified by Positional Gene Enrichment Analysis (24) were also

converted to with Rnor_6 annotation for reporting and comparison with our genetic findings by cross-referencing the encompassed genes with available coordinates (https://www.ncbi.nlm.nih.gov/genome/gdv/?org=rattus-norvegicus).

Gene Set Enrichment Analysis (GSEA): The gene lists were ranked according to the beta coefficient values (β) representing the magnitude and direction of effect estimated by the meta-analysis. We performed GSEA using the *fgsea()* function (R package *fgsea* (25); settings: 100,000 permutations, maxSize=1000) and gene set matrix files (.gmt) containing either standard gene ontology for rats (*GO2MSIG*; (26): *"Rattus_norvegicus_GSEA_GO_sets_all_symbols_highquality_April_2015.gmt"* downloaded 6/2017), or customized hippocampal-specific gene sets **(Table S1)** from: *1)* human hippocampal co-expression modules from post-mortem and freshly-resected tissue (27); *2)* mouse hippocampal co-expression modules from behaviorally-profiled animals in the hybrid mouse diversity panel (28); *3)* sets of genes with expression specific to hippocampal neuronal subtypes or subregions (*Hipposeq:* (29)).

Expression Within Particular Cell Types: To add additional annotation for the top differentially expressed genes (FDR<0.10), we used the new mousebrain.org database (30), which contains information about basal gene expression within a wide variety of cell types in the nervous system as identified using a large (n>500,000) mouse single-cell RNA-Seq dataset. To perform this analysis, we only considered the 81 cell types (clusters) that were likely to be present in our whole hippocampal dissections (as defined by a tissue source, probable location, and/or region of expression that included the hippocampus). These 81 cell types were subcategories within six general cellular classes (oligodendrocytes, vascular, immune, astrocytes, ependymal, neurons). We determined whether each of our top differentially-expressed genes were reliably expressed in a particular hippocampal cell type (trinarization score for cluster >0.95) and then

assessed the enrichment of that expression as compared to cell types within the other classes of cells.

Detailed Methods for the Comparison of bHR vs. bLR DE Results with Hippocampal DE Results from Other Bred Rat Models:

The Selection of Transcriptional Profiling Studies for Comparison: To determine which genes might show a similar pattern of DE across bred rat models targeting behavioral traits that represent extremes on the internalizing/externalizing spectrum, we located hippocampal transcriptional profiling results using two strategies. First, we searched both Pubmed and Gene Expression Omnibus (GEO) using the terms “rat*”, “selectiv*”, breed*” OR “bred”, “hippocamp*”, “microarray” OR “RNA-Seq*” to identify publications that might be of interest. We then added to this list of publications using an additional search for specific bred rat models targeting behavioral traits that represent extremes on the internalizing/externalizing spectrum that were overviewed in relevant review papers (31,32) along with “microarray” OR “RNA-Seq*”. The results from this search were then triaged manually based on whether: 1) we considered the targeted phenotype to reflect some aspect of internalizing or externalizing behavior that paralleled the bHR/bLR model, such as anxiety, stress-susceptibility, depressive-like behavior, hypoactivity/hyperactivity, risk-taking, novelty-seeking, or impulsivity; 2) the results reflected macro-dissected hippocampus (whole, dorsal or ventral hippocampus) versus a microdissection of the hippocampus (such as the Dentate Gyrus, CA1, or purified cell type). Finally, an additional publication was added via suggestion by a reviewer (33).

Following this selection process, we identified nine publications that compared gene expression using transcriptional profiling in macro-dissected hippocampal samples from bred rat

models with phenotypes reflecting some aspect of internalizing versus externalizing behavior ((33–41), **Figure S10).**

Re-analysis of Blaveri et al. (2010, (34)***): Pre-processing and Differential Expression Analysis:*** One of the publications (34) characterizing hippocampal gene expression in Flinders Sensitive Line (FSL) vs. Flinders Resistant Line (FRL) had publicly released their hippocampal microarray data (GEO Accession# GSE20388), allowing us to re-analyze their data using methodology similar to that used to process the bHR/bLR microarray data for the meta-analysis.

We extracted relevant metadata from the GEO website using the *getGEO()* function (package *GEOQuery* (42)), and filtered the data to only include hippocampal samples in our analysis. The *ReadAffy()* function (R package *affy* version 1.54.0; (2)) was used to read the data from the hippocampal Affymetrix .CEL files into R studio (version 1.0.153). An expression set was generated using the Robust Multi-Array Average method (RMA: (3). Using a custom .cdf (4) we summarized the probe signals into probesets for the Rat Genome 230 2.0 chip ("rat2302rnentrezgcdf_22.0.0") which mapped the probe sequences to Entrez Gene IDs (downloaded from http://mbni.org/customcdf/22.0.0/entrezg.download/ on Jan 2020). We re-annotated the data according to Official Gene Symbol using the function *org.Rn.egSYMBOL()* from the R package *org.Rn.eg.db* (version 3.4.1; (5)), annotating 14,074 rows with gene symbols. Quality control did not identify any outliers, leaving the full sample size (FRL: n=22, FSL: n=17).

As expected, principal components analysis revealed that the top two principal components of variation in the gene expression data (PC1, PC2) correlated strongly with Cohort. Therefore, differential expression in respect to bred Line (FSL vs. FRL) was assessed using a linear regression model that included Cohort (1 vs. 2) as a co-variate. P-values were corrected for false discovery rate (FDR) using the Benjamini-Hochberg method (*multtest* package: (43)).

Preparation of a master database of DE genes identified in bred rat models: With the exception of Blaveri et al. (34) (*see discussion above*), the only DE results available from previous studies performing hippocampal transcriptional profiling in bred rats were in the form of published tables identifying the top DE genes, as defined using publication-specific statistical criteria. This meant that a formal enrichment analysis or meta-analysis could not be performed. However, we were able to determine which of the top DE genes overlapped with those identified in our bHR/bLR meta-analysis. To do this, we combined the results from each of the 9 publications into a master database. We made formatting consistent, removed results for probes that lacked gene symbol annotation or that were multi-annotated, and updated the gene symbol annotation for three of the studies that appeared to have older annotation (39–41). We then assigned the direction of effect so that the bred animals that we considered to have the greatest internalizing behavior were the comparison group, so that a positive direction of effect (1) would signal relatively higher expression in this group and a negative direction of effect (-1) would signal relatively lower expression in this group. Finally, we consolidated results from probes representing the same gene symbol within a dataset by averaging their direction of effect. Using this method, we found that a small percentage of the differentially-expressed genes showed a different direction of effect for probes representing the same gene symbol within their original dataset (0.8%, 24/2948), and thus ended up with a direction of effect that represented an average value falling between -1 and 1.

The final data frame (**Table S4)** provides the direction of effect for all 2598 gene symbols that were identified as DE in at least one of these 8 publications (using publication-specific criteria), in our reanalysis of Blaveri et al. ((34), FDR<0.05), or in our bHR/bLR meta-analysis (FDR<0.10).

Supplemental Results

Additional Co-expression Networks Enriched with bHR/bLR Differentially Expressed Genes

Co-expression modules can capture regionally-important cell types and functions that remain undocumented in traditional ontology databases (44). We observed an enrichment of bHR/bLR effects within six hippocampal co-expression modules within the P14 meta-analysis (FDR<0.05) and within five co-expression modules within the adult meta-analysis (FDR<0.05, **Fig S6**). For the sake of conciseness, within the main text we only highlighted the results from two of these co-expression modules (*lightcyan, sienna3*)– additional results are described here.

M11, a medium-sized co-expression module (149 genes) present in both healthy and diseased hippocampi in humans, as well as in mouse hippocampi (27), was interesting because it showed an enrichment of bHR/bLR effects both within the P14 and adult meta-analyses (FDR<0.10) with expression generally elevated in bLRs compared to bHRs (NES<0). However, during further investigation, we could not assign a clear function to the *M11* module – it contained no genes exhibiting robust bHR/bLR effects (FDR<0.05), no enrichment of documented PPI or traditional functional pathways, and only two genes with differential expression in both development and adulthood (Fbox31, Arhgap39, p<0.05).

Of the remaining seven hippocampal co-expression modules that exhibited an enrichment of bHR/bLR effects in either P14 or adulthood, four seemed particularly provocative. Within the P14 meta-analysis, the *skyblue* module (top genes: Mobp and Mag) showed higher expression in bLRs compared to bHRs (NES<0). This module was highly-enriched for oligodendrocyte-specific gene expression, as indicated by overlap with the *BrainInABlender* database (64%, (45)). The large module, *M1* (1,225 genes), was also upregulated in bLRs in the P14 meta-analysis. It had

been previously identified in both healthy and diseased hippocampi in humans, as well as in mouse hippocampi, and showed an enrichment of pathways related to synaptic processing (27). This module was led by the Apba2 gene (β=-1.03, p=0.0015, FDR=0.58) which was recently shown to contain both a segregating variant in relationship to bHR/bLR phenotype and a strong QTL for locomotor behavior (22). This module also contained many genes with a strong relationship to bHR/bLR phenotype in adulthood (FDR<0.05, Fxyd7, Tmem2, Rltpr, Tubg1, Slc9a3r1).

Within the adult meta-analysis, the *paleturquoise* module (top gene: Adamts2) showed greater expression in bHRs (NES>0) and displayed a strong cell type association, with 34% of the genes in the module specifically expressed in either astrocytes or vasculature, as indicated by overlap with the *BrainInABlender* database (45). The *darkred* module (top gene: Dusp11), which showed greater expression in bLRs in the adult-meta-analysis, showed a nominal positive correlation with contextual fear immobility in a previous study (p=0.01; (28)).

Comparison of bHR vs. bLR DE Results with Hippocampal DE Results from Other Bred Rat Models

In general, we found that approximately one third (26/76) of our top bHR vs. bLR DE genes (FDR<0.05) had been previously identified as DE in other transcriptional profiling studies examining hippocampal gene expression in bred rat models targeting behavioral traits that could be considered extremes on the internalizing/externalizing spectrum ((33–41), **Fig 3, Fig S11**), and approximately 10% (8/76) had been previously identified by more than one study. Amongst these instances of overlap with our top bHR/bLR DE genes, the majority (20/35) showed a consistent direction of effect. Within these results, the gene Tmem144 stood out in particular, because it was the only gene identified as DE in three other studies in addition to our own and

the direction of effect was consistently up-regulated in bred rat models targeting behavioral traits related to internalizing behavior.

This overlap between our top DE genes and those from other bred rat models seemed intriguing, but it is difficult to fully interpret because each study used different statistical thresholds, sample sizes, and transcriptional profiling platforms. For example, unsurprisingly, we saw greater overlap between our top bHR/bLR DE genes and the top DE genes identified by studies profiling Flinders Sensitive Line (FSL) vs. Flinders Resistant Line (FRL) (34), FSL vs. Sprague-Dawley (35), and Wistar Kyoto (WKY) vs. Fischer 344 (F344) rats (36), which all identified relatively long lists of DE genes (908, 210, and 864 unique genes, respectively), whereas we saw less overlap with studies that identified relatively fewer (<50) DE genes (38–41). A comparison with the DE genes identified in studies profiling Wistar More Immobile (WMI) vs. Wistar Less Immobile rats (37) and NIH Heterogeneous Stock High Anxiety vs. Low Anxiety rats (33) stood out as a notable exception to that rule – each of these studies identified a relatively large number of DE genes (460 and 209 unique gene symbols, respectively), but showed almost no overlap with our top findings (1 and 0 genes, respectively), although some similar pathways were identified (*e.g.,* thyroid signaling in (33)). This lack of overlap could have arisen from technical factors that differentiate the studies or from genuine differences between our phenotypes, background strains, or experimental manipulations.

Overall, it seems extremely unlikely that we would observe this great of an overlap due to random chance alone, but strong conclusions cannot be drawn without access to the full results for all genes from each study to formally calculate enrichment. One study did fully release their raw data (Blaveri et al. (34)), therefore, as a test case, we reanalyzed their results using methodology similar to that used for the bHR/bLR studies and ran a full in depth comparison with our own findings. We found that 908 genes in the Blaveri et al. dataset were DE between the FSL and FRL lines (FDR<0.05), and this overlapped with 14 of our 76 top bHR/bLR DE genes (**Fig 3**).

Then, to formally calculate enrichment, for the sake of equivalency we only used the 74 genes that were DE in the adult bHR/bLR meta-analysis as comparison. We also filtered the results from both datasets to eliminate genes that were unlikely to be truly expressed in the adult rat hippocampus, but that were still included within both datasets due to representation on the microarray platforms because any non-expressed genes were likely to lack DE in both datasets, inflating our enrichment calculation. To filter out these genes, we eliminated genes that were not represented in our the bHR/bLR adult hippocampal RNA-Seq meta-analysis. This left differential results from 10,804 genes in each dataset, 787 of which were DE in FSL vs. FRL (7.3%), and 72 of which were DE in the adult bHR/bLR meta-analysis (0.67%). Thirteen of these genes were shared between the two datasets (18% of the adult bHR/bLR DE genes), representing a significant 2.84x enrichment in overlap over what would be expected by random chance (95% CI: 1.42-5.26, Fisher's Exact Test: p=0.00186).

Altogether, these results indicate that the DE genes identified by the bHR/bLR meta-analyses are likely to represent a rich collection of novel and previously-identified candidates that could provide broad insight into genetic predisposition for internalizing/externalizing behavior.

Supplemental Table Legends

See separate .gmt and .xls files for these tables

Table S1. For the Gene Set Enrichment Analyses: a .gmt containing 69 gene sets custom-designed to reflect hippocampal-specific functions.

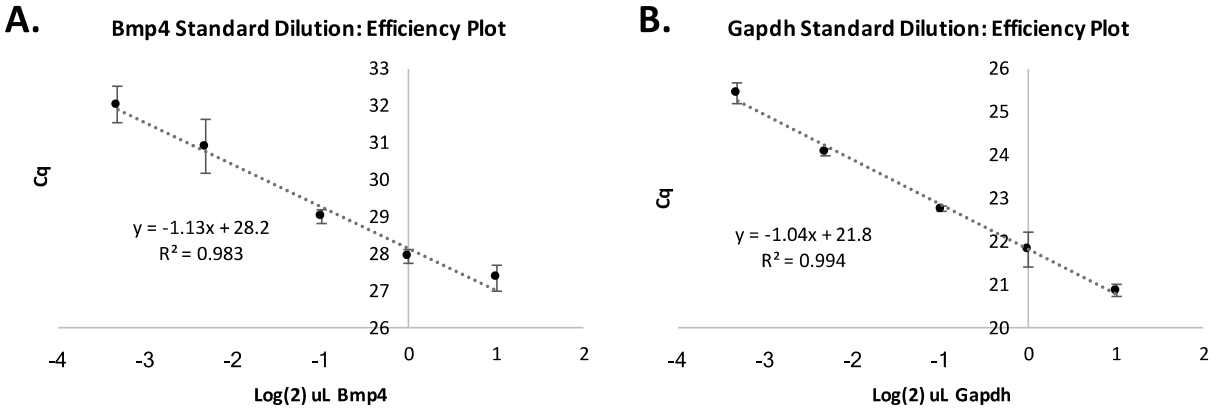
Table S2. The full meta-analysis results for each age group (adult, P7, P14, P21), as well as a meta-analysis performed using just the adult RNA-Seq data from the latest generations (F37 & F43).

Table S3. The full output from the Gene Set Enrichment Analyses conducted using the results from the adult and P14 meta-analyses, including a version of the adult meta-analysis conducted using just RNA-Seq data from the two latest generations (F37 and F43).

Table S4. **A database of DE genes identified by transcriptional profiling in 10 comparisons of bred rat models targeting behavioral traits that could be considered extremes on the internalizing/externalizing spectrum.** The provided data frame (.csv) identifies the direction of effect for all 2598 gene symbols that showed DE in at least one of 8 publications (using publication-specific criteria, (33,35–41)), in our reanalysis of Blaveri et al. ((29), FDR<0.05), or in our bHR/bLR meta-analysis (FDR<0.10). We assigned the direction of effect so that the bred animals that we considered to have the greatest internalizing behavior were the comparison group, so that a positive direction of effect (1) would signal relatively higher expression in this group and a negative direction of effect (-1) would signal relatively lower expression in this group. A small percentage of the DE genes showed a different direction of effect for probes representing the same gene symbol within their original dataset (0.8%, 24/2948), and thus ended up with a direction of effect that represented an average value falling between -1 and 1. Abbreviations:

WKY=Wistar Kyoto rats, F344=Fischer 344 rats, WMI=Wistar More Immobile, WLI=Wistar Less Immobile, FSL=Flinders Sensitive Line, FRL=Flinders Resistant Line, cLH=Congenitally Helpless rats, cNLH=Helpless Resistant rats, RLA=Roman Low Avoidance, RHA=Roman High Avoidance, SLA=Syracuse Low Avoidance, SHA=Syracuse High Avoidance, NIH-HS=National Institutes of Health Heterogeneous Stock.

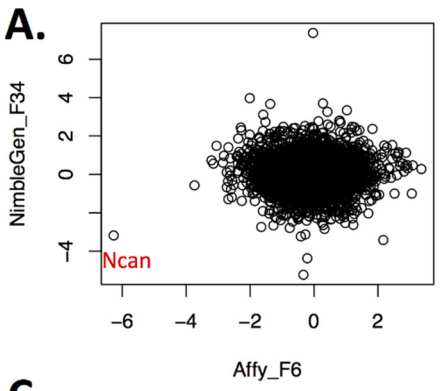
Supplementary Figures



C.

Postnatal Age	Generation	Sample Size	Gapdh Mean Cq	Bmp4 Mean Cq
P14	F55	bHR=6, bLR=6	18.6	24.4
P90	F51	bHR=6, bLR=6	16.6	24.3

Fig S1. Efficiency curves and methodological detail to accompany the qPCR validation results. A-B) The calibration curves using stock cDNA revealed efficiencies close to 1 for each probe (Bmp4: *ß*= -1.13, R^2=0.98, Gapdh: *ß*=-1.04 R^2=0.99), implying a doubling of PCR product with each cycle. C_q= the number of PCR cycles necessary to reach threshold. **C)** Overview of the samples used for the qPCR validation experiment. Note that the mean Cq for both the target and reference probes (Gapdh, Bmp4) was less than the Cq's measured within our calibration experiment, indicating a greater concentration for both transcripts in our samples than in the stock cDNA. Therefore, our estimates of the concentration for both transcripts should have safely occurred during the exponential phase of the PCR procedure, prior to depletion of PCR reagents (18).



B.

	Affy_F6	Affy_F15	Illumina_F15	RNASeq_F29	NimbleGen_F34
Affy_F6	1.00	0.00	0.02	0.17	-0.22
Affy_F15	0.00	1.00	-0.01	-0.02	-0.01
Illumina_F15	0.02	-0.01	1.00	0.06	0.03
RNASeq_F29	0.17	-0.02	0.06	1.00	-0.04
NimbleGen_F34	-0.22	-0.01	0.03	-0.04	1.00

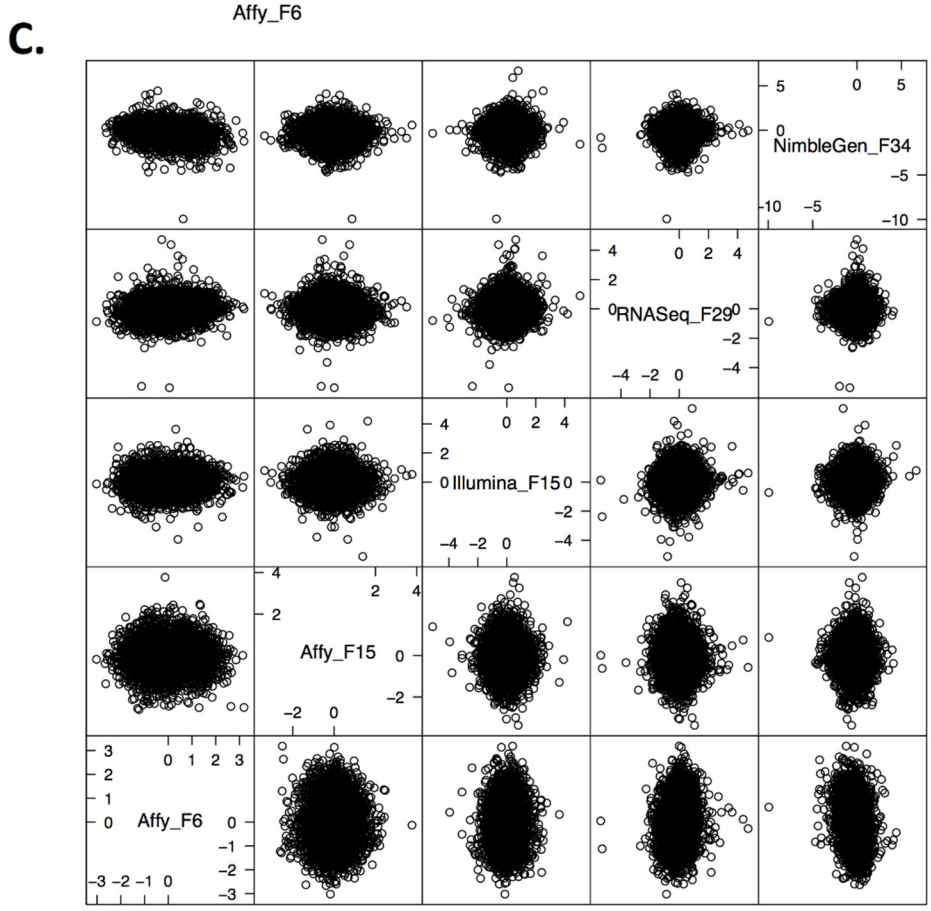


Fig S2. The results from small (n<6) developmental bHR/bLR transcriptomic studies are noisy and produce few reliable results when analyzed individually. A) A scatterplot illustrating the lack of correlation between the bHR vs. bLR effect sizes (Cohen's d) for all genes included in both P7 microarray studies. Neurocan (Ncan) is the only gene showing a strong effect in both studies. **B)** The pairwise correlation coefficients between each of the bHR/bLR P14 transcriptomic studies are weak (R<0.17). The correlation coefficients were calculated using the

bHR vs. bLR effect sizes for all genes present in each pair of datasets. **C)** Scatterplots illustrating the lack of correlation between the P14 transcriptomic studies. Note that the correlations are not noticeably better between studies with similar platforms or identical generations – most of the noise is likely to originate directly from the usage of small sample sizes.

A.

	Affy_F4	RNASeq_F29	Agilent_F34	RNASeq_F37	RNA_SeqF43
Affy_F4	1.00	-0.06	-0.05	0.11	0.03
RNASeq_F29	-0.06	1.00	0.05	0.00	0.11
NimbleGen_F34	-0.05	0.05	1.00	-0.01	0.08
RNASeq_F37	0.11	0.00	-0.01	1.00	0.40
RNA_SeqF43	0.03	0.11	0.08	0.40	1.00

B.

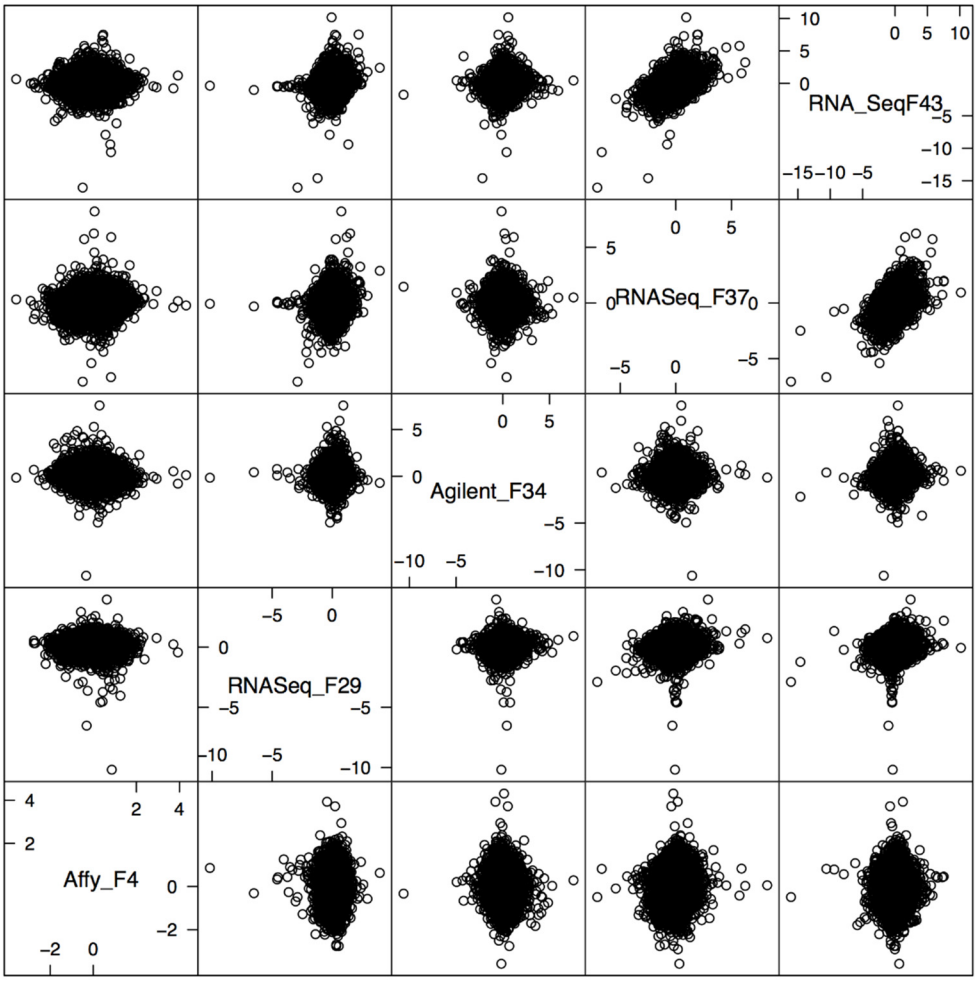


Fig S3. The results from small (n<6) adult bHR/bLR transcriptomic studies are noisy and produce few reliable results when analyzed individually until the latest generations of selective breeding. A) The pairwise correlation coefficients between each of the bHR/bLR adult

transcriptomic studies are weak (R<0.11) until the latest generations (F37 and F43, R=0.40). The correlation coefficients were calculated using the bHR vs. bLR effect sizes (Cohen's D) for all genes present in each pair of datasets. **B)** Scatterplots illustrating the lack of correlation between the adult transcriptomic studies until the latest generations.

A.

Age	Number of Genes in Meta-Analysis	Number of Datasets in Meta-Analysis	Number of Significant Genes (p-value < 0.05)	Number of Significant Genes after FDR (<0.05)
P7	3527	2	392	1
P14	15682	5	921	0
P21	3527	2	263	0
Adult	16269	5	1881	74

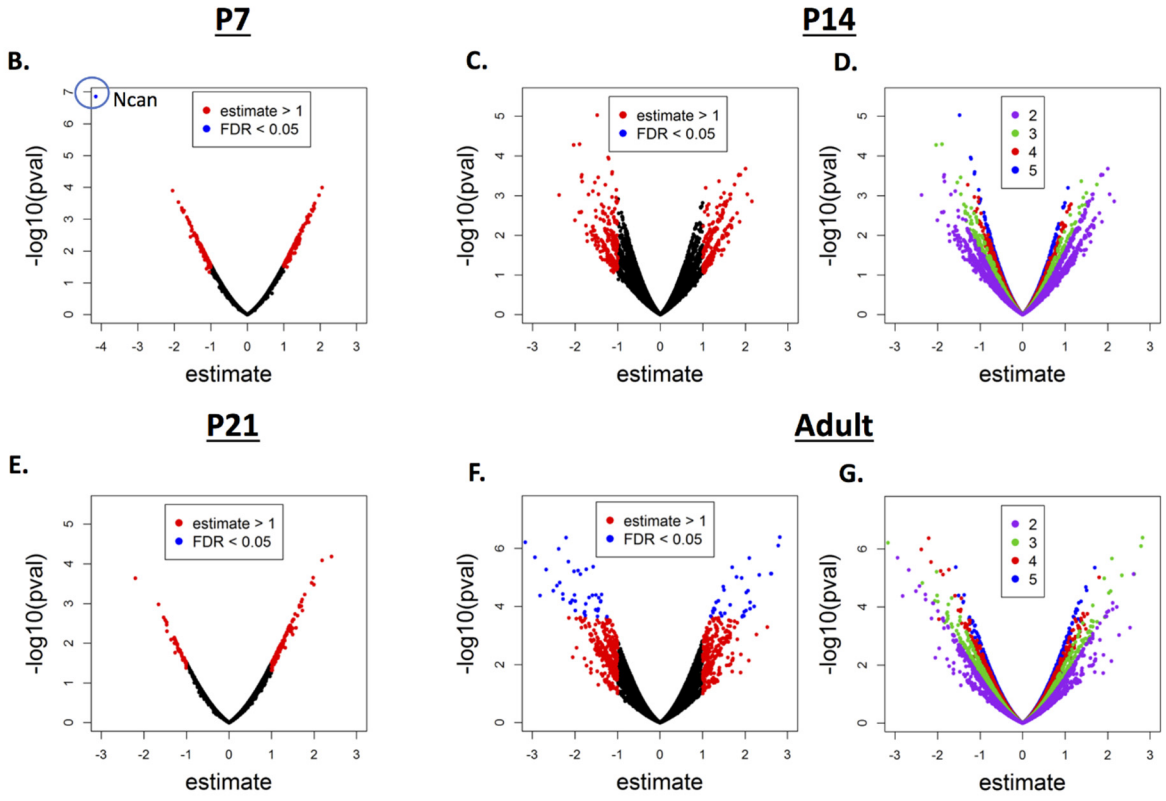


Fig S4. Meta-analysis uncovers hippocampal gene expression related to bHR/bLR phenotype: An overview. ***A)*** A table overviewing the results of the meta-analyses conducted for each of the four age groups. Genes were only included in the meta-analysis if they were represented in at least two datasets. As a result, the P14 and Adult meta-analyses contained a much larger number of genes because they contained five datasets, whereas the P7 and P21 meta-analyses each contained two datasets. The most DE genes were detected within the adult meta-analysis, which predominantly included data from later generations (F29-F43). **B-G.** Volcano plots illustrating the distribution of the results for all genes included in the meta-analysis for each age group. The x-axis is the estimated effect size for each gene (i.e., the difference in

expression between bHR and bLR rats in units of standard deviation, positive=higher expression in bHRs), the y-axis is the -log(10) nominal p-value from the meta-analysis (larger=more significant). These plots are colored to either illustrate: **B-C & E-F)** the number of genes in each meta-analysis with an estimated effect size greater than one (red) or with p-values surpassing false detection correction (FDR<0.05, blue), **D&G)** The number of datasets included in the meta-analysis for each gene (for the P14 and adult meta-analyses). We found that the estimated effect sizes tended to be more extreme for genes present in fewer datasets, most likely because those genes were often only represented in the RNA-Seq datasets from later generations.

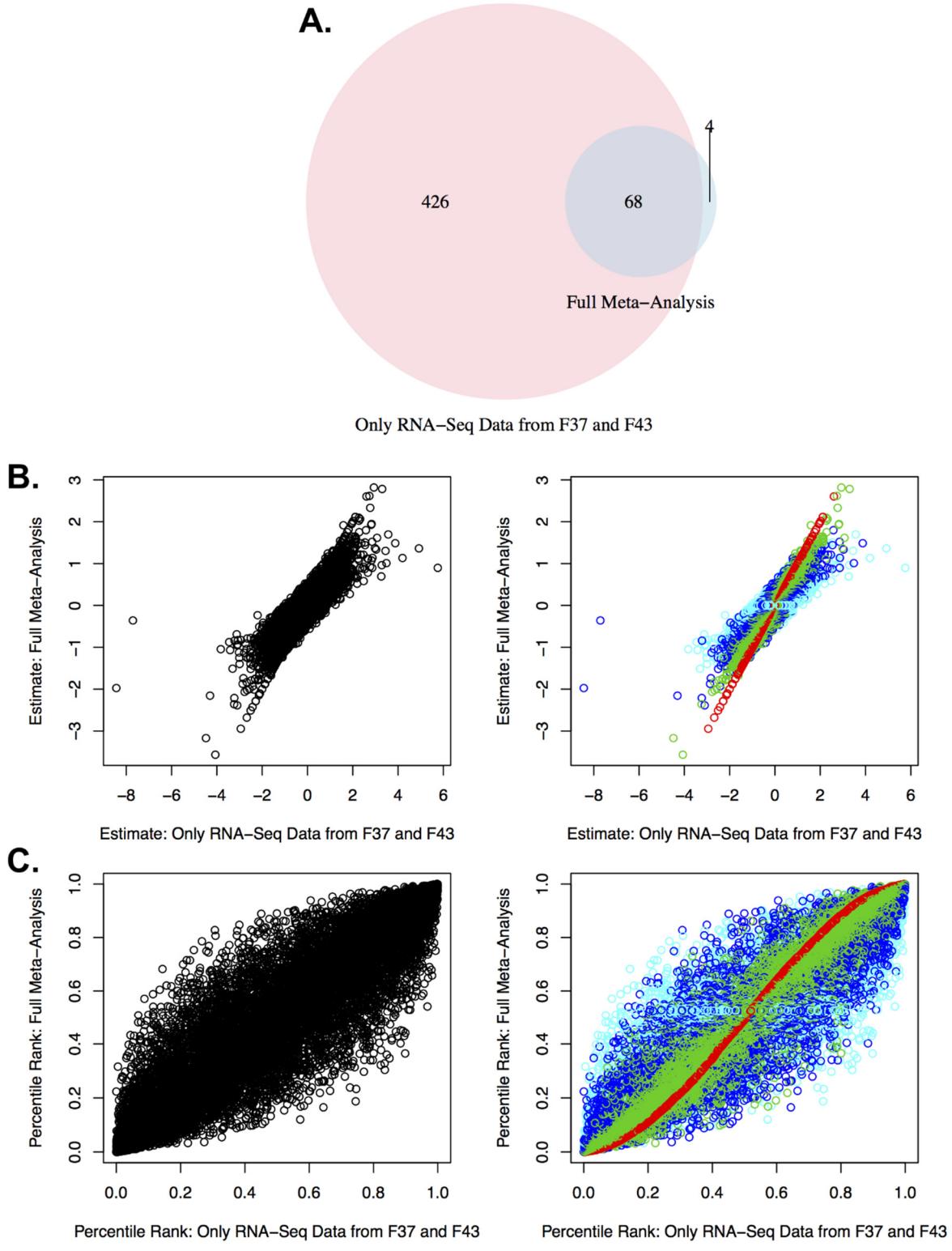


Fig S5. A meta-analysis of the adult RNA-Seq data from only the latest generations (F37

and F43) re-identifies almost all (68/74) of the top genes discovered in the full adult meta-analysis. A) A Venn diagram illustrating the overlap between the top genes (FDR<0.05) identified within the adult meta-analysis when using only the two RNA-Seq datasets from the most recent generations (F37 and F43) versus all datasets available for any particular gene. There are more genes identified as related to bHR/bLR phenotype when only considering the RNA-Seq data from the most recent generations. **B-C)** Scatterplots illustrating the strong correlation between the estimated effect sizes (B) and percentile ranks (C) produced by the two versions of the meta-analysis. The scatterplots on the right are colored to illustrate the number of datasets included in the full meta-analysis (turquoise=5, blue=4, green=3, red=2).

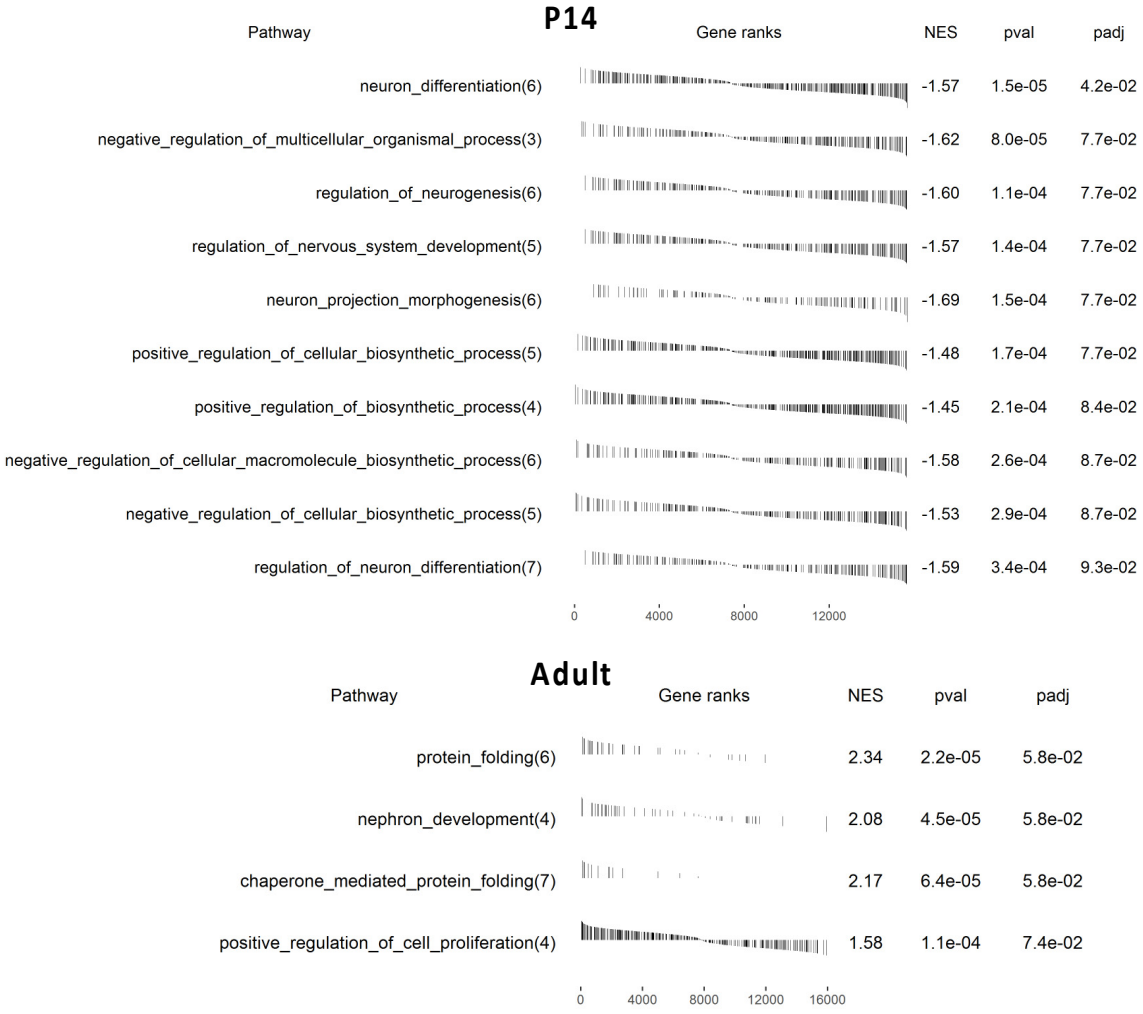


Fig S6. The top results (FDR<0.10) from gene set enrichment analysis for the P14 and adult meta-analysis performed using a database (.gmt) of traditional functional ontology pathways. The gene ranks were determined using the estimate from the meta-analysis (lowest rank=highly upregulated in bHRs (largest positive estimate), highest rank=highly upregulated in bLRs (largest negative estimate)) and the dark lines illustrate the rank and estimate for each gene in the respective gene set. NES=normalized enrichment score for the gene set (positive=upregulated in bHRs, negative=upregulated in bLRs), pval=nominal p-value, padj=false detection rate.

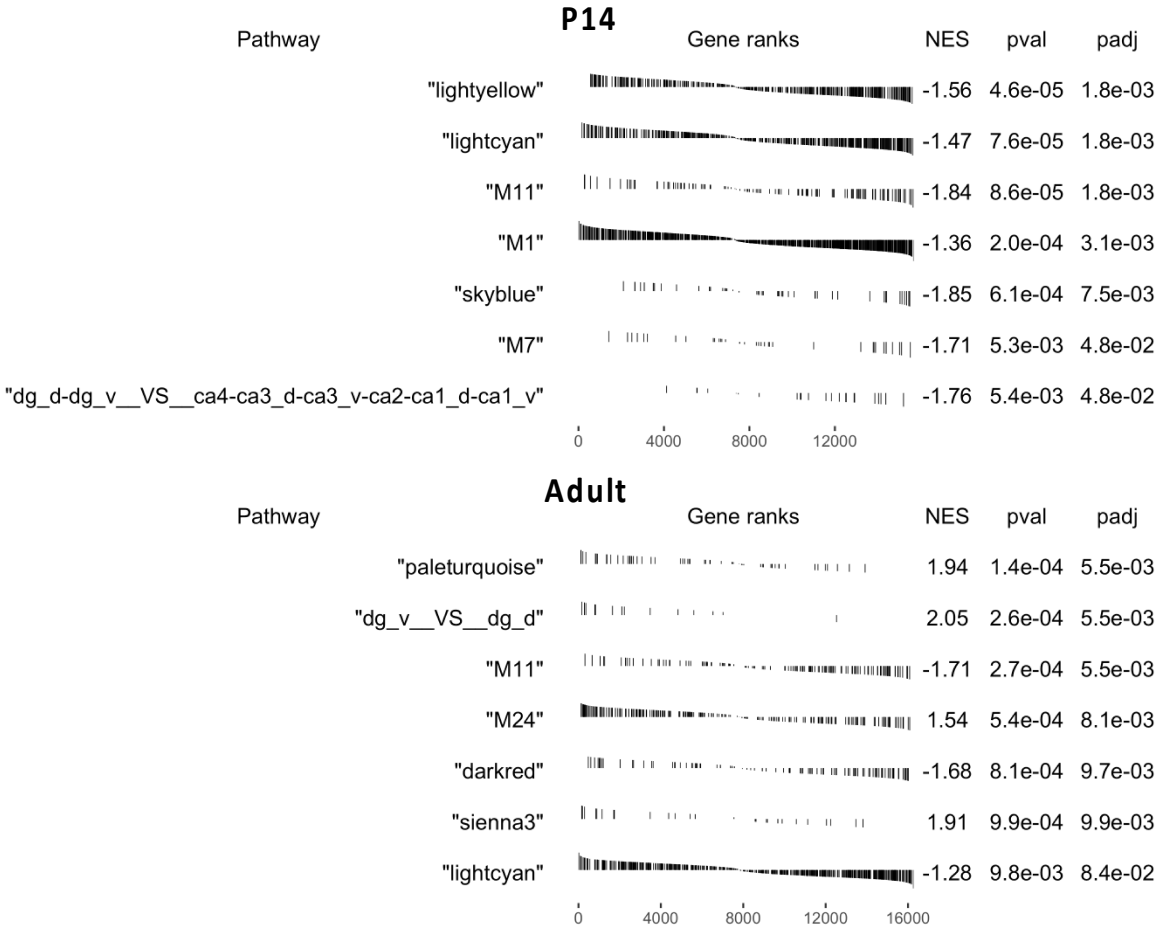


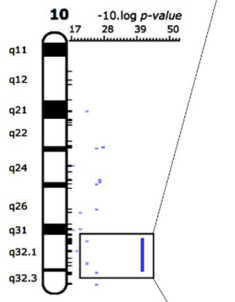
Fig S7. The top results (FDR<0.10) from gene set enrichment analysis for the P14 and adult meta-analysis performed using a database (.gmt) of hippocampal-specific co-expression networks and gene sets. Formatting is identical to Fig. S6. Gene sets named after colors represent co-expression networks previously identified in a large mouse study (28), whereas the gene sets identified with module numbers (M#) represent co-expression networks previously identified in a large human study (27). The other gene sets are derived from hippocampal regional comparisons of gene expression ((29); dg=dentate gyrus, d=dorsal, v=ventral, ca=Cornu Ammonis regions).

A. Genetic Variant within Etv4 Segregates bHR and bLR Rats

SNV within Etv4:
10_89485668 (Rnor5)

	TT	TC	CC
bHR	0	1	11
bLR	7	5	0

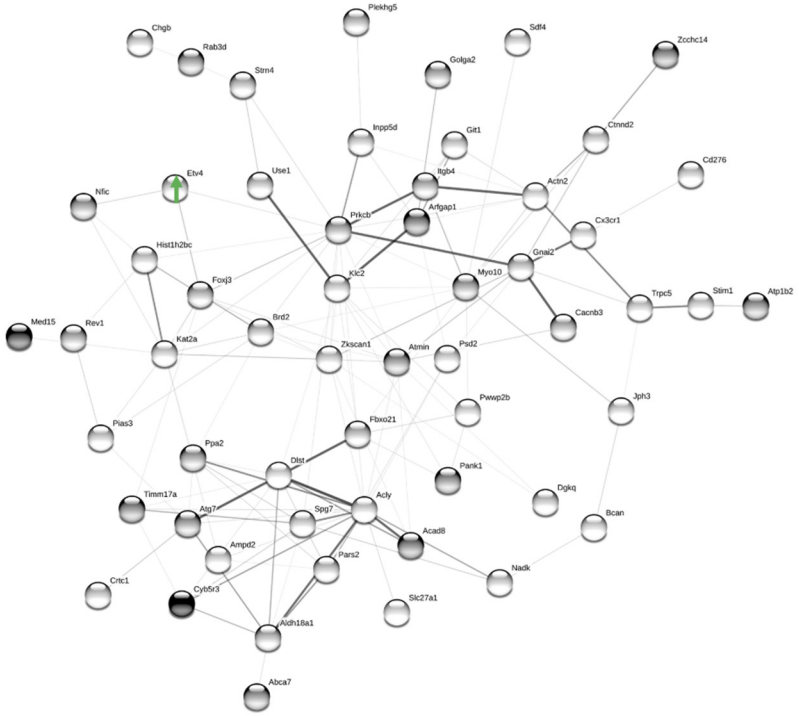
B. Etv4 is the Top Result in a Region of Chr 10 Enriched for Differentially Expressed Genes



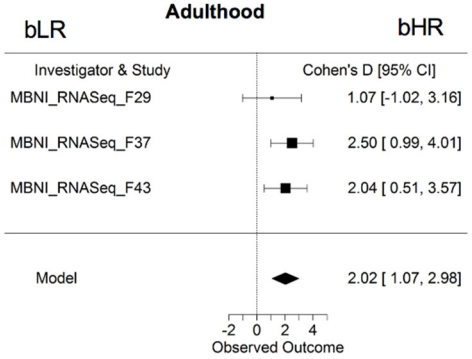
Chr10: 88,392,000-103,731,000

Gene Symbol	Adult_estimate	P14_estimate
Acly	-0.87	0.13
Kat2a	-1.27	NA
Ghdc	-1.97	0.60
Ptrf	1.54	-0.70
Tubg1	*-1.40*	-0.04
Cntnap1	-0.84	-0.22
Etv4	*2.02*	-0.11
Sost	0.64	**1.38**
Tmem101	1.31	-0.02
Rundc3a	0.22	**-1.30**
Acbd4	-1.19	0.56
Kcnh6	-0.92	0.09
Ddx42	-1.30	0.04
Cd79b	-0.09	**-1.04**
Arsg	-1.13	0.63
Kcnj16	0.82	0.02
Sox9	1.60	**-1.30**
Cd300lf	-0.96	-0.48
Slc9a3r1	*1.32*	-0.23

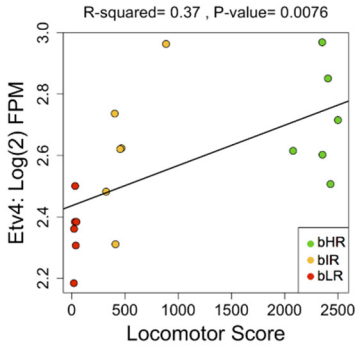
C. Etv4 is the Top Gene in a Hippocampal-Specific Co-expression Network Related to Neuronal Projections that is Enriched with Differentially Expressed Genes



D. Etv4 is More Highly Expressed in bHRs in Adulthood



E. Etv4 Correlates with Exploratory Activity



F. Etv4 Correlates with Anxiety

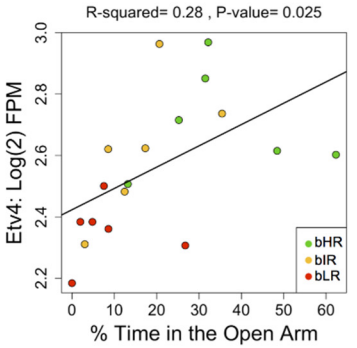


Fig S8. ETS Variant 4 (Etv4) was the top gene within a hippocampal specific gene set and within a genomic loci on Chromosome 10 enriched for bHR/bLR differential expression. A) Etv4 was the strongest result within a segment of Chromosome 10 enriched for differentially expressed genes that overlapped QTLs related to internalizing and externalizing behaviors (22,46,47). The table illustrates the differentially expressed genes within this region: estimate=estimated effect size (green/positive=more highly expressed in bHRs), bold=p<0.05, bold+italic=FDR<0.05. **B)** Etv4 was one of the top genes in a hippocampal specific co-expression network (*lightcyan*) that was enriched for bLR-upregulated genes in both development (FDR<0.05) and adulthood (FDR<0.10). Illustrated above is a PPI network (STRINGdb: confidence setting=0.40) constructed using the genes from this co-expression module that showed bHR/bLR differential expression in adulthood (n=74, p<0.05). This network was enriched with genes related to cell projections, neurons, synapses, and cation binding (FDR<0.05). **C)** A forest plot showing that Etv4 showed consistently higher expression in bHR rats (3 datasets, boxes=Cohen's D from each study +/-95% confidence interval, "Model"=estimated effect size +/-95% confidence intervals provided by the meta-analysis, β=2.02, p=3.30E-05, FDR= 1.92E-02). **D-E)** In the behavioral data accompanying the MBNI_RNASeq_F37 dataset, Etv4 (units: log(2) fragments per million (FPM)) showed a positive relationship with **D)** exploratory locomotor activity (β=0.000131, R^2= 0.37, p=0.0076) **E)** percent of time spent in the open arms of the elevated plus maze (β=0.00694, R^2= 0.28, p=0.0249).

Gene Symbol	Selectivity to Neurons	Description of Enrichment Pattern In Hippocampus (HC)	Top HC Cell Type	Top HC Cell Type: Mousebrain.org Description
Rltpr	52	Neurons	TEGLU23	Excitatory neurons, hippocampus CA3
Egfem1	21	Neurons	TEINH12	Non-border Cck interneurons, cortex/hippocampus
Fn3k	2	Neurons	TEINH10	R-LM border Cck interneurons, cortex/hippocampus
Zfp821	4	Neurons	TEGLU7	Excitatory neurons, cerebral cortex
Ddx42	2	Neurons	TEGLU23	Excitatory neurons, hippocampus CA3
Rabif	2	Neurons	TEINH11	R-LM border Cck interneurons, cortex/hippocampus
Chrm3	110	Neurons	TEGLU5	Excitatory neurons, cerebral cortex; Probable location: Entorhinal superficial layers
Zfp612	5	Neurons	TEGLU14	Excitatory neurons, cerebral cortex; Probable location: Subiculum
Tubg1	2	Neurons & neuroblasts	TEINH10	R-LM border Cck interneurons, cortex/hippocampus
Pkib	12	Neurons, esp. excitatory	TEGLU9	Excitatory neurons, cerebral cortex; Probable location: Cingulate/Retrosplenial area, layer 5
Fxyd7	6	Neurons, esp. excitatory	TEGLU23	Excitatory neurons, hippocampus CA3
R3hdm4	6	Neurons, esp. excitatory	TEGLU7	Excitatory neurons, cerebral cortex; Probable location: Cortical pyramidal layer 2/3
Fbxo31	4	Neurons, esp. excitatory	TEGLU7	Excitatory neurons, cerebral cortex; Probable location: Cortical pyramidal layer 2/3
Shf	7	Neurons, esp. excitatory	TEGLU5	Excitatory neurons, cerebral cortex; Probable location: Entorhinal superficial layers
Medag	3	Neurons, esp. excitatory	TEGLU12	Excitatory neurons, cerebral cortex; Probable location: Lateral cortex layer 6: gustatory, barrel field, auditory
Cnksr2	9	Neurons, esp. excitatory	TEGLU23	Excitatory neurons, hippocampus CA3
Prss12	77	Neurons, esp. excitatory	TEGLU20	Excitatory neurons, cerebral cortex; Probable location: Cortex pyramidal layer 6
Ddn	28	Neurons, esp. excitatory	TEGLU14	Excitatory neurons, cerebral cortex; Probable location: Subiculum
Amer2	3	Neurons, esp. inhibitory	TEINH8	Interneuron-selective interneurons, hippocampus
Mettl16	3	Neurons, esp. inhibitory	TEINH12	Non-border Cck interneurons, cortex/hippocampus
Tmem59l	20	Neurons, esp. inhibitory	TEINH7	Interneuron-selective interneurons, hippocampus
Fkbp4	2	Neurons, esp. inhibitory	TEINH10	R-LM border Cck interneurons, cortex/hippocampus
Cct4	2	Neurons, esp. inhibitory & neuroblasts	TEINH7	Interneuron-selective interneurons, hippocampus
Gal	83	Neurons, granule neuroblasts	DGNBL2	Granule neuroblasts, dentate gyrus
Tnnt1	11	Neurons, neurogliaform	TEINH15	CGE-derived neurogliaform cells, cortex/hippocampus

Fig S9. Many of the top differentially expressed genes are predominantly expressed in particular hippocampal cell types, including neurons. A comparison of our top meta-analysis results (FDR<0.10) with the new mousebrain.org database (Ziesel et al. 2018) suggested that many of our top genes (31%: 59/191 included in the analysis) have enriched (>2X) expression in a hippocampal cell type as compared to other hippocampal cell types. Shown above are the top bHR/bLR DE genes that have enriched expression in hippocampal neuronal cell types, which could not be properly interrogated within our cell type deconvolution analysis due to the dependence of *BrainInABlender* on cortically-derived datasets. The gene symbols are formatted to illustrate the results from our

bHR/bLR meta-analysis (green=more highly expressed in bHRs, red=more highly expressed in bLRs, bold=FDR<0.05). “Selectivity to neurons” provides the ratio of the average expression value within the top hippocampal neuronal cell type vs. the top non-neuronal cell type. The hippocampal cell type with the most enriched expression is identified using the cluster name and description provided by mousebrain.org.

Rat Line with Greater Internalizing Behavior	Reference Group	Citation	Source of Top Gene List
Flinders Sensitive Line (FSL)	Flinders Resistant Line (FRL)	Blaveri et al. 2010	*We reanalyzed the raw data (GSE20388)*
Flinders Sensitive Line (FSL)	Sprague-Dawley (SD)	Wilhelm et al. 2013	Table S2
Wistar Kyoto (WKY)	Fischer 344 (F344)	Meckes et al. 2018	Table S3
Wistar More Immobile (WMI) - Males	Wistar Less Immobile (WLI) - Males	Andrus et al. 2012	Table S4
Wistar More Immobile (WMI) - Females	Wistar Less Immobile (WLI) - Females	Raghavan et al. 2017	Table 1, Table 4
Congenitally Learned Helpless (cLH)	Congenitally No Learned Helpless (cNLH)	Garafola and Henn 2014	Table 1
Roman Low Avoidance (RLA)	Roman High Avoidance (RHA)	Sabariego et al. 2013	Table 1
Syracuse Low Avoidance (SLA)	Syracuse High Avoidance (SHA)	Zhang et al. 2005	Table S1
NIH Heterogeneous Stock: High Anxiety (NIH-HS HA)	NIH Heterogeneous Stock: Low Anxiety (NIH-HS LA)	Diaz-Moran et al. 2013	Table S1 and Table S2

Figure S10. Nine previous publications compared gene expression using transcriptional profiling in macro-dissected hippocampal samples from bred rat models targeting behavioral traits that could be considered to represent extremes on the internalizing/externalizing spectrum (33–41)**.** The results from these studies were used in an overlap analysis to determine which of our bHR/bLR DE genes were likely to be associated with a genetic predisposition for internalizing and externalizing behavior.

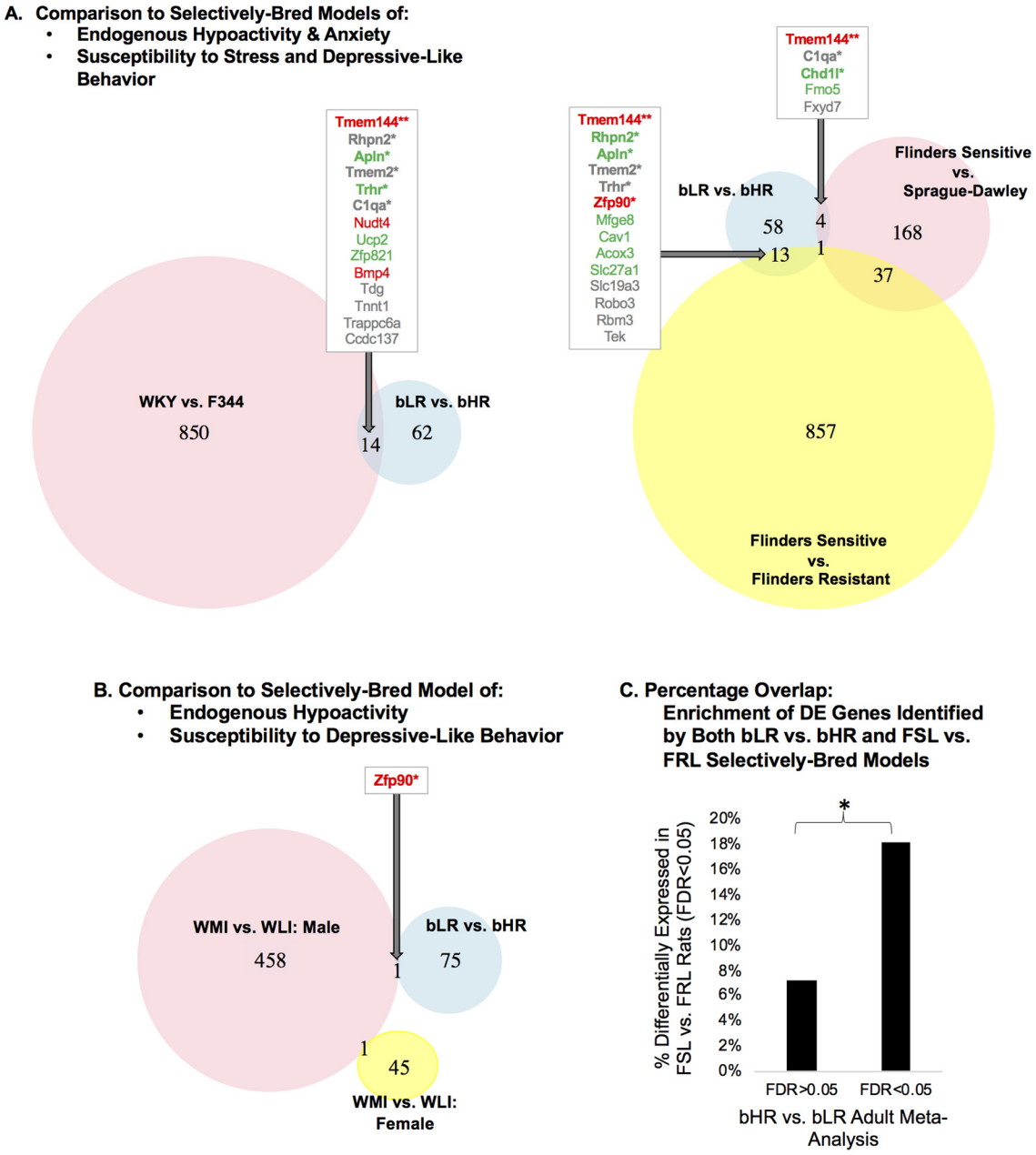


Figure S11. A high percentage of the top bHR vs. bLR DE genes are also DE in the hippocampus of other bred models targeting behavioral traits that represent extremes on the internalizing/externalizing spectrum. In general, we found that approximately one third (26/76) of our top bHR vs. bLR DE genes (FDR<0.05) had been previously identified as DE in at least one of the nine transcriptional profiling studies that we reviewed characterizing hippocampal

gene expression in other bred rat models targeting behavioral traits that could be considered to represent extremes on the internalizing/externalizing spectrum. **A-B)** Venn Diagrams illustrating the overlap between the top DE genes in bHR vs. bLR rats (full list in Fig 3) and DE genes identified in five other bred model comparisons. The genes present in each overlap are listed above the intersection, with red text indicating greater expression in both bLRs and in the rat strain/line that we considered to have the greatest internalizing behavior, green text indicating greater expression in both bHRs and in the rat strain/line that we considered to have greater externalizing behavior, and grey indicating that the direction of effect differed between our model comparison and the other model comparison shown. Asterisks demark the number of model comparisons in addition to our own identifying a gene as DE (*=2, **=3). Abbreviations: WKY=Wistar Kyoto rats, F344=Fischer 344 rats, WMI=Wistar More Immobile, WLI=Wistar Less Immobile. Four other bred model comparisons are not depicted due to little overlap with the DE genes identified in bHR vs. bLR rats: *1 gene overlap:* Congenitally Helpless (cLH) and Helpless Resistant (cNLH) rats (overlapping gene: Chd1l); *no overlap:* Roman Low Avoidance vs. Roman High Avoidance, Syracuse Low Avoidance vs. Syracuse High Avoidance, National Institutes of Health Heterogeneous Stock (NIH-HS) High Anxiety vs. NIH-HS Low Anxiety. **C)** One study comparing FSL vs. FRL rats released raw data (34), so we were able to re-analyze their results and perform a full enrichment analysis. We found that hippocampal-expressed genes that were DE in the bHR vs. bLR adult meta-analysis (FDR<0.05) were 2.84X more likely to be DE in FSL vs. FRL rats (Fisher's Exact Test: p=0.00186).

Supplemental References

1. Stead JDH, Clinton S, Neal C, Schneider J, Jama A, Miller S, *et al.* (2006): Selective breeding for divergence in novelty-seeking traits: Heritability and enrichment in spontaneous anxiety-related behaviors. *Behav Genet* 36: 697–712.

2. Gautier L, Cope L, Bolstad BM, Irizarry RA (2004): affy--analysis of Affymetrix GeneChip data at the probe level. *Bioinformatics* 20: 307–315.

3. Irizarry RA, Hobbs B, Collin F, Beazer-Barclay YD, Antonellis KJ, Scherf U, Speed TP (2003): Exploration, normalization, and summaries of high density oligonucleotide array probe level data. *Biostatistics* 4: 249–264.

4. Dai M, Wang P, Boyd AD, Kostov G, Athey B, Jones EG, *et al.* (2005): Evolving gene/transcript definitions significantly alter the interpretation of GeneChip data. *Nucleic Acids Res* 33: e175.

5. Carlson M (2017): *Org.Rn.Eg.Db: Genome Wide Annotation for Rat*, version 3.5.0. Retrieved May 23, 2018, from http://bioconductor.org/packages/org.Rn.eg.db/

6. Clinton SM, Stead JDH, Miller S, Watson SJ, Akil H (2011): Developmental underpinnings of differences in rodent novelty-seeking and emotional reactivity. *Eur J Neurosci* 34: 994–1005.

7. Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, Shi W, Smyth GK (2015): limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res* 43: e47.

8. Bolstad B (2016): *PreprocessCore: A Collection of Pre-Processing Functions*, version 1.40.0. Retrieved May 23, 2018, from https://github.com/bmbolstad/preprocessCore

9. Kim D, Pertea G, Trapnell C, Pimentel H, Kelley R, Salzberg SL (2013): TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. *Genome Biol* 14: R36.

10. Langmead B, Salzberg SL (2012): Fast gapped-read alignment with Bowtie 2. *Nat Methods* 9: 357–359.

11. Trapnell C, Roberts A, Goff L, Pertea G, Kim D, Kelley DR, *et al.* (2012): Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. *Nat Protoc* 7: 562–578.

12. McCoy CR, Golf SR, Melendez-Ferro M, Perez-Costas E, Glover ME, Jackson NL, *et al.* (2016): Altered metabolic activity in the developing brain of rats predisposed to high versus low depression-like behavior. *Neuroscience* 324: 469–484.

13. Aurbach EL, Inui EG, Turner CA, Hagenauer MH, Prater KE, Li JZ, *et al.* (2015): Fibroblast growth factor 9 is a novel modulator of negative affect. *Proc Natl Acad Sci USA* 112: 11953–11958.

14. Liao Y, Smyth GK, Shi W (2013): The Subread aligner: fast, accurate and scalable read mapping by seed-and-vote. *Nucleic Acids Res* 41: e108.

15. Liao Y, Smyth GK, Shi W (2014): featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics* 30: 923–930.

16. Aydin C, Frohmader K, Akil H (2015): Revealing a latent variable: individual differences in affective response to repeated injections. *Behav Neurosci* 129: 679–682.

17. Livak KJ, Schmittgen TD (2001): Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. *Methods* 25: 402–408.

18. Yuan JS, Reed A, Chen F, Stewart CN (2006): Statistical analysis of real-time PCR data. *BMC Bioinformatics* 7: 85.

19. Clinton SM, Watson SJ, Akil H (2014): High novelty-seeking rats are resilient to negative physiological effects of the early life stress. *Stress* 17: 97–107.

20. Re AD (2014): *Compute.Es: Compute Effect Sizes*, version 0.2-2. Retrieved May 23, 2018, from https://CRAN.R-project.org/package=compute.es

21. Viechtbauer W (2010): Conducting Meta-Analyses in R with The metafor Package. *Journal of Statistical Software* 36. https://doi.org/10.18637/jss.v036.i03

22. Zhou Z, Blandino P, Yuan Q, Shen P-H, Hodgkinson CA, Virkkunen M, *et al.* (2019): Exploratory locomotion, a predictor of addiction vulnerability, is oligogenic in rats selected for this phenotype. *Proc Natl Acad Sci USA* 116: 13107–13115.

23. Shimoyama M, De Pons J, Hayman GT, Laulederkind SJF, Liu W, Nigam R, *et al.* (2015): The Rat Genome Database 2015: genomic, phenotypic and environmental variations and disease. *Nucleic Acids Res* 43: D743-750.

24. De Preter K, Barriot R, Speleman F, Vandesompele J, Moreau Y (2008): Positional gene enrichment analysis of gene sets for high-resolution identification of overrepresented chromosomal regions. *Nucleic Acids Res* 36: e43.

25. Sergushichev A (2016): An algorithm for fast preranked gene set enrichment analysis using cumulative statistic calculation. *bioRxiv*. https://doi.org/10.1101/060012

26. Powell JAC (2014): GO2MSIG, an automated GO based multi-species gene set generator for gene set enrichment analysis. *BMC Bioinformatics* 15: 146.

27. Johnson MR, Shkura K, Langley SR, Delahaye-Duriez A, Srivastava P, Hill WD, *et al.* (2016): Systems genetics identifies a convergent gene network for cognition and neurodevelopmental disease. *Nat Neurosci* 19: 223–232.

28. Park CC, Gale GD, de Jong S, Ghazalpour A, Bennett BJ, Farber CR, *et al.* (2011): Gene networks associated with conditional fear in mice identified using a systems genetics approach. *BMC Syst Biol* 5: 43.

29. Cembrowski MS, Wang L, Sugino K, Shields BC, Spruston N (2016): Hipposeq: a comprehensive RNA-seq database of gene expression in hippocampal principal neurons. *Elife* 5: e14997.

30. Zeisel A, Hochgerner H, Lönnerberg P, Johnsson A, Memic F, van der Zwan J, *et al.* (2018): Molecular Architecture of the Mouse Nervous System. *Cell* 174: 999-1014.e22.

31. Wegener G, Mathe AA, Neumann ID (2012): Selectively bred rodents as models of depression and anxiety. *Curr Top Behav Neurosci* 12: 139–187.

32. Hayward A, Tomlinson A, Neill JC (2016): Low attentive and high impulsive rats: A translational animal model of ADHD and disorders of attention and impulse control. *Pharmacol Ther* 158: 41–51.

33. Díaz-Morán S, Palència M, Mont-Cardona C, Cañete T, Blázquez G, Martínez-Membrives E, *et al.* (2013): Gene expression in hippocampus as a function of differential trait anxiety levels in genetically heterogeneous NIH-HS rats. *Behav Brain Res* 257: 129–139.

34. Blaveri E, Kelly F, Mallei A, Harris K, Taylor A, Reid J, *et al.* (2010): Expression profiling of a genetic animal model of depression reveals novel molecular pathways underlying depressive-like behaviours. *PLoS ONE* 5: e12596.

35. Wilhelm CJ, Choi D, Huckans M, Manthe L, Loftis JM (2013): Adipocytokine signaling is altered in Flinders sensitive line rats, and adiponectin correlates in humans with some symptoms of depression. *Pharmacol Biochem Behav* 103: 643–651.

36. Meckes JK, Lim PH, Wert SL, Luo W, Gacek SA, Platt D, *et al.* (2018): Brain region-specific expression of genes mapped within quantitative trait loci for behavioral responsiveness to acute stress in Fisher 344 and Wistar Kyoto male rats. *PLoS ONE* 13: e0194293.

37. Andrus BM, Blizinsky K, Vedell PT, Dennis K, Shukla PK, Schaffer DJ, *et al.* (2012): Gene expression patterns in the hippocampus and amygdala of endogenous depression and chronic stress models. *Mol Psychiatry* 17: 49–61.

38. Raghavan NS, Chen H, Schipma M, Luo W, Chung S, Wang L, Redei EE (2017): Prepubertal Ovariectomy Exaggerates Adult Affective Behaviors and Alters the Hippocampal Transcriptome in a Genetic Rat Model of Depression. *Front Endocrinol (Lausanne)* 8: 373.

39. Garafola CS, Henn FA (2014): A change in hippocampal protocadherin gamma expression in a learned helpless rat. *Brain Res* 1593: 55–64.

40. Sabariego M, Morón I, Gómez MJ, Donaire R, Tobeña A, Fernández-Teruel A, *et al.* (2013): Incentive loss and hippocampal gene expression in inbred Roman high- (RHA-I) and Roman low- (RLA-I) avoidance rats. *Behav Brain Res* 257: 62–70.

41. Zhang S, Amstein T, Shen J, Brush FR, Gershenfeld HK (2005): Molecular correlates of emotional learning using genetically selected rat lines. *Genes Brain Behav* 4: 99–109.

42. Davis S, Meltzer PS (2007): GEOquery: a bridge between the Gene Expression Omnibus (GEO) and BioConductor. *Bioinformatics* 23: 1846–1847.

43. Pollard KS, Dudoit S, Laan MJ van der (2005): Multiple Testing Procedures: the multtest Package and Applications to Genomics. *Bioinformatics and Computational Biology Solutions Using R and Bioconductor*. Springer, New York, NY, pp 249–271.

44. Oldham MC, Konopka G, Iwamoto K, Langfelder P, Kato T, Horvath S, Geschwind DH (2008): Functional organization of the transcriptome in human brain. *Nat Neurosci* 11: 1271–1282.

45. Hagenauer MH, Schulmann A, Li JZ, Vawter MP, Walsh DM, Thompson RC, *et al.* (2018): Inference of cell type content from human brain transcriptomic datasets illuminates the effects of age, manner of death, dissection, and psychiatric diagnosis. *PLoS ONE* 13: e0200003.

46. Klimes I, Weston K, Gasperíková D, Kovács P, Kvetnanský R, Jezová D, *et al.* (2005): Mapping of genetic determinants of the sympathoneural response to stress. *Physiol Genomics* 20: 183–187.

47. Potenza MN, Brodkin ES, Joe B, Luo X, Remmers EF, Wilder RL, *et al.* (2004): Genomic regions controlling corticosterone levels in rats. *Biol Psychiatry* 55: 634–641.