Supplementary Information

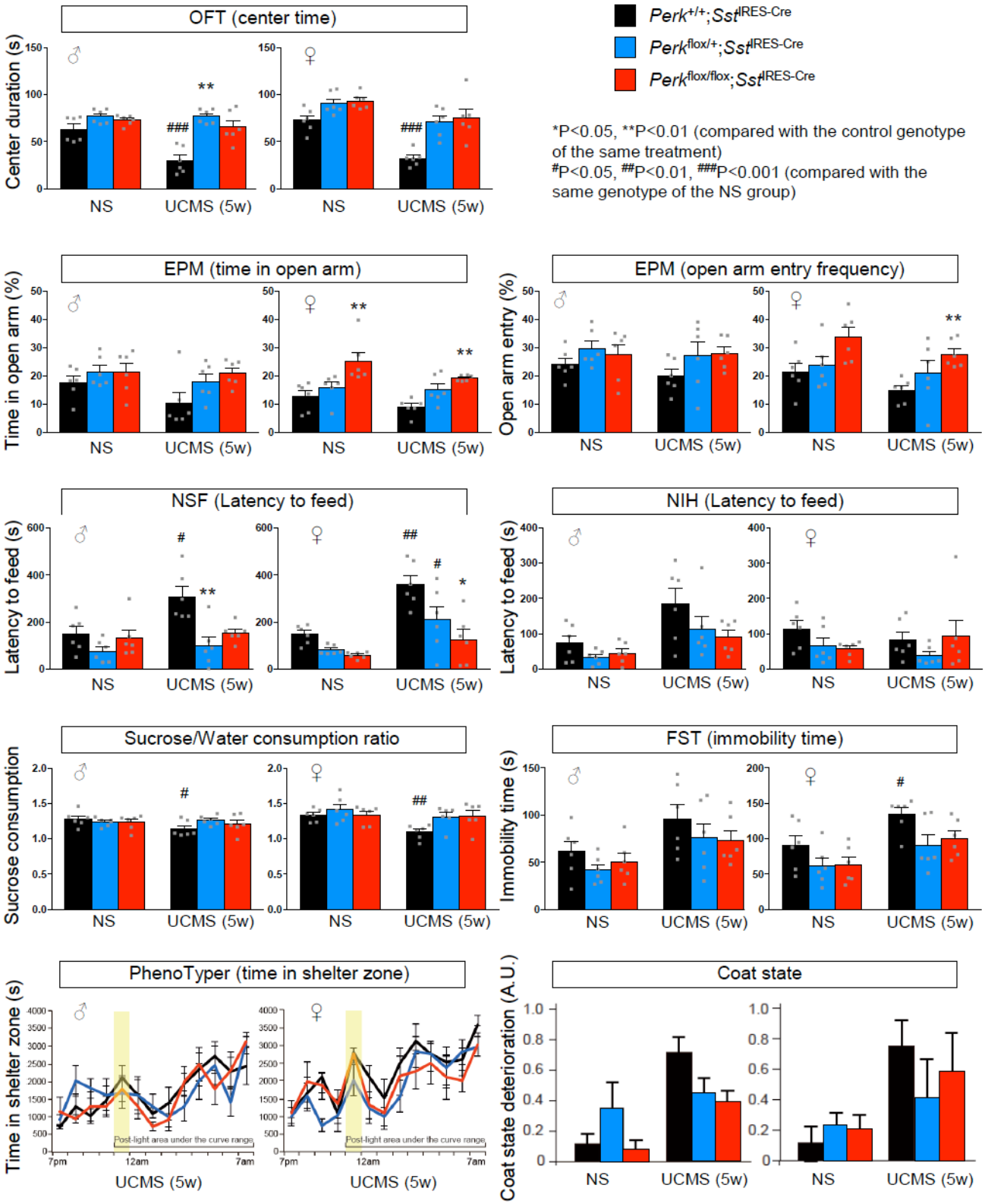
for

Molecular origin of somatostatin-positive neuron vulnerability

by

Toshifumi Tomoda, Akiko Sumitomo, Dwight Newton, and Etienne Sibille

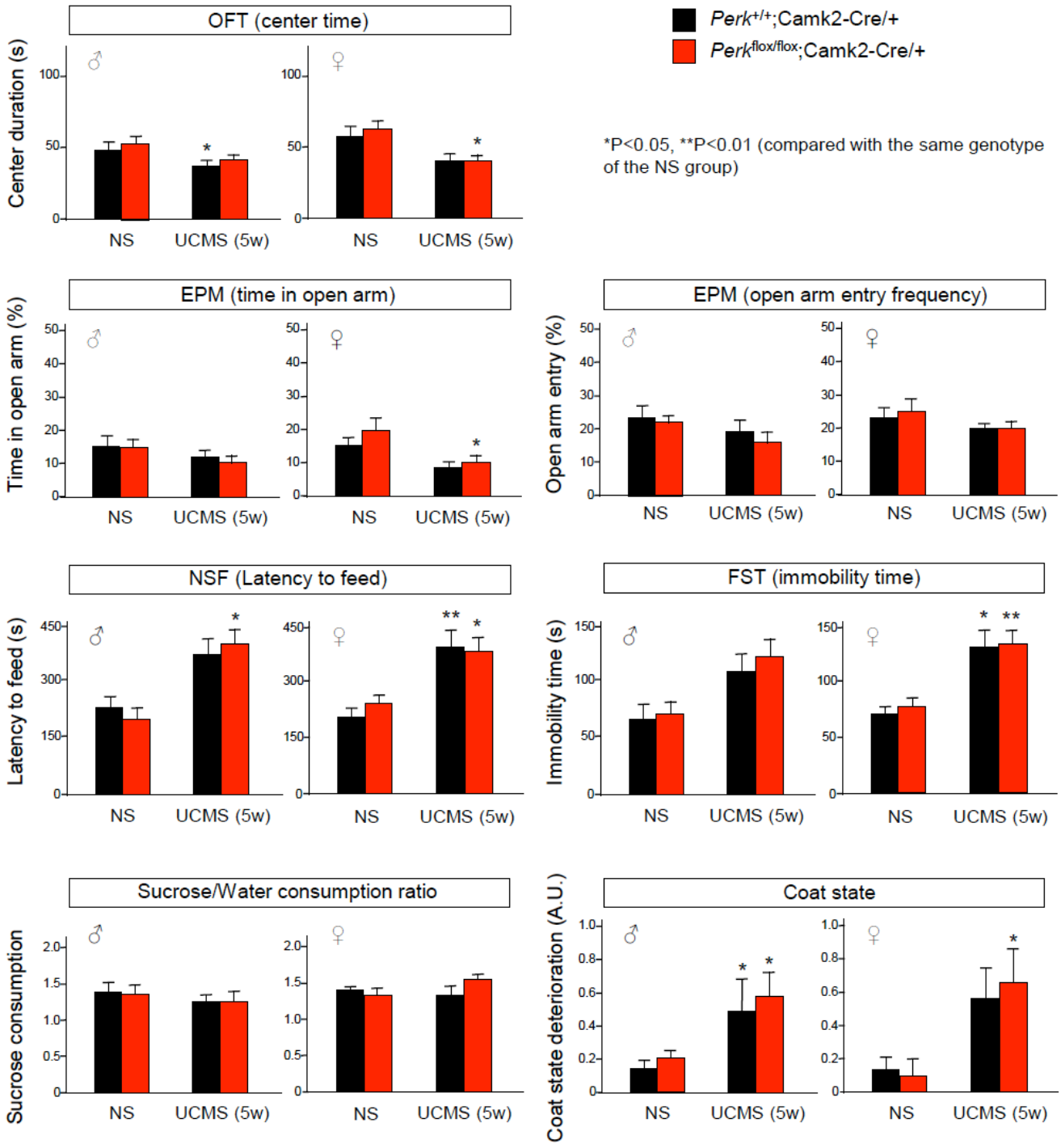
Supplementary Fig. 1



Supplementary Fig. 1. Genetic suppression of ER stress in SST^{+} neurons ameliorates UCMS-induced behavioral emotionality in mice.

Mice (6♂+6♀/genotype, 3–4 months old) were subjected to 5 weeks of UCMS or kept under no stress (NS) conditions, and assayed for open field test (OFT), elevated plus maze (EPM), PhenoTyper (PT), novelty-suppressed feeding (NSF), novelty-induced hypophagia (NIH), sucrose consumption (SC), forced-swim test (FST) and coat state. Statistical significance was evaluated by two-way ANOVA with repeated measures, followed by Bonferroni post-hoc tests; *P<0.05, **P<0.01 (as compared to the control genotype (*Perk*$^{+/+}$;*Sst*$^{IRES\text{-}Cre/+}$) of the same treatment group (i.e., either NS or 5 weeks of UCMS)); $^{\#}$P<0.05, $^{\#\#}$P<0.01, $^{\#\#\#}$P<0.001 (as compared with the same genotype of the NS group).

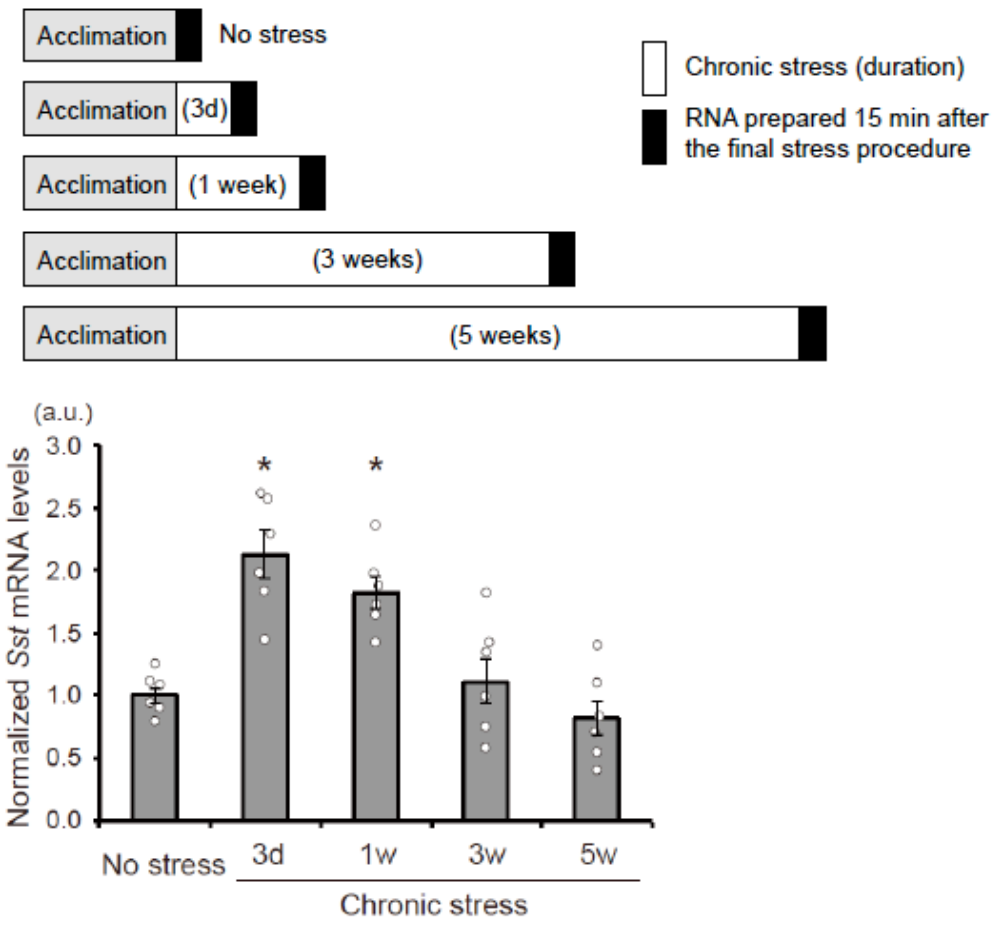
Supplementary Fig. 2



Supplementary Fig. 2. No rescue of UCMS-induced behavioral emotionality in mice with genetic suppression of ER stress in CaMKII^{+} pyramidal neurons.

Mice (6♂+6♀/genotype, 3–4 months old) were subjected to 5 weeks of UCMS or kept under no stress (NS) conditions, and assayed for open field test (OFT), elevated plus maze (EPM), novelty-suppressed feeding (NSF), sucrose consumption (SC), forced-swim test (FST) and coat state. Statistical significance was evaluated by two-way ANOVA with repeated measures, followed by Bonferroni post-hoc tests; *P<0.05, **P<0.01 (as compared with the same genotype of the NS group).

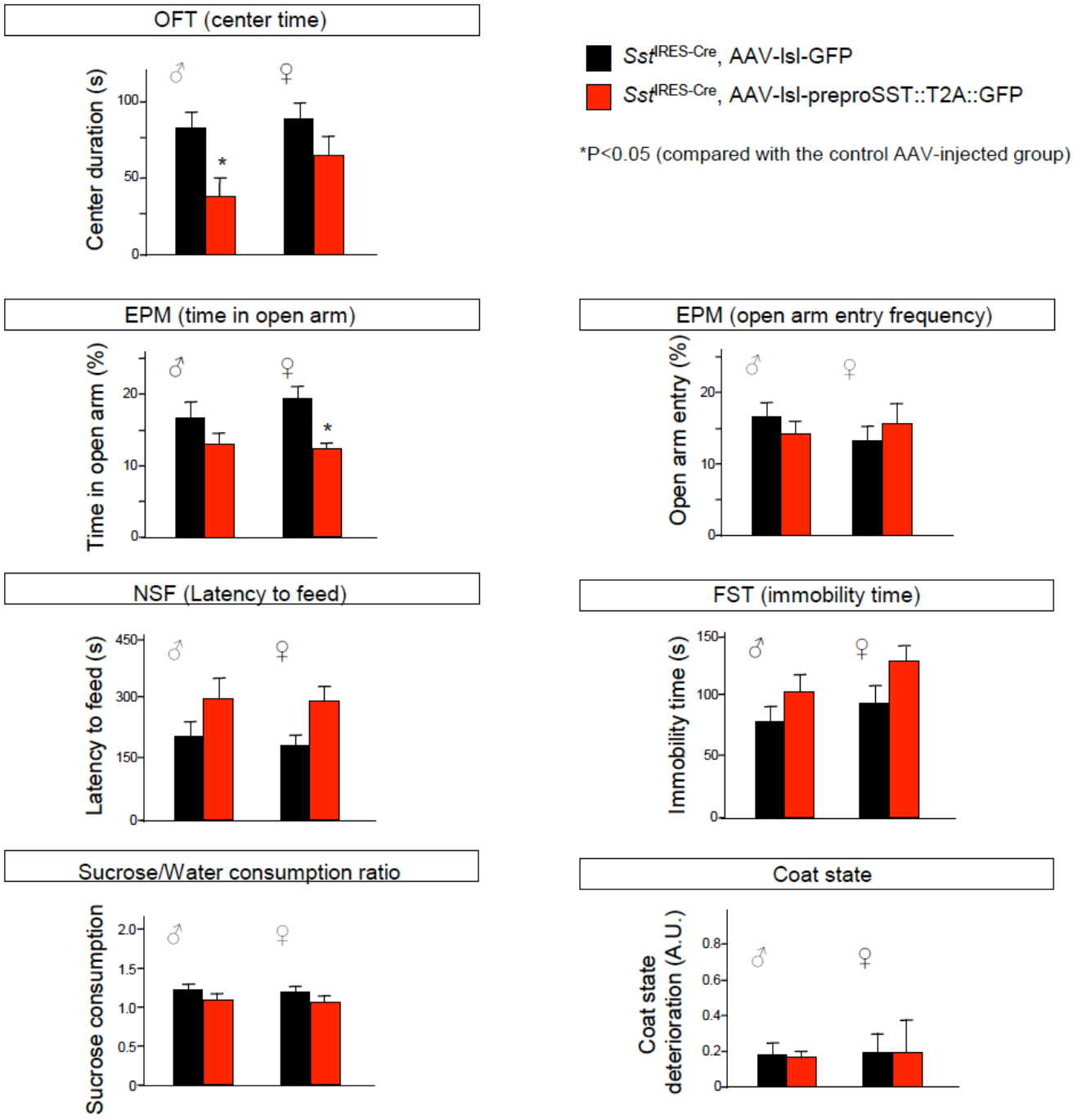
Supplementary Fig. 3



Supplementary Fig. 3. *Sst* mRNA levels in wild-type mice during 5 weeks of chronic stress.

Mice (N=6/group, 3-4 months old male) were subjected to chronic stress (physical restraint in a Falcon tube for 1 h, twice per day) for indicated periods and the PFC was sampled for quantitative PCR analysis immediately (15 min) after the last stress procedure. Data are plotted as mean±SEM. Kruskal-Wallis test with Dunn's multiple comparisons; *P<0.05 (as compared with the no stress condition).

Supplementary Fig. 4



Supplementary Fig. 4. Forced expression of preproSST in SST^{+} neurons induces behavioral emotionality.

Mice (*Sst*$^{IRES\text{-}Cre/+}$, 7♂+7♀/group, 3 months old) were stereotaxically injected into PFC with either AAV-lsl-GFP control or AAV-lsl-preproSST::T2A::GFP virus, and assayed 3 weeks post-surgery for open field test (OFT), elevated plus maze (EPM), novelty-suppressed feeding (NSF), sucrose consumption (SC), forced-swim test (FST) and coat state. Statistical significance was evaluated by two-way ANOVA with repeated measures, followed by Bonferroni post-hoc tests; *P<0.05 (as compared with control AAV of the same sex). The effects of preproSST transgene expression on both OFT and EPM (time in open arm) were significant when all cohorts (males and females combined) were evaluated (P=0.046 for OFT, P=0.048 or EPM). When the effects were evaluated for each sex, OFT in males (P=0.040) and EPM (time in open arm) in females (P=0.042) showed significant effects.

Sample Collection, Library Preparation, and RNA-Sequencing

After 5 weeks of UCMS or control conditions and subsequent anxiety- and anhedonia-like behavioral tests, mouse brains were freshly isolated, frozen, and cut through cingulate areas (24a, 24b, and 32; bregma: +1.53 –1.93, representing the medial prefrontal cortex) into 12 µm-thick coronal sections at –20°C using a Cryostat (CM1950, Leica Microsystems, Wetzlar, Germany). Sections were thaw-mounted onto RNase-free polyethylene-naphthalate slides (Leica) for laser-capture microdissection (LCM) applications.

A rapid fluorescence *in situ* hybridization (FISH) protocol was used to stain sections with cDNA probes (Advanced Cell Diagnostics, Newark, CA) specific to *Slc17a7* (vesicular glutamate transporter 1), *Sst*, *Pvalb*, or *Vip* to visualize major cortical neuronal cell types. Each cell type was detected in separate, consecutive tissue sections and visualized using ATTO 550, in order to prevent detection differences across fluorescence channels. FISH protocols for each cell type were staggered to ensure consistent timing between the protocol start and cell collection, thus minimizing cross-sample differences in RNA degradation. From each mouse, 130 cells of each cell type were collected using LCM (LMD7 system, Leica). Cells were identified and traced manually across all cortical layers. Microdissection progress was actively monitored, with incompletely dissected cells flagged and dissected manually, thus each sample represented precisely.

RNA was extracted using a PicoPure RNA isolation kit (Thermo Fisher Scientific, Waltham, MA), and libraries were prepared using the SMARTer Stranded Total RNA-Seq Kit v2-Pico Input Mammalian (Clontech, Mountain View, CA). Importantly, this kit employs random hexamer, which is better suited than oligo-dT priming to handle the degraded RNA resulting from LCM. Indeed, **Figure A** demonstrates consistent coverage across transcript lengths. Library fragment size distribution was determined using a Bioanalyzer (Agilent Technologies, Santa Clara, CA), and sequencing was performed on an Illumina HiSeq 2500 sequencer (Illumina, San Diego, CA) to generate 2x125bp paired-end reads. Reverse reads showed lower and more variable Q30 scores and were thus excluded. Forward reads were trimmed (5':9bp, 3':21bp) to remove regions with the highest variation in read quality. Reads passing quality control were aligned to the GRCm38 mouse reference genome (ftp.ensembl.org/pub/release-86/fasta/mus_musculus/dna/) using HiSat2 [1] and GenomicAlignments [2].

RNA-Sequencing Technical Measures and Quality Control (QC)

Per-sample sequencing depth, alignment statistics, and Phred scores (mean across read length) were reported previously [3]. Independent-samples t-tests revealed no significant differences in these measures between UCMS and control mice. RNA quality proved difficult to assess in our samples. Though the starting RNA quality is very high in snap-frozen brain tissue, this is not indicative of true RNA quality after FISH and LCM collection. Here, we inferred RNA quality in our libraries by examining the sequencing coverage across transcript lengths using Picard tools (using the *CollectRnaSeqMetrics* function), with greater 5' bias indicative of greater RNA degradation. Normalized coverage scores (corrected for GC bias) are calculated proportionally (i.e. normalized for transcript length) along the 1,000 most highly expressed transcripts. Coverage distributions were compared across cell types and between experimental groups using Kolmogorov-Smirnov (KS) tests, with significantly different distributions being further analyzed with post-hoc t-tests in the 5' and 3' regions (with Bonferroni correction). 5' and 3' were arbitrarily chosen as position 11-30 and 71-90, respectively, as the extreme 5' and 3' ends showed much lower coverage than the main portion of the transcript (**Figure A**, upper panel).

Figure A shows the coverage distributions across cell-types (upper panel), and across groups (lower panel). **Figure A** demonstrates a slight 5' bias in our libraries, though this bias is quite modest with an average of 11% greater coverage of the 5' end versus the 3' end. Given the RNA degradation inherent in LCM, some degree of 5' bias is expected in these samples. KS tests showed no significant difference between any combination of coverage distributions (across cell-types or experimental groups), except for UCMS PYR-cells and UCMS VIP-cells ($p=0.0243$). Post-hoc tests identified that this difference was driven by greater 5' coverage, and lower 3' coverage, in PYR-cells vs VIP-cells ($p<0.05$ at positions: 18-21, 26-30, 71, 76-77, 84-90). Differences in coverage across cell-types may reflect biological differences amongst cell-types and are modest here regardless. Importantly, no significant differences in RNA integrity were observed as a result of UCMS.

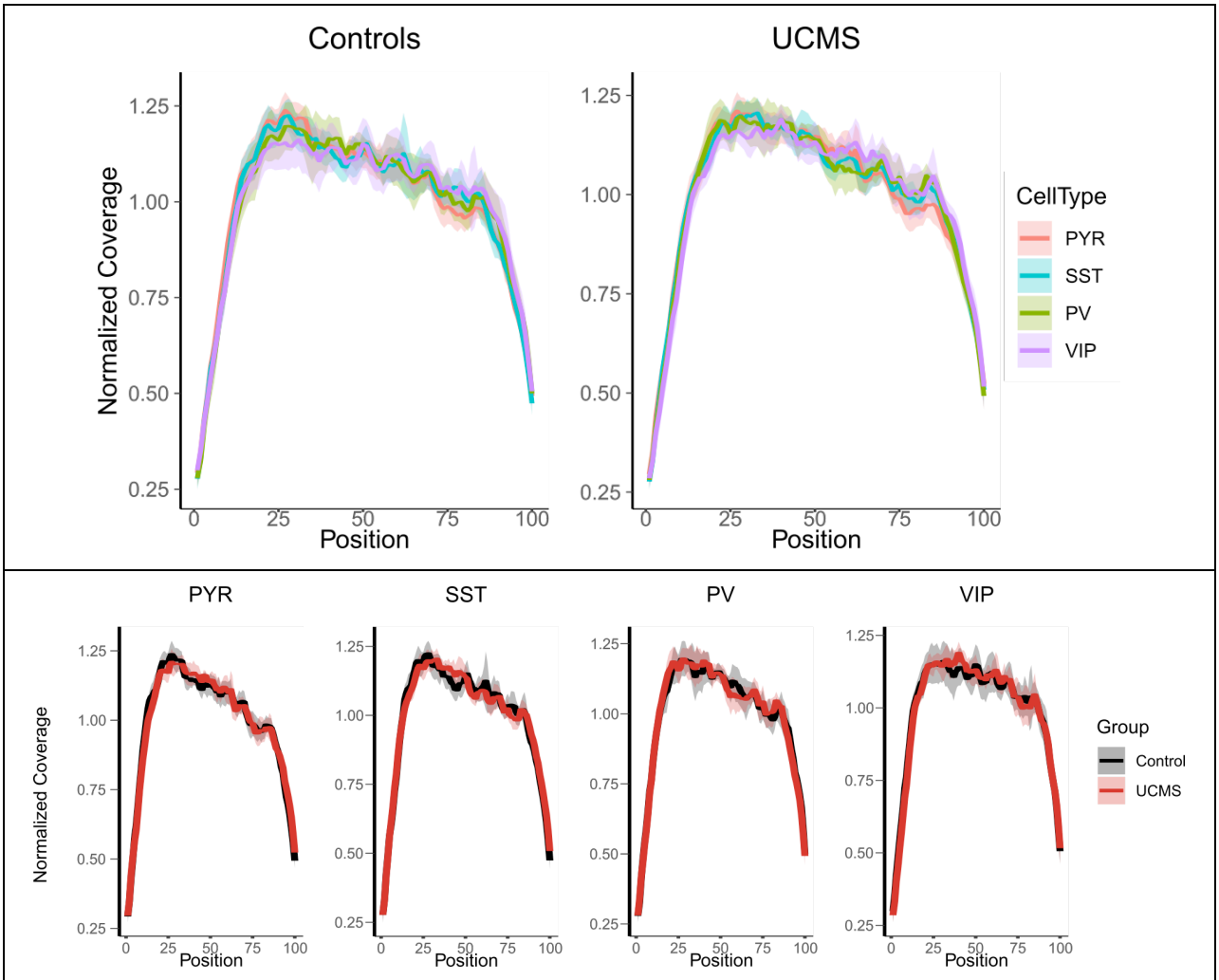


Figure A. Normalized coverage distributions along the length of the 1,000 most highly expressed transcripts. **Upper panel**: Distributions shown within experimental groups, coloured by cell type. The distributions are significantly different only between PYR and VIP-cells in the UCMS group (p=0.0243). **Lower panel**: Distributions shown within cell types, across experimental groups, no significant differences were observed between controls and UCMS mice in any cell type.

Cell type Specificity Assessment

Cell type specificity was assessed by determining relative expression of known molecular markers and data-driven methods. **Table B** outlines the known cell type markers used [4], and the expected cell types in which they are expressed. Expression of each marker gene of interest was expressed in reads per kilobase of transcript per million mapped reads (RPKM) and compared across cell types using one-way ANOVA with a Bonferroni correction of post-hoc independent-samples t-tests.

Table B. Cell type specific markers

Gene	Expected cell type(s) showing specific expression
Slc17a7	PYR-cells
Sst	SST-cells
Pvalb	PV-cells
Vip	VIP-cells
Gad1	All interneurons
Gad2	All interneurons
Calb1	SST and PV-cells
Calb2	SST and VIP-cells
Stmn2	All cell types

Data-driven cell type specificity was assessed by principal component analysis (PCA). PCA was performed on the top 500 most variable genes across all samples, and samples were plotted by the first and second components (accounting for 46% of the variance in our data). K-means clustering (k=4, 10,000 iterations, 10,000 starting points) was performed, and the known cell type identity of each sample was used as an external validation measure to determine percentage accuracy.

Differential Expression (DE) and Biological Pathway Enrichment Analysis

After alignment of reads to exons, DE analysis was performed using DESeq2 [5]. To account for differences in cell-specific presence or expression levels of genes, we filtered genes before DE analysis on a cell type-wise basis. In each cell type, genes were excluded if they had less than 10 reads and greater than 2/3 of samples with a 0 count. Contrasts for the effect of UCMS in each cell type were examined. FDR correction was performed using the *fdrtool* package, with a cell type-wise correction for multiple comparisons. Differential expression significance was set at 15% false discovery rate (FDR), and pathway enrichment significance was set at $p<0.05$.

Gene Set Enrichment Analysis (GSEA) [6] was performed using Wald statistic-ranked gene lists, EnrichmentMap gene-set database [7], and default GSEA parameters, permuted 10,000 times. GSEA provides the advantage of being less reliant on a small number of highly significant DE genes to provide enrichment output, as it is designed to capture biological pathways with multiple members consistently biased towards to tails of the DE distribution [6].

Supplementary References

1. Kim D, Langmead B, Salzberg SL. HISAT: a fast spliced aligner with low memory requirements. Nat Methods. 2015; 12: 357.

2. Lawrence M, Huber W, Pagès H, Aboyoun P, Carlson M, Gentleman R, et al. Software for computing and annotating genomic ranges. PLoS Comput Biol. 2013; 9: e1003118.

3. Newton DF, Oh H, Shukla R, Misquitta K, Fee C, Banasr M, Sibille E. Chronic stress induces coordinated cortical microcircuit cell-type transcriptomic changes consistent with altered information processing. Biol Psychiatry. 2021; doi: 10.1016/j.biopsych.2021.10.015.

4. Tremblay R, Lee S, Rudy B. GABAergic interneurons in the neocortex: From cellular properties to circuits. Neuron. 2016; 91: 260–292.

5. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol. 2014; 15: 550.

6. Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, et al. Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles. Proc Natl Acad Sci U S A. 2005; 102: 15545–15550.

7. Merico D, Isserlin R, Stueker O, Emili A, Bader GD. Enrichment map: A network- based method for gene-set enrichment visualization and interpretation. PLoS One. 2010; 5: e13984.