# Supplemental Information

# Modulation of Neuronal Excitability and Plasticity by BHLHE41 Conveys Lithium Non-Responsiveness

**Marius Stephan[1,2*], Sergi Papiol[1,3,4*], Mingyue Zhang[5], Jie Song[6,7], Samuel M. Frommeyer[5], Helen Haupt[5], Niels Jensen[1], Nirmal Kannaiyan[8], Rajinder Gupta[1], Philipp Schuler[1], Pia Picklmann[1], Ruofan Li[1], Michael McCarthy[9,10], Eva Schulte[3,11,12], Mikael Landen[6,13], Peter Falkai[1], Volker Scheuss[1,14], Thomas Schulze[3,15,16,17], Weiqi Zhang[5], Moritz J. Rossner[1,8]**

## Table of Contents

# Supplemental Methods

## Human Genetics

### Definition of circadian gene sets

Core clock gene sets: genes centrally involved in the cellular time-keeping mechanism [1–3]. Since it is difficult to define a universally accepted consensus list of core clock genes, we opted for three different gene sets ranging from a more restrictive ("Coreclock minimal") to a more relaxed ("Coreclock extended") definition.

Clock modulator gene set: *Clock modulators human* includes genes that serve as likely regulators of the core clock based on a genome-wide small interfering RNA (siRNA) screen in a human cellular clock model [4]. The primary hit list included i) double hits (siRNA) for high amplitude, short period and long period, and ii) single hits for high amplitude and short period.

Circadian oscillating/output gene sets: genes that exhibit oscillatory expression but make no known contribution to the core time keeping mechanism. The *Circadian 6 mouse tissues* gene set was based on those genes showing a circadian oscillation expression pattern in at least 6 different mouse tissues [5]. The *Circadian 4 CNS tissues (CircaDB)* gene set was based on the data available in CircaDB (http://circadb.hogeneschlab.org/) and included genes showing a circadian expression pattern in at least 4 mouse CNS tissues[6]. The *Circadian brain human* gene set included the top cyclic genes in the human brain across six different brain areas [7]. Likewise, the *Circadian prefrontal cortex human* gene set represented the top 50 clock-regulated genes in the human prefrontal cortex (Brodmann areas 11 and 47) [8]. A similar selection of the top 50 clock-regulated genes was done for the dorsolateral prefrontal cortex. *Circadian DL prefrontal cortex human* and the striatum (*Circadian nucleus accumbens/caudate/putamen human)* [9,10]. Finally, a gene set including the top 50 circadian genes most commonly found in different human tissues was also defined (*Human ubiquitous circadian gene* [11]. See Table S1 for a list of the specific genes contained in each gene set.

A curated homology list (Mouse Genome Informatics, The Jackson Laboratory) was used to perform mouse-to-human gene mapping.

### Target phenotypes

GWAS summary statistics available for European-origin populations for schizophrenia (SCZ) [12 13], post-traumatic stress disorder (PTSD) [14], major depressive disorder (MDD) [15], bipolar disorder (BD) [16], autism spectrum disorders (ASD) [17], attention-deficit/hyperactivity disorder (ADHD) [18], lithium response defined as continuous or dichotomous trait (ConLiGen) [20], and chronotype as positive control [21]. For the independent replication study, lithium response GWAS summary statistics from a Swedish cohort was used [22]. The sample size of each study is shown in Table S1.

### Lithium response: samples and phenotypes

The international Consortium on Lithium Genetics (www.ConLiGen.org) has recruited 2,563 BD patients treated with lithium and includes 22 sites across four continents (Europe, America, Asia and Australia). Only European-origin ConLiGen samples (N = 2,343) were used for gene-set association (summary statistics) and polygenic analyses (N = 2,153). The "Retrospective Criteria of Long-Term Treatment Response in Research Subjects with Bipolar Disorder" (Alda scale) [23] was used for the evaluation of long-term treatment response to lithium. This scale was designed to assess long-term response, while considering confounding factors. The A score (range 0 – 10) quantifies symptom improvement over the lithium treatment period. The A score is then weighted against five criteria (B score) that assess confounding factors (disease course, compliance, and add-on treatments) each scored 0, 1, or 2. The total score is derived by subtracting the total B score from the A-score. Dichotomous response criteria were defined with a total score ≥ 7 for "responders" (< 7 "non-responders"). The continuous response was defined using A score but excluding subjects with a total B score > 4 [24].

In the replication Swedish cohort of BD patients, lithium response was defined on the basis of one subjective and one objective assessment [22]. Subjective assessment of lithium response in participants who were treated for at least 1-year (N=1,120) was measured by means of a telephone interview, using a standardized interview protocol. Patients who reported complete remission on lithium were considered "responders". Objective assessment of lithium response was based on recurrence data of patients (N = 942) at yearly longitudinal follow-ups extracted from the Swedish Quality Register for BD (BipoläR). Responders were defined as having no mood episodes over at least 1 year of follow-up.

**Gene set analyses**

The European panel of 1000 Genomes Project, phase 3 was used as reference for all processing/analysis steps. NCBI build 37.3 genomic annotation was used for gene locations.

MAGMA [25] and INRICH [26] were used to carry out competitive gene set analyses. Both methods were chosen because of their good performance in keeping a low type-1 error rate and dealing with confounding factors like gene size, gene density and linkage disequilibrium (LD) between genes [27]. Together they provide complementary approaches because each method captures different facets of the association signal [27]. The built-in procedures for empirical multiple testing correction were not used. Instead, more conservative Bonferroni corrections were used for the set of traits and gene sets after enrichment analyses. Stratified LD Score Regression [28,29] was not used for enrichment analyses due to its low performance when using small annotations covering less than 0.5%-1% of the genome (https://www.medrxiv.org/content/10.1101/2021.03.13.21249938v1).

MAGMA (Multi-marker Analysis of Genomic Annotation, https://ctg.cncr.nl/software/magma) v1.09a was used for formal enrichment analyses using the summary statistics of the previously mentioned major psychiatric disorders, response to lithium treatment, and chronotype. After SNP-to-gene annotation, we performed gene-level analyses (± 10 kb) with the SNPwise-mean model. Gene-based P-values were obtained by combining the SNP P-values in a gene (± 10 kb) into a gene test statistic (mean $\chi 2$) corrected for linkage disequilibrium between SNPs. MAGMA competitive gene-set analysis was used within a regression framework to test whether the circadian-related gene sets were more strongly associated with the target phenotypes than other genes. For gene set competitive testing, linear regression analyses were conditioned to gene size, log(gene size), gene density, log(gene density), inverse mac, and log(inverse mac).

INRICH analyses (https://atgu.mgh.harvard.edu/inrich/) is a permutation-based competitive gene set analysis tool and takes a set of genomic regions (intervals) that are independently associated with a trait, and then tests for the enrichment of predefined gene sets. LD-independent associated genomic intervals for each trait were generated using SNP tagging (--tag-kb 1000 --tag-r2 0.2) in PLINK (v1.9) [30]. Intervals were calculated at three different P-value thresholds ($P_{TH}$ = 0.00001, $P_{TH}$ = 0.0001, $P_{TH}$ = 0.001). The next step was to count the number of associated intervals overlapping with genes in the circadian gene sets. Interval overlap was limited to ±10 kb up/downstream of a gene. Random interval sets were generated 10,000 times (first-pass permutations), with each set approximately matching the associated intervals regarding the number of SNPs, overlapping genes, and SNP density per kb. The overlap of circadian gene sets was also determined in these random intervals to generate a distribution of the enrichment statistics for such target gene sets under the null hypothesis. Finally, the empirical P-value was calculated as the proportion of random replicates where the enrichment statistic is as large as that of the original interval set.

**Data analysis and visualization**

Additional statistical analyses and data management were carried out using R v3.5.0 and SPSS v25. Plots were generated using *ggplot2* or *lattice* R packages.

# Behavioral Animal Study

All animal experiments were carried out in accordance with institutional guidelines and regulations approved by the Regierungspräsidium Oberbayern, ROB Munich, Germany under the license ROB-55.2-1-54-2532-179-2016.

## Animals and Husbandry

A wildtype colony of C57Bl/6NCrl mice was used for backcrossing. Chow and water were provided *ad libitum*, except for the IntelliCage system-based experiments for spatial learning. *Bhlhe40/41*$^{-/-}$ (DKO) and wildtype (WT) mice were no more than two to three generations apart from littermates and weaned after three weeks. Male and female *Arntl* (*Bmal1*) heterozygous and littermate wild-type control mice were used at an age of 8 weeks for behavioral analyses. Both, male and female mice were tested. The experiment mice were kept in type IV cages (Tecniplast 2000, 612 x 435 x 216 mm, 2065 cm²) in groups of litters of 12 to 16 mice. Experiments with young adult mice started at 8 weeks (Open Field Test) and ended at 19 weeks (Remote Fear Memory) of age.

## Handling and General Experimental Procedures

Home cages were equipped with handling tunnels. All experiments outside of the IntelliCage system (TSE Systems) were conducted during the light phase, randomizing the trials across group to ensure a uniform distribution, and excluding circadian effects. For habituation, the home cage was placed in the test room for ten minutes before the actual experiment. Before and after each trial, the arena and other equipment in direct contact with test mice was cleaned with a dilute SDS solution, wiped dry, and disinfected with 70 % ethanol, unless stated otherwise. Subjects were excluded from the analysis of individual learning phases, if they did not drink in the respective phase and, in consequence, had to be hand-fed. All behavioral test details have been described previously [34].

## Lithium treatment

Starting three weeks before the first behavioral test, mice were treated chronically with 600 mg/L lithium chloride (LiCl) in the drinking water throughout the experiments. Water was available *ad libitum*. No significant differences in drinking behavior were observed between groups when monitored with the IntelliCage system.

## Transponder Implantation

All experiment mice were equipped with an RFID chip for tracking in the IntelliCage and blinded identification throughout the experiments. Before the actual surgery, the neck region was shaved, cleaned with 70 % alcohol for disinfection, and the transponder (1.4 x 11 mm) was injected subcutaneously.

## Behavioral Test

All behavioral tests, statistical analyses and visualizations were performed blinded and essentially as described in detail previously [34], and are described in detail below.

## Camera Tracking

The following tests were done using the camera-based tracking software ANY-maze (Stoelting Europe): open field test, Y-maze test, tail suspension test, and fear conditioning test. In addition to the published standard fear conditioning protocol, the mice were placed in the original arena three weeks after the conditioning stage to assess remote context-related fear memory. All other tests were performed essentially as described in detail previously [34–36].

## Prepulse Inhibition Test

For the assessment of prepulse inhibition of the startle response to auditory stimuli, we used the SDI Startle Response System and SRLab software (San Diego Instruments).

## IntelliCage-Based Tests

For circadian activity tracking, spatial learning tests and the sucrose preference test, we used the IntelliCage system (TSE Systems). The device allows for home cage monitoring in social group and conditioning in four corners, each with a water bottle in a corner. The access to these water bottles can be denied in learning experiments by closing plastic doors, opening it only in case of “correct” nose pokes at the respective door. The tests were run as described [34–36] with the following minor modification: Activity was first assessed in a five-day period, during which the mice were first accustomed to the new environment (two days of free adaptation), the door opening (one day doors opened on visit), and to learn opening the doors by nose poking (two days nose-poke adaptation). After that, mice could access water in only one of the four corners for spatial positive reinforcement learning (two days place learning).

Next the corner was changed pseudo-randomly for five days. Finally, one of the two bottles in each corner was filled with 4 % sucrose water for the sucrose preference test (one day).

### Data Analysis and Visualization

Data were aggregated and processed using *FlowR* (XBehavior) as the user interface for the PsyCoP bundle. Additionally, R scripts were used to conduct all custom non-pipeline data analyses using *R 4.0.4* [37] in *RStudio IDE* (RStudio). Box and whisker plots and dimension plots were created with *ggplot2* [38] or *lattice* [39] R packages. Heatmaps were plotted with *pheatmap* [40]*.* A multivariate linear model was fitted and used in a multivariate analysis of variance (ANOVA) from a Wilk's lambda distribution using *car* [41]. Univariate contrasts were derived from the same model. For univariate contrasts with a significant interaction, a unifactorial "simple-effects" ANOVA was computed to independently test the treatment effect for each genotype.

The dimensionality reduction procedure was described in detail earlier [34]. In brief, missing values were imputed using the non-linear iterative partial least squares (NIPALS) algorithm in *ade4* [42]. The resulting reconstituted matrix was used for canonical discriminant analysis (CDA) with *candisc* [43]. CDA provides an optimal solution for group separation in phenotypic space, by finding linear combinations of single variables, called *canonical components,* similar to the *principal components* in a principal component analysis (PCA). Based on the CDA results of the collapsed factors, we generated a dimension plot with overlayed data ellipses, which include 75 % of the respective sample. A heatmap was created, showing the weights (*canonical coefficients*) of each dependent variable in the *canonical component* of each of the three terms (genotype, treatment, and interaction) of the multivariate linear model. Moreover, a CDA was computed for each of the two "simple-effects" models (DKO and WT).

Each variable was classified *a priori* in accordance with the Research Domain Criteria (RDoC) framework [44]. The heatmap was sorted and grouped, accordingly.

## Single-Nucleus Transcriptomics

### Sample Processing

For single-nucleus RNA sequencing (snRNAseq), mouse brains were isolated at zeitgeber time (ZT) 4, at the minimum of BHLHE41 expression [45,46], and kept in ice-cold phosphate-buffered solution (PBS; Gibco). Target tissue encompassing the anterior cingulate gyrus (bregma -0.1 – 1 mm) was isolated from 300 µm slices. Tissue from three mice was collected and pooled: two females, one male per group in the untreated pilot experiment at 3 months of age and one female, two males in the lithium treated main experiment at 4 months old. The freshly isolated tissue was homogenized in nucleus isolation buffer (0.32 M sucrose, 3 mM $Mg(Ac)_2$, 5 mM $CaCl_2$, 10 mM Tris-HCl pH 8.1 with 0.1 % Triton-X100 freshly added; all Sigma) with a 2 ml Dounce-type tissue homogenizer. The sample was overlayed on top of a sucrose cushion (1.8 M sucrose, 3 mM $Mg(Ac)_2$, 10 mM Tris-HCl, pH 8.1, all Sigma) and centrifuged at 4 °C and 17,000 x g for 70 min. The supernatant was removed, the pellet washed with PBS and resuspended in resuspension buffer (PBS, 1 % bovine serum albumin (BSA) and 0.04 U/µL Protector RNAse inhibitor; all Sigma). The suspension was passed through 30 µm pre-separation filters (Miltenyi Biotec) and cells were kept in ice-cold resuspension buffer. Small aliquots of the samples were stained with trypan blue and Hoechst 33258 to determine nucleus density, integrity, and multiplet rate in an improved Neumann counting chamber. The nucleus density was adjusted to 2500 nuclei/µL (±10 %). Nuclei aggregation rate was significantly below 10 % in all samples.

### Library Preparation

Single-cell libraries were prepared according to manufacturer protocols (Illumina) and isolated using a ddSEQ Single-Cell Isolator (Bio-Rad) in the pilot experiment or the chromium controller (10x Genomics) in the main treatment study. Libraries were prepared following the manufacturer's recommendations (SureCell 3'WTA (Illumina/BioRad) or Chromium Next GEM 5' v2 (10xGenomics). Average fragment size, quality, and yield of the final libraries were assessed on a Bioanalyzer 2100 (Agilent). Libraries were sequenced at a loading concentration of 300 pM on a NovaSeq6000 instrument (Illumina) in paired-end mode.

**Upstream Analysis of snRNAseq Data**

The BCL files were demultiplexed with *bcl2fastq* (Illumina) and the sequence quality was checked with *FastQC* (v0.11.9). For 10x Genomics Next GEM libraries, barcode tagging, mapping and cell calling were done using CellRanger (v7.0; 10xGenomics). The reference genome was constructed from release 102 of the *Ensembl* annotation and the *GRCm38* mouse genome. For SureCell libraries, cell barcodes and unique molecular identifiers (UMIs) were isolated and tagged with *ddSeeker* [47]. After tagging, the data were processed with *Drop-seq tools* (v2.3.0). A metabundle was created from the same genome assembly and annotation used for CellRanger. Corrupted cell barcodes were removed from the data. The reads were sorted by query name, and the data converted back to FASTQ for mapping. Read2 was mapped using *STAR aligner* (v2.7.5) [48] with default settings. Mapped data were merged with barcode-tagged data. The uniquely mapped reads were functionally annotated for counting. Digital gene expression data matrices were created for the 6000 most abundant cell barcodes in each sample, since no automated cell calling was implemented in the pipeline. The cut-off was chosen based on the knee plot made with *ddSeeker*'s *make_graphs.R*.

**Downstream Analysis of snRNAseq Data**

For downstream processing of snRNAseq data, we used the R package *Seurat v4* (v4.0.1) [49]. After import as a dgc matrix, Seurat objects were created for each sample. Ambient RNA contamination was assessed using *CellBender* (v0.3.2; 10'000 cells expected, 15'000 droplets considered, 150 epochs)[50] and *DecontX* (*celda v1.18.1*)[51]. These two methods were found to perform well across a variety of sxRNAseq datasets. Only cells with at least 200 features expressed and only features expressed in at least 3 cells were included in the object. Mitochondrial RNA content was determined by grepping for "^mt-" and hemoglobin contamination was quantified by grepping for "^Hb". The cells were filtered for hemoglobin contamination of less than 1 % of the total UMI counts, a minimum of 500 (untreated pilot) or 800 (lithium main) different expressed features to remove both empty droplets and low-quality nuclei. Aggregates were identified and removed using the *scds* package's combined approach [52].

**Normalization and Integration of snRNAseq Data**

The UMI count data matrices were normalized by the single cell transform procedure (*SCTransform*) [53]. Both, mitochondrial and hemoglobin transcript contents were regressed out as indicators of RNA background. For integration of the lithium data, we used *Seurat's* canonical correlation analysis (CCA) approach. For integration of the untreated dataset, we decided to use the reciprocal principal component analysis (rPCA) procedure implemented in *Seurat* with the *k.anchor* parameter set to 48, since we observed oversmoothing with CCA-based integration in that small dataset.

**Dimension Reduction and Clustering of snRNAseq Data**

The first 30 principal components (PCs) were computed using the 3000 most variably expressed features. Using all 30 PCs, we performed similar nearest neighbor (SNN) graph-based clustering. We identified a total of 6 cell clusters in the untreated and 15 clusters in the lithium treated samples. Next, we computed a uniform manifold approximation and projection (UMAP) embedding [54].

**Differential Expression Analysis**

Differential gene expression was analyzed with *Seurat v4* [49]. Features entering the analysis were filtered for a minimum of 1 % (untreated) or 10 % (lithium) expressing cells in the tested population. The filtered data were re-normalized to the lowest median UMI count sample but not integrated. The results were tested using the model-based analysis of single-cell transcriptomics (*MAST)* procedure [55] and P-values were Benjamini-Hochberg corrected. Differentially expressed features going into downstream analysis were filtered for a minimum average log2-transformed fold-change (avg_log2FC) of 0.1 (untreated) or 0.25 (lithium) as low coverage reduces the fold-change due to a high proportion of non-expressing cells.

**General Enrichment Analysis and Protein Interaction Network Analysis**

For gene annotation enrichment analysis and protein-protein interaction (PPI) network analysis, the Metascape platform was used [56]. Differentially expressed genes (DEGs) from layer 2/3 neurons from

the control samples were entered for custom analysis. Default parameters for annotation and membership were used. The background gene set parameter was set to the expressed features in the respective population and only selective gene ontology (GO) clusters were analyzed. For network analysis, the top 3 candidate genes, CSNK1E, BHLHE41, and RORB, as well as BHLHE41's paralog BHLHE40, were added to the gene list and the network was computed using all databases combined.

Extended Protein-protein interaction maps with layer 2/3 excitatory neurons from the lithium samples were queried using the DGE genes as input using the Bisogenet plugin from Cytoscape which consolidates data from DIP, BIOGRID, HPRD, BIND, MINT and INTACT databases [57,58]. Cytoscape network grafs were exported to Adobe Illustrator for final processing.

**Gene Set Enrichment Analysis (GSEA)**

The GSEA results were generated using an in-house developed wrapper for the clusterProfiler R package [59]. The wrapper allows for selecting the desired gmt file/source and streamlining multiple analysis steps from data input to generation of results (tables and figures). The presented GSEA analysis was performed for GO terms (BP, CC, and MF) and Reactome pathways. The gmt data was obtained from MSigDb [60] using the msigdbr R package [61]. The results were filtered for q-value ≤ 0.05 and normalized enrichment score (NES) > 0 for upregulated or NES < 0 for down regulated.

ClusterProfilerWrapper Github repo: https://github.com/MNB-Lab/clusterProfilerWrapper

**Enrichment Analysis of Synapse-Related Gene Ontology Terms**

For a more specific enrichment analysis of synapse asspciated associated transcripts, the SynGO database for synapse related GO annotations was used [62]. The input gene list was extended by relaxing the P-value adjustment and all expressed features in the respective cell population was used as background gene list.

# Reporter Assays

All reporter gene assays in primary cultured mouse cortical neurons were essentially performed as described previously [63]. For data analysis in R, the mean signal and the corresponding standard error of each group was computed for plotting with ggplot2. For the boxplots and statistics, the area under the curve (AUC) was approximated using the rolling mean between neighboring datapoints for each sample's stimulation peak and tested by two-way ANOVA using type 2 sum of squares and subsequent simple-effect (one-way) ANOVAs of the treatment effect of each genotype. Since a mean of means was tested and the small sample size of four, no test for normality was conducted.

**Cell culture**

Primary cortical neurons were isolated from brains of E15.5 C57BL6/NCrl WT and DKO mice. Cortical tissue was treated with a pre-activated papain suspension (Worthington) and gently dissociated by pipetting. $5 \times 10^5$ Cells were plated in neurobasal medium supplemented with 5 % fetal calf serum (FCS), 2 % B27 serum supplement, and 1 % GlutaMax (all Gibco) on 35 mm plastic tissue culture dishes (BD Falcon) coated with poly-D-lysine (MW 70,000 – 150,000; Sigma).

On day *in vitro* (DIV) 1, plating medium was replaced with 1.5 mL culture medium (neurobasal medium, 2 % B27, 1 % GlutaMax). 500 µL culture medium was added on DIV5 and DIV11. The cultures were kept in a humidified incubator at 37 °C and 5 % $CO_2$.

**AAV production**

AAV1/2 particles were produced following a previously published protocol [63]. In brief, HEK293FT cells were transfected with the ESARE reporter AAV vector (pAAV-ESARE-Luc2A) and pFdelta6 helper plasmid. A mix of the capsid plasmids pH21 (serotype 1) and pRV1 (serotype 2) was used to produce AAV1/2 mixed serotype particles. Polyethylenimine (PEI) transfection reagent (Polyscience) was used for transfection. Cells were lysed and treated with benzonase (Sigma). The lysate was filtered through a 0.45 µm syringe filter (Millipore) and concentrated using an Amicon Ultra-15 centrifugal filter unit (100 kDa membrane cut-off; Millipore). Each AAV sample was digested with TURBO DNase (Invitrogen) to remove double-stranded copies of the viral genome and successfully packaged AAVs were released

by proteinase K (Invitrogen) digestion and isolated by spin column-based purification (Macherey-Nagel). The titer was quantified by qRT-PCR with WPRE primers.

### Live-cell luciferase assays

For ESARE assays, cultures were infected by adding virus stock to the culture medium on DIV1 (MOI 1000). On DIV5, lithium chloride in PBS or PBS alone was added with 500 µL medium (final concentration 1.5 mM). On DIV11, the supernatant was supplemented with D-luciferin for baseline recordings. After 24 hours, cultures were stimulated by adding 1(*S*),9(*R*)(−)-bicuculline methochloride (BIC) to a final concentration of 50 µM, silenced by adding a mixture of tetrodotoxin (TTX) and (2*R*)-amino-5-phosphonopentanoate (AP-V) at a final concentration of 1 µM (TTX) and 100 µM (AP-V), or mock treated by adding fresh medium (vehicle only). During the recordings all cell cultures were kept in a LumiCycle 32 apparatus inside a dry incubator (37 °C, 5 % $CO_2$).

## Electrophysiological recordings

### Whole-cell patch-clamp recordings

Slices were prepared as described previously [64]. In brief, mice were decapitated, and the isolated brains transferred into ice-cold and pre-oxygenated artificial cerebrospinal fluid (aCSF; buffered solution of 1.25 mM $Na_2HPO_4$, 125 mM NaCl, 14 mM D-glucose, 2 mM $MgSO_4$, 1.5 mM $CaCl_2$, 2.5 mM KCl, and 26 mM $NaHCO_3$, pH 7.4). 300 µm Slices were cut using a vibratome (VT 1200, Leica). Slices containing either the ACC or the vHIP were transferred into the incubation chambers with oxygenated aCSF (pH 7.4, aerated with 95 % $O_2$, 5 % $CO_2$, 32 – 34 °C) for 1 h recovery period, and then to room temperature aCSF, oxygenated with carbogen at a rate of 4 mL/min.

All whole-cell recordings were performed on pyramidal cells in the cortical layer 2/3 in ACC or pyramidal neurons in the CA1 region in vHIP, as described previously [65,66]. Briefly, pyramidal neurons were identified by morphology. Access resistance was measured before and after recording using transient current responses to 5 mV hyperpolarization pulses. Patches with more than 10 % change in access resistance were excluded from downstream analysis. Additional exclusion criteria were leak currents of less than 150 pA, membrane resistance below 0.8 GΩ, or serial resistance of more than 20 MΩ.

Spontaneous excitatory postsynaptic currents (sEPSCs) were measured at a holding potential of −70 mV. sEPSC recordings in ACC were measured in the presence of 5 µM strychnine and 5 µM BIC. Recording pipettes were filled with recording solution (140 mM potassium gluconate, 0.5 mM $Na_2GTP$, 4 mM $Na_2ATP$, 10 mM EGTA, 1 mM $CaCl_2$, and 2 mM $MgCl_2$, buffered with 10 mM HEPES, pH 7.3).

Spontaneous inhibitory postsynaptic currents (sIPSCs) were measured in the presence of 10 µM 6-cyano-7-nitroquinoxaline-2,3-dione (CNQX, Sigma) and 50 µM AP-V. Here, the recording pipettes were filled with recording solution containing 0.5 mM $Na_2GTP$, 4 mM $Na_2ATP$, 10 mM EGTA, 140 mM KCl, 1 mM $CaCl_2$, and 2 mM $MgCl_2$, buffered with 10 mM HEPES (pH 7.2).

For mEPSC and mIPSC recordings, 0.5 µM TTX was added to the bath solution.

### Long-term potentiation measurements

Long-term potentiation (LTP) recordings were performed as described earlier[67–69]. In brief, transverse hippocampus slices (300 µm) were cut in ice-cold and pre-oxygenated slicing buffer (10 mM glucose, 218 mM sucrose, 3 mM KCl, 0.2 mM $CaCl_2$, 6 mM $MgSO_4$, 1.25 mM $NaH_2PO_4$, and 26 mM $NaHCO_3$). Slices were transferred to 32 °C warm aCSF oxygenated with carbogen and incubated for 1 h. For field potential recordings, Schaffer collaterals were stimulated in the stratum radiatum at the CA3/CA1 junction. Amplitude (baseline to peak) and slope (20 – 80 % level of the falling phase) of field excitatory post-synaptic potentials (fEPSPs) were quantified. The half-maximum response was defined as baseline fEPSP. Three trains of 1 s 100 Hz high-frequency stimulation (HFS) bouts in 20 s intervals were used for LTP induction. Every 20 s after the last train, the response to extracellular stimulation in the CA1 region was measured for 60 min. We quantified short-term potentiation (STP), defined as the mean

response in the first 5 min after HFS, and LTP, defined as mean response in the last 10 min of measurement.

**Action Potential Measurements and Kainic Acid Stimulations**

We measured action potentials (APs) on ex vivo AAC slices (see above) in the current clamp mode [64]. For quantification the interspike intervals (ISI) were plotted as mean instantaneous firing frequency (f) of two consecutive APs (f = 1/ISI). Kainic acid (1 µM) was applied via bath perfusion with a 2 min delay after application and a measurement time of 3-5 min.

**Data Analysis**

We quantified mean Amplitude and Frequency of mEPSC, sEPSC, mIPSC, as well as sIPSC in ACC and vHIP whole-cell recordings. Each variable was tested in a two-way ANOVA using type 2 sum of squares and, in case of a significant interaction, subsequent simple-effect (one-way) ANOVAs of the treatment effect for each genotype. Since a mean of means was tested, no test for normality was conducted. The same procedure was used for STP and LTP as well as Kainate stimulation experiments.

# Supplementary Figures

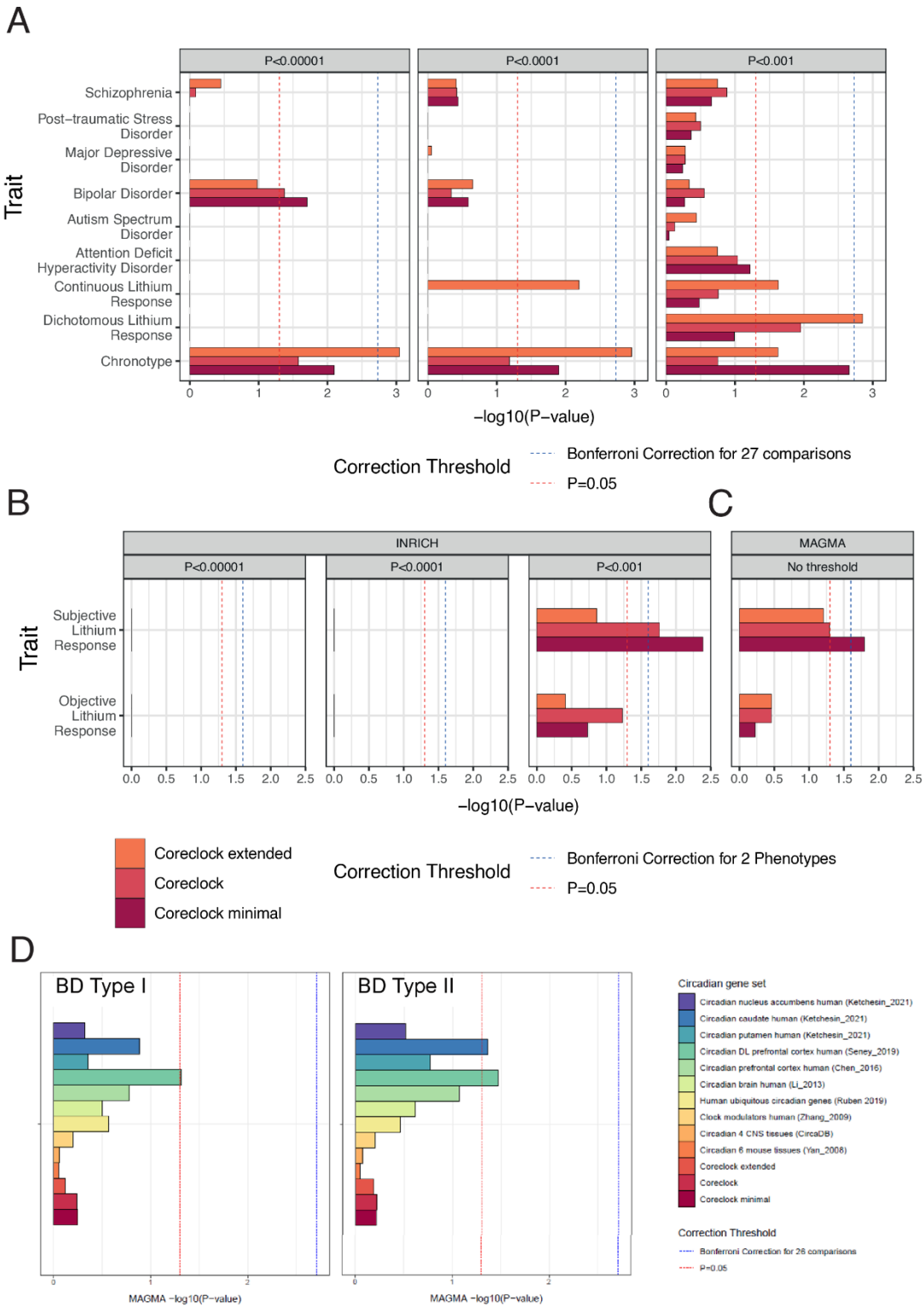


**Figure S1. INRICH analyses based on core clock gene-sets in psychiatric traits.** (A) INRICH analyses of data shown in Figure 1. Bonferroni significance thresholds: dashed blue line (P < 0.05/ (3 gene-sets x 9 traits)). Three different P-value thresholds were analyzed (P=0.00001, P=0.0001, P=0.001). The 'Coreclock extended' gene-set reaches corrected p-val-threshold at P=0.00001 and P=0.0001 in the Chronotype sample, whereas it reaches corrected p-val-threshold in the 'Dichotomous Lithium Response' sample at P=0.001. (B) INRICH analyses and (C) MAGMA enrichment analyses of swedish replication cohort with two phenotypes: subjective and objective lithium response. Bonferroni

significance thresholds: dashed blue line (P < 0.05/ (3 gene-sets x 2 traits)). (Competitive test P-values in the X-axis are –log10 converted. Dashed red line indicates the nominal (uncorrected) significance threshold P=0.05. (D) MAGMA enrichment analyses of Bipolar disorder Type I or Type II lifetime risk, based on circadian gene-sets. Competitive test P-values in the X-axis are –log10 converted. Dashed blue line indicates the Bonferroni significance threshold (P<0.05/8 gene-sets x 2 traits). Dashed red line indicates the nominal uncorrected) significance threshold P=0.05.

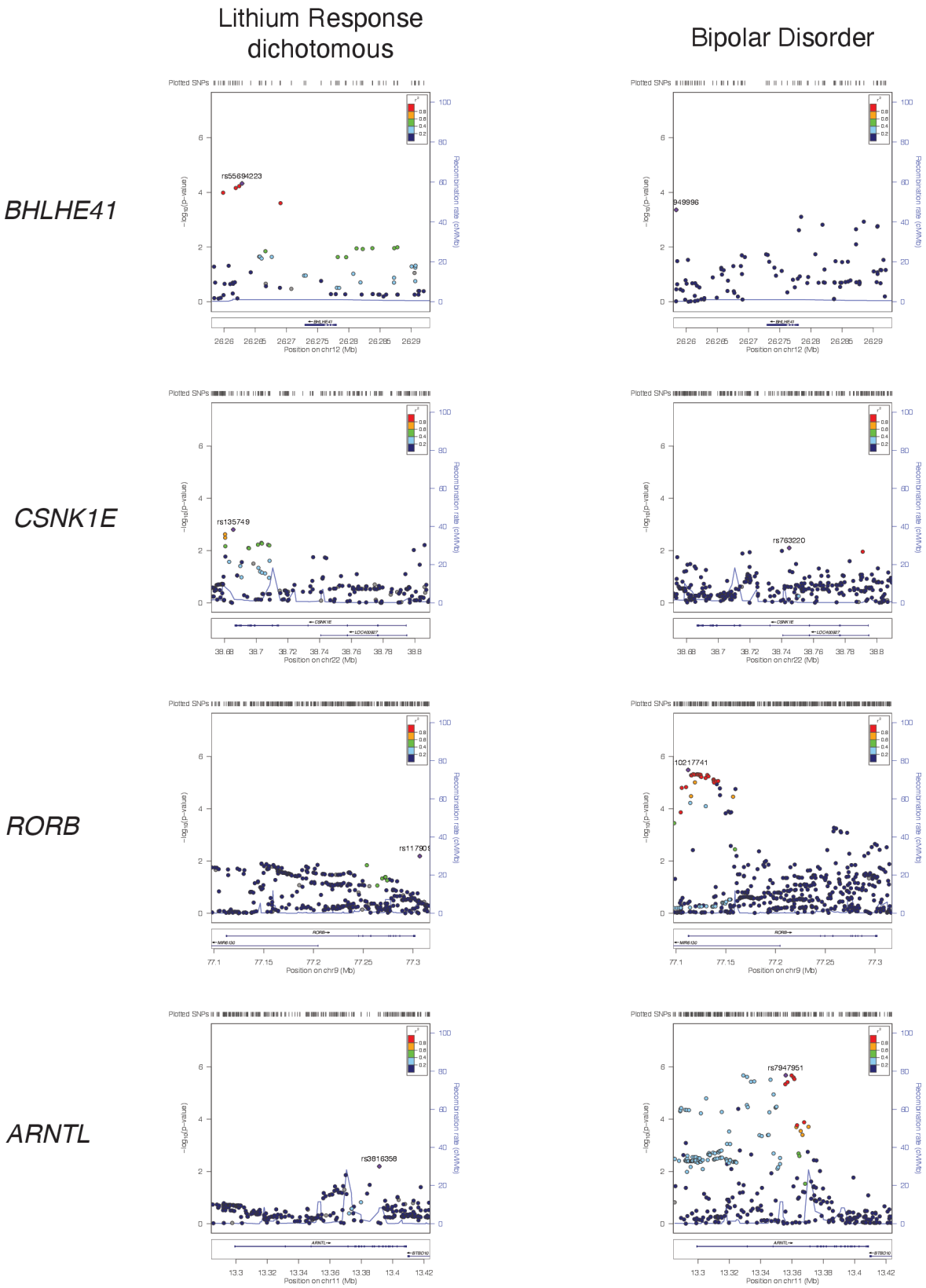


**Figure S2. Manhattan plots for nominally significant contributing genes in the core clock gene-set and ARNTL.** Manhattan plots from GWAS data on dichotomous lithium response (A, C, E, G) and bipolar disorder risk (B, D, F, H). For the top three (and nominally significant) lithium response-associated genes from the core clock gene-set: (A, B) *BHLHE41*, (C, D) *CSNK1E*, and (E, F) *RORB*. In addition, the same diagrams are shown for *ARNTL* (G, H), a GWAS associated BD risk gene also known as the core clock regulator *BMAL1*.

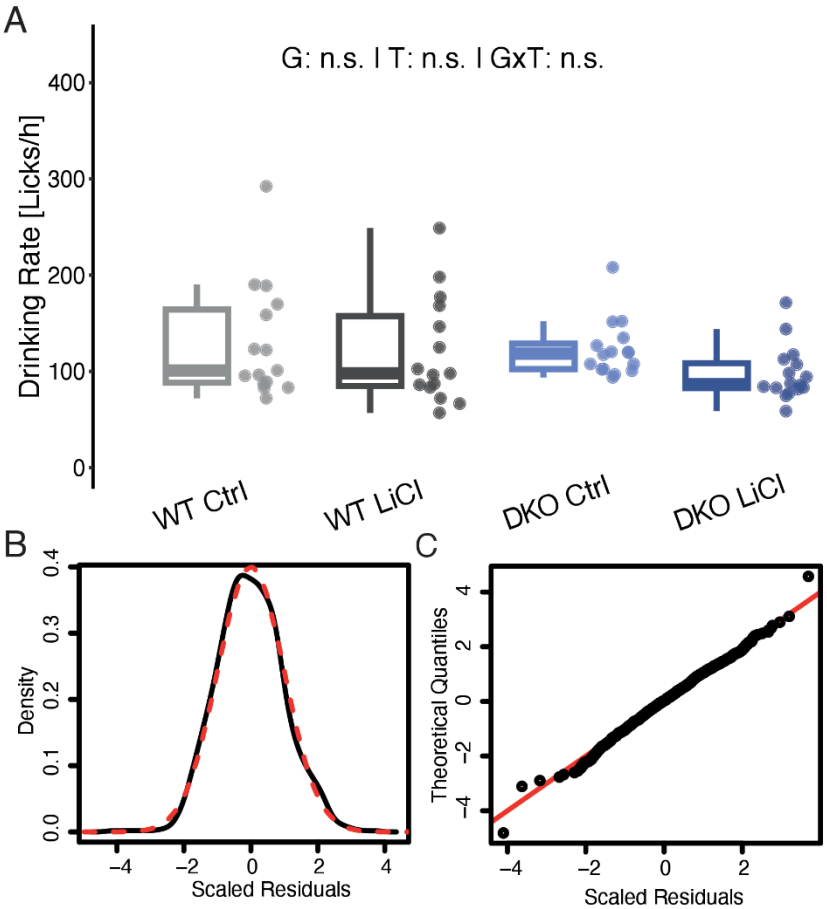


**Figure S3. Quality controls for the PsyCoP behavioral study.** (A) Drinking rate in the IntelliCage during the experiments was not significantly different in lithium chloride-treated mice compared to control. Data are shown as box plots with whiskers extending to no more than 1.5-fold IQR. (B) Density plot and QQ plot of the scaled residuals (black) of the multivariate linear model used for the ANOVA procedure of all PsyCoP measures compared to a normal distribution (red) show no major deviations. * P < 0.05, ** P < 0.01, *** P < 0.001, n.s. not significant; P-values are FDR-adjusted and refer to Wilk's lambda testing two-way ANOVA; G: genotype term; T: lithium treatment term; GxT: interaction term; WT: wildtype; DKO: Bhlhe40/41 double-knockout; Ctrl: vehicle control; LiCl: lithium-treated

**Figure S4. Remaining variables of the cognition and sensorimotor RDoC domains.** (A-C) display the variables of the Cognitive RDoC domain not shown in Figure 2. (A) There was no significant effect on working memory in the Y-Maze test measured with the rate of spontaneous alternations between arms. (B) In the first reversal lithium treated mice needed less trials to reach the learning criterion indicating improved learning flexibility. Although, the difference in DKO mice was smaller than in wildtypes, the interaction was not significant in this part of the spatial learning experiments. (C) Similarly, in the cued fear memory task, lithium treatment increased the time spent freezing, showing improved cued fear memory, with no significant influence in DKOs. (D) In the remote contextual fear memory, however, there was a significant Li-treatment dependent effect between the genotypes, as indicated. (E) The ribbon plot of this curve refers to the Serial Reversal Task analysis (Fig. 2E) and displays the mean number of trials each group needed to reach the spatial learning criterion (10% better success rate than random expectation) after each reversal of the correct corner. Smaller values equal less trials needed and indicate faster learning success. WT animals improve upon Lithium treatment, whereas DKO performance remains unaltered. (F-I) show the variables of the Sensorimotor domain with a strong genotype effect in all of them, but only minor influence of lithium treatment on the highest prepulse level (PPI 80dB) in panel (I). Data are shown as box plots with whiskers extending to no more than 1.5-fold IQR; * $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$, n.s. not significant; P-values are FDR-adjusted and refer to Wilk's lambda testing two-way ANOVA; G: genotype term; T: lithium treatment term; GxT: interaction term; WT: wildtype; DKO: Bhlhe40/41 double-knockout; Ctrl: Placebo control; LiCl: lithium-treated

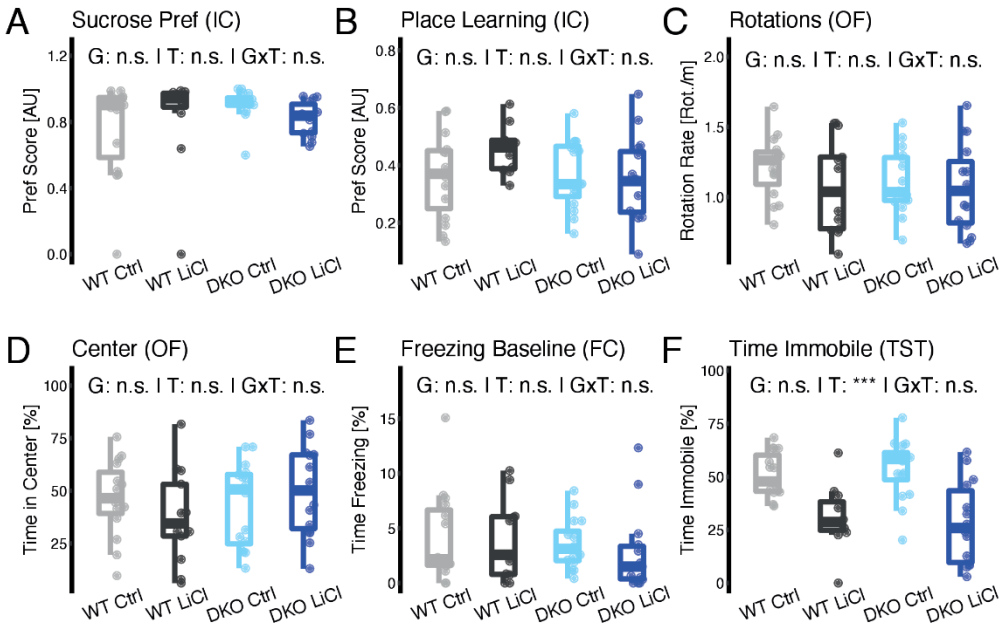


**Figure S5. Positive and Negative Valence RDoC domains.** The variables associated to positive valence systems did not show significant effects, neither in the Sucrose Preference Test (A) nor the place preference in Positive Reinforcement Learning (B). Notably, there was sex difference in sucrose preference detected, shown in Figure S7. In the negative valence domain parameters, there were no significant differences in rotation rate (C) and the center time (D) in the Open Field Test as well as the baseline freezing behavior (E) in Fear Conditioning either. However, lithium-treated mice were significantly less immobile (F) in the Tail Suspension Test independent of genotype. Data are shown as box plots with whiskers extending to no more than 1.5-fold IQR; * P < 0.05, ** P < 0.01, *** P < 0.001, n.s. not significant; P-values are FDR-adjusted and refer to Wilk's lambda testing two-way ANOVA; G: genotype term; T: lithium treatment term; GxT: interaction term; WT: wildtype; DKO: Bhlhe40/41 double-knockout; Ctrl: Placebo control; LiCl: lithium-treated

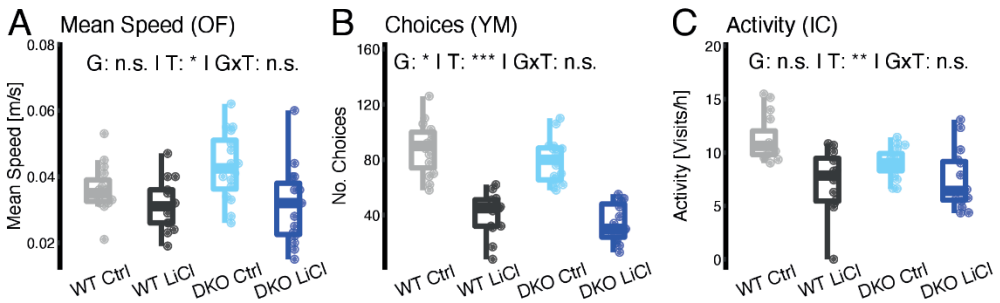


**Figure S6. Arousal and Regulatory RDoC domain.** Novelty-induced activity was reduced in lithium-treated animals (A) measured as Mean Speed in the Open Field Test and (B) the number of arm choices in the Y-Maze Test. In latter case Bhlhe40/41 DKO mice displayed a slightly reduced activity, as well, independent of treatment. General Activity (C) in the IntelliCage was similarly reduced in response to lithium treatment. Please note that there was a sex difference in activity level in response to lithium treatment shown in Figure S7 with males being impacted stronger than females. Data are shown as box plots with whiskers extending to no more than 1.5-fold IQR; * P < 0.05, ** P < 0.01, *** P < 0.001, n.s. not significant; P-values are FDR-adjusted and refer to Wilk's lambda testing two-way ANOVA. G: genotype term; T: lithium treatment term; GxT: interaction term; WT: wildtype; DKO: Bhlhe40/41 double-knockout; Ctrl: Placebo control; LiCl: lithium-treated

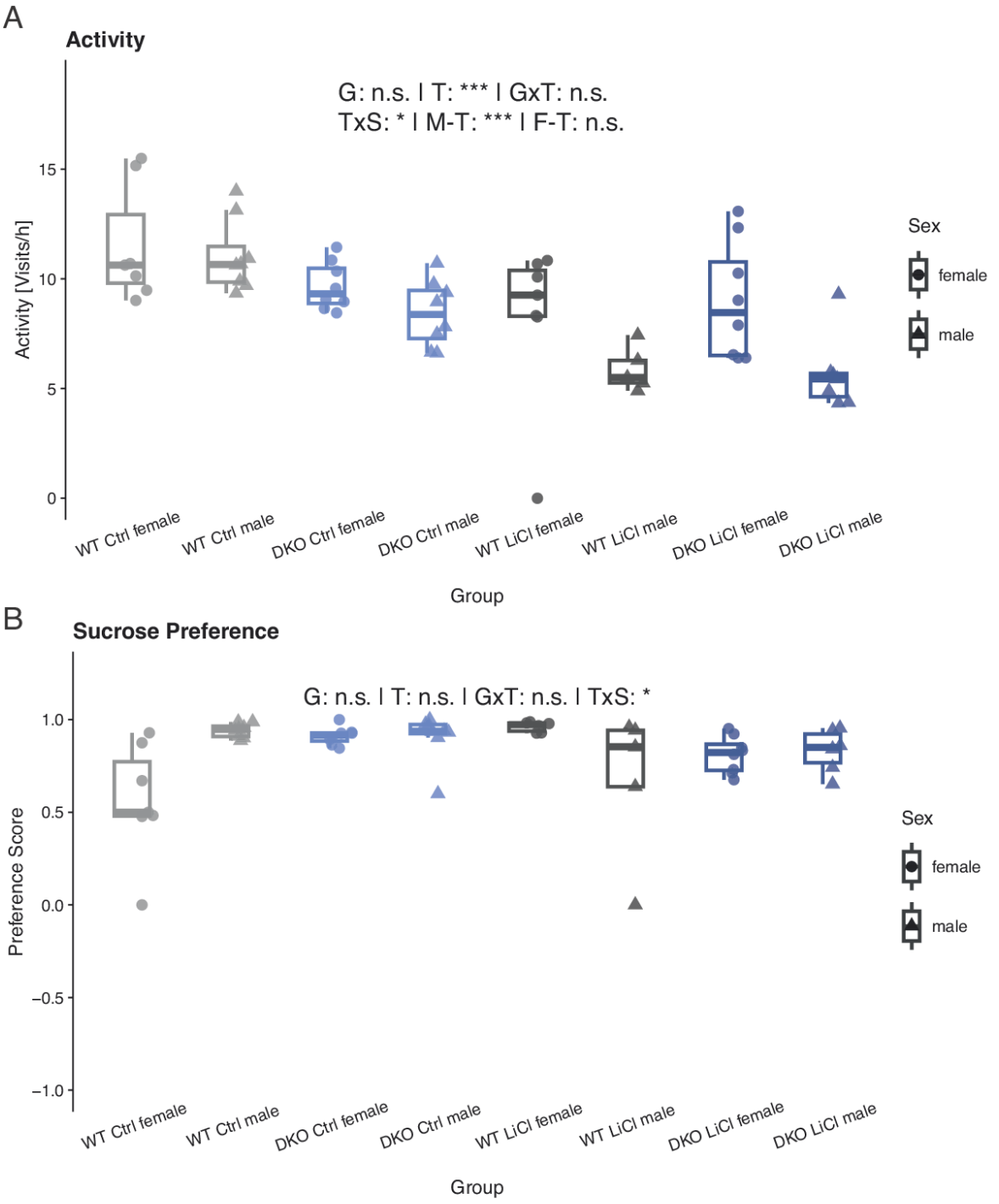


**Figure S7. Sex Differences in general activity and sucrose preference.** In most behavioral parameters assessed, we did not detect significant differences between sexes (Table S3). (A) However, male mice showed a stronger reduction in general activity in response to lithium treatment than female mice. (B) Moreover, the female wildtype placebo control (WT Ctrl female) group displayed a lower sucrose preference than their lithium-treated counterpart. This difference in lithium response between sexes was not found in DKO mice. Data are shown as box plots with whiskers extending to no more than 1.5-fold IQR; * P < 0.05, ** P < 0.01, *** P < 0.001, n.s. not significant; P-values are FDR-adjusted and refer to Wilk's lambda testing two-way ANOVA; n = ; G: genotype term; T: lithium treatment term; GxT: G by T interaction term; TxS: T by sex interaction term; M-T: male treatment effect; F-T: female treatment effect; WT: wildtype; DKO: Bhlhe40/41 double-knockout; Ctrl: Placebo control; LiCl: lithium-treated

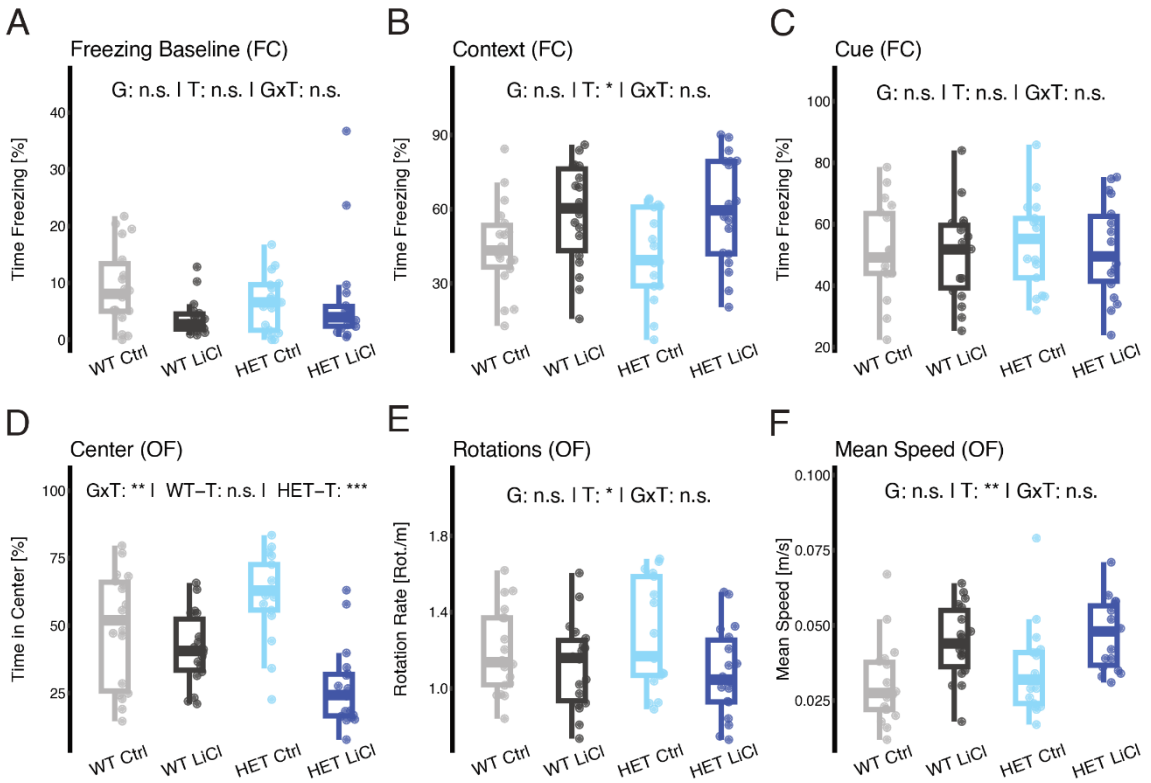


**Figure S8. Fear Conditioning and Open Field of Arntl mouse model.** Arntl heterozygous null mutant mice (Het) and littermate controls untreated or treated with lithium were subjected to fear conditioning (FC) paradigm (A-C) and open field test (D-F). (A) No differences were detected at baseline freezing. (B) Testing contextual fear revealed a treatment dependent significant increase in both genotypes. (C) Cued FC detected no changes. Data are shown as box plots with whiskers extending to no more than 1.5-fold IQR; * P < 0.05, ** P < 0.01, *** P < 0.001, n.s. not significant; P-values are FDR-adjusted and refer to Wilk's lambda testing two-way ANOVA; n = ; G: genotype term; T: lithium treatment term; GxT: G by T interaction term; TxS: T by sex interaction term; M-T: male treatment effect; F-T: female treatment effect; WT: wildtype; Het: Arntl heterozygous null mutants; Ctrl: Placebo control; LiCl: lithium-treated

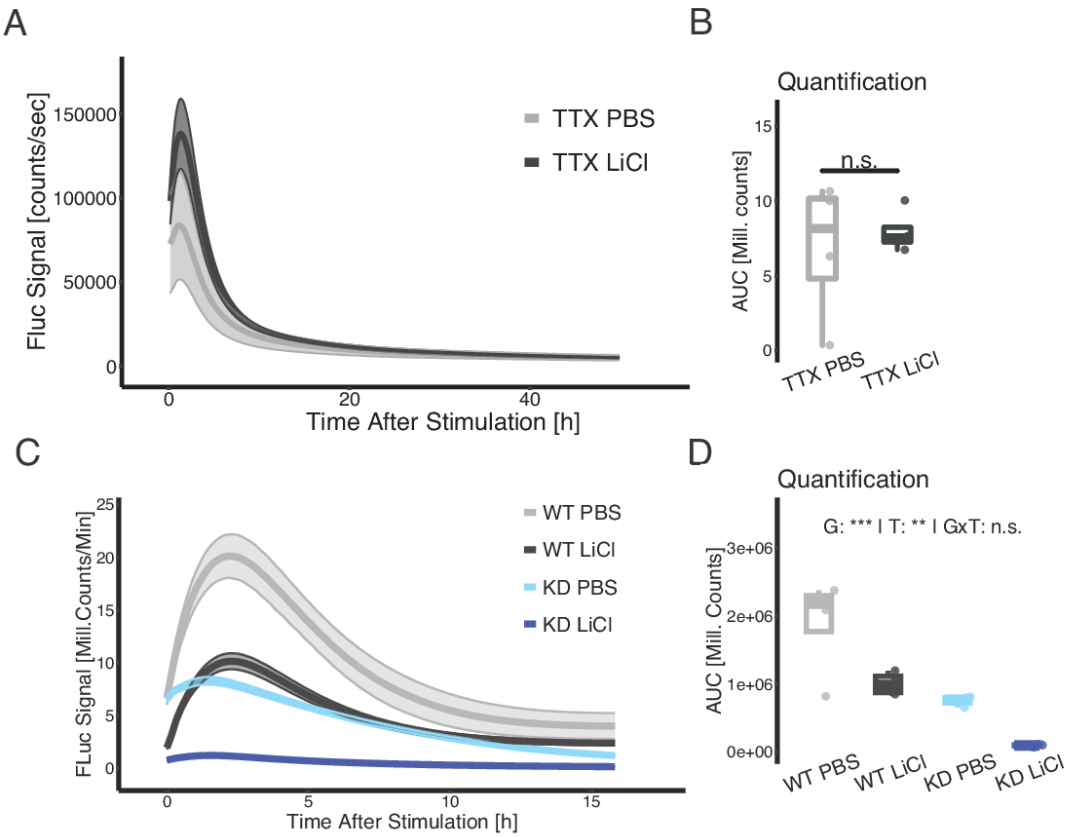


**Figure S9. Lithium toxicity and Bhlhe40 knockdown in primary cortical cultures.** (A) First 50 hours of tetrodotoxin (TTX)- and AP-V-silenced primary cortical neuron culture from wildtype (WT) embryos after lithium treatment. Cells were transduced with an ESARE-Firefly Luciferase (FLuc) reporter construct via AAV infection. After 7 days of pretreatment with lithium chloride (LiCl), a TTX/AP-V cocktail was added to silence neuronal network activity. At the start of the recording a typical artifact peak resulting from the time-lag of the transcriptional-translational reporter system, was seen. (B) The area under the curve of the first 100 hours of recording was quantified. No statistically significant difference was found in a Student's two-sided t-test, indicating that differences in ESARE-response are due to activity dampening not cytotoxic effects of lithium chloride. n.s.: not significant. (C) Timecourse of an ESARE-FLuc reporter assay with a shRNA-mediated knockdown of Bhlhe40 (KD) or a scrambled control (WT). After 7 days of LiCl pretreatment, the cultures were stimulated with bicuculline chloride (BIC) in PBS or plain PBS (PBS) as mock control. (D) Quantification of the ESARE-response peak as area under the curve (AUC) showing a significant knockdown effect (G) and lithium effect, but no significant statistical interaction (GxT) (Two-way ANOVA, Type II SS; G: P = 0.000138; T: P = 0.00139; GxT: P = 0.550).

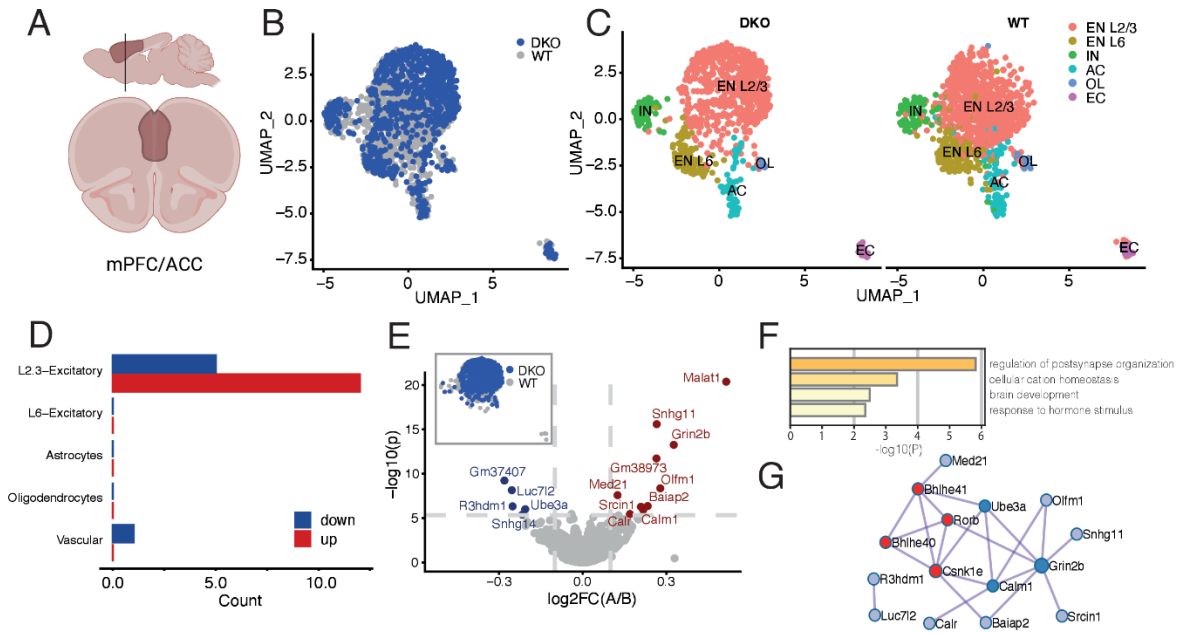


**Figure S10. Layer 2/3 excitatory neurons in the mPFC of BHLHE40-/-/41-/- mice show transcriptional changes in genes associated with the post-synapse.** (A) The anterior cingulate cortex (ACC) part of the medial prefrontal cortex (mPFC) was selected as region of interest and isolated at Zeitgeber Time (ZT) 4, at circadian trough of BHLHE41 expression for single-nucleus RNA sequencing (snRNAseq) to identify constitutively deregulated genes. (B) UMAP dimension plot of the integrated datasets from wildtype (WT) and BHLHE40$^{-/-}$/41$^{-/-}$ (DKO) shows successful integration. (C) Six major cell populations, shown in distinct colors, were identified in both WT and DKO tissue samples. (D) Barplot of numbers of significantly deregulated genes between the genotypes in all cell clusters reveals a selective response in layer 2/3 cells only. (E) A Differential expression analysis between WT and DKO Layer 2/3 excitatory neurons (EN L2/3) identified a small set of differentially expressed genes. The gray lines indicate a threshold of 0.1 average log2-transformed fold-change (log2FC) and the Sidak-corrected significance threshold of the -log10-transformed p-value, assuming independent tests. (F) Gene ontology clusters enriched in the differentially expressed gene (DEG) set find with MetaScape[70] hint to a postsynapse-related mechanism. (G) The top genes identified in the human GWGAS (red) form a protein-protein interaction (PPI) network with the DEGs identified in ACC-derived EN L2/3 (blue) with Grin2b, Calm1, and Ube3a (dark blue) at its center. EN L2/3: excitatory neurons layer 2/3 (upper layer); EN L6: excitatory neurons layer 6 (deeper layer); IN: inhibitory neurons; AC: astrocytes; OL: oligodendroglial cells; EC: endothelial cells

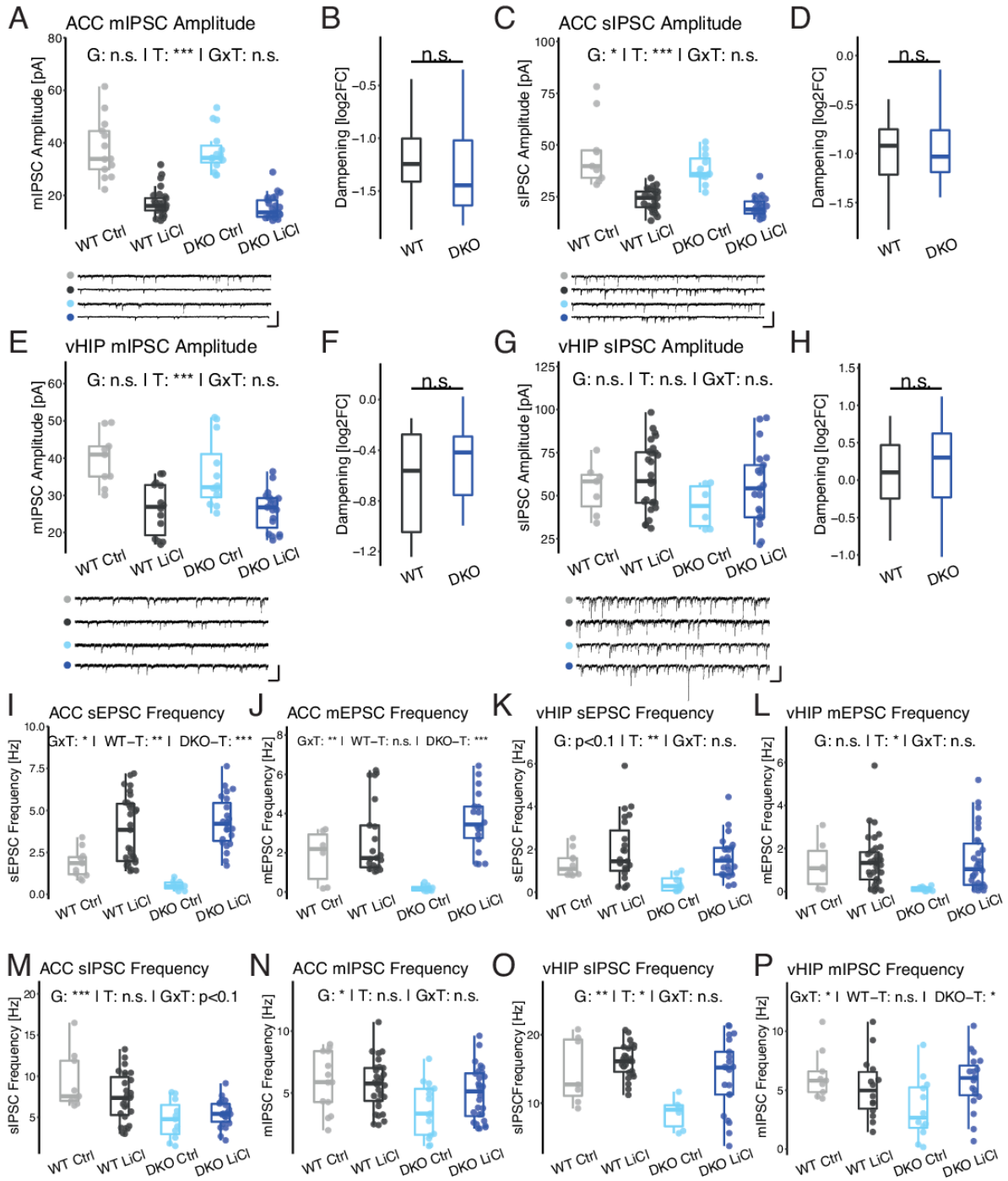


**Figure S11. Extended data from whole-cell recordings in ACC and vHIP.** Whole-cell voltage-clamp recordings in layer 2/3 pyramidal cells in the anterior cingulate cortex (ACC) area of the medial prefrontal cortex (mPFC) and pyramidal cells in the CA1 region of the ventral hippocampus (vHIP). (A-H) Average miniature and spontaneous inhibitory post-synaptic potential (mIPSC & sIPSC) amplitudes with example traces (scale bars x: 0.2 s & y: 0.1 nA) and the log2-transformed fold change of the respective amplitudes in LiCl compared to Ctrl mice. (I-L) Average miniature and spontaneous excitatory post-synaptic potential (mEPSC & sEPSC) frequency. (M-P) Average mIPSC and sIPSC frequency measured in ACC and vHIP. * P < 0.05; ** P < 0.01; *** P < 0.001; n.s. not significant; P-values refer to univariate two-way ANOVA with Type 2 sum-of-squares; simple effects were tested in a similar but unifactorial ANOVA procedure; WT: wildtype mice; DKO: Bhlhe40/41 double-knockout mice; Ctrl: vehicle control; LiCl: lithium-treated; G: genotype main effect; T: treatment main effect; GxT: interaction effect; simple T: simple effects; DKO-T: DKO simple treatment effect; WT-T: WT simple treatment effect.

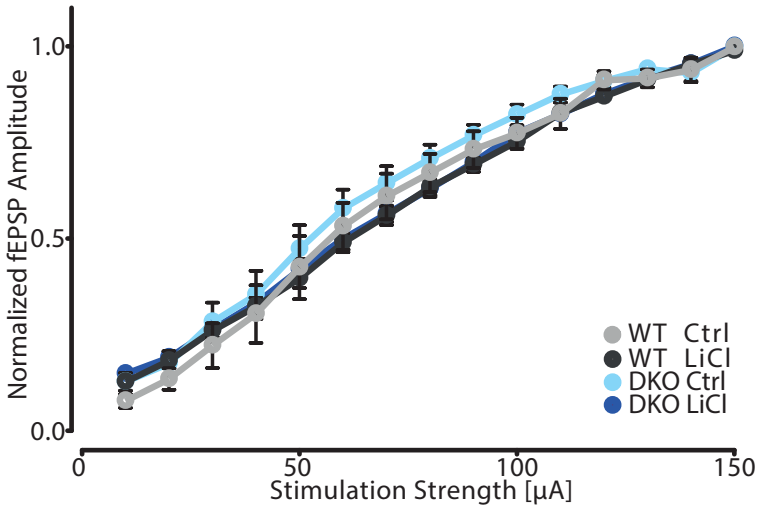


**Figure S12. Extended Data on LTP Recordings in the CA1 Region of the Hippocampus.** Input output (I/O) relationship of the mean (+/-SEM) normalized field excitatory post-synaptic potential (fEPSP) measured in the CA1 region plotted against the stimulation current injected at the Schaffer collaterals in the stratum radiatum at the CA3/CA1 junction did not show any differences in basic synaptic transmission between groups WT: wildtype mice; DKO: Bhlhe40/41 double-knockout mice; Ctrl: vehicle control; LiCl: lithium-treated

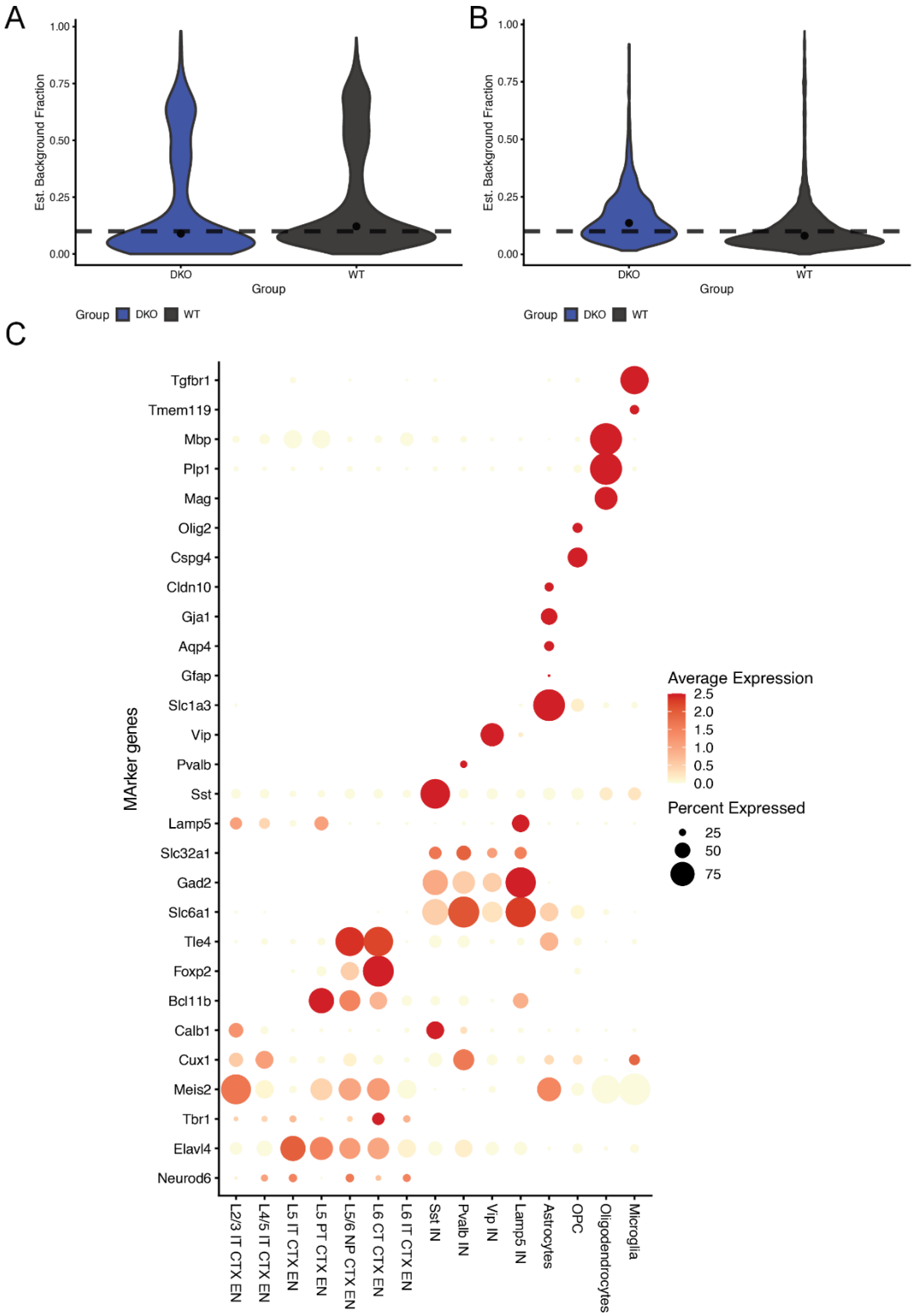


**Figure S13. Quality control for main snRNAseq**. (A, B) Violin plots showing estimated ambient RNA background per cell computed with (A) *CellBender* (9.0 % DKO, 12.2 % WT) and (B) *DecontX* (mean 12.2 % DKO, 8.6 % WT). Dots represent the median and dashed lines mark 10 % of ambient RNA counts. (C) Dotplot showing cell type marker gene expression levels. Normalized average expression is shown in color scale, abundance of expressing cells in percent as dot size. Minimum dot size cut-off is 10 % expressing cells.

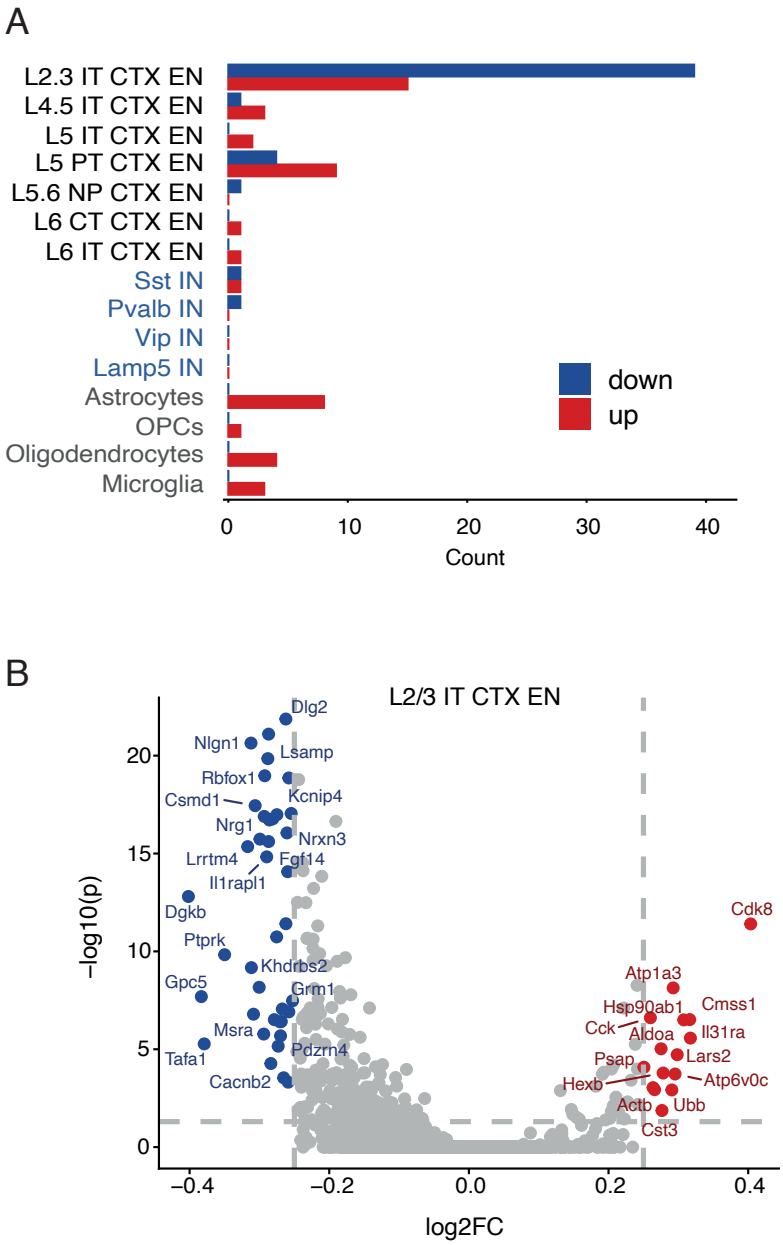


**Figure S14. Differential gene expression analysis in wt and DKO treated with lithium by single cell RNAseq.** (A) Barplot of numbers of significantly deregulated genes between the groups reveals a highly selective response in layer 2/3 excitatory neurons compared to all other cell types. (B) Volcano plot of deregulated genes in layer 2/3 excitatory neurons. On the Y axis the -log transformed adjusted P-value is shown and on the X-axis the average log2-transformed fold-change is plotted at the indicated thresholds (dashed lines).

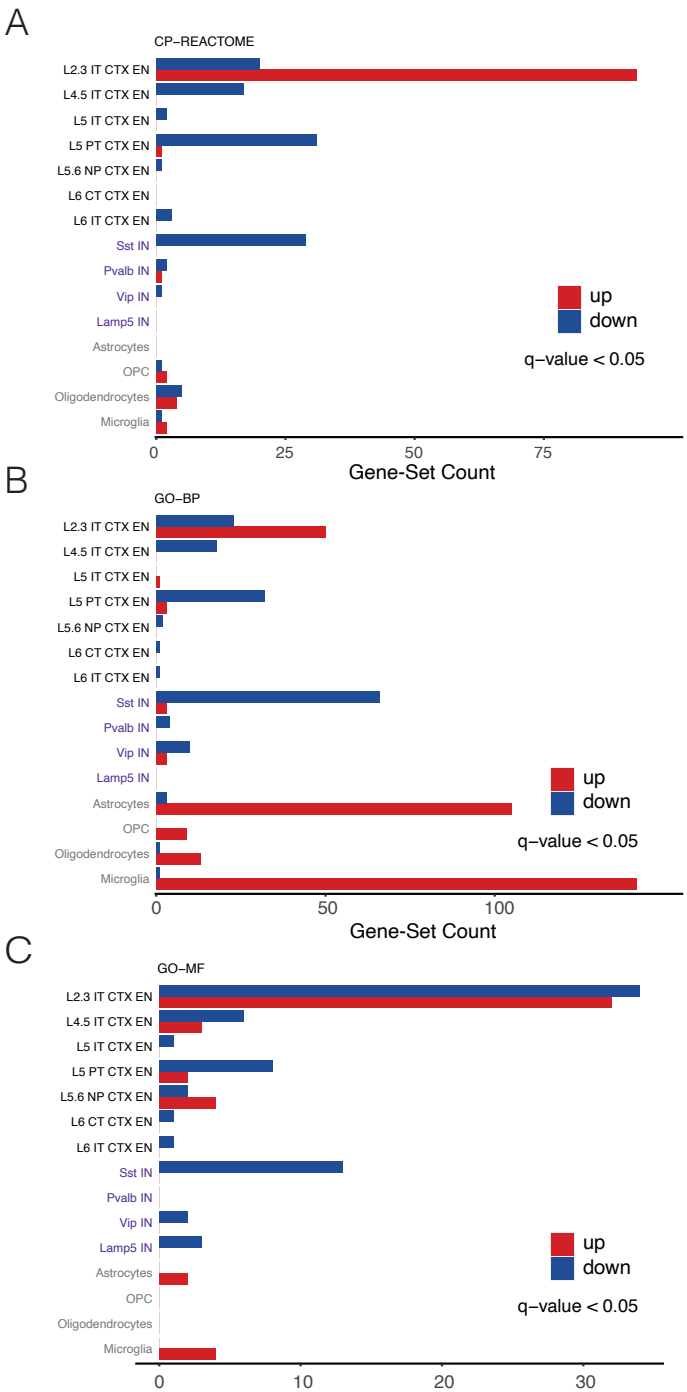


**Figure S15. Differentially regulated gene-sets in wt and DKO treated with lithium identified with GSEA.** (A-C) Barplot of numbers of significantly deregulated gene-sets (corrected q-val < 0.05) between the groups reveals a prominent response in layer 2/3 excitatory neurons in the pathway collection from reactome (A), gene ontology biological process (GO-BP) (B), and the gene ontology molecular function (GO-MF) (C) subcollections.

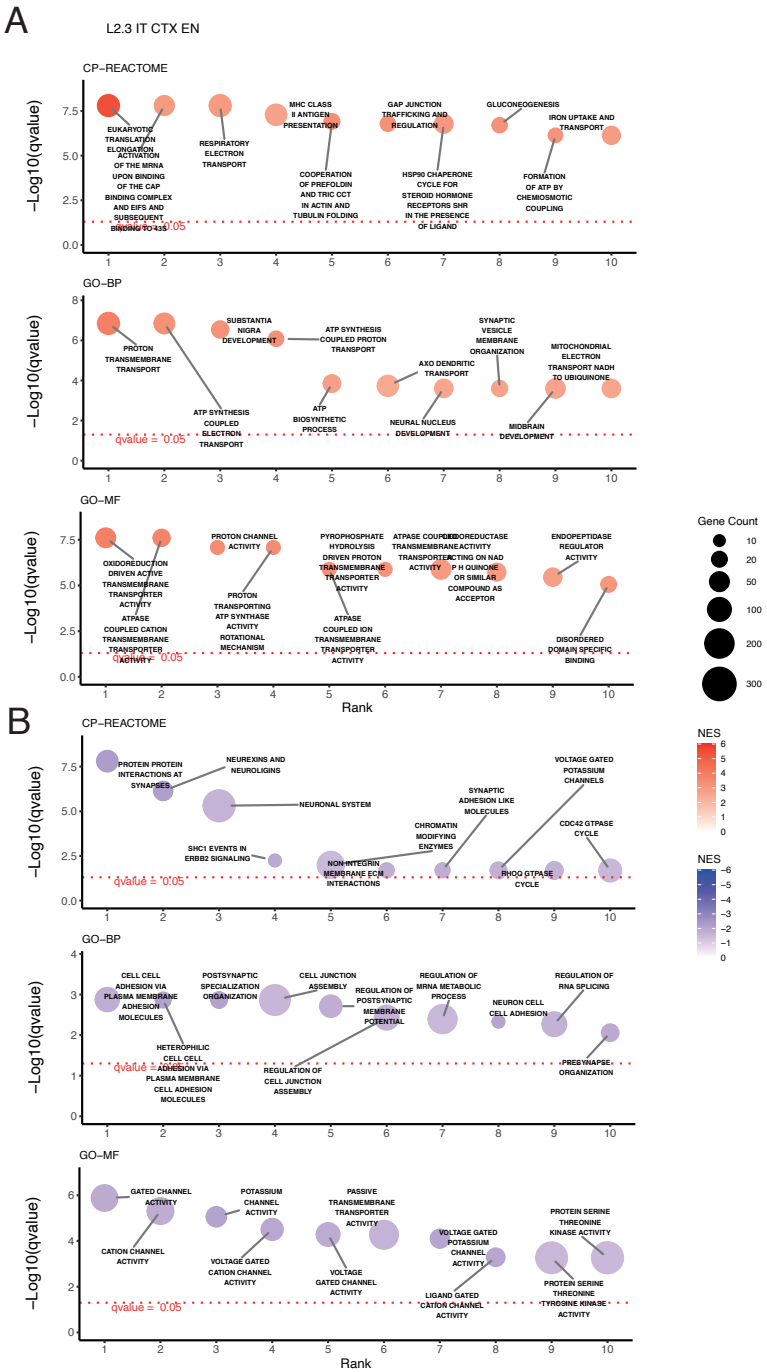


**Figure S16. Top deregulated gene-sets in layer 2/3 excitatory neurons from wt and DKO treated with lithium identified with GSEA.** (A) Ten most significantly upregulated and (B) downregulated gene-sets from the pathway databases reactome (top), the gene ontology biological process (GO-BP) (middle), and the gene ontology molecular function (GO-MF) (bottom) subcollections. Upregulated gene-sets are mainly associated with metabolism and mitochondrial ATP-synthesis. Down-regulated gene-sets are mainly associated with the synapse and cation-channel activities.

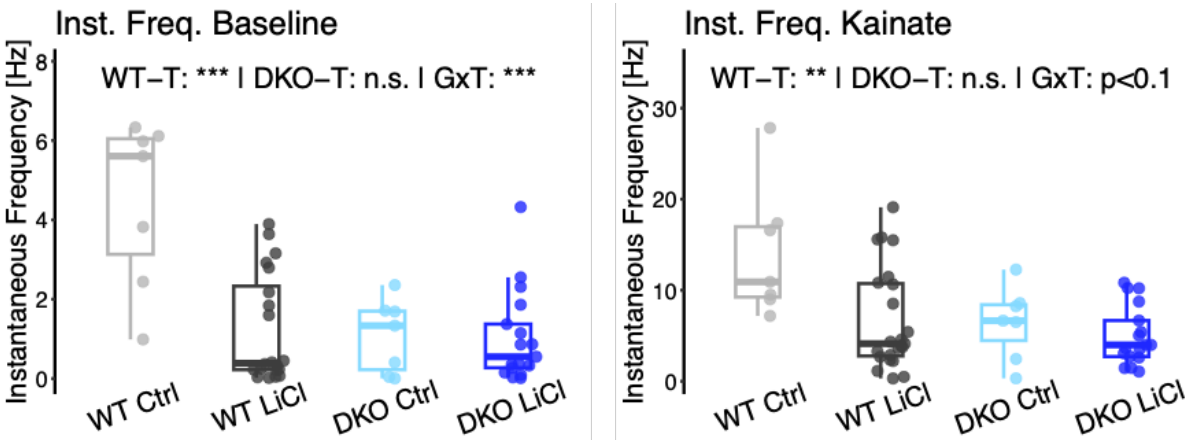


**Figure S17. Reduced lithium responsiveness of neuronal activity under baseline and upon kainic acid induced hyperexcitation.** Current clamp recordings of action potentials in layer 2/3 slices of the ACC isolated from untreated (Ctrl) and chronically lithium treated (LiCl) WT and DKO animals. Plotted are the instantaneous frequencies as a measure of neuronal activity. ** P < 0.01; *** P < 0.001; n.s. not significant; P-values refer to unifactorial ANOVA procedure; WT: wildtype mice; DKO: Bhlhe40/41 double-knockout mice; Ctrl: vehicle control; LiCl: lithium-treated; simple; WT-T: WT simple treatment effect; DKO-T: DKO simple treatment effect; GxT: interaction effect.

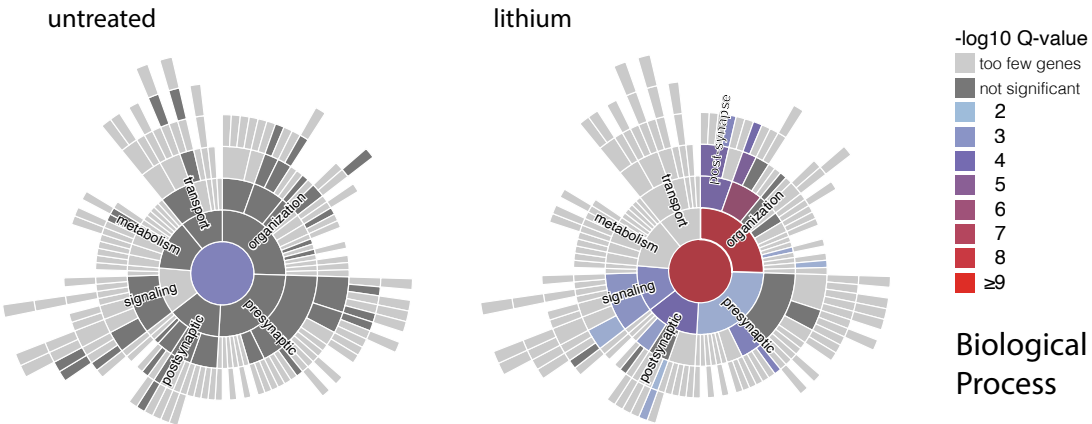


**Figure S18. SynGO enriched gene-sets in layer 2/3 excitatory neurons from untreated and lithium treated wt and DKO scRNAseq ACC samples.** Enrichment of synaptic gene-sets from the SynGO reveals most significant effects in lithium treated animals in biological processes associated with synaptic organization and post-synaptic functions.

# Supplemental References


1. McCarthy, M. J., Nievergelt, C. M., Kelsoe, J. R. & Welsh, D. K. A survey of genomic studies supports association of circadian clock genes with bipolar disorder spectrum illnesses and lithium response. *PloS One* **7**, e32091 (2012).
2. Byrne, E. M. *et al.* Testing the role of circadian genes in conferring risk for psychiatric disorders. *Am. J. Med. Genet. Part B Neuropsychiatr. Genet. Off. Publ. Int. Soc. Psychiatr. Genet.* **165B**, 254–260 (2014).
3. Takahashi, J. S. Transcriptional architecture of the mammalian circadian clock. *Nat. Rev. Genet.* **18**, 164–179 (2017).
4. Zhang, E. E. *et al.* A genome-wide RNAi screen for modifiers of the circadian clock in human cells. *Cell* **139**, 199–210 (2009).
5. Yan, J., Wang, H., Liu, Y. & Shao, C. Analysis of gene regulatory networks in the mammalian circadian rhythm. *PLoS Comput. Biol.* **4**, e1000193 (2008).
6. Pizarro, A., Hayer, K., Lahens, N. F. & Hogenesch, J. B. CircaDB: a database of mammalian circadian gene expression profiles. *Nucleic Acids Res.* **41**, D1009-1013 (2013).
7. Li, J. Z. *et al.* Circadian patterns of gene expression in the human brain and disruption in major depressive disorder. *Proc. Natl. Acad. Sci. U. S. A.* **110**, 9950–9955 (2013).
8. Chen, C.-Y. *et al.* Effects of aging on circadian patterns of gene expression in the human prefrontal cortex. *Proc. Natl. Acad. Sci. U. S. A.* **113**, 206–211 (2016).
9. Seney, M. L. *et al.* Diurnal rhythms in gene expression in the prefrontal cortex in schizophrenia. *Nat. Commun.* **10**, 3355 (2019).
10. Ketchesin, K. D. *et al.* Diurnal rhythms across the human dorsal and ventral striatum. *Proc. Natl. Acad. Sci. U. S. A.* **118**, e2016150118 (2021).
11. Ruben, M. D. *et al.* A database of tissue-specific rhythmically expressed human genes has potential applications in circadian medicine. *Sci. Transl. Med.* **10**, eaat8806 (2018).
12. Trubetskoy, V. *et al.* Mapping genomic loci implicates genes and synaptic biology in schizophrenia. *Nature* **604**, 502–508 (2022).
13. Schizophrenia Working Group of the Psychiatric Genomics Consortium. Biological insights from 108 schizophrenia-associated genetic loci. *Nature* **511**, 421–427 (2014).
14. Nievergelt, C. M. *et al.* International meta-analysis of PTSD genome-wide association studies identifies sex- and ancestry-specific genetic risk loci. *Nat. Commun.* **10**, 4558 (2019).
15. Wray, N. R. *et al.* Genome-wide association analyses identify 44 risk variants and refine the genetic architecture of major depression. *Nat. Genet.* **50**, 668–681 (2018).
16. Mullins, N. *et al.* Genome-wide association study of more than 40,000 bipolar disorder cases provides new insights into the underlying biology. *Nat. Genet.* **53**, 817–829 (2021).
17. Grove, J. *et al.* Identification of common genetic risk variants for autism spectrum disorder. *Nat. Genet.* **51**, 431–444 (2019).
18. Demontis, D. *et al.* Discovery of the first genome-wide significant risk loci for attention deficit/hyperactivity disorder. *Nat. Genet.* **51**, 63–75 (2019).
19. Jones, S. E. *et al.* Genome-Wide Association Analyses in 128,266 Individuals Identifies New Morningness and Sleep Duration Loci. *PLoS Genet.* **12**, e1006125 (2016).
20. Hou, L. *et al.* Genetic variants associated with response to lithium treatment in bipolar disorder: a genome-wide association study. *The Lancet* **387**, 1085–1093 (2016).
21. Jones, S. E. *et al.* Genome-wide association analyses of chronotype in 697,828 individuals provides insights into circadian rhythms. *Nat. Commun.* **10**, 343 (2019).
22. Song, J. *et al.* Genome-wide association study identifies SESTD1 as a novel risk gene for lithium-responsive bipolar disorder. *Mol. Psychiatry* **21**, 1290–1297 (2016).
23. Grof, P. *et al.* Is response to prophylactic lithium a familial trait? *J. Clin. Psychiatry* **63**, 942–947 (2002).

24. Manchia, M. *et al.* Assessment of Response to Lithium Maintenance Treatment in Bipolar Disorder: A Consortium on Lithium Genetics (ConLiGen) Report. *PloS One* **8**, e65636 (2013).

25. de Leeuw, C. A., Mooij, J. M., Heskes, T. & Posthuma, D. MAGMA: generalized gene-set analysis of GWAS data. *PLoS Comput. Biol.* **11**, e1004219 (2015).

26. Lee, P. H., O'Dushlaine, C., Thomas, B. & Purcell, S. M. INRICH: interval-based enrichment analysis for genome-wide association studies. *Bioinforma. Oxf. Engl.* **28**, 1797–1799 (2012).

27. de Leeuw, C. A., Neale, B. M., Heskes, T. & Posthuma, D. The statistical properties of gene-set analysis. *Nat. Rev. Genet.* **17**, 353–364 (2016).

28. Bulik-Sullivan, B. K. *et al.* LD Score regression distinguishes confounding from polygenicity in genome-wide association studies. *Nat. Genet.* **47**, 291–295 (2015).

29. Finucane, H. K. *et al.* Partitioning heritability by functional annotation using genome-wide association summary statistics. *Nat. Genet.* **47**, 1228–1235 (2015).

30. Chang, C. C. *et al.* Second-generation PLINK: rising to the challenge of larger and richer datasets. *GigaScience* **4**, 7 (2015).

31. Watanabe, K., Taskesen, E., van Bochoven, A. & Posthuma, D. Functional mapping and annotation of genetic associations with FUMA. *Nat. Commun.* **8**, 1826 (2017).

32. Hurst, J. L. & West, R. S. Taming anxiety in laboratory mice. *Nat. Methods* **7**, 825–826 (2010).

33. Hurst, J. L. & West, R. S. Taming anxiety in laboratory mice. *Nat. Methods* **7**, 825–826 (2010).

34. Volkmann, P., Stephan, M., Krackow, S., Jensen, N. & Rossner, M. J. PsyCoP - A Platform for Systematic Semi-Automated Behavioral and Cognitive Profiling Reveals Gene and Environment Dependent Impairments of Tcf4 Transgenic Mice Subjected to Social Defeat. *Front. Behav. Neurosci.* **14**, 618180 (2020).

35. Hühne, A., Volkmann, P., Stephan, M., Rossner, M. & Landgraf, D. An in-depth neurobehavioral characterization shows anxiety-like traits, impaired habituation behavior, and restlessness in male Cryptochrome-deficient mice. *Genes Brain Behav.* **19**, e12661 (2020).

36. Stephan, M. *et al.* Spironolactone alleviates schizophrenia-related reversal learning in Tcf4 transgenic mice subjected to social defeat. *Schizophr. Heidelb. Ger.* **8**, 77 (2022).

37. Team, R. C. *R: A Language and Environment for Statistical Computing*. (R Foundation for Statistical Computing, Vienna, Austria, 2022).

38. Wickham, H. Introduction. in *ggplot2: Elegant Graphics for Data Analysis* (ed. Wickham, H.) 1–7 (Springer, New York, NY, 2009).

39. Sarkar, D. *Lattice: Multivariate Data Visualization with R*. (Springer, New York, NY, 2008).

40. Kolde, R. Pheatmap: Pretty Heatmaps (version 1.0.12).https://CRAN.R-project.org/package=pheatmap. (2019).

41. Fox, J. & Weisberg, S. *Candisc: Visualizing Generalized Canonical Discriminant and Canonical Correlation Analysis (Version 0.8-6). Https://CRAN.R-Project.Org/Package=candisc.* (2019).

42. Dray, S. & Dufour, A.-B. The ade4 Package: Implementing the Duality Diagram for Ecologists. *J. Stat. Softw.* **22**, 1–20 (2007).

43. Friendly, M. & Fox, J. *Candisc: Visualizing Generalized Canonical Discriminant and Canonical Correlation Analysis (Version 0.8-6). Https://CRAN.R-Project.Org/Package=candisc.* (2021).

44. Anderzhanova, E., Kirmeier, T. & Wotjak, C. T. Animal models in psychiatric research: The RDoC system as a new framework for endophenotype-oriented translational neuroscience. *Neurobiol. Stress* **7**, 47–56 (2017).

45. Baier, P. C. *et al.* Mice Lacking the Circadian Modulators SHARP1 and SHARP2 Display Altered Sleep and Mixed State Endophenotypes of Psychiatric Disorders. *PLoS ONE* **9**, (2014).

46. Honma, S. *et al.* Dec1 and Dec2 are regulators of the mammalian molecular clock. *Nature* **419**, 841–844 (2002).

47. Romagnoli, D. *et al.* ddSeeker: a tool for processing Bio-Rad ddSEQ single cell RNA-seq data. *BMC Genomics* **19**, 960 (2018).

48. Dobin, A. *et al.* STAR: ultrafast universal RNA-seq aligner. *Bioinforma. Oxf. Engl.* **29**, 15–21 (2013).

49. Hao, Y. *et al.* Integrated analysis of multimodal single-cell data. *Cell* **184**, 3573-3587.e29 (2021).

50. Fleming, S. J. *et al.* Unsupervised removal of systematic background noise from droplet-based single-cell experiments using CellBender. *Nat. Methods* **20**, 1323–1335 (2023).

51. Yang, S. *et al.* Decontamination of ambient RNA in single-cell RNA-seq with DecontX. *Genome Biol.* **21**, 57 (2020).

52. Bais, A. S. & Kostka, D. scds: computational annotation of doublets in single-cell RNA sequencing data. *Bioinforma. Oxf. Engl.* **36**, 1150–1158 (2020).

53. Hafemeister, C. & Satija, R. Normalization and variance stabilization of single-cell RNA-seq data using regularized negative binomial regression. *Genome Biol.* **20**, 296 (2019).

54. McInnes, L., Healy, J. & Melville, J. UMAP: Uniform Manifold Approximation and Projection for Dimension Reduction. Preprint at https://doi.org/10.48550/arXiv.1802.03426 (2020).

55. McDavid, A., Finak, G. & Yajima, M. Model-based Analysis of Single Cell Transcriptomics. https://rglab.github.io/MAST/ (2020).

56. Zhou, Y. *et al.* Metascape provides a biologist-oriented resource for the analysis of systems-level datasets. *Nat. Commun.* **10**, 1523 (2019).

57. Martin, A. *et al.* BisoGenet: a new tool for gene network building, visualization and analysis. *BMC Bioinformatics* **11**, 91 (2010).

58. Shannon, P. *et al.* Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res.* **13**, 2498–2504 (2003).

59. Wu, T. *et al.* clusterProfiler 4.0: A universal enrichment tool for interpreting omics data. *Innov. Camb. Mass* **2**, 100141 (2021).

60. Subramanian, A. *et al.* Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc. Natl. Acad. Sci. U. S. A.* **102**, 15545–15550 (2005).

61. Dolgalev, I. msigdbr: MSigDB Gene Sets for Multiple Organisms in a Tidy Data Format_. R package version 7.5.1. https://CRAN.R-project.org/package=msigdbr (2022).

62. Koopmans, F. *et al.* SynGO: An Evidence-Based, Expert-Curated Knowledge Base for the Synapse. *Neuron* **103**, 217-234.e4 (2019).

63. Herholt, A. *et al.* Pathway sensor-based functional genomics screening identifies modulators of neuronal activity. *Sci. Rep.* **8**, 17597 (2018).

64. Teng, Z., Zhang, M., Zhao, M. & Zhang, W. Glucocorticoid exerts its non-genomic effect on IPSC by activation of a phospholipase C-dependent pathway in prefrontal cortex of rats. *J. Physiol.* **591**, 3341–3353 (2013).

65. Hu, W., Zhang, M., Czéh, B., Flügge, G. & Zhang, W. Stress Impairs GABAergic Network Function in the Hippocampus by Activating Nongenomic Glucocorticoid Receptors and Affecting the Integrity of the Parvalbumin-Expressing Neuronal Network. *Neuropsychopharmacology* **35**, 1693–1707 (2010).

66. Zhang, M., Weiland, H., Schöfbänker, M. & Zhang, W. Estrogen Receptors Alpha and Beta Mediate Synaptic Transmission in the PFC and Hippocampus of Mice. *Int. J. Mol. Sci.* **22**, 1485 (2021).

67. Agarwal, A. *et al.* Dysregulated Expression of Neuregulin-1 by Cortical Pyramidal Neurons Disrupts Synaptic Plasticity. *Cell Rep.* **8**, 1130–1145 (2014).

68. Saffari, R. *et al.* NPY+-, but not PV+-GABAergic neurons mediated long-range inhibition from infra- to prelimbic cortex. *Transl. Psychiatry* **6**, e736–e736 (2016).

69. Wehr, M. C. *et al.* Spironolactone is an antagonist of NRG1-ERBB4 signaling and schizophrenia-relevant endophenotypes in mice. *EMBO Mol. Med.* **9**, 1448–1462 (2017).

70. Zhou, Y. *et al.* Metascape provides a biologist-oriented resource for the analysis of systems-level datasets. *Nat. Commun.* **10**, 1523 (2019).

783