Dynamic regulation of subcellular transcriptomes in cortical projection neurons

Supplementary Table 1: List of GO terms differentially enriched in either soma or GC compartments for union of CPN and CThPN

Supplementary Table 2: Results of differential expression analysis at gene and transcript level, comparing expression across subtypes of cortical projection neurons, compartments, and across developmental stages for CPN.

Supplementary Table 3: Cross-referencing subcellular differential gene expression data of this report for CPN and CThPN with previously published, data sets of genes associated with neurodevelopmental and neuropsychiatric disorders.

Supplementary Table 4: List of GO terms differentially enriched in GC compartments of distinct subtypes of cortical projection neurons.
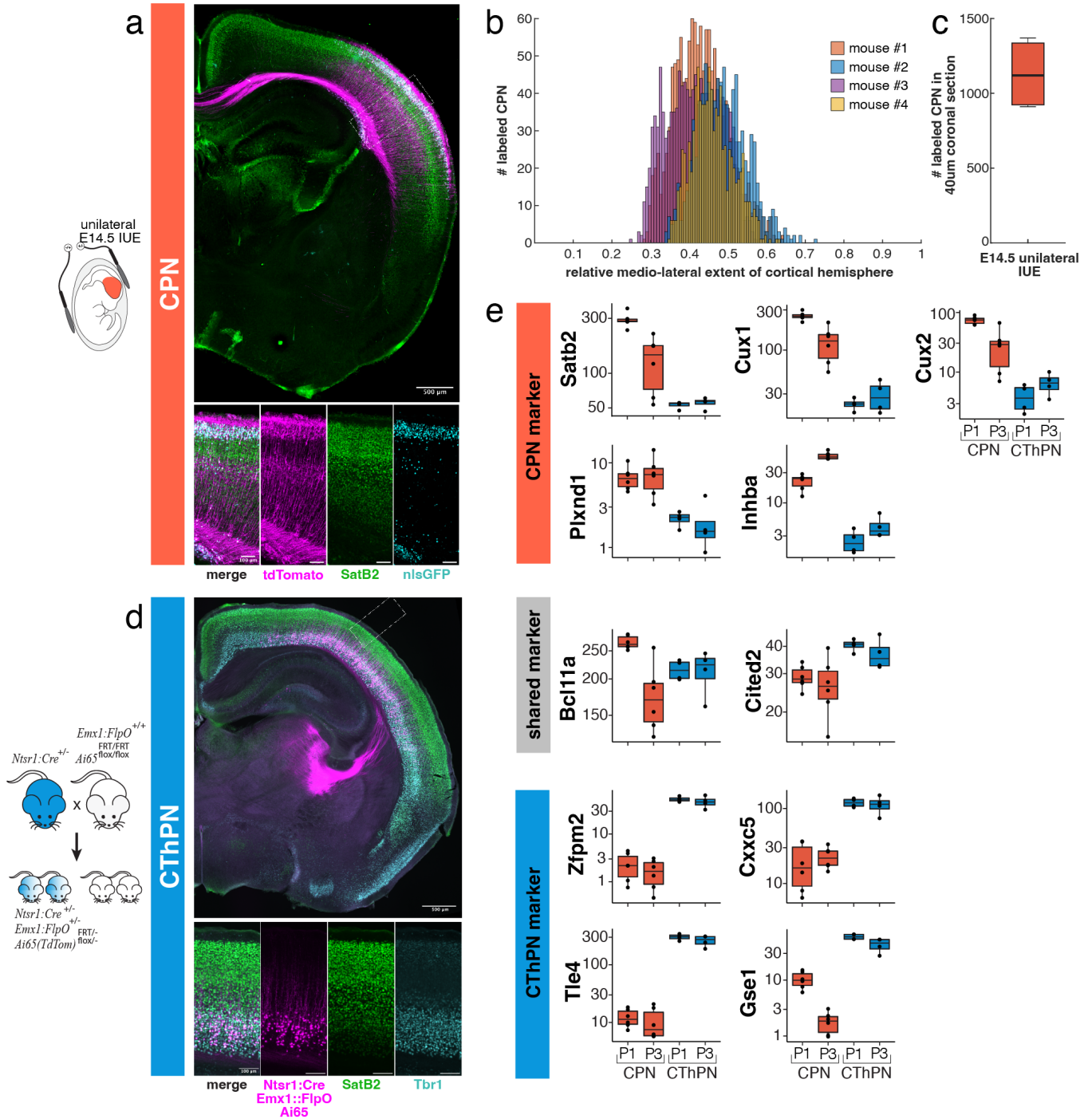
Figure S1: Subtype-specific labeling of CPN and CThPN. (a) *In utero* electroporation at embryonic day E14.5 in mice specifically labels CPN of cortical layer II/III. GFP-positive nuclei of CPN are localized mainly to cortical layer II, with limited labeling present in layer III, and are colocalized with the superficial layer marker SatB2 (inset). The multi-cistronically expressed, membrane-targeted tdTomato reveals the entire extent of CPN dendritic processes and axonal projections spanning all cortical layers, as well as across the corpus callosum. **(b)** Unilateral *in utero* electroporation at embryonic day E14.5 results in highly reproducible targeting of CPN in cortical layer II/III in the lateral somatosensory cortex. **(c)** Total number of labeled CPN in a single 40 μm coronal brain section is comparable across mice. **(d)** Intersectional mouse genetics (Ntsr1::Cre+/-, Emx1::FlpO+/-, Ai65(TdTom)FRT/wt, flox/wt) specifically labels CThPN of cortical layer VI. Labeled CThPN projections transverse the internal capsule and innervate the thalamus. TdTom-positive CThPN are colocalized with CThPN marker TBR1 but are negative for superficial layer marker SatB2 (inset). **(e)** RNA expression levels of subtype marker genes from flow cytometry sorted CPN and CThPN somata at developmental ages P1 and P3. CPN have markedly higher expression levels in superficial layer markers, whereas CThPN have appropriate enrichment for layer VI marker genes, as expected. Genes common to both subtypes are expressed at similar levels.
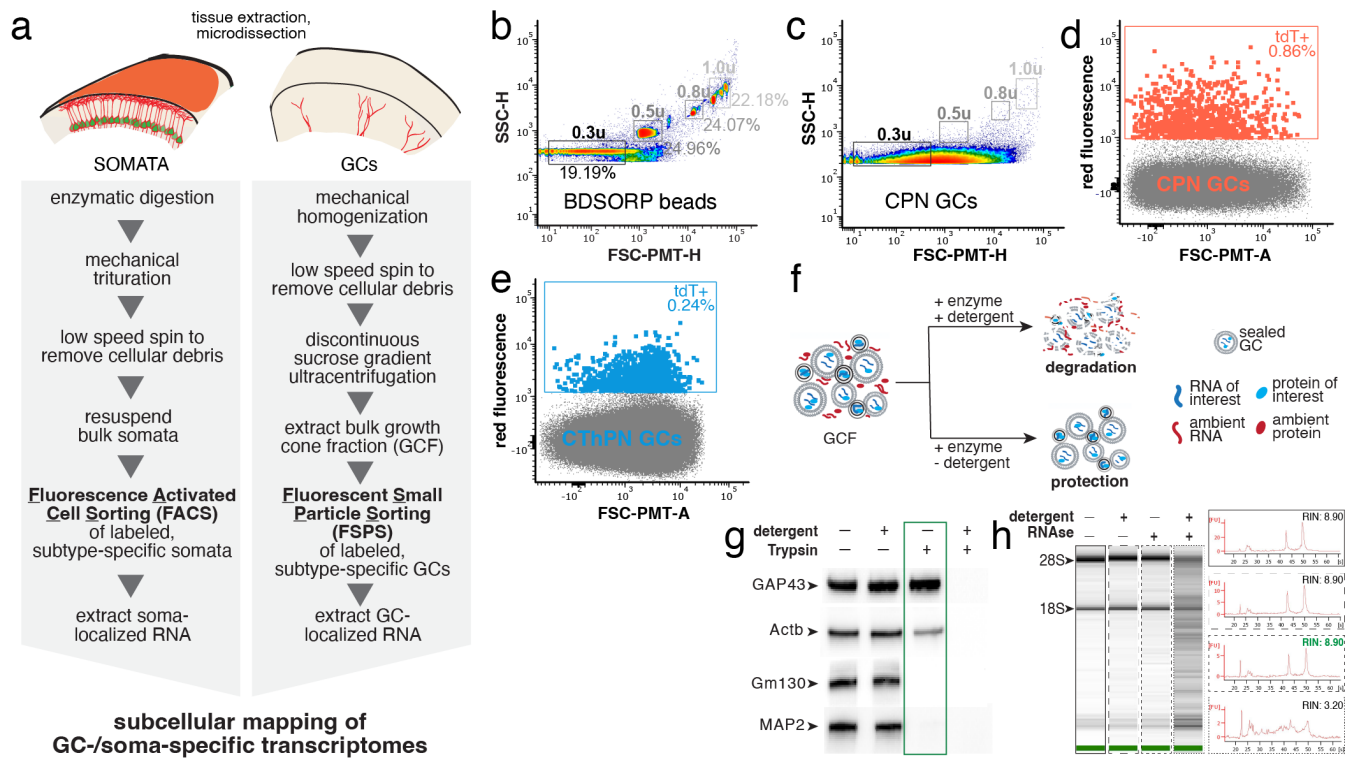
Figure S2: Parallel isolation of subtype-specific somata and GCs from the mouse neocortex. (a) Workflow enabling parallel extraction and subtype-specific sorting of CPN- or CThPN-specific somata and GCs. RNA isolated from respective samples sorter-purified by flow cytometry is then used for subcellular mapping of subtype-specific transcriptomes. A detailed description of this protocol can be found in[50]. **(b)** Side scatter over forward scatter data for size-calibrated BD SORP beads run on a flow cytometer in fluorescent small particle sorting (FSPS) configuration. FSPS enables detection, analysis, and efficient sorting of particles in the 100nm-1 μm diameter range. **(c)** Analysis of CPN GCs overlaid on top of BD SORP beads depicted in detail in panel (b). CPN GCs have diameters of ~200-800 nm, predominantly ~300-500 nm. **(d/e)** FSPS enables purification of subtype-specifically labeled GCs of CPN (d) and CThPN (f) from respective GCF samples. **(f)** Hydrolysis protection assays are used to assess the presence of intact GC particles with intact membranes, thereby protecting their specific molecular contents from RNase or proteinase digestion in the absence of detergents. **(g)** Western blot of representative proteinase protection assay of CPN GCF detected for multiple subcellular markers. Likely ambient contaminations like GM130 or Map2 are degraded in the absence of detergents (-/+), while GC markers such as GAP43 or Actin are preserved in the absence of detergents (-/+), and are only degraded once GC membranes are permeabilized by detergent (+/+). **(h)** Tape station results of representative RNase protection assay of CPN GC. When the sample is treated with RNase, RNA quality is preserved in the absence of detergents (-/+), and only upon permeabilization of membranes does pronounced RNA degradation occur (+/+).
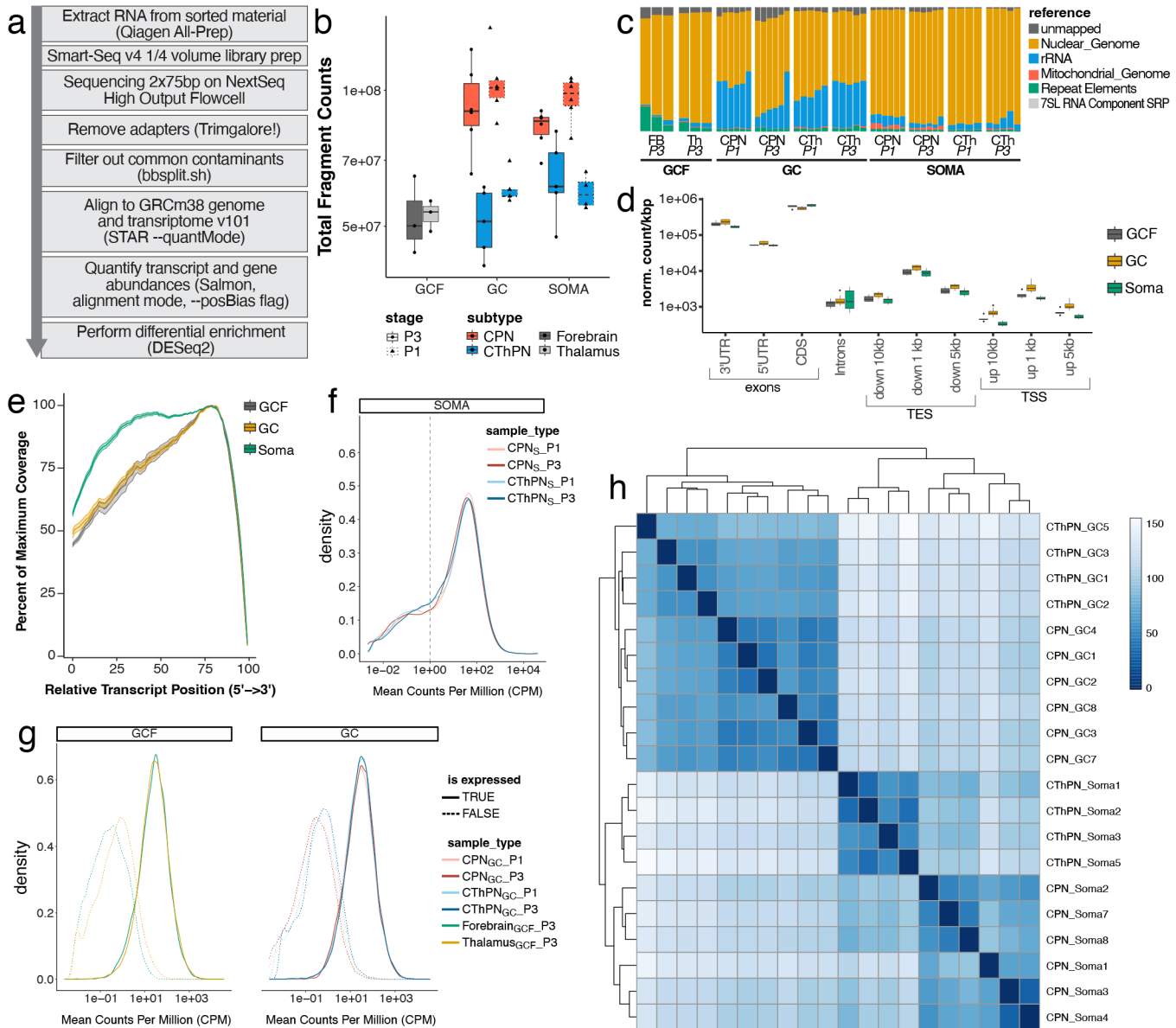
Figure S3: Subcellular RNA sequencing passes quality metrics and samples are highly correlated within compartments and subtypes. (a) Schematic of bioinformatics workflow to process transcriptomic data from reads to differential expression between samples. **(b)** Sequencing depth is >40M fragments for all libraries across subtypes and stages, as well as for background samples (GC fraction pre-sort, GCF). **(c)** Alignment of reads to contaminating features and the nuclear genome indicates that most reads, as expected, originate from the nuclear genome and that the sorted GC samples have a higher proportion of reads aligning to rRNAs. **(d)** Distribution of reads to genomic features shows that most reads align to UTRs and coding sequences (CDSs), and not genomic regions downstream of transcription end site (TES) or upstream of transcription start site (TSS), as expected; GC and GCF samples have higher representation in 3'UTRs. **(e)** The coverage over 3' ends is higher in GC and GCF samples relative to somata. **(f)** A cutoff of 2 counts per million (CPM) was used for soma samples, since this is the inflection point between signal and noise CPM distributions. **(g)** A CPM cutoff for sorted GC and GCF samples was defined as the 95th quantile of the distribution of CPMs of non-expressed genes (dotted lines), which were determined by the soma cutoff in panel (f). **(h)** Heatmap showing correlation between normalized transcript counts of all pairs of samples.

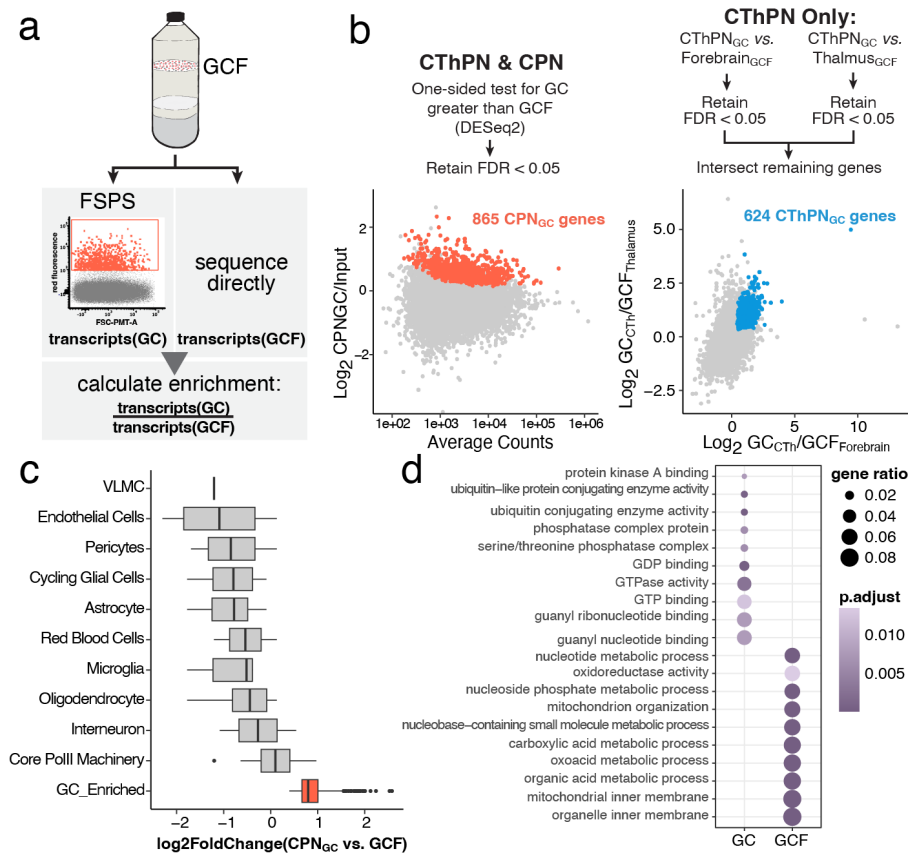*Dynamic regulation of subcellular transcriptomes in cortical projection neurons*

Figure S4: Transcriptome obtained from GC samples is corrected to minimize effect of potential ambient contamination. (a) To minimize effects of potential ambient contamination in the GC preparation, transcriptomes obtained from sorted GC samples are corrected for their respective input material, i.e. RNA extracted from the unsorted GCF. **(b)** GC transcripts for CPN and CThPN are defined as those having a higher abundance in sorted GCs than in GCF. **(c)** Correction for input material results in exclusion of likely ambient contaminants, e.g. marker genes for various non-neuronal cell types (grey), while preserving a class of robustly GC transcripts (red). **(d)** GO term enrichment comparing GC genes to likely contaminants from the GCF.
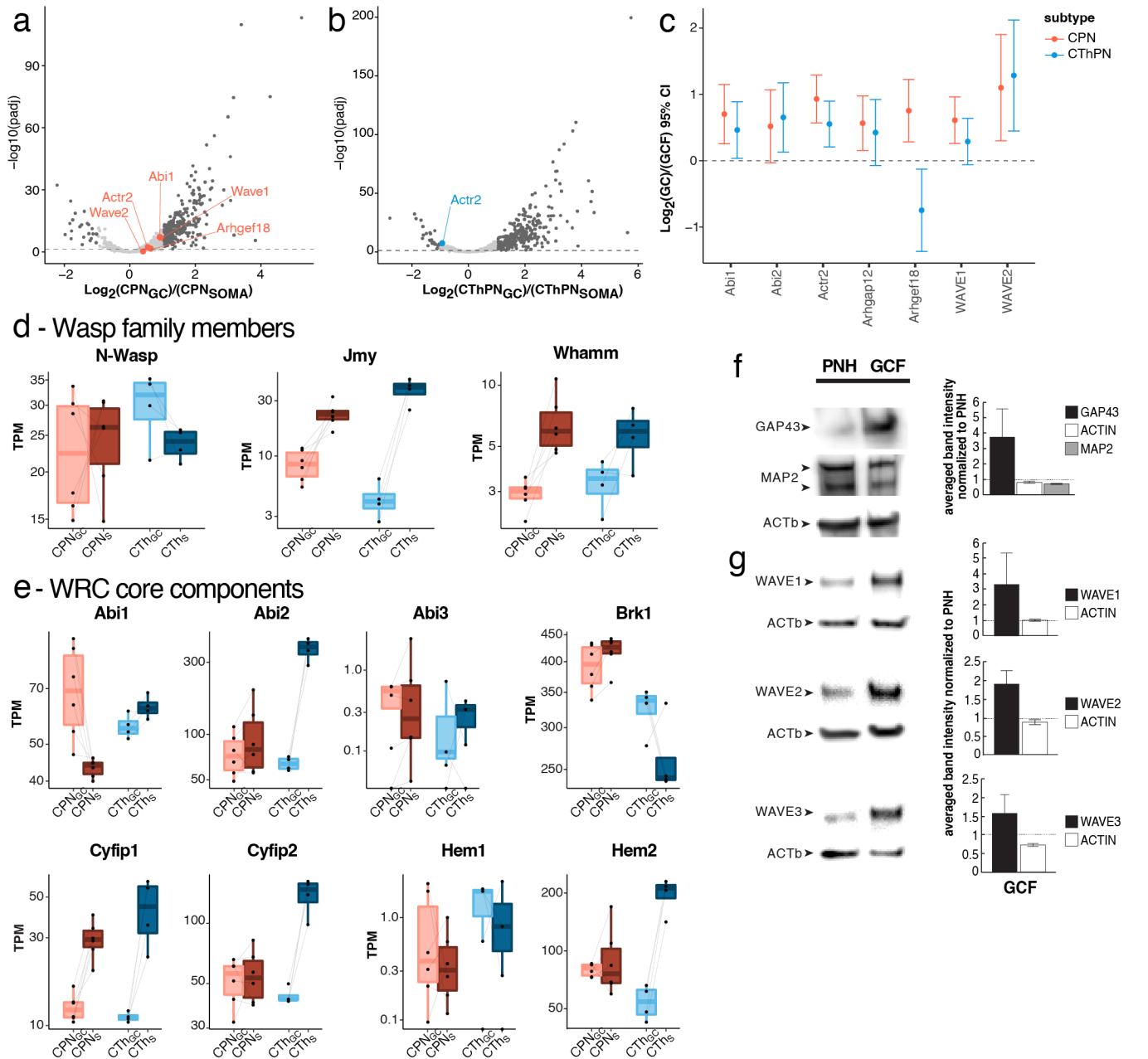
Figure S5: Expression and subcellular transcript localization of WRC core components in CPN and CThPN. (a) Volcano plot indicating CPN-specific subcellular transcript localization of core components of the WRC as orange dots (FDR < 0.05). **(b)** Volcano plot indicating CThPN-specific subcellular transcript localization of core components of the WRC as blue dots (FDR < 0.05). **(c)** 95% confidence intervals of log2 fold change GC/GCF for each of the WRC-associated transcripts. **(d)** Transcript abundance (transcripts per million, TPM) of the non-WRC associated members of the Wasp gene family *N-Wasp*, *Jmy*, and *Whamm* in somata and GCs of CPN (pink/brown) and CThPN (light/dark blue). **(e)** Transcript abundance (TPM) of WRC core components in somata and GCs of CPN (pink/brown) and CThPN (light/dark blue). **(f)** Representative western blot images and quantification of triplicates for each protein's abundance of WAVE paralogs WAVE1, WAVE2, WAVE3, and ACTIN in the post nuclear homogenate (PNH) and growth cone fraction (GCF) of CPN. All three paralogs are enriched in the GCF when compared to the PNH.
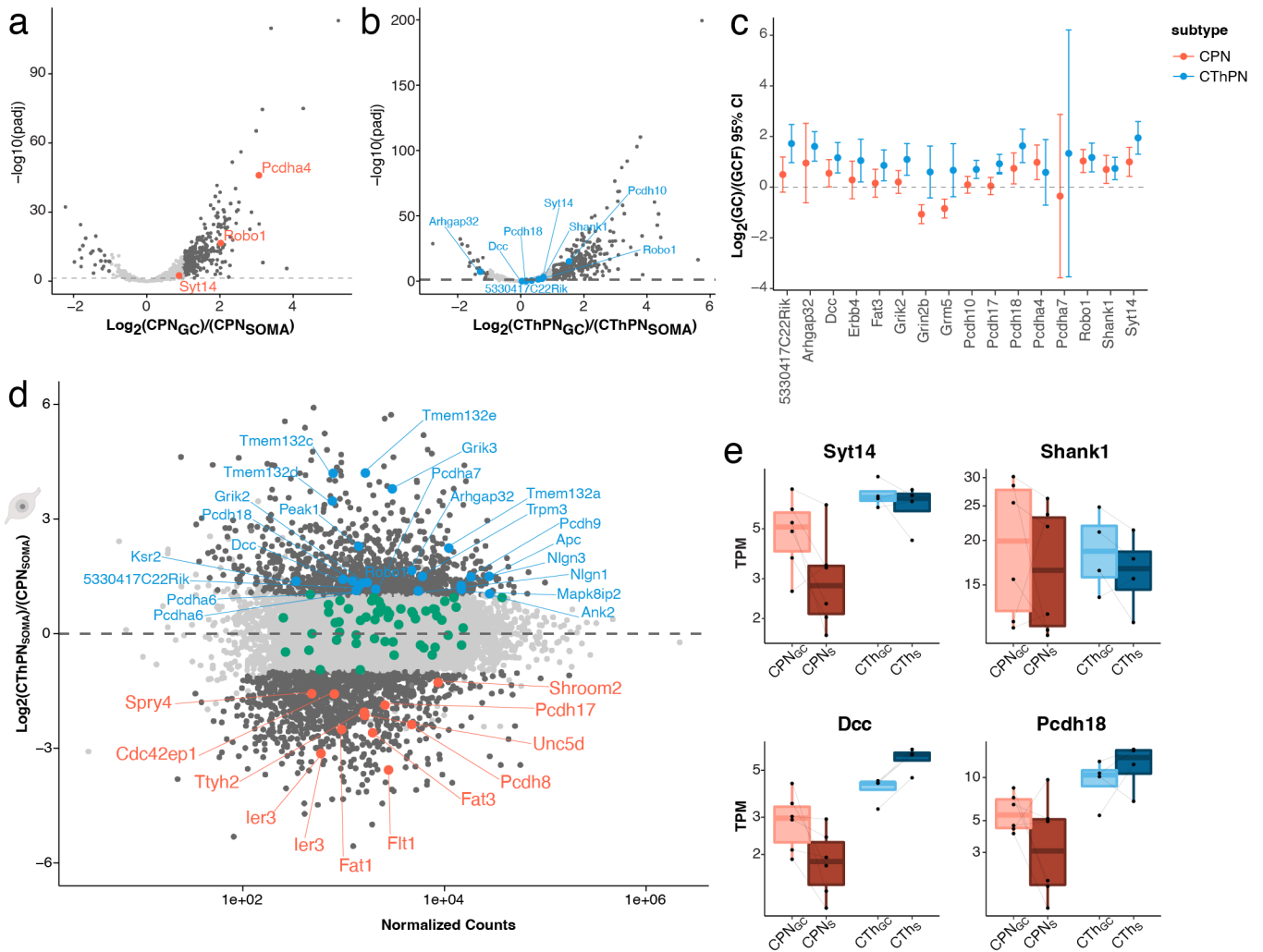
Figure S6: Expression and subcellular transcript localization of WIRS-containing receptors in CPN and CThPN. (a) Volcano plot indicating CPN-specific subcellular transcript localization of WIRS-containing receptors as orange dots (FDR < 0.05). **(b)** Volcano plot indicating CThPN-specific subcellular transcript localization of WIRS-containing receptors as blue dots (FDR < 0.05). **(c)** 95% confidence intervals of log2 fold change GC/GCF for each of the WIRS-containing transcripts. **(d)** Differential expression levels (FDR < 0.05, dark grey) of WIRS-containing genes in CPN (orange) vs. CThPN (blue) somata. WIRS-containing genes that are not differentially expressed between subtypes are shown in green. **(e)** Transcript abundance (transcripts per million, TPM) of examples of WIRS-containing receptors with GC-enriched transcript localization in GCs and somata of CPN (pink/brown) and CThPN (light/dark blue).

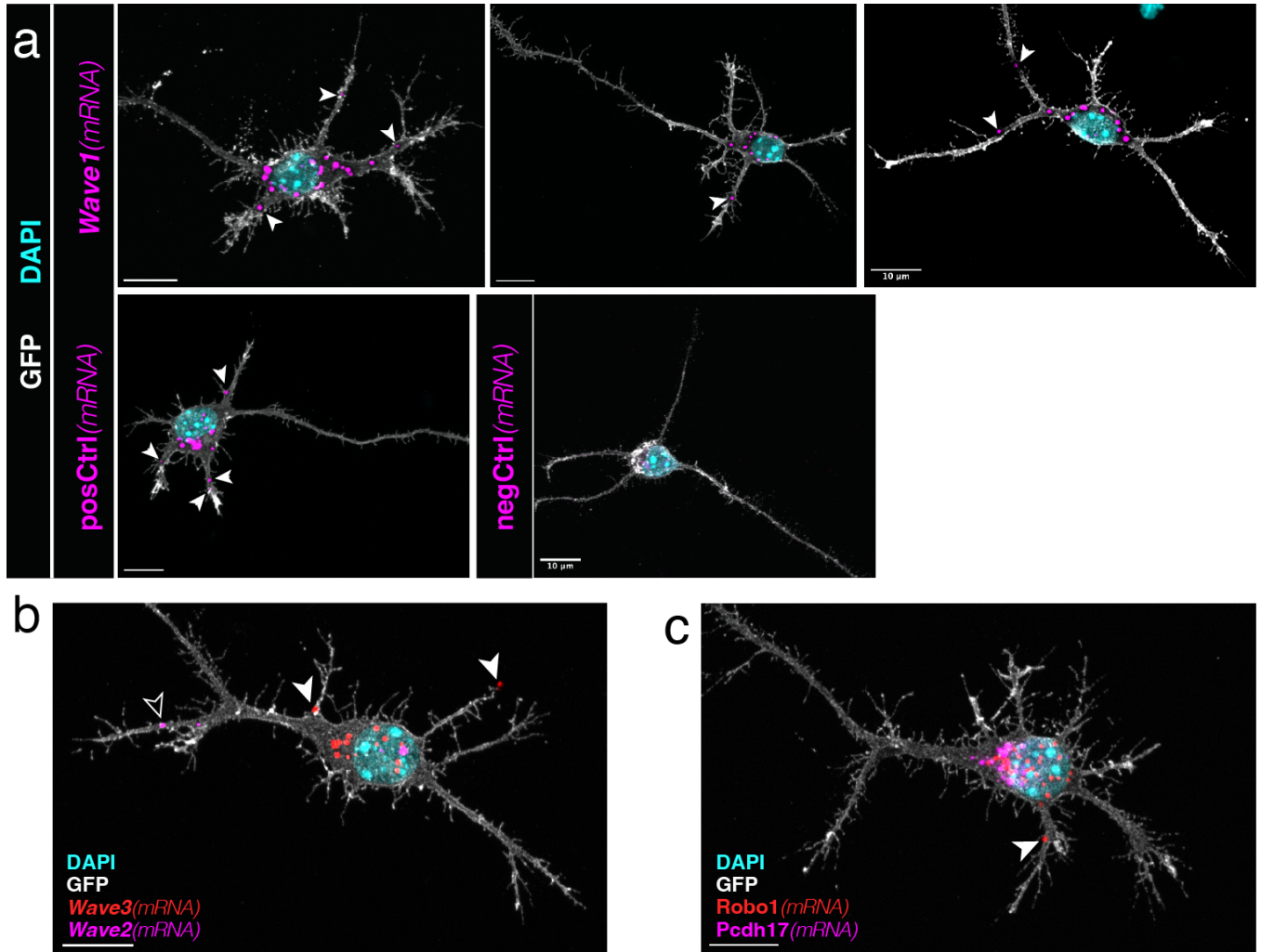*Dynamic regulation of subcellular transcriptomes in cortical projection neurons*

Figure S7: Single molecule FISH of *Wave1* and the WIRS containing receptors *Robo1* and *Pcdh17* in primary cultured CPN. (a) Maximum intensity projection (10 μm) of representative RNAscope confocal image of *Wave1*, the respective positive control (posCtrl, targeting *Ppib*), or negative control (negCtrl, targeting *DapB*) in GFP-positive cultured CPN. Solid arrowheads indicate RNAscope puncta localized to proximal or more distal neurites. **(b)** Maximum intensity projection (10 μm) of representative RNAscope confocal image of the *Wave* paralogs *Wave2* (magenta) and *Wave3* (red) in GFP-positive cultured CPN. Solid arrowheads indicate *Wave3* RNAscope puncta, while open arrow heads indicate *Wave2* RNAscope puncta localized to proximal or more distal neurites. **(c)** Maximum intensity projection (10 μm) of representative RNAscope confocal image of the WIRS receptors *Pcdh17* (magenta) and *Robo1* (red) in GFP-positive cultured CPN. Solid arrowhead indicates *Robo1* RNAscope puncta localized to proximal CPN neurite. Similar distribution of puncta were observed for 6-7 neurons per condition.

Dynamic regulation of subcellular transcriptomes in cortical projection neurons
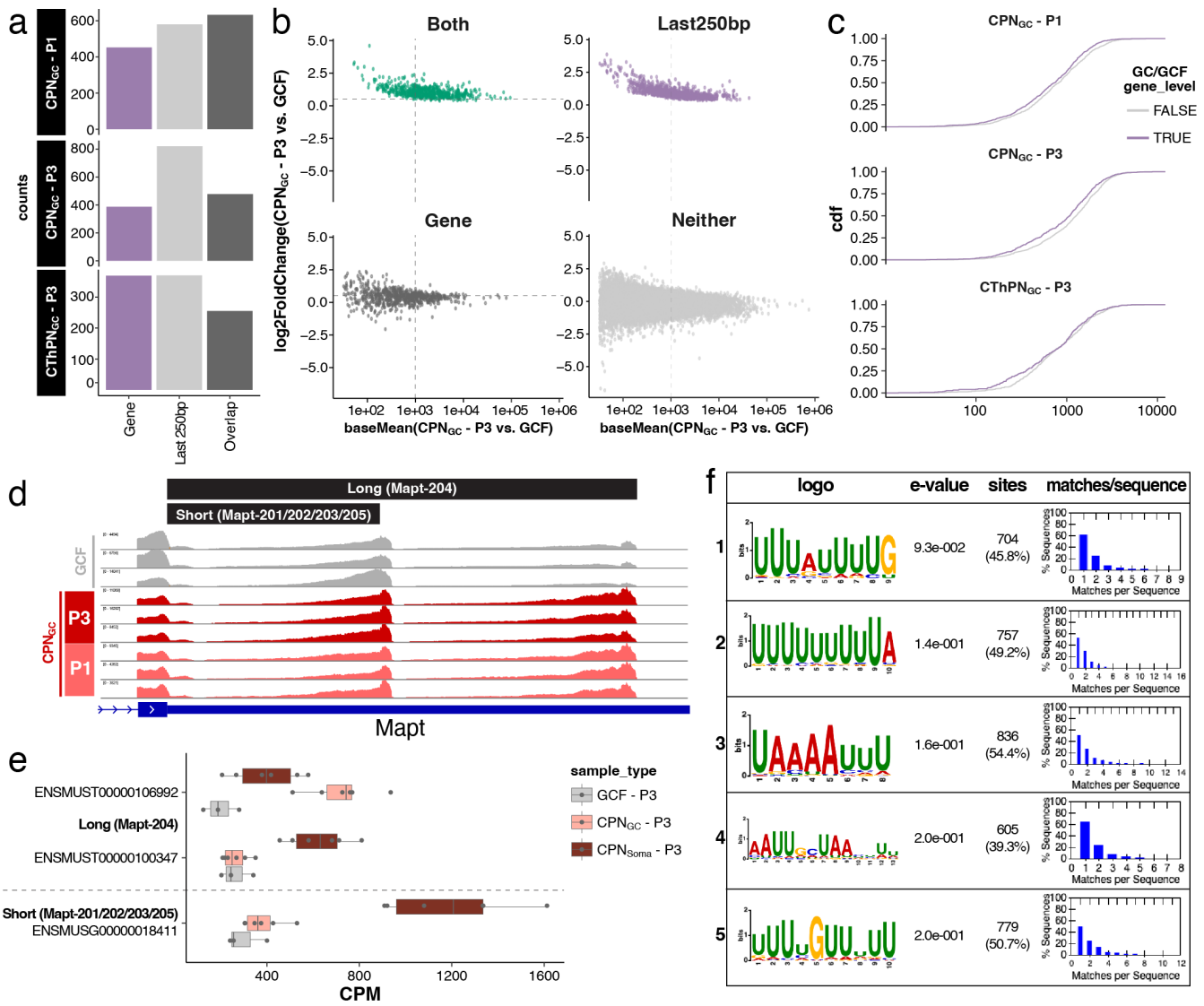

Figure S8: 3'UTR isoform analysis identifies enrichment of poly-U motifs in GC transcriptomes over those of somata. (a) Number of genes called by the gene-level approach for quantifying GC genes, last-250bp approach for quantifying GC 3'UTRs, or both. **(b)** MA plots depicting the log2 fold change and abundance of GC transcripts identified by the last-250bp approach. Each plot represents a subset of all transcripts: genes that are enriched in the last-250bp (purple), genes that were previously identified (black), isoforms of genes that were identified by both approaches (green), and as a control, isoforms of genes that were identified in neither approach (grey). Genes identified by only the gene-level approach (black) are generally of lower abundance. Last-250bp approach (purple) enables identification of GC-specific isoform variants of higher abundance genes. **(c)** GC 3'UTRs identified only in the last-250bp approach are on average longer (cdf – cumulative distribution function) (d) As an example, *Mapt* has one long 3'UTR enriched in GCs, but a shorter isoform not enriched in GCs, as seen by the read pileups; quantified in **(e)**. **(f)** De novo motif analysis of GC-enriched 3'UTRs identifies multiple distinct versions of poly-U rich motifs preferentially detected in the GC compartment.
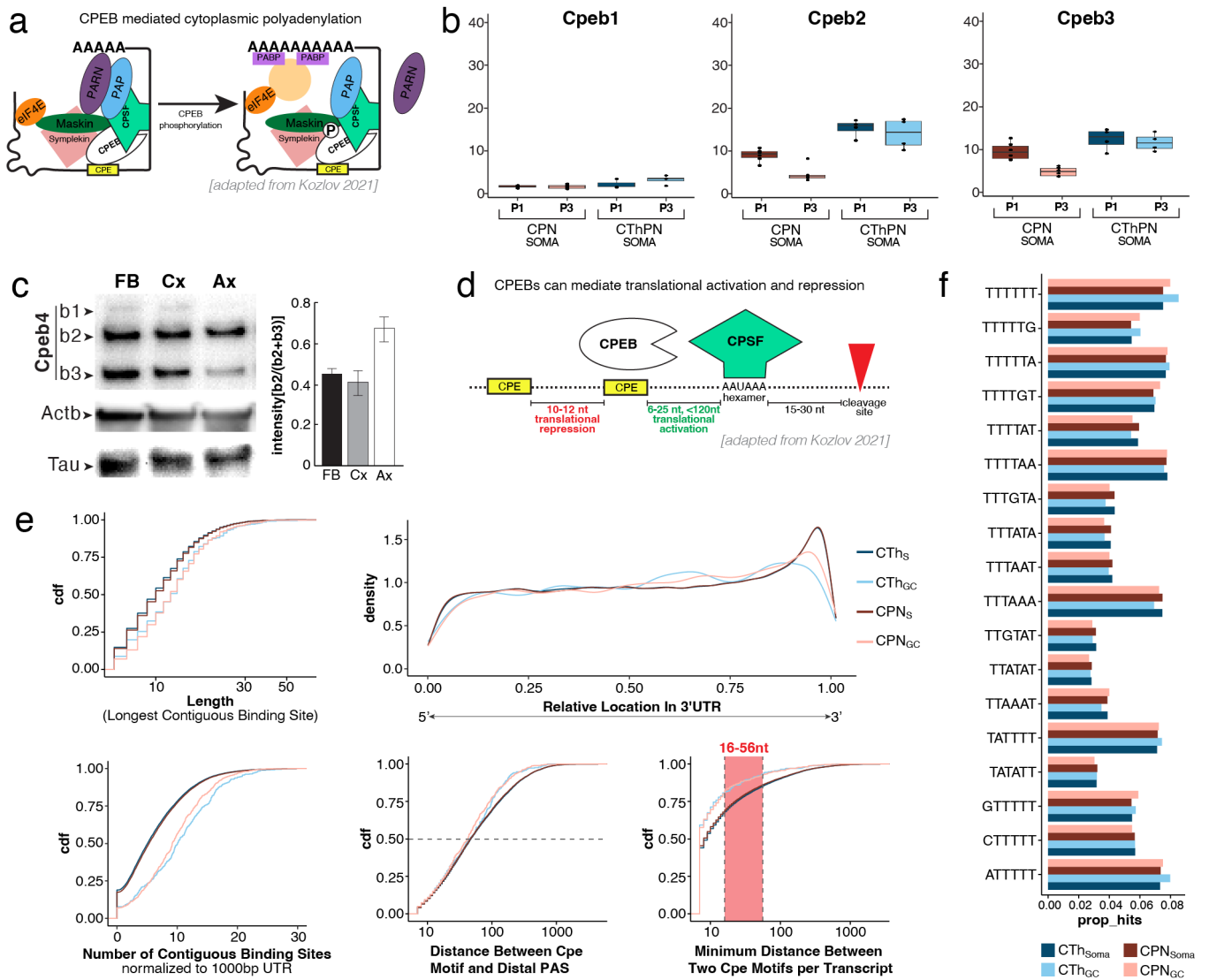
Figure S9: CPE location and spacing in GC 3'UTRs is consistent with baseline translational repression of targets. **(a)** CPEBs bind to the cytoplasmic polyadenylation element (CPE) located in the '3-UTR of transcripts and complex with factors that control polyA tail length and translation. Cue-induced CPEB phosphorylation mediates polyA polymerase-mediated lengthening of the poly-A tail in the cytosol, enabling eiF4E interaction with PABP, and thereby initiation of translational activation of the respective transcripts [adapted from[96]]. **(b)** Boxplots of expression levels in TPM of the three other *Cpeb* paralogs in CPN (pink/orange) and CThPN (light/dark blue) somata at P1 (brown, dark blue) and P3 (pink, light blue). **(c)** Representative western blot of protein samples obtained from whole forebrain (FB), micro-dissected cortex (Cx), and micro-dissected CPN axon bundles at corpus callosum (Ax), labeled for CPEB4, ACTIN, and TAU. Labeling for CPEB4 reveals 3 distinct bands, which are labeled as b1 (highest MW), b2 (middle MW), and b3 (lowest MW). Quantification of signal intensity in Cpeb4 b1-b3 from triplicates reveals relative enrichment of signal for micro-dissected axons to the middle band (b2). **(d)** Depending on the exact number and relative positioning of CPE motifs within the 3'-UTR, CPEBs can mediate either translational activation (as schematized in (a)) or translational repression [adapted from[96]]. **(e)** Quantification of length (top left), number (bottom left), density (top right), distance from the polyadenylation signal (PAS) (bottom middle), and minimum distance (bottom right) between starts of CPE motifs in 3'-UTRs enriched in GCs (orange, light blue) and somata (brown, dark blue) of CPN (orange/brown) and CThPN (light/dark blue), cdf – cumulative distribution function. **(f)** Relative proportions of distinct variations of Cpeb4 motif, detected in 3'UTRs enriched in GCs (orange, light blue) and somata (brown, dark blue) of CThPN (light/dark blue) and CPN (orange/brown).
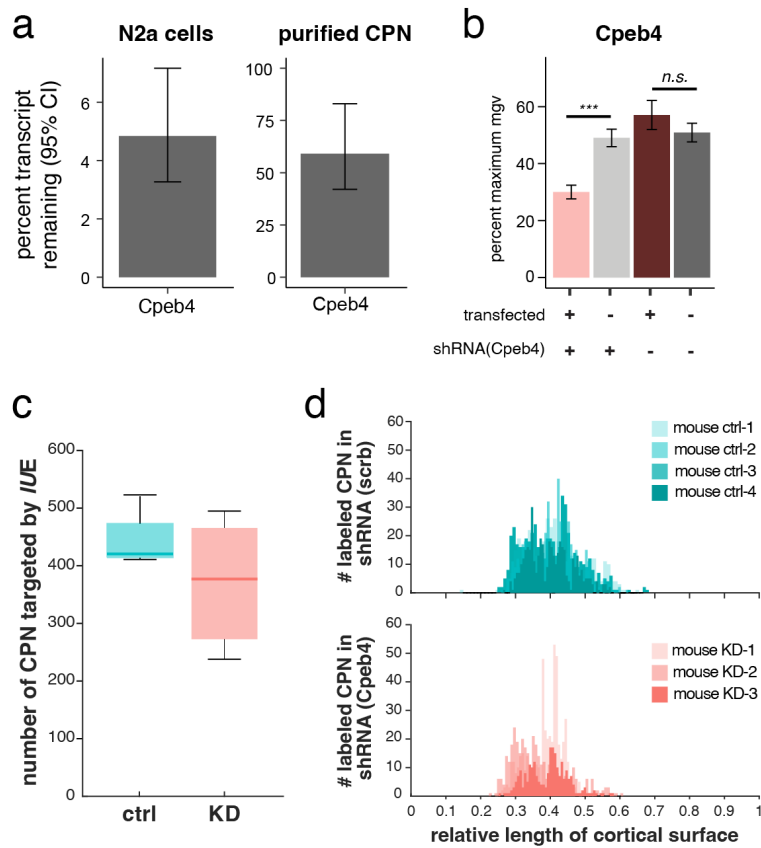
Figure S10: Validation of shRNA-mediated *Cpeb4* knockdown. (a/b) Efficiency of shRNA-mediated knockdown of *Cpeb4*. (a) Knockdown efficiency quantified via qPCR in transfected N2a cells and in *in utero* electroporated and FACS-purified CPN. (b) Knockdown efficiency assessed via immunohistochemistry in transfected N2a cells. **(c)** Total number of labeled CPN in a single 40 μm coronal brain section of mice *in utero* electroporated with a scrambled control construct (ctrl) or an shRNA targeting *Cpeb4* for knockdown (KD). **(d)** Unilateral *in utero* electroporation at E14.5 with a scrambled control construct (ctrl, upper panel, n = 4), or shRNA targeting *Cpeb4* (KD, lower panel, n = 3) results in reproducible targeting of CPN in cortical layer II/III in the lateral somatosensory cortex.
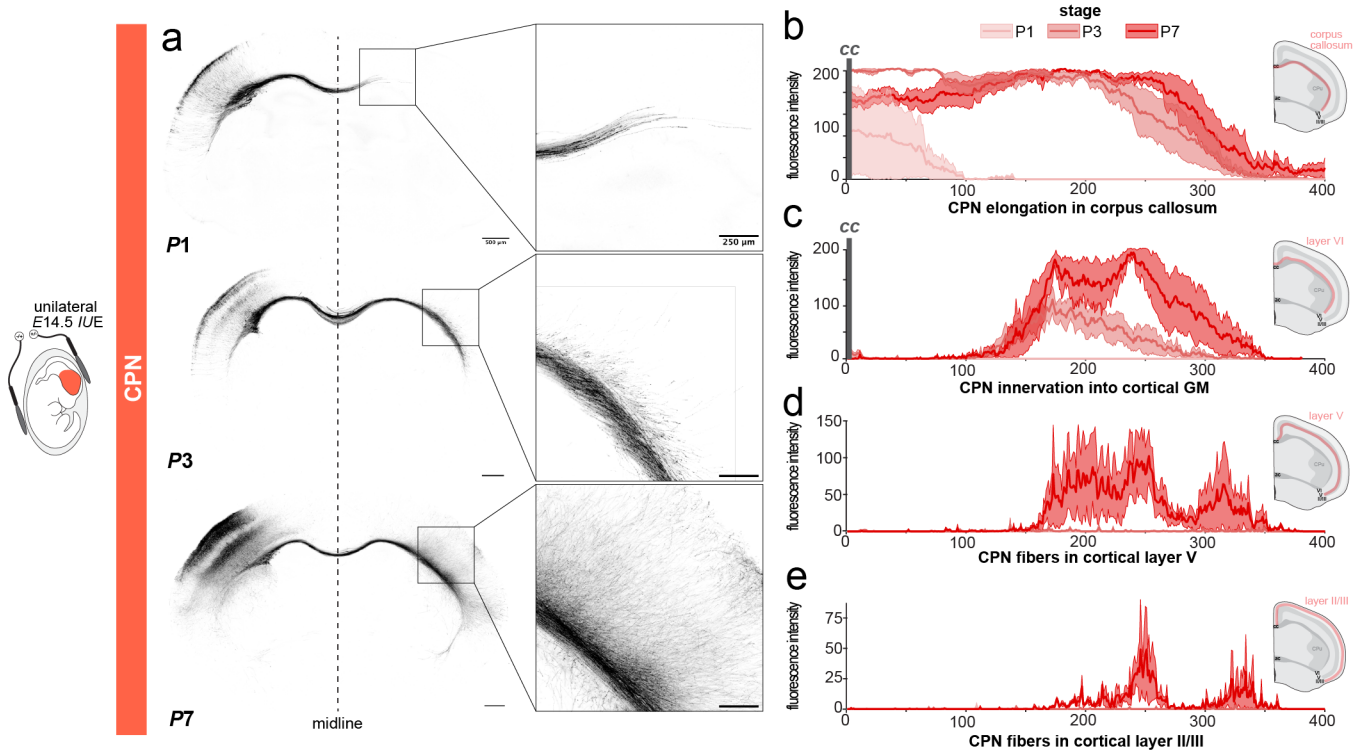
Figure S11: CPN circuit formation occurs during the first postnatal week. (a) Somatosensory area CPN extend their axons across the midline around P1, innervate their homotopic target area in the contralateral cortex around P3, and collateralize within the cortical grey matter and likely start formation of synapses around P7. **(b-e)** Quantification of labeled axons/collaterals in progressively entered regions of interest (highlighted in schematic insets to the right, n = 3 for each developmental time), focusing on (b) CPN axon/collateral elongation in the subcortical white matter, (c) innervations into cortical grey matter (GM), (d) axon/collateral density in layer V or (e) in layers II/III at P1 (light pink), P3 (dark pink), and P7 (red).
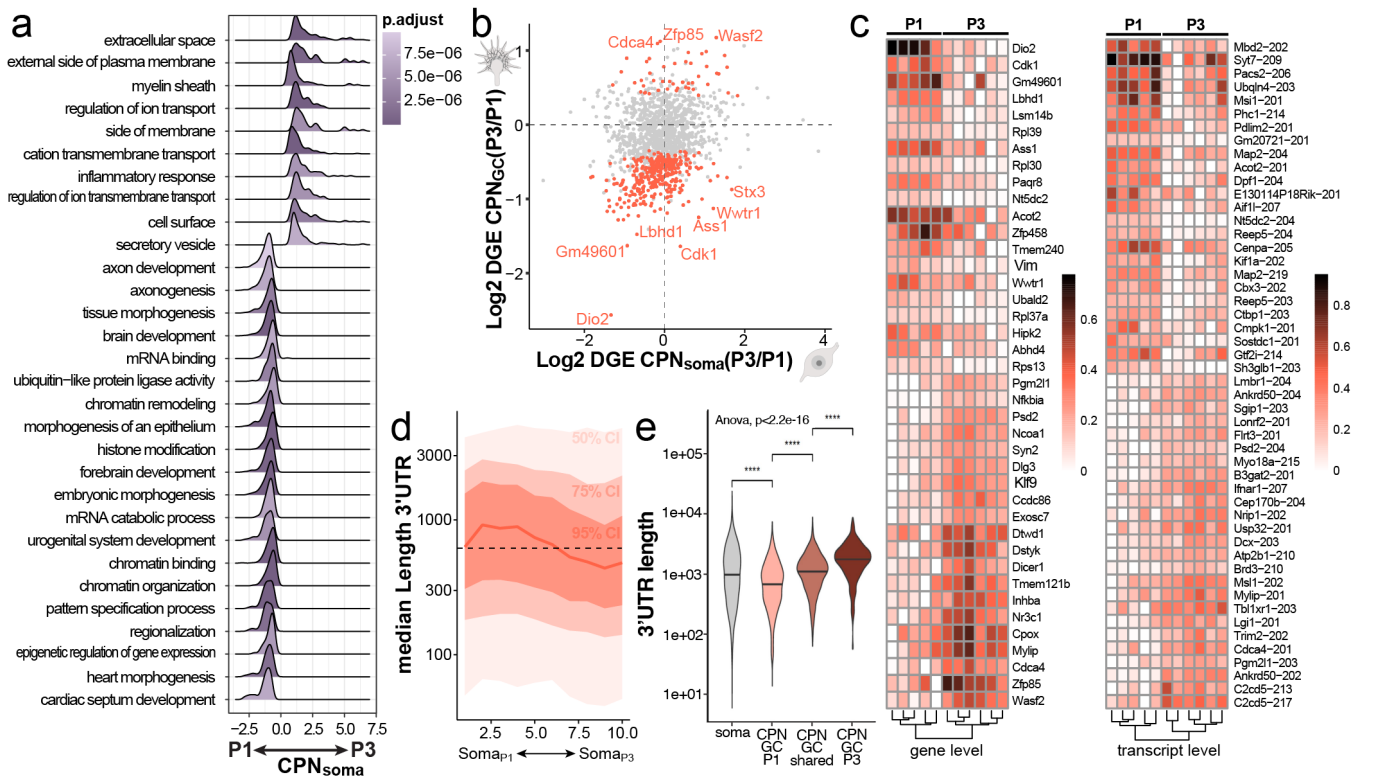
Figure S12: Localized transcriptomes change dynamically from P1 to P3, as CPN switch from axon elongation to grey matter innervation and synapse formation. (a) Gene set enrichment analysis for CPN somata comparing P1 and P3. **(b)** Quadrant plot of differential gene expression by CPN at P3 *vs.* P1, and transcript localization by subcellular compartments: in somata (x axis) and GCs (y-axis). **(c)** Scaled transcript abundance of most significantly regulated CPN$_{GC}$ genes comparing P1 and P3, analyzed at gene level (left) and transcript level (right). **(d)** Median length of 3'UTR for transcripts detected in CPN and enriched at P1 or P3. Color ribbons indicate 50%, 75%, and 95% confidence intervals (CI). **(e)** Length of 3'UTRs of CPN$_{GC}$ genes increases from developmental stage P1 to P3.
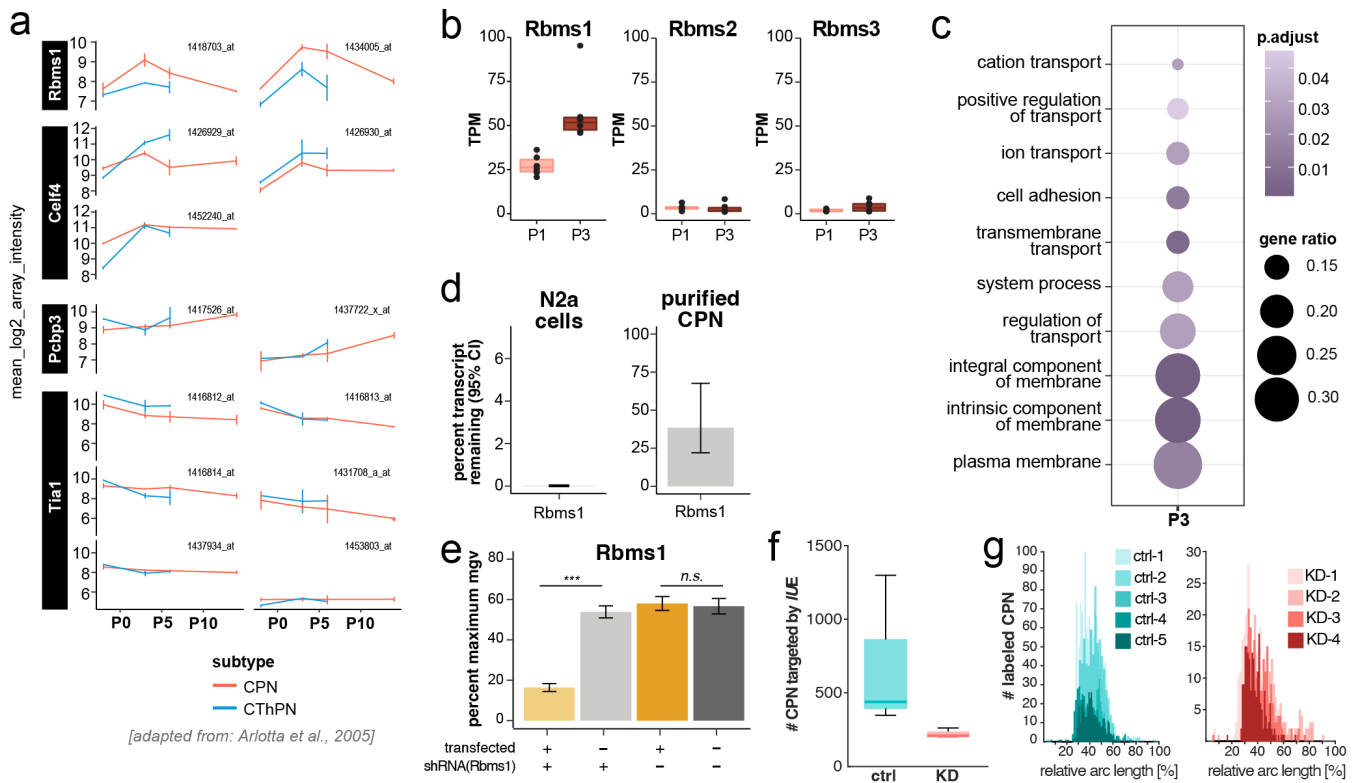
Figure S13: The RBP *Rbms1* increases transcript levels from P0 to P5, and is likely involved in localization and stabilization of CPN$_{GC}$ transcripts. (a) Micro array data for RBPs *Rbms1*, *Celf4*, *Pcbp3*, and *Tia1* at P0, P5, and P10, analyzed for CPN (red) and CThPN (blue), from Arlotta et al. 2005[51]. **(b)** Boxplots highlighting RNA expression changes for the paralogs *Rbms1*, *Rbms2*, and *Rbms3* in CPN somata at P1 (pink) and P3 (brown). **(c)** GO term enrichment for subset of CPN$_{GC}$ transcripts containing RBMS1 binding motifs at developmental stage P3. **(d/e)** Efficiency of shRNA-mediated knockdown of *Rbms1*, assessed via (d) qPCR in transfected N2a cells and *in utero* electroporated and FACS-purified CPN, as well as (e) via immunocytochemistry in transfected N2a cells (e). **(f)** Total number of labeled CPN in the cortex of a single 40 μm coronal brain section of mice electroporated with a scrambled control construct (ctrl) or a shRNA targeting *Rbms1* (KD). **(g)** Unilateral *in utero* electroporation at E14.5 with a scrambled control construct (ctrl, left panel, n = 5), or shRNA targeting *Rbms1* (KD, right panel, n = 4) results in comparable positioning of targeted CPN in cortical layer II/III in the lateral somatosensory cortex. The number of labeled CPN was higher in control mice, compared to KD mice.