Received: 29 November 2019 | Revised: 20 January 2020 | Accepted: 23 January 2020

DOI: 10.1002/wrna.1590

ADVANCED REVIEW

WIREs RNA WILEY

Conservation of a core neurite transcriptome across neuronal types and species

Nicolai von Kügelgen | **Marina Chekulaeva**

Non-coding RNAs and Mechanisms of Cytoplasmic Gene Regulation, Berlin Institute for Medical Systems Biology, Max Delbrück Center for Molecular Medicine, Berlin, Germany

Correspondence
Marina Chekulaeva, Non-coding RNAs and Mechanisms of Cytoplasmic Gene Regulation, Berlin Institute for Medical Systems Biology, Max Delbrück Center for Molecular Medicine, Berlin, Germany.
Email: marina.chekulaeva@mdc-berlin.de

Funding information
Deutsche Forschungsgemeinschaft, Grant/Award Number: CH 1737/6-1; EU Joint Programme - Neurodegenerative Disease Research, Grant/Award Number: localMND; Max Delbrück Center

Abstract

The intracellular localization of mRNAs allows neurons to control gene expression in neurite extensions (axons and dendrites) and respond rapidly to local stimuli. This plays an important role in diverse processes including neuronal growth and synaptic plasticity, which in turn serves as a foundation for learning and memory. Recent high-throughput analyses have revealed that neurites contain hundreds to thousands of mRNAs, but an analysis comparing the transcriptomes derived from these studies has been lacking. Here we analyze 20 datasets pertaining to neuronal mRNA localization across species and neuronal types and identify a conserved set of mRNAs that had robustly localized to neurites in a high number of the studies. The set includes mRNAs encoding for ribosomal proteins and other components of the translation machinery, mitochondrial proteins, cytoskeletal components, and proteins associated with neurite formation. Our combinatorial analysis provides a unique resource for future hypothesis-driven research.

This article is categorized under:

RNA Export and Localization > RNA Localization

RNA Evolution and Genomics > Computational Analyses of RNA

RNA Methods > RNA Analyses in Cells

KEYWORDS

local translation, mRNA localization, neuron, RNA-binding proteins

1 | INTRODUCTION

Neurons are highly polarized cells consisting of a cell body (soma) and neurites (axons and dendrites). Neurites can extend for long distances and function quite autonomously from cell bodies, mostly due to subcellular mRNA localization (reviewed in Glock, Heumuller, & Schuman, 2017). This targeting allows neurons to respond rapidly to local stimuli by regulating the translation of localized mRNAs as one means of modifying the local proteome (reviewed in Chekulaeva & Landthaler, 2016). This process contributes to numerous neuronal processes including axon guidance, dendritic arborization, and synaptic plasticity.

Specific patterns of localization for these mRNAs are mediated by *cis*-regulatory elements (zipcodes) usually found in their 3′UTRs. These elements are bound by specific RNA-binding proteins (RBPs) that link their targets to the

transport machinery and direct their trafficking to the sites where they function. Many of these "zipcodes" and RBPs are conserved across diverse cell types and species, indicating that RNA localization likely has fundamental biological functions. RBPs can both control mRNA localization, for example, by interacting with motor proteins, and regulate translation of localized mRNAs.

Recent high-throughput studies have identified hundreds to thousands of mRNAs that specifically localize to neurites (Briese et al., 2016; Cajigas et al., 2012; Ciolli Mattioli et al., 2019; Farris et al., 2019; Feltrin et al., 2012; Lein et al., 2007; Maciel et al., 2018; Middleton, Eberwine, & Kim, 2019; Minis et al., 2014; Poon, Choi, Jamieson, Geschwind, & Martin, 2006; Poulopoulos et al., 2019; Rotem et al., 2017; Saal, Briese, Kneitz, Glinka, & Sendtner, 2014; Taliaferro et al., 2016; Taylor et al., 2009; Toth et al., 2018; Tushev et al., 2018; Zappulo et al., 2017). There has yet to be a systematic analysis comparing these datasets to determine the extent to which they overlap, and might thus provide important functional insights, as well as factors that might contribute to differences (e.g., neuronal type, neurite isolation method, library preparation techniques, etc.). Here we compare 20 datasets on neuronal mRNA localization from the literature, all of which carried out analysis of separated neuronal compartments, to identify a core neurite transcriptome shared across different experimental setups. We found that this common set encodes components of the translation machinery and cytoskeleton, and proteins associated with mitochondria and neurite formation. The analysis clearly distinguished the transcriptomes of cell lines and neurons generated in vitro from those of primary neurons, but surprisingly, did not reveal clear differences in those obtained from different types of primary neurons. Our comprehensive analysis provides a valuable resource for future studies on mRNA localization.

2 | STUDYING mRNA LOCALIZATION: FROM IMAGING TO IN SITU SEQUENCING

Studies of neurite-specific transcriptomes have been based on different approaches, each of which entails particular methodological challenges. In situ hybridization is based on the use of fluorescent probes complementary to mRNAs of interest and has been long used to study mRNA localization. Single-molecule fluorescence in situ hybridization (smFISH) represents a further improvement of this technique; it applies multiple fluorescent probes for each transcript to be analyzed and permits the detection of individual RNA molecules. Many of the first mRNAs known to be localized to neurites were identified using these techniques, and they are still frequently used as markers of mRNA localization. However, determining the full complement of neurite-localized transcripts required transcriptome-wide sequencing methods.

Only now have advances in imaging methods permitted high-throughput smFISH, which has been a challenge due to the limited number of available fluorophores that can be distinguished. To overcome this and increase the number of potential parallel targets, labs have developed sequential hybridization and combinatorial labeling techniques including seqFISH (La Manno et al., 2016; Lubeck, Coskun, Zhiyentayev, Ahmad, & Cai, 2014; Shah et al., 2016; Shah, Lubeck, Zhou, & Cai, 2016, 2017) and multiplexed error-robust FISH (MERFISH; K. H. Chen, Boettiger, Moffitt, Wang, & Zhuang, 2015). Multiplexing is achieved via multiple cycles of hybridization and stripping or photobleaching, with each hybridization step applying distinct fluorescent probes. MERFISH (K. H. Chen et al., 2015) further introduced a two-step labeling procedure in which transcripts of interest are first hybridized with non-fluorescent probes that contain a sequence complementary to the selected transcript and two arms that are complementary to fluorescent readout probes. These readout probes are then used in multiple hybridization rounds, alternating with photobleaching, to detect the transcripts. Multiplexed smFISH permits evaluating both the copy numbers and localizations of thousands of mRNAs and holds great potential for studies of their localization in neurons.

Another challenge in detecting mRNA localization through light microscopy is its limited resolution. To address this issue, F. Chen, Tillberg, and Boyden (2015) developed a method called expansion microscopy (ExM), which permits enlarging biological specimens. The authors introduced a polymer gel into fixed samples, triggering a chemically induced swelling that expands them by almost two orders of magnitude. ExM permits conventional microscopes to achieve nanoscale imaging.

Another approach to combine the spatial context of imaging with the high throughput aspect of RNA sequencing (RNA-seq) is spatial transcriptomics (Ke et al., 2013; Lee et al., 2014; Rodriguez-Munoz et al., 2015; Stahl et al., 2016; Vickovic et al., 2019). This approach relies on either spatial barcoding or in situ sequencing (ISS). In spatial barcoding (Rodriguez-Munoz et al., 2015; Stahl et al., 2016; Vickovic et al., 2019), tissue sections are placed on a glass slide coated with oligonucleotides that bear localization-specific barcodes and oligo(dT) to capture mRNAs. After their capture and

reverse transcription, cDNA is released from the slide, sequenced and mapped to the tissue using barcodes. ISS involves the generation of cDNA within fixed tissue, followed by rolling-circle amplification and sequencing (ISS; Ke et al., 2013; fluorescence in situ sequencing, FISSEQ; Lee et al., 2014). Further advances in these methodologies and their applications to subcellular compartments promise to yield important new insights into questions of mRNA localization.

3 | ISOLATION OF SUBCELLULAR NEURONAL COMPARTMENTS FOR RNA SEQUENCING

While microscopy-based methods excel at analyzing subcellular patterns of mRNA localization, RNA-seq provides the best unbiased transcriptome-wide analysis. A number of approaches have been developed to isolate subcellular neuronal compartments for further RNA-seq to carry out transcriptome-wide analyses of mRNA localization. (a) Manual (Cajigas et al., 2012; Tushev et al., 2018) or laser capture microdissection (Farris et al., 2019; Simone, Bonner, Gillespie, Emmert-Buck, & Liotta, 1998; Zivraj et al., 2010) permits the isolation of subcellular compartments including dendrites, axons or axonal growth cones directly from the brain. However, material obtained from tissue samples is heterogeneous, containing not only different types of neurons but also non-neuronal cells. This limitation can be partially overcome by using immunopanning or fluorescence-activated cell sorting (FACS) to purify neuronal populations genetically labeled with a fluorescent marker (Lobo, Karsten, Gray, Geschwind, & Yang, 2006; Zhang et al., 2014). (b) Density gradient centrifugation has been used to isolate axonal growth cones (Poulopoulos et al., 2019) and synaptosomes, structures composed of pre- and postsynaptic compartments (Dunkley, Jarvie, & Robinson, 2008). This method relies on the homogenization of brain tissue to break off nerve terminals, which are then isolated by centrifugation. Similar to microdissection, this approach produces a heterogeneous population of neuronal types and can be combined with FACS (Poulopoulos et al., 2019). Subcellular compartments of neurons can also be separated by culturing the cells (c) in compartmentalized microfluidic chambers (Briese et al., 2016; Taylor, Dieterich, Ito, Kim, & Schuman, 2010) or (d) on microporous membranes (Ciolli Mattioli et al., 2019; Ludwik, von Kuegelgen, & Chekulaeva, 2019; Pertz, Hodgson, Klemke, & Hahn, 2006; Taliaferro et al., 2016; Zappulo et al., 2017) The latter separate cell bodies, which grow on top of the membrane, from neurites, which stretch through the pores and emerge on the lower side of the membrane. This method can be easily scaled up to produce enough material for multiple omics analyses, but does not always permit a separation of axons from dendrites.

Neurons can be isolated from the brain (yielding primary neurons) or generated through in vitro procedures from embryonic or induced pluripotent stem cells (iPSCs), using small molecules that direct neuronal differentiation (Wichterle & Peljto, 2008) or through the expression of neurogenic transcription factors (Chanda et al., 2014; Heinrich et al., 2011; Liu et al., 2015; Zappulo et al., 2017). Each system has distinct advantages. While primary cultures are more likely to recapitulate the properties of neuronal cells in vivo, stem cell-derived neurons are more homogeneous, easy to genetically modify, and reduce animal use. Moreover, generation of neurons from patient-derived iPSCs is the only way to obtain neurons with the genetic background of patients.

An alternative to separating subcellular compartments is proximity-dependent RNA labeling (Fazal et al., 2019; Wang et al., 2019). This method labels RNA pools using the peroxidase enzyme APEX2 (Fazal et al., 2019) which can be targeted to specific subcellular compartments, or the light-activated, proximity-dependent photo-oxidation of RNA nucleobases (Wang et al., 2019). The RNA pools localized in these ways can then be isolated and analyzed through RNA-seq. This approach has the advantage of providing subcellular high-throughput data from intact samples of tissues. Further improvement of precise subcellular targeting approaches will be important to future studies of RNA localization.

4 | COMPARATIVE ANALYSIS OF TRANSCRIPTOME-WIDE NEURONAL mRNA LOCALIZATION DATASETS

4.1 | Local transcriptome shared across multiple datasets

Advances in RNA-seq and imaging technologies over the last decade have yielded a number of datasets pertaining to neurite-localized transcriptomes. But these datasets have not been subjected to a comparative analysis, which hampers their interpretation; such an analysis would represent an essential resource for researchers in the fields of RNA

localization and neurobiology. We have collected 20 datasets pertaining to levels of RNA expression in the subcellular compartments of neurons from mice, rats, and humans (Figure 1). These include two datasets from cortical tissues (Lein et al., 2007; Poulopoulos et al., 2019), two datasets from cultured cortical neurons (Taliaferro et al., 2016; Taylor et al., 2009), three datasets from hippocampal tissue slices (Cajigas et al., 2012; Farris et al., 2019; Tushev et al., 2018), two datasets from cultured hippocampal neurons (Middleton et al., 2019; Poon et al., 2006), one dataset from dorsal root ganglia (DRG; Minis et al., 2014), three datasets from primary motor neurons (Briese et al., 2016; Rotem et al., 2017; Saal et al., 2014), two datasets from mouse embryonic stem cell (mESC)-derived neurons (Ciolli Mattioli et al., 2019; Zappulo et al., 2017), two datasets from human induced pluripotent stem cells (hiPSC) derived neurons (motor neurons, Maciel et al., 2018, as well as a mixture of GABAergic and glutamatergic neurons, Toth et al., 2018), and three neuroblastoma cell lines (Feltrin et al., 2012; Taliaferro et al., 2016). These datasets are derived from three different species, multiple model systems and different methods for separating subcellular compartments. This suggests that widely shared patterns likely represent a core set of components that have been conserved in a transcriptome localized to neurites.

Since these RNA-seq datasets were produced with custom settings that differed from study to study, we re-analyzed them all using the same pipeline (Wurmus et al., 2018). An exception was made for three RNA-seq datasets for which raw sequencing data were unavailable (Maciel et al., 2018; Poulopoulos et al., 2019; Rotem et al., 2017). Depending on the study, from 54 to ~10,000 transcripts were detected with high confidence in neurites (transcripts per million [TPM] of library reads >10, Figure 1); this reflects the sensitivity of methods used over the last decade. The vast majority of transcripts were detected in multiple datasets (≥3, Figure 1, dark bars).

This analysis yielded a set we call the core neurite transcriptome, comprising 70 transcripts detected in >15 datasets (Table 1 and extended online Table 1, see www.chekulaevalab.org/resources). Based on Uniprot annotations this set

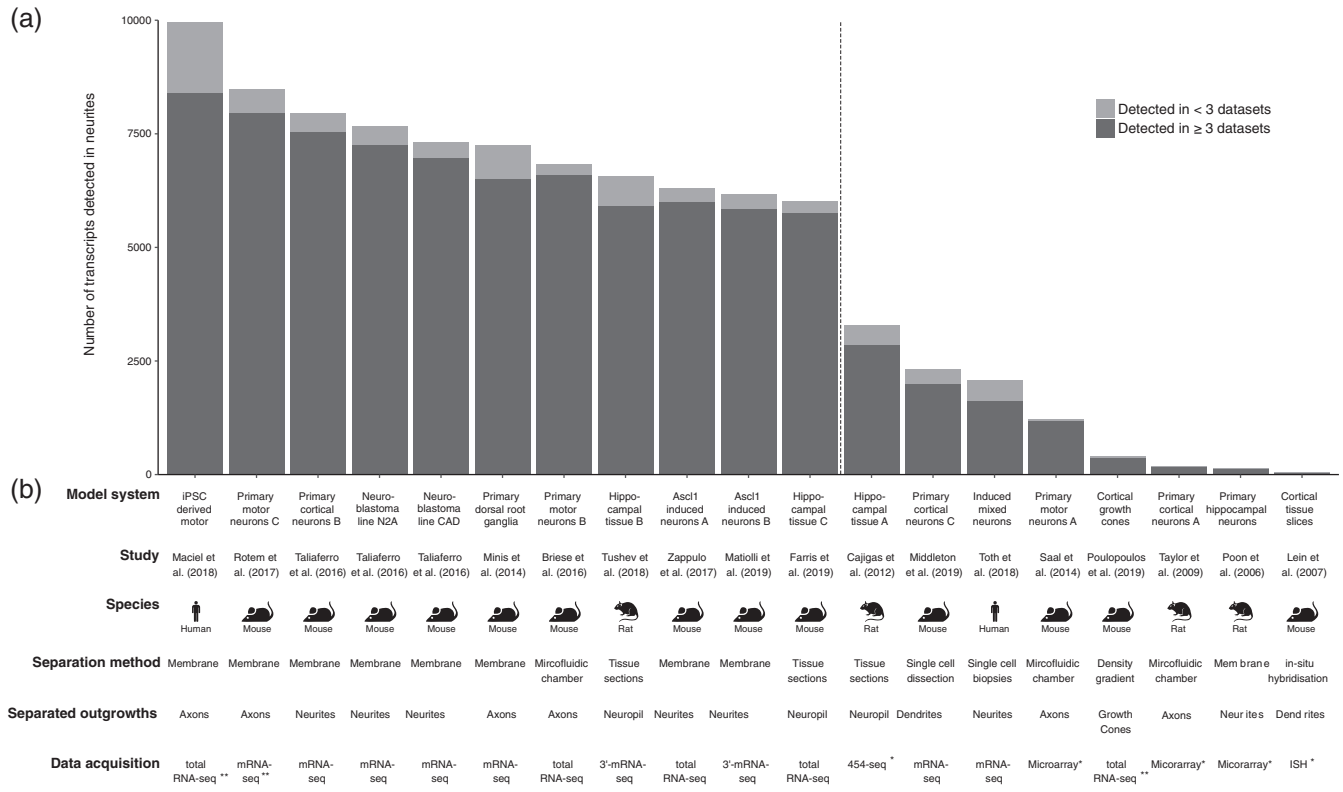

FIGURE 1 Datasets included in this study. (a) Bar graph showing the number of transcripts detected at a high confidence expression level (TPM > 10) in neurites for each dataset. The color of the bars indicates if transcripts are identified in at least three independent datasets (dark gray) or in less than three datasets (light gray). A dotted line separates datasets with high coverage (>5,000 transcripts) from the other datasets. (b) Annotation of the datasets: model system and cell type; citation; the species of a dataset (human, mouse, or rat); the separation method for outgrowths and cell bodies; the method of data acquisition. RNA-seq studies were re-analyzed using the PiGx RNA-seq pipeline (Wurmus et al., 2018). If studies used different analysis methods (*) or raw data was unavailable (**), transcript detection levels were obtained from supplementary data of the respective studies

TABLE 1 Most abundant neurite transcripts across at least 16 datasets

Gene name	Function	Datasets with neurite TPM > 10	Mean TPM in neurites	Studies with neurite ribosome association
Actb	Cytoskeleton	18	1,201.03	1
Tpt1	Outgrowth formation; mitochondria regulation (Roque, Wong, Lin, & Holt, 2016)	18	1,106.49	2
Rpl4	Ribosomal protein	18	811.03	2
Ybx1	RNA binding; multifunctional	17	660.04	0
Rps12	Ribosomal protein	17	635.16	1
Rps8	Ribosomal protein	17	616.48	1
Atp5b	Mitochondrial function	17	577.66	2
Ywhae	Outgrowth formation (Cornell, Wachi, Zhukarev, & Toyo-Oka, 2016)	17	568.77	1
Rpl6	Ribosomal protein	17	538.44	0
Npm1	Nuclear protein; ribosome-associated (Maggi et al., 2008)	17	490.34	1
Map1b	Cytoskeleton	17	489.71	0
Fau	Ribosomal protein	17	390.98	1
Calm1	Ca^{2+} regulation	17	350.40	2
Rps3a1	Ribosomal protein	17	349.43	2
Kif5c	Synaptic function (Willemsen et al., 2014)	17	259.35	1
Gap43	Outgrowth formation; ZBP1 target (Donnelly et al., 2011)	17	192.16	0
Kif5a	Axonal transport (Xia et al., 2003)	17	140.66	2
Park7	Oxidative stress protection	17	103.02	1
Arl3	Membrane trafficking	17	69.59	0
Vdac3	Mitochondrial regulation	17	60.90	0
Eef1a1	Translation machinery	16	2,234.90	2
Actg1	Cytoskeleton	16	1,055.28	1
Eef2	Translation machinery	16	951.90	2
Rplp1	Ribosomal protein	16	802.69	2
Rpl23	Ribosomal protein	16	707.27	1
Rps2	Ribosomal protein	16	648.43	0
Rps23	Ribosomal protein	16	618.93	0
Rps14	Ribosomal protein	16	572.98	1
Tubb5	Cytoskeleton	16	549.63	0
Rpl3	Ribosomal protein	16	536.79	2
Hsp90aa1	Protein folding	16	530.99	2
Eif4g2	Translation machinery	16	505.60	1
Rps21	Ribosomal protein	16	499.95	1
Rpl7	Ribosomal protein	16	480.81	2
Rps11	Ribosomal protein	16	444.06	2
Cox6a1	Mitochondrial function	16	436.29	0
Ywhaz	Outgrowth formation (Wachi et al., 2016)	16	417.89	1
Calm2	Ca^{2+} regulation	16	417.60	1
Rpl36	Ribosomal protein	16	392.79	1

(Continues)

TABLE 1 (Continued)

Gene name	Function	Datasets with neurite TPM > 10	Mean TPM in neurites	Studies with neurite ribosome association
Rpl15	Ribosomal protein	16	381.53	1
Stmn1	Cytoskeleton; outgrowth formation	16	270.92	1
Hint1	NMDA receptor regulation (Rodriguez-Munoz, Cortes-Montero, Pozo-Rodrigalvarez, Sanchez-Blazquez, & Garzon-Nino, 2015)	16	262.80	1
Set	Nuclear protein	16	243.58	0
H3f3b	Nuclear protein	16	230.98	0
Ctnnb1	Wnt signaling; outgrowth formation (Armstrong, Ryu, Chieco, & Kuruvilla, 2011)	16	229.71	1
Slc25a4	Mitochondrial regulation	16	216.78	1
Mapt	Cytoskeleton	16	212.45	1
Cpe	Secretory pathway	16	202.66	2
Fkbp1a	TGF-beta and SMAD signaling	16	195.73	1
Akr1a1	Oxidative stress protection	16	186.94	0
Ywhaq	Signaling	16	185.82	1
Gnb1	Signaling	16	184.81	0
Basp1	Outgrowth formation (Frey, Laux, Xu, Schneider, & Caroni, 2000)	16	176.43	2
Tubb2a	Cytoskeleton	16	175.27	1
Vdac1	Mitochondrial regulation	16	174.54	1
Skp1a	Proteasomal pathway	16	149.18	0
Hmgn1	Nuclear protein	16	143.65	1
Dynlrb1	Transport	16	140.44	1
Gdi2	Rab regulation	16	138.08	0
Arpc2	Cytoskeleton	16	137.44	0
Ywhah	Signaling	16	134.44	0
Atp5o	Mitochondrial function	16	102.35	1
Stmn3	Cytoskeleton	16	102.15	2
Tcf4	Nuclear protein	16	91.99	1
Rtn3	Membrane trafficking; amyloid processing (He et al., 2004)	16	84.89	1
Psmd14	Proteasomal pathway	16	75.00	0
Ppp3ca	Ca^{2+} signaling	16	68.74	0
Actr2	Cytoskeleton	16	67.67	0
Stx7	Synaptic function	16	52.08	0
Ubqln2	Proteasomal pathway	16	38.62	1

Note: Table of all transcripts that are detected in neurites in at least 16 out of the 20 studies included in our analysis. Listed are Ensembl gene names, manually annotated functions (Uniprot, unless otherwise stated), number of datasets detecting a transcript with transcripts per million (TPM) > 10, average TPM value from all datasets, and the number of studies reporting local translation of a transcript.

revealed an enrichment for mitochondrial mRNAs (*Atp5b*, *Apt5o*, *Vdac1*, *Vdac3*), mRNAs encoding ribosomal proteins and other translation-associated proteins (translation initiation and elongation actors), components of the cytoskeleton (*β-actin*, *Map1b*, *Tubulinβ2A*, *Tau/Mapt*, *Arpc5*, *Dynlrb1*), calcium-binding proteins (*Calm1*, *Calm2*), and proteins with roles in axon and dendrite formation (*Ywhaz*, *Ywhae*, *Gap43*, *Tpt1*, *Stmn1*). While well studied transcripts such as *β-actin* have long been known to localize to neurites (Bassell et al., 1998; Micheva, Vallee, Beaulieu, Herman, & Leclerc, 1998), the detection of mRNAs encoding ribosomal and mitochondrial proteins has mostly come through RNA-seq

studies (Briese et al., 2016; Cajigas et al., 2012; Gumy et al., 2011; Saal et al., 2014; Taylor et al., 2009; Zivraj et al., 2010). Additionally, the localization of transcripts encoding ribosomal proteins has also been confirmed by FISH (Poulopoulos et al., 2019; Zivraj et al., 2010).

Recent work by our lab has shown that up to a half of the local proteome in neurites is likely to be established through translation of localized mRNAs (Zappulo et al., 2017). To determine whether the mRNAs we identified as components of the core neurite transcriptome are actually translated there, we checked for their presence in four datasets reporting locally translated mRNAs (Ainsley, Drane, Jacobs, Kittelberger, & Reijmers, 2014; Ouwenga et al., 2017; Shigeoka et al., 2016; Zappulo et al., 2017). Notably, ~70% of transcripts consistently detected in neurites were also determined to undergo local translation in at least one of these studies (Table 1 and extended online Table 1).

Surprisingly, some of the transcripts commonly detected in neurites encode proteins with nuclear functions: these include histone *H3F3b*, chromosomal protein *Hmgn1* and the transcription factors *Tcf4* and *Rnf10*. For *H3F3b*, imaging-based analysis showed that the transcript is indeed present in dendrites but only at the proximal end (Cajigas et al., 2012). Such unexpected localization patterns might reflect regulatory mechanisms, such as an activity-dependent transport of proteins from neurites to the nucleus. In fact, this mechanism has been reported for the protein RNF10 (Dinamarca et al., 2016). The consistent detection of *Rnf10* mRNA both in neurites and among ribosome-associated transcripts (extended online Table 1) suggests that this postsynaptic protein is locally translated and then transported to the nucleus upon synaptic stimulation. It remains to be determined whether a similar mechanism also regulates other transcripts that have known nuclear functions and are localized to neurites.

4.2 | Differences between the subcellular transcriptomes of neurons

The datasets included in our comparison are derived from neurites of different types of neurons. This means that some components of the transcriptomes are likely to be cell-type specific for major classes such as motor neurons or hippocampal neurons. To analyze the differences between the datasets in more detail and in a statistically unbiased manner, we performed principal component analysis (PCA, see Box 1; Figure 2a,b). This type of analysis requires a large overlap between the datasets it draws on, so our analysis included only datasets with high coverage and complexity (Figure 1, datasets on the left of the dashed line). PCA produced the following clusters: (a) rat hippocampal neuropil (principal component 2 or PC2); (b) two neuroblastoma cell lines, CAD and N2A (PC1 and PC3); (c) mESC-derived neurons (PC1 and PC3); (d) all other datasets, including different types of mouse primary neurons and hiPSC-derived motor neurons (PC1 and PC3). To understand which differences of expression underlie this clustering, we generated heatmaps for the transcripts which made the greatest contributions to each PC (Figure 2c). This analysis showed that the dataset from rat hippocampal neuropils lacks a number of transcripts detected in others, separating it from other datasets. This is most likely due to the fact that the annotation of the rat genome is still incomplete, and the study in question (Tushev et al., 2018) applied 3′-mRNA-seq, which sequences only the very ends of 3′-UTRs (Figure 2d, Figure S1). Datasets

BOX 1 RNA-seq DATA COMPARISON USING PCA

Comparison of properly normalized RNA-seq data can be performed with different methods: clustered heatmaps visualize groups of transcripts with similar expression values; PCA as a dimensionality reduction method can show similarity of samples or datasets based on underlying numerical values, most often expression data. PCA (Mardia, Kent, & Bibby, 1979; Venables, Ripley, & Venables, 2002) produces multiple independent orthogonal principle components (PCs), represented as axes along which samples can be grouped. This step requires values to be measured for each sample and transcript or gene; therefore, only transcripts that are detected in all samples can be analyzed. Because PCA performs a linear dimensionality reduction, it is possible to calculate the contribution of each transcript to each PC. Nonlinear dimensionality reduction methods, like t-SNE (van der Maaten & Hinton, 2008) or uniform manifold approximation and projection (UMAP) (Becht et al., 2018), allow the compression of the multidimensional relationships, calculated in PCA, into two or three dimensions. This enables visualization of all results on the same plot. However, due to the nonlinearity of transformation, the contribution of individual transcripts to specific PCs or relationships between the clusters cannot be extracted anymore.

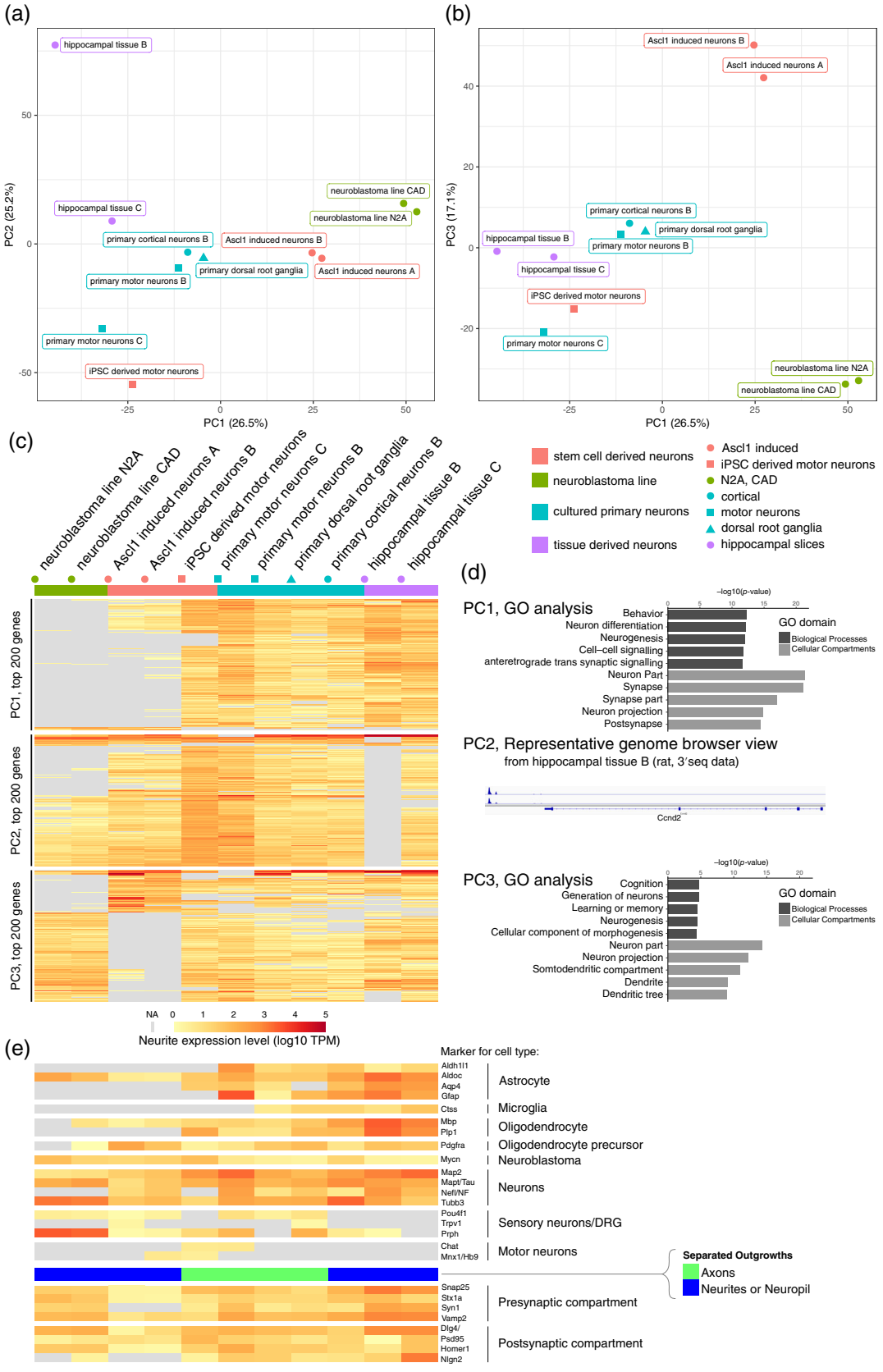

FIGURE 2 Legend on next page.

derived from the neuroblastoma lines and, to a lesser degree, mESC-derived neurons, differ most significantly from primary neurons in terms of transcripts associated with neurogenesis and synaptic activity (Figure 2d; full transcript list in subsets of extended online Table 1).

Contrary to our expectations, different types of primary neurons and tissue-derived samples (motor neurons, cortical neurons, hippocampal neurons and tissues, DRG) cluster together and cannot be clearly separated by PCA (Figure 2a, b). To determine whether neuronal cell type-specific differences are detectable in cell bodies, we applied the PCA to somatic expression levels (Figure S2a,b). Here, too, we were unable to distinguish different types of primary neurons and differences in expression are related to similar functional terms as in neurites (Figure S2c). We next applied t-distributed Stochastic Neighbor Embedding (t-SNE), which collects all PCs into a two-dimensional space. Here, too, we were unable to distinguish neuronal cell types (Figure S2d,e).

To determine whether these effects were due to contaminations by non-neuronal cell types, we assessed levels of the expression of genes specific to particular neuronal and non-neuronal cell types (Figure 2e). We observed higher levels of specific non-neuronal markers (for astrocytes and microglia) in primary neurons. This suggested that glia contamination might be a confounding factor that prevented us from discriminating neuronal cell types. Others might include an intrinsic heterogeneity among primary neurons, differences in the approaches used to separate the transcriptomes of neurites and soma, or in library preparation methods or general experimental handling. Notably, well-defined motor neuron (*Hb9*, *Chat*) and sensory neuron (*Prph*, *Trpv1*) markers were detected in the corresponding datasets. However, we did not observe clear signatures for pre- or postsynaptic markers (Figure 2e). Possible reasons might include the technical challenges involved in separating axons from dendrites and the fact that our analysis focused on levels of expression of mRNAs, while the best-established markers rely on immunostaining assays for synaptically localized proteins.

One possible strategy to deal with glial contamination, employed by Tushev et al. (2018), is to computationally filter out all transcripts that do not show enrichment in cultured primary neurons when compared with astrocyte cultures. To test whether the removal of non-neuronal contaminants improves neuronal cell type clustering, we applied the same filtering procedure to all of the datasets before performing the PCA (Figure S3). This filtering did not permit PCA to produce a clearer division of datasets according to cell type. It is important to note that this filtering removes not only astroglial transcripts, but also transcripts shared between neurons and astroglia that might have functions in neurites. So it drastically reduced the overall number of transcripts shared between the datasets (from 8,809 to only 2,352) (Figure S3).

4.3 | Robustly neurite-enriched transcriptomes

The term "mRNA localization" has been used in two different ways: (a) to signify the mere presence of a given mRNA in neurites; (b) as an enrichment of an mRNA in neurites compared to soma. Given that enrichment points to an active localization process, we decided to identify transcripts, which are consistently enriched in neurites versus soma across multiple datasets. We detected 61 transcripts significantly enriched in neurites in at least nine out of the analyzed 11 high coverage datasets (Table 2; given this small number of shared transcripts we were not able to perform further analyses, such as PCA). Strikingly, the majority of these transcripts (41 out of 61) encode ribosomal proteins. Interestingly, these transcripts are also ribosome-associated (Ainsley et al., 2014; Ouwenga et al., 2017; Shigeoka et al., 2016; Zappulo et al., 2017; Table 2 and extended online Table 2). These factors strongly hint at their local translation in

FIGURE 2 Comparison of high coverage datasets. (a, b) Principle component analysis (PCA) of RNA expression levels in neurites. Normalized transcripts per million (TPM) values (log10) for transcripts detected in at least three datasets with TPM > 10 (8,809 transcripts) were used for PCA. Plots show the distribution of datasets along the principle components (PC) 1 and 2 (a) or 1 and 3 (b). The way of obtaining neurons is indicated by color and type of neurons by the shape of the dot. (c) Heatmaps for the top 200 genes contributing to the first three PCs. Normalized TPM values (log10) were used for the heatmaps. Transcripts are sorted by hierarchical clustering using Euclidian distance and complete method. The specific 200 transcripts from PC are listed in subsets of extended online Table 1. (d) Gene ontology (GO) analysis of the top 200 transcripts contributing to the PC1 and PC3. Five most significant GO terms are shown. For PC2, the genome browser view of 3′-mRNA-seq peak for *Ccnd2*, mapping outside the annotated gene regions, is shown. Similar pattern is characteristic for 16 of the top 25 genes contributing to PC2 (Figure 1). (e) Heatmap of cell type and synaptic markers. Normalized TPM values (log10) were used for the heatmaps. A colored bar indicates which type of neuronal extensions were analyzed in each dataset (green: axons; blue: neurites, or neuropil)

TABLE 2 Transcripts enriched in neurites across at least nine datasets

Gene name	Function	Datasets with significant neurite enrichment ($p < .1$)	Average of significant ($p < .1$) log2 enrichment	Datasets with neurite TPM > 10	Detection rank (extended online Table 1)	Studies with neurite ribosome association
Rps18	Ribosomal protein	11	1.18	15	74	1
Fau	Ribosomal protein	11	1.13	17	12	1
Rab13	Signaling; neurite outgrowth (Sakane, Honda, & Sasaki, 2010)	10	2.45	10	2,662	1
Rplp1	Ribosomal protein	10	1.70	16	22	2
Rps12	Ribosomal protein	10	1.68	17	5	1
Rps2	Ribosomal protein	10	1.64	16	24	0
Rps28	Ribosomal protein	10	1.61	14	207	1
Lgals1	Lectin	10	1.55	10	2,509	1
Rpl35	Ribosomal protein	10	1.50	14	226	1
Rps19	Ribosomal protein	10	1.46	14	200	2
Rps23	Ribosomal protein	10	1.35	16	25	0
Rpl18a	Ribosomal protein	10	1.33	15	76	1
Rpl11	Ribosomal protein	10	1.30	14	204	2
Rps8	Ribosomal protein	10	1.28	17	6	1
Rpl28	Ribosomal protein	10	1.27	13	432	0
Rps14	Ribosomal protein	10	1.23	16	26	1
Rps9	Ribosomal protein	10	1.21	14	193	2
Rps16	Ribosomal protein	10	1.19	14	203	1
Cox6a1	Mitochondrial function	10	1.17	16	34	0
Rps26	Ribosomal protein	10	1.10	14	217	1
Rps11	Ribosomal protein	10	1.10	16	33	2
Cox8a	Mitochondrial function	10	1.05	14	197	1
Usf2	Ca^{2+} responsive transcription (W. G. Chen et al., 2003)	10	0.79	12	1,262	0
Nes	Outgrowth formation; cytoskeleton (Bott et al., 2019)	9	1.98	6	6,304	2
Eif4ebp1	Translation machinery	9	1.86	11	2,138	0
Anp32b	Nuclear protein; anti-apoptotic	9	1.83	13	507	1
Cyb5r3	Mitochondrial function	9	1.78	12	903	1
Rpl37	Ribosomal protein	9	1.61	15	80	1
S100a13	Ca^{2+} responsive protein secretion (Kathir et al., 2007)	9	1.51	8	4,682	1
Rpl36	Ribosomal protein	9	1.48	16	37	1
Rpl32	Ribosomal protein	9	1.42	14	196	2
Rpl38	Ribosomal protein	9	1.39	15	73	0

(Continues)

TABLE 2 (Continued)

Gene name	Function	Datasets with significant neurite enrichment ($p < .1$)	Average of significant ($p < .1$) log2 enrichment	Datasets with neurite TPM > 10	Detection rank (extended online Table 1)	Studies with neurite ribosome association
Tpt1	Outgrowth formation; mitochondria regulation (Roque et al., 2016)	9	1.30	18	2	2
Rhoc	Plasma membrane	9	1.27	11	1,732	1
Rps17	Ribosomal protein	9	1.26	14	192	2
Ybx1	RNA binding; multifunctional	9	1.25	17	4	0
Rps3	Ribosomal protein	9	1.19	15	88	0
Rps13	Ribosomal protein	9	1.19	14	245	1
Rps6	Ribosomal protein	9	1.13	15	77	1
Gnl1	Signaling	9	1.07	13	694	2
Eif3f	Translation machinery	9	1.02	15	111	2
Gpx4	Oxidative stress protection; mitochondrial regulation	9	0.97	12	894	1
Rpl3	Ribosomal protein	9	0.91	16	28	2
Atox1	Oxidative stress protection; copper transport	9	0.90	12	1,021	1
Cox4i1	Mitochondrial function	9	0.88	14	209	2
Rpl6	Ribosomal protein	9	0.85	17	9	0

Note: Table of all transcripts that are significantly enriched in at least nine out of the 11 analyzed datasets (Figure 1, left of dashed line). The data are presented largely as in Table 1, with the following extra columns: number of datasets reporting significant neurite enrichment, average of significant log2 neurite enrichment, rank of enriched transcript in extended online Table 1.

neurites. Although ribosome assembly is generally assumed to happen in the nucleolus (reviewed in Klinge & Woolford, 2019), recent studies report that the cytosolic replacement of ribosomal proteins can serve as a mechanism of ribosome maintenance (Mathis et al., 2017; Shigeoka et al., 2019). Moreover, it has recently became clear that ribosomes differ in their protein composition, resulting in heterogenous ribosome pools that translate specific subsets of mRNAs (Genuth & Barna, 2018; Shi et al., 2017). Given consistent enrichment of mRNAs encoding for ribosomal proteins in neurites, it is tempting to speculate that local transcriptome is translated by specialized ribosomes.

Neurite-enriched transcripts include mRNAs that encode mitochondrial proteins (*Cox6a1*, *Cox8a*, *Cyb5r3*). These, alongside ribosomal proteins, were also the most abundant transcripts in neurites (Table 1). An analysis that emphasizes enrichment over mere presence points out a number of transcripts which are less abundant but consistently enriched in neurites (Table 2). These include the Ca^{2+}-binding protein S100-a13, which facilitates signal peptide-independent protein secretion (Kathir et al., 2007); Nestin, involved in axon growth cone formation and early axon guidance (Bott et al., 2019); *Upstream stimulatory factor 2* (*Usf2*) implicated in response to Ca^{2+}-activated signaling pathways in neurons (W. G. Chen et al., 2003); and Rab-13, a key regulator of membrane trafficking and neurite outgrowth (Sakane et al., 2010) (Table 2).

Noteworthy, most transcripts commonly used as markers of axons, dendrites, and growth cones (such as *Actb*, *Map1b*, *Map2*, *Dlg4*, *Camk2a*, *Fmr1*, *Gria1*, *Nrgn*, *Grin1*, *Bdnf*, *Arc*) are not robustly enriched in neurites across multiple datasets, but rather equally distributed or even somatically enriched (Table 3). This is not surprising, taking into account that most of these markers have been established based on their detection in axons and dendrites using microscopy rather than through comparisons of their enrichment in neurites versus soma. Thus, depending on the

TABLE 3 Relative neurite/soma expression levels of established dendritic and axonal markers

Gene name	Detection rank (Extended online Table 1)	Datasets with neurite TPM > 10	Mean TPM in neurites	Enrichment rank (Table S2)	Datasets with significant neurite enrichment ($p < .1$)	Average of significant ($p < .1$) log2 enrichment	Studies with neurite ribosome association
Actb	1	18	1,201.03	3,787	2	0.05	1
Map1b	11	17	489.71	6,281	1	−0.68	0
Map2	222	14	285.00	768	5	−0.16	1
Dlg4	939	12	118.08	2,709	3	−0.50	1
Fmr1	2,925	10	36.47	5,319	1	−0.14	0
Gria1	3,596	9	58.45	8,353	0	−1.42	1
Camk2a	6,333	6	125.58	5,178	1	−0.05	2
Nrgn	7,320	5	158.28	NA	0	NA	0
Arc	7,509	5	44.22	8,714	0	NA	2
Grin1	8,457	4	71.39	8,280	0	−1.33	2
Bdnf	11,634	2	33.34	NA	0	NA	0

Note: The data are presented largely in Table 2, with the extra column: rank in extended online Table 2. For information on enrichment in specific datasets see extended online Table 2.

methodology, the term "localized mRNA" could mean either presence or enrichment of the given mRNA in the specified subcellular compartment.

5 | CONCLUDING REMARKS

Advances in sequencing and imaging techniques have made it possible to carry out transcriptomic profiling at subcellular resolutions. Our comprehensive analysis of relevant datasets has identified a core set of transcripts localized and enriched in neurites that is common to multiple types of neurons from three species. This set includes mRNAs encoding ribosomal and mitochondrial proteins, cytoskeletal components, factors involved in the formation of neural processes, and nuclear proteins that may regulate neuronal gene expression in response to localized activity in neurites (Tables 1 and 2). The commonly accepted synaptic mRNA markers are mostly detected, but not enriched in neurites. Remarkably, differences in neurite transcriptomes do not suffice to distinguish different types of primary neurons (Figure 2). This might be explained in two ways: (a) the core transcriptome of neurites we have identified is highly conserved between different neuronal types, hinting at potentially important general functions of localized transcripts; (b) differences between cell types might be masked by sample heterogeneity, through factors such as the presence of multiple neuronal types, contamination with non-neuronal cells, the maturity of the neurons analyzed and other confounders which are more highly represented in the dataset than type-specific factors. Our analysis of soma transcriptomes (Figure S2) suggests that the second explanation is more likely. Clarifying this issue will require studying diverse types of neurons using the same experimental workflow. Advances in new technologies that can generate high-throughput data from intact tissue at a subcellular resolution will surely provide more insights into processes that govern RNA localization in cell-type-specific ways.

In addition to the core neurite transcriptome, our comparative analysis allows informed choice of the test system for future localization studies. Due to their homogeneity, ease of genetic manipulation and compliance with the replacement principle aimed at reducing animal use in research, in vitro differentiated neurons and neuronal cell lines represent a system of choice for general mechanistic studies. They, however, do not express as many synaptic markers as primary neurons (Figure 2e) and should be carefully evaluated before using them for studies on biological functions of specific neuronal genes. Our extended online Table 1 allows to test whether the expression pattern of a transcript of interest is recapitulated in one of the neuronal cell lines or stem cell-derived neurons and whether they would therefore represent a proper test system, or studies should be preferably performed in primary cells. Our analysis thus provides an indispensable resource for future localization studies.

6 | METHODS

Raw RNA-seq data were downloaded from the NCBI-GEO or EBI ArrayExpress databases (Table S1) and processed using the PiGx RNA-seq pipeline version 0.0.10 (Wurmus et al., 2018) using default settings and ENSEMBLE genome assemblies for mouse (GRCm38.p6, release 96), human (GRCh38.p13, release 97), and rat (Rnor_6.0, release 98). Based on the initial analysis reports generated by the PiGx pipeline individual replicates that did not separate well between neurite and soma samples were excluded from further analysis (Table S1). Transcript count data based on genome mapped reads (STAR counts) generated by the pipeline were gathered, normalized to TPM and averaged for neurite and soma samples.

Transcript count data or detection status from studies that did not deposit raw datasets (Maciel et al., 2018; Poulopoulos et al., 2019; Rotem et al., 2017) or did not use RNA-seq methods (Cajigas et al., 2012; Feltrin et al., 2012; Lein et al., 2007; Poon et al., 2006; Saal et al., 2014; Taylor et al., 2009) were obtained from the corresponding supplementary materials of those studies and normalized accordingly or treated as binary detected or not detected. All Ensembl gene IDs from human or rat were mapped to orthologous mouse IDs using Ensembl Biomart database. To prevent species-specific genes from affecting analysis, genes without orthologs between mouse, human and rat were excluded from comparative analysis.

All further analyses were performed using custom R scripts: PC analysis was performed with log10 normalized average neurite expression values. Hierarchical clustering of genes and datasets for heatmaps was performed using an Euclidian distance metric and the "complete" method. Gene ontology analysis was performed using the gProfileR package (Reimand, Kull, Peterson, Hansen, & Vilo, 2007), functional terms with more than 2000 proteins were excluded. We considered those transcripts detected in neurites that showed an average neurite expression level of TPM greater than 10 in at least three independent studies (extended online Table 1).

For the comparative analysis of different datasets and calculation of transcript enrichment in neurites only those datasets with high coverage (>5,000 transcripts with TPM > 10 detected) were chosen. We considered those transcripts for PCA analysis that had average expression levels of TPM > 10 both in neurite and soma samples of at least three independent datasets (8,809 transcripts). We decided to exclude transcripts detected in only one or two studies to ensure that differences between all included transcripts represent biological variability and not technical differences from different experimental setups. Calculation of transcript enrichment in neurites was performed using the DEseq2 R package (Love, Huber, & Anders, 2014), using raw counts for all transcripts with TPM > 1 and accounting for pairing of neurite and soma samples as a covariate. We considered those transcripts enriched in neurites that passed criteria for inclusion in PCA analysis and showed significant ($p < 0.1$) enrichment (log2-fold change >0) in at least one dataset (extended online Table 2).

ACKNOWLEDGMENTS

We thank Vedran Franke for discussions on data analysis strategies, Russ Hodge and members of Chekulaeva lab for helpful comments on the manuscript and data representation. Data analysis was performed by N.K. and the manuscript written by N.K. and M.C.

CONFLICT OF INTEREST

The authors have declared no conflicts of interest for this article.

AUTHOR CONTRIBUTIONS

Nicolai von Kügelgen: Formal analysis and visualization. **Marina Chekulaeva**: Conceptualization, writing, and editing.

ORCID

Nicolai von Kügelgen https://orcid.org/0000-0001-8116-9032
Marina Chekulaeva https://orcid.org/0000-0002-6868-7887

RELATED WIREs ARTICLES

Classical and emerging techniques to identify and quantify localized RNAs
Local translation of mRNAs in neural development

REFERENCES

Ainsley, J. A., Drane, L., Jacobs, J., Kittelberger, K. A., & Reijmers, L. G. (2014). Functionally diverse dendritic mRNAs rapidly associate with ribosomes following a novel experience. *Nature Communications*, *5*, 4510. https://doi.org/10.1038/ncomms5510

Armstrong, A., Ryu, Y. K., Chieco, D., & Kuruvilla, R. (2011). Frizzled3 is required for neurogenesis and target innervation during sympathetic nervous system development. *The Journal of Neuroscience*, *31*(7), 2371–2381. https://doi.org/10.1523/JNEUROSCI.4243-10.2011

Bassell, G. J., Zhang, H., Byrd, A. L., Femino, A. M., Singer, R. H., Taneja, K. L., ... Kosik, K. S. (1998). Sorting of beta-actin mRNA and protein to neurites and growth cones in culture. *The Journal of Neuroscience*, *18*(1), 251–265.

Becht, E., McInnes, L., Healy, J., Dutertre, C. A., Kwok, I. W. H., Ng, L. G., ... Newell, E. W. (2018). Dimensionality reduction for visualizing single-cell data using UMAP. *Nature Biotechnology*, *37*, 38–44. https://doi.org/10.1038/nbt.4314

Bott, C. J., Johnson, C. G., Yap, C. C., Dwyer, N. D., Litwa, K. A., & Winckler, B. (2019). Nestin in immature embryonic neurons affects axon growth cone morphology and Semaphorin3a sensitivity. *Molecular Biology of the Cell*, *30*(10), 1214–1229. https://doi.org/10.1091/mbc.E18-06-0361

Briese, M., Saal, L., Appenzeller, S., Moradi, M., Baluapuri, A., & Sendtner, M. (2016). Whole transcriptome profiling reveals the RNA content of motor axons. *Nucleic Acids Research*, *44*(4), e33. https://doi.org/10.1093/nar/gkv1027

Cajigas, I. J., Tushev, G., Will, T. J., tom Dieck, S., Fuerst, N., & Schuman, E. M. (2012). The local transcriptome in the synaptic neuropil revealed by deep sequencing and high-resolution imaging. *Neuron*, *74*(3), 453–466. https://doi.org/10.1016/j.neuron.2012.02.036

Chanda, S., Ang, C. E., Davila, J., Pak, C., Mall, M., Lee, Q. Y., ... Wernig, M. (2014). Generation of induced neuronal cells by the single reprogramming factor ASCL1. *Stem Cell Reports*, *3*(2), 282–296. https://doi.org/10.1016/j.stemcr.2014.05.020

Chekulaeva, M., & Landthaler, M. (2016). Eyes on translation. *Molecular Cell*, *63*(6), 918–925. https://doi.org/10.1016/j.molcel.2016.08.031

Chen, F., Tillberg, P. W., & Boyden, E. S. (2015). Optical imaging. Expansion Microscopy. *Science*, *347*(6221), 543–548. https://doi.org/10.1126/science.1260088

Chen, K. H., Boettiger, A. N., Moffitt, J. R., Wang, S., & Zhuang, X. (2015). RNA imaging. Spatially resolved, highly multiplexed RNA profiling in single cells. *Science*, *348*(6233), aaa6090. https://doi.org/10.1126/science.aaa6090

Chen, W. G., West, A. E., Tao, X., Corfas, G., Szentirmay, M. N., Sawadogo, M., ... Greenberg, M. E. (2003). Upstream stimulatory factors are mediators of Ca^{2+}-responsive transcription in neurons. *The Journal of Neuroscience*, *23*(7), 2572–2581.

Ciolli Mattioli, C., Rom, A., Franke, V., Imami, K., Arrey, G., Terne, M., ... Chekulaeva, M. (2019). Alternative 3′ UTRs direct localization of functionally diverse protein isoforms in neuronal compartments. *Nucleic Acids Research*, *47*(5), 2560–2573. https://doi.org/10.1093/nar/gky1270

Cornell, B., Wachi, T., Zhukarev, V., & Toyo-Oka, K. (2016). Regulation of neuronal morphogenesis by 14-3-3epsilon (Ywhae) via the microtubule binding protein, doublecortin. *Human Molecular Genetics*, *25*(20), 4405–4418. https://doi.org/10.1093/hmg/ddw270

Dinamarca, M. C., Guzzetti, F., Karpova, A., Lim, D., Mitro, N., Musardo, S., ... Di Luca, M. (2016). Ring finger protein 10 is a novel synaptonuclear messenger encoding activation of NMDA receptors in hippocampus. *eLife*, *5*, e12430. https://doi.org/10.7554/eLife.12430

Donnelly, C. J., Willis, D. E., Xu, M., Tep, C., Jiang, C., Yoo, S., ... Twiss, J. L. (2011). Limited availability of ZBP1 restricts axonal mRNA localization and nerve regeneration capacity. *The EMBO Journal*, *30*(22), 4665–4677. https://doi.org/10.1038/emboj.2011.347

Dunkley, P. R., Jarvie, P. E., & Robinson, P. J. (2008). A rapid Percoll gradient procedure for preparation of synaptosomes. *Nature Protocols*, *3*(11), 1718–1728. https://doi.org/10.1038/nprot.2008.171

Farris, S., Ward, J. M., Carstens, K. E., Samadi, M., Wang, Y., & Dudek, S. M. (2019). Hippocampal subregions express distinct dendritic transcriptomes that reveal differences in mitochondrial function in CA2. *Cell Reports*, *29*(2), 522–539 e526. https://doi.org/10.1016/j.celrep.2019.08.093

Fazal, F. M., Han, S., Parker, K. R., Kaewsapsak, P., Xu, J., Boettiger, A. N., ... Ting, A. Y. (2019). Atlas of subcellular RNA localization revealed by APEX-Seq. *Cell*, *178*(2), 473–490 e426. https://doi.org/10.1016/j.cell.2019.05.027

Feltrin, D., Fusco, L., Witte, H., Moretti, F., Martin, K., Letzelter, M., ... Pertz, O. (2012). Growth cone MKK7 mRNA targeting regulates MAP1b-dependent microtubule bundling to control neurite elongation. *PLoS Biology*, *10*(12), e1001439. https://doi.org/10.1371/journal.pbio.1001439

Frey, D., Laux, T., Xu, L., Schneider, C., & Caroni, P. (2000). Shared and unique roles of CAP23 and GAP43 in actin regulation, neurite outgrowth, and anatomical plasticity. *The Journal of Cell Biology*, *149*(7), 1443–1454. https://doi.org/10.1083/jcb.149.7.1443

Genuth, N. R., & Barna, M. (2018). The discovery of ribosome heterogeneity and its implications for gene regulation and organismal life. *Molecular Cell*, *71*(3), 364–374. https://doi.org/10.1016/j.molcel.2018.07.018

Glock, C., Heumuller, M., & Schuman, E. M. (2017). mRNA transport & local translation in neurons. *Current Opinion in Neurobiology*, *45*, 169–177. https://doi.org/10.1016/j.conb.2017.05.005

Gumy, L. F., Yeo, G. S., Tung, Y. C., Zivraj, K. H., Willis, D., Coppola, G., ... Fawcett, J. W. (2011). Transcriptome analysis of embryonic and adult sensory axons reveals changes in mRNA repertoire localization. *RNA*, *17*(1), 85–98. https://doi.org/10.1261/rna.2386111

He, W., Lu, Y., Qahwash, I., Hu, X. Y., Chang, A., & Yan, R. (2004). Reticulon family members modulate BACE1 activity and amyloid-beta peptide generation. *Nature Medicine*, *10*(9), 959–965. https://doi.org/10.1038/nm1088

Heinrich, C., Gascon, S., Masserdotti, G., Lepier, A., Sanchez, R., Simon-Ebert, T., ... Berninger, B. (2011). Generation of subtype-specific neurons from postnatal astroglia of the mouse cerebral cortex. *Nature Protocols*, *6*(2), 214–228. https://doi.org/10.1038/nprot.2010.188

Kathir, K. M., Ibrahim, K., Rajalingam, D., Prudovsky, I., Yu, C., & Kumar, T. K. (2007). S100A13-lipid interactions-role in the non-classical release of the acidic fibroblast growth factor. *Biochimica et Biophysica Acta*, *1768*(12), 3080–3089. https://doi.org/10.1016/j.bbamem.2007.09.007

Ke, R., Mignardi, M., Pacureanu, A., Svedlund, J., Botling, J., Wahlby, C., & Nilsson, M. (2013). In situ sequencing for RNA analysis in preserved tissue and cells. *Nature Methods*, *10*(9), 857–860. https://doi.org/10.1038/nmeth.2563

Klinge, S., & Woolford, J. L., Jr. (2019). Ribosome assembly coming into focus. *Nature Reviews. Molecular Cell Biology*, *20*(2), 116–131. https://doi.org/10.1038/s41580-018-0078-y

La Manno, G., Gyllborg, D., Codeluppi, S., Nishimura, K., Salto, C., Zeisel, A., … Linnarsson, S. (2016). Molecular diversity of midbrain development in mouse, human, and stem cells. *Cell*, *167*(2), 566–580 e519. https://doi.org/10.1016/j.cell.2016.09.027

Lee, J. H., Daugharthy, E. R., Scheiman, J., Kalhor, R., Yang, J. L., Ferrante, T. C., … Church, G. M. (2014). Highly multiplexed subcellular RNA sequencing in situ. *Science*, *343*(6177), 1360–1363. https://doi.org/10.1126/science.1250212

Lein, E. S., Hawrylycz, M. J., Ao, N., Ayres, M., Bensinger, A., Bernard, A., … Jones, A. R. (2007). Genome-wide atlas of gene expression in the adult mouse brain. *Nature*, *445*(7124), 168–176. https://doi.org/10.1038/nature05453

Liu, Y., Miao, Q., Yuan, J., Han, S., Zhang, P., Li, S., … Cheng, L. (2015). Ascl1 converts dorsal midbrain astrocytes into functional neurons in vivo. *The Journal of Neuroscience*, *35*(25), 9336–9355. https://doi.org/10.1523/JNEUROSCI.3975-14.2015

Lobo, M. K., Karsten, S. L., Gray, M., Geschwind, D. H., & Yang, X. W. (2006). FACS-array profiling of striatal projection neuron subtypes in juvenile and adult mouse brains. *Nature Neuroscience*, *9*(3), 443–452. https://doi.org/10.1038/nn1654

Love, M. I., Huber, W., & Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biology*, *15*(12), 550. https://doi.org/10.1186/s13059-014-0550-8

Lubeck, E., Coskun, A. F., Zhiyentayev, T., Ahmad, M., & Cai, L. (2014). Single-cell in situ RNA profiling by sequential hybridization. *Nature Methods*, *11*(4), 360–361. https://doi.org/10.1038/nmeth.2892

Ludwik, K. A., von Kuegelgen, N., & Chekulaeva, M. (2019). Genome-wide analysis of RNA and protein localization and local translation in mESC-derived neurons. *Methods*, *162–163*, 31–41. https://doi.org/10.1016/j.ymeth.2019.02.002

Maciel, R., Bis, D. M., Rebelo, A. P., Saghira, C., Zuchner, S., & Saporta, M. A. (2018). The human motor neuron axonal transcriptome is enriched for transcripts related to mitochondrial function and microtubule-based axonal transport. *Experimental Neurology*, *307*, 155–163. https://doi.org/10.1016/j.expneurol.2018.06.008

Maggi, L. B., Kuchenruether, M., Dadey, D. Y. A., Schwope, R. M., Grisendi, S., Townsend, R. R., … Weber, J. D. (2008). Nucleophosmin serves as a rate-limiting nuclear export chaperone for the mammalian ribosome. *Molecular and Cellular Biology*, *28*(23), 7050–7065. https://doi.org/10.1128/Mcb.01548-07

Mardia, K. V., Kent, J. T., & Bibby, J. M. (1979). *Multivariate analysis*. London, England/ New York, NY: Academic Press.

Mathis, A. D., Naylor, B. C., Carson, R. H., Evans, E., Harwell, J., Knecht, J., … Price, J. C. (2017). Mechanisms of in vivo ribosome maintenance change in response to nutrient signals. *Molecular & Cellular Proteomics*, *16*(2), 243–254. https://doi.org/10.1074/mcp.M116.063255

Micheva, K. D., Vallee, A., Beaulieu, C., Herman, I. M., & Leclerc, N. (1998). Beta-actin is confined to structures having high capacity of remodelling in developing and adult rat cerebellum. *The European Journal of Neuroscience*, *10*(12), 3785–3798. https://doi.org/10.1046/j.1460-9568.1998.00391.x

Middleton, S. A., Eberwine, J., & Kim, J. (2019). Comprehensive catalog of dendritically localized mRNA isoforms from sub-cellular sequencing of single mouse neurons. *BMC Biology*, *17*(1), 5. https://doi.org/10.1186/s12915-019-0630-z

Minis, A., Dahary, D., Manor, O., Leshkowitz, D., Pilpel, Y., & Yaron, A. (2014). Subcellular transcriptomics-dissection of the mRNA composition in the axonal compartment of sensory neurons. *Developmental Neurobiology*, *74*(3), 365–381. https://doi.org/10.1002/dneu.22140

Ouwenga, R., Lake, A. M., O'Brien, D., Mogha, A., Dani, A., & Dougherty, J. D. (2017). Transcriptomic analysis of ribosome-bound mRNA in cortical neurites in vivo. *The Journal of Neuroscience*, *37*(36), 8688–8705. https://doi.org/10.1523/JNEUROSCI.3044-16.2017

Pertz, O., Hodgson, L., Klemke, R. L., & Hahn, K. M. (2006). Spatiotemporal dynamics of RhoA activity in migrating cells. *Nature*, *440*(7087), 1069–1072. https://doi.org/10.1038/nature04665

Poon, M. M., Choi, S. H., Jamieson, C. A., Geschwind, D. H., & Martin, K. C. (2006). Identification of process-localized mRNAs from cultured rodent hippocampal neurons. *The Journal of Neuroscience*, *26*(51), 13390–13399. https://doi.org/10.1523/JNEUROSCI.3432-06.2006

Poulopoulos, A., Murphy, A. J., Ozkan, A., Davis, P., Hatch, J., Kirchner, R., & Macklis, J. D. (2019). Subcellular transcriptomes and proteomes of developing axon projections in the cerebral cortex. *Nature*, *565*(7739), 356–360. https://doi.org/10.1038/s41586-018-0847-y

Reimand, J., Kull, M., Peterson, H., Hansen, J., & Vilo, J. (2007). g:Profiler—A web-based toolset for functional profiling of gene lists from large-scale experiments. *Nucleic Acids Research*, *35*, W193–W200. https://doi.org/10.1093/nar/gkm226

Rodriguez-Munoz, M., Cortes-Montero, E., Pozo-Rodrigalvarez, A., Sanchez-Blazquez, P., & Garzon-Nino, J. (2015). The ON:OFF switch, sigma1R-HINT1 protein, controls GPCR-NMDA receptor cross-regulation: Implications in neurological disorders. *Oncotarget*, *6*(34), 35458–35477. https://doi.org/10.18632/oncotarget.6064

Roque, C. G., Wong, H. H., Lin, J. Q., & Holt, C. E. (2016). Tumor protein Tctp regulates axon development in the embryonic visual system. *Development*, *143*(7), 1134–1148. https://doi.org/10.1242/dev.131060

Rotem, N., Magen, I., Ionescu, A., Gershoni-Emek, N., Altman, T., Costa, C. J., … Perlson, E. (2017). ALS along the axons—Expression of coding and noncoding RNA differs in axons of ALS models. *Scientific Reports*, *7*, 44500. https://doi.org/10.1038/srep44500

Saal, L., Briese, M., Kneitz, S., Glinka, M., & Sendtner, M. (2014). Subcellular transcriptome alterations in a cell culture model of spinal muscular atrophy point to widespread defects in axonal growth and presynaptic differentiation. *RNA*, *20*(11), 1789–1802. https://doi.org/10.1261/rna.047373.114

Sakane, A., Honda, K., & Sasaki, T. (2010). Rab13 regulates neurite outgrowth in PC12 cells through its effector protein, JRAB/MICAL-L2. *Molecular and Cellular Biology*, *30*(4), 1077–1087. https://doi.org/10.1128/Mcb.01067-09

Shah, S., Lubeck, E., Schwarzkopf, M., He, T. F., Greenbaum, A., Sohn, C. H., … Cai, L. (2016). Single-molecule RNA detection at depth by hybridization chain reaction and tissue hydrogel embedding and clearing. *Development*, *143*(15), 2862–2867. https://doi.org/10.1242/dev.138560

Shah, S., Lubeck, E., Zhou, W., & Cai, L. (2016). In situ transcription profiling of single cells reveals spatial organization of cells in the mouse hippocampus. *Neuron*, *92*(2), 342–357. https://doi.org/10.1016/j.neuron.2016.10.001

Shah, S., Lubeck, E., Zhou, W., & Cai, L. (2017). seqFISH accurately detects transcripts in single cells and reveals robust spatial organization in the hippocampus. *Neuron*, *94*(4), 752–758 e751. https://doi.org/10.1016/j.neuron.2017.05.008

Shi, Z., Fujii, K., Kovary, K. M., Genuth, N. R., Rost, H. L., Teruel, M. N., & Barna, M. (2017). Heterogeneous ribosomes preferentially translate distinct subpools of mRNAs genome-wide. *Molecular Cell*, *67*(1), 71–83 e77. https://doi.org/10.1016/j.molcel.2017.05.021

Shigeoka, T., Jung, H., Jung, J., Turner-Bridger, B., Ohk, J., Lin, J. Q., ... Holt, C. E. (2016). Dynamic axonal translation in developing and mature visual circuits. *Cell*, *166*(1), 181–192. https://doi.org/10.1016/j.cell.2016.05.029

Shigeoka, T., Koppers, M., Wong, H. H., Lin, J. Q., Cagnetta, R., Dwivedy, A., ... Holt, C. E. (2019). On-site ribosome remodeling by locally synthesized ribosomal proteins in axons. *Cell Reports*, *29*(11), 3605–3619 e3610. https://doi.org/10.1016/j.celrep.2019.11.025

Simone, N. L., Bonner, R. F., Gillespie, J. W., Emmert-Buck, M. R., & Liotta, L. A. (1998). Laser-capture microdissection: Opening the microscopic frontier to molecular analysis. *Trends in Genetics*, *14*(7), 272–276. https://doi.org/10.1016/s0168-9525(98)01489-9

Stahl, P. L., Salmen, F., Vickovic, S., Lundmark, A., Navarro, J. F., Magnusson, J., ... Frisen, J. (2016). Visualization and analysis of gene expression in tissue sections by spatial transcriptomics. *Science*, *353*(6294), 78–82. https://doi.org/10.1126/science.aaf2403

Taliaferro, J. M., Vidaki, M., Oliveira, R., Olson, S., Zhan, L., Saxena, T., ... Burge, C. B. (2016). Distal alternative last exons localize mRNAs to neural projections. *Molecular Cell*, *61*(6), 821–833. https://doi.org/10.1016/j.molcel.2016.01.020

Taylor, A. M., Berchtold, N. C., Perreau, V. M., Tu, C. H., Li Jeon, N., & Cotman, C. W. (2009). Axonal mRNA in uninjured and regenerating cortical mammalian axons. *The Journal of Neuroscience*, *29*(15), 4697–4707. https://doi.org/10.1523/JNEUROSCI.6130-08.2009

Taylor, A. M., Dieterich, D. C., Ito, H. T., Kim, S. A., & Schuman, E. M. (2010). Microfluidic local perfusion chambers for the visualization and manipulation of synapses. *Neuron*, *66*(1), 57–68. https://doi.org/10.1016/j.neuron.2010.03.022

Toth, E. N., Lohith, A., Mondal, M., Guo, J., Fukamizu, A., & Pourmand, N. (2018). Single-cell nanobiopsy reveals compartmentalization of mRNAs within neuronal cells. *The Journal of Biological Chemistry*, *293*(13), 4940–4951. https://doi.org/10.1074/jbc.M117.800763

Tushev, G., Glock, C., Heumuller, M., Biever, A., Jovanovic, M., & Schuman, E. M. (2018). Alternative 3′ UTRs modify the localization, regulatory potential, stability, and plasticity of mRNAs in neuronal compartments. *Neuron*, *98*(3), 495–511 e496. https://doi.org/10.1016/j.neuron.2018.03.030

van der Maaten, L., & Hinton, G. (2008). Visualizing data using t-SNE. *Journal of Machine Learning Research*, *9*, 2579–2605.

Venables, W. N., Ripley, B. D., & Venables, W. N. (2002). *Modern applied statistics with S* (4th ed.). New York, NY: Springer.

Vickovic, S., Eraslan, G., Salmen, F., Klughammer, J., Stenbeck, L., Schapiro, D., ... Stahl, P. L. (2019). High-definition spatial transcriptomics for in situ tissue profiling. *Nature Methods*, *16*(10), 987–990. https://doi.org/10.1038/s41592-019-0548-y

Wachi, T., Cornell, B., Marshall, C., Zhukarev, V., Baas, P. W., & Toyo-oka, K. (2016). Ablation of the 14-3-3gamma protein results in neuronal migration delay and morphological defects in the developing cerebral cortex. *Developmental Neurobiology*, *76*(6), 600–614. https://doi.org/10.1002/dneu.22335

Wang, P., Tang, W., Li, Z., Zou, Z., Zhou, Y., Li, R., ... Zou, P. (2019). Mapping spatial transcriptome with light-activated proximity-dependent RNA labeling. *Nature Chemical Biology*, *15*(11), 1110–1119. https://doi.org/10.1038/s41589-019-0368-5

Wichterle, H., & Peljto, M. (2008). Differentiation of mouse embryonic stem cells to spinal motor neurons. *Current Protocols in Stem Cell Biology*, *1*, 1H.1.1–1H.1.9. https://doi.org/10.1002/9780470151808.sc01h01s5

Willemsen, M. H., Ba, W., Wissink-Lindhout, W. M., de Brouwer, A. P., Haas, S. A., Bienek, M., ... Kleefstra, T. (2014). Involvement of the kinesin family members KIF4A and KIF5C in intellectual disability and synaptic function. *Journal of Medical Genetics*, *51*(7), 487–494. https://doi.org/10.1136/jmedgenet-2013-102182

Wurmus, R., Uyar, B., Osberg, B., Franke, V., Gosdschan, A., Wreczycka, K., ... Akalin, A. (2018). PiGx: Reproducible genomics analysis pipelines with GNU Guix. *GigaScience*, *7*(12), giy123. https://doi.org/10.1093/gigascience/giy123

Xia, C. H., Roberts, E. A., Her, L. S., Liu, X., Williams, D. S., Cleveland, D. W., & Goldstein, L. S. (2003). Abnormal neurofilament transport caused by targeted disruption of neuronal kinesin heavy chain KIF5A. *The Journal of Cell Biology*, *161*(1), 55–66. https://doi.org/10.1083/jcb.200301026

Zappulo, A., van den Bruck, D., Ciolli Mattioli, C., Franke, V., Imami, K., McShane, E., ... Chekulaeva, M. (2017). RNA localization is a key determinant of neurite-enriched proteome. *Nature Communications*, *8*(1), 583. https://doi.org/10.1038/s41467-017-00690-6

Zhang, Y., Chen, K., Sloan, S. A., Bennett, M. L., Scholze, A. R., O'Keeffe, S., ... Wu, J. Q. (2014). An RNA-sequencing transcriptome and splicing database of glia, neurons, and vascular cells of the cerebral cortex. *The Journal of Neuroscience*, *34*(36), 11929–11947. https://doi.org/10.1523/JNEUROSCI.1860-14.2014

Zivraj, K. H., Tung, Y. C., Piper, M., Gumy, L., Fawcett, J. W., Yeo, G. S., & Holt, C. E. (2010). Subcellular profiling reveals distinct and developmentally regulated repertoire of growth cone mRNAs. *The Journal of Neuroscience*, *30*(46), 15464–15478. https://doi.org/10.1523/JNEUROSCI.1800-10.2010

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of this article.

How to cite this article: von Kügelgen N, Chekulaeva M. Conservation of a core neurite transcriptome across neuronal types and species. *WIREs RNA*. 2020;e1590. https://doi.org/10.1002/wrna.1590