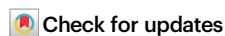


# Simultaneous profiling of native-state proteomes and transcriptomes of neural cell types using proximity labeling

Received: 31 January 2025

Accepted: 10 March 2026

Published online: 31 March 2026



Christina C. Ramelow<sup>1,2,3</sup>, Eric B. Dammer<sup>2,4</sup>, Hailian Xiao<sup>1,2</sup>, Lihong Cheng<sup>1,5</sup>, Prateek Kumar<sup>6</sup>, Claudia Espinosa-Garcia<sup>6</sup>, Maureen M. Sampson<sup>3</sup>, Dilpreet Kour<sup>6</sup>, Ruth S. Nelson<sup>6</sup>, Sneha Malepati<sup>6</sup>, Rashmi Kumari<sup>6</sup>, Wooyoung Eric Jang<sup>6</sup>, Qi Guo<sup>2,4</sup>, Pritha Bagchi<sup>7</sup>, Duc M. Duong<sup>2,7</sup>, Nicholas T. Seyfried<sup>1,2,4,7</sup>, Steven A. Sloan<sup>3,8</sup>✉ & Srikant Rangaraju<sup>1,5,8</sup>✉

Phenotyping cells at transcriptomic and proteomic levels is an essential step to understanding cellular contributions to development, aging, injury, and disease. Since proteome and transcriptome level abundances modestly correlate, complementary profiling of both is needed. We report a method called simultaneous protein and RNA -omics (SPARO) to capture the cell type-specific transcriptome and proteome simultaneously in vitro using BV2 microglial and HEK293 cell lines and in vivo using astrocytic and neuronal Cre driver mice crossed with Rosa26-TurboID knock-in mice. SPARO leverages TurboID to biotinylate RNA-interacting cytosolic proteins, enabling enrichment of proteins for proteomics and protein-associated RNA for transcriptomics. We validate SPARO first using well-controlled in vitro systems to verify that the proteomes and transcriptomes obtained reflect the global proteomes and transcriptomes. The effect of neuroinflammatory activation by lipopolysaccharide is also faithfully captured. We apply SPARO to obtain native-state proteomes and transcriptomes from astrocytes and neurons, thereby validating the approach in vivo. We interrogate mRNA-protein concordance and discordance, providing insights into molecular processes that exhibit uniform or cell type-specific patterns.

The transcriptome and proteome of a cell provide complementary molecular information about its phenotype, state, and cellular functions. Quantifying RNA and protein levels in cell types using high-throughput-omics approaches uncover mechanisms and molecular pathways that contribute to development, aging, injury, and disease. RNA-sequencing can be leveraged to record different transcript species, expression patterns, gene structure, splicing patterns, and post-

transcriptional modifications<sup>1</sup>. On the other hand, mass spectrometry (MS)-based quantitative proteomics provides information on protein abundances, localization, post-translational modifications, structure, cell-to-cell interactions, and effector functions<sup>2</sup>. Importantly, the number of RNA transcripts for a given gene does not always correlate with its protein abundance<sup>3,4</sup>. This is straightforward to explain from a biological standpoint by considering the dynamics of RNA stability,

<sup>1</sup>Department of Neurology, Emory University, Atlanta, GA, USA. <sup>2</sup>Center for Neurodegenerative Disease, Emory University, Atlanta, GA, USA. <sup>3</sup>Department of Human Genetics, Emory University, Atlanta, GA, USA. <sup>4</sup>Department of Biochemistry, Emory University, Atlanta, GA, USA. <sup>5</sup>Department of Pharmacology and Chemical Biology, Emory University, Atlanta, GA, USA. <sup>6</sup>Department of Neurology, Yale University, New Haven, CT, USA. <sup>7</sup>Emory Integrated Proteomics Core, Emory University, Atlanta, GA, USA. <sup>8</sup>These authors jointly supervised this work: Steven A. Sloan, Srikant Rangaraju. ✉e-mail: [steven.a.sloan@emory.edu](mailto:steven.a.sloan@emory.edu); [srikant.rangaraju@yale.edu](mailto:srikant.rangaraju@yale.edu)

translational regulation, protein half-life, chaperone proteins, cellular sequestration, as well as other post-transcriptional and post-translational events<sup>5–9</sup>. This contributes to observed discordances between mRNA and protein abundances and molecular signatures in cells and tissues.

Measuring RNA and protein levels simultaneously can reveal changes at the RNA-level that are not reflected at the protein-level, and vice versa, and enables exploration of mechanisms regulating mRNA-protein correlation. This can have biological implications as well as highly practical consequences that impact the choice of RNA or protein markers in experiments (e.g., the choice of RNA probe for in situ hybridization or the choice of protein marker for spatial studies). Also, the level of discordance between mRNA and protein, and the mechanisms that control mRNA and protein levels may vary across different cell types, tissues, and biological contexts, and cannot be ascertained from bulk tissue -omics approaches. Therefore, we need new approaches to jointly measure mRNA and protein levels in cell-type-specific contexts.

Several attempts have been made to quantify this general concordance (or discordance) between transcriptomic and proteomic levels<sup>4,10–18</sup>. Simultaneous sampling of the transcriptome and proteome can be readily accomplished under in vitro monoculture conditions. However, under in vivo settings, these methods involve collecting RNA and protein from separate physical samples and rely on cell-type purification using mechanical approaches, which, by themselves, impact transcriptomic and molecular characteristics of the cells<sup>19</sup>. The ability to purify intact cells from complex tissues like the brain also varies across cell types. For example, adult neurons are difficult to isolate without loss of their cellular and synaptic architecture<sup>20,21</sup>.

A general approach to obtain cell-type-specific transcriptomes or proteomes from tissues that are independent of cell-type isolation involves protein tagging. For cell-type-specific in vivo transcriptomics, ribosomal subunits (e.g., Rpl22) can be tagged (e.g., HA) in a cell-type-specific manner using Cre/lox genetics, allowing capture of ribosome-associated mRNA species (translatome). For cell-type-specific proteomics, metabolic labeling with non-canonical amino acids<sup>22–27</sup> and proximity-labeling by biotin ligases<sup>28–39</sup> have been employed. There have also been several efforts to achieve joint mRNA and protein measurements, using RiboTag<sup>40–43</sup> and Translating Ribosome Affinity Purification (TRAP)<sup>44–47</sup> to capture mRNAs and nascent peptides from a single tagged ribosomal protein under a cell-type-specific promoter. However, these approaches limit profiling to only ribosome-bound transcripts and nascent peptides rather than the broader proteome and transcriptome of the cell. Further, mammalian cells contain at least 79 different ribosomal proteins for cytosolic translation<sup>48</sup>, and ribosomes with different protein compositions translate different groups of mRNAs<sup>49</sup>. Therefore, ribosome affinity purification-based methods that target a single ribosomal protein may not capture the full complexity of the RNA species within a cell.

Another method to achieve joint mRNA and protein measurements is Single-Cell Protein And RNA Co-profiling (SPARC)<sup>50</sup>, which uses poly(A) oligo-dTs conjugated to magnetic beads that hybridize to mRNAs in conjunction with a homogeneous protein extension assay (PEA)<sup>51,52</sup>. SPARC, however, is limited by the 96-plex capability of the targeted proteins of the PEA assay. Cellular Indexing of Transcriptomes and Epitopes by sequencing (CITE-seq) also enables simultaneous measurement of transcriptomes and surface proteins in single cells, using antibodies conjugated to DNA oligonucleotides<sup>53</sup>. However, protein detection is restricted to the targeted antibody panel. Other proximity-labeling-based methods, such as APEX-seq<sup>54–56</sup> and CAP-seq<sup>57</sup>, have also been developed for simultaneous molecular profiling at the RNA and protein levels. These approaches leverage the APEX2 and miniSOG enzymes, which can directly biotinylate nucleic acids as well as proteins. A limitation of these methods is that labeling must be performed ex vivo, requiring either the addition of hydrogen

peroxide (APEX2) or light (miniSOG) to initiate the reaction. Due to the limitations of current methods, new tools for complementary molecular profiling of cell-type-specific transcriptomes and proteomes are needed to gain a more global characterization of cell types in their native states in a complex tissue environment.

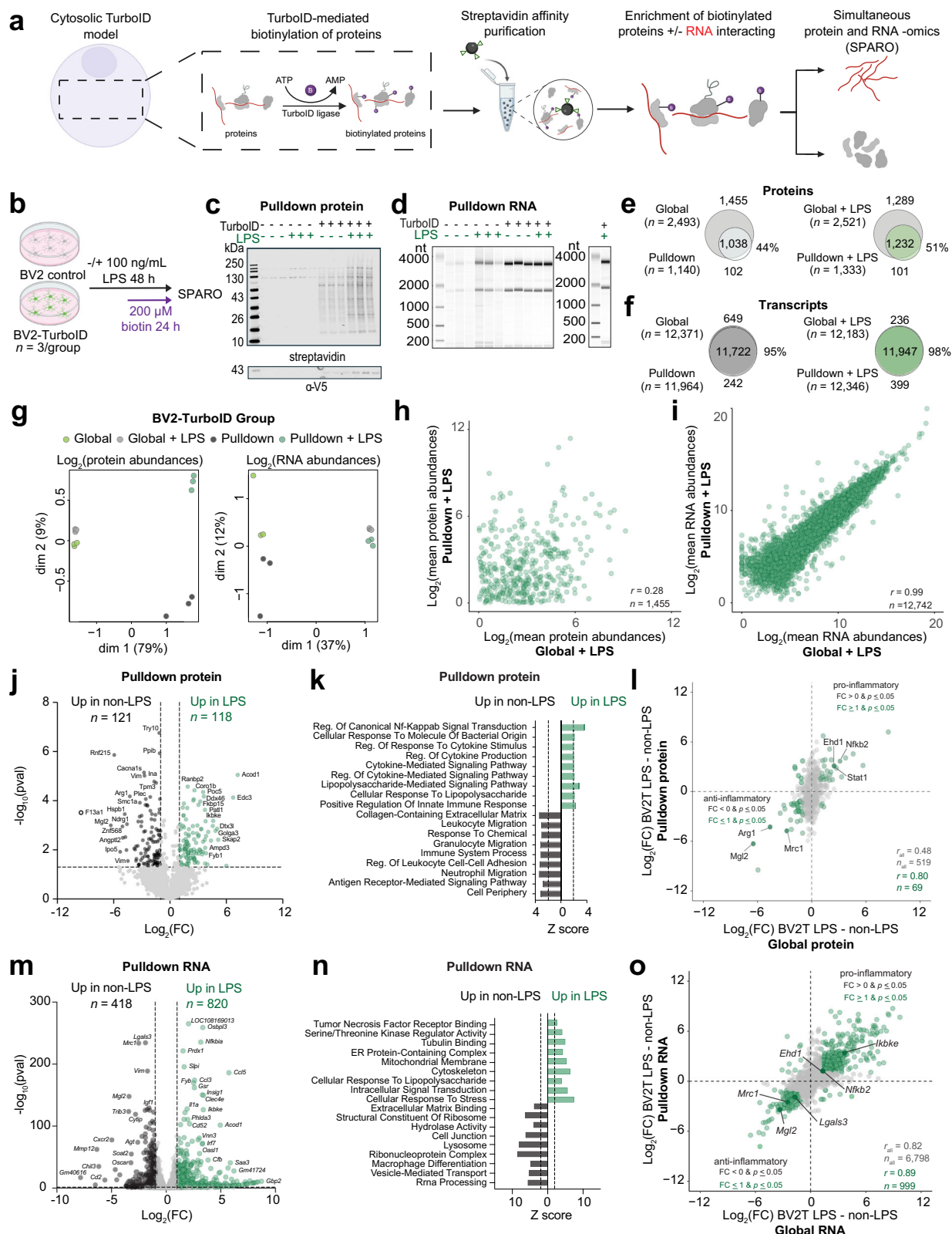
To develop an approach for Simultaneous Protein and RNA Omics (SPARO), we leveraged the biotin ligase, TurboID. Using MS-based proteomics, we previously found that cytosolic TurboID biotinylates many RNA-binding proteins (RBPs) and ribosomal proteins both in vitro and in vivo, including Rpl22 that is used in the RiboTag approach<sup>36,37</sup>. Therefore, we hypothesized that transcripts associated with these proteins could also be purified simultaneously with biotinylated proteins. We developed and validated the SPARO approach to capture cell-type-specific transcriptomes and proteomes using a microglial cell line in vitro, and neurons and astrocytes in vivo. This allowed us to investigate the concordance and discordance between mRNA and protein levels, and to examine how patterns of mRNA-protein discordance are conserved or vary across cell types. SPARO represents an approach to simultaneously quantify RNA and protein profiles with potential for broad applications in vitro and in vivo.

## Results

### Validation of SPARO in BV2-TurboID cells in vitro

To test whether we could leverage TurboID-based cytosolic proteomic labeling to simultaneously capture cellular proteomes and transcriptomes, we first used our previously validated, stably transduced mouse microglia BV2 cell line that expresses V5-TurboID-NES (BV2-TurboID)<sup>37,58</sup>. In this construct, the TurboID biotin ligase is restricted to the extranuclear compartment via a nuclear export sequence (NES) to bias towards biotinylation of the cytosolic proteome (Fig. 1a). By starting with a homogeneous cell line, we could quantitatively assess how the transcriptome and proteome of BV2 cells using the SPARO approach compare with whole-cell transcriptomes and proteomes from global cell lysates. Furthermore, to test whether SPARO could also capture cellular changes in microglia induced by an inflammatory stimulus, we treated BV2 control and BV2-TurboID cells with lipopolysaccharide (LPS) (Fig. 1b). Immunocytochemistry (ICC) confirmed that the TurboID enzyme (via V5) and biotinylation are primarily restricted to the cytosol, under unstimulated and LPS-stimulated conditions (Supplementary Fig. 1a). Following this validation, we lysed cells in a buffer that maintains RNA-protein interactions and then processed for simultaneous transcriptomics (mRNA-seq) and proteomics (label-free quantitative mass spectrometry or LFQ-MS). Immunoblot analyses confirmed the presence of biotinylated proteins (via streptavidin) and the TurboID protein (via V5) in BV2-TurboID cells, which were absent in non-TurboID control cells both before (Supplementary Fig. 1b) and after streptavidin bead affinity purification (henceforth “pulldown”) (Fig. 1c). We detected only endogenously biotinylated proteins in the BV2 control cells (Fig. 1c). After streptavidin pulldown, we confirmed protein enrichment using a silver stain (Supplementary Fig. 1b). Bioanalyzer RNA analysis showed comparably high levels of RNA in global BV2 lysates (Supplementary Fig. 1d). Bioanalyzer RNA analysis displayed RNA of high integrity in the global lysates (mean RIN = 9.9, Supplementary Fig. 1d) and pulldowns (mean RIN = 9.8), but low integrity in the non-TurboID controls (mean RIN = 6.9, Fig. 1d and Supplementary Fig. 2a, b). At the pulldown level, high-quality RNA was enriched only when proteomic biotinylation was present in BV2 microglia ( $p = 0.004$ , unpaired  $t$ -test).

We next performed LFQ-MS and RNA-seq on global BV2-TurboID samples and pulldowns. Raw intensity and processed intensity values are found in Supplementary Data 2–4. Raw RNA-seq and processed counts values are found in Supplementary Data 5 and 6, respectively. We first assessed the overlap of the number of proteins between the TurboID pulldowns and the global proteome. We observed 44% (untreated) and 51% (LPS-treated) overlap of TurboID pulldown



proteomes with the ground truth global BV2 proteomes, fairly consistent with prior observations in BV2-TurboID cells<sup>37</sup>. This is likely a result of the NES tag on TurboID that biases the proteome towards the cytosolic compartment and against the nuclear and mitochondrial compartments (Fig. 1e). Next, to benchmark the fidelity of our TurboID pulldown transcriptome, we compared our RNA-seq datasets between the global BV2 samples and our pulldown populations. We observed

between 95% (untreated) and 98% (LPS-treated) overlap of TurboID pulldown transcriptomes with the ground truth global BV2 transcriptome (Fig. 1f). We next asked whether the pulldown proteomes and transcriptomes could adequately capture changes in BV2 cells upon exposure to LPS (Fig. 1g, left). Multidimensional scaling (MDS) analysis of the transcriptomes identified distinct untreated and LPS-treated transcriptomes regardless of sample type (global vs pulldown)

**Fig. 1 | Simultaneous transcriptomics and proteomics profiling of BV2 microglia under homeostatic and LPS-stimulated conditions.** **a** Schematic of the Simultaneous Protein and RNA -Omics (SPARO) workflow. **b** Experimental design: BV2 control and BV2-TurboID cells ( $n = 3$  biological replicates/genotype and condition) were treated with LPS (100 ng/mL, 48 h) and biotin (200  $\mu$ M, final 24 h). **c** Immunoblot of biotinylated proteins (via streptavidin-680) and TurboID recombinant protein (via V5 tag) following streptavidin pulldown. **d** Bioanalyzer analysis of RNA isolated from biotinylated protein pulldowns. Venn diagrams of proteins identified by LFQ-MS (**e**) and genes identified by RNA-seq (**f**) in global and pulldown datasets. **g** Multidimensional scaling (MDS) plots of  $\log_2$ -transformed normalized LFQ-MS intensities (left) and RNA-seq counts (right). **h** Correlation of protein abundances between global and pulldown proteomes based on normalized mean LFQ-MS intensities ( $n = 1455$  proteins, Pearson's  $r = 0.28$ ,  $p < 0.0001$ ). **i** Correlation of  $\log_2$ -transformed *DESeq2*-normalized mean gene counts between

global and pulldown transcriptomes ( $n = 12,742$  genes, Pearson's  $r = 0.99$ ,  $p < 0.0001$ ). **j** Volcano plot of differentially enriched proteins (DEPs, two-sided  $t$ -test,  $p \leq 0.05$ ,  $\geq 1\text{-log}_2$ -fold change) in LPS-treated versus control BV2-TurboID pulldown samples ( $n = 3/\text{group}$ ). **k** Gene Ontology (GO) analysis of DEPs shown in (**j**). **l** Correlation of shared DEPs between global and pulldown datasets following LPS treatment ( $n = 519$  proteins, Pearson's  $r = 0.48$ ,  $p < 0.0001$ ). Highly enriched DEPs are highlighted in green ( $n = 69$ ,  $r = 0.80$ ,  $p < 0.0001$ ). Canonical pro- and anti-inflammatory microglial markers are labeled. **m** Volcano plot of differentially expressed genes (DEGs, two-sided Wald test,  $p \leq 0.05$ ,  $\geq 1\text{-log}_2$ -fold change) in LPS-treated versus control BV2-TurboID pulldown samples ( $n = 3/\text{group}$ ). **n** GO analysis of DEGs shown in (**m**). **o** Correlation of shared DEGs between global and pulldown transcriptomes after LPS treatment ( $n = 6798$  genes, Pearson's  $r = 0.82$ ,  $p < 0.0001$ ). Highly enriched DEGs are highlighted in green ( $n = 999$ ,  $r = 0.89$ ,  $p < 0.0001$ ). Canonical microglial markers are labeled. Schematics created with BioRender.com.

(Fig. 1g, right), suggesting that the LPS effect was equally captured by the TurboID pulldown and global transcriptomes. We also defined the background noise in the system by comparing mRNAs enriched from non-TurboID BV2 cells with those from BV2-TurboID cells (Supplementary Fig. 2). The “noise” that we found consisted primarily of many non-coding RNAs (e.g., *Gdap10*, *Xist*, and *Malat1*), pseudogenes (e.g., *H2-K2* and *Pisd-ps1*), and splicing factors (e.g., *Luc7l3* and *Rsrp1*) enriched in the BV2 non-TurboID LPS controls ( $p \leq 0.05$ ,  $\geq 1\text{-log}_2$ -fold change, Supplementary Fig. 2c, d, Supplementary Data 7).

To test the fidelity of the pulldowns to their respective global samples, we next performed correlation analysis of our LFQ-MS and RNA-seq data, respectively. At the proteomic level, we observed modest correlation between LPS-treated pulldown and the global LPS-treated BV2-TurboID proteomes (Pearson's  $r = 0.28$ ,  $p < 0.0001$ ,  $n = 1455$  shared proteins, Fig. 1h). Notably, we observed a much higher correlation between the LPS-treated pulldown transcriptomes and global BV2-TurboID transcriptomes (Pearson's  $r = 0.99$ ,  $p < 0.0001$ ,  $n = 12,742$  shared transcripts, Fig. 1i). We also compared the correlation between the untreated global and pulldown proteomes and transcriptomes, respectively (Supplementary Fig. 3a, d). To confirm the cytosolic bias of the V5-TurboID-NES recombinant protein, we differentially compared the pulldown and global BV2-TurboID transcriptomes and proteomes. Differential enrichment analysis of BV2 transcriptomic and proteomic data is found in Supplementary Data 7 and 8, respectively. At the transcriptomic level, we found that mitochondrial-encoded genes (e.g., *COX1*, *CYTb*, and *ND2*) were highly enriched in the global samples ( $p \leq 0.05$ ,  $\geq 1\text{-log}_2$ -fold change, Supplementary Fig. 4a) as compared to the pulldowns, consistent with the inability of cytosolic TurboID to label intra-mitochondrial proteins and their associated transcripts. At the proteomic level, the global proteome was enriched in nuclear, endoplasmic reticulum (ER), and mitochondrial-related terms, consistent with under-sampling of these subcellular compartments by TurboID labeling ( $p \leq 0.05$ ,  $\geq 1\text{-log}_2$ -fold change, Supplementary Fig. 5a).

SPARO also detected proteomic and transcriptomic changes driven by LPS treatment (Fig. 1j, m, Supplementary Figs. 3–5). As expected, GO analysis of differentially enriched proteins (DEPs) between TurboID pulldowns revealed overrepresented ontologies of LPS-treated cells, including Regulation of Lipopolysaccharide-mediated Signaling Pathway, Cellular Response to Molecule of Bacterial Origin, and Regulation of Cytokine Production (Fig. 1k, Supplementary Data 9). The changes that we observed upon LPS treatment measured by TurboID pulldown were still modestly correlated with LPS changes in the global proteome (Pearson's  $r = 0.48$ ,  $n = 519$  shared proteins,  $p \leq 0.05$ , Fig. 1l). However, this correlation was substantially improved when specifically considering proteins that exhibit high differential abundance between LPS-treated and control BV2-TurboID cells (Pearson's  $r = 0.80$ ,  $n = 69$  shared proteins with  $p \leq 0.05$  and  $\leq -1$  or  $\geq 1\text{-log}_2$ -fold change, Fig. 1l). We also found canonical pro-inflammatory

microglial activation markers (e.g., Stat1, Nfkb2, and Ehd1) enriched in the LPS-treated BV2-TurboID pulldowns and anti-inflammatory microglial activation markers (e.g., Mrc1/Cd206, Mgl2 and Arg1) enriched in the control pulldowns ( $p \leq 0.05$ ,  $\geq 1\text{-log}_2$ -fold change, Fig. 1l).

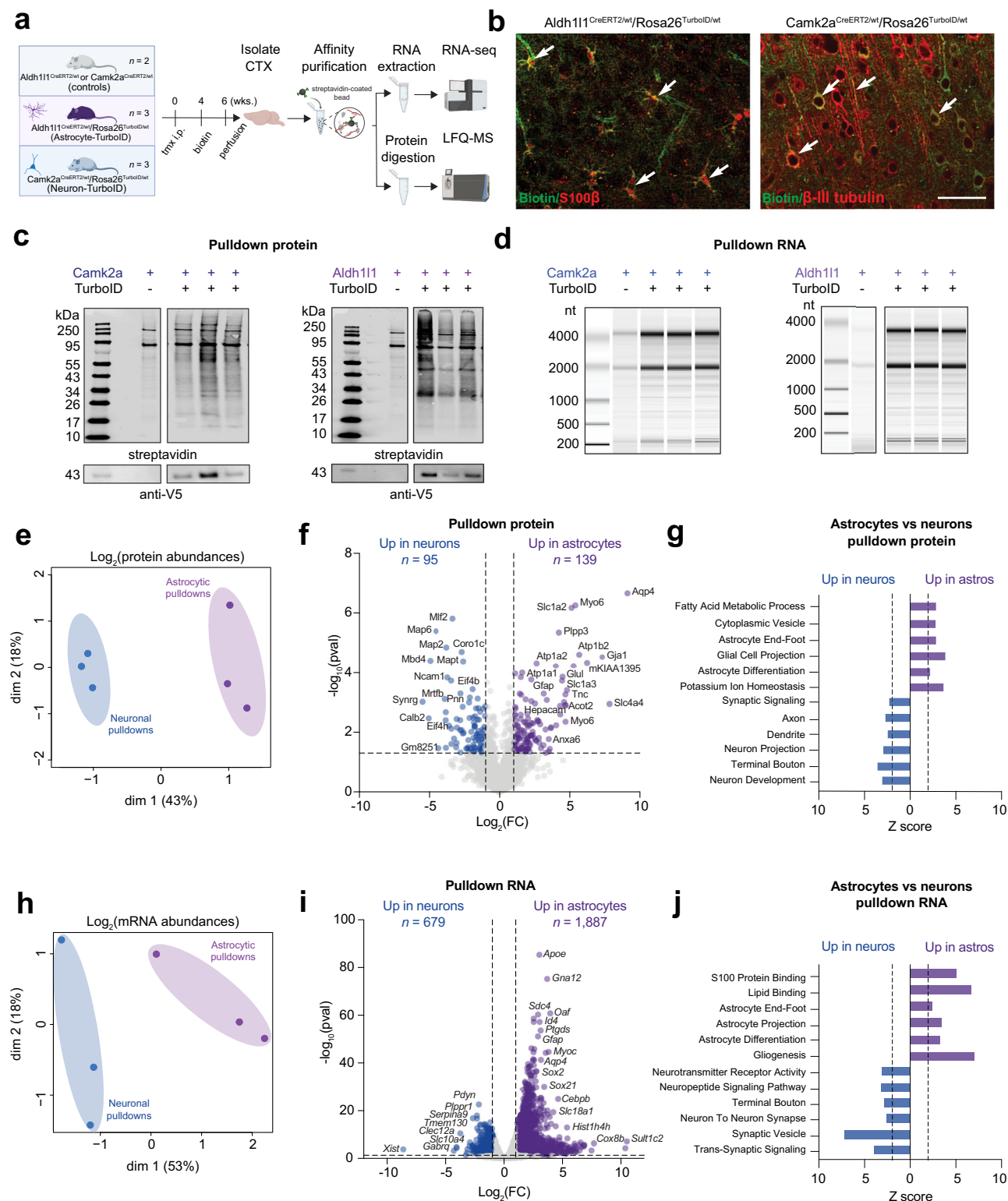
Similar to the proteome, we performed differential enrichment and GO analyses of differentially expressed genes (DEGs) in the LPS-treated and non-LPS-treated pulldown transcriptomes (Supplementary Data 10). We observed the same canonical reactive microglial response enriched in LPS-treated and untreated samples (Fig. 1m–o). Unlike at the protein level, the LPS-driven transcriptomic changes correlated highly between the TurboID RNA pulldown and the LPS-enriched global transcriptome (Pearson's  $r = 0.82$ ,  $n = 6798$  shared transcripts with  $p \leq 0.05$ , Fig. 1o). This correlation was further improved when specifically considering genes that exhibit high differential enrichment between LPS-treated and control BV2-TurboID cells (Pearson's  $r = 0.89$ ,  $n = 999$  shared genes,  $p \leq 0.05$  and  $\leq -1\text{-log}_2$  or  $\geq 1\text{-log}_2$ -fold change, Fig. 1o). Additional transcriptomic and proteomic GO analysis data comparing either LPS-treated versus untreated global or pulldown samples and pulldown versus global samples can be found in Supplementary Data 11–16.

Overall, these results suggest that SPARO can simultaneously capture the cellular transcriptome and proteome and effectively capture the inflammatory effects of LPS in an in vitro mammalian system.

### Validation of SPARO in vivo using cortical astrocytes and neurons

We next wanted to test whether we could capture cell-type-specific transcriptomes and proteomes simultaneously in vivo while retaining the native state of these cells. We crossed Rosa26<sup>TurboID/wt</sup> (TurboID) mice with appropriate inducible Cre-ERT2 mice to drive Cre expression specifically in Camk2a-expressing excitatory neurons or in Aldh1l1-expressing astrocytes. We chose the Aldh1l1-Cre/ERT2 mice and Camk2a-Cre/ERT2 mouse lines for their extensive validation with regard to the specificity of Cre recombination<sup>42,59</sup>. Our breeding scheme resulted in the generation of Aldh1l1<sup>CreERT2/wt</sup>/Rosa26<sup>TurboID/wt</sup> (astrocyte-TurboID) and Camk2a<sup>CreERT2/wt</sup>/Rosa26<sup>TurboID/wt</sup> (neuron-TurboID) mice. We used Aldh1l1<sup>CreERT2/wt</sup> or Camk2a<sup>CreERT2/wt</sup> as Cre-only controls. We induced Cre-mediated recombination by intraperitoneal tamoxifen at 7 weeks of age, followed by a 3-week gap, and then biotin water supplementation for 2 weeks (Fig. 2a). Based on our prior study, V5-TurboID-NES presence and biotin treatment do not have adverse effects on the mice and do not impact homeostatic cellular functions<sup>31,33,36</sup>. Following biotin treatment (age 2.5–3 months), we harvested and lysed cortical tissue in the same buffer as the in vitro studies to maintain RNA-protein interactions and then processed for simultaneous transcriptomics and proteomics using SPARO (Fig. 2a). Immunofluorescence imaging of the cortex showed robust biotinylation in TurboID brains compared to controls (Fig. 2b and Supplementary Fig. 6a). Moreover, the biotinylation signal co-localized with





the astrocytic marker, S100 $\beta$  (in astrocyte-TurboID mice) and neuronal marker,  $\beta$ -III-tubulin (in neuron-TurboID mice) confirming cell targeting (Fig. 2b). Importantly, our prior work validated that the biotinylation signal present in astrocyte- and neuron-TurboID animals is cell-type-specific via the absence of signal in other brain cell types, and without reactive gliosis<sup>36</sup>. Immunoblot analysis of cortical lysates at the global level (Supplementary Fig. 6b, c) and after streptavidin affinity purification displayed biotinylated proteins from astrocyte-TurboID and neuron-TurboID pulldowns not seen in non-TurboID controls

(Fig. 2c). The presence of the TurboID recombinant protein via V5 presence was also only seen in astrocyte-TurboID and neuron-TurboID pulldowns and not non-TurboID controls (Fig. 2c), providing biochemical verification of the presence of TurboID in these cells.

Following streptavidin pulldown, we determined protein enrichment using a silver stain (Supplementary Fig. 6d, e) and RNA integrity using a Bioanalyzer RNA quality control (QC) assay (Fig. 2d, Supplementary Fig. 6f). In the TurboID-expressing samples, we consistently found RNA of higher integrity (mean RIN = 8.0) when compared to the

**Fig. 2 | TurboID-based dual transcriptomic and proteomic analysis of cortical Aldh1l1-expressing astrocytes and Camk2a-expressing neurons in vivo.**

**a** Experimental overview of SPARO in vivo using Aldh1l1<sup>CreERT2/WT</sup> or Camk2a<sup>CreERT2/WT</sup> (control,  $n = 2$  M mice), Aldh1l1<sup>CreERT2/WT</sup>/Rosa26<sup>TurboID/WT</sup> (astrocyte-TurboID,  $n = 3$  M mice) and Camk2a<sup>CreERT2/WT</sup>/Rosa26<sup>TurboID/WT</sup> (neuron-TurboID,  $n = 3$  mice (2 F and 1 M)) mouse models. Following tamoxifen (75 mg/kg/day via intraperitoneal (i.p.) injections, 5 days and 3 weeks of Cre recombination, mice received biotin-containing water (37.5 mg/L) for two weeks. Cortical tissue was isolated for streptavidin affinity purification, followed by LFQ-MS and RNA-sequencing. **b** Representative sagittal brain immunofluorescence images of control, astrocyte-TurboID, and neuron-TurboID mice ( $n = 3$  mice/experimental group). Streptavidin-488 labeling (green) for biotinylation is shown alongside the astrocytic marker

S100 $\beta$  (red) or the neuronal marker  $\beta$ -III-tubulin (red). Nuclei are labeled with DAPI (blue). Scalebar = 50  $\mu$ m. **c** Immunoblot of biotinylated proteins (via streptavidin-680) and TurboID recombinant protein (via V5 tag) following streptavidin pull-down. **d** Bioanalyzer analysis of RNA isolated from biotinylated protein pull-downs. **e** MDS plot of log<sub>2</sub>-transformed normalized LFQ-MS intensities. **f** Volcano plot of DEPs (two-sided  $t$ -test,  $p \leq 0.05$  and  $\geq 1$ -log<sub>2</sub>-fold change) between astrocyte-TurboID and neuron-TurboID pull-downs ( $n = 3$ /group). **g** Proteome GO analysis of DEPs visualized in (f). **h** MDS plot of log<sub>2</sub>-transformed normalized RNA-seq count data pull-downs. **i** Volcano plot of DEGs (two-sided Wald test,  $p \leq 0.05$  and  $\geq 1$ -log<sub>2</sub>-fold change) between astrocyte-TurboID and neuron-TurboID pull-downs ( $n = 3$ /group). **j** GO analysis of DEGs visualized in (i). Schematic created with BioRender.com.

non-TurboID samples (mean RIN = 5.2,  $p = 0.04$ , unpaired  $t$ -test). We next prepared samples for LFQ-MS. Raw intensity and processed intensity values are found in Supplementary Data 17–19, respectively. Notably, MDS analysis of the proteomes showed separation between neuronal and astrocytic proteomes (Fig. 2e). We also confirmed the cell-type specificity of the proteomes, including enrichment of canonical astrocytic proteins (e.g., Hspcam, Glul, Aqp4, Plpp3,  $p \leq 0.05$ ,  $\geq 1$ -log<sub>2</sub>-fold change) and GO terms (e.g., astrocyte end-foot, glial cell projection, and astrocyte differentiation) in astrocytic pull-downs and neuronal proteins (e.g., Map2, Ncam1, Mapt,  $p \leq 0.05$ ,  $\geq 1$ -log<sub>2</sub>-fold change) and GO terms (e.g., synaptic signaling, dendrite, and neuron projection) in neuronal pull-downs (Fig. 2f, g, Supplementary Data 20 and 21). These data are consistent with our prior work using astrocyte-TurboID and neuron-TurboID systems<sup>36</sup>. It is important to highlight that the lysis buffer used in the current studies was a homogenization buffer intended to maintain RNA-protein interactions, as opposed to the 8 M urea lysis buffer in prior work<sup>36</sup>.

Like the proteome, we conducted MDS, differential enrichment, and GO analysis of the neuronal and astrocytic transcriptomes. Raw RNA-seq and processed counts values are found in Supplementary Data 22 and 23, respectively. Like the proteome, MDS analysis identified distinct neuronal and astrocytic transcriptomes (Fig. 2h), each of which was enriched for canonical cell-type-specific markers (Fig. 2i, j, Supplementary Data 24 and 25), suggesting that the pull-downs are cell-type-enriched. To further validate the cell-type specificity of the pull-down transcriptomes, we compared the relative gene expression of top astrocytic and neuronal genes from Zhang et al.<sup>60</sup> between the global cortex and the cell-type-enriched pull-downs (Supplementary Fig. 7). The cell-type-enriched gene lists can be found in Supplementary Data 26. The astrocyte-TurboID pull-down transcriptomes exhibited a high expression of astrocytic genes (e.g., *Tnc*, *Sox9*, and *Aqp4*) and a low expression of neuronal genes (e.g., *Tmem130*, *Slc10a4*, and *Npy*) (Supplementary Fig. 7a), and the converse was true for the neuron-TurboID pull-down transcriptomes (Supplementary Fig. 7b).

We included a 5-day repeated LPS paradigm in our original astrocyte SPARO study to examine effects of LPS-induced neuroinflammation on astrocyte profiles (Supplementary Fig. 8). We observed modest effects of LPS at both the proteome ( $p \leq 0.05$ ,  $\geq 1$ -log<sub>2</sub>-fold change,  $n = 11$  up with LPS,  $n = 35$  up in non-LPS controls, Supplementary Data 18) and transcriptome levels ( $p \leq 0.05$ ,  $\geq 1$ -log<sub>2</sub>-fold change,  $n = 141$  up with LPS,  $n = 41$  up in non-LPS controls, Supplementary Data 21), suggesting that the LPS effect may have been partially washed out during the one-week interval between the last LPS dose and euthanasia (Supplementary Fig. 8b, c). Despite this, we observed modest yet meaningful effects of LPS on the astrocyte pull-down transcriptome using the SPARO approach, including upregulation of canonical reactive genes (e.g., *Lcn2* and *Serpina3g*, Supplementary Fig. 8b). We also performed GO analysis of DEGs and found synapse- and translation-related terms enriched with LPS treatment (Supplementary Fig. 8d, Supplementary Data 27). In contrast, LPS-induced changes were minimal at the level of the astrocyte

proteome (Supplementary Fig. 8c), suggesting a discordance between LPS-induced effects on astrocytes at the transcriptomic and proteomic levels. We did not find any GO terms enriched from DEPs, due to the small number of DEPs within each sample group.

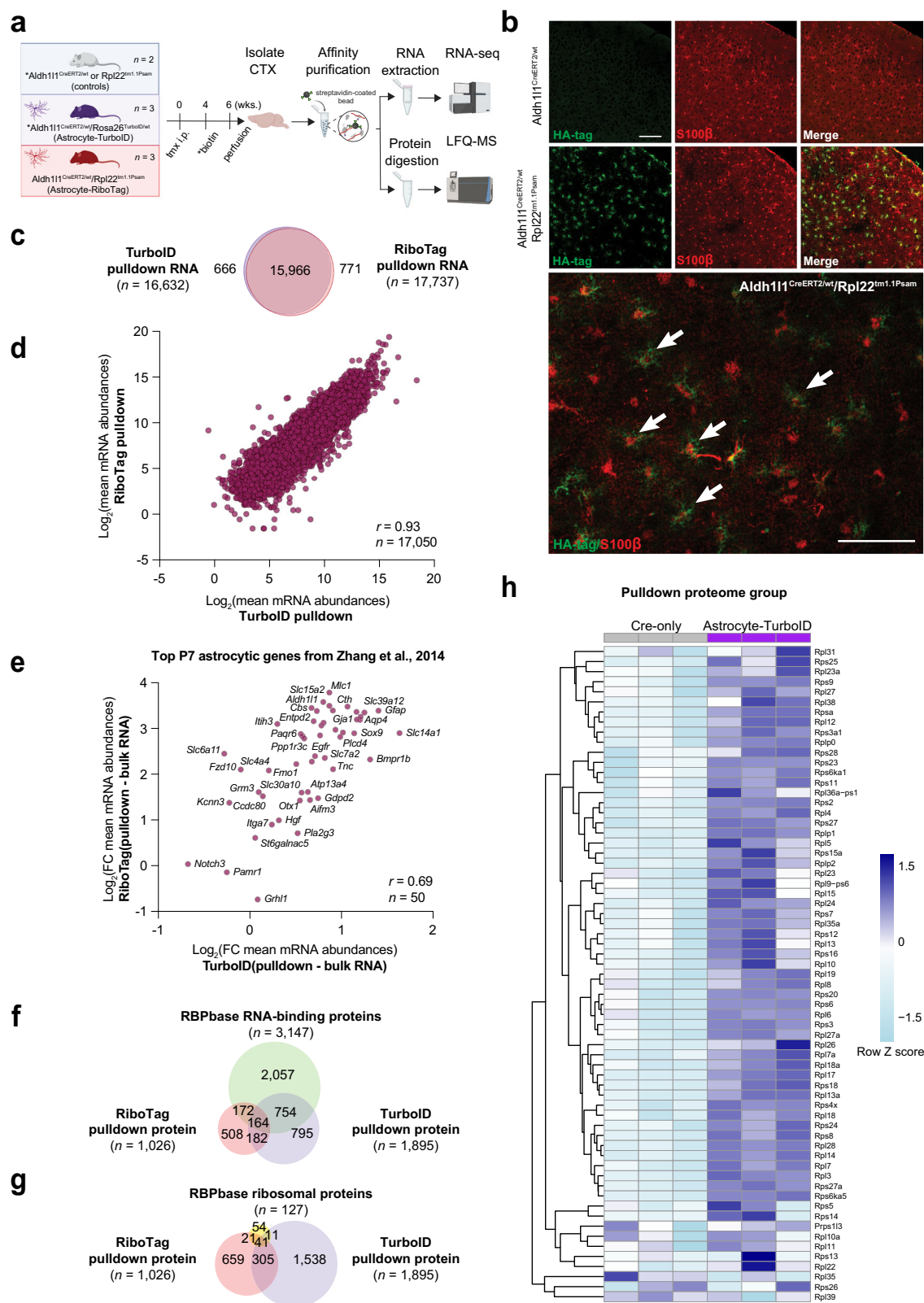
These results demonstrate the application of the SPARO approach to obtain biologically meaningful proteomes and transcriptomes of astrocytes and neurons while retaining their native states in the adult mouse cortex.

**The SPARO- and RiboTag-enriched astrocytic transcriptomes are similar**

We next benchmarked SPARO against RiboTag, a gold-standard in vivo cell-type-specific transcriptomic profiling method<sup>40,41,43</sup>. RiboTag mice express a hemagglutinin tag (HA) modified allele of the *Rpl22* gene (*Rpl22-HA*), a major component of the polyribosome complex<sup>40,41</sup>. In the presence of Cre recombinase, HA-tagged polyribosomes can be isolated from target cell types using an anti-HA antibody and immunoprecipitation methods. The mRNAs interacting with the ribosomes during translation can then be extracted and quantified to obtain a snapshot of actively translated transcripts (translatome) in specific cell types. To assess the similarity between the TurboID-enriched and RiboTag-enriched transcriptomes, we compared both methods using cortical Aldh1l1-expressing astrocytes in vivo (Fig. 3a). We bred Aldh1l1<sup>CreERT2/WT</sup> mice with *Rpl22*<sup>tm1.LPpsam</sup> or TurboID mice to generate astrocyte-RiboTag and astrocyte-TurboID mice, respectively. We administered tamoxifen to astrocyte-TurboID, astrocyte-RiboTag, and Aldh1l1<sup>CreERT2/WT</sup> or *Rpl22*<sup>tm1.LPpsam</sup> (control) mice at 7 weeks of age, followed by a 3-week gap, then performed cortical tissue lysis, pull-down of biotinylated or HA-tagged proteins, LFQ-MS, and RNA-seq (Fig. 3a).

Immunofluorescent imaging of the cortex showed the presence of HA recombinant protein in astrocyte-RiboTag brains, which co-localized with the astrocytic marker, S100 $\beta$  (Fig. 3b). TapeStation RNA QC analysis from bulk cortical lysates from control and astrocyte-RiboTag samples showed RNA of high integrity (mean RIN = 8.7) across all samples (Supplementary Fig. 9a). Astrocyte-RiboTag pull-downs also showed high-quality RNA (mean RIN = 8.7), whereas RiboTag-only controls did not produce reportable RIN values (Supplementary Fig. 9b). We next prepared samples for RNA-seq. Raw RNA-seq and processed counts values are found in Supplementary Data 28 and 29, respectively. Importantly, transcript overlap between the astrocyte-TurboID and astrocyte-RiboTag pull-downs was 95% (Fig. 3c). The cell-type specificity of RiboTag and SPARO transcriptomes (Supplementary Fig. 9c), as well as the correlation across datasets (Pearson's  $r = 0.93$ ,  $n = 17050$  shared transcripts, Fig. 3d), were both robust. Next, we examined whether highly cell-type-specific astrocytic genes from Zhang et al.<sup>60</sup> were correlated across the TurboID- and RiboTag-enriched transcriptomes. We found that the mRNA abundances were highly correlated (Pearson's  $r = 0.69$ ,  $n = 50$  genes, Fig. 3e).

We next performed differential enrichment analysis and GO analysis between the astrocyte-TurboID and astrocyte-RiboTag pull-down transcriptomes (Supplementary Fig. 9d–f, Supplementary



Data 30 and 31) and identified several differences in signatures. We hypothesized that these differences arise from the fact that RiboTag enriches all RNAs bound to ribosomal complexes that contain the ribosomal protein, Rpl22, whereas the SPARO approach captures mRNAs interacting with many biotinylated proteins that interact with RNA in astrocytes. To test this, we performed LFQ-MS on the astrocyte-RiboTag pulldowns and compared the results to the astrocyte-TurboID

pulldown proteome. The astrocyte-RiboTag raw intensity and processed intensity values are found in Supplementary Data 32 and 33, respectively. For this comparison, we first identified proteins in the astrocyte-TurboID or astrocyte-RiboTag enriched over Cre-only or RiboTag-only controls, respectively ( $\geq 1$ -log<sub>2</sub>-fold change) and then determined the number of RNA-binding proteins (RBPs and ribosomal proteins) enriched in each dataset (Supplementary Data 34). We found



**Fig. 3 | The TurboID-enriched transcriptome is comparable to the RiboTag-enriched transcriptome of cortical astrocytes.** **a** Experimental overview to compare the TurboID-enriched astrocyte transcriptome to the RiboTag-enriched astrocyte transcriptome. *Aldh1l1*<sup>CreERT2/wt</sup> or *Rpl22*<sup>tm1.1Psa</sup> (control,  $n = 3$  mice (1 M and 2 F)), *Aldh1l1*<sup>CreERT2/wt</sup>/*Rosa26*<sup>TurboID/wt</sup> (astrocyte-TurboID,  $n = 3$  M mice), *Aldh1l1*<sup>CreERT2/wt</sup>/*Rpl22*<sup>tm1.1Psa</sup> (astrocyte-RiboTag,  $n = 3$  F mice) mouse models (aged 2.5–3.5 months). Mice received tamoxifen (75 mg/kg/day via i.p. injections, 5 days). After 3 weeks, astrocyte-TurboID mice received biotin-containing water (37.5 mg/L) for two weeks. Cortical tissue was isolated and lysed for affinity purification of either biotinylated ribosomal and RNA-binding proteins from astrocyte-TurboID mice, or HA-tagged *Rpl22*-containing polyribosomes, followed by RNA extraction and RNA-sequencing. **b** Representative sagittal brain immunofluorescence images of *Aldh1l1*<sup>CreERT2/wt</sup> control and astrocyte-RiboTag mice ( $n = 2$ –3 mice/experimental group). HA-tag (green) is shown alongside the astrocytic marker *S100 $\beta$*  (red). Scalebar = 50  $\mu$ m. **c** Venn diagram of genes identified by RNA-seq. **d** Scatter plot comparing the expression of 17,050 transcripts present in the TurboID (x-axis) or

RiboTag (y-axis) pulldown transcriptomes (Pearson's  $r = 0.93$ ,  $p < 0.0001$ ). Genes that are not present in both lists were removed from the analysis. **e** Scatter plot visualization comparing the expression of the top 50 postnatal day 7 (P7) astrocyte-specific transcripts enriched in the TurboID (x-axis) or RiboTag (y-axis) pulldown transcriptomes after global RNA subtraction (Pearson's  $r = 0.69$ ,  $p < 0.0001$ ). The astrocyte-specific gene lists were acquired from Zhang et al.<sup>60</sup>. **f** Venn diagram of EMBL RBPbase-identified RNA-binding proteins identified by DDA LFQ-MS present in the astrocyte-TurboID pulldowns and DIA LFQ-MS present in the astrocyte-RiboTag IP proteomes. Proteins were included based on a  $\geq 1$ -log<sub>2</sub>-fold change over the non-TurboID and RiboTag-only control pulldowns. **g** Venn diagram of ribosomal proteins identified by LFQ-MS filtered from (f). **h** Heatmap of protein relative abundances (via Row Z score, purple = relative high abundance or enrichment and blue = relative low abundance or depletion) of ribosomal proteins ( $n = 65$ ) present in the *Aldh1l1*<sup>CreERT2/wt</sup> (Cre only) and astrocyte-TurboID pulldowns. Schematic created with BioRender.com.

that the TurboID approach captured many RBPs ( $n = 918$  total proteins with 754 unique to TurboID) when compared to the RiboTag approach ( $n = 336$  total proteins with 172 unique to RiboTag) and 164 RBPs shared (Fig. 3f). Next, we filtered the RBP list for only ribosomal proteins. We found that both approaches captured many ribosomal proteins, with slightly more ribosomal proteins detected in RiboTag ( $n = 11$  unique to TurboID and  $n = 21$  unique to RiboTag), with 41 ribosomal common proteins (Fig. 3g). Abundance analysis of ribosomal proteins in the astrocyte-TurboID pulldowns relative to the Cre-only controls (via row Z score) confirmed strong enrichment (Fig. 3h).

We next examined transcriptomic features, including mean transcript length, GC content, and exon number. DEGs ( $p \leq 0.05$ ,  $\geq 1$ -log<sub>2</sub>-fold change, Supplementary Fig. 9d, Supplementary Data 35) between astrocyte-RiboTag and astrocyte-TurboID pulldowns showed significant differences across all three features. Mean transcript length ( $p < 0.0001$ , Mann-Whitney test), and exon number ( $p = 0.029$ , Mann-Whitney test) were higher in the TurboID pulldowns while percent GC content ( $p = 0.002$ , Mann-Whitney test) was higher in the RiboTag pulldowns (Supplementary Fig. 9g). These differences may reflect the broader nature of TurboID labeling, which captures many RBPs not limited to those associated with active translation.

Because TurboID biotinylates many RNA-interacting proteins, we tested whether SPARO could measure small non-coding RNAs (ncRNAs). Small RNA extraction and sequencing of astrocyte-TurboID pulldowns and bulk cortex detected over 400 miRNAs in both sample types (Supplementary Fig. 10a, b, Supplementary Data 36 and 37). MDS analysis revealed clear separation of bulk cortex and astrocyte pulldown miRNAomes (Supplementary Fig. 10c). Differential enrichment analysis identified 126 miRNAs enriched in astrocyte pulldowns and 102 miRNAs enriched in bulk cortex ( $p \leq 0.05$ ,  $\geq 1$ -log<sub>2</sub>-fold change, Supplementary Fig. 10d, Supplementary Data 38). The most enriched astrocyte pulldown miRNA was *mmu-miR-760*.

Together, these findings could explain why the astrocyte-RiboTag and astrocyte-TurboID pulldown transcriptomes show overall similarities in transcript detection and abundance yet diverge in specific signatures.

### Correlation of the SPARO transcriptomes and proteomes of cortical astrocytes and neurons in vivo

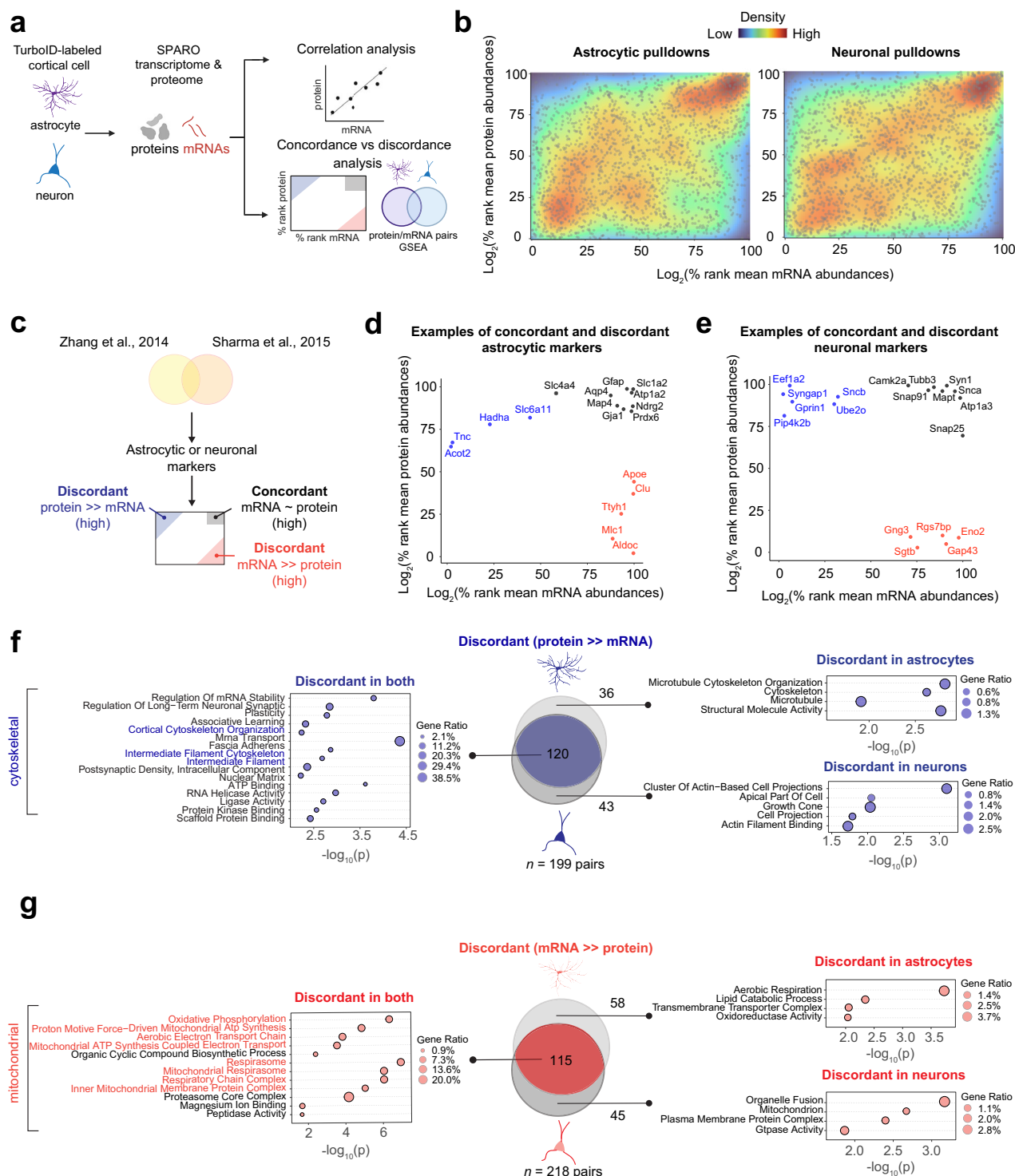
We next evaluated the concordance between the in vivo proteomes and transcriptomes of astrocytes and neurons obtained using SPARO (Fig. 4a). The raw correlation between 1934 mRNA and protein pairs was modest in both astrocytes (Pearson's  $r = 0.29$ , Supplementary Fig. 11a, Supplementary Data 39) and neurons (Pearson's  $r = 0.27$ , Supplementary Fig. 11b, Supplementary Data 40). To determine if these generally low correlations were simply a product of the complexity of an in vivo brain environment, we also performed similar transcriptome and proteome correlation analysis of the in vitro BV2-

TurboID cells and found similar results (Pearson's  $r = 0.24$ , Supplementary Fig. 8c, Supplementary Data 41). We did observe the highest concordance between global BV2-TurboID cells (Pearson's  $r = 0.66$ , Supplementary Fig. 11c, Supplementary Data 42) and bulk cortical transcriptomes and proteomes (Pearson's  $r = 0.47$ , Supplementary Fig. 11f, Supplementary Data 43), which lack any cell-type specificity. To benchmark our in vivo SPARO correlations to published datasets, we used RNA data from the Zhang et al.<sup>60</sup> transcriptomic study of mouse brain cell types and Sharma et al.<sup>61</sup> proteome of acutely isolated mouse astrocytes and neurons. Like our findings in SPARO, the correlation between mRNA and protein abundances was modest for all pairs (Pearson's  $r = 0.36$  for astrocytes and  $r = 0.29$  in neurons) as well as after limiting the analysis to mRNA-protein pairs found in SPARO ( $r = 0.41$  in astrocytes and  $r = 0.21$  in neurons, Supplementary Fig. 11e, Supplementary Data 44 and 45).

Quantifying concordance is further complicated by innate technical differences between LFQ-MS and RNA-sequencing. For example, the dynamic range and sensitivity of both approaches differ vastly. Therefore, to better normalize against these biases, we calculated percentile ranked abundances for proteins and transcripts in each dataset (1934 gene and protein pairs that overlap) and classified these mRNA-protein pairs into distinct groups based on level of mRNA-protein rank discordance (Fig. 4b and Supplementary Fig. 11g). In both astrocytes and neurons, and in the bulk cortex, we observed that most mRNA-protein pairs fell into either concordantly low or high groups (Log<sub>2</sub> % rank values between 75 and 100 for both the x and y-axis). Together, these findings verify modest levels of concordance between protein and mRNA abundances<sup>61–63</sup>, although mRNAs that tend to be highly expressed in neurons and astrocytes, and in the cortex in general, also tend to be highly abundant at the proteomic level.

We next classified mRNA and protein pairs from the astrocyte and neuronal pulldowns into either concordant (mRNA ~ protein,  $>0.75$  in black) or discordant (protein  $\gg$  mRNA,  $>-0.50$  in blue, or mRNA  $\ll$  protein levels,  $>0.50$  in red) groups based on a rank differential of abundance values (Supplementary Fig. 11h, Supplementary Data 40 and 41). We wondered whether there were examples of canonical cell-type-specific markers that fall into either highly concordant or discordant quadrants. To designate astrocytic and neuronal-enriched canonical markers, we used a unionized marker list<sup>64–66</sup> generated from the Zhang et al.<sup>60</sup> transcriptomic data and the Sharma et al.<sup>61</sup> proteomic data from astrocytes and neurons (Fig. 4c and Supplementary Fig. 11i, j, Supplementary Data 46). A subset of example markers belonging to each category is shown in Fig. 4d. We repeated the same analysis for the neuronal pulldown proteomes and transcriptomes (Fig. 4e). These findings nominate cell-type-specific astrocytic and neuronal markers that may be preferentially suitable for either transcriptomic or protein-based studies, as well as those that are highly abundant at both mRNA and protein levels.





**Fig. 4 | Correlation analysis of TurboID-based transcriptomes and proteomes of *Aldhl1*-expressing astrocytes and *Camk2a*-expressing neurons in vivo.**

**a** Schematic of the analysis approach to compare SPARO of cortical astrocytes and neurons ( $n = 3$  animals/group). **b** Percent rank transformation and density plot visualization of 1934 gene and protein pairs between the astrocyte-TurboID (left) and neuron-TurboID (right) pulldown transcriptome (x-axis, via normalized,  $\text{log}_2$ -transformed and percent ranked RNA-seq mean count values) and proteome (y-axis, via normalized,  $\text{log}_2$ -transformed and percent ranked LFQ-MS mean intensity values). **c** Schematic of analysis approach to nominate astrocytic and neuronal markers into highly concordant (mRNA ~ protein in black) or highly discordant (protein >> mRNA in blue or mRNA >> protein in red) groups. The top astrocyte-

specific and neuronal-specific markers were acquired from the unionization of the Zhang et al.<sup>60</sup> and Sharma et al.<sup>61</sup> datasets. Scatter plot of examples of astrocytic markers (**d**) and neuronal markers (**e**) that are concordant where mRNA and protein abundances are high and similar. Venn diagrams of mRNA and protein pair quantities present in the discordant groups: (**f**) protein >> mRNA and (**g**) mRNA >> protein, between astrocyte-TurboID and neuron-TurboID paired pulldown transcriptomes and proteomes. Gene-set enrichment analysis of GO terms was performed using a one-tailed Fisher's exact test for enrichment. Enriched GO terms are visualized as  $-\log_{10}(p\text{-value})$ , with bubble size corresponding to the gene ratio (number of hits/total genes). Only terms with a Z score >1.96 were included. Schematics created with BioRender.com.

The observed mRNA-protein discordance can arise from many biological phenomena, which may also vary across cell types. To determine whether discordant proteins and mRNAs were consistent across astrocytes and neurons, we assessed the level of overlap in mRNA-protein pairs across astrocytes and neurons in 3 quadrants (high mRNA and high protein, high mRNA and low protein, low mRNA and high protein, Supplementary Data 47). Across all 3 groups, most concordant mRNA-protein pairs (high RNA, high protein) were present in both cell types (Supplementary Fig. 11k). Many discordant mRNA-protein pairs were also shared across cell types (Fig. 4f, g). Interestingly, we observed a pattern in which the types/categories of discordant genes were generally conserved across cell types, with specific subtypes of genes that are cell-type-specific. For example, within the discordant group where protein exceeded mRNA (protein > mRNA) abundance, we found overrepresented ontologies of cytoskeletal pathways in both astrocytes and neurons. However, within astrocytes, only terms related to microtubules were enriched, as opposed to actin machinery in neurons (Fig. 4f). For the discordant group of genes in which mRNA exceeded protein (mRNA > protein) abundance, we found a consistent signature in both neurons and astrocytes of ontologies related to mitochondria. However, within astrocytes, this signature comprised genes involved in aerobic respiration, whereas neurons were biased towards genes involved with mitochondrial fusion and GTPase activity (Fig. 4g). This pattern most likely reflects the metabolic differences that exist between these two cell types. After performing GO analyses on the mRNA and protein pairs in each concordant and discordant quadrant, we also looked at the overlap of the GO terms across the two cell types. The cell-type-specific GO enrichment patterns of mRNA and protein pairs in the concordant and discordant groups generally matched our results from GO analyses at the level of mRNA-protein pairs (Fig. 4f, g, Supplementary Fig. 12, Supplementary Data 48).

To validate our discordant mRNA-protein findings, we performed concordance/discordance analyses on the BV2-TurboID globals and SPARO pulldowns (Supplementary Fig. 13a, Supplementary Data 49 and 50), and the Zhang et al.<sup>60</sup> Sharma et al.<sup>61</sup> published astrocytic and neuronal datasets (Supplementary Fig. 13b, Supplementary Data 51 and 52). Across the independent experiments and analyses of published work—the BV2-TurboID samples (global and pulldown), Zhang vs Sharma astrocyte and neuron global datasets—we found cytoskeletal-related GO terms enriched in the discordant protein > mRNA blue group (Supplementary Fig. 13c, d). We also found some mitochondrial-related GO terms enriched in the discordant mRNA > protein groups from the BV2-TurboID SPARO pulldowns and the Zhang vs Sharma astrocyte and neuron global datasets (Supplementary Fig. 13c, d). These findings show that the SPARO discordant mRNA-protein phenotypes are likely not due to the labeling nature of TurboID and may be due to biological processes in these cells.

After evaluating the correlation between the paired proteomes and transcriptomes of each TurboID model, we aimed to uncover mechanisms mediating concordance and discordance between mRNA and protein abundances in the bulk cortex, astrocytes, and neurons. We found no relationship between the level of mRNA-protein discordance and the average RNA features: % GC content, number of exons, and transcript length or protein half-life<sup>67</sup> (Supplementary Fig. 14, Supplementary Data 53 and 54), suggesting that these mRNA characteristics and protein synthesis dynamics are unlikely on their own to explain the observed discordance.

Together, these findings emphasize an important observation about discordant mRNAs and proteins across these two cell types. In general, the broad classes of mRNAs and proteins that are highly discordant in abundance are well conserved between neurons and astrocytes. It is only in the nuanced subtypes of these gene classes that each cell type exhibits a uniquely discordant repertoire of proteins and transcripts.

## Removal of NES from TurboID does not alter the correlation between mRNA-protein abundances in vitro

After identifying concordant and discordant phenotypes of cortical astrocytes and neurons using SPARO in vivo, we wanted to further confirm that the discordant mRNA-protein phenotypes are not due to the labeling nature of cytosolic TurboID. To investigate this, we infected HEK293 cells with a TurboID lentivirus with or without the NES and performed SPARO (Supplementary Fig. 15a, b). IF and immunoblot analysis confirmed the presence of biotinylated proteins (via streptavidin) in HEK293 TurboID-NES and noNES cells and only endogenously biotinylated proteins in the non-TurboID control cells. (Supplementary Fig. 15c). TapeStation RNA QC analysis was otherwise comparable between constructs (Supplementary Fig. 15d).

We next performed DIA LFQ-MS and RNA-seq on global HEK293 samples and pulldowns (Supplementary Fig. 16–18). Raw intensity and processed intensity values are found in Supplementary Data 55 and 56, respectively. Raw RNA-seq and processed counts values are found in Supplementary Data 57 and 58, respectively. We observed 92% overlap between the TurboID-NES and TurboID-noNES pulldown transcriptomes (Supplementary Fig. 15e, top), as well as 85% overlap of detected peptides between the TurboID-NES and TurboID-noNES pulldowns (Supplementary Fig. 15e, bottom). However, MDS analysis of the transcriptomes and proteomes identified distinct TurboID-NES and TurboID-noNES pulldown groups (Supplementary Fig. 16d).

To assess differences between the TurboID-NES and noNES pulldowns, we also examined the correlation between the transcriptomes and proteomes, respectively. Surprisingly, we found a very high correlation between the mRNA and protein abundances between the respective TurboID pulldowns (Supplementary Fig. 15f, Pearson's  $r = 0.99$ ,  $n = 17,998$  shared genes and Supplementary Fig. 15g, Pearson's  $r = 0.93$ ,  $n = 7333$  shared proteins), suggesting that TurboID-enriched pulldowns with and without the NES are highly similar. To define the labeling bias of the TurboID-NES and TurboID-noNES enzymes, we differentially compared the pulldown transcriptomes and proteomes. Differential enrichment analysis of HEK293 transcriptomic and proteomic data is found in Supplementary Data 59 and 60, respectively. Proteomic GO analyses can be found in Supplementary Data 61–63, and transcriptomic GO analyses can be found in Supplementary Data 64–66. At the transcriptomic level, we found few DEGs in the TurboID-noNES pulldown group and nearly triple the amount in the TurboID-NES pulldown group ( $n = 81$  Up in the TurboID-noNES and  $n = 261$  Up in TurboID-NES,  $p \leq 0.05$ ,  $\geq 1\text{-log}_2\text{-fold change}$ , Supplementary Fig. 15h). We found nuclear and endoplasmic reticulum (ER)-associated genes (e.g., *HSPA5*, *SDF2L1*, and *H2BC21*) to be the most highly enriched in the TurboID-noNES group ( $p \leq 0.05$ ,  $\geq 1\text{-log}_2\text{-fold change}$ , Supplementary Fig. 15h). GO analyses also confirmed nuclear, ER and mitochondrial-associated cellular processes to be enriched in the TurboID-noNES pulldowns (Supplementary Fig. 16g). Notably, some nuclear-associated genes (e.g., *ZNF75A*, *HPS6* and *EID2*) also appeared in the TurboID-NES pulldowns ( $p \leq 0.05$ ,  $\geq 1\text{-log}_2\text{-fold change}$ , Supplementary Fig. 15h). At the proteomic level, we found many DEPs in the TurboID-noNES pulldowns and very few in TurboID-NES pulldowns ( $n = 791$  Up in the TurboID-noNES and  $n = 45$  Up in TurboID-NES,  $p \leq 0.05$ ,  $\geq 1\text{-log}_2\text{-fold change}$ , Supplementary Fig. 15i). We found DEPs of many different cellular subcompartments (e.g., RABL3 (ER), SCP2 (peroxisomes and nucleoplasm), GANAB (ER), NFKBIA (cytosol), and ETFA (mitochondria)) to be enriched in the TurboID-noNES group ( $p \leq 0.05$ ,  $\geq 1\text{-log}_2\text{-fold change}$ , Supplementary Fig. 15i). In contrast, many nucleoplasm-associated proteins appeared in the TurboID-NES pulldowns (e.g., SREK1IP1, RP9, HOXD13 and NGFR). Many mitochondrial carboxylases (e.g., PCCA and MCCC1) were also enriched, and these have been identified as endogenously biotinylated proteins important for cell metabolism<sup>68,69</sup>. These findings are in alignment with the BV2-TurboID-NES versus global differential enrichment analyses

(Supplementary Fig. 2 and 3) and further confirm that TurboID restricted to the extranuclear compartment can still interact and biotinylate some proteins in other cellular compartments. We also examined the correlation between mRNA and protein abundances in the TurboID-NES and TurboID-noNES pulldowns. In both cases, we observed modest correlations (TurboID-NES: Pearson's  $r=0.31$ ,  $n=6748$  pairs, TurboID-noNES:  $r=0.35$ , Supplementary Fig. 15l, m). These values are comparable, though slightly higher, than those observed in the in vitro BV2-TurboID-NES and in vivo astrocyte- and neuron-TurboID pulldowns ( $r=0.24$ – $0.27$ , Supplementary Fig. 11a–c). We also differentially compared the pulldown and global transcriptome and proteomes with or without the NES, respectively (Supplementary Figs. 17 and 18, Supplementary Data 59).

Finally, we tested if the removal of the NES on the TurboID enzyme would yield the same mRNA-protein discordant phenotypes as seen in the BV2-TurboID cells in vitro, astrocytes and neurons in vivo, and published datasets. The raw correlation between 6,748 mRNA and protein pairs was modest in the global (Pearson's  $r=0.47$ , Supplementary Fig. 19a), TurboID-NES pulldowns (Pearson's  $r=0.31$ , Supplementary Fig. 19b), and TurboID-noNES pulldowns (Pearson's  $r=0.35$ , Supplementary Fig. 19c). Again, we performed the percentile rank transformation and differential calculation to classify mRNA-protein pairs as concordant or discordant (Supplementary Fig. 19d, Supplementary Data 67–69). Using GO analysis, we observed some cytoskeletal-associated terms in the discordant protein-mRNA blue group across all datasets (Supplementary Fig. 19e, Supplementary Data 70–72), like the SPARO and Zhang versus Sharma published datasets analyses (Supplementary Fig. 13). We also again observed the discordant mRNA>protein mitochondrial phenotype across all datasets (Supplementary Fig. 19f, Supplementary Data 70–72).

## Discussion

### An approach for simultaneous transcriptome and proteome profiling of mammalian cells in their native state

Measuring transcriptomic and proteomic levels of gene products can provide important insights into cell state, response to injury or stress, and cell function under homeostatic and disease conditions. With rapid advances in -omics approaches, transcriptomics analyses of tissues and cells, and more recently proteomics analyses, are being broadly adopted in biology. There are several post-transcriptional and post-translational steps between gene expression at the mRNA level and functional protein levels, leading to a well-recognized modest correlation between the proteome and transcriptome of a cell<sup>61–63</sup>. Therefore, complementary profiling of both levels of abundance can be more meaningful than an assessment of each, individually. Traditional approaches to cell-type-specific transcriptomics and proteomics rely on dissociation and purification of individual cell types from their native tissues as a first step. However, this induces cell damage, stress, and death, leading to biased sampling and artifactual effects on their molecular profiles. Many cell types, such as neurons and astrocytes, have complex cellular architecture and processes that are disrupted or lost during isolation. While the transcriptome of a cell can be obtained from profiling its nucleus, the proteome of a cell cannot be fully captured without retaining its non-nuclear cellular components. Labeling of the proteome of the cell in vivo, using approaches such as proximity-labeling with biotin ligases, provides an avenue for phenotyping the proteome of cells without requiring cell-type purification. We describe an approach to capture the cell-type-specific transcriptome and proteome simultaneously that is applicable to in vitro and in vivo models. This method takes advantage of the biotin ligase, TurboID, to biotinylate proteins that interact with RNA, including ribosomal and RNA-binding proteins, which allows for enrichment of biotinylated proteins for proteomics as well as protein-associated RNA species for transcriptomics. In this study, we establish the SPARO approach to obtain valid proteomes and transcriptomes from BV2 microglia in vitro and further capture the effect of a

neuroinflammatory stimulus on microglia, using whole-cell global profiles of BV2 cells as gold-standard references. We also determine that the non-TurboID “noise” pulldown transcriptome consists primarily of many non-coding RNAs, pseudogenes, and splicing factors. After demonstrating the feasibility and validity of this approach in vitro, we extended SPARO to the in vivo setting, using *Camk2a*-expressing neurons and *Aldh1l1*-expressing astrocytes as brain cell types of interest, further validating the method for in vivo applications to measure both coding and non-coding RNAs. By benchmarking the SPARO-derived astrocyte transcriptome with RiboTag-derived astrocyte transcriptomes, we further confirmed that both approaches yield comparable results, although each approach has some relative advantages and disadvantages. Leveraging simultaneously enriched transcriptomes and proteomes of neurons and astrocytes, we also investigated patterns of mRNA-protein concordance across cell types. The SPARO approach addresses a major methodological gap in the field by providing a single pipeline for dual transcriptomics and proteomics from a desired cell type, while retaining the cell's native state in the tissue.

### SPARO to assess the concordance between paired proteomes and transcriptomes

One of the major advantages of SPARO is the ability to assess concordance between proteomes and transcriptomes of the same cells. In our in vivo astrocyte and neuron datasets, we observed only modest concordance between the pulldown proteomes and transcriptomes, which is in line with prior studies<sup>61–63</sup> and our own analyses of published datasets<sup>60,61</sup>. There are many possible reasons as to why the concordance is modest. From a technical perspective, the sensitivity of MS detection limits likely contributes to challenges of accurately measuring protein abundances and, therefore, the correlation with transcript abundances. Additionally, transcriptomic and proteomic datasets are usually analyzed in isolation with different data processing and statistical analysis approaches that may contribute to the poor correlation between mRNA and protein levels. Short-read bulk RNA-seq alone cannot resolve protein isoform differences that result from alternative splicing or post-translational modifications. Moreover, the dynamic range of abundance is much higher for proteins than for transcripts<sup>3,4,15</sup>.

From a biological standpoint, other possibilities for the mRNA-protein discordance are post-transcriptional regulatory mechanisms such as mRNA and protein stability, mRNA half-life, transcription and translation rates and efficiency, post-translational modifications, and vesicle-bound secretory pathways. We observed that the HEK293 TurboID-NES, HEK293 TurboID-noNES, BV2-TurboID global, and respective pulldown transcriptomes in vitro are highly correlated, suggesting that the TurboID-enriched transcriptome (regardless of the NES presence) is generally representative of the global. At the proteomic level, the global and respective pulldown groups were modestly correlated. This dichotomy likely reflects TurboID's ability to biotinylate numerous RNA-binding proteins, thereby capturing a more holistic transcriptomic cell state. In contrast, the modest proteome correlations may be attributed to protein localization and trafficking. Because we did not find a pattern associated with protein half-life and mRNA-protein discordance, we hypothesized that the cytosolic localization of TurboID and the biotinylation of a subset of the proteome may contribute to the modest correlation. However, when we removed the NES from the TurboID cells in vitro so that TurboID could move freely throughout the cell, the correlation remained modest. We also found that implementing a more sensitive MS approach, DIA, which yielded more mRNA-protein pairs in the analysis, or subletting the analysis to only cytosolic proteins, also did not change the correlation between the paired-omes. We therefore predict that other mechanisms or a collection of many cellular and molecular processes contribute to the discordance between mRNA and protein levels.



We identified candidate astrocytic and neuronal markers that are highly concordant or discordant. Some of the discordant markers are derived from secreted proteins (e.g., Apoe and Gap43). Although cytosolic TurboID has been found to biotinylate secreted proteins<sup>31,33,36,37,58</sup>, it is not well understood if TurboID is able to localize within extracellular vesicles (EV) found in the endolysosomal secretory pathway. However, there is evidence that proteins can be secreted without being vesicle-bound<sup>70</sup>, which would enable cytosolic TurboID to biotinylate such proteins. Future studies using different TurboID constructs localized to various subcellular compartments will allow for a more organelle-specific quantification of the concordance between paired proteomes and transcriptomes using SPARO. Future studies that will investigate post-transcriptional regulatory processes within astrocytes and neurons in their native environment will also likely shed light on the determinants of discordance between mRNA and protein levels.

We also found that many of the cellular processes involved in the concordant and discordant mRNA and protein pair groups are shared between astrocytes and neurons. For example, in the discordant group (protein  $\gg$  mRNA), astrocytes are enriched in microtubule proteins, and neurons are enriched in actin proteins. It is not fully clear whether the abundance of these cytoskeletal proteins differs across the two cell types in general, across subcellular compartments, or across different brain regions. It has been well reported that actin filaments are an important part of the neuronal cytoskeleton and contribute to the structure of axons and dendrites<sup>71</sup>. In contrast to our study in the cortex, a cytosolic BioID2-based LFQ-MS proteomics study found actin-filament-based processes enriched in mouse striatal astrocytes<sup>35</sup>, suggesting that there may be brain region differences. Another study found that microtubules are present in perivascular astrocytic end-feet<sup>72</sup>. Perhaps the high enrichment of microtubule proteins in the SPARO astrocytic pulldowns is due to capturing more end-feet-associated proteins, as suggested by the high abundance of Aqp4 protein as well. In the discordant group (mRNA  $\gg$  protein), different mitochondrial-related terms were enriched in astrocytes and neurons. The transcripts associated with mitochondrial function are nuclear-encoded, as opposed to mitochondrial DNA-encoded. Therefore, we predict that the mitochondrial transcripts are higher across both cell types because SPARO utilizes a cytosolic TurboID and undersamples mitochondrial proteins. Interestingly, we found ganglioside-induced differentiation-associated protein 1 (Gdap1), a neuronal mitochondrial outer membrane protein implicated in Charcot-Marie-Tooth disease<sup>73–75</sup>, to be enriched in the SPARO neuronal transcriptome, suggesting that SPARO may be a useful tool to study disease-associated genes and proteins. Although the cellular processes are similar across the two cell types, our findings suggest that astrocytes and neurons may use different molecular machinery. The differences are likely due to the unique functions astrocytes and neurons play in the brain<sup>76</sup>.

Furthermore, in our HEK293 TurboID-NES and TurboID-noNES study, we found that the pulldowns closely reflect their respective global transcriptomes and proteomes and are overall highly similar to each other. Differential enrichment analyses further revealed distinct subcellular biases. Despite such biases, the correlation between the paired transcriptomes and proteomes remained modest even after removing the NES. Most notably, the discordant phenotypes found using the TurboID-NES enzyme in the BV2-TurboID cells in vitro, and astrocytes and neurons in vivo, were still present in the HEK293 TurboID-noNES analysis. Together, this suggests that the discordant findings are capturing biological phenomena.

### The SPARO with RiboTag transcriptomes are similar and complementary

Our benchmarking of SPARO with RiboTag noted some similarities and differences. Both approaches captured similar transcript abundances broadly, and astrocyte-specific genes were highly abundant in both

datasets. When we differentially compared the RiboTag and SPARO-enriched transcriptomes, we found overrepresented ontologies involved in translation-, splicing-, and mitochondrial-related terms in the RiboTag pulldowns as compared to synapse- and ion transport-related terms overrepresented in the SPARO pulldowns. Notably, the highest DEP in our SPARO astrocytic pulldown proteome is Aqp4, which is a highly abundant water channel found in the astrocytic end-feet<sup>77</sup>. Perhaps the differences highlight methodological differences between the two approaches. For example, we observed that biotinylation by TurboID in astrocytes preferentially labels astrocytic end-feet, which is where Aqp4 is mostly localized.

When we investigated transcriptomic features (average transcript length, percent GC content, and number of exons) of DEGs between the two approaches, we found all three features to be significantly different. These differences may be attributed to SPARO capturing mRNA species beyond only those that are actively being translated. Our proteomic study shows that the astrocyte-RiboTag pulldown captures a smaller number of RNA-binding proteins and a larger proportion of ribosomal (translational) proteins, whereas SPARO captures many more RNA-binding proteins while also recovering a substantial number of ribosomal proteins. We predict that percent GC content may be indirectly related to transcript length<sup>78</sup>, and that the average GC content may also reflect the classes or functions of transcripts captured by each method. Because RiboTag primarily enriches ribosome-associated, actively translated mRNAs, these transcripts may be biased towards housekeeping genes with higher GC content<sup>79</sup>. In contrast, TurboID labeling likely recovers a broader pool of RNA-interacting proteins, including those associated with longer, less GC-rich transcripts at earlier stages of RNA processing. Consistent with this, TurboID has been shown to biotinylate proteins involved in transcription, splicing, and mRNA processing in our prior studies<sup>36,37</sup>, indicating that it can label RNA-interacting proteins at multiple stages of the mRNA life cycle. This multi-stage labeling may explain why SPARO transcriptomes include pre-spliced transcripts, which are generally longer and contain more exons than ribosome-associated, spliced mRNAs.

Despite these differences, we show that the SPARO and RiboTag astrocytic transcriptomes are indeed comparable and serve as appropriate tools to measure astrocytic genes in adult cortical astrocytes in vivo.

### Broader applications of SPARO

**Applying SPARO to study other cell types, brain regions, and subcellular compartments.** SPARO can be applied to study many neural cell types, brain regions, and subcellular compartments. To accomplish this, the Rosa26<sup>TurboID</sup> mouse can be bred with cell-type-specific Cre lines. Studies from our lab have already validated the use of the Rosa26<sup>TurboID</sup> mouse and the cell-type-specific in vivo biotinylation of proteins (CIBOP) approach in *Aldh1l1*-expressing astrocytes, *Camk2a*-expressing excitatory neurons, and parvalbumin (PV)-expressing interneurons<sup>31,33,36</sup>. The Jackson Laboratory has hundreds of Cre lines commercially available to study other CNS cell types. SPARO can also be utilized for other peripheral tissues and cell types, making it a broad tool for many fields.

Using CIBOP in astrocytic and neuronal TurboID models, regional differences across CNS tissues (e.g., cortical, subcortical, and spinal) have been reported<sup>36</sup>. These findings suggest that SPARO may be a valuable tool for investigating in vivo regional transcriptomic differences in these specific cell populations, or others. There is already substantial evidence of transcriptomic variation across mouse brain regions from single-cell RNA-sequencing and spatial transcriptomics data available through the Allen Brain Atlas<sup>80</sup>. Integrating SPARO datasets with Allen Brain Atlas resources offers a powerful opportunity to further characterize neural cell types. In contrast to our findings in the cortex, where we found astrocytes have high levels of microtubule

proteins, a cytosolic BioID2-based LFQ-MS proteomics study found actin-filament-based processes enriched in mouse striatal astrocytes<sup>35</sup>, suggesting that there may be brain region differences. Another study found that microtubules are present in perivascular astrocytic end-feet<sup>72</sup>. Perhaps the high enrichment of microtubule proteins in the SPARO astrocytic pulldowns is due to capturing more end-feet-associated proteins, as suggested by the high abundance of Aqp4 protein as well. These findings suggest that there may be regional differences in astrocytic cytoskeletal organization, which may contribute to unique astrocytic functions in different brain regions. Understanding regional differences in the brain is particularly important in the context of injury and disease, where immune responses, cellular damage, or pathology may be localized to specific areas. Using *in vivo* SPARO to study these differences offers opportunities to explore concordance and discordance between mRNA and protein abundances across regions. These patterns of concordance or discordance may be linked to injury or disease phenotypes, aid in identifying aberrant molecular processes, and highlight potential molecular pathways for therapeutic intervention. One experimental opportunity involves injecting pathological proteins, such as human tau or amyloid beta (A $\beta$ ), both implicated in Alzheimer's disease, into the brains of cell-type-specific TurboID mouse models and conducting SPARO studies. The burden of tau and A $\beta$  can be correlated with observed concordance or discordance patterns across regions, potentially revealing molecular processes that promote or inhibit pathology.

Furthermore, with the rapid evolution of MS instrumentation and pipelines, it is anticipated that the SPARO approach can be extended to small brain regions, lowly abundant cell types, and subcellular compartments with significantly deeper proteomic coverage in future studies<sup>81,82</sup>. Future modifications of TurboID *in vivo* models using both transgenic<sup>83</sup> and AAV-based approaches<sup>35,84</sup>, may provide opportunities to restrict TurboID to other cellular compartments or remove the cytosolic restriction by excluding the NES tag *in vivo*. These future endeavors are likely to extend the SPARO approach to interrogate compartment-specific protein and mRNA processes, such as local translation that occurs in synapses, dendrites, or glial processes<sup>85–94</sup> or expand our understanding of axonal structure and function<sup>95,96</sup>. A TurboID-mediated EV approach was recently developed<sup>97</sup> and a TurboID-EV-SPARO approach will be an exciting future tool to better understand the RNA and protein compositions of EVs.

**Measuring abundances of different RNA species.** Because TurboID biotinylates many RNA-interacting proteins, SPARO can also measure diverse non-coding RNAs (ncRNAs). Studying ncRNAs is particularly important because they play a significant role in various cellular processes, including transcription, translation, metabolism, and signaling<sup>98</sup>. miRNAs, for example, play key roles in binding mRNAs to control gene expression<sup>99</sup>, have been implicated in many diseases<sup>100</sup>, and are avenues for therapeutics<sup>101</sup>. In cortical astrocytes, SPARO detected six classes of small RNAs, including miRNAs, snRNAs, snoRNAs, piRNAs, tRNAs, and circRNAs. Among these, miR-760 was the most enriched in astrocytic pulldowns relative to the bulk cortex. miR-760 has been implicated in matrix metalloproteinase-2 (MMP-2) signaling and glioma proliferation, migration, and invasion<sup>102</sup>, suggesting that SPARO can uncover small RNAs with potential roles in disease. Additionally, our miRNA findings also allow for the extension of the SPARO pipeline to other approaches, such as miRNA tagging and the Affinity Purification method, miRAP<sup>103</sup>. Extending SPARO to systematically profile the biological significance of additional ncRNAs represents an important future direction.

**Ribosomal profiling and quantifying ribosome versus non-ribosomal bound transcripts.** Mammalian cells contain at least 79 different ribosomal proteins involved in cytosolic translation<sup>48</sup>, and

ribosomes with distinct protein compositions can translate different subsets of mRNAs<sup>49</sup>. The full extent of ribosome heterogeneity across cell types is still being uncovered. One promising application of SPARO is to investigate this heterogeneity and to distinguish between ribosome-bound and non-ribosome-bound transcripts. A common experimental approach is ribosomal fractionation using a sucrose gradient, which separates ribosomes and polyribosomes (ribosomes bound to mRNA) based on size and density<sup>104</sup>. These fractions can then be analyzed using SPARO-based proteomics and transcriptomics to gain insight into ribosomal dynamics in specific cell populations.

SPARO can also be used to compare ribosome profiles between global and cell-type-enriched samples, revealing differences in ribosomal protein abundance and cell-type-specific translational phenotypes. Once these profiles are established, perturbation studies, such as overexpression or knockout of ribosomal proteins, can be performed to examine how mRNA translation changes in a cell-type-specific context. This can help identify conserved or compensatory mechanisms in ribosomal function across different cell types. Additionally, distinguishing ribosome-bound from non-ribosome-bound transcripts offers a more precise understanding of gene regulation and protein synthesis, particularly in cases of mRNA-protein abundance discordance.

**Exploring mechanisms of discordance between mRNA and protein levels.** We found modest concordance between the pull-down transcriptomes and proteomes of astrocytes and excitatory neurons *in vivo* using SPARO. To report possible biological reasons for this, experiments that investigate post-transcriptional regulatory mechanisms will be required. Ways to accomplish this include performing mRNA half-life experiments by inhibiting transcription (e.g., via actinomycin D treatment or 4-thiouridine (4SU))<sup>105–107</sup>, or using a pulse-chase approach<sup>108–110</sup>. Moreover, because SPARO can capture ncRNAs, post-transcriptional processes that involve ncRNAs can be experimentally investigated. For example, miRNAs can bind to mRNAs and either prevent translation or promote the degradation of mRNAs. Because it is estimated that >60% of the human coding genes contain predicted miRNA target sites<sup>111</sup>, there are numerous avenues for exploration of miRNA-mediated processes (e.g., phenotypes related to upregulation/overexpression<sup>112–115</sup> or downregulation/inhibition<sup>101,114,116,117</sup>).

Furthermore, the ubiquitin-proteasome system for protein degradation can also be targeted for cell-type-specific SPARO studies. Ubiquitination is a PTM that is attached to lysine residues of proteins, which in turn leads to the recognition and downstream degradation of the protein by the proteasome<sup>118</sup>. There are small molecule inhibitors, activity-based probes, or peptides that can target the E1, E2, or E3 enzymes that are important for ubiquitination<sup>119–122</sup>. Although we did not find an association between protein half-life and mRNA-protein discordance in our SPARO studies, additional experiments will be helpful to validate this finding. One validation approach would be to perform stable isotope labeling studies *in vitro* and *in vivo*<sup>15,123–135</sup>. Future studies that will investigate post-transcriptional regulatory processes using SPARO will help uncover the determinants of discordance between mRNA and protein levels.

### Limitations of this study

Although most protein translation occurs in the cytosol, TurboID-NES likely labels only a subset of the cellular proteome, potentially missing proteins localized to other compartments such as the nucleus or mitochondria. In both the current study and prior work, TurboID-NES labeled approximately 40–50% of the global proteome in a monoculture BV2 system using DDA LFQ-MS<sup>37</sup>. This partial coverage is biologically expected, as many proteins are either synthesized in or transported to organelles like mitochondria, where they may be inaccessible to cytosolic TurboID. Consistent with its cytosolic localization,

the SPARO transcriptomic data show a strong bias against capturing mitochondrial-encoded transcripts. However, this trend is less evident in the SPARO proteomic data. We attempted to address the labeling biases of TurboID in our HEK293 experiments comparing the SPARO-omes with and without the NES. We found that pulldowns closely reflect their respective global transcriptomes and proteomes and are overall highly similar to each other, regardless of the NES. Differential enrichment analyses further revealed distinct subcellular biases. Despite such biases, the correlation between the paired transcriptomes and proteomes remained modest even after removing the NES.

Another biological factor contributing to incomplete proteome coverage is the presence of membrane-bound or secreted proteins, which may evade labeling by TurboID-NES. Nonetheless, TurboID can still label vesicle-associated proteins that have cytosol-facing domains, such as peripheral membrane proteins. Additionally, vesicle fusion events may allow transient exposure of vesicular contents to TurboID. Furthermore, TurboID labels proteins with biotin at lysine residues<sup>136</sup>; however, some proteins may contain very few lysines, or “lysine deserts”<sup>137</sup>, which could lead to lower labeling efficiency and reduced detection sensitivity by MS. Additionally, lysine residues are also targets of other post-translational modifications (PTMs), such as ubiquitination and acetylation<sup>138</sup>. It remains unclear to what extent these PTMs can coexist with biotinylation at the same lysine residue, and how the presence of other modifications or structural factors may affect TurboID’s accessibility to specific lysines.

From a technical standpoint, several factors may influence how comprehensively the TurboID-labeled proteome is captured. First, the dosage and duration of exogenous biotin supplementation can affect the yield of biotinylated proteins. For in vitro applications, biotin dosage and timing have been well characterized<sup>30</sup>. For in vivo mouse models, two weeks of biotin-supplemented water has been used to successfully label the proteomes of astrocytes and neurons<sup>31,33,36</sup>. However, shorter labeling windows will require optimization. Additionally, there is variability across animals in the extent of protein biotinylation, likely due to differences in biotin water intake. Another possible contributing factor is variation in the abundance of biotin transport proteins<sup>139</sup> across cell types, which may affect intracellular biotin uptake. Moreover, the efficiency of streptavidin-based enrichment and the MS approach are factors influencing the depth of the cellular proteome. Further investigation is needed to determine how different MS strategies, such as LFQ versus TMT labeling, or data-dependent versus data-independent acquisition, affect proteomic depth.

In conclusion, SPARO is an approach that can be used to study many facets of cellular and molecular biology. For example, SPARO in vitro can be used for cause-and-effect mechanistic studies using reductionist models in human multi-cellular model systems in vitro. Additionally, SPARO can be used to study how transcriptional and translational processes are fully regulated, changes in the efficiency of protein biosynthesis, processes that contribute to the discordance between mRNAs and proteins, mRNA and protein degradation, and much more. A major advantage of SPARO is that it can be adapted to study different cell types, brain regions, and other tissues outside the brain, making the approach a highly versatile and broadly applicable tool for the field of molecular biology.

## Methods

### Cloning, plasmid, and lentivirus generation

Through Emory’s Integrated Genomics Core Custom Cloning Division, the TurboID-NES construct (Addgene plasmid #107169) used in our prior study<sup>37</sup>, was edited to generate a construct without the NES (TurboID-noNES). Restriction enzyme sites were introduced with respective forward and reverse primers (Supplementary Data 1). PCR

products were then subcloned into the TurboID-NES backbone. TurboID-NES and TurboID-noNES plasmids were transformed into DH5- $\alpha$  competent *E. coli* cells (NEB, C29871) according to the manufacturer’s instructions. Plasmid DNA was purified following the Miracle-prep (Miraprep) instructions<sup>140</sup> (QIAprep Spin Miniprep Kit), and the resulting concentration was measured using a Qubit. Lentivirus (LV) using the TurboID-NES and TurboID-noNES plasmids was generated through Emory’s Viral Vector Core.

### In vitro studies

Mouse microglia BV2 cells were cultured in 0.2  $\mu$ m vacuum filter-sterilized Dulbecco’s Modified Eagle Medium (DMEM) supplemented with high glucose and L-glutamine, containing 1% penicillin-streptomycin, and 10% fetal bovine serum. BV2 cells were incubated in a 100 mm cell culture dish at 37 °C and 5% CO<sub>2</sub> until reaching 80–90% confluency. Cells were treated with 100 ng/mL of LPS for 48 h and 200  $\mu$ M of biotin for 24 h during the second day of LPS treatment.

Human embryonic kidney 293 cells (HEK293) were cultured in DMEM containing 1% penicillin-streptomycin and 10% fetal bovine serum. HEK293 cells were incubated in 6-well culture dishes under the same conditions as the BV2 cells described above. At 30% confluency, HEK293 cells were infected with 3  $\mu$ L of the two LVs generated (TurboID-NES, viral titer:  $3.9 \times 10^8$  IU/mL or TurboID-noNES, viral titer:  $1.7 \times 10^9$  IU/mL). LV-containing media were removed after 3 days, and successful infection was determined by cellular GFP signal using the ECHO Revolve fluorescence microscope (Echo, USA) prior to lysis. HEK293 cells were grown to 80–90% confluency before biotin treatment (200  $\mu$ M for 24 h).

The 48 h LPS and 24 h biotin treatment paradigms were used to model similar treatment and labeling kinetics for in vivo SPARO studies.

### Immunocytochemistry

BV2 control and BV2-TurboID cells were seeded at a density of  $1 \times 10^5$  cells onto autoclaved and ethanol-treated coverslips placed in a 12-well plate. Seeded cells were treated with 100 ng/mL LPS. After 24 h, the cells were supplemented with external biotin (200  $\mu$ M), along with LPS, for an additional 24 h. After treatments, cells were washed three times with PBS and then fixed with 4% paraformaldehyde (Thermo Scientific, J19943-K2) in PBS for 30 min at RT. Fixed cells were permeabilized and blocked simultaneously using 5% goat serum diluted in 0.25% TBS-T for 1 h at room temperature. Each well was then incubated with 1 mL of antibody cocktail containing rabbit anti-V5 (1:500, Abcam, 206566) and Streptavidin-DyLight 488 (1:500, Invitrogen, 21832) prepared in 1% Goat serum diluted in TBS, overnight at 4 °C. The next day, cells were washed three times with 0.25% TBS-T, followed by incubation with anti-rabbit Alexa Fluor 594 secondary antibody (1:500, Invitrogen, A21207) prepared in 1% goat serum in TBS for 2 h at room temperature. Nuclei were stained with 4’,6-diamidino-2-phenylindole (DAPI) for 5 min. After final washes, coverslips were mounted onto glass slides using Fluoromount-G (Invitrogen, 00-4958-02). Images were captured at 60 $\times$  using a Keyence BZ-X810 microscope.

### In vivo studies

Mice were housed in the Division of Animal Resources vivarium at Emory University under a standard environment (12 h light/ 12 h dark cycle, temperature 72 °F, humidity range 40–50%) with access to food and water ad libitum. All animal-related studies (PROTO-201700821) were conducted with the approval of Emory’s Institutional Animal Care and Use Committee following the National Institute of Health’s “Guide for the Care and Use of Laboratory Animals” and reported in accordance with the ARRIVE guidelines.

Transgenic approach for Cre recombinase expression in neurons and astrocytes using Rosa26<sup>TurboID/wt</sup> mice:



**Neuronal TurboID cohort.** Rosa26<sup>TurboID/wt</sup> (Jackson Labs, Strain No. 037890) were crossed with Camk2a<sup>CreERT2</sup> (Jackson Labs, Strain No. 012362) mice to obtain heterozygous Rosa26<sup>TurboID/wt</sup>/Camk2a<sup>CreERT2/wt</sup> (“neuron-TurboID”,  $n = 3$ , 1 M and 2 F littermate mice). Heterozygous littermate mice Camk2a<sup>CreERT2/wt</sup> (“Camk2a”,  $n = 2$ , 1 M and 1 F) were used as controls.

**Astrocytic TurboID cohort.** Rosa26<sup>TurboID/wt</sup> (Jackson Labs, Strain No. 037890) were crossed with Aldh1l1<sup>CreERT2</sup> (Jackson Labs, Strain No. 031008) mice to obtain heterozygous Rosa26<sup>TurboID/wt</sup>/Aldh1l1<sup>CreERT2/wt</sup> (“astrocyte-TurboID”,  $n = 8$ , 6 M and 2 F). Heterozygous littermate mice Aldh1l1<sup>CreERT2/wt</sup> (“Aldh1l1”,  $n = 3$ , 2 M and 1 F) were used as controls. For astrocyte-TurboID LPS studies, systemic LPS at 0.75/kg/dose  $\times$  5 days was given via i.p. injection during the first week of biotin water supplementation. Based on previous work and our own, we validated neuron- and astrocyte-specificity and nonleaky Cre activity of the Camk2a<sup>CreERT2</sup> and Aldh1l1<sup>CreERT2</sup> models, respectively<sup>42,59</sup>.

**Astrocytic RiboTag cohort.** RiboTag mice, Rpl22<sup>tm1.1P<sub>Sam</sub></sup> (Jackson Labs, Strain No. 029977), were crossed with Aldh1l1<sup>CreERT2</sup> (Jackson Labs, Strain No. 031008) mice to obtain heterozygous Rpl22<sup>tm1.1P<sub>Sam</sub></sup>/Aldh1l1<sup>CreERT2/wt</sup> (“astrocyte-RiboTag”,  $n = 3$ , 3 F). Heterozygous littermate mice Rpl22<sup>tm1.1P<sub>Sam</sub></sup> (“RiboTag”,  $n = 2$ , 1 M and 1 F) were used as controls.

All mice were given tamoxifen (75 mg/kg) intraperitoneally (i.p.) for 5 days at 6 weeks of age and allowed 3 weeks of Cre recombination. After Cre recombination, TurboID mice were given water supplemented with biotin (37.5 mg/L) for 2 weeks until euthanasia at 3–4 months of age. RiboTag mice were euthanized at 2.5 months of age. Mice from each cohort were anesthetized using ketamine/xylazine (ketamine 87.5 mg/kg, xylazine 12.5 mg/kg) followed by transcardial perfusion with ice-cold 1 $\times$  PBS. The brain was immediately removed and hemisected along the mid-sagittal line. The left hemisphere was immediately dissected to obtain the cortex and then snap frozen using dry ice. The right hemisphere was used for immunohistological studies.

### Immunohistochemistry

For immunohistochemistry studies, mice were euthanized and perfused with ice-cold 1 $\times$  PBS, and the brain was dissected. One hemisphere was immediately transferred into 4% PFA (Cat No. J19943.K2) and fixed overnight for immunohistological analysis. After PFA fixation, the brain was washed with 1 $\times$  PBS 3 times to remove residual PFA and transferred into 30% sucrose solution until sectioning. After removing extra sucrose, the fixed and sucrose-saturated hemisphere was embedded in the Tissue-Tek optimal cutting temperature compound (Cat No. 4583, Sakura) and frozen on dry ice. Serial 40  $\mu$ m sagittal sections of the brain were generated using a cryostat (Leica Biosystems, CM1850) and stored free-floating in cryoprotective media (Glycerin, Ethylene Glycol, and 0.1 M Phosphate Buffer in a ratio of 2.5:3.0:5.0) at  $-20^{\circ}\text{C}$ . For immunofluorescence (IF) staining, 3–4 brain sections from each mouse were washed 3 times with 1 $\times$  TBS and then blocked and permeabilized by incubating in 1 $\times$  TBS containing 0.25% Triton X-100 (TBS-T) and 5% goat serum for 1 h at room temperature. The following sections were incubated with primary antibodies diluted in 1 $\times$  TBS-T containing 1% goat serum overnight at  $4^{\circ}\text{C}$ . Next, the sections were washed 3 $\times$  with 1 $\times$  TBS and then incubated with fluorophore-conjugated secondary antibodies for 2 h at room temperature in the dark. All primary and secondary antibodies were used at a dilution of 1:500, including HA-tag (Cell Signaling Technology, 3724S), S100 $\beta$  (Proteintech, 15146-1-AP),  $\beta$ III-Tubulin (Biolegend, 802001), Streptavidin Protein, DyLight<sup>TM</sup> 488, and Donkey anti-Rabbit-Alexa Fluor<sup>TM</sup> 594. Brain sections were then incubated for 5 min in nuclear staining reagent DAPI (1  $\mu$ g/mL) dissolved in 1 $\times$  TBS. The samples were further washed 3 $\times$  with 1 $\times$  TBS, mounted onto glass

slides, and coverslipped using ProLong Diamond Antifade Mountant (P36965, Thermo Fisher). Images of the same brain regions for each mouse were captured using a Keyence fluorescence microscope (Keyence BZ-X810). Images were processed and analyzed using ImageJ software (Fiji Version 2.14).

### Cell and tissue homogenization

BV2 control, BV2-TurboID, HEK293 control, HEK293 TurboID-NES, and HEK293 TurboID-noNES cells were washed three times with cold 1 $\times$  PBS containing 100  $\mu$ g/mL of cycloheximide. Next, cells were incubated with 1 mL of cold homogenization buffer (10 mM HEPES pH 7.4, 150 mM KCl, 10 mM MgCl<sub>2</sub>, supplemented with 0.5 mM DTT, 0.1% v/v RNasin, 0.1% v/v SUPERasin, 1 $\times$  Halt<sup>TM</sup> protease and phosphatase inhibitors, 100  $\mu$ g/mL cycloheximide) for 10 min, scraped and collected into a 1.5 mL Eppendorf LoBind tube. For brain samples, frozen Camk2a and Aldh1l1 controls and neuron-TurboID and astrocyte-TurboID mouse cortices (up to 50 mg) were lysed in the same cold homogenization buffer as the BV2 cells with RNase-free glass beads and a Bullet Blender Homogenizer (Next Advance). RiboTag control and astrocyte-RiboTag mouse cortices were lysed in cold RiboTag homogenization buffer (50 mM Tris-HCl, pH 7.5, 100 mM KCl, 12 mM MgCl<sub>2</sub>, 1% NP-40, 1 mM DTT, 1% v/v sodium deoxycholate, 0.1% v/v RNasin, 0.1% v/v SUPERasin, 1 $\times$  Halt protease and phosphatase inhibitors, 100  $\mu$ g/mL cycloheximide) similarly to the TurboID samples. All homogenates underwent 3 rounds of sonication (5 s of active sonication at 25% amplitude followed by a 10 s incubation on ice) followed by centrifugation for 10 min at 2000  $\times g$  at  $4^{\circ}\text{C}$ . The supernatants were transferred to a new tube, and protein concentrations were calculated using a Pierce BCA assay for proteomic studies. To account for variability in infection efficiency, the HEK293 cell lysates per condition (control uninfected, TurboID-NES, and TurboID-noNES) were pooled prior to streptavidin affinity purification.

### Immunoblotting

To confirm the presence of biotinylated proteins, 17  $\mu$ g of cell or tissue lysates were resolved on a 4–12% Bis-Tris gel, transferred onto a nitrocellulose membrane, blocked with StartingBlock for 30 min and then probed with a streptavidin-Alexa Fluor 680 antibody (Thermo Fisher, S32358, dilution: 1:10,000 in StartingBlock) for 1 h at room temperature. To confirm TurboID construct presence, the membrane was probed with a rabbit anti-V5 primary antibody (dilution: 1:500 in StartingBlock) overnight and then incubated with a goat anti-rabbit HRP-conjugated secondary antibody (Jackson ImmunoResearch, 111-035-003, dilution: 1:10,000 in StartingBlock). Immunoblots were imaged with an Odyssey Infrared Imaging System (LI-COR Biosciences) or a Bio-Rad chemiluminescence system.

### Biotinylated protein and ribosomal protein enrichment for transcriptomics

Following confirmation of biotinylated proteins and V5 presence via immunoblotting, biotin-tagged proteins were enriched for using streptavidin-coated magnetic beads<sup>36</sup>. For each sample in the BV2, astrocyte-TurboID and neuron-TurboID studies, 25  $\mu$ L of streptavidin beads (Thermo Fisher, 88817) in a 1.5 mL Eppendorf LoBind tube were washed 3 times with 1 mL of homogenization buffer on rotation for 2 min. The streptavidin beads were incubated with 300  $\mu$ g of protein lysate from each sample with additional wash buffer (10 mM HEPES pH 7.4, 150 mM KCl, 10 mM MgCl<sub>2</sub> supplemented with 0.5 mM DTT, 1% v/v NP-40 substitute, 0.1% v/v RNasin, 0.1% v/v SUPERasin, 1 $\times$  Halt<sup>TM</sup> protease and phosphatase inhibitors, 100  $\mu$ g/mL cycloheximide) to create a final volume of 250  $\mu$ L on rotation for 1 h at  $4^{\circ}\text{C}$ . For the HEK293 studies, 17  $\mu$ L of streptavidin beads and 200  $\mu$ g of protein lysate from each pooled sample were used. Next, the beads were washed four times with cold high salt buffer (10 mM HEPES pH 7.4, 350 mM KCl, 10 mM MgCl<sub>2</sub> supplemented with 0.5 mM DTT, 1% v/v NP-

40 substitute, 0.1% v/v RNasin, 0.1% v/v SUPERasin, 1× Halt™ protease and phosphatase inhibitors). After RiboTag control and Astrocyte-RiboTag cortical lysis, 800 µL of lysate was incubated with 4 µL of anti-hemagglutinin (HA) antibody (Biolegend, 901513) for 4 h on end-over-end rotation at 4 °C. For each sample, 25 µL of A/G beads (Thermo Fisher, 88803) in a 1.5 mL Eppendorf LoBind tube were washed 3 times with 400 µL of RiboTag homogenization buffer. The A/G beads were incubated with 100 µL of HA-tagged lysate overnight on end-over-end rotation at 4 °C. Next, the A/G beads were washed four times with cold RiboTag high salt buffer (50 mM Tris-HCl, pH: 7.5, 300 mM KCl, 12 mM MgCl<sub>2</sub>, 1% v/v NP-40, 0.50 mM DTT, 100 µg/mL cycloheximide). (10 mM HEPES pH 7.4, 350 mM KCl, 10 mM MgCl<sub>2</sub>, supplemented with 0.5 mM DTT, 1% v/v NP-40 substitute, 0.1% v/v RNasin, 0.1% v/v SUPERasin, 1× Halt™ protease and phosphatase inhibitors). After final washes, the streptavidin and A/G beads were resuspended in 100 µL of wash buffer and then added to 700 µL of Trizol and stored at −80 °C until RNA extraction. For global/input lysate/animal, 50 µL of lysate remained at 4 °C during the protein enrichment protocols.

### RNA extraction

After protein enrichment using streptavidin and A/B beads, RNA was extracted from the beads and/or global lysates using a miRNeasy Mini Kit (Qiagen Inc., 217004) following the manufacturer's instructions. RNA was eluted in RNase-free water and analyzed for purity and concentration using either an Agilent 2100 Bioanalyzer with the RNA 6000 Nano or Pico Kit, or Agilent 4200 TapeStation System prior to RNA-sequencing library preparation.

### RNA-sequencing and processing

After assessing RNA quality, cDNA library preparation was performed. The SMART-Seq® v4 PLUS Kit (Takara Bio) was used following the manufacturer's protocol. For small RNA-sequencing library preparation, a QIAseq miRNA NGS kit (Qiagen, 331565) was used. Global sequencing of 40 million paired-end reads per sample was completed using the Illumina-based platform at Admera Health. FASTQ files were evaluated for quality using FastQC<sup>41</sup> and then mapped to the mouse genome, GRCh38 (mm10), or GRCh38, human genome (hg38), using the Spliced Transcripts Alignment to a Reference (STAR)<sup>142</sup> (for mRNA-seq) or Bowtie 2<sup>143</sup> (for smRNA-seq) aligners with the paired-end option. The *featureCounts*<sup>144</sup> package was used to generate a raw counts data matrix from mapped reads. Lowly abundant transcripts across all samples (average count value ≤ 10) were filtered out. After generating the filtered, raw counts data matrices, we used *DESeq2*<sup>145</sup> to normalize the matrices and perform differential gene expression-based statistics. For astrocyte-TurboID versus astrocyte-RiboTag studies, we used the same astrocyte samples used in neuron-TurboID comparison. RNA-sequencing was completed independently, and the resulting FASTQ files were processed and analyzed together. For the Zhang et al.<sup>60</sup> transcriptomes versus the Sharma et al.<sup>61</sup> proteomes of acutely isolated astrocytes and neurons, lowly abundant transcripts were filtered out using an average FPKM value of ≤ 3.

### Biotinylated or HA-tagged polyribosome protein enrichment for proteomics

For each sample in the BV2, astrocyte-TurboID, and neuron-TurboID studies, 42 µL of streptavidin beads in a 1.5 mL Eppendorf LoBind tube were washed 2 times with 1 mL of RIPA lysis buffer (50 mM Tris, 150 mM NaCl, 0.1% SDS, 0.5% sodium deoxycholate, 1% Triton X-100) on rotation for 2 min. The beads were incubated with 500 µg of protein lysate from each sample with additional RIPA lysis buffer to create a final volume of 500 µL on rotation for 1 h at 4 °C. For the HEK293 studies, 13 µL of streptavidin beads and 150 µg of protein lysate from each pooled sample were used. Next, the beads were quickly centrifuged, placed on a magnetic rack, and the supernatant was transferred to a new 1.5 mL Eppendorf LoBind tube and stored at

−80 °C. The beads were washed with the following buffers on rotation at room temperature: twice with 1 mL of RIPA lysis buffer for 8 min, once with 1 M KCl for 8 min, once with 1 mL 0.1 M sodium carbonate for −10 s, once with 1 mL 2 M urea in 10 mM Tris-HCl (pH 7.6) for −10 s, and twice with 1 mL RIPA lysis buffer for 8 min. After the final RIPA wash, the beads were resuspended in 1 mL of 1× PBS, transferred to a new tube, and washed two more times with 1× PBS on rotation for 2 min. To confirm biotinylated protein enrichment, 10% of the streptavidin bead volume was transferred to a new 1.5 mL Eppendorf LoBind tube and boiled in 30 µL of 2× Laemmli protein loading buffer (Bio-Rad, 1610737) supplemented with 2 mM biotin and 20 mM dithiothreitol (DTT) at 95 °C for 10 min to elute the biotinylated proteins. Following, 1/3 of the eluate was resolved on a 4–12% Bis-Tris gel, transferred onto a nitrocellulose membrane, blocked with StartingBlock for 30 min, and incubated overnight with a rabbit anti-V5 antibody (Abcam, ab206566) on a shaker at 4 °C. Next, the membrane was incubated with a goat anti-rabbit HRP-conjugated secondary antibody (Jackson ImmunoResearch, 111-035-003, dilution: 1:10,000 in StartingBlock) and streptavidin-Alexa Fluor 680 antibody (Thermo Fisher, S32358, dilution: 1:10,000 in StartingBlock) for 1 h at room temperature. The remaining 2/3 of protein eluate was resolved on a 4–12% Bis-Tris gel for a silver stain (Thermo Fisher, 24612). Immunoblots were imaged with an Odyssey Infrared Imaging System (LI-COR Biosciences) or a Bio-Rad chemiluminescence system. Silver-stained gels were imaged using a Canon scanner (CanoScan LIDE 300).

For astrocyte-RiboTag-based proteomics of HA-tagged polyribosomes and associated proteins, 50 µL of A/G beads were incubated with 200 µL of anti-HA lysate per sample. This ratio was used to approximately match the astrocyte-TurboID bead to lysate ratio. The rest of the protocol matched the RNA pulldowns with the addition of RNA depletion via 50 µg/mL of RNaseA (Thermo Scientific, EN0531) in PBS for 15 min at room temperature on end-to-end rotation. After, the beads were washed 3 times with PBS prior to preparation for MS.

### Protein digestion

To prepare enriched biotinylated proteins from BV2, astrocyte-TurboID, and neuron-TurboID for MS-based proteomics, the remaining 90% volume of the streptavidin beads was resuspended in 50 mM ammonium bicarbonate (Na<sub>3</sub>CO<sub>3</sub>). Next, the bound proteins were reduced with 1 mM DTT for 30 min at room temperature and then alkylated with 5 mM iodoacetamide (IAA) in the dark for 30 min. Following, proteins were subsequently digested overnight with 0.5 µg of lysyl (Lys-C) endopeptidase (Wako, 127-06621) and then 1 µg of trypsin (Thermo Fisher, 90058) on the shaker at room temperature. After digestion, the resulting peptide mixtures were acidified to a final concentration of 1% formic acid and 0.1% trifluoroacetic acid, desalted with an HLB column (Waters, 186003908), and dehydrated using a vacuum centrifuge (SpeedVac Vacuum Concentrator). The same approach was used for HA-tagged polyribosomes from astrocyte-RiboTag samples. To prepare global lysates for MS proteomics, 50 µg of protein from each sample was reduced in 5 mM DTT for 30 min at room temperature and then alkylated with 10 mM IAA for 30 min in the dark. After, each sample was diluted (4-fold) with 50 mM ammonium bicarbonate (ABC) buffer and then digested with 1 µg of Lys-C endopeptidase on the shaker overnight at room temperature. Next, samples were diluted (4-fold) with 50 mM ABC buffer and digested with 2 µg of trypsin on the shaker overnight at room temperature. After digestion, the peptide mixtures were acidified, desalted, and dried down as described above.

To prepare enriched biotinylated proteins from HEK293 TurboID-NES and TurboID-noNES samples for MS-based proteomics, the remaining 90% of the streptavidin beads were resuspended in 5% SDS in 50 mM triethylammonium bicarbonate (TEAB). Proteins were reduced with 10 mM DTT for 1 h at room temperature, followed by alkylation with 20 mM iodoacetamide (IAA) in the dark for 30 min. The

proteins were then acidified with phosphoric acid. A mixture of 100 mM TEAB and 90% methanol (MeOH) was added, and the samples were transferred to S-Trap micro spin columns (Protifi). Binding and washing steps were performed according to the S-Trap protocol. Digestion was carried out by adding 0.5 µg Lys-C and 1 µg trypsin, followed by incubation at 37 °C for 18 h. Peptides were eluted sequentially with 50 mM TEAB, 0.2% formic acid, and 50% acetonitrile, then dried using a SpeedVac. The same protocol was followed for the global lysates, with 25 µg of protein from each sample used as input. Peptide concentration was determined by fluorescence-based BCA assay. Finally, samples were injected volumetrically into the mass spectrometer.

### Mass spectrometry

Mass spectrometry preparation and acquisition parameters were selected based on conditions established in our prior work<sup>31,33,36</sup> and detailed here. For BV2, astrocyte-TurboID, and neuron-TurboID studies, dried peptides were reconstituted in 15 µL of loading solution (0.1% formic acid and 0.03% trifluoroacetic acid in water), and 7–8 µL was injected onto a self-packed 25 cm analytical column (100 µm inner diameter) packed with 1.7 µm Waters CSH particles. Peptides were separated using either an Easy-nLC 1200 or a Dionex 3000 RSLCnano nanoLC system. The gradient began at 1% solvent B (80% acetonitrile containing 0.1% formic acid) and increased to 5% over 10 s, followed by a 55-min linear gradient to 35% solvent B. The run concluded with a 4-min, 50-s wash at 99% solvent B. Following, an Orbitrap Lumos Tribrid MS with a high-field asymmetric waveform ion mobility spectrometry (FAIMS Pro)<sup>146</sup> interface was used to obtain all mass spectra at a compensation voltage of −45V. BV2-TurboID in vitro and astrocyte-TurboID and neuron-TurboID in vivo studies were completed using data-dependent acquisition (DDA) LFIQ-MS using top-speed mode with a 3 s cycle time. Full MS scans were acquired in the Orbitrap at 120,000 resolution across an m/z range of 400–1600, with an automatic gain control (AGC) target of 400,000, a maximum injection time of 50 ms, and the RF lens set to 30%. MS/MS spectra were generated by higher-energy collision dissociation (HCD) and analyzed in the ion trap using a normalized collision energy of 35% and an isolation window of 0.7 m/z. The AGC target for MS/MS scans was 10,000 with a maximum injection time of 35 ms. Dynamic exclusion was enabled for 30 s with a mass tolerance window of 10 ppm.

For astrocyte-RiboTag and HEK293 studies, dried peptides were separated on a 75 µm × 60 cm analytical column (Thermo Fisher Scientific) using a 135-min (astrocyte-RiboTag) or 115-min (HEK293) gradient from 1–99% solvent B (80% acetonitrile, 0.1% formic acid) at 300 nL/min. MS analysis was performed on an Orbitrap Exploris 480 (Thermo Fisher Scientific) in positive ion mode. Full MS scans were acquired at 120,000 resolution (m/z 400–1000), and data-independent-acquisition (DIA) MS/MS scans at 30,000 resolution (m/z 150–2000) using 10 m/z isolation windows with 1 m/z overlap and HCD fragmentation at 28% normalized collision energy.

### Protein identification and quantification

DDA raw MS files for global and pulldown protein from BV2-TurboID in vitro and astrocyte- and neuron-TurboID in vivo studies, along with respective controls, were searched separately using the Andromeda<sup>147</sup> search engine integrated into MaxQuant<sup>148</sup> (version v2.4.2.0), against the 2020 downloaded mouse UniProt proteome (<https://www.uniprot.org/proteomes/UP000000589>), including sequences for V5 and TurboID. DIA Raw MS files from the astrocyte-RiboTag pulldown and the HEK293 TurboID (NES versus noNES) global and pulldown experiments were searched separately using the Pulsar search engine integrated into Spectronaut (version 19)<sup>149</sup>. The astrocyte-RiboTag search was performed against the same 2020 mouse UniProt proteome described above (with the HA sequence included). The HEK293 TurboID studies were searched against the 2025 downloaded human

UniProt proteome (<https://www.uniprot.org/proteomes/UP000005640>). Search parameters were determined by our prior work<sup>31,33,36,37,150</sup> and described here. Methionine oxidation (+15.9949 Da) and protein N-terminal acetylation (+42.0106 Da) were variable modifications (up to 5 allowed per peptide); cysteine was assigned as a fixed carbamidomethyl modification (+57.0215 Da). Only fully tryptic peptides were considered with up to 2 missed cleavages in the database search. A precursor mass tolerance of ±20 ppm was applied prior to mass accuracy calibration and ±4.5 ppm after internal calibration. Other search settings included a maximum peptide mass of 4600 Da, a minimum peptide length of 6 residues, 0.05 Da tolerance for Orbitrap and 0.6 Da tolerance for ion trap MS/MS scans. Quantification settings were as follows: re-quantify with a second peak finding attempt after protein identification has completed; match MS1 peaks between runs; a 0.7 min retention time match window was used after an alignment function was found with a 20 min RT search space. Peptide spectral match false discovery rate (FDR) was set to 1%. The MaxQuant or Spectronaut output data (raw intensity values) were uploaded to R (version 4.3.1) for normalization, log<sub>2</sub> transformation, and differential abundance analyses.

To account for endogenously biotinylated and non-specific proteins captured during streptavidin affinity purification, mean intensity values from TurboID-negative control pulldown samples were subtracted from the TurboID-positive pulldown samples by group. To account for variability in TurboID expression, biotinylated protein abundance, and streptavidin affinity purification efficiency, intensity values were normalized depending on the dataset. For BV2 in vitro studies comparing the abundance values from global protein and enriched pulldown samples, samples were normalized by taking each sample's log<sub>2</sub>-transformed summed protein intensity median and setting it to zero. The BV2-TurboID pulldown groups (untreated and LPS-treated) intensity values were normalized based on TurboID abundance. We did not find a difference in TurboID abundance in the other SPARO datasets, so this normalization approach was only completed on the BV2-TurboID pulldown groups. For astrocyte-TurboID and neuron-TurboID in vivo studies, enriched pulldown samples were normalized by subtracting the difference from the mean of the log<sub>2</sub>-transformed summed protein intensity for 71 ribosomal subunit proteins present in all TurboID-positive pulldown samples. Proteins were filtered out if they had missing values in 2/3 samples across groups. Missing values were imputed from a normal distribution for each sample. Nomination of proteins as either ribosomal or RNA-binding was determined using the EMBL RBPbase (v0.23 alpha), a database that integrates high-throughput RNA-binding protein detection studies (<https://apps.embl.de/rbpbase/>).

### Data analysis and visualization

We utilized differential enrichment analysis, multidimensional scaling (MDS), and gene-set enrichment analysis (GSEA) to analyze transcriptomic and proteomic data from in vitro and in vivo studies. Data analysis and visualization were completed using R software (version 4.3.1) and Prism (GraphPad, version 10). Venn diagrams were generated using either a protein missingness (present in 2/3 samples) or a low-abundant transcript (average count value ≤ 10) filter per sample group. For transcriptomic studies, the DESeq2<sup>145</sup> Wald test was used to identify DEGs between groups. For proteomic studies, unpaired two-sided (equal variance assumption) *t*-test equivalent calculations using F-value from ANOVA run strictly with 2 comparison groups were performed with the parANOVA suite of functions in R (<https://github.com/edammer/parANOVA>) to identify DEPs between groups. DEGs and DEPs are presented as volcano plots. GSEA was applied using the GOpallel function (<https://github.com/edammer/GOpallel>) that scrapes monthly updates. GMT formatted GO from the Bader Lab website (<https://baderlab.org/>) as gene symbol lists submitted to one-tailed Fisher's exact test for GO enrichment of overlapping symbols in



gene lists of interest, including DEPs and DEGs. GO terms meeting Fisher's exact significance of  $p$ -value  $\leq 0.05$  (i.e., a  $z \geq 1.96$ ) were considered. Protein and gene input lists were significantly differentially biochemically enriched over global or background, in addition to gene lists that were different by treatment condition ( $p \leq 0.05$ ,  $\geq 1$ -log<sub>2</sub>-fold change). To assess average GC content, number of exons, and average transcript length of transcriptomic data, we used the *biomaRt* R package<sup>151,152</sup> and the BioMart/Ensembl databases<sup>153</sup>. To examine protein half-life information of proteomic data, we used values from Fornasiero et al.<sup>67</sup>.

## Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

## Data availability

The MS proteomics data have been deposited to the ProteomeXchange Consortium<sup>154</sup> via the PRIDE<sup>155</sup> partner repository (<http://www.proteomexchange.org/>) with the dataset identifier **PXD059818** and **PXD073954**. Transcriptomics data is available on the NIH Gene Expression Omnibus (GEO) repository with the dataset identifier **GSE287770**. Source data are provided as a Source Data File. The acutely isolated astrocytic and neuronal mouse brain cell data were acquired from Zhang et al.<sup>60</sup> and Sharma et al.<sup>61</sup>. Source data are provided with this paper.

## Code availability

Data cleanup and processing code for proteomic and transcriptomic studies is available at <https://github.com/cramelow/SPARO>.

## References

- Han, Y., Gao, S., Muegge, K., Zhang, W. & Zhou, B. Advanced applications of RNA sequencing and challenges. *Bioinform. Biol. Insights* **9**, BBI.S28991 (2015).
- Mallick, P. & Kuster, B. Proteomics: a pragmatic perspective. *Nat. Biotechnol.* **28**, 695–709 (2010).
- Liu, Y., Beyer, A. & Aebersold, R. On the dependency of cellular protein levels on mRNA abundance. *Cell* **165**, 535–550 (2016).
- Franks, A., Airolidi, E. & Slavov, N. Post-transcriptional regulation across human tissues. *PLoS Comput. Biol.* **13**, e1005535 (2017).
- Velázquez-Cruz, A. et al. Post-translational control of RNA-binding proteins and disease-related dysregulation. *Front. Mol. Biosci.* **8**, 658852 (2021).
- Saibil, H. Chaperone machines for protein folding, unfolding and disaggregation. *Nat. Rev. Mol. Cell Biol.* **14**, 630–642 (2013).
- Goswami, S. K. Regulation of gene expression in mammals. In *Post-Transcriptional Gene Regulation in Human Disease* 1–31. <https://doi.org/10.1016/B978-0-323-91305-8.00019-3> (Elsevier, 2022).
- Buccitelli, C. & Selbach, M. mRNAs, proteins and the emerging principles of gene expression control. *Nat. Rev. Genet.* **21**, 630–644 (2020).
- Brito Querido, J., Díaz-López, I. & Ramakrishnan, V. The molecular basis of translation initiation and its regulation in eukaryotes. *Nat. Rev. Mol. Cell Biol.* **25**, 168–186 (2024).
- Bauernfeind, A. L. & Babbitt, C. C. The predictive nature of transcript expression levels on protein expression in adult human brain. *BMC Genom.* **18**, 322 (2017).
- Cenik, C. et al. Integrative analysis of RNA, translation, and protein levels reveals distinct regulatory variation across humans. *Genome Res.* **25**, 1610–1621 (2015).
- Chen, G. et al. Discordant protein and mRNA expression in lung adenocarcinomas. *Mol. Cell. Proteom.* **1**, 304–313 (2002).
- Perl, K. et al. Reduced changes in protein compared to mRNA levels across non-proliferating tissues. *BMC Genom.* **18**, 305 (2017).
- Salovska, B. et al. Isoform-resolved correlation analysis between mRNA abundance regulation and protein level degradation. *Mol. Syst. Biol.* **16**, e9170 (2020).
- Schwanhäusser, B. et al. Global quantification of mammalian gene expression control. *Nature* **473**, 337–342 (2011).
- Specht, H. et al. Single-cell proteomic and transcriptomic analysis of macrophage heterogeneity using SCoPE2. *Genome Biol.* **22**, 50 (2021).
- Wang, D. et al. A deep proteome and transcriptome abundance atlas of 29 healthy human tissues. *Mol. Syst. Biol.* **15**, e8503 (2019).
- Wei, Y.-N. et al. Transcript and protein expression decoupling reveals RNA binding proteins and miRNAs as potential modulators of human aging. *Genome Biol.* **16**, 41 (2015).
- Marsh, S. E. et al. Dissection of artifactual and confounding glial signatures by single-cell sequencing of mouse and human brain. *Nat. Neurosci.* **25**, 306–316 (2022).
- Brewer, G. J. & Torricelli, J. R. Isolation and culture of adult neurons and neurospheres. *Nat. Protoc.* **2**, 1490–1498 (2007).
- Brewer, G. J. Isolation and culture of adult rat hippocampal neurons. *J. Neurosci. Methods* **71**, 143–155 (1997).
- Alvarez-Castelao, B. et al. Cell-type-specific metabolic labeling of nascent proteomes in vivo. *Nat. Biotechnol.* **35**, 1196–1201 (2017).
- Alvarez-Castelao, B., Schanzenbächer, C. T., Langer, J. D. & Schuman, E. M. Cell-type-specific metabolic labeling, detection and identification of nascent proteomes in vivo. *Nat. Protoc.* **14**, 556–575 (2019).
- Dieterich, D. C., Link, A. J., Graumann, J., Tirrell, D. A. & Schuman, E. M. Selective identification of newly synthesized proteins in mammalian cells using bioorthogonal noncanonical amino acid tagging (BONCAT). *Proc. Natl. Acad. Sci. USA* **103**, 9482–9487 (2006).
- Hodas, J. J. L. et al. Dopaminergic modulation of the hippocampal neuropil proteome identified by bioorthogonal noncanonical amino acid tagging (BONCAT). *Proteomics* **12**, 2464–2476 (2012).
- Landgraf, P., Antileo, E. R., Schuman, E. M. & Dieterich, D. C. BONCAT: metabolic labeling, click chemistry, and affinity purification of newly synthesized proteomes. In *Site-Specific Protein Labeling* (eds Gautier, A. & Hinner, M. J.) vol. 1266 199–215 (Springer New York, New York, NY, 2015).
- Link, A. J. et al. Discovery of aminoacyl-tRNA synthetase activity through cell-surface display of noncanonical amino acids. *Proc. Natl. Acad. Sci. USA* **103**, 10180–10185 (2006).
- Branon, T. C. et al. Efficient proximity labeling in living cells and organisms with TurboID. *Nat. Biotechnol.* **36**, 880–887 (2018).
- Cho, K. F. et al. Split-TurboID enables contact-dependent proximity labeling in cells. *Proc. Natl. Acad. Sci. USA* **117**, 12143–12154 (2020).
- Cho, K. F. et al. Proximity labeling in mammalian cells with TurboID and split-TurboID. *Nat. Protoc.* **15**, 3971–3999 (2020).
- Goettemoeller, A. M. et al. Entorhinal cortex vulnerability to human APP expression promotes hyperexcitability and tau pathology. *Nat. Commun.* **15**, 7918 (2024).
- Kim, D. I. et al. An improved smaller biotin ligase for BioID proximity labeling. *MBoC* **27**, 1188–1196 (2016).
- Kumar, P. et al. Native-state proteomics of Parvalbumin interneurons identifies unique molecular signatures and vulnerabilities to early Alzheimer's pathology. *Nat. Commun.* **15**, 2823 (2024).
- Mair, A., Xu, S.-L., Branon, T. C., Ting, A. Y. & Bergmann, D. C. Proximity labeling of protein complexes and cell-type-specific organellar proteomes in Arabidopsis enabled by TurboID. *eLife* **8**, e47864 (2019).

35. Soto, J. S. et al. Astrocyte–neuron subproteomes and obsessive–compulsive disorder mechanisms. *Nature* **616**, 764–773 (2023).
36. Rayaprolu, S. et al. Cell type-specific biotin labeling in vivo resolves regional neuronal and astrocyte proteomic differences in mouse brain. *Nat. Commun.* **13**, 2927 (2022).
37. Sunna, S. et al. Cellular proteomic profiling using proximity labeling by TurboID–NES in microglial and neuronal cell lines. *Mol. Cell. Proteom.* **22**, 100546 (2023).
38. Dumrongprechachan, V. et al. Cell-type and subcellular compartment-specific APEX2 proximity labeling reveals activity-dependent nuclear proteome dynamics in the striatum. *Nat. Commun.* **12**, 4855 (2021).
39. Hobson, B. D. et al. Subcellular proteomics of dopamine neurons in the mouse brain. *eLife* **11**, e70921 (2022).
40. Sanz, E. et al. Cell-type-specific isolation of ribosome-associated mRNA from complex tissues. *Proc. Natl. Acad. Sci. USA* **106**, 13939–13944 (2009).
41. Sanz, E., Bean, J. C., Carey, D. P., Quintana, A. & McKnight, G. S. RiboTag: ribosomal tagging strategy to analyze cell-type-specific mRNA expression in vivo. *Curr. Protoc. Neurosci.* **88**, e77 (2019).
42. Srinivasan, R. et al. New transgenic mouse lines for selectively targeting astrocytes and studying calcium signals in astrocyte processes in situ and in vivo. *Neuron* **92**, 1181–1195 (2016).
43. Kang, S. S. et al. Microglial translational profiling reveals a convergent APOE pathway from aging, amyloid, and tau. *J. Exp. Med.* **215**, 2235–2245 (2018).
44. Heiman, M. et al. A translational profiling approach for the molecular characterization of CNS cell types. *Cell* **135**, 738–748 (2008).
45. Heiman, M., Kulicke, R., Fenster, R. J., Greengard, P. & Heintz, N. Cell type-specific mRNA purification by translating ribosome affinity purification (TRAP). *Nat. Protoc.* **9**, 1282–1291 (2014).
46. Doyle, J. P. et al. Application of a translational profiling approach for the comparative analysis of CNS cell types. *Cell* **135**, 749–762 (2008).
47. Dougherty, J. D., Schmidt, E. F., Nakajima, M. & Heintz, N. Analytical approaches to RNA profiling data for the identification of genes enriched in specific cells. *Nucleic Acids Res.* **38**, 4218–4230 (2010).
48. Zhou, X., Liao, W.-J., Liao, J.-M., Liao, P. & Lu, H. Ribosomal proteins: functions beyond the ribosome. *J. Mol. Cell Biol.* **7**, 92–104 (2015).
49. Shi, Z. et al. Heterogeneous ribosomes preferentially translate distinct subpools of mRNAs genome-wide. *Mol. Cell* **67**, 71–83.e7 (2017).
50. Reimegård, J. et al. A combined approach for single-cell mRNA and intracellular protein expression analysis. *Commun. Biol.* **4**, 624 (2021).
51. Darmanis, S. et al. Simultaneous multiplexed measurement of RNA and proteins in single cells. *Cell Rep.* **14**, 380–389 (2016).
52. Assarsson, E. et al. Homogenous 96-Plex PEA immunoassay exhibiting high sensitivity, specificity, and excellent scalability. *PLoS ONE* **9**, e95192 (2014).
53. Stoekius, M. et al. Simultaneous epitope and transcriptome measurement in single cells. *Nat. Methods* **14**, 865–868 (2017).
54. Fazal, F. M. et al. Atlas of subcellular RNA localization revealed by APEX-Seq. *Cell* **178**, 473–490.e26 (2019).
55. Padrón, A., Iwasaki, S. & Ingolia, N. T. Proximity RNA labeling by APEX-Seq reveals the organization of translation initiation complexes and repressive RNA granules. *Mol. Cell* **75**, 875–887.e5 (2019).
56. Li, R., Zou, Z., Wang, W. & Zou, P. Metabolic incorporation of electron-rich ribonucleosides enhances APEX-seq for profiling spatially restricted nascent transcriptome. *Cell Chem. Biol.* **29**, 1218–1231.e8 (2022).
57. Wang, P. et al. Mapping spatial transcriptome with light-activated proximity-dependent RNA labeling. *Nat. Chem. Biol.* **15**, 1110–1119 (2019).
58. Bowen, C. A. et al. Proximity labeling proteomics reveals Kv1.3 potassium channel immune interactors in microglia. *Mol. Cell. Proteom.* **23**, 100809 (2024).
59. Madisen, L. et al. A robust and high-throughput Cre reporting and characterization system for the whole mouse brain. *Nat. Neurosci.* **13**, 133–140 (2010).
60. Zhang, Y. et al. An RNA-sequencing transcriptome and splicing database of glia, neurons, and vascular cells of the cerebral cortex. *J. Neurosci.* **34**, 11929–11947 (2014).
61. Sharma, K. et al. Cell type- and brain region-resolved mouse brain proteome. *Nat. Neurosci.* **18**, 1819–1831 (2015).
62. Maier, T., Güell, M. & Serrano, L. Correlation of mRNA and protein in complex biological samples. *FEBS Lett.* **583**, 3966–3973 (2009).
63. Vogel, C. & Marcotte, E. M. Insights into the regulation of protein abundance from proteomic and transcriptomic analyses. *Nat. Rev. Genet.* **13**, 227–232 (2012).
64. Seyfried, N. T. et al. A multi-network approach identifies protein-specific co-expression in asymptomatic and symptomatic Alzheimer’s disease. *Cell Syst.* **4**, 60–72.e4 (2017).
65. Johnson, E. C. B. et al. Large-scale proteomic analysis of Alzheimer’s disease brain and cerebrospinal fluid reveals early changes in energy metabolism associated with microglia and astrocyte activation. *Nat. Med.* **26**, 769–780 (2020).
66. Johnson, E. C. B. et al. Large-scale deep multi-layer analysis of Alzheimer’s disease brain reveals strong proteomic disease-related changes not observed at the RNA level. *Nat. Neurosci.* **25**, 213–225 (2022).
67. Fornasiero, E. F. et al. Precisely measured protein lifetimes in the mouse brain reveal differences across tissues and subcellular fractions. *Nat. Commun.* **9**, 4230 (2018).
68. Chapman-Smith, A. & Cronan, J. E. Jr Molecular biology of biotin attachment to proteins. *J. Nutr.* **129**, 477S–484S (1999).
69. Tong, L. Structure and function of biotin-dependent carboxylases. *Cell. Mol. Life Sci.* **70**, 863–891 (2013).
70. Cohen, M. J., Chirico, W. J. & Lipke, P. N. Through the back door: unconventional protein secretion. *Cell Surf.* **6**, 100045 (2020).
71. Dogterom, M. & Koenderink, G. H. Actin–microtubule crosstalk in cell biology. *Nat. Rev. Mol. Cell Biol.* **20**, 38–54 (2019).
72. Eom, T.-Y. et al. Direct visualization of microtubules using a genetic tool to analyse radial progenitor–astrocyte continuum in brain. *Nat. Commun.* **2**, 446 (2011).
73. Pedrola, L. et al. GDAP1, the protein causing Charcot–Marie–Tooth disease type 4A, is expressed in neurons and is associated with mitochondria. *Hum. Mol. Genet.* **14**, 1087–1094 (2005).
74. Miressi, F. et al. GDAP1 involvement in mitochondrial function and oxidative stress, investigated in a Charcot–Marie–Tooth model of hiPSCs-derived motor neurons. *Biomedicine* **9**, 945 (2021).
75. León, M. et al. Rapid degeneration of iPSC-derived motor neurons lacking Gdap1 engages a mitochondrial-sustained innate immune response. *Cell Death Discov.* **9**, 217 (2023).
76. Mederos, S., González-Arias, C. & Perea, G. Astrocyte–neuron networks: a multilane highway of signaling for homeostatic brain function. *Front. Synaptic Neurosci.* **10**, 45 (2018).
77. Mader, S. & Brimberg, L. Aquaporin-4 water channel in the brain and its implication for health and disease. *Cells* **8**, 90 (2019).
78. Lopes, I., Altab, G., Raina, P. & De Magalhães, J. P. Gene size matters: an analysis of gene length in the human genome. *Front. Genet.* **12**, 559998 (2021).
79. Zhao, D. et al. Analysis of ribosome-associated mRNAs in rice reveals the importance of transcript size and GC content in translation. *G3* **7**, 203–219 (2017).

80. Yao, Z. et al. A high-resolution transcriptomic and spatial atlas of cell types in the whole mouse brain. *Nature* **624**, 317–332 (2023).
81. Breckels, L. M. et al. Advances in spatial proteomics: mapping proteome architecture from protein complexes to subcellular localizations. *Cell Chem. Biol.* **31**, 1665–1687 (2024).
82. Bhushan, V. & Nita-Lazar, A. Recent advancements in subcellular proteomics: growing impact of organellar protein niches on the understanding of cell biology. *J. Proteome Res.* **23**, 2700–2722 (2024).
83. Liu, J., Jang, J. Y., Pirooznia, M., Liu, S. & Finkel, T. The secretome mouse provides a genetic platform to delineate tissue-specific in vivo secretion. *Proc. Natl. Acad. Sci. USA* **118**, e2005134118 (2021).
84. Wei, W. et al. Cell type-selective secretome profiling in vivo. *Nat. Chem. Biol.* **17**, 326–334 (2021).
85. Vasek, M. J. et al. Local translation in microglial processes is required for efficient phagocytosis. *Nat. Neurosci.* **26**, 1185–1195 (2023).
86. Pilaz, L.-J., Lennox, A. L., Rouanet, J. P. & Silver, D. L. Dynamic mRNA transport and local translation in radial glial progenitors of the developing brain. *Curr. Biol.* **26**, 3383–3392 (2016).
87. Mazaré, N. et al. Local translation in perisynaptic astrocytic processes is specific and changes after fear conditioning. *Cell Rep.* **32**, 108076 (2020).
88. Hobson, B. D. et al. Subcellular and regional localization of mRNA translation in midbrain dopamine neurons. *Cell Rep.* **38**, 110208 (2022).
89. Herbert, A. L. et al. Dynein/dynactin is necessary for anterograde transport of Mbp mRNA in oligodendrocytes and for myelination in vivo. *Proc. Natl. Acad. Sci. USA* **114**, E9153–E9162 (2017).
90. Boulay, A.-C. et al. Translation in astrocyte distal processes sets molecular heterogeneity at the gliovascular interface. *Cell Discov.* **3**, 17005 (2017).
91. Sakers, K. et al. Astrocytes locally translate transcripts in their peripheral processes. *Proc. Natl. Acad. Sci. USA* **114**, E3830–E3838 (2017).
92. Hafner, A.-S., Donlin-Asp, P. G., Leitch, B., Herzog, E. & Schuman, E. M. Local protein synthesis is a ubiquitous feature of neuronal pre- and postsynaptic compartments. *Science* **364**, eaau3644 (2019).
93. Biever, A. et al. Monosomes actively translate synaptic mRNAs in neuronal processes. *Science* **367**, eaay4991 (2020).
94. Ouwenga, R. et al. Transcriptomic analysis of ribosome-bound mRNA in cortical neurites in vivo. *J. Neurosci.* **37**, 8688–8705 (2017).
95. Hamdan, H. et al. Mapping axon initial segment structure and function by multiplexed proximity biotinylation. *Nat. Commun.* **11**, 100 (2020).
96. Zhang, W. et al. Immunoproximity biotinylation reveals the axon initial segment proteome. *Nat. Commun.* **14**, 8201 (2023).
97. Li, Y., Kanao, E., Yamano, T., Ishihama, Y. & Imami, K. TurboID-EV: proteomic mapping of recipient cellular proteins proximal to small extracellular vesicles. *Anal. Chem.* **95**, 14159–14164 (2023).
98. Mattick, J. S. et al. Long non-coding RNAs: definitions, functions, challenges and recommendations. *Nat. Rev. Mol. Cell Biol.* **24**, 430–447 (2023).
99. O'Brien, J., Hayder, H., Zayed, Y. & Peng, C. Overview of MicroRNA biogenesis, mechanisms of actions, and circulation. *Front. Endocrinol.* **9**, 402 (2018).
100. Ardekani, A. M. & Naeini, M. M. The role of MicroRNAs in human diseases. *Avicenna J. Med. Biotechnol.* **2**, 161–179 (2010).
101. Rupaimoole, R. & Slack, F. J. MicroRNA therapeutics: towards a new era for the management of cancer and other diseases. *Nat. Rev. Drug Discov.* **16**, 203–222 (2017).
102. Qian, Z. et al. MiR-760 exerts a critical regulatory role in glioma proliferation, migration, and invasion by modulating MMP2 expression. *J. Cancer* **15**, 3076–3084 (2024).
103. He, M. miRNA tagging and affinity-purification (miRAP). *Bio Protoc.* **2**, e265 (2012).
104. Bhattarai, N., Cao, B., Zeng, S. X. & Lu, H. Ribosomal profiling by gradient fractionation of cell lysates. In *RNA-Protein Complexes and Interactions* (ed. Lin, R.-J.) vol. 2666 149–155 (Springer US, New York, NY, 2023).
105. Chan, L. Y., Mugler, C. F., Heinrich, S., Vallotton, P. & Weis, K. Non-invasive measurement of mRNA decay reveals translation initiation as the major determinant of mRNA stability. *eLife* **7**, e32536 (2018).
106. Wang, C. & Liu, H. Factors influencing degradation kinetics of mRNAs and half-lives of microRNAs, circRNAs, lncRNAs in blood in vitro using quantitative PCR. *Sci. Rep.* **12**, 7259 (2022).
107. Lugowski, A., Nicholson, B. & Rissland, O. S. Determining mRNA half-lives on a transcriptome-wide scale. *Methods* **137**, 90–98 (2018).
108. Xu, N., Loflin, P., Chen, C.-Y. A. & Shyu, A.-B. A broader role for AU-rich element-mediated mRNA turnover revealed by a new transcriptional pulse strategy. *Nucleic Acids Res.* **26**, 558–565 (1998).
109. Loflin, P. T., Chen, C.-Y. A., Xu, N. & Shyu, A.-B. Transcriptional pulsing approaches for analysis of mRNA turnover in mammalian cells. *Methods* **17**, 11–20 (1999).
110. Chen, C. A., Ezzeddine, N. & Shyu, A. Chapter 17 Messenger RNA half-life measurements in mammalian cells. In *Methods in Enzymology* vol. 448 335–357 (Elsevier, 2008).
111. Friedman, R. C., Farh, K. K.-H., Burge, C. B. & Bartel, D. P. Most mammalian mRNAs are conserved targets of microRNAs. *Genome Res.* **19**, 92–105 (2009).
112. Wu, J., Bonsra, A. N. & Du, G. pSM155 and pSM30 vectors for miRNA and shRNA expression. In *siRNA and miRNA Gene Silencing: From Bench to Bedside* 1–15 (Springer, 2009).
113. Wu, P., Wilmarth, M. A., Zhang, F. & Du, G. miRNA and shRNA expression vectors based on mRNA and miRNA processing. In *MicroRNA Protocols* 195–207 (Springer, 2013).
114. Shu, Y. et al. A simplified system to express circularized inhibitors of miRNA for stable and potent suppression of miRNA functions. *Mol. Ther. Nucleic Acids* **13**, 556–567 (2018).
115. Fan, J. et al. A simplified system for the effective expression and delivery of functional mature microRNAs in mammalian cells. *Cancer Gene Ther.* **27**, 424–437 (2020).
116. Bak, R. O., Hollensen, A. K. & Mikkelsen, J. G. Managing microRNAs with vector-encoded decoy-type inhibitors. *Mol. Ther.* **21**, 1478–1485 (2013).
117. Bak, R. O. & Mikkelsen, J. G. miRNA sponges: soaking up miRNAs for regulation of gene expression. *Wiley Interdiscip. Rev. RNA* **5**, 317–333 (2014).
118. Leestemaker, Y. & Ovaa, H. Tools to investigate the ubiquitin proteasome system. *Drug Discov. Today. Technol.* **26**, 25–31 (2017).
119. Lu, X. et al. Designed semisynthetic protein inhibitors of Ub/Ubl E1 activating enzymes. *J. Am. Chem. Soc.* **132**, 1748–1749 (2010).
120. An, H. & Statsyuk, A. V. Development of activity-based probes for ubiquitin and ubiquitin-like protein signaling pathways. *J. Am. Chem. Soc.* **135**, 16948–16962 (2013).
121. Mund, T., Lewis, M. J., Maslen, S. & Pelham, H. R. Peptide and small molecule inhibitors of HECT-type ubiquitin ligases. *Proc. Natl. Acad. Sci. USA* **111**, 16736–16741 (2014).
122. An, H. & Statsyuk, A. V. Facile synthesis of covalent probes to capture enzymatic intermediates during E1 enzyme catalysis. *Chem. Commun.* **52**, 2477–2480 (2016).



123. Price, J. C. et al. Measurement of human plasma proteome dynamics with  $2\text{H}_2\text{O}$  and liquid chromatography tandem mass spectrometry. *Anal. Biochem.* **420**, 73–83 (2012).
124. Zhang, Y. et al. Proteome scale turnover analysis in live animals using stable isotope metabolic labeling. *Anal. Chem.* **83**, 1665–1672 (2011).
125. Toyama, B. H. et al. Identification of long-lived proteins reveals exceptional stability of essential cellular structures. *Cell* **154**, 971–982 (2013).
126. Swovick, K. et al. Cross-species comparison of proteome turnover kinetics. *Mol. Cell. Proteom.* **17**, 580–591 (2018).
127. Savas, J. N., Park, S. K. & Yates III, J. R. Proteomic analysis of protein turnover by metabolic whole rodent pulse-chase isotopic labeling and shotgun mass spectrometry analysis. In *Quantitative Proteomics by Mass Spectrometry* 293–304 (Springer, 2016).
128. Rolfs, Z. et al. An atlas of protein turnover rates in mouse tissues. *Nat. Commun.* **12**, 6778 (2021).
129. Price, J. C., Guan, S., Burlingame, A., Prusiner, S. B. & Ghaemmaghami, S. Analysis of proteome dynamics in the mouse brain. *Proc. Natl. Acad. Sci. USA* **107**, 14508–14513 (2010).
130. Pratt, J. M. et al. Dynamics of protein turnover, a missing dimension in proteomics. *Mol. Cell. Proteom.* **1**, 579–591 (2002).
131. Mathieson, T. et al. Systematic analysis of protein turnover in primary cells. *Nat. Commun.* **9**, 689 (2018).
132. Lau, E. et al. A large dataset of protein dynamics in the mammalian heart proteome. *Sci. Data* **3**, 1–15 (2016).
133. Guan, S., Price, J. C., Ghaemmaghami, S., Prusiner, S. B. & Burlingame, A. L. Compartment modeling for mammalian protein turnover studies by stable isotope metabolic labeling. *Anal. Chem.* **84**, 4014–4021 (2012).
134. Doherty, M. K., Hammond, D. E., Clague, M. J., Gaskell, S. J. & Beynon, R. J. Turnover of the human proteome: determination of protein intracellular stability by dynamic SILAC. *J. Proteome Res.* **8**, 104–112 (2009).
135. Cambridge, S. B. et al. Systems-wide proteomic analysis in mammalian cells reveals conserved, functional protein turnover. *J. Proteome Res.* **10**, 5275–5284 (2011).
136. Ingaramo, M. & Beckett, D. Biotinylation, a post-translational modification controlled by the rate of protein-protein association. *J. Biol. Chem.* **286**, 13071–13078 (2011).
137. Kampmeyer, C. et al. Lysine deserts prevent adventitious ubiquitylation of ubiquitin-proteasome components. *Cell. Mol. Life Sci.* **80**, 143 (2023).
138. Shukri, A. H., Lukinović, V., Charih, F. & Biggar, K. K. Unraveling the battle for lysine: A review of the competition among post-translational modifications. *Biochim. Biophys. Acta Gene Regul. Mech.* **1866**, 194990 (2023).
139. Sakurai-Yageta, M. & Suzuki, Y. Molecular mechanisms of biotin in modulating inflammatory diseases. *Nutrients* **16**, 2444 (2024).
140. Pronobis, M. I., Deutch, N. & Peifer, M. The Miraprep: a protocol that uses a Miniprep kit and provides Maxiprep yields. *PLoS ONE* **11**, e0160509 (2016).
141. Andrews, S. *FastQC: A Quality Control Tool for High Throughput Sequence Data*. <https://www.bioinformatics.babraham.ac.uk/projects/fastqc/> (2010).
142. Dobin, A. et al. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29**, 15–21 (2013).
143. Langmead, B. & Salzberg, S. L. Fast gapped-read alignment with Bowtie 2. *Nat. Methods* **9**, 357–359 (2012).
144. Liao, Y., Smyth, G. K. & Shi, W. featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics* **30**, 923–930 (2014).
145. Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **15**, 550 (2014).
146. Hebert, A. S. et al. Comprehensive Single-shot Proteomics with FAIMS on a hybrid orbitrap mass spectrometer. *Anal. Chem.* **90**, 9529–9537 (2018).
147. Cox, J. et al. Andromeda: a peptide search engine integrated into the MaxQuant environment. *J. Proteome Res.* **10**, 1794–1805 (2011).
148. Cox, J. & Mann, M. MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. *Nat. Biotechnol.* **26**, 1367–1372 (2008).
149. Bruderer, R. et al. Extending the limits of quantitative proteome profiling with data-independent acquisition and application to acetaminophen-treated three-dimensional liver microtissues. *Mol. Cell. Proteom.* **14**, 1400–1410 (2015).
150. Trautwig, A. N. et al. Network analysis of the cerebrospinal fluid proteome reveals shared and unique differences between sporadic and familial forms of amyotrophic lateral sclerosis. *Mol. Neurodegener.* **20**, 58 (2025).
151. Durinck, S. et al. BioMart and Bioconductor: a powerful link between biological databases and microarray data analysis. *Bioinformatics* **21**, 3439–3440 (2005).
152. Durinck, S., Spellman, P. T., Birney, E. & Huber, W. Mapping identifiers for the integration of genomic datasets with the R/Bioconductor package biomaRt. *Nat. Protoc.* **4**, 1184–1191 (2009).
153. Hubbard, T. J. P. et al. Ensembl 2009. *Nucleic Acids Res.* **37**, D690–D697 (2009).
154. Deutsch, E. W. et al. The ProteomeXchange consortium at 10 years: 2023 update. *Nucleic Acids Res.* **51**, D1539–D1548 (2023).
155. Perez-Riverol, Y. et al. The PRIDE database resources in 2022: a hub for mass spectrometry-based proteomics evidences. *Nucleic Acids Res.* **50**, D543–D552 (2022).

## Acknowledgements

Research reported in this publication was supported by the National Institute on Aging and the National Institute of Mental Health of the National Institutes of Health: 1F31AG079597-01A1 (C.C.R.), R01AG075820 (N.T.S., S.R.), R01AG071587 (S.R.), and R01MH125956 (S.A.S.). Research was also supported in part by the Emory Integrated Proteomics Core (EIPC), which is subsidized by the Emory University School of Medicine and is one of the Emory Integrated Core Facilities. Additional support was provided by the Georgia Clinical & Translational Science Alliance of the National Institutes of Health under Award Number UL1TR002378 (EIPC). Additional support was provided by the National Center for Advancing Translational Sciences of the National Institutes of Health under Award Number UL1TR000454. Thank you to the members of the Rangaraju, Sloan, and Seyfried labs for their supportive scientific discussions that strengthened this manuscript, including Sarah Shapley, Christine Bowen, Sydney Sunna, Juliet Santiago, Caroline Watson, Adam Trautwig, Ananth Shantaraman, Eddie Fox, Anson Sing, Emily Hill, Melissa Cadena, Alexia King, Nardos Kebede, Tarun Bhatia and Aditya Natu. We also wish to thank Victor Faundez and Yue Feng from C.C.R.'s dissertation committee.

## Author contributions

Conceptualization: C.C.R., S.A.S., S.R., Methodology: C.C.R., E.B.D., H.X., L.C., P.K., C.E.G., M.M.S., D.K., R.S.N., S.M., R.K., W.E.J., Q.G., P.B., D.M.D., N.T.S., S.A.S., S.R., Investigation: C.C.R., E.B.D., N.T.S., S.A.S., S.R., Writing-Original draft: C.C.R., S.A.S., S.R., Writing-Review and Editing: C.C.R., S.A.S., S.R., Methodology: C.C.R., E.B.D., H.X., L.C., P.K., C.E.G., M.M.S., D.K., R.S.N., S.M., R.K., W.E.J., Q.G., P.B., D.M.D., N.T.S., S.A.S., S.R., Funding acquisition: C.C.R., N.T.S., S.A.S., S.R., Resources: N.T.S., S.A.S., S.R., Supervision: S.A.S., S.R.

## Competing interests

N.T.S. and D.M.D. are co-founders of Emtherapro and Arc Proteomics. N.T.S. is the co-founder of Stitch Rx. The remaining authors declare no competing interests.

## Additional information

**Supplementary information** The online version contains supplementary material available at <https://doi.org/10.1038/s41467-026-71098-4>.

**Correspondence** and requests for materials should be addressed to Steven A. Sloan or Srikant Rangaraju.

**Peer review information** *Nature Communications* thanks Tess Branon and the other anonymous reviewer(s) for their contribution to the peer review of this work. A peer review file is available.

**Reprints and permissions information** is available at <http://www.nature.com/reprints>

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

**Open Access** This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

© The Author(s) 2026