Integrated transcriptome and proteome analysis in human brain organoids reveals posttranscriptional regulation of ribosomal genes

Jaydeep Sidhaye[1,a], Philipp Trepte[1,a], Natalie Sepke[a], Maria Novatchkova[a], Michael Schutzbier[a], Gerhard Dürnberger[a], Karl Mechtler[a], and Jürgen A. Knoblich[a,b,✉]

[1]Equal contribution
[a]Institute of Molecular Biotechnology of the Austrian Academy of Sciences (IMBA), Vienna BioCenter (VBC), Dr. Bohr-Gasse 3, 1030, Vienna, Austria
[b]Department of Neurology, Medical University of Vienna, Vienna, Austria

During development of the human cerebral cortex, multipotent neural progenitors generate excitatory neurons and glial cells. Investigations of the transcriptome and epigenome have revealed important gene regulatory networks underlying this crucial developmental event. However, the post-transcriptional control of gene expression and protein abundance during human corticogenesis remains poorly understood. We addressed this issue by using human telencephalic brain organoids grown using a dual reporter cell line to isolate neural progenitors and neurons and performed cell class and developmental stage-specific transcriptome and proteome analysis. Integrating the two datasets revealed modules of gene expression during human corticogenesis. Investigation of one such module uncovered mTOR-mediated regulation of translation of the 5'TOP element-enriched translation machinery in early progenitor cells. We show that in early progenitors partial inhibition of the translation of ribosomal genes prevents precocious translation of differentiation markers. Overall, our multiomics approach reveals novel posttranscriptional regulatory mechanisms crucial for the fidelity of cortical development.

Keywords: Gene regulation | brain organoids | neurodevelopment | regulation of translation | mTOR signaling

Correspondence: ***juergen.knoblich@imba.oeaw.ac.at***

Introduction

The development of the human cerebral cortex is a highly elaborate process occurring over several months. During this process multipotent neural progenitors give rise initially to excitatory neurons of the different layers of the cerebral cortex and later to glial cells. Studies using mouse model systems, human fetal samples as well as brain organoids have revealed elaborate spatiotemporal gene regulatory networks that orchestrate mammalian corticogenesis (Cadwell et al., 2019; Greig et al., 2013; Shibata et al., 2015; Vaid and Huttner, 2020). However, an emerging theme from mouse corticogenesis studies highlights the additional role of post-transcriptional gene regulatory mechanisms, such as alternative splicing and translational repression, in determining progenitor cell fate and neuronal migration (Hoye and Silver, 2020; Lennox et al., 2017; Zahr et al., 2019). Whether similar mechanisms play a role in human neurodevelopment remains elusive. Hence, investigating the post-transcriptional control of gene expression and protein abundance is crucial to comprehensively understand gene regulation during human neurodevelopment.

Transcriptome analyses have placed a lot of importance on transcript abundance, however, the relationship between transcript and protein abundance remains elusive. While previous studies have investigated post-transcriptional regulation of specific genes or the impact of loss of key RNA-binding proteins (Hoye and Silver, 2020), proteome-scale analyses of gene regulation during human corticogenesis have been lacking. Currently available proteome data from hiPSC-derived neural progenitors and neurons (Djuric et al., 2017; Varderidou-Minasian et al., 2020), cerebral organoids (McClure-Begley et al., 2018; Melliou et al., 2022; Nascimento et al., 2019), and the fetal brain (Djuric et al., 2017; Kim et al., 2014; Melliou et al., 2022) are an important step towards establishing a human neurodevelopmental proteomic repertoire. However, these studies suffer from some of the following limitations: a) due to the inherent diversity of cell types present during corticogenesis, bulk tissue data do not provide cell type-specific information, b) cell sorting strategies result in datasets that suffer from low number of successfully detected proteins, and c) importantly, a direct comparison of the transcript and protein expression to understand posttranscriptional gene regulation has been missing.

We addressed these issues by using a dual reporter cell line to separate progenitors and neurons from the human brain organoid tissue and performed cell class- and developmental stage-specific transcriptome and proteome analysis. We integrated the two datasets to identify gene expression modules active during neural progenitor proliferation and subsequent neurogenesis in the developing cortical tissue. This dataset is available as a resource for the community in the form of a Shiny app (https://organoid.multiomics.vbc.ac.at). We followed up on one gene expression module in detail and found that the genes of the core translational machinery, containing the ‘5’ Terminal Oligo Pyrimidine’ (5’TOP) motif in their 5’UTR, are coregulated post-transcriptionally during neurodevelopment. We show that in contrast to neurons, translation of 5’TOP transcripts is partially inhibited in early progenitors, resulting in discordant RNA-protein relationship between the two cell classes. This regulation of the 5’TOP element-enriched translational machinery is due to lower mTOR activity in early progenitors and is crucial for the fidelity of cortical development.

Results

RNA-protein multiomics of neural progenitors and neurons isolated from telencephalic organoids.

To investigate progenitor and neuron-specific gene regulatory programs, we generated a dual reporter H9 human embryonic stem cell (hESC) line that enables sorting of neural progenitors and neurons (Figure 1A). To label neural progenitors, we replaced the stop codon of one allele of the pan-neural progenitor marker gene SOX2 with P2A-EGFP (Figure 1-figure supplement 1A). On the other hand, to label neurons, we inserted dTomato driven by human Synapsin1 (hSYN1) promoter in the AAVS1 safe harbor locus (Figure 1 -figure supplement 1B). Using this reporter line, we generated brain organoids enriched in dorsal telencephalic tissue using a previously published organoid culture protocol (Esk et al., 2020) (Figure 1 -figure supplement 1C) to recapitulate the key stages of corticogenesis, when multipotent progenitors give rise to diverse excitatory neurons of distinct layer identities. Immunostaining for EGFP and dTomato confirmed expression of the fluorophores in the progenitor and neuronal zones, respectively (Figure 1B). Furthermore, analysis of sparsely labeled organoids set up using 5% dual report hESCs and 95% non-reporter, wild-type hESCs confirmed expression of EGFP and dTomato in the progenitor-rich ventricular zone (VZ) (SOX2), and the surrounding neuron-rich region (NeuN), respectively (Figure 1-figure supplement 1D-F).

To study progenitor and neuron-specific gene regulatory programs, we aimed to perform whole transcriptome and proteome analysis of the respective cell classes. Dissociation of the organoids followed by fluorescence activated cell sorting (FACS) confirmed the occurrence of EGFP positive (EGFP+) and dTomato positive (dTomato+) cell populations across different stages of organoid development (Days 40, 60, 90 and 120) (Figure 1-figure supplement 2A). Additionally, an EGFP/dTomato double positive cell population was detected (Figure 1-figure supplement 2A). RNA-seq-based whole transcriptome analysis of these sorted cell populations and subsequent principal component analysis (PCA) revealed that the double positive population at day 40 and day 60 matched a transition stage between progenitors and neurons, while at day 90 and day 120 it was more similar to EGFP+ cells (Figure 2-figure supplement 2B). However, we focused on cells exclusively positive for single fluorophores as stringent criteria for bona fide progenitors (SOX2::EGFP+) and neurons (hSYN1::dTomato+) and did not include the double positive cells for further analysis. PCA showed that SOX2::EGFP+ cells and hSYN1::dTomato+ cells clustered separately and showed transcriptomic signatures of progenitors and neurons, respectively (PC1, Figure 1-figure supplement 3A,B). These samples further separated according to the organoid developmental stages with little batch effect (PC2, Figure 1-figure supplement 3A,B).

For proteome analysis, proteins from the cells sorted from the same organoid culture batches and time points as used for RNA-seq were tandem mass tag (TMT) labeled and combined for each batch separately. In total, 6,740 proteins were detected by mass spectrometry of the sorted cell populations. Initial analysis confirmed that similar to the transcriptome, the proteome of sorted neurons (hSYN1::dTomato+) and progenitor cells (SOX2::EGFP+) segregated according to the cell class and age (Figure 1-figure supplement 3C,D).

Transcriptome and proteome analysis (Figure 1C-E) confirmed that EGFP and dTomato expression matched the expression of endogenous SOX2 and SYN1 (Figure 1C, D). Furthermore, this analysis revealed that the SOX2::EGFP+ population encompassed all different classes of cortical progenitors including ventricular radial glia (expression of SOX1, PAX6), intermediate progenitors (EOMES) as well as outer radial glia (HOPX, MOXD1, TNC). The expression patterns of these markers followed the expected temporal trajectories of cortical development. While the markers of ventricular radial glia were expressed at all stages, markers of intermediate progenitors and outer radial glia showed enrichment at day 60 and day 90, respectively (Figure 1C). Similarly, the hSYN1::dTomato+ population included excitatory neurons of all different cortical layers and followed the expected temporal expression patterns (Figure 1D), with deep-layer markers (TBR1, BCL11B) being expressed before upper-layer markers (SATB2, CUX2). The enrichment of dorsal telencephalic fate was confirmed by the expression of above stated markers of cortical progenitors and neurons along with high expression of FOXG1 and EMX2 (Figure 1E) and undetectable levels of ventral markers (NKX2.1, LHX6) (TPM values below 1). Thus, we generated a time-resolved transcriptome and proteome dataset of neural progenitors (SOX2::EGFP+) and neurons (hSYN1::dTomato+) sorted from telencephalic organoids (Figure 1-source data 1, 2).

Integrated transcriptome and proteome analysis reveals gene expression modules active during neurodevelopment.

Next, we examined the general degree of correlation between mRNA and protein levels. We detected mean TPM values greater than 0 and at least one peptide for 6,714 genes. Similar to a previous report (Schwanhäusser et al., 2011), we observed only a limited correlation between absolute mRNA and protein levels, which seemed comparable between cell classes and developmental timepoints (Figure 1-figure

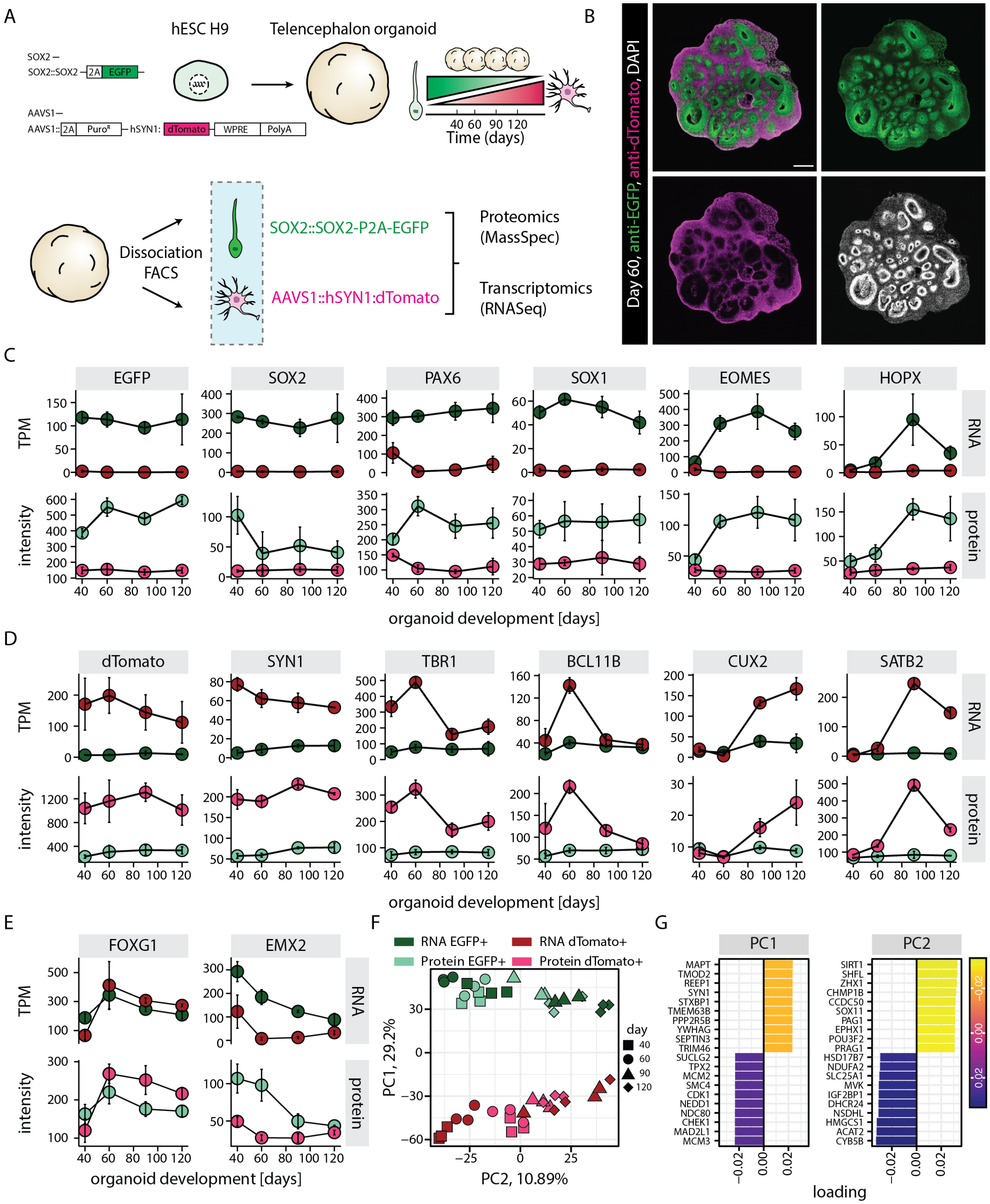


Fig. 1. RNA-protein multi-omics of progenitors and neurons in human brain organoids.
A) Schematic representation of the experimental design. A dual reporter line was generated in H9 hESC background. Cells from brain organoids grown using this reporter line were sorted and collected by FACS in EGFP and dTomato positive fractions. Gene expression of the sorted cell populations were analyzed by RNA-seq and Mass-spec. B) Confocal scan of a dual reporter organoid cryosection (day 60) stained with anti-EGFP, anti-dTomato antibodies and DAPI. Scale bar = 500 µm. C,D,E) Plots showing RNA (top) and protein (bottom) abundance of key progenitor-specific markers (C), neuronal markers (D) and dorsal telencephalic fate markers (E) in SOX2::EGFP+ (green) and hSYN1::dTomato+ positive (red) cells at different stages of organoid development. RNA abundance is measured as transcript per million (TPM). F,G) Principal component analysis of z-score normalized integrated transcriptome and proteome of progenitors (EGFP+, green) and neurons (dTomato+, red) at different stages of organoid development with (G) loading scores for PC1 and PC2 top contributing genes.

supplement 4A). Given the very different measurement units of RNA-seq and Mass-spec, a PCA on the combined datasets revealed that almost 80% of the variation was explained by the omics method (PC1, Figure 1-figure supplement 4B) and not by cell class or developmental time point highlighting the importance of data normalization. Furthermore, to identify groups of genes that follow similar relative gene expression profiles at RNA and protein level in progenitors and neurons over time, we combined the two data sets upon rescaling using gene-by-gene z-score normalization (Figure 1-source data 3). PCA of the combined datasets confirmed that the scaled data grouped by cell classes and developmental age and not by the omics method (Figure 1F,G). Using the z-score scaled dataset we first fit a global regression model to identify genes differentially expressed between the two cell classes on RNA and/or protein level. In a second regression step, genes showing significant temporal profile differences for RNA and protein expression in progenitors and neurons were identified (Nueda et al., 2014). Out of the 6,714 genes, we could successfully fit 5,978 genes, 3,668 of which showed significant temporal expression changes between cell classes on RNA and/or protein levels. Finally, we performed hierarchical clustering to identify genes with underlying common trends of RNA-protein expression. To estimate an optimal number of biologically meaningful clusters, we used the “within cluster sums of squares”, “average silhouette” and “gap statistics” methods (Figure 1-figure supplement 4C). In the end, a total of 3,368 genes were classified into 9 modules (Figure 2A,B). The remaining 2,310 genes were considered as “not clustered” and 736 genes for which the first regression fit failed as “not fit” (Figure 1-figure supplement 4D,E and Figure 1-source data 3). These genes lacked specific expression patterns and were enriched in house-keeping genes, including core components of the proteasome, spliceosome and RNA polymerase (Figure 1-figure supplement 4F). Thus, we were able to identify modules of genes showing distinct temporal and cell class-specific expression patterns (Figure 1-source data 3).

Functional enrichment analysis helped to characterize the members of each module (Figure 2C, Figure 2-figure supplement 1A,B). Furthermore, the average global trends of relative expression of RNA and proteins of each module helped to postulate potential regulatory mechanisms for each module (Figure 2B). Additionally, we examined the trend of a few individual member genes, to verify if the global trends are reflected for individual genes (Figures 1C-E and Figure 2-figure supplement 2A). Modules 1, 2, 7, 8 and 9 showed very strong cell class-specific expression patterns matching the sorting strategy for progenitors and neurons (defined here as “C” modules) (Figure 2B). Modules 1 and 2 showed higher relative expression in the progenitors than neurons across all time points and were enriched in DNA replication and cell activation-related genes, respectively (Figure 2C and Figure 2-figure supplement 1A,B). For module 2, progenitor-enriched expression increased with time, in line with the observation that many outer radial glia markers were members of this cluster. Genes in modules 7, 8 and 9 showed higher expression in neurons than progenitors across all timepoints. Module 7 was enriched in genes related to mitochondrial respiration (Figure 2C, Figure 2-figure supplement 1A,B), which is in agreement with the metabolic shift from glycolysis to oxidative phosphorylation observed during neurogenesis (Iwata and Vanderhaeghen, 2021). Modules 8 and 9 were enriched in axonal and pre- and postsynaptic genes (Figure 2C and Figure 2-figure supplement 1A,B). Importantly, in “C” modules, global relative expression of RNA and protein followed largely similar trends, indicating that the regulation of these genes might occur via cell class-specific transcription or transcript retention and decay.

Interestingly, the global profiles of relative RNA and protein expression did not follow similar trends for other modules, which instead showed a temporal or cell class-specific discrepancy. Modules 4 and 6 showed relative gene expression changing with time (“T” modules) (Figures 2B). In Module 4, expression decreased with time and the average RNA trends for these genes were similar for progenitors and neurons. However, progenitors expressed relatively higher amounts of protein. Such a trend might be explained by less effective translation or reduced protein stability of these genes in neurons. This module includes known early fate specification genes such as FEZF2, as well as genes related to glycolysis, a process more prominent in progenitors (Iwata and Vanderhaeghen, 2021). In contrast, module 6 genes were expressed more at later stages. The average RNA trends of these genes were similar for progenitors and neurons, but protein levels were higher in neurons (Figure 2B), suggesting less effective translation or reduced protein stability in progenitors over neurons. This cluster includes upper-layer fate regulators such as SATB2 and CUX2.

The remaining modules 3 and 5 showed more ambiguous patterns of expression (“A” modules) (Figures 2B). Genes from module 3 showed considerably higher transcript levels in progenitors than neurons, although protein levels between the two cell classes were more similar. This observation suggested that translation of these genes is either less effective in progenitors or enhanced in neurons, or that protein stability is different between the two cell classes. Gene enrichment analysis suggested enrichment of translational machinery proteins including those of ribosomes (Figure 2C, Figure 2-figure supplement 1A,B). Lastly, module 5 genes showed opposite patterns for RNA and protein relation in progenitors and neurons. The relative RNA levels showed a temporal increase in the two cell classes, with late stage neurons expressing even higher amounts than corresponding progenitors. Despite this trend, the relative protein levels were higher in progenitors than neurons and with the difference being highest at early stages. This pattern suggested higher yield of translation or higher protein stability of these genes in progenitors over neurons (Figure 2B). This module is enriched in RNA-binding proteins like DEAD box

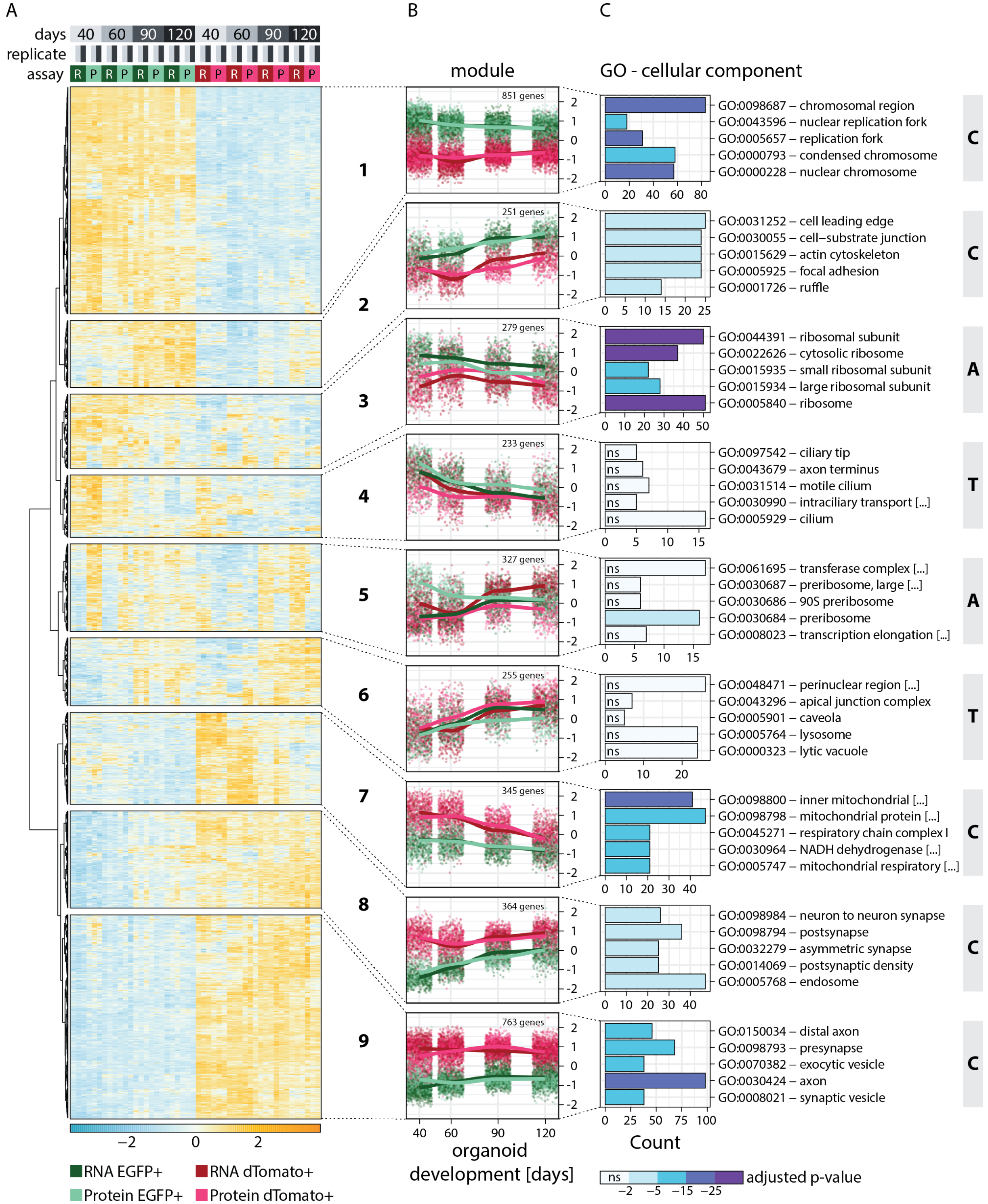


Fig. 2. Clustering of genes based on RNA-protein expression patterns in progenitors and neurons in human brain organoids.
A) Heatmap showing clustering results for the nine gene expression modules. Z-score normalized abundance of RNA (R) and protein (P) in progenitors (EGFP+, green) and neurons (dTomato+, red) at different stages of organoid development. B) Z-score normalized RNA and protein abundance in progenitors (EGFP+, green) and neurons (dTomato+, red) at different stages of organoid development for the nine modules. For each dataset, each dot displays the relative abundance of one gene and a trendline was fit through all data points. C) Enrichment analysis for gene ontology (GO) cellular component (CC) for genes in the nine modules. Shown are the top five enriched terms per module and the number of genes (x-axis) belonging to the respective GO term. Bar fill colors show binned adjusted p-values and not significant (ns) terms are indicated. Modules are categorized according to their global expression pattern: cell class-specific expression ("C" modules), temporal expression ("T" modules) and ambiguous expression ("A" modules).

proteins and proteins important for rRNA processing and ribosome biogenesis (Figure 2C, Figure 2-figure supplement 1A,B). Finally, we analyzed the correlation of the relative abundance of RNA to protein for each gene expression module (Figure 2-figure supplement 2B). Modules that show a high cell class-specific expression ("C" modules) showed a higher correlation than modules with broader expression patterns ("T" and "A" modules) reflected in the coherency between RNA and protein trends. Thus, by combining RNA and protein expression data, we identified gene modules with cell class-specific and temporal gene expression patterns, as well as modules with seemingly ambiguous expression patterns that show divergent coherency between relative RNA and protein abundance.

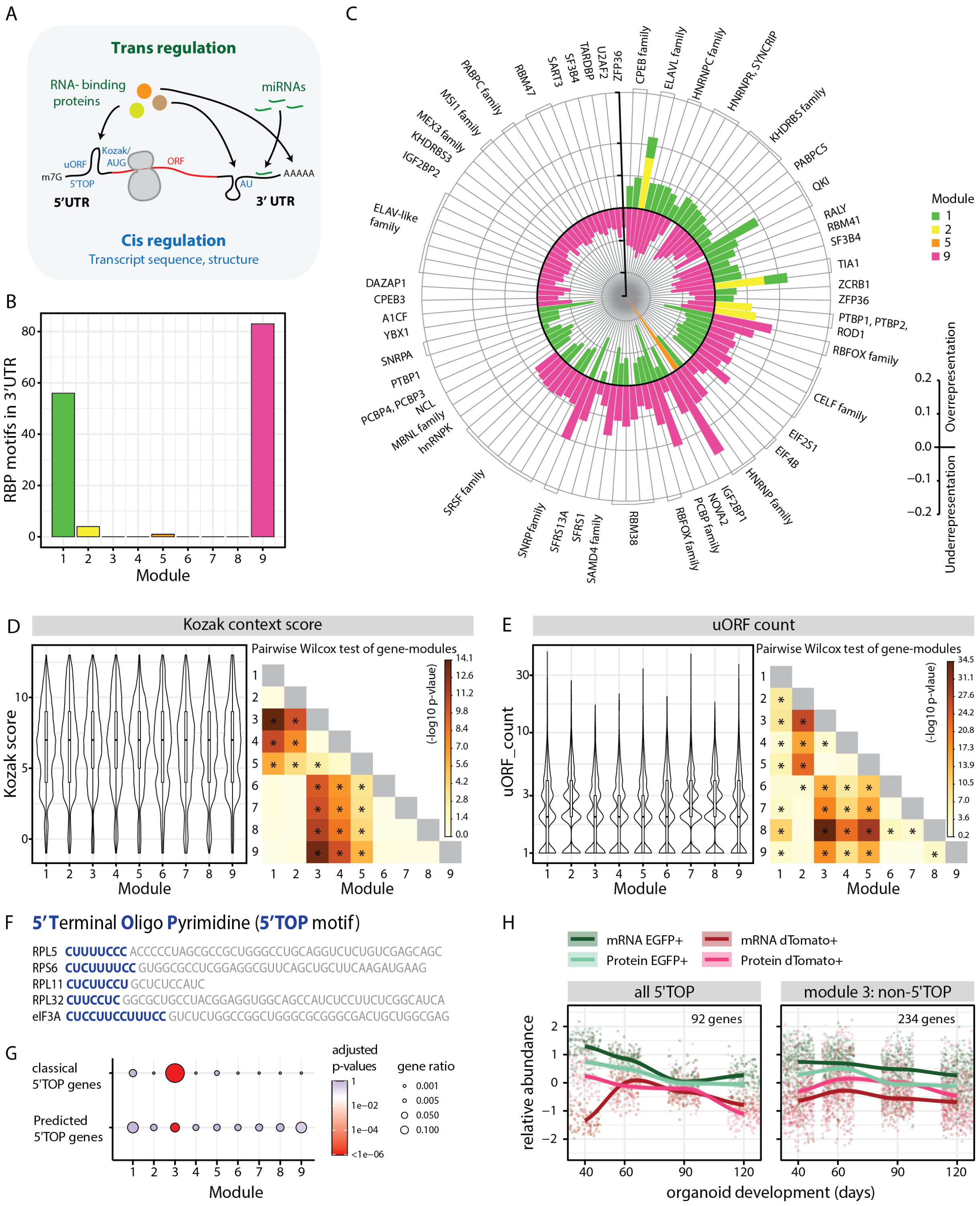
Analysis of module-specific RNA regulatory features highlights the 5'TOP gene module.

The emergence of highly distinct expression modules raises the possibility that the genes within each module are regulated transcriptionally and post-transcriptionally by common cell class- or stage-specific mechanisms. The post-transcriptional mechanisms can include regulation of translation as well as stability of the mRNAs and proteins, either uniformly or partially through compartmentalisation. As transcript features of the 5' and 3' UTR elements are postulated to be a major factor that directly influences RNA stability and translation (Blair et al., 2017; Floor and Doudna, 2016), we performed a comprehensive transcript feature analysis of the members of the 9 modules (Figure 3A, Figure 3-source data 1,2).

Firstly, we analyzed trans-regulatory features that can potentiate binding of RNA binding proteins (RBPs) and miRNAs (Figure 3A). To this end, we performed module-wise over- and under-representation analysis for RBP-binding motifs in the 5' and 3' UTR elements of the member genes (Figure 3B-C, Figure 3-figure supplement 1A-B, Figure 3-source data 1). Modules 1 and 9, whose member genes showed most cell class-specific expression, showcased high numbers of RBP motifs that were significantly enriched or depleted in the 3'UTR (Figure 3B). Upon comparing over- and underrepresented motifs among different modules, we observed that there was a reciprocal relationship between motifs enriched in the progenitor modules 1 and 2 and the neuronal module 9 (Figure 3C). These data suggest that the RBP motifs we identified and their associated proteins might play a prominent role in cell class-specific expression to promote progenitor or neuron-specific gene regulation. For instance, we observed that motifs for the RNA-binding protein RBFOX1, which was shown to promote neuronal expression of synaptic genes (Lee et al., 2016), were enriched in the 3'UTRs of module 9 (Figure 3C) in line with the enrichment for GO terms for synaptic genes (Figure 2C). Next, we searched for miRNA binding sites in the 3'UTR, which are known to repress the expression of transcripts they bind to. Module-wise analysis showed a significantly higher density of miRNA binding sites for genes of modules 6 and 8, indicating that the expression of some fate regulators and neuronal genes might be regulated by miRNAs (Figure 3-figure supplement 1C). For example, the upper-layer neuronal fate regulator gene SATB2, a member of module 6 is regulated through the binding of miR-541 and miR-92a/b to its 3'UTR thereby preventing its translation (Martins et al., 2021). Altogether, these data indicated that trans-regulatory factors play a major role in the regulation of transcripts with cell class-specific expression ("C" modules).

Next, we analyzed cis-regulatory features inherent in the mRNA sequence (Figure 3A, Figure 3-source data 2). We compared the lengths of coding sequences (CDS) and the 5' and 3'UTRs (Figure 3-figure supplement 1D-F). Compared to all other modules, module 3 transcripts were overall shorter in all these features, whereas modules 8 and 9 seemed to have longer 5' and 3'UTRs (Figure 3-figure supplement 1D-F). Secondary structures in a transcript can be estimated by prediction of minimum free energy (MFE) (Lorenz et al., 2011; Mathews and Turner, 2006), which is positively correlated with translation rate and a measure of RNA stability (Nomura et al., 1984). Our analysis revealed that the MFE for the CDS as well as the 5' and 3'UTRs were overall higher for module 3, indicative of less stable secondary structures (Figure 3-figure supplement 1G-I). In contrast, the 5'UTRs of modules 8 and 9 genes showed overall significantly lower MFE indicating more stable structures (Figure 3-figure supplement 1H). These observations suggest that the long transcripts of neuron-enriched genes and their complex UTRs might contribute to complex gene regulatory mechanisms. On the contrary, short and simpler transcripts in the ribosome-enriched module might circumvent some of the gene regulatory mechanisms and govern their unique expression pattern (Figure 3-figure supplement 1K).

Translation initiation efficiency is directly influenced by the Kozak sequence, which refers to the six nucleotides preceding the AUG start codon and the nucleotide immediately downstream. We computationally calculated a Kozak context score (Floor and Doudna, 2016) that describes the overall similarity to the consensus sequence GCCRCCAUGG (R = A or G) for all transcripts. Interestingly, genes in "C" modules have weaker Kozak context scores, suggesting reduced translational initiation rates that potentially allow precise translational regulation of these transcripts via additional mechanisms (Figure 3D). In contrast, genes in "A" modules have the overall highest Kozak scores (Figure 3D). Similar differences between "C" and "A" modules were also observed for the occurrence of upstream open reading frames (uORFs) in the 5' UTR. It has been suggested that uORFs do not affect mRNA levels but can significantly reduce translation (Barbosa et al., 2013). We predicted uORFs by occurrence of [ACT]TG sequence upstream of the main ORF (Floor and Doudna, 2016). We observed a reduced number of uORFs predicted in modules 3, 4 and 5 (Figure 3E) suggesting a limited role of uORFs in their expression pattern in contrast to gene expression in the "C"



RPL5 **CUUUUCCC** ACCCCCUAGCGCCGCUGGGCCUGCAGGUCUCUGUCGAGCAGC
RPS6 **CUCUUUUCC** GUGGCGCCUCGGAGGCGUUCAGCUGCUUCAAGAUGAAG
RPL11 **CUCUUCCU** GCUCUCCAUC
RPL32 **CUUCCUC** GGCGCUGCCUACGGAGGUGGCAGCCAUCUCCUUCUCGGCAUCA
eIF3A **CUCCUUCCUUUCC** GUCUCUGGCCGGCUGGGCGCGGGCGACUGCUGGCGAG

Fig. 3. Analysis of module-specific RNA-regulatory features.

A) Schematic showing trans-regulation of mRNA translation by RBPs and miRNAs and cis-regulation by mRNA sequence features and secondary structure. B) Barplot showing number of RBP motifs with significant over/underrepresentation in the 3'UTRs of genes for each module. Significance threshold: adjusted p-value <0.05. C) Circular barplot showing RBP motifs that are significantly over/underrepresented in modules in Figure 3B. Black circle marks zero position. Positive values outside the black circle indicate overrepresentation. Negative values inside the black circle indicate underrepresentation. RBPs for specific motifs are mentioned on the circular axis ticks. Bars are colored according to the module. D, E) Violin plots of module-wise distribution of Kozak context scores (D) and number of uORFs (E) for transcripts of member genes. Inner boxplots mark median Kozak context scores and interquartile range (IQR). Whiskers extend to 1.5x IQR. The heatmap on the right shows the -log10 p-value of the pairwise wilcox test of gene modules. Significant comparisons are marked by an asterisk (p < 0.01). F) 5' terminal sequences of five selected 5'TOP transcripts. Pyrimidine-rich 5'TOP sequences are highlighted in blue. G) Dot plot showing overrepresentation analysis of 5'TOP genes across the 9 modules. Significance threshold: Adjusted p-value <0.05. Gene lists refer to Classical 5'TOP genes (Philippe et al., 2020) and predicted 5'TOP mRNAs (Yamashita et al., 2008) (Figure 3-source data 3). H) Z-score normalized relative abundance of transcripts and proteins in progenitors and neurons of all 5'TOP genes and of module 3 excluding 5'TOP genes.

modules. Lastly, we observed that the 3'UTRs of genes in module 6 are enriched in adenylate/uridylate (AU)-rich sequences that promote cellular context-dependent mRNA decay or stability (Figure 3-figure supplement 1J) (Otsuka et al., 2019). This mode of regulation probably supports the expression pattern of this module, where transcripts are translationally silenced in late progenitors and active in neurons (Figure 2B). On the other hand, module 9 has significantly less AU-rich elements in the 3'UTR (Figure 3-figure supplement 1J) that corroborates the observed depletion of binding motifs for the ELAV-like family (Figure 3C), which is a major binding partner of AU-rich sequences (Ma et al., 1997). Overall, this analysis revealed that cell class-specific gene expression seems to be regulated by diverse posttranscriptional mechanisms: trans-regulatory factors, secondary structures in UTRs, imperfect Kozak and uORFs (Figure 3-figure supplement 1K). These mechanisms might contribute to the high correlation of relative RNA and protein abundances. However, despite having perfect Kozak sequences, "A" module genes have the lowest correlation of relative levels of RNA and protein (Figure 2-figure supplement 2B).

GO-term analysis showed that "A" modules (Module 3 and 5) were enriched in ribosome and ribosomal biogenesis genes. (Figure 2C, Figure 2-figure supplement 1A,B). This was especially striking for module 3 which was enriched in ribosomal proteins and eukaryotic translation initiation factors that contain the unique regulatory 5' terminal oligopyrimidine (5'TOP) motif in their 5'UTR (Figure 3F,G; Figure 3-source data 3) (Cockman et al., 2020; Philippe et al., 2020; Yamashita et al., 2008). This observation raised the possibility that 5'TOP genes exhibit a unique RNA-protein expression pattern during corticogenesis.

Relative protein yield of 5'TOP mRNAs is lower in early progenitors compared to early born neurons.

Analyzing the expression pattern of all 5'TOP genes showed that although expressed highly in both progenitors and neurons, 5'TOP transcripts were relatively more abundant in progenitors compared to neurons, with this effect being the most striking at early stage of organoid development at day 40 (Figure 3H). Analyzing the transcript distribution by RNA-FISH for two example 5'TOP genes RPL5 and RPL11 in 40 day old organoids confirmed this observation (Figure 4A), with high intensity of fluorescence being present in the progenitor-rich VZ, and low in the surrounding neuron-rich region. This pattern was not observed by RNA-FISH for the immature neuron marker DCX and in the no-probe control (Figure 4-figure supplement 1A,B). At protein level, however, a relative difference in progenitor- and neuron-rich regions was not detectable (Figure 4B). This discrepancy was also replicated in the absolute levels of RNA and protein for RPL5 and RPL11 (data available in the R shiny app). To verify this RNA-protein discrepancy at cellular resolution, we grew sparsely labeled organoids using 95% of unlabeled control H9 cells and 5% of the dual reporter cell line (Figure 1-figure supplement 1D). Next, we stained cryosections from these organoids for transcripts and protein of the 5'TOP genes RPL5 and RPL11 by RNA-FISH and by immunostainings, respectively. We quantified the average RNA and protein intensity per cell of individual progenitors labeled by EGFP and neurons labeled by dTomato (Figure 4-figure supplement 1C). Indeed, we could validate that RPL5 and RPL11 show higher transcript levels in progenitors than neurons, but protein levels were not different between the two cell classes (Figure 4-figure supplement 1D,E). To confirm that this observation is also found in vivo, we analyzed a published human fetal brain scRNA-seq dataset (Eze et al., 2021). Similar to our results, 5'TOP transcripts show higher expression in neuroepithelial, radial glia and intermediate progenitor cells compared to neurons at gestational weeks 6 to 10 (Figure 4-figure supplement 1F).

To investigate possible molecular mechanisms that contribute to the observed discrepancy between RNA and protein levels of the 5'TOP genes, we analyzed a published ribosome profiling dataset from in vitro 2D cultures of hPSC-derived human neural progenitors and neurons (Blair et al., 2017) (Figure 4-figure supplement 1G). This dataset corroborated our observation that progenitors have higher levels of cytosolic 5'TOP mRNAs compared to neurons. Interestingly, while for 5'TOP mRNAs the actively translated (polysome-bound) mRNA pool was similar between the cell classes, progenitors contained significantly higher amounts of monosome-bound mRNA that is considered translationally less active (Figure S8H). In contrast, transcripts of cell class-specific modules 1 and 9 showed not only a cell class-specific expression in the cytosol but also coherently their specific occurrence in the translated polysome fraction (Figure 4-figure supplement 1H). Hence, we hypothesized that some of the excess 5'TOP transcripts present in the progenitors might reside in a translationally-inhibited stage.

To test if similar mechanisms operates in the 3-D organoid tissue and if the translation of 5'TOP genes is indeed different between early progenitors and neurons, we generated hPSC lines carrying a reporter that contained a 5'TOP motif in the 5'UTR of a doxycycline-inducible tagBFP fluorescent protein. As a negative control we generated non-5'TOP control reporter lines, in which the 5'TOP nucleotides were mutated (Figure 4C). Initially, the fidelity of these reporter lines was evaluated at hPSC stage (Figure 4-figure supplement 2A). Untreated hPSCs did not show leaky tagBFP expression (Figure 4-figure supplement 2B), whereas doxycycline treatment induced strong tagBFP fluorescence in all six reporter cell lines and no differences between the 5'TOP and non-5'TOP reporters were observed (Figure 4-figure supplement 2B). As the translation of 5'TOP mRNAs is known to be inhibited upon mTOR inhibition (Cockman et al., 2020), we treated cells with the mTOR inhibitor Everolimus. As expected, we observed reduced tagBFP expression in the 5'TOP compared to the non-5'TOP reporter lines (Figure 4-figure supplement 2C). This observa-

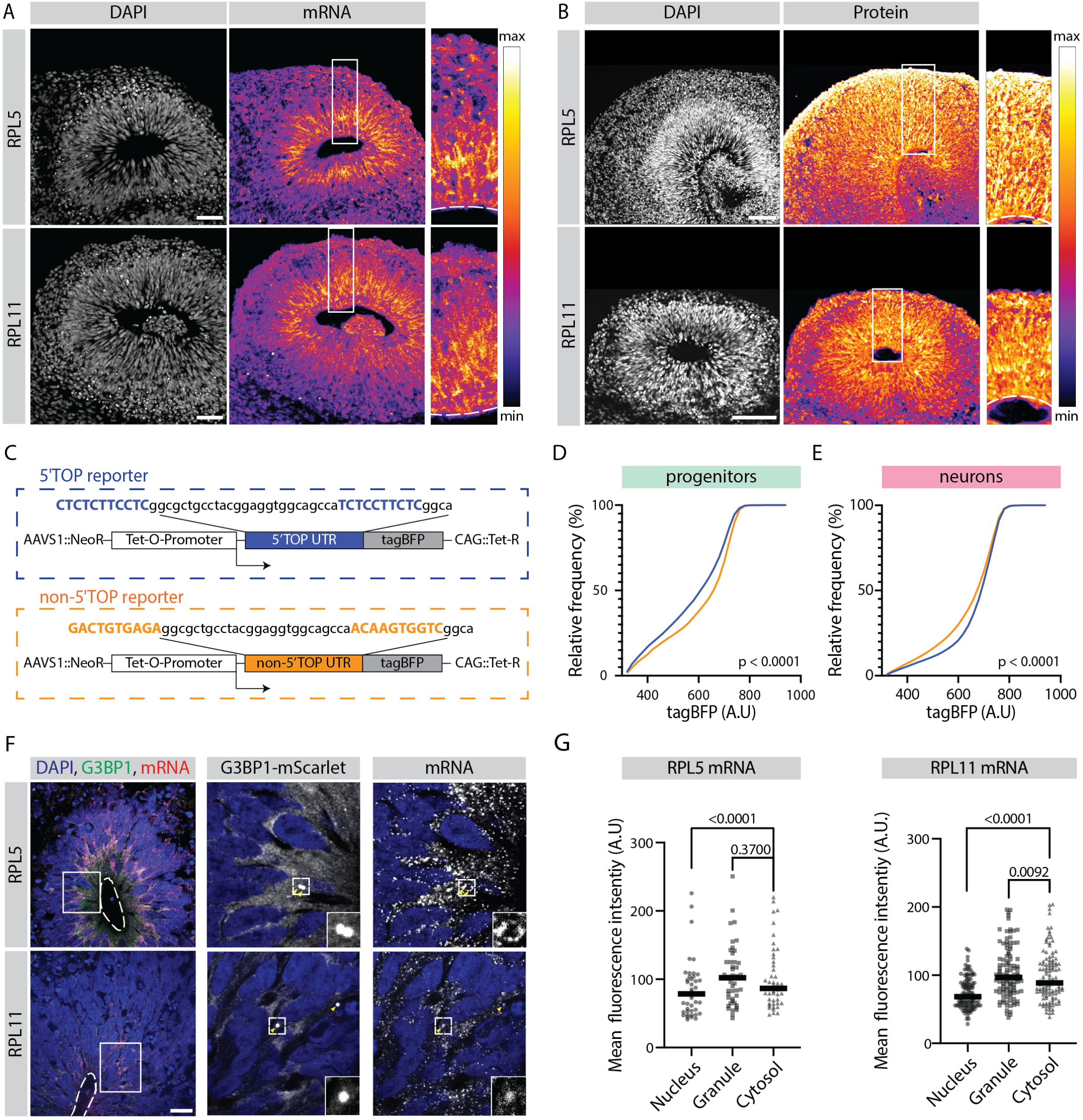


Fig. 4. Translation of 5'TOP mRNAs is partially inhibited in early progenitors compared to early born neurons.
A) RNA-FISH for the 5'TOP transcripts RPL5 and RPL11 in 40-day old organoids. Images show a typical ventricular zone (VZ)-like structure in the brain organoid with a zoomed-in image of the boxed area on the right. Dotted line marks the apical side of VZ. Scale bar = 50 µm. B) Immunostaining for the 5'TOP proteins RPL5 and RPL11 in 40-day old organoids. Images show a typical VZ-like structure in the brain organoid with a zoomed-in image of the boxed area on the right. Dotted line marks the apical side of VZ. Scale bar = 50 µm. C) Schematic representation of the dox-inducible 5'TOP (blue) and non-5'TOP (orange) tagBFP reporter constructs. D, E) Representative cumulative distribution of reporter tagBFP levels in neural progenitors (D) and in neurons at day 40 (E) from 5'TOP reporter (blue) and non-5'TOP reporter (orange). F) Confocal scan of a typical VZ-like structure in a 40-day old organoid expressing G3BP1-mScarlet stained for RPL5 and RPL11 transcripts by RNA-FISH. Zoomed in image of the inlay shows G3BP1 granules pointed by yellow arrows. Dotted line marks the apical side of VZ. Scale bar = 20 µm. G) Quantification of RPL5 and RPL11 RNA-FISH signal intensities in the nucleus, G3BP1 positive granules and cytosol of progenitors in the VZ. P-value of Mann-Whitney test. Each dot represents a cell. (Number of cells analyzed: RPL5: n=45; RPL11: n=34).

tion confirmed that tagBFP expression in the reporter lines is indeed mTOR pathway-regulated and thus suitable to study 5'TOP translation during organoid development. Hence, we used these reporter lines to grow organoids that we treated at day 40 with doxycycline to induce the expression of tagBFP. Three days after doxycycline induction, the organoids were dissociated to make a single cell suspension and immunostained for progenitor (SOX2) and neuron (TUJ1) markers (Figure 4-figure supplement 3A). The expression of tagBFP and cell class-specific markers was analyzed by flow cytometry (Figure 4-figure supplement 3B-C). We observed that only 30% of the cells in the organoids expressed tagBFP, indicating that it is difficult to achieve uniform doxycycline-mediated activation of the construct in 3D organoid tissue (Figure 4-figure supplement 3B). We therefore only analyzed cells that expressed tagBFP and were SOX2-PE or TUJ1-488 positive (Figure 4-figure supplement 3C). In progenitors (SOX2+), the 5'TOP reporter showed more cells with lower tagBFP fluorescence compared to the non-5'TOP reporter (Figure 4D). On the other hand, in neurons (TUJ1+), the 5'TOP reporter showed the tendency towards higher tagBFP expressing cells compared to the non-5'TOP reporter (Figure 4E). This was observed consistently between different cell lines and experiments (Figure 4-figure supplement 3D,E) indicating that in general, the presence of 5'TOP motif lowers the effective tagBFP protein abundance in progenitors. Together, we provide multiple pieces of evidence indicating that despite higher mRNA levels, early progenitors exhibit less effective translation of 5'TOP element-containing mRNAs than neurons.

Early progenitors exhibit a stress-associated 5'TOP translational control mechanism in a developmental context.

5'TOP translation is known to be regulated during cellular stress (Cockman et al., 2020). During stress conditions, 5'TOP RNAs localize to stress granules (SGs), which are dense RNA-protein condensates (Damgaard and Lykke-Andersen, 2011; Wilbertz et al., 2019). Hence, we asked if we also observe stress granule-like (SGL) structures during corticogenesis in organoids. We immunostained for the stress granule marker G3BP1 in untreated organoids and organoids treated with sodium arsenite (NaAs), a strong inducer of cellular stress (Figure 4-figure supplement 4A). As expected, more G3BP1-positive puncta (Figure 4-figure supplement 4A) with higher fluorescence intensity (Figure 4-figure supplement 4B) were observed in the NaAs treated organoids. However, the occurrence of weak G3BP1 puncta in control organoids, particularly in the VZ was supportive of our hypothesis (Figure 4-figure supplement 4C).

Next, we asked if 5'TOP transcripts localize to G3BP1-positive SGL structures. For this, we generated a G3BP1 reporter line in which we endogenously tagged G3BP1 with the fluorescent protein mScarlet (Figure 4-figure supplement 4D). We generated organoids from this line and observed mScarlet-positive puncta that co-stained with an anti-G3BP1 antibody (Figure 4-figure supplement 4E) mostly in the VZ (Figure 4-figure supplement 4F). Furthermore, we performed RNA-FISH and observed localization of RPL5 and RPL11 transcripts to these mScarlet-positive SGL structures (Figure 4F). The intensity of the FISH signal was similar in the SGL structures and the cytoplasm (Figure 4G), which would suggest that only a small fraction of the mRNA is sequestered in the granules, in agreement with reduced effective translation and not a translational block. In line with previous studies (Damgaard and Lykke-Andersen, 2011; Wilbertz et al., 2019) RPL5 and RPL11 transcripts were strongly enriched in G3BP1 positive granules in NaAs-treated organoids (Figure 4-figure supplement 4G).

To reinforce our observation that 5'TOP transcripts are partially translationally silenced, we estimated the relative RNA stability from the fold change of exonic to intronic reads using our RNA-seq data (Alkallas et al., 2017) (Figure 4-source data 1). This analysis revealed that 5'TOP RNAs are more stable in the progenitors than neurons (Figure 4-figure supplement 5A) and that this difference is most prominent at day 40 (Figure 4-figure supplement 5B). In contrast, the cell class-specific modules 1 and 9 showed higher transcript stabilities in progenitors and neurons, respectively, in line with our transcript and protein expression data (Figure 4-figure supplement 5B). Thus, early progenitors seem to have excess mature 5'TOP mRNAs with higher RNA stability and a fraction of these is not actively translated and sequestered into G3BP1-positive SGL structures.

It is known that 5'TOP translation is regulated by the mTOR pathway and is inhibited upon mTOR pathway inhibition (Cockman et al., 2020). Therefore we assessed the status of mTOR pathway by immunostaining for phosphorylated forms of two of its substrates: 4E-BP1 and RPS6 (S6). The uniform distribution of phospho-4E-BP1 (p4E-BP1) across the tissue indicated generally active mTOR signaling (Figure 5-figure supplement 1A). However, as reported previously for organoid and fetal samples (Andrews et al., 2020; Eze et al., 2021) phospho-S6 (pS6) staining showed enrichment outside the VZ in neuronal regions (Figure 5-figure supplement 1B). These observations suggest that differential mTOR activity in early progenitors, especially ventricular radial glia, and neurons might regulate differential 5'TOP translation. Based on these results we propose that the mTOR-5'TOP translational regulation axis is active in early neural progenitors and might play a crucial role in neuronal differentiation.

mTOR-mediated regulation of 5'TOP mRNA translation is crucial for the fidelity of cortical development.

We hypothesized that mTOR pathway overactivation should result in increased translation of ribosomal and other 5'TOP-containing genes in neural progenitors. To upregulate the mTOR pathway during organoid development, we generated hPSCs with loss of function mutations in TSC2 (TSC2-/-), a

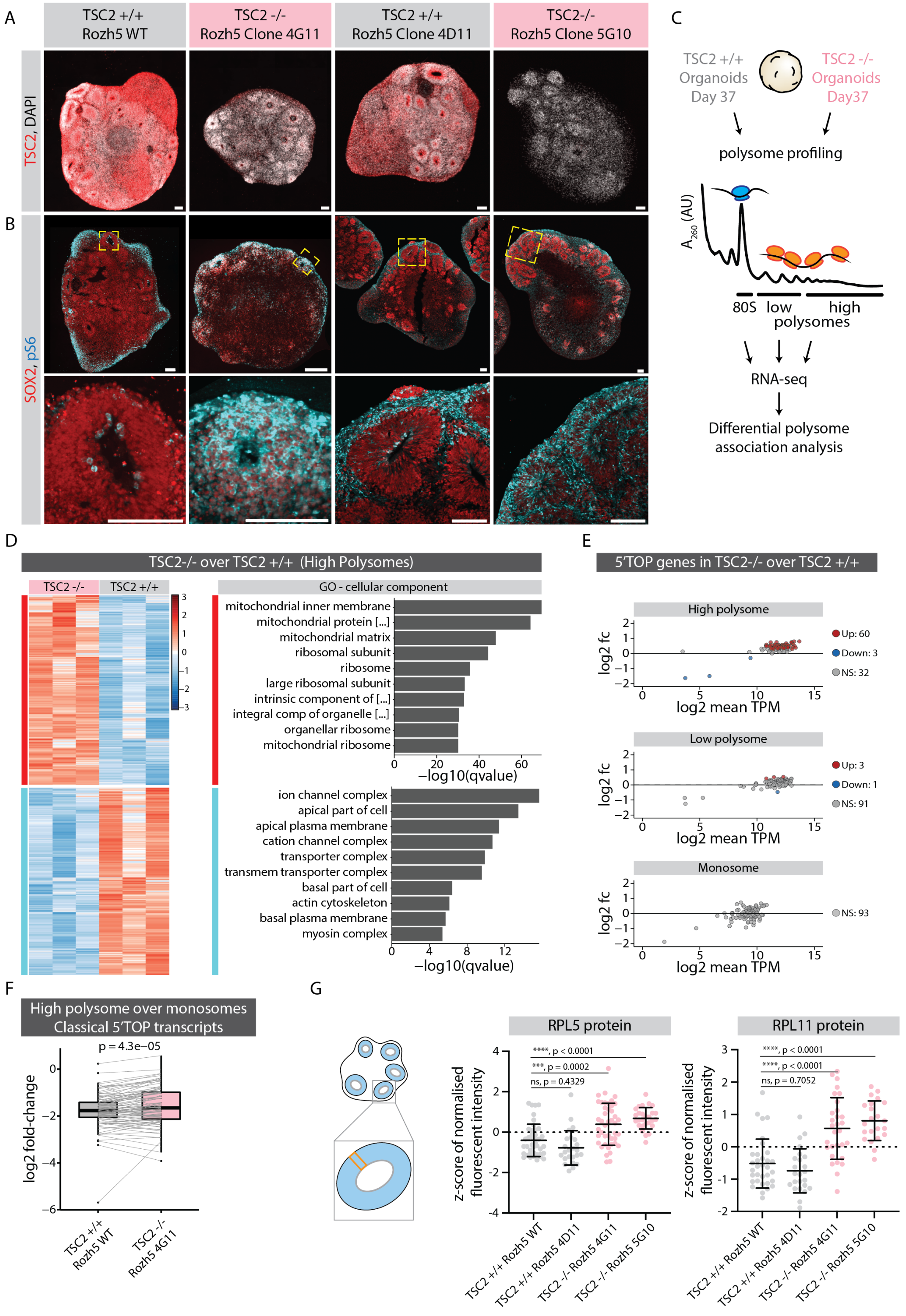


Fig. 5. mTOR overactivation causes increased translation of 5'TOP transcripts leading to precocious differentiation.
A) Immunostaining of 40 days old TSC2+/+ and TSC2-/- organoid sections with anti-TSC2 antibody. Scale bar = 100 µm. B) Immunostaining of 40 days old TSC2+/+ and TSC2-/- organoid sections stained with anti-pS6 and anti-SOX2 antibodies. Zoomed in image of the inlay on the bottom. Scale bar = 100 µm. C) Schematic representation of polysome profiling of 37 days old TSC2+/+ and TSC2-/- organoids followed by RNA-seq and differential polysome association analysis of 80s monosome, low-polysome and high-polysome fractions. Data was generated from 3 batches of organoid differentiation for Rozh5 WT and Rozh5 TSC2 -/- 4G11 clones. D) Heat map of hierarchical clustering of genes differentially associated with high-polysome fractions of TSC2+/+ and TSC2-/- organoids (left). Top 10 most significant cellular component GO terms for TSC2-/- enriched DAGs and TSC2+/+ enriched DAGs (right). E) MA plot of 5'TOP genes for high-polysome, low-polysome and monosome fractions shows log2 fold change of TSC2-/- over TSC2+/+ versus log2 mean TPM (expression levels). Significant genes (threshold: log2 fold change = 1, FDR = 0.1) are color coded. F) Log2 fold change of classical 5'TOP genes associated with high-polysome fractions over monosome fractions of TSC2+/+ and TSC2-/- organoids. P-value of paired t test. G) Quantification of mean z-score normalized fluorescent intensity of RPL5 and RPL11 proteins in the ventricular zones (VZs) of TSC2+/+ and TSC2-/- organoids. Error bars mark standard deviation. P-values of unpaired t-test are shown. Each dot represents a VZ. Data from 3 batches of organoid differentiation for each clone. Schematic on the left shows a typical ROI analyzed for a VZ.

negative regulator of mTOR signaling (Crino, 2016) (Figure 5-figure supplement 1C-E). Immunostainings of organoid tissue confirmed the absence of TSC2 protein (Figure 5A and Figure 5-figure supplement 1F). Furthermore, TSC2-/- organoids showed significantly higher signal of pS6 throughout the VZ when compared to the control tissue (Figure 5B and Figure 5-figure supplement 1G), indicating overactivation of the mTOR pathway. To check if increased mTOR signaling indeed leads to excess translation of 5'TOP genes, we performed polysome profiling from control and TSC2-/- organoids (Figure 5-figure supplement 2A) and sequenced the ribosome-bound transcripts in the monosome (80S), low- and high-polysome fractions (Figure 5C). Due to the large amount of material needed for polysome profiling, we used 37 day old intact organoids that contain a vast majority of cells in progenitor state. Initially, we performed differential gene association analysis between TSC2-/- and control samples for each fraction (Figure 5D and Figure 5-figure supplement 2B,C, Figure 5-source data 1). We indeed found an enrichment for ribosomal genes specifically in the high-polysome fraction of TSC2-/- over control organoids (Figure 5D). When specifically analyzing 5'TOP transcripts we found overall increased levels in the high-polysome fractions in TSC2-/- organoids compared to control organoids, whereas similar levels were detected in the low-polysome and monosome fractions (Figure 5E). Accordingly, the relative fraction of 5'TOP transcripts actively translated in high-polysomes over the translationally less active monosomes was increased in TSC2-/- over control organoids (Figure 5F, Figure 5-source data 1). To confirm that the upregulation of 5'TOP transcripts in the polysome fraction results in increased protein levels, we performed RNA-FISH and immunostainings in TSC2-/- and control organoids for RPL5 and RPL11. While we could not detect differences in their RNA levels (Figure 5-figure supplement 3A,B), we observed an increase in protein levels in the VZ of TSC2 -/- organoids (Figure 5G and Figure 5-figure supplement 3C,D). Taken together, these results verify an increase in 5'TOP mRNA translation upon mTOR activation and indicate that physiological mTOR activity is crucial to regulate protein levels of 5'TOP genes in early progenitors. Thus, our results suggest that the translational machinery itself is uniquely regulated post-transcriptionally during neurodevelopment.

Next, we asked if mTOR overactivation and the concomitantly increased translation of ribosomes impacts gene regulation of other genes and potentially the tissue development. We checked if the genes differentially associated with ribosome fractions in polysome profiling were enriched in any of the 9 gene expression modules (Figure 2B) by over- and under-representation analysis (Figure 5-figure supplement 4A, B). For genes, which were significantly higher in the high-polysome fraction of TSC2-/- over those of control, we observed a significant enrichment for modules 4, 7 and 9 (Figure 5-figure supplement 4A). Module 7 is enriched in mitochondrial genes known to be differentially expressed in TSC2-/- at transcript level (He et al., 2020; Koyanagi et al., 2011). Importantly, the enrichment of neuronal module 9 to be highly translated in TSC2-/- is striking. In contrast, the genes which were less translated in the TSC2-/- organoids over control were significantly enriched in progenitor-specific modules 1 and 2 (Figure 5-figure supplement 4A). This data suggests that mTOR overactivation and the subsequent impact on the translation machinery shifts the state of the tissue towards precocious differentiation. These observations were also corroborated in the underrepresentation analysis. Genes with higher translation in TSC2-/- were significantly depleted in modules 1 and 2 (progenitor-specific expression), whereas genes with lower translation in TSC2-/- were significantly depleted in module 9 (neuron-specific expression) as well as module 3, which is enriched in 5'TOP genes (Figure 5-figure supplement 4B).

We next tested if the increased translation of differentiation-related genes is reflected by an increased neuronal differentiation using the dual reporter line. For this, we performed flow cytometry using dual reporter TSC2-/- and TSC2+/+ organoids (Figure 5-figure supplement 4C,D). In agreement with the polysome profiling data, we observed a significant reduction in the fraction of EGFP+ cells (progenitors) and a significant increase in the fraction of dTomato+ cells (neurons) (Figure 5-figure supplement 4D). Overall, the ratio of dTomato+ cells to EGFP+ was increased in the TSC2-/-, indicative of an increased differentiation (Figure 5-figure supplement 4D). Apart from an increase in neuronal markers, we also observed precocious occurrence of the glial marker S100B in the TSC2-/- tissue at day 40 (Figure 5-figure supplement 4E), reinforcing our observation that the progenitors are shifted towards a precocious differentiation state. This was already observed at day 40 and possibly explains the premature gliogenesis observed at later stages of TSC2-/- cortical spheroids (Blair et al., 2018). While it is difficult to separate the primary and secondary effects of mTOR overactivation, we propose that the partial translational inhibition of 5'TOP mRNAs in early neural progenitors is crucial for cortical development.

Discussion

In this study, we show that multiple transcripts encoding components of the translational machinery that contain 5'TOP motifs are regulated on a post-transcriptional level during neurodevelopment. 5'TOP transcripts are abundantly expressed in all cells and yet show a unique regulation compared to other cellular transcripts (Cockman et al., 2020; Fonseca et al., 2014; Meyuhas and Kahan, 2015). During mitosis, 5'TOP transcripts skip the global translational inhibition (Park et al., 2016) whereas in stress context these transcripts are the first ones to be translationally inhibited to conserve the resources and energy of the cell by downregulating translation (Cockman et al., 2020; Fonseca et al., 2014; Meyuhas and Kahan, 2015). While translational regulation of 5'TOP transcripts is intensely studied in the context of stress, we show that it plays a major role during corticogenesis. Our observation of SG-like RNA-granules in early ventricular radial glia further supports the occurrence of stress-associated processes during corticogenesis. A recent report suggests that the energetic stage of a cell can impact the formation of SGLs under physiological conditions (Wang et al., 2022). Thus, the well documented difference in the metabolic states of glycolysis-dependent early progenitors and oxidative phosphorylation-heavy neurons (Iwata and Vanderhaeghen, 2021) might contribute to the regulation of 5'TOP mRNAs through the formation of SGLs. Another stress-related pathway, the unfolded protein response (UPR) is physiologically active at early stages of mouse corticogenesis to promote neurogenesis (Laguesse et al., 2015). Overactivation of the UPR beyond its physiological levels affects the generation of intermediate progenitors in the mouse and human cortical tissue, resulting in a microcephalic phenotype (Esk et al., 2020; Laguesse et al., 2015; Pașca et al., 2019). Perhaps, similar mechanisms regulating 5'TOP mRNA translation are utilized during development and during stress conditions. Recently, a role of the translational regulation of 5'TOP transcripts has been suggested in germline stem cell differentiation in Drosophila (Martin et al., 2022) and in mouse adult neurogenesis (Baser et al., 2019). It will be exciting to investigate if stress-associated pathways play a physiological role during the development of not only the cerebral cortex but also other tissues.

Regulation of the 5'TOP mRNA translation ultimately determines the availability of ribosomes in the cells. Ribosome availability is proposed to be a critical factor determining the translation of mRNAs with complex 5'UTRs with secondary structures (Hetman and Slomnicki, 2019). Supporting this hypothesis, we found that Module 9 neuronal genes that feature more complex 5'UTRs (Figure 3-figure supplement 1H) are particularly deregulated upon 5'TOP translation deregulation accompanying mTOR overactivation (Figure 5-figure supplement 4A). Our data suggests that regulating the number of ribosomes is particularly important in the early progenitors. An imbalance of ribosomal components caused by deregulation of 5'TOP transcripts likely causes aberrant translation of differentiation markers and affects the fidelity of cortical development. Reduced ribosome biogenesis and availability is also indicated at an early stage of mouse corticogenesis (Chau et al., 2018; Harnett et al., 2021). While many gene-specific examples of neural priming in radial glia are known (Yang et al., 2014; Zahr et al., 2018), our data suggest that radial glia regulate the translation of the translational machinery itself. We speculate that this regulation is crucial to maintain the multipotency of early progenitors to prevent aberrant translation of pre-existing transcripts of differentiation genes. A global regulation of the translational machinery has been recently reported for mid-gestation stages of mouse corticogenesis, with gradual reduction in the translational efficiency of 5'TOP genes and the availability of ribosomes (Harnett et al., 2021). It would be interesting to test if similar changes occur during late corticogenesis in humans, which includes a complex repertoire of neural progenitors, especially the outer radial glia.

The mTOR pathway is a major regulator of 5'TOP mRNA translation (Cockman et al., 2020; Fonseca et al., 2014; Meyuhas and Kahan, 2015) and is a key feature of any healthy cell. The role of the mTOR pathway in progenitor proliferation and neuronal morphology and function is well documented (Switon et al., 2017). However, the cell type- and developmental stage-specific roles of the mTOR pathway are becoming clear only recently. Our results in human brain organoid tissue show that early radial glia show lower mTOR activity in interphase. Studies using mouse, 2D neuronal cultures and cortical spheroid cultures have also reported a drop in mTOR activity during early stages of corticogenesis (Blair et al., 2018, 2017; Chau et al., 2018). Analysis of human fetal transcriptome of early stages has also suggested a similar trend (Eze et al., 2021). On the other hand, human outer radial glia that arise at later stages of corticogenesis exhibit higher mTOR activity, both in primary fetal tissue and brain organoids (Andrews et al., 2020). Thus, despite its general role, mTOR pathway activity undergoes modulation during distinct neurodevelopmental stages. Additionally, downstream effects of mTOR signaling are linked to diverse cellular processes such as translation, cell cycle and actin biology. Therefore, linking the defects in the pathway to a final cellular phenotype remains challenging. We provide an alternative molecular role of this pathway to directly regulate ribosome availability and thus indirectly control many other downstream cellular pathways. Mutations in mTOR pathway-related genes are associated with many genetic disorders such as tuberous sclerosis, focal cortical dysplasia and megalencephaly (Crino, 2016). Thus, to understand the disease mechanisms, it will be crucial to consider cell type- and stage-specific activity of the mTOR pathway.

Beyond 5'TOP genes of the translational machinery, our dataset highlights distinct gene expression modules based on transcript and protein abundance traits. This includes cell class-specific "C" modules with coherent RNA-protein expression as well as "T" and "A" modules where transcript and protein do not follow the same trajectories. Additionally, analysis of cis and trans regulatory factors alludes to distinct RNA regulation mechanisms that are potentially predominant for these genes and contribute to their expression patterns. For instance, the reciprocal relation in the RBP motif enrichment in 3'UTRs of "C" modules 1 and 9 highlights interesting roles of RBPs to regulate cell class-specific gene expression. Thus, our approach opens new avenues to study RNA regulation during corticogenesis.

Finally, our integrated dataset describes cell class- and developmental stage-specific gene expression, both at RNA and protein level and is available to browse for the wider community in the form of a R-based Shiny app (https://organoid.multiomics.vbc.ac.at). To our knowledge, this is the first time RNA and protein datasets for human corticogenesis have been integrated. Only looking at transcript or protein data wouldn't have revealed the post-transcriptional regulation of the ribosomal genes, which are usually considered housekeeping genes and ignored for example in sc-RNA-seq analysis and often used as normalizers for gene expression. This rich resource can be used to browse RNA-protein expression patterns of various other genes across the human organoid developmental timeline. For instance, our data indicates that the gene regulation of the upper layer marker SATB2 observed during mouse corticogenesis (Harnett et al., 2021) also holds true in humans. As characterized by the general trend of module 6, where the RNA expression of SATB2 is high in progenitors and neurons of later stages, the trends at protein level suggest translational inhibition of the SATB2 transcripts in progenitors. Integration with information on protein abundance is crucial for the comprehensive understanding of gene expression. This approach opens a new avenue to study the regulation and expression pattern of individual genes and gene modules. Thus, integrative omics approaches can reveal new biological mechanisms underlying diverse developmental events.

Acknowledgements

We would like to thank Daniel Matějů and the Knoblich lab members for their feedback on the manuscript. We thank the IMBA stem cell core facility for their service, IMBA/IMP/GMI BioOptics facility for flow cytometry and microscopy services; IMBA/IMP/GMI Bioinformatics for sequencing analysis; the VBCF Sequencing unit for sequencing the IMBA/IMP/GMI protein chemistry core facility for mass spectrometry. We thank the group of Andrea Pauli, especially Katrin Friederike Leesch, for sharing the resources and expertise in polysome profiling.

Work in the Knoblich laboratory is supported by the Austrian Academy of Sciences, the Austrian Science Fund (FWF), (Special Research Programme F7804-B and Stand-Alone grants P 35680 and P 35369), the Austrian Federal Ministry of Education, Science and Research, the City of Vienna, and a European Research Council (ERC) Advanced Grant under the European Union's Horizon 2020 programs (no. 695642 and no. 874769). JS was supported by EMBO long term fellowship (EMBO ALTF 794-2018). This project also received funding from the European Union's Horizon 2020 research and innovation program under the Marie Skłodowska-Curie fellowship agreement 841940 awarded to JS and 897137 to PT. Work in the Mechtler lab is supported by the EPIC-XS, Project Number 823839, funded by the Horizon 2020 Program of the European Union, by the project LS20-079 of the Vienna Science and Technology Fund and by ERA-CAPS I 3686, P35045-B, P32054 (FB) and P33380 (FB) projects of the Austrian Science Fund.

Author contributions

Conceptualization: JS, PT, JAK; Methodology: JS, PT; Investigation: JS, PT, NS, MS; Validation: JS, PT, NS; Visualization: JS, PT, NS; Data curation: JS, PT, MN, GD; Software: PT, MN, GD; Resources: JAK, KM; Writing-original draft: JS, PT; Writing- review editing: JS, PT, JAK; Funding acquisition: JS, PT, JAK; Supervision: JAK.

Conflict of interest

J.A.K. is inventor on a patent describing cerebral organoid technology and co-founder and scientific advisory board member of a:head bio AG.

Data and code availability

The RNA-seq data discussed in this publication have been deposited in NCBI's Gene Expression Omnibus (Edgar et al., 2002) and are accessible through GEO Series accession numbers GSE214654 and GSE214652. The mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium via the PRIDE (Perez-Riverol et al., 2021) partner repository with the dataset identifier PXD037106.

Material and methods

Cell lines and culture.

Feeder-free hESChuman embryonic stem cell (hESCs) line WA09 (H9) and feeder-free human induced pluripotent stem cell (hiPSC) line HPSI0114i-rozh_5 (Rozh-5) were used in this study. H9 was obtained from WiCell and Rozh-5 from HipSci. Cells were verified to display a normal karyotype and regularly tested for mycoplasma. All cells were cultured on hESCs-qualified Matrigel (Corning, 354277) coated plates with a modified in-house medium based on the E8 culture system (Chen et al., 2011). The original E8 recipe was supplemented with 0.5% BSA (Europa Bioproducts, EQBAH62-0500), 200 ng/ml in-house produced FGF2 and 1.8 ng/ml TGFβ1 (RD Systems, RD-240-B-010). Cells were passaged every 3-4 days using 0.5 mM EDTA. For generating TSC2 KO, cells were cultured using Cellartis DEF-CS 500 culture system (Takara Bio, cat. no. Y30012).

Dorsal tissue-enriched telencephalic organoid generation.

Dorsal forebrain-enriched telencephalic organoids were generated as previously described with slight modifications (Esk et al., 2020). Briefly, hPSCs were grown to 60-80% confluency and dissociated with Accutase (Merck, A6964) to obtain a single cell suspension. 4000-9000 cells were resuspended in 150µl of E8 supplemented with Rock inhibitor and seeded in each well of the 96-well ultra-low-attachment U-bottom plate to form embryoid bodies (EBs). On day 3 the medium was changed to E8. From day 6 onwards the EBs were cultured in neural induction medium (NI) (Lancaster et al., 2017). On day 10, EBs were embedded in Matrigel and transferred to 10-cm cell-culture dishes coated with anti-adherence rinsing solution (Stemcell technologies, P2443) in 15ml NI. From day 13 to 25 the organoids were grown in a differentiation medium lacking vitamin A (Diff-A). During day 13 to 15 they were treated with 2 pulses of 3µM CHIR99021 (Merck, 361571) to dorsalise the tissue. On day 18 dishes were transferred to an orbital shaker at 57 rpm. From day 25 onwards the organoids were grown in differentiation medium with vitamin A (Diff+A). The medium was supplemented with 1% Matrigel from day 40 onwards and with BDNF from day 55 onwards. See the media composition in Supplementary file 2.

For Tet-induction, the medium was supplemented with Doxycycline-hyclate (Merck, D9891) dissolved in water at a final concentration of 1.7 µg/ml medium of hPSCs and 3 µg/ml for organoids. For mTOR inhibition, hPSC culture medium was supplemented with Everolimus (Abcam, ab142151) dissolved in DMSO at a final concentration of 20 nM. For sodium arsenite (NaAs) treatment, the medium was supplemented with NaAs (Merck, S7400) dissolved in water to achieve a final concentration of 500 µM. To ensure diffusion in the 3D organoid tissue, treatment with NaAs was performed for 1 h.

Organoid dissociation.

Organoids were washed with DPBS without calcium and magnesium and added to a 9:1 mixture of Accutase (Merck, A6954) and 10x Trypsin (Thermo Fisher Scientific, 15400054) supplemented with 2 units/ml TURBO™ DNase (Thermo Fisher Scientific, AM2238). The dissociation was performed using the NTDK1 protocol on a gentleMACS™ Dissociator (Miltenyi Biotec, 130-093-235). The dissociated cells were pelleted, washed with DPBS without calcium and magensium (Thermo Fisher Scientific, 14190250) and filtered through a 70 µm cell strainer to remove residual tissue chunks. The filtered single cell suspension was then used either for live cell FACS or for immunostaining.

FACS.

For live cell sorting cells were resuspended in FACS buffer (DPBS without calcium and magensium with 2% BSA), filtered through a 35 µm cell strainer and then sorted on a FACS ARIAIII (BD Biosciences) controlled by FACSDiva software. For detecting GFP signal, a 488 nm laser was used with a 530/30 nm filter. For detecting dTomato signal, a 561 nm laser was used with 582/15 nm filter. The sort was done in PBS at low pressure using a 100 µm nozzle.

Immunostaining of dissociated cells.

For immunostaining cells, the dissociated single cells were fixed with 4% paraformaldehyde (PFA) for 30 mins on ice. The cells were permeabilized with 0.1% Saponin (Merck, 47036) in PBS either during or after fixation. Fixed cells were then pelleted and washed with a wash solution (0.1% Saponin + 1% BSA in PBS) to remove traces of PFA. Next, the cells were split for appropriate controls and resuspended in a staining solution (Fluorophore conjugated primary antibody in 0.1% Saponin + 1% BSA in PBS) for 30 mins on ice. The stained cells were then pelleted and washed twice with wash solution and resuspended in the final resuspension buffer (0.5%BSA in PBS). Antibodies used in this study are summarized in Supplementary file 2.

Flow cytometry analysis.

For analyzing the expression in live cells or in immunostained single cells, flow cytometry was performed on LSR Fortessa (BD Biosciences) controlled by FACSDiva software. Data was analyzed using FlowJo. For detecting tagBPF signal, 405nm laser was used with 442/46 nm filter. For detecting Alexa-488 signal, 488nm laser was used with 530/30 nm filter. For detecting PE signal, 561nm laser was used with 582/15 nm filter.

Immunohistochemistry.

Organoids were fixed in 4% paraformaldehyde (PFA) for 3-4 hours at room temperature. After extensive washes with PBS, organoids were immersed overnight, first in 30% sucrose (Merck, 84097) in PBS and thereafter in 1:1 mixture of 30% sucrose and OCT (Sakura, 4583). These organoids were then embedded in OCT using suitable cryomolds and frozen at -70°C. Samples were sectioned at 20 µm thickness using a Epredia Cryostar NX70 cryostat (Thermo Fisher, 957000H). The slides were dried overnight and then stored at -20°C.

After washing the slides extensively with PBS to remove the OCT, cryo-sections were permeabilized and blocked with blocking solution (5% BSA containing 0.3% Triton X-100 (Merck, 93420)) for 1-2 h at room temperature. Sections were then incubated with primary antibodies diluted in antibody solution (5% BSA, 0.1% Triton X-100 in PBS) overnight at 4°C. After three washes of 10 min with PBST (0.01% Triton X-100 in PBS), sections were incubated with secondary antibodies diluted in antibody solution containing 2 µg/ml DAPI at room temperature for 2 h. After three washes of 10 min with PBST (0.01% Triton X-100 in PBS) and one wash with PBS, the samples were mounted in fluorescence mounting medium (Agilent, S302380-2) and imaged with a spinning disk confocal microscope. Antibodies used in this study are summarized in Supplementary file 2. AlexaFluor 488, 568 or 647-conjugated secondary donkey antibodies were used at a 1:500 dilution.

RNA-FISH.

FISH probes for RPL5, RPL11 (Wilbertz et al., 2019) and DCX (Stellaris, VSMF-2504-5) were ordered from Stellaris. RPL5 and RPL11 probes were labelled with Quasar® 570 Dye. Sequences of the FISH-probes and corresponding fluorophores are listed in Supplementary file 1. DCX probes were labelled with Quasar® 670 Dye. Slides with cryosections were thawed and washed twice with DEPC-treated PBS, once with nuclease-free water and then immersed in a 1x TEA buffer (Triethylammonium acetate buffer) for 10 min. For permeabilization, sections were immersed in 2x SSC buffer and subsequently in 70%, 95% and 100% ethanol. After air drying for 1.5 h, sections were incubated in hybridization solution (10% Dextran sulfate (Merck, S4030), 1 mg/ml E.coli tRNA (Merck, 10109541001), 2mM Vanadyl ribonucleoside complex (New England Biolabs, S1402S), 2x SCC, 10% formamide (Merck, F9037), 0.5% BSA)) containing 250 nM of the FISH probe at 37°C overnight. After incubation, the slides were washed with wash buffer A (Stellaris, SMF-WA1-60) for 30 min at 37°C and subsequently stained with Hoechst for 30 min at 37°C in the dark. Next, the slides were washed with wash buffer B (Stellaris, SMF-WB1-20), subsequently immersed in 50%-, 85%-, and 100% ethanol for three minutes and air dried. Finally, the samples were mounted in fluorescence mounting medium (Agilent, S302380-2) and imaged with a spinning disk confocal microscope.

Imaging and image analysis.

Fluorescence images of immunostainings and RNA-FISH were obtained on an Olympus spinning disk confocal based on an Olympus IX3 Series (IX83) inverted microscope, equipped with a dual-camera Yokogawa W1 spinning disk (SD) allowing fast confocal acquisition. All components were controlled by CellSense software. Objectives used with the spinning disk confocal were 4x/0.16 (Air) WD 13mm, 10x/0.4 (Air) WD 3.1mm (for organoid overview), 20x/0.8 (Air) WD 0.6mm (for imaging individual VZs), 40x/1.25 (Silicon Oil) WD 0.3mm and 100x/1.45 (Oil) WD 0.13mm (for imaging SGL).

Preparation of image panels and Image analysis was performed using Fiji (Schindelin et al., 2012). For measuring average intensity per cell, progenitors and neurons were identified in mosaic reporter organoids, and the cell outline was traced to mark a ROI. Average intensities for each ROI were measured using the multi measure tool. For measuring average intensity in the VZ, a rectangular ROI was marked in the maximum intensity projection of the VZ. Average intensities for each ROI were measured using the multi measure tool. To enable integration of data from different experiments, results were z-score normalized per experiment.

Integration of transcriptomic and proteomic dataset.

For the 6,740 UniProt protein identifiers from the proteomics dataset, 6,732 could be mapped to an ensembl gene identifier using the BioMart data mining tool as well as manual annotation. Accordingly, for 20,158 out of 33,849 ensembl gene identifiers from the transcriptomic dataset a UniProt protein identifier could be identified. For data integration, the theoretically observable tryptic peptide normalized protein abundances were log10 transformed followed by z-score normalization. TPM values for genes that encode for identical proteins were summed, log10 transformed and finally z-score normalized. Finally, the normalized proteomics and transcriptomics data sets were combined by UniProt identifiers resulting in gene expression data for 6,714 genes.

To identify temporal gene expression modules, the R bioconductor package maSigPro version 1.68.00 for time-series was used (Nueda et al., 2014). Briefly, a first regression fit for each of the 6,714 genes was performed and significant genes selected at a false discovery rate of 0.05. For the remaining 5,978 significant genes, a second stepwise regression fit was performed to identify profile differences between experimental groups. Regression models at R-squared > 0.6 were found for 3,668 genes, which were hierarchical clustered. The optimal number of clusters were estimated using the factoextra package for R and the ‘average silhouette width’, ‘total within sum of square’ and ‘gap statistics’ methods. Finally, different numbers of clusters (6 to 12) were manually evaluated with regards to their biologically meaningfulnes using for example gene gene enrichment analysis tools. Finally, the 3,668 genes were clustered into nine gene expression modules and gene enrichment analysis performed using the R bioconductor package clusterProfiler (Yu et al., 2012).

Transcript feature analysis.

For each cluster, we considered all the transcripts of the member genes detected in our RNA-seq dataset (mean TPM >1) for the analysis. For transcript feature analysis, the Ensembl release sep2019.archive.ensembl.org in biomart R package was used to derive lengths of transcripts, coding sequence and UTRs. RNA structure features were determined using a Python program available through GitHub at `https://github.com/stephenfloor/tripseq-analysis` (Blair et al., 2017). The structure was computed using RNALfold from the ViennaRNA package (Lorenz et al., 2011) in a 75-nt window.

Analysis of RNA-binding protein motif enrichment.

Analysis of enrichment or depletion of RBP motifs in 5’UTR and 3’UTR sequences was performed by transcript set motif analysis (TSMA) analysis with k-mer method and standard settings of Transite tool (Krismer et al., 2020) `https://transite.mit.edu/`. Enrichment was calculated per module over the entire gene-set. Adjusted p-value of 0.05 was used as the significance threshold.

Analysis of published tripseq data.

We analyzed the previously published data for quantified expression values for TrIP-seq polysome profiling ((Blair et al., 2017), Supplementary Table 4). The data was filtered for neural progenitors (progenitors) and 50-day old neurons (neurons) as well as for 5’TOP genes or genes in modules 1 or 9 and all TPM values were log2 transformed.

Analysis of relative stability of transcripts.

The analysis was performed according to a method that removes the bias from expression level differences (Alkallas et al., 2017). In brief, mRNA stability estimates were calculated using REMBRANDTS `https://github.com/csglab/REMBRANDTS`. Preprocessing was performed following the CIRES workflow `https://github.com/csglab/CRIES` applying hisat2 v.2.1.0 in read alignment, samtools v1.10 with parameters -F 1548 -q 30 in alignment-filtering, htseq v0.11.2 with parameter "intersection-strict" for exonic read summarization, and with parameter "union" for intronic read counting. REMBRANDTS.sh with linear biasMode, 0.99 stringency cutoff, Rv3.6.2 and DESeq2 v1.26.0 was run to obtain the per sample differential mRNA stability estimates.

References

AAlkallas, R., Fish, L., Goodarzi, H., and Najafabadi, H.S. (2017). Inference of RNA decay rate from transcriptional profiling highlights the regulatory programs of Alzheimer's disease. Nat Commun 8, 909.

Andrews, M.G., Subramanian, L., and Kriegstein, A.R. (2020). mTOR signaling regulates the morphology and migration of outer radial glia in developing human cortex. Biorxiv 27, 2020.05.14.095117.

Bagley, J.A., Reumann, D., Bian, S., Levi-Strauss, J., and Knoblich, J.A. (2017). Fused cerebral organoids model interactions between brain regions. Nat Methods 14, 743–751.

Barbosa, C., Peixeiro, I., and Romão, L. (2013). Gene Expression Regulation by Upstream Open Reading Frames and Human Disease. Plos Genet 9, e1003529.

Baser, A., Skabkin, M., Kleber, S., Dang, Y., Balta, G.S.G., Kalamakis, G., Göpferich, M., Ibañez, D.C., Schefzik, R., Lopez, A.S., et al. (2019). Onset of differentiation is post-transcriptionally controlled in adult neural stem cells. Nature 566, 100–104.

Benito-Kwiecinski, S., Giandomenico, S.L., Sutcliffe, M., Riis, E.S., Freire-Pritchett, P., Kelava, I., Wunderlich, S., Martin, U., Wray, G.A., McDole, K., et al. (2021). An early cell shape transition drives evolutionary expansion of the human forebrain. Cell 184, 2084-2102.e19.

Bindels, D.S., Haarbosch, L., Weeren, L. van, Postma, M., Wiese, K.E., Mastop, M., Aumonier, S., Gotthard, G., Royant, A., Hink, M.A., et al. (2017). mScarlet: a bright monomeric red fluorescent protein for cellular imaging. Nat Methods 14, 53–56.

Blair, J.D., Hockemeyer, D., Doudna, J.A., Bateup, H.S., and Floor, S.N. (2017). Widespread Translational Remodeling during Human Neuronal Differentiation. Cell Rep 21, 2005–2016.

Blair, J.D., Hockemeyer, D., and Bateup, H.S. (2018). Genetically engineered human cortical spheroid models of tuberous sclerosis. Nat Med 24, 1568–1578.

Boom, R., Sol, C.J., Salimans, M.M., Jansen, C.L., Dillen, P.M.W., and Noordaa, J. van der (1990). Rapid and simple method for purification of nucleic acids. J Clin Microbiol 28, 495–503.

Cadwell, C.R., Bhaduri, A., Mostajo-Radji, M.A., Keefe, M.G., and Nowakowski, T.J. (2019). Development and Arealization of the Cerebral Cortex. Neuron 103, 980–1004.

Chau, K.F., Shannon, M.L., Fame, R.M., Fonseca, E., Mullan, H., Johnson, M.B., Sendamarai, A.K., Springel, M.W., Laurent, B., and Lehtinen, M.K. (2018). Downregulation of ribosome biogenesis during early forebrain development. Elife 7, e36998.

Chen, G., Gulbranson, D.R., Hou, Z., Bolin, J.M., Ruotti, V., Probasco, M.D., Smuga-Otto, K., Howden, S.E., Diol, N.R., Propson, N.E., et al. (2011). Chemically defined conditions for human iPSC derivation and culture. Nat Methods 8, 424–429.

Cockman, E., Anderson, P., and Ivanov, P. (2020). TOP mRNPs: Molecular Mechanisms and Principles of Regulation. Biomol 10, 969.

Crino, P.B. (2016). The mTOR signalling cascade: paving new roads to cure neurological disease. Nature Reviews. Neurology 12, 379–392.

Damgaard, C.K., and Lykke-Andersen, J. (2011). Translational coregulation of 5TOP mRNAs by TIA-1 and TIAR. Gene Dev 25, 2057–2068.

Djuric, U., Rodrigues, D.C., Batruch, I., Ellis, J., Shannon, P., and Diamandis, P. (2017). Spatiotemporal Proteomic Profiling of Human Cerebral Development. Molecular Cellular Proteomics 16, 1548–1562.

Dorfer, V., Pichler, P., Stranzl, T., Stadlmann, J., Taus, T., Winkler, S., and Mechtler, K. (2014). MS Amanda, a Universal Identification Algorithm Optimized for High Accuracy Tandem Mass Spectra. J Proteome Res 13, 3679–3684.

Edgar, R., Domrachev, M., and Lash, A.E. (2002). Gene Expression Omnibus: NCBI gene expression and hybridization array data repository. Nucleic Acids Res 30, 207–210.

Esk, C., Lindenhofer, D., Haendeler, S., Wester, R.A., Pflug, F., Schroeder, B., Bagley, J.A., Elling, U., Zuber, J., Haeseler, A. von, et al. (2020). A human tissue screen identifies a regulator of ER secretion as a brain-size determinant. Science 370, 935–941.

Eze, U.C., Bhaduri, A., Haeussler, M., Nowakowski, T.J., and Kriegstein, A.R. (2021). Single-cell atlas of early human brain development highlights heterogeneity of human neuroepithelial cells and early radial glia. Nat Neurosci 24, 584–594.

Floor, S.N., and Doudna, J.A. (2016). Tunable protein synthesis by transcript isoforms in human cells. Elife 5, 1276.

Fonseca, B.D., Smith, E.M., Yelle, N., Alain, T., Bushell, M., and Pause, A. (2014). The ever-evolving role of mTOR in translation. Seminars in Cell and Developmental Biology 36, 102–112.

Greig, L.C., Woodworth, M.B., Galazo, M.J., Padmanabhan, H., and Macklis, J.D. (2013). Molecular logic of neocortical projection neuron specification, development and diversity. Nat Rev Neurosci 14, 755–769.

Hagemann-Jensen, M., Ziegenhain, C., Chen, P., Ramsköld, D., Hendriks, G.-J., Larsson, A.J.M., Faridani, O.R., and Sandberg, R. (2020). Smart-seq3 Protocol v3.

Harnett, D., Ambrozkiewicz, M.C., Zinnall, U., Rusanova, A., Borisova, E., Dannenberg, R., Imami, K., Münster-Wandowski, A., Fauler, B., Mielke, T., et al. (2021). A critical period of translational control during brain development at codon resolution. Biorxiv 2021.06.23.449626.

He, Z., Maynard, A., Jain, A., Gerber, T., Petri, R., Lin, H.-C., Santel, M., Ly, K., Dupré, J.-S., Sidow, L., et al. (2020). Lineage recording in human cerebral organoids. Nat Methods 19, 90–99.

Hockemeyer, D., Wang, H., Kiani, S., Lai, C.S., Gao, Q., Cassady, J.P., Cost, G.J., Zhang, L., Santiago, Y., Miller, J.C., et al. (2011). Genetic engineering of human pluripotent cells using TALE nucleases. Nat Biotechnol 29, 731–734.

Hoye, M.L., and Silver, D.L. (2020). ScienceDirect Decoding mixed messages in the developing cortex: translational regulation of neural progenitor fate. Current Opinion in Neurobiology 66, 93–102.

Iwata, R., and Vanderhaeghen, P. (2021). Regulatory roles of mitochondria and metabolism in neurogenesis. Curr Opin Neurobiol 69, 231–240.

Käll, L., Canterbury, J.D., Weston, J., Noble, W.S., and MacCoss, M.J. (2007). Semi-supervised learning for peptide identification from shotgun proteomics datasets. Nat Methods 4, 923–925.

Kim, M.-S., Pinto, S.M., Getnet, D., Nirujogi, R.S., Manda, S.S., Chaerkady, R., Madugundu, A.K., Kelkar, D.S., Isserlin, R., Jain, S., et al. (2014). A draft map of the human proteome. Nature 509, 575–581.

Koyanagi, M., Asahara, S., Matsuda, T., Hashimoto, N., Shigeyama, Y., Shibutani, Y., Kanno, A., Fuchita, M., Mikami, T., Hosooka, T., et al. (2011). Ablation of TSC2 Enhances Insulin Secretion by Increasing the Number of Mitochondria through Activation of mTORC1. Plos One 6, e23238.

Krismer, K., Bird, M.A., Varmeh, S., Handly, E.D., Gattinger, A., Bernwinkler, T., Anderson, D.A., Heinzel, A., Joughin, B.A., Kong, Y.W., et al. (2020). Transite: A Computational Motif-Based Analysis Platform That Identifies RNA-Binding Proteins Modulating Changes in Gene Expression. Cell Rep 32, 108064.

Laguesse, S., Creppe, C., Nedialkova, D.D., Prévot, P.-P., Borgs, L., Huysseune, S., Franco, B., Duysens, G., Krusy, N., Lee, G., et al. (2015). A Dynamic Unfolded Protein Response Contributes to the Control of Cortical Neurogenesis. Dev Cell 35, 553–567.

Lancaster, M.A., Corsini, N.S., Wolfinger, S., Gustafson, E.H., Phillips, A.W., Burkard, T.R., Otani, T., Livesey, F.J., and Knoblich, J.A. (2017). Guided self-organization and cortical plate formation in human brain organoids. Nat Biotechnol 35, 659–666.

Lee, J.-A., Damianov, A., Lin, C.-H., Fontes, M., Parikshak, N.N., Anderson, E.S., Geschwind, D.H., Black, D.L., and Martin, K.C. (2016). Cytoplasmic Rbfox1 Regulates the Expression of Synaptic and Autism-Related Genes. Neuron 89, 113–128.

Lennox, A.L., Mao, H., and Silver, D.L. (2017). RNA on the brain: emerging layers of post-transcriptional regulation in cerebral cortex development. Wiley Interdisciplinary Reviews: Developmental Biology 7, e290-18.

Lorenz, R., Bernhart, S.H., Siederdissen, C.H. zu, Tafer, H., Flamm, C., Stadler, P.F., and Hofacker, I.L. (2011). ViennaRNA Package 2.0. Algorithm Mol Biol 6, 26.

Ma, W.-J., Chung, S., and Furneaux, H. (1997). The Elav-like proteins bind to AU-rich elements and to the poly(A) tail of mRNA. Nucleic Acids Res 25, 3564–3569.

Markmiller, S., Soltanieh, S., Server, K.L., Mak, R., Jin, W., Fang, M.Y., Luo, E.-C., Krach, F., Yang, D., Sen, A., et al. (2018). Context-Dependent and Disease-Specific Diversity in Protein Interactions within Stress Granules. Cell 172, 590-604.e13.

Martin, E.T., Blatt, P., Nguyen, E., Lahr, R., Selvam, S., Yoon, H.A.M., Pocchiari, T., Emtenani, S., Siekhaus, D.E., Berman, A., et al. (2022). A translation control module coordinates germline stem cell differentiation with ribosome biogenesis during Drosophila oogenesis. Dev Cell 57, 883-900.e10.

Martins, M., Galfrè, S., Terrigno, M., Pandolfini, L., Appolloni, I., Dunville, K., Marranci, A., Rizzo, M., Mercatanti, A., Poliseno, L., et al. (2021). A eutherian-specific microRNA controls the translation of Satb2 in a model of cortical differentiation. Stem Cell Rep 16, 1496–1509.

Mathews, D.H., and Turner, D.H. (2006). Prediction of RNA secondary structure by free energy minimization. Curr Opin Struc Biol 16, 270–278.

McClure-Begley, T.D., Ebmeier, C.C., Ball, K.E., Jacobsen, J.R., Kogut, I., Bilousova, G., Klymkowsky, M.W., and Old, W.M. (2018). Cerebral organoid proteomics reveals signatures of dysregulated cortical development associated with human trisomy 21. 9, 181–63.

Melliou, S., Sangster, K.T., Kao, J., Zarrei, M., Lam, K.H.B., Howe, J., Papaioannou, M.-D., Tsang, Q.P.L., Borhani, O.A., Sajid, R.S., et al. (2022). Regionally defined proteomic profiles of human cerebral tissue and organoids reveal conserved molecular modules of neurodevelopment. Cell Reports 39, 110846.

Meyuhas, O., and Kahan, T. (2015). The race to decipher the top secrets of TOP mRNAs. BBA - Gene Regulatory Mechanisms 1849, 801–811.

Nascimento, J.M., Saia-Cereda, V.M., Sartore, R.C., Costa, R.M. da, Schitine, C.S., Freitas, H.R., Murgu, M., Reis, R.A. de M., Rehen, S.K., and Martins-de-Souza, D. (2019). Human Cerebral Organoids and Fetal Brain Tissue Share Proteomic Similarities. Frontiers in Cell and Developmental Biology 7, 303.

Nomura, M., Ohsuye, K., Mizuno, A., Sakuragawa, Y., and Tanaka, S. (1984). Influence of messenger RNA secondary structure on translation efficiency. Nucl Acid S 173–176.

Nueda, M.J., Tarazona, S., and Conesa, A. (2014). Next maSigPro: updating maSigPro bioconductor package for RNA-seq time series. Bioinformatics 30, 2598–2602.

Otsuka, H., Fukao, A., Funakami, Y., Duncan, K.E., and Fujiwara, T. (2019). Emerging Evidence of Translational Control by AU-Rich Element-Binding Proteins. Frontiers Genetics 10, 332.

Parekh1, S., Ziegenhain, C., Vieth, B., Enard, W., and Hellmann, I. (2018). zUMIs - A fast and flexible pipeline to process RNA sequencing data with UMIs. Gigascience 7, giy059-.

Park, J.-E., Yi, H., Kim, Y., Chang, H., and Kim, V.N. (2016). Regulation of Poly(A) Tail and Translation during the Somatic Cell Cycle. Mol Cell 62, 462–471.

Pașca, A.M., Park, J.-Y., Shin, H.-W., Qi, Q., Revah, O., Krasnoff, R., O'Hara, R., Willsey, A.J., Palmer, T.D., and

Paşca, S.P. (2019). Human 3D cellular model of hypoxic brain injury of prematurity. Nat Med 25, 784–791.

Perez-Riverol, Y., Bai, J., Bandla, C., García-Seisdedos, D., Hewapathirana, S., Kamatchinathan, S., Kundu, D.J., Prakash, A., Frericks-Zipper, A., Eisenacher, M., et al. (2021). The PRIDE database resources in 2022: a hub for mass spectrometry-based proteomics evidences. Nucleic Acids Res 50, D543–D552.

Philippe, L., Elzen, A.M.G. van den, Watson, M.J., and Thoreen, C.C. (2020). Global analysis of LARP1 translation targets reveals tunable and dynamic features of 5' TOP motifs. Proceedings of the National Academy of Sciences of the United States of America 117, 5319–5328.

Polioudakis, D., Torre-Ubieta, L. de la, Langerman, J., Elkins, A.G., Shi, X., Stein, J.L., Vuong, C.K., Nichterwitz, S., Gevorgian, M., Opland, C.K., et al. (2019). A Single-Cell Transcriptomic Atlas of Human Neocortical Development during Mid-gestation. Neuron 103, 785-801.e8.

Ran, F.A., Hsu, P.D., Wright, J., Agarwala, V., Scott, D.A., and Zhang, F. (2013). Genome engineering using the CRISPR-Cas9 system. Nat Protoc 8, 2281–2308.

Schindelin, J., Arganda-Carreras, I., Frise, E., Kaynig, V., Longair, M., Pietzsch, T., Preibisch, S., Rueden, C., Saalfeld, S., Schmid, B., et al. (2012). Fiji: an open-source platform for biological-image analysis. Nat Methods 9, 676–682.

Schwanhäusser, B., Busse, D., Li, N., Dittmar, G., Schuchhardt, J., Wolf, J., Chen, W., and Selbach, M. (2011). Global quantification of mammalian gene expression control. Nature 473, 337–342.

Shibata, M., Gulden, F.O., and Sestan, N. (2015). From trans to cis: transcriptional regulatory networks in neocortical development. Trends Genet 31, 77–87.

Stuart, T., Butler, A., Hoffman, P., Hafemeister, C., Papalexi, E., Mauck, W.M., Hao, Y., Stoeckius, M., Smibert, P., and Satija, R. (2019). Comprehensive Integration of Single-Cell Data. Cell 177, 1888-1902.e21.

Switon, K., Kotulska, K., Janusz-Kaminska, A., Zmorzynska, J., and Jaworski, J. (2017). Molecular neurobiology of mTOR. Neuroscience 341, 112–153.

Taus, T., Kocher, T., Pichler, P., Paschke, C., Schmidt, A., Henrich, C., and Mechtler, K. (2011). Universal and Confident Phosphorylation Site Localization Using phosphoRS. J Proteome Res 10, 5354–5362.

Vaid, S., and Huttner, W.B. (2020). Transcriptional Regulators and Human-Specific/Primate-Specific Genes in Neocortical Neurogenesis. International Journal of Molecular Sciences 21, 4614–4619.

Varderidou-Minasian, S., Verheijen, B.M., Schatzle, P., Hoogenraad, C.C., Pasterkamp, R.J., and 0000-0001-5093-5945, M.A.O. (2020). Deciphering the Proteome Dynamics during Development of Neurons Derived from Induced Pluripotent Stem Cells. Journal of Proteome Research 19, 2391–2403.

Wang, T., Tian, X., Kim, H.B., Jang, Y., Huang, Z., Na, C.H., and Wang, J. (2022). Intracellular energy controls dynamics of stress-induced ribonucleoprotein granules. Nat Commun 13, 5584.

Wilbertz, J.H., Voigt, F., Horvathova, I., Roth, G., Zhan, Y., and Chao, J.A. (2019). Single-Molecule Imaging of mRNA Localization and Regulation during the Integrated Stress Response. Mol Cell 73, 946-958.e7.

Yamashita, R., Suzuki, Y., Takeuchi, N., Wakaguri, H., Ueda, T., Sugano, S., and Nakai, K. (2008). Comprehensive detection of human terminal oligo-pyrimidine (TOP) genes and analysis of their characteristics. Nucleic Acids Research 36, 3707–3715.

Yang, G., Smibert, C.A., Kaplan, D.R., and Miller, F.D. (2014). An eIF4E1/4E-T Complex Determines the Genesis of Neurons from Precursors by Translationally Repressing a Proneurogenic Transcription Program. Neuron 84, 723–739.

Yu, G., Wang, L.-G., Han, Y., and He, Q.-Y. (2012). clusterProfiler: an R Package for Comparing Biological Themes Among Gene Clusters. Omics J Integr Biology 16, 284–287.

Zahr, S.K., Yang, G., Kazan, H., Borrett, M.J., Yuzwa, S.A., Voronova, A., Kaplan, D.R., and Miller, F.D. (2018). A Translational Repression Complex in Developing Mammalian Neural Stem Cells that Regulates Neuronal Specification. Neuron 97, 520-537.e6.

Zahr, S.K., Kaplan, D.R., and Miller, F.D. (2019). Translating neural stem cells to neurons in the mammalian brain. Cell Death Differentiation 1–18.

Supplementary methods

Plasmid constructs and cloning.

All the oligos used in the manuscript are listed in Supplementary file 1. All PCRs were performed using Thermo Scientific™ Phusion Hot Start II DNA-Polymerase (New England Biolabs, M0535S).

Plasmids for CRISPR-Cas9 gene editing.

For cloning guide sequences against SOX2, G3BP1 and TSC2 loci, corresponding DNA oligos containing the guide sequences were phosphorylated and hybridized to make double stranded DNA (ds-DNA) fragments. These ds-DNA fragments were cloned into pSpCas9(BB)-2A-GFP plasmid (Addgene, 48138, RRID:Addgene_48138), modified to express eCas9 instead of WT-Cas9 and dTomato instead of GFP (Esk et al., 2020), using the Bbs1 cloning strategy of Ran et al. (Ran et al., 2013). See Supplementary file 1 for the guide and oligo sequences.

Homology plasmids.

Sox2-P2A-EGFP homology construct:
Homology construct for tagging SOX2 locus with P2A-EGFP was assembled by Gibson cloning. A vector backbone containing Diphtheria toxin A cassette, as a negative selection marker, was used. Left and right homology arms of the SOX2 locus were amplified from H9 gDNA with primers in which the native stop codon was removed. The P2A-EGFP fragment was amplified with the mentioned primers. The intermediate plasmid was used for mutagenesis PCR to mutate the guide cutting base from G to A to avoid cutting of the repaired clones.

AAVS1 hSyn1::dtomato homology construct:
Donor plasmid was constructed to insert the following cassette SA-2A-puro-PA-2xCHS4-hSYN1-INTRON-tdTomato-WPRE-SV40-2xCHS4. For this purpose, a previously generated backbone containing 2xCHS4-EF1a-tdTomato-SV40-2xCHS4 was used (Bagley et al., 2017). The promoter was exchanged with hSYN1 promoter.

tagBFP reporter constructs (5'TOP reporter and non 5'TOP reporter):
Donor plasmids for Doxycycline-inducible 5'TOP and non5'TOP reporters were generated by gibson assembly to insert the following cassettes 1) Neo-TetO-minCMV-5'TOPUTR-INTRON-tagBFP-bGHpA-CAG-TetR and 2) Neo-TetO-minCMV-non5'TOPUTR-INTRON-tagBFP-bGHpA-CAG-TetR respectively. pAAVS1-Neo-TRE-CMV-Cre-rtTA, containing the Tet-response element and CAG:Tet repressor sequence, was a gift from Madeline Lancaster and was used as backbone upon digestion with Kpn1 and Sal1 (Addgene, 165457, RRID:Addgene_165457) (Benito-Kwiecinski et al., 2021). The 5'TOP sequence of RPL32 with CMV minimal promoter and chimeric intron was amplified from plasmid pSF4-TetCMV-5'TOP-intron-20xGCN4-Renilla-FKBP-Stop-24xMS2v5-SV40-CTE-polyA (Addgene, 19946, RRID:Addgene_119946), which was a gift from Jeffrey Chao (Wilbertz et al., 2019). For the non-5'TOP, the pyrimidine-rich sequences in the 5' terminus of the RPL32 sequence were mutated and ordered as a gblock.

G3BP1 homology construct:
The plasmid HR_G3BP1-V5-APEX2-GFP (Addgene, 105284, RRID:Addgene_105284) was a gift from the Eugene Yeo lab (San Diego, USA)(Markmiller et al., 2018) and the plasmid pmScarlet-i_C1 was a gift from Dorus Gadella (Addgene, 85044) (Bindels et al., 2017). To generate the plasmid pHDR3BP1-V5-APEX2-mScarlet-I, used for CRISPR mediated endogenous tagging of G3BP1 with APEX2-mScarlet, the coding sequence for CopGFP was exchanged with the sequence of the monomeric red fluorescent protein mScarlet (Bindels et al., 2017).

Reporter line generation.

All reporter lines were generated by nucleofection of the plasmid DNA with the Amaxa nucleofector 2b device (Lonza, AAB-1001) and Human Stem Cell Nucleofector Kit 1 (Lonza, VPH-5012) following manufacturer's guidelines. For each nucleofection one million single cells were used. For genotyping, DNA was extracted using the QuickExtract DNA Extraction Solution (Cambio QE09050) and a PCR assay was performed to identify correctly edited clones. See Table S9 for oligo sequences used for genotyping. Genomic integrity of the final clones was confirmed by SNP arrays.

Insertion into the AAVS1 safe-harbor locus was performed using TALEN technology as described before (Bagley et al., 2017; Hockemeyer et al., 2011). For this purpose, the nucleofection mix containing 0.5 µg of each of the TALEN plasmids and 1.5 µg of each of the donor plasmids was used. Nucleofected cells were grown for four to seven days and then selected with appropriate antibiotics. For puromycin resistance reporter: 1 µg/ml Puromycin dihydrochloride (Thermo Fisher Scientific, A1113803) and for neomycin resistance reporter 250 µg/ml Geneticin/G418 Sulfate (Thermo Fisher Scientific, 10131035). Surviving colonies were picked manually, transferred into 24-well plates and further expanded for genotyping and cryopreservation.

For generating FF H9 SOX2::SOX2-P2A-EGFP reporter line, the nucleofection mix included 1 µg guide-Cas9 plasmid and 4 µg homology repair construct. 7 days post nucleofection, surviving cells were FACSed to obtain GFP positive cells which were sorted as single cells and further expanded for genotyping and cryopreservation. Positive clones were selected by performing PCR to verify genomic insertion of P2A-EGFP.

For generating the dual reporter line, FF H9 SOX2::SOX2-P2A-EGFP reporter line clone was used. Clone 10.1 and clone 10.2 were used for the transcriptome and proteome datasets. For generating 5'TOP and non-5'TOP reporter lines, FF H9 hESCs and Rozh-5 iPSCs carrying monoallelic insertion of pAAVS1(LH)-SA-2A-puro-7xTetO-minCMV-iRFP-pA-AAVS1(RH) were used. Cells resistant to puromycin and neomycin were selected for tag-BFP reporter assay. This strategy ensured monoallelic insertion of the tag-BFP reporter.

For generating G3BP1 reporter line, FF H9 cells were electroporated with 0.8 µg guide-Cas9 plasmid and 1 µg homology repair construct. The HDR plasmid carried in addition to the APEX2-mScarlet fusion, also a floxed puromycin resistance cassette under the control of the EF1a core promoter. 48 h post electroporation, cells were selected with 0.5 µg/ml puromycin and correct gene editing was confirmed by genotyping PCRs and Sanger sequencing. Following, a pool of the correct clones was nucleofected as before with 2 µg Cre-recombinase-2A-EGFP expressing plasmid and 1 µg of pmaxGFP to remove the EF1a-puromycin cassette. Successfully electroporated cells were FACS sorted for GFP and mScarlet double positive cells. Excision of the EF1a-puromycin cassette was confirmed by genotyping PCRs and Sanger sequencing.

For generating TSC2 KO hPSCs, nucleofection was performed with 3 µg guide-Cas9 plasmid. The surviving cells were FACSed the next day for dTomato fluorescence and sorted as single cells into 96-well plates. The single cell clones were expanded and genotyped to identify TSC2 homozygous mutants by Sanger sequencing.

RNA extraction, library generation and RNA-seq.

FACSed cells were pelleted and used for RNA isolation. Total RNA was isolated using a lysis step based on guanidine thiocyanate (adapted from (Boom et al., 1990)) with DNaseI digestion using magnetic beads (GE Healthcare, 65152105050450) applied on the KingFisher instrument (Thermo Fisher Scientific). The purified RNA was used to generate the RNA-seq library.

RNA sequencing libraries were generated using 500 ng of RNA using NEBNext® Ultra™ II Directional RNA Library Prep Kit for Illumina (New England BIolabs, E7760). Isolation of mRNA was done using the NEBNext® Poly(A) mRNA Magnetic Isolation Module (New England BIolabs, E7490). The size of libraries was assessed using NGS HS analysis kit and a fragment analyzer system (Agilent). Library concentrations were quantified with KAPA Kit (Roche) and 75-bp paired end-read sequencing was performed using the Illumina NextSeq550 platform.

RNA-seq reads were trimmed using trim-galore v0.5.0, and reads mapping to abundant sequences included in the iGenomes UCSC hg38 reference (human rDNA, human mitochondrial chromosome, phiX174 genome, adapter) were removed using bowtie2 v2.3.4.1 alignment. Remaining reads were analyzed using genome and gene annotation for the GRCh38/hg38 assembly obtained from Homo sapiens Ensembl release 94. Reads were aligned to the genome using star v2.6.0c and reads in genes were counted with featureCounts (subread v1.6.2). Differential gene expression analysis on raw counts and variance-stabilized transformation of count data for heatmap visualization were performed using DESeq2 v1.18.1.

Published scRNA-seq dataset (Eze et al., 2021) was analysed using Seurat package (Stuart et al., 2019) to obtain 5'TOP module scores in fetal cell types.

Proteomics.

A minimum of 2 million FACSed cells were pelleted, snap frozen in liquid nitrogen and stored at -80°C until processed. Each cell pellet was resuspended in 72.5 µL 10M Urea 50mM HCl and incubated 10 min at room temperature. Then 7.5 µL 1M TEAB (triethylammonium bicarbonate) buffer (pH 8) was added and total protein amount determined by Bradford assays. Following, 1 µL Benzonase and 1.25 µL 1M DTT (dithiothreitol) was added to each sample and incubated for 37°C for 1 h. Protein alkylation was performed with 2.5 µL of 1M IAA (iodoacetamide) and incubated for 30 min in the dark at room temperature and the reaction quenched with 0.6 µL 1M DTT. All samples were diluted with 100 mM TEAB Buffer to a urea concentration of 6M and proteins digested with LysC (Wako) for 3 h at 37ºC using an enzyme to protein ratio of 1:50. After the first digestion with LysC the samples were diluted with 100 mM TRIS buffer to a final Urea concentration of 2 M. The final tryptic digestion (Trypsin Gold, Promega) was performed with an enzyme to protein ratio of 1:50 overnight at 37ºC. To evaluate the digest efficiency 100 ng of all samples were analyzed on a monolithic HPLC system. Following, peptides were desalted using C18 cartridges (Sep-Pak Vac 1cc (50mg), Waters) that were equilibrated with 500 µL methanol/water 50:50, 2x 0.5mL 80% acetonitrile (ACN) 0.1% formic acid (FA) and 3x 0.5 mL 0.1% trifluoroacetic acid (TFA) applying gentle pressure with compressed air. The pH of all samples was adjusted to pH<2 with 10% TFA before they were loaded on the columns. All columns were washed with 6x 0.5 mL 0.1% TFA applying gentle pressure using compressed air. To elute the sample 2x 200 µL of 80% ACN and 0.1% FA were used and to remove the solvent completely from the column, gentle pressure was applied with compressed air. The organic content of the eluates was removed by using a SpeedVac vacuum concentrator and samples were diluted to a volume of 200 µL with 0.1% FA, snap frozen with liquid nitrogen and lyophilized overnight. The dried samples were dissolved in 100 µL 100 mM HEPES and quantified using a monolithic HPLC system. One bridge channel was produced from all FACS sorted cells. In these channels the corresponding samples were mixed in a 1 to 1 ratio.

For TMT labeling, 20 µg of peptides (in 100 µl 100 mM Hepes pH 7.6) were labeled with distinct channels of the TMTpro16 plex reagent (Thermo Fisher, TMT16plex) according to the manufacturer's description. Labeling efficiency was determined by LC-MS/MS on a small aliquot of each sample. Samples were mixed in equimolar amounts and equimolarity was evaluated again by LC-MS/MS. The mixed sample was acidified to a pH below 2 with 10% TFA and was desalted using C18 cartridges (Sep-Pak Vac 1cc (50mg), Waters). Peptides were eluted with 3 x 150 ml 80% Acetonitrile (ACN) and 0.1% Formic Acid (FA), followed by freeze-drying. The dried samples were dissolved in 70 µl of SCX Buffer A (5 mM NaH2PO4, pH 2.7, 15% ACN) and 200 µg of peptide were loaded on the column. SCX was performed using a custom-made TSK gel SP-2PW SCX column (5 µm particles, 12.5 nm pore size, 1 mm i.d. x 250 mm, TOSOH) on an Ultimate system (Thermo Fisher Scientific) at a flow rate of 35 µl/min. For the separation, a ternary gradient was used starting with 100% buffer A for 10 min, followed by a linear increase to 10% buffer B (5 mM NaH2PO4, pH 2.7, 1M NaCl, 15% ACN) and 50% buffer C (5 mM Na2HPO4, pH 6, 15% ACN) in 10 min, to 25% buffer B and 50% buffer C in 10 min, 50% buffer B and 50% buffer C in 5 min and an isocratic elution for further 15 min. The flow-through was collected as single fractions, along with the gradient fractions that were collected every minute. In total 60 fractions were collected, and low abundant fractions were pooled. ACN was removed by vacuum centrifugation and the samples were acidified with 0.1% TFA and analyzed by LC-MS/MS. The UltiMate 3000 HPLC RSLC nano system (Thermo Scientific) was coupled to a Q Exactive HF-X mass spectrometer (Thermo Scientific), equipped with a Proxeon nanospray source (Thermo Scientific). Peptides were loaded onto a trap column (Thermo Scientific, PepMap C18, 5 mm × 300 µm ID, 5 µm particles, 100 Å pore size) at a flow rate of 25 µL/min using 0.1% TFA as mobile phase. After 10 min, the trap column was switched in line with the analytical column (Thermo Scientific, PepMap C18, 500 mm × 75 µm ID, 2 µm, 100 Å). Peptides were eluted using a flow rate of 230 nl/min and a binary 120 min gradient. The two-step gradient started with the mobile phases: 98%A (water/formic acid, 99.9/0.1, v/v) and 2% B (water/acetonitrile/formic acid, 19.92/80/0.08, v/v/v) increased to 35% B over the next 120 min, followed by a gradient in 5 min to 90% B, stayed there for five min and decreased in 2 min back to the gradient 95% A and 2% B for equilibration at 30°C.

The Q Exactive HF-X mass spectrometer was operated in data-dependent mode, using a full scan (m/z range 375-1650, nominal resolution of 120,000, target value 3e6) followed by MS/MS scans of the 10 most abundant ions. MS/MS spectra were acquired using normalized collision energy of 35, isolation width of 0.7 m/z, resolution of 45.000, a target value of 1e5 and maximum fill time of 250 ms. For the detection of the TMT reporter ions a fixed first mass of 110 m/z was set for the MS/MS scans. Precursor ions selected for fragmentation (exclude charge state 1, 7, 8, >8) were put on a dynamic exclusion list for 60s. Additionally, the minimum AGC target was set to 1e4 and intensity threshold was calculated to be 4e4. The peptide match feature was set to preferred and the exclude isotopes feature was enabled.

Proteomics data processing.

For peptide identification, the RAW-files were loaded into Proteome Discoverer (version 2.4.0.305, Thermo Scientific). All hereby created MS/MS spectra were searched using MSAmanda v2.0.0.13248 (Dorfer et al., 2014). The RAW-files were searched against the human uniprot-reference-database (20,541 sequences; 11,395,640 residues). The following search parameters were used: Iodoacetamide derivative on cysteine was set as fixed modification, deamidation on asparagine and glutamine, oxidation on methionine, TMTpro-16plex tandem mass tag on lysine, phosphorylation on serine, threonine and tyrosine as well as carbamylation on peptide N-Terminus, TMTpro-16plex tandem mass tag on peptide N-Terminus, acetylation on protein N-Terminus were set as variable modifications. Monoisotopic masses were searched within unrestricted protein masses for tryptic enzymatic specificity. The peptide mass tolerance was set to ±5 ppm and the fragment mass tolerance to ±15 ppm. The maximal number of missed cleavages was set to 2. The result was filtered to 1% FDR on peptide-spectrum-match and protein level using Percolator (Käll et al., 2007) algorithm integrated in Thermo Proteome Discoverer. The localization of the modification sites within the peptides was performed with the tool ptmRS, which is based on phosphoRS (Taus et al., 2011). Peptides were quantified based on Reporter Ion intensities extracted by the "Reporter Ions Quantifier"-node implemented in Proteome Discoverer. Proteins were quantified by summing reporter intensities of spectra with less than 50% isolation interference from unique peptides. Inspired by the iBAQ algorithm (Schwanhäusser et al., 2011), protein abundances were normalized by dividing summarized reporter intensities by the number of theoretically observable tryptic peptides with a length between 7 and 30 amino acids. TMT batches were normalized using a "bridge-channel" as reference which contains a mix of all samples.

Polysome profiling by sucrose gradient fractionation.

Polysome profiling was performed for 3 batches of organoid differentiation using Rozh5 WT and Rozh5 TSC2-/- 4G11 hiPSC clones. For this procedure, all solutions and materials contained 100 µg/ml cycloheximide (CHX), were DPEC-treated if possible, and were pre-chilled at 4°C for at least 30 min before use. Before harvest, organoids were treated with 100 µg/ml CHX at 37°C for 10 min to arrest ribosomes and washed with ice-cold DPBS without calcium and magnesium (Thermo Fisher Scientific, 14190250). Organoids were then dissociated as described above with Accutase/Trypsin mix supplemented with 100 µg/ml CHX. All cells were pelleted at 1000 g for 10 min at 4°C. The supernatant was removed, cells were washed with 15 ml ice-cold DPBS without calcium and magnesium and pelleted at 1000 g at 4°C for 10 min. Next the pellet was resuspended in lysis buffer (20 mM TrisHCl, 30 mM MgCl2, 100 mM NaCl, 1% Triton X, 100 µg/ml CHX, 0.5 mM DTT, 1 mg/ml heparin, 100 µ/ml RNase inhibitor (Promega, NA2615) and Turbo DNase (ThermoFisher Scientific, AM2238)). Samples were triturated 10 times with a 26G needle, incubated on ice for 10 min, and spun at 21,000 g at 4°C for 20 min. 400 µl of lysate was loaded onto the sucrose gradient. Sucrose gradients of 10-50% range were prepared in 20 mM TrisHCl, 5 mM MgCl2, 140 mM NaCL with Gradient master unit (BIOCOMP, B108). Tubes were spun at 35,000 rpm for 2 h in a SW-40 rotor (Optima L-90K, Beckman Coulter). Each sample was passed through a gradient fractionator (TriaxTM Flow Cell Manual Gradient Station ip, BIOCOMP) and the A260 profile was monitored. 24 fractions with 500 µl of sample each were collected and immediately processed for RNA extraction.

RNA extraction from sucrose gradient.

Fractions were pooled into monosome, low- and high-polysome-associated mRNAs (Floor Doudna, 2016). RNA was extracted by adding equal volumes of acid Phenol/Chloroform (pH 4.7, Thermo Fisher Scientific, AM9720) to the sucrose fractions. The mix was vortexed briefly and spun down at 14,000 g at 4°C for 20 min. Subsequently, the aqueous phase was transferred to a new tube. RNA was precipitated with 300 mM NaAcetate, Ethanol and 20 µg Glycoblue (Thermo Fisher Scientific, AM9515). The sample was mixed by inversion and stored at -80°C overnight. The RNA was spun down at 14,000 g at 4°C for 20 min. The pellet was washed with 75% ethanol, vortexed briefly and pelleted at 14,000 g at 4°C for 5 min. Subsequently, the pellet was air-dried and resuspended in RNase-free water. The purified RNA was quality controlled and used for RNA-seq library preparation.

RNA-sequencing for ribosome fractions.

RNA sequencing libraries were generated using Smart-seq3 V3 protocol suitable for Illumina sequencing from low input total RNA. (Hagemann-Jensen et al., 2020) `https://www.protocols.io/view/smart-seq3-protocol-bcq4ivyw`. The method uses a template switch oligo which has UMIs and a binding site for Nextera read 1 side. Therefore the method creates UMI labelled, stranded 5' fragments and unstranded fragments all over the transcript body. The size of libraries was assessed using NGS HS analysis kit and a fragment analyzer system (Agilent). Library concentrations were quantified with KAPA Kit (Roche) and 50-bp paired-read sequencing was performed using the Illumina NovaSeq S1 platform.

Analysis of RNA-seq of ribosome fractions.

Smart-seq3 read analysis was performed using zUMI v2.9.7 (Parekh1 et al., 2018), providing the expected barcodes file, using STAR 2.7.7a with additional STAR parameters "–limitOutSJcollapsed 50000000 –limitIObufferSize 1500000000 –limitSjdbInsertNsj 2000000 –clip3pAdapterSeq CTGTCTCTTATACACATCT" and Homo sapiens Ensembl GRCh38 release 94 as a reference. zUMI generated inex UMI counts were used for downstream analyses. Differential gene association analysis between groups was performed using DESeq2 wald tests on the umi count tables after minimal pre-filtering where only protein-coding genes with at least 2 reads in multiple samples of a condition were retained.

Figure Supplements

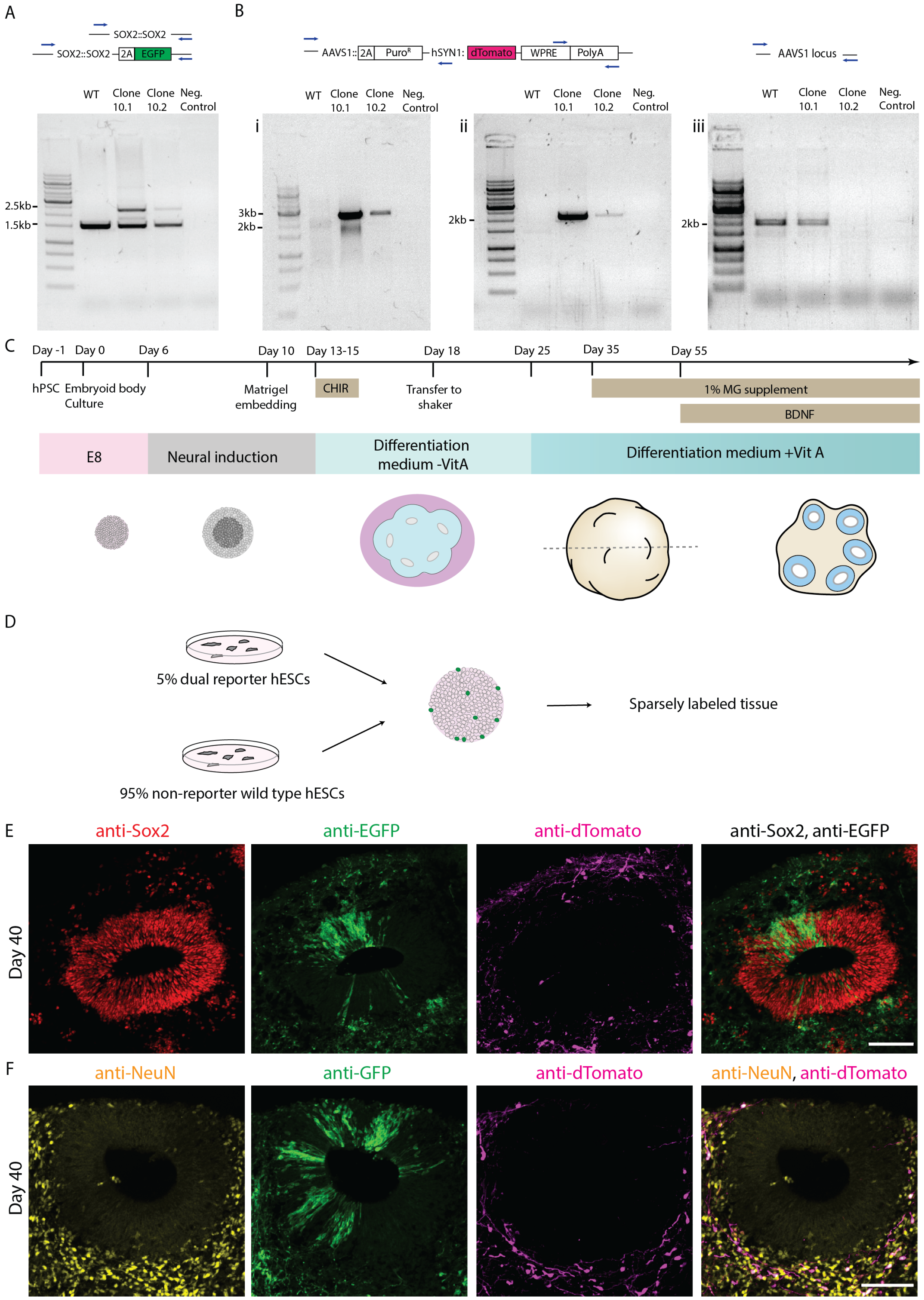


Figure 1 -figure supplement 1. Generation of a dual reporter hESC line to differentially label neural progenitors and neurons. A) Results of genotyping PCR for H9 WT hESCs, dual reporter H9 hESC clones and negative control samples to test integration of the P2A-EGFP fragment in the SOX2 genomic locus. Integration of EGFP leads to the 2.5kb band. Unedited SOX2 genomic locus is marked by the 1.5kb band. B) Results of genotyping PCRs for H9 WT hESCs, dual reporter H9 hESC clones and negative control samples to test integration of the hSYN1:dTomato-WPRE-polyA transgene into the AAVS1 locus. PCRs (i) and (ii) test the presence of the transgene. For PCR (iii) a 2kb band indicates an unedited AAVS1 locus and thus reports the zygosity of AAVS1 locus after transgene insertion. C) Schematic representation of the organoid culture protocol used in this study. D) Schematic representation of sparse labeling in organoid culture using 95% unlabeled control H9 hESCs and 5% dual reporter hESCs for embryoid body formation. E,F) Confocal scan of 40 day old sparsely labeled organoid sections immunostained with anti-SOX2 (D), anti-NeuN (E), anti-EGFP and anti-dTomato antibodies. The image shows a typical ventricular zone and surrounding neuronal zone in the organoid tissue. Scale bar = 100 µm.

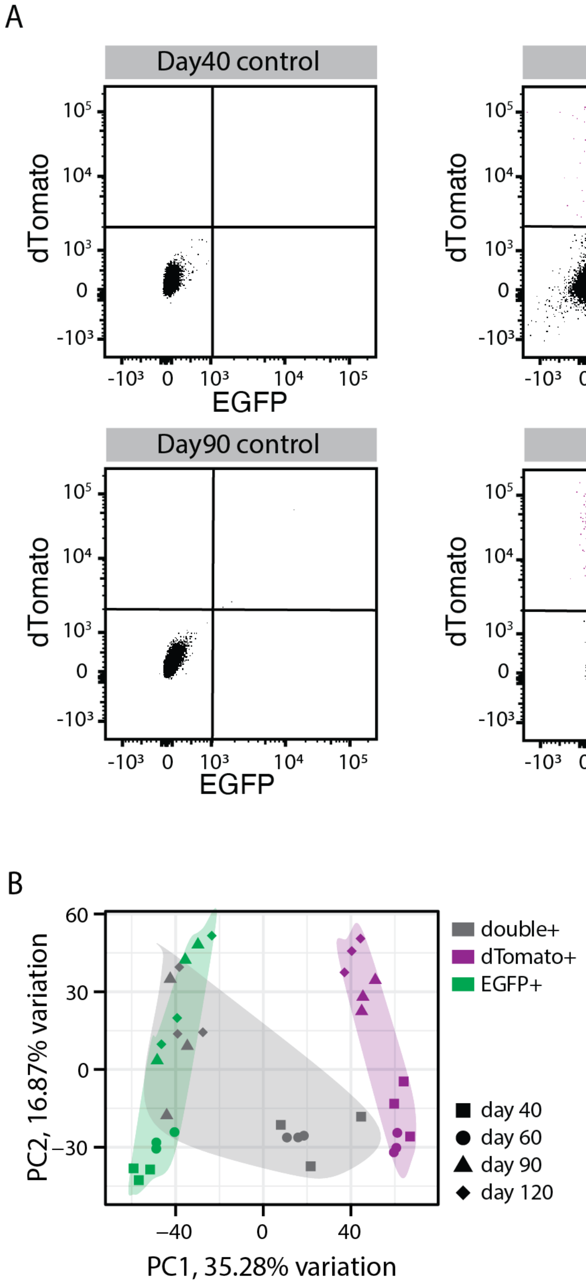


Figure 1 -figure supplement 2. Fluorescent activated cell sorting from the dual reporter organoids.
A) Gating criteria for FACS to sort different cell populations from dual reporter organoids at different organoid development stages. Negative (black), EGFP+ (green), dTomato+ (magenta), EGFP/dTomato double+ (grey) cells. B) Principal component analysis of transcriptome for sorted EGFP+ (green), dTomato+ (magenta), EGFP/dTomato double+ (grey) cells at 40, 60, 90 and 120 days of organoid development.

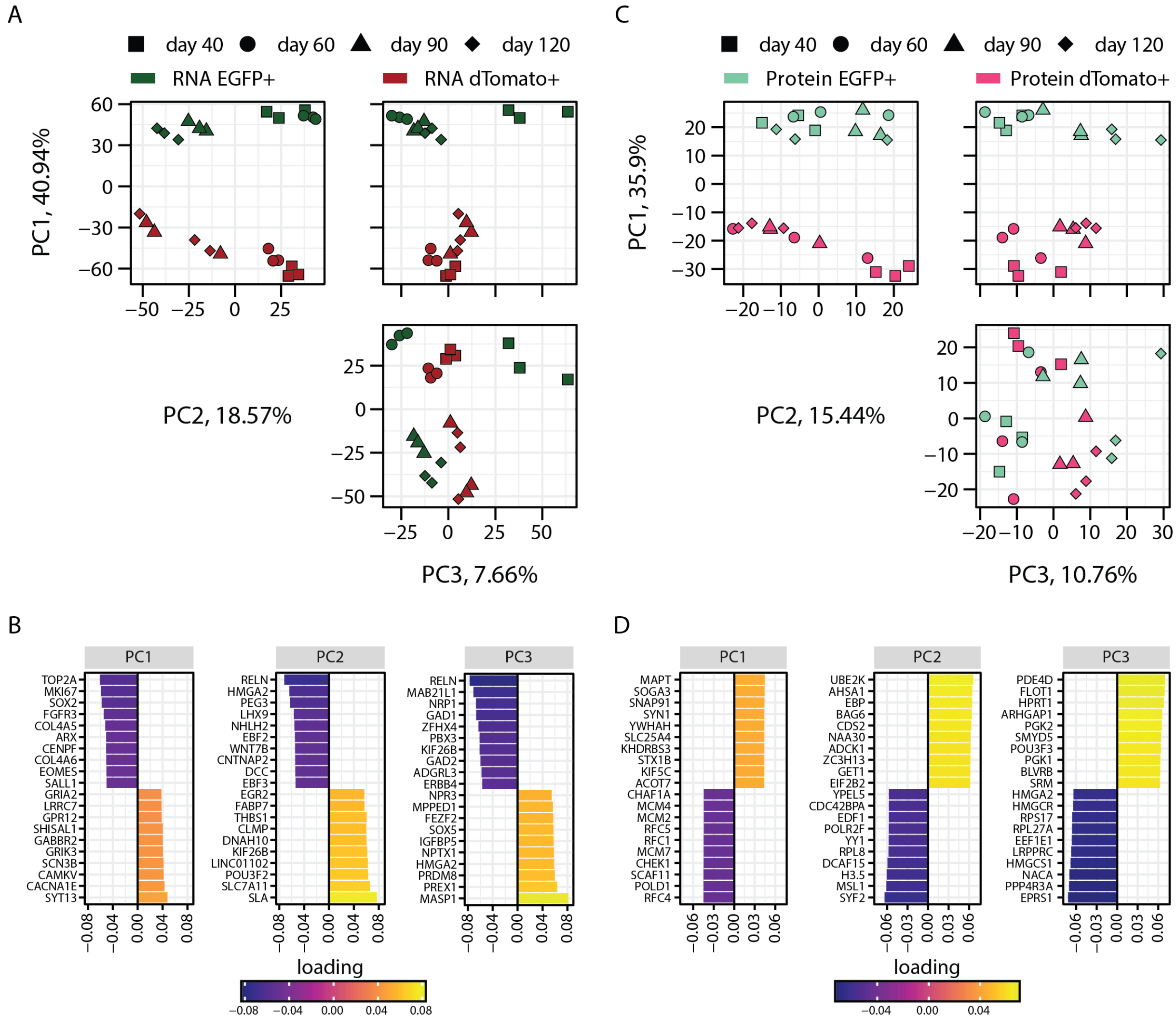


Figure 1 -figure supplement 3. PCA analysis of transcriptome and proteome of progenitors and neurons sorted from dual reporter organoids.

A) Principal component analysis of transcriptome of EGFP+ (dark green) and dTomato+ (red) cells at 40, 60, 90 and 120 days of organoid development (B) with loading scores for PC1, PC2 and PC3 showing top contributing genes. C) Principal component analysis of proteome of EGFP+ (light green) and dTomato+ (magenta) cells at 40, 60, 90 and 120 days of organoid development with (D) loading scores for PC1, PC2 and PC3 showing top contributing genes.

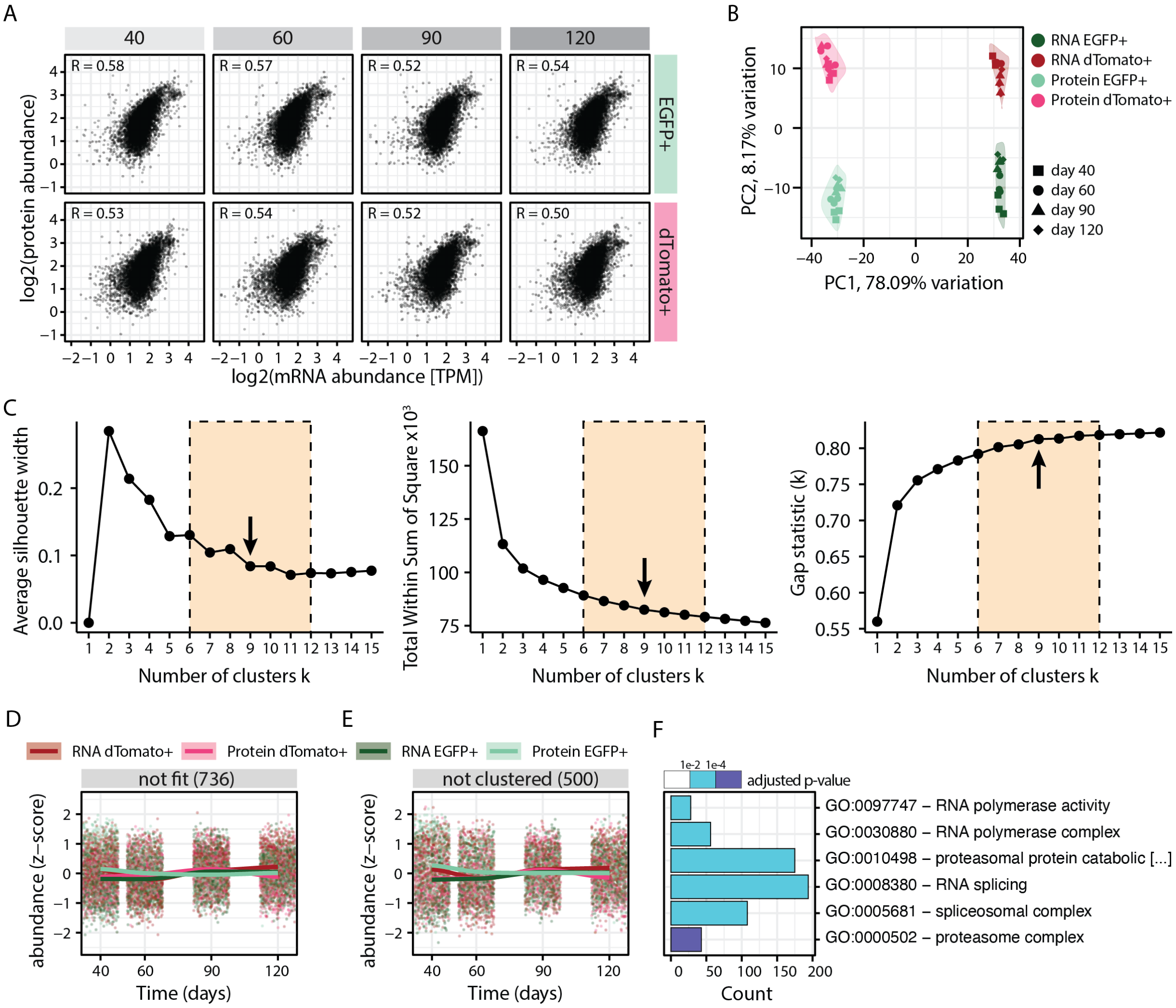


Figure 1 -figure supplement 4. Integration and clustering of the RNA-protein multi-omics data.

A) Plots showing correlation between protein and RNA abundance in progenitors (EGFP+, green) and neurons (dTomato+, red) at different stages of organoid development. B) Principal component analysis of transcriptome and proteome of progenitors (EGFP+, green) and neurons (dTomato+, red) at different stages of organoid development before z-score normalization. C) Optimal number of clusters were estimated for z-score normalized data using the 'average silhouette width', 'total within sum of square' and 'gap statistics' methods. Six to twelve clusters were manually evaluated with regards to their biologically meaningfulness and a subjective decision made to separate the data into nine modules. D, E) Z-score normalized RNA and protein abundance in progenitors (EGFP+, green) and neurons (dTomato+, red) at different stages of organoid development for genes for which the first regression fit failed (D) and 500 randomly sampled genes that were not clustered into any module (E). For each dataset, each dot displays the relative abundance of one gene and a trendline was fit through all data points. F) Enrichment analysis for gene ontology (GO) for genes with no regression fit and unclustered genes. Shown are six selected GO terms and the number of genes (x-axis) belonging to the respective GO term. Bar fill colors show binned adjusted p-values.

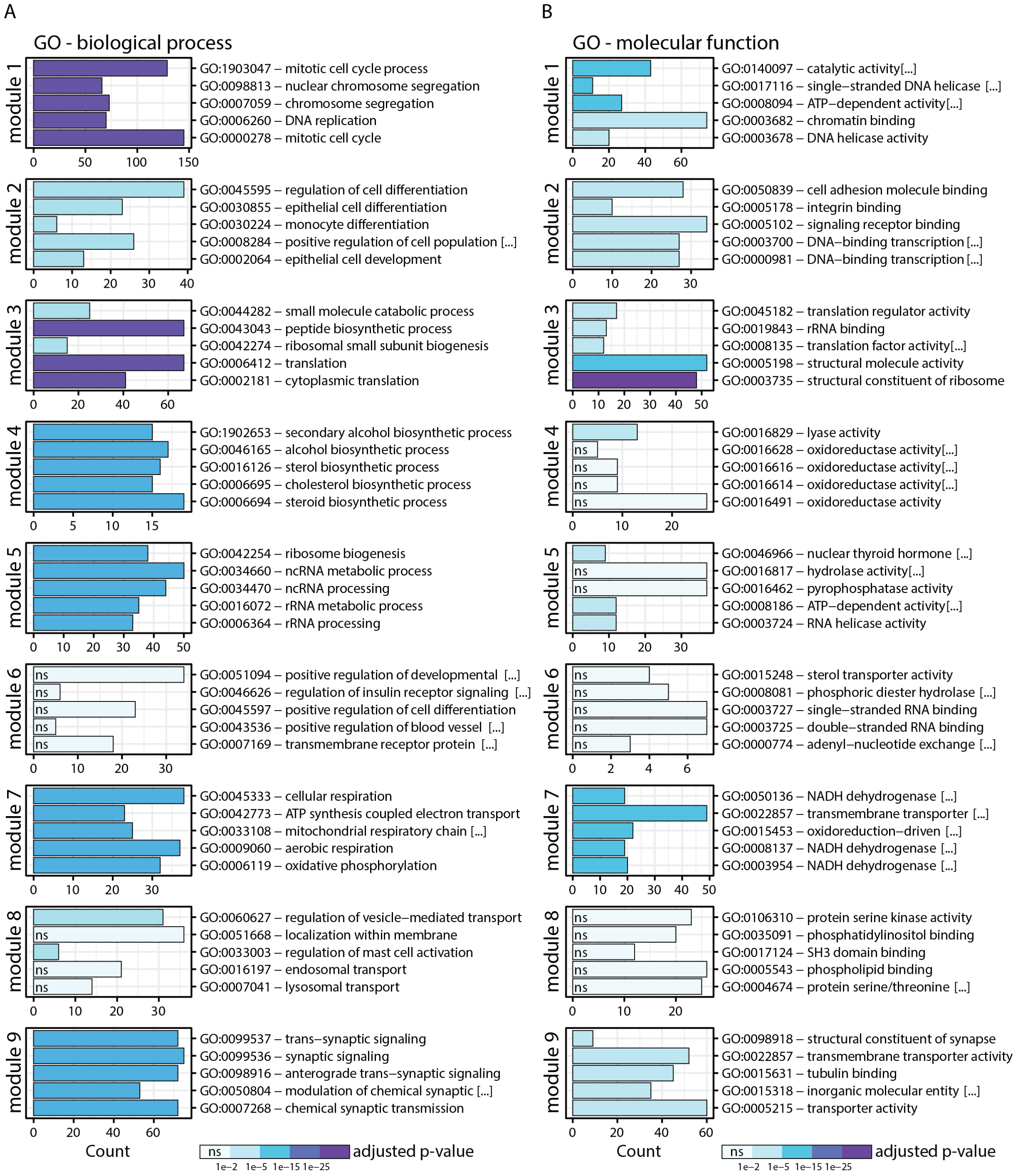


Figure 2 -figure supplement 1. GO analysis of gene expression modules.
A,B) Enrichment analysis of genes in the nine modules for (A) GO biological process (BP) and (B) GO molecular function (MF). Shown are the top five enriched terms per module and the number of genes (x-axis) belonging to the respective GO term. Bar fill colors show binned adjusted p-values.

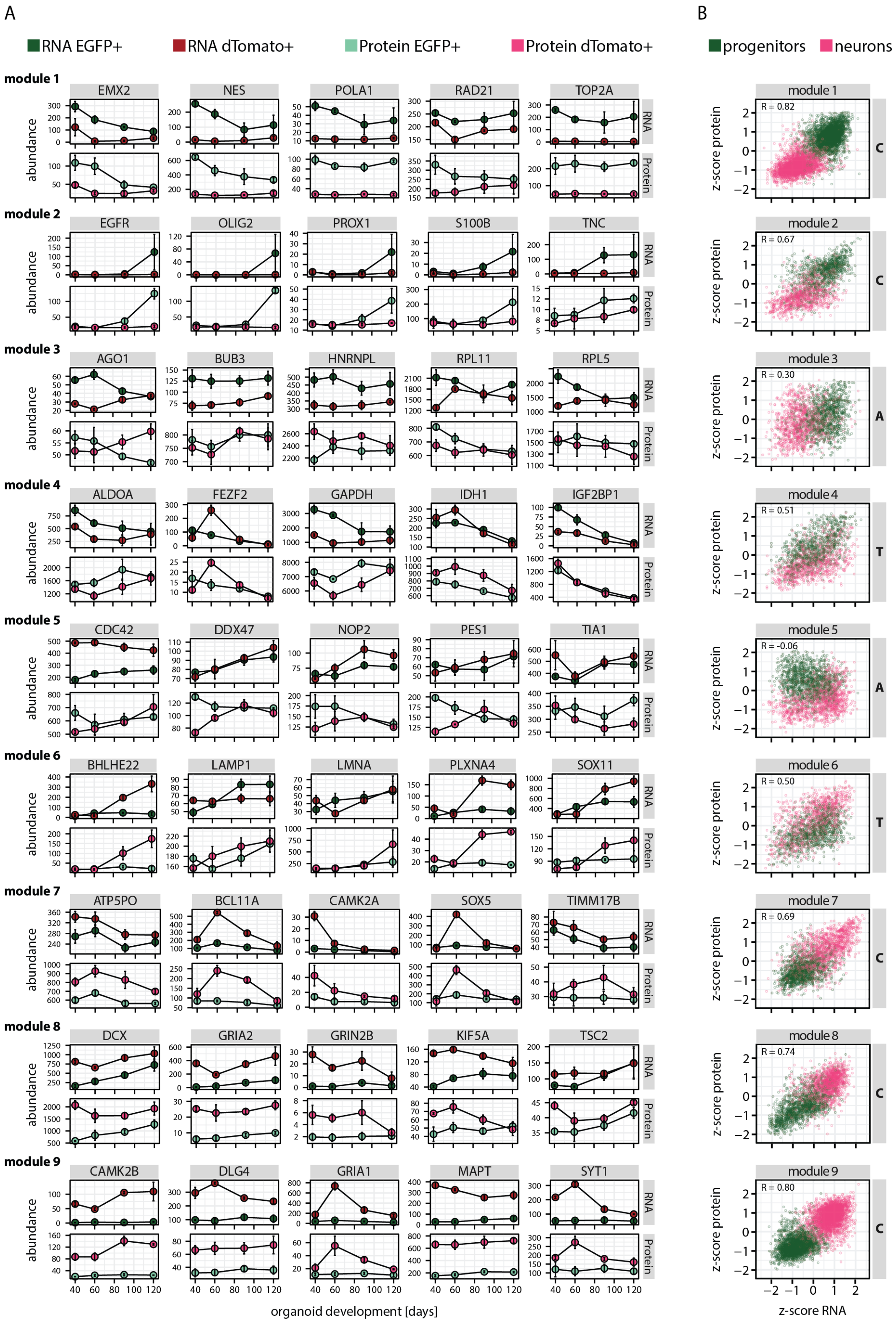
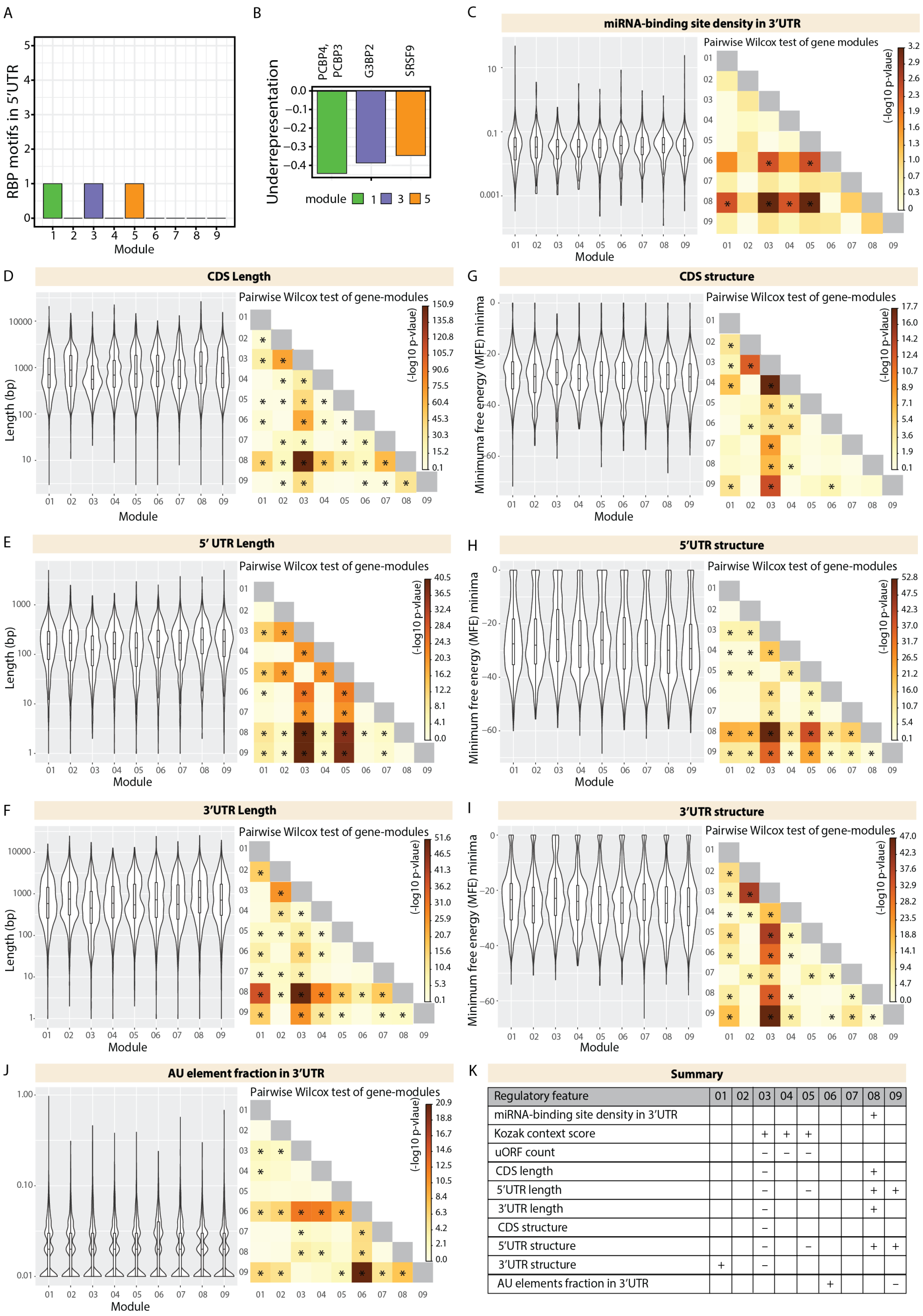


Figure 2 -figure supplement 2. Analysis of gene expression modules and correlation between relative RNA-protein abundance. A) RNA and protein expression in progenitors (EGFP+, green) and neurons (dTomato+, red) of five manually selected member genes per module. B) Plots showing module-wise correlation between relative abundance (z-score) of RNA and protein in progenitors (green) and neurons (red).



Regulatory feature	01	02	03	04	05	06	07	08	09
miRNA-binding site density in 3'UTR								+	
Kozak context score			+	+	+				
uORF count			–	–	–				
CDS length			–					+	
5'UTR length			–		–			+	+
3'UTR length			–					+	
CDS structure			–						
5'UTR structure			–		–			+	+
3'UTR structure	+		–						
AU elements fraction in 3'UTR						+			–

Figure 3 -figure supplement 1. Analysis of module-specific trans- and cis-regulatory features.

A) Barplot showing number of RBP motifs with significant over/under representation in 5'UTR of member genes of each gene module. Significance threshold: adjusted p-value <0.05. B) Barplot showing enrichment of RBP motifs that are significantly over/underrepresented in specific modules in Figure S7A. Black line marks zero position. Negative values indicate underrepresentation. RBPs for specific motifs are mentioned in the axis ticks. Bars are colored according to the module. C) Violin plots of module-wise distribution of miRNA-binding site density for transcripts of member genes. D,E,F) Violin plots of module-wise distribution of length of (D) coding sequences, (E) 5'UTRs and (E) 3'UTRs for transcripts of member genes. G,H,I) Violin plots of module-wise distribution of minimum free energy minima of (G) coding sequences, (H) 5'UTRs and (I) 3'UTRs for transcripts of member genes. J) Violin plots of module-wise distribution of AU element fraction in the 3'UTRs for transcripts of member genes. For all violin plots (C-J), inner boxplots mark median scores and interquartile range (IQR). Whiskers extend to 1.5x IQR. The heatmap on the right shows the -log10 p-value of the pairwise wilcox test of gene modules. Significant comparisons are marked by an asterisk ($p < 0.01$). K) Summary table highlighting the cis-regulatory transcript features for the gene expression modules. (+) indicates a significantly higher contribution. (-) indicates a significantly lower contribution.

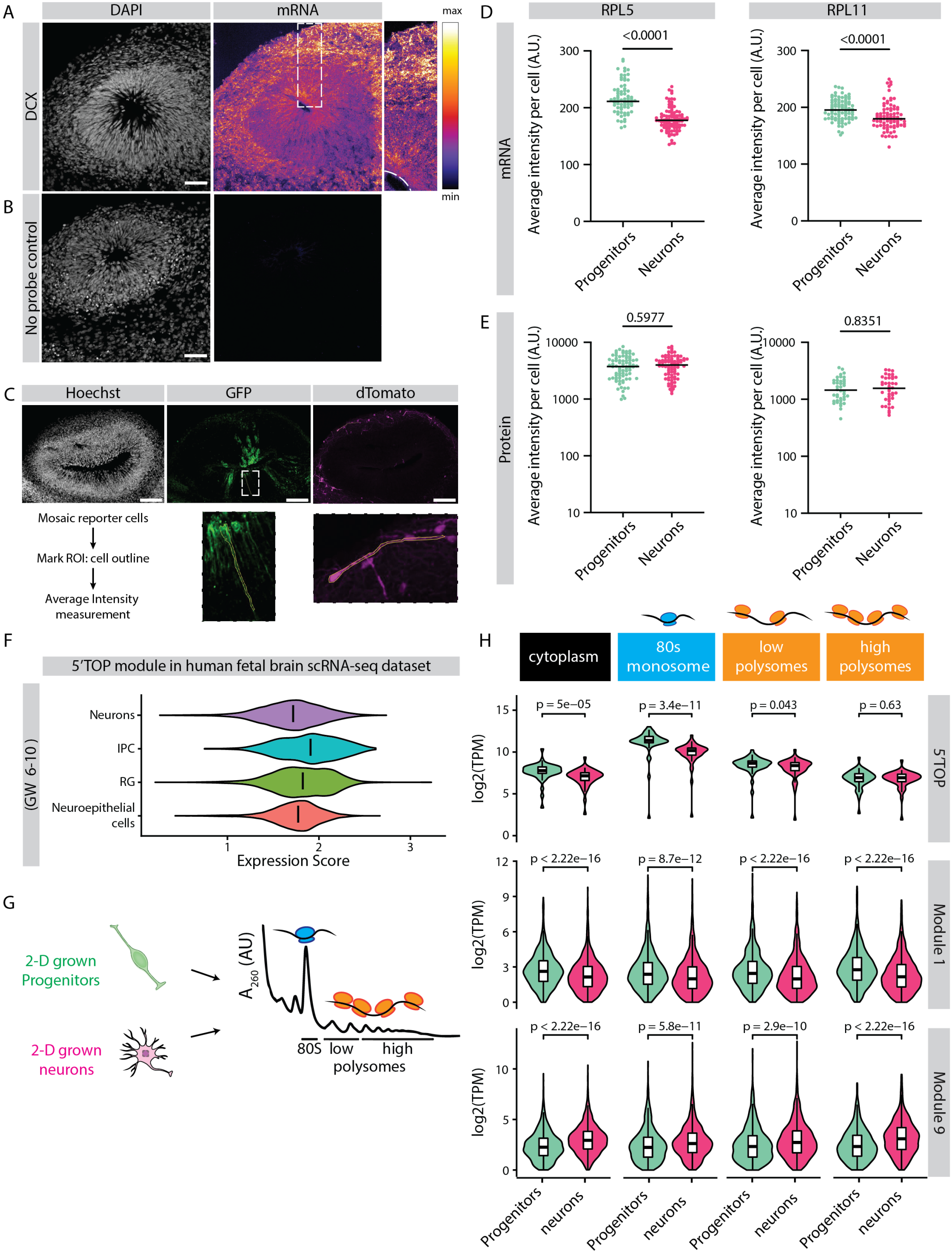


Figure 4 -figure supplement 1. Relative protein yield of 5'TOP mRNAs is lower in early progenitors compared to early born neurons. A,B) RNA-FISH for (A) DCX transcripts and (B) no probe control in 40 days old organoids. Image shows a typical ventricular zone (VZ)-like structure in the brain organoid with a zoomed-in image of the boxed area on the right. Dotted line marks the apical side of the VZ. Scale bar = 50 µm. C) Schematic representation of cell outlines of mosaic reporter cells marked as regions of interest (ROI) for intensity measurements. D, E) Average intensities of RPL5 and RPL11 transcripts (D) and protein (E) measured per cell. Each dot represents a cell. Black line marks the median. P-value of Mann-Whitney test is indicated. F) Expression score of 5'TOP genes in distinct cell clusters from scRNA-seq data of fetal tissues of gestation week 6-10 (Eze et al., 2021). G) Schematic representation and analysis of polysome profiling of 2D grown progenitors and neurons performed by (Blair et al., 2017). H) Plot represents RNA abundance of genes across different polysome fractions from progenitors and neurons. All 5'TOP genes and genes from modules 1 and 9 are plotted. Each dot represents a gene. Data from (Blair et al., 2017). P values of Wilcoxon test are shown.

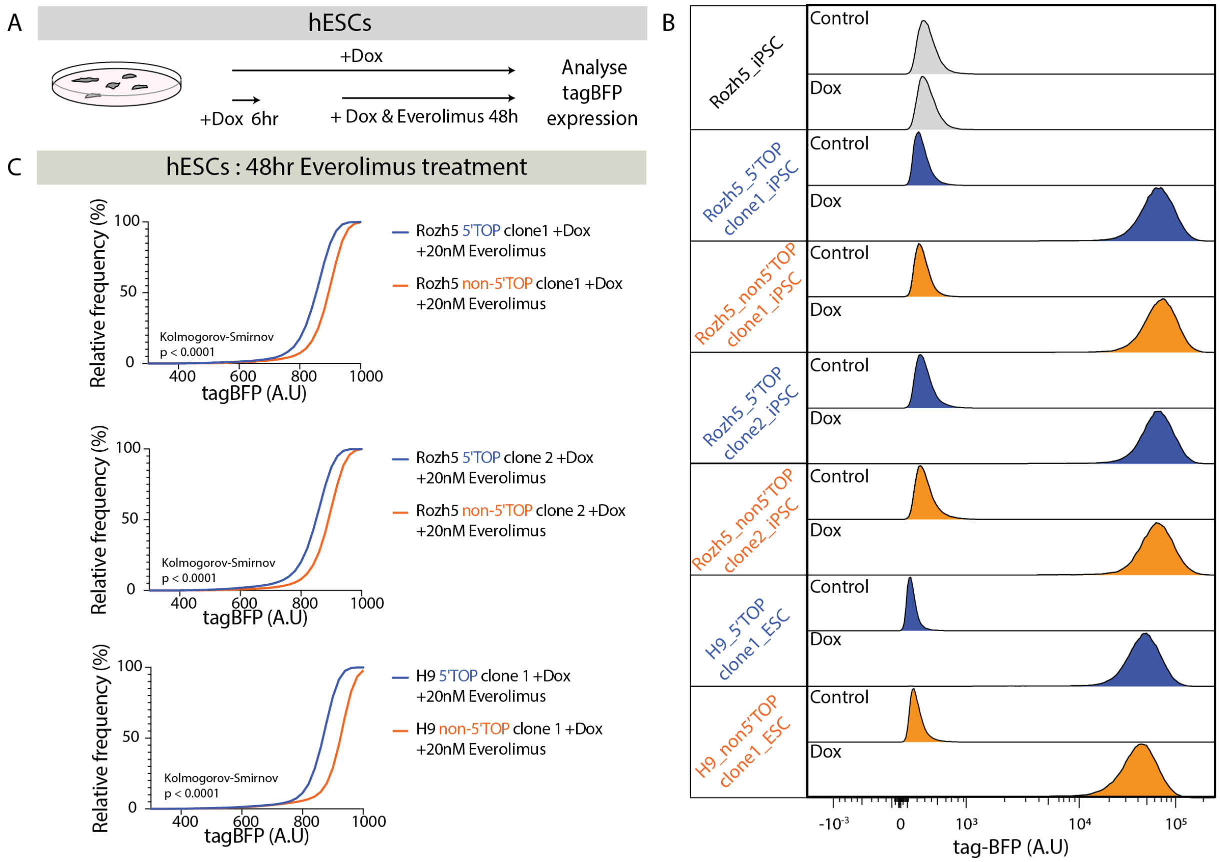


Figure 4 -figure supplement 2. Doxycycline-inducible 5'TOP-reporter hPSC lines

A) Schematic representation of the doxycycline-inducible reporter assay in hPSCs. B) Flow cytometry analysis of tagBFP expression in hPSCs carrying 5'TOP (blue) and non-5'TOP (orange) sequences in 5'UTR of the reporter construct. Parental cell line without reporter transgene (grey). Untreated control hPSCs and hPSCs treated with doxycycline (Dox) for 48 h were analyzed for multiple clones. C) Cumulative distribution of reporter tagBFP levels in doxycycline-induced 5'TOP (blue) and non-5'TOP (orange) reporter hPSC pairs treated with 20 nM everolimus.

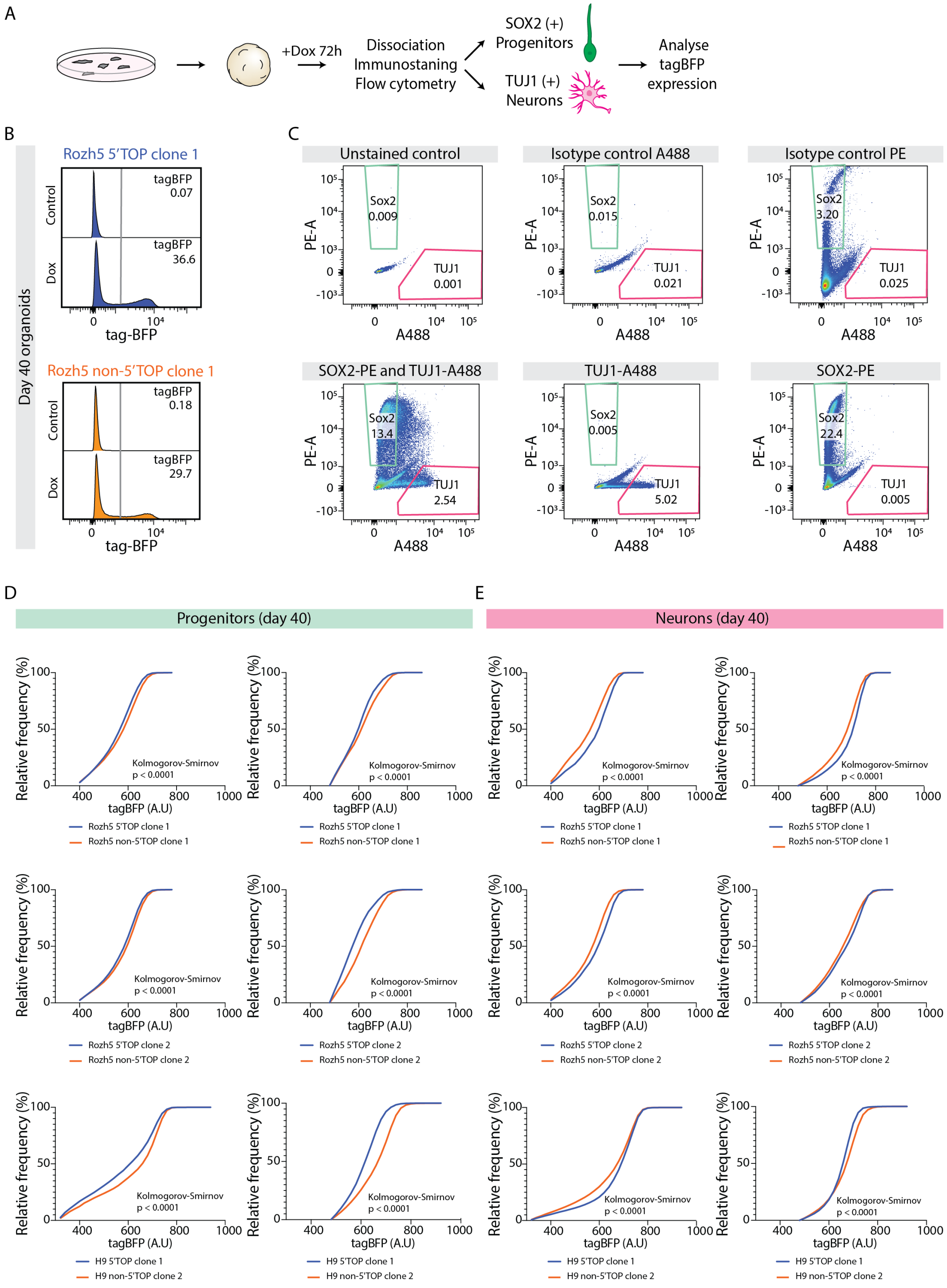


Figure 4 -figure supplement 3. 5'TOP reporter assay in 40 day old brain organoids.

A) Schematic representation of the dox-inducible reporter assay. hPSCs carrying 5'TOP (blue) and non-5'TOP (orange) sequences in 5'UTR of tagBFP were used to grow brain organoids. Doxycycline treated 40 days old organoids were dissociated and cells stained for SOX2 and TUJ1 to measure tag-BFP levels in progenitors and neurons, respectively. B) Representative image of flow cytometry analysis of tagBFP expression in untreated control organoids and 40 days old organoids after 72 h doxycycline treatment. Cells carrying 5'TOP (blue) and non-5'TOP (orange) sequences in 5'UTR of tagBFP are depicted. Percentage of single cells expressing tagBFP according to the gating parameters is indicated. C) Flow cytometry analysis of cells from 40 days old organoids immunostained with anti-SOX2-PE and anti-TUJ1-A488 antibodies. Panel includes isotype control antibody staining controls used to evaluate the staining and to define gating parameters. Cells stained with SOX2 (green box), TUJ1 (red box) were defined as shown in the figure. D,E) Cumulative distribution of reporter tagBFP levels in (D) progenitors (SOX2) and (E) neurons (TUJ1) dissociated from doxycycline-treated 40 days old organoids. Each graph represents a separate experiment with a 5'TOP (blue) and non-5'TOP (orange) reporter pair. Statistical significance was calculated using the non-parametric Kolmogorov-Smirnov test.

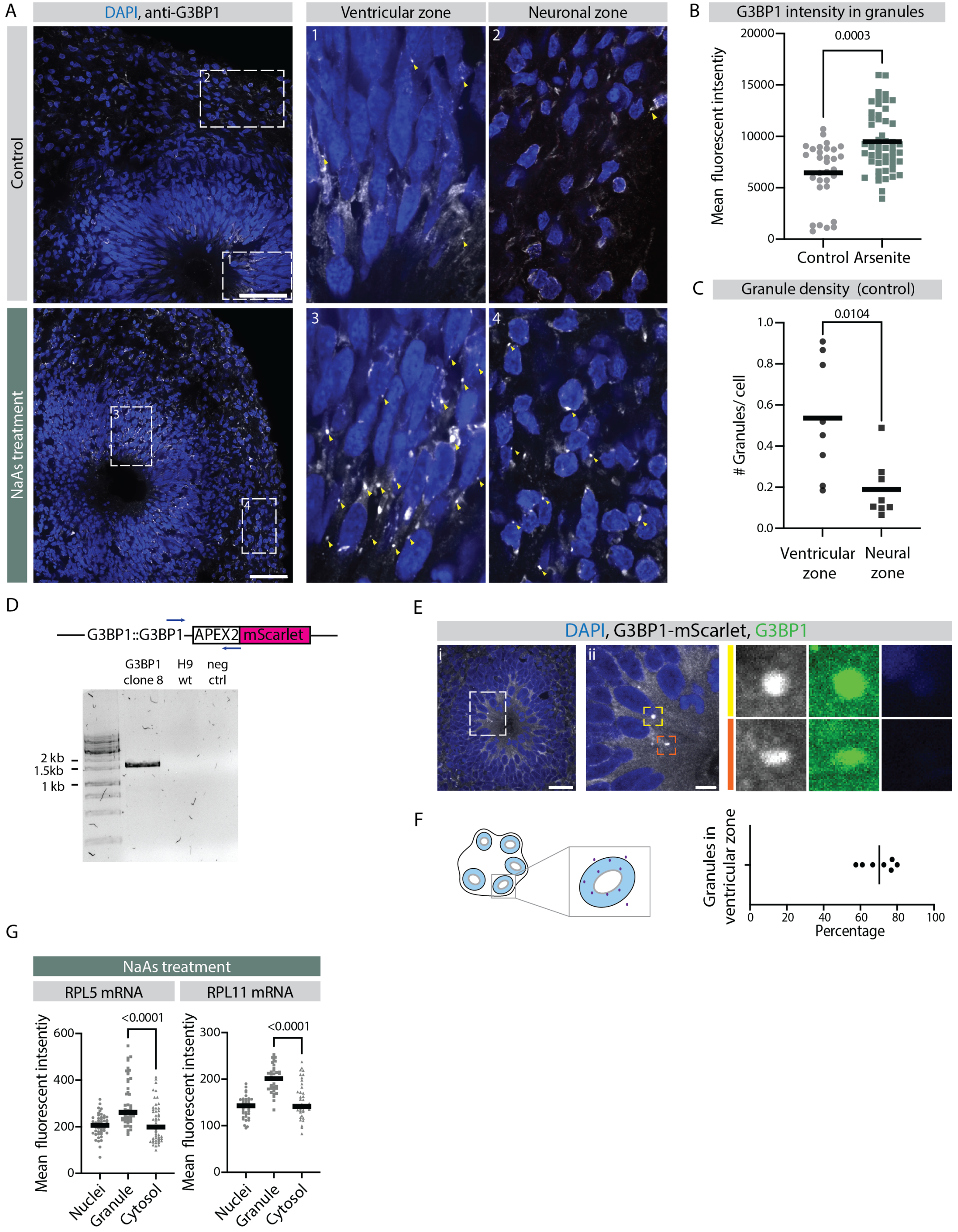


Figure 4 -figure supplement 4. Early ventricular radial glia exhibit stress granule-like structures.

A) Confocal scan of ventricular zone (VZ) in a control and sodium arsenite-treated organoid stained with anti-G3BP1 antibody and DAPI. Zoomed-in image of the inlay shows G3BP1 granules in VZ and neuronal zones. Scale bar = 50 µm. B) Quantification of G3BP1 signal intensity in the granules observed in the control and sodium arsenite-treated organoids. P-value statistic of t test. Each dot represents a granule. (Number of organoids analyzed, control: n=3; NaAs: n=4; Number of granules analyzed: control: n=31; NaAs: n=51) C) Quantification of average number of G3BP1 granules per cell in VZ and neuronal zone of 40 days old control organoids. P-value statistic of t test. Each dot represents an area from an organoid. Quantification performed in 3 organoids with total field of views: n=8 and number of cells counted: NPCs: n=829; Neurons: n=804. D) Results of genotyping PCR to check insertion of the G3BP1-mScarlet transgene in the G3BP1 locus. Samples include G3BP1 reporter hESCs, control H9 hEScs and negative control. E) Confocal scan of VZ in a 40 day old G3BP1-mScarlet organoid stained with anti-G3BP1 antibody and DAPI. Zoomed-in images of the inlay show granules in VZ marked by mScarlet (grey) and G3BP1 (green). Scale bars = 20 µm (i), 5 µm (ii). F) Quantification of percentage of total G3BP1 granules observed in a field that are located in the VZ region. Schematic on the left shows G3BP1 granules in the tissue and quantification strategy. G) Quantification of RPL5 and RPL11 RNA-FISH signal intensities in the nucleus, G3BP1-positive granules and the cytosol of progenitors in the VZ of sodium arsenite treated organoids. The p-values of Mann-Whitney tests are shown. Each dot represents a cell (Number of granules analyzed: RPL5: SA: n-51; RPL11 SA: n=111).

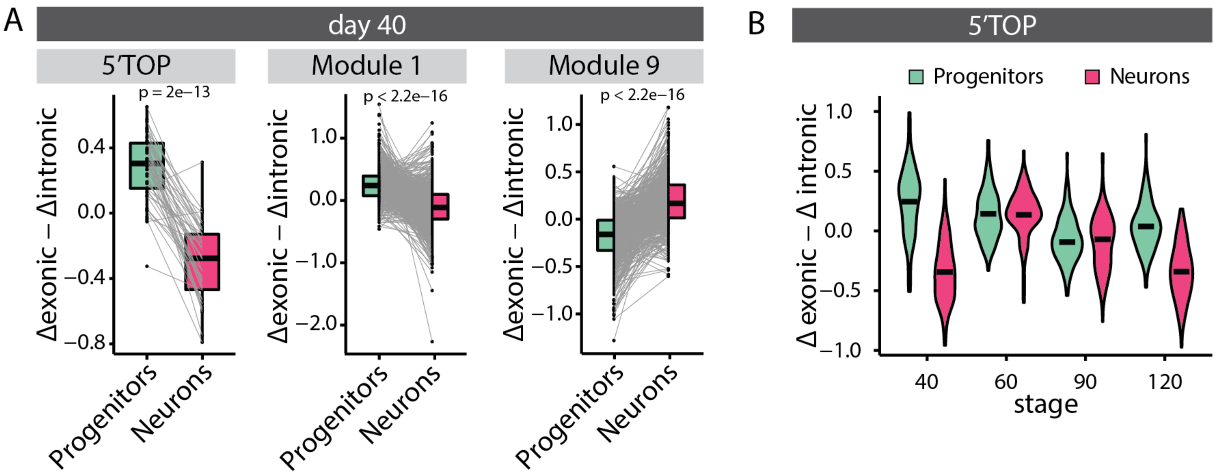
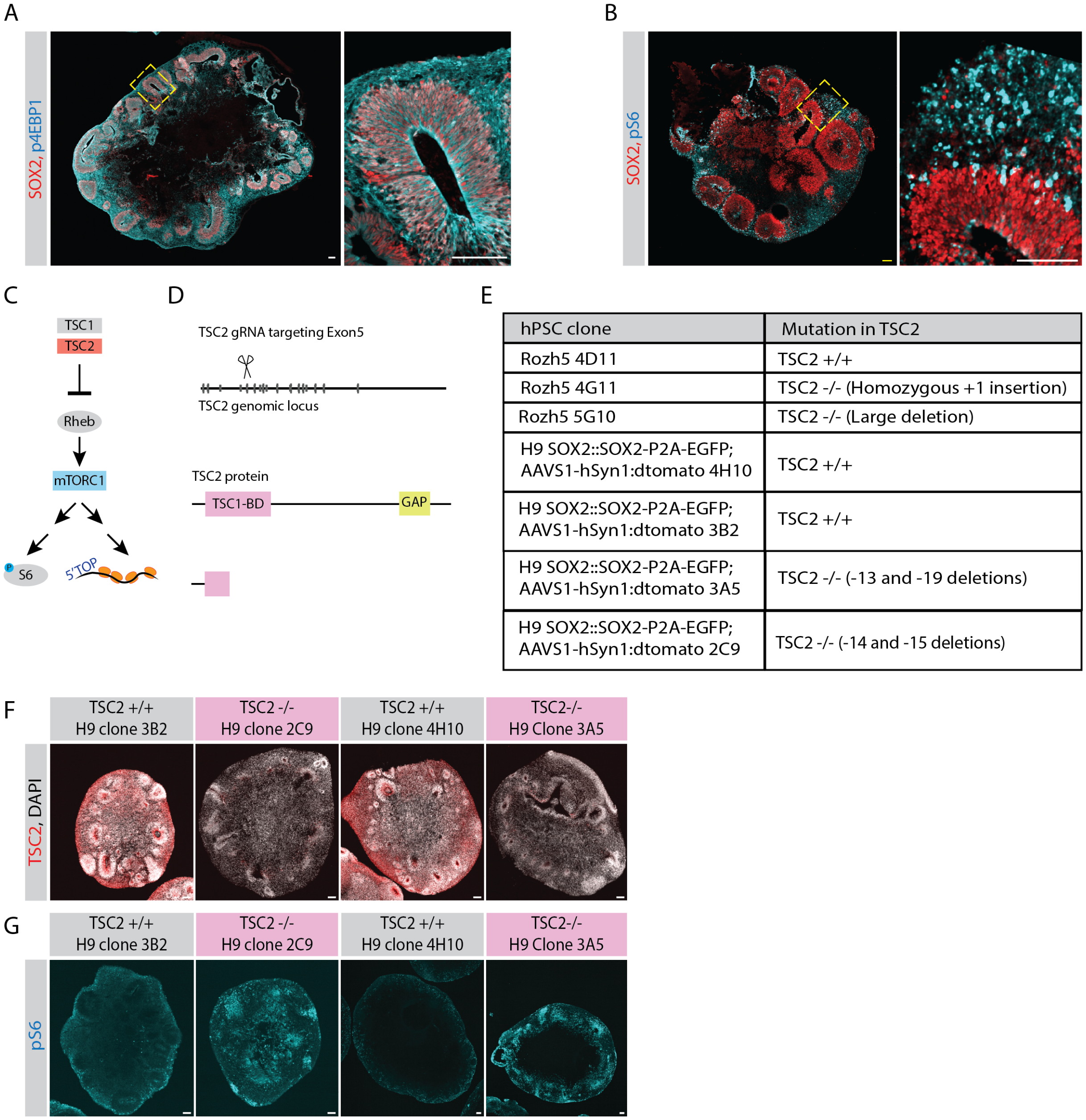


Figure 4 -figure supplement 5. Analysis of relative mRNA stability.

A) Paired relative stability of mRNAs in progenitors and neurons of 40 days old organoids measured by estimating exonic versus intronic reads from the RNA-seq data. Black dash marks the mean relative stability. Results are shown for all classical 5'TOP genes and genes in modules 1 and 9. The p-values of the Wilcoxon test are shown. B) Relative stability of mRNAs of classical 5'TOP genes in progenitors and neurons at different stages of organoid development. Black dash marks the mean relative stability.



hPSC clone	Mutation in TSC2
Rozh5 4D11	TSC2 +/+
Rozh5 4G11	TSC2 -/- (Homozygous +1 insertion)
Rozh5 5G10	TSC2 -/- (Large deletion)
H9 SOX2::SOX2-P2A-EGFP; AAVS1-hSyn1:dtomato 4H10	TSC2 +/+
H9 SOX2::SOX2-P2A-EGFP; AAVS1-hSyn1:dtomato 3B2	TSC2 +/+
H9 SOX2::SOX2-P2A-EGFP; AAVS1-hSyn1:dtomato 3A5	TSC2 -/- (-13 and -19 deletions)
H9 SOX2::SOX2-P2A-EGFP; AAVS1-hSyn1:dtomato 2C9	TSC2 -/- (-14 and -15 deletions)

Figure 5 -figure supplement 1. Analysis and perturbation of mTOR activity.
AA) Confocal scan of a 40 day old organoid stained with anti-phospho-4EBP1 (cyan) and anti-SOX2 (red) antibodies. Zoomed-in image of the inlay shows a ventricular zone (VZ). Scale bar = 100 µm. B) Confocal scan of a 40 day old organoid stained with anti-phospho-S6 (cyan) and anti-SOX2 (red) antibodies. Zoomed-in image of the inlay shows a VZ. Scale bar = 100 µm. C) Schematic representation of the mTOR pathway. D) Schematic representation of genome editing of TSC2 genomic locus by CRISPR-Cas9 approach. A guide targeting Exon5 was used to generate TSC2-/- hPSCs. E) Table summarizing TSC2 control and knockout hPSC clones used in this study. F) Confocal scan of 40 days old TSC2+/+ and TSC2-/- organoid sections stained with anti-TSC2 antibody and DAPI. Scale bar = 100 µm. G) Confocal scan of 40 days old TSC2+/+ and TSC2-/- organoid sections stained with anti-pS6 and antibody. Scale bar =100 µm.

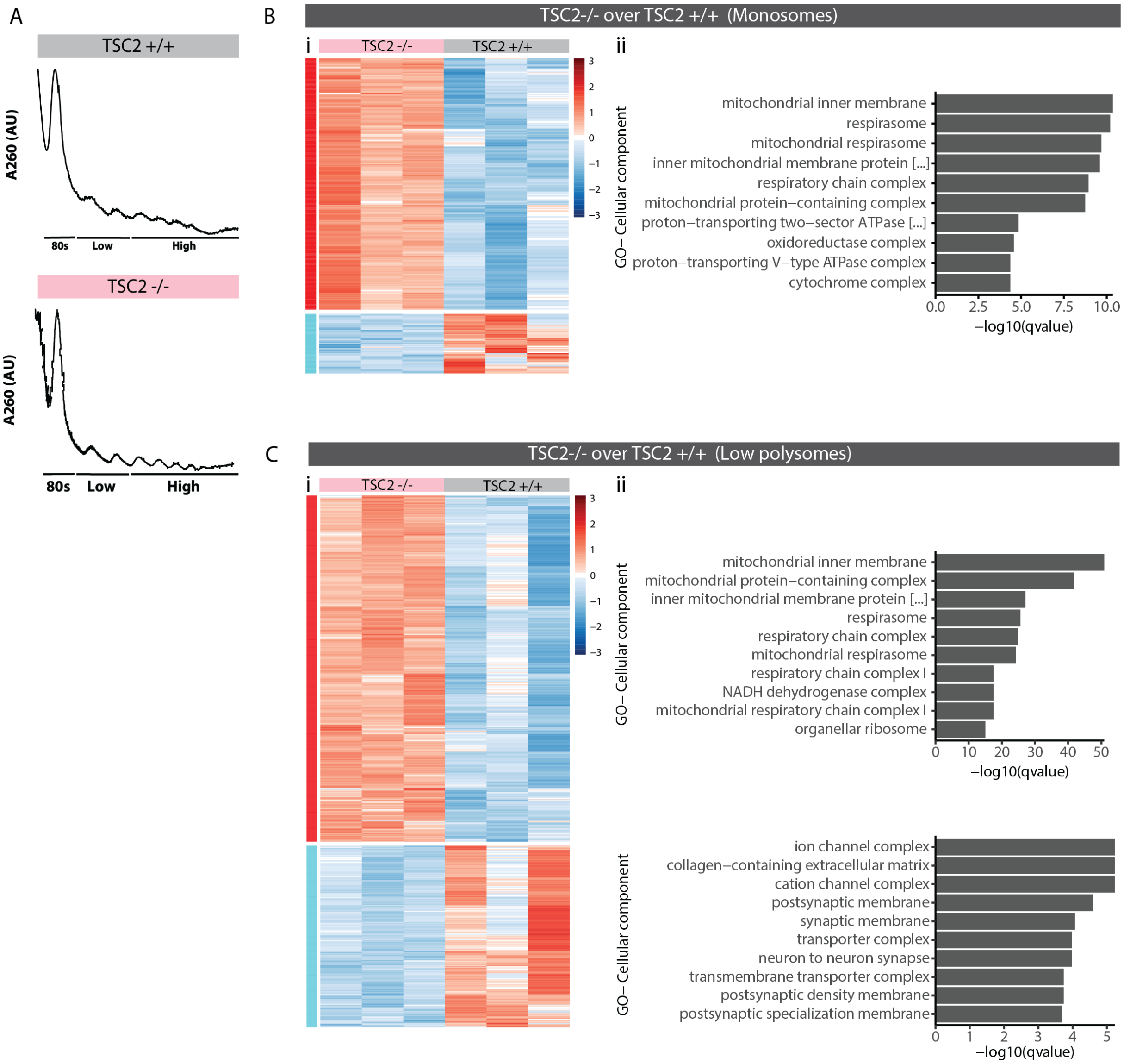


Figure 5 -figure supplement 2. Polysome profiling in mTOR overactivated organoids.

A) Representative traces of polysome profiling of TSC2+/+ and TSC2-/- organoids. B) Hierarchical clustered heat map of differentially associated genes (DAGs) with monosome fractions of TSC2+/+ and TSC2-/- organoids (i). Top 10 most significant cellular component GO terms for TSC2-/- enriched DAGs (ii). There was no significant GO term observed for TSC2+/+ enriched DAGs. C) Hierarchical clustered heat map of genes differentially associated with low-polysome fractions of TSC2+/+ and TSC2-/- organoids (i). Top 10 most significant cellular component GO terms for TSC2-/- enriched DAGs and TSC2+/+ enriched DAGs (ii).

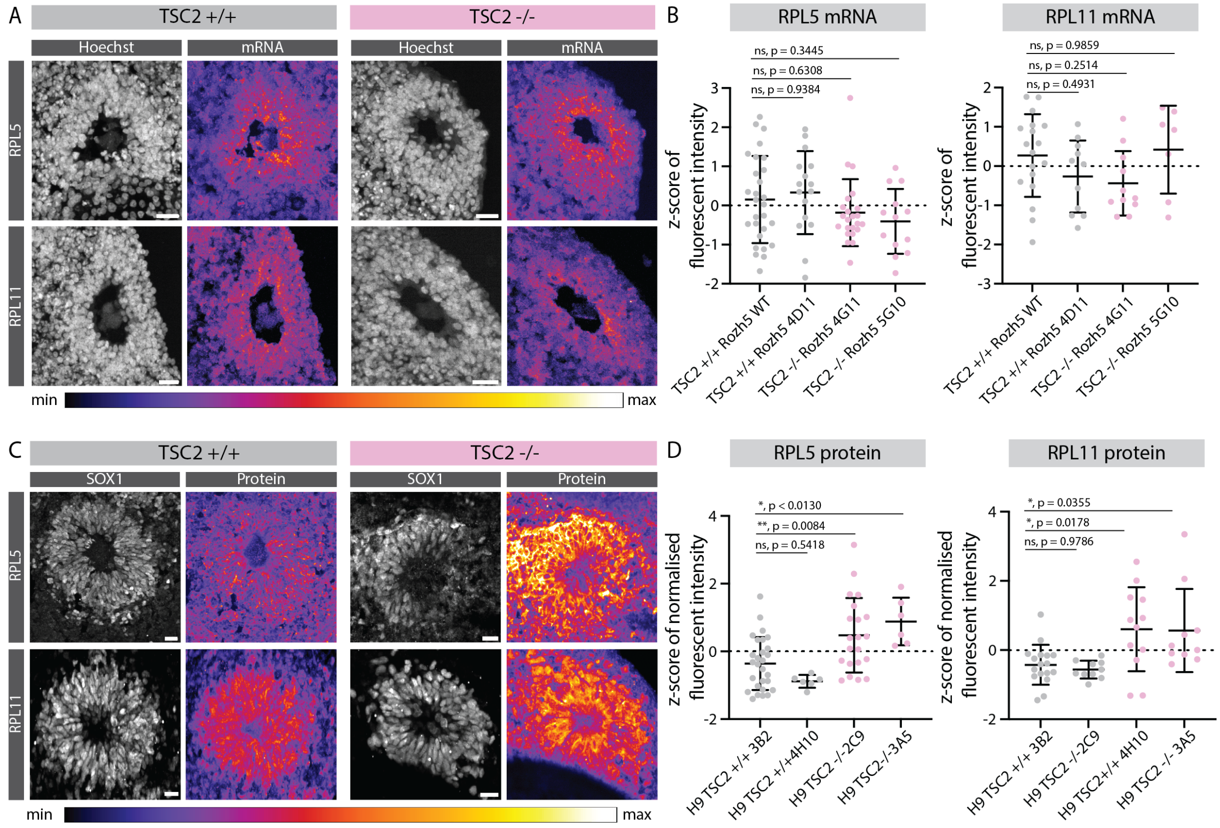


Figure 5 -figure supplement 3. Effect of mTOR overactivation on 5'TOP RNA and protein levels.
A) RNA-FISH for the 5'TOP transcripts RPL5 and RPL11, in 40 days old TSC2+/+ and TSC2-/- organoids. Images show a typical ventricular zone (VZ)-like structure in the brain organoid. Scale bar = 20 µm. B) Quantification of z-score normalized fluorescent intensity of RPL5 and RPL11 FISH signal in the VZs of TSC2+/+ and TSC2-/- organoids. Error bars mark standard deviation. P-values of unpaired t test are shown. Each dot represents a VZ. Data from 3 batches of organoid differentiation for each clone. C) Immunostaining for the 5'TOP proteins RPL5 and RPL11, in 40 days old TSC2+/+ and TSC2-/- organoids. Images show a typical VZ-like structure in the brain organoid. Scale bar = 20 µm. D) Quantification of z-score normalized fluorescent intensity of RPL5 and RPL11 proteins in the VZs of TSC2+/+ and TSC2-/- organoids. Error bars mark standard deviation. P-values of unpaired t test are shown. Each dot represents a VZ. Data from 3 batches of organoid differentiation for each clone.

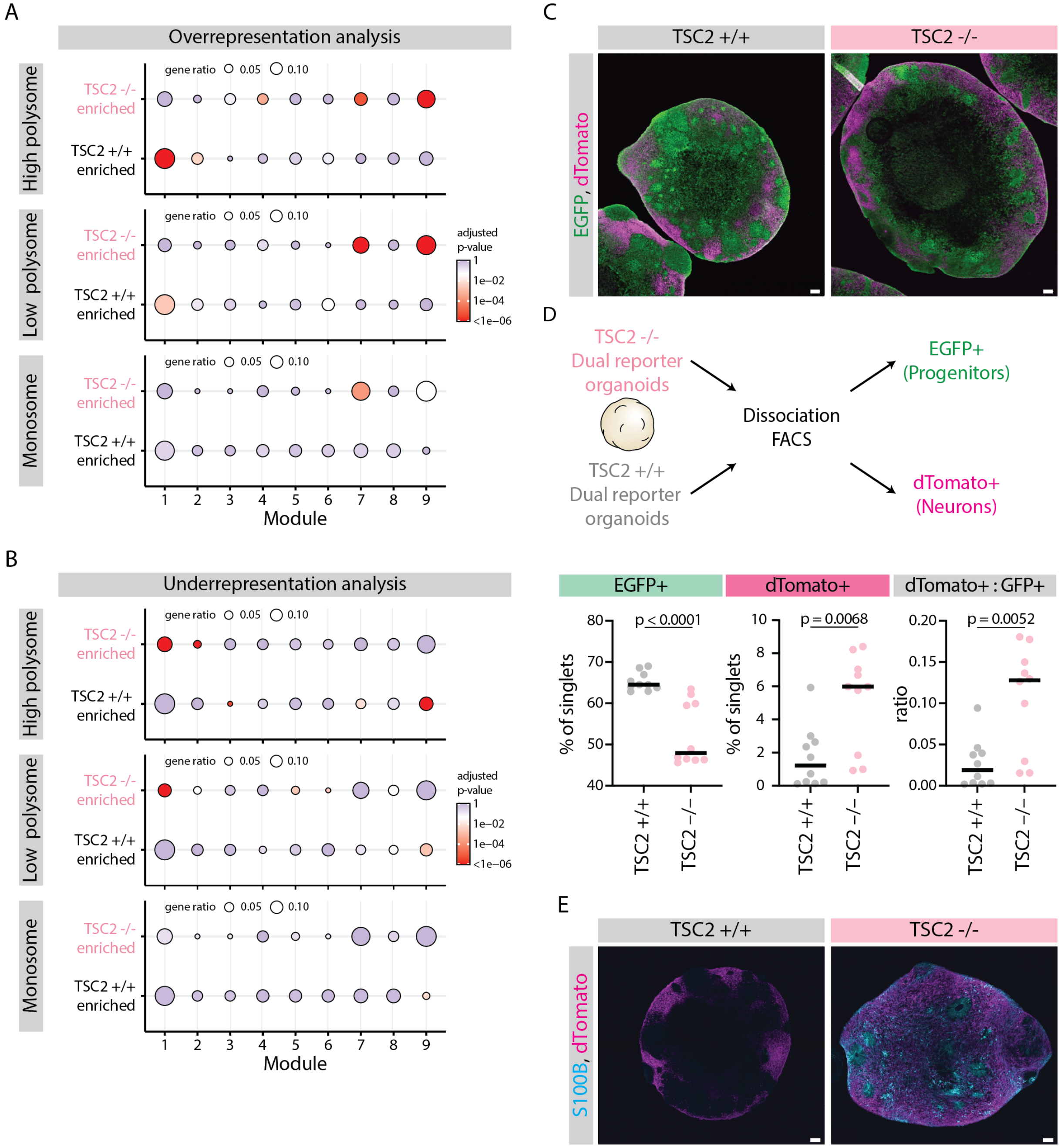


Figure 5 -figure supplement 4. Effect of mTOR overactivation on tissue development.

A) Dot plot showing overrepresentation analysis of genes differentially associated with ribosome fractions of TSC2-/- and TSC2+/+ organoids across the 9 gene modules. Significance threshold: Adjusted p-value <0.05. Gene lists refer to DAGs from clusters from Figure 5D. B) Dot plot showing underrepresentation analysis of genes differentially associated with ribosome fractions of TSC2-/- and TSC2+/+ organoids across the 9 gene modules. Significance threshold: Adjusted p-value <0.05. Gene lists refer to DAGs from clusters from Figure 5D. C) Confocal scan of cryosections of 40 days old TSC2+/+ and TSC2-/- dual reporter organoids showing EGFP and dTomato distribution. Scale bar = 100 µm. D) Schematic representation of the flow cytometry analysis of dual reporter TSC2+/+ and TSC2-/- organoids. Plots show the percentage of single cells that are EGFP+ (green), dTomato+ (magenta) and the ratio of dTomato:EGFP cells (grey). Horizontal line marks the median values. P-values of Mann Whitney tests are shown. Samples from 2 batches of organoid differentiation with 2 clones per condition. E) Confocal scan of cryosections of 40 days old TSC2+/+ and TSC2-/- dual reporter organoids immunostained for dTomato and S100B. Scale bar = 100 µm.