Transcriptional buffering and 3′UTR lengthening are shaped during human neurodevelopment by shifts in mRNA stability and microRNA load

Marat Mufteev*[1,2], Deivid C. Rodrigues*[1], Kyoko E. Yuki[3], Ashrut Narula[2,4], Wei Wei[1], Alina Piekna[1], Jiajie Liu[1], Peter Pasceri[1], Olivia S. Rissland[4,5], Michael D. Wilson[2,3], and James Ellis[1,2].

* *equal contribution*

Affiliations:

[1] Developmental & Stem Cell Biology, Hospital for Sick Children, Toronto, Ontario M5G 0A4, Canada

[2] Department of Molecular Genetics, University of Toronto, Toronto, Ontario M5S 1A8, Canada

[3] Genetics & Genome Biology, Hospital for Sick Children, Toronto, Ontario M5G 0A4, Canada

[4] Molecular Medicine, Hospital for Sick Children, Toronto, Ontario M5G 0A4, Canada

[5] RNA Bioscience Initiative and Department of Biochemistry & Molecular Genetics, University of Colorado School of Medicine, Aurora, Colorado 80045, USA

Corresponding author: jellis@sickkids.ca

Key words: miRNA load, transcription rate, mRNA stability, RATE-seq, small RNA-seq, transcriptional buffering, post-transcriptional regulation, 3′UTR.

Abstract

The contribution of mRNA half-life is commonly overlooked when examining changes in mRNA abundance during development. mRNA levels of some genes are regulated by transcription rate only, but others may be regulated by mRNA half-life only shifts. Furthermore, transcriptional buffering is predicted when changes in transcription rates have compensating shifts in mRNA half-life resulting in no change to steady-state levels. Likewise, transcriptional boosting should result when changes in transcription rate are accompanied by amplifying half-life shifts. During neurodevelopment there is widespread 3′UTR lengthening that could be shaped by differential shifts in the stability of existing short or long 3′UTR transcript isoforms. We measured transcription rate and mRNA half-life changes during induced human Pluripotent Stem Cell (iPSC)-derived neuronal development using RATE-seq. During transitions to progenitor and neuron stages, transcriptional buffering occurred in up to 50%, and transcriptional boosting in up to 15%, of genes with changed transcription rates. The remaining changes occurred by transcription rate only or mRNA half-life only shifts. Average mRNA half-life decreased two-fold in neurons relative to iPSCs. Short gene isoforms were more destabilized in neurons and thereby increased the average 3′UTR length. Small RNA sequencing captured an increase in microRNA copy number per cell during neurodevelopment. We propose that mRNA destabilization and 3′UTR lengthening are driven in part by an increase in microRNA load in neurons. Our findings identify mRNA stability mechanisms in human neurodevelopment that regulate gene and isoform level abundance and provide a precedent for similar post-transcriptional regulatory events as other tissues develop.

Introduction

Regulation of mRNA stability is essential for pluripotent stem cell homeostasis[1,2] and neuronal development[3]. mRNA stability also regulates many genes that impact cortical neuron activity, including autism spectrum disorder-risk genes[4] and Fragile X Messenger Ribonucleoprotein 1 (FMRP) targets[5]. During development, active modulation of mRNA stability can reinforce transcriptional programs by boosting transcript accumulation of weakly transcribed genes[6]. More broadly, coordination between mRNA transcription and stability has been shown to buffer transcript abundance from yeast to human cells in response to environmental or genetic manipulations [7,8].

To understand the dynamics of transcriptome regulation as healthy cells differentiate through developmental stages, it is crucial to characterize transcription rates and mRNA stability on a genome-wide scale. To search for evidence of changes in mRNA half-life that can buffer or boost changes in transcription rates during human neurodevelopmental stages, pluripotent stem cells can be differentiated *in vitro* into Neural Progenitor Cells (NPC) and cortical neurons[9]. We previously used iPSC-derived cortical neurons from an individual with Rett syndrome compared to their isogenic control to demonstrate pronounced mRNA half-life only shifts and transcriptional buffering of hundreds of genes in this neurodevelopmental disorder [10].

Most eukaryotic genes contain multiple 3′-end cleavage and polyadenylation sites (PAS). It is well established in several organisms that neurogenesis is accompanied by the utilization of longer 3′UTR isoforms[11–13]. Although unusually long 3′UTR isoforms are found in neurons[14], a recent single-cell RNA-seq analysis during mouse embryogenesis revealed a gradual increase in 3′UTR length from E9.5 to E13.5 in most tissues[15]. A functional role of 3′UTR length regulation during development is supported by evidence that knockdown of the mRNA cleavage machinery component *NUDT21* increases the reprogramming efficiency of human fibroblasts into induced Pluripotent Stem Cells (iPSCs)[16]. The observed increase in 3′UTR isoform lengths is currently attributed to a switch from a proximal to a more distal cleavage site in the 3′UTR by the process of alternative polyadenylation (APA)[17].

Differences in the stability of 3′UTR isoforms can also control their relative abundance, as is the case with the critical neuronal transcription modulator *MECP2*. In human embryonic stem cells (ESCs), the RNA binding protein (RBP) PUM1 and pluripotent-specific miRNAs actively degrade *MECP2* mRNAs containing the long 3′UTR. This degradation halts in neurons leading to an increase in abundance of the long 3′UTR isoform of *MECP2*[18]. This observation suggests that 3′UTR isoform length can be determined by shifts in mRNA stability.

This novel model complements and extends the APA-based model by proposing that 3′UTR length is altered by the differential degradation dynamics of mRNA isoforms in the cytoplasm during neurodevelopment. Potential mechanisms could act in *cis* on long isoforms that have more RNA motifs available as targets than short isoforms, or in *trans* by altering the cell-specific levels of motif-binding molecules such as microRNAs (miRNA). To examine the potential for these *trans*-effects during neurodevelopment, changes in miRNA load can be defined as the number of available miRNAs per cell relative to target mRNA molecules. The miRNA load could be a driving feature that controls both mRNA stability and isoform length in different cell types.

The overall copy number of miRNAs affects the silencing machinery and has functional consequences during neurodevelopment. First, genetic inhibition of the miRNA silencing machinery responsible for mRNA stability regulation causes altered self-renewal, differentiation, and apoptosis in human and mouse ESCs[19–23]. Correspondingly, reintroducing pluripotency-specific mature miR-302 or miR-372 is sufficient to rescue *DICER* knock-out deficits in human ESCs[23]. In parallel to pluripotent stem cells, a conditional *Dicer* KO in neurons of mouse cerebral cortex leads to neurite outgrowth and soma size deficits, suggesting a crucial role of miRNAs and mRNA stability in healthy neuronal development[24]. We predict that miRNA load increases as stem cells differentiate into neurons and this would make target transcripts less stable. In the context of *Dicer* deficient mouse brain, a decreased miRNA load would result in transcripts that are less bound by AGO2 and have changes in 3′UTR isoform length.

Powerful techniques for genome-wide measurement of mRNA half-life at high resolution use metabolic labeling of RNAs followed by sequencing. These methods include thiol (SH)-linked alkylation for the metabolic sequencing of RNA (SLAM-seq)[25] and RNA approach to equilibrium

sequencing (RATE-seq)[26]. RNA half-life is then measured from decay or accumulation of mRNAs following label incorporation of up to 24 hours or pulse-and-chase of modified RNA nucleotides. In addition, RNA half-life can be deduced from the equilibrium relationship between steady-state abundance and transcription rate. In this case, transcription rate is measured using samples from the short labeling times before RNA degradation[27]. Moreover, 3′-end based sequencing of metabolic labeled RNAs accurately defines each 3′UTR isoform of a gene allowing the analysis of mRNA half-life at the isoform level. Genome-wide mRNA half-life measurements have been performed across different cell types and organisms. However, most studies measured mRNA half-life in a single cell type or compared perturbed cell conditions[27–29]. The developmental dynamics of transcription rate and mRNA half-life shifts during human neurodevelopment and their effects on transcriptional buffering and 3′UTR isoform length are unknown.

Here, we systematically measured mRNA transcription rate and half-life at the 3′UTR isoform level using RATE-seq at three stages of human neurodevelopment: pluripotent stem cells, NPC, and cortical neurons. We revealed a widespread influence of mRNA stability shifts on the landscape of mRNA isoforms initially set by transcription rates in pluripotent stem cells and during the early stages of human neuronal development. Transcriptional buffering by compensatory half-life changes was observed in up to 50% of genes with changed transcription rates, and transcriptional boosting by additive half-life changes occurs in up to 15% of genes. Half-life only changes had roughly the same frequency as transcription rate only changes. Overall, 80% of altered genes were subjected to mRNA stability changes during neurodevelopment, and 50% of genes targeted by pluripotent transcription factors in iPSC were buffered or boosted by half-life changes. We also report reduced global transcript half-life in neurons and a destabilization of short isoforms in neurons relative to pluripotent cells that is coincident with the apparent lengthening of 3′UTR isoforms during human neurodevelopment. Furthermore, by measuring absolute miRNA abundance in all cell types, we found that an increased miRNA load in neurons mirrored the reduced mRNA half-life. Genes whose short isoforms were more unstable in neurons were more bound by AGO2, consistent with a miRNA-based mechanism. An increased miRNA load model that preferentially reduces short isoform half-life and abundance in neurons is proposed to explain the reduced global mRNA stability and 3′UTR lengthening events.

Results

RATE-seq captures global shifts in transcription rate and RNA steady-state during human neurodevelopment

To directly determine the contribution of transcription rates and RNA stability on mRNA steady-state abundance and 3′UTR isoform usage, we performed RATE-seq using a 3′-end sequencing method (QuantSeq) on cells collected and processed at different stages of iPSC-based neuronal development. The original RATE-seq approach measures absolute mRNA half-life from the incorporation kinetics of the modified nucleotide 4sU into newly transcribed nascent mRNAs over time points of up to 24 hours[26]. We have recently adapted RATE-seq to estimate a relative mRNA half-life by calculating the fold-change ratio between steady-state and transcription rate measured from early time points of 0.5 and 1 hour[10] (Fig 1A).

Cells were collected and processed simultaneously at three stages of human neurodevelopment *in vitr*o. We used a control iPSC line (d3-4 #37) and NPCs derived from them using a dual SMAD differentiation protocol. Cortical neurons (Neu) were matured from the NPCs for 4 weeks. The neurons were magnetically sorted based on the presence of specific cell surface proteins to exclude contaminant glia and progenitor cells[30]. RATE-seq results of these same control neuron samples have already been reported in comparison to isogenic Rett syndrome neurons[10]. Total RNA from all three cell types was collected at 5 different time points (0.5, 1, 4, 8 and 24 hours) in two replicate experiments, biotinylated and pulled down using streptavidin beads[26]. Exogenous 4sU labeled fly and unlabeled yeast whole-cell RNAs were spiked into the samples to reconstruct the human 4sU labeled fraction within the total mRNA pool and to control for pull-down efficiency and background contamination, respectively (Fig S1A-C). ERCC spike-in controls were added during the 3′-end sequencing library preparations. Transcription rate was measured at 0.5-hour and 1-hour timepoints when mRNA degradation has a negligible effect on the labeled RNAs for most genes[27]. In addition, steady-state abundance was measured from the input sample obtained at the 24-hour time point, matching the 4sU incubation conditions between samples. As expected, incubation with 4sU for 24 hours did not affect the viability of cells[31] (Fig S1D).

We found approximately 2100-6000 genes with changes in transcription rate or steady-state, or both, after neurodevelopmental transitions of iPSC to NPC, and NPC to Neu (Fig 1B) (Supplementary table 1). Supporting the ability of the differentiation protocol to produce NPCs and neurons, we observed a decrease in transcription rate in pluripotency marker genes and an increase in progenitor marker genes in the first transition, and then a decrease in progenitor marker genes and an increase in neuron marker genes in the second transition (Fig 1B). Similarly, mRNA abundance of progenitor and then neuron marker genes increased respectively at the steady-state upon the transitions while housekeeping genes were unchanged (Fig 1B). Principal component analysis (PCA) separated cell types into distinct clusters, and the abundance of endogenous and ERCC spike-in RNAs was highly reproducible between replicates (Fig 1C and Fig S1E-F). Furthermore, transcription rate shifts were validated for a selection of genes by qRT-PCR on nuclear fractions collected from independently differentiated cells (Fig 1D). Finally, we checked whether gene-body length, k-mers (dinucleotides or 6-mers) or promoter occupancy by transcription factors (ENCODE TFs[32] obtained from Harmonizome[33]) can predict the direction of the transcription rate and steady-state shifts. We used a sequence content-based random forest classifier to predict the effect of these features on steady-state abundance and transcription rate. No effect was observed for gene-body length, but an approximately 60% accuracy was found for 6-mers and the ENCODE TFs on the steady-state during both transitions (Fig 1E). The best predictor of changes in the direction of transcription rate during the iPSC to NPC transition was ENCODE TFs at 65% accuracy, but 6-mers were more accurate than ENCODE TFs in the NPC to Neu transition.

Importantly, change in transcription rate for 2453 and 2028 (approximately 40%-60% of all genes with transcription rate changes in Fig 1F) led to no change in their steady-state levels in both cell transitions. This result emphasizes that about half of transcription rate changes led to altered steady-state abundance, while the other half were buffered at the steady-state level. Similarly, 1729 and 740 genes, respectively, had altered steady-state abundance but no corresponding shift in transcription rate, suggesting that these genes were differentially regulated at the level of mRNA stability (Fig 1F). Together, the observed discrepancy between transcription

rate and steady-state shifts suggests a widespread contribution of mRNA stability in the regulation of genes during human neurodevelopment.

Absolute and relative mRNA half-life changes identify transcriptional buffering and boosting gene sets

Our profiling of transcription rate and steady-state changes indicated a genome-wide effect of mRNA stability during both neurodevelopmental transitions. To quantify this effect, we used the 4sU saturation kinetics over the entire 24-hour time course (Fig S1A). By measuring the absolute change in mRNA half-life in hours, we observed a mean mRNA half-life of 5.4 hours in iPSCs and NPCs, compared to 2.9 hours in neurons[10] (Fig 2A). The observed global mRNA destabilization suggests that RNA content might be reduced overall in neurons. To measure the per-cell RNA content and validate the global mRNA half-life decrease, we extracted and quantified RNA yields from whole-cell, nuclear, and cytoplasmic fractions from an equal number of cells (10 million). We found a striking ~3-fold reduction in cytoplasmic and whole-cell RNA during the transition from iPSC to neurons (Fig 2B). Moreover, the reduction occurs with no concomitant change in nuclear RNA, consistent with increased mRNA degradation in the cytoplasm (Fig 2B). These results are consistent with 4sU-labeling mRNA half-life measurement, which showed destabilization of mRNAs in neurons (Fig 2A).

We also used steady-state and transcription rate ratios to derive the relative shifts in mRNA half-life during neurodevelopment. The ratio and saturation curve methods employed to measure mRNA half-life were concordant with each other for the iPSC and NPCs, similar to our findings with the Neu samples[10] (Fig S2A-C) (Supplementary table 2). However, the absolute method (mRNA half-life in hours) required a sufficiently high transcription rate to measure the saturation curves accurately and yielded fewer quantified genes (Fig S2A). This was in contrast to the ability of the ratio method to measure relative fold-changes in genes with low transcription rates, as already shown in the neuron samples[10]. We observed that shifts in mRNA half-life during both cell transitions result in four modes of gene regulation: 1) steady-state changes caused by transcription rate only or 2) steady-state changes caused by mRNA half-life only, 3) transcription rate changes buffered or 4) boosted by mRNA half-life (Fig 2C-D).

We found that for approximately half of the regulated genes the mRNA half-life compensates for shifts in transcription rate resulting in buffering of steady-state levels after both developmental transitions. We observed 3013 genes (39%) after iPSC to NPC and 2204 genes (53%) after NPC to Neu transitions were fully buffered where mRNA stability shifts counteracted the changes in transcription (Fig 2D). Interestingly, another 1729 (23%) and 740 (18%) genes regulated steady-state through half-life only mechanisms after each developmental transition (Fig 2D). A further 1148 (15%) and 374 (9%) genes boosted transcription rate shifts through mRNA half-life, leading to amplified steady-state changes (Fig 2D). Finally, we detected 1795 (23%) and 804 (20%) genes with changes in steady-state mediated by transcription rate only changes.

Since ENCODE TFs partially predicted the direction of transcription rate and steady-state changes in the iPSC to NPC transition (Fig 1E), we examined whether networks of genes targeted by core pluripotency TFs escaped buffering and boosting. We first obtained gene targets of *NANOG*, *OCT4*, and Polycomb Repressive Complex 2 (PRC2) target genes (overlap of SUZ12, EED, H3K27 targets[33]) defined by previously published ChIP-seq experiments in hESC[33]. Approximately 30-40% of target genes bound by core pluripotency TFs were regulated by transcription rate only changes as iPSC differentiate into NPC (Fig 2E). In support of our transcription rate data, we found that the PRC2 targets were predominantly upregulated in NPCs, supporting the repressive role of these TFs in iPSCs (Fig 2E). As observed above with the 4 global modes of gene regulation, mRNA half-life changes still buffered or boosted transcription rate changes in nearly 50% of genes that are TF targets of NANOG, OCT4, or PRC2 (Fig 2E). The remaining 20% of NANOG and OCT4 target genes were regulated by mRNA half-life only. PRC2 had a >2-fold reduction in target genes controlled by mRNA half-life only (down to 8%). Together, our data highlight the widespread contribution of half-life in buffering and boosting mRNA levels (as observed in Fig 1E) as stem cells differentiate into progenitors and neurons. Buffering caused by shifts in half-life in different cell types leads to a substantial decoupling of steady-state levels from the transcription rate changes of these genes.

RNA stability controls 3′UTR length in pluripotent stem cells and during neurodevelopment

To define the contribution of mRNA stability to the specification of the 3′UTR usage during human neurodevelopment, we quantified mRNA transcription rate and half-life for all known 3′-end mRNA isoforms. In this analysis, the transcription rate of a 3′UTR isoform is the net result of gene transcription and a cleavage frequency at the corresponding pA site, measured in the first time points of 4sU labeling (0.5h and 1h). In addition, the high-resolution annotation defined by the roughly 200,000 pA sites previously compiled by Wang *et al.*[34] allowed the comprehensive comparison between 3′UTR isoforms of the same gene. In particular, we detected that approximately 10,000 genes contained multiple 3′UTR isoforms in their last exon. For most genes, two 3′UTR isoforms with a large distance between the pA sites accounted for at least 80% of the total mRNA abundance (Fig S3A-B). Based on this observation, we focused the analysis on the 2 most abundant 3′UTR isoforms for each gene.

We define short or long 3′UTR-containing mRNAs as those cleaved and polyadenylated at the most proximal or distal sites to the stop-codon, respectively (Fig 3A). We observed that, on average, processed short and long 3′UTR-containing mRNAs have equal transcription rates in iPSCs. However, the short 3′UTR-containing mRNA isoforms are generally more abundant than long isoforms at the steady-state (Fig 3A-B and Fig S3C). These data suggest that active degradation of mRNAs with long 3′UTRs in iPSC is widespread, extending the published observation of the *MECP2* long 3′UTR-containing mRNA degradation in human ESC[18] (Fig S3C).

In agreement with published results[35], we observed an accumulation of long 3′UTRs in NPCs and neurons (Fig 3C). We next sought to quantify the contribution of mRNA stability to the apparent switch to long 3′UTR isoforms during neurodevelopment. We observed in iPSCs that the half-life of mRNAs containing long 3′UTRs is less than in isoforms with short 3′UTRs when the difference in length is roughly equal to or greater than 500 nucleotides (Fig 3D). A typical long 3′UTR isoform is about 2 to 4 times less stable than short 3′UTR mRNAs in iPSCs (Fig 3D). Although less prominent in NPCs and neurons, a similar trend of less stable long isoforms was observed in those cell types (Fig 3D).

As previously described[18,36], we expected that a reduction in the half-life of the long 3′UTR isoforms in NPCs and neurons would explain the increased abundance of these mRNAs relative to the short isoforms. Surprisingly, we observed a gradual destabilization of the short 3′UTR-containing mRNAs after both transitions (Fig 3E). Long 3′UTR-containing mRNAs decreased their half-life only after the transition from NPCs to neurons (Fig 3E). Overall, the reduced half-life of the short 3′UTR isoforms was observed in the first transition to NPCs but is most evident as an approximately 2-fold decrease in the mRNA stability relative to the long 3′UTR isoforms when comparing iPSC to neurons. This effect on the short isoforms was observed when the difference in length between 3′UTR isoforms was > 500 nucleotides (Fig 3E) but not when it was < 500 nucleotides (Fig S3D). These data suggest that mRNA half-life has a greater impact on short compared to long 3′UTR-containing mRNAs during neurodevelopment (Fig 3E).

MicroRNA load accumulation in neurons co-occurs with destabilization of short 3′UTRs

To validate the reduced half-life of long 3′UTR isoforms in pluripotent stem cells using an alternative system, we reanalyzed a published mouse ESC RNA-seq dataset that compared *Dicer1*-KO with control[37]. *Dicer1* is part of the miRNA processing machinery, and its ablation causes a significant decrease in global miRNA activity[23]. After quantifying all 3′UTR isoforms using QAPA[38], we observed that in the absence of the miRNA-mediated degradation activity (*Dicer1-KO*), the mRNAs bound by AGO2 accumulate significantly (Fig 4A). Moreover, we found that long 3′UTR isoforms increased their abundance more than short 3′UTR isoforms in the *Dicer1*-KO ESCs (Fig 4A). These data suggest that miRNAs preferentially degrade long 3′UTRs in mouse ESCs.

To probe the contribution of miRNAs to the regulation of 3′UTR isoforms at a greater depth in our human cells, we performed small RNA sequencing spiked with small exogenous RNA normalized by cell numbers to quantify relative and absolute fold-changes in miRNA abundance. Reassuringly, pluripotent stem cell (miR-302, miR-200)[39,40] and neuron (miR-137, miR-7)[41] specific miRNAs accumulated in the appropriate cell types after both transitions (Fig 4B) (Supplementary table 3). Small RNA-seq results of the same control neuron samples have already been reported in comparison to isogenic Rett syndrome neurons[10]. As expected, we observed a

remodeling of the miRNA landscape with 676 (25%) and 585 (22%) miRNAs changing steady-state abundance after iPSC to NPC and NPC to Neu transitions, respectively (Fig 4B). Interestingly, coincident with the shutdown of pluripotent stem cell-specific miRNAs, the number of miRNAs with a large fold-change (>5) is 7 times higher after the first transition from iPSCs to NPCs (14%) compared with NPC to Neu (2.2%).

As the miRNA profile changes during neurodevelopment, the half-life of their mRNA targets is expected to change. Indeed, the half-lives of mRNAs defined in TargetScan[42] as targeted by pluripotent stem cell-specific miRNAs like miR-302 and miR-200 are the most increased upon transition to NPCs (Fig S4A). Further, normalization of miRNA levels using the small RNA spike-ins showed a ~4-fold global increase in miRNA abundance per cell from iPSC to Neu, with the main increase during the transition to NPC (Fig 4C and Fig S4B). In addition, a ~4-fold per cell reduction in RNA yield is detected when comparing iPSC to Neu, with the main reduction during the transition to NPCs (Fig S4C). These observations demonstrate a drastic 20-fold increase in available miRNAs per cell for each mRNA molecule in neurons relative to iPSC, defined as miRNA load (Fig S4D). The 6-fold increased miRNA load in NPCs relative to iPSC is not accompanied by a change in mean half-life. However, the higher miRNA load in neurons is accompanied by the observed reduction of mean mRNA half-life (Fig 2A).

To independently test the increase in miRNA load during neurodevelopment, we verified whether there was an increase in the association of mRNAs with AGO2 in human neurons relative to iPSCs. AGO2 is part of the miRNA-induced silencing complex (miRISC) and an increased association with mRNAs indicate an increased miRNA degradation activity. We performed AGO2 RNA immunoprecipitation (RIP) followed by qRT-PCR of four genes with short 3′UTR-containing mRNAs with reduced half-life upon the transition to neurons. Anti-IgG antibody was used as a negative control. We observed a ~3-fold increase in AGO2 binding upon transition to neurons in all four genes with minimal signal for the IgG control (Fig 4D). This enhanced binding of AGO2 to mRNAs indicates an increase in the chance of miRNA annealing with its target motif, being proportional to the miRNA load.

Discussion

We studied mRNA stability regulation to identify post-transcriptional mechanisms leading to transcriptional buffering and mRNA isoform usage during human neurodevelopment. We used RATE-seq to calculate relative transcript half-life changes as the ratio of standard RNA-seq steady-state and transcription rate measured using short 4sU incorporation times[27]. We found that roughly 20% of differentially regulated genes during transitions of iPSC to NPC, and NPC to Neu, were regulated by transcription rate only changes. Strikingly, roughly 80% of differentially regulated genes experienced shifts in mRNA half-life. The predominant mode of half-life regulation was to buffer transcription rate changes by mRNA half-life leading to no change in steady-state levels, with approximately half of the genes in this set. The remaining genes were either regulated through half-life only or by boosting transcription changes. These findings demonstrate that mRNA stability is an equal partner with transcription rate in establishing the steady-state levels of the transcriptome in human neurodevelopment. Such stability-dependent regulation in pluripotent cells agrees with widespread *AGO2* binding to the transcriptome and the functional significance of miRNA biogenesis machinery on proper self-renewal and differentiation[19,20,37,43]. Given the consistently high frequency of buffered, boosted and half-life only gene sets in iPSC, NPC and Neu, it is highly likely that equivalent modes of gene regulation will be found when differentiating iPSC into other human cell types to model development and disease. Importantly, the same modes of gene regulation should be widely found in vivo, as we have already verified buffered and other gene regulation modes in a high-confidence RNA-seq resource of control and *Mecp2* null mouse brain samples[10].

We found evidence that similar RNA stability mechanisms cooperate with the core pluripotency transcription factors and adjust the abundance of their targets. In agreement with our data, a recent study showed that a knockdown of *OCT4* or *NANOG* alters the expression of many RBP genes[44]. In addition, we detected buffering of roughly half of the transcriptionally upregulated targets. We used the ratio method to calculate mRNA half-life fold-changes for genes with low transcription rates, circumventing the detection limits of the standard saturation curve method. Buffering of lowly transcribed genes may facilitate the elimination of leaky transcripts by mRNA stability mechanisms. Overall, our data indicate the role of mRNA stability in buffering genes

controlled by pluripotency transcription factors. Strikingly, only 30-40% of their target genes were regulated by transcription rate only. Our datasets may prove useful to define the transcription rate changes on target genes of other epigenetic or transcription factors implicated in human neurodevelopment and its disorders.

Average mRNA half-life analysis revealed a synchronized genome-wide increase in degradation capacity in cortical neurons compared to iPSCs and NPC by roughly 2-fold. We observed a global 4-fold increase in miRNA molecules per cell in neurons compared to iPSC and a concurrent 5-fold decrease in total RNA content. Rapid miRNA turnover previously observed in the mammalian brain relative to ESC and NPC aligns with our data on higher degradation capacity, considering that the degradation of target mRNAs and miRNAs are a linked process[45,46]. We propose that increasing miRNA load synchronizes the mRNA degradation of most genes in neurons.

We detected a genome-wide active degradation of long 3'UTR-containing mRNAs in pluripotent stem cells. This observation supports and expands the previously reported findings of the *MECP2* gene[18]. In that report, the RBP PUM1 and pluripotent-specific miR-200a and miR-302c destabilized the long 3'UTR *MECP2* isoform in human ESCs[18]. Here, we focused on the role of miRNAs because they can be quantitatively assessed by RNA-seq on a per-cell basis. Our results suggest that miRNAs degrade most long 3'UTR isoforms in pluripotent stem cells. Our conclusion is corroborated by the observation of widespread AGO2 binding in mouse ESCs[37] and our re-analysis of the changes in the abundance of short versus long 3'UTR isoforms in *Dicer1* knockout mouse ESCs. Potential roles for RBPs in this process cannot be excluded, but calculating RBP load changes during differentiation at the level of absolute protein molecule number per cell will require exquisitely sensitive proteomics methods, and their consensus binding sites are more degenerate than miRNA target sites for target analyses.

We also observed a striking destabilization of the short 3'UTR-containing mRNAs upon differentiation to neurons that led ultimately to equal stability between long and short 3'UTR isoforms. Recent data from mouse hippocampal neurons corroborates the negligible difference between long and short 3'UTR isoforms in neurons compared to the difference observed in our

iPSCs[47]. In addition, measurements of 3'UTR isoforms in mouse fibroblasts also showed no difference in stability[48]. Overall, our data suggest that differential short 3'UTR isoform stability is related to the degree of stemness, with the most prominent difference present in pluripotent stem cells.

It has been previously proposed that a switch to more distal 3′-end cleavage sites through APA causes the overall increase in 3'UTR length consistently observed during neurodevelopment of different species[11–13]. We confirmed the relative accumulation of longer 3'UTR isoforms in neurons leading to an increase in the average 3'UTR length in these cells. However, in contrast to the APA model, we detected a substantial contribution of mRNA stability through the destabilization of short 3'UTR isoforms upon differentiation to neurons. In addition, we probed the accumulation of long 3'UTR isoforms at a greater temporal resolution than previously studied by including the measurements at the NPC stage. We observed a considerable increase in relative long 3'UTR isoform abundances upon transition to NPC, again driven by an mRNA stability mechanism. Overall, we show that mRNA stability shifts contribute to the apparent lengthening of 3'UTRs in neurodevelopment that is first detected at the NPC stage and is maximal in neurons.

Taken together, we propose a miRNA load model for the RNA stability-based regulation of 3'UTR isoform usage during human neurodevelopment. In human pluripotent stem cells, short and long 3'UTR-containing mRNAs are equally transcribed and processed. However, the long isoforms are degraded more efficiently and closer to a maximum capacity than the short isoforms because there is a limited number of miRNA molecules. Neuronal differentiation leads to an accumulation of miRNAs, expanding the pool of the available miRNAs per mRNA molecule. This leads to preferential degradation of the short 3'UTR-containing mRNAs in neurons since long 3'UTR-containing mRNAs have a smaller margin for a further decrease in RNA stability.

Acknowledgments

This study was funded by grants from the Canadian Institutes of Health Research (CIHR; PJT-148746, PJT-168905, and ERARE Team ERT 161303 to J.E.); the Canada First Research Excellence Fund (Medicine by Design Cycle I: J.E.); the Col. Harland Sanders Rett Syndrome Research Fund at the University of Toronto (J.E.); the Ontario Brain Institute (POND Network:

J.E.); and John Evans Leaders Fund/Canada Foundation for Innovation (JELF/CFI: J.E.); Canada Research Chairs Program (M.D.W. and J.E.); Beta Sigma Phi International Endowment Fund (J.E.); Early Researcher Award from the Ontario Ministry of Research and Innovation (M.D.W.); Genome Canada Disruptive Innovation in Genomics Grant (M.D.W to support K.E.Y.); NIH grant R35GM128680 and the University of Colorado RNA Bioscience Initiative (O.R.); David Steven Cant Scholarship (M.M.).

Author contributions

Conceptualization, M.M., D.C.R., A.N., O.S.R., and J.E.; Investigation, M.M., D.C.R., A.N., K.Y., W.W., A.P., and J.L.; Software, M.M., and A.N.; Writing – original draft, M.M., D.C.R., and J.E.; Writing - Review & Editing, M.M., D.C.R., A.N., K.Y., P.P., O.S.R., M.D.W., and J.E.; Visualization, M.M., and D.C.R.; Supervision & Funding Acquisition, O.S.R., M.D.W., and J.E.

Data availability

The original sequencing data that support the findings of this study can be accessed from GEO using the access number GSE191168 (wild type Neu [10]) and wild type iPSC and NPC will be made available on publication. The mouse data reanalyzed in this study can be downloaded from GSE139349 (whole-cell RNA-seq and Ago2 HEAP).

Code availability

The computational code used in this manuscript is available at GitHub https://github.com/JellisLab/stabilome-neurodevo

Figures

Figure 1. Changes in transcription rate during neurodevelopment do not automatically result in altered mRNA steady states.

A, schematics of experimental outline (created with BioRender.com) for simultaneous quantification of transcription rate, mRNA half-life, and steady-state mRNA level. iPSC and iPSC-derived neuronal progenitor cells and cortical neurons were pulse-labeled with 4sU, and at designated time points total RNA was harvested. 4sU-labeled *Drosophila melanogaster* (fly) and unlabeled *Saccharomyces cerevisiae* (yeast), and ERCC spike-in RNAs were added as indicated and used as pull-down efficiency, non-specific binding, library preparation, and sequencing controls, respectively. Steady-state mRNA levels were quantified from an aliquot of the 24-hour time point (non-biotinylated and unprocessed). This experiment was repeated for a total of two replicates. B, scatter-plots depicting genome-wide changes in transcription rate and steady-state after both iPSC to NPC and NPC to Neu transitions. Cell type marker genes for iPSC, NPC, Neu are highlighted by green, red, and blue colours, respectively. C, principal component analysis of sequencing data from iPSC, NPC, and Neu showing significant separation between cell types. D, transcription rate fold-changes determined by RATE-seq (X-axis) were validated using an alternative approach. Nuclear RNA fold-change measured by qRT-PCR (Y-axis) was used as a proxy for transcription rate shifts of genes selected to cover a large spectrum of fold-changes, including genes with no changes (n= 2). E, random forest classifier for predicting steady-state and transcription rate direction based on gene-body sequence and promoter occupancy by ENCODE transcription factors. An individual point represents a classifier percent accuracy trained on a particular sample of training and test data. A boxplot for a model is obtained from 50 resamplings. F, overlap of genes altered at transcription rate and/or steady-state during neuronal development.

Figure 2. Widespread changes in mRNA half-life directly alter steady-state abundance to buffer or boost transcription rate changes during neuronal development.

A, global decrease in the absolute half-life after transition to Neu. The mean half-life for each cell type is shown in the figure panel (a decrease in the mean half-life from 5.4 hours in iPSC and NPC to 2.9 hours in Neu was detected). B, RNA yield in μg per 1 million cells from whole-cell, cytoplasmic and nuclear fractions measured at iPSC, NPC and Neu stages of differentiation. C,

example of genes with changes in transcription rate or half-life only, buffering or boosting by half-life changes. Fold-change (Y-axis, log_2) of corresponding assay. D, number and % of all genes with changes in the 4 modes of regulation during cell transitions: buffering, half-life only, boosting and transcription rate only. The height of each group corresponds to the number of genes. E, number and % of genes targeted by the pluripotent TF denoted on top with changes in the 4 modes of regulation during the iPSC to NPC transition.

Figure 3. RNA stability controls the 3′UTR landscape in pluripotent stem cells and during neurodevelopment.

A, representative sequencing read peaks in the iPSC transcription rate (upper) and steady-state (lower) samples, corresponding to polyadenylation sites (arrowheads) in the 3′UTR as an example of genes with active degradation of long 3′UTR-containing mRNAs. Y-axis represents the number of sequencing reads. B-C, comparison in detected abundance of long and short 3′UTR-containing mRNAs. Genes with a distance >500 nucleotides between polyadenylation sites corresponding to long and short 3′UTR-containing mRNAs are included in the boxplots. (B) steady-state and transcription rate read counts fold-change (Y-axis) between long and short 3′UTR-containing mRNAs in iPSC. (C) steady-state read counts fold-change (Y-axis) between long and short 3′UTR-containing mRNAs during neuronal development. D, mRNA half-life fold-change (Y-axis) during neuronal development split according to distance (X-axis) between long and short 3′UTR-containing mRNAs. E, mRNA half-life fold-change (Y-axis) separately for long and short 3′UTR-containing mRNAs split by developmental transitions. Genes with a distance >500 nucleotides between polyadenylation sites corresponding to long and short 3′UTR-containing mRNAs are included in the boxplots.

Figure 4. miRNA accumulation co-occurs with the destabilization of short 3′UTR in cortical neurons.

A, effect of a decrease in global miRNA activity by *Dicer1* ablation in mESC. Left panel, mRNA fold-change (Y-axis) of Ago2 targets split according to Ago2 binding signal (X-axis) measured in counts. Right panel, number of genes with a great mRNA fold-change for either long or short 3′UTR-containing mRNAs. Ago2 with mRNA diagram is created with BioRender.com. B, scatter plot showing changes in miRNA abundance during both developmental transitions. X-axis

represents the basal abundance of each mature miRNA detected, and Y-axis represents their fold-change. Cell-type marker miRNAs for iPSC, NPC, Neu are highlighted by green, red, and blue colours, respectively. C, DESeq2 miRNA steady-state level fold-change after spike-in normalization. Total RNA was extracted from the same number of cells, to which we added the same mass of a commercially available small RNA spike-in mixture. This spike-in mixture contains 54 different small RNA molecules covering a wide range of sequence composition and molar concentrations. The absolute abundance of miRNAs is increasing for most miRNAs during neuronal development. D, AGO2 fold enrichment in cortical neurons relative to iPSCs measured by RIP-qRT-PCR. The 3′UTR-containing mRNA of the selected genes is destabilized after the transition to Neu.

Supplementary Figures

Figure S1. Quality control of the RATE-seq experiment.

A, 4sU incorporation kinetic curves of human mRNA normalized to fly spike-in RNA. B, representative agarose gels showing 4sU-labeled pulled-down RNA for each time-point in iPSCs and NPCs. The presence of human (*Hs*) and fly ribosomal RNAs are denoted by arrows. C, read counts of spike-in RNAs: unlabeled yeast relative to fly RNA for all time points in iPSC, NPC, Neu. Unlabeled yeast RNA spike-in was used as a negative pull-down control (background control). The absence of yeast RNA indicates that the streptavidin-biotin pull-down of 4sU labeled RNAs had minimal contamination of unlabeled human RNAs but was readily detectable in the steady-state samples. D, 4sU does not cause changes in neuronal viability or cell numbers at the dose and times used in the experiment. E, Pearson's correlation between replicates of the human RNA for each time-point. F, Pearson's correlation of ERCC spike-in RNAs between replicates (relative), and spike-in concentration and sequencing measured abundance (absolute) used for all time points in iPSC, NPC, Neu. ERCC spike-in RNA was used as a control for library prep quality. High Pearson's correlations indicate the high quality of the library samples.

Figure S2. Half-life measured with 4sU saturation curve and ratio method.

A, percentage of genes with measured half-life depending on transcription rate. B, the accuracy of half-life shown against half-life magnitude. CI_{50} stands for 50% confidence interval. Black points denote genes with poorly fit saturation curves and are removed from analysis in panel C. C, comparison of half-life estimated with two methods for well-measured genes selected in panel B.

Figure S3. Summary of long and short 3′UTR isoform quantification.

A, fraction of transcription rate reads assigned to the highest or top 2 highest in abundance 3′UTR isoforms. B, number of genes (Y-axis) depending on distance (X-axis) between long and short 3′UTR-containing mRNAs. C, representative sequencing read peaks in the iPSC transcription rate (upper) and steady-state (lower) samples, corresponding to polyadenylation sites (arrowheads) in the 3′UTR as an example of genes with active degradation of long 3′UTR-containing mRNAs. Y-axis represents the number of sequencing reads. D, mRNA half-life fold-change (Y-axis) separately for long and short 3′UTR-containing mRNAs split by developmental transitions. Genes with a distance <500 nucleotides between polyadenylation sites corresponding to long and short 3′UTR-containing mRNAs are included in the boxplots.

Figure S4. Impact of miRNA on mRNA half-life.

A, mRNA fold-change of miRNA targets relative to all genes in iPSC to NPC transition depending on miRNA fold-change magnitude. B, fold-change for endogenous miRNAs and small RNA spike-ins after developmental transitions. C, RNA yield in μg per 1 million of cells from whole-cell measured at iPSC, NPC and Neu stages of differentiation normalized to iPSCs. D, fold-change for absolute miRNA abundance (from Fig 4C), RNA yield (from Fig S4C) and derived from them miRNA load quantified during iPSC to NPC, NPC to Neu, and iPSC to Neu transitions.

Methods

iPSC cultures, NPC, and neuronal differentiation

iPSC line #37 (WT) was previously described[49]. This cell line was generated and cultured under the approval of the SickKids Research Ethics Board and the Canadian Institutes of Health Research Stem Cell Oversight Committee. Methods to culture and differentiate the cell types are exactly as described[10]. In brief, iPSCs were cultured on BD hESC-qualified matrigel (BD) in mTeSR medium (STEMCELL Technologies). To generate Neural Progenitor Cells (NPCs), iPSCs were aggregated as Embryoid Bodies (EBs) and replated after 7 days for growth in N2 media + laminin (1 ml/ml). After 7 days, neural rosettes were manually picked and transferred to poly-L-ornithine + laminin-coated wells and picked again after another 7 days. NPCs were grown as a monolayer and split every 5-7 days in NPC media containing DMEM/F12, N2, B27, 1x non-essential amino acid (NEAA), 2 mg/ml Heparin, 1 mg/ml laminin. To generate neurons, NPCs were plated on poly-L-ornithine + laminin-coated plates and cultured for 3 weeks in neural differentiation medium (Neurobasal, N2, B27, 1 mg/ml laminin, 1x penicillin-streptomycin, 10ng/ml BDNF, 10ng/ml GDNF, 200 mM ascorbic acid, and 10 mM cAMP).

4sU metabolic labeling and RNA extractions

The RATESeq experiment was performed exactly as described[10]. In brief, culture media of all three cell types was replaced with media supplemented with 100 µM 4sU (Sigma-Aldrich) reconstituted in DMSO. Cells were incubated at 37°C, 5% CO_2, and total RNA was harvested at 0.5, 1, 4, 8, and 24 hours after the addition of 4sU. Metabolic labeling was designed such that all time points were collected together. Cells were quickly washed twice in ice-cold PBS 1x and lysed on the plate by addition of 1 mL of Trizol (Thermo Fisher Scientific). Total RNA was extracted according to manufacturer instructions. The steady-state samples were prepared from a 5µg aliquot of the 24 h time-point to which we added 0.5 µg of both unlabeled *yeast* and 4sU-labeled S2 *fly* spike-in RNAs. Cellular viability in the presence of 100 µM 4sU was monitored up to 24 h of treatment on parallel cultures by using Trypan blue staining and live/dead cell counting.

Biotinylation and pull down of 4sU-labeled RNAs

Biotinylation and pull down of labelled RNA was performed exactly as described[10]. In brief, 50 µg of total RNA was mixed with 5 µg unlabeled *yeast* RNA and 5 µg 4sU-labeled

S2 *fly* RNA. 4sU Biotinylation was performed by adding 120 µL of 2.5× citrate buffer (25 mM citrate, pH 4.5, 2.5 mL EDTA) and 60 µL of 1 mg/mL HPDP-biotin (ThermoFisher Scientific) to the RNA samples for each time point. The solution was incubated at 37°C for 2 h. Samples were extracted twice with acid phenol, and once with chloroform. RNA was precipitated, pelleted and resuspended in 1× wash buffer (10 mM Tris-HCl, pH 7.4, 50 mM NaCl, 1 mM EDTA). Biotinylated RNAs were purified using the µMACS Streptavidin microbeads system (Miltenyi Biotec). 50 µL Miltenyi beads per sample were pre-blocked, applied to microcolumns and washed 5x. Beads were demagnetized and eluted off the column, and columns placed back on the magnetic stand. A total of 200 µL beads were mixed with each sample of biotinylated RNA and rotated at room temperature for 20 min. Samples were loaded into the microcolumns, washed 3x with wash buffer A (10 mM Tris-HCl, pH 7.4, 6 M urea, 10 mM EDTA), and washed 3x with wash buffer B (10 mM Tris-HCl, pH 7.4, 1 M NaCl, 10 mM EDTA). Final RNA samples were eluted with 5 washes of 1x wash buffer supplemented with 0.1 M DTT, and the flow-through was collected. Purified RNA samples were precipitated at −20°C overnight. Samples were spun and resuspended in 20 µL water. RNA quality was assessed by running 3 µL of samples on a 1.5% agarose gel.

Cellular fractionation and yield quantification

Cytoplasmic and nuclear RNA fractions were extracted from 10^7 cells of all cell types using the Cytoplasm and Nuclear RNA Purification Kit (Norgen) following the manufacturer's instructions for cultured animal cells. After the membrane lysis step and separation of cytoplasmic and nuclear fractions, 1 µg of each yeast and fly total spike-in RNAs was added to each fraction before proceeding to column binding and RNA elution. This procedure kept a ratio of 1 µg of spike-in RNAs per cellular fraction of 10^7 cells, which was later used for cell number normalizations by qRT-PCR. Whole-cell RNA lysates (total) were extracted from 10^7 cells using the Direct-zol RNA miniprep Plus (Zymo Research) following the manufacturer's instructions. 1 µg of each yeast and fly total spike-in RNAs were added to the cellular lysate (in Tryzol) before proceeding to the column binding step. The optional on-column DNase digestion steps of both kits were included.

cDNA synthesis and qRT-PCR

cDNAs were synthesized from the nuclear RNA fractions using SuperScript III reverse transcriptase (ThermoFisher) with random hexamer primers according to the manufacturer's instructions. For qRT-PCR, we used SYBR Select PCR Master Mix (ThermoFisher). ΔCt values of the target genes were normalized against the spike-in RNAs from yeast and fly genes. Absolute fold-change measurements relative to cell numbers were calculated using the $2^{-(\Delta\Delta Ct)}$ method, averaged between technical and subsequently biological replicates to achieve an average fold difference. The primers used are described in the primers table.

AGO2 RIP-qRT-PCR

For each AGO2 RNA-immunoprecipitation, iPSCs and Neurons were grown on two 10 cm dishes each. Lysate collection and RNA immunoprecipitations were performed using the Magna RIP™ RNA-Binding Protein Immunoprecipitation Kit (Millipore, #17-700) following the manufacturer's instructions. Briefly, cells were washed twice in ice-cold PBS and lysed in RIP Lysis Buffer complemented with Protease and RNase Inhibitor Cocktails. Before immunoprecipitation, magnetic protein A/G beads were prepared with 5 µg of rabbit anti-AGO2 (Millipore, #03-110) or normal rabbit IgG antibodies (from the Magna RIP kit). The RIP lysates were centrifuged at 12,000 g for 10 min at 4 °C. The lysates were then split into three fractions: 10 µl was set aside as an input sample and frozen at − 80 °C until the RNA purification step, and 100 µl of RIP lysate was added to the tube containing the AGO2 or IgG-bound beads. The RIP lysates were incubated with the antibody-bead complexes overnight at 4 °C. Total RNA was isolated and purified from the input lysates and the immunoprecipitation samples using phenol/chloroform and resuspended in 10 µl of RNase-free water.

RNA was reversed transcribed to cDNA using SuperScript III reverse transcriptase (ThermoFisher) and random hexamer primers according to the manufacturer's instructions. qRT-PCR was performed using SYBR Select PCR master mix on a ViiA7 Real-Time PCR System (ThermoFisher). Fold changes were calculated by the $2^{-(\Delta\Delta Ct)}$ method and 18S rRNA as a normalizing control. The primers used are described in the primers table. For comparison of mRNA levels between input and IP samples, results are displayed as a fold-change relative to the iPSC input sample.

miRNA extraction and spike-in strategy

To calculate relative and absolute differences in the miRNA population in iPSC, NPC and Neurons, small RNAs were extracted from two replicates of both lines using the same number of cells followed by the addition of a set of spike-in RNAs exactly as described[10]. Small RNAs were extracted from 500,000 cells of each line using the SPLIT RNA extraction Kit (Lexogen). A set of 52 RNA spike-ins (QIAseq miRNA Library QC Spike-Ins – Qiagen) spanning a wide range of molar concentrations were added to the recovered RNAs. Sequencing libraries were made using the Small RNA library preparation kit NEBNext (NEB). Sequencing was performed on the Illumina HiSeq 2500 using the Rapid Run mode.

Library preparation and RNA-sequencing

RNA-seq libraries were prepared for each time-point and steady-state sample exactly as described[10]. In brief, the QuantSeq 3′ mRNA-Seq Library Prep Kit FWD for Illumina (Lexogen) was automated on the NGS WorkStation (Agilent) at The Centre for Applied Genomics (TCAG). PCR cycle numbers were determined using the PCR Add-on Kit for Illumina (Lexogen). All steady-state samples were processed with 250 ng of total RNA input. To minimize variability between time-points within a batch, RNA samples were processed with the same total RNA input with a minimum of 100 ng of total RNA used. Each sample was spiked-in with ERCC RNA Spike-In Control Mix 1 (Ambion) prior to the start of library preparation. Library quality and quantity were measured at The Centre for Applied Genomics (TCAG) with Bioanalyzer (Agilent) and KAPA qPCR (Roche). Sequencing was performed at TCAG on the Illumina HiSeq 2500 with single-end 100bp read length yielding 40 to 50 million reads.

Processing of raw sequencing reads

Processing was performed exactly as described[10]. In brief we trimmed reads in 4 steps using cutadapt version 1.10[50]. First, we removed adapters exactly at the 3′-end of the reads. Second, we removed internal or long stretches of adapter. Third, we trimmed low-quality bases at the 3′-end of the reads. Finally, we removed poly-A tail at the 3′-end of the reads.

Generation of custom hybrid genome index and reads alignment with STAR

We generated a custom genome index to accommodate the quantification of yeast, fly, and ERCC spike-in RNA exactly as described[10]. Annotations (gencode version 29, flybase version all-r6.22, saccharomyces_cerevisiae.gff from yeastgenome.org, custom for ERCC) and genomes (hg38, dm6, sacCer3, ERCC from ThermoFisher) for all species and ERCC were combined and then processed with STAR version 2.6.0c (--sjdbOverhang 100). Finally, reads are aligned to hybrid genome with STAR version 2.6.0c (default settings)[51].

Quantification of RNA abundance

Poly-A sites were obtained from PolyA_DB version 3 and converted to hg38 coordinates with *liftOver* (UCSC)[34,52]. Reads with MAPQ < 2 were filtered out. Finally, usage of poly-A sites was defined as a sum of reads whose 3′-ends are falling within 20bp upstream and 10bp downstream of the poly-A sites. The sum was counted with a custom Python script using *pybedtools*, *pysam*, *pypiper*[53–55]. The abundance of mature miRNAs was quantified with *mirdeep2* pipeline[56]. Reads were preprocessed and collapsed with *mapper.pl* script (-e -h -j -k AGATCGGAAGAGCACA -l 18 -m -v) and quantified with *quantifier.pl* script, using hairpin and mature sequences obtained from miRbase[57].

Normalization of human read counts with fly spike-ins

First, reads were assigned as originating from either human, fly, yeast or ERCC, based on alignment to hybrid genome. Then, human raw counts were divided by the sum of all fly spike-in raw counts. Since both human and fly RNAs are 4sU-labeled, this normalization to fly spike-ins reconstructs the fraction of 4sU-labeled human RNA at each time point.

Spike-in RNA usage clarification

Fly 4sU spike-ins were used to normalize pulled down human counts for all time points (0.5-hour to 24-hour) to assist with transcription rate and absolute half-life calculations. Yeast spike-ins were used as quality control for contamination in the pull-downs. ERCC spike-ins were used to control for sequencing quality. First, it directly estimates sequencing error magnitude, excluding biological variation. Second, it shows the capacity of 3′-end QuantSeq to reconstruct

molar concentrations. Note that the steady-state sample was not normalized with spike-ins and is separate from the 24-hour time point sample.

Transcription rate and half-life measurements

The transcription rate was estimated from 0.5-hour and 1-hour time points exactly as described[10]. This estimate assumes that RNA degradation is negligible for most genes at time points before 1 hour. First, human normalized counts at 1-hour time point are divided by 2 to create a new approximate replicate at 0.5-hour time point. Division by 2 is done to account for a twice longer period of transcription in a 1-hour time point sample. To clarify, the addition of an approximate replicate was motivated by the reduction in the standard error of $\log_2$FC estimated with *DESeq2*. To compare the transcription rate between cell types using *DESeq2*, human normalized counts are further quantile normalized between replicates of the same cell type with *normalize.quantiles* (*preprocessCore* package)[58,59]. Half-life was estimated in 2 separate ways: fit of the 4sU saturation curve and the ratio of steady-state to transcription rate. For the 4sU saturation curve method, the half-life is estimated in a 2-step procedure. First, normalized counts are fit with *nls* (nonlinear least squares from *stats* package) to approximate the true number of counts Y at each timepoint. Then, in a second pass, normalized counts are fit again with *nls*, but now correcting for the increase in variance using weights set as 1/Y. Confidence intervals are estimated with *confint* function (*stats* package). For the ratio method, the half-life is estimated with *DESeq2* using raw human counts from 0.5-hour, 1-hour and steady-state samples (design =~ assay). The assay is a 2-level factor one for transcription rate and one for steady-state. 0.5-hour and 1-hour samples correspond to the transcription rate.

The shift in average mRNA half-life between cell types

To describe the global absolute mRNA half-life shift, we filtered out unreliable genes from the saturation curve method. First, both gene transcription rate and steady-state should be above the bottom 10% quantile. Second, the ratio of 50% confidence interval and estimate of the half-life should be below 0.75. Then, the mean half-life was calculated from the filtered set of genes.

Random forest prediction of up and down-regulated genes in transcription rates

Features of the classifier are frequencies of k-mers in 3′UTR or gene-body, calculated using *oligonucleotideFrequency* (*Biostrings* package)[60]. Predicted variable denotes genes that are either up or downregulated in transcription rate. Thresholds for the human data were:

1. TR_{up}: $\log_2 FC_{TR} > 1$ & $padj < 0.1$
2. TR_{down}: $\log_2 FC_{TR} < -1$ & $padj < 0.1$

Data was split into 75% and 25% for training and test sets. The classifier is trained using *randomForest* (*randomForest* package)[61].

Processing of mouse datasets

Whole-cell RNA-seq data for wild-type and *Dicer1* knock-out mESC were downloaded from *GSE139348*. Ago2 peaks data in mESC were downloaded from the supplementary materials of the Li *et al* study[37]. The abundance of 3′UTR isoforms for all samples is estimated using the QAPA standard pipeline[62].

Transcription factors targets data

Targets of transcription factors NANOG, OCT4, and a repressive complex PRC2 defined by ChIP-seq were downloaded from molecular signatures database (MSigDB)[33,63,64].

Computer code

All computer code used for sequencing analysis is available on GitHub at https://github.com/JellisLab/stabilome-neurodevo

Primer list for qRT-PCRs

DUSP26	Forward	CTGCCGACTTCATCCACC
	Reverse	TGACTTTCTTGATGGCCTCC
NNAT	Forward	AAACCCTCTTTCTCGACCAC
	Reverse	ATCCTACCCAGTAAATGCAGC
BTG3	Forward	AGTGAAACCCAGTTCGGTG
	Reverse	GAGGATAGTGATTCTGATGGCC
PLPP3	Forward	GACCTCTTCTGCCTCTTCATG
	Reverse	CAGAGCACAGCGTCATTTATTG
SVIP	Forward	AAAACAAATTGCTACATCCGGG
	Reverse	AGCCCACTTGATTGCAGAT
HEYL	Forward	GAAACGCAGAGGGATCATAGAG
	Reverse	ATCCACCGTCATCTGCAAG
ARL2	Forward	AGATGAAGCAGAAAGAGCGG
	Reverse	AGGGTCTTGATGTTGAAGCC
LHFPL2	Forward	TCTTCAATGTCTGTGGGCTG
	Reverse	CTCCAGGTTTGTAGGCAGATG
VASH	Forward	CCAGAAGGAGACAGGCAAG
	Reverse	TCGGATTTGATAGCCCACTTC
TUBB3 - Fly	Forward	TGTCGCGTGTGAAACACTTC
	Reverse	AGCAGGCGTTTCCAATCTG
ALG9 - Yeast	Forward	CACGGATAGTGGCTTTGGTGAACAATTAC
	Reverse	TATGATTATCTGGCAGCAGGAAAGAACTTGGG
TAF10 - Yeast	Forward	ATATTCCAGGATCAGGTCTTCCGTAGC
	Reverse	GTAGTCTTCTCATTCTGTTGATGTTGTTGTTG

18S	Forward	GATGGGCGGCGGAAAATAG
	Reverse	GCGTGGATTCTGCATAATGGT
FLNB	Forward	CCTTCAAGGTGGCTGTCACTGA
	Reverse	CCCTCAACAGTTATGCCAAGCC
HTT	Forward	CTCTGGTGTCAGATACTGCTGC
	Reverse	CTCCTCTTCTCCAGACATCTGG
PDHB	Forward	TGTAACTGTGGAAGGAGGCTGG
	Reverse	CATCAGCACCAGTGACACGAAC
RNF126	Forward	GGAAGAGACCAGGAGCACAGAA
	Reverse	GATGCCGAAAGCAAACTGTCCG
DUSP26	Forward	CTGCCGACTTCATCCACC
	Reverse	TGACTTTCTTGATGGCCTCC
NNAT	Forward	AAACCCTCTTTCTCGACCAC
	Reverse	ATCCTACCCAGTAAATGCAGC
BTG3	Forward	AGTGAAACCCAGTTCGGTG
	Reverse	GAGGATAGTGATTCTGATGGCC

References

1. Li, T. *et al.* Smg6/Est1 licenses embryonic stem cell differentiation via nonsense-mediated mRNA decay. *EMBO J.* 34, 1630–1647 (2015).

2. Jia, J. *et al.* Regulation of pluripotency and self- renewal of ESCs through epigenetic-threshold modulation and mRNA pruning. *Cell* 151, 576–589 (2012).

3. Prashad, S. & Gopal, P. P. RNA-binding proteins in neurological development and disease. *RNA Biol.* 18, 972–987 (2021).

4. Lee, J.-A. *et al.* Cytoplasmic Rbfox1 regulates the expression of synaptic and autism-related genes. *Neuron* 89, 113–128 (2016).

5. Shu, H. *et al.* FMRP links optimal codons to mRNA stability in neurons. *Proc. Natl. Acad. Sci. U. S. A.* 117, 30400–30411 (2020).

6. Ebert, M. S. & Sharp, P. A. Roles for microRNAs in conferring robustness to biological processes. *Cell* 149, 515–524 (2012).

7. Timmers, H. T. M. & Tora, L. Transcript buffering: A balancing act between mRNA synthesis and mRNA degradation. *Mol. Cell* 72, 10–17 (2018).

8. Lee, J. E. *et al.* The PARN Deadenylase Targets a Discrete Set of mRNAs for Decay and Regulates Cell Motility in Mouse Myoblasts. *Plos Genet* 8, e1002901 (2012).

9. Brennand, K. J. *et al.* Modelling schizophrenia using human induced pluripotent stem cells. *Nature* 473, 221–225 (2011).

10. Rodrigues, D. C. *et al.* Buffering of transcription rate by mRNA half-life is a conserved feature of Rett syndrome models. *BioRxiv* (2021) doi:10.1101/2021.12.11.472181.

11. Hilgers, V. *et al.* Neural-specific elongation of 3' UTRs during Drosophila development. *Proc. Natl. Acad. Sci. U. S. A.* 108, 15864–15869 (2011).

12. Smibert, P. *et al.* Global patterns of tissue-specific alternative polyadenylation in Drosophila. *Cell Rep.* 1, 277–289 (2012).

13. Ulitsky, I. *et al.* Extensive alternative polyadenylation during zebrafish development. *Genome Res.* 22, 2054–2066 (2012).

14. Miura, P., Shenker, S., Celia, A.-A., Westholm, J. O. & Lai, E. C. Widespread and extensive lengthening of 3' UTRs in the mammalian brain. *Genome Res.* 23, 812–825 (2013).

15. Agarwal, V., Lopez-Darwin, S., Kelley, D. R. & Shendure, J. The landscape of alternative polyadenylation in single cells of the developing mouse embryo. *Nat. Commun.* 12, 5101 (2021).

16. Brumbaugh, J. *et al.* Nudt21 Controls Cell Fate by Connecting Alternative Polyadenylation to Chromatin Signaling. *Cell* 172, 106-120.e21 (2018).

17. Mayr, C. Regulation by 3′–Untranslated Regions. *Annu Rev Genet* 51, 1–24 (2016).

18. Rodrigues, D. C. *et al.* MECP2 Is Post-transcriptionally Regulated during Human Neurodevelopment by Combinatorial Action of RNA-Binding Proteins and miRNAs. *Cell Reports* 17, 720–734 (2016).

19. Kanellopoulou, C. *et al.* Dicer-deficient mouse embryonic stem cells are defective in differentiation and centromeric silencing. *Genes Dev.* 19, 489–501 (2005).

20. Wang, Y., Medvid, R., Melton, C., Jaenisch, R. & Blelloch, R. DGCR8 is essential for microRNA biogenesis and silencing of embryonic stem cell self-renewal. *Nat. Genet.* 39, 380–385 (2007).

21. Su, H., Trombly, M. I., Chen, J. & Wang, X. Essential and overlapping functions for mammalian Argonautes in microRNA silencing. *Genes Dev.* 23, 304–317 (2009).

22. Qi, J. *et al.* microRNAs regulate human embryonic stem cell division. *Cell Cycle* 8, 3729–3741 (2009).

23. Teijeiro, V. *et al.* DICER1 is essential for self-renewal of human embryonic stem cells. *Stem Cell Reports* 11, 616–625 (2018).

24. Hong, J., Zhang, H., Kawase-Koga, Y. & Sun, T. MicroRNA function is required for neurite outgrowth of mature neurons in the mouse postnatal cerebral cortex. *Front. Cell. Neurosci.* 7, 151 (2013).

25. Herzog, V. A. *et al.* Thiol-linked alkylation of RNA to assess expression dynamics. *Nature Methods* 14, 1198–1204 (2017).

26. Neymotin, B., Athanasiadou, R. & Gresham, D. Determination of in vivo RNA kinetics using RATE-seq. *RNA* 20, 1645–1652 (2014).

27. Muhar, M. *et al.* SLAM-seq defines direct gene-regulatory functions of the BRD4-MYC axis. *Science* 360, 800–805 (2018).

28. Schwalb, B. *et al.* TT-seq maps the human transient transcriptome. *Science* 352, 1225–1228 (2016).

29. Schofield, J. A., Duffy, E. E., Kiefer, L., Sullivan, M. C. & Simon, M. D. TimeLapse-seq: Adding a temporal dimension to RNA sequencing through nucleoside recoding. *Nature Methods* 15, 221–225 (2018).

30. Rodrigues, D. C. *et al.* Shifts in Ribosome Engagement Impact Key Gene Sets in Neurodevelopment and Ubiquitination in Rett Syndrome. *Cell Reports* 30, 4179-4196.e11 (2020).

31. Friedel, C. C. & Dölken, L. Metabolic tagging and purification of nascent RNA: implications for transcriptomics. *Mol. Biosyst.* 5, 1271–1278 (2009).

32. Consortium, E. P. A user's guide to the encyclopedia of DNA elements (ENCODE). *PLoS Biol.* 9, e1001046 (2011).

33. Ben-Porath, I. *et al.* An embryonic stem cell-like gene expression signature in poorly differentiated aggressive human tumors. *Nat. Genet.* 40, 499–507 (2008).

34. Wang, R., Nambiar, R., Zheng, D. & Tian, B. PolyA_DB 3 catalogs cleavage and polyadenylation sites identified by deep sequencing in multiple genomes. *Nucleic Acids Research* 46, D315–D319 (2018).

35. Mayr, C. Evolution and biological roles of alternative 3'UTRs. *Trends Cell Biol.* 26, 227–237 (2016).

36. Zheng, D. *et al.* Cellular stress alters 3′UTR landscape through alternative polyadenylation and isoform-specific degradation. *Nature Communications* 9, 2268 (2018).

37. Li, X. *et al.* High-resolution in vivo identification of miRNA targets by Halo-enhanced Ago2 pull-down. *Mol. Cell* 79, 167-179.e11 (2020).

38. Ha, K. C. H., Blencowe, B. J. & Morris, Q. QAPA: a new method for the systematic analysis of alternative polyadenylation from RNA-seq data. *Genome Biol.* 19, 45 (2018).

39. Gill, J. G. *et al.* Snail and the microRNA-200 family act in opposition to regulate epithelial-to-mesenchymal transition and germ layer fate restriction in differentiating ESCs. *Stem Cells* 29, 764–776 (2011).

40. Judson, R. L., Babiarz, J. E., Venere, M. & Blelloch, R. Embryonic stem cell-specific microRNAs promote induced pluripotency. *Nature Biotechnology* 27, 459–461 (2009).

41. Zahr, S. K., Kaplan, D. R. & Miller, F. D. Translating neural stem cells to neurons in the mammalian brain. *Cell Death Differ.* 26, 2495–2512 (2019).

42. Agarwal, V., Bell, G. W., Nam, J.-W. & Bartel, D. P. Predicting effective microRNA target sites in mammalian mRNAs. *eLife* 4, 101 (2015).

43. Leung, A. K. L. *et al.* Genome-wide identification of Ago2 binding sites from mouse embryonic stem cells with and without mature microRNAs. *Nature Structural and Molecular Biology* 18, 237–244 (2011).

44. Dvir, S. *et al.* Uncovering the RNA-binding protein landscape in the pluripotency network of human embryonic stem cells. *Cell Rep.* 35, 109198 (2021).

45. Kim, C. K. & Pak, T. R. miRNA degradation in the mammalian brain. *Am. J. Physiol. Cell Physiol.* 319, C624–C629 (2020).

46. Krol, J. *et al.* Characterizing light-regulated retinal microRNAs reveals rapid turnover as a common property of neuronal microRNAs. *Cell* 141, 618–631 (2010).

47. Tushev, G. *et al.* Alternative 3′ UTRs Modify the Localization, Regulatory Potential, Stability, and Plasticity of mRNAs in Neuronal Compartments. *Neuron* 98, 495-511.e6 (2018).

48. Spies, N., Burge, C. B. & Bartel, D. P. 3′ UTR-isoform choice has limited influence on the stability and translational efficiency of most mRNAs in mouse fibroblasts. *Genome Res* 23, 2078–2090 (2013).

49. Cheung, A. Y. L. *et al.* Isolation of MECP2-null Rett Syndrome patient hiPS cells and isogenic controls through X-chromosome inactivation. *Human Molecular Genetics* 20, 2103–2115 (2011).

50. Martin, M. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet.journal* 17, 10–12 (2011).

51. Dobin, A. *et al.* STAR: ultrafast universal RNA-seq aligner. *Bioinformatics (Oxford, England)* 29, 15–21 (2013).

52. Karolchik, D. *et al.* The UCSC Table Browser data retrieval tool. *Nucleic Acids Res* 32, D493–D496 (2004).

53. Dale, R. K., Pedersen, B. S. & Quinlan, A. R. Pybedtools: a flexible Python library for manipulating genomic datasets and annotations. *Bioinformatics* 27, 3423–3424 (2011).

54. Danecek, P. *et al.* Twelve years of SAMtools and BCFtools. *Gigascience* 10, giab008 (2021).

55. Sheffield, N. C., Stolarczyk, M., Reuter, V. P. & Rendeiro, A. F. Linking big biomedical datasets to modular analysis with Portable Encapsulated Projects. *Gigascience* 10, giab077 (2021).

56. Friedländer, M. R., Mackowiak, S. D., Li, N., Chen, W. & Rajewsky, N. miRDeep2 accurately identifies known and hundreds of novel microRNA genes in seven animal clades. *Nucleic Acids Res* 40, 37–52 (2012).

57. Kozomara, A., Birgaoanu, M. & Griffiths-Jones, S. miRBase: from microRNA sequences to function. *Nucleic Acids Res* 47, D155–D162 (2019).

58. Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome biology* 15, 550 (2014).

59. Bolstad, B. *preprocessCore: A collection of pre-processing functions. R package version 1.56.0.* (https://github.com/bmbolstad/preprocessCore, 2021).

60. Pagès, H., Aboyoun, P., Gentleman, R. & DebRoy., S. *Biostrings: Efficient manipulation of biological strings. R package version 2.62.0*. (https://bioconductor.org/packages/Biostrings, 2021).

61. Liaw, A. & Wiener, M. Classification and Regression by randomForest. *R News* 3, 18–22 (2002).

62. Ha, K. C. H., Blencowe, B. J. & Morris, Q. QAPA: a new method for the systematic analysis of alternative polyadenylation from RNA-seq data. *Genome biology* 19, 45 (2018).

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

64. Liberzon, A. *et al.* Molecular signatures database (MSigDB) 3.0. *Bioinformatics* 27, 1739–1740 (2011).

Fig 1.

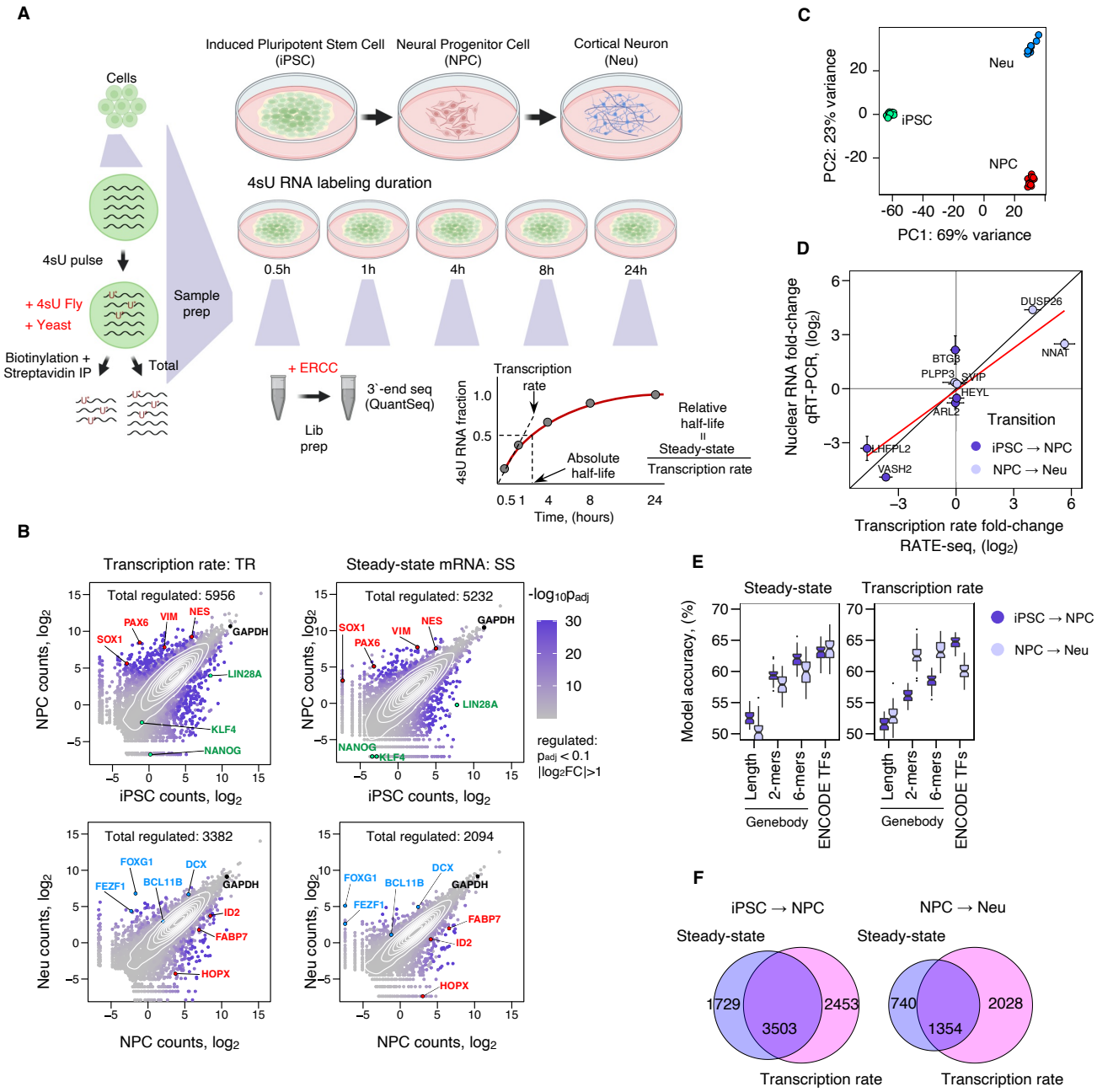

A
Induced Pluripotent Stem Cell (iPSC)
Neural Progenitor Cell (NPC)
Cortical Neuron (Neu)
Cells
4sU RNA labeling duration
0.5h
1h
4h
8h
24h
4sU pulse
+ 4sU Fly
+ Yeast
Sample prep
Biotinylation + Streptavidin IP
Total
+ ERCC
3`-end seq (QuantSeq)
Lib prep
Transcription rate
4sU RNA fraction
Absolute half-life
Relative half-life
Steady-state Transcription rate
Time, (hours)
C
PC2: 23% variance
PC1: 69% variance
Neu
iPSC
NPC
D
Nuclear RNA fold-change qRT-PCR, (log2)
Transcription rate fold-change RATE-seq, (log2)
DUSP26
NNAT
BTG3
PLPP3
SVIP
HEYL
ARL2
LHFPL2
VASH2
Transition
iPSC → NPC
NPC → Neu
B
Transcription rate: TR
Steady-state mRNA: SS
Total regulated: 5956
Total regulated: 5232
Total regulated: 3382
Total regulated: 2094
NPC counts, log2
iPSC counts, log2
Neu counts, log2
-log10padj
regulated: padj < 0.1 |log2FC|>1
SOX1
PAX6
VIM
NES
GAPDH
LIN28A
KLF4
NANOG
FOXG1
DCX
FEZF1
BCL11B
ID2
FABP7
HOPX
E
Steady-state
Transcription rate
Model accuracy, (%)
iPSC → NPC
NPC → Neu
Length
2-mers
6-mers
Genebody
ENCODE TFs
F
iPSC → NPC
NPC → Neu
Steady-state
1729
3503
2453
740
1354
2028
Transcription rate

Fig 2.

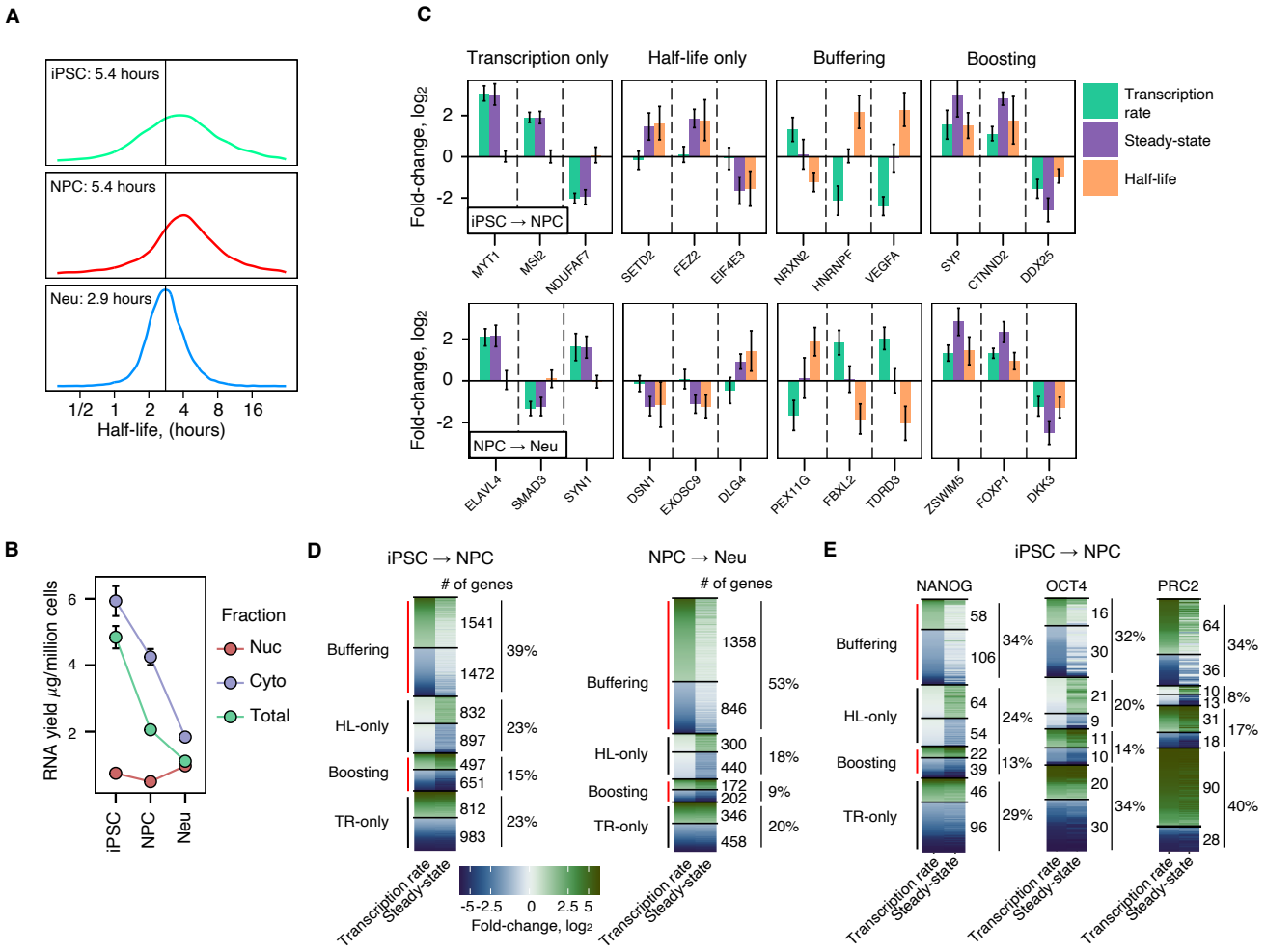

A
iPSC: 5.4 hours
NPC: 5.4 hours
Neu: 2.9 hours
1/2 1 2 4 8 16
Half-life, (hours)
C
Transcription only
Half-life only
Buffering
Boosting
Transcription rate
Steady-state
Half-life
Fold-change, log2
iPSC → NPC
MYT1 MSI2 NDUFAF7 SETD2 FEZ2 EIF4E3 NRXN2 HNRNPF VEGFA SYP CTNND2 DDX25
NPC → Neu
ELAVL4 SMAD3 SYN1 DSN1 EXOSC9 DLG4 PEX11G FBXL2 TDRD3 ZSWIM5 FOXP1 DKK3
B
RNA yield µg/million cells
Fraction
Nuc
Cyto
Total
iPSC NPC Neu
D
iPSC → NPC
NPC → Neu
# of genes
Buffering 1541 1472 39%
HL-only 832 897 23%
Boosting 497 651 15%
TR-only 812 983 23%
Buffering 1358 846 53%
HL-only 300 440 18%
Boosting 172 202 9%
TR-only 346 458 20%
Transcription rate
Steady-state
-5 -2.5 0 2.5 5
Fold-change, log2
E
iPSC → NPC
NANOG
OCT4
PRC2
Buffering 58 106 34% 16 30 32% 64 36 34%
HL-only 64 54 24% 21 9 20% 10 13 8%
Boosting 22 39 13% 11 10 14% 31 18 17%
TR-only 46 96 29% 20 30 34% 90 28 40%
Transcription rate
Steady-state

Fig 3.

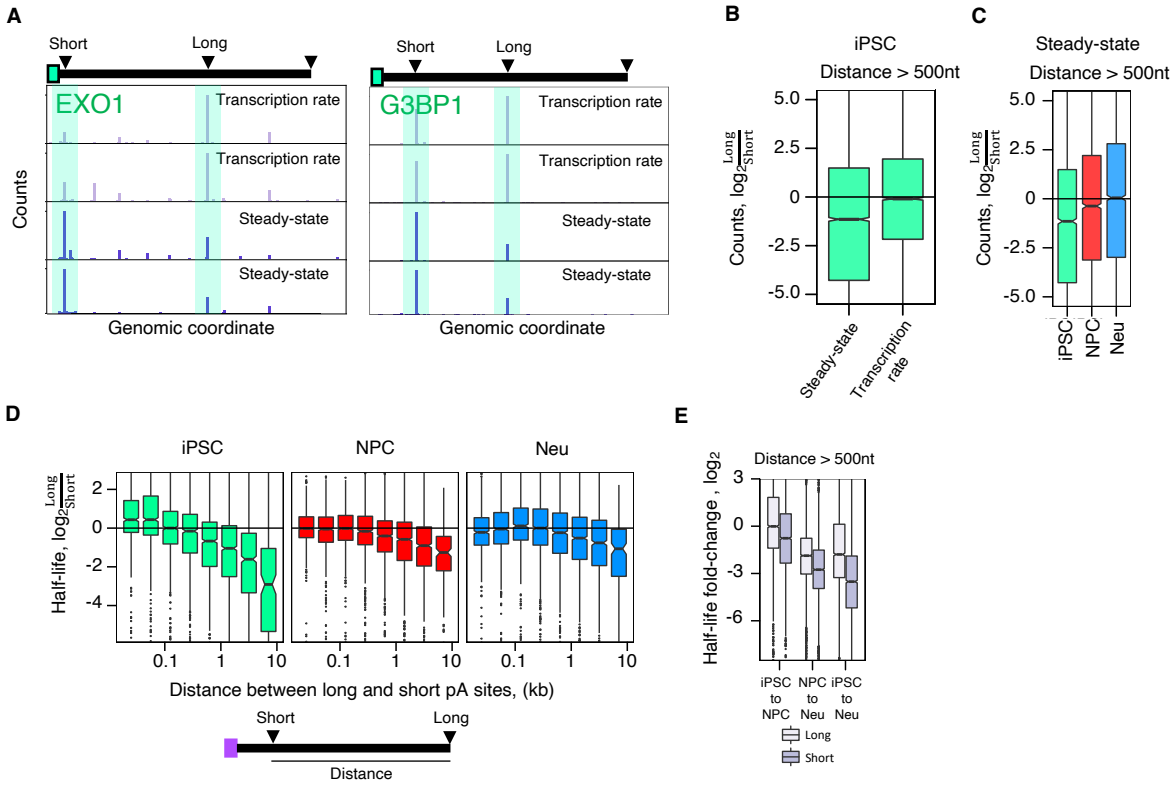
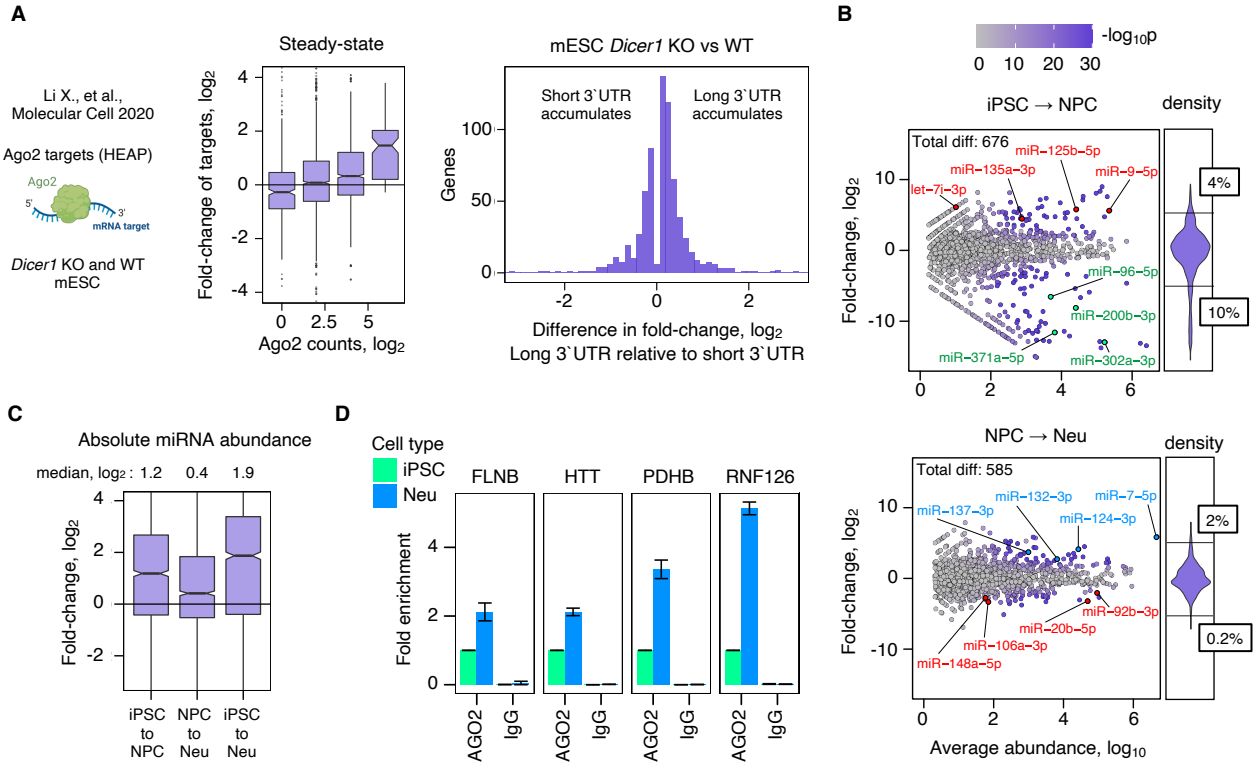

A
Short
Long
EXO1
G3BP1
Transcription rate
Steady-state
Counts
Genomic coordinate
B
iPSC
Distance > 500nt
Counts, log2 Long/Short
Steady-state
Transcription rate
C
Steady-state
Distance > 500nt
iPSC
NPC
Neu
D
iPSC
NPC
Neu
Half-life, log2 Long/Short
Distance between long and short pA sites, (kb)
Distance
E
Distance > 500nt
Half-life fold-change , log2
iPSC to NPC
NPC to Neu
iPSC to Neu
Long
Short

Fig 4.


A
Li X., et al.,
Molecular Cell 2020
Ago2 targets (HEAP)
Ago2
mRNA target
Dicer1 KO and WT
mESC
Steady-state
Fold-change of targets, log2
Ago2 counts, log2
mESC Dicer1 KO vs WT
Short 3`UTR
accumulates
Long 3`UTR
accumulates
Genes
Difference in fold-change, log2
Long 3`UTR relative to short 3`UTR
B
-log10p
iPSC → NPC
density
Total diff: 676
miR−125b−5p
miR−135a−3p
miR−9−5p
let−7i−3p
miR−96−5p
miR−200b−3p
miR−371a−5p
miR−302a−3p
4%
10%
Fold-change, log2
NPC → Neu
Total diff: 585
miR−132−3p
miR−7−5p
miR−137−3p
miR−124−3p
miR−92b−3p
miR−20b−5p
miR−106a−3p
miR−148a−5p
2%
0.2%
Average abundance, log10
C
Absolute miRNA abundance
median, log2 : 1.2 0.4 1.9
Fold-change, log2
iPSC
to
NPC
NPC
to
Neu
iPSC
to
Neu
D
Cell type
iPSC
Neu
FLNB
HTT
PDHB
RNF126
Fold enrichment
AGO2
IgG