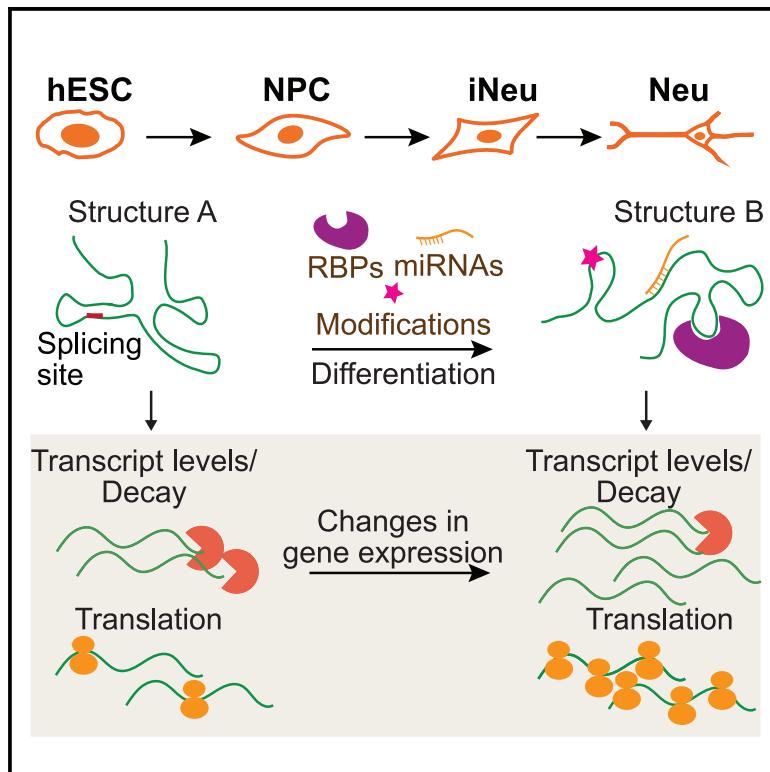


Genome-wide RNA structure changes during human neurogenesis modulate gene regulatory networks

Graphical abstract



Authors

Jiaxu Wang, Tong Zhang, Zhang Yu, ..., Finnlay R.P. Lambert, Roland G. Huber, Yue Wan

Correspondence

wany@gis.a-star.edu.sg

In brief

Wang et al. performed global structure mapping, expression, translation, and decay analysis during human neuronal differentiation. They observed that the hESC transcriptome is structurally more accessible than that of differentiated cells, and structure changes are associated with cellular factors, highlighting the complex interplay among RBPs, miRNAs, RNA structure, and gene regulation.

Highlights

- Systematically assayed the role of RNA structures and gene regulation during neurogenesis
- The hESC transcriptome is structurally more accessible than that of differentiated cells
- Structure changes in differentiation are associated with cellular factors, including RBPs
- Demonstrated interplay among RBP, structure, and gene regulation on LIN28A



Resource

Genome-wide RNA structure changes during human neurogenesis modulate gene regulatory networks

Jiaxu Wang,^{1,5} Tong Zhang,^{1,5} Zhang Yu,¹ Wen Ting Tan,¹ Ming Wen,¹ Yang Shen,¹ Finnlay R.P. Lambert,^{1,2} Roland G. Huber,³ and Yue Wan^{1,4,6,*}

¹Stem Cell and Regenerative Biology, Genome Institute of Singapore, A*STAR, Singapore 138672, Singapore

²Division of Biomedical Sciences, Warwick Medical School, University of Warwick, Coventry, UK

³Bioinformatics Institute, A*STAR, Singapore 138671, Singapore

⁴Department of Biochemistry, National University of Singapore, Singapore 117596, Singapore

⁵These authors contributed equally

⁶Lead contact

*Correspondence: wan_y@gis.a-star.edu.sg

<https://doi.org/10.1016/j.molcel.2021.09.027>

SUMMARY

The distribution, dynamics, and function of RNA structures in human development are under-explored. Here, we systematically assayed RNA structural dynamics and their relationship with gene expression, translation, and decay during human neurogenesis. We observed that the human ESC transcriptome is globally more structurally accessible than differentiated cells and undergoes extensive RNA structure changes, particularly in the 3' UTR. Additionally, RNA structure changes during differentiation are associated with translation and decay. We observed that RBP and miRNA binding is associated with RNA structural changes during early neuronal differentiation, and splicing is associated during later neuronal differentiation. Furthermore, our analysis suggests that RBPs are major factors in structure remodeling and co-regulate additional RBPs and miRNAs through structure. We demonstrated an example of this by showing that PUM2-induced structure changes on LIN28A enable miR-30 binding. This study deepens our understanding of the widespread and complex role of RNA-based gene regulation during human development.

INTRODUCTION

Neuronal development is an extremely complex process that involves extensive gene regulation. A comprehensive understanding of how cells are regulated is needed to decipher the complexity of our brain. Beyond the primary sequence and RNA expression levels, recent developments have shown the importance of RNA structures in almost every step of the RNA life cycle (Ganser et al., 2019), regulating cellular processes including transcription (Dethoff et al., 2012), translation (Kozak, 2005; Mao et al., 2014), localization (Martin and Ephrussi, 2009), and decay (Garneau et al., 2007). In addition to the static landscape of RNA structures inside cells, studying RNA structure dynamics and their regulators and cellular functions is key to understanding RNA structure function relationships. Recent high-throughput structure probing have interrogated RNA structures across different vertebrate processes, including during zebrafish embryogenesis (Beaudoin et al., 2018; Shi et al., 2020), upon virus infections (Mizrahi et al., 2018), and across cellular compartments (Sun et al., 2019b). Although RNA structures have been shown to be associated with severe neurological diseases (Bernat and Disney, 2015; Kolb and Kissel, 2011) and important for diverse neuronal processes, including directing mRNA localiza-

tion to the tips of axons in neurons (Jung et al., 2012), the full extent of RNA structure-based gene regulation during neuronal development is still largely unclear.

Here, to understand the role of RNA structure dynamics during human neuronal development, we used high-throughput RNA structure mapping together with RNA sequencing, ribosome profiling, and RNA decay studies for combinatorial analysis. We compared different analytic strategies to best identify structural changes between cellular states using RNA footprinting experiments as ground truth and showed that RNA structures are highly accessible in human embryonic stem cells (hESCs) and are dynamic between cell types. We further identified additional regulators important for structure changes and demonstrated the importance and complexity of structure-based gene regulation during human neuronal development.

RESULTS

To study RNA structural dynamics during human neurogenesis, we differentiated hESCs into neurons using an established differentiation protocol (STAR Methods). We performed high-throughput RNA structure probing experiments *in vivo* using icSHAPE (Spitale et al., 2015), as well as other functional



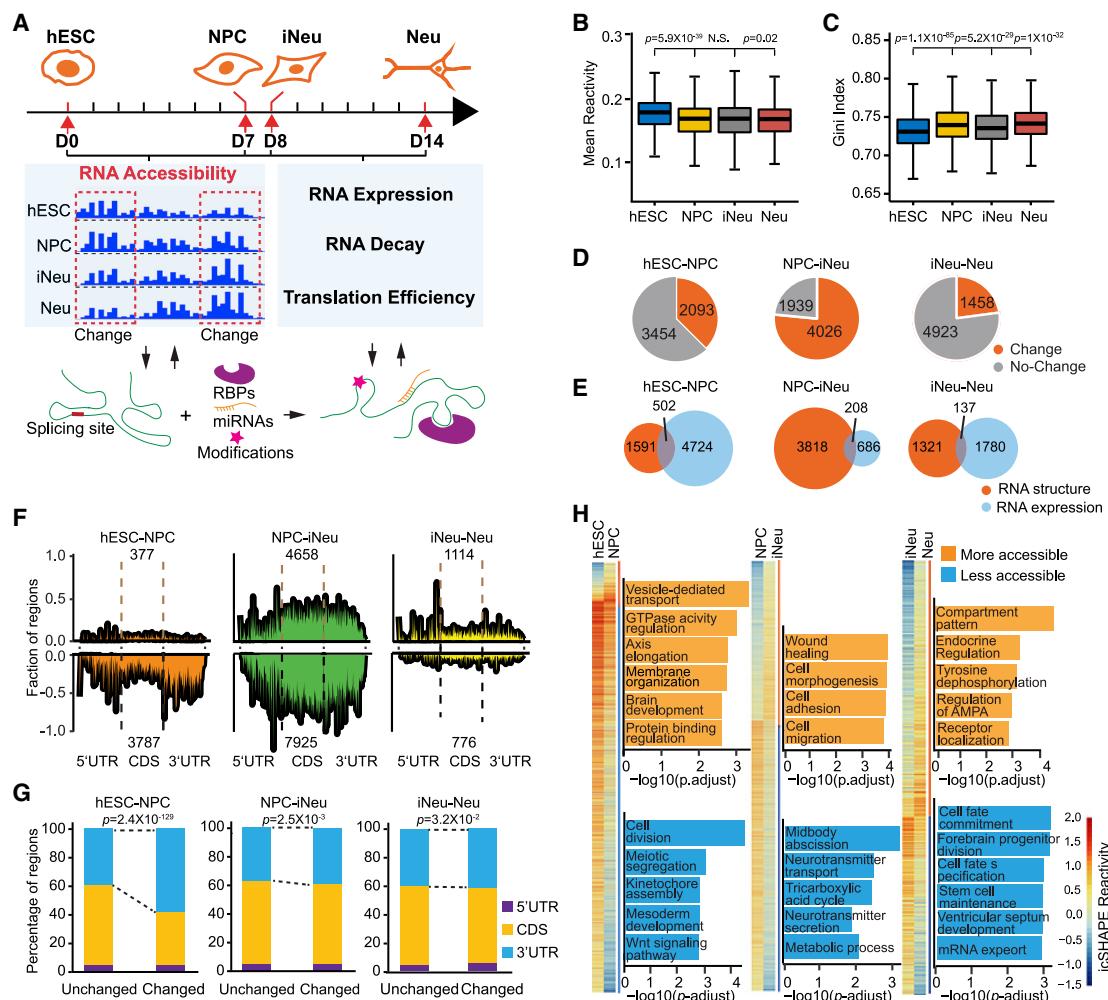


Figure 1. RNA structure profiling during neural development

(A) Schematic of our experimental workflow.

(B and C) Boxplots showing the distribution of icSHAPE reactivity (B) and Gini index (C) per gene at the different time points of neuronal differentiation. The three comparisons were performed between neighboring time points. p values were computed using two-sided Student's t tests.

(D) Pie charts showing the numbers of genes with (orange) and without (gray) significant reactivity changes during neuronal differentiation.

(E) Venn diagram showing the overlap of RNA structure changed genes and RNA expression level changed genes during neurogenesis.

(F) The relative proportions of significantly changed regions on mRNAs that are becoming more accessible (top panel) and less accessible (bottom panel) between hESC-NPC (orange), NPC-iNeu (green), and iNeu-Neu (yellow). The number of changing regions is stated for both the top and bottom panels. The lengths of 5' UTR, CDS, and 3' UTR are scaled to the same length.

(G) Stacked bar plots showing the numbers of windows without (left bar in each plot) and with (right bar in each plot) significant reactivity changes located in the 5' UTR (purple), CDS (yellow), and 3' UTR (blue). The p values were computed using hypergeometric tests.

(H) Left: heatmap showing the reactivity of structure-changing windows at different stages. Right: GO enrichment of windows with significant reactivity changes between stages. Enriched GO terms in windows that are becoming more accessible are shown in yellow, while enriched GO terms in windows that are becoming less accessible are shown in blue.

genomic experiments, including RNA sequencing (Stark et al., 2019), ribosome profiling (Brar and Weissman, 2015), and RNA decay (Chen et al., 2008) in hESCs (day 0), neural precursor cells (NPCs; day 7), immature neurons (iNeu; day 8), and neurons (Neu; day 14; Figure 1A; Figure S1A; Table S1; STAR Methods). The combination of these high-throughput experiments enabled us to study the impact of RNA structure on gene expression, translation, and decay as the stem cells differentiate.

We observed that RNA reactivity signals from NAI-N3- and DMSO-treated cells cluster independently from each other, as expected (Figure S1B). icSHAPE reactivity, gene expression, translation, and decay experiments showed high correlation between biological replicates ($R \geq 0.97, 0.99, 0.89$, and 0.99 , respectively) at each of the different time points, indicating that our high-throughput measurements are reproducible (Figures S1C–S1F). Additionally, Gene Ontology (GO) term analysis of transcripts that undergo TE (translation efficiency), decay, and

gene expression changes during neuronal differentiation showed that they are enriched for neuronal processes, as expected (Figure S1G). We detected 3,867, 4,881, 5,598, and 5,521 genes at hESC, NPC, iNeu, and Neu stages using all four methods, respectively (Figures S2A and S2B; Table S2). We observed a total of 2,910 genes that are shared and successfully detected in all stages using the four methods (Figures S2C and S2D), indicating that we are interrogating a large fraction of the transcriptome.

Metagene analysis of icSHAPE reactivities on the hESC transcriptome showed that we could capture previously reported structural patterns in human transcriptomes (Kertesz et al., 2010), including the presence of a 3 nt periodicity in the coding region and a negative correlation between Gini index and TE for the transcriptome (Figures S3A and S3B). The Gini index is a measure of structural heterogeneity along a transcript, whereby a high Gini index indicates increased heterogeneity in icSHAPE reactivity and increased structure. Interestingly, our structure probing experiments showed that hESC RNAs have the highest icSHAPE reactivities and lowest Gini index compared to RNAs from differentiated cells (Figures 1B and 1C), indicating that hESC RNAs are more structurally accessible. This result holds even when we calculate the icSHAPE reactivities for transcripts that are highly expressed across all four stages (Figures S4A and S4B). We also did not observe a good correlation between average reactivity and gene expression along a transcript ($R = 0.11$), indicating that the increased reactivity of transcripts in hESCs is not due to their increased transcript abundance (Figure S4C). Additionally, we validated this increase in RNA accessibility in hESCs versus differentiated cells using two orthogonal structure probing methods, SHAPE-MaP and DMS-MaP-Seq (Figures S4D–S4F), confirming that hESC RNAs are indeed more structurally accessible in hESCs than in differentiated cells.

Although the static picture of RNA structure landscapes has been mapped in many transcriptomes, being able to accurately detect structure changes remains to be a challenge. To find a statistically robust way to identify RNA structure changes sensitively and accurately across various time points, we compared different strategies of analyzing icSHAPE changes using RNAs that are known to change structure. This includes six riboswitch and two riboSNitch RNAs (Figure S5A) that are known to change structure in the presence of ligands or mutations, respectively. We performed icSHAPE and traditional footprinting (the gold standard for in-solution structure probing) to determine the analytic method that best captures the structure changes observed by footprinting (Figure S5B; Table S3). Both diff-BUMHMM and the t test have been previously used to identify RNA structure changes in the transcriptome, but we observed that NOISeq outperforms both methods in its ability to identify structure-changing regions sensitively and accurately (Figure S5C; Table S3). We further tested different window sizes and the extent of fold changes required to call a region as differentially changing, and we observed that a window size of 20 bases and a fold change of 1.5 times is optimal for calling differential reactivities, as these parameters enable sensitive detection of structure dynamics with low amount of false positives (Figures S5D and S5E). We used NOI-seq with these parameters for all of our downstream analyses.

Upon neuronal differentiation, we observed that 2,093 genes (4,164 regions), 4,026 genes (12,538 regions), and 1,458 genes (1,890 regions) showed reactivity changes between hESC-NPC, NPC-iNeu, and iNeu-Neu differentiation, respectively (Figure 1D; Figure S6A). Most of the transcripts had one or two regions of reactivity changes (Figure S6B), and the reactivity changes rarely exceed 20 bases in each region (Figures S6C and S6D). Transcripts with structure changes show poor overlap with transcripts that undergo gene expression changes during differentiation, indicating that RNA structure dynamics add an additional layer of information during differentiation (Figure 1E). Most of the RNA structural changes from hESCs to NPCs showed a decrease in reactivity (Figure 1F), agreeing with our observation that hESCs are highly structurally accessible (Figure 1B; Figure S4). These reactivity-changing regions are enriched in the 3' UTRs (Figure 1G), indicating that the 3' UTRs either become more double stranded or contain additional cellular factors such as microRNA (miRNA) and RNA binding proteins (RBPs) bound to them during differentiation. Interestingly, we observed the largest number of RNA reactivity changes (12,538 regions) as NPCs differentiate into iNeu, when cells become fixed in their path toward neuronal lineage. Reactivity-changing regions during NPC-to-iNeu differentiation are still enriched in the 3' UTRs and become less accessible during differentiation (Figures 1F and 1G). As the cells differentiate from immature to more mature neurons, we observed fewer reactivity changes in the transcriptome and an increased propensity for RNA regions to become more accessible.

To determine whether the reactivity changes along transcripts have functional consequences, we analyzed changes in translation, expression, and decay of these transcripts during neuronal differentiation. We observed that reactivity changes during hESC-to-NPC differentiation is associated with changes in translation (Figure S7A), agreeing with previous literature that transcripts with an increase in accessibility are more highly translated. We also performed GO term enrichments to determine the functions of the transcripts with reactivity changes. Structurally dynamic RNAs are enriched for GO terms including axon elongation, brain development, neurotransmitter processes, and cell division (Figure 1H), and transcripts with large reactivity changes (two or more windows) are enriched for cell cycle, proliferation, and stem cell maintenance (Figure S7B). These results further indicate that RNA structures may play important roles during neurogenesis.

To identify RNA elements that show continuous reactivity changes during neuronal differentiation, we performed k-means clustering to group the reactivity changes into five clusters (Figure S8A). We then correlated these clusters with gene regulatory processes including gene expression, translation, and decay. Cluster 1 contains RNA regions that show a decrease in icSHAPE reactivities (Figure 2A), suggesting that they become more double-stranded over time. Similar to our above observations, regions in cluster 1 are significantly enriched in 3' UTRs (Figure 2B) and the transcripts they belong to show an enrichment for faster RNA decay and decreased translation during differentiation (Figure 2C; Figure S8B). These transcripts with changes in decay and translation are enriched for GO terms including regulation of progenitor cell differentiation, mRNA destabilization and cell cycle phase transition (for faster decay),

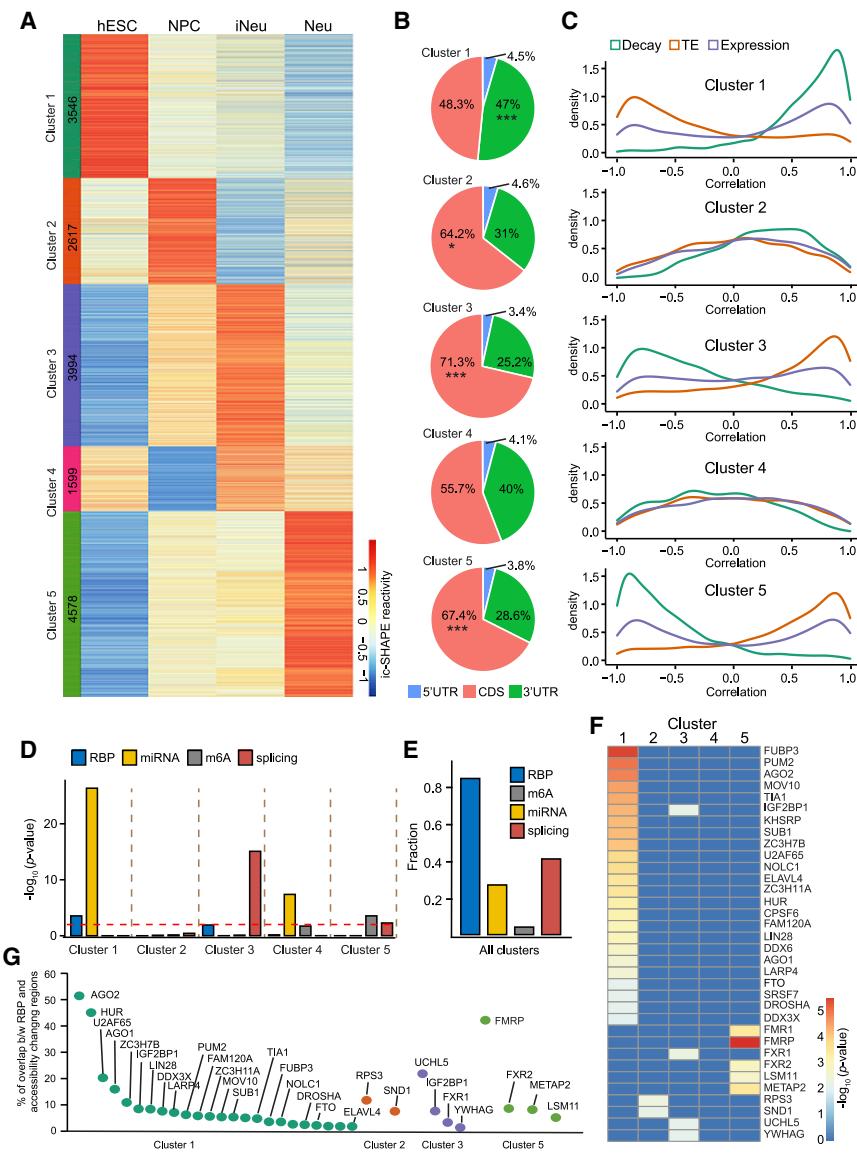


Figure 2. Functional analysis of structure clusters during neuronal differentiation

(A) Heatmap showing k-means clustering of reactivity-changing windows across the four time points. The clustering was performed using Ward's method on the basis of the Euclidean distances on the normalized reactivity per window.

(B) Pie charts showing the proportion of windows localized in 5' UTR (blue), CDS (red), and 3' UTR (green) regions for each cluster. The enrichment test was performed using hypergeometric test. (*p < 0.05 and ***p < 0.001).

(C) Density plots of Pearson's correlation coefficient between normalized reactivity per windows and (1) RNA decay (green), (2) translation efficiency (orange), and (3) gene expression (purple) in each cluster.

(D) The enrichment of regulator binding sites including RBP binding sites (blue), miRNA binding sites (yellow), m6A sites (gray), and splicing sites (red) in each cluster. p value is calculated using binomial test. The expected proportion is calculated as the overlapping proportion of different regulators to all windows in five clusters.

(E) The proportion of reactivity-changing windows that overlap with RBP binding sites (blue), miRNA binding sites (yellow), m6A sites (gray), and splicing sites (red) in each cluster. Some of the reactivity-changing windows overlap with two or more factors.

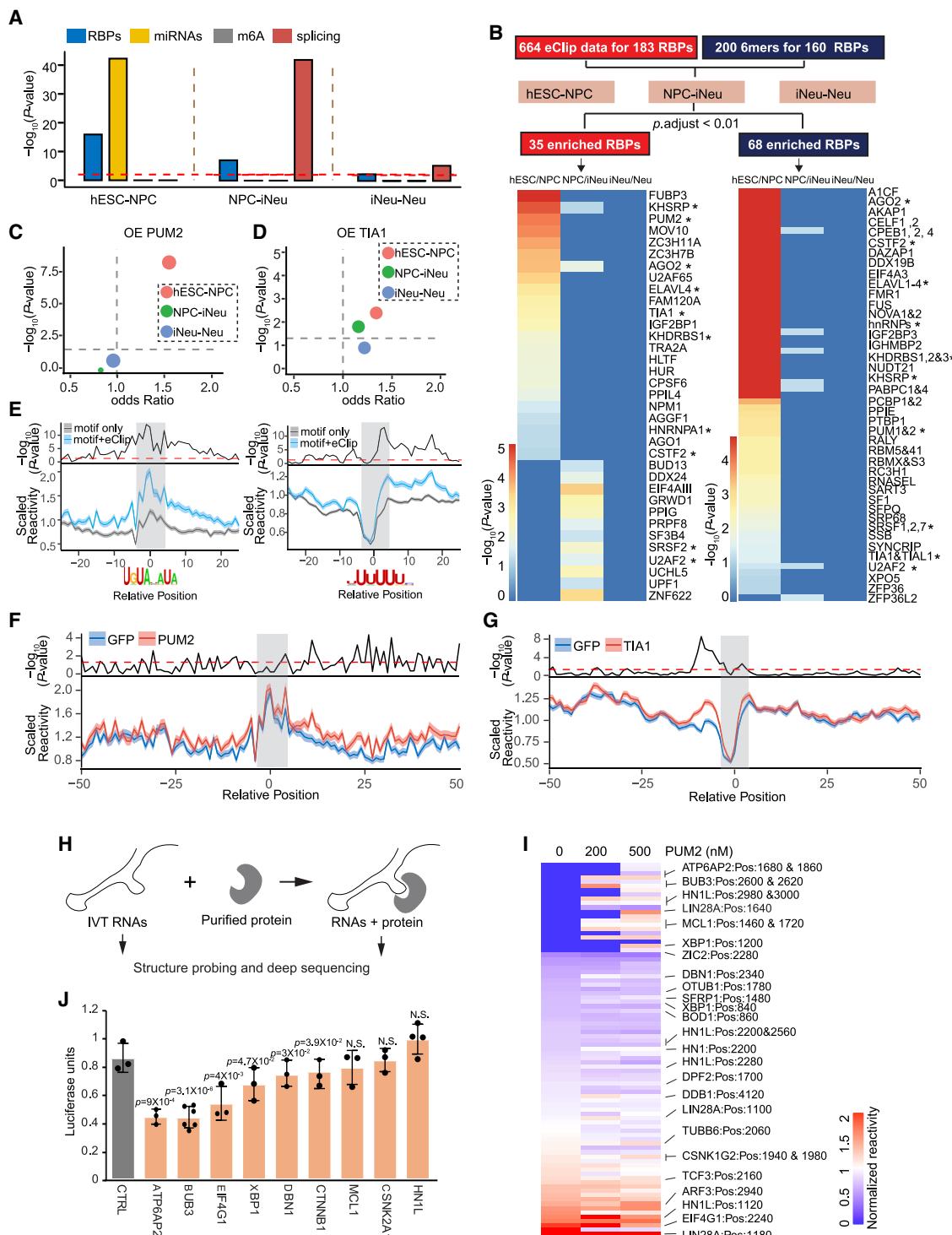
(F) Heatmap showing the $-\log_{10}$ of the enrichment p values of RBP in each cluster. p value was computed using hypergeometric test.

(G) The fraction of structure-changing windows in each cluster that overlaps with the different enriched RBPs in that cluster.

and cell-cell adhesion (for decreased translation; Figure S8C). Clusters 2 and 4 show a biphasic change in reactivity whereby the RNA regions either become more accessible and then less accessible (cluster 2) or less accessible and then more accessible during differentiation (cluster 4; Figure 2A). These two clusters are not enriched for translation or decay. Clusters 3 and 5 contain regions that show an increase in reactivities, indicating more single-strandedness over time (Figure 2A). These regions are enriched in the coding region (Figure 2B) and show a positive correlation with translation and a negative correlation with decay (Figure 2C), and the transcripts they belong to are enriched for GO terms associated with neurogenesis-related pathways (Figure S8C). We did not see an enrichment with gene expression for any of the clusters (Figure 2C; Figure S8D), agreeing with our above observation that RNA structure dynamics are gene expression independent (Figure 1E; Figure S4C).

Many cellular factors, including miRNAs and RBPs, regulate and are also in turn regulated by RNA structures (Kedde et al., 2010; Lewis et al., 2017; Liu et al., 2015; Spitale et al., 2015; Van Nostrand et al., 2020). To study the underlying mechanisms behind icSHAPE reactivity changes, we tested for the enrichment

of miRNA, m6A modification, splicing, and RBP binding sites near the reactivity-changing windows in each cluster, compared with all reactivity-changing windows. Enrichments in the different clusters are independent of the window size used to overlap with the factors (Figure S9A). Cluster 1 is strongly enriched for miRNA and RBP sites (Figure 2D; Figure S9A), agreeing with our observation that it is associated with increased RNA decay (Figures 2A–2C). Several of the enriched miRNAs, including miR-30 and miR-302, are important regulators of stem cell differentiation and stem cell pluripotency (Li et al., 2017; Rosa and Brivanlou, 2011) (Figure S9B). We also observed that cluster 3 is enriched for splice sites, indicating that splicing could be positively associated with increased translation during differentiation. Additionally, cluster 5, which shows a strong increase in reactivity over time (Figure 2A), is enriched for m6A modifications (Figure 2D; Figure S9A), again suggesting that

**Figure 3. Reactivity-changing regions are enriched for cellular factors**

(A) The enrichment of binding sites of cellular factors in reactivity-changing windows during hESC-NPC, NPC-iNeu, and iNeu-Neu differentiation. The number of reactivity-changing regions that fall ± 50 bases of the binding sites of RBPs (blue), miRNAs (yellow), m6A (gray), and splicing sites (red) are compared with all (changing and non-changing) windows that fall within ± 50 bases of these factors. p value is calculated using binomial test.

(B) Heatmaps showing RBP enrichment in reactivity-changing regions using eCLIP binding sites (left) and motif sequences (right). The ten shared RBPs that are predicted by both methods are labeled with asterisks.

(legend continued on next page)

m6A modifications can regulate translation during differentiation (Mao et al., 2019; Yu et al., 2018).

To determine the relative contributions of the different regulators in altering RNA structures, we calculated the proportion of reactivity changes that are attributed to each regulator. In the order of prevalence, 82%, 40%, 27%, and 4% of the reactivity-changing regions are associated with RBP, splicing, miRNA, and m6A modifications, respectively (Figure 2E), indicating that RBPs are a major contributor to reactivity changes. To identify the RBPs that are enriched in each cluster, we mined publicly available eCLIP datasets (STAR Methods) (Van Nostrand et al., 2016) and observed many enriched RBPs in different clusters (Figure 2F). Some of these RBPs could contribute to more than 10% of structural changes in a cluster (Figure 2G), suggesting that they play major roles in the remodeling of RNA structures. We found enrichment for RBPs that are functionally important for neuronal development, including PUM2, TIA1, IGF2BP1, ELAVL4, FTO, FMRP, FXR1, FXR2, Mov10, and FAM120A (Figure 2F). PUM2, TIA1, and ELAVL4, which were known to preferentially bind to 3' UTRs to regulate RNA stability, are enriched in cluster 1 (Ince-Dunn et al., 2012; Meyer et al., 2018; Wang et al., 2002), consistent with our observations that cluster 1 is positively correlated with RNA decay (Figure 2C). In addition, FMRP, which preferentially binds to the coding regions of RNA to regulate translation and decay in prior studies (Darnell et al., 2011; Greenblatt and Spradling, 2018), is enriched in our cluster 5 (Figures 2F and 2G), which is associated with the coding region and with transcripts that show increased TE and decreased decay (Figure 2C). Some of these RBPs, including FMRP, Mov10, and FAM120A, have been reported to be important in neurological diseases such as fragile X syndrome, suggesting that dysregulation of these RBP could play roles in those diseases through structure.

We also observed that cellular factors are enriched in different stages of the differentiation process, such that RBPs and miRNAs are enriched in reactivity-changing regions during hESC-to-NPC differentiation, while splice sites are enriched as NPCs differentiate to neurons (Figure 3A). To comprehensively identify the RBPs that are associated with reactivity-changing regions, we performed a more detailed search using sequence motif enrichments and binding site enrichments using eCLIP (Van Nostrand et al., 2016) datasets and binding sites predicted from the program PrismNet (Sun et al., 2021). We found a total of 35 enriched RBPs using eCLIP datasets, 68 RBPs using motif searches, and 51 using PrismNet across different time points (Figure 3B; Figure S10A; STAR Methods). Overlap between

eCLIP and motif sequences identified a total of ten shared RBPs, while four RBPs (PUM2, TIA1, CSTF2, and U2AF2) are consistently enriched across all three methods (Figure S10B). Nine out of ten RBPs enriched using eCLIP and motif sequences showed an increase in either gene expression or TE during hESC-to-NPC differentiation (Figure S10C), suggesting that an increase in their protein product could regulate the RNA structures of their targets.

To understand the relationship between RBP and RNA structures, we focused our studies on two neurologically important RBPs, PUM2 and TIA1, which are enriched during hESC-to-NPC differentiation. Western blot experiments showed that the protein levels of both PUM2 and TIA1 are elevated upon differentiation (Figure S10D). To determine the impact of these two RBPs on transcript structures, we individually overexpressed them in hESCs using tetracycline-controlled transcription activation (Tet-on system; STAR Methods; Figures S11A–S11C) and performed structure probing in RBP overexpressed and control hESC. Overexpression of PUM2 or TIA1 in hESC resulted in reactivity changes that significantly overlapped with naturally occurring reactivity changes during hESC-to-NPC differentiation (Figures 3C and 3D), indicating that PUM2 and TIA1 could regulate RNA reactivities during differentiation.

Studying the structural contexts of RBP targets enables us to understand the substrate specificity of RBPs (Dominguez et al., 2018). In addition, RBP binding can result in structure changes that have important functional consequences for their targets. Although PUM2 is known to bind to single-stranded regions (Jarmoskaite et al., 2019), TIA1 is not known to have a clear structural context for its substrate. By comparing “real” binding sites (motifs with evidence of eCLIP binding) with “artificial” binding sites (motifs without evidence of eCLIP binding), we observed that real PUM2 motifs are more accessible around the binding motifs compared with artificial motifs (Figure 3E). This agrees with previous literature that PUM2 prefers binding to more single-stranded regions (Jarmoskaite et al., 2019). Interestingly, real TIA1 binding sites are also more accessible than artificial sites, and we also observed that TIA1 binding motifs contain lower reactivity regions compared with its neighboring sequences (Figure 3E). This suggests that an accessible structural context of TIA1 motifs could facilitate its binding to its targets.

To determine the effect of protein binding on RNA structure, we performed metagene analysis of RNA reactivities at “real” PUM2 or TIA1 binding sites before and after their overexpression. Metagene analysis showed that PUM2 binding resulted in an increase in accessibility downstream of PUM2 motifs

(C and D) Enrichment of significant reactivity-changing windows during hESC-NPC, NPC-iNeu, and iNeu-Neu with reactivity-changing windows upon PUM2 (C) or TIA1 (D) overexpression.

(E) Bottom: Metagene analysis of normalized icSHAPE reactivity of real (with eCLIP and motif) and artificial (with motif only) PUM2 (left panel) and TIA1 (right panel) binding sites. Top: p values above each subplot were calculated for every nucleotide using single-sided t tests. The gray-shaded region indicates the location of the binding motif of PUM2 (Hafner et al., 2010) and TIA1 (Ray et al., 2013). Motif sequences are indicated below each subplot.

(F and G) Metagene analysis of normalized icSHAPE reactivity around real PUM2 (F) or TIA1 (G) binding sites before and after PUM2 (F) or TIA1 (G) overexpression. Real binding sites are sites with both motif sequences and eCLIP binding evidence. p values above plot were calculated per nucleotide using single-sided t tests. The gray-shaded box indicates the location of the motif.

(H) Schematic of our *in vitro* RNA-protein interaction assay.

(I) Heatmap showing the normalized SHAPE-MaP reactivity of 87 RNA fragments incubated with different concentrations of PUM2 (0, 200, and 500 nM).

(J) Ratio of relative luciferase units of nine randomly selected transcript regions cloned behind luciferase, upon PUM2 or GFP overexpression in HEK293T cells. p values were calculated using the one-tailed t test (replicate number n = 3–6).

(Figure 3F), with 77% of PUM2 binding sites becoming more accessible upon PUM2 overexpression and 70% of sites becoming less accessible upon PUM2 knockdown (Figures S11C–S11E). Overexpression of TIA1, on the other hand, resulted in a sharp, local increase in accessibility of 10 bases upstream of the “real” TIA1 motif (Figure 3G), with 65% of its targets becoming more accessible with increased TIA1 levels and 79% of its targets becoming less accessible with decreased TIA1 levels (Figures S11A, S11B, S11D, and S11F). We confirmed that overexpression of PUM2 and TIA1 resulted in reactivity changes at their respective binding sites using predicted binding sites from the program PrismNet (Figure S11G).

To test the proportion of *in vivo* reactivity changes that are caused by direct RBP binding, we cloned 87 RNA regions from genes that showed different extents of structural changes upon PUM2 overexpression (Figure S12). We *in vitro* transcribed and refolded the RNAs as a pool and performed RNA-protein interaction experiments by incubating the RNAs with different concentrations of purified PUM2 protein *in vitro* (Figure 3H; STAR Methods). Fifty-three percent of our selected RNA regions showed increases in reactivity upon PUM2 protein incubation *in vitro*, confirming our *in vivo* data that PUM2 protein binding indeed results in greater structural accessibility (Figures 3F and 3I). For nine of the regions that showed significant reactivity changes upon PUM2 binding *in vitro*, we additionally tested the effect of PUM2 binding on their gene expression inside cells. Using luciferase reporter assays, we observed gene repression in six out of nine RNA regions, indicating that PUM2 binding indeed alters the gene regulation of these RNA regions (Figure 3J).

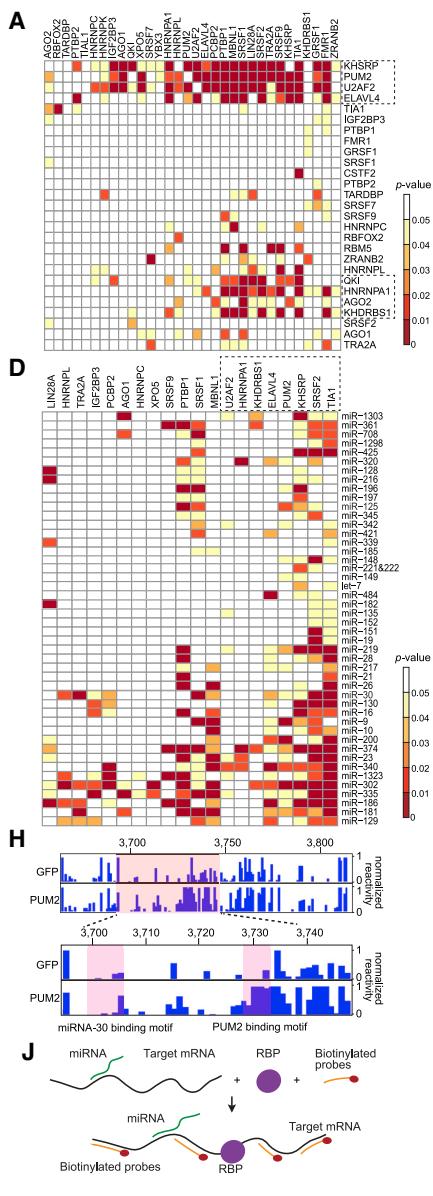
Although multiple RBPs can bind to an RNA to coordinate gene regulation, the extent to which their coordination is structure dependent is largely unknown. To determine whether RBP-induced reactivity changes could alter the accessibility of nearby regulators, we calculated the enrichment of RBP pairs near reactivity-changing windows versus non-reactivity-changing windows in our dataset (Figure 4A; STAR Methods; Table S4). Enriched RBP pairs share similar transcript targets (Figure S13A), and their structure-changing regions are enriched in the 3' UTR of their targets (Figure 4B). We observed that enriched RBP pairs include PTBP1 and QKI (Figure 4A), which are master regulators of neuronal cell fates (Hardy et al., 1996; Shu et al., 2017; Xue et al., 2013), agreeing with their importance in our data.

Although some RBPs are enriched for other RBPs, we also identified RBPs that are promiscuously enriched for many other RBPs (Figure 4A). Enriched RBPs, in particular promiscuously binding RBPs, are associated with functions such as regulation of RNA decay (Figure 4C). To determine whether target transcripts with RBP pairs located in reactivity-changing regions are regulated differently from transcripts with RBP pairs found in non-reactivity-changing regions, we determined changes in expression, decay, and translation of these targets. We observed changes in target gene regulation when RBP pairs are associated with structure changes, highlighting that structure-based co-regulation contributes to an important layer of complexity in neuronal development (Figure S13B).

Besides RBPs, miRNAs also play important roles in gene regulation during neurogenesis. To determine whether RBP-induced reactivity changes could alter nearby miRNA binding as hESCs differentiate into NPCs, we analyzed the enrichment of top expressed miRNAs using sites predicted from TargetScan (STAR Methods; Table S4). We observed 71 miRNAs that are enriched near RBP-associated reactivity-changing regions (Figure 4D). These include miR-302, miR-30, miR-200, and miR-130, which are important miRNAs in stem cell maintenance, differentiation and iPS reprogramming (Li et al., 2017; Subramanyam et al., 2011; Wang et al., 2013). Enriched RBP-miRNA pairs have differential changing windows that are enriched in the 3' UTRs (Figure 4E) and are associated with functions including regulation of translation and decay (Figure 4F). These data suggest that RBP-induced structure changes could be a common mechanism for RBPs to regulate the binding of a second modulator to modulate gene expression.

To demonstrate the functional consequences of reactivity changes during neuronal differentiation, we focused our studies on RNAs that are functionally important and show large reactivity changes during the process. LIN28A, which is a key gene in stem cell maintenance, demonstrated the greatest number of structure-changed regions upon PUM2 overexpression (Figure S12) and showed increased decay during differentiation (Shyh-Chang and Daley, 2013) (Figures S14A and S14B), consistent with previous reports of PUM2 induces decay of its targets (Goldstrohm et al., 2018; Vessey et al., 2006; Wickens et al., 2002). Although LIN28A has two isoforms, we confirmed that the same isoform is dominant in hESCs and NPCs and validated that PUM2 indeed binds to LIN28A in hESCs and NPCs using eCLIP (Figure S14C). PUM2 overexpression results in reactivity changes in more than 30 regions on LIN28A (Figure S14D), including an increase in icSHAPE reactivities around the PUM2 binding sites and two miRNA binding sites, miR-30 and miR-125, in the 3' UTR (Figures 4G and 4H). To confirm that these reactivity changes are indeed due to PUM2 binding to LIN28A, we incubated LIN28A with PUM2 *in vitro*. We observed that the presence of PUM2 increases RNA accessibility at PUM2, miR-30, and miR-125 binding sites (Figures S14E and S14F).

To determine the effect of PUM2 binding on LIN28A, we measured LIN28A levels in hESC and NPC upon PUM2 knockdown. PUM2 knockdown resulted in an increase in LIN28A expression (Figures S15A and S15B), suggesting that PUM2 downregulates LIN28A. Additionally, we cloned each of the two LIN28A reactivity-changing regions that are close to the PUM2 binding sites (Figure 4G) behind a luciferase reporter gene. Luciferase assays showed that overexpression of PUM2 decreased gene expression (Figure 4I), while downregulation of PUM2 increased gene expression of the two LIN28A regions, again indicating that LIN28A is downregulated in the presence of PUM2. As the PUM2-induced reactivity changes overlap with miR-30 and miR-125 binding sites, we hypothesized that reactivity changes during PUM2 overexpression or neuronal differentiation could affect the accessibility of these miRNAs to bind to LIN28A. Indeed, pull-down of LIN28A using a pool of biotinylated antisense oligos upon PUM2 overexpression or during neuronal differentiated significantly increased the amount of



(H) icSHAPE reactivity profile of PUM2 binding site on LIN28A (F-3729, in orange background) upon GFP overexpression (top panel) and PUM2 overexpression (bottom panel). The detailed reactivity profiles in the region from 3,700 to 3,740 are expanded below.

(I) Ratio of relative luciferase units of two LIN28A regions cloned behind luciferase, upon PUM2 versus GFP overexpression (left) and upon PUM2 knockdown versus control (right) in HEK293T cells, empty vector was used as control. p values were calculated using the one-tailed t test (replicate number n = 3).

(J) Schematic showing the experiment design of LIN28A pull-down assay.

(K) Ratio of miR-30 (left) or miR-125 (right) amount pulled down with LIN28A under PUM2 and GFP overexpression. p values were calculated using the one-tailed t test (replicate number n = 4 [left] and n = 3 [right]).

miR-30 and miR-125 that is bound to LIN28A (Figures 4J and 4K; Figures S15C and S15D).

To visualize the structure changes induced by PUM2 binding and their impact on the accessibility of miRNA binding on LIN28A, we focused on miR-30, as it was enriched in our data and is an important miRNA during stem cell differentiation. We incorporated icSHAPE reactivity as constraints in the Vienna RNAfold package to model LIN28A RNA secondary

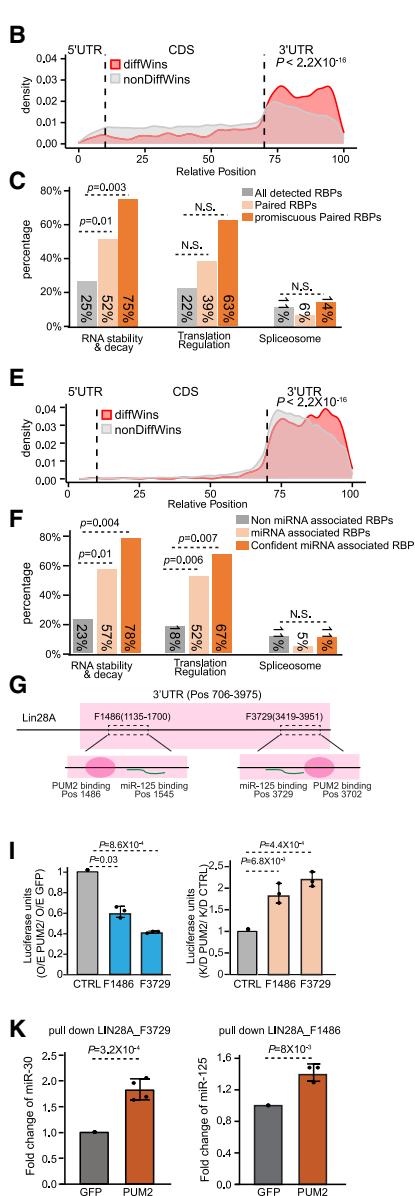


Figure 4. RBPs regulate the binding of other cellular factors through structure

(A) Heatmap showing the enrichment of expressed secondary RBPs (shown as row names) in reactivity-changing versus non-changing regions near primary RBP binding sites (shown as column names). p values were calculated using hypergeometric test for RBPs that are enriched in reactivity-changing regions \pm 100 bases around our enriched real RBP binding sites. The dotted boxes indicate the promiscuous RBPs (Table S4).

(B) Density plot of the locations of structure-changing windows (red) and -non-structure-changing windows (gray) in significantly enriched RBP-RBP pairs along the transcriptome. The representative transcript is separated into 5' UTR, CDS, and 3' UTR with a scaling ratio of 10:60:30. p value was calculated using bootstrap Kolmogorov-Smirnov test.

(C) Bar plots showing the function annotation of detected RBPs, paired RBPs, and promiscuously paired RBPs. The functional annotation of individual RBPs is shown in Table S4.

(D) Heatmap showing the enrichment of top 100 expressed miRNAs in HESCs. Forty-seven miRNAs with enrichment to at least two RBPs are shown in heatmap. p values were calculated using hypergeometric test for miRNAs enriched in reactivity-changing regions \pm 100 bases around enriched real RBP binding sites. The dotted box indicates the confident RBPs that are enriched from both eCLIP and motif enrichment analysis from Figure 3B.

(E) Density plot of the locations of structure-changing windows (red) and -non-structure-changing windows (gray) in significantly enriched RBP-miRNA pairs along the transcriptome. The representative transcript is separated into 5' UTR, CDS, and 3' UTR with a scaling ratio of 10:60:30. p value was calculated using bootstrap Kolmogorov-Smirnov test.

(F) Bar plots showing the function annotation of non-miRNA-associated RBPs, miRNA-associated RBPs, and confident miRNA-associated RBPs (Table S4).

(G) Schematic showing two -PUM2 binding sites (named F-1486 and F-3729, in pink circles) and two miRNA binding sites (miR-125 and miR-30, in green curves) on gene LIN28A.

(H) icSHAPE reactivity profile of PUM2 binding site on LIN28A (F-3729, in orange background) upon GFP overexpression (top panel) and PUM2 overexpression (bottom panel). The detailed reactivity profiles in the region from 3,700 to 3,740 are expanded below.

(I) Ratio of relative luciferase units of two LIN28A regions cloned behind luciferase, upon PUM2 versus GFP overexpression (left) and upon PUM2 knockdown versus control (right) in HEK293T cells, empty vector was used as control. p values were calculated using the one-tailed t test (replicate number n = 3).

(J) Schematic showing the experiment design of LIN28A pull-down assay.

(K) Ratio of miR-30 (left) or miR-125 (right) amount pulled down with LIN28A under PUM2 and GFP overexpression. p values were calculated using the one-tailed t test (replicate number n = 4 [left] and n = 3 [right]).

structures before and after PUM2 binding (Lorenz et al., 2011) (STAR Methods). Interestingly, the structure models show extensive rearrangement, resulting in miR-30 binding site becoming more accessible, upon PUM2 binding (Figure 5A). To demonstrate that it is indeed PUM2-induced structure changes that affect miR-30 binding, we performed mutations on the LIN28A secondary structures to either destabilize the pre-PUM2 bound structure or to stabilize the post-PUM2

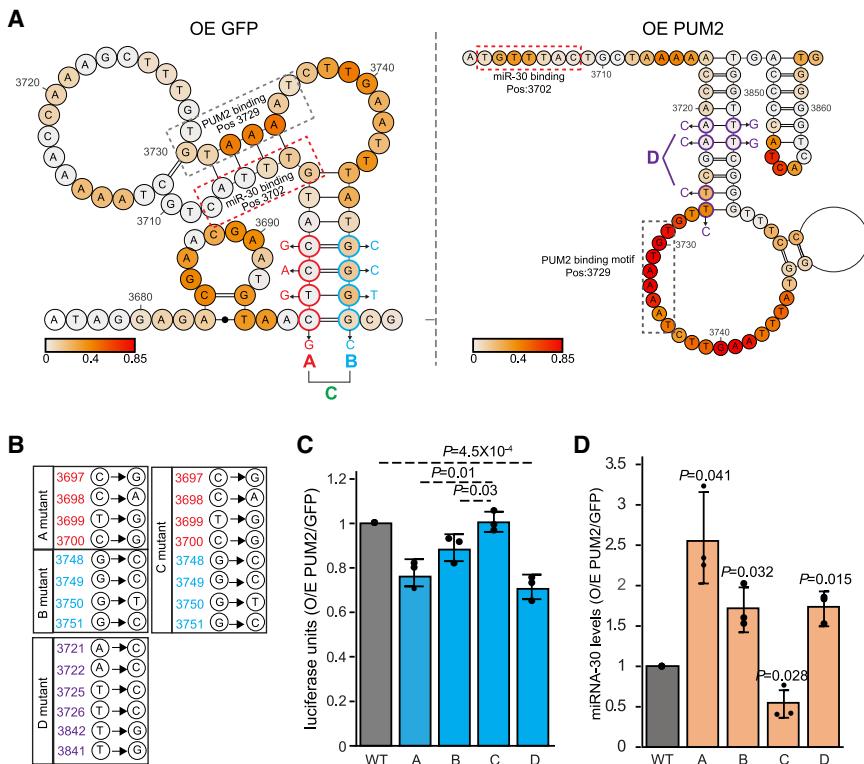


Figure 5. PUM2 and miR-30 co-regulate LIN28A through structure

(A) Structural models of the pre-PUM2 binding LIN28A structure (left) and post-PUM2 binding LIN28A structure (right) by using RNA reactivities as constraints. PUM2 (gray dotted box) and miR-30 (red dotted box) binding sites are marked on the secondary structure. The mutational target nucleotides are labeled in bold circles, and the designed mutations are shown side by side. A, B, C, and D mutants are labeled in red, cyan, green, and purple color, respectively.

(B) The position and sequences of designed mutants as shown in (A).

(C) Relative luciferase units of LIN28A WT and mutants, cloned behind luciferase genes, upon overexpression of PUM2 versus GFP. p values were calculated using the one-tailed t test (replicate number n = 3).

(D) Ratio of miR-30 pulled down with WT LIN28A or its mutants upon PUM2 versus GFP overexpression. miR-30 levels were determined using qRT-PCR. p values were calculated using the one-tailed t test (replicate number n = 3).

bound structure (Figure 5B). We confirmed that the mutations do disrupt RNA structures, as expected from footprinting experiments (Figure S16). Indeed, mutations that disrupted either side of the stem in the pre-PUM2 structure resulted in a further decrease in luciferase activity (Figure 5C, mutations A and B) and an increase in miR-30 binding (Figure 5D), indicating a shift toward post-PUM2 structure. Importantly, compensatory mutations that restored the pre-PUM2 binding structure rescued this decrease in luciferase activity and the increase in miR-30 binding (Figures 5C and 5D, mutation C). Additionally, mutations that stabilized the post-PUM2 structure by increasing the number of GC base pairs along the stem also decreased luciferase activity and increased miR-30 binding (Figures 5C and 5D, mutation D). This is consistent with our conclusion that PUM2 miR-30 co-regulation on LIN28A is mediated through structure.

DISCUSSION

Neuronal development is an extremely complex process that is extensively regulated. By using high-throughput structure probing strategies, we probed RNA structures and their dynamics during neuronal differentiation, providing an understanding of an additional layer of regulation to gene expression. One of the challenges of structure probing inside cells is that although highly reactive regions indicate the presence of single-stranded bases, low-reactivity regions could either be a result of double-stranded bases or binding due to other cellular factors, such as RBP and miRNA. Additionally, confidently identifying reac-

tivity-changing regions between two biological states can be challenging. Here, we compared different analytic strategies to identify reactivity differences by using footprinting data as the ground truth.

We showed that NOISeq performs better than other methods, and believe that NOISeq can be widely used to identify reactivity changes in other cellular conditions.

Globally, we observed that RNA structures are more accessible in hESCs than in differentiated cells, reminiscent of open chromatin and pervasive transcription in hESCs (Efroni et al., 2008). This suggests that RNA structures could also be “poised” for transcript regulation in hESCs depending on the direction of cellular lineage in which they are heading. Recent high-throughput studies show that ribosomes are a major factor for structure remodeling inside cells. Here, we show that dynamic RNA structures during differentiation are associated with miRNA, splicing, m6A modification, and RBP binding, with m6A modification being enriched in CDS and associated with increased translation. This agrees with existing literature that m6A modification increases structure accessibility (Liu et al., 2015) and that modifications on CDS enhances translation.

Although *in silico* modeling has suggested the presence of extensive RBP co-regulation through secondary structures (Lin and Bundschuh, 2013), only individual examples of this have been shown experimentally in a couple of cellular states, including in quiescent (Kedde et al., 2010) and hypoxic (Ray et al., 2009) conditions. Whether RBPs could coordinate gene regulation through RNA structure remodeling in a widespread manner in human development is still largely unknown. Our studies show that RBPs play pervasive roles in regulating RNA structures during development and modulate co-regulation networks through structure. The complexity of their regulation is demonstrated by our LIN28A-PUM2-miR-30 example, showing

that PUM2 can use structure dynamics to control the miR-30 interaction with LIN28A. This study deepens our understanding of the connectivity and complexity of structure-mediated gene regulation during neuronal development.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- **KEY RESOURCES TABLE**
- **RESOURCE AVAILABILITY**
 - Lead contact
 - Materials availability
 - Data and code availability
- **EXPERIMENTAL MODEL AND SUBJECT DETAILS**
 - Cell culture and inducible gene expression in hESCs
- **METHOD DETAILS**
 - RNA structure probing, icSHAPE library preparation and data analysis
 - Riboswitch & RiboSNich benchmarks
 - SHAPE-MaP library preparation and data analysis
 - DMS-MaP-Seq library preparation and data analysis
 - Enhanced CLIP (eCLIP) library preparation
 - RNA pull down experiment
 - *In vitro* RNA-protein interaction experiment
 - Luciferase assay
 - Identification of structure changing regions in the transcriptome
 - Filtering of windows associated with alternative splicing
 - Ribosome profiling library preparation and analysis
 - RNA sequencing library preparation and analysis
 - RNA decay library preparation and analysis
 - RBP enrichment and PrismNet analysis
 - miRNA enrichment analysis
 - Analysis of m6A binding sites
 - RNA structure modeling
- **QUANTIFICATION AND STATISTICAL ANALYSIS**

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at <https://doi.org/10.1016/j.molcel.2021.09.027>.

ACKNOWLEDGMENTS

We thank members of the Wan lab for helpful discussions. Y.W. is supported by funding from the A*STAR Investigatorship, the EMBO Young Investigatorship, the National Research Foundation, and the CIFAR Azrieli global scholar fellowship. FRPL is supported by a doctoral scholarship from the Warwick-A*STAR research attachment programme.

AUTHOR CONTRIBUTIONS

Y.W. conceived the project. Y.W. and J.W. designed the experiments. J.W. and W.T.T. performed all the experiments. T.Z., M.W., and Y.S. performed the computational analysis. R.G.H. performed the structure modeling of LIN28A. Y.W. and J.W. organized and wrote the paper with all other authors.

DECLARATION OF INTERESTS

The authors declare no competing interests.

Received: March 18, 2021

Revised: August 1, 2021

Accepted: September 24, 2021

Published: October 15, 2021

REFERENCES

- Agarwal, V., Bell, G.W., Nam, J.-W., and Bartel, D.P. (2015). Predicting effective microRNA target sites in mammalian mRNAs. *eLife* 4, e05005.
- Batista, P.J., Molinie, B., Wang, J., Qu, K., Zhang, J., Li, L., Bouley, D.M., Lujan, E., Haddad, B., Daneshvar, K., et al. (2014). m(6)A RNA modification controls cell fate transition in mammalian embryonic stem cells. *Cell Stem Cell* 15, 707–719.
- Beaudoin, J.D., Novoa, E.M., Vejnar, C.E., Yartseva, V., Takacs, C.M., Kellis, M., and Giraldez, A.J. (2018). Analyses of mRNA structure dynamics identify embryonic gene regulatory programs. *Nat. Struct. Mol. Biol.* 25, 677–686.
- Bernat, V., and Disney, M.D. (2015). RNA structures as mediators of neurological diseases and as drug targets. *Neuron* 87, 28–46.
- Brar, G.A., and Weissman, J.S. (2015). Ribosome profiling reveals the what, when, where and how of protein synthesis. *Nat. Rev. Mol. Cell Biol.* 16, 651–664.
- Bushnell, B. (2014). BBMap: A Fast, Accurate, Splice-Aware Aligner. <https://www.osti.gov/servlets/purl/1241166>.
- Chen, C.Y., Ezzeddine, N., and Shyu, A.B. (2008). Messenger RNA half-life measurements in mammalian cells. *Methods Enzymol.* 448, 335–357.
- Darnell, J.C., Van Driesche, S.J., Zhang, C., Hung, K.Y., Mele, A., Fraser, C.E., Stone, E.F., Chen, C., Fak, J.J., Chi, S.W., et al. (2011). FMRP stalls ribosomal translocation on mRNAs linked to synaptic function and autism. *Cell* 146, 247–261.
- Darty, K., Denise, A., and Ponty, Y. (2009). VARNA: interactive drawing and editing of the RNA secondary structure. *Bioinformatics* 25, 1974–1975.
- Das, R., Laederach, A., Pearlman, S.M., Herschlag, D., and Altman, R.B. (2005). SAFA: semi-automated footprinting analysis software for high-throughput quantification of nucleic acid footprinting experiments. *RNA* 11, 344–354.
- Deigan, K.E., Li, T.W., Mathews, D.H., and Weeks, K.M. (2009). Accurate SHAPE-directed RNA structure determination. *Proc. Natl. Acad. Sci. U S A* 106, 97–102.
- Dethoff, E.A., Chugh, J., Mustoe, A.M., and Al-Hashimi, H.M. (2012). Functional complexity and regulation through RNA dynamics. *Nature* 482, 322–330.
- Dobin, A., Davis, C.A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M., and Gingeras, T.R. (2013). STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* 29, 15–21.
- Dominguez, D., Freese, P., Alexis, M.S., Su, A., Hochman, M., Palden, T., Bazile, C., Lambert, N.J., Van Nostrand, E.L., Pratt, G.A., et al. (2018). Sequence, structure, and context preferences of human RNA binding proteins. *Mol. Cell* 70, 854–867.e9.
- Efroni, S., Duttagupta, R., Cheng, J., Dehghani, H., Hoeppner, D.J., Dash, C., Bazett-Jones, D.P., Le Grice, S., McKay, R.D., Buetow, K.H., et al. (2008). Global transcription in pluripotent embryonic stem cells. *Cell Stem Cell* 2, 437–447.
- Flynn, R.A., Zhang, Q.C., Spitale, R.C., Lee, B., Mumbach, M.R., and Chang, H.Y. (2016). Transcriptome-wide interrogation of RNA secondary structure in living cells with icSHAPE. *Nat. Protoc.* 11, 273–290.
- Ganser, L.R., Kelly, M.L., Herschlag, D., and Al-Hashimi, H.M. (2019). The roles of structural dynamics in the cellular functions of RNAs. *Nat. Rev. Mol. Cell Biol.* 20, 474–489.

- Garneau, N.L., Wilusz, J., and Wilusz, C.J. (2007). The highways and byways of mRNA decay. *Nat. Rev. Mol. Cell Biol.* 8, 113–126.
- Giudice, G., Sánchez-Cabo, F., Torroja, C., and Lara-Pezzi, E. (2016). ATtRACT—a database of RNA-binding proteins and associated motifs. *Database (Oxford)* 2016, baw035.
- Goldstrohm, A.C., Hall, T.M.T., and McKenney, K.M. (2018). Post-transcriptional regulatory functions of mammalian Pumilio proteins. *Trends Genet.* 34, 972–990.
- Greenblatt, E.J., and Spradling, A.C. (2018). Fragile X mental retardation 1 gene enhances the translation of large autism-related proteins. *Science* 361, 709–712.
- Hafner, M., Landthaler, M., Burger, L., Khorshid, M., Hausser, J., Berninger, P., Rothballer, A., Ascano, M., Jr., Jungkamp, A.C., Munschauer, M., et al. (2010). Transcriptome-wide identification of RNA-binding protein and microRNA target sites by PAR-CLIP. *Cell* 141, 129–141.
- Hardy, R.J., Loushin, C.L., Friedrich, V.L., Jr., Chen, Q., Ebersole, T.A., Lazzarini, R.A., and Artzt, K. (1996). Neural cell type-specific expression of QKI proteins is altered in quakingviable mutant mice. *J. Neurosci.* 16, 7941–7949.
- Hunt, S.E., McLaren, W., Gil, L., Thormann, A., Schuilenburg, H., Sheppard, D., Parton, A., Armean, I.M., Trevanion, S.J., Flücke, P., and Cunningham, F. (2018). Ensembl variation resources. *Database (Oxford)* 2018, bay119.
- Ince-Dunn, G., Okano, H.J., Jensen, K.B., Park, W.Y., Zhong, R., Ule, J., Mele, A., Fak, J.J., Yang, C., Zhang, C., et al. (2012). Neuronal Elav-like (Hu) proteins regulate RNA splicing and abundance to control glutamate levels and neuronal excitability. *Neuron* 75, 1067–1080.
- Jarmoskaite, I., Denny, S.K., Vaidyanathan, P.P., Becker, W.R., Andreasson, J.O.L., Layton, C.J., Kappel, K., Shivasankar, V., Sreenivasan, R., Das, R., et al. (2019). A quantitative and predictive model for RNA binding by human Pumilio proteins. *Mol. Cell* 74, 966–981.e18.
- Jung, H., Yoon, B.C., and Holt, C.E. (2012). Axonal mRNA localization and local protein synthesis in nervous system assembly, maintenance and repair. *Nat. Rev. Neurosci.* 13, 308–324.
- Kedde, M., van Kouwenhove, M., Zwart, W., Oude Vrielink, J.A., Elk, R., and Agami, R. (2010). A Pumilio-induced RNA structure switch in p27-3' UTR controls miR-221 and miR-222 accessibility. *Nat. Cell Biol.* 12, 1014–1020.
- Kertesz, M., Wan, Y., Mazor, E., Rinn, J.L., Nutter, R.C., Chang, H.Y., and Segal, E. (2010). Genome-wide measurement of RNA secondary structure in yeast. *Nature* 467, 103–107.
- Kolb, S.J., and Kissel, J.T. (2011). Spinal muscular atrophy: a timely review. *Arch. Neurol.* 68, 979–984.
- Kozak, M. (2005). Regulation of translation via mRNA structure in prokaryotes and eukaryotes. *Gene* 361, 13–37.
- Langmead, B., and Salzberg, S.L. (2012). Fast gapped-read alignment with Bowtie 2. *Nat. Methods* 9, 357–359.
- Lewis, C.J., Pan, T., and Kalsotra, A. (2017). RNA modifications and structures cooperate to guide RNA-protein interactions. *Nat. Rev. Mol. Cell Biol.* 18, 202–210.
- Li, N., Long, B., Han, W., Yuan, S., and Wang, K. (2017). microRNAs: important regulators of stem cells. *Stem Cell Res. Ther.* 8, 110.
- Liao, Y., Smyth, G.K., and Shi, W. (2014). featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics* 30, 923–930.
- Lin, Y.H., and Bundschuh, R. (2013). Interplay between single-stranded binding proteins on RNA secondary structure. *Phys. Rev. E Stat. Nonlin. Soft Matter Phys.* 88, 052707.
- Liu, N., Dai, Q., Zheng, G., He, C., Parisien, M., and Pan, T. (2015). N(6)-methyladenosine-dependent RNA structural switches regulate RNA-protein interactions. *Nature* 518, 560–564.
- Lorenz, R., Bernhart, S.H., Höner Zu Siederdissen, C., Tafer, H., Flamm, C., Stadler, P.F., and Hofacker, I.L. (2011). ViennaRNA Package 2.0. *Algorithms Mol. Biol.* 6, 26.
- Low, J.T., and Weeks, K.M. (2010). SHAPE-directed RNA secondary structure prediction. *Methods* 52, 150–158.
- Mao, Y., Liu, H., Liu, Y., and Tao, S. (2014). Deciphering the rules by which dynamics of mRNA secondary structure affect translation efficiency in *Saccharomyces cerevisiae*. *Nucleic Acids Res.* 42, 4813–4822.
- Mao, Y., Dong, L., Liu, X.M., Guo, J., Ma, H., Shen, B., and Qian, S.B. (2019). m⁶A in mRNA coding regions promotes translation via the RNA helicase-containing YTHDC2. *Nat. Commun.* 10, 5332.
- Marangio, P., Law, K.Y.T., Sanguinetti, G., and Granneman, S. (2020). Differential BUM-HMM: a robust statistical modelling approach for detecting RNA flexibility changes in high-throughput structure probing data. *bioRxiv*. <https://doi.org/10.1101/2020.07.30.229179>.
- Martin, M. (2011). Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet. J.* 17, 10–12.
- Martin, K.C., and Ephrussi, A. (2009). mRNA localization: gene expression in the spatial dimension. *Cell* 136, 719–730.
- Meyer, C., Garzia, A., Mazzola, M., Gerstberger, S., Molina, H., and Tuschl, T. (2018). The TIA1 RNA-binding protein family regulates EIF2AK2-mediated stress response and cell cycle progression. *Mol. Cell* 69, 622–635.e6.
- Mizrahi, O., Nachshon, A., Shitrit, A., Gelbart, I.A., Dobesova, M., Brenner, S., Kahana, C., and Stern-Ginossar, N. (2018). Virus-induced changes in mRNA secondary structure uncover cis-regulatory elements that directly control gene expression. *Mol. Cell* 72, 862–874.e5.
- Nowicka, M., and Robinson, M.D. (2016). DRIMSeq: a Dirichlet-multinomial framework for multivariate count outcomes in genomics. *F1000Res.* 5, 1356.
- Patro, R., Duggal, G., Love, M.I., Irizarry, R.A., and Kingsford, C. (2017). Salmon provides fast and bias-aware quantification of transcript expression. *Nat. Methods* 14, 417–419.
- Ray, P.S., Jia, J., Yao, P., Majumder, M., Hatzoglou, M., and Fox, P.L. (2009). A stress-responsive RNA switch regulates VEGFA expression. *Nature* 457, 915–919.
- Ray, D., Kazan, H., Cook, K.B., Weirauch, M.T., Najafabadi, H.S., Li, X., Gueroussou, S., Albu, M., Zheng, H., Yang, A., et al. (2013). A compendium of RNA-binding motifs for decoding gene regulation. *Nature* 499, 172–177.
- Rosa, A., and Brivanlou, A.H. (2011). A regulatory circuitry comprised of miR-302 and the transcription factors OCT4 and NR2F2 regulates human embryonic stem cell differentiation. *EMBO J.* 30, 237–248.
- Shi, B., Zhang, J., Heng, J., Gong, J., Zhang, T., Li, P., Sun, B.-F., Yang, Y., Zhang, N., Zhao, Y.-L., et al. (2020). RNA structural dynamics regulate early embryogenesis through controlling transcriptome fate and function. *Genome Biol.* 21, 120.
- Shu, P., Fu, H., Zhao, X., Wu, C., Ruan, X., Zeng, Y., Liu, W., Wang, M., Hou, L., Chen, P., et al. (2017). MicroRNA-214 modulates neural progenitor cell differentiation by targeting Quaking during cerebral cortex development. *Sci. Rep.* 7, 8014.
- Shyh-Chang, N., and Daley, G.Q. (2013). Lin28: primal regulator of growth and metabolism in stem cells. *Cell Stem Cell* 12, 395–406.
- Siegfried, N.A., Busan, S., Rice, G.M., Nelson, J.A., and Weeks, K.M. (2014). RNA motif discovery by SHAPE and mutational profiling (SHAPE-MaP). *Nat. Methods* 11, 959–965.
- Spitale, R.C., Flynn, R.A., Zhang, Q.C., Crisalli, P., Lee, B., Jung, J.-W., Kuchelmeister, H.Y., Batista, P.J., Torre, E.A., Kool, E.T., and Chang, H.Y. (2015). Structural imprints in vivo decode RNA regulatory mechanisms. *Nature* 519, 486–490.
- Stark, R., Grzelak, M., and Hadfield, J. (2019). RNA sequencing: the teenage years. *Nat. Rev. Genet.* 20, 631–656.
- Subramanyam, D., Lamouille, S., Judson, R.L., Liu, J.Y., Bucay, N., Deryck, R., and Blelloch, R. (2011). Multiple targets of miR-302 and miR-372 promote reprogramming of human fibroblasts to induced pluripotent stem cells. *Nat. Biotechnol.* 29, 443–448.
- Sun, A.X., Yuan, Q., Fukuda, M., Yu, W., Yan, H., Lim, G.G.Y., Nai, M.H., D'Agostino, G.A., Tran, H.D., Itahana, Y., et al. (2019a). Potassium channel

- dysfunction in human neuronal models of Angelman syndrome. *Science* 366, 1486–1492.
- Sun, L., Fazal, F.M., Li, P., Broughton, J.P., Lee, B., Tang, L., Huang, W., Kool, E.T., Chang, H.Y., and Zhang, Q.C. (2019b). RNA structure maps across mammalian cellular compartments. *Nat. Struct. Mol. Biol.* 26, 322–330.
- Sun, L., Xu, K., Huang, W., Yang, Y.T., Li, P., Tang, L., Xiong, T., and Zhang, Q.C. (2021). Predicting dynamic cellular protein-RNA interactions by deep learning using *in vivo* RNA structures. *Cell Res.* 31, 495–516.
- Tan, S.M., Kirchner, R., Jin, J., Hofmann, O., McReynolds, L., Hide, W., and Lieberman, J. (2014). Sequencing of captive target transcripts identifies the network of regulated genes and functions of primate-specific miR-522. *Cell Rep.* 8, 1225–1239.
- Tarazona, S., Furió-Tarí, P., Turrà, D., Pietro, A.D., Nueda, M.J., Ferrer, A., and Conesa, A. (2015). Data quality aware analysis of differential expression in RNA-seq with NOISeq R/Bioc package. *Nucleic Acids Res.* 43, e140.
- Trapnell, C., Hendrickson, D.G., Sauvageau, M., Goff, L., Rinn, J.L., and Pachter, L. (2013). Differential analysis of gene regulation at transcript resolution with RNA-seq. *Nat. Biotechnol.* 31, 46–53.
- Van Nostrand, E.L., Pratt, G.A., Shishkin, A.A., Gelboin-Burkhart, C., Fang, M.Y., Sundaraman, B., Blue, S.M., Nguyen, T.B., Surka, C., Elkins, K., et al. (2016). Robust transcriptome-wide discovery of RNA-binding protein binding sites with enhanced CLIP (eCLIP). *Nat. Methods* 13, 508–514.
- Van Nostrand, E.L., Freese, P., Pratt, G.A., Wang, X., Wei, X., Xiao, R., Blue, S.M., Chen, J.Y., Cody, N.A.L., Dominguez, D., et al. (2020). A large-scale binding and functional map of human RNA-binding proteins. *Nature* 583, 711–719.
- Vessey, J.P., Vaccani, A., Xie, Y., Dahm, R., Karra, D., Kiebler, M.A., and Macchi, P. (2006). Dendritic localization of the translational repressor Pumilio 2 and its contribution to dendritic stress granules. *J. Neurosci.* 26, 6496–6508.
- Wang, X., McLachlan, J., Zamore, P.D., and Hall, T.M. (2002). Modular recognition of RNA by a human pumilio-homology domain. *Cell* 110, 501–512.
- Wang, J., He, Q., Han, C., Gu, H., Jin, L., Li, Q., Mei, Y., and Wu, M. (2012). p53-facilitated miR-199a-3p regulates somatic cell reprogramming. *Stem Cells* 30, 1405–1413.
- Wang, G., Guo, X., Hong, W., Liu, Q., Wei, T., Lu, C., Gao, L., Ye, D., Zhou, Y., Chen, J., et al. (2013). Critical regulation of miR-200/ZEB2 pathway in Oct4/Sox2-induced mesenchymal-to-epithelial transition and induced pluripotent stem cell generation. *Proc. Natl. Acad. Sci. U.S.A.* 110, 2858–2863.
- Wang, J., Jenjaroenpun, P., Bhinge, A., Angarica, V.E., Del Sol, A., Nookae, I., Kuznetsov, V.A., and Stanton, L.W. (2017). Single-cell gene expression analysis reveals regulators of distinct cell subpopulations among developing human neurons. *Genome Res.* 27, 1783–1794.
- Wickens, M., Bernstein, D.S., Kimble, J., and Parker, R. (2002). A PUF family portrait: 3'UTR regulation as a way of life. *Trends Genet.* 18, 150–157.
- Xue, Y., Ouyang, K., Huang, J., Zhou, Y., Ouyang, H., Li, H., Wang, G., Wu, Q., Wei, C., Bi, Y., et al. (2013). Direct conversion of fibroblasts to neurons by reprogramming PTB-regulated microRNA circuits. *Cell* 152, 82–96.
- Yu, J., Chen, M., Huang, H., Zhu, J., Song, H., Zhu, J., Park, J., and Ji, S.J. (2018). Dynamic m6A modification regulates local translation of mRNA in axons. *Nucleic Acids Res.* 46, 1412–1423.
- Zubradt, M., Gupta, P., Persad, S., Lambowitz, A.M., Weissman, J.S., and Rouskin, S. (2017). DMS-MaPseq for genome-wide or targeted RNA structure probing *in vivo*. *Nat. Methods* 14, 75–82.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
anti PUM2 antibody	Axil Scientific Pte Ltd	A300-202A; RRID:AB_2173752
anti TIA1 antibody	Scientific Hub Services PTE LTD	RN014P; RRID:AB_1570650
anti Flag antibody	Sigma-Aldrich Pte Ltd	F1804-50UG; RRID:AB_262044
anti Flag Magnetic beads	Merck Pte Ltd	M8823-5ML
Chemicals, peptides, and recombinant proteins		
PUM2 protein	Gene-Ethics (Asia) Private Limited	TP311307
BDNF	PEPROTECH	450-02
GDNF	PEPROTECH	450-10
EGF	Invitrogen	PHG0314
bFGF	Invitrogen	13256-029
mTeSR™1	stem cell technologies	#05851
5XSupplement	stem cell technologies	#05852
DMEM/F12	Thermo Fisher Scientific	11330-032
Neurobasal Medium	Thermo Fisher Scientific	21103-049
cAMP	Sigma-Aldrich Pte Ltd	A9501
Vitamin C	Sigma-Aldrich Pte Ltd	A4544
Glutamax	Invitrogen/GIBCO	35050-061
MEM-NEAA	Thermo Fisher Scientific	11140-050
N2	Invitrogen/GIBCO	17502-048
B27 minus A	Invitrogen/GIBCO	12587-010
B27	Invitrogen/GIBCO	17504-044
SB431542	Cellagentech	C7243-25
CHIR99021	Cellagentech	C2447-10
Dispase	stem cell technologies	#7923
Accutase	Corning	# 25-058-CI
ROCK inhibitor	Tocris	#1254
poly-L-lysine	Sigma-Aldrich Pte Ltd	P4707
laminin	Sigma-Aldrich Pte Ltd	L2020
Actinomycin D	Sigma-Aldrich Pte Ltd	A1410
NAI-N3	MedChemExpress	HY-103006
DMS	Sigma-Aldrich Pte Ltd	D186309
Trizol LS reagent	Life Technologies	10296-028
Click-IT biotin DIBO alkyne	Life Technologies	C-10412
RiboLock RNase inhibitor	Thermo Scientific	EO0384)
RNA fragmentation reagents	Ambion	AM8740
T4 polynucleotide kinase	New England BioLabs	M0201L
FastAP	Thermo Scientific	EF0651
T4 RNA ligase 1, high concentration	New England BioLabs	M0437M
RNase cocktail enzyme mix	Ambion, cat. no	AM2286
RNaseH	Enzymatics, cat. no.	Y9220L
CircLigase II ssDNA ligase	Epicenter	CL9025K
Phusion high-fidelity PCR master mix	New England BioLabs	M0531L
Dynabeads® MyOne Streptavidin C1	Thermo Fisher Scientific	#65001

(Continued on next page)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Critical commercial assays		
NEB ultra-directional RNA sequencing	NEB	#E7760
Nextera XT DNA Library Prep kit	Illumina	FC-131-1096
Illumina TruSeq Ribo Profile kit	Illumina	RPHMR12126
Dual-Luciferase Reporter Assay System	Promega	#E1980
TruSeq Ribo Profile Library Prep kit	Illumina	RPHMR12126
High Capacity cDNA Reverse Transcription Kit	Thermo Fisher Scientific	4368813
SYBR Green PCR Master Mix	Thermo Fisher Scientific	4309155
Deposited data		
All sequencing data have been uploaded to GEO	This Study	GSE156671
Unprocessed Images	This Study	https://doi.org/10.17632/m68gpy66yy.1
Experimental models: Cell lines		
H9 cells	Lawrence W. STANTON's lab	N/A
NPCs	Cells were differentiated from H9 cells	N/A
early Neuron and Neuron	Cells were differentiated from NPCs	N/A
HEK293T cells	Lawrence W. STANTON's lab	N/A
Oligonucleotides		
miR-30-5p RT primer	GTCGTATCCAGTGCCTGTCGTGGAG TCGGCAATTGCACTGGATACGACCT CCAGT	N/A
miR-30-5p Forward	CAGTGCCTGTCGTGGAGT	N/A
miR-30-5p Reverse	GCCTGTAAACATCCYYGAC	N/A
miR-125 RT primer	GTCGTATCCAGTGCCTGTCGTGGAG GTCGGCAATTGCACTGGATACGAC TCACAGGT	N/A
miR-125 Forward	CAGTGCCTGTCGTGGAGT	N/A
miR-125 Reverse	CAGTCCTGAGACCTTTA	N/A
U6 RT primer	CGCTTCACGAATTGCGTGTCA	N/A
U6 Forward	GCTTCGGCACATATACTAAAAT	N/A
Continued		
U6 Reverse	CGCTTCACGAATTGCGTGTCA	N/A
Lin28A qPCR Forward	GGGGAATCACCCTACAACCT	N/A
Lin28A qPCR Reverse	CTTGGCTCCATGAATCTGGT	N/A
Recombinant DNA		
pmirGLO Dual-Luciferase expression Vector	Promega	9PIE133
teton plasmid	Alfred Xuyang Sun's lab	N/A
teon-GFP plasmid	constructed in teton backbone	N/A
teon-PUM2 plasmid	constructed in teton backbone	N/A
teon-TIA1 plasmid	constructed in teton backbone	N/A
Software and algorithms		
icSHAPE	(Spitale et al., 2015)	https://github.com/qczhang/icSHAPE
cutadapt	(Martin, 2011)	https://cutadapt.readthedocs.io/en/stable/
bowtie2	(Langmead and Salzberg, 2012)	http://bowtie-bio.sourceforge.net/bowtie2/index.shtml
STAR	(Dobin et al., 2013)	https://github.com/alexdobin/STAR
Salmon	(Patro et al., 2017)	https://github.com/COMBINE-lab/salmon

(Continued on next page)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Cufflinks	(Trapnell et al., 2013)	http://cole-trapnell-lab.github.io/cufflinks/cuffdiff/
NOISeq	(Tarazona et al., 2015)	https://bioconductor.org/packages/release/bioc/html/NOISeq.html
DRIMSeq	(Nowicka and Robinson, 2016)	https://bioconductor.org/packages/release/bioc/html/DRIMSeq.html
featureCounts	(Liao et al., 2014)	http://subread.sourceforge.net/
R	R	https://www.r-project.org/
PrismNet	(Sun et al., 2021)	https://github.com/kuixu/PrismNet
Others		
genome annotation	(Hunt et al., 2018)	ensembl.org
RBP binding regions by eCLIP	(Van Nostrand et al., 2016)	https://www.encodeproject.org/
RBP binding motifs	(Giudice et al., 2016)	https://attract.cnic.es/

RESOURCE AVAILABILITY**Lead contact**

Further information and requests for reagents may be directed to and will be fulfilled by the Lead Contact, Yue Wan (wany@gis.a-star.edu.sg).

Materials availability

This study did not generate new unique reagents.

Data and code availability

- All sequence data have been deposited at GEO and are publicly available as of the date of publication. Accession numbers are listed in the [Key resources table](#). This paper analyzes existing, publicly available data. These accession numbers for the data-sets are listed in the [Key resources table](#).
- This paper does not report original code.
- Any additional information required to reanalyze the data reported in this paper is available from the lead contact upon request.

EXPERIMENTAL MODEL AND SUBJECT DETAILS**Cell culture and inducible gene expression in hESCs**

HEK293T cells were kindly gifted from Lawrence W. STANTON's lab and cultured in DMEM with 10% FBS and 1% Pen-Strep. Human ESCs, neural stem cell induction and neuronal differentiation were performed as previously described (Wang et al., 2017). Briefly, human embryonic stem cells (H9) were cultured in mTeSR™1 (#05851, stem cell technologies) medium, in the presence of 1X Supplement (#05852, stem cell technologies). Tet-on system was used to generate gene expression stable cell line by following previous study (Sun et al., 2019a). Briefly, H9 cells were infected with lentiviruses expressing tet-GFP, -PUM2 or -TIA1 together with rtTA. 1ug/ml puromycin was used to screen GFP, PUM2 and TIA1 expressing cells. 1ug/ml Doxycycline was used to induce gene expression of GFP, PUM2 and TIA1.

Neuronal differentiation was performed as in the previous publication (Wang et al., 2017). Briefly, H9 cells were dissociated using dispase (#07923, stem cell technologies), and seeded in a new cell culture dish to 20%-30% confluence. After 1-2 days, H9 cells were then treated with CHIR99021, SB431542 and Compound E in neural induction medium. We changed the culture medium every 1-2 days. After seven days, we split the cells 1:3 using accutase (# 25-058-CI, Corning) and seeded the cells on matrigel-coated plates. We then added the ROCK inhibitor (1254, Tocris) to a final concentration of 10μM when we were passaging the cells.

The derived 2 X10⁵ neural stem cells were subsequently seeded on poly-L-lysine (P4707, Sigma) and laminin (L2020, Sigma) coated 6-well plates in neural cell culture medium for 1-2 days. The cells were then cultured in neuronal differentiation medium: DMEM/F12 (11330-032), 1X N2 (17502-048), 1X B27 (17504-044), 300ng/ml cAMP (A9501), 0.2mM vitamin C (A4544-25), 10ng/ml BDNF (450-02), 10ng/ml GDNF (450-10) for 7 days.

METHOD DETAILS**RNA structure probing, icSHAPE library preparation and data analysis**

icSHAPE libraries were performed as previously described (Flynn et al., 2016). Briefly, cultured cells were dissociated and transferred to a 1.5ml Eppendorf tube. The cells were then washed with PBS and treated with DMSO or 50mM NAI-N3 for 10min at 37°C with

gentle rotation. RNA was then extracted using Trizol and poly(A)⁺ selected. We add on biotin on 0.5-1ug of poly(A)⁺ RNA using DIBO-biotin, before performing fragmentation, ligation, reverse transcription and PCR amplification. We performed 6-8 PCR cycles for DMSO treated samples, and 9-11 PCR cycles for NAI-N3 treated samples.

Raw sequencing reads from icSHAPE library were first trimmed to remove adaptors using cutadapt (version 1.12) (Martin, 2011), collapsed to remove PCR duplicates and trimmed to remove the leading 15-nt containing both the unique molecular identifiers and index sequences. The processed reads were then mapped to an in-house reference transcriptome using bowtie2 (version: 2.3.4.3) (Langmead and Salzberg, 2012). The in-house reference transcriptome contains the longest transcript of each gene, and was built based on the human reference genome GRCh38 and the gene annotation data from ENSEMBL version 92 (Hunt et al., 2018). The SHAPE reactivity score was calculated as previously suggested by (Spitale et al. 2015). Briefly, the RT stop counts for each base were counted and replicates were combined. The combined RT Stop for each transcript were first adjusted by the background base density from the corresponding DMSO libraries after subtracting the background RTstop obtained from DMSO libraries:

$$R = (RTstop_{NAI-N3} - \alpha RTstop_{DMSO}) / background_base_density_{DMSO}$$

Where the α is set to 0.25 according to Spitale et al. The R scores were then scaled to [0,1] with a 90% winsorization, where the top 5% of R were set to 1, and bottom 5% of R were set to 0. All remaining R scores were scaled accordingly.

The transcript was chosen for analysis if it has a mean RT stop count higher than 2 and a mean background base density higher than 100.

Riboswitch & RiboSNich benchmarks

1ug of each IVT RNA [6 riboswitches (with or without metabolite) and 2 riboSNiches] were refolded separately in PCR machine by heating the RNA at 90°C for 2 min, chilling on ice for 2 min, and then ramping up to 37°C at 0.1°C/s. We then incubated the RNA for another 20 min at 25°C, before treating the refolded RNA with NAI-N3 for 10 min. DMSO treatment was used as a negative control. NAI-N3 and DMSO RNA was purified using phenol:chloroform:isoamyl alcohol (25:24:1). Half of the RNA was used to run radioactive gels while the other half of the samples were used to make icSHAPE libraries. The working concentrations of the ligands for the riboswitches are: 75uM GUA (Guanine), 100uM FMN (flavin mononucleotide), 150uM SAM (S-adenosylmethionine), 100uM TPP (Thiamine pyrophosphate)

SHAPE-MaP library preparation and data analysis

We performed SHAPE-MaP following the previously published literature (Siegfried et al., 2014) with a few modifications. Briefly, we extracted the RNAs after the cells were treated with 100mM NAI for 10min. Poly(A)⁺ RNA were then purified using the Poly(A)Purist MAG Kit (AM1922). 1ug of poly(A)⁺ RNA was fragmented and used to perform reverse transcription in the presence of 6mM MnCl₂. This is then followed by second-strand synthesis and NEB ultra-directional RNA sequencing kit (NEB #E7760).

Raw sequencing reads from SHAPE-MaP library were first trimmed to remove adaptors using cutadapt (version 1.12) (Martin, 2011). Processed reads were then mapped to an in-house reference transcriptome using bowtie (version: 0.12.8) (Langmead and Salzberg, 2012). The mutation rate was computed as the ratio of non-reference alleles in a particular location. The reactivity was defined as

$$r_{NAI-N3} - r_{DMSO}$$

Whereby r_{NAI-N3} is the mutation rate from the NAI-N3 treated library and the r_{DMSO} is the mutation rate from the DMSO treated library. Reactivities were then normalized by dividing by the mean reactivity of the top 10% of reactivities gene-wise, after reactivities above a threshold are excluded. The threshold is given as suggested by Low and Weeks (2010).

$$\max[1.5 * IQR, 90\% \text{ percentile if gene length} \geq 100 \text{ nt. OR } 95\% \text{ percentile if gene length} < 100 \text{ nt.}]$$

where IQR is the inter quantile range for the reactivities. The normalized reactives were then used for following analysis.

DMS-MaP-Seq library preparation and data analysis

We performed DMS-MaP-Seq following the previously published literature (Zubradt et al., 2017) with a few modifications. Briefly, total RNA was extracted after the cells were treated with 4% DMS for 5min and quenched using 30% BME. Poly(A)⁺ RNA were then purified using the Poly(A)Purist MAG Kit (AM1922). 1ug of poly(A)⁺ RNA was fragmented and used to perform reverse transcription by SSII in the presence of 6mM MnCl₂. This is then followed by second-strand synthesis and NEB ultra-directional RNA sequencing kit (NEB #E7760).

The first 5 bases were trimmed from the raw sequencing data from DMS-MaP-Seq libraries as suggested by Zubradt et al. (2017) for SSII generated libraries. The trimmed reads were then collapsed to remove PCR duplicates using the BBMap package (Bushnell, 2014). The collapsed reads were further cleaned to remove adaptors using cutadapt (version 1.12) (Martin, 2011). The resulting clean reads were mapped to our in-house reference transcriptome using Bowtie2 (version: 2.3.5.1) (Langmead and Salzberg, 2012). The best mapping location for multiple mapping reads were selected based on the best alignment score or randomly if the alignment scores are same. The mutation rate was computed as the ratio of non-reference alleles in a particular location. Only the nucleotides whose reference allele is A or C and the mutation rate is $\leq 20\%$ were considered. The raw reactivity was defined as:

$$r_{DMS} - r_{control}$$

Whereby r_{DMS} is the mutation rate from the DMS treated library and the $r_{control}$ is the mutation rate from the untreated library. To reduce noise, we filtered out the nucleotides with negative raw reactivity values. For each gene, the raw reactivity values were winsorized to 5%–95% and further scaled to [0,1]. Transcripts need to have 1) a mean depth of coverage (for A/C only) $>= 100$, 2) more than 50 A/C detected and 3) more than 50% of the total A/C detected to be included in analysis. The reactivity per gene is calculated as the mean reactivities across all the remaining A/C nucleotides.

Enhanced CLIP (eCLIP) library preparation

eCLIP libraries were prepared as described in the previous publication (Van Nostrand et al., 2016). Briefly, cells were grown on 10cm plates, washed twice with PBS and cross-linked at 254-nm UV with an energy setting of 400 mJoules/cm². The cells were then harvested, lysed and sonicated. 2% of the cell lysate was used as input, and the rest of the cell lysate was incubated with anti-Flag M2 Magnetic beads (M8823-5ML) overnight at 4°C. The incubated beads mixture (IP samples) was then washed for 5 times. Input and IP RNA samples were size selected using SDS-PAGE gel electrophoresis, purified and reversed transcribed into a cDNA library. The cDNA library was then amplified using PCR and deep sequenced.

RNA pull down experiment

We pulled down LIN28A using a protocol that was previously described (Tan et al., 2014). Briefly, pcDNA-PUM2 was co-transfected with pmirglo- LIN28A-3729 or LIN28A-1485 into HEK293T cells in 6-well plates for 48 hours. The cells were then harvested and lysed. We added biotinylated probes against LIN28A into the cell lysate and allowed the probe to hybridize by incubating overnight at 4°C with continuous rotation. We then added 20ul of Dynabeads® MyOne Streptavidin C1 (#65001) into each sample and rotated at room temperature for 45 minutes. We pulled down the RNA by placing the magnetic beads on a magnetic stand and washing the beads as previously described. Pulled down RNA is eluted and extracted from the beads using Trizol. We then performed qRT-PCR for miR-30 as previously described (Wang et al., 2012). Briefly, cDNA was synthesized from 1ug RNA using High Capacity cDNA Reverse Transcription Kit (Thermo Fisher Scientific, #4368813) according to the manufacturer's guide. The specific reverse transcription primers of miR-30: GTC GTA TCC AGT GCG TGT CGT GGA GTC GGC AAT TGC ACT GGA TAC GAC CTT CCA GT. qPCR was then performed using SYBR Green PCR Master Mix (Thermo Fisher Scientific, #4309155). qPCR primer of miR-30: forward: CAG TGC GTG TCG TGG AGT, Reverse: GCC TGT AAA CAT CCY YGA C. U6 was used as loading control.

In vitro RNA-protein interaction experiment

We obtained individual RNA fragments (see supplemental table) using PCR amplification and *in vitro* transcription. We then pooled the RNA fragments together and added 12 pmol of RNA pool into 60ul of structure probing buffer (final concentration: 50mM Tris PH7.4, 10mM MgCl₂ and 150mM NaCl). We aliquoted 20ul of RNA into 3 tubes, and refolded the RNAs in thermocycler by heating the RNA sample at 90°C for 2min, cooling down the sample at 4°C for 2min, and then slowly ramping up the sample to 37°C at 0.1°C/s. We incubated the sample for another 20min at 37°C before adding 0, 4 and 10 pmol of purified PUM2 proteins (#TP311307) into the 3 aliquots respectively and incubating for 1 hour. We then added NAI-N3 (final concentration of 50mM) or DMSO to each reaction for 10min at 37°C for structure probing. The RNAs were purified using phenol:chloroform:isoamyl alcohol (25:24:1) before being converted into a cDNA library following icSHAPE protocol (Siegfried et al., 2014).

Luciferase assay

Target regions of PUM2 were PCR amplified and cloned into pmirGLO Dual-Luciferase Expression Vector (# 9PIE133). pcDNA-PUM2/pcDNA-GFP and target pmirGLO-candidates were then co-transfected into HEK293T cells. After 48 hours post-transfection, cells were then lysed and luciferase assays were performed using Dual-Luciferase Reporter Assay System (#E1980).

Identification of structure changing regions in the transcriptome

We designed a computational pipeline to identify the regions with significant structural changes. First, we split our reference transcriptome into individual windows of 20 nucleotides in length, and counted all the RT stops within the windows for each of the NAI-N3 treated libraries. We tested the optimal window size using data from two technical replicates. A window size of 20 nucleotides could obtain a per-gene correlation between technical replicates that is not significantly different from that of larger window sizes (Figure S5E). Next, we filtered out windows with a RT stop count of ≤ 50 to remove poorly expressed transcripts. We then normalized the RT stop counts by calculating the RPM (reads per million mapped reads) for each window within each NAI-N3 library. Only windows with higher RPM in the NAI-N3 treated library than in the corresponding DMSO library were kept for downstream analysis. To enable data comparison between the different libraries, we performed an upper quantile normalization for each library. The RTstop counts in NAI-N3 treated library was further scaled by the ratio of RTstop counts between the corresponding DMSO libraries. To further reduce noise in our data, the windows were filtered out if (1) the average RT stop counts in the same window of the corresponding DMSO libraries was less than 20 reads; or (2) the coefficient of variation of the RT stop counts within the DMSO replicates was higher than 0.3. Finally, we applied NOISeq with biological replicates (noiseqbio function) (Tarazona et al., 2015) on the normalized RPMs of the windows in NAI-N3 treated libraries. NOIseq does not require the data to follow a particular distribution to perform

differential analysis, making it suitable to analyze data from different systems and distributions. The regions with significant structural changes were defined as windows with (1) an adjusted *p*-value of not more than 0.05, and (2) having a fold-change of not less than 1.5 fold.

We validated our analysis pipeline using 6 riboSWitches and 2 riboSNiches, which are known to change structures with the presence of metabolites or mutations. To identify the structure changing regions along these RNAs, we performed gel electrophoresis in the presence and absence of metabolites or mutations. We then quantitated the relative intensity of each band on the gel using the program SAFA (Das et al., 2005). For each riboSWitch or riboSNitch, the per-nucleotide intensity was normalized to the average intensity per aptamer as follows:

$$\hat{I}_i = \frac{I_i}{\sum_{i=1}^N I_i / N}$$

Where the \hat{I}_i is the normalized intensity of nucleotide i ; I_i is the raw intensity of nucleotide i from SAFA; N is length of the aptamer used to run gel. Based on the assumption that the majority of the nucleotides in a riboswitch and riboSNitch will not change its secondary structure in the presence of metabolites or mutations, the distribution of N_i before and after structural change should be approximately the same. The difference of N_i will hence follow a normal distribution. Therefore, we calculated a Z-score, Z_i , for each nucleotide as follows:

$$Z_i = \frac{(\Delta \hat{I}_i - \mu)}{s}$$

Where the $\Delta \hat{I}_i$ is the difference of \hat{I}_i before and after structural changes; μ is the average of $\Delta \hat{I}_i$; and s is the standard deviation of $\Delta \hat{I}_i$. A nucleotide is defined as having a significant secondary structure change when the corresponding Z_i is $>= 1.96$ or ≤ -1.96 .

We next performed our pipeline to identify the structural changed regions in riboswitches and riboSNitches using icSHAPE data. All the cutoff and filters were applied as described above, except that the RTstop count per nucleotide is used as input, and the minimum RTstop count was set to 2.5 (based on the RTstop cutoff of 50 in a 20-nt window). This is due to the short length of riboswitches and riboSNitches used in footprinting and SAFA analysis. To benchmark the performance of our pipeline, we also tested other metrics to quantify the structural changes: (1) “delta” by taking the absolute difference between SHAPE reactivities per nucleotide. (2) “delta-win” by taking the absolute difference between average SHAPE reactivities in a 10-nt sliding windows with a step size of 1-nt. (3) “T-test” by considering the p value from a t test using SHAPE reactivities in a 10-nt sliding windows with a step size of 1-nt. The method is adopted from Shi et al. (2020). (4) “correlation” by taking the pearson correlation between SHAPE reactivities in a 10-nt sliding windows with a step size of 1-nt. (5) “diff-BUMHMM” by taking the posterior probability of modification changes. The method is adopted from Marangio et al. (2020). An AUC score was obtained for each method by comparing the variables described above with the structure changed nucleotides in riboswitches and riboSNitches from SAFA analysis. For our pipeline (termed as NOISeqRT), the p values from NOISeq were used to obtain AUC score. Among all of the methods tested, NOISeq obtained the highest AUC score of 0.777 (Figure S5C). In addition, adding an absolute fold changes filter of 1.5 further increases the AUC score for NOISeq to 0.824.

Filtering of windows associated with alternative splicing

As we are using a redundant reference transcriptome whereby only the longest transcript for each gene is included, differential splicing of the same gene might result in windows being artificially called to have significant structural changes. In order to filter these windows out, we identified the genes with different transcripts between the two time points. We first used DRIMseq to find the genes with significant differential transcript usage. Then, we identified the genes with dominant isoform changes based on TPM calculations using Salmon. The genes with both differential transcript usage by DRIMseq and dominant isoform changes by RPM were defined as genes with “significant dominant isoform changes.” The windows within the genes were removed for the downstream analysis. For the genes that show significant differences in isoform abundance between different time points using DRIMseq, but had the same isoform as the dominant isoform across time points, we filtered out the windows located ± 40 bp to the splice sites of the genes.

Ribosome profiling library preparation and analysis

Ribosome profiling was performed by following Illumina TruSeq Ribo Profile (mammalian) Library Prep kit (RPHMR12126). The raw sequencing reads from Ribosome profiling library were first processed using cutadapt (version 1.12) (Martin, 2011) to remove adaptor sequences and reads with low quality or short length. The processed reads were then mapped to the reference genome (UCSC hg19 with annotations from GENCODE v.19) using program STAR (version 2.5.0c) (Dobin et al., 2013). The expression level per gene was quantified by Cuffdiff (version 2.2.1) (Trapnell et al., 2013). The gene translation efficiency (TE) was defined as the \log_2 value of the ratio between RPKM of the footprinting library and RNA sequencing library. The genes with significant differences between TE were identified using NOISeq optimized for the use on biological replicates (Tarazona et al., 2015).

RNA sequencing library preparation and analysis

RNA sequencing library was performed using the NEBNext® Ultra II Directional RNA Library Prep Kit for Illumina (NEB #E7760L), following manufacturer’s instructions. The raw sequencing data from RNaseq library were first processed using cutadapt (version

1.12) (Martin, 2011) to remove adaptor sequences and reads of low quality or short length. The processed reads were then mapped to the reference genome (UCSC hg19 with annotations from GENCODE v.19) using STAR (version 2.5.0c) (Dobin et al., 2013). The differentially expressed genes were identified using Cuffdiff (version 2.2.1) (Trapnell et al., 2013). In addition, we also quantified the level of transcripts using the program Salmon (version 0.11.2) (Patro et al., 2017). Transcript annotations were obtained from GRCh38 from ENSEMBL (version 92) (Hunt et al., 2018). The differentially expressed transcripts were identified using DRIMseq from Bioconductor (Nowicka and Robinson, 2016).

RNA decay library preparation and analysis

To determine the half-life of cellular RNAs, we stopped transcription by treating cells with 5 μ M of Actinomycin D (A1410, Sigma) for 0, 1, 2, 4 and 8 hours by following previous report (Mizrahi et al., 2018). We then extracted the RNA using Trizol, used 1-2ng RNA for reverse transcription and performed 8 cycles of PCR amplification. The derived PCR products were then used to make into a sequencing library using the Nextera XT DNA Library Prep kit (FC-131-1096).

The raw sequencing reads from mRNA decay library were mapped to our reference genome (GRCh38) using STAR (version 2.5.0c) (Dobin et al., 2013). The gene annotation data was collected from ENSEMBL (2020). The RPM values of each gene were quantified using featureCounts (version 1.6.3) (Liao et al., 2014), and then used to estimate mRNA half-life. We used the mean RPM values of a few house-keeping genes (GAPDH, ACTB, PGK1, PPIA, RPL13A, RPLP0) that are known to have consistent expression upon Actinomycin D treatment as reference, and normalized the expression levels of other genes to them across different time points post treatment (0h, 1h, 2h, 4h and 8h after treatment). The normalized expression levels were used to fit into an exponential decay model for each gene to calculate half-life:

$$\text{expression}_t = a * \exp(-b * t)$$

where **a** is the expected expression before treatment, **b** is the decay constant and **t** is time in hours after treatment. A minimum correlation of 0.75 between fitted and observed gene levels was used to filter out genes with poor fit. Genes with significant differences in estimated half-lives between two time points were identified using NOISeq (Tarazona et al., 2015).

RBP enrichment and PrismNet analysis

The binding sites of 183 different RNA binding proteins (RBP) determined by eCLIP (Van Nostrand et al., 2016) were downloaded from ENCODE. We then overlapped the binding sites of each RBP with reactivity changing and non-changing regions. The enrichment of RBP in the reactivity changing versus non-changing regions were tested using hypergeometric tests. The resulting *p*-values were FDR corrected using the bonferroni method. RBPs with an adjusted *p*-value of smaller than or equal to 0.01 are considered as enriched.

Independent of eCLIP, we also developed a *de novo* approach using sequence motifs collected from the ATtRACT database (Giudice et al., 2016) to identify RBP binding motifs that are enriched in the reactivity changing regions. First, we identified 200 different 6-mer sequences enriched in the reactivity changing regions using hypergeometric test with bonferroni correction (adjusted *p*-value cutoff = 0.01). Then the enriched 6-mers were compared to the known regulatory motif sequences using pairwise local sequence alignment. A motif was defined as enriched when its sequence is similar to the 6-mer with at most 1 base difference. The enriched RBP by motif were defined as the RBPs with at least 1 motif sequence enriched in the structure changed regions. Last, we identified 10 RBPs that are significantly enriched in structure changing regions by using both the binding sites and motif sequence enrichment analysis. In addition to eCLIP and Motif based methods, we also used the PrismNet predicted RBP binding sites to perform the RBP enrichment analysis in the structural changing regions. The predicted RBP binding sites in H9 cells were adopted from Sun et al. (2021)

miRNA enrichment analysis

The potential miRNA binding sites are obtained from the predicted non-conserved miRNA sites in TargetScan release 7.2 (Agarwal et al., 2015). The miRNA used for analysis were the highly expressed miRNA in hESC and NPC (top 100 RPM using small RNA sequencing from both developmental stages). The enrichment of miRNA near the RBP binding sites with reactivity changing regions were computed using hypergeometric test.

Analysis of m6A binding sites

m6A binding sites in hESC were obtained from previous literature by Batista et al. (2014). The binding sites are presented as 100-nucleotide regions in genomic coordinates. We converted the genomic coordinates into transcriptomic coordinates by matching the sequences.

RNA structure modeling

Secondary structures of LIN28A were predicted by incorporating *in vivo* icSHAPE data into the program RNAstructure (Deigan et al., 2009). Use of raw icSHAPE data in structure modeling was precluded, as the discontinuous nature of the data caused divergence in partition function calculations. Hence, we proceeded to smooth the raw icSHAPE data in a 5 nucleotide window moving average before incorporating it into secondary structure modeling. We used default parameters for shape intercept ($-0.6 \text{ kcal mol}^{-1}$) and

shape slope ($1.8 \text{ kcal mol}^{-1}$) respectively. Separate partition functions were calculated for LIN28A regions 1390-2000, 2200-2600 and 3600-4014 after evaluating the likely structure boundaries of a full-length LIN28A model. After obtaining partition functions, pairing probabilities for all bases were extracted and Shannon entropy calculated for each position. Concurrently, maximum likelihood structures were calculated based on the partition functions. VARNA (Darty et al., 2009) was used for visualization and moving average icSHAPE scores were added as a color map.

Separate structural models were constructed from icSHAPE data obtained under conditions of overexpression GFP control and PUM2 respectively and changes in LIN28A structures under these conditions are shown in Figure 5A. Mutants for testing PUM2 binding were subsequently designed based on these structure models (Figures 5A and 5B).

QUANTIFICATION AND STATISTICAL ANALYSIS

All statistical analysis was performed in R (version 3.6.3) unless otherwise stated in the methods. Student's T Tests were performed using the `t.test()` function in R. The hypergeometric tests were performed using the `phyper()` function in R. The binomial tests were performed using the `binom.test()` function in R. All other statistics data is specified in the figure legends.

Supplemental information

**Genome-wide RNA structure changes
during human neurogenesis modulate
gene regulatory networks**

**Jiaxu Wang, Tong Zhang, Zhang Yu, Wen Ting Tan, Ming Wen, Yang Shen, Finnlay R.P.
Lambert, Roland G. Huber, and Yue Wan**

SUPPLEMENTAL INFORMATION

Figure S1 Data quality of high throughput libraries. Related to Figure 1

Figure S2 Summary of detected genes by different methods. Related to Figure 1

Figure S3 Global features of RNA structures using icSHAPE. Related to Figure 1

Figure S4 Genes in hESC are more accessible than in NPC. Related to Figure 1

Figure S5 Benchmarking of icSHAPE analytic methods using riboswitches and riboSNitches Related to Figure 1

Figure S6 Summary of the significantly structural changing windows. Related to Figure 1

Figure S7 The association of structural changes with other methods. Related to Figure 1

Figure S8 Clustering of significant structural changing windows. Related to Figure 2

Figure S9 Overlap of reactivity changing windows with cellular regulators. Related to Figure 2

Figure S10 Enrichment of RBPs in significant structure changing windows. Related to Figure 3

Figure S11 Overexpression of PUM2 and TIA1 in hESC. Related to Figure 3

Figure S12 Number of changing windows per gene upon PUM2 over-expression. Related to Figure 3

Figure S13 Regulation of RBP-RBP and RBP-miRNA interactions through structure. Related to Figure 4

Figure S14 PUM2 and miR-30 regulated LIN28A through structure. Related to Figure 4

Figure S15 Impact of PUM2 binding on LIN28A Related to Figure 4

Figure S16 Foot printing of mutated RNA sequences. Related to Figure 5

Supplementary Figure Legends

Figure S1. Data quality of high throughput libraries. **A**, Workflow of the multi-omics analysis framework. **B**, Heatmap of correlation between the biological replicates of hESC, NPC, iNeu and Neu treated with both DMSO and NAI-N3. **C**, Correlation of RT stop counts between 2 biological replicates of icSHAPE libraries for hESC, NPC, iNeu and Neu, from left to right. **D**, Correlation of RPKM values per gene between 2 biological replicates of ribosomal profiling libraries for hESC, NPC, iNeu and Neu, from left to right. **E**, Correlation of TPM values per gene between 2 biological replicates of RNA sequencing libraries for hESC, NPC, iNeu and Neu, from left to right. **F**, Correlation of RPM values per gene between 2 biological replicates of RNA decay libraries at different time points post treatment (left to right), and from hESC, NPC, iNeu and Neu (top to bottom). The Pearson correlation (R) and the p -value are shown inside the subplots **C-F**. **G**, GO enrichment analysis of genes with differential expression, translation efficiency and RNA half-life between neighbor time points.

Figure S2. Summary of detected genes by different methods. **A**, Number of genes detected at each stage of neuronal differentiation by icSHAPE (RNA structure), ribosomal profiling, RNA decay and RNA sequencing, respectively. **B**, Upset plot showing the number of shared genes detected by different methods for hESC, NPC, iNeu and Neu, respectively. **C**, Venn diagram showing the overlapping of genes detected by all 4 methods in all stages of differentiation. **D**, Upset plot showing the number of genes detected by all 4 methods in all stages during differentiation.

Figure S3. Global features of RNA structures using icSHAPE. **A, Bottom:** Metagene analysis of icSHAPE reactivity centered on translation start and stop sites showed a three-nucleotide periodicity in the coding region across all differentiation stages. **Top:** Autocorrelation of the reactivity in the 4 stages from hESC to Neu (Red-hESC, green-NPC, light blue-iNeu and purple-Neu). **B**, Correlation between the Gini index of reactivity from icSHAPE and TE calculated from ribosome profiling for hESC, NPC, iNeu and Neu, from left to right. The Pearson correlation (R) and the p -value are shown inside the plots.

Figure S4. Genes in hESC are more accessible than in NPC. **A, B**, Boxplots showing the distribution of the mean icSHAPE reactivities (**A**) and Gini index (**B**) per gene for hESC, NPC, iNeu and Neu. The genes used here are detected by icSHAPE in all 4 stages. P -value is calculated using Student's T- Test. **C**, Correlation between gene expression levels and the mean icSHAPE reactivity in hESC, NPC, iNeu and Neu from left to right. The Pearson correlation (R) and the p -value are shown inside the plots. **D**, Boxplot of the average reactivities per gene obtained by SHAPE-MaP in all 4 stages. **E**, The AUC-ROC curves of the normalized reactivities obtained by DMS-MaP-Seq in hESC and NPC cells using the 18S

rRNA. The AUC-ROC values are labeled in legend. **F**, Boxplot of the average reactivities per gene obtained by DMS-MaP-Seq in hESC and NPC cells.

Figure S5. Benchmarking of icSHAPE analytic methods using riboswitches and riboSNitches. **A**, Table of riboswitches and riboSNitches used for benchmarking. **B**, Footprinting images of riboswitches and riboSNitches. For each gel, we are displaying G ladder (lane 1), DMSO treated RNA (lane 2), NAI-N3 treated RNA without ligand (riboswitch) or WT (riboSNitch) (lane 3) and NAI-N3 treated RNA with ligand (riboswitch) or Mutant (riboSNitch) (lane 4). **C**, The AUC-ROC curves of the different methods (NOISeqRT in red, delta in blue, delta-window in black, T-test in green, correlation in cyan, and diff-BUMHMM in light blue) used for calculating reactivity changes. The AUC-ROC scores are labeled in legend box. **D**, The AUC-ROC curves of different fold-change cutoffs (1X is in blue, 1.5X is in red, and 2X is in green) used in the NOISeqRT method. The AUC-ROC scores are labeled in legend box. **E**, Boxplots of per gene correlations between hESC replicates using different window sizes ranging from 5 to 50 nucleotides from left to right.

Figure S6. Summary of the significantly structural changing windows. **A**, Pie charts showing the numbers of windows with (orange) and without (grey) significant reactivity changes during neuronal differentiation. **B**, Histogram showing the number of genes with 1 or more significant structural changing windows during hESC-NPC (left), NPC-iNeu (middle) and iNeu-Neu (right). **C**, Table showing the number of significantly structural changing regions with different sizes after merging neighboring significant changing windows in hESC-NPC, NPC-iNeu and iNeu-Neu. **D**, Bar plot showing the percentage of significantly structural changing regions with the size of 20nt, 40nt, 60nt and more during hESCs-NPC (blue), NPC-iNeu (orange) and iNeu-Neu (grey).

Figure S7. The association of structural changes with other methods. **A**, Correlation of changes in Gini index with changes in TE (top row), half-life (middle row), and gene expression (bottom row) during hESCs-NPC, NPC-iNeu and iNeu-Neu differentiation, from left to right. **B**, Enriched biological processes obtained by GO analysis for the genes with ≥ 2 significant structural changing regions during hESCs-NPC, NPC-iNeu and iNeu-Neu differentiation from left to right.

Figure S8. Clustering of significant structural changing windows. **A**, Line chart showing the number of optimal number of K means clusters predicted using Gap statistics. The Y-axis indicates the Gap statistic and the X-axis indicate the number of clusters. **B**, Histogram showing correlations between normalized RT stop counts by icSHAPE and half-life estimated from RNA decay (blue), gene expression level from RNA sequencing (yellow) or translation efficiency from ribosomal profiling (grey) for cluster 1 to 5 (top to bottom panels). The red

dashed lines indicate the correlation values at -0.8 and 0.8 in each subplot. **C**, Barplot showing *p*-values of enriched biological processes from GO analysis using the corresponding genes of windows from cluster 1 (top row), 3 (middle row) and 5 (bottom row) with either positive correlation (colored in blue, $R \geq 0.8$) or negative correlation (colored in orange, $R \leq -0.8$) with half-life (Decay) or TE (Translation Efficiency). **D**, Boxplot showing the expression levels of genes from clusters 1 to 5 from left to right for hESCs (blue), NPC (yellow), iNeu (grey) and Neu (red).

Figure S9. Overlap of reactivity changing windows with cellular regulators. **A**, *P*-value of enrichments of cellular factor binding sites in reactivity changing windows in clusters 1-5 and all clusters (from top to bottom). We calculated the enrichment of predicted binding sites of RBPs (blue), miRNAs (yellow), m6A (grey) and splicing sites (red) that overlap with 10-50 nucleotides upstream and downstream from the reactivity changing region. *P*-value is calculated using the binomial test. The expected proportion of overlap for each regulator is calculated from all overlapping proportions of different regulators to all windows in 5 clusters (bottom row). **B**, Barcharts showing significantly enriched miRNAs in cluster 1 using with varying flanking regions from 10 to 50 nucleotides (from left to right) of reactivity changing regions.

Figure S10. Enrichment of RBPs in significant structure changing windows. **A**, Heatmap of *p*-values (transformed by $-\log_{10}$) of RBP binding sites enrichments in significant structural changing windows during hESC-NPC, NPC-iNeu and iNeu-Neu. **B**, Venn diagram of the significantly enriched RBPs in structural changing windows using binding sites predicted by eCLIP, by RBP binding motifs and by PrismNet. **C**, Barcharts showing the fold changes of expression (left) and translation efficiency (right) of the 10 enriched RBPs from hESCs to NPC(Figure 3B). **D**, Western blot experiments showing the protein levels of PUM2 (left) and TIA1 (right) in hESC and NPCs. Beta-actin is used as loading control.

Figure S11. Overexpression of PUM2 and TIA1 in hESC. **A, B**, Western blot experiments using antibodies against FLAG tag (**A**) or TIA1 (**B**) when cells are over-expressed with 3X FLAG-GFP or 3X FLAG TIA1. B-actin is used as loading control. **C**, Western blot experiments using antibodies against FLAG tag (top) or PUM2 (middle) when cells are over-expressed with 3X FLAG-GFP or 3X FLAG PUM2. B-actin is used as loading control. **D**, Western blot experiments using antibodies against PUM2 and TIA1 in control, PUM2 knockdown cells and TIA knockdown cells. B-actin is used as loading control. **E**, Barcharts showing the percentages of PUM2-targeting windows becoming significantly more accessible (orange) or less accessible (blue) in knockdown (KD) and overexpression (OE) of PUM2. **F**, Barcharts showing the percentages of TIA1-targeting windows becoming significantly more accessible (orange)

or less accessible (blue) in knockdown (KD) and overexpression (OE) of TIA1. **G**, Metagene analysis of normalized icSHAPE reactivity around predicted binding sites of PUM2 (left) or TIA1 (right) by PrismNet before and after PUM2 (left) or TIA1 (right) overexpression. *P*-values on top of plot were calculated per nucleotide using the single-sided T-test. The center of each PrismNet predicted binding sites were treated as relative position =0.

Figure S12. Number of changing windows per gene upon PUM2 over-expression. The distribution of number of structural changing windows per gene upon overexpression of PUM2 in hESC. Representative genes are shown below the barchart. The genes selected for luciferase assays are highlighted in red.

Figure S13. Regulation of RBP-RBP and RBP-miRNA interactions through structure. **A**, Boxplot showing the Jaccard similarity of the RBP targeting windows between enriched RBP pairs (blue) and non-enriched RBP pairs (yellow). The enriched and non-enriched RBP pairs are shown in Figure 4A. **B**, Heatmap showing changes in gene regulation (translation, decay and gene expression) of targets of RBP pairs when there are structure changes versus when there are no structure changes. *P*-value of enrichment is calculated using T-test. ‘Dec’ represents decrease in translation efficiency, half-life or expression, ‘Inc’ represents increase in TE, half-life or expression. *P*-value of enrichment is calculated using hypergeometric test.

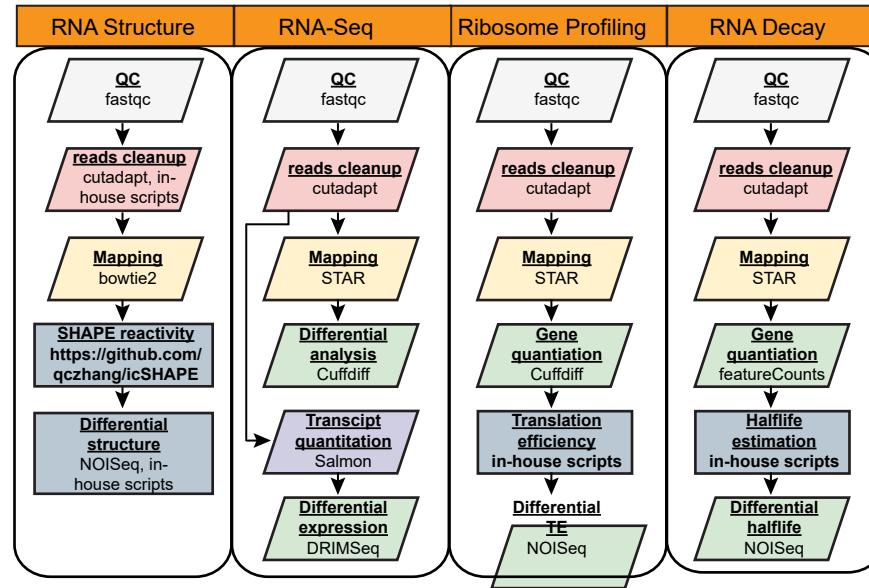
Figure S14. PUM2 and miR-30 regulated LIN28A through structure. **A**, Bar chart showing the relative expression level of LIN28A in hESC (H9, yellow) and NPC (orange). **B**, Gene expression of LIN28A upon 0, 1, 2, 4, 8 hours post actinomycin D treatment in hESC (black line) and in NPC (red line). **C**, Top, different isoforms of LIN28A based on UCSC genome browser. Bottom left, the depth of coverages obtained from eCLIP experiments of PUM2 in hESC and NPC cells on the 3'UTR region of gene LIN28A (from positions 3400 to 3800) were shown. The PUM2 binding site (position 3729) and miR-30 binding site (position 3702) are labeled. Bottom right, bar chart showing the expression levels of LIN28A isoforms in hESC and NPCs from RNA sequencing data. **D**, icSHAPE reactivity profile of LIN28A upon overexpression of GFP (top) or PUM2 (bottom). The grey bars are the reactivity changing regions between GFP and PUM2 overexpression. **E, F**, in vitro SHAPE-MaP reactivity profiles of PUM2 and miR-30 (**E**) or miR-125 (**F**) binding regions along LIN28A in the presence of increasing concentrations of PUM2 (0, 200nM, 500nM).

Figure S15. Impact of PUM2 binding on LIN28A. **A, B**, Barplots showing the expression levels of LIN28A after PUM2 knock down in hESC cells (**A**) and NPC cells (**B**). **C, D**, Barplots showing the amount of miR-30 (**C**) and miRNA-125 (**D**) associated with LIN28A upon pulldown of LIN28A, using biotinylated probes, in hESC and NPC cells.

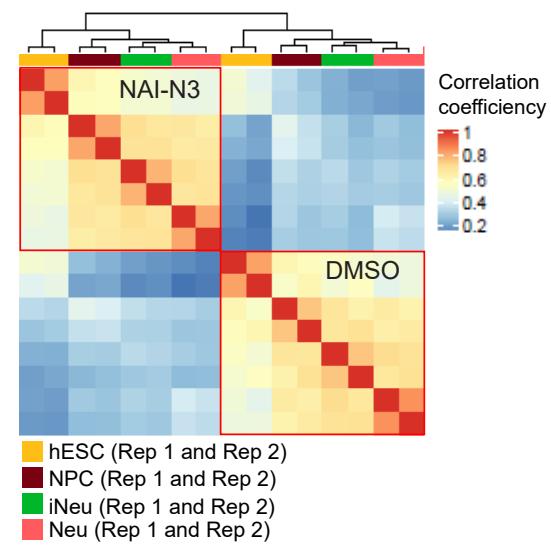
Figure S16. Foot printing of mutated RNA sequences. Sequencing gel showing RT stoppages from WT and mutant DMSO and NAI-N3 treated RNAs. The G ladder is in lane 11, WT DMSO and NAI-N3 treated samples are in lanes 1 and 2 respectively, Mutant A DMSO and NAI treated samples are in lanes 3 and 4, Mutant B DMSO and NAI treated samples are in lanes 5 and 6, compensatory Mutant C DMSO and NAI treated samples are in lanes 7 and 8, and Mutant D DMSO and NAI treated samples are in lanes 9 and 10. miR-30 binding motif and PUM2 binding motif region were boxed in red and pink respectively. The band intensity of the boxed images were measured using ImageJ. The intensity of the miR-30 RNA binding site in WT and mutants are shown as barcharts on the top right. The intensity of the PUM2 RNA binding site in WT and mutants are shown as barcharts on the bottom right.

Figure S1 Data quality of high throughput libraries. Related to Figure 1

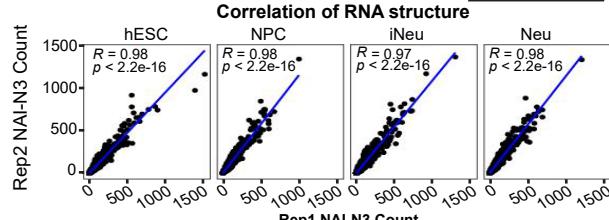
A



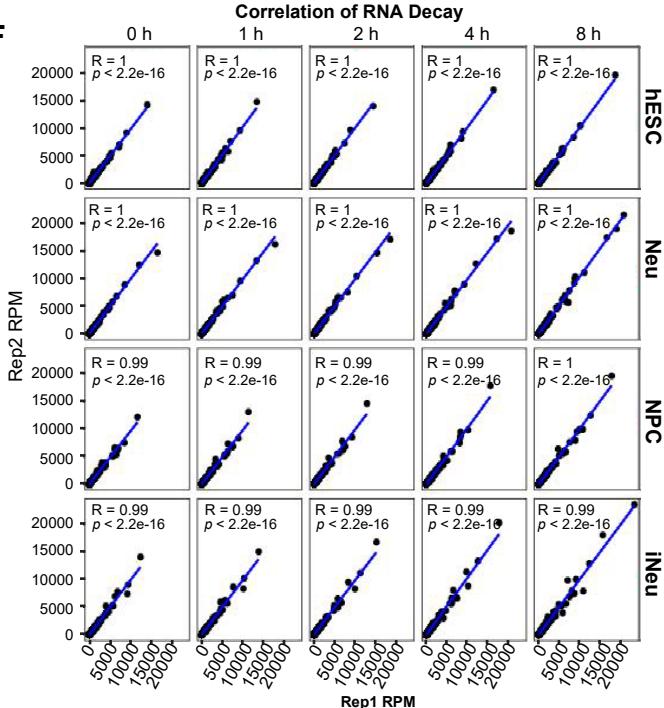
B



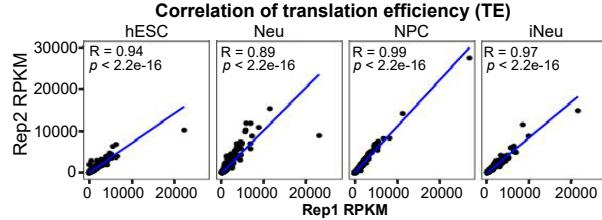
C



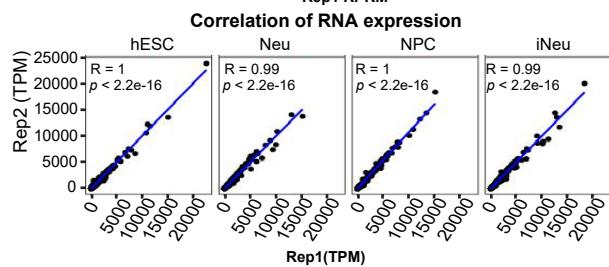
F



D



E



G

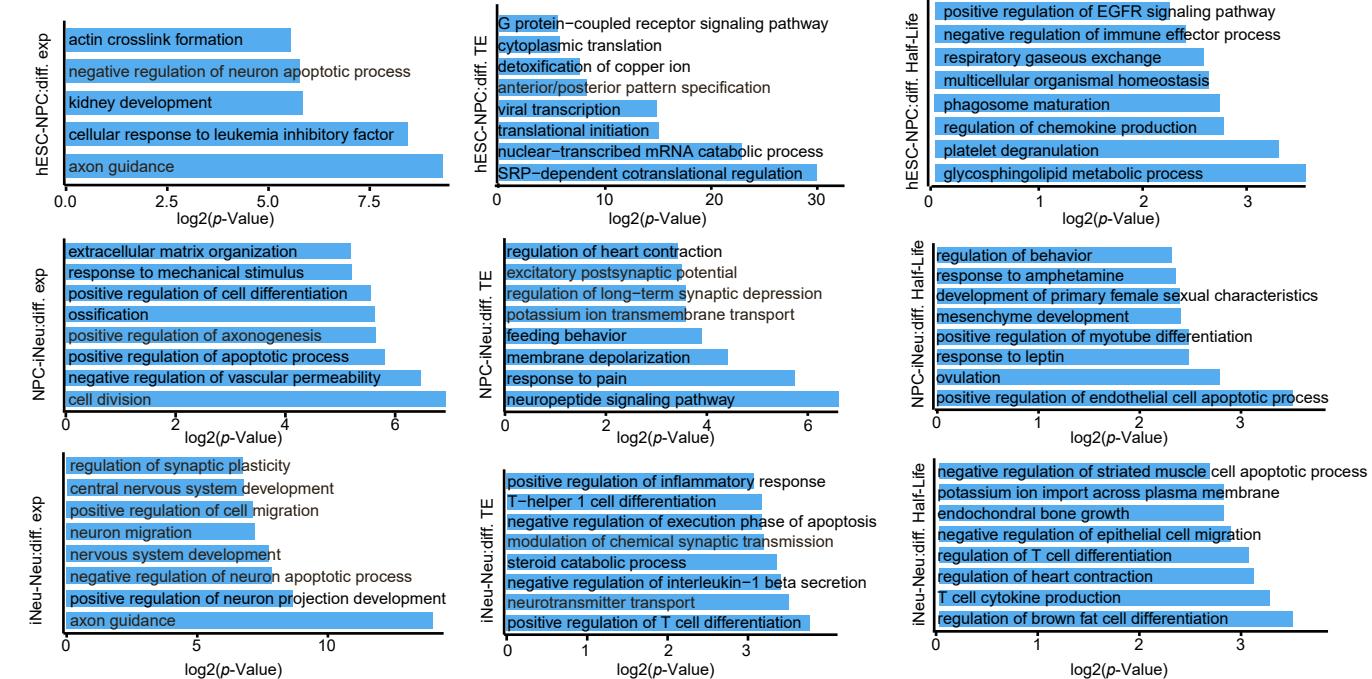
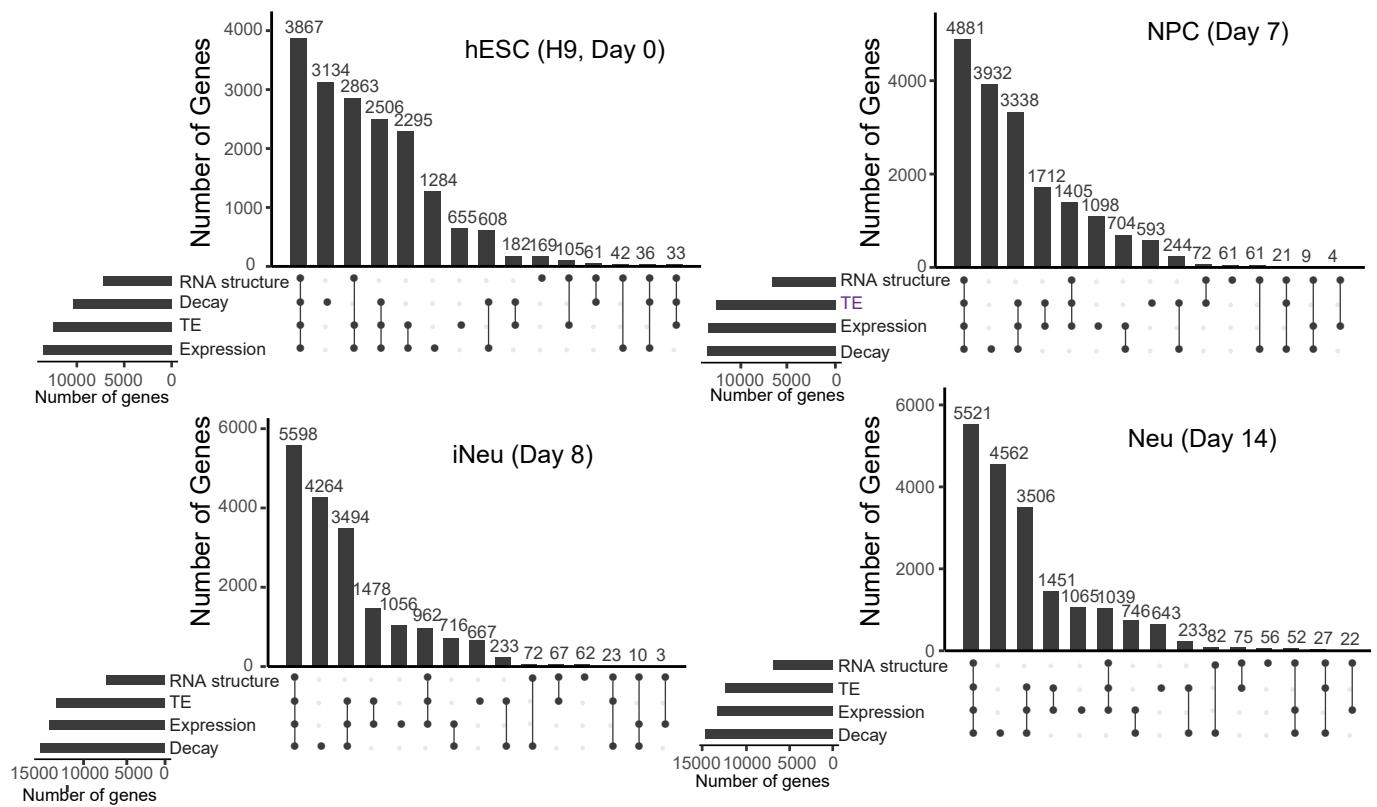


Figure S2 Summary of detected genes by different methods. Related to Figure 1

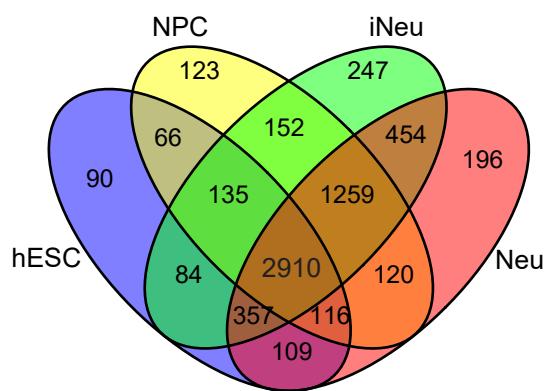
A

	No. of genes detected			
	hESC	NPC	iNeu	Neu
RNA Structure	7176	6514	6797	6874
ribosome profiling	12505	12266	12522	12495
RNA decay	10427	13190	14410	14729
RNA sequencing	13501	13151	13317	13402

B



C



D

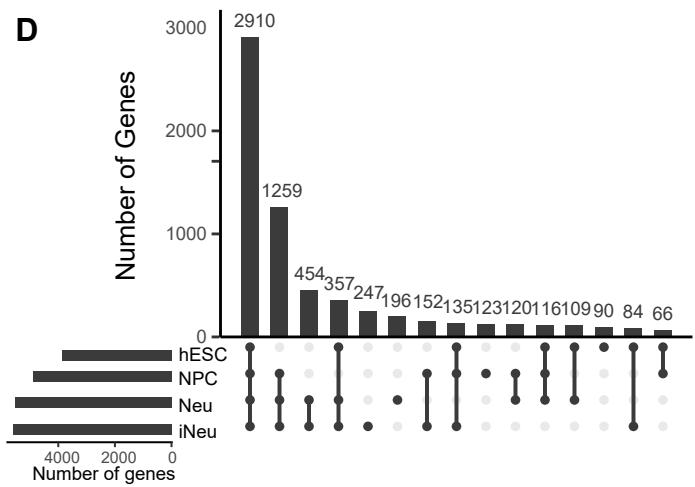
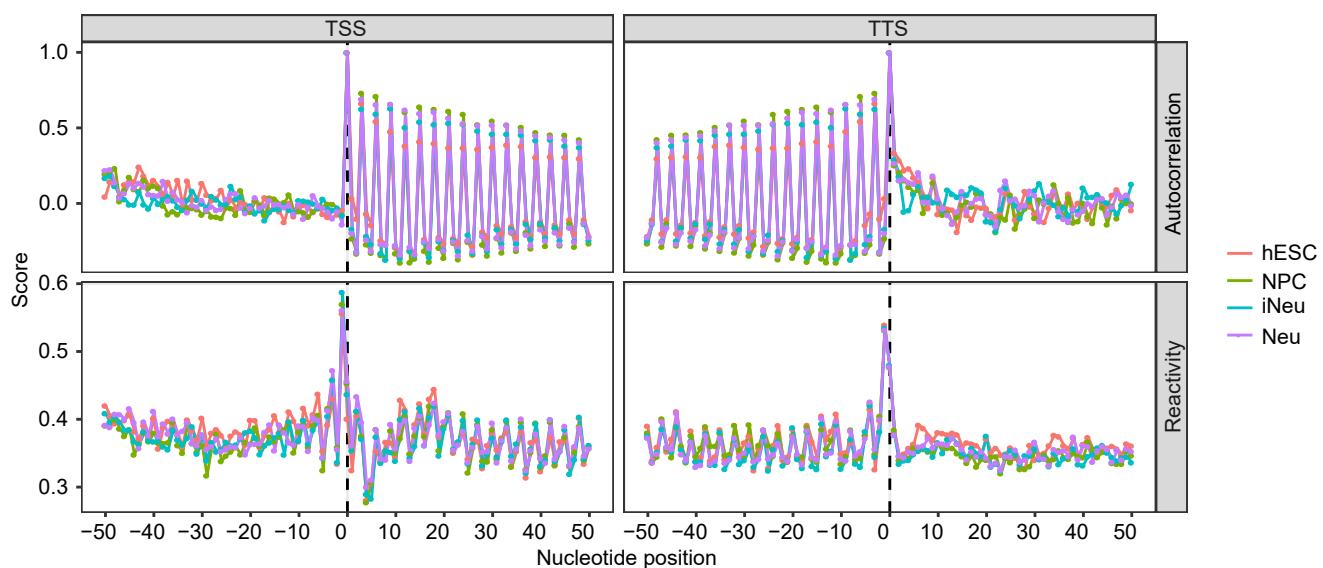


Figure S3 Global features of RNA structures using icSHAPE. Related to Figure 1

A



B

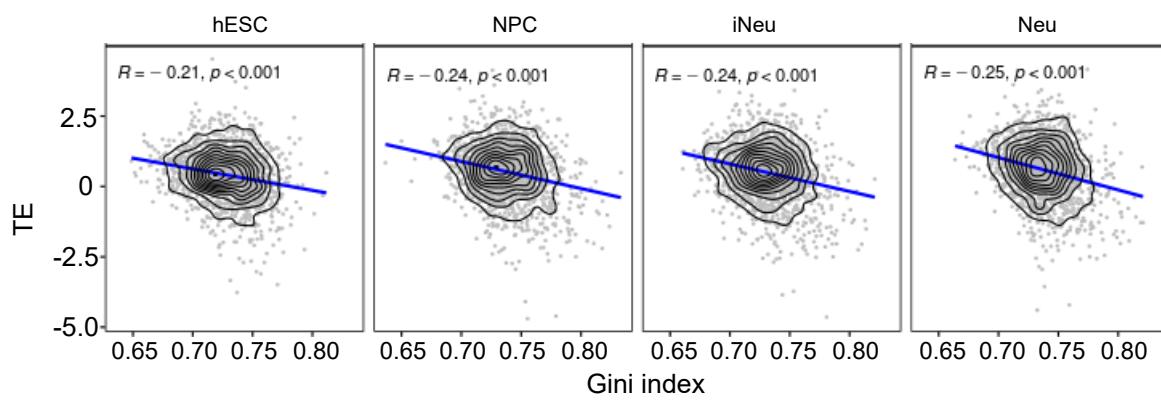


Figure S4 Genes in hESC are more accessible than in NPC. Related to Figure 1

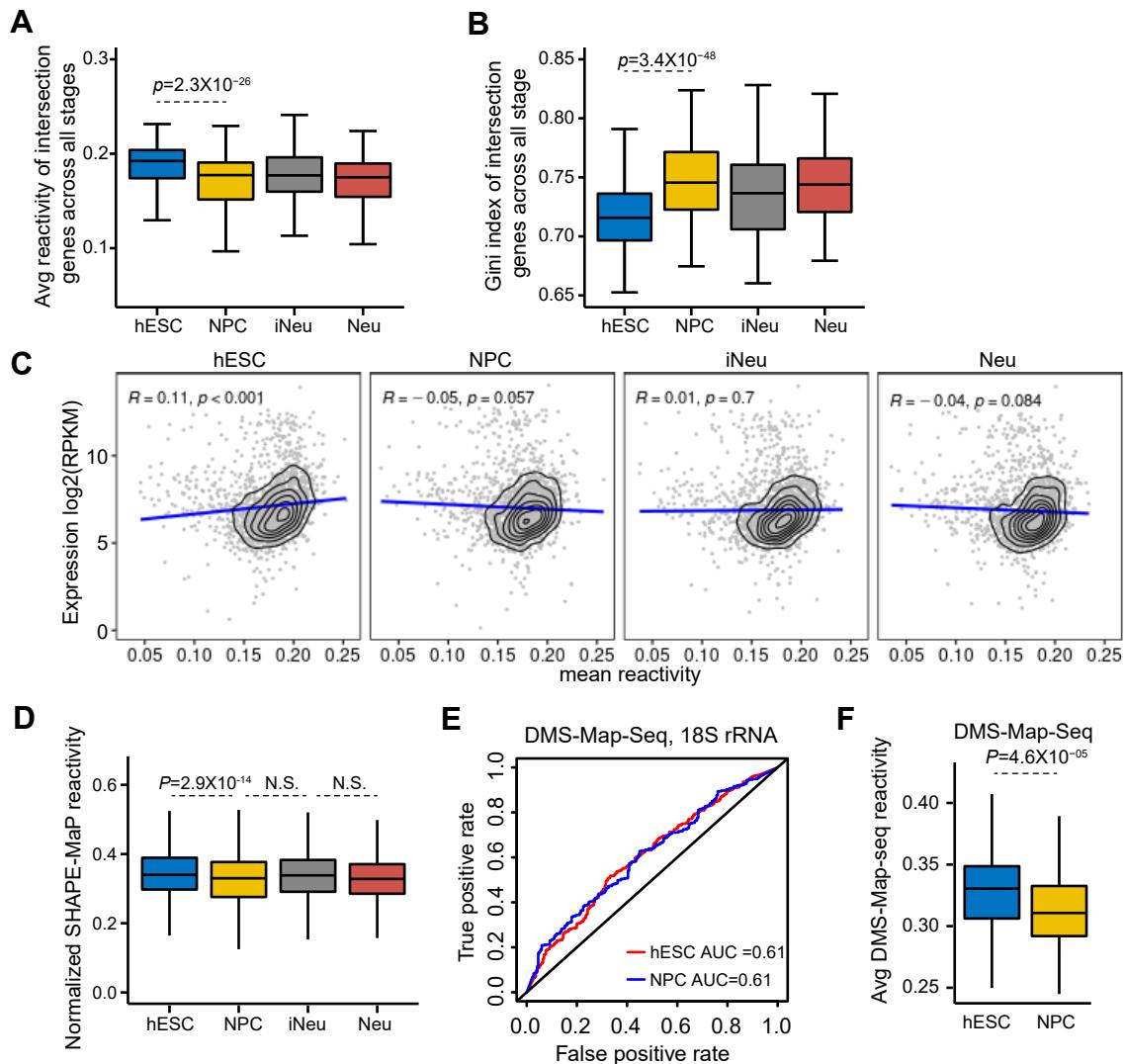
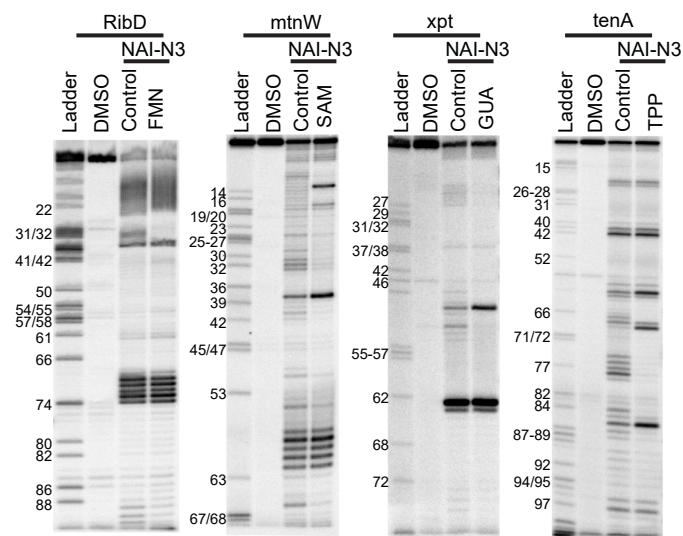


Fig.S5 Benchmarking of icSHAPE analytic methods using riboswitches and riboSNitches.
Related to Figure 1

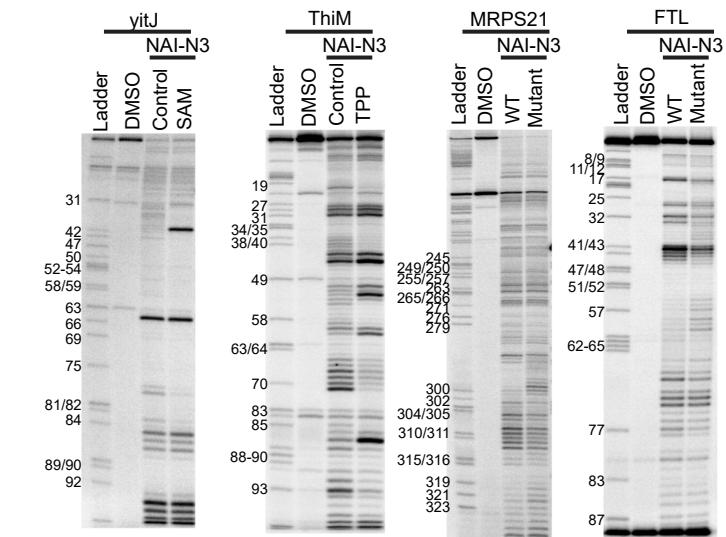
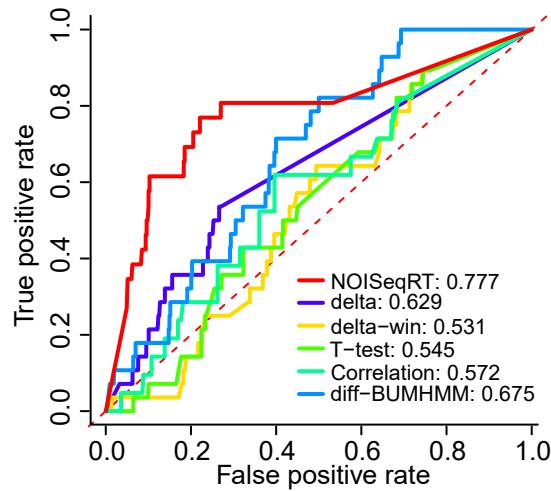
A

Target gene	Metabolite	
xpt	Control	GUA
RibD	Control	FMN
yitJ	Control	SAM
mtnW	Control	SAM
ThiM	Control	TPP
tenA	Control	TPP
Target gene	riboSNich	
MRPS21	WT	Mutation
FTL	WT	Mutation

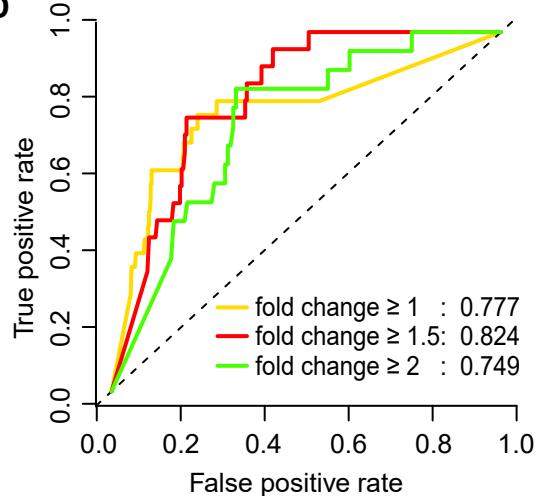
B



C



D



E

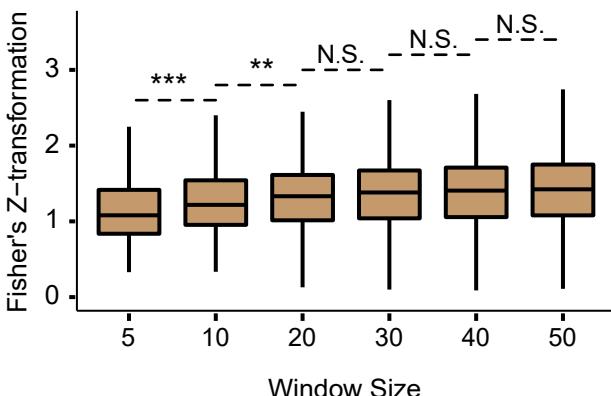
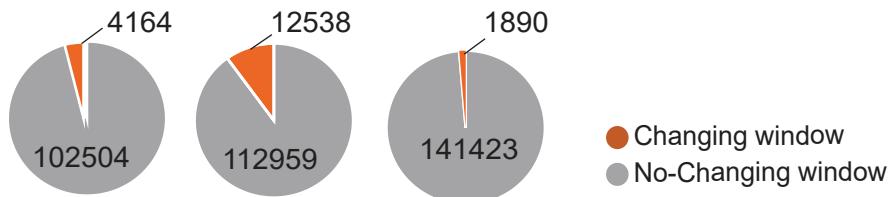
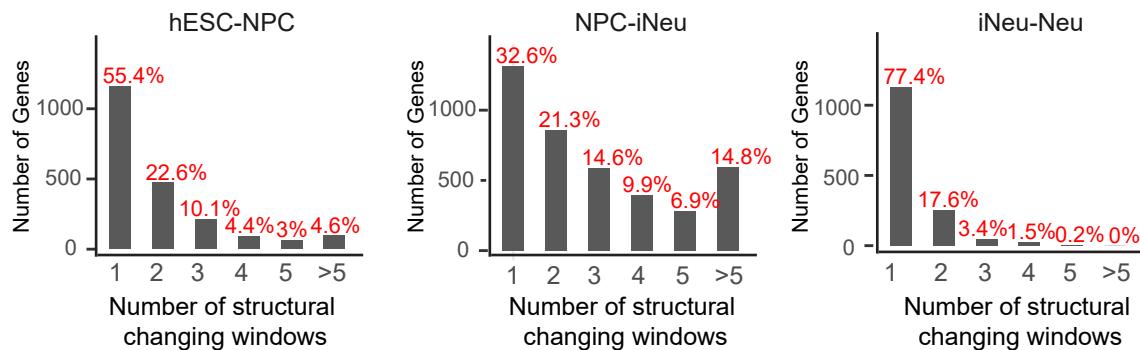


Figure S6 Summary of the significantly structural changing windows. Related to Figure 1

A



B



C

	hESC-NPC	NPC-iNeu	iNeu-Neu
20nt	3897	11669	1880
40nt	121	430	5
60nt	7	18	0
>60nt	1	0	0
Total number of windows	4164	12538	1890

D

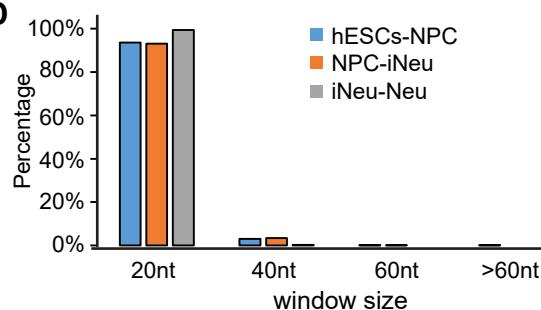


Figure S7 The association of structural changes with other methods. Related to Figure 1

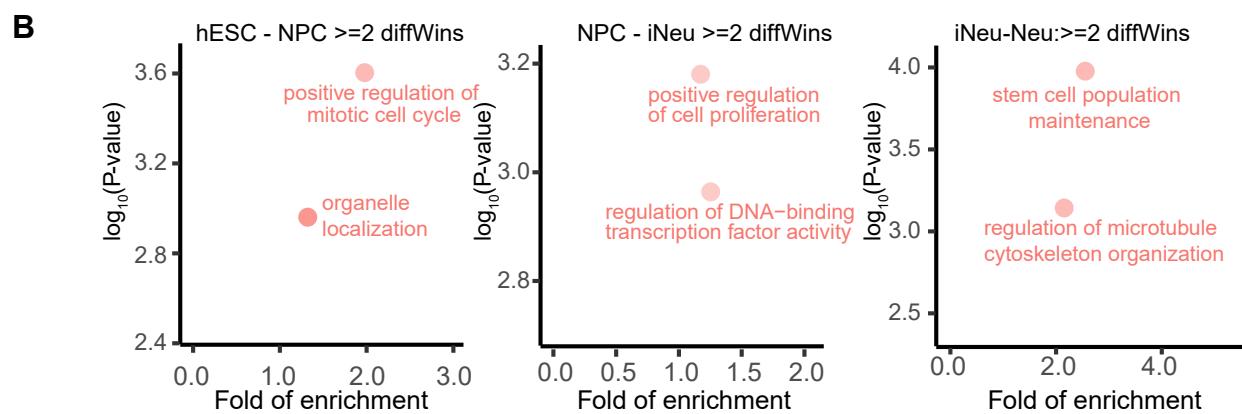
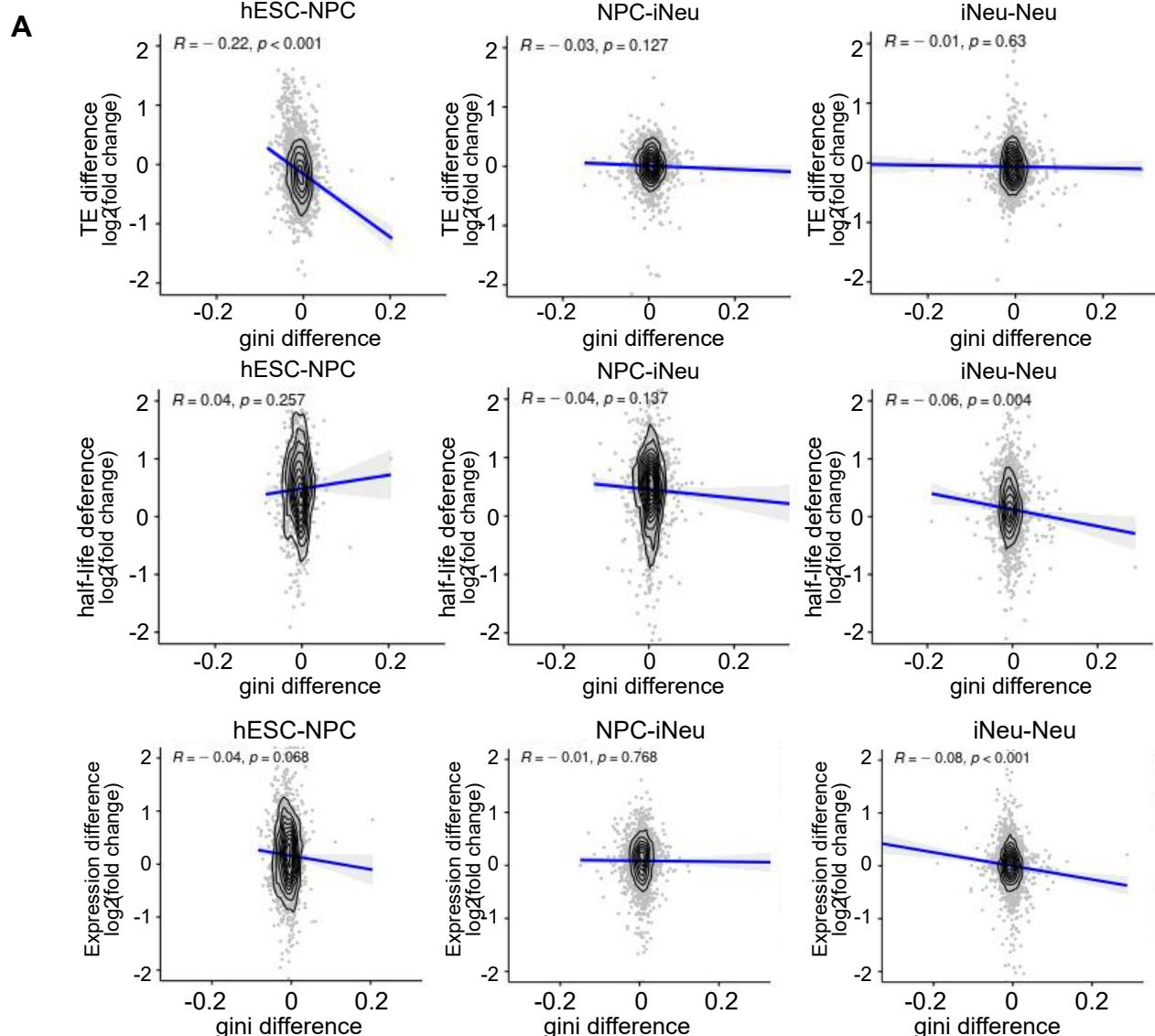
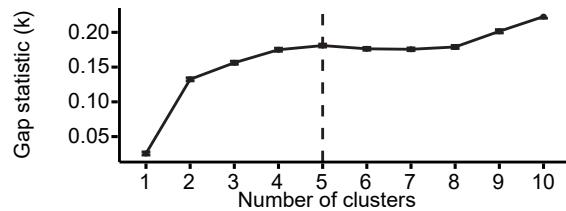
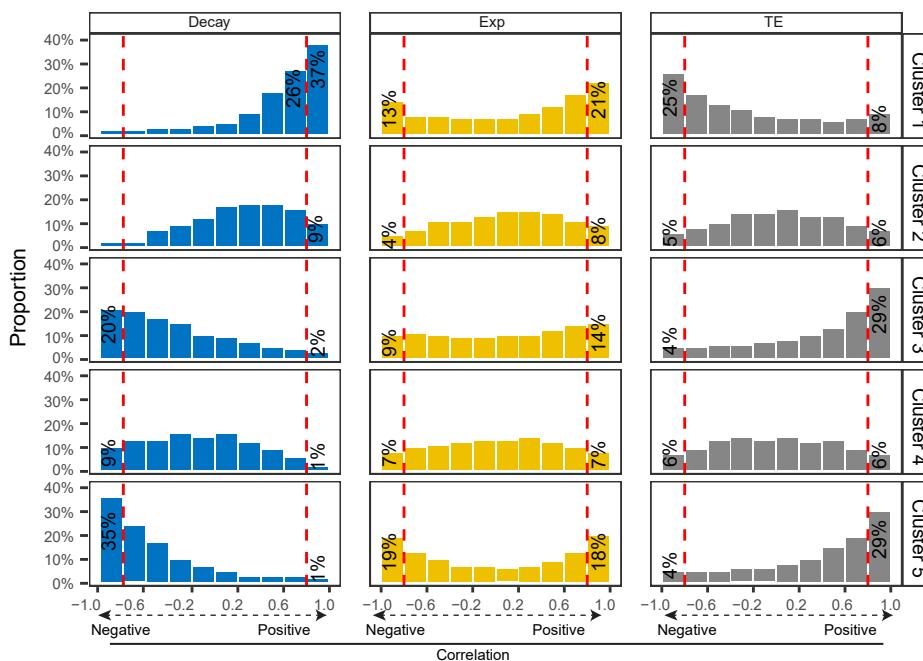


Figure S8 Clustering of significant structural changing windows. Related to Figure 2

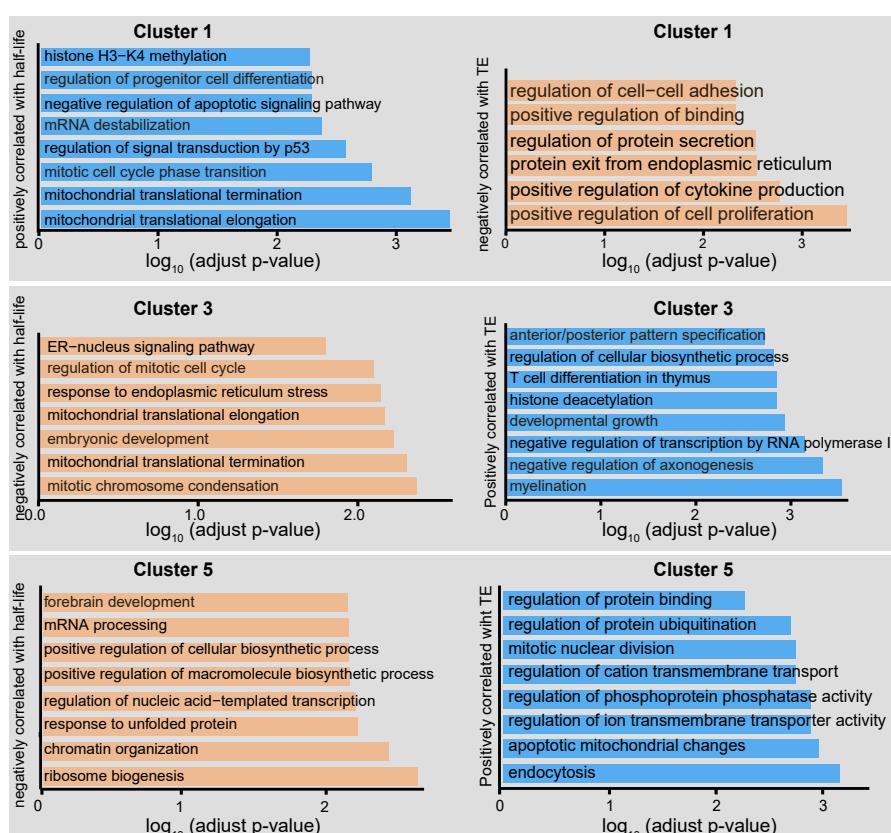
A



B



C



D

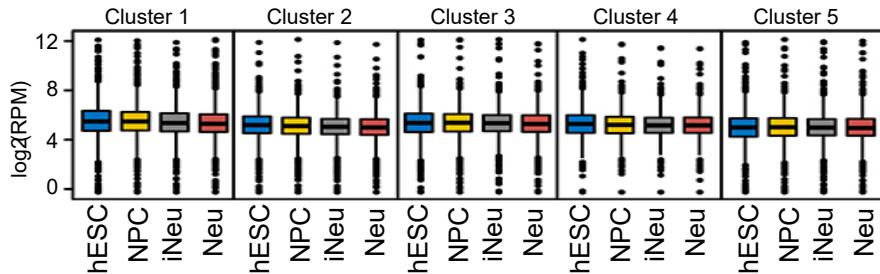


Figure S9 Overlap of reactivity changing windows with cellular regulators. Related to Figure 2

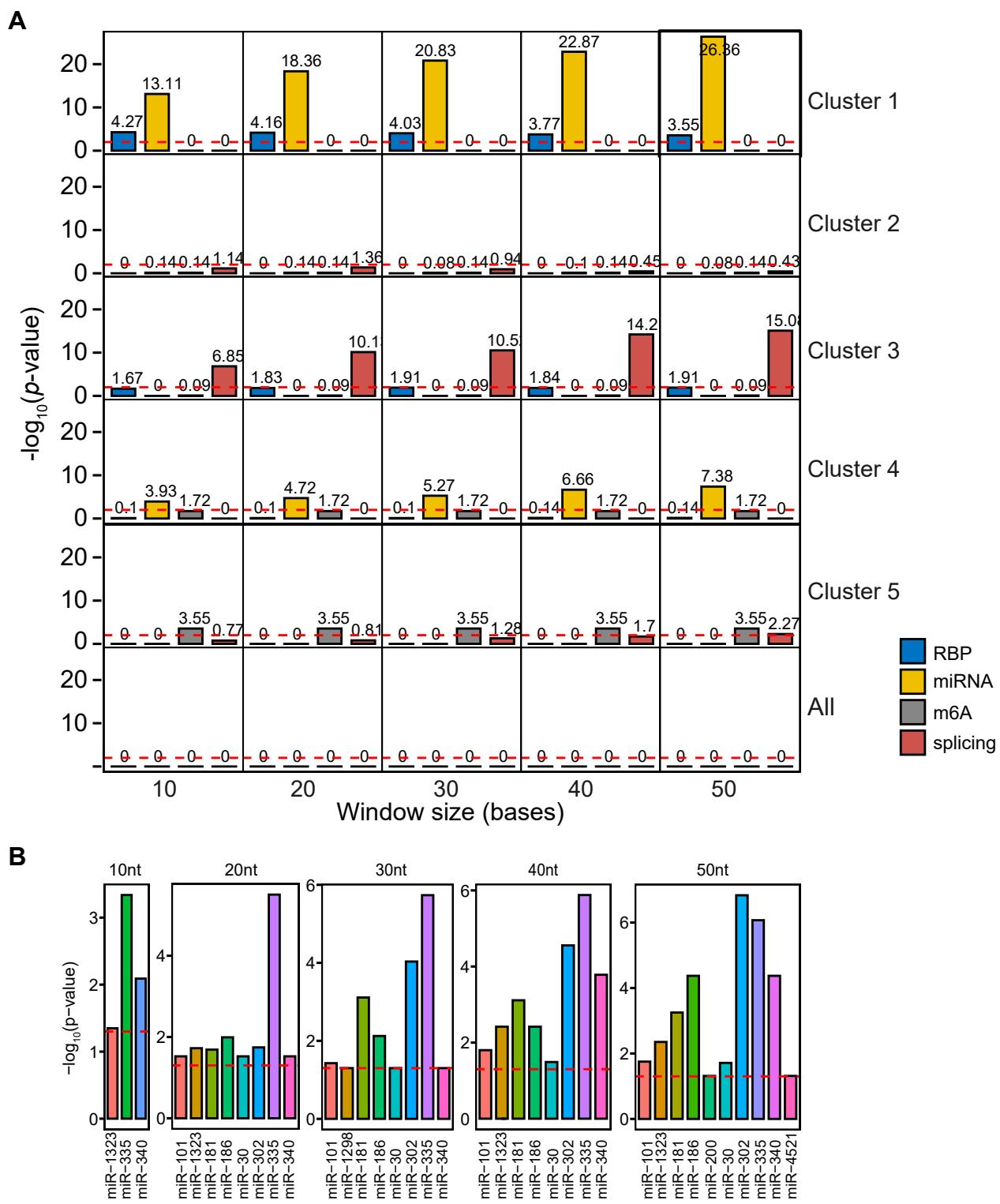


Figure S10 Enrichment of RBPs in significant structure changing windows. Related to Figure 3

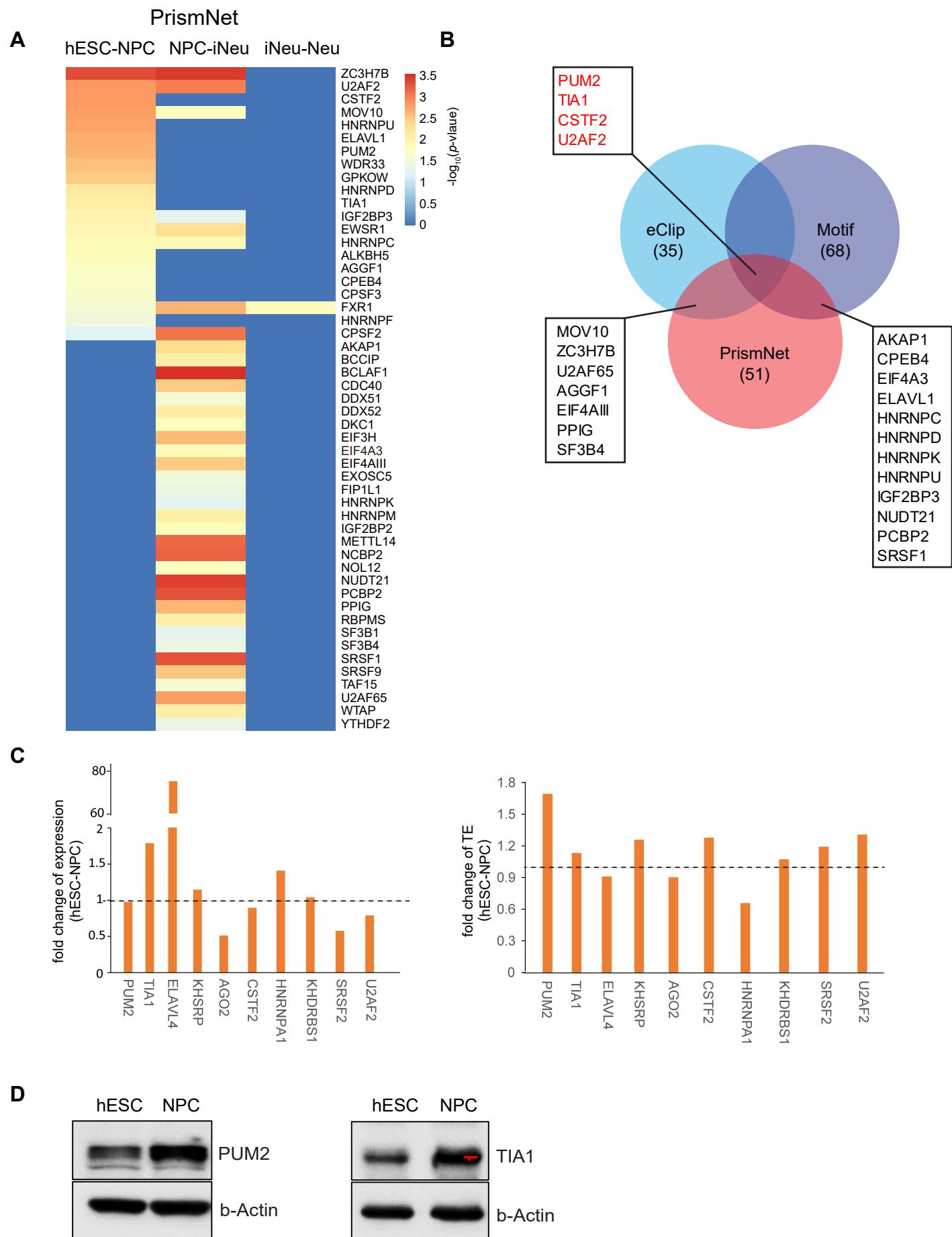


Figure S11 Overexpression of PUM2 and TIA1 in hESC. Related to Figure 3

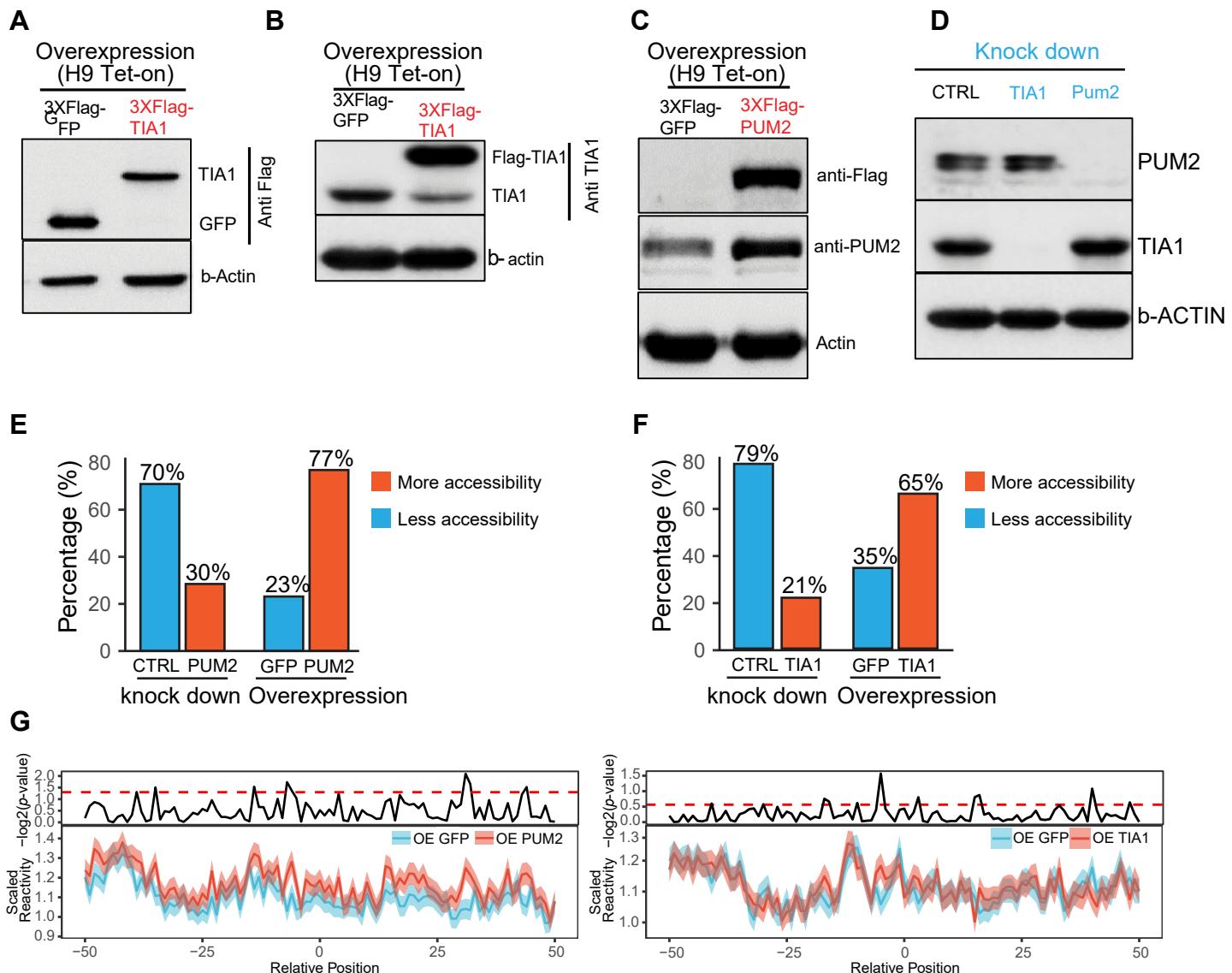


Figure S12 Number of changing windows per gene upon PUM2 over-expression Related to Figure 3

Number of changing windows per gene upon overexpression of PUM2

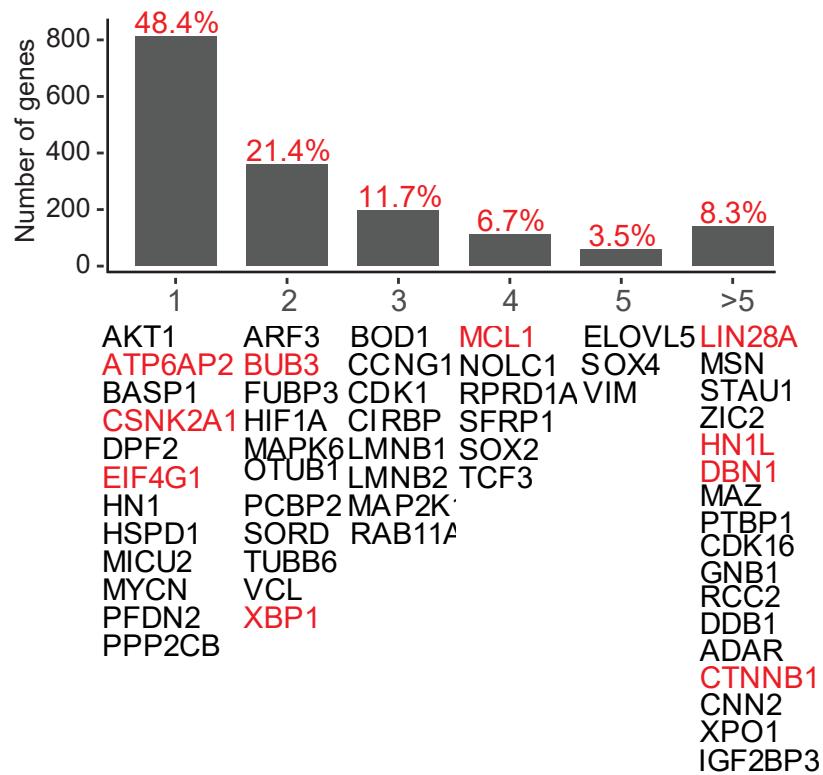


Figure S13 Regulation of RBP-RBP and RBP-miRNA interactions through structure Related to Figure 4

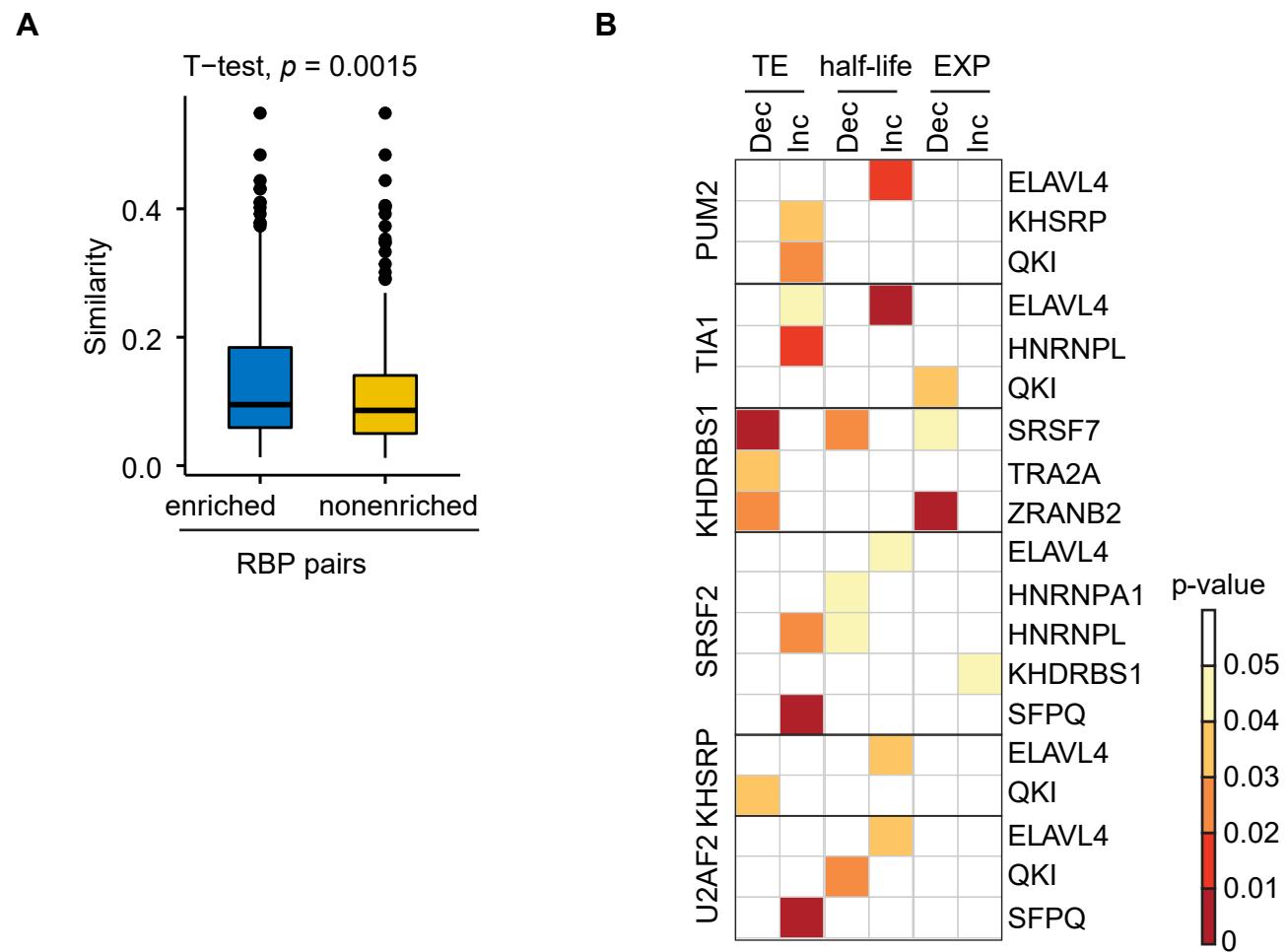
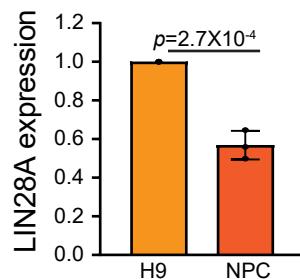
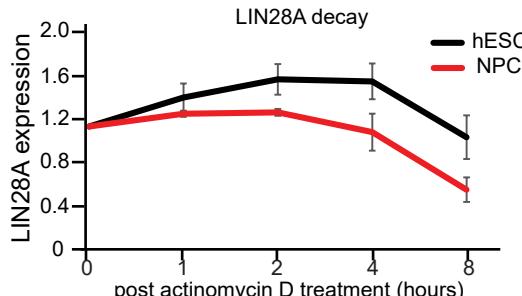


Figure S14 PUM2 and miR-30 regulated LIN28A through structure. Related to Figure 4

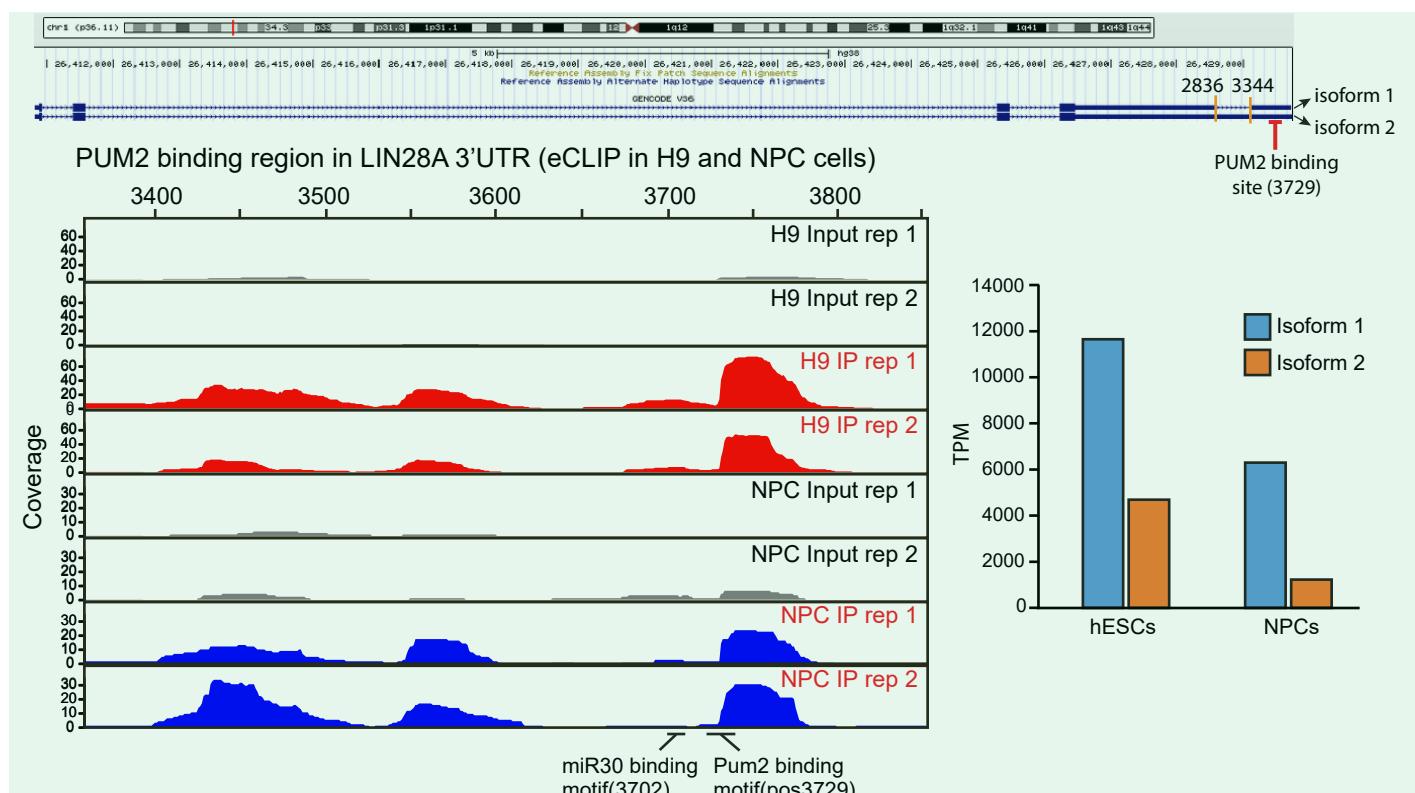
A



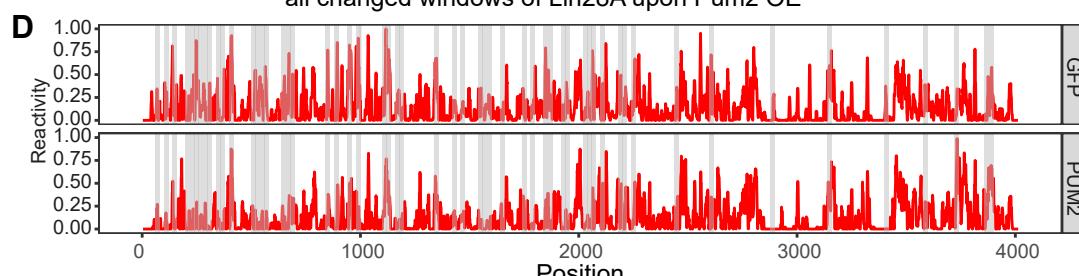
B



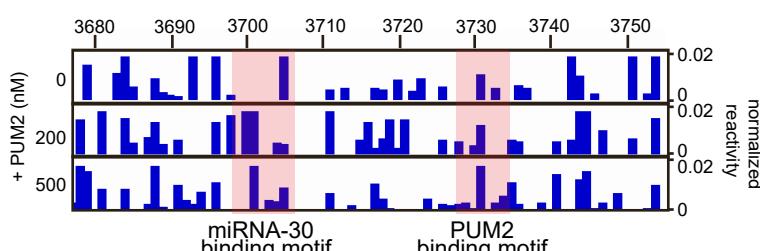
C



all changed windows of Lin28A upon Pum2 OE



E



F

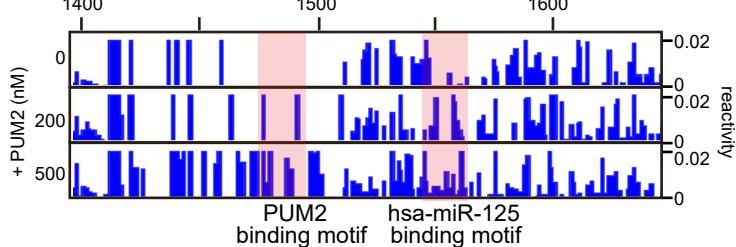


Figure S15 Impact of PUM2 binding on LIN28A. Related to Figure 4

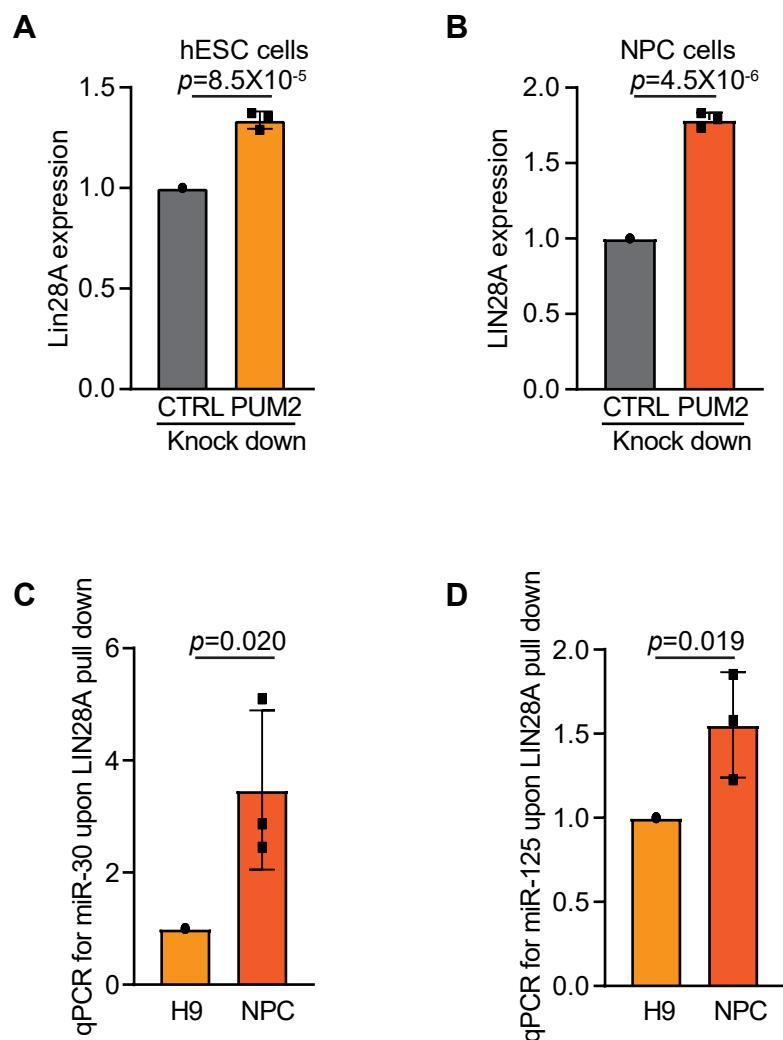


Figure S16 Foot printing of mutated RNA sequences. Related to Figure 5

