Probing enzyme-dependent pseudouridylation using direct RNA sequencing to assess neuronal epitranscriptome plasticity

Oleksandra Fanari[1,3], Sepideh Tavakoli[1,3], Stuart Akeson[1], Amr Makhamreh[1], Keqing Nian[1], Caroline A. McCormick[1], Yuchen Qiu[1], Dylan Bloch[1], Miten Jain[1, 2], Meni Wanunu[1, 2], and Sara H. Rouhanifard[1,4,*]

[1]Dept. of Bioengineering, Northeastern University, Boston, MA
[2]Dept. of Physics, Northeastern University, Boston, MA
[3]These authors contributed equally
[4]Lead contact
[*]Correspondence: s.rouhanifard@northeastern.edu

Summary

Chemical modifications in mRNAs such as pseudouridine (psi) can regulate gene expression, although our understanding of the functional impact of individual psi modifications, especially in neuronal cells, is limited. We apply nanopore direct RNA sequencing to investigate psi dynamics under cellular perturbations in SH-SY5Y neuroblastoma cells. We identified psi sites and assigned them to psi synthase enzymes using siRNA-based knockdown, finding that TRUB1 knockdown slightly upregulates psi occupancy in PUS7 sites. Next, we studied psi occupancy changes upon differentiation and lead exposure of SH-SY5Y cells. A simple steady-state enzyme-substrate model reveals a strong correlation between psi synthase and mRNA substrate levels and psi modification frequencies. Finally, comparing psi sites across all three cellular states reveals stable and plastic modifications that are either condition-dependent or condition-independent. Our psi occupancy analysis across cell states, integrated with knockdown validation of modification sites, allows robust investigations into the dynamics and plasticity of RNA modifications.

Keywords

Pseudouridine, RNA sequencing, mRNA modification, Nanopore DRS

Introduction

RNA modifications are enzyme-mediated chemical changes to the canonical structure of RNA nucleotides. Over 170 types of RNA modifications have been discovered in all types of RNAs[1] and play roles in diverse biological processes such as RNA metabolism[2], translational control[3], gene expression[4], RNA-protein interactions[4], and immune response[5]. Of the total uridines in mammalian mRNA, 0.2% - 0.6% are pseudouridine (psi) [6,7], whereas 0.15-0.6% of all adenosines are estimated to bear the m^6A modification[8,9]. Psi is an isomer of uridine[10] in which a new hydrogen bond is available to base-pair with adenosine. However, it is known to base-pair with other nucleobases in a duplex and stabilize them[11]. Substitution of U with psi can stabilize the overall RNA structure[12–14] and duplex formation, which is likely to modulate cellular interactions with proteins and other biomolecules[12,15]. Psi is prominent in transfer RNAs and ribosomal RNAs,[16–18], and modulation of psi levels in cells affects RNA metabolism[12], gene expression[12], translation[19,20], and splicing[12,15].

Various methods involving molecular amplification and next-generation sequencing have recently been utilized for psi mapping in mRNAs[6,20–24]; however, these methods all need chemical conversion steps and reverse transcription to cDNA before amplification and sequencing. We and others have recently developed algorithms to classify psi sites from nanopore direct RNA sequencing (DRS)[25–28]. Our method, Mod-*p* ID[28], compares the frequency of systematic basecalling errors at the modification site to an *in vitro* transcribed (IVT)

unmodified transcriptome[29]. Mod-*p* ID accounts for the sequence context surrounding individual psi modifications and coverage at a given site to determine a statistical probability of a modification and provides a low-limit occupancy value. However, a caveat of this and other nanopore-based methods for psi-calling is that systematic basecalling errors are insufficient to validate psi modifications, necessitating exhaustive orthogonal validation approaches such as synthetic controls[30–32] or biochemical assays[33]. A suitable route for transcriptome-wide validation is to utilize knockdown/knockout of pseudouridine synthases (PUS) and measure changes in psi occupancies[15,24] at the sites that match the motif of the respective PUS enzymes.

In this work, we aim to understand how psi occupancies in mRNAs of neurons respond to changes in cellular state. Dysregulation in genes that encode PUS enzymes has been associated with neuronal impairment.[34,35] However, the environmental factors that affect the mRNA substrates of PUS enzymes (i.e., targets that get modified) in neurons are unknown. For our model system, we chose undifferentiated SH-SY5Y neuroblastoma cells, which continuously express markers similar to immature catecholaminergic neurons while maintaining the ability to divide[36]. Upon retinoic acid differentiation, SH-SY5Y cells become morphologically similar to primary neurons; their proliferation rate is decreased (similar to mature neurons), and the activity of enzymes specific to neuronal tissues is increased[37]. First, we investigated how retinoic acid, a differentiation-inducing agent, affects the transcriptional landscape of psi modifications. Next, we focused on the impact of exposure to lead (Pb^{2+}) on the psi landscape of undifferentiated SH-SY5Y cells. Pb^{2+} is an environmental toxin that has been shown to adversely affect neuron functionality, especially during the developmental stages of the human brain.[38,39,40]

We specifically focus on differential analysis, i.e., changes in psi levels at validated psi sites, by comparing our data with SH-SY5Y cells in which we have used siRNA to knock down two key PUS enzymes, TRUB1 (motif GUUCN)[41] and PUS7 (motif UNUAR)[42]. Knockdowns were used as controls to validate that the uridine modification is a psi site. Differential analysis was then used to compare cell states at specific positions – this is a precise measurement of changes in psi occupancy at a given position. Here, we are interested in understanding the relative plasticity of psi sites in response to environmental cues for differentiation and Pb^{2+} exposure. Such plasticity has been observed with m^6A, for which modification levels increase significantly throughout brain development, which is suggested as a mechanism to achieve higher-order brain function[43].

Results

Our experimental strategy is outlined in **Figure 1**. We used DRS to perform a transcriptome-wide survey of the impact of three types of perturbations on psi modifications: 1. siRNA-based knockdown of the predominant mammalian psi synthase (PUS) enzymes, TRUB1 and PUS7; 2. retinoic acid-induced differentiation into neuron-like cells, and 3. environmental toxin exposure. We quantify the differential occupancy of psi at detected sites for unperturbed and perturbed cells across Pb^{2+} treatment and differentiation to characterize plastic and homeostatic sites.

Knockdown and nanopore sequencing of human psi synthases *TRUB1* and *PUS7*

We generated siRNA knockdown (KD) sequencing libraries for the two prevalent PUS enzymes acting on mRNAs, *TRUB1*[41] and *PUS7*[44]. We performed DRS to assign a specific enzyme to our detected psi positions (**Fig. 2a**). SH-SY5Y cells were treated with *TRUB1* siRNA, *PUS7* siRNA,

and a scrambled siRNA control for three days (**Supplementary Fig. 1**). Subsequently, qPCR analysis was conducted to evaluate the extent of knockdown achieved by *TRUB1* KD and *PUS7* KD compared to the scrambled siRNA control. The results indicated substantial mRNA KD for *TRUB1* (82.3% ± 5.6%) and *PUS7* (57.1% ± 5.6%; **Fig. 2b**). Next, we stained the KD and scrambled control cells with anti-Pus7 and anti-Trub1 antibodies to evaluate protein expression following the knockdown (**Supplementary Fig. 1**). We observed a substantial fold-decrease in mean fluorescence intensity for Pus7 (0.56-fold ± 0.0275) as well as Trub1 (0.45-fold ± 0.0321) KD cells as compared to the scrambled siRNA control. We extracted polyA+ RNA and prepared libraries for each sample, including a scrambled siRNA control library for comparison (see ***Methods*** section for details).

RNA expression profiling for *TRUB1* and *PUS7* KD of SH-SY5Y cells

To assess the effects of *TRUB1* and *PUS7* KD on other psi-synthases, we evaluated expression levels for 13 different PUS enzymes from *TRUB1* KD, *PUS7* KD, and scrambled siRNA control libraries (**Fig. 2c**). For the *PUS7* KD library we observed a 21.5 fold ± 0.07 (p = 0.0004) decrease in mRNA (TPMs) compared to the scrambled control. For the *TRUB1* KD library, we observed an 8.9 fold ± 0.11 (p = 0.00005) decrease in *TRUB1* mRNA (TPMs) compared to the scrambled control. We also observed significant alterations in other PUS enzymes in response to *PUS7* and *TRUB1* KD. For the *PUS7* KD library compared to the scrambled control library, we observed a 1.8 fold ± 0.35 (p = 0.006) reduction in *PUSL1* mRNA, a 3.4 fold ± 0.14 (p = 0.0002) reduction in *RPUSD1* mRNA, a 1.6 fold ± 0.50 (p = 0.03) decrease in *TRUB1* mRNA, a 2.1 fold ± 0.14 reduction (p = 0.00001) in *DKC1* mRNA, a 2.6 ± 0.35 fold (p = 0.0007) reduction in *RPUSD4* mRNA, and a 3.3 fold ± 0.43 (p = 0.0001) reduction in RPUSD3 mRNA. For the *TRUB1* KD library compared to the scrambled control library, we observed a 1.9 fold ± 0.39 (p = 0.007) reduction in *RPUSD1* mRNA.

Transcriptome-wide mapping using Mod-*p* ID for *TRUB1* and *PUS7* KD libraries

We generated a paired, unmodified transcriptome for SH-SY5Y cells to identify putative psi positions and enriched this data set using IVT libraries from 5 other human cell lines. Using the pan-human IVT[29] as a baseline, we compared the U-to-C basecalling error between the KD and scrambled siRNA control libraries. We identified putative psi sites based on significant differences in U-to-C basecalling error (p < 0.001) between the unmodified and modified (KD and scrambled (control) libraries using Mod-p ID[28] (**Supplementary Table 1**).

First, we selected positions for which we had measured $p < 0.001$ in the scrambled control, and the position was present in both the scrambled and KD libraries (**Supplementary Table 1**). To compare individual points between data sets, we needed to give more weight to samples with higher coverage or a more robust U-to-C error signature indicative of a psi site. To this end, we computed the log marginal likelihood ratio of DRS and IVT modeled by separate distributions versus a single combined distribution (See **Methods**). This calculation modeled each state as a beta-binomial distribution and used Jeffrey's noninformative prior. We used this metric as the signal-to-noise ratio (SNR) (**Fig. 2d; Supplementary Table 2**). We made a cutoff at SNR = 1 and observed that 10636 sites exceeded this threshold. Of these sites, 232 fell within a TRUB1 motif sequence (GUUCN). For this step, we selected only the sites that showed a decrease in U-to-C error in response to knockdown.

We defined sites as knocked down if they met the SNR = 1 criteria and also met a specific read count threshold (>10) and a specified difference between U-to-C error in the KD sample and the control (**Fig. 2e; Supplementary Table 3**). For positions with more reads (>30 reads), we

set the cutoff at a 15% difference between the samples. For positions with <30 reads, we set the cutoff at a 30% difference between the samples. We observed 74 sites that met these criteria and defined these as targets of TRUB1. We derived a sequencing logo to identify the motifs that fell within this criteria (**Fig. 2e**).

We performed the same analysis for PUS7 KD sites and observed that 6,070 sites exceeded the SNR = 1 threshold, of which 120 sites fell within a PUS7 motif sequence (**Fig. 2f-g; Supplementary Table 3**). We observed nine sites that met the read and mismatch error cutoffs and defined these as targets of PUS7. We derived a sequencing logo to identify the motifs that fell within this criteria (**Fig. 2g**).

Modeling the cooperative effects of TRUB1 and PUS7 KD on mRNA substrates
We observed a positive population shift of positions with U-to-C error within PUS7 motifs (i.e., higher expression of psi within PUS7 motifs as a function of TRUB1 KD (**Fig. 2d**). To explore this observation further, we first assessed the changes in PUS7 mRNA levels upon TRUB1 KD and observed a -1.35 fold ± 1.2 (p = 0.3) reduction in *PUS7* mRNA compared to the scrambled control (**Fig. 2c**). To test the hypothesis that TRUB1 KD has a global effect on psi levels within PUS7 motifs, we compared the differences in positional occupancy for the TRUB1 KD with the PUS7 KD for the same positions to a random distribution model of positions that were constrained on the parameters of the observed sequencing data. We found that the data is inconsistent with a random model (Mann–Whitney–Wilcoxon test, $p < 0.001$) in the region where the position shows a decreased positional occupancy for both conditions.

Steady-state enzyme-substrate model of pseudouridylation
Since DRS provides access to both transcript levels and relative occupancy of site-specific RNA modifications, we reason that changes in levels of a PUS enzyme specific to a U-site in a transcript would impact psi occupancies at that site. Since assigning psi levels based on U-to-C mismatch errors is semi-quantitative, we can probe differential psi levels as cells change from conditions A to B. We have observed a subset of mRNAs (i.e., validated substrates for their corresponding PUS enzyme). While enzyme-substrate action requires substantial correction when the enzyme and substrate are within small compartments (<0.1 μm) due to noise from a small number of statistics[45] in the broader compartments of the cell, we tested the following first-order steady-state hypothesis:

$$[\mathrm{mRNA_U}]_{\mathrm{Cond\ A}}[\mathrm{mRNA_{PUS}}]_{\mathrm{Cond\ A}} \gtrless [\mathrm{mRNA_U}]_{\mathrm{Cond\ B}}[\mathrm{mRNA_{PUS}}]_{\mathrm{Cond\ B}} \rightarrow \mathrm{mRNA_{\psi A}} \gtrless \mathrm{mRNA_{\psi B}},$$

where $[\mathrm{mRNA_U}]_{\mathrm{Cond\ x}}$ is the number of substrate transcripts in Condition X; $[\mathrm{mRNA_{PUS}}]_{\mathrm{Cond\ x}}$ is the value in transcripts per millions (TPMs) for a transcript that encodes for a PUS enzyme in Condition X (used as a proxy for PUS levels in that cell line); $\mathrm{mRNA_{\psi X}}$ is the number of ψ transcripts in condition X (**Fig. 2i**). We use $\mathrm{mRNA_{PUS}}$ counts as a proxy for enzyme expression. Here, we assume that the encoded protein levels will increase or decrease with mRNA levels because 1. the cell types are similar (SH-SY5Y cells, perturbed and unperturbed), and 2. there is active protein degradation[46]. We applied our data (i.e., sequencing reads) to this model and compared it to the difference in psi occupancy (as determined by U-to-C error at a position of validated psi) and found a positive correlation for TRUB1 sites when using the TRUB1 enzyme levels (R^2 = 0.76; **Fig. 2j**). As a control, we compared PUS7 sites to TRUB1 mRNA levels and found a decreased correlation (R^2 = 0.56; **Fig. 2j**), as expected for sites not meant to interact with the TRUB1 enzyme.

Retinoic acid-induced differentiation of SH-SY5Y cells leads to a change in cell state
We differentiated SH-SY5Y cells into neuron-like cells by supplementing them with retinoic acid[47], according to Kovalevich et al.[48] (**Fig. 3a**). We compared the cellular morphology to assess the change in cell state from undifferentiated to differentiated. We observed the elongation and branching of neurite-like processes from the differentiated cells (**Fig. 3b**). We also assessed gene expression differences to define a cellular state change. Total RNA was extracted, and poly-A selection was performed on differentiated and untreated SH-SY5Y cells. Subsequently, nanopore direct RNA libraries were prepared, sequenced, and aligned to a reference. Next, we used our acquired Nanopore DRS data sets to perform a comparative analysis of mRNA expression using DESeq2. We observed differential mRNA expression when comparing the two groups, supporting a change in cell state (**Fig 3c**). Importantly, we observed the expected differential mRNA expression of known differentiation markers[49–51]*. CRABP2*, *RARB*, *RGS2*, *RET*, and *DKK2* exhibited upregulation and *ISOC1*, *MYC*, *SPRY2*, and *ASCL1* displayed decreased RNA expression in differentiated SH-SY5Y cells compared to the untreated counterparts (**Supplementary Fig. 2)**. To assess the effects of retinoic-acid-mediated differentiation on psi machinery, we evaluated expression levels for 13 different psi-synthases from untreated and differentiated libraries (**Fig. 3d**). We observed a significant difference in the RNA expression levels for the psi synthases PUS7L in response to differentiation ($p = 0.03$).

Transcriptome-wide mapping of psi-modifications before and after differentiation
We identified pseudouridine's position in differentiated and untreated SH-SY5Y cells to explore psi-modification changes. We generated a paired, unmodified transcriptome for untreated SH-SY5Y cells to identify putative psi positions and compared the U-to-C basecalling error between the untreated and differentiated libraries. We applied Mod-*p* ID to identify putative psi sites based on significant differences in U-to-C basecalling error ($p < 0.001$) in the untreated and differentiated libraries compared to the IVT control library. We selected positions that were represented in both the untreated and differentiated libraries and defined the SNR for each site by calculating the number of standard deviations separating the mismatch error for sites observed in both the untreated and differentiated libraries from that of the IVT control (**Fig. 3e**). We made a cutoff at SNR = 1. We observed that 1786 sites exceeded this threshold (**Supplementary Table 4**).

We defined sites as changed in response to a change in the cellular state if they met three criteria: SNR ≥ 1, > ten reads in both samples and a ≥ 5% difference in U-to-C error between the untreated and the differentiated libraries (**Fig. 3f, Supplementary Table 4**). We identified several of these sites as psi sites based on the knockdown of their corresponding psi synthase (**Supplementary Table 5**). *FBXO5* (chr6:152975604), HECTD1 (chr14:31101239), NKAIN1 (chr1:31181249), CCDC22 (chrX:49249707), *EFEMP2* (chr11:65866534), *IDI1* (chr10:1044099; Trub1 substrates), and NES (chr11:56669130), *MCM5* (chr22:35424407), *RPL22* (chr1:6185970), and *RHBDD2* (chr7:75888787; PUS7 substrates) showed decreased positional occupancy in response to differentiation. HPS4 (chr22:26452838), *TAF9B* (chrX:78129969), *CDC6* (chr17:40295391), *TTYH3* (chr7:2663889)*, RNF167* (chr17:4944900)*, SCP2* (chr1:53027999), *ZNF317* (chr19:9161020), *THY1* (chr11:119418679; Trub1 substrates), each showed increased positional occupancy in response to differentiation. The most substantial change was for *ZNF317* (chr19:9161020), which went from a 15% U-to-C error in the untreated

samples to a 52% in the differentiated sample, and *THY1* (chr11:119418679), which went from a 34% U-to-C error in the untreated samples to a 61% U-to-C error in the differentiated sample.

Interestingly, we find that two knockdown-validated pseudouridine sites that show plasticity after cellular differentiation are involved in neuronal processes: *THY1* encodes a cell surface glycoprotein involved in cell adhesion processes and modulates neurite outgrowth[52]; *NKAIN1* encodes a protein involved in sodium-potassium transport in the brain[53]. We exported the sequencing logo for positions categorized into three groups: those with higher U-to-C base-calling errors during differentiation, positions showing no difference between differentiated and untreated samples, and positions with lower U-to-C base-calling errors during differentiation (**Fig. 3f**). We observed that for the positions in which the positional occupancy decreased after differentiation, the +1 nucleotide neighboring the psi position was frequently uridine. We found 26 positions with differential expression of psi that were assigned to a specific PUS enzyme using our KD experiments (**Fig. 2**) and cross-validated these using orthogonal controls (**Fig. 3g**).

We calculated TPMs for each TRUB1 and PUS7 target with differences in psi levels to explore differences in mRNA expression for transcripts that harbor a psi. We found that only *IDI1* (chr10:1044099; 15.3 TPM in the untreated sample and 58.9 TPM in the differentiated sample) showed a significant difference in mRNA expression between the two conditions ($p < 0.001$; **Fig. 3h**). The other mRNA levels remained unchanged while having differences in psi occupancy. Furthermore, we examined the protein expression levels in cellular compartments for the two dominant PUS enzymes for humans, PUS7 and TRUB1, in untreated and differentiated SH-SY5Y cells. Our analysis revealed no significant differences in the subcellular distribution of these two PUS enzymes (**Supplementary Fig. 3**). For these datasets, we also tested the steady-state enzyme: substrate model (**Figure 2i,j**) as we have done for the PUS knockdowns. Our results (see **Fig. 3i**) show a positive correlation between the changes in U-C mismatches with the product of PUS enzyme and mRNA site concentrations, which further supports that TRUB1-mediated pseudouridylation of putative TRUB1 sites with the motif GUUCN follows a simple enzyme: substrate model ($R^2 = 0.73$), whereas for PUS7-mediated pseudouridylation of UNUAR motifs, it is less clear that the relationship follows a simple steady-state model ($R^2 = 0.33$).

SH-SY5Y cell state changes in response to lead (Pb^{2+}) exposure

We introduced the neurotoxicant, Pb^{2+}, to SH-SY5Y cells as an alternative change in cellular state. The cells were cultured in growth media for 24 hours before changing to growth media supplemented with 50 μM Pb^{2+} for six days (**Fig. 4a**). Previous studies have shown that 5 μM is close to the Pb^{2+} levels in human blood that can cause encephalopathy in children[54]; however, we chose to use a relatively high Pb^{2+} concentration (50 μM) because the *in vitro* tolerance for cytotoxicity is higher than *in vivo*[55]. RNA was extracted from the untreated and Pb^{2+}-exposed cells, and each sample underwent library preparation and subsequent analysis. As anticipated, the cellular morphology remained unchanged following exposure to Pb^{2+} (**Fig. 4b**).

We conducted gene expression analysis to compare untreated and Pb^{2+}-exposed libraries to demonstrate a change in cellular state (**Fig. 4c**). Gene ontology (GO) analysis of the genes that exhibit significant downregulation reveals an association with the TGF-beta signaling pathway (**Supplementary Table 9**). The upregulated gene *PTN* plays a role in Leukocyte Chemotaxis Involved In Inflammatory Response (GO:0002232) and Dendrite Arborization (GO:0140059); together with the gene ID2 PTN also shows an association with Positive Regulation Of Glial

Cell Differentiation (GO:0045687). The upregulated genes CUX2 and PLXNA2 are associated with Positive regulation of neurogenesis (GO:0050769). The genes *CHGA* and *RAMP1* exhibit downregulation and are associated with the pathways *ADORA2B* Mediated Anti-Inflammatory Cytokine Production (R-HSA-9660821) and Anti-inflammatory Response Favoring Leishmania Infection (R-HSA-9662851; **Supplementary Table 9**). The downregulated gene *SPOCK1* is associated with Nervous System Development (GO:0007399), Neurogenesis (GO:0022008), and Central Nervous System Neuron Differentiation (GO:0021953). To assess the effects of Pb^{2+} exposure on other psi-synthases, we evaluated expression levels for 13 different psi-synthases from Pb^{2+} treated and untreated SH-SY5Y cell libraries, and we found no significant differences in expression levels (**Fig. 4d**).

Transcriptome-wide mapping of psi-modifications following Pb^{2+} exposure

To explore changes in psi modification between untreated and Pb^{2+} treated cells, we generated a paired, unmodified transcriptome for untreated SH-SY5Y cells to identify putative psi positions. We applied Mod-*p* ID to identify putative psi sites based on significant differences in U-to-C basecalling error ($p < 0.001$) in the untreated and Pb^{2+} treated libraries compared to the IVT control library. We selected positions found in both the untreated and Pb^{2+} treated libraries and defined the SNR for each site by calculating the number of standard deviations separating the mismatch error for sites observed in both the untreated and differentiated libraries from that of the IVT control (**Fig. 4e**). We made a cutoff at SNR = 1. We observed that 946 sites exceeded this threshold.

We defined sites as changed in response to Pb^{2+} treatment if they met the same three criteria described for differentiated cells: SNR ≥ 1, >10 reads in both samples and ≥ a 5% difference in U-to-C error between the untreated and the differentiated libraries (**Fig. 4f, Supplementary Table 6**). We identified several of these sites as psi sites based on the knockdown of their corresponding psi synthase (**Supplementary Table 7**). *SH2B1* (chr16:28872589), *HECTD1* (chr14:31101239), *FBXO5* (chr6:152975604), *IDI1* (chr10:1044099), *SLC30A5* (chr5:69123368), *AK2* (chr1:33014553), *PMPCB* (chr7:103311810; Trub1 substrates), and *CKAP5* (chr11:46743829), *HDGFL2* (chr19:4502197), *RPL22* (chr1:6185970), and *MCM5* (chr22:35424407; PUS7 substrates) showed decreased positional occupancy in response to Pb^{2+} treatment. *RAD21* (chr8:116863212), *SPPL3* (chr12:120764051), *ERH* (chr14:69380270), *SCP2* (chr1:53027999), *PSMB2 (*chr1:35603333), *PABPC4* (chr1:39565149), CDC6 (chr17:40295391), *AMFR* (chr16:56362360), *AMOTL1* (chr11:94874364), *INSIG2* (chr2:118109417), *ST13* (chr22:40824900), *SCAF1* (chr19:49654781), *PHF13* (chr1:6623607), *THY1* (chr11:119418679), *ASDHPPT* (chr11:106090565), *PSMB2* (chr1:35603333), and *GTF3C2* (chr2:27326828; TRUB1 substrates) showed increased positional occupancy in response to Pb^{2+} treatment. The biggest change was for *PHF13* (chr1:6623607), which increased from 26% U-to-C error in the untreated library to 56% U-to-C error following Pb^{2+} treatment. We exported the sequencing logo for positions categorized into three groups: those with higher U-to-C base-calling errors following Pb^{2+} treatment, positions showing no difference between Pb^{2+} treated and untreated samples, and positions with lower U-to-C base-calling errors following Pb^{2+} treatment (**Fig. 4f**). We found that psi sites with increased positional occupancy following Pb^{2+} treatment tend to be flanked by two uridines, while those with decreased positional occupancy following Pb^{2+} treatment have uridine in the N+1 position. We found 28 positions with differential expression of psi that were assigned to a specific PUS

enzyme using our KD experiments (**Fig. 1**) and cross-validated these using orthogonal controls (**Fig. 4g**).

To test for differences in mRNA expression for the psi targets that change in response to Pb^{2+} treatment, we calculated transcripts per million for each TRUB1 and PUS7 target with differences in psi levels. We found that only *ERH* (chr14:69380270; 158 TPM in the untreated sample and 207 TPM in the Pb^{2+} treated sample) showed a significant difference in mRNA expression between the two conditions ($p < 0.05$ (**Fig. 4h**). We examined the protein expression levels in cellular compartments for the two dominant PUS enzymes for humans, PUS7 and TRUB1, in untreated and Pb^{2+} treated SH-SY5Y cells and found no significant differences in the subcellular distribution of these two PUS enzymes (**Supplementary Fig. 4**). Finally, we tested the steady-state enzyme: substrate correlation with psi levels for Pb-treated and untreated cells (**Fig. 4i**). We found that the correlation of changes in U-C mismatches upon Pb treatment is not as linear as in the case of SH-SY5Y differentiation. For TRUB1-mediated pseudouridylation of putative TRUB1 sites, we find a correlation of $R^2 = 0.62$, whereas for PUS7-mediated pseudouridylation of UNUAR motifs, we observe a weak correlation ($R^2 = 0.23$).

Plasticity of pseudouridylation of mRNAs in response to changes in the cellular state
To assess differences in pseudouridylation at identified positions between differentiated and Pb^{2+}-exposed samples, we compared the U-to-C basecalling errors for psi-modified positions for each perturbation (**Fig. 5a, b**). We divided the positions into different zones based on the relationship between the conditions: 1. Single condition-dependent positional occupancy whereby the relative occupancy of a given site changes for either differentiation or Pb^{2+} exposure, demonstrating the plasticity of a given site; 2. Double condition-dependent positional occupancy whereby the positional occupancy changes for both differentiation or Pb^{2+} exposure (demonstrating plasticity); 3. Condition-independent positional occupancy, whereby the positional occupancy does not change in response to the different conditions (these modifications are present and stable between different perturbations, i.e., static). Interestingly, we observed a cluster of targets (541 sites) that were decreased for both Pb^{2+} exposure and differentiation. We compared this distribution to a random distribution. We found that the two distributions were significantly different (Mann–Whitney–Wilcoxon test, $p < 0.001$), meaning that the concomitant decrease in both libraries is inconsistent with a random effect. We calculated the number of pseudouridine sites detected by Mod-*p* ID across cell conditions and normalized it by the total number of reads in each of the three samples. We found that the Pb^{2+}-exposed sample has the higher normalized percentage of psi sites (0.07%), followed by the differentiated (0.04%) and untreated conditions (0.02%).

We selected positions that met the following criteria to evaluate the changes to individual positions between the three conditions: 1. Detected as a modified position with *p*-value < 0.001 in at least one condition; 2. The U-to-C mismatch was >40% in at least one of the conditions. We rank-ordered these positions based on a calculated similarity score (**Fig. 5c**). The most similar position between the three conditions is *ARF6* (chr14:49895568), with a similarity score of 0 and >90% U-to-C error for each condition. The conditions with similarity scores below five are considered the most static positions. Interestingly, 22 out of 77 positions with similarity scores below five fall within a TRUB1 motif, and 4 out of 77 positions fall within a PUS7 motif. The least similar position is *SZRD1* (chr1:16396654), with a similarity score of 25.2 and U-to-C errors of 67.3%, 31,6%, and 53.8%, respectively. Positions with a similarity score of >5 are considered to have higher plasticity. Among the positions detected as psi with the

highest score across the three conditions (i.e., high plasticity) is *YTHDF1* (chr20:63202504), which encodes a well-characterized m^6A reader protein that promotes protein synthesis in response to neuronal stimuli[56].

The abundance of psi synthase may account for some of the differences in psi percentage within that condition. Hence, we compared mRNA levels for 11 psi synthases across all three conditions and found no significant differences in mRNA expression (**Fig. 5d**). We also hypothesized that for the sites with higher plasticity, mRNA expression of the target may explain the discrepancy (**Fig. 5e**). However, we did not observe an association between mRNA expression levels and the percentage of U-to-C error for targets with >11 similarity score. For example, *THY1* (chr11:119418679) has similar U-to-C error for the differentiated and Pb^{2+} treated conditions (61.3% and 45.1%, respectively) and 33.5% for the untreated; however, the mRNA expression levels for this target across each condition are not statistically different.

Discussion

This study assessed sites of U-to-C basecalling error in SH-SY5Y cells based on direct RNA nanopore sequencing. We assigned them psi modifications by siRNA knockdown of the predominant psi synthases acting on human mRNAs, TRUB1 and PUS7. Interestingly, in TRUB1 KD libraries, certain positions displaying TRUB1 motifs with high U-to-C basecalling errors did not exhibit changes following TRUB1 KD, and most of these positions have been validated through chemical-based methods. This finding suggests that other PUS enzymes may compensate for knocked-down TRUB1 expression.

We tested the validity of a steady-state mRNA site/PUS enzyme action model by correlating changes in PUS enzyme and mRNA U-site levels with the corresponding changes in U-to-C mismatch frequencies (i.e., ψ levels for these validated motif sites). First, we observed a strong correlation between the enzyme-substrate levels and U-to-C mismatch frequencies for *TRUB1* and *PUS7* knockdown experiments. In contrast, we found a weaker correlation when comparing *PUS7* to *TRUB1* mRNA targets and *TRUB1* levels to *PUS7* mRNA targets. This supports the cross-specificity of Trub1 and Pus7 to their consensus motifs (GUUCN for Trub1 and UNUAR for Pus7), as well as the observation that steady-state enzyme-substrate levels are a good predictor of ψ levels (with the caveat that actual U-to-C mismatch frequencies are not quantitative proxies of ψ levels [28]. While we observe a similar steady-state enzyme-substrate response for SH-SY5Y differentiation using retinoic acid, we find that for the Pb^{2+}-treated undifferentiated SH-SY5Y cells, there is a weaker correlation between enzyme-substrate product levels and psi levels. This suggests that for Pb^{2+} treatment, there are additional factors (e.g., a trans-acting factor or other PUS enzyme dysregulation mechanisms) that disrupt the simple steady-state enzyme-substrate pseudouridylation.

We differentiated SH-SY5Y cells into neuron-like cells using retinoic acid and compared the relative occupancy of psi at the same positions. Upon comparing SH-SY5Y cells in undifferentiated and differentiated states, we noted a change in U-to-C basecalling at several positions following differentiation. We then applied Pb^{2+} to the SH-SY5Y cells. Pb^{2+} is a neurotoxicant chosen because exposure to a neurotoxicant should change the state of the cell differently than a differentiation. Our goal was to observe changes to psi at positions in response to changes in cell state. Our analysis of SH-SY5Y cells in untreated and Pb^{2+}-treated states revealed distinct variations in U-to-C basecalling errors at certain positions between differentiated and untreated SH-SY5Y cells. Interestingly, we did not find any significant differences in the expression level of these positions. Moreover, we found no notable

differences in PUS enzyme expression or localization between differentiated and untreated libraries (**Supplementary Fig. 3**), which have been shown previously to relocalize in response to stress[42].

Finally, our analysis of multiple cellular conditions (untreated, differentiated, and Pb^{2+} treated) revealed various types of site responses: we found numerous positions with high similarity scores (i.e., low similarity), which indicates the plasticity of those sites. Interestingly, the most static psi positions were frequently targets of TRUB1, while the more plastic psi positions were often targets of PUS7. Among these, *NES* (chr1:156669130) encodes the neuronal marker and cytoskeletal protein Nestin, which plays a role in differentiation and self-renewal. This *NES* psi site has decreased psi levels following differentiation, while it does not change upon Pb^{2+} treatment. Noteworthy, mRNA levels for *NES* are unchanged across conditions (**Fig. 5e**), suggesting a possible mechanism of psi-mediated translational control. In contrast, both *THY1* (chr11:119418679) and *NKAIN1*(chr1:31181249) exhibit high upregulation of psi occupancy for both Pb^{2+} treatment and differentiation, and both encode for proteins that are involved in neuronal processes. Possible functions for sites highly sensitive to environmental factors include translational control in response to cellular stress and maintenance of cellular fitness.

This study was the first to determine static positions of psi modifications across cellular states and plastic sites sensitive to specific environmental stimuli. Future analysis will determine whether static sites play critical roles in the cell's biological function and whether the plastic sites are responses to external cues for fine-tuning gene expression.

Methods

Cell culture

Human neuroblastoma SH-SY5Y cells were cultured in EMEM/F12 (EMEM from Quality Biological Inc and Cytiva HyClone Ham's Nutrient Mixture F12 Media) supplemented with 10% Fetal Bovine Serum (FisherScientific, FB12999102). For untreated SH-SY5Y cells, the culture remained in this medium for seven days at 37C and 5% CO_2, refreshed every three days. For differentiated SH-SY5Y cells, after 24h, the media changed to differentiation media, which is Neurobasal media (Gibco Neurobasal-A Medium, minus phenol red) supplemented with 10uM all-trans-retinoic acid (Fisher, AC207341000), 1X B27(Fisher, A3582801), and 1X Glutamax (Fisher, 35-050-061). The differentiation media was renewed every other day. For lead exposure SH-SY5Y cells, after 24 hours, the culture media was removed, and a 50 μM Pb^{2+}-supplemented (Lead(II) acetate trihydrate, Sigma) untreated media was added to the cells. The media was replaced every three days.

Immunofluorescence (IF)

For fixing SH-SY5Y cells, half of the culture media was removed, and an equal volume of 4% formaldehyde (Fisher, F79500) in PBS was added to each well for the final of 2% formaldehyde. After 2 min incubation at room temperature, the solution was aspirated, replaced by 4% formaldehyde, and incubated for 10 mins. The cells were then washed with PBS and permeabilized by incubating in PBS-Triton (0.1%) for 10 min. The cultures were blocked by incubation in 2% bovine serum albumin (BSA) in BS-Triton (0.1%) for one hour, followed by three times washed with PBS-Tween 20 (0.1%). The cells were then incubated with 1ug/ul primary antibody (For TRUB1 staining, TRUB1 Rabbit anti-human polyclonal, 50-172-8037, Protein tech; for PUS7 staining, PUS7 Rabbit anti-human, HPA024116, Sigma) in 1% BSA/PBS-Triton (0.1%) overnight at 4C. The following day, the cells were washed with PBS-Tween 20 (0.1%) and incubated for one hour in 1:1000 secondary antibody (Mouse anti-rabbit

IgG-PE-Cy7, NC1569698, Fisher) in 1% BSA/PBS-Triton (0.1%) at room temperature, and stained using DAPI.

siRNA Knockdown (KD)
SH-SY5Y cells were cultured in untreated media for 24h for KD and control samples. The media was replaced with siRNA delivery media (Horizon, B-005000-500) with 1uM of Accell Non-targeting Control Pool (Horizon, D-001910-10-05), PUS7 siRNA (Horizon, E-015341-00-0050) or TRUB1 siRNA (Horizon, E-016391-00-0050) in delivery media for KD samples and cultured for 72h. Total RNA was extracted for qPCR KD confirmation or fixed for IF imaging.

Total RNA extraction
Total RNA was extracted from cells by combining a TRIzol (Invitrogen,15596026) RNA extraction and the PureLink RNA Mini Kit (Invitrogen, 12183025). Cells were first washed with ice-cold PBS, followed by incubation for 5 min in TRIzol at RT (2ml for 10mm dishes and 300ul for 8-well plates). Then, the solution was transferred to Eppendorf tubes, and 200ul chloroform (Acros Organics, 423555000) was added to each 1ml of TRIzol. The solution was mixed by shaking it for 15 sec and incubated at RT for 3 min, followed by centrifugation for 15 min at 12000 x g at 4C. The aqueous supernatant was then transferred to a new tube, and the manufacturer's recommended protocol was followed for PureLink RNA Mini Kit RNA extraction. RNA was quantified using the RNA Qubit assay.

Poly-A RNA isolation
According to the manufacturer's protocol, poly-A mRNA was selected using the NEBNext Poly(A) mRNA Magnetic Isolation Module (E7490L). RNA was quantified using the RNA Qubit assay.

RT-qPCR
The extracted total RNA was treated with TURBO DNase (Fisher, AM2238) following the manufacturer’s protocol. The RNA is then reverse transcribed using SuperScript III RT kit (Fisher,18080044) using the target primers and housekeeping gene HPRT. qPCR was performed using Luna qPCR master mix (NEB, M3004).

Direct RNA library preparation and sequencing
A direct RNA sequencing library (SQK-RNA002) was prepared following the manufacturer’s instructions. Briefly, 500ng poly-A tailed RNA was ligated to ONT RT adaptor (RTA) using T4 DNA ligase (NEB, M0202M) and reverse transcribed by SuperScriptTM III Reverse transcriptase (Invitrogen, 18080044). The product was then purified using Agencourt RNAClean XP beads (Beckman, A63987) ligated to the RNA adaptor (RMX) and purified by Agencourt RNAClean XP beads, followed by washing with wash buffer (WSB) and eluted in elution buffer (ELB). The final product was mixed with an RNA running buffer and loaded into R9.4.1 FLO-MIN106D flow cells from ONT. For the KD samples and scrambled control, the samples were loaded onto PromethION flow cells (ONT, FLO-PRO004RA).

Base-calling and alignment
Fast5 files were basecalled using Guppy version 6.4.2 and aligned to the human reference genome (hg38) using Minimap2 version 2.17 with the ‘‘-ax splice -uf -k 14’’ option. The aligned .sam files were converted to .bam and indexed using samtools version 2.8.13.

SNR Calculation

We modeled the IVT and DRS data separately for each target position with a beta-binomial distribution and Jeffrey's prior. We used U counts and C counts to parameterize the beta of the beta-binomial distribution and calculate the log marginal likelihood of the posterior distribution. Additionally, we modeled a combined distribution of IVT and DRS U and C counts with the same beta-binomial and Jeffrey's prior. The ratio of these log marginal likelihoods approximates the degree to which the U to C mismatch at a position is better modeled with two independent distributions instead of a single joined distribution. According to Kass and Raftery's model [57], we selected a value of SNR >1 to guarantee a strong DRS signal level compared to the IVT.

$$SNR = \log\left(\frac{\Theta_{DRS}(X_{DRS}) * \Theta_{IVT}(X_{IVT})}{\Theta_{DRS+IVT}(X_{DRS+IVT})}\right)$$

$$\Theta = BetaBinom(n, \alpha, \beta)$$

Supplemental Information
Supplemental information can be found online at xxx.

Resource availability
Lead contact
Further information and requests for resources should be directed to and will be fulfilled by the lead contact, Sara H. Rouhanifard (*s.rouhanifard@northeastern.edu*).

Accessing Publicly Available Data Sets
All IVT Libraries used in this work were sourced from NIH NCBI SRA under BioProject accession PRJNA947135.

Data and code availability
All fast5 raw data for Direct Libraries generated in this work has been made publicly available in NIH NCBI SRA under the BioProject accession PRJNA1092333.

All code can be found at github.com/RouhanifardLab/NeuronalEpitranscriptomePlasticity.

Acknowledgments
We acknowledge support from the National Institutes of Health (R01HG011087, O.F., S.T., C.A.M, S.H.R., M.W.) and R01HG012856 (K.N., Y.Q., A.M., M.W., S.H.R).

Author contributions
Conceptualization S.T., S.H.R, and M.W.; Methodology O.F., S.H.R, and M.W.; Software O.F., S.T., and S.A.; Formal analysis O.F.; Investigation S.T., D.B., O.F., and Y.Q.; Resources S.T. and D.B.; Writing - Original Draft S.T., S.H.R.; Writing - Rewiev & Editing S.H.R., M.W., O.F., D.B., S.A, C.A.M, and K.Q.; Visualization O.F., S.T., S.H.R., K.N., and C.A.M.; Supervision S.H.R., M.W., and M.J.; Project administration S.H.R. and M.W.; Funding acquisition S.H.R.

Declaration of Interest
The authors declare no competing interests.

FIGURES

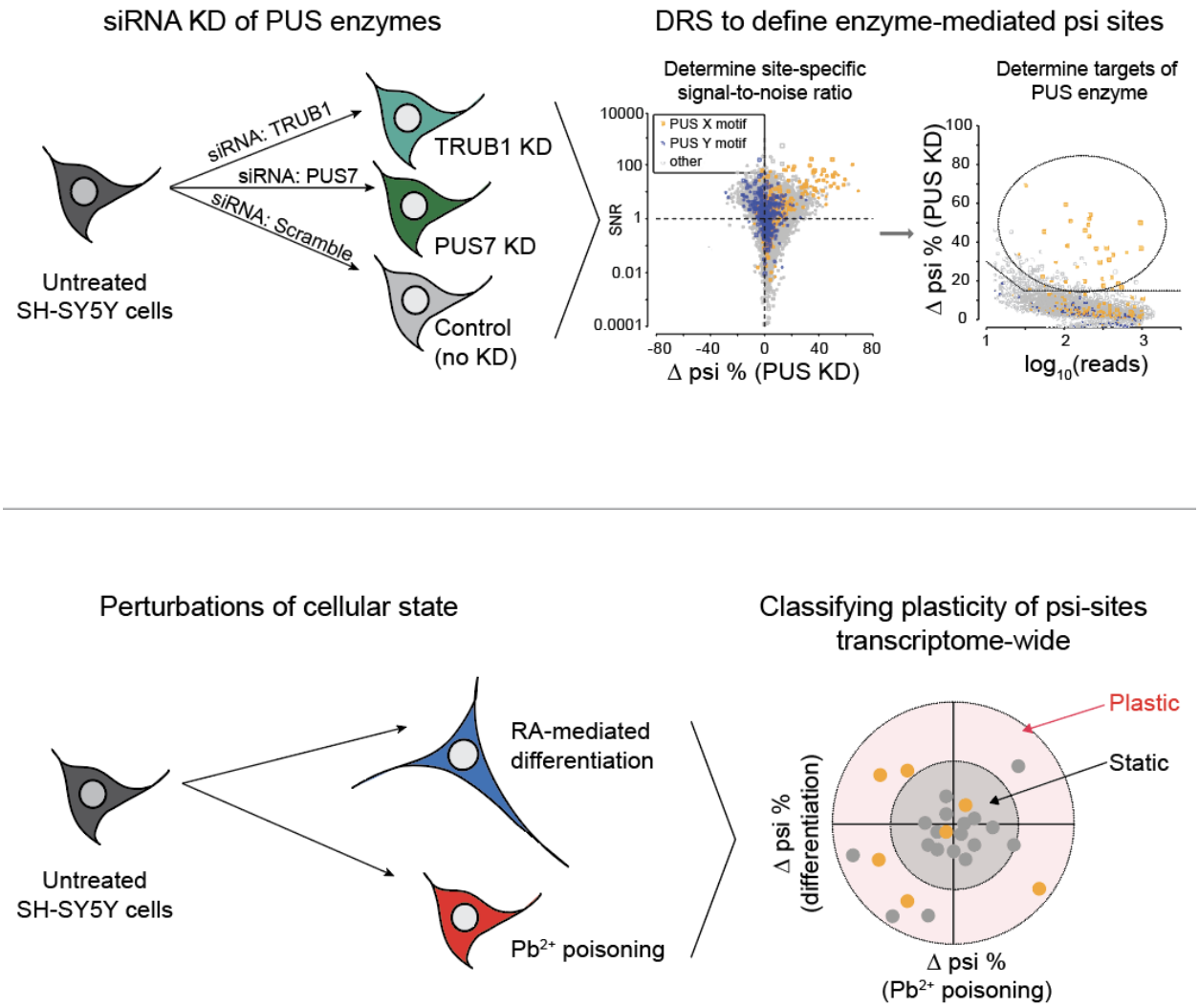

Figure 1 (Graphical Abstract)

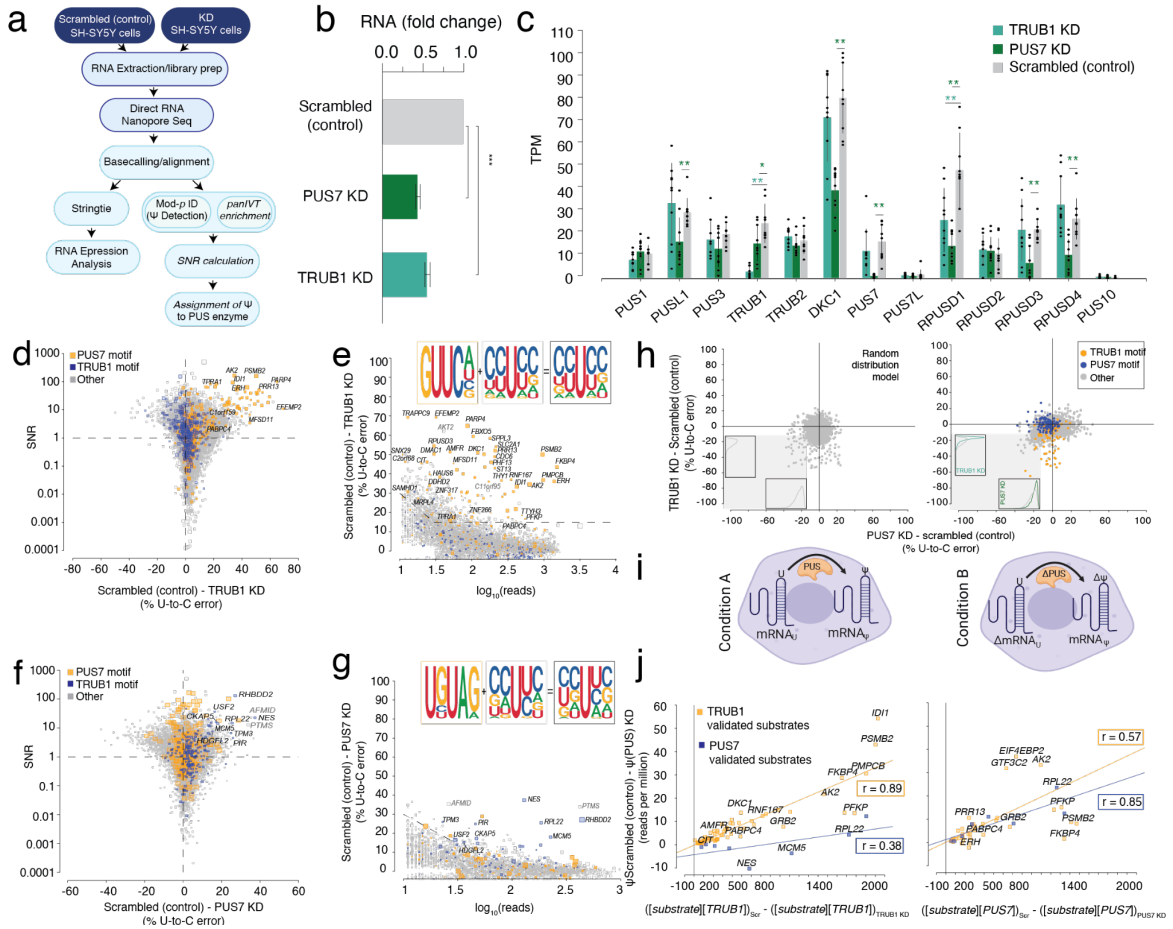

Figure 2. Enzymatic KD of PUS enzymes and DRS is used to determine enzyme-mediated psi-sites.

a. Schematic workflow of siRNA knockdown (KD), DRS, and analysis.

b. The concentration of PUS7 and TRUB1 mRNA in SH-SY5Y cells for the scrambled (control), PUS7 KD, and TRUB1 KD, respectively, following siRNA KD, was quantified by RT-qPCR.

c. TPM of various PUS enzymes following TRUB1 knockdown (KD), PUS7 knockdown (KD), and scrambled (control) determined by DRS. Individual colored bars represent each experimental condition, with error bars describing the standard error of the mean (SEM) across downsampled replicates. Individual replicates are shown as black dots.

d. SNR was calculated using beta-binomial distribution and Jeffrey's prior, and particular points were plotted against the difference in U-to-C error between the Scrambled (control) and TRUB1 KD. Orange dots represent uridine positions within a TRUB1 motif, and blue dots represent uridine positions within a PUS7 motif.

e. Putative psi-positions determined by Mod-p ID are plotted according to the difference in U-to-C basecalling error in the scrambled (control) and the TRUB1 KD against the reads for each position. The inlet shows the sequencing logo for positions within the TRUB1 motif, grey points above the threshold line, and total points above the threshold line.

f. SNR was calculated using beta-binomial distribution and Jeffrey's prior, and particular points were plotted against the difference in U-to-C error between the Scrambled (control) and PUS7 KD. Orange dots represent uridine positions within a TRUB1 motif, and blue dots represent uridine positions within a PUS7 motif.

g. Putative psi-positions determined by Mod-p ID are plotted according to the difference in U-to-C basecalling error in the scrambled (control) and the TRUB1 KD against the reads for each position. The inlet shows the sequencing logo for positions within the PUS7 motif, grey points above the threshold line, and total points above the threshold line.

h. (left) Random distribution model of points constrained by the data parameters. Histograms indicate the distribution of points in the lower left quadrant as a function of each simulated knockdown. (right) distribution of TRUB1 KD plotted against PUS7 knockdown. Histograms indicate the distribution of points in the lower left quadrant as a function of each knockdown.

i. (Left) Steady-state model of PUS enzyme concentration affecting psi deposition at a given position. (Right) Steady-state model of the difference in pus concentration between two conditions being proportional to the substrate concentration (transcript).

j. Steady-state PUS enzyme:mRNA substrate correlation with psi levels for knockdown cell lines. Left: plot of number of reads per million that contain U-C mismatches vs. the difference in the product of TRUB1 substrate and the TRUB1 enzyme concentrations between Scrambled and TRUB1 knockdown cell lines (plotted for both TRUB1 and PUS7 sites). Right: Similar plot as in Left, except the comparison was made for PUS7 enzyme concentrations. The lines shown are Pearson correlations with R^2 and ρ values indicated in boxes.

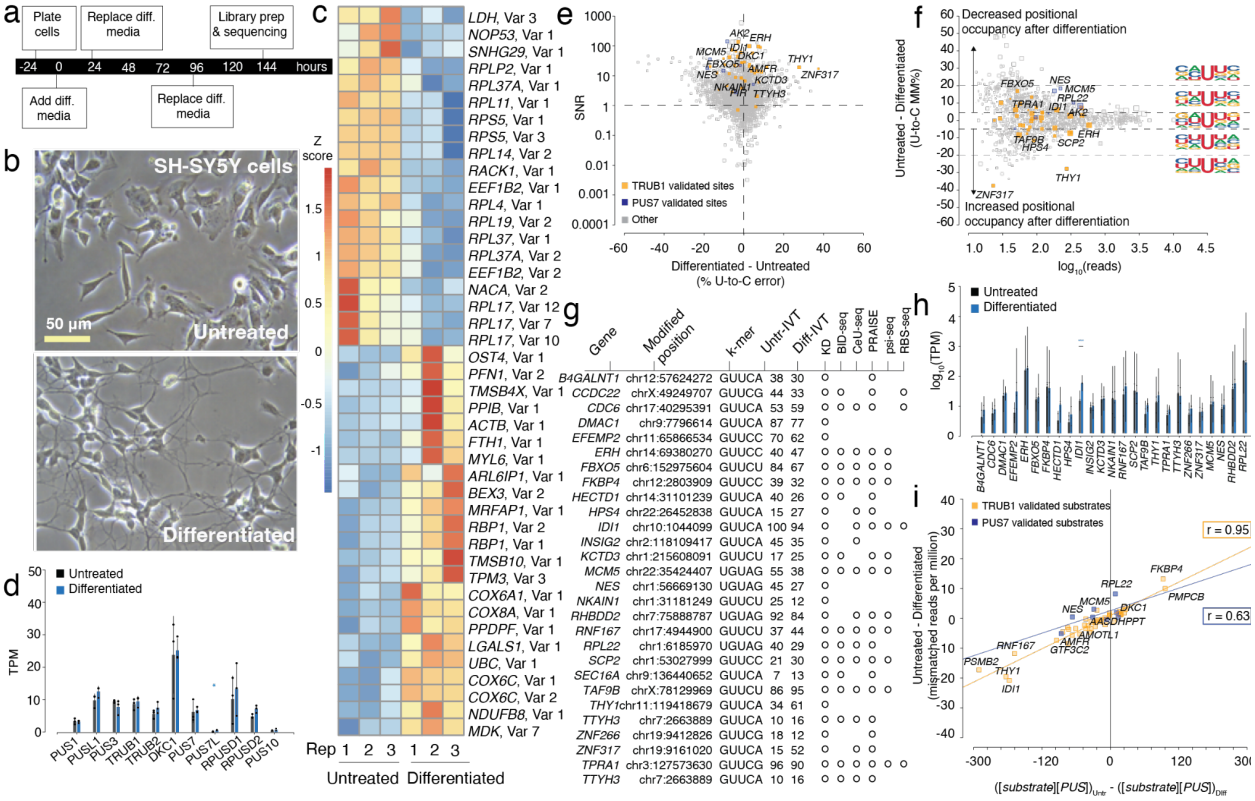

Figure 3. Effects of RA-mediated differentiation on mRNA psi modification and machinery in SH-SY5Y cells

a. Timeline illustrating the stages and duration of the RA treatment applied to SH-SY5Y cells.

b. A representative photomicrograph of untreated and differentiated SH-SY5Y cells is shown.

c. We used Deseq2 to identify the transcripts with the highest fold change between the untreated and differentiated samples. Three biological replicates for each condition were used. The color scale shows a Z score based on the relative fold change.

d. TPM of various PUS enzymes in untreated and differentiated SH-SY5Y cells determined by DRS. Individual colored bars represent each experimental condition, with error bars describing the standard error of the mean (SEM) across downsampled replicates. Individual replicates are shown as black dots.

e. SNR was calculated using beta-binomial distribution and Jeffrey's prior, plotting particular points against the difference in U-to-C error between the untreated and differentiated libraries. Orange dots represent uridine positions that are validated TRUB1 substrates, and blue dots represent uridine positions that are validated PUS7 substrates.

f. Putative psi-positions determined by Mod-*p* ID are plotted according to the difference in U-to-C basecalling error in the untreated and differentiated samples against the reads for each position. A dotted line at the +5% and -5% marks indicates the cutoff for a position to be changed in response to perturbation. The inlet shows the sequencing logo for positions within the TRUB1 motif, grey points above the threshold line, and total points above the threshold line.

g. Annotation of genes containing a psi modification validated by PUS7 or TRUB1 KD (**Figure 2**) and orthogonal methods.

h. TPM of the transcripts bearing a validated psi modification that had the most significant differences between conditions, determined by DRS. Individual colored bars represent each experimental condition, with error bars describing the standard error of the mean (SEM) across downsampled replicates. Individual replicates are shown as black dots.

i. Correlation of the differential U-C mismatch number of reads vs. the enzyme: substrate product for untreated and differentiated SH-SY5Y cells. A positive correlation was observed for both PUS enzymes (R^2=0.73 for TRUB1 and R^2=0.33 for PUS7).

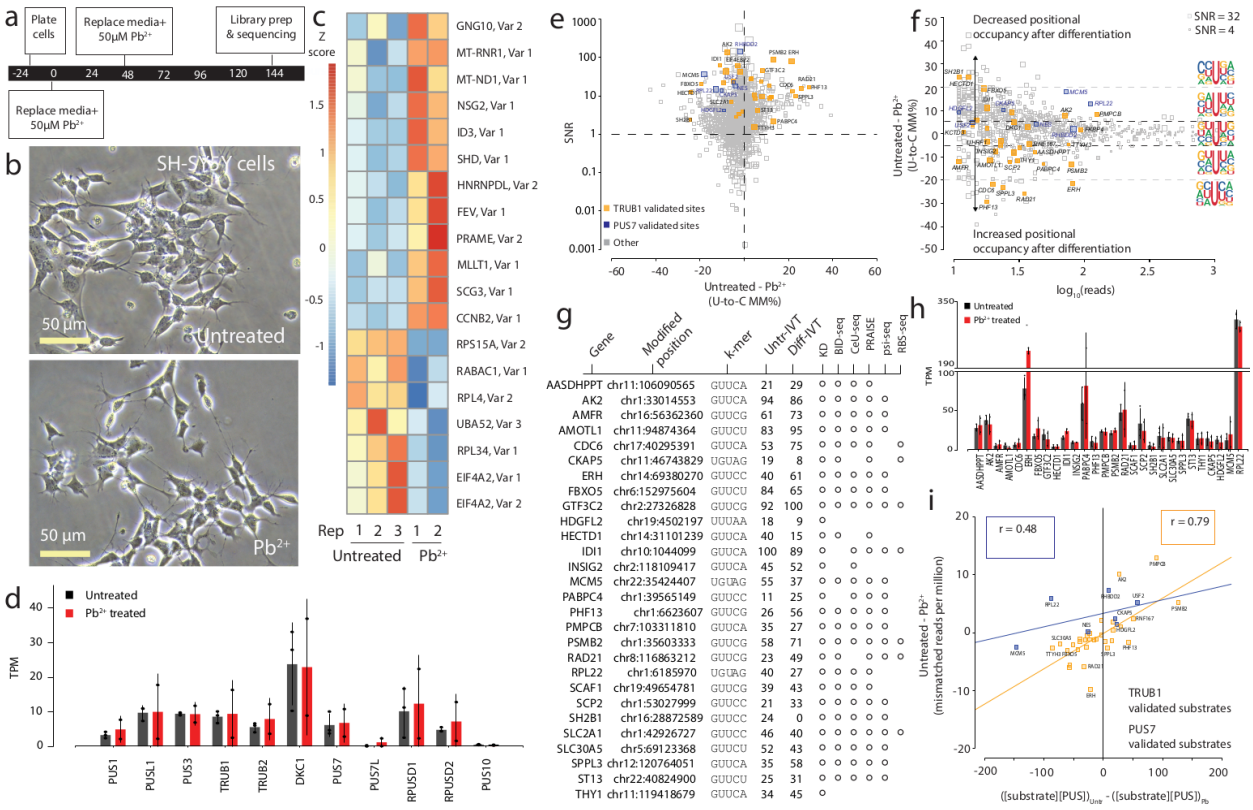

Figure 4. Effects of Pb^{2+} treatment on mRNA psi modification and machinery in SH-SY5Y cells

a. Timeline illustrating the stages and duration of the Pb^{2+} treatment applied to SH-SY5Y cells.

b. A representative photomicrograph of untreated and Pb^{2+} treated SH-SY5Y cells is shown.

c. We used Deseq2 to identify the transcripts with the highest fold change between the untreated and Pb^{2+} treated samples. 3 biological replicates of each condition were used. The color scale shows a Z score based on the relative fold change.

d. DRS determined the TPM of various PUS enzymes in untreated and Pb^{2+}-treated SH-SY5Y cells. Individual colored bars represent each experimental condition, with error bars describing the standard error of the mean (SEM) across downsampled replicates. Individual replicates are shown as black dots.

e. SNR was calculated by using beta-binomial distribution and Jeffrey's prior, plotting particular points against the difference in U-to-C error between the untreated and Pb^{2+} treated libraries. Orange dots represent uridine positions that are validated TRUB1 substrates, and blue dots represent uridine positions that are validated PUS7 substrates.

f. Putative psi-positions determined by Mod-*p* ID are plotted according to the difference in U-to-C basecalling error in the untreated and Pb^{2+} treated samples against the reads for each position. A dotted line at the +5% and -5% marks indicates the cutoff for a position to be changed in response to perturbation. The inlet shows the sequencing logo for positions within the TRUB1 motif, grey points above the threshold line, and total points above the threshold line.

g. Annotation of genes containing a psi modification that changed in response to perturbation and validated by PUS7 or TRUB1 KD (**Figure 2**) and orthogonal methods.

h. TPM of the transcripts bearing a validated psi modification that had the most significant differences between conditions determined by DRS. Individual colored bars represent each experimental condition, with error bars describing the standard error of the mean (SEM) across downsampled replicates. Individual replicates are shown as black dots.

i. Correlation of the differential U-C mismatch number of reads vs. the enzyme: substrate product for untreated and Pb^{2+} treated SH-SY5Y cells. Lines are Pearson correlation fits showing weak positive correlations for both PUS enzymes (R^2=0.62 for TRUB1 and R^2=0.23 for PUS7).

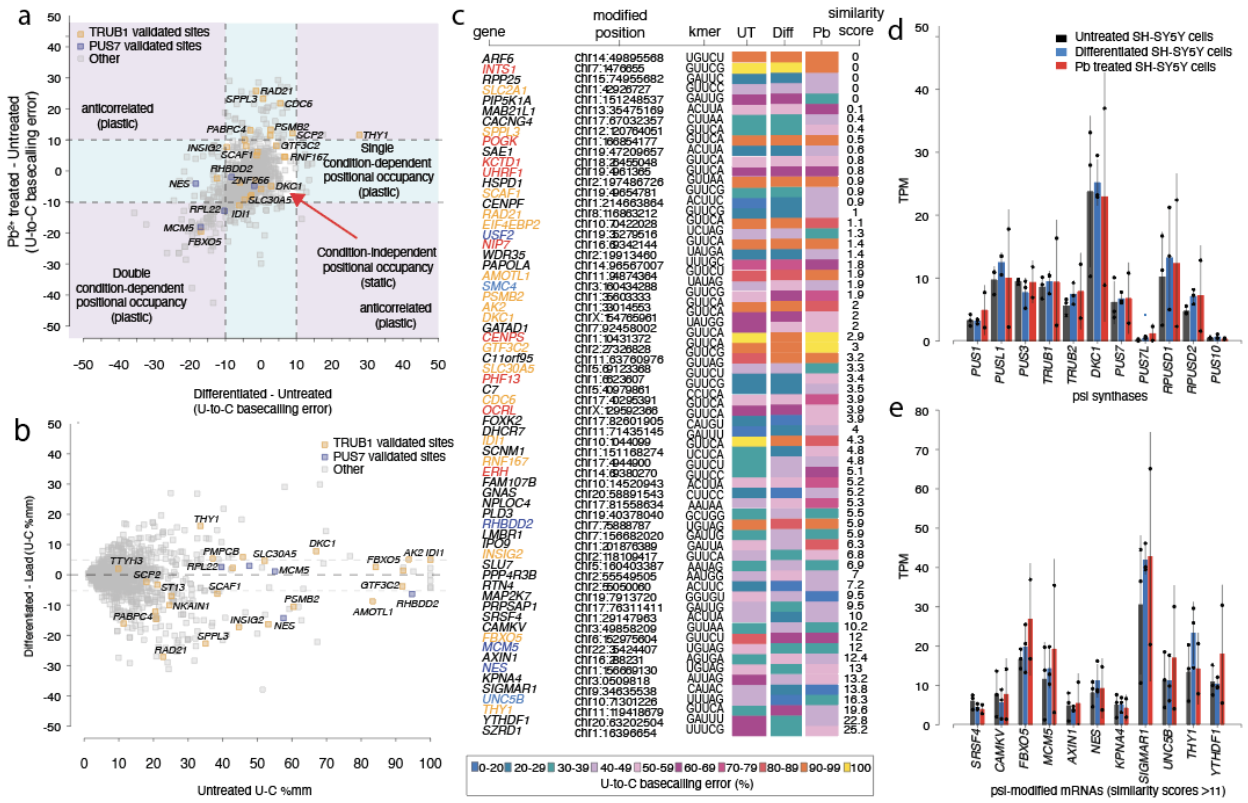

Figure 5. Psi analysis across three cellular states enables the classification of plastic and static sites of modification, transcriptome-wide

a. Putative psi-positions determined by Mod-*p* ID are plotted according to the difference in U-to-C basecalling error in the untreated and Pb^{2+} treated samples against the difference in U-to-C basecalling error between the differentiated and untreated for each position. A dotted line at the +10% and -10% marks indicates the cutoff for a position to be changed in response to perturbation. Condition-independent positional occupancy (static) positions are located in the center square. Single condition-dependent positional occupancy is shown in the blue stripes. Double condition-dependent positional occupancy is shown in the purple areas.

b. Putative psi-positions determined by Mod-*p* ID are plotted according to the difference in U-to-C basecalling error in the Differentiated and Pb^{2+} treated samples against the U-to-C basecalling error for the untreated sample. A dotted line at the +5% and -5% marks indicates the cutoff for a position to be significantly different between the two perturbations.

c. Heatmap of sites with at least 40% U-to-C basecalling error in one of three conditions. Colors indicate the percentage of U-to-C basecalling errors. A similarity score is calculated for each of these positions, and they are ranked in order, with the most similar at the top and the least similar at the bottom. Positions validated substrates for TRUB1 are shown in orange, and TRUB1 motifs that have not yet been validated are shown in red. Positions that are validated substrates for PUS7 are shown in dark blue, and PUS7 motifs that have not been validated are shown in light blue.

d. DRS determined the TPM of various PUS enzymes in untreated, differentiated, and Pb^{2+}-treated SH-SY5Y cells. Individual colored bars represent each experimental condition, with error bars describing the standard error of the mean (SEM) across downsampled replicates. Individual replicates are shown as black dots.

e. TPM of the transcripts bearing a validated psi modification that had the most significant differences between conditions determined by DRS. Individual colored bars represent each experimental condition, with error bars describing the standard error of the mean (SEM) across downsampled replicates. Individual replicates are shown as black dots.

REFERENCES

1. Roundtree, I.A., Evans, M.E., Pan, T., and He, C. (2017). Dynamic RNA Modifications in Gene Expression Regulation. Cell *169*, 1187–1200.

2. Jia, G., Fu, Y., Zhao, X., Dai, Q., Zheng, G., Yang, Y., Yi, C., Lindahl, T., Pan, T., Yang, Y.-G., et al. (2011). N6-methyladenosine in nuclear RNA is a major substrate of the obesity-associated FTO. Nat. Chem. Biol. *7*, 885–887.

3. Franco, M.K., and Koutmou, K.S. (2022). Chemical modifications to mRNA nucleobases impact translation elongation and termination. Biophys. Chem. *285*, 106780.

4. Carlile, T.M., Martinez, N.M., Schaening, C., Su, A., Bell, T.A., Zinshteyn, B., and Gilbert, W.V. (2019). mRNA structure determines modification by pseudouridine synthase 1. Nat. Chem. Biol. *15*, 966–974.

5. Parr, C.J.C., Wada, S., Kotake, K., Kameda, S., Matsuura, S., Sakashita, S., Park, S., Sugiyama, H., Kuang, Y., and Saito, H. (2020). N 1-Methylpseudouridine substitution enhances the performance of synthetic mRNA switches in cells. Nucleic Acids Res. *48*, e35.

6. Li, X., Zhu, P., Ma, S., Song, J., Bai, J., Sun, F., and Yi, C. (2015). Chemical pulldown reveals dynamic pseudouridylation of the mammalian transcriptome. Nat. Chem. Biol. *11*, 592–597.

7. Li, X., Xiong, X., and Yi, C. (2016). Epitranscriptome sequencing technologies: decoding RNA modifications. Nat. Methods *14*, 23–31.

8. Dominissini, D., Moshitch-Moshkovitz, S., Schwartz, S., Salmon-Divon, M., Ungar, L., Osenberg, S., Cesarkas, K., Jacob-Hirsch, J., Amariglio, N., Kupiec, M., et al. (2012). Topology of the human and mouse m6A RNA methylomes revealed by m6A-seq. Nature *485*, 201–206.

9. Liu, J., Yue, Y., Han, D., Wang, X., Fu, Y., Zhang, L., Jia, G., Yu, M., Lu, Z., Deng, X., et al. (2013). A METTL3–METTL14 complex mediates mammalian nuclear RNA N6-adenosine methylation. Nat. Chem. Biol. *10*, 93–95.

10. Cortese, R., Kammen, H.O., Spengler, S.J., and Ames, B.N. (1974). Biosynthesis of pseudouridine in transfer ribonucleic acid. J. Biol. Chem. *249*, 1103–1108.

11. Kierzek, E., Malgowska, M., Lisowiec, J., Turner, D.H., Gdaniec, Z., and Kierzek, R. (2014). The contribution of pseudouridine to stabilities and structure of RNAs. Nucleic Acids Res. *42*, 3492–3501.

12. Borchardt, E.K., Martinez, N.M., and Gilbert, W.V. (2020). Regulation and Function of RNA Pseudouridylation in Human Cells. Annu. Rev. Genet. *54*, 309–336.

13. Arroyo, J.D., Jourdain, A.A., Calvo, S.E., Ballarano, C.A., Doench, J.G., Root, D.E., and Mootha, V.K. (2016). A Genome-wide CRISPR Death Screen Identifies Genes Essential for Oxidative Phosphorylation. Cell Metab. *24*, 875–885.

14. Bakin, A., and Ofengand, J. (1993). Four newly located pseudouridylate residues in Escherichia coli 23S ribosomal RNA are all at the peptidyltransferase center: analysis by the application of a new sequencing technique. Biochemistry *32*, 9754–9762.

15. Martinez, N.M., Su, A., Burns, M.C., Nussbacher, J.K., Schaening, C., Sathe, S., Yeo, G.W., and Gilbert, W.V. (2022). Pseudouridine synthases modify human pre-mRNA co-transcriptionally and affect pre-mRNA processing. Mol. Cell *82*, 645–659.e9.

16. Addepalli, B., and Limbach, P.A. (2011). Mass spectrometry-based quantification of pseudouridine in RNA. J. Am. Soc. Mass Spectrom. *22*, 1363–1372.

17. Anderson, B.R., Muramatsu, H., Nallagatla, S.R., Bevilacqua, P.C., Sansing, L.H., Weissman, D., and Karikó, K. (2010). Incorporation of pseudouridine into mRNA enhances translation by diminishing PKR activation. Nucleic Acids Res. *38*, 5884–5892.

18. Antonicka, H., Choquet, K., Lin, Z.-Y., Gingras, A.-C., Kleinman, C.L., and Shoubridge, E.A. (2017). A pseudouridine synthase module is essential for mitochondrial protein synthesis and cell viability. EMBO Rep. *18*, 28–38.

19. Eyler, D.E., Franco, M.K., Batool, Z., Wu, M.Z., Dubuke, M.L., Dobosz-Bartoszek, M., Jones, J.D., Polikanov, Y.S., Roy, B., and Koutmou, K.S. (2019). Pseudouridinylation of mRNA coding sequences alters translation. Proc. Natl. Acad. Sci. U. S. A. *116*, 23068–23074.

20. Carlile, T.M., Rojas-Duran, M.F., Zinshteyn, B., Shin, H., Bartoli, K.M., and Gilbert, W.V. (2014). Pseudouridine profiling reveals regulated mRNA pseudouridylation in yeast and human cells. Nature *515*, 143–146.

21. Zhang, M., Jiang, Z., Ma, Y., Liu, W., Zhuang, Y., Lu, B., Li, K., Peng, J., and Yi, C. (2023). Quantitative profiling of pseudouridylation landscape in the human transcriptome. Nat. Chem. Biol. *19*, 1185–1195.

22. Lovejoy, A.F., Riordan, D.P., and Brown, P.O. (2014). Transcriptome-wide mapping of pseudouridines: pseudouridine synthases modify specific mRNAs in S. cerevisiae. PLoS One *9*, e110799.

23. Khoddami, V., Yerra, A., Mosbruger, T.L., Fleming, A.M., Burrows, C.J., and Cairns, B.R. (2019). Transcriptome-wide profiling of multiple RNA modifications simultaneously at single-base resolution. Proc. Natl. Acad. Sci. U. S. A. *116*, 6784–6789.

24. Dai, Q., Zhang, L.-S., Sun, H.-L., Pajdzik, K., Yang, L., Ye, C., Ju, C.-W., Liu, S., Wang, Y., Zheng, Z., et al. (2022). Quantitative sequencing using BID-seq uncovers abundant pseudouridines in mammalian mRNA at base resolution. Nat. Biotechnol. 10.1038/s41587-022-01505-w.

25. Begik, O., Lucas, M.C., Pryszcz, L.P., Ramirez, J.M., Medina, R., Milenkovic, I., Cruciani, S., Liu, H., Vieira, H.G.S., Sas-Chen, A., et al. (2021). Quantitative profiling of pseudouridylation dynamics in native RNAs with nanopore sequencing. Nat. Biotechnol. 10.1038/s41587-021-00915-6.

26. Huang, S., Zhang, W., Katanski, C.D., Dersh, D., Dai, Q., Lolans, K., Yewdell, J., Eren,

A.M., and Pan, T. (2021). Interferon inducible pseudouridine modification in human mRNA by quantitative nanopore profiling. Genome Biol. *22*, 330.

27. Huang, S., Wylder, A.C., and Pan, T. (2024). Simultaneous nanopore profiling of mRNA m6A and pseudouridine reveals translation coordination. Nat. Biotechnol. 10.1038/s41587-024-02135-0.

28. Tavakoli, S., Nabizadeh, M., Makhamreh, A., Gamper, H., McCormick, C.A., Rezapour, N.K., Hou, Y.-M., Wanunu, M., and Rouhanifard, S.H. (2023). Semi-quantitative detection of pseudouridine modifications and type I/II hypermodifications in human mRNAs using direct long-read sequencing. Nat. Commun. *14*, 334.

29. McCormick, C.A., Akeson, S., Tavakoli, S., Bloch, D., Klink, I.N., Jain, M., and Rouhanifard, S.H. (2023). Multicellular, IVT-derived, unmodified human transcriptome for nanopore direct RNA analysis. bioRxiv. 10.1101/2023.04.06.535889.

30. Makhamreh, A., Tavakoli, S., Gamper, H., Nabizadehmashhadtoroghi, M., Fallahi, A., Hou, Y.-M., Rouhanifard, S.H., and Wanunu, M. Messenger-RNA Modification Standards and Machine Learning Models Facilitate Absolute Site-Specific Pseudouridine Quantification. Preprint, 10.1101/2022.05.06.490948 10.1101/2022.05.06.490948.

31. Gamper, H., McCormick, C., Tavakoli, S., Wanunu, M., Rouhanifard, S.H., and Hou, Y.-M. Synthesis of Long RNA with a Site-Specific Modification by Enzymatic Splint Ligation. Preprint, 10.1101/2022.09.17.508400 10.1101/2022.09.17.508400.

32. Fleming, A.M., Mathewson, N.J., Howpay Manage, S.A., and Burrows, C.J. (2021). Nanopore Dwell Time Analysis Permits Sequencing and Conformational Assignment of Pseudouridine in SARS-CoV-2. ACS Cent. Sci. 10.1021/acscentsci.1c00788.

33. Zhang, W., Eckwahl, M.J., Zhou, K.I., and Pan, T. (2019). Sensitive and quantitative probing of pseudouridine modification in mRNA and long noncoding RNA. RNA *25*, 1218–1225.

34. Shaheen, R., Han, L., Faqeih, E., Ewida, N., Alobeid, E., Phizicky, E.M., and Alkuraya, F.S. (2016). A homozygous truncating mutation in PUS3 expands the role of tRNA modification in normal cognition. Hum. Genet. *135*, 707–713.

35. Jung, Y., and Goldman, D. (2018). Role of RNA modifications in brain and behavior. Genes Brain Behav. *17*, e12444.

36. Shipley, M.M., Mangold, C.A., and Szpara, M.L. (2016). Differentiation of the SH-SY5Y Human Neuroblastoma Cell Line. J. Vis. Exp., 53193.

37. Påhlman, S., Ruusala, A.I., Abrahamsson, L., Mattsson, M.E., and Esscher, T. (1984). Retinoic acid-induced differentiation of cultured human neuroblastoma cells: a comparison with phorbolester-induced differentiation. Cell Differ. *14*, 135–144.

38. Santa Maria, M.P., Hill, B.D., and Kline, J. (2019). Lead (Pb) neurotoxicology and cognition. Appl. Neuropsychol. Child *8*, 272–293.

39. Shih, R.A., Hu, H., Weisskopf, M.G., and Schwartz, B.S. (2007). Cumulative lead dose and

cognitive function in adults: a review of studies that measured both blood lead and bone lead. Environ. Health Perspect. *115*, 483–492.

40. D'souza, H.S., Dsouza, S.A., Menezes, G., and Venkatesh, T. (2011). Diagnosis, evaluation, and treatment of lead poisoning in general population. Indian J. Clin. Biochem. *26*, 197–201.

41. Safra, M., Nir, R., Farouq, D., Vainberg Slutskin, I., and Schwartz, S. (2017). TRUB1 is the predominant pseudouridine synthase acting on mammalian mRNA via a predictable and conserved code. Genome Res. *27*, 393–406.

42. Schwartz, S., Bernstein, D.A., Mumbach, M.R., Jovanovic, M., Herbst, R.H., León-Ricardo, B.X., Engreitz, J.M., Guttman, M., Satija, R., Lander, E.S., et al. (2014). Transcriptome-wide mapping reveals widespread dynamic-regulated pseudouridylation of ncRNA and mRNA. Cell *159*, 148–162.

43. Meyer, K.D., Saletore, Y., Zumbo, P., Elemento, O., Mason, C.E., and Jaffrey, S.R. (2012). Comprehensive analysis of mRNA methylation reveals enrichment in 3' UTRs and near stop codons. Cell *149*, 1635–1646.

44. Purchal, M.K., Eyler, D.E., Tardu, M., Franco, M.K., Korn, M.M., Khan, T., McNassor, R., Giles, R., Lev, K., Sharma, H., et al. (2022). Pseudouridine synthase 7 is an opportunistic enzyme that binds and modifies substrates with diverse sequences and structures. Proceedings of the National Academy of Sciences *119*, e2109708119.

45. Grima, R. (2009). Investigating the robustness of the classical enzyme kinetic equations in small intracellular compartments. BMC Syst. Biol. *3*, 101.

46. Raj, A., Peskin, C.S., Tranchina, D., Vargas, D.Y., and Tyagi, S. (2006). Stochastic mRNA Synthesis in Mammalian Cells. PLoS Biol. *4*, e309.

47. Korecka, J.A., van Kesteren, R.E., Blaas, E., Spitzer, S.O., Kamstra, J.H., Smit, A.B., Swaab, D.F., Verhaagen, J., and Bossers, K. (2013). Phenotypic characterization of retinoic acid differentiated SH-SY5Y cells by transcriptional profiling. PLoS One *8*, e63862.

48. Kovalevich, J., and Langford, D. (2013). Considerations for the use of SH-SY5Y neuroblastoma cells in neurobiology. Methods Mol. Biol. *1078*, 9–21.

49. Harasym, E., McAndrew, N., and Gomez, G. (2017). Sub-micromolar concentrations of retinoic acid induce morphological and functional neuronal phenotypes in SK-N-SH neuroblastoma cells. In Vitro Cell. Dev. Biol. Anim. *53*, 798–809.

50. Khazeem, M.M., Casement, J.W., Schlossmacher, G., Kenneth, N.S., Sumbung, N.K., Chan, J.Y.T., McGow, J.F., Cowell, I.G., and Austin, C.A. (2022). TOP2B Is Required to Maintain the Adrenergic Neural Phenotype and for ATRA-Induced Differentiation of SH-SY5Y Neuroblastoma Cells. Mol. Neurobiol. *59*, 5987–6008.

51. Janesick, A., Wu, S.C., and Blumberg, B. (2015). Retinoic acid signaling and neuronal differentiation. Cell. Mol. Life Sci. *72*, 1559–1576.

52. Rege, T.A., and Hagood, J.S. (2006). Thy-1 as a regulator of cell-cell and cell-matrix

interactions in axon regeneration, apoptosis, adhesion, migration, cancer, and fibrosis. FASEB J. *20*, 1045–1054.

53. Gorokhova, S., Bibert, S., Geering, K., and Heintz, N. (2007). A novel family of transmembrane proteins interacting with beta subunits of the Na,K-ATPase. Hum. Mol. Genet. *16*, 2394–2410.

54. Abadin, H., Ashizawa, A., Llados, F., and Stevens, Y.-W. (2007). Toxicological profile for lead.

55. Gülden, M., and Seibert, H. (2006). In vitro-in vivo extrapolation of toxic potencies for hazard and risk assessment—problems and new developments. ALTEX *23*, e225.

56. Shi, H., Zhang, X., Weng, Y.-L., Lu, Z., Liu, Y., Lu, Z., Li, J., Hao, P., Zhang, Y., Zhang, F., et al. (2018). m6A facilitates hippocampus-dependent learning and memory through YTHDF1. Nature *563*, 249–253.

57. Kass, R.E., and Raftery, A.E. (1995). Bayes Factors. J. Am. Stat. Assoc. *90*, 773.