Ischemic stroke rewires the neuronal translatome via stress-induced stop codon readthrough, frameshifting, and codon-biased translation

Sherif Rashad[1,2]#*, Yuki Kitamura[1-3]*, Tomohito Nagai[1-3], Daisuke Ando[1,2,4,5], Abdulrahman Mousa[1], Hajime Ikenouchi[1,4], Hidenori Endo[3], Kuniyasu Niizuma[1-3,5]

1- Department of Neurosurgical Engineering and Translational Neuroscience, Tohoku University Graduate School of Medicine, Sendai, Japan.
2- Department of Neurosurgical Engineering and Translational Neuroscience, Graduate School of Biomedical Engineering, Tohoku University, Sendai, Japan.
3- Department of Neurosurgery, Tohoku University Graduate School of Medicine, Sendai, Japan.
4- Department of Neurology, Tohoku University Graduate School of Medicine, Sendai, Japan
5- Division of Development and Discovery of Interventional Therapy, Tohoku University Hospital

*** Equal contribution**

#Corresponding author:

Sherif Rashad, MD, PhD

Sherif.mohamed.rashad.e3@tohoku.ac.jp

Abstract

Ischemic stroke triggers rapid molecular programs that are poorly captured by transcriptomics. We present a temporal ribosome-profiling atlas of mouse ischemic brain and uncover a staged translational injury response. During the hyperacute phase, mRNA abundance and ribosome occupancy are uncoupled, translation transiently shifts toward GC-ending codons, and hundreds of transcripts show a burst of stop-codon readthrough with 3′UTR ribosome footprints, consistent with termination stress. Machine-learning analysis identified pre-stop ribosome queuing and frame-specific 3′UTR ribosome usage, with contributions from local RNA structure, as key determinants of readthrough dynamics. From 6-24 h, translation exhibits progressive A-site pausing and widespread disruption of ribosomal reading frame. A subset of events localizes to discrete CDS loci with characteristic sequence/structure signatures and predicted PTC/NMD propensity, whereas most coincide with extensive alternative ORF usage and uORF remodeling of translation. These data reveal translational mechanisms that evolve across stroke and provide a framework to interrogate translation dysregulation in diseases.

Keywords

Ischemic stroke; Ribo-seq; Translation efficiency; Stop-codon readthrough; Translation termination; Frameshifting; upstream open reading frames; ORF; translational reprogramming.

Introduction

Ischemic stroke (IS) remains one of the leading causes of long-term disability and mortality worldwide[1]. Although rapid advances in reperfusion therapies over the past two decades have markedly improved patient outcomes[2,3], the global incidence of IS continues to rise, particularly in low- and middle-income countries[4]. The pathophysiology of ischemic stroke is highly dynamic and involves a coordinated yet complex interplay among neurons, microglia, astrocytes, endothelial cells, and other brain-resident cell types[5]. Neuronal death after ischemia is driven predominantly by excitotoxicity, oxidative stress, and microglia-mediated neuroinflammation[5,6], While regenerative approaches are being actively pursued[7], there remains an urgent unmet need for neuroprotective strategies capable of preserving neuronal integrity and preventing irreversible tissue loss[5,8]. The limited success of prior neuroprotective interventions is often attributed to an incomplete understanding of the molecular programs engaged during ischemia.

Despite extensive transcriptomic work, including bulk and single-cell RNA-seq studies[9-11], our understanding of the translatome in stroke remains severely underdeveloped. This represents a major gap, as accumulating evidence shows a pronounced uncoupling between mRNA abundance and protein synthesis during acute cellular stress[12]. Technologies that directly quantify ribosome-engaged mRNAs, such as ribosome profiling (Ribo-seq), offer a far more accurate representation of the cellular proteome and functional state compared to gene-expression–centric approaches[13,14]. Yet, remarkably, Ribo-seq has never been applied to the study of ischemic stroke, leaving the translational landscape of injured neurons essentially unexplored.

In parallel, oxidative stress activates extensive epitranscriptional and post-transcriptional regulatory programs that modulate both mRNA translation and tRNA function[12,15]. The tRNA epitranscriptome has emerged as a potent regulator of codon decoding and selective translation during stress[12,15-19]. We previously demonstrated that ischemia induces widespread tRNA cleavage and production of tRNA-derived fragments (tDRs)[20,21]. hinting at a deeper layer of translational and epitranscriptomic remodeling during stroke. However, no studies to date have systematically integrated translatome profiling with these regulatory processes in vivo.

Here, we address this critical knowledge gap by performing the first temporal Ribo-seq analysis in a mouse model of ischemic stroke. Our findings uncover multiple layers of translational deregulation, spanning codon-biased translation, stress-induced ribosomal reprogramming, stop-codon readthrough (SCRT), and frameshifting, that collectively shape the neuronal response to ischemia. By defining these translational adaptive and maladaptive mechanisms, our work provides previously inaccessible insights into stroke biology and highlights new molecular pathways that may be leveraged for the development of next-generation neuroprotective therapeutics.

Results

➢ Temporal dynamics of translation reflect stroke pathophysiological evolution over time

To profile translational responses to stroke, we used the distal middle cerebral artery occlusion (dMCAO) model by occluding the cortical distal M4 branch of the MCA in male mice (Figure 1a). Cortical infarction was evident at 24 h (Figure 1b). We therefore sampled tissue at 1, 6, and 24 h to capture early, intermediate, and established injury states. Immunofluorescence of the infarct region showed a progressive loss of neurons and astrocytes in the core, whereas microglia were evident at 24 h (Figure 1c).

We next profiled molecular changes using matched RNA-seq and Ribo-seq at each timepoint. Ribo-seq libraries showed the expected 3-nt periodicity, confirming capture of translating ribosomes (Figure 1d). Stroke induced progressive changes in both mRNA abundance (RNA-seq) and ribosome occupancy (Ribo-seq) relative to sham controls (Figure 1e). Translational efficiency (TE; Ribo-seq normalized to RNA-seq, see Methods) was significantly altered at 1 h (320 genes; $|\log2FC| \geq 1$, FDR ≤ 0.05) and increased in intensity at later timepoints (927 genes at 6 h; 849 genes at 24 h) (Supplementary Figure 1a). Consistent with early transcription-translation decoupling, RNA-seq and Ribo-seq log2FC values were essentially uncorrelated at 1 h and 6 h (Figure 1f-g; Spearman's $\rho = -0.056$ and 0.109, respectively), but became more aligned by 24 h, although still modestly (Figure 1h; Spearman's $\rho = 0.345$). This suggests partial synchronization of transcriptional and translational programs as injury consolidates.

To interpret these temporal patterns, we performed pre-ranked GSEA using Gene Ontology Biological Process (GOBP) sets for each dataset at each timepoint. At 1 h, RNA-seq revealed induction of unfolded protein response pathways, consistent with early translational stress, and downregulation of differentiation/development programs (Supplementary Figure 1b). In contrast, Ribo-seq already showed enrichment of immune-related pathways (Supplementary Figure 1c), despite minimal overt immune-cell presence at this stage (Figure 1c). TE changes at 1 h highlighted pathways related to vascular function (e.g., regulation of coagulation) and transmembrane transport, alongside reduced bioenergetic and dendritic morphogenesis programs, consistent with early neuronal dysfunction (Supplementary Figure 1d). By 6 h, immune activation strengthened at both the transcriptional and translational levels (Supplementary Figure 1e-f), accompanied by marked suppression of bioenergetic/metabolic pathways (Supplementary Figure 1g) and broad perturbation of neuronal programs, cell-cycle processes, and pattern-recognition responses (Supplementary Figure 1e-g). At 24 h, immune signatures became dominant at both RNA and Ribo levels and neuronal pathway repression was more pronounced (Supplementary Figure 2a-b). In parallel, TE downregulation of translation-related pathways intensified, consistent with global translational impairment at later stages (Supplementary Figure 2c).

In summary, stroke induces a pronounced early mismatch between transcriptional and translational responses, with partial convergence by 24 h. Notably, immune pathway enrichment precedes overt neuronal loss and microglial accumulation, suggesting early activation of cell-intrinsic immune programs in response to acute ischemic stress.

➢ Stroke induces early synonymous codon usage changes and progressive ribosome pausing.

To characterize translational dynamics after stroke, we quantified ribosome A-site pausing. Pausing increased progressively from 6 h onward and affected nearly all codons, consistent with escalating translational stress (Figure 2a-c).We next assessed codon optimality and synonymous codon usage using two complementary metrics: GC3 scores and isoacceptors codon frequencies[22]. GC3, which captures third-position nucleotide bias (G/C versus A/T[22]) differed between up- and down-regulated genes only at 1 h after stroke and only in the Ribo-seq dataset (Figure 2d-f). Specifically, codon optimality showed a transient shift toward G/C-ending codons at 1 h, suggesting an early codon-selective translation program. Consistently, isoacceptors codon frequency analysis (synonymous codon usage) identified significant codon-biased translation only at 1 h and only at the translational level (Ribo-seq), with no comparable signal in RNA-seq or TE (Figure 2g).

Together, these data indicate that stroke elicits a brief hyperacute phase of codon-biased translation, potentially reflecting an acute oxidative/metabolic stress response, followed by a later phase dominated by widespread ribosome stalling that coincides with progressive disruption of translational efficiency, including reduced TE of translation-related pathways.

- **Stroke induces changes in ribosomal reading frames**

Altered ribosomal reading frames can introduce premature termination codons (PTCs) and trigger nonsense-mediated decay (NMD) (Figure 3a) or generate alternative proteoforms[23-25]. Ribosomal frameshifting has been linked to neurodegeneration[26], but has not been described in stroke. Given the marked perturbations in translation dynamics after stroke, we asked whether reading-frame usage changes over time. We assigned each RPF to a coding frame per gene in each sample, then computed a frameshift score (FS) reflecting deviation from the canonical 0 frame (see Methods). Using Limma-trend to compare each condition to sham, we detected a progressive increase in genes with significant shifts in dominant frame usage, emerging primarily at 6 and 24 h after stroke (Figure 3b–d, Supplementary Figure 3a). FS changes were bidirectional: some genes became more out-of-frame relative to sham (FS “UP”), whereas others became more in-frame (FS “DOWN”). We therefore focused subsequent analyses on the 6 h and 24 h timepoints. Genes with increased frameshifting were enriched for synaptic and energy metabolism pathways (Supplementary Figure 3b–c), whereas genes with decreased frameshifting were enriched for cell–cell adhesion/communication, protein localization, mitochondrial membrane organization and synaptic assembly (Supplementary Figure 3d–e). Codon-usage comparisons between FS UP and FS DOWN genes showed no major differences (Supplementary Figure 3f–h), although frameshifted genes were skewed toward specific G/C-ending codons relative to background (Supplementary Figure 3h).

To localize where frame changes occurred, we scanned the CDS of significantly frameshifted genes using an unsupervised hidden Markov model (HMM) and identified approximate windows where the dominant frame differed between stroke and sham (see Methods) (Figure 3e–f). At 6 h, we detected 228 loci in FS UP genes and 338 loci in FS DOWN genes: at 24 h, 232 and 420 loci, respectively. Some transcripts contained multiple loci. Notably, most significantly frameshifted genes did not yield a discrete locus, suggesting either diffuse frame disruption that is not captured by the HMM or, alternatively, changes driven by differential ORF usage. Across timepoints and directions, −1 frame shifts were most frequent (Figure 3g). We used the subset of localized loci for downstream consequences and feature analyses.

Using the HMM-defined windows, we modeled the impact of a frame shift on downstream coding output, focusing on PTC formation, predicted NMD susceptibility (PTC >50–55 nt upstream of the last exon–exon junction[27]), and the potential for SCRT-like C-terminal extensions via altered frame usage near stop codons. Across events, the dominant predicted outcome was PTC formation with inferred NMD susceptibility, followed by SCRT-like extension (Figure 3h–i). These results match canonical expectations for frameshifting and highlight the heterogeneity of its predicted consequences at transcriptome scale.

- **Frameshifting is largely decoupled from RNA/Ribo changes**

To test whether FS changes tracked with RNA abundance or ribosome occupancy, we correlated FS log2FC values with log2FC from RNA-seq, Ribo-seq and TE analyses. FS changes showed little to no correlation with gene expression, translation, or TE (Figure 3j, Supplementary Figure 4a), and overlap between FS-significant genes and conventional differential gene lists was limited (Supplementary Figure 4b–c). We next asked whether specific predicted fates (PTC, NMD, etc.) were associated with systematic shifts in RNA/Ribo metrics. K–S tests comparing log2FC distributions of fate-stratified genes to non-frameshifted genes identified modest, time- and dataset-specific effects (Supplementary Figure 4d). For example, NMD-prone events were associated with lower RNA-seq log2FC at 6 and 24 h, whereas at 6 h they were associated with higher TE log2FC. Overall, these effects were subtle and typically below thresholds for differential expression/translation, indicating that frameshifting impact cannot be inferred from RNA-seq/Ribo-seq summaries alone.

- **Sequence and structural features define frameshift loci**

To identify determinants of localized frameshifting, we trained XGBoost models with SHAP interpretation using features derived from FS loci, including RNAfold minimum free energy, amino-acid tracts, codon and nucleotide composition, entropy, CpG metrics, and others (see Methods). We trained models to classify (i) FS UP versus FS DOWN loci and (ii) FS loci versus background windows sampled either from the same transcript (to control for transcript context) or from global genomic background (see Methods). All tasks achieved strong discrimination (AUC ≥ 0.8) (Figure 3k). Across timepoints, FS loci exhibited characteristic sequence/structure signatures compared with transcript-matched and global backgrounds. FS loci were consistently enriched for higher GC and GC3 content and for specific codons encoding arginine and acidic residues (e.g., CGA/AGG/CGG, GAA/GAC), as well as embedded stop codons, at both 6 h and 24 h (Figure 4a–b, Supplementary Figure 5a–d). SHAP analyses ranked these features among the strongest contributors, indicating that localized frameshifting preferentially occurs in GC-rich contexts enriched for Arg/acidic codons and stop motifs. Determinants separating FS UP from FS DOWN loci also shifted over time. At 6 h, classification was driven mainly by CDS position (mid-CDS versus toward the end), nucleotide entropy, CpG observed/expected, and specific codon/amino acid composition (Figure 4c, Supplementary Figure 5c). The positional signal suggests FS UP events tend to occur earlier in the CDS than FS DOWN events (which are enriched in sham), potentially distinguishing stress-induced from baseline/regulated frame usage. By 24 h, local RNA structure (mfe), GC/GC3 content, hydrophobic amino-acid content, and defined codon/AA motifs (e.g., ACG, CCG, Thr, Gln, Tyr) became dominant predictors (Figure 4d, Supplementary Figure 5d), consistent with later selective maintenance or suppression of structurally and compositionally distinct FS-prone regions.

To probe whether frameshifting is associated with altered protein abundance, we examined several significantly frameshifted genes by western blot across time. Sephs2, which showed increased frameshifting after stroke, exhibited a modest (non-significant) reduction in protein abundance at 6 and 24 h (Figure 4e), despite no significant change in RNA-seq, Ribo-seq, or TE (Supplementary Figure 5e). Given prior evidence linking Sephs2 to neuronal oxidative stress resistance[28], this trend may reflect reduced stress-buffering capacity. Smad2 showed increased frameshifting but no major protein change (Figure 4f); it was modestly downregulated at 24 h in Ribo-seq and TE (log2FC = −0.4267 and −0.527; FDR < 0.05; Supplementary Figure 5f) but not at the RNA level and not earlier, despite significant frameshifting at 6 h. Dcx was markedly reduced at 24 h, coinciding with increased frameshifting (Figure 4g) and decreased RNA expression (Supplementary Figure 5g); TE increased but did not reach significance. Overall, we did not observe a simple, consistent relationship between frameshifting and steady-state protein abundance, mRNA levels, or TE.

Taken together, these results show that stroke induces robust, time-dependent perturbations in ribosomal frame usage. A subset of events localizes to discrete CDS loci with identifiable sequence/structure determinants and diverse predicted consequences, whereas many gene-level frame shifts likely reflect diffuse disruption and/or alternative ORF usage. Importantly, the proteomic impact of frameshifting is not captured by conventional RNA/Ribo fold changes, underscoring the need for locus- and ORF-aware analyses to resolve translation-derived proteome remodeling during stroke.

- **Stroke induces hyperacute stop codon readthrough events caused by termination failure**

Because frameshifting can generate SCRT-like C-terminal extensions (Figure 3h), we next asked whether stroke induces stop codon readthrough (SCRT). SCRT is a well-documented but comparatively understudied phenomenon[29]. It has been most extensively characterized in simpler organisms[30,31], although several functional mammalian examples exist[32-34]. Despite this, large-scale efforts to quantify SCRT dynamics in multicellular physiology or disease remain limited. Here, we quantified SCRT after stroke using a custom pipeline that identifies bona fide 3′UTR RPFs and models 3′UTR-to-CDS ribosome occupancy using a GLM–beta-binomial framework to detect differential SCRT relative to sham controls (Figure 5a). Metagene analysis revealed clear 3-nt periodic P-site signal in the 3′UTR, most prominently 1 h after dMCAO (Figure 5b).

Differential SCRT analysis uncovered a striking hyperacute burst at 1 h, with 815 genes showing increased SCRT relative to sham and only 36 showing decreased SCRT (Figure 5c). From 6 h onward, SCRT changes were more balanced, with a shift toward predominately decreased SCRT by 24 h (Figure 5d–e). Pathway enrichment indicated that genes with increased SCRT at 1 h were enriched for synaptic and neuronal functions (Figure 5f). At 6 h, genes with reduced SCRT were enriched for synaptic pathways, whereas genes with increased SCRT were enriched for metal ion transport, axon myelination, and metabolic processes (Supplementary Figure 6a–b). By 24 h, reduced SCRT was associated with higher cognition and synaptic functions, whereas increased SCRT was enriched for synaptic activity and nervous system development pathways (Supplementary Figure 6c–d). Together, these results argue that stroke-induced SCRT is pathway-structured rather than random, consistent with a regulated contribution to post-stroke proteome output.

239 ➢ **SCRT is largely independent of CDS frameshifting and reflects hyperacute termination**
240 **stress**

241 We next examined the mechanisms underlying SCRT, particularly at 1 h when the signal
242 peaked. Because frame changes near the stop codon could produce apparent readthrough, we first
243 assessed reading-frame usage patterns in 3′UTRs. Genes with significant SCRT showed subtle
244 shifts in 3′UTR frame distributions after stroke (Supplementary Figure 6e). However, across
245 samples, SCRT rate (Calculated as Ribosome Readthrough Score; RRS) and CDS frameshifting
246 (FS) were uncorrelated (Supplementary Figure 6f), indicating that gene-level CDS frameshifting is
247 not the primary driver of SCRT after stroke, despite the fact that some localized FS events can yield
248 SCRT-like consequences (Figure 3h). Consistent with this, the major SCRT burst occurred at 1 h,
249 before robust CDS frameshifting is observed, supporting the view that these are distinct (though
250 potentially overlapping) phenomena in stroke.

251 We also evaluated whether changes in release factor abundance could explain the
252 hyperacute SCRT surge. eRF1 protein levels did not show major changes early after stroke, with an
253 observable upregulation 24 hours after stroke induction (Figure 5g), arguing against release-factor
254 depletion as an explanation for the observed SCRT burst at 1 hour. More broadly, genes with
255 increased SCRT at 1 h displayed a sharp pre-stop ribosome peak compared to a matched-coverage
256 background set (Supplementary Figure 7a), a pattern absent in sham (Supplementary Figure 7b).
257 This pre-stop aggregation was not a dominant feature at later time points (Supplementary Figure 7c–
258 f), consistent with a transition from hyperacute, stress-driven termination failure to later, more
259 balanced and potentially regulated readthrough dynamics.

260 ➢ **Sequence context and local 3′UTR structure shape SCRT dynamics**

261 To explore how sequence and structure influence SCRT, we first compared synonymous
262 codon usage in SCRT-significant genes to genome-wide background using Welch's t-test. The
263 strongest codon usage bias occurred at 1 h, where genes with increased SCRT preferentially used
264 A/T-ending codons (Figure 6a). Codon usage bias persisted at 6 h and 24 h, but the dominant signal
265 shifted: genes with decreased SCRT overused G/C-ending codons (Figure 6a). Because 3′UTR
266 structure downstream of the stop codon can modulate readthrough[35], we evaluated local RNA
267 secondary structure in the first 100 nt of the 3′UTR using RNAfold. At 1 h, genes with increased
268 SCRT showed significantly different 3′UTR minimum free energy (MFE) relative to background,
269 whereas genes with decreased SCRT did not (Figure 6b–c). These structure differences were not
270 observed at 6 h or 24 h (Supplementary Figure 8a–b), again suggesting mechanistically distinct early
271 versus late phases of SCRT regulation after stroke.

272 ➢ **Machine learning identifies termination stress and 3′UTR frame usage as dominant**
273 **SCRT determinants**

274 Given the number of potentially interacting factors, we applied a machine learning (ML)
275 framework to identify determinants of SCRT across time. After benchmarking generalized linear
276 models, random forests, and XGBoost, XGBoost performed best for SCRT classification (not
277 shown). We trained XGBoost models with SHAP interpretation using transcript features at H1, H6,
278 and H24 after ischemia. Model performance was highest at H1 and H6 (AUC = 0.80 and 0.79,
279 respectively) and declined at H24 (AUC = 0.66) (Figure 6d), consistent with SCRT being most
280 predictable during the early stress response. Across timepoints, the strongest predictors were the

281 fractions of 3′UTR P-sites in each of the three downstream reading frames, accounting for ~40–60%
282 of total model gain (Figure 6e, Supplementary Figure 8c–d). SHAP analyses confirmed these 3′UTR
283 frame-occupancy signatures as dominant global contributors, consistent with ribosomes entering
284 specific 3′UTR reading frames during termination stress.

285 At H1, pre-stop ribosome density was the single strongest SHAP feature (Figure 6f), supporting
286 a model in which acute termination failure and ribosome queuing drive hyperacute SCRT. This effect
287 weakened by H6 and was nearly absent by H24 (Supplementary Figure 8e–f), consistent with partial
288 recovery of termination fidelity over time. Additional informative features included 3′UTR length, CDS
289 length, GC3 scores, and 3′UTR MFE, suggesting that longer and more structured transcripts with
290 distinct elongation kinetics are more permissive to SCRT. The contribution of GC3 also aligns with
291 the A/T-biased codon usage in genes with increased SCRT at 1 h (Figure 6a). By contrast, stop
292 codon identity contributed minimally (<1% gain at all timepoints), indicating that stroke-induced
293 SCRT is not primarily driven by canonical readthrough motifs (e.g., UGA-CUAG)[36] but rather by
294 stress-dependent modulation of termination and downstream ribosome behavior.

295 Together, these results support a model in which stroke induces SCRT through hyperacute
296 disruption of translation termination, followed by selective, frame-dependent ribosome occupancy in
297 the 3′UTR. As injury progresses, SCRT becomes more heterogeneous and less predictable,
298 consistent with dynamic remodeling and additional regulatory inputs. Importantly, SCRT changes at
299 1 h showed no significant correlation with RNA-seq or Ribo-seq differential signals (Supplementary
300 Figure 8g–i), indicating that the hyperacute SCRT burst is largely orthogonal to conventional
301 expression/translation changes. At 6 h and 24 h, the strongest association was a modest negative
302 correlation between SCRT and Ribo-seq log2FC ($\rho = -0.289$ and -0.282, respectively; $p < 0.0001$;
303 not shown). To assess functional coupling to downstream regulation, we compared RNA-seq, Ribo-
304 seq, and TE log2FC distributions for SCRT-altered genes using Wilcoxon rank-sum and
305 Kolmogorov–Smirnov tests with Benjamini–Hochberg correction (Supplementary Figure 9). SCRT
306 changes had no detectable effect on RNA abundance (Supplementary Figure 9a) but were
307 associated with changes in translation (Supplementary Figure 9b) beginning at 6 h and in TE
308 (Supplementary Figure 9c) beginning at 1 h. Increased SCRT tended to reduce TE, whereas
309 decreased SCRT tended to increase TE, indicating that SCRT-associated effects are subtle yet
310 remain primarily translational.

311 ➢ **Stroke induces changes in ORF dominance and translation**

312 We next examined how stroke alters open reading frame (ORF) usage and translation. Our
313 gene-level frameshifting analysis indicated that although thousands of genes showed altered frame
314 usage, discrete frameshift loci could be localized in only a subset, suggesting that much of the signal
315 may reflect altered ORF usage rather than ribosomal slippage. We therefore used RiboCode[37] to
316 identify translated ORFs in each sample. ORFs were filtered for length ≥10 amino acids (≥30 nt), P-
317 sites ≥20, $p \leq 0.05$, and ≥50% of reads in frame 0. Across all samples we detected ~40,000
318 translated ORFs, dominated by annotated ORFs, followed by novel ORFs and downstream ORFs
319 (dORFs) in the 3′UTR (Figure 7a). Annotated ORFs were the longest (mean 660 aa; median 504 aa)
320 (Figure 7b). dORFs and overlap-dORFs were the next longest, whereas novel ORFs were ~one-third
321 the length of annotated ORFs on average (Figure 7b).

322 We first performed an ORF-level differential analysis using total P-site counts per ORF as input.
323 At 1 h, most significant ORFs were downregulated with few upregulated (Figure 7c). At 6 h, we

detected essentially no differentially translated ORFs (Figure 7d). In contrast, 24 h showed robust ORF-level remodeling with many ORFs significantly upregulated (Figure 7e). In both the 1 h and 24 h comparisons, annotated ORFs were the most frequent differentially translated class, followed by novel ORFs and dORFs (Supplementary Figure 10a–b). Gene Ontology (GOBP) over-representation analysis of parent genes revealed that ORFs downregulated at 1 h were enriched for growth hormone signaling, focal adhesions, and cell junction pathways (Supplementary Figure 10c), and KEGG analysis highlighted axon guidance as the top pathway (Supplementary Figure 10d), linking early ORF suppression to neuronal circuitry. By 24 h, downregulated ORFs were enriched for neuronal and synaptic functions (Figure 7f), whereas upregulated ORFs were enriched for apoptotic signaling, regulation of protein metabolism, phagocytosis, and cell–substrate adhesion (Figure 7g), consistent with concurrent neuronal dysfunction/death and immune activation/clearance as injury consolidates.

- **Stroke drives progressive differential ORF usage (DOU) independent of RNA abundance**

Because a single gene can encode multiple translated ORFs (and multiple isoforms can expand this space), we next tested whether stroke shifts dominant ORF usage. We defined “canonical” ORFs as those accounting for ≥60% of all P-sites attributed to a gene within a sample. We then used DRIMSeq[38] to detect genes with significant shifts in canonical ORF usage. Differential ORF usage (DOU) increased progressively from 1 h and peaked at 24 h (Figure 7h). Notably, DOU increased at 6 h relative to 1 h despite the absence of ORF-level differential translation at 6 h (Figure 7c–d), indicating that ORF remodeling can occur even when total ORF translation changes are minimal.

Across DOU events, annotated and novel ORFs dominated, reflecting their prevalence in the discovery set, followed by dORFs. Unlike our SCRT analysis, switching toward dORF dominance was significantly increased at 24 h, suggesting regulated downstream initiation rather than deregulatory readthrough. To determine whether DOU tracked with conventional regulatory modes, we tested enrichment of DOU among genes with significant RNA-seq and/or Ribo-seq changes using Fisher’s exact tests with FDR correction. At 1 h and 6 h, DOU was not significantly enriched in any regulatory mode (Supplementary Figure 10e). At 24 h, DOU was modestly enriched among genes without significant RNA or Ribo changes (OR = 1.5, q = 0.017; 589 DOU genes not significant), whereas transcription-only and combined RNA+Ribo modes were depleted for DOU (OR = 0.16 and 0.1; q = 6.27E−07 and 0.015, respectively), although the absolute counts were small (5/1075 and 1/340). These results indicate that ORF usage shifts are largely orthogonal to changes in mRNA abundance or bulk ribosome occupancy. Consistent with this interpretation, ~30% of genes with significant DOU at 24 h also scored significant in our frameshifting analysis (Figure 7i), indicating that gene-level frame changes capture a substantial component of ORF switching after stroke.

- **uORF engagement increases at 24 h and associates with reduced coding translation**

Within the DOU results, we observed an increase in uORF usage at 24 h. uORFs can repress downstream CDS translation[39,40]. We therefore stratified genes showing increased uORF usage, decreased uORF usage, or switching to uORF dominance (Figure 7j). To relate these patterns to gene regulation, we compared RNA-seq, Ribo-seq, and TE log2FC distributions for uORF-changing genes versus stable genes using Wilcoxon rank-sum and Kolmogorov–Smirnov tests with Benjamini–Hochberg correction. The strongest association was at the translational level: genes with

either increased or decreased uORF usage showed lower median Ribo-seq log2FC than stable genes beginning at 6 h (Supplementary Figure 11a). TE effects were most pronounced for uORF-increased genes, which showed reduced TE at 24 h (Supplementary Figure 11b). RNA-seq effects were weakest and only evident at 24 h (Supplementary Figure 11c). Although the number of genes switching toward uORFs was modest, these genes were enriched for synaptic functions at 24 h (not shown), supporting biological specificity.

- **Gene-level upstream translation increases progressively and represses CDS translation**

To complement ORF calling, we quantified upstream translation at the gene level using a GLM–beta-binomial framework analogous to SCRT, comparing 5′UTR P-sites to CDS P-sites. This analysis revealed progressive changes in upstream translation that peaked at 24 h (Figure 8a–c), mirroring the ORF-based remodeling. Clustering of the upstream translation ratio (UPR; 5′UTR/CDS reads per gene) separated late timepoints from sham and 1 h (Figure 8d). Pathway analyses highlighted enrichment for neuronal and synaptic programs (Figure 8e–f, Supplementary Figure 12a–d). We also observed overlap between genes with significant frameshifting at 24 h and genes with significant upstream translation changes (Figure 8g; significance: $|\log2FC| \geq 1$, FDR ≤ 0.05).

To visualize locus-level examples, we examined Dcx and Smad2, both of which showed frameshift and Dcx showed reduction in protein abundance (Figure 4f–g). For Dcx, 1 h showed a reduced 5′UTR P-site peak relative to sham (Figure 8h), followed by the strongest relative increase at 6 h (Figure 8i) that persisted at 24 h (Figure 8j). Smad2 showed a similar pattern, with 5′UTR peaks comparable to sham at 1 h and increased peaks from 6 h onward (Supplementary Figure 12e–g). Finally, consistent with uORF-mediated gating, differential upstream translation (DUT) was associated with reciprocal shifts in CDS translation and TE: genes with increased DUT showed lower Ribo-seq and TE log2FC values (and vice versa), detectable from 1 h onward (Supplementary Figure 13a–b). No consistent association was observed with RNA abundance (Supplementary Figure 13c), indicating that upstream translation remodeling primarily acts at the translational level.

Collectively, these results indicate that stroke induces a time-dependent ORF remodeling program that is modest in the hyperacute phase but becomes prominent from 6–24 h. This program includes shifts in dominant ORF usage, increased uORF engagement, and rising upstream translation that inversely tracks CDS translation and TE (without parallel RNA changes), consistent with translational gating that diversifies coding outputs and contributes to proteome remodeling during stroke.

Discussion

In this work, we provide a time-resolved analysis of translational remodeling during stroke that reveals a staged translational response program not apparent from RNA abundance analysis alone. We show that changes in mRNA expression are largely disconnected from mRNA translation during the hyperacute and acute phases after stroke. We also identify a short-lived, codon-biased translational program in the hyperacute phase, distinct from what is observed in cellular oxidative stress models[16,41]. In parallel, we detect pervasive stop codon readthrough events consistent with translation termination failure, with the potential to alter C-terminal structures of neuronal and synaptic proteins before major changes in their expression. At later stages, we observe broad changes in ribosomal reading frames. A subset of frameshifting events can be localized to discrete CDS regions, consistent with ribosomal slippage triggered by stalling or collision, coinciding with progressive A-site pausing. These localized events often introduce downstream premature termination codons (PTCs) and, in some cases, predict nonsense-mediated decay. However, most frameshifting signals could not be localized, and overlapped substantially with progressive alternative ORF usage, suggesting that differential ORF selection rather than slippage explains the majority of apparent frame changes. Finally, using two orthogonal approaches, we observe progressive increases in uORF usage and upstream translation initiation. Increased upstream translation (increased 5′UTR RPFs) is inversely associated with CDS translation and translational efficiency, consistent with a regulatory role[39,40]. Collectively, these findings indicate that translational changes in the ischemic brain extend well beyond mRNA abundance, reflecting complex regulatory and deregulatory processes that shape translational efficiency and proteome stability and diversity. To our knowledge, these phenomena have not been systematically explored in stroke, and rarely in mammalian disease contexts. Thus, our analysis provides a framework for extracting hidden translational mechanisms from Ribo-seq datasets, applicable beyond stroke and neurological disorders.

Despite modern improvements in stroke therapy, stroke remains the second leading cause of morbidity and mortality worldwide[4], underscoring the need for novel neuroprotective strategies[42]. A major barrier to therapeutic development is an incomplete understanding of the evolving molecular dynamics of stroke pathophysiology. Stroke is temporally dynamic and involves interactions among multiple cell types[5]. Nevertheless, stroke molecular studies remain dominated by transcriptomics and proteomics[43,44], with recent shifts toward single-cell and spatial transcriptomics[45,46]. While these approaches are crucial for defining cellular states, they do not resolve the mechanisms controlling mRNA translation, particularly under acute oxidative and metabolic stress[47,48], conditions central to ischemic injury. Our observation of transcription–translation discordance further emphasizes the importance of studying translation regulation directly. To date, only a few studies have examined translation in stroke[49,50], primarily through translatome capture in Ribo-tag mice[49,50]. Although informative for cell-specific ribosome loading, these approaches lack nucleotide-resolution ribosome footprinting and therefore cannot capture the breadth of translational regulatory and deregulatory phenomena uncovered here.

Using ribosome profiling, we identify multiple translational behaviors that are not accessible to transcriptomics or standard proteomics. Importantly, these phenomena change dynamically over time, revealing a nuanced landscape as neurons respond to acute ischemic stress, remodel the translatome, and ultimately fail. In the hyperacute phase, we observe stress-responsive codon-biased translation with a shift toward G/C-ending codons[12,16] and a pronounced spike in stop codon readthrough. SCRT generates C-terminally extended proteoforms and has been described in viral

infection, hypoxia, and oxidative stress, particularly in prokaryotes and yeast, but has not been systematically studied in multicellular organisms or in acute brain injury[51-53]. Our data reveal widespread readthrough acutely after stroke, consistent with an early stress response[54] followed by a shift toward more balanced readthrough patterns at later timepoints, suggestive of additional regulatory processes. ML analysis implicates termination failure and altered ribosome behavior around stop codons as key drivers of hyperacute SCRT[54,55], with contributions from 3′UTR structure[35]. SCRT has been proposed as a mechanism that expands proteome diversity[56], and functional examples exist in neurons, including AQP4 and its readthrough isoform AQP4ex[33]. We detect later SCRT remodeling in neuronal and synaptic genes, with both increases and decreases relative to sham. However, these later dynamics differ from the hyperacute burst, which appears dominated by termination stress. Notably, SCRT-associated genes exhibit significant changes in translational efficiency without detectable changes in mRNA expression, reinforcing that SCRT is primarily translational. Whether these events yield stable proteoforms and alter gene function will require targeted proteomics and functional studies.

After the hyperacute phase, we observe progressive frameshifting beginning 6 h after stroke. A subset of events can be localized to discrete CDS loci, consistent with ribosome slippage or stalling and/or collision, coinciding with progressive A-site pausing[57]. Ribosome stalling and collisions canonically trigger the ribotoxic stress response[47,58,59], which limits further ribosome loading. We did not detect overt activation of this pathway in the stroke model(data not shown), suggesting that collision signaling may be subtle, transient, or restricted to specific transcripts, consistent with the fact that localizable frameshift loci are detected in only a subset of frameshifted genes. ML analyses further indicate that frameshifting is strongly shaped by local sequence features, including nucleotide and codon composition, amino-acid content, and RNA structure, and distinguish loci with increased versus decreased frameshifting relative to sham. Consequence modeling predicts frequent downstream PTC formation and potential NMD engagement. However, most gene-level frameshifting signals could not be mapped to discrete loci, raising the possibility that frame changes are diffuse or, more plausibly, reflect alternative ORF usage. Indeed, differential ORF usage (DOU) increases progressively with stroke, and overlaps substantially with frameshifting calls, supporting the interpretation that alternative ORF selection explains most apparent frame disruption, whereas collision/slippage-driven frameshifting represents a minority subset in stroke. Consistent with this remodeling, our ORF analyses reveal progressive uORF engagement. uORFs are established regulators of downstream translation[39,40]. Using a GLB-beta-binomial model, we identify genes with increased 5′UTR RPF proportions relative to the CDS after stroke, enriched for neuronal and synaptic pathways. Importantly, increased or decreased upstream translation correlates inversely with translational efficiency, consistent with prior work linking upstream initiation to suppression of CDS translation[39,40,60,61].

Collectively, our data expand the molecular framework of ischemic stroke by placing translation control and translation failure mechanisms alongside transcriptional programs and inflammatory remodeling. Our results support a temporally ordered cascade of termination stress, elongation/collision stress, and ORF rewiring—translation-layer mechanisms extensively characterized in cell stress models but rarely mapped in vivo in mammalian disease. More broadly, our study highlights the need for complementary tools to validate these mechanisms in stroke and other disorders, and suggests that orthogonal biochemical, genetic, and proteomic approaches will be essential to identify regulators and establish causal links.

Several limitations should be considered. First, ribosome profiling was performed on bulk infarct-containing tissue; thus, observed changes may reflect both cell-intrinsic regulation and evolving cell-type composition, and future cell-type-resolved translatomics or spatial approaches will refine attribution. Second, while SCRT-like signals and 3′UTR ribosome occupancy are consistent with termination failure, rescue, and/or impaired recycling, ribosome footprints alone cannot distinguish productive readthrough that yields stable C-terminally extended proteoforms from rescue-associated post-termination ribosomes or scanning/reinitiation in 3′UTRs, as documented in other systems[55]. Third, localized out-of-frame signatures are consistent with slippage/collision at specific loci, but we do not directly measure collisions or ribotoxic signaling; disome profiling or biochemical readouts would strengthen mechanistic specificity[48,62]. Fourth, consequence annotation (e.g., PTC/NMD susceptibility) and ORF assignment depend on transcript models and general NMD rules and should be interpreted as predicted propensity rather than direct measurements; isoform-resolved transcription/decay profiling and targeted validation would refine these predictions[63]. Finally, we focus on discrete timepoints in a single model, using adolescent mice, and only in male sex; expanded time courses, reperfusion paradigms, additional regions, female stroke models, using aged mice, and other stroke subtypes will be important to generalize the staged model.

In summary, time-resolved integration of RNA-seq and ribosome profiling shows that translational remodeling during ischemic stroke is not a uniform suppression of protein synthesis, but a staged program evolving from early termination/3′UTR ribosome stress and codon-biased translation to later elongation stress, out-of-frame decoding signatures, and progressive rewiring of ORF and uORF usage. These layers help explain why transcript abundance incompletely predicts translational output during acute brain injury and highlight translational control nodes, termination/recycling, collision surveillance, and ORF selection, that may shape neuronal proteostasis and fate as ischemia progresses. More broadly, our framework demonstrates how nucleotide-resolution ribosome footprinting can reveal hidden translational regulatory and deregulatory mechanisms in vivo, providing a foundation for mechanistic testing across diverse pathologies beyond stroke.

517 **Methods**

518 ➢ **Animals**

519 Distal middle cerebral artery occlusion (dMCAO) was induced on the left side in 12-week-old, 20-25
520 g male C57BL/6JJmsSlc mice as previously described[64]. Animals were bought from a local vendor
521 (Kumagai-Shoten Sendai) and housed in a controlled environment with a 12 h light/dark cycle and
522 temperature regulated at 23°C. Animals were allowed to acclimatize for at least 1 week before
523 surgery. All experiments were conducted according to protocols approved by the animal care facility
524 of Tohoku University (Ethical approval number 2024-Ikoudou-005-01), the Animal Research:
525 Reporting of In Vivo Experiments guidelines, and the Code of Ethics of the World Medical
526 Association.

527 ➢ **dMCAO surgery**

528 The permanent dMCAO model was performed according to previous reports[64,65]. Briefly, mice were
529 anesthetized with 2% isoflurane (30% oxygen, 70% nitrous oxide). Rectal temperature was
530 maintained at 3°C using a heating pad with feedback control during surgery. The temporalis muscle
531 between the left eye and ear was retracted along a 4-mm skin incision. A small bone window (~1 mm
532 diameter) was drilled to expose the distal MCA. The MCA was cauterized using bipolar
533 electrocoagulation distal to the lenticulostriate branches. The incision was closed with 6-0 nylon
534 sutures. Animals were monitored in a cage placed on a 3°C -heating plate for 1 h during anesthesia
535 recovery. The left common carotid artery (CCA) was temporarily occluded with 6-0 nylon immediately
536 before MCA cauterization and was released after 1 h.

537 Animals were euthanized at 1 h, 6 h, or 24 h after MCA cauterization. Sham-control animals
538 underwent the same surgical procedure without MCA cauterization and CCA ligation and were
539 euthanized after 1 h.

540 ➢ **TTC stain**

541 At the designated time points after induction of dMCAO, mice were euthanized using isoflurane
542 overdose inhalation followed by trans-cardiac perfusion with ice-cold saline to flush out the blood.
543 Brains were dissected and 2-mm coronal sections of the brains were cut on a brain matrix and
544 stained with 0.1% 2,3,5-triphenyltetrazoliumchloride (TTC; Sigma, Cat# T8877) dissolved in normal
545 saline at 37°C for 15 min. Images were acquired using a flat-head scanner.

546 ➢ **Western blotting**

547 At the designated time points, animals were sacrificed and trans-cardiac saline perfusion
548 performed. Brains were dissected and the infarct area extracted and homogenized in T-PER tissue
549 protein extraction reagent (Thermo Fisher, Cat# 78510) containing cOmplete protease inhibitor
550 cocktail (Roche, Cat# 4693116001) and PhosSTOP phosphatase inhibitor (Roche, Cat#
551 4906837001). After centrifugation, the proteins were isolated from the supernatant and measured
552 using Pierce BCA Protein Assay Kits (Thermo Fisher Scientific, Cat# 23227). Next, on Mini-
553 PROTEAN TGX Precast Protein Gels (Bio-Rad, Cat# 4561096), equal protein loads were separated
554 and then transferred to Trans-Blot Turbo Mini 0.2 μm PVDF Transfer Packs (Bio-Rad, catalog
555 #1704156). Afterward, membranes were blocked with 5% skim milk (Fujifilm Wako, Cat# 190-12865)
556 in phosphate-buffered saline with Tween 20 (PBS-T) (Takara, Cat# T9183). The primary antibody
557 was then incubated overnight at 4°C, followed by a one-hour incubation at room temperature with

558 Western BLoT Rapid Detect v2.0 (Takara, Cat# T7122A). Signal detection was performed using
559 Pierce ECL substrate reagent (Thermo Fisher, Cat# 32106) on a ChemiDoc Touch system (Bio-
560 Rad). The detected bands were analysed using Image Lab Software (BioRad). Beta-actin was used
561 as a loading control. To reprobe the membrane, Restore stripping buffer (Thermo Fisher, Cat#
562 21059) was utilized.

563 **List of antibodies used:**

564 beta-Actin Rabbit Monoclonal (1:3000, Cell signaling, Cat# 4970).

565 eRF1 Rabbit Polyclonal (1:1000, Invitrogen, Cat# PA5-28777)

566 Sephs2 Rabbit Polyclonal (1:1000, Invitrogen, Cat# PA5-27950)

567 Doublecortin (DCX) Rabbit Polyclonal (1:500, Abcam, Cat# ab18723)

568 Smad2 Rabbit Polyclonal (1:1000, Invitrogen, Cat# PA5-88144)

569 ➢ **Immunofluorescence staining**

570 Mice were euthanized by aspirating excess isoflurane at the designated timepoints after dMCAO
571 procedure. Mice were transcardially perfused with saline to remove the blood and subsequently with
572 4% paraformaldehyde in phosphate buffer (Nakarai Tesque, Cat# 09154-85). The fixed middle third
573 of the brain, including the ischemic region, was removed and embedded in OCT compound (Sakura
574 Finetek, Cat# 4583) to be frozen by liquid nitrogen. The prepared brain was then sectioned at a 10
575 µm thickness with a cryostat. The sections were washed with 1 x PBS, incubated with blocking buffer
576 with 0.8% Block Ace (UKB, Cat# UK880), 1% BSA (Fujifilm Wako , Cat# 010-25783), and 0.3%
577 Triton X-100 (Nacalai Tesque, Cat# 35501-15) in PBS for 1 h at room temperature before adding
578 primary antibodies (details below) diluted in blocking buffer, and incubated overnight at 4 °C.
579 Subsequently, they incubated with conjugated secondary antibodies of AlexaFluor 568 (1:500;
580 Invitrogen, Cat# A11041, A10042, and A10037) and AlexaFluor 647 (1:500; Invitrogen, Cat# A21244,
581 A21136, and 21244) in antibody-dilution buffer with 0.2% Block Ace, 0.2% BSA and 0.3% Triton X-
582 100 in PBS and DAPI (1:500; Invitrogen, Cat# D1306) was used for counterstaining. Samples were
583 inspected under a laser confocal microscope (OLYMPUS, FV3000).

584 **Primary antibodies used:**

585 Iba-1 Rabbit Polyclonal (1:500, Fujifilm Wako, Cat# 019-19741)

586 NeuN Rabbit Monoclonal (1:1500, Abcam, Cat# ab177487)

587 GFAP Rabbit Monoclonal (1:200, Invitrogen, Cat# 130300)

588 ➢ **Ribo-seq library preparation:**

589 Ribo-seq was conducted as previously described[28,66]. Briefly, at the designated time point,
590 animals were sacrificed and perfused using ice-cold PBS containing 100 µg/ml cycloheximide
591 (Sigma, Cat# 01810), followed by extraction of the brain. Brains were extracted and sectioned into
592 three equal parts in the sagittal plane. The leftmost third of the brain, corresponding to Bregma -2
593 mm to +2 mm, where the infarct occurs, was immediately snap-frozen in liquid nitrogen. Dissection
594 was conducted on ice, and all tools were decontaminated with RNase AWAY(Thermo Fisher, Cat#
595 7002PK) prior to use. Tissues were homogenized in polysome lysis buffer (20 mM Tris-Cl, pH 7.5,

596 150 mM NaCl, 5 mM $MgCl_2$, 1 mM dithiothreitol (Sigma, Cat# D9779), 100 µg/ml cycloheximide, and
597 1% Triton X-100) and centrifuged at 13000 x g at 4°C for 15 min. Ribosome foot-printing was
598 performed on the supernatant collected by adding 1.25 U/µl RNase I (NEB, Cat# M0243) to 500 µl
599 clarified lysate and incubating samples on a rotator mixer for 45 min at room temperature. QIAzol
600 (QIAGEN, Cat# 79306) reagent was added, and RNA was extracted using miRNeasy Mini Kit
601 (QIAGEN, Cat# 217004). Ribosome-protected fragments (RPFs) were selected by isolating RNA
602 fragments of 27-35 nucleotides (nt) using TBE-Urea PAGE gel. rRNA depletion was conducted using
603 NEBNext rRNA depletion kit v2 (NEB, Cat# E7400) followed by end-repair using T4 PNK after
604 purification of the rRNA depleted samples using Oligo Clean & Concentrator Kits (Zymo Research,
605 Cat# D4060). The preparation of sequencing libraries for ribosome profiling was conducted via the
606 NEBNext Multiplex Small RNA Library Prep kit for Illumina according to the manufacturer's protocol
607 (NEB, Cat# E7300S). Library quality was assessed on an Agilent Bioanalyzer 2100 using the Agilent
608 DNA 1000 kit (Cat# 5067-1504). Pair-end sequencing reads of size 150 bp were produced for Ribo-
609 seq on the Illumina HiSeq X-ten system. Four biological replicates were prepared per group.

610 ➢ **Ribo-seq data analysis**

611 Quality control for Raw Fastq was performed using FastQC. Next, adapter trimming and collapsing
612 the pair-end reads into one was done using Seqprep. Trimmomatic[67] was used to further clean low-
613 quality reads. Bowtie2[68] was used to align the reads to a reference of rRNA and tRNA genes (mm10)
614 to remove contaminants. After that, reads were aligned to the genome (mm39 downloaded from
615 UCSC), and reads counted using FeatureCounts[69] with the tag "-t CDS" to generate counts of protein
616 coding genes only. RiboWaltz[70] was used for metagene analysis, P-site offset determination, and
617 quality control of Ribo-seq samples. Pausepred offline was used for ribosome pause sites
618 determination (https://github.com/romikasaini/Pausepred_offline). Visualization of Pausepred results
619 was performed in *R* (version 4.5.0).

620 ➢ **RNA-seq library preparation:**

621 The ischemic brain regions described in Ribo-seq section, or anatomically matched areas from
622 sham animals, were lysed using QIAzol Lysis Reagent (QIAGEN, Cat# 79306), and total RNA was
623 extracted with the miRNeasy Mini Kit (QIAGEN, Cat# 217004) according to the manufacturer's
624 instructions. RNA purity and concentration were measured using a NanoDrop spectrophotometer
625 (Thermo Fisher, Cat# ND-ONE-W), and RNA integrity was assessed using the RNA 6000 Nano Kit
626 (Agilent, Cat# 5067-1511) on the Agilent 2100 Bioanalyzer. Only samples with an RNA Integrity
627 Number (RIN) of 7 or higher are used for RNA sequencing (RNA-seq).

628 RNA-seq was performed using three biological replicates per group, with mRNA enriched using
629 the NEBNext Poly(A) mRNA Magnetic Isolation Module (NEB, Cat# E7490) and libraries prepared
630 using the Ultra II Directional RNA Library Prep Kit for illumina (NEB, Cat# E7760) according to the
631 manufacturer's instructions. Library quality was assessed on an Agilent Bioanalyzer 2100 using the
632 Agilent DNA 1000 kit (Cat# 5067-1504), and library concentrations were quantified with the NEBNext
633 Library Quant Kit for Illumina (NEB, Cat# E7630). Sequencing was conducted on an Illumina
634 NovaSeq X Plus platform using 150 bp paired-end reads. Raw Fastq files underwent quality control
635 with FastQC, followed by trimming with Trimmomatic[67] to remove adapter sequences and low-quality
636 reads. Clean reads were aligned to the Mus musculus (mm39) reference genome (UCSC) using the
637 splice-aware aligner HISAT2[71], and gene-level counts were generated with FeatureCounts[69] using
638 UCSC Refseq annotation.

➢ Differential gene expression and translational efficiency analysis

Differential gene expression and translational efficiency analysis were performed in *R* using DESeq2[72] and (https://github.com/smithlabcode/riborex) with Limma-voom based approach respectively. Pathway enrichment analysis, correlation analysis, and codon analysis of sequencing datasets were all performed using *R* language programming.

➢ GC3 and isoacceptors codon analysis

Isoacceptors frequency and total codon frequency were analyzed as we previously reported[28,66,73,74]. We used the protein coding transcript fasta from Gencode (mouse V37) to extract the coding sequences. We selected the longest transcript per gene as representative transcript for codon analysis. We used the R package coRdon to generate codon tables.

GC3 scores for each gene were calculated using this formula:

$$GC3 = \frac{N_{GC3}}{N_T}$$

Where N_{GC3} is the number of codons ending with G or C in the mRNA sequence and N_T is the total number of codons in the mRNA sequence. The GC3 score varies from 0 (no G/C-ending codons) to 1 (All codons are G/C-ending).

Isoacceptors codon frequencies were calculated as follows:

$$\text{frequency of codon } i \text{ in gene } j = \frac{\text{number of occurrences of codon } i \text{ in gene } j}{\sum_{i=1}^{n} \text{number of occurrences of isoacceptor}_{ij}}$$

where isoacceptors refers to the synonymous codons for the amino acid encoded by codon *i* and n refers to the number of isoacceptors for a given amino acid ($2 \leq n \leq 6$). The isoacceptors scores ranged from 0 to 1, representing the ratio of each synonymous codon in the mRNA sequence.

The t-statistic (T-stat) describing the codon frequency was calculated as:

$$\text{T-statistic for codon } i = \frac{x_i - \mu_i}{s_i/\sqrt{n}}$$

Where x_i refers to the mean frequency for codon *i* in the sample, μ_i refers to the mean frequency for codon *i* across the Mus musculus genome, s_i refers to the standard deviation of the frequency for codon *i* in the sample, and n refers to the sample size.
T-statistics value of > 2 or < -2 indicates statistical significance ($p < 0.05$) versus the background (i.e. all genes)

➢ Gene level frameshifting analysis

To analyze frameshifted RPFs, we created a custom python script that first estimates the P-site offset using the same logic of RiboWaltz[70]. After P-site offset calculation per RPF length for each bam file, reads were counted and mapped to the P-site to determine their frame, based on their unique P-site offset in each bam file, whether it is 0, +1, or +2 frame. Frameshift Score (FS) for a gene was calculated as:

$$FS = \frac{Counts\ in\ frames + 1\ and + 2}{Counts\ in\ frame\ 0}$$

Thus, FS > 1 means frameshifting for a specific gene, while FS < 1 means in-frame reads for a gene.

After calculating FS scores and RPF counts for all samples, genes with total RPF < 20 in at least 3 biological replicates per group were discarded and the results merged into a single matrix. Gene with 0 RPF reads in the 0 frame and enough reads in the +1 and +2 frames to cross the filtering threshold were given an arbitrary FS score of 10. Limma-Trend[75] was used to compare between each condition and the sham group.

Identification of ribosomal frameshift loci from Ribo-seq data

To identify sites of aberrant translational frameshifting, we developed a computational framework that quantifies local deviations from the expected 0-frame periodicity in gene with statistically significant frameshifting in stroke conditions versus sham. For each coding sequence (CDS), codon-resolved RPF densities were modeled using an unsupervised two-state Hidden Markov Model (HMM) representing in-frame (0-frame) and out-of-frame (+1/-1) reading states. Transition probabilities were estimated genome-wide, and posterior decoding was used to identify genomic windows (typically ~3-6 codons) exhibiting statistically enriched +1 or -1 periodicity relative to background. Candidate loci were retained if they satisfied (i) a minimum out-of-frame posterior probability threshold, (ii) enrichment in RPF density above transcript-level background, and (iii) reproducibility across biological replicates.

Sequence-context characterization of frameshift loci

For each identified locus, we extracted a fixed-length, strand-aware RNA window centered on the frameshift event. We quantified known frameshift-promoting features, including slippery heptamers (XXXYYYZ), homopolymeric A/U tracts, codon runs (e.g., poly-Pro, poly-Lys), and viral-like motifs. Local secondary structure was analyzed using minimum free energy (MFE), ensemble free energy, and MFE frequency computed by RNAfold[76]. GC/AU composition, CpG frequency, nucleotide entropy, and G-quadruplex-forming sequences were extracted. Positional features (relative location within CDS, exon boundaries, and codon span) and amino-acid–level patterns (poly-Pro, poly-basic tracts, acidic–proline transitions) were quantified.

We quantified local nucleotide diversity using Shannon entropy of base composition in the FS-centered window ($H = -\sum p_i \log_2 p_i$ for $i \in (A,U,G,C)$). Lower entropy reflects strong compositional bias (e.g. AU-rich or GC-rich tracts), whereas higher entropy indicates compositionally balanced sequences. Frameshift loci tended to reside in windows with modestly reduced entropy compared to matched background windows, consistent with enrichment of homopolymeric tracts and biased sequence contexts.

Mapping frameshift events to transcript structures and functional consequences

To determine the consequences of each frameshift event on coding output, genomic coordinates were mapped to transcript-specific CDS coordinates using a GTF annotation. For each gene, all annotated transcripts containing a CDS were considered. A transcript was selected preferentially if (i) it matched the transcript ID carried from the upstream analysis, or (ii) its CDS encompassed the frameshift coordinate. For the selected transcript, we reconstructed its CDS and 3′UTR sequence directly from the genome. Frameshifted open reading frames (ORFs) were simulated by applying a +1 or −1 shift at the exact nucleotide where the out-of-frame state was detected. We determined:

714 • **PTC (premature termination codon)** if the shifted ORF terminated upstream of the
715 annotated stop codon;

716 • **SCRT (stop-codon readthrough–like extension)** if the shifted ORF terminated
717 downstream of the annotated stop;

718 • **NMD susceptibility** if the PTC occurred >50 nt upstream of the last exon–exon junction in
719 the CDS.

720 The canonical and frameshifted CDS and protein sequences were generated for all events and
721 classified into these functional categories.

722 **XGBoost modeling of frameshift loci features**

723 **Definition of FS windows:**

724 For each FS locus identified by the Ribo-seq pipeline, we defined a fixed-length genomic window
725 centered on the FS signal using the locus genomic coordinates. All windows were extracted as
726 strand-oriented sequences so that features were computed consistently with transcript directionality.

727 **Background (negative) window sampling:**

728 We generated two negative sets to evaluate whether FS loci exhibit distinctive sequence/structural
729 properties:

730 1. **Same-transcript background (local control)**
731 For each transcript, we defined a transcript-specific FS span as the minimum-to-maximum
732 coordinates covered by that transcript's FS windows. For each positive FS window, we
733 sampled up to **k = 5** non-overlapping windows of identical length within the same transcript
734 span, excluding any region overlapping any FS window. Background sequences were
735 extracted and oriented using the same procedure as positive FS windows.

736 2. **Global genomic background (genome-wide control)**
737 For the FS vs global analysis, we sampled length-matched windows from anywhere in the
738 genome. For each positive FS window, we sampled a chromosome with probability
739 proportional to chromosome length, drew a uniform random start coordinate within valid
740 bounds, and extracted the corresponding segment (treated as the + strand for feature
741 computation). (This negative set tests whether FS windows differ from generic genomic
742 sequence of matched length.)

743 **Feature extraction:**

744 All features were computed from the oriented window sequence using a dedicated feature extractor.
745 Features fell into five categories: composition, codon/AA usage, CpG/entropy, positional context, and
746 predicted RNA structure/motifs. Unless stated otherwise, counts were converted to fractions by
747 dividing by the relevant sequence length.

748 **(1) Window length and base composition**

749 • **Window length (nt)**: number of nucleotides in the extracted window.

750 • **GC fraction**: (G + C) / window length.

- **AU fraction**: (A + T) / window length.
- **GC3 fraction (GC at codon position 3)**: fraction of G/C among every third nucleotide when the window is parsed in-frame from position 1 (positions 3, 6, 9, …). When fewer than one codon is present, GC3 was set to 0.

(2) Codon usage features

Each oriented DNA window was split into non-overlapping codons from position 1 (truncating any remainder). For all 64 codons, we computed relative codon frequencies (codon count divided by total codon count in the window). Each codon frequency was used as a separate feature.

(3) Amino acid composition and hydrophobicity

The coding-strand DNA window was translated in-frame using the standard genetic code. From the amino acid sequence we computed:

- **Amino acid frequencies** for each residue (fraction of total amino acids).
- **Hydrophobic fraction**, using the hydrophobic amino acids set: **H = A, G, I, L, M, F, P, W, V**.
- **Longest hydrophobic run length**: maximum number of consecutive hydrophobic amino acids.
- **Proline fraction**: fraction of residues that are Pro (P).

(4) CpG and nucleotide entropy metrics

- **CpG count**: number of “CG” dinucleotides in the window.
- **CpG observed/expected ratio**:
  (CpG count × window length) / (C count × G count), with denominator protected against zero.
- **Shannon nucleotide entropy**: computed from nucleotide frequencies (A/C/G/T) in the window using the standard definition; entropy was set to 0 if the window contained only one nucleotide type.

In addition, several summary metrics (e.g., CpG O/E and entropy) derived from the master FS context file during preprocessing were retained as candidate predictors, particularly for the UP vs DOWN models (see below).

(5) Positional, structural, and motif features

- **Relative CDS position metrics**: location of the FS window (start, midpoint, end) scaled to CDS length (0 = start codon; 1 = stop codon), and the number of codons spanned by the window.
- **Local RNA structure (RNAfold)**: minimum free energy (MFE), ensemble free energy, frequency/probability of the MFE structure, and a success flag indicating valid folding output.
- **Putative G-quadruplex content**: count of predicted PQS motifs within the window.

- **Motif flags**: binary indicators for presence of predefined local sequence motifs, including canonical stop codon motifs and homopolymer tracts (polyA/polyU ≥ 6 nt), as defined in the FS pipeline.

Construction of the feature matrix and modeling tasks:

For each timepoint, we combined positive FS windows and background windows into a single feature matrix. For XGBoost, we retained numeric predictors and excluded identifiers (e.g., gene ID, transcript ID, locus coordinates). Missing numeric values were imputed as 0.

We trained models for two classification tasks:

1. **FS vs background** (same-transcript background or global background)
   To avoid circularity, we excluded "pass-through" context variables that were used during upstream FS filtering/preselection (e.g., certain motif flags, positional metrics, or context-level structure/CpG/entropy variables depending on the pipeline stage). These variables were retained for post hoc interpretation but not used for training in FS-vs-background models.

2. **FS UP vs DOWN**
   To identify features associated with differential regulation among FS loci, we restricted to FS windows only and trained a binary classifier (UP = 1, DOWN = 0). In this setting, we allowed the full numeric feature set (including CpG/entropy, positional and RNA structure features), as these were not used to define the UP/DOWN labels.

XGBoost training and model interpretation:

All models were implemented in R using the xgboost package. For each timepoint and classification task, we trained a gradient-boosted tree classifier with a logistic objective. Models were trained for up to 400 boosting rounds with early stopping (patience = 20) based on AUC. Hyperparameters were:

- objective = binary:logistic
- eval_metric = auc
- max_depth = 4
- eta = 0.05
- subsample = 0.8
- colsample_bytree = 0.8

Feature importance was assessed using (i) tree-based gain importance and (ii) permutation-based global importance. To interpret model predictions at the feature level, we also computed **SHAP values** to obtain global and locus-specific contributions of individual features.

➢ **Stop codon readthrough (SCRT) analysis**

RiboWaltz was used to count per gene CDS and 3'UTR P-sites (In order to avoid counting non-RPF reads in the 3'UTR) and create the count matrix. Next, GLM-betabinomial model was implemented on the counts to compare each time point to the Sham group. Significance level was set as |log2FC|

821 ≥ 1 and FDR ≤ 0.05. Metagene plots for selected genes were created using RiboWaltz, extending
822 the plots to the 3'UTR as needed. GC3 and isoacceptors codon frequencies analysis was performed
823 as explained above on significant genes. Per gene Ribosome Readthrough Score (RRS) was
824 calculated as:

825 $$RRS = \frac{Reads\ in\ 3'UTR}{Reads\ in\ CDS}:$$

826 **Stop-Proximal Metagene Profiling and pre-stop ribosome occupancy calculation.**

827 To compare genes with significant changes in SCRT events in stroke vs control, we extracted the
828 reads mapped to a window 60 nucleotides upstream of the stop codon. The window was binned in
829 the codon space, to generate normalized P-sites at each codon site upstream of the stop codon. For
830 the background genes, we selected a random set of genes with comparable CDS coverage to the
831 foreground set to avoid inflation of pre-stop p-site density of the background genes due to large
832 number of genes selected or differences in coverage.

833 Pre-stop ribosome occupancy was calculated 30 nucleotides (i.e. 10 codons) upstream of the stop
834 codon. This metric was used for machine learning modeling.

835 **3′UTR Frame-Usage Analysis**

836 To evaluate the dominant frame in the 3'UTR reads, we assigned a frame for each read based on
837 the P-site offset calculated for the CDS. Frame-usage matrices were aggregated across all genes
838 and timepoints to detect frame biases associated with stop codon readthrough (SCRT).

839 **RNA Structural Features Around the Termination Codon**

840 We extracted a window of 100 nucleotides downstream of the stop codon and calculated minimum
841 free energy (MFE) using Vienna RNAfold[76].

842 **Machine learning based identification of SCRT Determinants**

843 We used extreme gradient boosting (XGBoost; xgboost R package) to identify transcript features
844 associated with stroke-induced stop-codon readthrough (SCRT) at each timepoint (H1, H6, H24).
845 For each timepoint, we trained a binary classifier where positives were genes classified as SCRT-
846 upregulated by the SCRT differential model versus sham group (logFC > 1 and FDR < 0.05), and
847 negatives were an equal number of non-significant genes in the same timepoint table. Feature
848 matrices included: (i) stop-proximal ribosome occupancy (mean normalized P-site density in the last
849 30 nt of CDS upstream of the stop with a coverage matching background genes selected), (ii)
850 fractions of early 3'UTR P-sites in each of the three reading frames, (iii) CDS and 3'UTR length
851 (log1p-transformed), (iv) CDS GC3 score, (v) stop-window sequence composition (GC fraction), (vi)
852 stop codon identity, and (vii) local RNA structure (RNAfold minimum free energy of the first 100 nt
853 downstream of the stop). Models were trained with a logistic objective (binary:logistic) and evaluated
854 using area under the ROC curve (AUC). We used 5-fold stratified cross-validation with early stopping
855 (20 rounds; maximum 400 rounds) to select the optimal number of boosting iterations.
856 Hyperparameters were max_depth=4, eta=0.05, subsample=0.8, and colsample_bytree=0.8.
857 Missing numeric values were imputed as 0 before model fitting. Feature importance was assessed
858 using tree-based gain, and feature attributions were examined using SHAP-like Shapley values
859 implemented via the *iml* framework (global summaries and per-gene local explanations).

860 ➢ **ORF analysis**

861 Translated open read frames (ORFs) were calculated using RiboCode[37]. We fixed the start codon as
862 ATG to reduce complexity and limited the ORFs to a minimum 6 codons for ORF identification.
863 RibCode's results were then imported into R, and additional filtering (minimum P-sites per ORF ≥ 10,
864 minimum codons/amino acids ≥ 10, frame 0 P-site counts ≥ 50% of all ORF P-site counts, and
865 RiboCode significance (*p* value) ≤ 0.05) were implemented prior to downstream analysis. Differential
866 ORF expression was calculated using EdgeR[77] using ORF P-site counts as input. The canonical
867 ORF for each gene was defined as the longest ORF with the highest 0 frame P-sites. ORFs
868 designated as annotated and protein coding were primarily defined as canonical. If multiple
869 annotated and protein coding ORFs were detected per gene, the longest with the highest P-site
870 counts was selected as the canonical. The dominant translated ORF was determined by calculating
871 the ratio of ORF P-site counts for all ORFs derived from the same gene and selecting the ORF that
872 has a ratio of ≥ 60%. Differential ORF usage (DOU) was performed using DRIMSeq[38] using a
873 Dirichlet–multinomial model.

874 **Analysis of differential upstream translation**

875 To analyze genes with changes in their 5'UTR to CDS translation ratio, we used the same approach
876 employed in the SCRT analysis but rather than comparing CDS and 3'UTR P-site counts with GLM-
877 betabinomial modeling, we compared CDS and 5'UTR P-site counts between stroke and sham
878 groups. Significance level was set as |log2FC| ≥ 1 and FDR ≤ 0.05. Upregulation indicates more
879 reads in the 5'UTR compared to CDS in the condition versus sham and vice versa.

880 **Acknowledgments:** The authors report no conflict of interest regarding this work. Large language
881 mode (ChatGPT) was used for language editing and as assistant in code development. The authors
882 would like to thank Ms Natsumi Konno and Dr Michitoshi Watanabe for their assistance in the
883 various stages of this work. This work was supported by Biomedical Research Core of Tohoku
884 University Graduate School of Medicine.

885 **Funding:** This work was funded by the Japan society for the promotion of science (JSPS) grant
886 number 23H02741 for S.R.

887 **Authors contributions: S.R.:** Conception and study design. Writing. Bioinformatics. Funding.
888 Administration. Supervision. **Y.K., T.N., D.A., A.M.,** and **H.I:** Experimentation and animal modeling.
889 **H.E.:** critically reviewed the manuscript. Supervision. **K.N.:** Critically reviewed the manuscript.
890 Administration. Supervision.

891 **Data availability:** Raw Ribo-seq and RNA-seq fastq files were deposited in sequence read archive
892 (SRA; PRJNA1220332 and PRJNA1402676).

893 **Code availability:** The python scripts used to calculate frameshifting and analyze the frameshift
894 location and impact are available through GitHub (https://github.com/Rashad-lab/Gene-level-
895 frameshifting). All other analysis were performed using R language programming and available
896 libraries. Further information can be provided by the corresponding author upon reasonable
897 requests.

898 **References**

899 1 Benjamin, E. J. *et al.* Heart Disease and Stroke Statistics-2019 Update: A Report
900 From the American Heart Association. *Circulation* **139**, e56-e528,
901 doi:10.1161/cir.0000000000000659 (2019).
902 2 Moritoyo, T. *et al.* A first-in-human study of the anti-inflammatory profibrinolytic
903 TMS-007, an SMTP family triprenyl phenol. *Br J Clin Pharmacol* **89**, 1809-1819,
904 doi:10.1111/bcp.15651 (2023).
905 3 Xiong, Y., Wakhloo, A. K. & Fisher, M. Advances in Acute Ischemic Stroke Therapy.
906 *Circ Res* **130**, 1230-1251, doi:10.1161/circresaha.121.319948 (2022).
907 4 Fan, J. *et al.* Global Burden, Risk Factor Analysis, and Prediction Study of Ischemic
908 Stroke, 1990-2030. *Neurology* **101**, e137-e150, doi:10.1212/wnl.0000000000207387
909 (2023).
910 5 Campbell, B. C. V. *et al.* Ischaemic stroke. *Nat Rev Dis Primers* **5**, 70,
911 doi:10.1038/s41572-019-0118-8 (2019).
912 6 Sekerdag, E., Solaroglu, I. & Gursoy-Ozdemir, Y. Cell Death Mechanisms in Stroke
913 and Novel Molecular and Cellular Treatment Options. *Curr Neuropharmacol* **16**,
914 1396-1415, doi:10.2174/1570159x16666180302115544 (2018).
915 7 Niizuma, K. *et al.* Randomized placebo-controlled trial of CL2020, an allogenic
916 muse cell-based product, in subacute ischemic stroke. *J Cereb Blood Flow Metab*
917 **43**, 2029-2039, doi:10.1177/0271678x231202594 (2023).
918 8 Faheem, H. *et al.* Neuroprotective effects of minocycline and progesterone on white
919 matter injury after focal cerebral ischemia. *J Clin Neurosci* **64**, 206-213,
920 doi:10.1016/j.jocn.2019.04.012 (2019).
921 9 Zheng, K. *et al.* Single-cell RNA-seq reveals the transcriptional landscape in
922 ischemic stroke. *J Cereb Blood Flow Metab* **42**, 56-73,
923 doi:10.1177/0271678x211026770 (2022).
924 10 Zhao, Y. *et al.* Single-Cell RNA-Seq Reveals Changes in Cell Subsets in the Cortical
925 Microenvironment during Acute Phase of Ischemic Stroke Rats. *J Integr Neurosci* **22**,
926 128, doi:10.31083/j.jin2205128 (2023).
927 11 Guo, K. *et al.* Single-Cell RNA Sequencing With Combined Use of Bulk RNA
928 Sequencing to Reveal Cell Heterogeneity and Molecular Changes at Acute Stage of
929 Ischemic Stroke in Mouse Cortex Penumbra Area. *Front Cell Dev Biol* **9**, 624711,
930 doi:10.3389/fcell.2021.624711 (2021).
931 12 Rashad, S. *et al.* Translational response to mitochondrial stresses is orchestrated by
932 tRNA modifications. *bioRxiv*, 2024.2002.2014.580389,
933 doi:10.1101/2024.02.14.580389 (2024).
934 13 Ingolia, N. T., Ghaemmaghami, S., Newman, J. R. & Weissman, J. S. Genome-wide
935 analysis in vivo of translation with nucleotide resolution using ribosome profiling.
936 *Science* **324**, 218-223, doi:10.1126/science.1168978 (2009).
937 14 Brar, G. A. & Weissman, J. S. Ribosome profiling reveals the what, when, where and
938 how of protein synthesis. *Nat Rev Mol Cell Biol* **16**, 651-664, doi:10.1038/nrm4069
939 (2015).

15 Dedon, P. C. & Begley, T. J. Dysfunctional tRNA reprogramming and codon-biased translation in cancer. *Trends Mol Med* **28**, 964-978, doi:10.1016/j.molmed.2022.09.007 (2022).
16 Huber, S. M. *et al.* Arsenite toxicity is regulated by queuine availability and oxidation-induced reprogramming of the human tRNA epitranscriptome. *Proc Natl Acad Sci U S A* **119**, e2123529119, doi:10.1073/pnas.2123529119 (2022).
17 Chan, C. T. Y. *et al.* Reprogramming of tRNA modifications controls the oxidative stress response by codon-biased translation of proteins. *Nature Communications* **3**, 937, doi:10.1038/ncomms1938 (2012).
18 Torrent, M., Chalancon, G., de Groot, N. S., Wuster, A. & Madan Babu, M. Cells alter their tRNA abundance to selectively regulate protein synthesis during stress conditions. *Sci Signal* **11**, doi:10.1126/scisignal.aat6409 (2018).
19 Rashad, S. *et al.* The stress specific impact of ALKBH1 on tRNA cleavage and tiRNA generation. *RNA Biol* **17**, 1092-1103, doi:10.1080/15476286.2020.1779492 (2020).
20 Rashad, S., Tominaga, T. & Niizuma, K. The cell and stress-specific canonical and noncanonical tRNA cleavage. *J Cell Physiol* **236**, 3710-3724, doi:10.1002/jcp.30107 (2021).
21 Sato, K., Rashad, S., Niizuma, K. & Tominaga, T. Stress Induced tRNA Halves (tiRNAs) as Biomarkers for Stroke and Stroke Therapy; Pre-clinical Study. *Neuroscience* **434**, 44-54, doi:10.1016/j.neuroscience.2020.03.018 (2020).
22 Rashad, S. & Niizuma, K. Third-nucleotide codon bias and synonymous codon bias define functional translational programs that shape human tissue and cancer proteomes. *bioRxiv*, 2025.2010.2031.685942, doi:10.1101/2025.10.31.685942 (2025).
23 Ren, G. *et al.* Ribosomal frameshifting at normal codon repeats recodes functional chimeric proteins in human. *Nucleic Acids Res* **52**, 2463-2479, doi:10.1093/nar/gkae035 (2024).
24 Demo, G. *et al.* Structural basis for +1 ribosomal frameshifting during EF-G-catalyzed translocation. *Nat Commun* **12**, 4644, doi:10.1038/s41467-021-24911-1 (2021).
25 Kishor, A., Fritz, S. E. & Hogg, J. R. Nonsense-mediated mRNA decay: The challenge of telling right from wrong in a complex transcriptome. *Wiley Interdiscip Rev RNA* **10**, e1548, doi:10.1002/wrna.1548 (2019).
26 Hol, E. M. *et al.* Frameshifted beta-amyloid precursor protein (APP+1) is a secretory protein, and the level of APP+1 in cerebrospinal fluid is linked to Alzheimer pathology. *J Biol Chem* **278**, 39637-39643, doi:10.1074/jbc.M302295200 (2003).
27 Popp, M. W. & Maquat, L. E. Leveraging Rules of Nonsense-Mediated mRNA Decay for Genome Engineering and Personalized Medicine. *Cell* **165**, 1319-1322, doi:10.1016/j.cell.2016.05.053 (2016).
28 Zhou, Y. *et al.* Dynamic mRNA Stability Buffer Transcriptional Activation During Neuronal Differentiation and Is Regulated by SAMD4A. *J Cell Physiol*, doi:10.1002/jcp.31477 (2024).
29 Schueren, F. & Thoms, S. Functional Translational Readthrough: A Systems Biology Perspective. *PLoS Genet* **12**, e1006196, doi:10.1371/journal.pgen.1006196 (2016).

30 Poncová, K. *et al.* uS3/Rps3 controls fidelity of translation termination and programmed stop codon readthrough in co-operation with eIF3. *Nucleic Acids Res* **47**, 11326-11343, doi:10.1093/nar/gkz929 (2019).
31 Dunn, J. G., Foo, C. K., Belletier, N. G., Gavis, E. R. & Weissman, J. S. Ribosome profiling reveals pervasive and regulated stop codon readthrough in Drosophila melanogaster. *Elife* **2**, e01179, doi:10.7554/eLife.01179 (2013).
32 Pati, R. *et al.* The Readthrough Isoform AQP4ex Is Constitutively Phosphorylated in the Perivascular Astrocyte Endfeet of Human Brain. *Biomolecules* **12**, doi:10.3390/biom12050633 (2022).
33 Palazzo, C. *et al.* Tissue Distribution of the Readthrough Isoform of AQP4 Reveals a Dual Role of AQP4ex Limited to CNS. *Int J Mol Sci* **21**, doi:10.3390/ijms21041531 (2020).
34 Eswarappa, S. M. *et al.* Programmed translational readthrough generates antiangiogenic VEGF-Ax. *Cell* **157**, 1605-1618, doi:10.1016/j.cell.2014.04.033 (2014).
35 Mangkalaphiban, K. *et al.* Extended stop codon context predicts nonsense codon readthrough efficiency in human cells. *Nat Commun* **15**, 2486, doi:10.1038/s41467-024-46703-z (2024).
36 Loughran, G. *et al.* Evidence of efficient stop codon readthrough in four mammalian genes. *Nucleic Acids Res* **42**, 8928-8938, doi:10.1093/nar/gku608 (2014).
37 Xiao, Z. *et al.* De novo annotation and characterization of the translatome with ribosome profiling data. *Nucleic Acids Res* **46**, e61, doi:10.1093/nar/gky179 (2018).
38 Nowicka, M. & Robinson, M. D. DRIMSeq: a Dirichlet-multinomial framework for multivariate count outcomes in genomics. *F1000Res* **5**, 1356, doi:10.12688/f1000research.8900.2 (2016).
39 Sun, Y. *et al.* Upstream open reading frames buffer translational variability during Drosophila evolution and development. *Elife* **14**, doi:10.7554/eLife.104074 (2025).
40 Young, S. K. & Wek, R. C. Upstream Open Reading Frames Differentially Regulate Gene-specific Translation in the Integrated Stress Response. *J Biol Chem* **291**, 16927-16935, doi:10.1074/jbc.R116.733899 (2016).
41 Rashad, S. *et al.* Translational response to mitochondrial stresses is orchestrated by tRNA modifications. *bioRxiv*, doi:10.1101/2024.02.14.580389 (2024).
42 Zhao, M., Qiao, Y., Weiss, A. & Zhao, W. Neuroprotective strategies in acute ischemic stroke: A narrative review of recent advances and clinical outcomes. *Brain Circ* **10**, 296-302, doi:10.4103/bc.bc_165_24 (2024).
43 Androvic, P. *et al.* Decoding the Transcriptional Response to Ischemic Stroke in Young and Aged Mouse Brain. *Cell Rep* **31**, 107777, doi:10.1016/j.celrep.2020.107777 (2020).
44 Hochrainer, K. & Yang, W. Stroke Proteomics: From Discovery to Diagnostic and Therapeutic Applications. *Circ Res* **130**, 1145-1166, doi:10.1161/circresaha.122.320110 (2022).
45 Zucha, D. *et al.* Spatiotemporal transcriptomic map of glial cell response in a mouse model of acute brain ischemia. *Proc Natl Acad Sci U S A* **121**, e2404203121, doi:10.1073/pnas.2404203121 (2024).

46 Weber, R. Z. *et al.* A molecular brain atlas reveals cellular shifts during the repair phase of stroke. *J Neuroinflammation* **22**, 112, doi:10.1186/s12974-025-03437-z (2025).
47 Snieckute, G. *et al.* Ribosome stalling is a signal for metabolic regulation by the ribotoxic stress response. *Cell Metab* **34**, 2036-2046 e2038, doi:10.1016/j.cmet.2022.10.011 (2022).
48 Wu, C. C., Peterson, A., Zinshteyn, B., Regot, S. & Green, R. Ribosome Collisions Trigger General Stress Responses to Regulate Cell Fate. *Cell* **182**, 404-416 e414, doi:10.1016/j.cell.2020.06.006 (2020).
49 Shim, B. *et al.* RiboTag RNA Sequencing Identifies Local Translation of HSP70 in Astrocyte Endfeet After Cerebral Ischemia. *Int J Mol Sci* **26**, doi:10.3390/ijms26010309 (2025).
50 Arbaizar-Rovirosa, M. *et al.* Transcriptomics and translatomics identify a robust inflammatory gene signature in brain endothelial cells after ischemic stroke. *J Neuroinflammation* **20**, 207, doi:10.1186/s12974-023-02888-6 (2023).
51 Wangen, J. R. & Green, R. Stop codon context influences genome-wide stimulation of termination codon readthrough by aminoglycosides. *Elife* **9**, doi:10.7554/eLife.52611 (2020).
52 Mangkalaphiban, K. *et al.* Transcriptome-wide investigation of stop codon readthrough in Saccharomyces cerevisiae. *PLoS Genet* **17**, e1009538, doi:10.1371/journal.pgen.1009538 (2021).
53 Zhang, H. *et al.* Metabolic stress promotes stop-codon readthrough and phenotypic heterogeneity. *Proc Natl Acad Sci U S A* **117**, 22167-22172, doi:10.1073/pnas.2013543117 (2020).
54 Gerashchenko, M. V., Lobanov, A. V. & Gladyshev, V. N. Genome-wide ribosome profiling reveals complex translational regulation in response to oxidative stress. *Proc Natl Acad Sci U S A* **109**, 17394-17399, doi:10.1073/pnas.1120799109 (2012).
55 Guydosh, N. R. & Green, R. Dom34 rescues ribosomes in 3' untranslated regions. *Cell* **156**, 950-962, doi:10.1016/j.cell.2014.02.006 (2014).
56 Manjunath, L. E., Singh, A., Som, S. & Eswarappa, S. M. Mammalian proteome expansion by stop codon readthrough. *Wiley Interdiscip Rev RNA* **14**, e1739, doi:10.1002/wrna.1739 (2023).
57 Smith, A. M., Costello, M. S., Kettring, A. H., Wingo, R. J. & Moore, S. D. Ribosome collisions alter frameshifting at translational reprogramming motifs in bacterial mRNAs. *Proc Natl Acad Sci U S A* **116**, 21769-21779, doi:10.1073/pnas.1910613116 (2019).
58 Vind, A. C. *et al.* ZAKα Recognizes Stalled Ribosomes through Partially Redundant Sensor Domains. *Mol Cell* **78**, 700-713.e707, doi:10.1016/j.molcel.2020.03.021 (2020).
59 Juszkiewicz, S. *et al.* Ribosome collisions trigger cis-acting feedback inhibition of translation initiation. *Elife* **9**, doi:10.7554/eLife.60038 (2020).
60 Wu, H. W. *et al.* Noise reduction by upstream open reading frames. *Nat Plants* **8**, 474-480, doi:10.1038/s41477-022-01136-8 (2022).

1071 61 Morris, D. R. & Geballe, A. P. Upstream open reading frames as regulators of mRNA
1072 translation. *Mol Cell Biol* **20**, 8635-8642, doi:10.1128/mcb.20.23.8635-8642.2000
1073 (2000).
1074 62 Sinha, N. K. *et al.* EDF1 coordinates cellular responses to ribosome collisions. *Elife*
1075 **9**, doi:10.7554/eLife.58828 (2020).
1076 63 Kurosaki, T., Popp, M. W. & Maquat, L. E. Quality and quantity control of gene
1077 expression by nonsense-mediated mRNA decay. *Nat Rev Mol Cell Biol* **20**, 406-420,
1078 doi:10.1038/s41580-019-0126-2 (2019).
1079 64 Kato, Y. *et al.* Intravenous administration of muse cells improves cerebral ischemia
1080 outcome via immunomodulation in the spleen. *J Cereb Blood Flow Metab*,
1081 271678x241290363, doi:10.1177/0271678x241290363 (2024).
1082 65 Taguchi, A. *et al.* A Reproducible and Simple Model of Permanent Cerebral Ischemia
1083 in CB-17 and SCID Mice. *J Exp Stroke Transl Med* **3**, 28-33, doi:10.6030/1939-067x-
1084 3.1.28 (2010).
1085 66 Ando, D. *et al.* Decoding Codon Bias: The Role of tRNA Modifications in Tissue-
1086 Specific Translation. *Int. J. Mol. Sci.* **26**, 706 (2025).
1087 67 Bolger, A. M., Lohse, M. & Usadel, B. Trimmomatic: a flexible trimmer for Illumina
1088 sequence data. *Bioinformatics* **30**, 2114-2120, doi:10.1093/bioinformatics/btu170
1089 (2014).
1090 68 Langmead, B. & Salzberg, S. L. Fast gapped-read alignment with Bowtie 2. *Nat*
1091 *Methods* **9**, 357-359, doi:10.1038/nmeth.1923 (2012).
1092 69 Liao, Y., Smyth, G. K. & Shi, W. featureCounts: an efficient general purpose program
1093 for assigning sequence reads to genomic features. *Bioinformatics* **30**, 923-930,
1094 doi:10.1093/bioinformatics/btt656 (2014).
1095 70 Lauria, F. *et al.* riboWaltz: Optimization of ribosome P-site positioning in ribosome
1096 profiling data. *PLoS Comput Biol* **14**, e1006169, doi:10.1371/journal.pcbi.1006169
1097 (2018).
1098 71 Kim, D., Paggi, J. M., Park, C., Bennett, C. & Salzberg, S. L. Graph-based genome
1099 alignment and genotyping with HISAT2 and HISAT-genotype. *Nat Biotechnol* **37**, 907-
1100 915, doi:10.1038/s41587-019-0201-4 (2019).
1101 72 Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and
1102 dispersion for RNA-seq data with DESeq2. *Genome Biol* **15**, 550,
1103 doi:10.1186/s13059-014-0550-8 (2014).
1104 73 Rashad, S. *et al.* Codon Usage and mRNA Stability are Translational Determinants of
1105 Cellular Response to Canonical Ferroptosis Inducers. *Neuroscience* **501**, 103-130,
1106 doi:10.1016/j.neuroscience.2022.08.009 (2022).
1107 74 Tumu, S., Patil, A., Towns, W., Dyavaiah, M. & Begley, T. J. The gene-specific codon
1108 counting database: a genome-based catalog of one-, two-, three-, four- and five-
1109 codon combinations present in Saccharomyces cerevisiae genes. *Database* **2012**,
1110 doi:10.1093/database/bas002 (2012).
1111 75 Ritchie, M. E. *et al.* limma powers differential expression analyses for RNA-
1112 sequencing and microarray studies. *Nucleic Acids Res* **43**, e47,
1113 doi:10.1093/nar/gkv007 (2015).

76 Lorenz, R. *et al*. ViennaRNA Package 2.0. *Algorithms Mol Biol* **6**, 26, doi:10.1186/1748-7188-6-26 (2011).

77 Chen, Y., Chen, L., Lun, A. T. L., Baldoni, P. L. & Smyth, G. K. edgeR v4: powerful differential analysis of sequencing data with expanded functionality and improved support for small counts and larger datasets. *Nucleic Acids Res* **53**, doi:10.1093/nar/gkaf018 (2025).

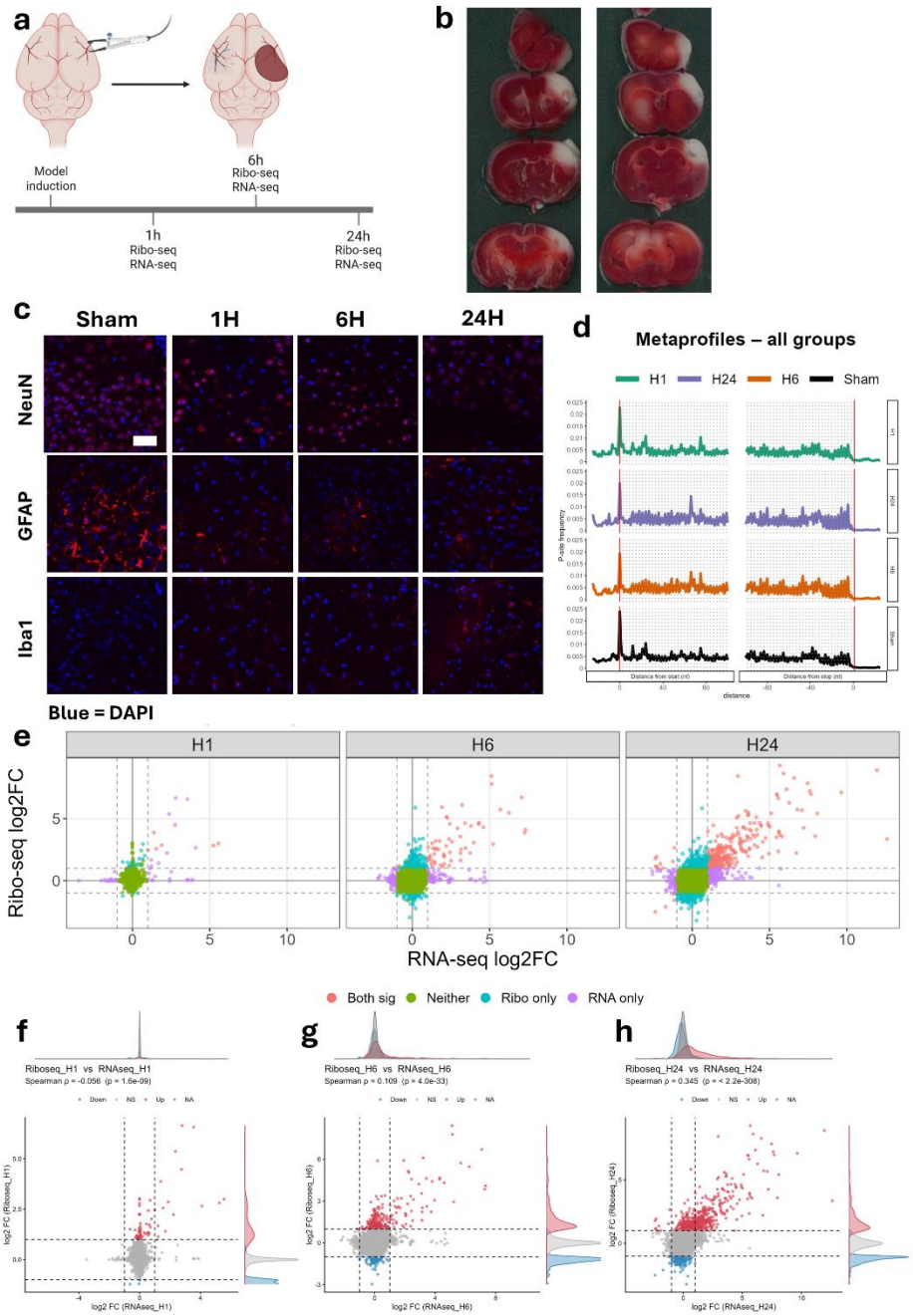


Figure 1: Ischemic stroke leads to transcription-translation discordance. a: Schematic showing the method for inducing stroke and the analysis performed at different timepoints after stroke. **b:** TTC stain showing the area of ischemic infarction 24 hours after dMCAO model induction. **c:** Immunofluorescence staining showing neurons (stained by NeuN), astrocytes (stained by GFAP), and microglia (stained by Iba1) after dMCAO model induction. bar = 50μm. **d:** Group-level metagene plot showing the cumulative 3-nucleotide periodicity indicating detection of translating ribosomes in the Ribo-seq analysis. **e:** Quadrant scatter plots showing the significant genes in the RNA-seq and Ribo-seq analyses at each time point. **f-h:** Spearman's rank correlation analysis scatter plots showing the correlation between gene RNA-seq and Ribo-seq log2 fold change values at each time points (1 hour, 6 hours, and 24 hours respectively).

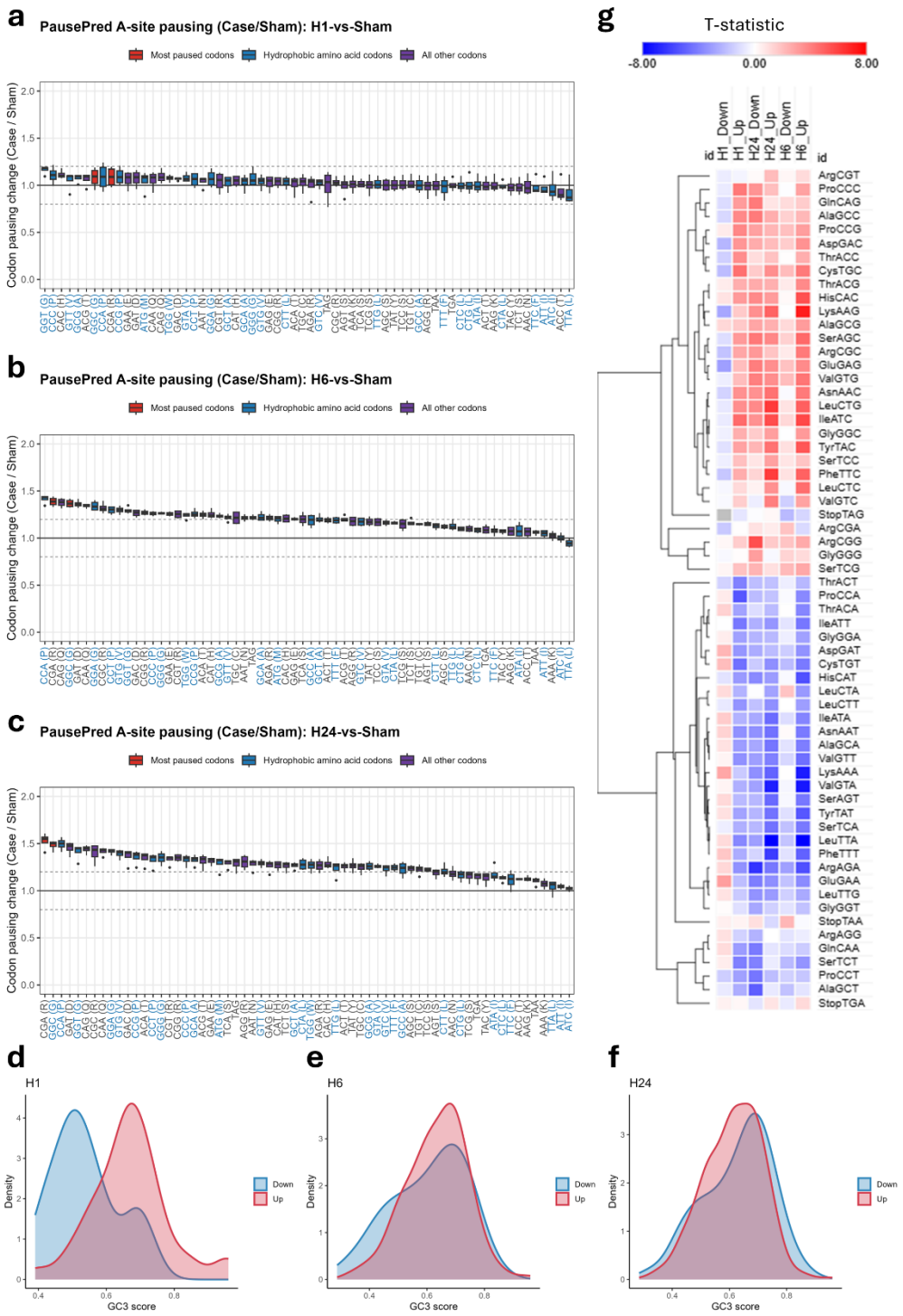


Figure 2: Ischemic stroke leads to altered codon decoding. a-c: Box plots showing A-site ribosome stalling after dMCAO model at each time point compared to Sham. **d-f:** Density plot of GC3 scores of the up and downregulated genes in the Ribo-seq dataset across time points (1 hour, 6 hours, and 24 hours respectively). **g:** Isoacceptors codon frequencies analysis of the up and downregulated genes in the Ribo-seq dataset across time points. T-statistics indicates significance compared to the genome average codon usage (|T-stat| ≥ 2 indicates $p \leq 0.05$).

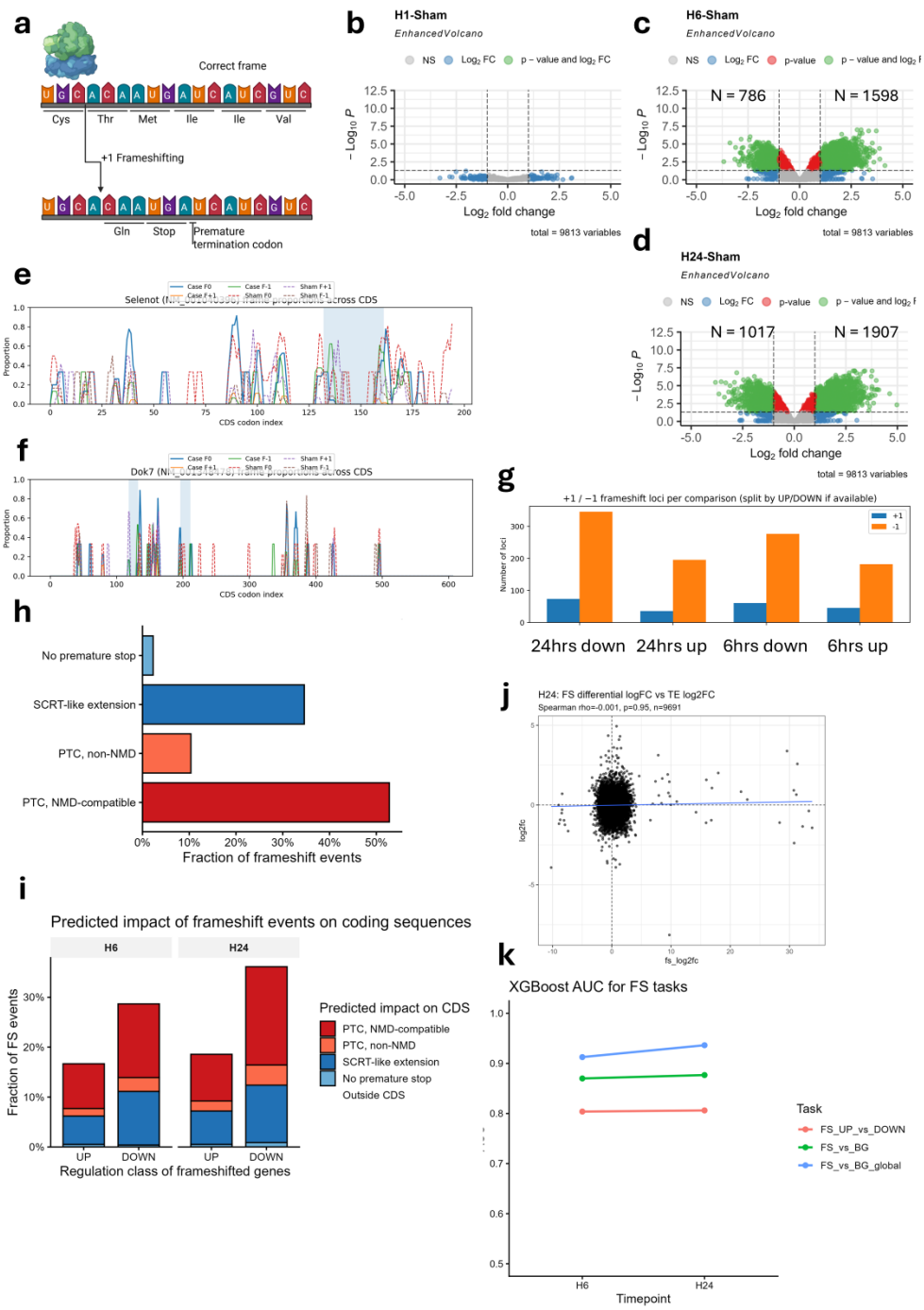


Figure 3: Ischemic stroke leads to frameshifting events across the coding sequences. a: Schematic showing the influence of frameshifting on the CDS reading frame. **b-d:** Volcano plots showing the changes in gene frameshifting at each time point after stroke compared to sham. Upregulated genes indicate more frameshifting (i.e. more RPFs in +1 and/or +2 frames compared to 0 frame) and vice versa. Significance was designated at |log2FC| ≥ 1 and FDR ≤ 0.05. **e:** Identification of single frameshift locus in Selenot gene, which was upregulated in the FS analysis after 24 hours. **f:** Identification of multiple frameshifting loci in Dok7 gene, which was upregulated in

1148 the FS analysis after 24 hours. **g:** Bar plot showing the number of identified +1 and -1 frameshifting
1149 events at each time point. **h:** Bar plot showing the potential impact of frameshifting on mRNA
1150 translation. This bar plot summarizes all frameshifted genes across all time points and all directions.
1151 **i**: Frameshifting impact on mRNA translation at each time point and direction of change. **j**:
1152 Spearman's rank correlation between gene log2FC of frameshifting analysis and TE analysis at 24
1153 hours timepoint. **k:** XGBoost model area under the curve (AUC) showing the model's ability to
1154 discriminate between frameshift loci features and the background (same transcript or global) and
1155 between up and down genes.

1156

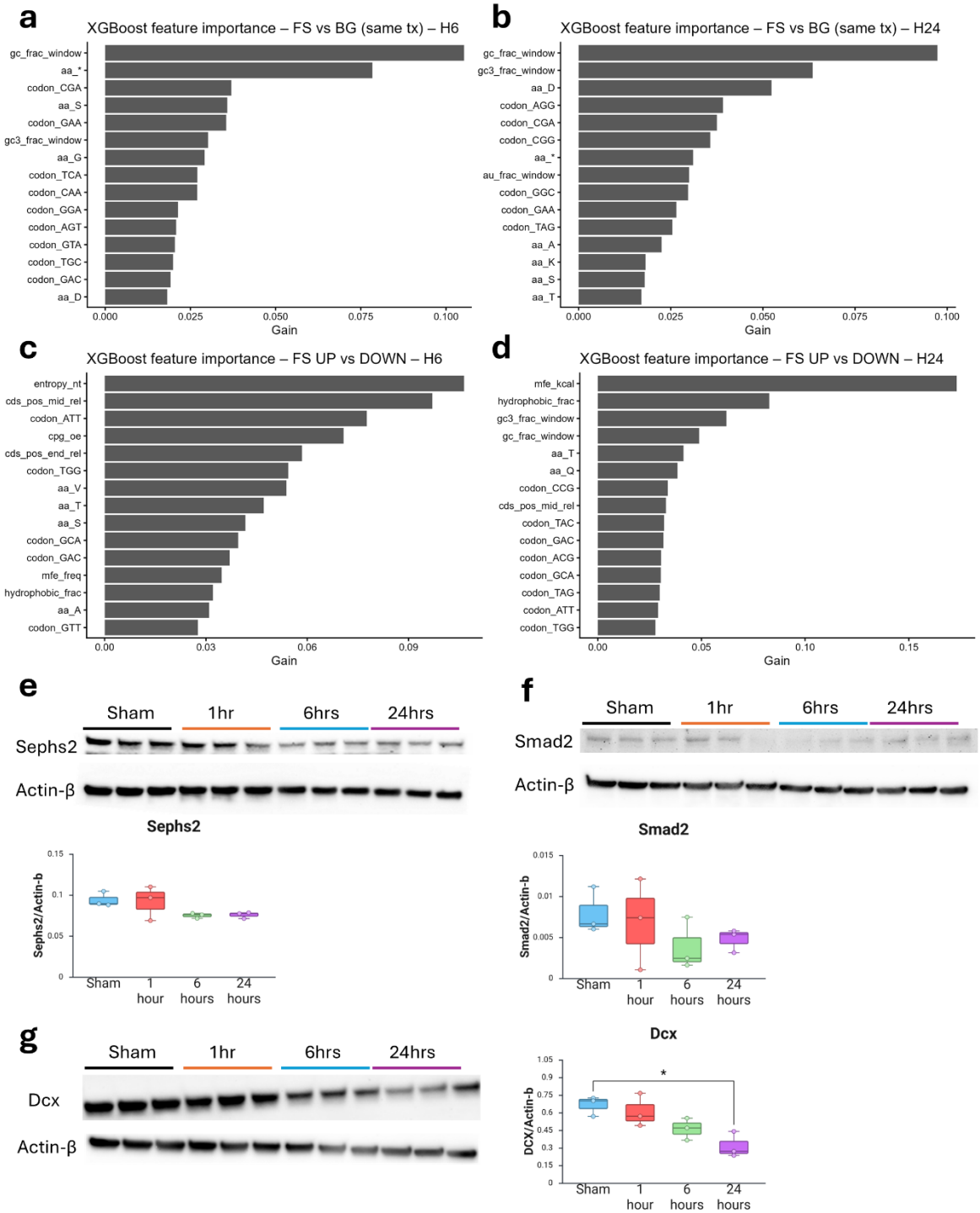


Figure 4: Sequence features define frameshifted loci. a-d: XGBoost feature importance plots at 6 and 24 hours timepoints showing the important features discriminating between frameshift loci and same transcript background (different regions of the same transcript) and up and down genes. **e-f:** Western blot analysis of significantly frameshifted genes showing the impact of frameshifting on protein levels. Graphs show quantification of western blot data. One-way ANOVA with Turkey's post-hoc analysis was used for statistical analysis.

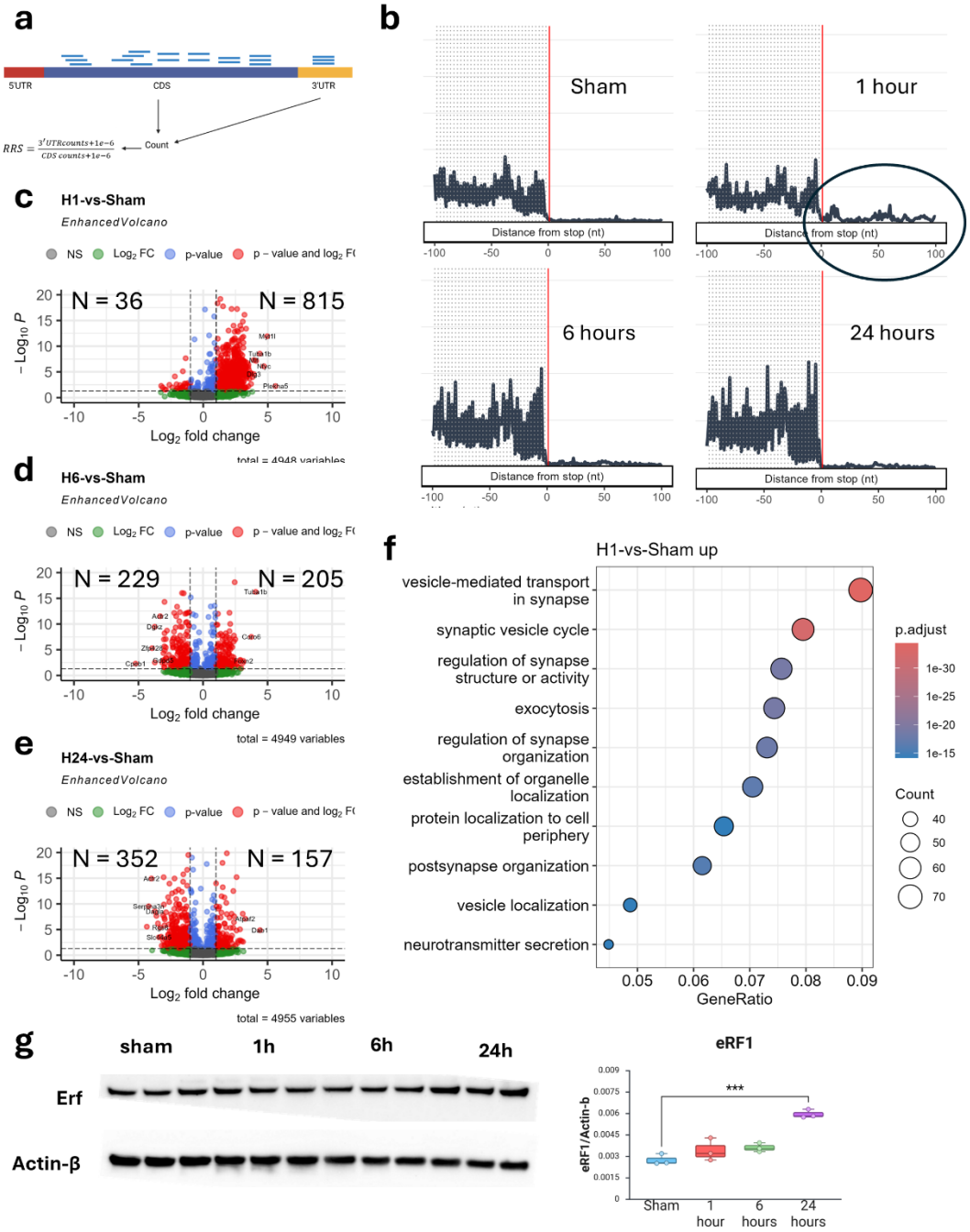
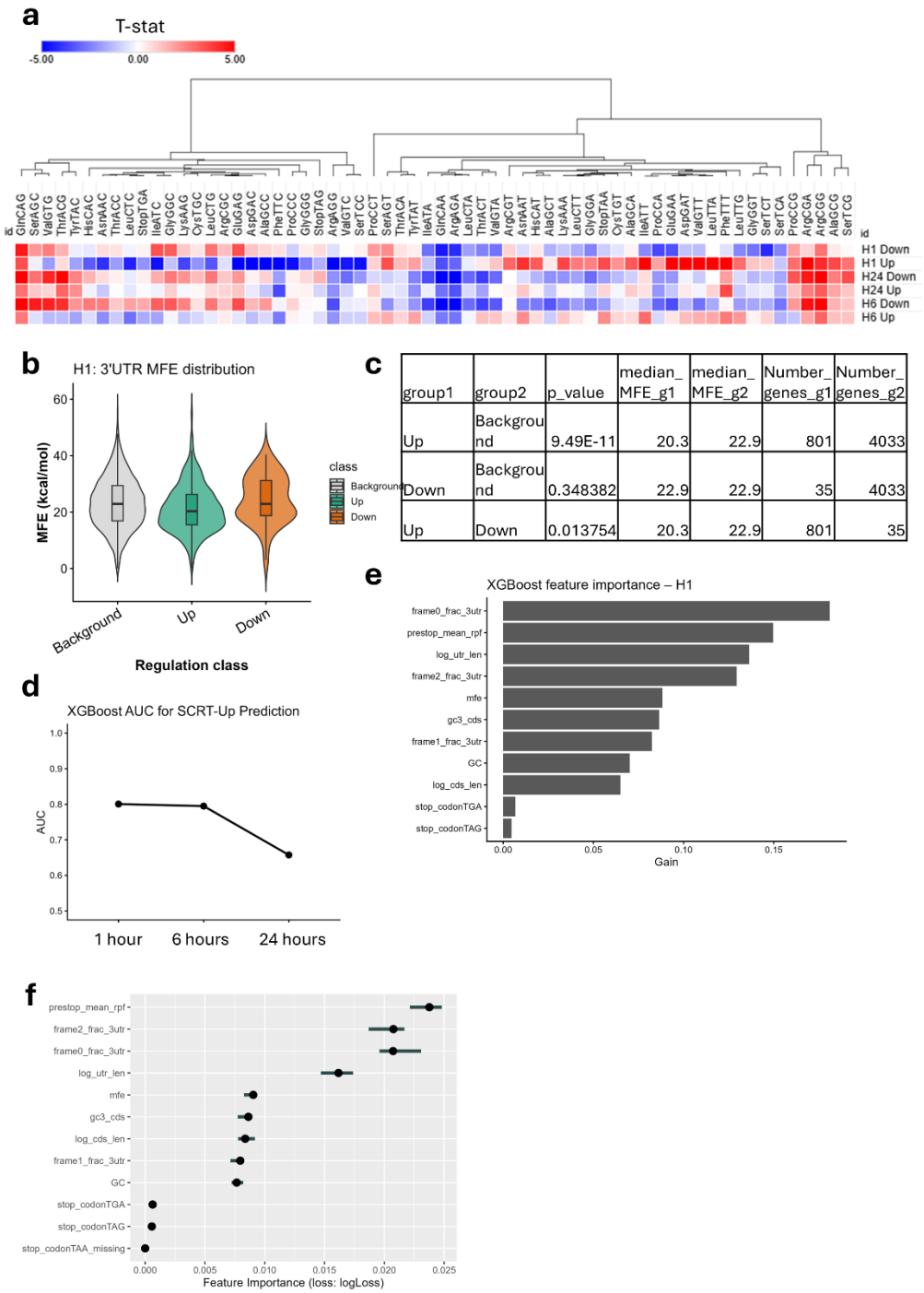


Figure 5: Ischemic stroke leads to early burst of strop codon readthrough events. a: Schematic representation of the approach used to identify SCRT events. **b:** Representative metagene plots showing 3-nucleotide periodicity in the 3'UTR especially in the 1-hour group. **c-d:** Volcano plots of gene showing significant SCRT events. Upregulated genes indicate more SCRT events compared to Sham and vice versa. Statistical significance cut off was |log2FC| ≥ 1 and FDR ≤ 0.05. **f:** ORA GOBP pathway analysis of upregulated genes in the SCRT analysis 1-hour after stroke. **g:** Western blot showing eukaryotic release factor (Erf or Etf1) protein levels. Protein levels remained unchanged after stroke.



group1	group2	p_value	median_MFE_g1	median_MFE_g2	Number_genes_g1	Number_genes_g2
Up	Background	9.49E-11	20.3	22.9	801	4033
Down	Background	0.348382	22.9	22.9	35	4033
Up	Down	0.013754	20.3	22.9	801	35

Figure 6: Translational dynamics and 3'UTR structures predict SCRT potential. a: Isoacceptors codon frequencies analysis of genes with significant SCRT events. **b:** Violin plot of minimum free energy (MFE) analysis of RNAfold using the first 100 nucleotides of the 3'UTR as input. **c:** Wilcoxon test analysis of MFE comparing the up and downregulated genes to the background or to each other. **d:** XGBoost model AUC results for predicting upregulated SCRT gene features at each time point. **e:** XGBoost feature importance for predicting upregulated SCRT genes 1-hour after stroke. **f:** SHAP feature importance of XGBoost model at the 1-hour timepoint.

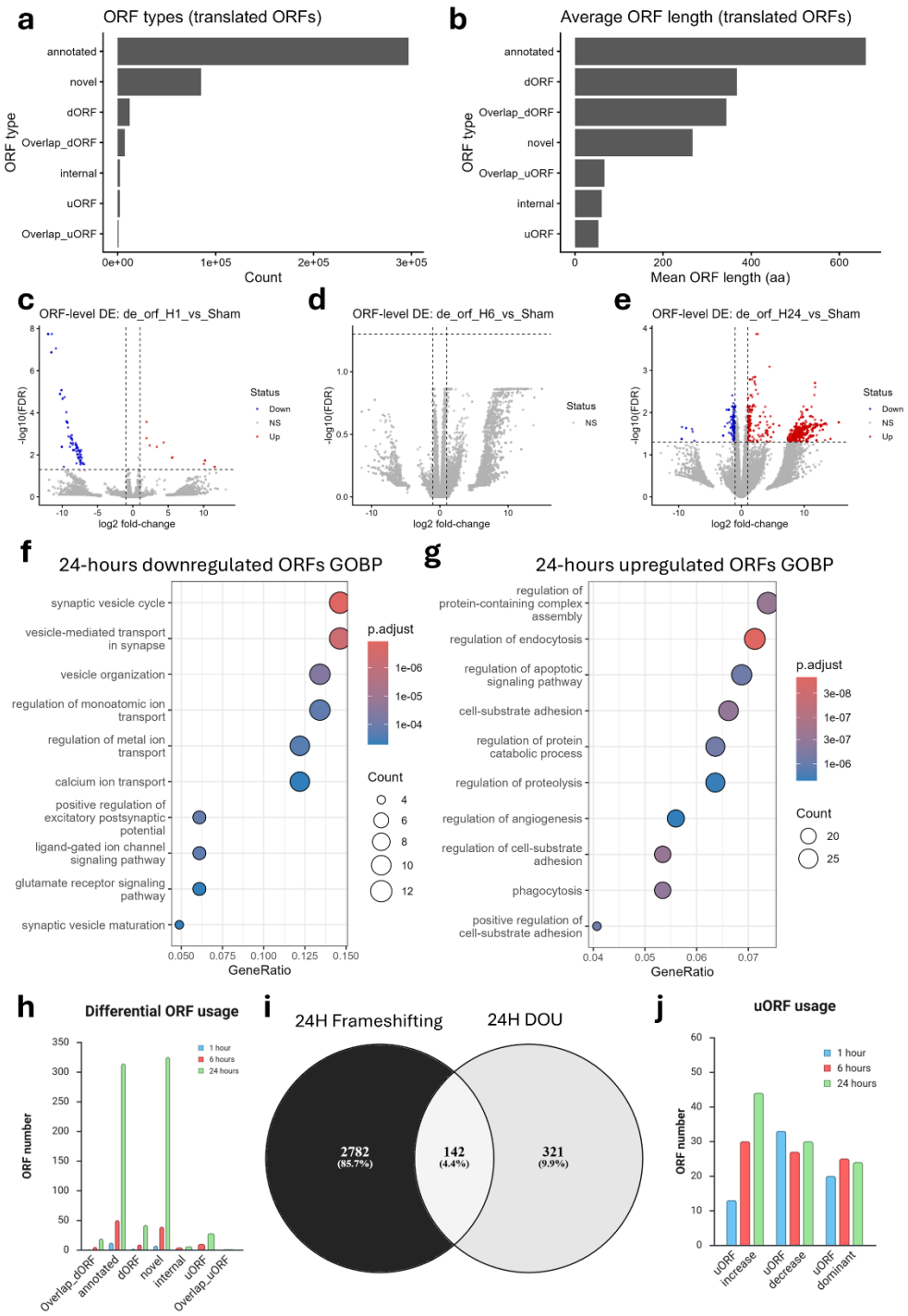


Figure 7: Ischemic stroke leads to transcriptome-wide ORF changes. a: Bar plot showing the ORF subtype counts detected in the Ribo-seq dataset. **b:** Bar plot showing the average translated ORF length for each subtype detected. **c-e:** Volcano plot showing the differentially translated ORFs at each timepoint compared to sham. **f-g:** ORA GOBP pathway analysis of downregulated and upregulated ORFs (using the Gene symbols as input) at the 24-hours timepoint. **h:** Bar plot showing the number of genes with differential ORF usage (DOU) detected by DRIMSeq and the subtypes of the ORFs. **i:** Venn diagram showing the overlap between genes having significant DOU and

1189 frameshifting at 24-hour timepoint. **j:** Bar plot showing genes with uORF usage changes across time
1190 points. uORF dominance is a separate entity from uORF increase and decrease, it shows genes
1191 with their dominant ORF being uORF regardless of change in usage after stroke.

1192

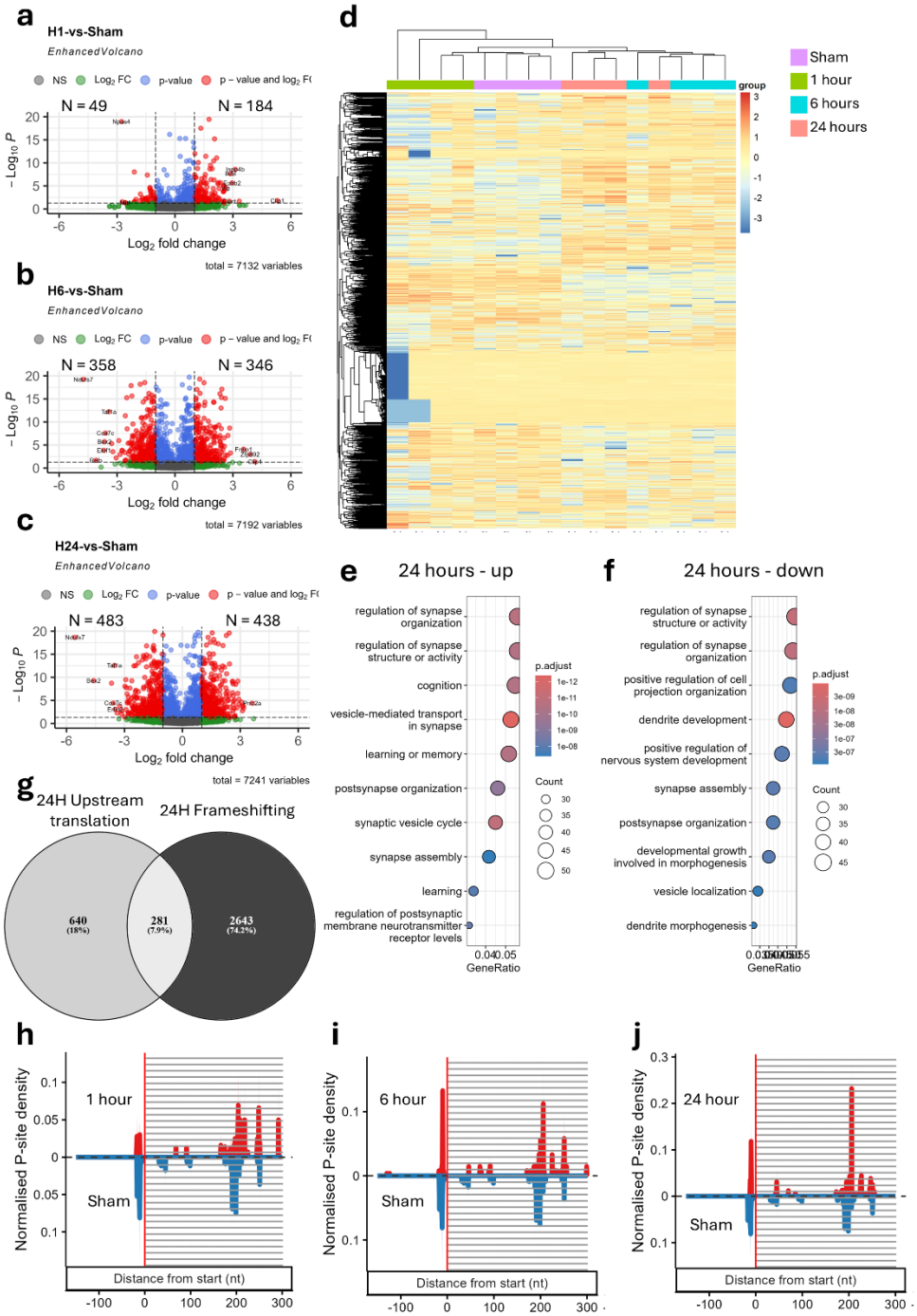


Figure 8: Ischemic stroke leads to progressive upstream translation starts. a-c: Volcano plots showing the number of genes with increased or decreased upstream (5'UTR) RPF ratio compared to CDS RPFs at each time points. **d:** Heatmap showing the gene upstream translation ratio (UTR; ratio of 5'UTR reads to CDS reads per gene). **e-f:** ORA GOBP analysis of up and downregulated genes in the upstream translation analysis at the 24-hour timepoint. **g:** Venn diagram showing the overlap between genes having significant frameshifting and those significant in the upstream translation

1201 analysis. **h-j:** Normalized P-side density mirror plots of DCX gene, which is significant in both
1202 frameshifting analysis and upstream translation analysis across timepoints.