A spatiotemporally resolved atlas of mRNA decay in the *C. elegans* embryo reveals differential regulation of mRNA stability across stages and cell types

Felicia Peng, C Erik Nordgren and John Isaac Murray

Genome Res. published online August 14, 2024
Access the most recent version at doi:10.1101/gr.278980.124

A spatiotemporally resolved atlas of mRNA decay in the *C. elegans* embryo reveals differential regulation of mRNA stability across stages and cell types

Felicia Peng[1], C Erik Nordgren[2], John Isaac Murray[1]*

[1] Department of Genetics, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA, USA

[2] Department of Biology, School of Arts and Sciences, University of Pennsylvania, Philadelphia, PA, 19104, USA

* Corresponding author. Email: jmurr@pennmedicine.upenn.edu

Running title: Atlas of mRNA decay in *C. elegans* embryo

Abstract

During embryonic development, cells undergo dynamic changes in gene expression that are required for appropriate cell fate specification. Although both transcription and mRNA degradation contribute to gene expression dynamics, patterns of mRNA decay are less well-understood. Here we directly measured spatiotemporally resolved mRNA decay rates transcriptome-wide throughout *C. elegans* embryogenesis by transcription inhibition followed by bulk and single-cell RNA sequencing. This allowed us to calculate mRNA half-lives within specific cell types and developmental stages and identify differentially regulated mRNA decay throughout embryonic development. We identified transcript features that are correlated with mRNA stability and found that mRNA decay rates are associated with distinct peaks in gene expression over time. Moreover, we provide evidence that, on average, mRNA is more stable in the germline compared to in the soma and in later embryonic stages compared to in earlier stages. This work suggests that differential mRNA decay across cell states and time helps to shape developmental gene expression, and it provides a valuable resource for studies of mRNA turnover regulatory mechanisms.

Introduction

Development is a highly regulated process that relies on the precise spatial and temporal control of gene expression patterns. Although transcriptional regulation is heavily studied, the regulation of mRNA decay can also be complex and necessarily influences mRNA abundance (Alonso 2012). Transcripts are often specified for decay through the binding of RNA and protein factors to *cis*-regulatory elements within the 3' untranslated region (UTR) (Ray et al. 2013; Vejnar et al. 2019). Cross-talk between factors can lead to differential mRNA decay, such as RNA-binding proteins stabilizing and protecting transcripts from miRNA-mediated repression (Bhattacharyya et al. 2006; Young et al. 2012). Furthermore, computational analyses predict widespread antagonistic or synergistic actions between RNA-binding proteins and miRNAs in

regulating gene expression (Jiang et al. 2013). Translational efficiency and features that influence it, such as codon optimality, are also key determinants of mRNA stability. From yeast to humans, greater translational efficiency is associated with greater transcript stability (Presnyak et al. 2015; Wu et al. 2019). Once transcript turnover is specified, transcripts are ultimately targeted to one of many distinct decay pathways, such as those mediated by the major exoribonucleases XRN1 (5'-3' decay) and the RNA exosome (3'-5' decay) (Łabno et al. 2016). Given the potential complexity in the regulation of mRNA degradation, differential mRNA decay in different lineages, cell types, or developmental stages could contribute to distinct gene expression patterns. The degree to which such regulation occurs throughout development is unclear. Measuring how mRNA degradation varies across different cell types and developmental stages would determine the extent of differential mRNA decay during development and provide the basis for identifying mechanisms for this regulation.

The developmental importance of mRNA degradation is underscored by the maternal-to-zygotic transition, in which maternal gene products in the early embryo must be degraded for the control of development to switch to zygotically-encoded products (Vastenhouw et al. 2019). Global studies in both vertebrate and invertebrate model organisms have measured maternal mRNA decay and identified key RNA-binding proteins and small RNAs involved in the clearance of maternal products (Giraldez et al. 2006; Tadros et al. 2007). In addition, studies of zygotic mRNA degradation have uncovered cell differentiation events that are influenced by mRNA decay. For example, timely decay of *glial cells missing* transcripts in the fly embryo is important for appropriate nervous system differentiation (Soustelle et al. 2008), and stabilization of muscle-specific mRNAs by the RNA-binding protein Human antigen R promotes myogenesis (van der Giessen and Gallouzi 2007). The regulation of mRNA degradation for developmental regulators may therefore be important in patterning fates. As studies of zygotic mRNA decay tend to be on a gene-by-gene basis or in the context of cell culture systems, global studies in

complex organisms will be necessary to establish the extent of developmentally regulated zygotic decay.

The nematode *Caenorhabditis elegans* provides an ideal system for the global study of zygotic mRNA decay due to its invariant cell lineage, experimental tractability, and conservation of RNA decay machinery with humans. In this study, we undertook the first global measurement of mRNA decay rates throughout *C. elegans* embryogenesis at high resolution by pairing a transcription inhibition approach with both bulk and single-cell RNA sequencing. Using this atlas of decay rates, we investigated the contribution of mRNA decay toward specific gene expression patterns across developmental stages and cell types and how differential mRNA degradation may be regulated.

Results

Transcription inhibition by actinomycin D allows mRNA decay measurements in C. elegans embryonic cells

To measure zygotic mRNA decay throughout *C. elegans* embryogenesis, we measured changes in gene expression after transcription inhibition (**Fig 1A**). Primary cell cultures of *C. elegans* embryos divide and differentiate in ways that mirror development of intact embryos, with a similar range of cell types differentiating at roughly the same time they would in intact embryos (Edgar 1995; Christensen et al. 2002). We took advantage of this by dissociating mixed-stage embryos into single cells and treating the cells with the transcription inhibitor actinomycin D (actD) in a time course experiment. As accurately staging or synchronizing large numbers of *C. elegans* embryos is nontrivial, using populations of mixed-staged embryos allowed us to measure mean mRNA decay rates in a population of cells spanning a broad range of embryonic stages, possibly at the cost of reduced detection of very stage-specific genes. Transcription inhibition was efficient, as measured by the rapid decrease in levels of unspliced

transcripts for several housekeeping genes within minutes of actD treatment (**Supplemental Fig S1A**). On the other hand, levels of unspliced transcripts were maintained in cultured cells not treated with actD (**Supplemental Fig S1B**). After an hour of actD treatment, unspliced transcripts remained low and cell viability remained high (> 90%).

To determine half-lives transcriptome-wide, we measured mRNA levels in cells after 0, 10, 20, 40, and 60 minutes of actD treatment. After each time point, samples were spiked with External RNA Controls Consortium (ERCC) control transcripts (Baker et al. 2005) before RNA was isolated and analyzed using total RNA sequencing after ribosomal RNA depletion (**Supplemental Table S1**). An mRNA half-life was calculated for each gene by fitting ERCC-normalized gene expression over time to an exponential decay equation across all replicates. The measured half-lives had a median of 54 minutes and varied by more than an order of magnitude from gene to gene, with most ranging from ~20 to 200 minutes (**Fig 1B, Supplemental Table S2**).

Several pieces of evidence indicate that these half-lives are reproducible and biologically meaningful. First, calculating half-lives separately for each of three biological replicates yielded reproducible half-life estimates (Spearman's ρ > 0.7 for all pairwise comparisons) (**Supplemental Fig S1C-S1E**). We focused our downstream analyses on high confidence half-life measurements by implementing a filtering strategy based on an initial count of 30 mapped reads for each gene and coefficient of variation and 95% confidence interval thresholds for half-lives across biological replicates (**Supplemental Fig S1F, see Methods**).

Second, genes with the slowest or fastest mRNA decay rates were enriched for different functional categories. Gene Ontology analysis found an enrichment among the most stable transcripts for housekeeping functions, such as ribosomal and cytoskeletal proteins (**Fig 1C,**

Supplemental Fig S2A). In contrast, among the most unstable mRNAs there was an enrichment for dynamic developmental processes, such as transcription factor activity and the cell cycle. (**Fig 1C, Supplemental Fig S2B**). These functional associations with RNA stability generally agree with those found in previous studies of decay in other systems (Yang et al. 2003; Narsai et al. 2007; Thomsen et al. 2010; Burow et al. 2015).

Third, we found that RNA stability correlates with temporal dynamics in untreated embryonic cells. High mRNA stability should result in persistent expression over time, while faster mRNA degradation could allow genes to have transient expression (**Fig 1D**). We categorized genes as having either transient or persistent expression over time in a previously published *C. elegans* embryo single-cell atlas (Packer et al. 2019) based on the magnitude and frequency of expression decrease from mother to daughter cell states (see **Methods**). For example, the transcription factor genes *zip-7* and *hlh-14* have highly transient expression, with rapid loss of expression from parent to daughter cell in multiple lineages (**Supplemental Fig S2C-S2D**). On the other hand, the genes *rpl-35* and *pab-1* have highly persistent expression maintained from parent to daughter cell across multiple lineages (**Supplemental Fig S2E-S2F**). These genes show similar temporal dynamics in a separate *C. elegans* whole-embryo time series dataset (Hashimshony et al. 2015); *zip-7* and *hlh-14* increase to peaks in expression around 200 minutes past the four-cell stage before dropping off rapidly, while *rpl-35* and *pab-1* have fairly constant expression over time (**Fig 1E**). In agreement with our expectations, we found that *zip-7* and *hlh-14* transcripts have much shorter half-lives than those of *rpl-35* and *pab-1* (**Fig 1E**). Across the transcriptome, highly transient genes generally had the shortest mRNA half-lives while highly persistent genes generally had the longest mRNA half-lives (**Fig 1F**). These findings are consistent with gene-specific mRNA decay rates having an important contribution to gene expression dynamics.

156 ***Transcript stability correlates with specific sequence features***

157 To identify features that differ between stable and unstable transcripts, we tested for
158 correlations between our mRNA stability measurements and various transcript attributes. This
159 identified numerous features associated with transcript stability.

160

161 Codon optimality concerns the nonuniform decoding rate of codons by the ribosome, with a
162 codon being considered optimal when its cognate tRNA is prevalent and allows for rapid
163 decoding and ribosome movement (Bae and Coller 2022). In examining codon optimality scores
164 between the most stable and unstable 15% of transcripts, we found that the most stable
165 transcripts had significantly greater codon optimality than the least stable transcripts (**Fig 2A**).

166

167 Since the presence of introns has been associated with higher translation, we asked whether
168 intron number is correlated with mRNA stability (Shaul 2017). We found that unstable mRNAs
169 had significantly fewer introns compared to stable mRNAs, with a median of 4.5 and 5.7 mean
170 number of introns per gene, respectively (**Fig 2B**). This was associated with differences in
171 overall gene length; stable transcripts had longer mean coding sequence lengths than unstable
172 transcripts (**Fig 2F**).

173

174 Since 3' UTRs can influence mRNA stability (Mayr 2019), we tested for 3' UTR characteristics
175 correlated with half-lives. Unstable transcripts had significantly shorter 3' UTR lengths compared
176 to stable transcripts, with a median difference of 98 base pairs (**Fig 2C**). Nucleotide composition
177 of both the 3' UTR and coding sequence also varied with stability, with higher GC content for
178 stable transcripts (**Fig 2D-2E**). In line with the findings for 3' UTR and coding sequence length,
179 overall transcript length was significantly longer in stable transcripts compared to unstable
180 transcripts (**Supplemental Fig S3C**). We also found a relationship between mRNA half-lives

and number of 3' UTR isoforms, with half-lives being significantly longer for genes with a higher number of 3' UTR isoforms (**Supplemental Fig S3D-3DE**).

We tested whether these different transcript features provide independent information about mRNA half-lives by using linear regression modeling. Codon optimality explained the largest fraction of variation in mRNA stability (~21%) (**Fig 2G**), while other features such as number of introns, 3' UTR length, and GC content in the coding sequence provided independent information about mRNA stability (**Fig 2H**). Most of the variation in mRNA stability, however, remains unexplained by the examined sequence features.

To identify sequence motifs associated with differences in transcript stability, we used the *de novo* motif-finding program MEME (Bailey et al. 2015) to identify motifs differentially enriched in the 3' UTRs of the most stable and unstable transcripts. No statistically significant motifs were enriched for unstable transcripts, but several motifs were differentially enriched for stable transcripts (**Fig 2I, Supplemental Fig S3A**). To identify potential regulators of RNA stability, we compared these motifs against a database of RNA-binding protein motifs, which were identified by incubating RNA-binding proteins from many eukaryotic species within a complex pool of RNAs (Ray et al. 2013). The motifs resembled poly(C)-binding protein (PCBP), poly(A)-binding protein (PABP), and heterogeneous nuclear ribonucleoprotein (hnRNP) family binding sites (**Fig 2I, Supplemental Fig S3B**). RNA-binding proteins from these families have previously been shown to have transcript-stabilizing effects in other systems (Makeyev and Liebhaber 2002; Geuens et al. 2016; Passmore and Coller 2021). Future experiments will be needed to determine whether they are responsible for the increased stability of transcripts carrying these motifs in *C. elegans* embryos.

These findings highlight that stable and unstable transcripts are enriched for distinct sequence features, consistent with multiple mechanisms regulating mRNA stability. Codon optimality appears to be the single strongest predictor of mRNA stability in the *C. elegans* embryo. Additional complexity to the regulation of mRNA decay may result from other binding sites for regulatory RNAs and proteins that act in specific contexts.

High transcript accumulation is associated with increased mRNA stability

Throughout *C. elegans* embryogenesis, many genes accumulate to high transcript levels. While this requires high transcription rates, we asked whether rapid transcript accumulation is also associated with increased mRNA half-life, as such accumulation would be difficult to achieve in the presence of rapid RNA turnover. Genes with lower accumulation rates, on the other hand, could be more tolerant of RNA turnover rates (**Fig 3A**). For example, moderate transcription rate with moderate mRNA half-life or high transcription rate with high RNA turnover could yield similar slopes of RNA versus time.

Using a whole-embryo RNA sequencing time series (Hashimshony et al. 2015), we identified 985 dynamic genes whose expression peaks at ~200 minutes past the four-cell stage. We categorized these genes as having high (top 20%), medium (middle 20%), or low (bottom 20%) transcript accumulation rates (**Fig 3B**). For example, the high-accumulation gene *cht-1* peaks at an expression level of around 6,300 transcripts per million (TPM), which is in the top 1% of maximum expression across all genes and accounts for ~0.6% of transcripts in the embryo. On the other hand, the medium-accumulation gene *ced-11* peaks at around 150 TPM, and the low-accumulation gene *sec-10* peaks at around 60 TPM.

We found that genes with high accumulation had a median mRNA half-life about 27% longer than that of genes with low accumulation (median $t_{1/2}$ of 52 vs 41 minutes; **Fig 3D**). Similar

results were observed for genes whose expression peaks later, around 350 minutes past the four-cell stage (**Fig 3C**). At this stage, high-accumulation genes had a median mRNA half-life about 45% longer than that of low-accumulation genes (median $t_{1/2}$ of 81 vs 56 minutes) (**Fig 3E**).

We identified motifs differentially enriched in the 3' UTRs of high-accumulation genes compared to the 3' UTRs of low-accumulation genes using MEME (Bailey et al. 2015) (**Fig 3F-3G; Supplemental Fig S4A, S4C**). High-accumulation genes peaking at ~200 minutes were enriched for potential embryonic lethal abnormal vision-like (ELAVL) and RNA-binding motif (RBM) family binding sites, those peaking at ~350 minutes were enriched for potential hnRNP and T cell intracellular antigen 1 (TIA1) family binding sites, and high-accumulation genes at both stages were enriched for potential PCBP binding sites (**Supplemental Fig S4B, S4D**). Several of the corresponding RNA-binding protein genes have *C. elegans* homologs that have distinct peaks in mRNA expression throughout embryogenesis (**Fig 3H**).

Overall, dynamically expressed genes with high transcript accumulation rates tend to have longer mRNA half-lives. This suggests that regulation of both transcription and mRNA decay help these genes accumulate to high transcript levels.

Single-cell RNA sequencing allows measurement of mRNA half-lives at high resolution throughout C. elegans embryogenesis

While our bulk RNA sequencing data supports the idea that many zygotic genes vary in their decay rates, it does not allow us to identify changes in transcript stability for the same gene in different cell types or developmental stages. To globally identify developmentally regulated mRNA decay in the *C. elegans* embryo, we combined actD transcription inhibition with single-cell RNA sequencing (**Fig 4A**).

We dissociated mixed-stage populations of *C. elegans* embryos into single cells and cultured them for a time course in the presence or absence of actD as before. mRNA levels were then measured in single cells using the 10x Genomics single-cell RNA sequencing platform. Cells from 0-minute untreated, 20-minute actD-treated, and 40-minute actD-treated cells were used to calculate mRNA half-lives. An additional 40-minute untreated time point served as a control for the impact of cell culturing. We used this slightly shorter time course to allow for more accurate annotation of actD-treated cells, as modeling from the bulk data indicated that these reduced time points were sufficient to accurately estimate mRNA half-lives for most genes (**Supplemental Fig S5D**). We performed the entire experiment across three biological replicates and in total captured 183,777 cells (**Supplemental Table S1**).

We integrated the three biological replicates using Seurat (Stuart et al. 2019), and both manual and automated annotation approaches were used to identify each cell's terminal cell type or lineage identity and developmental stage (in minutes from the four-cell stage) (**Supplemental Fig S5J**, see **Methods**). Cell type annotations were automatically transferred to this dataset from an existing *C. elegans* embryo single-cell atlas (Packer et al. 2019), using a principal component projection approach implemented in Seurat (Stuart et al. 2019). Manual annotation to refine the epidermal cell type annotations (see **Methods**) was based on cell type-specific marker genes reported in WormBase (Lee et al. 2018). For each cell, the stage of the embryo from which it came from was estimated by correlating its transcriptome with a whole-embryo bulk RNA-seq time series (Hashimshony et al. 2015; Packer et al. 2019). The cells here cover a similar range of stages as those in the existing *C. elegans* embryo single-cell atlas (from gastrulation, around the 50-cell stage, through the early stages of terminal differentiation) (Packer et al. 2019).

We projected the integrated dataset to two dimensions using the Uniform Manifold Approximation and Projection (UMAP) algorithm (Becht et al. 2019; McInnes et al. 2020), which organized the cells by type and stage (**Fig 4B**). From a central group of progenitor cells, trajectories formed that generally reflected progression through embryogenesis and that branched out into major cell types. Projecting data from individual biological replicates using the UMAP algorithm revealed similar cell clustering (**Supplemental Fig S5A-S5C**). These UMAPs resembled the global UMAP generated from our previous *C. elegans* single-cell RNA sequencing atlas, in which cells were immediately captured for sequencing rather than subjected to short-term culturing (Packer et al. 2019) (**Supplemental Fig S6B**). Cells from our mRNA decay atlas that were subjected to short-term culturing also co-embedded well with cells from this previous single-cell RNA sequencing atlas (**Supplemental Fig S6C-S6F, S7A-S7C**). Importantly, transcription inhibition or short-term cell culture did not appear to significantly affect the fraction of cells annotated as each major cell type across biological replicates (**Supplemental Fig S5K**), consistent with past studies showing that cultured embryonic cells develop similarly to their *in vivo* fates (Edgar 1995; Christensen et al. 2002).

Half-lives were calculated in a manner similar to those calculated for our bulk data. However, we were not able to use ERCC control transcripts with the single-cell RNA sequencing approach, so expression of each gene was normalized to the expression of ribosomal protein gene transcripts and further adjusted to account for decay of those ribosomal transcripts (see **Methods**, **Supplemental Fig S5E-S5F**, **Supplemental Table S3**). Final pseudobulk half-lives for well-measured transcripts from each of three single-cell biological replicates were well-correlated with one another (Spearman's ρ > 0.7 for all pairwise comparisons) (**Supplemental Fig S5G-S5I**) and with half-lives calculated from the bulk data (Spearman's ρ = 0.76, **Fig 4C**). This indicates that the single-cell RNA sequencing approach captures comparable half-lives to those determined from the bulk approach.

To examine how mRNA half-lives change across development, we focused first on broad developmental stages and cell types. Half-lives for these cell subsets were calculated in pseudobulk as before. To examine mRNA stability across developmental stages, we divided embryogenesis into three major phases using estimated embryo time: Early (50-200 minutes), Middle (200-350 minutes), and Late (350+ minutes). Early and Middle stages correspond to embryonic cleavage and mostly contain progenitor cell states. The Early stage largely corresponds to gastrulation, the Middle stage corresponds to later rounds of cell division and lineage specification, and Late stage consists mostly of terminally differentiating cells. Overall, we observed increased average mRNA stability over time; the median half-life of transcripts in Late cells was about 38% higher than in Early cells, with Middle cells having an intermediate range of half-lives (**Fig 4D, Supplemental Table S4-6**).

Similarly, we compared global half-lives in broad cell classes (muscle, germline, epidermis, neuron, and pharynx) (**Supplemental Table S7-12**). The most notable difference observed was the overall longer half-lives in the germline compared to in somatic cell types (**Fig 4E, Supplemental Fig S6A**). Half-life distributions were significantly different between somatic cell types as well, but the magnitude of these differences was generally smaller (**Fig 4E**, **Supplemental Fig S6A**). Together these observations suggest that mRNA turnover rates are globally regulated across both developmental stages and cell types.

Small RNA pathways such as the microRNA pathway are thought to target specific subsets of genes for destabilization and translational inhibition (Chekulaeva and Filipowicz 2009). Our cell type-specific mRNA half-life measurements also allowed us to examine the decay of known miRNA targets. *C. elegans* genes encoding subunits of the vacuolar adenosine triphosphatase (V-ATPase) complex, which controls intracellular and extracellular pH (Pamarthy et al. 2018),

are ubiquitously expressed (Gutiérrez-Pérez et al. 2021). However, miR-1, a deeply conserved muscle-specific miRNA whose loss results in muscle cell defects, was found to repress multiple subunits of the V-ATPase complex in muscle cells (Gutiérrez-Pérez et al. 2021). We found that the half-lives of transcripts encoding such subunits (*vha-8*, *vha-12*, *vha-14*) are shorter in muscle cells compared to all other cell types, consistent with the model that miR-1 targets these transcripts for more rapid decay to adjust their levels specifically in muscle cells (Gutiérrez-Pérez et al. 2021) (**Supplemental Fig S7G**). Similarly, we examined the miR-35 and miR-51 families, which are the two miRNA families necessary for *C. elegans* embryonic development (Dexheimer et al. 2020). In each of the Early, Middle, and Late stages, the predicted targets of these miRNA families had significantly shorter mRNA half-lives overall compared to all other transcripts (**Supplemental Fig S7D-S7F**). Our data is thus consistent with the idea that the miR-35 and miR-51 families target transcripts for more rapid degradation.

Differential mRNA decay occurs throughout different developmental stages of C. elegans embryogenesis

While transcripts on average increased in stability over time (**Fig 4D**), we identified a subset of transcripts with faster decay at later stages of embryogenesis (**Fig 5A**). These genes were enriched for terms such as structural constituent of chromatin, transporter activity, and cilium organization (**Fig 5B, Supplemental Fig S8A-S8C**). MEME identified several motifs differentially enriched in the 3' UTR of these genes compared to the 3' UTR of genes for which mRNA decay slows the most over time (Bailey et al. 2015). From Early to Middle stages and Middle to Late stages, motifs were identified that had potential similarities to motifs bound by the human factors HNRNPR, PCBP2, PABPC4, and HNRNPC (Ray et al. 2013) (**Supplemental Fig S9D-S9E**).

To illustrate how this resource can be used to characterize differential mRNA decay, we examined the dynamics of 70 previously identified core cilia component genes (Brocal-Ruiz et al. 2023) These genes generally peaked in expression around 400 minutes, around the end of our Middle stage, and genes annotated with ciliary functions were enriched among genes with faster mRNA decay over time (**Supplemental Fig S9A**). This peak corresponds well to the time when many sensory neurons generate their cilia (Nechipurenko and Sengupta 2017). While the average transcript becomes more stable in Late- vs Middle-stage cells, the average cilia transcript becomes less stable or maintains a similar half-life during this period (**Supplemental Fig S9A**). This pattern was especially pronounced for transcripts encoding ciliary transition zone (Williams et al. 2011) and intraflagellar transport proteins (Blacque et al. 2006; Ou et al. 2007) (**Fig 5C-5D**). The ciliary transition zone is a domain at the base of cilia that controls the entry and exit of ciliary proteins needed for signal transduction (Li et al. 2016). On the other hand, intraflagellar transport mediates protein trafficking along a microtubular axoneme for the assembly and maintenance of cilia (Wang et al. 2021). Our observations suggest that these cilia components are produced at high levels during ciliogenesis, with transcripts destabilized once they are no longer needed at high levels.

Many putative transcription factor genes (Fuxman Bass et al. 2016) displayed changes in mRNA decay over time. We identified 16 transcription factor transcripts with decreasing stability in Middle-stage cells and 14 with decreasing stability in Late-stage cells compared with the prior stage. On average these genes showed decreasing expression during the time when their transcript stability is lower (**Fig 5E-5F**). Thus, differential mRNA decay may allow transcripts of developmental regulators to degrade in a timely manner after accumulating to a certain threshold.

As the median scaled expression patterns of these ciliary and transcription factor genes have distinct peaks, differential mRNA degradation may contribute to their specific patterns of expression throughout embryogenesis. Indeed we found that on average, zygotic-only expressed genes with faster mRNA decay over time have peaks in expression corresponding to the time when stability decreases (**Supplemental Fig S9B**). Importantly, the expression patterns of genes with increased mRNA turnover over time were distinct from those observed for genes with decreased mRNA turnover over time and did not peak as strongly or in the same developmental stage (**Supplemental Fig S9C**).

These results highlight that differential mRNA decay over time is associate with dynamic gene expression patterns throughout embryogenesis. This suggests that the regulation of mRNA degradation makes important contributions to developmental gene expression.

mRNA degradation may contribute to transcription factor dynamics at both the RNA and protein levels

Because transcription factors drive developmental gene expression, their mRNA dynamics are of particular interest in the context of fate specification. Across all developmental stages, transcripts encoding transcription factors (Fuxman Bass et al. 2016) decay more rapidly compared to those encoding other genes (**Fig 6A**). This suggests that transcripts encoding transcription factors may need to achieve faster turnover on average throughout embryogenesis.

To test whether mRNA decay dynamics are associated with protein expression over time, we took advantage of an existing single-cell transcription factor protein expression atlas of the *C. elegans* embryo (Ma et al. 2021). This study used confocal microscopy for the live-imaging of strains expressing almost 300 transcription factor-GFP fusion proteins. Automated cell lineage

tracing of this imaging data allowed for the quantification of protein expression at single-cell and ~1-minute temporal resolution. We categorized transcription factor proteins as either transiently or persistently expressed based on the occurrence of loss or maintenance of expression from parent to daughter cells. Similarly, we categorized transcription factors as transiently or persistently expressed at the mRNA level using a previous single-cell RNA sequencing dataset (Packer et al. 2019). Overall, among the 91 transcription factors well-measured in both protein and mRNA datasets, there was a significant enrichment for concordant protein and RNA dynamics: 36.3% had both transient protein and mRNA expression, and 26.4% had both persistent protein and mRNA expression (**Fig 6B**, P = 0.017). A smaller number had discordant dynamics, including 12.1% with transient protein but persistent mRNA, and 25.3% with persistent protein but transient mRNA (**Supplemental Table S13**). These discordant cases likely reflect distinct types of post-transcriptional regulation. Transient protein expression despite the persistent presence of mRNA would suggest that post-translational events are targeting proteins for rapid degradation or that their translation is limited to specific temporal stages. This could occur for example for proteins only needing to act for a short time window or if it were detrimental if they persist for too long. Potentially, the persisting mRNA could allow cells to rapidly create new protein in the future as needed. The reverse pattern, persistent protein despite transient mRNA dynamics, suggests that the encoded proteins are highly stable.

We hypothesized that transcription factors with concordant RNA and protein dynamics might reflect stronger regulation by RNA degradation. Consistent with this, transcription factors with transient protein and mRNA had shorter mRNA half-lives on average compared to transcription factors with persistent protein and mRNA (**Fig 6C**). This suggests that transcript turnover may contribute to protein dynamics by influencing overall mRNA expression dynamics for these genes. For example, the transcription factors REF-2 and CEH-83 were both transiently expressed, with protein expression being lost across many lineages (**Fig 6D**). The

corresponding mRNAs were also transiently expressed (**Supplemental Fig S10A-S10B**), with unstable mRNA half-lives of 19.7 and 22.0 minutes, respectively. The proteins MEP-1 and LSY-2, on the other hand, were both persistently expressed (**Fig 6E**). The corresponding mRNAs were also persistently expressed (**Supplemental Fig S10C-S10D)**, but *lsy-2* mRNA was much more stable with a half-life of 62 minutes (vs 28.6 minutes for *mep-1*).

These results suggest that mRNA degradation may contribute to the protein dynamics of developmental regulators. In many cases, the regulation of transcription and mRNA decay appear to act in concert to have a coherent impact on corresponding protein levels. However, it is not uncommon for RNA dynamics to differ from protein dynamics, highlighting the complexity associated with the regulation of protein expression.

mRNA stability is correlated with cell type-specific functions

As described above, different somatic cell types have more similar mRNA half-life distributions to one another than they do to germline cells (**Fig 4E**). We asked what categories of genes have faster or slower mRNA decay within each cell class. We identified cell type-specific genes within each class and tested whether their stability differed from broadly expressed genes (see **Methods**). For muscle cells and neurons, cell type-specific transcripts had significantly longer half-lives compared to broadly expressed transcripts, while for the other somatic cell classes, specific and broadly expressed genes had similar stability distributions (**Fig 7A**).

To determine what may drive these differences in half-life distributions, we examined the mRNA half-lives of genes that fell within specific functional categories in muscle and neuronal cells (**Supplemental Fig S12A-S12B**). Within muscle cells, transcripts encoding transcription factors with roles in mesoderm or muscle fate specification (such as *ceh-34*, *tbx-2*, *ceh-20*, *eya-1*, and *hlh-1*) had significantly shorter half-lives overall compared to transcripts encoding muscle

structural proteins (such as *deb-1*, *unc-78*, *unc-112*, *sgn-1*, *tmd-2*) (**Fig 7B**). Similarly, in neurons, transcripts encoding neural fate transcription factors (such as *ceh-5*, *egl-46*, *unc-86*, *ceh-43*, and *hlh-2*) had significantly shorter half-lives compared to transcripts associated with synapses, dendrites, or axons (such as *lnp-1*, *unc-37, mig-2*, *ced-10*, and *unc-53*) **(Fig 7C**). While the overall mRNA half-life distributions for cell type-specific and broadly expressed genes were similar within epidermis and pharynx, in these cell classes we similarly saw faster decay of cell type-enriched transcription factor transcripts and slower decay of structural transcripts such as those encoding cuticle-associated proteins in the epidermis and peptidase inhibitors in the pharynx (**Supplemental Fig S11A-S11F, S12C-S12D**).

We used MEME to identify motifs associated with slow or fast mRNA decay within neurons and muscle (Bailey et al. 2015) (**Supplemental Fig S13A-S13D**). Similar motifs were identified for the most stable and unstable transcripts in the soma as a whole (**Supplemental Fig S13E-S13F**), suggesting that these motifs may not be unique to muscle and neurons. While the motifs found in stable transcripts included those that resemble PCBP and hnRNP family binding sites (**Supplemental Fig S13A, S13C, S13E**), the CSGGU motif associated with unstable transcripts had no significant matches when examined against a database of known RNA-binding protein motifs (Ray et al. 2013).

To begin to examine how extensive differential mRNA decay may be across somatic cell types, we focused on comparing half-lives between muscle, epidermis, and neuronal cells. We focused on these cell types as they were the most abundant in our dataset, allowing us to accurately measure half-lives for more genes. We focused on the Middle stage to avoid differences in half-life due to tissue-specific differences in expression timing, with the Middle stage having a large number of cells. Differential decay between cell types during that stage could conceivably contribute to both fate specification and terminal differentiation.

Comparing genes with half-lives measured in muscle and epidermis, ~3% (58/1926) had half-lives at least 1.5-fold longer in the muscle and ~6% (115/1926) in the epidermis. Between muscle and neurons, ~18% of genes (442/2455) had mRNA half-lives at least 1.5-fold longer in muscle and ~1.2% of genes (30/2455) had mRNA half-lives at least 1.5-fold longer in neurons. Lastly, between epidermis and neurons, ~21.3% of genes (539/2531) had mRNA half-lives at least 1.5-fold longer in epidermis and ~0.9% of genes (23/2531) had mRNA half-lives at least 1.5-fold longer in neurons. While mRNA half-lives of shared genes were overall shorter in neurons compared to in muscle and epidermis, the impact of differences in half-lives for individual transcripts across the different cell types is unclear. The expression of many of these genes in different cell types, however, may be shaped in part by differences in mRNA decay. For example, *zig-8* transcripts were more stable in neurons than in muscle, with half-lives of 37.5 and 24.3 minutes, respectively; *jac-1* transcripts were more stable in neurons than in epidermis, with half-lives of 37.6 and 23.5 minutes, respectively; and *vha-20* transcripts were more stable in epidermis than in muscle, with half-lives of 92.2 and 60.8 minutes, respectively. The expression of these genes in our existing *C. elegans* embryo single-cell atlas (Packer et al. 2019) highlights that they are well-expressed in both cell types of interest; furthermore, gene expression in the cell type with a shorter mRNA half-life decreases in later embryogenesis while gene expression in the cell type with a longer mRNA half-life tends to persist or increase, suggesting a role of regulated mRNA decay in cell type-specific persistence (**Fig 7D**).

Differential mRNA decay occurs across germline and somatic cells in the C. elegans embryo

The most striking cell type-specific difference in mRNA degradation was between the germline (median $t_{1/2}$ = 62 minutes) and soma (median $t_{1/2}$ = 36 minutes) (**Figs 4E, 8A**). When comparing

mRNA half-lives of shared genes between these cell classes, about 80% of transcripts were more stable in the germline than they were in the soma (**Fig 8B**).

In the *C. elegans* embryo, the germline is derived from the P4 blastomere, which is born at the 24-cell stage after a series of asymmetric cell divisions (Wang and Seydoux 2013). At the 88-cell stage, P4 divides into the primordial germ cells Z2 and Z3, which do not divide further throughout embryogenesis. P4, Z2 and Z3 are thought to be largely transcriptionally quiescent (Wang and Seydoux 2013). Despite this, our *C. elegans* embryo single-cell atlas (Packer et al. 2019) identified substantial quantitative differences in gene expression between early, mid-embryo, and late germline cells. This raises the question of whether differences in mRNA stability govern the quantitative maturation of the germline transcriptome during embryogenesis.

To test this, we compared mRNA stability between transcripts that increase or decrease in relative abundance across time in the germline. If expression differences result from differential mRNA decay, we would expect genes with decreasing relative abundance to have shorter mRNA half-lives compared to genes with constant expression (**Fig 8C**). Similarly, we would expect genes with increasing relative abundance to have longer mRNA half-lives. On the other hand, if changes in relative abundance result from previously unrecognized transcription, we might expect comparable half-life distributions for each group (**Fig 8D**).

We found that changes in germline mRNA abundance over time correlated with mRNA stability. Transcripts that decrease in abundance over time had shorter half-lives overall, and transcripts that increase in abundance over time had longer half-lives overall (**Fig 8E**). This effect was stronger than the increase in mRNA half-lives we saw for genes with greater soma-specific expression over time (**Fig 8F**). These results are consistent with a model where changes in

mRNA levels throughout embryogenesis result from mRNA stability differences in the germline and from a mix of transcriptional and post-transcriptional mechanisms in the soma.

The changes observed in transcript abundance in the germline raise the question of what types of genes increase or decrease over time. Gene Ontology analysis revealed that germline-specific genes were enriched for genes associated with mRNA 3' UTR-binding (**Supplemental Fig S12E**). We found that genes with this annotation had significantly shorter mRNA half-lives in the germline compared to those of mRNA-binding genes not associated with 3' UTRs or genes not annotated as RNA-binding (**Fig 8G**). This was not the case in the soma (**Fig 8H**), suggesting that germline maturation might be specifically regulated by decay of transcripts encoding RNA-binding proteins. Transcripts encoding mRNA-binding proteins with faster decay in the germline than in the soma included *mex-5* and *mex-6* transcripts, for which half-lives in the soma were more than twice as long as they were in the germline (**Fig 8I**). MEX-5 and MEX-6 are nearly identical proteins that inhibit the translation of some germline proteins in the soma, and they become more highly expressed in somatic cells versus germline cells over time in the early embryo (Schubert et al. 2000). Our results suggest that maturation of the germline in part results from germline-specific turnover of maternally provided somatic regulators like *mex-5* and *mex-6.*

Gene Ontology analysis further found an enrichment among the most stable transcripts in the germline for those encoding proteins with functions related to peptide biosynthetic process (**Supplemental Fig S14A**). While transcripts in the germline encoding proteins involved in translational elongation and the positive regulation of translation were significantly more stable compared to all other transcripts, the same general trend was observed in neuron, muscle, epidermis, and pharynx cells (**Supplemental Fig S14B-S14F**). Thus it is unclear whether such transcripts play specific roles in promoting the germline fate during postembryonic development

or are just generally stable across all cell types. To characterize transcripts preferentially stable in the germline, we identified genes with mRNA half-lives that were more than two-fold longer in the germline compared to in the soma. These genes were largely maternally expressed, as 148/176 genes were detected in one-cell embryos in a *C. elegans* whole-embryo RNA sequencing time series (threshold of TPM > 5) (Hashimshony et al. 2015). The expression over time of genes whose transcripts were preferentially stable in the germline was generally more persistent in the germline compared to in the soma. This is consistent with an important role for the differential stability of such transcripts in determining the late-embryonic germline transcriptome (**Supplemental Fig S14G-S14H**).

Discussion

Using a transcription inhibition approach and RNA sequencing, we measured mRNA half-lives transcriptome-wide throughout *C. elegans* embryogenesis. We found that global mRNA decay rates differ between cell types and developmental stages, identified gene attributes and sequence motifs associated with different rates of decay, and showed that regulated RNA turnover is associated with dynamic gene expression patterns. Our findings emphasize that developmental gene expression patterns result from regulation of both transcription and RNA decay.

A major advantage to the transcription inhibition approach is that it can be readily applied to any RNA sequencing platform, and the downstream analyses to calculate mRNA half-lives transcriptome-wide are relatively straightforward in a dynamic situation like early embryogenesis. In addition, general concordance in measured mRNA decay rates has been observed between transcription inhibition approaches and approaches that maintain transcription, such as metabolic labeling, within mammalian cells and fly embryos (Burow et al. 2015; Herzog et al. 2017). A limitation to our approach is that transcription inhibition is best

applied to primary cell cultures, as the chitin eggshell of the *C. elegans* embryo is impermeable to drugs like actD. Thus our sequencing-based mRNA quantification cannot be readily applied to or validated by alternative approaches that would be best applied to whole embryos, such as fluorescence *in situ* hybridization. However previous work has shown that in *C. elegans* embryos transcript abundance measured by single molecule RNA FISH generally agrees well with that determined by single-cell RNA sequencing (Raj et al. 2010; Nair et al. 2013; Tintori et al. 2016; Sivaramakrishnan et al. 2023).

Our analyses suggest that mRNA half-lives for most genes in the *C. elegans* embryo range from ~20 to 200 minutes . Previous studies in cell culture systems found that median half-life is roughly proportional to cell cycle length (Vejnar et al. 2019). A study of decay in early *Drosophila melanogaster* embryos (nc11 through cephalic furrow formation), which have comparable cell cycle lengths (~10-180 minutes) to those in the *C. elegans* embryo (Sulston et al. 1983; Bao et al. 2008), observed a similar range of half-lives to ours for ~260 zygotic transcripts (Beadle et al. 2023). In older fly embryos (stage 12 to 15), which have longer cell cycles, the median half-life was longer at 73 minutes (Burow et al. 2015). This was longer than our *C. elegans* median half-life of 54 minutes, perhaps reflecting that older fly embryos no longer have the short mitotic cycles and rapid development characteristic of early fly embryos. We similarly found that mRNA half-lives increase over time during *C. elegans* embryogenesis. It is important to emphasize that some genes go against this trend, however, and have faster mRNA decay at later embryonic stages. The fact that these genes are enriched for specific categories of genes, such as transcription factor genes or genes involved in building the cilium, suggests that the rate of RNA turnover across time is heavily regulated.

While these changes are consistent with a relationship between the cell cycle and RNA turnover, they could also reflect the need for progenitor cells to degrade mRNAs in a timely

manner to allow cell differentiation to proceed appropriately, as has been observed in other systems (Soustelle et al. 2008; Abbadi et al. 2019). Consistent with this, we observed that mRNAs encoding transcription factors had faster mRNA degradation compared to other transcripts, and fast RNA turnover was significantly associated with transient transcription factor protein expression. These results highlight that mRNA degradation likely plays a key role in modulating the protein expression of developmental regulators.

We found that specific functional categories of genes had distinct mRNA decay dynamics in different cell types. For example in neurons, transcripts encoding neural fate-specifying transcription factors had relatively short half-lives, while those associated with synapses, dendrites, or axons had relatively long half-lives. Similar findings were observed for neuron-specific mRNA half-lives in the fly embryo (Burow et al. 2015). Perhaps the unique architecture of neuronal cells can help explain these differences in the regulation of mRNA degradation. Neurons are known to localize some transcripts distant from the cell body such as at synapses, their stabilization may facilitate regulated local translation. We also found that transcripts vary substantially in their germline mRNA stability, and this is correlated with changes in germline expression over time. This suggests that the regulation of mRNA degradation plays an important role in shaping the germline transcriptome. The relatively rapid decay of transcripts encoding mRNA 3' UTR-binding proteins in the germline further points to the importance of post-transcriptional regulation in this cell type.

From yeast to human cells, greater translational efficiency is associated with greater transcript stability (Presnyak et al. 2015; Wu et al. 2019). Our observation that stable transcripts had significantly higher optimal codon usage compared to unstable transcripts in *C. elegans* embryos is consistent with this. We also found that stable transcripts tend to have more introns than unstable transcripts, in line with the presence of introns having a stabilizing effect in

644 mammals and plants (Shaul 2017) and increasing protein production in *C. elegans* (Crane et al.
645 2019). The relationship between 3' UTR length and mRNA stability appears to differ depending
646 on the context. In zebrafish, longer 3' UTRs conferred resistance to codon-mediated
647 deadenylation for maternal transcripts (Mishima and Tomari 2016). Our data align with this
648 finding, as stable transcripts had significantly longer 3' UTRs compared to unstable transcripts.
649 Future studies examining the effect of altering such features will test the potentially
650 combinatorial causal relationships between these features and mRNA stability.
651
652 Examining the mRNA half-lives of genes that accumulate to high transcript levels revealed that
653 such genes undergo overall slower transcript turnover compared to genes that accumulate to
654 low transcript levels. Some genes accumulate so rapidly in early embryogenesis that their
655 transcription rates must be high, irrespective of their mRNA turnover rates (Sivaramakrishnan et
656 al. 2023). This emphasizes that the regulation of both transcription and mRNA degradation can
657 work together to achieve high gene expression. The 3' UTRs of genes with high transcript
658 accumulation were significantly enriched for several motifs, including poly(C) motifs. In
659 mammals, poly(C)-binding proteins have been implicated in the stabilization of transcripts
660 (Makeyev and Liebhaber 2002), and transcripts that contain a poly(C) motif may be stabilized in
661 *C. elegans* oocytes (Stoeckius et al. 2014). Future studies will establish whether poly(C)-binding
662 protein homologs act as transcript stabilizers in the *C. elegans* embryo.
663
664 Differential isoform usage, particularly among 3' UTR isoforms, could be another potential
665 contributor to differential mRNA decay of a gene across developmental stages or cell types. In
666 *C. elegans*, hundreds of genes have been identified to change whether the proximal or distal 3'
667 UTR isoform is more dominantly expressed throughout various stages of embryogenesis (Levin
668 et al. 2020). Alternative polyadenylation has also been probed across intestine, muscle,
669 neurons, arcade and intestinal valve cells, seam cells, and hypodermal tissues in mixed stage

cultures of *C. elegans*, with an average of 31% of genes found to express 3' UTR isoforms in a tissue-specific manner (Blazie et al. 2015; Blazie et al. 2017). Technical obstacles, such as incomplete isoform annotations and ambiguous mapping of RNA sequencing reads, currently make the measurement of isoform-specific half-life measurements difficult; overcoming these challenges to add additional context to mRNA decay rates would be an intriguing future direction.

Methods

Embryonic cell isolation and culturing. *C. elegans* adults (N2 strain) were grown on large, enriched peptone plates seeded with NA22 bacteria. Mixed-stage embryos were released from adult worms using hypochlorite treatment followed by two washes with previously described complete L-15 cell culture media (Bianchi and Driscoll 2006). To obtain cell suspensions, embryos were treated with 0.5 mg/ml chitinase in egg buffer on ice until the eggshells were dissolved (about 10 minutes). Then embryonic cells were dissociated using a 3 ml syringe fitted with 21 1/2 gauge needle until >80% of embryos were disrupted. The cell suspension was passed through a 10 µM filter before being washed then resuspended in complete L-15.

Bulk RNA sequencing transcription inhibition time course. Following embryonic cell isolation, cells were treated with trypan blue and counted using a hemocytometer. Cell cultures at a concentration of ~1 million cells/ml were treated with 2 µg/ml of the transcription inhibitor actinomycin D (actD) over a time course of 0, 10, 20, 40, and 60 minutes at 20 degrees C. Each time point had 2 ml of cell suspension (~2 million cells). After each time point, RNA was isolated from cells using QIAzol that contained a 1:350,000 dilution of the ERCC spike-in kit (Baker et al.

2005). Libraries were prepared using the Illumina Ribo-Zero Plus rRNA Depletion Kit and sequenced on an Illumina NextSeq 500.

Bulk RNA sequencing computational analysis. The RNA sequencing data were processed using the STAR (Spliced Transcripts Alignment to a Reference) alignment tool (Dobin et al. 2013) against the WormBase WS277 reference transcriptome with 3' UTR extensions as previously carried out (Packer et al. 2019). VERSE was used to get read counts for genes (Zhu et al. 2016). Genes had to have a count greater than 30 at the 0 minute time point across all biological replicates for their expression data to be used in downstream analyses. For half-life comparisons between biological replicates, only well-measured genes were examined. These were genes with a count greater than 30 at the 0 minute time point whose expression throughout the transcription inhibition time course fit an exponential decay model $R^2 \geq 0.75$ within each replicate.

Single-cell RNA sequencing transcription inhibition time course. Cell cultures at a concentration of ~1 million cells/ml were treated with 2 µg/ml of the transcription inhibitor actinomycin D (actD) over a time course of 0, 20, and 40 minutes at 20 degrees C. Each time point had 2 ml of cell suspension (~2 million cells). After each time point, cells were immediately put on ice and kept on ice until the remaining time points were collected, and processed on ice in all subsequent steps until 10x collection. At the end of the 40 minute time course, cells were washed with complete L-15, washed with 1x PBS with 0.4% BSA, then resuspended in 1x PBS with 0.4% BSA. Single cell capture and library preparation followed 10x Genomics published protocols for the Chromium Next GEM Single Cell 3' Reagent Kits v3.1, with the objective of recovering ~10,000 cells for each time point. Libraries were sequenced on an Illumina NextSeq 500.

721 **Single-cell RNA sequencing computational analysis.** The data were processed with the 10x
722 Genomics CellRanger pipeline, aligning reads to the WormBase WS277 reference
723 transcriptome with 3' UTR extensions as previously carried out (Packer et al. 2019). Ambient
724 RNA contamination was removed using the SoupX algorithm (Young and Behjati 2020) with
725 default settings. The data were then visualized using dimensionality reduction methods for
726 either merged datasets within each biological replicate or an integrated dataset of all biological
727 replicates. Single-cell transcriptomes were first projected into 150 dimensions using PCA and
728 then projected into two dimensions using the UMAP algorithm. For each cell, the age of the
729 embryo from which it came from was estimated by correlating its transcriptome with a whole-
730 embryo bulk RNA-seq time series as previously described (Hashimshony et al. 2015). Cells
731 were then automatically annotated with their corresponding cell type or manually annotated
732 based on cell type-specific marker genes reported in WormBase (Lee et al. 2018). Automated
733 annotation was done using Seurat (Stuart et al. 2019) to project the PCA structure of a
734 reference, an existing *C. elegans* embryo single cell atlas (Packer et al. 2019), to our single-cell
735 data. Epidermal cells were manually annotated to better include progenitor cells that the
736 reference dataset did not include. Within the cell types we examined, doublets were manually
737 removed. This was done by identifying clusters of doublets in iterated UMAP projections of the
738 data on the basis of co-expression of cell type-specific marker genes. Genes with a UMI count
739 greater than 30 at the zero minute time point combined across all biological replicates were
740 used in downstream analyses. For half-life comparisons between biological replicates, only well-
741 measured genes were examined. These were genes with an UMI greater than 30 at the 0
742 minute time point whose expression throughout the transcription inhibition time course fit an
743 exponential decay model $R^2 \geq 0.75$ within each replicate.
744
745 **mRNA half-life calculations.** For the bulk RNA sequencing data, gene counts were normalized
746 to the sum of counts corresponding to ERCC RNAs. Normalized gene counts were then fit to an

exponential decay model ($A(t) = A_0e^{-kt}$) using nonlinear least squares regression in R (R Core Team 2021), allowing a decay rate constant (k_{decay}) to be calculated for each gene. A half-life was then calculated for each gene using the following equation: $t_{1/2} = \ln(2)/k_{decay}$. 95% confidence intervals for k_{decay} were determined using the R package nlstools (Baty et al. 2015). For the single-cell RNA sequencing data, half-lives were calculated in a similar manner with the number of UMIs per gene. Transcript abundance was normalized to the sum of UMIs corresponding to ribosomal protein genes, since their transcripts were highly stable in our spike-in control-normalized bulk data (median $t_{1/2}$ = 295.5 minutes). This normalized expression was further adjusted by a correction factor to account for the decay of ribosomal protein genes over time ($0.5^{\text{time point/295.5 minutes}}$).

A combination of coefficient of variation and confidence intervals was used to filter out poorly measured half-lives. Gene half-lives were included in downstream analyses if their coefficient of variation (standard deviation/mean × 100) across biological replicates was ≤ 50% or the fold-change between the upper limit of its 95% confidence interval and measured half-life was ≤ 3. To better include high-stability mRNAs, genes with half-lives > 100 minutes were allowed a looser filtering strategy. Such genes were kept if their half-lives had a coefficient of variation ≤ 75% or fold-change between the upper limit of its 95% confidence interval and measured half-life ≤ 4.

Gene Ontology analysis. Gene Ontology analysis was performed using the WormBase Enrichment Analysis tool (Angeles-Albores et al. 2018) against relevant background sets of genes. For example, for the bulk RNA sequencing data examining the top 15% slow- and fast-decaying genes, the background set of genes used was all genes that met our moderate half-life filtering metric.

Transcript sequence feature analysis. 3' UTRs across the *C. elegans* transcriptome were retrieved from the WormBase ParaSite database (Howe et al. 2017) on March 9, 2023. Only 3' UTRs with ≥ 8 nucleotides were used for downstream analyses. To examine percent GC content, the longest 3' UTR was used for genes with multiple isoforms. To examine the correlation between 3' UTR length and half-life, mean UTR length across isoforms for each gene was calculated.

Coding sequences across the *C. elegans* transcriptome were retrieved from the WormBase ParaSite database(Howe et al. 2017) on March 4, 2023. To examine percent GC content, the longest coding sequence was used for genes with multiple isoforms. To examine the correlation between coding sequence length and half-life, mean coding sequence length across isoforms for each gene was calculated. *C. elegans*-specific tRNA adaptation index (tAI) weights were used from stAIcalc (Sabi et al. 2017) (http://tau-tai.azurewebsites.net/) to calculate a tAI value for each gene, with the longest coding sequence being used for genes with multiple isoforms. The tAI value for each gene was the geometric mean of tAI weights for all codons in the coding sequence (Yoon et al. 2018).

Introns were identified using the TxDb.Celegans.UCSC.ce11.refGene annotation package (Team BC, Maintainer BP 2019) for TxDb objects. To examine the correlation between the number of introns and half-life, the mean number of introns across isoforms for each gene was calculated.

The 'regsubsets' function from the R package 'leaps' (Miller TLboFcbA 2024) was used for linear multiple regression analysis. mRNA half-life was the modeled variable and the sequence features of interest (tAI, mean intron, mean 3' UTR length, mean coding sequence length, GC content in 3' UTR, GC content in coding sequence) the predictors. The individual contribution of each sequence feature to variation in half-lives was evaluated by having each feature as the only predictor in the model.

Motif analysis. MEME differential enrichment was used for motif analysis (https://meme-suite.org/meme/tools/meme). The minimum width for motifs was set to 5 and the maximum width for motifs was set to 15. For motif analysis in 3' UTRs, the longest 3' UTR was used for genes with multiple isoforms. To identify motifs in the 3' UTRs of the top 15% slow-decaying genes in our bulk data, the 3' UTRs of the top 15% fast-decaying genes were used as control sequences (and vice versa).

Query motifs were compared to a database of RNA-binding protein motifs, which were identified from in vitro assays (Ray et al. 2013). The Tomtom motif comparison tool was used to search each query motif against this database of target motifs, with an alignment produced for each significant match (Gupta et al. 2007). The E-value compensates for multiple testing by multiplying each p-value by twice the number of target motifs, and the default Tomtom significance threshold for the E-value is < 10, though we only highlighted those with an E-value < 3. Pearson correlation coefficient was used for the motif comparison function. For more details on the workings of the motif comparison tool, see the Tomtom manual: https://meme-suite.org/meme/doc/tomtom.html?man_type=web

Determining transient or persistent gene expression. We identified genes with highly transient, transient, persistent or highly persistent expression over time using our previously published *C. elegans* embryo single cell atlas (Packer et al. 2019). Highly transient genes were defined as genes whose expression was lost from parent to daughter cell in ≥ 30 comparisons or ≥ 30% of comparisons. Transient genes were defined as genes whose expression was lost from parent to daughter cell in ≥ 10 comparisons or ≥ 10% of comparisons, excluding highly transient genes. Highly persistent genes were defined as genes whose expression was maintained from parent to daughter cell in ≥ 30 comparisons or ≥ 30% of comparisons. Furthermore, these genes had to have ≤ 1 comparison and ≤ 1% of comparisons with a loss in expression from parent to daughter. Persistent genes were defined as genes whose expression

was maintained from parent to daughter cell in ≥ 10 comparisons or ≥ 10% of comparisons, excluding highly persistent genes. Furthermore, these genes had to have ≤ 5 comparisons and ≤ 5% of comparisons with a loss in expression from parent to daughter.

Determining transient or persistent protein expression for transcription factors. To characterize transcription factor protein expression dynamics, we used expression data provided from a single-cell transcription factor protein expression atlas of the *C. elegans* embryo (Ma et al. 2021). A transcription factor was considered expressed in a cell if its expression was > 1.5xIQR below the first quartile for protein expression across all cells. Transient proteins were defined as proteins whose expression was lost from parent to daughter cell in ≥ 10 comparisons or ≥ 10% of comparisons. Persistent proteins were defined as proteins whose expression was maintained from parent to daughter cell in ≥ 10 comparisons or ≥ 10% of comparisons. Proteins with data from more than one reporter strain were only characterized as transient or persistent if the strains agreed with one another.

Hierarchical clustering of gene expression patterns. Gene expression was first scaled using the ‘scale’ function in R (R Core Team 2021), then a distance metric was calculated between each pair of genes using the ‘dist’ function. Hierarchical clustering was performed on these distance values using the ‘hclust’ function. Cluster number was chosen based on visualization with a dendrogram.

Determining genes with high transcript accumulation. To characterize genes in the staged embryo RNA sequencing data (Hashimshony et al. 2015) whose expression peaks around 200 minutes past the four-cell embryo stage as having high, medium, or low transcript accumulation, we calculated a slope value for each gene. This slope = (maximum expression from embryo stages 10 to 200 minutes - minimum expression from embryo stages 10 to 200 minutes) /

maximum stage in minutes - minimum stage in minutes. Slopes from 0 to 0.2 quantile were considered low-accumulating genes, slopes from 0.4 to 0.6 quantile were considered medium-accumulating genes, and slopes from 0.8 to 1.0 quantile were considered high-accumulating genes. A similar analysis was done for genes whose expression peaks around 350 minutes past the four-cell embryo stage.

Determining cell type-specific and broadly expressed genes. Data from our *C. elegans* embryo single cell atlas (Packer et al. 2019) was used to determine gene expression within muscle, germline, epidermis, neuron, and pharynx cells in transcripts per million (TPM). A gene was considered to be specific to a given cell type if it was expressed more than twofold greater in that cell type compared to all other cell types. For example, a gene was considered muscle-specific if it was expressed more than twofold greater in muscle than it was in germline, epidermis, neuron, and pharynx. All other genes were considered broadly expressed. 10 TPM was added to each gene to avoid issues in comparing unexpressed or poorly measured genes.

Examining gene-specific differential mRNA decay between cell types. To examine mRNA half-lives between muscle, epidermis, and neuron, we used Middle stage-specific half-lives. To avoid comparing high stability transcripts that could have artificially large differences in half-lives between cell types, we only considered genes with an mRNA half-life < 75 minutes in at least one cell type. A gene was considered to have differential mRNA decay between cell types if it had an mRNA half-life at least 1.5-fold longer in one cell type compared to the other.

Defining maternal and zygotic-only genes. Zygotic-only genes were determined using a whole-embryo time series RNA sequencing dataset of the *C. elegans* embryo (Hashimshony et al. 2015). Genes with transcript per million (TPM) > 5 within one-cell embryos were considered to be maternally-expressed genes. All other genes were considered to be zygotic-only genes.

miRNA analyses. Predicted mRNA targets of the miR-35 and miR-51 families were identified from TargetScanWorm, release 6.2 (https://www.targetscan.org/worm_52/). Half-lives for these predicted targets were calculated separately for Early, Middle, and Late stage cells. Half-lives for the V-ATPase complex transcripts were calculated from Middle and Late stage cells.

Data access

All RNA sequencing data generated in this study have been submitted to the NCBI BioProject database (https://www.ncbi.nlm.nih.gov/bioproject) under accession number PRJNA1061171. All R codes regarding data visualization and analyses are available at GitHub (https://github.com/fe-peng/celegans_embryo_mRNA_decay) and as Supplemental Code. A VisCello object for visualization of the three single-cell RNA sequencing biological replicates in an integrated dataset is available at Zenodo (https://doi.org/10.5281/zenodo.10499334).

Competing interest statement

The authors declare no competing interests.

Acknowledgements

We thank members of the Murray lab, the Penn Worm Group, Naveen Jain (University of Pennsylvania, Arjun Raj lab), Calvin Huang (University of California, Davis, Bo Liu lab), and

Chenxin Li (University of Georgia, Robin Buell lab) for providing valuable discussion and comments on the manuscript. We also thank Jean Rosario and Catherine Lucey (University of Pennsylvania, Junhyong Kim lab) for their help in sequencing RNA libraries. This work was funded by R35GM127093, F31HD10785, and T32GM008216.

Author contributions: F.P. and J.I.M. conceived and designed the study; F.P. performed the experiments; C.E.N. created new genome and transcriptome references and processed the bulk and single-cell RNA sequencing data; F.P. performed analyses and visualization; J.I.M. supervised the analyses; F.P. wrote the original draft; F.P., C.E.N., and J.I.M. reviewed and edited the manuscript.

Figure legends

Fig 1. Transcription inhibition by actinomycin D allows mRNA decay measurements in *C. elegans* embryonic cells. (A) A schematic representation of the approach to measure mRNA half-lives in the *C. elegans* embryo using transcription inhibition and bulk RNA sequencing. (B) Distribution of mRNA half-lives across three biological replicates that met a moderate filtering strategy. Median half-life was 54 minutes (black line). 72 genes with half-lives greater than 200 minutes not shown. (C) Select Gene Ontology categories enriched among the top 15% stable and unstable transcripts. Background set of genes used was all genes that met our moderate mRNA half-life filtering metric. (D) A schematic representation of a gene with highly persistent expression and a gene with highly transient expression and the predicted overall stability of their transcripts. (E) *Left.* Expression over time of highly persistent (*rpl-35*, *pab-1*) and highly transient (*zip-7*, *hlh-14*) genes from a whole-embryo RNA sequencing time series (Hashimshony et al. 2015). *Right.* The measured mRNA decay of the corresponding genes from our bulk RNA

928 sequencing data, with each point representing normalized transcript abundance from one of
929 three biological replicates. (F) Box plots showing the mRNA half-life distributions of genes
930 characterized as highly transient, transient, persistent, or highly persistent transcriptome-wide.
931 Numbers to the left of the box plots are median half-lives within each group. Numbers above the
932 box plots are the number of genes within each group. P-values comparing median half-lives
933 were calculated using the Wilcoxon rank sum test. Outliers not shown: from left to right, 1, 5, 50,
934 and 57 genes within each group had mRNA half-lives > 150 minutes.
935
936 **Fig 2. Transcript stability correlates with specific sequence features.** The relationship
937 between transcript stability and different sequence features was examined using the top 15%
938 stable and unstable transcripts identified in our bulk RNA sequencing data. The box plots
939 compare different features between all, stable, and unstable transcripts. Numbers to the left of
940 the box plots are median values within each group. Numbers above the box plots are the
941 number of genes within each group. P-values comparing median values were calculated using
942 the Wilcoxon rank sum test. Outliers not shown. (A) Distribution of codon optimality scores, as
943 measured using tRNA adaptation index (tAI) values for each gene. Higher tAI value
944 corresponds to greater codon optimality. (B) Distribution of the number of introns averaged
945 across all splice isoforms of a gene. (C) Distribution of 3' UTR length averaged across all 3'
946 UTR isoforms of a gene. (D) Distribution of percent GC content in the coding sequence of
947 genes. For genes with multiple coding sequence isoforms, the longest isoform was used. (E)
948 Distribution of percent GC content in the 3' UTR of genes. For genes with multiple 3' UTR
949 isoforms, the longest isoform was used. (F) Distribution of coding sequence length averaged
950 across all splice isoforms of a gene. (G) Linear regression was used to identify the percent of
951 variation in mRNA half-lives explained by individual sequence features. (H) Multiple linear
952 regression was used to identify the percent of variation in mRNA half-lives explained by
953 combinations of different sequence features. (I) MEME (Bailey et al. 2015) identified three

motifs enriched in the 3' UTRs of the top 15% stable transcripts compared to the 3' UTRs of the top 15% unstable transcripts. For genes with multiple 3' UTR isoforms, the longest isoform was used. The best match of each *de novo* motif to known motifs in mammals (Ray et al. 2013) is noted.

Fig 3. High transcript accumulation is associated with increased mRNA stability. (A) A schematic representation of genes that accumulate high, medium, or low transcript levels and the expected range of transcription and mRNA decay rates that could contribute to such accumulation. (B) Examples of genes that peak in expression ~200 minutes after the four-cell stage with high, medium, and low transcript accumulation, from left to right. Expression data taken from a whole-embryo RNA sequencing time series (Hashimshony et al. 2015). (C) Examples of genes that peak in expression ~350 minutes after the four-cell stage with high, medium, and low transcript accumulation, from left to right. Expression data taken from a whole-embryo RNA sequencing time series (Hashimshony et al. 2015). (D) Box plots showing the mRNA half-life distributions of genes that peak in expression ~200 minutes after the four-cell stage to low, medium, and high transcript levels. Numbers to the left of the box plots are median half-lives within each group. Numbers above the box plots are the number of genes within each group. P-values comparing median half-lives were calculated using the Wilcoxon rank sum test. Outliers not shown: from left to right, 1, 0, and 7 genes within each group had mRNA half-lives > 125 minutes. (E) Box plots showing the mRNA half-life distributions of genes that peak in expression ~350 minutes after the four-cell stage with low, medium, and high transcript levels. Numbers to the left of the box plots are median half-lives within each group. Numbers above the box plots are the number of genes within each group. P-values comparing median half-lives were calculated using the Wilcoxon rank sum test. Outliers not shown: from left to right, 0, 0, and 2 genes within each group had mRNA half-lives > 175 minutes. (F) MEME (Bailey et al. 2015) identified three motifs enriched in the 3' UTRs of genes with high transcript accumulation

compared to the 3' UTRs of genes with low transcript accumulation ~200 minutes after the four-cell stage. For genes with multiple 3' UTR isoforms, the longest isoform was used. The best match of each *de novo* motif to known motifs in mammals (Ray et al. 2013) is noted. (G) MEME (Bailey et al. 2015) identified three motifs enriched in the 3' UTRs of genes with high transcript accumulation compared to the 3' UTRs of genes with low transcript accumulation ~350 minutes after the four-cell stage. For genes with multiple 3' UTR isoforms, the longest isoform was used. The best match of each *de novo* motif to known motifs in mammals (Ray et al. 2013) is noted. (H) The gene expression patterns for three putative RNA-binding protein genes in *C. elegans* from a whole-embryo RNA sequencing time series (Hashimshony et al. 2015). The genes are homologs of mammalian RNA-binding protein genes whose corresponding proteins are known to bind motifs similar to those discovered in (F) and (G).

Fig 4. Single-cell RNA sequencing allows measurement of mRNA half-lives at high resolution throughout *C. elegans* embryogenesis. (A) A schematic representation of the approach to measure mRNA half-lives in the *C. elegans* embryo using transcription inhibition and single-cell RNA sequencing. (B) UMAP projection of the integrated dataset of three biological replicates. Cells are colored by the age of the embryo from which a cell was produced, estimated from correlations to a whole-embryo RNA sequencing time series (Hashimshony et al. 2015). Trajectories corresponding to major cell types are labeled. (C) Scatter plot comparing the mRNA half-lives calculated in pseudobulk from the single-cell data and the half-lives calculated from the bulk-cell data on a log-log scale. Spearman's correlation coefficient = 0.76. Dashed line is the x = y line. (D) Box plots showing the stage-specific mRNA half-life distributions of genes within Early-, Middle-, and Late-stage cells (50-200, 200-350, and 350+ minutes after the four-cell embryo stage, respectively). (E) Box plots showing the cell type-specific mRNA half-life distributions of genes within muscle, germline, epidermis, neuron, and pharynx. Numbers to the left of the box plots are median half-lives within each group. Numbers

above the box plots are the number of genes within each group. P-values comparing median half-lives were calculated using the Wilcoxon rank sum test. Outliers not shown: from left to right for (D), 203, 345, and 323 genes within each group had mRNA half-lives > 100 minutes; from left to right for (E), 141, 36, 150, 70, and 125 genes within each group had mRNA half-lives > 150 minutes.

Fig 5. Differential mRNA decay occurs throughout different developmental stages of *C. elegans* embryogenesis. (A) Scatter plots comparing mRNA half-lives specific to Early-, Middle-, and Late-stage cells in all pairwise comparisons. Each point represents a gene. Blue points correspond to the top 5% of genes with faster mRNA decay in the later stage compared to in the earlier stage. Pink points correspond to the top 5% of genes with slower mRNA decay in the later stage compared to in the earlier stage. Dashed line is the x = y line. (B) Select Gene Ontology categories enriched among the top 5% of genes with faster mRNA decay in the later stage of embryogenesis compared to the earlier stage of embryogenesis. Background set of genes used in each comparison was shared genes we were able to calculate stage-specific mRNA half-lives for between the relevant stages. (C, D, E, F) *Left*. Median scaled expression of gene subset using data from a whole-embryo RNA sequencing time series (Hashimshony et al. 2015). Blue shading spans the Early stage, purple shading spans the Middle stage, and pink shading spans the late stage. *Right*. Plot displaying the change in mRNA half-lives from earlier to later stage for gene subset.

Fig 6. mRNA degradation may contribute to transcription factor dynamics at both the RNA and protein levels. (A) Box plots showing the mRNA half-life distributions of transcription factor genes and all other genes in Early-, Middle-, and Late-stage cells. Numbers to the left of the box plots are median half-lives within each group. Numbers above the box plots are the number of genes within each group. P-values comparing median half-lives were calculated

using the Wilcoxon rank sum test. Outliers not shown: from left to right, 198, 5, 336, 9, 314, and 9 genes within each group had mRNA half-lives > 100 minutes. (B) Grid displaying the percentage of transcription factors with transient or persistent mRNA expression and transient or persistent protein expression. The probability that RNA and protein dynamics are independent of one another was calculated using Fisher's exact test. (C) Box plots showing the single-cell RNA sequencing pseudobulk mRNA half-life distributions of transcription factors with transient mRNA and protein expression, transient mRNA and persistent protein expression, persistent mRNA and transient protein expression, and persistent mRNA and protein expression. Numbers to the left of the box plots are median half-lives within each group. Numbers above the box plots are the number of genes within each group. P-values comparing median half-lives were calculated using the Wilcoxon rank sum test. (D, E) Sublineages with coloring representing reporter GFP expression from a single-cell transcription factor protein expression atlas of the *C. elegans* embryo (Ma et al. 2021). Protein and mRNA dynamics are characterized below, along with pseudobulk measured mRNA half-life.

Fig 7. mRNA stability is correlated with cell type-specific functions. (A) Box plots showing the mRNA half-life distributions of cell type-specific and broadly expressed genes within muscle, germline, epidermis, neuron, and pharynx cells. Numbers to the left of the box plots are median half-lives within each group. Numbers above box plots are the number of genes within each group. P-value comparing median half-lives was calculated using the Wilcoxon rank sum test. Outliers not shown: from left to right, 16, 125, 8, 28, 1, 149, 1, 69, 5, and 120 genes within each group had mRNA half-lives > 150 minutes. (B) (i) Box plots showing the muscle-specific mRNA half-life distributions of mesoderm/muscle fate-specifying transcription factor genes, muscle structure genes, and all other genes. Numbers to the left of the box plots are median half-lives within each group. Numbers above box plots are the number of genes within each group. P-value comparing median half-lives was calculated using the Wilcoxon rank sum test. Outliers

not shown: from left to right, 0, 4, and 137 genes within each group had mRNA half-lives > 150 minutes. (ii, iii) Scatter plots of the normalized transcript abundance of the mesoderm/muscle fate-specifying transcription factor genes *ceh-34*, *tbx-2*, *ceh-20*, *eya-1*, *hlh-1* and the muscle structure genes *deb-1*, *unc-78*, *unc-112*, *sgn-1*, *tmd-2* throughout a 40 minute transcription inhibition time course in muscle cells. Each point represents normalized transcript abundance from one of three biological replicates. (C) (i) Box plots showing the neuron-specific mRNA half-life distributions of neural fate-specifying transcription factor genes, synapse-/dendrite-/axon-associated genes, and all other genes. Numbers to the left of the box plots are median half-lives within each group. Numbers above box plots are the number of genes within each group. P-value comparing median half-lives was calculated using the Wilcoxon rank sum test. Outliers not shown: from left to right, 0, 0, and 70 genes within each group had mRNA half-lives > 150 minutes. (ii, iii) Scatter plots of the normalized transcript abundance of the neural fate-specifying transcription factor genes *ceh-5*, *egl-46*, *unc-86*, *ceh-43*, *hlh-2* and the synapse-/dendrite-/axon-associated genes *lnp-1*, *unc-37*, *mig-2*, *ced-10*, *unc-53* throughout a 40 minute transcription inhibition time course in neuronal cells. Each point represents normalized transcript abundance from one of three biological replicates. (D) Cell type-specific expression in Middle and Late stages for genes with differential mRNA decay between (i) muscle and neuron, (ii) epidermis and neuron, and (iii) muscle and epidermis. Cell type-specific expression in TPM was obtained from a *C. elegans* embryo single-cell atlas (Packer et al. 2019).

Fig 8. Differential mRNA decay occurs across germline and somatic cells in the *C. elegans* embryo. (A) Box plots showing the cell type-specific mRNA half-life distributions of genes across the germline and soma. (B) Scatter plot comparing soma- and germline-specific mRNA half-lives to one another. Each point represents a gene. Pink points correspond to the top 10% of genes with longer mRNA decay in the germline compared to in the soma. Blue points correspond to the top 10% of genes with faster mRNA decay in the germline compared to

in the soma. Dashed line is the x = y line. (C) A cartoon of box plots representing the expected mRNA half-life distribution of genes with less, constant, or greater expression over time if mRNA decay primarily contributes to gene expression. (D) A cartoon of box plots representing the expected mRNA half-life distribution of genes with less, constant, or greater expression over time if transcription primarily contributes to gene expression. (E) Box plots showing the germline-specific mRNA half-life distributions for genes that decrease or increase in expression over time in the germline from a fold-change of 1-2 or greater than 2. (F) Box plots showing the soma-specific mRNA half-life distributions for genes that decrease or increase in expression over time in the soma from a fold-change of 1-2 or greater than 2. (G) Box plots showing the germline-specific mRNA half-life distributions of mRNA 3' UTR-binding genes, other mRNA-binding genes, and genes not annotated as RNA-binding. (H) Box plots showing the soma-specific mRNA half-life distributions of mRNA 3' UTR-binding genes, other mRNA-binding genes, and genes not annotated as RNA-binding. (I) Plot displaying the soma- and germline-line specific mRNA half-lives for mRNA-binding genes whose decay is more rapid in the germline than in the soma. Numbers to the left of the box plots are median half-lives within each group. Numbers above the box plots are the number of genes with half-lives greater than 150 minutes (A, G, H) or 325 minutes (E, F) within each group. P-value comparing median half-lives was calculated using the Wilcoxon rank sum test.

References

Abbadi D, Yang M, Chenette DM, Andrews JJ, Schneider RJ. 2019. Muscle development and regeneration controlled by AUF1-mediated stage-specific degradation of fate-determining checkpoint mRNAs. Proc Natl Acad Sci. 116(23):11285–11290. doi:10.1073/pnas.1901165116.

Alonso CR. 2012. A complex 'mRNA degradation code' controls gene expression during animal development. Trends Genet. 28(2):78–88. doi:10.1016/j.tig.2011.10.005.

Angeles-Albores D, Lee RYN, Chan J, Sternberg PW. 2018 Mar 2. Two new functions in the WormBase Enrichment Suite. MicroPublication Biol. doi:10.17912/W25Q2N. [accessed 2023 Apr 13]. https://www.micropublication.org/journals/biology/w25q2n.

Bae H, Coller J. 2022. Codon optimality-mediated mRNA degradation: Linking translational elongation to mRNA stability. Mol Cell. 82(8):1467–1476. doi:10.1016/j.molcel.2022.03.032.

Bailey TL, Johnson J, Grant CE, Noble WS. 2015. The MEME Suite. Nucleic Acids Res. 43(Web Server issue):W39–W49. doi:10.1093/nar/gkv416.

Baker SC, Bauer SR, Beyer RP, Brenton JD, Bromley B, Burrill J, Causton H, Conley MP, Elespuru R, Fero M, et al. 2005. The External RNA Controls Consortium: a progress report. Nat Methods. 2(10):731–734. doi:10.1038/nmeth1005-731.

Bao Z, Zhao Z, Boyle TJ, Murray JI, Waterston RH. 2008. Control of Cell Cycle Timing during C. elegans Embryogenesis. Dev Biol. 318(1):65–72. doi:10.1016/j.ydbio.2008.02.054.

Baty F, Ritz C, Charles S, Brutsche M, Flandrois J, Delignette-Muller M. 2015. A Toolbox for Nonlinear Regression in R: The Package nlstools. Journal of Statistical Software. 66(5), 1–21. doi:10.18637/jss.v066.i05.

Beadle LF, Love JC, Shapovalova Y, Artemev A, Rattray M, Ashe HL. 2023. Combined modelling of mRNA decay dynamics and single-molecule imaging in the Drosophila embryo uncovers a role for P-bodies in 5′ to 3′ degradation. PLOS Biol. 21(1):e3001956. doi:10.1371/journal.pbio.3001956.

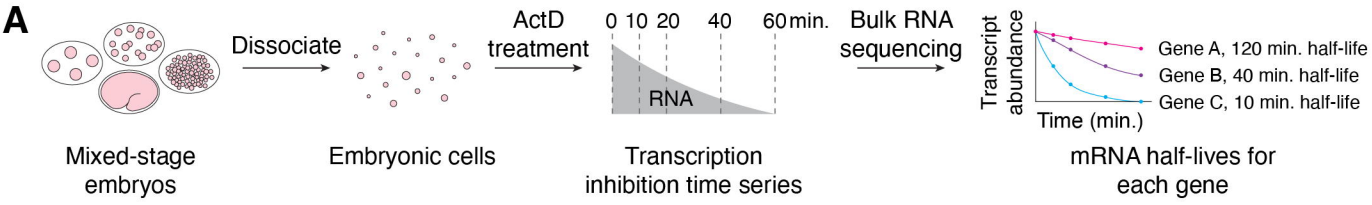
Becht E, McInnes L, Healy J, Dutertre C-A, Kwok IWH, Ng LG, Ginhoux F, Newell EW. 2019.

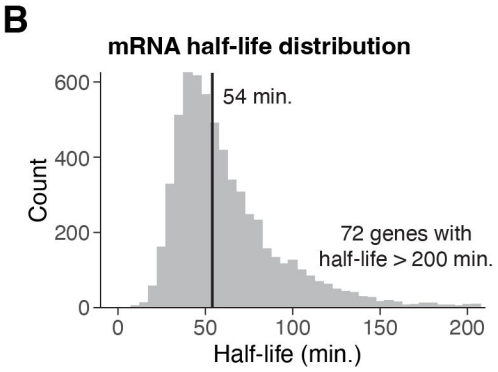
Dimensionality reduction for visualizing single-cell data using UMAP. Nat Biotechnol. 37(1):38–44. doi:10.1038/nbt.4314.
Bhattacharyya SN, Habermacher R, Martine U, Closs EI, Filipowicz W. 2006. Relief of microRNA-mediated translational repression in human cells subjected to stress. Cell. 125(6):1111–1124. doi:10.1016/j.cell.2006.04.031.
Bianchi L, Driscoll M. 2006. Culture of embryonic C. elegans cells for electrophysiological and pharmacological analyses. WormBook. [accessed 2023 Apr 12]. https://www.ncbi.nlm.nih.gov/books/NBK19713/.
Blacque OE, Li C, Inglis PN, Esmail MA, Ou G, Mah AK, Baillie DL, Scholey JM, Leroux MR. 2006. The WD Repeat-containing Protein IFTA-1 Is Required for Retrograde Intraflagellar Transport. Mol Biol Cell. 17(12):5053–5062. doi:10.1091/mbc.E06-06-0571.
Blazie SM, Babb C, Wilky H, Rawls A, Park JG, Mangone M. 2015. Comparative RNA-Seq analysis reveals pervasive tissue-specific alternative polyadenylation in Caenorhabditis elegans intestine and muscles. BMC Biol. 13(1):4. doi:10.1186/s12915-015-0116-6.
Blazie SM, Geissel HC, Wilky H, Joshi R, Newbern J, Mangone M. 2017. Alternative Polyadenylation Directs Tissue-Specific miRNA Targeting in *Caenorhabditis elegans* Somatic Tissues. Genetics. 206(2):757–774. doi:10.1534/genetics.116.196774.
Brocal-Ruiz R, Esteve-Serrano A, Mora-Martínez C, Franco-Rivadeneira ML, Swoboda P, Tena JJ, Vilar M, Flames N. 2023. Forkhead transcription factor FKH-8 cooperates with RFX in the direct regulation of sensory cilia in Caenorhabditis elegans. Portman D, Sengupta P, editors. eLife. 12:e89702. doi:10.7554/eLife.89702.
Burow DA, Umeh-Garcia MC, True MB, Bakhaj CD, Ardell DH, Cleary MD. 2015. Dynamic regulation of mRNA decay during neural development. Neural Develop. 10(1):11. doi:10.1186/s13064-015-0038-6.
Chekulaeva M, Filipowicz W. 2009. Mechanisms of miRNA-mediated post-transcriptional regulation in animal cells. Curr Opin Cell Biol. 21(3):452–460. doi:10.1016/j.ceb.2009.04.009.
Christensen M, Estevez A, Yin X, Fox R, Morrison R, McDonnell M, Gleason C, Miller DM, Strange K. 2002. A primary culture system for functional analysis of C. elegans neurons and muscle cells. Neuron. 33(4):503–514. doi:10.1016/s0896-6273(02)00591-3.
Crane MM, Sands B, Battaglia C, Johnson B, Yun S, Kaeberlein M, Brent R, Mendenhall A. 2019. In vivo measurements reveal a single 5′-intron is sufficient to increase protein expression level in Caenorhabditis elegans. Sci Rep. 9(1):9192. doi:10.1038/s41598-019-45517-0.
Dexheimer PJ, Wang J, Cochella L. 2020. Two MicroRNAs Are Sufficient for Embryonic Patterning in *C. elegans*. Curr Biol. 30(24):5058-5065.e5. doi:10.1016/j.cub.2020.09.066.
Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, Batut P, Chaisson M, Gingeras TR. 2013. STAR: ultrafast universal RNA-seq aligner. Bioinformatics. 29(1):15–21. doi:10.1093/bioinformatics/bts635.
Edgar LG. 1995. Blastomere culture and analysis. Methods Cell Biol. 48:303–321. doi:10.1016/s0091-679x(08)61393-x.
Fuxman Bass JI, Pons C, Kozlowski L, Reece-Hoyes JS, Shrestha S, Holdorf AD, Mori A, Myers CL, Walhout AJ. 2016. A gene-centered C. elegans protein–DNA interaction network provides a framework for functional predictions. Mol Syst Biol. 12(10):884. doi:10.15252/msb.20167131.
Geuens T, Bouhy D, Timmerman V. 2016. The hnRNP family: insights into their role in health and disease. Hum Genet. 135(8):851–867. doi:10.1007/s00439-016-1683-5.
van der Giessen K, Gallouzi I-E. 2007. Involvement of Transportin 2–mediated HuR Import in Muscle Cell Differentiation. Mol Biol Cell. 18(7):2619–2629. doi:10.1091/mbc.E07-02-0167.
Giraldez AJ, Mishima Y, Rihel J, Grocock RJ, Van Dongen S, Inoue K, Enright AJ, Schier AF. 2006. Zebrafish MiR-430 promotes deadenylation and clearance of maternal mRNAs. Science. 312(5770):75–79. doi:10.1126/science.1122689.
Gupta S, Stamatoyannopoulos JA, Bailey TL, Noble WS. 2007. Quantifying similarity between

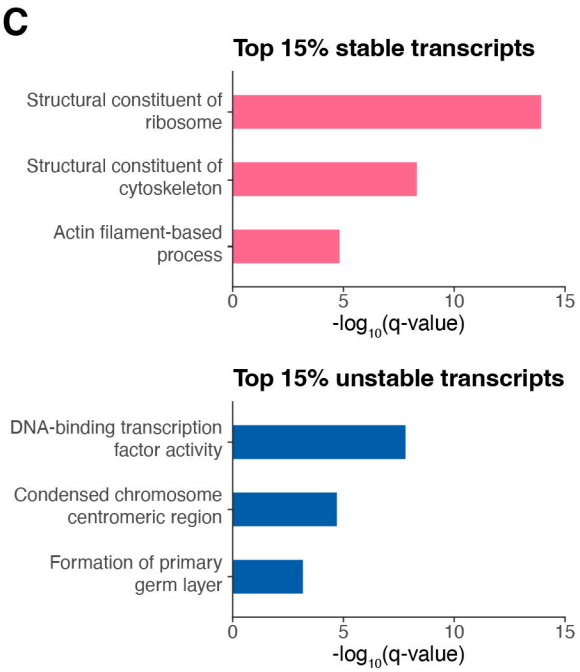
motifs. Genome Biol. 8(2):R24. doi:10.1186/gb-2007-8-2-r24.
Gutiérrez-Pérez P, Santillán EM, Lendl T, Wang J, Schrempf A, Steinacker TL, Asparuhova M, Brandstetter M, Haselbach D, Cochella L. 2021. miR-1 sustains muscle physiology by controlling V-ATPase complex assembly. Sci Adv. 7(42):eabh1434. doi:10.1126/sciadv.abh1434.
Hashimshony T, Feder M, Levin M, Hall BK, Yanai I. 2015. Spatiotemporal transcriptomics reveals the evolutionary history of the endoderm germ layer. Nature. 519(7542):219–222. doi:10.1038/nature13996.
Herzog VA, Reichholf B, Neumann T, Rescheneder P, Bhat P, Burkard TR, Wlotzka W, von Haeseler A, Zuber J, Ameres SL. 2017. Thiol-linked alkylation of RNA to assess expression dynamics. Nat Methods. 14(12):1198–1204. doi:10.1038/nmeth.4435.
Howe KL, Bolt BJ, Shafie M, Kersey P, Berriman M. 2017. WormBase ParaSite − a comprehensive resource for helminth genomics. Mol Biochem Parasitol. 215:2–10. doi:10.1016/j.molbiopara.2016.11.005.
Jiang P, Singh M, Coller HA. 2013. Computational assessment of the cooperativity between RNA binding proteins and MicroRNAs in Transcript Decay. PLoS Comput Biol. 9(5):e1003075. doi:10.1371/journal.pcbi.1003075.
Łabno A, Tomecki R, Dziembowski A. 2016. Cytoplasmic RNA decay pathways - Enzymes and mechanisms. Biochim Biophys Acta BBA - Mol Cell Res. 1863(12):3125–3147. doi:10.1016/j.bbamcr.2016.09.023.
Lee RYN, Howe KL, Harris TW, Arnaboldi V, Cain S, Chan J, Chen WJ, Davis P, Gao S, Grove C, et al. 2018. WormBase 2017: molting into a new stage. Nucleic Acids Res. 46(Database issue):D869–D874. doi:10.1093/nar/gkx998.
Levin M, Zalts H, Mostov N, Hashimshony T, Yanai I. 2020. Gene expression dynamics are a proxy for selective pressures on alternatively polyadenylated isoforms. Nucleic Acids Res. 48(11):5926–5938. doi:10.1093/nar/gkaa359.
Li C, Jensen VL, Park K, Kennedy J, Garcia-Gonzalo FR, Romani M, Mori RD, Bruel A-L, Gaillard D, Doray B, et al. 2016. MKS5 and CEP290 Dependent Assembly Pathway of the Ciliary Transition Zone. PLOS Biol. 14(3):e1002416. doi:10.1371/journal.pbio.1002416.
Ma X, Zhao Z, Xiao L, Xu W, Kou Y, Zhang Y, Wu G, Wang Y, Du Z. 2021. A 4D single-cell protein atlas of transcription factors delineates spatiotemporal patterning during embryogenesis. Nat Methods. 18(8):893–902. doi:10.1038/s41592-021-01216-1.
Makeyev AV, Liebhaber SA. 2002. The poly(C)-binding proteins: a multiplicity of functions and a search for mechanisms. RNA. 8(3):265–278.
Mayr C. 2019. What Are 3′ UTRs Doing? Cold Spring Harb Perspect Biol. 11(10):a034728. doi:10.1101/cshperspect.a034728.
McInnes L, Healy J, Melville J. 2020. UMAP: Uniform Manifold Approximation and Projection for Dimension Reduction. doi:10.48550/arXiv.1802.03426. [accessed 2023 Oct 3]. http://arxiv.org/abs/1802.03426.
Miller TLboFcbA. 2024. leaps: Regression Subset Selection. R package version 3.2, https://CRAN.R-project.org/package=leaps.
Mishima Y, Tomari Y. 2016. Codon Usage and 3′ UTR Length Determine Maternal mRNA Stability in Zebrafish. Mol Cell. 61(6):874–885. doi:10.1016/j.molcel.2016.02.027.
Nair G, Walton T, Murray JI, Raj A. 2013. Gene transcription is coordinated with, but not dependent on, cell divisions during C. elegans embryonic fate specification. Development. 140(16):3385–3394. doi:10.1242/dev.098012.
Narsai R, Howell KA, Millar AH, O'Toole N, Small I, Whelan J. 2007. Genome-Wide Analysis of mRNA Decay Rates and Their Determinants in *Arabidopsis thaliana*. Plant Cell. 19(11):3418–3436. doi:10.1105/tpc.107.055046.
Nechipurenko IV, Sengupta P. 2017. The rise and fall of basal bodies in the nematode Caenorhabditis elegans. Cilia. 6(1):9. doi:10.1186/s13630-017-0053-9.

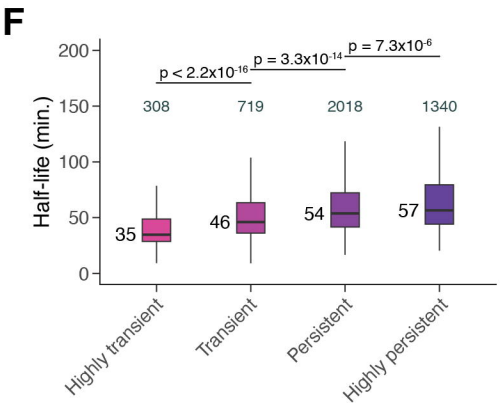
Ou G, Koga M, Blacque OE, Murayama T, Ohshima Y, Schafer JC, Li C, Yoder BK, Leroux MR, Scholey JM. 2007. Sensory Ciliogenesis in Caenorhabditis elegans: Assignment of IFT Components into Distinct Modules Based on Transport and Phenotypic Profiles. Mol Biol Cell. 18(5):1554–1569. doi:10.1091/mbc.E06-09-0805.
Packer JS, Zhu Q, Huynh C, Sivaramakrishnan P, Preston E, Dueck H, Stefanik D, Tan K, Trapnell C, Kim J, et al. 2019. A lineage-resolved molecular atlas of C. elegans embryogenesis at single-cell resolution. Science. 365(6459):eaax1971. doi:10.1126/science.aax1971.
Pamarthy S, Kulshrestha A, Katara GK, Beaman KD. 2018. The curious case of vacuolar ATPase: regulation of signaling pathways. Mol Cancer. 17(1):41. doi:10.1186/s12943-018-0811-3.
Passmore LA, Coller J. 2021 Sep 30. Roles of mRNA poly(A) tails in regulation of eukaryotic gene expression. Nat Rev Mol Cell Biol.:1–14. doi:10.1038/s41580-021-00417-y.
Presnyak V, Alhusaini N, Chen Y-H, Martin S, Morris N, Kline N, Olson S, Weinberg D, Baker KE, Graveley BR, et al. 2015. Codon Optimality Is a Major Determinant of mRNA Stability. Cell. 160(6):1111–1124. doi:10.1016/j.cell.2015.02.029.
Raj A, Rifkin SA, Andersen E, van Oudenaarden A. 2010. Variability in gene expression underlies incomplete penetrance. Nature. 463(7283):913–918. doi:10.1038/nature08781.
Ray D, Kazan H, Cook KB, Weirauch MT, Najafabadi HS, Li X, Gueroussov S, Albu M, Zheng H, Yang A, et al. 2013. A compendium of RNA-binding motifs for decoding gene regulation. Nature. 499(7457):172–177. doi:10.1038/nature12311.
R Core Team. 2021. R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna. https://www.R-project .org/.
Sabi R, Volvovitch Daniel R, Tuller T. 2017. stAIcalc: tRNA adaptation index calculator based on species-specific weights. Bioinformatics. 33(4):589–591. doi:10.1093/bioinformatics/btw647.
Schubert CM, Lin R, Vries CJ de, Plasterk RHA, Priess JR. 2000. MEX-5 and MEX-6 Function to Establish Soma/Germline Asymmetry in Early C. elegans Embryos. Mol Cell. 5(4):671–682. doi:10.1016/S1097-2765(00)80246-4.
Shaul O. 2017. How introns enhance gene expression. Int J Biochem Cell Biol. 91:145–155. doi:10.1016/j.biocel.2017.06.016.
Sivaramakrishnan P, Watkins C, Murray JI. 2023. Transcript accumulation rates in the early Caenorhabditis elegans embryo. Sci Adv. 9(34):eadi1270. doi:10.1126/sciadv.adi1270.
Soustelle L, Roy N, Ragone G, Giangrande A. 2008. Control of gcm RNA stability is necessary for proper glial cell fate acquisition. Mol Cell Neurosci. 37(4):657–662. doi:10.1016/j.mcn.2007.11.007.
Stoeckius M, Grün D, Kirchner M, Ayoub S, Torti F, Piano F, Herzog M, Selbach M, Rajewsky N. 2014. Global characterization of the oocyte-to-embryo transition in Caenorhabditis elegans uncovers a novel mRNA clearance mechanism. EMBO J. 33(16):1751–1766. doi:10.15252/embj.201488769.
Stuart T, Butler A, Hoffman P, Hafemeister C, Papalexi E, III WMM, Hao Y, Stoeckius M, Smibert P, Satija R. 2019. Comprehensive Integration of Single-Cell Data. Cell. 177, 1888-1902. doi:10.1016/j.cell.2019.05.031, https://doi.org/10.1016/j.cell.2019.05.031.
Sulston JE, Schierenberg E, White JG, Thomson JN. 1983. The embryonic cell lineage of the nematode *Caenorhabditis elegans*. Dev Biol. 100(1):64–119. doi:10.1016/0012-1606(83)90201-4.
Tadros W, Goldman AL, Babak T, Menzies F, Vardy L, Orr-Weaver T, Hughes TR, Westwood JT, Smibert CA, Lipshitz HD. 2007. SMAUG Is a Major Regulator of Maternal mRNA Destabilization in Drosophila and Its Translation Is Activated by the PAN GU Kinase. Dev Cell. 12(1):143–155. doi:10.1016/j.devcel.2006.10.005.
Team BC, Maintainer BP. 2019. TxDb.Hsapiens.UCSC.hg38.knownGene: Annotation package for TxDb object(s). R package version 3.4.6.
Thomsen S, Anders S, Janga SC, Huber W, Alonso CR. 2010. Genome-wide analysis of mRNA

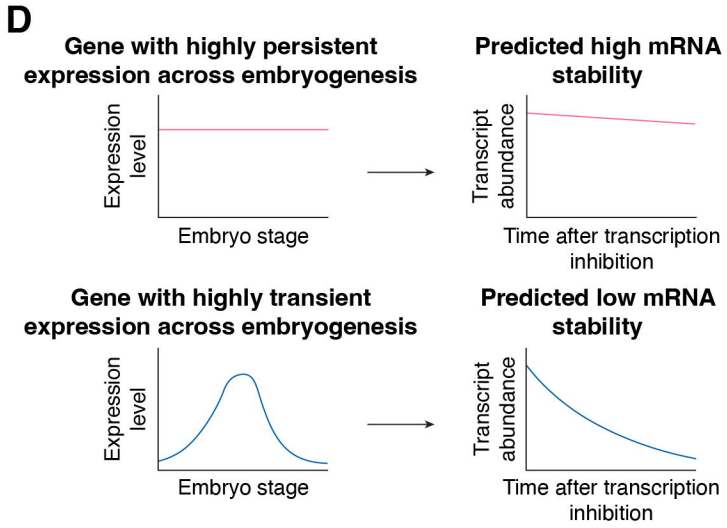
decay patterns during early Drosophiladevelopment. Genome Biol. 11(9):R93. doi:10.1186/gb-2010-11-9-r93.
Tintori SC, Osborne Nishimura E, Golden P, Lieb JD, Goldstein B. 2016. A Transcriptional Lineage of the Early C. elegans Embryo. Dev Cell. 38(4):430–444. doi:10.1016/j.devcel.2016.07.025.
Vastenhouw NL, Cao WX, Lipshitz HD. 2019. The maternal-to-zygotic transition revisited. Development. 146(11):dev161471. doi:10.1242/dev.161471.
Vejnar CE, Abdel Messih M, Takacs CM, Yartseva V, Oikonomou P, Christiano R, Stoeckius M, Lau S, Lee MT, Beaudoin J-D, et al. 2019. Genome wide analysis of 3' UTR sequence elements and proteins regulating mRNA stability during maternal-to-zygotic transition in zebrafish. Genome Res. 29(7):1100–1114. doi:10.1101/gr.245159.118.
Wang JT, Seydoux G. 2013. Germ Cell Specification. Adv Exp Med Biol. 757:17–39. doi:10.1007/978-1-4614-4015-4_2.
Wang W, Jack BM, Wang HH, Kavanaugh MA, Maser RL, Tran PV. 2021. Intraflagellar Transport Proteins as Regulators of Primary Cilia Length. Front Cell Dev Biol. 9. [accessed 2023 Oct 10]. https://www.frontiersin.org/articles/10.3389/fcell.2021.661350.
Williams CL, Li C, Kida K, Inglis PN, Mohan S, Semenec L, Bialas NJ, Stupay RM, Chen N, Blacque OE, et al. 2011. MKS and NPHP modules cooperate to establish basal body/transition zone membrane associations and ciliary gate function during ciliogenesis. J Cell Biol. 192(6):1023–1041. doi:10.1083/jcb.201012116.
Wu Q, Medina SG, Kushawah G, DeVore ML, Castellano LA, Hand JM, Wright M, Bazzini AA. 2019. Translation affects mRNA stability in a codon-dependent manner in human cells. Sonenberg N, Struhl K, Weissman JS, editors. eLife. 8:e45396. doi:10.7554/eLife.45396.
Yang E, van Nimwegen E, Zavolan M, Rajewsky N, Schroeder M, Magnasco M, Darnell JE. 2003. Decay Rates of Human mRNAs: Correlation With Functional Characteristics and Sequence Attributes. Genome Res. 13(8):1863–1872. doi:10.1101/gr.1272403.
Yoon J, Chung Y-J, Lee M. 2018. STADIUM: Species-Specific tRNA Adaptive Index Compendium. Genomics Inform. 16(4):e28. doi:10.5808/GI.2018.16.4.e28.
Young LE, Moore AE, Sokol L, Meisner-Kober N, Dixon DA. 2012. The mRNA stability factor HuR inhibits microRNA-16 targeting of COX-2. Mol Cancer Res MCR. 10(1):167–180. doi:10.1158/1541-7786.MCR-11-0337.
Young MD, Behjati S. 2020. SoupX removes ambient RNA contamination from droplet-based single-cell RNA sequencing data. GigaScience. 9(12):giaa151. doi:10.1093/gigascience/giaa151.
Zhu Q, Fisher SA, Shallcross J, Kim J. 2016. VERSE: a versatile and efficient RNA-Seq read counting tool. :053306. doi:10.1101/053306. [accessed 2023 Oct 6]. https://www.biorxiv.org/content/10.1101/053306v1.

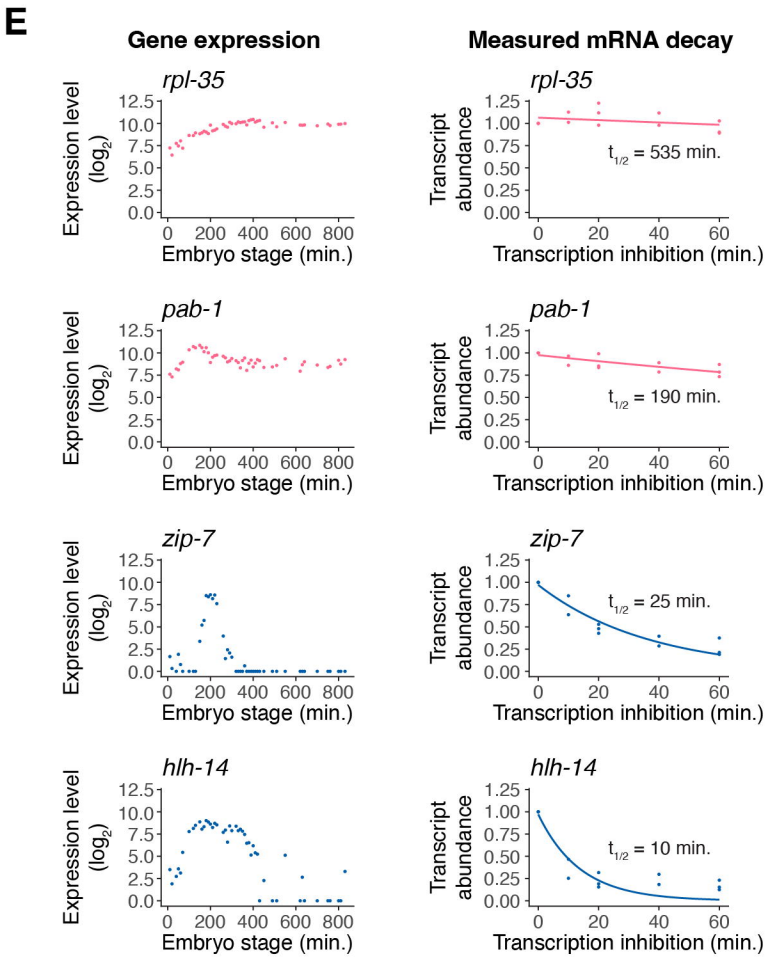


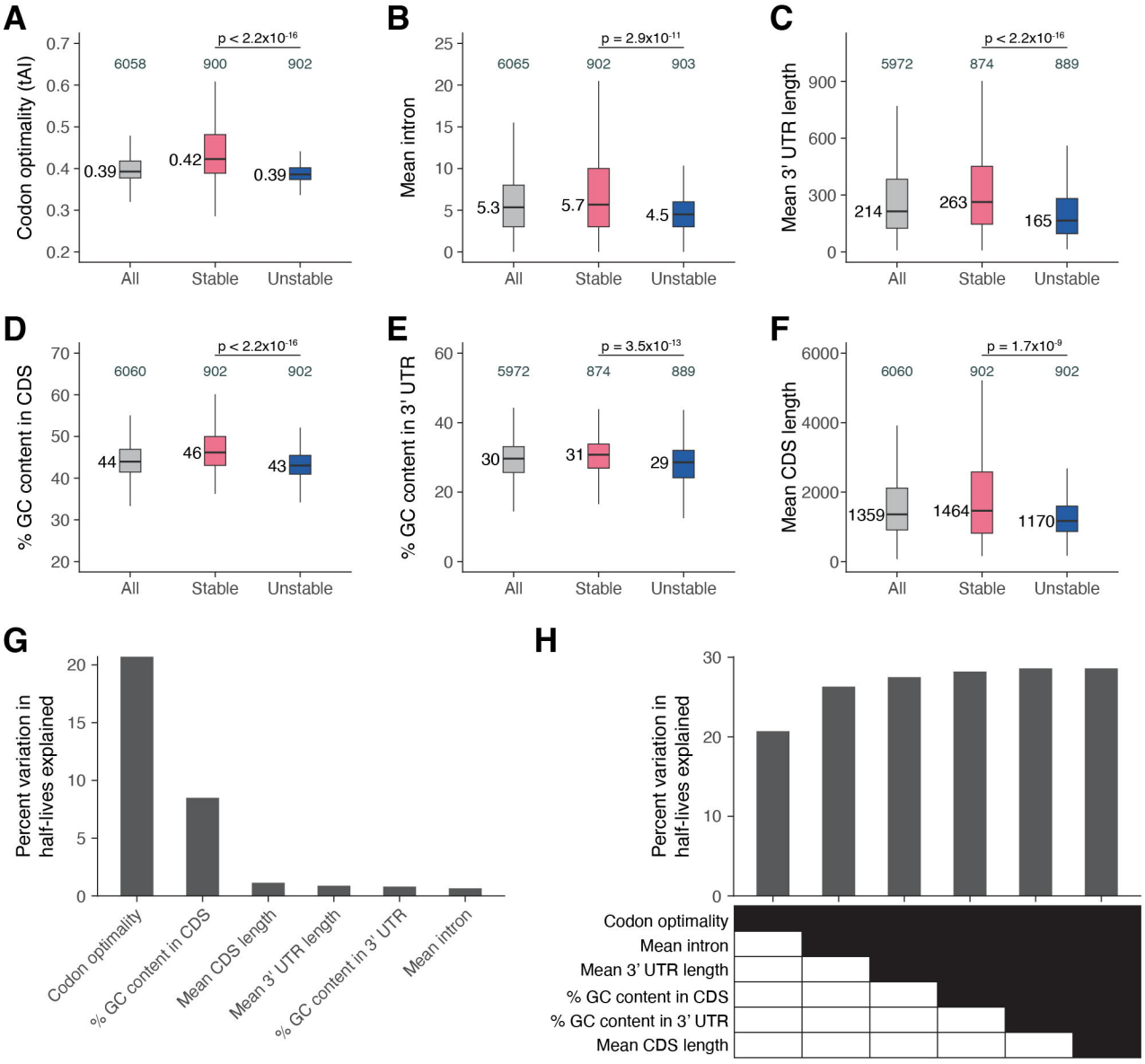






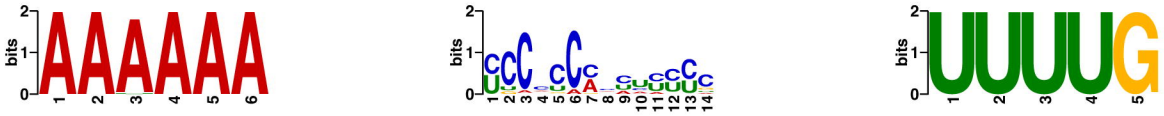






I

Enriched motifs in 3' UTR of stable transcripts



E-value: 2.8×10^{-15}
PABPC4 motif, E-value: 3.4×10^{-2}

E-value: 4.1×10^{-11}
PCBP2 motif, E-value: 3.7×10^{-3}

E-value: 1.5×10^{-7}
RALY motif, E-value: 0.33