Expanded View Figures

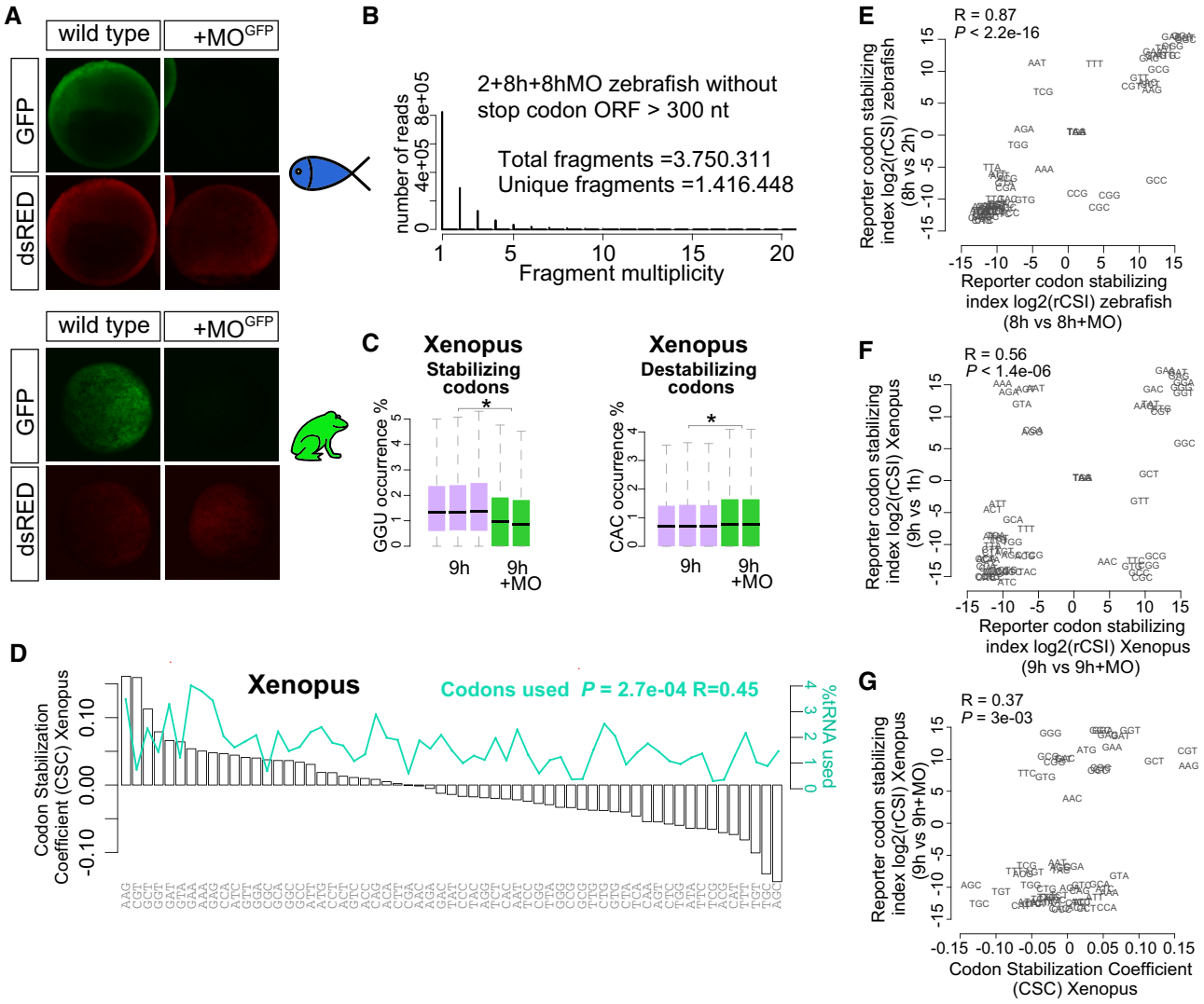


Figure EV1. Translation modulates stability of mRNA in zebrafish and *Xenopus*.

A GFP reporter expression (green) and control dsRed expression (red) at 8 hpf and 9.5 hpf in zebrafish and *Xenopus*, respectively. The GFP reporter is repressed by co-injection with the morpholino which recognizes the GFP translation initiation sequence (MO). dsRed is not recognized by the MO. This morpholino is complementary to the translation initiation site in the reporter mRNA library used in Fig 1A.

B Bar plot shows distribution of fragment multiplicity for fragments longer that 300 nt without stop codons at 2, 8 h, and 8 h + MO in zebrafish embryos, showing a high degree of complexity of the reporter library mRNA.

C Box plots show the occurrence of the optimal codon GGU and non-optimal codon CAC in injected mRNA at 1, 9 h, and 9 h + MO in *Xenopus* embryos (**P* < 1e-300 and **P* < 1e-300, respectively, one-way ANOVA). The box defines the first and third quartiles, with the median indicated with a thick black line and vertical lines indicate the variability outside the upper and lower quartiles. The analysis of variance (ANOVA) for the occurrence of each codon in the presence (−MO) or the absence (+MO) of translation is used to calculate the reporter codon stabilizing index (rCSI). If codons cause differential mRNA stability in a translation-dependent manner, potential stabilizing codons should be enriched in translated mRNAs (−MO versus +MO) (+rCSI), and destabilizing codons depleted (−MO versus +MO) (−rCSI).

D Bar plots ranking the 61 coding codons from highest to lowest correlation with mRNA half-life (CSC). The relative abundance of each codon present in the expressed transcriptome at 9 hpf in *Xenopus* is shown (green).

E, F Biplots of the rCSI derived by comparing codon enrichment or depletion in two different conditions: *x*-axis, in the presence and the absence of translation (±MO); *y*-axis, in a late versus earlier time points (E) in zebrafish or (F) in *Xenopus*. Both comparisons yield similar results, ruling out nonspecific effects of the MO. One hour or 2 h is too early to observe the full effects of translational regulation in mRNA stability.

G Biplots of the rCSI (9 h/9 h + MO) from *Xenopus* embryos and the codon stabilization coefficient (CSC) in *Xenopus* for each codon, *P* = 3e-3, *R* = 0.37, Spearman rank correlation.

Figure EV2. Translation of specific codons modulates mRNA stability in vertebrates.

A Cumulative distributions of the half-life of the maternal mRNAs in the absence of zygotic transcription in zebrafish. Represented are mRNAs with the highest fraction of strong optimal codons (red, CSC value > 0.03), weak optimal codons (light red, $0.03 > \text{CSC} > 0$), weak non-optimal (light blue $-0.03 < \text{CSC} < 0$), or strong non-optimal (blue $\text{CSC} < -0.03$). These four groups are non-overlapping. The control group contains all the maternal transcripts (gray) for which we could calculate the half-life. This result shows that the effects of the codon optimality (half-life depending in the codon composition) are gradual and depend on the strength of the codons.

B Cumulative distributions of the half-life of the maternal mRNAs in the absence of zygotic transcription in zebrafish. The red line represents genes enriched in optimal codons quantified 2 nucleotides out frame. The blue one is enriched non-optimal codons also counting in 2 nucleotides out frame. The gray lines represented all mRNAs. This result shows that the effect of codon optimality depends on the reading frame those codons are analyzed and as such depends on translation rather than a sequence motif independent of the reading frame. The *P*-values are shown, and the color code determines the conditions that were compared using a Wilcoxon rank-sum test.

C Cumulative plots shown as described in (B), for *Xenopus* embryos, counting the codons in frame (left panel), 1 nucleotide out frame (middle) and 2 nucleotides out frame (right panel). These results indicate codon optimality shapes mRNA half-life in a codon-dependent manner in *Xenopus* embryos. The group of transcripts enriched in either optimal or non-optimal in each of the three frames is different, meaning that the 364 optimal transcripts in frame are different from the 364 enriched in optimal codons when the codons composition was calculated in 1 or 2 nucleotides out of frame.

D, E Diagram of two transcripts that only differ in a single nucleotide (G in yellow) which creates a frameshift. One is enriched in optimal codons and the other in non-optimal codons due to a change in the reading frame. A time-course Northern blot shows that non-optimal (N) transcripts decay faster than their optimal (O) counterparts. These differences are lost when translation is blocked by cycloheximide (CHX) (D). The radioactive intensity is shown in italics and normalized to the 1 hpi sample for each injection. Right, qRT–PCR for the two exogenous transcripts with opposite optimality from Fig 1I. Non-optimal transcripts decay faster than their optimal counterparts (GFP was used as internal control). Results are shown as the averages $\pm$ standard error of the mean for three replicates. $^*P < 0.0001$.

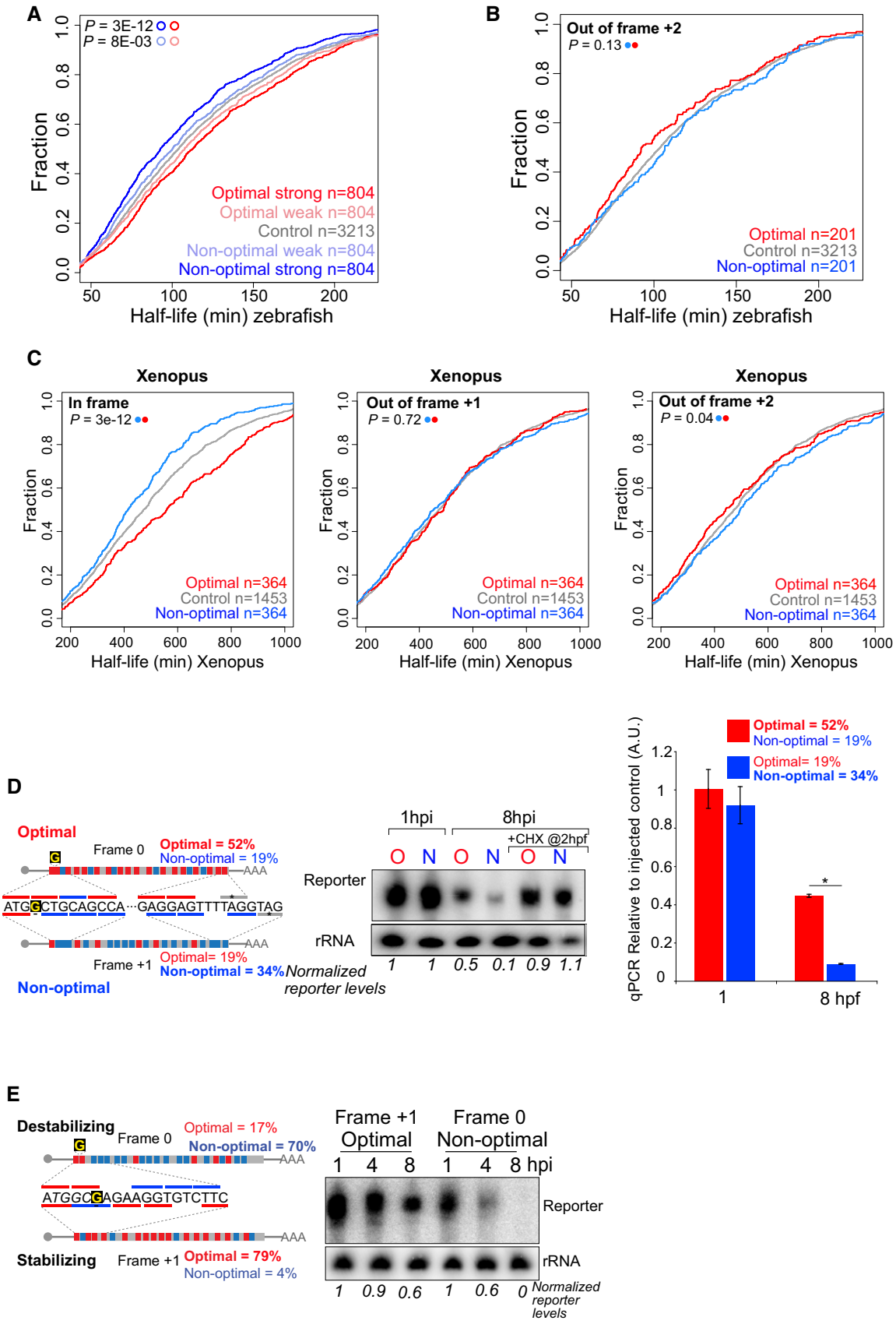


Figure EV2.

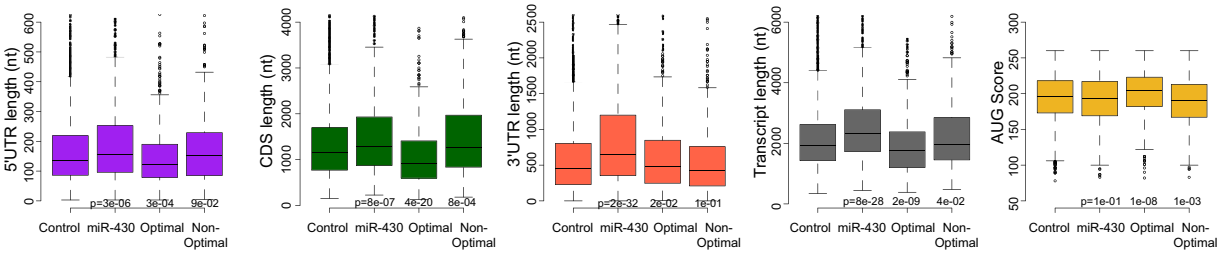
Figure EV3. Characterization of genes enriched in optimal and non-optimal codons.

A Box plot showing different transcript features of genes enriched in optimal codons, in non-optimal codons and genes with predicted miR-430 target sites. For example, miR-430 targets have significantly longer 3′UTRs. The box defines the first and third quartiles, with the median indicated with a thick black line and vertical lines indicate the variability outside the upper and lower quartiles.

B Enrichment for Gene ontology functions among optimal, non-optimal, and miR-430 target sites genes.

C Gene ontology functions enriched in orthologous genes in zebrafish, *Xenopus*, mouse, and *Drosophila* that share the same level of optimality.

A



B

Optimal

Regulation Of Protein Complex Assembly
Regulation Of Cellular Component Biogenesis
Positive Regulation Of Protein Complex Assembly
Ribonucleoprotein Complex Assembly
Ribonucleoprotein Complex Subunit Organization
Cellular Macromolecular Complex Assembly
Positive Regulation Of Cellular Component Biogenesis
Regulation Of Actin Filament Polymerization
Peptide Biosynthetic Process
Lipid Metabolic Process
Translation
Regulation Of Protein Polymerization
Cellular Component Organization Or Biogenesis
Cellular Component Organization
Protein Folding
Regulation Of Actin Filament Length
Regulation Of Actin Polymerization Or Depolymerization
Positive Regulation Of Cellular Component Organization
Peptide Metabolic Process
Macromolecular Complex Subunit Organization
Regulation Of Transcription Initiation From Rna Polymerase II Promoter
Positive Regulation Of Transcription Initiation From Rna Polymerase II Promoter
Regulation Of Dna-Templated Transcription, Initiation
Positive Regulation Of Dna-Templated Transcription, Initiation
Response To Oxidative Stress
Amide Biosynthetic Process

Non-optimal

Phosphatidylinositol Biosynthetic Process
Gpi Anchor Metabolic Process
Gpi Anchor Biosynthetic Process
Membrane Lipid Biosynthetic Process
Membrane Lipid Metabolic Process
Glycolipid Biosynthetic Process
Glycerophospholipid Biosynthetic Process
Small Molecule Metabolic Process
Phospholipid Biosynthetic Process
Cytoplasmic Transport
Cellular Process
Nucleobase-Containing Small Molecule Metabolic Process
Glycerolipid Biosynthetic Process
Glycolipid Metabolic Process
Liposaccharide Metabolic Process

miR-430

Lipid Modification
Single-Organism Process
Biological Regulation
Signal Transduction
Regulation Of Biological Process
Regulation Of Cellular Process
Single-Organism Cellular Process
Fatty Acid Catabolic Process

C

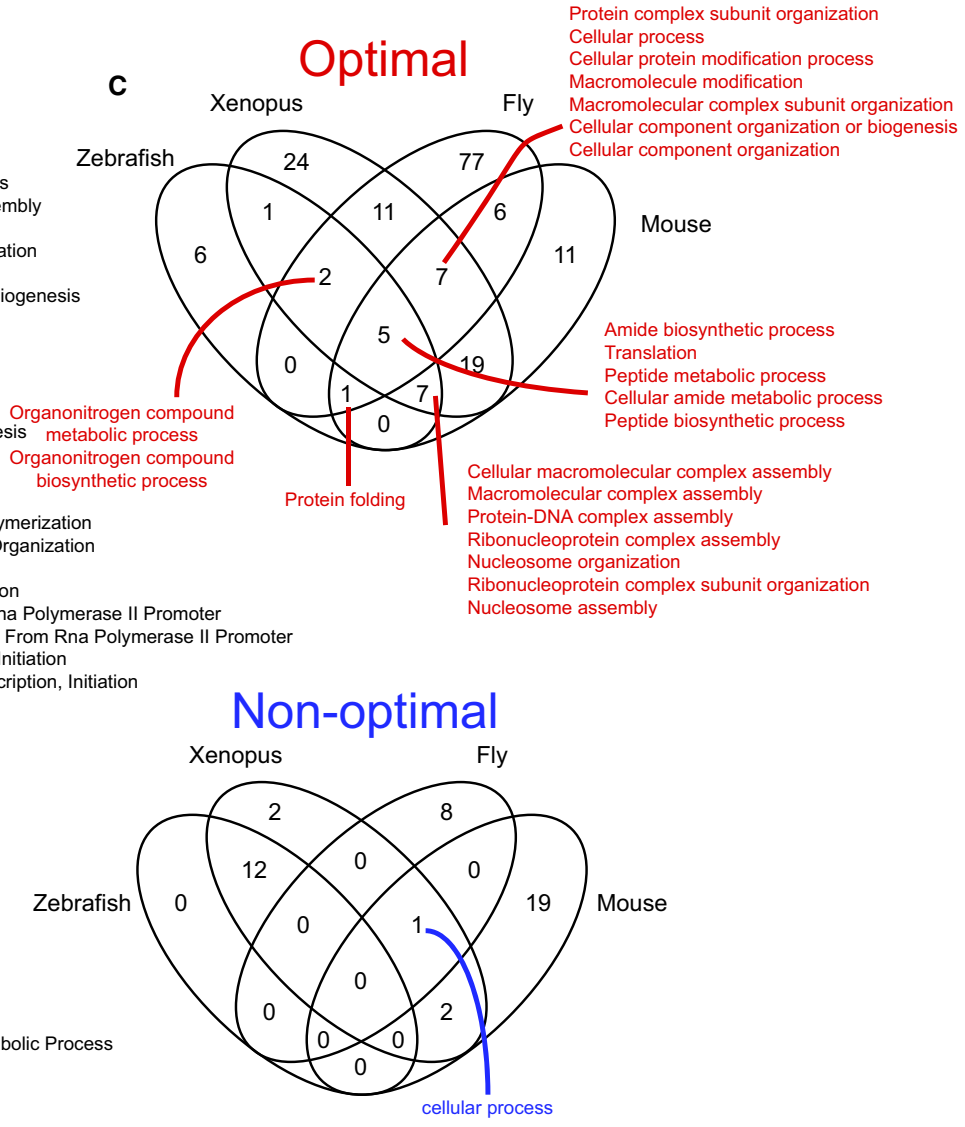


Figure EV3.

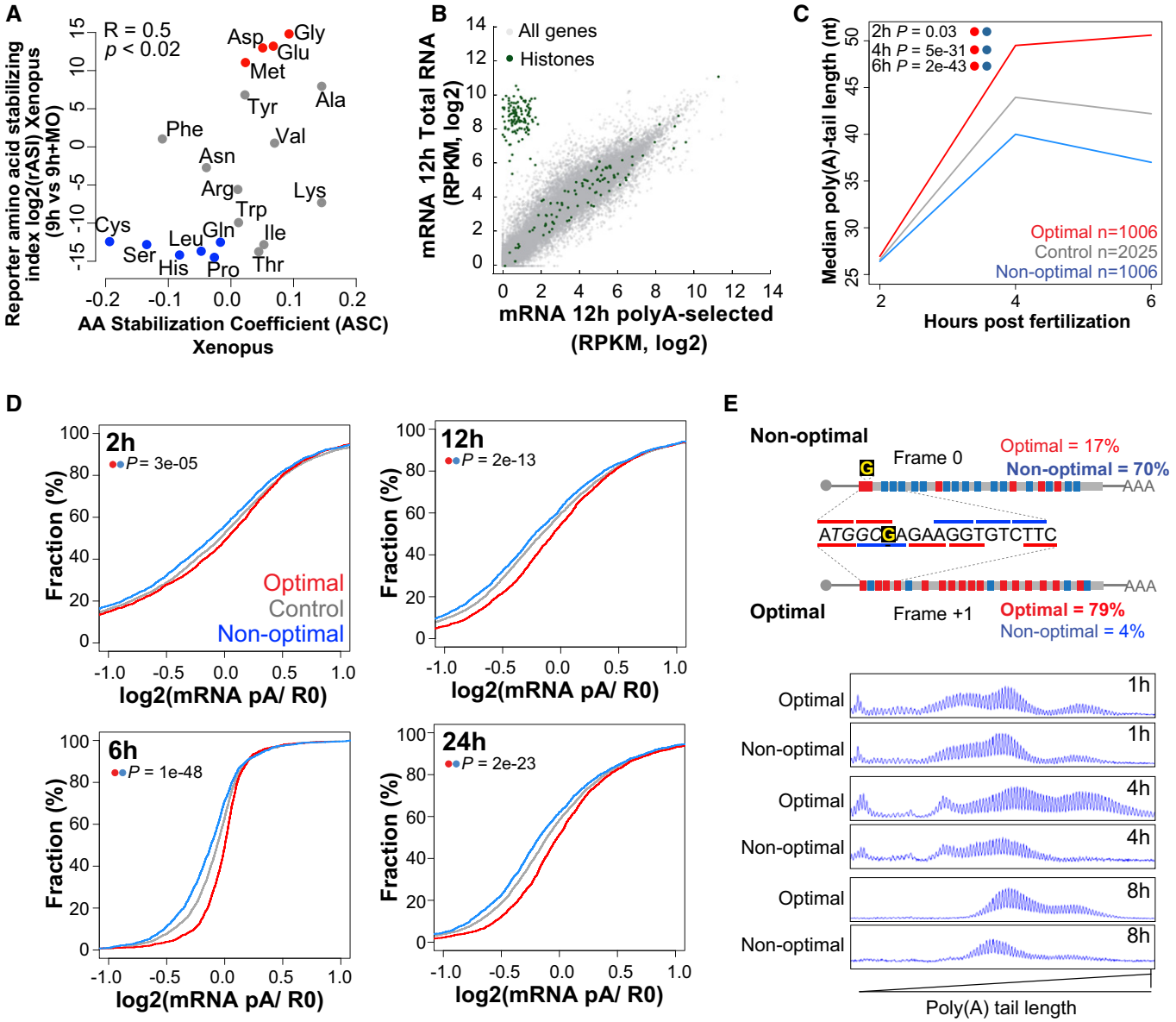


Figure EV4. Codon optimality regulates poly(A) tail length.

A Biplots of the reporter amino acid stabilizing index (rASI) comparing 9 h versus 9 h + MO and amino acid stabilization coefficient (ASC) in *Xenopus* embryos, $P < 0.02$, $R = 0.5$, Spearman rank correlation. Based on the consistent identification by both methods, we defined optimal amino acids highlighted in red and non-optimal amino acid in blue.

B Scatter plot showing expression levels (log2 RPKM) of purified poly(A) mRNA and total RNA at 12 hpf. Histone genes (that lack a poly(A) tail) are highlighted in green.

C Line plot showing the poly(A) tail length of control mRNAs (gray) based on PAL-seq data, mRNAs enriched in optimal codons (red), or in non-optimal (blue) at 2, 4, and 6 hpf ($P = 0.03$, $P = 5e\text{-}31$, $P = 2e\text{-}43$, Wilcoxon rank-sum test, respectively).

D Cumulative distributions of the ratio of purified poly(A)+/total RNA for genes enriched in optimal codons (red) ($n = 1{,}943$ at 2 h, 1,791 at 6 h, 2,207 at 12 h, and 2,469 at 24 h) or non-optimal codons (blue) ($n = 1{,}941$ at 2 h, 1,790 at 6 h, 2,208 at 12 h, and 2,462 at 24 h) or genes that were not included in any of the other two groups (gray) at 2, 6, 12, and 24 hpf ($P = 3e\text{-}05$, $P = 1e\text{-}48$, $P = 2e\text{-}13$, $P = 2e\text{-}23$, Wilcoxon rank-sum test, respectively).

E Single nucleotide resolution electrophoresis for poly(A) length for two sets of transcripts that only differ in a single nucleotide insertion G, but contains opposite codon optimality due to a change in the reading frame (Fig EV2E). This result shows that at 1 h the poly(A) tail length of both constructs is similar; however, by 4 and 8 h the length of the poly(A) tail is shorter in the transcript enriched for non-optimal codons.

Figure EV5. Codon optimality regulates translation efficiency.

A Cumulative distributions of mRNA levels and translation efficiency at 2, 6, 12, and 24 hpf for genes enriched in optimal (red) or non-optimal codons (blue). Also plotted are two control sets of genes that have similar RNA level distribution to those enriched in optimal or non-optimal codons (yellow for optimal and purple for non-optimal), and the rest of the analyzed genes (gray). Number of analyzed genes and *P*-values obtained after comparing the cumulative distribution between of the experimental set and the control set as indicated by the colors, Wilcoxon rank-sum test.

B Subcodon profile plot showing ribosome protected fragments (RPF) and input reads (total RNA) aligned to each of the transcripts shown. Reads are colored based on the frame (1, 2, or 3) position relative to the transcript. All putative ORFs (distal AUG-Stop) were also colored for each respective frame (blue, pink, and green boxes). The annotated ORF is marked. Note that the scale for both transcripts is the same. The gene enriched in optimal codons presents a higher translation efficiency because contain a higher ration of RPF/RNA than the gene enriched in non-optimal codons.

C Speculative model representing how different factors could modulate mRNA expression due to codon optimality. The length of the poly(A) tail as well as factors such microRNA or RNA binding protein can modulate translation and consequently modulate the effects of codon optimality on mRNA stability. Alternatively, specific cellular factors may affect the codon optimality code and so change the gene expression across tissues by modifying mRNA stability. For example, in the oocyte the level of translation is low. After egg activation and/or fertilization translation is activated, consequently a gene enriched in non-optimal codons may become unstable after translation is activated.

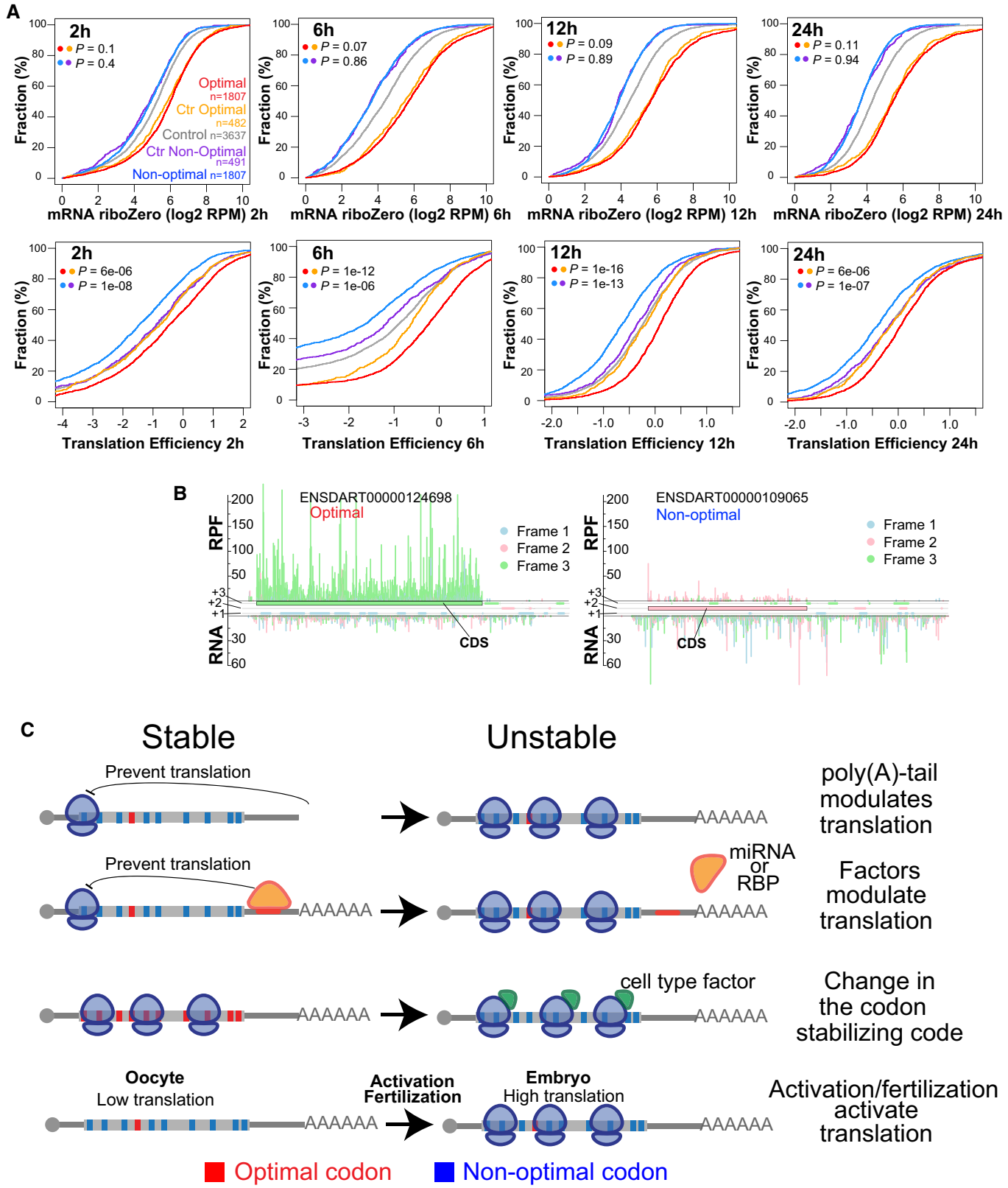

A
2h
P = 0.1
P = 0.4
Optimal n=1807
Ctr Optimal n=482
Control n=3637
Ctr Non-Optimal n=491
Non-optimal n=1807
Fraction (%)
mRNA riboZero (log2 RPM) 2h
6h
P = 0.07
P = 0.86
mRNA riboZero (log2 RPM) 6h
12h
P = 0.09
P = 0.89
mRNA riboZero (log2 RPM) 12h
24h
P = 0.11
P = 0.94
mRNA riboZero (log2 RPM) 24h
2h
P = 6e-06
P = 1e-08
Translation Efficiency 2h
6h
P = 1e-12
P = 1e-06
Translation Efficiency 6h
12h
P = 1e-16
P = 1e-13
Translation Efficiency 12h
24h
P = 6e-06
P = 1e-07
Translation Efficiency 24h
B
ENSDART00000124698
Optimal
RPF
RNA
Frame 1
Frame 2
Frame 3
CDS
ENSDART00000109065
Non-optimal
C
Stable
Unstable
Prevent translation
AAAAAA
poly(A)-tail modulates translation
miRNA or RBP
Factors modulate translation
cell type factor
Change in the codon stabilizing code
Oocyte
Low translation
Activation
Fertilization
Embryo
High translation
Activation/fertilization activate translation
Optimal codon
Non-optimal codon

Figure EV5.