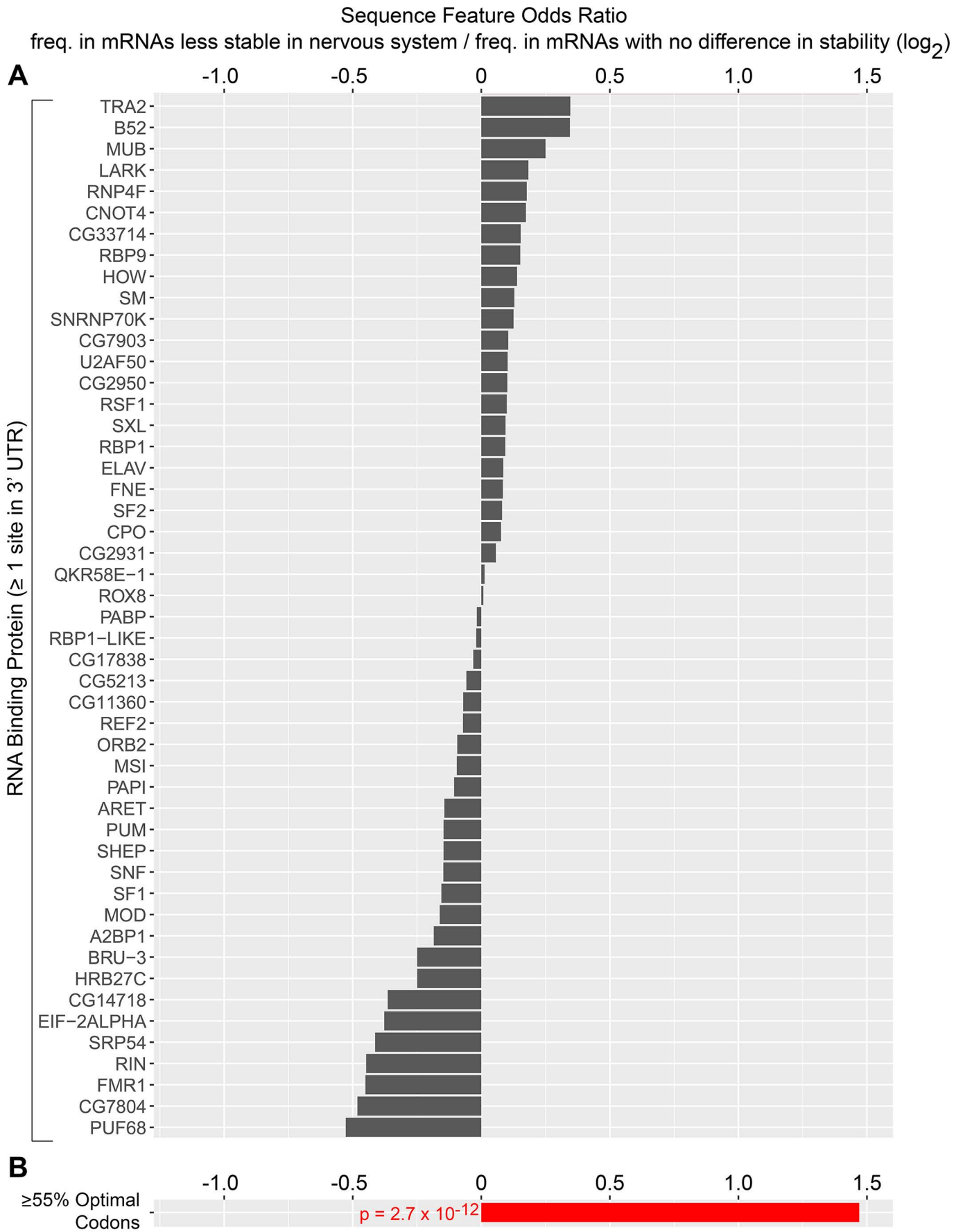
Cell Reports, Volume *24*

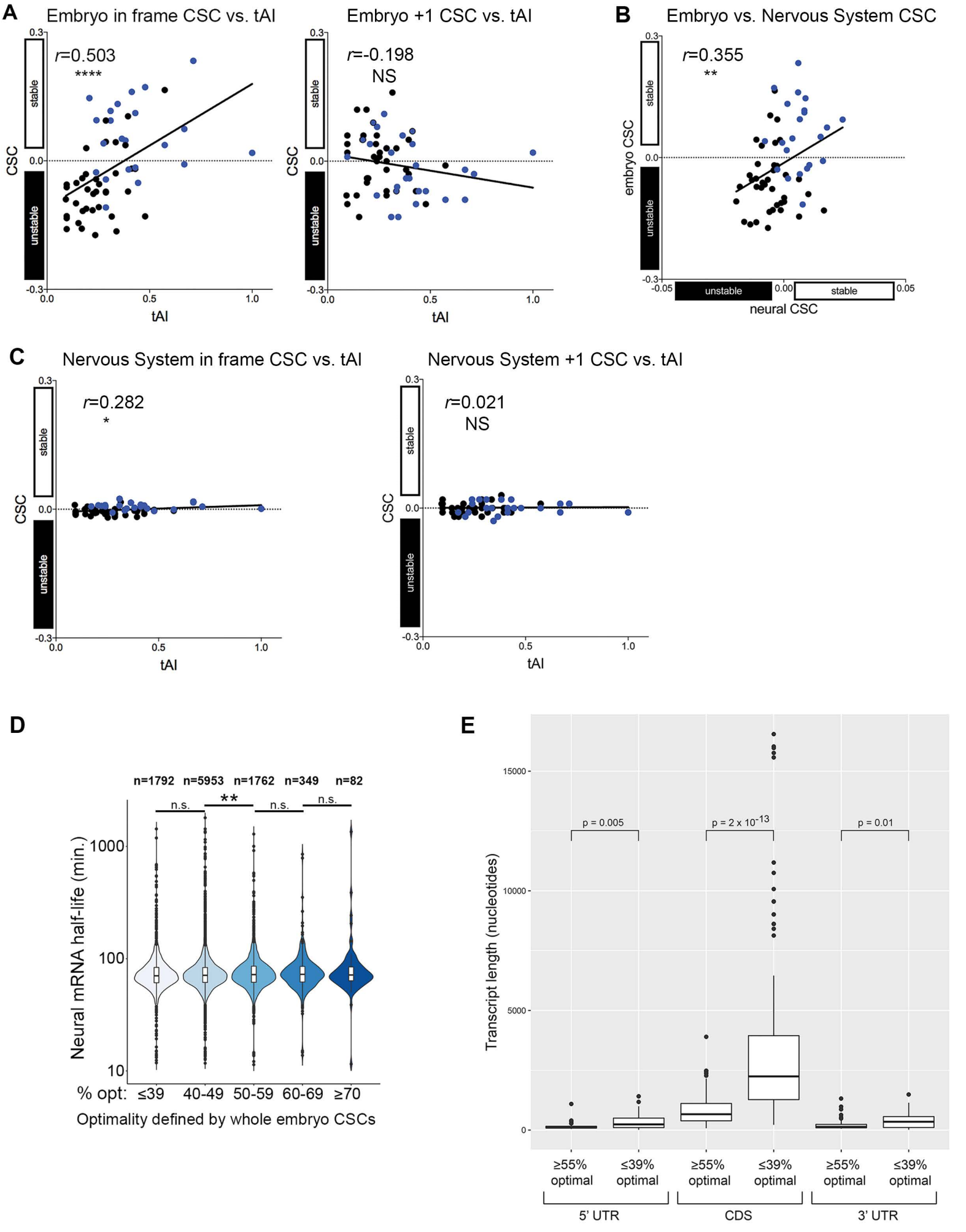
Supplemental Information

Attenuated Codon Optimality Contributes to Neural-Specific mRNA Decay in *Drosophila*

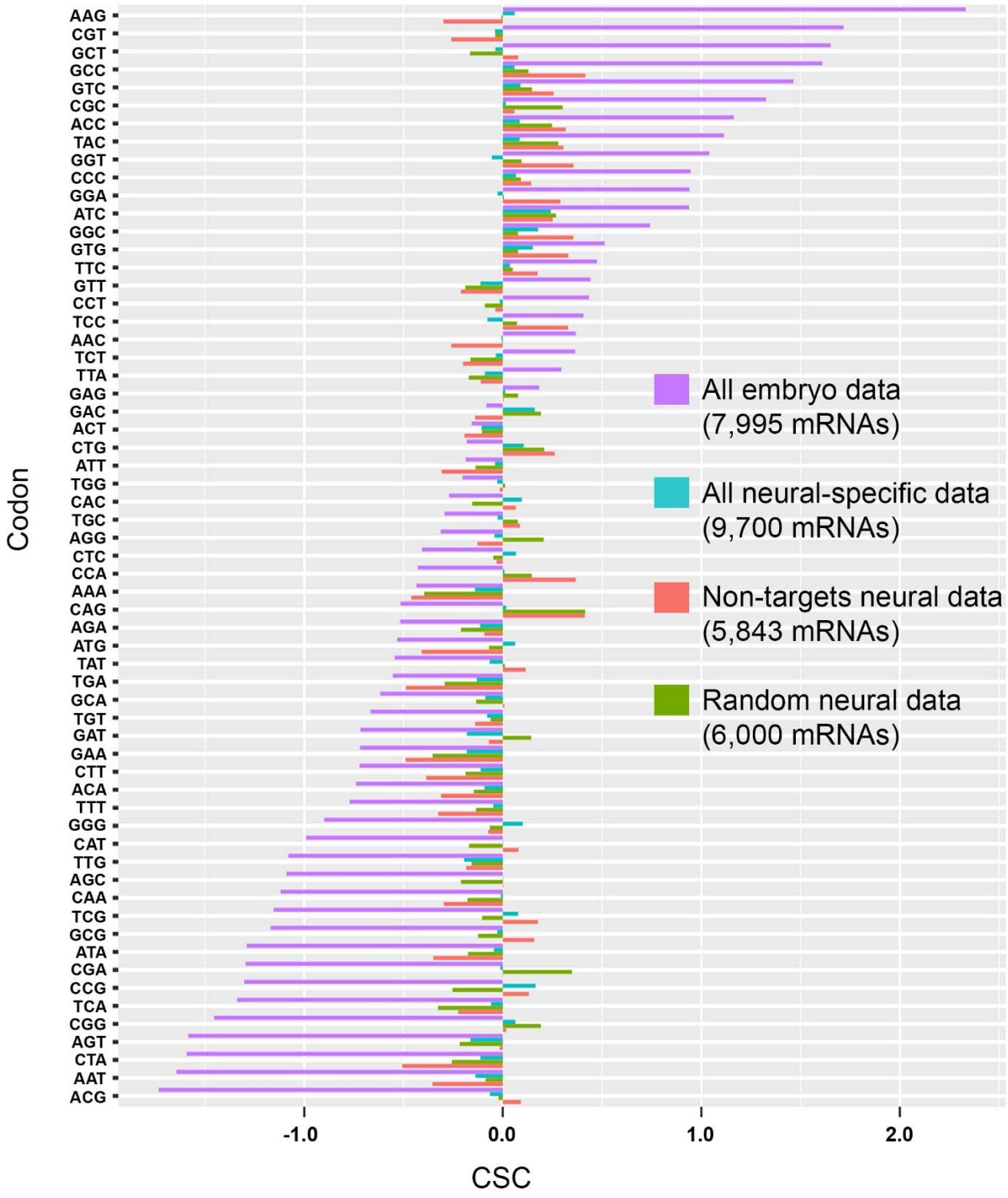
Dana A. Burow, Sophie Martin, Jade F. Quail, Najwa Alhusaini, Jeff Coller, and Michael D. Cleary


Sequence Feature Odds Ratio
freq. in mRNAs less stable in nervous system / freq. in mRNAs with no difference in stability (log2)
A
-1.0
-0.5
0
0.5
1.0
1.5
RNA Binding Protein (≥ 1 site in 3' UTR)
TRA2
B52
MUB
LARK
RNP4F
CNOT4
CG33714
RBP9
HOW
SM
SNRNP70K
CG7903
U2AF50
CG2950
RSF1
SXL
RBP1
ELAV
FNE
SF2
CPO
CG2931
QKR58E-1
ROX8
PABP
RBP1-LIKE
CG17838
CG5213
CG11360
REF2
ORB2
MSI
PAPI
ARET
PUM
SHEP
SNF
SF1
MOD
A2BP1
BRU-3
HRB27C
CG14718
EIF-2ALPHA
SRP54
RIN
FMR1
CG7804
PUF68
B
-1.0
-0.5
0
0.5
1.0
1.5
≥55% Optimal Codons
p = 2.7 x 10-12

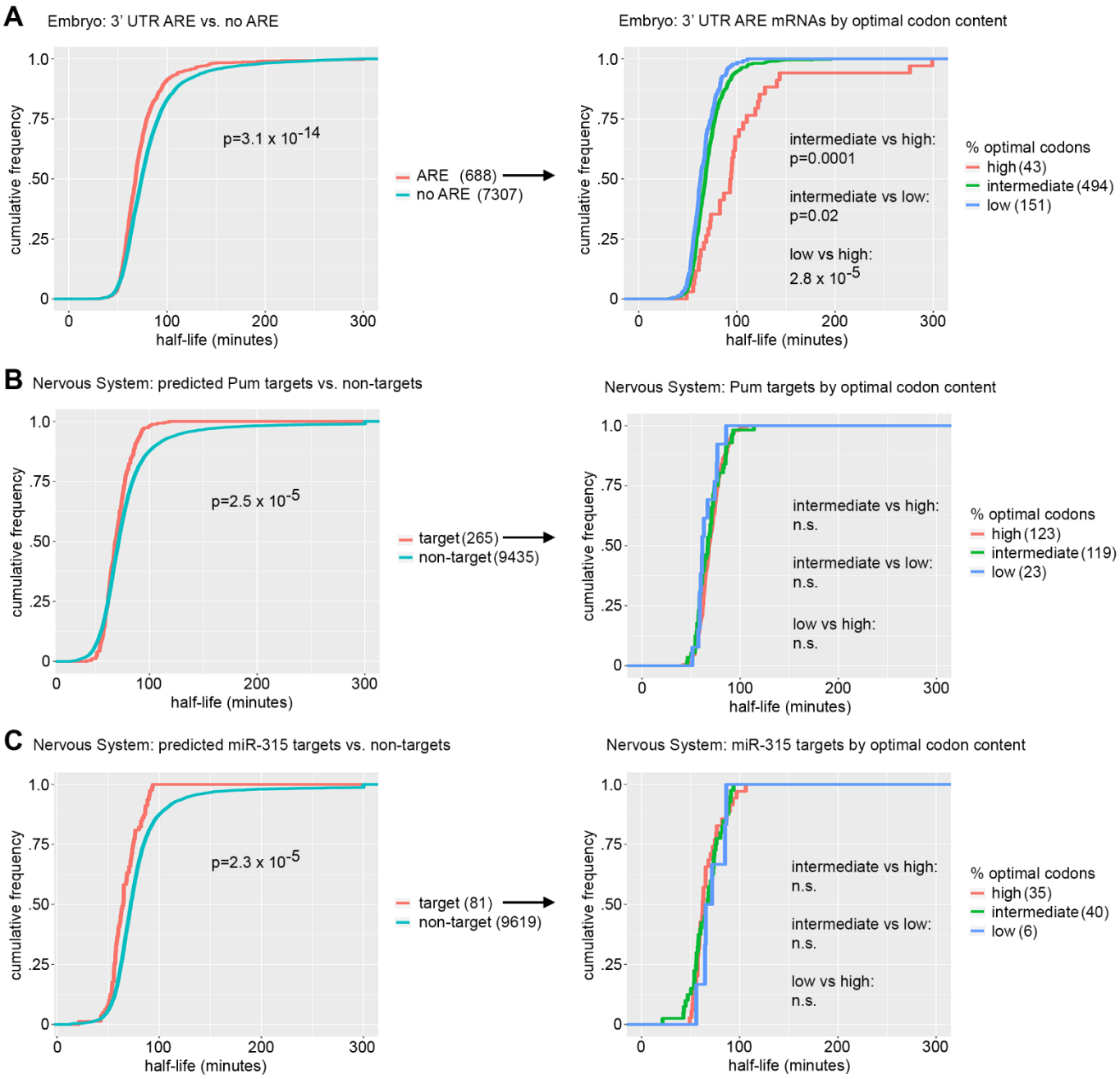
Supplemental Figure S1. Sequence feature enrichment or depletion among mRNAs that are less stable in the nervous system. ***Related to Figure 1.*** **A**. Enrichment or depletion of RNA-binding protein sites in mRNAs with shorter half-life in the nervous system. RBPmap (Paz et al., 2014) was used to identify RNA-binding protein (RBP) sites in the 3' UTR of 331 mRNAs with ≥ 1.5-fold decreased stability in the nervous system and 1,658 mRNAs with no difference in stability (neural / whole embryo half-life ratio between 0.75 and 1.25). RBP binding site frequencies in each group of mRNAs were used to calculate the odds ratio. Significance was tested using Fisher's Exact Test and no RBPs showed significant enrichment or depletion. **B**. Odds ratio analysis as described in part A using optimal codon content ≥ 55% as a sequence feature.


A
Embryo in frame CSC vs. tAI
r=0.503
****
Embryo +1 CSC vs. tAI
r=-0.198
NS
stable
unstable
CSC
tAI
B
Embryo vs. Nervous System CSC
r=0.355
**
embryo CSC
neural CSC
C
Nervous System in frame CSC vs. tAI
r=0.282
*
Nervous System +1 CSC vs. tAI
r=0.021
NS
D
n=1792 n=5953 n=1762 n=349 n=82
n.s.
**
n.s.
n.s.
Neural mRNA half-life (min.)
% opt: ≤39 40-49 50-59 60-69 ≥70
Optimality defined by whole embryo CSCs
E
Transcript length (nucleotides)
p = 0.005
p = 2 x 10-13
p = 0.01
≥55% optimal
≤39% optimal
5' UTR
CDS
3' UTR

Supplemental Figure S2. **Correlations between CSC, tAI, transcript stability and transcript length**. ***Related to Figures 1 and 2.*** **A**. Correlations based on whole embryo mRNA decay measurements. Left panel: in frame CSC. Right panel: +1 frameshifted CSC. **B**. Correlation between whole embryo CSCs and neural-specific CSCs. **C.** Correlations based on neural-specific mRNA decay measurements (as in part A). Pearson *r* values are listed. **** indicates $p < 0.0001$, ** indicates $p < 0.001$, * indicates $p < 0.01$, NS indicates no significant correlation. Predicted preferred codons are indicated as blue dots. **D**. Neural-specific half-life of mRNAs binned according to optimal codon content (% opt) as defined by whole embryo CSC calculations. Significant differences among categories were detected by Kruskal-Wallis test and p-values (** < 0.001 or no significant difference (n.s.)) for the indicated pairwise comparisons are based on Dunn's test. **E**. 5' UTR, coding sequence (CDS), and 3' UTR lengths were compared for all mRNAs present in the neural-specific and whole embryo datasets with ≥ 55% optimal codon content and ≤ 39% optimal codon content. Significance was determined using Kruskal-Wallis followed by Dunn's test to calculate p-values.



Supplemental Figure S3. **"Non-target" CSC calculations.** ***Related to Figure 1***. CSC calculations based on whole embryo data, neural-specific data, neural data that excludes all predicted targets of Pumilio, Fmrp, Orb2B, miR-124 and miR-315 (non-targets), and a random set of genes selected from the neural data.



Supplemental Figure S4. **Optimal codon content influences the stability of predicted targets of decay-promoting factors in whole embryos but not in the nervous system.** ***Related to Figures 1 and 2.*** **A.** Left panel: cumulative distribution plot of whole embryo mRNA half-lives for transcripts with a 3' UTR ARE element and transcripts with no ARE. Right panel: cumulative distribution plot of the predicted ARE targets grouped by optimal codon content, high = ≥55% optimal codons, intermediate = between 40 and 54% optimal codons, low = ≤39% optimal codons. **B** and **C**. The same analyses described for part A were performed using neural-specific mRNA decay data and predicted targets of Pumilio (B) or miR-315 (C). P-values are based on Kolmogorov-Smirnov tests.