

UC Merced

UC Merced Electronic Theses and Dissertations

Title

Dynamic regulation of RNA stability during Drosophila neurogenesis

Permalink

<https://escholarship.org/uc/item/2qv5n9xj>

Author

Burow, Dana

Publication Date

2016

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA, MERCED

DYNAMIC REGULATION OF RNA STABILITY DURING
DROSOPHILA NEUROGENESIS

A DISSERTATION
SUBMITTED TO THE SCHOOL OF NATURAL SCIENCES
AND THE DIVISION OF GRADUATE STUDIES
OF UNIVERSITY OF CALIFORNIA, MERCED
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

by

Dana Burow

August 2016

© Copyright by Dana Burow 2016
All Rights Reserved

I certify that I have read this dissertation and that, in my opinion, it is fully adequate in scope and quality as a dissertation for the degree of Doctor of Philosophy.

(Professor Michael Cleary) Principal Advisor

I certify that I have read this dissertation and that, in my opinion, it is fully adequate in scope and quality as a dissertation for the degree of Doctor of Philosophy.

(Professor David Ardell) Committee Chair

I certify that I have read this dissertation and that, in my opinion, it is fully adequate in scope and quality as a dissertation for the degree of Doctor of Philosophy.

(Professor Fred Wolf)

I certify that I have read this dissertation and that, in my opinion, it is fully adequate in scope and quality as a dissertation for the degree of Doctor of Philosophy.

(Professor Masashi Kitazawa)

Approved for the University of California, Merced Graduate Studies

For my family.

Learn from yesterday, live for today, hope for tomorrow.
The important thing is not to stop questioning.

-Albert Einstein

Acknowledgments

I would like to acknowledge the support of my family, colleagues, UC Merced and my funding agencies for their vital contributions to this work. First and foremost, I wish to thank my advisor, Professor Michael Cleary, for his generous and patient support of even the craziest of ideas. Immense thanks to my committee, especially Professor David Ardell, who introduced me to my first command line prompt, and always pushed me to think outside the box. I would like to acknowledge the important contributions of Maxine Umeh and Professor Jeff Collier's group at CASE (Vladimir Presnyak, Sophie Martin, Najwa Al-Husaini), which have set the stage for this study and inspired more exciting work to come. I would also like to thank the UC Davis Stem Cell Training Program, California Institute of Regenerative Medicine (TG2-01163), the UC Merced Graduate Division, and the Quantitative and Systems Biology Graduate Group for their generous support of my graduate career through training grants, fellowships, and travel funds. Thanks to Michael O'Connor and Robin Wharton for reagents and the TRiP at Harvard Medical School (NIH/NIGMS R01-GM084947) for providing transgenic RNAi fly stocks used in this study.

Special thanks to my wonderful family and friends; without your support I would not be who I am today. To my parents, Karen, Andy and Graham, you inspire me to never stop questioning and strive to understand the beautiful complexity around me. To my brothers, Paul and Carson, your love, support, and friendly competition drive me to always do my best. My beautiful friends, Terza and Leily, you are amazing women and role models for me in science and in life. Last but not least, to Cyprian, your love and sense of humor has kept me going through the toughest of times.

Dana A. Burow
5200 N Lake Road, SE1 325
Merced, Ca 95343
(916) 837-7909
dburow@ucmerced.edu

RESEARCH INTERESTS

I am interested in understanding how a single genome produces the diversity of cell types required for metazoan life. I have a solid background in biochemistry, molecular and cellular biology theory and research techniques with specific experience in RNA and stem cell biology fields. I am a graduate student in the Quantitative and Systems Biology group at UC Merced where I aim to use a systems biology approach to study mRNA decay dynamics during neural development. The goal of my research is to understand the patterns and mechanisms of post-transcriptional control of gene expression that contribute to a highly dynamic tunable gene expression program required for development.

EDUCATION

- 2012–2016 **PhD Quantitative and System Biology**, *The University of California, Merced.*
2010–2012 **MA Stem Cell Biology**, *California State University, Sacramento.*
2005–2009 **BS Microbiology**, *The University of California, Berkeley.*

EXPERIENCE

- 2012–Present **Graduate Researcher**, LAB OF DR. MICHAEL CLEARY, QUANTITATIVE AND SYSTEMS BIOLOGY GROUP, UC MERCED, Merced, CA.
PhD Dissertation: *Dynamic regulation of mRNA stability during Drosophila neural development*
Helped establish a novel tissue-specific RNA tagging assay using microarrays to study mRNA decay kinetics in vivo. Developed a novel mRNA decay reporter assay and next generation sequencing based approach using 4-thiouridine tagging in vivo to model global RNA decay kinetics. Evaluated equipment, advised purchasing committee and performed initial set-up and training for core equipment (e.g. Bioanalyzer, Covaris ultrasonicator). Lab and project management experience, including launching collaborations and advising junior researchers.
- 2012–2013 **Teaching Assistant**, BIOLOGICAL SCIENCES, UC MERCED, Merced, CA.
 - Bio 120L: Microbiology Lab
 - Bio 150: Embryos, Genes and Development
- 2011–2012 **Graduate Researcher**, LAB OF DR. PAUL KNOEPFLER, DEPT. OF CELL BIOLOGY AND HUMAN ANATOMY, UC DAVIS, Sacramento, CA.
Masters thesis: *Myc-interacting-zinc-finger-protein-1 (Miz-1): DNA and protein interactions in hESCs*
Demonstrated proficiency with human pluripotent stem cell culture and manipulation for various in vitro assays. Generated an in vitro expression system for the human Miz-1 protein and used this to assess novel Miz-1 protein and DNA binding interactions. Applied an in vitro next generation sequencing based assay to identify novel DNA binding motifs for the transcription factor Miz-1.
- 2011–2012 **Teaching Assistant**, BIOLOGICAL SCIENCES, CSU SACRAMENTO, Sacramento, CA.
 - Bio 2: Cells, Molecules and Gene
 - Bio 221A: Advanced Molecular Techniques Lab
- 2009 **Research Intern**, NOVOZYMES, INC., Fungal Expression Group, Davis, CA.
Developed a transformation protocol and reporter assay for assessing heterologous protein expression in a novel species of thermophilic filamentous fungus.

CERTIFICATIONS

- 2013 Next Generation Sequencing and Bioinformatics Short Course, UC Davis Bioinformatics Core, Davis, CA
- 2011 cGMP Stem Cell Manufacturing, UC Davis, Sacramento, CA
- 2011 Stem Cell Techniques Course, Buck Institute for Age Research, Novato, CA
- 2010 Graduate Teaching Associate Training Program, CSUS, Sacramento, CA

COMPUTER SKILLS

- Operating systems **UNIX, Windows, Mac**
- Languages **R, perl**
- Other tools **Microsoft Office, LaTeX, Galaxy, GraphPad Prism, Image J, VectorNTI, SnapGene, Tuxedo Suite, Rsubread, limma, Voom!, edgeR, DESeq, SAMtools, codonW**

FELLOWSHIPS AND AWARDS

- 2016 Graduate Dean's Dissertation Year Fellowship, UC Merced
- 2013–2016 CIRM Stem Cell Training Program Fellowship, UC Davis and UC Merced
- 2014 Outstanding Student Poster Presentation Award, Fall QSB Retreat, UC Merced
- 2013,2015 QSB Summer Research Fellowship, UC Merced
- 2012 QSB First Year Recruiting Scholarship, UC Merced
- 2010–2012 CIRM Bridges to Stem Cell Research Fellowship, CSUS and UC Davis

PUBLICATIONS

Burow, D. A., Umeh-Garcia, M. C., True, M. B., Bakhaj, C. D., Ardell, D. H., Cleary, M. D. (2015). Dynamic regulation of mRNA decay during neural development. *Neural Development*, 10(11). <http://doi.org/10.1186/s13064-015-0038-6>

Barrilleaux, B. L., Burow, D., Lockwood, S. H., Yu, A., Segal, D. J., Knoepfler, P. S. (2014). Miz-1 Activates Gene Expression via a Novel Consensus DNA Binding Motif. *PLoS ONE*, 9(7), e101151. <http://doi.org/10.1371/journal.pone.0101151>

PRESENTATIONS

Dana Burow, Maxine Umeh, Vladimir Presnyak, Sophie Martin, Jeff Coller, Dave Ardell, and Mike Cleary. Dynamic regulation of mRNA stability during Drosophila neural development. *Society for Developmental Biology West Coast Regional Meeting*, Fish Camp, CA (poster) 3/2015

Dana Burow, Maxine Umeh, and Michael Cleary. Neural developmental regulation of mRNA decay in Drosophila, *QSB Fall Retreat*, Merced, CA (poster) 9/2014

Dana Burow, Maxine Umeh, and Michael Cleary. Neural developmental regulation of mRNA decay in Drosophila, *FASEB, Post-transcriptional control of gene expression: mechanisms of mRNA decay Conference*, Big Sky, MT (poster) 7/2014

Abbreviations

RNA-seq: RNA-sequencing	XRN1: pacman	mRNP: messenger ribonucleo- protein complex
ChIP-chip: Chromatin im- munoprecipitation on microar- ray chip	UTR: untranslated region	Wnt: wingless
mRNA: messenger RNA	NMD: nonsense-mediated de- cay	AD: Alzheimer's disease
RBP: RNA binding protein	siRNA: small interfering RNA	ALS: Amyotrophic Lateral Sclerosis
ncRNA: non-coding RNA	AGO2: Argonaute 2	FMRP: Fragile X mental re- tardation protein
NMDA: N-methyl-D- aspartate receptor	PTC: premature termination codon	ActD: actinomycin D
Arc: activity-regulated cytoskeleton-associated protein	EJC: exon junction complex	TU: 4-thiouracil
ELAV/Hu: embryonic lethal abnormal vision	UPF: up-frameshift	GAL4: galactosidase induced activator
ARE: AU-rich element	ARF1: ADP ribosylation fac- tor 1	UAS: upstream activating se- quence
PABP: poly(A) binding pro- tein	SLPB: Stem-loop binding pro- tein	UPRT: uracil-phosphoribosyltransferase
PAN2/3: positive and nega- tive regulation of poly(A) nu- clease	miRNA: micro-RNA	TUd: 4-thiouridine
CCR4-NOT: CCR4-NOT transcription complex	MAPK: p38 mitogen- Activated Protein kinase	eIF-4a: eukaryotic initiation factor 4E
Dcp: death caspase-1	ERK: Extracellular regulated kinase (rolled)	CNS: central nervous system
	JNK: c-Jun N-terminal kinase (basket)	iPSC: induced pluripotent stem cell
		ESC: embryonic stem cell

HFF : human foreskin fibroblast	PRE : Pumilio recognition element	tAI : tRNA adaptive index
PCBP : poly(rC) binding protein	RNAi : RNA interference	ENC : effective number of codons
Hey : Hairy/E(spl)-related with YRPW motif	Fas2 : Fasciclin 2	CAI : codon adaptation index
DI : dorsal	Eph : Eph receptor tyrosine kinase	GC : guanine-cytosine
ind : intermediate neuroblasts defective	otk : off-track	RpL : ribosomal large protein subunit
vnd : ventral nervous system defective	Sema-1a : semaphorin-1a	RpS : ribosomal small protein subunit
Dr : Drop	MAGUK : Membrane-associated guanylate kinase	par-6 : abnormal embryonic partitioning of cytoplasm
btd : buttonhead	Dlg1 : discs large 1	baz : bazooka
ems : empty spiracles	wdb : widerborst	His : histone
chinmo : Chronologically inappropriate morphogenesis	nmo : nemo	Syb : Synaptobrevin
ap : apterous	heph : hephaestus	lola : longitudinals lacking
SoxN : SoxNeuro	step : steppke	rRNA : ribosomal RNA
dpn : deadpan	Imp : IGF-II mRNA-binding protein	FPKM : Fragments Per Kilo-base of transcript per Million mapped reads
pros : prospero	msi : musashi	TMM : Trimmed Mean of M-values
pb : proboscipedia	tyf : twenty-four	LOWESS : locally weighted scatterplot smoothing
CaMKII : Calcium/calmodulin-dependent protein kinase II	GFP : green fluorescent protein	tUMP : thiouracil monophosphate
Ef1alpha100E : Elongation factor 1 α 100E	hsp70 : heat-shock protein 70	Vps15 : Vacuolar protein sorting 15
CASK : Calcium/calmodulin-dependent protein kinase ortholog	SAM : Significance analysis of microarrays	ef1beta : Elongation factor 1 β
PUF : Pumilio family	FDR : false discovery rate	FLuc : Firefly luciferase
Pum : Pumilio	GO : gene ontology	cpm : counts per million mapped reads
	DAVID : Database for Annotation, Visualization and Integrated Discovery	
	tRNA : transfer RNA	
	CSC : codon stabilization coefficient	

Summary of Dissertation

High-throughput gene expression and functional genomics technologies have revolutionized the study of gene expression during development. Yet traditional genome-wide approaches fail to address complex patterns of gene expression, which include both synthesis and decay. Post-transcriptional regulation is essential for temporal and spatial control of protein expression. Although candidate studies have revealed important mechanisms of post-transcriptional regulation of gene expression via mRNA decay, there is a lack of global information on mRNA decay during development. TU-decay technology allows measurement of genome-wide mRNA decay in intact *Drosophila* embryos, across all tissues and specifically in the nervous system. Using this approach neural-specific decay kinetics were observed, including stabilization of transcripts encoding regulators of axonogenesis and destabilization of transcripts encoding ribosomal proteins and histones. These data also demonstrate mRNA stability is correlated to physiologic properties of mRNAs; mRNAs that are predicted to be translated within axon growth cones or dendrites have long half-lives while mRNAs encoding transcription factors that regulate neurogenesis have short half-lives. Searches for candidate *cis*-regulatory elements identified enrichment of the Pumilio recognition element (PRE) in mRNAs encoding regulators of neurogenesis. Decreased expression of the RNA-binding protein, Pumilio, stabilizes predicted neural mRNA targets and presence of a 3'UTR PRE is sufficient to trigger mRNA decay in the nervous system. Clustering of mRNAs by codon content reveals a bias in codon usage related to physiological function and neural-specific mRNA decay kinetics; ribosomal proteins and histones have higher fractions of optimal codons while transcriptional regulators and mRNAs predicted to be localized to synapses have lower fractions of

optimal codons. In *Drosophila* whole embryos, codon optimality is correlated with mRNA decay, while in the nervous system, codon-mediated effects on mRNA decay are attenuated. This work demonstrates a dynamic post-transcriptional program including mRNA decay allows fine-tuning of gene expression during neural development.

Contents

Acknowledgments	vii
Curriculum Vitae	viii
Abbreviations	x
Summary of Dissertation	xii
1 Introduction	1
1.1 Post-transcriptional control of gene expression	1
1.1.1 mRNA stability and neural development	4
1.1.2 mRNA stability and neuronal disease	6
1.2 Approaches for monitoring mRNA decay	7
1.2.1 Modeling mRNA decay	8
1.3 Objective of the Study	9
2 <i>In vivo</i> mRNA decay dynamics	11
2.1 Introduction	11
2.2 Results	12
2.2.1 TU-decay measures tissue-specific mRNA decay	12
2.2.2 Neural-specific decay dynamics	15
2.2.3 Neural mRNA stability is associated with gene function	21
2.2.4 Patterns of regulation related to neural mRNA decay dynamics	22
2.2.5 <i>Pumilio</i> knockdown stabilizes neural PRE-containing mRNAs	25

2.2.6	PRE 3'UTR deletion is sufficient to stabilize GFP reporter mRNAs	28
2.3	Discussion	29
2.4	Approach	31
2.4.1	Fly lines	31
2.4.2	Embryo permeabilization and pulse-chase labeling	31
2.4.3	TU-RNA purification	32
2.4.4	cRNA probe synthesis and array hybridization	32
2.4.5	mRNA half-life calculations	32
2.4.6	Statistical analysis	33
2.4.7	Gene ontology analysis	33
2.4.8	Immunofluorescent imaging	33
2.4.9	RT-qPCR	34
3	Codon-mediated mRNA decay dynamics	35
3.1	Introduction	35
3.2	Results	37
3.2.1	Codon optimality is correlated with mRNA stability in <i>Drosophila</i> embryos	37
3.2.2	Codon-mediated decay effects are attenuated in neural tissues	41
3.3	Discussion	45
3.4	Approach	46
3.4.1	Fly lines	46
3.4.2	Codon analysis	46
3.4.3	Transcript CDS reporter assay	47
4	Unbiased TU-decay via RNA-seq	52
4.1	Introduction	52
4.2	Preliminary Results	53
4.2.1	Strategy for transcriptome-wide TU-decay by RNA-seq	53
4.2.2	Comparison of TU-decay by microarrays and RNA-seq	57
4.3	Discussion	61

4.4	Approach	62
4.4.1	Library preparation and sequencing	62
4.4.2	Read mapping, summarization and normalization	62
4.4.3	mRNA half-life calculations	62
5	Synthesis	64
A	Appendix	67
A.1	Supplemental methods	67
A.1.1	Western blot	67
A.1.2	RNA-seq analysis pipeline	68
A.2	Supplemental data	72
	Bibliography	79

List of Tables

2.1	Functional annotation of whole embryo high and low-stability mRNAs	14
2.2	Functional annotation of neural-specific high and low-stability mRNAs	17
2.3	Functional annotation of mRNAs with altered stability in the nervous system versus whole embryo	20
3.1	Functional annotation of mRNAs with high or low-optimal codon content	41
A.1	Custom primers used in the study.	72

List of Figures

2.1	TU-decay measures genome-wide mRNA decay	13
2.2	<i>Pros-UPRT</i> TU-tagging isolates neural mRNAs from intact embryos	15
2.3	Comparison of neural and whole embryo mRNA decay	18
2.4	Neural mRNA decay dynamics are linked to gene function	23
2.5	<i>Pumilio</i> regulates neural mRNA stability	25
2.6	<i>Pumilio</i> knockdown does not lead to gross defects in embryonic motor neuron development	27
2.7	PRE 3'UTR deletion is sufficient to stabilize GFP reporter mRNAs .	29
3.1	Heatmap of codon usage across 3631 genes in the <i>Drosophila</i> embryo	38
3.2	Codon optimality is associated with mRNA stability in the whole embryo	39
3.3	Codon composition is associated with stability and gene function . .	48
3.4	Codon optimality is not associated with mRNA stability in the embry- onic nervous system	49
3.5	Codon composition among classes of neural destabilized and stabilized mRNAs	50
3.6	Candidate CDS reporter mRNAs	51
4.1	Strategy for transcriptome-wide TU-decay by RNA-seq	54
4.2	Mean-variance relationship of TU-decay RNA-seq and microarray data	56
4.3	Modeling differential expression and decay via limma in TU-decay RNA-seq data	58
4.4	Distributions of RNA half-life and R^2 for microarray and RNA-seq based decay models	59

4.5	Whole embryo RNA decay models via RNA-seq	60
A.1	Initial mRNA concentration, 3'UTR length and uracil content do not correlate with transcript stability	73
A.2	<i>Pumilio</i> knockdown leads to reduction of Pumilio protein	74
A.3	CAI and CSC values have distinct distributions	74
A.4	tRNA copy number and CSC values are correlated only in the whole embryo	75
A.5	Neural destabilized classes of mRNAs show the strongest codon usage bias by multiple measures	76
A.6	Initial and lifetime mRNA expression (cpm) do not correlate with tran- script stability.	77
A.7	Overall tRNA expression does not correlate with CSC	78

Chapter 1

Introduction

High-throughput gene expression and functional genomics technologies like microarrays [1], RNA-sequencing (RNA-seq) [2], and chromatin immunoprecipitation (e.g. ChIP-chip [3]) have implicated a complex network of *cis*-regulatory elements and their coordinate transcription factors in controlling important aspects of development. However, differential transcription alone fails to consider the role of post-transcriptional regulation of gene expression, such as transcript stability, in determining expression levels. Post-transcriptional regulation via messenger RNA (mRNA) decay is essential for precise temporal and spatial control of protein expression, especially during developmental transitions [4]. Advances in biosynthetic tagging of RNA *in vivo* coupled with high throughput gene expression technologies like microarrays and RNA-seq have allowed genome-wide quantitation of mRNA decay *in vivo* [5–7]. Integration of transcript decay and *cis*-element information is the basis for establishing global *in vivo* models of the regulation of mRNA stability. These models are essential for strengthening our understanding of post-transcriptional control of gene regulation during development.

1.1 Post-transcriptional control of gene expression

Development requires a dynamic tunable gene expression program to generate a high diversity of cell types from a single genome. Expression patterns are genetically

programmed and subject to regulation at the gene (DNA), transcript (RNA) and protein level via input from distinct cellular signals. Post-transcriptional control of gene expression is mediated by protein-RNA and RNA-RNA interactions which direct RNA processing, localization, decay and translation.

Transcript decay and protein expression are intimately related processes; genetic features within transcripts direct both translation and decay. Eukaryotic mRNAs are transcribed with a 5' 7-methylguanosine cap and 3' poly(A) tail varying in length from 25-200 residues. These two general features are determinants of transcript stability protecting the transcript from the major exonucleolytic decay pathways and promoting translation through interaction with the translation initiation factor, eIF4E, cap-binding complex, and poly(A) binding protein (PABP) [8,9]. Transcripts become vulnerable to decay when either of these terminal structures are compromised or the transcript is subject to endonucleolytic cleavage. The majority of mRNAs undergo $5' \rightarrow 3'$ or $3' \rightarrow 5'$ exonucleolytic decay via poly(A) tail shortening, decapping and cleavage. Deadenylation is an important first step in bulk mRNA decay pathways. However, this first step is reversible and transcripts may be readenylated and return to a polysome state. Deadenylation is important in nascent mRNA processing via the PABP-independent PAN2-PAN3 poly(A) nuclease complex and large-scale mRNA decay via the CCR4-NOT complex deadenylation [10–12]. Although these deadenylation complexes are implicated in bulk cellular deadenylation, there is considerable redundancy in poly(A) nucleases, likely important in fine tuning local control of translation. Decapping occurs via the $5' \rightarrow 3'$ decay pathway and is catalyzed by the decapping enzyme (Dcp) complex [13]. Then, exoribonuclease, XRN1, catalyzes $5' \rightarrow 3'$ degradation. A large complex of $3' \rightarrow 5'$ exonucleases, termed the *exosome*, mediates $3'$ end decay. Work in yeast has strengthened the links between bulk mRNA decay and translation, demonstrating that $5' \rightarrow 3'$ decay occurs co-translationally on polysomes [14] and decay kinetics are tightly linked to codon optimality [15]. Codon optimality reflects the interplay between transcript codon usage and the availability of coordinate charged tRNAs, affecting both the rate of translation elongation and decay. In zebrafish embryos codon optimality is implicated in bulk maternal transcript decay via CCR4-NOT complex deadenylation, where 3'UTR length, modulates this effect [16].

This recent work suggests codon optimality allows a highly context-specific means of regulating global mRNA translation and decay.

There are a number of alternate routes of mRNA decay including endoribonucleolytic decay, nonsense-mediated decay (NMD), non-stop decay, and no-go decay [17]. These additional pathways serve as a means of both transcript quality control and offer alternative avenues for post-transcriptional gene regulation. Endonucleolytic cleavage is a potent means of transcript decay; short interfering RNA (siRNA) mediated cleavage by argonaute protein-2 (AGO2) [18, 19] and a number of cellular endonucleases have been implicated in endonucleolytic attack [20, 21]. Transcript surveillance pathways (i.e. NMD, non-stop decay, and no-go decay) are another translation dependent means of mRNA decay. NMD is responsible for destroying transcripts containing premature termination codons (PTCs), which may arise from genetic mutation, altered translation initiation or 3'UTR lengthening [22]. NMD complex proteins (i.e. UPF proteins) detect the irregular mRNP, which may contain exon-junction complex (EJC) proteins downstream of the PTC or abnormally long 3'UTRs [23–25]. In cases where cellular mRNAs are targeted by NMD as a means of normal gene regulation, mRNAs contain features that alter mRNP conformation including: upstream open reading frames (ORFs), 3'UTR introns, or frameshifts [26]. Regulatory RBPs may be recruited to transcript 3'UTRs to induce UPF1-dependent decay that is distinct from classical NMD. RBPs that bind stem-loop structures in the 3'UTRs of histones and ADP-ribosylation factor-1 (ARF1) mRNAs (stem-loop-binding protein (SLBP) [27] and Staufen [28] respectively) recruit UPF1 to these transcripts to initiate decay. Cellular mRNA decay can occur in cytoplasmic foci known as *P-bodies* [29]. Components involved in many aspects of mRNA processing, translation and decay have been found to localize to P-bodies [29–36], although P-bodies may vary considerably in composition. In addition to canonical 5' → 3' decay components P-bodies may also contain components of the noncanonical microRNA (miRNA) mediated mRNA-decay pathway [37] or nonsense-mediated decay (NMD) pathway [34, 35]. There is evidence to suggest P-bodies assemble in response to demands on mRNA sequestration and turnover, for example, under stress conditions [38, 39]. P-bodies appear to be dynamic mRNA-dependent protein-RNA aggregates that coordinate regulation

between mRNA decay and translation.

Cellular signaling pathways including p38 mitogen-activated protein kinase (MAPK) [40,41], extracellular signal-regulated kinase (ERK) [42], Jun N-terminal kinase (JNK) [43,44], and Wnt [45] affect the formation of messenger ribonucleoprotein complexes (mRNPs). The composition of these mRNPs is flexible, creating a potent system for efficiently altering protein expression. mRNP composition is encoded within transcripts through stability-determining sequence elements and RBP recruitment to target mRNAs. *Cis*-acting elements may be located throughout the transcript in the 5'UTR, coding sequence, 3'UTR, or even within intronic regions of a nascent transcript. *Cis*-acting elements in the 3'UTR are of particular interest owing to their accessibility to regulatory proteins. AU-rich elements (AREs) are a class of sequence element that affect transcript stability and are primarily located within 3'UTRs. These elements contain the signature *AUUUA* pentamer, but vary widely in composition and position. AREs are known to regulate transcript stability in genes requiring precise tuning of gene expression like transcription factors, cytokines and proto-oncogenes [46]. A class of RBPs called Puf proteins (e.g. Pumilio) bind to UG-rich sequence elements in target mRNAs silencing protein expression through translation inhibition and accelerated decay [47–49].

1.1.1 mRNA stability and neural development

Global gene expression analyses in the context of cellular stress response [50, 51] and development [52] demonstrate an essential role for post-transcriptional control of gene expression via regulation of mRNA decay. Modulation of mRNA stability allows for precise regulation of protein abundance and localization [53, 54]. This level of regulation is particularly important in the developing nervous system, where control of mRNA stability and localization is important for cell fate determination and neuron structure [55–57]. Regulation of mRNA stability plays a crucial role in neurodevelopmental processes including: axon guidance [58], synaptic architecture [59], plasticity [60] and function [61].

Tissue-specific expression of *trans*-acting RBPs, ncRNAs, and *cis*-acting elements

in target mRNAs direct regulation of mRNA stability with temporal and spatial specificity. For example, the embryonic lethal abnormal vision (ELAV/Hu) family is a well-studied class of *trans*-acting RBPs involved in neural-specific regulation of transcript stability, localization and translation [62]. ELAV/Hu family RNA recognition motifs are highly conserved and bind preferentially to AREs [63]. ELAV RBPs mediate neural-specific 3'UTR isoform expression, which alter target mRNA stability and translation [64]. Recent studies suggest distinct neural mRNA isoform expression plays an important role in *trans*-acting factor recognition and transcript stability [65,66]. In another example, neural-specific miRNA, miR-128, modulates activity of the NMD pathway through direct targeting of NMD factors: RNA helicase, UPF1, and MLN51, a component of the exon-junction complex (EJC) involved in NMD [55]. While miR-128 has modest effects on direct target expression, it represses NMD activity resulting in an up-regulation of neural NMD targets involved in neural differentiation, maturation, and function [55]. These types of candidate studies have provided functional and mechanistic insights into a neural developmental mRNA decay regulatory network; however, comprehensive evaluation of the mechanisms and functions of post-transcriptional regulation via mRNA decay remains relatively unknown.

Subcellular regulation of protein expression is crucial for proper neuron structure and function. Local regulation of transcript stability is influenced by external stimuli via neuronal signaling pathways (e.g. NMDA signaling in the synapse and Slit secreted axonal guidance cues) and has implications for development including axon guidance, and synaptic plasticity, learning and memory. Growth cones direct axon patterning and growth in response to interaction with extracellular cues. These extracellular signals are mediated in part through their control on axonal mRNA stability and translation. Robo3.2 is an axonal growth cone-associated receptor that responds to secreted Slit guidance proteins. *Robo3.2* mRNA is degraded via the NMD pathway in axons that reach the floor plate [58]. *Arc* (activity-regulated cytoskeleton-associated protein) is a dendrite-localized mRNA and immediate early gene. Synaptic activation is sufficient to stabilize *Arc* mRNA in active synapses while allowing mRNA decay to occur in inactive dendrites [67]. Precise post-transcriptional

regulation of *Robo3.2* and *Arc* are mediated by neural receptor activation, linking post-transcriptional regulation via mRNA decay to important neural developmental processes.

1.1.2 mRNA stability and neuronal disease

Aberrant regulation of mRNA decay is associated with human neurodegenerative and developmental disorders including Alzheimer’s disease (AD) [68, 69], Amyotrophic Lateral Sclerosis (ALS) [70], Fragile-X syndrome [71], and intellectual disability [72]. Altered miRNA expression profiles are strongly associated with AD, serving as putative biomarkers for the disease [73, 74]. Misregulation of specific neural miRNA networks are linked to amyloid beta [75–77] and tau [78, 79] pathologies leading to neurodegeneration associated with AD. While the NMD pathway is important in regulating RBPs associated with RNA and protein dysfunction in ALS [80, 81]. These include the autoregulated nuclear protein, TDP-43, encoded by *TARDBP*, which is normally degraded by NMD in the cytoplasm of spinal motor neurons via alternative polyadenylation and cleavage [82]. Mutations including copy number variants in multiple genes encoding regulators of NMD (i.e. *UPF2*, *UPF3B*, *UPF3A*, *SMG6*, *EIF4A3* and *RNPS1*) are also associated with patients exhibiting various forms of intellectual disability [83, 84]. Fragile-X mental retardation protein (FMRP) is a well-studied RPB that regulates synaptic mRNA localization, translation and stability [85, 86]. Misregulation of FMRP-directed RNA decay is implicated in Fragile-X syndrome, the most common form of inherited mental retardation [87]. Human neurodegenerative and developmental diseases remain challenging to understand and treat. Misregulation of mRNA decay is an important mechanism contributing to the pathology of these complex neurological diseases. Establishing global *in vivo* models of post-transcriptional regulation via mRNA decay will help to define the functions of mRNA decay regulatory networks and establish novel regulatory mechanisms. A systemic understanding of the regulation of mRNA stability in the nervous system will further our understanding of both normal physiology and disease.

1.2 Approaches for monitoring mRNA decay

Monitoring post-transcriptional control of gene expression requires the development of new tools to gather data akin to the types of transcriptional level data that are now widely available. Generating quantitative tissue-specific *in vivo* RNA decay models within the context of normal physiology remains challenging. Addressing this problem requires development and integration of methods to track RNA populations over time with spatial specificity. The studies establishing methods to monitor transcript decay kinetics in eukaryotic cells relied on transcriptional inhibition via Actinomycin D (ActD) [50,88,89] or inducible RNA polymerase II mutants [15,90,91] to monitor transcripts over time. ActD treatment is a simple approach that does not require exogenous genetic manipulation, however there are significant effects on cellular physiology including mRNA stability [88,92,93]. RNA polymerase II shut-off offers advantages over ActD, but this approach is limited to monitoring a steady-state population of RNA [94]. Transcriptional pulse-chase labeling approaches provide the best means to track a synchronized population of RNA over time and have been applied to monitor global mRNA decay in eukaryotic cells [5,51,95]. Traditional technologies for measuring RNA decay require cell sorting or manual dissection to achieve spatial specificity and are not suitable for *in vivo* tissue-specific measurements. A variation of *in vivo* biosynthetic tagging methodology, 4-thiouracil (TU)-tagging [96], overcomes these technical challenges. TU-tagging uses the uracil phosphoribosyltransferase (*UPRT*) gene of the protozoan parasite *Toxoplasma gondii* to convert the modified uracil, 4-thiouracil (TU), into the nucleotide form, 4-thiouridine-monophosphate, for subsequent incorporation into newly synthesized RNA. Multicellular eukaryotes have limited endogenous UPRT activity, thus targeted expression of UPRT allows cell type-specific tagging of mRNAs with TU at defined times during development. TU-tagging can be modified to allow measurement of mRNA decay, termed *TU-decay*, by following an initial TU-pulse with exposure to a molar excess of uridine [5,97]. Following TU-tagging, RNA from a mixture of cells or a whole organism can be reacted with a thio-specific biotinylation reagent and mRNA from the cells of interest

can then be purified using streptavidin-magnetic beads [5, 96, 98]. RNA isolated using this technique is suitable for quantitation by conventional strategies including qRT-PCR, microarrays and RNA-seq.

1.2.1 Modeling mRNA decay

Analyzing mRNA decay dynamics in the context of development is analogous to studying gene or protein abundance, but requires efficient methods to summarize this complex process. Although theoretically an mRNA can become vulnerable to decay via decapping, deadenylation or endonucleolytic attack, most eukaryotic degradation mechanisms rely on deadenylation as the first and rate-limiting step [99]. In bulk cellular mRNA decay, deadenylation is followed by decapping resulting in co-translational 5' $\rightarrow$ 3' exonucleolytic degradation [14], although many routes of mRNA decay are shown to be deadenylation-dependent (e.g. NMD [100, 101], miRNA [102, 103], ARE-mediated [104–106]). In metazoans, deadenylation is shown to exhibit biphasic kinetics. Initially deadenylases display distributive kinetics, hydrolyzing relatively few nucleotides before dissociating and repeating the process, while in the second phase, deadenylation becomes more progressive, with hydrolyzation of a majority of the poly(A) tail before dissociation [107, 108]. These two phases are also characterized by poly(A) tail length and susceptibility of the mRNA body to decay; in the first phase, poly(A) tails show more uniform distribution around 110 bases and decay of the mRNA body is not observed, while in the second phase, tails can range widely from 110 to 10 bases and the mRNA body is susceptible to degradation [102, 108]. In this way, stable mRNAs exhibit slow kinetics, characterized by the first phase of distributive deadenylation, while unstable mRNAs exhibit rapid kinetics, characterized by the second phase of progressive deadenylation [106]. Owing to the complex multi-step process of mRNA decay, recent work has attempted to recapitulate this using a Markov chain model in order to derive age-dependent decay rates and a residual lifetime of an mRNA [109]. This work allows classification of mRNAs into distinct groups characterized by their kinetics: those with slow initial kinetics where decay become more efficient over time, those with fast initial kinetics where decay becomes

less efficient over time, and those with more uniform decay kinetics over time, which follow first order (exponential) kinetics. While more complex models can recapitulate the multiphasic nature of mRNA decay, these models necessitate higher temporal resolution and more complex summarization (i.e. rate dependent on time) than simpler, exponential models. Thus, summarization of exponential decay models using a half-life measure is a convenient means of modeling decay, and owing to the largely deadenylation-dependent rate-limiting step of bulk mRNA decay, these models are sufficient to categorize mRNAs based on their decay kinetics.

Relative transcript quantities of TU-tagged RNA (a synchronized population) are ideal to model transcript decay over time. Decay of this synchronized population of transcripts is modeled using first order kinetics, shown in Equation (1.1), where k_d is the decay rate constant:

$$d[RNA]/dt = -k_d[RNA] \quad (1.1)$$

It is useful to summarize transcript decay using a half-life measure, calculated by deriving the equation (1.2), and substituting 1/2 as the proportion of the present RNA concentration $[RNA]$ to the initial RNA concentration $[RNA_0]$ (1.3):

$$\ln(RNA/RNA_0) = -k_d t \quad (1.2)$$

$$\ln 1/2 = -k_d t_{1/2} \quad (1.3)$$

Thus, half-life $t_{1/2}$ is inversely proportional to the decay rate constant k_d (1.4):

$$t_{1/2} = \ln 2/k_d \quad (1.4)$$

1.3 Objective of the Study

This work addresses the hypothesis that neural-specific regulation of mRNA decay is an important part of a dynamic gene expression program required for development. Recent advances in biosynthetic tagging of RNA in intact organisms coupled with high-throughput gene expression technologies like RNA-sequencing have allowed

genome-wide quantitation of mRNA decay. Integration of novel mRNA decay models and existing gene annotation information will be the basis for establishing neural-specific patterns of the regulation of mRNA stability. The specific aims of this project are:

Aim 1: Establish genome-wide *Drosophila* whole embryo and neural-specific mRNA decay models

Aim 2: Identify intrinsic elements that direct neural-specific mRNA decay dynamics in *Drosophila* embryos through integration of novel mRNA decay and existing *cis*-element and gene structure information

Aim 3: Develop methods for modeling mRNA decay using TU-decay and transcriptome-wide RNA-seq based measurements

Chapter 2

In vivo mRNA decay dynamics

2.1 Introduction

Interest in post-transcriptional control of gene expression in the context of neural development is growing and has implications for neurodevelopmental and degenerative diseases alike [69–71, 83]. However, global analysis of neural mRNA decay requires the ability to track a synchronized population of mRNA over time with cell-type specificity. The *TU-decay* technique was established to model global transcript decay kinetics in *Drosophila* embryos using microarrays [7]. TU-decay offers temporal and spatial specificity for measuring global transcript decay *in vivo* and does not rely on transcriptional inhibition or manual separation of tissues. TU-decay is based on thiouracil tagging (TU-tagging) of nascent RNA transcripts in tissues expressing a protozoan uracil salvage enzyme, uracil phosphoribosyltransferase (UPRT). In *Drosophila*, cell type-specific expression of UPRT is easily accomplished using the *GAL4/UAS* expression system [110]. By restricting UPRT expression and exposure to thiouracil, a synchronized population of transcripts can be followed at a precise developmental time frame.

Work from *Burow et al.* [7] is referenced in the following results (included in Figures 2.1, 2.2, 2.3, 2.4B,C, 2.5A,B,D, 2.7A,B, Tables 2.1, 2.2, 2.3). Maxine Umeh completed all the ActD and whole embryo TU-decay experiments and initiated the *pros > UPRT* neural-specific TU-decay experiments (included in Figures 2.1, 2.2,

2.3, Tables 2.1, 2.2, 2.3).

2.2 Results

2.2.1 TU-decay measures tissue-specific mRNA decay

TU-decay is a pulse-chase assay developed to monitor zygotic mRNA decay in intact *Drosophila* embryos. The nucleoside 4-thiouridine (TUd) is incorporated into nascent mRNAs independent of UPRT activity, allowing tagging of a synchronized population of mRNAs across all embryonic tissues. Using this approach, stage 12 to 15 embryos were incubated with 1mM TUd for a 1-hour pulse then transferred to another media containing 10mM uridine for a 1 hour or 3 hour chase period. TU-tagged RNA (TU-RNA) samples were purified after the 1-hour pulse, 1-hour chase, and 3-hour chase. To rule out bias in decay measurements due to tagged uracil recycling or incomplete uracil saturation in the chase, parallel chase experiments were performed on embryos using ActD. TU-RNA abundance was measured using microarrays and mRNA half-life was calculated by fitting these measures to the linear transform of a first order exponential decay model ($\ln(E(t)) = E_0 - kt$). The majority of mRNAs from the preliminary uridine plus ActD experiments fit this decay model with $R^2 \geq 0.75$. mRNA half-life distributions for experiments performed using uridine-only or uridine plus ActD were similar and a majority of mRNAs have reproducible half-lives, including cases of known high-stability mRNAs (i.e. *eIF-4a* [111]) and known low-stability mRNAs (i.e. *Myc* [112]) (Figure 2.1A,B).

Further analysis through gene ontology (GO) classification of high and low stability mRNAs confirms the reproducibility of the TU-decay (uridine-only chase) data compared to the conventional approach using ActD transcriptional shut-off and similar studies conducted in other models [113, 114]. GO classes enriched among high-stability mRNAs are those involved in core cellular functions including: ribosomal proteins, cytoskeletal proteins, chromatin-associated proteins, and electron transport proteins. GO classes enriched among low-stability mRNAs are those involved in more transient processes including: transcription factors, kinases, and regulators of specific

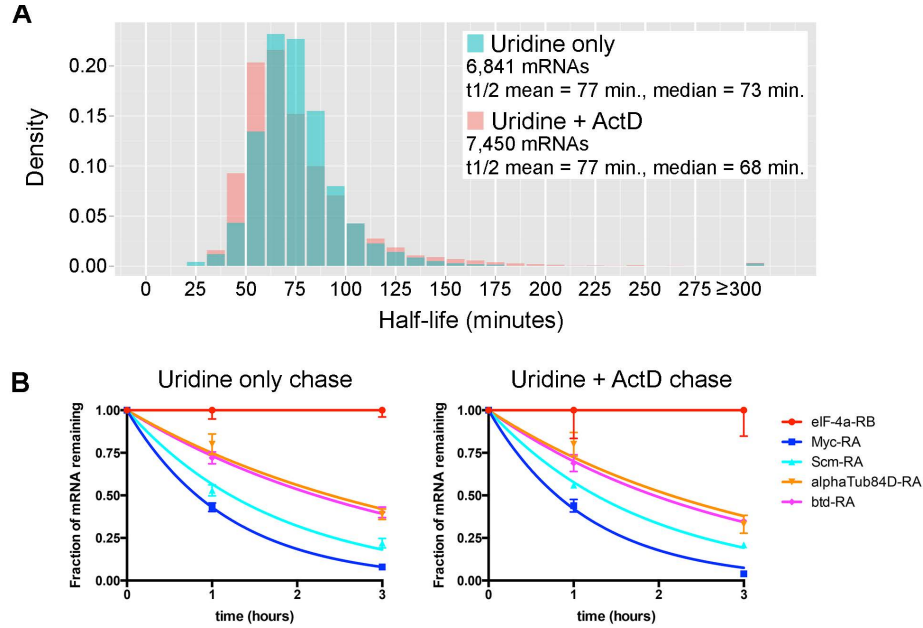


Figure 2.1: (A) Distribution of whole embryo half-life measurements between uridine-only and uridine plus ActD chase conditions. (B) Transcript decay curves between candidate high and low-stability mRNAs, showing mean and standard deviation of measurements.

developmental precesses (e.g. neuron morphogenesis).

To validate the TU-tagging technique for measurement of neural-specific mRNA abundance in *Drosophila* embryos, the expression of *UAS-UPRT* was driven by *prospero-GAL4* (Figure 2.2A). Prospero is a well-characterized neural transcription factor involved in establishing cell polarity during asymmetric cell division and was previously used to identify novel neural-expressed genes in *Drosophila* embryos [115]. *prospero-GAL4* drives expression of *GAL4* in neural stem and progenitor cells (neuroblasts and ganglion mother cells), glia of the central nervous system (CNS), and sensory neurons. Microarray comparison of *prospero-GAL4;UAS-UPRT* TU-tagged mRNA to untagged mRNA was the basis for establishing neural cell-type specificity for TU-decay. The majority (88%) of TU-tagged mRNAs enriched 1.5-fold or greater in the *prospero-GAL4;UAS-UPRT* microarrays were verified for neural expression patterns via published data from the Berkeley *Drosophila* Genome Project (BDGP) or *in situ* hybridization (Figure 2.2B). Comparison of pulse labeled TU-RNA from

Whole embryo measurements: high-stability mRNAs					
GO Term	Definition	P value	FDR	Count	Fold-enrichment
022626	Cytosolic ribosome <i>RpL4, RpL23, RpS16, RpS24</i>	2×10^{-57}	5×10^{-56}	67	10.3
000022	Microtubule cytoskeleton <i>Tub84b, Tub56D, ssp4, robl</i>	8×10^{-17}	3×10^{-14}	57	3.5
022900	Electron transport chain <i>mt:Cyt-b, mt:CoI, mt:CoII, mt:ND1</i>	1×10^{-14}	1×10^{-12}	29	5.9
0031497	Chroamtin assembly <i>His1, His2A, His2Av, His2B, His3</i>	2×10^{-9}	2×10^{-7}	15	7.8
0045941	Positive regulation of transcription <i>DSIF, HP1, Trl, ash2, lid, arm, btd</i>	4×10^{-3}	0.14	15	2.4
0048024	Regulation of mRNA splicing <i>Hrb27c, Hrb87f, B52, SC35, snf</i>	7×10^{-3}	0.19	10	2.9
Whole embryo measurements: low-stability mRNAs					
GO Term	Definition	P value	FDR	Count	Fold-enrichment
0004672	Protein kinase activity <i>S6kII, bubR1, mkk4, dTOR, wts, htl</i>	2×10^{-9}	1×10^{-6}	41	2.6
0048812	Neuron projection morphogenesis <i>Con, trio, Ephrin, Eph, otk, drl, rl</i>	2×10^{-8}	2×10^{-6}	40	2.7
0048812	Transcription <i>svp, dpn, Myc, grn, kn, Trx, eve</i>	1×10^{-5}	4×10^{-4}	47	1.8
0019226	Transmission of nerve impulse <i>CASK, cha, synaptogyrin, cpx, shi</i>	3×10^{-5}	9×10^{-4}	26	2.6
0019226	Striated muscle tissue development <i>Hem, Mef2, Mlp84B, flw, mbc</i>	3×10^{-5}	1×10^{-3}	11	5.3
0051674	Localization of cell <i>cbl, APC, Dl, Mhc, baz, dome, cv-c</i>	5×10^{-4}	0.01	23	2.3

Table 2.1: Gene ontology analysis of high and low stability mRNAs in whole embryos by DAVID.

the nervous system ($pros > UPRT$) and whole embryo shows significant depletion for many mesoderm associated genes and moderate but significant enrichment of nervous system mRNAs (Figure 2.2C). Further gene ontology analysis of mRNAs absent in nervous system measurements reveals significant depletion of transcripts involved in tissue specification outside the nervous system (Figure 2.2D). Given the specificity of

pros > *UPRT* tagging of neural transcribed mRNAs, this system was implemented for measuring neural cell-type specific mRNA decay in *Drosophila* embryos.

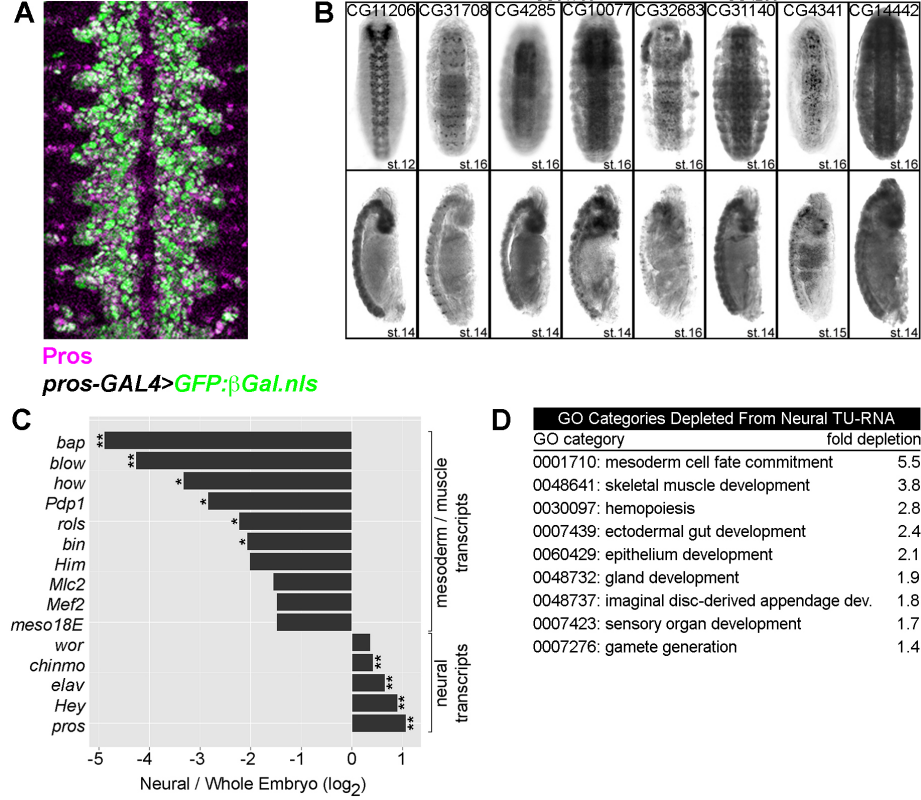


Figure 2.2: Immunohistochemistry of ventral *prospero-GAL4>UAS::lac.nls* embryo view showing *prospero-GAL4* expression pattern (A). *In situ* hybridization for 8 neural-enriched transcripts identified by *prospero-GAL4>UAS-UPRT* TU-tagging showing ventral (top panel) and lateral (bottom panel) views (B). Comparison of neural and whole embryo pulse-labeled mRNA shows depletion of muscle/mesoderm-specific transcripts and enrichment of neural-specific transcripts (C) ** = $FDR \leq 0.020$, * = $FDR \leq 0.040$ by SAM analysis. Gene enrichment analysis of transcripts below detection in neural pulse-labeled mRNA by DAVID analysis (D) $P < 0.0001$.

2.2.2 Neural-specific decay dynamics

Neural-specific TU-decay was performed by TU pulse-chase labeling of stage 12 to 15 *prospero-GAL4;UAS-UPRT* embryos and analyzed as previously described for whole

embryos. Neural cell-type specific half-life measurements were similarly reproducible to whole embryo measurements (having only 13% of mRNAs with standard deviation $\geq 30\%$) and fit first-order decay kinetics well with more than 73% of mRNAs having $R^2 \geq 0.75$. Half-life distributions show similar median half-life, while whole embryo mean half-life is slightly greater, due to the presence of more high-stability mRNAs (Figure 2.3A). Analysis of classes of mRNAs enriched in the top 500 high-stability or bottom 500 low-stability nervous system measurements show distinct differences from the whole embryo high-stability and low-stability sets (Table 2.2). Indeed mRNAs involved in transient processes including, transcription, cell-cycle and neuronal cell fate regulation are enriched among the nervous system low-stability set. Rapid turnover of these low-stability classes of mRNAs could be tied to tight temporal regulation of protein expression required for these transient processes. Notably, there is also enrichment of ribosomal protein mRNAs in this nervous-system low-stability set, which is contrasting to observations in whole embryos. Among the nervous system high-stability set of mRNAs there is distinct enrichment of classes related to neural development processes including: neuron projection, signal transduction, apoptosis, neurotransmission and cellular localization. Many high-stability nervous system mRNAs, especially those related to neuron morphogenesis, show inverse decay kinetics in whole embryos, suggesting a distinct program of nervous system transcript decay dynamics.

To assess differential transcript decay dynamics during neural development, comparisons of mRNA half-lives between nervous system (neural)-specific and whole embryo decay measurements were performed. A large proportion of mRNAs (24%) had statistically significant differences in half-life of ≥ 1.5 -fold between neural and whole embryo measurements. Of the set of 3,776 mRNAs with neural and whole embryo half-life measurements, 484 show decreased stability in the nervous system and 409 show increased stability in the nervous system (Figure 2.3B). This 24% of transcripts showing differential decay is striking considering other studies assessing mRNA decay involved in regulating neural development (via the NMD pathway [55, 116]), show far fewer transcripts affected (generally 10%) [117]. Though it is important to note that these studies rely on measurements of steady state mRNA levels when the NMD

Neural-specific embryo measurements: high-stability mRNAs					
GO Term	Definition	<i>P</i> value	FDR	Count	Fold-enrichment
0048812	Neuron projection morphogenesis <i>trc, fusl, Fas2, Appl, Fmr1, otk, ben</i>	5×10^{-6}	6×10^{-4}	33	2.4
0003924	GTPase activity <i>Rab23, Rab30, Rab11, Rho1, Ras64B</i>	2×10^{-4}	0.02	17	2.9
0012501	Programmed cell death <i>grim, pten, scyl, egr, mod</i>	2×10^{-3}	0.06	17	2.4
0004672	Protein kinase activity <i>CamK1, Cdk4, Ror, ald, aur, nmo, trbl</i>	9×10^{-3}	0.24	23	1.8
0019226	Transmission of nerve impulse <i>CASK, CanB, Dap160, cpx, Sap47</i>	0.02	0.25	17	1.8
0051674	Localization of cell <i>cbl, APC, cortactin, par-6, baz, spri</i>	0.01	0.16	18	1.9
Neural-specific embryo measurements: low-stability mRNAs					
GO Term	Definition	<i>P</i> value	FDR	Count	Fold-enrichment
0006351	Transcription <i>svp, dpn, pros, vvl, grh, gsb, onecut</i>	4×10^{-10}	2×10^{-7}	58	2.5
0030182	Neuron differentiation <i>chinmo, Dl, tup, Ephrin, Eph, fax, shot</i>	2×10^{-7}	4×10^{-5}	48	2.3
0007067	Mitosis <i>cycB3, cdc27, Sse, cnn, glu, mts</i>	6×10^{-5}	3×10^{-3}	21	2.9
0018024	Histone methyltransferase activity <i>G9a, Set1, Set2, Su(z)12, ash1, egg</i>	2×10^{-4}	0.02	7	6.9
0006897	Endocytosis <i>Arpc4, shi, CalpA, pat1, faf, scb, Rac1</i>	4×10^{-3}	0.08	25	1.9
0005840	Ribosome <i>RpL17, RpL35, mRpL11, mRpS35</i>	2×10^{-3}	0.06	21	2.0

Table 2.2: Gene ontology analysis of high and low stability mRNAs in the embryonic nervous system by DAVID.

pathway is perturbed, which likely underrepresent the proportion of transcripts affected by tissue-specific decay dynamics. Studies directly assaying mRNA decay upon

cellular differentiation provide evidence and possible mechanisms of differential transcript decay, but do not provide an *in vivo* context for development. Thus, this study provides for the first time, global data on tissue-specific mRNA decay during *in vivo* development.

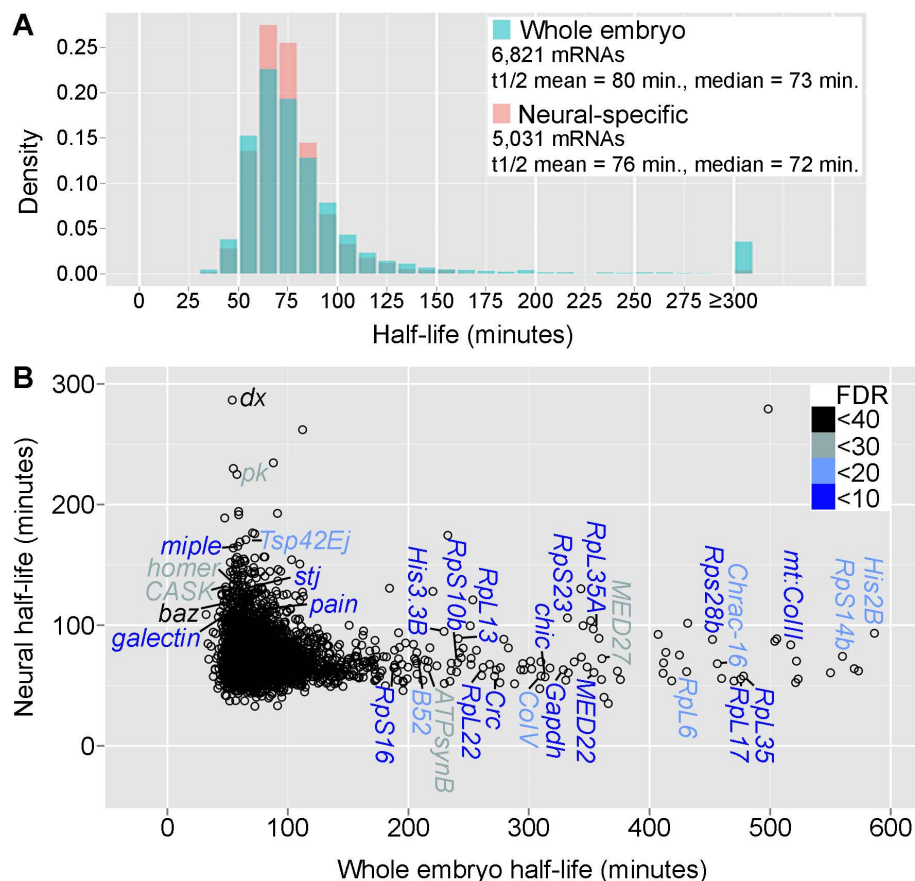


Figure 2.3: Distributions of mRNA half-lives for neural and whole embryo measurements (A). Comparison of mean half-lives for individual mRNAs in the nervous system and whole embryos (B). mRNAs with 1.5-fold half-life differences are indicated with the corresponding FDR.

To define the relationship between neural-specific decay and gene function, gene ontology analysis was performed. Functional annotation of mRNAs that exhibit stabilization or destabilization in the nervous system compared to whole embryo strongly suggest a distinct program of neural-specific transcript decay related to gene function (Table 2.3). Transcripts encoding proteins involved in translation (ribosomal

proteins and translation regulators), DNA replication (cell cycle regulators and histones) and cellular metabolism (proteasome and oxidative phosphorylation) were among the most enriched classes of mRNAs exhibiting destabilization in the nervous system, while transcripts encoding proteins involved in neural development functions (including neurotransmission, axonogenesis, axonal transport, and maintenance of neuroblast polarity) were among the most highly enriched classes of mRNAs exhibiting stabilization in the nervous system. Members of these enriched classes of stabilized or destabilized transcripts are expressed in tissues other than the nervous system (based on expression data available on multiple databases: FlyBase, modENCODE, BDGP *in situ*), corroborating a role for neural-specific regulation of transcript decay and not simply differential transcription. *In vitro* studies of mRNA decay in mammalian cell lines support a role for differential decay in cellular differentiation. Studies of mRNA decay in human induced pluripotent stem (iPS) cells [118] and mouse embryonic stem (ES) cells [119] show similar stabilization of protein synthesis, core metabolic and structure-associated genes in pluripotent cell types. Analysis of transcript decay rates in mouse ES cells following a crude differentiation protocol show similar trends, but most notably, a destabilization of transcripts encoding ribosomal proteins [119]. Analysis of differential decay dynamics between human iPS cells and their fully differentiated counterparts (human foreskin fibroblasts (HFF)) shows stabilization of mRNAs encoding chromatin and nucleosome-associated proteins (most notably histone proteins), and master transcriptional regulators of pluripotency (i.e. zinc-finger proteins), in the undifferentiated iPS cells. The authors of the iPS cell study [118] sought to define mechanisms of differential decay for stabilized (e.g. zinc-finger proteins) and destabilized mRNAs in iPS cells. They demonstrate that the iPS cell stabilized zinc-finger protein encoding mRNAs are largely targeted by *trans*-acting regulatory miRNAs, which were much more abundant in the differentiated HFF cells compared to iPS cells. Destabilized transcripts in iPS cells showed an enrichment of C-rich binding elements, and while an abundance of *trans*-acting poly(C)-binding protein-4 (PCBP4) in the differentiated HFF cells may explain these data, other PCBPs showed disparate (PCBP3) or similar expression patterns (PCBP2 and PCBP3) between cell types. In order to better understand the mechanisms by

which transcript decay is differentially regulated during development, it will be important to understand the contribution of both regulatory *trans*-factor expression and intrinsic features in determining tissue-specific decay dynamics.

mRNAs stabilized in the nervous system					
GO Term	Definition	<i>P</i> value	FDR	Count	Fold-enrichment
007409	Axonogenesis <i>APC, CadN, trio, caps, fray, otk, daw</i>	2×10^{-3}	0.06	13	2.8
003924	GTPase activity <i>Rab21, Rab30, Ras64B, RabX1</i>	2×10^{-3}	0.31	10	3.5
0009321	Oligosaccharide biosynthesis <i>mgat2, GalNAc-T1, pgant2, pgant5</i>	9×10^{-3}	0.15	4	9.2
0045196	Establishment of neuroblast polarity <i>par-6, baz, Lgl</i>	0.01	0.16	3	18.7
0007269	Neurotransmitter secretion <i>CanB, Lgl, CASK, Dap160, cpx, rop</i>	0.04	0.56	6	2.6
mRNAs destabilized in the nervous system					
GO Term	Definition	<i>P</i> value	FDR	Count	Fold-enrichment
0003735	Structural constituent of ribosome <i>RpL11, RpL30, RpL6, RpS16, RpS23</i>	5×10^{-17}	5×10^{-15}	26	9.2
0000278	Mitotic cell cycle <i>dap, cycB3, ball, glt, eff, ncd, mts</i>	4×10^{-8}	4×10^{-5}	30	3.2
0031497	Chromatin assembly <i>H2B, H3, H3.3A/B, HP2, Df31, Set</i>	3×10^{-4}	0.02	7	7.5
0005839	Proteasome core complex <i>Pros26, Prosα7, Prosβ2, Prosβ7</i>	8×10^{-3}	0.07	5	6.0
0006119	Oxidative phosphorylation <i>ATPsynB, ATPsynD, mt:ColIII</i>	3×10^{-4}	0.03	12	3.5

Table 2.3: Gene ontology analysis of stabilized (neural/whole embryo half-life > 1.5) and destabilized (neural/whole embryo half-life < 0.06) mRNAs in the embryonic nervous system by DAVID.

2.2.3 Neural mRNA stability is associated with gene function

First, it is important to assess the relationship between gene function and neural-specific decay dynamics. There is evidence of coordinated gene organization and expression dynamics for genes with related functions in eukaryotes like *Drosophila* [120–122], and whole embryo decay data suggest similar coordination at the level of mRNA decay. It is plausible that genes with coordinated functions in neural cell types would exhibit neural-specific decay kinetics. For example, neural cell fate associated mRNAs that are transiently expressed will be rapidly decayed, while locally translated mRNAs that require stabilization for transport to distal sites, will be stabilized in neural tissues. To address the hypothesis that regulators of neural cell fate will show rapid decay, 39 mRNAs encoding transcription factors involved in neurogenesis were identified based on Interactive Fly annotations [123] and to test the hypothesis that locally translated transcripts will exhibit stabilization, 36 *Drosophila* orthologs of mRNAs localized to axon growth cones [124] or dendrites [125] in mammals were identified. Comparison of half-lives between the designated neural fate, distally localized, and all neural mRNAs show significant differences in stability between all groups ($P < 0.0001, q < 0.0001$); with neural fate mRNAs showing collective decreased stability and distally localized mRNAs showing collective increased stability (Figure 2.4A). Rapidly degraded neural fate encoding transcription factor mRNAs have documented dynamic temporal and spatial patterns of protein expression including: those whose expression is coupled to the timing of neuroblast cell divisions (*cas*, *svp*, *pdm-1*) [126], those that are differentially expressed within neuroblast lineages (*Hey*, *Dl*, *ind*, *vnd*, *Dr*, *btd*, *ems*, *chinmo*, *ap*, *SoxN* [127–130]), those that are differentially expressed among neuroblasts, GMCs, and neurons (*dpn*, *pros* [131]), and those whose expression is coupled to temporally regulated neuronal maturation (e.g. regulation of synapse formation *Dl*, *pb*, *en* [132]) (Figure 2.4B). Among the class of distally localized stable mRNAs, is a known dendrite-localized mRNA (*CaMKII* [133]) and growth cone-localized (pre-synaptic) mRNAs (*synaptogyrin* and *Ef1 α 100E* [132, 134]) in *Drosophila*, and orthologs of mammalian mRNAs that are

locally translated in dendrites (*Arc1* [67] and *Ef1 α 100E* [135]) (Figure 2.4C). Other examples of distally localized mRNAs at synaptic sites in *Drosophila* neurons include *homer* [136] and *CASK* [137], although conclusive evidence of their site of translation is unknown. The increased stability of these putative distally translated mRNAs may be related to their subcellular patterns of expression, although direct evidence is lacking.

2.2.4 Patterns of regulation related to neural mRNA decay dynamics

Differences in global patterns of polyadenylation in *Drosophila* can yield transcripts of varying lengths, which are often associated with tissue-specific gene expression. RNA-seq analysis of alternative polyadenylation in various dissected adult tissues revealed widespread lengthening of transcripts in the nervous system, particularly those encoding RBPs and transcription factors [65]. Another study using high-density tiling microarrays and RNA-seq showed similar, but progressive 3'UTR lengthening during embryonic development, with the longest UTRs present in neural-expressed transcripts [66]. Transcripts encoding RBPs were notable among the genes with long UTR isoforms and these sequences enriched in the Pumilio (Pum) recognition element [66]. Global analysis of transcript decay rates in differentiating mouse embryonic stem cells showed decreased stability among isoforms retaining introns and those containing mammalian orthologs of the highly-conserved *pumilio* family (*PUF*) recognition element (PRE) [119]. 3'UTR lengthening is also implicated in NMD-specific regulation of transcript decay in *Drosophila* S2 cells, especially among regulators of translation and cell division (i.e. RBPs, ribosomal proteins and cell cycle proteins) [138]. Alternative cleavage and polyadenylation resulting in long 3'UTR transcript isoforms is a conserved mechanism in mammals [139] and was recently attributed to directing transcript localization to distal sites within neurites *in vitro* [140]. Additionally, among transcripts with long 3'UTR isoforms, those with distinct distal localization patterns showed increased expression in conjunction with mammalian neural differentiation *in vitro* [140]. As a general mechanism, extended 3'UTR isoforms may allow

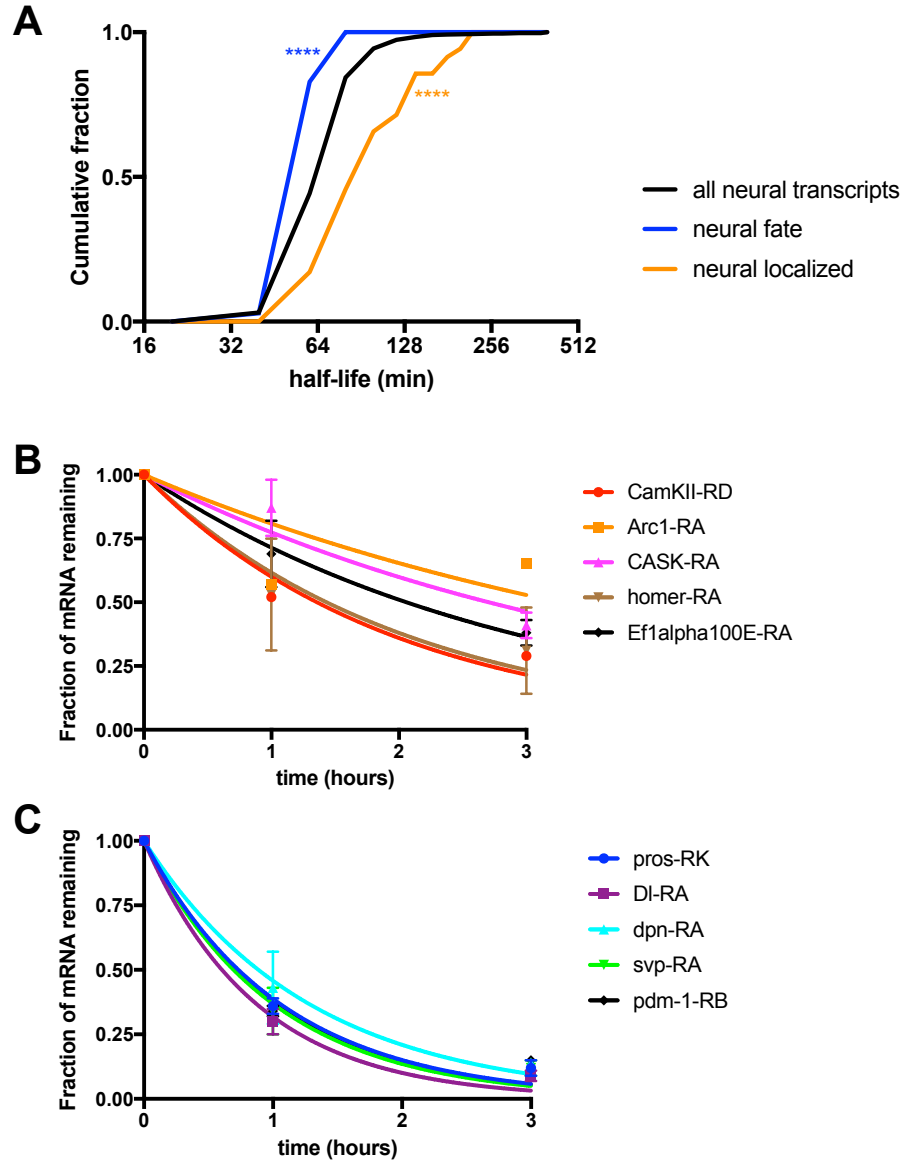


Figure 2.4: Cumulative distributions of mRNA half-lives for neural cell fate, neural localized and all neural transcripts, P -values are computed between all groups by Kruskal-Wallis test (A). Decay curves of select neural cell fate associated mRNAs (B) and select distally localized mRNAs (C).

for greater targeting by *trans*-acting regulatory factors whereby both *cis*-element and *trans*-factor expression is highly context-dependent.

Investigation of 3'UTR length and stability showed no direct correlation among distally localized transcripts, neural fate transcripts or the entire neural mRNA decay data (Figure A.1A,B). Assessing extended 3'UTR isoform-specific stability effects by these microarray-based decay data is challenging, owing to the only recent wealth of information on tissue-specific alternative cleavage and polyadenylation (based largely on RNA-sequencing approaches). The relationship between stability and 3'UTR characteristics are likely complex—relating to both function and context. Thus, candidate *trans*-acting factors that have documented roles in neural development (Pum [141], AREs [142]), neural-specific expression patterns (*miR-124* [143]), and available databases containing target information were selected for specific investigation. Analysis of enrichment or depletion of candidate *cis*-regulatory elements among the classes of stable distally localized mRNAs or unstable neural fate mRNAs revealed enrichment of PREs, defined by Gerber *et al.* (*UGUAHAUA*) [141], in the class of neural fate mRNAs. Of the neural fate mRNAs, 20 out of 36 transcripts with half-life ≤ 75 min contain a PRE, while there was a slight depletion from the set of high-stability distally localized mRNAs, although not at a statistically-significant level (1 out of 29 mRNAs with half-life ≥ 80 min contain a PRE) (Figure 2.5A). Nearly half (14 of 30) of the set of transcripts characterized as having neural 3'UTR extensions in *Drosophila* [65] also contained one or more PREs in the extended 3'UTR isoforms, which include transcriptional (Nej [144]) and post-transcriptional regulators of neural development (ELAV [145, 146], Imp [147], msi [148], Pum [149]). Analysis of the neural-specific mRNA decay data revealed 439 transcripts containing the PRE motif; however, this subset of PRE-containing mRNAs showed no significant trend toward low or high-stability classes (8.5% of mRNAs with half-life ≤ 75 min contain a PRE, and 8.8% of mRNAs with half-life ≥ 80 min contain a PRE). Gerber *et al.* also identified an additional 156 putative Pumilio targets by tandem-affinity purification (TAP)-tagged Pumilio assay [141]. These data did not include any transcripts from the distally localized or neural fate classes and did not show bias related to stability in the entire neural mRNA decay data (0.9% of mRNAs with half-life ≤ 75 min were predicted targets by TAP, and 1.5% of mRNAs with half-life ≥ 80 min were predicted targets by TAP). Taken together these results indicate that the presence of a PRE

motif is associated with decreased mRNA stability among the class of neural fate mRNAs.

2.2.5 *Pumilio* knockdown stabilizes neural PRE-containing mRNAs

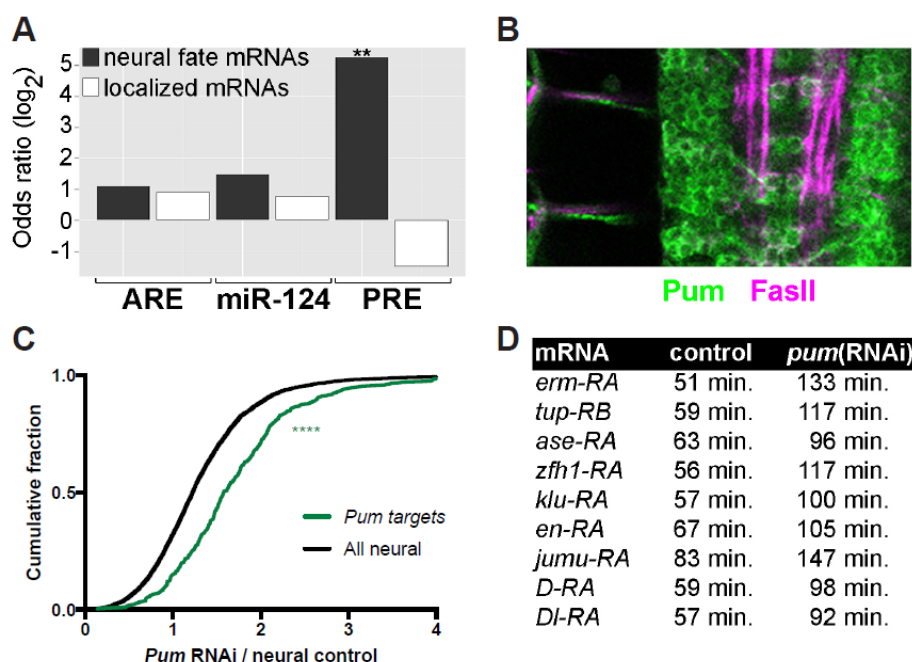


Figure 2.5: Enrichment or depletion of candidate *cis*-regulatory elements as follows: AREs, miR-124 target sites, and PREs. Odds ratios were calculated as the frequency of each element within the low-stability ($t_{1/2}$ 75min) neural fate mRNAs or high-stability ($t_{1/2} \geq 80$ min) distally localized mRNAs versus the frequency of that element in all low-stability or high-stability neural mRNAs (A) $**P < 1.0 \times 10^{-15}$ (Fisher's test). *Pum* expression in a stage 15 embryo detected with an anti-*Pum* antibody (green). Anti-Fas2 antibody stain (magenta) shows a subset of longitudinal axon fiber tracts and axons leaving the ventral nerve cord (B). Cumulative distribution of *pum*RNAi versus control mRNA half-lives in the embryonic nervous system (C) *P*-values are calculated by Kolmogorov-Smirnov test. Examples of predicted neural fate *Pum* targets with increased mRNA stability in *Pum*RNAi embryos (D), these transcripts had reproducible changes in half-life, $FDRs \leq 0.30$ by SAM analysis.

The significant enrichment of PREs among low-stability neural fate mRNAs

indicated a novel role for Pum in regulating neural development. Pum has well-documented roles in synapse form and function [149], and memory formation [150] in *Drosophila*. Investigation of Pum in the embryonic nervous system shows widespread localization to the cytoplasm of neural progenitors, neurons and specific axon projections throughout the ventral nerve cord (Figure 2.5B). To assess phenotypic defects in embryonic neural development, neural-specific RNA interference (RNAi) [151] was performed to reduce *Pum* expression using *pros-GAL4 x UAS-pumRNAi* embryos. Western blot analysis reveals an approximate 50% decrease in expression from wild-type levels (Figure A.2). *Pum* knockdown did not cause gross neural development defects (Figure 2.6A-E), based on analysis of *Fas2*⁺ axon projection patterns and production of *Eve*⁺ motor neurons [152]. Lack of obvious defects in motor neuron development in the embryonic nervous system may not preclude defects in other aspects of embryonic neural development or later stages of neural development, which were not evaluated in this study. Although partial knockdown of *pum* did not show obvious embryonic neural development defects, it should be sufficient to induce changes in mRNA stability if Pum regulates mRNA decay.

Neural-specific TU-decay and *pum* knockdown was performed by *pros-GAL4* driving expression of *UAS-pumRNAi;UAS-T.g.UPRT* in stage 12 to 15 embryos. Half-lives for neural-specific *pum* knockdown embryos were compared to control neural-specific half-lives (*pros > UPRT*). Analysis of 3,008 mRNAs with reproducible half-life measurements showed a significant trend of stabilization for the set of *Gerber et al.* [141], PRE-containing, Pum targets (Figure 2.5C). Half (132 of 265) of the predicted Pum targets from these data had statistically significant increases in mRNA half-life of ≥ 1.3 -fold. Many of the stabilized Pum targets were among the set of neural cell-fate transcription factors or other cell-fate determinants (Figure 2.5D). Stabilized Pum targets also include regulators of axon pathfinding (*Eph*, *otk*, *Sema-1a*) and putative distally localized, *dlg1*, which is a member of the membrane-associated guanylate kinase homologs (*MAGUK*). *MAGUK* proteins are important for proper synaptic architecture [153,154]. *Dlg1* mRNA was previously identified as a target of translation repression by Pum [155], and its mammalian ortholog is known to localized to dendrites [156]. *Dlg1* mRNA is relatively unstable in control embryos, but

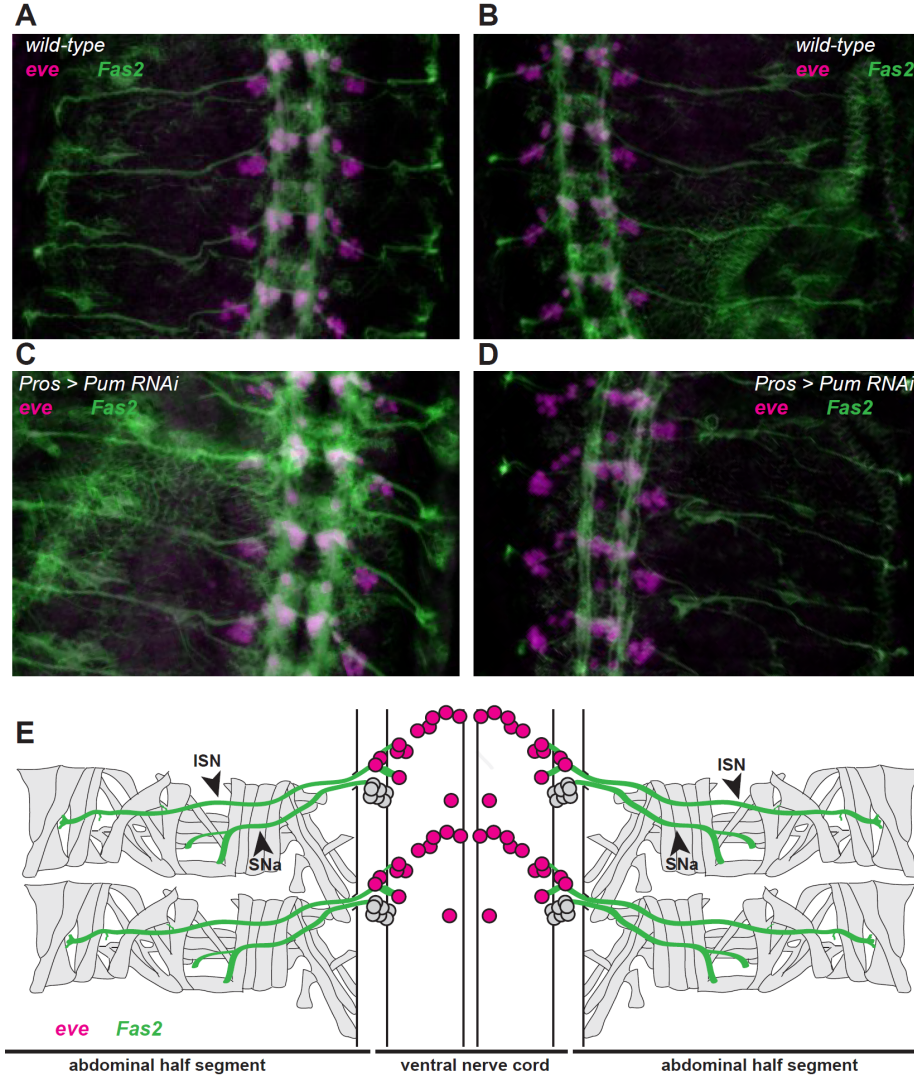


Figure 2.6: *Wild-type* (A), (B) and *pros-GAL4 x UAS-pumRNAi* (C), (D) stage 15/16 embryos stained with an anti-Fas2 antibody (green) and anti-eve antibody (magenta) shows eve⁺ motor neurons (magenta), longitudinal nerve fibers (green), and the intersegmental (ISN) and segmental (SNa) nerves (green) leaving the ventral nerve cord. A model (E) of the proper orientation of motor neurons and peripheral nerves (ISN and SNa) that can be visualized by anti-Fas2 antibody (green) and anti-eve antibody staining in stage 16 *Drosophila* embryo abdominal segments, based on Landgraf *et al.* [152].

is significantly stabilized in Pum RNAi embryos (half-life of 65 versus 127 min). Of the 14 mRNAs with neural-specific extended 3'UTR isoforms identified by Smibert

et al. [65], 9 exhibit stabilization upon *pum* knockdown (*wdb*, *heph*, *step*, *elav*, *nmo*, *Imp*, *msi*, *tyf*, *Glut4EF*). These novel Pum targets function in diverse processes related to neural development including: signaling (*Wdb*, *step*, *nmo*), RNA-binding and translation (*heph*, *Imp*, *ELAV*, *msi*, *tyf*). *Wdb* and *nmo* are required for planar cell polarization in epithelial cell types [157,158], while *Wdb* [159] and *Step* [160] are implicated in regulating the insulin signaling pathway that controls energy metabolism and growth. *Tyf* is a novel translation initiation factor functioning in the regulation of gene expression associated with circadian rhythms in pacemaker neurons [161]. These putative neural Pum targets are subject to neural-specific regulation by Pum via tissue-specific isoform expression; only the extended 3'UTR isoforms, which are expressed exclusively in neural tissues, contain PREs.

2.2.6 PRE 3'UTR deletion is sufficient to stabilize GFP reporter mRNAs

In order to determine if a 3'UTR PRE is necessary for destabilization of a putative Pum target in the embryonic nervous system, an *in vivo* green fluorescent protein (GFP) reporter system was constructed. Reporter transcripts were designed with a *hsp70* 5'UTR, GFP open reading-frame, and various 3'UTR sequences with endogenous polyadenylation signals; *RPS9* 3'UTR as an endogenous non-Pum target (no PRE), *Hey* 3'UTR as an endogenous Pum target (PRE), and *Hey* 3'UTR with a PRE 8-mer deletion as a mutant non-Pum target (Figure 2.7A). All constructs were integrated into the same genomic locus to generate transgenic fly lines. Neural-specific expression of these reporter mRNAs and TU-decay was achieved by *pros-GAL4* driving expression of *UAS-GFP3'UTR;UAS-UPRT* in stage 12 to 15 embryos. The reporter transcript containing the *RPS9* 3'UTR had similar though slightly above average stability, compared to the endogenous *RPS9* transcript (Figure 2.7B). The *Hey* 3'UTR reporter was more rapidly decayed, although similar to the endogenous *Hey* transcript. While deletion of the PRE from the *Hey* 3'UTR led to a striking stabilization of the reporter transcript. These results are compatible with findings in the *pum* RNAi decay data that showed increased stability of the endogenous *Hey* transcript

upon *pum* knockdown (110 min compared to 74 min). Taken together these data indicate that a PRE within the *Hey* 3'UTR is necessary to target this transcript for Pum-mediated degradation in the embryonic nervous system.

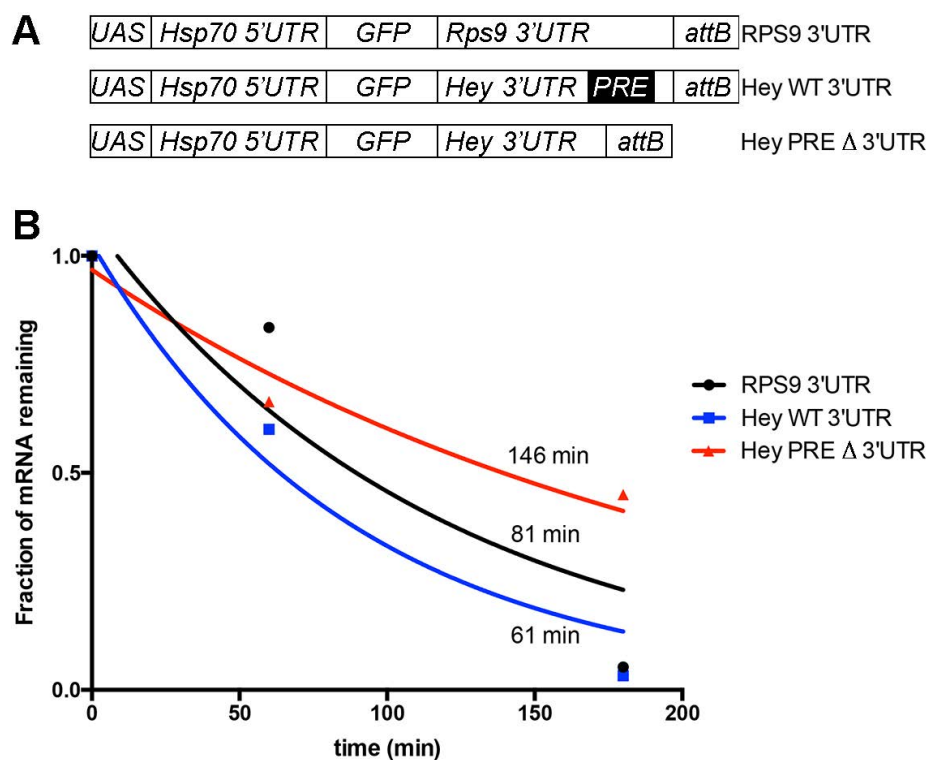


Figure 2.7: Reporter construct design. Hey WT 3'UTR contains the single endogenous PRE. Hey PRE Δ 3'UTR has only the PRE sequence removed. (A) Decay curves based on reverse-transcription, quantitative-PCR measurements of reporter transcript decay (using GFP primers). Chase times are plotted on the x-axis (0-hour = end of 1-hour pulse). Mean chase/pulse values at the 1-hour chase and 3-hour chase time points are plotted based on multiple measurements. (B) Half-lives in minutes are listed next to the decay curve for each transcript.

2.3 Discussion

Regulation of gene expression at the transcriptional and post-transcriptional level determine final protein abundance and localization. Post-transcriptional control of gene expression via mRNA decay has often been overlooked, but can provide information

on spatial and temporal dynamics of protein expression during development *in vivo*. The TU-decay technique is developed to measure global zygotic mRNA decay rates in *Drosophila* embryos and the embryonic nervous system via microarray analysis. Global data on mRNA decay rates in whole embryos reveals links between mRNA stability and gene function; highly stable genes are enriched in constitutively expressed housekeeping and core metabolic genes, while highly unstable genes are enriched in regulatory proteins, transcriptional regulators, and regulators of neurogenesis. Global data on mRNA decay rates in the embryonic nervous system reveals links between mRNA stability and gene function that are distinct from the whole embryo; stable genes are enriched in regulators of neurogenesis, regulatory proteins, and regulators of cell death, while unstable genes are enriched in transcriptional regulators, cell-cycle regulators, histones, and ribosomal proteins. These differences are the basis for investigating neural-specific mRNA decay dynamics. One-quarter of genes detected in whole embryos and the embryonic nervous system had differential decay kinetics (≥ 1.5 -fold differences in half-life) between tissues. The whole embryo context provides an average half-life measure across all tissues compared to the nervous system, which is more uniform. This is reflected in the global distributions of mRNA half-lives between tissues; while the average and median half-lives are similar, the whole embryo has a larger range, which may reflect the heterogeneous nature of these measurements. Neural-specific mRNA decay dynamics likely contribute to a finely-tuned program of gene expression driving neural development. This is evidenced—at least in part—by the observation that the Pumilio RBP regulates cohorts of neural development associated mRNAs (i.e. neural cell fate transcription factors and mRNAs with neural-specific extended 3'UTR isoforms) via PREs in the 3'UTR. Destabilization of cell-cycle regulators, histones, and ribosomal proteins and stabilization of synaptic, axonogenesis and neuroblast polarity associated genes in neural tissues signify a global shift in gene expression at the post-transcriptional level that may reflect a changing program from a more proliferative state to a differentiative state. *Prospero* (used in this study to drive neural-specific mRNA decay measurements) encodes a protein that regulates this shifting cellular program during neural development through its effects on cell cycle regulators [162]. Regulation by specific *trans*-acting factors, like

Pumilio, have the ability to target cohorts of transcripts in a context-dependent manner, but the underlying mechanisms driving large-scale changes in mRNA stability upon neural development in the *Drosophila* embryo remain unclear.

2.4 Approach

2.4.1 Fly lines

The following fly lines were used: *UAS-T.g.UPRT* (3rd chromosome), *UAS-T.g.UPRT* (2nd chromosome, provided by M. O'Connor), *prospero-GAL4* (Bloomington Drosophila Stock Center), FlyLight project line GMR45H06 (*lola(ubiquitous)-GAL4*) (Bloomington Drosophila Stock Center). Whole embryo TU-decay experiments were performed using embryos of either *UAS-T.g.UPRT* line without any *GAL4* driver. *Pumilio* knockdown was performed using the TRiP project line HMS01685 (Bloomington Drosophila Stock Center). *UAS-GFP-3'UTR* reporter constructs were made by modifying *pUAST-3XemGFP* (Drosophila Genetics Resource Center) to remove two copies of *emGFP* and replace the *SV40* terminator with the *Drosophila Hsp70* terminator and an *attB* site. *RPS9*, *Hey* wild-type or *Hey PRE deletion 3'UTR* DNA were chemically synthesized (Integrated DNA Technologies) and ligated into the modified *pUAST-emGFP-attB* backbone. Constructs were injected into *Pnos-phi-C31int.NLSX*, *PCaryPattP40* embryos for PhiC31- mediated integration (Rainbow Transgenic Flies, Inc.).

2.4.2 Embryo permeabilization and pulse-chase labeling

Stage 12 to 15 embryos were used for pulse-chase labeling of RNA. Embryos were permeabilized as previously described [96] and pulse-labeled in tagging media [D22 insect media, 5% FBS, and either 1 mM 4-thiouridine (whole embryo tagging) or 1 mM 4-thiouracil (neural-specific tagging)] for 1 h at 30°C. Pulse-labeled embryos were then transferred for 1 or 3 h to chase media containing uridine [D22 insect media, 5% FBS, 10 mM uridine] at 30°C. Embryos were homogenized for 30 sec and frozen in TRIzol® (ThermoFisher) at 80°C.

2.4.3 TU-RNA purification

Multiple 100 μ g samples for each time point were pooled for TRIzol[®] (ThermoFisher) RNA extraction and TU-RNA purification, as previously described [96].

2.4.4 cRNA probe synthesis and array hybridization

Biological replicates (independent RNA collections, TU-RNA purifications, and microarray hybridizations) were performed for all experiments. Purified TU-RNA plus a spike-in control (One-Color RNA Spike-In Kit, Agilent) was used to generate Cy3 labeled cRNA (Low Input Quick Amp Labeling Kit, Agilent) and hybridized to Agilent 4x44K *Drosophila* gene expression microarrays. Signal intensities within a pulse-chase series were normalized using the spike-in control. For comparisons between pulse-chase series, 1-hour chase/1-hour pulse and 3-hour chase/1-hour pulse ratios were normalized so that the mean of the ratios at each time-point was the same across experiments (the top and bottom 10% ratio values were excluded from normalization factor calculations). For comparisons of signal intensities in 1-hour pulse samples, spots with values below background were excluded from mean normalization calculations but were included in calculations of average signal intensity per gene.

2.4.5 mRNA half-life calculations

Normalized signal intensities for time 0, 1, and 3 h were fitted to the linear transform of a single order exponential decay model, $\ln((t)) =_0 kt$, by least squares regression analysis in R. E_0 is the transcript expression value for the pulse-labeled sample and k is the decay rate constant. Fit of the decay model was assessed by an adjusted R^2 value. Half-life was determined by the equation, $t_{1/2} = \ln 2/k$. Many mRNAs whose levels decreased slightly or not at all during the chase had R^2 0.75. These high-stability mRNAs (typically with half-lives estimated at ≥ 300 min) were excluded from calculations of mean and median half-life for all datasets. mRNAs in this category that had reproducible half-life measurements across biological replicates (standard deviation/mean $\leq 60\%$) were included as high-stability mRNAs for gene ontology analyses and comparisons of whole embryo versus neural mRNA half-lives.

2.4.6 Statistical analysis

Statistical analyses of genome-wide mRNA levels and genome-wide half-life values were made using mean-normalized ratio values for each dataset and absolute difference calculations using the significance analysis of microarrays (SAM) program [163]. Two-class, unpaired sample analysis and the T-statistic were used to calculate q values (a multiple testing corrected FDR). FDR values listed in figure legends are the q value (for statistical analyses performed using SAM). All other statistical analyses were made using the R environment. For comparisons of half-life values between samples (neural-specific versus whole embryo), only transcripts with reproducible measurements across biological replicates were included. Reproducible measurements were defined as those with standard deviation in half-life $\leq 60\%$ of the mean half-life.

2.4.7 Gene ontology analysis

Gene ontology analysis was performed using DAVID functional annotation clustering [164,165], with default settings and high classification stringency. Unnamed genes (those identified by CG number only) were excluded from GO analysis. Representative GO terms for each annotation cluster are listed in the results tables. Functionally redundant annotation clusters were excluded.

2.4.8 Immunofluorescent imaging

Embryo fixation and antibody staining were performed using standard methods. The following antibodies were used: rat anti-Pum (1:200, provided by Robin Wharton), mouse anti-Fas2 (1:100, Developmental Studies Hybridoma Bank), mouse anti-prospero (1:20, Developmental Studies Hybridoma Bank), rabbit anti- β -gal (1:1000, Abcam). RNA probes were made using cDNA clones obtained from the Drosophila Genome Resource Center, Indiana University, USA. Digested plasmids were used as template for in vitro transcription of DIG-labeled RNA probes (Roche Applied Science). *In situ* hybridization was performed using standard methods.

2.4.9 RT-qPCR

Purified TU-RNA quality was assessed by Bioanalyzer RNA Pico Eukaryotic total RNA assay (Agilent) prior to whole transcriptome amplification and cDNA synthesis (MessageBOOSTER Whole-Transcriptome cDNA Synthesis Kit for qPCR, Epicentre). Real-time PCR quantitation was performed on a Rotor-Gene Q (Qiagen) in 20 μ L reactions using QuantiTect Primer Assays (Qiagen) or custom designed primers (Table A.1), and SYBR green detection (QuantiTech SYBR green PCR kit, Qiagen). C_t values were normalized to an *RpL32* internal reference and relative abundance calculated by the equation, fold-change = $2^{\Delta\Delta C_t}$. Transcript decay was modeled as described above using relative abundance at time 0, 1 and 3 h.

Chapter 3

Codon-mediated mRNA decay dynamics

3.1 Introduction

Evaluation of the PUF family ortholog, Pumilio, demonstrates a mode of RBP-mediated regulation of a class of functionally related mRNAs in neural tissues of the developing *Drosophila* embryo. This approach evaluating candidate *trans*-acting factor regulation of target mRNA stability only accounts for differential decay of subsets of mRNAs, but fails to explain the bulk differences in mRNA stability seen between tissues. Roughly 25% of mRNAs show tissue-specific decay dynamics; house-keeping and constitutively expressed genes enriched among high-stability mRNAs in the whole embryo were not present in the high stability set of neural mRNAs, while ribosome and electron transport associated mRNAs, which are highly stable in the whole embryo, show high enrichment among mRNAs with the greatest destabilization in neural cell-types (Table 2.3). *In vitro* studies of mRNA decay dynamics highlight the relationship between transcript stability and gene function [118, 119], but argue intrinsic features of transcripts most strongly correlate with stability. Intrinsic features of mRNAs including 3'UTR length [64, 65, 119] and codon bias [15, 16] have demonstrated ability to modulate global decay dynamics, but these studies fail to address differential decay dynamics across different cellular contexts *in vivo*.

Differences in synonymous codon usage, termed *codon bias*, has long been studied in population genetics. Despite well-defined measures of codon bias across many species, most synonymous sites show weak selection [166], which can be explained largely by GC-content bias [167]. Selective pressures on codon bias can be due to translation efficiency and accuracy [168–172], though these effects are likely highly context-dependent [173]. Frequently used codons in highly expressed genes correlate well with measured tRNA levels in prokaryotes and tRNA gene copy numbers in yeast [173–177]. These observations support the explanation that codon usage bias is a result of selection for more efficient translation. It is plausible that codon bias in multicellular eukaryotes may be linked to context-specific expression. For example, highly expressed genes of a particular tissue would show a bias toward optimal codons, reflected in the availability of their cognate isoaccepting tRNAs in that tissue. However, evaluation of selective pressures on codon bias in *Drosophila* revealed similar bias in genes with disparate mRNA expression patterns [178]. There are correlations between codon usage bias and putative translation efficiency (via tRNA availability) [179] and accuracy [168] in *Drosophila*. The ties between mRNA decay and translation elongation are demonstrated by the discovery that mRNA decay occurs $5' \rightarrow 3'$ on active polysomes [14], and optimal codon content is a major determinant of transcript stability in yeast [15] and zebrafish embryos [16]. It is plausible that similar mechanisms driving codon-mediated mRNA decay and translation kinetics in other eukaryotes are present in *Drosophila*.

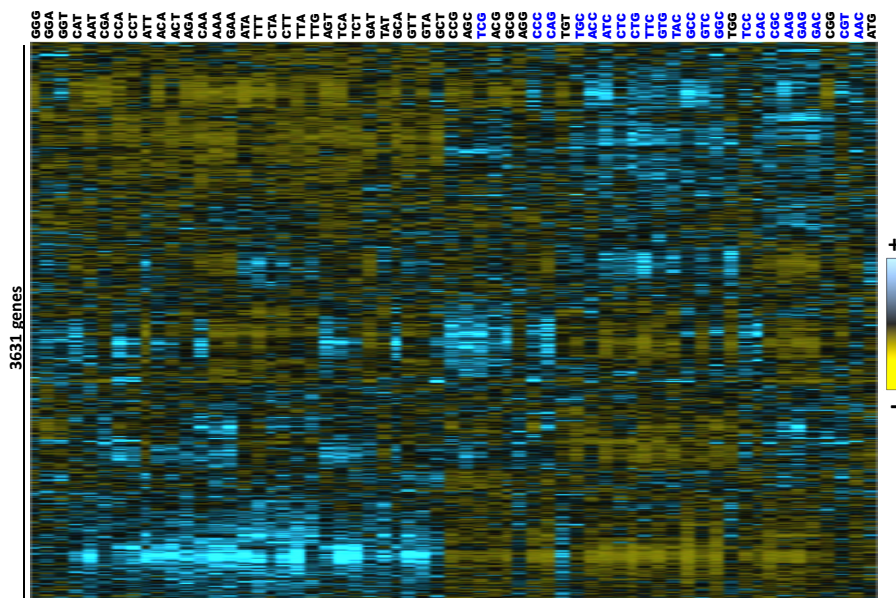
The codon optimality work is a collaboration with Jeff Collier’s group at Case Western Reserve University. Vladimir Presnyak and Sophie Martin calculated codon stabilization coefficients (CSC) for the 3631 mRNAs with reproducible half-life measurements for both whole embryos and the embryonic nervous system, and produced the heatmap used in Figure 3.1. The Collier group designed the mutant CDS and Najwa Al-Husaini performed ligation of the wild-type and mutant CDS into the modified *pUAST* backbone (see methods for detailed information).

3.2 Results

3.2.1 Codon optimality is correlated with mRNA stability in *Drosophila* embryos

Analysis of 3631 genes (the set of mRNAs with reproducible half-life measurements for both whole embryos and the embryonic nervous system) confirms global trends of codon usage bias (Figure 3.1) in the *Drosophila* embryo. There are clear classes of genes with shared codon usage; for example, genes across the top portion of the heatmap show an enrichment of the set of codons (shown in blue) designated as preferred in *Drosophila melanogaster* by multiple independent studies [168, 178, 180, 181]. These studies assessed codon usage bias across the *Drosophila* genome with preferred codons being those that are favored among highly expressed genes and correlate with higher tRNA abundance. Although there is a consensus of preferred codons in the studies described above, measures of codon optimality such as the tRNA adaptive index (tAI), codon adaptation index (CAI), tRNA abundance and tRNA gene copy number, which perform well for yeast [15, 182], are not as well-suited for more complex metazoans like *Drosophila*. Interestingly, among the set of genes containing more putative *optimal* codons, 55% are highly stable in the whole embryo but ≥ 1.5 -fold destabilized in the embryonic nervous system. Alternately, genes across the bottom portion of the heatmap show the opposite codon usage trend, with enrichment of codons that are not among the preferred set. Among this set of genes containing more putative *non-optimal* codons, 25% are highly stable in the embryonic nervous system (≥ 3 -fold stabilized compared to the whole embryo).

To assess the relationship between codon usage bias and mRNA decay dynamics, and define codon optimality using an independent approach from those detailed above, the codon stabilization coefficient (CSC) was calculated for all codons in *Drosophila melanogaster* based on whole embryo mRNA decay rates (Figure 3.2A). CSC values are a measure of the Pearson correlation between codon usage and transcript stability. The distribution of CSC values across all codons in the whole embryo shows a definite bias in codon usage linked to decay and many of the previously described



preferred codons (shown in blue) are among the most optimal codons by CSC. The maximum and minimum CSC values and proportion of optimal versus non-optimal codons observed for *Drosophila* whole embryos is similar to that seen in the original yeast study [15] (23 optimal versus 38 non-optimal codons by CSC in both organisms), although the optimal and non-optimal sets of codons are distinct for each organism. The similar distribution of CSC values in the *Drosophila* whole embryo and yeast is unsurprising considering both organisms show similar magnitudes of codon usage bias based on other measures [178]. Further analysis of the relationship between codon optimality and mRNA stability reveals strong links between optimal codon content and mRNA decay in the whole embryo. Genes classified by optimal codon content exhibit similar decay rates, especially among genes with high optimal codon content (Figure 3.2B). Additionally, CSC values correlate with other measures of

codon adaptation like tAI ($r = 0.485, p < 0.0001, 3.2C$) and tRNA gene copy number ($r = 0.601, p < 0.0001, A.4A$). Preferred codons in *Drosophila* tend to cluster toward positive CSC values and higher tAI values in the whole embryo.

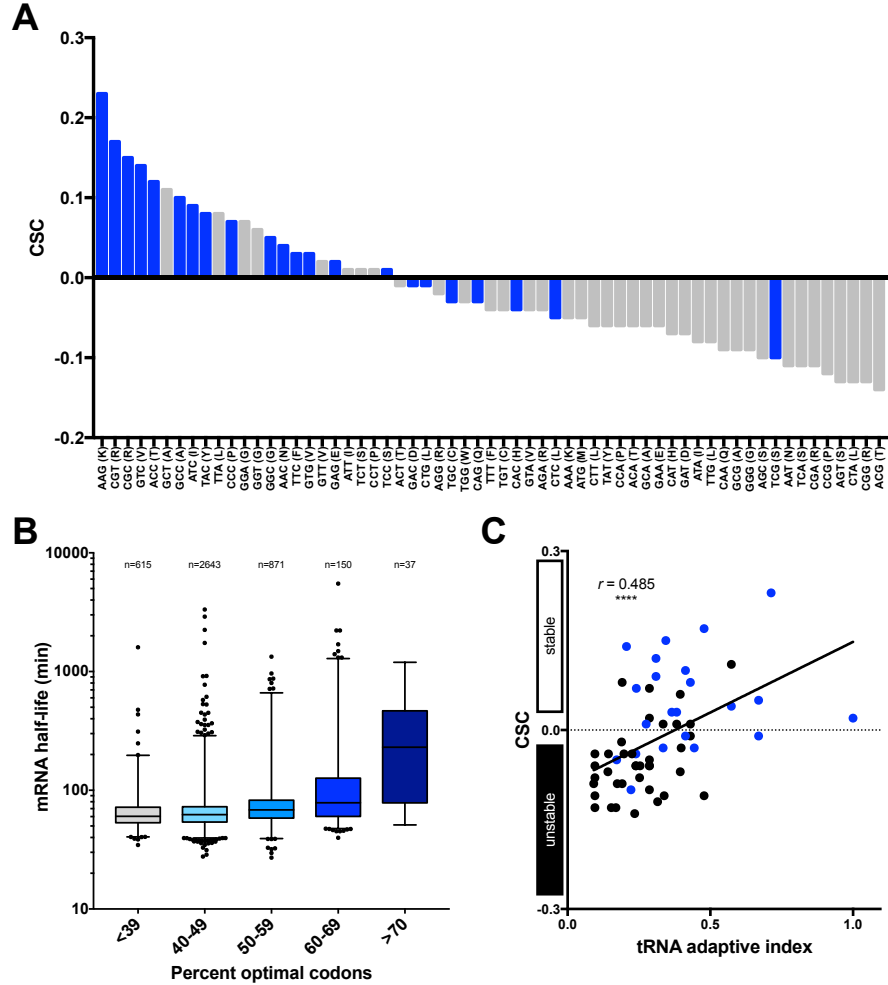


Figure 3.2: Codons in the whole embryo are arranged by CSC values (A), putative optimal codons are designated as those with positive CSC values. Box plots of mRNA half-lives categorized by optimal codon content in the whole embryo (A), half of the data fall within the box and the whiskers represent the 1st and 99th percentile. Correlations between tAI and CSC in the whole embryo (C), Pearson r values are indicated and **** designates $p < 0.0001$. Codons designated in blue are preferred for *Drosophila melanogaster* [168, 178, 180, 181].

Functional annotation of genes containing a high fraction of optimal codons and

those containing a low fraction of optimal codons was performed by DAVID analysis, with the optimal set of codons determined based on a positive CSC value. Core metabolic and house-keeping genes were among the most highly enriched classes among genes with a high fraction (> 0.6) of optimal codons (Table 3.1). While transcriptional regulators, synapse-associated proteins, and regulators of cell death were among the most enriched classes of genes with a low fraction (< 0.4) of optimal codons (Table 3.1). The stability of these classes of high-optimal and low-optimal genes is reflected by the whole embryo half-life; high-optimal codon genes have longer half-lives and low-optimal codon genes have shorter half-lives (Figure 3.3A). A more striking comparison can be made when looking at the relative stability in the embryonic nervous system versus whole embryo, where high-optimal genes show a trend of destabilization (mean neural/whole embryo half-life ratio of 0.90, $p < 0.0001$) and low-optimal genes show no significant difference in stability from the set of all genes (mean neural/whole embryo half-life ratio of 1.12 for low-optimal versus 1.13 for all genes, *N.S.*) (Figure 3.3B). Multiple measures of codon bias confirm the same trend of neural-destabilized genes showing stronger bias toward optimal codons defined in the whole embryo (Figure A.5A-C). There are significant differences in the fraction of optimal codons categorized by CSC between the neural-destabilized set of genes in relation to all genes or neural-stabilized genes (Figure 3.3C). The same trend of relatively high optimal codon content among destabilized genes in embryonic neural tissues is observed for other measures of codon bias: optimal codons categorized by CAI (Figure 3.3D) and overall codon bias defined by ENC (Figure 3.3E). These trends are not seen when assessing other sequence features that can account for codon usage bias such as GC content at the codon's 3rd position (Figure 3.3F) or overall GC content (Figure A.5D), indicating selection for optimal codons is not merely a result of mutational bias.

mRNAs with high-fraction optimal codons					
GO Term	Definition	<i>P</i> value	FDR	Count	Fold-enrichment
0003735	Structural constituent of ribosome <i>RpS15Aa, RpL23, RpL9, RpL13</i>	3.2×10^{-21}	2.1×10^{-19}	26	14.6
0015992	Proton transport <i>Vha16-1, Vha36-1, Vha55, blw</i>	9.9×10^{-6}	4.3×10^{-5}	7	7.1
0006099	Tricarboxylic acid cycle <i>Mdh1, Mdh2, Scsα</i>	7.6×10^{-4}	2×10^{-2}	5	3.1
0005784	Translocon complex <i>Sec61α, Sec61γ, Tapδ</i>	7×10^{-4}	9.5×10^{-3}	6	2.5
0019395	Fatty acid oxidation <i>Thiolase, yip2</i>	2.6×10^{-3}	0.05	4	2.1
0031497	Chromatin assembly <i>His3.3A, His3.3B, His2B, Nlp, Df31</i>	0.02	0.19	5	1.3
mRNAs with low-fraction optimal codons					
GO Term	Definition	<i>P</i> value	FDR	Count	Fold-enrichment
0045449	Regulation of transcription <i>sv, sav, osa, sqz, GATAd, Nca</i>	2.7×10^{-6}	3.9×10^{-3}	77	5.2
0008270	Zinc ion binding <i>chif, G9a, Adar, Laspl, scrt</i>	2.9×10^{-7}	1.2×10^{-4}	93	4.2
0031327	Negative regulation of cellular biosynthetic process <i>SuUR, pho, crol, ph-d, Dsp1, brk</i>	2.3×10^{-3}	0.57	23	2.4
0045202	Synapse <i>Cals, Syt7, l(2)gl, unc-13, nAChR</i>	0.07	0.82	8	1.7
0010941	Regulation of cell death <i>pnt, fkh, sav, Dark, skl, morgue</i>	0.08	0.99	11	1.1

Table 3.1: Gene ontology analysis of high-optimal mRNAs (optimal codon fraction > 0.6) and low-optimal mRNAs (optimal codon fraction < 0.4) by DAVID analysis.

3.2.2 Codon-mediated decay effects are attenuated in neural tissues

Given the relationship between high optimal codon content among destabilized mRNAs in the embryonic nervous system, it is plausible that codon optimality is tunable

and highly context-dependent. To assess the relationship between codon optimality and neural-specific mRNA decay dynamics, CSC values were calculated for all codons in *Drosophila* based on embryonic nervous system mRNA decay rates. Strikingly, the trend of CSC values seen in the whole embryo context is not seen in the embryonic nervous system, where there is flattening of CSC values hovering around zero (Figure 3.4A). The distribution of preferred codons (shown in blue) is also more random with no clear indication of optimal codons in the nervous system based on CSC values. Classification of mRNAs based on optimal codon content reveals no strong ties to neural mRNA stability. Additionally, no correlation between CSC and other measures of codon adaptation like tAI ($r = -0.090$, *N.S.*, Figure 3.4C) or tRNA gene copy number ($r = 0.019$, *N.S.*, Figure A.4B) is seen in the nervous system. Taken together these data indicate that codon optimality is an indicator of mRNA stability in the whole embryo context and codon-mediated effects on neural mRNA decay dynamics are attenuated.

Although codon-mediated effects on mRNA stability are strongest in the whole embryo, these data may reflect optimization of expression dynamics for important core housekeeping and metabolic genes required for sustained rapid growth in this context. Alternately, attenuated codon-mediated effects on neural mRNA stability may reflect changing global mRNA stability dynamics, which require more fine-tuned regulation of protein expression via alternate mechanisms. mRNAs with higher fractions of optimal codons, long half-lives in the whole embryo and destabilization in the developing nervous system are enriched in constitutively-expressed ribosomal protein-coding genes and chromatin-associated genes including: *RpL6*, *RpS23*, *RpL23*, *His3.3B*, *His4r* (Figure 3.5A). While mRNAs with lower fractions of optimal codons, short half-lives in the whole embryo and stabilization in the embryonic CNS are enriched in genes involved in transcriptional regulation, cell death and neural-development functions including: *CadN*, *par-6*, *baz*, *CASK* (Figure 3.5B). High-optimal neural-destabilized mRNAs contain core housekeeping genes like ribosomal protein and non-canonical histone genes, which are ubiquitously expressed. Members of the non-canonical replication-independent histone 3 and 4 families (*His3.3A*, *His3.3B*, *His4r*) all exhibit higher degrees of codon usage bias [183] and specifically higher fractions

of optimal codons (from this study) than their replication-dependent counterparts (though both classes are destabilized in the embryonic nervous system). These differences may be explained by core structural differences between genes, which is indicative of vastly different modes of regulation. Replication-independent histones are single or few-copy genes with introns and their mRNAs are polyadenylated [184]; in contrast, replication-dependent histones (*H1*, *H2A*, *H2B*, *H3*, *H4*) are intronless genes contained within a tandemly duplicated histone gene unit, and whose mRNAs are not generally polyadenylated [185, 186]. Replication-independent histone family members evolved under different pressures than their replication-dependent counterparts, which require synchronous expression during S-phase of the cell cycle. This may be why replication-independent histones exhibit a greater degree of selection for optimal codons, which in turn influences their stability and translation. Members of the enriched classes of low-optimal neural-stabilized mRNAs are expressed in tissues other than the nervous system though the specific genes indicated (Figure 3.5B) have well-defined roles in neural-development. Cadherin-N (CadN) is a neural cell adhesion protein that functions in axon guidance through control of fast filopodial dynamics and growth cone stabilization [187]. Although *CadN* is widely transcribed in mesodermal and neural cells of the developing *Drosophila* embryo, the protein accumulates at axon terminals in later embryonic stages [188], suggesting the necessity of neural-specific post-transcriptional control of protein expression. *par-6* and *baz* exhibit similar decay dynamics and codon bias (Figure 3.5B) and are members of the apical polarity complex in neuroblasts, colocalizing to the apical cortex. Absence of either *par-6* or *baz* causes diffuse mislocalization to the cytoplasm and asymmetric cell divisions in neuroblasts become disordered [189]. Synchronous regulation of these proteins appears to be important for their proper function and this is reflected in their mRNA stability. *CASK* is a kinase and scaffolding protein that is widely expressed [190], but localizes to synapses and regulates glutamatergic synaptic transmission [137]. *CASK* has low-optimal codon content and is stabilized in the embryonic nervous system similar to the previously mentioned neural development associated mRNAs. It is plausible that codon-mediated mRNA decay dynamics are distinct in the developing nervous system; there appears to be less selective pressure toward usage of

optimal codons in neural-stabilized transcripts. Additionally, ubiquitously expressed core housekeeping genes like ribosomal protein mRNAs and non-canonical histones, which have optimized codon usage in the whole embryo context, show destabilization in the nervous system. This suggests codon-mediated decay dynamics are highly context-specific and have the ability to drive bulk changes in mRNA stability.

To confirm codon-mediated mRNA decay dynamics in the *Drosophila* whole embryo and embryonic nervous system, codon reporter constructs were made from select neural-destabilized and neural-stabilized classes of genes. Wild-type and mutant mRNA CDS were ligated into a vector containing the *UAS* enhancer and endogenous *Hsp70* UTR sequences. These reporters were expressed in the whole embryo (under the control of *lola(ubiquitous)-GAL4*) or embryonic nervous system (under the control of *prospero-GAL4*). Stage 12 to 15 embryos were pulse-labeled with 4-thiouridine and mRNA abundance measured at the initial pulse, time zero, and after a two hour uridine chase. *RpS15Aa*, a ribosomal protein-coding gene, has a high fraction of optimal codons (0.63) and is destabilized in the embryonic nervous system (Figure 3.6A). The wild-type *RpS15Aa* reporter exhibits stabilization in whole embryos, while being destabilized in the nervous system (Figure 3.6E). A mutant *RpS15Aa* reporter comprised of all non-optimal codons via synonymous mutations in the CDS (Figure 3.6C) exhibits destabilization in both the whole embryo and neural contexts (Figure 3.6E). *Syb*, a ubiquitously expressed membrane protein associated with vesicle secretion, has a slightly above-average fraction of optimal codons (0.53 compared to 0.46), and is stabilized in the embryonic nervous system (Figure 3.6B). A mutant *Syb* reporter comprised of all optimal codons via synonymous mutations in the CDS (Figure 3.6D) exhibits stabilization in both the whole embryo and neural contexts (Figure 3.6F). The altered stability of the *Syb* wild-type reporter compared to the endogenous mRNA in the nervous system may be due to effects the UTRs have on stability that are not recapitulated in the reporter assay. Interestingly the optimal *Syb* reporter transcript (80% optimal codons) is stabilized in neural tissues while the *RpS15Aa* wild-type (63% optimal codons) remains unstable. This difference may be explained by the higher fraction of optimal codons in the optimized *SyB* reporter.

3.3 Discussion

These results indicate codon-mediated mRNA decay dynamics are highly context-specific. *In vitro* studies of mRNA stability upon cellular differentiation in mammals corroborate these findings. Analysis of mRNA decay in human iPS cells and their fully differentiated counterparts (human foreskin fibroblasts) shows destabilization of histones and master transcriptional regulators of pluripotency (i.e. zinc-finger proteins) upon differentiation [118]. Mouse ES cells also show decreases in stability of ribosomal protein mRNAs upon differentiation, although there was no correlation between mRNA abundance and changes in stability upon differentiation [119]. Our data also show no correlation between mRNA abundance and mRNA stability between the embryonic nervous system and whole embryo (Figure A.1C,D), which may suggest codon-mediated decay dynamics are primarily a function of translation efficiency. Recent evidence in zebrafish embryos demonstrate that maternally deposited mRNAs are subject to codon-mediated mRNA decay dynamics dependent on *CCR4-NOT* deadenylation (the first step in bulk 5' $\rightarrow$ 3' decay) with 3'UTR length having a modulatory effect [16]; longer 3'UTRs protected maternal transcripts from deadenylation-dependent 5' $\rightarrow$ 3' decay. While long 3'UTRs may protect against deadenylation-dependent decay, it does not rule out regulation via alternative mRNA decay pathways (e.g. NMD) or translation inhibition, where long 3'UTRs may allow for targeting by *trans*-acting RBPs (e.g. Pum) or ncRNAs. Ultimately the interplay between specific regulatory mechanisms mediated through *cis*-element/*trans*-acting factor recognition (e.g. 3'UTR lengthening) and global regulatory mechanisms affecting decay and translation (i.e. codon optimality, translation-dependent decay) determine the final protein expression pattern. This mode of global regulation, occurring at the level of protein translation via regulation of the ribosome and tRNAs, would allow bulk changes in both mRNA decay and translation in response to a shifting cellular program. This shift may reflect the differences between a program of rapid proliferation and growth, and a program of differentiation and cellular maturation. Bulk regulation of gene expression at the transcriptional level is achieved through modulation of gene expression by master transcriptional regulators (e.g. via

chromatin and transcription factors); in the same way, bulk regulation of protein expression can be achieved through regulation of mRNA decay and translation efficiency. It is plausible that the most efficient means of changing bulk decay and translation dynamics is via the translation machinery itself: ribosomes and tRNAs. To corroborate this, it will be necessary to measure tissue-specific translation and tRNA dynamics.

3.4 Approach

3.4.1 Fly lines

UAS-GFP-3'UTR reporter constructs were made by modifying *pUAST-3XemGFP* (Drosophila Genetics Resource Center) to remove two copies of *emGFP* and replace the *SV40* terminator with *Drosophila Hsp70* terminator and an *attB* site. The reporter backbone was shipped to the Collier lab (Case Western Reserve University, Cleveland, OH) for ligation of chemically synthesized *Rps15Aa wild-type*, *Rps15Aa non-optimal*, *Syb-RB wild-type*, or *Syb-RB optimal* DNA (Integrated DNA Technologies, Inc.) into the modified *pUAST-emGFP-attB* backbone (performed by Najwa Al-Husaini). Constructs were mailed back to UC Merced and submitted for injection into *Pnos-phi- C31int.NLSX*, *PCaryPattP40* embryos for PhiC31- mediated integration (Rainbow Transgenic Flies, Inc.).

3.4.2 Codon analysis

Preferred codons are those previously designated in multiple independent studies [168, 178, 180, 181]. Codon stabilization coefficients (CSC), a Pearson correlation between codon usage and transcript stability, were calculated as previously described [15] by Vladimir Presnyak and Sophie Martin (Case Western Reserve University, Cleveland, OH). Optimal codons were determined to be those with positive CSC values. Effective number of codons (ENC), codon adaptation index (CAI), frequency of optimal codons (Fop), and GC content were calculated using **CodonW** (version 1.4.4) [191] for the set of all CDS corresponding to genes with reproducible half-live measurements for the

whole embryo and embryonic CNS. *Drosophila melanogaster* tRNA gene copy number is based on the Genomic tRNA database (GtRNAdb) [192] and Berkeley Drosophila Genome Project release 6 (dm6). tRNA adaptive index (tAI) values were calculated using CUA available on CPAN [193].

3.4.3 Transcript CDS reporter assay

UAS-CDS reporter lines were crossed to either *prospero-GAL4* or FlyLight project line GMR45H06 (*lola(ubiquitous)-GAL4*) (Bloomington Drosophila Stock Center) for embryo neural or whole embryo reporter expression respectively. St. 12-15 embryos were permeabilized as previously described [96] and pulse-labeled in tagging media [D22 insect media, 5% FBS, 1 mM 4-thiouridine (whole embryo tagging)] for 1 h at 25°C. Pulse-labeled embryos were then transferred for 2 h to chase media containing uridine [D22 insect media, 5% FBS, 10 mM uridine] at 25°C. Embryos were homogenized for 30 sec and frozen in TRIzol[®] (ThermoFisher) at 80°C. 50 μ g samples for each time point were pooled for TRIzol[®] (ThermoFisher) RNA extraction and TU-RNA purification as previously described [96] with the modification of MTSEA-biotinylation [98]. Purified TU-RNA quality was assessed by Bioanalyzer RNA Pico Eukaryotic total RNA assay (Agilent) prior to transcriptome amplification and cDNA synthesis (MessageBOOSTER cDNA Synthesis Kit for qPCR, Epicentre). Real-time PCR quantitation was performed on a Rotor-Gene Q (Qiagen) in 20 uL reactions using QuantiTect Primer Assays (Qiagen) or custom designed primers (Table A.1), and SYBR green detection (QuantiTect SYBR Green PCR Kit, Qiagen). C_t values were normalized to an *FLuc* spike-in reference and relative abundance calculated by the equation, fold-change = $2^{\Delta\Delta C_t}$.

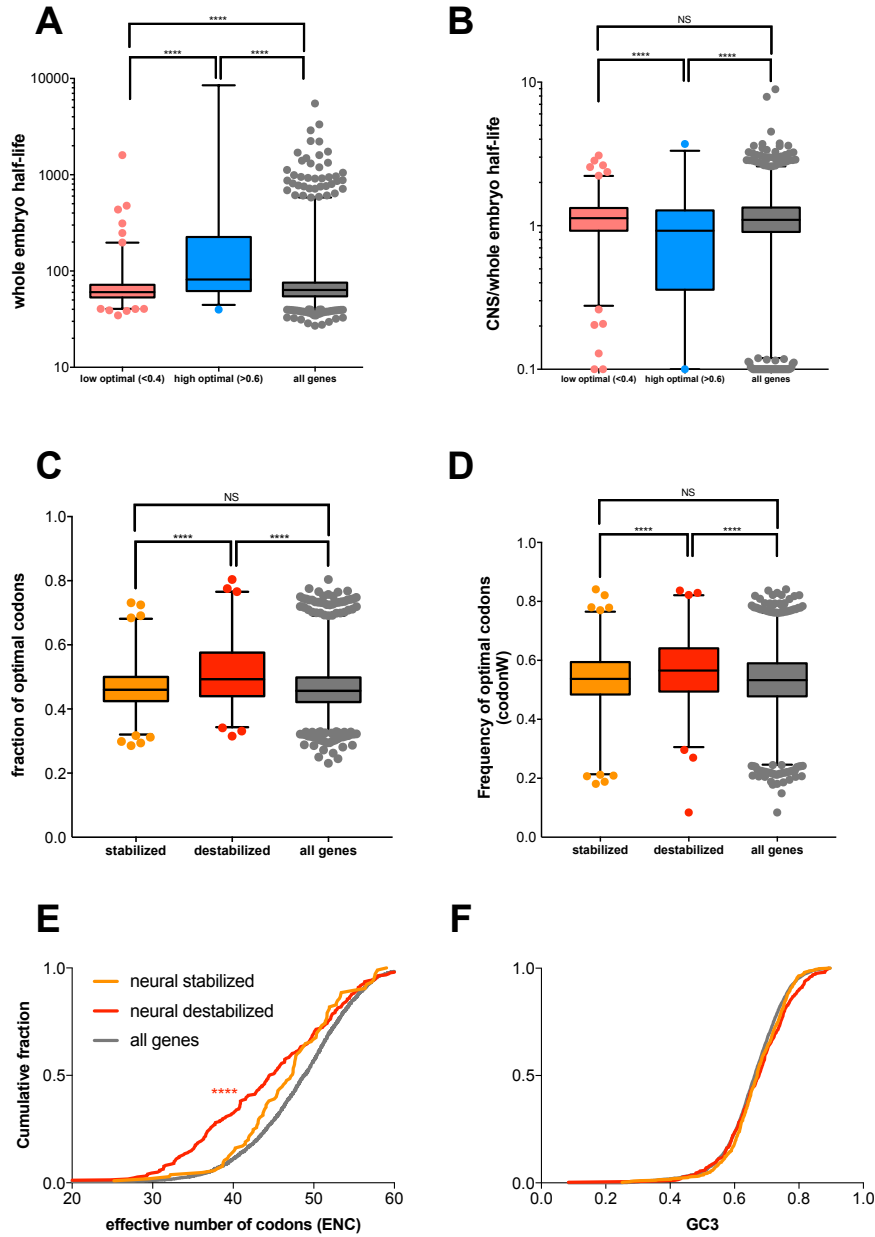


Figure 3.3: Box plots of mRNA half-lives categorized by low-optimal (A) or high-optimal (B) codon content. Box plot of optimal codon content, with optimal codons classified by CSC values, categorized by CNS stabilized or CNS-destabilized mRNAs (C). Box plot of optimal codon content, with optimal codons classified by CAI values (codonW), categorized by neural-stabilized or neural-destabilized mRNAs (D). Whiskers represent 1st and 99th percentile. Cumulative distributions of effective number of codons (ENC) (E) and GC content (F) for neural-stabilized (orange), neural-destabilized (red) or all genes (gray). Group comparisons are assessed by Kruskal-Wallis test corrected for multiple comparisons, **** designates $p < 0.0001$.

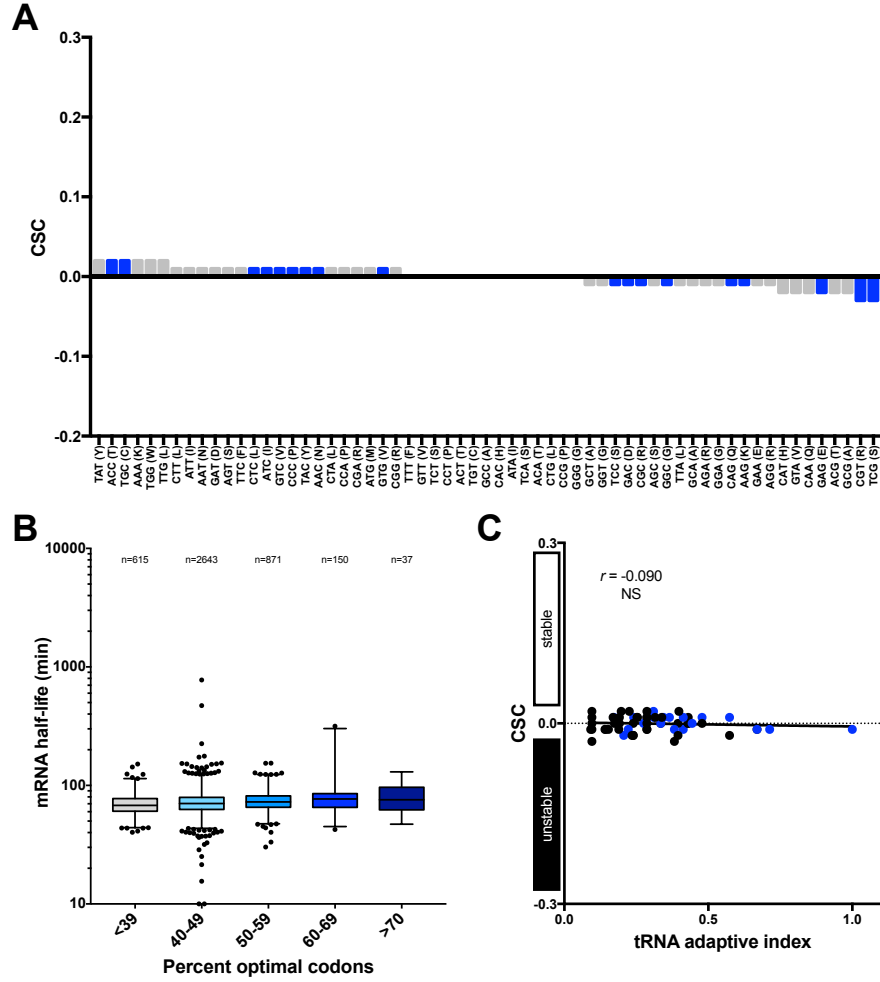


Figure 3.4: Codons in the embryonic nervous system are arranged by CSC values (A). Box plots of mRNA half-lives categorized by optimal codon content in the embryonic nervous system (B), half of the data fall within the box and the whiskers represent the 1st and 99th percentile. Correlations between tAI and CSC in the embryonic nervous system (D), Pearson r values are indicated and **** designates $p < 0.0001$. Codons designated in blue are preferred for *Drosophila melanogaster* [168,178,180,181].

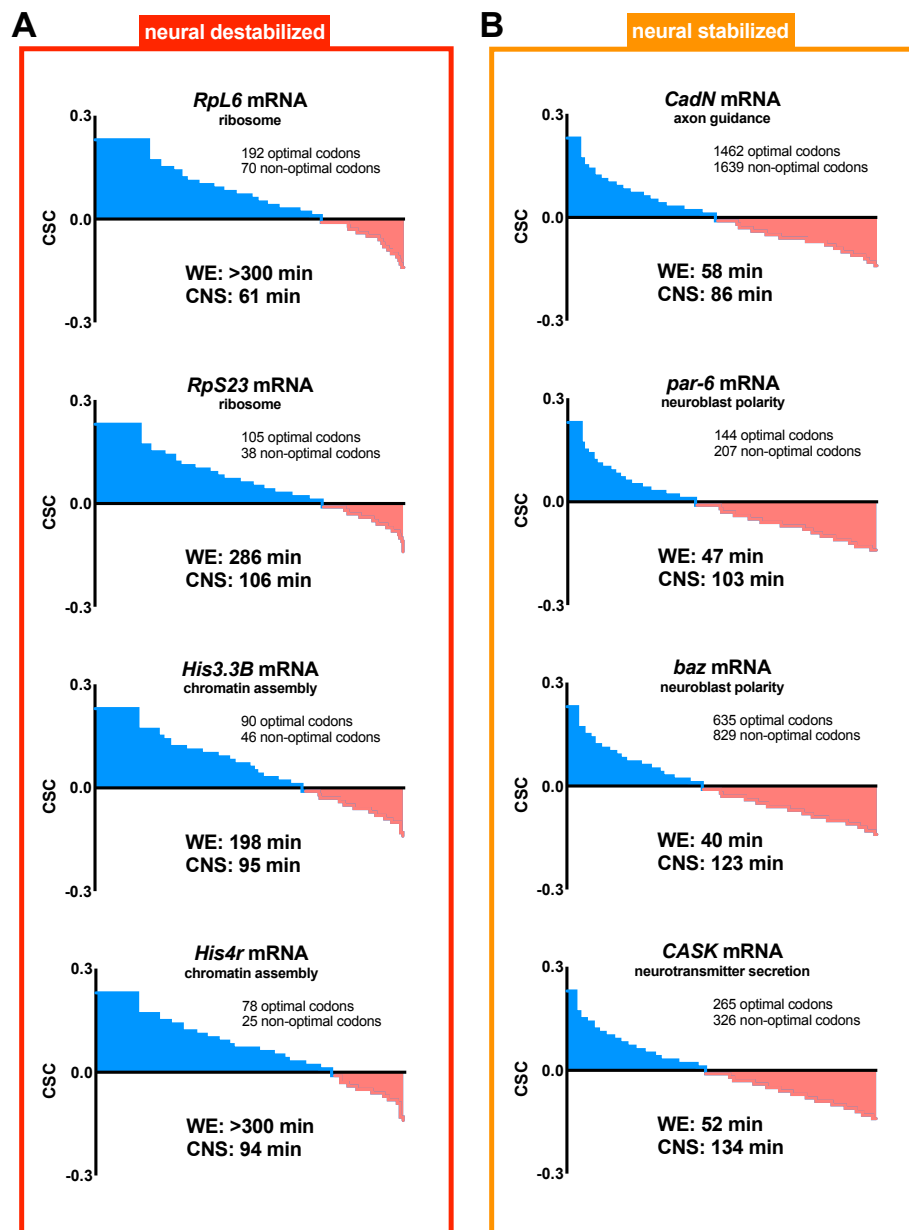


Figure 3.5: Select neural-destabilized ribosomal protein encoding mRNA and chromatin associated mRNA (i.e. histones) (A), and select neural-stabilized neural development associated mRNA (B) coding sequences are illustrated in order of their codon optimality. Optimal codons are shown in blue and non-optimal codons in pink.

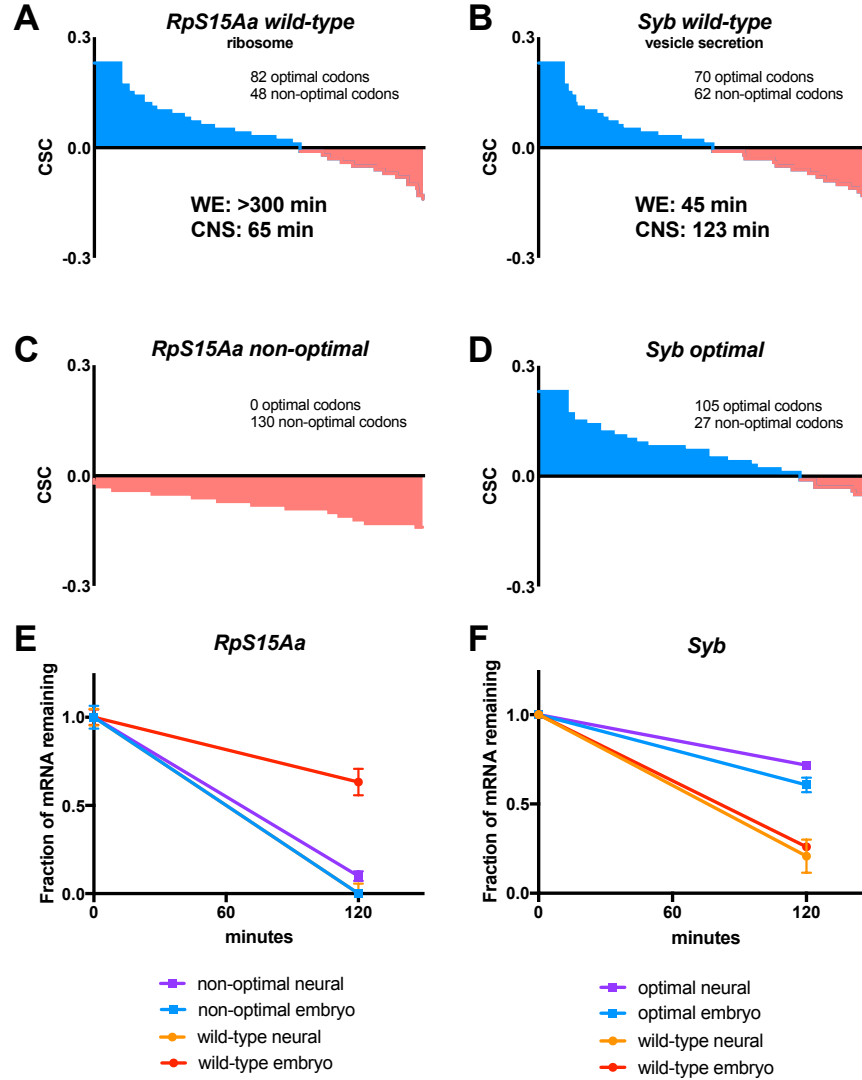


Figure 3.6: *RpS15Aa* wild-type (A), *Syb* wild-type (B), *RpS15Aa* non-optimal (C), and *Syb* optimal (D) coding sequences are illustrated in order of their codon optimality. Optimal codons are shown in blue and non-optimal codons in pink. TU-decay and RT-qPCR quantification of *RpS15Aa* wild-type and *RpS15Aa* non-optimal codon reporters expressed in either the whole embryo or embryonic nervous system (E). TU-decay and RT-qPCR quantification of *Syb* wild-type and *Syb* optimal codon reporters expressed in either the whole embryo or embryonic nervous system (F).

Chapter 4

Unbiased TU-decay via RNA-seq

4.1 Introduction

Global analysis of RNA decay by TU-decay can be expanded with the inclusion of new ways to measure genome-wide RNA abundance. Highly parallel systems for monitoring global expression patterns arose with complementary DNA microarray technologies [1] and continue to evolve with new sequencing-based technologies (e.g. RNA-seq) [2]. RNA-sequencing has gained popularity due to its efficiency, nucleotide-level resolution and high level of flexibility. RNA-seq is advantageous over array technologies because it allows a more unbiased approach, offering direct monitoring of the transcriptome through sequencing. Though RNA-seq has many advantages, it is still important to address experimental noise and bias related to the sequencing technology, experimental design, and procedures for assessing differential expression. Some inherent errors in massively parallel sequencing technologies include: discrepancies in base calling and sequence quality [194, 195], variability in sequencing depth and nucleotide composition that affect transcriptome coverage [196], and technical and biological variability across samples [2, 197]. Unlike microarrays, which correlate a continuous fluorescence intensity measure with expression, RNA-seq data are discrete counts of sequence aligning to a particular genomic feature. These discrete count data require a different set of statistical tools and procedures for modeling variance and

determining differential expression. Despite growing popularity of RNA-seq based differential expression profiling there is not a strong consensus on how to address sources of noise and bias, and detect differential expression [198,199]. Considerations in both experimental design and statistical modeling to detect differential expression are important for generating informative RNA-seq data. These decisions on experimental design and analysis are highly dependent on the system and hypotheses under consideration. Sequencing depth, biological and technical replicates, sequencing library construction and normalization procedures are important factors in the experimental design (Figure 4.1). Samples of synchronized (TU-tagged) populations of RNA will be the basis for establishing methods to monitor transcriptome-wide RNA decay using TU-decay by RNA-seq. These data can provide novel information on the stability of non-coding RNAs, underrepresented mRNAs, and RNA decay intermediates.

4.2 Preliminary Results

4.2.1 Strategy for transcriptome-wide TU-decay by RNA-seq

To achieve transcriptome-wide RNA-seq measurements for modeling transcript decay, libraries were generated from duplicate purified TU-RNA samples containing a spike-in control. Ribosomal RNA (rRNA) transcript depletion was achieved using *Drosophila*-specific complementary oligonucleotides (detailed in Section 4.4.1). Libraries were sequenced using the Illumina HiSeq platform, which provides the highest throughput. Though increased sequencing depth allows greater coverage of all transcripts [200], it is more costly and does not guarantee a more sensitive measurement [199]. Paired-end sequencing was performed in order to achieve the best detection of transcript isoforms. Technical replication of RNA-seq data is fairly robust, especially for designs that allow for multiplexing of samples, where the majority of variation is related to library construction [201]. Technical replicates were not included in this study, however, addressing biological variation through replication is crucial. Optimal gains in sensitivity to detect of differential expression are achieved

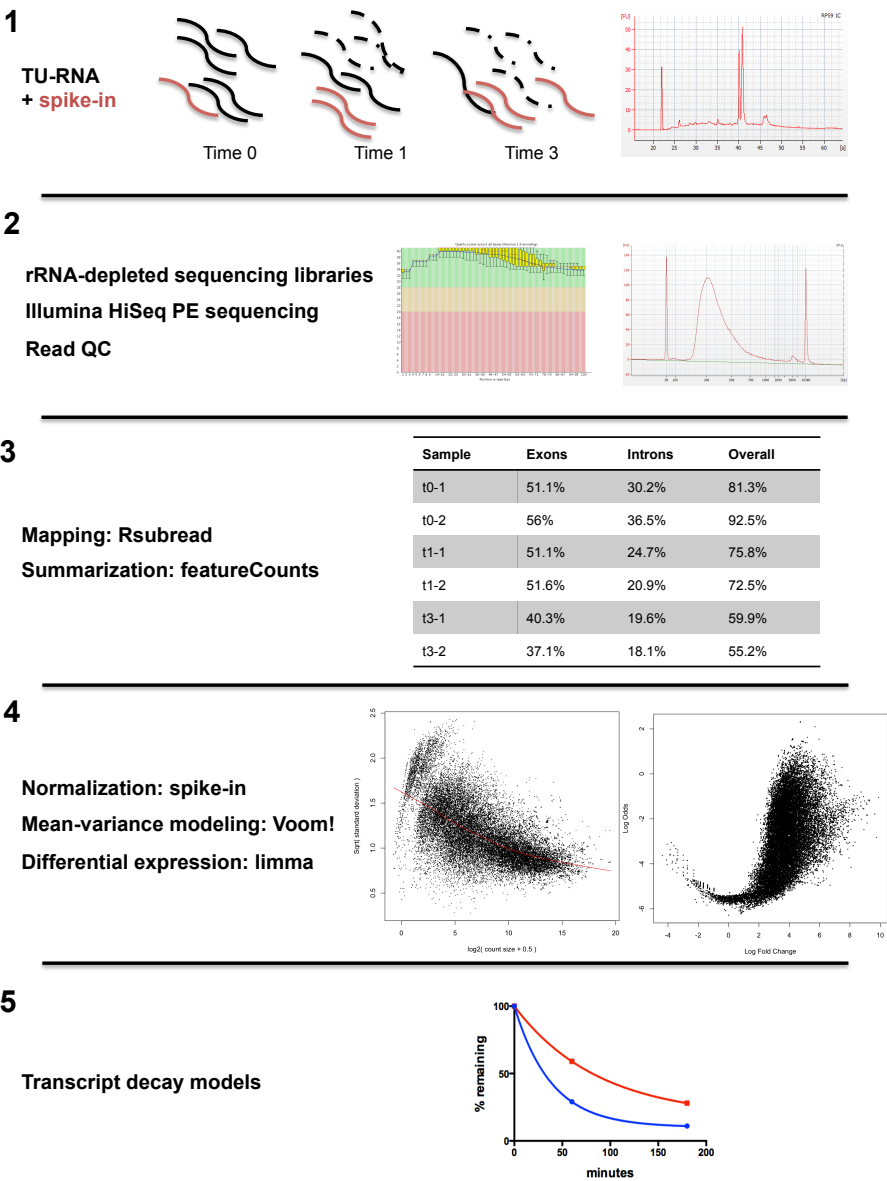


Figure 4.1: Strategy for transcriptome-wide TU-decay by RNA-seq

with $n = 5$ replicates [202]. Although $n = 5$ replicates may be ideal for assessing differential expression, a large number of replicates is less crucial when studying systems with low biological variability such as cell lines and inbred animals, including the laboratory fly lines used in this study.

In addition to developing a suitable experimental design, selection of a powerful analysis approach, including normalization and variance modeling procedures, will influence the sensitivity and specificity of decay measurements. Variability in library size (i.e. sequencing depth) and RNA composition determine appropriate normalization procedures. For example, size-based normalization procedures (e.g. fragments Per Kilobase of transcript per Million mapped reads (FPKM)) assume similar RNA composition across samples. Implicit differences in RNA composition across TU-decay libraries presents a challenge for size-based normalization procedures and will require additional normalization through use of spike-in controls. Many differential expression analysis tools (e.g. edgeR, DESeq, and voom) employ a scaling factor like the trimmed mean of M-values (TMM) or quantile normalization for library size to minimize log fold-change between samples [203–205]. Early statistical models assume a Poisson distribution of reads, setting variance equal to the mean, however this approach fails to model the higher levels of variance seen in discrete RNA-seq count data, termed *overdispersion* [203, 206–208]. Many methods (e.g. DESeq) now use a negative binomial distribution to model this overdispersion and estimate variance [207, 209]. Another tool, voom [210], estimates the mean-variance of the log counts per million (log-cpm) by locally weighted scatterplot smoothing (LOWESS regression), giving a residual standard deviation for each gene. Overdispersion in the discrete RNA-seq count data versus continuous microarray fluorescent intensity data can be visualized by the mean-variance relationship. The microarray data has fairly uniform variation (Figure 4.2B) while RNA-seq data has relatively more variation for smaller counts than larger counts (Figure 4.2A). Using voom, fitted log-cpm are turned into a predicted count and assigned a weight, which is the inverse of the predicted variance [210]. The advantage of this approach is that these measures are suitable for integration with tools (i.e. limma [211]) used for assessing expression in continuous expression data, like microarray data. limma provides a powerful and flexible framework for assessing differential expression using linear models.

Using limma, differential expression across the TU-decay time series was assessed by two approaches. Since limma is not specifically designed to model decay, these approaches essentially provide different insights into the RNA-seq decay data. In this

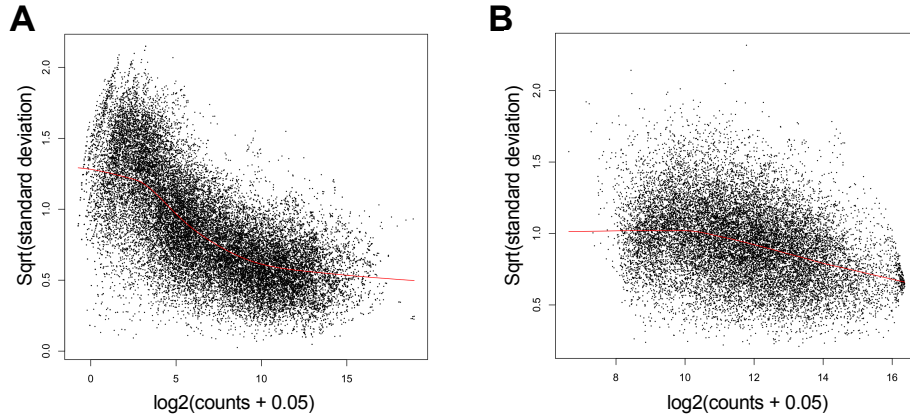


Figure 4.2: The mean (log counts or fluorescence intensity) and variance (sqrt standard deviation) relationship of the TU-decay RNA-seq (A) and microarray data (B) are plotted for each transcript with a LOWESS trend shown in red.

first case, differential expression contrasts between all time points were assessed, taking into account biological replicates, and subsequently log-transformed abundances were used for linear regression to estimate decay. In the second scenario, differential expression over time was assessed with respect to two sets of replicate timecourse data. A volcano plot of the log fold-change plotted against the log odds of differential expression with respect to the whole genome shows a majority of genes in the RNA-seq data are decreasing over time, which is especially evident in the contrasts design (Figure 4.3A). Negative log fold-change in these scenarios corresponds with an increase in expression from time 0 hours to 3 hours (i.e. 0hr-3hr). These log odds scores are associated with a corrected p value (i.e. the Benjamini-Hochberg FDR [212]) of differential expression over time with respect to all measurements; meaning, since a vast majority of genes exhibit decay the log odds of differential decay is small ($\log \text{ odds} < 1$). This concept is illustrated when comparing half-life to the *decay* p value, which corresponds to fold-change in expression over time for a single observation (i.e. transcript) (Figure 4.3C). Transcripts that exhibit stability (larger half-lives) have correspondingly larger p values representing less significant decay over time. The alternate (time course) design does not explicitly take into account replicates, so time values are jittered over the actual time measurement and

applied to a set of samples. This method is essentially comparing differential expression between replicates though it treats them as separate treatments. Again, a plot of the log fold-change versus log odds of differential expression over time with respect to replicates shows a majority of genes in the RNA-seq data are decreasing over time (Figure 4.3B); although in this case, the log odds represents the variation between one iteration of replicates. For example, those transcripts with log odds of 1 corresponds to a 50/50 chance they are differentially expressed between replicate time courses. This approach can be improved by summarizing many iterations of randomly selected replicate time courses. Since the *decay p* value shown in this scenario is related to differential expression between replicate time courses over time, it does not correlate with half-life measurements (Figure 4.3D). These types of comparisons could easily be applied to looking at differential decay over time between tissues or treatments (instead of across replicates), in which case, differential decay between samples could be assessed by log odds scores or the adjusted *p* value.

4.2.2 Comparison of TU-decay by microarrays and RNA-seq

Comparison of decay models by microarray and RNA-seq based measurements shows similar distributions for adjusted R^2 values (Figure 4.4B), although half-lives of the RNA-seq data are largely smaller (Figure 4.4A). These differences are likely due to differences in normalization procedures with spike-in controls. The spike-ins for the RNA-seq experiments are *in vitro* synthesized with thio-UMP, allowing addition at the step of TU-RNA purification versus the arrays where spike-ins only control for cRNA synthesis and hybridization. It is possible that the array measurements are not adjusted to account for differential purification efficiencies of the pulse-labeled versus chase TU-RNA fractions. Although global differences in half-lives are seen between methods, the relationships between stability and gene function are largely the same. Similar to the microarray measurements, no significant correlations between initial transcript abundance or lifetime abundance and half-lives were seen (Figure A.6A,B).

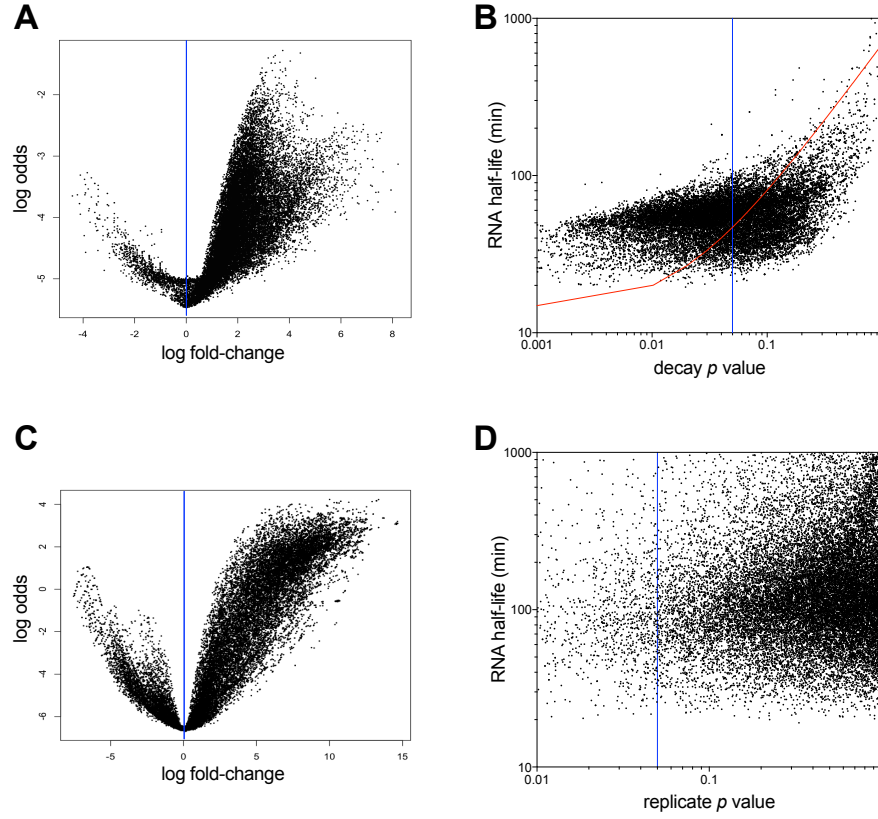


Figure 4.3: Log odds of differential expression over time (decay) with respect to all genes versus the change in abundance over time (A) and the genewise adjusted p value of differential expression over time (decay) versus RNA half-life (B) for the contrasts design. A linear fit is shown in red and the blue line designates $x = 0.05$. Log odds of differential expression over time (decay) with respect to all genes versus the change in abundance over time (C) and the genewise adjusted p value of differential expression over time (replicate) versus RNA half-life (D) for the time course design. The blue line designates $x = 0.05$. Adjusted p values are the Benjamini–Hochberg FDR [212].

However, comparison of half-lives of cohorts of functionally related genes reveals similar ties seen in the microarray based measurements (Figure 4.5A). Transcripts encoding ribosomal proteins, non-canonical histones, mitosis associated transcripts, and translation initiation factors have relatively higher stability, while canonical histones, cell-cycle associated, and neural development associated transcripts have relatively smaller half-lives. An advantage of RNA-seq based measurements of stability are

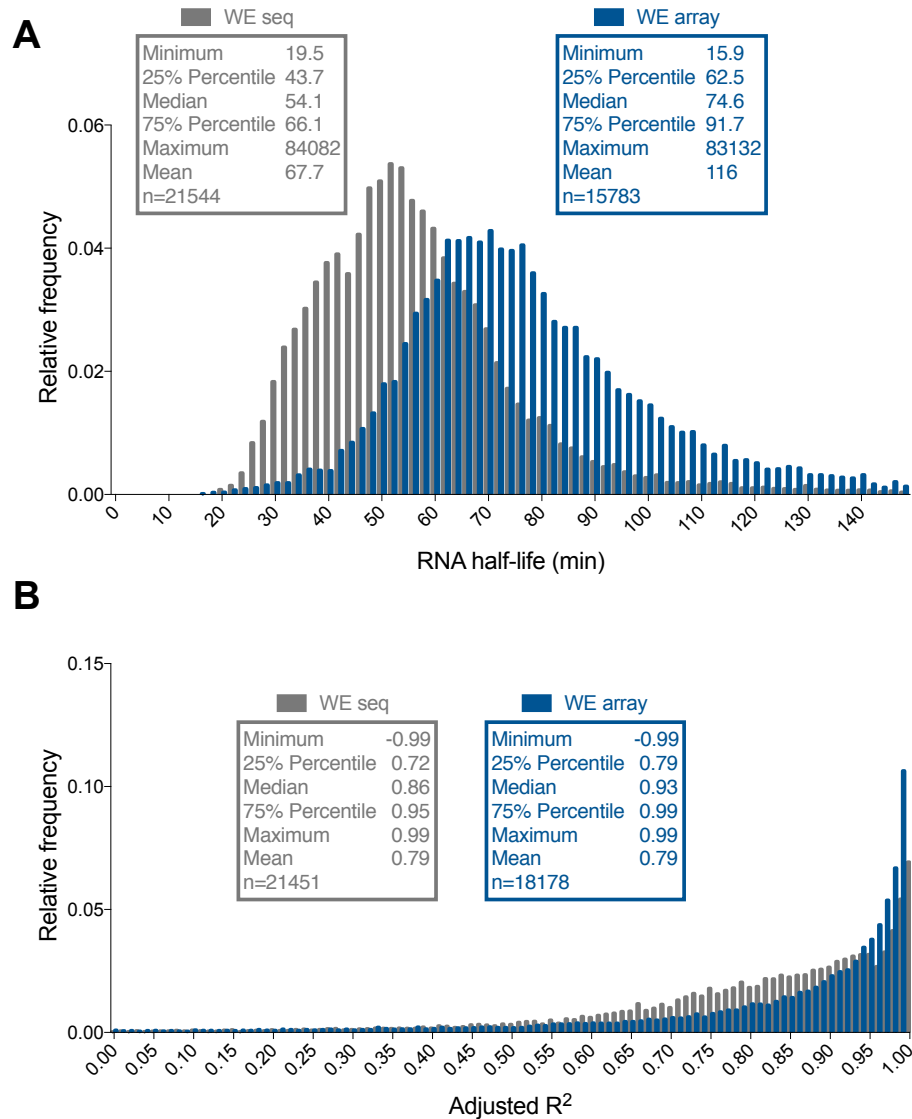


Figure 4.4: Histograms of mRNA half-life (A) and adjusted R^2 values (B) for RNA-seq and microarray based measurements.

they allow observation of ncRNAs. tRNAs and miRNAs have fairly broad distributions of half-lives compared to 5S rRNAs. The major species of *Drosophila* rRNAs were efficiently targeted for transcript depletion, but analysis of mitochondrial rRNAs reveals very long half-lives of nearly 700 minutes each. A number of *Drosophila* tRNAs were detected, although this experiment was not designed to efficiently capture

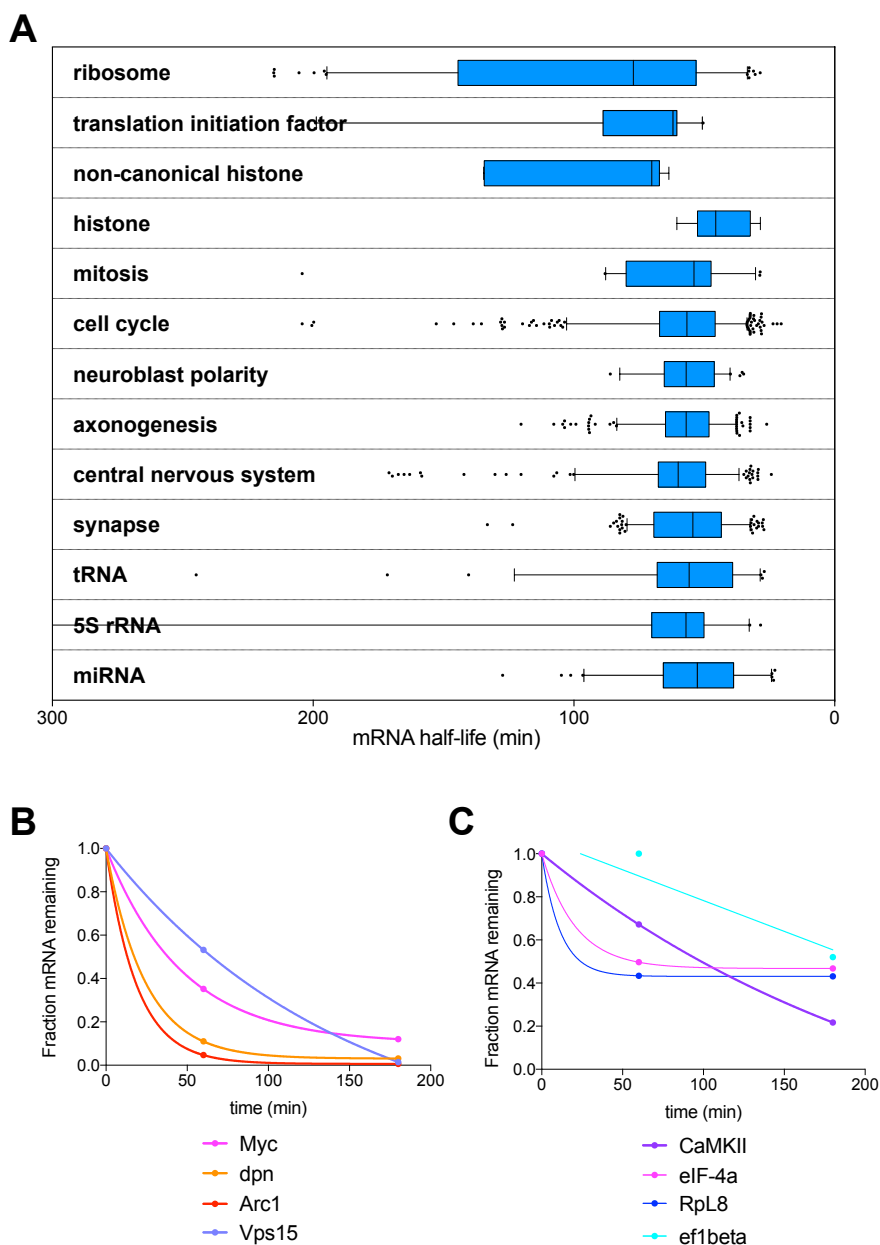


Figure 4.5: Box pots of RNA half-life of various coding and non-coding gene categories, whiskers represent the 5th and 95th percentile (A). Decay models for low (B) and high (C) stability RNAs.

small RNAs. There were no significant correlations between initial tRNA abundance,

lifetime tRNA abundance or tRNA stability and corresponding codon optimality (Figure A.7A-C). However, these data do not include the full repertoire of tRNAs and are not specific to charged tRNAs. Decay models of select low (Figure 4.5B) and high (Figure 4.5C) stability RNAs reveals different modes of decay. Among the low stability RNAs, there are transcripts that decay quickly initially (e.g. *Arc1*, *dpn*) versus those that decay more slowly initially (*Myc*, *Vps15*). Among the high stability set of RNAs, there are transcripts that decay initially then stabilize (e.g. *eIF-4a*, *RpL8*) versus those that decay slowly initially then more rapidly (*ef1beta*, *CaMKII*). An advantage of the `limma` framework is that these subtleties can be assessed through pairwise comparison between time points and to the overall decay model (i.e. by using the contrasts design).

4.3 Discussion

Development of a more unbiased approach to model global *in vivo* RNA decay is possible through TU-decay by RNA-seq and application powerful tools to assess differential expression and decay. New tools, like `voom`, that address the unique variance seen with discrete RNA-seq count data allow integration into existing pipelines for analyzing differential expression (i.e. `limma`). Preliminary data using this approach to assess differential expression over the TU-decay time course in whole embryos confirms the trends seen in the whole embryo decay data by microarray analysis. The RNA-seq based approach also provides novel *in vivo* decay data on ncRNAs including many miRNAs and tRNAs. Although only preliminary whole embryo data is presented here, this approach will be especially useful for complex experimental designs like assessing decay (via differential expression over time) between tissues or treatments (e.g. neural specific/whole embryo or *PumRNAi*/control) using TU-decay by RNA-seq.

4.4 Approach

4.4.1 Library preparation and sequencing

St. 12-15 *UAS-T.g.UPRT* embryos were permeabilized as previously described [96] and pulse-labeled in tagging media [D22 insect media, 5% FBS, 1 mM 4-thiouridine (whole embryo tagging)] for 1 h at 25°C. Pulse-labeled embryos were then transferred for 2 h to chase media containing uridine [D22 insect media, 5% FBS, 10 mM uridine] at 25°C. Embryos were homogenized for 30 sec and frozen in TRIzol® (ThermoFisher) at 80°C. 50µg samples for each time point were pooled for TRIzol® (ThermoFisher) RNA extraction and TU-RNA purification as previously described [96] with the addition of a firefly luciferase (FLuc) *in vitro* synthesized and TU-tagged spike-in control. Purified TU-RNA quality was assessed by Bioanalyzer RNA Pico Eukaryotic total RNA assay (Agilent) prior to generating randomly-primed dedicated-strand RNA-seq libraries using Ovation FFPE RNA-seq Library System (NuGen) with custom primers to deplete *Drosophila* rRNA sequences. Replicate libraries for each time point were pooled for 100bp paired-end sequencing by HiSeq2000 (Illumina) at BGI@CHOP.

4.4.2 Read mapping, summarization and normalization

After demultiplexing and quality control steps (to remove adapter contamination and poor quality sequence) were performed, reads were mapped to the *dmel* BDGP r5.25 reference genome and summarized by exons into transcripts using **Rsubread** package, **subjunc** function in order to best estimate transcript-level expression. A Firefly luciferase spike-in transcript was used to normalize raw counts between TU-decay libraries.

4.4.3 mRNA half-life calculations

Spike-in normalized counts were used for voom mean-variance modeling and linear model fitting in **limma** to determine differential expression between time points and across all transcripts with respect to time. Relative, scaled transcript abundance (log₂ fold-changes) at time 0, 1 hr and 3 hrs were used to fit a linear model of decay,

transcripts with < 1 cpm at the initial time 0 were excluded from downstream analysis. Fit of the decay model was assessed by an adjusted R^2 value. Half-life was determined by the equation, $t_{1/2} = \ln 2/k$. Detailed methodology is provided in Section A.1).

Chapter 5

Synthesis

The work presented in this study provides for the first time, global *in vivo* models of RNA decay during development. These data give a wealth of information on patterns of post-transcriptional regulation of gene expression via transcript decay, highlighting the role of neural-specific regulation of mRNA decay. Transcripts that encode core housekeeping, metabolism and proliferation associated proteins are highly stable in the developing *Drosophila* embryo. Inspection of neural-specific decay dynamics in the same stage embryos, reveals is widespread destabilization of these core housekeeping, metabolism and proliferation associated mRNAs. These data suggest a distinct program of post-transcriptional regulation in developing neural tissues compared to all tissues. Specific factor regulation has been proposed as a major means of controlling specific transcript decay in many contexts. The discovery that Pumilio regulates a cohort of neural fate transcription factors and neural-development associated long 3'UTR isoforms is evidence for the specific regulation of mRNA by *trans*-acting factors. While these data provide compelling evidence, they do not explain the large scale shifts in mRNA stability seen between all tissues and neural tissues. The possibility that inherent features within transcripts (e.g. codon content) could regulate decay on this scale, drove the investigation into codon-mediated decay. In the whole embryo, multiple measures of codon bias confirm the same relationships seen in yeast and zebrafish; optimal codon content is indicative of mRNA stability. However, these effects are attenuated in neural tissues, suggesting there could be neural-specific codon

optimality. Investigation of codon bias related to neural mRNA stability did not reveal the same relationship between decay and codon optimality seen in the whole embryo. For example, genes that show the highest codon usage bias (i.e. ribosomal protein and non-canonical histones), had significant decreases in stability in neural tissues. These results could be indicative of a widespread attenuation of growth, evidenced by the downregulation of replication and translation associated mRNAs, in the developing nervous system. This would reflect the shifting program from proliferation to differentiation that is expected in the neural tissues observed in this study (i.e. prospero-expressing cells). The observation that the translation machinery itself (e.g. ribosomal protein coding mRNAs) are less stable may present a mechanism for global post-transcriptional regulation at both the level of mRNA stability and translation. This attenuation of codon associated effects on stability may allow for a more specific means of regulation by *trans*-acting factors, like Pumilio. Ultimately, the interplay between global regulatory mechanisms affecting decay and translation (i.e. codon optimality) and specific regulatory mechanisms mediated through *cis*-element and *trans*-acting factor recognition (e.g. Pum targeting long 3'UTR isoforms) determine the final protein expression pattern. In a context where codon-mediated decay is paramount, decay dynamics likely reflect a gene expression program tuned to meet the demands of growth and proliferation. Alternately, in neural tissues, where codon-mediated decay is attenuated, decay dynamics likely reflect a gene expression program tuned for the demands of neural development and differentiation. The mechanisms driving the attenuation of codon-mediated decay in neural tissues is still not known, but could be related to ribosome and charged tRNA pools, which will require further study to confirm. Assessing charged tRNA expression levels and translation dynamics (through polysome analysis) in neural tissues can provide clues to the mechanisms controlling context-specific codon optimality, and are part of future plans to expand on this work.

Development of tools to model decay in a more unbiased manner is important for extending the application of the TU-decay assay for measuring transcriptome-wide RNA decay. RNA-seq provides a flexible and sensitive means of quantifying RNA abundance, but development of a powerful analysis approach is needed to account

for sources of noise and bias in these measurements. Voom mean-variance modeling and `limma` linear modeling tools provide a robust and flexible framework for modeling transcriptome-wide RNA decay. These methods provide enhanced sensitivity to detect low-abundance, poorly annotated, and non-coding transcripts, and provide a statistical framework for assessing differential decay in complex experimental designs. Although only whole embryo TU-decay via RNA-seq data are provided as a proof of concept, these methods will be useful in assessing RNA stability across different tissues or treatments.

Appendix A

Appendix

A.1 Supplemental methods

A.1.1 Western blot

Embryos were dechorinated in bleach for 3 minutes, washed with PBS and homogenized in protein extraction buffer [50mM HEPES (pH 7.5), 150mM NaCl, 10% glycerol, 1% Triton X-100, 1.5mM MgCl, 1mM EGTA, 1 mini tab cOmplete protease inhibitor cocktail (Roche)]. Homogenate was cleared by centrifugation three times, quantified by Bradford Protein Assay (Thermo) and run in protein loading buffer [1X LDS, 100mM DTT] on a 4-20% Tris-Glycine polyacrylamide gel (Thermo) and transferred to a Hybond nitrocellulose membrane (GE Healthcare Life Sciences) according to manufacturers recommendations. Membranes were washed in TBST buffer [50mM Tris HCL (pH 7.5), 150 mM NaCl, 0.05% TWEEN-20] and blocked in 5% milk in TBST. Primary rat anti-Pum (1:200) and rat anti- Tubulin (1:300) were applied in blocking buffer overnight at 4°C. Secondary HRP conjugated anti-rat secondary antibodies were applied at 1:5000 (Pumilio) and 1:1000 (Tubulin) for 2 hours at room temperature. Blots were developed using SuperSignal West Pico Chemiluminescent Substrate (Thermo) for 5 minutes and imaged on a ChemiDoc MP System (BioRad).

A.1.2 RNA-seq analysis pipeline

Read mapping and summarization

All read mapping and summarization procedures can be carried out in R using the **Rsubread** package. The first step requires indexing the reference genome:

```
>buildindex(basename="dmelr5.25_index",
indexSplit=FALSE,
gappedIndex=FALSE,
memory=200,
reference="genome.fa" )
```

Read mapping is performed using the **subjunc** function, which takes into account informative exon-exon spanning reads when mapping to transcript isoforms. This function also produces the corresponding BED junction file which can be used for visualization of read coverage over the genome.

```
subjunc(index="dmelr5.25_index",
phredOffset=66,
readfile1="WE1P 1_1.fq.gz",
readfile2="WE1P 1_2.fq.gz",
PE_orientation="rf",
nthreads=4,
input_format="gzFASTQ",
output_file="WE1P 1.bam" )
```

Read summarization can then be performed on the mapped read files in SAM/BAM format.

```
fc <- featureCounts(
files=c("WE1P 1.bam", "WE1P 2.bam", "WE1C 1.bam",
"WE1C 2.bam", "WE3C 1.bam", "WE3C 2.bam"),
annot.ext="genes.gtf",
isPairedEnd=TRUE,
isGTFAnnotationFile=TRUE,
GTF.featureType="exon",
GTF.attrType="transcript_id",
countMultiMappingReads=TRUE,
strandSpecific=2,
allowMultiOverlap=TRUE,
ignoreDup=TRUE )
```

Voom mean-variance modeling and limma linear modeling of decay

A design must be specified for designating replicates and making comparisons in `limma`. The first design shown below is for a time course jittering the time variable around the actual time since `limma` is not designed to handle replicates in this type of design.

```
>
>time <- c(
  jitter(0, factor=100, amount=.003),
  jitter(1, factor=100, amount=.003),
  jitter(3, factor=100, amount=.003),
  jitter(0, factor=100, amount=.003),
  jitter(1, factor=100, amount=.003),
  jitter(3, factor=100, amount=.003) )
>
>trt <- factor( rep( c(0,1), each=3) )
>
>design <- model.matrix(~time+trt)
>
> design
      (Intercept)          time trt1
1             1 0.0005360065      0
2             1 1.0024468810      0
3             1 2.9980313689      0
4             1 0.0020590678      1
5             1 1.0014600015      1
6             1 3.0008571724      1
attr(,"assign")
[1] 0 1 2
attr(,"contrasts")
attr(,"contrasts")\ $trt
[1] "contr.treatment"
>
```

The following is the setup for the contrasts design, which requires a targets frame, `targets.txt`.

File	Target
1	wt.0 hr
2	wt.0 hr
3	wt.1 hr
4	wt.1 hr
5	wt.3 hr
6	wt.3 hr

```

>targets <- readTargets()
>
>lev <- c("wt.0 hr","wt.1 hr","wt.3 hr")
>
>f <- factor(targets$Target, levels=lev)
>
>design <- model.matrix(~0+f)
>
>colnames(design) <- lev
>
> design
      fwt.0 hr fwt.1 hr fwt.3 hr
1           1          0          0
2           1          0          0
3           0          1          0
4           0          1          0
5           0          0          1
6           0          0          1
attr(,"assign")
[1] 1 1 1
attr(,"contrasts")
attr(,"contrasts")$f
[1] "contr.treatment"
>

```

Next, spike-in normalized raw count data are imported into R, where voom can model mean-variance and assign a predicted count. Optional normalization procedures in voom are designed for minimizing log-fold changes between samples, which is what we expect to see in the TU-decay data, so no additional normalization procedures are selected.

```
>
>v <- voom(expression_data , design , plot=TRUE)
>
```

limma's `lmFit` function performs the initial linear (least squares) fit for replicates in the contrasts design, or for the data over the time course.

```
>
>fit <- lmFit(v, design , method="ls")
>
```

`eBayes` calculates log-odds of differential expression by empirical Bayes moderation of standard errors. `topTable` summarizes the output of `eBayes`, providing p values for differential expression and optional gene ranking.

```
>
>fit2 <- eBayes(fit)
>
>z <- topTable(fit1 , coef="time", sort.by="none", number=34718)
>
```

For the contrasts design, an additional fit is required before running `eBayes` and summarizing with `topTable`.

```
>cont.t <- makeContrasts("wt.0 hr wt.1 hr", "wt.1 hr wt.3 hr", "wt.0 hr wt.3 hr",
  levels=design)
>
>fit1 <- contrasts.fit(fit , cont.t)
>
>fit2 <- eBayes(fit1)
>
>z <- topTable(fit2 , sort.by="none", number=34718)
>
```

Data from all the fits can be exported along with gene annotation data for downstream processing.

A.2 Supplemental data

Primers		
Hsp70-attB F	cloning	AAAAAAAGATCTTCTAGTAGGACGTCTATCGATACCG
Hsp70-attB R	cloning	AAAAAAAGATCTGTTCGACATGCCCCGCGTGACC
Hey F	cloning	AAAAAAGCGGCCGCGTAGTTGTGGGCATAAGCCTCAGA
Hey R	cloning	GGGGGCTCGAGGGAGGCCGAAATGCCATTGTA
RPS9 F	cloning	AAAAAAGCGGCCGCGTAGGCGACAACATCGTCGCCAGGC
RPS9 R	cloning	GGGGGCTCGAGAACAACGTGGACCCAAGAGG
GFP F	qPCR	AGCGTTCAACTAGCAGACCA
GFP R	qPCR	TGCCATGTGTAATCCCAGCA
3' generic	qPCR	ACCGTGACCTACATCGTCGA
5' Syb WT	qPCR	GCGTGTGAACGTGGAGAAGG
5' Syb O	qPCR	CGAGAAGGTCTTAGAGCGTG
5' RpS15Aa WT	qPCR	AGGCTAAACAAGTGCGGCGT
5' RpS15Aa NO	qPCR	TAAATTTGACGGGGCGGCTA
FLuc F	qPCR	ACTGGGACGAAGACGAACAC
FLuc R	qPCR	GGCGACGTAATCCACGATCT

Table A.1: Custom designed primers used in the study.

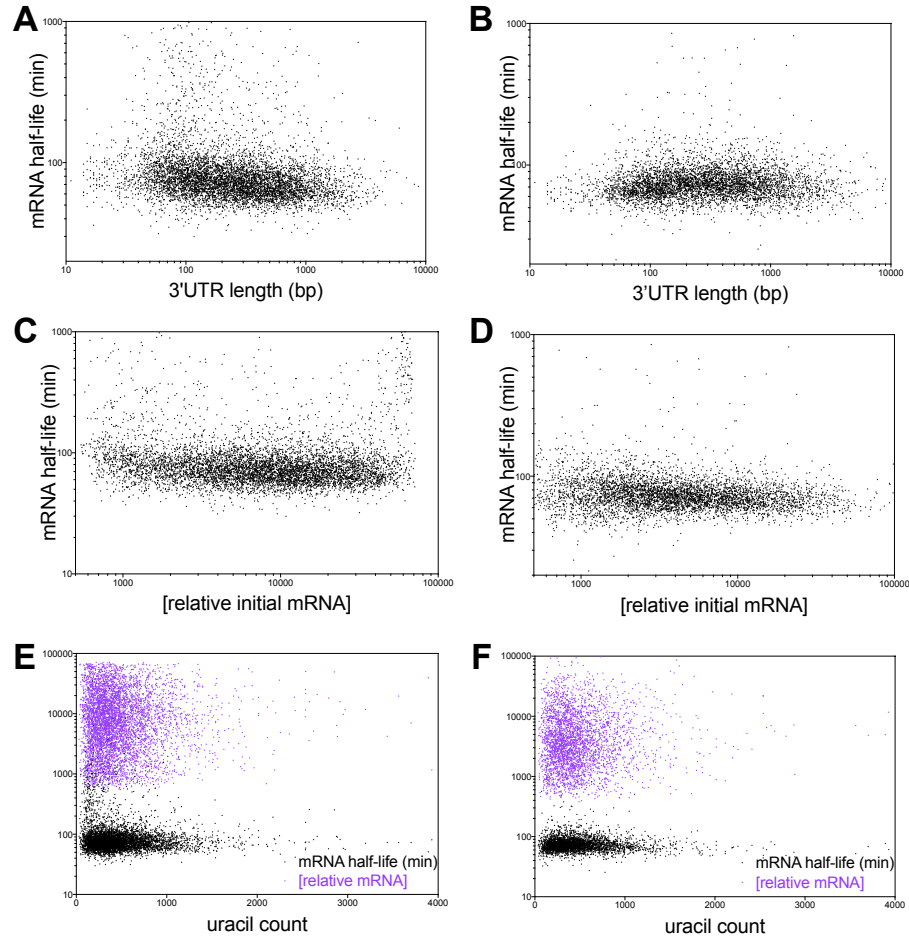


Figure A.1: 3'UTR length (bp) plotted as a function of mean mRNA half-life (min) for whole embryos (A) and the nervous system (B). Initial relative mRNA concentration plotted as a function of mean mRNA half-life (min) for whole embryos (C) and the nervous system (D). Uracil count per mRNA transcript plotted as a function of mean initial relative mRNA concentration (purple) or mean mRNA half-life (black) for whole embryos (E) and the nervous system (F).

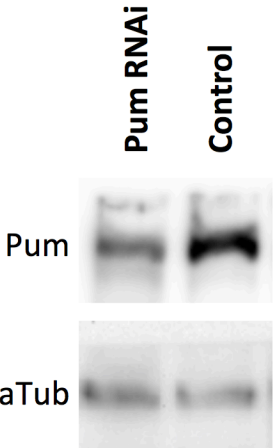


Figure A.2: Decreased *Pumilio* expression in stage 12-15 *lola(Ubiquitous)-GAL4 ; UAS-PumRNAi* embryos were used for protein extraction and western blot against Pumilio and α Tubulin.

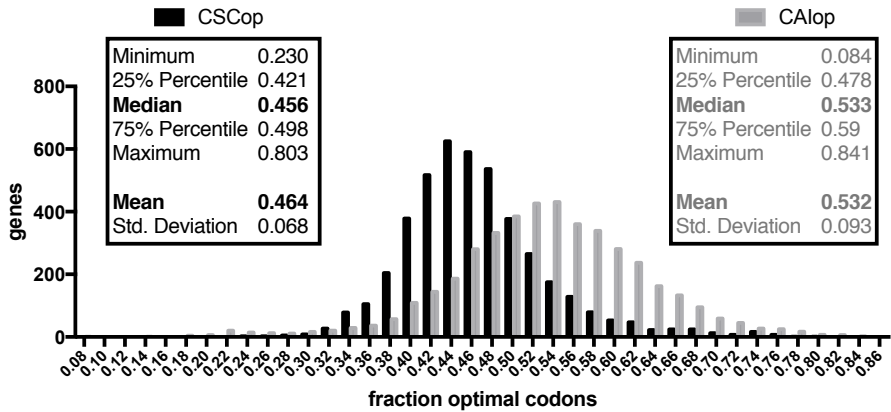


Figure A.3: The distributions of measures of codon bias, CSC and CAI are shown with the corresponding statistics. Although distributions are significantly different, gene ranks and ratios between the data are correlated.

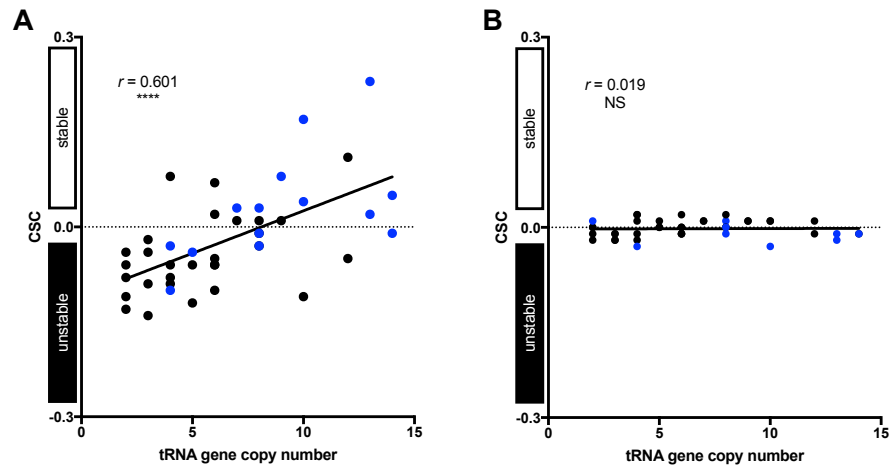


Figure A.4: Correlations between tRNA gene copy number and CSC in the whole embryo (A) embryonic nervous system (B), Pearson r values are indicated and **** designates $p < 0.0001$. Codons designated in blue are preferred for *Drosophila melanogaster* [168, 178, 180, 181].

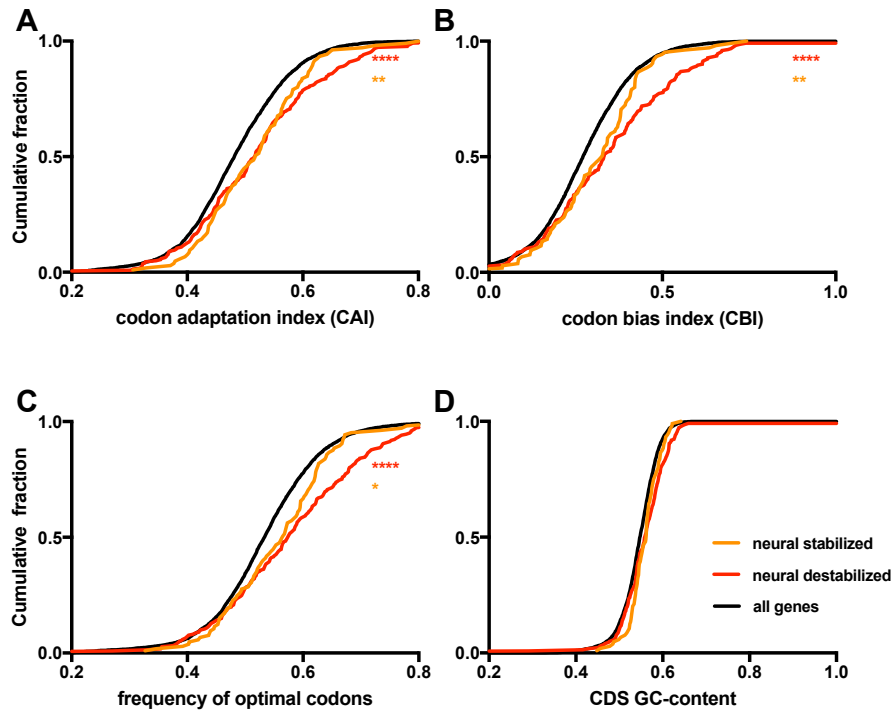


Figure A.5: Cumulative distributions of codon adaptation index (CAI) (A), GC content (B) for neural stabilized (orange), neural destabilized (red) or all genes (gray), **** $p < 0.0001$, ** $p < 0.001$, * $p < 0.01$ by Kruskal-Wallis test corrected for multiple comparisons.

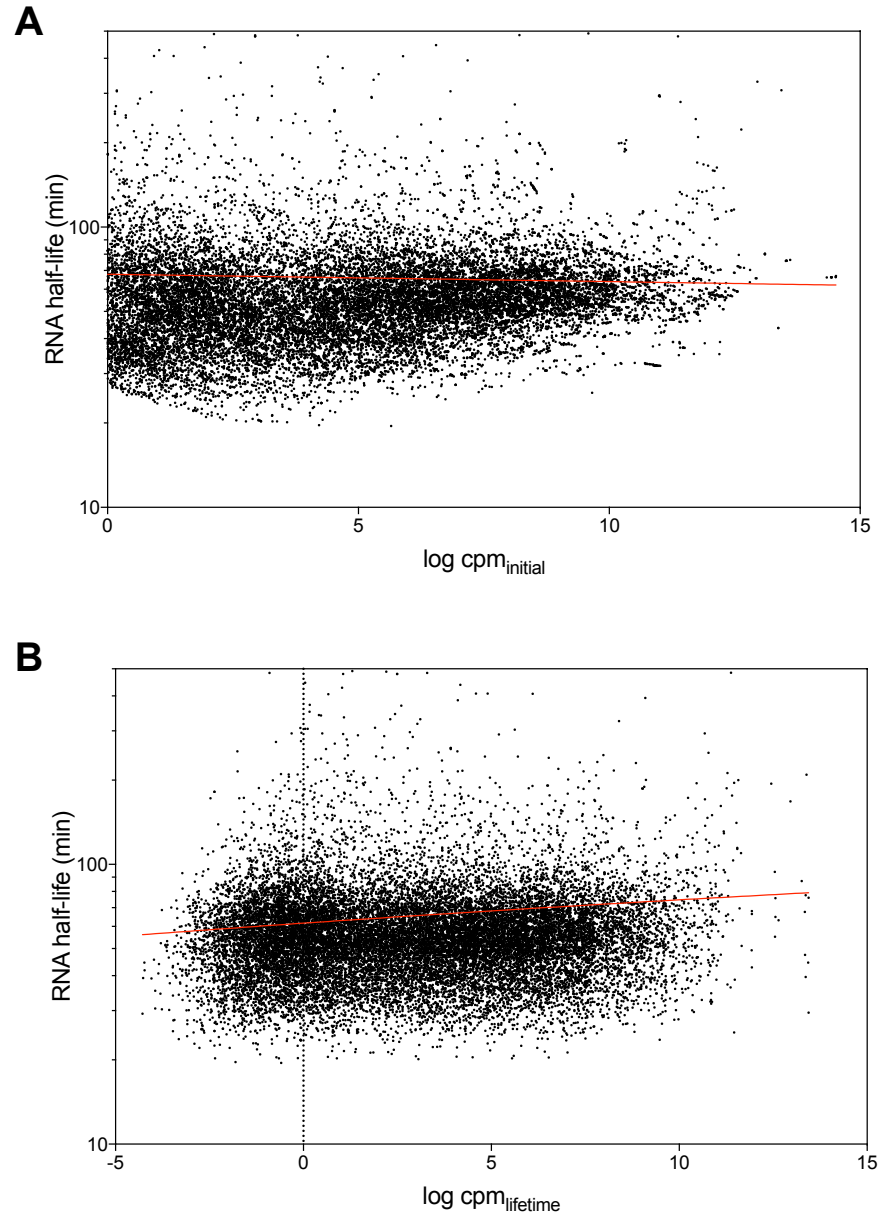


Figure A.6: Initial (A) and lifetime (B) normalized abundance are plotted against corresponding RNA half-life. A linear fit is shown although there is no significant correlation observed.

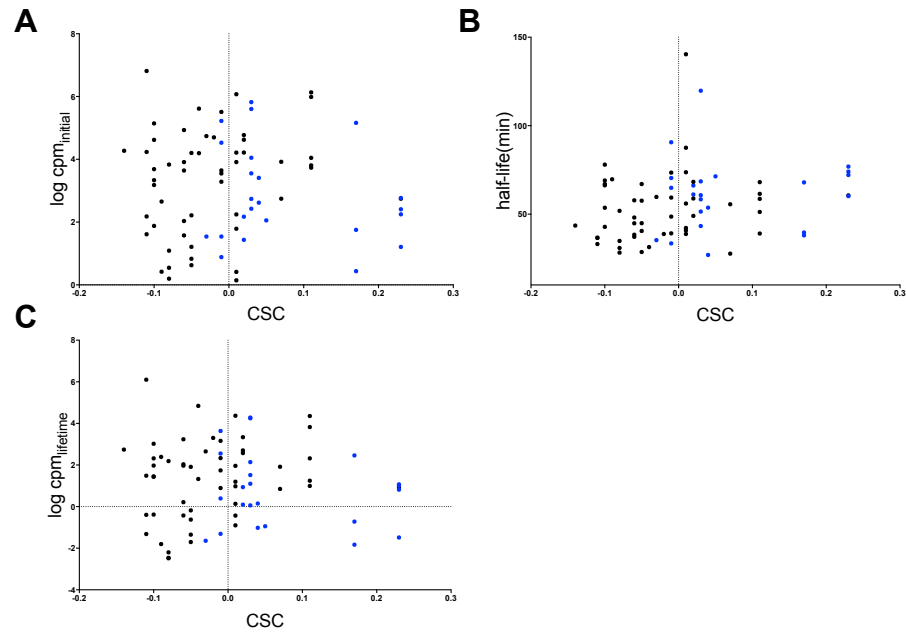


Figure A.7: Initial (A) and lifetime (B) normalized tRNA abundance and tRNA decay (C) are plotted against corresponding codon optimality by CSC. Codons designated in blue are preferred for *Drosophila melanogaster* [168, 178, 180, 181]. No significant trends are observed.

Bibliography

- [1] Mark Schena, Dari Shalon, Ronald W Davis, and Patrick O Brown. Quantitative monitoring of gene expression patterns with a complementary DNA microarray. *Science*, 270(5235):467–470, 1995.
- [2] Ali Mortazavi, Brian A Williams, Kenneth McCue, Lorian Schaeffer, and Barbara Wold. Mapping and quantifying mammalian transcriptomes by RNA-Seq. *Nature methods*, 5(7):621–628, 2008.
- [3] Vishwanath R Iyer, Christine E Horak, Charles S Scafe, David Botstein, Michael Snyder, and Patrick O Brown. Genomic binding sites of the yeast cell-cycle transcription factors SBF and MBF. *Nature*, 409(6819):533–538, 2001.
- [4] Claudio R Alonso. A complex ‘mRNA degradation code’ controls gene expression during animal development. *Trends in Genetics*, 28(2):78–88, 2012.
- [5] Michael D Cleary, Christopher D Meiering, Eric Jan, Rebecca Guymon, and John C Boothroyd. Biosynthetic labeling of RNA with uracil phosphoribosyltransferase allows cell-specific microarray analysis of mRNA synthesis and decay. *Nature biotechnology*, 23(2):232–237, 2005.
- [6] Leslie Gay, Michael R Miller, P Britten Ventura, Vidusha Devasthali, Zer Vue, Heather L Thompson, Sally Temple, Hui Zong, Michael D Cleary, Kryn Stankunas, et al. Mouse TU tagging: a chemical/genetic intersectional method for purifying cell type-specific nascent RNA. *Genes & development*, 27(1):98–115, 2013.

- [7] Dana A Burow, Maxine C Umeh-Garcia, Marie B True, Crystal D Bakhaj, David H Ardell, and Michael D Cleary. Dynamic regulation of mRNA decay during neural development. *Neural development*, 10(1):11, 2015.
- [8] Allan Jacobson. 16 poly (A) metabolism and translation: The closed-loop model. *Cold Spring Harbor Monograph Archive*, 30:451–480, 1996.
- [9] Alan Sachs. 10 physical and functional interactions between the mRNA cap structure and the poly (A) tail. *Cold Spring Harbor Monograph Archive*, 39:447–465, 2000.
- [10] Christine E Brown, SZ Tarun, Ronald Boeck, and Alan B Sachs. PAN3 encodes a subunit of the Pab1p-dependent poly (A) nuclease in *saccharomyces cerevisiae*. *Molecular and cellular biology*, 16(10):5744–5753, 1996.
- [11] Akio Yamashita, Tsung-Cheng Chang, Yukiko Yamashita, Wenmiao Zhu, Zhenping Zhong, Chyi-Ying A Chen, and Ann-Bin Shyu. Concerted action of poly (A) nucleases and decapping enzyme in mammalian mRNA turnover. *Nature structural & molecular biology*, 12(12):1054–1063, 2005.
- [12] Morgan Tucker, Robin R Staples, Marco A Valencia-Sanchez, Denise Muhlrud, and Roy Parker. Ccr4p is the catalytic subunit of a Ccr4p/Pop2p/Notp mRNA deadenylase complex in *saccharomyces cerevisiae*. *The EMBO journal*, 21(6):1427–1436, 2002.
- [13] Michelle Steiger, Anne Carr-Schmid, David C Schwartz, Megerditch Kiledjian, and ROY Parker. Analysis of recombinant yeast decapping enzyme. *Rna*, 9(2):231–238, 2003.
- [14] Wenqian Hu, Thomas J Sweet, Sangpen Chamnongpol, Kristian E Baker, and Jeff Collier. Co-translational mRNA decay in *saccharomyces cerevisiae*. *Nature*, 461(7261):225–229, 2009.

- [15] Vladimir Presnyak, Najwa Alhusaini, Ying-Hsin Chen, Sophie Martin, Nathan Morris, Nicholas Kline, Sara Olson, David Weinberg, Kristian E Baker, Brenton R Graveley, et al. Codon optimality is a major determinant of mRNA stability. *Cell*, 160(6):1111–1124, 2015.
- [16] Yuichiro Mishima and Yukihide Tomari. Codon usage and 3'UTR length determine maternal mRNA stability in zebrafish. *Molecular Cell*, 61(6):874–885, 2016.
- [17] Nicole L Garneau, Jeffrey Wilusz, and Carol J Wilusz. The highways and byways of mRNA decay. *Nature reviews Molecular cell biology*, 8(2):113–126, 2007.
- [18] Jidong Liu, Michelle A Carmell, Fabiola V Rivas, Carolyn G Marsden, J Michael Thomson, Ji-Joon Song, Scott M Hammond, Leemor Joshua-Tor, and Gregory J Hannon. Argonaute 2 is the catalytic engine of mammalian RNAi. *Science*, 305(5689):1437–1441, 2004.
- [19] Ji-Joon Song, Stephanie K Smith, Gregory J Hannon, and Leemor Joshua-Tor. Crystal structure of argonaute and its implications for RISC slicer activity. *science*, 305(5689):1434–1437, 2004.
- [20] David Gatfield and Elisa Izaurralde. Nonsense-mediated messenger RNA decay is initiated by endonucleolytic cleavage in drosophila. *Nature*, 429(6991):575–578, 2004.
- [21] Audrey Stevens, Yang Wang, Kirsten Bremer, Jing Zhang, Robert Hoepfner, Michael Antoniou, Daniel R Schoenberg, and Lynne E Maquat. β -globin mRNA decay in erythroid cells: Ug site-preferred endonucleolytic cleavage that is augmented by a premature termination codon. *Proceedings of the National Academy of Sciences*, 99(20):12741–12746, 2002.
- [22] Meenakshi K Doma and Roy Parker. RNA quality control in eukaryotes. *Cell*, 131(4):660–668, 2007.

- [23] David Gatfield, Leonie Unterholzner, Francesca D Ciccarelli, Peer Bork, and Elisa Izaurralde. Nonsense-mediated mRNA decay in drosophila: at the intersection of the yeast and mammalian pathways. *The EMBO journal*, 22(15):3960–3970, 2003.
- [24] Marc Bühler, Silvia Steiner, Fabio Mohn, Alexandra Paillusson, and Oliver Mühlemann. EJC-independent degradation of nonsense immunoglobulin- μ mRNA depends on 3'UTR length. *Nature structural & molecular biology*, 13(5):462–464, 2006.
- [25] Nadia Amrani, Robin Ganesan, Stephanie Kervestin, David A Mangus, Shubendu Ghosh, and Allan Jacobson. A faux 3'UTR promotes aberrant termination and triggers nonsense-mediated mRNA decay. *Nature*, 432(7013):112–118, 2004.
- [26] Joshua T Mendell, Neda A Sharifi, Jennifer L Meyers, Francisco Martinez-Murillo, and Harry C Dietz. Nonsense surveillance regulates expression of diverse classes of mammalian transcripts and mutes genomic noise. *Nature genetics*, 36(10):1073–1078, 2004.
- [27] Handan Kaygun and William F Marzluff. Regulated degradation of replication-dependent histone mRNAs requires both ATR and Upf1. *Nature structural & molecular biology*, 12(9):794–800, 2005.
- [28] Yoon Ki Kim, Luc Furic, Luc DesGroseillers, and Lynne E Maquat. Mammalian Staufen1 recruits Upf1 to specific mRNA 3'UTRs so as to elicit mRNA decay. *Cell*, 120(2):195–208, 2005.
- [29] Ujwal Sheth and Roy Parker. Decapping and decay of messenger RNA occur in cytoplasmic processing bodies. *Science*, 300(5620):805–808, 2003.
- [30] Maria A Ferraiuolo, Sanjukta Basak, Josee Dostie, Elizabeth L Murray, Daniel R Schoenberg, and Nahum Sonenberg. A role for the eIF4E-binding protein 4E-T in P-body formation and mRNA decay. *The Journal of cell biology*, 170(6):913–924, 2005.

- [31] Maria Alexandra Andrei, Dierk Ingelfinger, Heintzmann Rainer, Tilmann Achel, Rolando Rivera-Pomar, and Reinhard Lührmann. A role for eIF4E and eIF4E-transporter in targeting mRNPs to mammalian processing bodies. *Rna*, 11(5):717–727, 2005.
- [32] Nicolas Cougot, Sylvie Babajko, and Bertrand Séraphin. Cytoplasmic foci are sites of mRNA decay in human cells. *The Journal of cell biology*, 165(1):31–40, 2004.
- [33] Vladimir I Bashkirov, Harry Scherthan, Jachen A Solinger, Jean-Marie Buerstedde, and Wolf-Dietrich Heyer. A mouse cytoplasmic exoribonuclease (mXRN1p) with preference for G4 tetraplex substrates. *The Journal of cell biology*, 136(4):761–773, 1997.
- [34] Leonie Unterholzner and Elisa Izaurralde. SMG7 acts as a molecular link between mRNA surveillance and mRNA decay. *Molecular cell*, 16(4):587–596, 2004.
- [35] Ujwal Sheth and Roy Parker. Targeting of aberrant mRNAs to cytoplasmic processing bodies. *Cell*, 125(6):1095–1109, 2006.
- [36] George L Sen and Helen M Blau. Argonaute 2/RISC resides in sites of mammalian mrna decay known as cytoplasmic bodies. *Nature cell biology*, 7(6):633–636, 2005.
- [37] Isabelle Behm-Ansmant, Jan Rehwinkel, Tobias Doerks, Alexander Stark, Peer Bork, and Elisa Izaurralde. mrna degradation by miRNAs and GW182 requires both CCR4: NOT deadenylase and DCP1: DCP2 decapping complexes. *Genes & development*, 20(14):1885–1898, 2006.
- [38] Nancy Kedersha, Georg Stoecklin, Maranatha Ayodele, Patrick Yacono, Jens Lykke-Andersen, Marvin J Fritzler, Donalyn Scheuner, Randal J Kaufman, David E Golan, and Paul Anderson. Stress granules and processing bodies are dynamically linked sites of mRNP remodeling. *The Journal of cell biology*, 169(6):871–884, 2005.

- [39] Valérie Hilgers, Daniela Teixeira, and Roy Parker. Translation-independent inhibition of mRNA deadenylation during stress in *saccharomyces cerevisiae*. *Rna*, 12(10):1835–1845, 2006.
- [40] Paola Briata, Sonia Vanina Forcales, Marco Ponassi, Giorgio Corte, Ching-Yi Chen, Michael Karin, Pier Lorenzo Puri, and Roberto Gherzi. p38-dependent phosphorylation of the mRNA decay-promoting factor KSRP controls the stability of select myogenic transcripts. *Molecular cell*, 20(6):891–903, 2005.
- [41] Angélica Figueroa, Ana Cuadrado, Jinshui Fan, Ulus Atasoy, George E Muscat, Pura Munoz-Canoves, Myriam Gorospe, and Alberto Munoz. Role of HuR in skeletal myogenesis through coordinate regulation of muscle differentiation genes. *Molecular and Cellular Biology*, 23(14):4991–5004, 2003.
- [42] Zhong-Jian Shen, Stephane Esnault, and James S Malter. The peptidyl-prolyl isomerase pin1 regulates the stability of granulocyte-macrophage colony-stimulating factor mRNA in activated eosinophils. *Nature immunology*, 6(12):1280–1287, 2005.
- [43] Ching-Yi Chen, Fabienne Del Gatto-Konczak, Zhenguo Wu, and Michael Karin. Stabilization of interleukin-2 mRNA by the c-Jun NH2-terminal kinase pathway. *Science*, 280(5371):1945–1949, 1998.
- [44] Ching-Yi Chen, Roberto Gherzi, Jens S Andersen, Guido Gaietta, Karsten Jürchott, Hans-Dieter Royer, Matthias Mann, and Michael Karin. Nucleolin and YB-1 are required for JNK-mediated interleukin-2 mRNA stabilization during T-cell activation. *Genes & development*, 14(10):1236–1248, 2000.
- [45] Paola Briata, Cristina Ilengo, Giorgio Corte, Christoph Moroni, Michael G Rosenfeld, Ching-Yi Chen, and Roberto Gherzi. The Wnt/ β -catenin Pitx2 pathway controls the turnover of Pitx2 and other unstable mRNAs. *Molecular cell*, 12(5):1201–1211, 2003.

- [46] Khalid SA Khabar. The AU-rich transcriptome: more than interferons and cytokines, and its role in disease. *Journal of interferon & cytokine research*, 25(1):1–10, 2005.
- [47] Michael Parisi and Haifan Lin. Translational repression: a duet of Nanos and Pumilio. *Current Biology*, 10(2):R81–R83, 2000.
- [48] Aaron C Goldstrohm, Brad A Hook, Daniel J Seay, and Marvin Wickens. PUF proteins bind Pop2p to regulate messenger RNAs. *Nature structural & molecular biology*, 13(6):533–539, 2006.
- [49] Chase A Weidmann, Nathan A Raynard, Nathan H Blewett, Jamie Van Etten, and Aaron C Goldstrohm. The rna binding domain of Pumilio antagonizes polyadenosine binding protein and accelerates deadenylation. *RNA*, 20(8):1298–1319, 2014.
- [50] Jinshui Fan, Xiaoling Yang, Wengong Wang, William H Wood, Kevin G Becker, and Myriam Gorospe. Global analysis of stress-regulated mRNA turnover by using cDNA arrays. *Proceedings of the National Academy of Sciences*, 99(16):10611–10616, 2002.
- [51] Sarah E Munchel, Ryan K Shultzaberger, Naoki Takizawa, and Karsten Weis. Dynamic profiling of mRNA turnover reveals gene-specific and system-wide regulation of mRNA decay. *Molecular biology of the cell*, 22(15):2787–2795, 2011.
- [52] Chris Cheadle, Jinshui Fan, Yoon S Cho-Chung, Thomas Werner, Jill Ray, Lana Do, Myriam Gorospe, and Kevin G Becker. Control of gene expression during T cell activation: alternate regulation of mRNA transcription and mRNA stability. *BMC genomics*, 6(1):75, 2005.
- [53] Graydon B Gonsalvez and Roy M Long. Spatial regulation of translation through RNA localization. *F1000 biology reports*, 4, 2012.

- [54] Caroline Medioni, Kimberly Mowry, and Florence Besse. Principles and roles of mRNA localization in animal development. *Development*, 139(18):3263–3276, 2012.
- [55] Ivone G Bruno, Rachid Karam, Lulu Huang, Anjana Bhardwaj, Chih H Lou, Eleen Y Shum, Hye-Won Song, Mark A Corbett, Wesley D Gifford, Jozef Gecz, et al. Identification of a microRNA that activates gene expression by repressing nonsense-mediated RNA decay. *Molecular cell*, 42(4):500–510, 2011.
- [56] Antonia Ratti, Claudia Fallini, Lidia Cova, Roberto Fantozzi, Cinzia Calzarossa, Eleonora Zennaro, Alessia Pascale, Alessandro Quattrone, and Vincenzo Silani. A role for the ELAV RNA-binding proteins in neural stem cells: stabilization of Msi1 mRNA. *Journal of cell science*, 119(7):1442–1452, 2006.
- [57] Hosung Jung, Christos G Gkogkas, Nahum Sonenberg, and Christine E Holt. Remote control of gene function by local translation. *Cell*, 157(1):26–40, 2014.
- [58] Dilek Colak, Sheng-Jian Ji, Bo T Porse, and Samie R Jaffrey. Regulation of axon guidance by compartmentalized nonsense-mediated mRNA decay. *Cell*, 153(6):1252–1265, 2013.
- [59] A Ashleigh Long, Cecon T Mahapatra, Elvin A Woodruff, Jeff Rohrbough, Hung-Tat Leung, Shikoh Shino, Lingling An, Rebecca W Doerge, Mark M Metzstein, William L Pak, et al. The nonsense-mediated decay pathway maintains synapse architecture and synaptic vesicle cycle efficacy. *Journal of cell science*, 123(19):3303–3315, 2010.
- [60] Juan Ji An, Kusumika Gharami, Guey-Ying Liao, Newton H Woo, Anthony G Lau, Filip Vanevski, Enrique R Torre, Kevin R Jones, Yue Feng, Bai Lu, et al. Distinct role of long 3'UTR BDNF mRNA in spine morphology and synaptic plasticity in hippocampal neurons. *Cell*, 134(1):175–187, 2008.
- [61] Sharon A Swanger and Gary J Bassell. Making and breaking synapses through local mRNA regulation. *Current opinion in genetics & development*, 21(4):414–421, 2011.

- [62] Nora Perrone-Bizzozero and Federico Bolognani. Role of HuD and other RNA-binding proteins in neural development and plasticity. *Journal of neuroscience research*, 68(2):121–126, 2002.
- [63] Todd D Levine, Fenbia Gao, PH King, LG Andrews, and Jack D Keene. Hel-N1: an autoimmune RNA-binding protein with specificity for 3'uridylate-rich untranslated regions of growth factor mRNAs. *Molecular and Cellular Biology*, 13(6):3494–3504, 1993.
- [64] Valérie Hilgers, Sandra B Lemke, and Michael Levine. ELAV mediates 3'UTR extension in the drosophila nervous system. *Genes & development*, 26(20):2259–2264, 2012.
- [65] Peter Smibert, Pedro Miura, Jakub O Westholm, Sol Shenker, Gemma May, Michael O Duff, Dayu Zhang, Brian D Eads, Joe Carlson, James B Brown, et al. Global patterns of tissue-specific alternative polyadenylation in drosophila. *Cell reports*, 1(3):277–289, 2012.
- [66] Valérie Hilgers, Michael W Perry, David Hendrix, Alexander Stark, Michael Levine, and Benjamin Haley. Neural-specific elongation of 3'UTRs during drosophila development. *Proceedings of the National Academy of Sciences*, 108(38):15864–15869, 2011.
- [67] Shannon Farris, Gail Lewandowski, Conor D Cox, and Oswald Steward. Selective localization of arc mRNA in dendrites involves activity-and translation-dependent mRNA degradation. *The Journal of Neuroscience*, 34(13):4481–4493, 2014.
- [68] Walter J Lukiw, Tatiana V Andreeva, Anastasia P Grigorenko, and Evgeny I Rogaev. Studying micro-RNA function and dysfunction in alzheimer's disease. *Frontiers in genetics*, 3:327, 2013.
- [69] Elysia M Hollams, Keith M Giles, Andrew M Thomson, and Peter J Leedman. mRNA stability and the control of gene expression: implications for human disease. *Neurochemical research*, 27(10):957–980, 2002.

- [70] Magdalini Polymenidou, Clotilde Lagier-Tourenne, Kasey R Hutt, C Frank Bennett, Don W Cleveland, and Gene W Yeo. Misregulated RNA processing in amyotrophic lateral sclerosis. *Brain research*, 1462:3–15, 2012.
- [71] Xia-Lian Xu, Yan Li, Fay Wang, and Fen-Biao Gao. The steady-state level of the nervous-system-specific microRNA-124a is regulated by dFMR1 in drosophila. *The Journal of Neuroscience*, 28(46):11883–11889, 2008.
- [72] Francesca Sartor, Jihan Anderson, Colin McCaig, Zosia Miedzybrodzka, and Berndt Müller. Mutation of genes controlling mRNA metabolism and protein synthesis predisposes to neurodevelopmental disorders. *Biochemical Society Transactions*, 43(6):1259–1265, 2015.
- [73] John P Cogswell, James Ward, Ian A Taylor, Michelle Waters, Yunling Shi, Brian Cannon, Kevin Kelnar, Jon Kemppainen, David Brown, Caifu Chen, et al. Identification of miRNA changes in Alzheimer’s disease brain and CSF yields putative biomarkers and insights into disease pathways. *Journal of Alzheimer’s disease*, 14(1):27–41, 2008.
- [74] Petra Leidinger, Christina Backes, Stephanie Deutscher, Katja Schmitt, Sabine C Mueller, Karen Frese, Jan Haas, Klemens Ruprecht, Friedemann Paul, Cord Stähler, et al. A blood based 12-miRNA signature of Alzheimer’s disease patients. *Genome biology*, 14(7):1, 2013.
- [75] Justin M Long, Balmiki Ray, and Debomoy K Lahiri. MicroRNA-339-5p down-regulates protein expression of β -site amyloid precursor protein-cleaving enzyme 1 (bace1) in human primary brain cultures and is reduced in brain tissue specimens of alzheimer disease subjects. *Journal of Biological Chemistry*, 289(8):5184–5198, 2014.
- [76] Justin M Long, Balmiki Ray, and Debomoy K Lahiri. MicroRNA-153 physiologically inhibits expression of amyloid- β precursor protein in cultured human fetal brain cells and is dysregulated in a subset of alzheimer disease patients. *Journal of Biological Chemistry*, 287(37):31298–31310, 2012.

- [77] Jaekwang Kim, Hyejin Yoon, Takahiro Horie, Jack M Burchett, Jessica L Restivo, Noemi Rotllan, Cristina M Ramírez, Philip B Verghese, Masafumi Ihara, Hyang-Sook Hoe, et al. MicroRNA-33 regulates ApoE lipidation and amyloid- β metabolism in the brain. *The Journal of Neuroscience*, 35(44):14717–14726, 2015.
- [78] Pierre Lau, Koen Bossers, Evgenia Salta, Carlo Sala Frigerio, Shahar Barbash, Roy Rothman, Annerieke SR Sierksma, Amantha Thathiah, David Greenberg, Aikaterini S Papadopoulou, et al. Alteration of the microRNA network during the progression of Alzheimer’s disease. *EMBO molecular medicine*, 5(10):1613–1634, 2013.
- [79] Julia Banzhaf-Strathmann, Eva Benito, Stephanie May, Thomas Arzberger, Sabina Tahirovic, Hans Kretzschmar, André Fischer, and Dieter Edbauer. MicroRNA-125b induces tau hyperphosphorylation and cognitive deficits in Alzheimer’s disease. *The EMBO journal*, page e201387576, 2014.
- [80] Shuo-Chien Ling, Magdalini Polymenidou, and Don W Cleveland. Converging mechanisms in ALS and FTD: disrupted RNA and protein homeostasis. *Neuron*, 79(3):416–438, 2013.
- [81] Michael J Strong. The evidence for altered RNA metabolism in amyotrophic lateral sclerosis (ALS). *Journal of the neurological sciences*, 288(1):1–12, 2010.
- [82] Akihhide Koyama, Akihiro Sugai, Taisuke Kato, Tomohiko Ishihara, Atsushi Shiga, Yasuko Toyoshima, Misaki Koyama, Takuya Konno, Sachiko Hirokawa, Akio Yokoseki, et al. Increased cytoplasmic TARDBP mRNA in affected spinal motor neurons in ALS caused by abnormal autoregulation of TDP-43. *Nucleic acids research*, page gkw499, 2016.
- [83] Lam S Nguyen, Hyung-Goo Kim, Jill A Rosenfeld, Yiping Shen, James F Gusella, Yves Lacassie, Lawrence C Layman, Lisa G Shaffer, and Jozef Gécz. Contribution of copy number variants involving nonsense-mediated mrna decay

- pathway genes to neuro-developmental disorders. *Human molecular genetics*, 22(9):1816–1825, 2013.
- [84] Lachlan A Jolly, Claire C Homan, Reuben Jacob, Simon Barry, and Jozef Gecz. The UPF3B gene, implicated in intellectual disability, autism, ADHD and childhood onset schizophrenia regulates neural progenitor cell behaviour and neuronal outgrowth. *Human molecular genetics*, 22(23):4673–4687, 2013.
- [85] Luyuan Pan, Yong Q Zhang, Elvin Woodruff, and Kendal Broadie. The drosophila fragile X gene negatively regulates neuronal elaboration and synaptic differentiation. *Current Biology*, 14(20):1863–1870, 2004.
- [86] Cheryl L Gatto and Kendal Broadie. Temporal requirements of the fragile X mental retardation protein in the regulation of synaptic structure. *Development*, 135(15):2637–2648, 2008.
- [87] Olga Penagarikano, Jennifer G Mulle, and Stephen T Warren. The pathophysiology of fragile X syndrome. *Annu. Rev. Genomics Hum. Genet.*, 8:109–129, 2007.
- [88] Sharon Harrold, Christine Genovese, Barry Kobrin, Sherie L Morrison, and Christine Milcarek. A comparison of apparent mRNA half-life using kinetic labeling techniques vs decay following administration of transcriptional inhibitors. *Analytical biochemistry*, 198(1):19–29, 1991.
- [89] Jennifer L Shock, Kael F Fischer, and Joseph L DeRisi. Whole-genome analysis of mRNA decay in *Plasmodium falciparum* reveals a global lengthening of mRNA half-life during the intra-erythrocytic development cycle. *Genome Biol*, 8(7):R134, 2007.
- [90] Paul Loflin, Chyi-Ying A Chen, and Ann-Bin Shyu. Unraveling a cytoplasmic role for hnRNP D in the in vivo mRNA destabilization directed by the AU-rich element. *Genes & development*, 13(14):1884–1897, 1999.

- [91] Yulei Wang, Chih Long Liu, John D Storey, Robert J Tibshirani, Daniel Herschlag, and Patrick O Brown. Precision and functional specificity in mRNA decay. *Proceedings of the National Academy of Sciences*, 99(9):5860–5865, 2002.
- [92] CY Chen and Ann-Bin Shyu. Selective degradation of early-response-gene mRNAs: functional analyses of sequence features of the AU-rich elements. *Molecular and Cellular Biology*, 14(12):8471–8482, 1994.
- [93] Christian Seiser, Markus Posch, Nancy Thompson, et al. Effect of transcription inhibitors on the iron-dependent degradation of transferrin receptor mRNA. *Journal of Biological Chemistry*, 270(49):29400–29406, 1995.
- [94] Jeff Collier. Methods to determine mRNA half-life in *Saccharomyces cerevisiae*. *Methods in enzymology*, 448:267–284, 2008.
- [95] Lars Dölken, Zsolt Ruzsics, Bernd Rädle, Caroline C Friedel, Ralf Zimmer, Jörg Mages, Reinhard Hoffmann, Paul Dickinson, Thorsten Forster, Peter Ghazal, et al. High-resolution gene expression profiling for simultaneous kinetic parameter analysis of RNA synthesis and decay. *Rna*, 14(9):1959–1972, 2008.
- [96] Michael R Miller, Kristin J Robinson, Michael D Cleary, and Chris Q Doe. TU-tagging: cell type-specific RNA isolation from intact complex tissues. *Nature methods*, 6(6):439–441, 2009.
- [97] Michael D Cleary. Cell type-specific analysis of mRNA synthesis and decay in vivo with uracil phosphoribosyltransferase and 4-thiouracil. *Methods in enzymology*, 448:379–406, 2008.
- [98] Erin E Duffy, Michael Rutenberg-Schoenberg, Catherine D Stark, Robert R Kitchen, Mark B Gerstein, and Matthew D Simon. Tracking distinct RNA populations using efficient and reversible covalent chemistry. *Molecular cell*, 59(5):858–866, 2015.
- [99] Chyi-Ying A Chen and Ann-Bin Shyu. Mechanisms of deadenylation-dependent decay. *Wiley Interdisciplinary Reviews: RNA*, 2(2):167–183, 2011.

- [100] Dan Cao and Roy Parker. Computational modeling and experimental analysis of nonsense-mediated decay in yeast. *Cell*, 113(4):533–545, 2003.
- [101] SAVERIO BROGNA. Nonsense mutations in the alcohol dehydrogenase gene of drosophila melanogaster correlate with an abnormal 3' end processing of the corresponding pre-mRNA. *Rna*, 5(04):562–573, 1999.
- [102] Chyi-Ying A Chen, Dinghai Zheng, Zhenfang Xia, and Ann-Bin Shyu. Ago–TNRC6 triggers microRNA-mediated decay by promoting two deadenylation steps. *Nature structural & molecular biology*, 16(11):1160–1166, 2009.
- [103] Marc R Fabian, Géraldine Mathonnet, Thomas Sundermeier, Hansruedi Mathys, Jakob T Zipprich, Yuri V Svitkin, Fabiola Rivas, Martin Jinek, James Wohlschlegel, Jennifer A Doudna, et al. Mammalian miRNA RISC recruits CAF1 and PABP to affect PABP-dependent deadenylation. *Molecular cell*, 35(6):868–880, 2009.
- [104] Tim Wilson and Richard Treisman. Removal of poly (A) and consequent degradation of c-fos mRNA facilitated by 3' AU-rich sequences. 1988.
- [105] Ann-Bin Shyu, Joel G Belasco, and Michael E Greenberg. Two distinct destabilizing elements in the c-fos message trigger deadenylation as a first step in rapid mRNA decay. *Genes & Development*, 5(2):221–231, 1991.
- [106] Dinghai Zheng, Nader Ezzeddine, Chyi-Ying A Chen, Wenmiao Zhu, Xiangwei He, and Ann-Bin Shyu. Deadenylation is prerequisite for P-body formation and mRNA decay in mammalian cells. *The Journal of cell biology*, 182(1):89–101, 2008.
- [107] Claudia Temme, Sophie Zaessinger, Sylke Meyer, Martine Simonelig, and Elmar Wahle. A complex containing the CCR4 and CAF1 proteins is involved in mRNA deadenylation in drosophila. *The EMBO journal*, 23(14):2862–2871, 2004.

- [108] Akio Yamashita, Tsung-Cheng Chang, Yukiko Yamashita, Wenmiao Zhu, Zhenping Zhong, Chyi-Ying A Chen, and Ann-Bin Shyu. Concerted action of poly (A) nucleases and decapping enzyme in mammalian mRNA turnover. *Nature structural & molecular biology*, 12(12):1054–1063, 2005.
- [109] Carlus Deneke, Reinhard Lipowsky, and Angelo Valleriani. Complex degradation processes lead to non-exponential decay patterns and age-dependent decay rates of messenger RNA. *PloS one*, 8(2):e55442, 2013.
- [110] Andrea H Brand and Norbert Perrimon. Targeted gene expression as a means of altering cell fates and generating dominant phenotypes. *development*, 118(2):401–415, 1993.
- [111] Gabriela Galicia-Vázquez, Regina Cencic, Francis Robert, Aouod Quang Agenor, and Jerry Pelletier. A cellular response linking eIF4AI activity to eIF4AII transcription. *Rna*, 18(7):1373–1384, 2012.
- [112] CH Dani, JM Blanchard, M Piechaczyk, S El Sabouty, L Marty, and Ph Jean-teur. Extreme instability of myc mRNA in normal and transformed human cells. *Proceedings of the National Academy of Sciences*, 81(22):7046–7050, 1984.
- [113] Christian Miller, Björn Schwalb, Kerstin Maier, Daniel Schulz, Sebastian Dümcke, Benedikt Zacher, Andreas Mayer, Jasmin Sydow, Lisa Marcinowski, Lars Dölken, et al. Dynamic transcriptome analysis measures rates of mRNA synthesis and decay in yeast. *Molecular systems biology*, 7(1):458, 2011.
- [114] Hidenori Tani, Rena Mizutani, Kazi Abdus Salam, Keiko Tano, Kenichi Ijiri, Ai Wakamatsu, Takao Isogai, Yutaka Suzuki, and Nobuyoshi Akimitsu. Genome-wide determination of RNA stability reveals hundreds of short-lived noncoding transcripts in mammals. *Genome research*, 22(5):947–956, 2012.
- [115] Miguel C Manansala, Sarah Min, and Michael D. Cleary. The drosophila {SERTAD} protein taranis determines lineage-specific neural progenitor proliferation patterns. *Developmental biology*, 376(2):150–162, 2011.

- [116] Chih H Lou, Ada Shao, Eleen Y Shum, Josh L Espinoza, Lulu Huang, Rachid Karam, and Miles F Wilkinson. Posttranscriptional control of the stem cell and neurogenic programs by the nonsense-mediated RNA decay pathway. *Cell reports*, 6(4):748–764, 2014.
- [117] Evangelos D Karousis, Sofia Nasif, and Oliver Mühlemann. Nonsense-mediated mRNA decay: novel mechanistic insights and biological impact. *Wiley Interdisciplinary Reviews: RNA*, 2016.
- [118] Ashley T Neff, Ju Youn Lee, Jeffrey Wilusz, Bin Tian, and Carol J Wilusz. Global analysis reveals multiple pathways for unique regulation of mRNA decay in induced pluripotent stem cells. *Genome research*, 22(8):1457–1467, 2012.
- [119] Lioudmila V Sharova, Alexei A Sharov, Timur Nedorezov, Yulan Piao, Nabeebi Shaik, and Minoru SH Ko. Database for mRNA half-life of 19 977 genes obtained by DNA microarray analysis of pluripotent and differentiating mouse embryonic stem cells. *DNA research*, 16(1):45–58, 2009.
- [120] Paul T Spellman and Gerald M Rubin. Evidence for large domains of similarly expressed genes in the drosophila genome. *Journal of biology*, 1(1):1, 2002.
- [121] Jennifer M Lee and Erik LL Sonnhammer. Genomic gene clustering analysis of pathways in eukaryotes. *Genome research*, 13(5):875–882, 2003.
- [122] Laurence D Hurst, Csaba Pál, and Martin J Lercher. The evolutionary dynamics of eukaryotic gene order. *Nature Reviews Genetics*, 5(4):299–310, 2004.
- [123] Brody Thomas B. The interactive fly, 1995,1996.
- [124] Krishna H Zivraj, Yi Chun Loraine Tung, Michael Piper, Laura Gumy, James W Fawcett, Giles SH Yeo, and Christine E Holt. Subcellular profiling reveals distinct and developmentally regulated repertoire of growth cone mrnas. *The Journal of Neuroscience*, 30(46):15464–15478, 2010.

- [125] Jun Zhong, Theresa Zhang, and Lisa M Bloch. Dendritic mRNAs encode diversified functionalities in hippocampal pyramidal neurons. *BMC neuroscience*, 7(1):17, 2006.
- [126] Xin Li, Zhenqing Chen, and Claude Desplan. Temporal patterning of neural progenitors in drosophila. *Current topics in developmental biology*, 105:69, 2013.
- [127] Maria Monastirioti, Nikolaos Giagtzoglou, Konstantinos A Koumbanakis, Evanthia Zacharioudaki, Myrto Deligiannaki, Irmgard Wech, Mara Almeida, Anette Preiss, Sarah Bray, and Christos Delidakis. Drosophila hey is a target of notch in asymmetric divisions during embryonic and larval neurogenesis. *Development*, 137(2):191–201, 2010.
- [128] HJ Kim, HJ Ahn, S Lee, JH Kim, J Park, S-H Jeon, and SH Kim. Intrinsic dorsoventral patterning and extrinsic EGFR signaling genes control glial cell development in the drosophila nervous system. *Neuroscience*, 307:242–252, 2015.
- [129] Ching-Po Yang, Chi-Cheng Fu, Ken Sugino, Zhiyong Liu, Qingzhong Ren, Ling-Yu Liu, Xiaohao Yao, Luke P Lee, and Tzumin Lee. Transcriptomes of lineage-specific drosophila neuroblasts profiled via genetic targeting and robotic sorting. *Development*, pages dev–129163, 2015.
- [130] Haluk Lacin and James W Truman. Lineage mapping identifies molecular and architectural similarities between the larval and adult drosophila central nervous system. *Elife*, 5:e13399, 2016.
- [131] Chris Q Doe. Neural stem cells: balancing self-renewal with differentiation. *Development*, 135(9):1575–1587, 2008.
- [132] Kelvin Xi Zhang, Liming Tan, Matteo Pellegrini, S Lawrence Zipursky, and Jason M McEwen. Rapid changes in the transcriptome during the conversion of growth cones to synaptic terminals. *Cell reports*, 14(5):1258–1271, 2016.

- [133] Shovon I Ashraf, Anna L McLoon, Sarah M Sclarsic, and Sam Kunes. Synaptic protein synthesis associated with memory is regulated by the RISC pathway in drosophila. *Cell*, 124(1):191–205, 2006.
- [134] Robin J Stevens, Yulia Akbergenova, Ramon A Jorquera, and J Troy Littleton. Abnormal synaptic vesicle biogenesis in drosophila synaptogyrin mutants. *The Journal of Neuroscience*, 32(50):18054–18067, 2012.
- [135] Fen Huang, Jennifer K Chotiner, and Oswald Steward. The mRNA for elongation factor 1 α is localized in dendrites and translated in response to treatments that induce long-term depression. *The Journal of neuroscience*, 25(31):7199–7209, 2005.
- [136] Melissa M Rolls, Daisuke Satoh, Peter J Clyne, Astra L Henner, Tadashi Uemura, and Chris Q Doe. Polarity and intracellular compartmentalization of drosophila neurons. *Neural development*, 2(1):7, 2007.
- [137] Kaiyun Chen and David E Featherstone. Pre and postsynaptic roles for drosophila CASK. *Molecular and Cellular Neuroscience*, 48(2):171–182, 2011.
- [138] Kasper Daniel Hansen, Liana F Lareau, Marco Blanchette, Richard E Green, Qi Meng, Jan Rehwinkel, Fabian L Gallusser, Elisa Izaurralde, Donald C Rio, Sandrine Dudoit, et al. Genome-wide identification of alternative splice forms down-regulated by nonsense-mediated mRNA decay in drosophila. *PLoS Genet*, 5(6):e1000525, 2009.
- [139] Pedro Miura, Sol Shenker, Celia Andreu-Agullo, Jakub O Westholm, and Eric C Lai. Widespread and extensive lengthening of 3'UTRs in the mammalian brain. *Genome research*, 23(5):812–825, 2013.
- [140] J Matthew Taliaferro, Marina Vidaki, Ruan Oliveira, Sara Olson, Lijun Zhan, Tanvi Saxena, Eric T Wang, Brenton R Graveley, Frank B Gertler, Maurice S Swanson, et al. Distal alternative last exons localize mRNAs to neural projections. *Molecular cell*, 61(6):821–833, 2016.

- [141] André P Gerber, Stefan Luschnig, Mark A Krasnow, Patrick O Brown, and Daniel Herschlag. Genome-wide identification of mRNAs associated with the translational regulator Pumilio in drosophila melanogaster. *Proceedings of the National Academy of Sciences of the United States of America*, 103(12):4487–4492, 2006.
- [142] Fatima Cairrao, Anason S Halees, Khalid SA Khabar, Dominique Morello, and Nathalie Vanzo. AU-rich elements regulate drosophila gene expression. *Molecular and cellular biology*, 29(10):2636–2643, 2009.
- [143] Kailiang Sun, Jakub Orzechowski Westholm, Kazuya Tsurudome, Joshua W Hagen, Yubing Lu, Minoree Kohwi, Doron Betel, Fen-Biao Gao, A Pejmun Haghighi, Chris Q Doe, et al. Neurophysiological defects and neuronal gene deregulation in drosophila mir-124 mutants. *PLoS Genet*, 8(2):e1002515, 2012.
- [144] Justin P Kumar, Tazeen Jamal, Alex Doetsch, F Rudolf Turner, and Joseph B Duffy. CREB binding protein functions during successive stages of eye development in drosophila. *Genetics*, 168(2):877–893, 2004.
- [145] Sandhya P Koushika, Matthias Soller, and Kalpana White. The neuron-enriched splicing pattern of drosophila erect wing is dependent on the presence of ELAV protein. *Molecular and Cellular Biology*, 20(5):1836–1845, 2000.
- [146] Emanuela Zaharieva, Irmgard U Haussmann, Ulrike Bräuer, and Matthias Soller. Concentration and localization of coexpressed ELAV/Hu proteins control specificity of mRNA processing. *Molecular and cellular biology*, 35(18):3104–3115, 2015.
- [147] Kristin LM Boylan, Sarah Mische, Mingang Li, Guillermo Marqués, Xavier Morin, William Chia, and Thomas S Hays. Motility screen identifies drosophila IGF-II mrna-binding protein—zipcode-binding protein acting in oogenesis and synaptogenesis. *PLoS Genet*, 4(2):e36, 2008.

- [148] Makoto Nakamura, Hideyuki Okano, Julie A Blendy, and Craig Montell. Musashi, a neural RNA-binding protein required for drosophila adult external sensory organ development. *Neuron*, 13(1):67–81, 1994.
- [149] Bing Ye, Claudia Petritsch, Ira E Clark, Elizabeth R Gavis, Lily Yeh Jan, and Yuh Nung Jan. Nanos and Pumilio are essential for dendrite morphogenesis in drosophila peripheral neurons. *Current biology*, 14(4):314–321, 2004.
- [150] Josh Dubnau, Ann-Shyn Chiang, Lori Grady, Jody Barditch, Scott Gossweiler, John McNeil, Patrick Smith, Francois Buldoc, Rod Scott, Uli Certa, et al. The staufen/pumilio pathway is involved in drosophila long-term memory. *Current Biology*, 13(4):286–296, 2003.
- [151] Jian-Quan Ni, Michele Markstein, Richard Binari, Barret Pfeiffer, Lu-Ping Liu, Christians Villalta, Matthew Booker, Lizabeth Perkins, and Norbert Perrimon. Vector and parameters for targeted transgenic RNA interference in drosophila melanogaster. *Nature methods*, 5(1):49–51, 2008.
- [152] Matthias Landgraf, Sudipto Roy, Andreas Prokop, K VijayRaghavan, and Michael Bate. even-skipped determines the dorsal growth of motor axons in drosophila. *Neuron*, 22(1):43–52, 1999.
- [153] Francisco J Tejedor, Amr Bokhari, Oscar Rogero, Michael Gorczyca, Jiangwen Zhang, Eunjoon Kim, Morgan Sheng, and Vivian Budnik. Essential role for dlg in synaptic clustering of shaker K⁺ channels in vivo. *The Journal of neuroscience*, 17(1):152–159, 1997.
- [154] Timothy Lahey, Michael Gorczyca, Xi-Xi Jia, and Vivian Budnik. The drosophila tumor suppressor gene dlg is required for normal synaptic bouton structure. *Neuron*, 13(4):823–835, 1994.
- [155] Gengxin Chen, Wanhe Li, Qing-Shuo Zhang, Michael Regulski, Nishi Sinha, Jody Barditch, Tim Tully, Adrian R Krainer, Michael Q Zhang, and Josh Dubnau. Identification of synaptic targets of drosophila pumilio. *PLoS Comput Biol*, 4(2):e1000026, 2008.

- [156] Francesca Zalfa, Boris Eleuteri, Kirsten S Dickson, Valentina Mercaldo, Silvia De Rubeis, Alessandra di Penta, Elisabetta Tabolacci, Pietro Chiurazzi, Giovanni Neri, Seth GN Grant, et al. A new function for the fragile X mental retardation protein in regulation of PSD-95 mRNA stability. *Nature neuroscience*, 10(5):578–587, 2007.
- [157] Michael Hannus, Fabian Feiguin, Carl-Philipp Heisenberg, and Suzanne Eaton. Planar cell polarization requires widerborst, a B' regulatory subunit of protein phosphatase 2A. *Development*, 129(14):3493–3503, 2002.
- [158] Kwang-Wook Choi and Seymour Benzer. Rotation of photoreceptor clusters in the developing drosophila eye requires the nemo gene. *Cell*, 78(1):125–136, 1994.
- [159] Natalia Vereshchagina, Marie-Christine Ramel, Emmanuelle Bitoun, and Clive Wilson. The protein phosphatase PP2A-B' subunit widerborst is a negative regulator of cytoplasmic activated Akt and lipid metabolism in drosophila. *Journal of cell science*, 121(20):3383–3392, 2008.
- [160] Bernhard Fuss, Thomas Becker, Ingo Zinke, and Michael Hoch. The cytohesin steppke is essential for insulin signalling in drosophila. *Nature*, 444(7121):945–948, 2006.
- [161] Chunghun Lim, Jongbin Lee, Changtaek Choi, Valerie L Kilman, Juwon Kim, Sung Mi Park, Sung Key Jang, Ravi Allada, and Joonho Choe. The novel gene twenty-four defines a critical translational step in the drosophila clock. *Nature*, 470(7334):399–403, 2011.
- [162] Ling Li and Harald Vaessin. Pan-neural prospero terminates cell proliferation during drosophila neurogenesis. *Genes & development*, 14(2):147–151, 2000.
- [163] Virginia Goss Tusher, Robert Tibshirani, and Gilbert Chu. Significance analysis of microarrays applied to the ionizing radiation response. *Proceedings of the National Academy of Sciences*, 98(9):5116–5121, 2001.

- [164] Da Wei Huang, Brad T Sherman, and Richard A Lempicki. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nature protocols*, 4(1):44–57, 2009.
- [165] Da Wei Huang, Brad T Sherman, and Richard A Lempicki. Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. *Nucleic acids research*, 37(1):1–13, 2009.
- [166] Paul M Sharp, Elizabeth Cowe, Desmond G Higgins, Denis C Shields, Kenneth H Wolfe, and Frank Wright. Codon usage patterns in *escherichia coli*, *bacillus subtilis*, *saccharomyces cerevisiae*, *schizosaccharomyces pombe*, *drosophila melanogaster* and *homo sapiens*; a review of the considerable within-species diversity. *Nucleic Acids Research*, 16(17):8207–8211, 1988.
- [167] Joshua B Plotkin and Grzegorz Kudla. Synonymous but not the same: the causes and consequences of codon bias. *Nature Reviews Genetics*, 12(1):32–42, 2011.
- [168] Hiroshi Akashi. Synonymous codon usage in *drosophila melanogaster*: natural selection and translational accuracy. *Genetics*, 136(3):927–935, 1994.
- [169] Chien-Hung Yu, Yunkun Dang, Zhipeng Zhou, Cheng Wu, Fangzhou Zhao, Matthew S Sachs, and Yi Liu. Codon usage influences the local rate of translation elongation to regulate co-translational protein folding. *Molecular cell*, 59(5):744–754, 2015.
- [170] Kristin S Koutmou, Aditya Radhakrishnan, and Rachel Green. Synthesis at the speed of codons. *Trends in biochemical sciences*, 40(12):717–718, 2015.
- [171] Joel D Richter and Jeff Collier. Pausing on polyribosomes: Make way for elongation in translational control. *Cell*, 163(2):292–300, 2015.
- [172] Florian Buhr, Sujata Jha, Michael Thommen, Joerg Mittelstaet, Felicitas Kutz, Harald Schwalbe, Marina V Rodnina, and Anton A Komar. Synonymous codons

- direct cotranslational folding toward different protein conformations. *Molecular cell*, 61(3):341–351, 2016.
- [173] Kimberly A Dittmar, Michael A Sørensen, Johan Elf, Måns Ehrenberg, and Tao Pan. Selective charging of tRNA isoacceptors induced by amino-acid starvation. *EMBO reports*, 6(2):151–157, 2005.
- [174] Toshimichi Ikemura. Codon usage and tRNA content in unicellular and multicellular organisms. *Molecular biology and evolution*, 2(1):13–34, 1985.
- [175] Shigehiko Kanaya, Yuko Yamada, Makoto Kinouchi, Yoshihiro Kudo, and Toshimichi Ikemura. Codon usage and tRNA genes in eukaryotes: correlation of codon usage diversity with translation efficiency and with cg-dinucleotide usage as assessed by multivariate analysis. *Journal of molecular evolution*, 53(4-5):290–298, 2001.
- [176] Shigehiko Kanaya, Yuko Yamada, Yoshihiro Kudo, and Toshimichi Ikemura. Studies of codon usage and tRNA genes of 18 unicellular organisms and quantification of bacillus subtilis trnas: gene expression level and species-specific diversity of codon usage based on multivariate analysis. *Gene*, 238(1):143–155, 1999.
- [177] Fumiaki Yamao, Yoshiki Andachi, Akira Muto, Toshimichi Ikemura, and Syozo Osawa. Levels of tRNAs in bacterial cells as affected by amino acid usage in proteins. *Nucleic acids research*, 19(22):6119–6122, 1991.
- [178] Jeffrey R Powell and Etsuko N Moriyama. Evolution of codon usage bias in drosophila. *Proceedings of the National Academy of Sciences*, 94(15):7784–7790, 1997.
- [179] Etsuko N Moriyama and Jeffrey R Powell. Codon usage bias and tRNA abundance in drosophila. *Journal of molecular evolution*, 45(5):514–523, 1997.

- [180] Laurent Duret and Dominique Mouchiroud. Expression pattern and, surprisingly, gene length shape codon usage in caenorhabditis, drosophila, and arabidopsis. *Proceedings of the National Academy of Sciences*, 96(8):4482–4487, 1999.
- [181] Saverio Vicario, Etsuko N Moriyama, and Jeffrey R Powell. Codon usage in twelve species of drosophila. *BMC evolutionary biology*, 7(1):1, 2007.
- [182] Renana Sabi and Tamir Tuller. Modelling the efficiency of codon–tRNA interactions based on codon usage bias. *DNA Research*, 21(5):511–526, 2014.
- [183] Yoshinori Matsuo. Evolutionary change of codon usage for the histone gene family in drosophila melanogaster and drosophila hydei. *Molecular phylogenetics and evolution*, 15(2):283–291, 2000.
- [184] Anna S Akhmanova, Petra CT Bindels, Jie Xu, Koos Miedema, Hannie Kremer, Wolfgang Hennig, Jie Xu, and Wolfgang Hennig. Structure and expression of histone H3. 3 genes in drosophila melanogaster and drosophila hydei. *Genome*, 38(3):586–600, 1995.
- [185] RP Lifton, ML Goldberg, RW Karp, and DS Hogness. The organization of the histone genes in drosophila melanogaster: functional and evolutionary implications. In *Cold Spring Harbor symposia on quantitative biology*, volume 42, pages 1047–1051. Cold Spring Harbor Laboratory Press, 1978.
- [186] C Stauber, B Lüscher, R Eckner, E Lötscher, and D Schümperli. A signal regulating mouse histone H4 mrna levels in a mammalian cell cycle mutant and sequences controlling RNA 3’processing are both contained within the same 80-bp fragment. *The EMBO journal*, 5(12):3297, 1986.
- [187] Mehmet Neset Özel, Marion Langen, Bassem A Hassan, and P Robin Hiesinger. Filopodial dynamics and growth cone stabilization in drosophila visual circuit development. *eLife*, 4:e10721, 2015.

- [188] Youichi Iwai, Tadao Usui, Shinji Hirano, Ruth Steward, Masatoshi Takeichi, and Tadashi Uemura. Axon patterning requires D N-cadherin, a novel neuronal adhesion receptor, in the drosophila embryonic CNS. *Neuron*, 19(1):77–89, 1997.
- [189] Mark Petronczki and Juergen A Knoblich. DmPAR-6 directs epithelial polarity and asymmetric cell division of neuroblasts in drosophila. *Nature Cell Biology*, 3(1):43–49, 2001.
- [190] Yutaka Hata, Stefan Butz, and Thomas C Sudhof. CASK: a novel dlg/PSD95 homolog with an N-terminal calmodulin-dependent protein kinase domain identified by interaction with neuroligins. *The Journal of neuroscience*, 16(8):2488–2494, 1996.
- [191] John Peden. Correspondence analysis of codon usage, 2005.
- [192] Patricia P Chan and Todd M Lowe. GtRNAdb: a database of transfer RNA genes detected in genomic sequence. *Nucleic acids research*, 37(suppl 1):D93–D97, 2009.
- [193] Zhenguo Zhang. Cua: a flexible and comprehensive codon usage analyzer. *bioRxiv*, page 022814, 2015.
- [194] Juliane C Dohm, Claudio Lottaz, Tatiana Borodina, and Heinz Himmelbauer. Substantial biases in ultra-short read data sets from high-throughput DNA sequencing. *Nucleic acids research*, 36(16):e105–e105, 2008.
- [195] Kasper D Hansen, Steven E Brenner, and Sandrine Dudoit. Biases in illumina transcriptome sequencing caused by random hexamer priming. *Nucleic acids research*, 38(12):e131–e131, 2010.
- [196] Edward Sandler, Graham D Johnson, and Stephen A Krawetz. Local and global factors affecting RNA sequencing analysis. *Analytical biochemistry*, 419(2):317–322, 2011.

- [197] Bingjian Lü, Jiyang Yu, Jing Xu, Jian Chen, and Maode Lai. A novel approach to detect differentially expressed genes from count-based digital databases by normalizing with housekeeping genes. *Genomics*, 94(3):211–216, 2009.
- [198] Paul L Auer, Sanvesh Srivastava, and RW Doerge. Differential expression—the next generation and beyond. *Briefings in functional genomics*, page elr041, 2011.
- [199] Zhide Fang and Xiangqin Cui. Design and validation issues in RNA-seq experiments. *Briefings in bioinformatics*, page bbr004, 2011.
- [200] Ying Wang, Noushin Ghaffari, Charles D Johnson, Ulisses M Braga-Neto, Hui Wang, Rui Chen, and Huaijun Zhou. Evaluation of the coverage and depth of transcriptome by RNA-Seq in chickens. *Bmc Bioinformatics*, 12(Suppl 10):S5, 2011.
- [201] Adam Roberts, Cole Trapnell, Julie Donaghey, John L Rinn, Lior Pachter, et al. Improving RNA-Seq expression estimates by correcting for fragment bias. *Genome Biol*, 12(3):R22, 2011.
- [202] Vanessa M Kvam, Peng Liu, and Yaqing Si. A comparison of statistical methods for detecting differentially expressed genes from RNA-seq data. *American journal of botany*, 99(2):248–256, 2012.
- [203] Mark D Robinson, Davis J McCarthy, and Gordon K Smyth. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics*, 26(1):139–140, 2010.
- [204] Xiaobei Zhou, Helen Lindsay, and Mark D Robinson. Robustly detecting differential expression in RNA sequencing data using observation weights. *Nucleic acids research*, 42(11):e91–e91, 2014.
- [205] Davis J McCarthy, Yunshun Chen, and Gordon K Smyth. Differential expression analysis of multifactor RNA-Seq experiments with respect to biological variation. *Nucleic acids research*, page gks042, 2012.

- [206] Likun Wang, Zhixing Feng, Xi Wang, Xiaowo Wang, and Xuegong Zhang. DEGseq: an R package for identifying differentially expressed genes from RNA-seq data. *Bioinformatics*, 26(1):136–138, 2010.
- [207] Simon Anders and Wolfgang Huber. Differential expression analysis for sequence count data. *Genome biol*, 11(10):R106, 2010.
- [208] Paul L Auer and RW Doerge. Statistical design and analysis of RNA sequencing data. *Genetics*, 185(2):405–416, 2010.
- [209] Simon Anders and Wolfgang Huber. Differential expression of RNA-Seq data at the gene level—the DESeq package. *Heidelberg, Germany: European Molecular Biology Laboratory (EMBL)*, 2012.
- [210] Charity W Law, Yunshun Chen, Wei Shi, and Gordon K Smyth. Voom: precision weights unlock linear model analysis tools for RNA-seq read counts. *Genome Biol*, 15(2):R29, 2014.
- [211] Gordon K Smyth. Limma: linear models for microarray data. In *Bioinformatics and computational biology solutions using R and Bioconductor*, pages 397–420. Springer, 2005.
- [212] Yoav Benjamini and Yosef Hochberg. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the royal statistical society. Series B (Methodological)*, pages 289–300, 1995.