

ENDOGENOUS AND ENGINEERED EXPRESSION OF tRNA INTRON-DERIVED CIRCULAR
RNAS

John Joseph Noto

A dissertation submitted to the faculty of the University of North Carolina at Chapel Hill in
partial fulfillment of the requirements for the degree of Doctor of Philosophy in the
Curriculum in Genetics and Molecular Biology.

Chapel Hill
2017

Approved by:

A. Gregory Matera

Stephen T. Crews

William F. Marzluff

Daniel J. McKay

© 2017
John Joseph Noto
ALL RIGHTS RESERVED

ABSTRACT

John Joseph Noto: Endogenous and Engineered Expression of tRNA Intron-derived Circular RNAs
(Under the direction of A. Gregory Matera)

Sophisticated methods of ribonucleic acid (RNA) detection including RNA-sequencing (RNA-seq) have been used to rapidly expand the known transcriptomic landscape, and formerly underappreciated RNA species including circular RNAs (circRNAs) have emerged as important subjects of research. Our lab recently developed a bioinformatics tool called Vicinal used to map RNA-seq reads that cannot be mapped end-to-end to the genome. Using a modified version of this software, we discovered abundant reads mapping to the 5'- and 3'-ends of tRNA introns in *Drosophila* that could best be explained by circularization of introns. We have termed these tRNA intronic circular RNAs (tricRNAs). We have used numerous assays to show that the predominant form of a particular excised intron, tric31905, is circular in animals. This tricRNA is unusual in length, displays sequence conservation among *Drosophilids*, and could be processed into small RNAs. We have shown that tric31905 has a dynamic expression profile during fly development, and is more abundant in certain tissues. Further research is necessary to conclude whether the cell has evolved functions for tricRNAs. Additionally, we have established that the cell uses a conserved mechanism to generate tricRNAs, and have harnessed the endogenous machinery to robustly generate stable, “designer” circRNAs in cell cultures and in living animals. We utilize a powerful imaging technique, using the

aptamer sequence Broccoli as a reporter to readily track expression of engineered tricRNAs. This system could be a powerful tool for research in tRNA processing, and could be a platform for the delivery of effector RNAs, therapeutic or otherwise.

To my family and friends. Thank you for your support past, present and in the future.

ACKNOWLEDGEMENTS

Thank you to my friends, family, and lab members. In particular, Zhipeng Lu and Casey Schmidt have offered a great deal of help to me in completion of this dissertation. Their contributions have been essential for me to find my footing as a researcher. The guidance of my committee and my mentor, Greg Matera, has also had a profound influence on my maturation.

TABLE OF CONTENTS

LIST OF FIGURES	xi
LIST OF ABBREVIATIONS	xiii
CHAPTER 1: INTRODUCTION	14
RNAs: masters of the ‘warm little pond’, regulators of life.....	14
Advances in RNA-seq shed light on circular RNAs.....	15
Circular RNA function.....	15
Circular RNA biogenesis	18
Exploring tricRNAs.....	19
Figures.....	21
CHAPTER 2: A FRUIT FLY tRNA INTRON IS PREDOMINANTLY PROCESSED INTO A CIRCULAR RNA WITH A TISSUE-SPECIFIC AND DYNAMIC EXPRESSION PROFILE THROUGHOUT DEVELOPMENT	31
Introduction	31
Discovery and mechanism of tricRNA biogenesis in Archaea.....	31
Splicing and circularization of tRNA introns in eukarya	33
<i>Drosophilid</i> tric31905	35
Research goals	36
Results.....	36
Divergent PCR of tric31905	36

RNase R	37
RNase H.....	37
Aberrent electrophoretic mobility	38
RtcB knockdown experiments	38
Developmental expression pattern.....	39
Discussion	39
Methods and Materials	41
RT-PCR	41
Northern blotting.....	42
RNase H Digestion and Northern blotting.....	42
RNase R Digestion and Northern blotting.....	43
Fly genetics and preparation of RNA samples	43
Developmental Northern blot	44
Figures.....	45
CHAPTER 3: ENGINEERED EXPRESSION AND TISSUE- AND AGE-SPECIFIC ACCUMULATION OF tricRNAs IN ANIMALS	63
Introduction	63
Methods of ectopic circRNA expression and detection.....	64
Summary.....	65
Results	66
Circular RNA expression systems comparison.....	66
Transgenic fly creation.....	67
Age-dependent accumulation	68

Discussion	68
Methods and Materials	69
Expression Comparison: Plasmid construction.....	69
Expression Comparison: DFHBI and Ethidium Bromide staining.....	70
RT-PCR	70
Expression Comparison: Northern blotting.....	70
tricBroc transgene insertion: phiC31 integration	71
Fly dissection and Northern blotting for age and tissue comparison.....	71
Cell culture and RNA isolation	71
Figures.....	72
CHAPTER 4: CONCLUSIONS AND FUTURE DIRECTIONS	88
tricRNA expression is conserved and results in a diverse expression pattern in tissues and throughout animal development.....	88
tRNA splicing as an engineered platform for robust expression of stable circRNAs	88
Can we engineer an inducible tricRNA expression system?	89
Do endogenous tricRNAs have functions?.....	90
Can we engineer expression of functional tricRNAs?	91
Methods.....	92
Cloning pVALIUM20 plasmid for TRiP line generation, injection/integration selection to make fly stock, and crossing to expression driver	92
Northern blotting.....	92
Figures.....	93
APPENDIX 1: miR-21 SPONGE tricRNA EXPRESSION	101

Goals and results	101
Methods	102
mIR-21 sponge.....	102
Cloning pSICHECK-2 luciferase reporter for mIR-21 activity	103
Luciferase assay	103
Figures	104
APPENDIX 2: SMN SPLICING MODULATOR tricRNA EXPRESSION	110
Goals and results	110
Methods	111
Cloning SMN splice modifier tricRNA expression vector	111
RT-PCR	111
Cell culture and RNA preparation	111
Figures	113
References	119

LIST OF FIGURES

Figure 1.1: The mammalian genome controls expression of a multitude of RNA classes including many non-coding subtypes	21
Figure 1.2: A simplified evolutionary history of life on Earth	23
Figure 1.3: circRNA functions	25
Figure 1.4: Translation of back-spliced circRNAs.....	27
Figure 1.5: Mechanism of back-splicing with protein effectors.....	29
Figure 2.1: Production of circular RNAs via tRNA splicing.	45
Figure 2.2: RNA-seq reads spanning the circularized junction of tRNA introns	47
Figure 2.3: Processing of tric31905 into small RNAs	49
Figure 2.4: Divergent PCR validation of tric31905 circularity.....	51
Figure 2.5: RNase R validation of tric31905 circularity.....	53
Figure 2.6: Products of RNase H digestion of tric31905	55
Figure 2.7: Aberrant electrophoretic mobility of tric31905 in gels with different agarose concentrations	57
Figure 2.8: Knockdown of tRNA ligase RtcB results in decreased tricRNA expression in flies	59
Figure 2.9: Developmental pattern of tric31905 expression in fruit flies	61
Figure 3.1: tricRNA expression can be detected by RT-PCR.....	72
Figure 3.2: Comparison of <i>in vivo</i> methods of circular RNA expression using the Broccoli RNA aptamer as a reporter	74
Figure 3.3: phiC31 integration of transgenes.....	76
Figure 3.4: Schematic for tricBroccoli transgene insertion using phiC31 in flies	78
Figure 3.5: Plasmid map of pAttB-tricRNA phi-C31 vector.....	80

Figure 3.6: Expression of tricBroccoli in flies can be detected by in-gel DFHBI-1T staining.....	82
Figure 3.7: Gene dosage alters abundance of expression of tricBroccoli in fruit flies	84
Figure 3.8: Accumulation of tricBroccoli in aged fly tissues	86
Figure 4.1: Inducible tricRNA expression system.....	93
Figure 4.2: The effects of B-Box point mutations on transcription of a <i>Drosophila</i> tRNA^{Arg} gene in (A) <i>Drosophila</i> and (B) HeLa cytoplasmic extracts.....	95
Figure 4.3: tric31905 shRNA knockdown in flies is ineffective	97
Figure 4.4: Proposed CRISPR-Cas9 tRNA-mutant genotypes for study of tric31905 function.....	99
Figure Appendix 1.1: Schematic of miR-21-responsive luciferase reporter assay for miR-21 sponge effectiveness	104
Figure Appendix 1.2: miR-21 sponge is expressed after transient transfection in HeLa cells	106
Figure Appendix 1.3: Luciferase assay for miR-21 sponge effectiveness	108
Figure Appendix 2.1: Use of antisense oligonucleotides to alter SMN splicing in SMA model mice	113
Figure Appendix 2.2: SMN splicing modifier tricRNA design.....	115
Figure Appendix 2.3: RT-PCR to detect SMN splice modifier tricRNAs and exon 7 inclusion in spliced SMN transcripts in cells transfected with SMN splice modifier tricRNA construct.....	117

LIST OF ABBREVIATIONS

ADAR	Adenosine deaminase acting on RNA
cDNA	Complementary DNA
circRNA	Circular RNA
DNA	Deoxyribonucleic acid
hnRNP	Heterogeneous nuclear ribonucleoprotein
IRES	Internal ribosome entry site
mRNA	Messenger RNA
miRNA	MicroRNA
PCR	Polymerase chain reaction
RNA	Ribonucleic acid
RNase	Ribonuclease
RNA-seq	RNA-sequencing
RT-PCR	Reverse transcription-PCR
snRNA	Small nuclear RNA
sRNA	Small RNA
SR protein	Serine arginine protein
tricRNA	tRNA intronic circular RNA
tRNA	Transfer RNA
rRNA	Ribosomal RNA
UTR	Untranslated region

CHAPTER 1: INTRODUCTION¹

RNAs: masters of the ‘warm little pond’, regulators of life

The origin of life on earth is one of the most fundamental questions in modern science. Many researchers now share the belief that life’s emergence depended on the spontaneous formation of ribonucleic acid (RNA), which according to the ‘RNA world’ hypothesis (Gilbert 1986) was the primordial biomolecule making life possible. Several factors led to this conclusion, including the presence of RNA in every living cell, its ability to store information and catalyze chemical reactions, and estimations of the prebiotic conditions and chemical composition of the early earth that would allow for the formation of RNA polymers. Though this subject is still an active debate, it is clear that RNA regulates the biological world we currently inhabit to an immense degree. Not only does RNA serve as a functional messenger between deoxyribonucleic acid (DNA) and proteins; myriad types of RNA are transcribed from the genome with various functions (**Figure 1.1**). It’s now appreciated that these many categorizations of RNAs control the flow of information essential for life, described in the ‘central dogma of biology’. This control is exerted in an increasingly complex manner along the evolutionary timescale (**Figure 1.2**).

¹ Portions of this chapter previously appeared in *RNA Biology*. The original citation is as follows: Noto, J. J., Schmidt, C. A. and Matera, A. G. (2017). Engineering and expressing circular RNAs via tRNA splicing. *RNA Biology*, 14(8), 978-984.

Advances in RNA-seq shed light on circular RNAs

Modern researchers have benefitted greatly from advances in high-throughput sequencing technologies that have allowed in-depth characterization of cellular transcriptomes. This has helped to reveal a multitude of actively transcribed and functionally relevant RNA species that had been previously underappreciated, particularly with regard to non-coding RNAs. Advances in many disciplines have contributed to this burgeoning science, including powerful hardware and software development, along with new and improved molecular biological methods. RNA-seq has helped to determine that > 85% of the human genome is in fact transcribed into RNA, whereas only 3% is known to code for protein (Hangauer, Vaughn and McManus, 2013). Though the exact percentage of the genome that has function is a matter of debate (Graur, 2017), the discovery and characterization of new functions for RNA species continues at an impressive rate. Circular RNAs (circRNAs) represent one such type of RNA that was mostly overlooked until recently. Growth in the study of circRNA processing and function has been propelled largely by these advances in high-throughput sequencing technology and refined bioinformatic pipelines (Jeck *et al.*, 2013; Jeck and Sharpless, 2014). In addition, the use of advanced molecular biological protocols, such as treatment using enzymes such as RNase R, have aided in the enrichment and detection of RNAs from total cellular RNA samples.

Circular RNA function

CircRNAs in eukaryotes were once thought of as rare molecules resulting from aberrant messenger RNA (mRNA) splicing (Coquercelle *et al.*, 1993). It has since

become clear that circRNA biogenesis in many eukaryotes proceeds through several diverse and evolutionarily conserved biochemical pathways (Lasda and Parker, 2014; Petkovic and Müller, 2015; Ebbesen *et al.*, 2017; Fischer and Leung, 2017), resulting in an abundant class of RNA molecules with a host of proposed functions (Qu *et al.*, 2016; Salzman, 2016). CircRNAs have been shown to be remarkably stable in living cells (Kjems and Garrett, 1988; Dalgaard and Garrett, 1992; Lu *et al.*, 2015), potentially due to their unique structure and lack of free termini on which the exonucleolytic RNA turnover machinery relies. Along with being generally abundant (Jeck *et al.*, 2013), circRNAs can be differentially expressed in a tissue-dependent manner and can have expression profiles that differ from their parental mRNA (Chen, 2016), suggesting regulated expression and diversity of function. CircRNAs are enriched in mammalian and *Drosophila* brains (Hansen *et al.*, 2013; Memczak *et al.*, 2013; Salzman *et al.*, 2013; Rybak-Wolf *et al.*, 2014; Westholm *et al.*, 2014), are positively correlated with aging (Westholm *et al.*, 2014), and can be associated with disease risk (Burd *et al.*, 2010). The most well known example of circular RNA function is their ability to act as molecular “sponges,” competing for binding of molecules that affect gene regulation (**Figure 1.3**). For example, the human circular transcript CDR1as/ciRS-7 contains approximately 70 seed target sequences for microRNA-7 (miR-7) and associates with exosomal Argonaute proteins *in vivo* in a miR-7 dependent manner (Hansen *et al.* 2013, Memczak *et al.* 2013). CDR1as/ciRS-7 represents a striking example of the controversial “competing endogenous RNA” theory, which posits that transcripts can “compete” for binding of microRNAs (miRNAs) and other molecules to affect gene expression (Seitz, 2009;

Salmena *et al.*, 2011; Tay *et al.*, 2011). This transcript has overlapping expression patterns with miR-7 in mouse brains, suggesting a potential *in vivo* regulatory interaction (Hansen *et al.*, 2013; Memczak *et al.*, 2013). Interestingly, either inhibition of miR-7 or ectopic expression of CDR1as causes midbrain development defects in zebrafish (Memczak *et al.*, 2013), indicating that disruption of this type of regulation could have significant consequence.

In addition to miRNAs, circRNAs can also bind and sequester proteins (**Figure 1.3**). In human, mouse, and *Drosophila*, the RNA-binding protein Muscleblind (Mbl) has been shown to promote circularization of several transcripts, including the second exon of the Mbl pre-mRNA itself (Ashwal-Fluss *et al.*, 2014). In all three species, circMbl RNA contains conserved Mbl protein binding sites, and pulldown experiments in *Drosophila* S2 cells revealed an *in vivo* interaction between Mbl protein and circMbl (Ashwal-Fluss *et al.*, 2014), suggesting a possible feedback mechanism by which Mbl protein and circMbl interact to control splicing of the parent transcript. The relative stability of circRNAs could perhaps contribute to a unique cellular role in initiating and maintaining these types of regulatory interactions.

Circular RNAs can also serve as templates for protein translation (**Figure 1.4**). While it has been known for some time that eukaryotic ribosomes could translate protein from *in vitro* synthesized circRNAs (Chen and Sarnow, 1995), until very recently it was up for debate as to whether endogenous circRNAs could be translated. Initiation of translation in eukaryotes includes recruitment of RNA to an assembling ribosome through association with the 7-methylguanosine cap moiety on the transcript's 5'-end,

which is not present in a circRNA. However, some linear transcripts include elements such as an internal ribosome entry site (IRES) that can promote cap-independent translation by recruiting the small ribosomal subunit to an internal initiator codon. It was reported that inclusion of an IRES can enable translation from engineered circRNAs (Chen and Sarnow, 1995; Perriman and Ares, 1998; Wang and Wang, 2015). More recently, it was shown that circRNAs lacking an IRES can be translated into proteins. Some linear mRNAs can be translated in a cap-independent fashion relying on N⁶-methyladenosine (m⁶A) base-modification sites in their 5'- untranslated regions (UTRs) (Meyer *et al.*, 2015). CircRNAs harboring even a single m⁶A site can also be templates for translation (Yang *et al.*, 2017). Since m⁶A is the most common RNA base modification (Linder *et al.*, 2015), this could dramatically increase the number of known transcripts with protein-coding potential. Furthermore, endogenous eukaryotic circRNAs have recently been measured in polysome fractions, detected by ribosome footprinting, and circRNA protein products have been detected by mass-spectrometry, suggesting that circRNA translation could be widespread in eukaryotes (Legnini *et al.*, 2017; Pamudurti *et al.*, 2017; Yang *et al.*, 2017).

Circular RNA biogenesis

The most common pathway for RNA circularization occurs via a process termed “back-splicing.” Back-splicing proceeds via a canonical 2-step transesterification reaction, except that a downstream splice donor sequence becomes ligated to an upstream splice acceptor site, resulting in a non-canonical alternative splicing event (**Figure 1.5**).

Importantly, back-splicing has been shown to compete with forward splicing of certain pre-mRNA transcripts (Ashwal-Fluss *et al.*, 2014; Starke *et al.*, 2015). RNA elements that promote juxtaposition of splice sites close to one another in physical space are thought to promote back-splicing. For example, back-splicing of circular transcripts is associated with the presence of repetitive elements (e.g. Alu repeats in humans or different repetitive elements in other species) that flank the circularized region (Jeck *et al.*, 2013; Liang and Wilusz, 2014; Kramer *et al.*, 2015). Additionally, back-splicing can be facilitated by RNA-binding proteins such as the aforementioned Mbl, QKI (quaking), heterogeneous nuclear ribonucleoproteins (hnRNPs) and serine/arginine (SR) proteins (Ashwal-Fluss *et al.*, 2014; Conn *et al.*, 2015; Kramer *et al.*, 2015). Back-splicing can also be inhibited by the enzyme adenosine deaminase acting on RNA (ADAR). ADAR catalyzes adenosine to inosine base modification, resulting in loss of A-T base pairing at that residue. This loss of base-pairing can relax the secondary structure formed by long inverted repeats flanking a circularized exon in back-splicing, and thereby inhibit circularization (Ivanov *et al.*, 2015).

Exploring tricRNAs

Researchers have made strides in shining a light on transcriptomic “dark matter”. There has been an increased emphasis on the importance of non-coding portions of the transcriptome, as well as discoveries and functional analyses of RNAs that do not map contiguously to the genome and were perhaps overlooked previously. It is becoming clear that circRNAs do not represent a mere error in splicing, but are likely regulated and have

function within cells. The focus of this dissertation is to explore a novel method of RNA circularization, where excised introns from eukaryotic transfer RNAs (tRNAs) are ligated to form stable circRNAs. Here I detail my efforts to show conclusively that the phenomenon of tRNA intron circularization to form tRNA-intronic circular RNAs (tricRNAs) occurs in eukaryotic cells and in animals, and show that we can engineer a platform using this molecular pathway to drive the expression of “designer” circRNAs that could be of use as a research tool or potential therapeutic.

Figures

Figure 1.1: The mammalian genome controls expression of a multitude of RNA classes including many non-coding subtypes. This figure shows the transcriptional repertoire encoded in the mammalian genome, which controls expression of coding and non-coding transcripts. These include housekeeping RNAs such as rRNAs, tRNAs, snRNAs, snoRNAs, a number of coding and noncoding transcripts such as mRNAs and lncRNAs, as well as small regulatory RNAs including miRNAs, piRNAs, tiRNAs and spliRNAs, snoRNA-derived small RNAs, and tRNA-derived small RNAs. This figure was adapted (Morris and Mattick, 2014).

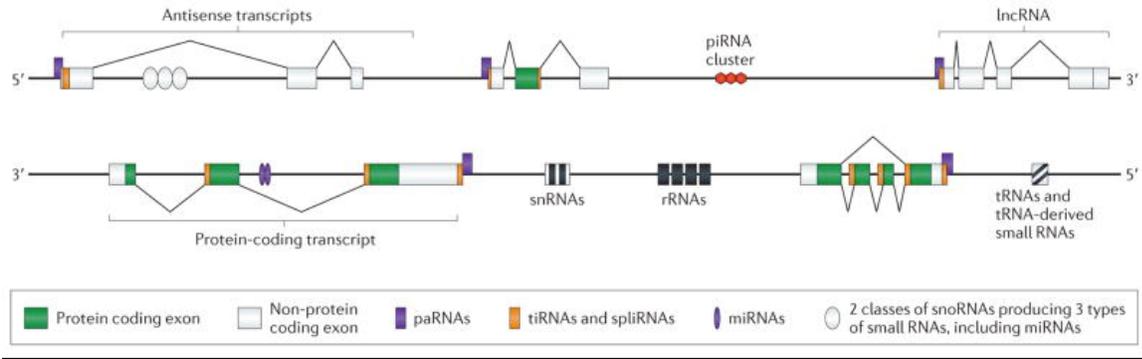


Figure 1.2: A simplified evolutionary history of life on Earth. This graph presents an overview of some important evolutionary leaps over time (some dates are still an open debate) contributing to an increase in diversity and complexity of life. The scale of complexity shown here is arbitrary. The inclusions of introns in genomes as well as the formation of RNA regulatory networks correlate with and could be causative for increases in biological complexity. This figure was adapted (Mattick, 2004).

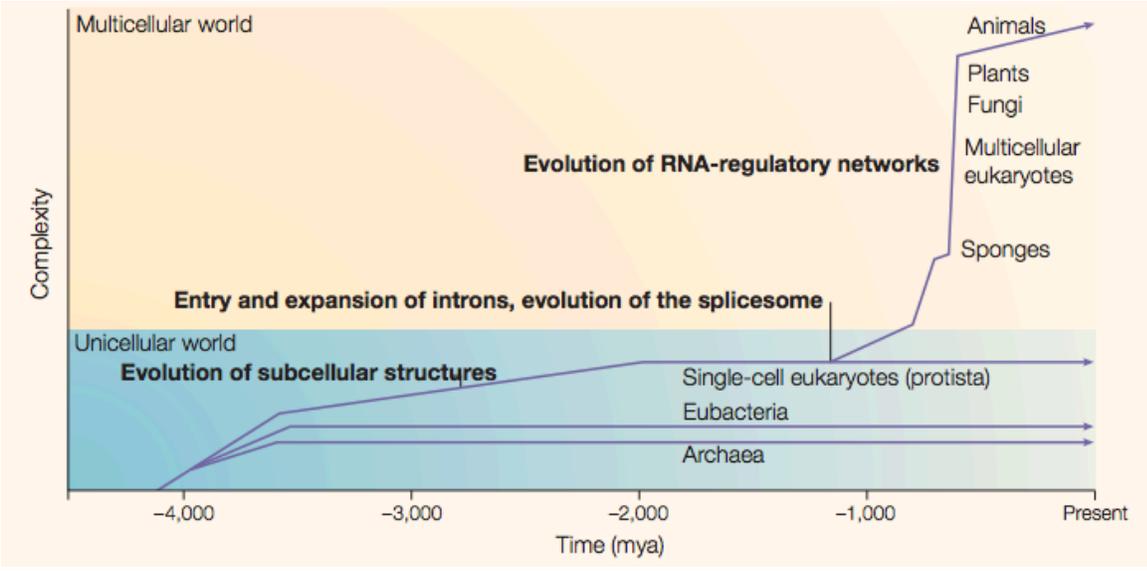


Figure 1.3: circRNA functions: circRNAs can function as molecular sponges containing binding sites for particular miRNAs or RBPs and compete for binding (dashed arrows) with their target RNAs. This figure was adapted (Barrett and Salzman, 2016).

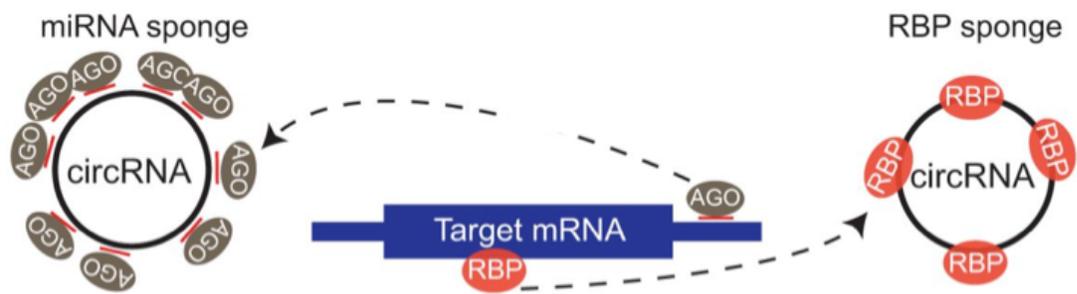
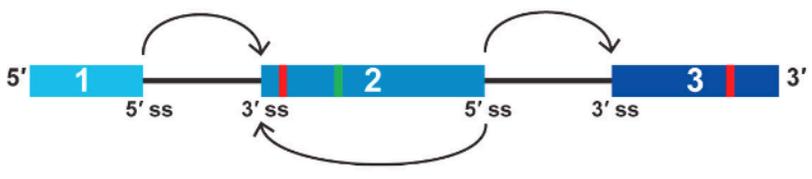


Figure 1.4: Translation of back-spliced circRNAs. A pre-mRNA can be spliced to generate a linear mRNA or a circular RNA. When the pre-mRNA splice sites (ss) are joined in the canonical order (top), a linear mRNA is generated that is subsequently translated to produce the annotated full-length protein. Alternatively, backsplicing (bottom) can generate a circular RNA whose ends are covalently joined at the backspliced junction (purple). Translation can initiate at the AUG and elongate past the backspliced junction until a stop codon is reached. This results in the generation of a truncated protein that has a unique C terminus that is not present in the protein generated from the linear mRNA. **This figure was adapted** (Tatomer and Wilusz, 2017).

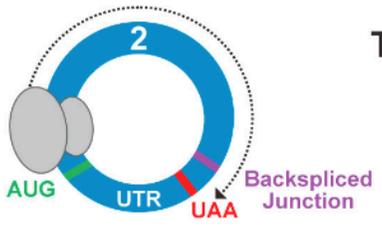
Linear mRNA



Canonical Splicing

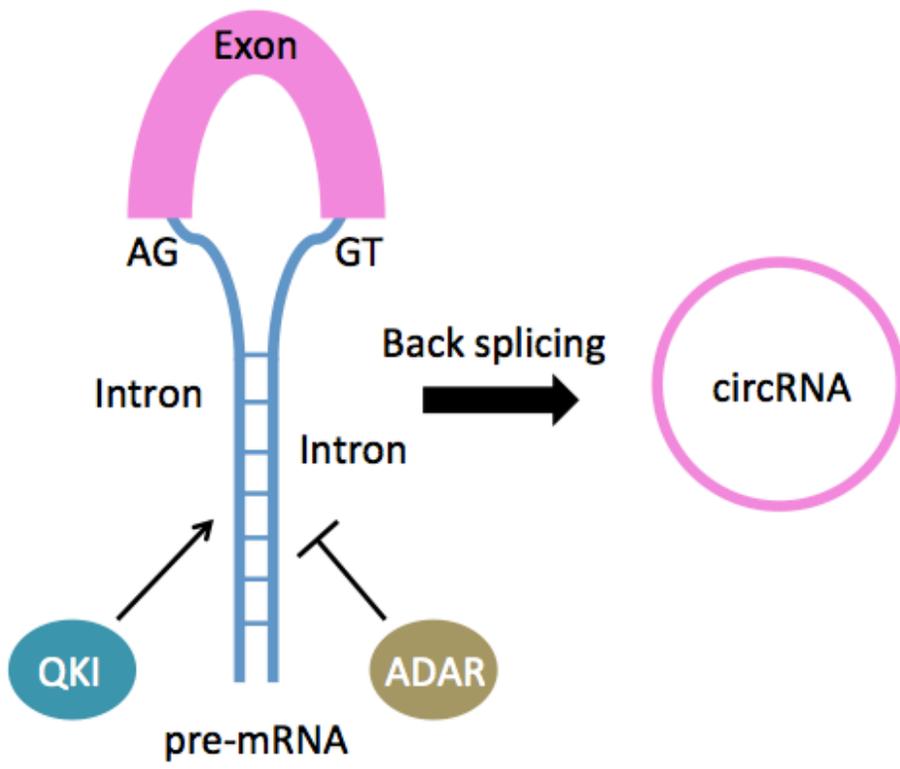


Backsplicing



Circular RNA

Figure 1.5: Mechanism of back-splicing with protein effectors. Intron base-pairing and/or RNA-binding proteins facilitate pairing of a downstream splice site and an upstream splice site, bringing them into close proximity. Negative regulation by proteins like ADAR that reduce base-pairing is also possible. Non-canonical “back-splicing” of these sites results in a circularized exon. This figure was adapted (Noto *et al.*, 2017).



CHAPTER 2: A FRUIT FLY tRNA INTRON IS PREDOMINANTLY PROCESSED INTO A CIRCULAR RNA WITH A TISSUE-SPECIFIC AND DYNAMIC EXPRESSION PROFILE THROUGHOUT DEVELOPMENT²

Introduction

Discovery and mechanism of tricRNA biogenesis in Archaea

In these sections, I detail a conserved mode of RNA circularization occurring in archaea and eukaryotes, in which tRNA introns are spliced to form circles. Despite the absence of the spliceosomal machinery in Bacteria and Archaea, intron-containing genes have been observed in all 3 domains of life (Johnson, 2002). Bacterial introns are of the self-splicing variety (Woodson, 1998). In Archaea, however, introns are fairly common among tRNA and ribosomal RNA (rRNA) genes, and their removal is catalyzed by a conserved splicing endonuclease complex (Tocchini-Valentini, Fruscoloni and Tocchini-Valentini, 2005; Calvin and Li, 2008). Interestingly, archaeal RNA splicing typically produces a circularized intron following its removal from the primary transcript (Salgia *et al.*, 2003). Splicing of pre-tRNAs is a 2-step process wherein the transcript is first cleaved at defined positions by the archaeal splicing endonuclease complex, which recognizes a bulge-helix-bulge (BHB) sequence motif (**Figure 2.1**) (Xue, Calvin and Li, 2006; Yoshihisa, 2014).

² Portions of this chapter were previously published in *RNA*. The original citation is as follows: Lu, Z., Filonov, G. S., Noto, J. J., Schmidt, C. A., Hatkevich, T. L., Wen, Y., Jaffrey, S. R. and Matera, A. G. (2015). Metazoan tRNA introns generate stable circular RNAs in vivo. *RNA*, 21(9), 1554–1565.

The exon halves are then ligated together to form a mature tRNA, while the intron termini are also joined to form a circle (**Figure 2.1**). Although the two ligation events appear to use the same processing factors, the reactions can occur independently (Salgia *et al.*, 2003).

Circular RNAs derived from multiple tRNA isodecoder transcripts have been detected (Chan and Lowe, 2009) though it is not known if stable circles are made from every tRNA intron. To distinguish these circRNAs from their back-spliced cousins in eukaryotes, we use the term tRNA intronic circular RNAs, or tricRNAs. In the archaeon *Haloferax volcanii*, a tricRNA was detected in vivo from tRNA^{Trp} by reverse-transcription PCR (RT-PCR) and Northern blot analysis (Salgia *et al.*, 2003). In this same study, introns derived from elongator tRNA^{Met} appear to undergo end-joining that would be characteristic of circularization, but tricRNA expression could not be confirmed by blotting. This lack of detectability could be due to differences in overall parental tRNA expression: in *Haloferax volcanii*, there is only one tRNA^{Trp} gene, whereas there are 3 elongator tRNA^{Met} genes, and only one of them contains an intron (Chan and Lowe, 2009). Thus, reduced expression of the parent tRNA^{Met} gene might lead to lower amounts of its tricRNA that are detectable only by RT-PCR. In certain archaeal species, tricRNAs are highly expressed and/or are quite stable. For example, a whole-genome sequencing approach in *Sulfolobus solfataricus* revealed that among circRNAs, only those generated from 16S and 23S rRNA were more abundant than tricRNAs (Danan *et al.*, 2012). Circles were only detected for 5 of the 19 intron-containing tRNAs, but this is most likely due to an ascertainment bias, as the remaining 14 introns are quite small and would thus require specific methods of library preparation to enrich for them.

Several groups have speculated on the functions of circularized introns in archaea. Circularized rRNA introns have been found to contain putative open reading frames, and it is possible that these encode peptides (Kjems and Garrett, 1988; Dalgaard and Garrett, 1992; Burggraf *et al.*, 1993). Additionally, a functional C/D box RNA is contained in the tRNA^{Trp} intron of *H. volcanii*, responsible for 2'-O-methylation of its parent tRNA, though both circular and linear excised introns are capable of guiding this modification (Singh *et al.*, 2004). Similarly, the tRNA^{Trp} intron of *S. solfataricus* contains regions of sequence complementarity to its parent tRNA, suggesting a role for the excised intron in tRNA modification (Danan *et al.*, 2012). Whether tricRNAs from other parental tRNAs and other species also function in RNA modification remains to be determined.

Splicing and circularization of tRNA introns in eukarya

Our first indication that metazoan tRNA introns were probably processed into circular molecules in eukaryotes was the discovery that in *Drosophila* RNA-seq experiments there were an abundance of reads generated whose 5' and 3' ends mapped to opposite ends of tRNA introns (**Figure 2.2**). These reads did not map to the genome end-to-end, and thus required mapping using modified software Vicinal (Lu and Matera, 2014; Lu *et al.*, 2015). The density of reads spanning the putative circular junction was similar to that of conventionally mapped reads (using Bowtie2 software), which was an indication that these introns are mostly processed into circles in a similar fashion to that seen in archaea. In addition, circular intron junction-spanning reads were found in sequencing datasets from *C. elegans*.

Eukaryotic pre-tRNA cleavage is slightly different from archaeal in that it is

performed by a complex of 4 proteins rather than one (Trotta *et al.*, 1997). The heterotetrameric eukaryotic tRNA splicing endonuclease (TSEN) complex is functionally homologous to the archaeal protein; however, the RNA elements that define the sites of cleavage are different (Yoshihisa, 2014). Many eukaryotic pre-tRNAs indeed contain a BHB-like motif, and the eukaryotic endonuclease complex has been shown to be capable of recognizing and splicing pre-tRNAs with archaeal features (Di Segni *et al.*, 2005), implying conservation of the structural elements underlying substrate recognition by TSEN. However, the currently accepted model of splicing by eukaryotic TSEN relies on recognition of the tRNA body, independent of a BHB, and splicing is dictated by a “ruler” mechanism whereby splice sites are chosen based on distance from the recognized tRNA body (Greer, Soll and Willis, 1987; Reyes and Abelson, 1988). Unlike archaeal tRNA introns, which can be found at several positions in the tRNA, eukaryotic introns are nearly always found between the canonical positions 37 and 38 of the mature tRNA, adjacent to the anticodon (Chan and Lowe, 2009). This difference in intron positioning is thought to result from co-evolution of introns and their processing enzymes (Tocchini-Valentini, Fruscoloni and Tocchini-Valentini, 2005). After cleavage by TSEN, tRNA halves are ligated to form a mature tRNA. There are 2 distinct ligation pathways in eukaryotes. The 5'-ligation pathway involves the multifunctional Trl1 protein and is considered to be the predominant one in fungi and plants (Phizicky, Schwartz and Abelson, 1986; Englert and Beier, 2005; Wang *et al.*, 2006). Also known as the “healing and sealing” pathway, this process incorporates an external phosphate for use as the splice junction phosphate (Yoshihisa, 2014). In vertebrates as well as archaea, tRNAs are ligated by the RtcB enzyme (called HSPC117 in humans) in the 3'-ligation or “direct ligation” pathway, where an external phosphate is not

incorporated (Englert *et al.*, 2011; Popow *et al.*, 2011). In at least one case, ligation of mRNA splice products by tRNA processing enzymes can also result in intron circularization in vitro (Peschek *et al.*, 2015). The unfolded protein response in metazoans and yeast causes stress signaling, triggering an unconventional splicing of XBP1 (HAC1 in yeast) mRNA. The transcript is cleaved by IRE1 and ligated by RtcB in metazoans (Jurkin *et al.*, 2014; Kosmaczewski *et al.*, 2014; Lu *et al.*, 2015) or Trl1 in yeast (Sidrauski, Cox and Walter, 1996). In vitro reconstitution of these elements resulted in a circularized intron following human XBP1 splicing, and interestingly a circular intron was produced upon addition of either RtcB or the yeast Trl1 ligase (Peschek *et al.*, 2015), which has homologues in some animal species. It is unclear whether this circularization occurs in vivo, if there are other mRNAs whose splicing is dependent upon tRNA ligase activity, and whether the removed introns have any cellular function.

Drosophilid tric31905

Sequence conservation could indicate an important functional property for certain tricRNAs. In *Drosophilid* species, many tRNA introns are highly conserved at the sequence level. One tricRNA (tric31905) even displays compensatory mutation and preservation of base-pairing motifs, indicating a conserved secondary structure (Lu *et al.*, 2015).

Additionally, small RNA (sRNA) sequencing analysis indicates that this intron is further processed into microRNA-sized fragments (**Figure 2.3**) (Lu *et al.*, 2015). Analysis of several sRNA sequencing datasets revealed prominent read pileups in a 45-nt region in the 5' half of the intron (**Figure 2.3** Genome Browser track), corresponding to one side of the double-stranded duplex (**Figure 2.3** Structure track). Inspection of read start positions within this

pile of reads (**Figure 2.3** sRNAs track) showed evidence of phased duplex production, with 2-nt 3' overhangs. That is, the position where one abundant read ends is precisely where the next read begins, and the cuts are staggered. Such a pattern is generally taken as evidence for some sort of endonucleolytic cleavage event (e.g., dicing). Looking at sRNA sequencing from different cell lines and developmental stages (**Figure 2.3** Heatmap), this phasing is present in most samples although the sites of processing are not always precise. It is currently unclear whether these sRNAs have a function in the cell or are merely degradation intermediates. Inspection of sequence and secondary structure of other *Drosophilid* introns does not reveal common sequence motifs or elements, suggesting that processing of tric31905 is not necessarily applicable to all tricRNAs.

Research goals

In this chapter, I focus on proving that the predominant form of the intron (tric31905), processed from the tyrosine pre-tRNA transcript CR31905, is indeed circular, and that its expression is dependent upon ligase activity of the enzyme RtcB. I also sought to characterize the pattern of expression of tric31905 throughout the fruit fly life cycle and in various tissues, given its interesting properties such as a high degree of sequence conservation and potential to be processed into small RNAs, along with its novelty as a bona fide tricRNA found in animals.

Results

Divergent PCR of tric31905

To confirm that tric31905 is indeed circular, we carried out reverse transcription

(RT)-PCR using total RNA from third-instar larvae and four different primer pairs, two of which are divergent and two that are convergent (**Figure 2.4**). Consistent with expectations for a circular RNA, the divergent primer pairs were also able to generate PCR products. Moreover, we detected the presence of concatameric PCR products that are known to form as a result of reverse transcription on circular templates (Danan *et al.*, 2012).

RNase R

I performed Northern blotting on total larval RNA using an oligoprobe that spans the putative circular junction. The probe detects a prominent band running at ~200nt in a 10% polyacrylamide gel. In order to further demonstrate that this band represents a circular RNA, we carried out RNase R digestion of total larval RNA followed by Northern blotting (**Figure 2.5**). RNase R is a 3'-5' exoribonuclease that degrades most linear RNAs with the important exception of tRNAs (Cheng and Deutscher, 2002). The presumptive circular RNA is resistant to digestion with RNase R, whereas the control linear RNAs, U1 and U4 small nuclear RNAs (snRNAs), are degraded.

RNase H

I also used RNase H digestion as a diagnostic assay to determine if the prominent band seen in Northern blotting was likely to represent a circular intron. RNase H cleaves phosphodiester bonds of RNA in a double-stranded RNA:DNA hybrid. I incubated total larval RNA with an oligonucleotide (dr1) that is complementary to the tric31905 intron sequence in order to generate an RNA:DNA hybrid that will allow cleavage at a specified

sequence. Following RNase R digestion, I performed Northern blotting to visualize the products of the reaction using the oligoprobe cr1. RNase R can cleave randomly along the length of the hybrid, and I expected to see several patterns of digestion depending on the identity of the cleaved molecule (**Figure 2.6**). The cleavage reaction produced a smear at ~100 nt (**Figure 2.6**), indicating that the cleaved RNA is most likely either a circular intron or a pre-tRNA containing the intronic sequence.

Aberrent electrophoretic mobility

Circular RNAs have been shown to have generally retarded electrophoretic mobility with increasing concentrations of polyacrylamide when compared to linear species (Salgia *et al.*, 2003). I performed Northern blotting on total larval RNA using an oligoprobe that spans the putative circular junction. The probe detects a prominent band running at ~200 nt in a 10% polyacrylamide gel, whereas this band runs at ~125 nt in a 6% gel (**Figure 2.7**). This observation is consistent with the documented migration patterns of circular RNAs.

RtcB knockdown experiments

In metazoa, ligation of the two exonic tRNA halves is thought to be carried out by a complex that contains RtcB-like proteins (Popow *et al.*, 2011; Kosmaczewski *et al.*, 2014). To determine whether RtcB is also important for formation of tricRNAs, we carried out RNA interference (RNAi) in vivo. Database analysis identified CG9987 as the putative *Drosophila* RtcB ortholog. Using the GAL4:UAS system (Dietzl *et al.*, 2007), ubiquitous expression of dsRNA targeting CG9987 resulted in pupal lethality. Northern

blot and RT-PCR analysis of total larval and pupal RNA from these animals revealed a significant reduction in tric31905 expression, relative to controls (**Figure 2.8**). Despite the fact that the fruit fly genome contains 10 tRNA:Tyr_{GUA} genes (each of which bears an intron), expression of mature tRNA:Tyr_{GUA} was also reduced following depletion of CG9987 (**Figure 2.8 B**). These results provide strong support for the hypothesis that CG9987 is the *Drosophila* RtcB ortholog, and that this protein is required for proper formation of tricRNAs and mature tRNAs.

Developmental expression pattern

I examined expression of tric31905 over developmental time and found that its levels increase substantially from embryos to adults (**Figure 2.9**). In contrast, expression of the mature tRNA (Tyr_{GUA}) was relatively constant. Notably, tric31905 expression was slightly lower in adult females, perhaps because the ovary is such a large organ and expression levels in these tissues were relatively low (**Figure 2.9**). Consistent with their embryonic origin, Schneider2 cells also showed relatively low levels. Additional experiments will be required to determine if expression of tric31905 is actively regulated; it is possible that the RNA simply builds up over time in the organism, peaking in adults.

Discussion

We originally discovered the existence of tricRNAs from analysis of RNA-seq reads using the modified read-mapping software Vicinal. Many reads were found to map in a chiasitic manner to the 5' and 3' ends of tRNA intronic sequences, and one logical

interpretation of these data are that the reads are generated from circularization of these introns. Since it has been reported that tRNA introns are circularized in Archaea, we hypothesized that this was also the case in metazoan species but that this process had not been revealed due to strict read-mapping constraints that are typical in many RNA-seq pipelines. This hypothesis is also supported by the evolutionary conservation of tRNA splicing machinery between Archaea and metazoans.

We first sought to thoroughly prove the existence of circular RNAs derived from *Drosophila* tRNA introns using a variety of molecular biology techniques. The first technique used was amplification of complementary DNA (cDNA) products of reverse transcription of these molecules using divergent primers, which face away from each other, that would specifically amplify cDNA from a circular template. Furthermore, reverse transcription of circular RNAs results in cDNA concatemers that yield a “ladder” of PCR products when amplified. The sizes of these products can be predicted based on the length of the circular RNA. Divergent primer RT-PCR resulted in the expected pattern of amplification products for tric31905, giving us our second indication that these molecules were circular. Following up on this, I used Northern blotting to visualize tric31905. Blotting revealed a prominent band migrating at ~200 nt which we suspected was tric31905 based on the approximately equal density of sequencing reads crossing the putative circular junction in relation to those mapping to the rest of the body of the intron. To further demonstrate its circularity, we treated total RNA with RNAses (RNase R or RNase H) that would give clues to the physical nature of these RNAs. RNase R is known to degrade primarily linear RNA species, however structured RNAs such as tRNAs are resistant to RNase R similarly to circular RNAs. The results of the RNase H experiment suggest that the

RNA targeted for degradation is likely either a circularized intron or a pre-tRNA. Finally, I varied the concentration of polyacrylamide during gel electrophoresis to determine if these molecules migrated aberrantly in relation to linear standards, as is characteristic of circular RNAs. The electrophoretic migration of the prominent band in Northern analysis is retarded in higher percentage polyacrylamide, which is consistent with the documented migration pattern of circular RNAs. Taken together, these results indicate that the predominant form of this particular tRNA intron is circular. Additional experiments revealed that the fly homolog of RtcB is necessary for proper generation of tricRNAs and mature tRNAs.

Having been successful in identifying tric31905 as a bona fide tricRNA and visualizing it via Northern blot, I sought to determine its expression pattern in specific tissues and through fly development. I found that tric31905 is expressed at low levels in embryos and in the embryo-derived Schneider 2 cell line, as well as in ovaries. Expression seems to increase over developmental time, plateauing in adults while tRNA_{Tyr} remains relatively constant in all samples. I observed a bias in expression with males having higher expression than females. This could be explained by the large size of the ovaries in the female, which are shown to have low expression. Further experimentation is necessary to determine if expression of tric31905 is regulated or if it tends to accumulate with organismal age as a result of continued tRNA metabolism through the life of the animal.

Methods and Materials

RT-PCR

The following primers were used for testing the circularity of the tricRNAs. Note

that in the primer names, c is convergent, d is divergent, f is forward, and r is reverse.
tric31143cf: GCATT CAGTC TTGCC CTTCGC. tric31143cr: CGCTC GTACA CGCCC TTCT.
tric31143df: AGAAG GGCGT GTACG AGCG. tric31143dr: GCGAA GGGCA AGACT GAATGC.
PCR using both convergent and divergent primers for tric31143 generates 40-bp products.
tric31905cf1: GGCTT GGTGA TGCGA GCAGAG. tric31905cf2: TGATG CGAGC AGAGG
CAGGAC. tric31905cr1: ACTGA TGAAA CAAGA GCAAGCGC. tric31905df1: CGCTT GCTCT
TGTTT CATCAGTT. tric31905dr2: CTGCC TCTGC TCGCA TCACCA. tric31905dr1: TGGTA
TTGTT TGGCC CAGAGTC. tric31905cf1 and tric31905cr1 generates an 82-bp product.
tric31905cf2 and tric31905cr1 generates a 75-bp product. tric31905df1 and tric31905dr1
generates a 100-bp product. tric31905df1 and tric31905dr2 generates a 78-bp product.

Northern blotting

RNA samples were electrophoresed through 6% or 10% TBE-urea gels (Life Technologies). Following electrophoresis, RNA was transferred to a nylon membrane. For the analysis of tricRNAs, the following oligonucleotides were 5'-end labeled with ³²P as probes for Northern blotting: tric31905cr1 and tric31905dr1 for tric31905; dmtRNALr1 for tRNA:Leu_{CAA} and dmtRNAYr1 for tRNA: Tyr_{GUA}. dmtRNALr1: GGGAT TCGAA CCCAC GCCCTC. dmtRNAYr1: GCCGG ATTTG AACCA GCGAC CTATG.

RNase H Digestion and Northern blotting

Total RNA was isolated from third instar Oregon-R larvae using TRIzol reagent and dissolved in TE buffer. 10 µg of total RNA was mixed with 100 pmol of the complementary oligodeoxynucleotide (dr1, sequence: TGGTA TTGTT TGGCC CAGAGTC) and the mixture

was heated to 70C for 10 minutes and allowed to cool to room temperature. After cooling, 10X RNase H buffer and 5 U of RNase H (NEB cat#M0297S) were added and the mixture was incubated at 37C for 30 minutes. Samples were then used for electrophoresis and northern blotting as described above using the tric31905 junction probe.

RNase R Digestion and Northern blotting

Total RNA was isolated from third instar Oregon-R larvae using TRIzol reagent and dissolved in TE buffer. Equal amounts of RNA (10 µg) were mixed with 10x RNase R buffer, 10 U of RNase R (Epicentre Biotechnologies, Cat# RNR07250) and additional MgCl₂ to bring reactions to 1 mM [Mg²⁺]. RNA samples were then incubated for 20 min at 37°C, after which they were electrophoresed in a 10% TBE-urea gel (Life Technologies). Gels were stained for total RNA and imaged. RNA was then transferred to a nylon membrane. The following oligonucleotides were 5'-end labeled and used as probes for Northern blotting: tric31905cr1, tric31905dr1, tric31905junction: AAGCC TCGGA AACTT TCAAC ATATCGC; U1 snRNA: GAATA ATCGC AGAGG TCAAC TCAGC CGAGGT; U6 snRNA: CTTCT CTGTA TCGTT CCAAT TTTAG TATAT GTTCT GCCGA AGCAAGA.

Fly genetics and preparation of RNA samples

Two RNAi lines for RtcB (CG9987) were obtained from VDRC: line 36198 and line 36494. These two RNAi lines were crossed with an Actin-Gal4 driver to express the dsRNAs. Three samples were collected from each genotype and developmental stage. For VDRC36198 Actin-Gal4, VDRC36198/Tm3 (Tubby, without driver), VDRC36494 Actin-Gal4, and VDRC36494/Tm3 (Tubby, without driver), RNA was extracted from wandering third

instar larvae and early-to-medium stage pupae. RtcB PCR primers are as follows: RtcB.f1, CGTTCG GTGGA ACCAT GGGCA; RtcB.r1, AGATT GCGCC GGGAT TTGGCT.

Developmental Northern blot

To analyze the expression of tricRNAs during fly development, fly embryos, larvae, pupae, adults, ovaries, and S2 cells were collected. The developmental samples were collected daily as follows: embryo, 1- to 4-d-old larvae, 1- to 4-d-old pupae, male adults, and female adults. The stages prior to eclosion were collected on a ~16-h window. Male and female adults were mixed ages. Ovaries were dissected from 2- to 6-d-old females. Samples were analyzed by Northern blot as described above.

Figures

Figure 2.1: Production of circular RNAs via tRNA splicing. The tRNA splicing endonuclease (TSEN) recognizes pre-tRNA and cleaves at a characteristic bulge-helix-bulge (BHB) motif. The tRNA ligase then acts on both the cleaved tRNA halves and the intron to produce a mature tRNA and a circularized intron. This figure was adapted (Noto *et al.*, 2017).

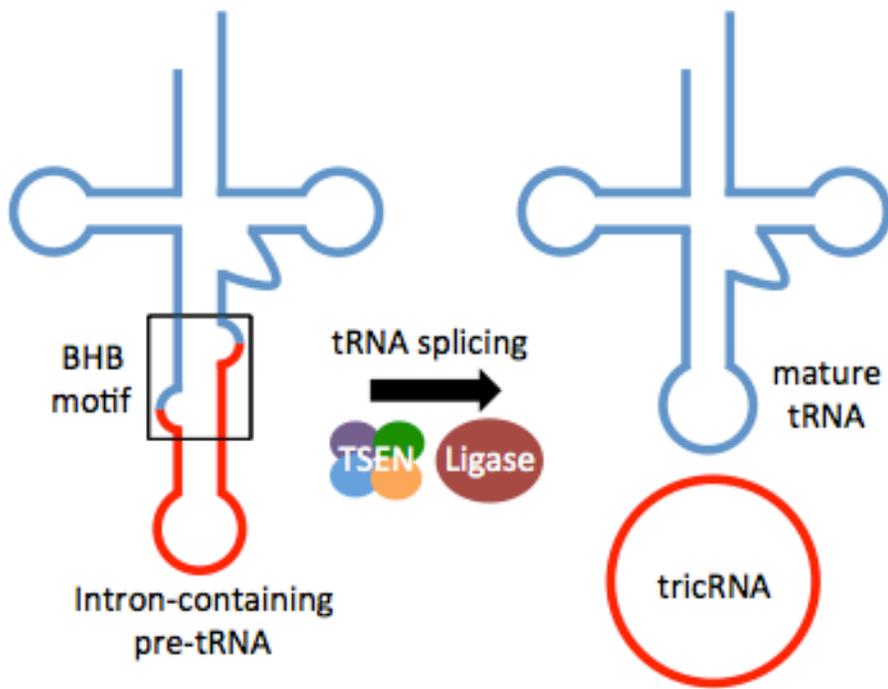


Figure 2.2: RNA-seq reads spanning the circularized junction of tRNA introns. A schematic of a tRNA gene body is shown with reads, mapped using Vicinal, aligned and that span a putative circular junction of the intron.

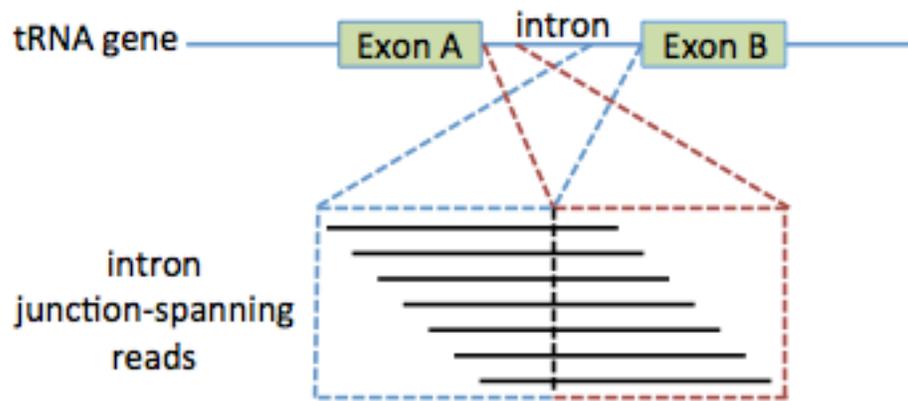


Figure 2.3: Processing of tric31905 into small RNAs. Structure track: the predicted secondary structure of tric31905 in linear format. sRNAs track: diagram of the RNAs observed in small RNA-seq data. The solid arrows represent the most prominent species, whereas the dashed arrows represent the opposite strands. Genome Browser track: UCSC Genome Browser view of the read pileup (vertical scale 0–15,000). Read starts: numbers of reads starting at each position in the Genome Browser track (vertical scale: 0–15,000). Heatmap: the numbers of reads that start at each position for every small RNA-seq data set. Each row of the heatmap is normalized to a range of 0–1. (Right side) Cartoon showing locations of small RNAs along the secondary structure of tric31905, revealing a processing pattern similar to that of miRNAs. This figure was adapted (Lu *et al.*, 2015).

Figure 2.4: Divergent PCR validation of tric31905 circularity. Cartoon shows schematic of an intronic tRNA gene with convergent (Con) and divergent (Div) primer pairs. Four pairs of primers (two Div and two Con) were tested. The ladders of PCR products correspond to amplification of concatemers of the cDNA via rolling circle reverse transcription. Experimental results agree well with the expected lengths (in base pairs) of the PCR products. Lane 1: $82 + 113n$; lane 2: $75 + 113n$; lane 3: $100 + 113n$; lane 4: $78 + 113n$; $n = 0, 1, 2, \dots$ etc. This figure was adapted (Lu *et al.*, 2015).

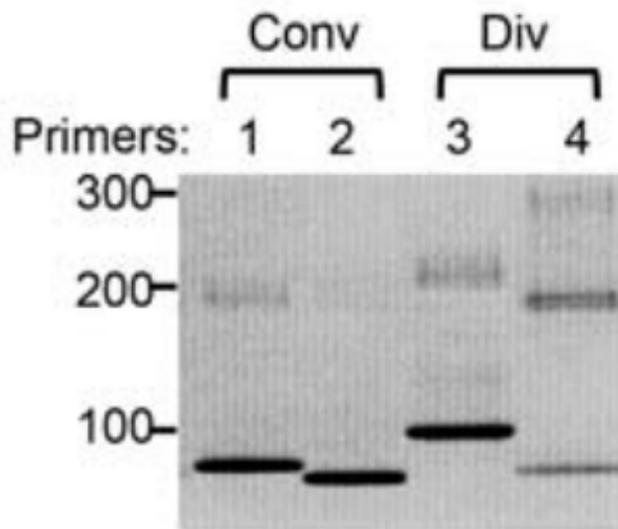
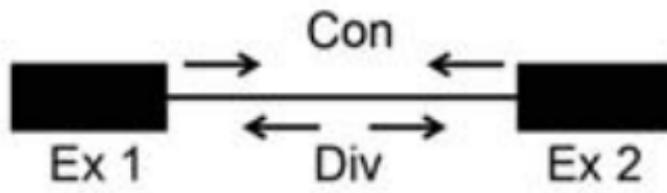


Figure 2.5: RNase R validation of tric31905 circularity. Total larval RNA samples were treated with or without RNase R and run on a 10% TBE-urea gel. RNA was imaged using SYBR Gold (left panel), whereas tric31905 (right upper panel), U1 and U4 snRNAs (right lower panel) were detected by Northern blotting. This figure was adapted (Lu *et al.*, 2015).

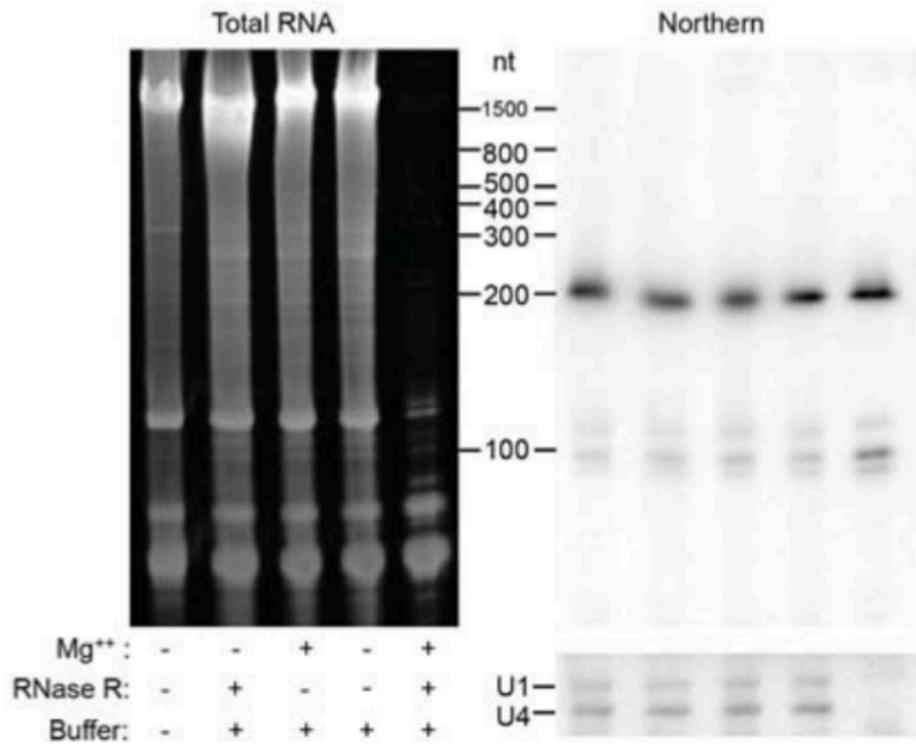


Figure 2.6: Products of RNase H digestion of tric31905. (A-C) Expected sizes of products of RNase H treatment depending on the identity of the degraded molecule: pre-tRNA (intron in orange), circular intron, or linear intron. A DNA oligonucleotide dr1 (green) was used to form RNA:DNA hybrid which RNase H cleaves, and cr1 (red) was used as a probe for Northern blot analysis. (D) Map of intron types with oligonucleotide probe binding sites shown. (E) Northern blot of RNase H products probed with tric31905 cr1 oligonucleotide.

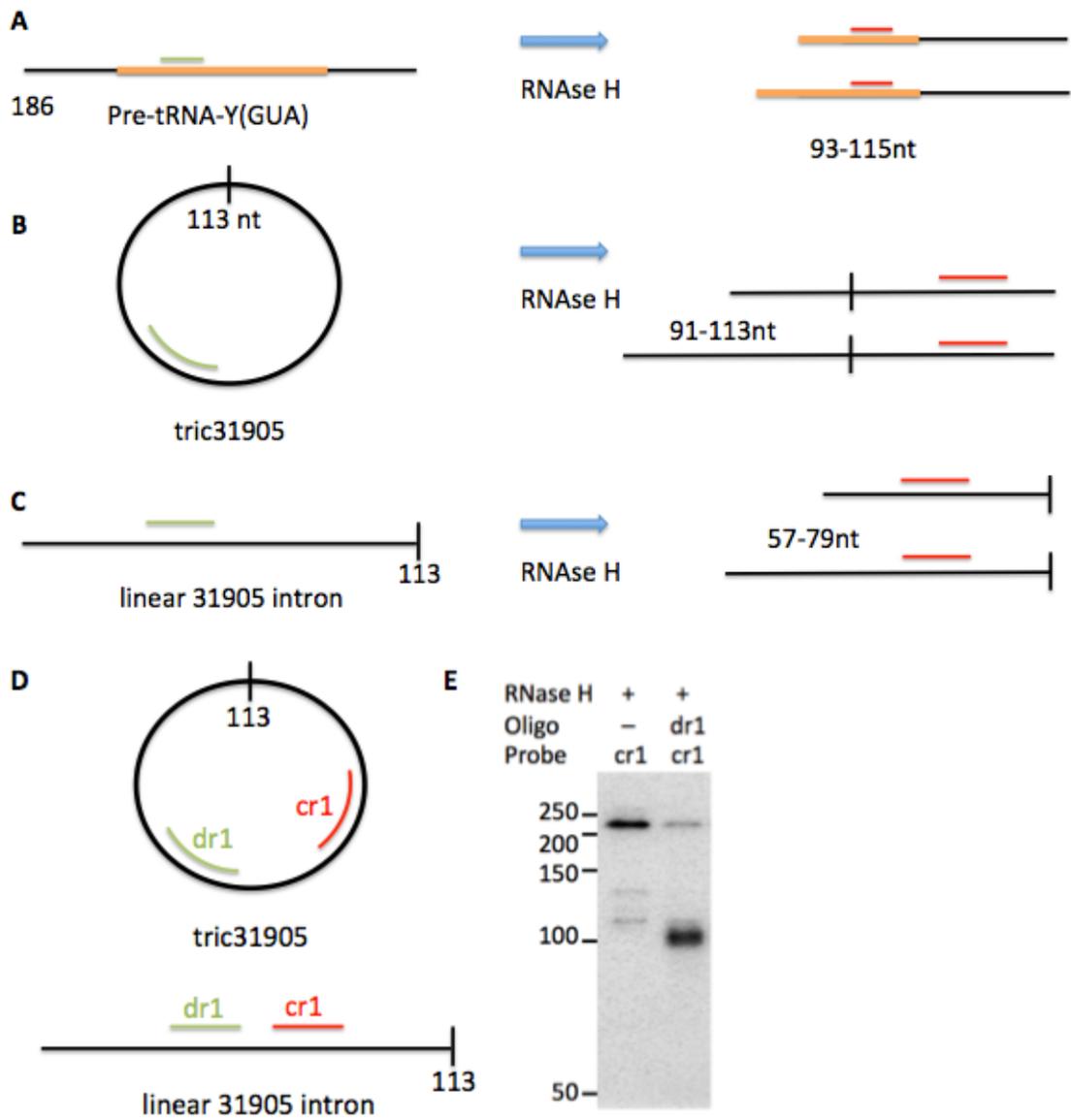


Figure 2.7: Aberrant electrophoretic mobility of tric31905 in gels with different agarose concentrations. Total larval RNA samples were electrophoresed through 10% or 6% TBE-urea gels and tric31905 was detected by Northern blotting using a radiolabeled circular RNA junction-spanning oligonucleotide probe. This figure was adapted (Lu *et al.*, 2015).

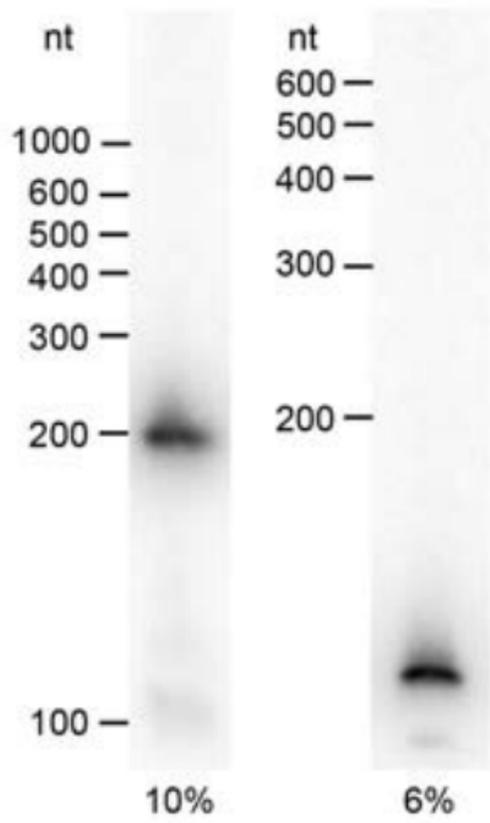
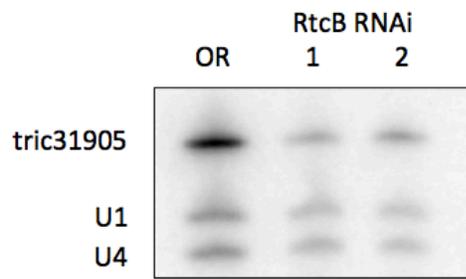


Figure 2.8: Knockdown of tRNA ligase RtcB results in decreased tricRNA expression in flies. (A) Northern blot analysis shows tric31905 expression is lower in wandering third instar larvae expressing either of two shRNAs targeting CG9987 (RtcB). (B) Knockdown of CG9987 reduces expression of tric31905 as well as that of mature tRNA:Tyr_{GUA}. RT-PCR was performed on total RNA samples from larvae or pupae. Error bars represent standard deviations from three biological replicates. (*) P < 0.05, (**) P < 0.01, (***) P < 0.001, Student's t-test. Panel B was adapted (Lu *et al.*, 2015).

A



B

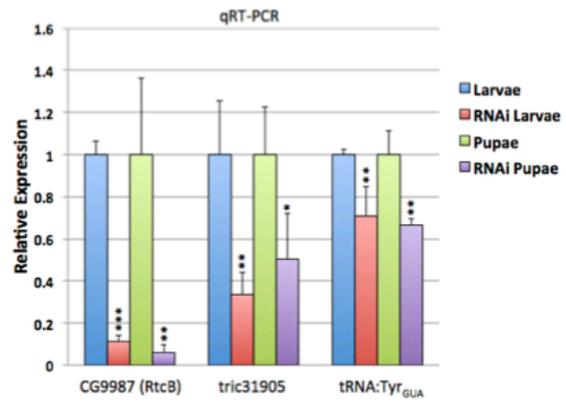
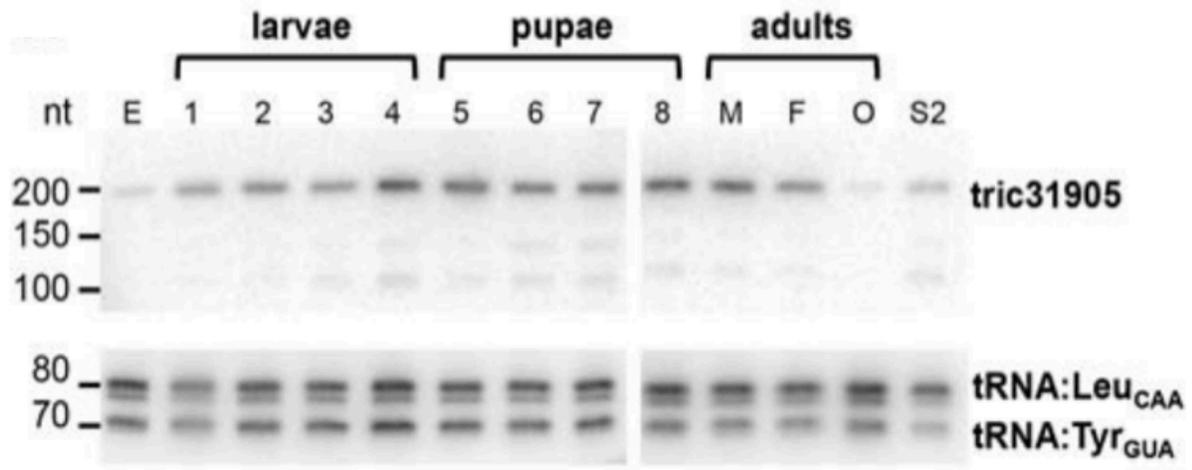


Figure 2.9: Developmental pattern of tric31905 expression in fruit flies. Northern blot assay to show tric31905 expression at various fly developmental stages and in different cell types. tRNA:LeuCAA and tRNA:TyrGUA were used as loading controls. (E) embryo, (1-4) 1- to 4-d-old larvae, (5-8) 1- to 4-d-old pupae, (M) male, (F) female, (O) ovary, (S2) S2 cells. This figure was adapted (Lu *et al.*, 2015).



CHAPTER 3: ENGINEERED EXPRESSION AND TISSUE- AND AGE-SPECIFIC ACCUMULATION OF tricRNAs IN ANIMALS³

Portions of this chapter were previously published (Noto *et al.*, 2017)

Introduction

Synthetic biology represents a fast growing and highly impactful field of research. Genetic manipulation through selection of desirable traits has had far reaching consequences for humanity. Synthetic biologists are now ushering in a new era of genetic manipulation, using sophisticated methods to tinker with and even synthesize complete, functional genomes. Many initiatives have been undertaken to reduce genomic size and complexity in model organisms to facilitate creation of modular platforms for synthetic biology design. In addition to genome-scale modulation, researchers are also introducing exogenous nucleic acids by means such as gene therapy in order to treat many human diseases. Gene delivery is accomplished by myriad techniques, and many nucleotide-based therapies including delivery of therapeutic RNAs have been established as effective drugs. Molecular stability within cells is essential to attain a favorable therapeutic outcome in many cases. Due to the increased stability of circRNAs relative to linear, they represent an

³ Portions of this chapter previously appeared in *RNA Biology*. The original citation is as follows: Noto, J. J., Schmidt, C. A. and Matera, A. G. (2017). Engineering and expressing circular RNAs via tRNA splicing. *RNA Biology*, 14(8), 978-984.

intriguing candidate to serve as functional biological effector molecules, therapeutic or otherwise.

Methods of ectopic circRNA expression and detection

To date, few tools exist to direct generation of RNA circles *in vivo*. Several groups have constructed mini-gene vectors that contain complementary sequences flanking the region of interest, designed to facilitate its circularization in a mechanistically similar fashion to endogenous exonic back-splicing (Hansen *et al.*, 2013; Kramer *et al.*, 2015; Wang and Wang, 2015). Effective circularization has been achieved using both endogenous and designed elements, or some combination thereof. Such engineered mini-genes have also shed light on cis-acting sequences that either promote or inhibit circularization (Liang and Wilusz, 2014; Kramer *et al.*, 2015; Wang and Wang, 2015). These ectopic circRNAs are capable of serving as a template for protein translation following inclusion of methyl-6-adenosine RNA modification sites (Yang *et al.*, 2017) or an internal ribosomal entry site (IRES) in the circularized transcript (Wang and Wang, 2015).

Back-spliced circRNA expression systems are typically driven by an RNA pol II promoter such as the one present in human cytomegalovirus (CMV) (Kramer *et al.*, 2015; Wang and Wang, 2015). Ectopic expression of tricRNAs, in contrast, relies on an RNA pol III promoter to drive transcription (**Figure 3.2 A**). Robust expression can be achieved with an external pol III promoter, such as the one from the U6 snRNA gene (Lu *et al.*, 2015; Schmidt *et al.*, 2016). Pol III transcription can exceed that of pol II, and produces more consistent levels of expression across cell types as compared with the CMV promoter (Qin *et al.*, 2010; Lebbink *et al.*, 2011).

Accurate characterization of circRNA expression can be difficult, however, due to a wide variety of circRNA sizes. Divergent primers are used to specifically amplify circRNA-derived cDNA (**Figure 3.1 A-C**). During cDNA preparation, reverse transcriptase is capable of rolling circle cDNA synthesis on a circular template RNA, resulting in concatemeric cDNA (**Figure 3.1 B-C**) that could lead to exaggerated measurement of circRNA levels by sequencing or RT-PCR. Indeed, concatemeric reads containing multiple copies of a tRNA intron were detected in *Drosophila* and *C. elegans* RNA-seq data sets (Lu *et al.*, 2015). For larger circles, concatemerization is less likely due to the limited processivity of reverse transcriptase. However, RT-PCR of a short circRNA can produce concatemers that are many times longer than the original transcript (Schmidt *et al.*, 2016). Thus, circRNA length must be carefully considered when using RT-based methods to quantify transcript abundance, and direct measurement of circRNA abundance by fluorescence or northern blotting may be preferred. Indeed, we have chosen to use Northern and fluorescence to assay expression of our engineered circRNAs. We have successfully expressed ‘designer’ tricRNAs greater than 800 nt in length (**Figure 3.1 D**), the primary transcript of which is well over 900 nt. To our knowledge, this is the longest known pol III-based transcription unit in eukaryotes. Hence, pol III processivity should not be a limiting factor for many applications aimed at expression of engineered tricRNAs *in vivo*.

Summary

In this chapter, we quantitatively compare methods of RNA circularization with application in mind; for most conceivable applications, abundance of expression of circles will probably be a key determinant in the effectiveness. In addition, I have established a

method for transgenic expression of tricRNAs, and tracked their accumulation over time and in specific tissues.

Results

Circular RNA expression systems comparison

To directly compare the two circRNA expression systems, we generated reporter constructs using the Broccoli fluorescent RNA aptamer system (**Figure 3.2A**) (Filonov *et al.*, 2014). Each construct is designed to express an equivalently sized aptamer-containing circle. To differentiate between the two systems, we refer to the pol II/back-spliced Broccoli reporter as circBroc, and to the tRNA-spliced Broccoli reporter as tricBroc. Following transient transfection in human HeLa and HEK293T cells, RNA was extracted and analyzed using a convenient and powerful in-gel staining assay (**Figure 3.2B**) (Filonov *et al.*, 2014, 2015; Schmidt *et al.*, 2016). RT-PCR (which is more sensitive, but less quantitative) was also performed (**Figure 3.2C**). As shown, the circBroc back-splicing reporter RNA is detectable by RT-PCR but not by in-gel fluorescence, whereas tricBroc was robustly detected by both methods. The in-gel fluorescence assay allows for detection of as little as 100 pg of Broccoli-containing RNA (Filonov *et al.*, 2015); thus, we can conclude that circBroc is expressed at levels below this detection threshold. In agreement with these data, only tricBroc was detectable by Northern blot analysis in 293T cells (**Figure 3.2D**). We note that the size of the circBroc exon used here is smaller than those used in previous circRNA expression vectors (76bp compared with 300– 1500 bp) (Liang and Wilusz, 2014; Kramer *et al.*, 2015; Wang and Wang, 2015), although it is certainly on par with the size of many endogenous exons (roughly 80% of human exons are shorter than 200 bp)

(Sakharkar, Chow and Kanguene, 2004). However, back-spliced exons average 690 bp in humans (Jeck *et al.*, 2013) and so it remains unclear whether back-splicing efficiency depends on exon size. Nevertheless, our results show that the pol III-based expression system provides more robust expression of circular RNAs (at least for smaller sized ones) than does the pol II-based system. Thus, the tricRNA expression system described here provides a compact and powerful method for *in vivo* expression of circular RNAs.

Transgenic fly creation

After testing our tricRNA expression system in cell culture, we set out to establish that this method could be used to drive the expression of designer circular RNAs in living animals. We used phi-C31 recombination (**Figure 3.3**) in flies to insert a transgene carrying a tricRNA reporter construct (U6+27-tricY-Broccoli) (**Figure 3.4 and 3.5**). After screening adult flies for transgene insertion by selecting for red eye marker, I crossed transgenic flies to balancer lines (TM6B/*tubby*). I then analyzed RNA from transgenic flies and found that expression levels of tricBroccoli was sufficient to visualize in a DFHBI-1T-stained gel (**Figure 3.6**). Propagation of these balanced stocks allowed me to then collect adult flies that were either homozygous or heterozygous for the tricBroccoli transgene. I prepared RNA from these flies and used RT-PCR and Northern blotting to show that doubling the gene dosage in the fly results in roughly twice the level of tricBroccoli expression (**Figure 3.7**).

Age-dependent accumulation

CircRNAs are thought to be generally resistant to degradation cellular exonucleases, and no general mechanism for the turnover of circRNAs has been described as of yet. Several groups have reported the age-related accumulation of circular RNAs, particularly in post-mitotic tissues such as neurons (Westholm *et al.*, 2014; Gruner *et al.*, 2016; Cortés-López *et al.*, 2017). We sought to determine whether engineered tricRNAs also accumulate with ageing, and whether this phenomenon was specific to certain tissue types. We dissected fly heads from 1-day old and 48-day old animals in order to enrich for RNA from neuronal cell types, and compared expression of our tricRNA reporter in this sample to that of headless bodies from both ages using Northern blot analysis. We found that tricBroccoli is 1.74-fold higher in RNA from heads in aged flies, compared to a 1.47-fold increase in bodies, suggesting a subtle, general accumulation of ectopically expressed tricRNAs during the life of an adult fly that may be more pronounced in certain cell types.

Discussion

Despite their conservation across great evolutionary distances, the function of tRNA introns has, for the most part, yet to be elucidated. Many interesting questions remain regarding their role in tRNA biosynthesis (as DNA elements), as well as their fate and function in the cell (as tricRNAs) following removal during processing. As with many non-coding RNAs, characterization of circular RNAs has been greatly aided by the progress of detection technologies and molecular genetic approaches. As detailed here, tricRNA vectors provide an effective means for the orthogonal expression of a wide variety of genetically engineered circRNAs both in cell culture and in animals. Use of the broccoli aptamer has

allowed us to track the expression of exogenous tricRNAs, and this technique could be used in the future to rapidly assess the processing of tRNAs as well to further determine the fate of tRNA introns in cells and animals. This work represents the first known example of genetically-encoded expression and detection of fluorescent RNA aptamers in living animals.

Methods and Materials

Expression Comparison: Plasmid construction

Annealed oligonucleotides were ligated into the pCircGFP plasmid (Zefeng Wang, PICB Shanghai) following digestion with BamHI and SacII to generate a new vector, called pCircBroc. Top oligo:

ggctcttttctccagGAGCGGCCGCGAGACGGTCCAGATATTCGTATCTGTGCGAGTAGAGTGTG
GGCTCGTGGCCGCGGTGTTGAGgtaagtctcgacg, bottom oligo: gatcc

gtcgagacttacctcaacaccgcgccacgagcccacactctactcgacagatacgaatat

ctggaccgaccgtctcgggcgctctggaagaaagagaccgc. For the top strand oligo, the exon is shown in uppercase letters and the intronic sequence is in lower case. The nucleotides corresponding to Broccoli are underlined.

Expression Comparison: DFHBI and Ethidium Bromide staining

Ten micrograms of total RNA from transfected cells was run on a Novex 10% TBE-Urea gel (Thermo Fisher Scientific cat#- EC68572BOX) for 40 minutes at 300V. After washing 3 times with ddH₂O, the gel was stained with DFHBI-1T staining solution (40 mM HEPES pH 7.4, 100 mM KCl, 1 mM MgCl₂, 10 mM DFHBI-1T [Lucerna cat#410-1 mg]) and

fluorescence imaged using a GE Typhoon Trio variable mode imager. To visualize total RNA, the gel was then stained with ethidium bromide and fluorescence imaged.

RT-PCR

Total RNA was treated with Turbo DNase (Thermo Fisher Scientific cat#AM2238) and used to prepare cDNA using SuperScript III (Invitrogen cat#18080051) with random hexamer primers. Divergent primers used for circBroc PCR amplification were as follows: F:5'-GTGGAGTAGAGTGTGGGCTCGT-3', R:5'-GATACGAATATCTGGACCCGACCGTC-3'

Expression Comparison: Northern blotting

Ten micrograms of total RNA was run on a 10% TBE-Urea gel for 40 minutes at 300V and transferred to a nylon membrane at 10 mA overnight. The membrane was dried and UV-crosslinked, followed by pre-hybridization in Rapid-hyb Buffer (GE Healthcare Life Sciences cat#RPN1635) at 42C for 30 minutes. Oligonucleotide probes were radiolabeled using T4 PNK (NEB) and then purified using Illustra Microspin G-50 columns (GE Healthcare Life Sciences cat#27-5330-71). They were then boiled for 5 minutes and cooled on ice for 2 minutes, and then added to the Rapid-hyb Buffer. The membrane was hybridized at 42C for one hour with rocking, and then washed at ambient temperature twice for 15 minutes with 2X SSC with 0.1% w/v SDS and twice for 15 minutes with 0.1X SSC with 0.1% w/v SDS at 42C before overnight PhosphorImager screen exposure. The blot was then imaged using a GE Typhoon Trio variable mode imager. Probe sequences: anti-Broccoli-5' -gagcccacactctactcgacagatacgaatatctggacccgaccgtctc-3' , anti-5S rRNA- 5' -ccctgcttagcttccgagatcagacgagat-3' .

tricBroc transgene insertion: phiC31 integration

PattB-tricBroccoli plasmid (Casey Schmidt) was injected into fly embryos (Bloomington Stock Center stock #9750, VK33 attP chromosome 3L landing site) by BestGene Incorporated. Flies were screened for transgene integration by red eye marker upon eclosion. Flies carrying the transgene were crossed to TM6B balancer stocks.

Fly dissection and Northern blotting for age and tissue comparison

For the age-related accumulation experiment, flies were aged for 1 or 48 days and frozen for 10 minutes in ethanol in dry ice. They were then frozen at -80C for two hours, after which they were vortexed to remove heads from bodies. Total RNA was prepared using Trizol reagent. Northern blot was performed as described above.

Cell culture and RNA isolation

All cells were grown at 37C with 5% CO₂ in DMEM (Gibco) supplemented with 10% fetal bovine serum and 1% penicillin and streptomycin. One million HEK293T or HeLa cells were plated in T-25 flasks and incubated at 37C. After 24 hours, cells were transfected with equimolar amounts of either circBroc or tricBroc expression vector using Fugene HD transfection reagent (Promega cat#E2311). After 72 hours, total RNA was isolated from cells using Trizol reagent (Thermo Fisher Scientific cat#15596026).

Figures

Figure 3.1: tricRNA expression can be detected by RT-PCR. (A) Schematic of the tricRNA expression construct, showing the PCR primer binding sites (red arrows) in the intron and an external pol III promoter (blue arrow). Exon A and Exon B represent the sequences present in the mature tRNA. The dotted line indicates a variably-sized intronic region. (B) Reverse transcriptase can transcribe around a tricRNA template many times, resulting in a concatameric cDNA with numerous tandem repeats. (C) RT-PCR primers can bind to multiple sites along the cDNA concatamer, resulting in a ladder of potential PCR products. The formula for the sizes of the bands is: [size of the circle] - [distance between the divergent primers] + [size of circle] n , where n is the number of tandem repeats. (D) The ladder of PCR products can be seen for circles of several different sizes. In this experiment, the distance between the divergent primers is 105 nt, and the sizes of the circles are listed above the gel. For the 259 nt circle, the formula for the ladder is therefore: $259 - 105 + 259n$. The bands detected on the gel are at 154, 413, 672, and 931 nt. This figure was adapted (Noto *et al.*, 2017).

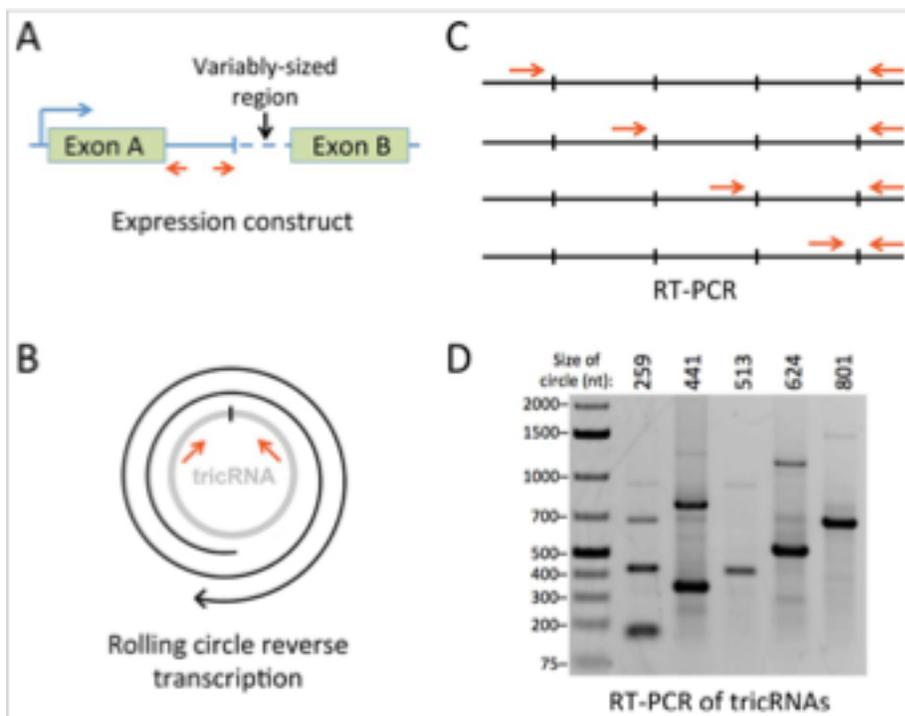


Figure 3.2: Comparison of *in vivo* methods of circular RNA expression using the Broccoli RNA aptamer as a reporter. Testing *in vivo* methods of RNA circularization using the Broccoli RNA aptamer as a reporter. (A) Schematic of the 2 constructs, where the bright green box indicates the placement of the Broccoli aptamer sequence. The dotted arcs indicate the splice junctions. The U6* promoter includes the first 27 nucleotides of U6 snRNA in the transcriptional unit. This promotes 5' capping (Good et al 1997) and enhances stability of the expressed RNA, resulting in a higher yield of tricRNAs (Lu et al 2015). (B) In-gel imaging following transient transfection of the reporter constructs into HeLa and 293T cells. The left hand image is the DFHBI-1T stain, which binds to all Broccoli-containing RNAs. In this image, the top band is the pre-tRNA and the lower band is the circular RNA. The doublet of bands in the first lane is most likely due to transcription beginning from both the external U6 promoter, which has a longer 5' leader sequence, and the internal tRNA promoter. The right hand image is the ethidium bromide stain of the same gel, which marks total RNA. (C) RT-PCR was performed on cDNA generated from the RNA used in Fig. 3B. Diverging PCR primers were used, such that they only generate products of the appropriate size from a circularized template. In this experiment, the lengths for tricBrocc and circBrocc are 77 nt and 76 nt, respectively. (D) Northern blot analysis was performed to quantitatively assess circRNA expression in 293T cells. This figure was adapted (Noto *et al.*, 2017).

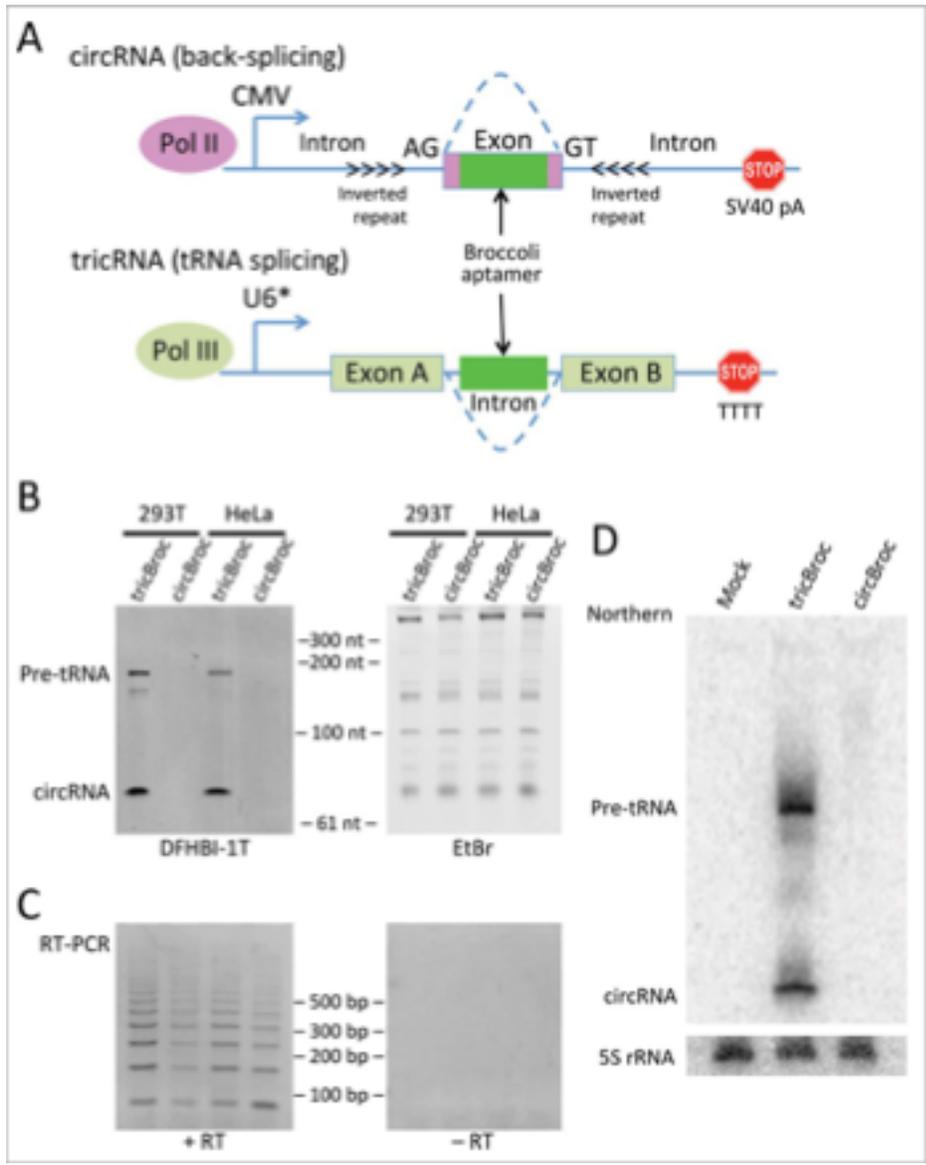


Figure 3.3: phiC31 integration of transgenes. Wild-type integrase (Int) mediates recombination between the attP and attB sequences. The transgene becomes stably integrated into the chromosome by recombining the attP of a landing site (introduced into the genome via a transposon) with the attB on the plasmid. This figure was adapted (Knapp, Chung and Simpson, 2015).

Integration

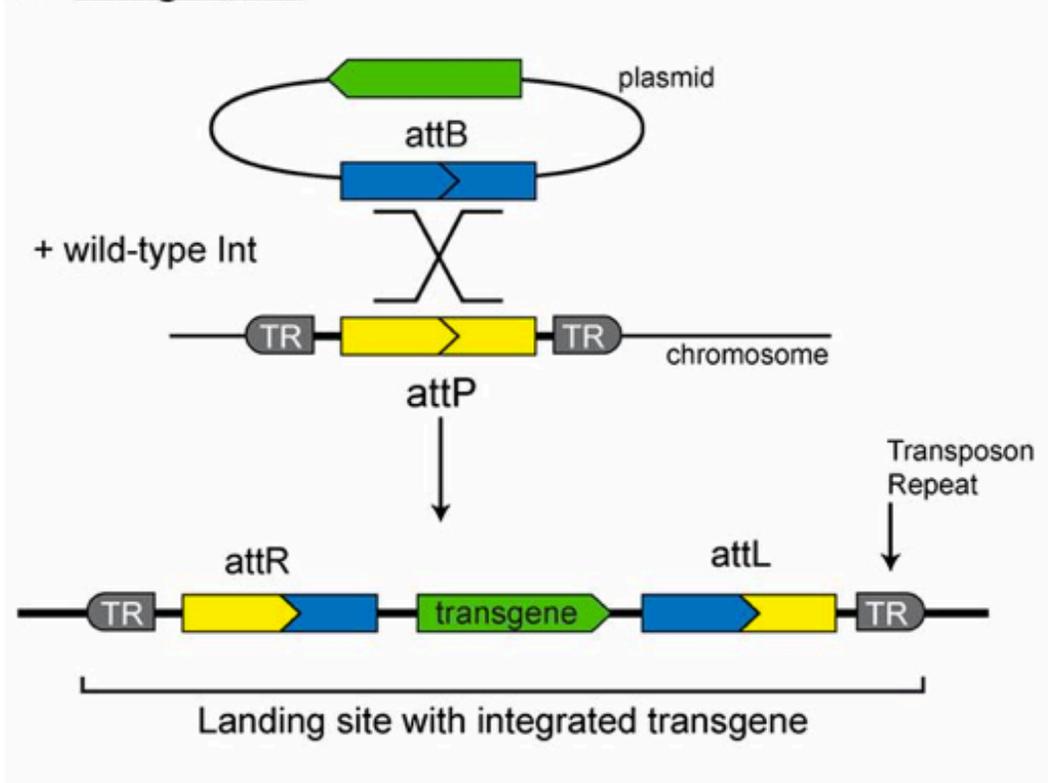
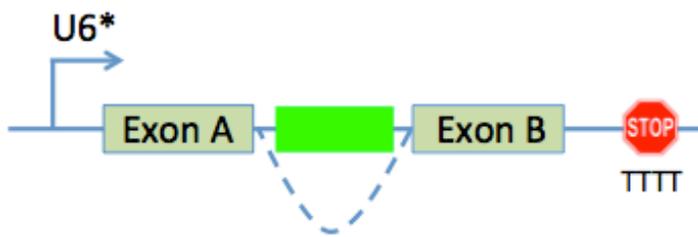


Figure 3.4: Schematic for tricBroccoli transgene insertion using phiC31 in flies. The U6*-tricBroccoli transgene was recombined into the fly genome on chromosome 3L (VK33 landing site) using phiC31 integration.



Tyrosine 31905 tRNA scaffold

phiC31 integration
(VK33, chromosome 3L)

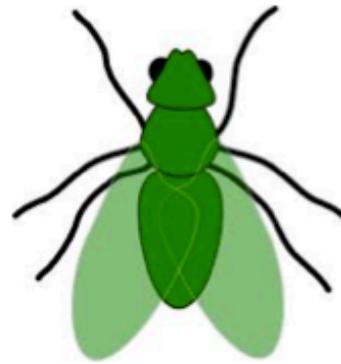


Figure 3.5: Plasmid map of pAttB-tricRNA phi-C31 vector.

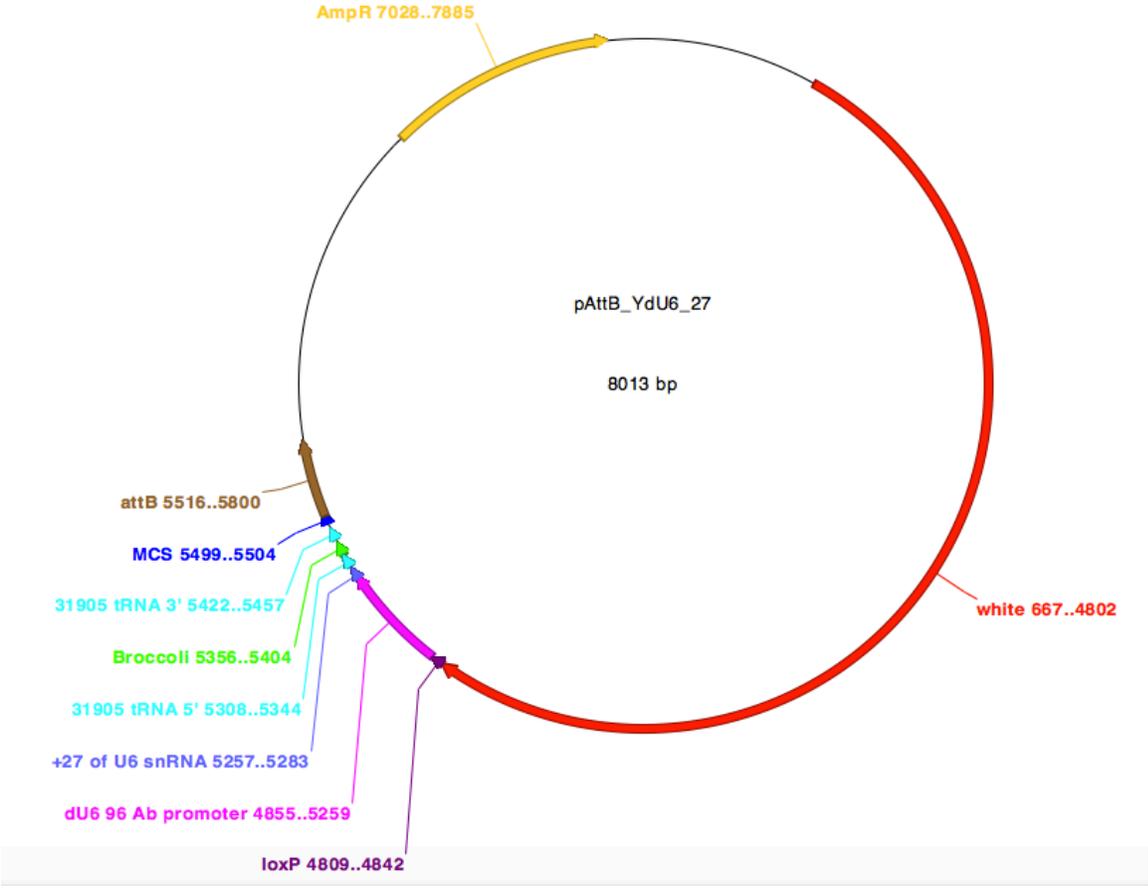


Figure 3.6: Expression of tricBroccoli in flies can be detected by in-gel DFHBI-1T staining. RNA from flies expressing the tricBroccoli transgene was electrophoresed on a polyacrylamide gel and stained with DFHBI-1T (top). Total RNA was observed after staining with ethidium bromide. RNA from S2 cell transient transfection of tricBroc construct was used as a positive control. RNA from Oregon-R flies was used as a negative control.

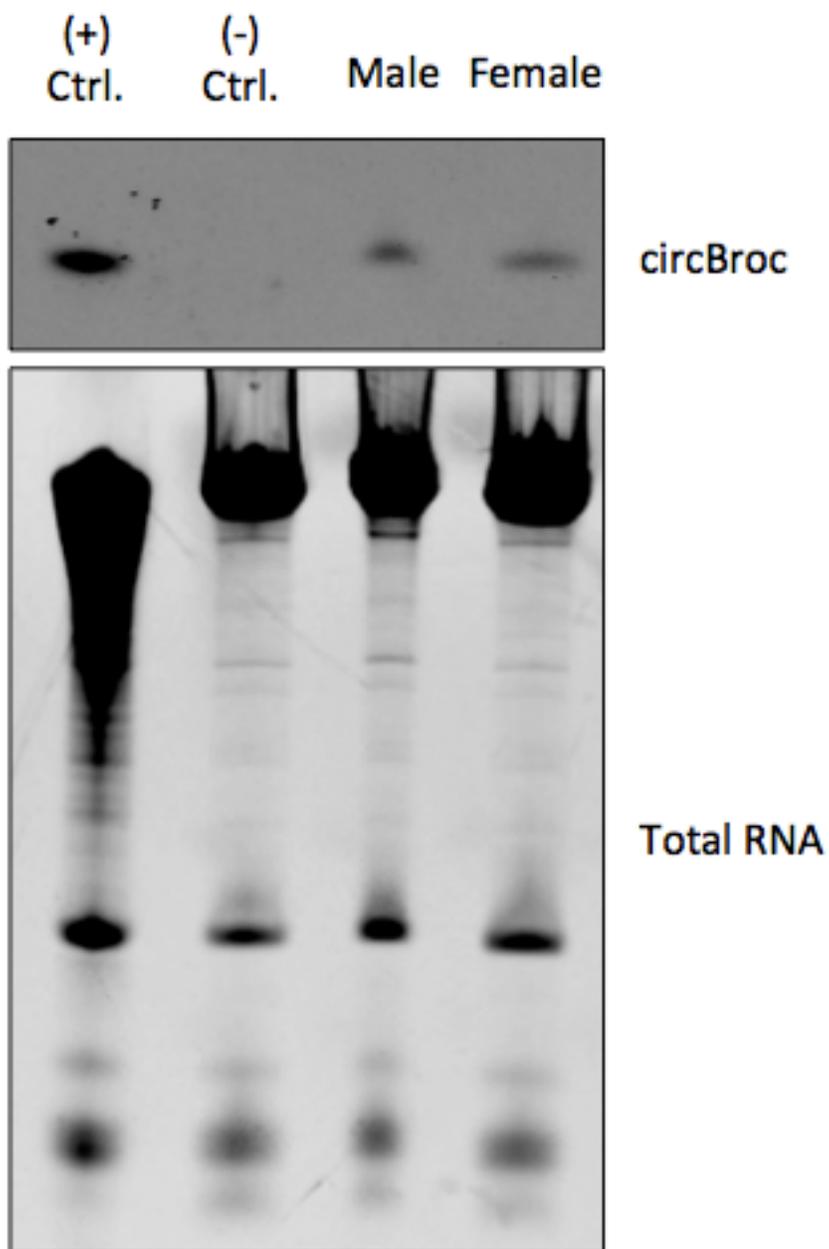


Figure 3.7: Gene dosage alters abundance of expression of tricBroccoli in fruit flies. Northern blotting was used to show expression of transgenic tricBroccoli from total RNA in male or female homo/heterozygotes, using an oligoprobe spanning the circular junction. Relative band intensities were calculated using ImageJ and normalized to U4 snRNA loading controls.

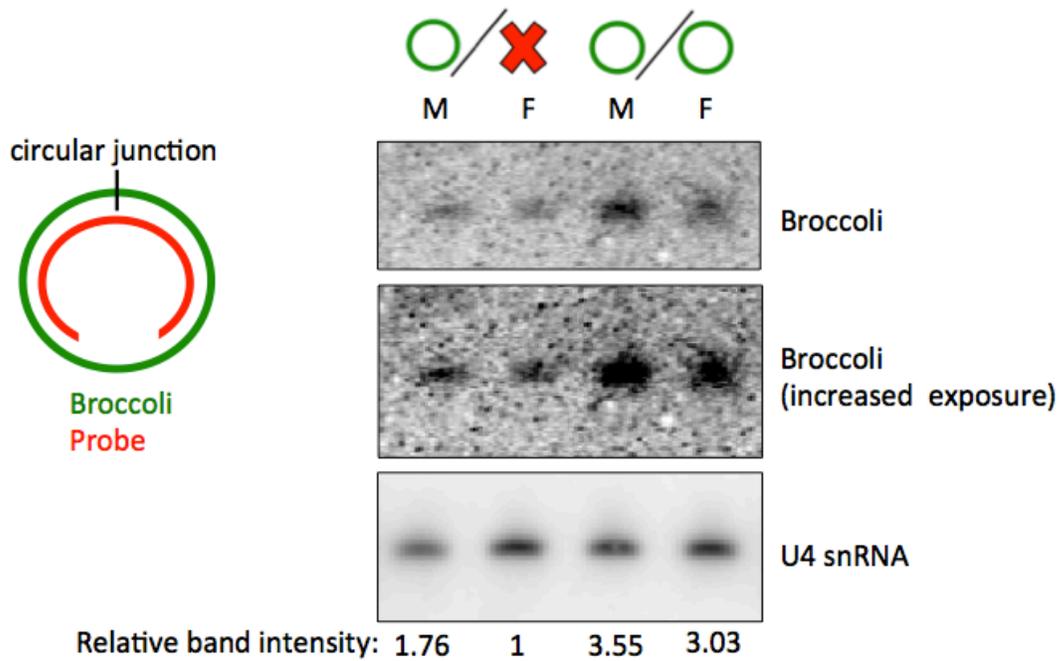
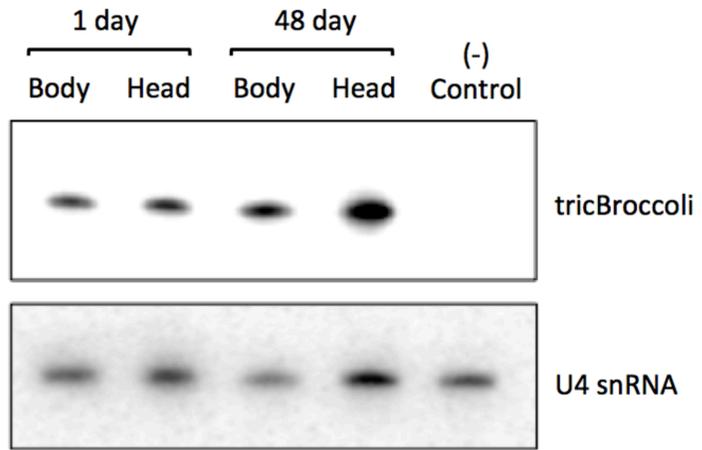
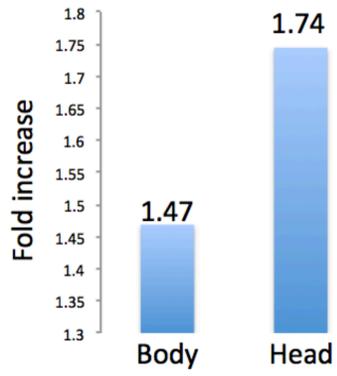


Figure 3.8: Accumulation of tricBroccoli in aged fly tissues. Northern blotting was used to quantify expression of tricBroccoli in total RNA from dissected heads or bodies from male flies 1-day post pupation or 48 days post pupation. Band intensities were calculated using ImageJ and levels were normalized to U4 snRNA loading control. Graph shows fold increase in Broccoli expression in heads or bodies in aged flies.



CHAPTER 4: CONCLUSIONS AND FUTURE DIRECTIONS

tricRNAs expression is conserved and results in a diverse expression pattern in tissues and throughout animal development

We have shown that tricRNA biogenesis occurs in metazoans through an evolutionarily conserved mechanism. This information adds to an increasingly complex portrait of the transcriptomic landscape in eukaryotes, and provides further insight into the subject of tRNA processing. Due to their stability and presence in extracellular fluids, it is speculated that circRNAs could be good biomarkers for disease. Several human diseases related to defects in tRNA processing/metabolism have been reported (Levinger *et al.*, 2004; Wilson, 2004; Karaca *et al.*, 2014), and tricRNAs could potentially be related to pathogenesis or serve as disease biomarkers. Additionally, as with some other circRNAs, tricRNAs tend to accumulate during the process of ageing and display tissue-specific expression profiles. Further studies are needed to determine whether tricRNA function and accumulation is linked to biological processes, or whether they could be good biomarkers for tRNA metabolic or age-related disorders.

tRNA splicing as an engineered platform for robust expression of stable circRNAs

Engineered systems for ectopic RNA expression can be used in many different ways, including the study of endogenous RNA function, expression of guide RNAs for CRISPR-

related applications, and delivery of therapeutic RNAs. We have shown here that tricRNA biogenesis can be harnessed to express stable circRNAs in both in cells and in animals. Our data show that this technique is advantageous over the use of a back-splicing method of circular RNA expression with regard to volume of circles produced, at least for small (<100 nt) circles. Whether this also applies to expression of large circles, such as those that can code for proteins, has yet to be tested. We have not yet met a “size barrier” in expression of engineered tricRNAs, and have thus far surpassed the length in nucleotides of any other known RNA polymerase III transcription event.

Can we engineer an inducible tricRNA expression system?

To add utility to our expression system, we can theoretically engineer an inducible tricRNA expression system using FLP/FRT recombination to remove an inserted poly-U tract that will otherwise inhibit transcription of the tRNA body and thereby inhibit expression of the tricRNA (**Figure 4.1**). Upon addition of Flippase, for instance, by crossing flies with our inducible system to flies carrying a Flippase gene, we can remove the poly-U transcription stop signal and induce tricRNA expression. Transcription of endogenous tRNAs occurs from promoter elements within the tRNA exons, called the A and B-boxes respectively (Sharp *et al.*, 1981). For a truly inducible system (where transcription does not occur at a baseline level due to built-in promoter elements), we would likely need to mutate bases in the tRNA A- or B-box that would inhibit promoter activity but preserve functionality of the mature tRNA in order to limit any deleterious effects of expression of mutant tRNAs. One study has examined the effects of B-box mutations on transcription

efficiency in *Drosophila* and HeLa cell cultures (Gaëta, Sharp and Stewart, 1990) (**Figure 4.2**). This information could provide a good starting point for mutational analysis.

Do endogenous tricRNAs have functions?

The function of tRNA introns, which are widespread in eukaryotes, is largely unknown. Several studies have shown functions such as direction of tRNA modification, but no broad explanation for their presence currently exists, whether at the level of DNA in the tRNA genes or at the level of RNA during or post-processing. Our data suggest that tricRNAs could be processed into smaller RNAs and could have downstream function, perhaps in a miRNA-like pathway. CircRNAs can also be packaged into extracellular vesicles and taken up by neighboring cells, and it is possible that tricRNAs could also function in this manner, perhaps as cellular signaling molecules. I have made an attempt to knock down endogenous tric31905 in flies using transgenic expression (driven by Actin- or Tubulin-GAL4) of dsRNA to promote RNA-interference by targeting the circularized junction (**Figure 4.3 A**), but this method was not effective in reducing levels of the tricRNA (**Figure 4.3 B**). It is possible that the high degree of base pairing of tric31905 sterically hinders accessibility of the target site. For this reason, it may be more effective to study tric31905 function by CRISPR-mediated deletion of the intron from the genome. The advent of CRISPR/Cas9 genome editing has provided researchers a quick and reliable means of specifying genomic alterations, such as deletions or specific mutations, in model organisms such as *Drosophila*. In order to determine if tRNA introns are required for development, we could use CRISPR to generate a group of mutants (**Figure 4.4**). Along with a mutant lacking the entire tRNA, we can generate mutants lacking introns but retaining

the ability to decode the codon associated with the CRISPR-mutated tRNA gene. This can be accomplished by co-injection of a synthetic covering ssDNA oligonucleotide with the sequence of interest for repair. The replacement tRNAs would also contain a point mutation in the variable loop which will allow assessment of expression of the tRNA itself using RNA-seq, and remove its loss from consideration for possible causes of observed phenotypes.

Can we engineer expression of functional tricRNAs?

Potential exists for both endogenous and engineered tricRNAs to be functional within cells. Aside from any speculation on endogenous tricRNA function, engineered tricRNAs could be used to bind and sequester cellular molecules to abrogate their function and thus be used to study functional biomolecules (miRNAs, RNA-binding proteins, etc.) without perturbing their genomic site of origin. For instance, many functional miRNAs are processed from primary transcripts encoding proteins. Mutation or deletion of miRNA-encoding genomic sequence to study their function can have effects unrelated to the processed miRNA, making it difficult to assign causation of any observed cellular consequences.

Methods

Cloning pVALIUM20 plasmid for TRiP line generation, injection/integration selection to make fly stock, and crossing to expression driver

The following oligonucleotides were annealed and ligated into *NhoI*- and *EcoRI*-linearized pVALIUM20 vector. The 21-nt sequence targeting the tric31905 circular junction is in uppercase.

Top oligo:

ctagcagtGATATGTTGAAAGTTTCCGAGtagttatattcaagcataCTCGGAAACTTTCAACATATCgcg

Bottom oligo:

aattcgcCTATACAACCTTTCAAAGGCTCtatgcttgaatataactaGAGCCTTTGAAAGTTGTATAGactg

This vector was then injected into embryos (Bloomington Stock #27510) and integration was selected by vermilion marker. Flies were crossed to Actin- or Tubulin-GAL4 lines to drive expression of the shRNA in the whole animal.

Northern blotting

Total RNA from flies was prepared and assayed by Northern blot as described above for tric31905 using a junction-spanning oligoprobe.

Figures

Figure 4.1: Inducible tricRNA expression system. Here I show a theoretical inducible tricRNA expression construct containing FRT recombination sites flanking a poly-U tract upstream of the tRNA gene, which will prevent tRNA transcription. Addition of Flippase enzyme will catalyze excision of the poly-U tract based on the position and directionality of the FRT sites, allowing transcription of tRNA to proceed.

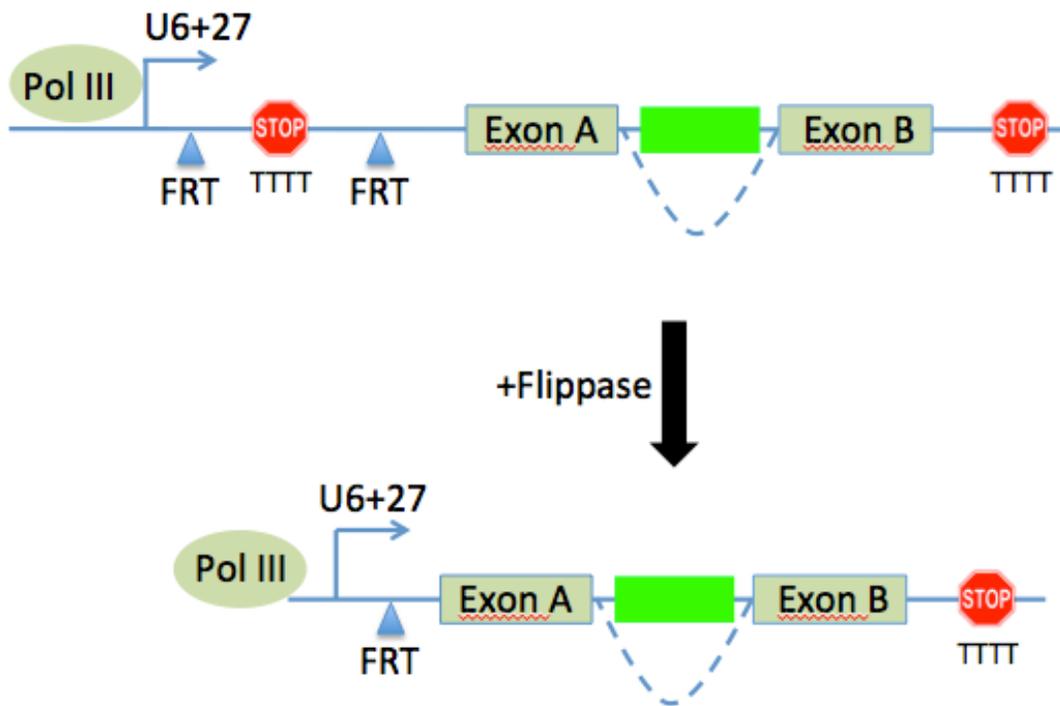


Figure 4.2: The effects of B-Box point mutations on transcription of a *Drosophila* tRNA^{Arg} gene in (A) *Drosophila* and (B) HeLa cytoplasmic extracts. This figure was adapted (Gaëta, Sharp and Stewart, 1990).

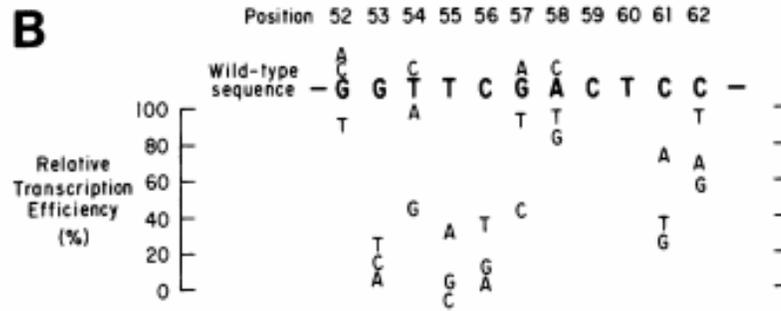
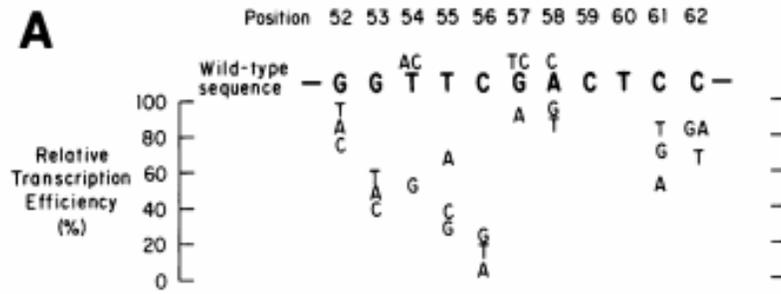


Figure 4.3: tric31905 shRNA knockdown in flies is ineffective. (A) Cartoon showing that shRNA processing generates siRNA that targets the circular junction of tric31905. (B) Driving shRNA expression in five transgenic fly lines (shRNA transgene expressed from either chromosome 2 or 3) with either Actin- or Tubulin-GAL4 does not affect tric31905 expression in flies as assayed by Northern blotting.

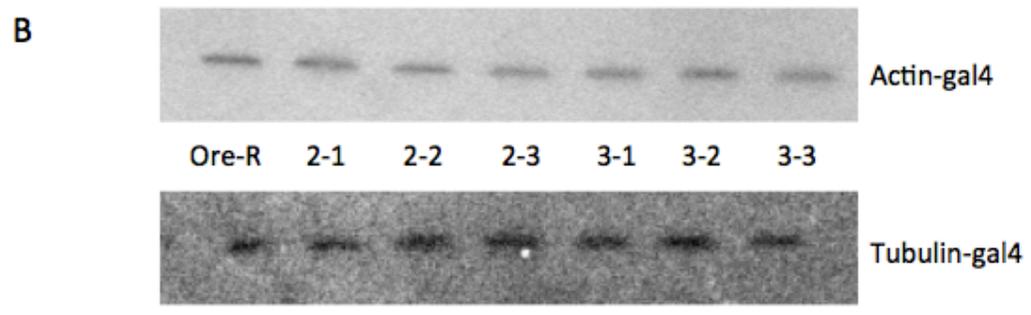
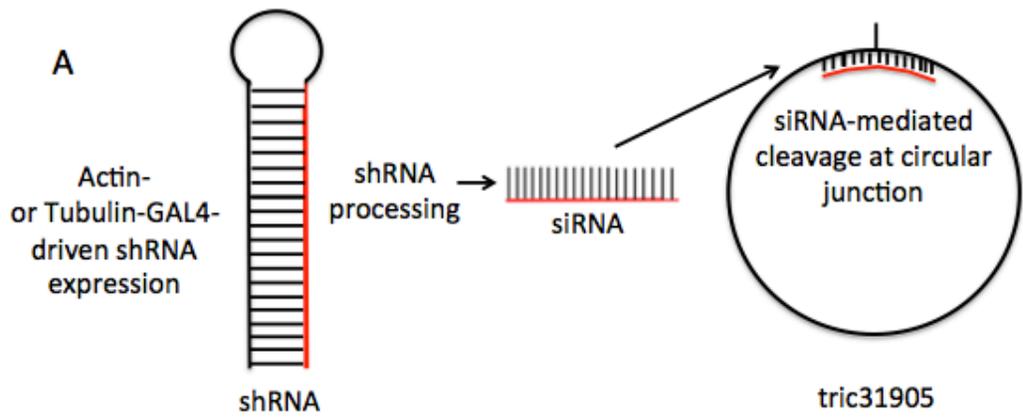
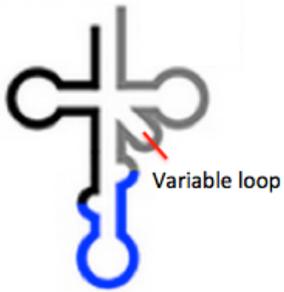
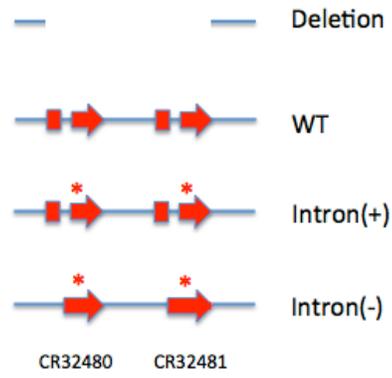


Figure 4.4: Proposed CRISPR-Cas9 tRNA-mutant genotypes for study of tric31905 function. Diagram of intronic pre-tRNA (black/gray are tRNA halves, blue is intronic sequence) with variable loop to be mutated (for tracking of mutant tRNA) indicated. Also shown: genotypes to be generated with CRISPR-Cas9 system.

Intronic tRNA



CRISPR-generated genotypes



APPENDIX 1: miR-21 SPONGE tricRNA EXPRESSION

Goals and results

MicroRNAs associate with Argonaute protein complex and cause inhibition of target protein translation through destabilization of the primary mRNA transcript or preventing the ribosomal apparatus from association and elongation (Coller and Parker, 2004). Given that endogenous circRNAs can “sponge” miRNAs and inhibit their function, and taking into consideration that our ectopically expressed tricRNAs accumulate both the nucleus and cytoplasm (Lu *et al.*, 2015), we hypothesized that we could engineer a tricRNA harboring miRNA-binding sites and inhibit a specific miRNA (or miRNA-Argonaute complex), causing its targets to be upregulated. We chose to target miR-21, a miRNA that targets many tumor suppressors and is thought to be involved in oncogenesis. This miRNA is upregulated in many cancers and in cancer cell lines, including HeLa cells (Yao *et al.*, 2009). We designed a miRNA sponge harboring four miR-21 binding sites and transfected this construct into HeLa cells, along with a luciferase reporter. This reporter had been tagged with a segment of the 3'-UTR of the tumor suppressor gene PDCD4, which contains a miR-21 binding site that targets it for translational inhibition. Thus, our luciferase reporter should become negatively regulated by miR-21 in HeLa cells due to miR-21 upregulation in this cell type. If our sponge were to be effective in inhibition miR-21 function, co-transfection of the sponge construct should result in relief of miR-21 regulation causing an increase in luciferase

signal (**Figure Appendix 1.1**). I used RT-PCR to show that the amplification product gives the laddering effect characteristic of reverse transcription of circRNAs. The results of the luciferase assay were varied (**Figure Appendix 1.3**); in all trials, tagged luciferase is repressed relative to untagged. However, the degree to which this inhibition is relieved upon co-transfection of our miR-21 sponge ranges from almost complete inhibition (trial 1) to little or no inhibition of miR-21 repression of luciferase signal. Further testing is necessary to determine whether this was the result of varying degrees of transfection efficiency, leading to decreased expression of our sponge construct, or if a re-design of the sponge is necessary (including more miRNA binding sites, for example) in order to achieve effective inhibition of miRNAs.

Methods

miR-21 sponge

The following oligonucleotides were annealed and ligated into NotI/SacII linearized pGEM-tricY tricRNA expression vector (Casey Schmidt) to make the expression construct for the miR-21 sponge:

Top: 5'-

CTGTgcgccgcTCAACATCAGGACATAAGCTATCGCAAATCGCGCGGCATTCATCAACATCAGGAC

ATAAGCTActtcTCAACATCAGGACATAAGCTACGACCGTGCAACGATCGCGTTCAACATCAGGAC

ATAAGCTAgtggccgcgACAG-3'; Bottom: 5'-

CTGTccgcgccacTAGCTTATGTCCTGATGTTGAACGCGATCGTTGCACGGTCGTAGCTTATGTCC

TGATGTTGAgaagTAGCTTATGTCCTGATGTTGATGAATGCCGCGGATTTGCGATAGCTTATGT

CCTGATGTTGAgcgccgcACAG-3'.

Primers for divergent PCR to test circularity were as follows: Sponge.df1: 5'-CGACCGTGCAACGATCGCGT-3'; Sponge.dr1: 5'-TGAATGCCGCGCGATTTGCGA-3'.

Cloning pSICHECK-2 luciferase reporter for miR-21 activity

pSICHECK-2 is a dual luciferase reporter plasmid. The following primers were used to amplify cDNA from whole flies to generate a PCR product carrying a miR-21 seed site from the 3'-UTR of PDCD4. Forward Xho1: 5'-GGCTCTCGAGGGGACTCTGGCAGAG-3'; Reverse Not1 :5'-AATGCGGCCGCGATCCCACCAGT-3'. The PCR product was ligated into XhoI/NotI linearized pSICHECK-2.

Luciferase assay

HeLa cells were transfected with 2 ug of pSICHECK-2 and miR-21 sponge construct using Fugene Reagent. After 3 days incubation, cells were treated using Promega Dual Luciferase Reporter Assay System. Luminescence was measured using a Tecan M1000 Microplate reader.

Figures

Figure Appendix 1.1: Schematic of miR-21-responsive luciferase reporter assay for miR-21 sponge effectiveness. Luciferase tagged with PDCD4 3'-UTR should be negatively regulated by miR-21, which is upregulated in the HeLa cell type. If the sponge is effective, an increase in luminescence should be observed as the sponge competing for miR-21 binding relieves this repression.

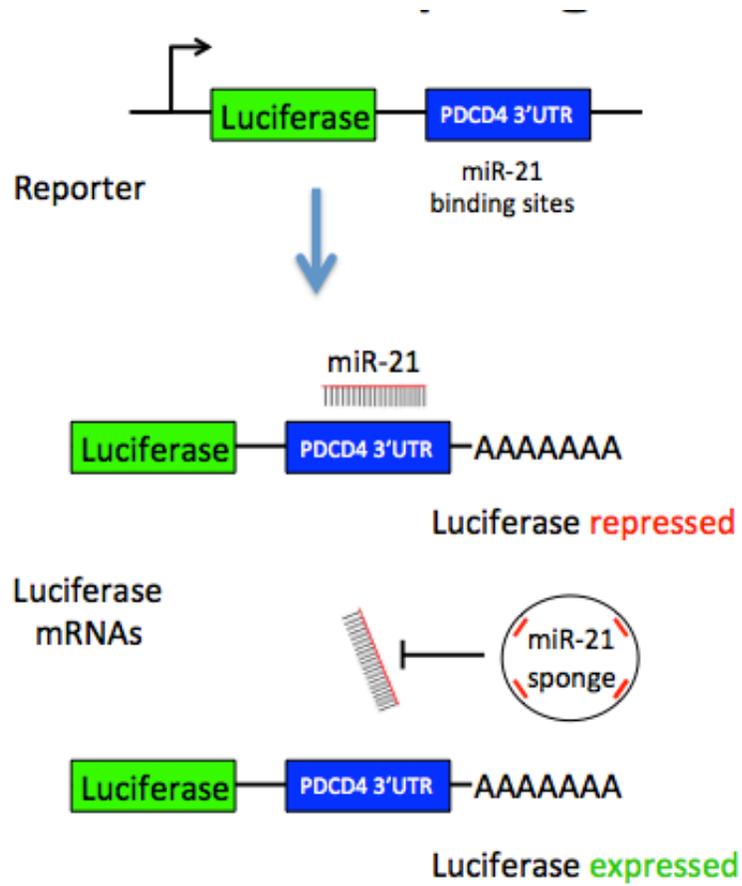
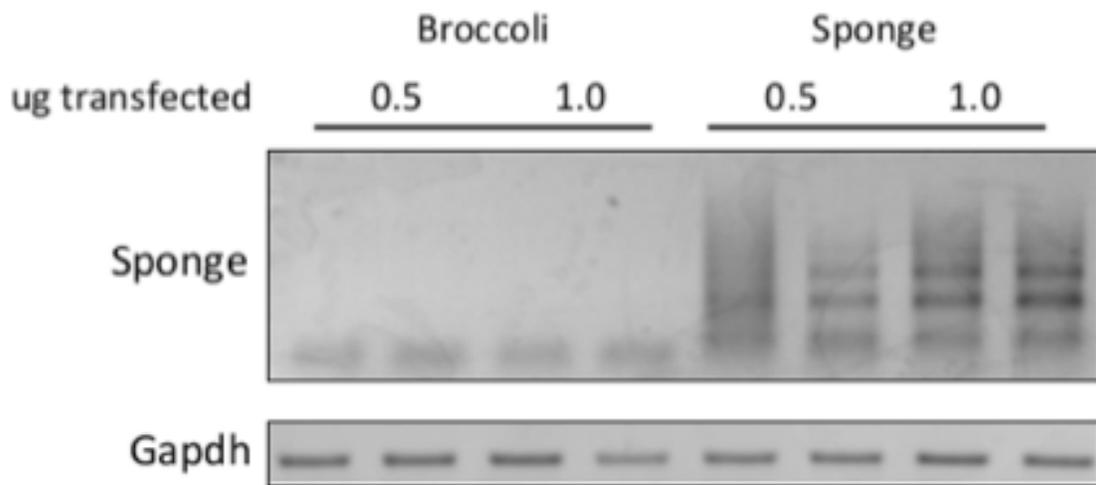
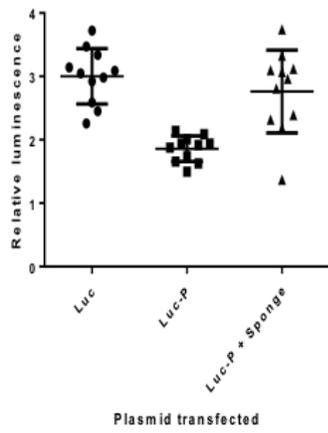


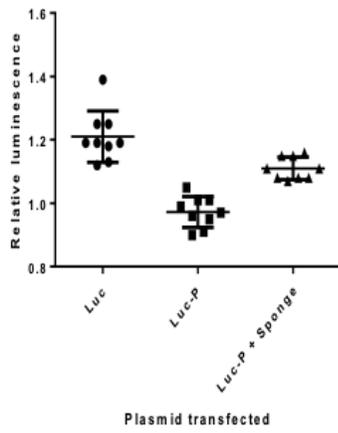
Figure Appendix 1.2: miR-21 is expressed after transient transfection in HeLa cells. RT-PCR confirms that the miR-21 sponge is expressed as a circle with concatameric amplification products in transiently transfected cells. TricBroccoli was used as a negative control, with GAPDH as a loading control.



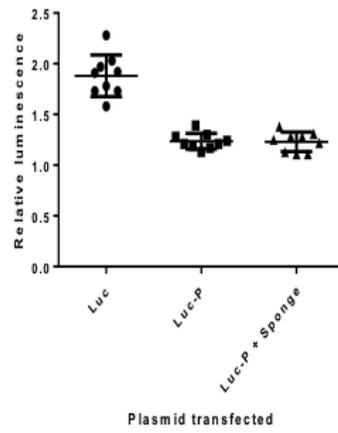
Appendix Figure 1.2: Luciferase assay for miR-21 sponge effectiveness. Untagged luciferase (Luc), tagged luciferase with PDCD4 3'-UTR and miR-21 binding site (P-luc), and miR-21 tricRNA sponge (Sponge) were transfected into HeLa cells. After three days, luminescence was measured.



Trial 1



Trial 2



Trial 3

APPENDIX 2: SMN SPLICING MODULATOR tricRNA EXPRESSION

Goals and results

I also wanted to determine whether we could use tricRNAs as an antisense effector of splicing. In the disease spinal muscular atrophy (SMA), exon 7 of the *SMN2* gene is skipped during splicing leading to an unstable, truncated protein (Cartegni and Krainer, 2002), which sensitizes motor neurons and causes them to degenerate and can lead to the onset of loss of motor function and death. Antisense oligonucleotides (ASOs) have been successfully used therapeutically, which bind the intronic splice silencer ISS-N1 in the intron directly downstream of exon 7 (**Figure Appendix 2.1 A**). This binding abrogates exon-skipping (**Figure Appendix 2.1 B**) and allows translation of stable and functional SMN (Skordis *et al.*, 2003; Geib and Hertel, 2009; Zhou *et al.*, 2013). ASOs are typically chemically modified to make them more stable in cells and thereby more effective in treatment. Since circRNAs are relatively stable molecules, we speculate that an antisense circRNA could be a good therapeutic agent in the context of SMA. To determine if I could express a tricRNA with the splice-modifying potential, I designed a tricRNA with 25 bases of complementarity with ISS-N1, similar to the ASO PMO25 (**Figure Appendix 2.2 A**), which is an effective SMN therapeutic oligonucleotide. I expressed this construct in HEK293T cells and verified circular RNA expression by RT-PCR, observing a cDNA concatamer PCR product characteristic of RT-PCR of circRNAs. When expressed in

HEK293T cells via transient transfection of our tricRNA construct did not appear to shift splicing toward full-length SMN. Boosting expression of the tricRNA and/or re-design of the construct may be necessary to affect SMN splicing.

Methods

Cloning SMN splice modifier tricRNA expression vector

The following oligos were annealed and ligated into NotI/SacII linearized pTricY (pGEM or pAV backbone): Top:

5'-ggccgagagcactgtacagtaagattcactttcataatgctggcccgactgcatctcgtggccgc-3';

Bottom:

5'-ggccacgagatgcagtcgggcccagcattatgaaagtgaatcttactgtacagtgctctcgc-3' where the sequence complimentary to SMN intronic splice silencer is underlined.

RT-PCR

The following divergent primers were used to amplify and detect the SMN splice modifier triRNA: Fwd: 5'- tggcccgactgcatctcg-3'; Rev: 5'- actgtacagtgctctcgcgg-3'.

The following primer pair was used to amplify SMN and SMN2 revealing splicing ratios:

Ex5 Fwd (5'-CTA TCA TGC TGG CTG CCT-3') and Ex8 Rev (5'-CTA CAA CAC CCT TCT CAC AG-3').

Cell culture and RNA preparation

All cells were grown at 37C with 5% CO2 in DMEM (Gibco) supplemented with 10% fetal bovine serum and 1% penicillin and streptomycin. One million HEK293T were plated

in T-25 flasks and incubated at 37C. After 24 hours, cells were transfected with equimolar amounts of either circBroc or SMN splice modifier tricRNA expression vector using Fugene HD transfection reagent (Promega cat#E2311). After 72 hours, total RNA was isolated from cells using Trizol reagent (Thermo Fisher Scientific cat#15596026).

Figures

Figure Appendix 2.1: Use of antisense oligonucleotides to alter SMN splicing in SMA model mice. (A) Diagrammatic representation of the positions of AOs targeting the ISS-N1 region. The sequence of exon 7 is shown in upper case and that of intron 7 in lower case. The sequence of ISS-N1 and the annealing sites for the 18-mer ASO-10-27, 20-mer HSMNEx7D, and 25-mer PMO25 are highlighted. (B) A representative reverse transcription-PCR image shows the effect of PMOs on SMN2 exon 7 inclusion in cultured SMA fibroblasts. Cells were treated at 100 and 500 nM in each PMO group. This figure was adapted (Zhou *et al.*, 2013).

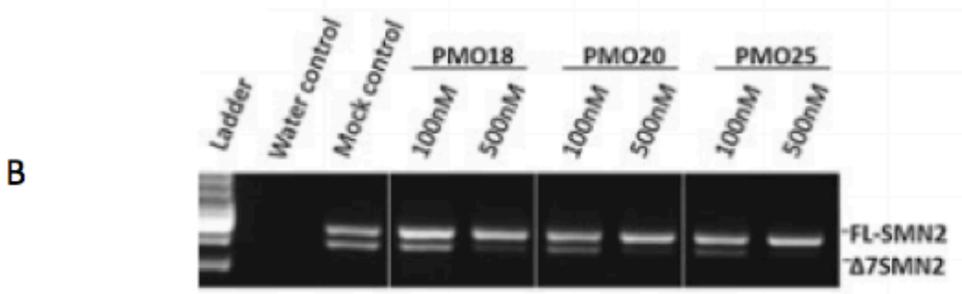
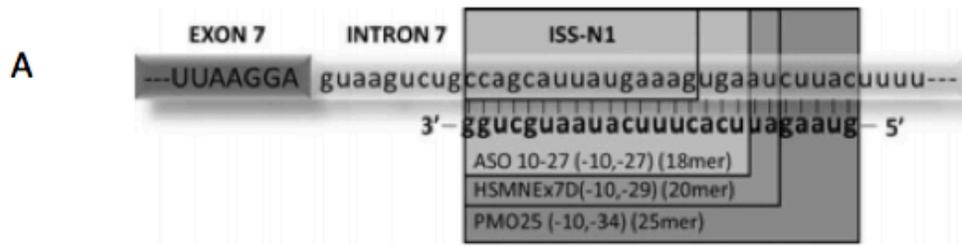


Figure Appendix 2.2: SMN splicing modifier tricRNA design. This design includes a 12 base pair stem loop and 25 nucleotides of complementarity to the SMN intron 7 intronic splice silencer (ISS).

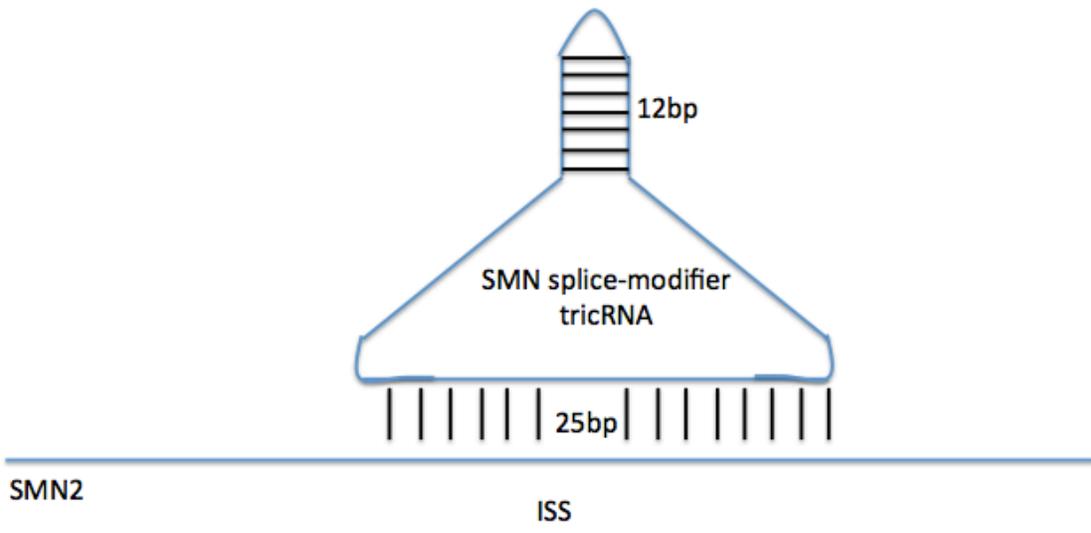
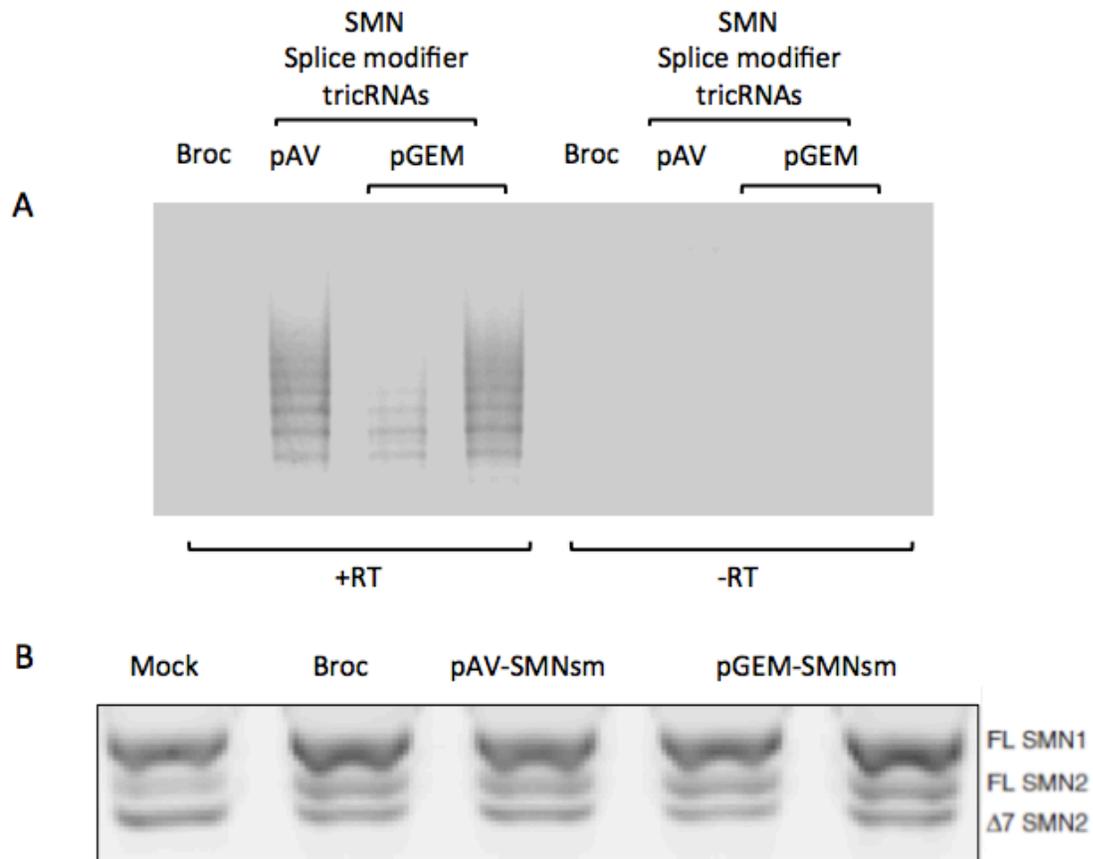


Figure Appendix 2.3: RT-PCR to detect SMN splice modifier tricRNAs and exon 7 inclusion in spliced SMN transcripts in cells transfected with SMN splice modifier tricRNA construct. (A) SMN splice modifier in either pAV or pGEM vector backbone was transfected into HEK293T cells and tricRNAs were detected by RT-PCR, using tricBroccoli as a negative control. (B) SMN splice modifier in either pAV or pGEM vector backbone was transfected into HEK293T cells and RT-PCR was used to determine ratios of products of SMN splicing, with mock transfected and tricBroccoli transfected cellular RNA input as negative controls.



REFERENCES

- Ashwal-Fluss, R., Meyer, M., Pamudurti, N. R., Ivanov, A., Bartok, O., Hanan, M., Evantal, N., Memczak, S., Rajewsky, N. and Kadener, S. (2014). CircRNA Biogenesis competes with Pre-mRNA splicing', *Molecular Cell*, 56(1), 55-66.
- Barrett, S. P. and Salzman, J. (2016). Circular RNAs: analysis, expression and potential functions. *Development*, 143(11), 1838–1847.
- Burd, C. E., Jeck, W. R., Liu, Y., Sanoff, H. K., Wang, Z. and Sharpless, N. E. (2010). Expression of linear and novel circular forms of an INK4/ARF-associated non-coding RNA correlates with atherosclerosis risk. *PLoS Genetics*, 6(12), 1–15.
- Burggraf, S., Larsen, N., Woese, C. R. and Stetter, K. O. (1993). An intron within the 16S ribosomal RNA gene of the archaeon *Pyrobaculum aerophilum*. *Proceedings of the National Academy of Sciences of the United States of America*, 90(March), 2547–2550.
- Calvin, K. and Li, H. (2008). RNA-splicing endonuclease structure and function. *Cellular and Molecular Life Sciences*, 65(7–8), 1176–1185.
- Cartegni, L. and Krainer, A. R. (2002). Disruption of an SF2/ASF-dependent exonic splicing enhancer in SMN2 causes spinal muscular atrophy in the absence of SMN, *Nature Genetics*, 30(4), 377–384.
- Chan, P. P. and Lowe, T. M. (2009) GtRNAdb: A database of transfer RNA genes detected in genomic sequence. *Nucleic Acids Research*, 37(SUPPL. 1), 93–97.
- Chen, C. and Sarnow, P. (1995). Initiation of protein synthesis by the eukaryotic translational apparatus on circular RNAs. *Science*, 268(5209), 415–417.
- Chen, L. (2016). The biogenesis and emerging roles of circular RNAs. *Nature Reviews. Molecular Cell Biology*, 17(4), 205–11.
- Coller, J. and Parker, R. (2004). Eukaryotic mRNA Decapping. *Annual Review of Biochemistry*, 73(1), 861–890.
- Conn, S. J., Pillman, K. A., Toubia, J., Conn, V. M., Salmanidis, M., Phillips, C. A., Roslan, S., Schreiber, A. W., Gregory, P. A. and Goodall, G. J. (2015). The RNA binding protein quaking regulates formation of circRNAs. *Cell*, 160(6), 1125–1134.
- Coqquerelle, Claude; Mascrez, Benedicte; Hetuin, Dominique; Bailleul, B. (1993). Mis-splicing yields circular RNA molecules. *FASEB J*, 7(1), 155–160.

Cortés-López, M., Gruner, M., Cooper, D. A., Gruner, H. N., Voda, A.-I., van der Linden, A. and Miura, P. (2017). Global accumulation of circRNAs during aging in *Caenorhabditis elegans*. *Doi.Org*, p. 175026. doi: 10.1101/175026.

Dalgaard, J. Z. and Garrett, R. A. (1992). Protein-coding introns from the 23S rRNA-encoding gene form stable circles in the hyperthermophilic archaeon *Pyrobaculum organotrophum*. *Gene*, 121(1), 103–110.

Danan, M., Schwartz, S., Edelheit, S. and Sorek, R. (2012). Transcriptome-wide discovery of circular RNAs in Archaea. *Nucleic Acids Research*, 40(7), 3131-3142.

Dietzl, G., Chen, D., Schnorrer, F., Su, K. C., Barinova, Y., Fellner, M., Gasser, B., Kinsey, K., Oettel, S., Scheiblaue, S., Couto, A., Marra, V., Keleman, K. and Dickson, B. J. (2007). A genome-wide transgenic RNAi library for conditional gene inactivation in *Drosophila*. *Nature*, 448(7150), 151–156.

Di Segni, G., Borghese, L., Sebastiani, S. and Tocchini-Valentini, G. P. (2005). A pre-tRNA carrying intron features typical of Archaea is spliced in yeast. *RNA*, 11(1), 70–76.

Ebbesen, Karoline K.; Hansen, Thomas B.; Kjems, J. (2017). Insights into circular RNA biology. *RNA Biology*, 14(8), 1035-1045.

Englert, M. and Beier, H. (2005). Plant tRNA ligases are multifunctional enzymes that have diverged in sequence and substrate specificity from RNA ligases of other phylogenetic origins. *Nucleic Acids Research*, 33(1), 388–399.

Englert, M., Sheppard, K., Aslanian, A., Yates, J. R. and Söll, D. (2011). Archaeal 3'-phosphate RNA splicing ligase characterization identifies the missing component in tRNA maturation. *Proceedings of the National Academy of Sciences of the United States of America*, 108(4), 1290–1295.

Filonov, G. S., Kam, C. W., Song, W. and Jaffrey, S. R. (2015). In-Gel Imaging of RNA Processing Using Broccoli Reveals Optimal Aptamer Expression Strategies. *Chemistry & Biology*, 22(5), 649–660.

Filonov, G. S., Moon, J. D., Svensen, N. and Jaffrey, S. R. (2014). Broccoli: Rapid selection of an RNA mimic of green fluorescent protein by fluorescence-based selection and directed evolution. *Journal of the American Chemical Society*, 136(46), 16299–16308.

Fischer, J. W. and Leung, A. K. L. (2017). CircRNAs: a regulator of cellular stress. *Critical Reviews in Biochemistry and Molecular Biology*, 9238(February), 1–17.

Gaëta, B. A., Sharp, S. J. and Stewart, T. S. (1990). Saturation mutagenesis of the *Drosophila* tRNA^{Arg} gene B-box intragenic promoter element: Requirements for transcription activation and stable complex formation. *Nucleic Acids Research*, 18(6), 1541–1548.

Geib, T. and Hertel, K. J. (2009). Restoration of full-length SMN promoted by adenoviral vectors expressing RNA antisense oligonucleotides embedded in U7 snRNAs. *PLoS ONE*, 4(12).

Graur, D. (2017). An upper limit on the functional fraction of the human genome. *Genome Biology and Evolution*, 9(7), 1880–1885.

Greer, C. L., Soll, D. and Willis, I. (1987). Substrate recognition and identification of splice sites by the tRNA-splicing endonuclease and ligase from *Saccharomyces cerevisiae*. *Molecular Cell Biology*, 7(1), 76–84.

Gruner, H., Cortés-López, M., Cooper, D. A., Bauer, M. and Miura, P. (2016). CircRNA accumulation in the aging mouse brain. *Scientific Reports*, 6(November), 1–14.

Hangauer, M. J., Vaughn, I. W. and McManus, M. T. (2013). Pervasive Transcription of the Human Genome Produces Thousands of Previously Unidentified Long Intergenic Noncoding RNAs. *PLoS Genetics*, 9(6).

Hansen, T. B., Jensen, T. I., Clausen, B. H., Bramsen, J. B., Finsen, B., Damgaard, C. K. and Kjems, J. (2013). Natural RNA circles function as efficient microRNA sponges. *Nature*, 495, 384–388.

Ivanov, A., Memczak, S., Wyler, E., Torti, F., Porath, H. T., Orejuela, M. R., Piechotta, M., Levanon, E. Y., Landthaler, M., Dieterich, C. and Rajewsky, N. (2015). Analysis of intron sequences reveals hallmarks of circular RNA biogenesis in animals. *Cell Reports*, 10(2), 170–177.

Jeck, W. R. and Sharpless, N. E. (2014). Detecting and characterizing circular RNAs. *Nature Biotechnology*, 32(5), 453–61.

Jeck, W. R., Sorrentino, J. A., Wang, K., Slevin, M. K., Burd, C. E., Liu, J., Marzluff, W. F. and Sharpless, N. E. (2013). Circular RNAs are abundant, conserved, and associated with ALU repeats. *RNA*, 19(2), 141–57.

Johnson, P. J. (2002). Spliceosomal introns in a deep-branching eukaryote: The splice of life. *Proceedings of the National Academy of Sciences of the United States of America*, 99(6), 3359–3361.

Jurkin, J., Henkel, T., Nielsen, A. F., Minnich, M., Popow, J., Kaufmann, T., Heindl, K., Hoffmann, T., Busslinger, M. and Martinez, J. (2014). The mammalian tRNA ligase complex mediates splicing of XBP1 mRNA and controls antibody secretion in plasma cells. *The EMBO journal*, 33(24), 2922–36.

Karaca, E., Weitzer, S., Pehlivan, D., Shiraishi, H., Gogakos, T., Hanada, T., Jhangiani, S. N., Wiszniewski, W., Withers, M., Campbell, I. M., Erdin, S., Isikay, S., Franco, L. M., Gonzaga-Jauregui, C., Gambin, T., Gelowani, V., Hunter, J. V., Yesil, G., Koparir, E., Yilmaz, S., Brown, M., Briskin, D., Hafner, M., Morozov, P., Farazi, T. a, Bernreuther, C., Glatzel, M., Trattinig, S., Friske, J., Kronnerwetter, C., Bainbridge, M. N., Gezdirici, A., Seven, M., Muzny, D. M., Boerwinkle, E., Ozen, M., Clausen, T., Tuschl, T., Yuksel, A., Hess, A., Gibbs, R., Martinez, J., Penninger, J. M. and Lupski, J. R. (2014). Human CLP1 mutations alter tRNA biogenesis, affecting both peripheral and central nervous system function.. *Cell*, 157(3), pp. 636–650.

Kjems, J. and Garrett, R. A. (1988). Novel splicing mechanism for the ribosomal RNA intron in the archaeobacterium *desulfurococcus mobilis*. *Cell*, 54(5), 693–703.

Knapp, J. M., Chung, P. and Simpson, J. H. (2015). Generating customized transgene landing sites and multi-transgene arrays in *Drosophila* using phiC31 integrase. *Genetics*, 199(4), 919–934.

Kosmaczewski, S. G., Edwards, T. J., Han, S. M., Eckwahl, M. J., Meyer, B. I., Peach, S., Hesselberth, J. R., Wolin, S. L. and Hammarlund, M. (2014). The RtcB RNA ligase is an essential component of the metazoan unfolded protein response. *EMBO Reports*, 15(12), 1278–1285.

Kramer, M. C., Liang, D., Tatomer, D. C., Gold, B., March, Z. M., Cherry, S. and Wilusz, J. E. (2015). Combinatorial control of *Drosophila* circular RNA expression by intronic repeats, hnRNPs, and SR proteins. *Genes and Development*, 29, 2168-2182.

Lasda, E. and Parker, R. (2014). Circular RNAs: diversity of form and function. *RNA*, 20(12), 1829–42.

Lebbink, R. J., Lowe, M., Chan, T., Khine, H., Wang, X. and McManus, M. T. (2011). Polymerase II promoter strength determines efficacy of microRNA adapted shRNAs. *PLoS ONE*, 6(10).

Legnini, I., Di Timoteo, G., Rossi, F., Morlando, M., Briganti, F., Sthandier, O., Fatica, A., Santini, T., Andronache, A., Wade, M., Laneve, P., Rajewsky, N. and Bozzoni, I. (2017). Circ-ZNF609 Is a Circular RNA that Can Be Translated and Functions in Myogenesis. *Molecular Cell*, 66(1), p. 22–37.

Levinger, L., Oestreich, I., Florentz, C. and Mörl, M. (2004). A Pathogenesis-associated Mutation in Human Mitochondrial tRNA^{Leu}(UUR) Leads to Reduced 3'-End Processing and CCA Addition. *Journal of Molecular Biology*, 337(3), 535–544.

Liang, D. and Wilusz, J. E. (2014). Short intronic repeat sequences facilitate circular RNA production. *Genes and Development*, 28, 2233-2247.

- Linder, B., Grozhik, A. V., Olarerin-George, A. O., Meydan, C., Mason, C. E. and Jaffrey, S. R. (2015). Single-nucleotide-resolution mapping of m6A and m6Am throughout the transcriptome. *Nature Methods*, 12(8), 767–772.
- Lu, Z., Filonov, G. S., Noto, J. J., Schmidt, C. A., Hatkevich, T. L., Wen, Y., Jaffrey, S. R. and Matera, A. G. (2015). Metazoan tRNA introns generate stable circular RNAs in vivo. *RNA*, 21(9), 1554–1565.
- Lu, Z. and Matera, A. G. (2014). Vicinal: A method for the determination of ncRNA ends using chimeric reads from RNA-seq experiments. *Nucleic Acids Research*, 42(9).
- Mattick, J. S. (2004). RNA regulation: A new genetics?. *Nature Reviews Genetics*, 5(4), 316–323.
- Memczak, S., Jens, M., Elefsinioti, A., Torti, F., Krueger, J., Rybak, A., Maier, L., Mackowiak, S. D., Gregersen, L. H., Munschauer, M., Loewer, A., Ziebold, U., Landthaler, M., Kocks, C., Le Noble, F. and Rajewsky, N. (2013). Circular RNAs are a large class of animal RNAs with regulatory potency. *Nature*, 495, 333–338.
- Meyer, K. D., Patil, D. P., Zhou, J., Zinoviev, A., Skabkin, M. A., Elemento, O., Pestova, T. V., Qian, S.-B. and Jaffrey, S. R. (2015). 5' UTR m6A Promotes Cap-Independent Translation. *Cell*, 163(4), 999–1010.
- Morris, K. V. and Mattick, J. S. (2014). The rise of regulatory RNA. *Nature Reviews Genetics*, 15(6), 423–437.
- Noto, J. J., Schmidt, C. A. and Matera, A. G. (2017). Engineering and expressing circular RNAs via tRNA splicing. *RNA Biology*, 14(8), 978–984.
- Pamudurti, N. R., Bartok, O., Jens, M., Ashwal-Fluss, R., Stottmeister, C., Ruhe, L., Hanan, M., Wyler, E., Perez-Hernandez, D., Ramberger, E., Shenzis, S., Samson, M., Dittmar, G., Landthaler, M., Chekulaeva, M., Rajewsky, N. and Kadener, S. (2017). Translation of CircRNAs. *Molecular Cell*, 66(1), p. 9–21.
- Perriman, R. and Ares, M. (1998). Circular mRNA can direct translation of extremely long repeating-sequence proteins in vivo. *RNA*, 4(9), 1047–1054.
- Peschek, J., Acosta-Alvear, D., Mendez, A. S. and Walter, P. (2015). A conformational RNA zipper promotes intron ejection during non-conventional XBP1 mRNA splicing. *EMBO Reports*, 16, 1688–1698.
- Petkovic, S. and Müller, S. (2015). RNA circularization strategies in vivo and in vitro. *Nucleic Acids Research*, 43(4), 2454–2465.
- Phizicky, E. M., Schwartz, R. C. and Abelson, J. (1986). *Saccharomyces cerevisiae* tRNA ligase. Purification of the protein and isolation of the structural gene. *Journal of Biological Chemistry*, 261(6), 2978–2986.

- Popow, J., Englert, M., Weitzer, S., Schleiffer, A., Mierzwa, B., Mechtler, K., Trowitzsch, S., Will, C. L., Lührmann, R., Söll, D. and Martinez, J. (2011). HSPC117 is the essential subunit of a human tRNA splicing ligase complex. *Science*, 331(6018), 760–764.
- Qin, J. Y., Zhang, L., Clift, K. L., Hular, I., Xiang, A. P., Ren, B. Z. and Lahn, B. T. (2010). Systematic comparison of constitutive promoters and the doxycycline-inducible promoter. *PLoS ONE*, 5(5), pp. 3–6.
- Qu, S., Zhong, Y., Shang, R., Zhang, X., Song, W., Kjems, J. and Li, H. (2016). The emerging landscape of circular RNA in life processes. *RNA Biology*, 14(8), 992–999.
- Reyes, V. M. and Abelson, J. (1988). Substrate recognition and splice site determination in yeast tRNA splicing. *Cell*, 55(4), 719–730.
- Rybak-Wolf, A., Stottmeister, C., Glažar, P., Jens, M., Pino, N., Hanan, M., Behm, M., Bartok, O., Ashwal-Fluss, R., Herzog, M., Schreyer, L., Papavasileiou, P., Ivanov, A., Hman, M., Refojo, D., Kadener, S. and Rajewsky, N. (2014). Circular RNAs in the Mammalian Brain Are Highly Abundant, Conserved, and Dynamically Expressed. *Molecular Cell*, 58(5), 870–885.
- Sakharkar, M. K., Chow, V. T. K. and Kanguene, P. (2004). Distributions of Exons and Introns in the Human Genome. *In Silico Biology*, 4(4), 387–393.
- Salgia, S. R., Singh, S. K., Gurha, P. and Gupta, R. (2003). Two reactions of *Haloferax volcanii* RNA splicing enzymes: joining of exons and circularization of introns. *RNA*, 9(3), 319–330.
- Salmena, L., Poliseno, L., Tay, Y., Kats, L. and Pandolfi, P. P. (2011). A ceRNA hypothesis: The rosetta stone of a hidden RNA language?. *Cell*, 146(3), 353–358.
- Salzman, J. (2016). Circular RNA Expression: Its Potential Regulation and Function. *Trends in Genetics*, 32(5), 309–316.
- Salzman, J., Chen, R. E., Olsen, M. N., Wang, P. L. and Brown, P. O. (2013). Cell-Type Specific Features of Circular RNA Expression. *PLoS Genetics*, 9(9).
- Schmidt, C. A., Noto, J. J., Filonov, G. S. and Matera, A. G. (2016). A Method for Expressing and Imaging Abundant, Stable, Circular RNAs in Vivo Using tRNA Splicing. *Methods in Enzymology*, 572, 215–236.
- Seitz, H. (2009). Redefining MicroRNA Targets. *Current Biology*, 19(10), 870–873.
- Sharp, S., DeFranco, D., Dingermann, T., Farrell, P. and Söll, D. (1981). Internal control regions for transcription of eukaryotic tRNA genes. *Proceedings of the National Academy of Sciences of the United States of America*, 78(11), 6657–61.

Sidrauski, C., Cox, J. S. and Walter, P. (1996). tRNA ligase is required for regulated mRNA splicing in the unfolded protein response. *Cell*, 87(3), 405–413.

Singh, S. K., Gurha, P., Tran, E. J., Maxwell, E. S. and Gupta, R. (2004). Sequential 2'-O-methylation of archaeal pre-tRNA Trp nucleotides is guided by the intron-encoded but trans-acting box C/D ribonucleoprotein of pre-tRNA. *Journal of Biological Chemistry*, 279(46), 47661–47671.

Skordis, L. A., Dunckley, M. G., Yue, B., Eperon, I. C. and Muntoni, F. (2003). Bifunctional antisense oligonucleotides provide a trans-acting splicing enhancer that stimulates SMN2 gene expression in patient fibroblasts. *Proceedings of the National Academy of Sciences of the United States of America*, 100(7), 4114–9.

Starke, S., Jost, I., Rossbach, O., Schneider, T., Schreiner, S., Hung, L. H. and Bindereif, A. (2015). Exon circularization requires canonical splice signals. *Cell Reports*, 10(1), 103–111.

Tatomer, D. C. and Wilusz, J. E. (2017). An Uncharted Journey for Ribosomes: Circumnavigating Circular RNAs to Produce Proteins. *Molecular Cell*, 66(1), 1–2.

Tay, Y., Kats, L., Salmena, L., Weiss, D., Tan, S. M., Ala, U., Karreth, F., Poliseno, L., Provero, P., Di Certo, F., Lieberman, J., Rigoutsos, I. and Pandolfi, P. P. (2011). Coding-independent regulation of the tumor suppressor PTEN by competing endogenous mRNAs. *Cell*, 147(2), 344–357.

Tocchini-Valentini, G. D., Fruscoloni, P. and Tocchini-Valentini, G. P. (2005). Coevolution of tRNA intron motifs and tRNA endonuclease architecture in Archaea. *Proceedings of the National Academy of Sciences of the United States of America*, 102(43), 15418–22.

Trotta, C. R., Miao, F., Arn, E. A., Stevens, S. W., Ho, C. K., Rauhut, R. and Abelson, J. N. (1997). The yeast tRNA splicing endonuclease: a tetrameric enzyme with two active site subunits homologous to the archaeal tRNA endonucleases. *Cell*, 89(6), 849–58.

Wang, L. K., Schwer, B., Englert, M., Beier, H. and Shuman, S. (2006). Structure-function analysis of the kinase-CPD domain of yeast tRNA ligase (Trl1) and requirements for complementation of tRNA splicing by a plant Trl1 homolog. *Nucleic Acids Research*, 34(2), 517–527.

Wang, Y. and Wang, Z. (2015). Efficient backsplicing produces translatable circular mRNAs. *RNA*, 21(2), 172–179.

Westholm, J. O., Miura, P., Olson, S., Shenker, S., Joseph, B., Sanfilippo, P., Celniker, S. E., Graveley, B. R. and Lai, E. C. (2014). Genome-wide Analysis of Drosophila Circular RNAs Reveals Their Structural and Sequence Properties and Age-Dependent Neural Accumulation. *Cell Reports*, 9(5), 1966–1980.

Wilson, F. H. (2004). A Cluster of Metabolic Defects Caused by Mutation in a Mitochondrial tRNA. *Science*, 306(5699), 1190–1194.

Woodson, S. A. (1998). Ironing out the kinks: Splicing and translation in bacteria. *Genes and Development*, 12(9), 1243–1247.

Xue, S., Calvin, K. and Li, H. (2006). RNA recognition and cleavage by a splicing endonuclease. *Science*, 312(5775), 906–10.

Yang, Y., Fan, X., Mao, M., Song, X., Wu, P., Zhang, Y., Jin, Y., Yang, Y., Chen, L. L., Wang, Y., Wong, C. C. L., Xiao, X. and Wang, Z. (2017). Extensive translation of circular RNAs driven by N⁶-methyladenosine. *Cell Research*, 27(5), 626–641.

Yao, Q., Xu, H., Zhang, Q. Q., Zhou, H. and Qu, L. H. (2009). MicroRNA-21 promotes cell proliferation and down-regulates the expression of programmed cell death 4 (PDCD4) in HeLa cervical carcinoma cells. *Biochemical and Biophysical Research Communications*, 388(3), 539–542.

Yoshihisa, T. (2014). Handling tRNA introns, archaeal way and eukaryotic way. *Frontiers in Genetics*, 5(JUL), 1–16.

Zhou, H., Janghra, N., Mitrpant, C., Dickinson, R. L., Anthony, K., Price, L., Eperon, I. C., Wilton, S. D., Morgan, J. and Muntoni, F. (2013). A novel morpholino oligomer targeting ISS-N1 improves rescue of severe spinal muscular atrophy transgenic mice. *Human gene therapy*, 24(3), 331–342.