THE NECESSITY OF TRNA-VAL-CAC-1-1 DERIVED TRNA FRAGMENTS (TRFS) IN *DROSOPHILA*

By

NICHOLAS GATTONE

A thesis submitted to the

Graduate School- Camden

Rutgers, The State University of New Jersey

In partial fulfillment of the requirements for

For the degree of Master of Science

Graduate Program in Biology

Written under the direction of

Dr. Nir Yakoby

And approved by

_____

Nir Yakoby

_____

Andrey Grigoriev

_____

Eric Klein

Camden, NJ

January 2021

ABSTRACT OF THE THESIS

The Necessity of tRNA-Val-CAC-1-1 Derived tRNA Fragments (tRFs) in *Drosophila*

By: NICHOLAS GATTONE

Thesis Director
Dr. Nir Yakoby

Small non-coding RNAs, such as micro-RNAs (miRNAs) or small interfering RNAs (siRNAs) are critical regulators of gene expression during tissue development. tRNA-derived fragments are a newly discovered class of small non-coding RNA that participate in different biological processes, some of which are attributed to post-transcriptional regulation similar to miRNAs. Using a computational pipeline designed to find tRF targets by mirroring a miRNA-like seed-based targeting mechanism, we identified that the tRNA-Val-CAC-1-1 contains a seed sequence that may target epidermal growth factor receptor (EGFR) inhibitor *Sprouty*. Using CRISPR/Cas9 genome engineering, this tRNA was deleted. The transgenic fly remained heterozygous, suggesting a necessary role for fly survival. The RNA-seq data from the heterozygous fly did demonstrate small transcriptional changes for target genes in both the ovary and embryo. In addition, a global change in the distribution of small RNAs was found, suggesting a possible stress reaction on the small RNA level from the targeted deletion. Further CRISPR experiments are underway to determine the reason for the unviability of the tRNA-Val-CAC-1-1 homozygous mutant.

ACKNOWLEDGMENTS

I would first like to thank all those who provided reagents and flies for this work: to the Siomi Lab at the University of Tokyo for providing the anti-*Drosophila* AGO1 and AGO2 antibodies, J. Kadrmas for the *ics* and *PINCH* antibodies. I was also partially supported by the Busch Biomedical Grant.

I would also like to thank Nicole Revaitis and Spyros Karaiskos, whose hard work laid the groundwork for this project I've been lucky enough to inherit, as well as Lingyu Guan whose computational analysis of our RNA-seq data has proven invaluable. I'd like to thank all members of the Yakoby Lab during my time here for generous feedback and for fruitful discussions. A special thank you, however, should go to Cody Stevens and Chris Sottolano, who I've become fortunate enough to be able to call friends first and lab-mates second.

Last but not least, I'd like to thank Nir Yakoby, whose knowledge and mentorship have been essential to my growth over the last few years. Through his encouragement, friendship, and his contagious drive, Nir makes his lab feel like a family where everyone's voice matters.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

Introduction

Type of RNAs:

The role of RNA in cells is often primarily viewed in the context of Francis Crick's central dogma of molecular biology- DNA is transcribed to RNA, and RNA is translated to protein (Crick, 1958). While this has long described an essential framework to understand the relationship between the genetic code of DNA and the synthesis of biopolymers, it only accurately describes the behavior of messenger RNA (mRNA), which constitutes approximately 4-5% of total cellular RNA (Lodish et al, 2000). The vast majority of RNA is non-coding RNA (ncRNA) and is accordingly not translated into protein. Roughly 80% of cellular RNA, by far the largest portion, is ribosomal RNA (rRNA), around 15% is transfer RNA (tRNA), and less than 1%, falls under the umbrella of small RNA (Warner et al, 1999.)

Small RNA, including species of RNA smaller than 200 nt, perform a wide variety of functions in organisms, including plant development (Chen, 2012), animal development (Stefani et al., 2008), and microbial immunity (Guo et al., 2019). Among these functions is the regulation of gene expression through RNA degradation or inhibition of translation (Hutwagner et al. 2008). Several species of small RNA have been extensively studied, including microRNA (miRNA), small interfering RNA (siRNA), and Piwi-interacting RNA (piRNA (Shin et al., 2010). Recently, high-throughput small RNA sequencing methods have detected a novel class of small RNAs which align preferentially to certain mature tRNAs found throughout the genome (Tuck et al., 2011). These small RNAs, called tRNA-derived fragments (tRFs), are not random

byproducts of degradation- they are formed in a precise manner through tRNA cleavage and have drawn attention to the possibility tRNAs may act as a regulatory molecule (Li et al. 2008, Wei et al. 2012).

The canonical role of tRNA in cells is well understood. Their anticodon loop, located on the bottom of their 2-D cloverleaf-like structure, binds codon triplets on an mRNA strand; an amino acid loaded onto the 3' end of the tRNA forms a peptide bonds with other amino acids, creating a growing polypeptide chain which will eventually become a functional protein (**Figure 1**) (Schimmel 2018). Eukaryotic organisms, however, often have hundreds of tRNA-encoding genes in their genomes (Haeusler 2006). The requirement for 64 codon triplets suggests that the great redundancy of tRNAs may have roles other than the canonical protein synthesis.

tRNA fragments:

The formation of tRNA fragments (tRF) is a structured process that occurs in one of three distinct patterns of cleavage from the mature tRNA molecule (Figure 2); 5’ tRFs originate from the 5’ end of the tRNA, 3’ tRFs originate from the 3’ end, and precursor tRFs originate from the extension on the 3’ end that exists prior to the loading of the amino acid (Lee et al. 2009, Sobala et al. 2011). Several attempts have been made to understand the pathways of biogenesis, as well as potential functional roles of tRFs (Gebetsberger et al. 2012, Sobala et al. 2011).

While the precise mechanisms of tRNA cleavage are not fully understood, there is still evidence regarding the roles of specific nucleases like Dicer and angiogenin in their formation (Shen et al., 2018). This is of particular relevance to the potential developmental roles of tRFs because of Dicer's role in miRNA biogenesis and the RISC pathway (Li et al., 2018). Certain species of tRFs in each of the three cleavage patterns have been demonstrated to be formed in a Dicer-dependent manner in humans (Cole et al., 2009), mice (Babiarz et al., 2008), and in *Drosophila* (Kawamura et al., 2008). While it is atypical for Dicer to recognize and cut molecules with a tRNA's secondary structure, there is evidence some tRNA molecules can form a hairpin-like structure more suited for Dicer processing (Babiarz et al., 2008). Other tRFs are generated in a Dicer-independent manner. RNase Z cleavage has been shown to produce 3' tRFs from pre-tRNAs (Lee et al., 2009). Additionally, angiogenin, a RNase A family nuclease, was previously discovered to produce tRFs in cellular stress conditions, but has also been shown to cleave the T-loop of tRNAs *in vitro* (Li Z et al., 2012).

The exact targeting mechanisms of tRFs remain unclear, but current evidence suggests tRFs act as regulatory molecules in a similar fashion to miRNAs (Li et al. 2008, Gebetsberger et al. 2012) and sometimes may compete with miRNAs (Tuck et al. 2011). The targeting mechanisms of miRNAs are well understood. Specifically, plant miRNAs have sequences fully complementary to their mRNA targets, whereas animal miRNAs have a highly conserved 7 nucleotide "seed sequence" that binds its targets (Grimson et al. 2007, Figure 3). Multiple models suggest a similar seed-based targeting mechanism for tRFs (Wang et al. 2013). tRFs have also been shown to associate with the Argonaute

family of proteins and enter the RISC pathway, similar to the miRNA mode of action (Kumar et al., 2014).

While there are many biological roles in which tRFs seem to be involved, such as viral protection, carcinogenesis, and neurodegeneration (Venkatesh et al. 2016, Goodarzi et al. 2015, Maute et al., 2013), their potential role as post-transcriptional modifiers in

tissue development and cell fate determination remains mostly unexplored. These can be investigated in the fruit fly *Drosophila melanogaster*, which is a well-established model system to investigate changes in tissue morphogenesis (Berg, 2005). The plethora of genetic, molecular, and genomic tools to manipulate the Drosophila genome (Hales et al., 2015), makes it an ideal system to investigate the role of tRFs in tissue development. We can also monitor changes in *Drosophila* RNA profiling, an approach which has been successfully used by the Yakoby lab in the past (Revaitis et al. 2017).

The experimental system: *Drosophila* oogenesis:

Drosophila oogenesis is a well-established model system for the study of developmental processes like morphogenesis, cell fate determination, and tissue patterning (Berg, 2005). With respect to our small-RNA focused experiments, this system has also been used to study gene regulation by small RNAs targeting of genes through the inhibition of small RNA production (Nakahara et al., 2005). Oogenesis, the process of egg development, occurs in an assembly line-like fashion in the ovary and is comprised of 14 morphologically distinct stages (Spradling, 1993). The developing egg chamber has three main parts; the oocyte, nurse cells, and follicle cells. The oocyte becomes the developing embryo after fertilization, nurse cells provide essential RNAs and proteins to the oocyte, and the follicle cells form a two-dimensional layer which is patterned and stretches to envelop the oocyte and form a three-dimensional eggshell structure (Berg, 2005; Spralding 1993). Follicle cell patterning and the formation of dorsal eggshell structures is reliant on the epidermal growth factor receptor (EGFR) pathway (Neuman-Silberberg and Schupbach, 1993; Neuman-Silberbeg and Schupbach, 1994). During stages 7-10A of oogenesis, TGF-α like ligand Gurken (GRK) is the major ligand of EGFR; GRK activates EGFR and triggers signaling through the Ras/Raf/MEK/ERK

phosphorylation cascade (Ray and Schupbach, 1996; Wasserman et al., 1998). In the absence of this signaling, the dorsal fate of the eggshell is not established and thus dorsal structures are not present (Niepielko and Yakoby, 2014).

A computational pipeline developed by the Grigoriev Lab predicted a *Drosophila* tRF derived from tRNA-Val-CAC-1-1 to target *sprouty (sty),* a potent regulator of the EGFR pathway (Hacohen et al., 1998; Casci et al., 1999). Using CRISPR/Cas9, tRNA-Val-CAC-1-1 was removed from the genome. Interestingly, this mutation could not become homozygous, despite the presence of six additional tRNA-Val-CAC genes in the *Drosophila* genome. In an RNA-seq based transcriptional profiling experiment, genes ACXD, IP3K1, and *sta* show significant differential expression in both the ovary and embryo of the heterozygous tRNA-Val-CAC-1-1 mutant. This is in line with the potential activity of tRFs as regulators of gene expression in development. Additionally, small RNA-seq demonstrated global changes in tRF populations in the heterozygous mutant, though the exact reason for this finding is unclear.

Materials and Methods

Flies: The following fly stocks were used in this study: CRE recombinase (BDSC #34516), wild-type strain (OreR; BDSC, #25211), y[1]w[*]; if / SM6a,Cy; D/TM3,sb (a generous gift from Miki Fujioka, Thomas Jefferson University, Philadelphia, PA, USA), and y,sc,v; nos-Cas9; +/+ (Kondo et al., 2013) used for CRISPR injections by Rainbow Transgenics . Flies were grown on corn meal agar and all crosses were carried out at 23 C.

Prediction of tRFs: A computational pipeline was developed by the Grigoriev Lab to predict mRNA targets of tRFs using a 7-mer sliding window to account for all possible locations of a seed sequence. Exact matches of the reverse complements of these seeds were found in the introns, exons, 5' UTR, or 3' UTR of genes and compared to the expected frequency of "hits" that occur by chance to estimate the significance of the predicted targets. The Grigoriev Lab previously employed this technique with success in aging rat brains (Karaiskos et al., 2016.)

PCR: Wing clip PCR protocol was carried out according to Carvalho et al., 2009, with the exception of the amount of Buffer A used, which was raised to 20µl from 10µl. PCR was also carried out using Phusion HF polymerase (NEB) and reaction buffer.

CRISPR: Donor vector for CRISPR injection was generated with pHd-DsRed-attP (Addgene #51019). Two flanking 1kb homology arms were cloned into this vector. Co-injected gRNAs (Gratz et al., 2014) G1 (CACAGGAAACTAGAGTCATG) and G2 (AATGCCTTGCTTTGCTAGT) had calculated efficiencies of 7.27 and 8.01, respectively (https://www.flyrnai.org/evaluateCrispr/). Plasmids were injected by Rainbow Transgenics (CA). Emerging flies were crossed to a y[1]w[*]; if / SM6a,Cy; D/TM3,sb, and screened for the marker dsRed, driven by the 3xP3 promoter in the adult eye. Transgenic animals were PCR validated by amplifying OreR and mutant gDNA phenol-chloroform extracted from whole flies (50 ng/µl tube) with specific primers (Forward: ACTTCGGCTCTCCAAATCTAGT Reverse: CGGCGCTGAGAGTATTCTTC) at 62 C annealing temperature and 45 second extension time at 72 C.

Microscopy: Eggshell and eye morphologies were evaluated using a JEOL NeoScope Standard Electron Microscope (JCM-6000). To collect eggs, flies were kept on grape juice agar plates with yeast overnight. Eggshells were collected and mounted on double-sided carbon tape and imaged directly. Immunoassay images were taken on a Leica SP8 confocal microscope.

Immunoassay: Ovary dissection and fixations were carried out as previously described (Pacquelet and Rorth, 2005). Ovaries were dissected in cold Schneider's medium .The primary antibodies used were mouse anti-BR core (25E9.D7; 1:100, DSHB) mouse anti-FasIII (7G10, 1:100, DSHB), and rabbit anti-*icarus* (1:100, a gift J. Kadrmas at the Beckerle Lab). DAPI (1:10,000) was used to stain for nuclear DNA. The secondary antibodies used were 488 anti-mouse and 568 anti-rabbit (Invitrogen) (1:1000).

RNA extraction: For ovary RNA, 3-7 day-old flies were incubated in the presence of yeast for 24 hours prior to dissection directly into Trizol reagent. RNA isolation was carried out according to Direct-zol RNA Miniprep Kit (Zymo) protocol. For embryo RNA, embryos less than one hour old were collected from grape agar plates and stabilized in Trizol reagent until 100 embryos were collected. RNA isolation was then carried out as above. Control and mutant flies were genetically identical (y[1]w[*]; if / SM6a,Cy; D/TM3,sb) with the exception of the deleted tRNA-Val-CAC-1-1 on chromosome 3R of the mutant.

Small RNA size exclusion: 1μg total RNA dissolved in nuclease-free water was combined with equal volume RNA Gel Loading Dye (ThermoFisher #R0641) and run on 15% TBE-Urea Denaturing Polyacrylamide Gel for 90 minutes at 100V. Partitions of the gel located in between the bromophenol blue and xylene cyanol tracking dyes (marking

the 15-50 b range) were excised with a razor blade, then dissolved in 400µl PAGE elution buffer (20mM pH 7.5 Tris-HCl, 0.25M Sodium acetate, 1mM EDTA, 0.25% SDS) overnight at 4 C. Centrifuge the tube at 20,000xG for 10 minutes then draw off supernatant, which will contain RNA. Extract with 1:1 volumes and phenol: chloroform, then precipitate with 2 volumes ethanol. After 70% ethanol wash pellet was air-dried and resuspended in 21µl DEPC-treated water. 1µl of sample was used for analysis of purity on NanoDrop 2000.

qRT-PCR: cDNA libraries were generated from 1µg total RNA from all samples according to NEB Protoscript II First Strand cDNA Synthesis Kit (New England Biolabs E6560S) protocol with random hexamers from the kit. qPCR primers were designed to create a 80-120bp amplicon on an exon region of target gene. A 384 well plate was prepared using 5µl SYBR Green Reagent, 1.25µl cDNA, and primers at 500mM in a total well volume of 10µl. Reaction was run on an Applied Biosystems QuantStudio 6 for 35 cycles. CT values were analyzed and fold change calculations were performed on Microsoft Excel. All qPCR primers listed in supplementary Table 1

Results

Prediction of tRF targets

Based on the assumption tRFs use a seed sequence analogous to animal miRNAs to bind mRNA targets, the Grigoriev Lab developed an algorithm to find potential gene targets of a given tRF (Karaiskos et al., 2017). Specifically, using a 7-mer sliding window (Figure 4A), they searched for reverse complements in the 5’ UTR, 3’ UTR, and exon sequences in the *D. melanogaster* genome. This resulted in a list of 51 potential target

genes for tRNA-Val-CAC-1-1, which includes the EGFR pathway inhibitor *sty* (Figure 4B).

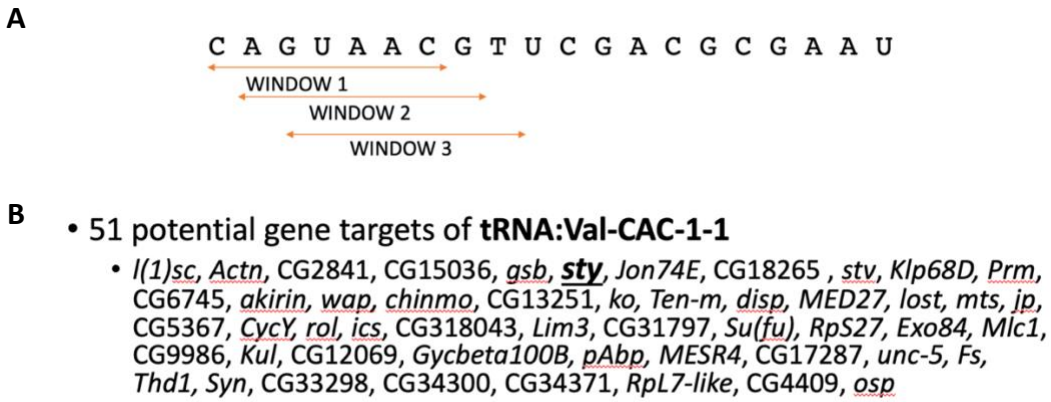


Figure 1: Computational prediction of tRF targets: A) Visualization of the 7-mer sliding window employed by the Grigoriev lab to obtain possible tRF targets B) List of 51 predicted target-genes for a tRF derived from the tRNA gene tRNA-Val-CAC-1-1.

CRISPR/Cas9 mediated deletion of tRNA-Val-CAC-1-1

Using the CRISPR/Cas9 genome engineering method, we cut tRNA-Val-CAC-1-1 from the genome (Figure 5). Two guide RNAs (gRNAs) with predicted efficiencies of 7.27 and 8.01, and no predicted off target sites, directed this Cas9 excision. We generated a donor plasmid containing an *attp* landing site, a dsRed positive selection marker, and the tRNA-Pro-TGG-1-1 gene which lies adjacent to tRNA-Val-CAC-1-1 and would otherwise be removed from the genome entirely. Two 1kb homologous regions, or homology arms, on the donor plasmid align the vector to the desired region and integrate the landing site and dsRed marker via homology-directed repair. A total of 3 injections of these plasmids yielded four positive lines of flies, all validated by PCR and Sanger sequencing (Figure 5D).

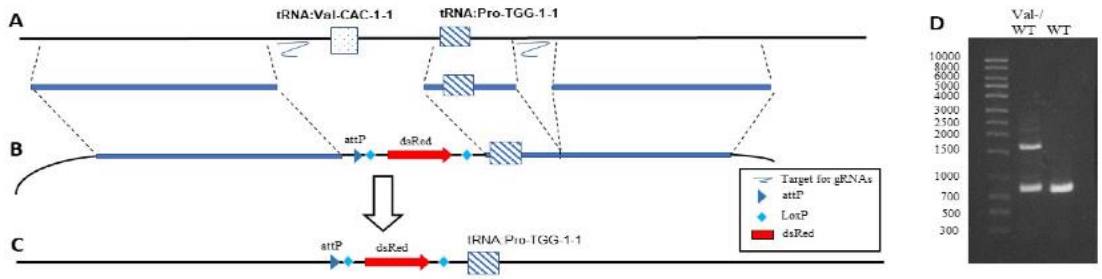


Figure 2: CRISPR/Cas9 strategy. (A) Genome model of the target tRNA:Val-CAC-1-1 and the adjacent tRNA:Pro-TGG-1-1. (B) Two guides with predicted efficiencies 7.27 and 8.01. Since the right guide RNA removes the tRNA:Pro-TGG-1-1, this fragment was cloned into the donor plasmid adjacent to the right homology arm. (C) The expected genome organization after the removal of the tRNA:Val-CAC-1-1. (D) PCR validation confirms Val tRNA null heterozygous fly line. The null fly is expected to produce a band size of 1719 bp, whereas the wild type fly is expected to generate a 828 bp band size. Bands were gel purified and sequenced to validate the expected changes (Revaitis, unpublished)

Necessity of tRNA-Val-CAC-1-1

Six other Val-CAC encoding genes exist in the *D. melanogaster* genome and five more non-CAC Valine tRNAs can compensate via wobbling. It was expected the deletion of tRNA-Val-CAC-1-1 would be tolerated as a homozygous mutation since the other six would be able to compensate for the lost tRNA activity. Three of these six are confirmed to be expressed during oogenesis (Figure S2). However, the fly did not tolerate a homozygous mutation. Flies were initially screened in bulk genomic preps (10-15 flies), then single fly genomic preps, and eventually wing-clip PCR reactions to avoid potential contamination from mixed populations of hetero and homozygotes. Since the tRNA-Val-CAC-1-1 deletion included an eye marker, the DsRed, we aimed to test whether the exogenous protein was interfering and if removing it would allow the fly to become homozygous null for the deletion. DsRed is flanked by two loxP sites in our donor vector, which allowed the removal through CRE-mediated recombination. Even without the DsRed, flies would not become homozygous.

One theory regarding tRNA-Val-CAC-1-1's potential necessity revolves around the predicted structures of these tRNA molecules, which can be calculated using *RNAFold* (Kerpedjiev et al., 2015), an open source program which constructs the most thermodynamically preferable model secondary RNA structures from given sequences. tRNA-Val-CAC-1-1 is the only Val-CAC encoding tRNA gene of its exact sequence in the genome. The other six Val-CAC genes (tRNA-Val-CAC-2-1 through tRNA-Val-CAC-2-6) differ by four base pairs. *RNAFold*'s model predicts a different, atypical tRNA structure for tRNA-Val-CAC-1-1 compared to the others, which are predicted to form a typical tRNA cloverleaf (Figure 6).

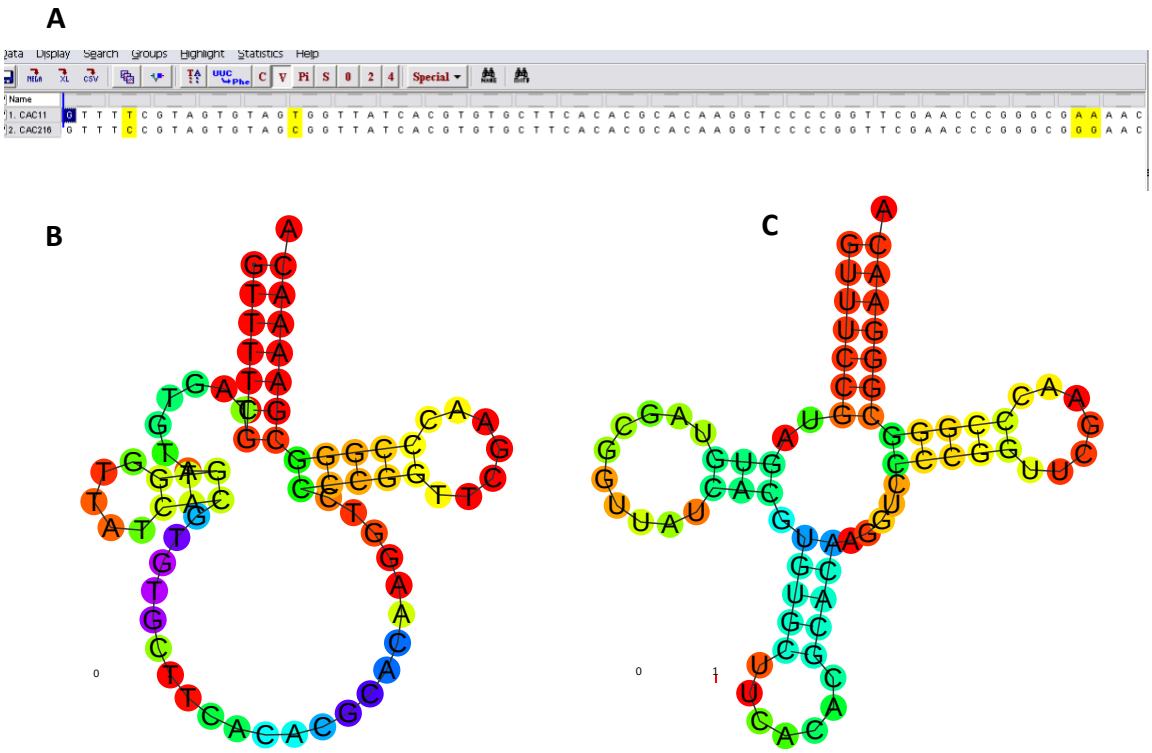


Figure 3: Predicted tRNA structures: A) MEGA alignment of tRNA-Val-CAC-1-1 and tRNA-Val-CAC-1-(1-6) B) RNAFold's predicted minimum free energy (MFE) structure for tRNA-Val-CAC-1-1 C) RNAFold's predicted MFE structure for tRNA-Val-CAC-2-1 and its five identical copies. Sliding color scale represents probability of location

The effect of heterozygous tRNA-Val-CAC-1-1 on transcriptional profile

RNA-seq data was obtained in triplicates for both the heterozygous tRNA-Val-CAC-1-1 CRISPR fly and a control fly which did not have the tRNA mutation but was otherwise identically balanced. This data was subsequently analyzed by the Grigoriev Lab, who found several differentially expressed genes in both tissues (Figure S1). While these data do not closely match up with our original list of 51 target genes (Figure 4), they do show several genes which demonstrate significant differential expression between our heterozygous and control conditions. Multiple genes, including ACXD, IP3K1, and *sta* demonstrate up or downregulation in both ovary and embryo samples. Read counts are also largely consistent between triplicate repeats. Additionally, a series of up-regulated gene targets in the ovary were then experimentally validated by qPCR (Figure 7). Previous qPCR experiments identified *Icarus (ics)* as a potentially upregulated target; however, this was not supported by RNA seq data. Repeat qPCR with different primers for *ics* demonstrated barely detectable levels of transcript, consistent with the single-digit reads in the RNA-seq data.

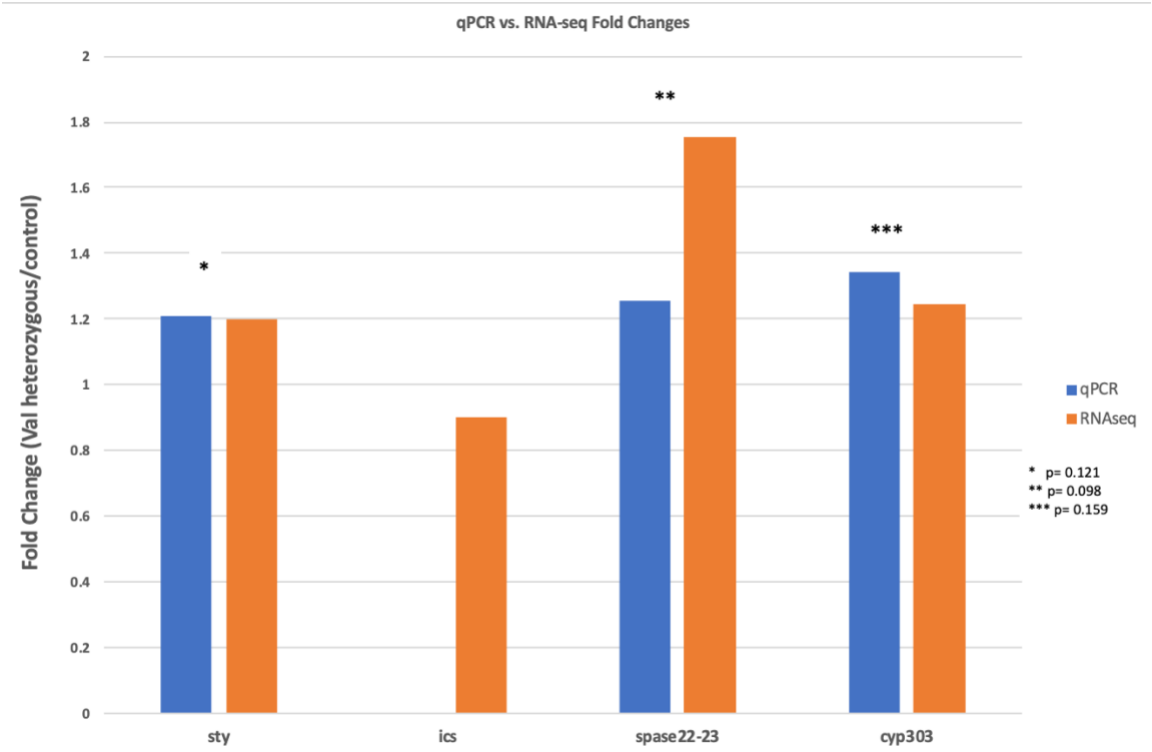


Figure 4: qPCR and RNA-seq comparison. Four genes found in the embryo were experimentally validated via qPCR. Fold change levels were comparable for *sty, spase22-23, and cyp303a1*. *Ics*, which was thought to have elevated expression levels in previous qPCR experiments showed no amplification in this run, comparable to its single digit reads in sequencing data. All p-values reflect are >0.05, reflecting insignificant differences between the two measurements of differential expression.

Small RNA Distribution changes in tRNA-Val-CAC-1-1 heterozygotes

Small RNA seq data for RNA molecules between 15-50bp was also obtained via PAGE of total RNA samples described above. The small RNA reads were also analyzed by the Grigoriev Lab. Results demonstrated few changes in Val-CAC tRFs themselves, but found widespread differential expression of other species of tRFs between the heterozygous and wild-type flies (Figure S3). However, the reason for these findings is unclear. Deletion of their parent molecule did not obviously reduce the number of tRF reads from Val-CAC-1-1. Additionally, global changes in small RNA populations were unexpected, suggesting a potential dynamic regulation of the entire small RNA transcriptome.

Discussion

The emerging field of tRFs represent the potential for the study of a new layer of gene regulation. Their role as post-transcriptional modifiers in tissue development is especially interesting, as there is currently only limited insight onto the effects of tRFs in this role (Guzzi et al., 2020). As the field expands and more experimentally based evidence is obtained, computational target prediction should also improve, creating a positive feedback loop of prediction and validation that should push the field forward.

Drosophila oogenesis, a long-standing model for the study of tissue development, is an excellent system in which to study tRFs. miRNAs, a biological analog of tRFs, have been studied extensively in the *Drosophila* ovary and have revealed critical insights into small RNA roles in cell division, stress response, and cell differentiation (Shcherbata, 2019). Additionally, previous studies were able to characterize the role of singular miRNA species in the development of other tissues, such as the eye (Li et al., 2005) and the wing discs (Kennel et al., 2008). Given the computational tools at our disposal (Karaiskos et al., 2016) for target prediction and its successful history of studying the roles of small RNAs, *Drosophila* is the ideal system in which tRF studies can take place.

Non-viability of a homozygous mutant

Many *Drosophila* gene mutations can be tolerated as both heterozygous and homozygous mutants without lethal effects (Eanes 1985). tRNA encoding genes, in

particular, seem like good candidates to tolerate mutations due to the redundancy of 100s of predicted tRNAs in the genomes that are lost or gained throughout evolutionary time (Rogers et al., 2010). Given this and the 6 additional Val-CAC genes in the *D. melanogaster* genome, it was surprising the CRISPR-mediated deletions failed to produce a fly homozygous for a null tRNA-Val-CAC-1-1, even after removal of the DsRed cassette. Analysis of previous RNA-seq data from the Yakoby Lab shows other Val-CAC-encoding tRNAs are expressed in oogenesis (Figure S2, Revaitis et al., 2017). Thus, availability of Val-CAC alternatives should not be a concern if tRNA-Val-CAC-1-1 is acting as a tRNA, as well as the other six. However, the inability of the fly to become homozygous null for the gene raises a question about the function of the gene, and whether the different predicted non-classical shape of the predicted structure is an indicator that this tRNA has a regulatory function beyond the tRNA activity. It also may be necessary for the development of other tissues.

One of the many significant outcomes of the Human Genome Project was the discovery RNA transcripts are significantly more abundant than protein encoding genes (van Bakel et al., 2010). This has led to the current understanding of small non-coding RNAs and their extensive roles as gene regulators. miRNAs, more than any other class of ncRNA, are a central figure in many approaches to therapeutic treatment for a variety of human diseases (Hanna et al., 2019). This includes the use in treatment for various types of cancer (Mercatelli et al., 2008, Reid et al., 2016), amyloidosis (Kristen et al., 2018), and fighting drug resistance (Wang et al., 2015). It's possible to imagine a similar path of discovery for tRFs. Current evidence suggests their targeting and gene silencing mechanisms are similar to miRNAs, and even in this small-scale study we may have

unearthed a critical regulatory mechanism. It is possible, given our evidence, tRNA molecules have an important biological role outside of their canonical role, and use tRFs to modify genes post-transcriptionally.

Transcriptional changes in heterozygous mutant

Our list of target genes does not closely represent the list of genes which demonstrate differential expression in Val-CAC-1-1 mutant heterozygotes. This may be in part due to the low stringency of the original prediction, which was designed to catch as many tRNA-Val-CAC-1-1-related genes as possible. In the case of *sty*, specifically, both RNA-seq and qPCR data demonstrate a fold change of around 1.2, indicative of, potentially, a small amount of lost transcriptional targeting. However, with only one functional copy of the tRNA, compensatory mechanisms for gene expression are likely to activate. It's also possible the tRFs loaded onto Argonaute proteins, for which we do not have a profile, were able to silence this gene sufficiently. Alternatively, other small RNAs which target *sty* may have been recruited to compensate. While it's encouraging one of the predicted targets did show changes, albeit insignificant, in transcript level, it's likely our heterozygous mutation wasn't enough to impact gene expression to a significant degree in many of our targets.

Unsurprisingly, we do not see any of the originally expected eggshell or eye phenotypes associated with reduced EGFR signaling. The EGFR pathway is one with many points of regulation, and it's likely able to compensate for whatever changes the minor increase in *sty* may have caused. For example, *sty* plays a similar role to Kekkon-1 (kek1) in EGFR signaling, but acts reciprocally (Hacohen et al., 1998; Musacchio and Perrimon, 1996; Zartman et al., 2011)- since both genes act in a negative feedback to

reduce EGFR pathway activation, possibly the amount of *kek1* may have been down-regulated to compensate for the increased *sty* (Figure 8).

Changes in small RNA populations

Analysis of the small RNA distributions in our heterozygous and wild-type conditions did not show significant changes in the numbers of Val-CAC-derived tRFs, a counterintuitive finding. However, there are many significantly up and downregulated tRFs derived from other tRNA molecules- this differential expression of tRFs is seen consistently across small RNA-seq replicates, and thus seems unlikely to fall within a standard deviation of small RNA populations. It is not entirely clear why this is the case. It's possible the deletion of tRNA-Val-CAC-1-1 created a stress response on the small RNA level, leading to the misregulation of other tRF species.

It is also interesting to note tRNA-Val-CAC derived tRFs are a recurring object of interest in tRF studies. In humans (Hogg M. et al., 2020; Skeparnias et al., 2020) and in plants (Byeon et al., 2019) tRFs from Val-CAC are significantly upregulated as a response to various external stressors, specifically disease and heat. It is possible the cell may require a minimum number of specific tRF species, such as Val-CAC tRFs, to mediate these stress responses. In this case, the single functional tRNA-Val-CAC-1-1 gene may have been a sufficient source for this requisite number of tRFs.

The answer may also lie in which tRFs are loaded onto Argonaute (Ago) proteins that enter the RISC pathway. tRFs can be loaded onto Ago proteins and sequestered to target mRNAs; they have been shown to sometimes even compete directly with miRNAs for these roles (Tuck 2011, Kumar et al., 2014). These tRFs would not be picked up on small RNA sequencing data from total cellular RNA, as riboprotein complexes will be dissolved by phenol RNA extractions. Thus, our current view of the small RNAs present in our tissues of interest is incomplete.

Analysis of tRF interactions with Ago proteins is an ongoing topic of interest. It has been previously found tRFs associate with Ago 1 and 2 in a Dicer-independent manner to regulate gene expression, and that amounts of tRFs loaded onto Ago proteins does increase with overexpression of their parent tRNA molecule (Kuscu et al., 2018). Additionally, tRFs loaded onto Ago proteins are dynamically regulated in an age dependent manner; Ago1 is preferentially loaded with a greater number of tRFs in 3 day old flies, but this preferential loading shifts dramatically to Ago2 in 30 day old flies (Karaiskos et al., 2015). Given the parameters of this experiment, which was restricted to

3-7 day old flies, we would expect to see any downshift in loaded tRNA-Val-CAC-1-1 tRF's reflected in our Ago1 RNA-IP data.

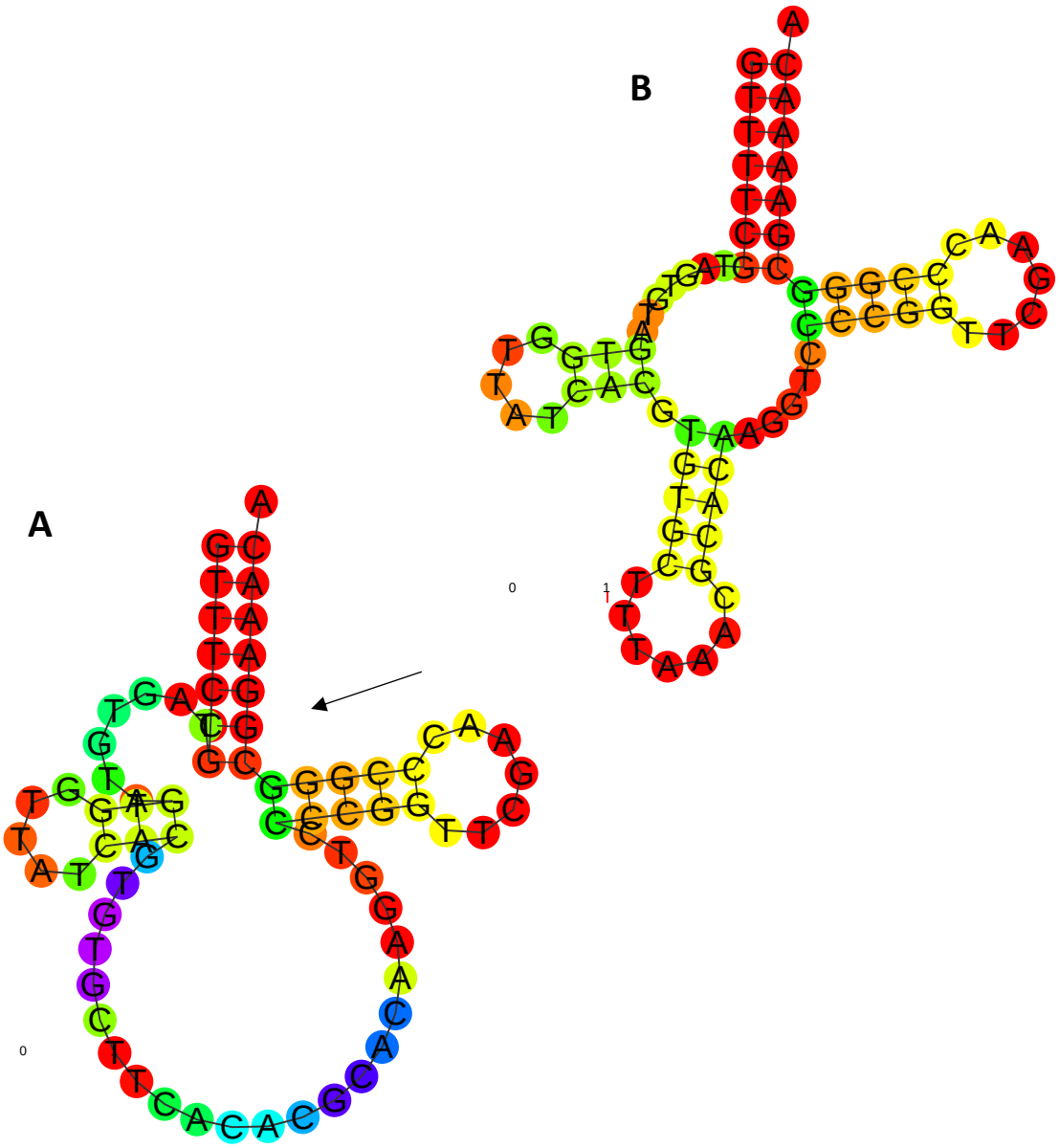
Future Directions

To further pursue the potential necessity of tRNA-Val-CAC-1-1 in developmental, and to determine which role- tRNA or tRF- is required if this is the case, we have planned the following experiments.

Three injections are designed (Figure 9) to determine if tRNA-Val-CAC-1-1 acts as a tRNA molecule or functions as a tRF. The donor vectors contain modified versions of the tRNA-Val-CAC-1-1 genes. One has two modified nucleotides which constitute one base pair found in the acceptor stem of the tRNA molecule. This region contains the operative seed sequence of the tRNA molecule; the changes made in this vector should cause it to have no predicted tRF activity, as its sequence will mirror that of the tRNA-Val-CAC-2 analogs. If this fly is homozygous viable, it indicates Val-CAC-1-1's lost tRNA activity was responsible for its necessity in our original mutant. Conversely, two vectors are designed with rarely used anticodon loop sequences for other amino acids (Leu-TAA and Ser-AGA, respectively) based on codon usage bias in *Drosophila* species (Vicario et al., 2007.) This will restore Val-CAC-1-1's tRF activity but will not restore

functional Val-tRNA activity. If this fly can tolerate a homozygous mutation, our missing regulatory tRF activity was the culprit in our original mutant.

As a control, we will inject the tRNA-Val-CAC-1-1, putting the gene back into the genome with our positive selection marker DsRed- this will control the possibility the observed effects are from an effect from the CRISPR separate from the activity tRNA-Val-CAC-1-1. Rescue with the endogenous gene should provide an effectively wild-type fly. These flies have been successfully injected with the appropriate donor vectors and are currently being screened for our positive selection marker, DsRed.



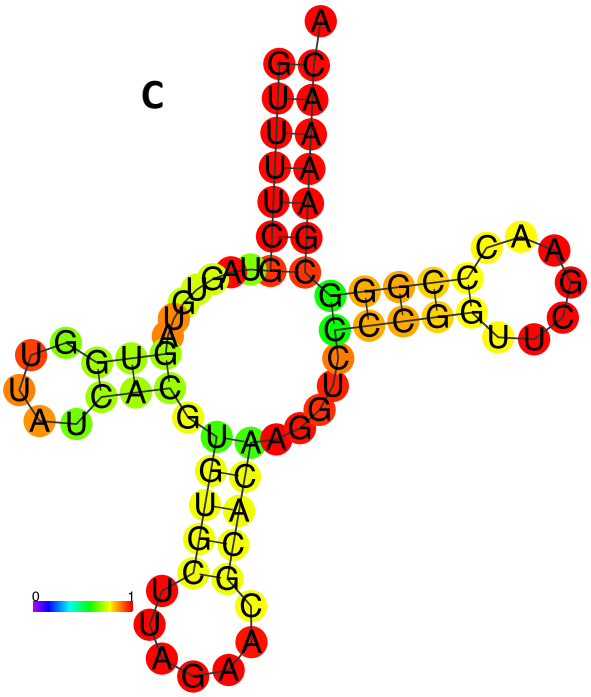


Figure 9: RNAFold Predicted Structures for current tRNA CRISPR experiments A) tRF seed sequence change- arrow indicates CG pairing changed from AT in the endogenous gene B) Val-CAC tRF with Leucine-AGA anticodon C) Val-CAC tRF with Serine-AGA anticodon

Additionally, our current profile of tRF populations is limited to those which are not loaded onto Ago proteins and is thus incompelete. tRFs other than the species we have reads for are being used by the cellular machinery of the RISC pathway. This is being addressed with another currently underway experiment. In this, RNA is extracted from Ago1 and Ago2 immunoprecipitation products. Sequencing of these RNA extracts should complete our picture of small RNA populations and allow us to see the full effect of the tRNA-Val-CAC-1-1 deletion on the distribution of small RNAs.

Further experiments may also help define the specifics of the effects of the lost tRNA-Val-CAC-1-1. Namely, overexpression of tRNA-Val-CAC-1-1 through UAS/Gal4 and subsequent screening for differentially expressed genes affected by the greater number of tRFs may further elucidate seed-target interaction mechanisms. Additionally, it may be beneficial to attempt our injected CRISPR endogenous tRNA-Val-CAC-1-1 rescue with the gene inserted at a different location in the genome, rather than its original site. This would rule out the possibility of our effects being a result of a missing

regulatory region between tRNA-Val-CAC-1-1 and tRNA-Pro-TGG-1-1, rather than our gene of interest.

Supplementary Figures

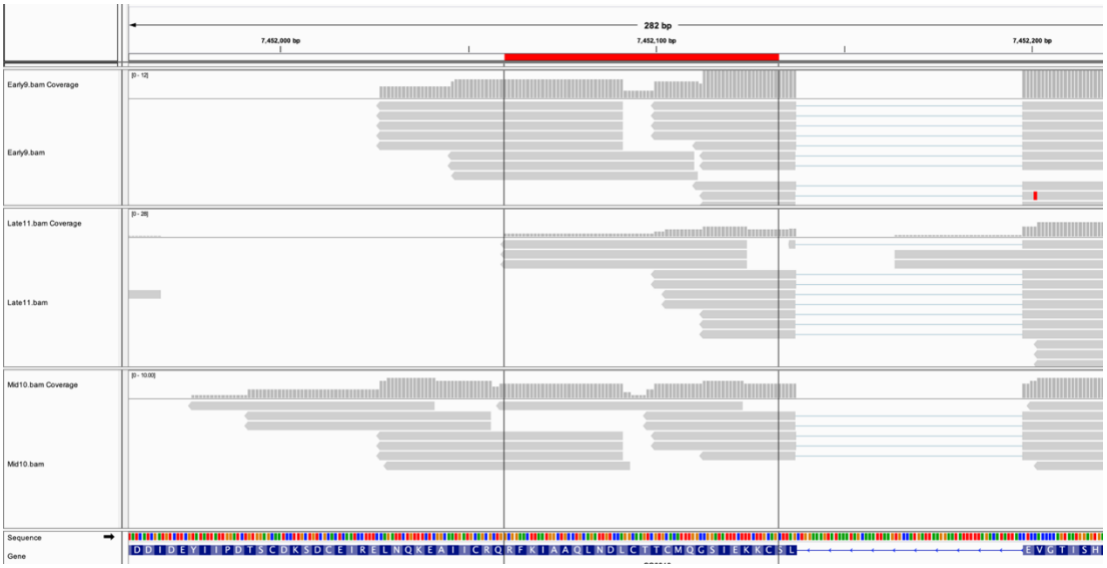
downregulated, embryo														
#gene_id	gene_name	biotype	EC1	EC2	EC3	EV1	EV2	EV3	baseMean	log2FoldCha	lfcSE	stat	pvalue	padj
FBgn0026179	siz	protein_coding	188	410	453	21	13	16	190.13208	-4.6266405	0.3868319	-11.960339	5.73E-33	5.28E-29
FBgn0086913	Rab26	protein_coding	162	207	211	4	1	0	102.29778	-7.1367027	0.7210838	-9.8971894	4.28E-23	1.97E-19
FBgn0037133	CG7370	protein_coding	64	42	68	8	0	0	30.777224	-4.7788695	0.7005219	-6.8218698	8.99E-12	1.38E-08
FBgn0267033	mamo	protein_coding	80	82	70	27	7	9	46.893351	-2.7737554	0.4433829	-6.2558916	3.95E-10	4.55E-07
FBgn0023416	Ac3	protein_coding	55	48	59	9	9	6	31.964346	-2.9564967	0.4936389	-5.9891896	2.11E-09	1.77E-06
downregulated, ovary														
#gene_id	gene_name	biotype	OC1	OC2	OC3	OV1	OV2	OV3	baseMean	log2FoldCha	lfcSE	stat	pvalue	padj
FBgn0032166	CG4619	protein_coding	1039	1044	1026	438	446	434	744.02473	-1.3721801	0.0742613	-1.85E+01	3.12E-76	3.21E-73
FBgn0030562	CG9400	protein_coding	1677	1537	1624	818	941	894	1251.5783	-1.0020185	0.0586471	-1.71E+01	1.90E-65	1.40E-62
FBgn0031343	CG18131	protein_coding	853	840	847	354	385	466	627.36471	-1.2125222	0.0851804	-1.42E+01	5.58E-46	3.58E-43
FBgn0250755	CG42233	protein_coding	457	428	445	188	185	201	319.79556	-1.3466295	0.1112229	-1.21E+01	9.64E-34	4.96E-31
FBgn0035533	Cip4	protein_coding	208	210	211	67	75	61	140.51671	-1.7657161	0.1732927	-1.02E+01	2.22E-24	8.13E-22
upregulated, embryo														
#gene_id	gene_name	biotype	EC1	EC2	EC3	EV1	EV2	EV3	baseMean	log2FoldCha	lfcSE	stat	pvalue	padj
FBgn0263036	CG43331	protein_coding	0	0	0	50	27	18	13.87032	7.1701732	1.3144328	5.4549562	4.90E-08	2.51E-05
FBgn0033160	Dhx15	protein_coding	32	37	50	254	129	99	90.412214	1.7673444	0.3490757	5.0629256	4.13E-07	0.0001586
FBgn0004034	y	protein_coding	1	0	1	43	34	22	15.515873	5.4557009	1.1612122	4.6982807	2.62E-06	0.0006597
FBgn0038247	Cad88C	protein_coding	3	1	7	36	30	38	17.784407	3.1048705	0.6762017	4.59E+00	4.40E-06	0.0010232
FBgn0000150	awd	protein_coding	110	49	70	374	135	303	155.09947	1.5357289	0.3346074	4.5896448	4.44E-06	0.0010232
upregulated, ovary														
#gene_id	gene_name	biotype	OC1	OC2	OC3	OV1	OV2	OV3	baseMean	log2FoldCha	lfcSE	stat	pvalue	padj
FBgn0262617	CG43143	protein_coding	945	911	911	2073	2316	2338	1534.1283	1.1460285	0.052249	2.19E+01	1.23E-106	3.17E-103
FBgn0031235	CG13693	protein_coding	579	544	576	985	1021	1090	779.81736	0.7308681	0.0707335	1.03E+01	5.01E-25	1.98E-22
FBgn0000083	AnxB9	protein_coding	463	474	513	852	920	947	677.51864	0.7716546	0.07655	1.01E+01	6.75E-24	2.31E-21
FBgn0039663	CG2321	protein_coding	989	1079	945	1472	1657	1556	1256.4484	0.5015833	0.0617554	8.12E+00	4.58E-16	9.06E-14
FBgn0032167	CG5853	protein_coding	111	127	139	275	311	318	207.03741	1.1254368	0.14013	8.03E+00	9.64E-16	1.77E-13

downregulated, matching 7mer(CCGGGCG), embryo														
#gene_id	gene_name	biotype	EC1	EC2	EC3	EV1	EV2	EV3	baseMean	log2FoldCha	lfcSE	stat	pvalue	padj
FBgn0011206	bol	protein_coding	42	52	63	21	6	5	31.549621	-2.6075053	0.4995564	-5.2196416	1.79E-07	7.87E-05
FBgn0003435	sm	protein_coding	58	64	55	30	8	9	37.66492	-2.2605695	0.4622969	-4.8898649	1.01E-06	0.0003445
FBgn0024913	Actbeta	protein_coding	30	20	21	1	2	4	13.621701	-3.5439057	0.7879217	-4.4977892	6.87E-06	0.001376
FBgn0014388	sty	protein_coding	42	59	58	34	6	6	33.790951	-2.1803759	0.5167992	-4.2190004	2.45E-05	0.003187
FBgn0029587	CG14797	protein_coding	216	282	460	158	137	148	231.55443	-1.2583671	0.3007079	-4.1846833	2.86E-05	0.0036059
downregulated, maching 7mer(CCGGGCG), ovary														
#gene_id	gene_name	biotype	OC1	OC2	OC3	OV1	OV2	OV3	baseMean	log2FoldCha	lfcSE	stat	pvalue	padj
FBgn0262570	CG43110	protein_coding	827	794	823	514	557	543	675.50944	-0.7333473	0.0744914	-9.8447299	7.22E-23	2.32E-20
FBgn0039908	Asator	protein_coding	1838	1851	1848	1567	1798	1808	1767.8649	-0.2341398	0.0483187	-4.8457382	1.26E-06	7.20E-05
FBgn0032147	IP3K1	protein_coding	241	243	263	188	193	199	220.14659	-0.4996459	0.1289192	-3.8756527	0.0001063	0.0035037
FBgn0031183	CG14621	protein_coding	370	382	340	288	313	307	331.21065	-0.401181	0.1063647	-3.7717495	0.0001621	0.0049304
FBgn0040507	ACXD	protein_coding	14	9	16	2	0	0	7.0564457	-4.4081538	1.2033207	-3.6633241	0.000249	0.007149
upregulated, matching 7mer(CCGGGCG), embryo														
#gene_id	gene_name	biotype	EC1	EC2	EC3	EV1	EV2	EV3	baseMean	log2FoldCha	lfcSE	stat	pvalue	padj
FBgn0263036	CG43331	protein_coding	0	0	0	50	27	18	13.87032	7.1701732	1.3144328	5.4549562	4.90E-08	2.51E-05
FBgn0003517	sta	protein_coding	407	250	346	1156	394	819	507.4923	0.9399025	0.2684961	3.500619	0.0004642	0.0289108
FBgn0001120	gnu	protein_coding	32	23	32	110	45	97	51.270432	1.2711638	0.3736695	3.4018398	0.0006693	0.0352569
FBgn0031145	Ntf-2	protein_coding	33	30	52	139	67	93	62.979439	1.1433144	0.3372991	3.3896157	0.0006999	0.0358839
FBgn0033566	CG18004	protein_coding	11	5	16	77	20	37	23.55455	1.768503	0.5347299	3.3072825	0.0009421	0.0423605
upregulated, matching 7mer(CCGGGCG), ovary														
#gene_id	gene_name	biotype	OC1	OC2	OC3	OV1	OV2	OV3	baseMean	log2FoldCha	lfcSE	stat	pvalue	padj
FBgn0262617	CG43143	protein_coding	945	911	911	2073	2316	2338	1534.1283	1.1460285	0.052249	21.933977	1.23E-106	3.17E-103
FBgn0039172	Spase22-23	protein_coding	283	286	309	518	554	558	407.76597	0.7576313	0.0964473	7.855388	3.99E-15	6.40E-13
FBgn0001992	Cyp303a1	protein_coding	2961	2806	2789	3553	3856	3725	3226.9465	0.2453974	0.0378975	6.4752869	9.46E-11	1.08E-08
FBgn0260744	Tango9	protein_coding	2321	2356	2398	2780	3236	3433	2703.5522	0.2802738	0.0464578	6.0328644	1.61E-09	1.51E-07
FBgn0033421	CG1888	protein_coding	7	5	3	25	40	27	17.005121	2.4792701	0.5537934	4.4768864	7.57E-06	0.0003605

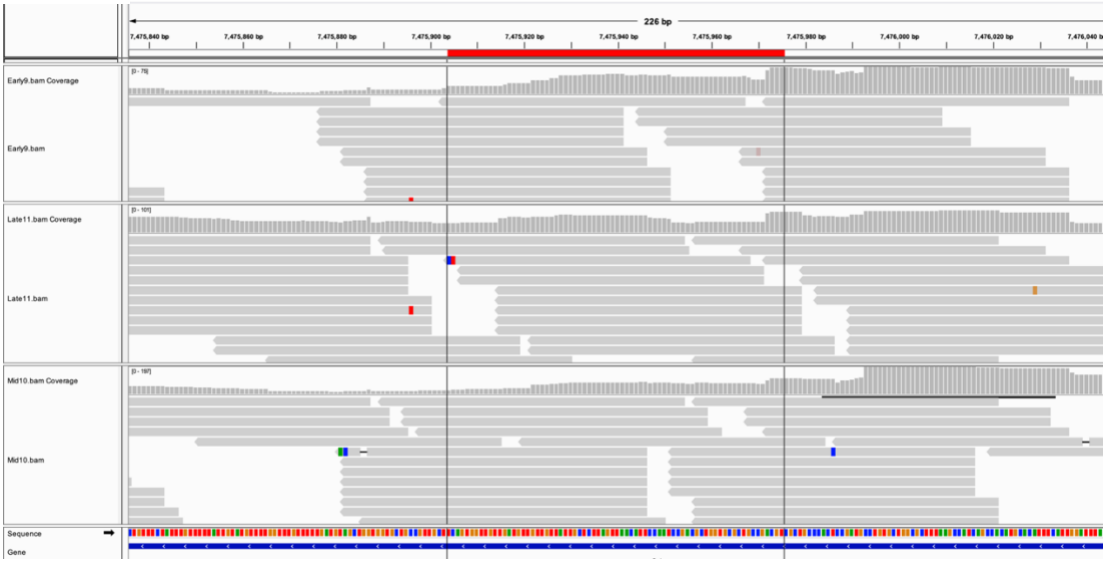
downregulated, matching 7mer+A_in_5nt_flanking, embryo														
#gene_id	gene_name	biotype	EC1	EC2	EC3	EV1	EV2	EV3	baseMean	log2FoldCha	lfcSE	stat	pvalue	padj
FBgn0024913	Actbeta	protein_coding	30	20	21	1	2	4	13.621701	-3.5439057	0.7879217	-4.4977892	6.87E-06	0.001376
FBgn0014388	sty	protein_coding	42	59	58	34	6	6	33.790951	-2.1803759	0.5167992	-4.2190004	2.45E-05	0.003187
FBgn0029587	CG14797	protein_coding	216	282	460	158	137	148	231.55443	-1.2583671	0.3007079	-4.1846833	2.86E-05	0.0036059
FBgn0052647	CG32647	protein_coding	160	128	216	54	83	21	112.69954	-1.7249371	0.4437034	-3.8875906	0.0001012	0.0093327
FBgn0004595	pros	protein_coding	39	29	27	11	9	5	20.641463	-2.1566288	0.562609	-3.8332636	0.0001265	0.010894
downregulated, matching 7mer+A_in_5nt_flanking, ovary														
#gene_id	gene_name	biotype	OC1	OC2	OC3	OV1	OV2	OV3	baseMean	log2FoldCha	lfcSE	stat	pvalue	padj
FBgn0262570	CG43110	protein_coding	827	794	823	514	557	543	675.50944	-0.7333473	0.0744914	-9.8447299	7.22E-23	2.32E-20
FBgn0039908	Asator	protein_coding	1838	1851	1848	1567	1798	1808	1767.8649	-0.2341398	0.0483187	-4.8457382	1.26E-06	7.20E-05
FBgn0032147	IP3K1	protein_coding	241	243	263	188	193	199	220.14659	-0.4996459	0.1289192	-3.8756527	0.0001063	0.0035037
FBgn0040507	ACXD	protein_coding	14	9	16	2	0	0	7.0564457	-4.4081538	1.2033207	-3.6633241	0.000249	0.007149
FBgn0262866	S6kII	protein_coding	178	154	165	128	136	129	147.5667	-0.4726683	0.1574837	-3.0013794	0.0026876	0.0527261
upregulated, matching 7mer+A_in_5nt_flanking, embryo														
#gene_id	gene_name	biotype	EC1	EC2	EC3	EV1	EV2	EV3	baseMean	log2FoldCha	lfcSE	stat	pvalue	padj
FBgn0263036	CG43331	protein_coding	0	0	0	50	27	18	13.87032	7.1701732	1.3144328	5.4549562	4.90E-08	2.51E-05
FBgn0003517	sta	protein_coding	407	250	346	1156	394	819	507.4923	0.9399025	0.2684961	3.500619	0.0004642	0.0289108
FBgn0031145	Ntf-2	protein_coding	33	30	52	139	67	93	62.979439	1.1433144	0.3372991	3.3896157	0.0006999	0.0358839
FBgn0033566	CG18004	protein_coding	11	5	16	77	20	37	23.55455	1.768503	0.5347299	3.3072825	0.0009421	0.0423605
FBgn0037686	RpL34b	protein_coding	368	245	352	1023	428	581	456.85985	0.7976693	0.2435322	3.2754161	0.0010551	0.0448184
upregulated, matching 7mer+A_in_5nt_flanking, ovary														
#gene_id	gene_name	biotype	OC1	OC2	OC3	OV1	OV2	OV3	baseMean	log2FoldCha	lfcSE	stat	pvalue	padj
FBgn0262617	CG43143	protein_coding	945	911	911	2073	2316	2338	1534.1283	1.1460285	0.052249	21.933977	1.23E-106	3.17E-103
FBgn0039172	Spase22-23	protein_coding	283	286	309	518	554	558	407.76597	0.7576313	0.0964473	7.855388	3.99E-15	6.40E-13
FBgn0001992	Cyp303a1	protein_coding	2961	2806	2789	3553	3856	3725	3226.9465	0.2453974	0.0378975	6.4752869	9.46E-11	1.08E-08
FBgn0033421	CG1888	protein_coding	7	5	3	25	40	27	17.005121	2.4792701	0.5537934	4.4768864	7.57E-06	0.0003605
FBgn0028577	hfp	protein_coding	4	2	8	34	26	26	16.01564	2.4911453	0.5773484	4.3148042	1.60E-05	0.0006786
downregulated, matching 8mer(CCGGGCGA), embryo														
#gene_id	gene_name	biotype	EC1	EC2	EC3	EV1	EV2	EV3	baseMean	log2FoldCha	lfcSE	stat	pvalue	padj
FBgn0024913	Actbeta	protein_coding	30	20	21	1	2	4	13.621701	-3.5439057	0.7879217	-4.4977892	6.87E-06	0.001376
FBgn0037238	CG1090	protein_coding	35	11	5	3	0	0	9.5467964	-4.5025119	1.1925658	-3.775483	0.0001597	0.0128008
FBgn0032147	IP3K1	protein_coding	531	475	524	312	293	215	396.03855	-1.0657752	0.2888018	-3.6903341	0.000224	0.0165157
FBgn0004198	ct	protein_coding	28	29	43	17	10	5	21.884479	-1.8757398	0.5335666	-3.5154743	0.000439	0.0275267
FBgn0026411	Lim1	protein_coding	19	17	7	6	2	3	9.3828963	-2.3104045	0.8171242	-2.8274827	0.0046916	0.110691
downregulated, matching 8mer(CCGGGCGA), ovary														
#gene_id	gene_name	biotype	OC1	OC2	OC3	OV1	OV2	OV3	baseMean	log2FoldCha	lfcSE	stat	pvalue	padj
FBgn0032147	IP3K1	protein_coding	241	243	263	188	193	199	220.14659	-0.4996459	0.1289192	-3.8756527	0.0001063	0.0035037
FBgn0040507	ACXD	protein_coding	14	9	16	2	0	0	7.0564457	-4.4081538	1.2033207	-3.6633241	0.000249	0.007149
FBgn0266756	btsz	protein_coding	70	79	78	40	46	65	62.813788	-0.7274591	0.2506737	-2.9020157	0.0037077	0.0668687
FBgn0038320	Sra-1	protein_coding	229	209	216	176	169	206	199.44304	-0.3821103	0.1378609	-2.7717086	0.0055763	0.0895692
FBgn0023550	FarO	protein_coding	38	33	44	21	20	23	29.909149	-0.9797279	0.3541265	-2.7666047	0.0056643	0.0901384
upregulated, matching 8mer(CCGGGCGA), embryo														
#gene_id	gene_name	biotype	EC1	EC2	EC3	EV1	EV2	EV3	baseMean	log2FoldCha	lfcSE	stat	pvalue	padj
FBgn0033566	CG18004	protein_coding	11	5	16	77	20	37	23.55455	1.768503	0.5347299	3.3072825	0.0009421	0.0423605
FBgn0038877	CG3308	protein_coding	68	95	110	266	148	147	129.94616	0.7982343	0.298344	2.6755501	0.0074607	0.1414696
FBgn0039733	CG11504	protein_coding	19	26	37	82	40	54	39.636286	0.8609401	0.3969227	2.1690371	0.0300799	0.2863682
FBgn0028697	RpL15	protein_coding	117	60	79	369	70	148	122.09013	0.8046675	0.3829194	2.1014019	0.0356057	0.3027798
FBgn0053170	CG33170	protein_coding	107	114	134	271	153	187	151.96174	0.5456751	0.2642709	2.0648324	0.0389389	0.3156241
upregulated, matching 8mer(CCGGGCGA), ovary														
#gene_id	gene_name	biotype	OC1	OC2	OC3	OV1	OV2	OV3	baseMean	log2FoldCha	lfcSE	stat	pvalue	padj
FBgn0085478	CG34449	protein_coding	6	11	8	42	34	30	21.090122	1.9550829	0.4676416	4.180729	2.91E-05	0.001176
FBgn0029922	CG14431	protein_coding	302	299	261	383	383	421	335.39287	0.3269452	0.1079341	3.0291179	0.0024527	0.0491724
FBgn0032402	CG14945	protein_coding	1507	1428	1469	1597	1846	1955	1606.9647	0.1569521	0.0539014	2.9118386	0.0035931	0.0655597
FBgn0261556	CG42674	protein_coding	96	86	67	136	105	174	108.18856	0.602404	0.2131978	2.8255645	0.0047197	0.079021
FBgn0038877	CG3308	protein_coding	130	109	129	169	180	186	147.56793	0.4050924	0.1581854	2.5608712	0.010441	0.1470323

Figure S1:RNA Seq data for heterozygous tRNA-Val-CAC-1-1 mutant: Top 5 up and down regulated targets between conditions, per tissue. Genes are grouped according to matches to different categories of seed predictions. Group where seed is not indicated is overall most up/down regulated. Genes which appear in multiple tissues are highlighted in orange. *Sty* is highlighted in yellow. Ribosomal genes are highlighted in green, as tRFs may affect their translation.

A


282 bp
7,452,000 bp
7,452,100 bp
7,452,200 bp
Early9.bam Coverage
Early9.bam
Late11.bam Coverage
Late11.bam
Mid10.bam Coverage
Mid10.bam
Sequence
Gene

B


226 bp
7,475,840 bp
7,475,860 bp
7,475,880 bp
7,475,900 bp
7,475,920 bp
7,475,940 bp
7,475,960 bp
7,475,980 bp
7,476,000 bp
7,476,020 bp
Early9.bam Coverage
Early9.bam
Late11.bam Coverage
Late11.bam
Mid10.bam Coverage
Mid10.bam
Sequence
Gene

C

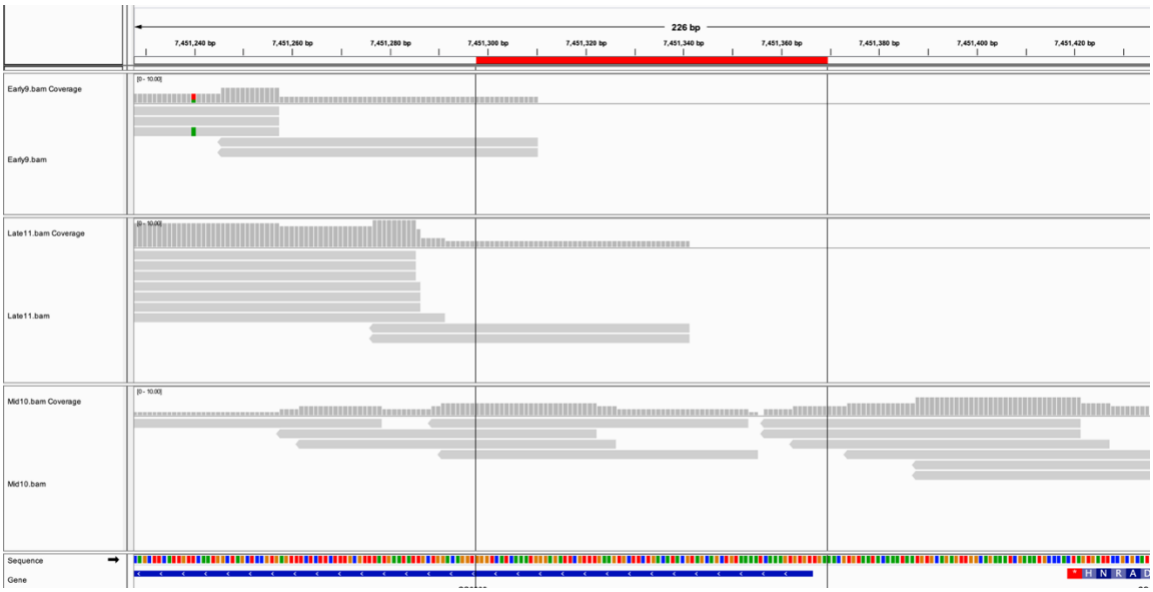


Figure S2: Expression of other tRNA-Val-CAC genes according to RNA-seq data: Using IGV software, regions encoding for tRNA-Val-CAC-2-1 (A) Val-CAC-2-2 (B) and Val-CAC-2-4 (C) were confirmed to be expressed in oogenesis. Data is adapted from Revaitis et al., 2017, where RNA-seq was done for early (stage 1-9), middle (stage 10), and late (stage 11-14) egg chambers.

Primer Name	Sequence
styF	GTGGATAATCCCTGCTCCTG
styR	ATGGGCCAGTAGAACCACAG
icsF	CAGATACTCGGATTGCGTGA
icsR	AGCACCTGCAGCCTATTGTT
Spase22-23F	GAAGGACAACCGCAACGTAT
Spase22-23R	TGAACAGGTGCTTTCCAGTG
Cyp303a1F	CATGTTTTGCAAGGTGATCG
Cyp303a1R	ATGGCATCGTTCGTATAGGC

Table 1: Primer sequences for qPCR

Supplementary Figure 3 Small RNA Distribution- included as hyperlink due to size

References

Babiarz, J. E., Ruby, J. G., Wang, Y., Bartel, D. P., & Blelloch, R. (2008). Mouse ES cells express endogenous shRNAs, siRNAs, and other Microprocessor-independent, Dicer-dependent small RNAs. *Genes Dev, 22*(20), 2773-2785. doi:10.1101/gad.1705308

Berg, C. A. (2005). The Drosophila shell game: patterning genes and morphological change. *Trends Genet, 21*(6), 346-355. doi:10.1016/j.tig.2005.04.010

Byeon, B., Bilichak, A., & Kovalchuk, I. (2019). Transgenerational Response to Heat Stress in the Form of Differential Expression of Noncoding RNA Fragments in Brassica rapa Plants. *Plant Genome, 12*(1). doi:10.3835/plantgenome2018.04.0022

Carvalho, G. B., Ja, W. W., & Benzer, S. (2009). Non-lethal PCR genotyping of single Drosophila. *Biotechniques, 46*(4), 312-314. doi:10.2144/000113088

Casci, T., Vinos, J., & Freeman, M. (1999). Sprouty, an intracellular inhibitor of Ras signaling. *Cell, 96*(5), 655-665. doi:10.1016/s0092-8674(00)80576-0

Chen, X. (2012). Small RNAs in development - insights from plants. *Curr Opin Genet Dev, 22*(4), 361-367. doi:10.1016/j.gde.2012.04.004

Cole, C., Sobala, A., Lu, C., Thatcher, S. R., Bowman, A., Brown, J. W., . . . Hutvagner, G. (2009). Filtering of deep sequencing data reveals the existence of abundant Dicer-dependent small RNAs derived from tRNAs. *RNA, 15*(12), 2147-2160. doi:10.1261/rna.1738409

Duhart, J. C., Parsons, T. T., & Raftery, L. A. (2017). The repertoire of epithelial morphogenesis on display: Progressive elaboration of Drosophila egg structure. *Mech Dev, 148*, 18-39. doi:10.1016/j.mod.2017.04.002

Eanes, W. F., Hey, J., & Houle, D. (1985). Homozygous and Hemizygous Viability Variation on the X Chromosome of DROSOPHILA MELANOGASTER. *Genetics, 111*(4), 831-844. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/17246306

Garcia-Silva, M. R., Cabrera-Cabrera, F., Guida, M. C., & Cayota, A. (2012). Hints of tRNA-Derived Small RNAs Role in RNA Silencing Mechanisms. *Genes (Basel), 3*(4), 603-614. doi:10.3390/genes3040603

Gebetsberger, J., & Polacek, N. (2013). Slicing tRNAs to boost functional ncRNA diversity. *RNA Biol, 10*(12), 1798-1806. doi:10.4161/rna.27177

Gebetsberger, J., Zywicki, M., Kunzi, A., & Polacek, N. (2012). tRNA-derived fragments target the ribosome and function as regulatory non-coding RNA in Haloferax volcanii. *Archaea, 2012,* 260909. doi:10.1155/2012/260909

Goodarzi, H., Liu, X., Nguyen, H. C., Zhang, S., Fish, L., & Tavazoie, S. F. (2015). Endogenous tRNA-Derived Fragments Suppress Breast Cancer Progression via YBX1 Displacement. *Cell, 161*(4), 790-802. doi:10.1016/j.cell.2015.02.053

Gratz, S. J., Ukken, F. P., Rubinstein, C. D., Thiede, G., Donohue, L. K., Cummings, A. M., & O'Connor-Giles, K. M. (2014). Highly specific and efficient CRISPR/Cas9-catalyzed homology-directed repair in Drosophila. *Genetics, 196*(4), 961-971. doi:10.1534/genetics.113.160713

Grimson, A., Farh, K. K., Johnston, W. K., Garrett-Engele, P., Lim, L. P., & Bartel, D. P. (2007). MicroRNA targeting specificity in mammals: determinants beyond seed pairing. *Mol Cell, 27*(1), 91-105. doi:10.1016/j.molcel.2007.06.017

Guo, Z., Li, Y., & Ding, S. W. (2019). Small RNA-based antimicrobial immunity. *Nat Rev Immunol, 19*(1), 31-44. doi:10.1038/s41577-018-0071-x

Guzzi, N., & Bellodi, C. (2020). Novel insights into the emerging roles of tRNA-derived fragments in mammalian development. *RNA Biol, 17*(8), 1214-1222. doi:10.1080/15476286.2020.1732694

Hacohen, N., Kramer, S., Sutherland, D., Hiromi, Y., & Krasnow, M. A. (1998). sprouty encodes a novel antagonist of FGF signaling that patterns apical branching of the Drosophila airways. *Cell, 92*(2), 253-263. doi:10.1016/s0092-8674(00)80919-8

Haeusler, R. A., & Engelke, D. R. (2006). Spatial organization of transcription by RNA polymerase III. *Nucleic Acids Res, 34*(17), 4826-4836. doi:10.1093/nar/gkl656

Hales, K. G., Korey, C. A., Larracuente, A. M., & Roberts, D. M. (2015). Genetics on the Fly: A Primer on the Drosophila Model System. *Genetics, 201*(3), 815-842. doi:10.1534/genetics.115.183392

Hamm, D. C., & Harrison, M. M. (2018). Regulatory principles governing the maternal-to-zygotic transition: insights from Drosophila melanogaster. *Open Biol, 8*(12), 180183. doi:10.1098/rsob.180183

Hanna, J., Hossain, G. S., & Kocerha, J. (2019). The Potential for microRNA Therapeutics and Clinical Research. *Front Genet, 10*, 478. doi:10.3389/fgene.2019.00478

Hutvagner, G., & Simard, M. J. (2008). Argonaute proteins: key players in RNA silencing. *Nat Rev Mol Cell Biol, 9*(1), 22-32. doi:10.1038/nrm2321

Karaiskos, S., & Grigoriev, A. (2016). Dynamics of tRNA fragments and their targets in aging mammalian brain. *F1000Res, 5*. doi:10.12688/f1000research.10116.1

Karaiskos, S., Naqvi, A. S., Swanson, K. E., & Grigoriev, A. (2015). Age-driven modulation of tRNA-derived fragments in Drosophila and their potential targets. *Biol Direct, 10*, 51. doi:10.1186/s13062-015-0081-6

Kawamura, Y., Saito, K., Kin, T., Ono, Y., Asai, K., Sunohara, T., . . . Siomi, H. (2008). Drosophila endogenous small RNAs bind to Argonaute 2 in somatic cells. *Nature, 453*(7196), 793-797. doi:10.1038/nature06938

Kennell, J. A., Gerin, I., MacDougald, O. A., & Cadigan, K. M. (2008). The microRNA miR-8 is a conserved negative regulator of Wnt signaling. *Proc Natl Acad Sci U S A, 105*(40), 15417-15422. doi:10.1073/pnas.0807763105

*Kerpedjiev P, Hammer S, Hofacker IL (2015). Forna (force-directed RNA): Simple and effective online RNA secondary structure diagrams***.** Bioinformatics 31(20):3377-9.

Kondo, S., & Ueda, R. (2013). Highly improved gene targeting by germline-specific Cas9 expression in Drosophila. *Genetics, 195*(3), 715-721. doi:10.1534/genetics.113.156737

Kristen, A. V., Ajroud-Driss, S., Conceicao, I., Gorevic, P., Kyriakides, T., & Obici, L. (2019). Patisiran, an RNAi therapeutic for the treatment of hereditary transthyretin-mediated amyloidosis. *Neurodegener Dis Manag, 9*(1), 5-23. doi:10.2217/nmt-2018-0033

Kumar, P., Anaya, J., Mudunuri, S. B., & Dutta, A. (2014). Meta-analysis of tRNA derived RNA fragments reveals that they are evolutionarily conserved and associate with AGO proteins to recognize specific RNA targets. *BMC Biol, 12*, 78. doi:10.1186/s12915-014-0078-0

Kumar, P., Kuscu, C., & Dutta, A. (2016). Biogenesis and Function of Transfer RNA-Related Fragments (tRFs). *Trends Biochem Sci, 41*(8), 679-689. doi:10.1016/j.tibs.2016.05.004

Kuscu, C., Kumar, P., Kiran, M., Su, Z., Malik, A., & Dutta, A. (2018). tRNA fragments (tRFs) guide Ago to regulate gene expression post-transcriptionally in a Dicer-independent manner. *RNA, 24*(8), 1093-1105. doi:10.1261/rna.066126.118

Lee, Y. S., Shibata, Y., Malhotra, A., & Dutta, A. (2009). A novel class of small RNAs: tRNA-derived RNA fragments (tRFs). *Genes Dev, 23*(22), 2639-2649. doi:10.1101/gad.1837609

Li, X., & Carthew, R. W. (2005). A microRNA mediates EGF receptor signaling and promotes photoreceptor differentiation in the Drosophila eye. *Cell, 123*(7), 1267-1277. doi:10.1016/j.cell.2005.10.040

Li, Y., Luo, J., Zhou, H., Liao, J. Y., Ma, L. M., Chen, Y. Q., & Qu, L. H. (2008). Stress-induced tRNA-derived RNAs: a novel class of small RNAs in the primitive eukaryote Giardia lamblia. *Nucleic Acids Res, 36*(19), 6048-6055. doi:10.1093/nar/gkn596

Li, Z., Ender, C., Meister, G., Moore, P. S., Chang, Y., & John, B. (2012). Extensive terminal and asymmetric processing of small RNAs from rRNAs, snoRNAs, snRNAs, and tRNAs. *Nucleic Acids Res, 40*(14), 6787-6799. doi:10.1093/nar/gks307

Li, Z., Xu, R., & Li, N. (2018). Correction to: MicroRNAs from plants to animals, do they define a new messenger for communication? *Nutr Metab (Lond), 15*, 74. doi:10.1186/s12986-018-0311-x

Maute, R. L., Schneider, C., Sumazin, P., Holmes, A., Califano, A., Basso, K., & Dalla-Favera, R. (2013). tRNA-derived microRNA modulates proliferation and the DNA damage response and is down-regulated in B cell lymphoma. *Proc Natl Acad Sci U S A, 110*(4), 1404-1409. doi:10.1073/pnas.1206761110

Mercatelli, N., Coppola, V., Bonci, D., Miele, F., Costantini, A., Guadagnoli, M., . . . Ciafre, S. A. (2008). The inhibition of the highly expressed miR-221 and miR-222 impairs the growth of prostate carcinoma xenografts in mice. *PLoS One, 3*(12), e4029. doi:10.1371/journal.pone.0004029

Musacchio, M., & Perrimon, N. (1996). The Drosophila kekkon genes: novel members of both the leucine-rich repeat and immunoglobulin superfamilies expressed in the CNS. *Dev Biol, 178*(1), 63-76. doi:10.1006/dbio.1996.0198

Nakahara, K., Kim, K., Sciulli, C., Dowd, S. R., Minden, J. S., & Carthew, R. W. (2005). Targets of microRNA regulation in the Drosophila oocyte proteome. *Proc Natl Acad Sci U S A, 102*(34), 12023-12028. doi:10.1073/pnas.0500053102

Neuman-Silberberg, F. S., & Schupbach, T. (1993). The Drosophila dorsoventral patterning gene gurken produces a dorsally localized RNA and encodes a TGF alpha-like protein. *Cell, 75*(1), 165-174. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/7691414

Neuman-Silberberg, F. S., & Schupbach, T. (1994). Dorsoventral axis formation in Drosophila depends on the correct dosage of the gene gurken. *Development, 120*(9), 2457-2463. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/7956825

Niepielko, M. G., & Yakoby, N. (2014). Evolutionary changes in TGFalpha distribution underlie morphological diversity in eggshells from Drosophila species. *Development, 141*(24), 4710-4715. doi:10.1242/dev.111898

Peng, H., Shi, J., Zhang, Y., Zhang, H., Liao, S., Li, W., . . . Duan, E. (2012). A novel class of tRNA-derived small RNAs extremely enriched in mature mouse sperm. *Cell Res, 22*(11), 1609-1612. doi:10.1038/cr.2012.141

Ray, R. P., & Schupbach, T. (1996). Intercellular signaling and the polarization of body axes during Drosophila oogenesis. *Genes Dev, 10*(14), 1711-1723. doi:10.1101/gad.10.14.1711

Reid, G., Kao, S. C., Pavlakis, N., Brahmbhatt, H., MacDiarmid, J., Clarke, S., . . . van Zandwijk, N. (2016). Clinical development of TargomiRs, a miRNA mimic-based treatment for patients with recurrent thoracic cancer. *Epigenomics, 8*(8), 1079-1085. doi:10.2217/epi-2016-0035

Revaitis, N. T., Marmion, R. A., Farhat, M., Ekiz, V., Wang, W., & Yakoby, N. (2017). Simple Expression Domains Are Regulated by Discrete CRMs During Drosophila Oogenesis. *G3 (Bethesda), 7*(8), 2705-2718. doi:10.1534/g3.117.043810

Rogers, H. H., Bergman, C. M., & Griffiths-Jones, S. (2010). The evolution of tRNA genes in Drosophila. *Genome Biol Evol, 2*, 467-477. doi:10.1093/gbe/evq034

Schimmel, P. (2018). The emerging complexity of the tRNA world: mammalian tRNAs beyond protein synthesis. *Nat Rev Mol Cell Biol, 19*(1), 45-58. doi:10.1038/nrm.2017.77

Shcherbata, H. R. (2019). miRNA functions in stem cells and their niches: lessons from the Drosophila ovary. *Curr Opin Insect Sci, 31*, 29-36. doi:10.1016/j.cois.2018.07.006

Shen, L., Gan, M., Tan, Z., Jiang, D., Jiang, Y., Li, M., . . . Zhu, L. (2018). A Novel Class of tRNA-Derived Small Non-Coding RNAs Respond to Myocardial Hypertrophy and Contribute to Intergenerational Inheritance. *Biomolecules, 8*(3). doi:10.3390/biom8030054

Shen, Y., Yu, X., Zhu, L., Li, T., Yan, Z., & Guo, J. (2018). Transfer RNA-derived fragments and tRNA halves: biogenesis, biological functions and their roles in diseases. *J Mol Med (Berl), 96*(11), 1167-1176. doi:10.1007/s00109-018-1693-y

Shin, C., Nam, J. W., Farh, K. K., Chiang, H. R., Shkumatava, A., & Bartel, D. P. (2010). Expanding the microRNA targeting code: functional sites with centered pairing. *Mol Cell, 38*(6), 789-802. doi:10.1016/j.molcel.2010.06.005

Skeparnias, I., Anastasakis, D., Grafanaki, K., Kyriakopoulos, G., Alexopoulos, P., Dougenis, D., . . . Stathopoulos, C. (2020). Contribution of miRNAs, tRNAs and tRFs to Aberrant Signaling and Translation Deregulation in Lung Cancer. *Cancers (Basel), 12*(10). doi:10.3390/cancers12103056

Sobala, A., & Hutvagner, G. (2011). Transfer RNA-derived fragments: origins, processing, and functions. *Wiley Interdiscip Rev RNA, 2*(6), 853-862. doi:10.1002/wrna.96

Spradling, A., 1993. Developmental genetics of oogenesis. In: Bate, M., Martinez-Arias, A. (Eds.), The Development of Drosophila melanogaster. Cold Spring Harbor Laboratory Press, New York, pp. 1–70.

Stefani, G., & Slack, F. J. (2008). Small non-coding RNAs in animal development. *Nat Rev Mol Cell Biol, 9*(3), 219-230. doi:10.1038/nrm2347

Tuck, A. C., & Tollervey, D. (2011). RNA in pieces. *Trends Genet, 27*(10), 422-432. doi:10.1016/j.tig.2011.06.001

van Bakel, H., Nislow, C., Blencowe, B. J., & Hughes, T. R. (2010). Most "dark matter" transcripts are associated with known genes. *PLoS Biol, 8*(5), e1000371. doi:10.1371/journal.pbio.1000371

Venkatesh, T., Suresh, P. S., & Tsutsumi, R. (2016). tRFs: miRNAs in disguise. *Gene, 579*(2), 133-138. doi:10.1016/j.gene.2015.12.058

Vicario, S., Moriyama, E. N., & Powell, J. R. (2007). Codon usage in twelve species of Drosophila. *BMC Evol Biol, 7*, 226. doi:10.1186/1471-2148-7-226

Wang, J., Tan, L., Tan, L., Tian, Y., Ma, J., Tan, C. C., . . . Yu, J. T. (2015). Circulating microRNAs are promising novel biomarkers for drug-resistant epilepsy. *Sci Rep, 5*, 10201. doi:10.1038/srep10201

Wang, Q., Lee, I., Ren, J., Ajay, S. S., Lee, Y. S., & Bao, X. (2013). Identification and functional characterization of tRNA-derived RNA fragments (tRFs) in respiratory syncytial virus infection. *Mol Ther, 21*(2), 368-379. doi:10.1038/mt.2012.237

Warner, J. R. (1999). The economics of ribosome biosynthesis in yeast. *Trends Biochem Sci, 24*(11), 437-440. doi:10.1016/s0968-0004(99)01460-7

Wasserman, J. D., & Freeman, M. (1998). An autoregulatory cascade of EGF receptor signaling patterns the Drosophila egg. *Cell, 95*(3), 355-364. doi:10.1016/s0092-8674(00)81767-5

Yakoby, N., Bristow, C. A., Gouzman, I., Rossi, M. P., Gogotsi, Y., Schupbach, T., & Shvartsman, S. Y. (2005). Systems-level questions in Drosophila oogenesis. *Syst Biol (Stevenage), 152*(4), 276-284. doi:10.1049/ip-syb:20050039

Zartman, J. J., Cheung, L. S., Niepielko, M. G., Bonini, C., Haley, B., Yakoby, N., & Shvartsman, S. Y. (2011). Pattern formation by a moving morphogen source. *Phys Biol, 8*(4), 045003. doi:10.1088/1478-3975/8/4/045003

Zartman, J. J., Kanodia, J. S., Cheung, L. S., & Shvartsman, S. Y. (2009). Feedback control of the EGFR signaling gradient: superposition of domain-splitting events in Drosophila oogenesis. *Development, 136*(17), 2903-2911. doi:10.1242/dev.039545