# Programmed Translational Readthrough in *Drosophila melanogaster*

Dissertation

For the award of the degree

"Doctor rerum naturalium" (Dr. rer. nat.)

of the Georg-August-Universität Göttingen

within the doctoral program IMPRS Molecular Biology

of the Georg-August University School of Science (GAUSS)

submitted by

Prajwal Karki

from Kathmandu, Nepal

Göttingen, 2019

Thesis Committee

Prof. Dr. Marina V. Rodnina
Department of Physical Biochemistry
Max Planck Institute for Biophysical Chemistry
Göttingen, Germany

Prof. Dr. Halyna Shcherbata
Institute of Cell Biochemistry
Hannover Medical School
Hannover, Germany

Prof. Dr. Heike Krebber
Department of Molecular Genetics
Institute for Microbiology and Genetics
Göttingen, Germany

Members of the Examination Board

Prof. Dr. Marina V. Rodnina (Referee)
Department of Physical Biochemistry
Max Planck Institute for Biophysical Chemistry
Göttingen, Germany

Prof. Dr. Halyna Shcherbata (2nd Referee)
Institute of Cell Biochemistry
Hannover Medical School
Hannover, Germany

Further members of the Examination Board

Prof. Dr. Wolfgang Wintermeyer
Department of Physical Biochemistry
Max Planck Institute for Biophysical Chemistry
Göttingen, Germany

Prof. Dr. Herbert Jäckle
Department of Molecular Developmental Biology
Max Planck Institute for Biophysical Chemistry
Göttingen, Germany

Dr. Alexis Caspar Faesen
Department of Biochemistry of Signal Dynamics
Max Planck Institute for Biophysical Chemistry
Göttingen, Germany

Date of the oral examination: 12th June, 2019

# Affidavit

I hereby declare that the presented thesis entitled “Programmed Translational Readthrough in *Drosophila melanogaster*” has been written independently and with no other sources and aids than quoted.

Göttingen, April 30th, 2019

Prajwal Karki

## Related Publications


1. Beißel, C., Neumann, B., Uhse, S., Hampe, I., Karki, P. and Krebber, H., 2019. Translation termination depends on the sequential ribosomal entry of eRF1 and eRF3. *Nucleic Acids Research*, in press.

2. Adio, S., Sharma, H., Senyushkina, T., Karki, P., Maracci, C., Wohlgemuth, I., Holtkamp, W., Peske, F. and Rodnina, M.V., 2018. Dynamics of ribosomes and release factors during translation termination in *E. coli*. *eLife*, *7*, p.e34252.

3. Florin, T., Maracci, C., Graf, M., Karki, P., Klepacki, D., Berninghausen, O., Beckmann, R., Vázquez-Laslop, N., Wilson, D.N., Rodnina, M.V. and Mankin, A.S., 2017. An antimicrobial peptide that inhibits translation by trapping release factors on the ribosome. *Nature Structural & Molecular Biology*, *24*(9), p.752.

# Table of contents

## ABSTRACT

Translational readthrough (TR) is extensively used by viruses to expand their limited genome capacity. The pervasiveness of TR across eukaryotic genomes is only recently being explored. Systems biology approaches such as comparative phylogenetics, combined with ribosome profiling studies have led to the identification of several hundred genes that can undergo TR in *Drosophila melanogaster*. However, only limited number of these genes have been experimentally studied. Due to the lack of systemic biochemical studies, the actual mechanism of TR regulation and the role of regulated TR in proteome expansion in *D. melanogaster* remains largely unexplored.

In this study, we develop a dual luciferase reporter assay system to analyze a set of genes from *D. melanogaster* that have been phylogenetically predicted to undergo TR. We provide experimental validation and quantification of the extent of TR in these genes and also analyze the impact of immediate stop codon context on TR modulation. For one particular candidate, *aPKC*, we use systematic mutational analysis to delineate the minimal primary sequence motif responsible for driving efficient TR. Furthermore, in order to understand the biological significance and phenotypic outcome of gene-specific TR, we employ CRISPR/Cas9-based genome editing to create genetic mutants of the large Maf transcription factor, *traffic jam* (*tj*), that exhibit constitutive TR or abolished it. We identify tissue-specific regulation of TR in *tj*, wherein the expression of Tj-TR isoform is restricted to the nervous tissues. The TR extension selectively attenuates the native Tj function, specifically in the determination of morphogenetic behavior and spatial distribution of cap cells in adult ovaries. Conversely, the TR extension positively amplifies the selective gene regulatory function of native Tj in downregulation of the adhesion protein *Fasciclin III*. Using high-throughput RNA sequencing, we further dissect the role of TR in *tj* in shaping the transcriptome profile in adult brains. Our results hint towards a complex mode of regulation of Tj function by TR that operates via conditional fine-tuning of its specific gene regulatory functions.

# 1. INTRODUCTION

The flow of genetic information from DNA to RNA to protein governs the cellular basis of life. The information present in the DNA is transcribed into RNA by RNA polymerases. Some RNAs can further propagate the information as coding molecules called messenger RNAs (mRNAs). Other non-coding RNAs act as functional components of the ribosome, the protein synthesizing machinery in a cell (ribosomal RNA, rRNA), or as adapter molecules (transfer RNA, tRNAs). The ribosome scans the information encoded in mRNAs as triplet codons and translates it into a sequence of polypeptide with the help of tRNAs carrying amino acids.

The ribosome follows strict rules of decoding by scrutinizing the complementarity between the codons in the mRNA and the anticodons of the tRNAs, which dictates the fidelity of translation. However, non-canonical decoding events might take place that defy the conventional genetic code resulting in the production of altered protein products. At times, such reprogramming aids in the fitness of the organism by allowing an extra step of post-transcriptional regulation as well as expanding the coding capacity of the genome. Recoding events have garnered great scientific interest owing to their implication in clinically relevant genetic disorders.

## 1.1 Protein synthesis

### 1.1.1 The translational machinery

Translation is the final step of gene expression. The key player of translation is the ribosome, a large ribonucleoprotein complex consisting of two unequal subunits. In bacteria, the large 50S subunit (LSU) (S, Svedberg unit which denotes the rate of sedimentation) and the small 30S subunit (SSU) constitute the 2.5 megadalton (MDa) 70S ribosome. In eukaryotes, the 80S ribosome is composed of a 60S and a 40S subunit. The size of eukaryotic ribosomes ranges from 3.3 MDa in lower eukaryotes to 4.3 MDa in higher eukaryotes (Melnikov et al., 2012).

Each ribosomal subunit is composed of highly structured rRNA and a set of ribosomal proteins (Fig. 1). The interface between the ribosomal subunits defines three stable tRNA binding sites: the acceptor site (A site), which accepts the incoming aminoacylated tRNA (aa-tRNA), the peptidyl site (P site), which houses the tRNA bearing the growing polypeptide chain, and the exit site (E site), which holds the deacylated tRNA. The functional centers of the ribosome, the decoding site in the SSU and the peptidyl transferase center (PTC) in the LSU, are mostly composed of rRNA. The decoding center plays an important role in ensuring translational fidelity by recognizing the Watson-Crick base pairing between the three bases of the codon in mRNA at the A site with the anticodon bases of aa-tRNA. In order to adapt to the degeneracy of the genetic code, the third position (wobble position) of the codon can accept non-Watson-Crick pairing. The formation of the correct

codon-anticodon interaction leads to conformational changes in the decoding center that are crucial for discrimination between near-cognate and cognate tRNAs (Battle and Doudna, 2002; Ogle et al., 2001; Ogle et al., 2002; Rodnina et al., 2017; Saint-Leger and Ribas de Pouplana, 2015).

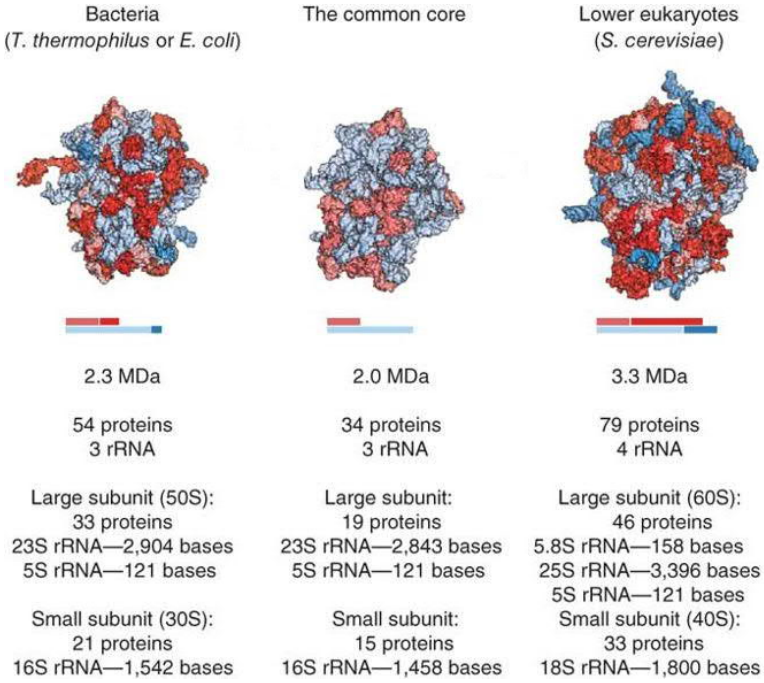


Figure 1. Conserved common core of bacterial and eukaryotic ribosomes.
The architecture of bacterial and eukaryotic ribosomes with conserved rRNA core (light blue) and conserved proteins (light red) are depicted. Conserved rRNA and protein domains specific to each kingdom are highlighted in red and blue, respectively. The molecular weight, protein and RNA composition of LSU and SSU for bacterial as well as eukaryotic ribosomes are listed below the structures. The figure is adapted from Melnikov et al. (2012).

### 1.1.2 The translation cycle

The cycle of translation can be divided into four stages: initiation, elongation, termination and recycling (Fig. 2). Translation initiation is the rate-limiting step of protein synthesis and is strictly regulated. In prokaryotes, three initiation factors (IF1, IF2 and IF3) act together to recruit the mRNA and the initiator tRNA, fMet-tRNA$^{fMet}$, to the P site of the SSU (Milon and Rodnina, 2012). Initiation in eukaryotes is more complex and involves a minimum of eleven different initiation factors (eIFs) (Rodnina and Wintermeyer, 2009). Eukaryotic initiation starts with the recruitment of a ternary complex (TC) composed of Met-tRNA$^{Met}$, eIF2 and GTP to the SSU with the help of eIF1, eIF1A and eIF3 to form a 43S preinitiation complex (43S PIC) (Aylett et al., 2015; Hashem et al., 2013). Binding of these factors induces an open state of the SSU, which facilitates the binding of TC (Aitken and Lorsch, 2012). The cap-binding complex formed by eIF4F, eIF4A and eIF4B, bound at the 7-methylguanosine-capped 5' end of the mRNA, recruits the 43S PIC (Jackson et al., 2010). The poly(A)-binding protein (PABP) binds to the 3' poly(A) tail of mRNA and circularizes it, synergistically promoting translation (Munroe and Jacobson, 1990; Nicholson and Pasquinelli, 2018). This

complex, in a conformation with partially accommodated Met-tRNA$^{Met}$, scans the mRNA until it encounters the start codon AUG. AUG recognition induces a conformational change of the SSU head relative to its body leading to complete accommodation of Met-tRNA$^{Met}$ in the P site (closed state), thus forming the 48S initiation complex (IC) (Hinnebusch and Lorsch, 2012; Llacer et al., 2015). eIF5 binds to eIF2 and stimulates the GTPase activity of the γ-subunit of eIF2 (Paulin et al., 2001), which results in dissociation of both factors. eIF5B binding to the complex promotes the joining of the LSU as well as the dissociation of eIF1, eIF1A and eIF3 to form the 80S complex (Pestova et al., 2000). GTP hydrolysis by eIF5B is required for its release from the 80S complex which is then committed to the elongation step (Lee et al., 2002).

Translation elongation is a cyclic process, in which the three steps of aa-tRNA selection (decoding), peptide bond formation and translocation repeat until the ribosome encounters a stop codon on the mRNA (Fig. 2). Translation elongation is highly conserved across all kingdoms of life and is mediated by two elongation factors (EFs): eEF1A and eEF2 in eukaryotes, and the functional homologs EF-Tu and EF-G in prokaryotes, respectively. EF-Tu/eEF1a forms a tight ternary complex with aa-tRNA and GTP and delivers aa-tRNA to the A site of the ribosome. The cognate codon-anticodon base pairing between the mRNA and the aa-tRNA triggers EF-Tu/eEF1a to hydrolyze GTP (Rodnina et al., 1995). This results in the release of EF-Tu/eEF1a (Rodnina, 2012), which allows the aa-tRNA to accommodate into the PTC in the LSU (Dever et al., 2016; Pape et al., 1998; Schuller and Green, 2018). The amino group of the aa-tRNA in the A site attacks the ester bond on the peptidyl-tRNA in the P site to form the peptide bond (Beringer and Rodnina, 2007). The ribosome undergoes an inter-subunit ratcheting motion to adopt a 'hybrid' state (Behrmann et al., 2015; Budkevich et al., 2011; Frank and Agrawal, 2000). EF-G/eEF2 facilitates the translocation of the ribosome to the next codon (Ferguson et al., 2015; Ling and Ermolenko, 2016; Taylor et al., 2007). The deacylated tRNA is released from the E site and the complex is ready to undergo subsequent cycles of elongation.

The cycle of elongation continues until the ribosome encounters one of the universal stop codons UAG, UAA or UGA, which triggers the termination of protein synthesis. Translation termination involves recognition of the stop codon, peptide hydrolysis and dissociation of release factors (RFs). Recognition of stop codons is mediated by class I release factors (RF1 and RF2 in bacteria and eRF1 in eukaryotes) that act as tRNA mimics. RF1 recognizes UAG and UAA, RF2 recognizes UGA and UAA (Freistroffer et al., 2000; Kisselev et al., 2003), while eRF1 recognizes all three stop codons (Bertram et al., 2000; Dever and Green, 2012). Class I RFs catalyze the hydrolysis of the ester bond of the peptidyl-tRNA with the help of a universally conserved GGQ motif that reaches the PTC (Frolova et al., 1999; Jin et al., 2010; Laurberg et al., 2008; Seit-Nebi et al., 2001)

and shields nucleophiles larger than water (Shaw and Green, 2007). Class II RFs (RF3 in prokaryotes and eRF3 in eukaryotes) are GTPases with little homology limited to their GTP binding domains (Kisselev and Buckingham, 2000). While RF3 accelerates the dissociation of RF1 and RF2 from posttermination complexes (postTCs) (Adio et al., 2018; Koutmou et al., 2014; Pallesen et al., 2013; Peske et al., 2014; Shi and Joseph, 2016; Zavialov et al., 2001; Zavialov et al., 2002), eRF3 facilitates binding of eRF1 to pretermination complexes (preTCs) by forming a ternary complex eRF1-eRF3-GTP (Dever and Green, 2012; Frolova et al., 1996; Mitkevich et al., 2006; Pisareva et al., 2006).

The final step of translation is ribosome recycling, which allows the use of ribosomal subunits for subsequent rounds of translation. In prokaryotes, the ribosome recycling factor (RRF) binds to the A site of the ribosome, mediating the recruitment of EF-G (Gao et al., 2005). GTP hydrolysis by EF-G and Pi release result in the splitting of the ribosome into its subunits. Binding of IF3 promotes dissociation of the tRNA from the SSU and provides the first step of the new round of translation (Peske et al., 2005; Savelsbergh et al., 2009; Seo et al., 2004). In eukaryotes, the highly conserved ATPase, ABCE1 mediates recycling (Franckenberg et al., 2012; Khoshnevis et al., 2010; Pisarev et al., 2010); ribosome splitting depends on the presence of eRF1 in the A site and dissociation of eRF3 (Pisarev et al., 2010). ABCE1 also promotes eRF1-mediated peptide hydrolysis in an ATP-independent manner (Shoemaker and Green, 2011). ABCE1 seems to serve additional roles in regulating initiation (Mancera-Martinez et al., 2017). The deacylated tRNA and the mRNA that are bound to the SSU are released by eIF1, eIF1A and eIF3 along with its weakly associated eIF3j subunit (Fraser et al., 2007; Pisarev et al., 2007; Pisarev et al., 2010).

Efficient translation relies on the coordinated dynamics of the ribosome and of the translation factors. The major conformational rearrangements during translation involve the relative rotation of the ribosomal subunits, the swiveling of the head of the SSU, the movement of tRNAs and of ribosomal protein L1. These motions aid the ribosome to progress through different steps of translation and are crucial for the maintenance of translational accuracy and processivity (Adio et al., 2015; Belardinelli et al., 2016; Frank and Gonzalez, 2010; Ling and Ermolenko, 2015; Myasnikov et al., 2005; Rodnina et al., 2017; Sharma et al., 2016).

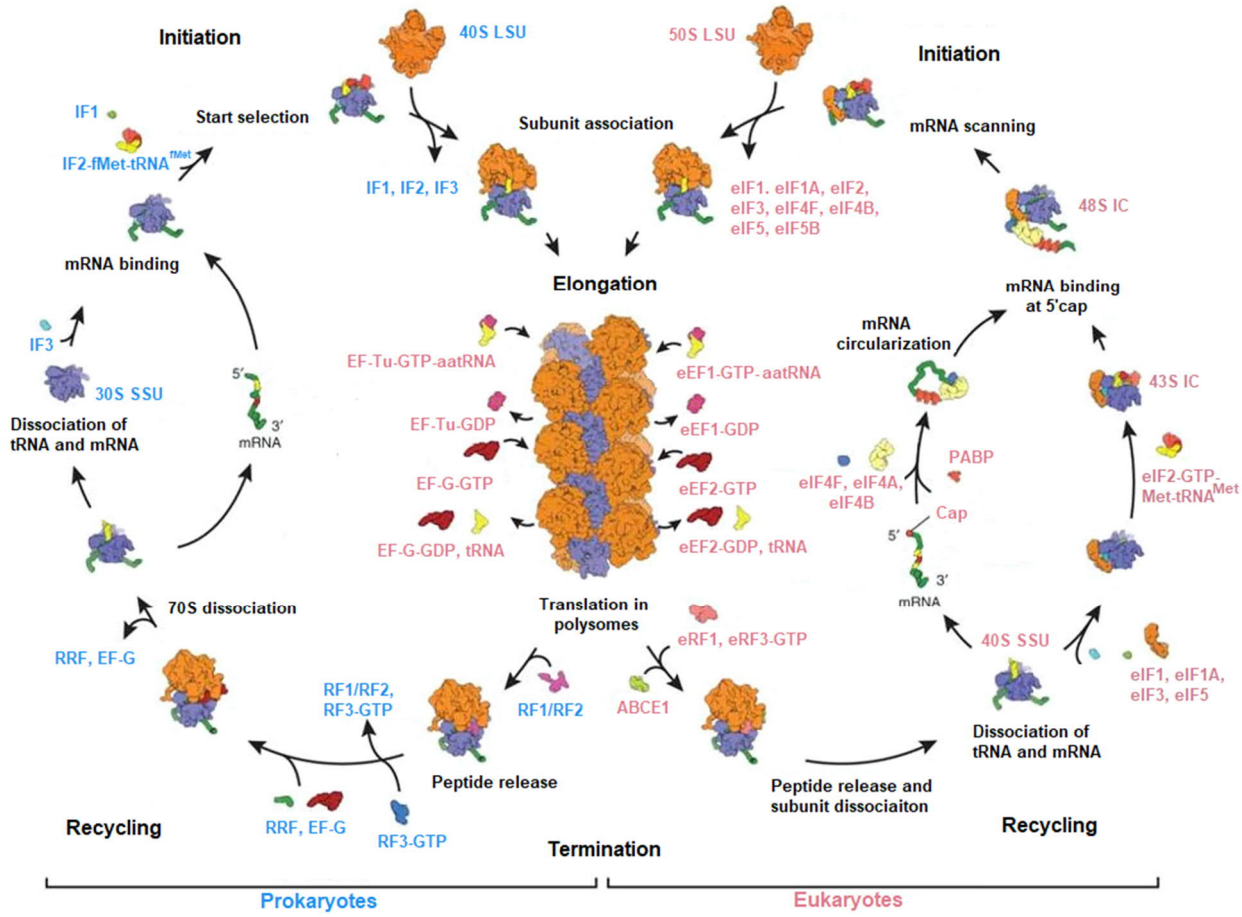


Figure 2. Scheme of prokaryotic and eukaryotic translation cycles.
The key steps of the translation cycle: initiation, elongation, termination and recycling are depicted. The protein factors that mediate each step of the process are shown. The figure is adapted from (Melnikov et al., 2012)

### 1.1.3 Prokaryotic translation termination

Translation termination in prokaryotes utilizes the dynamic properties of the ribosome, induced by binding of the release factors to navigate through loosely coupled motions in order to achieve peptide hydrolysis and release. Structural and smFRET studies have shown that binding of RF1 stabilizes the non-rotated state of the ribosome (Laurberg et al., 2008; Petry et al., 2005; Weixlbaumer et al., 2008) while RF2 bound preTCs are dynamic with higher fraction of rotated states (Adio et al., 2018). RF1 and RF2 recognize their respective stop codons with the help of the conserved PVT and SPF motifs in domain 2 respectively (Korostelev et al., 2008; Korostelev et al., 2010; Laurberg et al., 2008; Weixlbaumer et al., 2008). When bound to the A site, RF1 and RF2 adopt an open conformation where the GGQ motif extends into the PTC and catalyzes the hydrolysis of the peptidyl-tRNA (Laurberg et al., 2008; Petry et al., 2005; Weixlbaumer et al., 2008). The role of RF3 in termination and its mechanism of action is controversial. Binding of RF3-GTP alone induces the rotated state of the ribosome (Gao et al., 2007; Jin et al., 2011; Sternberg et al.,

2009). Previous studies highlighted the role of RF3 in the GTP dependent release of RF1 and RF2 (Koutmou et al., 2014; Peske et al., 2014; Shi and Joseph, 2016; Zavialov et al., 2001; Zavialov et al., 2002) while recently, it has been shown that RF2 is less dependent of RF3 for dissociation (Adio et al., 2018). Comparable affinity of RF3 towards GTP and GDP suggests that, at cellular concentrations, RF3 exists predominantly in GTP bound state (Koutmou et al., 2014; Peske et al., 2014). Recently, we have shown that the recruitment of RFs to the preTCs follows a stochastic order: RF1 can bind to preTCs and catalyze peptide hydrolysis, after which RF3-GTP is recruited or both factors can be recruited simultaneously after which peptide is hydrolyzed (Adio et al., 2018). RF3-induced subunit rotation as well as conformational adjustments conferred by peptide release jointly accelerate the dissociation of RF1 and RF3-GTP from the postTCs. Overall, translation termination in prokaryotes is non-deterministic and driven by multiple steps of subunit rotation, peptide hydrolysis, conformational rearrangements and GTP hydrolysis that exhibit minimal dependence on kinetic coupling (Fig. 3).

Profound differences with the eukaryotic mechanism (see section 1.1.4) make the bacterial termination step an optimal target for drug development. Only one inhibitor has been characterized so far that targets bacterial termination complexes. Apidaecin 37 (Api37) belongs to the proline-rich antimicrobial peptides (PrAMPs), a class of natural peptides that targets the translational machinery to inhibit bacterial growth (Gagnon et al., 2016; Li et al., 2014). While most PrAMPs bind to the peptide exit tunnel and prevent aa-tRNA binding (Roy et al., 2015b; Seefeldt et al., 2016); Api37 traps RF1 and RF2 on the ribosome after the hydrolysis of the nascent peptide (Florin et al., 2017). RF3 is unable to facilitate the recycling of RF1 or RF2 trapped on the ribosome, which results in the depletion of free release factors in the cell such that the translating ribosomes are stalled at the stop codon.

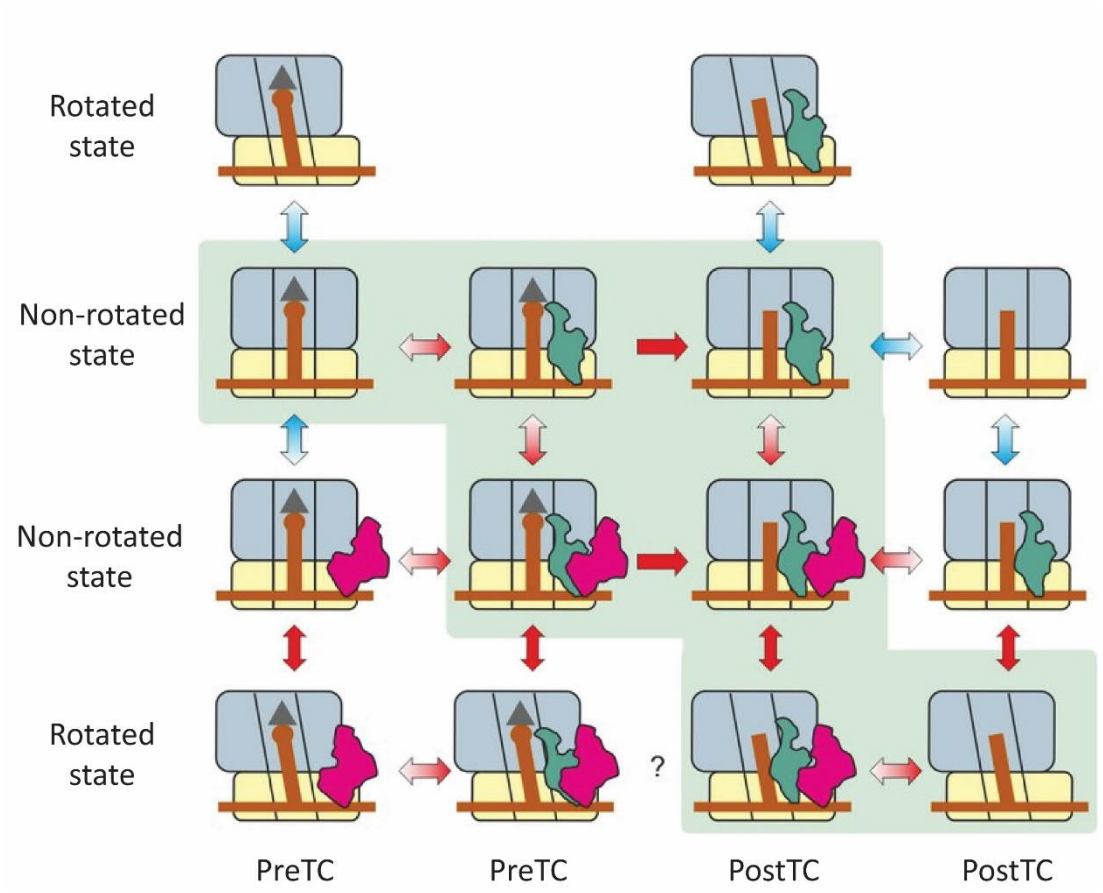


Figure 3. Model of prokaryotic translation termination
The conformation of ribosome in the rotated state and the non-rotated state are indicated. RF1 is depicted in green, RF3 in red and nascent peptide chain in grey (triangle). Red arrows indicate fast reactions, blue arrows indicate static reactions while single headed arrows indicate irreversible step of peptide hydrolysis. The color gradient in the arrows depicts the preferred direction of reaction. The figure is adapted from Adio et al. (2018).

### 1.1.4 Translation termination and stop codon recognition in eukaryotes

Eukaryotic termination largely differs from prokaryotic termination in the sequence of binding of RFs, interactions between class I and class II RFs, mechanism of stop codon recognition and role of GTP hydrolysis by class II RF.

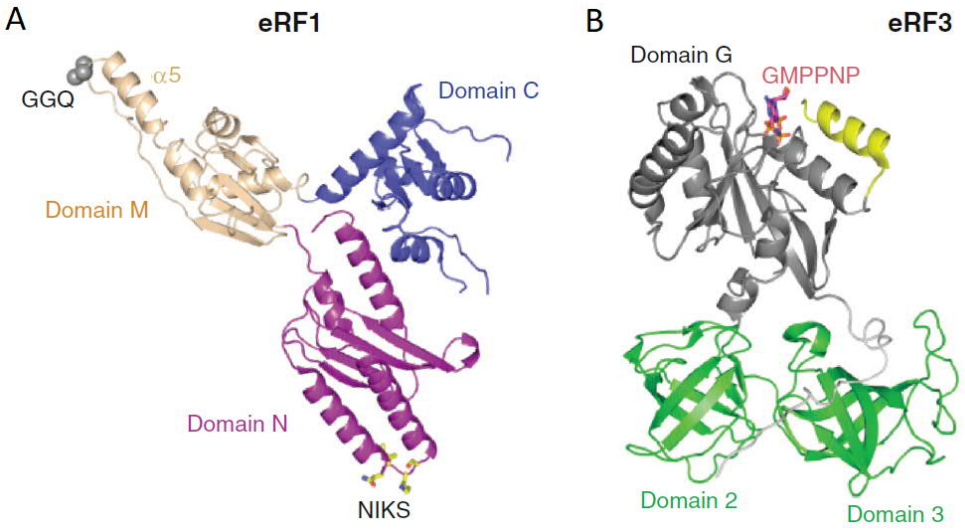


Figure 4. Structures of eRF1 and eRF3.
(A) Ribbon diagram of human eRF1 with GGQ motif in domain M and NIKS motif in domain N highlighted (PDB 1DT9). (B) *S. pombe* eRF3 (215-662) with GMPPNP in stick model (PDB 1R5B). The figure is adapted from Jackson et al. (2012).

eRF1 is composed of three domains (Fig. 4A). The N-terminal domain binds to the decoding center of the ribosome and recognizes the stop codon (Bertram et al., 2000; Brown et al., 2015; Preis et al., 2014). The middle domain contains the universally conserved GGQ motif that extends into the PTC and catalyzes the hydrolysis of the nascent polypeptide chain (Frolova et al., 1999; Song et al., 2000). The C-terminal domain interacts with eRF3 as well as ABCE1 (Preis et al., 2014; Zhouravleva et al., 1995). eRF3 consists of an N-terminal domain that is dispensable for its function in termination but is able to interact with PABP (Kozlov and Gehring, 2010) and with UPF3B, which is implicated in the non-sense mediated decay pathway machinery (Neu-Yilik et al., 2017) (Fig. 4B). The functional C-terminal region consists of the GTP-binding domain (domain G) and two β-barrel domains that share significant homology with translational GTPases such as EF-Tu and eEF1A (Andersen et al., 2000; Kong et al., 2004; Song et al., 1999).

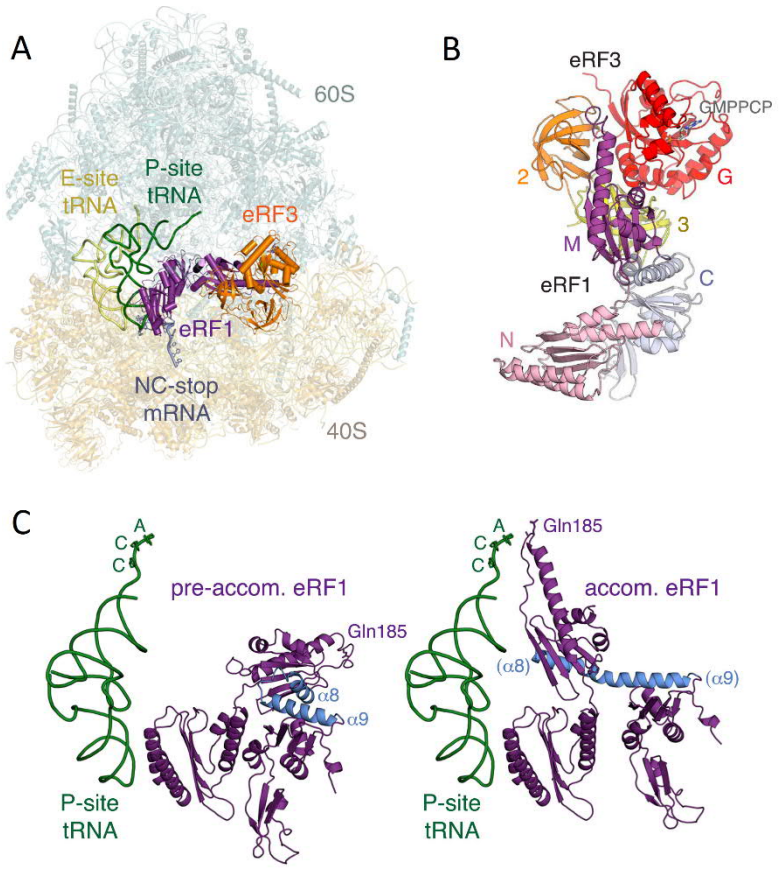


Figure 5. Overview of the conformations attained by eRF1 and eRF3 during termination.
(A) PreTC assembled with eRF1 (purple) and eRF3 (orange). (B) The middle domain (M) of eRF1 (purple) is tucked between the G domain (orange) and domain 2 (yellow) of eRF3 in the pre-accommodated state. (C) eRF1 (purple) in pre-accommodated state (left) and post-accommodated state (right) where the middle domain undergoes a 140° rotation to position the GGQ motif in the PTC. The image is adapted from Shao et al. (2016).

eRF1 and eRF3 interact with each other via their C-terminal domains (Cheng et al., 2009; Ito et al., 1998; Kononenko et al., 2008; Merkulova et al., 1999) and enhance their translational activities. The stable binding occurs in solution, irrespective of the presence of the ribosomal machinery. eRF1 stabilizes the binding of GTP to eRF3 by lowering its dissociation rate constant and forming a stable eRF1-eRF3-GTP ternary complex (Hauryliuk et al., 2006; Mitkevich et al., 2006; Pisareva et al., 2006). The ternary complex binds to the empty A site of the preTC (Fig. 5A), where eRF1 is positioned in a pre-accommodation state until GTP is hydrolyzed by eRF3 (Fig. 5C). In this conformation, the middle domain of eRF1 is tucked into the cleft between the G-domain and domain 2 of eRF3 (Fig. 5B), such that the catalytic GGQ motif is located >80 Å away from the ester bond of the P site tRNA in the PTC. eRF3 is docked to the GTPase associated center of the ribosome, between the sarcin-ricin loop of the 60S subunit and helices h5 and h14 of the 18S rRNA on the SSU (des Georges et al., 2014; Preis et al., 2014; Shao et al., 2016).

Upon binding, the N-domain of eRF1 creates a tight pocket that accommodates the stop codon as well as the following nucleotide (+4) in a geometry that resembles an RNA U-turn motif (Fig. 6A) (Brown et al., 2015; Matheisl et al., 2015; Shao et al., 2016). Induction of such conformation is a distinct property of the eukaryotic translation machinery, and provides a basis for discrimination against sense codons. The mRNA compaction depends on the flipping out of base A1825 in h44 of the 18S rRNA, which allows stacking with the second base of the stop codon that in turn stacks on the third base (Fig. 6B). This configuration allows stacking of the +4 base with G626 of the 18S rRNA, which leads to compaction of the mRNA (Fig. 6B), in agreement with toeprinting data (Alkalaeva et al., 2006). The interaction of G626 with the +4 nucleotide is favored in case of purines, which also supports the significant bias for A and G at the position +4 (Brown et al., 1990a).

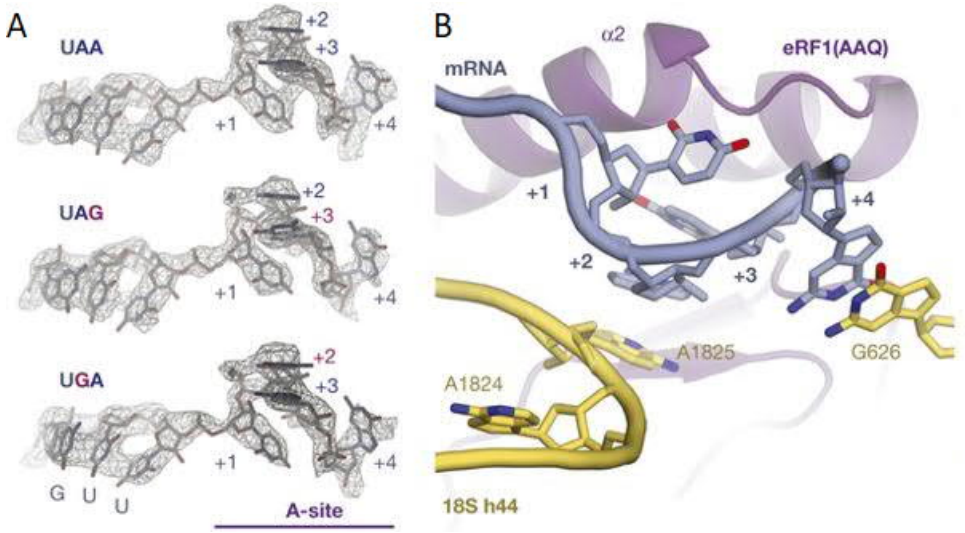


Figure 6. Stop codon configuration in the decoding center.
(A) Cryo-EM densities of the mRNA in mammalian termination complexes containing UAA, UAG and UGA stop codons. The +1 to +3 bases of the stop codons and the following +4 base are indicated. (B) Interactions formed by the stop codon at the decoding center. eRF1 (AAQ) recognizes +1 to +4 bases (grey) in the A site. A1825 of h44 stacks with the +2 and +3 bases of the stop codon while +4 base stacks with G626 of 18S rRNA. Image is adapted from Brown et al. (2015).

The N-domain of eRF1 reaches deep into the decoding center where it establishes contacts with the components of the ribosomal complex that govern the recognition and decoding of the stop codon. The TAS-NIKS (residues 58-64), YxCxxxF (residues 125-131), E(55) and the GTS (31-33) motifs (human eRF1 numbering) are crucial for stop codon recognition (Blanchet et al., 2015; Bulygin et al., 2010; Chavatte et al., 2002). The invariant +1 U is recognized by residues in the TASNIKS motif (Fig. 7A), whereas only purines in the +2 and +3 position can interact with residues of the YxCxxF motif and E55 (Fig. 7B) (Brown et al., 2015). Finally, the conformation adopted by the GTS motif explains how UAG and UGA codons are recognized via discrete interactions. In the case of UAG, T32 faces the +3 base and forms hydrogen bond with N2 of guanosine. The +2 G in UGA is accommodated by a movement of the YxCxxxF motif that is relayed into a 4 Å movement of GTS

motif such that T32 faces away from the stop codon and does not interact with +3 base. The UGG codon for tryptophan is discriminated against by steric repulsion between the O6 atoms of G with E55 of eRF1. The differences in interactions of eRF1 with the three stop codons form the basis for the fidelity of stop codon recognition (Fig. 7C).

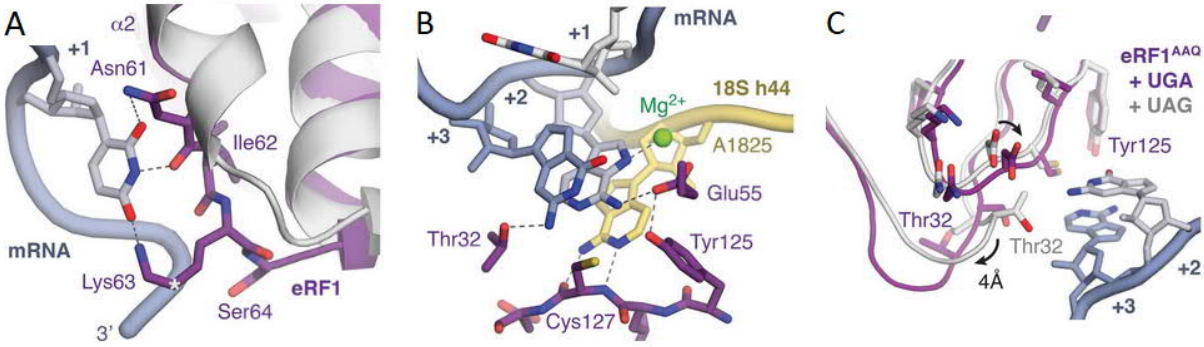


Figure 7. Interactions between eRF1 and stop codon.
(A) Hydrogen bonds between +1 U of the stop codon with the NIKS motif of preTC bound eRF1 (purple) compared to eRF1 crystal structure (grey) (PDB 1DT9). (B) Interactions between UAG stop codon with YxCxxxF, E55 and T32 motifs of eRF1. (C) Differences in T32 interaction between eRF1 bound to UGA (purple) compared to eRF1 bound to UAG (grey). Adapted from (Brown et al., 2015).

The GTPase activity of eRF3 is ribosome-dependent, and is accelerated by stop codon recognition (Hellen, 2018) (Fig. 8). Following GTP hydrolysis, the middle domain of eRF1 undergoes a 140° rotation relative to the N-domain thereby assuming an extended conformation and positioning the catalytic GGQ motif in the PTC of the preTC (Matheisl et al., 2015; Shao et al., 2016) (Fig. 5C). Although eRF1 can perform peptide hydrolysis alone, its activity is strongly enhanced by eRF3 (Alkalaeva et al., 2006; Eyler et al., 2013), either due to a more efficient recruitment of eRF1 or an enhanced rate of peptide hydrolysis. The GTPase activity of eRF3 couples stop codon recognition and peptidyl-tRNA hydrolysis by eRF1 and also increases the kinetic fidelity of termination by introducing an irreversible step (Hellen, 2018; Salas-Marco and Bedwell, 2004). The DEAD-box RNA helicase Dbp5 is known to genetically interact with the release factors and modulate the efficiency of termination in eukaryotes (Gross et al., 2007). Its role in in stepwise assembly of termination complex at the stop codon and prevention of premature dissociation of eRF1 and eRF3 from the termination complex has recently been identified (Beissel et al., 2019).

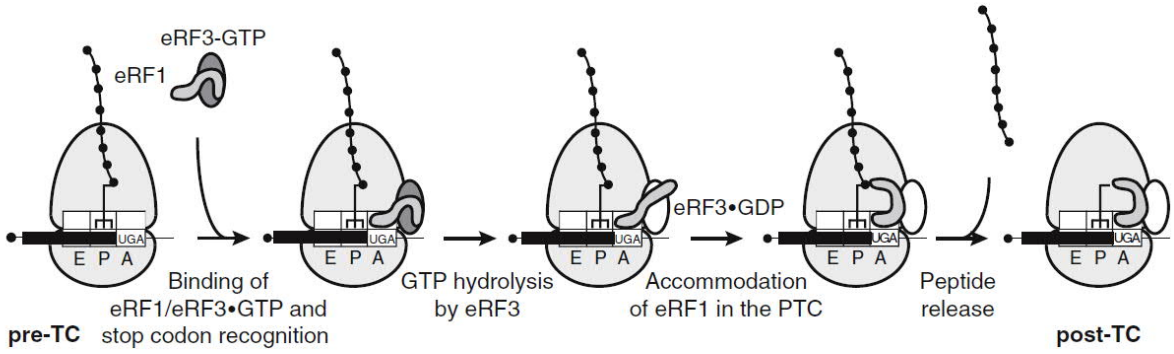


Figure 8. Scheme of canonical translation termination in eukaryotes.
The TC formed by eRF1, eRF3 and GTP binds to preTC. Stop codon recognition involves compaction of the mRNA that results in accommodation of the +4 base into the pocket formed by eRF1 and the SSU. Stop codon recognition is followed by structural changes within the TC that result in GTP hydrolysis by eRF3. Accommodation of the GGQ motif into the PTC induces the hydrolysis and release of the nascent peptide chain, resulting in the formation of the post-TC. The figure is adapted from Hellen (2018).

## 1.2 Translational recoding

In order to maintain cellular homeostasis, translation needs to be fast and accurate. Protein products formed erroneously might fold and function aberrantly, which is detrimental to cell viability. The frequency of amino acid misincorporation in eukaryotic organisms has been reported to be between $10^{-6}$ to $10^{-3}$ (Kramer et al., 2010; Stansfield et al., 1998). Several proofreading steps ensure accurate aminoacylation of tRNA and mRNA decoding during translation elongation (Guo and Schimmel, 2012; Hati et al., 2006; Ibba and Soll, 1999; Moras, 2010; Rodnina, 2012; Rodnina et al., 1996; Rodnina and Wintermeyer, 2001). Errors that occur due to misacylation or miscoding generally result in a point mutation that does not pose deleterious effects unless it affects the active site or the residues crucial for folding. Under normal conditions, such errors are non-redundant, stochastic and relatively rare. However, the translational machinery can adopt surprisingly high levels of flexibility when it comes to coding of certain mRNAs containing signal sequences that interfere with standard decoding. In some cases, the deviations from the standard genetic code can be exceptionally high, reaching up to 80% (Caliskan et al., 2015; Grentzmann et al., 1998; Huang et al., 1988; Loughran et al., 2014; Namy et al., 2001; Pennell et al., 2008; Tsuchihashi and Kornberg, 1990); such events are known as ‘translational recoding’.

Recoding events such as ribosomal frameshifting, translational readthrough and ribosomal bypassing (Fig. 9) have been identified in all domains of life and serve to expand the coding capacity of the genome. The majority of these events are guided by the presence of *cis*-acting elements embedded in the mRNA and the resulting polypeptide possess altered biological properties.

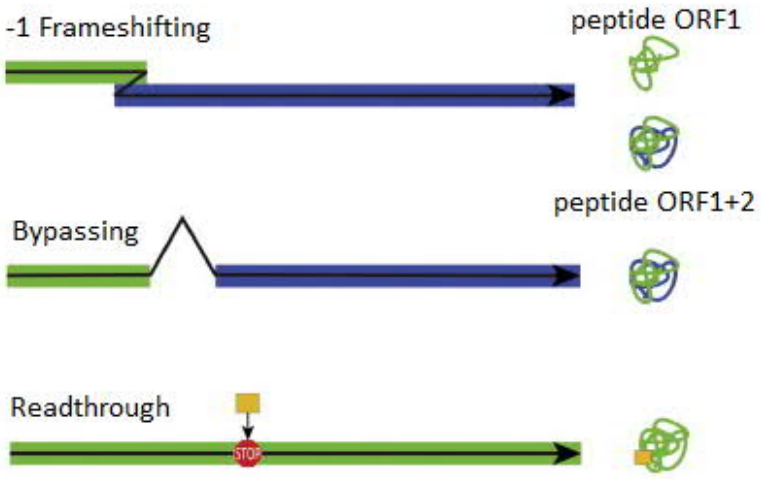


Figure 9. Recoding events during translation.
The green and blue regions indicate different reading frames in case of frameshifting and bypassing. In readthrough, an amino acid (yellow) is incorporated at the stop codon (red). The black arrow depicts the direction of translation. The figure is adapted from Caliskan et al. (2015).

Programmed ribosomal frameshifting (PRF) leads to the synthesis of a polypeptide from an altered frame, slipping one base either forwards (+1 PRF) or more commonly backwards (-1 PRF) on the mRNA. Specific slippery sites induce stalling of the translating ribosome leading to slippage (Caliskan et al., 2015; Plant and Dinman, 2006). PRF was initially identified in viral genomes where it plays an important role in viral propagation by modulating synthesis of viral proteins in specific stoichiometric ratios (Jacks and Varmus, 1985; Plant et al., 2010). Over the recent years, -1 PRF has been found to occur in all three domains of life (Dinman, 2012; Namy et al., 2004).

Ribosomal bypassing involves skipping of a portion of the mRNA by the translating ribosome leading to the production of an altered polypeptide from a discontinuous frame. Bypassing was first identified in the translation of the *gene 60* of bacteriophage T4 (Huang et al., 1988; Weiss et al., 1990), where the ribosome translates the mRNA until it reaches a GGA (Gly) codon, positioned upstream of a UAG stop codon. Stimulatory signals in the mRNA, together with the nascent peptide chain and the presence of the stop codon induce the peptidyl-tRNA to disengage its pairing with the GGA codon, which triggers the sliding of the ribosome on a 50-ntd non-coding mRNA gap. The ribosome "lands" on another GGA codon, where decoding is resumed (Agirrezabala et al., 2017; Herr et al., 2000; Samatova et al., 2014). Bypassing has also been identified in the mitochondrial genome of the yeast *Magnusiomycetes* (Lang et al., 2014).

Translational readthrough or stop codon readthrough involves decoding of a stop codon as a sense codon by a near-cognate tRNA (nc-tRNA), or a natural suppressor tRNA. TR does not alter the reading frame of translation, but rather extends the polypeptide C-terminally. Readthrough has evolved as a highly programmed and regulated mode of translational recoding. Particular examples

of TR include the recoding of the UGA stop codon to selenocysteine or of UAG to pyrrolysine, both of which expand the capacity of the genetic code by introducing non-canonical aminoacids and require dedicated tRNAs and protein factors (Blight et al., 2004; Bock et al., 1991; Forchhammer et al., 1991).

### 1.2.1 Translational readthrough (TR)

Translation termination is an efficient process with an intrinsic error rate of ≤ 0.1%. (Floquet et al., 2012; Harrell et al., 2002; Namy et al., 2001; Schueren and Thoms, 2016). Despite the high fidelity, the ribosome occasionally decodes the stop codon as a sense codon, leading to the incorporation of a standard amino acid at the stop signal. In such cases, translation continues until the next in-frame stop codon is encountered. This phenomenon is known as 'Translational Readthrough' (TR). Stop codon suppression via TR appends a C-terminal extension to the native protein allowing the production of two protein isoforms from the same transcript. TR relies on the competition between decoding of stop codons by release factors and by nc-tRNAs. nc-tRNAs are cellular tRNAs that act as suppressors by pairing with the stop codon at two out of the three positions of the codon-anticodon duplex. The extent of misincorporation of nc-tRNA at the stop codon is influenced by many factors (described in detail in section 1.2.2). Under normal circumstances, a gene will undergo TR at a rate dictated by normal translation fidelity. However, in some genes, where the stop codon is present in a context that favors termination suppression, TR levels are elevated by several hundred-fold, ranging from 1% to 40% (Loughran et al., 2014; Namy et al., 2001). TR, in these cases, is highly programmed and gives rise to specific stoichiometric ratios of functional protein isoforms (Csibra et al., 2014; Irigoyen et al., 2018).

TR is widely employed by viruses to expand the coding potential of their limited genome (Felsenstein and Goff, 1988; Firth et al., 2011; Hofstetter et al., 1974; Pelham, 1978). TR was first detected in *E. coli* phage *Qβ* (Weiner and Weber, 1973), following which many viral genes were identified that utilize TR to produce extended proteins with functional roles (Table 1). In recent years, extensive TR has been reported in higher eukaryotes where the addition of a C-terminal extension adds cellular localization signals, homo/heterodimerization domains, alters ligand binding properties or even confers antagonistic properties (Table 1).

Table 1. Examples of TR in genes from different kingdoms of life.

| | Gene | RT % | Function of TR | Reference |
|---|---|---|---|---|
| Virus | | | | |
| Coliphage *Qβ* | Minor coat protein *A1* | 5% | Formation of infectious particles | Hofstetter et al. (1974); Weiner and Weber (1973) |
| *Tobacco mosaic virus* | Replicase | 10-35% | RNA polymerase domain | Beier et al. (1984); Pelham (1978) |
| *Sindbis virus* | *nsP4* | 10% | Viral replication | Li and Rice (1993) |
| *Luteovirus* (BYDV, BWYV) | Coat protein | | Aphid transmission | Brault et al. (1995); Brown et al. (1996); Filichkin et al. (1994) |
| MuLV | *gag-pol* | 5% | Replication by Gag-pol fusion protein | Csibra et al. (2014); Yoshinaka et al. (1985) |
| Bacteria | | | | |
| *B. subtilis* | *sacB* levansucrase | | Modification of enzymatic properties | Chambert et al. (1992) |
| Eukaryotes | | | | |
| *S. cerevisae* | *PDE2* | 0.5-2.2% | Proteasome dependent degradation | Namy et al. (2002) |
| *U. maydis*<br>*A. nidulans* | *PGK*<br>*GAPDH* | | PTS1<br>PTS1 | Freitag et al. (2012) |
| Rabbit | *β-globin* | | | Chittum et al. (1998); Geller and Rich (1980); Hatfield et al. (1988) |
| Vertebrates | *MPZ* | 14% | Role in myelination | Yamaguchi and Baba (2018); Yamaguchi et al. (2012) |
| Mammals | *VEGFA*<br>*MTCH2*<br>*AGO1* | 10-85%<br>13%<br>24% | Anti-angiogenic activity | Eswarappa et al. (2014) |
| | *OPRL1*<br>*OPRK1*<br>*MAPK10*<br>*AQP4* | 31%<br>13%<br>14%<br>7% | | Loughran et al. (2014) |
| Humans | *LDHB*<br>*MDH1* | 1.5-5% | PTS1<br>PTS1 | Schueren et al. (2014); Stiebler et al. (2014) |
| | *VDR* | 6.7% | Reduced transcriptional response to calcitriol | Loughran et al. (2018) |

### 1.2.2 Factors influencing TR

TR is an evolutionarily conserved phenomenon. A large number of biochemical experiments in several systems, backed by comparative genomics, have identified many factors that modulate TR (Dabrowski et al., 2015; Harrell et al., 2002; Jungreis et al., 2011; McCaughan et al., 1995; Schueren and Thoms, 2016; Tork et al., 2004). These factors include *cis* elements on the mRNA as well as *trans* factors (Fig. 10).

Not surprisingly the first element that influences TR is the identity of the stop codon: UAA was reported to have the highest fidelity in termination, while UGA demonstrated the highest potential to undergo TR (Cridge et al., 2018; Howard et al., 2000; Loughran et al., 2014; Manuvakhova et al., 2000). Curiously, although UAA is the most represented stop codon in genes with higher expression and housekeeping functions; UGA is the most frequent stop codon in humans (Trotta, 2016).

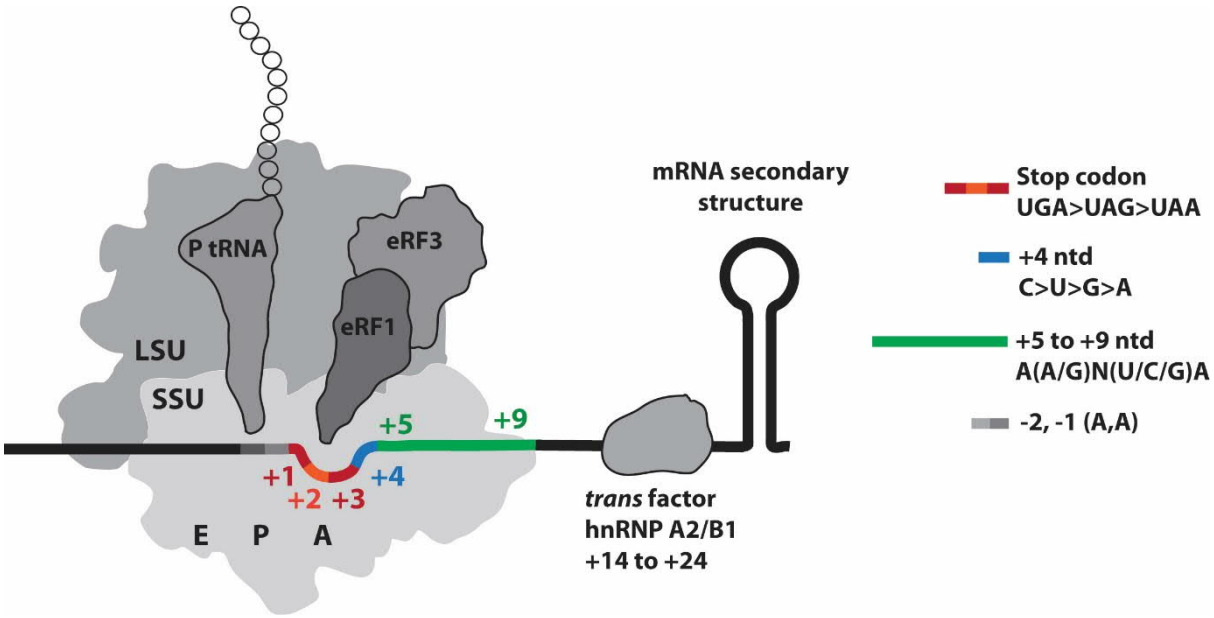


Figure 10. Schematic representation of factors affecting translational readthrough.
*cis* factors that affect TR include sequences upstream of stop codon (light grey), identity of the stop codon (red), the +4 nucleotide (blue) and the downstream sequences that occupy the mRNA channel (green). Distal *cis* element includes downstream mRNA secondary structure. Among several *trans* factors that affect TR, a specific case of hnRNP A2/B1 is depicted in the figure. hnRNP A2/B1 binds a *cis* element in 3' UTR of mammalian gene *VEGFA* and promotes TR.

The 5' context of stop codons shows a non-random distribution of nucleotides in *E. coli* and in humans (Arkov et al., 1995). Initial studies in yeast found an influence of -2 codon relative to the stop codon on termination efficiency, which suggested that acidic amino acids would favor efficient termination. The influence of -1 codon was attributed to the identity of P site tRNA (Mottagui-Tabar et al., 1998). However, a later study in yeast suggested that the major effect was imposed by only two bases immediately upstream of the stop codon with no correlation with the penultimate and ultimate amino acid residue in the polypeptide chain nor the structure of P site tRNA (Tork et al., 2004). The presence of two adenines immediately upstream of the stop codon induces elevated levels of readthrough, perhaps by modifying the mRNA structure in the P site which in turn alters decoding at the A site through distortion of the ribosome structure (Tork et al., 2004). U at the -1 position is associated with the lowest levels of TR (Cassan and Rousset, 2001; Loughran et al., 2014). Interestingly, adenines at -1 and -2 positions are conserved in the majority of plant and animal viral genes that undergo TR (Bonetti et al., 1995; Tork et al., 2004).

The 3' context of the stop codon has a major role in determining TR. In the phage *Qβ* coat protein (Hofstetter et al., 1974; Weiner and Weber, 1973) as well as *nsP4* polymerase in *Sindbis virus* (Li and Rice, 1993), a single cytidine downstream of the UGA stop codon, i.e at the +4 position was found to be important for TR. This observation is true in both prokaryotes and eukaryotes, despite the different mechanisms of stop codon recognition (Brown et al., 1990a, b; McCaughan et

al., 1995; Pedersen and Curran, 1991; Tate et al., 1995). In higher eukaryotes, the effect of the +4 nucleotide on TR modulation was found to vary between the three stop codons (Dabrowski et al., 2015). Nevertheless, the presence of a C immediately downstream of the stop codon consistently increased TR with the most striking effect exerted on UGA. Notably, the stop codon context with a C downstream of UGA or UAG are rarely used as termination contexts in mammals (McCaughan et al., 1995). The identity of the stop codon together with the +4 nucleotide is referred to as the immediate stop codon context hereon.

Context-dependent effects can often be attributed to a broader stretch of nucleotides downstream of the stop codon. In *Tobacco Mosaic Virus* (TMV), the consensus sequence CARYYA (purines: R, pyrimidines: Y) triggers TR at all stop codons (Skuzeski et al., 1991). The nucleotides +4 to +6 in the context of UGA-CUA or UGA-CGG are essential to induce TR at the UGA stop codon in a number of plant and animal viral genes using natural suppressor tRNAs (Beier and Grimm, 2001; Urban et al., 1996). UGA-CUA seems to be the most efficient autonomous TR signal in mammalian cell lines (Cridge et al., 2018). The stretch of six nucleotides downstream of a stop codon is inferred to be occupying the mRNA channel (Ingolia et al., 2012; Ingolia et al., 2009) where it makes extensive contacts with rRNA and ribosomal proteins until the solvent front is reached (Cridge et al., 2018; Shao et al., 2016). The nucleotides occupying the mRNA channel, along with the two nucleotides upstream of the stop codon (-2 to +9) comprise the minimal *cis* primary sequence motif that modulates TR (Fig. 10).

In addition to the immediate context, more distal stimulatory 3' *cis* elements involving mRNA structures have been identified in several viruses as well as eukaryotic mRNAs that play an important role in regulating termination efficiency. (Feng et al., 1992; Firth et al., 2011; Jungreis et al., 2011; Wills et al., 1991, 1994). Well studied examples of TR stimulatory RNA structures do not show conservation in the nucleotide span that comprise the secondary structure (pseudoknot) or the spacer region between the stop codon and the stem loop akin the mRNA secondary structure elements enhancing programmed ribosome frameshifting (Brown et al., 1996; Firth et al., 2011). *Cis*-acting RNA structures are proposed to modulate TR by (i) interfering with release factor recruitment to preTC by steric hindrance or sequestration, (ii) modulating ribosome function by direct interaction with ribosomal proteins or rRNAs, (iii) inducing ribosomal stalling during termination by impeding mRNA unwinding or (iv) recruiting *trans* factors (Eswarappa et al., 2014; Firth et al., 2011; Napthine et al., 2012). In any case, such interference allows enough time for the aa-tRNA to decode the stop codon thereby leading to TR. It is worthy to note that the sequences downstream of the primary stop codon are subject to evolutionary selection against stop codon

readthrough, as in-frame stop codons are significantly over-represented immediately downstream of the primary stop signal, which limits leaky termination (Williams et al., 2004).

In addition to the elements in the mRNA, several *trans* factors may influence the efficiency of termination by various mechanisms. Recently, eIF3 was proposed to promote TR at all three stop codons in leaky context by preventing eRF1 from recognizing the third position of the stop codon (Beznoskova et al., 2015). Depletion of termination factors eRF1 and/or eRF3 results in increased levels of stop suppression in humans (Carnes et al., 2003; Chauvin et al., 2005). Readthrough by altering eRF levels is independent of the codon context. The [Psi$^+$] strain of *S. cerevisiae* exhibits the epigenetically inherited prion state of termination factor eRF3 where translation termination is heavily compromised. In these strains, eRF3 forms amyloid fibrils that aggregate, thereby sequestering a part of the release factor pool and rendering them non-functional (Liebman and Sherman, 1979; Paushkin et al., 1996). The amyloid state of eRF3 causes increased misreading of stop codons and can be propagated to daughter cells. The switch of the normally soluble eRF3 protein to prion state possibly reprograms gene expression to aid survival and fitness during stress conditions (True and Lindquist, 2000; Tyedmers et al., 2008). Similarly, the abundance of specific suppressor tRNAs also influences the cellular levels of TR (Beznoskova et al., 2016; Blanchet et al., 2014; Roy et al., 2015a). For example, the relative abundance of the major tRNA$^{Gln}$ isoacceptor with 5'-UUG-3' anticodon compared to the minor tRNA$^{Gln}$ with 5'-CUG-3' explains why glutamine is selectively incorporated at UAA compared to UAG, despite the same non-conventional G-U base pairing that occurs with both tRNAs. Furthermore, modification of the bases within the anticodon or in its vicinity affects the ability of tRNAs to read stop codons (Beier and Grimm, 2001). In the case of the mammalian vascular endothelial growth factor A (VEGFA) gene, the heterogenous nuclear ribonucleoprotein (hnRNP) A2/B1 was identified as a *trans* factor that binds the hnRNP A2/B1 recognition element (A2RE) in the readthrough region to positively regulate TR (Eswarappa et al., 2014).

## 1.3 TR in *Drosophila*

While biochemical assays and comparative genomics have aided the identification of TR candidates in yeast and in mammals, a major breakthrough was achieved by studies performed in the *Drosophila* species, which have the highest number of TR genes reported among eukaryotes. *Drosophila melanogaster*, colloquially known as the fruit fly, has been used as a model organism for more than a hundred years. Low cost of maintenance, short generation time, ease of genetic manipulation and extensive exchange of knowledge and resources within the fly research community has allowed scientists to use *Drosophila* as a model system to perform analytical investigations to address large number of biomedical problems. In addition, with the advancement

in genome-wide surveys and gene engineering techniques in *Drosophila*, novel findings in modern biology are often pioneered using *Drosophila* and then tested and generalized to higher organisms such as humans.

The first gene reported to undergo TR in *Drosophila* was *kelch* which encodes a short native protein and a longer extended TR protein via suppression of a UGA codon (Xue and Cooley, 1993). The regulation of TR in *kelch* was found to occur temporally during development with the ratio of the extended protein to the native protein reaching 1:1 during metamorphosis. The increase in expression of Kelch-TR isoform also occurs in a tissue-specific manner with highest expression observed in the imaginal discs (Robinson and Cooley, 1997). Successively, *synapsin* (*syn*), *headcase* (*hdc*), non-sense alleles of *embryonic lethal abnormal vision* (*elav*) and *wingless* (*wg*) were also identified to undergo TR (Chao et al., 2003; Klagges et al., 1996; Samson et al., 1995). The native Hdc protein was found to be four times more abundant that the longer TR isoform in embryos. (Steneberg et al., 1998).

### 1.3.1 System biology of *Drosophila* genes

A major advancement in the identification of TR genes in *Drosophila* was brought about by the availability of sequenced genomes from 12 *Drosophila* species (Clark et al., 2007; Stark et al., 2007). This allowed the identification of recurrent patterns of evolutionary signatures that are specific to protein coding sequences. Using these signatures to examine the 3' untranslated region (UTR) of genes across the entire *Drosophila* genome, 283 TR candidates were identified (Jungreis et al., 2011; Lin et al., 2007). Analysis of the stop codon context of the putative TR candidates showed that UGA-C is the most common stop codon context (32.2%) compared to non-TR transcripts (3.1%). Transcripts containing UGA-C are 10 fold more likely to be TR candidates. The frequency of stop codons in the TR candidates was found to be in the order UGA>UAG>UAA while that of 4' nucleotide is C>U>G>A. UGA is the stop codon in 64% of these genes, while C is present in the +4 position in 51% of these genes (Jungreis et al., 2011).

Algorithm-based phylogenetic tools employ cross-species comparisons to identify evolutionarily conserved signatures. Therefore, they are incapable of identifying evolutionarily novel TR events. The use of ribosome profiling, which involves deep sequencing of ribosome-protected footprints along the mRNA (Ingolia et al., 2009), addressed these issues and made it possible to annotate 350 functional TR candidates in the *Drosophila* genome (Dunn et al., 2013) . By comparing footprint densities obtained from Schneider 2 cells (S2 cells) and early embryos, several transcripts were found to undergo TR to different extents, indicating differential regulation across cell types.

While phylogenetics and ribosome profiling are complementary methods that identify TR events, they do not provide information on the spatio-temporal regulation of TR or the functional relevance of TR extension in an organism. Only few of the predicted TR extensions in *D. melanogaster* contain annotated protein domains, such as nuclear localization signals (NLS), peroxisomal targeting signal 1 (PTS1) or transmembrane domains, which hints that the majority of the TR isoforms function in unknown ways (Dunn et al., 2013). Computational analyses of TR genes have revealed that TR mostly affects long, modular proteins that possess intrinsically disordered C-termini with low sequence complexity (Kleppe and Bornberg-Bauer, 2018; Pancsa et al., 2016). Lack of structurally ordered C-terminus could provide conformational pliability and accessibility that allows the TR extensions to perform functions without distorting the native protein. The majority of TR genes identified in *D. melanogaster* have regulatory roles, and appending a functional C-terminal extension may confer conditional advantage to the gene function. The abundance of ribosomal components and translation initiation factors among TR candidates suggests that TR might help to fine tune the function of ribosomes by affecting its stability, fidelity of translation or even giving rise to specialized ribosomes that preferentially translate specific mRNAs (Pancsa et al., 2016). Our understanding of the functional significance of TR extensions is largely impeded by the lack of biochemical studies that characterize TR isoforms.

## 1.4 *traffic jam* as a TR candidate

*Traffic jam (tj)* is the only large Maf transcription factor in *Drosophila* and has been predicted to be a TR candidate by phylogenetic studies (Jungreis et al., 2011). *tj* was initially identified as a regulator of multiple processes during gonad morphogenesis including stem cell specification during spermatogenesis and collective cell migration during oogenesis (Gunawan et al., 2013; Li et al., 2003; Wingert and DiNardo, 2015). Tj is translated from a single-exon open reading frame (ORF) comprised of 509 amino acids that terminates with the leakiest stop codon context of UGA-C. The TR extension would append an additional 44 amino acids to the native Tj.

Maf transcription factors have been identified in a wide range of higher metazoans. Tj is a homolog of the retroviral oncoprotein v-Maf and large Mafs of vertebrates (Blank and Andrews, 1997). Tj shows extensive sequence similarity with its mammalian orthologues c-Maf and MafB especially within the conserved domains (Li et al., 2003). They contain a highly conserved, basic leucine zipper structure (bZip) and an extended homology region (Motohashi et al., 2002) (Fig. 11). bZip transcription factors can form homo- or heterodimers via their leucine zipper domain in order to bind target DNA (Franza et al., 1988; Lamb and McKnight, 1991). Maf factors are known to modulate tissue-specific gene expression and cell differentiation by binding to the regulatory

regions of target genes and by interacting with other transcription factors (Blank and Andrews, 1997; Kataoka et al., 2002; Kurokawa et al., 2009; Ogino and Yasuda, 1998; Rehemtulla et al., 1996).

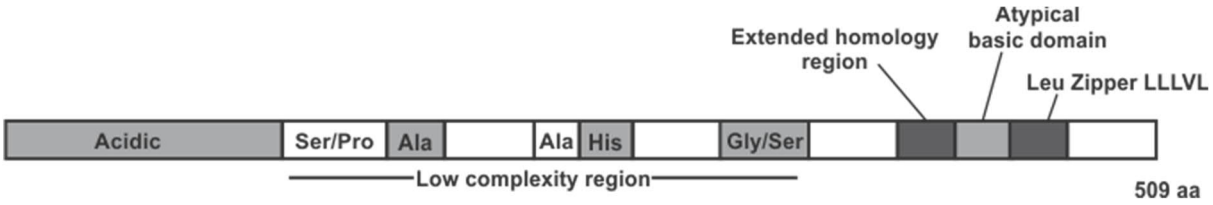


Figure 11. Protein domains in Tj.
Tj contains an N-terminal acidic domain followed by a low complexity domain consisting of Ser/Pro rich sequence, two Ala repeats, region containing His repeats and Gly/Ser rich sequence. C-terminal domain contains an extended homology domain, a basic domain and a leucine zipper motif.

### 1.4.1 Function of *tj* in *Drosophila* gonad development

Tj is a nuclear protein that is extensively studied in the somatic cells in the gonads. Somatic gonadal cells interact with germline cells throughout development and the interaction is crucial for normal development and differentiation of the germline. The somatic and germline cells in fly gonads arise from stem cells that reside in a specialized microenvironment called the 'niche'. The stem cell niche provides instructive cues that guide stem cells to self-renew or differentiate (Morrison and Spradling, 2008). In male gonads, the stem cell niche is located at the apical tip of the testis and is composed of approximately ten post-mitotic somatic cells called the hub (Hardy et al., 1979) (Fig. 12). The germline stem cells (GSCs) and somatic cyst stem cells (CySCs) are in direct contact with the hub cells. The hub cells secrete protein factors and cytokines that maintain the self-renewal of GSCs and CySCs (Wingert and DiNardo, 2015). Each GSC is ensheathed by two CySCs via cytoplasmic extensions. The GSCs divide asymmetrically to generate one stem cell that maintains contact with the hub and one gonialblast (GB) that initiates differentiation. A GB undergoes four rounds of divisions to produce a cluster of 16 spermatogonial cells that further develop into spermatocytes, spermatids and eventually mature sperms (White-Cooper, 2010). The CySCs also divide asymmetrically to give rise to CySCs and two somatic cyst cells that ensheath the GB. The cyst cells grow but do not divide while maintaining the encasement around GB and its progeny throughout spermatogenesis (de Cuevas and Matunis, 2011). Tj is mainly expressed in the CySCs and early cyst cells that encapsulate mitotically active germline. Weak expression of Tj is also found in hub cells (Li et al., 2003; Okegbe and DiNardo, 2011; Voog et al., 2008).

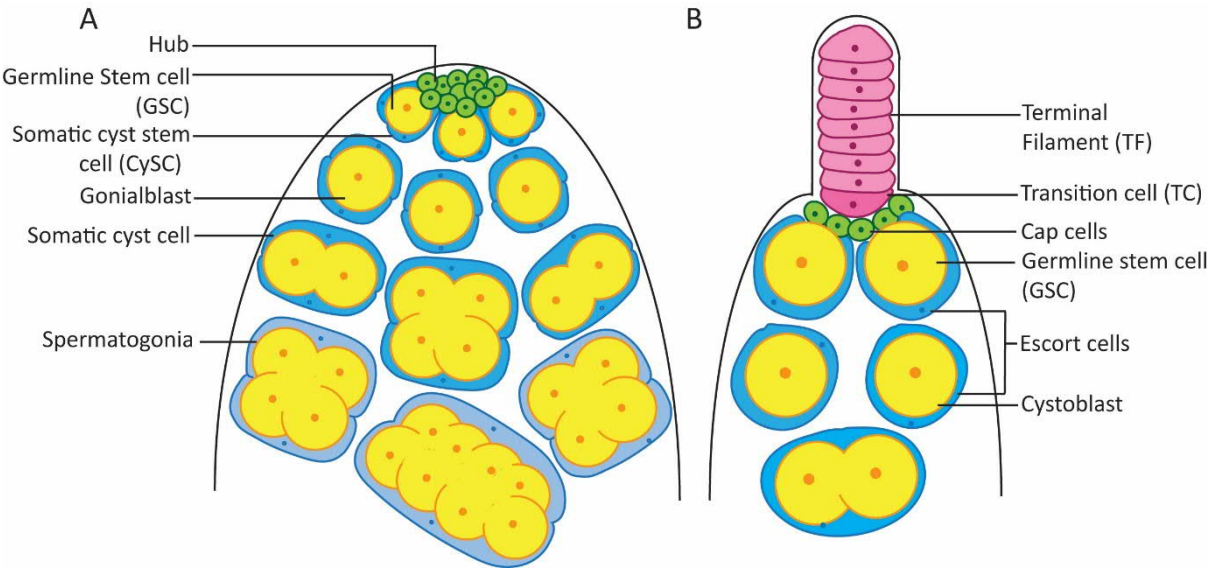


Figure 12. Schematic drawing of *Drosophila* GSC niche.
The organization of the GSC niche depicting hub/cap cells (green), germline cells (yellow) and somatic cells (blue) in apical tip of testis (A) and in the germarium of ovaries (B). Terminal filament cells at the tip of germarium are shown in pink.

The female GSC niche is located at the anterior-most tissue in the *Drosophila* ovary called the germarium. The female GSC niche consists of cap cells, escort cells and terminal filament (TF) cells (Fig. 12). GSCs are in direct contact with the cap cells. Cap cells secrete different factors (similar to the hub cells) that repress differentiation of GSCs (Song et al., 2004; Xie and Spradling, 1998). GSCs divide asymmetrically to produce daughter cystoblasts that lose contact with the cap cells. Cystoblasts divide four times with incomplete cytokinesis to form a germline cyst composed of 16 cystocytes. Escort cells ensheath the GSCs and cystoblasts with their cytoplasmic projections. In the germarium, Tj is expressed in cap cells, escort cells and also weakly in newly identified transition cells that lie between cap cells and TF cells. The 16-cell cyst passes the follicle stem cells where the follicle cells encapsulate the cyst to form an egg chamber. The egg chamber buds off from the posterior end of the germarium whilst still being attached via stalk cells. One of the 16 cells in the cyst assumes the oocyte fate while the remaining 15 become nurse cells (Waghmare and Page-McCaw, 2018). Tj is also expressed abundantly in the ovarian follicle cells (Gunawan et al., 2013).

Expression of Tj in SGCs begins at stage 12 of embryogenesis (Li et al., 2003). Tj expression is concomitant with the timeline when the primordial germ cells (PGCs) travel through the gut and establish first contacts with the somatic gonadal precursor cells (SGPs). After coalescence, PGCs intermingle with the SGPs to form the presumptive gonad. The intermingling of PGCs and SGPs is crucial for proper development and differentiation of the germline and is mediated by a differential expression of several cell adhesion molecules such as Fascilin III (Fas3), DE-cadherin (DEcad) and

Neurotactin (Nrt) (Li et al., 2003). The expression of Tj in somatic gonadal cells is maintained throughout development.

*tj* is essential for gonad morphogenesis, as *tj* null mutants ($tj^{-/-}$) were found to be viable but sterile (Li et al., 2003). *tj* mutants display highly defective gonads without other morphological and developmental anomalies. Adult testes in $tj^{-/-}$ mutants are reduced to distorted ball-like structures with very few germ cells that cluster together and are arrested in the pre-spermatocyte state (Fig. 13). The somatic cells are segregated from the germ cells and are distinct from the wild-type testes in that the contact between the somatic cyst cells and differentiating germline cyst via cellular ensheathing is completely absent. About 25% of the ovaries of 1-2 day old $tj^{-/-}$ flies lack germline cells completely and the remaining ones lose them over time. The persisting germ cells form randomly placed, irregular clusters that reach different levels of early differentiation. The mutant ovaries also lack somatic follicle cells, thereby the germ cells that reach follicle stage cannot get enveloped by follicle cells.

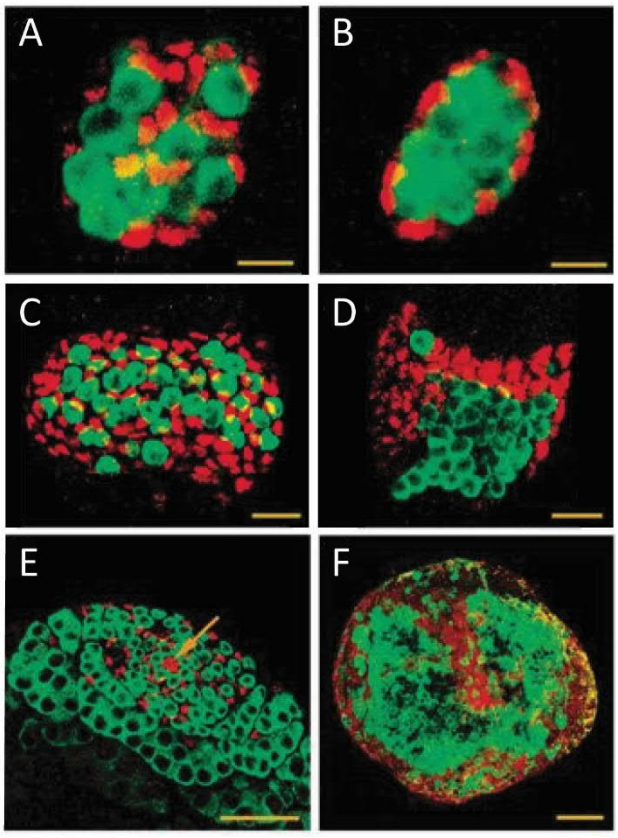


Figure 13. Defects in the interaction between somatic cells and germ cells in *tj* mutants.
(A,B) Embryonic gonads at stage 15-16. Tj (red) labels nuclei of SGPs and Vasa (green) labels PGCs. In wild-type (A), the SGPs and PGCs are intermingled while in $tj^{-/-}$ mutants (B), the SGPs are restricted to the periphery of the gonad. (C, D) Late third instar larval ovaries. Tj (red) labels nuclei of somatic interstitial cells and Vasa (green) labels germ cells. In wild-type larval ovaries (C) the germ cells and interstitial cell form an interspersed population while in $tj^{-/-}$ ovaries (D), they form segregated populations with germ cells clustered at the center. (E, F) Late third instar testes. Tj (red) stains hub cells and somatic cyst cells. Vasa (green) stains germline cells. (E) Magnified view of apical tip of wild-type testis shows Tj-positive hub cells at the center (arrow) with somatic cyst cells ensheathing mitotically active early germ cells. (F) Cross-section of $tj^{-/-}$ mutant testis with small germ cells clustered together and somatic gonadal cells restricted to the periphery. The figure is adapted from Li et al. (2003).

In stage 15-16 $tj^{-/-}$ embryos, the PGCs and SGPs coalesce and form the embryonic gonad. SGPs are properly specified as they express several SGP-specific markers. However, the number of PGCs does not increase in a normal way and the mixing between PGCs and SGPs is hindered such that the SGPs remain in the periphery of the gonad (Fig. 13B). The segregation of SGPs and germ cells is

even more distinct in late third instar larval testes and ovaries (Fig. 13 D, F). The germline cells cluster together, distinctly separated from Tj-positive somatic cells.

The intermingling of somatic and germ cells is guided by principles of cell sorting that originate from differences in the expression levels of cell adhesion molecules in the participating cell lines. Differential adhesive properties of cells allow them to sort into a defined three-dimensional structure (McNeill, 2000, 2003). *tj*$^{-/-}$ testes express Fas3 not only in the hub cells but also ectopically in the cyst cells (Fig. 14B). *tj*$^{-/-}$ follicle cells show ectopic expression of Fas3, DEcad and Nrt, which are all cell adhesion molecules (Fig. 14 D,F). Thus, *tj* regulates the expression of several cell adhesion molecules that are crucial for cell sorting and gonad morphogenesis. Removal of *tj* function results in upregulation of adhesion molecules and defective germline-soma interaction. However, it is not known whether the interaction between the transcription factor and its targets is direct.

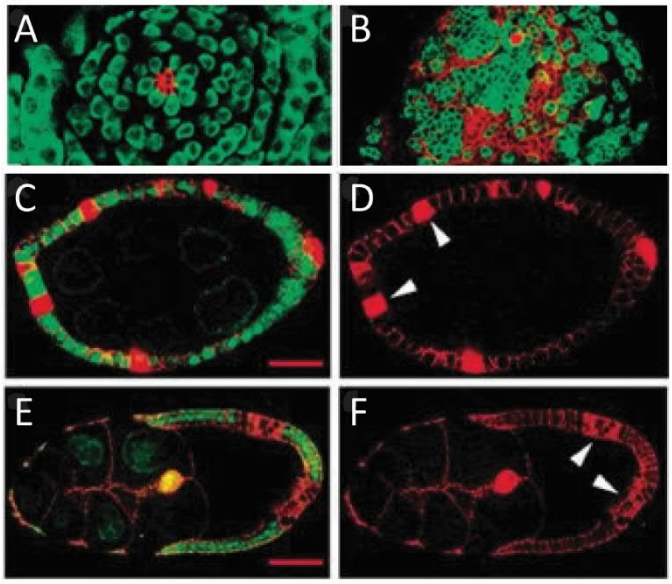


Figure 14. Effect of *tj* on the expression of cell adhesion molecules.
(A, B) Late third instar larval testes. In wild-type (A), Fas3 (red) is only expressed in the hub cells while in *tj*$^{-/-}$ mutants (B), Fas3 is ectopically expressed in somatic cells along with hub cells. (C-F) Ovarian follicles containing *tj*$^{-/-}$ mutant follicle cells recognized by lack of GFP (green) (arrows). (C, D) Ovarian follicles at stage 7 of oogenesis where Fas3 (red) is specifically upregulated in *tj*$^{-/-}$ follicle cells that are GFP-negative. (E, F) Ovarian follicles at stage 9 where only polar cells and *tj*$^{-/-}$ follicle cells exhibit elevated levels of DEcad expression (red). The figure is adapted from Li et al. (2003).

*tj* plays an additional important role in the specification of the stem cell niche in both male and female gonads (Panchal et al., 2017; Wingert and DiNardo, 2015). In male gonads, hub cell specification begins after *Notch* activation in SGPs located at the anterior end of embryonic gonads. Activation of *Notch* results in downregulation of *tj,* which is crucial for hub cell specification. Loss of *tj* function results in generation of dispersed ectopic hub cells that are capable of recruiting stem cells. Thus, *tj* functions downstream of *Notch* in suppressing hub cell fate (Wingert and DiNardo, 2015). In ovaries, *tj* controls specification and morphogenetic behavior of cap cells. *tj* depletion results in development of additional terminal filament cells at the expense of cap cells. Hypo-expression of *tj* causes cap cells to acquire shape and spatial organization of terminal filament cells. In ovaries, *tj* and *Notch* are proposed to act in distinct ways to specify cap cells. *tj* promotes cap cell fate and regulates niche architecture by blocking terminal filament cell fate, while *Notch* supports cap cell fate by preventing escort cell fate (Panchal et al., 2017).

The 3'UTR of *tj* codes for a Piwi-interacting RNA (piRNA) cluster that gives rise to sense-oriented piRNAs via a conserved primary piRNA biogenesis pathway. The *tj*-3'UTR derived piRNAs are then loaded on to Piwi, whose expression is positively controlled by Tj, and together, they silence specific target genes, usually transposable elements (Robine et al., 2009; Saito et al., 2009). The target genes of *tj*-derived piRNAs have not been discovered yet, however some of the piRNAs have been shown to demonstrate strong complementarity to *Fas3* primary transcript. *Fas3* has already been identified as a target whose expression is negatively regulated by *tj* (Li et al., 2003). Thus, it is probable that the downregulation of Fas3 is achieved by a composite action of Tj protein together with *tj*-3'UTR derived piRNAs. Tj is also known to activate the expression of Piwi in ovarian follicle cells (Saito et al., 2009).

### 1.4.3 *tj* expression in nervous tissue

Apart from somatic gonadal cells, restricted expression of *tj* transcript has been observed in embryonic and larval central nervous system (CNS) (Fig. 15), adult heads as well as adult fat bodies (Gelbart and Emmert, 2013; Li et al., 2003), where its function is largely unknown.

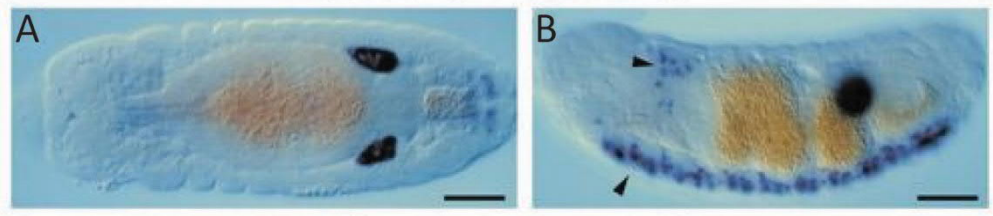


Figure 15. *tj* transcript distribution in embryos stage 15-16.
(A) Dorsal and (B) ventral view of wild-type embryos where *tj* transcript can be detected in embryonic gonads as well as in a subset of cells of the central nervous system (arrowheads). Scale bars represent 50 µm. The figure is adapted from Li et al. (2003)

*tj* was predicted to be responsible for glutamatergic neuronal fate in the optic lobe of adult brain (Konstantinides et al., 2018). RNAi-based knock-down of *tj* resulted in downregulation of Vesicular glutamate transporter (VGlut) in specific synaptic boutons in the medulla region of the brain. Tj is also expressed in glutamatergic, cholinergic and GABAergic interneurons of the ventral nerve cord (VNC) from embryogenesis to third instar larval stage. The Tj-positive cholinergic neurons control body posture in *Drosophila* larvae, while Tj-positive glutamatergic neurons control locomotion and Tj-positive GABAergic neurons control the speed of locomotion (Babski et al., 2018). Tj is also known to control the cell fate of photoreceptor type 8 (pR8) by regulating the expression levels of three different genes: *melted* (*melt*), *warts* (*wts*) and *Rhodopsin 6* (*Rh6*) (Jukam et al., 2013).

## 1.5 Scope of the thesis

Translational readthrough (TR) is a recoding mechanism that is extensively utilized by viruses, bacteria and eukaryotes. Comparative phylogenetic analysis of 12 *Drosophila* species has identified 283 putative TR candidates, revealing the pervasiveness of TR within the *Drosophila* genome. Despite the abundance, only limited number of predicted TR genes have been verified experimentally and the biological significance of TR in *Drosophila* remains unknown. In this work, we developed an *in vitro* reporter assay to study TR in a selected set of candidate genes that have been predicted to undergo TR in *Drosophila*. We used this assay to experimentally validate TR as well as quantify the extent of TR in these genes using S2 cells. In addition, we analyzed the impact of the immediate stop codon context on the extent of TR in these genes. In the case of one particular candidate, *aPKC*, we used systematic mutational analysis to delineate the minimal sequence context that is responsible for driving efficient levels of readthrough. Furthermore, to understand the biological significance of gene-specific readthrough in another candidate gene, *traffic jam*, we utilized CRISPR/Cas9 genome editing to create genetic mutants that exhibit constitutive or abolished TR. We utilized immunohistochemistry to study how TR affects *tj*-specific phenotypes in embryonic as well as adult gonads. Finally, by using high throughput RNA sequencing combined with qPCR gene expression analysis, we studied the role of TR in *tj* in shaping the transcriptome profile in adult fly brains.

# 2. MATERIALS AND METHODS

## 2.1 Materials

### 2.1.1 Equipment

Table 2. List of equipment

| Device | Manufacturer |
|---|---|
| Milli-Q Advantage | Millipore |
| pH-meter, pH electrode | WTW |
| Water bath E100 | Lauda |
| Benchtop centrifuge 5415R and 5810R | Eppendorf |
| Thermomixer comfort | Eppendorf |
| Vortex Genie 2 | Scientific industries |
| Centrifuge Avanti J-26 XP | Beckmann Coulter |
| Centrifuge Avanti J-30I | Beckmann Coulter |
| Dri-Block DB-2D heater | Techne |
| Mini star Silverline Microcentrifuge | VWR |
| JA-30.1 rotor | Beckmann Coulter |
| JLA-8.100 rotor | Beckmann Coulter |
| Nanodrop 2000C | Peqlab Biotechnologie |
| Electrophoresis power supply EV261 | Peqlab Biotechnologie |
| Agarose gel electrophoresis chamber | Peqlab Biotechnologie |
| PCR thermocycler Peqstar | Peqlab Biotechnologie |
| SDS PAGE electrophoresis equipment | BioRad |
| Incubator shaker series Innova 44 | New Brunswick |
| ÄKTA FPLC | GE Healthcare |
| Bio-vision imaging system | Peqlab Biotechnologie |
| Bio-5000 plus | Microtek |
| Incubator INE600 | Memmert |
| Branson Sonifier W-250D | G. Heinemann |
| Luminometer | Sirius Single, Berthold |
| Confocal LSM 700 | Zeiss |
| ZEN microscope | Zeiss |
| Step One Plus Real Time PCR System | Applied Biosystems |

### 2.1.2 Chemicals and consumables

Chemicals were purchased from Sigma Aldrich (Steinheim, Germany), Merck (Darmstadt, Germany), Carl Roth (Karlsruhe, Germany), Roche Diagnostics (Mannheim, Germany) and Serva (Heidelberg, Germany), unless indicated otherwise. DNA SmartLadder was from Eurogentec (Seraing, Belgium), Stain G was from PeqLab (Erlagen, Germany), 2-log DNA ladder and DNA gel loading dye purple (6x) were from New England Biolabs (NEB) (Frankfurt, Germany), Oligo (dT)20 primer was from Jena Bioscience (Jena, Germany). Kits for DNA purification, Nucleospin® plasmid and NucleoSpin® Gel and PCR clean up kits were from Macherey-Nagel (Düren, Germany). DNA oligonucleotides were purchased from Eurofins Genomics (Ebersberg, Germany). Substrates for

luciferase assay, Beetle Juice and Renilla Glow Juice were from PJK GmBH (Saarland, Germany). Invitrogen™ TRIzol™ reagent was from Thermo Fisher Scientific (Braunschweig, Germany).

### 2.1.3 Enzymes

Table 3. List of enzymes

| Cloning Enzymes | Manufacturer |
|---|---|
| Phusion® High-Fidelity DNA Polymerase | NEB |
| DpnI | NEB |
| T4 Polynucleotide Kinase | NEB |
| T4 Quick Ligase | NEB |
| T4 DNA Ligase | NEB |
| T5 Exonuclease | NEB |
| Taq DNA ligase | Biozymes |
| **Reverse Transcription and qPCR** | |
| DNaseI | Jena Bioscience |
| High Capacity Reverse Transcriptase | Beckmann Coulter |
| Fast SYBR™ Green Master mix | Thermo Fisher |

### 2.1.4 Media and buffers

Table 4. List of Solutions

| Media | Recipe |
|---|---|
| LB medium | 10 g/l NaCl, 10 g/l tryptone, 5 g/l yeast extract |
| LB agar | 10 g/l NaCl, 10 g/l tryptone, 5 g/l yeast extract, 15 g/l agar |
| **Buffers** | **Recipe** |
| TBE | 0.1 M Tris-HCl, 0.1 M boric acid, 2 mM EDTA (disodium salt) |
| PBS | 0.137 M Nacl, 2.7 mM KCl, 10 mM $Na_2HPO_4$, 1.8 mM $KH_2PO_4$ |
| PBT | 0.2% Triton™ X-100 in PBS |
| PBTB | 2 g/l BSA, 5% Normal goat serum, 0.5 g/l $NaN_3$ in PBT |

### 2.1.5 Bacterial strains and plasmids

All strains used are classified under biosafety level BSL-1. Genetic sub-cloning and mutagenetic experiments were performed in *E. coli* NovaBlue Singles™ competent cells.

Table 5. List of bacterial strains

| *E. coli* Strain | Genotype | Company |
|---|---|---|
| NovaBlue | *endA1 hsdR17* ($r_{K12}^{-}$ $m_{K12}^{+}$) *supE44 thi-1 recA1 gyrA96 relA1 lac* F′[*proA*$^{+}$*B*$^{+}$ *lacI*$^{q}$ZΔ*M15*::Tn*10*] ($Tet^{R}$) | Novagen, Merck |

Table 6. List of plasmid vectors used

| Vectors | Description | Company |
|---|---|---|
| psiCHECK™-2 | SV40 promoter - Rluc, HSV TK promter - Fluc, AmpR | Promega |
| pHD-DsRed | LoxP-3xP3-DsRed-LoxP, AmpR | Addgene |
| pU6-BbsI-chiRNA | DmU6 promoter-gRNA scaffold, AmpR | Addgene |

### 2.1.5 Software

Softwares used for DNA sequence analysis, data representation, qPCR data analysis, image build up and processing are listed in Table 7.

Table 7. List of software

| Software | Provider |
|---|---|
| Adobe Illustrator CS5 | Adobe Systems |
| CorelDRAW X7 | Corel |
| DNA Star Lasergene 12 | DNASTAR Inc. |
| SnapGene Viewer v4.1.3 | GSL Biotech |
| GraphPad Prism 5.0 | GraphPad Software |
| Fiji (ImageJ) | NIH |
| StepOne Software v2.3 | ThermoFisher Scientific |

### 2.1.6 Services

DNA sequencing services of Microsynth Seqlab, Göttingen were used. RNA sequencing services were provided by Transcriptome and Genome Analysis Laboratory (TAL), Göttingen. CRISPR/Cas9 injection and recombinant screening services were provided by BestGene, Chino Hills, CA, USA.

## 2.2 Molecular biology protocols

### 2.2.1 Polymerase chain reaction

Phusion polymerase was used for all molecular cloning procedures. Each 50 µl PCR reaction consisted of 100 ng template DNA (plasmid or gDNA or cDNA), 0.5 mM forward and reverse primers, 0.2 mM dNTPs and 1 unit of Phusion polymerase. The PCR reaction was performed in a thermocycler with an initial DNA denaturing step at 98 °C for 5 min followed by 30-35 cycles of denaturation at 98 °C for 3 min, annealing at a suitable temperature for 30 s and elongation at 72 °C for 30 s per kb of amplicon. Following each PCR reaction, the template DNA in the reaction mix was digested using DpnI. The PCR products were analyzed by agarose gel electrophoresis and visualized using Stain G. When required, the desired PCR products were purified using NucleoSpin® Gel and PCR clean up kit.

### 2.2.2 Site-directed mutagenesis using QuikChange protocol

Point mutations were introduced into plasmid vectors using The Agilent QuikChange II site-directed mutagenesis protocol. The first five rounds of amplification were performed separately for each template using forward or reverse primers in a final volume of 25 µl each. The reaction mixes were then pooled together and subsequent rounds of amplification were performed. The PCR products were digested with DpnI and transformed directly into chemically competent cells.

### 2.2.3 Mutagenesis using blunt-end ligation

Insertional and deletion mutagenesis was performed using a blunt-end ligation method. For insertional mutagenesis (up to 32 bp), forward and/or reverse primers containing a 5' overhang corresponding to the sequence to be inserted were designed. For deletion mutagenesis, primers complementary to the sites flanking the region to be deleted were designed. Standard PCR amplification was performed and the template was digested using DpnI. The blunt-end linear products were purified using NucleoSpin® Gel and PCR clean up kit. Phosphorylation and ligation reactions were performed using T4 Polynucleotide kinase and T4 ligase according to manufacturer's protocol. The ligated products were transformed into competent cells.

### 2.2.4 DNA cloning using Gibson assembly

Molecular cloning of TR test cassettes into psiCHECK™-2 vector, homology arms (HA) and template for recombination (TfR) into pHD-DsRed was achieved by isothermal assembly (Gibson assembly) (Gibson et al., 2009). Insert sequences were amplified from gDNA or cDNA using primers containing 18 bp overhangs that overlap with the blunt-ends of PCR amplified linearized vectors. Isothermal assembly was performed by incubating the purified inserts and 100 ng linearized vectors in a molar ratio of 3:1 with 'lab made' Gibson assembly mix for 1 h at 50 °C in a total volume of 15 µl. 1 µl of end product was transformed into competent cells.

### 2.2.5 Bacterial transformation

50 µl of chemically competent NovaBlue cells were used for transformation. Thawed cells were incubated with 1 to 5 µl of PCR, ligation or Gibson assembly products on ice for 20 min. The cells were heat shocked by incubating at 42 °C for 45 s. The transformed cells were kept on ice for 2 min, supplemented with 450 µl Luria-Bertani (LB) media and incubated at 37 °C for 1 hour. 100 µl of the culture was plated on LB agar containing the appropriate selection antibiotic. Positive colonies were screened via plasmid DNA sequencing.

## 2.3 Luciferase assay

### 2.3.1 Construct design for readthrough reporter assay

To quantify the TR efficiency in putative candidate genes, an *in vitro Drosophila* Schneider 2 (S2) cell culture-based dual reporter assay was developed. psiCHECK™-2 (Promega) vectors contain Renilla (Rluc) and Firefly (Fluc) luciferases under the control of SV40 and HSV TK promoters respectively. The commercial vector was modified by deleting the Rluc Poly(A) signal and the Fluc HSV TK promoter using blunt-end ligation, such that both the reporters are transcribed as a single transcriptional unit from a monocistronic mRNA controlled by SV40 promoter. A self-cleaving P2A

sequence (66 bp) was inserted between the two reporters using two-step insertional blunt-end ligation mutagenesis. The start codon of Fluc was deleted using blunt-end ligation mutagenesis. The primers used for vector modification are listed in Table 8. TR motifs (105 bp) from the candidate genes containing the leaky stop codons were amplified from $w^{1118}$ cDNA using primers with 18 bp overhangs and inserted into modified linearized vector using Gibson assembly. UGA to UUC and UAAA point mutations were introduced into each construct using blunt-end ligation mutagenesis. The primers for cloning of TR motifs and point mutations are listed in Supplementary Table S1.

Table 8. List of primers used for psiCHECK™-2 vector modification

| Name | Primer Sequence 5' to 3' | Comments |
|---|---|---|
| CM135_F | ATGGCCGATGCTAAGAACATTAAG | Amplification of psiCHECK™-2 template at Fluc start. |
| CM138_R | GTTGGTGGCGCCGGAGCCCTGCTCGTTCTTCAGCAC | Insertion of 1-18 bp of P2A after Rluc stop. |
| CM141_F | GAGGAGAACCCCGGCCCCATGGCCGATGCTAAGAAC | Insertion of 49-66 bp of P2A before Fluc start. |
| CM146_F | CAGGCCGGCGACGTGGAGGAGAACCCCGGCC | Insertion of 34-48 bp of P2A. |
| CM147_R | CTTCAGCAGGGAGAAGTTGGTGGCGCCGGAGCC | Insertion of 19-33 bp of P2A |
| PK27_F | GCCGATGCTAAGAACATTAAGAAGGGC | Deletion of Fluc AUG from modified psiCHECK™-2 P2A constructs |
| PK28_R | GGGGCCGGGGTTCTCC | |

### 2.3.2 Maintenance of *Drosophila* Schneider 2 cells

*Drosophila* S2 cells were cultured in 25 $cm^2$ flasks at 25°C in a $CO_2$ incubator in Schneider's *Drosophila* medium (Gibco®), supplemented with 10% heat-inactivated fetal bovine serum (GE healthcare), 100 units/ml penicillin and 100 µg/ml streptomycin (Gibco®) in a total volume of 6 ml. The cells were passaged every 3-4 days by splitting them in 1:6 ratio.

### 2.3.3 Transfection of S2 cells

Prior to transfection, S2 cells were split 1:6, and 150 µl of the cell suspension was seeded into flat bottomed 96-well plates (Sigma Aldrich) and incubated overnight. Transfection reactions were prepared using Effectene® Transfection Kit (Qiagen) according to the manufacturer's protocol. 100 ng of dual reporter construct was used to transfect each well. Transfections were carried out when the cells reached approximately 70% confluency. The cells were incubated for 72 hours before measuring the reporter activities. Transfections were performed in three technical replicates. After 72 hours, cells from the 96 well plates were resuspended and transferred to microfuge tubes and pelleted at 800 xg. The cell pellets were shock frozen in liquid nitrogen and stored at -80 °C.

### 2.3.4 Measurement of luciferase reporter activity

For measurement of luciferase activity, cell pellets were thawed and lysed using 40 µl Lysis-juice (PJK GmbH). To measure Fluc activity, 20 µl of cell lysate was mixed with 100 µl of Beetle juice and incubated for 10 min at room temperature (RT). To measure Rluc activity, 20 µl of same cell lysate preparation was mixed with 100 µl Renilla Glow juice and incubated for 10 min at RT.

Measurements were performed in a luminometer with a delay time of 2 s and an integration time of 5 s. The background luminescence obtained from cell lysates prepared from S2 cells transfected with empty transfection mixes was subtracted from the raw readouts of the luminescence signals. The ratio of Fluc:Rluc signal was calculated for each construct containing the native and the UAA-A stop codon context in their TR motif, as well as for the corresponding constructs where the stop codon is mutated to UUC. To calculate TR efficiency of test constructs with native and UAA-A stop codon context, their respective Fluc:Rluc values were divided with Fluc:Rluc values of constructs containing UUC codon, which serve as positive controls. Non-paired two-tailed Student's t-test was used to analyze the results.

## 2.4 *Drosophila* handing and maintenance

### 2.4.1 Stock maintenance

Fly stocks were maintained in vials containing cornmeal-agar prepared with 6.25 g/L agar (Serva), 18g/L dry yeast (Saf-Instant), 80 g/L corn flour (Zieler & Co.), 22 g/L beet syrup (Ferdinand Kreutzer Sabamühle GmbH), 80 g/L malt (Ulmer Spatz), 0.625 % propionic acid (Merck), 1.5 g/L methylparaben (Sigma). Flies were kept in a controlled environment with constant temperature of 25 °C, constant humidity and 12 hr-12 hr light-dark cycle.

### 2.4.2 Creation of transgenic flies

CRISPR/Cas9-based genome editing was employed to create three different genetic mutants of *D. melanogaster* that harbor mutations in and around the stop codon of the *traffic jam* (*tj*) gene (sequence location 2L:19,64,267 to 19,467,758), using the RH genome engineering services offered by Best Gene Inc. Chino Hills, CA, USA for injection. The following constructs were designed to introduce the desired mutants:

a) Two pU6-BbsI-chiRNA vectors (Addgene) expressing the guide RNA (gRNA) scaffold with 5'-appended protospacer adjacent motif (PAM) sites under the control of a DmU6 promoter. The gRNA scaffold primes to specific PAM sites in the fly genome and recruits Cas9 nuclease that introduces a double stranded break in the chromosome.
b) pHD-DsRed vector (Addgene) containing two ~1 kb long homology arms (HA) derived from *Drosophila* gDNA that flank the region that provides the template for recombination (TfR). The TfR harbors all the desired mutations and also includes a loxP-flanked DsRed-SV40 poly(A) region that enables the screening of recombinants by eyes fluorescence.

The fly strain used for injection had the following genotype: y1w1118; attP2(nos-cas9)/TM6C, Sb Tb. These flies express Cas9 protein under the control of *nos* regulatory sequences inserted at the chromosome III and bear white eyes.

### 2.4.3 Construct design for CRISPR/Cas9 injection

The CRISPR target finder tool (http://tools.flycrispr.molbio.wisc.edu/targetFinder/) was used to find optimal CRISPR target sites or PAM sites on the *tj* gene that flank the readthrough region between the first and the second stop codon of the *tj* ORF. The proximal PAM site was 5' AGAGCTTT|GGCTATCGCCGC CGG 3' and the distal PAM site was 5' ACACAATG|TATAAGGTAAAT TGG 3', where the NGG motifs are highlighted in bold. The 20 bp proximal and distal PAM regions were introduced separately upstream of gRNA scaffold in separate pU6-BbsI-chiRNA vectors via blunt-end ligation mediated insertional mutagenesis using primer pairs PK241_F/PK243_R and PK242_F/PK243_R respectively (Table 9).

pHD-DsRed vectors carrying the homology arm 1 (HA1), the TfR and the homology arm 2 (HA2) were generated in subsequent steps using Gibson assembly. HA1 (1100 bp) + TfR (250 bp) was amplified from gDNA obtained from $w^{1118}$ as a single fragment and inserted upstream of loxP-DsRed-SV40poly(A)-loxP sequence. HA2 (1144 bp) was amplified and inserted immediately downstream of this sequence. QuikChange mutagenesis protocol was used to introduce synonymous mutations into the proximal PAM sequence that borders HA1 and TfR, in order to prevent Cas9 from cleaving the vector once injected into the embryos. UGA to UUC mutation was then introduced in the TfR at the *tj* stop codon by QuikChange mutagenesis and 3xUAA was inserted downstream of the *tj* stop codon by blunt-end ligation method. 3xFlag was inserted upstream of the second stop codon by Gibson assembly. These cloning steps were performed in pHD-DsRed vector where loxP1 site had been deleted in order to avoid complications associated with redundant primer binding sites. Finally, the loxP1 site was reinserted. Additionally, the dispensable phage pC31 attP site was removed from the pHD-DsRed vector during PCR amplification. The primers used for Gibson assembly, point mutations and blunt-end ligation cloning are listed in Table 9. Due to the introduction of an independent SV40 transcription termination signal in the TfR, the biogenesis of *tj* derived piRNAs in the CRISPR-derived recombinants is inhibited. To overcome this limitation, the loxP-flanked *DsRed*-SV40 poly(A) marker cassette was removed by Cre-Lox recombination, which restored the native *tj* 3' UTR in *tj*-TR mutants ($tj^{mut}$). The introduction of the desired mutations was verified via sequencing of a genomic DNA derived amplicon. *DsRed* deletion was confirmed via screening of eyes for negative fluorescence as well as sequencing.

Table 9. List of primers used for preparing constructs for CRISPR/Cas9 injections

| Name | Primer sequence 5' to 3' | Comments |
|---|---|---|
| PK241_F | GAGAGCTTTGGCTATCGCCGCGTTTTAGAGCTAGAAATAGC | Insertion of proximal PAM site at 5' end of gRNA scaffold in pU6-BbsI-chiRNA vector |
| PK242_F | GACACAATGTATAAGGTAAATGTTTTAGAGCTAGAAATAGC | Insertion of distal PAM site at 5' end of gRNA scaffold in pU6-BbsI-chiRNA vector |
| PK243_R | GAAGTATTGAGGAAAACATA | Reverse amplification of pU6-BbsI-chiRNA for PAM insertion |
| PK77_F | ATATGCACACCTGCGATCGGTGAACACATCTTCGGG | Amplification of HA1+TfR from gDNA with 18 bp overhangs for Gibson assembly |
| PK135_R | TGAATTAGATCCCGTACGTACCTTATACATTGTGTCTAGGA AAAGC | |
| PK132_F | CGTACGGGATCTAATTCAATTAGAGACTAATTCAATTAGAG | Amplification of pHD-DsRed without loxP1 and attP site for HA1+TfR insertion |
| PK71_R | GATCGCAGGTGTGCATATGTCCG | |
| PK153_F | TAAGTAGAGAGCGTTCCGTGTTTAAGG | Amplifcation of pHD-DsRed for 3xFlag insertion |
| PK154_R | GTTGACCAGCTGCTGGGGATTC | |
| PK147_F | CCCCAGCAGCTGGTCAACGACTACAAGGACCACGACGGTG ACTACAAGGACCACGACATCGACTACAAGGACGACGACGA CAAGTAAGTAGAGAGCGTTCCG | Insertion of 3xFlag upstream of tj second stop codon by Gibson assembly |
| PK148_R | CGGAACGCTCTCTACTTACTTGTCGTCGTCGTCCTTGTAGTC GATGTCGTGGTCCTTGTAGTCACCGTCGTGGTCCTTGTAGT CGTTGACCAGCTGCTGGGG | |
| PK207_F | TGTATGCTATACGAAGTTATAATTGGTTTGATTCCAAGAAT GTTTTC | Amplification of HA2 from gDNA with 18 bp overhangs for Gibson assemby |
| PK208_R | ATCTTTACTAGTGCTCTTCTCGCGTGTGTTTCTTCTAG | |
| PK209_F | AGAAGAGCACTAGTAAAGATCTCCATGC | Amplification of pHD-DsRed for HA2 insertion |
| PK210_R | ATAACTTCGTATAGCATACATTATACGAAGTTATACCG | |
| PK83_F | CAACCGCGGGCGGAGATAGCCAAAG | Introduction of synonymous point mutations at proximal PAM site in the HA1+TfR cloned pHD-DsRed vector by QuikChange mutagenesis |
| PK84_R | CTTTGGCTATCTCCGCCCGCGGTTG | |
| PK21_F | GGAATTCTACCTCTTCCGCCAGCTGGCGG | Mutation of *tj* stop codon UGA to sense codon UUC coding for Phe by QuikChange mutagenesis |
| PK22_R | CCGCCAGCTGGCGGAAGAGGTAGAATTCC | |
| PK229_F | TAATAATAACGCCAGCTGGCGGTGG | Insertion of UAAUAAUAA after tj stop codon UGA by blunt-end ligation |
| PK230_R | TCAGAGGTAGAATTCCGGAGAGCTTTGGC | |
| PK277_F | CGAAATCTAAGAAACCGGCATCGAAG | Generation of gDNA amplicon for sequencing |
| PK278_R | GGTGGTAATGGGAATGCACTTCTCTTG | |
| PK279_F | GCGACGCACCCTGAAGAATCG | Sequencing primer for genotyping mutants |

### 2.4.4 *Drosophila* genetics

DsRed-positive CRISPR mutants were crossed with Sco/CDY balancer lines to obtain CDY balanced mutant lines for second chromosome. $tj^{mut}$(+DsRed)/CDY lines were crossed with Sco/Cre lines in order to achieve Cre recombinase mediated removal of DsRed marker. The progenies, $tj^{mut}$(±DsRed)/Cre were back crossed with Sco/CDY balancer lines to obtain DsRed deleted $tj^{mut}$/CDY flies that served as stocks. DsRed deletion was confirmed by screening individual balanced flies for the absence of DsRed. The $tj^{mut}/tj^{mut}$ obtained by back crossing of $tj^{mut}$/CDY flies were used for experimental purposes. gDNA was extracted from the homozygous mutant flies (section 2.5.1). Using it as template, the genomic region flanking PAM sites was amplified using primers PK277_F and PK278_R and sequenced using primers PK277_F and PK279_F to confirm the introduced mutations (Table 9). $w^{1118}$ flies were used as wild-type controls as they have the closest genetic background to the mutants. The fly stocks used are listed in Table 10.

Table 10. List of fly stocks

| Short name | Genotype | Ref./ Source |
|---|---|---|
| wt-control | *w[1118]; +/+; +/+* | VDRC60000 |
| Sco/CDY | w[*]; sna[Sco]/CyO, P(w[+mC]=Dfd-EYFP)2 | BDSC 8578 |
| Sco/Cre | y[1] w[67c23]; sna[Sco]/CyO, P(w[+mC]=Crew)DH1 | BDSC1092 |

## 2.5 Gene expression analysis

### 2.5.1 gDNA extraction

Standard miniprep protocol for *D. melanogaster* genomic DNA extraction was followed as described in Huang et al. (2009). 10 adult male flies were used for each extraction.

### 2.5.2 RNA extraction and cDNA synthesis

Total RNA was extracted from heads and ovaries of 3-4 days old adult flies of each genotype. 15-20 heads and five pairs of ovaries were dissected in cold PBS and immediately transferred to 200 µl TRIzol™ reagent and homogenized. Total RNA extraction was performed using manufacturer's protocol. Extracted RNA was quantified and treated with DNaseI (2 units per µg of RNA). Total cDNA was prepared using random primers with High Capacity Reverse Transcriptase following manufacturer's instructions. 20 µl RT reactions were set up for 1 µg RNA template.

### 2.5.3 Real-time quantitative PCR (RT-qPCR)

RT-qPCR was performed for gene expression analysis using Quantitect SYBR Green PCR kit. Each reaction was performed in 15 µl volume using 20 ng cDNA template and 200 nM primers. All reactions were performed in triplicates. Control reactions were set up for each target gene using non-RT templates. The primers used for each transcript quantification were obtained from DRSC primer bank and are listed in Table 11.

The qPCR reaction conditions used were according to manufacturer's instructions. *αTub84B* was used as endogenous control and $w^{1118}$ flies were used as control samples. The analysis of the acquired threshold cycle ($C_T$) values was performed using StepOne Software. $C_T$ value denotes the fractional cycle number at which the fluorescence signal for each test sample passes a defined threshold. Average $C_T$ values from three technical replicates of respective genes were subtracted from that of the *αTub84B* control to obtain $\Delta C_T$. These $\Delta C_T$ values for each gene was normalized again by subtracting the $\Delta C_T$ of the control sample from the $\Delta C_T$ of the test sample. The ΔΔCT values thus obtained were used to calculate gene expression levels by using the formula $RQ = 2^{-\Delta\Delta CT}$. Non-paired two-tailed Student's t-test was used for calculating p values.

Table 11. List of primers used for qPCR

| Flybase ID | Gene | Primer sequence (5' to 3') |
|---|---|---|
| FBgn0000964 | *tj* | Forward: GGCGGTTAAATGGACGACAAT<br>Reverse: AAGGACCTCAGCTTGATGTGC |
| FBgn0015609 | *CadN* | Forward: CCCCGCTTGTCAATCGTCC<br>Reverse: CCACTGGCATCGAGTAGGGA |
| FBgn0003391 | *DEcad* | Forward: CCATCAGCCTTCTGTCACCC<br>Reverse: CCTCTTTGACCTTAGGAGCGTAT |
| FBgn0000636 | *Fas3* | Forward: ATATGTCTCGCAGCCATCTTAAC<br>Reverse: GGATCTGCCGTACCGACAG |
| FBgn0004108 | *Nrt* | Forward: GTCTCAAACTGCCTGGTTTCT<br>Reverse: TCTCCATTGGCTTCCTTATCCT |
| FBgn0031424 | *VGlut* | Forward: CCTTCGGCATGAGGTGCAATA<br>Reverse: CGAGTCCACATGGCTCTCC |
| FBgn0003884 | *αTub84B* | Forward: CACACCACCCTGGAGCATTC<br>Reverse: CCAATCAGACGGTTCAGGTTG |
| FBgn0033257 | *sand* | Forward: GGTTTATAGCACGGAACTTCAGT<br>Reverse: GGTGGTCGAAGAAGCTGATGT |
| FBgn0035789 | *mthl6* | Forward: AACTGGCGGATGGATCACAG<br>Reverse: AGCAGAATCGGATACAAGGTTTT |
| FBgn0027348 | *bgm* | Forward: TGGACAAGATTCACGCCATTC<br>Reverse: CGACCACCTGTAGTAGCCATC |
| FBgn0032706 | *Irk3* | Forward: CTGCCACGGATTCCCTAACC<br>Reverse: CCGTCTCCTTTTCGGAGGAAC |
| FBgn0002939 | *ninaD* | Forward: TGTGGGGTGACCCAACAAAAG<br>Reverse: CCCTGAGTCTATAAAGCCAGGC |
| FBgn0039678 | *Obp99a* | Forward: TTGCCATCTGCGTGCTGATT<br>Reverse: TTGGGGTACTCCCACTTCTGG |
| FBgn0038914 | *fit* | Forward: ATGCACAAGGTGGATAAGCGA<br>Reverse: ACTGCCTCAACTGATTGACGG |
| FBgn0010381 | *Drs* | Forward: CTGGGACAACGAGACCTGTC<br>Reverse: ATCCTTCGCACCAGCACTTC |
| FBgn0260446 | *GABA-B-R1* | Forward: AACCGCAAAAGCTGATGCTG<br>Reverse: CCGTAGCAGAGCACAATTAGATT |
| FBgn0053310 | *CG33310* | Forward: GAGCAACGCGAATCAACTAACG<br>Reverse: ATCTTGGAACCCTTCACTTCATC |
| FBgn0053200 | *VepD* | Forward: CCAGGAACATACACGCTCCAC<br>Reverse: CAAGGGCCTCCCAGTGAAG |
| FBgn0029823 | *Shmt* | Forward: CTTGACGCACGGTTTCTTCAC<br>Reverse: TCTCCGGGTTCACTTTGTACG |
| FBgn0001187 | *Hex-C* | Forward: CCCGGTGTGGACCTATTCG<br>Reverse: GTGGCAGATATGCGGTCTTCA |
| FBgn0044810 | *TotX* | Forward: AGAGAATACCGGGCAGTTTTG<br>Reverse: ATATACCGGGTTCCGACTCTG |

## 2.6 Immunohistochemistry

### 2.6.1 Embryo collection

Flies were kept in embryo collection cages and placed on apple juice-agar plates, smeared with fresh yeast paste at the center. The agar plates provide a substrate for egg laying and fresh yeast promotes the process. The flies were kept in the cages for 2-3 days at 25 °C to allow them to adjust. After the habituation period, the agar plates were replaced with fresh ones to start embryo collection. To obtain embryos staged 15-16 (12-16 h), timed collection was carried out for 4 hours after which the agar plates with eggs were removed from the collection cages and set aside for 12 hours at 25 °C. The egg-laden plates were then cleared off any dead flies. The eggs were gently dislodged from the plates and transferred to a microcentrifuge tube containing wash buffer (120 mM NaCl, 0.03% TritonX). Embryos were rinsed two times with wash buffer and two times with water. 50% commercial bleach was added for 2-3 min. Wash buffer was added until the embryos started to sink. Two more washes were performed with water. 4% paraformaldehyde (in PBS): n-Heptane (1:1) was added to the embryos and shaken vigorously for 30 s. Embryos were then placed on a rotating wheel for 25 min to allow fixation. The lower phase containing the fixing solution (4% paraformaldehyde) was removed and equal amount of methanol was added and shaken immediately for 15 s. The upper liquid and floating embryos were removed. The embryos were rinsed with methanol two more times. The fixed embryos were stored at -20 °C and rehydrated by washing with PBT.

### 2.6.2 Tissue dissection and antibody staining

Tissue dissection was performed in cold PBS. All tissue samples, brains, testes and ovaries were collected from 3-4 days old adult flies that were kept on food vials containing yeast. Dissected tissues were fixed in 4% formaldehyde in PBT (pH 7.4). Ovaries were fixed for 13 min on high speed rotator, brains were fixed for 25 min on nutator while testes were fixed for 20 min horizontally on bench top. Post fixing, the tissues were washed in PBT. Washed tissues were blocked in PBTB for 1 h at room temperature. Incubation with primary antibodies (diluted in PBTB to desired concentrations) was allowed overnight at 4 °C on a nutator. The specimens were then washed with PBT, followed by 1 hour incubation with PBTB at RT. Secondary antibody solution (diluted in PBTB) was added and incubation was allowed overnight at 4 °C in a nutator. The samples were washed in PBT, followed by a 10 min wash in 10 mg/l DAPI (Sigma) in PBT. The samples were washed in PBT and mounted in VECTASHIELD medium (VectorLabs) on glass slides (76x26 mm, Thermo Scientific) and imaged. The primary and secondary antibodies used for immunohistochemistry studies are listed in Table 12.

Table 12. List of antibodies used for immunohistochemistry

| Type | Specificity | Dilution | Source |
|---|---|---|---|
| Guinea pig | Anti-Tj | 1:10,000 | D. Godt, University of Toronto, Canada |
| Rabbit | Anti-Vasa | 1:5000 | Herbert Jäckle, MPI-bpc, Göttingen |
| Mouse M2 | Anit-Flag | 1:500 | Sigma Aldrich |
| Secondary goat | Anti-guinea Alexa 647 | 1:500 | Life Technologies, A-21450 |
| Secondary goat | Anti-mouse Alexa 488 | 1:500 | Molecular Probes |
| Secondary goat | Anti-rabbit Alexa 568 | 1:500 | Molecular Probes |

### 2.6.3 Imaging

Fluorescence images were taken using a confocal laser-scanning microscope (Zeiss LSM 700). The images were processed with Fiji (ImageJ) and Adobe Illustrator.

# 3. RESULTS

## 3.1 Dual luciferase reporter construct design

Dual luciferase reporter assays have long been used as a method to study translational recoding in different systems (Grentzmann et al., 1998; Harger and Dinman, 2003; Sherf et al., 1996). We developed a psiCHECK™-2-based dual luciferase reporter assay to quantify TR efficiency in *D. melanogaster* genes *in vitro* using S2 cell lines. S2 cells are commonly used *Drosophila* cell lines that are derived from primary culture of late stage embryos and can be subjected to transient transfection with several vectors. We modified the psiCHECK™-2 vector by deleting the *Rluc* poly(A) signal, the promoter of the *Fluc* gene and the start codon of *Fluc* ORF to generate a single ORF that codes for both luciferases (Fig. 16B). In traditionally used dual luciferase reporter constructs, a positive recoding event produces a minor ratio of Rluc-Fluc fusion protein, whereas a standard decoding event produces native Rluc protein. Rluc enzymatic activities in such constructs arise from a mixture of native Rluc and Rluc-Fluc fusion proteins. Normalization with Rluc activity from constructs with constitutive recoding event eliminates the probability of errors that arise from differential transfection efficiency, technical experimental errors and cell viability. However, the differences in reporter activity in a fused protein product and the contribution of translated test sequences on enzymatic activities might lead to discrepancies in data calculation and interpretation. To address these issues, we inserted a self-cleaving P2A peptide (Donnelly et al., 2001; Doronina et al., 2008; Ryan and Drew, 1994), upstream of *Fluc* reporter or downstream of test sequences to allow the expression of luciferase genes as independent polypeptide products (Fig 16B).

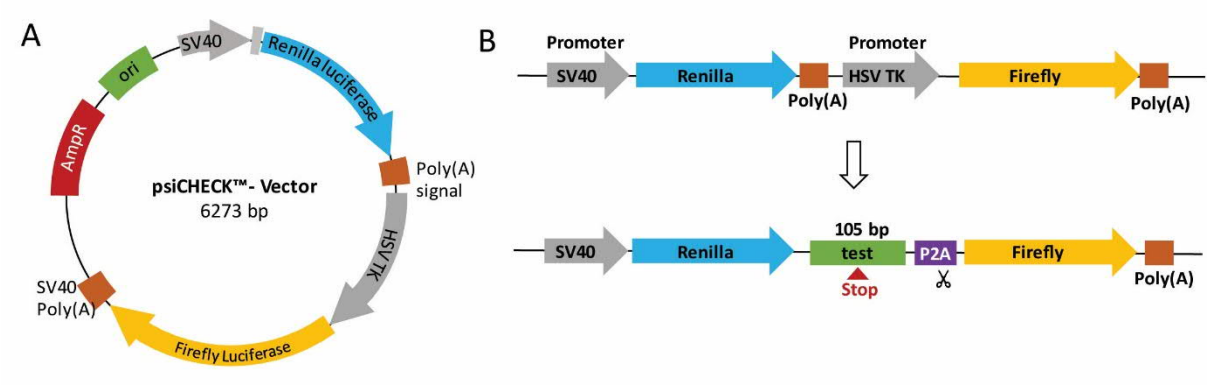


Figure 16. Dual luciferase reporter constructs for TR quantification in S2 cells.
(A) Plasmid map of commercial psiCHECK™-2 vector coding for *Renilla* and *Firefly* luciferases under the transcriptional control of SV40 and HSV TK promoters, respectively. (B) Modification of psi-CHECK™-2 vector to replace HSV TK promoter with TR test sequences and self-cleaving P2A sequence.

## 3.2 Quantification of TR in phylogenetically predicted genes in *Drosophila*

Phylogenetic analyses have predicted a total of 283 TR candidates in *Drosophila* (Jungreis et al., 2011), the majority of which have not been experimentally verified. We narrowed down our study to a set of 11 candidate genes that perform biologically significant functions during fly development. These candidates have varied TR length, ranging from 11 to 236 amino acids, several of which possess distinct peptide profiles (Table 13). With the only exception of *wishful thinking* (*wit*), the 3' UTR of these genes do not form secondary structures; furthermore, their gene functions have been well characterized and are associated to traceable phenotypes. Among the selected genes, *klumpfuss* (*klu*), *doublesex* (*dsx*), *traffic jam* (*tj*), *seven up* (*svp*), *chronologically inappropriate morphogensis* (*chinmo*), *fruitless* (*fru*) and *broad* (*br*) encode transcription factors or transcriptional regulators. *wit*, *atypical protein kinase C* (*aPKC*) and *discs large 1* (*dlg1*) encode protein kinases involved in cell signaling. *Kinesin heavy chain 73* (*Khc-73*) encodes a motor protein that regulates cell polarity.

Table 13. List of TR candidates selected for TR validation

| Gene | TR length, codons | Region Profile | Peptide feature | Expression |
|---|---|---|---|---|
| *br* | 131 | Ala/Gly/His rich | Disordered | Embryonic/larval CNS |
| *klu* | 15 | - | - | Embryonic neuroblasts, larval CNS |
| *chinmo* | 236 | Thr rich | BTB domain, disordered | Embryonic/larval nervous system, eye disc, adult testes |
| *wit* | 10 | - | - | Embryonic/larval/adult CNS, mid gut, eye, salivary gland |
| *dsx* | 23 | - | Non-cytoplasmic/ signal peptide | Embryonic gonad, embryonic/larval/adult CNS, testis |
| *Khc-73* | 58 | - | - | Enriched in larval/pupal CNS, ubiquitous |
| *fru* | 187 | Gln/Asn rich | Polar, disordered | Ubiquitous in embryos, larval/pupal/adult CNS |
| *svp* | 11 | - | | Embryonic neuroblasts, larval photoreceptor cells, fat body, adult optic lobe, photoreceptors |
| *aPKC* | 131 | Asn/Gln rich | Polar, disordered | Ubiquitous in early embryos, larval/pupal/adult CNS |
| *dlg1* | 41 | - | - | Embryonic/larval/adult CNS, salivary glands, fat bodies |
| *tj* | 44 | - | Disordered | Gonadal somatic cells, embryonic/larval CNS |

The expression patterns of the majority of these genes indicate that their gene products are enriched in embryonic neuroblasts as well as specific neurons and glia of the CNS (central nervous system) and PNS (peripheral nervous system) from larval to adult stages (Gelbart and Emmert, 2013). These genes perform important functions associated with the maintenance of neuroblast polarity during proliferation and self-renewal, synaptic homeostasis, neurotransmission, neurogenesis, etc. (Table 13).

We created dual luciferase reporter constructs for the 11 candidate genes in order to validate and quantify TR *in vitro* in S2 cell lines. The sequence comprising the test cassette cloned into the dual luciferase reporter vector are listed in Table 14. For each gene construct, constitutive TR controls were generated by mutating their native stop codons to UUC sense codon, coding for phenylalanine. Additionally, the +4 stop codon context for each of these genes were mutated to UAA-A to obtain constructs with highly efficient translation termination. *alpha-Tubulin 84B* (*αTub84B*), a globular protein that performs housekeeping functions related to cytoskeletal organization, was used as negative control. In our assay system, *αTub84B* undergoes TR with an efficiency of 0.35% (Fig. 17). Basal translation termination error rate of 0.02-1.4% has been observed in control non-readthrough reporter constructs in yeast and mammalian cells lines (Bonetti et al., 1995; Fearon et al., 1994; Firth et al., 2011; Keeling et al., 2004; Namy et al., 2002; Napthine et al., 2012). Mutating the native stop codon context of *αTub84B* from UAA-G to UGA-C did not increase TR values, indicating that the 105 bp *αTub84B* test cassette represents a robust sequence with efficient termination, independent of the immediate stop codon context.

Table 14. Test sequence of putative TR candidates cloned into dual luciferase reporter constructs
The primary TR motif is highlighted in red with the leaky stop codon in bold. Sequence involved in forming a potential secondary structure is highlighted in blue.

| Gene | Sequence |
|---|---|
| *Khc-73* | ATGACACGCTCCAAGAGCCGCGGCGATTCGCTAAACCTTTCGGCGCGTAAATGATTGTACCCAAAGTGTTCGCA TCAGCTGCAGCGTTGGACTAATTGCAGG |
| *fru* | CAATTGCAGCCGCAACACCACCAGCAGCAGCAGCACAATGCCACAAGTGAATGATACAGTCAGTACCTGGGCT GGAACTACGGCGCCTGGCGTCCCGATGACCCC |
| *chinmo* | GCAGCTGCAGCATCAGCAGCGGCTACAGCGGCAGCCAGCAATCATTCACCATAGGAAGCAGCCGCAACAGCAG CGGCATCAGTAGTCATGAACGCCAACAAGGAG |
| *wit* | CTGAGCTTGTACGACGATCGGATGATGGACTCCTCCCTGCTCAACATTCTCTAGCATGAGGAGGTTCTGCTGCTG GAGCTATAAACTAATCGAATGCAGCAGAAC |
| *dsx* | AACGGAGCCTACCACCACGGCCACCACCTGGTCAGCTCCACGGCTGCCACGTAGCAGTATCGCAACGTTGCTGC CGCCGTGGCAGCAGCAGCAGCGGCCGCTGTC |
| *br* | CAGCAGCAGCAACAGTCGTCGCCGGGCATTGTCAAACCGTGCATGGACTTCTTATAAGATCAGCAGCAACTCTT GCAGCAGCTGTTCCACGTGGCGCTCAACAACTCC |
| *klu* | CTAACCCTGGGTGGACCCATGCCACTGGTTGTCAAGACGGAGAGCGCCTAACGGTGTCTGTATGCAGCAGCATC CAGACCATTTATGACCTGT |
| *aPKC* | GAGTATGTAAACCCCTTGCTGATGTCTCTGGAGGATTGCGTCTGACACCACGAAATGTGCGACTTACATCCGTAT AACATGCGGCTCTATGGCGAG |
| *dlg-1* | ATTTGGTCCCAGTCGGGACCAACCATTTGGGTACCTTCCAAGGAATCTCTATGACCAACAGCCACCACAACTTGG ACACTGCCGCCTCGAGTTCGATGTCGACCA |
| *tj* | GGAGGTGCAACCGCCGGCGGCGATAGCCAAAGCTCTCCGGAATTCTACCTCTGACGCCAGCTGGCGGTGGCCA GCGGCTCCTCCGCATCCTCATCCTCCTCCTCG |
| *svp* | GATATGCTGCTGAGCGGCAACAGTTTCTCCTGGCCCTATCTGCCTTCGATGTGACACACGATGTGGCGCCAATTG ACAACAACT |
| ***α**Tub84B* | GAGGTCGGCATGGACTCCGGTGACGGCGAGGGTGAGGGCGCTGAGGAGTACTAAGCGTCACGCCACTTCAAC GCTCGATGGGAGCGTCATTGGTGGGCGGGG |

Among the candidate genes selected, three genes, *dsx*, *Khc-73* and *fru,* showed basal TR levels, indicating that the list of phylogenetically predicted TR genes contains false positives, or that the translational machinery in S2 cells does not accommodate specific cases of recoding that are spatio-temporally regulated in the living organism.

*br* and *klu* harbor a UAA stop codon followed by G and C, respectively. The dual luciferase assay showed that *br* undergoes TR with an efficiency of 2% and *klu* with an efficiency of 4.5% respectively. Mutating the stop codon context to UAA-A had minimal effect on the levels of TR in both the genes (Fig 17). The identity of the nucleotides beyond the +4 context in these cases probably modulates TR efficiency in highly sequence specific manner.

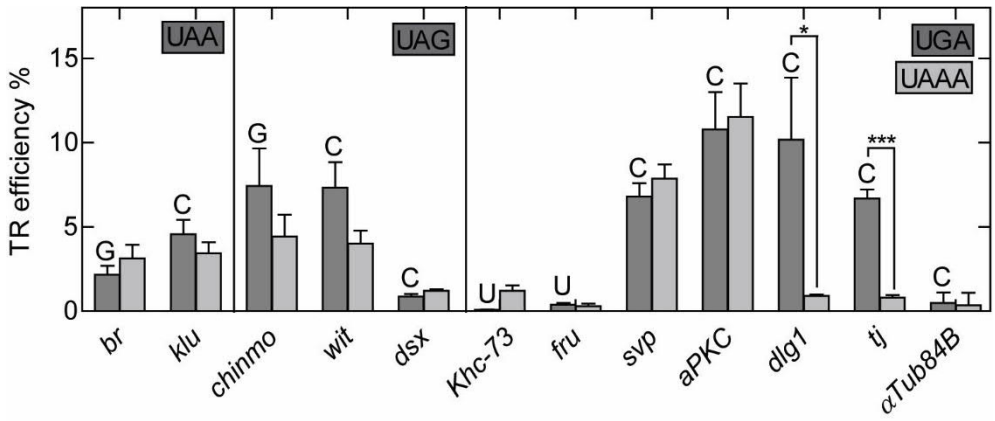


Figure 17. Translational readthrough efficiencies for putative candidate genes determined by dual luciferase reporter assay in S2 cells.
TR efficiencies of putative genes with UAA, UAG and UAA stop codons are represented by dark grey bars. The +4 nucleotide comprising the stop codon context for each gene is indicated by the letter above each bar. Light grey bars represent TR efficiencies for corresponding genes upon mutating the native stop codon context to UAAA. *αTub84B* represents a non-readthrough negative control. The bar graphs indicate (avg ± SD). p-values are calculated using two-tailed unpaired Student's t-test. (*p<0.05, ***p<0.0005).

Of the three genes containing a UAG stop codon, *chinmo* has the stop codon context of UAG-G and undergoes TR with an efficiency of 7.4%. Replacing UAG-G with UAA-A reduced TR in *chinmo* to 4.4 %. Possibly, the presence of -1 A (Table 14), together with the stop codon context, plays a role in modulating TR in *chinmo*. The presence of A in -1 and/or -2 position has been previously linked to high levels of TR (Cassan and Rousset, 2001; Mottagui-Tabar et al., 1998; Tork et al., 2004). UAG-C to UAA-A mutation reduced the TR values in *wit,* from 7.3 % to 4% (Fig 17). The 3'UTR of *wit* forms a potentially stable stem loop structure that is accommodated within the test cassette (Jungreis et al., 2011). Introduction of a strong stop codon context exerted only modest effects in lowering TR values in all genes ending with UAG.

We also studied TR in six genes containing UGA as stop codons. As expected, the four genes with a stop codon context of UGA-C: *svp*, *aPKC*, *dlg1* and *tj,* exhibited the highest TR efficiencies, ranging from 6.7% to 11%. Mutating the stop codon context to UAA-A in *dlg1* and *tj* abolished TR, indicating that the immediate nucleotide context is the only requirement to drive TR in these cases. TR levels in *svp* and *aPKC* were unaffected upon mutating UGA-C to UAA-A. TR in these two genes might be modulated by wider stretches of mRNA signals that extend beyond the +4 context.

## 3.3 Analysis of sequence determinants for TR in *aPKC*

Because the relatively high levels of TR in *svp* and *aPKC* were unaffected by the mutation to a strong stop codon context, we decided to look further into the sequence elements that promote TR in these candidates. Both *svp* and *aPKC* possess an in-frame AUG downstream of their primary stop codon at positions +10 and +13 respectively (Table 14). Although the phylogenetically predicted TR candidates were manually curated to rule out possible events of alternative splicing, RNA editing and internal ribosome entry site dependent initiation (Jungreis et al., 2011), we cannot exclude initiation at an in-frame AUG present in close proximity to the primary stop codon. We decided to focus on *aPKC* and performed systematic mutational analysis in order to delineate the primary sequence requirement responsible for efficient TR.

Nucleotide sequence-dependent effects on translation termination efficiency might originate from sequences upstream or downstream of the stop codon. Proximal 5' nucleotides (-1, -2 positions) potentially affect termination by conformational modification of the mRNA structure (Tork et al., 2004), whereas distal 5' nucleotides might affect termination by coding for nascent arrest peptides (3 to 99 amino acids) (Chiba et al., 2009; Cymer et al., 2015; Gumbart et al., 2012; Sarker et al., 2000). Specific cases of translational arrest have been identified in *TnaC*, *AAP*, *CMV* and *SAM-DC* uORF, where the nascent chain interacts with the components of the ribosomal peptide tunnel leading to ribosomal stalling at the stop codon (Ito and Chiba, 2013; Wilson et al., 2016). The influence from downstream nucleotides can originate from interactions of nucleotides within the mRNA channel with components of rRNA or recruitment of *trans* factors (Cridge et al., 2018; Eswarappa et al., 2014; Namy et al., 2001). In order to rule out what part of the mRNA is important for TR in *aPKC*, we created two chimeric dual luciferase constructs by alternatively replacing the 51-ntd long upstream and downstream segments of *aPKC* with that of *αTub84B* (Fig. 18A). In the first construct, we introduced the 5' nucleotide sequence *of aPKC* (-51 to -1) into the *αTub84B* control construct. The TR efficiency in the 5'-*aPKC*-STOP-*αTub84B*-3' construct was extremely low indicating that the nascent chain peptide comprising the C-terminus of *aPKC* does not play a role in modulating TR. Interestingly, the introduction of the 3' sequence (+4 to +54) of *aPKC* into the control construct containing the *αTub84B* 5' segment was able to induce efficient readthrough. The TR values obtained for the 5'-*αTub84B*-STOP-*aPKC*-3' constructs harboring UGA-C and UAA-A contexts were 5.8% and 7.4% respectively, confirming that TR is not due to the stop codon contexts in this particular context.

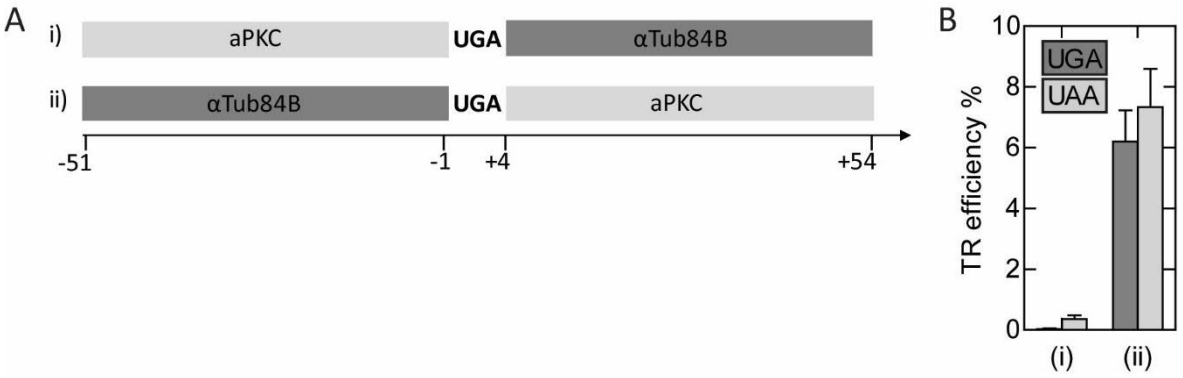


Figure 18. Analysis of the sequence determinants for TR in *aPKC*.
(A) Chimeric dual luciferase constructs with *aTub84B* TR motif containing (i) -51 to -1 *aPKC* sequence and (ii) +1 to +54 *aPKC* sequence. (B) TR efficiencies for constructs (i) and (ii). Dark grey bars represent TR efficiencies in constructs containing UGA stop codon and light grey bars represent TR efficiencies in constructs where the stop codon context is mutated to UAA-A. Error bars represent the SD of three technical replicates.

Next, we sought to delineate the minimal sequence required to induce TR in *aPKC*. We created sequential deletions in nucleotide the sequence downstream of the *aPKC* stop codon. Progressive deletions of +48, +30, +21, +18 and +15 nucleotides did not result in the attenuation of TR in the *aPKC* constructs with both UGA-C and UAA-A stop codon contexts, resulting in TR values of ~20% (Fig. 19A). This indicates that the distal downstream sequence does not affect translation termination in *aPKC*. Further truncation up to +12 position drastically reduced TR efficiency to 1.3 %. Stepwise truncations to +9 and +6 nucleotides showed basal levels of TR indicating that the minimal context to drive efficient TR in *aPKC* comprises a stretch of nucleotides downstream of the stop codon at positions +5 to +15.

A minimal +7 sequence context of UGA-CUAG drives efficient readthrough of up to 31% in mammals (Loughran et al., 2014). In yeast, a consensus sequence immediately downstream of the stop codon: -CARYYA (R: A/G, Y: C/U) is known to upregulate readhtrough (Namy et al., 2001). The +9 downstream sequence context of *aPKC*, UGA-CACCAC shows limited nucleotide conservation with known TR promoting primary motifs. The nucleotides up to +9 position occupy the mRNA channel and influence termination efficiency by potentially establishing interactions with the components of the translation machinery (Cridge et al., 2018; Namy et al., 2001). In the case of *aPKC,* the minimal determinant for TR extends beyond the +9 context and thus is exposed to the solvent front while the stop codon is positioned in the A site. In order to preclude the possibility of an initiation event at the in-frame AUG at position +13, we created point mutations at all three bases of the AUG codon. The TR values were unaffected by the identity of the codon (Fig. 19C), indicating that the high levels of Fluc expression are in fact solely due to TR and not because of an independent initiation at the in-frame AUG.

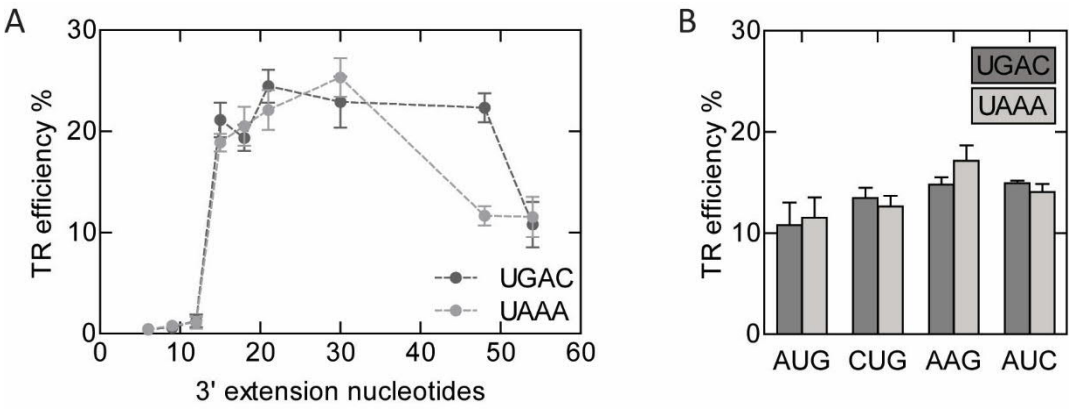


Figure 19. Delineating of the minimal sequence determinant for TR in *aPKC*.
(A) TR efficiencies for truncation variants of the dual-luciferase constructs containing the TR motif from *aPKC*. The X-axis represents the extension beyond the *aPKC* stop codon. Dark grey dots represent TR efficiency for constructs with the native *aPKC* stop codon context. Light grey dots represent the TR efficiency for constructs where UGA-C is mutated to UAA-A. (B) TR efficiency when the +13 AUG is mutated to CUG, AAG, and AUC respectively. Error bars represent the SD of three technical replicates.

In order to test whether the *aPKC* sequence downstream of the stop codon can independently induce TR in an otherwise stringent termination context, we introduced the +4 to +21 sequence from *aPKC* into *αTub84B* construct (Fig 20A). This short nucleotide stretch was able to induce efficient TR in the chimeric construct indicating that this sequence can independently regulate termination irrespective of the stop codon context, upstream and distal downstream sequences (Fig 20B). To further understand the mechanism of TR regulation by this sequence, we looked for potential RNA binding protein (RBP) binding sites within the +4 to +21 sequence in *aPKC* using RBPmap tool (http://rbpmap.technion.ac.il/index.html). We were able to map several short regions that served as potential binding sites for RBPs (Fig 20C). From our mutational analysis, the minimal sequence context driving efficient TR in *aPKC* is +15. RNA-binding protein 1 (RBP1) has a binding site spanning +5 to +11 region from aPKC. However, because the truncation variant comprising +12 *aPKC* sequence was unable to induce significant readthrough, the implication of a potential *trans* factor in TR modulation in *aPKC* is unlikely. The TR-inducing properties of the *aPKC* downstream sequence can be attributed to the conformational limitations exerted on the mRNA in the decoding center that interferes with stop codon recognition by eRF1. Structural studies would facilitate further understanding of the mechanistic principle of TR regulation by *aPKC* TR motif.

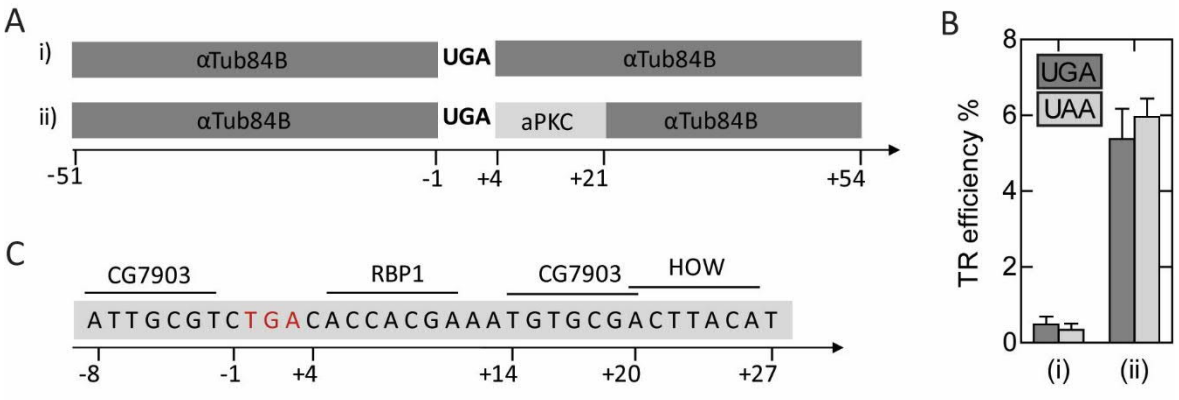


Figure 20. Induction of TR in ***α****Tub84B* by *aPKC* TR sequence.
(A) (i) Native TR sequence of *αTub84B*, (ii) *αTub84B* sequence where +4 to +21 nucleotides are replaced with that of *aPKC*. (B) TR efficiencies of constructs (i) and (ii). (C) Predicted binding sites for potential RNA binding proteins within the *aPKC* TR motif.

## 3.4 CRISPR/Cas9 induced mutations of *tj* stop codon in *D. melanogaster*

So far, very little is known about the physiological relevance of TR in the identified candidates in *D. melanogaster*. Based on sequence predictions, signal motifs such as nuclear localization signal, peroxisomal targeting signal 1 and transmembrane domain signal have been detected in the TR region of several candidates (Dunn et al., 2013). However, the significance of the appended TR motifs in biological contexts and how they affect the native protein function has not been studied. Because the stop codon context in *tj* seems to be the only element required to induce TR (Fig. 17), and because of the absence of introns in its coding sequence, which avoids the complications associated with genetic manipulation involving multiple splice isoforms, we chose *tj* as a candidate to study the biological relevance of TR *in vivo. tj* encodes a large Maf transcription factor whose role in gonad development has been well documented (Li et al., 2003; Panchal et al., 2017; Wingert and DiNardo, 2015). Furthermore, the phenotypes associated with the loss of function or hypomorphic alleles of *tj* have been described, making the effect of TR potentially interesting and phenotypically traceable. The TR sequence in Tj does not contain any conserved signal motifs.

We created three *D. melanogaster* mutants that harbor chromosomal mutations in and around the *tj* stop codon using CRISPR/Cas9-based genome editing (Fig. 21). All three mutants were designed to code for a 3xFlag epitope tag, upstream of the stop codon that terminates the Tj-TR isoform. The first mutation, $tj^{nat}$ does not alter the primary stop signal that terminates the *tj* ORF; TR event in this mutant is expected to occur at the same frequency as that of the native *tj*. The $tj^{TR}$ mutation replaces the primary *tj* stop codon with a UUC sense codon, such that the mutants produce only the 3xFlag-tagged TR isoform. Finally, the $tj^{nTR}$ mutation introduces multiple stop codons immediately after the primary *tj* stop codon, which leads to complete abolition of TR in *tj*.

The homozygotic flies for three genomic *tj*-TR mutations do not exhibit any defects in growth and viability.

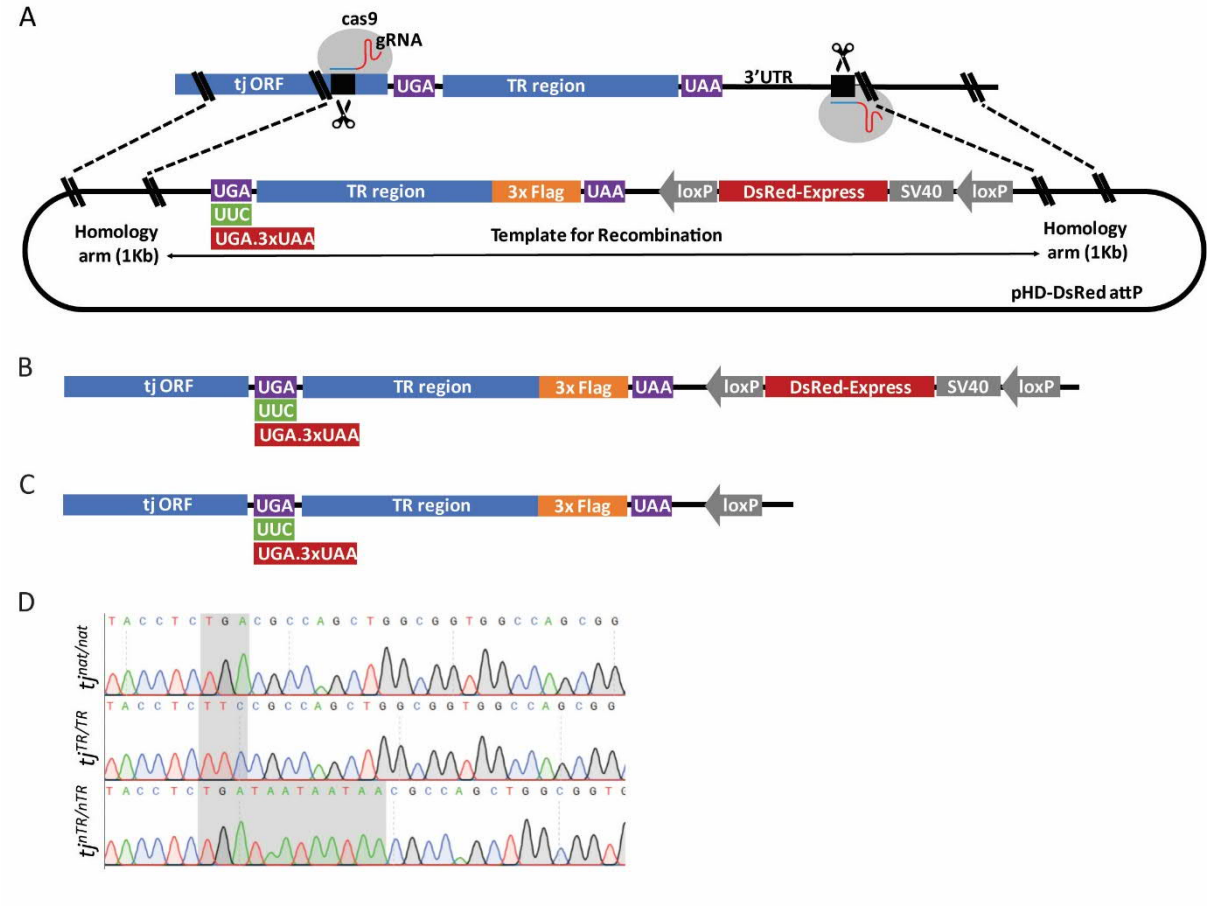


Figure 21. Construct design for CRISPR/Cas9 mediated genome editing to create *tj*-TR mutants
(A) Gene locus surrounding the TR region of *tj* with proximal and distal PAM sites for gRNA directed Cas9 cleavage is depicted above the modified pHD-DsRed attP vector containing the Template for Recombination (TfR) flanked by 1-Kb homology arms. Dotted lines indicate the region of homology between the gene locus and the modified vector. The TfR contains modifications to introduce the desired mutations at the primary *tj* stop codon, 3xFlag upstream of the second stop codon and a loxP-flanked *DsRed* marker. (B) Sequence depicting the modifications introduced in the *tj* locus in *tj*-TR mutants post CRISPR/Cas9 editing. (C) Cre recombinase-mediated removal of loxP flanked DsRed marker restores the native 3' UTR in *tj*-TR mutants. (D) Sequence verification of $tj^{nat}$, $tj^{TR}$ and $tj^{nTR}$ mutations in homozyotic mutant flies.

## 3.5 Tissue-specific regulation of TR in *tj* during embryogenesis

Expression of Tj begins at stage 12 of embryogenesis, when somatic gonadal precursors (SGPs) first establish contacts with the primordial germ cells (PGCs) (Li et al., 2003). Accumulation of Tj in SGPs has been used as a marker for somatic cell specification in stage 13-16 embryos (Jemc et al., 2012; Okegbe and DiNardo, 2011; Wingert and DiNardo, 2015). Disruption of the *tj* function in SGPs has been associated with defects in the arrangement of SGPs and PGCs that eventually lead to defective gonad development and sterility. In order to study the effect of TR on the SGP-specific function of *tj*, we analyzed embryonic gonads at stage 15-16. We stained embryos for Tj, Flag and

Vasa, an RNA helicase widely used as germline specific marker. In all three *tj*-TR mutants, the embryonic gonads coalesce into a round organ, as reported for wild-type embryonic gonads. Furthermore, the SGPs and PGCs are properly specified (Fig. 22).

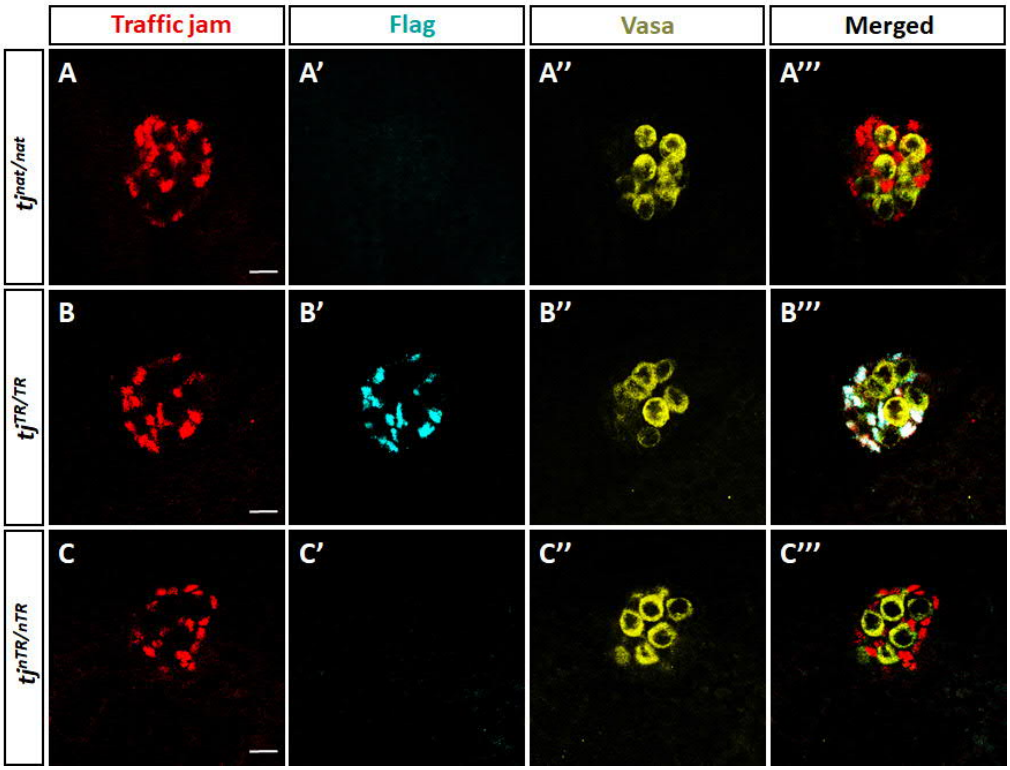


Figure 22. Embryonic gonad development in *tj*-TR mutants.
Stage 15-16 embryos immunostained with anti-Tj (red), anti-Flag (cyan) and anti-Vasa (yellow). (A-A') Under native stop codon context, $tj^{nat/nat}$ embryos express Tj in SGPs but not the Flag-tagged Tj-RT isoform. (B-B') $tj^{TR/TR}$ embryos exhibit constitutive expression of Tj-TR isoform. (C-C') $tj^{nTR/nTR}$ embryos only express native Tj in SGPs. PGCs in the embryos of all mutants express the germline marker Vasa (A'', B'', C''). All three mutants exhibit proper intermingling of SGPs and PGCs (A''', B''', C'''). Scale bars represent 10 µM in all panels.

Furthermore, the SGPs were nicely intermingled with the PGCs, indicating proper incorporation of PGCs into the embryonic gonads. Constitutive TR or complete abolition of TR did not affect the intermingling of the embryonic somatic and germ cells. Interestingly, we could not detect Flag expression in $tj^{nat/nat}$ embryos, indicating that *tj* does not undergo readthrough in the embryonic gonads. High termination fidelity during the translation of *tj* transcript in $tj^{nat/nat}$ embryonic gonads and the absence of phenotypic effect of TR abolition in $tj^{nRT/nRT}$ embryos demonstrate that TR in *tj* does not play a role in gonad development during embryogenesis. Furthermore, constitutive TR in *tj* does not affect the gonadal soma-specific *tj* function during embryogenesis.

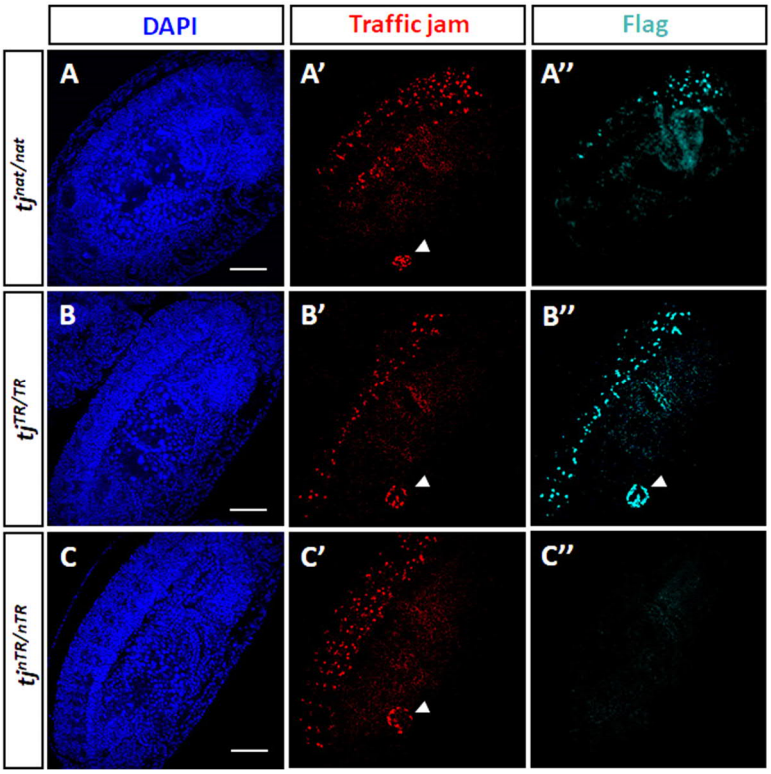


Figure 23. Tissue-specific regulation of TR in *tj* during embryogenesis.
Embryos at stage 15-16, stained with nuclear stain DAPI (blue), anti-Tj (red) and anti-Flag (cyan). (A-A'') Under native stop codon context, $tj^{nat/nat}$ embryos express Tj in the neural cells of embryonic VNC as well as SGPs of embryonic gonads (arrow) (A'). Flag tagged Tj-TR isoform is selectively expressed in embryonic VNC and excluded from SGPs in these embryos (A''). (B-B'') $tj^{TR/TR}$ embryos exhibit constitutive expression of Tj-TR isoform in VNC, as well as SGPs. (C-C'') $tj^{nTR/nTR}$ embryos only express native Tj in VNCs as well as SGPs (C') and do not express Tj-TR isoform in any embryonic tissues (C''). Scale bars represent 50 µm in all panels.

Expression of the *tj* transcript in the embryonic CNS has been observed in *in situ* hybridization studies of stage 14-16 embryos (Li et al., 2003). The embryonic CNS is composed of a ventral nerve cord (VNC) and the embryonic central brain. The VNC runs ventrally along the anterior-posterior axis. The function of Tj in CNS during embryogenesis is however largely unexplored. We analyzed the expression pattern of Tj in the embryonic CNS of *tj*-TR mutants. Tj is expressed in the nuclei of a subset of neural cells in the VNC along the cephalic, thoracic and abdominal region (Fig. 23), as well as in the central brain (data not shown). Interestingly, we could observe remarkable expression of the Tj-TR isoform in the VNC in $tj^{nat/nat}$ embryos (Fig. 23). Expression of the Tj-TR isoform is completely excluded from the SGPs in the embryonic gonads of the same mutants. $tj^{TR/TR}$ mutants constitutively express the Tj-TR isoform in the VNC as well as SGPs while $tj^{nTR/nTR}$ mutants only express the native Tj. Selective expression of the Tj-TR isoform in the CNS of $tj^{nat/nat}$ mutants suggests that TR in *tj* is regulated in a tissue-specific manner during embryogenesis.

## 3.6 Regulation of TR in *tj* in adult tissues

In order to assess whether the tissue-specific regulation of TR is maintained throughout development, we examined the brains of 3- to 4-days old adult flies. We could identify Tj-positive neurons in adult brains of all *tj*-TR mutants. Tj-positive neurons were scattered throughout the optic lobe with specific enrichment at the junction of the lobula and the protocerebrum. Tj-positive neurons were also scattered throughout the mid brain (data not shown). The neurons expressing Tj in the $tj^{nat/nat}$ brains also expressed the TR isoform, indicating that the nervous tissue-specific regulation of TR in *tj* is maintained until adulthood (Fig. 24). We could not detect any abnormal morphological or anatomical phenotypes associated with constitutive induction or abolition of TR in *tj* in *tj*-TR mutants.

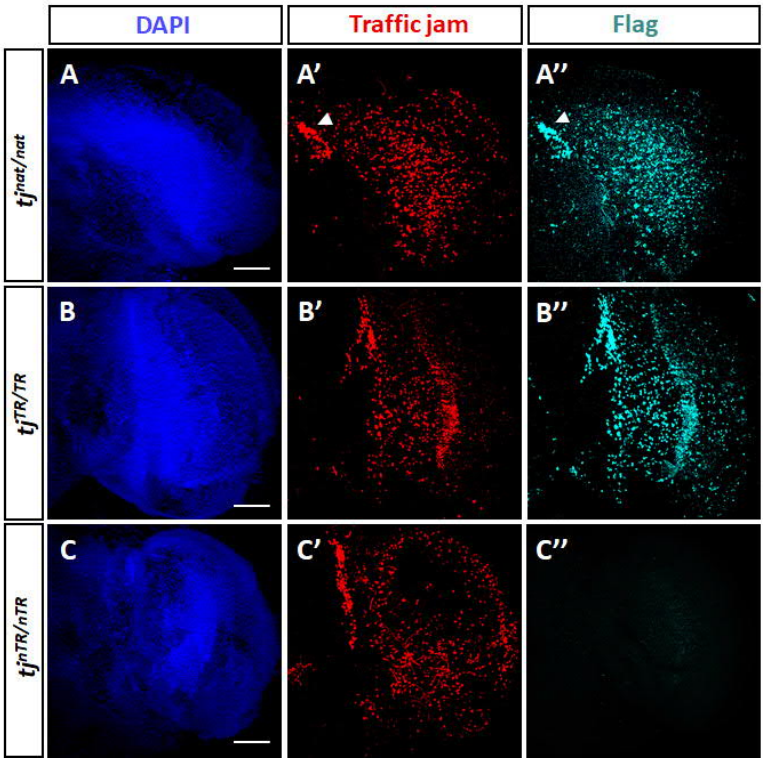


Figure 24. Regulation of TR in *tj* in the optic lobes of adult brains.
Optic lobes of adult brains immunostained with DAPI (blue), anti-Tj (red) and anti-Flag (cyan). (A-A'') $tj^{nat/nat}$ flies express Tj (A'), along with the flag tagged Tj-TR isoform (A'') in the neurons of the optic lobe. Tj-positive neurons are particularly enriched in the region between the lobula and the ventrolateral protocerebrum (arrow). (B-B'') Brains from $tj^{TR/TR}$ flies exhibit constitutive neuronal expression of Tj-TR isoform. (C-C'') $tj^{nTR/nTR}$ brains express only the native Tj protein (C') and do not express Tj-RT isoform in any neurons (C'') Scale bars represent 50 µm in all panels.

Next, we checked if the TR isoform of *tj* is expressed in the adult gonads. We first tested the fertility of mutant adult males by crossing them with females from a wild-type strain: all *tj*-TR mutant males were fertile. Furthermore, the testes did not exhibit any visible anatomical defects. Adult testes expressed Tj in the CySCs as well as early cyst cells (Fig. 25). Consistent with our

observations in embryonic gonads of *tj*$^{nat/nat}$ flies, the gonadal somatic cells in the testes of these mutants did not express the Tj-TR isoform; thus the tissue specific exclusion of TR in *tj* in the gonads persists until adulthood. Constitutive TR or abolition of TR in *tj* did not affect the development of germline as observed from the distribution and developmental pattern of Vasa-positive germ cells. This further supports our observation that TR in *tj* does not play a role in gonad morphogenesis of adult males.

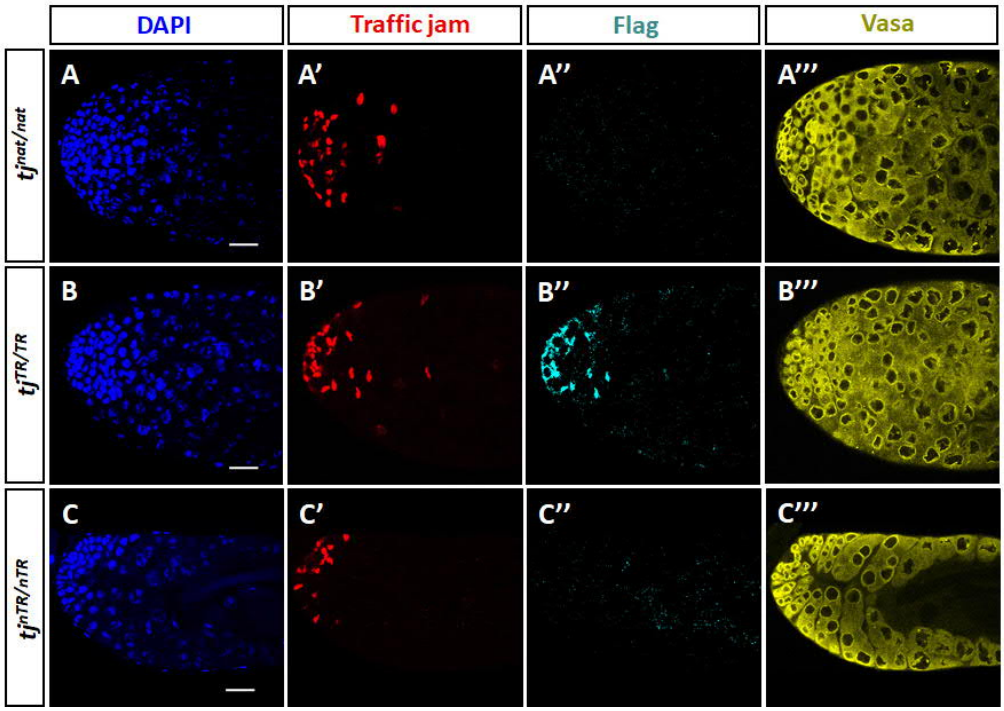


Figure 25. Exclusion of TR in *tj* from somatic cells in adult testes.
DAPI (blue) stains the nuclei, anti-Tj (red) stains the CySCs and early cyst cells, anti-Flag (cyan) stains the Tj-TR isoforms in somatic cells and anti-Vasa (yellow) stains the germline. (A-A') *tj*$^{nat/nat}$ flies express native Tj in the early somatic cells of the testes. (A'') Expression of flag tagged Tj-TR isoform is excluded from somatic cells in testes. (B-B''') *tj*$^{TR/TR}$ testes constitutively express Tj-TR isoform. (C-C'') *tj*$^{nTR/nTR}$ ovaries only express the native Tj (C') and do not express Tj-RT isoform (C''). The Vasa-positive germline cells exhibit wild-type like developmental pattern (A''', B''', C'''). Scale bars represent 25 µm in all panels.

Finally, we analyzed the germaria of *tj*-TR mutant ovaries for possible defects in the organization of the female GSC niche. The female GSC niche architecture in *tj*$^{nat/nat}$ ovaries resembled that of the wild-type (Fig. 26). The cap cells were arranged in a symmetrical streak at the base of the terminal filament. Two somatic cell types of the female GSC niche, the cap cells and escort cells, exhibited expression of Tj. Consistent with the embryonic gonads and adult male testes, *tj*$^{nat/nat}$ ovaries did not express the Tj-TR isoform in the somatic gonadal cells, confirming the observation that TR in *tj* does not occur naturally in the gonadal soma. Strikingly, the GSC niche of *tj*$^{TR/TR}$ ovaries revealed several defects in the architectural properties of the Tj-positive cap cells. Such defects resemble the phenotypic outcomes observed in the ovaries of hypomorphic *tj* (*tj*$^{hypo}$) mutants, where the reduction of Tj activity leads to the development of abnormally long terminal

filaments (Panchal et al., 2017). Some of the Tj-positive cells in the germarium from the *tj*$^{hypo}$ mutants were integrated into the terminal filaments, while the remaining were clustered at the tip of the germarium. We observed similar phenotypes in the *tj*$^{TR/TR}$ ovaries (Fig. 26). While most of the Tj-positive cap cells retained their native cell fate, several of them were clustered at the tip of the germaria and a few attained the spatial organization and morphogenetic behavior of terminal filaments cells. Similar to *tj*$^{hypo}$ mutants, several Tj-positive cells in *tj*$^{TR/TR}$ germaria lost cap cell characteristics, and the displaced cap cells adopted the fate of terminal filaments cells.

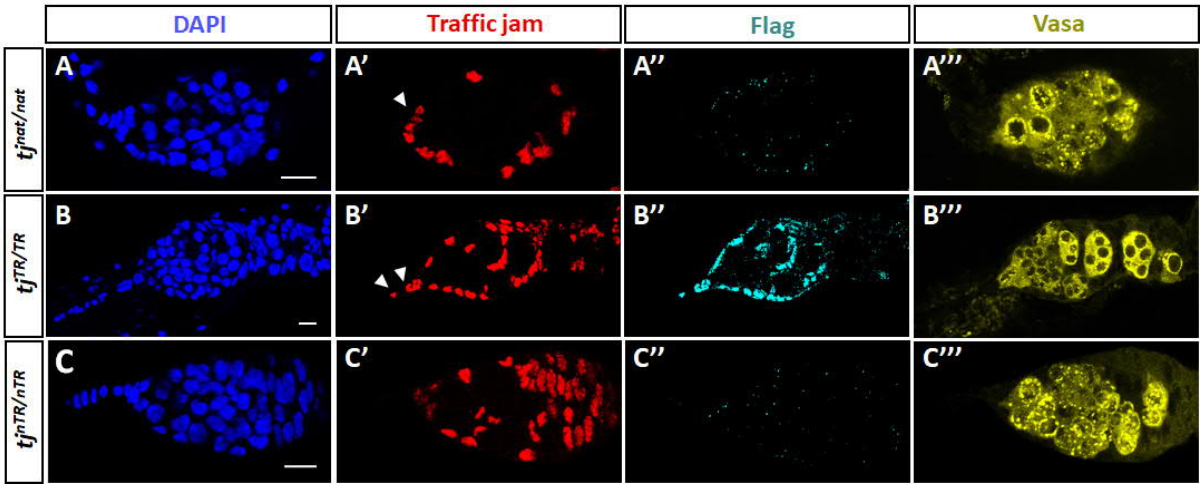


Figure 26. Exclusion of TR in *tj* from somatic cells in the germarium.
DAPI (blue) stains the nuclei, anti-Tj (red) stains the somatic cap cells and escort cells, anti-Flag (cyan) stains the Tj-TR isoform and anti-Vasa (yellow) stains the germline. (A-A') *tj*$^{nat/nat}$ flies express native Tj in the cap cells (arrow) and escort cells in the germarium. (A'') Expression of flag tagged Tj-TR isoform is excluded from the germarium. (B-B''') *tj*$^{TR/TR}$ ovaries constitutively express Tj-TR isoform in the somatic cells within the germarium and additionally in few terminal filament cells (arrow). The cap cells seem to adopt the spatial organization of terminal filament cells. (C-C'') *tj*$^{nTR/nTR}$ ovaries only express native Tj (C') and do not express Tj-RT isoform (C''). Scale bars represent 10 µm in all panels.

Our results show that the constitutive expression of Tj-TR isoform results in phenotypic defects that mimic the hypomorphic allelic mutation of *tj.* Thus, the Tj-TR isoform represents a conditional weaker variant of the native Tj protein. Since constitutive TR did not affect *tj* function during embryogenesis and male gonad development, hypomorphism due to the appendage of TR motif is prominent in the context of specification of cap cell fate in the female GSC; the Tj-TR isoform otherwise retains the functional properties of the native Maf factor. Crossing *tj* mutant females with wild-type males revealed that constitutive TR causes sub-fertility in adult females wherein only few egg follicles develop to maturity (data not shown), a phenotype which is shared by *tj*$^{hypo}$ mutants. *tj*$^{nTR/nTR}$ mutants did not show defects in fertility or in the organization of the female GSC niche, further substantiating the dispensability of TR in *tj* in gonad morphogenesis.

## 3.7 TR in *tj* affects the transcriptome profile in adult brains

Because *tj* undergoes TR specifically in the neural cells of CNS, we wanted to enquire how disruption in TR affects the expression of genes regulated by *tj* in the adult brains. To this aim, we first tested the transcript expression levels of known Tj targets in adult heads. Tj is known to downregulate the expression of several adhesion molecules in ovarian follicle cells (Li et al., 2003). $tj^{-/-}$ follicle cells ectopically overexpress Fas3, DEcad and Nrt; however, whether the regulation of these molecules by Tj is direct and the regulation is sufficient to explain the phenotypes associated with loss of *tj* function is not yet clear. $tj^{-/-}$ follicles do not show changes in the expression levels of another adhesion molecule DN-cadherin (CadN). DEcad is ubiquitously expressed in large number of tissues throughout development. Fas3, CadN and Nrt are particularly enriched in the larval CNS and to some extent in the adult heads (Gelbart and Emmert, 2013). Tj also regulates the expression of the glutamate transporter *VGlut* (Konstantinides et al., 2018) and determines the fate of the photoreceptor subtype 8 by promoting the expression of *melted* (*melt*) and repressing *warts* (*wts*) and *Rhodopsin 6* (*Rh6*) (Jukam et al., 2013).

We performed RT-qPCR analysis with adult head tissues from *tj*-TR mutants to assess the effect of constitutive and abolished TR in *tj* on the transcript levels of the known targets (Fig. 27A). The expression levels of *CadN*, *Nrt*, *VGlut*, *DEcad, Rh6, melt and wts* did not show significant changes in $tj^{TR/TR}$ and $tj^{nTR/nTR}$ mutants, indicating that the loss of TR does not affect Tj function associated with the transcriptional regulation of these genes. Additionally, when constitutively expressed, the Tj-TR isoform can efficiently regulate the expression of these target transcripts at levels comparable to the $tj^{nat/nat}$. Remarkably, *Fas3* was highly downregulated (RQ=0.025) in $tj^{TR/TR}$ flies, compared to the $tj^{nat/nat}$, while *Fas3* expression in $tj^{nTR/nTR}$ remained unaffected indicating that the TR isoform is more efficient than the native Tj in downregulating *Fas3* expression. Our previous experiments suggested that the Tj-TR isoform acts as a hypomorphic variant in specifying cap cell fate in the female GSC niche (Fig. 26). Conversely, in adult heads the Tj-TR isoform seems to mimic a hypermorphic variant that efficiently downregulates the expression of *Fas3*. Our results, thus, hint towards a complex mode of regulation of Tj activity by programmed TR.

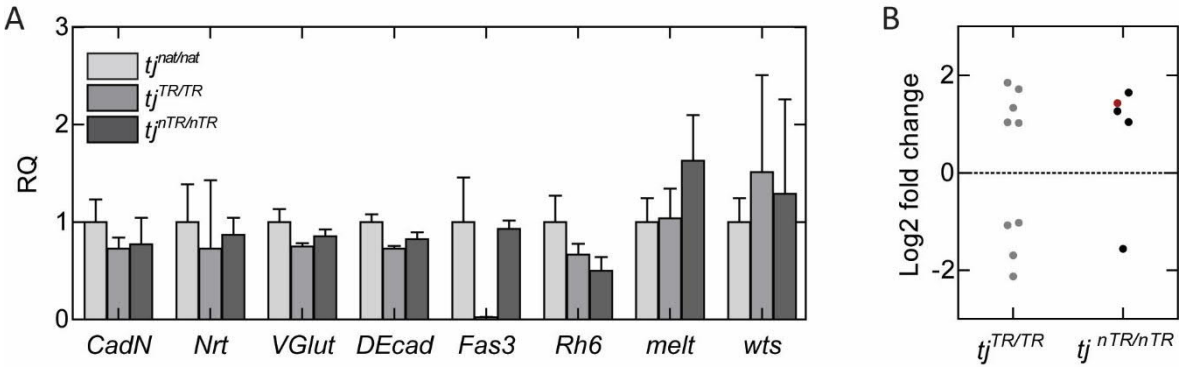


Figure 27. Effect of TR in *tj* on the transcriptome profile in adult CNS.
(A) RT-qPCR analysis of gene expression of targets known to be regulated by Tj in *tj*-TR mutant heads. Error bars represent the range of possible RQ values (RQ Min and Max) defined by the standard error of $\Delta C_T$ from three technical replicates. (B) RNA sequencing studies on adult fly brains identified several genes that are dysregulated in $tj^{TR/TR}$ (light grey dots) and $tj^{nTR/nTR}$ mutants (dark grey dots). Each dot represents a gene that demonstrates a log2 fold change in expression of >1 or <-1, compared to the endogenous control $tj^{nat/nat}$. *tj* transcript is upregulated in the brains of $tj^{nTR/nTR}$ flies (red dot).

Maf transcription factors are involved in the regulation of a large number of genes. In order to expand the set of *tj* targets in our study, we performed high throughput RNA sequencing (RNAseq) on adult brain samples. We could identify several genes that are specifically dysregulated in $tj^{TR/TR}$ or $tj^{nTR/nTR}$ brain samples compared to the reference sample $tj^{nat/nat}$. Several genes were found to be upregulated or downregulated by a log2 fold change of >1 or <-1 (Fig. 27B, Table S2). Among the genes that were upregulated in $tj^{TR/TR}$ brains, *Inwardly rectifying potassium channel 3* (*Irk3*) and *sandman* (*sand*) are particularly interesting candidates because they both encode potassium channel subunits, which participate in critical functions related to electrical signaling among neurons in the brain. *Irk3* is primarily expressed in Malpighian tubules, where it functions in osmoregulation, and is moderately expressed in the adult CNS, where its function is not well understood (Doring et al., 2002; Evans et al., 2005). *sand* is implicated in the circadian rhythm in flies and it acts as a wake-promoting ion channel (Pimentel et al., 2016). *GABA-B-R1,* which belongs to G-protein-coupled receptors family of receptors that binds to inhibitory neurotransmitter GABA (γ-aminobutyric acid), was downregulated in the brains of $tj^{TR/TR}$ mutants. Similarly, several genes were dysregulated in the $tj^{nTR/nTR}$ mutants: *Tektin-C*, *methuselah-like 6* (*mthl6*) and the long non-coding RNA *lncRNA*:*CR45973* were upregulated while *Hexokinase C* (*Hex-C*) was downregulated. Interestingly, *tj* itself was upregulated in the brains of $tj^{nTR/nTR}$ mutants.

In order to reproduce and verify our results, we performed RT-qPCR experiments in which we included few other genes that showed marked differential expression between $tj^{nTR/nTR}$ and $tj^{TR/TR}$ mutants and that carry out important regulatory functions in the cell. We tested the expression of *neither activation nor afterpotential D* (*ninaD*), *Odorant-binding protein 99a* (*Obp99a*), *female-specific independent of transformer* (*fit*), *Drosomycin* (*Drs*), *bubble gum* (*bgm*) and *CG33310*. *ninaD*

codes for a scavenger receptor that is essential for the uptake of carotenoids in adult brains to support the synthesis of 3-hydroxyretinal chromophore which is transported to the retina for rhodopsin biogenesis (Voolstra et al., 2006; Yang and O'Tousa, 2007). *Obp99a* belongs to OBP family of proteins that are components of the olfactory system. They may act as carriers that transport odorants and facilitate their delivery to olfactory receptors on the sensory neurons (Vieira et al., 2007; Wang et al., 2010). *ninaD* and *Obp99a* are thus implicated in sensory perceptions involving phototransduction and olfacto-reception respectively. *fit*, on the other hand is a sexually dimorphic gene that codes for a hormone that controls protein consumption-induced insulin release and protein-specific satiety behavior (Sun et al., 2017). *Drs* constitutes the most expressed antimicrobial peptide in *D. melanogaster* which is regulated by the toll pathway (Lemaitre et al., 1996), as well as constitutively expressed in some parts of fly body during all stages of development (Cohen et al., 2009; Ferrandon et al., 1998).

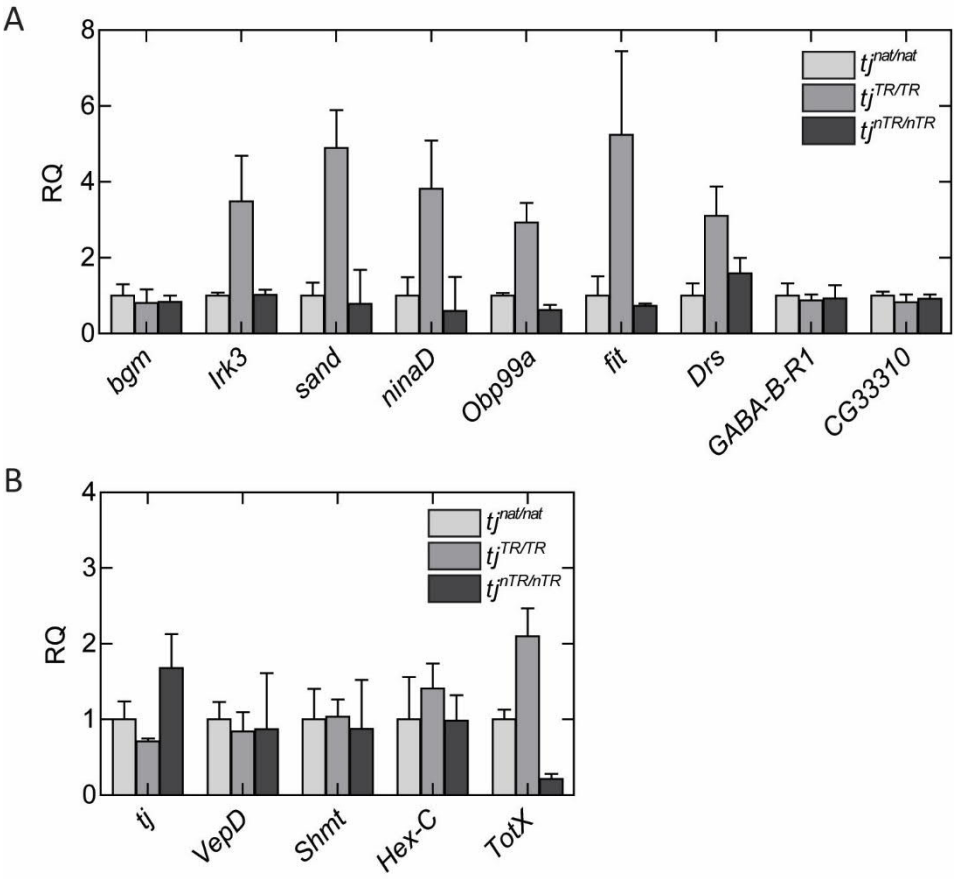


Figure 28. qPCR analysis of target genes identified by RNAseq.
Genes dysregulated in $tj^{TR/TR}$ (A) and $tj^{nTR/nTR}$ (B) mutant heads were analyzed by qPCR using cDNA prepared from head tissues of all three *tj*-TR mutants. Error bars represent the range of possible RQ values (RQ Min and Max) defined by the standard error of $\Delta C_T$ from three technical replicates.

We confirmed upregulation of several genes: *Irk3*, *sand*, *ninaD*, *Obp99a*, *fit* and *Drs* in $tj^{TR/TR}$ heads by 3-5 folds compared to $tj^{nat/nat}$ heads (Fig. 28A). These genes represent potential transcriptional targets that are positively regulated by the Tj-TR isoform. However, the expression

levels of these genes in $tj^{nTR/nTR}$ mutants were comparable to $tj^{nat/nat}$, indicating that under natural TR conditions, the regulation is ensured by the native Tj isoform. Alternatively, the native Tj acts as a repressor of these genes and the TR isoform is less active in their transcriptional repression. The TR extension may allow fine tuning of the activity of the transcription factor by regulating its capacity to enhance or repress the expression of specific targets. Among the genes reported by RNAseq to be dysregulated in $tj^{TR/TR}$ mutants, *bgm*, *GABA-B-R1 and CG33310,* did not show marked differential expression.

Among the candidate genes identified by RNAseq to be dysregulated *in* $tj^{nTR/nTR}$ mutants, we could verify the upregulation of *tj* and the downregulation of *Turandot X* (*TotX*), while we could not reproduce the dysregulation of *Ventrally-expressed protein D* (*VepD*), *Serine hydroxymethyl transferase* (*Shmt*) and *Hex-C*. (Fig. 28B). *TotX* belongs to Turandot family of poorly characterized peptides. It is mainly implicated in cellular response to stress conditions such as heat, bacteria or oxidative stress (Ekengren and Hultmark, 2001). Upregulation of *TotX* in $tj^{TR/TR}$ mutants and downregulation in $tj^{nTR/nTR}$ mutants suggests that it might be one of the specific transcriptional targets of the tj-TR isoform. To eliminate sample heterogeneity, qPCR studies would need to be repeated with cDNA prepared from total RNA of brain lysates. Additional targets can be tested to further expand the list of candidates that are influenced by disruption of TR in *tj*. Overall, our studies identified several genes that were preferentially upregulated upon constitutive induction of readthrough in *tj*. We also identified upregulation of *tj* itself upon abolition of TR, which suggest a mechanism of feedback regulation of *tj*. Based on the transcript expression profile, Tj-TR isoform most likely affects CNS functions related to perception of external stimuli and homeostatic cellular response to such stimuli. Such stimuli may include light, olfactory molecules, stress, dietary components or even pathogenic load. The exact mechanism of regulation of such functions by Tj-TR isoform requires further study.

# 4. DISCUSSION

Translational readthrough is an important recoding mechanism that has been widely described in viruses (Beier et al., 1984; Firth et al., 2011; Hofstetter et al., 1974). TR offers functional advantage by expanding the genetic diversity of the limited viral genomes and by creating protein isoforms that regulate important functions involving viral replication, transmission and infectivity. Breakthrough studies over the last decade have revealed that TR can be highly pervasive across eukaryotic genomes (Dunn et al., 2013; Jungreis et al., 2011; Loughran et al., 2014; Namy et al., 2003) such as yeast, insects and humans. Eukaryotes employ several mechanisms to expand their proteomes such as alternative splicing, alternative polyadenylation and alternative initiation (Kim et al., 2007; Kornblihtt et al., 2013; Tian and Manley, 2017; Touriol et al., 2003). TR is one such mechanism that gives rise to sub-stoichiometric ratios of C-terminally extended protein isoforms with altered physiological functions. Some TR motifs are known to regulate protein function by appending subcellular targeting motifs, reducing ligand binding properties, targeting the protein for proteasomal degradation or even conferring antagonistic properties (Dunn et al., 2013; Eswarappa et al., 2014; Loughran et al., 2018; Stiebler et al., 2014).

Advancements in system biology approaches have led to the identification of several hundred TR candidates in *Drosophila melanogaster*, however, only few of them have been experimentally validated. We have utilized an *in vitro* dual luciferase reporter assay system in S2 cells to study the TR potential of a set of candidate genes in *D. melanogaster*. We selected 11 candidate genes that were predicted by comparative phylogenetic analyses to undergo TR, based on the evolutionarily conserved protein-coding constraints in their putative TR region (Jungreis et al., 2011; Lin et al., 2007). We were able to validate, as well as quantify, the extent of TR in several of these candidates. Consistent with reports in yeast and mammalian systems (Bonetti et al., 1995; Cassan and Rousset, 2001; Cridge et al., 2018; Loughran et al., 2014; McCaughan et al., 1995), we observed high levels of readthrough in genes ending with UGA-C. This includes four genes: *svp*, *aPKC*, *dlg1* and *tj*. The minimal UGA-C sequence was the major determinant that drives TR in two of these candidates (*dlg1* and *tj*). Stop codon contexts favoring leaky termination are highly under-represented in the eukaryotic genome, especially among highly expressed genes (Bonetti et al., 1995; McCaughan et al., 1995). UGA-C represents the least common context in non-TR transcripts while comprising a significant 32.2% of the putative TR candidates in *Drosophila* (Jungreis et al., 2011).

The relative frequencies of stop codons in the putative TR candidates follows the order UGA>UAG>UAA and that of the +4 nucleotide follows the order C>U>G>A (Jungreis et al., 2011). However, biochemical studies are inconsistent in demonstrating the effect of the +4 nucleotide for contexts other than UGA-C (Bonetti et al., 1995; Floquet et al., 2012; Manuvakhova et al., 2000;

McCaughan et al., 1995). Indeed, among the candidates screened, we obtained lowest TR efficiencies for genes *br* and *klu* that possess UAA-G and UAA-C as their stop codon contexts. Termination efficiencies in these genes were minimally affected by the identity of +4 base. *chinmo* and *wit,* which have UAG-G and UAG-C as their stop codon contexts, undergo efficient TR. Mutating the native contexts to UAA-A in these genes lead to modest reduction in TR efficiencies.

The interaction of the translational machinery with different stop codons contexts explains the observed differences in the termination fidelity. The decoding of stop codons during termination has been proposed to occur in two steps, where an initial recognition is followed by the stabilization of the interaction after GTP hydrolysis by eRF3. It has been proposed that the second step occurs at a much slower rate for UGA stop codons (Fan-Minogue et al., 2008) , which allows facultative mode of stop recognition to compete with canonical decoding by eRF1. eRF1 utilizes distinct interactions involving three separate peptide regions to decode different stop codons, which further contributes to the rate and fidelity of canonical decoding. UAA is decoded with the help of the interactions between Glu55 and Tyr125 of YxCxxxF motif in eRF1, whereas recognition of UAG involves the GTS motif. UGA recognition, on the other hand, requires elaborate accommodation of the YxCxxxF motif and involves fewer hydrogen bonds with eRF1 (Brown et al., 2015).

The impact of the +4 nucleotide on termination fidelity might be explained by the interactions it forms with the components of the translational machinery. Stop codon recognition in eukaryotes involves the compaction of mRNA to form a stable U-shaped geometry that pulls the fourth base into the A-site (Brown et al., 2015; Matheisl et al., 2015). Eukaryotic stop signals exploit the +4 base to stabilize the compaction of the mRNA at the decoding center. The +4 nucleotide stacks with G626 of the 18S rRNA; the stacking is more favorable for purines, which corroborates the statistical bias for +4 nucleotide in eukaryotes (Brown et al., 1990b) and ensures termination efficiency. Crosslinking experiments have demonstrated that the +4 nucleotide interacts with eRF1 (Bulygin et al., 2002). This observation has been supported by structural studies that have identified the close proximity of the TASNIKS motif of eRF1 with the +4 nucleotide (Brown et al., 2015; des Georges et al., 2014). Biochemical studies have thus proposed that the Lys63 of TASNIKS motif of eRF1 might form distinct interactions with a +4 C, compared to the other three bases in the same position (Beznoskova et al., 2016). Thus, weak interactions formed by +4 C with G626 of 18S rRNA and Lys63 of eRF1 possibly explains its why the +4 C influences stop codon decoding by eRF1. The inefficiency of +4 C to form stabilizing interactions combined with weaker decoding of UGA by eRF1 might explain why this tetranucleotide signal can serve as an independent trigger to direct readthrough in selected genes in eukaryotes.

Stop codon recognition by eRF1 utilizes a combination of probing the identities of individual bases as well as monitoring the exquisite U-turn like geometry, possibly via an induced-fit mechanism. Recognition of the U-turn geometry allows eRF1 to act as a universal release factor capable of decoding three stop codons that differ in positions +2 and +3. This also rationalizes the divergent evolution of eRF1 compared to bacterial RF1 and RF2 that solely rely on extensive hydrogen bonding for decoding. Any perturbation that hinders the formation of U-turn is likely to impede the decoding capacity of eRF1, thereby leading to TR.

Interestingly, for some of the genes we analyzed, UGA-C alone was not sufficient for triggering TR. Systematic mutational analysis of the TR extension of *aPKC,* revealed that the minimal sequence driving TR comprises a stretch of 12 bp downstream of the stop codon that extends beyond the mRNA channel. Introduction of this short sequence from *aPKC,* downstream of non-readthrough control *αTub84B,* lead to efficient readthrough, indicating the self-sufficiency of this sequence in affecting termination fidelity. The interactions formed between the mRNA beyond the immediate stop codon context and the components of the ribosomal machinery, particularly the rRNA can influence stop codon decoding. In yeast, base pairing between a downstream hexanucleotide stimulatory motif and regions close to h18 and h44 of 18S rRNA have been proposed to modulate termination efficiency (Namy et al., 2001). Structural studies have identified that the +5 base can form stacking interactions with C1695 of 18S rRNA which contributes to the stability of stacking interactions already formed by +4 purines (Shao et al., 2016). The nucleotides in the mRNA channel, specifically at +6 and +8 positions, influence termination fidelity in a complex manner (Cridge et al., 2018). Any dynamic structural changes conferred by downstream or even upstream nucleotides might be might result in subtle distortions in the positioning of RNA bases in the A site that lead to perturbations in the eRF1-nucleotide interactions that interfere with mRNA compaction in the A site. Thus, beyond the stop codon context, the contributions from individual nucleotides in the mRNA can be translated into a more composite effect conferred by a broader stretch of nucleotide sequences. Such effects are sequence-specific and hence it might be difficult to ascertain the precise contribution of nucleotides in a certain position towards modulation of termination efficiency.

While many studies focus on dissecting the molecular contribution of nucleotides that surround the stop codon, there are limited studies that dissect the biological aspect of readthrough and its regulation at the cellular level. For example, several TR candidates were identified in the *Drosophila* genome, but very little is known about the mechanisms by which TR is regulated and the role it plays in the expansion of proteome. Given the diversity in gene ontology functions of the TR candidates identified in *Drosophila*, it would be interesting to explore how TR regulates protein

function at a gene-specific level as well as at a genome-wide level. In order to understand the biological significance of gene-specific TR in *Drosophila*, we chose *traffic jam* as our candidate. *tj* codes for the only large Maf transcription factor in *Drosophila*, whose role in gonad morphogenesis has been widely studied (Li et al., 2003; Panchal et al., 2017; Wingert and DiNardo, 2015). *tj* is also expressed in neuronal tissues at different stages of development but its functions in these tissues remain largely unexplored (Babski et al., 2018; Konstantinides et al., 2018; Li et al., 2003). Our initial dual-luciferase experiment revealed that the minimal UGA-C context plays the determinant role in promoting efficient levels of readthrough in *tj*. In the absence of extended mRNA elements and *trans* factors, leaky termination depends solely on the competition between eRF1 and one or several near-cognate tRNAs in decoding the leaky stop signal.

We employed CRISPR/Cas9-based genome editing in *Drosophila* to create genetic mutants that exhibit constitutive or abolished TR in *tj*. Despite the constitutive expression of Tj in the somatic cells of both male and female gonads, we could not detect the Tj-TR isoform in the embryonic or in the adult gonads, indicating that translation termination occurs with highest fidelity in these tissues. On the other hand, our immunohistochemistry data show that *tj* undergoes readthrough, specifically in the neuronal tissues. The tissue-specific regulation of *tj* was observed starting from embryogenesis and was maintained during adulthood. Tissue-specific regulation of readthrough has been previously reported in *Drosophila* (Dunn et al., 2013; Robinson and Cooley, 1997), however, the mechanism of such a regulatory control and its physiological significance remains unclear.

To prevent physiological aberrations that might affect the germline and the developing gametes, the cells that comprise the stem cell niche in gonads (e.g CySCs) might employ rigid fidelity control during gene expression that would limit mutations in the genome, transcriptome or proteome. It is possible that in specific neuronal cell types, translation termination is globally less strictly regulated, and therefore, TR genes with leaky contexts such as UGA-C are more prone to TR. This hypothesis is supported by the observation that neuron-specific genes are highly enriched in the phylogenetically predicted list of TR candidates in *Drosophila* (Jungreis et al., 2011). Because UAA is the preferred stop codon in genes with housekeeping functions and with high breadth and level of expression, such leaky termination profile is unlikely to pose drastic effects on basic homeostatic functions in a cell (Trotta, 2016). Additionally, the over-abundance of in-frame stop signals downstream of the primary stop signal (up to 3 codons downstream) in non-TR genes decrease the occurrence of TR (Williams et al., 2004). Understanding the mechanistic aspects of TR regulation in *tj* would require identification of the amino acid that is incorporated at the primary

stop codon and profiling the abundance of specific tRNA isoforms that might code for the identified amino acid via near-cognate pairing with the UGA stop codon.

Given that the Tj-TR isoform is not expressed in the gonadal somatic cells and that disruption of TR did not affect gonad specific functions of *tj*, we can conclude that TR does not play a role in gonad morphogenesis and gametogenesis. Additionally, the Tj-TR isoform, when expressed constitutively, can perform the gonadal soma-specific functions of native Tj during embryonic gonad development as well as male gametogenesis, indicating that the TR extension does not affect most of the regulatory functions of Tj. However, constitutive expression of the Tj-TR isoform lead to defects in the specification of cap cells in the ovaries, such that they assume the morphology and spatial arrangement of terminal filament cells. A similar phenotype has been previously reported in hypomorphic *tj* mutants that exhibit reduced Tj expression (Panchal et al., 2017). It has been proposed that the gene(s) involved in controlling the morphogenetic behavior and spatial orientation of cap cells respond to high concentrations of Tj in the cap cells. Because the Tj-TR isoform selectively affects the morphology and spatial distribution of cap cells, it mimics a hypomorphic or weaker variant of the native protein. The TR extension might thus affect the DNA binding properties of native Tj, such that it can no longer regulate the expression of genes that control the morphogenetic behavior and arrangement of cap cells in the germarium.

Because TR in *tj* is specific to neuronal tissues, we tested the effect of TR on the expression levels of known putative *tj* targets in adult heads. We did not observe differences in the transcript expression levels of *CadN*, *DEcad*, *nrt*, *VGlut*, *melt*, *wts* and *Rh6* upon constitutive induction or abolition of TR. However, we observed significant downregulation of *Fas3* in $tj^{TR/TR}$ mutant adult heads. *Fas3* is negatively regulated by Tj, as selective depletion of *tj* function in ovarian follicle cells leads to ectopic overexpression of Fas3 (Li et al., 2003). It is likely that the Tj-TR isoform resembles a condition-specific hypermorphic variant of native Tj that can actively repress the expression of *Fas3* by direct transcriptional regulation. Because we do not see differences in the levels of *Fas3* in $tj^{nat/nat}$ and $tj^{nTR/nTR}$ mutants, transcriptional repression of *Fas3* by Tj-TR might be dose dependent.

To expand our knowledge on the effect of TR in *tj* on the transcriptome profile in adult CNS, we performed high throughput RNAseq on adult brains. We were able to identify and confirm dysregulation of several neuronal genes upon constitutive induction or abolition of TR in *tj*. Constitutive expression of Tj-TR lead to upregulation of genes such as *Irk3*, *sand*, *ninaD*, *Obp99a*, *fit* and *Drs*. The Tj-TR isoform can thus act as efficient activator of these targets. Furthermore, abolition of TR in *tj* lead to severe downregulation of the *TotX* transcript expression, indicating that this gene might be a dedicated target for the Tj-TR isoform. Interestingly, the *tj* transcript itself was upregulated in the brains of $tj^{nTR/nTR}$ mutants, which hints towards a negative feedback regulation

of *tj* expression by the Tj-TR isoform. The genes that are dysregulated upon constitutive induction or abolition of TR in *tj* seem to be mostly involved in CNS functions that are involved in perception of external stimuli such as photo stimuli, odorants, dietary cues, stress and pathogens and regulation of bodily response towards such stimuli. It is not known whether the TR in *tj* is further regulated in response to external cues, such that the expression of the target genes depend on the effective concentration of Tj-TR isoform in the cell. It would be interesting to quantify the expression levels of the Tj-TR isoform upon stress induction to correlate it with changes in the expression profile of target genes. For example, *sand* is a wake-promoting ion channel that is upregulated in $tj^{TR/TR}$ mutants (Pimentel et al., 2016). Analysing the sleep-wake pattern of the *tj*-TR mutants under normal light-dark cycles as well as dark-dark cycles and correlating it with the expression pattern of Tj-TR induced overexpression of *sand* would provide valuable phenotypic information on the physiological function of TR in *tj*.

Combining our gene expression studies with immunohistochemistry in ovarian cap cells, we conclude that TR in Tj confers diverse functional properties to the native transcription factor. The constant nuclear expression of Flag-tagged Tj-TR isoform confirms that TR does not affect protein stability and nuclear localization properties of Tj. The TR extension might have an overall negative effect in cap cell specific functions of Tj, while simultaneously amplifying *Fas3* repressor function of Tj in adult heads. The repression of *Fas3* by the Tj-TR isoform in the gonads needs to be verified via RT-qPCR and immunohistochemistry analysis. Nevertheless, the Tj-TR isoform can function like the native protein without disrupting many of its intrinsic transcriptional functions.

The exact mechanism of regulation of Tj function by TR is not clear yet. Given the proximity of the TR extension to the DNA binding motifs (leucine zipper motif) in the C-terminal region of Tj (Fig. 11), it is possible that the juxtaposition of TR motif impedes or confers novel DNA binding properties, or that the TR motif alters the protein-binding properties of Tj. Novel protein binding properties might increase its affinity towards transcription activators leading to increased activity in gene regulation (e.g. *Fas3* downregulation) or sequester the protein, thereby depleting its cellular pool. It would be thus be interesting to analyze the impact of TR motif on the interactome profile of Tj. Furthermore, bZip transcription factors are known to form homo/heterodimers using their dimerization domains (Kataoka et al., 1994; Kurokawa et al., 2009); therefore, the TR extension might affect the dimerization affinity of Tj, which would lead to attenuation of gene regulatory functions of the native Tj. TR mostly affects proteins with structurally disordered C-termini (Kleppe and Bornberg-Bauer, 2018; Pancsa et al., 2016). This would allow increased flexibility and accessibility to TR extensions to engage in independent interactions without affecting the native protein function. Given the high number of TR genes involved in regulatory functions in *Drosophila*,

an interaction-prone C-terminal segment would allow fine-tuning of the protein function in a regulated manner. Such regulation can be temporal or spatial in effect. TR in *tj* serves as a spatially regulated recoding mechanism that widens the gene regulatory functions of the native Maf factor. While the TR extension does not result in detrimental effects affecting protein stability and localization, it primarily helps to fine tune the transcription factor/enhancer activity of Tj in a tissue specific manner.

Thus, the study of regulation of TR and its biological consequences provides new insights into how organisms alter gene expression to yield relatively low yet constant amounts of gene products with novel protein functions at low evolutionary cost. TR provides means to adapt to stress conditions as well as enhance the overall fitness of the organism. Given that most of the putative TR candidates in *D. melanogaster* are not conserved with their yeast orthologues TR events might have appeared independently during the course of evolution, allowing proteins to sample selective fitness by evolving at the C-terminus.

# 5. REFERENCES


Adio, S., Senyushkina, T., Peske, F., Fischer, N., Wintermeyer, W., and Rodnina, M.V. (2015). Fluctuations between multiple EF-G-induced chimeric tRNA states during translocation on the ribosome. Nat Commun *6*, 7442.

Adio, S., Sharma, H., Senyushkina, T., Karki, P., Maracci, C., Wohlgemuth, I., Holtkamp, W., Peske, F., and Rodnina, M.V. (2018). Dynamics of ribosomes and release factors during translation termination in E. coli. Elife *7*.

Agirrezabala, X., Samatova, E., Klimova, M., Zamora, M., Gil-Carton, D., Rodnina, M.V., and Valle, M. (2017). Ribosome rearrangements at the onset of translational bypassing. Sci Adv *3*, e1700147.

Aitken, C.E., and Lorsch, J.R. (2012). A mechanistic overview of translation initiation in eukaryotes. Nat Struct Mol Biol *19*, 568-576.

Alkalaeva, E.Z., Pisarev, A.V., Frolova, L.Y., Kisselev, L.L., and Pestova, T.V. (2006). In vitro reconstitution of eukaryotic translation reveals cooperativity between release factors eRF1 and eRF3. Cell *125*, 1125-1136.

Andersen, G.R., Pedersen, L., Valente, L., Chatterjee, I., Kinzy, T.G., Kjeldgaard, M., and Nyborg, J. (2000). Structural basis for nucleotide exchange and competition with tRNA in the yeast elongation factor complex eEF1A:eEF1Balpha. Mol Cell *6*, 1261-1266.

Arkov, A.L., Korolev, S.V., and Kisselev, L.L. (1995). 5' contexts of Escherichia coli and human termination codons are similar. Nucleic Acids Res *23*, 4712-4716.

Aylett, C.H., Boehringer, D., Erzberger, J.P., Schaefer, T., and Ban, N. (2015). Structure of a yeast 40S-eIF1-eIF1A-eIF3-eIF3j initiation complex. Nat Struct Mol Biol *22*, 269-271.

Babski, H., Surel, C., Shingo, Y., Valmier, J., Thomas, J.B., Carroll, P., and Garces, A. (2018). A GABAergic Maf-expressing interneuron subset regulates the speed of locomotion in Drosophila. bioRxiv, 421057.

Battle, D.J., and Doudna, J.A. (2002). Specificity of RNA-RNA helix recognition. Proc Natl Acad Sci U S A *99*, 11676-11681.

Behrmann, E., Loerke, J., Budkevich, T.V., Yamamoto, K., Schmidt, A., Penczek, P.A., Vos, M.R., Burger, J., Mielke, T., Scheerer, P., *et al.* (2015). Structural snapshots of actively translating human ribosomes. Cell *161*, 845-857.

Beier, H., Barciszewska, M., Krupp, G., Mitnacht, R., and Gross, H.J. (1984). UAG readthrough during TMV RNA translation: isolation and sequence of two tRNAs with suppressor activity from tobacco plants. EMBO J *3*, 351-356.

Beier, H., and Grimm, M. (2001). Misreading of termination codons in eukaryotes by natural nonsense suppressor tRNAs. Nucleic Acids Res *29*, 4767-4782.

Beissel, C., Neumann, B., Uhse, S., Hampe, I., Karki, P., and Krebber, H. (2019). Translation termination depends on the sequential ribosomal entry of eRF1 and eRF3. Nucleic Acids Res.

Belardinelli, R., Sharma, H., Peske, F., Wintermeyer, W., and Rodnina, M.V. (2016). Translocation as continuous movement through the ribosome. RNA Biol *13*, 1197-1203.

Beringer, M., and Rodnina, M.V. (2007). The ribosomal peptidyl transferase. Mol Cell *26*, 311-321.

Bertram, G., Bell, H.A., Ritchie, D.W., Fullerton, G., and Stansfield, I. (2000). Terminating eukaryote translation: domain 1 of release factor eRF1 functions in stop codon recognition. RNA *6*, 1236-1247.

Beznoskova, P., Gunisova, S., and Valasek, L.S. (2016). Rules of UGA-N decoding by near-cognate tRNAs and analysis of readthrough on short uORFs in yeast. RNA *22*, 456-466.

Beznoskova, P., Wagner, S., Jansen, M.E., von der Haar, T., and Valasek, L.S. (2015). Translation initiation factor eIF3 promotes programmed stop codon readthrough. Nucleic Acids Res *43*, 5099-5111.

Blanchet, S., Cornu, D., Argentini, M., and Namy, O. (2014). New insights into the incorporation of natural suppressor tRNAs at stop codons in Saccharomyces cerevisiae. Nucleic Acids Res *42*, 10061-10072.

Blanchet, S., Rowe, M., Von der Haar, T., Fabret, C., Demais, S., Howard, M.J., and Namy, O. (2015). New insights into stop codon recognition by eRF1. Nucleic Acids Res *43*, 3298-3308.

Blank, V., and Andrews, N.C. (1997). The Maf transcription factors: regulators of differentiation. Trends Biochem Sci *22*, 437-441.

Blight, S.K., Larue, R.C., Mahapatra, A., Longstaff, D.G., Chang, E., Zhao, G., Kang, P.T., Green-Church, K.B., Chan, M.K., and Krzycki, J.A. (2004). Direct charging of tRNA(CUA) with pyrrolysine in vitro and in vivo. Nature *431*, 333-335.

Bock, A., Forchhammer, K., Heider, J., Leinfelder, W., Sawers, G., Veprek, B., and Zinoni, F. (1991). Selenocysteine: the 21st amino acid. Mol Microbiol *5*, 515-520.

Bonetti, B., Fu, L., Moon, J., and Bedwell, D.M. (1995). The efficiency of translation termination is determined by a synergistic interplay between upstream and downstream sequences in Saccharomyces cerevisiae. J Mol Biol *251*, 334-345.

Brault, V., van den Heuvel, J.F., Verbeek, M., Ziegler-Graff, V., Reutenauer, A., Herrbach, E., Garaud, J.C., Guilley, H., Richards, K., and Jonard, G. (1995). Aphid transmission of beet western yellows luteovirus requires the minor capsid read-through protein P74. EMBO J *14*, 650-659.

Brown, A., Shao, S., Murray, J., Hegde, R.S., and Ramakrishnan, V. (2015). Structural basis for stop codon recognition in eukaryotes. Nature *524*, 493-496.

Brown, C.M., Dinesh-Kumar, S.P., and Miller, W.A. (1996). Local and distant sequences are required for efficient readthrough of the barley yellow dwarf virus PAV coat protein gene stop codon. J Virol *70*, 5884-5892.

Brown, C.M., Stockwell, P.A., Trotman, C.N., and Tate, W.P. (1990a). Sequence analysis suggests that tetra-nucleotides signal the termination of protein synthesis in eukaryotes. Nucleic Acids Res *18*, 6339-6345.

Brown, C.M., Stockwell, P.A., Trotman, C.N., and Tate, W.P. (1990b). The signal for the termination of protein synthesis in procaryotes. Nucleic Acids Res *18*, 2079-2086.

Budkevich, T., Giesebrecht, J., Altman, R.B., Munro, J.B., Mielke, T., Nierhaus, K.H., Blanchard, S.C., and Spahn, C.M. (2011). Structure and dynamics of the mammalian ribosomal pretranslocation complex. Mol Cell *44*, 214-224.

Bulygin, K.N., Khairulina, Y.S., Kolosov, P.M., Ven'yaminova, A.G., Graifer, D.M., Vorobjev, Y.N., Frolova, L.Y., Kisselev, L.L., and Karpova, G.G. (2010). Three distinct peptides from the N domain of translation termination factor eRF1 surround stop codon in the ribosome. RNA *16*, 1902-1914.

Bulygin, K.N., Repkova, M.N., Ven'yaminova, A.G., Graifer, D.M., Karpova, G.G., Frolova, L.Y., and Kisselev, L.L. (2002). Positioning of the mRNA stop signal with respect to polypeptide chain release factors and ribosomal proteins in 80S ribosomes. Febs Lett *514*, 96-101.

Caliskan, N., Peske, F., and Rodnina, M.V. (2015). Changed in translation: mRNA recoding by -1 programmed ribosomal frameshifting. Trends Biochem Sci *40*, 265-274.

Carnes, J., Jacobson, M., Leinwand, L., and Yarus, M. (2003). Stop codon suppression via inhibition of eRF1 expression. RNA *9*, 648-653.

Cassan, M., and Rousset, J.P. (2001). UAG readthrough in mammalian cells: effect of upstream and downstream stop codon contexts reveal different signals. BMC Mol Biol *2*, 3.

Chambert, R., Rain-Guion, M.C., and Petit-Glatron, M.F. (1992). Readthrough of the Bacillus subtilis stop codon produces an extended enzyme displaying a higher polymerase activity. Biochim Biophys Acta *1132*, 145-153.

Chao, A.T., Dierick, H.A., Addy, T.M., and Bejsovec, A. (2003). Mutations in eukaryotic release factors 1 and 3 act as general nonsense suppressors in Drosophila. Genetics *165*, 601-612.

Chauvin, C., Salhi, S., Le Goff, C., Viranaicken, W., Diop, D., and Jean-Jean, O. (2005). Involvement of human release factors eRF3a and eRF3b in translation termination and regulation of the termination complex formation. Mol Cell Biol *25*, 5801-5811.

Chavatte, L., Seit-Nebi, A., Dubovaya, V., and Favre, A. (2002). The invariant uridine of stop codons contacts the conserved NIKSR loop of human eRF1 in the ribosome. EMBO J *21*, 5302-5311.

Cheng, Z., Saito, K., Pisarev, A.V., Wada, M., Pisareva, V.P., Pestova, T.V., Gajda, M., Round, A., Kong, C., Lim, M., *et al.* (2009). Structural insights into eRF3 and stop codon recognition by eRF1. Genes Dev *23*, 1106-1118.

Chiba, S., Lamsa, A., and Pogliano, K. (2009). A ribosome-nascent chain sensor of membrane protein biogenesis in Bacillus subtilis. EMBO J *28*, 3461-3475.

Chittum, H.S., Lane, W.S., Carlson, B.A., Roller, P.P., Lung, F.D., Lee, B.J., and Hatfield, D.L. (1998). Rabbit beta-globin is extended beyond its UGA stop codon by multiple suppressions and translational reading gaps. Biochemistry *37*, 10866-10870.

Clark, A.G., Eisen, M.B., Smith, D.R., Bergman, C.M., Oliver, B., Markow, T.A., Kaufman, T.C., Kellis, M., Gelbart, W., Iyer, V.N., *et al.* (2007). Evolution of genes and genomes on the Drosophila phylogeny. Nature *450*, 203-218.

Cohen, L., Moran, Y., Sharon, A., Segal, D., Gordon, D., and Gurevitz, M. (2009). Drosomycin, an innate immunity peptide of Drosophila melanogaster, interacts with the fly voltage-gated sodium channel. J Biol Chem *284*, 23558-23563.

Cridge, A.G., Crowe-McAuliffe, C., Mathew, S.F., and Tate, W.P. (2018). Eukaryotic translational termination efficiency is influenced by the 3' nucleotides within the ribosomal mRNA channel. Nucleic Acids Res *46*, 1927-1944.

Csibra, E., Brierley, I., and Irigoyen, N. (2014). Modulation of stop codon read-through efficiency and its effect on the replication of murine leukemia virus. J Virol *88*, 10364-10376.

Cymer, F., Hedman, R., Ismail, N., and von Heijne, G. (2015). Exploration of the arrest peptide sequence space reveals arrest-enhanced variants. J Biol Chem *290*, 10208-10215.

Dabrowski, M., Bukowy-Bieryllo, Z., and Zietkiewicz, E. (2015). Translational readthrough potential of natural termination codons in eucaryotes--The impact of RNA sequence. RNA Biol *12*, 950-958.

de Cuevas, M., and Matunis, E.L. (2011). The stem cell niche: lessons from the Drosophila testis. Development *138*, 2861-2869.

des Georges, A., Hashem, Y., Unbehaun, A., Grassucci, R.A., Taylor, D., Hellen, C.U., Pestova, T.V., and Frank, J. (2014). Structure of the mammalian ribosomal pre-termination complex associated with eRF1.eRF3.GDPNP. Nucleic Acids Res *42*, 3409-3418.

Dever, T.E., and Green, R. (2012). The elongation, termination, and recycling phases of translation in eukaryotes. Cold Spring Harb Perspect Biol *4*, a013706.

Dever, T.E., Kinzy, T.G., and Pavitt, G.D. (2016). Mechanism and Regulation of Protein Synthesis in Saccharomyces cerevisiae. Genetics *203*, 65-107.

Dinman, J.D. (2012). Control of gene expression by translational recoding. Adv Protein Chem Struct Biol *86*, 129-149.

Donnelly, M.L., Luke, G., Mehrotra, A., Li, X., Hughes, L.E., Gani, D., and Ryan, M.D. (2001). Analysis of the aphthovirus 2A/2B polyprotein 'cleavage' mechanism indicates not a proteolytic reaction, but a novel translational effect: a putative ribosomal 'skip'. J Gen Virol *82*, 1013-1025.

Doring, F., Wischmeyer, E., Kuhnlein, R.P., Jackle, H., and Karschin, A. (2002). Inwardly rectifying K+ (Kir) channels in Drosophila. A crucial role of cellular milieu factors Kir channel function. J Biol Chem *277*, 25554-25561.

Doronina, V.A., Wu, C., de Felipe, P., Sachs, M.S., Ryan, M.D., and Brown, J.D. (2008). Site-specific release of nascent chains from ribosomes at a sense codon. Mol Cell Biol *28*, 4227-4239.

Dunn, J.G., Foo, C.K., Belletier, N.G., Gavis, E.R., and Weissman, J.S. (2013). Ribosome profiling reveals pervasive and regulated stop codon readthrough in Drosophila melanogaster. Elife *2*, e01179.

Ekengren, S., and Hultmark, D. (2001). A family of Turandot-related genes in the humoral stress response of Drosophila. Biochem Biophys Res Commun *284*, 998-1003.

Eswarappa, S.M., Potdar, A.A., Koch, W.J., Fan, Y., Vasu, K., Lindner, D., Willard, B., Graham, L.M., DiCorleto, P.E., and Fox, P.L. (2014). Programmed translational readthrough generates antiangiogenic VEGF-Ax. Cell *157*, 1605-1618.

Evans, J.M., Allan, A.K., Davies, S.A., and Dow, J.A. (2005). Sulphonylurea sensitivity and enriched expression implicate inward rectifier K+ channels in Drosophila melanogaster renal function. J Exp Biol *208*, 3771-3783.

Eyler, D.E., Wehner, K.A., and Green, R. (2013). Eukaryotic release factor 3 is required for multiple turnovers of peptide release catalysis by eukaryotic release factor 1. J Biol Chem *288*, 29530-29538.

Fan-Minogue, H., Du, M., Pisarev, A.V., Kallmeyer, A.K., Salas-Marco, J., Keeling, K.M., Thompson, S.R., Pestova, T.V., and Bedwell, D.M. (2008). Distinct eRF3 requirements suggest alternate eRF1 conformations mediate peptide release during eukaryotic translation termination. Mol Cell *30*, 599-609.

Fearon, K., McClendon, V., Bonetti, B., and Bedwell, D.M. (1994). Premature translation termination mutations are efficiently suppressed in a highly conserved region of yeast Ste6p, a member of the ATP-binding cassette (ABC) transporter family. J Biol Chem *269*, 17802-17808.

Felsenstein, K.M., and Goff, S.P. (1988). Expression of the gag-pol fusion protein of Moloney murine leukemia virus without gag protein does not induce virion formation or proteolytic processing. J Virol *62*, 2179-2182.

Feng, Y.X., Yuan, H., Rein, A., and Levin, J.G. (1992). Bipartite signal for read-through suppression in murine leukemia virus mRNA: an eight-nucleotide purine-rich sequence immediately downstream of the gag termination codon followed by an RNA pseudoknot. J Virol *66*, 5127-5132.

Ferguson, A., Wang, L., Altman, R.B., Terry, D.S., Juette, M.F., Burnett, B.J., Alejo, J.L., Dass, R.A., Parks, M.M., Vincent, C.T., *et al.* (2015). Functional Dynamics within the Human Ribosome Regulate the Rate of Active Protein Synthesis. Mol Cell *60*, 475-486.

Ferrandon, D., Jung, A.C., Criqui, M., Lemaitre, B., Uttenweiler-Joseph, S., Michaut, L., Reichhart, J., and Hoffmann, J.A. (1998). A drosomycin-GFP reporter transgene reveals a local immune response in Drosophila that is not dependent on the Toll pathway. EMBO J *17*, 1217-1227.

Filichkin, S.A., Lister, R.M., McGrath, P.F., and Young, M.J. (1994). In vivo expression and mutational analysis of the barley yellow dwarf virus readthrough gene. Virology *205*, 290-299.

Firth, A.E., Wills, N.M., Gesteland, R.F., and Atkins, J.F. (2011). Stimulation of stop codon readthrough: frequent presence of an extended 3' RNA structural element. Nucleic Acids Res *39*, 6679-6691.

Floquet, C., Hatin, I., Rousset, J.P., and Bidou, L. (2012). Statistical analysis of readthrough levels for nonsense mutations in mammalian cells reveals a major determinant of response to gentamicin. PLoS Genet *8*, e1002608.

Florin, T., Maracci, C., Graf, M., Karki, P., Klepacki, D., Berninghausen, O., Beckmann, R., Vazquez-Laslop, N., Wilson, D.N., Rodnina, M.V., *et al.* (2017). An antimicrobial peptide that inhibits translation by trapping release factors on the ribosome. Nat Struct Mol Biol *24*, 752-757.

Forchhammer, K., Boesmiller, K., and Bock, A. (1991). The function of selenocysteine synthase and SELB in the synthesis and incorporation of selenocysteine. Biochimie *73*, 1481-1486.

Franckenberg, S., Becker, T., and Beckmann, R. (2012). Structural view on recycling of archaeal and eukaryotic ribosomes after canonical termination and ribosome rescue. Curr Opin Struct Biol *22*, 786-796.

Frank, J., and Agrawal, R.K. (2000). A ratchet-like inter-subunit reorganization of the ribosome during translocation. Nature *406*, 318-322.

Frank, J., and Gonzalez, R.L., Jr. (2010). Structure and dynamics of a processive Brownian motor: the translating ribosome. Annu Rev Biochem *79*, 381-412.

Franza, B.R., Jr., Rauscher, F.J., 3rd, Josephs, S.F., and Curran, T. (1988). The Fos complex and Fos-related antigens recognize sequence elements that contain AP-1 binding sites. Science *239*, 1150-1153.

Fraser, C.S., Berry, K.E., Hershey, J.W., and Doudna, J.A. (2007). eIF3j is located in the decoding center of the human 40S ribosomal subunit. Mol Cell *26*, 811-819.

Freistroffer, D.V., Kwiatkowski, M., Buckingham, R.H., and Ehrenberg, M. (2000). The accuracy of codon recognition by polypeptide release factors. Proc Natl Acad Sci U S A *97*, 2046-2051.

Freitag, J., Ast, J., and Bolker, M. (2012). Cryptic peroxisomal targeting via alternative splicing and stop codon read-through in fungi. Nature *485*, 522-525.

Frolova, L., Le Goff, X., Zhouravleva, G., Davydova, E., Philippe, M., and Kisselev, L. (1996). Eukaryotic polypeptide chain release factor eRF3 is an eRF1- and ribosome-dependent guanosine triphosphatase. RNA *2*, 334-341.

Frolova, L.Y., Tsivkovskii, R.Y., Sivolobova, G.F., Oparina, N.Y., Serpinsky, O.I., Blinov, V.M., Tatkov, S.I., and Kisselev, L.L. (1999). Mutations in the highly conserved GGQ motif of class 1 polypeptide release factors abolish ability of human eRF1 to trigger peptidyl-tRNA hydrolysis. RNA *5*, 1014-1020.

Gagnon, M.G., Roy, R.N., Lomakin, I.B., Florin, T., Mankin, A.S., and Steitz, T.A. (2016). Structures of proline-rich peptides bound to the ribosome reveal a common mechanism of protein synthesis inhibition. Nucleic Acids Res *44*, 2439-2450.

Gao, H., Zhou, Z., Rawat, U., Huang, C., Bouakaz, L., Wang, C., Cheng, Z., Liu, Y., Zavialov, A., Gursky, R.*, et al.* (2007). RF3 induces ribosomal conformational changes responsible for dissociation of class I release factors. Cell *129*, 929-941.

Gao, N., Zavialov, A.V., Li, W., Sengupta, J., Valle, M., Gursky, R.P., Ehrenberg, M., and Frank, J. (2005). Mechanism for the disassembly of the posttermination complex inferred from cryo-EM studies. Mol Cell *18*, 663-674.

Gelbart, W.M., and Emmert, D.B. (2013). FlyBase High Throuput Expression Pattern Data. FlyBase analysis *(flybase.org/reports/FBrf0221009.html)*.

Geller, A.I., and Rich, A. (1980). A UGA termination suppression tRNATrp active in rabbit reticulocytes. Nature *283*, 41-46.

Gibson, D.G., Young, L., Chuang, R.Y., Venter, J.C., Hutchison, C.A., 3rd, and Smith, H.O. (2009). Enzymatic assembly of DNA molecules up to several hundred kilobases. Nat Methods *6*, 343-345.

Grentzmann, G., Ingram, J.A., Kelly, P.J., Gesteland, R.F., and Atkins, J.F. (1998). A dual-luciferase reporter system for studying recoding signals. RNA *4*, 479-486.

Gross, T., Siepmann, A., Sturm, D., Windgassen, M., Scarcelli, J.J., Seedorf, M., Cole, C.N., and Krebber, H. (2007). The DEAD-box RNA helicase Dbp5 functions in translation termination. Science *315*, 646-649.

Gumbart, J., Schreiner, E., Wilson, D.N., Beckmann, R., and Schulten, K. (2012). Mechanisms of SecM-mediated stalling in the ribosome. Biophys J *103*, 331-341.

Gunawan, F., Arandjelovic, M., and Godt, D. (2013). The Maf factor Traffic jam both enables and inhibits collective cell migration in Drosophila oogenesis. Development *140*, 2808-2817.

Guo, M., and Schimmel, P. (2012). Structural analyses clarify the complex control of mistranslation by tRNA synthetases. Curr Opin Struct Biol *22*, 119-126.

Hardy, R.W., Tokuyasu, K.T., Lindsley, D.L., and Garavito, M. (1979). The germinal proliferation center in the testis of Drosophila melanogaster. J Ultrastruct Res *69*, 180-190.

Harger, J.W., and Dinman, J.D. (2003). An in vivo dual-luciferase assay system for studying translational recoding in the yeast Saccharomyces cerevisiae. RNA *9*, 1019-1024.

Harrell, L., Melcher, U., and Atkins, J.F. (2002). Predominance of six different hexanucleotide recoding signals 3' of read-through stop codons. Nucleic Acids Res *30*, 2011-2017.

Hashem, Y., des Georges, A., Dhote, V., Langlois, R., Liao, H.Y., Grassucci, R.A., Hellen, C.U., Pestova, T.V., and Frank, J. (2013). Structure of the mammalian ribosomal 43S preinitiation complex bound to the scanning factor DHX29. Cell *153*, 1108-1119.

Hatfield, D., Thorgeirsson, S.S., Copeland, T.D., Oroszlan, S., and Bustin, M. (1988). Immunopurification of the suppressor tRNA dependent rabbit beta-globin readthrough protein. Biochemistry *27*, 1179-1183.

Hati, S., Ziervogel, B., Sternjohn, J., Wong, F.C., Nagan, M.C., Rosen, A.E., Siliciano, P.G., Chihade, J.W., and Musier-Forsyth, K. (2006). Pre-transfer editing by class II prolyl-tRNA synthetase: role of aminoacylation active site in "selective release" of noncognate amino acids. J Biol Chem *281*, 27862-27872.

Hauryliuk, V., Zavialov, A., Kisselev, L., and Ehrenberg, M. (2006). Class-1 release factor eRF1 promotes GTP binding by class-2 release factor eRF3. Biochimie *88*, 747-757.

Hellen, C.U.T. (2018). Translation Termination and Ribosome Recycling in Eukaryotes. Cold Spring Harb Perspect Biol *10*.

Herr, A.J., Gesteland, R.F., and Atkins, J.F. (2000). One protein from two open reading frames: mechanism of a 50 nt translational bypass. EMBO J *19*, 2671-2680.

Hinnebusch, A.G., and Lorsch, J.R. (2012). The mechanism of eukaryotic translation initiation: new insights and challenges. Cold Spring Harb Perspect Biol *4*.

Hofstetter, H., Monstein, H.J., and Weissmann, C. (1974). The readthrough protein A1 is essential for the formation of viable Q beta particles. Biochim Biophys Acta *374*, 238-251.

Howard, M.T., Shirts, B.H., Petros, L.M., Flanigan, K.M., Gesteland, R.F., and Atkins, J.F. (2000). Sequence specificity of aminoglycoside-induced stop condon readthrough: potential implications for treatment of Duchenne muscular dystrophy. Ann Neurol *48*, 164-169.

Huang, A.M., Rehm, E.J., and Rubin, G.M. (2009). Quick preparation of genomic DNA from Drosophila. Cold Spring Harb Protoc *2009*, pdb prot5198.

Huang, W.M., Ao, S.Z., Casjens, S., Orlandi, R., Zeikus, R., Weiss, R., Winge, D., and Fang, M. (1988). A persistent untranslated sequence within bacteriophage T4 DNA topoisomerase gene 60. Science *239*, 1005-1012.

Ibba, M., and Soll, D. (1999). Quality control mechanisms during translation. Science *286*, 1893-1897.

Ingolia, N.T., Brar, G.A., Rouskin, S., McGeachy, A.M., and Weissman, J.S. (2012). The ribosome profiling strategy for monitoring translation in vivo by deep sequencing of ribosome-protected mRNA fragments. Nat Protoc *7*, 1534-1550.

Ingolia, N.T., Ghaemmaghami, S., Newman, J.R., and Weissman, J.S. (2009). Genome-wide analysis in vivo of translation with nucleotide resolution using ribosome profiling. Science *324*, 218-223.

Irigoyen, N., Dinan, A.M., Brierley, I., and Firth, A.E. (2018). Ribosome profiling of the retrovirus murine leukemia virus. Retrovirology *15*, 10.

Ito, K., and Chiba, S. (2013). Arrest peptides: cis-acting modulators of translation. Annu Rev Biochem *82*, 171-202.

Ito, K., Ebihara, K., and Nakamura, Y. (1998). The stretch of C-terminal acidic amino acids of translational release factor eRF1 is a primary binding site for eRF3 of fission yeast. RNA *4*, 958-972.

Jacks, T., and Varmus, H.E. (1985). Expression of the Rous sarcoma virus pol gene by ribosomal frameshifting. Science *230*, 1237-1242.

Jackson, R.J., Hellen, C.U., and Pestova, T.V. (2010). The mechanism of eukaryotic translation initiation and principles of its regulation. Nat Rev Mol Cell Biol *11*, 113-127.

Jackson, R.J., Hellen, C.U., and Pestova, T.V. (2012). Termination and post-termination events in eukaryotic translation. Adv Protein Chem Struct Biol *86*, 45-93.

Jemc, J.C., Milutinovich, A.B., Weyers, J.J., Takeda, Y., and Van Doren, M. (2012). raw Functions through JNK signaling and cadherin-based adhesion to regulate Drosophila gonad morphogenesis. Dev Biol *367*, 114-125.

Jin, H., Kelley, A.C., Loakes, D., and Ramakrishnan, V. (2010). Structure of the 70S ribosome bound to release factor 2 and a substrate analog provides insights into catalysis of peptide release. Proc Natl Acad Sci U S A *107*, 8593-8598.

Jin, H., Kelley, A.C., and Ramakrishnan, V. (2011). Crystal structure of the hybrid state of ribosome in complex with the guanosine triphosphatase release factor 3. Proc Natl Acad Sci U S A *108*, 15798-15803.

Jukam, D., Xie, B., Rister, J., Terrell, D., Charlton-Perkins, M., Pistillo, D., Gebelein, B., Desplan, C., and Cook, T. (2013). Opposite feedbacks in the Hippo pathway for growth control and neural fate. Science *342*, 1238016.

Jungreis, I., Lin, M.F., Spokony, R., Chan, C.S., Negre, N., Victorsen, A., White, K.P., and Kellis, M. (2011). Evidence of abundant stop codon readthrough in Drosophila and other metazoa. Genome Res *21*, 2096-2113.

Kataoka, K., Han, S.I., Shioda, S., Hirai, M., Nishizawa, M., and Handa, H. (2002). MafA is a glucose-regulated and pancreatic beta-cell-specific transcriptional activator for the insulin gene. J Biol Chem *277*, 49903-49910.

Kataoka, K., Noda, M., and Nishizawa, M. (1994). Maf nuclear oncoprotein recognizes sequences related to an AP-1 site and forms heterodimers with both Fos and Jun. Mol Cell Biol *14*, 700-712.

Keeling, K.M., Lanier, J., Du, M., Salas-Marco, J., Gao, L., Kaenjak-Angeletti, A., and Bedwell, D.M. (2004). Leaky termination at premature stop codons antagonizes nonsense-mediated mRNA decay in S. cerevisiae. RNA *10*, 691-703.

Khoshnevis, S., Gross, T., Rotte, C., Baierlein, C., Ficner, R., and Krebber, H. (2010). The iron-sulphur protein RNase L inhibitor functions in translation termination. EMBO Rep *11*, 214-219.

Kim, E., Magen, A., and Ast, G. (2007). Different levels of alternative splicing among eukaryotes. Nucleic Acids Res *35*, 125-131.

Kisselev, L., Ehrenberg, M., and Frolova, L. (2003). Termination of translation: interplay of mRNA, rRNAs and release factors? EMBO J *22*, 175-182.

Kisselev, L.L., and Buckingham, R.H. (2000). Translational termination comes of age. Trends Biochem Sci *25*, 561-566.

Klagges, B.R., Heimbeck, G., Godenschwege, T.A., Hofbauer, A., Pflugfelder, G.O., Reifegerste, R., Reisch, D., Schaupp, M., Buchner, S., and Buchner, E. (1996). Invertebrate synapsins: a single gene codes for several isoforms in Drosophila. J Neurosci *16*, 3154-3165.

Kleppe, A.S., and Bornberg-Bauer, E. (2018). Robustness by intrinsically disordered C-termini and translational readthrough. Nucleic Acids Res *46*, 10184-10194.

Kong, C., Ito, K., Walsh, M.A., Wada, M., Liu, Y., Kumar, S., Barford, D., Nakamura, Y., and Song, H. (2004). Crystal structure and functional analysis of the eukaryotic class II release factor eRF3 from S. pombe. Mol Cell *14*, 233-245.

Kononenko, A.V., Mitkevich, V.A., Dubovaya, V.I., Kolosov, P.M., Makarov, A.A., and Kisselev, L.L. (2008). Role of the individual domains of translation termination factor eRF1 in GTP binding to eRF3. Proteins *70*, 388-393.

Konstantinides, N., Kapuralin, K., Fadil, C., Barboza, L., Satija, R., and Desplan, C. (2018). Phenotypic Convergence: Distinct Transcription Factors Regulate Common Terminal Features. Cell *174*, 622-635 e613.

Kornblihtt, A.R., Schor, I.E., Allo, M., Dujardin, G., Petrillo, E., and Munoz, M.J. (2013). Alternative splicing: a pivotal step between eukaryotic transcription and translation. Nat Rev Mol Cell Biol *14*, 153-165.

Korostelev, A., Asahara, H., Lancaster, L., Laurberg, M., Hirschi, A., Zhu, J., Trakhanov, S., Scott, W.G., and Noller, H.F. (2008). Crystal structure of a translation termination complex formed with release factor RF2. Proc Natl Acad Sci U S A *105*, 19684-19689.

Korostelev, A., Zhu, J., Asahara, H., and Noller, H.F. (2010). Recognition of the amber UAG stop codon by release factor RF1. EMBO J *29*, 2577-2585.

Koutmou, K.S., McDonald, M.E., Brunelle, J.L., and Green, R. (2014). RF3:GTP promotes rapid dissociation of the class 1 termination factor. RNA *20*, 609-620.

Kozlov, G., and Gehring, K. (2010). Molecular basis of eRF3 recognition by the MLLE domain of poly(A)-binding protein. PLoS One *5*, e10169.

Kramer, E.B., Vallabhaneni, H., Mayer, L.M., and Farabaugh, P.J. (2010). A comprehensive analysis of translational missense errors in the yeast Saccharomyces cerevisiae. RNA *16*, 1797-1808.

Kurokawa, H., Motohashi, H., Sueno, S., Kimura, M., Takagawa, H., Kanno, Y., Yamamoto, M., and Tanaka, T. (2009). Structural basis of alternative DNA recognition by Maf transcription factors. Mol Cell Biol *29*, 6232-6244.

Lamb, P., and McKnight, S.L. (1991). Diversity and specificity in transcriptional regulation: the benefits of heterotypic dimerization. Trends Biochem Sci *16*, 417-422.

Lang, B.F., Jakubkova, M., Hegedusova, E., Daoud, R., Forget, L., Brejova, B., Vinar, T., Kosa, P., Fricova, D., Nebohacova, M., *et al.* (2014). Massive programmed translational jumping in mitochondria. Proc Natl Acad Sci U S A *111*, 5926-5931.

Laurberg, M., Asahara, H., Korostelev, A., Zhu, J., Trakhanov, S., and Noller, H.F. (2008). Structural basis for translation termination on the 70S ribosome. Nature *454*, 852-857.

Lee, J.H., Pestova, T.V., Shin, B.S., Cao, C., Choi, S.K., and Dever, T.E. (2002). Initiation factor eIF5B catalyzes second GTP-dependent step in eukaryotic translation initiation. Proc Natl Acad Sci U S A *99*, 16689-16694.

Lemaitre, B., Nicolas, E., Michaut, L., Reichhart, J.M., and Hoffmann, J.A. (1996). The dorsoventral regulatory gene cassette spatzle/Toll/cactus controls the potent antifungal response in Drosophila adults. Cell *86*, 973-983.

Li, G., and Rice, C.M. (1993). The signal for translational readthrough of a UGA codon in Sindbis virus RNA involves a single cytidine residue immediately downstream of the termination codon. J Virol *67*, 5062-5067.

Li, M.A., Alls, J.D., Avancini, R.M., Koo, K., and Godt, D. (2003). The large Maf factor Traffic Jam controls gonad morphogenesis in Drosophila. Nat Cell Biol *5*, 994-1000.

Li, W., Tailhades, J., O'Brien-Simpson, N.M., Separovic, F., Otvos, L., Jr., Hossain, M.A., and Wade, J.D. (2014). Proline-rich antimicrobial peptides: potential therapeutics against antibiotic-resistant bacteria. Amino Acids *46*, 2287-2294.

Liebman, S.W., and Sherman, F. (1979). Extrachromosomal psi+ determinant suppresses nonsense mutations in yeast. J Bacteriol *139*, 1068-1071.

Lin, M.F., Carlson, J.W., Crosby, M.A., Matthews, B.B., Yu, C., Park, S., Wan, K.H., Schroeder, A.J., Gramates, L.S., St Pierre, S.E., *et al.* (2007). Revisiting the protein-coding gene catalog of Drosophila melanogaster using 12 fly genomes. Genome Res *17*, 1823-1836.

Ling, C., and Ermolenko, D.N. (2015). Initiation factor 2 stabilizes the ribosome in a semirotated conformation. Proc Natl Acad Sci U S A *112*, 15874-15879.

Ling, C., and Ermolenko, D.N. (2016). Structural insights into ribosome translocation. Wiley Interdiscip Rev RNA *7*, 620-636.

Llacer, J.L., Hussain, T., Marler, L., Aitken, C.E., Thakur, A., Lorsch, J.R., Hinnebusch, A.G., and Ramakrishnan, V. (2015). Conformational Differences between Open and Closed States of the Eukaryotic Translation Initiation Complex. Mol Cell *59*, 399-412.

Loughran, G., Chou, M.Y., Ivanov, I.P., Jungreis, I., Kellis, M., Kiran, A.M., Baranov, P.V., and Atkins, J.F. (2014). Evidence of efficient stop codon readthrough in four mammalian genes. Nucleic Acids Res *42*, 8928-8938.

Loughran, G., Jungreis, I., Tzani, I., Power, M., Dmitriev, R.I., Ivanov, I.P., Kellis, M., and Atkins, J.F. (2018). Stop codon readthrough generates a C-terminally extended variant of the human vitamin D receptor with reduced calcitriol response. J Biol Chem *293*, 4434-4444.

Mancera-Martinez, E., Brito Querido, J., Valasek, L.S., Simonetti, A., and Hashem, Y. (2017). ABCE1: A special factor that orchestrates translation at the crossroad between recycling and initiation. RNA Biol *14*, 1279-1285.

Manuvakhova, M., Keeling, K., and Bedwell, D.M. (2000). Aminoglycoside antibiotics mediate context-dependent suppression of termination codons in a mammalian translation system. RNA *6*, 1044-1055.

Matheisl, S., Berninghausen, O., Becker, T., and Beckmann, R. (2015). Structure of a human translation termination complex. Nucleic Acids Res *43*, 8615-8626.

McCaughan, K.K., Brown, C.M., Dalphin, M.E., Berry, M.J., and Tate, W.P. (1995). Translational termination efficiency in mammals is influenced by the base following the stop codon. Proc Natl Acad Sci U S A *92*, 5431-5435.

McNeill, H. (2000). Sticking together and sorting things out: adhesion as a force in development. Nat Rev Genet *1*, 100-108.

McNeill, H. (2003). Stuck without Traffic Jam. Nat Cell Biol *5*, 948-949.

Melnikov, S., Ben-Shem, A., Garreau de Loubresse, N., Jenner, L., Yusupova, G., and Yusupov, M. (2012). One core, two shells: bacterial and eukaryotic ribosomes. Nat Struct Mol Biol *19*, 560-567.

Merkulova, T.I., Frolova, L.Y., Lazar, M., Camonis, J., and Kisselev, L.L. (1999). C-terminal domains of human translation termination factors eRF1 and eRF3 mediate their in vivo interaction. Febs Lett *443*, 41-47.

Milon, P., and Rodnina, M.V. (2012). Kinetic control of translation initiation in bacteria. Crit Rev Biochem Mol Biol *47*, 334-348.

Mitkevich, V.A., Kononenko, A.V., Petrushanko, I.Y., Yanvarev, D.V., Makarov, A.A., and Kisselev, L.L. (2006). Termination of translation in eukaryotes is mediated by the quaternary eRF1*eRF3*GTP*Mg2+ complex. The biological roles of eRF3 and prokaryotic RF3 are profoundly distinct. Nucleic Acids Res *34*, 3947-3954.

Moras, D. (2010). Proofreading in translation: dynamics of the double-sieve model. Proc Natl Acad Sci U S A *107*, 21949-21950.

Morrison, S.J., and Spradling, A.C. (2008). Stem cells and niches: mechanisms that promote stem cell maintenance throughout life. Cell *132*, 598-611.

Motohashi, H., O'Connor, T., Katsuoka, F., Engel, J.D., and Yamamoto, M. (2002). Integration and diversity of the regulatory network composed of Maf and CNC families of transcription factors. Gene *294*, 1-12.

Mottagui-Tabar, S., Tuite, M.F., and Isaksson, L.A. (1998). The influence of 5' codon context on translation termination in Saccharomyces cerevisiae. Eur J Biochem *257*, 249-254.

Munroe, D., and Jacobson, A. (1990). mRNA poly(A) tail, a 3' enhancer of translational initiation. Mol Cell Biol *10*, 3441-3455.

Myasnikov, A.G., Marzi, S., Simonetti, A., Giuliodori, A.M., Gualerzi, C.O., Yusupova, G., Yusupov, M., and Klaholz, B.P. (2005). Conformational transition of initiation factor 2 from the GTP- to GDP-bound state visualized on the ribosome. Nat Struct Mol Biol *12*, 1145-1149.

Namy, O., Duchateau-Nguyen, G., Hatin, I., Hermann-Le Denmat, S., Termier, M., and Rousset, J.P. (2003). Identification of stop codon readthrough genes in Saccharomyces cerevisiae. Nucleic Acids Res *31*, 2289-2296.

Namy, O., Duchateau-Nguyen, G., and Rousset, J.P. (2002). Translational readthrough of the PDE2 stop codon modulates cAMP levels in Saccharomyces cerevisiae. Mol Microbiol *43*, 641-652.

Namy, O., Hatin, I., and Rousset, J.P. (2001). Impact of the six nucleotides downstream of the stop codon on translation termination. EMBO Rep *2*, 787-793.

Namy, O., Rousset, J.P., Napthine, S., and Brierley, I. (2004). Reprogrammed genetic decoding in cellular gene expression. Mol Cell *13*, 157-168.

Napthine, S., Yek, C., Powell, M.L., Brown, T.D., and Brierley, I. (2012). Characterization of the stop codon readthrough signal of Colorado tick fever virus segment 9 RNA. RNA *18*, 241-252.

Neu-Yilik, G., Raimondeau, E., Eliseev, B., Yeramala, L., Amthor, B., Deniaud, A., Huard, K., Kerschgens, K., Hentze, M.W., Schaffitzel, C.*, et al.* (2017). Dual function of UPF3B in early and late translation termination. EMBO J *36*, 2968-2986.

Nicholson, A.L., and Pasquinelli, A.E. (2018). Tales of Detailed Poly(A) Tails. Trends Cell Biol.

Ogino, H., and Yasuda, K. (1998). Induction of lens differentiation by activation of a bZIP transcription factor, L-Maf. Science *280*, 115-118.

Ogle, J.M., Brodersen, D.E., Clemons, W.M., Jr., Tarry, M.J., Carter, A.P., and Ramakrishnan, V. (2001). Recognition of cognate transfer RNA by the 30S ribosomal subunit. Science *292*, 897-902.

Ogle, J.M., Murphy, F.V., Tarry, M.J., and Ramakrishnan, V. (2002). Selection of tRNA by the ribosome requires a transition from an open to a closed form. Cell *111*, 721-732.

Okegbe, T.C., and DiNardo, S. (2011). The endoderm specifies the mesodermal niche for the germline in Drosophila via Delta-Notch signaling. Development *138*, 1259-1267.

Pallesen, J., Hashem, Y., Korkmaz, G., Koripella, R.K., Huang, C., Ehrenberg, M., Sanyal, S., and Frank, J. (2013). Cryo-EM visualization of the ribosome in termination complex with apo-RF3 and RF1. Elife *2*, e00411.

Panchal, T., Chen, X., Alchits, E., Oh, Y., Poon, J., Kouptsova, J., Laski, F.A., and Godt, D. (2017). Specification and spatial arrangement of cells in the germline stem cell niche of the Drosophila ovary depend on the Maf transcription factor Traffic jam. PLoS Genet *13*, e1006790.

Pancsa, R., Macossay-Castillo, M., Kosol, S., and Tompa, P. (2016). Computational analysis of translational readthrough proteins in Drosophila and yeast reveals parallels to alternative splicing. Sci Rep *6*, 32142.

Pape, T., Wintermeyer, W., and Rodnina, M.V. (1998). Complete kinetic mechanism of elongation factor Tu-dependent binding of aminoacyl-tRNA to the A site of the E. coli ribosome. EMBO J *17*, 7490-7497.

Paulin, F.E., Campbell, L.E., O'Brien, K., Loughlin, J., and Proud, C.G. (2001). Eukaryotic translation initiation factor 5 (eIF5) acts as a classical GTPase-activator protein. Curr Biol *11*, 55-59.

Paushkin, S.V., Kushnirov, V.V., Smirnov, V.N., and Ter-Avanesyan, M.D. (1996). Propagation of the yeast prion-like [psi+] determinant is mediated by oligomerization of the SUP35-encoded polypeptide chain release factor. EMBO J *15*, 3127-3134.

Pedersen, W.T., and Curran, J.F. (1991). Effects of the nucleotide 3' to an amber codon on ribosomal selection rates of suppressor tRNA and release factor-1. J Mol Biol *219*, 231-241.

Pelham, H.R. (1978). Leaky UAG termination codon in tobacco mosaic virus RNA. Nature *272*, 469-471.

Pennell, S., Manktelow, E., Flatt, A., Kelly, G., Smerdon, S.J., and Brierley, I. (2008). The stimulatory RNA of the Visna-Maedi retrovirus ribosomal frameshifting signal is an unusual pseudoknot with an interstem element. RNA *14*, 1366-1377.

Peske, F., Kuhlenkoetter, S., Rodnina, M.V., and Wintermeyer, W. (2014). Timing of GTP binding and hydrolysis by translation termination factor RF3. Nucleic Acids Res *42*, 1812-1820.

Peske, F., Rodnina, M.V., and Wintermeyer, W. (2005). Sequence of steps in ribosome recycling as defined by kinetic analysis. Mol Cell *18*, 403-412.

Pestova, T.V., Lomakin, I.B., Lee, J.H., Choi, S.K., Dever, T.E., and Hellen, C.U. (2000). The joining of ribosomal subunits in eukaryotes requires eIF5B. Nature *403*, 332-335.

Petry, S., Brodersen, D.E., Murphy, F.V.t., Dunham, C.M., Selmer, M., Tarry, M.J., Kelley, A.C., and Ramakrishnan, V. (2005). Crystal structures of the ribosome in complex with release factors RF1 and RF2 bound to a cognate stop codon. Cell *123*, 1255-1266.

Pimentel, D., Donlea, J.M., Talbot, C.B., Song, S.M., Thurston, A.J.F., and Miesenbock, G. (2016). Operation of a homeostatic sleep switch. Nature *536*, 333-337.

Pisarev, A.V., Hellen, C.U., and Pestova, T.V. (2007). Recycling of eukaryotic posttermination ribosomal complexes. Cell *131*, 286-299.

Pisarev, A.V., Skabkin, M.A., Pisareva, V.P., Skabkina, O.V., Rakotondrafara, A.M., Hentze, M.W., Hellen, C.U., and Pestova, T.V. (2010). The role of ABCE1 in eukaryotic posttermination ribosomal recycling. Mol Cell *37*, 196-210.

Pisareva, V.P., Pisarev, A.V., Hellen, C.U., Rodnina, M.V., and Pestova, T.V. (2006). Kinetic analysis of interaction of eukaryotic release factor 3 with guanine nucleotides. J Biol Chem *281*, 40224-40235.

Plant, E.P., and Dinman, J.D. (2006). Comparative study of the effects of heptameric slippery site composition on -1 frameshifting among different eukaryotic systems. RNA *12*, 666-673.

Plant, E.P., Rakauskaite, R., Taylor, D.R., and Dinman, J.D. (2010). Achieving a golden mean: mechanisms by which coronaviruses ensure synthesis of the correct stoichiometric ratios of viral proteins. J Virol *84*, 4330-4340.

Preis, A., Heuer, A., Barrio-Garcia, C., Hauser, A., Eyler, D.E., Berninghausen, O., Green, R., Becker, T., and Beckmann, R. (2014). Cryoelectron microscopic structures of eukaryotic translation termination complexes containing eRF1-eRF3 or eRF1-ABCE1. Cell Rep *8*, 59-65.

Rehemtulla, A., Warwar, R., Kumar, R., Ji, X., Zack, D.J., and Swaroop, A. (1996). The basic motif-leucine zipper transcription factor Nrl can positively regulate rhodopsin gene expression. Proc Natl Acad Sci U S A *93*, 191-195.

Robine, N., Lau, N.C., Balla, S., Jin, Z., Okamura, K., Kuramochi-Miyagawa, S., Blower, M.D., and Lai, E.C. (2009). A broadly conserved pathway generates 3'UTR-directed primary piRNAs. Curr Biol *19*, 2066-2076.

Robinson, D.N., and Cooley, L. (1997). Examination of the function of two kelch proteins generated by stop codon suppression. Development *124*, 1405-1417.

Rodnina, M.V. (2012). Quality control of mRNA decoding on the bacterial ribosome. Adv Protein Chem Struct Biol *86*, 95-128.

Rodnina, M.V., Fischer, N., Maracci, C., and Stark, H. (2017). Ribosome dynamics during decoding. Philos Trans R Soc Lond B Biol Sci *372*.

Rodnina, M.V., Pape, T., Fricke, R., Kuhn, L., and Wintermeyer, W. (1996). Initial binding of the elongation factor Tu.GTP.aminoacyl-tRNA complex preceding codon recognition on the ribosome. J Biol Chem *271*, 646-652.

Rodnina, M.V., Pape, T., Fricke, R., and Wintermeyer, W. (1995). Elongation factor Tu, a GTPase triggered by codon recognition on the ribosome: mechanism and GTP consumption. Biochem Cell Biol *73*, 1221-1227.

Rodnina, M.V., and Wintermeyer, W. (2001). Ribosome fidelity: tRNA discrimination, proofreading and induced fit. Trends Biochem Sci *26*, 124-130.

Rodnina, M.V., and Wintermeyer, W. (2009). Recent mechanistic insights into eukaryotic ribosomes. Curr Opin Cell Biol *21*, 435-443.

Roy, B., Leszyk, J.D., Mangus, D.A., and Jacobson, A. (2015a). Nonsense suppression by near-cognate tRNAs employs alternative base pairing at codon positions 1 and 3. Proc Natl Acad Sci U S A *112*, 3038-3043.

Roy, R.N., Lomakin, I.B., Gagnon, M.G., and Steitz, T.A. (2015b). The mechanism of inhibition of protein synthesis by the proline-rich peptide oncocin. Nat Struct Mol Biol *22*, 466-469.

Ryan, M.D., and Drew, J. (1994). Foot-and-mouth disease virus 2A oligopeptide mediated cleavage of an artificial polyprotein. EMBO J *13*, 928-933.

Saint-Leger, A., and Ribas de Pouplana, L. (2015). The importance of codon-anticodon interactions in translation elongation. Biochimie *114*, 72-79.

Saito, K., Inagaki, S., Mituyama, T., Kawamura, Y., Ono, Y., Sakota, E., Kotani, H., Asai, K., Siomi, H., and Siomi, M.C. (2009). A regulatory circuit for piwi by the large Maf gene traffic jam in Drosophila. Nature *461*, 1296-1299.

Salas-Marco, J., and Bedwell, D.M. (2004). GTP hydrolysis by eRF3 facilitates stop codon decoding during eukaryotic translation termination. Mol Cell Biol *24*, 7769-7778.

Samatova, E., Konevega, A.L., Wills, N.M., Atkins, J.F., and Rodnina, M.V. (2014). High-efficiency translational bypassing of non-coding nucleotides specified by mRNA structure and nascent peptide. Nat Commun *5*, 4459.

Samson, M.L., Lisbin, M.J., and White, K. (1995). Two distinct temperature-sensitive alleles at the elav locus of Drosophila are suppressed nonsense mutations of the same tryptophan codon. Genetics *141*, 1101-1111.

Sarker, S., Rudd, K.E., and Oliver, D. (2000). Revised translation start site for secM defines an atypical signal peptide that regulates Escherichia coli secA expression. J Bacteriol *182*, 5592-5595.

Savelsbergh, A., Rodnina, M.V., and Wintermeyer, W. (2009). Distinct functions of elongation factor G in ribosome recycling and translocation. RNA *15*, 772-780.

Schueren, F., Lingner, T., George, R., Hofhuis, J., Dickel, C., Gartner, J., and Thoms, S. (2014). Peroxisomal lactate dehydrogenase is generated by translational readthrough in mammals. Elife *3*, e03640.

Schueren, F., and Thoms, S. (2016). Functional Translational Readthrough: A Systems Biology Perspective. PLoS Genet *12*, e1006196.

Schuller, A.P., and Green, R. (2018). Roadblocks and resolutions in eukaryotic translation. Nat Rev Mol Cell Biol *19*, 526-541.

Seefeldt, A.C., Graf, M., Perebaskine, N., Nguyen, F., Arenz, S., Mardirossian, M., Scocchi, M., Wilson, D.N., and Innis, C.A. (2016). Structure of the mammalian antimicrobial peptide Bac7(1-16) bound within the exit tunnel of a bacterial ribosome. Nucleic Acids Res *44*, 2429-2438.

Seit-Nebi, A., Frolova, L., Justesen, J., and Kisselev, L. (2001). Class-1 translation termination factors: invariant GGQ minidomain is essential for release activity and ribosome binding but not for stop codon recognition. Nucleic Acids Res *29*, 3982-3987.

Seo, H.S., Kiel, M., Pan, D., Raj, V.S., Kaji, A., and Cooperman, B.S. (2004). Kinetics and thermodynamics of RRF, EF-G, and thiostrepton interaction on the Escherichia coli ribosome. Biochemistry *43*, 12728-12740.

Shao, S., Murray, J., Brown, A., Taunton, J., Ramakrishnan, V., and Hegde, R.S. (2016). Decoding Mammalian Ribosome-mRNA States by Translational GTPase Complexes. Cell *167*, 1229-1240 e1215.

Sharma, H., Adio, S., Senyushkina, T., Belardinelli, R., Peske, F., and Rodnina, M.V. (2016). Kinetics of Spontaneous and EF-G-Accelerated Rotation of Ribosomal Subunits. Cell Rep *16*, 2187-2196.

Shaw, J.J., and Green, R. (2007). Two distinct components of release factor function uncovered by nucleophile partitioning analysis. Mol Cell *28*, 458-467.

Sherf, B.A., Navarro, S.L., Hannah, R.R., and Wood, K.V. (1996). Novel reporter gene technology integrating firefly and Renilla luciferase assays for internal control. Mol Biol Cell *7*, 925-925.

Shi, X., and Joseph, S. (2016). Mechanism of Translation Termination: RF1 Dissociation Follows Dissociation of RF3 from the Ribosome. Biochemistry *55*, 6344-6354.

Shoemaker, C.J., and Green, R. (2011). Kinetic analysis reveals the ordered coupling of translation termination and ribosome recycling in yeast. Proc Natl Acad Sci U S A *108*, E1392-1398.

Skuzeski, J.M., Nichols, L.M., Gesteland, R.F., and Atkins, J.F. (1991). The signal for a leaky UAG stop codon in several plant viruses includes the two downstream codons. J Mol Biol *218*, 365-373.

Song, H., Mugnier, P., Das, A.K., Webb, H.M., Evans, D.R., Tuite, M.F., Hemmings, B.A., and Barford, D. (2000). The crystal structure of human eukaryotic release factor eRF1--mechanism of stop codon recognition and peptidyl-tRNA hydrolysis. Cell *100*, 311-321.

Song, H., Parsons, M.R., Rowsell, S., Leonard, G., and Phillips, S.E. (1999). Crystal structure of intact elongation factor EF-Tu from Escherichia coli in GDP conformation at 2.05 A resolution. J Mol Biol *285*, 1245-1256.

Song, X., Wong, M.D., Kawase, E., Xi, R., Ding, B.C., McCarthy, J.J., and Xie, T. (2004). Bmp signals from niche cells directly repress transcription of a differentiation-promoting gene, bag of marbles, in germline stem cells in the Drosophila ovary. Development *131*, 1353-1364.

Stansfield, I., Jones, K.M., Herbert, P., Lewendon, A., Shaw, W.V., and Tuite, M.F. (1998). Missense translation errors in Saccharomyces cerevisiae. J Mol Biol *282*, 13-24.

Stark, A., Lin, M.F., Kheradpour, P., Pedersen, J.S., Parts, L., Carlson, J.W., Crosby, M.A., Rasmussen, M.D., Roy, S., Deoras, A.N., *et al.* (2007). Discovery of functional elements in 12 Drosophila genomes using evolutionary signatures. Nature *450*, 219-232.

Steneberg, P., Englund, C., Kronhamn, J., Weaver, T.A., and Samakovlis, C. (1998). Translational readthrough in the hdc mRNA generates a novel branching inhibitor in the drosophila trachea. Genes Dev *12*, 956-967.

Sternberg, S.H., Fei, J., Prywes, N., McGrath, K.A., and Gonzalez, R.L., Jr. (2009). Translation factors direct intrinsic ribosome dynamics during translation termination and ribosome recycling. Nat Struct Mol Biol *16*, 861-868.

Stiebler, A.C., Freitag, J., Schink, K.O., Stehlik, T., Tillmann, B.A., Ast, J., and Bolker, M. (2014). Ribosomal readthrough at a short UGA stop codon context triggers dual localization of metabolic enzymes in Fungi and animals. PLoS Genet *10*, e1004685.

Sun, J., Liu, C., Bai, X., Li, X., Li, J., Zhang, Z., Zhang, Y., Guo, J., and Li, Y. (2017). Drosophila FIT is a protein-specific satiety hormone essential for feeding control. Nat Commun *8*, 14161.

Tate, W.P., Poole, E.S., Horsfield, J.A., Mannering, S.A., Brown, C.M., Moffat, J.G., Dalphin, M.E., McCaughan, K.K., Major, L.L., and Wilson, D.N. (1995). Translational termination efficiency in both bacteria and mammals is regulated by the base following the stop codon. Biochem Cell Biol *73*, 1095-1103.

Taylor, D.J., Nilsson, J., Merrill, A.R., Andersen, G.R., Nissen, P., and Frank, J. (2007). Structures of modified eEF2 80S ribosome complexes reveal the role of GTP hydrolysis in translocation. EMBO J *26*, 2421-2431.

Tian, B., and Manley, J.L. (2017). Alternative polyadenylation of mRNA precursors. Nat Rev Mol Cell Biol *18*, 18-30.

Tork, S., Hatin, I., Rousset, J.P., and Fabret, C. (2004). The major 5' determinant in stop codon read-through involves two adjacent adenines. Nucleic Acids Res *32*, 415-421.

Touriol, C., Bornes, S., Bonnal, S., Audigier, S., Prats, H., Prats, A.C., and Vagner, S. (2003). Generation of protein isoform diversity by alternative initiation of translation at non-AUG codons. Biol Cell *95*, 169-178.

Trotta, E. (2016). Selective forces and mutational biases drive stop codon usage in the human genome: a comparison with sense codon usage. BMC Genomics *17*, 366.

True, H.L., and Lindquist, S.L. (2000). A yeast prion provides a mechanism for genetic variation and phenotypic diversity. Nature *407*, 477-483.

Tsuchihashi, Z., and Kornberg, A. (1990). Translational frameshifting generates the gamma subunit of DNA polymerase III holoenzyme. Proc Natl Acad Sci U S A *87*, 2516-2520.

Tyedmers, J., Madariaga, M.L., and Lindquist, S. (2008). Prion switching in response to environmental stress. PLoS Biol *6*, e294.

Urban, C., Zerfass, K., Fingerhut, C., and Beier, H. (1996). UGA suppression by tRNACmCATrp occurs in diverse virus RNAs due to a limited influence of the codon context. Nucleic Acids Res *24*, 3424-3430.

Vieira, F.G., Sanchez-Gracia, A., and Rozas, J. (2007). Comparative genomic analysis of the odorant-binding protein family in 12 Drosophila genomes: purifying selection and birth-and-death evolution. Genome Biol *8*, R235.

Voog, J., D'Alterio, C., and Jones, D.L. (2008). Multipotent somatic stem cells contribute to the stem cell niche in the Drosophila testis. Nature *454*, 1132-1136.

Voolstra, O., Kiefer, C., Hoehne, M., Welsch, R., Vogt, K., and von Lintig, J. (2006). The Drosophila class B scavenger receptor NinaD-I is a cell surface receptor mediating carotenoid transport for visual chromophore synthesis. Biochemistry *45*, 13429-13437.

Waghmare, I., and Page-McCaw, A. (2018). Wnt Signaling in Stem Cell Maintenance and Differentiation in the Drosophila Germarium. Genes (Basel) *9*.

Wang, P., Lyman, R.F., Mackay, T.F., and Anholt, R.R. (2010). Natural variation in odorant recognition among odorant-binding proteins in Drosophila melanogaster. Genetics *184*, 759-767.

Weiner, A.M., and Weber, K. (1973). A single UGA codon functions as a natural termination signal in the coliphage q beta coat protein cistron. J Mol Biol *80*, 837-855.

Weiss, R.B., Huang, W.M., and Dunn, D.M. (1990). A nascent peptide is required for ribosomal bypass of the coding gap in bacteriophage T4 gene 60. Cell *62*, 117-126.

WeixIbaumer, A., Jin, H., Neubauer, C., Voorhees, R.M., Petry, S., Kelley, A.C., and Ramakrishnan, V. (2008). Insights into translational termination from the structure of RF2 bound to the ribosome. Science *322*, 953-956.

White-Cooper, H. (2010). Molecular mechanisms of gene regulation during Drosophila spermatogenesis. Reproduction *139*, 11-21.

Williams, I., Richardson, J., Starkey, A., and Stansfield, I. (2004). Genome-wide prediction of stop codon readthrough during translation in the yeast Saccharomyces cerevisiae. Nucleic Acids Res *32*, 6605-6616.

Wills, N.M., Gesteland, R.F., and Atkins, J.F. (1991). Evidence that a downstream pseudoknot is required for translational read-through of the Moloney murine leukemia virus gag stop codon. Proc Natl Acad Sci U S A *88*, 6991-6995.

Wills, N.M., Gesteland, R.F., and Atkins, J.F. (1994). Pseudoknot-dependent read-through of retroviral gag termination codons: importance of sequences in the spacer and loop 2. EMBO J *13*, 4137-4144.

Wilson, D.N., Arenz, S., and Beckmann, R. (2016). Translation regulation via nascent polypeptide-mediated ribosome stalling. Curr Opin Struct Biol *37*, 123-133.

Wingert, L., and DiNardo, S. (2015). Traffic jam functions in a branched pathway from Notch activation to niche cell fate. Development *142*, 2268-2277.

Xie, T., and Spradling, A.C. (1998). decapentaplegic is essential for the maintenance and division of germline stem cells in the Drosophila ovary. Cell *94*, 251-260.

Xue, F., and Cooley, L. (1993). kelch encodes a component of intercellular bridges in Drosophila egg chambers. Cell *72*, 681-693.

Yamaguchi, Y., and Baba, H. (2018). Phylogenetically Conserved Sequences Around Myelin P0 Stop Codon are Essential for Translational Readthrough to Produce L-MPZ. Neurochem Res *43*, 227-237.

Yamaguchi, Y., Hayashi, A., Campagnoni, C.W., Kimura, A., Inuzuka, T., and Baba, H. (2012). L-MPZ, a novel isoform of myelin P0, is produced by stop codon readthrough. J Biol Chem *287*, 17765-17776.

Yang, J., and O'Tousa, J.E. (2007). Cellular sites of Drosophila NinaB and NinaD activity in vitamin A metabolism. Mol Cell Neurosci *35*, 49-56.

Yoshinaka, Y., Katoh, I., Copeland, T.D., and Oroszlan, S. (1985). Murine leukemia virus protease is encoded by the gag-pol gene and is synthesized through suppression of an amber termination codon. Proc Natl Acad Sci U S A *82*, 1618-1622.

Zavialov, A.V., Buckingham, R.H., and Ehrenberg, M. (2001). A posttermination ribosomal complex is the guanine nucleotide exchange factor for peptide release factor RF3. Cell *107*, 115-124.

Zavialov, A.V., Mora, L., Buckingham, R.H., and Ehrenberg, M. (2002). Release of peptide promoted by the GGQ motif of class 1 release factors regulates the GTPase activity of RF3. Mol Cell *10*, 789-798.

Zhouravleva, G., Frolova, L., Le Goff, X., Le Guellec, R., Inge-Vechtomov, S., Kisselev, L., and Philippe, M. (1995). Termination of translation in eukaryotes is governed by two interacting polypeptide chain release factors, eRF1 and eRF3. EMBO J *14*, 4065-4072.

# 6. APPENDIX

## 6.1 Supplementary tables

Table S1. Genes identified to be dysregulated in *tj*-TR mutants

| $tj^{nat/nat}$ vs $tj^{TR/TR}$ | | | |
|---|---|---|---|
| Gene ID | Gene Name | Log2 fold change | p-value |
| FBgn0013673 | mt:ATPase8 | -2.12 | 4.44E-67 |
| FBgn0013680 | mt:ND2 | -1.69 | 6.85E-47 |
| FBgn0033792 | CG13325 | -1.43 | 2.58E-34 |
| FBgn0029831 | CG5966 | -1.16 | 3.84E-21 |
| FBgn0013683 | mt:ND4L | -1.08 | 5.83E-19 |
| FBgn0260446 | GABA-B-R1 | -1.03 | 2.58E-29 |
| FBgn0033830 | CG10814 | -1.00 | 3.77E-20 |
| FBgn0027348 | bgm | 1.02 | 2.72E-26 |
| FBgn0033760 | CG8785 | 1.03 | 3.57E-17 |
| FBgn0031435 | Elba2 | 1.08 | 9.90E-19 |
| FBgn0033683 | CG18343 | 1.31 | 2.24E-31 |
| FBgn0032620 | CG12288 | 1.33 | 2.17E-29 |
| FBgn0032706 | Irk3 | 1.71 | 4.87E-93 |
| FBgn0033257 | sand | 1.84 | 3.42E-51 |

| $tj^{nat/nat}$ vs $tj^{nTR/nTR}$ | | | |
|---|---|---|---|
| Gene ID | Gene name | Log2 fold change | p-value |
| FBgn0037801 | CG3999 | -1.63 | 4.07E-36 |
| FBgn0001187 | Hex-C | -1.56 | 1.46E-74 |
| FBgn0264979 | CG4267 | -1.46 | 9.77E-26 |
| FBgn0033792 | CG13325 | -1.30 | 1.15E-23 |
| FBgn0039241 | CG11089 | -1.27 | 1.18E-25 |
| FBgn0038467 | AdSL | -1.20 | 1.73E-26 |
| FBgn0029831 | CG5966 | -1.16 | 1.57E-17 |
| FBgn0029823 | Shmt | -1.16 | 4.32E-28 |
| FBgn0033885 | DJ-1alpha | -1.03 | 1.77E-18 |
| FBgn0035638 | Tektin-C | 1.04 | 9.79E-14 |
| FBgn0031434 | insv | 1.06 | 6.34E-16 |
| FBgn0260475 | CG30059 | 1.23 | 6.62E-21 |
| FBgn0035789 | mthl6 | 1.26 | 6.91E-21 |
| FBgn0031435 | Elba2 | 1.41 | 7.74E-25 |
| FBgn0000964 | tj | 1.43 | 1.85E-63 |
| FBgn0267635 | CR45973 | 1.64 | 4.59E-33 |
| FBgn0267160 | CR45600 | 2.12 | 4.80E-53 |

| $tj^{TR/TR}$ vs $tj^{nTR/nTR}$ | | | |
|---|---|---|---|
| Gene ID | Gene name | Log2 fold change | p-value |
| FBgn0033257 | sand | -2.48 | 1.39E-53 |
| FBgn0031701 | TotM | -1.72 | 1.05E-26 |
| FBgn0262881 | CG43236 | -1.70 | 8.72E-27 |
| FBgn0032706 | Irk3 | -1.54 | 9.95E-39 |
| FBgn0032620 | CG12288 | -1.52 | 2.91E-32 |
| FBgn0038083 | CG5999 | -1.48 | 1.21E-26 |
| FBgn0001187 | Hex-C | -1.44 | 2.66E-55 |
| FBgn0032754 | CG10700 | -1.42 | 5.76E-20 |
| FBgn0032684 | CG10178 | -1.26 | 2.42E-23 |
| FBgn0050090 | CG30090 | -1.13 | 1.80E-12 |
| FBgn0034512 | CG18067 | -1.10 | 5.10E-36 |
| FBgn0044810 | TotX | -1.10 | 4.69E-12 |
| FBgn0040349 | CG3699 | -1.09 | 5.51E-13 |
| FBgn0002939 | ninaD | -1.08 | 5.47E-12 |
| FBgn0020513 | ade5 | -1.01 | 3.76E-20 |
| FBgn0053310 | CG33310 | 1.00 | 2.67E-10 |

| | | | |
|---|---|---|---|
| FBgn0013688 | mt:srRNA | 1.00 | 3.25E-10 |
| FBgn0260446 | GABA-B-R1 | 1.09 | 3.45E-20 |
| FBgn0051832 | CG31832 | 1.10 | 2.44E-12 |
| FBgn0050083 | CG30083 | 1.31 | 4.84E-16 |
| FBgn0000964 | tj | 1.61 | 6.37E-59 |
| FBgn0013680 | mt:ND2 | 1.74 | 1.17E-29 |
| FBgn0013683 | mt:ND4L | 1.79 | 4.80E-29 |
| FBgn0267635 | CR45973 | 1.92 | 1.16E-32 |
| FBgn0035638 | Tektin-C | 2.33 | 1.52E-51 |
| FBgn0013673 | mt:ATPase8 | 2.76 | 9.95E-76 |

Table S2. List of primers used for generating dual luciferase constructs for candidate TR genes

| Name | Primer sequence 5' to 3' | Comments |
|---|---|---|
| PK96_F | GTGCTGAAGAACGAGCAGCTGAGCTTGTACGACGATCGGATG | Insertion of *wit* TR motif into psiCHECK™ -2 vector with 18 bp overhang for Gibson cloning |
| PK97_R | GTTGGTGGCGCCGGAGCCGTTCTGCTGCATTCGATTAGTTTATAGCTCC | |
| PK100_F | GTGCTGAAGAACGAGCAGAACGGAGCCTACCACCACGG | Insertion of *dsx* TR motif into psiCHECK™ -2 vector with 18 bp overhang for Gibson cloning |
| PK101_R | GTTGGTGGCGCCGGAGCCGACAGCGGCCGCTGC | |
| PK102_F | GTGCTGAAGAACGAGCAGCAATTGCAGCCGCAACAC | Insertion of *fru* TR motif into psiCHECK™ -2 vector with 18 bp overhang for Gibson cloning |
| PK103_R | GTTGGTGGCGCCGGAGCCGGGGTCATCGGGACGC | |
| PK104_F | GTGCTGAAGAACGAGCAGATGACACGCTCCAAGAGCC | Insertion of *khc-73* TR motif into psiCHECK™ -2 vector with 18 bp overhang for Gibson cloning |
| PK105_R | GTTGGTGGCGCCGGAGCCCCTGCAATTAGTCCAACGCTGCAGC | |
| PK106_F | GTGCTGAAGAACGAGCAGGCAGCTGCAGCATCAGCAGCGG | Insertion of *chinmo* TR motif into psiCHECK™ -2 vector with 18 bp overhang for Gibson cloning |
| PK107_R | GTTGGTGGCGCCGGAGCCCTCCTTGTTGGCGTTCATGACTACTGA | |
| PK108_F | GTGCTGAAGAACGAGCAGCTAACCCTGGGTGGACCCATG | Insertion of *klu* TR motif into psiCHECK™ -2 vector with 18 bp overhang for Gibson cloning |
| PK109_R | GTTGGTGGCGCCGGAGCCACAGGTCATAAATGGTCTGGATGCTG | |
| PK110_F | GTGCTGAAGAACGAGCAGCAGCAGCAGCAACAGTC | Insertion of *br* TR motif into psiCHECK™ -2 vector with 18 bp overhang for Gibson cloning |
| PK111_R | GTTGGTGGCGCCGGAGCCGGAGTTGTTGAGCGCCAC | |
| PK114_F | GTGCTGAAGAACGAGCAGGATATGCTGCTGAGCGGCAAC | Insertion of *svp* TR motif into psiCHECK™ -2 vector with 18 bp overhang for Gibson cloning |
| PK115_R | GTTGGTGGCGCCGGAGCCAGTTGTTGTCAATTGGCGCCACATCGTG | |
| PK156_F | AATCAGCAGCAACTCTTGCAGC | UAAG to UAAA mutation in *br* TR motif using blunt end ligation |
| PK155_R | TTATAAGAAGTCCATGCACGGTTTGACAATGC | |
| PK158_F | TCGATCAGCAGCAACTCTTGCAGC | UAA to UUC mutation in *br* TR motif using blunt end ligation |
| PK157_R | ATAAGAAGTCCATGCACGGTTTGACAATGC | |
| PK160_F | AAAAGCAGCCGCAACAGC | UAGG to UAAA mutation in *chinmo* TR motif using blunt end ligation |
| PK159_R | TATGGTGAATGATTGCTGGCTGCC | |
| PK162_F | TCGAAGCAGCCGCAACAG | UAA to UUC mutation in *chinmo* TR motif using blunt end ligation |
| PK161_R | ATGGTGAATGATTGCTGGCTGC | |
| PK164_F | AAAGTATCGCAACGTTGCTGC | UAGC to UAAA mutation in *dsx* TR motif using blunt end ligation |
| PK163_R | TACGTGGCAGCCGTGGAG | |
| PK166_F | TCCAGTATCGCAACGTTGCTG | UAG to UUC mutation in *dsx* TR motif using blunt end ligation |
| PK165_R | ACGTGGCAGCCGTGGA | |
| PK168_F | AAAACAGTCAGTACCTGGGCTGGA | UGAU to UAAA mutation in *fru* TR motif using blunt end ligation |
| PK167_R | ATTCACTTGTGGCATTGTGCTGC | |
| CM183_F | TCTACAGTCAGTACCTGGGCTGGAACTACGGCG | UGA to UUC mutation in *fru* TR motif using blunt end ligation |
| CM184_R | ATTCACTTGTGGCATTGTGCTGCTGCTG | |
| PK176_F | AAATGTACCCAAAGTGTTCGCATCAG | UGAU to UAAA mutation in *khc-73* TR motif using blunt end ligation |
| PK175_R | ATTTACGCGCCGAAAGGTTTAGC | |
| PK178_F | TCTTGTACCCAAAGTGTTCGCATCAGC | UGA to UUC mutation in *khc-73* TR motif using blunt end ligation |
| PK177_R | ATTTACGCGCCGAAAGGTTTAGC | |
| PK180_F | AGGTGTCTGTATGCAGCAGC | UAAC to UAAA mutation in *klu* TR motif using blunt end ligation |
| PK179_R | TTAGGCGCTCTCCGTCTTGACAAC | |
| PK182_F | TCCGGTGTCTGTATGCAGCAGC | UAA to UUC mutation in *klu* TR motif using blunt end ligation |
| PK181_R | AGGCGCTCTCCGTCTTGAC | |
| PK188_F | AAATGCCTTCGATGTGACACACGA | UGAC to UAAA mutation in *svp* TR motif using blunt end ligation |
| PK187_R | AGGGCCAGGAGAAACTGTTGC | |
| PK190_F | TCCTGCCTTCGATGTGACACACG | UGA to UUC mutation in *svp* TR motif using blunt end ligation |
| PK189_R | AGGGCCAGGAGAAACTGTTGC | |
| PK192_F | AAAATGAGGAGGTTCTGCTGC | UAGC to UAAA mutation in *wit* TR motif using blunt end ligation |
| PK191_R | AGAGAATGTTGAGCAGGGAGGAGT | |

| | | |
|---|---|---|
| PK194_F | TCCATGAGGAGGTTCTGCTGC | UAG to UUC mutation in *wit* TR motif using blunt end ligation |
| PK193_R | AGAGAATGTTGAGCAGGGAGGAGT | |
| PK122_F | GCGAGGGTGAGGGCGCTGAGGAGTACTGACACCACGAAATGTGC | Replacing sequences upstream of stop codon in aPKC TR motif with *αTub84B* in constructs with UGA-C SCC |
| PK123_R | CGTCACCGGAGTCCATGCCGACCTCCTGCTCGTTCTTCAGC | |
| PK137_F | GCGAGGGTGAGGGCGCTGAGGAGTACTAAAACCACGAAATGTGCGAC | Replacing sequences upstream of stop codon in aPKC TR motif with *αTub84B* in constructs with UAA-A SCC |
| PK123_R | CGTCACCGGAGTCCATGCCGACCTCCTGCTCGTTCTTCAGC | |
| PK120_F | TGGGAGCGTCATTGGTGGGCGGGGGGCTCCGGCGC | Replacing sequences downstream of stop codon in *aPKC* TR motif with αTub84B in constructs with UGA-C SCC |
| PK121_R | TCGAGCGTTGAAGTGGCGCGACGCTCAGACGCAATCCTCCAGAGACATC | |
| PK120_F | TGGGAGCGTCATTGGTGGGCGGGGGGCTCCGGCGC | Replacing sequences downstream of stop codon in *aPKC* TR motif with αTub84B in constructs with UAA-A SCC |
| PK136_R | TCGAGCGTTGAAGTGGCGCGACGTTTAGACGCAATCCTCCAGAGACATC | |
| PK199_F | CGCTGAGGAGTACTTCCACCACGAAATGTGCG | UGA to UUC mutation in 5'-*αTub84B-aPKC*-3' TR constructs |
| PK200_R | CGCACATTTCGTGGTGGAAGTACTCCTCAGCG | |
| PK201_F | GGATTGCGTCTTCGCGTCGCGCCAC | UGA to UUC mutation in 5'-*aPKC- αTub84B*-3' TR constructs |
| PK202_R | GTGGCGCGACGCGAAGACGCAATCC | |
| PK244_F | GGCTCCGGCGCCACCAAC | Forward primer for truncations in *aPKC* TR motif |
| PK245_R | GTGTCAGACGCAATCCTCCAGAGAC | Truncation of *aPKC* +6 TR motif with UGAC SCC |
| PK246_R | GTGGAAGACGCAATCCTCCAGAGAC | Truncation of *aPKC* +6 TR motif with UUC SCC |
| PK247_R | GTTTTAGACGCAATCCTCCAGAGACATCAG | Truncation of *aPKC* +6 TR motif with UAAA SCC |
| PK248_R | GTGGTGTCAGACGCAATCCTCCAG | Truncation of *aPKC* +9 TR motif with UGAC SCC |
| PK249_R | GTGGTGGAAGACGCAATCCTCCAG | Truncation of *aPKC* +9 TR motif with UUC SCC |
| PK250_R | GTGGTTTTAGACGCAATCCTCCAGAGAC | Truncation of *aPKC* +9 TR motif with UAAA SCC |

## 6.2 List of abbreviations

| | |
|---|---|
| A site | Aminoacyl site |
| A2RE | hnRNP A2/B1 responsive element |
| ATP | Adenosine Triphosphate |
| aa-tRNA | Aminoacyl-tRNA |
| ABCE1 | ATP-binding cassette sub-family E member 1 |
| bZip | Basic Leucine Zipeer |
| CDY | Cyo-DfD-YFP |
| CNS | Central nervous system |
| Cre | Cre recombinase |
| CRISPR | Clustered regularly interspaced short palindromic repeats |
| CySC | Cyst stem cells |
| dNTP | Deoxyribonucleotide triphosphate |
| E site | Exit site |
| EF/eEF | Elongation factor/ eukaryotic elongation factor |
| Fluc | Firefly luciferase |
| fMet | Formyl-methionine |
| GABA | Gamma-amino butyric acid |
| GB | Gonialblast |
| gDNA | Genomic DNA |
| GMPPNP | 5'-guanylyl imidodiphosphate |
| GSC | Germline stem cell |
| gRNA | Guide RNA |
| GTP | Guanosine Triphosphate |
| HA | Homology arms |
| hnRNP | Heteronuclear ribonucleoprotein |
| IC | Initiation complex |
| IF/eIF | Initiation factor/ eukaryotic initiation factor |
| LB | Luria-Bertani |
| LSU | Large subunit |
| MDa | Mega Dalton |
| mRNA | Messenger RNA |
| MuLV | Murine Leukemia Virus |
| N state | Non-rotated state |
| Nc-tRNA | Near cognate tRNA |

| | |
|---|---|
| NLS | Nuclear Localization signal |
| NMD | Nonsense-mediated decay |
| NTP | Nucleoside triphosphate |
| ORF | Open reading frame |
| P site | Peptidyl site |
| PABP | Poly(A) binding protein |
| PAM | Protospacer adjacent motif |
| PCR | Polymerase chain reaction |
| PDB | Protein data bank |
| PGC | Primordial germline cells |
| PIC | Preinitiation complex |
| piRNA | Piwi-interacting RNA |
| PNS | Peripheral nervous system |
| postTC | Posttermination complex |
| preTC | Pretermination complex |
| PRF | Programmed frameshifting |
| PTC | Peptidyl transferase center |
| PTS | Peroxisomal targeting signal |
| R state | Rotated state |
| RBP | RNA binding protein |
| RF/eRF | Release factor/ eukaryotic release factor |
| Rluc | Renilla luciferase |
| RRF | Ribosome release factor |
| rRNA | Ribosomal RNA |
| RT | Room temperature |
| RT-qPCR | Real time quantitative PCR |
| S | Svedberg Unit |
| S2 | Schneider 2 |
| SCC | Stop codon context |
| Sco | Scutoid |
| SGP | Somatic gonadal precursor |
| smFRET | Single molecule fluorescence resonance energy transfer |
| SSU | Small subunit |
| TC | Ternary complex |
| TfR | Template for recombination |

| | |
|---|---|
| TMV | *Tobacco mosaic virus* |
| tRNA | Transfer RNA |
| UTR | Untranslated region |
| VEGFA | Vascular endothelial growth factor A |
| VNC | Ventral nerve cord |

## 6.3 List of figures

## 6.4 List of tables

# ACKNOWLEDGMENTS

I am extremely grateful to my supervisor, Prof. Dr. Marina V. Rodnina for giving me the opportunity to pursue my doctoral studies in her department. I thank Marina for her guidance and support and most importantly for giving me the freedom and encouragement to venture into different areas of scientific research that has really helped me hone my expertise and grow as a scientist. I would like to thank my thesis committee members Prof. Dr. Halyna Shcherbata and Prof. Dr. Heike Krebber for their invaluable scientific input and fruitful collaborations. I would also like to thank members of my extended examination committee Prof. Dr. Wolfgang Wintermeyer, Prof. Dr. Herbert Jäckle and Dr. Alex Faesen.

I am exceptionally thankful to Dr. Cristina Maracci for being my mentor throughout the years. Cris has supervised me from my lab rotation days and I truly thank her for her wonderful guidance, intellectual support and constant motivation during my time as a master and a PhD student. Be it brainstorming ideas, setting up experiments, time management or scientific writing, I have truly learned a lot from her. A special thanks goes to Dr. Travis Carney from the Shcherbata lab for his insightful discussions and for teaching me everything fly related. I really appreciate Travis for always being willing to talk about flies, helping me set up experiments and providing me innumerable fly food vials at any given time.

I would like to thank everybody involved in the bacterial release project, Sarah, Tamara, Wolf, Heena, Frank, Ingo and Vedran. I started my work in this lab with the release project, which was challenging at parts but I am happy I was able to learn a lot from it. I am also glad I was a part of the Apidaecin project lead by Prof. Dr. Alexander Mankin and his team from University of Illinois. I was lucky to have worked on the yeast translation project with Namit, Sandra and Theo. I would like to thank Christian from the Krebber lab for the brief yet productive collaboration on the yeast termination project. I want to extend my appreciation to all my colleagues and friends from the Rodnina lab with whom I've shared wonderful times, both inside and outside the lab. It has been a pleasure working in such a friendly and scientific atmosphere where there is easy exchange of ideas and methodologies. I thank Sung Hui, Jakob and Abirami for their contribution in my project during their lab rotations. A special mention goes out to the technical staff who really deserve the appreciation for their hard work and behind-the-scenes contribution towards all the scientific discoveries.

Besides my lab, I would like to thank Halyna and her team for their expert technical as well as intellectual support and for being wonderful collaborators. I am grateful to Omer and Jasmin for their advice and expertise during the initial years of my PhD. I thank Sharif for providing me the S2 cell lines when needed.

It was fantastic being a part of the IMPRS Molecular Biology program. I would like to remember all my MolBio and Neuro classmates with whom I have shared memorable experiences during my Masters. I truly thank Steffen and Kerstin for making my transition to Germany smooth and easy and for taking care of us students at every step of our way.

On a personal note, a very special gratitude goes out to Shruti who has been a very dear friend and a confidante. Talking to her has always made me feel a part of home, away from home. I want to thank my family, my mom, dad and my sister who have always been supportive of me. Finally, I would like to thank mero baba for making the past few years of my life beautiful. I have truly lived the happiest moments of my life with you and I will always cherish our time together.