Open Journal of Animal Sciences, 2015, 5, 418-428
Published Online October 2015 in SciRes. http://www.scirp.org/journal/ojas
http://dx.doi.org/10.4236/ojas.2015.54044

Regulation of Expression for the RNP-4F Splicing Assembly Factor in the Fruit-Fly *Drosophila melanogaster*

Sushmita Ghosh, Shelby E. Thomas, Lindsey M. Abraham, Jack C. Vaughn*

Department of Biology, Cell Molecular and Structural Biology Program, Miami University, Oxford, USA
Email: *vaughnjc@MiamiOH.edu

Received 9 July 2015; accepted 6 October 2015; published 9 October 2015

Copyright © 2015 by authors and Scientific Research Publishing Inc.
This work is licensed under the Creative Commons Attribution International License (CC BY).
http://creativecommons.org/licenses/by/4.0/

Open Access

Abstract

Intron splicing in eukaryotic organisms requires the interactions of five snRNAs and numerous different proteins in the spliceosome. Although the molecular mechanism behind splicing has been well studied, relatively little is known about regulation of expression for these splicing factor proteins. One of these proteins is the evolutionarily-conserved *Drosophila* RNP-4F splicing assembly factor. This protein is transcribed from a single gene into two developmentally regulated mRNAs that differ in their 5'-UTR structure. In the longer isoform, known to be abundant in the developing fly central nervous system, a conserved retained intron which folds into a stem-loop has been implicated in expression control of the mRNA. Here, we describe construction and utilization of several new *rnp-4f* gene expression study vectors using a GFP reporter in the ΦC31 system. The results confirm our previous observation that presence of the regulatory stem-loop enhances RNP-4F protein expression. However, in that study, the enhancement factor protein was not identified. We show here that overexpression of the RNP-4F transgene compared to the control results in additional translation, as indicated by the GFP reporter in the fluorescent images. These results are interpreted to show that RNP-4F protein acts back on its own mRNA 5'-UTR regulatory region *via* a feedback pathway to enhance protein synthesis in the developing fly central nervous system. A model is proposed to explain the molecular mechanism behind *rnp-4f* gene expression control.

Keywords

***rnp-4f* Gene, Gene Expression Control, ΦC31 Transgenic Vectors, UAS-GAL4 System, Fluorescence Microscopy**

*Corresponding author.

How to cite this paper: Ghosh, S., Thomas, S.E., Abraham, L.M. and Vaughn, J.C. (2015) Regulation of Expression for the RNP-4F Splicing Assembly Factor in the Fruit-Fly *Drosophila melanogaster*. *Open Journal of Animal Sciences*, **5**, 418-428.
http://dx.doi.org/10.4236/ojas.2015.54044

1. Introduction

A major long-term objective of our research is to understand evolutionarily-conserved cellular, developmental, molecular and genetic mechanisms behind regulation of genes which encode intron splicing assembly factor proteins, a topic about which relatively little is known. The model system we are currently using to address our research questions [1]-[9] is the single-copy nuclear *rnp*-4*f* gene. This gene is located near the distal end of the X-chromosome in the fruit-fly *Drosophila melanogaster* [10], an important model genetic organism. It has been shown that *Drosophila* RNP-4F is an ortholog of human p110/SART3 and *Saccharomyces cerevisiae* Prp24, which function as snRNP recycling factors playing direct roles in assembly of U4 and U6 into di-snRNPs required for subsequent pre-mRNA splicing [11] [12]. The 5'-UTR and 3'-UTR of mRNAs have been shown to play a variety of *cis*-regulatory roles in control of eukaryotic gene expression, including translational modulation, and most translational control elements are located in these regions [13]-[15]. Stem-loop structures which function as binding sites for regulatory proteins have been shown to play a key role in 5'-UTR functions [16] [17], and in 3'-UTR stability [18]. To understand the genetic regulation of a gene such as *Drosophila rnp*-4*f*, it is important to elucidate the functional significance of both *cis*- and *trans*-acting pre-mRNA and mRNA protein binding factors.

Two major *rnp*-4*f* mRNA isoform classes have been described, arising due to alternative splicing in the 5'-UTR [4]. We have designated these two isoform classes as "long" (unspliced at intron 0 in the 5'-UTR) and "short" (alternatively spliced in the 5'-UTR by utilizing a secondary 3'-splice junction located in downstream exon 2), and their relative abundance is developmentally regulated during embryogenesis [5]. Quantification during development shows that both isoform classes are abundant at the 0 - 2 h embryo stage and are probably maternally inherited, followed by rapid declines in mRNA levels at about the time of the mid-blastula transition stage. The longer isoform then increases in abundance to peak at 8 - 12 h, but declines by the first larval instar stage. In contrast, the shorter isoform is present at only very low levels following the mid-blastula transition and within embryonic stages.

The two major *rnp*-4*f* mRNA isoforms are identical in their coding region, but a highly conserved primary and secondary structure has been observed in the 5'-UTR of the unspliced variant [4]. These observations have suggested that the key to understand the molecular basis for regulation of the alternative splicing decision, and ultimately the role of the 5'-UTR in *rnp*-4*f* gene regulation, may reside in this conserved region. We have shown that the entire 177-nt region which is removed by alternative splicing to produce the short isoform can potentially fold into a single long, stable stem-loop in many *Drosophila* species [6]. This interaction occurs by base-pairing between nucleotides of intron 0 and those in part of downstream exon 2, forming a structure closely resembling an intron Editing site Complementary Sequence or ECS [19]. We earlier proposed that the functional significance behind intron retention in this system may lie in a negative feedback control mechanism for regulation of RNP-4F protein abundance, where a long stable stem-loop could potentially interfere with small ribosome subunit scanning to reach the AUG start codon and initiate translation [4]. However, it was subsequently shown using transgenic flies and a GFP reporter system [8] that retention of the intron in the developing central nervous system (CNS) results in an unexpected increase in RNP-4F protein levels, completely the opposite of what had been expected. The cited work proposed that an unidentified protein binds to the 5'-UTR stem-loop in the regulatory region within the CNS, unwinding the long stable stem-loop and removing the barrier to small ribosome subunit scanning to reach the start codon AUG. In the work presented here, we utilize a newly developed set of transgenic fly lines, and show that overexpression of RNP-4F protein enhances translation for the corresponding gene *via* a feedback pathway, as indicated by a GFP reporter. These results, taken together with what is currently known about *rnp*-4*f* gene expression control, have resulted in a model to explain how this gene's expression is regulated during development of the *Drosophila* CNS.

2. Materials and Methods

2.1. Fly Stocks and Embryo Preparation

Drosophila red eye wild-type strain Oregon R, balancer stocks *CyO* and Tm_6Tb for chromosome II and III, respectively, with the markers *Spi* and Tm_2 for the corresponding chromosomes, and *elav*-*Gal*4 (stock #458) GAL4 promoter driver fly line which directs expression of genes in the embryonic nervous system, were obtained from the Bloomington, IN Stock Center. All the fly strains were cultured in standard media [3] at 25˚C in

a walk-in incubator which has 12 h light-dark cycle. Flies were grown in net cages on apple juice plates to collect staged embryos [20] which were stored at −80˚C for subsequent protein isolations. The recombinant plasmid pUAS-Stem WT [8] and pUASTattB vector, a gift of Johannes Bischof (Gen Bank no. EF362409), were used for making three new expression vectors: pUASTattB-SCR, pUASTattB-SL and pUASTattB-*rnp*-4*f*.

2.2. Protein Isolation from Fly Embryos

Total embryo protein was isolated from 100 mg of staged embryos, as described [6], and stored at −80˚C. Protein concentration was estimated by OD_{280} spectrophotometry (NanoDrop Technologies).

2.3. Preparation of Modified Wild-type, Control ("Scrambled") *rnp*-4*f* Stem-Loop, and Full-Length *rnp*-4*f* DNA Fragments

The wild-type 177-nt *rnp*-4*f* 5'-UTR containing an intron 0 and exon 2 stem-loop regulatory region attached to a downstream eGFP reporter gene DNA fragment was PCR-amplified from plasmid pUAS-Stem WT [8]. The forward primer used for the amplification was designed with an adapter which had a restriction site for *Eco*RI (shown in bold): 5'-TAT**GAATTC**G<u>G</u>AA<u>C</u>TGCTTGG AAAGTTTTT-3'. Two nucleotides (underlined) within the DNA fragment were mutated (U→G and G→C) to prevent *in vivo* splicing mediated by U1-snRNP association [21]. The reverse primer used for the amplification was designed with an adapter which had a restriction site for *Xba*I (shown in bold): 5'-TAT**TCTAGA**TGGACAAACCACAACTAGAAT-3'. The resulting DNA fragment with the modified wild-type stem-loop and downstream eGFP was designated SL-eGFP.

A control DNA fragment of random scrambled sequence but identical in length to that of the wild-type 177-nt *rnp*-4*f* 5'-UTR intron 0 and exon 2 stem-loop regulatory region was synthesized commercially (GenScript) and made available from the supplier in a pUC57 plasmid. The sequence was 5'- CCTTGTCGAACCTTTCAAAAA GGTTCTTTTTCACGGGTTATCCCTTTGGTTTGGACTTTGGTTTATTAAAAGGTGCTATTCGCGTTGTT GCGCTATAATCTAACCGGTCGGGACGCAAACGCCAAAAAAAAAACAAAAAAATAAAGGACATAA AAAAAACCCCCCCTTTTTAACGTCAAGGTATTA-3'. This sequence was designed using a computer-assisted program such that no stem within the RNA transcript longer than eight nucleotides is predicted [22]. The forward primer used to amplify this fragment was designed with an adapter which had a restriction site for *Eco*RI (shown in bold): 5'-TAT**GAATTC**CCTTGTCGAACCTTTCAAAAA-3'. The reverse primer used for amplification was designed with an adapter which had a restriction site for *Bam*HI (shown in bold): 5'-TAT**GGA TCC-**TAATACCTTGACGTTAAAAA-3'. The resulting DNA fragment with the scrambled sequence was designated SCR. pUAS-Neostinger vector [8] was then linearized at the multiple cloning site (MCS) by cutting with *Eco*RI and *Bam*HI. The SCR sequence with *Eco*RI and *Bam*HI restricted ends was inserted by ligation into pUAS-Neostinger to make the pUAS-SCR vector. The control DNA fragment of random scrambled sequence was PCR-amplified from the plasmid pUAS-SCR. The forward primer used to amplify this fragment was designed with an adapter which had a restriction site for *Eco*RI (shown in bold): 5'-TAT**GAATTC**CCTTGTC GAACCTTTCAAA-3'. The reverse primer used for amplification was designed with an adapter which had a restriction site for *Xba*I (shown in bold): 5'-TAT**TCTAGA**TGGACAAACCACAACTAGAAT-3'. The resulting DNA fragment with the scrambled sequence attached to a downstream eGFP gene was designated SCR-eGFP.

The full-length, sequence verified *rnp*-4*f* cDNA was obtained from the "gold collection" of the *Drosophila* Genomics Resource Center, and made available in a pOT2 plasmid vector (cDNA clone no. LD32008). The *rnp*-4*f* cDNA fragment was PCR-amplified using primers to specifically eliminate the 5'-UTR region which forms the conserved regulatory stem-loop in the *rnp*-4*f* mRNA. The forward primer used for amplification was designed with an adapter which had a restriction site for *Eco*RI (shown in bold): 5'-TAT**GAATTC**AAATT GCAGTTCCACGGAAA-3'. The reverse primer used for amplification was designed with an adapter which had a restriction site for *Xba*I (shown in bold): 5'-TAT**TCTAGA**TTCACTTGGTTCATCAAGAA-3'. The resulting DNA fragment with the slightly modified *rnp*-4*f* sequence was designated RNP-4F. All DNA fragments were sequenced and found to have no errors or mispositioned start codons (AUG) that could interfere with subsequent full-length translation *in vivo*.

2.4. Construction of pUASTattB-SCR, pUASTattB-SL and pUASTattB-*rnp*-4*f* Expression Vectors in ΦC31

The SL-eGFP DNA fragment amplified from the plasmid pUAS-Stem WT [8], the SCR-eGFP DNA fragment

amplified from the plasmid pUAS-SCR and the *rnp*-4*f* fragment amplified from the pOT2 plasmid vector using forward and reverse primers with adapters for restriction sites *Eco*RI and *Xba*I, were digested using *Eco*RI and *Xba*I and the correct DNA fragment gel purified. The pUASTattB vector was linearized at the multiple cloning site by cutting with *Eco*RI and *Xba*I and the three DNA fragments were separately ligated with the pUASTattB vector. The three resulting vectors were designated as pUASTattB-SL, pUASTattB-SCR and pUASTattB-*rnp*-4*f*. A quantity of each final transformed vector DNA, required for the fly embryo transfection process, was prepared (Wizard Midi-Prep kit, Promega).

2.5. Transgenic Fly Lines, Balancing, DNA Sequencing and Genetic Crosses

The fly transformation step using the ΦC31 transformation system was carried out using a commercial service (Rainbow Transgenic Flies), wherein the three different plasmid vectors were separately microinjected into very early stage *Drosophila* w^{1118} fly (white eyes) embryos for insertion of constructs into chromosomes. The φC31 transformation system in flies has the advantages of an endogenous recombinase and pre-determined landing site, which increases transformation efficiency and eliminates the need for mapping constructs in the transgenic fly genomes [23]. The pUASTattB-SCR and pUASTattB-SL constructs were inserted into chromosome 3 and the pUASTattB-*rnp*-4*f* construct was inserted into chromosome 2. The red eye progeny flies were crossed with balancer flies containing *CyO* and Tm_6Tb on chromosome II and III, respectively, with the markers *Spi* and Tm_2 for the corresponding chromosomes. Several fly crosses were set up to arrive at the three final transgenic fly lines: homozygous for SCR (chromosome 3), homozygous for SL (chromosome 3) and doubly homozygous for SL (chromosome 3) and for RNP-4F (chromosome 2).

The nucleotide sequences of all constructs inserted into the fly genome were verified by DNA sequencing. The forward primer 5'- GCCGGAGTATAAATAGAGGC-3' and the reverse primer 5'- CGTGACCTACATC GTCGAC-3' were used to PCR-amplify the transforming DNA constructs within the fly genomes. The PCR products were electrophoresed to purify the band and use it for sequencing constructs derived from the host genome.

The UAS-GAL4 system was used to express the RNP-4F transgene and eGFP reporter gene [24] [25] wherein the GAL4 protein from the driver fly binds to the UAS in the transgenic fly to drive expression of the downstream transgene. Virgin transgenic female flies were crossed with driver male flies, *elav-Gal*4 (stock #458, Bloomington, IN Stock Center).

2.6. Fluorescence Microscopy for *in Situ* Localization of GFP Reporter Gene Expression in Embryos

For all the three genetic crosses, virgin females (transgenic fly line) were crossed with driver fly line males *elav-Gal*4 (stock #458), a GAL4 promoter driver fly line which directs expression of genes in the embryonic nervous system, to collect embryos for *in situ* fluorescence microscopy. Embryos were then collected that were 0-16 h old. The sample preparation for microscopy was done using a modified protocol as described [26]. Embryos were treated with 50% bleach, heptane, 3.7% formaldehyde and methanol to remove the chorion and vitelline membrane for viewing and imaging. The images were taken using a Nikon Eclipse 80i fluorescence microscope fitted with a digital SPOT RT camera and processed using computer-assisted version 4.1 SPOT RT software. All images were captured using identical microscope and camera settings, as described [8].

3. Results and Discussion

3.1. Construction of pUASTattB-SCR, pUASTattB-SL and pUASTattB-*rnp*-4*f* Expression Vectors in ΦC31

The publicly available GFP expression vectors were not suitable for our experimental goals and thus we developed new transformation vectors with experimentally desirable characteristics. The three source plasmids that we used were pUAS-Stem WT, pUAS-Neostinger [8] and pUASTattB [23] (**Figure 1**). pUAS transformation vectors use the P-element transformation system which results in random integration of the DNA constructs into the host genome. This is an advantage in mutational analysis studies but poses considerable challenge in the case of precise functional analysis. The random integration makes it necessary to map the insertion site to eliminate the possibility of false positive results caused by differential positioning of control and experimental DNA con-

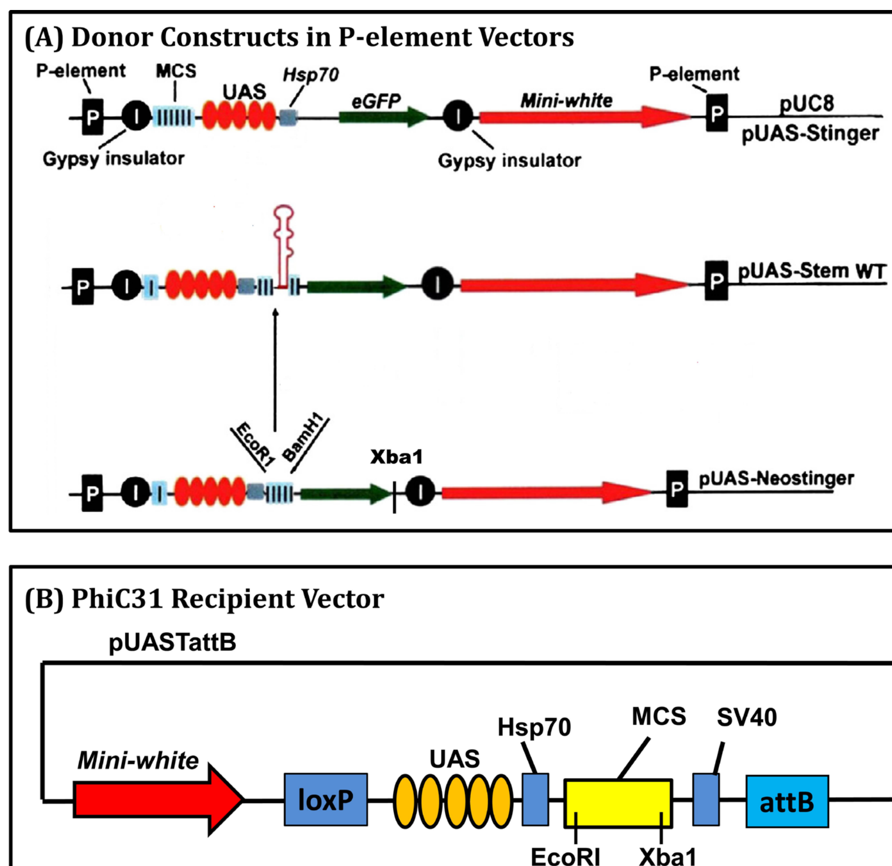

Figure 1. Source plasmids. (A) pUAS-Neostinger and pUAS-Stem WT were two of the source plasmids used for construction of the transforming vectors [8]. A commercially synthesized scrambled sequence (random sequence but same length as the 177-nt stem-loop in wild-type) containing *Eco*RI and *Bam*HI adapters was PCR-amplified, restricted and then ligated into the pUAS-Neostinger plasmid within the multiple cloning site (MCS). (B) pUASTattB was the third source plasmid used for constructing the transforming vectors [23]. This plasmid is provided with a *mini-white* gene for selection of red eye transformed flies, UAS, Hsp70 promoter and a SV40 region to enhance genomic integration of transforming DNA constructs.

structs. In an attempt to bypass these problems we opted to use the ΦC31 transforming vector, pUASTattB, which because of its endogenous integrase and precise integration makes the transformation more efficient and optimal for *in vivo* functional studies [23]. Two of the plasmids were constructed with a downstream reporter eGFP which is an enhanced GFP coding sequence with a nuclear localization signal to optimize the GFP expression and visualization process. The transformed flies could be easily selected based on their red eye phenotype because of the *mini-white* gene present in the transforming vector, pUASTattB. The construction process resulted in three plasmid vectors: pUASTattB-SCR, pUASTattB-SL and pUASTattB-*rnp*-4*f* (**Figure 2**).

3.2. *Drosophila* Fly Lines Successfully Transformed Using the ΦC31 Transformation Vector

The three ΦC31-based plasmid vectors were microinjected into very early stage *Drosophila* w^{1118} (white eye) fly embryos. The successfully transformed flies were crossed with balancer fly lines to make the three desired fly lines: homozygous for SCR (chromosome 3), homozygous for SL (chromosome 3) and doubly homozygous for SL (chromosome 3) and for RNP-4F (chromosome 2). The successful integration of the DNA constructs within the host genome was verified by PCR analysis. PCR analysis followed by gel electrophoresis showed that flies homozygous for SCR or SL produce a single band sized at about 1300-bp and flies homozygous for *rnp*-4*f* produce a single band sized at about 3100-bp (**Figure 3**), as expected. PCR analysis of genomic DNA from flies doubly homozygous for SL and *rnp*-4*f* produces two bands sized about 1300-bp and 3100-bp. The DNA constructs within the fly genome were verified by DNA sequencing.

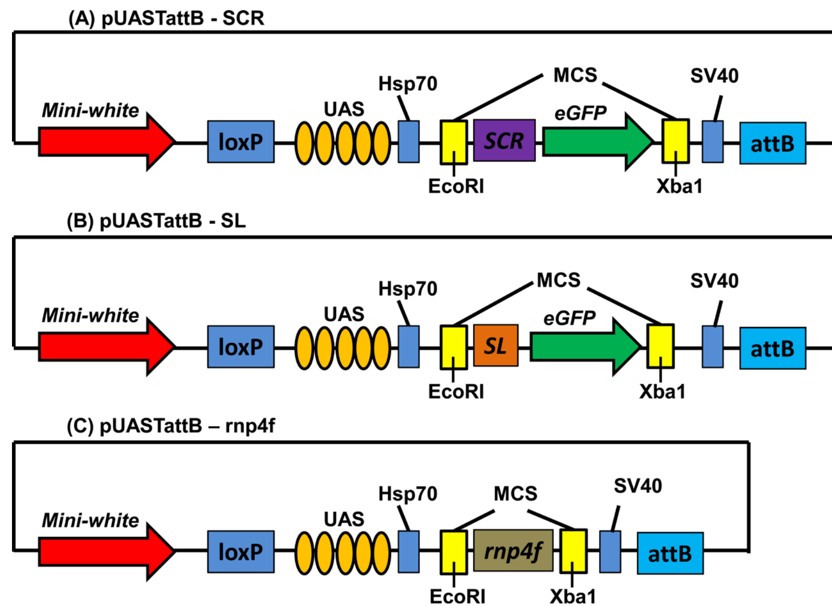

Figure 2. Vector constructs used for making the transgenic flies. (A) The scrambled control sequence (SCR) inserted within the MCS of pUAS-Neostinger and the downstream eGFP was PCR-amplified using primers with *Eco*RI and *Xba*I adapters, restricted and then ligated with pUASTattB to construct a vector with the following organization as verified by DNA sequencing: *mini-white*-loxP-UAS-Hsp70-MCS with DNA insert-SV40-attB and was designated pUASTattB-SCR. (B) The modified wild-type stem-loop sequence (SL) inserted within the MCS of pUAS-Stem WT and the downstream eGFP was PCR-amplified using primers with *Eco*RI and *Xba*I adapters, restricted and then ligated with pUASTattB to construct a vector with the following organization as verified by DNA sequencing: *mini-white*-loxP-UAS-Hsp70-MCS with DNA insert-SV40-attB and was designated pUASTattB-SL. (C) The *rnp-4f* cDNA sequence inserted within the MCS of pOT2 vector was PCR-amplified using primers with *Eco*RI and *Xba*I adapters, restricted and then ligated with pUASTattB to construct a vector with the following organization as verified by DNA sequencing: *mini-white*-loxP-UAS-Hsp70-MCS with DNA insert-SV40-attB and was designated pUASTattB-*rnp4f*.

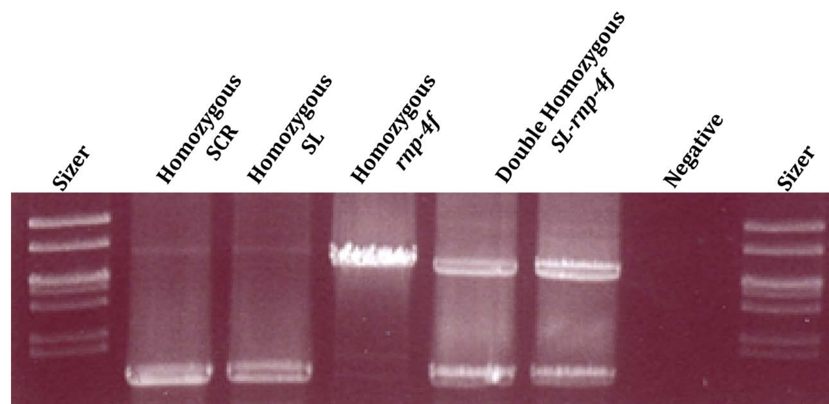

Figure 3. PCR analysis of DNA constructs amplified from the genomes of transgenic flies**.** The genomic DNAs of transgenic flies were used to PCR-amplify the transforming DNA constructs using primers against a region of the Hsp70 and attR (common to all constructs). PCR products were then electrophoresed. Flies homozygous for SCR or SL gave a single band (~1300-bp), flies homozygous for *rnp-4f* gave a single band (~3100-bp) and flies doubly homozygous for SL and *rnp-4f* gave two bands (~1300-bp and ~3100-bp), as expected.

3.3. Fluorescence Microscopy Study of eGFP Reporter Gene Expression Pattern in Transgenic Fly Embryos

During *Drosophila* development, the CNS is well formed in 14 - 16 h staged embryos [20]. The *rnp-4f* mRNA is most abundant in the developing CNS in fly embryos and RNAi knockdown of the long *rnp-4f* isoform leads to CNS deformities in fly embryos [5]. Hence, to study the role of excess RNP-4F on reporter gene expression in

the developing CNS, 14 - 16 h staged embryos were selected for fluorescence microscopy analysis of eGFP reporter expression profiles (**Figure 4**). A dull green autofluorescence is always observed in the developing gut [27] and the embryonic polyploid salivary glands also display bright green fluorescence. These distracting results are not to be confused with those arising from GFP reporter fluorescence. In control embryos from transgenic fly line *UAS-SCR-GFP* crossed with *elav-Gal*4 driver **(A-D)**, faint GFP green fluorescence is observed in the developing CNS. In embryos from transgenic fly line *UAS-SL-GFP* crossed with *elav-Gal*4 driver **(E-H)**, GFP green fluorescence observed in the developing CNS is enhanced. GFP expression in progeny embryos from transgenic fly lines *UAS-RNP/UAS-WT-GFP* crossed with *elav-Gal*4 **(I-L)** is further enhanced. These results clearly show that presence of the *rnp*-4*f* 5'-UTR stem-loop enhances downstream reporter gene expression, confirming work previously reported [8]. However, in that study, the enhancement factor protein was not identified. Our results in the present study show that overexpression of the RNP-4F transgene leads to an additional level of translation, as indicated by the GFP reporter in the fluorescent images. These results are interpreted to show that RNP-4F protein functions as an activator to upregulate gene expression by binding to the stem-loop in the 5'-UTR of its own long mRNA isoform in the developing fly CNS *via* a feedback pathway.

Our intention in this study was to quantify the qualitative reporter gene expression levels observed in the fluorescent images using Western technology, as in our related previous work [8]. However, repeated attempts to grow adequate numbers of embryos to enable preparation of sufficient protein amounts from some transgenic fly lines were unsuccessful. This was due to a tendency for the RNP-4F overexpression stocks to be less vigorous and in general the flies were smaller than in the wild-type stock. The males grew better than the females, perhaps owing to the location of the *rnp*-4*f* gene on the X-chromosome, which is only present in single copy in males.

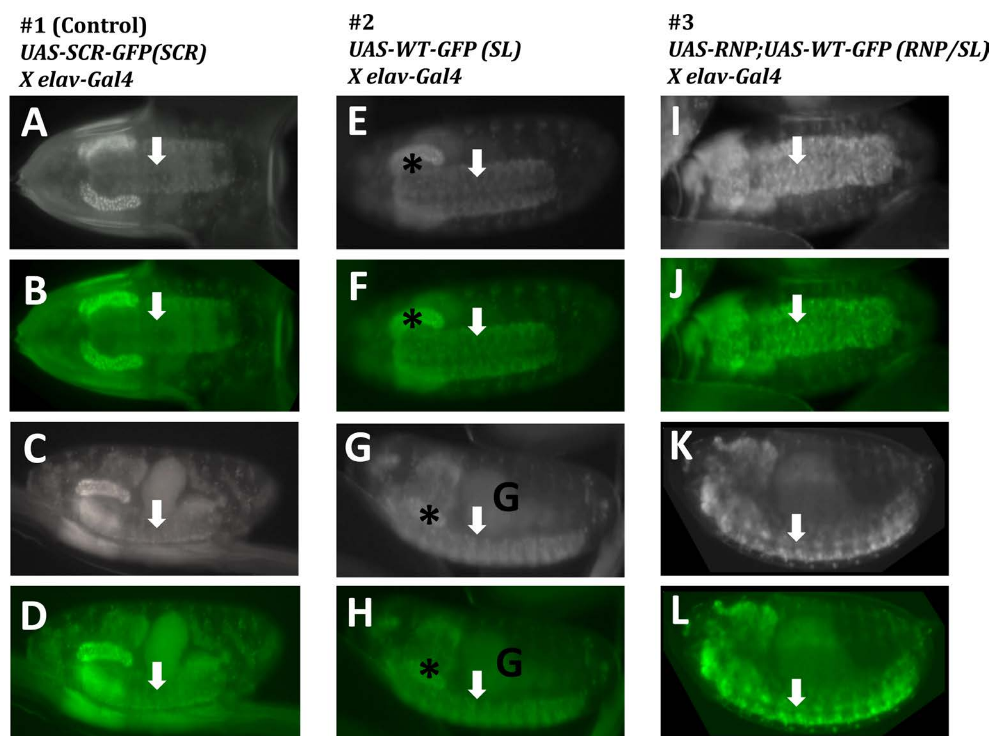

Figure 4. GFP reporter gene expression in *Drosophila* embryos**.** Fluorescence micrographs of GFP reporter gene protein expression in stage 14 - 16 embryos showed that in control embryos (column #1, (A)-(D)), UAS-SCR-GFP crossed with *elav-Gal4,* there is weak fluorescence in the gut and CNS, in the presence of the stem-loop upstream of the GFP reporter. In column #2 ((E)-(H)), UAS-WT-GFP crossed with *elav-Gal*4, the GFP fluorescence in the developing CNS is enhanced compared to the control. In column #3 ((I)-(L)), in the presence of the stem-loop upstream of the GFP reporter and excess RNP-4F in embryos, GFP fluorescence in the developing CNS is further enhanced. Individual images are printed in color and also black/white, for greater clarity in visualization. Figures in the two upper panels of each column show dorsal views and those in the two lower panels of each column show lateral views with ventral side facing down. Anterior is to the left in all panels. Arrows show ventral nerve cord and G is gut. Asterisks mark salivary glands.

3.4. A Model for *rnp*-4*f* Gene Expression Control

The *Drosophila rnp*-4*f* gene codes for two mRNA isoforms which have identical coding potential but differ by a 177-nt sequence that forms an evolutionarily-conserved stem-loop caused by alternative splicing in the 5'-UTR region within intron 0 and exon 2 [5] [6]. Northern blots and RT-PCR studies suggest that the relative abundance of the two isoforms is controlled developmentally [4] [5]. Elevated levels of the longer *rnp*-4*f* isoform correspond to stages of fly development when the central nervous system develops [4]. The 5'-UTR secondary structure is evolutionary-conserved across several *Drosophila* species [6], and this together with the developmental regulation of alternative transcript levels suggests functional significance of the stem-loop structure. *In situ* localization studies using a DIG-labeled RNA probe have shown that in early stages of fly development, the longer *rnp*-4*f* isoform is present in the mesodermal and ectodermal cells which are precursors of the ventral nerve cord. In later stages of embryo development, the long isoform is located primarily in the brain and the dorsal roof of the ventral nerve cord [5]. These results, when compared to the localization of the RNP-4F protein and dADAR protein, have clearly shown that the long mRNA isoform and the two proteins colocalize in developing fly embryos [5]. RNA electrophoretic mobility shift assay (REMSA), using *in vitro*-transcribed RNA (*rnp*-4*f* 5'-UTR 177-nt stem-loop) and whole embryo protein extract from wild-type embryos and *dADAR* mutant embryo protein combined with qPCR analysis, suggest that dADAR is one of two proteins that bind to the stem-loop to affect levels of long *rnp*-4*f* mRNA by interfering with splicing [7]. RNAi studies have shown that it is the short isoform of dADAR which binds to the *rnp*-4*f* stem-loop to regulate alternative splicing [9]. However, the identity of the other protein detected by REMSA which binds the *rnp*-4*f* 5'-UTR regulatory region, remained unknown.

A structural study has revealed that there is a conserved sequence in U6-snRNA to which yeast Prp24, an RNP-4F ortholog, preferably binds in its role as a splicing assembly factor [11]. This binding to U6-snRNA, initially having a long stem-loop secondary structure, has been shown to melt the structure, functioning as a chaperone to enable nucleotide base-pairing with U4-snRNA during spliceosome assembly [28]. A 12-nucleotide long sequence within the 5'-UTR 177-nt region of the *rnp*-4*f* mRNA shares considerably high sequence similarity with the binding site of RNP-4F on U6-snRNA [29]. An additional similarity is that in both cases the consensus sequence lies within a long stable stem-loop secondary structure. The findings from the study reported here show that presence of the *rnp*-4*f* 5'-UTR stem-loop upstream of the reporter eGFP enhances gene expression, which is further enhanced in the presence of excess RNP-4F. Taken together, these observations lead us to propose a model to explain the regulation of expression of the RNP-4F splicing assembly factor during *Drosophila* development (**Figure 5**). According to this model, in the developing CNS, within the nucleus, the short dADAR protein isoform binds to the stem-loop in the 5'-UTR of the long *rnp*-4*f* mRNA. This binding prevents removal of the stem-loop by splicing, perhaps using steric hindrance to interfere with snRNP bonding. The long *rnp*-4*f* isoform is then transported into the cytoplasm. In the presence of RNP-4F protein, a nuclear protein which however has been detected in the cytoplasm of *Drosophila* cells [5], protein binds to the *rnp*-4*f* stem-loop secondary structure in its own mRNA, acting as an enhancement factor and increasing RNP-4F expression *via* a feedback pathway.

4. Conclusions

Controlled mRNA translation and mRNA degradation both play important roles in regulating gene expression [30]. Control sequences present in both the 5'- and 3'-UTR of mRNAs serve as *cis* elements with important roles in regulating the translation process [13]. Splicing enhancer and silencer sequences have also been identified within introns and exons [31]-[33]. For example, secondary structures within the 5'-UTR of mRNAs regulate translation as in the human and murine *p*53 mRNA with a stable stem-loop structure in the 5'-UTR [34] [35]. In another example, *Drosophila* SXL protein binds to the 5'-UTR of *msl*2 mRNA and through interactions with the STAR protein HOW represses *msl*2 expression by nuclear retention of the *msl*2 mRNA [36]. About 4000 of the approximately 14,000 protein coding genes in *Drosophila* contain introns of unknown functional significance in their 5'-UTR (J. Carlson, personal communication).

The *Drosophila rnp*-4*f* gene codes for two mRNA isoforms which differ by a 177-nt sequence caused by alternative splicing in the 5'-UTR within intron 0 and exon 2. The mRNA isoforms are "long" (unspliced) and "short" (alternatively spliced), making up 15% and 85% of total *rnp*-4*f* transcripts in adults, respectively [4]. The computer-assisted M-fold program shows that the 5'-UTR 177 nucleotides in the unspliced *rnp*-4*f* mRNA

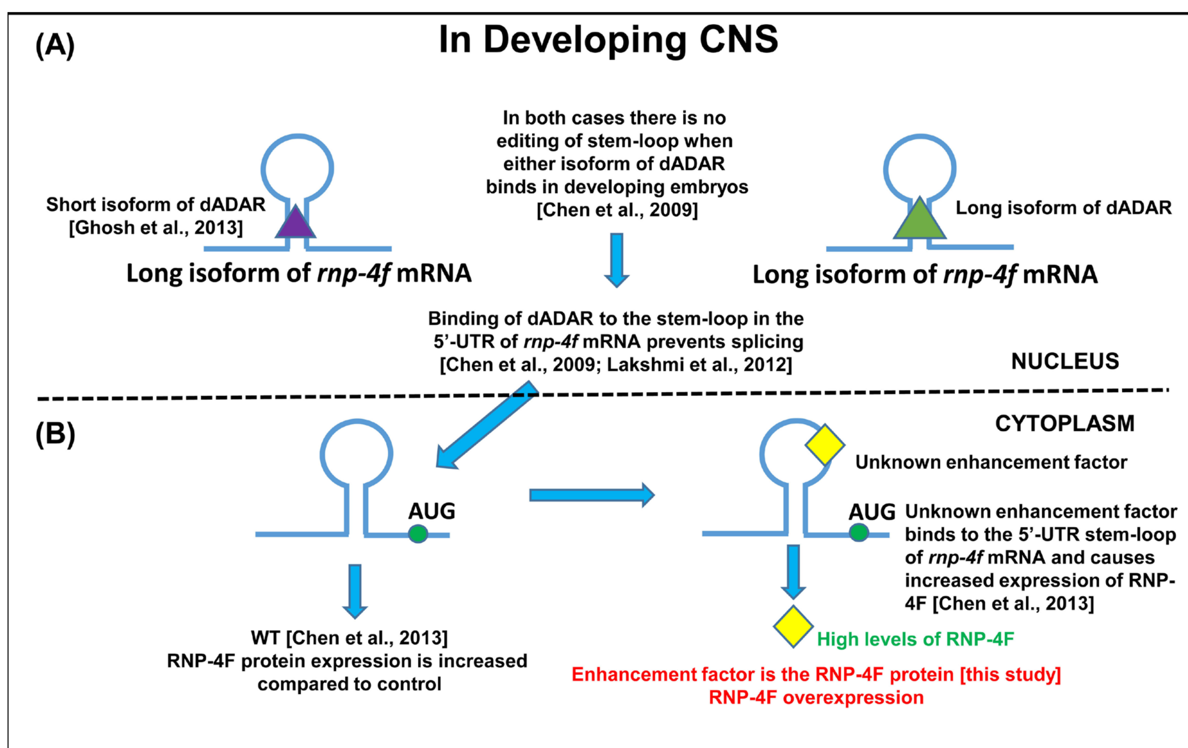

Figure 5. Model for *rnp-4f* gene expression regulation**.** (A) In the developing central nervous system (CNS), a dADAR isoform binds to the long *rnp-4f* stem-loop in the 5'-UTR regulatory region [6] [7] [9]. This interferes with splicing of the intron, perhaps by using a steric hindrance mechanism, after which the longer *rnp-4f* mRNA isoform exits the nucleus. (B) In the cytoplasm of developing CNS cells, the retained stem-loop recruits RNP-4F protein, which unwinds the structure and results in efficient small ribosomal subunit scanning to locate the AUG start codon, enhancing RNP-4F translation *via* a feedback pathway.

forms an evolutionarily-conserved stem-loop secondary structure due to complementary base-pairing between intron 0 and exon 2 [4] [6]. REMSA studies suggest that more than one protein (*trans*-acting factor) binds to the 5'-UTR stem-loop structure in *rnp*-4*f* mRNA [7]. *In vivo* localization studies in *Drosophila* embryos have shown that the long *rnp*-4*f* mRNA and the RNP-4F protein are both present in the developing CNS. RNAi mediated significant reductions in the level of long *rnp*-4*f* mRNA in the CNS and caused severe deformities of the CNS, which suggests a role of this isoform in CNS development [6]. Hence, the objective of this research was to understand the significance of the preferential expression of the long isoform in the CNS. In this study, using transgenic flies and tissue-directed transgene expression tools, it has been observed that presence of the *rnp*-4*f* 5'-UTR stem-loop upstream of a reporter gene, enhances downstream gene expression. Significantly, the CNS gene reporter expression is further enhanced in the CNS compared to the control in the presence of excess RNP-4F, which is most likely due to an autoregulatory mechanism of gene expression regulation of *rnp*-4*f* in the developing fly CNS. Taken together, these observations have led us to propose a working model to explain the molecular regulation of *Drosophila rnp*-4*f* during fly development (**Figure 5**).

Acknowledgements

We wish to thank Xiaoyun Deng from the Miami University Center for Bioinformatics and Functional Genomics for her help in running the ABI 3100 DNA sequencings. Matt Duley from the Miami University Electron Microscopy and Optical Microscopy Facility is thanked for helping with the embryo imaging studies. The Bloomington, IN *Drosophila* Stock Center and the *Drosophila* Genomics Resource Center provided valuable fly stocks and plasmids for this work. A preliminary report of the work presented here was given at the 56th Annual *Drosophila* Research Conference, Chicago, IL, in March 2015. This work was supported by National Institutes of Health (NIH) Grant 1-R15-GM093895-01 to J. Vaughn.

References

[1] Petschek, J.P., Mermer, M.J., Scheckelhoff, M.R., Simone, A.A. and Vaughn, J.C. (1996) RNA Editing in *Drosophila* 4*f-rnp* Gene Nuclear Transcripts by Multiple A-to-G Conversions. *Journal of Molecular Biology*, **259**, 885-890. http://dx.doi.org/10.1006/jmbi.1996.0365

[2] Petschek, J.P., Scheckelhoff, M.R., Mermer, M.J. and Vaughn, J.C. (1997) RNA Editing and Alternative Splicing Generate mRNA Transcript Diversity from the *Drosophila* 4*f-rnp* Locus. *Gene*, **204**, 267-276. http://dx.doi.org/10.1016/S0378-1119(97)00465-4

[3] Peters, N.T., Rohrbach, J.A., Zalewski, B.A., Byrkett, C.M. and Vaughn, J.C. (2003) RNA Editing and Regulation of *Drosophila* 4*f-rnp* Expression by *sas*-10 Antisense Readthrough mRNA Transcripts. *RNA*, **9**, 698-710. http://dx.doi.org/10.1261/rna.2120703

[4] Fetherson, R.A., Strock, S.B., White, K.N. and Vaughn, J.C. (2006) Alternative Pre-mRNA Splicing in *Drosophila* Spliceosomal Assembly Factor RNP-4F during Development. *Gene*, **371**, 234-245. http://dx.doi.org/10.1016/j.gene.2005.12.025

[5] Chen, J., Concel, V.J., Bhatla, S., Rajeshwaran, R., Smith, D.L.H., Varadarajan, M., Backscheider, K.L., Bockrath, R.A., Petschek, J.P. and Vaughn, J.C. (2007) Alternative Splicing of an *rnp*-4*f* mRNA Isoform Retaining an Evolutionarily-Conserved 5'-UTR Intronic Element Is Developmentally Regulated and Shown *via* RNAi to Be Essential for Normal Central Nervous System Development in *Drosophila melanogaster*. *Gene*, **399**, 91-104. http://dx.doi.org/10.1016/j.gene.2007.04.038

[6] Chen, J., Lakshmi, G.G., Hays, D.L., McDowell, K.M., Ma, E. and Vaughn, J.C. (2009) Spatial and Temporal Expression of *dADAR* mRNA and Protein Isoforms during Embryogenesis in *Drosophila melanogaster*. *Differentiation*, **78**, 312-320. http://dx.doi.org/10.1016/j.diff.2009.08.003

[7] Lakshmi, G.G., Ghosh, S., Jones, G.P., Parikh, R., Rawlins, B.A. and Vaughn, J.C. (2012) An RNA Electrophoretic Mobility Shift and Mutational Analysis of *rnp*-4*f* 5'-UTR Intron Splicing Regulatory Proteins in *Drosophila* Reveals a Novel New Role for a dADAR Protein Isoform. *Gene*, **511**, 161-168. http://dx.doi.org/10.1016/j.gene.2012.09.088

[8] Chen, J., Yang, J.T., Doctor, D.L., Rawlins, B.A., Shields, B.C. and Vaughn, J.C. (2013) 5'-UTR Mediated Translational Control of Splicing Assembly Factor RNP-4F Expression during Development of the *Drosophila* Central Nervous System. *Gene*, **528**, 154-162. http://dx.doi.org/10.1016/j.gene.2013.07.027

[9] Ghosh, S., Wang, Y., Cook, J.A., Chhiba, L. and Vaughn, J.C. (2013) A Molecular, Phylogenetic and Functional Study of the *dADAR* mRNA Truncated Isoform during *Drosophila* Embryonic Development Reveals an Editing-Independent Function. *Open Journal of Animal Sciences*, **3**, 20-30. http://dx.doi.org/10.4236/ojas.2013.34A2003

[10] Hess, K.A., Simone, A.A. and Petschek, J.P. (1996) Spatial and Temporal Expression of the 4*f-rnp* Gene in *Drosophila melanogaster*. *Differentiation*, **61**, 103-111. http://dx.doi.org/10.1046/j.1432-0436.1996.6120103.x

[11] Bell, M., Schreiner, S., Damianov, A., Reddy, R. and Bindereif, A. (2002) p110, a Novel Human U6-snRNP Protein and U4/U6-snRNP Recycling Factor. *The EMBO Journal*, **21**, 2724-2735. http://dx.doi.org/10.1093/emboj/21.11.2724

[12] Rader, S.D. and Guthrie, C. (2002) A Conserved Lsm-Interaction Motif in Prp24 Required for Efficient U4/U6 di-snRNP Formation. *RNA*, **8**, 1378-1392. http://dx.doi.org/10.1017/S1355838202020010

[13] Mignone, F., Gissi, C., Liuni, S. and Pesole, G. (2002) Untranslated Regions of mRNAs. *Genome Biology*, **3**, 0004.1-0004.10.

[14] Pickering, B.M. and Willis, A.E. (2005) The Implications of Structured 5'-Untranslated Regions on Translation and Disease. *Seminars in Cell and Developmental Biology*, **16**, 39-47. http://dx.doi.org/10.1016/j.semcdb.2004.11.006

[15] Chatterjee, S. and Pal, J.K. (2009) Role of 5'- and 3'-Untranslated Regions of mRNAs in Human Diseases. *Biology of the Cell*, **101**, 251-262. http://dx.doi.org/10.1042/BC20080104

[16] Klausner, R.D., Rouault, T.A. and Harford, J.B. (1993) Regulating the Fate of mRNA: The Control of Cellular Iron Metabolism. *Cell*, **72**, 19-28. http://dx.doi.org/10.1016/0092-8674(93)90046-S

[17] Gebauer, F., Grskovic, M. and Hentze, M.W. (2003) *Drosophila* Sex-Lethal Inhibits the Stable Association of the 40S Ribosomal Subunit with *msl*2 mRNA. *Molecular Cell*, **11**, 1397-1404. http://dx.doi.org/10.1016/S1097-2765(03)00176-X

[18] Marzluff, W.F., Wagner, E.J. and Duronio, R.J. (2008) Metabolism and Regulation of Canonical Histone mRNAs: Life without a Poly(A) Tail. *Nature Review Genetics*, **9**, 843-854. http://dx.doi.org/10.1038/nrg2438

[19] Higuchi, M., Single, F.N., Kohler, M., Sommer, B., Sprengel, R. and Seeburg, P.H. (1993) RNA Editing of AMPA Receptor Subunit GluR-B: A Base-Paired Intron-Exon Structure Determines Position and Efficiency. *Cell*, **75**, 1361-1370. http://dx.doi.org/10.1016/0092-8674(93)90622-W

[20] Campos-Ortega, J.A. and Hartenstein, V. (1997) The Embryonic Development of *Drosophila*. Springer-Verlag, Berlin. http://dx.doi.org/10.1007/978-3-662-22489-2

[21] Lewin, B. (1994) Genes V. Oxford University Press, New York.

[22] Zuker, M. (2003) Mfold Web Server for Nucleic Acid Folding and Hybridization Prediction. *Nucleic Acids Research*, **31**, 3406-3415. http://dx.doi.org/10.1093/nar/gkg595

[23] Bischof, J., Maeda, R.K., Hediger, M., Karch, F. and Basler, K. (2007) An Optimized Transgenesis System for *Drosophila* Using Germ-Line-Specific ϕC31 Integrases. *Proceedings of the National Academy of Sciences of the United States of America*, **104**, 3312-3317. http://dx.doi.org/10.1073/pnas.0611511104

[24] Brand, A.H. and Perrimon, N. (1993) Targeted Gene Expression as a Means of Altering Cell Fates and Generating Dominant Phenotypes. *Development*, **118**, 401-415.

[25] Duffy, J.B. (2002) GAL4 System in *Drosophila*: A Fly Geneticist's Swiss Army Knife. *Genesis*, **34**, 1-15. http://dx.doi.org/10.1002/gene.10150

[26] Edoff, K., Dods, J.S. and Brand, A.H. (2007) Detection of GFP during Nervous System Development in *Drosophila melanogaster*. *Methods in Molecular Biology*, **411**, 81-98. http://dx.doi.org/10.1007/978-1-59745-549-7_6

[27] Sun, B., Xu, P. and Salvaterra, P.M. (1999) Dynamic Visualization of Nervous System in Live *Drosophila*. *Proceedings of the National Academy of Sciences of the United States of America*, **96**, 10438-10443. http://dx.doi.org/10.1073/pnas.96.18.10438

[28] Jandrositz, A. and Guthrie, C. (1995) Evidence for a Prp24 Binding Site in U6 snRNA in a Putative Intermediate in the Annealing of U6 and U4 snRNA. *The EMBO Journal*, **14**, 820-832.

[29] Vaughn, J.C., Ghosh, S. and Chen, J. (2013) A Phylogenetic Study of *Drosophila* Splicing Assembly Chaperone RNP-4F Associated U4-/U6-snRNA Secondary Structure. *Open Journal of Animal Sciences*, **3**, 36-48. http://dx.doi.org/10.4236/ojas.2013.34A2005

[30] Semotok, J.L., Cooperstock, R.L., Pinder, B.D., Vari, H.K. and Lipshitz, H.D. (2005) Smaug Recruits the CCR4/POP2/NOT Deadenylase Complex to Trigger Maternal Transcript-Localization in the Early *Drosophila* Embryo. *Current Biology*, **15**, 284-294. http://dx.doi.org/10.1016/j.cub.2005.01.048

[31] Ladd, A.N. and Cooper, T.A. (2002) Finding Signals That Regulate Alternative Splicing in the Post-Genomic Era. *Genome Biology*, **3**, 0008.1-0008.16.

[32] Zheng, Z.M. (2004) Regulation of Alternative RNA Splicing by Exon Definition and Exon Sequences in Viral and Mammalian Gene Expression. *Journal of Biomedical Science*, **11**, 278-294. http://dx.doi.org/10.1007/BF02254432

[33] Wang, Z., Rolish, M.E., Yeo, G., Tung, V., Mawson, M. and Burge, C.B. (2004) Systematic Identification and Analysis of Exonic Splicing Silencers. *Cell*, **119**, 831-845. http://dx.doi.org/10.1016/j.cell.2004.11.010

[34] Mosner, J., Mummenbrauer, T., Bauer, C., Sczakiel, G., Grosse, F. and Deppert, W. (1995) Negative Feedback Regulation of Wild-Type p53 Biosynthesis. *The EMBO Journal*, **14**, 4442-4449. http://dx.doi.org/10.1007/bf02572252

[35] Takagi, M., Absalon, M.J., McLure, K.G. and Kastan, M.B. (2005) Regulation of p53 Translation and Induction after DNA Damage by Ribosomal Protein L26 and Nucleolin. *Cell*, **123**, 49-63. http://dx.doi.org/10.1016/j.cell.2005.07.034

[36] Graindorge, A., Carre, C. and Gebauer, F. (2013) Sex-Lethal Promotes Nuclear Retention of *msl2* mRNA via Interactions with the STAR Protein HOW. *Genes and Development*, **27**, 1421-1433. http://dx.doi.org/10.1101/gad.214999.113