The *Drosophila melanogaster* *rnp-4f* and *sas10* genes produce diverse converging mRNAs by differential 3’-end processing of *sas10* during specific developmental stages. The longer *sas10* isoforms overlap with *rnp-4f* mRNAs to form substrates that have been shown to be A-to-G edited during development, which indicates that double-stranded RNAs are formed, since the dADAR editase requires duplex molecules for its activity. Our lab has previously reported the simultaneous appearance during mid-to-late embryo stages of a long readthrough *sas10* isoform and the disappearance of both *rnp-4f* and *sas10* mRNAs, which suggests a posttranscriptional model for control of gene expression in *D. melanogaster*. In this Thesis, it is asked whether or not this model is generally applicable. A genomics approach shows that the two genes are closely adjacent and predicts converging, overlapping transcripts in all *Drosophila* species available for comparison, extending at least to the mosquito *Anopheles gambiae* with 250 million years divergence time from *Drosophila*. Developmental Northerns show two abundant *sas10* isoforms the longer one of which overlaps *rnp-4f* mRNA. Isoform-specific DIG-labeled RNA probes show that the longer *sas10* isoform co-localizes with *rnp-4f* transcripts and is broadly distributed in developing embryos, with heavy staining appearing in the ventral nerve cord in later stages. This pattern suggests a potential role for *sas10* in central nervous development. We are exploring the hypothesis that the simultaneous destruction of *rnp-4f* and *sas10* mRNAs during development may be via the RNAi pathway. Given that *sas10* and *rnp-4f* are predicted to produce converging transcripts leading to formation of double-stranded RNAs in diverse *Drosophila* species, an antisense mechanism of gene expression control for *rnp-4f* may be general. Future research utilizing both genetic and reverse genetic experimental approaches may clarify some of the questions raised during this research.
EVOLUTIONARY CONSERVATION OF A POTENTIAL ANTISENSE MECHANISM FOR SAS10 AND RNP-4F

GENE EXPRESSION CONTROL IN DROSOPHILA

A Thesis

Submitted to the
Faculty of Miami University
in partial fulfillment of
the requirements for the degree
Master of Science
Department of Zoology

by

Rajkumar Rajeshwaran
Miami University
Oxford, OH

2006

Advisor____________________________

Dr. Jack C. Vaughn

Reader____________________________

Dr. Susan M. Hoffman
# Table of Contents

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Detailed Table of Contents</td>
<td>iii</td>
</tr>
<tr>
<td>List of Figures</td>
<td>vii</td>
</tr>
<tr>
<td>List of Appendices</td>
<td>ix</td>
</tr>
<tr>
<td>Acknowledgements</td>
<td>x</td>
</tr>
<tr>
<td>Introduction</td>
<td>1</td>
</tr>
<tr>
<td>Materials and Methods</td>
<td>10</td>
</tr>
<tr>
<td>Results</td>
<td>17</td>
</tr>
<tr>
<td>Discussion</td>
<td>23</td>
</tr>
<tr>
<td>Future Research Directions</td>
<td>28</td>
</tr>
<tr>
<td>Appendices</td>
<td>60</td>
</tr>
<tr>
<td>Literature Cited</td>
<td>69</td>
</tr>
</tbody>
</table>
## Detailed Table of Contents

### Introduction

The *sas10* and *rnp-4f* genes in *Drosophila melanogaster* .................................................. 1

The *sas10* and *rnp-4f* isoforms in *D. melanogaster*
produce naturally occurring overlapping transcripts
and form sense/antisense pairs................................................................. 2

Antisense RNAs in posttranscriptional regulation of gene expression.......... 2

RNA editing in posttranscriptional gene expression control .................... 5

RNA silencing in posttranscriptional gene expression control ................. 6

*D. melanogaster* long *sas10* and *rnp-4f* transcript levels are inversely related.... 7

Role of the 3’-UTR in mRNA localization and alternative 3’-termini selection
in *D. melanogaster* .................................................................................. 7

### Materials and Methods

Fly stocks and propagation........................................................................... 10

Embryo preparation and storage for hybridizations.................................... 10

Total cell RNA isolation............................................................................. 10

Affinity chromatography for mRNA purification....................................... 10
Agarose gel electrophoresis of RNAs................................. 11
Polyacrylamide gel electrophoresis of RNAs.......................... 12
Northern blot analysis....................................................... 12
Polymerase chain reaction (PCR)........................................... 13
DNA fragment isolation by agarose gel electrophoresis.............. 14
DNA fragment cloning into plasmid vector/transformation/colony selection........ 14
Small scale plasmid isolation.............................................. 14
Bulk plasmid isolation...................................................... 15
DNA sequencing............................................................ 15

In vitro transcription...................................................... 15

In situ DIG hybridizations ............................................. 16
Microscopy........................................................................... 16

Comparative genomic analysis............................................. 16
Results

Quantification of \textit{sas10} transcript levels across development……………….. 17

\textit{rnp-4f} and \textit{sas10} genes are closely adjacent to each other and
transcribed in opposite direction in diverse \textit{Drosophila} species………………. 18

Attempted identification of different \textit{sas10} isoforms in
diverse \textit{Drosophila} species using Northern hybridization………………….. 20

Attempted evaluation of RNAi hypothesis for regulation of
\textit{sas10} mRNA abundance in \textit{D. melanogaster} during embryogenesis……….. 21

Localization of different \textit{sas10} mRNA isoforms across \textit{D. melanogaster}
development using DIG-labeled RNA probes………………………………….. 21

Discussion

Control of gene expression by antisense RNAs……………………………… 23

Role of 3’-UTR in gene expression, localization and mRNA stability………… 25

\textit{rnp-4f} gene expression is developmentally regulated by its antisense
\textit{sas10} transcript ………………………………………………………………… 26

Future Research Directions………………………………………………….. 28
Appendices

Appendix I: Dechorionation and devitellinization of *Drosophila* embryos................................................................. 60

Appendix II: Hybridization vs long targets (>100 -nt)

using random prime-labeled probe ........................................ 62

Appendix III: Hybridization vs short targets (~25-nt)

using random prime-labeled probe......................................... 63

Appendix IV: *In vitro* transcription........................................ 64

Appendix V: DIG-labeled RNA probe *in situ* hybridization of

*Drosophila* embryos............................................................... 66

Literature Cited....................................................................... 69
List of Figures

1. Orthologues of yeast splicing assembly factor PRP24 in diverse organisms .................................................. 30-31

2. Orientation of sas10 and rnp-4f genes in D. melanogaster ........ 32-33

3. RNAi silencing pathways ..................................................... 34-35

4. The inverse relationship between long readthrough sas10 isoform and rnp-4f ......................... 36-37

5. Expression profile of strictly maternal Drosophila genes ...... 38-39

6. Denaturing agarose gel electrophoresis of D. melanogaster poly (A\(^{+}\))-selected RNAs from different developmental stages… 40-41

7. Quantification of major sas10 transcript levels across development ............................................................... 42-43

8. Phylogenetic tree for dipteran insects utilized in this study....... 44-45

9A. Sequence alignment across the rnp-4f/sas10 intergenic spacer..... 46-47

9B. Sequence alignment overviews across the rnp-4f/sas10 intergenic Spacer................................................................. 48-49

10. Denaturing 1% agarose gel electrophoresis of adult total cell RNA from different Drosophila species....................... 50-51
11. Non-denaturing 1% agarose gel electrophoresis of adult total cell RNA from different *D. melanogaster* developmental stages ................................................................. 52-53

12. Denaturing polyacrylamide gel electrophoresis of *D. melanogaster* total cell RNAs from different developmental stages ................................................................. 54-55

13. *In vitro* transcription gels showing DIG-labeled RNA products... 56-57

14. *sas10* and *rnp-4f* mRNAs co-localize during development....... 58-59
List of Appendices

Appendix I: Dechorionation and devitellinization of *Drosophila* embryos

Appendix II: Hybridization vs long targets (>100 -nt)

Appendix III: Hybridization vs short targets (~25-nt)

Appendix IV: *In vitro* transcription

Appendix V: DIG-labeled RNA probe *in situ* hybridization of *Drosophila* embryos
Acknowledgements

I would like to thank Dr. Vaughn for giving me the opportunity to work in his lab and for being a great advisor. I have learnt so much from him in the past couple of years. It was a wonderful experience working in his lab. I would also like to thank the members of my Thesis committee, Dr. Susan Hoffman and Dr. Quinn Li, for their constructive criticism and valuable suggestions.

I would also like to thank my lab mates Sunetra Bhatla and Malini Varadarajan for all their help and support. I would like to thank Steve Strock, Kristen White, Kim Backscheider, Katie Niese and Robyn Bockrath who helped at various stages of my research. My special thanks to Chris wood for help with sequencing and Matt Duley for letting me use the electron microscopy facility.

Finally, I would like to thank my family and friends here and especially those back home in India. If not for their support and sacrifices, I wouldn’t have been here.

Graduate assistantship from Miami University and a grant from National institute of general medical sciences awarded to Dr. Vaughn supported this work.
Introduction

The \textit{sas10} and \textit{rnp-4f} genes in \textit{Drosophila melanogaster}

The \textit{sas10} gene in \textit{D. melanogaster} was initially designated CG4202 during genomic sequencing, and later shown to be an orthologue to the yeast gene when a BLAST analysis was carried out (Peters \textit{et al.}, 2003). As in yeast, the fly Sas10 protein contains a bipartite nuclear localization signal within its basic C-terminal domain. In fly, the protein encoded by the \textit{sas10} gene is 428 amino acids long. The gene is closely adjacent to the \textit{D. melanogaster} gene \textit{rnp-4f}, with a separation of only 95-nt (Peters \textit{et al.}, 2003), and contains no known introns. The \textit{sas10} gene (something about silencing) was first described in the yeast \textit{Saccharomyces cerevisiae} (Kamakaka and Rine, 1998) where it encodes a 610 amino acid protein (70,200 kDa) that is especially rich in charged amino acids. The protein has a bipartite nuclear localization signal, which is consistent with the finding that Sas10 protein in yeast is localized to the nucleus. In yeast, Sas10 protein is involved in releasing the repressed state of certain genes, such as those at the rDNA locus and telomeres, by altering chromatin structure. It has also been shown that \textit{sas10} is essential for cell viability in yeast, which suggests that it must have other functions in the cell in addition to silencing, as disruption of silencing is not required for viability.

Protein database searches revealed that the C-terminal domain of yeast Sas10 protein is well conserved in other eukaryotes like mouse and humans (Kamakaka and Rine, 1998).

In \textit{D. melanogaster}, the \textit{sas10} and \textit{rnp-4f} genes produce converging transcripts (Peters \textit{et al.}, 2003). These genes are single-copy and are located on the distal end of the X-chromosome (Hess \textit{et al.}, 1996). The \textit{rnp-4f} gene encodes the RNP-4F protein (943 amino acids), which has domains including a nuclear localization signal and at least two RNA recognition motifs, among others. The yeast orthologue, PRP24, plays a role in spliceosomal assembly (Rader and Guthrie, 2002). Human p110, an orthologue of PRP24, interacts with U6 and U4-snRNPs in the assembly of the spliceosome (Bell \textit{et al.}, 2002). PRP24 orthologues have been identified in several eukaryotic species (Fig. 1).
The *sas10* and *rnp-4f* isoforms in *D. melanogaster* produce naturally occurring overlapping transcripts and form sense/antisense pairs

The *rnp-4f* gene in *D. melanogaster* contains nine introns, and although its transcripts are alternatively spliced (Fetherson *et al.*, 2006), all transcripts thus far characterized appear to have the same 3’-terminus (Petschek *et al.*, 1997). This result was also obtained (Vaughn and Bhatla, unpublished) by sequencing EST clones for this gene in the FlyBase Annotation (http://flybase.bio.indiana.edu/). It has been shown via RT-PCR and cDNA library studies that *sas10* produces at least three different transcript isoforms in *D. melanogaster*: short, intermediate and long readthrough (Peters *et al.*, 2003). The long readthrough transcript makes use of a different promoter (P2) that is upstream of the promoter (P1) utilized by the other two transcript classes, and its transcription begins in late embryo stage. The longer readthrough and intermediate *sas10* isoforms overlap the *rnp-4f* mRNA transcript (Fig. 2). It is not yet known whether or not the synthesis of converging overlapping *sas10* and *rnp-4f* transcripts is restricted to *D. melanogaster*, or if this arrangement is conserved among diverse *Drosophila* species.

Converging overlapping mRNA transcripts have the potential to form double-stranded RNA, which is a potential target for modification or degradation pathways such as RNA interference (RNAi) and RNA editing (reviewed in Hutvagner and Zamore, 2002; Bass 2002, respectively). It has been observed that some *rnp-4f* transcripts are extensively A- to -G edited (Petschek *et al.*, 1996; 1997), which has been interpreted to reflect pairing of *rnp-4f* and *sas10* transcripts (Peters *et al.*, 2003), because the *Drosophila* editing enzyme dADAR (adenosine deaminases that act on RNA) has been shown to require duplex RNA as a substrate (Palladino *et al.*, 2000).

**Antisense RNA in posttranscriptional gene expression control**

Antisense regulation of gene expression has been an important area of study since the discovery of a large number of overlapping transcripts in mice and humans (Yelin *et al.*, 2003), rice (Osato *et al.*, 2003), and *Drosophila*. Our lab has identified 141 examples of overlapping transcripts from the *D. melanogaster* X-chromosome alone (Peters *et al.*, 2003). It is of considerable interest to determine whether or not potentially pairing natural antisense transcripts actually do pair, and if they form duplex RNAs, to determine the potential relationship of such pairing to the control of gene expression.

Antisense RNAs may alter the transcription, elongation, processing, stability, location and translation of mRNAs (Kiyosawa *et al.*, 2003). The α-thyroid hormone receptor
(TRα) gene in mammals produces two different classes of proteins, TRα1 and TRα2, by alternative splicing of the 3’-terminal exon (exon 10) (Hastings et al., 2000). The TRα1 protein binds thyroid hormone while the TRα2 does not. The TRα1/ TRα2 balance is tissue specific and important for the effective signaling of the thyroid hormone, which plays a key role in growth, differentiation and metabolism. Transcripts of the Orphan receptor RevErbAa (NR1D1) gene, transcribed in the opposite direction, overlap with the 3’-terminal exon region of TRα2 transcripts (Harding and Lazar, 1993). It has been reported that the levels of TRα2 mRNAs and RevErb mRNAs are inversely related. The exact mechanism behind this inverse relationship remains to be investigated, but it may be because of inhibition of transcription of the 3’-terminal exon (exon 10) or repression of TRα2 mRNA processing, or due to the stability of TRα2 mRNAs (Hastings et al., 2000).

The expression of translation elongation factor, eIF-2α, is posttranscriptionally regulated by its antisense RNA transcribed from the initiator promoter (Inr) downstream of the sense promoter region in T-lymphocytes (Noguchi et al., 1994). Here it is predicted that the duplex RNAs formed are degraded by nucleases leading to a decrease in the levels of eIF-2α. Since the overlap covers an intron/exon junction, splicing inhibition could also lead to degradation of the dsRNA (Noguchi et al., 1994). In mice, expression of a non-coding region called Xist is required for X-chromosome inactivation (Ogawa and Lee, 2002). Tsix is transcribed from the opposite DNA strand and it overlaps the entire Xist transcript. Tsix represses the expression of Xist. A number of models have been suggested to explain the mechanism behind this antisense repression, including RNAi. The repression may also occur at the transcriptional level or through repressive DNA elements. Another posttranscriptional model of regulation suggests that the stability of Xist mRNAs is increased in the absence of Tsix transcript (Ogawa and Lee, 2002). The exact mechanism of regulation is still not clear. The examples discussed above highlight the role of overlapping transcripts in the regulation of gene expression. The sas10 and rnp-4f genes, which produce overlapping mRNAs, may also be post-transcriptionally regulated.

MicroRNAs (miRNAs) are small 21-25 nt long RNAs (reviewed in Carthew, 2006). Biogenesis of miRNA starts with processing of long double-stranded RNAs into 70–100 nt stem-loop forming RNAs in the nucleus by a complex that includes the RNA-specific ribonuclease Drosha and a protein called Pasha (Sanchez et al., 2006). These hairpin
RNAs are subsequently cleaved into smaller miRNAs by the ribonuclease Dicer-1. These small RNAs are then loaded onto the RNA induced silencing complex (RISC), that directs mRNA cleavage. The miRNAs have been identified in diverse organisms that down-regulate gene expression by binding to the 3'-UTR region of complementary mRNA (reviewed in Munroe, 2004). The miRNAs may regulate gene expression through target mRNA cleavage, or by translational repression. High sequence complementarity results in target mRNA cleavage, otherwise repression takes place at the translational level (Rhoades et al., 2002). Some miRNAs are highly conserved across species (Bartel, 2004). In animals, miRNAs mostly regulate gene expression by translational repression, while in plants gene expression is regulated predominantly by mRNA degradation (Hartman et al., 2004). In plants, the majority of miRNAs have a high degree of complimentarity with their target and hence regulate gene expression by cleavage of the target mRNA (Rhoades et al., 2002). Recently, in the plant Arabidopsis, an miRNA called miR172 has been implicated in early flowering by regulating expression of a set of transcription factors called APETAL2 (AP2) during seedling development (Aukerman et al., 2003). The transcript levels of AP2 remained constant, but there was decrease in the AP2 protein levels, suggesting that translational repression is taking place. This is paradoxical because mi172 and AP2 have mismatches only at two places. The presence of small RNA fragments from the cleaved target suggests that the expression is regulated posttranscriptionally as well (Aukerman et al., 2003).

In Drosophila, miRNAs encoded from a gene called bantam negatively regulate the expression of the proapoptotic gene hid. The bantam gene plays a role in tissue growth by increasing proliferation of cells and inhibiting apoptosis via translational repression of hid (Brennecke et al., 2003). In C. elegans, let 7 and lin 4 encode two miRNAs complementary to a set of heterochronic genes. The lin 4 gene is expressed in the late larval moult to repress the expression of LIN-14 and LIN-28, which is important for transition to the next larval stage and also for differentiation. The let 7 miRNA coordinates the transition to adult stage by negatively regulating expression of another heterochronic gene, lin 41. Thus, miRNAs play an important role in signal transductions during the development of C. elegans (Reinhart et al., 2000). However, the functions of many miRNAs still remain to be investigated (Sevignani et al., 2006). Thus, antisense RNAs can regulate gene expression via different mechanisms. Also, given that miRNAs play a role in development, it will be interesting to see whether or not they have any role...
in the regulation of *sas10/rnp-4f* gene expressions, when they produce overlapping transcripts during certain periods of development.

**RNA editing in posttranscriptional gene expression control**

The A- to -G RNA editing pathway acts on double-stranded RNAs that may arise as a result of sense-antisense pairing. RNA editing by ADAR involves deamination of adenosine-to-inosine and has been observed in many multicellular organisms (reviewed in Bass, 2002). Hyperedited RNAs are retained in the nucleus and degraded by the conserved protein p54<sup>nrB</sup>, leading to a decrease in gene expression in *Xenopus* oocytes (Zhang and Carmichael, 2001). Although three ADAR genes have been identified in vertebrates, only one ADAR gene (*dADAR*) is present in *D. melanogaster* (Palladino et al., 2000).

RNA editing has also been implicated in modification of protein structure and function by altering mRNA codons and splicing patterns (Bass, 2002). A- to -G RNA editing in mammalian glutamate receptor (*gluR*) gene mRNAs facilitates the production of two protein isoforms from the same mRNA. Here a codon for arginine (AGA) is converted to a codon for glycine (GGA), which results in a new protein isoform. The change in the protein isoform alters the kinetic properties of the ionic channel formed from the gluR subunits (Aruscavage and Bass, 2000).

The *bFGF* transcript encoding basic fibroblast growth factor-2 in *Xenopus* oocytes was one of the first examples of the A- to -G class of editing to be described in an animal RNA. In this example, sense/antisense RNA pairing occurs when the smallest antisense bFGF transcript, named bFGF-AS, from the opposite DNA strand overlaps with part of the largest bFGF transcript to form a duplex. Pairing is followed by transcript degradation in this case (Kimelman and Kirschner, 1989; Volk et al., 1989). bFGF gene expression in mammals, including humans, is also regulated by the same antisense RNA mechanism (Knee et al., 1994). Recent work in trypanosomes has shown that RNA editing is behind correction of errors in open reading frames and changes to start and stop codons of mitochondrial RNAs (Golden and Hajduk, 2006). In mouse liver tissue, A- to -G editing in the 3’-UTR region of *CTN RNA* is required for its interaction with p54<sup>nrB</sup> and retention in the nucleus. This nuclear retention increases the stability of mRNA. Under stress, the cells cleave the 3’-UTR region of this RNA to produce the protein coding *mCAT2* transcript (Prasanth et al., 2005). It is not yet known whether the *rnp-4f* transcripts that are edited are degraded or not. It will be interesting to see if there is any functional change at the rnp-4f protein level as a result of this editing
The RNAi or RNA silencing pathway is present in many eukaryotic groups, including the embryonic stages of *D. melanogaster* (Tuschl et al., 1999). Long double-stranded RNAs are processed by the dicer enzyme into small 21-25 nt RNA fragments called short interfering RNA (siRNA). These small RNAs affect translational rate and chromatin structure. Three subclasses of siRNA termed repeat associated siRNA (rasiRNA), transcript associated RNA (tasiRNA) and small scan RNA (scnRNA) have recently been identified (reviewed in Kim, 2005). rasiRNA regulates gene expression by histone modification (Xie et al., 2004) or DNA modification in *Arabidopsis* (Zilberman et al., 2004). The function of tasiRNA remains unclear and is found only in organisms that have the enzyme RNA dependent RNA polymerase, including nematodes and plants (Kim, 2005). In *Tetrahymena*, scnRNAs derived from cleavage of dsRNA by dicer (Dcr) are primarily involved in genome rearrangements during conjugation by associating with Twi1p, a PAZ-Piwi Domain (PPD) protein that is required for internal eliminated sequences (IES) elimination (Mochizuki and Gorovsky, 2004).

Two dicer variants, Dcr-1 and Dcr-2, are present in *Drosophila*. Dcr-1 is involved in processing of precursors of miRNA, while Dcr-2 is responsible for cleavage of dsRNAs to give siRNAs (Tomari and Zamore, 2005). RNA induced silencing complex (RISC) is the effector complex into which both siRNA (reviewed in Hutvagner and Zamore, 2002) and miRNA (Kim, 2005) are loaded for subsequent degradation of mRNA that is complementary to the RNA in the RISC. The only proteins that are common to the RNA silencing complexes of siRNAs and miRNAs are the highly conserved Argonaute (Ago), the Vasa intronic gene (VIG) protein, Tudor-SN and the *Drosophila* ortholog of human fragile X mental retardation protein (FMRP), dFXR (Caudy and Hannon, 2004). So far five Ago proteins have been identified in *Drosophila* (Filipowicz et al., 2005). Ago-1 is associated with translational repression by miRNA, in contrast to Ago-2, which plays a role in siRNA mediated cleavage (Okamura et al., 2004). The different models for RNA silencing pathways are shown in Fig 3. In *Drosophila*, the only known example of a naturally occurring RNAi mechanism controlling gene expression involves suppression of testis expressed *Stellate* genes by the antisense transcripts of *Suppressor of Stellate* (Aravin et al., 2004).

The RNA editing and RNA silencing pathways interact with each other and it has...
been reported that hyperediting suppresses the RNAi pathway (Scadden and Smith, 2001). Recently, it has been shown that editing of miRNA seen in hematopoietic tissues leads to a decrease in their population following degradation by Tudor-SN, a component of RISC (Yang et al., 2006). It is also thought that the editing of miRNAs is important in increasing the diversity of miRNAs (Blow et al., 2006). The rnp-4f and sas10 duplex RNA could be a target for a number of the antisense RNA or double stranded RNA pathways discussed above.

**D. melanogaster long sas10 and rnp-4f transcript levels are inversely related**

It has been shown that the long readthrough sas10 isoform and rnp-4f levels are inversely related during the late embryo stages in *D. melanogaster* development (Peters et al., 2003), as shown in Fig. 4. Degradation of sas10 and rnp-4f mRNAs here may not arise as a consequence of RNA editing, because the levels of dADAR editase activity are relatively low during fly embryonic stages when these transcripts are degraded (Hanrahan et al., 2000; Ma et al., 2002). RNAi could however be behind the observed simultaneous degradation of sas10 and rnp-4f mRNA levels during fly development. The hypothesis that the RNAi pathway is behind the observed mRNA level declinations awaits testing during *D. melanogaster* development.

**Role of the 3’-UTR in mRNA localization and alternative 3’-termini selection in D. melanogaster**

The 3’-UTR region in eukaryotic mRNAs has important regulatory functions, which include roles in mRNA localization, stability and translation (reviewed in Nibuya et al., 2003; Hesketh, 2004). In *Drosophila*, the 3’-UTR region of some transcripts are involved in posttranscriptional control (Chen et al., 2004). Specific *cis*-acting regulatory elements have been identified in species including humans that play a role in polyadenylation site selection (Hall-Pogar et al., 2005). The variable 3’-UTR regions of the three *hts* transcript classes in *Drosophila* are important in directing each of these mRNAs to specific cytoplasmic locations (Whittaker et al., 1999). mRNA localizations in *Drosophila* are critical for embryogenesis (Jantsch et al., 1992). For example, the 3’-UTR directed *nos* mRNA localization to the posterior pole of embryos is important for abdomen formation (Gavis and Lehmann, 1994).

*Cis*-acting signals have been identified in the 3’-UTR of several maternal mRNAs which, during animal embryogenesis, play an important role in localization of these
mRNAs to specific locations (reviewed in Bashirullah et al., 2001). Localization patterns are established by both protection and degradation of mRNAs (Tadros et al., 2003). For example, mouse cytoplasmic polyadenylation elements in the 3’-UTR, which are required for degradation of the maternal mRNAs, have been identified (Alizadeh et al., 2005). In Drosophila, a 3’-UTR destabilization signal has been identified for fushi tarazu mRNA that is important for establishment of body plan (Riedl and Jacobs-Lorena, 1996). At the midblastula transition (MBT) stage, which occurs about three hours after fertilization, maternal mRNAs are degraded, shifting developmental control to the zygotically encoded mRNAs. This shift happens after thirteen synchronous syncytial mitotic divisions and is regulated by maternal gene activity (Sibon et al., 1999). For example, degradation of the maternally encoded mRNAs twine and string is required for this transition from a synchronous mitotic stage to an asynchronous zygotically controlled mitotic stage (Edgar and Datar, 1996). Maternal transcripts, including Hsp83, string and nanos, are degraded very early in embryogenesis, in contrast to rpA1 transcripts, which are protected (Bashirullah et al., 1999). Several trans-acting proteins involved in the protection and degradation of these mRNAs have been identified (reviewed in Bashirullah et al., 1998). Recently, it has been shown that Smaug, a known translational repressor of nanos, is involved in degradation of Hsp83 by recruiting CCR4/POP2/NOT deadenylase (Semotok et al., 2005). Two degradation pathways have been described in D. melanogaster that facilitate the midblastula transition (Bashirullah et al., 1999). The first pathway is maternally encoded and becomes active 4-5 h after egg activation. The second pathway is zygotically programmed and is activated 1.5-2 h after fertilization. Both pathways acting alone are sufficient for degradation of maternal transcripts, but they must function together to ensure maternal mRNA transcript degradation by the MBT stage.

A study focusing on gene expression during the life cycle of D. melanogaster identified 27 strictly maternal transcripts, including rnp-4f (Artbeitman et al., 2002). Interestingly, the timing of degradation of rnp-4f differs from most of the other transcripts (Fig. 5), suggesting that degradation may be via a different pathway. The role of RNAi or other pathways that act on double-stranded RNAs containing rnp-4f awaits investigation.

As discussed above, alternative sas10 transcripts in D. melanogaster have variable 3’-termini. Regulatory elements in sas10 responsible for alternative 3’-UTR selection have not yet been identified. Furthermore, it is the 3’-UTR region which determines whether
or not sas10 transcripts overlap rnp-4f mRNA, and hence analyzing the sas10 gene for potential regulatory signals is one of the goals of my Thesis. We do not yet know if the intermediate length sas10 isoform, like the long readthrough variant, waits to appear until the late embryo stages of development. Since the intermediate length variant also overlaps the rnp-4f transcript, it is anticipated that this mRNA class will pause before its initial appearance. If not, some mechanism may prevent its pairing with rnp-4f mRNA during early embryogenesis. We do not yet know the mechanism behind dramatic simultaneously declines in sas10 and rnp-4f mRNA levels during late embryo stages in fly development. If overlapping mRNAs do form duplexes and this is important in observed simultaneous sas10 and rnp-4f mRNA level declinations, then it would be predicted that the classes of sas10 which potentially pair with rnp-4f mRNA would co-localize with the latter. It will be interesting to know whether rnp-4f and sas10 are adjacent to each other and produce converging transcripts in diverse Drosophila species. If so, then the recently proposed model involving antisense regulation for D. melanogaster rnp-4f gene expression control (Peters et al., 2003) may be general among insects.

One of the goals of my Thesis work will be to determine the variety of sas10 mRNA isoforms in D. melanogaster across development via Northern hybridizations. The levels of each sas10 mRNA isoform observed in D. melanogaster with reference to constitutively expressed rp49 mRNA will be quantified. I will also localize different sas10 mRNA isoforms across fly development by using DIG- labeled RNA probes. I will be using a genomics approach to determine whether or not sas10 and rnp-4f genes are in close proximity and of opposite polarity in diverse Drosophila species, as observed for D. melanogaster. In addition, I will use this approach to identify potential 3’-UTR signals for regulation of sas10 alternative transcript termini selection. My goal will also be to initiate evaluation of the RNAi hypothesis for regulation of sas10 mRNA abundance in D. melanogaster by observing the appearance or lack of appearance of 21-25 nt probe-specific RNA fragments across development. Finally, I will utilize an RNAi approach to knock down the expression level of at least one of the longer sas10 mRNA isoforms, and determine the corresponding phenotype.
Materials and Methods

Fly stocks and propagation
For this study wild-type Drosophila melanogaster strain Oregon R, originally obtained from the National Drosophila Stock Center, was used. Several other species obtained from the Drosophila Species Stock Center in Tucson, AZ were also utilized. D. persimilis and D. pseudoobscura were grown on media containing opuntia cactus powder. Standard culture media consisting of cornmeal, agar, dark corn syrup, Brewers yeast, propionic acid and mold inhibitor was used to culture the flies. The flies were incubated in half-pint milk bottles under 12 h light-dark cycles in a 23°C environmentally controlled chamber. Harvesting embryos for RNA isolation from different developmental stages was done using 60 mm diameter Petri dish media plates, consisting of dextrose, sucrose, apple juice, propionic acid, phosphoric acid and agar. The collected embryos and other developmental stages, including adults, were stored at -70°C for subsequent RNA isolations.

Embryo preparation and storage for hybridizations
Chorion and vitelline membranes surrounding the embryos were removed (B.W. Jones, personal communication), prior to in situ hybridizations (described below). The embryos were stored in 100% methanol until hybridizations were performed (Appendix I).

Total cell RNA isolation
RNA from different developmental stages was isolated by following Nakagawa and Cohen (1967), with modifications which included DNase I treatment (Koller et al., 1987). RNAs were quantified by absorbance determination at 260 nm, and stored at -70°C in diethylpyrocarbonate (DEPC)-treated water. The quality of the RNA was verified via 1% agarose gel electrophoresis with a brine shrimp total cell RNA standard, using intactness of bands as the criterion.

Affinity chromatography for mRNA purification
Poly (A⁺) mRNAs were obtained from total cell RNAs by using 10 ml columns (0.4 x 4 cm poly-prep columns, Bio-Rad) packed with 0.3 ml dry oligo-dT cellulose powder (Amersham), as described in Chabot (1994). The total RNA was run through the
columns in batches of 3 mg. Prior to running the sample through the column, the RNA was dissolved in 9 ml of RNA elution buffer (10 mM Tris-HCl pH 7.5, 1 mM disodium EDTA, 0.05% SDS) and was heat denatured (10 min at 65°C). After addition of 0.9 ml 5 M LiCl, samples were run through the column, which had been pre-equilibrated with RNA binding buffer (10 mM Tris-HCl pH 7.5, 400 mM LiCl, 1 mM disodium EDTA, 0.5% SDS). The run-through RNA samples were collected and passed over the column a second time. The column was then washed with 5 ml of RNA binding buffer. RNA elution was carried out by using 1.5 ml of elution buffer. The collected mRNA was then precipitated using 75 μl of 5 M NaCl and two volumes of 100% ETOH. The mRNA recovery was about 1-2% of the input total cell RNA. Quantifications were done by A$_{260nm}$ measurement, and RNAs were stored at -70°C in DEPC-treated water.

**Agarose gel electrophoresis of RNAs**

**(A) mRNAs**

Poly (A$^+$) mRNAs (5 μg) were dissolved in 10 μl DEPC-treated water and then combined with 5 μl 10 X MOPS buffer, 20 μl formamide, 5 μl 37% formaldehyde and 4.4 μl RNA tracker dye. RNAs were heat denatured for 15 min at 65°C, then loaded into 120 ml, 1.2% denaturing formaldehyde agarose gels (SeaKem GTG agarose, Cambridge Bioscience). The gels were 7 mm thick and prepared in 1 X MOPS buffer (20 mM MOPS pH 7.0, 5 mM sodium acetate, 1 mM EDTA). The gels were run at 75 V at 4°C until the bromophenol blue tracker dye was about 2.5 cm from the end of the gel (Ausubel et al., 1994-1997). The flanking lanes in the gels contained 2000-nt and 5000-nt bovine cell rRNAs (Sigma) and *Drosophila* rRNAs (2,112-nt, 1,995-nt and 1,788-nt) for use as size markers. Gels were then stained in 100 ml of 1 X MOPS buffer with 10 μl 1:10000 diluted SYBR Green II dye (Molecular Probes). Blue fluorescence mode was used in the Molecular Dynamics Storm 860 phosphorimager for scanning the stained gel to visualize the rRNA sizers. The images were saved in the computer for analysis using the Image Quant Version 5.2 program.

**(B) Total cell RNAs**

Total cell RNAs (30 μg) from different *Drosophila* species were electrophoresed using a BIO-RAD Sub-cell GT submerged gel system at 10 mA for 48 h until the rRNA size marker had moved approximately 75% of the total gel length.
**Polyacrylamide gel electrophoresis of RNAs**

In RNAi experiments, 40 μg of total cell RNAs isolated from various developmental stages was dissolved in 20 μl DEPC-treated water and then combined with 40 μl formamide-dye solution (80% formamide). RNAs were heat denatured for 5 min at 65°C, then loaded into 16% denaturing polyacrylamide gels of 60 ml volume. The gels were prepared by combining 24 g urea, 9.1 g acrylamide, 0.5 g bis-acrylamide, 6 ml 10X TBE buffer pH 8.3, and 22 ml DEPC-treated water. Gel polymerization was promoted by adding 80 mg ammonium persulphate and 10 μl TEMED. Gels were pre-run in 1X TBE buffer for 1 h at 200 V. RNA samples were added after their denaturation for 5 min at 65°C and the gel was initially run at 50V for 30 min, then at 300 V at room temperature until the bromophenol blue tracking dye had run to the bottom of the gel (Aravin et al., 2001). These gels contained single-stranded DNA oligonucleotide markers in the size range 18-31 nt. After electrophoresis, gels were stained with SYBR Green II and scanned in the phosphorimager to visualize the markers.

**Northern blot analysis**

**(A) Agarose gels**

Poly(A⁺) mRNAs which had been separated in agarose gels were blotted onto nylon membrane by overnight capillary transfer in 10X SSPE (0.1 M sodium phosphate buffer pH 7.4, 10 mM EDTA, 1.5 M NaCl), Sambrook et al. (1989). After the transfer, membranes were air dried for 20 min followed by another 20 min in the 80°C oven. The mRNA was then cross-linked to the nylon membrane by UV light exposure. The membranes were then baked in the 80°C oven for two more hours. Membranes were stored at 4°C until further use. Hybridization was carried out at 40°C overnight, as described in Appendix II. Probes for hybridization were constructed by random priming using ³²P-dCTP with a commercially available “Prime-a-Gene” kit (Promega), following the manufacturer’s directions.

**(B) Polyacrylamide gels**

RNA transfer to positively charged nylon membranes was carried out by electroblotting at 200 mA using a high amp power supply overnight, followed by UV cross-linking.
Hybridization was carried out overnight at 37°C as described in Appendix III. Probes were prepared as described above.

**Polymerase chain reaction (PCR)**

Specific regions of the *sas10* gene were amplified from genomic DNA (already available in the lab) via PCR, for use in making hybridization probes or to clone prior to carrying out *in vitro* transcriptions. PCR (Mullis and Faloona, 1987) was done in 25 μl reaction volumes containing 10 mM Tris-HCl pH 8.3, 50 mM KCl, 1.5 mM MgCl₂, 0.05% NP-40, 200 μM of each dNTP, 1 μM each primer, 1 μl (0.1 μg) genomic DNA and 0.5-1 ul (2.5-5 U) Taq polymerase (Promega). Thermocycler parameters utilized were as follows: 2 min initial denaturation at 97 °C, followed by 30-35 cycles of denaturation at 95°C for 30 sec, 1 min annealing at gradient-optimized temperature and 1 min extension at 72°C. The appropriate annealing temperature for each primer set was established by doing a preliminary gradient PCR. Primers amplifying specific regions of the *D. melanogaster* genome were designed and utilized, based on the completely sequenced *sas10* gene (Peters et al., 2003).

**Primers utilized for PCR**

(1) *sas10* primer sets

A. Full length *sas10* gene
   - Fwd 5’-[GACGACTACGTGATGAG]-3’
   - Rev 5’-[AACTAGGTGCGGAACTTG]-3’

B. Specific to intermediate isoform 3’ UTR
   - Fwd 5’-[CAAATGCCGGCGGCTGA]-3’
   - Rev 5’-[CTTGGTTCATCAAGAAACCA]-3’

C. Specific to *rnp-4f/sas10* overlap region for RNAi fragment probe
   - Fwd 5’-[AGCGGCCATGGTTAGTTTGT]-3’
   - Rev 5’-[GCTCCGATGATGATAATGAT]-3’

(2) *rp49* primer set

   - Fwd 5’-[GCTTCAAGATGACCATCC]-3’ (57 °C)
   - Rev 5’-[AAGTCCGTTATATTACGTT]-3’
**DNA fragment isolation by agarose gel electrophoresis**

DNA fragments produced by PCR were separated on low melting temperature agarose gels (Promega) and the band of interest was excised from the gel. The DNA fragment was then purified from the gel using Agar-ACE enzyme (Promega) according to the manufacturer’s directions, and the DNA was ethanol precipitated. In some cases isolated fragments were directly $^{32}$P labeled by random priming for use as hybridization probes. In other cases, fragments were cloned into a pGEM-T Easy (Promega) vector for *in vitro* transcription of DIG-labeled probes (described below).

**DNA fragment cloning into plasmid vector/transformation/colony selection**

PCR generated DNA fragments were cloned into the pGEM-T Easy vector by overnight incubation at 4°C, following the manufacturer’s directions. The ligated plasmid was transformed into competent *E.coli* cells by the heat shock method (Sambrook *et al*., 1989). Individual white colonies expected to contain the desired insert were picked from ampicillin drugged (60 μg/ml final concentration) and X-Gal (100 μl, concentration 20 mg/ml) treated LB nutrient agar plates. Small scale plasmid isolation was carried out for these white colonies following Davis *et al.* (1980), as described below. Alternatively, a colony PCR procedure was utilized to identify cells containing the desired recombinant vector. In colony PCR, individual positive white colonies were picked using sterile pipette tips and dispersed in 10 μl distilled autoclaved water by vortexing for 30 sec (with the tips still in the Eppendorf tubes). The colonies were numbered to keep track of the positive clones. A PCR was done using primer sets specific to the insert using 1 μl of the dissolved DNA. The PCR products were run on a 1% agarose gel with an appropriate size marker to confirm the presence of the insert. Plasmids from the selected individual colonies were isolated and purified for sequencing verification using a Wizard Plus midipreps (Promega) kit, as described below.

**Small scale plasmid isolation**

The isolated white colonies were grown in 1 ml LB medium overnight at 37°C. Plasmid isolation was done by following Davis *et al.* (1980). The isolated DNA was ETOH precipitated and treated with RNase A at 37°C for 1 h. Following this, the DNA was subjected to shaking for removal of proteins in chloroform/isoamyl alcohol (24:1, v/v) and ETOH precipitated. *EcoRI* enzyme was used to cut on either side of the insert to
release it from the plasmid. The digest was run on a 1% agarose gel. In the presence of an insert in the plasmid two bands were visible, one for the plasmid and another for the insert.

**Bulk plasmid isolation**

Cells which were verified to contain the desired insert via a gel check following *EcoRI* digestion were cultured overnight in a flask containing 400 ml LB medium. The cells were transferred to 500 ml bottles and centrifuged for 15 min at 8000 rpm in the Beckman JA-17 rotor at 4°C. DNA was isolated by using a Wizard Plus Midipreps DNA purification system (Promega), following the manufacturer’s instructions. The precipitated DNA was stored at -70°C.

**DNA sequencing**

The plasmid with the insert was sequenced using the ABI automated sequencer (ABI-3100 or ABI Prism 310) in the Miami DNA Core Facility to verify that the correct fragment had been cloned, and to determine the orientation of the insert in the vector. The sequencing in these machines was done using the dideoxynucleotide triphosphate (ddNTP) method (Sanger *et al.* 1977). Prior to sequencing, DNA amplification of the region of interest was carried out in the thermocycler, following the directions of the supplier (Amersham-Pharmacia). The total reaction volume (20 μl) contained ~0.1 ug cDNA template, 1 μl of 1 μM forward or reverse primer, 2 μl ET or Big Dye sequencing mix, 6 μl 2.5 X sequencing buffer, and water. The following parameters were used for the PCR reaction: 25 cycles of denaturation at 95°C for 20 sec, 15 sec annealing at 50°C, and 1 min extension at 60°C. The amplified DNA was ETOH precipitated and stored at -20°C. Before loading in the sequencer the DNA samples were dissolved in 20 μl of megabase loading solution.

**In vitro transcription**

Recombinant plasmids containing an insert of known orientation were linearized by cutting with an appropriate restriction enzyme prior to transcription from either the T7 or SP6 vector promoter. DIG-labeled RNA probes were synthesized by using the appropriate template DNA, using a commercially available “Riboprobe” kit (Promega), following which the RNA probe was precipitated and quantified using the nanodrop
instrument. RNA was stored at -70°C for future use. The full procedure is described in Appendix IV.

**In situ DIG hybridizations**

The embryos were hybridized with DIG-labeled RNA probes, after which anti-DIG antibody conjugated to alkaline phosphatase enzyme was used as secondary antibody. A colorimetric analysis was done following the action of alkaline phosphatase upon BCIP/NBT (substrate) to reveal primary antibody binding sites. In these procedures, a protocol provided by B.W Jones (personal communication) was used. The full procedure is described in Appendix V.

**Microscopy**

The hybridized embryos were initially examined under a 80i Nikon microscope in the Vaughn lab. A Nikon DR 6 camera attached to the Olympus SZX12 microscope in the Miami University Electron Microscopy Facility was used to prepare photomicrographs of sas10 mRNA localizations.

**Comparative genomic analysis**

The entire genomes of several different *Drosophila* species are now being sequenced, and a few have been completed. These raw, non-annotated sequences are being deposited at the NCBI Trace Archives (http://www.ncbi.nlm.nih.gov/blast/mmtrace.shtml) and can be accessed at the Welcome Trust/Sanger Institute Trace Server (http://trace.ensembl.org/pet/traceviews). Comparative genomic analysis was utilized to determine whether or not the sas10 and rnp-4f genes are adjacent to each other in diverse *Drosophila* species and of opposite polarity as in *D. melanogaster* (Peters *et al.*, 2003). The 3’-ends of rnp-4f and sas10 have been observed to be evolutionarily conserved in our preliminary studies, and this feature was used to design query sequences for BLAST searches. Relevant sequences were identified from different species for further analysis. A multiple sequence alignment on these sequences was carried out using with the ClustalW program (http://align.genome.jp), which was then studied to determine if conserved potential regulatory sequence elements such as the poly (A) signal could be recognized, for specification of alternate 3’-termini within sas10 transcripts.
Results

Quantification of \textit{sas10} transcript levels across development

\textit{Sas10} mRNA isoform levels were quantified in a previous study (Peters \textit{et al}., 2003) via an RT-PCR approach. In that study, the upstream PCR primer was located within the \textit{sas10} coding region and the downstream primer was moved progressively to fall within (a) the short isoform 3’-UTR, (b) the intermediate length isoform 3’-UTR, or (c) positions further downstream. The results were of limited value, in that the levels of short and intermediate length \textit{sas10} isoforms could not be individually determined and also RT-PCR does not permit determination of transcript size. These limitations could be overcome using the developmental Northern experimental approach.

\textit{sas10} transcript levels across development were quantified using Northern hybridizations in which constitutively expressed \textit{rp49} was employed as a standard of reference. mRNAs extracted from different developmental stages were separated on the basis of size (5 µg per lane) by electrophoresis through a 1% denaturing agarose gel. SYBR Green II dye was used to stain the gel for visualization of RNA (Fig 6A). rRNA contaminants in the mRNAs provided additional size references. After electrophoresis, the RNAs were transferred to nylon membranes. Northern hybridizations were carried out with both full length \textit{sas10} and also \textit{rp49} DNA probes. The probes were made by the random-priming method using template DNAs that were generated by PCR with primers corresponding to most of the coding sequences of \textit{sas10} and \textit{rp49}. Upon hybridization, two major \textit{sas10} transcripts were visualized; a 1,550-nt short length isoform and a 1,750-nt intermediate length isoform (Fig 6B). The long readthrough \textit{sas10} isoform previously detected via RT-PCR (Peters \textit{et al}., 2003) and expected to be about 6000 -nt in length (Hunt-Smith, 1999) was not detected, perhaps due to its very low abundance.

Quantifications for the two major \textit{sas10} isoforms and for total \textit{sas10} transcripts (all isoforms combined) expression levels were carried out using the phosphorimager Image Quant program. The levels of \textit{sas10} were expressed relative to \textit{rp49}, which is a constitutively expressed gene (O’Connell and Rosbash, 1984).

The levels of \textit{sas10} for each developmental stage were calculated from three different Northern blots and averaged. Standard deviations for each determination were also calculated (Fig. 7). The two combined major \textit{sas10} isoforms (Fig. 7A) are at their highest levels in the early embryo stages, suggesting that both isoforms may be maternal
transcripts. This point will be revisited in the section on in situ sas10 mRNA localizations (below). The levels of both individual major isoforms then decrease simultaneously (Fig. 7B, C), reaching relatively low levels by the 6-12 h embryo stage, somewhat in contrast to the previously determined RT-PCR results in which low transcript levels were not seen until after about 15-21 h. The reason for this observed timing difference is not known. The results suggest a more rapid and more pronounced degradation for the intermediate length sas10 isoform in comparison to the shorter length isoform, which may reflect different mechanisms for their degradation.

rnp-4f and sas10 genes are closely adjacent to each other and transcribed in opposite direction in diverse Drosophila species

It has previously been reported (Peters et al., 2003) that in D. melanogaster, the rnp-4f and sas10 genes are closely adjacent, and transcribe in opposite directions, and that the longer sas10 transcripts pair with those arising from rnp-4f. This observation led to a model for control of gene expression in which the resulting double-stranded mRNAs are recognized by either the A-to-G RNA editing or RNAi pathway, followed by degradation of both transcripts during embryogenesis. Insofar as this model was developed solely from observations made on D. melanogaster, it would be interesting to determine if these two genes follow a similar pattern in other species. An initial step was taken to see if this was the case.

Using comparative genomic analysis for several of the newly sequenced Drosophila species total genomic sequences (Fig. 8), it was asked whether or not sas10 and rnp-4f are closely adjacent to each other and also if they would be predicted to transcribe mRNAs that are opposite in polarity. The conserved 3’-ends of rnp-4f and sas10 were used as query sequences in two different BLAST searches against the different Drosophila genomic sequences. The BLAST searches and subsequent alignment (Fig. 9A, B) revealed that the sas10 and rnp-4f genes are closely adjacent to each other and are also predicted to transcribe mRNAs that are opposite in polarity in every Drosophila species for which the region of interest could be identified, using the as yet not completely assembled genomic sequence data. Interestingly, the analysis also showed that these two genes are closely adjacent and transcribe in opposite directions in Anopheles gambiae, which diverged 250 million years ago from the Drosophila lineage.

The alignment was then examined more closely in an effort to recognize potential regulatory sequence elements, with the goal of predicting whether or not sas10 and rnp-4f
transcripts overlap in other species as they do in *D. melanogaster*. In the case of this species, previous work including total sequencing of full length cDNAs obtained by screening libraries had identified stop codons and potential poly (A) signals. In the case of the two major *sas10* isoforms, poly (A) signals for the short and intermediate length transcripts had been identified (Peters *et al.*, 2003). These already determined identifications of potential regulatory elements were utilized to make predictions about comparable elements in other species. It was found that the stop codons for both *rnp-4f* and *sas10* occupy comparable positions in every species studied including the distantly related mosquito (*Anopheles gambiae*). As expected considerable sequence conservation is present in the 3’-terminal coding region in both genes for all the species. Although the canonical poly (A) signal sequence is A-A-U-A-A-A in both vertebrates and *Drosophila* (Graber *et al.*, 1999; Zhang *et al.*, 2005), the less frequent alternative signal U-A-U-A-A-A had previously been identified in *rnp-4f* for *D. melanogaster* (Petschek *et al.*, 1997). In *rnp-4f* for the closely related sister species *D. simulans*, A-A-U-A-A-A occurs at the corresponding position and also at a position further downstream. This identical sequence occurs in the same *rnp-4f* downstream position for every other species examined. The *rnp-4f* poly (A) sequences for *D. melanogaster* and the closely related sister species, *D. simulans*, are found to be closer to the coding sequence than other distantly related species. Also, in *D. simulans* two poly (A) sites have been identified, which suggests that there may be two *rnp-4f* isoforms in this species. The *rnp-4f* poly (A) sites for other species are further away from the gene, and this may be the ancestral state.

The *sas10* gene proximal canonical poly (A) signal A-A-U-A-A-A-A had previously been identified for *D. melanogaster* (Peters *et al.*, 2003), and the identical sequence occurs at a corresponding positions in the alignment for species closely related to this species, while the non-canonical predicted poly (A) signal A-U-U-A-A-A for the distantly related species *D. pseudoobscura* and *D. persimilis* occurs further downstream. The *sas10* poly (A) site for *A. gambiae* is at the same spot as in *D. pseudoobscura* and *D. persimilis*, which suggests that this is the ancestral state. The *sas10* poly (A) signal for the species *D. simulans*, *D. yakuba* and *D. erecta*, which are closely related to *D. melanogaster*, are closer to the gene and lie outside the overlapping region. The non-canonical *sas10* gene distal poly (A) signal A-A-U-A-U-A had previously been identified in a long transcript class for *D. melanogaster* (Peters *et al.*, 2003), but this sequence element is not observed at the corresponding position for any other species.
studied (Fig 9 B), including the closely related sister species D. simulans. If multiplesas10 isoforms exist in species other than D. melanogaster, their identification will require direct experimental verification. An attempt was made to do so (see below). This analysis suggests that rnp-4f and sas10 transcripts must certainly overlap in D. pseudoobscura, D. persimilis and mosquito, which are distantly related to D. melanogaster, and that this is likely also the case for the other Drosophila species studied.

**Attempted identification of different sas10 isoforms in diverse Drosophila species using Northern hybridization**

It is confirmed in this Thesis that in D. melanogaster there are at least three different sas10 mRNA isoforms, as previously determined using an RT-PCR approach (Peters et al., 2003). The above described genomic analysis predicted that sas10 and rnp-4f produce overlapping transcripts in diverse Drosophila species (Fig. 9A, B). Northern hybridization experiments were designed with the aim of verifying this prediction. In the initial experimental design, total cell RNAs were isolated from adult flies of diverse Drosophila species, followed by preparation of poly (A+) mRNA for use in Northern hybridizations. Unexpectedly, mRNA recoveries were too low to be useful, for unknown reasons.

Insufficient quantities of total cell RNAs precluded repeating the mRNA selections, and it was decided to try using total cell RNAs for these experiments. Adult total cell RNAs, isolated from ten different Drosophila species, were electrophoresed on a 1% denaturing agarose gel. Ethidium bromide was used to visualize the RNA bands (Fig. 10), which were judged to be of high quality based on intactness of rRNA bands.

The stained total cell RNAs were photographed, then transferred to nylon membranes for hybridizations using 32P-labeled probes specific to all isoforms of the sas10 transcript. Insofar as sequence alignment shows that there is a high degree of similarity between sas10 sequences in different Drosophila species, it was judged that the same D. melanogaster sas10 probe could be used for all species studied. However, repeated Northern hybridizations using high specific activity probe failed to detect sas10 transcripts in any of the species, including D. melanogaster. One of the reasons for the failure could be due to inefficient transfer of RNA from the gel to the nylon membrane due to the utilization of ethidium bromide instead of SYBR Green II dye. The problem could have also been due to utilizing total cell RNA instead of mRNA, a potential
problem insofar as *rnp-4f* produces transcripts of very low abundance. The high content of ribosomal RNA in the total RNA samples could also potentially sterically interfere with hybridizations, since rRNAs and *sas10* mRNAs are of similar size. In the future, these experiments could be repeated using embryonic mRNAs.

**Attempted evaluation of RNAi hypothesis for regulation of *sas10* mRNA abundance in *D. melanogaster* during embryogenesis**

The molecular mechanism behind simultaneous decline of the two major *sas10* isoforms and also of *rnp-4f* transcripts during embryogenesis is not known, but the RNAi pathway is a likely one. The role of RNAi was evaluated by looking for the presence of 21-25 nt long RNA fragments derived from degradation of *rnp-4f* and *sas10*, because these fragments are a characteristic feature of the RNAi pathway. Total cell RNAs from different developmental stages were electrophoresed in a 16% polyacrylamide gel. SYBR Green II dye was used to stain the RNA for visualization (Fig. 12A), and after preparation of a photograph RNAs were transferred via electroblotting to positively charged nylon membranes. The flanking lanes were loaded with single-stranded DNA markers (18-31 nt size range). Initially, much of the entire *sas10* gene region (~1200-nt) was $^{32}$P-labeled to high specific activity by using the random priming method and used as a probe for detection of gene-specific RNA fragments in the size range 21-25 nt. The predicted 21-25 nt RNA fragments could not be visualized following phophorimager analysis. It was anticipated that the negative result may be due to the choice of the probe position within the gene, which excluded the 3’-UTR. Hence, a shorter probe (~220-nt) corresponding to the overlapping region between *rnp-4f* and *sas10* mRNAs was used to probe the Northern blots. This approach also failed to detect the predicted small RNA fragments (Fig. 12B). This result could have been due to high background levels, or to very low expression levels of the *rnp-4f* gene, although it was expected that the higher abundance *sas10* gene transcript fragments should have been detectable.

**Localization of different *sas10* mRNA isoforms across *D. melanogaster* development using DIG-labeled RNA probes**

RNA editing previously described during late embryo and adult stages in *D. melanogaster* was observed in the transcript overlap region of both *rnp-4f* and intermediate length *sas10* mRNAs (Peters *et al.*, 2003). Because the A- to -G RNA editase ADAR has been shown to require double-stranded RNA for a template (reviewed in Bass, 2000), the *D. melanogaster* example has been interpreted to show that these
transcripts do form a double-stranded structure. This would then predict that \textit{rnp-4f} and the longer \textit{sas10} isoforms occupy identical cells in the developing fly embryo. Although two studies have described \textit{rnp-4f} mRNA localizations during \textit{D. melanogaster} embryogenesis (Hess \textit{et al.}, 1996; Varadarajan, 2006), it is not known where individual \textit{sas10} isoforms localize with respect to the potentially interacting \textit{rnp-4f} mRNAs. To determine the localization pattern of the different major \textit{sas10} isoforms during embryogenesis, two different digoxigenin (DIG)-labeled RNA probes were constructed and utilized (Fig 13A, B). A \textasciitilde{1200}-nt long DIG-labeled antisense RNA probe was used to study localization of all combined isoforms of the \textit{sas10} gene. Technical difficulties prevented obtaining useful results. To determine the localization pattern of the longer \textit{sas10} isoform, a \textasciitilde{220}-nt DIG-labeled anti-sense RNA probe specific to the 3\textquotesingle-U TR region of the intermediate length \textit{sas10} isoform overlapping the \textit{rnp-4f} trailer sequence was constructed. A sense DIG-labeled RNA probe specific to the \textit{rnp-4f} intron 0 region, shown to be retained in about 50\% of the transcripts (Fetherson \textit{et al.}, 2006), was used as a negative control. All probes were made by \textit{in vitro} transcription from appropriate genomic DNA inserts cloned into a pGEM-T easy vector system. The localization pattern of the intermediate length \textit{sas10} isoform and \textit{rnp 4f} mRNA during embryogenesis is shown in Fig. 14. These results show that abundant levels of transcript are present throughout the yolk cytoplasm even at the earliest developmental stages, suggesting that \textit{sas10} mRNAs are maternally derived. The staining is observed to be generally distributed throughout the embryo as development progresses. Noticeably, a high concentrations of transcripts is observed along the dorsal roof of the developing ventral nerve cord and brain, very much resembling the pattern described for \textit{rnp-4f} mRNAs during development (Hess \textit{et al.}, 1996; Varadarajan, 2006).
Discussion

Posttranscriptional control of gene expression by overlapping antisense RNAs is becoming recognized as a more widespread phenomenon than originally thought, especially in the central nervous system (reviewed in Korneev and Shea, 2005). In *D. melanogaster*, the *rnp-4f* and *sas10* genes produce converging transcripts and the extent of their overlapping is dependent upon the class of *sas10* transcripts, whose 3'-UTRs vary during development (Peters *et al.*, 2003). This Thesis focuses on aspects of *rnp-4f* gene expression control by an antisense transcript encoded by the *sas10* gene.

The *sas10* gene produces two major abundance isoforms: a 1,550-nt short length isoform and a 1,750-nt intermediate length isoform. A third very long isoform, potentially overlapping the entire *rnp-4f* gene, has also been reported (Peters *et al.*, 2003), but is not sufficiently abundant to be detected in Northern hybridizations. The abundances of the two major isoforms, short and intermediate, decline to a very low level at the 6-12 h embryo stage from their highest levels at the early embryo stage. The *sas10* intermediate length isoform and also the combined isoforms localize to the dorsal roof of the ventral nerve cord in later stage embryos, which is identical to the localization pattern observed for *rnp-4f* isoforms (Fig. 14). It is known that *rnp-4f* and *sas10* form double-stranded RNA based on the fact that *rnp-4f* transcripts are edited by dADAR, which requires double-stranded RNA for a substrate. Taken together, it can be said that there is a distinct possibility that *rnp-4f* gene expression is post-transcriptionally regulated by one or more of the antisense RNA pathways. The finding that these two genes are closely adjacent and transcribed in opposite direction in diverse *Drosophila* species suggests that this regulation of *rnp-4f* gene expression by its antisense RNA may be a general phenomenon.

Control of gene expression by antisense RNAs

A number of mechanisms have been identified that are activated in the presence of double-stranded RNAs that arise as a result of sense/antisense pairing, including A- to -G RNA editing, silencing by siRNAs and translational modulation by miRNAs. These pathways can regulate gene expression at the level of transcription, or post-transcriptionally, or at the level of translation. It has been reported that the 3'/3' overlaps, as in our case, are more common than 5'/5' overlaps. These 3'/3' overlaps are more evolutionary conserved than 5'/5' overlaps in diverse species, and their
evolutionary conservation predicts a prominent role for antisense RNA in gene regulation (Sun et al., 2005). In our case, we have shown that \textit{sas10} and \textit{rnp-4f} are closely adjacent to each other and probably produce converging overlapping transcripts in diverse \textit{Drosophila} species, extending to the mosquito \textit{Anopheles gambiae}, which diverged from \textit{Drosophila} 250 million years ago (Gaunt and Miles, 2002). This observation is not unique to our case. The orientation of many other genes overlapping in their 3’-regions are conserved among vertebrates, including mammals (Sun et al., 2005).

For example, in \textit{Arabidopsis}, \(\Delta^1\)-pyroline-5-carboxylate dehydrogenase (\textit{P5CDH}) is involved in cell growth, regulation of stress related genes and production of reactive oxygen species (Borsani et al., 2005). \textit{RO5} is a gene of unknown function that is transcribed in the opposite direction to produce an antisense transcript in response to salt stress, leading to the formation of double-stranded RNA. The expression of \textit{RO5} can be compared to that of the overlapping longer \textit{sas10} isoform. The difference between these two cases is that \textit{RO5} expression is a function of stress while \textit{sas10} is developmentally regulated. The \textit{RO5/ P5CDH} double-stranded RNA is then cleaved into siRNAs which are 24-nt long. These 24-nt naturally occurring RNAs lead to the production of 21-nt long siRNAs. Presence of 21-nt siRNAs leads to downregulation of \textit{P5CDH} gene expression by mRNA degradation. The decreased expression of \textit{P5CDH} results in reduction of salt stress, but there is accumulation of ROS and toxic metabolites. It has been suggested that \textit{SRO5} protein might play role in negating the effects of this toxicity.

The A- to- G class of RNA editing has been reported in the mammalian 2C subtype serotonin receptor, which is G-protein coupled and involved in downstream signaling affecting various physiological processes. Here, editing leads to a change in the protein coding sequence. The serotonin mRNAs form a stem-loop, which acts as the substrate for the editase ADAR. The change in amino acid sequence decreases interaction between the serotonin receptor and the G-protein. Thus, this post-transcriptional modification can be used to modulate the strength of serotonin signal transduction (Burns et al., 1997). It has been reported that mRNAs which are extensively A- to-G edited are retained in the nucleus and degraded. In the \textit{Xenopus} oocyte, \textit{p54\textsuperscript{arb}}, PSF, and matrin 3 are part of a protein complex that recognizes inosine residues in the mRNA arising due to editing and then degrades the RNA (Zhang and Carmichael, 2001). It is not yet known whether \textit{rnp-4f} transcripts that are edited are degraded by this mechanism or not.

Gene silencing by antisense RNAs can take place at the translational level, with or without degradation of the target mRNA. In \textit{C. elegans}, \textit{let 7} -encoded miRNAs
regulate the expression of LIN-41, a protein required for transition from pupae to adult, by translational repression (Bagga et al., 2005). This translational repression results in the destabilization of lin-41 mRNA by deadenylation (Wang et al., 2006). In mammalian cells, the fibroblast growth factor-2 (FGF-2) gene is bi-directionally transcribed. FGF-2 antisense RNA is complimentary to the 3’-UTR region and its expression leads to down regulation of FGF-2 protein levels, but its mRNA levels remain constant. Thus, here the regulation is through translational repression (Li and Murphy, 2000). In contrast to this, the regulation of LIN 14 protein levels by the antisense RNA encoded by lin 4 takes place at the level of translation without any changes in lin 14 mRNA levels. Examples where gene expression involves decreases in mRNA levels are more relevant to our experimental results. These examples illustrate the different pathways acting on different levels of gene expression which a double-stranded RNA could follow, leading to degradation of the transcripts. The rnp-4f/sas10 duplex could be degraded after translational repression or earlier by any of the mechanisms described above.

**Role of the 3’-UTR in gene expression, mRNA localization and transcript stability**

The different *Drosophila sas10* isoforms vary in their 3’-UTRs. Consequences of having different 3’-UTRs include altered gene expression, mRNA localization and stability. In *Drosophila* ovary, the *hu-li tai shao (hts)* locus produces three classes of mRNAs (R2, N32 and N4) which differ in their 3’-UTR (Whittaker et al., 1999). This difference in the 3’-UTRs is critical for the differential localization of these different mRNA isoforms. The R2 and N32 transcripts are restricted to the nurse cells, in contrast to N4 mRNA, which is localized to the anterior pole of the embryos. A 345-nt long 3’-UTR has been identified in N4 transcripts that is required for this localization. Secondary structure formation in the 3’-UTRs of mRNAs also seems to be critical for localization. One such example involves the *bicoid* mRNA, which is localized to the posterior pole of *Drosophila* embryos, where it plays a role in patterning (Gamberi et al., 2002). It has been shown that binding of Stau, a double-stranded RNA binding protein involved in localization of many mRNAs (Clark et al., 2002), requires loop–loop interactions between two *bcd* mRNA molecules (Ferrandon et al., 1997). Oskar protein is required for posterior cell patterning in *Drosophila*. The 3’-UTR here plays a key role in the localization-coupled translational control of the encoding mRNAs (Nakamura et al., 2001).
miRNAs regulate gene expression by recognizing the 3’-UTR of their targets. Giraldez et al. (2006) observed that miR-430 predominantly targets degradation of maternal mRNAs in a 3’-UTR dependent manner in zebrafish. In Drosophila, miRNAs (miR-11, miR-2b) have been found complimentary to known negative post-transcriptional regulatory elements such as the two 3’-UTR sequence motifs, the K box (cUGUGAUa) and the Brd box (AGCUUUA) (Lai, 2002). The 3’-UTRs of sas10 and rnp-4f get masked due to the overlap; hence, it will be interesting to investigate the potential effect of this in the localization of these mRNAs and gene expression by miRNAs.

Localization of rnp-4f and the intermediate length sas10 isoform is uniform at the early embryo stages. By the 12-15 h stage, intermediate length sas10 isoform is found primarily in the dorsal roof of the ventral nerve cord along with the rnp-4f transcripts. This localization may arise as a result of an overall selective protection/degradation mechanism. In Drosophila, Hsp83, nanos and Pgc are degraded everywhere except at the posterior of the embryo at the midblastula transition stage (Bashirullah et al., 1999). It has been shown that the signal for degradation may reside in the 3’-UTR, as in the case of maternally encoded transcripts of Hsp83. It has been suggested that a protective mechanism is active in the posterior which prevents degradation of the transcript (Bashirullah et al., 2001). It is not yet known which proteins are important for such selective protective functions.

Similarly, mechanisms must clearly be present that protect overlapping sas10 intermediate length and rnp-4f mRNAs when they co-localize in the syncytial blastoderm stage and later in the CNS. This mechanism may repress degradation pathways such as RNAi to protect the double-stranded RNAs.

**rnp-4f gene expression is developmentally regulated by its antisense sas10**

rnp-4f is a maternally encoded mRNA that is present during the very early embryo stages (Hess et al., 1996; Varadarajan, 2006). However, as discussed previously in this Thesis, the rnp-4f and sas10 expression profiles are different from those of most other maternally expressed genes, which are degraded during the midblastula transition stage at about 2.75 h post-fertilization (Artbeitman et al., 2002). This observation suggests that the mechanism of degradation for these two transcripts may be unique.

It has been shown that sas10 and rnp-4f transcripts form duplex RNA molecules during development (Peters et al., 2003), and lead to the activation of double-stranded RNA pathways described above. These pathways can act independently of each other or
together in a concerted manner. For example, it has been reported that RNAi is antagonized by A- to -G RNA editing (Scadden and Smith, 2001). In this Thesis, the decline in levels of \textit{rnp-4f} during embryogenesis cannot be a consequence of the RNA editing mechanism, for the levels of the editase enzyme responsible for this process are reportedly very low during \textit{Drosophila} embryo stages (Hanrahan \textit{et al.}, 2000; Ma \textit{et al.}, 2002). Thus, we could very well attribute the rapid simultaneous decline in both \textit{rnp-4f} and \textit{sas10} mRNA levels during embryogenesis to the activity of some other mechanism, potentially one that acts on double-stranded RNAs.

RNA silencing by siRNAs or miRNAs may play a role in regulating \textit{rnp-4f} gene expression. It is known that the RNAi pathway is operational in even very early \textit{Drosophila} embryonic cell extracts (Tuschl \textit{et al.}, 1999). Results presented in this Thesis show that both the intermediate length \textit{sas10} isoform and \textit{rnp-4f} co-localize in the syncytial blastoderm embryo stage, and thus share a common cytoplasm. Insofar as these molecules can potentially form a duplex molecule at that stage, it is surprising that they are not immediately degraded by the RNAi pathway. This observation suggests that the RNAi pathway itself must either be developmentally regulated, or factors must be present which protect the double-stranded RNA during specific stages of development. As previously mentioned, it is unlikely that RNA editing protects these mRNAs from RNAi degradation. Future research may therefore be directed at learning not only the molecular mechanism for simultaneous degradation of both \textit{sas10} and \textit{rnp-4f} during mid-early stages, and the potential role of antisense RNA in this process, but also the molecular mechanism behind the inhibition of transcript degradation during very early development and within the developing central nervous system (CNS).
Future Research Directions

Both genetic and reverse genetic experimental approaches will be valuable in future research to answer some of the questions posed in this Thesis. To investigate the potential role of RNAi as the cause of simultaneous declines in levels of Drosophila sas10 and rnp-4f mRNA levels during development, a dcr-2 mutant fly line recently produced in the R. Carthew lab, which is defective in this pathway, may be valuable. If it is observed that sas10 and rnp-4f transcript levels do not decline during embryogenesis in these mutant flies, then it could be concluded that RNAi does act on the sas10/rnp-4f double-stranded RNA to control their expression. The gene high risk (hri) encodes a protein that protects some double-stranded RNA from the RNAi pathway (R.Carthew, personal communication). rnp-4f and sas10 mRNA levels could be quantified in a hri mutant to evaluate the potential role of this protective protein from the effects of the RNAi pathway in syncytial stage embryos. The role of RNA editing in this decline in transcript levels cannot be completely ruled out. Hence, a Drosophila dADAR mutant fly line produced by the R. Reenan lab group could be used to test this hypothesis. Recently, it was shown by our lab using in situ hybridizations that dADAR transcripts are present in the earliest embryo stages (Varadarajan, 2006). However, nothing is known about whether or not editing is active during very early embryo stages. This could be clarified by sequencing cDNA clones obtained from very early embryo stages.

RNAi knockdown using exogenously injected dsRNA is a powerful tool that could be used to probe the role of sas10 in rnp-4f gene expression control, mRNA localization, and the potential for a sas10 role in central nervous system (CNS) development. Comparison between hypomorph embryos in which levels of combined sas10 (all isoforms) are reduced versus those in which intermediate sas10 transcript levels are reduced would be an excellent approach that may lead to clarification of the roles of sas10 isoforms in rnp-4f gene expression control. The embryos in which intermediate length sas10 isoforms are knocked down could be used to measure the effect on rnp-4f mRNA levels during embryogenesis to determine whether or not rnp-4f gene expression is regulated by antisense RNA. We will be using an RNAi bullet that will target the intermediate isoform sas10 isoform region that does not overlap with the rnp-4f mRNA and not present in the short sas10 isoform. Given that the extent of RNA/RNA pairing would be greatly reduced following intermediate length sas10 isoform knockdown, rnp-4f expression profiles would be expected to be different from the control embryos if
antisense gene regulation is taking place. Preliminary studies show that knockdown of *rnp-4f* long isoform mRNAs in embryos using RNAi leads to development of defects in the CNS (Bhatla, 2006). It will be interesting to now determine if the same phenotype is observed in embryos in which intermediate length *sas10* isoforms have been knocked down using appropriately constructed RNAi duplex molecules for microinjection. The role of 3′-UTR in RNA localization is well established. Sense/antisense pairing between *sas10* and *rnp-4f* may mask the 3′-UTR region of both molecules. It will be interesting to see if there is any change in localization pattern of *rnp-4f* mRNAs in the *sas10* intermediate length RNAi knockdown embryos. These experiments will help to clarify the role of *sas10* in *rnp-4f* gene expression control and central nervous system development in *Drosophila*. 
**Figure 1. Orthologues of yeast splicing assembly factor PRP24 in diverse organisms**

Alignment of protein domains shows that the PRP24 orthologue in *D. melanogaster* is *rnp-4f*. Protein domains include: repeated tetratricopeptide repeats (heavy black arrows), potential nuclear localization signal (short blue arrow), RNA recognition motifs (RRM) (green ovals), and a conserved C-terminal motif (black triangle). Figure is modified from Rader and Guthrie (2002).

**Reprinted with permission from publisher**
Figure 2. Orientation of sas10 and rnp-4f genes in D. melanogaster
Known transcripts of varying length are shown. The two genes are only 95-nt apart (S) and have opposite transcriptional polarities (arrows). Putative alternative sas10 poly(A⁺) signals are indicated by location and sequence. Short and intermediate sas10 cDNAs have been completely sequenced and contain poly (A) tails. Full length of transcripts labeled “RT-PCR” is not known. Alternate sas10 promoters P1 and P2 are shown (redrawn from Peters et al., 2003).
Figure 3. RNAi silencing pathways
Shown here are the different RNA silencing pathways. The miRNAs are processed by a two-step mechanism involving the two ribonucleases Drosha and Dicer from hairpin-forming RNA. The siRNAs are processed from long double-stranded RNA by Dicer-2. Figure is from Kim (2005).

Reprinted with permission from publisher
**Figure 4.** The inverse relationship between long readthrough *sas10* isoform and *rnp-4f*

Transcript abundances are relative to those of the constitutively expressed gene *rp49*. *sas10* primer set 1 detects all *sas10* transcript isoforms, whereas primer set 3 detects only the long readthrough isoforms (note differences in scales for A and C). Figure is redrawn from Peters *et al.* (2003).
**Figure 5.** Expression profile of strictly maternal *D. melanogaster* genes

Shown here are the gene expression profiles of strictly maternal genes. These genes were identified based on the expression patterns of already known maternal genes. Notice the unique expression profile of *rnp-4f* (CG3312). The indicated high expression levels for *rnp-4f* in males vs females is a probable error, since it has previously been shown that this gene is maternally expressed with high transcript levels in ovary. This figure was compiled from the Yale *Drosophila* Gene Expression Database available online at http://genome.med.yale.edu/Lifecycle/ (Artbeitman *et al.*, 2002).
Yellow color represents high relative levels of expression while blue represents low levels. The brightest color is three-fold or greater differential from the reference black.
**Figure 6.** Denaturing agarose gel electrophoresis of *D. melanogaster* poly (A⁺)-selected RNAs from different developmental stages

A. Poly (A⁺)-selected RNAs (5 μg) prepared from total cell RNAs across development were run in a denaturing 1% agarose gel, which was then stained with SYBR Green II. The quality of the poly (A⁺) RNAs was indirectly verified using the intactness of the three contaminating rRNA bands (arrows) as the criterion. Total bovine cell RNA and total *Drosophila* cell RNA were used as size markers.

B. Following electrophoresis in 1% denaturing agarose gels and Northern transfer, selected mRNAs were detected by hybridization to gene-specific probes.

- *rp49* (~600-nt) is the constitutively expressed gene that was used for quantification of *sas10* isoforms.

- Two major *sas10* transcript isoforms are seen: short (S) and intermediate (I) are visualized, sized at about 1,550-nt and 1,750-nt, respectively.
Figure 7. Quantification of major sas10 transcript levels across development

Results were obtained via phosphorimager-based quantifications of Northern hybridizations, and show transcript levels relative to rp49. Standard deviation bars are indicated. Each data point represents the average of three independent determinations.

A. Total sas10 transcript levels (all isoforms combined).

B, C. Individual sas10 isoform levels across development. Note that levels of both isoforms decline at approximately the same stage.
Figure 8. Phylogenetic tree for dipteran insects utilized in this study

*Drosophila* species highlighted in red were utilized for *rnp-4f/sas10* intergenic spacer sequence alignment. The mosquito *Anopheles gambiae* is an early evolving dipteran insect. Tree is from Russo *et al.* (1995) and Gaunt and Miles (2002).
**Figure 9 A. Sequence alignment across the rnp-4f/sas10 intergenic spacer**

Regulatory sequence elements identified experimentally for *D. melanogaster* or predicted for other species are color highlighted. The *rnp-4f* and *sas10* transcripts are predicted to converge (wavy lines) in all species shown and to be encoded by the corresponding closely adjacent genes. Consensus sequence is shown in blue letters, and comparable nucleotide position numbers are indicated. The sequences of *sas10* poly (A) signals are shown regardless of DNA strand.
Poly (A) signals and their sequences are indicated, and have been identified experimentally only for *D. melanogaster*. The arrows indicate the direction of transcription of the genes. The *rnp-4f* and *sas10* are predicted to converge in all species studied, and to be encoded by closely adjacent homologous genes. At least one canonical poly (A) signal (A-A-U-A-A-A) for the *sas10* gene was identified in all the *Drosophila* species studied. For the *sas10* gene in *D. melanogaster*, a proximal and distal poly (A) signal was identified. U-A-U-A-A-A, a previously reported non-canonical poly (A) signal for the *rnp-4f* gene was also identified in *D. melanogaster*. The *rnp-4f* and *sas10* mRNAs are predicted to be capable of forming double-stranded mRNAs for *D. melanogaster*, *D. persimilus*, *D. pseudoobscura* and *A. gambiae*. The non-canonical poly (A) signal (U-A-A-U-U-U) for the *A. gambiae sas10* gene is shown. It is not yet known whether or not additional longer isoforms exists for *D. simulans*, *D. yakuba* or *D. erecta* (represented by the question marks). The evolutionary distances between the organisms studied is indicated by the phylogenetic tree on the left (not drawn to scale). The sequences of *sas10* poly (A) signals are shown regardless of DNA strand. IGS=Intergenic spacer.
**Figure 10.** Denaturing 1% agarose gel electrophoresis of adult total cell RNA from different *Drosophila* species

The three rRNA bands (2,112-nt, 1,995-nt and 1,788-nt) indicated by arrows can be visualized in all the species. The high quality of the RNAs is indicated by the intactness of the three rRNA bands.
Figure 11. Non-denaturing 1% agarose gel electrophoresis of adult total cell RNA from different D. melanogaster developmental stages

Total cell RNAs (5 μg) isolated from developmental stages were evaluated for RNA quality, as indicated from appearance of the two major rRNA bands. Electrophoresis was in 1% non-denaturing agarose gel, which was stained with ethidium bromide. Gel verifies the quality of total cell RNAs which were eventually used to check for the presence of 21-25 nt long RNA fragments as an indication of RNAi activity.
<table>
<thead>
<tr>
<th>Adult</th>
<th>Pupa</th>
<th>3rd Instar</th>
<th>2nd Instar</th>
<th>1st Instar</th>
<th>12-18 h</th>
<th>6-12 h</th>
<th>4-8 h</th>
<th>0-4 h</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
**Figure 12. Denaturing polyacrylamide gel electrophoresis of *D. melanogaster* total cell RNAs from different developmental stages**

A. Total cell RNAs (40 μg) prepared from different developmental stages were run in a denaturing 16% polyacrylamide gel, which was then stained with SYBR Green II. Arrow shows position of the small 30-nt, 2S rRNA band (Pavlakis *et al.*, 1979). The flanking lanes contain single-stranded DNA markers (18-31 nt size range).

B. Following gel electrophoresis and Northern blotting, a double-stranded DNA probe prepared by labeling with $^{32}$-P and potentially capable of detecting both short *rnp-4f* and *sas10* RNA fragments was hybridized to the nylon membrane. RNA fragments in the size range 21-25 nt were not detectable via hybridization experiments, possibly due to high background levels, although some of the high abundance size markers did hybridize in the flanking lanes.
Figure 13. *In vitro* transcription gels showing DIG-labeled RNA products

A. Top panel shows the ~200 -nt DIG-labeled RNA probe (arrow), run alongside the double-stranded pBR322/Alu I-cut DNA marker. This probe is specific to the extended 3’-UTR region unique to the intermediate length *sas10* isoform which overlaps the *rnp-4f* trailer.

B. Bottom panel shows the ~1200 -nt DIG-labeled RNA probe (arrow), run alongside the double-stranded λ/EcoRI+BamHI-cut DNA marker. This probe is predicted to recognize all *sas10* isoforms combined.
**Figure 14. sas10 and rnp-4f mRNAs co-localize during development**

Upper panels I and II A-F show the localization pattern of the *rnp-4f* mRNA and intermediate length *sas10* isoform respectively. Both *sas10* and *rnp-4f* are distributed throughout the embryonic cytoplasm and central yolk region during earliest developmental stages, consistent with the conclusion that both represent maternal mRNAs. Both transcripts localize to the developing ventral nerve cord later in embryogenesis. All views are lateral, with the anterior to the left and the ventral side down.  

**A.** Stage 3 (0-2 h) showing dense uniformly distributed hybridization signal.  

**B.** Stage 4 (0-2), syncytial blastoderm showing staining in the overall embryo and central unstained egg cytoplasm.  

**C.** Stage 7 (2-4 h) showing staining of the cephalic furrow (cf).  

**D.** Stage 9,10 (4-8 h) showing the extending germ band and heavy staining of the mesoderm and ectoderm.  

**E.** Stage 12 embryo (8-12 h) showing staining of the dorsal component of segmentally repeated neuromeres in the ventral nerve cord (arrow).  

**F.** Stage 13 (12-15 h) showing continued heavy localization of *rnp-4f* and intermediate length *sas10* transcripts to the dorsal roof of the ventral nerve cord (arrow).  

Lower panels G-L show the negative control, in which a sense strand corresponding to the intron 0 region of *rnp-4f* was used as the probe. The *sas10* panel is from this Thesis and that of *rnp-4f* is from Varadarajan (2006).
Appendix I

Dechorionation and Devitellinization of *Drosophila* Embryos

1. Grow embryos to desired developmental stages on apple juice agar plates with a dab of yeast paste
2. Loosen the embryos from the plate using distilled water and a brush
3. Suspend the embryos in distilled water on the plates
4. Pour the embryos in distilled water through a nylon sieve (100µ mesh)
5. Rinse the embryos well with distilled water to remove residual yeast paste
6. Blot the bottom of the sieve to remove distilled water
7. Dismantle the sieve and transfer embryos using a brush to a 20 ml glass scintillation vial containing freshly prepared 50% bleach solution
8. Allow embryos to sit for 5 min in the bleach solution
9. Remove the bleach solution using a Pasteur pipette and rinse the embryos in distilled water several times to ensure removal of all the bleach
10. Wash the embryos in 0.7% NaCl, 0.02% Triton X-100 and remove any excess solution by using a Pasteur pipette [0.35 g NaCl, 10 µl Triton X-100, 50 ml distilled water]
11. Add 5 ml of heptane followed by 5 ml of 3.7% formaldehyde fixative solution to the washed embryos [3.7 % formaldehyde: 5 ml stock 37% formaldehyde + 45 ml distilled water]
12. Rock the vial at room temperature for 20 min on the nutator
13. The dechorionated embryos remain at the interface between the fix (lower phase) and the heptane (upper phase)
14. Remove and discard as much of the lower phase as possible including the embryos that have settled in the bottom
15. Save the embryos at the interface
16. Replace the fix solution with an equal (5 ml) volume of 100% methanol maintaining the level of heptane at 5 ml (add more if required)
17. Shake the vial vigorously (manually) for one minute such that the embryos pop out of their vitelline membrane and sink to the bottom of the methanol
18. Transfer de-vitellinized embryos from the bottom of the vial using a Pasteur pipette to another scintillation vial containing about 5 ml of 100% methanol
19. Wash the embryos 3 times in 100% methanol for 2 min each
20. Store embryos in 100% methanol at 4°C (tightly capped)

References

B.W. Jones, University of Mississippi [personal communication]
Appendix II
Hybridization vs long targets (>100 -nt) Using Random Prime-Labeled Probe

1. Float filter on double distilled water to wet it completely, then submerge it
2. Agitate filter in pan for 2 h at 40° C containing 150 ml of 1st pre-hybridization solution [100 ml “Aravin mix” solution, 15 ml 50X Denhardt’s, 6 ml 25% SDS, and 29 ml water]
3. Make 25 ml 2nd pre-hybridization solution [13.4 ml “Aravin mix”, 2 ml 50X Denhardt’s, 800 μl 25% SDS, 9 ml deionized formamide, and 400 μl tRNA [10 mg/ml stock i.e. 4,000μg]]
4. Seal filter in Seal-a-Meal bag with 20 ml 2nd pre-hybridization solution
5. Incubate 2 h at 40° C, with agitation
6. Denature DNA probe by boiling for 3 min and add to bag
7. Continue hybridization at 40° C overnight, with agitation
8. Open bag in radiation sink and discard remaining probe into proper vessel
9. Two 15 min room temp gentle agitation washes [10 ml 20X SSC, 2 ml of 25% SDS, and 188 ml water]
10. Background counts should noticeably diminish from out-of-bag to post first wash
11. Two 15 min room temp gentle agitation washes of [2.5 ml 20X SSC, 400 μl 25% SDS, and 197 ml water]
12. Photograph in phosphorimager under “storage phosphor”
Appendix III
Hybridization vs Short Targets (~25-nt) Using Random Prime-Labeled Probe

1. Float filter on double distilled water to wet it completely, then submerge it
2. Agitate filter in pan for 2 h at 37° C containing 150 ml of 1st hybridization solution [50 ml “Aravin mix” solution, 3 ml 50X Denhardt’s, 42 ml 25% SDS, and 55 ml water]
3. Make 21 ml 2nd pre-hybridization solution [6.7 ml “Aravin mix” solution, 400μl 50X Denhardt’s, 5.6 ml 25% SDS, 8 ml deionized 100% formamide, and 400 μl tRNA [10 mg/ml stock i.e. 4000 μg]]
4. Seal filter in Seal-a-Meal bag with 21 ml 2nd pre-hybridization solution
5. Incubate 2 h at 40° C, with agitation
6. Denature double-stranded DNA probe by boiling for 3 min and add to bag
7. Continue hybridization at 37° C overnight, with agitation
8. Open bag in radiation sink and discard remaining probe into proper vessel
9. Two 5 min room temp gentle agitation washes [20 ml 20X SSC and 180 ml water]
10. Background counts should noticeably diminish from out-of-bag to post first wash
11. Two 15 min washes at 42° C, with gentle agitation [20 ml 20X SSC, 4ml 25% SDS, and 176 ml water]
12. Photograph in phosphorimager under “storage phosphor”
Appendix IV

*In vitro* Transcription for *in situ* Hybridization

**Restriction of the DNA template**
1. Linearize 50 ug of template DNA with appropriate restriction enzyme
2. Check for completion of the restriction digestion on a 1-2% agarose gel

**Purification of the DNA template**
3. To the digested DNA, add 1/10 volume of 1M NaCl followed by 2 volumes of 100% ethanol and mix well
4. Leave in -20°C for 1 h or -70°C for 15 min
5. Centrifuge at maximum speed for 15 min at 4°C
6. Remove all of the supernatant and air dry the pellet for at least 20 min in the cold room
7. Resuspend the DNA in 25 ul of DEPC treated water
8. Nanodrop the DNA to check the precise concentration of the DNA (approx 0.5 ug/ul)

**In vitro transcription**
9. Set up a reaction tube in a 100 ul PCR tube depending on the probe to be transcribed as per the following table

<table>
<thead>
<tr>
<th>Template DNA (0.5 ug) Make it up with DEPC water</th>
<th>5.5 ul</th>
</tr>
</thead>
<tbody>
<tr>
<td>T7/SP6 10 X buffer</td>
<td>1 ul</td>
</tr>
<tr>
<td>T7/SP6 polymerase</td>
<td>1.5 ul</td>
</tr>
<tr>
<td>DIG-NTP mix (Roche)</td>
<td>1.5 ul</td>
</tr>
<tr>
<td>RNasin</td>
<td>0.5 ul</td>
</tr>
<tr>
<td>Total</td>
<td>10 ul</td>
</tr>
</tbody>
</table>

10. Incubate at 37°C for 2 h (water bath)
11. Add 10 ul of DNase I mix:
<table>
<thead>
<tr>
<th>DEPC water</th>
<th>2 ul</th>
</tr>
</thead>
<tbody>
<tr>
<td>DNase I buffer</td>
<td>2 ul</td>
</tr>
<tr>
<td>5 Units of DNase I enzyme (Promega)</td>
<td>5 ul</td>
</tr>
<tr>
<td>0.1 M DTT (Promega)</td>
<td>1 ul</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>10 ul</strong></td>
</tr>
</tbody>
</table>

12. Incubate at $37^\circ$C for 20 min (water bath)
13. Repeat the DNase I treatment again
14. Add 50 ul of 7.5 M ammonium acetate and mix well
15. Add 375 ul cold 100 % EtOH. Mix well.
16. Incubate in crushed ice slurry for 10 min to precipitate RNA
17. Centrifuge 15 min at top speed, 4$^\circ$ C to pellet RNA probe
18. Drain tube well and invert to dry for 10 min in cold room
19. Resuspend RNA pellet to 25 ug/ml concentration in RNA suspension buffer
20. Store dissolved RNA probe at 70 $^\circ$ C
Appendix V
DIG-labeled RNA Probe in situ Hybridization of Drosophila Embryos

Rehydrating and fixing the embryos
1. Prepare dechorionated, devitellinized embryos and store in 100% methanol at 4°C
2. a. Transfer ~100 ul of test embryos in methanol to a 1.5 ml Eppendorf tube [also process ~100 ul of control embryos and 100 ul of embryos for use in the preabsorption step (below)]
   b. Alternatively, aspirate methanol such that 500 ul of the solution is left
3. Add an equal volume of 1X PBT solution
4. Rock in the nutator for 5 min at room temperature
5. Allow the embryos to settle down; remove all of the 1X PBT/ methanol solution
6. Add 500 ul of 1X PBT and wash 3 times on the rocking nutator for 5 min each at room temperature
7. Post-fix in 1 ml of 1X PBT containing 3.7% formaldehyde for 20 min in a rocking nutator at room temperature
8. Remove all of the fixing solution and wash the embryos in 500 ul of 1X PBT about 4 times with rocking in the nutator at room temperature for 5 min each
9. Refix in 500 ul of 3.7% formaldehyde/1X PBT for 20 min with rocking at room temperature
10. Wash in 500 ul 1X PBT for 4 times, 5 min each with rocking on the nutator at room temperature

Prehybridization of the embryos
11. Take 1 ml of the hybridization buffer, add to it 20 ul of t-RNA (10 mg/ml stock conc’n) and 100 ul of heat-inactivated normal goat serum
12. Add 500 ul of this mixture to each of the tubes containing embryos
13. Prehybridize at 52°C with rocking in the nutator for 3 h

Addition of primary (DIG-labeled) probe
14. Add 5-7 ul of the DIG-antisense RNA probe (25 ug/ml stock conc’n) to one tube and 5-7 ul of the (negative control) DIG-sense RNA probe control RNA (25 ug/ml stock conc’n) to another tube
15. Hybridize overnight at 52°C with rocking on the nutator

**Preabsorption of secondary antibody**
16. While doing step #15, bring the embryos to be used for preabsorption into 1X PBT, then remove as much solution as possible. Add 1 ul of sheep Anti- DIG-AP (secondary) antibody (conjugated to alkaline phosphatase, Roche cat. #1-093-274) in 200 ul of 1X PBT to the embryos (1:200)
17. Allow them to mix overnight at 4°C on nutator
18. Retrieve and save the supernatant and store at 4°C for future use

**Washing of the embryos-I**
19. Rinse at least 3 times with 500 ul wash buffer at room temperature for 1 h each on a rocking nutator
20. Finally, wash with 500 ul of wash buffer at room temperature for 1 h with rocking
21. Rinse twice with 500 ul of 1X PBT for 5 min each at room temperature with rocking
22. Rinse with 500 ul of 1X PBT for 30 min with rocking at room temperature

**Blocking step**
23. Add 500 ul of 1X PBT/ 1% BSA to both the test and control embryos
24. Allow embryos to mix on nutator for 4 h at room temperature or overnight at 4°C

**Addition of preabsorbed secondary antibody**
25. Remove as much as the 1X PBT/1% BSA as possible
26. The final required dilution of the preabsorbed secondary antibody is 1: 2000
27. Add 70 ul of 1X PBT, 20 ul of heat-inactivated normal goat serum, and 10 ul of preabsorbed secondary antibody to the test and control embryos from step #25
28. Allow the embryos to bind preabsorbed secondary antibody for 4 h at room temperature or overnight at 4°C on the nutator

**Washing of the embryos- II**
29. Rinse briefly for 5 min with 500 ul of 1X PBT at room temperature on nutator
30. Rinse twice with 500 ul of alkaline phosphatase (AP) buffer 5 min each at room temperature
Staining of the embryos

31. Prepare 1 ml of the color developer solution in a separate tube
   1 ml AP buffer 7 ul NBT solution (100 mg/ml stock conc’n) [Roche cat. # 1383-213] 7 ul BCIP solution (50 mg/ml stock conc’n) [Roche cat. # 1-383-221]
32. Add 500 ul of the color developer solution to each tube
33. Cover the tubes with foil (light-sensitive reaction) and mix them on the nutator at room temperature
34. The development of color depends on the intensity of signal and takes between 20 min to 24 h
35. Monitor the color change by observing small aliquots of embryos under the dissecting microscope
36. When the appropriate color intensity is reached, stop the reaction by adding 500 ul of 1X PBT
37. Rinse at least 3 times with 1X PBT for 3 min each on the nutator at room temperature until all the blue color is lost from the solution
38. Remove as much of the 1X PBT as possible
39. Add 1 ml of 70% ethanol and rinse on rocking nutator for 5 min at room temperature
40. Remove all of the 70% ethanol and add 1 ml of 100% ethanol.
41. Rinse for 5 min on rocking nutator in 100% ethanol at room temperature
42. Remove all of the 100% ethanol
43. Add 1 ml of 1X PBT, wash on rocking nutator 3 times, 5 min each at room temperature
44. Remove all of the 1X PBT

Glycerol mount preparation of the embryos

45. Add approximately 200 ul of 70% glycerol/1X PBS to the embryos
46. Allow the embryos to mix overnight at 4°C
47. Store embryos at 4°C in the dark
48. The embryos are ready for whole mounts/ dissections when they settle down to the bottom of the Eppendorf tube [store slide mounts at 4°C in the dark]

References: B.W. Jones, University of Mississippi [personal communication]
Literature Cited


Knee RS, Pitcher SE, Murphy PR. 1994. Basic fibroblast growth factor sense (FGF) and antisense (gfg) RNA transcripts are expressed in unfertilized human oocytes and in differentiated adult tissues. Biochem Biophys Res Commun 205:577-583.


