ANALYSIS OF TEMPORAL AND SPATIAL EXPRESSION OF DROSOPHILA EMBRYONIC SMALL RNAS BY DEEP-SEQUENCING METHOD

A Thesis Submitted to
the Graduate School of Engineering and Sciences of
İzmir Institute of Technology
in Partial Fulfillment of the Requirements for the Degree of

MASTER OF SCIENCE

in Molecular Biology and Genetics

by

Mehmet İlyas COŞACAK

July 2012

İZMİR

We approve the thesis of **Mehmet İlyas COŞACAK**

Examining Committee Members:

_____

Assist. Prof. Dr. Bünyamin AKGÜL
Department of Molecular Biology and Genetics
İzmir Institute of Technology

_____

Prof. Dr. Şermin GENÇ
Institute of Health Sciences
Dokuz Eylül University

_____

Assist. Prof. Dr. Ayten NALBANT
Department of Molecular Biology and Genetics
İzmir Institute of Technology

06 August 2012

_____

Assist. Prof. Dr. Bünyamin AKGÜL
Supervisor, Department of Molecular
Biology and Genetics
İzmir Institute of Technology

_____

Assoc. Prof. Dr. Ahmet KOÇ
Head of Department of Molecular
Biology and Genetics

_____

Prof. Dr. R. Tuğrul SENGER
Dean of the Graduate School of
Engineering and Sciences

ACKNOWLEDGEMENTS

First of all, I would like to thank my supervisor Assist. Prof. Dr. Bünyamin Akgül for his guidance, understanding, encouragement and excellent support during my undergraduate and graduate studies.

Also I would like to express my grateful thanks to my committee members, Assist. Prof. Dr. Ayten Nalbant and Prof. Dr. Şermin Genç for their suggestions and contributions.

I would like to thank my labmates for their invaluable help and their kindness in the laboratory. I want to thank all one by one, Özge Tüncel, İpek Erdoğan, Tahsin Saygılı, Hatice Yiğit, Çağdaş Göktaş, Semih Can Akıncılar, Cenk Dağlıoğlu, Selçuk Özdemir, Hamid Syed, İbrahim Çelik, Saniye Elvan Başkurt, Caner Akıl and Özgür Öksüz.

Finally, I gratefully thank my family for all their motivation, encouragement and support throughout not only this study but also all through my life.

ABSTRACT

ANALYSIS OF TEMPORAL AND SPATIAL EXPRESSION OF DROSOPHILA EMBRYONIC SMALL RNAS BY DEEP-SEQUENCING METHOD

The world of small RNAs is expanding and new types of small RNAs are being identified. By using deep-sequencing techniques in addition to most abundant small RNAs; miRNA and siRNA, piRNA and tRFs were further characterized and shown to be functional. The global behavior of small RNAs during MZT and their location in the cytoplasmic complexes has not been shown. By combining polysomal fractionation and deep-sequencing technique as well as the highly regulated developmental stages in *Drosophila* we have shown that the temporal and spatial expression of small RNAs changes during MZT. We have shown that each small RNA group has unique behaviour in cytoplasm and is enriched in specific polysomal fractions which shows that their local function in cytoplasmic complexes is mainly translational machinery.

ÖZET

DROZOFİLA EMBRYONİK KÜÇÜK RNA'LARININ ZAMANA VE KONUMA BAĞLI İFADELERİNİN DERİN SEKANSLAMA METODU İLE ANALİZİ

Küçük RNA'larının dünyası gittikçe büyüyor ve yeni küçük RNA türleri keşfedilmektedir. Derin sekanslama tekniği kullanarak en çok bulunan küçük RNA'lar olan miRNA ve siRNA'ya ek olarak piRNA ve tRF'ler de keşfedilmiş ve fonksiyonel oldukları gösterilmiştir. Küçük RNA'ların genel olarak MZT sırasında nasıl davrandıkarı ve sitoplazmik komplekslerdeki yerleri daha gösterilmemiştir. Polizomal fraksiyonasyonu ve derin sekanslama tekniklerini birleştirerek ve bunlara ek olarak yüksek derecede kontrol edilen Drozofila gelişim aşamalarını kullanarak küçük RNA'ların MZT sırasında zamana bağlı yerlerinin değitiştiğini gösterdik. Her Küçük RNA grubunun sitoplazmada unik davranışlara sahip olduğunu ve spesifik polizomal fraksiyonlarda yer aldıklarını buda bunların sitoplazmadaki yerlerinin ve fonksiyonlarının daha çok translasyon bölümü olduğunu gösterir.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

AGO.....Argonaute
Aub.....Aburgine
DGCR8.....Di George Critical Syndrome Region 8
dsRNA.....double stranded RNA
Endo-siRNA.....Endogenous small interfering RNA
MBT.....Mid-Blastula Transition
miRNA.....microRNA
MZT.....Maternal-to-Zygotic Transition
piRNA.....Piwi Interacting RNAs
Pre-miRNA.....Precursor MicroRNA
Pri-miRNA.....Primary MicroRNA
RdRNA Pol.....RNA dependent RNA Polymerase
RISC.....RNA Induced Silencing Complex
RPM.....Read Per Million
siRNA.....small interfering RNA
ssRNA.....single stranded RNA
tRFs.....tRNA derived RNA Fragments
TE.....Transposable Elements
UTR.....UnTranslated Region

CHAPTER 1

INTRODUCTION

1.1. miRNA

microRNAs (miRNAs) are 20-23 nucleotide (nt) in length and their biogenesis and function are well characterised in many organsims. Post-transcriptional gene regulation by miRNA is one of the most important steps that plays role in many cellular processes including cell fate determination, apoptosis, metabolism (Bartel, *et al.* 2004). To date, many small RNAs that play role in transcriptional and post-transcriptional regulations have been identified in various organsims (for review; Aalto*, et al.* 2012). In *Drosophila*, the well common known small RNAs playing role in gene silencing are miRNAs, siRNA and piRNAs (Okamura, *et al.* 2008; Siomi, *et al.* 2009; Ghildiyal, *et al.* 2009). These small RNAs have similar features in their length, biogenesis and mode of action. Thus, these small RNAs has certain length of 20-30 nt and processed from longer transcript for their maturation by RNase III enzymes at multiple, sequential steps (Kim, *et al.* 2009). Additionally, these small RNAs associate with Argonautes (AGO) proteins and loaded into protein complexes named RNA-Induced Silencing Complex (RISC) and play role in gene silencing (Okamura, *et al.* 2008; Hutvagner, *et al.* 2008; Siomi, *et al.* 2009; Ghildiyal, *et al.* 2009; Kim, *et al.* 2009). Eventhough their processing and function have not been well characterised in *Drosophila*, tRNA-derived fragments (tRF) have also been shown to play role in gene silencing and cell proliferation in mammals (Lee, *et al.* 2009; Cole, *et al.* 2009; Haussecker, *et al.* 2010).

Since the discovery of the first miRNA, *lin-4*, in 1993 (Lee, *et al.* 1993) and the RNA interference (RNAi) in 1998 (Fire, *et al.* 1998), the role of small non-coding RNAs has been expanding from transcriptional to post-trancriptional regulation**.** It has been estimated that 30-50% of mRNAs are regulated by miRNAs (John, *et al*., 2004; Krek, *et al*., 2005; Lewis, *et al*., 2005; Brennecke, *et al*., 2005). miRNAs are 20-23 nt the most well studied small RNA group, whose biogenesis and function are well caracterised. miRNAs are transcribed from miRNA genes or miRNA clusters by RNA Pol II/III or from intron of protein coding genes after splicing. The first trancribed

primary miRNA (pri-miRNA) transcript has 5'-cap and poly(A) tail and folds to form stem-loop hairpin structures. In nucleus, the microprocessing complex (Drosha/Pasha with protein partner DGCR8) recognizes the hairpin structure and processes the primary transcript into 70 nt hairpin struture releasing 3'-2 nt overhang that is recognised by exporting proteins (exportin-5) and transported to the cytoplasm. The pre-miRNA is converted to a duplex by dicer-depended or indepedent canonical miRNA pathway. One of the strands from the dublex is selected to be loaded into the RISC complex containing Ago and auxiliary proteins (Lee, *et al.* 2003; Cai *et al.* 2004; Borchert, *et al.* 2006; Ruby *et al.* 2007; Okamura, *et al.* 2007; Berezikov, *et al.* 2007). The miRNA directs the RISC complex to their target and depending on the complementarity between the miRNA and its target, the target can be repressed or degraded. (for review, Kim V. N., etal, 2009; Huntzinger, E., etal, 2011; Djuranovic, S., etal, 2011). In addition to translation repression, it has been shown that the repressed mRNAs are degraded as well, which contradict that partially complementarity repress translation (Guo, *et al.* 2010). miRNAs can repress translation at multiple steps; (1) translation initiation (Pillai, *et al.* 2005; Humphreys, *et al.* 2005; Kirikidou, *et al.* 2007), (2) preventing 60S joining (Chendrimada, *et al.* 2007; Wang, *et al.* 2008), (3) post-initiation repression that includes premature ribosome "drop-off", co-translational nascent polypeptide degradation and impaired elongation (Petersen, *et al.* 2006; Nottrott, *et al.* 2006; Gu, *et al.* 2009).

1.2. siRNA

small interfering RNAs (siRNA) are another class of small RNAs that are derived from experimentally introduced dsRNA or endogeously expressed mRNA and transposons. The exogenous siRNA are processed from dsRNA from viral RNAs or experimentally introduced dsRNA (Fire, *et al.* 1998; Timons and Fire, 1998; Elbashir, *et al.* 2001; Wilkins, *et al.* 2005). In *Drosophila*, the dsRNA is processed by Dicer into ~21 nt dsRNA with such complementarity that leaves 3'-2 nt overhang, and one of these strands is loaded into RISC complex (Zamore, *et al.* 2000; Elbashir, *et al.* 2001a; Elbashir, *et al.* 2001b). While mammals and *C. elegans* have one Dicer and Ago proteins that play role in both miRNA and siRNA pathways, flies have two Dicers (Dicer-1, -2) and Argonautes (AGO-1, -2) that play role in miRNA or siRNA. Because

flies are highly exposed to viral infections, Dicer-2/AGO-2 are specialized to interact with siRNA while Dicer-1/AGO-1 interact with miRNA (Hutvágner, *et al.* 2001; Grishok, *et al.* 2001; Ketting, *et al.* 2001; Lee, *et al.* 2004). Additionally, plant and *C. elegans* has evolved to convert single-stranded RNA (ssRNA) into double stranded RNA (dsRNA) by RNA-depended RNA Polymerase (RdRNA Pol), while RdRNA Pol is not found in flies and mammals. The RdRNA Pol play role in endogenous or viral single stranded RNA degradation by converting ssRNA to dsRNA and subsequently its degradation by siRNA pathway (Voinnet, *et al.* 2008; Gazzani, *et al.* 2004). The mammals evolved to deal with dsRNA and viral infection by using protein based immune system (Williams, *et al.* 1999; Kunzi, *et al.* 2005; Vilcek, *et al.* 2006). Because mammals does not have interferon response in their germ cells (e.g., oocyte), it has been shown that endogenous siRNAs (endo-siRNA) are generated in the oocyte of mouse (Watanabe, *et al.* 2008).

1.3. piRNA

Piwi-interacting RNAs (piRNA) are 23-30 nt small RNAs derived from transposon and mainly play role in transposon silencing in germ cells (Malone CD, *et al.* 2009; Siomi, *et al.* 2011) as well as clearance of maternal mRNA during maternal-to-zygotic transition (MZT) (Rouget, *et al.* 2010). piRNAs in *Drosophila* differs from siRNA and miRNAs as they interact with PIWI familiy proteins and processed independent of Dicer. piRNAs are found in plants, *C. elegans*, mammals and flies. The biogenesis of piRNAs is well characterised in flies. piRNAs interact with PIWI family proteins. There are three PIWI proteins in *Drosophila*, AGO3, Auburgine and PIWI which interact with a different subset of small RNAs that are classified as piRNAs (Saito, *et al.* 2006; Nishida, *et al.* 2007; Gunawardane, *et al.* 2007; Brennecke, et al, 2007). AGO3 and Aub are mainly cytoplasmic in oocyte of *Drosophila* and AGO3 is more dense in nuage bodies than Aub while PIWI is nuclear in both somatic (such as follicle cells) and ovary (Brennecke, et al, 2007). Deep-sequencing studies have shown that Aub/PIWI interacts with piRNA that has 5' U preference (Brennecke, *et al.* 2007). Bioinformatic analyses of small RNA interacting with PIWI has shown that, there is 10 nt complementarity at the 5' end of piRNAs and the 5' U at position 1 is complement to A at position 10 on the complement piRNA (Brennecke, *et al.* 2007). This leads to the

ping pong cycle hypothesis, as a piRNA binds to the target sequence and generates a new piRNA. In the ping pong cycle, the primary transcipt from Transposable Elements (TEs) is single stranded and does not generate dsRNA, the slicer acticvity of AGO3 and Aub leads to the degradation of TE transcript in nuage bodies, a perinuclear cytoplasmic structure, that generates new piRNAs (Gunawardane, *et al.* 2007; Brennecke, *et al.* 2007; Li, *et al.* 2009; Malone, *et al.* 2009). The piRNAs that initiates the ping pong cycle are believed to be inherited maternally and called as primary piRNAs (Brennecke, *et al.* 2008) and the newly generated piRNAs in nuage bodies are called secondary piRNAs. In addition to ping pong cycle another pathway has been identified in *Drosophila* which is named as primary piRNA pathway. The primary piRNA are exclusively derived from the flamenco locus around ~180 kb genomic sequence that generates also a trasncript in that size (Brennecke, *et al.* 2007). During the primary piRNA, the ping pong cycle is not seen and slicer activity of PIWI is not required. The primary pathway takes play in Yb body, non-membranus cytoplasmic structure, and three proteins involve; Armitage (Armi), the Tudor Domain containing RNA helicase Yb and Zucchini (Malone, *et al.* 2009; Saito, *et al.* 2009; Olivieri, *et al.* 2010). The piRNAs processed in the cytoplasmic structure Yb, are loaded to PIWI and PIWI-piRNA complexes are transported to the nucleus (Qi, *et al.* 2011).

1.4. tRF

Transfer RNA (tRNA) are ~70 nt in length and one of the fundemantal component of translation machinary. It has been shown that under stress conditions various fragments (17-55 nt) are derived from tRNAs (Thompson, *et al*. 2008; Hsieh, *et al*. 2009; Cole, *et al*. 2009). In many unicellular organism, it has been shown that under starvation and oxidative stress, tRNA fragments in length of 30-50 bp are generated (Haiser, *et al*. 2008; Jochl, *et al*. 2008; Lee, *et al*. 2005; Thompson, *et al*. 2008; Thompson, *et al*. 2009). In *Drosophil*a, the presence ~35 nt fragments derived from the 5' end of tRNAs has been cloned from mRNA fractions (Hatice Yiğit, Unpublished Data). During developmental processes of *Drosophila melanogaster*, tRNA fragments of 16-26 nt has been identified, no processing positions specified (Aravin, *et al.* 2003). By high-throughput sequencing, tRNA derived fragments (tRF) were shown to be abundant as musch as miRNAs and tRF-1001 (19 nt) was shown to be essential for cell

proliferation (Lee, *et al.* 2009). tRF are shown to be processed by Dicer (Cole, *et al*. 2009). From a tRF gene region various fragments can be generated as (1) after transcription from 5' leader and 3' end trailer, (2) 5' and 3' end of mature miRNAs (with 3' CCA) (for review; Haussecker *et al*. 2010; Sobala, et al., 2011). The function of tRF in *Drosophila* has not been shown and remains to be elucidated. Even though there are studies showing that various tRFs may interact with different Ago proteins (Haussecker, *et al*. 2010) but the role of tRFs in development, apoptosis and many cellular processes remains to be elucidated.

1.5. Smaug, miRNA, siRNA and piRNA in Early Embryo Development

The early development after fertilization involves several important regulatory steps before and after the genome of zygote is activated. In almost all organisms, the majority of embryonic components are obatained maternally through oocyte cytoplasm. The majority of embryonic mRNAs maternally obtained are stored in oocyte during maturation or after fertilization (Aoki, *et al.* 1997; Bachvarova, *et al.* 1985; Bachvarova, *et al.* 1992; Braude, *et al.* 1988; Crosby, *et al.* 1988; Jeanblanc, *et al.* 2008; Kopecny, *et al.* 1989; Leandri, *et al.* 2009). The stored mRNAs in the oocyte, are used for protein sythesis as the transcription silenced when the oocyte has fully matured (Miyara, *et al.* 2003; Fulka, *et al.* 2009; De La Fuente, *et al.* 2001; Fair, *et al.* 1995; Fair, *et al.* 1996). The maternally inherited mRNAs are removed when the zygote genome is activated for trancription (Schier, 2007; Hyttel, *et al.* 1997; Memili, *et al.* 1998). The duration of genome activation may takes hours (e.g., *Drosophila*) or days (e.g., human).

For the maternal-to-zygote transition (MZT) two processes are important; (1) the removal of maternal transcripts and (2) the activation of the zygotic genome. In *Drosophila*, the activation of zygotic genome is required for the blastoderm cellularisation (Merrill, *et al.* 1988; Poulson, *et al.* 1937; Wieschaus, *et al.* 1988) and this takes around 2.5 to 3.0 hour post-fertilisation. The completion of MZT is followed by mid-blastula transition (MBT). Recent studies show the important role of timing and regulation of MZT by the mother as around 50% of all protein coding genes that can be encoded by the genome are loaded to the oocyte (De Renzis *et al*., 2007; Tadros *et al*., 2007). Each study has shown that (1) in acitivated but unfertilised egg which is transcriptionally silent, %20 of maternal mRNA are degraded (Tadros *et al*., 2007), (2)

the zygotic trasncription also leads to the degradation of maternal mRNAs and transcription of newly expressed mRNAs represent around 18% of transcipts (De Renzis *et al*., 2007). These studies have shown that after the maternal mRNA clearance, the zygotic mRNAs are also trancribed to replace with the maternal mRNAs that are degraded, consistant with previous studies (Heifetz, *et al.* 1999; Bashirullah, *et al.* 2001; Tadros,*et al.* 2003). These studies identified an upstream consensus heptamer sequence (TAGteam sites) that may be required for the transcriptional activation (Liang, *et al.* 2008).

The degradation mechanisms of mRNAs in early embryo depend on different mechansims that destabilise the mRNAs such as Smaug and small RNAs, thus mainly translational repression and coupled by degradation (Jacobson, *et al.* 1996). The spatial expresion of mRNA in the embryo depends on "degradation/protection" mechanism as an mRNA is degraded in some cells, it may be protected in another cell. In *Drosophila* embryo, the spatio/temporal expression of nanos and Hsp83 is determined by the mechanism described above. The nanos gradient in *Drosophila* embryo depends on the translational regulation (Dahanukar, *et al.* 1996), the localisation of Hsp83 depends on Smaug (Semotok, *et al.* 2005) by translational regulation as well, which is shown that Hsp83 has binding site for Smaug (Semtook, *et al.* 2008). The role of Smaug in global gene expression in MZT was shown (Tadros, *et al.* 2007). These studies show that the destabilisation of mRNAs during MZT is regulated by translational regulation and destabilisation of mRNAs by decapping of Poly(A).

In many organisms (e.g., *Drosophila*, Mouse, Xenopus and Zebrafish), it has been shown that maternal mRNA degradation depends on translation regulation by miRNAs. The first studies from Zebrafish have shown that miR-430 is involved in the maternal mRNA degradation and clearance of these mRNAs during MZT (Giraldez, *et al.* 2006; Kedde, *et al.* 2007). The same miRNA family was shown to act similarly in Xenopus to repress and destabilise the maternal mRNAs during MBT (Lund, *et al.* 2009). In *C. elegans*, miR-35-42 family (maternal), miR-51-56 family (zygotic) and miR-58/80-82 (bantam) family (zygotic) has been shown to have the same role as miRNAs mentioned in destablisation of maternal mRNAs (Wu, *et al.* 2010). The same miRNA family (miR-290) has been predicted to exist in Mouse and enriched in early embryo but not yet characterised experimentally to degrade maternal mRNAs (Zeng, *et al.* 2005; Tang, *et al.* 2005).

In *Drosophila*, the miRNA cluster (miR-6-1, miR-6-2, miR-6-3, miR-5, miR-4, miR-286, miR-3 and miR-309) modulates clearance of maternal mRNAs (Bushati, *et al.* 2008; Benoit, *et al.* 2009). Thus, there is a reciprocal change in the expression level of miRNAs and their targets. Additionally the expression of these miRNAs depends on Smaug. Smaug acts in two ways, (1) binding to the target mRNAs 3' UTR and destabilise, (2) activating the transcription of miR-309 cluster in zygote.

In addition to miRNAs, piRNAs (generated from roo and 412 retrotranposon) act together with Smaug to destablize maternal mRNAs (nanos) in *Drosophila* (Rouget, *et al.* 2010). Additionally, in mouse spermatoids MIWI are shown to interact with polysomes (Grivna, *et al.* 2006). In another study, it has been shown that MILI is required for translation in mouse spermatocytes (Unhavaithaya, *et al.* 2009). In addition to the role of piRNAs and miRNAs in maternal mRNA clearance, it has been shown that siRNAs also act in the germ cells in mouse and *Drosophila* (Watanabe, *et al.* 2008; Tchurikov, *et al.* 2011).

The role of small RNAs in MZT seems to be evolutionary conserved and acts in insects, *C. elegans*, Zebrafish, Xenopus and mouse. In studies discussed so far, there is information about the role of small RNAs in translational repression of maternal mRNAs, but to what extent these small RNAs are found in translational machinary has not been studied genome wide. Here, we aimed to analyse the temporal expression of small RNA in early *Drosophila* embryo and their localisation in translational machinary in the cytoplasmic RNAs.

1.6. Aim of The Study

In this study, (1) it was aimed to analyse deep-sequencing data from fractionated and unfractionated of *Drosophila melanogaster* embryo, (2) comparing the expression level of various small RNAs from miRNA, piRNA, siRNA and tRNA, (3) localisation of small RNAs in translational machinary.

CHAPTER 2

MATERIAL AND METHOD

In this study, the bioinformatic analyses of deep-sequencing data and supporting experiment are given. The experiements including embryo collection, RNA isolation and RNA quality control are described elsewhere (Hatice Yiğit, Master Thesis) but summarised here.

2.1. Embryo Collection and Sucrose Density Gradient

2.1.1. Embryo Collection

The Drosophila Oregon R strain were grown at 25 ºC on standard agar, corn meal, sucrose, yeast medium prepared. 30 g agar and 60 g yeast extract were dissolved and boiled in 1,8 L dH_2O and another mix of 110 gram corn meal and 96 gram sucrose were added to the mixture. Then, the total mixture boiled for extra 20 min and after cooling 25 ml propionic acid (Merc) and 25 ml nipagyn were added to the mixture.

When *Drosophila* population reached sufficient density, the whole population was transferred to a common and large cage. Embryo collection plates were prepared as follows. 22.5 g agar was boiled in 700 mL dH2O, 94 g sucrose was dissolved in 150 mL dH2O and mixed with 330 mL fruit juice. 0-1h embryos was collected by placing embryo collection plates in a large cage and waiting for an hour. 8 hour embryos were collected by incubating the plate in an incubator at 25ºC for 7 hours after the plate incubated for one hour in the large *Drosophila* case. Embryos were immediately washed with 0,7% NaCl and 0.1% Triton-X and then stored at -80 ºC until use.

2.1.2. Polysome Profile of 0-1 Hour and 8 Hour Embryos

By Polysome profiling, cytoplasmic mRNP complexes can be fractionated based on their translational status such as mRNP, 60S, monosome and polysome. There is no

ribosomal subunits. 40S or 60S status represent that mRNPs bounded with small ribosomal subunits in the mRNP fractions. At monosome or 80S levels, messenger RNAs are bounded with only one full ribosome. Polysomes contain at least more than one ribosomes. Polysome profiles of 0-1 and 8 h embryos were obtained successfully with an increasing size and volume (Figure 2.1).

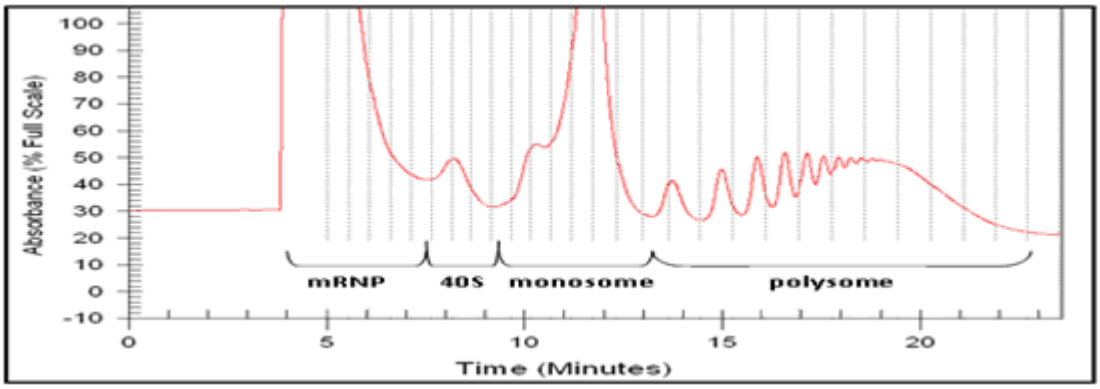

Figure 2.1. Polysome profilings of 1 and 8 hour embryos. The lysate of embryos was centrifuged and fractionated by Density Sucrose Gradient Fractionation (DSGF) system. The first highest peak represents mRNP status and second one represents 60S subunits of ribosome and in the of the profile, the highest peak represents monosome. Polysome is heaviest fraction relative to monosome, 60S and mRNP. Each increasing peak presents the ribosome number on mRNA and polysome volume (Source: Çağdaş Göktaş, Master Thesis).

2.1.3. Total RNA Isolation

Total RNA was isolated by miRVana kit accroding to the Kit procedure (Hatice Yiğit, Master Thesis)

2.2. Deep-Sequencing of Small RNAs

1 µg of total RNA from three replicates were mixed and sequenced by Fasteris Illumunia (Switzerland). cDNA clones in length of 15-30 bp were sequenced.

2.3. Data Analyses of Deep-sequencing Data

2.3.1 Barcode Selection and 3' Adapter Trimming

Deep-sequencing yielded fragments of 36 bp in length. Because we used multiplexed sequencing, sequences were splitted to their corresponding samples according to their barcode sequences. In order to choose the barcode, we allowed single mismatch. As a result, single mismatch caused in some sample the same barcode. We removed these kinds of sequences from further analyses. Only sequences without any ambiguity were used for further analyses.

After barcode selection, the 3' adapter sequence were removed. The 3' adaptor sequences were trimmed from the raw reads in 4 steps by using 3' Adapter sequence (5'-ATCTCGTATGCCGTCTTCTGCTTGT-3'). (1) At the end of the raw reads, the full adapter sequence searched, which permits to identify inserts equal to 7 nt. (2) If no adapter sequence was found, the last base of the 3' adapter was removed and the sequence was searched at the end of the reads. The minimum adapter size of 4 nt permits identifying inserts up to 28 nt. (3) Finally, the mismatch of the adapter was searched in the remaining reads. The first 3 bases of the adapter was searched in the reads. (a) If the first 3 bases were found only one time in the full read, then the homology of the aligned nucleotides were calculated. >=75% homology (including the first 3 bases) permits identifying inserts upto 29 nucleotides. (b) If the first three bases found more than one times in the full read, then the homology for each possible adapter were calculated. >=75% homology of only one possible adapter in the full reads permits identifying up to 29 nucleotide inserts. If >=75% homology was found for more than one possible adapter, then the highest homology (the most-left position) in nucleotide number were used to identify the insert. *Because only a single mismatch at position 10 of the adapter generates two possible homology with >=75%, the alignment of such inserts to the model genome shows that the most-left adapter is most-likely the adapter start position.* (4) inserts with poly(A) and/or with more than 1 "N" were also removed from the data. Inserts in length of 15-29 bp were used for further analyses.

After insert selection, all samples were collapsed and only unique sequences were obtained. Then the cloning frequency of each sequence was added to the header

line of the sequence into fasta format. This allows easy manipulation of the data in subsequent analyses.

2.3.2 Alignment of Sequences to the *Drosophila melanogaster* Genome

After converting the sequences into fasta format, the sequences were aligned to the *Drosophila mealnogaster* genome, release dmel-r5.39 (flybase.org). For alignment, mature tRNA sequences were downloaded from flybase and 3'-CCA appended. Then, nexalign program (de Hoon, *et al*. 2009) was used to align all sequences to the genome and mature tRNAs. We used three options for alignment, (1) aligning all the reads for exact match (EMM), (2) using remaining sequences and aligning reads for one error (insertion or deletion or mismatch), (MINDEL) (3) using remaining sequences at each step and aligning for two mismatches (M2M) and then three mismatches (M3M). The remaining sequences were grouped as unmapped (UNM). After genome alignment for each unique reads number of mapping site in the genome and the mapping type were appended to the header line for further data analyses.

2.3.3. Alignment of Sequences to the *Drosophila* Known RNAs

After aligning sequences to the genome, the number of genome mapping were also added to the header line of the fasta sequences. All known RNAs sequences downloaded from flybase (dmel-r5.39) except (i) hairpin and miRNAs sequences were downloaded fom mirbase (www.mirbase.org) (Release 17), (ii) rRNA sequences (5.8S, 18S and 28S) from NCBI (M21017.1), (iii) Repbase collection (Jurka, *et al.* 2000; Jurka, *et al.* 2005) and (iv) piRNA clusters genomic coordinates from a previously published article (Brennecke, *et al.* 2007) and the piRNA clusters from sense strand of genome were extracted by an in-house-algorithm from dmel-r5.39.

The known RNA names were arranged for further data analyses and all reads was aligned to the known RNA for exact match, one error, 2 and 3 mismatches as done for the genome alignment. The reads aligned to each known RNA groups as; (i) align sequences to a known RNA group, (ii) the matched ones are grouped as the knwon RNAs the reads mapped, (iii) aligning remaining sequences to the next knwon RNAs

groups and classifying the matched as the RNA group the reads until all known RNA group were used, (iv) repeat this steps for exact match, then one error and finally for 2 and 3 mismatches. The groups of known RNAs were ordered as rRNA, hairpin, tRNA, miscRNA, ncRNA, transposon, transcript, intron, pseudogene and intergenic region.

In order to analyse the miRNA cluster and Repbase collection another alignment done. For Repbase collection, the transposon mapped sequences were remapped for exact matches to the repbase collection, for piRNA clusters the intergenic region mapped reads and transposon mapped reads were realigned to the piRNA clusters.

2.3.4 Mature miRNA Expression Levels

The mature miRNAs sequences and hairpin sequences were downloaded from mirbase (Release 17). The 5'- seed sequence is important for the miRNA target recognition. Thus, we used the 5'-end of mature miRNAs from mirbase, we extended the mature miRNA sequences 3 bp at the 3' end, as some mature miRNAs has 1,2 and 3 nucleotide appended as the 3' is variable. Then we only used hairpin exact matched sequences to realign mature miRNAs with 3 bp appended (mature_3bp). After aligning sequences to the mature_3bp, each sample in 7-8h embryo was compared with the same sample in 0-1h. Thus, the read per million (RPM) for each mature_3bp, we used the below formula ; (fold = ((7-8h_RPM + 10)/(0-1h_RPM + 10)), then the log2 of fold were calculated by excel. By this way we decrease the fold changes for lower Read Per Million frequencies such as 8 versus 2 will be 4 fold, but if we add 10 this will be around 1 fold as a result not significant change.

2.3.5 tRF Expression Levels

For tRNA derived fragments we only calculated the nucleotide distribution on the mature tRNAs to determine the most abundant fragments and the length distribution of tRFs to identify the average tRF in *Drosophila*. In order to calculate the each nucleotide cloned on the mature tRNA, we converted each read frequency to RPM and increased the frequency of the nucleotide on the mature tRNAs that the read mapped. By this way we calculated the frequency of each nucleotide in 10 samples and find the major processing sites on mature tRNAs for each one and for all tRNA by

superimposing. Also, one of tRNA fragment from the tRNA- Glycine was further confirmed by Northern Blot (Goktas, *et al*., unpublished data).

2.3.6. Transposon Derived Fragments (siRNA and piRNA)

2.3.6.1 Repbase Collection Expression Level

The reads mapped perfectly to the transposon and mapped only one time in the genome were realigned to the Repbase collection and the number of reads in RPM mapped to each transposon were calculated and the log2 ratio were calculated as above. The transposons from flybase which is calculated based on the Repbase were used to classify small RNAs transposon derived ones, while Repbase colection were used for the expression level of transposon in 0-1h versus 7-8h. For Repbase expression level we only used 23-29 nt reads as piRNAs are in these length but siRNAs are 21 nt. Here we are interested in the expression level of piRNAs that are derived from transposons. The siRNA expression level is too low and siRNAs map many sites in the genome. Therefore we excluded in our analyses.

2.3.6.2 Classification of Transposon Derived small RNAs

In *Drosophila* so far there are two small RNA groups that were shown to derive from tranpsons siRNA and piRNA. siRNA are 21 nt in length and piRNAs are 23-30 nt in length. In order to check whether 23-29 nt reads has piRNA features we checked two features as (i) the frequency of reads that has 10 nt at 5'-5', (ii) if a read has 10 nucleotide complementarity the frequency of nucleotide at position 1 and 10 were calculated. Thus, piRNA has 5' U preference and A at position in the complement piRNAs (Brennecke, *et al*, 2007). In order to do such analyses we used all piRNAs that mapped to the transposon in length of 23-29 nt and mapping at most 10 times in the genome (this can be done for others as well but it takes longer time so we ignore them). Then, for each transposon mapped reads (i) the start position for the sense strand were determined, (ii) then end position of all reads in the antisense strand were determined, (iii) the distance between sense/start and antisense/end were calculated, and (iv) if the distance were less than equal to 100, the distance frequency were calculated by the

frequency of the piRNAs. The by using excel, the ferquency versus distance were shown on XY-chart. Also, if the distance is equal to 10 then the frequency of the nucleotide at position 1 and position 10 were calcualted.

2.3.7. piRNA Clusters Analyses

piRNA clusters are the major piRNA processing sites.The piRNA cluster coordinates determined previously (Brennecke, et al. 2007) were used to get genomic sense from flybase (dmel-r5.39) and all reads mapped to the transposon and intergenic regions were realigned to these genomic strands.

CHAPTER 3

RESULTS AND DISCUSSION

3.1. Data Analyses

3.1.1. Barcode Selection and Adapter Trimming

The barcode selection and 3' adapter trimming was performed by an in-house-written algorithm and the summary of the insert selection is given in Table-3.1. The sequences in length of 15-29 bp were chosen for further data analyses. The percentage of inserts with length of 15-29 bp range from 88.8 to 99.5. The remaining reads for Poly_7-8h is the highest with 75%. The minimum adapter size used for adapter sequence is 3 bp. If 2 bp allowed then the remaining sequences are generated from 2S rRNA with more than 95%. Thus, in our-data set we have major 2S rRNA sequences as expected. Because we did not remove 2S rRNA, the enrichment mainly found in Polysomal fractions.

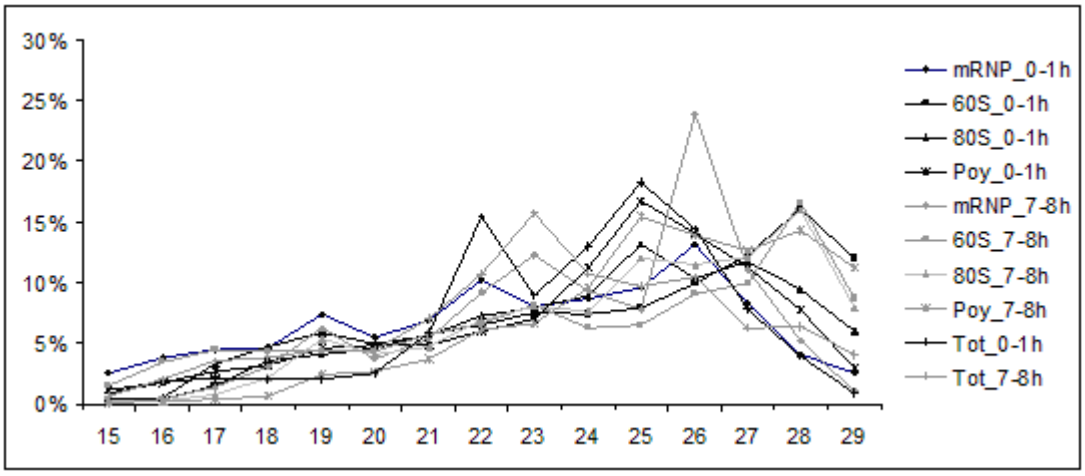

Figure 3.1. The lengh distribution of inserts. The inserts in length of 15-29 bp used and the percentage of reads length calculated in each sample. (0-1h/7-8h: 0-1h/7-8h embryonic RNA; Pol: Polysomal RNA, Tot: Total RNA) .

After the insert selection, the length distribution of inserts in each sample were calculated (Figure-3.1). The relative length distribution of our data is enriched at 22-29 bp. Because we used the whole embryo, we have different cell population (especially, in 7-8h), and the length distribution may have slight changes between samples. miRNAs are 22-23 bp in length in *Drosophila* as a result the enrichment at 22-23 bp is because there are miRNAs at this length that constitute the majority of 22-23 bp reads. The reads longer than 23 bp were found to be mainly derived from transposon as they are classified as piRNAs which are in length of 23-30 with an average lenght of 24-25 bp (Brennecke, et al, 2007). Thus, the relative difference in reads lenght distribution is not because of any experimental error, rather because of the increase or the decrease in the amount functional small RNAs (e.g., miRNA, siRNA).

3.1.2. Alignment of Sequences to the *Drosophila* Genome

After the insert selection, the reads were aligned to the *Drosophila melanogaster* genome (r5.39) (Figure-3.2). After aligning the sequences to the genome, for each insert the number of mapping site in the genome was appended to the header of each insert. By this way, in further data analyses the mapping sites were calculated easily. Also, the percentage of inserts that mapped exactly (EMM), with one error (MINDEL), 2 and 3 mismatches were also determined. 72-82% of inserts mapped perfectly to the genome, while more than 90% of inserts mapped to the genome with exact and single error.

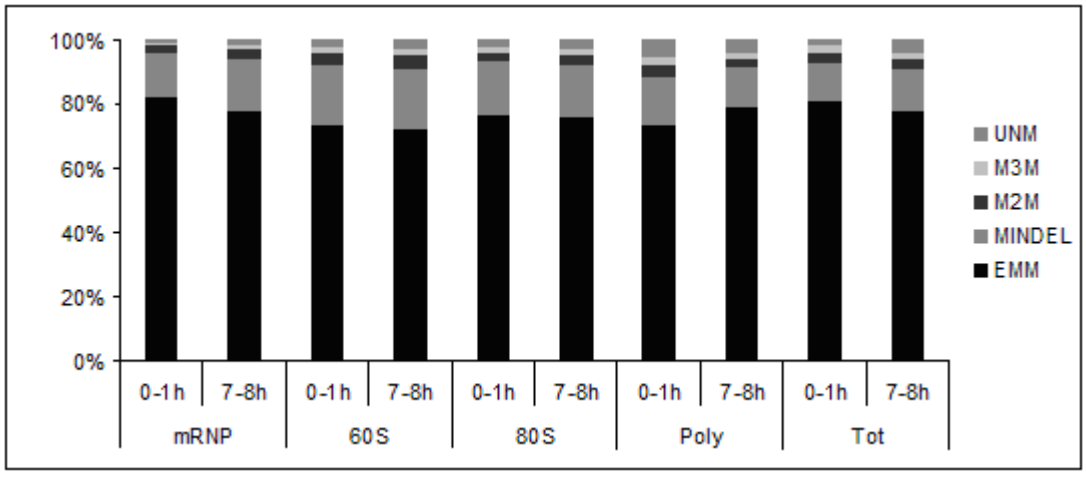

Figure 3.2. Alignment of the sequences to the genome. Inserts were aligned to the dmel-r5.39 genome and the percentage of inserts that mapped to the genome determined.

The alignment of the reads to the genome gave a relatively similar exact match from 72-80%. For our experiment we used *Drosophila melanogaster* Oregon R strain, while the *Drosophila* genome is sequenced from the "y1; cn1 bw1 sp1" isogenic strain (Adams, et al, 2000). The effect of strain differences is discussed elsewhere (Brennecke, et al, 2007, Supplemental Data). The single error that macth to the miRNAs seems to be an SNP in *Drosophila* genome as we manulay sequenced Oregon R, w1118, S2 and Ago^{114} strains and we showed for a single miRNA that the single error is not because of error in the sequencing but because of SNP in *Drosophila* and cell lines (Data not shown). Overall this indicates the quality of our data that can be used for data analyses, as the SNP is common. In order to be consistent in data samples, we only used sequences that mapped exactly to the known RNAs, as thre are sequences that unmapped to the genome (exon-exon junction may mapped exactly, to the transcripts CDS).

3.1.3. Analysis of Known RNAs

3.1.3.1 Alignment of Sequences to the Known RNAs

After aligning sequences to the dmel-r5.39 and determining number of candidate mapping sites, the inserts were aligned to the known RNA database. The nexalign program was used to align sequences with sense or antisense sequences to the known RNAs. A pre-alignemnt to the all known RNAs, directed us to align sequences in an order described in method. The alignment results are given in Figure-3.3. 71-82% of reads mapped exactly to the known RNAs. Thus, the majority of small RNAs mapped exactly to the known RNAs and more than 90% of sequences mapped with exact and single error. Except for rRNA, hairpin sequences we used all other sequences from the flybase as a result we expected the same mapping percentage as genome alignment.

The alignment of the sequences gave novel data about the intracellular localisation of small RNAs. While analysing such data, researchers generally used the most abundant small RNA group while indetifying novel RNA group (Lee, et al, 2009). Thus, while analysing a data if a group of small RNA is enriched in a sample, this leads to the assumption that the RNA group has functional role. So far, the total RNAs are used for deep-sequencing or immunoprecipitation of protein complexes has been done

Table 3.1. Barcode Selection and 3'-adapter trimming. The table shows the number of reads assigned to each sample after barcode selection and given as "reads". After 3'-adaper trimming the percentage of reads in all reads is given as "%all" and the percentage of reads length in total inserts is given as "%ins".

	mRNP_0-1h			**60S_0-1h**			**80S_0-1h**			**Poly_0-1h**			**Tot_0-1h**		
	reads	%al	%in	reads	%al	%in	reads	%al	%in	reads	%al	%in	reads	%al	%ins
0	651	0	0,0	19352	1,1	1,3	3952	0,1	0,3	725	0,1	0,2	71257	2,2	2,4
1_14	33275	1,7	1,8	44531	2,5	3,0	6340	0,1	0,5	18888	2	4,0	255272	7,8	8,7
15_29	179366	94	98,1	143708	82	95,7	121894	24,	99,2	45000	48	95,8	259416	79	88,8
Remainin	81981	4,3		246708	14		369988	75,		47712	50		370156	11	
All_Reads	190957			174767			492912			94675			329084		
	mRNP_7-8h			**60S_7-8h**			**80S_7-8h**			**Poly_7-8h**			**Tot_7-h**		
	reads	%al	%in	reads	%al	%in	reads	%al	%in	reads	%al	%in	reads	%al	%ins
0	49	0	0,0	33756	1,7	1,8	6499	0,2	0,4	330	0	0,2	14156	1,3	2,5
1_14	3781	0,5	0,5	47615	2,3	2,5	39238	1,0	2,2	740	0,1	0,4	7397	0,7	1,3
15_29	762090	94	99,5	179783	88	95,7	177803	45,	97,5	17427	25	99,4	547815	52	96,2
Remainin	41668	5,2		154570	7,6		204818	52,		51491	75		487770	46	
All_Reads	807588			203377			387195			69026			105713		

that allows the researcher to enrich a small RNA group that is found in a particular protein complexes (Brennecke, et al, 2007). In our data we used both total and cytoplasmic RNA and enriched small RNAs according to their translational status.

As can be seen in Figure-3.3, for each fraction specific small RNA groups are enriched. For mRNP fraction the majority of small RNAs matched to the tRNAs, with 30% and 37% in 0-1h and 7-8h, respectively. In Polysomal fraction the majority of small RNAs are derived from rRNA and transposon. Because polysomal fraction are enriched with polysome, it is possible to have such small RNAs, either degradation products or biogenesis (not known). While we compared the total RNAs, the majority of 0-1h embryonic total and polysomal RNAs are derived from transposons compared with 7-8h. This indicates the advantage of our polysomal techniques to enrich similar size complexes either translation machinary or complexes in that sizes.

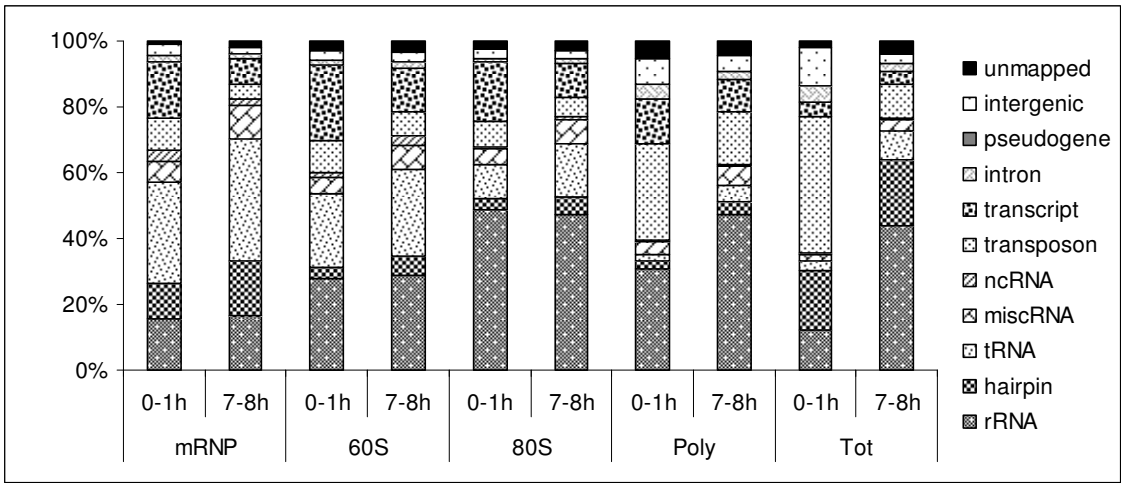

Figure 3.3. Alignment of the sequences to the known RNAs. The sequences aligned to the order given in legends from down to top, and this figure shows the alignment for the exact, mindel, 2 and 3 mismatch to the known RNAs. If a sequence matched to a known RNA group, we assumed the sequence to be derived from that groups, as the sequence may also mapped to other known RNAs.

The relative percentage of rRNA derived fragments in lower complexes is equal, while there is a slight increase in 7-8h polysomal fractions and total RNAs. Most probably there is a direct degradation or biogenesis of rRNAs in later stage of development. As the increase in small RNAs derived or degraded from rRNAs increased in 7-8h and this increase is found in polysomal fraction. Additionally, the

transcript derived fragments are highly seen in early embryo, which probably indicates either clerance of maternal mRNAs or biogenesis of new small RNAs. In additon to known RNAs also small RNAs mapped to the intron and intergenic regions. All these small RNAs have features of piRNA, as their length are enriched in 23-29 bp (discussed below).

3.1.3.2. miRNA Expression Analyses

In order to calculate the expression level of miRNAs, we firstly aligned sequences to the hairpin with perfect matches. Considering the top 10 highest hairpin sequences in each sample, 18 hairpin sequences consists of 82.52-93.6% of hairpin exact matched sequences. Thus, for some miRNAs the expression level is highly increased. For instance, dme-bantam has 30% in total 7-8h while dme-miR-1 has 43% of 0-1h. The enrichment of dme-bantam in 60S and 80S fraction is similar in 0-1h and 7-8h. Thus, eventhough their fold changes are different their abundance in fractions are, similar, which show that the position of the miRNAs does not change but the amount changes. Because we do not have a direct normalisation factor, the abundance can not be compared with other fractions.

The mature miRNA expression level was determined by aligning sequences to the miRNA mature core sequence (deposited in mirbase) and 3 bp appended to the 3'. Because the 5' seed sequence play important role in miRNA/target recognition, the 5' end deposited in mirbase was used. As can be seen in Table-3.2, the percentage of sequences that mapped to the hairpin or mature_3bp appened are almost the same, which indicates the majority of processing is at mature_sequences with 3bp appended.

In order to calculate the fold changes, the number of reads in read per million mapped to the mature_3bp appended were increased by 10 and the log2 ratio of 7-8h frequency over 0-1h were calcualted. Then by using cluster 3.0, the samples were clustered and visualised by java Tree View (http://jtreeview.sourceforge.net/).

In order to analyse the expression level of miRNAs, we chose the ones with higher frequency. Thus, a miRNA with at least 50 RPM in at least one sample were used for further analyses.As the remaining were assumed to expressed with lower frequency (Figure-3.4).

In total 426 mature miRNAs sequences deposited in mirbase were used for analyses. The sequences are annotated as 3p or 5p depending on the location of the mature sequence on the hairpin. But in Release 17, the mature and star form are not given. Of 426 mature sequences, 256 (60%) mature miRNAs has been detected with perfect matches. We only analysed the expression level of mature miRNAs whose frequency in RPM is above 50 in at least sample. 162 miRNAs (63,28%) of 256 has expression level below 50 RPM in all samples, while 94 (36,72%) of 256 miRNAs has 50 RPM at least in one sample. Thus 94 of 426 (22,07%) of mature miRNAs were represented in our data. As we analysed the data elsewhere (Malone, *et al*, 2007), we found the same results (not shown).

We calculated the fold change in all samples. Because we do not have a normalisation factor we only compared each fraction of 0-1h/7-8h with each other. By this way, we expected to observe changes in the fractions. To our expectation, we observed changes in different fractions. To understand each miRNA behaviour we divided each samples into three different categories, changes (increase or decrease) or no change. As a result we had 4 major groups each has its unique subgroups. The first attempt to group miRNAs was to compare their expression level in total RNAs, as we have 3 groups with increase, decrease or no change. The 4th group behaves differently, as the expression level in total RNAs is contrast to the fractions.

The first group of small RNAs we observed is the one highly expressed in early embryo. Of 94 miRNAs analysed, 29 miRNAs are highly expressed in 0-1h while their expression level decrease toward 7-8h. The majority of these changes is found in mRNP with 20 miRNAs, while this changes decreases towards polysomes (9,6,3, respectively, in 60S, 80S and polysome). This indicates the majority of small RNAs, acts at mRNP complexes in early embryo compared with 7-8h. Moreover the changes in total RNA may not be significant found in fractions for 7 miRNAs (Table-3.3, Figure-3.4). The second group of miRNAs are highly expressed in 7-8h embryo (41 of 94 miRNAs). While 17 miRNAs are highly expressed in mRNP, the majority of miRNAs are highly expressed in 60S. Moreover the fold changes in higher complexes increased (17,25,15,14 in mRNP, 60S, 80S and Polysome). This indicates that in addition to repressed translation at the intiation step, these miRNAs also repress translation at the elongation step of translation. Similar to 0-1h, the increase in total RNAs is not found in fractions for 9 miRNAs (Table-3.3, Figure-3.4). The third group of miRNAs expression

level does not change in total RNAs, while increasing in fractions. Thus, in 7-8h embryo an increase was found in heavier complexes,while in lower complexes in 0-1h embryo. Also some miRNAs are equally expressed in total RNAs as well as in fractions (Table-3.3, Figure-3.4). The fourth group of miRNAs acts differently, as their expression in fraction is opposite in total RNAs. For that miRNAs, we propose two main mechanisms; (1) the elimination of small RNAs at the isolation, if they are in higher complexes such as nucleus or unknown structures , (2) because we are interested in cytoplasmic miRNAs (Table-3.3, Figure-3.4).

3.1.4. Transposon derived small RNAs.

The majoirty of 0-1h embryonic total RNAs matches to the transposon either in sense or antisense orientation. This indicates the function of these small RNAs in early development. So far, it has been shown that piRNAs deposited maternally or generated in the embryo play important role in early development. Also it has been shown that there is a maternal effect in the production of piRNAs derived from transposon. In mouse, MILI/MIWI are associated with polysomal fractions (Grivna, *et al.* 2006; Unhavaithaya, *et al.* 2009), and maternal mRNAs are degraded from the embryo by piRNAs (Rouget, *et al.* 2010). Thus, piRNA acts in early development and germ cells protection as well as has been shown to regulate non-transposon mRNAs as well. The presence of MIWI/MILI in polysome indicates the possible role of piRNAs in translational regulation miRNAs. To check whether, transposon derived small RNAs are found in polysomal fraction, the percentage of transposon mapped sequences were calculated. To our expectation, we found that the majority of polysomal fraction contains piRNAs. However it remains to be elucidated that these piRNAs are associated with polysome but not heavy complexes. In order to determine the expression level of transposon the Repbase colection was used. We used the same calculation as previously except, we normalised our data to total number of reads (Brennecke, et al. 2007). The expression level of transposon decreased or not change towards 7-8h. This is a direct indication of the role of piRNAs in early development.

For piRNAs matched to the abovementioned transposon we only used reads in length of 23-29 bp. The main important feature of piRNAs, studied in *Drosophila* so far, is the presence of U the 1st position and A at the 10th position. Additionally, the

Table 3.2. The percentage of top 10 miRNAs in each sample. The mature sequences name also indicates the hairpin name. The percentage of sequences that mapped to all hairpin sequences and mature sequences in all sequences that mapped to hairpin sequences are given.

	hairpin										mature_3bp									
	mRNP		60S		80S		Poly		Tot		mRNP		60S		80S		Poly		Tot	
	0-1h	7-8h	0-1h	7-8h	0-1h	7-8h	0-1h	7-8h	0-1h	7-8h	0-1h	7-8h	0-1h	7-8h	0-1h	7-8h	0-1h	7-8h	0-1h	7-8h
dme-bantam-3p	0,3	3,4	36,3	27,3	35,9	26,4	2,9	1,0	1,5	30,3	0,3	3,3	35,7	26,2	35,7	26,1	2,6	0,9	1,5	30,1
dme-miR-1-3p	7,7	5,5	3,1	7,9	14,8	19,9	1,0	7,5	43,9	13,7	7,7	5,5	3,1	7,9	14,8	19,9	1,0	7,4	43,9	13,7
dme-miR-184-3p	17,5	9,6	7,6	4,6	10,4	6,1	27,0	25,9	11,1	3,6	17,5	9,6	7,6	4,6	10,4	6,1	27,0	25,9	11,1	3,6
dme-miR-263a-5p	18,0	24,7	1,5	4,2	3,4	5,1	33,1	22,2	3,5	9,0	18,0	24,7	1,5	4,2	3,4	5,1	33,1	22,1	3,5	9,0
dme-miR-283-5p	0,2	0,3	0,9	1,3	1,3	1,9	1,4	1,8	0,2	1,4	0,2	0,3	0,9	1,2	1,3	1,8	1,4	1,8	0,1	1,4
dme-miR-286-3p	2,0	10,8	7,1	18,1	7,5	18,9	6,8	10,3	2,2	11,6	2,0	10,8	7,1	18,1	7,5	18,9	6,8	10,3	2,2	11,5
dme-miR-305-5p	1,0	0,0	1,5	0,3	0,4	0,1	0,3	0,0	1,7	0,1	1,0	0,0	1,3	0,2	0,3	0,0	0,3	0,0	1,7	0,1
dme-miR-306-5p	2,7	0,0	0,1	0,0	0,1	0,0	0,1	0,0	2,4	0,1	2,7	0,0	0,1	0,0	0,1	0,0	0,1	0,0	2,4	0,1
dme-miR-311-3p	2,1	0,0	0,6	0,1	1,0	0,1	0,2	0,1	2,6	0,1	2,1	0,0	0,6	0,1	1,0	0,1	0,2	0,1	2,6	0,1
dme-miR-315-5p	0,2	0,8	0,6	0,8	0,3	0,5	1,4	0,8	0,2	2,0	0,2	0,8	0,6	0,8	0,3	0,5	1,4	0,8	0,2	2,0
dme-miR-5-5p	16,4	27,2	3,4	4,3	2,4	3,5	5,9	7,0	3,3	4,5	16,4	27,2	3,1	3,9	2,3	3,3	5,9	6,9	3,2	4,2
dme-miR-7-5p	0,1	1,5	0,5	1,2	0,6	1,6	1,1	1,4	0,3	1,0	0,1	1,5	0,5	1,2	0,6	1,6	1,1	1,4	0,3	1,0
dme-miR-8-3p	0,9	1,2	3,0	2,9	3,3	3,0	1,2	1,7	1,3	2,5	0,5	1,1	2,7	2,5	3,1	2,7	1,2	1,6	1,2	2,2
dme-miR-92b-3p	0,3	0,0	1,6	1,0	2,0	1,2	0,2	0,3	0,8	0,1	0,3	0,0	1,6	0,9	2,0	1,2	0,2	0,3	0,8	0,1
dme-miR-958-3p	0,0	0,1	4,2	5,1	0,5	0,6	0,1	0,2	0,0	0,8	0,0	0,1	4,1	5,0	0,5	0,6	0,0	0,1	0,0	0,8
dme-miR-9a-5p	6,0	5,3	1,6	3,5	1,7	2,9	1,7	8,6	4,1	6,4	6,0	5,3	1,6	3,4	1,7	2,9	1,7	8,6	4,1	6,4
dme-miR-9c-5p	11,5	4,1	8,8	2,4	6,0	1,8	9,1	4,4	9,4	2,6	11,5	4,1	8,8	2,4	6,0	1,8	9,0	4,4	9,4	2,5
dme-miR-iab-4-5p	1,2	0,1	0,1	0,1	0,0	0,0	0,1	0,0	0,5	0,1	1,6	0,1	0,0	0,1	0,0	0,0	0,2	0,1	0,7	0,1

Table 3.3. miRNA expression level grouping. 4 major miRNA groups were found and are labeled as 1-4, and the miRNA mature names are given. The fold changes in red indicates increase, while green decrease in 7-8h, and blue indicates no significant change.

G	miRNA	Tot	mRNP	60S	80S	Poly	G	miRNA	Tot	mRNP	60S	80S	Poly
1	dme-miR-286-3p	2,6	3,1	2,2	2,1	1,6	2	dme-miR-11-3p	0,8	1,3	0,8	0,5	2,4
1	dme-miR-283-5p	3,6	1,5	1,3	1,3	1,3	2	dme-miR-9b-5p	0,4	0,8	-0,4	-0,2	1,7
1	dme-miR-7-5p	2,0	4,5	2,0	2,1	1,3	2	dme-miR-1012-3p	-0,8	-0,8	-0,5	-1,0	1,4
1	dme-miR-957-3p	2,6	1,3	1,2	1,2	0,8	2	dme-miR-1010-3p	0,2	-0,6	0,4	0,5	1,4
1	dme-miR-263a-5p	1,6	1,1	2,3	1,3	0,4	2	dme-miR-5-5p	0,6	1,4	1,2	1,3	1,2
1	dme-miR-315-5p	3,9	2,6	1,3	1,3	0,2	2	dme-miR-12-5p	0,3	0,4	0,6	1,2	1,1
1	dme-miR-314-3p	5,6	2,1	1,3	1,0	0,0	2	dme-miR-13a-3p	0,2	0,6	1,3	0,2	0,0
1	dme-bantam-5p	2,6	1,7	1,5	1,2	-0,8	2	dme-miR-31a-3p	0,2	0,0	1,0	0,7	-0,3
1	dme-miR-958-3p	7,0	3,6	1,1	0,8	1,8	2	dme-miR-124-3p	0,0	-0,3	1,2	0,2	-0,3
1	dme-miR-956-3p	3,3	1,2	1,0	0,6	0,7	2	dme-miR-965-5p	0,9	-0,4	0,7	0,8	0,8
1	dme-miR-998-3p	1,3	1,2	1,5	0,5	0,4	2	dme-miR-1003-3p	-0,8	-0,5	0,1	0,0	0,0
1	dme-miR-2c-5p	1,4	1,2	1,6	0,2	0,2	2	dme-miR-190-5p	0,4	0,1	0,0	0,1	-0,7
1	dme-miR-304-5p	6,0	1,5	0,9	0,5	0,1	2	dme-miR-282-3p	-0,1	-1,9	-1,7	-0,7	-0,7
1	dme-miR-8-3p	1,1	1,8	0,8	0,5	1,3	2	dme-miR-2a-3p	0,6	0,2	0,8	0,5	-1,5
1	dme-miR-983-5p	2,0	1,2	0,7	0,3	0,0	2	dme-miR-33-5p	-0,3	-2,4	-1,6	-1,1	-1,9
1	dme-bantam-3p	4,6	4,3	0,4	0,3	-0,5	2	dme-miR-281-2-5p	-0,5	-1,3	0,3	-0,1	0,6
1	dme-miR-2b-3p	1,6	1,2	0,4	0,8	-1,1	2	dme-miR-2b-2-5p	-0,8	-1,6	0,5	1,0	0,5
1	dme-miR-276a-3p	3,3	0,9	1,4	1,0	0,8	2	dme-miR-375-3p	0,1	-2,3	0,8	-0,2	0,1
1	dme-miR-5-3p	2,7	0,8	1,7	1,0	2,2	3	dme-miR-3-5p	-2,8	0,0	-0,1	-0,1	0,0
1	dme-miR-263b-5p	1,5	0,7	1,9	1,2	1,5	3	dme-miR-184-3p	-1,4	-0,2	0,1	0,0	0,9
1	dme-miR-137-3p	1,3	-0,6	1,6	1,3	1,4	3	dme-miR-312-5p	-3,8	-0,2	-1,1	-0,6	-0,6
1	dme-miR-1002-5p	3,1	-0,2	1,2	1,9	1,4	3	dme-miR-6-1-5p	-4,2	-0,3	-0,7	-0,8	0,0

(Cont. on next page)

Table 3.3. (Cont.)

1	dme-miR-10-5p	1,9	0,2	1,2	0,7	2,2	3	dme-miR-3-3p	-1,8	-0,4	0,4	0,5	0,7
1	dme-miR-252-5p	1,5	0,8	1,7	0,7	0,5	3	dme-miR-999-3p	-1,7	-0,4	-0,2	-0,1	0,6
1	dme-miR-316-5p	3,4	0,0	1,3	0,9	-0,1	3	dme-miR-305-3p	-1,0	-0,7	-0,2	-0,6	0,0
1	dme-miR-1006-3p	2,5	0,5	1,1	0,5	0,0	3	dme-miR-9c-5p	-1,7	-0,8	-1,0	-1,0	0,0
1	dme-miR-968-5p	4,2	0,6	1,1	0,3	0,0	3	dme-miR-13b-2-5p	-1,4	-0,8	0,7	0,6	0,1
1	dme-miR-1000-5p	1,9	0,1	1,1	0,6	0,7	3	dme-miR-9b-3p	-2,5	-1,1	0,0	-0,1	-0,3
1	dme-miR-987-5p	2,1	0,7	0,2	0,1	1,1	3	dme-miR-79-3p	-2,3	-1,1	-0,6	0,0	0,7
1	dme-miR-31a-5p	1,3	0,9	0,5	0,1	0,8	3	dme-miR-92a-3p	-2,3	-1,3	0,3	-0,3	1,4
1	dme-miR-13b-3p	1,1	0,1	0,9	0,7	0,5	3	dme-miR-124-5p	-1,5	-1,3	0,5	0,5	0,4
1	dme-miR-927-3p	1,8	0,6	0,6	0,6	0,5	3	dme-miR-1002-3p	-1,6	-1,6	0,3	0,2	0,0
1	dme-miR-996-3p	1,0	-0,2	-0,2	0,2	0,1	3	dme-miR-308-3p	-1,4	-1,7	-0,8	-0,1	-0,5
1	dme-miR-284-5p	2,4	0,4	0,6	0,1	0,0	3	dme-miR-310-3p	-3,0	-1,7	-1,4	-1,7	-0,3
1	dme-miR-927-5p	2,1	-0,6	0,5	0,4	0,0	3	dme-miR-92b-3p	-2,5	-1,9	0,1	0,0	1,2
1	dme-miR-87-3p	2,6	-0,3	0,7	0,0	-0,1	3	dme-miR-968-3p	-2,1	-2,2	0,3	0,5	0,7
1	dme-miR-281-3p	1,6	-0,2	-0,2	-0,7	-0,1	3	dme-miR-995-3p	-1,4	-2,4	-0,1	-0,4	0,4
1	dme-miR-1008-3p	1,4	-0,2	0,5	0,3	-0,5	3	dme-miR-312-3p	-5,2	-2,5	-1,2	-0,9	-0,2
1	dme-miR-8-5p	1,1	-1,6	1,0	1,0	2,6	3	dme-miR-279-3p	-3,9	-2,6	0,3	0,0	0,8
1	dme-miR-10-3p	2,2	-1,0	0,6	1,2	2,0	3	dme-miR-6-2-5p	-4,9	-2,7	-1,0	-0,1	0,4
1	dme-miR-14-3p	1,0	-1,4	1,5	-0,3	1,1	3	dme-miR-iab-8-3p	-2,3	-3,2	0,8	0,5	-0,3
4	dme-miR-1-3p	-1,5	0,2	2,2	1,2	3,8	3	dme-miR-iab-4-5p	-2,3	-3,3	1,0	0,7	-0,3
4	dme-miR-988-3p	-1,0	-0,1	1,4	1,2	0,0	3	dme-miR-275-3p	-2,9	-3,6	-1,2	-1,4	0,2
4	dme-miR-4-3p	-1,1	-0,5	2,2	0,7	1,4	3	dme-miR-282-5p	-2,5	-3,7	-1,7	-1,2	-1,4
4	dme-miR-31b-5p	-1,2	-0,9	0,1	-0,9	1,7	3	dme-miR-311-3p	-5,2	-4,8	-1,2	-2,0	-0,2
4	dme-miR-981-3p	1,0	-1,9	-0,1	0,6	0,4	3	dme-miR-305-5p	-4,0	-4,9	-1,8	-1,8	-1,3
2	dme-miR-9a-5p	0,8	0,5	1,9	1,5	3,3	3	dme-miR-306-5p	-5,1	-5,2	-0,6	-0,8	-1,1

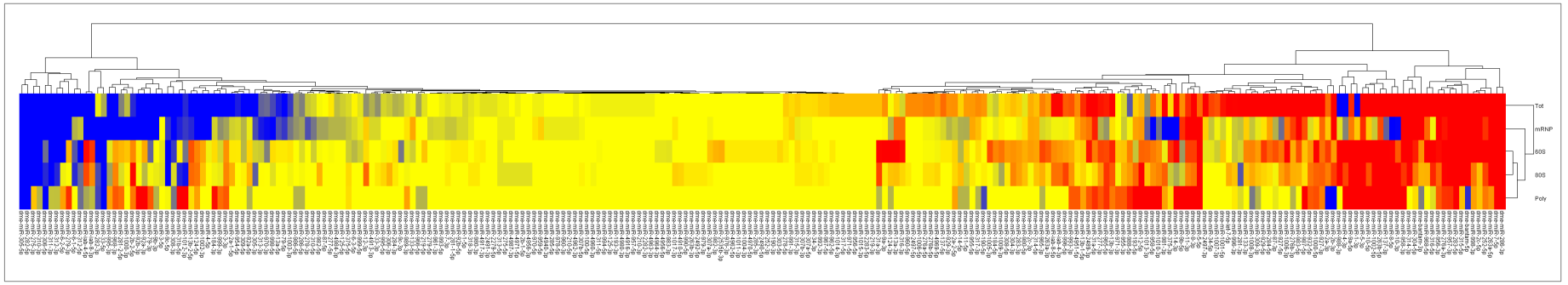

Figure 3.4. The Clustering of miRNA expression level. The log2 fold chages were cluster by Gene Cluster 3.0 and viewed by Java TreeView. Here only the mature_3bp (94 miRNAs) that has at least 50 RPM in at least one sample are shown as the fold changes of the remaing are not changed.

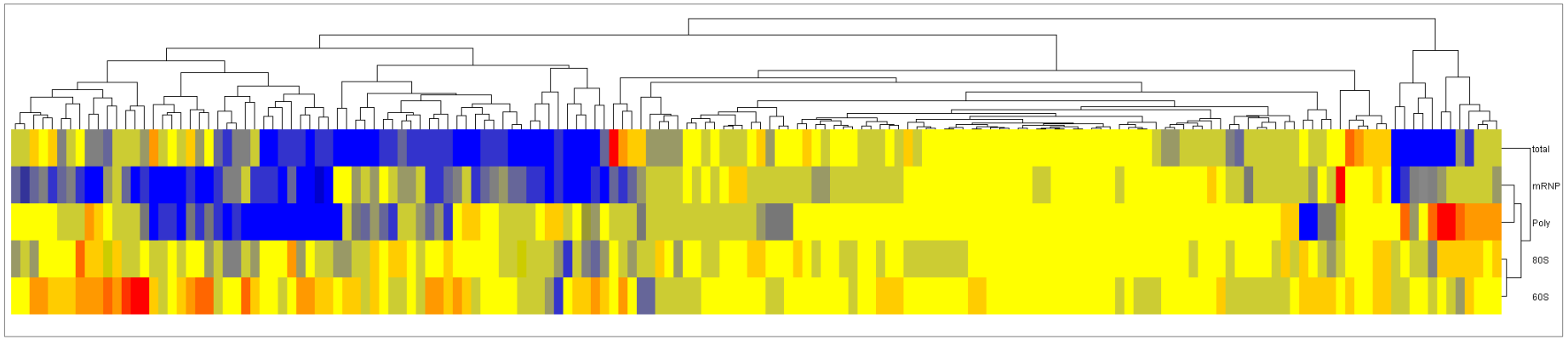

Figure 3.5. The fold changes of Repbase Collection. The fold changes of each transposon were converted to log2 and clsutered by Gene Cluster3.0 and viewed by Java TreeView.

piRNAs with U at the first position has preference for Aub while the ones with A at the position 10th are found in PIWI/AGO3 (Brennecke, et al. 2007). Morover piRNAs has 10 nucleotide complementarity at the 5' end and are generated by ping pong cycle. In order check whether these reads are piRNAs or not, we aligned sequences to the transposon from flybase. Then, for each piRNA on the transposon we calculated the distance to antisense reads to each sense reads frequency. We found that, there is a major pick at the position 10 which supports that these small RNAs are piRNAs (Figure-3.6). We further analyses the frequency of nucleotide at the position 1 and 10. We found that the majority of position 1 is U while the position 10 are A (Figure-3.7). Thus, the small RNA population matched to the transposons has unique features of piRNAs.

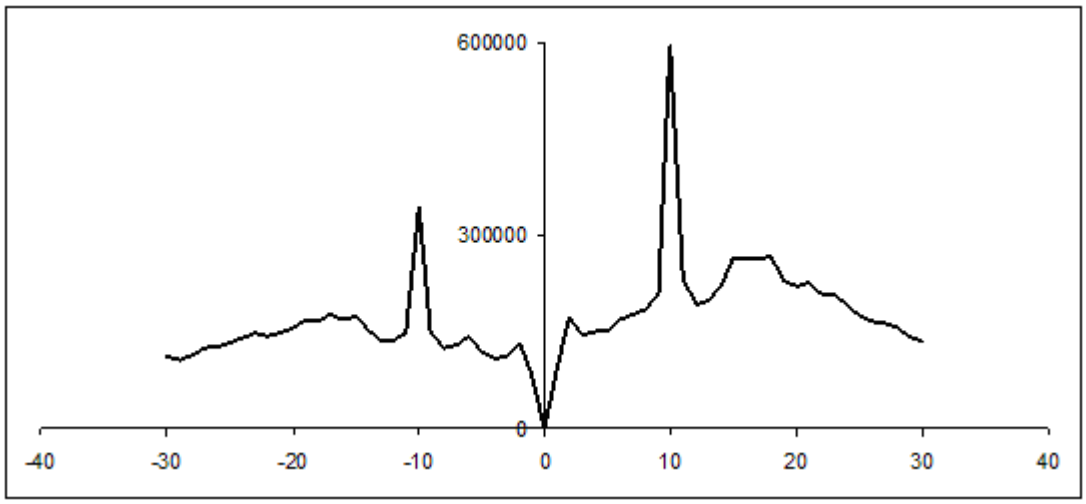

Figure 3.6. The 5'-5' complemetarity among the transposon matched small RNAs. The distance of each small RNAs from each other were calculated and the frequency of each distance were determined for all reads.

Because the majority of piRNAs are found in polysomal fractions we next determined the length distribution of transposon matched sequences in all samples. Interestingly, there is a peak at 21 bp in total 7-8h embryo, with hight frequency compared with 0-1h. Moreover, the 21 bp frequency is high in mRNP fraction while dereasing towards polysomal fractions. Comparing in all samples, we found that 21 bp are higher in 7-8h embryo. This data supports that siRNAs are important in later develeopmental stages compared with piRNAs. Additionally, their inracellular localisation differs. Thus, the amound of siRNA decreases gradually towards polysome, while of piRNAs increases (Figure-3.8, Figure-3.9).

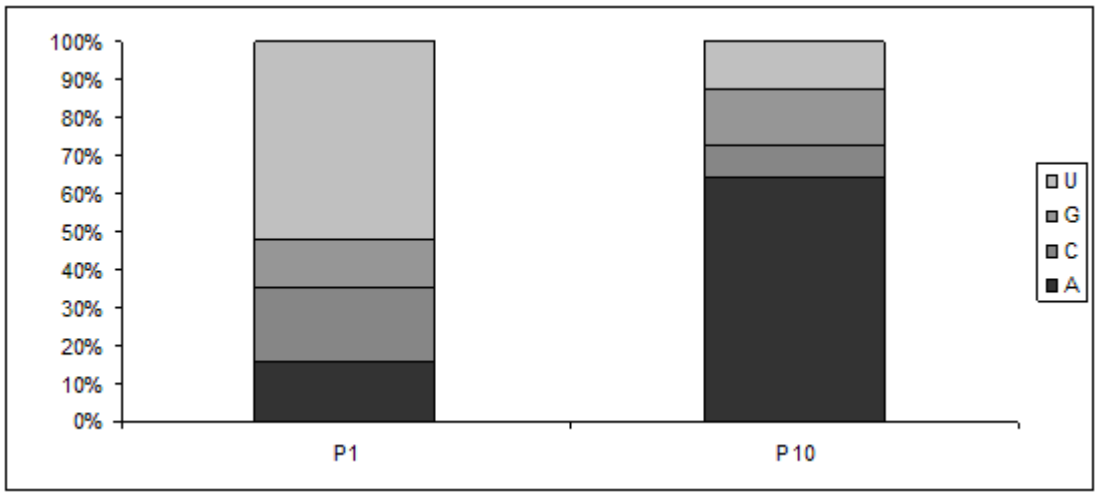

Figure 3.7. The nucleotide bias at position 1 and 10 th of the transposon mapped small RNAs. To check whether these reads has U or A at position 1 and 10 th we calculated the number of reads that has A,C,G or U at Pos 1 and 10, only using the small RNAs that has 10 nucleotide complementarity.

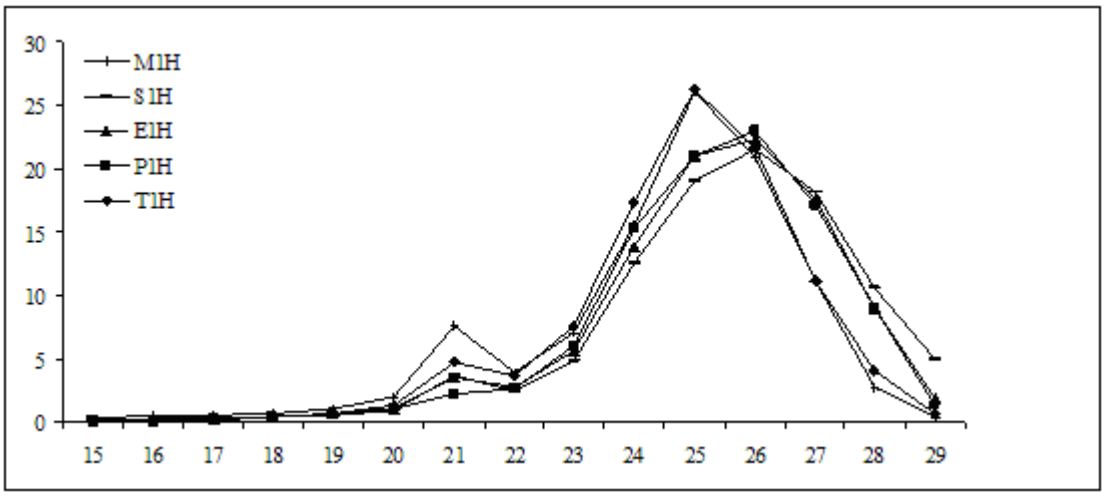

Figure 3.8. The nucleotide distribution of transposon derived small RNAs in 0-1h. The percentage of small RNAs in total exact matched reads to the transposon were calcculated in each fraction.

Moreover, we collected deep-sequencing data from different developmental stages of *Drosophila melanogaster*.We found that, the siRNA contents increased significantly during development. The majority of S2 and Kc cell line transposons matched small RNAs are in length of 21 bp. This is a direct evidence of the role of siRNAs in somatic cells. Evenmore, the transposon derived small RNA contents decreases in later development, as shown in Figure-3.11.

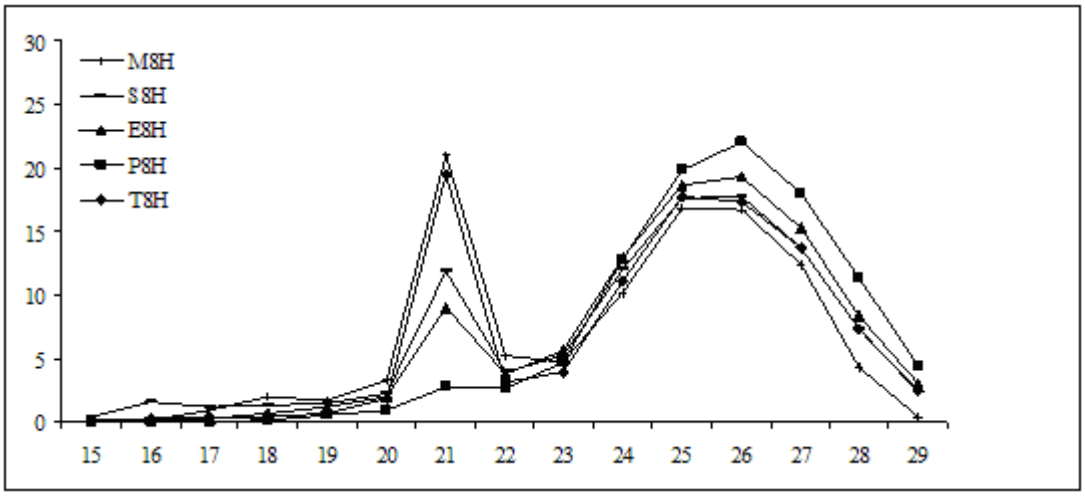

Figure 3.9. The nucleotide distribution of transposon derived small RNAs in 7-8h. The percentage of small RNAs in total exact matched reads to the transposon were calcculated in each fraction.

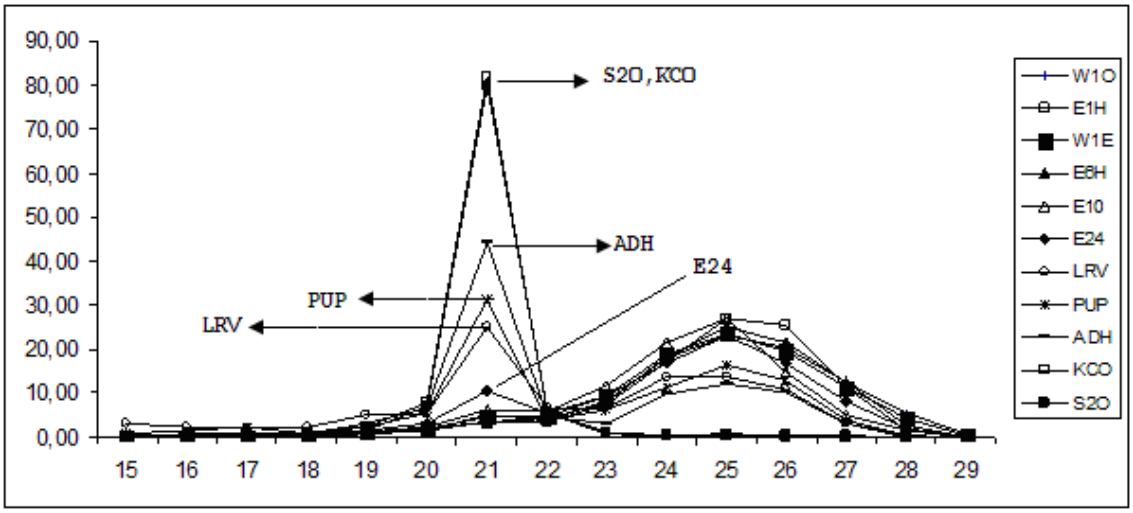

Figure 3.10. The percentage of transposon derived small RNAs in different developmental stages of *Drosophila*. The percentage of small RNAs in total exact matched reads to the transposon were calcculated in each fraction.

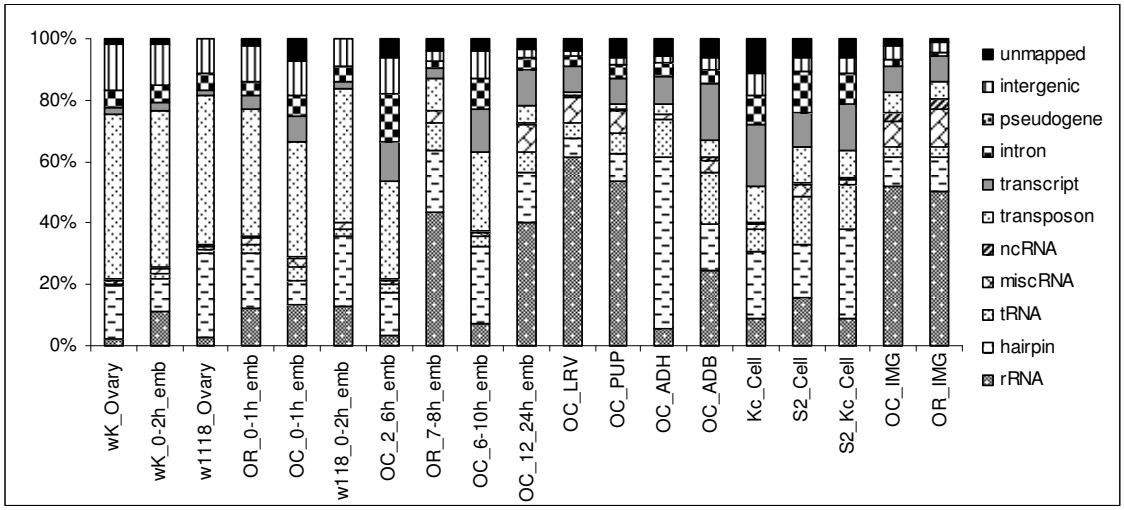

Figure 3.11. The known RNAs content in different developmental stages of Drosphila melanogaster. The strain names are wK: wild type caugth, OR: laboratory wild type, W1118: white eye, OC: Oregon Cansas, Ovary: oocyte, emb: embryo, Kc and S2: Drosphila cell lines, ADH: Adult Head, ADB: Adult Body, LRV: Larva, PUP: Pupa, IMG: Imaginal Discs. (These data sets were collected from previously published articles and raw sequences downloaded from GEO according to the GEO Accession Numbers in these articles; Aravin, *et al*, 2003; Brennecke, *et al*, 2007; Chung, *et al*, 2008).

We also check the relative percentage of *Drosophila* small RNAs in all developmental stages described in Figure-3.8, 3.9. As can be seen the transposon mathced small RNAs percentage decreases in later developmental stages, while miRNA expression level increases (Figure-3.10, 3.11).

3.1.5 tRF Expression Levels

The detailed data about tRF analyses were descirebed elsewhere (Goktas, *et al.*, unpublished data). Here, a brief information about the most abundant tRNA Glycine is given and the summary of length distribution is described.

For each the tRNA, the nucleotide position were determined and the frequency of the reads in each sample were calculated in RPM. Then for each nucleotide at a particular the number of times the nucleotide cloned is calcuted and shown in Figure-3.12. We found that the majority of tRF (~95%) are derived from the first 29 nt of mature tRNAs. It should be noted that the most abundant tRF are derived from the 5'

end of tRNAs, also there are fragments that are derived from middle or 3’ end of tRNA on individual tRNA (not shown).

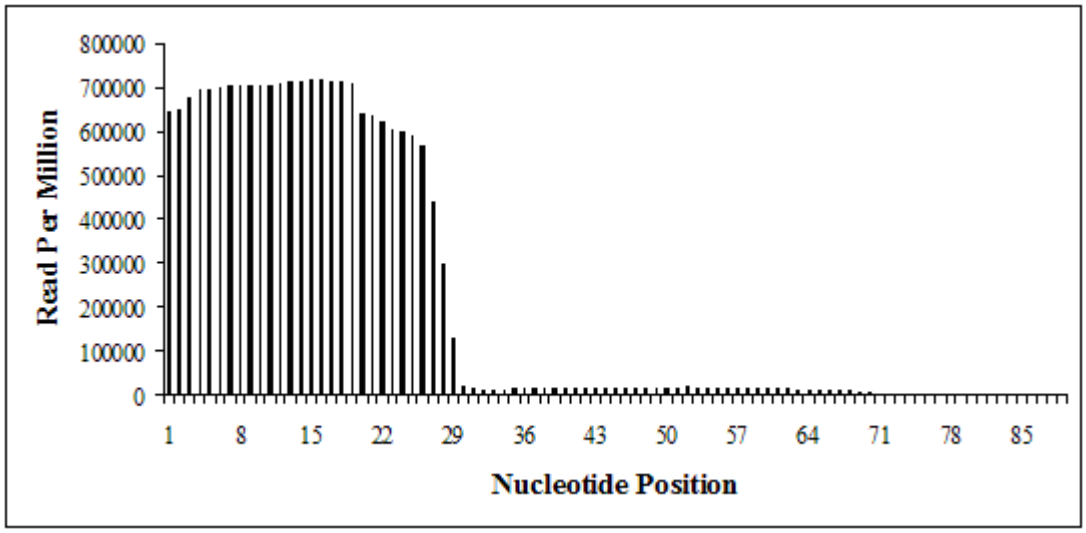

Figure 3.12. The nucleotide distribution on a superimposed tRNA.

Also the 3' end processing of all tRNAs were determined on the superimposed tRNA (Figure-3.12) and on a clover-leaf structure of tRNAs (Figure-3.13). The 3’ end of tRF fragments are at position 19 and 26-29 with highest peak at 28 nt. Around 25% of the mitochondrial tRNA (mt:tRNA:S:AGY) is 19 nt from the first 19 nt of the 5’ end. Thus, except for mitochondrial tRNAs the tRF from *Drosophila* is derived from the 26-29 nt at the 5’ end.

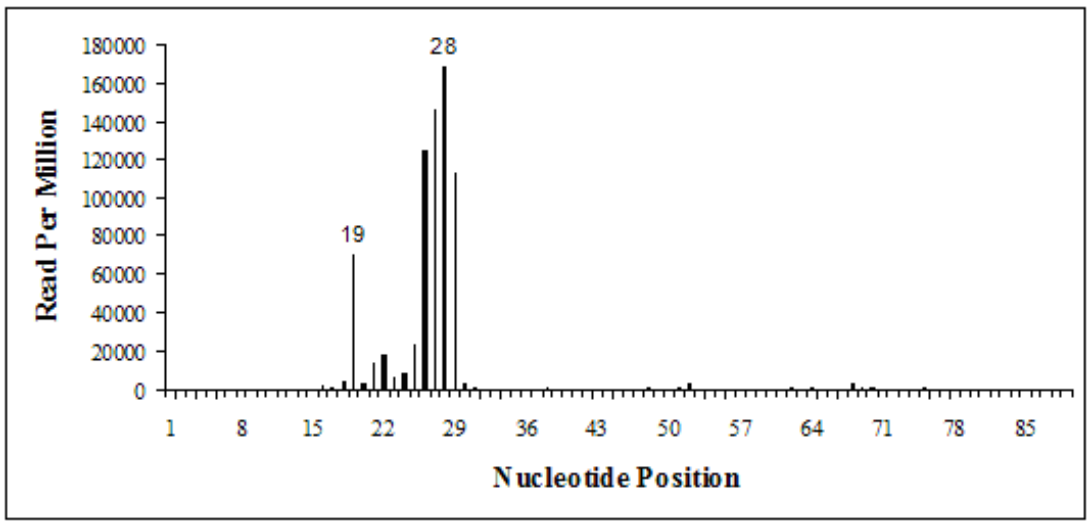

Figure 3.13. The 3' end processing of tRNAs. The frequency of each read's 3'-end were calculated for all tRNAs on a superimposed tRNAs. In Figure-10, all nucleotides were calculated while here only the 3' end nucleotide were calculated as it gives the averege length distribution of tRFs.

In order to check the 3’ end position on the cloverleaf of mature tRNAs, we divided a tRNA into 15 regions (Figure-3.14) and calculated the 3’ end position of the tRF in this regions. The mt:tRNA:S:AGY does not have a cloverleaf structure as a result we excluded this tRNA from this analyses. Thus, around 75% of tRF were used to determine the location 3’ end processing.

We found that the majority of tRF are derived at region 7 and at position 2,3,4. Thus, tRF 3’ end are processedd from the 2nd, 3rd and 4th nt at the anticodon stem close to D-loop and enriched at 4th nt. This is interesting as there four 3’ position on primary structure (26-29 nt), 3 position are enrcihed on the secondary structure, which indicates specific processing on the anticodon stem.

We choose one of the tRNA that has the most abundant tRF for further detailed analyses. Majority of the tRF are derived from tRF glycine and we calculated the frequency of each nucleotide in each farctions. As can be seen, tRFs from glycine are derived from the 5’ end and higher in 0-1h/7-8h fractions compared with other fractions (Figure-3.15).

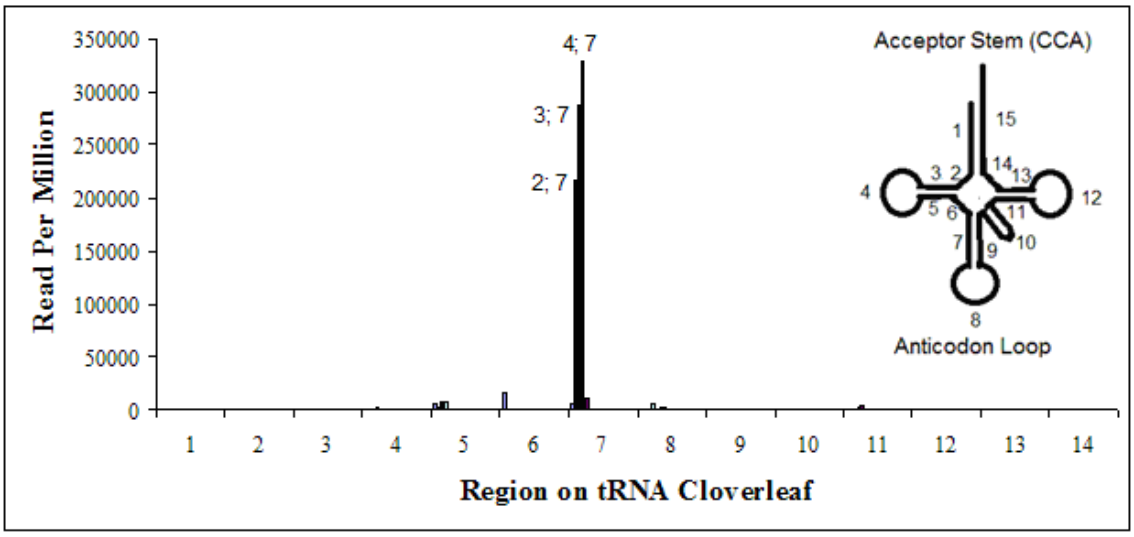

Figure 3.14. The 3'-end processing sites on the 2D-tRNA structure. The cloverleaf tRNA sturcture were divided into 15 regions and the 3'-end processing sites in each region were calculated.

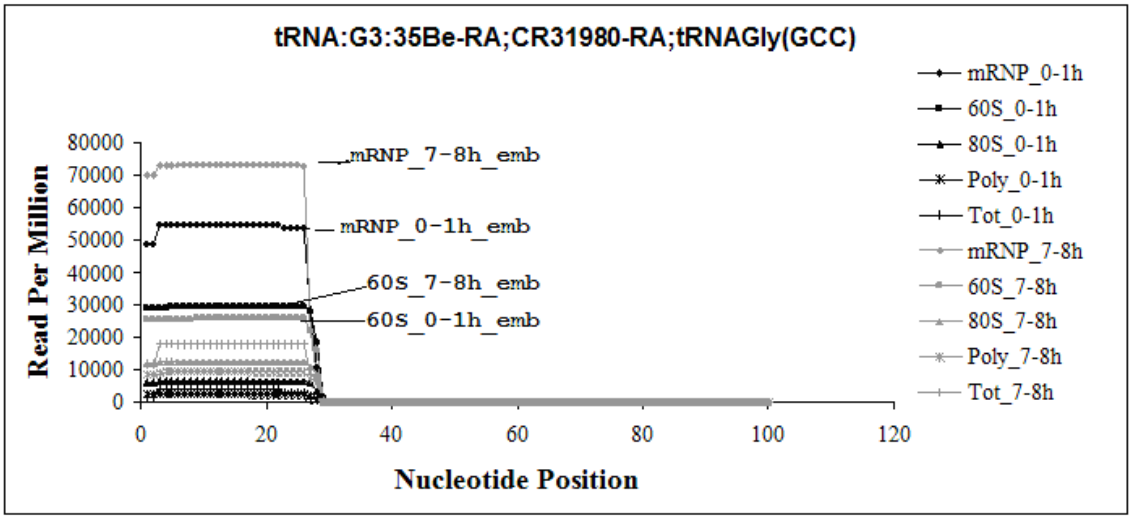

Figure 3.15. The frequency of tRNA:G:35Be for all samples. The most abundant tRFs derived from the tRNA:G: were shown on the tRNA by calculating the frequncy of each nucleotide in each sample seperately.

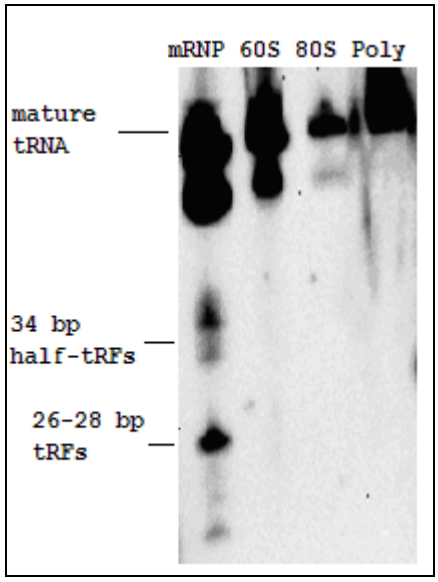

Figure 3.16. Polysomal fractions of 7-8h embryo and Northern Blot of tRNA:G:35Be. A probe against the first 35 nucleotide of tRNA:G: were used to show tRF derived from the 5' end of the tRNA:G:. (Source: Master Thesis of Çağdaş Göktaş).

We choose one of the tRNA that has the most abundant tRF for further detailed analyses. Majority of the tRF are derived from tRF glycine and we calculated the frequency of each nucleotide in each farctions. As can be seen, tRFs from glycine are derived from the 5' end and higher in 0-1h/7-8h fractions compared with other fractions (Figure-3.15).

In order to confirm these observation in 7-8h embryo northern blot was performed by using bioting labelled probe against the first 40 bp of the tRNA:G3:35Be. We have shown ~27 nt tRF are enrcihed in mRNP frcaction.

CHAPTER 4

CONCLUSION

During the maternal-to-zygotic transition two crucial events takes place, (i) clearance of maternal mRNAs, (ii) activation of the genome and transcription of zygotic mRNA. The clearance of maternal mRNAs has been shown in different organisms that depends on RNA-Binding proteins (Gallo, *et al*. 2008; Schubert, *et al*. 2000; D'Agostino, *et al*. 2006; Smibert, *et al*., 1996; De Rensiz, *et al*. 2007; Tadros, et al., 2007; Aviv, *et al*., 2006) or small RNAs including miRNA, siRNA and piRNAs (Rouget, *et al.* 2010; Bushati, *et al*. 2008; Giraldez, *et al.* 2006; Tchurikov, *et al*. 2011). The maternal mRNA degradation depends on translatioal repression and destabilisation of mRNA by removing of poy(A) tail (Semotook, et al.2005; Semotook, et al. 2007). One of the location for mRNA decay or storage is in processing bodies (Flemr, et al. 2010; Sheth, et al. 2007), where as a miRNA can stop translation at any step of translation (for review; Fabian, et al. 2012). To answer the temporal and spatial expression of miRNAs, we used polysomal fractionation and deep-sequencing method together. We have shown that miRNAs expression level changes during maternal-to-zygotic transition. Interestingly, the localisation of miRNAs seems to does not change during the MZT. If a miRNA acts in mRNP fraction it seems that the miRNA are highly found in these samples for two embryonic stages. For instance, the developmental important bantam are highly expressed in 7-8h embryo (30-fold), while the localisation of the miR-bantam does not change, it acts in the 60S and 80S but not mRNP or polysomes. This is interesting as an miRNA transcription increases the complexes the miRNA interacts does not change. On the other hand, as the amount of the miRNA increases the increase is distributed in all samples not evenly. We have defined 4 miRNA groups whose expression level differs from each other. As expected, three miRNA groups were determine in total RNAs as increase or decrease or no changes. But these three groups has been further grouped according to their behaviour in translational machinary. Our data shows that not all miRNAs behaves in the same way in the cytoplasm but different groups of miRNAs behaves the sames. The main mechanism behind such behaviour is not known. By assuming that a miRNAs is

enriched where its target is found, the behaviour of these miRNAs should be further charcaterised by miRNA and its target reciprocal location in during translation.

Another group of small RNAs are transposon derived small RNAs that are divided into two major groups as siRNA (21 nt) and piRNAs (23-29 nt).The behaviour of these small RNAs differs from each other in two ways, (i) siRNA are highly expressed in 7-8h, while piRNAs are expressed in 0-1h, (ii) siRNA mainly found in mRNP while piRNAs in polysomes. We further analysed from other datas available at different developmnetal stages that temporal expression of siRNA and piRNAs differs in Drosphila. The siRNA population derived from transposon significanly increases at later development and most probably in somatic cells. It should be noted that we used only the early embryo which i enriched still with piRNAs. Either the increase or decrease is because of dilution of the germ cells towards either remains to be elucidated.

The third major group of small RNAs are tRFs and mainly found in mRNP and 60S fractions. The main functions of tRFs is not known, except the tRF-1001 which is derived from 3'-trailer of tRNA has been shown to be functional. In our data we have shown that tRF are in non-polysomal fractions, but whether these tRFs has a functional role remains to be elucidated.

Overall, the behaviour of four most abundant small RNAs differ in their localisation in polysomal fractions. While miRNAs can act in all 4 fractions, siRNAs and tRFs act in non-polysomal fraction and piRNAs mainly in polysomal frcations.

REFERENCES

Aalto A. P. and Pasquinelli A. E. (2012). **Small non-coding RNAs mount a silent revolution in gene expression.** *Curr. Opp. in Cell Bio.*, **24**., 1-8.

Adams M. D., Celniker S. E., Holt R. A., Evans C. A., Gocayne J. D., Amanatides P. G., Scherer S. E., Li P. W., Hoskins R. A., Galle R. F., et al. (2000). **The genome sequence of *Drosophila melanogaster*.** *Science.,* **287**., 2185-95.

Aoki F., Worrad D. M., Schultz R. M. (1997). **Regulation of transcriptional activity during the first and second cell cycles in the preimplantation mouse embryo.** *Dev. Biol.,* **181**., 296-307.

Aravin A.A., Lagos-Quintana, M., Yalcin, A., Zavolan, M., Marks, D., Snyder, B.,Gaasterland, T., Meyer, J., Tuschl, T. 2003. **The small RNA profile during *Drosophila melanogaster* development**. *Dev. Cell* 5, 337–350.

Aviv T., Lin Z., Ben-Ari G., Smibert C. A., Sicheri F. (2006). **Sequence specific recognition of RNA hairpins by the SAM domain of Vts1p.** *Nat. Struct. Mol. Biol.*, **13**,168-176.

Bachvarova R., De Leon V., Johnson A., Kaplan G., Paynton B. V. (1985). **Changes in total RNA., polyadenylated RNA., and actin mRNA during meiotic maturation of mouse oocytes.** *Dev. Biol.,* **108**., 325-31.

Bachvarova R. F. (1992) **A maternal tail of poly(A): the long and the short of it.** *Cell*., **69**., 895-97.

Bartel D. P. (2004). **MicroRNAs: genomics., biogenesis., mechanism., and function.** *Cell.,* **116**., 281–97.

Bashirullah A., Halsell S. R., Cooperstock R. L., Kloc M., Karaiskakis A., Fisher W.W.,Fu W., Hamilton J. K., Etkin L. D., Lipshitz H. D. (1999). **Joint action of two RNA degradation pathways controls the timing of maternal transcript elimination at the midblastula transition in *Drosophila melanogaster*.** *EMBO J.*, **18**., 2610–20.

Benoit B., He C. H., Zhang F., Votruba S. M., Tadros W., Westwood J. T., Smibert C. A., Lipshitz H. D., Theurkauf W. E. (2009). **An essential role for the RNA-binding protein Smaug during the *Drosophila* maternal-to-zygotic transition.** Development.,**36**., 923-32.

Berezikov E., Chung W. J., Willis J., Cuppen E., Lai E. C. (2007). **Mammalian mirtron genes.** *Mol. Cell*., **28**., 328–36.

Borchert G. M., Lanier W., Davidson B. L. (2006). **RNA polymerase III transcribes human microRNAs.** *Nat. Struct. Mol. Biol.,* **13.,** 1097–101.

Braude P., Bolton V., Moore S. (1988). **Human gene expression first occurs between the four- and eight-cell stages of preimplantation development.** *Nature*., **332.,** 459-61.

Brennecke J., Aravin A. A., Stark A., Dus M., Kellis M., Sachidanandam R., Hannon G. J. (2007). **Discrete small RNA-generating loci as master regulators of transposon activity in *Drosophila*.** *Cell*., **128**., 1089-103.

Brennecke J., Malone C. D., Aravin A. A., Sachidanandam R., Stark A., Hannon G. J. (2008) **An epigenetic role for maternally inherited piRNAs in transposon silencing.** *Science.,* **322**., 1387-92.

Brennecke J., Stark A., Russell R. B. and Cohen S. M. (2005). **Principles of microRNA- target recognition.** *PLoS Biol.,* **3**., e85.

Bushati N., Stark A., Brennecke J., Cohen S. M. (2008). **Temporal reciprocity of miRNAs and their targets during the maternal-to-zygotic transition in *Drosophila*.** *Curr. Biol*., **18**., 501-6.

Cai X., Hagedorn C. H., Cullen B. R. (2004). **Human microRNAs are processed from capped., polyadenylated transcripts that can also function as mRNAs.** *RNA.,* **10**., 1957–66.

Chendrimada T. P., Finn K. J., Ji X., Baillat D., Gregory R. I., Liebhaber S. A., Pasquinelli A., E Shiekhattar R. (2007). **microRNA silencing through RISC recruitment of eIF6**. *Nature*., **447** ., 823–828.

Chung WJ, Okamura K, Martin R, Lai EC. **Endogenous RNA interference provides a somatic defense against Drosophila transposons.** *Curr Biol* 2008 Jun 3;18(11):795-802.

Cole C., Sobala A., Lu C., Thatcher S. R., Bowman A., Brown J. W., Green P. J., Barton G. J., Hutvagner G. (2009). **Filtering of deep sequencing data reveals the existence of abundant Dicer-dependent small RNAs derived from tRNAs.** RNA., **15**., 2147-60.

Crosby I. M., Gandolfi F., Moor R. M. (1988). **Control of protein synthesis during early cleavage of sheep embryos.** J. Reprod. Fertil., **82**., 769-75.

Dahanukar., A., and Wharton., R. P. (1996). **The Nanos gradient in *Drosophila* embryos is generated by translational regulation.** Genes Dev., **10**., 2610–20.

D'Agostino I., Merritt C., Chen P. L., Seydoux G., Subramaniam K. (2006). **Translational repression restricts expression of the C. elegans Nanos homolog NOS-2 to the embryonic germline.** *Dev. Biol.*, 292:244-252.

de Hoon M. J. L., Taft R. J., Hashimoto T., Kanamori-Katayama., M., Hideya Kawaji H., Kawano M., Kishima M., Lassmann T., Faulkne G. F., Mattick J. S., Daub C. O., Carninci P., Kawai J.,Harukazu Suzuki H., Hayashizaki Y. (2009). **Cross-mapping and the identification of editing sites in mature**

microRNAs in high-throughput sequencing libraries. *Genome Res.,* **20**., 257-64.

De La Fuente R., Eppig J. J. (2001). **Transcriptional activity of the mouse oocyte genome: companion granulosa cells modulate transcription and chromatin remodeling.** *Dev Biol.,* **229.,** 224-36.

De Renzis S., Elemento O., Tavazoie S., Wieschaus E. F. (2007). **Unmasking activation of the zygotic genome using chromosomal deletions in the *Drosophila* embryo.** *PLoS Biol.,* **5**., e117.

Elbashir S. M., Lendeckel W., Tuschl T. (2001). **RNA interference is mediated by 21- and 22-nucleotide RNAs**. *Genes Dev.*, **15**., 188–200.

Elbashir S. M., Martinez J., Patkaniowska A., Lendeckel W., Tuschl T. (2001). **Functional anatomy of siRNAs for mediating efficient RNAi in *Drosophila melanogaster* embryo lysate.** *EMBO J.,* **20.,** 6877–88.

Fair T., Hyttel P., Greve T., Boland M. (1996). **Nucleus structure and transcriptional activity in relation to oocyte diameter in cattle.** *Mol. Reprod. Dev.*, **43**., 503-12.

Fair T., Hyttel P., Greve T. (1995). **Bovine oocyte diameter in relation to maturational competence and transcriptional activity.** *Mol. Reprod. Dev.,***42**., 437-42.

Filipowicz W., Bhattacharyya S. N., Sonenberg N. (2008). **Mechanisms of post-transcriptional regulation by microRNAs: are the answers in sight?** *Nature*., **9**., 102-14.

Fire A., Xu S., Montgomery M. K., Kostas S. A., Driver S. E., Mello C. C. (1988). **Potent and specific genetic interference by double-stranded RNA in Caenorhabditis elegans.** *Nature*., **391**(6669)., 806-11.

Fulka H., Novakova Z., Mosko T., Fulka J. Jr. (2009). **The inability of fully grown germinal vesicle stage oocyte cytoplasm to transcriptionally silence transferred transcribing nuclei.** *Histochem. Cell Biol*., **132**, 457-68.

Gallo C. M., Munro E., Rasoloson D., Merritt C., Seydoux G. (2008). **Processing bodies and germ granules are distinct RNA granules that interact in C. elegans embryos**. *Dev Biol*., **323**, 76-87.

Gazzani S., Lawrenson T., Woodward C., Headon D., Sablowski R. A. (2004). **link between mRNA turnover and RNA interference in Arabidopsis.** Science, **306**, 1046–48.

Ghildiyal M. and Zamore P. D. (2009). **Small silencing RNAs: an expanding universe.** *Nat Rev Genet.*, **10,** 94–108.

Giraldez A. J., Mishima Y., Rihel J., Grocock R. J., Van Dongen S., Inoue K., Enright A. J., Schier A.F. (2006). **Zebrafish MiR-430 promotes deadenylation and clearance of maternal mRNAs.** *Science, **312***, 75-79.

Grishok A., Pasquinelli A. E., Conte D., Li N., Parrish S., Ha I., Baillie D. L., Fire A., Ruvkun G., Mello C. C. (2001). **Genes and Mechanisms Related to RNA Interference Regulate Expression of the Small Temporal RNAs that Control C. elegans Developmental Timing.** *Cell*., **106**, 23–34.

Grivna., S. T., Pyhtila B., and Lin H. (2006). **MIWI associates with translational machinery and PIWI-interacting RNAs (piRNAs) in regulating spermatogenesis.** *PNAS*., **103**(36) ,13415–20.

Gu S., Jin L., Zhang F., Sarnow P., Kay MA (2009). **Biological basis for restriction of microRNA targets to the 3' untranslated region in mammalian mRNAs.** *Nat Struct.Mol. Biol.*, **16**, 144-50.

Gunawardane L. S., Saito K., Nishida K. M., Miyoshi K., Kawamura Y., Nagami T., Siomi H., Siomi M. C. (2007). **A slicer-mediated mechanism for repeat-associated siRNA 5' end formation in *Drosophila*.** *Science,* 315, 1587-90.

Guo H., Ingolia N. T., Weissman J. S., Bartel D.P. (2010). **Mammalian microRNAs predominantly act to decrease target mRNA levels.** *Nature,* **466**, 835-40.

Haussecker D., Huang Y., Lau A., Parameswaran P., Fire A. Z., Kay M. A. (2010). **Human tRNA-derived small RNAs in the global regulation of RNA silencing**. RNA 2010., 16:673-695.

Heifetz Y., Yu J., and Wolfner., M. F. (2001). **Ovulation triggers activation of *Drosophila* oocytes.** *Dev. Biol.*, **234**, 416–24.

Hsieh,L.C., Lin, S.I., Kuo, H.F., and Chiou, T.J. (2010) **Abundance of tRNA-derived small RNAs in phosphate-starved Arabidopsis roots.** *Plant Signal Behav.* 5, Epub.

Jochl, C., Referstorff, M., Hertel, J., Stadler, P.F., Hofacker, I.L., Schrettle, M., Hass, H., and Huttenhofer, A. (2008). **Small ncRNA trancsriptome analysis from Aspergillus fumigatus suggests a novel mechanism for regulation of protein synthesis**. *Nucleic Acids Res* 36, 2677-2689.

Humphreys., D. T., Westman., B. J., Martin., D. I. & Preiss., T. (2005). **MicroRNAs control translation initiation by inhibiting eukaryotic initiation factor 4E/cap and poly(A) tail function.** *Proc. Natl Acad. Sci.*,102 ., 16961–16966.

Hutvágner G., et al. **A cellular function for the RNA-interference enzyme Dicer in the maturation of the let-7 small temporal RNA**. Science 2001;293:834–838.

Hutvagner G., Simard MJ. **Argonaute proteins: key players in RNA silencing**. Nat Rev Mol Cell Biol 2008;9:22–32

Hyttel P., Fair T., Callesen H., Greve T. (1996). **Oocyte growth., capacitation and final maturation in cattle.** Theriogenology 1997.,47:23-32.

Jacobson., A., and Peltz., S.W. (1996). **Interrelationships of the pathways of mRNA decay and translation in eukaryotic cells.** *Annu. Rev. Biochem.* 65., 693–739.

Jeanblanc M., Salvaing J., Mason K., Debey P., Beaujean N. (2008). **Embryonic genome activation.** Gynecol Obstet Fertil., 2008.,36:1126-1132.

Jennifer L. Semotok.,Ramona L. Cooperstock.,Benjamin D. Pinder.,Heli K. Vari.,Howard D. Lipshitz., and Craig A. Smibert., **Smaug Recruits the CCR4/POP2/NOT Deadenylase Complex to Trigger Maternal Transcript Localization in the Early *Drosophila* Embryo.** Current Biology., 15., 284–294.

John B., Enright A.J., Aravin A., Tuschl T., Sander C. and Marks D.S. (2004) **Human MicroRNA targets.** PLoS Biol., 2., e363.

Jurka., J., Kapitonov., V.V., Pavlicek., A., Klonowski., P., Kohany., O., Walichiewicz., J. (2005) **Repbase Update., a database of eukaryotic repetitive elements.** Cytogentic and Genom.e Research 110:462-467.

Jurka., J. Repbase (2000) **Update: a database and an electronic journal of repetitive elements**. Trends Genet. ***9***:418-420.

Kedde M., Strasser MJ., Boldajipour B., Oude Vrielink JA., Slanchev K., le Sage C., Nagel R., Voorhoeve PM., van Duijse J., Orom UA et al.: **RNA-binding protein Dnd1 inhibits microRNA access to target mRNA.** Cell 2007., 131 :1273-1286.

Ketting RF., et al. **Dicer functions in RNA interference and in synthesis of small RNA involved in developmental timing in C. elegans.** Genes Dev 2001;15:2654–2659.

Kim VN., Han J., Siomi MC. **Biogenesis of small RNAs in animals**. Nat Rev Mol Cell Biol 2009;10:126–39.

Kiriakidou M., Tan GS., Lamprinaki S., De Planell-Saguer M., Nelson PT., Mourelatos Z. 2007., **An mRNA m7G cap binding-like motif within human Ago2 represses translation.** Cell. 2007 Jun 15;129(6):1141-51.

Knight SW., Bass BL. **A role for the RNase III enzyme DCR-1 in RNA interference and germ line development in Caenorhabditis elegans**. Science 2001;293:2269–2271.

Kopecny V., **High-resolution autoradiographic studies of comparative nucleologenesis and genome reactivation during early embryogenesis in pig., man and cattle.**Reprod Nutr Dev., 1989.,29:589-600.

Krek, A., Grun, D., Poy, M.N., Wolf, R., Rosenberg, L., Epstein, E.J., Macmenamin.,P., da Piedade.,I., Gunsalus.,K.C., Stoffel.,M. et al. (2005) **Combinatorial microRNA target predictions**. Nature Genet., 37., 495–500.

Kunzi MS., Pitha PM. **Interferon research: a brief history**. Methods Mol Med 2005;116:25–35.

Leandri RD., Archilla C., Bui LC., Peynot N., Liu Z., Cabau C., Chastellier A., Renard JP., Duranthon V: **Revealing the dynamics of gene expression during embryonic genome activation and first differentiation in the rabbit embryo with a dedicated array screening.** Physiol Genomics., 2009., 36:98-113.

Lee, S.R., Collins, K. 2005. **Starvation-induced cleavage of the tRNA anticodon loop in *Tetrahymena thermophila***. *J Biol Chem* 280, 42744-42749.

Lee YS., Shibata Y., Malhotra A., Dutta A: **A novel class of small RNAs: tRNA-derived RNA fragments (tRFs) .** Genes Dev 2009., 23:2639-2649.

Lee Y., Ahn C., Han J., Choi H., Kim J., et al. 2003. **The nuclear RNase III Drosha initiates microRNA processing.** Nature 425:415–19.

Lee YS., et al. **Distinct roles for *Drosophila* Dicer-1 and Dicer-2 in the siRNA/miRNA silencing pathways.** Cell 2004;117:69–81.

Lee., R.C., Feinbaum., R.L., and Ambros., V. (1993). **The C. elegans heterochronic gene lin-4 encodes small RNAs with antisense complementarity to lin-14.** Cell 75., 843– 854.

Lewis.,B.P., Burge.,C.B. and Bartel.,D.P. (2005) **Conserved seed pairing., often flanked by adenosines., indicates that thousands of human genes are MicroRNA targets.** Cell., 120., 15–20.

Li C., Vagin VV., Lee S., Xu J., Ma S., Xi H., Seitz H., Horwich MD., Syrzycka M., Honda BM et al.: **Collapse of germline piRNAs in the absence of Argonaute3 reveals somatic piRNAs in flies .** Cell 2009., 137 :509-521.

Liang HL., Nien CY., Liu HY., MM., Kirov N., Rushlow., C., 2008., **The zinc-finger protein Zelda is a key activator of the early zygotic genome in *Drosophila*.** Nature., 456., 400-404.

Lund E., Liu M., Hartley RS., Sheets MD., Dahlberg JE: **Deadenylation of maternal mRNAs mediated by miR-427 in Xenopus laevis embryos.** RNA (New York., NY) 2009., 15:2351-2363.

Malone CD., Brennecke J., Dus M., Stark A., McCombie WR., Sachidanandam R., Hannon GJ: **Specialized piRNA pathways act in germline and somatic tissues of the *Drosophila* ovary .** Cell 2009., 137 :522-535.

Malone CD., Hannon GJ: **Small RNAs as guardians of the genome** . Cell 2009., 136 :656-668.

Memili E., Dominko T., First NL: **Onset of transcription in bovine oocytes and preimplantation embryos.** Mol Reprod Dev1998.,51:36-41.

Merrill., P.T., Sweeton., D., and Wieschaus., E. (1988). **Requirements for autosomal gene activity during precellular stages of *Drosophila melanogaster*.** *Development*, 104 ., 495–509.

Miyara F., Migne C., Dumont-Hassan M., Le Meur A., Cohen-Bacrie P., Aubriot FX., Glissant A., Nathan C., Douard S., Stanovici A., Debey P: **Heterochromatin configuration and transcriptional control in human and mouse oocytes.** Mol Reprod Dev., 2003.,64:458-470.

Nishida KM., Saito K., Mori T., Kawamura Y., Nagami-Okada T., Inagaki S., Siomi H., Siomi MC: **Gene silencing mechanisms mediated by Aubergine piRNA complexes in *Drosophila* male gonad.** RNA 2007., 13:1911-1922.

Nottrott S., Simard MJ., Richter JD: **Human let-7a miRNA blocks protein production on actively translating polyribosomes.** Nat Struct Mol Biol 2006., 13:1108-1114.

Okamura K., Lai EC. **Endogenous small interfering RNAs in animals**. Nat Rev Mol Cell Biol 2008;9:673–8.

Okamura., K., Hagen., J. W., Duan., H., Tyler., D. M. & Lai., E. C. **The mirtron pathway generates microRNA-class regulatory RNAs in *Drosophila* .** Cell 130 ., 89–100 (2007).

Olivieri D., Sykora MM., Sachidanandam R., Mechtler K., Brennecke J: **An in vivo RNAi assay identifies major genetic and cellular requirements for primary piRNA biogenesis in *Drosophila*.** EMBO J 2010., 29:3301-3317.

Petersen CP., Bordeleau ME., Pelletier J., Sharp PA: **Short RNAs repress translation after initiation in mammalian cells.** Mol Cell 2006., 21:533-542.

Pillai., R. S. et al. **Inhibition of translational initiation by let-7 microRNA in human cells.** Science 309 ., 1573–1576 (2005).

Poulson., D.F. (1937). **Chromosomal deficiencies and embryonic development of *Drosophila melanogaster*.** Proc. Natl. Acad. Sci. USA23., 133–137.

Qi H., Watanabe T., Ku HY., Liu N., Zhong M., Lin H: **The Yb body., a major site for Piwi-associated RNA biogenesis and a gateway for Piwi expression and transport to the nucleus in somatic cells.** J Biol Chem 2011., 286 :3789-3797.

Rouget C., Papin C., Boureux A., Meunier AC., Franco B., Robine N., Lai EC., Pelisson A., Simonelig M: **Maternal mRNA deadenylation and decay by the piRNA pathway in the early *Drosophila* embryo .** Nature 2010., 467 :1128-1132.

Rouget C., Papin C., Boureux A., Meunier AC., Franco B., Robine N., Lai EC., Pelisson A., Simonelig M., 2010., **Maternal mRNA deadenylation and decay by the piRNA pathway in the early *Drosophila* embryo.,** Nature., 467., 1128–1132.

Ruby., J. G., Jan., C. H. & Bartel., D. P. **Intronic microRNA precursors that bypass drosha processing.** Nature 448 ., 83–86 (2007).

Sachs., A. B. & Davis., R. W. **The poly(A) binding protein is required for poly(A) shortening and 60S ribosomal subunit-dependent translation initiation.** Cell 58., 857–867 (1989).

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

Saito K., Ishizu H., Komai M., Kotani H., Kawamura Y., Nishida KM., Siomi H., Siomi MC: **Roles for the Yb body components Armitage and Yb in primary piRNA biogenesis in *Drosophila***. Genes Dev 2010., 24:2493-2498.

Saito K., Nishida KM., Mori T., Kawamura Y., Miyoshi K., Nagami T., Siomi H., Siomi MC: **Specific association of Piwi with rasiRNAs derived from retrotransposon and heterochromatic regions in the *Drosophila* genome.** Genes Dev 2006., 20:2214-2222.

Schier AF: **The maternal-zygotic transition: death and birth of RNAs**. Science 2007.,316:406-407.

Schubert C. M., Lin R., de Vries C. J., Plasterk R. H., Priess J. R. (2000). **MEX-5 and MEX-6 function to establish soma/germline asymmetry in early C. elegans embryos.** *Mol. Cell*, 5:, 671-82.

Semotok JL., Luo H., Cooperstock RL., Karaiskakis A., Vari HK., Smibert CA., Lipshitz HD: ***Drosophila* maternal Hsp83 mRNA destabilization is directed by multiple SMAUG recognition elements in the open reading frame.** Mol Cell Biol 2008., 28:6757-6772.

Smibert C. A., Wilson J. E., Kerr K., Macdonald P. M. (1996). **smaug protein represses translation of unlocalized nanos mRNA in the Drosophila embryo.** Genes Dev 1996, 10:2600-2609.

Siomi MC., Sato K., Pezic D., Aravin AA: **PIWI-interacting small RNAs: the vanguard of genome defence .** Nat Rev Mol Cell Biol 2011., 12:246-258.

Siomi H., Siomi MC. **On the road to reading the RNA-interference code**. Nature 2009;457:396–404.

Sheth U., Roy P. (2003). **Decapping and Decay of Messenger RNA Occur in Cytoplasmic Processing Bodies.** *Science*, **300**(5620), 805–8.

Tadros W., Goldman AL., Babak T., Menzies F., Vardy L., Orr-Weaver T., Hughes TR., Westwood JT., Smibert CA., Lipshitz HD: **SMAUG is a major regulator of maternal mRNA destabilization in *Drosophila* and its translation is activated by the PAN GU kinase.** Dev. Cell 2007., 12:143-155.

Tadros., W., Goldman., A.L., Babak., T.,Menzies., F., Vardy., L., Orr-Weaver., T., Hughes., T.R., Westwood., J.T., Smibert., C.A., and Lipshitz., H.D. (2007). **SMAUG is a major regulator of maternal mRNA destabilization in *Drosophila* and its translation is activated by the PAN GU kinase.** Dev. Cell 12., 143–155.

Tadros., W., Houston., S.A., Bashirullah., A., Cooperstock., R.L., Semotok., J.L., Reed., B.H., and Lipshitz., H.D. (2003). **Regulation of maternal transcript destabilization during egg activation in *Drosophila*.** Genetics 164 ., 989–1001.

Tam OH et al (2008) **Pseudogene derived small interfering RNAs regulate gene expression in mouse oocytes.** Nature 453: 534–538.

Tang F., Kaneda M., O'Carroll D., Hajkova P., Barton SC., Sun YA., Lee C., Tarakhovsky A., Lao K., Surani MA: **Maternal microRNAs are essential for mouse zygotic development.** Genes Dev 2007., 21:644-648.

Tchurikov NA., Kretova OV (2011) **Both piRNA and siRNA Pathways Are Silencing Transcripts of the Suffix Element in the *Drosophila melanogaster* Germline and Somatic Cells.** PLoS ONE 6(7): e21882.

Thompson, D.M., Lu, C., Green, P.J., and Parker, R. (2008) tRNA cleavage is a conserved response to oxidative stres in eukaryotes. *RNA* 14, 2095-2103.

Thompson, D.M., and Parker, R. (2009). **The RNase Rny1p cleaves tRNAs and promotes cell death during oxidative stres in *Saccharomyces cerevisiae*.** *J Cell Biol* 185, 43-50.

Vilcek J**. Fifty years of interferon research: aiming at a moving target.** Immunity 2006;25:343–348.

Voinnet O. Use**., Tolerance and avoidance of amplified RNA silencing by plants**. Trends Plant Sci 2008;13:317–328.

Wang., B., Yanez., A and Novina., C D., 2008., **MicroRNA-repressed mRNAs contain 40S but not 60S components.** PNAS., 105(1)4 5343-5348.

Watanabe Ti ., Totoki Y., Toyoda A., Kaneda M., Kuramochi-Miyagawa S., Obata Y., Chiba H., Kohara Y., Kono T., Nakano T., Surani MA., Sakaki Y., Sasaki H., **Endogenous siRNAs from naturally formed dsRNAs regulate transcripts in mouse oocytes.,** 2008., Nature., 453., 539-544.

Wieschaus., E., and Sweeton., D. (1988). **Requirements for X-linked zygotic activity during cellularization of early *Drosophila* embryos.** Dev. 104 ., 483–493.

Williams B. R**. PKR; a sentinel kinase for cellular stress**. Oncogene 1999;18:6112–6120.

Wu E., Thivierge C., Flamand M., Mathonnet G., Vashisht AA., Wohlschlegel J., Fabian MR., Sonenberg N., Duchaine TF: **Pervasive and cooperative deadenylation of 3' UTRs by embryonic microRNA families** . Mol Cell 2010., 40:558-570.

Wu., L., Belasco., J G., 2008., **Let Me Count the Ways: Mechanisms of Gene Regulation by miRNAs and siRNAs.,** Molecular Cell., 29., 1-7.

Yingdee Unhavaithaya .,YiHao ., Ergin Beyret., Hang Yin., Satomi Kuramochi-Miyagawa., Toru Nakano and Haifan Lin., **MILI., a PIWI-interacting RNA-binding Protein., Is Required for Germ Line Stem Cell Self-renewal and Appears to Positively Regulate Translation.** The Jou. of Biologic. Chem., 284., NO. 10., pp. 6507–6519. 2009.

Zamore PD., Tuschl T., Sharp PA., Bartel DP. **RNAi: double-stranded RNA directs the ATP-dependent cleavage of mRNA at 21 to 23 nucleotide intervals.** Cell 2000;101:25–33.

Zeng F., Schultz RM: **RNA transcript profiling during zygotic gene activation in the pre-implantation mouse embryo** . *Dev Biol* 2005., 283 :40-57.