Dissertation zur Erlangung des Doktorgrades
der Fakultät für Chemie und Pharmazie
der Ludwig-Maximilians-Universität München

The role of small RNAs in cell cycle regulation and transposon defense in *Drosophila melanogaster*

Milijana Mirkovic-Hösle

aus Uzice, Serbien

2013

Erklärung

Diese Dissertation wurde im Sinne von § 7 der Promotionsordnung vom 28. November 2011 von Herrn Professor Dr. Klaus Förstemann betreut

Eidesstattliche Versicherung

Diese Dissertation wurde eigenständig und ohne unerlaubte Hilfe erarbeitet.

München, den 08.02.2013

_____

Milijana Mirkovic-Hösle

Dissertation eingereicht am 08.02.2013

1. Gutachter: Prof. Dr. Klaus Förstemann
2. Gutachter: Prof. Dr. Mario Halic

Mündliche Prüfung am 26.03.2013

1 SUMMARY

Small RNAs like microRNAs (miRNAs) and endogenous short interfering RNAs (endo-siRNAs) are 21 to 23 nucleotide long single-stranded molecules that are involved in post-transcriptional gene silencing (PTGS) by inhibiting protein translation or by inducing the degradation of messenger RNA. Small RNA expression profiling in genome-wide studies revealed miRNAs as important players during the progression of the cell cycle. They repress target mRNAs and can either act pro-carcinogenic to accelerate the cell cycle or anti-carcinogenic to slow down the cell cycle. Furthermore, the RNAi pathway is implicated in heterochromatin-based transcriptional gene silencing (TGS) thereby affecting different steps during the cell cycle.

We showed in the cultured *Drosophila* S2 cell line that depletion of miRNA biogenesis factors e.g. *ago1* and *drosha* resulted in changes in the cell cycle distribution. This was in contrast to siRNA biogenesis factors, where we observed no changes. Deep sequencing of cell cycle synchronized S2 cells demonstrated that miRNAs do not vary in abundance during the cell cycle. The global miRNA analysis was limited due to adapter and bar coding bias in small RNA libraries but despite this technical issue, qRT-PCR as an independent method confirmed the absence of miRNA oscillation during the cell cycle. The same conclusion was drawn for endo-siRNA expression profiles. Further studies of small RNAs within the DNA repair pathway might elucidate their implication with regard to the cell cycle.

Internal serious threats for genome integrity are transposable elements (TEs). *Drosophila melanogaster* has two silencing mechanisms that repress TEs expression: endo-siRNAs and Piwi-interacting small RNAs (piRNAs). The biogenesis of endo-siRNAs involves Loqs-PD, which acts predominantly during processing of dsRNA by Dcr-2, and R2D2 that primarily loads siRNAs into Ago2. With the help of mutant flies, we demonstrated that during both biogenesis steps R2D2 and Loqs-PD function at least partially redundant. We could not reveal a common principle why certain transposons differ in their requirements for Loqs-PD and R2D2 but we could show that their dependence is neither based on the abundance of small RNAs, nor on specific transposon classes, nor on their presence in specific master control loci. Furthermore, the endo-siRNA biogenesis pathway in the germline operates according to the same principles as the existing model for the soma, and its impairment does not significantly affect piRNAs. Expanding the analysis, we revealed that miRNA* as well as exo-siRNAs were also loaded into RISC at least partially independently of R2D2.

Finally, we confirmed the occurrence of somatic piRNA-like RNAs (pilRNAs) that show a ping-pong signature resembling the one of their germline relatives. No correlation was noticed between endo-siRNA and pilRNA pathway in *Drosophila* S2 cells. Further tissue-specific studies might elucidate the origin of pilRNAs as well as their biogenesis pathway.

2 INTRODUCTION

2.1 Classes of small RNAs

The discovery of small RNA silencing in the late 1990s revealed a so far unknown field in understanding of RNAs as regulatory molecules (Fire, Xu et al. 1998). The most important families of small RNA guides include microRNAs (miRNAs), small interfering RNAs (siRNAs) and piwi-interacting RNAs (piRNAs). These different classes of regulatory RNAs differ in the proteins required for their biogenesis, their modes of target regulation and the biological functions in cellular processes. There is a high conservation from the yeast *S. pombe* to plants and animals but not bacteria or archea (Ghildiyal, Seitz et al. 2008; Ghildiyal and Zamore 2009).

2.2 Biogenesis of small RNAs

2.2.1 miRNAs

miRNAs are encoded in the genome and derive from precursor transcripts called primary miRNAs (pri-miRNAs) which are ubiquitously transcribed by RNA polymerase II (Figure 2.1A). First the pri-miRNA is cleaved in the nucleus by the RNase III endonuclease Drosha and dsRNA-binding domain (dsRBD) partner Pasha generating long precursor miRNA (pre-miRNA) about 70 nt in length (Lee, Jeon et al. 2002; Lee, Nakahara et al. 2004; Kim, Han et al. 2009). The resulting pre-miRNA has a partially base-paired stem with a single-stranded loop structure. The nuclear export protein Exportin 5 carries the pre-miRNA to the cytoplasm through the nuclear pore (Yi, Qin et al. 2003; Bohnsack, Czaplinski et al. 2004; Aleman, Doench et al. 2007). It is then further processed by removing the hairpin loop via a complex containing the RNase III enzyme Dicer-1 with its dsRBD partner protein Loquacious (Loqs, isoform Loqs-PB) (Grishok, Pasquinelli et al. 2001; Hutvagner, McLachlan et al. 2001; Ketting, Fischer et al. 2001; Forstemann, Tomari et al. 2005; Jiang, Ye et al. 2005; Saito, Ishizuka et al. 2005; Park, Liu et al. 2007). A 20-24 nt mature miRNA/miRNA* duplex is then loaded into an effector complex with Argonaute proteins named RNA-induced silencing complex (RISC). In flies the biogenesis of small RNAs is uncoupled from the subsequent loading step into RISC (Forstemann, Horwich et al. 2007). Loading is governed by the structure of the duplex. The majority of miRNA duplexes carries bulges and mismatches and therefore binds Ago1. After loading, one strand named miRNA star (miRNA*) is discarded resulting in formation of mature RISC. Strand selection is determined by the thermodynamic properties of the small RNA duplex. Depending on the specific Argonaute protein within RISC and the extent of complementarity between the miRNA and the mRNA, gene silencing is performed by inhibition of translation or induction of degradation of the mRNA (Ghildiyal and Zamore 2009).

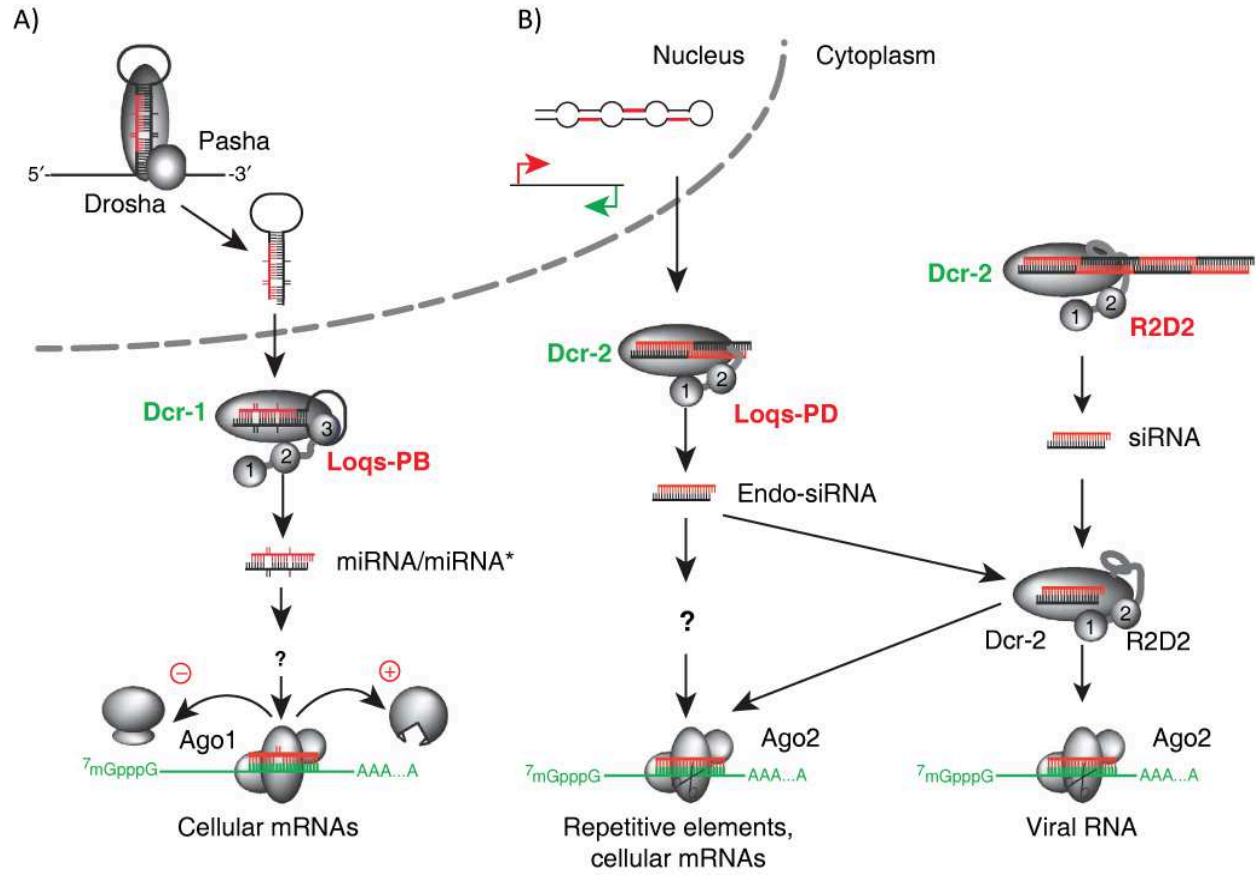


Figure 3.1 The biogenesis of miRNAs and siRNAs.

(A) miRNA pathway: miRNAs are encoded in the genome and are transcribed by RNA polymerase II into primary transcript (pri-miRNA) which is cleaved by the RNaseIII enzyme Drosha together with Pasha. The resulting pre-miRNA is exported into the cytoplasm and further processed by a second RNaseIII enzyme Dcr-1 with dsRBD protein Loqs-PB into double stranded duplex miRNA/miRNA* with partial complementarity. This is loaded usually into RISC comprising Ago1. miRNA guide strand is selected, the duplex is unwound and miRNA* degraded. After binding of mature RISC to its target sequence, silencing is executed by translational repression.

(B) siRNA pathway (left: endo-siRNA; right: exo-siRNAs): Dcr-2 generate endo- and exo-siRNAs. Endogenous RNAs are encoded in the genome and derive from long hairpin transcription of inverted repeats, convergent transcription or antisense transcription. Exo-siRNAs derive from long dsRNAs generated during viral replication or experimentally introduced long dsRNAs. Loqs-PD is required for the production of endo-siRNAs while R2D2 participate in the production of exo-siRNAs. The resulting perfectly complementary duplex is loaded via RLC (Dcr-2 and R2D2) into RISC containing Ago2. After the generation of mature RISC by cleavage of the passenger strand, siRNA are methylated at its 3' end. Silencing is executed by cleavage of targets with perfect complementarity. Figure adapted from (Hartig, Esslinger et al. 2009).

2.2.2 siRNAs

Small interfering RNAs (siRNAs) are divided into two classes depending on their origin. siRNAs with endogenous origin are named endo-siRNAs while exogenously derived siRNAs are referred as exo-siRNAs. Endo-siRNAs can be produced from RNA transcripts with extensive hairpin structures (Okamura, Chung et al. 2008), convergent transcription units (Czech, Malone et al. 2008; Okamura, Balla et al. 2008; Okamura and Lai 2008) or from the annealing of sense and antisense RNAs from unlinked loci (reviewed in (Berretta and Morillon 2009). The long double-stranded precursors of exo-siRNAs derive from viral replication intermediates or are introduced into the cell to induce RNA interference experimentally (Liu, Rand et al. 2003; Lee, Nakahara et al. 2004; Ghildiyal and Zamore 2009).

All siRNAs are derived from long double-stranded perfectly complementary RNA precursors (Czech, Malone et al. 2008; Okamura, Chung et al. 2008; Okamura and Lai 2008). Endo-siRNAs are processed through cleavage by Dicer-2 in collaboration with its dsRBD co-factor Loqs (isoform Loqs-PD) into 21 nt siRNA duplexes in contrast to exo-siRNAs which depend on Dcr-2 and another dsRBD protein R2D2 (Hartig, Esslinger et al. 2009; Zhou, Czech et al. 2009). The sorting step is executed by RISC-loading-complex (RLC), consisting of Dcr-2 and R2D2. It senses the thermodynamic asymmetry of the siRNA duplex (Liu, Rand et al. 2003; Tomari, Matranga et al. 2004). R2D2 serves as the differentiating factor by binding the more stable 5’ end because the more unstable 5’ end is usually selected as the mature guide strand (Khvorova, Reynolds et al. 2003; Schwarz, Hutvagner et al. 2003). After this, RNA duplexes with a high degree of basepairing is delivered to RISC containing Ago2. Ago2 cleaves the passenger strand and the endonuclease C3PO converts pre-RISC into mature RISC by degrading the nicked passenger fragments (Liu, Ye et al. 2009; Ye, Huang et al. 2011). In *Drosophila* the guide strand is 2’-*O*-methylated at its 3’-end by the S-adenosyl-methionine (SAM)-dependent methyltransferase DmHen1 to increase its stability (Horwich, Li et al. 2007; Saito, Sakaguchi et al. 2007). Finally the mature RISC binds the target sequence via perfect complementarity and induces cleavage of the mRNA with its own endonucleolytic activity.

Subsequent studies of endo- and exo-siRNAs yielded some controversial findings. Recently, *in vitro* processing and loading reactions of dsRNA in fly mutant embryo extracts implicated Loqs to function in dsRNA-triggered silencing. Furthermore RNAi was impaired in both *loqs* and *r2d2* mutants after injection of dsRNA into embryos as well as in case of artificial endogenous dsRNA corresponding to one part of the *white* (*w*) gene. Hence Loqs-PD was suggested to associate the processing of endo- as well as exo-siRNA while the products of both siRNAs are loaded by Dcr-2/R2D2 into Ago2-complex. Thus a sequential model was proposed separating the processing from the loading in a common pathway triggered by either exogenous or endogenous dsRNA (Marques, Kim et al. 2010). In contrast to the previous study, the next

study analyzed Loqs-PD, a specific Loqs isoform, which allowed separating effects coming from miRNA or endo-siRNA biogenesis. In cultured *Drosophila* cells depletion of R2D2 resulted in an increased repression of GFP-based endo-siRNA reporter while depletion of Loqs-PD increased RNAi efficiency (Hartig, Esslinger et al. 2009; Hartig and Forstemann 2011). Therefore R2D2 and Loqs-PD are functional antagonists during both endo- and exo-siRNA mediated silencing resulting in the competition of both proteins for Dcr-2 binding. Their mutually antagonistic activities are incompatible with the sequential model. Furthermore deep sequencing data of fly mutants showed that the nucleolytic processing of hairpin-derived and transposon-targeting endo-siRNAs does not depend on R2D2. Next in *r2d2* mutant endo-siRNAs are reduced but still detectable and retain their thermodynamic asymmetry showing that the RISC loading complex Dcr-2/R2D2 can be bypassed for some substrates. All in all, both endo- and exo-siRNA pathways are proposed to proceed at least partially in parallel (Hartig and Forstemann 2011).

2.2.3 piRNAs

piRNA biogenesis is distinct from the other small RNA silencing pathways since it is Dicer-independent. Instead, piRNAs associate with the germline-specific Piwi clade of Argonaute proteins (Aravin, Gaidatzis et al. 2006; Girard, Sachidanandam et al. 2006; Lau, Seto et al. 2006; Vagin, Sigova et al. 2006; Brennecke, Aravin et al. 2007; Gunawardane, Saito et al. 2007; Batista, Ruby et al. 2008). They comprise of Piwi, Aubergine (Aub) and Ago3 in flies and perform processing and target cleavage in a self-amplifying feed-forward mechanism (Brennecke, Aravin et al. 2007; Gunawardane, Saito et al. 2007). In contrast to miRNAs and siRNAs, piRNAs arise from long single-stranded precursor RNAs. Sense piRNAs originating from transposon mRNAs associate with Ago3 complex and bind long antisense RNA transcripts mostly derived from clusters of selfish genetic elements, named piRNA master loci (Figure 2.2). Hereafter the long precursor is cleaved by guiding formation of the 5’ end of antisense piRNAs which then associates with Aub and Piwi and vice versa resulting in a 10 nt overlap of corresponding piRNAs. They are finally 2’-*O*-methylated at their 3’ termini, unlike miRNAs but similar to siRNAs in flies. Piwi- and Aub-interacting piRNAs have a U bias at their 5’ end while Ago3-interacting piRNAs often have A at the 10th nucleotide from their 5’ end based on the overlap during biogenesis. These characteristics are referred to as the ping-pong signature. The initiating triggers for of the ping-pong cycle are supplied by maternally deposited or primary processed piRNAs. Little is known about the primary processing at mechanistic level.

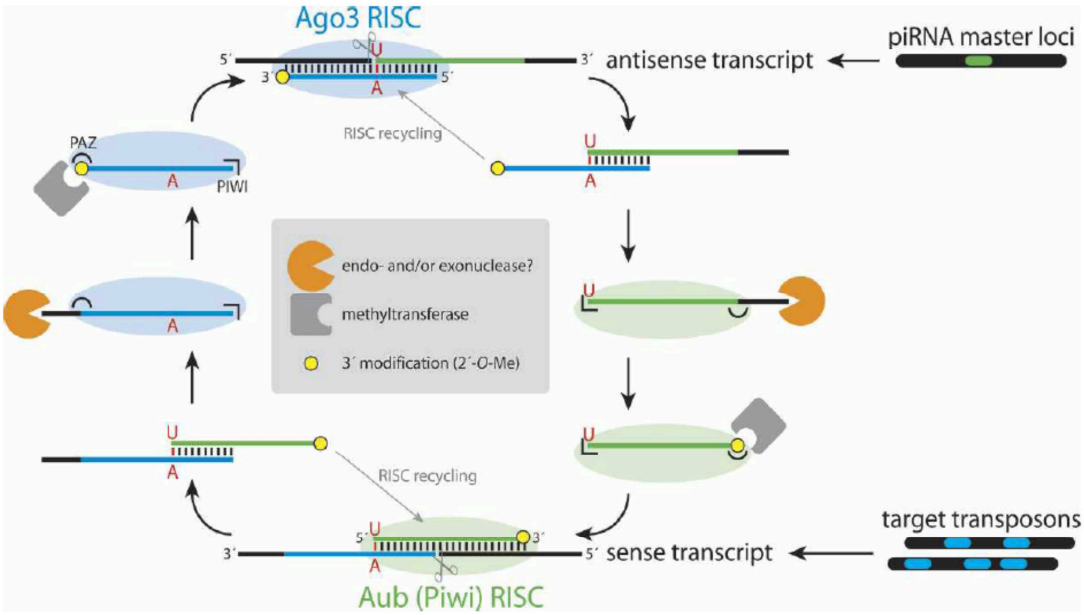


Figure 3.2 The ping-pong model for piRNA biogenesis.

Piwi- and Aub-RISC complexes are loaded with a piRNA guide. They cleave sense transcripts derived from transposon mRNA. The cleaved transcript is further used to program Ago3-RISC complex which in turn cleaves the antisense transcripts that originate from the master control loci. Again, the cleaved RNA serves to program Piwi or Aub-RISC serving for the amplification cycle. 5' ends of piRNAs are defined by cleavage, while 3' ends are shortened by a 3'-exonuclease. Subsequently piRNAs are stabilized by 2'-*O*-methylation at the 3' end. Figure adapted from (Hartig, Tomari et al. 2007).

The PIWI proteins show different subcellular localization and expression patterns suggesting distinct roles (Cox, Chao et al. 2000; Nishida, Saito et al. 2007). In the context of the cell, Piwi is localized in the nucleus contrary to Aub and Ago3 (Cox, Chao et al. 2000). Its role in primary silencing appears to be in the nucleus (Saito, Ishizu et al. 2010). A deeper look at different tissue types showed that Piwi can be detected in somatic and germ cells in ovaries whereas Aub and Ago3 are absent in gonadal somatic tissue (Cox, Chao et al. 2000; Brennecke, Aravin et al. 2007; Nishida, Saito et al. 2007; Saito, Ishizu et al. 2010). Consistently, piRNAs in the somatic support cells of the gonad do not show the ping-pong signature. Therefore, these piRNAs seem to be generated exclusively via primary processing carried out by Piwi.

Are any piRNAs present outside the germ line? Somatic piRNA-like small RNAs (pilRNAs) have been observed in *ago2* mutant fly heads (Ghildiyal, Seitz et al. 2008). Next, specific pilRNAs were described in human Natural Killer (NK) cells implicated in transcriptional silencing of Ig-like receptros (Cichocki, Lenvik et al. 2010). Very recently (Yan, Hu et al. 2011) showed a widespread presence of pilRNAs in various somatic tissues of fly, mouse and rhesus macaque samples. They displayed all known characteristics of piRNA based

on their sequence feature and genomic origin and resemble germline piRNAs rather than primary piRNAs found in fly somatic ovarian follicle cells and a derived cell line, ovarian somatic sheet (OSS) cells.

2.3 The significance of small RNAs

miRNAs regulate many key biological processes, including developmental, differentiation, cell growth, death, metabolic homeostasis, oncogenesis and memory (reviewed in (Ghildiyal and Zamore 2009; Kim, Han et al. 2009); (Grishok, Pasquinelli et al. 2001; Ketting, Fischer et al. 2001; Bernstein, Kim et al. 2003; Lee, Seong et al. 2004; Poy, Eliasson et al. 2004; Harfe 2005; Kanellopoulou, Muljo et al. 2005; Li and Carthew 2005; Ashraf and Kunes 2006; Schratt, Tuebing et al. 2006; Teleman, Maitra et al. 2006). 1% of genomic transcripts in mammalian cells are estimated to encode for miRNAs while nearly one third of the encoded genes are regulated by miRNAs. This simply demonstrates their integral role in genome-wide regulation of gene expression (Stark, Brennecke et al. 2003; Rajewsky and Socci 2004; Krek, Grun et al. 2005; Lewis, Burge et al. 2005).

Cellular defense against viruses evolved formation of exo-siRNAs which are produced from viral long dsRNA. This machinery can be exploited for RNA interference (RNAi) to monitor the gene silencing by artificial introduced double-stranded RNA into the organism (reviewed in (Ghildiyal and Zamore 2009; Kim, Han et al. 2009).

The internal serious threats are transposable elements. They insert themselves at other non-homologous regions in the genome and their transposition can cause mutations, deletions, duplications and changes in gene expression at the site of insertion or in nearby genes. piRNAs counteract the mobilization of transposons in the germline and maintain the genomic integrity of the offspring (reviewed in (Hartig, Tomari et al. 2007; Malone, Brennecke et al. 2009). The endo-siRNA pathway contributes to transposon silencing through control of transposon activity in somatic tissues and germline. Furthermore, they are involved in the regulation of cellular gene expression (Chung, Okamura et al. 2008; Czech, Malone et al. 2008; Ghildiyal, Seitz et al. 2008; Kawamura, Saito et al. 2008).

2.3.1 The role of small RNAs and their silencing machineries in the cell cycle

Most small RNAs mentioned above are involved in post-transcriptional gene silencing (PTGS) by inhibiting protein translation or degradation of messenger RNA. This endows them with an essential role in the regulation of gene expression. miRNA expression profiling in genome-wide studies showed that miRNA expression levels are altered in primary human tumors (Calin, Liu et al. 2004; Lu, Getz et al. 2005). Unique miRNA signatures are associated with different types of tumors and various miRNA species are differentially

expressed in cancer cells which allow identifying a number of miRNAs with potential diagnostic and prognostic applications (Lu, Getz et al. 2005; Calin and Croce 2006). miRNA gene deletions or amplifications inhibit tumor suppressor genes or inappropriately activate oncogenes initiating the cancer process by uncontrolled cell proliferation (Cho 2007). The general model for their implication in the cell cycle is that by repressing target mRNAs, miRNAs act pro-carcinogenic and accelerate the cell cycle e.g. via inhibition of the cyclin dependent kinase inhibitor p21CIP through miR-106 (Ivanovska, Ball et al. 2008). Otherwise miRNA can act also anti-carcinogenic and slow down the cell cycle e.g. via repression of Ras by let-7 (Johnson, Grosshans et al. 2005).

The RNAi pathway is furthermore involved in transcriptional gene silencing (TGS) executed by the RNA-induced transcriptional gene silencing (RITS) complex in the nucleus (Verdel and Moazed 2005). Many publications implicated RITS in the assembly of a repressive chromatin structure called heterochromatin by promoting DNA or histone modifications (Mochizuki, Fine et al. 2002; Reinhart and Bartel 2002; Cam, Sugiyama et al. 2005). In mouse stem cells, Dicer is involved in the maintenance of centromeric heterochromatin structure and centromeric silencing (Kanellopoulou, Muljo et al. 2005). In *Drosophila*, components of the RNAi pathway, like Piwi, Aubergine and Spindle-E, have been as well implicated in heterochromatin-mediated gene silencing (Pal-Bhadra, Bhadra et al. 1999; Pal-Bhadra, Bhadra et al. 2002; Kogan, Tulin et al. 2003; Pal-Bhadra, Leibovitch et al. 2004). Next, In *S. pombe* where no miRNAs have been identified, Dicer and the RNAi pathway have been implicated in the generation of heterochromatic siRNAs that mediate TGS of centromeric repeats (Hall, Shankaranarayana et al. 2002; Reinhart and Bartel 2002; Volpe, Kidner et al. 2002; Volpe, Schramke et al. 2003; Verdel and Moazed 2005).

In addition, heterochromatin is crucial for functional organization of chromosomal structures such as centromeres and telomeres. Its important function is to protect genome integrity by maintaining the repetitive DNA elements inert and inhibiting potentially mutagenic transposition events (Plasterk 2002). Mutations within the RITS pathway affect centromere function by defects in sister-chromatid cohesion and chromosomal segregation (Volpe, Schramke et al. 2003). In the yeast *Schizosaccharomyces pombe* as well as in the trypanosome *Trypanosoma brucei* Ago1 and Dcr1 are critical for proper mitosis (Durand-Dubief and Bastin 2003; Volpe, Schramke et al. 2003). *Ago2* mutants in *Drosophila melanogaster*, which are defective in siRNA silencing but still viable and fertile, show a number of abnormalities, among them the peculiar fact that in early embryogenesis, the nuclear replication and division cycles become asynchronous despite a common cytoplasm (Deshpande, Calhoun et al. 2005). Furthermore Piwi (Cox, Chao et al. 1998), as well as Ago1 (Yang, Chen et al. 2007) both mediate a somatic signalling mechanism to regulate the division and maintenance of germline stem cells (GSCs) in *Drosophila*. The same was demonstrated for Loquacious (isoform Loqs-PB), a dsRBD protein required for processing of pre-miRNA (Forstemann, Tomari

et al. 2005). Hatfield et al. have shown that the RNase III-family enzyme Dcr-1 is essential for stem cells to bypass the normal G1/S checkpoint implicating that the miRNA pathway might be part of a mechanism that makes stem cells insensitive to environmental signals that normally stop the cell cycle at the G1/S transition. Given that aforementioned factors serve as key components of the RNAi pathway, all findings strongly support the idea that small RNA silencing participates in the regulation of the cell cycle.

The cell cycle of proliferating cells is comprised of chromosome condensation with the subsequent cell division mitosis (M), quiescent stage referred to as G1, which is followed by the DNA synthesis referred as S phase (S) and the second period of apparent quiescence G2.

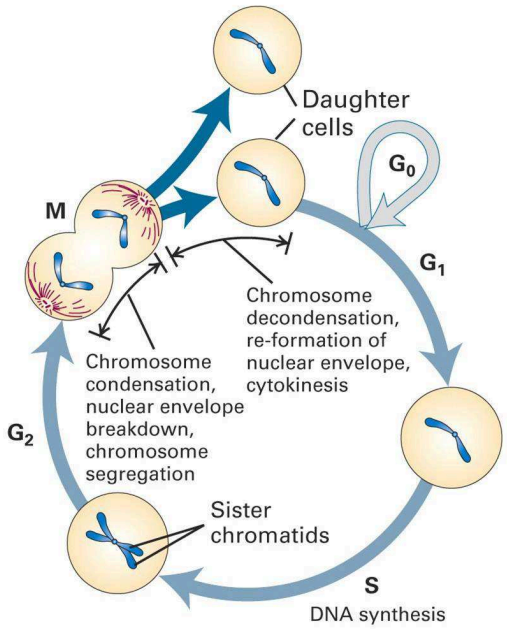


Figure 3.3 Cell division cycle.

The cell cycle also named cell division cycle is divided in two periods: interphase and mitosis. During the interphase known as preparatory phase the cell grows, accumulate nutrients and duplicate (replicate) the DNA. It proceeds in three stages G1, S and G2. During mitosis (M) the mother cell is divided into two daughter cells genetically identical to each other and to their parent cell. The nuclear envelope breaks down, the pairs of chromosomes condense and attach to fibers which pull the sister chromatids to opposite sides of the cell. The nuclei, cytoplasm, organelles and cell membrane are equally divided. G_0 is termed post-mitotic and used to refer to both quiescent and senescent nonproliferative cells. Figure adapted from: The cell cycle: Principles of Control by David O Morgan 2007.

Studies of regulatory mechanisms and understanding of cell cycle events require methods for isolating cells at specific positions of the cell cycle and determination where in the cell cycle they reside. There are various techniques based on either a chemical or a physical strategy. By means of chemicals cells are arrested at a certain stage of cell cycle (e.g. inhibition of DNA replication in S phase with hydroxyurea, thymidine, methotrexate or aphidicolin and collected at the time they enter the stage of interest (Tobey and Crissman 1972; Vogel, Schempp et al. 1978; Fox, Read et al. 1987; Matherly, Schuetz et al. 1989). The

major disadvantage of chemical usage is that the metabolism of cells is often perturbed, which makes it difficult to ensure no drug artifacts affecting the experimental observations. Counterflow centrifugal elutriation (CCE) is one of various physical methods which fractionate the cell population regarding to sedimentation properties influenced primarily by the size of the particle, whereas the effect of density is much smaller (Lindahl 1948; Lindahl 1956; Sanderson, Bird et al. 1976). Cell growth takes place in S phase when the cell size and DNA content increase synchronously. Consequently, cell size generally correlates with cell cycle stage and centrifugal elutriation is the method of choice to separate cells in various stages with minimal perturbation of cellular functions.

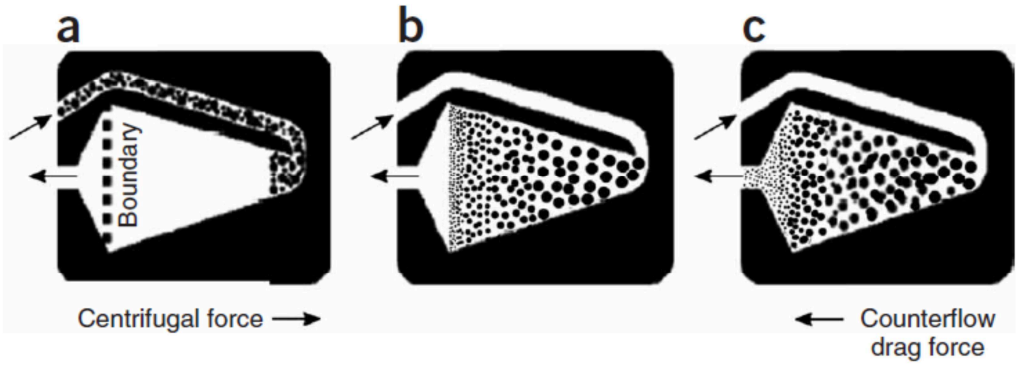


Figure 3.4 Schematic view of cell synchronization by centrifugal elutriation.

(a) Unfractionated asynchronous cells are loaded into the elutriation chamber. (b) The centrifugal and the drag forces are acting in opposite directions. As the size gradient is balanced by the centrifugal force and the counterflow of the elutriation fluid the cells stay inside the chamber. The smaller sized cells line up toward the elutriation boundary within the chamber whereas larger, faster-sedimenting cells migrate to an area of the chamber having the greatest centrifugal force. (c) Increasing flow rate of elutriation fluid result first in elution of smaller cells while the larger ones follow. Figure adapted from (Banfalvi 2008).

2.3.2 The cellular defense against transposable elements

Mobile genetic elements, also known as transposons, are selfish nucleic acids that replicate and proliferate parasitically in a host genome. They can be either DNA-based, resulting in a cut-and-past mode of insertion, or belong to the class of retrotransposons, which generate an RNA intermediate resulting in a copy-and-paste mode of integration. The latter are further classified into LTRs (long terminal repeats), LINES (long interspersed elements) and SINEs (short interspersed elements) differing in transcription by RNA polymerase II or III and encoding of their own reverse transcriptase. The insertion of a transposable element into a new genomic location can have profound effects not only if they hit the coding sequence of a gene but also by modifying the local expression pattern. Most transposition events are believed to have detrimental consequences, but the host species can also benefit from selfish DNA. In bacteria, transposons can transfer antibiotic resistance genes (Bennett 2008). More general, transposon mobility contributes to

genetic variation within a population upon which evolutionary selection may act and the species can adapt. This can be rationalized if e.g. the promotor activity associated with some transposon sequences confers a selective advantage. Other examples of beneficial transposon activity include the participation of transposase enzyme in genome rearrangements during macronuclear development of *Oxytricha trifallax* (Nowacki, Higgins et al. 2009) or the domestication of the *Het-A*, *Tart* and *Tahre* retrotransposons for telomere maintenance in *Drosophila melanogaster* and certain other insects.

To restrict the activity of transposons, several defense mechanisms can be deployed by the host organism. Their common principle is to prevent the generation or accumulation of transposon transcripts, which either occur as an intermediate of transposition or serve as mRNA encoding the transposase enzyme of DNA-based transposons. Germ line cells and the derived oocytes are maternally primed with highly abundant and diverse 24 to 30 nt long piRNAs that associate with a Piwi-family protein. These ribonucleoprotein complexes can both cleave perfectly complementary transposon RNAs and instruct the formation of transcriptionally silenced heterochromatin at corresponding genomic loci (Hall, Shankaranarayana et al. 2002; Paddison, Caudy et al. 2002; Pal-Bhadra, Bhadra et al. 2002; Volpe, Kidner et al. 2002). Generation of piRNAs requires active transposon mRNA transcription as their biogenesis relies on single stranded sense and antisense transcripts originating from transposon loci and so-called master regulatory regions in the host genome, respectively. piRNAs are very efficient but very slow to adapt towards a new transposons threat. Crossing a naïve female fly with a male fly carrying a new transposon results in sterile offspring referred as hybrid dysgenesis syndrome (Picard, Bucheton et al. 1972; Bucheton, Lavige et al. 1976; Khurana, Wang et al. 2011). Crossing in the opposite orientation has no phenotype due to the inheritance of a pool of maternally transmitted piRNAs (Blumenstiel and Hartl 2005; Brennecke, Malone et al. 2008). The phenotype can eventually be overcome due to insertion of the new element into transposon into one of the master regulatory regions and the generation of new piRNAs (Ronsseray, Lehmann et al. 1991; Ronsseray, Lehmann et al. 1996; Todeschini, Teysset et al. 2010; Khurana, Wang et al. 2011; Kawaoka, Mitsutake et al. 2012).

In addition to piRNAs, 21 nt long siRNAs that target transposons are generated. The trigger for production of siRNAs is the generation of dsRNA. In the germ line, piRNA and siRNA biogenesis therefore both utilize a combination of sense and antisense transcripts, potentially competing for precursors. Whether such competition occurs *in vivo* has not been analyzed so far. The endogenous siRNA pathway is highly active in somatic and germ line cells of *Drosophila melanogaster*. While in the case of piRNAs the targeted sequences are primarily defined by the presence of corresponding sequences in the master control loci, our understanding of how the cells decide to deploy siRNAs against transposons is very limited. Nonetheless, siRNA are the first defense mechanism that responds to a new transposon challenge (Hartig, Esslinger et al.

2009) and they are generated against artificial high copy transgenes in cell culture, implying that neither pre-existing genomic template nor passage through the germ line appear to be required. While genetic deficiency in piRNA pathway components leads to sterility, the RNAi pathway, which can be genetically separated from the miRNA pathway in *Drosophila*, is not essential. Therefore, although siRNAs also target transposons, they cannot compensate for the loss of piRNAs.

2.4 Deep sequencing – ligase dependent approach

A revolution for the discovery of new RNAs and transcriptome expression profiling came trough methods developed for massive parallel sequencing referred as deep sequencing. Their advantage is that no requirement of prior sequence information limits the analysis, in contrast to microarray analysis. At present, four different technologies are mostly used: 454 (Roche), Solexa (Illumina), SOLiD (Life Technologies) and Ion Torrent (Life Technologies). In this thesis, we preformed deep sequencing based on the Solexa technology. The generation of small RNA libraries includes sequential ligation of adapter oligonucleotides by usage of RNA ligases (ΔRnl 2, Rnl 1) that introduce primer-binding sites for subsequent reverse transcription (RT) and PCR amplification prior to deep sequencing (Figure 2.5A).

RNA ligases were originally discovered in bacteriophage T4 (Silber, Malathi et al. 1972). Under T4 phage attack, bacteria nick their tRNA to block translation while thereafter T4 phage ligases together with a polynucleotide kinase repair the nick (Amitsur, Levitz et al. 1987; Ho and Shuman 2002). Two ligase enzymes are encoded in T4 phage, Rnl 1 and Rnl 2. Both enzymes mediate tRNA repair but they differ structurally in the nucleotidyl transferase domain (Pascal 2008).

RNA ligases function in the ATP-dependent ligation of the 5'-phosphate of donor RNA to the 3'-hydroxyl terminus of the acceptor RNA (Figure 2.5B). The reaction proceeds in three nucleotidyl transfer steps. First the RNA ligase interacts with ATP by being self-adenylated while pyrophosphate is released. Subsequently the adenylyl group is transferred to the 5'-phosphate of the donor RNA. Finally the 3'-hydroxyl of the acceptor RNA attacks the activated donor RNA forming the new phosphodiester bond and releasing the adenylyl group (Walker, Uhlenbeck et al. 1975) .

The desired 3'-adapter ligation competes with side reactions, e.g. the reverse reaction from the adenylation of the 3'-adapter or a circularization of input/acceptor RNA (Figure 2.5C). The latter is based on the ligase-catalyzed adenylylate transfer to the acceptor RNA. To reduce the side reaction, ATP is excluded and a truncated form of Rnl 2 (1-249 aa) missing the nucleotidyl transferase domain was used during 3' adapter ligation. Under these conditions, chemically pre-adenylylated 3' adapter oligonucleotides must be employed. There is no risk of side reactions in the 5' adapter ligation since the 3' hydroxyl of the ligated

RNA/3’ adapter is modified and the 5’ adapter does not have a reactive 5’-phosphate. Therefore ordinary ligation in the presence of ATP can be used.

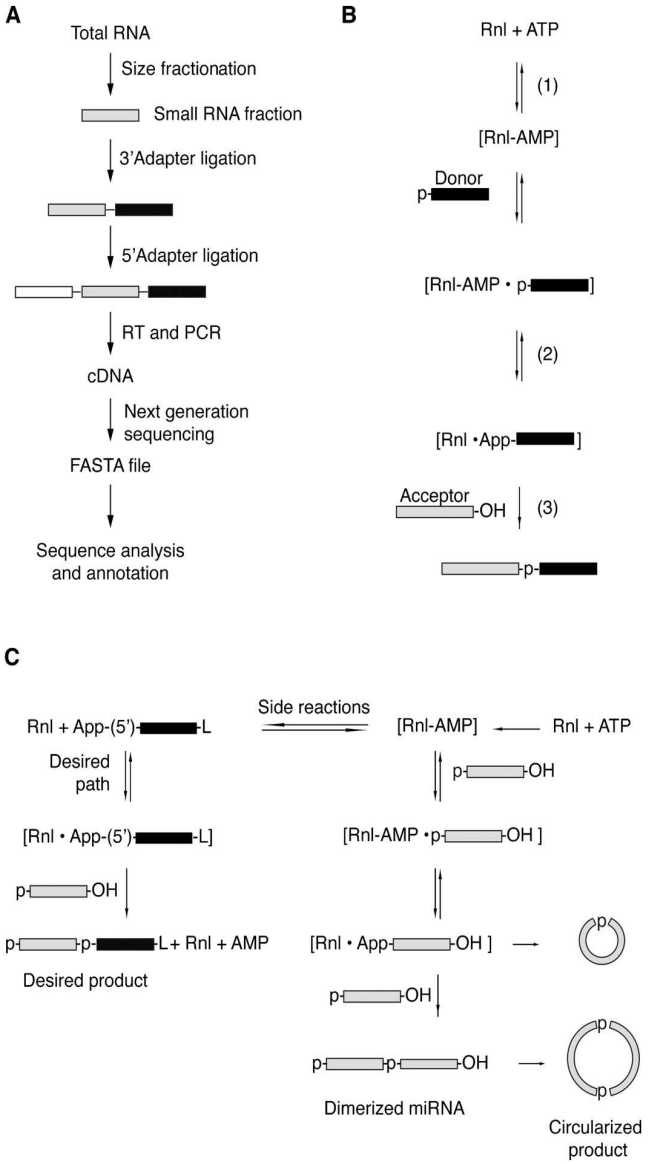


Figure 3.5 Schematic overview of the preparation of small RNA libraries.

(A) Overview of the workflow for the analysis of small RNAs by generation of cDNA libraries with subsequent sequencing. (B) ATP-dependent ligation reaction of the 5’-phosphate of donor RNA to the 3’-hydroxyl of the acceptor RNA involves three nucleotidyl transfer steps: 1) RNA ligase interacts with ATP and forms a covalent Rnl-(lysyl-N)-AMP intermediated while pyrophosphate is released; 2) the AMP is transferred from the ligase to the 5’-phosphate of the donor RNA (black box) to form an RNA-adenylate (AppRNA); 3) the 3’-hydroxyl of the acceptor RNA (light gray box) attacks the adenylylated adapter forming a new phosphodiester bond and releasing the AMP. (C) Overview of the desired ligation reaction (left panel) and the side reaction (right panel). In the latter the back reaction proceeds by ligase-catalyzed adenylylate transfer to the acceptor RNA resulting in undesired circulation or concatenation of small RNAs. Figure adapted from (Hafner, Renwick et al. 2011).

2.5 Aims of this thesis

PART I

1. Which miRNA and endo-siRNA biogenesis factors are required for the cell cycle progression?

2. Do small RNAs, miRNAs and endo-siRNAs, set the timing of cell cycle phases and what changes occur in the small RNA profiles across the cell cycle in Schneider S2 cells of *Drosophila*?

3. Can further types of small RNA species be identified that oscillate in abundance with the cell cycle?

PART II

1. Do the requirements for Loqs-PD and R2D2 differ between soma and germline?

2. What is the requirement of transposons for Loqs-PD and R2D2 during processing and loading based upon?

3. Does impaired endo-siRNA biogenesis change the profile of transposon-targeting piRNAs?

4. Can occurrence of somatic piRNA-like small RNAs (pilRNAs) be confirmed in *Drosophila* somatic tissue? If so, do pilRNAs resemble their germline relatives?

5. Are miRNA* loaded into Ago2 complex in absence of R2D2?

3 MATERIALS AND METHODS

3.1 Materials

3.1.1 Laboratory equipment

Agarose gel running chamber	Carl Roth GmbH; Karlsruhe, Germany
Beckman J-6M/E centrifuge	Beckman Coulter GmbH; Krefeld, Germany
BioLogic LP System	BioRad; Hercules, USA
BioPhotometer	Eppendorf AG; Hamburg, Germany
Branson Sonifier 250	Heinemann Ultraschall Labortechnik
Centrifuge 5417 R	Eppendorf AG; Hamburg, Germany
Centrifuge, Rotanta 460 R	Hettich GmbH; Tuttlingen, Germany
Desk Centrifuge, 220/230 VAC	Stuart
Flow buddy CO_2-distributer	Genesee Scientific; San Diego, USA
Fly anesthetic pad and pistol	Genesee Scientific; San Diego, USA
Fraction Collector Model 2110	BioRad; Hercules, USA
Gel Photometer	Intas INTAS; Göttingen, Germany
HiTrap Chelating HP	GE Healthcare; Freiburg, Germany
Heater	HLC
Incubator	WTC binder
Incubator Shaker Series	New Brunswick Scientific
INTAS UV Imaging System	INTAS; Göttingen, Germany
LAS 3000 mini Western Imager	Fujifilm; Tokyo, Japan
Leica MZ7 stereomicroscope	Leica Microsystems; Wetzlar, Germany
Magnetic Stirrer, MR 3001	Heidolph
Microplate Reader Infinite® F500	Tecan
Overhead Shaker, REAX 2	Heidolph
PAGE-electrophoresis	BioRad; Hercules, USA
Power supply	BioRad; Hercules, USA
PVDF Membrane (0.45 micron pore size)	Thermo Scientific
Roller Mixer, SRT9	Stuart
Shaker, Polymax 1040	Heidolph Instruments
SLC-6000 Centrifuge	Eppendorf AG; Hamburg, Germany
Spectra/Por® Dialysis Membrane, MWCO: 3.500	Spectrum Laboratories, Inc.

SpectroLinker XL1500 UV Crosslinker	Spectronics Corporation; Westbury, USA
SterilGARD cell culture workbench	The Baker Company; Sanford, USA
Table top centrifuge (5417R and 5415R)	Eppendorf AG; Hamburg, Germany
Tank-blotting chamber	BioRad; Hercules, USA
Thermocycler Sensoquest	Sensoquest; Göttingen, Germany
TOptical Thermocycler	Biometra; Jena, Germany
Vortex Genie 2	Scientific Industries; Bohemia, New York, USA
Water Bath	GFL; Burgwedel, Germany
Western Blot Imager LAS 3000 mini	Fujifilm

3.1.2 Laboratory chemicals

2% Triton	Sigma-Aldrich
Acetic acid	Carl Roth GmbH; Karlsruhe, Germany
Acrylamide 40%	Carl Roth GmbH; Karlsruhe, Germany
Agarose Biozym	Biozym Scientific GmbH; Oldendorf, Germany
Ammonium peroxodisulfate (APS)	Roth GmbH; Karlsruhe, Germany
Ampicillin	Carl Roth GmbH; Karlsruhe, Germany
Bacto Agar	Becton, Dickinson; Franklin Lakes, USA
Bradford Assay	BioRad; Hercules, USA
Chloramphenicol	Carl Roth GmbH; Karlsruhe, Germany
Chloroform	Merck Biosciences GmbH; Schwalbach, Germany
Complete® without EDTA (protease-inhibitor)	Roche Diagnostics; Mannheim, Germany
Coomassie G250	Carl Roth GmbH; Karlsruhe, Germany
Desoxyribonucleotides	Sigma Aldrich; Taufkirchen, Germany
Dimethyl sulfoxide (DMSO)	Carl Roth GmbH; Karlsruhe, Germany
Dithiothreitol (DTT)	Carl Roth GmbH; Karlsruhe, Germany
ECL-Solution	Thermo Scientific
Ethanol (p.a.)	Merck Biosciences GmbH; Schwalbach, Germany
FACS Flow/Clean/Rinse	Becton, Dickinson; Franklin Lakes, USA
Fetal bovine serum (FBS)	Thermo Fisher Scientific; Waltham, USA
Formaldehyde	Sigma Aldrich; Taufkirchen, Germany
Formamide	Sigma Aldrich; Taufkirchen, Germany
Fugene®HD transfection reagent	Roche Diagnostics; Mannheim, Germany
G418 sulphate (neomycin)	PAA, The Cell Culture company; Cölbe, Germany

Glycerin	Carl Roth GmbH; Karlsruhe, Germany
H_2O HPLC quality	Carl Roth GmbH; Karlsruhe, Germany
HEPES	Carl Roth GmbH; Karlsruhe, Germany
Isopropanol (p.a.)	Merck Biosciences GmbH; Schwalbach, Germany
Methanol (p.a.)	Merck Biosciences GmbH; Schwalbach, Germany
Phenol/Chloroform/Isoamylalcohol pH4.5-5	Roth; Karlsruhe, Germany
Polyacrylamide	National diagnostics
Powdered milk	Rapilait Migros; Zürich, Switzerland
Ribo LockTM RNase Inhibitor	Fermentas; St. Leon-Rot, Germany
Roti®Aqua Phenol/C/I	Carl Roth GmbH; Karlsruhe, Germany
Sodium dodecyl sulphate (SDS)	Merck Biosciences GmbH; Schwalbach, Germany
Syber Safe/Gold	Invitrogen; Karlsruhe, Germany
TEMED	Carl Roth GmbH; Karlsruhe, Germany
Triton X-100	Sigma Aldrich; Taufkirchen, Germany
Trizol	Invitrogen; Karlsruhe, Germany
Tween 20	Carl Roth GmbH; Karlsruhe, Germany

All other standard laboratory chemicals were purchased from the Gene Center in-house supply.

3.1.3 Enzymes

DNase I, RNAse free	Thermo Scientific; Wattham, USA
Mutant T4 Rnl2 RNA ligase	Laboratory stock (see 3.2.6)
Polynucleotidekinase with buffers	Fermentas, St. Leon-Rot, Germany
Proteinase K	New England Biolabs; Ipswich, USA
Phusion Hot Start DNA Polymerase	Thermo Scientific; Wattham, USA
Superscript II, Reverse Transcriptase	Invitrogen; Karlsruhe, Germany
T4 DNA Ligase	New England Biolabs; Ipswich, USA
T4 RNA ligase	Life Technologies; Carlsbad, USA
T7 polymerase	Laboratory stock
Taq DNA Polymerase	Laboratory stock

3.1.4 Kits

DyNAmo Flash SYBR Green qPCR Kit	Thermo Scientific; Wattham, USA
Clone JET TM PCR Cloning Kit (TA-cloning)	Fermentas; St. Leon-Rot, Germany
QIAGEN Gel extraction Kit	Qiagen; Hilden, Germany
QIAGEN PCR Purification Kit	Qiagen; Hilden, Germany
QIAGEN Plasmid Midi Kit	Qiagen; Hilden, Germany
QIAGEN Plasmid Mini Kit	Qiagen; Hilden, Germany

3.1.5 Marker

PageRulerTM Unstained Protein Ladder	Fermentas; St. Leon-Rot, Germany
PageRulerTM Prestained Protein Ladder	Fermentas; St. Leon-Rot, Germany
Gene RulerTM DNA Ladder Mix	Fermentas; St. Leon-Rot, Germany
50 bp DNA Ladder	New England Biolabs; Ipswich, USA
microRNA Marker	New England Biolabs; Ipswich, USA

3.1.6 Other materials

Mini Quick Spin Oligo Columns	Roche Diagnostics; Indianapolis, USA
Spin column (empty, for Solexa sequencing)	MoBiTec; Göttingen, Germany
SuperSignal West Dura Extended Duration	Thermo Fisher Scientific; Waltham, USA
Blotting paper	Machery-Nagel; Düren, Germany
RestoreTM Western Blot Stripping Buffer	Thermo Fisher Scientific; Waltham, USA
qPCR plates	Eppendorf AG; Hamburg, Germany
Polyvinylidenfluoride (PVDF) membrane	Millipore; Billerica, USA
Pistils for fly lysis	Sigma Aldrich; Taufkirchen, Germany
Protein G Plus/Protein A Agarose Suspension	Calbiochem, Germany
ANTI-FLAG®M2-Agarose from mouse	Sigma Aldrich; Taufkirchen, Germany

3.1.7 Bacterial cells

E. coli BL21(DE3)pLysS	Laboratory stock
E. coli XL2-blue $CaCl_2$-competent cells	Laboratory stock

All *E. coli* strains were cultivated in LB-medium or in SOC-medium following transformation.

Antibiotic containing agar plates were purchased from in-house supply.

SOB-medium	SOC-medium
0.5% (w/v) yeast extract	SOB-medium
2% (w/v) Tryptone	20 mM Glucose
10 mM NaCl	LB-medium
2.5 mM KCl	1% (w/v) Tryptone
10 mM $MgCl_2$	0.5% (w/v) yeast extract
10 mM $MgSO_4$	1% (w/v) NaCl
pH 7.0	pH 7.2

Antibiotics added to medium after autoclaving:

100 µg/ml ampicillin (100 mg/ml stock)

3.1.8 *Drosophila melanogaster* cells

S2 B2	parental cell line	laboratory stock
Ago2 Flag HA_4_2	stable Flag-Ago2 monoclonal expressing cells	kindly given by Katharina Elmer

Cell culture medium and additives for *Drosophila* Schneider cells was purchased from Bio & Sell (Nürnberg, Germany) and supplemented with 10% heat-inactivated Fetal Bovine Serum (FBS; Thermo Fisher; Waltham, USA).

3.1.9 Fly stocks and flyfood

genotype	**description**	**origin**
yw, hs-FLP/yw, hs-FLP; p{w+, $loqs^{KO2-48}$}, FRT40A/CyO; p{w+, Loqs-L (=PB)} $^{298\text{-}ba}$TM3, Sb	Loqs-PB rescue	(Park, Liu et al. 2007)
w/w; *r2d2*^{1}/CyO; 67-2/67-2	*r2d2* deletion	(Liu, Jiang et al. 2006)
w*; Kr/CyO; D/TM6C, Sb Tb	double balancer	Bloomington Stock Center (BL7199)
w^{1118}	recessive *white* mutation	Bloomington Stock center (BL6326)

Standard fly food was obtained from in-house supply.

5.8% corn meal

5.5% molasses

2.4% yeast extract

3.1.10 Plasmids

pKF63	constitutive myc-GFP expression under ubiquitin-promotor control (Forstemann, Tomari et al. 2005)	Amp
pHSneo	neomycin resistance selection of stable cell culture lines	Amp
pMMH2	pKF63 1x perfect match target sites for tsRNA in 3'-UTR	Amp
pMMH3	pKF63 2x perfect match target sites for tsRNA in 3'-UTR	Amp
pMMH4	pKF63 4x bulged match target sites for tsRNA in 3'-UTR	Amp
pMMH5	$tRNA_{Glu}$ with subsequent tRF in MCSof pBluescript II KS+	Amp

3.1.11 Oligonucleotides

3.1.11.1 Fly stock mapping

r2d2 mutant	herp_s	ACCGACACACCTATGAATCC
	r2d2_as	AACAGCGGCAAACCTTCTTA
	cdc14_as	ACGAGAGAGCGCTCTATCAA
loqs^{ko} mutant	loqs_s	CGCTCATCGACAAGCTGAT
	loqs_as	GAGCAGGCGATCGTAAAGAG

3.1.11.2 Oligonucleotides for dsRNA generation

ds dcr-1	T7 dcr-1_s	TAATACGACTCACTATAGGTTGGGCGACGTTTTCGAGTCGATC
	T7 dcr-1_as	TAATACGACTCACTATAGGTTTGGCCGCCGTGCACTTGGCAAT
ds dcr-2	T7 dcr-2_s	TAATACGACTCACTATAGGCTGCCCATTTGCTCGACATCCCTCC
	T7 dcr-2_as	TAATACGACTCACTATAGGTTACAGAGGTCAAATCCAAGCTTG
ds ago-1	ago-1_s	ATTTGATTTCTATCTATGCAGCCA (Forstemann, Horwich et al. 2007)
	ago-1_as	GCCCTGGCCATGGCACCTGGCGTA (Forstemann, Horwich et al. 2007)
ds ago-2	ago-2_s	CGCACCATTGTGCATCCTAACGAG (Forstemann, Horwich et al. 2007)
	ago-2_as	GGGGACAATCGTTCGCTTTGCGTA (Forstemann, Horwich et al. 2007)
ds loqs	T7 loqs 5'UTR_s	CGTAATACGACTCACTATAGGGCAACCACAAATATCAGT
	T7 loqs 5'UTR_as	CGTAATACGACTCACTATAGGTTGCACGGTTTTCGGGAG
ds loqs-PD	T7 loqs-PD_s	CGTAATACGACTCACTATGTGAGTATCATTCAAGACATCGATC
	T7 loqs-PD _as	CGTAATACGACTCACTATAGGTAAGGTGTAAGCATTATGTTAATT
ds r2d2	T7 r2d2_s	CGTAATACGACTCACTATAGGATTCAACTATTCTAGCTTA
	T7 r2d2_s	CGTAATACGACTCACTATAGGCTTTGATTACTAGCATTCCT
ds drosha	drosha_s	AGCAGCAGCAGTGATAGCGATGGC (Forstemann, Tomari et al. 2005)
	drosha_as	TCGGTTATTTTATTTGTTGCTTTAATG (Forstemann, Tomari et al. 2005)
ds gfp	T7 gfp_s	CGTAATACGACTCACTATAGGATGGTGAGCAAGGGCGAGGAGCTG

	T7 gfp_as	CGTAATACGACTCACTATAGGTTACTTGTACAGCTCGTCCA TG
ds DsRed	T7 DsRed_as	CGTAATACGACTCACTATAGGTGGTGTAGTCCTCGTTGTGG
	T7 DsRed_s	CGTAATACGACTCACTATAGGAGGACGGCTGCTTCATCTAC
ds la	T7 la_s	CGTAATACGACTCACTATAGGCCAGGAAGAGGTAGCACAGC
	T7 la_as	CGTAATACGACTCACTATAGGCTTTGTCGTAGTTGGCAGCA
ds jhl1_1	T7 jhI1_1_s	CGTAATACGACTCACTATAGGTTTCATCTGAGGCACAGCAC
	T7 jhI1_1_as	CGTAATACGACTCACTATAGGCTCCACAATCCAACAACACG
ds jhl1_2	T7 jhl1_2_s	CGTAATACGACTCACTATAGGAAACCTAGGGCAGACCCACT
	T7 jhI1_2_as	CGTAATACGACTCACTATAGGCGGGGTAGAGCTTGTGGTAG

3.1.11.3 Molecular Cloning

1x perfect tRF reporter

NotI_s_perfect	ggccAAAAAATGTCTCCACAGTAGT
XbaI_as_perfect	ctagACTACTGTGGAGACATTTTTT

2x perfect tRF reporter

NotI_s_2x_perfect	ggccAAAAAATGTCTCCACAGTAGTctgAAAAAATGTCTCCACAGTAGT
XbaI_as_2x_perfect	ctagACTACTGTGGAGACATTTTTTcagACTACTGTGGAGACATTTTTT

4x bulged tRF reporter

NotI_s_4x_buldge	ggccAAAAAATGTCgaaACAGTAGTctgAAAAAATGTCgaaACAGTAGTctgAAAAAA TGTCgaaACA GTAGTctgAAAAAATGTCgaaACAGTAGT
XbaI_as_4x_buldge	ctagACTACTGTttcGACATTTTTTcagACTACTGTttcGACATTTTTTcagACTACTGTtt cGACATTTTTTcagACTACTGTttcGACATTTTT

$tRNA_{Glu}$-tRF

Glu_tRNA_s	gcatgcggccgccgccacgtggtttaattctc
Glu_tRNA_as	cggatccactcgttgcggctaaaaaga

3.1.11.4 Northern Blotting

2S rRNA	TACAACCCTCAACCATATGTAGTCCAAGCA
bantam	AATCAGCTTTCAAAATGATCTCA
miR-277	TGTCGTACCAGATAGTGCATTTA
tRF	ACTACTGTGGAGACATTTTTT
as-tRF	AAAAAATGTCTCCACAGTAGT

3.1.11.5 Antisense oligonucleotides

as-Luciferase	CAUCACGUACGCGGAAUACUUCGAAAUGUCC
as-tRF	UCUUAAAAAAUGUCUCCACAGUAGUACCU

The constructs are 5’-cholesteryl-modified and all bases are 2’-O-methyl modified.

3.1.11.6 Other oligonucleotides

Oligo dT (EcoRI T18)	ACGAATTCTTTTTTTTTTTTTTTTT
random hexamers	NNNNNN

3.1.11.7 MicroRNA profiling

MicroRNA profiling plate was prepared by Integrated DNA Technologies (Coralville, USA).

Pos.	Name	Sequence 5'- 3'
A1	scrambled_bantam	AGTGCTAGTATTTACAGCTATAT
A2	dme-bantam	TGAGATCATTTTGAAAGCTGATT
A3	dme-let-7	TGAGGTAGTAGGTTGTATAGT
A4	dme-miR-1	TGGAATGTAAAGAAGTATGGAG
A5	dme-miR-1	TGGAATGTAAAGAAGTATGGAG
A6	dme-miR-10	ACCCTGTAGATCCGAATTTGT
A7	dme-miR-10*	AAATTCGGTTCTAGTGTGGTT
A8	dme-miR-1002	TTAAGTAGTGGATACAAAGGGCGA
A9	dme-miR-1003	TCTCACATTTACATATTCACAG
A10	dme-miR-1012	TTAGTCAAAGATTTTCCCCATAG
A11	dme-miR-1017	GAAAGCTCTACCCAAACTCATCC
A12	scrambled_dme-miR-184	AGTAGCGAGATGACATGCGGAC
B1	dme-miR-11	CATCACAGTCTGAGTTCTTGC
B2	dme-miR-12	TGAGTATTACATCAGGTACTGGT
B3	dme-miR-124	TAAGGCACGCGGTGAATGCCAAG
B4	dme-miR-125	TCCCTGAGACCCTAACTTGTGA
B5	dme-miR-133	TTGGTCCCCTTCAACCAGCTGT
B6	dme-miR-13a	TATCACAGCCATTTTGATGAGT
B7	dme-miR-13b	TATCACAGCCATTTTGACGAGT
B8	dme-miR-14	TCAGTCTTTTTCTCTCTCCTA
B9	dme-miR-184	TGGACGGAGAACTGATAAGGGC
B10	dme-miR-184*	CCTTATCATTCTCTCGCCCCG
B11	dme-miR-193	TACTGGCCTACTAAGTCCCAAC
B12	dme-miR-219	TGATTGTCCAAACGCAATTCTTG
C1	dme-miR-252	CTAAGTACTAGTGCCGCAGGAG
C2	dme-miR-263a	GTTAATGGCACTGGAAGAATTCAC
C3	dme-miR-274	TTTTGTGACCGACACTAACGGGT

C4	dme-miR-275	TCAGGTACCTGAAGTAGCGCGCG
C5	dme-miR-276*	CAGCGAGGTATAGAGTTCCTACG
C6	dme-miR-276a	TAGGAACTTCATACCGTGCTCT
C7	dme-miR-276b	TAGGAACTTAATACCGTGCTCT
C8	dme-miR-277	TAAATGCACTATCTGGTACGACA
C9	dme-miR-278	TCGGTGGGACTTTCGTCCGTTT
C10	dme-miR-279	TGACTAGATCCACACTCATTAA
C11	dme-miR-281	TGTCATGGAATTGCTCTCTTTGT
C12	dme-miR-282	AATCTAGCCTCTACTAGGCTTTG
D1	dme-miR-284	TGAAGTCAGCAACTTGATTCCAG
D2	dme-miR-285	TAGCACCATTCGAAATCAGTGC
D3	dme-miR-286	TGACTAGACCGAACACTCGTGCT
D4	dme-miR-289	TAAATATTTAAGTGGAGCCTGCG
D5	dme-miR-2a	TATCACAGCCAGCTTTGATGAGC
D6	dme-miR-2b	TATCACAGCCAGCTTTGAGGAGC
D7	dme-miR-2c	TATCACAGCCAGCTTTGATGGGC
D8	dme-miR-3	TCACTGGGCAAAGTGTGTCTCA
D9	dme-miR-305	ATTGTACTTCATCAGGTGCTCTG
D10	dme-miR-306	TCAGGTACTTAGTGACTCTCAA
D11	dme-miR-306*	GGGGGTCACTCTGTGCCTGTGC
D12	dme-miR-308	AATCACAGGATTATACTGTGAG
E1	dme-miR-309	GCACTGGGTAAAGTTTGTCCTA
E2	dme-miR-310	TATTGCACACTTCCCGGCCTTT
E3	dme-miR-311	TATTGCACATTCACCGGCCTGA
E4	dme-miR-312	TATTGCACTTGAGACGGCCTGA
E5	dme-miR-316	TGTCTTTTTCCGCTTACTGGCG
E6	dme-miR-317	TGAACACAGCTGGTGGTATCCAGT
E7	dme-miR-318	TCACTGGGCTTTGTTTATCTCA
E8	dme-miR-31a	TGGCAAGATGTCGGCATAGCTGA
E9	dme-miR-34	TGGCAGTGTGGTTAGCTGGTTGTG
E10	dme-miR-375	TTTGTTCGTTTGGCTTAAGTTA
E11	dme-miR-4	ATAAAGCTAGACAACCATTGA
E12	dme-miR-5	AAAGGAACGATCGTTGTGATATG

F1	dme-miR-7	TGGAAGACTAGTGATTTTGTTGT
F2	dme-miR-79	TAAAGCTAGATTACCAAAGCAT
F3	dme-miR-8	TAATACTGTCAGGTAAAGATGTC
F4	dme-miR-927	TTTAGAATTCCTACGCTTTACC
F5	dme-miR-92a	CATTGCACTTGTCCCGGCCTAT
F6	dme-miR-92b	AATTGCACTAGTCCCGGCCTGC
F7	dme-miR-932	TCAATTCCGTAGTGCATTGCAG
F8	dme-miR-956	TTTCGAGACCACTCTAATCCATT
F9	dme-miR-958	TGAGATTCTTCTATTCTACTTT
F10	dme-miR-965	TAAGCGTATAGCTTTTCCCCTT
F11	dme-miR-970	TCATAAGACACACGCGGCTAT
F12	dme-miR-977	TGAGATATTCACGTTGTCTAA
G1	dme-miR-980	TAGCTGCCTTGTGAAGGGCTTA
G2	dme-miR-981	TTCGTTGTCGACGAAACCTGCA
G3	dme-miR-984	TGAGGTAAATACGGTTGGAATTT
G4	dme-miR-986	TCTCGAATAGCGTTGTGACTGA
G5	dme-miR-987	TAAAGTAAATAGTCTGGATTGATG
G6	dme-miR-988	CCCCTTGTTGCAAACCTCACGC
G7	dme-miR-989	TGTGATGTGACGTAGTGGAAC
G8	dme-miR-992	AGTACACGTTTCTGGTACTAAG
G9	dme-miR-993	GAAGCTCGTCTCTACAGGTATCT
G10	dme-miR-994	CTAAGGAAATAGTAGCCGTGAT
G11	dme-miR-995	TAGCACCACATGATTCGGCTT
G12	dme-miR-996	TGACTAGATTTCATGCTCGTCT
H1	dme_mdg1	AACAGAAACGCCAGCAACAGC
H2	dme-miR-998	TAGCACCATGAGATTCAGCTC
H3	dme-miR-999	TGTTAACTGTAAGACTGTGTCT
H4	dme-miR-9a	TCTTTGGTTATCTAGCTGTATGA
H5	dme-miR-9b	TCTTTGGTGATTTTAGCTGTATG
H6	dme-miR-9c	TCTTTGGTATTCTAGCTGTAGA
H7	dme-CG4068_B	TTGACTCCAACAAGTTCGCTC
H8	dme-2S-rRNA	ACTACATATGGTTGAGGGTTG
H9	dme-tRNA-CR32359	CGTGGGTTCGAATCCCACTTC

H10	dme_snRNA_U6	CAAAATCGTGAAGCGTTCCAC
H11	dme_RP49	ATCGGTTACGGATCGAACA
H12	as_dme_2S-rRNA	CAACCCTCAACCATATGTAGT

3.1.11.8 mRNA analysis

blood	blood_s	GCAAAGAAAGCCGAATACCA
	blood_as	CCGGTGGAATCCTTTATCCT
copia	copia_s	AGCAAACAACCCCTCATGTC
	copia_as	GCAAACCCAATTTGTCTCGT
juan	juan_s	CAATGGGTTGACAACATTCG
	juan_as	CCCAAACAGGTGACCCATAC
qbert	qbert_s	CACATATACGGTCGCCTGTG
	qbert_as	GGTCAACGGACAAGGGATTA
tinker	tinker_s	CAAGGTCGGCCGAATAATAA
	tinker_as	GACTAGCGAGTCCGATCCAG
1731	1731_s	TCGTATGCGGTGATCTGAAG
	1731_as	CACAACGTGACCCTCTTTCA
Gypsy*	Gypsy_s	CCAGGTCGGGCTGTTATAGG
	Gypsy_as	GAACCGGTGTACTCAAGAGC
297*	297_s	AAAGGGCGCTCATACAAATG
	297_as	TGTGCACATAAAATGGTTCG
roo*	roo_s	CGTCTGCAATGTACTGGCTCT
	roo_as	CGGCACTCCACTAACTTCTCC
I-element*	I-element_s	TGAAATACGGCATACTGCCCCCA
	I-element_as	GCTGATAGGGAGTCGGAGCAGATA
mdg1*	mdg1_s	CACATGTTCTCATTCCCAACC
	mdg1_as	TTCGCTTTTTATATTTGCGCTAC
jockey*	jockey_s	TGCAGTTGTTTCCCCTAACC
	jockey_s	AGTTGGGCAAATGCTAGTGG
INE-1*	INE-1_s	GGCCATGTCCGTCTGTCC
	INE-1_as	AGCTAGTGTGAATGCGAACG
blood*	blood_G_s	TGCCACAGTACCTGATTTCG
	blood_G_as	GATTCGCCTTTTACGTTTGC
S-element*	S-element_s	TGAAAAGCGTCATTCATTCG
	S-element_as	TGTTTCTAGCGCACTCAACG

Doc*	doc_s	GGGTGACTATAACGCCAAGC
	doc_as	GCAAAATCGATCAGGTCTGG
1731*	1731_G_s	AGCAAACGTCTGTTGGAAGG
	1731_G_as	CGACAGCAAAACAACACTGC
F-element*	F-element_s	GCTGGTAGATACCGCTGAGG
	F-element_as	GTAGTCGTCCTCCGTTTTCG
412*	412_s	CACCGGTTTGGTCGAAAG
	412_as	GGACATGCCTGGTATTTTGG
NOF*	NOF_s	AGTTGGACCTGGAATTGTGG
	NOF_as	AATGCACACGGAAGAGGAAC
Idefix*	Idefix_s	AACAAAATCGTGGCAGGAAG
	Idefix_as	TCCATTTTTCGCGTTTACTG
Het-A*(Ghildiyal, Seitz et al. 2008)	Het_A_s	CGCGCGGAACCCATCTTCAGA
	Het_A_as	CGCCGCAGTCGTTTGGTGAGT
piwi	piwi_s	GCATAGGAAGCTGCCATCTC
	piwi_as	TCGTATCTCTCGGGCAGAGT
aub	aub_s	AGACCCAGGAATTTGTGCAG
	aub_as	CGAGGCGCGATAACTTTTAG
ago3	ago3_s	CCGCAGAGTTCTCCAAACAT
	ago3_as	GTAGGCATCGATTCGGTCAT
gapdh	gapdh_s	AATTTTTCGCCCGAGTTTTC
	gapdh_as	TGGACTCCACGATGTATTCG
rp49	rp49 A2	ATCGGTTACGGATCGAACA
	rp49 B2	ACAATCTCCTTGCGCTTCTT

All primer marked with * are adapted from (Ghildiyal, Seitz et al. 2008).

3.1.11.9 Solexa sequencing

Adapter

3’ ligation (Modban)	AMP-pCTGTAGGCACCATCAATdideoxyC
5’ ligation (Solexa linker)	rArCrArCrUrCrUrUrUrCrCrCrUrArCrArCrGrArCrGrC rUrCrUrUrCrCrGrArUrCrU Eurofins MWG – HPLC purified, 50 µM stock

Reverse transcription

3’ RT primer (BanOne)	ATTGATGGTGCCTACAG Eurofins MWG – HPSF purified, 5 µM stock

PCR

5'-Solexa	AATGATACGGCGACCACCGAACACTCTTTCCCT ACACGACG
3'-PCR BamHI	CAAGCAGAAGACGGCATAC**GGATCC**GATTGATGGTGCCTACAG
3'-PCR Pvu	CAAGCAGAAGACGGCATAC**CAGCTG**GATTGATGGTGCCTACAG-3'
3'-PCR Xba	CAAGCAGAAGACGGCATAC**TCTAGA**GATTGATGGTGCCTACAG-3'
3'-PCR Cla	CAAGCAGAAGACGGCATAC**ATCGAT**GATTGATGGTGCCTACAG-3'
3'-PCR BamHI (+2 nt)	CAAGCAGAAGACGGCATAC**gaGGATCC**GATTGATGGTGCCTACAG
3'-PCR Pvu (+2 nt)	CAAGCAGAAGACGGCATAC**gaCAGCTG**GATTGATGGTGCCTACAG
3'-PCR Cla (+2 nt)	CAAGCAGAAGACGGCATAC**gaATCGAT**GATTGATGGTGCCTACAG
3'-PCR Xba (+2 nt)	CAAGCAGAAGACGGCATAC**gaTCTAGA**GATTGATGGTGCCTACAG

All Eurofins MWG – HPSF purified, 10 µM stock

3.1.12 Antibodies

α-Ago1	mouse 1B8	(Okamura, Ishizuka et al. 2004)
α-flag	mouse α-flag M2	Sigma, F1804
α-myc	mouse α-myc 9E10	

3.1.13 Commonly used buffers and stock solutions

ATP-free T4 RNA ligase buffer	100 mM $MgCl_2$ 100 mM DTT 600 µg/ml BSA 500 mM Tris-HCl, pH 7.5
Buffer A (purification of T4 RNA ligase)	50 mM Tris-HCl, pH 7.5 1.2 M NaCl 15 mM imidazole 10 % glycerol
Buffer B (1) (purification of T4 RNA ligase)	50 mM Tris-HCl, pH 7.5 200 mM NaCl 15 mM imidazole 10 % glycerol
Buffer B (2) (purification of T4 RNA ligase)	50 mM Tris-HCl, pH 7.5 200 mM NaCl 200 mM imidazole 10 % glycerol
Citrat buffer	0.2 M Na_2HPO_4 0.1 M citric acid

Church buffer	1% (w/v) bovine serum albumine (BSA) 1 mM EDTA 0.5 M phosphate buffer, pH 7.2 7% (w/v) SDS
DNA loading buffer (6x)	0.25% (w/v) bromophenol blue 0.25% (w/v) xylene cyanol 30% (w/v) glycerol
Dialysis buffer (purification of T4 RNA ligase)	200 mM NaCl 50 mM HEPES pH 7.4 1 mM DTT
Elutriation buffer	1x PBS 0.25% EDTA 1% FBS
Formamide loading dye (2x)	80% (w/v) formamide 10 mM EDTA, pH 8 1 mg/ml xylene cyanol 1 mg/ml bromophenol blue
Laemmli SDS loading buffer (2x)	100 mM Tris/HCl, pH 6.8 4% (w/v) SDS 20% (v/v) glycerol 0.2% (w/v) bromophenol blue 200 mM DTT (freshly added)
Lysis buffer for protein extraction	100 mM KAc 30 mM HEPES, pH 7.4 2 mM $MgCl_2$ 1 mM DTT 1% (v/v) Triton X-100 2x Complete® without EDTA (=protease inhibitor cocktail)
Lysis Buffer (GST-purification)	1x PBS 2% (v/v) Triton 500 mM NaCl 2x Complete® without EDTA (=protease inhibitor cocktail) 1 mM DTT
Solexa elution buffer	0.4% NaCl 0.5% SDS 50 mM Tris-HCl, pH 8
SSC (20x)	3 M NaCl 0.3 M sodium citrate TAE (50x) 2 M Tris-base 0.9 M acetic acid 100 mM EDTA

TAE (50x)	2 M Tris-base 0,9 M acetic acid 100 mM EDTA
TBE (10x)	0.9 M Tris base 0.9 M boric acid 0.5 M EDTA (pH 8)
TBS (10x)	50 mM Tris, pH 7.4 150 mM NaCl

3.2 Methods

3.2.1 Methods of *Drosophila* S2 cell culture

3.2.1.1 Maintenance

Cells were cultured in Schneider´s Medium (Bio&Sell, Nürnberg, Germany) supplemented with 10% heat inactivated fetal bovine serum (FBS; Thermo Fisher Scientific, Waltham, USA) in appropriate cell culture dishes (Sarstedt, Nümbrecht, Germany). Cells were split once to twice a week into fresh medium.

3.2.1.2 Storage of cells in liquid nitrogen

Cell stocks were frozen by adding 500 µl cells to 100 µl Dimethylsulfoxide (DMSO) diluted in 400 µl cell culture medium (+10% FBS) in a Cryovial (Biozym; Oldendorf, Germany). Cryovials were slowly (1°C per hour) cooled to -80°C in an isopropanol freezing container (Nalgene/Thermo Fisher) and transferred into liquid nitrogen for long-term storage.

3.2.1.3 Depletion of individual genes by RNA interference (RNAi)

DsRNA for RNAi was generated using *in vitro* transcription (IVT) with T7-polymerase. To this end templates of the genes of interest were used in which T7-promotor sites were introduced by PCR and afterwards further amplified by PCR using T7-promotor primer (cgtaatacgactcactatagg). The resulting PCR products were precipitated with ethanol and applied for IVT at 37°C over night.

IVT-Mix:

T7-template DNA	10 µl
T7-buffer (10x)	10 µl
DTT (1 M)	0.5 µl
ATP (100 mM)	5 µl
CTP (100 mM)	5 µl
UTP (100 mM)	5 µl
GTP (100 mM)	8 µl
T7 polymerase	2 µl
H_2O (54,5 µl)	ad 100 µl

After *in vitro* transcription 1 µl of DNase I was added per 100 µl of reaction and incubated for 30 min at 37 °C. The precipitate of magnesium pyrophosphate, which formed during the reaction, was pelleted for 5 min at full speed. DsRNA was precipitated from the supernatant with 1x volume of isopropanol and

washed twice with 70% ethanol. The pellet was air-dried and redissolved in 100 µl of RNase-free H_2O. For proper strand annealing $MgCl_2$ was added to a final concentration of 5 mM, the sample was heated to 95°C for 5 minutes and slowly cooled down to room temperature. Concentration of dsRNA was estimated from an agarose gel in comparison to a DNA Ladder Mix (Fermentas; St. Leon-Rot, Germany).

To induce a knock down of a gene of interest cells were seeded at 0,5 x 10^6 cells/ml in 24-well plate and 10 µg of the corresponding dsRNA was added to the medium. After two days soaking with dsRNA was repeated and on day 5 the cells were harvested and stained with propidium iodide (see chapter 3.2.1.7). Finally they were analyzed on a flow cytometer (Becton Dickinson FACSCalibur) using an FL2 linear detector to determine DNA content.

3.2.1.4 Transfection of plasmid DNA

Transfections of S2 cells were carried out essentially as described in (Shah and Forstemann 2008). For each well of a 24-well cell culture dish 100-500 ng of the vector of interest in 50 µl medium (without serum) and 4 µl of Fugene Transfection Reagent (Roche Diagnostics; Mannheim, Germany) in 46 µl of medium (without serum) were mixed and incubated at RT for 1 hour. Cells were added to the transfection mix at 0.5 x 10^6 cells/ml medium (+10% FBS), split on day 3 after transfection and analyzed on day 5 or 6. For transfections in 6-wells all reagents were scaled up according to the culture volume.

3.2.1.5 Selection of clonal cell lines

To create cell lines that stably express a transgene the expression plasmid of interest was co- transfected with an antibiotic resistance plasmid into cells at 5-10 x 10^5 cells/ml. For native S2 cells 20 ng pHSneo (for neomycin resistance) were used together with 200 ng of the vector of interest. After 3 days, cells were split 1:5 into G418 containing medium, respectively. The concentration was 1.2 mg/ml of G418 for neomycin resistance. Cells were split 1:5 once a week for 4 weeks to obtain polyclonal stable cell lines. For clonal selection serial dilution steps in a 96-well plate were made and colonies derived from a single cell were picked.

The resulting GFP fluorescence of the reporter cell lines was determined in a Becton Dickinson FACSCalibur flow cytometer. For this analysis 100 µl of cells were added to 200 µl of FACS flow. For each sample 10 000 cells were measured. Analysis of fluorescence intensity was carried out with CellQuest software (Becton Dickinson; Franklin Lakes, USA). GFP-negative reporter cells were excluded from the analysis and the mean fluorescence value for each sample was determined.

3.2.1.6 Counterflow centrifugal elutriation

Exponential growing *Drosophila* S2 Schneider cells (10x 25 ml-plates: final concentration per 25 ml-plate: 2 x 10^6 cells/ml; total of 5 x 10^8 cells/elutriation approach) were harvested, centrifuged for 7 min at 1400 rpm and counted. The pellet was washed with 1x PBS (7 min, 1400 rpm) and afterwards resuspended in 10 ml elutriation medium (1x PBS, 0.25% EDTA, 1% FBS). The concentrated cells were loaded into a 10 ml syringe.

The preparation of the elutriation system (Figure 3.1) contained following steps: the centrifuge (Beckman J-6M/E) was turned on, the trapped air was removed and the system was equilibrated with elutriation medium. The valve at the pulse/bubble trap was set to bypass the trap while the flow rate was 9 ml/min. The loading of cells was performed by gently pushing the cells from the syringe into the pulse/bubble trap. The valve was opened to allow cells to be drawn into the medium stream. The loading step was carefully monitored through the view port of the centrifuge door. Cells were watched not to pack or flow over the top of the elutriation boundary by adapting the flow rate. After cells have equilibrated the flow rate was increased slowly to 12 ml/min. Gradual increase of pump speed by 0.1 ml/min increments allowed collection of fractions. Cells were placed directly on ice which cause sustained growth arrest. After the last fraction has been collected, any remaining cells were removed by continuing to pump by 70 ml/min. The entire system was sterilized by pumping 70% ethanol through. Any residual ethanol was rinsed by pumping sterile water and the finally dried off by N_2.

Collected cells were centrifuged at 2000 rpm for 5 min, washed with 1x PBS and counted using a counting chamber. Each fraction was analyzed for the cell cycle position by staining 0.5-1 x 10^6 cells with propidium iodide (see 3.2.1.7) while the remaining cells were resuspended with Trizol (Invitrogen; Carlsbad/CA, USA) to allow RNA to be extracted.

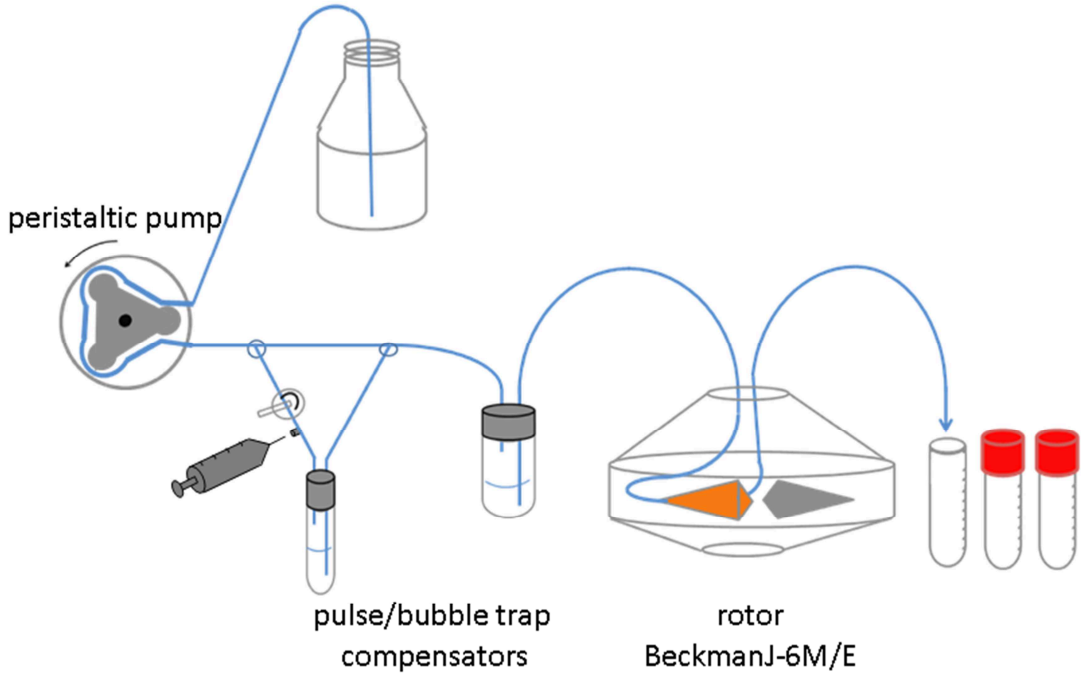


Figure 3.6 Schematic of the counterflow centrifugal elutriation system.

Fluid was drawn from a reservoir through a peristaltic pump. It proceeded through pressure gauge and two pulse/bubble trap compensators to the rotor while the second pulse/bubble trap was used for the loading of concentrated cells. The fluid finally proceeded into a collection vessel. The entire setup (with the exception of the centrifuge itself) is assembled in a fume hood. Figure is kindly given by Katarina Elmer.

3.2.1.7 Cell staining with propidium iodide and flow cytometric analysis

0.5-1 x 10^6 cells of each elutriated population were gently vortexted while gradually adding 1 ml ice-cold 70% ethanol and incubated over night at 4°C. After cells are fixed, they are centrifuged for 10 min at 1500 rpm and washed twice with 1x PBS and 1x Citrate buffer (0.2 M Na_2HPO_4, 0.1 M citric acid). Cell pellet was resuspended in 300 µl Citrate buffer. After addition of RNase (final conc: 20 µg/ml) and PI solution (final conc: 50 µg/ml), cells were incubated for 30 min at 30°C (protected from light). They were carefully vortexed before analysis and analyzed on a flow cytometer (Becton Dickinson FACSCalibur flow cytometer) using an FL2 linear detector.

3.2.2 Nucleic acid analysis

3.2.2.1 Analysis for viral infection of flies

RNA from mutant flies was reverse transcribed (see chapter 3.2.4.5). 1 µl of resulting cDNA was used according to the standard reaction mix (see chapter 3.2.4.9) to analyze for infection with Drosophila A virus. Standard protocol for gradient PCR (50–65°C) was performed to determine the appropriate annealing temperatures. Conditions were then adjusted accordingly:

10 min	95°C initial denaturation
30 cycles:	
30 sec	94°C denaturation
30 sec	59°C annealing
23 sec	72°C extension
5 min	72°C final extension
storage at 4°C	

PCR products were separated by agarose gel electrophoresis, excised, purified by QIAGEN Gel Extraction Kit (Qiagen; Hilden, Germany) and verified by DNA sequencing.

3.2.2.2 Agarose gel electrophoresis

According to the length of nucleotides to be separated 1 - 2% agarose gels were prepared with 1x TAE buffer and stained with 1x SybrSafe (Invitrogen; Karlsruhe, Germany). Gels were run at 55 V for 30 min and photographed in an Intas UV Imaging System. If higher sensitivity was required gels were re-stained in 1x SybrGold (Invitrogen; Karlsruhe, Germany) for 15 min.

3.2.3 RNA analysis

3.2.3.1 RNA extraction

The abdomen from 3 days old female flies was cut off and the remaining somatic portion comprising head and thorax was ground while the ovaries were hand-dissected and also isolated by grinding the tissue in Trizol using a pistil (Sigma Aldrich; Taufkirchen, Germany). RNA was extracted with Trizol (Invitrogen; Carlsbad/CA, USA) according to the manufacturer's instructions and quantified using spectrophotometry.

3.2.3.2 Beta (β)-elimination of RNA

40 µg total RNA dissolved in 40.5 µl H_2O was incubated with 12 µl 5x borate buffer (148 mM borax, 148 mM boric acid pH 8.6) and 7.5 µl $NaIO_4$ (200mM dissolved in H_2O) for 10 min at RT. The oxidation was quenched by addition of 6 µl 100% glycerol (10 min, RT). The elimination was performed by elevating the pH with 2M NaOH (5-7 µl) (ensure that pH=12). After 90 min at 45°C the sample was transferred to a Mini quick spin

oligo column (Roche Diagnostics; Mannheim, Germany), and centrifuged (12 000 g, 2 min). 20-40 µg glycogen were added and RNA was precipitated with 3x volume 100% ethanol (12 000 g, 15 min). RNA pellet was washed three times with 70% ethanol (last step 4°C, o/n) and dissolved in 20 µl 2x denaturating gel loading buffer. The samples were analyzed on a 15% Sequagel Acrylamide-Urea gel and subsequently used for generation of Solexa sequencing libraries.

3.2.3.3 Northern blotting

1-5 µg of RNA were separated on a 20% Sequagel Acrylamid-Urea Gel (National Diagnostics; Atlanta/USA) at 250 V for 90 min. RNA transfer was performed on a positively charged Nylon membrane (Roche Diagnostics; Mannheim, Germany) by semi dry blotting for 30 minutes at 20 V. Crosslinking of the RNA to the membrane was achieved by irradiation with UV-light. Membranes were transferred into hybridization tubes and pre-hybridized in Church buffer (1% (w/v) bovine serum albumine, 1 mM EDTA, 0.5 M phosphate buffer, 7% (w/v) SDS, pH 7.2) for at least 2 hours at 37°C in an oven under constant rotation. The probes were labeled by incubating 9 µl H_2O, 2 µl 10x PNK buffer, 2 µl 5 mM probe oligonucleotide (=10 pmol), 1 µl PNK (Fermentas) and 6 µl [γ-32P] ATP for 1h at 37°C. Unbound radioactive nucleotides were removed using a Sephadex G-25 spin column (Roche Diagnostics; Mannheim, Germany). Hybridization with labeled as-DNA-probes was performed overnight at 37°C in 5 ml Church buffer. Membranes were washed three times for 20 minutes with 2xSSC buffer with 0,1 % SDS and exposed on Phosphoimager Screens (FujiFilm; Tokio, Japan) for up to 1 week. Screens were scanned using a Typhoon scanner (Amersham Biosciences) and band intensities were analyzed using Multi Gauge software (Fujifilm; Tokyo, Japan).

Stripping of the membrane was achieved by dipping it into boiling 1% SDS solution by incubating it for 5 minutes in the solution. After a second pre-hybridization the membrane was reused for hybridization with further probes.

3.2.3.4 Analysis of miRNA and mRNA by quantitative RT-PCR (qPCR)

3.2.3.4.1 miRNA Profiling

The microRNA content of synchronized cells in various cell cycle phases (G1, late S and G2) was analyzed with qRT-PCR on an ABI Prism 7000 sequence detection system (Applied Biosystems Life Technologies, Carlsbad / CA, USA). Reverse transcription was performed using the miScript system (Qiagen, Hilden, Germany) according to the miScript protocol.

Reaction mix for reverse transcription:

miScript RT buffer (5x)	4 µl
100 ng RNA	0.3 – 0.7 µl
miScript enzyme mix	1 µl
H_2O	14.3 – 14.7 µl

Samples were incubated at 37°C for 60 min and then inactivated at 95°C for 5 min. 100 µl of water were added to make a final volume of 120 µl.

The qPCR reaction mixes for 14 reactions (for 1 row of 96-well plate):

SyBr-Green Mastermix (2x)	70 µl
miScript universal primer (5 µM)	14 µl
miScript specific primers (10 µM)	7 µl
H_2O	35 µl

9 µl of reaction mix and 1 µl of RT-reaction per well was amplified in an ABI PRISM 7000 qPCR cycler (Applied Biosystems; Foster City, USA) using the following conditions:

15 min 94°C	initial denaturation
---	---
40 cycles:	
20 sec, 94°C	denaturation
30 sec, 55°C	annealing
30 sec, 70°C	extension

The primer sequences miRNA amplification can be found in chapter 3.1.10.4. Cycle of threshold values (CT-values) were usually determined via the auto-CT function and manually adjusted if necessary. Expression was quantified with the $2^{-(\Delta\Delta Ct)}$ method (Livak and Schmittgen 2001).

3.2.3.4.2 mRNA levels

3.2.3.4.2.1 Digestion of DNA

Endonucleolytic digestion of DNA was carried out with endonuclease DNase I acquired from Thermo Fisher Scientific (Waltham, USA) and the buffer from Fermentas (St. Leon-Rot, Germany).

RNA	5 µg
DNase I	1 µl
DNase buffer (10x)	5 µl
RiboLock	1 µl
H_2O	add 50 µl

After incubation at 37°C for 30 min, 1 µl Proteinase K was added and incubated for 15 min at 65°C in shaking incubator at 600 rpm. The reaction mix was supplemented with equal volume of Phenol/Chloroform/IAA, pH 4.5-5 (Roth; Karlsruhe, Germany), vortexed and centrifuged (full speed, 20 min). The supernatant was precipitated with 80 µl isopropanol and 1 µl glycogen and incubated at RT for 10 min. The reaction mix was centrifuged at 12 000 g for 10 min. The pellet was washed with 100 µl 70% ethanol, dried at 55-60°C for 10 min and resuspended in100 µl RNase free water.

3.2.3.4.2.2 mRNA profiling

100 ng of total RNA after digestion of DNA was reverse transcribed according to the Superscript II Reverse Transcriptase protocol (Invitrogen; Karlsruhe, Germany) primed with random primer (Eurofins MWG Operon).

random primer (100 µM)	1.58 µl
100 ng RNA	x µl
dNTP Mix (10 mM each)	1 µl
H_2O	add 12 µl

The mixture was heated to 65°C for 5 min and quick chilled on ice. The contents of the tube were briefly centrifuged. Then following components were added:

First-Strand Buffer (5x)	4 µl
DTT (0.1 M)	2 µl
RiboLock RNase inhibitor	1 µl
SuperScript II RT	1 µl

The contents of the tube were mixed gently and incubated at 42°C for 50 min. The reaction was inactivated by heating at 70°C for 15 min. 100 µl of water were added to get a final volume of 120 µl. The qPCR reaction mix was as follows, according to the DyNAmo Flash SYBR Green qPCR Kit (Thermo Fisher Scientific; Waltham, USA).

Reaction mix for one well of a 96-well plate:

Dynamo Flash Master Mix	5 µl
oligo_s (10µM)	0.5 µl
oligo_as (10µM)	0.5 µl
H_2O	2.9 µl
xylencyanolblue (0.03%)	0.1 µl

9µl of the reaction solution was aliquoted in each well of a 96 well plate using an 8-canal pipette. 1 µl of the template was added and the samples cycled on a TOptical Thermocycler (Biometra; Jena, Germany) using the following PCR-program:

10 sec, 50°C	
3 min 95°C	initial denaturation

---

40 cycles:	
30 sec, 95°C	denaturation
30 sec, 59°C	annealing
42 sec, 72°C	extension

The primer sequences mRNA amplification can be found in chapter 3.1.11.8. Cycle of Threshold values (CT-values) were usually determined via the auto-CT function. Expression was quantified with the $2^{-(\Delta\Delta Ct)}$ method (Livak and Schmittgen 2001).

3.2.4 Generation of Solexa sequencing libraries

3.2.4.1 Gel purification of RNA

22-60 µg of RNA were separated on a 20% Sequagel Acrylamide-Urea gel (National Diagnostics; Atlanta, USA) at constant 250V for 45 to 60 min. 5 µl microRNA marker (New England Biolabs; Ipswich, USA) consisting of 17, 21 and 25 nt bands, was used as size control. After staining the gel in 1x SybrGold (Invitrogen; Karlsruhe, Germany) for 5 min, the bands of small RNAs were excised corresponding to the desired size from 17 to 30 nt. An 0.5 ml Eppendorf tube pierced with a .22 gauge needle was used to shredder the gel slice (full speed, 5 min). 500 µl of Solexa elution buffer (0.4M NaCl, 0.5% SDS, 50 mM Tris-HCl pH 8) and 1 µl Proteinase K was pipetted into the shred and shaken for at least 2 hours at 65°C to elute the RNA. The gel slices were eliminated by centrifuging (full speed, 2 min) through empty spin column

(MoBiTec; Göttingen, Germany). The eluted RNA was supplemented with 30 µg glycogen and 400 µl Phenol/Chloroform/IAA, pH 4.5-5 (Roth; Karlsruhe, Germany) and centrifuged full speed for 30 min at 4°C. The supernatant was transferred, precipitated with an equal volume of isopropanol, mixed well, incubated for 15 min at RT and centrifuged full speed for 20 min at RT. The supernatant was removed and washed twice with 150 µl of 70% ethanol. Finally RNA was dried for 1 min with the lid closed and resuspended in 8 µl RNase free water.

3.2.4.2 Linker ligation at 3' end of RNA

For linker ligation at 3' end of RNA, the reaction mix was as follows:

Gel purified RNA (resuspended in water)	6 µl
ATP-free T4 RNA ligase buffer (10x)	1 µl
DMSO	1 µl
Modban oligo (50 µM)	1 µl
Mutant RNA ligase (self-made)	1 µl

After incubation for 15 min at 37°C, the ligation reaction was mixed with 10 µl 2x formamide loading dye and inactivated at 95°C for 5 min.

The truncated T4 RNA ligase was taken from our own laboratory stock (see Methods 3.2.6). The corresponding ATP-free T4 RNA ligase buffer (500 mM Tris-HCl pH 7.5-7.6, 100 mM $MgCl_2$, 100 mM DTT, 600 µg/mL BSA) was aquired from New England Biolabs (Ipswich, USA).

3.2.4.3 Gel purification of ligated RNA product after 3' ligation

The ligated RNA products were separated on a 15% Acrylamide-Urea gel (National Diagnostics; Atlanta, USA) at 250 V for 45 to 60 min. After staining of the gel in 1x SybrGold, miRNA marker and 50 bp ladder (New England Biolabs; Ipswich, USA) were used as size control to excise the band corresponding to the desired size of small RNA of 36 to 41 nt. The RNA elution from the gel as well as the RNA precipitation and the final dissolving in water were carried out as explained above in 3.2.4.1.

3.2.4.4 Linker ligation at 5’ end of RNA

For linker ligation at 5’ end of RNA, the reaction mix was as follows:

Ligated product (resuspended in water)	6 µl
T4 RNA ligase buffer (10x)	1 µl
DMSO	1 µl
Solexa linker (50 µM)	1 µl
T4 RNA ligase	1 µl

After the incubation for 1 hour at 37°C, the T4 RNA ligase was inactivated at 95°C for 5 min. T4 RNA ligase and the appropriate buffer were acquired from Life Technologies; Carlsbad, USA.

3.2.4.5 Gel purification of ligated RNA product after 5’ ligation

The following gel purification step of RNA after 5’ ligation was used for the first Solexa sequencing run (cell cycle: G1, early S, late S and G2) while it was skipped for the later approaches.

The ligated RNA products were separated on a 10% Acrylamide-Urea gel (National Diagnostics; Atlanta, USA) at 250 V for 45 to 60 min. After staining of gel in 1x SybrGold, 50 bp ladder (New England Biolabs; Ipswich, USA) were used as size control to excise the band corresponding to the desired size of small RNA around 100 nt length. The RNA elution from the gel as well as the RNA precipitation and the final dissolving in water were carried out as explained above in 3.2.4.1.

3.2.4.6 Reverse transcription

The reverse transcription of ligated RNA is adapted to the Superscript II Reverse Transcriptase protocol (Invitrogen; Karlsruhe, Germany):

Ligated RNA product	9 µl
BanOne primer (5 µM)	2 µl

After incubation at 95°C for 2 min, the mix is cooled on ice for 2 min and centrifuged briefly at RT. The following components are added:

First strand buffer (5x)	4 µl
DTT (0.1 M)	2 µl
dNTP Mix (10 mM each)	1 µl
RiboLock RNase inhibitor	1 µl

The resulting reaction was mixed gently, split into two tubes comprising 9 µl each and incubated at 42°C for 3 min. After addition of 1 µl Superscript II RT (Invitrogen) to the sample or 1 µl H_2O as negative control, the content of the tubes were incubated at 42°C for 30 min. The reaction was inactivated by heating at 95°C for 5 min and cDNA stored at -20°C.

3.2.4.7 PCR amplification of cDNA

cDNA (+RT) or control (–RT)	5 µl
PCR buffer (5x; Mg^{2+} final conc.: 2.5 mM)	20 µl
dNTP Mix (10 mM each)	2 µl
5'-Primer - Solexa (10 µM)	1 µl
3'-Primer - PCR BamHI/Pvu/Xba/Cla (10 µM)	1 µl
Phusion Hot Start DNA Polymerase	1 µl
H_2O	70 µl

Thermocycler protocol:

OLD PROTOCOL			**NEW PROTOCOL**	
2 min	94°C	initial denaturation	2 min	94°C
5 cycles:				
15 sec	94°C	denaturation	----------	
30 sec	54°C	annealing	----------	
30 sec	72°C	extension	----------	
17 cycles:			**23 cycles:**	
15 sec	94°C	denaturation	15 sec	94°C
30 sec	60°C	annealing	30 sec	60°C
30 sec	72°C	extension	30 sec	72°C
2 min	72°C	final extension	2 min	72°C
storage at 4°C			storage at 4°C	

PCR products were separated by agarose gel electorophoresis (2% agarose gel), excised, purified by QIAGEN Gel Extraction Kit and finally eluted with 30 µl Elution buffer.

Furthermore different primers were used for old or new protocol as the bar codes were introduced at different steps during 5' ligation or PCR reaction, respectively (see Figure 3.2). The primer sequences can be found in chapter 3.1.11.9.

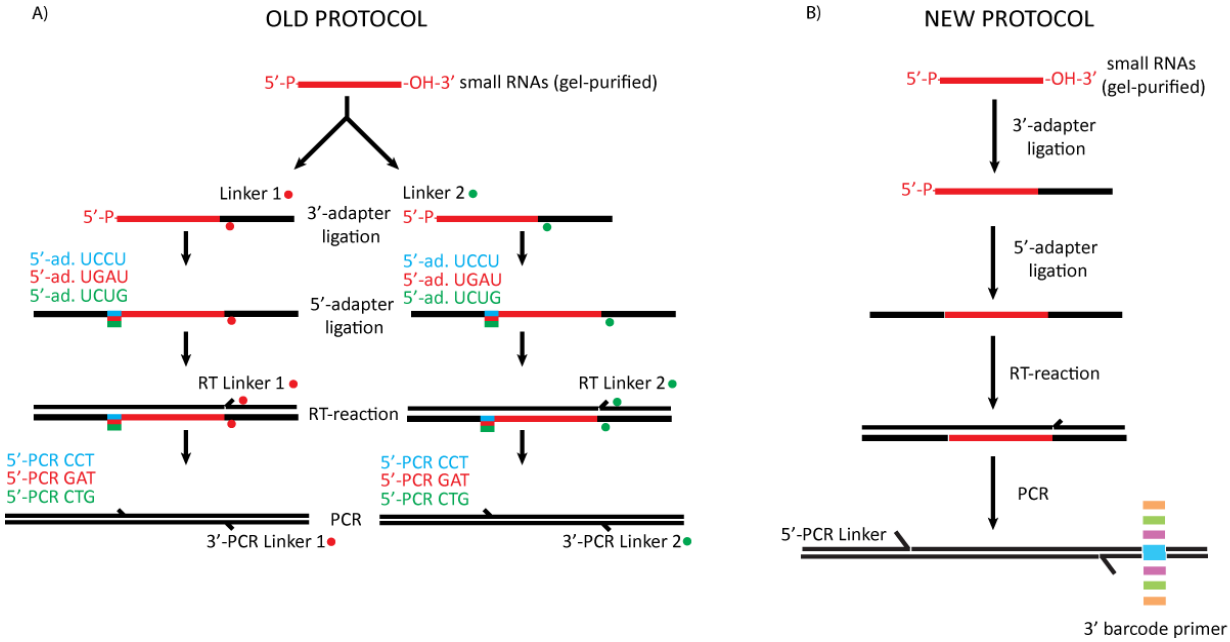


Figure 3.7 Overview of primer use for multiplexing experiment in generation of small RNA libraries.

3.2.4.8 Ligation of purified cDNA with pJET 1.2/blunt

Ligation of purified cDNA with pJET1.2/blunt was performed according to the CloneJET™ PCR Cloning Kit protocol.

PCR product	4 µl
pJET1.2/blunt cloning vector (50 ng/µl)	1 µl
Reaction buffer (2x)	10 µl
T4 DNA ligase	1 µl
H_2O	4 µl

The incubation time was extended up to 30 min at RT to obtain the maximal number of transformants.

3.2.4.9 Bacterial transformation

Transformation of competent bacteria was carried out by standard heat shock procedures. Briefly, 50 µl XL2-blue $CaCl_2$-competent cells were thawed on ice. 5-8 µl of ligation sample were added and the mixture was incubated on ice for 30 min, subjected to a 2 min heat shock at 42°C and returned to ice for 1 min. 1 ml SOC-medium was added and cells were allowed to grow for 1 h in at 37°C shaking incubator. Afterwards

cells were centrifuged for 30 sec at full speed in a table top centrifuge. The supernatant was removed and the resuspended cell pellet was streaked out on agarose plates with Ampicillin (Amp) antibiotic for selection of transformants.

3.2.4.10 Test for correct transformants by colony-PCR

Individual colonies were tested for their insert by colony-PCR with a primer pair contained in the CloneJET™ PCR Cloning Kit (pJET1.2 fw: 5'-cgactcactatagggagagcggc-3'; pJET1.2 rev: 5'-aagaacatcgattttccatggcag-3'). A following standard PCR reaction mix was inoculated with a single colony, which was subsequently streaked onto a fresh plate and labeled for later recognition. Standard amplification was carried out with 10 min initial denaturing for cell lysis of bacteria.

Taq buffer (+KCl, -$MgCl_2$) (10x)	2 µl
$MgCl_2$ (50 mM)	0.6 µl
dNTP Mix (10 mM each)	0.4 µl
pJET 1.2_s (10 µM)	0.4 µl
pJET 1.2_as (10 µM)	0.4 µl
Taq polymerase	0.1 µl
H_2O	16.1 µl

Thermocycler protocol:

10 min	95°C initial denaturation
---	---
35 cycles:	
30 sec	94°C denaturation
30 sec	55°C annealing
30 sec	72°C extension
---	---
storage at 4°C	

PCR products were separated by agarose gel electorophoresis, excised and purified by QIAGEN PCR Purification Kit.

3.2.4.11 DNA sequencing

The sequences of the obtained inserts were investigated by sequencing (Eurofins MWG, Ebersberg, Germany). Further analysis of the sequences and alignments were performed with ApE (A plasmid Editor; http://biologylabs.utah.edu/jorgensen/wayned/ape/) and the BLAST function of http://flybase.org.

3.2.4.12 Bioinformatic analysis of deep sequencing data

Solexa sequencing for total RNA libraries was carried out at Fasteris (Plan-Les-Ouates, Switzerland) while the sequencing of libraries consisting of beta-eliminated RNAs was performed at the Gene Center (Munich, Germany).

The sequences were mapped onto the target sequences using BOWTIE (http://bowtie-bio.sourceforge.net/) with the option –n0 to force selection of only perfectly matching sequences. Pre-processing of sequences and analysis of the BOWTIE output files were done using PERL scripts.

3.2.5 Protein analysis

3.2.5.1 Protein extraction

Fly protein was extracted by grinding flies in lysis buffer using a pistil (Sigma Aldrich; Taufkirchen, Germany) suitable for 1.5 ml reaction tubes and washed with PBS. The pellet was resuspended in lysis buffer (30 mM HEPES pH 7.4, 100 mM KAc, 2 mM $MgCl_2$, 1 mM fresh DTT, 2x protease inhibitor cocktail (Roche Diagnostics GmbH; Mannheim, Germany) and frozen in liquid nitrogen. Samples were thawed on ice and cell debris was pelleted in a refrigerated microcentrifuge (Eppendorf; Hamburg, Germany) at 13 000 rpm for 5 min. Protein concentrations were determined by Bradford assay (BioRad; Hercules, USA).

3.2.5.2 Co-Immunoprecipitation

For immunoprecipitation 50 µl Protein G Plus/Protein A Agarose beads were incubated with Ago1 (1B8) or myc (9E10) antibodies for 4 hours at 4°C and unbound antibody was removed by washing with 1x PBS twice. Anti-flag beads were pre-washed with 1x PBS twice. Prepared beads were incubated for 30 min with 0.5-2.5 µg protein extract in lysis buffer at 4°C on an overhead rotator. Flow-through and beads were separated by spin columns (MoBiTec; Göttingen, Germany) and washed three times with 500 µl lysis buffer. RNA was extracted by applying Trizol (Invitrogen; Karlsruhe, Germany) and following the manufacturer's instructions. 1.5 µl of glycogen was added before precipitation with isopropanol. RNA was dissolved in 20 µl H_2O.

3.2.6 Recombinant expression and purification of mutated T4 RNA ligase 2

3.2.6.1 Recombinant expression

The 100 ml pre-culture (LB medium, 100 µl ampicillin, 100 µl chloramphenicol, 0.5 % glucose) inoculated with *E. coli BL21(DE*)pLysS strain transformed with mutant pET-RNL2 plasmid and incubated at 250 rpm and 37°C over night. The 1 l expression culture, that contained appropriate antibiotics, was inoculated with 51 ml of pre-cultured transformed bacteria (dilution 1:20). After cell growth to OD600 = 0.7 in 2 hours at 37°C in baffled flasks, protein expression was induced by adding 1 mM IPTG. After incubation for 9 hours at 21°C, the bacteria were harvested by centrifugation for 30 min at 4500 rpm at 4°C. The obtained bacterial pellets were resuspended in 80 ml Buffer A (50 mM Tris-HCl pH 7.5, 0.2% Triton X-100, 1.2 M NaCl,15 mM imidazole, 10% glycerol, protease inhibitor cocktail without EDTA (Roche Diagnostics GmbH; Mannheim, Germany)) and stored at -80°C. After thawing 100 ml Buffer A were added to the bacterial pellet and incubated on ice for 1 hour. The pellet was resuspended with a sonicator twice for 1 min on ice (output 5-6, amplitude 20-30). Afterwards, the solution was centrifuged for 40 min at 17 000 rpm at 4°C and the supernatant was taken for the following step (see 3.2.6.2).

3.2.6.2 Affinity purification of mutant T4 RNA ligase 2

Purification of histidine-tagged mutant T4 RNA ligase 2 was achieved by binding to nickel (Ni^{2+}) immobilized to a sepharose matix (HiTrap Chelating HP) previously equilibrated with Buffer A. Then 180 ml protein solution was loaded at 4°C with a rate of 2ml/min. The column was washed four times to 5 ml Buffer A. Afterwards, the buffer was changed with Buffer B (50 mM Tris-HCl pH 7.5, 10% glycerol, 0.2 M NaCl, 20 mM imidazole) to start elution which is performed by increasing the imidazole concentration from 15 mM to 200 mM within 30 ml volume. The elution fractions containing the desired protein identified with SDS-PAGE were pooled and dialyzed against 2 l dialysis buffer (50 mM Tris-HCl pH 8.0, 0.2 M NaCl, 1 mM DTT, 50% glycerol) in a dialysis membrane with molecular weight cut off of 3.5 kDa, while stirring at 4°C over night. The final mutant T4 RNA ligase 2 was obtained in 2 ml with final concentration 0.5 mg/ml.

3.2.7 Methods with flies

3.2.7.1 Maintenance and handling

The fly stocks were kept on standard agar food at 25°C and transferred to new food once a week. For phenotype selection flies were anesthetized with CO_2 and sorted on a CO_2-emitting pad (Genesee Scientific; San Diego, USA) using a Leica MZ7 stereomicroscope (Leica Microsystems; Wetzlar, Germany). To slow

proliferation by reducing metabolic rates flies were kept at 18°C if they were not used for a current experiment and were transferred to new food every 4 weeks.

3.2.7.2 Crossing

To reduce the difference in transposon contents between *r2d2* mutant (Liu, Jiang et al. 2006) and *loqs*^{ko} mutant (Park, Liu et al. 2007), each transgenic fly line was crossed with wild type w^{1118} flies. The F1 offspring was each crossed to Kr/Cyo; D/TM6, Sb, Tb double balancer males, to obtain offspring with balanced 2nd (Cyo) and 3rd (TM6, Tb, Sb) autosomes. F2 offspring was selected for Cyo, TM6, Sb, Tb phenotypes. Siblings were then mated to produce homozygous stable lines.

Additionally the heterozygous *r2d2* and *loqs*^{ko} mutant balanced over Cyo were further crossed with wild type flies (*w*^{1118}) to remove the curly wing phenotype.

3.2.7.3 Characterization of *r2d2* and *loqs*^{ko} flies by genomic PCR

Genomic DNA was isolated for both mutants according to the Berkeley Drosophila Genome Project protocol. DNA was prepared from 30 anesthetized flies by freezing at -80°C and subsequent mechanical lysis in 400 µl Buffer A (100 mM EDTA, 100 mM NaCl, 0.5% SDS, 100 mM Tris-HCl, pH 7.5) by grinding flies using a pistil (Sigma Aldrich; Taufkirchen, Germany). After incubation at 65°C for 30 min, 800 µl LiCl/KAc-solution was added, cooled on ice for at least 10 min. Debris was pelleted and 1 ml supernatant was precipitated with 600 µl isopropanol, vortexted and centrifuged for 15 min at RT. The pellet was washed with 70% ethanol, dried and finally resuspended in 150 µl TE.

The characterization was performed by PCR while the standard reaction mix was as follows:

genomic DNA	1 µl
Taq buffer (+KCl, -$MgCl_2$) (10x)	2.5 µl
$MgCl_2$ (50 mM)	1 µl
dNTP Mix (10 mM each)	1 µl
oligo_s (10 µM)	1 µl
oligo_as (10 µM)	1 µl
Taq polymerase	1 µl
H_2O	16.5 µl

Thermocycler protocol:

10 min	95°C initial denaturation
---	---
30 cycles:	
30 sec	94°C denaturation
30 sec	55.9°C annealing
1 min per kb product size	72°C extension
---	---
storage at 4°C	

Amplification products were separated on a 1,5% agarose gel.

4 RESULTS

4.1 Part I

4.1.1 Optimization of the Solexa-based small RNA cloning protocol

4.1.1.1 Ligation at the 3' end of small RNAs

The Solexa-based small RNA cloning protocol (Figure 3.7) started with gel purification of RNA and proceeded with ligation of pre-adenylated DNA linker at the 3' end of small RNAs referred as 3' ligation. The reaction was catalyzed by truncated T4 RNA ligase 2 (ΔT4 Rnl2) which is compromised for adenylyltransferase activity. ΔT4 Rnl2 ligase was acquired from NEB (Ipswich, USA). In parallel ΔT4 Rnl2 was expressed from a plasmid in *E. coli*. The purification of histidine-tagged mutant T4 RNA ligase 2 was achieved by binding to Ni^{2+}-NTA-column and elution via increasing concentration of imidazole (see chapter 3.2.6). In the following, the commercial and the home-made enzymes were compared and different ligase concentrations were tested to optimize the ligation reaction. To do so, a synthesized miR-277 passenger oligo with 23 nt length was used as RNA substrate. We monitored the conversion of input miR-277 RNA oligo to higher-molecular-weight species in ligation reactions containing 2-fold molar excess of pre-adenylated linker. The concentration of commercial ligase was used according to the appropriate NEB protocol. The home-made ligase was analyzed at three different concentrations (Figure 4.1). The purchased ΔT4 Rnl2 ligase generated the ligated product at the expected size of 41 nt (lane 1). This ligation efficiency was comparable to the lowest concentration of home-made mutant ligase (lane 4). Three-fold enhancement in concentration increased the efficiency of small RNA ligation observed in depletion of RNA substrate (lane 3). No further changes could be detected after additional increase of the concentration (lane 2). Summarizing, home-made ΔT4 Rnl2 was shown to be more effective than commercial ligase under the analyzed conditions. Hence it was selected for further usage in the 3' ligation step of RNA with pre-adenylated linker during the RNA cloning protocol.

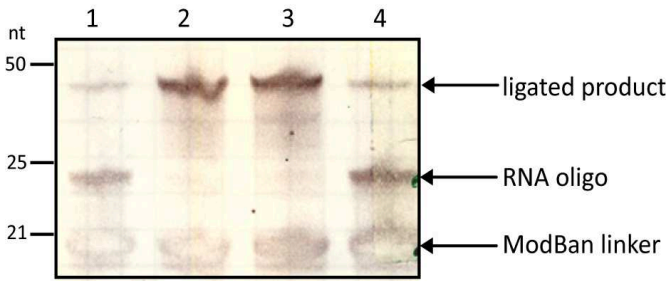


Figure 4.1 Optimization of the ligation at the 3' end of small RNAs.

The synthesized miR-277 passenger oligo with 23 nt length was used as RNA substrate. The pre-adenylated ModBan linker was added in a 2-fold molar excess while both ligases used ATP-free T4 RNA truncated ligase commercial buffer. The reaction products were quantified with a 15% Acrylamide-Urea gel. Lane 1: ΔT4 Rnl2 (NEB); lane 2: 6x ΔT4 Rnl2 (home-made); lane 3: 3x ΔT4 Rnl2 (home-made); lane 4: 1x ΔT4 Rnl2 (home-made).

4.1.1.2 Amplification step by polymerase chain reaction (PCR)

The advancement of deep sequencing technology progressed into an increase in sequencing depth. Thus multiplexing experiments could be performed by pooling various libraries into one sequencing run. To allow separation of samples in the later analysis, sample specific bar codes were inserted during the ligation at the 5' end of RNA referred as 5' ligation. After ligation and reverse transcription, cDNA was finally amplified by PCR program consisting of pre-amplification phase and specific amplification. First optimization was possible by usage of longer PCR primers which contained sequences against bar codes inserted in 5' ligation. Thereby small RNA sequences were recognized with higher specificity which made the pre-amplification cycles dispensable. The optimization of the annealing temperature for the specific amplification process which was performed by temperature gradient PCR from 56°C to 64°C yielded comparable product amounts and proposed an optimal annealing temperature of 60°C for further experiments.

Deep sequencing is the method of choice to quantitatively compare small RNAs differing in their expression levels. Thus already existing biases are exacerbated when libraries are over-amplified implicating that low expressed small RNAs are highly diminished and less detectable. Hence it is of critical importance to reveal the threshold to generate enough material without risk of over-amplification. To do so, PCR reactions were amplified between 12 to 28 cycles while 25 cycles were estimated as final set up (data not shown). Solexa libraries were generated according to the abovementioned optimized protocol and analyzed for miRNAs by aligning the reads to the reference of *Drosophila* miRNAs (miRBase). Unfortunately the libraries mainly consisted of bantam, miR-184 and miR-8 yielding together 71% to 93% of the miRNA matching reads (Table 4.1). To sum up, the observed result indicated that the cycle number in the amplification step was over-estimated.

A)

	bantam (% of miRNAs)	miR-184 (% of miRNAs)	miR-8 (% of miRNAs)	sum (bantam+ miR-184+miR-8	miRNAs total
G1	1711 **(3.2)**	41578 **(76.7)**	6018 **(11.1)**	**91.0**	54209
early S	231454 **(35.3)**	171984 **(26.2)**	106449 **(16.2)**	**77.7**	655994
late S	2005210 **(77.5)**	250169 **(9.7)**	142184 **(5.5)**	**92.6**	2588093
G2	210241 **(35.4)**	108168 **(18.2)**	107479 **(18.1)**	**71.7**	594114

B)

	bantam (% of miRNAs)	miR-184 (% of miRNAs)	miR-8 (% of miRNAs)	sum (bantam+ miR-184+miR-8	miRNAs total
G1	6252598 **(83.7)**	391741 **(5.2)**	385473 **(5.2)**	**94.1**	7468043
S	982389 **(37.9)**	326359 **(12.6)**	735220 **(28.4)**	**78.9**	2590416
G2	440469 **(49.0)**	104344 **(11.6)**	174112 **(19.4)**	**80.0**	898200

Table 4.1 Number of counts for bantam, miR-184 and miR-8.

(A) First sequencing round contained Solexa libraries made of small RNAs isolated from G1, early S, late S and G2 cell cycle phases. (B) Second sequencing round consisted of G1, S and G2 phases and was sequenced twice after amplification via 21 and 23 cycles, respectively. For each phase both libraries were clustered and used for further analysis. In general, deep sequencing reads were processed by selection into different cell cycle phases due to the cell cycle specific bar codes and selected for reads of 11 nt to 28 nt length. They were further mapped against each known *Drosophila* miRNA. The number of counts for bantam, miR-184 and miR-8 were normalized by the number of miRNA matching reads and displayed as the percentage of all miRNAs.

One possibility to obtain more heterogeneous libraries is to reduce the cycle number during the PCR. To elucidate to which extent the reduction has to be carried out, the yield of PCR product was analyzed in smaller steps at 15, 20, 25 and 30 cycles. The optimal time point was provided around 23 cycles (Figure 4.2). Thus the following libraries were generated by amplifying with 21 and 23 cycles, respectively. Additionally every gel purification step is accompanied with loss of RNA. To increase the concentration of small RNAs and sustain their original distribution in the sample, the gel purification step after 5' ligation was skipped. Taken together, to gain a diverse insight into analyzed RNA samples, libraries were generated without gel purification step after 5' ligation and amplified with 21 and 23 cycles, respectively, yielding two data sets for every RNA sample. The bioinformatic analysis demonstrated that both libraries for each RNA sample, generated with 21 or 23 cycles, contained the same percentages of miRNAs with no obvious differences (Table 4.2). The over-representation of bantam, miR-184 and miR-8 was hardly changed after the optimization procedure (Table 4.1B). Taken together, these results argue against the PCR as the sole source for over-production of enormous amounts of specific small RNA sequences which are favored over others during the cloning procedure. Hence the latter results suggested that the observed bias is introduced during the ligation reactions due to the high ligation efficiencies for specific small RNA sequences.

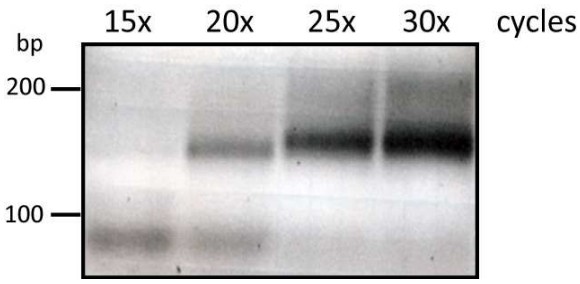


Figure 4.2 Optimization of PCR amplification cycles.

Small RNAs were ligated at their 3' end and 5' end, reverse transcribed and amplified. PCR was performed with 15, 20, 25 and 30 cycles yielding PCR product with expected size of around 120 bp.

As named above bantam was the most abundant miRNA species in all sequenced libraries with exception of G1 phase in the very first sequencing run (Table 4.1). Why is bantam over-represented to such an extent even in comparison to miR-184 and miR-8? Is it emerging from the high ligation efficiency of the linker used in the cloning procedure? In order to test this hypothesis, libraries from different sequencing runs made with the same combinations of linker were compared with each other. The overview of linkers used in 3' ligation and 5' ligation is depicted in Figure 4.3A. Each library was presented with its appropriate length distribution containing total amount of reads, exclusively miRNAs and solely bantam (Figure 4.3B). All combinations with linker 1 favored bantam heavily compared with all other miRNAs and even in respect to total reads. All libraries contained truncated bantam reads either at a length of 15 nt, 18 nt and/or 19 nt. These were more abundant than the most commonly known size of bantam at 23 nt (Figure 4.3B). A detailed look revealed that the shorter versions represented a defined part of bantam, rather than shifting along the bantam sequence (data not shown). Hence we wondered if these abbreviated bantam sequences have a biological relevance? The preference for either 15 nt or 18 nt or 19 nt peak was observed depending on which bar code linker was used in the 5' ligation (Figure 4.3B; GAT: 15nt and 19 nt, CTG: 18 nt, CCT: 18 nt). Taken together, it is obvious that the over-representation of bantam demonstrates a technical artifact based on the bar coding strategy.

A)

	3' ligation	5' ligation - barcode
G1	linker 2	CTG
early S	linker 1	CCT
late S	linker 1	GAT
G2	linker 1	CTG

G1	linker 1	GAT
S	linker 1	CTG
G2	linker 1	CCT

B)

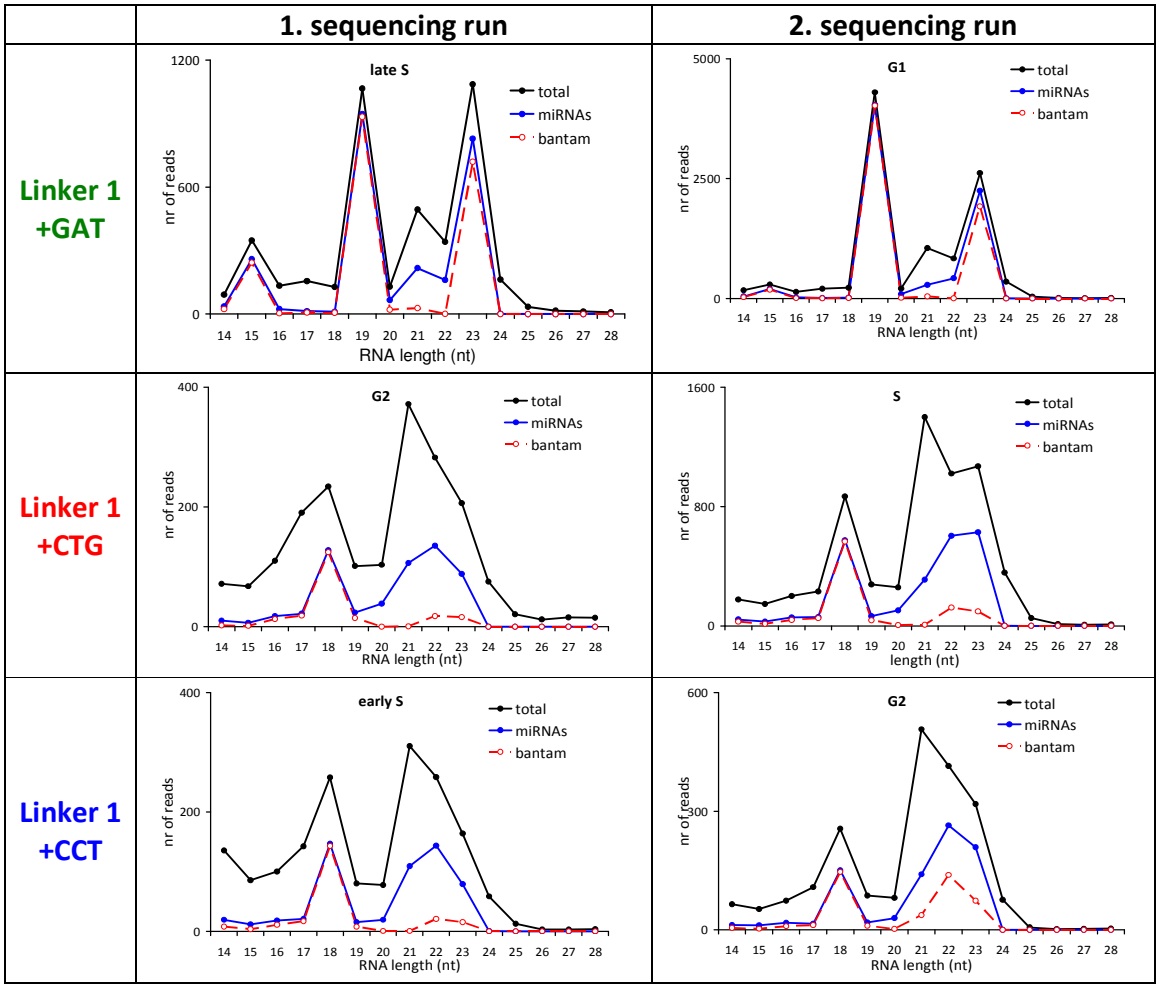


Figure 4.3 The analysis of libraries regarding to the linker used in the 3' and 5' ligation step.

(A) Overview of linker molecules used in the 3' and 5' ligation during the cloning procedure for each RNA sample. (B) Comparison of libraries generated with same linker combinations. Data are depicted in length distribution of absolute counts stretching from 14 nt to 28 nt and presented total number of reads, number of counts matching against all of miRNAs and exclusively bantam mapping reads.

4.1.2 Small RNA analysis with regard to the cell cycle

4.1.2.1 miRNA but not siRNA biogenesis factors are required for cell cycle progression

Which small RNA silencing pathway is influencing the cell cycle? To tackle this question we depleted cells for individual miRNA- and siRNA-pathway genes by means of RNA interference. After the double treatment with dsRNA, DNA was stained with DNA-binding propidium iodide (PI), an intercalating agent as well as a fluorescent molecule, which quantitatively stains DNA. The measurement was performed by means of the flow cytometric analysis. As the DNA content of cells duplicates during the S phase of the cell cycle, the fluorescence of cells in G2 phase will be twice as high as that of cells in G1 phase. If analyzed miRNA or

siRNA biogenesis factors participate in the cell cycle timing, cell cycle profiles will be perturbed. Differences in the cell cycle distribution were compared to the non-specific RNAi control cells, treated with dsRNA directed against DsRed and gfp, respectively. *dcr-1*, *loqs*, *drosha* and *ago1* genes were depleted to gain more insight about the role of miRNA pathway while *dcr-2*, *loqs-PD*, *r2d2* and *ago2* genes were tested for importance of the siRNA pathway. Cell cycle profiles did not change after impairment of siRNA biogenesis factors compared to both control knock downs (Figure 4.4). This indicated that siRNAs did not play any role in setting the timing of cell cycle phases which countered the *ago2 Drosophila* mutant with asynchronous nuclear replication and division cycles in early embryogenesis (Deshpande, Calhoun et al. 2005). Depletion of *ago1* and *drosha* caused a strong accumulation of cells in G1 phase (Figure 4.4). Interestingly depletion of *dcr-1* affected G1 arrest after the third treatment with dsRNA while the results in Figure 4.4 were depicted after double knock down. This finding could be possibly explained by inefficient RNAi. To sum up, depletion of *ago1*, *drosha* and eventually *dcr-1* genes clearly showed a loss of fidelity in cell cycle timing. Thus small RNAs, more precisely miRNA but not siRNAs, were required for the normal cell cycle progression.

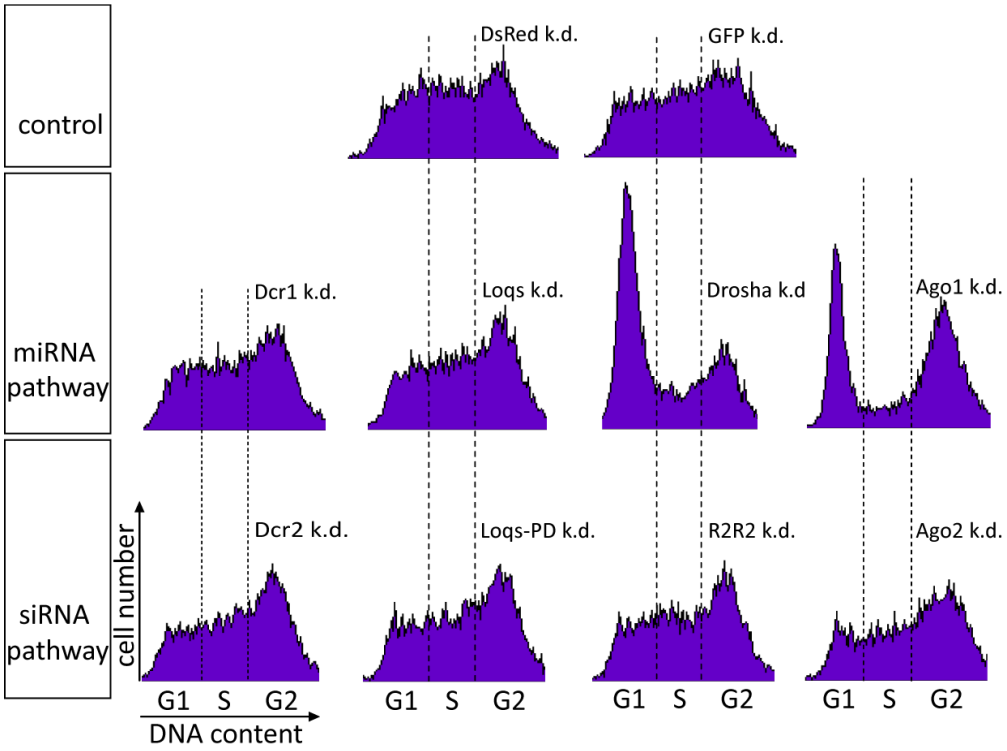


Figure 4.4 miRNA but not siRNA biogenesis factors are required for the cell cycle progression.

S2 cells were treated twice with dsRNA constructs against components of the small RNA silencing pathway. Ago2-RISC was primed with RNAi triggers against DsRed and gfp as a control, respectively. Hereafter cells were harvested, permeabilized with 70% ethanol, incubated with RNase and stained with propidium iodide (PI). The fluorescence was measured by flow cytometry using an FL2 linear detector. The diagrams present the DNA content against the cell number.

4.1.2.2 Overview and quality test of Solexa libraries generated from different cell cycle stages

The cell cycle of proliferating cells is comprised of chromosome condensation with the subsequent cell division named mitosis (M), quiescent stage referred to as G1, which is followed by the DNA synthesis named S phase (S) and the second period of apparent quiescence G2. To investigate the role of small RNAs in the cell cycle, *Drosophila* S2 Schneider cells were synchronized into various stages of the cell cycle by counterflow centrifugal elutriation functioning on the basis of the mass and size of cells. The collected cell fractions were verified for their cell cycle position by determination of DNA content using propidium iodide as explained above and measured with flow cytometry. Two elutriation approaches provided biological replicates consisting once of G1, early S, late S and G2 phase and secondly G1, S and G2 (Figure 4.5). RNA was isolated from selected fractions and applied for generation of Solexa-based small RNA libraries. The cloning protocols for both approaches differed mainly in the amplification step. Small RNAs isolated from G1, early S, late S and G2 cell cycle phase were amplified for 25 cycles. Small RNAs from the biological replicate containing G1, S and G2 were cloned by skipping the gel purification after 5' ligation and were sequenced twice after amplification with 21 and 23 cycles, respectively.

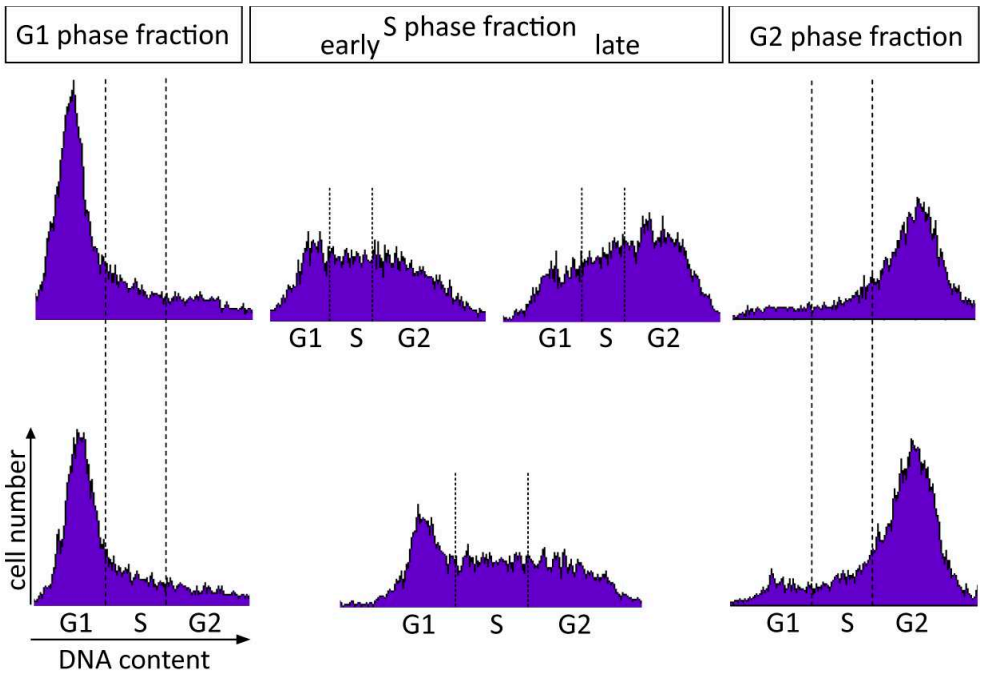


Figure 4.5 Overview of different cell cycle stages used for generation of Solexa-based small RNA libraries.

S2 cells were synchronized by counterflow centrifugal elutriation. The collected cell fractions were separated to be analyzed for the cell cycle position while the remainder of cells was added with Trizol for followed RNA isolation. For the analysis, cells were permeabilized with 70% ethanol, removed from RNA by incubation with RNase and stained with propidium iodide (PI). The fluorescence was measured by flow cytometry using an FL2 linear detector. The

diagrams present the DNA content against the cell number. The above shown fractions were selected to be applied for generation of Solexa-based small RNA libraries.

After sequencing, small RNA data sets were separated into different cell cycle phases due to the specific bar codes. Afterwards the linker introduced in 3' ligation and 5' ligation were discarded and the reads were selected in length from 11 nt to 28 nt. Mapping the libraries to the *Drosophila* genome reference convinced of good quality (73% to 92%) except for the G1 phase in the first run (40%) (Table 4.2). No mismatches were allowed during the mapping procedure. The second sequencing run which was performed with 21 and 23 amplification cycles, respectively, displayed nearly identical mapping pattern to the *Drosophila* genome reference as well as the databases of miRNAs and transposons. On the basis of this analysis, both data sets were clustered and used as one unit for further analysis unless otherwise stated.

	Library	Total no of insert 11-28 nt	Inserts matching the genome (% of total)	Inserts matching miRNAs (% of genome matching)	Inserts Matching transposons (% of genome matching)
1.seq run	**G1**	401565	159833 **(36.2)**	54209 **(33.9)**	14321 **(9.0)**
	early S	1954490	1659354 **(82.5)**	655994 **(39.5)**	269804 **(16.3)**
	late S	4259103	3709883 **(86.8)**	2588093 **(69.8)**	260257 **(7.0)**
	G2	2001897	1483444 **(73.2)**	594114 **(40.0)**	246990 **(16.6)**
2.seq run 21x	**G1**	5960166	5464320 **(90.7)**	4123923 **(75.5)**	427466 **(7.8)**
	S	3703505	3283836 **(86.9)**	1462429 **(44.5)**	634675 **(19.3)**
	G2	1358813	1207895 **(87.8)**	563770 **(46.7)**	236077 **(19.5)**
2.seq run 23x	**G1**	4774611	4318197 **(89.6)**	3344120 **(77.4)**	311155 **(7.2)**
	S	2748803	2394219 **(85.6)**	1127987 **(47.1)**	448355 **(18.7)**
	G2	800818	697574 **(86.1)**	334430 **(47.9)**	131154 **(18.8)**

Table 4.2 Analysis of deep sequencing libraries generated in this study.

First sequencing round contained Solexa libraries made of small RNAs isolated from G1, early S, late S and G2 cell cycle phases. Second sequencing round consisted of G1, S and G2 phases and was sequenced twice after amplification via 21 and 23 cycles, respectively. Both sequencing runs were processed by selection into different cell cycle phases due to the specific bar codes and selected for reads of 11 nt to 28 nt length. They were further mapped against the reference of *Drosophila* genome and listed as the percentage of total amount of reads to elucidate the quality of the libraries. Furthermore all libraries were mapped against the reference of miRNAs and transposons, respectively and displayed as the percentage of *Drosophila* genome mapping reads.

4.1.2.3 miRNAs stayed mainly unchanged during the cell cycle

Aberrant expression of miRNAs, e.g. miRNA gene deletions or amplifications which were reported in association with cancer, inhibit tumor suppressor genes or inappropriately activate oncogenes initiating the cancer process by uncontrolled cell proliferation (Cho 2007). This raised the questions: Do miRNAs set the timing of cell cycle phases and what changes occur in the small RNA profile across the cell cycle in Schneider S2 cells of *Drosophila*? In answer to that question, deep sequencing data sets were analyzed for miRNAs by aligning reads to the reference of known *Drosophila* miRNAs. As abovementioned, bantam was the most abundant miRNA in all analyzed cell cycle phases except for G1 which was performed in the first run. miR-184 and miR-8 were highly abundant but to a lesser extent than bantam. Together all three occupied 72% to 94% of all of miRNA matching reads (Table 4.1). In the following after exclusion of bantam, miR-184 and miR-8 artifacts, the remaining reads for each miRNA species were investigated for differences in abundance between different cell cycle phases. To do so, the clustered G1, S and G2 libraries were selected for further analysis. The read counts for each miRNA were normalized to the size of the library and depicted in a scatter plot. If specific miRNAs indeed fluctuate with the cell cycle, these small RNAs should be positioned distal to the regression line in the scatter plots. In contrast, most miRNAs in G2 and S phase demonstrated high correlation while G1 appeared to slightly differ from S and G2 (Figure 4.6A). Did the observed difference represent the real biological situation during the cell cycle? This was questionable as bantam, miR-184 and miR-8 diminished the size of the remaining miRNA pool and thereby potentially impeded reliable quantitative comparison.

To elucidate these findings by means of other methods, we generated expression profiles of cell cycle phases by qRT-PCR for a panel of the 80 most abundant *Drosophila* miRNAs. Therefore RNA extracted from G1, late S and G2 synchronized cells (RNA used for first sequencing run) was reverse transcribed and applied for miScript protocol. U6 snRNA control was used for normalization. Analyzed small RNAs were distributed from low to high expression levels. ΔCt values for each analyzed miRNA were compared between different phases via scatter plots in Figure 4.6B (G1 vs. late S, G2 vs. late S and G1 vs. G2). In comparison to the deep sequencing data, qRT-PCR demonstrated a higher correlation between G1, late S and G2. Regarding to the more abundant miRNAs demonstrated in the left corner of scatter plots, G1 was again slightly distinct from S and G2.

Summarizing, global miRNA analysis of deep sequencing libraries was limited due to the over-representation of specific miRNAs but still showed together with qRT-PCR that G1 phase displayed some very slight overall differences with no significant oscillation in abundance between different stages during the cell cycle.

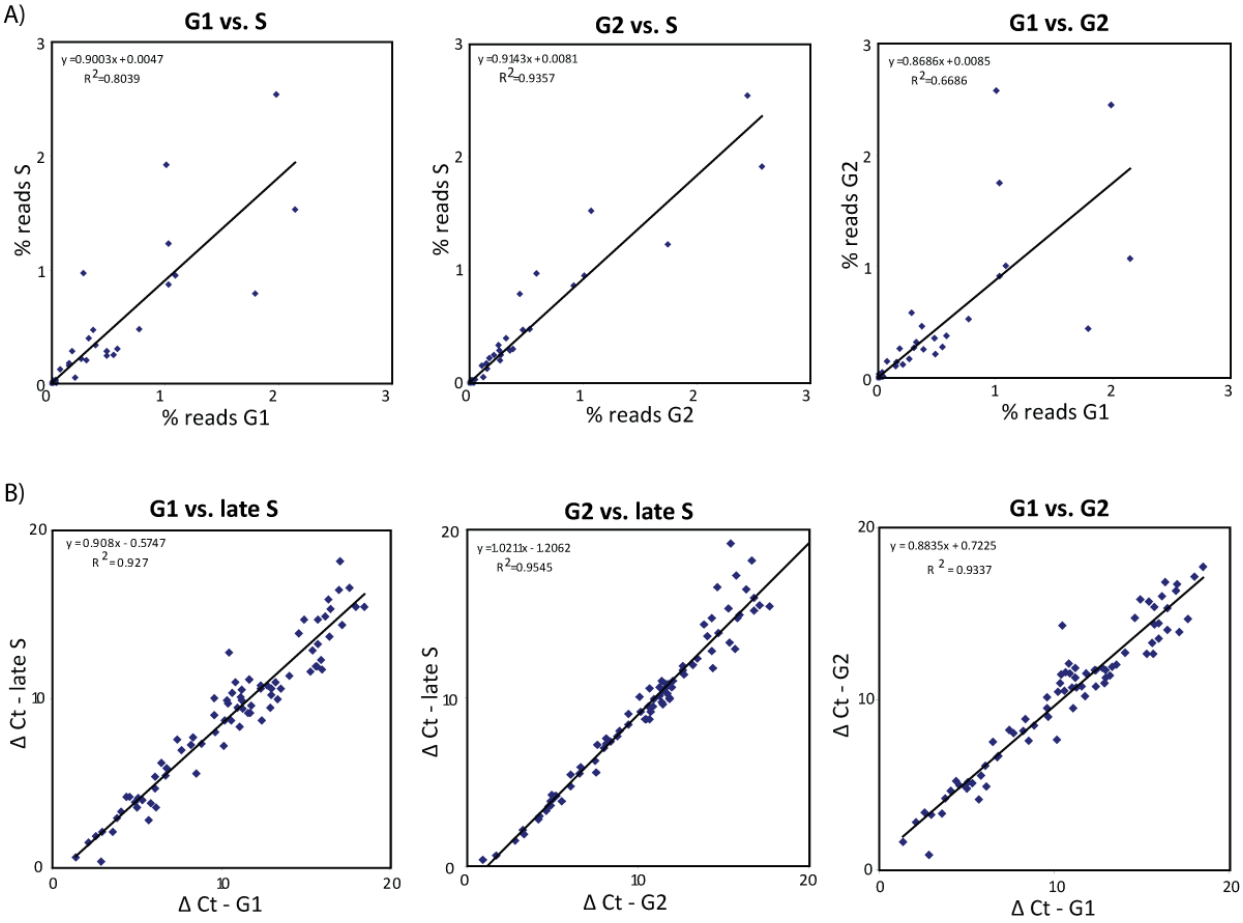


Figure 4.6 miRNAs stayed mainly unchanged during the cell cycle.

(A) Analysis of individual miRNAs was performed from G1, S and G2 libraries obtained from the second sequencing round. Bantam, miR-184 and miR-8 were excluded from this analysis due to their unfavorable cloning bias. The scatter plots present the abundance of remaining miRNAs from 11 nt to 28 nt most strongly represented by 21nt to 23 nt long miRNAs. They were normalized to the total *Drosophila* genome matching reads. In the scatter plots individual miRNAs from distinct cell cycle phases were compared with each other. (B) RNA which was applied for the first sequencing round was reverse transcribed and used for miRNA profiling via miScript protocol by qRT-PCR. The scatter plots represent the comparison of different cell cycle phases while each miRNA was normalized to the U6 snRNA control.

4.1.2.4 Small RNAs prefer heterochromatin over euchromatin site of origin

The aberrations in the nuclear division cycle were correlated with defects in the formation of centromeric heterochromatin (Deshpande, Calhoun et al. 2005). Furthermore *Drosophila* RNAi system must be intact to achieve targeted methylation of H3K9 and proper localization of HP1 in the heterochromatin formation (Pal-Bhadra, Bhadra et al. 2002; Pal-Bhadra, Leibovitch et al. 2004). Thus, the effector RNAs should be found in the S and G2 phases. In the following, data sets of G1, S and G2 (second sequencing run) were aligned to the reference of *Drosophila* transposons without permission of any mismatch. The size distribution normalized to the genome matching reads showed a clear peak at 21 nt as expected for endo-siRNAs (Figure 4.7A). The majority of 21 nt long TE-siRNAs were equivalently distributed between sense and antisense orientation which is explained due to their long dsRNA precursor structure (Figure 4.7B). Are specific transposons producing endo-siRNAs which oscillate in abundance with the cell cycle? The number of endo-siRNAs counts for each transposon was normalized to the total genome matching reads and finally compared between different cell cycle positions. All miRNAs were excluded from the normalization process due to the unfavorable cloning bias for specific miRNAs. Three transposons named *1731*, *297* and *blood* stood out as they produced the most abundant endo-siRNAs in descending order (Figure 4.7C). The remaining transposons did not present any outliers in expression pattern between G1, S and G2 phases. Taken together, no transposons were found to generate siRNAs which oscillate significantly with abundance during the cell cycle.

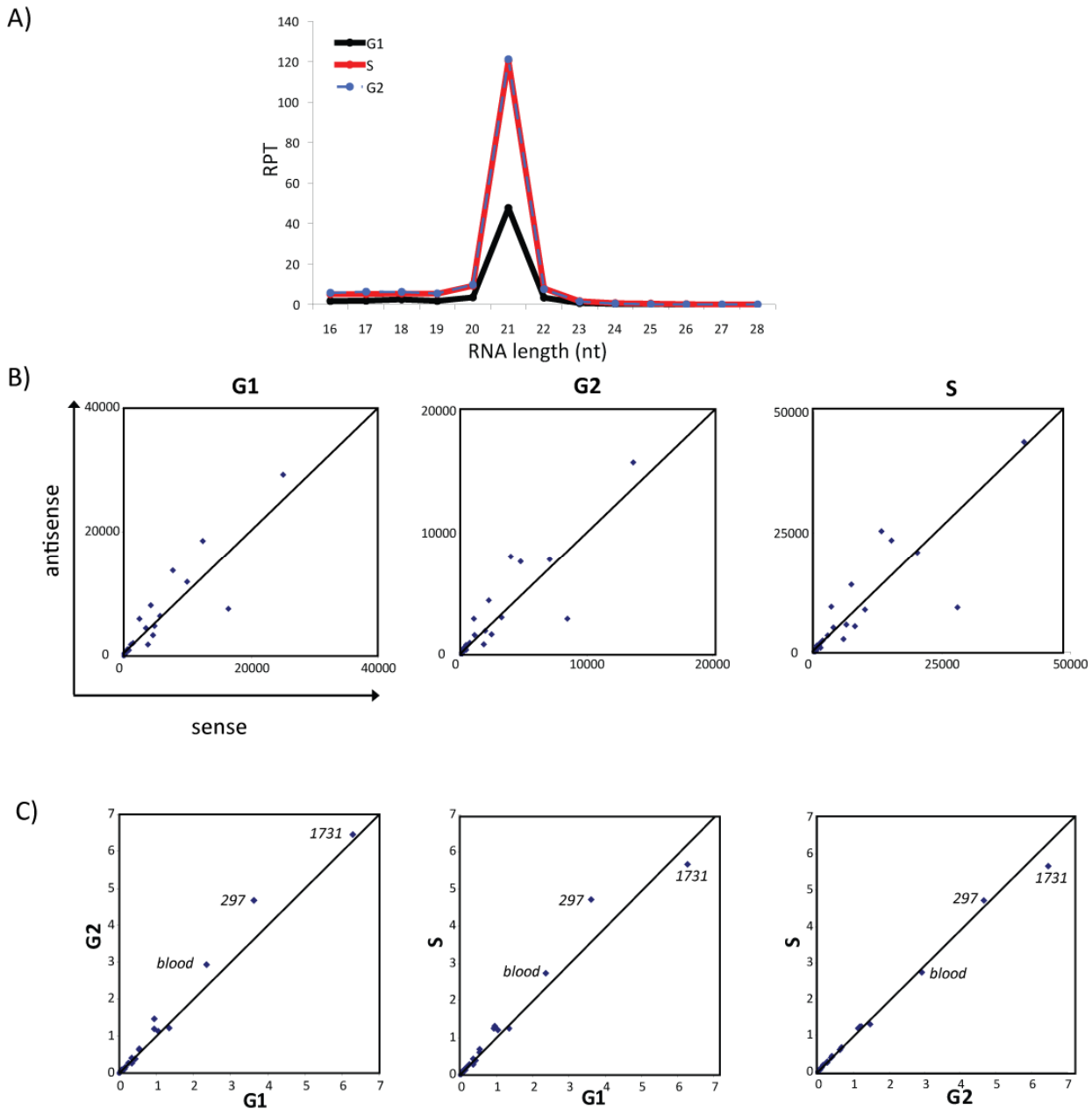


Figure 4.7 Endo-siRNA stayed mainly unchanged during the cell cycle.

(A) Analysis of endo-siRNAs was performed from G1, S and G2 libraries obtained from the second sequencing round. The counts matching against the reference for *Drosophila* transposons were normalized to genome matching reads and displayed as length distribution from 16 to 28 nt. (B) The endo-siRNA counts for individual transposons (*1731* and *297* excluded) were normalized to genome matching reads. Within scatter plots the sense reads are presented versus the antisense reads for each cell cycle phase. (C) The normalized endo-siRNA counts are depicted in scatter plots and the cell cycle phases were compared against each other.

Does small RNA production differ in genomic origin relating to different cell cycle positions? To achieve this, heterochromatic regions were analyzed versus euchromatic regions. *Drosophila melanogaster* has four pairs of chromosomes: the autosomes 2, 3 and 4 and X/Y sex chromosomes. The mapping indexes were generated for each chromosome while they were further separated into euchromatic and heterochromatic regions. For this analysis all miRNAs were excluded from the normalization process due to the unfavorable cloning bias for specific miRNAs. The remaining reads were mapped onto the genome. The resulting number of counts for small RNAs was normalized to the length of each mapping reference and the library size matching to *Drosophila* genome (Figure 4.8). The heterochromatin generated a higher production of small RNAs compared to the euchromatin corresponding to both sequencing runs. Endo-siRNA precursors derive from repetitive sequences, sense-antisense pairs or long stem-loop structures. The observed preference for heterochromatic origin was explained as the repetitive sequences were embedded in the heterochromatin. Furthermore all cell cycle phases from the second sequencing run demonstrated comparable amounts while G1 phase was slightly lower in Figure 4.8A. The difference in G1 in the first sequencing run is higher in Figure 4.8B due to the low depth and quality of the appropriate library.

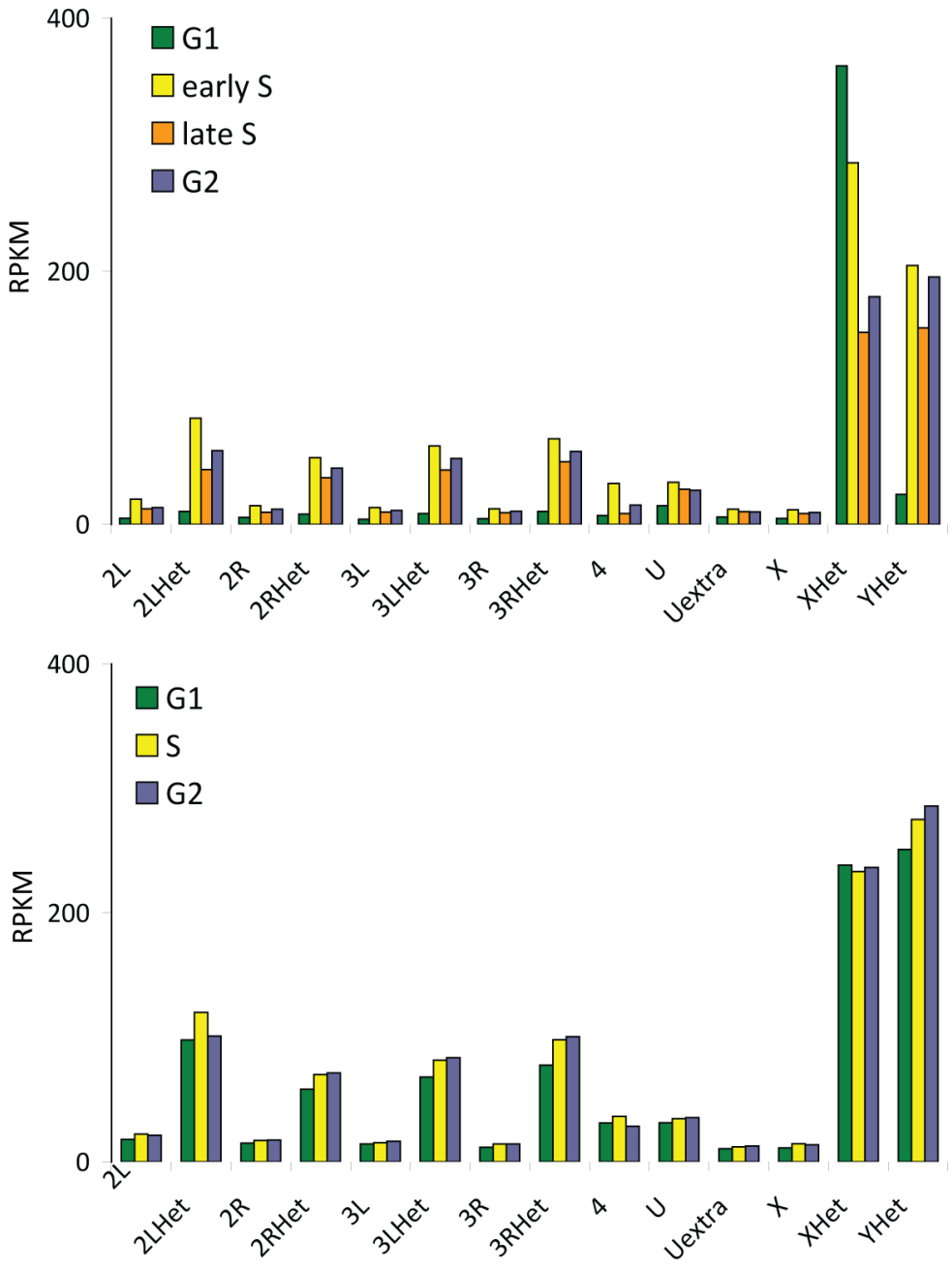


Figure 4.8 Small RNAs prefer heterochromatin over euchromatin site of origin.

Small RNAs mapped against transposons were analyzed for their genomic localization regarding to heterochromatin versus euchromatin structure. Therefore fasta files for all four chromosomes and sex chromosomes were applied from FlyBase (release5.21) and utilised for the mapping procedure of both sequencing rounds. The counts for every analyzed cell cycle phase were normalized to the length of each mapping reference and the library size matching to *Drosophila* genome referred as reads per kilobase of mapping reference per million mapped reads (RPKM).

4.1.2.5 tRNA-derived small RNA

Intronic regions were described to harbor noncoding small RNAs (Rodriguez, Griffiths-Jones et al. 2004). Their presence in introns has implications for the biogenesis of mature small RNAs and host mRNA indicating that the regulation of expression levels of small RNAs are highly important. We were interested in small RNAs originating from introns with regard to the cell cycle. To concentrate on sequences beside known miRNAs and siRNAs, we removed sequences matching against stemloops of miRNAs and transposons and mapped the remainder against *Drosophila* intron reference. A deeper look at the genomic localization of resulting RNA sequences allowed us to identify the most abundant RNA sequence with 21 nt length matching to a very specific position originating 3' to the tRNA (tRNA:E4:62Ad) (Table 4.3). We refer to this 5'-phosphate and 3'-hydroxylated tRNA-derived small RNA as tRNA fragment (tRF) in the remainder of this study. In order to verify the abundance observed in deep sequencing libraries (Table 4.3), RNA was extracted from G1, early S, late S and G2 synchronized cells , then reverse transcribed and applied for miScript protocol. For qPCR procedure primer for rp49-mRNA were used as control. The tRF was well detectable in comparison to the control (Table 4.3). Taken together, the identified tRF in the intronic region was evidently expressed via both methods while qPCR did not exhibit any differences in abundance during the cell cycle.

A)

	total no of insert 17-24 nt ø stemloops ø transposons	**Inserts matching introns (% of total)**	**tRF (tRNA:E4:62Ad) (% of total)**
G1	1268665	114872 (9.1)	11109 (0.9)
S	1155042	174004 (15.1)	26196 (2.3)
G2	390177	57092 (14.6)	7439 (1.9)

	total no of insert 20-25 nt ø stemloops ø transposons	**tRF (tRNA:E4:62Ad) (% of total)**
G1	119736	23 (0.02)
early S	268721	5500 (2.1)
late S	501053	10338 (2.1)
G2	358385	4655 (1.3)

B)

	Ct - tRF	**Ct - rp49**
G1	21.3	18.9
early S	21.6	19.9
late S	21.8	19.9
G2	21.5	19.7

Table 4.3 tRF abundance analyzed via deep sequencing and qRT-PCR.

(A) Both sequencing runs were analyzed for the abundance of tRF. Reads matching to stemloops and transposons were excluded from this analysis. The resulting data sets used for further investigation differ in length selected fraction (17 nt to 24 nt and 20 nt to 25 nt) which is meaningless as both comprise the tRF which is exactly 21 nt long. The counts were obtained after mapping to the reference of tRF. They were further normalized to the total amount of reads matching to *Drosophila* genome except for stemloops and transposons. (legend continued on p. 62)

(legend Table 4.3 continued) (B) RNA isolated from G1, early S, late S and G2 (first sequencing run) was reverse transcribed and endogenous tRF levels were measured by qRT-PCR. Ct values were listed from tRF and rp49 as a control.

4.1.2.5.1 Direct inhibition and over-expression of tRF does not influence the cell cycle

To further investigate the role of the specific tRF during the cell cycle, inhibition and over-expression of tRF were methods of choice. First we treated S2 cells with 2'-*O*-methyl-modified RNA oligonucleotides directed against the small RNA. Hereafter its cell cycle distribution was monitored by staining the cells with PI and measuring the fluorescence by flow cytometry. Inhibition of tRF did not lead to any changes in the distribution of cell cycle phases (Figure 4.9).

Pre-tRNAs are transcribed by RNA polymerase III. Due to the direct proximity of tRF to the 3' end of the mature tRNA, the expression of tRF could depend on the tRNA transcription. The construct for over-expression contained the sequence including tRNA and adjacent tRF within pBluescript KS^+. The control cells were transfected with pBluescript KS^+ without the insert. Over-expression of tRF did not perturb the cell cycle (Figure 4.9).

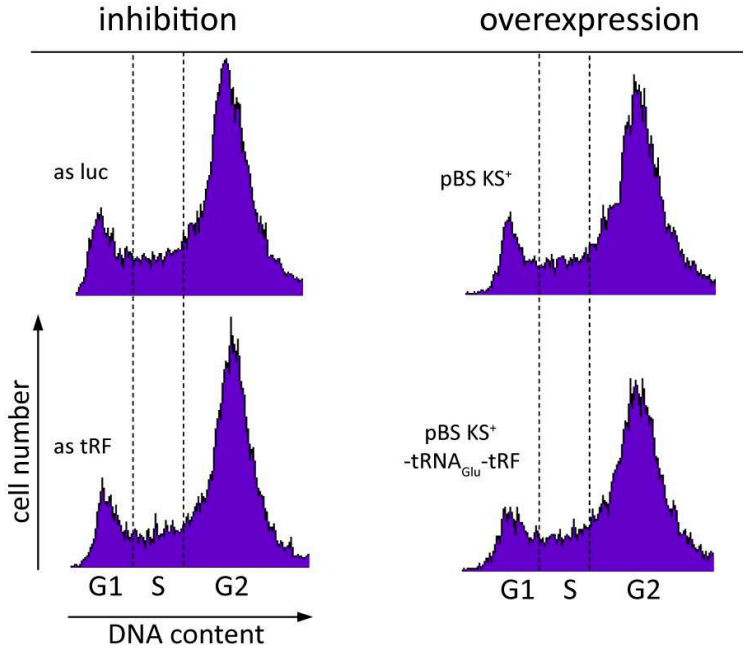


Figure 4.9 Direct inhibition and over-expression of tRF does not influence the cell cycle.

(A) tRF was inhibited by transfection of antisense oligonucleotides for 3 days. Control cells were directed against part of the firefly luciferase coding sequence derived from glow worm. After the cells were harvested and permeabilized with 70% ethanol, RNA was removed by incubation with RNase and DNA was stained with propidium iodide (PI). The fluorescence was measured by flow cytometry using an FL2 linear detector. The diagrams present the DNA content against the cell number. (B) pBS KS^+ was cloned with insert containing $tRNA_{Glu}$-tRF as a precursor of tRF. S2 cells were transiently transfected. pBS KS^+ without insert served as control. The further protocol for determination of cell cycle position is explained above.

4.1.2.5.2 tRF is processed by Jhl-1 and stabilized by La

Human tRNA-derived small RNAs were shown to act in the global regulation of RNA silencing (Haussecker, Huang et al. 2010). We were interested to characterize the functional relevance of tRF, possibly in novel modes of gene regulation. So far two distinct biogenesis pathways were identified, Dicer-dependent (Babiarz, Ruby et al. 2008) and Dicer-independent mode by usage of the tRNA processing machinery. In the latter case the 5' end of tRF was determined by the tRNA processing enzyme RNaseZ, an endonuclease leaving a 3'-hydroxyl and 5'-phosphate at the cleavage site (Mayer, Schiffer et al. 2000) and the 3' end generated by transcription termination of RNA polymerase III. Our aim was to investigate which mode of biogenesis applies for the identified tRF. Therefore we depleted both known isoforms of Dcr-1 and Dcr-2 protein generally processing miRNAs and siRNAs, respectively. Furthermore we depleted other miRNA and siRNA biogenesis factors including Ago1, Ago2, Loqs, specific isoform Loqs-PD, Drosha and R2D2. To analyze the Dcr-independent processing the sole *Drosophila* RNase Z ortholog named Jhl-1 was depleted. In addition, the human autoantigen La was published to bind 3' termini of all nascent polymerase III transcripts (Mathews and Francoeur 1984; Stefano 1984). Based on these findings, La was also depleted. Figure 4.10 displays a Northern Blot, after RNAi depletion of listed proteins, GFP and DsRed as controls , which was probed for tRF and miR-277. It presented no dependence on any mi/siRNA biogenesis factors suggesting that tRF is processed via tRNA processing machinery. This is evident as tRF precursor containing tRNA and tRF was highly accumulated after depletion of Jhl-1. Interestingly depletion of La did not accumulate for the precursor but diminished mature tRF amount indicating that La is not involved in the biogenesis of tRF but in the stabilization of its mature form. miR-277 probed Northern Blot showed that mir-277 is dependent on Drosha known to be the responsible enzyme in the first processing step of miRNAs. Furthermore depletion of Dcr-1 and Loqs but not Loqs-PD isoform accumulated the precursor, while depletion of Dcr-1 additionally resulted in loss of mature miR-277. The effect observed for depletions of Loqs is based on depleted Loqs-PB isoform. No loss of mature miR-277 is observed because Loqs-PB confers more specificity but is not involved in the dicing process of Dcr-1. Taken together, tRF identified in Solexa libraries is transcribed via polymerase III. Its 5' end is generated by Jhl-1 enzyme and the released mature small RNA is then stabilized by La.

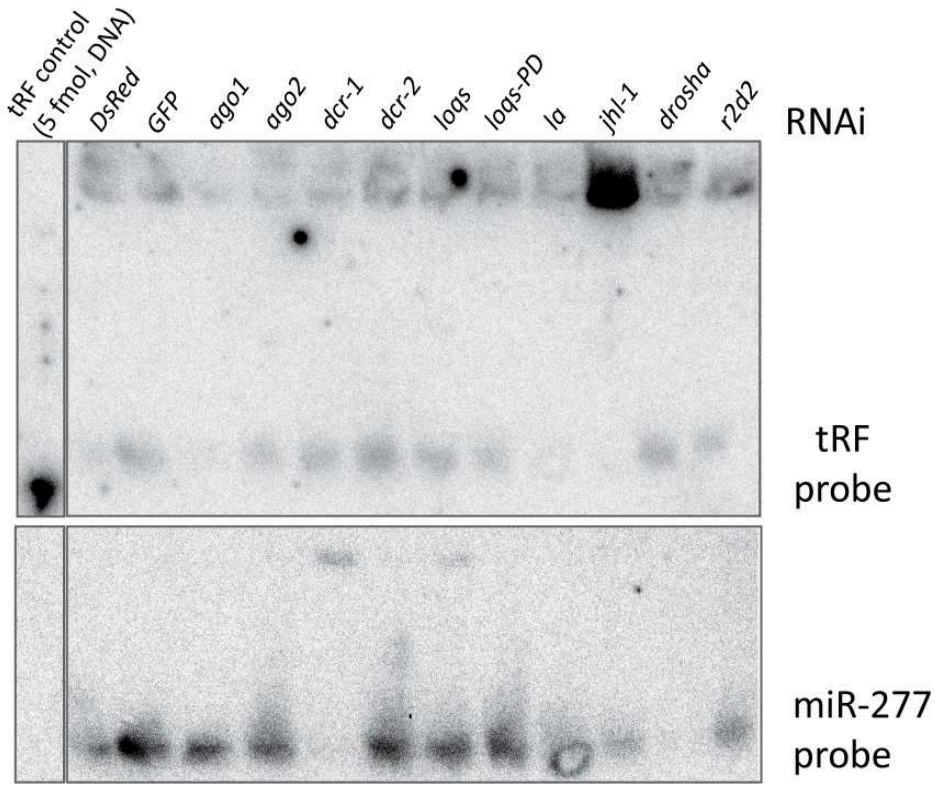


Figure 4.10 tRF is processed by Jhl-1 and stabilized by La.

S2 cells were treated twice with dsRNA constructs against the components of the small RNA pathways, *jhl-1* and *la*. Treatment with RNAi against DsRed and gfp served as control. RNA was isolated and analyzed by Northern blotting. 5 fmol of the tRF DNA oligo was also loaded. The analysis was performed with DNA probe against tRF (upper panel) while miR-277 (bottom panel) served as control for dependence on miRNA biogenesis factors.

4.1.2.5.3 tRF is not loaded into Ago2-RISC but effect Ago1-RISC silencing

Human tRFs were shown to be preferentially associated with the nonslicing Argonautes 3 and 4 (Haussecker, Huang et al. 2010). In order to investigate the association of identified tRF we immuno-purified Ago1 and Ago2 with Ago1- or Flag-tagged beads while myc-coated beads served as a control. The RNA was isolated from recovered fractions and applied for Northern blotting which was finally probed for tRF, bantam and miR-277. tRF slightly associated with Ago2 (Figure 4.11). The miRNA bantam is described to be predominantly loaded into RISC with Ago1 which was clearly confirmed in Figure 4.11 (Shah and Forstemann 2008). The miR-277/277* duplex is more extensively basepaired than typical miRNA duplex that are generally interrupted by mismatches which favor loading into Ago2 (Forstemann, Horwich et al. 2007). miR-277 was loaded into Ago2 to higher extent than bantam as already expected (Figure 4.11).

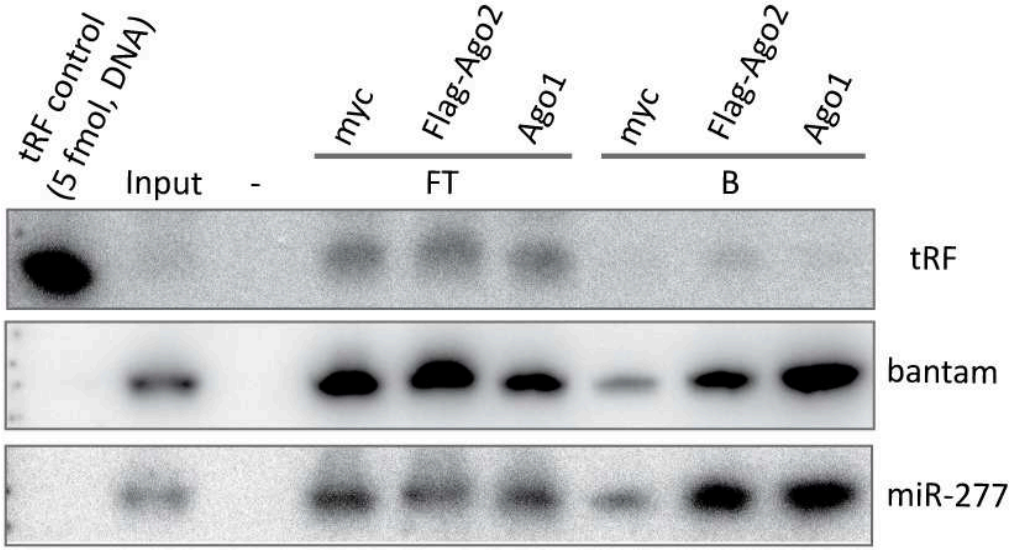


Figure 4.11 tRF slightly associates with Ago2.

Stable Flag-Ago2 expressing cells (Ago2FlagHA_4_2 kindly given by Katharina Elmer, monoclonal) were used for Ago1 and Ago2 co-immunopurification with Ago1- or Flag-coated beads, respectively (1 mg cell extract). Myc-coated beads served as control. RNA was isolated and the isolated fractions were analyzed by Northern blotting (input (10%), flow through (FT, 10%) and bound fractions (B)). 5 fmol of the tRF DNA oligo was also loaded onto 15% Acrylamide-Urea gel. The blot was probed with tRF, bantam and miR-277 DNA oligo to verify the association from tRF with Ago2 or Ago1.

We went on to test the post-transcriptional *trans*-silencing capacity of tRF. Ago1 and Ago2 differ due to distinct preferences for the architecture of their target sites. Ago2 targets a perfect match reporter while Ago1 will silence a bulged match reporter (Shah and Forstemann 2008). Therefore we generated reporter with inserted sequences in the 3' UTR of the GFP coding sequence which either harbored one or two perfectly matching sites for the tRF to test its role in Ago2-mediated silencing. Reporter generated with four bulged matching sites should give an insight if tRF is capable of silencing via Ago1-RISC complex.

Cells were transiently transfected with GFP-expression reporter constructs and then measured for the fluorescence over a time interval from day 2 to day 9. GFP fluorescence was normalized to pKF63, a control without any tRF-binding sites in the 3’ UTR of GFP. The GFP levels were unchanged in case of reporters with perfect matching sites. Hence tRF was not capable of Ago2-mediated silencing (Figure 4.12). To investigate the involvement in Ago1-RISC system, S2 cells were transiently transfected with reporter constructs harboring four partially complementary binding sites for tRF in the 3’ UTR of GFP. It resulted in repression of the reporter of about 40% which indicated tRF to affect Ago1-mediated silencing and not on Ago2-loaded RNAs. In contrast to this result, tRF was not shown to associate with Ago1 in the Northern Blot (Figure 4.11).

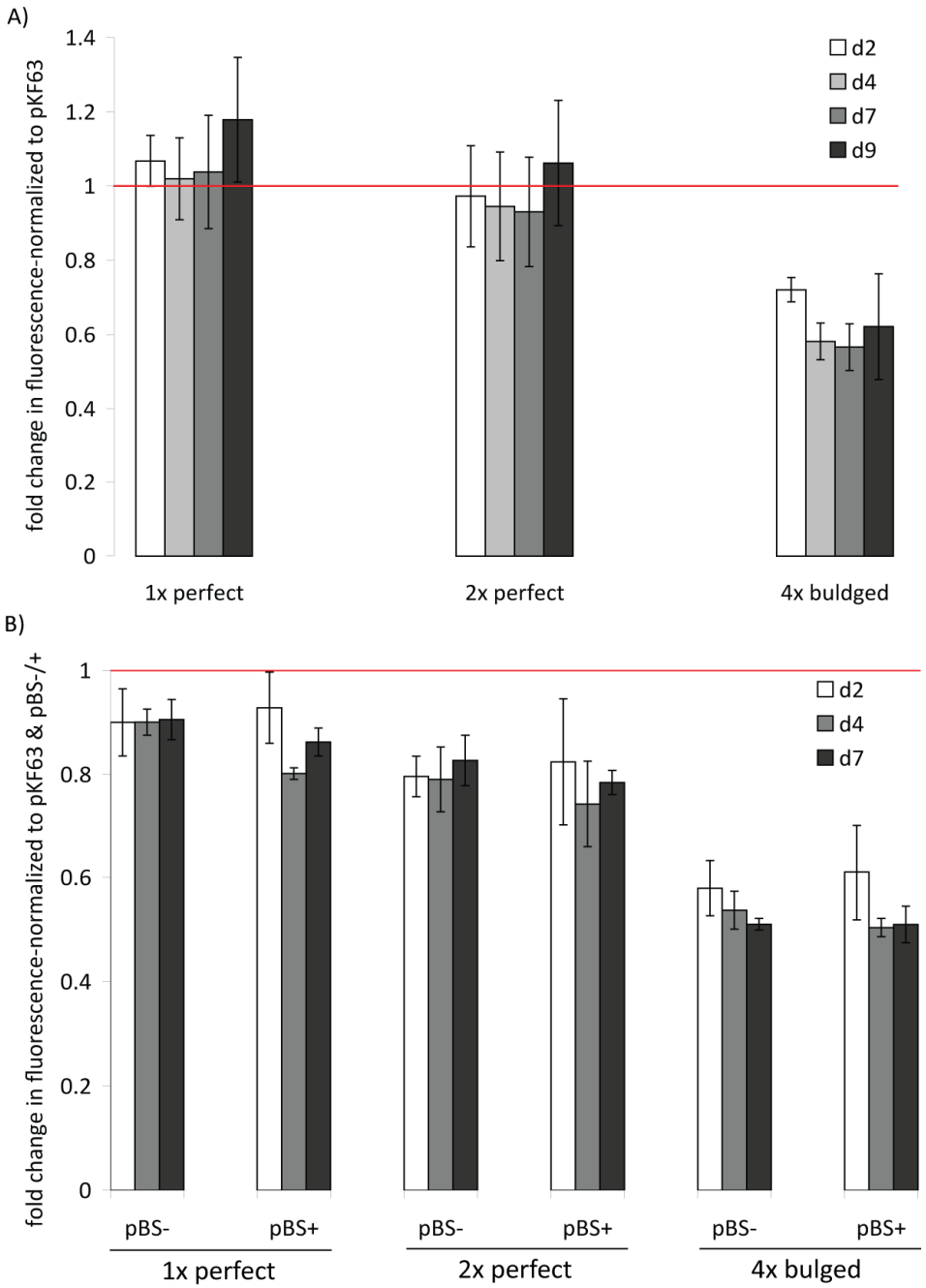


Figure 4.12 tRF is not loaded into Ago2-RISC but affect Ago1-RISC silencing.

(A) S2 cells were transfected with tRF reporter constructs. They contain one or two perfectly matching sites for the tRF in the 3' UTR of GFP (pKF63) to investigate Ago2-mediated silencing while four bulged reporter allow analysis of Ago1-RISC silencing. GFP fluorescence was measured by flow cytometry and normalized to the control transfected with pKF63 without binding sites in its 3' UTR. Values are the mean of 4 experiments. The horizontal line marks no change compared to the control. (B) In addition to the reporter constructs for tRF as above mentioned S2 cells were transfected with construct expressing tRF precursor containing $tRNA_{Glu}$ and tRF. GFP fluorescence was measured by flow cytometry and normalized to the control transfected with pKF63. Values are the mean of 4 experiments. The horizontal line marks when GFP levels were unaltered compared to the control.

Next we overexpressed tRF and analyzed repression effect on the reporter constructs which was expected to further repress the fluorescence. Therefore we transiently transfected the reporter and simultaneously overexpressed plasmid bearing the tRF precursor containing tRNA and tRF (pBS KS^{+}-$tRNA_{Glu}$-tRF). In case of Ago2- as well as Ago1-mediated silencing no difference was observed after expression of tRF precursor (Figure 4.12B). To verify the overexpression of tRF, RNA was isolated from cells after overexpression and applied for Northern blotting which showed that the precursor was accumulated but its processing did not proceeded as no increase of mature tRF was observed (Figure 4.13) This explained no difference in GFP expression after additional transfection of $tRNA_{Glu}$-tRF construct. Summarizing, the reporter construct with four partially complementary binding sites for tRF showed reproducible effect on Ago1-RISC mediated silencing which could be a direct effect of tRF or proceeded without any participation of tRF due to other cellular processes.

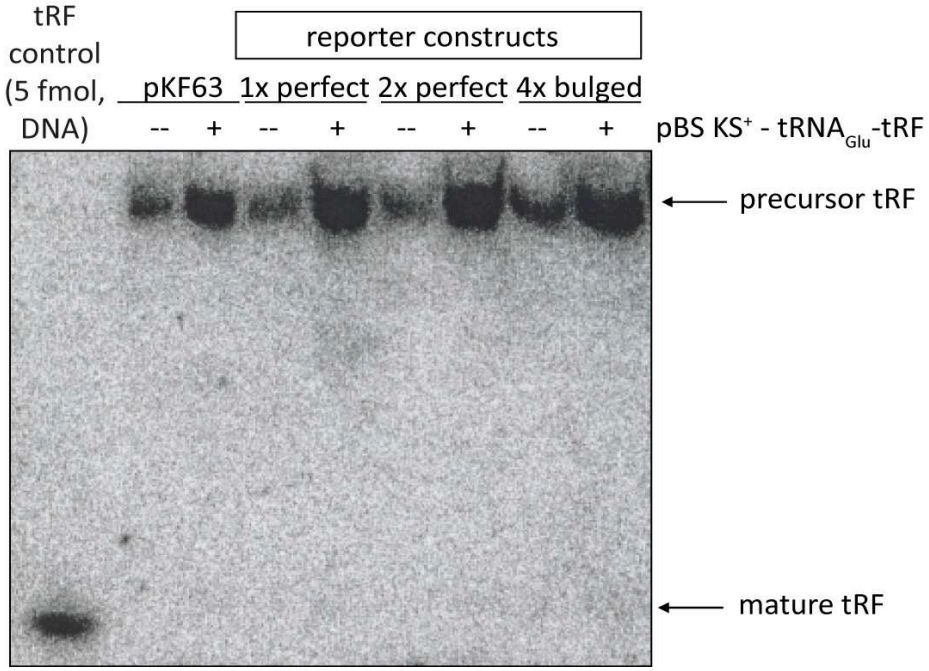


Figure 4.13 After over-expression tRF precursor was over-expressed but not further processed in mature tRF.

To verify the overexpression of tRF, S2 cells were transfected with reporter constructs for tRF (they contain one/two perfect or four bulged binding sites for tRF in 3' UTR of GFP) and construct expressing tRF precursor containing $tRNA_{Glu}$ and tRF. The construct without tRF precursor served as a control. RNA was isolated and applied for Northern blotting. 5 fmol of the tRF DNA oligo was also loaded onto 8% Acrylamide-Urea gel and the blot probed with tRF oligo.

4.2 Part II

4.2.1 Backcrossing of *loqs*^{ko} and *r2d2* mutants

Transposons are a major source of genome variability and their activity and genomic distribution may differ between fly strains. To facilitate our comparative analysis of the *loqs*^{ko} (Park, Liu et al. 2007) and *r2d2* mutants (Liu, Jiang et al. 2006), which derived from distinct genetic backgrounds, we performed one round of backcrossing for both mutations using *w*^{1118} stock.

4.2.1.1 Backcrossing scheme for *loqs*^{ko} mutant

The null allele *loqs*^{ko} was generated through ends-out homologous recombination (Park, Liu et al. 2007). As a result, the entire *loqs* open reading frame (ORF) was replaced with a *mini-white* transgene which serves as a marker gene. Park and colleagues described that the early development occurred normally while the viability dropped precipitously at the transition from pupa to adult stadium. 90% of the mutant flies died during eclosion and the remaining flies died after emerging. The *loqs* isoform Loqs-PB was shown to be essential to avoid defects in embryonic development and GSC maintenance (Forstemann, Tomari et al. 2005). As our interest is centered on Loqs-PD, we used a fly strain that also carried a Loqs-PB transgene on the 3rd chromosome. Virgins of a *loqs*^{ko} mutant with restored Loqs-PB function were crossed with wild type *w*^{1118} males (P) (Figure 4.14). After the selection against Cy and Sb, F1 offspring was further crossed with Kr/Cyo; D/TM6C, Sb, Tb double balancer males, to obtain offspring with balanced 2nd (CyO) and 3rd (TM6, Tb, Sb) autosomes. The following sibling mating (1 female + 1 male) allowed generation of stable balanced stocks producing homozygous and heterozygous *loqs*^{ko} mutants rescued with Loqs-PB. As balancer chromosomes carry recessive lethal mutations, individuals containing homozygous balancer chromosomes are nonviable.

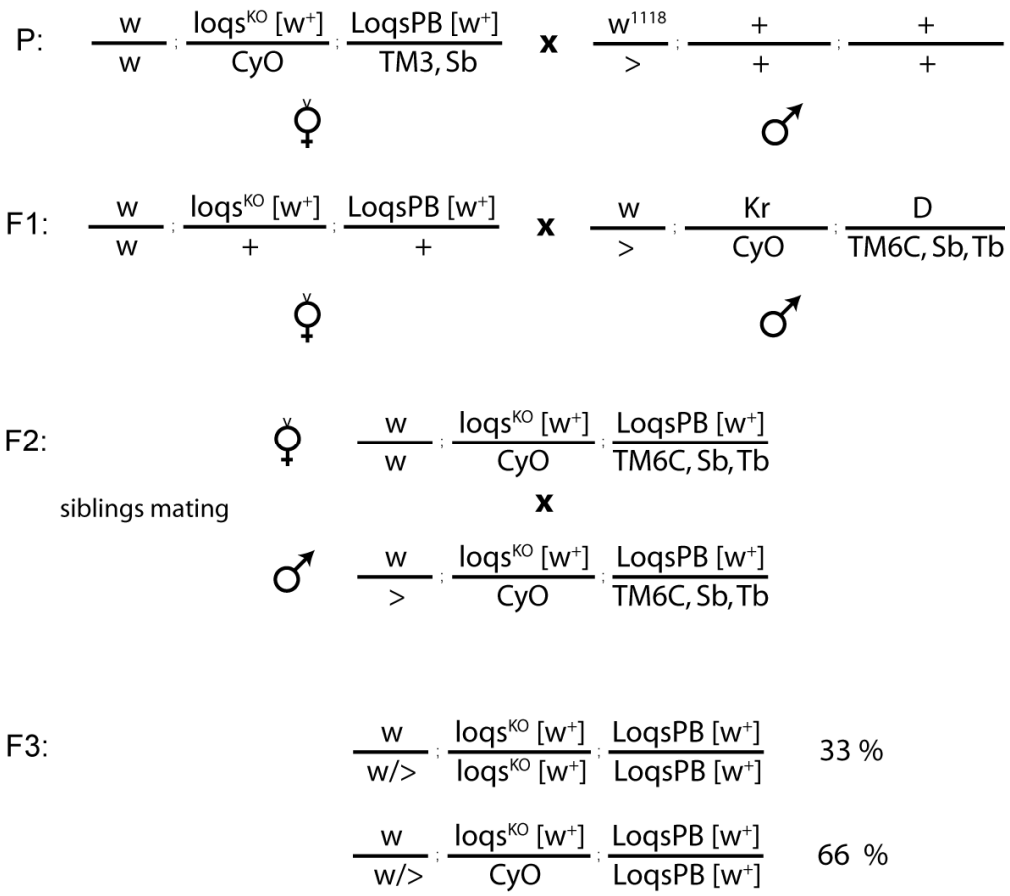


Figure 4.14 Backcrossing schema of *loqs*^{ko} mutant in *w*^{1118} genetic background.

Virgins of heterozygous *loqs*^{ko} mutant rescued with Loqs-PB via P-element insertion were mated with *w*^{1118} males (P). The offspring was selected for flies with *loqs*^{ko} deletion based on a *mini-white* transgene expressing red eye color pigment serving as a marker gene. Hereafter the offspring was mated with male double balancer flies to obtain balanced autosomes (F1). The third mating of siblings (F2) should produce homozygous and heterozygous *loqs*^{ko} mutants (F3).

w^{+} = gene for red eye color (intensity is additive); CyO = "Curly of Oster", curly wings; TM6, Sb, Tb = TM6 balancer chromosome with Sb (stubble = short thoracic bristles) and Tb (tubby = segmentation phenotype with short larval form) as phenotypic markers; ">" represents the male Y-chromosome.

The sibling mating resulted in a number of fly lines which were analyzed by genomic PCR with an appropriate primer pair within the first two exons of the *loqs* locus. The *loqs-PB* transgene consisted of cDNA lacking intronic sequences, therefore PCR products from wild type *loqs* locus and *loqs-PB* could be differentiated by different size of PCR products (320 nt versus 247 nt, respectively). All *loqs*^{ko} fly stocks (lane 1 to 9 except lane 5) showed knock out of both *loqs* alleles according to the missing product at 320 nt while Loqs-PB was present according to the 247 nt product (Figure 4.15A). As a control, the heterozygous *loqs* mutant, employed in the first crossing step (Figure 4.14, P), was verified and confirmed to the appropriate genotype (Figure 4.15A, lane 10). With regard to the following analysis two fly stocks were selected (lane 3 and 7) and their corresponding heterozygous mutants were further analyzed to complete the characterization of the genotype. They were recognized due to the curly wing marker (Cy). PCR products comprised one wild type copy of *loqs* locus (320 nt) as well as *loqs-PB* (247 nt) as already expected (Figure 4.15B).

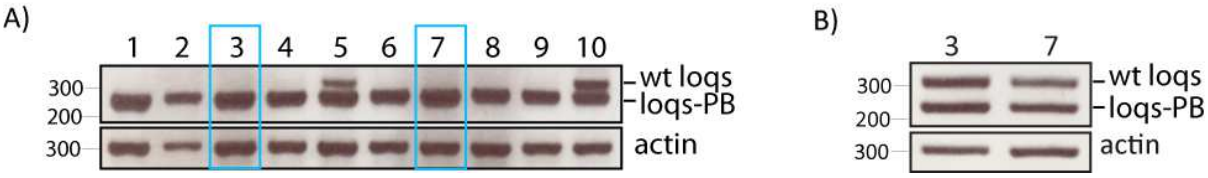


Figure 4.15 Fly stock mapping of *loqs*^{ko} mutants.

A) Nine fly stocks (lane 1 to 9) resulted from the sibling mating were mapped for *loqs* deletion by PCR. The genomic DNA was isolated according to the Berkeley Drosophila Genome Project protocol. Wild type *loqs* gene and *loqs-PB* were distinguished by size. *Loqs* allele was detected by 320 nt and *loqs-PB* at 247 nt length. In lane 10 heterozygous *loqs* mutant applied in the first crossing step was performed as a control to visualize *loqs-PB* and wt *loqs*. Actin served as DNA quality control. B) Flies with Cy marker in fly stocks 3 and 7 were tested for heterozygous *loqs*^{ko} genotype.

4.2.1.2 Backcrossing scheme for *r2d2* mutant

The *r2d2* deletion flies were generated by imprecise excision of a P-element inserted near the *r2d2* locus (Liu, Jiang et al. 2006). This resulted in a 4.9 kb deletion, which removed the entire *r2d2* ORF as well as 1 kb of upstream and 3 kb of downstream sequences, resulting in a null mutant allele. The deletion of the 4.9 kb region could be rescued by an intact *r2d2* gene (Liu, Jiang et al. 2006).

To obtain *r2d2* mutant flies with a genomic background more similar to the *loqs*^{ko} mutant described above, virgins of the *r2d2* mutant were crossed with the wild type *w*^{1118} males (P) (Figure 4.16A). F1 offspring was further crossed with Kr/Cyo; D/TM6C, Sb, Tb double balancer males, to obtain offspring with balanced 2nd (CyO) and 3rd (TM6, Tb, Sb) autosomes. The deletion of *r2d2* is not recognizable by a marker gene. The second crossing step produced heterozygous mutants where 2nd chromosome was balanced over CyO but contained wild type *r2d2* or deleted *r2d2*. The following sibling mating (1 female + 1 male fly) generated distinct genetic compositions, which could not be distinguished by visible markers (Figure 4.16B).

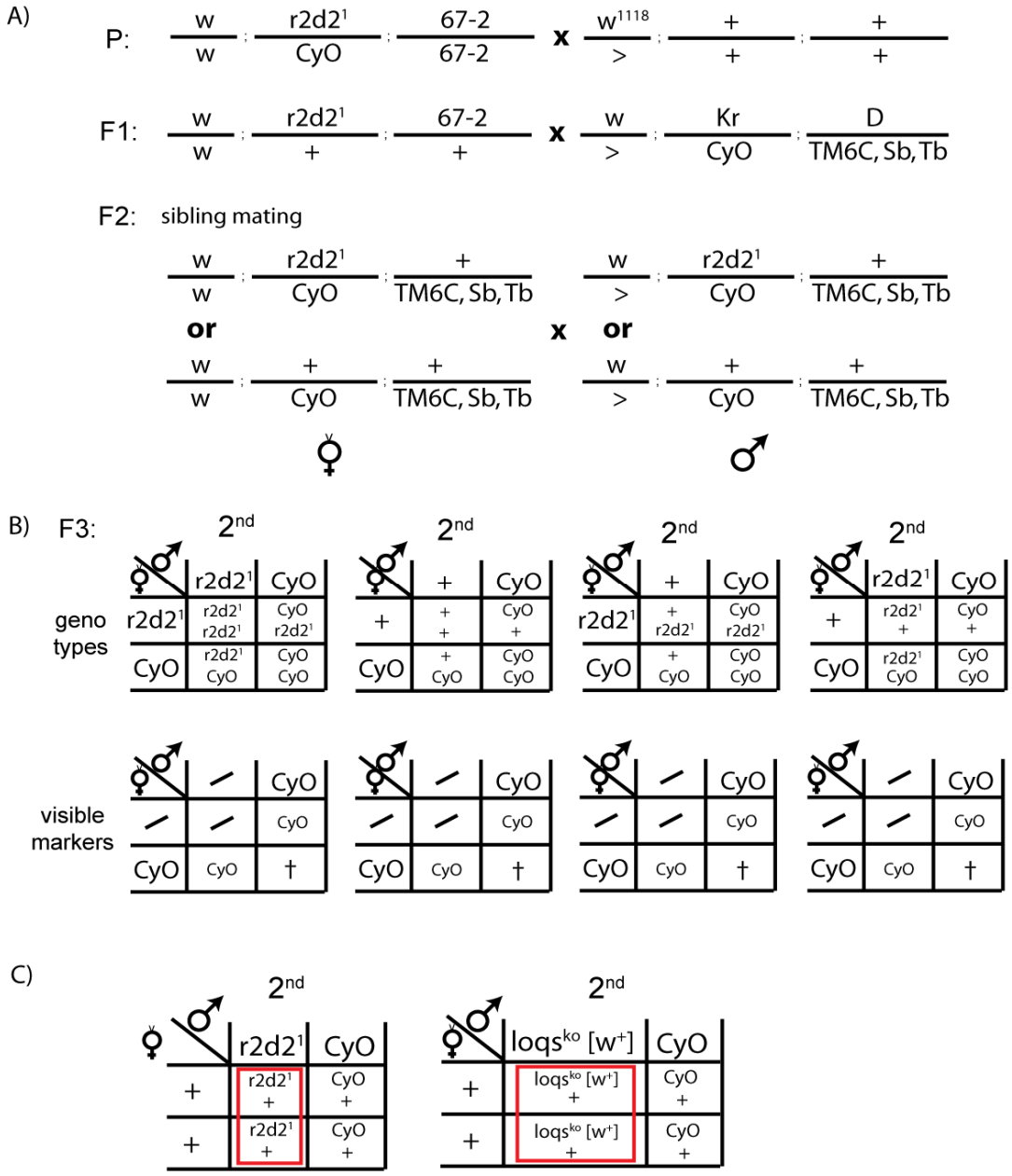


Figure 4.16 Backcrossing schema of *r2d2* mutant in *w*^{1118} genetic background.

A) Virgins of heterozygous *r2d2* mutant balanced over CyO were mated with *w*^{1118} males (P). The offspring was mated with male double balancer flies to obtain balanced autosomes (F1). As *r2d2* mutation was not selectable by a marker gene, the third mating of siblings (F2) resulted in four possible combinations of crossing (F3) shown for the relevant chromosome 2. The upper panel demonstrated the resulting genotypes of the offspring while the lower panel presented the corresponding visible phenotypic marker. C) Heterozygous *r2d2* flies balanced over Cyo were backcrossed with *w*^{1118} wt males. The genotype of the offspring for *r2d2* and *loqs*^{ko} mutant was depicted for the relevant 2nd chromosome. Flies marked in red were selected for further experiments.

w^{+} = gene for red eye color (intensity is additive); CyO = "Curly of Oster", curly wings; TM6, Sb, Tb = TM6 balancer chromosome with Sb (stubble = short thoracic bristles) and Tb (tubby = segmentation phenotype with short larval form) as phenotypic markers; ">" represents the male Y-chromosome.

Genomic PCR was the method of choice for recognition of deleted *r2d2* by using a combination of primer comprising *herp* (CG14536, upstream of *r2d2*), *r2d2* and *cdc14* (CG7134, downstream of *r2d2*) in a collection of derived fly stocks (lane 3 to 9). The PCR conditions used for *herp* and *cdc14* primer did not allow amplification of the 6.3 kb PCR product in the wild-type sample. From mutants with straight wings (possibly homozygous genotype), lane 4 and 6 were selected as the flies were devoid of a wt *r2d2* allele (Figure 4.17A, hepr_s + r2d2_as) and generated a correct size product with the *herp* and *cdc 14* primers (Figure 4.17A, hepr_s + cdc14_as). The corresponding heterozygous flies from stocks 4 and 6 yielded PCR products with both primer combinations as expected (Figure 4.17B). As a control, the heterozygous *r2d2* mutant (Figure 4.17A, lane 2) employed in the first crossing step (Figure 4.16, P) and the wild type w^{1118} stock were verified (Figure 4.17A, lane 1). Taken together, genomic PCR allowed fly stock mapping to characterize stably balanced homozygous and heterozygous *r2d2* mutants.

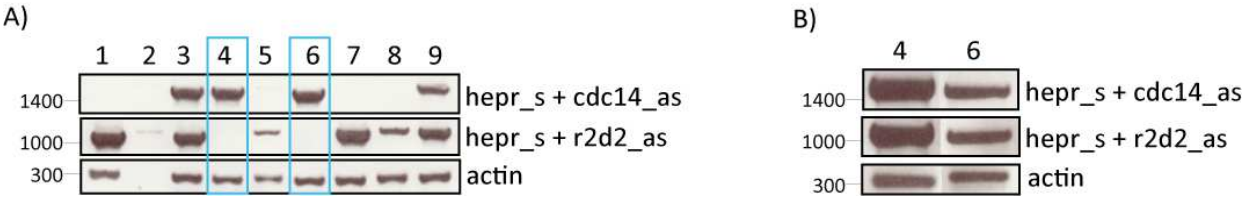


Figure 4.17 Fly stock mapping of *r2d2* mutants.

A) Seven fly stocks (lane 3 to 9) resulting from the sibling mating were mapped for *r2d2* deletion by PCR. The genomic DNA was isolated according to the Berkeley Drosophila Genome Project protocol. Primers were used for *herp* (upstream of *r2d2*), *r2d2* and *cdc14* (downstream of *r2d2*). As control w^{1118} and heterozygous *r2d2* mutant from first crossing step were analyzed in lane 1 and 2, respectively. Actin served as DNA quality control. B) Flies with Cy marker in fly stocks 4 and 6 were tested for heterozygous *r2d2* genotype.

4.2.2 Generation of small RNA libraries

The abovementioned homozygous $loqs^{ko}$ and *r2d2* mutants (see chapter 4.2.1) were further used for generation of small RNA libraries for deep sequencing. To obtain heterozygous control flies, the CyO balanced stocks were crossed to wild w^{1118} and non-Cy F1 animals were collected (Figure 4.16C). Libraries were differentiated in tissue types by utilizing RNA from head and thorax (soma) versus ovaries as representative for germ line (Figure 4.18).

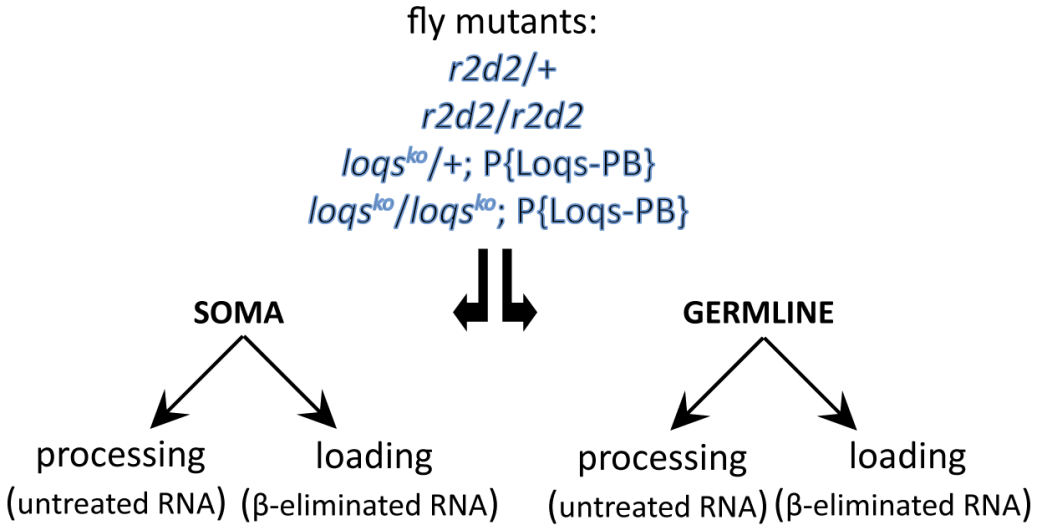


Figure 4.18 Overview of small RNA libraries.

After backcrossing of *loqs*^{ko} mutant rescued with Loqs-PB via P-element insertion as well as *r2d2* mutant into a *w*^{1118} background, homozygous and heterozygous mutants were used for generation of small RNA libraries. The tissue types were differentiated between soma and germline using head and thorax versus ovaries, respectively. In addition, we analyzed processing against loading by using untreated RNA versus β-eliminated RNA, respectively, as start material for generation of small RNA libraries.

In addition, we analyzed the loading state of the small RNAs (Figure 4.18). piRNAs and Ago2-loaded siRNAs bear a 2'-*O* methyl modification at their 3' end which is introduced by DmHen1, a S-adenoxyl-methionine-dependent methyltransferase after the loading into effector complex (Yu, Yang et al. 2005; Horwich, Li et al. 2007; Saito, Sakaguchi et al. 2007). To assess this modification state, RNA was oxidized with sodium periodate, then β-eliminated by a switch of pH into high basic range. Since vicinal diols are required for oxidation, the 2'-*O*-methyl end-modified small RNAs will not be affected by oxidation, while all other species (e.g. Ago1-loaded miRNAs, most RNA degradation products) are oxidized and the last nucleotide is removed due to the β-elimination. Deep sequencing libraries generated with β-eliminated RNA are depleted of shortened small RNAs as they are unable to ligate with the adaptor. β-elimination step was controlled for efficiency by gel-electrophoretic visualization of the 30 nt long 2S rRNA. β-eliminated 2S rRNA was visualized in terms of a shift in mobility by about 1 nt (Figure 4.19).

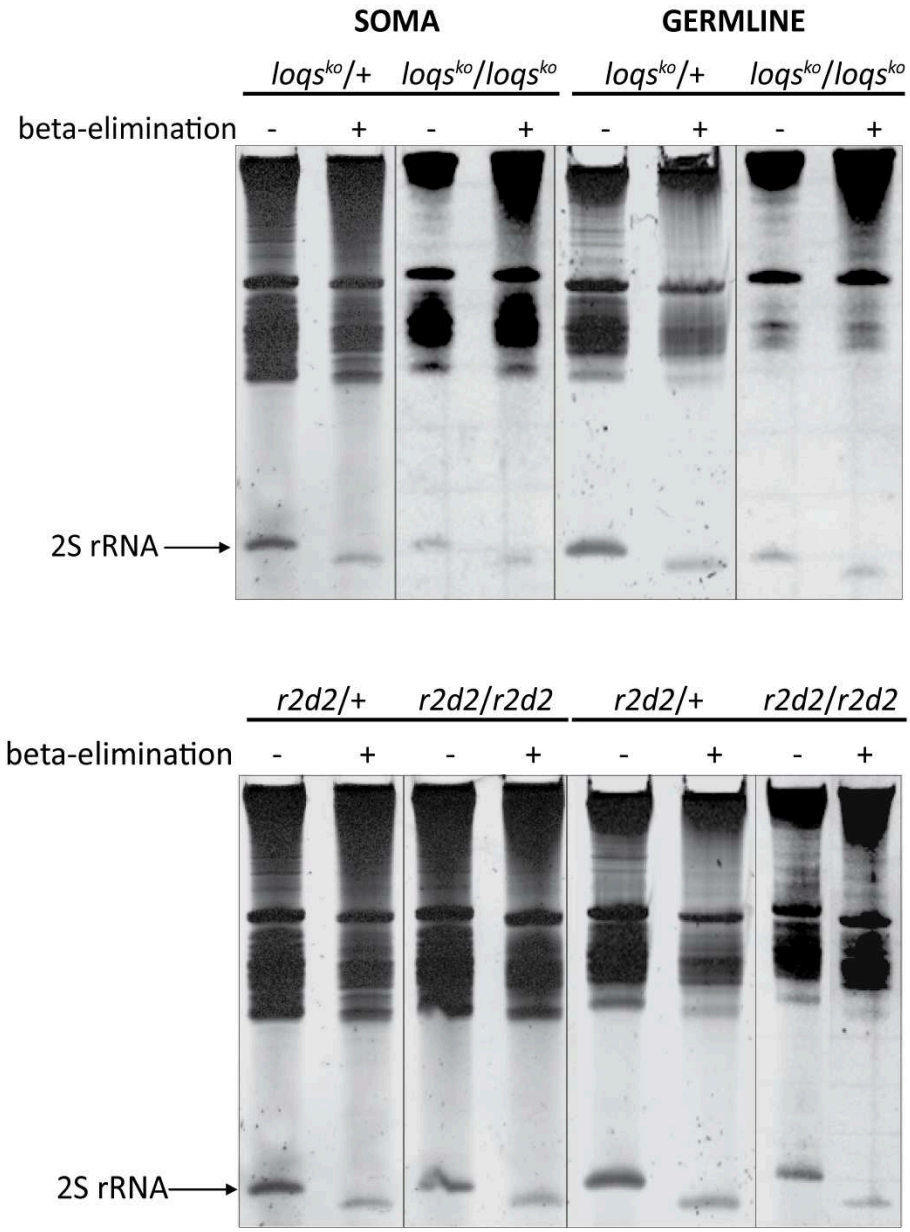


Figure 4.19 Verification of β-elimination efficiency.

RNA was isolated from heterozygous and homozygous *loqs*^{ko} and *r2d2* mutants originating from somatic and germline tissue, respectively. They were oxidized with sodium periodate and β-eliminated by raising pH into high basic range. Each RNA sample, before and after the treatment, was applied on 15% Acrylamide-Urea gel and stained with SybrGold. 2S rRNA served as control for β-elimination efficiency due to the high abundance.

For generation of small RNA libraries total RNA and β-eliminated RNA were gel-purified in size 17 to 30 nt and ligated at the 3' end with a 3' adaptor. The RNA was again gel-purified and ligated at the 5' end with a 5' adaptor. After ligation, RNA was reverse transcribed and amplified by PCR, different sequence bar codes were introduced at this step. The PCR products obtained from ovaries are in general slightly larger in size (Figure 4.20). This may be because germ line RNA contains a large fraction of piRNAs.

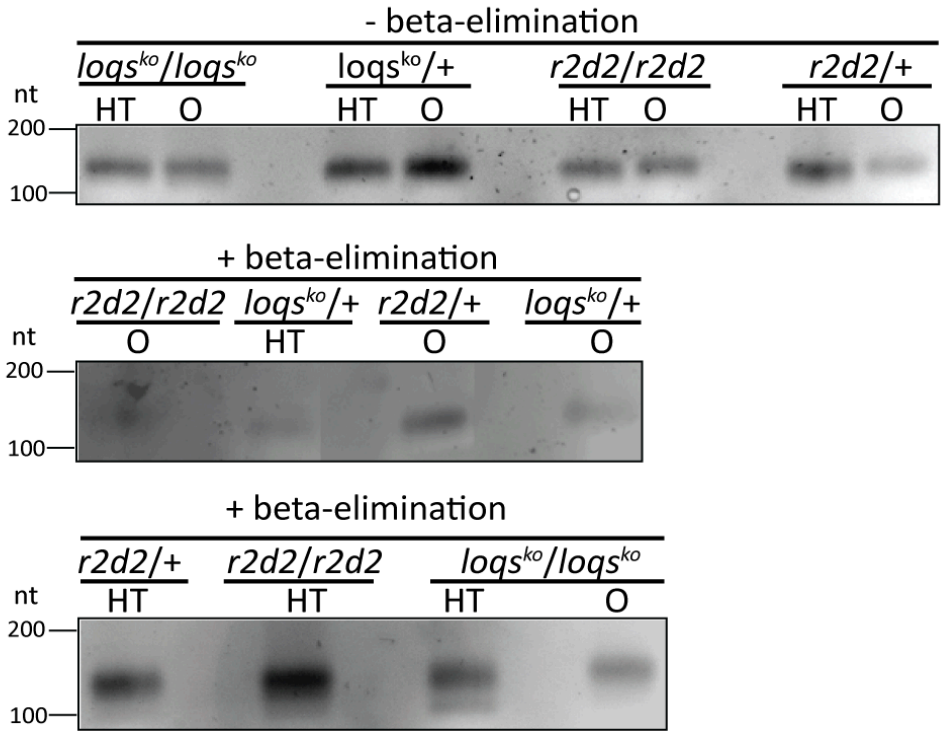


Figure 4.20 Quantification gels of DNA samples for each library.

Small RNA libraries were generated from untreated and β-eliminated RNA originating from heterozygous as well as homozygous *loqs*^{ko} and *r2d2* mutants. Furthermore RNA was isolated from different tissue types, soma (head and thorax, HT) and germline (ovaries, O). The final step of the small RNA library generation protocol is DNA amplification. The resulting product was gel purified by elution within 30 µl volume. Thereof 5 µl were used for quantification via agarose gel electrophoresis which is necessary for equal pooling of samples within one sequencing run.

Due to the advancement of deep sequencing technology and increasing sequencing depth, up to four libraries were pooled into one sequencing run. The bar codes allowed identification of each samples in the bioinformatic analysis after the sequencing procedure. Table 4.4 is an overview presenting the total number of reads, the proportion of reads matching *Drosophila* genome and the amount of reads mapped to transposon or miRNA sequence collections.

The β-eliminated somatic library of homozygous *r2d2* mutant matched with only 30.5% of all reads to the *Drosophila* genome. After permission of one or two mismatches during the mapping procedure, 60% or 71% could be mapped to the *Drosophila* genome, respectively. Furthermore the same library was slightly higher contaminated with *E. coli* matching reads (coming from ΔT4 RNA ligase purification) compared to other libraries but it still can not explain the low mapping percentage. Since all somatic libraries in the same sequencing run (marked with * in Table 4.4) mapped with somewhat lower efficiency to the *Drosophila* genome in comparison to other sequencing runs, we concluded that the low mapping percentage of homozygous *r2d2* mutant library can partly be explained by lower sequencing quality.

SOMA

Library	β-eliminated	Total no of insert 11-28 nt	Inserts matching the genome (% of total)	Inserts matching miRNAs (% of genome matching)	Inserts matching transposons (% of genome matching)
loqs-ko/+	-	9546079	7082781 **(74.2)**	3421132 **(48.3)**	36408 **(0.5)**
loqs-ko/+	+	3682719	2253071 **(61.2)**	464462 **(20.6)**	114366 **(5.1)**
loqs-ko/loqs-ko	-	22704677	18806689 **(82.8)**	10946243 **(58.2)**	80909 **(0.4)**
loqs-ko/loqs-ko*	+	3103232	1348791 **(43.5)**	415769 **(30.8)**	73543 **(5.5)**
r2d2/+	-	20954822	15961906 **(76.2)**	7213032 **(45.2)**	118789 **(0.7)**
r2d2/+*	+	4906737	2444080 **(49.8)**	568163 **(23.2)**	438137 **(17.9)**
r2d2/r2d2	-	4333692	2857262 **(65.9)**	1205092 **(42.2)**	30200 **(1.1)**
r2d2/r2d2*	+	3401343	1038959 **(30.5)**	273550 **(26.3)**	64192 **(6.2)**

GERMLINE

loqs-ko/+	-	14512820	10955066 **(75.5)**	2331242 **(21.3)**	2626560 **(24.0)**
loqs-ko/+	+	14167951	10343590 **(73.0)**	65066 **(0.6)**	4955844 **(47.9)**
loqs-ko/loqs-ko	-	12343141	10107095 **(81.9)**	1288227 **(12.7)**	3342406 **(33.1)**
loqs-ko/loqs-ko*	+	14963584	10910993 **(72.9)**	75443 **(0.7)**	4680467 **(42.9)**
r2d2/+	-	13982564	10624983 **(76.0)**	1805164 **(17.0)**	3076304 **(29.0)**
r2d2/+	+	5385640	4164837 **(77.3)**	25137 **(0.6)**	2012986 **(48.3)**
r2d2/r2d2	-	5715078	3617578 **(63.3)**	548556 **(15.2)**	1125305 **(31.1)**
r2d2/r2d2	+	6797261	5149557 **(75.8)**	15401 **(0.3)**	2244828 **(43.6)**

Table 4.4 Analysis of deep sequencing libraries generated in this study.

Small RNA libraries generated from untreated and β-eliminated RNA from heterozygous as well as homozygous *loqs*^{ko} and *r2d2* mutants were selected for reads of 11-28 nt length. They were further mapped against the reference of *Drosophila* genome and listed as the percentage of the total amount of reads to validate the quality of the libraries. Furthermore all libraries were mapped against the reference of miRNAs and transposons respectively and displayed as the percentage of *Drosophila* genome mapping reads. Libraries marked with * were pooled into one sequencing run.

4.2.3 Are germline piRNAs affected by an impaired endo-siRNA biogenesis?

Retrotransposons are transcriptionally very active in the germline and their efficient repression depends heavily on piRNAs. Nonetheless, transposon-targeting endo-siRNAs are also abundant in the germline. In the absence of a functional endo-siRNA system, it is conceivable that the piRNA pathway needs to adapt in order to ensure maximal repression. Furthermore, endo-siRNAs may directly affect piRNA biogenesis because they degrade transposon sense transcripts and potentially compete with the piRNA system for antisense transcripts, which are required for dsRNA generation as well as for the ping-pong amplification cycle. We therefore asked whether impaired endo-siRNA biogenesis could affect the germline piRNA profile. Since strains from distinct genetic backgrounds most likely differ in the transposon content (both quantitative and where insertions have occurred), we performed one round of backcrossing of the *r2d2* mutant as well as *loqs*^{ko} deletion allele into *w*^{1118} flies (see chapter 4.2.1). While a single round of

backcrossing certainly cannot create a homogenous background, in particular for the mutation-carrying second chromosome, it nonetheless will significantly reduce heterogeneity. Some variability among individuals is unavoidable due to the intrinsic mobility of transposons. Homozygous mutant animals were then compared with heterozygotes obtained by crossing the balanced stocks to w^{1118} flies. To assess loading versus processing, deep sequencing libraries were generated additionally with β-eliminated RNAs to enrich for piRNAs and Ago2-loaded siRNAs. Thus, the contribution of R2D2 and Loqs-PD to the processing of a certain small RNA species can be revealed by reduction in both untreated and β-eliminated libraries, while a functioning only during the loading step is evident by reduction only in the β-eliminated libraries. After mapping of the small RNAs to a collection of transposon sequences, the size distribution of the matching small RNA reads was profiled. We could distinguish peaks at 21 nt and at 24-27 nt, reflecting the presence of endo-siRNAs and piRNAs, respectively (Figure 4.21). Consistent with the published literature, piRNAs were more abundant than endo-siRNAs in the germ line (Vagin, Sigova et al. 2006; Li, Vagin et al. 2009).

A prevailing model is that Loqs-PD acts predominantly during processing of dsRNA by Dcr-2, while the function of R2D2 is to ensure that the siRNAs are loaded into Ago2, rather than Ago1 (Okamura, Robine et al. 2011) although exceptions to such a linear pathway clearly exist (Marques, Kim et al. 2010; Hartig and Forstemann 2011). In our study homozygous *loqs-PD* mutant RNA samples contained a reduced number of 21-mer transposon-targeting endo-siRNAs both before and after β-elimination, indicating that in this case their production by Dcr-2 is diminished (Figure 4.21, right panel). Interestingly, the production of endo-siRNAs was increased in homozygous *r2d2* mutant animals. In addition to genetic background effects, this may indicate a competition of both factors, R2D2 and Loqs-PD for Dcr-2 binding according to the finding that Dcr-2 is known to be limiting for transgenic RNAi (Hartig and Forstemann 2011). Upon β-elimination the 21 nt long transposon targeting endo-siRNAs derived from homozygous *r2d2* mutants were sensitive to β-elimination and declined substantially whereas the 21 nt size peak of RNA from the heterozygous controls remained. This has been attributed to the fact that the siRNAs are loaded into Ago1 in the absence of R2D2 (Okamura, Robine et al. 2011). Taken together, we could corroborate the existing model for the predominant endo-siRNA biogenesis pathway in the germline as well.

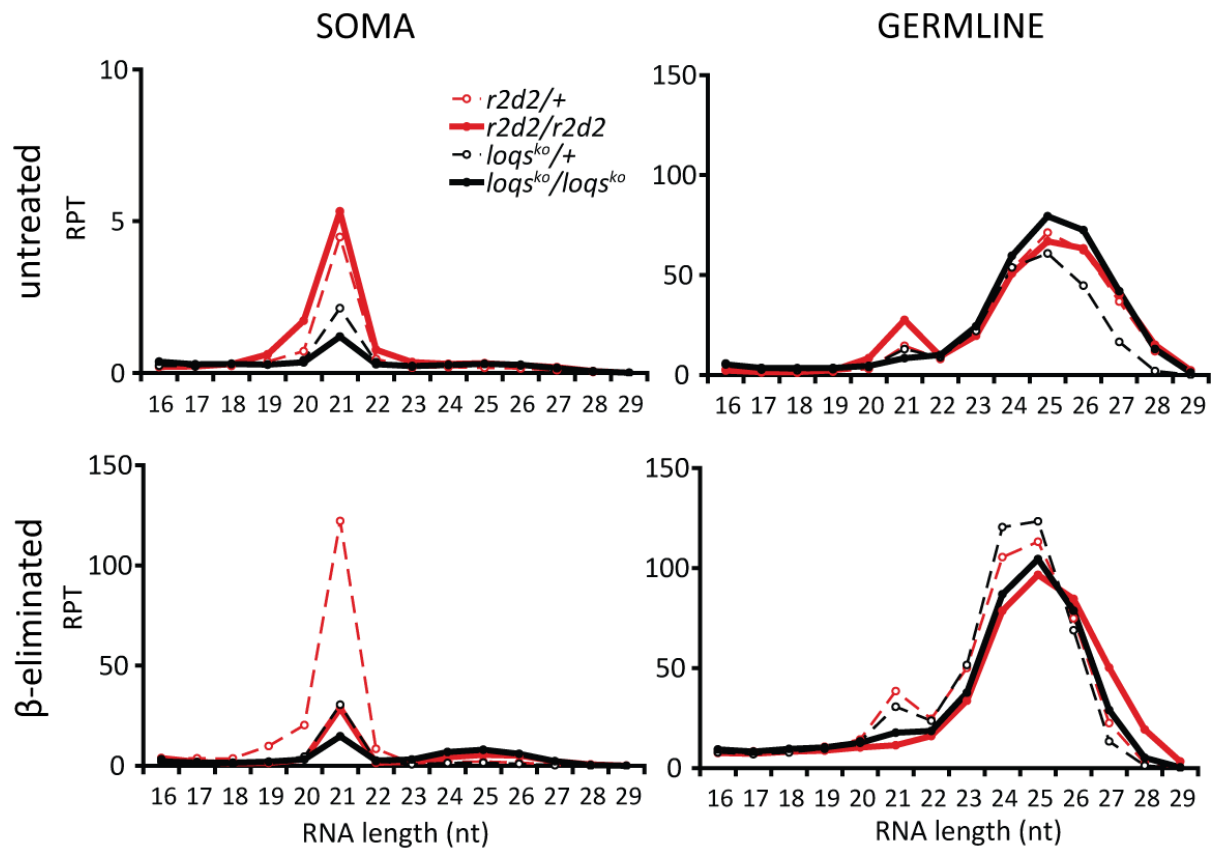


Figure 4.21 The length distribution of transposon matching small RNAs in *r2d2* and *loqs^{ko}* mutants.

Reads of each library originating from soma and germline were mapped to the reference containing a transposon sequence collection. Hereafter the transposon matching small RNAs were analyzed for their size distribution and normalized to the total genome matching reads. The normalized counts were expressed as reads per thousand (RPT).

By focusing on piRNAs in the ovary we noticed that the reads of untreated *loqs^{ko}*/*loqs^{ko}* mutant RNA were shortened by 1 nt in length. The size distribution from all reads of each library indicated that this shift results from technical variability during the library preparation and argues against any biological relevance (Appendix Figure 1). Despite the alterations observed for transposon-targeting endo-siRNAs in ovaries, the piRNA abundance showed no gross changes in abundance or size distribution, both before and after β-elimination, between heterozygous and homozygous mutant animals. We concluded that the biogenesis of germline piRNAs is qualitatively and quantitatively (Appendix Figure 2) unchanged when the endo-siRNA system functions with diminished activity.

4.2.4 Biogenesis of endo-siRNAs in the soma

In parallel to the germline analysis, we generated libraries from the somatic portions of flies to examine the function of R2D2 and Loqs-PD by differentiating between processing and loading. As for the ovary libraries, the reads were mapped to the transposon sequence collection and their size distribution profiled. To allow for quantitative comparisons, the libraries were normalized to the total number of reads matching to *Drosophila* genome with no mismatches allowed. A striking observation was that a large proportion of reads (0.6% to 14.5% of genome matching reads, 5.1% to 80.6% of transposons matching reads) could be attributed to only four transposable elements (*roo*, *297*, *TNFB* and *blood*) (Table 4.4). They were detected among the ten most actively siRNA generating transposons in all somatic libraries (Appendix Table 1). Diagrams that depicted the normalized length distribution for each one of these transposable elements individually are included in the appendix (Figure 3). The amount of endo-siRNAs against *roo*, *297*, *TNFB* and *blood* in libraries was disproportionate with respect to their steady state transcript levels in comparison to other transposons, indicating particularly efficient targeting by the endo-siRNA system (Figure 4.22).

SOMA		*r2d2/+*	*r2d2/r2d2*	*loqs^{ko}/+*	*loqs^{ko}/loqs^{ko}*
16-29nt	**untreated***				
	297	6.01	12.45	9.93	6.03
	TNFB	45.12	34.98	4.55	1.82
	roo	19.59	33.27	32.06	37.36
	blood	1.40	0.88	1.77	0.83
	sum	72.11	81.58	48.32	46.04
	β-eliminated*				
	297	7.52	2.55	15.26	5.85
	TNFB	53.41	7.56	6.94	2.02
	roo	17.92	5.80	29.99	15.73
	blood	1.75	37.19	2.33	1.19
	sum	80.61	53.10	54.52	24.79
	*normalized to transposons mapping reads [%]				
16-29nt	**untreated****				
	297	0.04	0.13	0.05	0.03
	TNFB	0.34	0.37	0.02	0.01
	roo	0.15	0.35	0.16	0.16
	blood	0.01	0.01	0.01	0.00
	sum	0.54	0.86	0.25	0.20
	β-eliminated**				
	297	1.35	0.16	0.77	0.32
	TNFB	9.57	0.47	0.35	0.11
	roo	3.21	0.36	1.52	0.86
	blood	0.31	2.30	0.12	0.06
	sum	14.45	3.28	2.77	1.35
	**normalized to genome mapping reads [%]				

GERMLINE		*r2d2/+*	*r2d2/r2d2*	*loqs^{ko}/+*	*loqs^{ko}/loqs^{ko}*
16-29nt	**untreated***				
	297	3.13	3.17	3.98	1.66
	TNFB	1.18	2.97	0.10	0.01
	roo	5.57	3.76	5.30	2.64
	blood	2.76	1.60	1.69	0.81
	sum	12.65	11.50	11.06	5.13
	β-eliminated*				
	297	2.95	2.97	2.96	1.70
	TNFB	1.54	0.06	0.06	0.01
	roo	3.58	4.49	3.14	2.48
	blood	2.05	1.76	1.46	0.89
	sum	10.13	9.28	7.63	5.08
16-29nt	**untreated****				
	297	0.91	0.99	0.95	0.55
	TNFB	0.34	0.92	0.02	0.00
	roo	1.61	1.17	1.27	0.87
	blood	0.80	0.50	0.40	0.27
	sum	3.66	3.58	2.65	1.70
	β-eliminated**				
	297	1.43	1.30	1.42	0.73
	TNFB	0.75	0.02	0.03	0.00
	roo	1.73	1.96	1.51	1.06
	blood	0.99	0.77	0.70	0.38
	sum	4.90	4.04	3.66	2.18

Table 4.5 The counts of *297*, *TNFB*, *roo* and *blood* mapping small RNAs.

All libraries with 16 nt to 29 nt long reads were mapped separately to *roo*, *297*, *TNFB* and *blood* transposon sequences. In the upper panel the counts of *roo*, *297*, *TNFB* and *blood* matching small RNAs were normalized to all transposons mapping reads and expressed as percentage. In the lower panel the counts were normalized to the genome matching reads and expressed as percentage. The percentages were summarized for all four transposons in each library to demonstrate which proportion of transposons- or genome- mapping reads was occupied by *roo*, *297*, *TNFB* and *blood* mapping small RNAs.

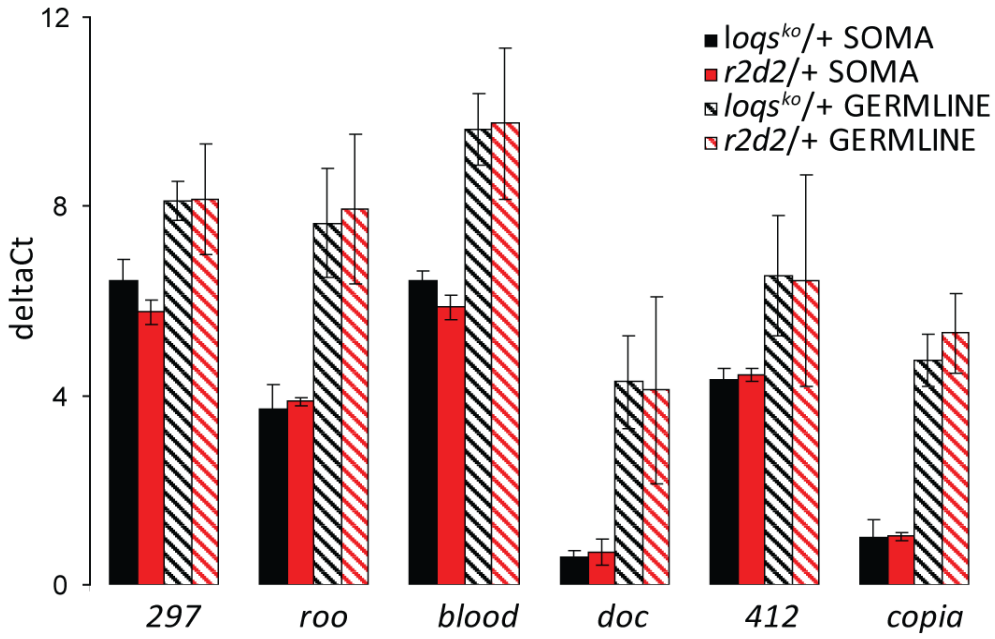


Figure 4.22 The steady state transcript levels of *297*, *TNFB*, *roo* and *blood* transposable elements.

RNA was isolated from three biological replicates of heterozygous *loqs*^{ko} and *r2d2* mutants separated in somatic and germline tissue, respectively. After that RNA was reverse transcribed and analyzed by qRT-PCR for mRNA levels of *297*, *roo*, *blood*, *doc*, *412* and *copia* transposons. Ct-values for each transposon were normalized to the rp49 control (delta Ct). Values are mean ± SD (n=3).

To allow for a more diversified representation of many distinct transposons, we present the remainder of the results with the reads matching those four mobile elements filtered out. We did not detect any differences of potential biological relevance between this group (*roo*, *297*, *TNFB* and *blood*) and other transposons (Appendix Figure 3). Before β-elimination, homozygous loss of Loqs-PD resulted in a 1.8-fold reduction of transposon-matching endo-siRNAs, consistent with the notion that its role is predominantly in siRNA production (Figure 4.23, left panel). While this was true for the analysis of all transposons in bulk, individual exceptions to this rule exist. For example, the transposons *F-element*, *412* and *Doc* were only slightly affected by loss of Loqs-PD (Figure 4.24). This general trend was also true after β-elimination, in agreement with the notion that small RNAs which are not produced also cannot be loaded. As expected, this situation was different for the *r2d2* mutation. In overall terms, no reduction of transposon matching endo-siRNAs could be detected when comparing heterozygous to homozygous mutants before β-elimination (Figure 4.23, left panel). But again, some exceptions exist. The production of endo-siRNA mapping against *F-element* was primarily affected by depletion of R2D2. However, after β-elimination even more of the 21-mer transposon matching small RNAs were eliminated. Our results are therefore in agreement with the published hypothesis that in the absence of R2D2, many endo-siRNAs are loaded into Ago1. Yet, a small fraction did not obey to this rule and remained loaded after loss of R2D2.

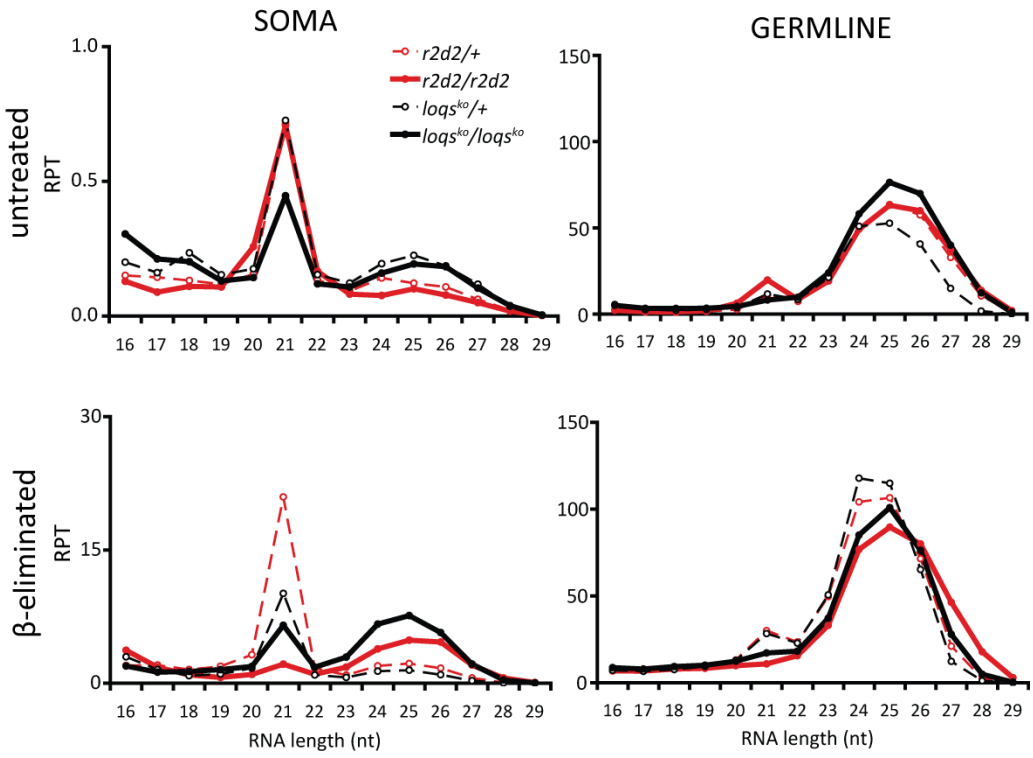
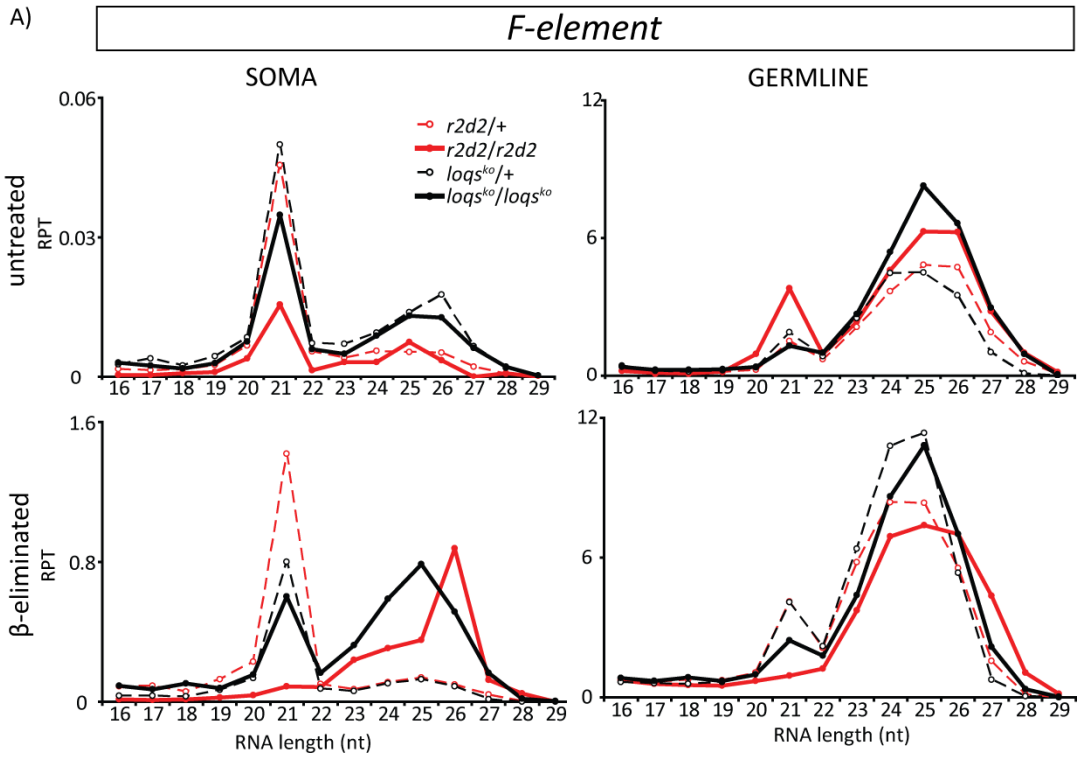


Figure 4.23 The length distribution of transposon matching small RNAs in *r2d2* and *loqs*^{ko} mutants after exclusion of *roo*, *297*, *TNFB* and *blood* transposons.

Reads of each library were mapped to the transposon sequence collection excluding *roo*, *297*, *TNFB* and *blood*. The transposon matching small RNAs were analyzed for their size distribution and normalized to the total genome matching reads. The normalized counts were expressed as reads per thousand (RPT).



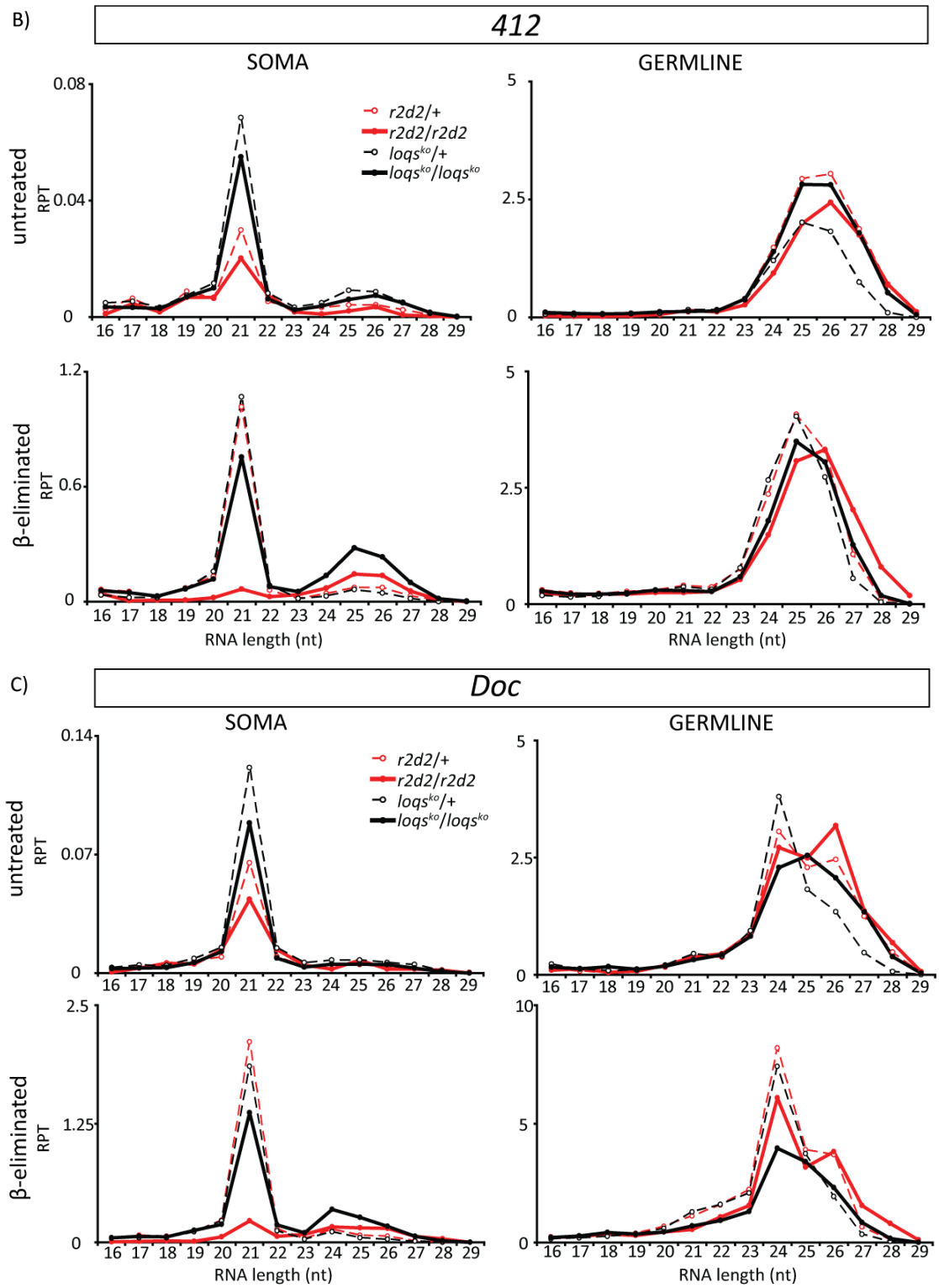


Figure 4.24 The length distribution of *412*, *F-element*, *doc* transposon mapping small RNAs in *r2d2* and *loqs*^{ko} mutants.

Reads of each library were mapped to A) *F-element*, B) *412* and C) *doc* transposon sequence separately and their size distribution was profiled. After normalization to total genome matching reads, the size distribution was expressed as reads per thousand (RPT).

Potentially, the small RNAs we sequenced in the homozygous *r2d2* mutant library after β-elimination might just represent Dicer products that somehow escaped the chemical treatment but remained unloaded as duplex intermediates. To rule out this possibility, we computed the ΔG values of the putative dsRNA precursors for Ago2 loading. To this end, we calculated the free energy of base pairing across the first 5 nt of each read at either end of the presumed duplex siRNA, then determined the difference between the 5' and 3' end of each presumed endo-siRNA precursor. We found a consistent excess of base-pairing stability of the 3'-end over the 5'-end of the small RNAs, indicating that they have been subject to a comparable extent of strand selection that follows the established rules according to base pairing stability (Figure 4.25). Thus, RLC action and strand selection can occur in certain instances in the absence of *r2d2*, strongly suggesting that Loqs-PD can substitute for the function of R2D2 in RLC.

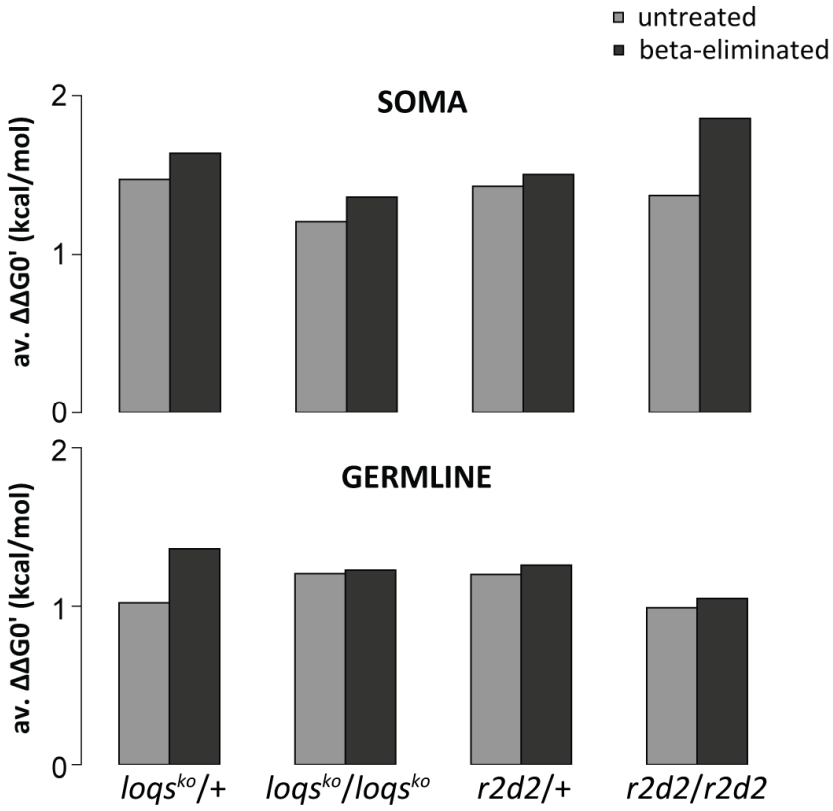


Figure 4.25 Analysis of strand asymmetry in deep sequencing data.

The thermodynamic asymmetry was calculated for transposon mapping endo-siRNAs of the indicated genotypes. We calculated the difference in free energy of duplex formation at either end of the presumed siRNA precursor for each sequence read using the nearest neighbor method (Xia, SantaLucia et al. 1998), then calculated the difference (ΔΔG0'). A positive value indicated that on average the 5' ends of the reads were less stably base paired than the opposite ends.

Is there any common principle that could explain why certain transposons differ from the bulk in their requirements for Loqs-PD and R2D2? Scatter plots depict the abundance of 21 nt long endo-siRNAs against individual transposons which were normalized to the total genome matching reads (Figure 4.26). Within the scatter plot it is possible to compare the normalized amounts of endo-siRNAs for each transposon between both mutants. As already described above, Loqs-PD was important for production of many endo-siRNAs. Not all transposons are equally affected, indicating that specificity is observed. This distinction is not based on their abundance since dependence on Loqs-PD for biogenesis or R2D2 for Ago2-loading does not correlate with the total amount of small RNAs (Figure 4.26).

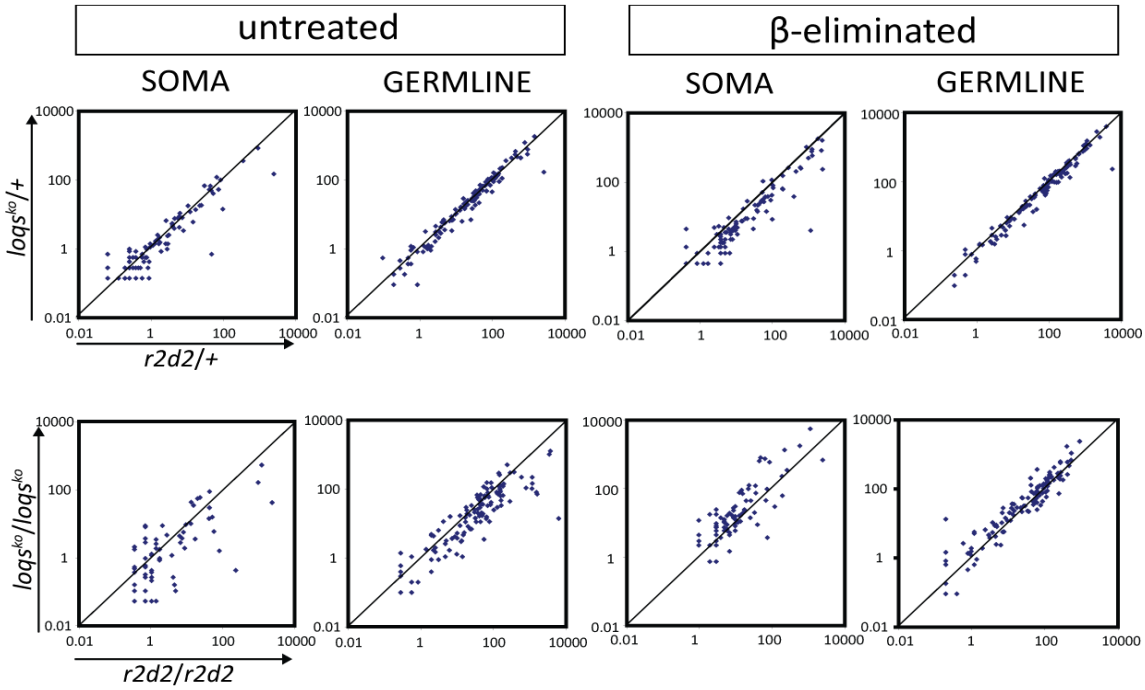
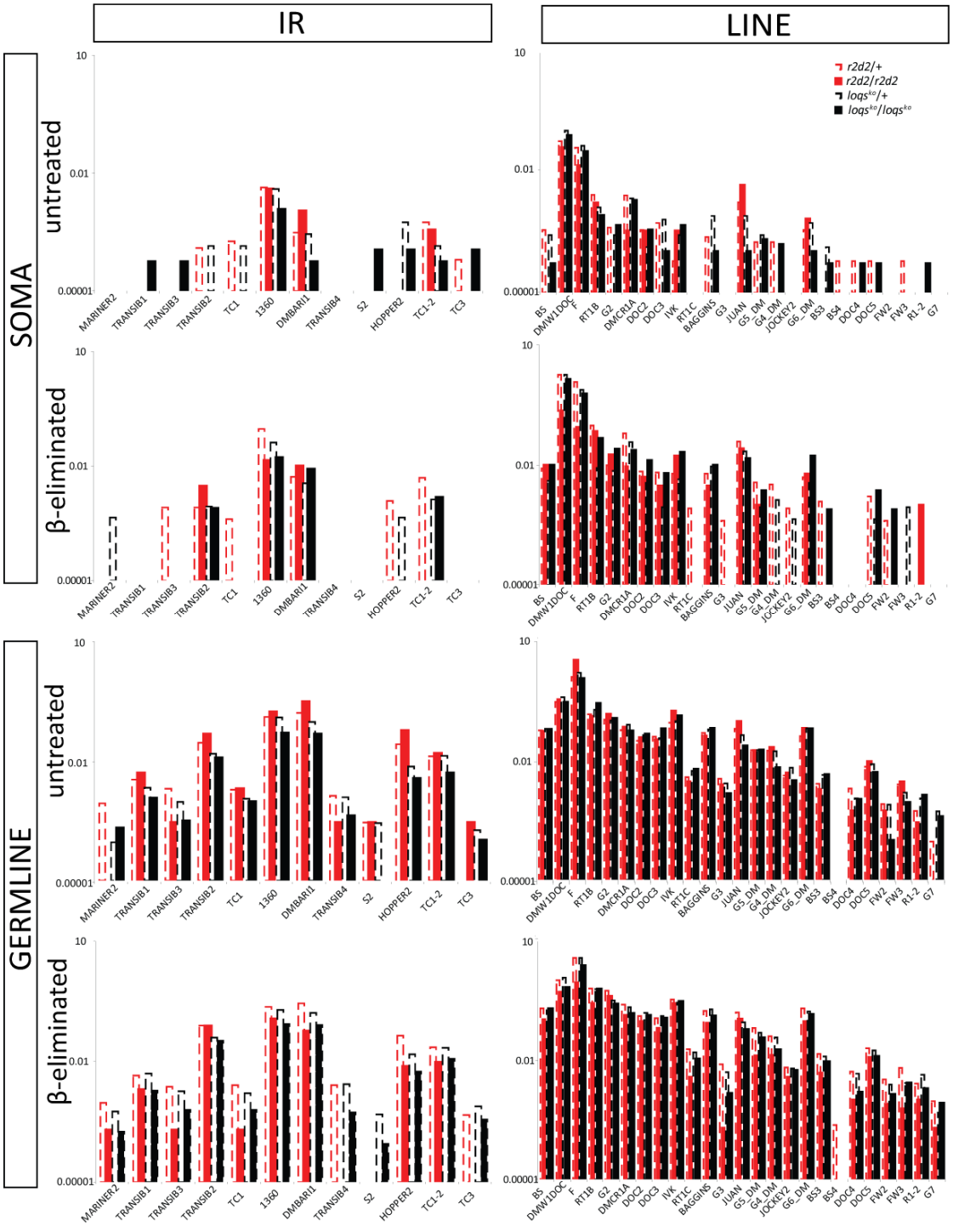


Figure 4.26 Changes in processing and loading of small RNAs matching to individual transposons in *r2d2* and *loqs*^{ko} mutants.

Transposon mapping endo-siRNA were normalized to total genome matching reads and expressed as reads per million (RPM). The upper panels of the scatter plots represent the comparison of heterozygous *r2d2* and *loqs*^{ko} mutants during the processing (left) and loading (right) for soma and germline. The lower panels compare homozygous *r2d2* with homozygous *loqs*^{ko} mutants. A higher amount of endo-siRNAs in *r2d2* homozygous mutant means that these endo-siRNAs are *r2d2* independent but *loqs* dependent and vice versa. Thus, classification of individual transposons is possible according to their dependence on Loqs-PD or R2D2.

Furthermore, when transposons were classified according to their general characteristics into long terminal repeats (LTRs), long interspersed elements (LINEs) and inverted repeats (IRs), we did not observe any particularity among transposons that could explain R2D2 versus Loqs-PD dependence (Figure 4.27).


IR
LINE
SOMA
GERMLINE
untreated
β-eliminated
r2d2/+
r2d2/r2d2
loqs^ko/+
loqs^ko/loqs^ko
10
0.01
0.00001
MARINER2
TRANSIB1
TRANSIB3
TRANSIB2
TC1
1360
DMBARI1
TRANSIB4
S2
HOPPER2
TC1-2
TC3
BS
DMW1DOC
F
RT1B
G2
DMCR1A
DOC2
DOC3
IVK
RT1C
BAGGINS
G3
JUAN
G5_DM
G4_DM
JOCKEY2
G6_DM
BS3
BS4
DOC4
DOC5
FW2
FW3
R1-2
G7

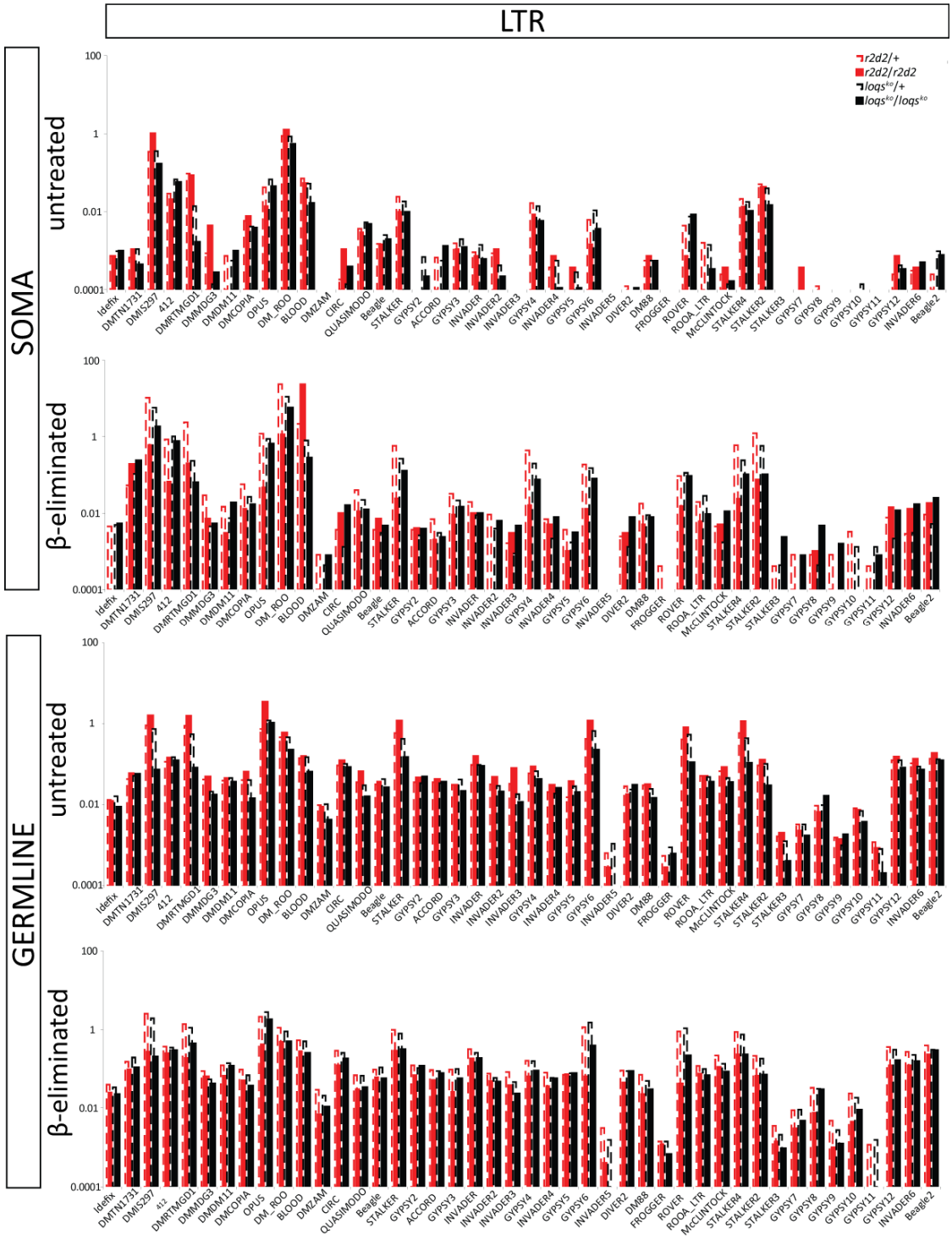


Figure 4.27 Analysis of endo-siRNAs classified in LTRs, LINEs and IRs transposons in *r2d2* and *loqs*^{ko} mutants.

The transposon sequence collection was classified into retrotransposons comprising of long terminal repeats (LTRs) and long interspersed elements (LINEs) while DNA transposons were represented by inverted repeats (IRs). Reads of each library were mapped to individual transposons and transposon matching endo-siRNAs were normalized to total genome matching reads. The effects of *r2d2* and *loqs*^{ko} mutants were examined during processing and loading within soma and germline.

As we could find no criterion intrinsic to the transposons, we analyzed whether a particular genomic origin of the reads could explain R2D2 versus Loqs-PD dependence. To this end, we mapped deep sequencing reads to a collection of transposon containing genomic clusters, allowing only those reads that mapped uniquely among these clusters (Brennecke, Aravin et al. 2007). Interestingly, one particular cluster on chromosome X (referred as cluster 2) generated a particularly high number of endo-siRNAs, which we detected before and after β-elimination. Presumably, this reflects active bi-directional transcription of this cluster in somatic cells. After β-elimination we noticed an increased endo-siRNA amount in soma in the absence of *r2d2* in contrast to other clusters (Figure 4.28). This difference was due to a unique sequence with 23275 counts at one location. We consider this sequence to be a technical artifact (e.g. particularly high ligation efficiency) and removed it from the analyzed data set (Figure 4.28 marked with **) resulting in a consistent decrease of endo-siRNAs upon mutation of *r2d2* for all clusters.

A detailed look at cluster 2 in soma demonstrated that endo-siRNAs were generated from defined regions while piRNAs were mapped along the entire cluster reference (data not shown) according to published literature (Li, Vagin et al. 2009; Malone, Brennecke et al. 2009). Furthermore sense and antisense endo-siRNAs were equally represented confirming their origin from a double stranded precursor, while piRNAs showed an orientation bias as expected. Interestingly, regions generating endo-siRNAs differed between soma and ovary. All in all, we saw no correlation between the site of genomic origin and dependence on Loqs-PD and R2D2 during either processing of the dsRNA precursor or loading of siRNA into Ago2.

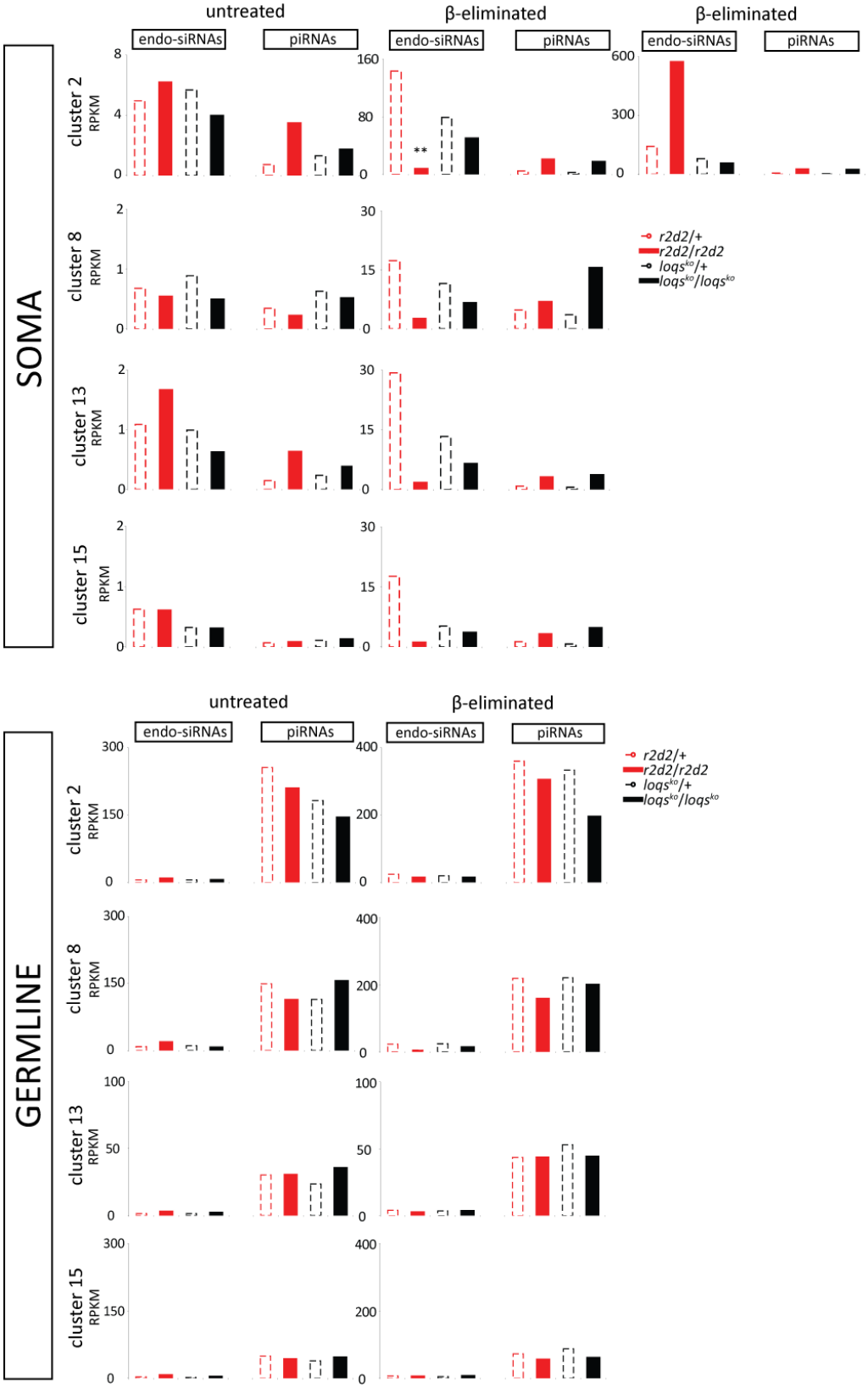


Figure 4.28 Analysis of endo-siRNAs and piRNAs regarding to transposon master loci in *Drosophila* genome (legend continued on p. 92).

(legend Figure 4.28 continued)
Fifteen transposons containing genomic regions were reported as master regulators of transposon activity (Brennecke, Aravin et al. 2007). Reads of all libraries separated in endo-siRNA (21 nt) and piRNAs (24-27 nt) were mapped allowing only those reads that mapped uniquely among these clusters. The counts were normalized to the cluster length as well as to total genome matching reads (reads per kilobase per million mapped reads, RPKM). Cluster 2 (chromosome X; 20A), 8 (chromosome X; 20 A-B), 13 (chromosome 3LHet) and 15 (chromosome 3LHet) are shown. In the soma cluster 2 showed an enormous amount of one unique sequence at a particular location in the homozygous mutant r2d2 sample after β-elimination. We depict the results after exclusion of the special sequence and marked with **.

Do the differences in endo-siRNA abundance in response to Loqs-PD or R2D2 deficiency lead to changes in the steady state level of transposons? We analyzed RNA isolated from soma and ovaries of heterozygous and homozygous flies, then determined the transcript levels of 22 distinct transposons by quantitative RT-PCR. The difference between homozygous and heterozygous mutants is presented as fold change in expression. In the somatic sample, loss of R2D2 resulted in a significant derepression of the transposons *mdg1*, *gypsy*, *297*, *roo*, *juan*, *idefix* and *412* (t-test, $p \leq 0{,}05$). Loss of Loqs-PD, in contrast, only resulted in derepression of *412*, *roo*, *INE-1* and *nof* (t-test, $p \leq 0{,}05$). Apparently, the redirection of endo-siRNAs into Ago1 upon loss of R2D2 represents a more severe loss of function than the reduced endo-siRNA biogenesis upon loss of Loqs-PD. In the ovarian RNA samples we detected significant changes only for *mdg1*, *het-A* and *F-element* upon loss of R2D2.

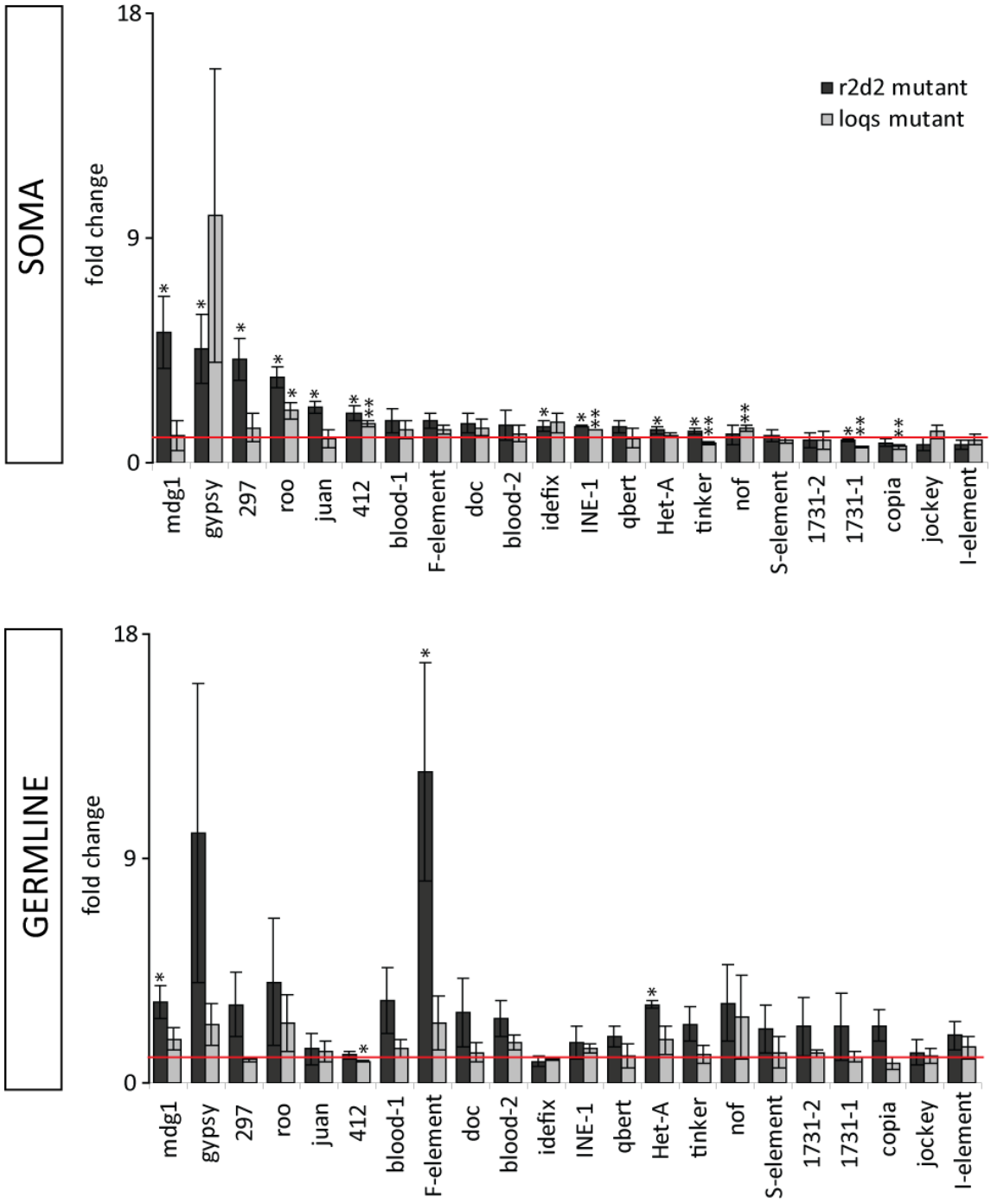


Figure 4.29 Analysis of steady state level of transposons.

RNA was isolated from heterozygous and homozygous *r2d2* and *loqs*^{ko} mutants. DNA was digested with DNase I, the RNA was reverse transcribed and used for transposon profiling by qRT-PCR. Each transposon was normalized to the average of rp49 and gapdh controls and depicted as the fold change of homozygous to heterozygous mutant in soma and germ line, respectively ($p \leq 0.05$(*),$p \leq 0.009$(**)student's t-test, n=3).

We then examined the correlation between the effects of *r2d2* or *loqs*^{ko} mutants on endo-siRNAs and the steady state levels of appropriate transposons. In scatter plots comparing the fold change of mRNA levels against the fold change in siRNAs levels from homozygous versus heterozygous *loqs*^{ko} mutants, the majority of transposons show a reduced amount of siRNAs but nevertheless transposon mRNA levels were hardly increased. This indicates that in the wild type context more siRNAs are generated than required to maintain repression against transposable elements.

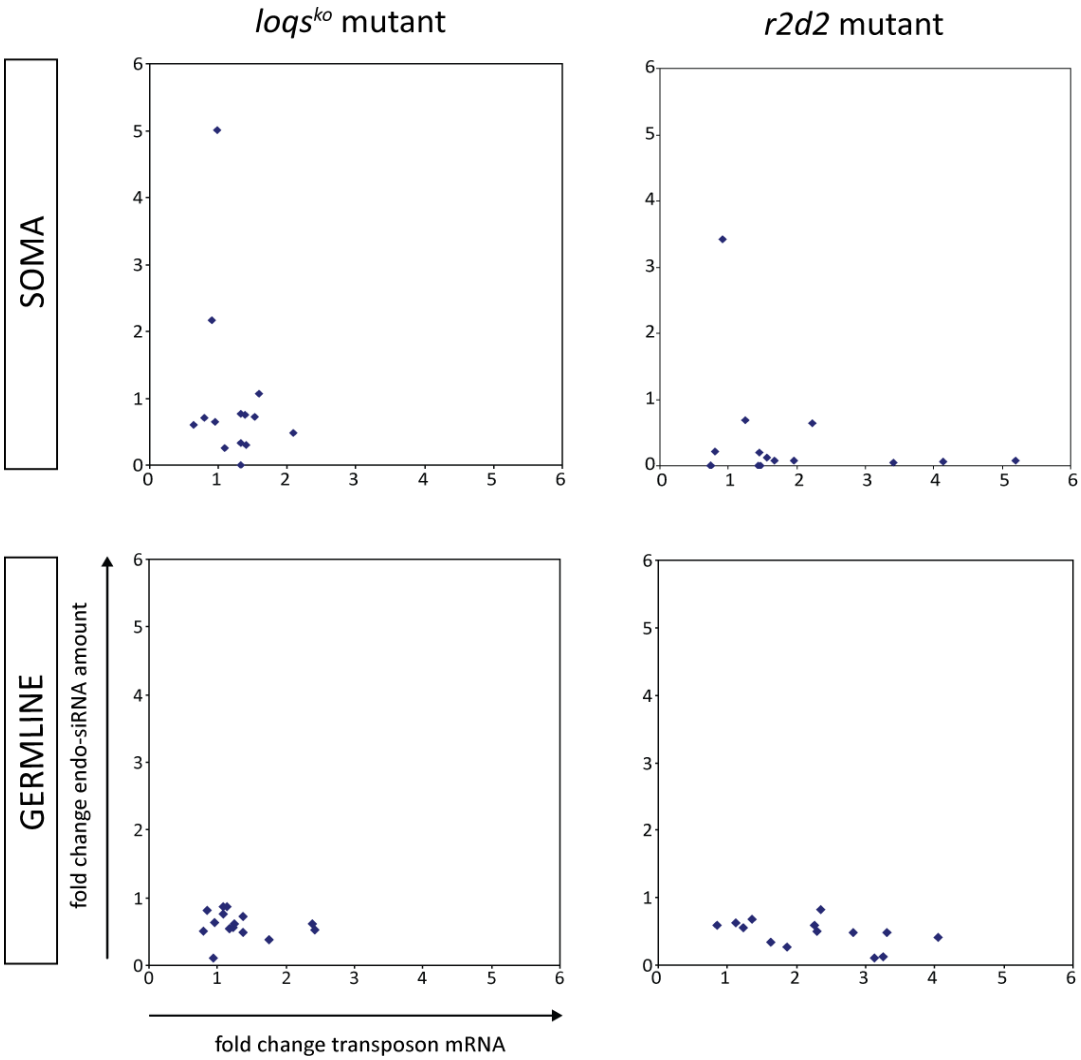


Figure 4.30 Comparison of steady state levels of transposons with loading of endo-siRNAs.

Scatter plots were performed to test the correlation of the change in steady state level of transposon mRNA against the change in loading of endo-siRNAs for specific transposon after deletion of *r2d2* or *loqs-PD*. To do so, we compared the fold change of homozygous to heterozygous mutants of deep sequencing data during the loading against qRT-PCR results in soma and germline. Deep sequencing data were normalized to the total genome matching reads while qRT-PCR values were normalized to the average of rp49 and gapdh controls.

4.2.5 Occurrence of somatic piRNAs

Upon mapping to the transposon consensus sequences, somatic small RNA libraries also produced matches in the size range of 24 to 27 nt. Such piRNA-sized species have been previously described in the heads of *ago2*^{414} mutant flies (Ghildiyal, Seitz et al. 2008) and also in mouse and rhesus macaque samples (Yan, Hu et al. 2011) and referred to as piRNA-like small RNAs (pilRNAs). If these small RNA species are loaded into either Ago2 or a PIWI-family effector protein, then they should be 3' end modified and resistant to β-elimination. Indeed, we found that transposon targeting 24 to 27 nt long RNAs were enriched in the β-eliminated small RNA libraries (Figure 4.23). Like germline piRNAs, they showed a particular orientation bias (Figure 4.31), which argues against dicer-dependent processing of their precursors. We refer to these somatic piRNA-like small RNAs as pilRNAs in the remainder of this study. Does a mutation in components of the endo-siRNA pathway could influence pilRNAs? Neither Loqs-PD nor R2D2 were involved in pilRNA production or loading (Figure 4.23). For further characterization of pilRNAs, we generated sequence logos of 24-27 nt long reads separately for sense and antisense matches. A strong preference for a 5'-U in the antisense matching reads could be seen whereas sense-matching reads in this size-range showed a preference for U at the first position and an A at 10th position (Figure 4.32). These features are obvious in all samples after β-elimination and characteristic of biogenesis via the ping-pong mechanism (Brennecke, Aravin et al. 2007; Gunawardane, Saito et al. 2007; Lau, Robine et al. 2009; Malone, Brennecke et al. 2009; Robine, Lau et al. 2009; Saito, Inagaki et al. 2009). In the case of germline piRNAs, the preference for A at position 10 of sense piRNAs can also be seen in the untreated sample, while the somatic samples likely contain transposon mRNA degradation products that mask this feature in the untreated libraries.

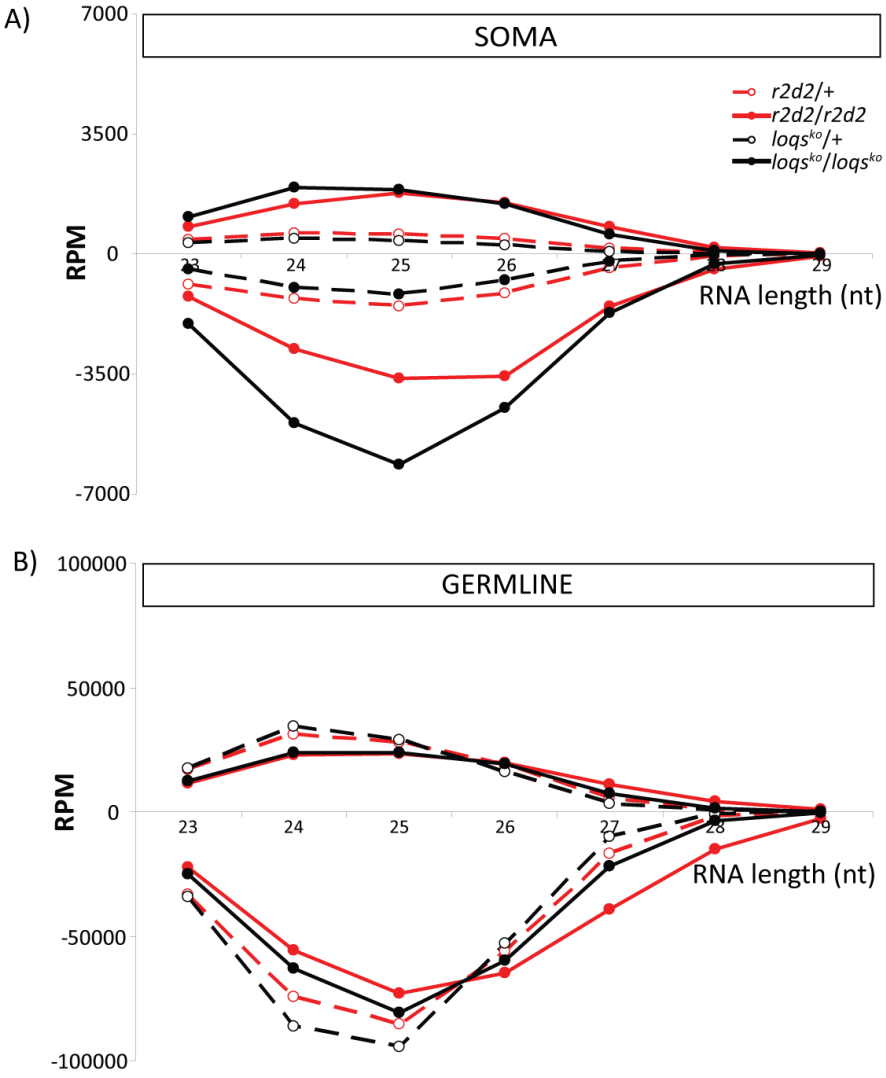
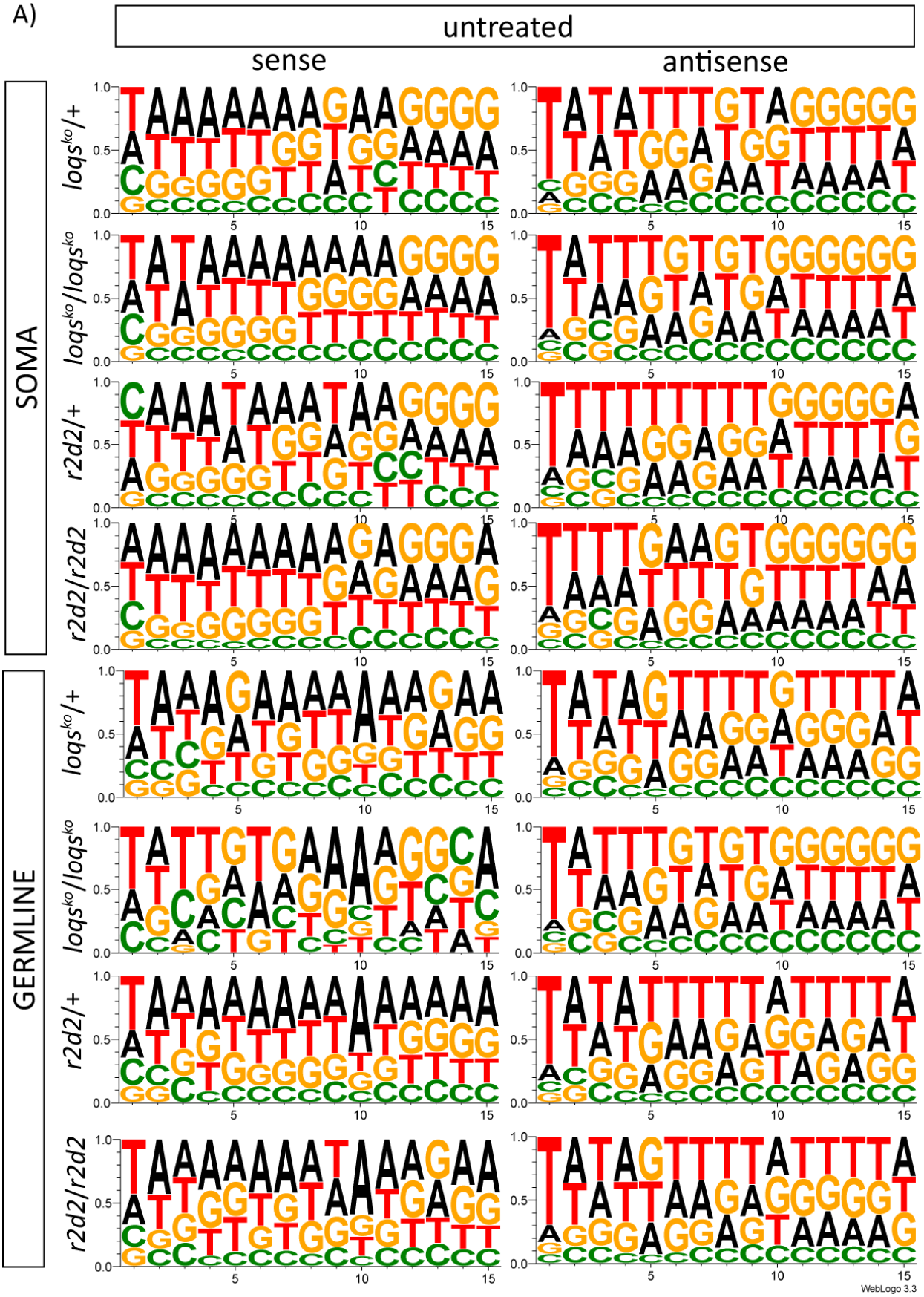


Figure 4.31 Orientation bias for pilRNAs in soma and piRNAs in germline.

Small RNA libraries generated with β-eliminated RNA samples were mapped to the reference of transposons sequence collection. The RPM for sense (+) and antisense (-) transposon matching small RNAs for 23 nt to 29 nt were depicted for soma (A) and germline (B) to demonstrate the orientation bias.


A)
untreated
sense
antisense
SOMA
GERMLINE
loqs^ko/+
loqs^ko/loqs^ko
r2d2/+
r2d2/r2d2
WebLogo 3.3

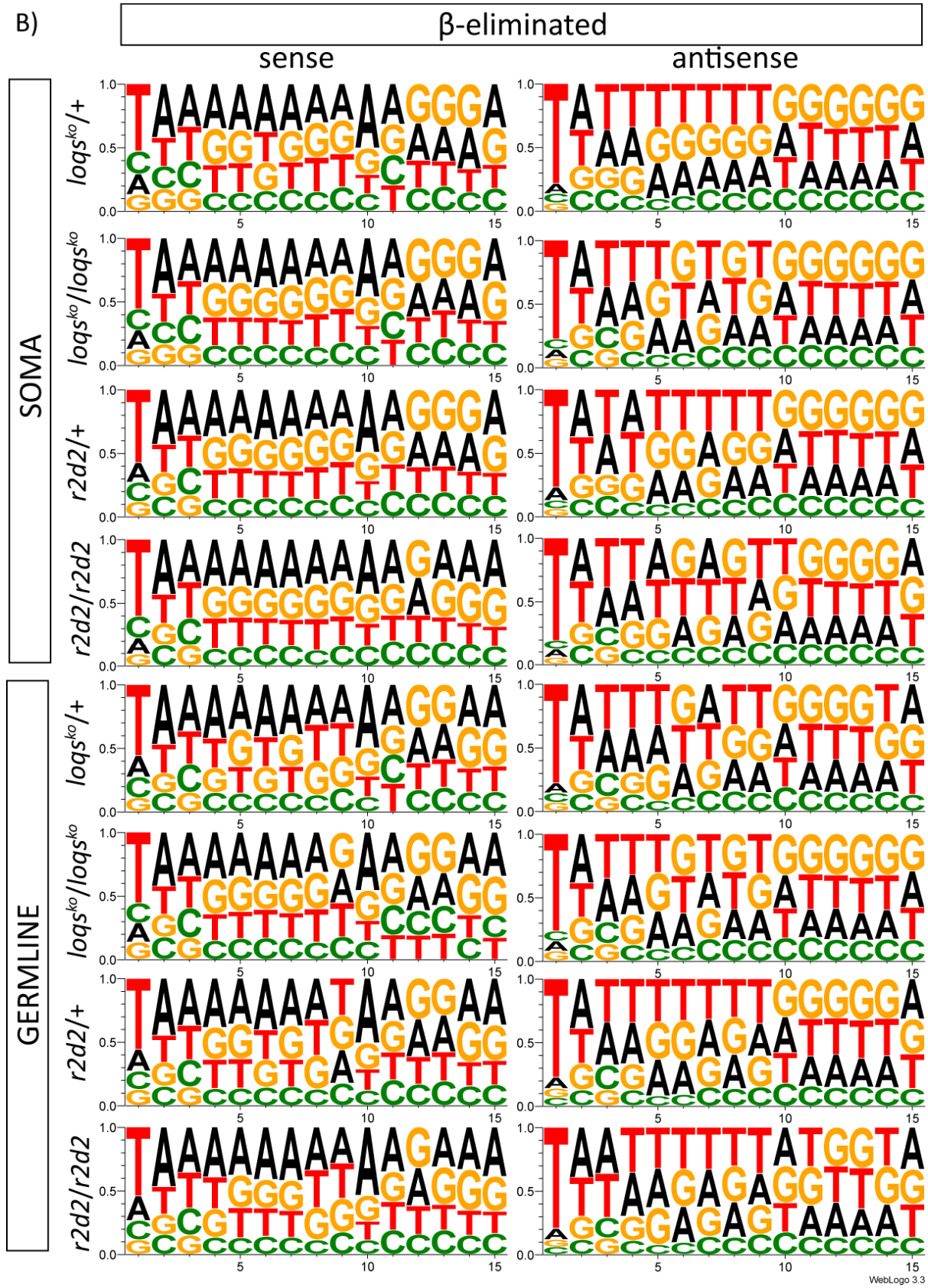


Figure 4.32 Analysis of ping-pong signature of pilRNAs and piRNAs (legend continued on p. 99).

(legend Figure 4.32 continued) Sequence logo plots of 24-27 nt long sense (+) and antisense (-) transposon matching RNAs were constructed with WebLogo 3.3. The height of symbols within the stack indicated the frequency of each nucleotide A, C, U or G (U is substituted by T) at this position for the first 15 nt in all analyzed reads. RNAs were oriented with their 5' end to the left. Untreated RNA samples in soma and germline were depicted in (A) while (B) demonstrated the β-eliminated samples after loading. Only the first 15 positions of the small RNAs are shown.

The biogenesis of piRNAs is based on Piwi-family proteins with Ago3 as the predominant carrier of sense piRNAs while Piwi and Aub bind antisense piRNAs. If somatic piRNAs exist, then Piwi-family proteins should be expressed in the soma as well. We tested for their presence by RT-PCR and only found expression close to background levels, whereas the transcripts were readily detectable in the germ line (Figure 4.33). This could either imply that most somatic cells express very low levels of Piwi-family genes, or that a small subset of cells in adult flies is proficient for the piRNA pathway. In the first scenario, a homogeneous somatic cell population should show an amount of piRNA-sized transposon-matching reads comparable to our somatic fly libraries. We analyzed published small RNA sequencing libraries from the somatic S2 cell line of embryonic origin but found no indication that pilRNA reads were present (Hartig and Forstemann 2011). The most likely explanation for the origin of somatic piRNAs is therefore that a small subset of cells with an active piRNA pathway including the ping-pong mechanism exists in the soma of flies.

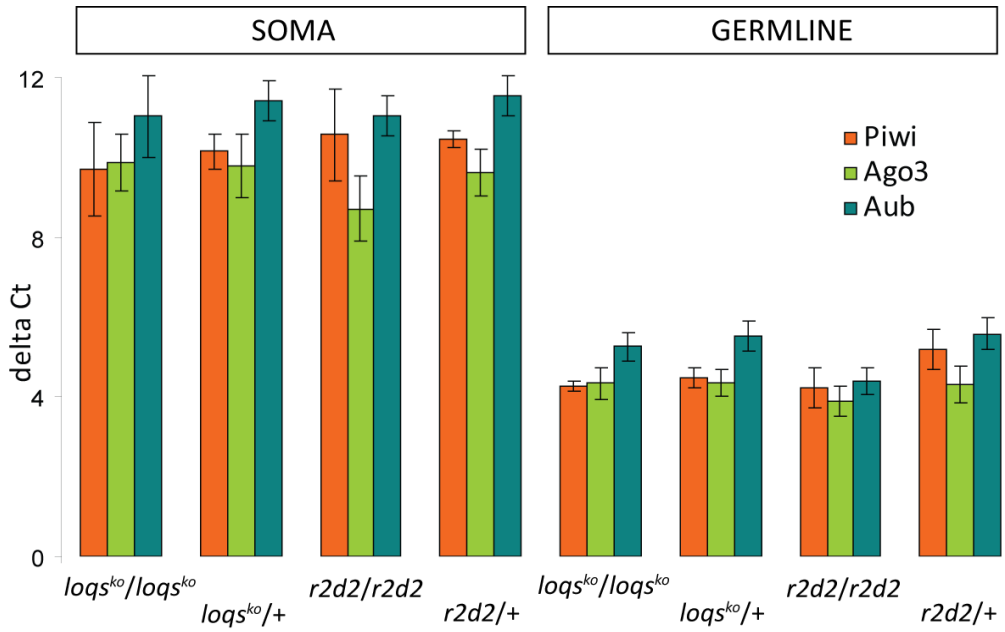


Figure 4.33 Transcript levels of *ago3*, *aub* and *piwi*.

Somatic RNA isolated from *r2d2* and *loqs*^{ko} mutants was analyzed for expression levels of *ago3*, *aub* and *piwi* by qRT-PCR. In addition germ line samples known for high abundance of Ago3, Aub and Piwi served as control. RNA was isolated from soma and germ line of heterozygous and homozygous *r2d2* and *loqs*^{ko} mutants. DNA degradation was performed by DNase I and RNA was reverse transcribed. All Ct values were normalized to the rp49 control (delta Ct). Values are mean ± SD (n=3).

While pilRNAs were enriched after β-elimination in soma, endo-siRNAs were reduced in *r2d2* and *loqs*^{ko} mutant. This genetic analysis demonstrates that both small RNA species depend on different biogenesis factors. If a cross talk exist between pilRNA and endo-siRNA pathway in somatic cells, *dcr-2*, *loqs* and *r2d2* mutants are expected to induce production of pilRNAs. However, analysis of published libraries from somatic S2 cells depleted for *dcr-2*, *loqs* and *r2d2* showed no pilRNAs and argued against any correlation between the endo-siRNA and piRNA-like piRNA pathway at least in S2 cells (data not shown). This result is consistent with the notion that pilRNA production is resticted to a particular subset of somatic cells.

4.2.6 miRNA* were loaded into Ago2 complex in absence of R2D2

R2D2 is required to load endo-siRNAs into Ago2 and prevent endo-siRNAs from binding to Ago1 (Okamura, Robine et al. 2011). The only exception to this rule were the miRNA* species. We asked if we could see the same phenomenon in our libraries. To visualize all alternative products from the pre-miRNA hairpin, reads matching the mature miRNAs were removed from the libraries and the remainder was mapped to the miRNA precursor hairpins. After normalization to total genome matching reads, 22 nt and 24 nt long reads were preferentially observed. Their processing seemed to proceed independently of R2D2 and Loqs-PD in soma. Interestingly, the 24 nt long reads were preferentially loaded into Ago2 in both *r2d2* or *loqs*^{ko} heterozygous mutants. The stronger reduction after depletion of *r2d2* indicated that R2D2 was primarily involved in the loading process but a subpopulation still remained loaded. Further investigation showed that only a minor fraction of pre-miRNA hairpins generated 24 nt long species including the precursor of miR-263, miR-284, miR-2a-1/2a-2/2b-1/2b-2 and miR-375. Reads originating from the precursor of miR-284 and miR-375 mapped to the miRNA* sequences while the stemloops of miR-263 and the collection of miR-2 species mapped to miRNA producing arm, thus these reads represent iso-miRNAs. Taken together, we could detect that isomiRs and miRNA* species were loaded in a manner that closely resembles other Ago2-loaded species.

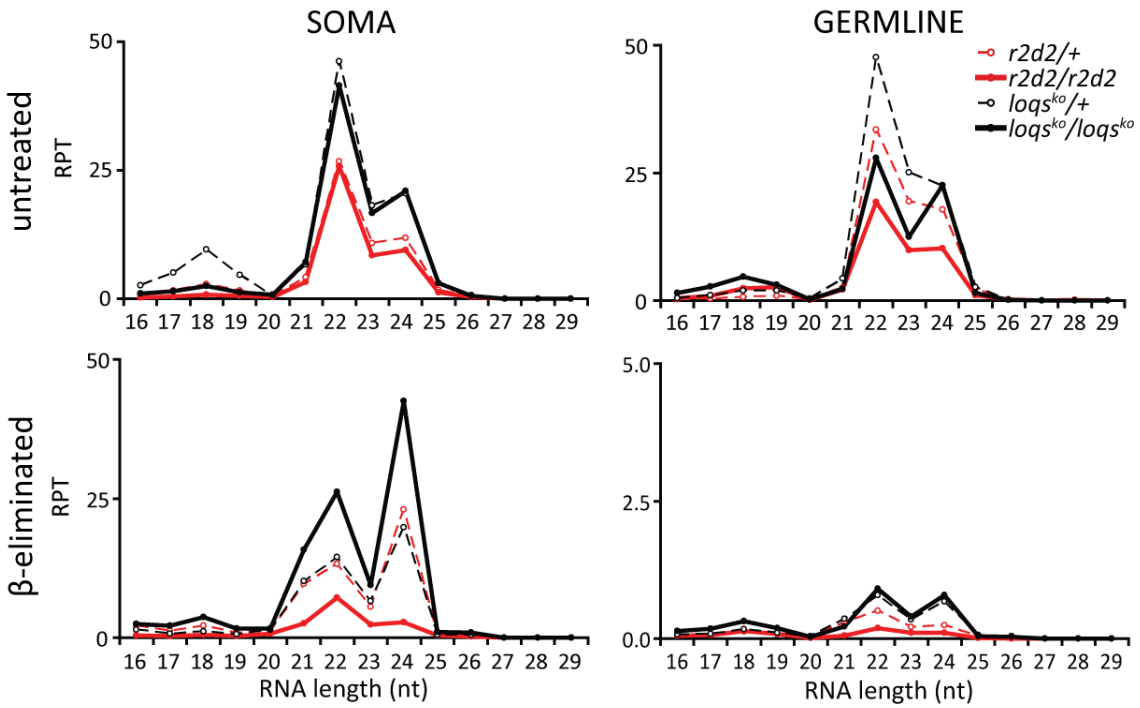


Figure 4.34 The length distribution of miRNA hairpin precursor matching small RNAs in *r2d2* and *loqs^{ko}* mutants.

Reads of each library originating from soma and germ line were filtered for mature miRNAs. The remainder was mapped to the miRNA hairpin precursor. The resulting small RNAs were analyzed for their size distribution and normalized to the total genome matching reads. The normalized counts were expressed as reads per thousand (RPT).

4.2.7 Exo-siRNAs are loaded into RISC independently of R2D2

Drosophila C virus (DCV) (Jousset, Plus et al. 1972), Drosophila A virus (DAV) and Drosophila P virus (DPV) (Plus and Duthoit 1969; Teninges and Plus 1972; Plus, Croizier et al. 1976) were described as natural pathogens of *D. melanogaster* with about 40% of wild populations being infected with one or more of these viruses (Plus, Croizier et al. 1975). They belong to the picorna-like RNA virus family and vary in their pathogenicity, with DCV being the most pathogenic and DAV the least. Additional RNA viruses were shown to infect *D. melanogaster* e.g. Cricket paralysis virus (CrPV) (Moore, Kearns et al. 1980), Flock house virus (FHV) (Chao, Lee et al. 2005).

The library of somatic homozygous *r2d2* mutant after β-elimination mapped with only 31% of reads against the *D. melanogaster* genome. What is the remaining part of reads mapping against? Is the reduction caused by a virus due to the high probability for viral infection in *D. melanogaster*? We scrutinized the *r2d2* homozygous and other libraries for reads matching the frequently occurring RNA viruses comprising DCV, DAV, CrPV and FHV. Exclusively the homozygous *r2d2* mutant in soma and germline revealed small RNAs matching against the genome of DAV but not DCV, CrPV and FHV before and after β-elimination. Although

in general the same RNA preparation was used for the untreated and β-eliminated libraries, for the somatic *r2d2/r2d2* sample two different RNA preparations were used.

In order to additionally confirm the infection with DAV, PCR was the method of choice by using oligos specific for DAV transcript. Three biological RNA samples from *r2d2* and *loqs*^{ko} mutants (n=1, n=2 and n=3) were isolated from soma and germ line, then reverse transcribed. The first replicate represented the RNA samples used for the generation of the small RNA libraries. Solely the homozygous *r2d2* mutants of the first replicate were infected with DAV in soma and germ line, consistent with the results of deep sequencing (Figure 4.35). PCR products were additionally verified by sequencing. Interestingly, the DNA amount in PCR correlated with the amount of viral siRNA reads in libraries. Only one other biological samples was positive exclusively in the germline (second replicate) while no infection was detected for the third replicate.

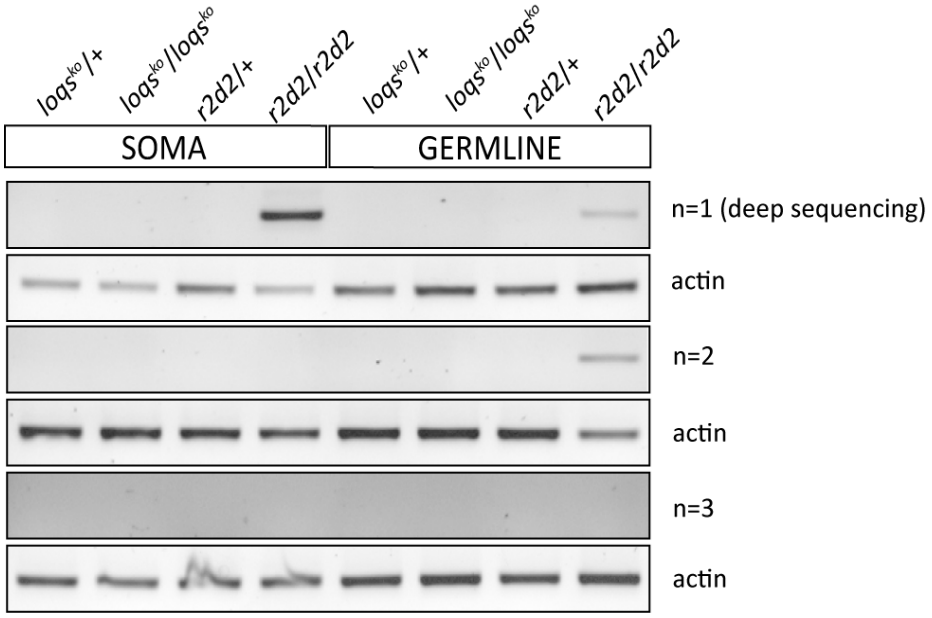


Figure 4.35 Detection of viral infection of *r2d2* and *loqs*^{ko} mutants.

r2d2 and *loqs*^{ko} mutant flies were analyzed for infection of Drosophila A virus (DAV) by detection of the presence of viral transcript by PCR and agarose gel electrophoresis. RNA from three biological replicates was reverse transcribed and the resulting cDNA was analyzed with a DAV specific primer pair. Actin served as DNA quality control.

Viral dsRNA is processed by Dcr-2 into 21 nt long siRNAs which are incorporated into Ago2-RISC complex, where they guide the recognition and endonucleic cleavage of viral target RNAs (Carthew and Sontheimer 2009; Kawamata and Tomari 2010; Blair 2011). If DAV RNA participated in this model by being processed by Dcr-2, then the observed small RNAs in *r2d2* mutant should be produced with preferential length of 21 nt. Indeed, the length distribution depicted the expected peak at 21 nt (Figure 4.36). Furthermore the viral small RNAs in *r2d2* mutants derived from both strands indicating that the dsRNA precursor is cleaved by Dcr-2 (Figure 4.36).

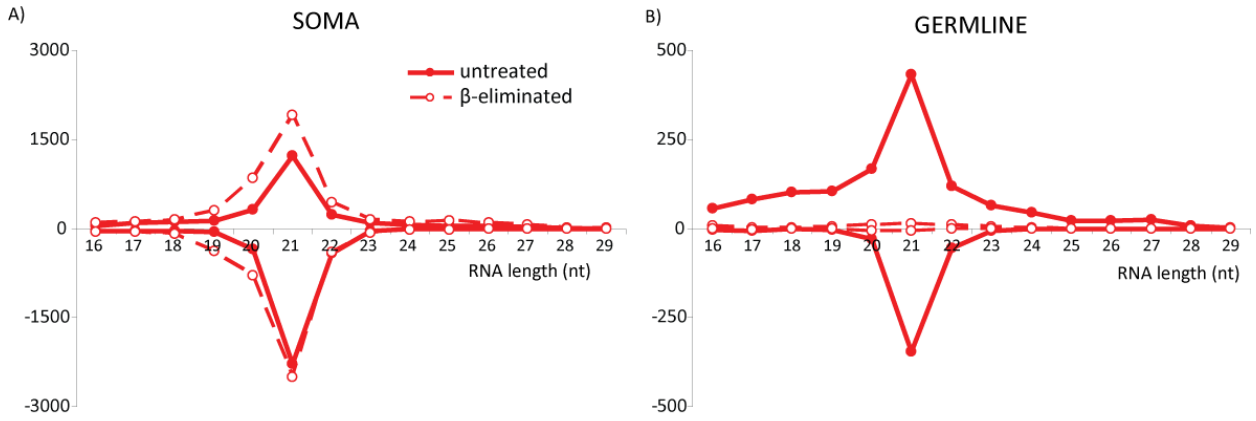


Figure 4.36 The length distribution of viral siRNAs regarding to sense and antisense orientation.

Small RNA libraries generated with untreated and β-eliminated RNA samples from homozygous *r2d2* mutant were mapped to the viral transcript of DAV. The absolute number for sense (+) and antisense (-) transposon matching small RNAs for 16 nt to 29 nt were depicted for soma (A) and germ line (B).

R2D2 is dispensable for converting exogenous dsRNA into siRNA (Liu, Rand et al. 2003) but it is essential for loading and function of exo-siRNAs (Liu, Rand et al. 2003; Tomari, Matranga et al. 2004). This led to the hypothesis that no protection could be ensured in a fly *r2d2* mutant. Taken together, viral siRNAs were observed in the library of homozygous *r2d2* mutant but not in the heterozygous *r2d2* mutant. Moreover libraries of both *loqs*^{ko} mutants lack viral siRNAs. In addition PCR did not detect a RNA transcript of DAV, which indicated that no infection of *loqs*^{ko} flies had occurred. In order to expand this investigation, we analyzed further libraries of total RNA from somatic tissue: *r2d2*/*r2d2*, *r2d2*/CyO, *loqs*^{ko}; P{Loqs-PB}/ *loqs*^{ko} P{Loqs-PB}, *loqs*^{ko}; P{Loqs-PB}/CyO (paper). In contrast, this analysis revealed that *loqs*^{ko} and *r2d2* mutants produced viral siRNAs except for *r2d2/r2d2* mutant. These results were confirmed using PCR to detect the RNA transcript of DAV. In agreement with our libraries, the processed reads showed sense and antisense orientation with length bias at 21 nt (data not shown). All in all, heterozygous and homozygous *r2d2* and *loqs*^{ko} mutants appear to produce viral siRNAs upon infection. Unfortunately the quantitative comparison between both mutants was impossible as no controlled infection was performed by adding the same amount of virus to obtain a defined multiplicity of infection (MOI). But it was obvious that viral siRNAs in homozygous *r2d2* mutant survived the treatment of β-elimination demonstrating that they carried the 2'-*O*-methylation at their 3' end. This revealed that exo-siRNAs can be loaded into Ago2-RISC independently of R2D2, which is in contrast with the current model.

5 DISCUSSION

5.1 Optimization of small RNA deep sequencing

Deep sequencing of small RNAs is a very powerful method for the discovery of new RNAs as well as the quantification of small RNA expression profiles. The cost of next-generation sequencing is still considerable, therefore efficient strategies for sequencing of pooled libraries are essential. The continuous technical advances are steadily increasing the number of reads obtained in an experiment, thus sequencing of multiplexed libraries can now yield sufficient sequencing depth for most applications. The reads can afterwards be sorted bioinformatically through the introduction of specific sequence tags called bar codes. These bar code sequences can be introduced either during the ligation steps or PCR amplification.

Our first multiplex libraries prepared to examine small RNAs during the cell cycle were constructed by introduction of bar codes within the linker appended to the 5’ end of small RNAs. We observed huge biases both according to the different adapter oligos introduced in the 3’ ligation as well as biases due to 5' bar codes (Table 4.1, Figure 4.3). The abundance of the miRNA bantam and its length profile agreed when comparing the same bar codes in different cell cycle phases but not vice versa (Figure 4.3). Consistent with our results, miRNA profiles with the same 5’ ligation bar codes in libraries from two different biological conditions (normal and diseased mouse heart) presented more agreement than miRNA profiles with different bar codes in the same tissue (Alon, Vigneault et al. 2011). Furthermore, after comparison of libraries with different 3’ ligation adapters but the same 5’ ligation barcode, we concluded that the miR-184 bias in the G1 cell cycle phase is caused by the 3’ ligation adapter. These results together indicate that there is sequence preference or possibly dependence on the ligation of adapters to small RNAs. Small RNA libraries from the 293T and mES cells, generated by using a pool of various 5’ and 3’ adapters, demonstrated that each miRNA seems to have a favored adapter pair confirming the previous hypothesis (Jayaprakash, Jabado et al. 2011).

Our first sequencing round consisted of 4 pooled libraries and resulted in an overrepresentation of bantam, miR-184 and miR-8 to different extents (Table 4.1A). We tried to reduce these artifacts by lowering the number of amplification cycles during PCR but did not succeed. Hence, PCR is not responsible for the generation of these biases. Supporting our results, no significant improvement was observed in total RNA libraries from 293T cells after reduction of the number of PCR cycles from 25 to 18 (Jayaprakash, Jabado et al. 2011).

In addition, neither the reverse transcription nor the sequencing technology are generating prominent biases as demonstrated by others, finally ending with the conclusion that the biases in the read distribution are caused primarily by the T4 RNA ligases (Hafner, Renwick et al. 2011). The enzymatic reactions are

sensitive to sequence and structure of their substrates by a varying degree. Indeed, it was shown that RNA secondary structure affects the efficiency of both 3' adapter and 5' adapter ligation steps (Hafner, Renwick et al. 2011; Sorefan, Pais et al. 2012). Thus, distinct RNA structures differ in their reactivity during adapter ligation, resulting in a significant impact on read frequencies. Small RNAs in a stable, nonreactive secondary structure are at risk of exclusion from the libraries. Both families of RNA end-joining enzymes differently impact the ligation bias as Rnl2 favor ss nucleotides downstream of the ligation site and ds nucleotides upstream of the ligation site while Rnl1 has a strong preference for ss ligation site (Sorefan, Pais et al. 2012). The thermodynamic stability of secondary structure also depends on nucleic acid backbone modifications. To reduce the effect of secondary structure to some extent, we used a modified DNA 3' adapter. It was shown that chemical pre-adenylation of 5'-phosphorylated donor molecules extends the range of substrates amenable to RNA ligation (England, Gumport et al. 1977). Further modifications of small RNAs, for example a 2'-*O*-methylation of the 3'-terminal nucleotide, was shown to negatively influence the ligation efficiency and reduce their representation in the library (Munafo and Robb 2010). Recently, a pooled adapter approach was suggested to overcome the limitations of the RNA ligase bias by using various 5' and 3' adapters with additions to the ligating 3' end of the 5' adapter and the 5' end of the 3' adapter (Jayaprakash, Jabado et al. 2011).

In the second part of the thesis, the bar codes were introduced during the PCR step. This greatly reduced the variability between different bar codes but certain artifacts remained. We observed a strikingly large proportion of reads originating from only four transposable elements (*roo*, *297*, *TNFB* and *blood*). This poses a question whether their abundance reflects the biological situation (Table 4.4). The *roo* transposon generates the most abundant ovarian piRNAs (Li, Vagin et al. 2009). We observed that *roo* generates highly abundant small RNAs in soma but their preference for sense orientation indicated that degraded *roo* mRNA might contribute to this. We concluded that *blood* represents a technical artifact since one specific sequence that mapped to a defined position existed exclusively in the homozygous *r2d2* mutant library. Therefore, these four mobile elements were filtered out and the remainder of the results was further analyzed. Taken together, the overrepresentation of specific small RNAs was significantly lower after introduction of bar codes during the PCR than during 5' ligation. One other study using PCR-based bar code introduction almost completely suppressed the bar code bias (Alon, Vigneault et al. 2011). In conclusion, the introduction of bar codes during PCR represents a method for more reliable detection of differentially expressed small RNAs.

5.2 Characterization of a tRNA derived small RNA

On the search for small RNAs which oscillate in abundance during the cell cycle, we identified a 21 nt small RNA derived from the 3' end of an immature glutamic acid (Glu) tRNA (referred as tRNA:E4:62Ad) (Table 4.3). High expression levels were shown via qRT-PCR, but this method did not allow distinguishing between the precursor and the mature sequence. Generally, tRNAs and rRNAs constitute most of the cellular non coding RNAs (ncRNA). It is therefore reasonable to assume that these RNAs generate much more degradation products than others. Those degradation products that harbor a 5'-phosphate and 3'-hydroxyl group will be included in our deep sequencing libraries. For this reason until a few years ago, small RNAs derived from ncRNAs were thought to represent solely random degradation intermediates during biogenesis and turnover. Recently a variety of tRNA derived small RNAs were discovered with cellular functions, which argues against the assumption of being degradation products and demonstrating that they belong to a novel class of small RNAs in a wide range of organisms from prokaryotes to eukaryotes (Lee and Collins 2005; Kawaji, Nakamura et al. 2008; Thompson, Lu et al. 2008).

The differences in abundance of the detected tRNA derived small RNA during the cell cycle were questionable, due to the high biases of our deep sequencing libraries. In addition, the attempted over-expression and inhibition of the tRNA derived small RNA did not impact the cell cycle (Figure 4.9). Recently, expression of a human tRNA derived RNA fragment referred to as tRF-1001 in a prostate cancer cell line was shown to correlate with cell proliferation while its deletion impaired cell cycle progression by specific accumulation of cells in G2 (Lee, Shibata et al. 2009). In our analysis a potential function in the cell cycle could be masked since S2 cells normally contain a high proportion of cells within the G2 cell cycle phase.

A variety of newly discovered small RNAs derived from tRNAs were classified into two groups based on their length and biogenesis. The first class comprises tRNA halves named tsRNAs with 28 to 36 nt length which are broadly conserved from bacteria to humans and play important role as regulators of gene expression nutritional, biological or physicochemical stress (Lee and Collins 2005; Li, Luo et al. 2008; Thompson and Parker 2009; Garcia-Silva, Frugier et al. 2010). The second class are 14 to 22 nt long tRNA fragments (tRF) (Lee, Shibata et al. 2009) which are further sub-classified into 3 groups and comprise sequences derived from the 5' end and 3' end of mature tRNA as well as the 3' end of tRNA precursors named the 3' trailers. tRFs are processed by either Dicer or RNase Z (Haussecker, Huang et al. 2010). We could show that the small RNA we identified resembled the tRFs due to the 21 nt length and its biogenesis. It is processed by the action of tRNase Z during the pre-tRNA processing and contains a track of uridines at its 3' end as a consequence of polymerase III transcription termination (Figure 4.10).

In general, tRFs function as gene regulators at different levels of post-transcriptional regulation (Elbarbary, Takaku et al. 2009; Lee, Shibata et al. 2009; Yamasaki, Ivanov et al. 2009; Haussecker, Huang et al. 2010).

The tRF in our study reproducibly affected Ago1-mediated silencing (Figure 4.12). The endogenous mature tRF could be detected via Northern blot but it was not associated with Ago1 (Figure 4.10, Figure 4.11). A technical reason could explain the result, for example unstable association during the procedure of co-immunopurification. However, the successful co-immunoprecipitation of the Ago1-loaded miRNA bantam makes this hypothesis unlikely. On the other hand, the effect of Ago1-RISC mediated silencing may occur without any participation of the tsRNA due to other cellular processes. The specificity has to be tested by further experiments. As Ago2-mediated silencing was unaffected by tRF using the reporter assay and the association with Ago2 was shown to be slightly over the background level, we followed that tRF is not capable of regulating the expression of target mRNAs similar to siRNA containing Ago2-RISC (Figure 4.11, Figure 4.12).

There may be alternative explanations for the down-regulation of our reporter in response to the tsRNA target site. For example, tRNAse Z was able to cleave a target mRNA bearing a complementary binding site under the direction of an artificial small guide RNAs (sgRNAs) (Tamura, Nashimoto et al. 2003; Nakashima, Takaku et al. 2007). Furthermore, a 5' half of $tRNA_{Glu}$ was reported to work as sgRNA using luciferase activity as a readout in 293 cells (Elbarbary, Takaku et al. 2009). Taken together, the physiological role of the tsRNA still needs to be elucidated in order to make it more than just a by-product of biogenesis.

5.3 R2D2 and Loqs-PD function at least partially redundant within the endo-siRNA pathway

The analysis of small RNAs involved in protection against transposable elements demonstrated that Loqs-PD acts predominantly during processing of dsRNA by Dcr-2, while the function of R2D2 is to ensure that siRNAs are loaded into Ago2. However, some exceptions still exist (Figure 4.24). First, a considerable amount of endo-siRNAs remained in the *loqs*^{ko} mutant, indicating that Loqs-PD may not be required for dicing of all endo-siRNA precursors. Consistent with our study, dsRNA processing activity remained in *loqs* mutant embryos or embryo extracts and was then partially dependent on R2D2 (Marques, Kim et al. 2010). Second, in absence of R2D2 our libraries showed that most but not all transposon matching endo-siRNAs could not be loaded into Ago2 but were redirected into Ago1 which is in agreement with the published hypothesis (Okamura, Robine et al. 2011). Thus, Ago2 loading appears predominantly but not exclusively dependent on R2D2. For a fraction of endo-siRNAs the RLC consisting of Dcr-2 and R2D2 can be therefore bypassed. An alternative explanation would be that if these small RNAs are Ago1-loaded small RNAs, which have escaped the sodium periodate treatment. The thermodynamic asymmetry, an indicator of strand selection, was slightly increased for transposon-matching siRNAs in the absence of R2D2 in the β-eliminated libraries compared to the untreated ones (Figure 4.25). This change suggests that Loqs-PD can substitute

for the function of R2D2 in RLC under certain circumstances. Taken together, the processing and loading steps show a partial redundancy between R2D2 and Loqs-PD.

Why do certain transposons differ from the bulk in their requirements for Loqs-PD and R2D2? They are not distinguished based on their abundance (Figure 4.26). Furthermore, we could exclude that the differential requirement for Loqs-PD or R2D2 is based on specific transposon classes or their presence in specific master control loci (Figure 4.27, Figure 4.28). We therefore could not identify the discriminating feature that imposes a requirement for Loqs-PD or R2D2. One potential reason may be that individual transposons depend on Loqs-PD or R2D2 to a different extent according to their expression level in different cell types. Since we isolated RNA from head and thorax as well as ovary to divide soma from germ line, respectively, we still worked with mixed tissues. Within a given type of cells, both the dsRBD proteins as well as the transposable elements could be differentially expressed due to tissue specific factors.

The role of R2D2 differed in the production of endo-siRNAs when comparing soma with germ line (Figure 4.23, Figure 4.27). In the soma, processing was mostly independent of R2D2 while the absence of R2D2 in germ line resulted in an increased production of endo-siRNAs. The latter effect was not caused by only a few transposons but was visible for most TEs which generated endo-siRNAs in germ line (Figure 4.27). Thus, R2D2 appears to reduce the yield of dsRNA processing, indicating a potential competition between R2D2 and Loqs-PD for Dcr-2. Both dsRBD proteins, R2D2 and Loqs-PD, were shown to interact with an equivalent position on Dcr-2, the helicase domain (Hartig and Forstemann 2011). Consistent with our results, depletion of R2D2 increased the efficiency of endo-siRNAs mediated silencing in *Drosophila* cell culture (Hartig, Esslinger et al. 2009). Analogously, the human dsRBD proteins TRBP and PACT have antagonistic effects on Dcr as TRBP stimulates miRNA dicing and stabilizes Dicer while PACT inhibits miRNA processing (Chendrimada, Gregory et al. 2005; Haase, Jaskiewicz et al. 2005; Ma, MacRae et al. 2008).

In the soma, the primary defense system against TEs consists of endo-siRNAs while piRNAs are more prevalent in the germ line. The amount of piRNAs is indeed tremendous compared to the abundance of endo-siRNAs in soma, indicating that transposons are far more strongly expressed in ovaries. We could detect that the steady-state transposon transcript levels were somewhat lower in comparison to soma, most likely due to the protection performed by such abundant piRNAs. In this context it is surprising, given the predominance of transposon repression by piRNAs, that impaired endo-siRNA biogenesis resulted in a measurable impact on the steady state transcript level for a small number of transposons. This indicates their biological significance despite the comparatively lower abundance (Figure 4.29). In soma the majority of transposons in *loqs*^{ko} and *r2d2* mutants were unchanged (Figure 4.29). This is in agreement with a potential redundancy between Loqs-PD and R2D2 during the biogenesis of endo-siRNA. Furthermore, endo-

siRNAs appear to be generated in excess as the reduction observed in *loqs*^{ko} mutants did not result in a clear increase of steady state transposon mRNA levels. In contrast, the redirection of endo-siRNAs into Ago1 in the *r2d2* mutant resulted in an increase in transposon transcript levels. A likely explanation for this observation is that Ago1 is a slower enzyme and dissociates inefficiently from the cleavage products (Ameres, Hung et al. 2010; Okamura, Robine et al. 2011).

5.4 Confirmation and characterization of somatic piRNA-like RNAs

The Piwi-interacting RNA pathway preserves the integrity of the genome in the germ line, guarding it against the activity of mobile elements. Our analysis of the ovarian tissue confirmed a high abundance of piRNAs (Figure 4.21). We further could detect piRNA-like RNAs (pilRNAs) with 23 to 27 nt length in soma matching transposons but present in significantly smaller quantity (Figure 4.23). These small RNAs were 2'-*O*-methyl modified as demonstrated by their enrichment after β-elimination. The *F-element* represents one transposable element with abundant production of pilRNAs in soma (Figure 4.24). Very few reports so far provided evidence for the occurrence of pilRNAs. A first description of pilRNAs was from libraries of *Drosophila ago2* mutant heads, including the characteristic 2'-*O*-methyl group at their 3' end (Ghildiyal, Seitz et al. 2008). Furthermore, pilRNAs were revealed in multiple somatic tissues of mouse and rhesus macaque as well as human natural killer cells (NK) (Ro, Park et al. 2007; Cichocki, Lenvik et al. 2010; Yan, Hu et al. 2011).

The majority of germ line piRNAs tend to be antisense to transposons (Brennecke, Aravin et al. 2007) and the same orientation bias was observed in soma for pilRNAs (Figure 4.31). Furthermore a ping-pong signature is conserved throughout animals, indicating its fundamental importance for piRNA biogenesis (Aravin, Sachidanandam et al. 2007; Houwing, Kamminga et al. 2007; Murchison, Kheradpour et al. 2008). Primary piRNAs serve as inputs for the piRNA pathway as they initiate the cycle of mutually cleaving interactions between piRNA clusters and transposon mRNAs (Brennecke, Malone et al. 2008). They are characterized by a 5' uridine bias (Brennecke, Aravin et al. 2007; Lau, Robine et al. 2009). The pilRNAs in soma resemble their relatives in germ line as they carry a ping-pong signature (Figure 4.32).

The full piRNA pathway in germ cells requires the presence of all three Piwi-family proteins: Piwi, Aub and Ago3 (Brennecke, Aravin et al. 2007; Gunawardane, Saito et al. 2007). Somatic piRNA-like RNAs are expected to require same set of proteins, though potentially at lower expression levels. We found that in mixed cell somatic samples the Piwi-family genes were expressed close to background levels (Figure 4.33). This could either imply that most somatic cells express very low levels of Piwi-family genes, or that a small subset of cells in adult flies is proficient for the piRNA pathway. As the somatic S2 cell line did neither

produce pilRNA reads nor show expression of Piwi-family proteins, the most likely explanation for the origin of somatic piRNAs is that a small subset of cells exists with an active full piRNA pathway within the soma of flies. Consistently, *in situ* hybridization in several adult macaque tissues indicated that pilRNA expression is localized to specific cell types (Yan, Hu et al. 2011). A possible source for this cell population may be the adult stem cells (also referred as somatic stem cells), which are in low abundance, undifferentiated, able to self-renew and generate the cell types of the organ from which they originate.

Our detected pilRNAs resemble the germ line piRNAs more than the primary piRNAs found in a fly ovarian somatic stem cell line. We anticipate that the pilRNAs we describe function in somatic tissue in a manner that is analogous to their germ line relatives. In addition to the well-known protection against transposable elements via posttranscriptional silencing, germ line piRNAs appear to be involved in heterochromatin formation by HP1 recruitment (Wang and Elgin 2011). Furthermore, pilRNA production detected in a somatic immune cell lineage (NK cells) correlates with CpG methylation, indicating their involvement in stable *cis* silencing (Cichocki, Lenvik et al. 2010). Alltogether, pilRNAs are involved in silencing at the transcriptional and posttranscriptional level. The association of pilRNAs with Piwi-clade proteins as well as the functional role has to be further elucidated. Finally, we did not detect any correlation between endo-siRNA and pilRNA pathways.

5.5 R2D2 and Loqs-PD function at least partially redundant within the exo-siRNA pathway

The antiviral defense mechanism in *Drosophila*, referred to as RNA-based antiviral immunity (RVI), is based on virus specific RNA sensor molecules which serve as the inducer of RNAi. Viral dsRNA, which originates e.g. during viral replication, is processed by Dcr-2 into 21 nt siRNA duplexes, then incorporated with the help of R2D2 into RISC leading to the specific cleavage of viral mRNAs and consequently suppression of virus infection (Galiana-Arnoux, Dostert et al. 2006; van Rij, Saleh et al. 2006; Zambon, Vakharia et al. 2006). Our analysis of deep sequencing libraries demonstrated that Drosophila A virus (DAV) RNA accumulated in *r2d2* homozygous mutants, revealing the unexpected presence of a virus and that virus resistance in adult flies may be compromised in absence of R2D2 (Figure 4.35). Consistent with this notion, *r2d2* homozygous mutant flies ($R2D2^{S165fsX}/R2D2^{1}$, both null alleles) were 10,000-fold more susceptible to FHVΔB2 in comparison to wt and thus inhibited for clearance of FHVΔB2 virus (Han, Luo et al. 2011). In contrast to the last reference, we detected Drosophila A virus (DAV) siRNAs by deep sequencing which resisted to β-elimination, indicating that they were loaded into Ago2-RISC independently of R2D2. Although we cannot make any quantitative statement here this serendipitous observation allows us to conclude that

also virus-derived bona fide exo-siRNAs can to some extent bypass the R2D2 requirement for loading. In addition, analysis of different deep sequencing libraries demonstrated that also heterozygous *loqs*^{ko} mutant flies can be persistently DAV infected, arguing that lack of r2d2 infection is not a prerequisite for DAV infection.

5.6 Outlook

5.6.1 Small RNAs and the cell cycle

Our small RNA analysis across the cell cycle was partly limited by technical obstacles during the generation of small RNA libraries, which prevented us from discovering new cell-cycle dependent small RNAs. A very recent development suggests participation of small RNAs in the double-stranded DNA break repair (Francia, Michelini et al. 2012; Michalik, Bottcher et al. 2012; Wei, Ba et al. 2012; Zhang, Chang et al. 2013). Since the genome is continuously challenged by a variety of genotoxic stresses, DNA double-strand breaks (DSBs) represent the most lethal type of damage and require efficient and accurate repair. To deal with DSBs, cells are equipped with non-homologous end-joining (NHEJ) and the homologous recombination (HR) pathway.
NHEJ joins two DNA ends irrespective of their sequence thus presenting an efficient but error-prone mode of repair, which takes place throughout the cell cycle. In contrast, HR is dependent on DNA resection of the DSB and a sister chromatid as template to perform error-free repair, which is limited to the late S and G2 phases. Thus cell cycle phases are the major determinant for which repair pathway will be used.
In further experiments, *Drosophila* S2 cells could be synchronized into different cell cycle stages by the gentle method of centrifugal elutriation and afterwards treated to generate DSBs. The induction of DSBs can be executed either unspecificly via treatment with ultraviolet (UV) radiation or DNA damaging agents or more specifically by light-inducible endonucleases (Schierling and Pingoud 2012). After induction of DNA damage, deep sequencing of small RNAs over a time course will allow to determine the time of occurrence of DSB-induced small RNAs as well as their abundance and change in context of the cell cycle. In parallel, synchronized cells after the induction of DSBs will be monitored for the cell cycle progression by usage of bromodeoxyuridine (BrdU) assay combined with other cell cycle phase specific DNA dyes. DSB-induced small RNAs are speculated to depend on the resection of DSB (Michalik, Bottcher et al. 2012; Wei, Ba et al. 2012), which occurs as part of the HR repair pathway mainly occurring during the S and G2 cell cycle phase. Therefore corresponding small RNAs are suggested to be detected in those cell cycle phases.
In case of identification of specific small RNAs generated at the flanking regions of a DSB, they can further be inhibited by transfecting LNA inhibitors or usage of small tandem target mimics (STTM) to finally monitor

how they will affect the repair of specific DNA damage sites. The cell cycle arrest might be released after re-addition of the specific small RNAs presenting direct association with cell cycle.

The recognition and signaling to the DSB repair machinery are induced by phosphorylation of histone H2AX which occurs independently of small RNAs (Wei, Ba et al. 2012), indicating that they act downstream as guide molecules directing chromatin modifications or the recruitment of protein complexes to DSB sites. Since a lot of changes occur after DNA damage, it is interesting to elucidate if the occurrence of DSB-induced small RNAs correlates with specific chromatin modification patterns. To do so, chromatin immunoprecipitation after specific introduction of DSB will allow capturing the fraction of the genome which carries the histone modification of interest. Afterwards bisulphate sequencing on the immunoprecipitated material can map the DNA methylation pattern. Comparing these results with the chromatin pattern from cells after inhibition DSB-induced small RNAs will define a potential correlation between both pathways.

5.6.2 Somatic piRNA-like RNAs (pilRNAs)

Somatic piRNA-like RNAs (pilRNAs) are supposed to be expressed in specific types of somatic cells. A recent method referred as TU-tagging was described to enable intact cell type-specific RNA isolation e.g. cell types from central nervous system which are hardly attainable by dissection (Miller, Robinson et al. 2009). TU-tagging is based on the cell type specific expression of UPRT which is achieved via mating flies encoding for the transcription factor Gal4 under the control of a tissue-specific promoter with GAL-4-inducible transgenic UAS-UPRT (upstream activating sequence-UPRT) flies. In the offspring, expressed UPRT couples ribose-5-phosphate to the N1 nitrogen of 4-thiouracil supplied in the food. The resulting product is incorporated into RNA, then coupled with thio-biotin *in vitro* and finally purified via streptavidin. It should be possible to isolate small RNAs via gel purification and generate small RNA libraries for subsequent deep sequencing. In addition, mRNAs from the isolated RNA can be checked for PIWI, AUB and AGO3 expression by qRT-PCR to examine the expression of pilRNA biogenesis factors.

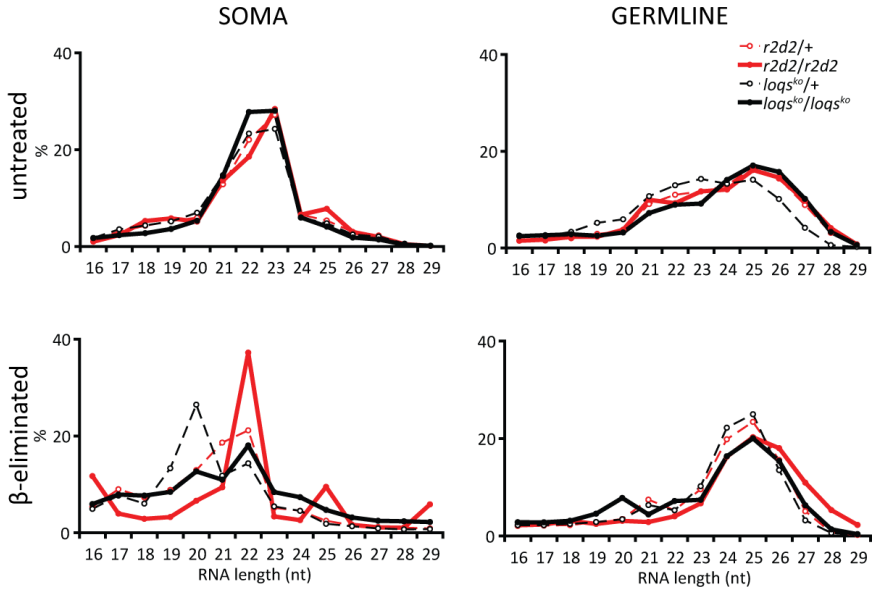
5.6.3 Loqs-PD and R2D2 in endo-siRNA pathway

The redundancy observed between R2D2 and Loqs-PD in endo-siRNA pathway could be further elucidated by focusing on the redundant functional role of different domains encoded in both proteins. To do so, hybrid proteins are generated by exchange of domains between Loqs and R2D2. In our laboratory a lot of work was already performed to generate a variety of such hybrid proteins, which could be introduced into

loqs or *r2d2* mutant flies. The recently published cytoplasmic D2 body was suggested to be the cellular location where endo-siRNAs are loaded onto Ago2 (Nishida, Miyoshi et al. 2013). Dcr-2 and R2D2 are required for D2 body formation but function distinctly as Dcr-2 stabilizes R2D2 whereas R2D2 localizes Dcr-2 to D2 bodies. Furthermore, the dsRNA-binding activity is necessary for R2D2 to localize to D2 bodies. Taken together, analysis of hybrid proteins in context of the D2 body formation will expand our understanding of the redundancy of Loqs-PD and R2D2 in the endo-siRNA pathway.

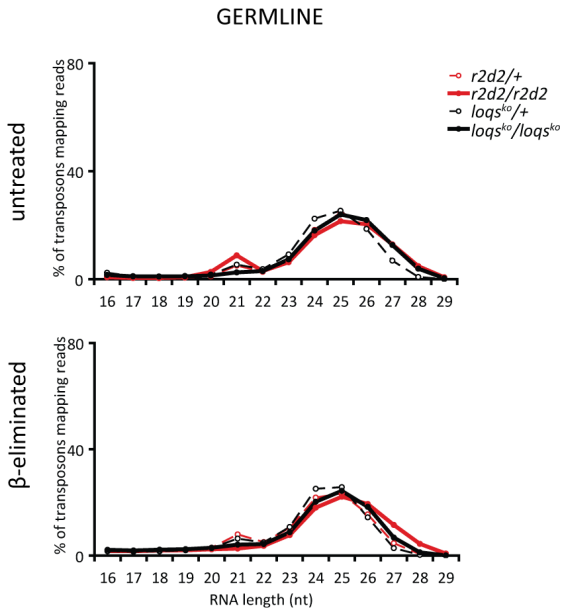
6 APPENDIX

6.1 Deep sequencing analysis



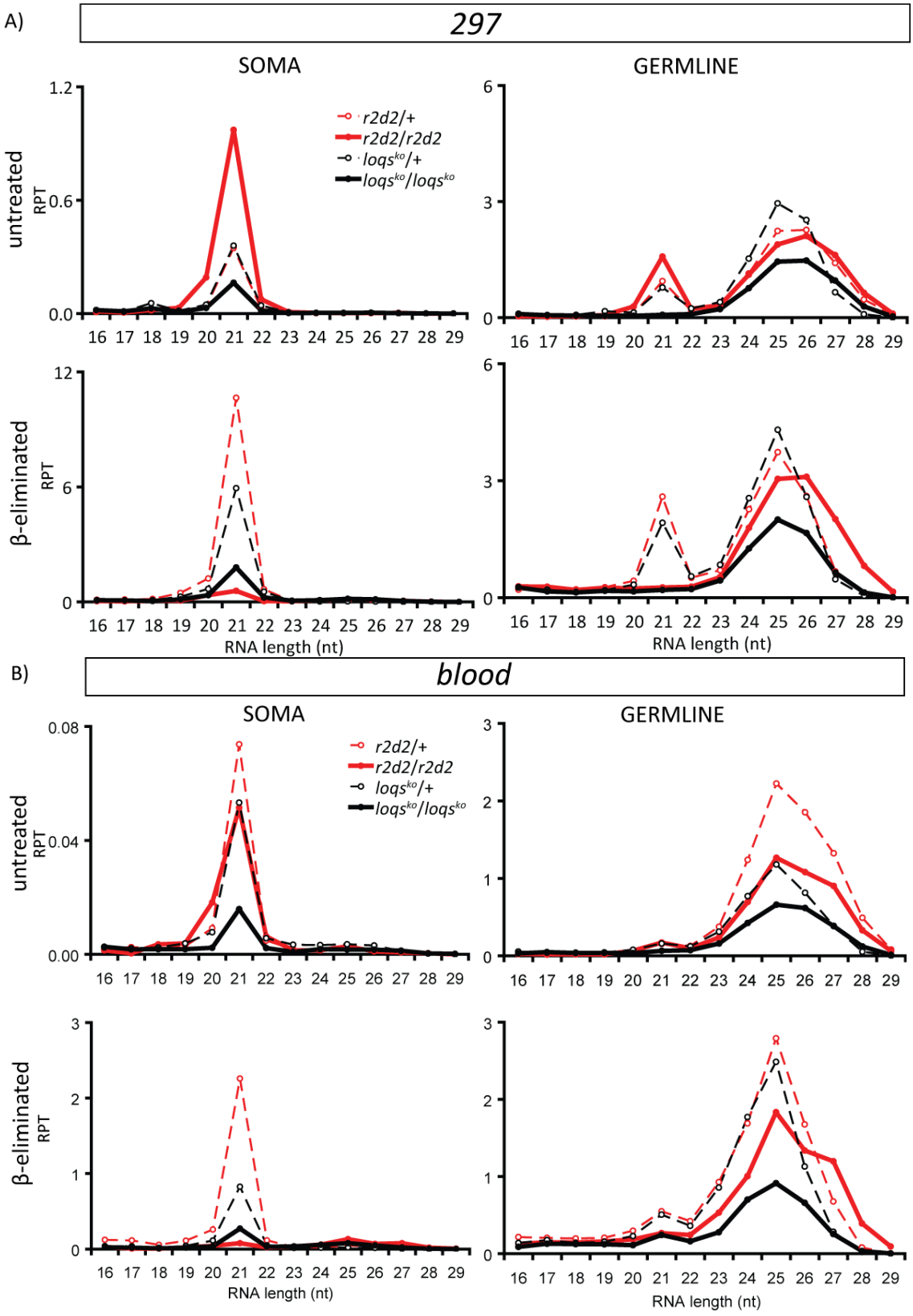
Appendix Figure 1 The length distribution of all reads from small RNA *r2d2* and *loqs*^{ko} mutants libraries.

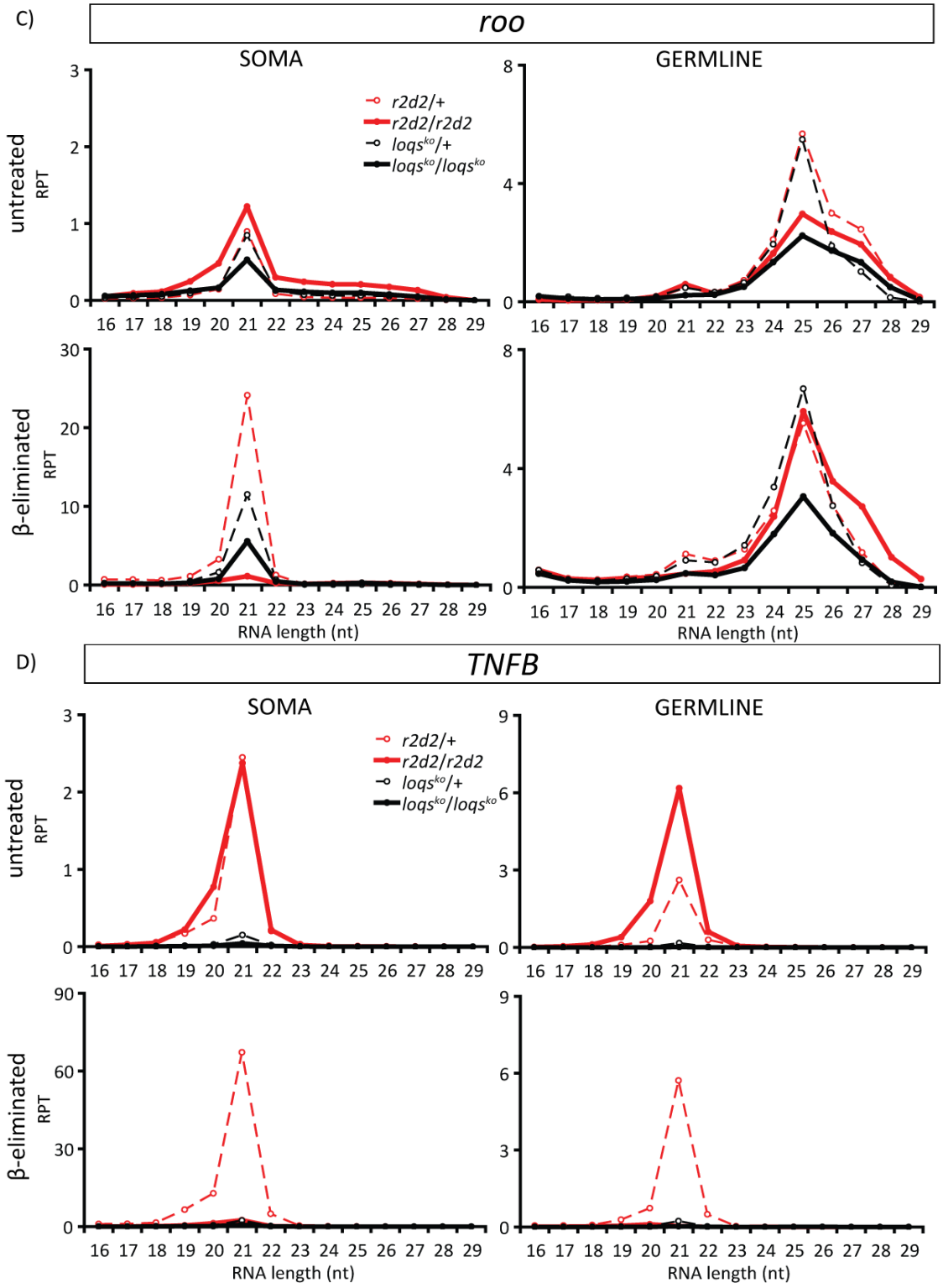
Reads of each *r2d2* and *loqs*^{ko} mutant library were analyzed for their size distribution and expressed as percentage of total reads of each library, respecitively.



Appendix Figure 2 Read length distribution of transposon matching small RNAs in *r2d2* and *loqs*^{ko} mutants.

Reads of each library were mapped to the reference containing transposon sequence collection. The resulting transposon matching small RNAs were analyzed for their size distribution and normalized to the transposons matching reads. The normalized counts were expressed as percentage of total transposon matching reads.


A)
297
SOMA
GERMLINE
r2d2/+
r2d2/r2d2
loqs^ko/+
loqs^ko/loqs^ko
untreated
β-eliminated
RPT
RNA length (nt)
B)
blood
SOMA
GERMLINE
r2d2/+
r2d2/r2d2
loqs^ko/+
loqs^ko/loqs^ko
untreated
β-eliminated
RPT
RNA length (nt)



Appendix Figure 3 Read length distribution of *roo*, *TNFB*, *blood* and *roo* transposon mapping small RNAs in *r2d2* and *loqs*^{ko} mutants.

Reads of each library were mapped to A) *297*, B) *blood* C) *roo* and D) *TNFB* transposon sequence separately and their size distribution was profiled. After normalization to genome matching reads, the counts were expressed as reads per thousand (RPT).

SOMA*

loqs^{ko}/+*				*loqs*^{ko}*/loqs*^{ko}				*r2d2/+*				*r2d2/r2d2*			
	s	as	total		s	as	total		s	as	total		s	as	total
DM_ROO	9139	2534	11673	DM_ROO	26197	4031	30228	DMTNFB	8972	44621	53593	DMTNFB	1462	9101	10563
DMIS297	1043	2572	3615	SPRINGER	4666	412	5078	DM_ROO	17235	6031	23266	DM_ROO	8805	1244	10049
SPRINGER	1731	260	1991	DMIS297	1420	3456	4876	DMIS297	2137	5001	7138	DMIS297	1203	2557	3760
DMTNFB	339	1318	1657	OPUS	2398	845	3243	SPRINGER	3087	390	3477	DMPOGOR11	554	429	983
OPUS	1066	444	1510	DMW1DOC	820	1963	2783	DMRTMGD1	1640	914	2554	SPRINGER	618	63	681
DMW1DOC	449	1008	1457	DMRER1DM	1800	506	2306	DMRER1DM	1869	136	2005	DMRTMGD1	409	99	508
DM23420	471	696	1167	412	992	1146	2138	DMW1DOC	765	1189	1954	DMGYPF1A	260	118	378
412	508	487	995	ROVER	1697	394	2091	DM23420	793	1107	1900	DMW1DOC	184	120	304
F	138	821	959	DM23420	955	1123	2078	OPUS	1135	689	1824	BLOOD	157	108	265
DMRER1DM	611	173	784	F	340	1658	1998	BLOOD	785	875	1660	OPUS	211	44	255
ROVER	486	213	699	DMTNFB	135	1337	1472	F	314	1081	1395	STALKER2	116	90	206
BLOOD	320	326	646	DMGYPF1A	1069	146	1215	STALKER2	575	719	1294	412	113	51	164
STALKER2	389	238	627	STALKER2	932	213	1145	412	718	530	1248	DMBLPP	137	6	143
DME487856	156	194	350	DME487856	348	490	838	DMPOGOR11	692	342	1034	DM23420	66	69	135
DMGYPF1A	272	68	340	DM14101	221	581	802	DMGYPF1A	728	216	944	DMCOPIA	117	7	124
DMCOPIA	302	27	329	DMCOPIA	680	29	709	DMCOPIA	752	44	796	F	25	92	117
STALKER	97	200	297	BLOOD	412	259	671	DMBLPP	685	29	714	DM14101	36	77	113
DMRTMGD1	171	107	278	QUASIMODO	29	640	669	STALKER	255	380	635	DMRER1DM	85	3	88
STALKER4	95	170	265	AF056940	0	661	661	STALKER4	231	345	576	STALKER4	42	45	87
DM14101	88	174	262	GYPSY12	144	464	608	RT1B	442	101	543	AF222049	71	7	78

SOMA**

loqs^{ko}/+				*loqs*^{ko}*/loqs*^{ko}				*r2d2/+*				*r2d2/r2d2*			
	s	as	total		s	as	total		s	as	total		s	as	total
DM_ROO	22051	12248	34299	DM_ROO	7298	4273	11571	DMTNFB	31750	202267	234017	BLOOD	23432	438	23870
DMIS297	5376	12071	17447	F	517	4346	4863	DM_ROO	52050	26470	78520	DMTNFB	628	4226	4854
DMTNFB	1021	6921	7942	DMIS297	1327	2977	4304	DMIS297	9886	23053	32939	DM_ROO	2561	1162	3723
DMW1DOC	1679	4160	5839	DMW1DOC	804	3164	3968	DM23420	3308	5246	8554	F	253	1976	2229
DM23420	2008	3251	5259	412	818	1803	2621	DMRTMGD1	4553	3563	8116	G5_DM	7	1722	1729
412	1650	1899	3549	DMRER1DM	1186	1409	2595	BLOOD	3644	4043	7687	DMIS297	380	1260	1640
F	711	2727	3438	OPUS	780	1697	2477	DMW1DOC	1815	4709	6524	AF009439	0	1225	1225
OPUS	1213	1729	2942	DM23420	967	1436	2403	F	1151	4223	5374	DMW1DOC	210	833	1043
AF009439	0	2692	2692	DME487856	444	1640	2084	OPUS	1876	2711	4587	DMGYPF1A	332	574	906
BLOOD	1267	1397	2664	DM14101	601	1173	1774	STALKER2	2046	2365	4411	DME487856	206	693	899
Tinker	47	2056	2103	RT1B	448	1189	1637	DMPOGOR11	2835	1183	4018	DM14101	296	578	874
DIVER2	59	2034	2093	DMTNFB	69	1418	1487	SPRINGER	1929	1793	3722	DIVER2	59	792	851
DMTRAM	0	1989	1989	IVK	110	1224	1334	412	1510	1974	3484	DMRER1DM	485	364	849
STALKER2	888	972	1860	SPRINGER	546	758	1304	STALKER4	925	1316	2241	DMTRAM	0	805	805
SPRINGER	878	914	1792	Beagle2	233	870	1103	STALKER	916	1283	2199	Tinker	87	692	779
STALKER	361	675	1036	AF009439	0	1098	1098	AF009439	0	1909	1909	OPUS	235	522	757
STALKER4	297	663	960	DME010298	170	908	1078	GYPSY4	677	1008	1685	SPRINGER	403	346	749
DME487856	262	655	917	G2	179	699	878	DMGYPF1A	628	975	1603	IVK	81	647	728
DMRTMGD1	410	461	871	BLOOD	358	514	872	DME487856	414	1146	1560	GYPSY12	308	402	710
DMRER1DM	409	411	820	STALKER	133	735	868	DMRER1DM	772	573	1345	G2	178	489	667

GERMLINE*

loqs^{ko}/+				*loqs*^{ko}*/loqs*^{ko}				*r2d2/+*				*r2d2/r2d2*			
	s	as	total		s	as	total		s	as	total		s	as	total
F	24099	191330	215429	F	27599	278213	305812	F	30745	183070	213815	F	11588	92122	103710
DM_ROO	26525	112580	139105	DMRER1DM	84067	127172	211239	DM_ROO	33082	138340	171422	OPUS	14256	30471	44727
DMW1DOC	14180	92495	106675	DME487856	35625	105980	141605	DMW1DOC	21357	98076	119433	DMW1DOC	6045	38364	44409
DMRER1DM	55817	48688	104505	DM14101	49182	88373	137555	412	14107	97380	111487	DM_ROO	11436	30878	42314
DMIS297	5533	98921	104454	RT1B	34435	92392	126827	DMIS297	8677	87701	96378	DMIS297	3443	32283	35726
RT1B	22146	61296	83442	DMW1DOC	16419	92692	109111	DMRER1DM	47665	45764	93429	DM14101	10610	24786	35396
DME487856	21050	59652	80702	OPUS	28144	79116	107260	BLOOD	7519	77448	84967	DMTNFB	6287	27145	33432
412	11334	62399	73733	412	9440	95004	104444	DME487856	20149	58461	78610	412	3489	26553	30042
OPUS	17448	55475	72923	DM_ROO	20596	67725	88321	DM14101	24550	52109	76659	IVK	2990	26768	29758
DM14101	27825	42351	70176	IVK	11078	77072	88150	DMIS176	15809	58896	74705	DMRER1DM	16242	12146	28388
DME010298	11452	47181	58633	Beagle2	16765	70716	87481	RT1B	17712	53439	71151	G2	6705	21144	27849
Beagle2	15137	40970	56107	DME010298	15505	58499	74004	Beagle2	17651	49656	67307	STALKER	4736	20671	25407
STALKER	6373	48064	54437	G2	11427	55246	66673	STALKER	9344	57772	67116	DMGYPF1A	7258	17812	25070
STALKER4	6423	47426	53849	STALKER	3972	58172	62144	STALKER4	9116	55934	65050	Beagle2	7224	17785	25009
DMIS176	11400	39048	50448	ROXELEMENT	15698	43733	59431	OPUS	17730	47273	65003	STALKER4	4615	18705	23320
GYPSY12	40078	8481	48559	DMIS297	2665	52892	55557	GYPSY12	49804	11273	61077	DMIS176	4029	18175	22204
BLOOD	7417	36888	44305	DM23420	25855	27455	53310	1360	33035	21425	54460	DME487856	5171	16385	21556
IVK	5507	35454	40961	DME278684	16559	36511	53070	DME010298	11652	42135	53787	INVADER	13502	5374	18876
1360	26064	13546	39610	INVADER	33597	16698	50295	AY561850	15221	36465	51686	RT1B	6564	11640	18204
G2	7614	31489	39103	SPRINGER	16194	32464	48658	G2	11356	39158	50514	BLOOD	2626	15331	17957

GERMLINE**

loqs^{ko}/+				*loqs*^{ko}*/loqs*^{ko}				*r2d2/+*				*r2d2/r2d2*			
	s	as	total		s	as	total		s	as	total		s	as	total
F	40950	399438	440388	F	33866	410795	444661	F	15957	142106	158063	F	15897	160920	176817
DMRER1DM	117482	115661	233143	DMRER1DM	126136	158619	284755	DMW1DOC	8346	84235	92581	DM_ROO	19356	81461	100817
DMW1DOC	21899	175723	197622	DME487856	46571	159485	206056	DMRER1DM	48966	37568	86534	DMW1DOC	11335	88769	100104
DM_ROO	36377	155755	192132	DM14101	63066	118466	181532	DME487856	16925	55634	72559	DM14101	23092	56616	79708
RT1B	52318	135243	187561	RT1B	49096	130612	179708	DM_ROO	16136	55951	72087	OPUS	13642	57882	71524
DME487856	44960	142361	187321	DMW1DOC	18344	145417	163761	RT1B	20151	47389	67540	DMRER1DM	33482	35229	68711
DMIS297	10839	135975	146814	OPUS	35018	122513	157531	DMIS297	4478	54948	59426	DMIS297	2703	64040	66743
OPUS	31504	110834	142338	IVK	13843	123406	137249	412	7748	46731	54479	IVK	6738	59396	66134
DM14101	50204	84895	135099	412	12175	117990	130165	DM14101	15087	35027	50114	412	6449	58020	64469
IVK	10837	123307	134144	Beagle2	22863	95391	118254	G2	11776	37583	49359	G2	12016	49710	61726
412	17035	106561	123596	DM_ROO	26938	89170	116108	OPUS	12017	37161	49178	Beagle2	15458	39264	54722
DME010298	18795	95795	114590	DME010298	18385	93217	111602	Beagle2	9946	32380	42326	DMGYPF1A	13130	38836	51966
Beagle2	23068	75776	98844	G2	16515	74283	90798	BLOOD	4765	36551	41316	DME487856	12651	37909	50560
STALKER	8541	83918	92459	ROXELEMENT	21182	60729	81911	DME010298	6555	34285	40840	RT1B	15059	33102	48161
G2	21104	66532	87636	DMIS297	3481	76243	79724	DMIS176	6658	32909	39567	DMIS176	7195	35548	42743
DMIS176	18389	68563	86952	DM23420	34569	39824	74393	AY561850	10609	26972	37581	STALKER	3262	37597	40859
BLOOD	12945	72325	85270	STALKER	5275	67570	72845	GYPSY12	26304	6559	32863	AY561850	12289	27961	40250
GYPSY12	67451	15312	82763	SPRINGER	19554	51408	70962	IVK	3597	28632	32229	BLOOD	4751	34725	39476
STALKER4	8232	70704	78936	DME278684	25410	44174	69584	STALKER	4640	26491	31131	INVADER	26523	12769	39292
DME278684	24040	47447	71487	DMRTMGD1	2922	62904	65826	DMTNFB	3930	27166	31096	STALKER4	2953	33928	36881

*=without beta-elimination
**=after beta-elimination

SOMA*

loqs^{ko}/+				*loqs*^{ko}/*loqs*^{ko}				*r2d2/+*				*r2d2/r2d2*			
	s	as	total		s	as	total		s	as	total		s	as	total
DM_ROO	4155	1839	5994	DM_ROO	7068	2880	9948	DMTNFB	6725	32342	39067	DMTNFB	1102	5683	6785
DMIS297	711	1829	2540	DMIS297	928	2128	3056	DM_ROO	9728	4596	14324	DM_ROO	2625	864	3489
DMTNFB	230	834	1064	DMW1DOC	454	1207	1661	DMIS297	1688	3854	5542	DMIS297	840	1936	2776
DMW1DOC	252	605	857	DM23420	407	697	1104	DMRTMGD1	857	664	1521	DMPOGOR11	363	297	660
DM23420	278	434	712	412	490	542	1032	DM23420	533	786	1319	DMRTMGD1	161	70	231
412	239	244	483	OPUS	427	382	809	BLOOD	540	620	1160	DMGYPF1A	80	83	163
OPUS	267	208	475	DMTNFB	75	710	785	DMW1DOC	283	750	1033	BLOOD	81	66	147
BLOOD	179	197	376	F	157	495	652	STALKER2	317	509	826	DMW1DOC	50	73	123
SPRINGER	232	142	374	SPRINGER	368	186	554	DMPOGOR11	516	230	746	SPRINGER	87	35	122
F	71	282	353	BLOOD	183	112	295	F	180	543	723	STALKER2	51	68	119
STALKER2	156	133	289	STALKER2	160	105	265	OPUS	331	355	686	DM23420	29	38	67
STALKER	59	72	131	STALKER4	83	106	189	SPRINGER	399	246	645	412	30	27	57
STALKER4	59	69	128	STALKER	66	110	176	412	213	260	473	DM14101	14	35	49
FB	118	8	126	DMRER1DM	151	14	165	STALKER	161	229	390	F	5	39	44
DMRTMGD1	60	41	101	AF056940	0	163	163	STALKER4	153	198	351	OPUS	25	13	38
GYPSY4	37	61	98	ROVER	80	71	151	GYPSY4	104	166	270	STALKER4	18	19	37
DMRER1DM	64	20	84	DMGYPF1A	87	24	111	DMGYPF1A	130	99	229	STALKER	11	16	27
GYPSY6	29	48	77	GYPSY4	42	64	106	DMRER1DM	175	15	190	DMU89994	11	14	25
AF222049	50	12	62	QUASIMODO	4	83	87	FB	155	11	166	GYPSY4	11	12	23
DMU89994	34	23	57	AF222049	66	16	82	DM14101	46	78	124	DMCOPIA	20	1	21

SOMA**

loqs^{ko}/+				*loqs*^{ko}/*loqs*^{ko}				*r2d2/+*				*r2d2/r2d2*			
	s	as	total		s	as	total		s	as	total		s	as	total
DM_ROO	16682	9272	25954	DM_ROO	4953	2558	7511	DMTNFB	23230	141042	164272	BLOOD	23296	39	23335
DMIS297	4071	9269	13340	DMIS297	921	1483	2404	DM_ROO	38892	20081	58973	DMTNFB	366	2198	2564
DMTNFB	747	4609	5356	DMW1DOC	499	1320	1819	DMIS297	8026	17963	25989	DM_ROO	830	336	1166
DMW1DOC	1236	2826	4062	DM23420	395	696	1091	DMRTMGD1	3429	2501	5930	DMIS297	192	407	599
DM23420	1407	2248	3655	412	505	498	1003	DM23420	2303	3517	5820	SPRINGER	141	124	265
412	1171	1169	2340	DMTNFB	46	880	926	BLOOD	2669	2819	5488	DMW1DOC	66	163	229
OPUS	893	1120	2013	OPUS	381	479	860	DMW1DOC	1276	3143	4419	DMRTMGD1	114	85	199
BLOOD	908	943	1851	F	171	630	801	STALKER2	1456	1616	3072	DMTN1731	149	44	193
F	471	1280	1751	SPRINGER	243	226	469	OPUS	1263	1754	3017	DMGYPF1A	60	101	161
STALKER2	638	682	1320	BLOOD	217	148	365	F	751	2216	2967	F	21	67	88
SPRINGER	542	580	1122	DMTN1731	291	21	312	DMPOGOR11	2064	707	2771	DMPOGOR11	41	36	77
STALKER	247	358	605	STALKER	67	101	168	SPRINGER	1241	1168	2409	STALKER2	36	39	75
FB	547	40	587	DMRER1DM	75	89	164	412	1032	1096	2128	RT1B	54	16	70
STALKER4	216	345	561	STALKER4	51	82	133	STALKER4	690	795	1485	412	37	29	66
DMRTMGD1	294	240	534	STALKER2	69	62	131	STALKER	657	775	1432	DM23420	16	34	50
GYPSY4	183	276	459	ROVER	27	95	122	GYPSY4	432	656	1088	OPUS	19	27	46
GYPSY6	163	184	347	GYPSY6	53	50	103	DMGYPF1A	421	532	953	DMU89994	17	24	41
ROVER	55	207	262	DM14101	47	53	100	FB	608	64	672	DM14101	14	23	37
DMTN1731	194	47	241	GYPSY4	41	55	96	DM14101	189	285	474	DMRER1DM	29	7	36
DME487856	89	93	182	DMRTMGD1	17	67	84	GYPSY6	206	264	470	DMLINEJA	8	26	34

GERMLINE*

loqs^{ko}/+				*loqs*^{ko}/*loqs*^{ko}				*r2d2/+*				*r2d2/r2d2*			
	s	as	total		s	as	total		s	as	total		s	as	total
F	5139	14888	20027	F	2779	10069	12848	DMTNFB	7527	20254	27781	DMTNFB	4689	17660	22349
OPUS	5400	7950	13350	OPUS	4518	5776	10294	F	4425	10895	15320	F	3754	9477	13231
DMIS297	2130	6349	8479	DMRER1DM	2832	2283	5115	DMIS297	2945	7043	9988	OPUS	5352	7044	12396
GYPSY6	3270	4169	7439	DMW1DOC	549	2611	3160	DMRTMGD1	5173	4581	9754	DMIS297	1566	4124	5690
DMRTMGD1	3036	3117	6153	DM14101	1392	1604	2996	OPUS	3490	4733	8223	DMRTMGD1	3473	1994	5467
ROVER	1579	4423	6002	RT1B	983	1968	2951	STALKER4	2936	3455	6391	STALKER	2301	1904	4205
DM_ROO	2537	2615	5152	DME487856	769	1845	2614	STALKER	2995	3385	6380	GYPSY6	2091	2086	4177
STALKER4	2178	2674	4852	DM_ROO	802	1444	2246	DM_ROO	2557	2489	5046	STALKER4	2310	1803	4113
STALKER	2059	2717	4776	GYPSY6	986	1239	2225	GYPSY6	2110	2479	4589	ROVER	813	2019	2832
DMW1DOC	977	3786	4763	DME278684	756	805	1561	ROVER	1232	3249	4481	DM_ROO	1249	877	2126
DMRER1DM	2506	1476	3982	DME010298	410	1131	1541	DMW1DOC	882	2692	3574	DMW1DOC	255	1071	1326
DM14101	1029	1461	2490	STALKER	292	1172	1464	DMRER1DM	1536	965	2501	DM14101	403	631	1034
DME487856	798	1642	2440	IVK	241	1194	1435	DM14101	761	1251	2012	DMBARI1	641	378	1019
RT1B	872	1472	2344	G2	312	934	1246	AY561850	650	1308	1958	DMRER1DM	571	285	856
DMTNFB	339	1519	1858	412	279	952	1231	DME487856	637	1234	1871	IVK	81	604	685
DM23420	714	1062	1776	Beagle2	338	889	1227	RT1B	712	1007	1719	Beagle2	276	381	657
BLOOD	429	1251	1680	ROXELEMENT	410	769	1179	DMBARI1	988	642	1630	DMGYPF1A	264	392	656
412	725	942	1667	DM23420	505	610	1115	BLOOD	366	1197	1563	AY561850	200	387	587
Beagle2	530	959	1489	ROVER	242	847	1089	Beagle2	551	936	1487	1360	325	253	578
GYPSY12	949	453	1402	STALKER4	281	764	1045	GYPSY12	871	494	1365	G2	133	424	557

GERMLINE**

loqs^{ko}/+				*loqs*^{ko}/*loqs*^{ko}				*r2d2/+*				*r2d2/r2d2*			
	s	as	total		s	as	total		s	as	total		s	as	total
F	9003	31421	40424	F	4727	21387	26114	DMTNFB	3020	20785	23805	F	912	3693	4605
OPUS	11805	17005	28810	OPUS	7576	11066	18642	F	3353	12887	16240	DMRER1DM	1425	1297	2722
DMIS297	4780	15035	19815	DMRER1DM	6149	5754	11903	DMIS297	2212	8543	10755	DMW1DOC	463	2220	2683
GYPSY6	7078	8509	15587	DMW1DOC	1015	6447	7462	OPUS	3665	5159	8824	DM_ROO	642	1734	2376
DMW1DOC	2285	10516	12801	RT1B	1902	4844	6746	DMRTMGD1	2802	2934	5736	G2	380	1750	2130
DMRTMGD1	4759	6951	11710	DM14101	2989	3575	6564	GYPSY6	2011	2735	4746	DM14101	878	1246	2124
DMRER1DM	6566	4784	11350	DME487856	1634	4674	6308	DM_ROO	1897	2764	4661	OPUS	521	1530	2051
ROVER	2631	8665	11296	DM_ROO	1967	3145	5112	DMW1DOC	800	3643	4443	DMGYPF1A	323	1291	1614
DM_ROO	3808	5575	9383	DMRTMGD1	110	4519	4629	DMRER1DM	2694	1690	4384	STALKER	136	1278	1414
STALKER	3095	5251	8346	GYPSY6	1736	2278	4014	STALKER	1754	2308	4062	IVK	146	1254	1400
STALKER4	3233	4442	7675	DME278684	1633	1699	3332	ROVER	753	3009	3762	DME487856	431	939	1370
DM14101	2926	3803	6729	IVK	371	2953	3324	STALKER4	1701	1950	3651	BLOOD	155	1214	1369
RT1B	2376	4171	6547	STALKER	582	2725	3307	DM14101	1140	1693	2833	RT1B	530	837	1367
DME487856	2049	4379	6428	412	644	2455	3099	DME487856	830	1979	2809	DMIS297	126	1204	1330
BLOOD	945	4233	5178	Beagle2	758	2326	3084	RT1B	1006	1741	2747	DMIS176	187	1108	1295
DMIS176	586	4116	4702	DME010298	759	2145	2904	G2	416	2056	2472	AY561850	475	778	1253
DME278684	1711	2431	4142	G2	653	2219	2872	AY561850	748	1535	2283	412	289	926	1215
DM23420	1652	2270	3922	DMIS176	341	2461	2802	DMIS176	280	2000	2280	Beagle2	312	837	1149
DME010298	950	2678	3628	ROXELEMENT	936	1742	2678	BLOOD	317	1934	2251	STALKER4	111	1030	1141
412	1323	2278	3601	BLOOD	1481	1167	2648	Beagle2	491	1142	1633	DMRTMGD1	75	846	921

*=without beta-elimination
**=after beta-elimination

Appendix Table 1 Abundance of *roo*, *297*, *TNFB* and *blood* transposon in libraries.

All libraries from *r2d2* and *loqs*^{ko} mutants were mapped to the transposons sequence collection and sorted in decreasing order. *roo*, *297*, *TNFB* and *blood* transposon matching small RNAs were marked in green. Counts from libraries comprising reads from 16-29 nt were listed in (A) and 21 nt in (B).

6.2 Abbreviations

°C	degrees Celsius
Δ	tuncated
Ago	Argonaute protein
Amp	ampicillin
AMP	adenosine monophosphate
APS	ammonium peroxodisulfate
ATP	adenosine triphosphate
bp	base pair(s)
BSA	bovine serum albumine
cDNA	complementary DNA
CCE	counterflow centrifugal elutriation
co-IP	co-immunoprecipitation
CT-value	cycle of threshold value in qPCR
d	day(s)
D. melanogaster	*Drosophila melanogaster*
dcr	dicer gene
Dcr	Dicer protein
DMSO	dimethyl sulfoxide
DNA	desoxy-ribonucleic acid
dNTP	desoxy-nucleotide-tri-phosphate
ds	double-stranded
dsRBD	double-stranded RNA binding domain
dsRBP	double-stranded RNA binding domain protein
dsRNA	double-stranded RNA
DTT	dithiothreitol
E. coli	*Escherichia coli*
ECL	Enhanced Chemiluminescence
EDTA	ethylenediaminetetraacetic acid
EGFP	Enhanced Green Fluorescent protein
endo-	endogenous
endo-siRNA	endogenous small interfering RNA
exo-	exogenous
exo-siRNA	exogenous small-interfering RNA

FACS	Fluorescence Activated Cell Sorting
FBS	Fetal Bovine Serum
G418	geneticin
GFP	Green Fluorescent Protein
gsRNA	guide small RNA
GST	glutathione S-transferase
h	hour(s)
H_2O	water
HRP	Horseradish Peroxidase
IP	immunoprecipitation
IPTG	Isopropyl-β-D-thiogalactopyranosid
ko	knock-out
luc	luciferase
mg	milligram
min	minute
miR	micro RNA
miRNA	micro RNA
ml	milliliter
mM	millimolar
mRNA	messenger RNA
N_2	nitrogen
ncRNA	non-coding RNA
N-term	protein N-terminus
Neo	neomycin
ng	nanogram
nt	nucleotide(s)
NTP	nucleotide-tri-phosphate
ORF	open reading frame
p.a.	pro analysis
PA/PB/PC/PD	protein isoform A/B/C/D
PAGE	Polyacrylamide Gel Electrophoresis
PAZ	Piwi-Argonaute-Zwille domain of Dicer and Argonaute proteins
PBS	phosphate buffered saline
PCR	Polymerase Chain Reaction
PI	propidium iodide

piRNA	Piwi-interacting RNA
PNK	polynucleotide kinase
Pol II	DNA polymerase II
Poly-A	poly-adenylation
PVDF	Polyvinylidenfluoride
qPCR	quantitative Polymerase Chain Reaction
RISC	RNA induced silencing complex
RITS	RNA-induced transcriptional silencing
RLC	RISC loading complex
RNA	ribonucleic acid
RNAi	RNA interference
RNaseIII	endoribonuclease class III
rpm	rotation per minute
rRNA	ribosomal RNA
RT	reverse transcription or real-time
S2 cell	Schneider-2 cell
SD	standard deviation
SDS	sodium dodecyl sulfate
siRNA	small interfering RNA
SSC	sodium chloride/sodium citrate
SV40	Simian Virus 40
tech.	technical
tRNA	transfer RNA
V	Volt
wt	wildtype
w/v	mass/volume concentration
α	anti
µ	micro
µg	microgram

6.3 Acknowledgements

An erster Stelle möchte ich mich ganz herzlich bei meinem Doktorvater Klaus Förstemann bedanken, der es mir ermöglicht an interessanten, spannenden und abwechslungsreichen Themen zu arbeiten. Die unerschöpflichen wissenschaftlichen Ideen, Anregungen und Ratschläge waren immer bereichernd. Danke für die Motivation und Begeisterung, die ich aus den Besprechungen schöpfen konnte.

Meinem Thesis Advisory Committee Prof. Dierk Niessing und Dr. Birgitta Beatrix danke ich für ihre hilfreichen Ideen und Tipps.

Prof. Mario Halic möchte ich für die Übernahme des Zweitgutachtens danken.

Ich möchte mich sehr herzlich beim Fond der Chemischen Industrie für das Stipendium bedanken. Außer einer internationalen Konferenz habe ich interne Meetings besucht, die für viel Austausch und Kontakt mit interessanten Menschen gesorgt haben.

Bei meinen Labormädels Romy, Steffi, Stephanie und Katha möchte ich mich für die tolle Zusammenarbeit, den wissenschaftlichen Austausch und viele Anregungen während der letzten Jahre bedanken. Das hat es nochmal spannender und interessanter gemacht. Zusätzlich möchte ich euch danken für eure Hilfe wann immer sie nötig war, für die vielen tolle Feste, die ganz besondere Atmosphäre und unzählige Lachkrämpfe, die alle zusammen meine Promotionszeit mit ganz viel Leben, Freude und guter Laune gefüllt haben.
Romy, danke, dass du für mich in jeder Lebenslage, beim Lachen und Weinen, da warst und dich immer so lieb um mich gesorgt hast...Möchte die Zeit mit dir nicht missen.
Steffi, danke dass du immer ein offenes Ohr für meine wilden und wirren Überlegungen hattest, und als tolle Nachbarin mir den Tag mit deinem Humor schöner gemacht hast. Danke für das Korrekturlesen der Arbeit. Unsere Nachbarschaft wird mir fehlen.

Katha (kleinster Cvrcak) möchte ich für ihre Aufmerksamkeit und immer bereichernde Ideen und Denkanstöße danken und das was sich daraus entwickelt hat.

Jens, danke dass du mich immer motivieren und aufmuntern konntest und immer an mich geglaubt hast.

Mein größter Dank gilt meinen Eltern, meinem Bruder und meiner Familie, die mich immer unterstützt haben und an mich geglaubt haben und mir immer wieder neue Kraft gegeben haben. Ohne euch wäre das nicht möglich gewesen!!! Mama i bato, hvala vam za vasu beskrajnu ljubav koja je svemu davala mnogo vise smisla. Hvala vam za vasu podrsku, savete i motivaciju kojima ste mi davali snagu i uvek stajali iza mene. Bez vas moj doktorski rad nebi bio to sto jeste!!

6.4 Curriculum vitae

Milijana Mirkovic-Hösle

Personal Data

Date of birth:	October 3rd, 1982
Place of birth:	Uzice, Serbia
Nationality:	German

Academic Training

2009 – 2013:	Graduate studies at the Gene Center of Ludwig-Maximilians-University Munich in the laboratory of Prof. Dr. Klaus Förstemann
2008 – 2009:	Master Thesis in Keith Humphries's Group, BCCRC, Vancouver, Canada
2006 – 2008:	Master studies in Biochemistry at the Ludwig-Maximilans-University Munich
2003 – 2006:	Bachelor studies in Chemistry and Biochemistry at the Ludwig-Maximilians-University Munich

Education

2001 – 2003:	Gymnasium-Marktoberdorf, Germany
1996 – 2001:	Marien-Gymnasium Kaufbeuren, Germany
1990 – 1996:	Grundschule Andrija Djurovic Uzice, Serbia

Scholarship

2009 – 2011:	PhD fellowship by Stiftung Stipendien-Fonds des Verbandes der chemischen Industrie

Munich, 20/01/2013

Milijana Mirkovic-Hösle

7 REFERENCES

Aleman, L. M., J. Doench, et al. (2007). "Comparison of siRNA-induced off-target RNA and protein effects." Rna **13**(3): 385-395.

Alon, S., F. Vigneault, et al. (2011). "Barcoding bias in high-throughput multiplex sequencing of miRNA." Genome Res **21**(9): 1506-1511.

Ameres, S. L., J. H. Hung, et al. (2010). "Target RNA-directed tailing and trimming purifies the sorting of endo-siRNAs between the two Drosophila Argonaute proteins." RNA **17**(1): 54-63.

Amitsur, M., R. Levitz, et al. (1987). "Bacteriophage T4 anticodon nuclease, polynucleotide kinase and RNA ligase reprocess the host lysine tRNA." EMBO J **6**(8): 2499-2503.

Aravin, A., D. Gaidatzis, et al. (2006). "A novel class of small RNAs bind to MILI protein in mouse testes." Nature **442**(7099): 203-207.

Aravin, A. A., R. Sachidanandam, et al. (2007). "Developmentally regulated piRNA clusters implicate MILI in transposon control." Science **316**(5825): 744-747.

Ashraf, S. I. and S. Kunes (2006). "A trace of silence: memory and microRNA at the synapse." Curr Opin Neurobiol **16**(5): 535-539.

Babiarz, J. E., J. G. Ruby, et al. (2008). "Mouse ES cells express endogenous shRNAs, siRNAs, and other Microprocessor-independent, Dicer-dependent small RNAs." Genes Dev **22**(20): 2773-2785.

Banfalvi, G. (2008). "Cell cycle synchronization of animal cells and nuclei by centrifugal elutriation." Nat Protoc **3**(4): 663-673.

Batista, P. J., J. G. Ruby, et al. (2008). "PRG-1 and 21U-RNAs interact to form the piRNA complex required for fertility in C. elegans." Mol Cell **31**(1): 67-78.

Bennett, P. M. (2008). "Plasmid encoded antibiotic resistance: acquisition and transfer of antibiotic resistance genes in bacteria." Br J Pharmacol **153 Suppl 1**: S347-357.

Bernstein, E., S. Y. Kim, et al. (2003). "Dicer is essential for mouse development." Nat Genet **35**(3): 215-217.

Berretta, J. and A. Morillon (2009). "Pervasive transcription constitutes a new level of eukaryotic genome regulation." EMBO Rep **10**(9): 973-982.

Blair, C. D. (2011). "Mosquito RNAi is the major innate immune pathway controlling arbovirus infection and transmission." Future Microbiol **6**(3): 265-277.

Blumenstiel, J. P. and D. L. Hartl (2005). "Evidence for maternally transmitted small interfering RNA in the repression of transposition in Drosophila virilis." Proc Natl Acad Sci U S A **102**(44): 15965-15970.

Bohnsack, M. T., K. Czaplinski, et al. (2004). "Exportin 5 is a RanGTP-dependent dsRNA-binding protein that mediates nuclear export of pre-miRNAs." RNA **10**(2): 185-191.

Brennecke, J., A. A. Aravin, et al. (2007). "Discrete small RNA-generating loci as master regulators of transposon activity in Drosophila." Cell **128**(6): 1089-1103.

Brennecke, J., C. D. Malone, et al. (2008). "An epigenetic role for maternally inherited piRNAs in transposon silencing." Science **322**(5906): 1387-1392.

Bucheton, A., J. M. Lavige, et al. (1976). "Non-mendelian female sterility in Drosophila melanogaster: quantitative variations in the efficiency of inducer and reactive strains." Heredity (Edinb) **36**(3): 305-314.

Calin, G. A. and C. M. Croce (2006). "MicroRNA signatures in human cancers." Nat Rev Cancer **6**(11): 857-866.

Calin, G. A., C. G. Liu, et al. (2004). "MicroRNA profiling reveals distinct signatures in B cell chronic lymphocytic leukemias." Proc Natl Acad Sci U S A **101**(32): 11755-11760.

Cam, H. P., T. Sugiyama, et al. (2005). "Comprehensive analysis of heterochromatin- and RNAi-mediated epigenetic control of the fission yeast genome." Nat Genet **37**(8): 809-819.

Carthew, R. W. and E. J. Sontheimer (2009). "Origins and Mechanisms of miRNAs and siRNAs." Cell **136**(4): 642-655.

Chao, J. A., J. H. Lee, et al. (2005). "Dual modes of RNA-silencing suppression by Flock House virus protein B2." Nat Struct Mol Biol **12**(11): 952-957.

Chendrimada, T. P., R. I. Gregory, et al. (2005). "TRBP recruits the Dicer complex to Ago2 for microRNA processing and gene silencing." Nature **436**(7051): 740-744.

Cho, W. C. (2007). "OncomiRs: the discovery and progress of microRNAs in cancers." Mol Cancer **6**: 60.

Chung, W. J., K. Okamura, et al. (2008). "Endogenous RNA interference provides a somatic defense against Drosophila transposons." Curr Biol **18**(11): 795-802.

Cichocki, F., T. Lenvik, et al. (2010). "Cutting edge: KIR antisense transcripts are processed into a 28-base PIWI-like RNA in human NK cells." J Immunol **185**(4): 2009-2012.

Cox, D. N., A. Chao, et al. (1998). "A novel class of evolutionarily conserved genes defined by piwi are essential for stem cell self-renewal." Genes Dev **12**(23): 3715-3727.

Cox, D. N., A. Chao, et al. (2000). "piwi encodes a nucleoplasmic factor whose activity modulates the number and division rate of germline stem cells." Development **127**(3): 503-514.

Czech, B., C. D. Malone, et al. (2008). "An endogenous small interfering RNA pathway in Drosophila." Nature **453**(7196): 798-802.

Deshpande, G., G. Calhoun, et al. (2005). "Drosophila argonaute-2 is required early in embryogenesis for the assembly of centric/centromeric heterochromatin, nuclear division, nuclear migration, and germ-cell formation." Genes Dev **19**(14): 1680-1685.

Durand-Dubief, M. and P. Bastin (2003). "TbAGO1, an argonaute protein required for RNA interference, is involved in mitosis and chromosome segregation in Trypanosoma brucei." BMC Biol **1**: 2.

Elbarbary, R. A., H. Takaku, et al. (2009). "Modulation of gene expression by human cytosolic tRNase Z(L) through 5'-half-tRNA." PLoS One **4**(6): e5908.

England, T. E., R. I. Gumport, et al. (1977). "Dinucleoside pyrophosphate are substrates for T4-induced RNA ligase." Proc Natl Acad Sci U S A **74**(11): 4839-4842.

Fire, A., S. Xu, et al. (1998). "Potent and specific genetic interference by double-stranded RNA in Caenorhabditis elegans." Nature **391**(6669): 806-811.

Forstemann, K., M. D. Horwich, et al. (2007). "Drosophila microRNAs Are Sorted into Functionally Distinct Argonaute Complexes after Production by Dicer-1." Cell **130**(2): 287-297.

Forstemann, K., Y. Tomari, et al. (2005). "Normal microRNA maturation and germ-line stem cell maintenance requires Loquacious, a double-stranded RNA-binding domain protein." PLoS Biol **3**(7): e236.

Fox, M. H., R. A. Read, et al. (1987). "Comparison of synchronized Chinese hamster ovary cells obtained by mitotic shake-off, hydroxyurea, aphidicolin, or methotrexate." Cytometry **8**(3): 315-320.

Francia, S., F. Michelini, et al. (2012). "Site-specific DICER and DROSHA RNA products control the DNA-damage response." Nature **488**(7410): 231-235.

Galiana-Arnoux, D., C. Dostert, et al. (2006). "Essential function in vivo for Dicer-2 in host defense against RNA viruses in drosophila." Nat Immunol **7**(6): 590-597.

Garcia-Silva, M. R., M. Frugier, et al. (2010). "A population of tRNA-derived small RNAs is actively produced in Trypanosoma cruzi and recruited to specific cytoplasmic granules." Mol Biochem Parasitol **171**(2): 64-73.

Ghildiyal, M., H. Seitz, et al. (2008). "Endogenous siRNAs derived from transposons and mRNAs in Drosophila somatic cells." Science **320**(5879): 1077-1081.

Ghildiyal, M. and P. D. Zamore (2009). "Small silencing RNAs: an expanding universe." Nat Rev Genet **10**(2): 94-108.

Girard, A., R. Sachidanandam, et al. (2006). "A germline-specific class of small RNAs binds mammalian Piwi proteins." Nature **442**(7099): 199-202.

Grishok, A., A. E. Pasquinelli, et al. (2001). "Genes and mechanisms related to RNA interference regulate expression of the small temporal RNAs that control C. elegans developmental timing." Cell **106**(1): 23-34.

Gunawardane, L. S., K. Saito, et al. (2007). "A slicer-mediated mechanism for repeat-associated siRNA 5' end formation in Drosophila." Science **315**(5818): 1587-1590.

Haase, A. D., L. Jaskiewicz, et al. (2005). "TRBP, a regulator of cellular PKR and HIV-1 virus expression, interacts with Dicer and functions in RNA silencing." EMBO Rep **6**(10): 961-967.

Hafner, M., N. Renwick, et al. (2011). "RNA-ligase-dependent biases in miRNA representation in deep-sequenced small RNA cDNA libraries." RNA **17**(9): 1697-1712.

Hall, I. M., G. D. Shankaranarayana, et al. (2002). "Establishment and maintenance of a heterochromatin domain." Science **297**(5590): 2232-2237.

Han, Y. H., Y. J. Luo, et al. (2011). "RNA-based immunity terminates viral infection in adult Drosophila in the absence of viral suppression of RNA interference: characterization of viral small interfering RNA populations in wild-type and mutant flies." J Virol **85**(24): 13153-13163.

Harfe, B. D. (2005). "MicroRNAs in vertebrate development." Curr Opin Genet Dev **15**(4): 410-415.

Hartig, J. V., S. Esslinger, et al. (2009). "Endo-siRNAs depend on a new isoform of loquacious and target artificially introduced, high-copy sequences." EMBO J **28**(19): 2932-2944.

Hartig, J. V. and K. Forstemann (2011). "Loqs-PD and R2D2 define independent pathways for RISC generation in Drosophila." Nucleic Acids Res **39**(9): 3836-3851.

Hartig, J. V., Y. Tomari, et al. (2007). "piRNAs--the ancient hunters of genome invaders." Genes Dev **21**(14): 1707-1713.

Haussecker, D., Y. Huang, et al. (2010). "Human tRNA-derived small RNAs in the global regulation of RNA silencing." RNA **16**(4): 673-695.

Ho, C. K. and S. Shuman (2002). "Bacteriophage T4 RNA ligase 2 (gp24.1) exemplifies a family of RNA ligases found in all phylogenetic domains." Proc Natl Acad Sci U S A **99**(20): 12709-12714.

Horwich, M. D., C. Li, et al. (2007). "The Drosophila RNA methyltransferase, DmHen1, modifies germline piRNAs and single-stranded siRNAs in RISC." Curr Biol **17**(14): 1265-1272.

Houwing, S., L. M. Kamminga, et al. (2007). "A role for Piwi and piRNAs in germ cell maintenance and transposon silencing in Zebrafish." Cell **129**(1): 69-82.

Hutvagner, G., J. McLachlan, et al. (2001). "A cellular function for the RNA-interference enzyme Dicer in the maturation of the let-7 small temporal RNA." Science **293**(5531): 834-838.

Ivanovska, I., A. S. Ball, et al. (2008). "MicroRNAs in the miR-106b family regulate p21/CDKN1A and promote cell cycle progression." Mol Cell Biol.

Jayaprakash, A. D., O. Jabado, et al. (2011). "Identification and remediation of biases in the activity of RNA ligases in small-RNA deep sequencing." Nucleic Acids Res **39**(21): e141.

Jiang, F., X. Ye, et al. (2005). "Dicer-1 and R3D1-L catalyze microRNA maturation in Drosophila." Genes Dev **19**(14): 1674-1679.

Johnson, S. M., H. Grosshans, et al. (2005). "RAS is regulated by the let-7 microRNA family." Cell **120**(5): 635-647.

Jousset, F. X., N. Plus, et al. (1972). "[Existence in Drosophila of 2 groups of picornavirus with different biological and serological properties]." C R Acad Sci Hebd Seances Acad Sci D **275**(25): 3043-3046.

Kanellopoulou, C., S. A. Muljo, et al. (2005). "Dicer-deficient mouse embryonic stem cells are defective in differentiation and centromeric silencing." Genes Dev **19**(4): 489-501.

Kawaji, H., M. Nakamura, et al. (2008). "Hidden layers of human small RNAs." BMC Genomics **9**: 157.

Kawamata, T. and Y. Tomari (2010). "Making RISC." Trends Biochem Sci **35**(7): 368-376.

Kawamura, Y., K. Saito, et al. (2008). "Drosophila endogenous small RNAs bind to Argonaute 2 in somatic cells." Nature **453**(7196): 793-797.

Kawaoka, S., H. Mitsutake, et al. (2012). "A role for transcription from a piRNA cluster in de novo piRNA production." RNA **18**(2): 265-273.

Ketting, R. F., S. E. Fischer, et al. (2001). "Dicer functions in RNA interference and in synthesis of small RNA involved in developmental timing in C. elegans." Genes Dev **15**(20): 2654-2659.

Khurana, J. S., J. Wang, et al. (2011). "Adaptation to P element transposon invasion in Drosophila melanogaster." Cell **147**(7): 1551-1563.

Khvorova, A., A. Reynolds, et al. (2003). "Functional siRNAs and miRNAs exhibit strand bias." Cell **115**(2): 209-216.

Kim, V. N., J. Han, et al. (2009). "Biogenesis of small RNAs in animals." Nat Rev Mol Cell Biol **10**(2): 126-139.

Kogan, G. L., A. V. Tulin, et al. (2003). "The GATE retrotransposon in Drosophila melanogaster: mobility in heterochromatin and aspects of its expression in germline tissues." Mol Genet Genomics **269**(2): 234-242.

Krek, A., D. Grun, et al. (2005). "Combinatorial microRNA target predictions." Nat Genet **37**(5): 495-500.

Lau, N. C., N. Robine, et al. (2009). "Abundant primary piRNAs, endo-siRNAs, and microRNAs in a Drosophila ovary cell line." Genome Res **19**(10): 1776-1785.

Lau, N. C., A. G. Seto, et al. (2006). "Characterization of the piRNA complex from rat testes." Science **313**(5785): 363-367.

Lee, D. W., K. Y. Seong, et al. (2004). "Properties of unpaired DNA required for efficient silencing in Neurospora crassa." Genetics **167**(1): 131-150.

Lee, S. R. and K. Collins (2005). "Starvation-induced cleavage of the tRNA anticodon loop in Tetrahymena thermophila." J Biol Chem **280**(52): 42744-42749.

Lee, Y., K. Jeon, et al. (2002). "MicroRNA maturation: stepwise processing and subcellular localization." EMBO J **21**(17): 4663-4670.

Lee, Y. S., K. Nakahara, et al. (2004). "Distinct roles for Drosophila Dicer-1 and Dicer-2 in the siRNA/miRNA silencing pathways." Cell **117**(1): 69-81.

Lee, Y. S., Y. Shibata, et al. (2009). "A novel class of small RNAs: tRNA-derived RNA fragments (tRFs)." Genes Dev **23**(22): 2639-2649.

Lewis, B. P., C. B. Burge, et al. (2005). "Conserved seed pairing, often flanked by adenosines, indicates that thousands of human genes are microRNA targets." Cell **120**(1): 15-20.

Li, C., V. V. Vagin, et al. (2009). "Collapse of germline piRNAs in the absence of Argonaute3 reveals somatic piRNAs in flies." Cell **137**(3): 509-521.

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

Li, Y., J. Luo, et al. (2008). "Stress-induced tRNA-derived RNAs: a novel class of small RNAs in the primitive eukaryote Giardia lamblia." Nucleic Acids Res **36**(19): 6048-6055.

Lindahl, P. E. (1948). "Principle of a counter-streaming centrifuge for the separation of particles of different sizes." Nature **161**(4095): 648.

Lindahl, P. E. (1956). "On counter streaming centrifugation in the separation of cells and cell fragments." Biochim Biophys Acta **21**(3): 411-415.

Liu, Q., T. A. Rand, et al. (2003). "R2D2, a bridge between the initiation and effector steps of the Drosophila RNAi pathway." Science **301**(5641): 1921-1925.

Liu, X., F. Jiang, et al. (2006). "Dicer-2 and R2D2 coordinately bind siRNA to promote assembly of the siRISC complexes." Rna **12**(8): 1514-1520.

Liu, Y., X. Ye, et al. (2009). "C3PO, an endoribonuclease that promotes RNAi by facilitating RISC activation." Science **325**(5941): 750-753.

Livak, K. J. and T. D. Schmittgen (2001). "Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method." Methods **25**(4): 402-408.

Lu, J., G. Getz, et al. (2005). "MicroRNA expression profiles classify human cancers." Nature **435**(7043): 834-838.

Ma, E., I. J. MacRae, et al. (2008). "Autoinhibition of human dicer by its internal helicase domain." J Mol Biol **380**(1): 237-243.

Malone, C. D., J. Brennecke, et al. (2009). "Specialized piRNA pathways act in germline and somatic tissues of the Drosophila ovary." Cell **137**(3): 522-535.

Marques, J. T., K. Kim, et al. (2010). "Loqs and R2D2 act sequentially in the siRNA pathway in Drosophila." Nat Struct Mol Biol **17**(1): 24-30.

Matherly, L. H., J. D. Schuetz, et al. (1989). "A method for the synchronization of cultured cells with aphidicolin: application to the large-scale synchronization of L1210 cells and the study of the cell cycle regulation of thymidylate synthase and dihydrofolate reductase." Anal Biochem **182**(2): 338-345.

Mathews, M. B. and A. M. Francoeur (1984). "La antigen recognizes and binds to the 3'-oligouridylate tail of a small RNA." Mol Cell Biol **4**(6): 1134-1140.

Mayer, M., S. Schiffer, et al. (2000). "tRNA 3' processing in plants: nuclear and mitochondrial activities differ." Biochemistry **39**(8): 2096-2105.

Michalik, K. M., R. Bottcher, et al. (2012). "A small RNA response at DNA ends in Drosophila." Nucleic Acids Res **40**(19): 9596-9603.

Miller, M. R., K. J. Robinson, et al. (2009). "TU-tagging: cell type-specific RNA isolation from intact complex tissues." Nat Methods **6**(6): 439-441.

Mochizuki, K., N. A. Fine, et al. (2002). "Analysis of a piwi-related gene implicates small RNAs in genome rearrangement in tetrahymena." Cell **110**(6): 689-699.

Moore, N. F., A. Kearns, et al. (1980). "Characterization of cricket paralysis virus-induced polypeptides in Drosophila cells." J Virol **33**(1): 1-9.

Munafo, D. B. and G. B. Robb (2010). "Optimization of enzymatic reaction conditions for generating representative pools of cDNA from small RNA." RNA **16**(12): 2537-2552.

Murchison, E. P., P. Kheradpour, et al. (2008). "Conservation of small RNA pathways in platypus." Genome Res **18**(6): 995-1004.

Nakashima, A., H. Takaku, et al. (2007). "Gene silencing by the tRNA maturase tRNase ZL under the direction of small-guide RNA." Gene Ther **14**(1): 78-85.

Nishida, Kazumichi M., K. Miyoshi, et al. (2013). "Roles of R2D2, a Cytoplasmic D2 Body Component, in the Endogenous siRNA Pathway in Drosophila." Molecular cell.

Nishida, K. M., K. Saito, et al. (2007). "Gene silencing mechanisms mediated by Aubergine piRNA complexes in Drosophila male gonad." Rna **13**(11): 1911-1922.

Nowacki, M., B. P. Higgins, et al. (2009). "A functional role for transposases in a large eukaryotic genome." Science **324**(5929): 935-938.

Okamura, K., S. Balla, et al. (2008). "Two distinct mechanisms generate endogenous siRNAs from bidirectional transcription in Drosophila melanogaster." Nat Struct Mol Biol **15**(6): 581-590.

Okamura, K., W. J. Chung, et al. (2008). "The Drosophila hairpin RNA pathway generates endogenous short interfering RNAs." Nature **453**(7196): 803-806.

Okamura, K., A. Ishizuka, et al. (2004). "Distinct roles for Argonaute proteins in small RNA-directed RNA cleavage pathways." Genes Dev **18**(14): 1655-1666.

Okamura, K. and E. C. Lai (2008). "Endogenous small interfering RNAs in animals." Nat Rev Mol Cell Biol **9**(9): 673-678.

Okamura, K., N. Robine, et al. (2011). "R2D2 organizes small regulatory RNA pathways in Drosophila." Mol Cell Biol **31**(4): 884-896.

Paddison, P. J., A. A. Caudy, et al. (2002). "Stable suppression of gene expression by RNAi in mammalian cells." Proc Natl Acad Sci U S A **99**(3): 1443-1448.

Pal-Bhadra, M., U. Bhadra, et al. (1999). "Cosuppression of nonhomologous transgenes in Drosophila involves mutually related endogenous sequences." Cell **99**(1): 35-46.

Pal-Bhadra, M., U. Bhadra, et al. (2002). "RNAi related mechanisms affect both transcriptional and posttranscriptional transgene silencing in Drosophila." Mol Cell **9**(2): 315-327.

Pal-Bhadra, M., B. A. Leibovitch, et al. (2004). "Heterochromatic silencing and HP1 localization in Drosophila are dependent on the RNAi machinery." Science **303**(5658): 669-672.

Park, J. K., X. Liu, et al. (2007). "The miRNA pathway intrinsically controls self-renewal of Drosophila germline stem cells." Curr Biol **17**(6): 533-538.

Pascal, J. M. (2008). "DNA and RNA ligases: structural variations and shared mechanisms." Curr Opin Struct Biol **18**(1): 96-105.

Picard, G., A. Bucheton, et al. (1972). "[A sterility phenomenon of nonmendelian determinism in Drosophila melanogaster]." C R Acad Sci Hebd Seances Acad Sci D **275**(8): 933-936.

Plasterk, R. H. (2002). "RNA silencing: the genome's immune system." Science **296**(5571): 1263-1265.

Plus, N., G. Croizier, et al. (1975). "Picornaviruses of laboratory and wild Drosophila melanogaster: geographical distribution and serotypic composition." Ann Microbiol (Paris) **126**(1): 107-117.

Plus, N., G. Croizier, et al. (1976). "A comparison of buoyant density and polypeptides of Drosophila P, C and A viruses." Intervirology **7**(6): 346-350.

Plus, N. and J. L. Duthoit (1969). "[A new Drosophila melanogaster virus, the P virus]." C R Acad Sci Hebd Seances Acad Sci D **268**(18): 2313-2315.

Poy, M. N., L. Eliasson, et al. (2004). "A pancreatic islet-specific microRNA regulates insulin secretion." Nature **432**(7014): 226-230.

Rajewsky, N. and N. D. Socci (2004). "Computational identification of microRNA targets." Dev Biol **267**(2): 529-535.

Reinhart, B. J. and D. P. Bartel (2002). "Small RNAs correspond to centromere heterochromatic repeats." Science **297**(5588): 1831.

Ro, S., C. Park, et al. (2007). "Cloning and expression profiling of testis-expressed piRNA-like RNAs." RNA **13**(10): 1693-1702.

Robine, N., N. C. Lau, et al. (2009). "A broadly conserved pathway generates 3'UTR-directed primary piRNAs." Curr Biol **19**(24): 2066-2076.

Rodriguez, A., S. Griffiths-Jones, et al. (2004). "Identification of mammalian microRNA host genes and transcription units." Genome Res **14**(10A): 1902-1910.

Ronsseray, S., M. Lehmann, et al. (1991). "The maternally inherited regulation of P elements in Drosophila melanogaster can be elicited by two P copies at cytological site 1A on the X chromosome." Genetics **129**(2): 501-512.

Ronsseray, S., M. Lehmann, et al. (1996). "The regulatory properties of autonomous subtelomeric P elements are sensitive to a Suppressor of variegation in Drosophila melanogaster." Genetics **143**(4): 1663-1674.

Saito, K., S. Inagaki, et al. (2009). "A regulatory circuit for piwi by the large Maf gene traffic jam in Drosophila." Nature **461**(7268): 1296-1299.

Saito, K., H. Ishizu, et al. (2010). "Roles for the Yb body components Armitage and Yb in primary piRNA biogenesis in Drosophila." Genes Dev **24**(22): 2493-2498.

Saito, K., A. Ishizuka, et al. (2005). "Processing of pre-microRNAs by the Dicer-1-Loquacious complex in Drosophila cells." PLoS Biol **3**(7): e235.

Saito, K., Y. Sakaguchi, et al. (2007). "Pimet, the Drosophila homolog of HEN1, mediates 2'-O-methylation of Piwi-interacting RNAs at their 3' ends." Genes Dev **21**(13): 1603-1608.

Sanderson, R. J., K. E. Bird, et al. (1976). "Design principles for a counterflow centrifugation cell separation chamber." Anal Biochem **71**(2): 615-622.

Schierling, B. and A. Pingoud (2012). "Controlling the DNA Cleavage Activity of Light-Inducible Chimeric Endonucleases by Bidirectional Photoactivation." Bioconjug Chem.

Schratt, G. M., F. Tuebing, et al. (2006). "A brain-specific microRNA regulates dendritic spine development." Nature **439**(7074): 283-289.

Schwarz, D. S., G. Hutvagner, et al. (2003). "Asymmetry in the assembly of the RNAi enzyme complex." Cell **115**(2): 199-208.

Shah, C. and K. Forstemann (2008). "Monitoring miRNA-mediated silencing in Drosophila melanogaster S2-cells." Biochim Biophys Acta **1779**(11): 766-772.

Silber, R., V. G. Malathi, et al. (1972). "Purification and properties of bacteriophage T4-induced RNA ligase." Proc Natl Acad Sci U S A **69**(10): 3009-3013.

Sorefan, K., H. Pais, et al. (2012). "Reducing ligation bias of small RNAs in libraries for next generation sequencing." Silence **3**(1): 4.

Stark, A., J. Brennecke, et al. (2003). "Identification of Drosophila MicroRNA targets." PLoS Biol **1**(3): E60.

Stefano, J. E. (1984). "Purified lupus antigen La recognizes an oligouridylate stretch common to the 3' termini of RNA polymerase III transcripts." Cell **36**(1): 145-154.

Tamura, M., C. Nashimoto, et al. (2003). "Intracellular mRNA cleavage by 3' tRNase under the direction of 2'-O-methyl RNA heptamers." Nucleic Acids Res **31**(15): 4354-4360.

Teleman, A. A., S. Maitra, et al. (2006). "Drosophila lacking microRNA miR-278 are defective in energy homeostasis." Genes Dev **20**(4): 417-422.

Teninges, D. and N. Plus (1972). "P virus of Drosophila melanogaster, as a new picornavirus." J Gen Virol **16**(1): 103-109.

Thompson, D. M., C. Lu, et al. (2008). "tRNA cleavage is a conserved response to oxidative stress in eukaryotes." RNA **14**(10): 2095-2103.

Thompson, D. M. and R. Parker (2009). "Stressing out over tRNA cleavage." Cell **138**(2): 215-219.

Tobey, R. A. and H. A. Crissman (1972). "Use of flow microfluorometry in detailed analysis of effects of chemical agents on cell cycle progression." Cancer Res **32**(12): 2726-2732.

Todeschini, A. L., L. Teysset, et al. (2010). "The epigenetic trans-silencing effect in Drosophila involves maternally-transmitted small RNAs whose production depends on the piRNA pathway and HP1." PLoS One **5**(6): e11032.

Tomari, Y., C. Matranga, et al. (2004). "A protein sensor for siRNA asymmetry." Science **306**(5700): 1377-1380.

Vagin, V. V., A. Sigova, et al. (2006). "A distinct small RNA pathway silences selfish genetic elements in the germline." Science **313**(5785): 320-324.

van Rij, R. P., M. C. Saleh, et al. (2006). "The RNA silencing endonuclease Argonaute 2 mediates specific antiviral immunity in Drosophila melanogaster." Genes Dev **20**(21): 2985-2995.

Verdel, A. and D. Moazed (2005). "RNAi-directed assembly of heterochromatin in fission yeast." FEBS Lett **579**(26): 5872-5878.

Vogel, W., W. Schempp, et al. (1978). "Comparison of thymidine, fluorodeoxyuridine, hydroxyurea, and methotrexate blocking at the G1/S phase transition of the cell cycle, studied by replication patterns." Hum Genet **45**(2): 193-198.

Volpe, T., V. Schramke, et al. (2003). "RNA interference is required for normal centromere function in fission yeast." Chromosome Res **11**(2): 137-146.

Volpe, T. A., C. Kidner, et al. (2002). "Regulation of heterochromatic silencing and histone H3 lysine-9 methylation by RNAi." Science **297**(5588): 1833-1837.

Walker, G. C., O. C. Uhlenbeck, et al. (1975). "T4-induced RNA ligase joins single-stranded oligoribonucleotides." Proc Natl Acad Sci U S A **72**(1): 122-126.

Wang, S. H. and S. C. Elgin (2011). "Drosophila Piwi functions downstream of piRNA production mediating a chromatin-based transposon silencing mechanism in female germ line." Proc Natl Acad Sci U S A **108**(52): 21164-21169.

Wei, W., Z. Ba, et al. (2012). "A role for small RNAs in DNA double-strand break repair." Cell **149**(1): 101-112.

Xia, T., J. SantaLucia, Jr., et al. (1998). "Thermodynamic parameters for an expanded nearest-neighbor model for formation of RNA duplexes with Watson-Crick base pairs." Biochemistry **37**(42): 14719-14735.

Yamasaki, S., P. Ivanov, et al. (2009). "Angiogenin cleaves tRNA and promotes stress-induced translational repression." J Cell Biol **185**(1): 35-42.

Yan, Z., H. Y. Hu, et al. (2011). "Widespread expression of piRNA-like molecules in somatic tissues." Nucleic Acids Res **39**(15): 6596-6607.

Yang, L., D. Chen, et al. (2007). "Argonaute 1 regulates the fate of germline stem cells in Drosophila." Development **134**(23): 4265-4272.

Ye, X., N. Huang, et al. (2011). "Structure of C3PO and mechanism of human RISC activation." Nat Struct Mol Biol **18**(6): 650-657.

Yi, R., Y. Qin, et al. (2003). "Exportin-5 mediates the nuclear export of pre-microRNAs and short hairpin RNAs." Genes Dev **17**(24): 3011-3016.

Yu, B., Z. Yang, et al. (2005). "Methylation as a crucial step in plant microRNA biogenesis." Science **307**(5711): 932-935.

Zambon, R. A., V. N. Vakharia, et al. (2006). "RNAi is an antiviral immune response against a dsRNA virus in Drosophila melanogaster." Cell Microbiol **8**(5): 880-889.

Zhang, Z., S. S. Chang, et al. (2013). "Homologous recombination as a mechanism to recognize repetitive DNA sequences in an RNAi pathway." Genes Dev **27**(2): 145-150.

Zhou, R., B. Czech, et al. (2009). "Processing of Drosophila endo-siRNAs depends on a specific Loquacious isoform." RNA **15**(10): 1886-1895.