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

Modeling Transposon Recognition in *Drosophila melanogaster*

Anna Katharina Elmer
Aus Freising, Deutschland

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, 12.03.2013

Anna Katharina Elmer

Dissertation eingereicht am 18.04.2013

1. Gutachter: Prof. Dr. Klaus Förstemann
2. Gutachterin: Dr. Katja Stässer

Mündliche Prüfung am 16.05.2013

CONTENTS

1 SUMMARY

In fly somatic cells deleterious invasion of transposons is repressed by endo-siRNAs via a post-transcriptional mechanism. Endo-siRNAs are 21nt long RNAs which derive from a double-stranded precursor and show perfect complementarity to their target mRNA. It is not clear how these dsRNA precursors, especially the antisense strands thereof, are produced and how their production is regulated. Stable integration of a reporter gene at high copy number can mimic the multiple occurrences of transposons. It was previously shown that the insertion of such artificial transgenes can lead to an endo-siRNA response, implicating that there is no need for a pre-existing pool of small RNAs, pseudogenes or related sequences in the master control loci as a template for their production. The source for endo-siRNA precursors rather seems to be low level antisense transcription that may be proportional to copy number.

Our data substantiate the copy number dependent increase of transgene targeting endo-siRNAs and support the existence of a threshold level for efficient transgene silencing.

Histone genes are encoded in repetitive clusters with up to 100 copies per haploid genome. However, they are only subject to a low-level endo-siRNA response. Although histones are transcribed by RNA polymerase II, their mRNA is not polyadenylated but ends with a characteristic hairpin, which might have a protective effect. We examined reporter clones carrying these hairpin structures instead of a canonical poly A signal regarding their transgene silencing capacity in cell culture and observed diminished targeting of these foreign genetic elements. At the same time we discovered an accumulation of siRNAs targeting an intron in close proximity to the transcription start side in several deep sequencing libraries of reporters with histone 3'UTR. These intron derived siRNAs depend on the absence of a canonical poly A signal, are loaded into Argonaute 2 and presumably mature from the debranched intron lariat. They could potentially play a role in the discrimination of unadjusted splice sites of horizontally transferred transposons or target partially intronized transposable elements.

2 INTRODUCTION

2.1 CLASSES OF SMALL RNAS

The degree of organismal complexity cannot solely be deduced from its sheer number of protein coding genes, as with increasing complexity non-protein coding sequences dominate the genome of metazoans (Amaral and Mattick, 2008) and alternative splicing also elevates complexity of genome interpretation. Furthermore increasing complexity relies on additional layers of fine control ensuring an increased diversity of gene regulation. In differentiated cells, transcriptional control of protein coding genes still is the major mechanism regulating gene expression. However, as we become more and more aware of the bulk of non-protein-coding RNAs posttranscriptional gene regulation moves to the center of attention.

Small RNAs already hold a great share of these post-transcriptional regulation mechanisms with further roles frequently being discovered (Chung et al., 2008; Fagegaltier et al., 2009; Francia et al., 2012; Lim et al., 2011; Michalik et al., 2012; Wei et al., 2012). The first species identified over a decade ago were micro RNAs (miRNAs) and exogenous small interfering RNAs (siRNAs) (Fire et al., 1998; Lee et al., 1993). The former derive from endogenously encoded usually RNA Polymerase II transcribed RNA hairpin precursors (pri-miRNAs) (Lee et al., 2004a). In flies, these pri-miRNAs are sequentially processed, first in the nucleus by the RNaseIII enzyme Drosha together with its RNA Binding Protein partner Pasha, then by the RNaseIII enzyme Dicer-1 (Dcr-1) which acts in concert with the isoform B of Loquacious (Loqs) in the cytosol (Lee et al., 2002; Saito et al., 2005). The resulting 21 to 24 nucleotide (nt) long duplexes are loaded predominantly into Argonaute1 (Ago1) (Okamura et al., 2004) and one of the strands, the so-called passenger strand, is expelled from the complex (Schwarz et al., 2003). Ago1 is a member of the Argonaute/PIWI protein family, but lacks multiple turnover endonucleolytic activity in its PIWI domain (Cox et al., 1998). It is the effector enzyme conferring translational silencing and mRNA destabilization to the miRNA target, which is recognized via sequence complementarity. Mismatches and bulges at the 3' end of the miRNA ensure passenger strand displacement and at the same time prevent target cleavage if the duplex is loaded into Argonaute2 (Ago2) bearing an RNA cleaving PIWI domain. miRNAs usually do not basepair with their targets throughout their entire length but rather depend on their seed region at the 5'end. Restriction of target specificity to the 6-8 nt long seed region moreover broadens the target spectrum of a single miRNA tremendously. In contrast to the regulatory function role of endogenous miRNAs, exogenous siRNAs (exo-siRNAs) usually protect their host from viral infection. Their precursors are long perfectly basepaired double stranded RNAs (dsRNA) from exogenous sources. In flies they are processed by the cytosolic RNaseIII protein Dicer-2 and its RNA binding protein partner R2D2 into 21 nt long perfectly matching duplexes (Lee et al., 2004b; Liu et al., 2003). These duplexes are loaded into the catalytically active effector enzyme Ago2 by the RNA induced silencing complex loading complex (RLC), which takes over the guide strand decision during loading (Tomari et al., 2004). The passenger strand is cleaved by the PIWI

domain of Ago2 and expelled; the remaining guide strand directs the RNA induced silencing complex (RISC) to its target, the viral mRNA.

As Ago2 bears an enzymatic activity to cleave its target and efficiently dissociates from the reaction products, siRNA silencing is a multiple-turnover mechanism (Hutvagner and Zamore, 2002). The extensive complementarity to its target, needed to execute its primary function, ensures great specificity. Usually several siRNAs against the same target are derived from a common precursor in a processive manner, in turn leading to a distribution of several points of attack on one mRNA.

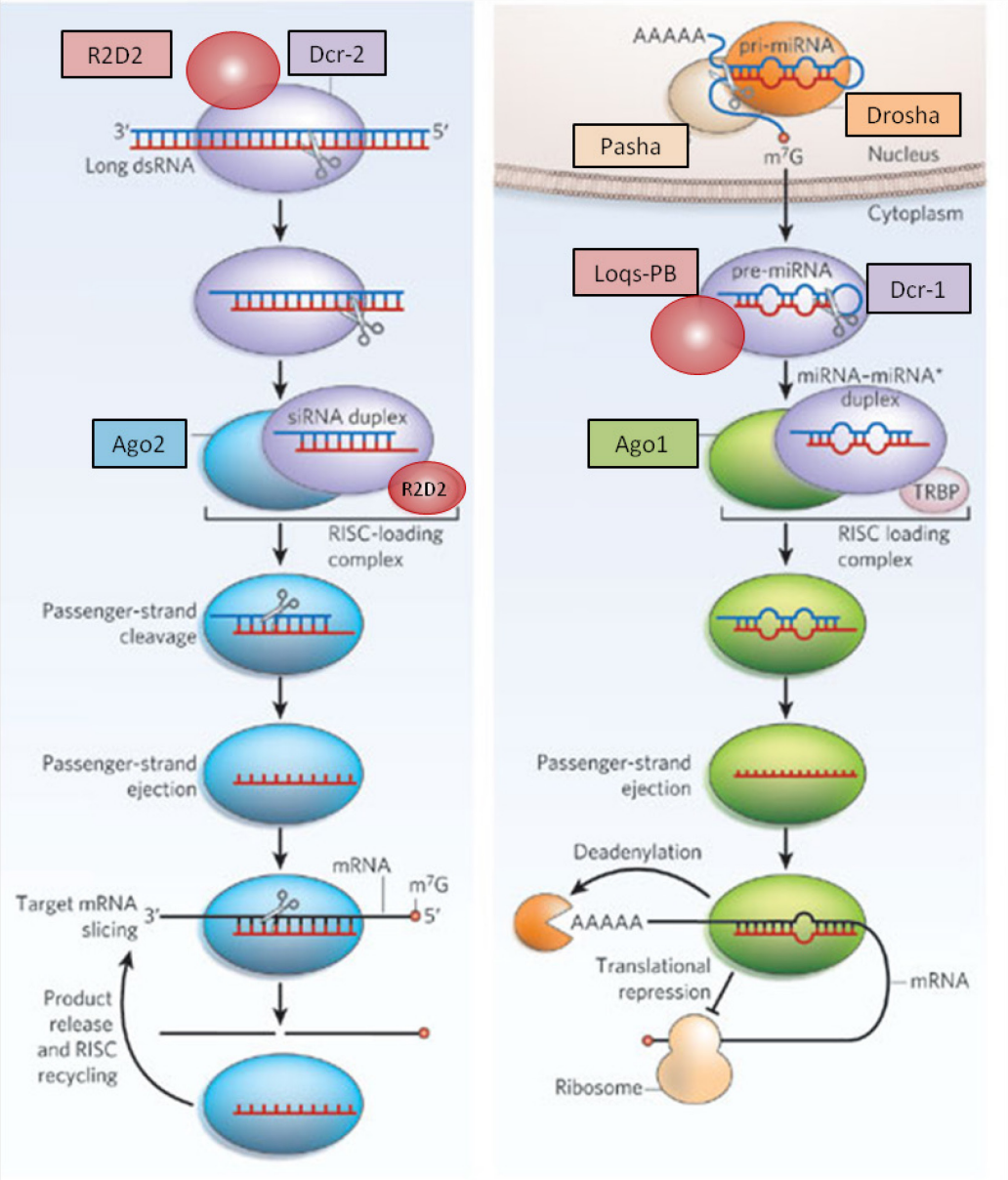


Figure 1 **Biogenesis pathways of miRNA and siRNAs** (adapted from Jinek and Doudna, 2009) (Left) siRNA precursors are long perfectly basepairing dsRNAs which are bound and cleaved into 21nt long duplexes by Dcr-2 and R2D2 in the cytosol. The siRNA duplex is loaded into Ago2 by the RLC composed of Dcr-2 and R2D2 and the passenger strand is cleaved by Ago2. Upon ejection the single strandes siRNA guide strand leads Ago2 to its target mRNA, which is silenced by mRNA cleavage, Ago2 can be recycled afterwards. (Right) miRNA precursors are polyadenylated and 5'capped Pol II transcripts forming 70-90nt long imperfectly matching hairpins. These hairpins are recognized by Drosha and Pasha, which cleave the single stranded 3' and 5' overhangs. The resulting pre-miRNA is, after export from the nucleus, bound by Dcr-1 and Loqs-PB and the loop is cleaved in a 21-24nt distance to the RNA ends to give a miRNA/miRNA* duplex. This duplex is loaded into Ago1, miRNA* is ejected and the miRNA directs Ago1 to its target mRNA. Silencing is accomplished by translational repression and recruitment of deadenylases which lead to mRNA degradation by the exosome.

A third class of small RNAs in Drosophila, which is predominantly expressed in germline cells are piRNAs. piRNAs are processed from long single stranded transcripts of piRNA clusters. These primary piRNAs are bound by Aubergine a member of the PIWI clade of the Argonaute/PIWI family of proteins. In an amplification cycle, termed the ping-pong cycle, the PIWI protein Argonaute3 (Ago3), generates 24-29 nt long secondary piRNAs in antisense and sense orientation to their target, which direct silencing of mobile genetic elements (Brennecke et al., 2008; Saito et al., 2006). As the ping-pong cycle involves transposon sense transcripts; the generation of secondary piRNAs at the same time reconstitutes a slicer dependent post-transcriptional silencing mechanism. This pathway is conserved in germline cells of sexually reproducing organisms to prevent genome disruption by transposition of mobile genetic elements.

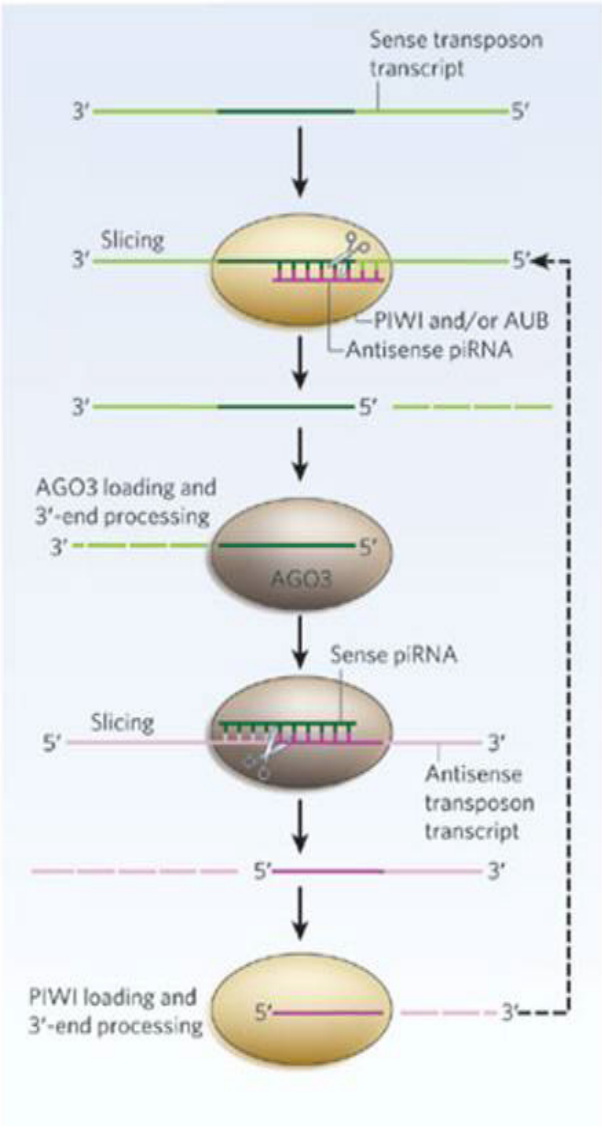


Figure 2 **Secondary piRNA biogenesis in the ping pong cycle.** (adapted from Jinek and Doudna, 2009) Secondary piRNAs derive from long single stranded precursors, for example transposon transcripts. The RNA is sliced by PIWI or Aubergine with the help of a primary piRNA to give the 5' end of the secondary piRNA. Upon loading into Ag3 the 3' end is trimmed and the mature 24-28nt long piRNA can target Ago3 to an antisense transposon transcript. Ago3 claves the RNA at position 10 and a new 5' end of a piRNA forms, which is in turn loaded into PIWI or Aubergine to slice a new sense transcript.

2.2 THE RISC COMPLEX

All minimal RISCs, mi-, pi-, and si-RISC are comprised of an Argonaute effector enzyme and a small RNA guiding the complex to its target sequence. This core-RISC is already competent in target suppression but holo-RISC can be as large as 80S (Pham, 2004). A heterogeneous set of proteins was observed to be associated with the Drosophila core-RISC. As such, the *glycine and tryptophan repeats containing 182kDa protein* (GW182) associates with Ago1 in flies and is indispensible from efficient gene silencing (Eystathioy et al., 2002; Rehwinkel et al., 2005), Other proteins often co-purified with the Ago1-RISC are *fragile X mental retardation analog* (dFXR), a RNA binding protein involved in neuronal growth and branching and the *Vasa intronic gene* (VIG) which is a target of *protein kinase C* (PKC) (Caudy et al., 2002). Furthermore the *Tudor Staphylococcal Nuclease homolog* (TudorSN), a nuclease cleaving DNA and RNA substrates (Caudy et al., 2003) and in oocytes Spindle E, a helicase like structured protein involved in RNA localization, were immunoprecipitated with RISC (Kennerdell, 2002). The exact function and stoichiometry of most of these interacting proteins remain unclear.

2.3 SMALL RNAS INVOLVED IN CELL DIVISION AND GROWTH CONTROL

Among targets silenced by the miRNA pathway, regulators of cell differentiation, development, growth control and metabolism are highly enriched (reviewed in Bartel, 2009). The first miRNAs discovered, lin-4 and let-7 both are heterochronic genes and regulate cell fate decisions in *C. elegans* embryonic stem cells (Lee et al., 1993; Pasquinelli et al., 2000). Several lines of evidence indicate that miRNAs show abnormal expression levels in malignantly transformed cells. Let-7 was also implicated in regulation of the nematodal let-60/ras protein (Johnson et al., 2005), which is highly conserved across species and a down-regulation of let-7 was observed in human lung cancer and melanoma cells (Schultz et al., 2008; Takamizawa et al., 2004) The miRNA *bantam* was shown to accelerate the cell cycle (Brennecke et al., 2003) and several oncogenes are de-repressed in human carcinoma cells by down-regulation of miRNAs (Zhang et al., 2007). Besides oncogenes, tumor suppressors can also be directly regulated by miRNAs, as for example p21, p27, p57 and PLK2 (Galardi et al., 2011; Li et al., 2008; Yu et al., 2009).

2.4 INFLUENCE OF THE CELL CYCLE ON RISC AND AIMS OF PART I OF THE THESIS

The aforementtioned observations indicate a direct role of the miRNA pathway in cell cycle control, but it is also conceivable that their expression profile might be governed by cell cycle progression. Since the half life of small RNAs was shown to be quite long in general (Bail et al., 2010; van Rooij et al., 2007) and the average cell cycle of the Drosophila Schneider cell line only takes 22 hours, a certain form of regulation fine-tuning would be conceivable at the level of RISC composition or modification. Such fine regulation might be mediated by a modifying factor, which temporarily prevents the target bound RISC from inhibiting the translation of a target mRNA. A modifying factor might as well bind directly to the target mRNA to protect it from the translation inhibiting signal transduction through Ago, or it might reverse silencing into a translation activating signal. In the first part of this thesis I tried to identify such RISC modifying factors in a screen, employing stable isotope labeling, cell cycle phase separation, RISC co-immunopurification and mass spectrometry.

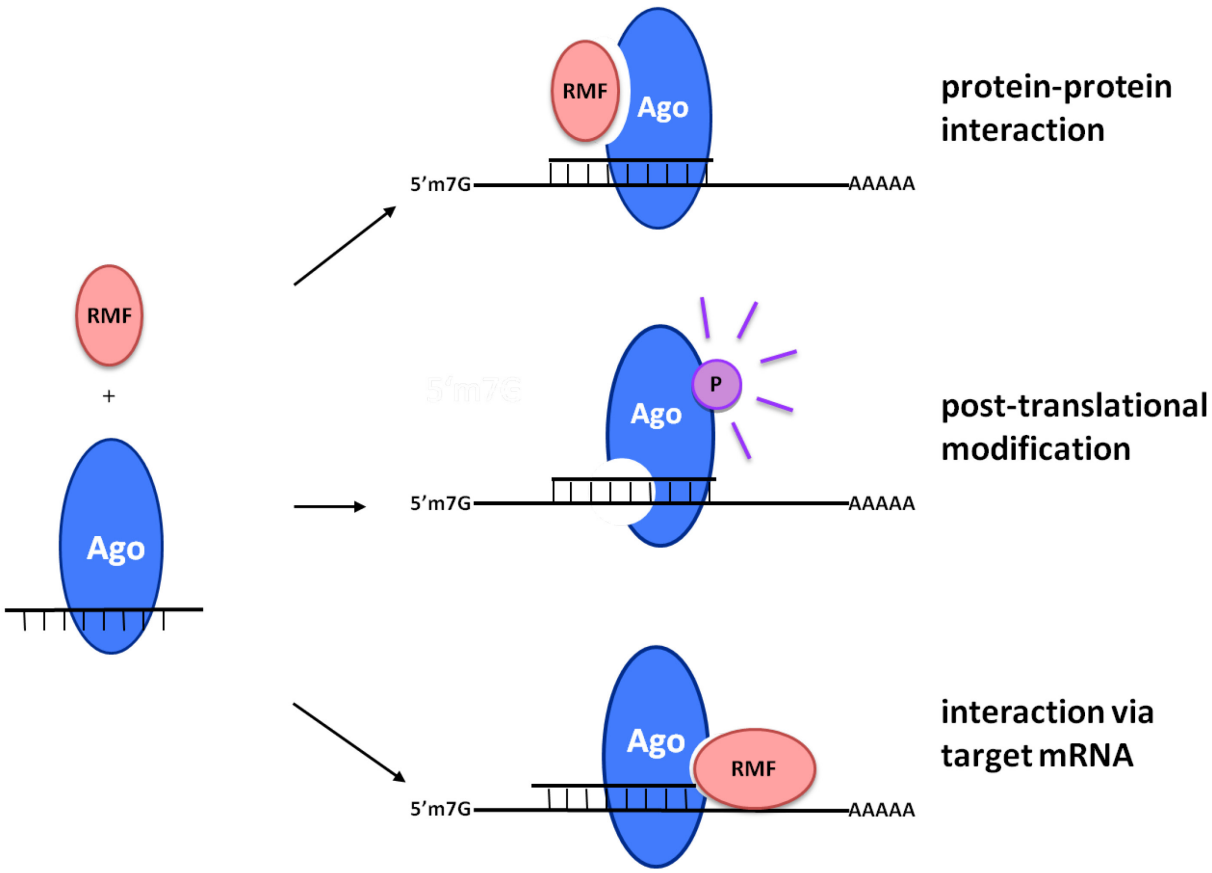


Figure 3 **Possible RISC modifications.** A RISC modifying Factor (RMF) could influence the target spectrum or silencing competence of Ago by direct protein-protein interaction, or could modify Ago with a posttranscriptional modification as for example phosphorylation. RMF could also interact with Ago via a common RNA to modify its silencing competence target specifically.

2.5 SELFISH GENETIC ELEMENTS

Selfish genetic elements , also called transposable elements (TEs) can be found in almost all organisms (Kaminker et al., 2002), typically integrated in the host's genome in large copy numbers. They reproduce by coding for replicative functions, such as reverse transcriptase or transposase, which they inherit from a common ancestor with retroviruses. These tools enable them to increase their number in the following generation by replication and random integration in a different locus (Werren, 2011). Transposition of mobile genetic elements can entail destructive as well as constructive consequences. Insertion into a gene can lead to loss of a certain gene function, modification of expression pattern or to rearrangement due to homologous recombination. At the same time withdrawing of transposable elements from the genome would lead to a reduction of 15 % of the genetic material (Kaminker et al., 2002), which in the course of evolution took a strategic role in genome regulation and maintenance. In Drosophila for example, telomers are constituted of three non-long terminal repeat retrotransposons, HeT-A, TAHRE, and TART, enabling Drosophila to reproduce without the otherwise essential protein telomerase.

Due to their potentially gene-disrupting activities, it is not surprising that TEs represent a serious threat to the integrity of a cell's genome. Especially germline cells need reliable mechanisms to inhibit transposon mobilization where the PIWI associated class of small RNAs mentioned before play a leading role in this defense in Drosophila and mice (reviewed in Ishizu et al., 2012) However, silencing of TEs must also be ensured in somatic cells, not only to prevent their decline but also because TEs can be transferred horizontally (Bartolome et al., 2009), thus endangering germline cells by soma derived transposons. Mammals solved this problem by establishing long lasting heterochromatin around transposon coding sites during embryogenesis, in a mechanism involving piRNAs (Kanellopoulou et al., 2005; Le Thomas et al., 2013) or endogenous siRNAs (Watanabe et al., 2008) to ensure transcriptional silencing. In Drosophila somatic cells transposons are silenced in a posttranscriptional mechanism by endo-siRNAs.

2.6 ENDOGENOUS-SIRNAS

The class of small RNAs the second part of this work focuses on is the class of endogenous-siRNAs. Endo-siRNAs derive from endogenous sources but their mechanism of action resembles exo-siRNAs. After export into the cytosol, long perfectly matching RNA precursor duplexes are processed by Dcr-2, but in this case Dcr-2 works together with isoform D of Loqs (Hartig et al., 2009; Zhou et al., 2009). It is not clear whether the Dcr-2/Loqs-PD complex is able to load the siRNA into Ago2 or if R2D2 participates in endo-siRISC loading as it does in exo-siRISC loading. As soon as the endo-siRNA is loaded into Ago2 the passenger strand is cleaved and the RISC becomes active. Endo-siRNAs in *Drosophila* mainly target transposable elements and therefore are regarded as "the guardian of the somatic genome" (Chung et al., 2008; Ghildiyal et al., 2008; Kawamura et al., 2008). As with exo-siRNAs, the main mechanism of silencing is mRNA cleavage and perfect complementarity of the guide strand to its target is needed. While the fate

of small RNAs in the cytoplasm is becoming more widely understood, the production of the initial precursors is still unclear. First studies on the origin of endo-siRNAs revealed that 86% of 21 nt and methylated RNAs map to transposon coding sites (Ghildiyal et al., 2008) exhibiting equal numbers of sense and antisense matching sequence and a peak at 21 nt length (Okamura et al., 2008a). A prominent amount of endo-siRNAs found in Drosophila mapped to piRNA clusters and two structured loci (CG18854 and GC4068) (Okamura et al., 2008a) and convergent transcripts from loci overlapping at their 3' end can form siRNA duplexes in a Dcr-2 dependent mechanism (Okamura et al., 2008b). Nevertheless, the exact mechanism leading to a perfect complementary double stranded RNA duplex arising from a transposon coding site and designated for further processed in the cytosol is still unknown.

In *C. elegans*, plants and yeast a RNA dependent RNA polymerase (RdRP) can synthesize a complementary strand using mRNA as template (Han et al., 2009; Martienssen et al., 2005). As flies lack a functional RdRP, they must depend on other mechanisms for dsRNA formation. In mouse oozytes endo-siRNA precursors can derive from pseudogenes (Tam et al., 2008), but in flies pseudogenes are rare (Harrison et al., 2003) and are no prerequisite for endo-siRNA silencing. Thus, the only self-evident explanation for endo-siRNA precursor production is antisense transcription directly from genomic DNA sequence. There is ample evidence that in different eukaryotic organisms antisense transcription widely occurs (Sun et al., 2006; Yelin et al., 2003), its regulation and control, however, are not clear. The intriguing selectivity and the absence of endo-siRNA silencing of important genomic material make this system a very elegant way of taming the amplification of selfish genetic elements.

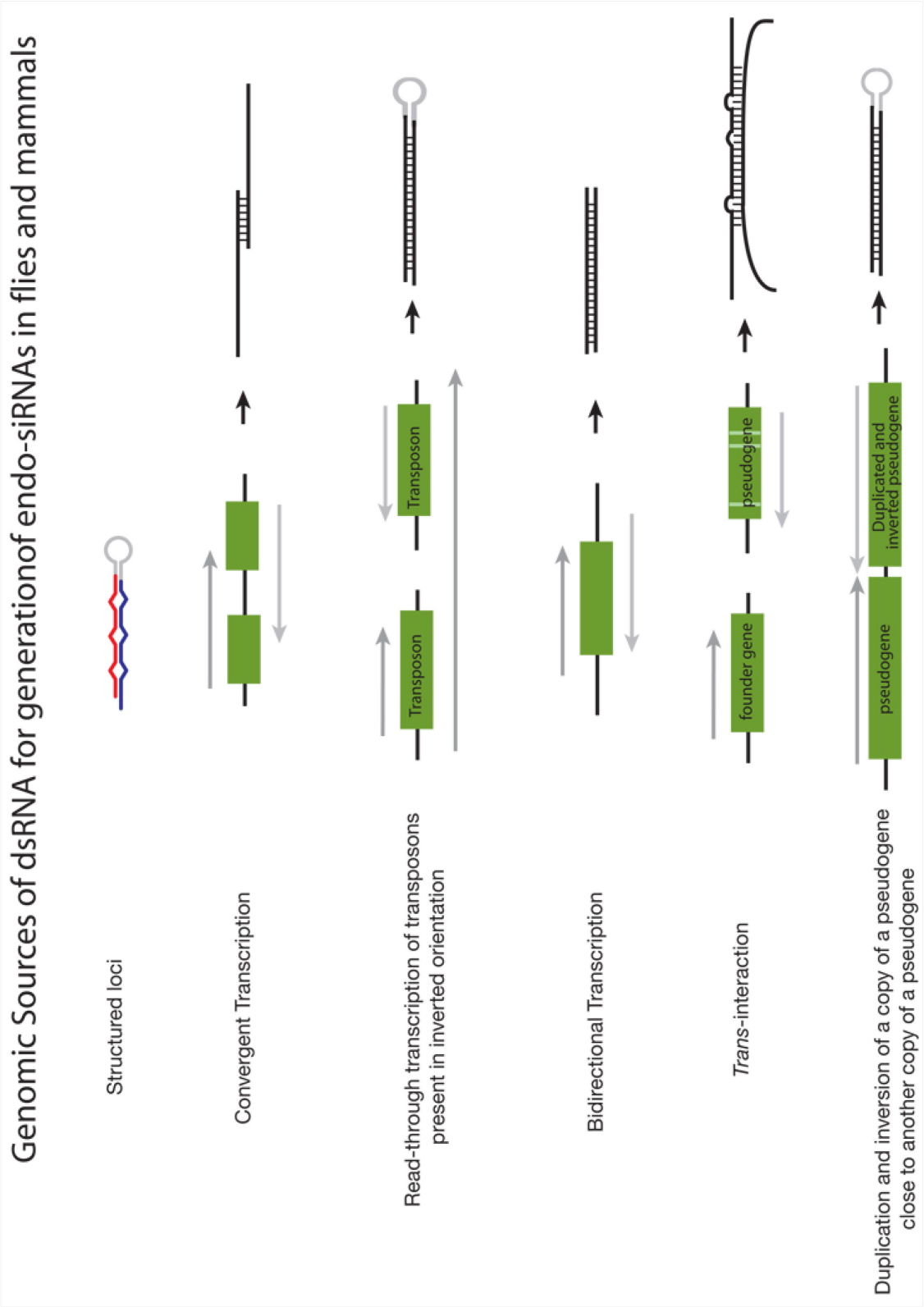


Figure 4 **Generation of endo-siRNA precursors in flies and mammals** (Ghildiyal and Zamore, 2009) Structured loci like CG4068 or CG18854 in flies can generate endo-siRNA. Endo-siRNA can also derive from annealing complementary transcripts (convergent or bidirectional transcription) or cis-NATs, another possibility for double strand formation as a prerequisite for endo-siRNA production might be pairing of a protein coding gene with a pseudogene or inverted repeat structures within pseudogenes.

2.7 MODELING ENDO-SIRNA REPRESSION AND AIM OF PART II OF THE THESIS

Several approaches in the past aimed to reveal mechanisms involved in transposon or transgene suppression. In 1997 Pal-Bhandra and colleagues observed a copy number dependent transgene suppression upon introduction of additional copies of a transgene in *Drosophila*. Thereby they could also observe a trans-effect on the endogenous gene, which they called co-suppression (Pal-Bhadra et al., 1997). A copy number dependent repression mechanism was also observed controlling the transposable element I-factor in Drosophila cells. Here the 5'UTR of the element was held responsible for the suppression but a RNA dependent mechanism was dismissed. In this study a threshold level for repression at 15 copies was suggested (Chaboissier et al., 1998). The Siomi lab used stably transfected multicopy GFP reporter constructs to analyze the involvement of Ago1 and Ago2 in transgene silencing. They could detect antisense transcripts complementary to GFP in their stably transfected cells and proclaimed that Ago2 depletion led to a shortening of the poly A tail resulting in mRNA stabilization in stably transfected cells but not in transiently transfected cells (Siomi et al., 2005).

We in our lab could previously observe that stable integration of a transgene in high copy numbers into the genome of somatic Drosophila Schneider 2 cells (S2) leads to an endo-siRNA response which can be abrogated by RNAi either against the endo-siRNA biogenesis components Dcr-2 and Loqs-PD or against the effector protein Ago2 (Forstemann et al., 2007; Hartig et al., 2009; Siomi et al., 2005). This concept was successfully used before as a reporter system to analyze proteins involved in processing, sorting and loading of small RNAs into different RNAi pathways but can also serve to investigate requirements for endo-siRNA production further upstream in the biogenesis pathway. In this thesis I aim to show whether the number of endo-siRNAs against artificial transgenes depends on their copy number and whether a threshold level might exist for silencing. Furthermore I want to investigate if silencing is influenced by the exchange of a canonical poly A signal to a histone stem loop.

Histone genes are a rare example of genes encoded repetitively in the genome without being excessively targeted by endo-siRNAs. Instead of a poly A signal, as all other metazoan RNA Pol II transcribed genes have, histone genes encode a histone stem loop structure within their 3' UTR. This structure consists of a highly conserved 6 nt long stem, a 4 nt long loop and a distinct purine rich histone downstream element (HDE) of 10-20 nt, recognized by the U7 snRNP and cleaved by the same protein components involved in polyadenylation, CstF and CPSF73 (Hentschel 1981; Townley-Tilson; 2006). Replication dependent transcription of histone genes is mediated by the stem loop binding protein (SLBP), which binds to the histone specific 3' end of the mature transcript (Wang 1996).

We hypothesized that this special 3' end formation might be involved in protecting histone genes from endo-siRNA mediated repression. To examine the effects of the histone stem loop sequence on transgene expression I exchanged the 3'UTR of our reporter construct with the Histone H2a or histone H3 3' UTR.

3. MATERIALS AND METHODS

3.1 MATERIALS

3.1.1 LABORATORY EQUIPMENT

Agarose gel running chamber	Carl Roth GmbH; Karlsruhe, Germany
Biometra Professional Thermocycler	Biometra; Jena, Germany
Biophotometer	Eppendorf AG; Hamburg, Germany
Branson Sonifier 250	Heinemann Ultraschall Labortechnik; Schwäbisch-Gmünd, Germany
Capillary Blotters MAXI	Carl Roth GmbH; Karlsruhe, Germany
Centrifuge Rotanta 460R	Andreas Hettich GmbG, Tuttlingen, Germany
Centro LB 960 Microplate Luminometer	Berthold Technologies; Bad Wildbad, Germany
FACSCalibur flow cytometer,	Becton, Dickinson; Franklin Lakes, USA
INTAS UV Imaging System,	INTAS; Göttingen, Germany
LAS 3000 mini Western Imager	Fujifilm; Tokyo, Japan
Leica TCS SP2 confocal microscope	Leica Microsystems; Wetzlar, Germany
Magnetic Stand	Kisker Biotech GmbH; Steinfurt, Germany
PAGE-electrophoresis	BioRad; Hercules, USA
Poly-lysine coated microscope slides	Carl Roth GmbH; Karlsruhe, Germany
Power supply	BioRad; Hercules, USA
Semi-dry blotter	BioRad; Hercules, USA
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
Typhoon 9400 Variable Mode Imager	GE Healthcare; Freiburg, Germany
Vortex Genie 2	Scientific Industries, Bohemia, USA
Water Bath	GFL, Burgwedel, Germany

3.1.2 ANALYSIS SOFTWARE

ApE plasmid Editor
BD Cell Quest Becton, Dickinson; Franklin Lakes, USA
BioEdit (see Hall, 1999)
BOWTIE
Multi Gauge V3.0 Fujifilm; Tokyo, Japan
PERL (see Hartig et al. 2009)

3.1.3 LABORATORY CHEMICALS

Acetic acid	Carl Roth GmbH; Karlsruhe, Germany
Acrylamide 40%	Carl Roth GmbH; Karlsruhe, Germany
Agarose	Biozym Scientific GmbH; Oldendorf, Germany
Ammonium peroxodisulfate (APS)	Carl Roth GmbH; Karlsruhe, Germany
Ampicillin	Carl Roth GmbH; Karlsruhe, Germany
Bacto Agar	Becton Dickinson; Franklin Lakes, USA
Boric acid	Sigma Aldrich; Taufkirchen, Germany
Bovine serum albumin (BSA)	New England Biolabs; Ipswich, USA
Bradford Assay	BioRad; Hercules, USA
Chloroform	Merck Biosciences GmbH; Schwalbach, Germany
Complete® without EDTA	Roche Diagnostics; Mannheim, Germany
Coomassie G250	Carl Roth GmbH; Karlsruhe, Germany
Desoxyribonucleotides	Sigma Aldrich; Taufkirchen, Germany
Dextransulfate	Carl Roth GmbH; Karlsruhe, Germany
Dimethyl sulfoxide (DMSO)	Carl Roth GmbH; Karlsruhe, Germany
Dimethylpimelimidate (DMP)	Carl Roth GmbH; Karlsruhe, Germany
Dithiothreitol (DTT)	Carl Roth GmbH; Karlsruhe, Germany
Ethanol (p.a.)	Carl Roth GmbH; Karlsruhe, 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	Carl Roth GmbH; Karlsruhe, Germany
Fugene® HD transfection reagent	Roche Diagnostics GmbH; Mannheim, Germany
G418 sulphate (neomycin)	PAA, The Cell Culture Company; Cölbe, Germany
H_2O HPLC quality	Carl Roth GmbH; Karlsruhe, Germany
Hepes	Carl Roth GmbH; Karlsruhe, Germany
Hygromycin B	Carl Roth GmbH; Karlsruhe, Germany
Isopropanol (p.a.)	Merck Biosciences GmbH; Schwalbach, Germany
Kanamycin	Carl Roth GmbH; Karlsruhe, Germany
L-Arginine:HCl: $^{13}C_6$, $^{15}N_4$	Euroisotope GmbH; Saint-Aubin Cedex, France
L-Lysine:2HCl: $^{13}C_6$, $^{15}N_2$	Euroisotope GmbH; Saint-Aubin Cedex, France
Methanol (p.a.)	Merck Biosciences GmbH; Schwalbach, Germany
Powdered milk	Rapilait Migros; Zürich, Switzerland
RiboLock RNase Inhibitor	Fermentas; St. Leon-Rot, German
Roti®Aqua Phenol/C/I	Carl Roth GmbH; Karlsruhe, Germany
Roti-liquid barrier marker	Carl Roth GmbH; Karlsruhe, Germany
Saponin	Fluka BioCemika; Ulm, Germany
Sequagel Sequencing System	National Diagnostics; Atlanta, USA
Sodium dodecyl sulphate (SDS)	Merck Biosciences GmbH; Schwalbach, Germany
Sodium periodate	Sigma Aldrich; Taufkirchen, Germany
Sodium tetraborate	Sigma Aldrich; Taufkirchen, 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
[α32P] CTP (SRP 505) 10 mCi/ml; 6000 Ci/mmol; 250 µCi	Hartmann Analytic; Braunschweig, Germany
[γ 32P] ATP (SRP 501) 10 mCi/ml; 6000 Ci/mmol; 250 µCi	Hartmann Analytic; Braunschweig, Germany

3.1.4 ENZYMES

DNAse I, RNAse free	Fermentas; St. Leon-Rot, Germany
Pfu DNA Polymerase	Fermentas; St. Leon-Rot, Germany
Phusion Hot Start DNA Polymerase	Finnzymes via Thermo Scientific, Waltham, USA
Polynucleotidekinase with Buffers	Fermentas; St. Leon-Rot, Germany
Proteinase K	Fermentas; St. Leon-Rot, Germany
Restriction enzymes	New England Biolabs; Ipswich, USA
RNaseA	Fermentas; St. Leon-Rot, Germany
Superscript II, Reverse Transcriptase	Invitrogen; Karlsruhe, Germany
T4-DNA Ligase	New England Biolabs; Ipswich, USA
T7-polymerase	laboratory stock
Taq DNA Polymerase	laboratory stock

3.1.5 TEST KITS AND OTHER MATERIAL

50 bp DNA Ladder	New England Biolabs; Ipswich, USA
Blotting paper	Machery-Nagel; Düren, Germany
Cell culture dishes and plates	Sarstedt; Nümbrecht, Germany
CloneJet PCR Cloning Kit	Fermentas; St. Leon-Roth, Germany
Cryovials	Biozym Scientific GmbH; Oldendorf, Germany
Dynabeads Protein G	Invitrogen Dynal AS, Oslo, Norway
DyNAmo Flash SYBR Green, qPCR Kit	Finnzymes; via Thermo Scientific, Waltham, USA
High Prime random labeling kit	Roche Diagnostics GmbH; Mannheim, Germany Isopropanol
freezing container	Nalgene via Thermo Fisher Scientific; Waltham, USA
microRNA Marker	New England Biolabs; Ipswich, USA
Nylon membrane, positively charged	Roche Diagnostics GmbH; Indianapolis, USA
Phosphoimager Screens	FujiFilm; Tokio, Japan
Polyacrylamide gradient gel	Pierce Thermo Fisher, Rockford, USA
Polyvinylidenfluoride membrane	Milipore; Billerica, USA
Prestained Protein Ladder	Fermentas; St. Leon-Rot, Germany
Promega Dual Luciferase Assay System	Promega; Madison, USA

Reliaprep gDNA Miniprep System	Promega; Madison, USA
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
RestoreTM WB Stripping Buffer	Thermo Fisher Scientific; Waltham, USA
Sephadex spin column (G25)	Roche Diagnostics GmbH; Mannheim, Germany
Spin column (empty, for IP)	MoBiTec; Göttingen, Germany
SuperSignal West Dura	Thermo Fisher Scientific; Waltham, USA
Whatman 595 ½ Folded Filters	Whatman GmbH; Dassel, Germany

3.1.6 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
Borax buffer	148 mM borax 148 mM boric acid pH 8.6
Buffer A for genomic DNA extraction	100 mM Tris/HCl, pH 7.5 100 mM EDTA 100 mM NaCl 0.5% SDS
Church buffer	1% (w/v) bovine serum albumine (BSA) 1 mM EDTA 0.5 M phosphate buffer, pH 7.2 7% (w/v) SDS
Colloidal Coomassie staining solution	50 g/l aluminum sulfate 2% (v/v) H3PO4 (conc.) 10% (v/v) ethanol 0.5% (v/v) Coomassie G250 stock
DMP/borate solution	0.1M Sodium Borate 20mM DMP
DNA loading buffer (6x)	0.25% (w/v) bromophenol blue 0.25% (w/v) xylene cyanol 30% (w/v) glycerol

Elutriation buffer	1xPBS 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 2 mM MgCl2 1 mM DTT
Probe hybridization solution for FISH	200 µl Formamide 200 Ml 50% Dextransulfat 40 µl 20xSSC 5 µl 100 mM random hexamers 10 µl BSA (Fermentas) 5µl 10 mM labeled oligonucleotide
SDS-running buffer (5x)	125 mM Tris/HCl, pH 7.5 1.25 M glycine 5% SDS

SILAC medium was prepared according to Bonaldi et al, 2008
Dialyzed Fetal Bovine Serum was purchased at Invitrogen, via Life Technologies, Carlsbad, USA;
Isotope labeled aminoacids were purchased at Euroisotope GmbH, Saint-Aubin Cedex, France

Solexa elution buffer	0.4% NaCl 0.5% SDS 50mM Tris-HCl, pH 8

Southern Blotting Buffer	0.4 M NaOH 1M NaCl
Southern Blot neutralizing buffer	1 M Tris pH 7.5 0.5 M NaCl
SSC (20x)	3 M NaCl 0.3 M sodium citrate
TAE (50x)	2 M Tris-base 5.71% acetic acid (0,9 M) 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
Western blotting stock (10x)	250 mM Tris/HCl, pH 7.5 1.92 M glycine
Western blotting buffer (1x)	10% Western blotting stock (10x) 20% methanol

3.1.7 OLIGONUCLEOTIDES AND PLASMIDS

3.1.7.1 Oligonucleotides for molecuar cloning and sequencing

pKF63	5prime P sense: CGT GCA CTG AAT TTA AGT GT
	3prime P as: GCA CTT ATT GCA AGC ATA CG
	ubi5prime exon s: GTA GAA AGT AAA GCG CAA TCA GCG
	Ubi intron s: GTA AGT TTT TAA CTC GCT GTT ACC
	Ubi intron as: GCG GGC AGA AAA TAG AGA TG
pJB47	Ago2 prom s: CGTCTATTAAAAGTCGTCAG
	Ago2 50nt ext s: GTTTTGTATTTCCCAAATTGCG
	Ago2 seq s2: TAATCGTCACGAACTGGATGATGGATA
	Ago2 seq as1: GGCTCCAATGTACATGGTGTTCTTCAT
	Flag s: GATTACAAGGATGACGATGACAAG

pKE6	RM62PB Xba s:TTC TAGAATGGCACCACACGATCGCGAC
	Rm62 Notas: GAAAGCGGCCGCCTAGTCGAAGCGCGAGTGTCTGC
pKE7	ncd Xba s: TTCTAGAATGGAATCCCGGCTACCGAAAC
	ncd Not as: GAAAGCGGCCGCTTATTTATCGAAACTGCCGCTGTTGTTG
pKE8	Tub Prom s: CTATGCTGCTGGAACGCTTC
	Ago1 Bam s: CGG ATC CAT GTA TCC AGT TGG ACA A
	Ago1 Not as: CAC GAG CGG CCG CTT TAG GCA AAG TAC ATG AC
pKE9	Ago2 Not s:CACGAGCGGCCGCTATGGGAAAAAAAGATAAGAAC
	Ago2 Sal as: CCATGGTCGACTCAGACAAAGTACATGGGC
	A2_nostop_SacII: TCCGCGGGACAAAGTACATGGGGTTTTTCTTC
pKE13	Tomato Not s: GTATGCGGCCGCTATGGTGAGCAAGGGCGAG
	Tomato EagI as: TACGGCCGAGGCTTGTACAGCTCGTCCATGC
pKE14	H2a Not s: ATTAGCGGCCGCTACGTTTCAAAGGCTAAGCTAAAAACC
	H2a Eco as: TAGAATTCCATCTTTACGTTAATAATTCTTTATGATTTACTAATTACAAC
pKE15	H3Not s: ATTAGCGGCCGCTGACACGGCATTAACTTGC
	H3 Eco as: TAGAATTCGAGCTGACAATTAAAATAAGATTATTTTTATTCCTTCTC

3.1.7.2 Oligonucleotides for quantitative PCR

1731 fw (Ghildiyal et al., 2008)	CCCAAACAGGTGACCCATAC
1731 rev (Ghildiyal et al., 2008)	CACAACGTGACCCTCTTTCA
297 fw (Ghildiyal et al., 2008)	GGTGATCCAGAAACCCTTCA
297 rev (Ghildiyal et al., 2008)	CTTTCGATGGCTCCCAGTAG
CG5599 fw (Hartig et al., 2009)	CTCCCGGTACTAACGTTCCA
CG5599 rev (Hartig et al., 2009)	TTGCATCAACTGGGTCATGT
CG1673 fw (Hartig et al., 2009)	ATGAACATGAACCGCATGAA
CG1673 rev (Hartig et al., 2009)	GGCTGAGGATCGTGTAGAGC
GAPDH fw (Hartig et al., 2009)	CTTCTTCAGCGACACCCATT
GAPDH rev (Hartig et al., 2009)	ACCGAACTCGTTGTCGTACC
GFP fw	ACGTAAACGGCCACAAGTTC
GFP rev	AAGTCGTGCTGCTTCATGTG
Juan fw	CAATGGGTTGACAACATTCG
Juan rev	CAATGGGTTGACAACATTCG
roo fw (Ghildiyal et al., 2008)	CGTCTGCAATGTACTGGCTCT
roo rev (Ghildiyal et al., 2008)	CGGCACTCCACTAACTTCTCC
Ubi Promoter fw (Hartig et al., 2009)	GCCGGTAGAGAAGACAGTGC
Ubi Promoter rev (Hartig et al., 2009)	ACTGACTTGACCGGCTGAAT
Tub Prom fw	CTATGCTGCTGGAACGCTTC
Tub prom rev	TAGCTACCTCTCTCACTCGC

3.1.7.3 Oligonucleotides for dsRNA generation

T7_dsred_fw	CGTAATACGACTCACTATAGGAGGACGGCTGCTTCATCTAC
T7_dsred_rev	CGTAATACGACTCACTATAGGTGGTGTAGTCCTCGTTG
T7_GFP_fw	CGTAATACGACTCACTATAGGATGGTGAGCAAGGGCGAG
T7_GFP_rev	CGTAATACGACTCACTATAGGTTACTTGTACAGCTCGTCCA
T7_Ago2_fw	CGCACCATTGTGCATCCTAACGAG (Forstemann et al., 2007)
T7_Ago2_rev	GGGGACAATCGTTCGCTTTGCGTA (Forstemann et al., 2007)
T7_Dcr2_fw	CTGCCCATTTGCTCGACATCCCTCC
T7_Dcr2_rev	TTACAGAGGTCAAATCCA AGCTTG
T7_LoqsPD_fw	CGTAATACGACTCACTATGTGAGTATCATTCAAGACATC
T7_LoqsPD_rev	CGTAATACGACTCACTATAGGTAAGGTGTAAGCATTATGT

3.1.7.4 Fluorescently labeled DNA probes for FISH

5'Cy3-UbiProm1	CTAAAGTGTTACGAACACTACGGTA
5'Cy3-UbiProm2:	GGTTTCTCAACAAAGTTGGCGTCG
5'Cy3-GFP1:	GAAGAAGACTTGGGCATGGTG
5'Cy3-GFP2:	CGACGGCAACTACAAGACCC
5'Cy3-GFP3:	CAAAGACCCCAACGAGAAGCG
5'Cy3-H2a-CDS:	CTGGCCGCTGAGGTTCTCGAG
5'Cy3-H3-CDS:	GGTGTGAAGAAGCCCCACCGC

3.1.7.5 Oligonucleotides for Northern Blotting

bantam probe	AATCAGCTTTCAAAATGATCTCA
miR-277 probe	TGTCGTACCAGATAGTGCATTTA
intron66 probe	AAA CTGCATTTCAAGGTCTTTGTTCGGCCA
as intron66	CGA ACA AAGACCTTGAAATGC

3.1.7.6 Oligonucleotides for SOLEXA library generation

Adapter	
3' ligation (Modban)	AMP-pCTGTAGGCACCATCAATdideoxyC
5' ligation (Solexa linker)	rArCrArCrUrCrUrUrUrCrCrCrUrArCrArCrGrArCrGrC rUrCrUrUrCrCrGrArUrCrU
	Eurofins MWG – HPLC purified, 50 µM stock
Reverse transcription	
3' RT primer	ATTGATGGTGCCTACAG
	Eurofins MWG – HPSF purified, 5 µM stock
PCR	
5'-Solexa	AATGATACGGCGACCACCGAACACTCTTTCCCTACACGACG
3'-PCR BamHI	CAAGCAGAAGACGGCATACGAGGATCCGATTGATGGTGCCTACAG
3'-PCR Pvu	CAAGCAGAAGACGGCATACGACAGCTGGATTGATGGTGCCTACAG
3'-PCR Cla	CAAGCAGAAGACGGCATACGAATCGATGATTGATGGTGCCTACAG
3'-PCR Xba	CAAGCAGAAGACGGCATACGATCTAGAGATTGATGGTGCCTACAG
	All Eurofins MWG – HPSF purified, 10 µM stock

3.1.7.7 Other Oligonucleotides

Oligo dT (EcoRI T18)	ACGAATTCTTTTTTTTTTTTTTTT
Random hexamers	NNNNNN

3.1.7.8 Plasmids

pSHNeo	Neomycin resistence plasmid, (Steller and Pirrotta, 1985)	Amp
pSHHyg	``	Amp
pKF63	pCasper2 with GFP (Förstemann et al 2005; Hartig et al, 2009)	Amp
pC5T	pCasper5 with Tubulin Promoter and 3'UTR	Amp
pCMV-tdTomato	comercial Clontech#632534	Kan
pJB47	pCasper with Ago2 genomic region	Amp
pBSIIKS+ GFP	GFP in reverse orientation 5' BamHI - 3' SpeI	Amp
pKE5	pBSIIKS+ with Ago2	Amp
pKE6	pC5T with RM62-GFP	Amp
pKE7	pC5T with ncd-GFP	Amp
pKE8	pC5T with Ago1-Tomato	Amp
pKE9	pC5T with Tomato-Ago2	Amp
pKE10	pC5T with Ago1	Amp
pKE13	pKF63 GFP exchanged by Tomato	Amp
pKE14	pKF63 SV40polyA exchanged by H2a 3'UTR	Amp
pKE15	pKF63 SV40polyA exchanged by H3 3'UTR	Amp
pKE16	pKF63, GFP exchanged by Rluc + SV40polyA exchanged by H2a 3'UTR	Amp
pKE17	pKF63, GFP exchanged by Rluc + SV40polyA exchanged by H3 3'UTR	Amp
pKE18	pKF63, H2a 3'UTR added upstream of SV40polyA	Amp
pKE19	pKF63, H3 3'UTR added upstream of SV40polyA	Amp
pKE20	pKF63, GFP exchanged by Fluc	Amp

3.1.8 Bacterial cells and culture coditions

E.coli XL2-blue CaCl2-competent cells were prepaired and maintained as a laboratory stock and were cultivated in LB-medium or in SOC-medium following transformation. Antibiotic containing agar plates were purchased from in-house supply.

SOB-medium	0.5% (w/v) yeast extract
	2% (w/v) Tryptone
	10 mM NaCl
	2.5 mM KCl
	10 mM $MgCl_2$
	10 mM $MgSO_4$
	pH 7
SOC-medium	SOB-medium
	20 mM Glucose

LB-medium	1% (w/v) Tryptone 0.5% (w/v) yeast extract 1% (w/v) NaCl pH 7.2

Antibiotics added to medium after autoclaving
100 µg/ml ampicillin (100mg/ml stock)
10 µg/ml kanamycin (10mg/ml stock)

3.1.9 DROSOPHILA MELANOGASTER CELLS CULTURE

Laboratory stocks
S2 B2 parental cell line Invitrogen; Karlsruhe, Germany
63N1 and 63N6 endo-siRNA reporter cell lines (Förstemann,2007; Hartig, 2009)

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).

3.1.10 IMMUNOCHEMICAL MATERIALS

3.1.10.1 Primary antibodies

α-Ago1	mouse monoclonal 1b8	1:1000 WB	(Okamura et al., 2004)
α-Ago2	rabibit polyclonal	1:500 WB	lab generated . peptide Ab, Davis
α- β-tubulin	mouse monoclonal E7	1:1000 WB	DSHB
α-Dcr-2	rabbit polyclonal	1:1000 WB	Abcam
α- dcp1			
α-Flag	mouse monoclonal M2	1:2000 WB	Sigma, F1804
α-dFXR	mouse monoclonal 5A11	1:1000 WB	DSHB
α-GW182			
α-Rm62	rat monoclonal	1:1000 WB	
α-myc	mouse monoclonal 9E10	1:1000 WB	
α-ncd	goat polyclonal dS-17	1:1000 WB	Santa Cruz Biotechnology
α-POLO	mouse monoclonal MA294	1:1000 WB	Sunkel-lab
α-Pur- α	rat monoclonal 7H8 and 1E10	1:1000 WB	gift of Elisabeth Kremmer, Helmholtz-Zentrum München, Germany

3.1.10.2 Secondary antibodies

Goat α-mouse IgG (H+L)	HRP-coupled 1:10 000	Pierce via Thermo Scientific
Goat α-rabbit IgG (H+L)	HRP-coupled 1:50 000	Pierce via Thermo Scientific
α-mouse IgG (H+L)	Cy3-coupeld1:200	Molecular Probes, Invitrogen
α-mouse IgG (H+L)	Alexa-Fluor488-coupeld 1:200	Molecular Probes, Invitrogen,
α-rat IgG (H+L)	Cy3-coupeld 1:200	Molecular Probes, Invitrogen

3.2 METHODS

3.2.1 METHODS FOR MOLECULAR CLONING

Inserts for expression plasmids were PCR amplified from S2 cDNA or a suitable plasmid template in a standard PCR reaction mix for molecular cloning in a standard thermocycler protocol using primer sequence extensions to introduce restriction sites as needed.

Standard reaction mixture:

1 µl Template DNA
0.5 µl forward primer (10µM)
0.5 µl reverse primer (10µM)
5 µl 10x Taq-buffer with $(NH_4)_2SO_4$
4 µl 25 mM $MgCl_2$
0.4 µl Taq-polymerase
0.2 µl Pfu-Polymerase
38.4 µl H_2O
50 µl final volume

Thermocycler protocol:

3’	95°C	initial denaturation	.
30”	95°C	denaturation	
30”	55°C	annealing	
1’ /kb	72°C	extension	x32
3’	72°C	final extension	
storage at 4°C			

After amplification the PCR-products were analyzed by separation on an agarose gel (0.5 - 2 % according to expected product size), excised and purified with Qiagen Gel Extraction Kit followed by digestion with the required restriction enzymes according to the manufacturer’s instructions. Vector used for cloning were digested with the corresponding enzymes, ligation was performed with T4 DNA ligase (NEB) according to manufacturer’s protocol and transformation into CaCl competent XL2 blue cells was accomplished by incubation on ice for 30 minutes and heat shock for 2 minutes at 42°C. Insert sequence and correctness of the open reading frame was verified by sequencing (Eurofins MWG) and for further analysis of the sequences ApE was used. The sequences of the oligonucleotides used for molecular cloning are listed in 3.1.7.1

3.2.2 DROSOPHILA SCHNEIDER 2 CELL CULTURE

Cells were cultured in Schneider´s Medium containing 10% heat inactivated fetal bovine serum in appropriate cell culture dishes. Cells were split twice a week into fresh medium for up to 20 passages.

3.2.2.1 Transfection of plasmid DNA

Transfection was essentially carried out as described in (Shah and Forstemann, 2008). 50 ng – 1 µg DNA of the relevant plasmid was transfected per 24-well cell-culture dish according to manufacturer's protocol using 4 µl Fugene Transfection reagents. Cells were harvested after 72 h after transfection to perform further experiments. For stable transfection, 50 ng of a plasmid containing a resistance marker against Hygromycin (pHSHygro) or G418 (pHSneo) was co-transfected together with the relevant plasmid. After three days of culture in Schneider's Medium without antibiotic additives the cells were split into medium containing 300 µg/ml Hygromycin or 1.2 mg/ml G418. The cells were split one a week into fresh medium containing the selection additive for 4 to 6 weeks. In parallel one well was transiently transfected with the relevant plasmid, without addition of a resistance marker to control for successive loss of non-resistant cells. After 4 to 6 weeks the cells were diluted to 6666 cells/ml into selection medium and plated into three columns of a 96 well plate. A 1:3 dilution row was pipetted in the following three columns and diluted further to the right with an 8-channel multipipette. The 96 well plates were sealed with parafilm and cultured at 25°C without moving. After two weeks the plates were checked for colony formation of single cells, colonies formed were carefully harvested using a 10 µl pipette with a 2 µl tip and placed in a 48 well plate with Schneider's medium without selective additive. After at least one week of culture the individual monoclonal cell cultures were analyzed by flow cytometry to select for monoclonal cultures which are characterized by a distinct peak in the fluorescence histogram.

3.2.2.2 RNA-Interference (RNAi)

Knock down of specific proteins in Drosophila Schneider cells was essentially carried out as described in (Michalik et al., 2012) DsRNA for RNAi was generated using in vitro transcription (IVT) with T7-polymerase. As a template for IVT a 300 nt to 600 nt long fragment of the relevant cDNA was amplified with primers introducing T7 promoter sequences on both sides. Oligonucleotides used are listed in 3.1.8.2. In further amplification reactions T7 promoter primers were used to generate cDNA template from specific templates. The resulting PCR products were directly applied for over-night in-vitro transcription at 37°C.

IVT-Mix:
- 25 µl T7-template DNA
- 10 µl 10x T7-buffer
- 0.5 µl 1 M DTT
- 20 µl 100 mM ATP
- 20 µl 100 mM CTP
- 20 µl 100 mM UTP
- 32 µl 100 mM GTP
- 2 µl T7-polymerase
- ad 400 µl H_2O

The sample was heated to 95°C for 5 min in a thermocycler and slowly cooled down to RT. Concentration of dsRNA was estimated from an agarose gel in comparison to a DNA Ladder Mix. Before application the

dsRNA preparation was centrifuged for 1 min to precipitate magnesium pyrophosphate emerging during the IVT reaction. To induce a knock down of a gene of interest cells were seeded at 0.5 x 10^6 cells / ml and 3 µg of the corresponding dsRNA was added per 100 µl cell suspension. After incubation with the dsRNA for 3 days the same amount of dsRNA was added after splitting the cells 1:5. GFP fluorescence of the reporter cell lines after dsRNA treatment was determined with a Becton Dickinson FACSCalibur flow cytometer. For each sample 5 000 or 10 000 cells were acquired and fluorescence intensity was analyzed with CellQuest software. All measurements were performed in triplicates to calculate mean and standard deviations.

3.2.2.3 Stable Isotope labeling with amino acids (SILAC)

Drosophila Schneider 2 cells stably expressing flagHA-Ago2 were split 1:10 into SILAC medium supplemented with heavy lysine, arginine and dialyzed FBS. For at least 7 days and 4 replication cycles cells (Bonaldi, 2008) were cultured in heavy SILAC medium and split 3 times. In parallel cells from the same passage of flagHA-Ago2 cells were cultured in SILAC medium supplemented with light lysine, arginine and dialyzed FBS. All conditions of the two parallel cell cultures were exactly the same except for the heavy and light isotopes. After 10 days the exponentially growing cells were harvested, washed with PBS twice and separated into 3 different cell cycle phases by centrifugal elutriation in two separate runs.

3.2.2.4 Counterflow Centrifugal elutriation

Centrifugal elutriation is a method based on the centripetal force of a centrifuge and the hydraulic thrust of a peristaltic pump to separate cells of different size as for example cells in G1 and G2 phase (Banfalvi, 2008). The elutriation centrifuge was set to 2600 RPM, the peristaltic pump was set to a constant rate of 9 ml/min and the centrifuge chamber and all rubber tubes were filled with elutriation buffer. After carefully inspection that no air bubbles were present in any tube, the bypass was closed and the harvested Drosophila cells labeled with light and heavy isotopes by SILAC were re-suspended in 10 ml of elutriation buffer (using dialyzed FBS) and injected into the elutriation loading chamber (Figure 5). The rate of the peristaltic pump was slowly increased to 12 ml/min to ensure that all the injected cells accumulated in the centrifuge chamber. Afterwards the bypass was reopened. While the speed of the centrifuge was held constant until the end of the elutriation, the rate of the peristaltic pump was slowly increased to collect highly pure cell cycle phases. As soon as the largest (G2) cells left the chamber all the fractions were centrifuged, re-suspended in HEPES lysis buffer and stored in liquid nitrogen. A small aliquot of every fraction was withdrawn before lysis, re-suspended in 1 ml 70 % EtOH for fixation, then re-suspended in 200 µl PBS and digested with 25 µg/ml RNaseA for 30 minutes at room temperature. By staining with propidium iodide the cell cycle separation was verified by flow cytometry.

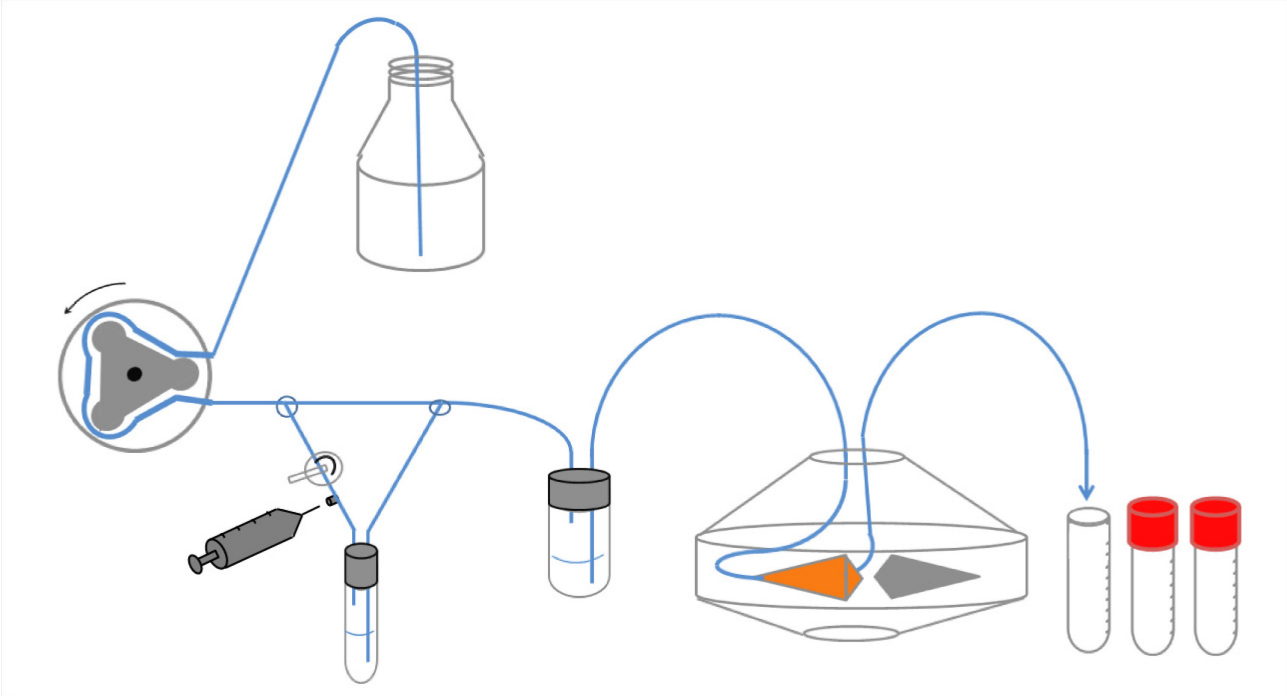


Figure 5 Schematic representation of the counterflow centrifugal elutriation setup. Buffer entered the system from a reservoir (top) through a peristaltic pump (left), proceeded trough a pulse/bubble trap compensator to enter the rotor and was finally collected in separate vessels (right). A second pulse/bubble trap was connected in series with the loading syringe used for loading of the asynchronous cells.

3.2.3 DNA ANALYSIS

3.2.3.1 DNA isolation

DNA for quantitative PCR was isolated according to the Berkley Drosophila genome project protocol, briefly, cells were harvested and lysed in 400 µl Buffer A by incubating at 65°C for 30 minutes. 800 µl of a mixture of 5M KAc and 6 M LiCl 1:2.5 was added and incubated for 15 minutes on ice. The mixture was centrifuged at 13000 RPM for 15 minutes at 4°C, the supernatant was transferred and precipitated with 600 µl Isopropanol. The pellet was washed with 70 % EtOH and re-suspended in 100 µl H_2O. Nucleic acid content was estimated with the Biophotometer without prior RNase treatment; the samples were diluted to 10 ng/µl and used as qPCR template.

DNA for Southern Blot analysis was isolated using the Promega gDNA isolation mini kit according to manufacturer's protocol with an additional RNaseA treatment 3 µg/ml for 20 minutes at 37°C and an additional Proteinase K treatment 0.1 µg/ml at 65°C for 20 minutes. DNA content was estimated in comparison to the DNA ladder mix after running on a 0.5 % agarose gel. The samples were diluted according to the estimated stock concentration and compared again on a 0.5 % agarose gel to adjust the concentrations.

3.2.3.2 Copy number determination by quantitative PCR

QPCR with genomic DNA was carried out as in (Hartig et al., 2009) To determine the number of genomic insertions of a certain sequence after stable transfection gDNA was isolated as described above. 10 ng of gDNA was amplified in a standard qPCR reaction.

Containing: 5 µl 2x Sybr-Green Mastermix (Dynamo Flash, Finnzymes)
2.9 µl H_2O
0.1 µl brilliant blue solution
0.5 µl forward primer
0.5 µl reverse primer

9 µl of the mastermix solution was aliquoted in each well of a 96 well plate using an 8-channal pipette. 1 µl of pre diluted gDNA was added and the samples cycled on the qPCR using the following

PCR-program				
	3’	95°C	initial denaturation	
	1’	95°C	denaturation	
	30”	55°C	annealing	
	42”	72°C	extension	x40
	storage at 4°C			

The primer sequences for copy number determination are listed in chapter 3.1.7.2. DNA levels were quantified with the $2^{-(\Delta\Delta Ct)}$ method (Livak and Schmittgen, 2001).

3.2.3.3 Southern Blot

30 µg of the isolated gDNA was diluted to 145 µl 1 x NEB restriction buffer 4 and stored at 4°C for at least 2 hours. For restriction digestion 50 U BamHI-HF (NEB) was added and the mix was incubated at 37°C and 300 RPM in a shaking heat block. After two hours of incubation 20 more U of BamHI-HF were added and incubated for at least 14 hours. 20 µg Proteinase K were added and the mix was heated to 65°C for 2 hours. Digested gDNA was precipitated with 150 µl isopropanol and 40 µg Glycogen at -20°C for at least 4 hours, washed with 70 % Ethanol once and re-suspended in 25 µl 1.5 x DNA loading dye for 30 minutes at room temperature.

The samples were loaded on a 0.7 % agarose gel without sybr safe in a large electrophoresis chamber and run over night at 20 V. The gel was stained with sybr gold, separation of the lanes and equal loading of the pockets was verified at the Intas UV gel imager and the gDNA was partly depurinated with 2 N HCl for 20 minutes. After neutralization with Southern Blotting Buffer for 20 minutes the capillary transfer chamber was packed as described in the manufacturer’s protocol using a positively charged nylon

membrane for transfer. During the transfer for 4 hours wet paper towels were exchanged every 30 minutes. After transfer the gDNA was crosslinked to the membrane with the UV crosslinker and pre-hybridized in church buffer at 65°C and rotating over night.

Probes for Southern Blot were radioactively labeled using the High Prime labeling kit according to manufacturer's protocol, 25 ng of a DNA probe were diluted to give a final volume of 12 µl, heated to 95°C for 10 minutes and chilled on ice quickly. 4 µl of high prime reaction mix and 4 µl of [α^{32}P] dCTP were added and incubated at 37°C for 10 minutes. The reaction was stopped by adding 2 µl of 0.2 M EDTA and unincorporated dCTPs were removed by centrifugation through an empty sephadex spin column.

The radioactively labeled probe was added into 5 ml of church buffer from the pre-hybridization reaction and incubated over night at 65°C. The blot was washed with 0.5 % SDS / 2 x SSC for 15 minutes at 65°C, followed by 30 minutes with 0.1 % SDS / 1 x SSC and 15 minutes with 0.2 % SDS / 2 xvSSC.

Excess washing buffer was removed with a paper towel; the blot was wrapped in plastic film and exposed to a phosphoimager screen for at least 24 hours. Stripping of the membrane was achieved by washing with boiling 1 % SDS solution for 5 minutes and after pre-hybridization a different probe could be hybridized. The exposed phoshoimager screen was read out at the Typhoon imager.

3.2.3.4 Fluorescence in situ hybridization (FISH)

5 x 10^6 cells were harvested, centrifuged, re-suspended in 10 ml of 0.06 M KCl and incubated at room temperature while softly shaking for 30 minutes. 1/6 Volume of a mixture of ice cold methanol and acetic acid were slowly added and the mixture was centrifuged at 2000 RPM at 4°C. The supernatant was reduced to 500 µl, the pellet was re-suspended and 500 µl of methanol/acetic acid was added. A polylysine coated microscope slides was prepared with a ring of Roti-liquid barrier marker and dipped into ice-water. Cells were applied to the slide by dropping from approximatly 30 cm onto the ice cold slide. The slide was dipped into ice cold ethanol for 10 seconds and airdried briefly. A denaturizing solution containing 2 x SSC and 70 % formamide was heated to 85°C and added to the dry slide, which was then incubated on a 85°C heat block for 10 minutes. After removing the denaturizing solution the slide was incubated under ice cold 70 % ethanol for 2 minutes, then moved under ice cold 100 % ethanol for 2 minutes. After airdrying the 85°C hot probe was added to the dry slide diluted in FISH probe buffer and incubated at 37°C in the dark for at least 14 hours. The slide was washed with 42°C warm 4 x SSC with 0.1% Triton X for 5 minutes and subsequently with 2 x SSC with 0.1 % Triton X. DNA content was stained with DAPI 1 : 10 000 diluted in 2 x SSC for 2 minutes, and the cells were shortly rinsed with 4 x SSC with 0.1 % Triton. Excess liquid was removed, the cells were mounted with DABCO glycerol mounting medium (Ono et al., 2001) and examined

with a Leica TCS SP2 confocal microscope (Leica Microsystems, Wetzlar, Germany).All fluorescently labeled DNA probes are listed in chapter 3.1.7.4.

3.2.4 RNA ANALYSIS

3.2.4.1 RNA isolation

RNA was extracted using Trizol from Drosophila Schneider 2 cells according to manufacturer's protocol.

3.2.4.2 Beta-elimination of RNA

Beta-elimination was carried out as in (Mirkovic-Hosle; 2013; submitted). 60 µg total RNA was diluted in 40.5 µl H_2O, 12 µl 5x borate buffer and 7.5 µl 200mM $NaIO_4$ was added and incubated at room temperature for 10 minutes. The oxidation reaction was stopped addition of 6 µl 100% glycerol for 20 minutes at room temperature, then the pH was elevated to 9.5 by adding approximately 6µl 2 N NaOH. After 90 min at 45°C the RNA sample was centrifuged through a Mini quick spin column and the flow through was precipitated with 3x volume of 100% ethanol and 30 µg glycogen. The RNA pellet was washed two times with 70% ethanol and dissolved in 20 µl H_2O. To confirm successful beta-elimination and removal of one nucleotide at the 3' end the samples were analyzed on a 15% Acrylamide-Urea gel before using 15 µg for a Northern Blot.

3.2.4.3 Northern Blotting

Northern Blot was carried out as described in (Helfer et al., 2012). 1 - 5 µg of RNA were separated on a 20 % Sequagel Acrylamid-Urea Gel at 250 V for ~1.5 hours. RNA transfer was performed on a nylon membrane by semi dry blotting for 30 minutes at 20 V. Crosslinking of the RNA to the membrane was achieved by UV-radiation with the UV crosslinker. Membranes were pre-hybridized in Church buffer for at least 2 hours at 37°C under constant rotation. The probes were labeled with [γ-32P] ATP using PNK according to manufacturer's protocol. 5 pmol DNA was mixed with 4 µl 5x exchange reaction buffer, 1 µl PNK and 6 µl [γ-^{32}P] ATP and added up to a total volume of 20 µl with water. The reaction mix was incubated at 37°C for 10 minutes and the reaction was stopped by heating to 65°c for 5 minutes. Hybridization with labeled antisense-probes was performed overnight in 5 ml Church buffer. Membranes were washed three times for 15 minutes with 2 x SSC 0.1 % SDS, wrapped in plastic film and exposed to a phosphoimager screens. Stripping of the membrane was achieved by washing with boiling 1 % SDS solution for 5 minutes. The membranes were re-hybridized after at least two hours of pre-hybridization. The phosphoimager screen was read out at the Typhoon imager.

3.2.4.4 Quantitative RT-PCR

Messenger RNA levels of individual genes were analyzed by quantitative PCR with the qPCR. After isolation of total RNA with Trizol, reverse transcription was performed with Superscript II reverse transcriptase primed with oligo(dT) or random hexamers (MWG).

Reaction mix for reverse transcription:
500 ng RNA
1 µl dNTP mix (10mM)
1 µl random hexamers /0.5 µl oligo dT primer
Add 12 µl

The mix was incubated for 5 minutes at 65°C, then split into two tubes and the following solutions were added to each one:

4 µl first strand buffer (5x)
2 µl DTT (0.1M)
1 µl Ribolock RNase Inhibitor
1 µl Super Script II / H_2O
The reverse transcription reaction mix was incubated for 50 minutes at 42°C followed by heat inactivation at 70°C for 5 minutes. 90 µl of nuclease free water was added and 1 µl of the diluted cDNA was amplified in a standard qPCR reaction as described in chapter 3.2.3.2.

3.2.4.5 Solexa sample preparation

Isolation of 17-30 nt small RNAs

50-60 µg of total RNA of individual monoclonal cell lines was isolated using Trizol, and separated according to size on a 20 % polyacrylamide/urea gel at 250 V for 1.5 hours. The gel was stained with sybr gold and a slice of gel corresponding to 17 – 30 nt small RNAs was cut out. To find the right localization of the relevant RNA a small RNA ladder (17 - 25 nt) and the 2S rRNA band at 30 nt, which is broadly visible, was used for orientation. The RNA was isolated from the gel by fragmentation and successive shaking in 400 µl SOLEXA elution buffer with Proteinase K for 2 hours at 65°C. The gel slices were separated from the eluate by centrifugation through an empty spin column. The eluate was complemented with 30 µg/ml Glycogen and RNA was extracted with 400 µl Phenol/Chloroform/IAA pH 8.0. RNA was precipitated from the aqueous phase with Isopropanol and the pellet washed with 70 % Ethanol twice. After one minute of drying with closed lid the RNA pellet was re-suspended in 6 µl of Nuclease free water.

3‘ Adapter Ligation

6 µl	RNA
1 µl	ATP free T4 RNA ligase buffer
1 µl	DMSO
1 µl	3’ RNA linker (µl/ml)
1 µl	truncated RNA Ligase
10 µl	final volume

The ligation mix was incubated at 37°C for 15 minutes, 10 µl of Formamide loading dye was added and the mix was heat inactivated for 5 minutes at 95°C.

Isolation of 3’ ligated RNAs

The sample was separated according to size on an 15% polyacrylamide/urea gel at 250V for 75 minutes, after sybr safe staining a slice according to 35 – 50 nt length was cut out, here a small RNA size marker and a 50 bp size ladder helped with orientation. RNA was isolated as described before and re-disolved in 6ul nuclease free water.

5‘ Adapter Ligation

6 µl	3’ligated RNAs
1 µl	10x T4 RNA ligase buffer (Ambion)
1 µl	DMSO
1 µl	5’linker
1 µl	T4 RNA ligase (Ambion)
10 µl	final volume

The reaction mix was incubated at 37°C for 1h and heat inactivated for 5 minutes at 95°C.

Reverse transcription

10 µl	3’and 5‘ ligated RNA
2 µl	3’RT linker

The sample was incubated at 95°C for 2 minutes, cooled on ice for 2 minutes then centrifuged for 1 minute and the following reagents were added:

4 µl	5x first strand buffer
2 µl	0.1 M DTT
1 µl	dNTP mix (10mM)
1 µl	RNase Inhibitor (fermantas)

The reaction mix was split into two microcentrifuge tubes and incubated for 3 minutes at 42°C. Finally 1ul of super script first strand enzyme or 1 up of nuclease free water was added and the + and – RT reactions were incubated at 42°C for 50 minutes followed by heat inactivation at 95°C for 5 minutes.

PCR amplification of the resulting cDNA

5 µl	first strand cDNA (+RT/-RT/H2O)
20 µl	5 x Phusion PCR buffer with MgCl2
2 µl	dNTP mix (10mM)
1 µl	universal 5' SOLEXA primer
1 µl	3' barcode primer
1 µl	hot start phusion polymerase
70 µl	H2O
100 µl	final concentration

Thermocycler protocol for SOLEXA sample amplification:

2'	94°C	initial denaturation	
15"	94°C	denaturation	
30"	60°C	annealing	
30"	72°C	extension	x23
2'	72°C	final extension	
storage at 4°C			

The PCR products were purified from a 2 % Agarose gel with the Quiagen Gel extraction kit.

Cloning of PCR fragments

3 µl	purified PCR product
10 µl	2 x PCR buffer
1 µl	pJet cloning vector
5 µl	water
1 µl	T4 DNA ligase
20 µl	final volume

After incubation for 30 minutes at room temperature the reaction mix was transformed into XL2 blue cells, 5 µl of ligation mix was incubated with 100 µl of CaCl2 competent XL2 blue cells on ice for 30 minutes, heated to 42°C for 2 minutes and cooled on ice for 2 minutes. 600 µl SOC medium was added and the mix was incubated in a 37°C shaker for 30 minutes. The mix was centrifuged briefly; the pellet was plated onto an ampicilin containing agar plate and incubated over night at 37°C. 20 Colonies were amplified in a standard colony PCR (standard PCR protocol with initial denaturation time extended to 10 minutes) using primers provided by the cloning kit to verify insertion and length distribution of correctly ligated small

noncoding RNAs. After gel electrophoresis 8 amplification product were chosen according to their length, purified with the PCR purification kit and sequenced (Eurofins MWG).

3.2.4.6 SOLEXA Data analysis

Solexa sequencing was carried out in house (Illumina Genome Analyzer IIx; LAFUGA), prior to sequencing, quality of libraries was checked by the Agilent 2100 Bioanalyzer. SOLEXA reads were pre-processed using perl scripts; the reads were mapped onto the target sequences using BOWTIE (http://bowtie-bio.sourceforge.net/) with the option –no to force selection of only perfectly matching sequences.

3.2.5 PROTEIN ANALYSIS

3.2.5.1 Protein extraction

S2 cells were harvested (2000 x g, 5 min) washed once in PBS and re-suspended in lysis buffer complemented with 2 x Preotease-Inhibitor. The suspension was frozen in liquid nitrogen, thawed on ice and centrifuged at 4°C in a micro-centrifuge at 16 400 rpm. Protein concentration of the supernatant was determined by Bradford assay.

3.2.5.2 Co-immunoprecipitation

For immunoprecipitation 30 µl magnetic protein A/G beads were covalently coated with antibody (http://www.lamondlab.com/pdf/CovConjAntibodyBeads.pdf) Therefore magnetic beads were incubated with Ago1 (1B8), flag (M2) or myc (9E10) antibodies for 4 hours at 4°C and unbound antibody was removed by washing with PBS twice. Covalent crosslink was accomplished by washing with 0.1 M Sodium Borate solution at pH 9.0, followed by two incubations times at RT for 30 minutes with freshly made ice cold DMP/borate solution. The beads were washed with 1 ml of 50 mM Glycin at pH 2.5 and subsequently with PBS. The covalently conjugated beads were stored as 50 % slurry in 20 % EtOH.

The pre-coated magnetic beads were incubated for 20 minutes with 0.5 - 2 µg protein extract at 4°C on a over head rolling wheel in a microcentrifuge tube. The tube was placed on a magnetic rack (Kisker) and the supernatant was discarded. The beads were re-suspended three times in 500 µl lysis buffer containing 1 % Triton, placed back to the magnetic rack and the supernatant was discarded, then the beads were rinsed with lysis buffer without Triton. The bound proteins were eluted by heating the samples to 95°C with 20 µl 1 x Laemmli buffer diluted with lysis buffer for 5 minutes; the eluate was pipetted from the tube placed in the magnetic rack.

3.2.5.3 Western-Blotting

Western Blotting was performed as described in (Aumiller et al, 2012) For Western Blotting (after co-IP) proteins were separated using 8 - 12 % polyacrylamide gels for 90 minutes at 150 V in a BioRad electrophoresis tank. Proteins were transferred to a polyvinylidene fluoride (PVDF) membrane by tank blotting (100 V for 60 minutes). Afterwards the membrane was blocked in 5 % milk for at least 20 minutes at room temperature. Incubation with primary antibodies was carried out over night at 4°C in a 50 ml Falcon tube using at least 3 ml antibody dilution in 1 x TBS-solution with 0,02 % Tween. Antibodies and dilutions used are listed in 3.1.11

After primary antibody binding the membrane was washed three times 10 minutes in TBS-T at room temperature and incubated with appropriate secondary antibody for 2 hours at room temperature. After the washing steps, Enhanced Chemiluminescence (ECL) substrate was applied and the resulting signal was measured in an LAS3000 mini Western Imager System.

Western blots were stripped with 10 ml of Restore Stripping Solution for 30 minutes at room temperature, rinsed with TBS-T and blocked with 5 % milk for second primary antibody incubation.

3.2.5.4 SILAC sample preparation and data analysis

The bound fractions after co-immunoprecipitation were separated on a 4 - 20 % polyacrylamide gradient gel and the lanes were slices into 5 individual size ranges. Protein identification was performed by LC-MS/MS (ZfP München). ZfP provided lists containing protein ID with peptide coverage and H/L intensity ratios for every IP, which were compared and sorted to find reproducibly copurified proteins with H/L intensities different from 1.

3.2.5.5 Analysis of Schneider cells by fluorescence microscopy

Fluorescent labeling of intracellular proteins was carried out as described in (Aumiller et al., 2012) Transfected cells were harvested and once washed with Schneider´s Medium without FBS. Then 100 µl of cell suspension was allowed to settle down within a Roti-Liquid Barrier marker painted circle on a Poly-lysine coated microscope slide for 30 minutes in a humid chamber. Cells were fixed for 15 minutes with 4 % formaldehyde and washed three times for 5 minutes in PBS + 0,2 % Triton X (PBT), then blocked for 1 hour with PBT containing 5 % bovine serum albumine (BSA) in the humid chamber.

Primary antibodies were diluted 1 : 200 – 1 : 50 in blocking solution and incubated overnight at 4°C. Cells on the slides were washed three times with PBT, secondary antibodies were diluted 1 : 200 in PBT and incubated 2 hours at room temperature, followed by three times washing for 5 minutes with PBT. DNA was stained with 0.5 µg/ml DAPI for 2 minutes at room temperature. Cells were washed with PBT, mounted with DABCO glycerol mounting medium (Ono et al., 2001) and examined with a Leica TCS SP2 confocal microscope. Primary and secondary antibodies used and the applied dilutions can be found in chapter 3.1.11.

3.2.5.6 Flow cytometry

For intracellular labeling of protein with specific antibodies harvested cells were washed with PBS with 2 mM EDTA and 2 % BSA, fixed with 1 % PFA for 15 minutes at room temperature, rinsed with PBS and permeabilized with saponine. Therefore the cells were sequentially washed with 0.1 % saponin in PBS and with 0.35 % saponin in PBS. Incubation with primary antibody was carried out for 20 minutes at room temperature and the secondary antibody was incubated for 20 minutes at room temperature as well after washing with 0.1 % saponin once. The cells were finally washed 3 times with 0.1 % Saponin, re-suspended in FACS flow and analyzed at the FACSCalibur.

3.2.5.7 Luciferase assay

Luciferase assay was carried out using the Promega Dual Luciferase Reporter Assay system and was analyzed at the Berthold Centro LB 960 Microplate Luminometer.

4 RESULTS

4.1 COMPARISON OF RISC COMPOSITION IN DIFFERENT CELL CYCLE PHASES

4.1.1 GENERATION OF A MONOCLONAL STABLE CELL LINE EXPRESSING FLAGHA AGO2

In order to understand possible alterations and modifications of RISC which might appear periodically during cell cycle progression, Ago1 and Ago2 were immunoprecipitated to analyze associated proteins. As for Drosophila Ago2 no suitable immunoprecipitating antibodies exist, a monoclonal stable cell line expressing flagHA-tagged Ago2 was generated. To this end the plasmid pJB47 (a gift of Dr. Julius Brennecke, see M&M 3.1.7.8) was transfected into S2 cells together with a G418 resistance plasmid and selected for stable transfection. Monoclonal selection was performed by limiting dilution and colony picking, then verified by intracellular fluorescence labeling of the flag-tag followed by flow cytometry (Figure 6a).

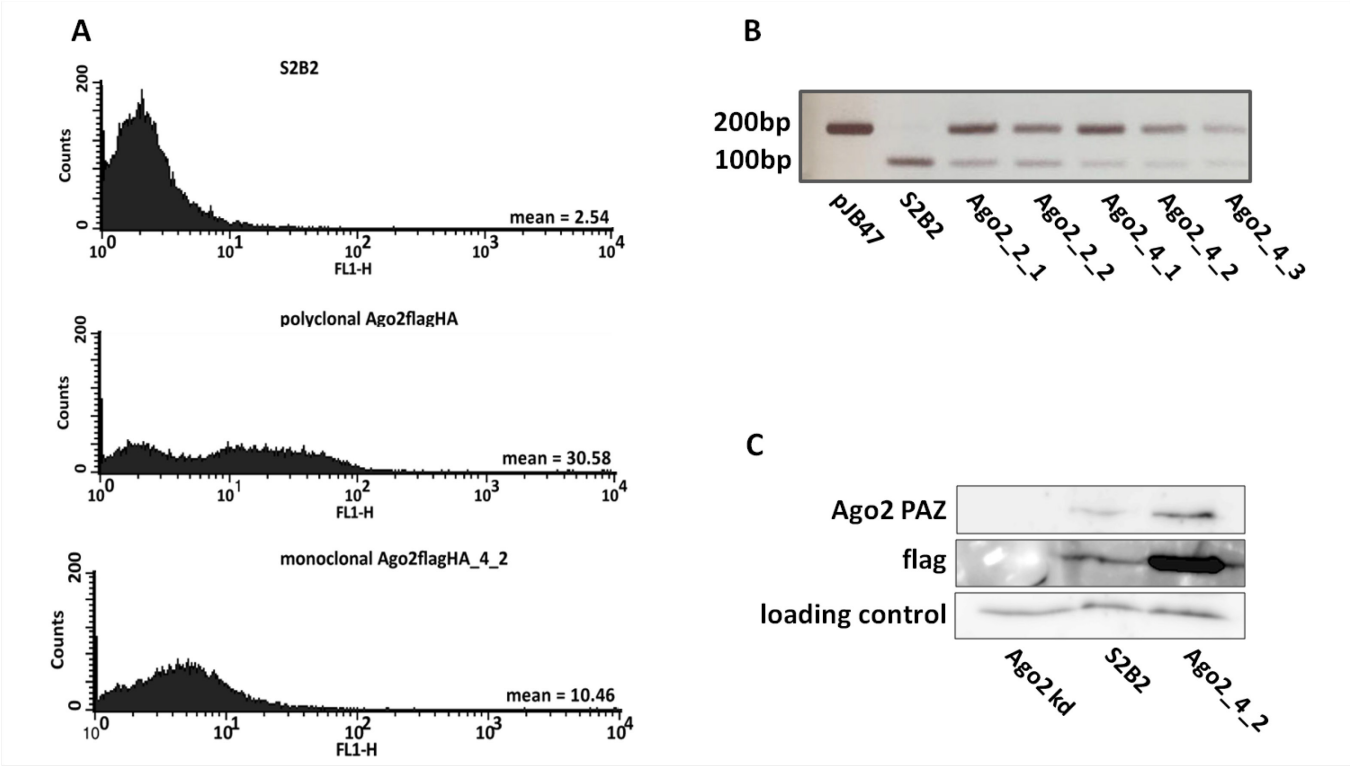


Figure 6 Selection of a stably transfected monoclonal cell line expressing pJB47 (flagHA-Ago2). (A) S2 cells were co-transfected with pJB47 and pSHNeo and selected for stable transfection. Cells were diluted; monoclonal cultures were picked, fluorescently labeled with flag antibody and analyzed by flow cytometry. The upper panel shows fluorescence intensity (FI) of the parental cell line (S2) in a histogram plot, the middle panel shows the FI of the polyclonal culture after stable transfection, the lower panel shows the FI of the culture used for further experiments (Ago2flagHA_4_2) after selection for monoclonal cell lines. (B) The ratio of endogenic (100 bp) to transgenic DNA material (200bp) was calculated. A primer pair amplifying fragments of different length from endogenous gDNA and transgenic DNA, which included a flagHA-tag was employed the number of transgene insertions was estimated to be about 3. (C) To reconfirm stable integration of the flagHA-Ago2 expressing plasmid, the stable cell line was cultured for 10 passages and flag and Ago2 expression was verified by Western Blot. The first lane shows protein extract from S2 cells after Ago2 knock down, the second lane shows the parental cell line (S2) and the right lane shows the stable cell line (Ago2flagHA_4_2) after 10 passages.

The amount of stably transfected, transgenic material was compared to the endogenous Ago2 genomic locus by PCR (Figure 6b), the same primer pair resulted in two fragments, a smaller fragment from the endogenous locus and a larger fragment including the flagHA-tag from the transgenic locus, a ratio of approximately 3 copies of the transgene was estimated based on relative intensity of the band for the cell line Ago2flagHA_4_2. Stable transfection and steady expression level of the transgenic construct was verified by Western Blot ten passages after selection (Figure 6c). The cell line Ago2flagHA_4_2 was used for further experiments.

4.1.2 CO-IMMUNOPRECIPITATION OF KNOWN INTERACTORS OF ARGONAUTE1 AND ARGONAUTE2

As a positive control and to optimize precipitation and washing conditions co-immunoprecipitation of Ago1 and Ago2 was performed and known interactors were detected via Western Blot. GW182 is known to be essential in Ago1 downstream function and its direct binding to Ago1 was shown before (Behm-Ansmant et al., 2006). Drosophila *Fragile X protein* (dFXR) was also already shown to interact with Ago1 (Ishizuka et al., 2002) and was applied as a second positive control for co-IP efficiency. Ago2 interacts with Dcr-2 during RISC loading (Liu et al., 2006) and with Pur-α (Aumiller et al., 2012). Ago1 could be highly enriched together with GW182 and dFXR by co-IP (Figure 7a) and Ago2 could be enriched moderately

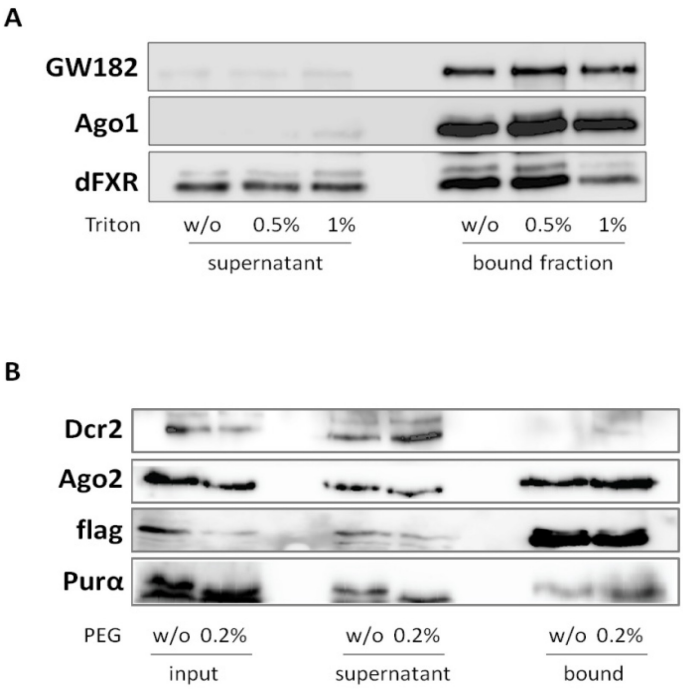


Figure 7 Ago1 and flag Ago2 IP. Ago2flagHA_4_2 cell extract was immunoprecipitated with either Ago1- (A) or flag- (B) antibody 500 mg of protein extract was used, input reflects 10% of total extract. The known interactors of Ago1, GW182 and dFXR (A), and the known interactors of Ago2, Dcr2 and purα (B) were detected in a Western Blot. For the Ago1-IP the supernatant, which was almost depleted from Ago1 and GW182 on the left was compared to the bound fraction on the right. For the flagAgo2 IP input, supernatant and bound fractions were compared. Different concentrations of Triton-X 100 in the washing buffer and PEG in the lysis buffer were tested. The following experiments were carried out with 0.2% PEG in the lysis buffer and adding 0.5% Triton-X 100 for the washing step.

together with a small fraction of pur-α and very little Dcr-2 (Figur 7b), which could be ascribed to a transient interaction. At the same time the Western Blot showed that 1 % Triton-X 100 in the washing buffer might be too stringent and could lead to loss of at least some interacting partners. Adding 0.2 % PEG into the co-IP buffer might help with protein enrichment. The following experiments were performed using 0.5 % Triton-X 100 and 0.02 % PEG in the washing buffer.

4.1.3 CELL CYCLE PHASE SEPARATION AND SILAC LABELING

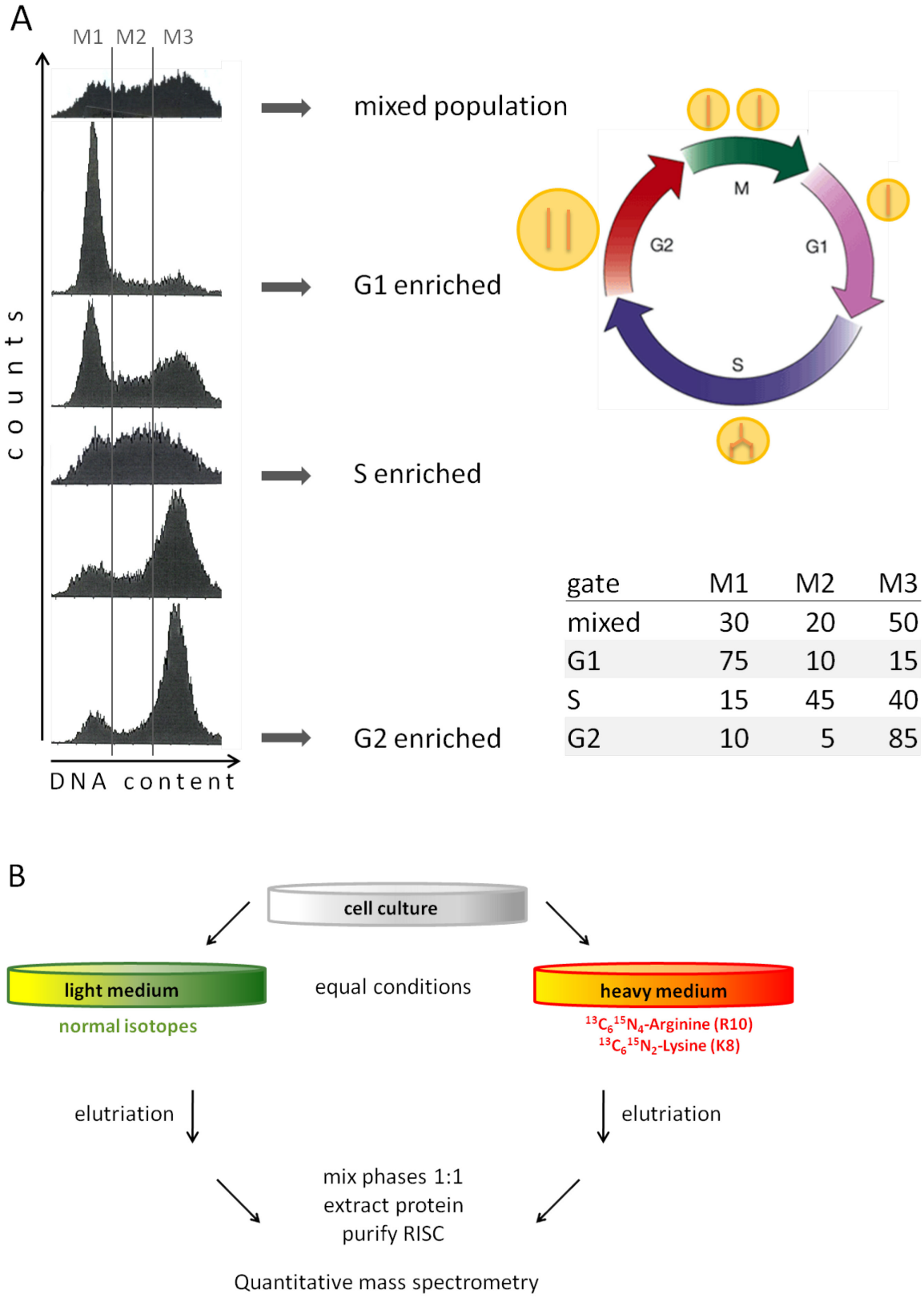
Centrifugal elutriation can enrich cells in different phases of the cell cycle. The method relies on a mechanical separation of the cells according to cell size rather than the induction of chemical blocks or cell cycle arrest (see M&M chapter 3.2.2.4). Avoiding any chemical or physical manipulation of the cells made centrifugal elutriation a suitable method for us to compare protein content and interaction of RISC in different stages of cell cycle progression.

Efficiency of phase separation was verified by staining of the genomic DNA with propidium iodide (PI) followed by flow cytometry. Figure 8a shows histograms of fluorescence intensity of PI. The first histogram plot represents asynchronous cells, while the following plots show elutriation fractions after separation. The first peak on the left (M1) represents G1 phase cells which have an 1n DNA content, the right peak represents G2 phase cells (M3), which have a duplicated chromosome set and thus twice the intensity of PI per cell. Between G1 and G2 the cells replicate their genome in S phase (Figure 8a).

Stable isotope labeling by amino acids in cell culture (SILAC) allows direct comparison of changes in protein expression between these closely related samples (Mann, 2006), therefore one population of cells was labeled with heavy isotopes and separated into three fractions each enriched in G1, S or G2 phase respectively (Figure 8b). A cell population incorporating light isotopes was treated the same way and fractionized. Now different combinations of cells could be mixed 1:1, for example, the G1 phase of the heavy population was mixed with the S phase of the light population. (Table 2) After protein extraction immunoprecipitation was performed to analyze differences in protein content associated to RISC.

lane	1	2	3	4	5	6	7	8	9
IP	myc	Ago1	flag		Ago1			flag	
light population		asynchronous		G1	S	G2	G1	S	G2
heavy population		asynchronous		G2	G1	S	G2	G1	S

Table 2 Combinations of heavy and light labeled cell extract from pooled elutriation fractions. The elutriation fractions from the two runs were gave three populations enriched in G1, S or G2 phase each. For the following Ago1- and flagAgo2-IPs a heavy and a light population was mixed 1 : 1 and employed in an Ago1- and a flagAgo2-IP. As a control, asynchronous heavy and light labeled cells were mixed 1 : 1 and applied for Ago1- flagAgo2- and myc-IP. Myc served as a negative control; the purpose of the Ago1 and flag IP was mainly to estimate intrinsic variations in intensity ratios.



gate	M1	M2	M3
mixed	30	20	50
G1	75	10	15
S	15	45	40
G2	10	5	85

Figure 8 Experimental outline combining counterflow centrifugal elutriation and stable isotope labeling in cell culture. (A) Cell cycle phase separation efficiency of centrifugal elutriation was verified by DNA staining with PI and flow cytometry. The histogram plots show fluorescence intensity of the populations before (first diagram) and after elutriation. G1 phase (M1), S phase (M2) and G2 phase (M3) are gated and the percentage of cells corresponding to M1, M2 and M3 for each elutriation fraction is shown in the table. (cell circle adapted from Coleman et al., 2004) (B) Schematic overview of SILAC labeling. Ago2_4_2 cells were cultivated medium supplemented with heavy or light labeled amino acids under equal conditions. Both cultures were elutriated separately and fractions each enriched in G1-, S- or G2-phase were mixed 1:1 from heavy and light cells. Protein was extracted and Ago1- or Ago2-RISC was purified, separated by SDS page and quantitatively analyzed by mass spectrometry.

4.1.4 DEFINITION OF CANDIDATE PROTEINS

The IPs were performed with combinations of heavy and light labeled cells from different elutriation fractions (see table 2), to enable comparison of as many cell cycle phases as possible. Copurified proteins bound to Ago1, flagAgo2 or as a negative control to mouse α-myc antibody were separated on a SDS page gel. Lanes were sliced in five separate patches and send to mass spectrometry for SILAC analysis to the ZfP (Zentrallabor für Proteinanalytik) of the LMU. Analysis was done by LC-MS/MS followed by maxquant analysis done by the core facility. Labeling with stable heavy isotopes was successful and fits a normal distribution as depicted in Figure 9.

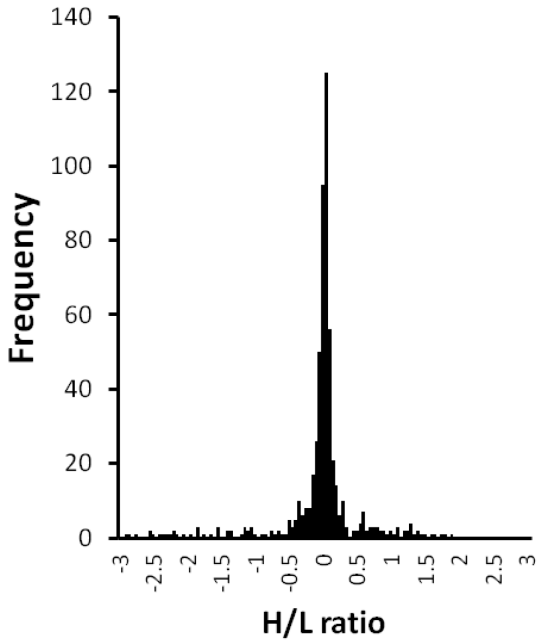


Figure 9 Histogram plot representing H/L ratios of all proteins identified by mass spectrometry (n=570) in 1:1 mixed fractions, H/L ratios are logarithmized. The overall labeling rate can be deduced from this plot, as all IP fractions are averaged and cell cycle dependent variations are qualified.

The list of bound proteins was sorted by peptide coverage. Ago1 and Ago2 were present in the respective IPs. Proteins associated with Ago1 or Ago2 in all three IPs are listed in table 3. Proteins also detected in the negative control (myc-IP) were excluded. Elongation factor 2 (EF2), *heat shock cognate 70-3* (Hsc70-3), the microtubule motor protein *non claret disjunctional* (ncd), *POLO kinase*, *eukaryotic initiation factor 4G* (eIF4G) and the *ribosomal protein S3a* could be detected in every Ago1- and flagAgo2-IP, but not in the myc-IP, the *deadbox helicase protein* RM62 and *the ribosomal protein L14* could be detected only in the three Ago1-IPs but not in the flagAgo2- and the myc-IP. No protein with sufficient peptide coverage could be reproducibly found associated with Ago2 but not with Ago1. It is important to notice that the positive controls used to optimize IP conditions could not be detected by mass spectrometry. Beside reproducible presence of a protein associated with Ago which is demonstrated by the appearance in all three corresponding IPs, a main goal of the experiment was to detect possible variations in the stoichiometry of

association in different cell cycle phases. Therefore the ratio of heavy to light isotopes was calculated in the fractions, which were precipitated from a 1 : 1 mixture of for example a light G1 and a heavy G2 population (see table 2). A factor heavy : light above 1 in this case would represent a more frequently association in G2 phase compared to G1. In table 3 the intensity ratios of heavy to light isotopes are listed. It can be concluded from the intensity ratios, that POLO and ncd are more frequently associated in G2 phase in this experiment, which can be explained by their role in mitosis. Rm62 is stronger associated with Ago1 in S phase. The ribosomal proteins, Hsc70-3 and the elongation and initiation factors were detected in roughly equal ratio in all cell cycle phases, as well as Ago1 and Ago2. In further experiments the association of ncd, POLO and Rm62 to Ago was sought to be clarified. The complete list of associated proteins obtained from this experiment is attached in the appendix.

	Ago1			Ago2		
Protein	G1/G2	S/G1	G2/S	G1/G2	S/G1	G2/S
EF2 (84.4kDa) peptid coverage	**27**	**29**	**29**	**18**	**20**	**20**
intensity ratio h/l	*1.0*	*1.0*	*1.5*	*1.2*	*1.1*	*1.0*
Hsc70-3 (72 kDa)	**19**	**18**	**20**	**15**	**18**	**16**
	1.4	*1.9*	*1.3*	*1.4*	*1.8*	*1.2*
ncd (70 kDa)	**9**	**10**	**5**	**11**	**10**	**10**
	0.6	*1.0*	*2.2*	*0.7*	*1.0*	*1.7*
POLO (57 kDa)	**1**	**3**	**1**	**1**	**10**	**3**
	0.4	*5.4*	*3.0*	*0.5*	*5.6*	*4.8*
eIF4G (211 kDa)	**6**	**8**	**4**	**5**	**9**	**6**
	1.0	*0.9*	*1.1*	*1.3*	*0.9*	*1.8*
Rp S3a (30 kDa)	**1**	**3**	**3**	**2**		**2**
	1.0	*1.1*	*1.1*	*0.9*		*0.3*
Argonaute1 (110 kDa)	**21**	**20**	**18**			
	1.0	*0.7*	*1*			
Rm62 (62 kDa)	**4**	**8**	**6**			
	0.8	*0.5*	*0.7*			
Rp L14 (19 kDa)	**3**	**2**	**3**			
	1.0	*0.9*	*1.1*			
Argonaute2 (130 kDa)				**12**	**13**	**13**
				1.0	*1.3*	*1.0*

Table 3 Peptide coverage and intensity ratio of the candidate proteins. The list of proteins received from ZfP was sorted according to peptide coverage. Proteins found in all three corresponding Ago1- or flagAgo2-IPs but not in the negative control IP (myc) are listed with their peptide coverage and intensity ratio (H/L). Rm62 and the ribosomal protein L14 was found to be associated with Ago1 only, while no protein with higher peptide coverage was found to be associated with Ago2 but not Ago1. Intensity ratios for Ago1 and Ago2 were near 1 and varied for the associated proteins.

4.1.5 VERIFICATION OF INTERACTION VIA WESTERN BLOTTING

To further confirm the association of the candidate proteins with Ago the experiment was repeated and detection of the candidates in the bound fraction was attempted via Western Blot. In Figure 10a a control IP performed with asynchronously growing cells is shown. All proteins can clearly be detected in the input. Both IPs efficiently precipitated the Ago protein, but the negative control myc also precipitated relative high levels of Ago1, ncd and Rm62 compared to 5% input. In the bound fraction of the Ago1- and flagAgo2-IP only very low levels of the candidate proteins could be detected. Figure 10b shows a Western Blot of the flagAgo2-IP, here ncd and POLO could only be detected in very low level in the bound fraction compared to 10% input. The amount of precipitated flagHA-tagged Ago2 did not change between the different cell cycle phases and also stays constant in the input (Figure 10d). The amount of ncd however increased towards G2 phase and in the bound fractions it peaks in S phase. Figure 10 c shows the Western Blot after Ago1-IP. Ago1 precipitation was very efficient (Figure 10d) but neither ncd and POLO nor Rm62 could be enriched in the bound fraction compared to 10 % input. To address the question, whether the interaction of the candidate proteins changes in the course of the cell cycle, the Western Blot shown in Figure 10c was quantified. Intensities of the bound fractions were normalized to input intensities and the amount detected in G1 phase was set to 1. Figure 10d illustrates the changes of POLO, ncd and RM62 associated with Ago1, adjusted to expression levels in the three cell cycle phases analyzed.

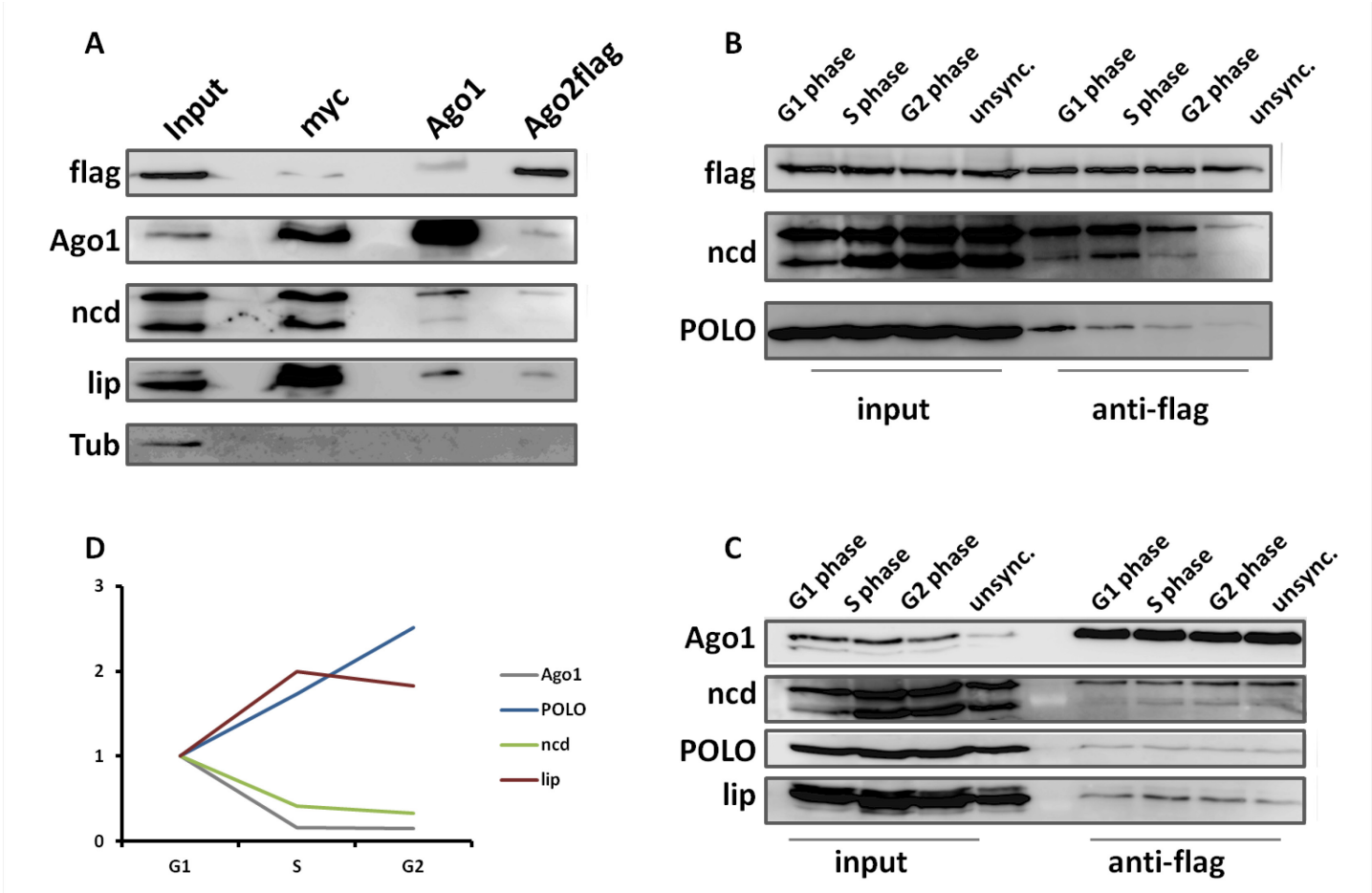


Figure 10 Interaction of ncd, POLO and RM62 with Ago could not be verified. Ago1- and flagAgo2-IPs were carried out under the same conditions as in Figure 7 and Table 2 and the bound fraction and input analyzed by Western Blot (A) Asynchronous cells were lysed and immunoprecipitated with myc-, Ago1- or flag-antibody. 2mg of total protein was used per IP, the first lane shows the input, and the following lanes show the bound fractions. (B) Elutriated cells were lysed and immunoprecipitated with Ago1-antibody, here 500 µg of total protein extract were used. Lane 1-4 shows the input from populations enriched in G1-, S- or G2-phase or asynchronous cells, lane 6-9 show the corresponding bound fractions. (C) Western Blot after flagAgo2-IP organized like in B. (D)The diagram shows a quantification of C. Proteins amounts detected in the bound fraction are normalized to the input and set in relation to G1.

4.1.6 CO-LOCALIZATION OF CANDIDATE PROTEINS WITH ARGONAUTE 1/2

To survey a possible protein-protein interaction of the candidate proteins with Ago more closely, GFP tagged versions of Rm62 and ncd were engineered and co-transfected together with Tomato tagged Argonaute versions to check for co-localization in immunofluorescence microscopy. Tomato is a red fluorescent monomeric mutant of Dsred (Shaner et al., 2004), constructs used for these experiments are listed in M&M chapter 3.1.7.8.

S2 cells were co-transfected with ncd-GFP and Ago1-Tomato, harvested after 3 days, stained with DAPI and analyzed at the confocal microscope. Sufficient ncd-GFP expression was only visible in mitotic cells and co-localization with Ago1 was not observed (Figure 11a). The same was true for Ago2-Tomato and ncd-GFP co-expression (Figure 11b). To avoid mis-localization and the formation of protein aggregates due to over-expression of Ago1/2-GFP, S2 cells or Ago2flaHA_4_2 cells were transfected with ncd-GFP, harvested after 3 days and labeled intracellularily with either the Ago1 antibody or with the flag antibody. Again only mitotic cells could be analyzed due to the low number of ncd expressing cells and no co-localization of ncd with Ago1 or Ago2 could be detected (Figure 11c and d).

Figure 12 shows fluorescence microscopy of cells co-transfected with Rm62-GFP and Ago1-Tomato. Rm62-GFP is located within the nucleus and could not be detected in the cytosol (Figure 12a). Ago1-Tomato is predominantly located in the cytosol and a co-localization of Ago1-Tomato and RM62-GFP could not be verified in these experiments (Figure 12b).

The main function of Ago proteins takes place in the cytosol. However, a role of Ago2 in chromatin modification and DNA double strand break detection becomes more and more evident. To actively participate in these processes Ago2 needs to be localized in the nucleus as well. Having all the constructs and techniques for Ago detection at hand, the localization of Ago1 and Ago2 in the nucleus was tested in further experiments. Figure 12c shows S2 cells labeled with Ago1 antibody and stained with DAPI. Figure 12d shows S2 cells transiently transfected with flaHA-Ago2, harvested after three days and labeled with flag antibody and DAPI. Both Figures show that Ago1 and Ago2 can be detected within the nucleus, but the main portion is localized in the cytosol.

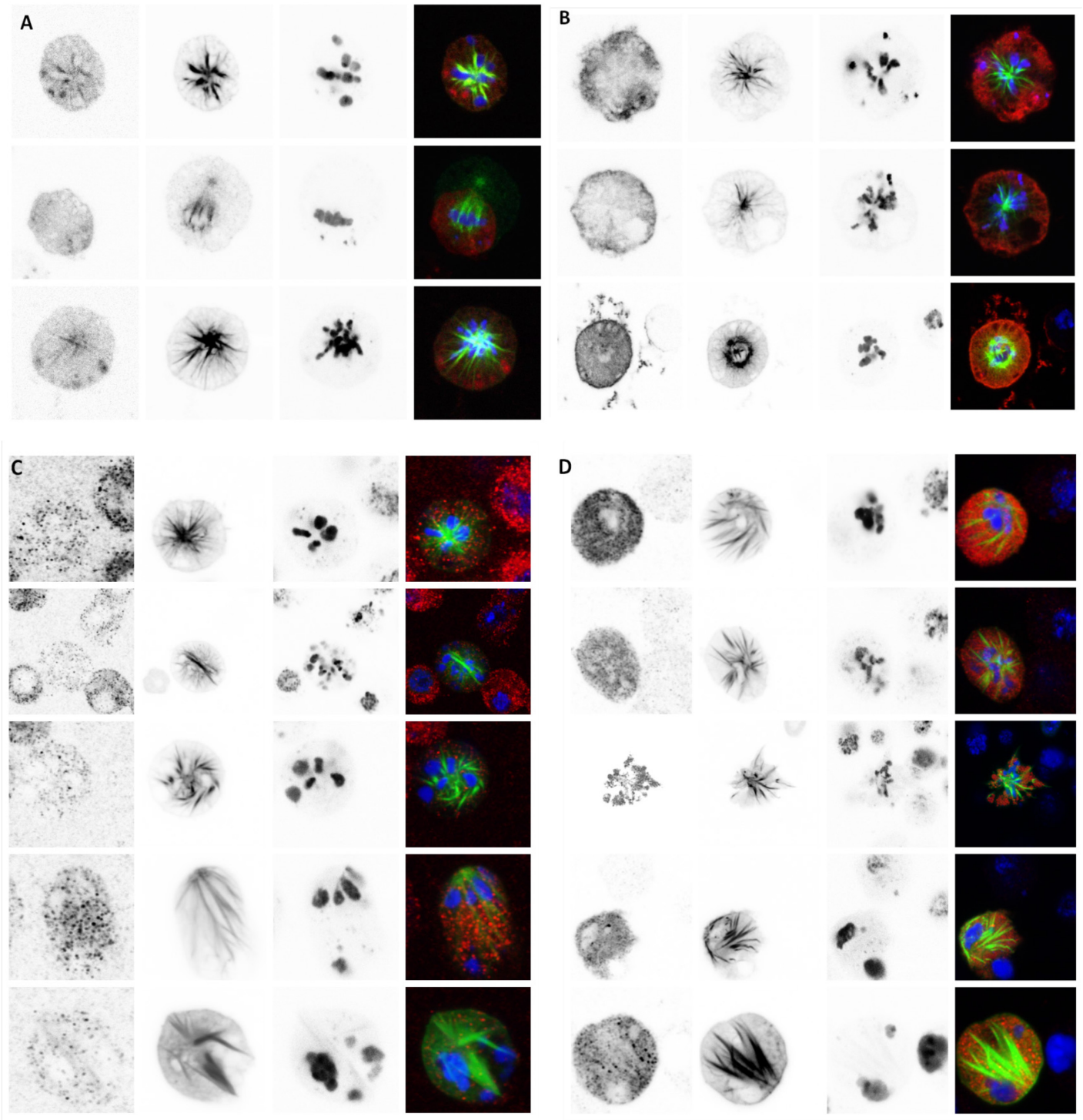


Figure 11 **Co-transfection of GFP tagged ncd and Tomato tagged Ago1 or Ago2.** (A) S2B2 cells were co-transfected with pKE7 (ncd-GFP) and pKE8 (Ago1-Tomato), stained with DAPI and analyzed via confocal microscopy. The left panels show the red channel detecting Ago1-Tomato in black and white and inverted for better visibility, all following confocal microscopy figures are visualized this way. The second panels show the green channel with ncd-GFP, the third panels show the blue channel with DAPI and the right panels show an overlay of all three channels. (B) The same experiment, except Tomato tagged Ago2(pKE9) was transfected instead of Ago1. (C) S2B2 cells were transfected with pKE7 (ncd-GFP), fluorescently labeled with Ago1-antibody, stained with DAPI and analyzed via confocal microscopy. The Figure is organized like (A), but the left panel show Cy3 labeled endogenous Ago1 instead of Tomato tagged overexpression. (D) Ago2flagHA_4_2 cells were transfected with pKE7 (ncd-GFP), fluorescently (Cy3) labeled with flag antibody, stained with DAPI and analyzed. The Figure is organized like (A) but the left panel shows stably transfected flagHA-Ago2 mildly overexpressed under the endogenous promotor.

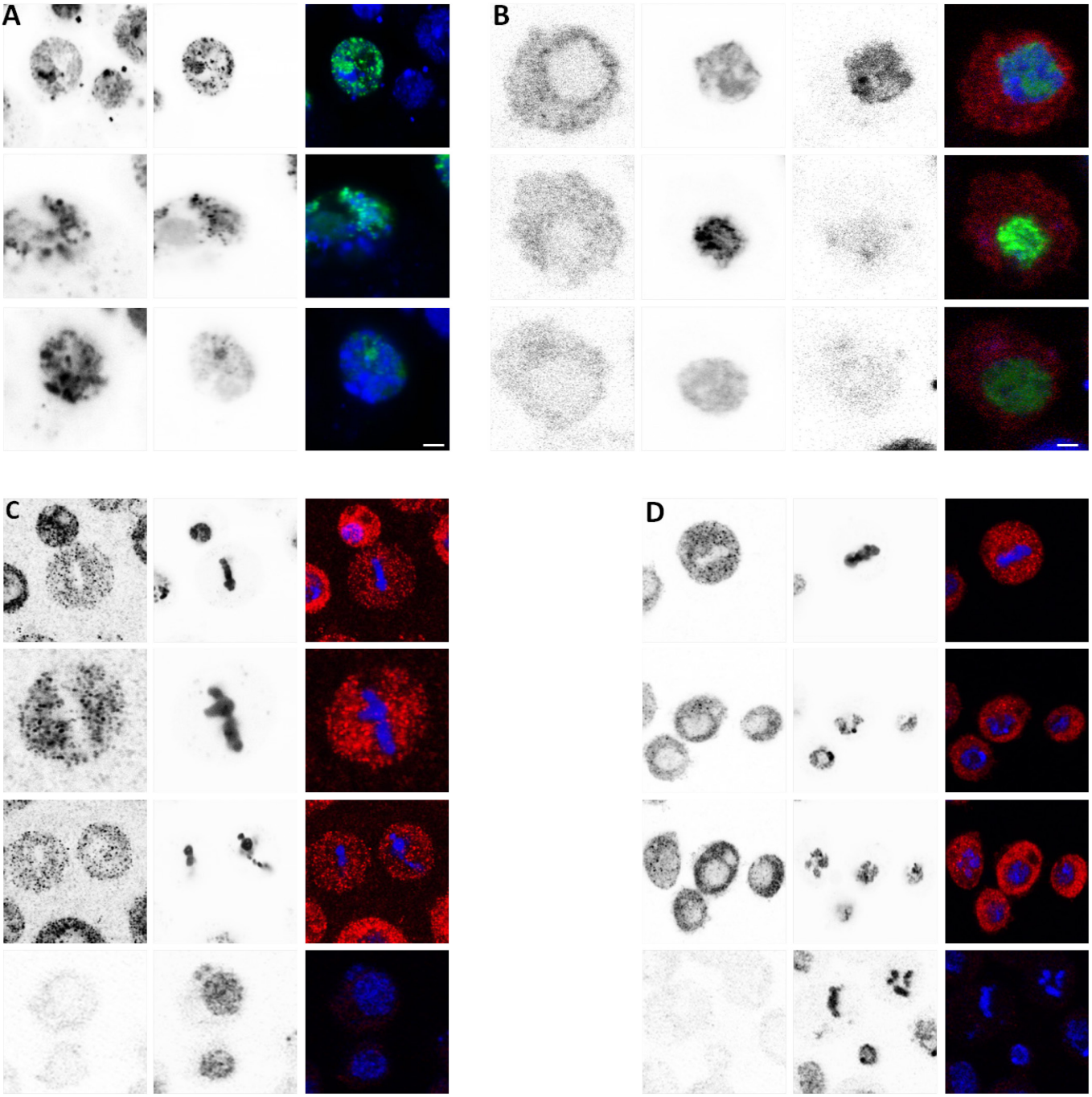


Figure 12 **Localization of RM62, Ago1 and Ago2.** (A) Transfection of pKE6 (RM62-GFP). The left panels shows the blue channel with DAPI, the middle panels shows the green channel with RM62-GFP and the right panels shows an overlays. RM62 is localized exclusively in the nucleus. (B) Co-transfection of pKE8 (Ago1-Tomato) and pKE6. On the left the red channel visualizing pKE8 is shown, the next panel shows the green channel representing Rm62 and the third channel shows DAPI. Scale bar represents 5µm. Ago1 and Ago2 both can be detected in the nucleus. S2 cells were transiently transfected with flagHA Ago2 with endogenous promoter and intracellularily labeled with Ago1 or flag antibody. Nuclei were stained with DAPI. (C) Ago1 can be detected in the cytosol in high levels, and also in the nucleus, the lowest panel shows cells incubated with secondary antibody only and stained with DAPI. (D) Ago2flag could be detected in the cytosol and the nucleus, the lowest panel shows cells incubated with secondary antibody only and and stained with DAPI.

4.2 RANDOM TRANSGENE INSERTION

The second part of the thesis will be concerned with the control of selfish genetic elements by endo-siRNA formation against them. To examine mechanisms involved in precursor formation of transposon directed endo-siRNAs, reporter cell lines which generated endo-siRNAs against a stably integrated reporter construct were analyzed. Altogether a diverse set of about 200 stable cell lines were generated and examined regarding their number of stable integration by flow cytometry and quantitative PCR of genomic DNA.

4.2.1 TRANSGENES INTEGRATED IN DIFFERENT COPY NUMBERS

Reporter constructs used for the establishment of stable cell lines all carry a coding sequence which can easily be measured by flow cytometry of luciferase assay and an Ubiquitin 64E promoter sequence, which is commonly used in our laboratory. Two of the constructs carry an heterologous simian vacuolating virus (SV40) poly A signal, which is also commonly used in fly expression vectors (pUAS-T expression plasmid or VALIUM system; Brand and Perrimon, 1993; Ni et al., 2008). Two of the reporter plasmids carry the *Drosophila* histone H2a or histone H3 3'UTR and in two of them both 3'UTRs follow the coding sequence sequentially. These construct were co-transfected with a resistance marker for stable selection and were cultured to result in monoclonal cell lines with stable reporter gene integration in different copy numbers. Monoclonal selection was verified with flow cytometry, only cell lines showing uniform GFP expression indicated by a narrow peak in an MFI histogram plot were used for further experiments.

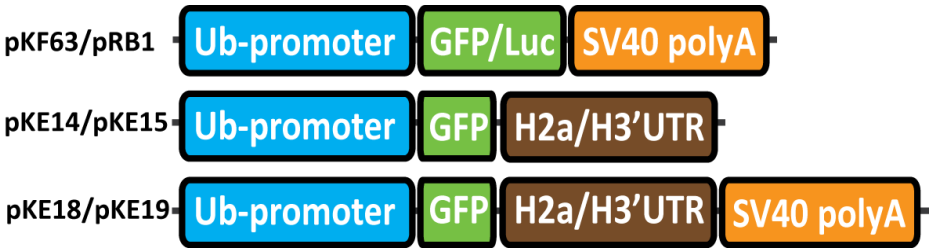


Figure 13 Constructs used for stable cell line generation. All reporter constructs carry an Ubiquitin Promoter and a coding region for an easily detectable protein. To compare the properties of reporter cells bearing a canonical poly A signal to reporters with histone stem loop the 3' UTR of the construct was exchanged with a histone H2a of histone H3 3'UTR or engineered in series. (not drawn to scale).

We analyzed a diverse set of cell lines by qPCR of genomic DNA as explained in M&M chapter 3.2.3.2. Figure 14 shows the fold change of Ubiquitin 64E over the parental cell line s2. qPCR was carried out with a primer detecting a sequence in the Ubiquitin (Ub) 64E Promoter region, which is present in the parental cell line S2 in one copy per haploid genome. Based on this background the fold change in qPCR of stably transfected clones reflect the copy number of the transgene carrying the Ub64E promoter. In another qPCR experiment adopted from Hartig et al., 2009 the quantity of detected copies of the Ub64E was normalized to the quantity of two other one-copy genes (CG5599 and CG1673), which lead to the same result (Hartig et al., 2009).

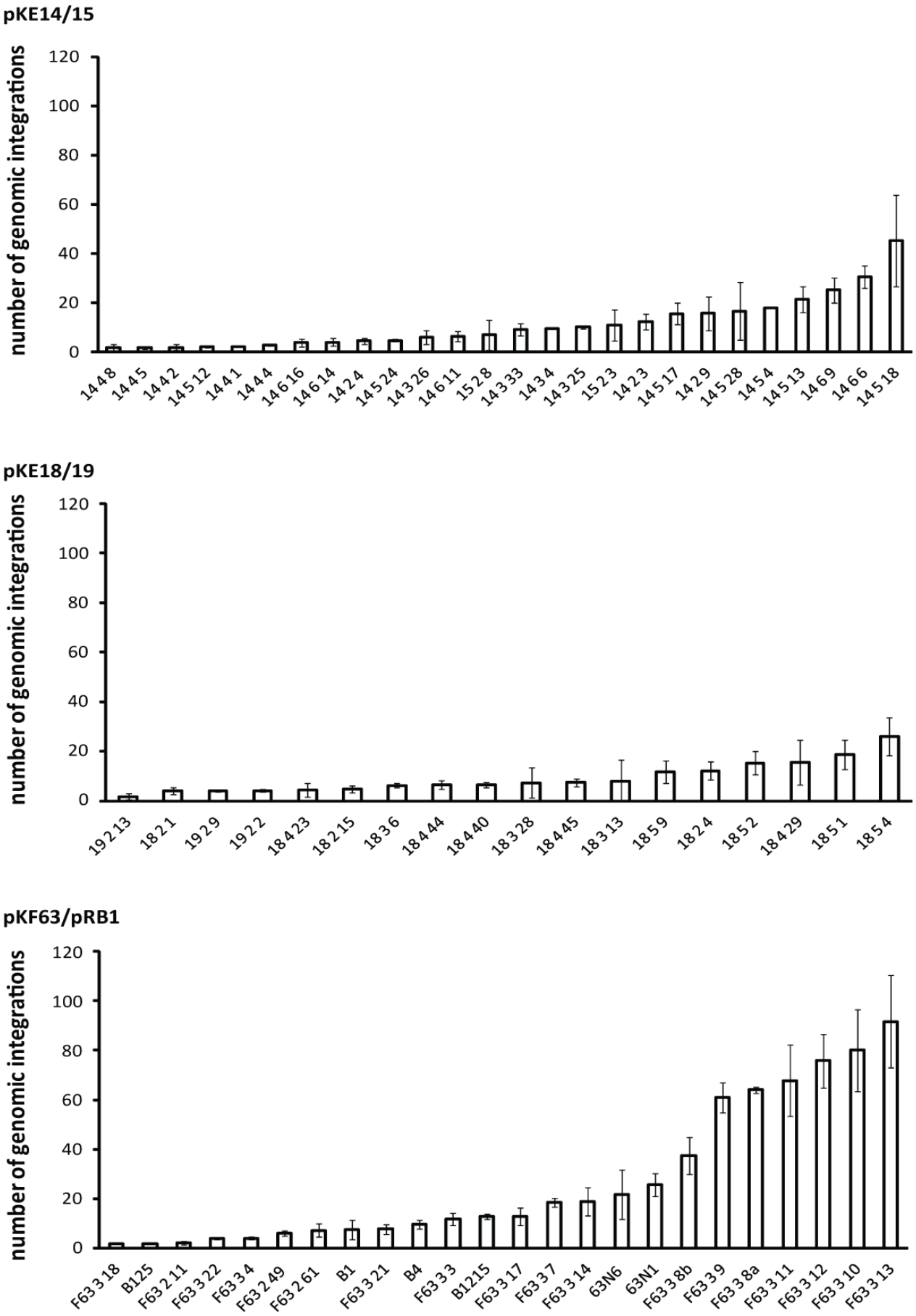


Figure 9/ Table 4. Copy number determination of stably transfected monoclonal cell lines carrying different transgenes. All qPCR results were determined at least in triplicates, error bars reflect standard deviation. Reporter clones carrying a construct with SV40 poly A signal (pKF63) are named starting with F63, clones carrying pKE14 start with 14 and clones with a construct bearing both 3'UTRs start with 18;clones carrying the Luciferase reporter pRB1 start with a B.

histone stem loop
pKE14/pKE15

clone	copy#	stdev
14 4 8	1.63	1.58
14 4 5	1.66	0.59
14 5 12	2.11	0.48
14 4 1	2.14	0.09
14 4 4	2.84	0.25
14 6 16	3.75	1.69
14 6 14	4.00	1.55
14 2 4	4.47	1.14
14 5 24	4.69	0.49
14 3 26	6.10	2.86
14 6 11	6.34	2.27
15 2 8	6.96	6.21
14 3 33	9.14	2.51
14 3 4	9.57	0.08
14 3 25	10.07	0.57
15 2 3	10.75	6.31
14 2 3	12.32	3.13
14 5 17	15.56	4.35
14 2 9	15.65	6.76
14 5 4	17.72	0.46
14 5 13	20.93	4.66
14 6 9	28.61	7.19
14 6 6	30.58	4.57
14 5 18	42.08	9.75

both 3' UTRs
pKE18/pKE19

clone	copy#	stdev
19 2 13	1.83	1.32
18 2 1	4.12	1.49
19 2 9	4.12	0.25
19 2 2	4.23	0.45
18 4 23	4.45	2.64
18 2 15	4.82	1.24
18 3 6	6.33	0.92
18 4 44	6.54	1.80
18 4 40	6.70	1.02
18 3 28	7.36	6.07
18 4 45	7.43	1.49
18 3 13	8.02	8.58
18 5 9	11.64	4.52
18 2 4	12.28	3.71
18 5 2	15.35	4.75
184 29	15.56	9.14
18 5 1	18.73	6.00
18 5 4	25.99	7.56

SV40 poly A signal
pKF63/pRB1

clone	copy#	stdev
F63 3 18	1.65	0.69
B125	1.73	0.38
F63 2 11	2.05	0.47
F63 3 22	3.84	1.16
F63 3 4	4.10	2.71
F63 2 49	6.09	3.94
F63 2 61	7.27	1.92
B1	7.67	1.84
F63 3 21	7.77	2.36
B4	9.55	1.12
F63 3 3	11.80	3.47
B1215	12.79	1.86
F63 3 17	12.86	5.77
F63 3 7	18.58	10.00
F63 3 14	18.88	4.72
63N6	21.77	7.38
63N1	25.63	6.06
F63 3 8b	37.46	1.27
F63 3 9	60.90	14.37
F63 3 8a	64.08	10.71
F63 3 11	67.79	16.45
F63 3 12	75.72	18.76
F63 3 10	80.00	24.90
F63 3 13	91.62	16.61

We further validated the relative copy number by Southern Blot for a subset of clones. Figure 15 shows a Southern Blot of genomic DNA after digestion with NotI and BamHI to cleave before and after the GFP coding region. The upper band shows the blot after hybridization with a radioactively labeled 700 nt long GFP probe (see M&M 3.2.3.3) the lower band serves as loading control and shows the blot after hybridization with a GAPDH probe. The intensity of the GFP bands were normalized to GAPDH. As in this experiment the coding sequence of the reporter was detected instead of the Ub promoter, it was not possible to define the intensity for one copy per cell. The data was therefore normalized to the first clone depicted in Figure 15 (F63 3 17) and set to the copy number determined by qPCR (12). Relative changes in the Southern Blot band intensity of other clones are shown next to relative changes determined by qPCR. Both methods give comparable results.

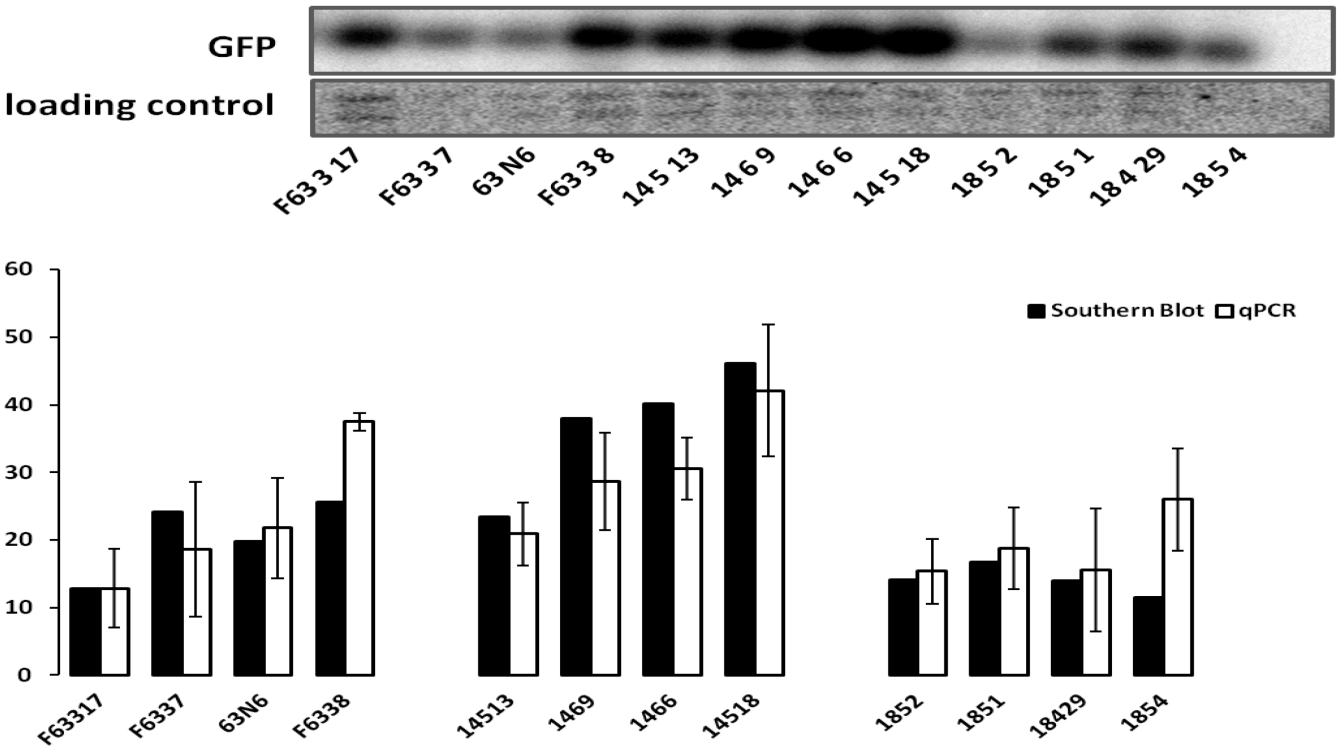


Figure 15 Comparison of copy number determination via qPCR and Southern Blot. genomic DNA was isolated from stably transfected clones, digested with BamHI and NotI and hybridized with a radioactively labeled probe corresponding to the GFP coding sequence in a Southern Blot. Copies numbers of the GFP coding sequence was determined normalized to GAPDH. The same clones were used as a template for qPCR of the plasmids promoter normalized to GAPDH. As qPCR in contrary to Southern Blot gives an absolute value for as copy number in our experiment, the Southern Blot outcome was normalized to one of the clones in the diagram (F63 3 17). Changes in detection intensity relative to F63 3 17 are depicted in the diagram next to qPCR values with error bars.

4.2.2 TRANSGENES INTEGRATED IN DIFFERENT INTEGRATION MODES

As the mode of integration could theoretically play a role in endo-siRNA generation against a certain sequence, a set of clones was analyzed regarding their ratio of disperse integration to concatemere integration. Stable integration of a transgene into the chromatin of a cell can either occur scattered throughout the genome or it can occur in a cluster confined to a small portion of the genome. Clustered concatemere integration is most likely a result of a rolling circle replication event during stable integration or catenation of the plasmid DNA during exponential growth of the host bacteria. If the transgene is

inserted into the genome in concatemeres, digestion with one restriction enzyme cutting shortly upstream of the hybridization site of the probe (as BamHI in this experiment) would result in fragments of the same size and corresponding to the linearized vector, which can be detected with the probe.

Figure 16 shows a Southern Blot of genomic DNA of different clones digested with BamHI and hybridized with a probe detecting the GFP coding sequence. The distinct band corresponding to the size of linearized pKF63 (lane 3) represents fragments with equal size formed by concatemere integration and digestion. These were normalized to the intensity of the whole lane and the percentage of concatemere integration was estimated. Some of the clones show a higher percentage of concatemere integration and some do not show an accumulation of fragments of the same size at all (F6337, 14518 and 1852).

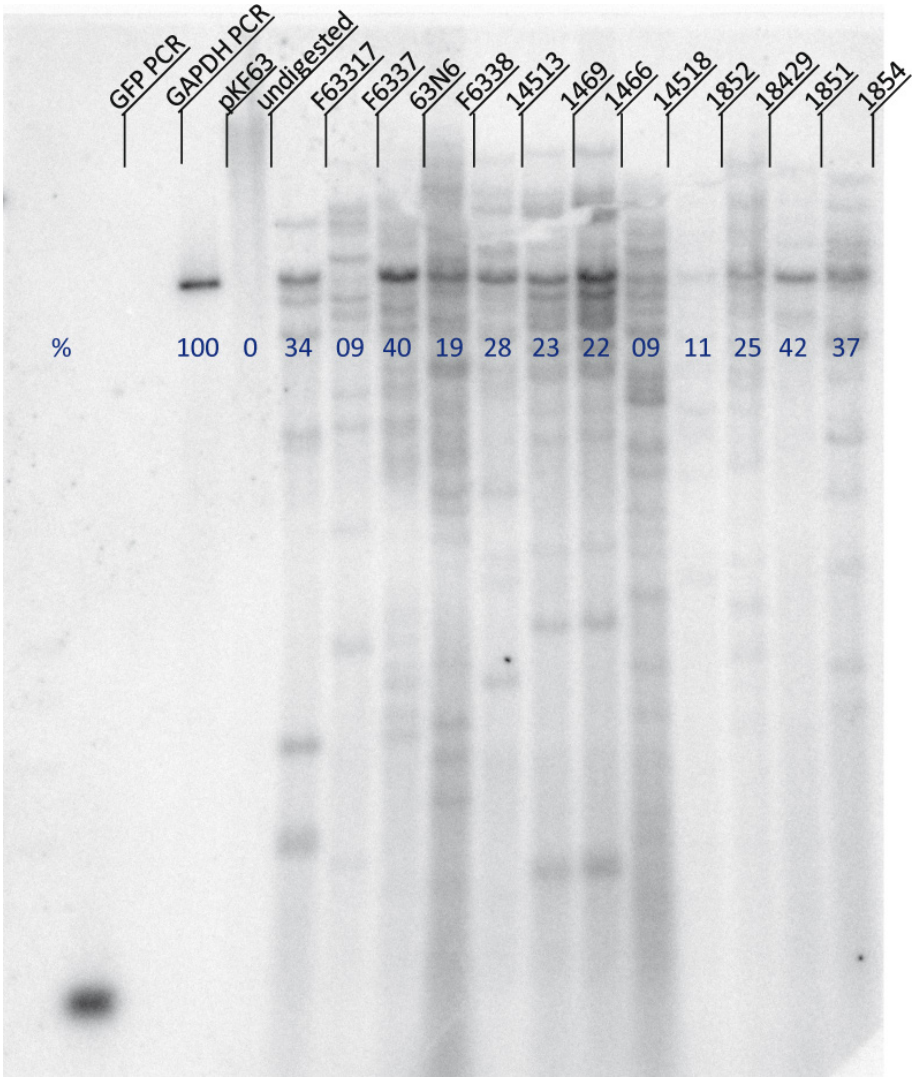
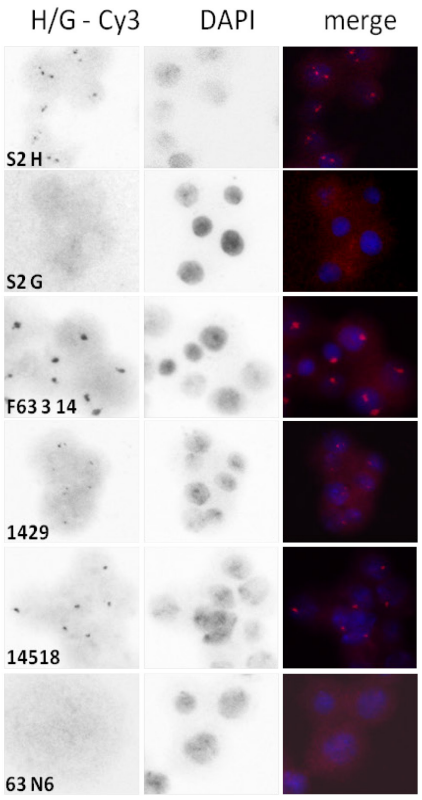


Figure 16 **Southern Blot illustrating a diverse integration mode of the reporter transgene.** Genomic DNA was isolated from different clones, digested with BamHI and hybridized with a probe corresponding to the GFP coding region. The first lane is loaded with a PCR fragment of the GFP coding region and the third lane is loaded with the reporter plasmid digested with BamHI. The blot is evaluated with multi gauge, background was substracted from the histograms of the lanes and the intensity of the prominent band corresponding to the linearized plasmid was expressed in percent of the whole lane.

Another possibility to examine the integration mode of stable transgenes is fluorescence in situ hybridization (FISH). Thereby clustered but not necessarily concatemere integration can be detected by hybridization of the chromatin with a fluorescently labeled probe. As a positive control for detection of a clustered gene locus probes complementary to the histone H2a and histone H3 coding sequence were employed. Histone genes are encoded in up to 100 copies in the histone locus cluster. In 67 out of 82 nuclei from 4 independent experiments two labeled clusters could be detected and altogether a mean value for cluster per cell of 1.9 was calculated (Figure 17). If the transgene integrated in a way analogous to the histone cluster a pattern similar to that of the histone genes should be visible with probes hybridizing with the inserted transgene (listed in M&M 3.1.7.4). In Figure 17 the number of nuclei employed for this experiment is listed together with the number of clusters counted per cell. FISH of every clone shown here was performed in independent experiments as described in M&M chapter 3.2.3.4, S2 cells treated with GFP probes were used as a negative control and hybridization with probes against the histone cluster were used as a positive control.



		clusters / cell						
clone	nuclei	0	1	2	3	experiments	mean	stdev
F63 3 3	44	3	35	4	2	3	1.1	0.58
F63 3 7	54	6	45	3	0	3	0.9	1.55
F63 3 8	32	3	24	5	0	2	1.1	0.5
F63 3 14	66	6	53	5	2	3	1	0.54
F63 3 17	35	11	22	2	0	2	0.7	0.56
63 N 6	56	54	2	0	0	4	0	0.19
14 2 9	24	5	16	2	1	2	1	0.68
14 5 13	23	18	2	0	3	2	0.5	0.7
14 5 17	63	15	47	1	0	3	0.8	0.44
14 5 18	62	13	43	6	0	3	0.9	0.55
14 6 6	43	35	8	0	0	3	0.2	0.38
14 6 9	42	7	31	4	0	3	0.9	0.5
18 2 4	35	34	1	0	0	3	0	0.17
18 5 1	41	6	31	4	0	3	1	0.49
18 5 4	40	9	31	0	0	3	0.8	0.42
S2 histone	82	0	10	67	5	4	1.9	0.4

Figure 17/Table 5 **Fluorescence in situ hybridization of the stably integrated transgene.** On the left, examples of the FISH experiments are shown; the very left column shows the red channel with the Cy3-labeled probe, the middle column shows DAPI and the right one an overlay of Cy3 and DAPI. In the first row the positive control, a probe hybridizing to two loci within the histone locus cluster is shown, two dots per cell can be detected, the second row shows the negative control, a probe hybridizing to the transgene brought into the parental cell line S2. Three clones showing dots of different sizes are depicted next together with one clone showing no clusters.

4.3 ENDO-SIRNA PRODUCTION CORRELATES WITH NUMBER OF GENOMIC INSERTIONS

4.3.1 GENERATION OF DSRNA DIRECTED AGAINST AGO2 AND DCR2

It was shown before (Hartig et al., 2009) that stably integrated transgene in high copy numbers can elicit an endo-siRNA response against its sequence. To investigate the correlation of endo-siRNA production against the inserted transgene with its number of insertion, proteins responsible for endo-siRNA biogenesis were depleted by RNAi to observe the resulting de-repression of GFP fluorescence. RNAi was performed by adding double stranded RNA (dsRNA) to the cell culture medium as described in M&M chapter 3.2.2.2. Reduction of the Ago2 level was verified by Western Blot (Figure 18). As a positive control for every knock down experiment GFP was depleted with the corresponding dsRNA, an unrelated dsRNA against the *Discosoma striata* red fluorescence protein (Dsred) served as a negative control. All knock down experiments were performed at least in triplicates.

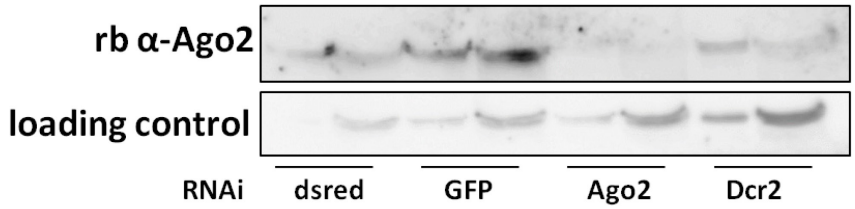


Figure 18 Western Blot of Ago2 after RNAi against different components. Ago, Dcr-2, GFP or Dsred were knocked down by soaking with dsRNA twice. Protein extract was analyzed by Western Blot against Ago2 to give information about protein content after knock down. An unrelated antibody was used for loading control.

4.3.2 KNOCK DOWN OF ENDO-SIRNA BIOGENESIS FACTORS LEADS TO INCREASED REPORTER ACTIVITY

If endo-siRNAs are generated against the transgene and GFP expression is repressed, depletion of the endo-siRNA biogenesis factors Dcr-2 and Loqs-PD or depletion of the effector nuclease Ago2 should lead to a de-repression of GFP fluorescence. In this experiment clones with different copy numbers of inserted transgene pKF63 were depleted of Ago2, Dcr-2 or Loqs-PD and analyzed by flow cytometry (see M&M chapter 3.2.5.7). The fold change of fluorescence intensity after depletion of Ago2 was compared to Dsred and plotted against the copy number determined by qPCR. No direct linear correlation between copy number and repression intensity could be assigned in these experiments, however, a threshold level could be defined at about 10 copies per cell, which was assigned based on a significant ($p<0.05$ students unpaired T-test) change in the mean fluorescence intensity (MFI) of GFP. Clones with no significant change in MFI upon Ago2 knockdown were classified as non-responders (black); clones which significantly increased in GFP intensity upon knockdown of Ago2 were classified as responders (red).

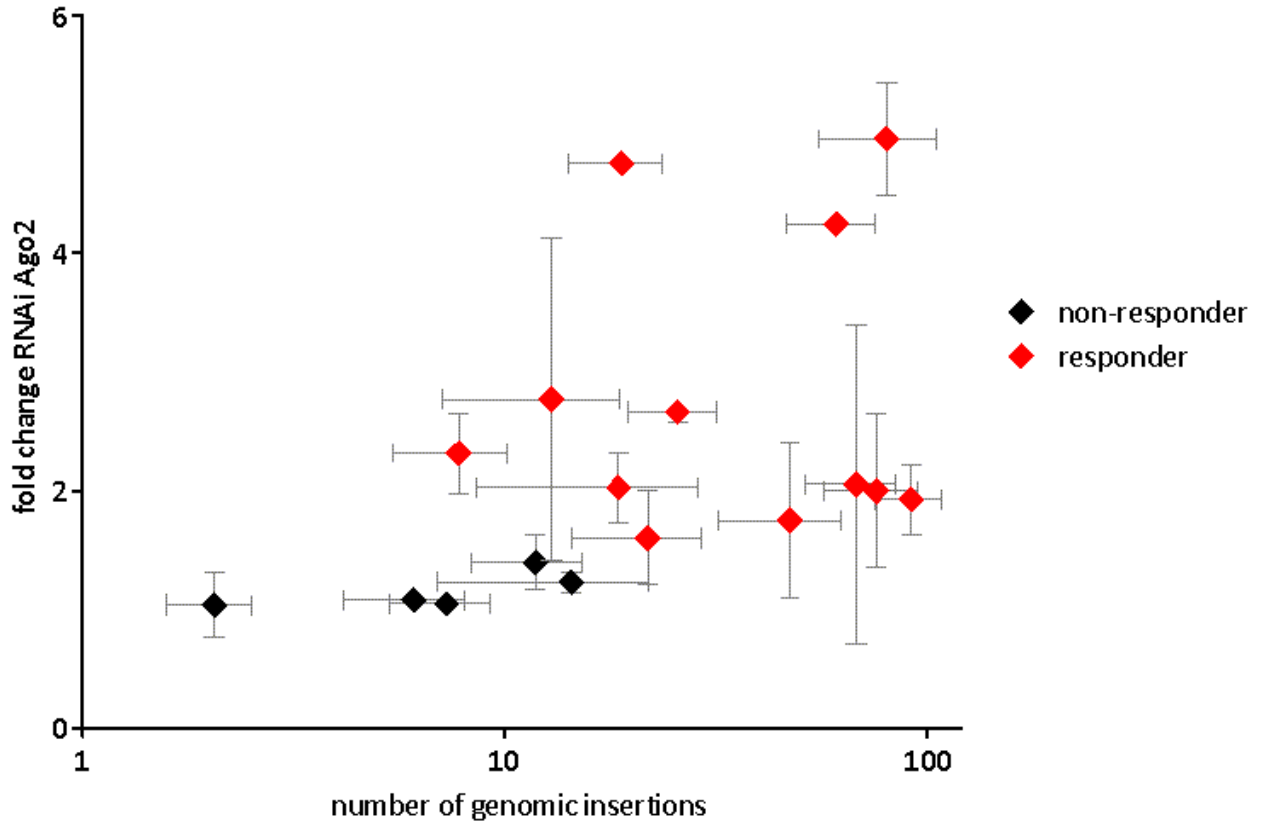


Figure 19 Scatter Plot correlating the number of genomic insertions to the level of de-repression after RNAi against Ago2. Different clones were depleted of Ago2 and as a control Dsred or GFP. The fold change of GFP intensity upon knock down of Ago2 compared to Dsred was calculated and plotted against copy number determines by qPCR. Clones marked in red showed a significant increase of GFP fluorescence intensity after RNAi Ago2 ($p<0.05$, students unpaired T-test),

4.3.3 REPORTER CELL LINES CARRYING HIGHER NUMBERS OF TRANSGENE INSERTIONS PRODUCE HIGHER LEVELS OF CORRESPONDING SMALL RNAS

Small RNAs generated at the transgene coding site in different clones were analyzed by SOLEXA deep sequencing. Total RNA was extracted from cultured cells and deep sequencing libraries were generated as described in M&M 3.2.4.4. Small RNAs were assigned to each of the clones based on the corresponding barcode and the length was restricted to 21 nucleotides (nt) as summarized in M&M 3.2.4.5. The resulting reads were normalized to genome matching 21nt long reads and mapped to the sequence of the stably integrated plasmid pKF63 or pRB1. The sums of these reads were plotted against the number of genomic insertions of the corresponding clone and are depicted in Figure 20. A linear correlation with a coefficient of determination $R2= 0.837$ was observed.

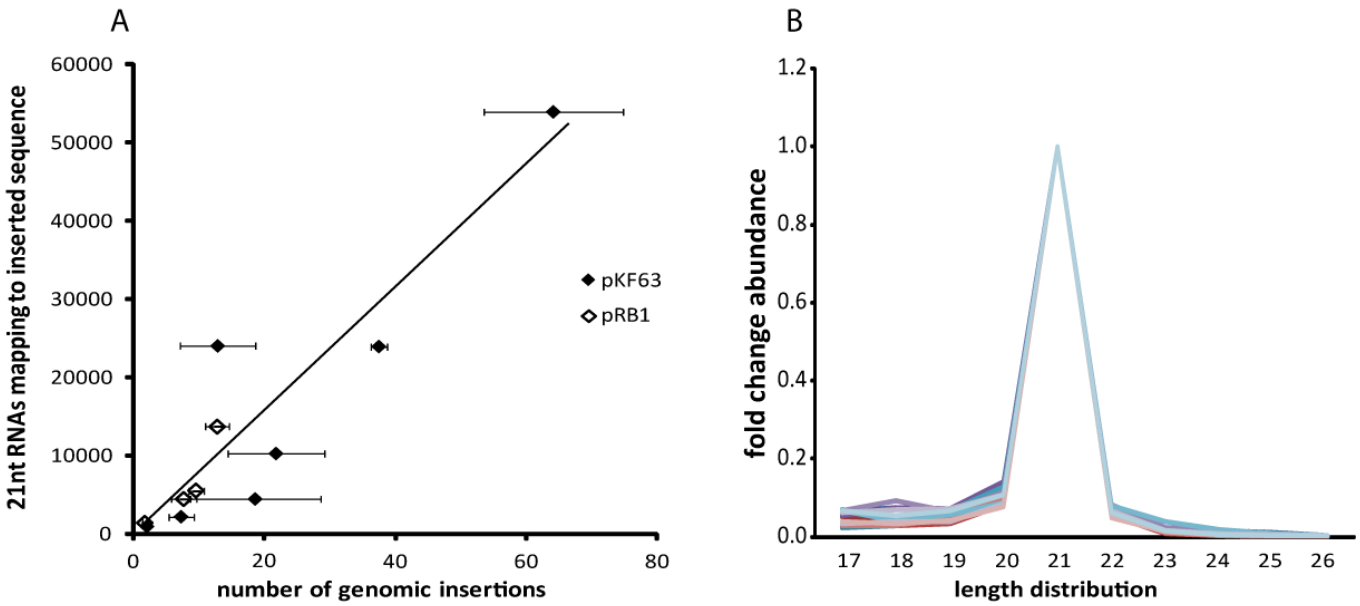


Figure 20 Linear correlation of small RNAs directed against the inserted plasmid and the copy number and length distribution. (A) 21nt deep sequencing reads mapping to the inserted construct were normalized to genome matching 21nt reads of each library and plotted against the number of genomic insertions. A linear regression with a coefficient of determination R^2 of 0.837 was observed, suggesting an increased endo-siRNA response with increasing copy numbers. (B) 17-26nt long RNAs from each library were mapped to the relevant sequence in sense and antisense orientation; all mapped libraries show a peak at 21nt length.

4.4 HISTONE GENES ARE LESS SUSCEPTIBLE TO ENDO-SIRNA REGULATION

Other repetitively encoded loci in the flies' genome besides transposable elements are the histone locus and the ribosomal RNA (rRNA) genes. Transposons are the main target of endo-siRNAs, but what about the other repeated clusters? The rRNA gene is encoded in tandem repeats and is transcribed by RNA polymerase I, whereas histone genes are transcribed by RNA Polymerase II. To investigate a potential endo-siRNA response against these repetitive motifs, the number of 21 nt long deep sequencing reads mapping to the relevant sequence were compared to transposon matching RNAs. Figure 21 shows 21 nt RNA reads from 10 experiments, each normalized to genome matching reads and mapped in a 10 nt interval to the LTR transposon 297, the LINE element juan, the rRNA gene and the histone gene cluster. The outcomes of all experiments were added up to give a representative distribution pattern. Small RNAs matching the sequence of both transposons show an equal distribution of sense and antisense reads; 297 shows a slight enrichment of reads matching the 5'LTR and 3'LTR but no enrichment of the protein coding region. In this example the LINE element juan elicits a lower level of endo-siRNAs with more or less equally distribution of reads along the sequence.

As the amount of rRNA in a eukaryotic cell makes up nearly 90 % of RNA it is inevitable that degradation products can be found in our deep sequencing libraries (< 1 % in every library) and these probably make up

most of the sense reads mapping to the rRNA locus. Antisense reads mapping to the rRNA locus, however show a peak at 21 nt length. Mapping the same pool of reads against the histone locus cluster reveals that these repeating elements are inefficiently targeted by endo-siRNAs, although copy number, transcribing polymerase and transcription rate is comparable to transposable elements. To determine a possible regulatory difference between transposon sequences and histone genes the 3'UTR of the reporter construct was replaced by a histone specific stem loop structure (pKE14/pKE15; see Figure 13). We compared reporter clones carrying the histone 3'UTR to SV40 poly A signal bearing reporters in the following experiments.

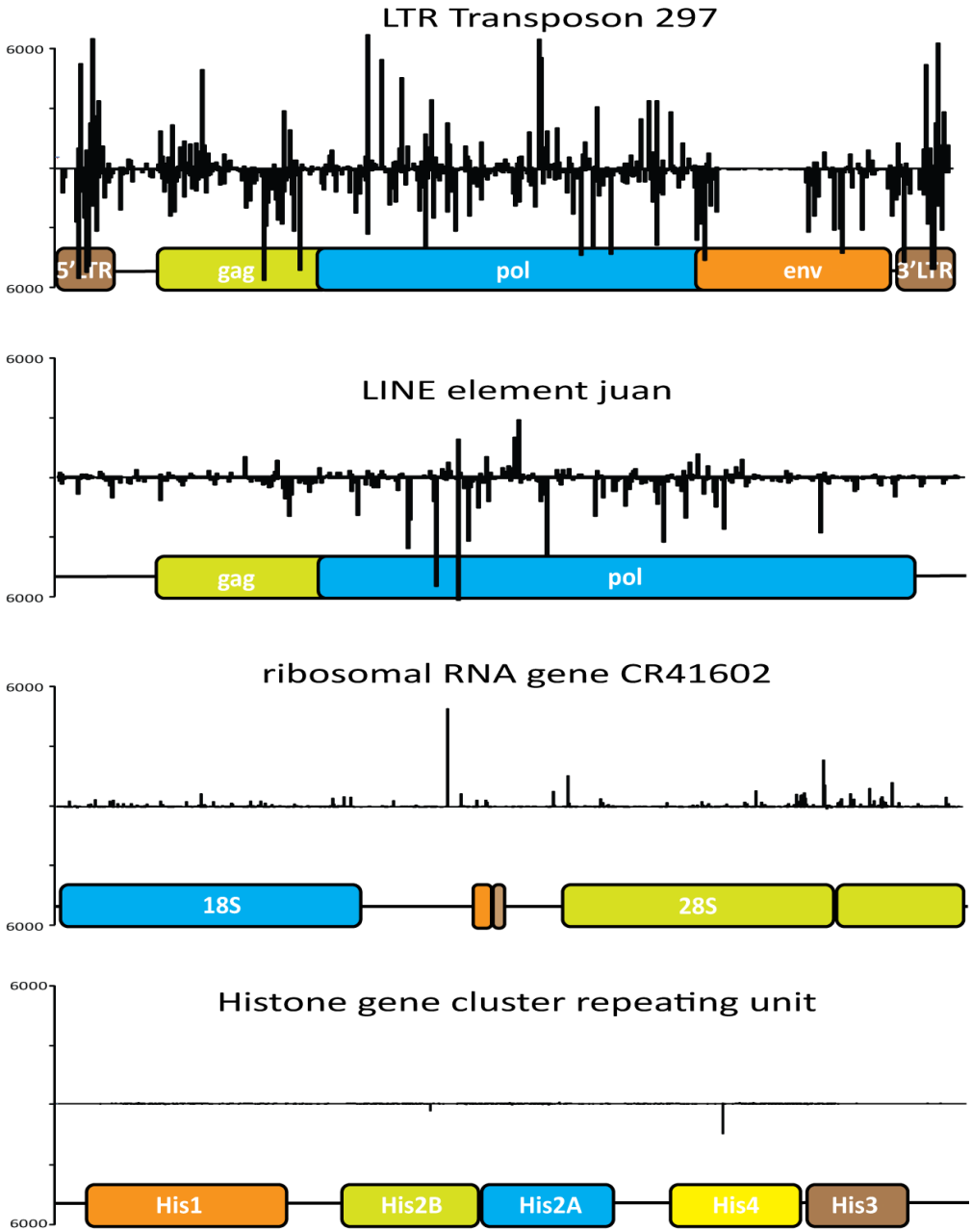


Figure 21 **Endo-siRNA production against repetitive loci in the fly genome.** 21 nt long deep sequencing reads, normalized to genome matching reads were summed up from 10 experiments and mapped in a 10 nt interval to the LTR transposon 297, the LINE element juan, the ribosomal RNA gene CR41602 and the histone gene cluster. The diagrams illustrate the distribution of sense and antisense reads along the indicated sequence, quantification of the sense reads are depicted above the x-axis, antisense reads below. Retrotransposons of the LTR and LINE element classes show a high density of 21 nt endo-siRNAs directed against the sequence with an equal distribution of sense and antisense reads. Reads mapping to the ribosomal RNA show a strong sense bias and therefore most likely represent degradation products. The histone gene cluster is targeted by a small yet clearly present amount of endo-siRNAs

4.4.1 60% OF HISTONE 3'UTR REPORTER PLASMID DERIVED TRANSCRIPTS ARE PROCESSED AT THEIR STEM LOOP

To validate the 3' end processing status of the histone reporter transcripts the ratio of polyadenylated transcripts compared to transcripts without poly A tail was quantified. Total RNA was isolated from a set of clones, genomic DNA was removed using DNase I and the RNA was reverse transcribed using either random hexamers or oligo dT for priming. Levels of the GFP coding sequence were determined by qPCR and normalized to GAPDH. Increased cDNA levels in RT reactions primed with random hexamers compared to oligo dT primed reactions demonstrates a pool of stable, non-polyadenylated transcripts. This change of detection after reverse transcription with random hexamers over oligo dT is depicted in Figure 22. In both histone 3' UTR reporter clones, pKE14 bearing the histone 3'UTR only and pKE18-clones carrying both 3' UTRs subsequently, about 60% of transcripts are processed from the histone stem loop.

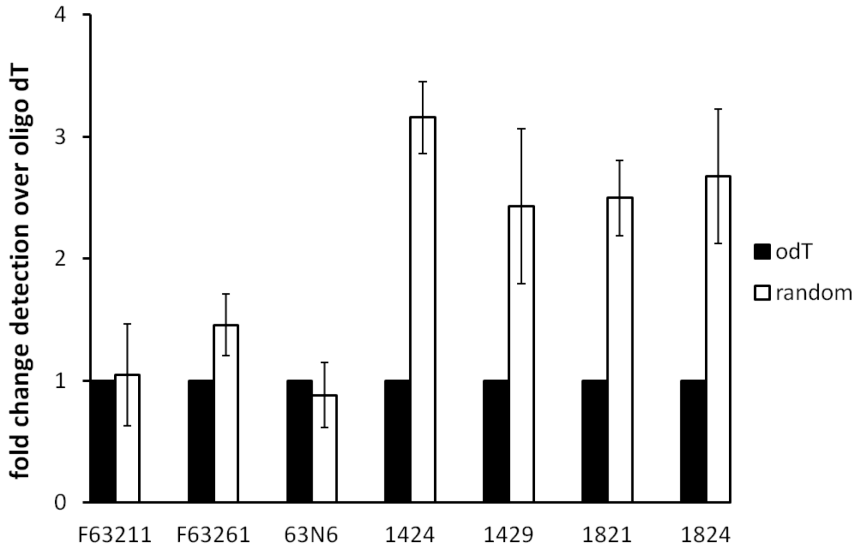


Figure 22 qPCR of Reporter transcripts comparing reverse transcription with oligo dT to reverse transcription with random hexamers. RNA from clones was reverse transcribed with oligo dT primer or random hexamers, a fragment within the GFP coding sequence was amplified in a q PCR and analyzed using the $2^{-\Delta\Delta Ct}$ method normalized to GAPDH. The diagram shows the average of three independent experiments and the data for each clone is shown normalized to the amount of transcripts detected after RT with oligo dT. About 60% of the transcripts from clones containing the pKE14 or pKE18 are processed at their histone stem loop. Transcripts isolated from pKF63-clones could be amplified in equals amounts using random hexamers or oligo dT.

4.4.2 HISTONE 3'UTR-REPORTER CONSTRUCTS SHOW DIMINISHED ENDO-SIRNA SILENCING

To determine the efficiency of an endo-siRNA response against histone stem loop reporters, the magnitude of de-repression was analyzed after depletion of Ago2 by RNAi. The fold change of fluorescence intensity was analyzed by flow cytometry compared to Dsred treated cells and plotted against the copy number of the reporter construct. Figure 22a shows a comparison of pKF63 to pKE14 (A) or pKE18 (B) regarding endo-siRNA silencing. While the increase in fluorescence of pKE14 bearing reporter cells did not exceed a 1.8 fold change upon Ago2 knockdown, the pKF63 reporter clones showed a stronger increase in reporter activity with increasing copy number. Clones carrying pKE18 also show a

reduced de-repression upon Ago2 depletion, but it is not clear whether a stronger de-repression would be observed at copy numbers above 30.

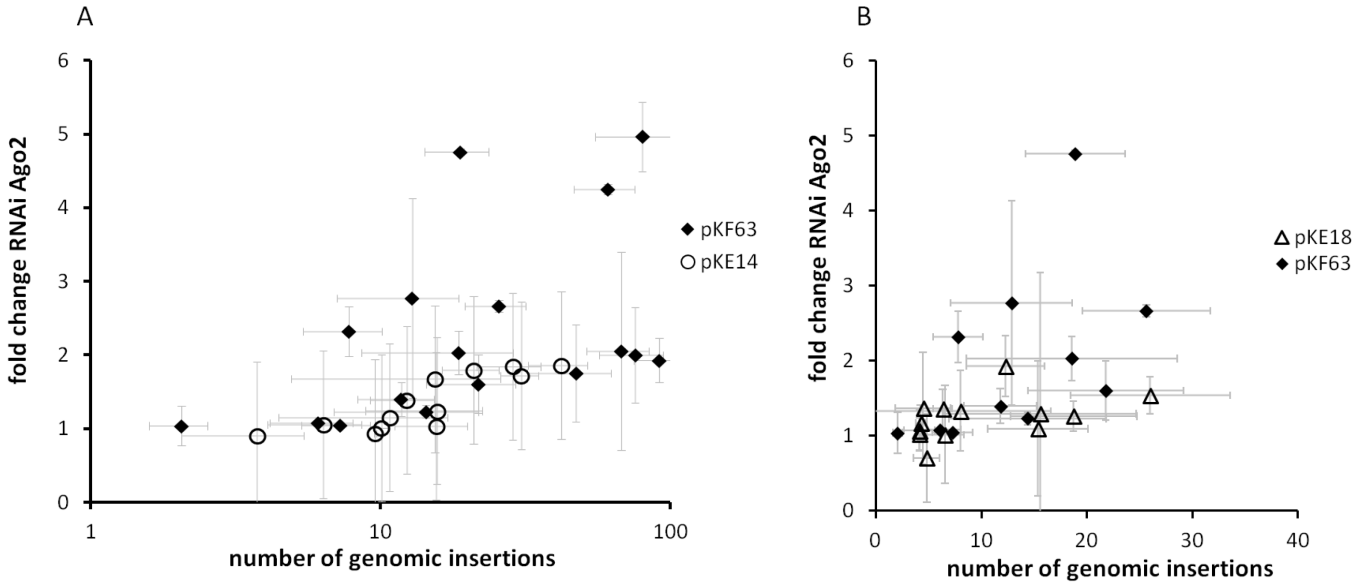
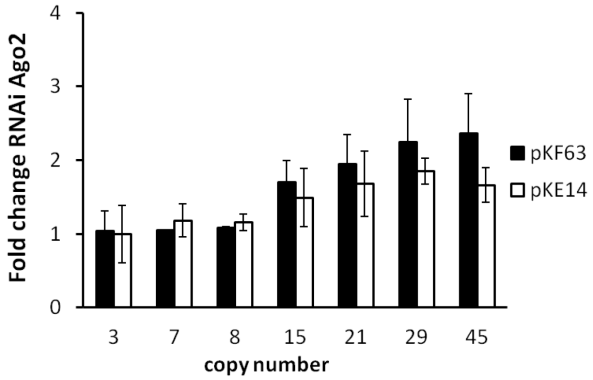


Figure 23 Correlation of copy number and the level of de-repression after RNAi against Ago2. (A) pKE14-clones compared to pKF63-clones (B) pKE18 clones compared to pKF63-clones. Cells were depleted of Ago2 and as a control Dsred or GFP by RNAi. The fold change of GFP intensity upon knock down of Ago2 compared to Dsred was calculated and plotted against copy number determined by qPCR. pKE14 and pKE18- clones do not exceed a de-repression above 2 fold, pKF63 clones can reach a change up to 5 fold.

Comparing reporters with similar copy numbers directly to each other revealed a consistent trend in a stronger response to knock down of Ago2 of pKF63-reporters, however, the difference to pKE14-reporters was not significant as determined by an unpaired student T-test of three independent experiments.



copy#	pKF63	fold change	pKE14	fold change	T-test
3	F63 2 11	1.04	1424	1.00	0.89
7	F63 2 49	1.05	14325	1.18	0.41
8	F63 2 61	1.08	1429	1.15	0.38
15	F6337	1.70	14517	1.49	0.51
21	63N6	1.95	14513	1.68	0.47
29	63N1	2.25	1469	1.85	0.36
45	F63 3 8b	2.37	14518	1.66	0.13

Figure 24 /Table 6 Direct comparison of pKE14 and pKF63 clones with equal numbers of genomic insertions. De-repression of GFP fluorescence intensity upon Ago2 depletion by RNAi was stronger in all pKF63 clones, but the difference did not prove to be significant.

4.4.3 NO CORRELATION OF ENDO-SIRNAS WITH COPY NUMBER IN HISTONE REPORTER CLONES

To further investigate the endo-siRNA response against reporter constructs with histone 3' UTR instead of a SV40 poly A signal, deep sequencing libraries were generated from a set of clones. 21 nt long RNAs were normalized to genome matching reads of each library and mapped to the reporter construct in sense and antisense direction. Sense and antisense reads showed an equal abundance and a peak at 21 nt length, which is congruent with the characteristics of endo-siRNA in the literature. Plotting the number of endo-siRNAs mapping to the inserted plasmid to its copy number in each clone did not a show copy number dependent increase in endo-siRNA production. Clones can rather be divided into responders and non-responders whereby very low copy numbers show no response (Figure 25 A). Although a minimal copy number is required for endo-siRNA production the correlation cannot be described with a linear slope. Only 4 clones of the group bearing both 3'UTRs were sequenced, for these both scenarios are thinkable, at lower copy numbers these clones behave like pKF63 clones, but due to the small number of sequenced clones the further slope of the curve cannot be predicted (Figure 25 B).

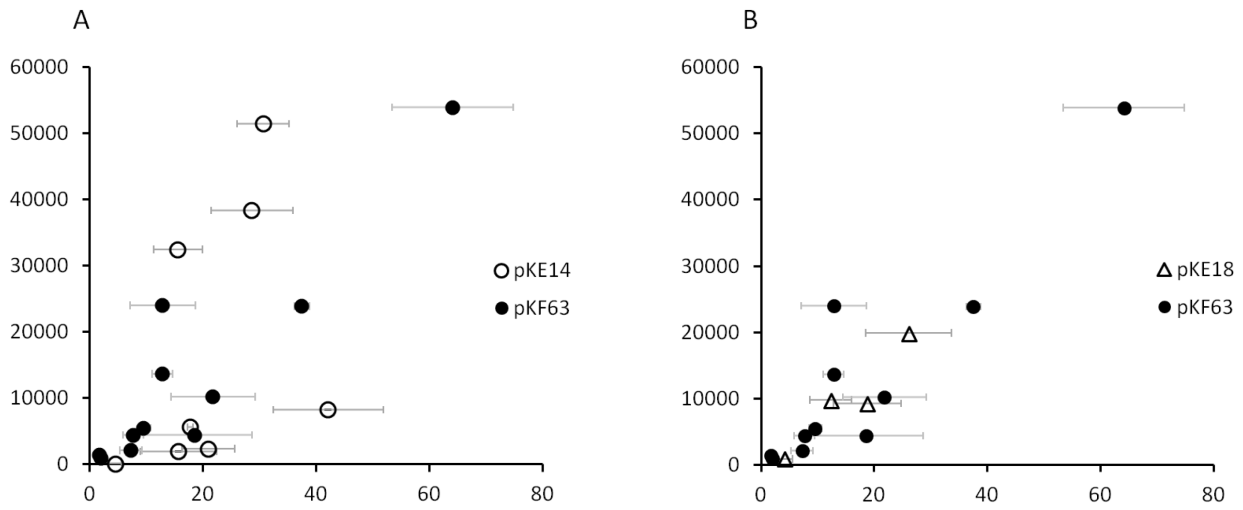
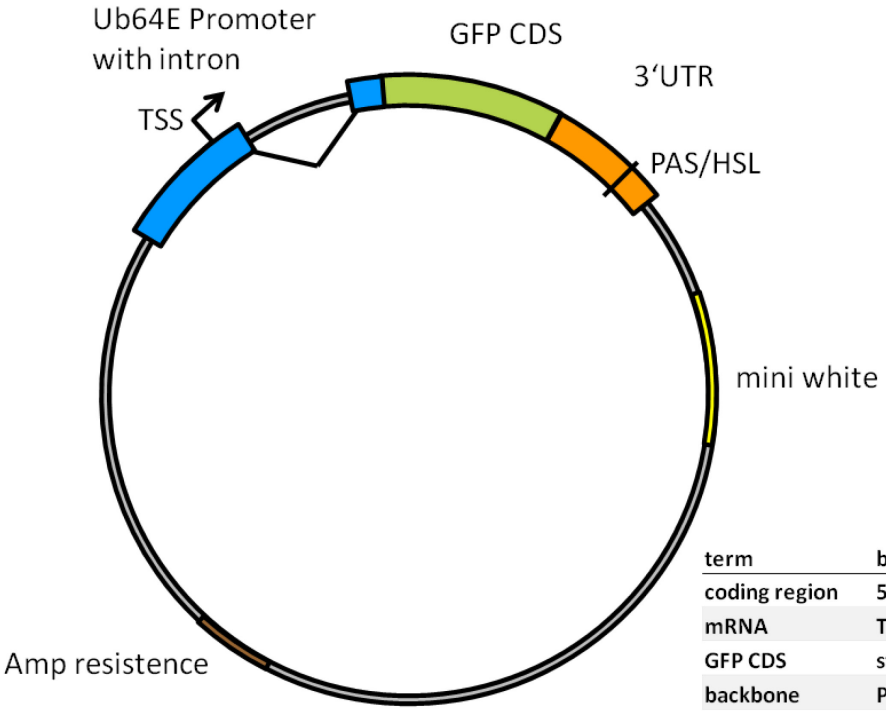


Figure 25 Correlation of small RNAs directed against the inserted plasmid with copy number. 21 nt deep sequencing reads mapping to the inserted construct pKE14 were normalized to genome matching 21 nt reads of each library and plotted against the number of genomic insertions. (A) pKE14-clones did not show a linear correlation of endo-siRNA production with increasing copy number compared to pKF63 clones. (B) No linear correlation could be definitely assigned to pKE18-clones, both scenarios are conceivable for the hybrid clones.

4.5 CHANGES IN 3' END PROCESSING SIGNAL AFFECT THE DISTRIBUTION OF ENDO-SIRNAS ALONG THE PLASMID SEQUENCE

Only small RNAs mapping to the mature target mRNA will lead to a post-transcriptional repression of the reporter by the cytoplasmatic RISC complex. We therefore divided the plasmid matching deep sequencing reads into siRNA either targeting the mature mRNA, the intron or the plasmid backbone.



term	beginning	end	omitting
coding region	5' end of Promoter	3' end of 3'UTR	
mRNA	TSS	PAS/HSL	intron
GFP CDS	start codon	stop codon	
backbone	PAS/HSL	5' end of Amp	

Figure 26 **Schematic representation of the reporter plasmid**. The Ubiquitin 64E Promoter region is depicted in blue, the transcription start side (TSS) and the intron within the promoter are labeled. The GFP protein coding sequence (CDS) is shown in green and the respect 3'UTRs are shown in orange with the approximate Polyadenylation signal (PAS) or histone stem loop. The backbone of the plasmid contains a mini white sequence and an ampicillin resistance cassette for selection.

The Ubiquitin promoter in our reporter constructs contains an intron so we designed the map file classifying the mRNA targeting siRNAs starting at the transcription start site, omitting the intron and ending at the poly A signal or histone stem loop respectively (Figure 26). When endo-siRNAs were analyzed matching the mature target mRNA a correlation of deep sequencing reads with the number of inserted copies could not be observed.

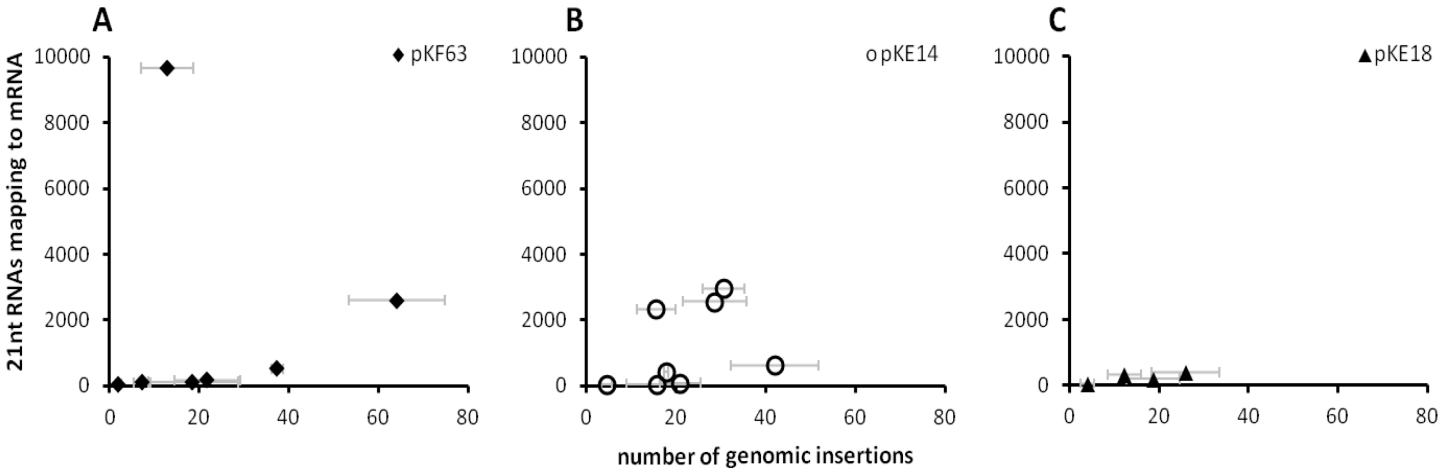


Figure 27 **Correlation of 21nt RNAs mapping to the mature mRNA.** (A) The amount of 21nt RNAs mapping to the mature target mRNA in sense and antisense orientation did not correlate with the copy number of stably transfected pKF63. Figure (B) shows the same experiment with pKE14 carrying cells and (C) with pKE18 carrying cells.

Deep sequencing reads matching the mature mRNA in most cases make up less than 25% of all reads matching the plasmid. We next sought to investigate whether a universal increase in endo-siRNA

production could be observed, correlating with the number of genomic integrations of the plasmid backbone and therefore not subject to cis effects within the transcriptional unit. As the ampicillin resistance cassette within the plasmid is of bacterial origin and the RNA ligase preparation used for library generation might contain traces of bacterial RNA, the map file used to select reads mapping to the plasmid backbone, only contained the sequence starting 3' of the relevant 3' UTR ranging to the 5'end of the ampicillin resistance (Figure 26). 21nt long deep sequencing reads of all clones were normalized to genome matching reads of each library and mapped to the reporter backbone. Sense and antisense matching reads were added and plotted against the number of genomic insertions. Both histone reporter groups, each compared to pKF63 are shown in Figure 28. Compared to pKF63 clones, reporters missing a canonical poly A signal (pKE14) show a reduced, but still roughly linear increase of endo-siRNAs derived from the plasmid backbone. Although correlation of the pKE14 clones was not significant, due to the small number of data points, a trend towards diminished small RNA production could be observed. Endo-siRNAs produced against the inserted pKE14 did not exceed 0.5 % of genome matching reads (5000 ppm) whereas clones with SV40 poly A show endo-siRNA levels against the backbone up to 1-2 %. Reporter clones containing both 3'UTRs react comparable to SV40-clones in this experiment, but due to the small number of data points no strong conclusion can be drawn (Figure 28B).

Altogether reporter clones carrying pKF63 show increased endo-siRNA targeting of the plasmid backbone alone as well as the whole plasmid. For reporter clones carrying pKE14 the amount of small RNAs targeting the backbone increase with reduced augmentation, but looking at the whole plasmid, the number of targeting siRNAs cannot be correlated with copy numbers. The property causing this difference therefore must lies within the insert of the plasmid, the coding region of the histone reporter transcript.

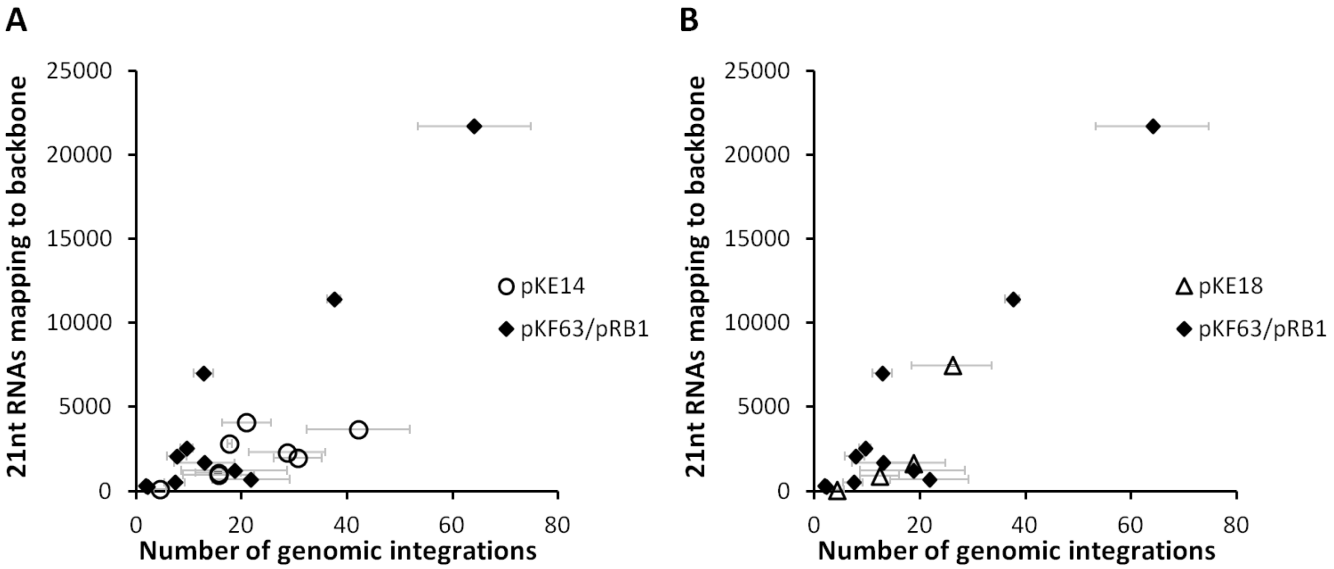


Figure 28 Correlation of small RNA production against the plasmid backbone with its copy number. The plasmid backbone elicits an increasing level of endo-siRNAs with increasing copy number regardless of the inserted reporter construct. A histone 3'UTR containing construct might show a flatter curve. A) Mapping 21nt reads from pKF63 containing clones to the plasmid backbone showed a correlation with number of genomic insertion with a Pearson's correlation coefficient of 0.96.Due to the low number of data points representing reporter clones containing the histone 3'UTR the correlation of small RNA production with copy number was not significant (Pearson's r= 0.68). (B) Clones containing both 3'UTRs show an increase of endo-siRNA production similar to pKF63 clones, with a Pearson's correlation of 0.88.

4.6 HISTONE REPORTERS ACCUMULATE ENDO-SIRNA AGAINST AN INTRON IN THE 5'UTR

Taking a closer look at the exact distribution of deep sequencing reads along the reporter sequence, we noticed that the intron within the 5' UTR gave rise to a tremendous amount of endo-siRNA. This phenomenon could predominantly be observed in clones containing the reporter with the histone stem loop. In contrast, previous deep sequencing libraries from our lab (Hartig et al., 2009) and others (Okamura et al., 2008a) revealed a reduced prevalence of endo-siRNA at intron coding sites, suggesting that in those cases spliced mature mRNA serves as precursor for siRNA generation. In Figure 29 the percentage of siRNAs mapping to the mature mRNA and the intron are depicted, normalized to sequence length. Thereby all 21 nt reads mapping to the whole plasmid constitute 100 % for each clone. The largest group of clones with over 75 % of intron mapping reads belongs to the pKE14 family of reporters, whereas clones mainly targeting the mRNA are exclusively pKF63-reporter cells. Figure 30 shows the distribution of sense and antisense matching deep sequencing reads along the reporter coding sequence of the transgene. A schematic representation shows the boundaries of the intron within the 5' UTR of the reporter. The transcription start site predominantly used was revealed by sequencing multiple cDNA clones and is marked with an arrow.

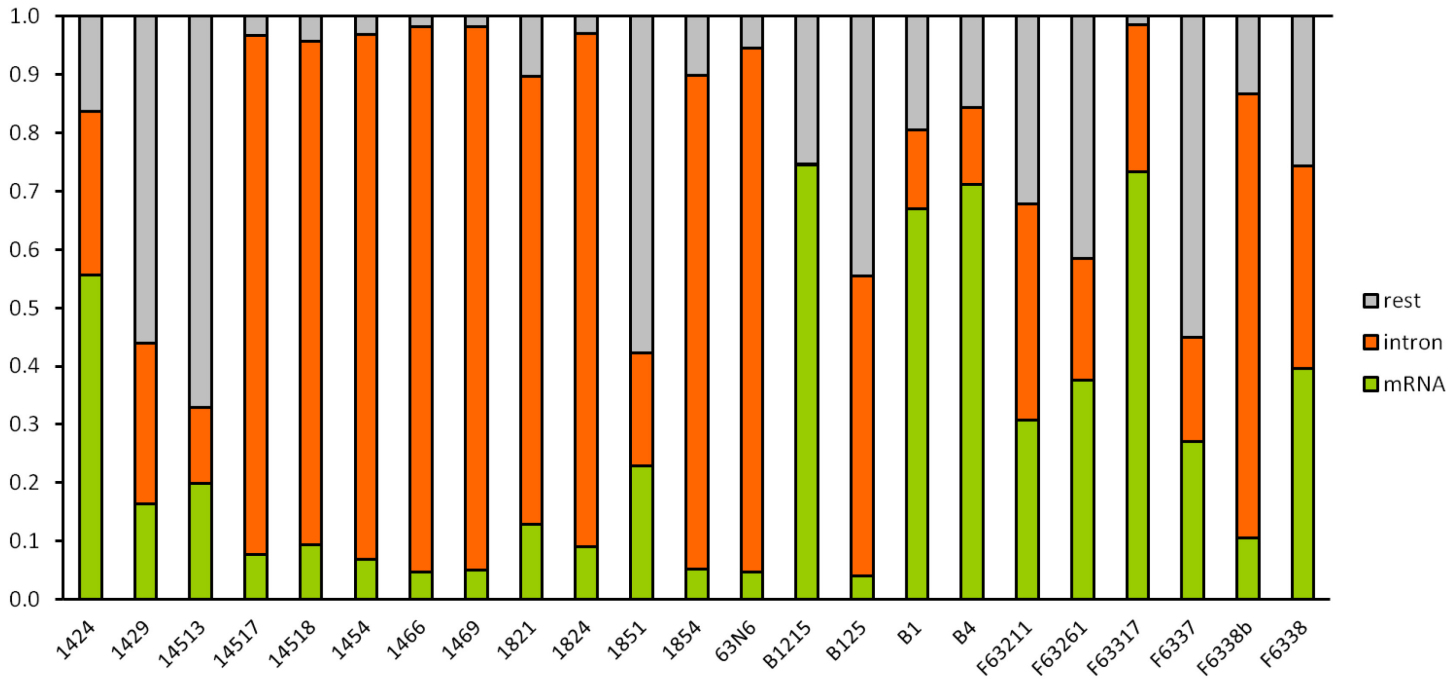


Figure 29 **Percentage or reads targeting the insert or mature mRNA.** Sense and antisense reads mapping to the plasmid were normalized to genome matching reads of each library. The percentage of reads mapping to the mature mRNA, the protein coding sequence of GFP or Luciferase or the intron are listed as percentage of plasmid matching reads. Numbers are normalized to length of each map file.

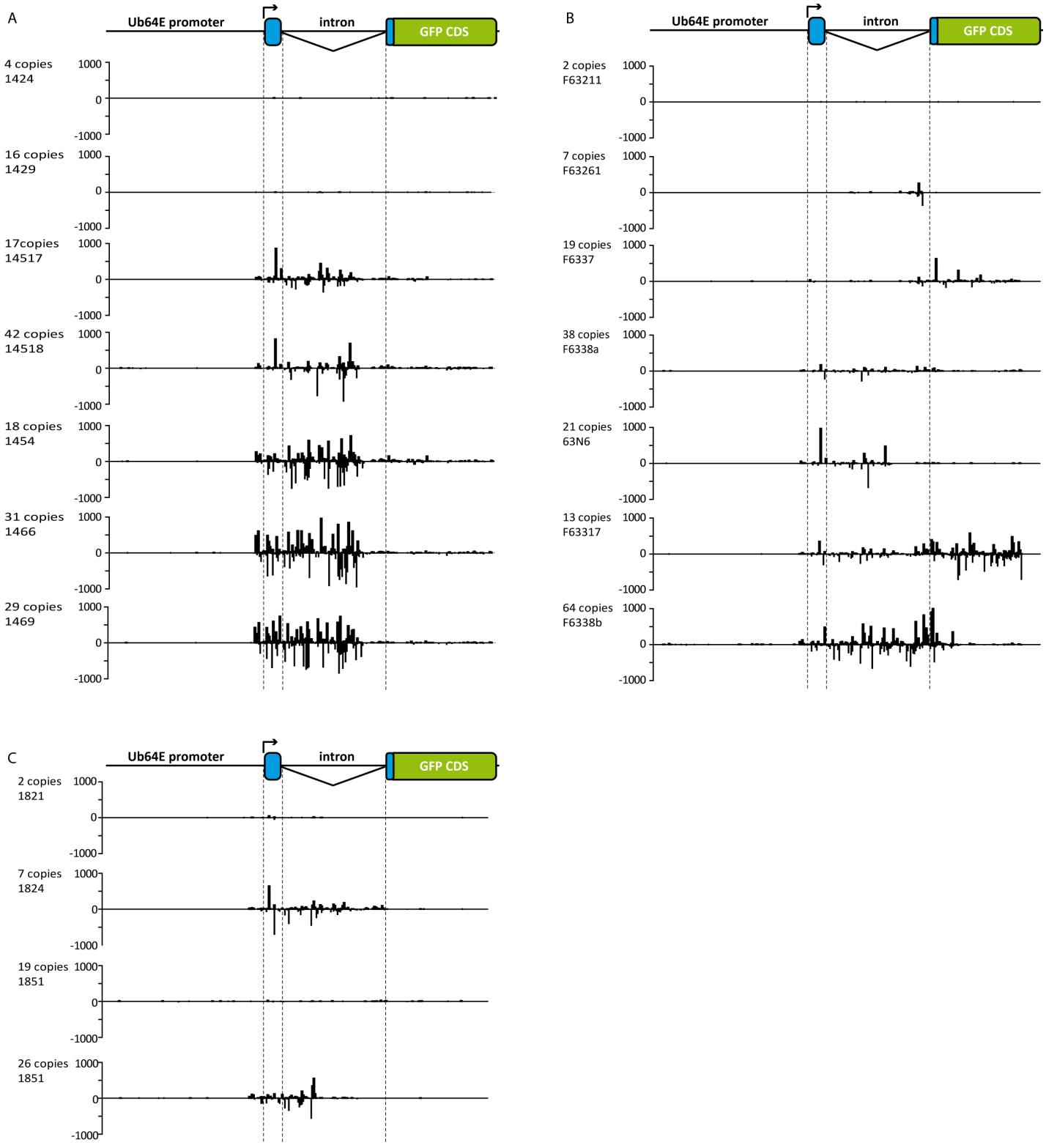


Figure 30 **Mapping of deep sequencing reads to the integrated plasmid sequence.** Quantification of sense reads targeting the illustrated sequence (x-axis) in a 10 nt interval are depicted above the x-axis, antisense reads are depicted below. The transcription start site and splicing patterns are schematically represented above the diagrams. (A) 21nt long RNAs sequenced from pKE14-clones (B) 21nt long RNAs sequenced from pKF63-clones (C) 21nt RNAs sequenced from pKE18 clones. pKE14-clones show a remarkable accumulation of siRNAs directed against the intronic sequence downstream of the transcription start site (TSS).

4.6.1 INTRON TARGETING SIRNAS CAN BE DETECTED BY NORTHERN BLOT

To exclude the possibility that these intron mapping small RNAs reflect a sequencing artifact, a Northern Blot was performed using a probe directed against 30 nt within the intron region eliciting the highest amount of small RNAs in most of the clones. The probe was designed hybridizing to antisense matching RNAs to avoid detection of pre-mRNA degradation products. The predominant length of intron matching reads was 21 nt and the band detected by Northern Blot correlated in length with the DNA oligonucleotide of 21 nt, which was loaded as a positive control. A weak band detecting intron matching siRNAs can be seen for high copy pKF63 reporter clones, higher levels of intron derived small RNAs can be observed in two clones of both histone reporter constructs, pKE14 and pKE18 each.

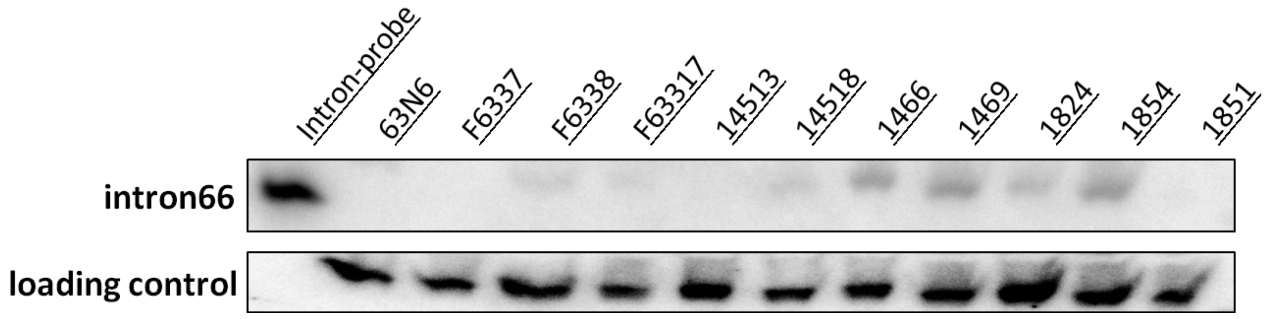


Figure 31 Northern Blot detecting small RNAs matching the intron. The probe detecting siRNAs derived from the intronic region was 30 nt long and designed detecting RNAs with antisense orientation. The region within the intron detected starts 66 nt downstream of the 5' splice site. Intron mapping small RNAs could be detected in the same clones, which showed an accumulation in deep sequencing data. As a positive control, a 21nt DNA oligonucleotide was used, mir277 served as loading control.

4.6.2 REPORTER MRNA IS SPLICED CORRECTLY

It was shown before that the combination of a histone stem loop is not compatible an intron in the same transcript leading to aberrant mRNA processing (Pandey et al., 1990). Pandey and collogues showed that most transcripts combining both features are spliced and splicing in turn directs the transcript to the polyadenylation pathway and usage of a cryptic poly A signal within the HDE. In these experiments the rarely occurring unspliced transcripts all have a histone 3'UTR and therefore Pandey and collogues concluded that the presence of a histone 3'UTR leads to inefficient or slower splicing. qPCR data of our reporter constructs reveal that about 30 % of the pKE14 reporter transcripts are polyadenylated (Figure 22). Correct splicing of our reporter transcripts was verified by sequencing cDNA from multiple clones, confirming the usage of the same exon-intron junction.

We also could not amplify unspliced transcripts from cDNA (Figure 32). As all fragments amplified in the +RT reaction exhibited the same length, we conclude that un-spliced remainders at least do not

accumulate in clones and thus cannot explain the elevated intron targeting phenotype. Nevertheless, we cannot exclude inefficient or stalled splicing in the histone reporter clones compared to reporters with canonical poly A signal. These mis-spliced transcripts and stalled snRNPs might be retained in the nucleus longer than correctly spliced mRNAs and therefore could be a preferred substrate for endo-siRNA generation at least in particular cases.

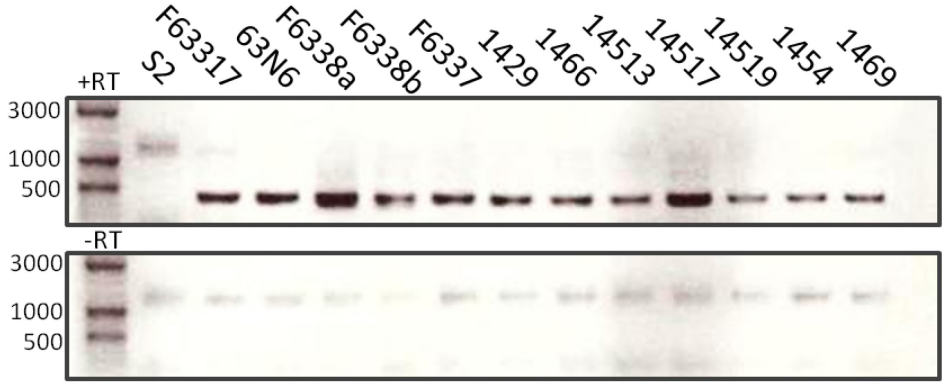


Figure 32 PCR amplifying a fragment containing the exon exon junction of spliced reporter transcripts. A forward primer complementary to a sequence upstream of the 5' splice side and a reverse primer complementary to a sequence downstream of the 3' splice site were used to detect unspliced transcripts beside correctly spliced mRNA. Template RNA from different clones was depleted of gDNA and reverse transcribed with random hexamers. PCR with sufficient amplification cycles resulted in fragments exhibiting the size of correctly spliced transcripts. From the –RT and the negative control, the parental cell line S2, a weak band near background was amplified, which showed the size of the un-spliced fragment.

4.6.3 INTRON DERIVED ENDO SIRNAS ARE SUBJECT TO STRAND SELECTION

In order to investigate whether the intron derived siRNAs are loaded into RISC we analyzed the base pair stability of 5 nt at every 5'-and 3'-end of all intron or mRNA mapping siRNAs in our deep sequencing libraries according to the nearest neighbor method (Xia et al., 1998). siRNA duplexes are loaded into Ago according to a thermodynamic asymmetry rule. R2D2 as a component of the RLC binds to the more stable 5' end of the duplex, determining the strand with the more instable 5' end as the guide strand (Tomari et al., 2004). Ago loaded siRNAs therefore exhibit a bias for weaker base pair stability at their 5' end. Comparing the free enthalpy (ΔG) of both ends revealed that the mRNA derived siRNAs as well as the intron derived siRNAs show a high propensity for a weaker 5' end, which can be deduced from a positive value for ΔΔG (Figure 33A). Comparing the percentage of weaker 5' ends among all ends, especially intron derived siRNAs show a bias towards a weaker 5' end (Figure 33B). To determine whether in addition to the thermodynamic asymmetry a nucleotide bias could be observed in intron targeting or mRNA targeting siRNAs the 5' ends of these RNAs were compared. Figure 33C shows the prevalence of every nucleotide at position 1 of siRNAs mapping to either the mRNA or the intron in sense and antisense orientation. mRNA targeting small RNAs do not show a nucleotide bias at all, whereas intron matching siRNAs show a weak

bias for A at their 5' end, which should be interpreted with caution, because of the low overall GC content (36 %) of the intron. siRNAs were characterized as preferentially starting with a C (Ghildiyal et al., 2008), but other studies predicted no nucleotide bias at any position for endo-siRNAs (Kawamura et al., 2008).

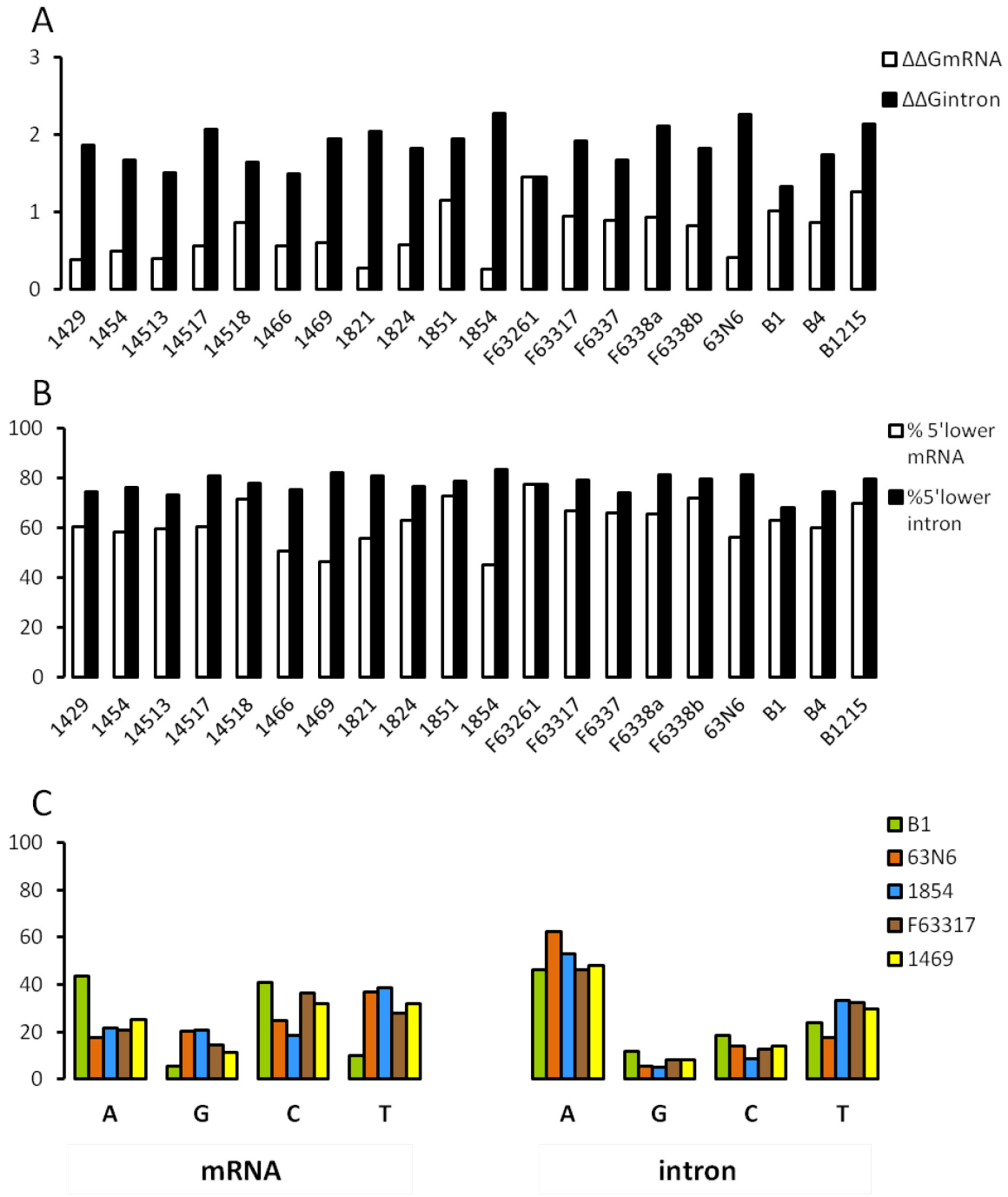


Figure 33 Characterization of base pair stability of siRNA ends and 5' end nucleotide bias. According the nearest neighbor method, the difference in free base pairing enthalpy at the 5' ends of the sequenced siRNAs is lower as at the 3' end, indicating that siRNAs mapping to the intron or the mature mRNA are loaded into the RISC complex. (A) Positive ΔΔG values indicate a weaker 5' end and therefore loading of the siRNA into RISC. (B) 50-80% of the intron and mRNA mapping siRNAs have a weaker 5' end and are presumably loaded. (C) No strong nucleotide bias for the 5' end of intron or mRNA targeting siRNAs could be observed in randomly tested samples.

4.6.4 INTRON DERIVED SMALL RNAS ARE LOADED INTO AGO2

To further corroborate loading of the intron derived siRNAs into RISC and to get a hint for their potential role we a wondered whether intron derived siRNAs are preferentially loaded into Ago1 or Ago2. Small RNAs loaded into Ago2 are 2'O-methylated at their 3' end by the methyltransferase Hen1, which makes them resistant to beta-elimination (Ameres et al., 2011). In contrast, Ago1 loaded small RNAs remain un-modified and are thus susceptible to an experimental removal of the last nucleotide by beta elimination. RNA was isolated from two clones, oxidized with periodate with subsequent pH elevation and hybridized with a probe detecting intron derived siRNAs by Northern Blot. While the Ago1 loaded miRNA bantam could be beta-eliminated (as seen by a shift in size due to removal of one nt) Northern Blot bands of the intron targeting siRNAs show no difference in size with or without beta elimination (Figure 34), suggesting loading of the intron derived siRNAs into Ago2.

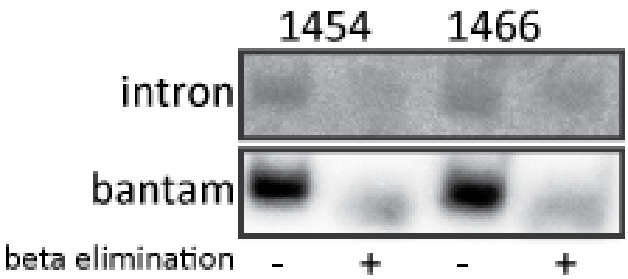


Figure 34 Northern Blot detecting intron derived siRNAs before and after beta-elimination. Total RNA was isolated from 1454 and 1466 clones, beta-eliminated and hybridized with a probe against the intron sequence. Intron derived siRNAs did not show a shift in size, whereas bantam, a miRNA loaded into Ago1 and used as a positive control was shortened by one nucleotide via beta-elimination.

4.7 DOES DEPLETION OF THE LARIAT DEBRANCHING ENZYME REDUCE INTRON DERIVED SMALL RNA ACCUMULATION?

Only recently Dumesic et al showed that in *Cryptococcus neoformans* inefficient splicing and stalled spliceosomes lead to an RNAi answer against the respective pre-mRNA (Dumesic et al., 2013).

Dumesic et al could also show that deletion of the lariat debranching enzyme abolished the increased production of small RNAs. We sought to test whether recently spliced intron-lariats are a preferred substrate for endo-siRNAs against a stable integrated reporter construct in our system. Therefore we depleted the lariat debranching enzyme (ldbr) from our reporter cell lines by RNAi with two different dsRNA designs and detected intron matching 21 nt long RNAs in a Northern Blot (Figure 35B+C). As a positive control Dcr-2 was depleted. Depletion of rrp6 resulted in a slight increase of intron derived small RNAs and both knock down events of ldbr lead to a small decrease of intron derived siRNAs to about 70% of the control. We verified accumulation of the intron-lariat in 4 different reporter clones upon depletion of ldbr. After RNAi-mediated knock down, RNA was extracted and reverse transcribed using random hexamers. qPCR of the intron-lariat revealed a 3 – 4 fold accumulation upon ldbr knock down compared to the negative control in all four clones (Figure 35A). Accumulation of the intron lariat confirms successful

reduction of the protein by RNAi, because debranching constitutes the rate limiting step in degradation of the intronic RNA sequence (Cheng and Menees, 2011) After debranching the RNA fragment is degraded by exonucleases from both sides and can give rise to intron derived snoRNAs or mirtrons. It must be clearly stated here that the changes in intron-derived endo-siRNAs after ldbr knockdown are small. Clearly the experiment should be repeated and further supplemented with deep sequencing data.

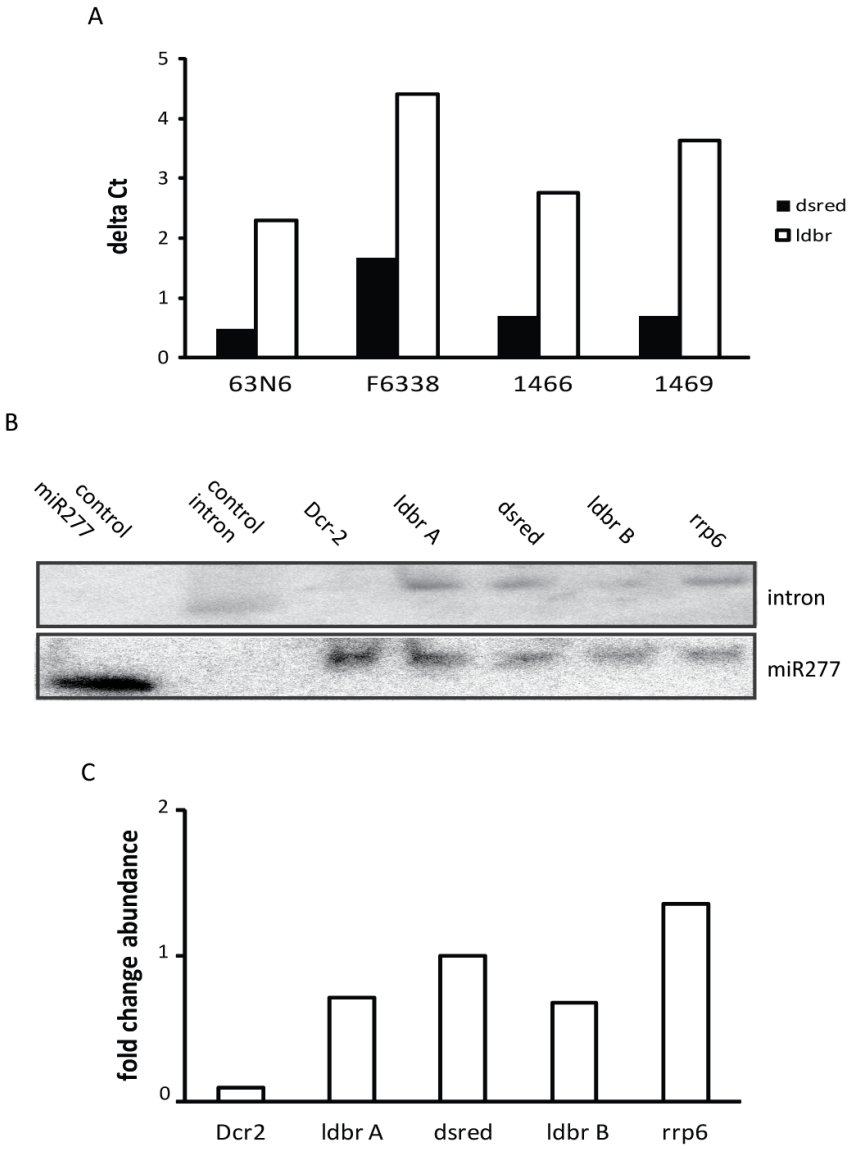


Figure 35 Depletion of the lariat debranching enzyme ldbr lead to an accumulation of intron-lariats and to a reduction of intron derived siRNAs. (A) qPCR amplification of a fragment within the 5'UTR intron of the reporter. cDNA was obtained by RNA extraction after treatment of the cells with dsRNA against ldbr or Dsred, digestion with DNaseI and reverse transcription with random hexamers. Delta Ct values of single experiments are shown for 4 different clones. (B) Northern Blot hybridizing a probe detecting intron derived siRNAs. RNA was isolated from the clone 1466 after treatment with dsRNA against Dcr-2, rrp6, Dsred or two differently regions of ldbr. The miRNA miR277 is used as loading control and for normalization. (C) Diagram analyzing the Northern Blot in B. The intensity of the bands detected with the intron probe was each normalized to miR277. Fold change in abundance of intron siRNAs compared to Dsred are shown for every knock down experiment. Rrp6 leads to increased abundance of siRNA (140%) Dcr2 and ldbr to reduced levels (10% and 70% respectively).

5 DISCUSSION

5.1 CELL CYCLE DEPENDENT MODIFICATION OF RISC COULD NOT BE CONFIRMED

In the first part of my thesis I sought to investigate oscillating patterns of the RNA induced silencing complex (RISC) over the course of the cell cycle. I used the SILAC labeling technique in combination with cell cycle phase separation to compare the entity of associated proteins in G1-, S- and G2-phase. As no suitable antibody precipitating Ago2 exists, I generated a stable cell line expressing low levels of flagHA-tagged Ago2 under the control of an endogenous promoter. I optimized immunoprecipitating conditions and cell cycle phase separation via counterflow centrifugal elutriation to perform an experiment combining elutriation, IP and SILAC labeling.

5.2 SILAC LABELING AND MS

Incorporation of heavy isotopes was sufficient and comparable to data in the literature. (Bonaldi et al., 2008). A histogram representing overall H/L ratios is shown in Figure 9b. The short list of proteins with adequate intensity as well as the low peptide coverage of the Ago proteins of only 10-20 unique peptides indicated a relatively low detection level in the MS data acquired. Furthermore proteins known to be associated with Ago could not be detected in the MS data, as for example *GW182, dFXR or Pur-alpha*. Nonetheless I decided to choose candidate proteins from the lists, which showed reproducible association in our data and a varying H/L ratio. In further IPs and immunofluorescence microscopy I tried to recapitulate a potential physical interaction.

5.3 CANDIDATE PROTEINS NCD, POLO AND RM62 NOT CONFIRMED

Candidate proteins chosen were two proteins with a role in cell cycle regulation, ncd and Polo and a protein which was implicated in RNAi function before, *Rm62*. The name of the protein *non claret disjunctional* (ncd) derives from the locus claret (ca) in *Drosophila simulans*. Mutation experiments by Chandra et al in this locus lead to non-disjunction and loss of chromosomes in meiosis, a corresponding mutant in *D. melanogaster* shows phenotypes involving the eye color, abnormal chromosome segregation in meiosis and frequent loss of maternal chromosomes in early embryonic cleavage divisions (Chandra et al., 1993). Ncd is a kinesin-related motor protein of the kinesin 14 family moving towards the minus end of microtubules (Sharp et al., 2000).

Polo is the first member discovered from the *polo like kinase* family (plk) comprised of serine/ threonine -kinases active during mitosis. Its predominant target is *string* an ortholog of the yeast *cdc25*, the phosphatase dephosphorylating cdc2/Cdk1 which in turn promotes entry into mitosis. (Llamazares et al.,

1991). Polo is required for spindle pole organization, sister chromatid separation and cytokinesis, *polo* mutants show an accumulation of cells with condensed chromosomes not aligned at the metaphase plate but no mitotic arrest (Donaldson et al., 2001).

Both proteins are clearly involved in cell cycle regulation and have never been implicated in RNAi function. Due to their essential role in cell vitality depletion studies are difficult to handle. As the SILAC experiment outcome was not reliable, because of the missing positive controls and further interaction studies did not convince us of physical interaction of ncd or Polo with RISC, we decided to concentrate on other subjects.

RM62 is a nuclear member of the large family of Asp-Glu-Ala-Asp (DEAD) box family of ATPases and helicases (Huang and Liu, 2002). Other names of RM62 are *p68*, or *lighten up* (lip); RM62 can unwind dsRNA in both 3′ → 5′ and 5′ → 3′ directions and plays a role in cell development and organ maturation. RM62 has been shown to interact with dFXR, it is necessary for RNAi function (Huang and Liu, 2002) and mutant flies show increased susceptibility to viral infection (Zambon et al., 2006). An interaction of RM62 with Ago2 and its role in efficient RNAi was also observed before (Ishizuka et al., 2002), but interaction with Ago1 in a cell cycle dependent manner has never been described so far. Rm62 also physically interacts with another protein detected in 4 out of 6 of our IPs, Su(var)3-9, a H3K9 specific methyltransferase (see appendix 6.2). Both proteins can mediate Hsp70 transcriptional shut down; Rm62 thereby presumably plays a functional role in recruiting a methyltransferase (Boeke et al., 2011). Altogether, as an oscillation of the interaction between RM62 and Ago1 over the cell cycle could not be confirmed we did not pursue the interaction any further.

Figure 10a shows a Western Blot of the candidate proteins after IP of Ago1, flagAgo2 and myc as a negative control. Thereby myc seemed to precipitate relatively high amounts of Ago1, ncd and Rm62 compared to Ago1 and flag Ago2 IP and 5 % input. As 2 mg of protein extract was used for these control IPs (500 µg for the other IPs shown in Figure 10b and c) these interactions are presumably unspecific. For the SILAC experiment, the control-IP resulted in a list of proteins of similar length as the Ago-IPs but with considerably lower peptide coverage, therefore this presumably is not the reason for the poor SILAC outcome.

In 1998 Rm62 was initially observed to interact with the U1 : 5′-splice site duplex during the splicing process (Liu et al., 1998) and its essential role for pre-mRNA splicing in vitro was demonstrated shortly after. Depletion of the RM62 helicase did not affect assembly of U1 and U2 (pre-spliceosome) but arrested transition from pre-spliceosome to spliceosome (Liu, 2002). In the light of our recent findings suggesting a correlation between impaired splicing and RNAi this confirmed interaction - albeit independent of cell cycle progression - might be of potential relevance in the mechanism of RNAi formation.

Blanks is another protein which was associated with Ago1 and Ago2 in good reliability in our data, but with a coonstant H/L ratio in each IP, therefore we did not choose it as a candidate protein in the first place. However, Blanks was recently implicated in RNAi and interacting with RM62 (Gerbasi et al., 2011). Blanks is a siRNA and dsRNA binding protein expressed in S2 cells and in Drosophila testes that forms a complex

distinct from the canonical RISC and is required for spermiogenesis. Due to its dsRNA binding properties and its nuclear localization it might be a candidate for endo-siRNA precursor formation or selection in the nucleus.

5.4 TRANSGENE INTEGRATION AS MODEL FOR TRANSPOSONS INSERTION

As a model for *de novo* genetic invasion of foreign DNA elements in *Drosophila* we used a GFP reporter construct, which we stably transfected in high copy numbers. Using this construct brings along two fundamental advantages; first, expression repression can be analyzed instantly by flow cytometry and secondly, as no related gene exists in Drosophila, GFP can mimic the horizontal transfer of a transposable element. Transposable elements integrate in the genome randomly and there is no scientific evidence for site specific integration of transposable elements in Drosophila, apart from a general preference of e.g. the P-element transposon for promoters and the first intron of a gene (Spradling et al., 1995). Transposable elements, although possessing certain characteristic features as LTRs or inverted repeats, are not predominantly targeted by siRNAs aiming at these features (see Figure 21). Our reporter construct, like many retrotransposons, possesses two open reading frames and the plasmid altogether exhibits a size of 10 kb. These characteristics allocate our GFP construct as a good model for transposable elements enabling us to look at a first stage of defense against foreign genetic elements in cell culture.

Endo-siRNA mediated repression of transgenes was clearly demonstrated before and it was assumed that a high copy number is necessary to initiate this response. The influence of the chromatic localization of the transgene, has not been examined before. In Southern Blot and Fluorescence in situ hybridization assays I observed a diverse integration mode of the transgene among the reporter clones, which did not influence the level of targeting endo-siRNA. In my thesis I demonstrated a copy number dependent increase in the amplitude of repression, which is independent from the chromatic localization of the transgene and I provide additional evidence for the reproducibility of the system.

5.5 ENDO-SIRNA REPRESSION OF TRANSGENES EXPANDS WITH COPY NUMBER

Some of the transposable elements in Drosophila's genome are present in very high copy numbers, implicating loose control of transposition. A potential explanation could be that repression of transposition only occurs if a certain threshold number of copies is reached. Chaboissier et al observed a threshold for suppression of the LINE like I-element at about 15 copies, without the knowledge of endo-siRNA mediated silencing (Chaboissier et al., 1998). Pal-Bhrada et al observed that as few as 2 to 6 copies of a transgene are enough to exhibit an inhibitory effect on expression levels (Pal-Bhadra et al., 1997). In our data we observed significant transgene silencing of the reporter construct in 12 out of 14 clones with

more than 10 copies. We therefore postulate the existence of a threshold level for transgene silencing for constructs with comparable structure. Deep sequencing of RNA isolated from these clones revealed that the production of siRNAs directed against the inserted sequence increase linearly with increasing copy number. Endo-siRNAs are produced all along the inserted region, with little accumulation in the reporter coding region compared to the plasmid backbone. All siRNAs perfectly complementary to the mature target mRNA potentially confer silencing and one functional RISC complex is sufficient for mRNA degradation. Protein occupancy and secondary structure of the target mRNA tend to result in variable silencing efficiency depending on the target site. Therefore a linear increase in endo-siRNAs does not necessarily entail reproducible linear repression levels. The number of RISC components presumably is not rate limiting in this system as siRNAs sequenced in our experiments exhibit equal distribution of sense and antisense orientation and show a bias for a less stable 5' end, hence we are looking at Ago2 loaded and therefore functional siRNAs.

Despite the fact that histone reporter clones accumulated siRNAs targeting a certain region within the plasmid sequence, which will be discussed later, the number of siRNAs targeting the backbone increased with the copy number. We hypothesized that with increasing number of genomic insertions the level of antisense transcripts, a prerequisite for forming double stranded precursors, increases in the nucleus independent from the localization of integration. These accumulating dsRNAs serve as substrate for the endo-siRNA biogenesis machinery and "feed" protective RISC complexes, targeting the foreign genetic material. Antisense transcription does not depend on an active promoter but is rather stimulated at any nucleosome free region or open chromatin structures (Finocchiaro et al., 2007). Antisense transcripts are rapidly degraded by the nuclear exosome and its exonucleolytic cofactor rrp6, but the endo-siRNA machinery competes for the same substrates, the cis-NATs (Okamura et al., 2008a). Due to the high local concentration of sense and antisense transcripts of high copy genes double stranded precursors have a high propensity to rapidly form, which in turn leads to stabilization of the duplex favoring the endo-siRNA pathway. In our experiments we could observe a copy number dependent increase of endo-siRNA levels targeting the transgene and confirmed these long standing hypotheses.

5.6 HISTONE GENES MIGHT BE IDENTIFIED BY MEANS OF THEIR 3'UTR

As transposons are predominantly inherited vertically and coevolved with their hosts for a long time (Agren and Wright, 2011) their detection is a difficult task for the genome's immune system and the discrimination between self and non self must be tackled from two sides. Protecting its own genes is as important as detecting foreign genetic material. Transposons stand out due to their high copy number but a response solely based on this attribute is delicate: two other genes in the Drosophila genome exhibit similar high copy numbers, ribosomal RNA genes and histone genes. In this thesis I focused on endo-siRNA based defense mechanisms and the underlying target discrimination on the basis of mRNA processing events.

Histone genes are encoded in (suspiciously) high copy numbers, they are transcribed by RNA Pol II, like most transposons and due to their alternating orientation they have a high risk for convergent transcription. Nevertheless, histone genes are targeted by endo-siRNAs in very low levels (Figure 21). A remarkable characteristic of histone genes is the 3' end processing mechanism and interaction with the stem loop binding protein (SLB).

We sought to examine the effects of the histone stem loop sequence on transgene repression, by exchanging the 3'UTR of our reporter construct pKF63 with the histone H2a or histone H3 3' UTR (pKE14/pKE15). Expression of these reporter constructs was successful and stably transfected cell lines in different copy numbers were generated. Knock down experiments of endo-siRNAs biogenesis factors and Ago2 revealed that the level of repression was considerably less compared to the reporters with canonical poly A signal (Figure 23).

A reporter construct carrying both 3'UTRs sequentially (pKE18/pKE19) showed the same results as the reporters with histone stem loop alone. In a publication from 1989 Liu and colleagues described that adding a poly A signal within 121 nt downstream of a histone H2a gene in mice resulted in polyadenylation of less than 5 % of the transcripts (Liu et al., 1989) This data suggests that the histone processing signal is dominant in competition with a downstream PAS. Histone 3' UTRs encode a cryptic PAS themselves within the HDE, which is used when an intron is added to the respective coding site. Histones are a rare example for genes without intervening sequences (Hentschel and Birnstiel, 1981) and as early as 1990 Pandey et al observed in chimeric human genes that adding an intron into the coding region alters the usage of the 3' processing signal and splicing of a nascent transcripts directs it to utilize the cryptic PAS and the polyadenylation pathway (Pandey et al., 1990). They observed polyadenylation of about 50% of the transcripts, while adding a canonical PAS close to the 3' end of an intron-containing histone gene did not shift 3' end formation further towards polyadenylation. As our reporter construct carries an intron in the Ubiquitin 64E promoter, residing in the 5'UTR of the encoded transcript, we reproduced the same situation and observed polyadenylation of about 30% of our transcripts (pKE14 and pKE18) irrespective of an additional PAS downstream of the histone stem loop. Creating this artificial situation of an intron containing histone reporter we serendipitously discovered that this non-natural combination of motifs, (which both standing alone are not unfamiliar in fly genes), elicits high levels of endo-siRNA targeting specifically within the intron.

In our data replacement of the canonical poly A signal with a histone stem loop led to a reproducible accumulation of endo-siRNAs directed against a stretch starting at the transcription start side and ending shortly upstream of the 3' splice side. As the intron in our reporter construct lies in close proximity to the transcription start site it is difficult to assign the accumulating siRNAs in our deep sequencing libraries to an exact localization. Two scenarios resulting in such an endo-siRNA targeting pattern are conceivable. Accumulation of transcription start site associated antisense transcripts or endo-siRNA accumulation at stalled spliceosomes.

5.7 REPLACEMENT OF THE POLY A SIGNAL MIGHT LEAD TO INCREASED LOCAL ANTISENSE TRANSCRIPTION

The presence of introns within eukaryotic genes is often associated with enhanced expression levels and inclusion of one intron near the 5'end of a gene increases transcription many folds in yeast and flies (Moabbi et al., 2012). A model explaining the phenomenon proposes that the presence of a splice site near the transcription start site recruits the transcription machinery by an interaction of the U1 snRNA and the transcription factor TFIIH to the promoter helping transcription initiation. Re-initiation of transcription is then promoted by the formation of a gene loop with the promoter region, which is in turn stimulated by an interaction of the 5'splices with the promoter and the 3' splice site with the terminator (Moabbi et al., 2012). Thereby active transcription and adoption of this gene loop confirmation reduces aberrant transcription. Tan-Wong et al identified a protein required for gene looping in yeast, ssu72 (Tan-Wong et al., 2012a). Depletion of ssu72 led to increased promiscuous transcription, higher RNA Pol II occupancy upstream of transcription start sites and an overall higher divergent transcription rate, suggesting a role of ssu72 in enforcing promoter directionality by facilitating gene looping. Interestingly, the group observed a phenotype mimicking ssu72 depletion when they replaced the poly A signal with an Rnt1 cleavage signal, which promotes termination and cleavage but not polyadenylation. Such modified plasmids did not form gene loop structures and the level of promoter associated noncoding RNAs was higher. In our experiments we detected sense and antisense RNAs equally distributed with a predominant length of 21 nt deriving from the sequence downstream of the transcription start side (TSS).

Promoter proximal non coding RNAs are frequently observed at the TSS, but these long noncoding RNAs (lncRNAs) deriving from divergent transcription are mostly single stranded, because RNAs deriving from the Watson strand align in a 250 bp window downstream of the TSS, and Crick oriented strands align upstream of the TSS (Layer and Weil, 2009).

If the missing poly A signal led to decreased promoter directionality and increased promiscuous (antisense)-transcription in our situation, the overall number of endo-siRNAs targeting the histone reporter construct should be higher than for poly A signal containing reporter cells. This is not the case and siRNA generation in our data is limited to an 800 nt stretch downstream of the transcription start site. Furthermore this model brings up a practical question for natural histone genes. These genes neither contain introns nor a canonical poly A signal, however as stated before these genes are not heavily targeted by endo-siRNAs although increased antisense transcription due to a defect in gene looping should generate plenty of precursors.

Altogether it would be interesting to see if we are able to mimic the phenotype we observe by depletion of the Drosophila ortholog of ssu72, CG14216. This would give us the opportunity to review the correlation between the histone stem loop 3'UTR and increased localized endo-siRNA production downstream of the transcription start site.

5.8 COMBINATION OF INTRON AND MISSING POLY A SIGNAL ACCELERATES ENDO-SIRNA PRODUCTION AGAINST THE INTRON

Intronic regions show a statistically significant enrichment of hairpins compared to exon regions, and this level is comparable to intergenic regions (Rearick et al., 2011), but this cannot explain the strong accumulation and the difference between reporters with canonical or cryptic poly A signal.

We suspected that the stem loop ending of the histone reporter transcripts must be involved in this phenomenon, most likely through an interference of the histone 3' processing with correct splicing. This hypothesis is mainly encouraged through the fact that histone genes usually do not contain introns and it would be plausible that recruitment of the SLBP, which usually does not come in proximity with the spliceosome during co-translational mRNA processing steps (Townley-Tilson et al., 2006) is involved.

Transgene silencing is also a well established phenomenon in *Neospora crassa* and *Cryptococcus neoformans* (Wang et al., 2012), known as quelling. This pathway targets repetitive transgene arrays in a post-transcriptional and homology dependent manner. Efficiency of the phenomenon was recently observed to mildly correlate with copy number of the transgene. However it was also speculated that qualitatively aberrant features of transgenic DNA or RNA trigger silencing rather than the high copy number (Wang et al., 2012).

Quelling is an RNAi related PTGS process induced by siRNAs and requiring the core RNAi components Ago, Dcr and RdRP. Only recently Dumesic et al described an RdRP containing complex recruited to stalled spliceosomes in *Cryptococcus neoformans*. The so called SCANR complex (spliceosome coupled and nuclear RNAi) is localized in the nucleus and is required for siRNA production directing unspliced mRNA precursors into the siRNA biogenesis pathway (Dumesic et al., 2013). In flies the existence of such a complex is very unlikely, due to the missing RdRP. However, Ago2 could be detected in the nucleus (Figure 12). This discovery substantiates our hypothesis that precursor mRNAs dwelling in the nucleus have a high propensity to form dsRNA with complementary antisense transcripts. These dsRNAs reflect potential substrates for the RNAi biogenesis machinery and impaired splicing could amplify such an effect. Dumesic et al found siRNAs spanning intro-exon boundaries in their deep sequencing reads (Dumesic et al., 2013), which was also the case in our deep sequencing data, a preferred substrate for endo-siRNA generation therefore might be the spliced intron or the nucleic pre-mRNA .

Hereupon two profound questions arise: How can endo-siRNAs that do not match the intended target - the mature and spliced mRNA - be beneficial for the organism and how does endo-siRNA mediated intron targeting make sense if retrotransposons do not contain introns?

A possible explanation might provide the overall correlation of introns with transposons. Based on structural information about involved proteins, group II introns in bacterial and organellar genomes were proposed a common ancestor of spliceosomal introns and retrotransposons in eukaryotes (Lambowitz and Zimmerly, 2011). These rybozymes catalyze their own reverse splicing from a transcript back into the DNA, enabling their proliferation within the genome. This process diverged into two fundamentally

different processes in eukaryotic systems over the time, intron splicing and transposition of retroelements, but could still be a hint for cooperation of these processes.

The evolutionary background of introns in eukaryotic genes has been matter of an unresolved debate since their discovery in the late 70s. One hypothesis indicates that introns evolved from transposable elements. Supporting evidence for this mechanism was shown in Drosophila by Yenerall 2012. As correct splicing implies a splice donor and acceptor site besides a branch point, transposons can be intronized completely once they insert into protosplice site (AGGT) (Yenerall and Zhou, 2012) and contain a pyrimidin rich branch point sequence 5-40 nt from their 3' end. They can thus be spliced out from a coding sequence without disrupting the gene's integrity. Transposons integrating into a sequence resembling only the splice donor site of the protosplice site might result in assembly of the spliceosome, stalling and endo-siRNA generation.

Spliceosome stalling of course should be an important prerequisite for endo-siRNA production and normal splicing activities including alternative splicing may not be affected. Under these circumstances this mechanism could serve as criterion to distinguish between self and non self.

Dumesic et al found that the majority of reads in their sequencing results mapped antisense to transposon and transposon like centromeric sequences with suboptimal introns (Dumesic et al., 2013). When an intron from siRNA targets was deleted or their 5' splice site mutated, the number of siRNAs decreased. In contrast, mutation of the 3' splice site led to higher occupancy of spliceosomes at these targets accompanied with increasing numbers of siRNAs targeting the region. As a compelling explanation for this observation they discussed that foreign genetic elements horizontally invading a cell are not pre-optimized for the species cellular machineries as for example the splicing apparatus (Dumesic et al., 2013; Kim Guisbert and Sontheimer, 2013) and this unadapted characteristic might alert the genome's immune system, the RNAi machinery.

Splicing factors were found as candidates in a phylogenetic approach interacting with the RNAi pathway by Tabach et al in early 2013. In this study proteins were analyzed for similar phylogenetic profiles as the *C. elegans* Argonaute protein RDE-1, based on similarity of patterns of conservation or divergence across different organisms. Among other proteins a poly A polymerase and orthologes of RNA splicing factors were identified (Tabach et al., 2013). A screen for proteins leading to RNAi defects, in correlation with double strand break induced small RNAs in our lab recently revealed a similar hint to a connection of RNA splicing and RNA interference. Earlier screens did not uncover to the same connection (Kim et al., 2005; Zhou et al., 2008).

5.9 ARE DEBRANCHED LARIATS A SUBSTRATE FOR THE ENDO-SIRNA PATHWAY?

Additional information about the biogenesis of intron derived siRNAs could be obtained analyzing the consequences of a depletion of the lariat debranching enzyme (ldbr). In *C. neoformans,* depletion of the debranching enzyme Dbr1 led to total abrogation of SCANR dependent siRNA production (Dumesic et al., 2013). To confirm intron lariats as a substrate for the siRNA biogenesis pathway and to clarify the order of events we knocked down *ldbr* using two differently designed dsRNAs in clones with strong intron targeting phenotype. We could observe a reduction of intron matching siRNAs in Northern Blot to 70 %. As the experiments in *C. neoformans* were carried out in genetically deficient yeast we did not expect to see similar substantial effect with knock down of the respective protein ldbr in *Drosophila* in cell culture. Closer inquiry of existing proteins in *Drosophila* moreover revealed a second lariat debranching enzyme expressed in S2 cells which might exhibit a redundant function. Altogether the current data is consistent with our hypothesis. Knock down of *ldbr* at the same time lead to a marked increase in detecting the lariat by qPCR, which presumably reflects accumulating branched introns in the nucleus. In our case the endo-siRNA pathway might compete with degradation of debranched introns, as accumulation of the precursor lariat is accompanied by a reduced existence of the endo-siRNA product. Branched lariats cannot be exported from the nucleus (Cheng and Menees, 2011) and it is not clear whether they are not further processed into endo-siRNAs because they are un-accessible due to their nuclear localization or their branched structure. Both is conceivable and should be further explored in the future, as Ago2 can partially localize to the nucleus and the function of other protein factors implicated in the RNAi pathway as RM62 (Figure 11 and 12) or blanks (appendix 6.2), both also localized in the nucleus, are not clear so far.

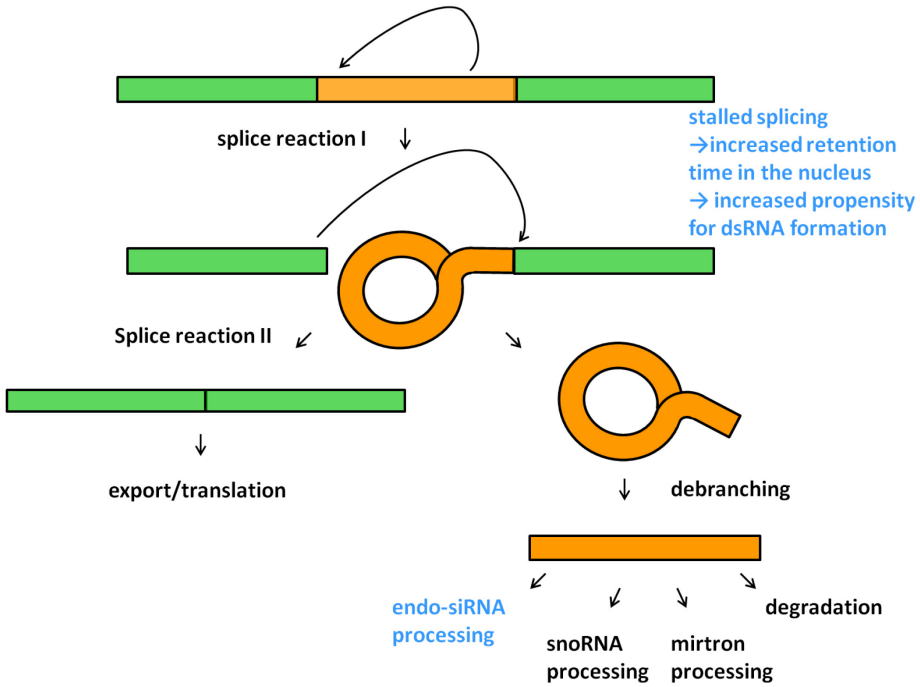


Figure 36 endo-siRNA production from debrached lariats. Increased retention time in the nucleus due to a stalled splicing reaction leads to increased propensity to encounter complementary antisense transcripts and dsRNA formation. Upon debranching of the partially double stranded lariat the linear precursor can be channeled into the endo-siRNA biogenesis pathway.

5.10 OUTLOOK

To model a potential protective effect of the histone 3' UTR on foreign genetic elements the Ubiquitin promoter should be exchanged with an intron-less promoter. The impact of the 3'UTR on endo-siRNA production should then be re-examined in stably transfected clones and deep sequencing experiments. Furthermore the distance of an intron to the 3' end could influence the production of intron derived siRNAs.

To further examine the phenomenon leading to an accumulation of intron derived siRNAs one could inhibit splicing in clones carrying pKF63 via RNAi of carefully selected splice factors and observe a potential increase in intron matching deep sequencing reads mapping to the intron.

At the same time it would be interesting if the phenotype could be induced by depletion of the Drosophila ortholog of ssu72, CG14216 (Tan-Wong et al., 2012b). Depletion of this protein led to defects in forming a gene loop structure, which helped to maintain promoter directionality. The same phenotype was observed to depend on a functional poly A tail which we might have mimicked with our histone reporter constructs.

Depletion of both Drosophila lariat debranching enzymes ldbr and CG7741 might reveal a coupling of debranching and channeling of intronic RNA into the RNAi pathway (Dumesic et al., 2013).

As RM62 and blanks, two proteins detected in the SILAC experiment, are located in the nucleus and their function is not fully explored yet, their involvement in precursor formation or selection in our specific situation should be included in further RNAi experiments.

6 APPENDIX

6.1 ABBREVIATIONS

°C	degrees Celsius
63N1	endo-siRNA cell culture reporter cell line
67-1D	siRNA cell culture reporter cell line
Aa	amino acid
Ago	Argonaute
Amp	Ampicillin
APS	ammonium peroxodisulfate
ATP	adenosine triphosphate
BLAST	Basic Local Alignment Search Tool
bp	base pair(s)
BSA	bovine serum albumine
C	cytosine
cDNA	complementary DNA
CG4068	hairpin forming endo-siRNA precursor gene
C. neoformans	*Cryptococcus neoformans*
co-IP	co-immunoprecipitation
C-term	protein C-terminus
CT-value	cycle of threshold value in qPCR
D. melanogaster	*Drosophila melanogaster*
Da	Dalton
Dcr	Dicer
dFXR	Fragile X Mental Retardation Protein
DMSO	dimethyl sulfoxide
DNA	desoxy-ribonucleic acid
dNTP	desoxy-nucleotide-tri-phosphate
ds	double-stranded
dsRBP	double-stranded RNA binding domain protein
Dsred	Discosoma striata red fluorescence protein
dsRNA	double-stranded RNA
DTT	dithiothreitol
E. coli	*Escherichia coli*
ECL	enhanced chemiluminescence
EGFP	Enhanced Green Fluorescent protein
endo-	endogenous
exo-	exogenous
FACS	Fluorescence Activated Cell Sorting
FBS	Fetal Bovine Serum
fw	forward

G	guanine
GFP	Green Fluorescent Protein
GO	gene ontology
h	hour(s)
HDE	histone downstream element
HEPES	(4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
HP-1	Heterochomatin Protein 1
HRP	Horseradish Peroxidase
Hygro	Hygromycin
IgG	immunoglobulin protein G
IP	immunoprecipitation
IPTG	Isopropyl-β-D-thiogalactopyranosid
k	kilo
k.d.	knock down
k.o.	knock-out
Kan	kanamycin
lncRNA	long noncoding RNA
Loqs	loquacious
Luc	luciferase
mg	milligram
min	minute
miR	micro RNA
miRNA	micro RNA
ml	milliliter
mRNA	messenger RNA
mRNP	messenger ribonucleoprotein
nc	non coding
N. crassa	*Neurosopra crassa*
Neo	Neomycin
ng	nanogram
nt	nucleotide(s)
N-term	protein N-terminus
NTP	nucleotide-tri-phosphate
ORF	open reading frame
p.a.	pro analysi
PA/PB/PC/PD	protein isoform A/B/C/D
PAGE	Polyacrylamide Gel Electrophoresis
PAS	polyadenylation signal
PAZ	Piwi-Argonaute-Zwille domain of Dicer and Argonaute proteins
PCR	Polymerase Chain Reaction
piRNA	Piwi-interacting RNA
PNK	polynucleotide kinase

Pol II	DNA polymerase II
Poly-A	poly-adenylation
PVDF	Polyvinylidenfluoride
qPCR	quantitative Polymerase Chain Reaction
R	Arginine
R2D2	2 dsRBD-containing protein interacting with Dcr-2
rb	rabbit
RdRP	RNA-dependent RNA-Polymerase
rel.	relative
rev	reverse
RISC	RNA induced silencing complex
RLC	RISC loading complex
RNA	ribonucleic acid
RNAi	RNA interference
RNaseIII	endoribonuclease class III
RNP	ribonucleoprotein particle
rpm	rounds per minute
rRNA	ribosomal RNA
RT	reverse transcription
rt	room temperature
S. pombe	*Schizosaccharomyces pombe*
S2 cell	Schneider-2 cell
SD	standard deviation
SDS	sodium docecyl sulfate
siRNA	small interfering RNA
SOB	Super Optimal Broth
ss	single stranded
SSC	sodium chloride/sodium citrate
SV40	Simian Virus 40
T	thymine
TAR	transactivation response RNA
tech.	technical
TRBP	TAR RNA binding protein
TRIS	Tris(hydroxymethyl)-aminomethane
tub.	tubulin
U	uracil
UTR	untranslated region
V	Volt
α	anti
Δ	deletion
µ	micro
µg	microgram

6.2 LIST OF PROTEINS OBTAINED FOR THE SILAC EXPERIMENT

Proteins are listed ranked by overall detection intensity, for every IP the number of uniue peptides detected is specified. Proteins detected in less than 4 out of the 6 experiments and proteins detected in the negative control (myc) in comparable levels are omitted in the list.

Proteins associated with Ago1 and Ago2 but also found in negative control

			Ago1			Ago2			
Protein ID		**Name**	**G1/G2**	**S/G1**	**G2/S**	**G1/G2**	**S/G1**	**G2/S**	**myc**
FBpp0085235	CG1483-PB	Map205	14	22	18	20	26	20	6
FBpp0083802	CG6779-PA	RpS3	1	2	5	1		1	1
FBpp0088242	CG2168-PA	RpS3A	1	3	3	2		2	2
FBpp0088439	CG1883-PA	RpS7	4	1	2		1		2
FBpp0070279		RpSA	1	1	1		1		1
FBpp0085717	CG7726-PA	RpL11	1	2	1		1	1	1
FBpp0087142	CG8280-PA	Ef1alpha48D	16	11	15	15	16	17	10
FBpp0111817	CG10811PB	eIF4G	6	8	4	5	9	6	1
FBpp0084623	CG5520-PA	Gp93		1	1		2	1	1
FBpp0082421	CG3379-PA	His4r	3	4	3	4	4	1	1
FBpp0081062	CG2512-PA	alphaTub84D	9	6	8	2	6	5	2

Proteins associated with Ago1 and Ago2, not found in negative control

			Ago1			Ago2			
Protein ID		**Name**	**G1/G2**	**S/G1**	**G2/S**	**G1/G2**	**S/G1**	**G2/S**	**myc**
FBpp0077720	CG4164-PA	CG4164	4	1	4	1	2	2	
FBpp0086769	CG6543-PA	CG6543	4	6	4	2	4	4	
FBpp0081618	CG8507-PA	CG8507	7	8	8	8	11	9	
FBpp0084761	CG11901PB	Ef1gamma	12	6	15	9	9	10	
FBpp0085265	CG2238-PA	Ef2b	27	29	29	18	20	20	
FBpp0087142	CG8280-PA	Ef1alpha48D	16	11	15	15	16	17	10
FBpp0111817	CG10811PB	eIF4G	6	8	4	5	9	6	1
FBpp0073445	CG4147-PB	Hsc70-3	19	18	20	15	18	16	
FBpp0082514	CG4264-PE	Hsc70-4	17	16	19	12	11	14	3
FBpp0071213	CG10701PB	Moe	9	6	7	3	3	3	
FBpp0073538	CG3989-PA	ade5	6	8	10	4	3	5	
FBpp0086190		eEF1beta	10	9	10	8	10	10	
FBpp0079542	CG4912-PB	eEF1delta	5	4	8	4	3	4	
FBpp0084900	CG7831-PA	ncd	9	10	5	11	10	10	
FBpp0083611	CG7070-PA	PyK	7	15	8	6	7	3	
FBpp0074608	CG12306PA	polo	1	3	1	1	10	3	
FBpp0085721	CG9277-PA	betaTub56D	8	6	11	2	9	9	

Proteins associated with Ago1 and Ago2 in 5 out of 6 IPs

			Ago1			Ago2			
Protein ID		**Name**	**G1/G2**	**S/G1**	**G2/S**	**G1/G2**	**S/G1**	**G2/S**	**myc**
FBpp0085585	CG8900-PB	RpS18	*6*	*2*	*5*	*2*		*2*	
FBpp0099686	CG7808-PC	RpS8	*3*	*3*	*3*	*1*		*2*	
FBpp0074075	CG9946-PA	eIF-2alpha	*3*	*2*		*1*	*2*	*2*	
FBpp0291000	CG11963PC	skap	*1*	*1*	*1*		*2*	*1*	
FBpp0086971	CG8759-PB	NACalpha		*2*	*2*	*1*	*1*	*1*	
FBpp0076856	CG10630PA	blanks	*1*	*4*	*3*	*4*	*5*		
FBpp0086747	CG18076PH	shot		*1*	*1*	*1*	*1*	*1*	
FBpp0081401	CG8351-PA	Tcp-1eta	*2*	*3*	*5*		*1*	*2*	
FBpp0073902	CG8231-PA	Tcp-1zeta	*3*	*4*	*8*		*2*	*4*	

Proteins associated with Ago1 and Ago2 in 4 out of 6 IPs

			Ago1			Ago2			
Protein ID		**Name**	**G1/G2**	**S/G1**	**G2/S**	**G1/G2**	**S/G1**	**G2/S**	
FBpp0081845	CG6684-PA	RpS25	*2*	*1*	*1*	*1*			
FBpp0075618	CG11276PA	RpS4		*2*	*5*	*2*		*1*	
FBpp0086897	CG3821-PA	Aats-asp	*5*	*5*	*9*			*1*	
FBpp0073513	CG2028-PB	CkIalpha	*1*	*1*	*1*		*1*		
FBpp0078450	CG11999PA	CG11999	*2*	*3*	*3*			*1*	
FBpp0080180	CG18095PA	CG18095		*1*		*1*	*1*	*1*	
FBpp0071226	CG7033-PA	CG7033	*1*	*3*	*6*			*2*	
FBpp0084968	CG7920-PA	CG7920	*7*	*7*	*4*	*2*			
FBpp0085281	CG17949PA	His2B	*3*	*4*	*3*	*3*			
FBpp0082584	CG43664	Su(var)3-9	*3*	*1*			*2*	*2*	
FBpp0074550	CG7176	IDH1	*1*	*1*	*2*		*2*		
FBpp0080146	CG8978-PB	Arpc1	*1*	*2*	*2*		*2*		
FBpp0082787	CG8977-PA	Cctgamma	*1*	*1*	*3*			*1*	

Proteins associated with Ago1

			Ago1		
Protein ID		**Name**	**G1/G2**	**S/G1**	**G2/S**
FBpp0079500	CG5920-PA	RpS2	*3*	*3*	*3*
FBpp0078134	CG7490-PA	RpLP0	*1*	*2*	*4*
FBpp0077716	CG4087-PA	RpLP1	*1*	*1*	*1*
FBpp0079484	CG4651-PA	RpL13	*2*	*1*	*3*
FBpp0076359	CG6253-PA	RpL14	*3*	*2*	*3*
FBpp0076602	CG8615-PA	RpL18	*1*	*1*	*1*
FBpp0072312	CG2746-PA	RpL19	*1*	*1*	*1*
FBpp0070143	CG7434-PA	RpL22	*2*	*1*	*1*
FBpp0072958		RpL28	*2*	*1*	*3*
FBpp0081822	CG4863-PA	RpL3	*1*	*2*	*2*
FBpp0084617	CG5502-PA	RpL4	*2*	*3*	*2*
FBpp0086738	CG6671-PC	AGO1	*21*	*20*	*18*
FBpp0086895	CG3644-PA	bic	*4*	*1*	*2*
FBpp0074345	CG32549	CG32549	*2*	*2*	*5*
FBpp0079454	CG4389-PA	Mtpalpha	*3*	*3*	*1*
FBpp0075754	CG5642-PA	CG5642	*1*	*1*	*1*
FBpp0073872	CG9281-PB	CG9281	*7*	*8*	*7*
FBpp0085915	CG5119-PC	pAbp	*1*	*4*	*3*
FBpp0087869	CG8705-PB	pnut	*1*	*1*	*2*
FBpp0082123	CG6148-PA	Past1	*1*	*9*	*6*
FBpp0111773	CG11661-PI	Nc73EF	*2*	*2*	*1*
FBpp0075764	CG7283-PA	RpL10Ab	*1*	*3*	*2*
FBpp0078301	CG10279-PA	Rm62	*4*	*8*	*6*
FBpp0083683	CG5374-PB	T-cp1	*1*	*2*	*3*
FBpp0079992	CG5525-PA	T-cp1	*1*	*2*	*2*

Proteins associated with Ago2

			Ago2		
Protein ID		**Name**	**G1/G2**	**S/G1**	**G2/S**
FBpp0075313	CG7439-PC	AGO2	*12*	*13*	*13*

6.3 ACKNOWLEDGEMENTS

An erster Stelle möchte ich mich ganz herzlich bei meinem Doktorvater Klaus Förstemann bedanken, der mir zwei sehr spannende und abwechslungsreiche Projekte anvertraut hat und mich dabei mit seiner freundschaftlichen und optimistischen Haltung auch in schwierigen Zeiten motivieren konnte.

Vielen Dank auch an mein Thesis Advisory Committee, Petra Wendler und Laurent Lariviere, danke für kritische und motivierende Denkanstöße und für die Zeit die Ihr dafür geopfert habt.

Katja Strässer danke ich sehr herzlich für die Übernahme des Zweitgutachtens.

Ganz besonderer Dank gilt auch Stefan Krebs und dem LAFUGA für die Übernahme zahlreicher SOLEXA Sequenizierung und die Unterstützung und Beratung am Miseq, an dieser Stelle auch herzlichen Dank an Kerstin Meier.

Ich möchte mich auf bei Katrin und Daniel bedanken, die in ihrem Praktikum sehr nützliche Experimente für meine Arbeit durchgeführt haben.

Danke auch Dir im Speziellen, Romy, für die Plasmide pRB1 und pRB2 (Bötch 1 und 2) für die Klone B1 und B4 und dafür, dass Du das eine oder andere Mal für mich in die Bresche gesprungen bist.

Auch allen anderen Förstegirls und Kollegen und Ex-Kollegen am Genzentrum möchte ich für die tolle Zeit danken, für Hilfsbereitschaft und Kollegialität, aber vor allem auch für die Parties, die speziellen Feiertage, Kaffeerunden und „Betriebsausflüge“. Ausserdem möchte ich mich bei meiner Doktorarbeit bedanken dass ich durch sie eine tolle Freundin wie die Steffi gefunden habe.

Auch meinem privaten Umfeld bin ich sehr dankbar, Mama und Papa, danke dass Ihr seit ich denken kann immer an mich geglaubt habt, mich dann im Studium seelisch, moralisch und auch finanziell großzügig unterstützt habt und euch immer noch interessiert meine Geschichten über die Fruchtfliege anhört. Ihr habt sehr dazu beigetragen, dass ich immer wieder nach vorne schauen konnte.

Michi, danke, dass Du mich immer unterstützt und motiviert hast, für das was du bist und dass ich Dich an meiner Seite habe.

6.4 CURRICULUM VITAE

ANNA KATHARINA ELMER

Email: katharina.elmer@gmail.com

ACADEMIC TRAINING

Since 2009	PhD in Biochemistry Ludwig-Maximilians-University Munich, Gene Center, Laboratory of Prof. Dr. Klaus Förstemann *Thesis title: "Modeling transposon recognition in Drosophila melanogaster"*
06/2008 – 12/2008	Master Thesis in Immunology Helmholtz Zentrum München, Institute of Molecular Immunology, Laboratory of PD Dr. Elfriede Nößner *Thesis title: "Characterization of a T cell population primed with heat shock protein-70"*
08/2007 – 12/2007	Internship in Cell Biology Cajal Institute Madrid, Department of Functional and Systems Neurobiology, Laboratory of Prof. Dr. Torres Alemán
2006-2008	Master in Biochemistry Ludwig-Maximilians-University Munich Major: Biochemistry and Cell biology - Minor: Immunology Final grade: 1.3 (A)
2003 -2006	Bachelor in Molecular Science Friedrich Alexander University Erlangen/Nuremberg Major: Molecular Life Science Final grade: 2.3 (B)

PUBLICATIONS

Peer reviewed journal	Jolesch A.*, **Elmer K.***, Bendz H., Issels R. D., Noessner E., **2012**. Hsp70, a messenger from hyperthermia for the immune system. Eur. J. Cell Biol. *91, 48 – 52* **contributed equally*
Poster presentation	Cold Spring Harbor Laboratory meeting: Regulatory and Non-coding RNA. 2012, Cold Spring Harbor, USA Anna Katharina Elmer & Klaus Förstemann, *Endo-siRNA silencing of artificial transposons in Drosophila somatic cells depends on copy number and the presence of a poly A signal.*

SCHOLARSHIPS

2012	Full traveling award of DAAD for the Cold Spring Harbor Laboratory meeting: Regulatory and Non-coding RNA, 2012

7. REFERENCES

Agren, J.A., and Wright, S.I. (2011). Co-evolution between transposable elements and their hosts: a major factor in genome size evolution? Chromosome Res 19, 777-786.

Amaral, P.P., and Mattick, J.S. (2008). Noncoding RNA in development. Mamm Genome 19, 454-492.

Ameres, S.L., Hung, J.H., Xu, J., Weng, Z., and Zamore, P.D. (2011). Target RNA-directed tailing and trimming purifies the sorting of endo-siRNAs between the two Drosophila Argonaute proteins. RNA 17, 54-63.

Aumiller, V., Graebsch, A., Kremmer, E., Niessing, D., and Forstemann, K. (2012). Drosophila Pur-alpha binds to trinucleotide-repeat containing cellular RNAs and translocates to the early oocyte. RNA Biol 9, 633-643.

Bail, S., Swerdel, M., Liu, H., Jiao, X., Goff, L.A., Hart, R.P., and Kiledjian, M. (2010). Differential regulation of microRNA stability. RNA 16, 1032-1039.

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

Bartel, D.P. (2009). MicroRNAs: target recognition and regulatory functions. Cell 136, 215-233.

Bartolome, C., Bello, X., and Maside, X. (2009). Widespread evidence for horizontal transfer of transposable elements across Drosophila genomes. Genome Biol 10, R22.

Behm-Ansmant, I., Rehwinkel, J., Doerks, T., Stark, A., Bork, P., and Izaurralde, E. (2006). mRNA degradation by miRNAs and GW182 requires both CCR4:NOT deadenylase and DCP1:DCP2 decapping complexes. Genes Dev 20, 1885-1898.

Boeke, J., Bag, I., Ramaiah, M.J., Vetter, I., Kremmer, E., Pal-Bhadra, M., Bhadra, U., and Imhof, A. (2011). The RNA helicase Rm62 cooperates with SU(VAR)3-9 to re-silence active transcription in Drosophila melanogaster. PLoS One 6, e20761.

Bonaldi, T., Straub, T., Cox, J., Kumar, C., Becker, P.B., and Mann, M. (2008). Combined use of RNAi and quantitative proteomics to study gene function in Drosophila. Mol Cell 31, 762-772.

Brand, A.H., and Perrimon, N. (1993). Targeted gene expression as a means of altering cell fates and generating dominant phenotypes. Development 118, 401-415.

Brennecke, J., Hipfner, D.R., Stark, A., Russell, R.B., and Cohen, S.M. (2003). bantam encodes a developmentally regulated microRNA that controls cell proliferation and regulates the proapoptotic gene hid in Drosophila. Cell 113, 25-36.

Brennecke, J., Malone, C.D., Aravin, A.A., Sachidanandam, R., Stark, A., and Hannon, G.J. (2008). An epigenetic role for maternally inherited piRNAs in transposon silencing. Science 322, 1387-1392.

Caudy, A.A., Ketting, R.F., Hammond, S.M., Denli, A.M., Bathoorn, A.M., Tops, B.B., Silva, J.M., Myers, M.M., Hannon, G.J., and Plasterk, R.H. (2003). A micrococcal nuclease homologue in RNAi effector complexes. Nature 425, 411-414.

Caudy, A.A., Myers, M., Hannon, G.J., and Hammond, S.M. (2002). Fragile X-related protein and VIG associate with the RNA interference machinery. Genes Dev 16, 2491-2496.

Chaboissier, M.C., Bucheton, A., and Finnegan, D.J. (1998). Copy number control of a transposable element, the I factor, a LINE-like element in Drosophila. Proc Natl Acad Sci U S A 95, 11781-11785.

Chandra, R., Salmon, E.D., Erickson, H.P., Lockhart, A., and Endow, S.A. (1993). Structural and functional domains of the Drosophila ncd microtubule motor protein. J Biol Chem 268, 9005-9013.

Cheng, Z., and Menees, T.M. (2011). RNA splicing and debranching viewed through analysis of RNA lariats. Mol Genet Genomics 286, 395-410.

Chung, W.J., Okamura, K., Martin, R., and Lai, E.C. (2008). Endogenous RNA interference provides a somatic defense against Drosophila transposons. Curr Biol 18, 795-802.

Coleman, M.L., Marshall, C.J., and Olson, M.F. (2004). RAS and RHO GTPases in G1-phase cell-cycle regulation. Nat Rev Mol Cell Biol 5, 355-366.

Cox, D.N., Chao, A., Baker, J., Chang, L., Qiao, D., and Lin, H. (1998). A novel class of evolutionarily conserved genes defined by piwi are essential for stem cell self-renewal. Genes Dev 12, 3715-3727.

Donaldson, M.M., Tavares, A.A., Ohkura, H., Deak, P., and Glover, D.M. (2001). Metaphase arrest with centromere separation in polo mutants of Drosophila. J Cell Biol 153, 663-676.

Dumesic, P.A., Natarajan, P., Chen, C., Drinnenberg, I.A., Schiller, B.J., Thompson, J., Moresco, J.J., Yates, J.R., 3rd, Bartel, D.P., and Madhani, H.D. (2013). Stalled Spliceosomes Are a Signal for RNAi-Mediated Genome Defense. Cell 152, 957-968.

Eystathioy, T., Chan, E.K., Tenenbaum, S.A., Keene, J.D., Griffith, K., and Fritzler, M.J. (2002). A phosphorylated cytoplasmic autoantigen, GW182, associates with a unique population of human mRNAs within novel cytoplasmic speckles. Mol Biol Cell 13, 1338-1351.

Fagegaltier, D., Bouge, A.L., Berry, B., Poisot, E., Sismeiro, O., Coppee, J.Y., Theodore, L., Voinnet, O., and Antoniewski, C. (2009). The endogenous siRNA pathway is involved in heterochromatin formation in Drosophila. Proc Natl Acad Sci U S A 106, 21258-21263.

Finocchiaro, G., Carro, M.S., Francois, S., Parise, P., DiNinni, V., and Muller, H. (2007). Localizing hotspots of antisense transcription. Nucleic Acids Res 35, 1488-1500.

Fire, A., Xu, S., Montgomery, M.K., Kostas, S.A., Driver, S.E., and Mello, C.C. (1998). Potent and specific genetic interference by double-stranded RNA in Caenorhabditis elegans. Nature 391, 806-811.

Forstemann, K., Horwich, M.D., Wee, L., Tomari, Y., and Zamore, P.D. (2007). Drosophila microRNAs are sorted into functionally distinct argonaute complexes after production by dicer-1. Cell 130, 287-297.

Francia, S., Michelini, F., Saxena, A., Tang, D., de Hoon, M., Anelli, V., Mione, M., Carninci, P., and d'Adda di Fagagna, F. (2012). Site-specific DICER and DROSHA RNA products control the DNA-damage response. Nature 488, 231-235.

Galardi, S., Mercatelli, N., Farace, M.G., and Ciafre, S.A. (2011). NF-kB and c-Jun induce the expression of the oncogenic miR-221 and miR-222 in prostate carcinoma and glioblastoma cells. Nucleic Acids Res 39, 3892-3902.

Gerbasi, V.R., Preall, J.B., Golden, D.E., Powell, D.W., Cummins, T.D., and Sontheimer, E.J. (2011). Blanks, a nuclear siRNA/dsRNA-binding complex component, is required for Drosophila spermiogenesis. Proc Natl Acad Sci U S A 108, 3204-3209.

Ghildiyal, M., Seitz, H., Horwich, M.D., Li, C., Du, T., Lee, S., Xu, J., Kittler, E.L., Zapp, M.L., Weng, Z., et al. (2008). Endogenous siRNAs derived from transposons and mRNAs in Drosophila somatic cells. Science 320, 1077-1081.

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

Han, T., Manoharan, A.P., Harkins, T.T., Bouffard, P., Fitzpatrick, C., Chu, D.S., Thierry-Mieg, D., Thierry-Mieg, J., and Kim, J.K. (2009). 26G endo-siRNAs regulate spermatogenic and zygotic gene expression in Caenorhabditis elegans. Proc Natl Acad Sci U S A 106, 18674-18679.

Harrison, P.M., Milburn, D., Zhang, Z., Bertone, P., and Gerstein, M. (2003). Identification of pseudogenes in the Drosophila melanogaster genome. Nucleic Acids Res 31, 1033-1037.

Hartig, J.V., Esslinger, S., Bottcher, R., Saito, K., and Forstemann, K. (2009). Endo-siRNAs depend on a new isoform of loquacious and target artificially introduced, high-copy sequences. EMBO J 28, 2932-2944.

Helfer, S., Schott, J., Stoecklin, G., and Forstemann, K. (2012). AU-rich element-mediated mRNA decay can occur independently of the miRNA machinery in mouse embryonic fibroblasts and Drosophila S2-cells. PLoS One 7, e28907.

Hentschel, C.C., and Birnstiel, M.L. (1981). The organization and expression of histone gene families. Cell 25, 301-313.

Huang, Y., and Liu, Z.R. (2002). The ATPase, RNA unwinding, and RNA binding activities of recombinant p68 RNA helicase. J Biol Chem 277, 12810-12815.

Hutvagner, G., and Zamore, P.D. (2002). A microRNA in a multiple-turnover RNAi enzyme complex. Science 297, 2056-2060.

Ishizu, H., Siomi, H., and Siomi, M.C. (2012). Biology of PIWI-interacting RNAs: new insights into biogenesis and function inside and outside of germlines. Genes Dev 26, 2361-2373.

Ishizuka, A., Siomi, M.C., and Siomi, H. (2002). A Drosophila fragile X protein interacts with components of RNAi and ribosomal proteins. Genes Dev 16, 2497-2508.

Jinek, M., and Doudna, J.A. (2009). A three-dimensional view of the molecular machinery of RNA interference. Nature 457, 405-412.

Johnson, S.M., Grosshans, H., Shingara, J., Byrom, M., Jarvis, R., Cheng, A., Labourier, E., Reinert, K.L., Brown, D., and Slack, F.J. (2005). RAS is regulated by the let-7 microRNA family. Cell 120, 635-647.

Kaminker, J.S., Bergman, C.M., Kronmiller, B., Carlson, J., Svirskas, R., Patel, S., Frise, E., Wheeler, D.A., Lewis, S.E., Rubin, G.M., et al. (2002). The transposable elements of the Drosophila melanogaster euchromatin: a genomics perspective. Genome Biol 3, RESEARCH0084.

Kanellopoulou, C., Muljo, S.A., Kung, A.L., Ganesan, S., Drapkin, R., Jenuwein, T., Livingston, D.M., and Rajewsky, K. (2005). Dicer-deficient mouse embryonic stem cells are defective in differentiation and centromeric silencing. Genes Dev 19, 489-501.

Kawamura, Y., Saito, K., Kin, T., Ono, Y., Asai, K., Sunohara, T., Okada, T.N., Siomi, M.C., and Siomi, H. (2008). Drosophila endogenous small RNAs bind to Argonaute 2 in somatic cells. Nature 453, 793-797.

Kim Guisbert, K.S., and Sontheimer, E.J. (2013). Quit Stalling or You'll Be Silenced. Cell 152, 938-939.

Kim, J.K., Gabel, H.W., Kamath, R.S., Tewari, M., Pasquinelli, A., Rual, J.F., Kennedy, S., Dybbs, M., Bertin, N., Kaplan, J.M., et al. (2005). Functional genomic analysis of RNA interference in C. elegans. Science 308, 1164-1167.

Lambowitz, A.M., and Zimmerly, S. (2011). Group II introns: mobile ribozymes that invade DNA. Cold Spring Harb Perspect Biol 3, a003616.

Layer, J.H., and Weil, P.A. (2009). Ubiquitous antisense transcription in eukaryotes: novel regulatory mechanism or byproduct of opportunistic RNA polymerase? F1000 Biol Rep 1, 33.

Le Thomas, A., Rogers, A.K., Webster, A., Marinov, G.K., Liao, S.E., Perkins, E.M., Hur, J.K., Aravin, A.A., and Toth, K.F. (2013). Piwi induces piRNA-guided transcriptional silencing and establishment of a repressive chromatin state. Genes Dev 27, 390-399.

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

Lee, Y., Jeon, K., Lee, J.T., Kim, S., and Kim, V.N. (2002). MicroRNA maturation: stepwise processing and subcellular localization. EMBO J 21, 4663-4670.

Lee, Y., Kim, M., Han, J., Yeom, K.H., Lee, S., Baek, S.H., and Kim, V.N. (2004a). MicroRNA genes are transcribed by RNA polymerase II. EMBO J 23, 4051-4060.

Lee, Y.S., Nakahara, K., Pham, J.W., Kim, K., He, Z., Sontheimer, E.J., and Carthew, R.W. (2004b). Distinct roles for Drosophila Dicer-1 and Dicer-2 in the siRNA/miRNA silencing pathways. Cell 117, 69-81.

Li, Z., Lu, J., Sun, M., Mi, S., Zhang, H., Luo, R.T., Chen, P., Wang, Y., Yan, M., Qian, Z., et al. (2008). Distinct microRNA expression profiles in acute myeloid leukemia with common translocations. Proc Natl Acad Sci U S A 105, 15535-15540.

Lim, D.H., Oh, C.T., Lee, L., Hong, J.S., Noh, S.H., Hwang, S., Kim, S., Han, S.J., and Lee, Y.S. (2011). The endogenous siRNA pathway in Drosophila impacts stress resistance and lifespan by regulating metabolic homeostasis. FEBS Lett 585, 3079-3085.

Liu, Q., Rand, T.A., Kalidas, S., Du, F., Kim, H.E., Smith, D.P., and Wang, X. (2003). R2D2, a bridge between the initiation and effector steps of the Drosophila RNAi pathway. Science 301, 1921-1925.

Liu, T.J., Levine, B.J., Skoultchi, A.I., and Marzluff, W.F. (1989). The efficiency of 3'-end formation contributes to the relative levels of different histone mRNAs. Mol Cell Biol 9, 3499-3508.

Liu, X., Jiang, F., Kalidas, S., Smith, D., and Liu, Q. (2006). Dicer-2 and R2D2 coordinately bind siRNA to promote assembly of the siRISC complexes. RNA 12, 1514-1520.

Liu, Z.R. (2002). p68 RNA helicase is an essential human splicing factor that acts at the U1 snRNA-5' splice site duplex. Mol Cell Biol 22, 5443-5450.

Liu, Z.R., Sargueil, B., and Smith, C.W. (1998). Detection of a novel ATP-dependent cross-linked protein at the 5' splice site-U1 small nuclear RNA duplex by methylene blue-mediated photo-cross-linking. Mol Cell Biol 18, 6910-6920.

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

Llamazares, S., Moreira, A., Tavares, A., Girdham, C., Spruce, B.A., Gonzalez, C., Karess, R.E., Glover, D.M., and Sunkel, C.E. (1991). polo encodes a protein kinase homolog required for mitosis in Drosophila. Genes Dev 5, 2153-2165.

Mann, M. (2006). Functional and quantitative proteomics using SILAC. Nat Rev Mol Cell Biol 7, 952-958.

Martienssen, R.A., Zaratiegui, M., and Goto, D.B. (2005). RNA interference and heterochromatin in the fission yeast Schizosaccharomyces pombe. Trends Genet 21, 450-456.

Michalik, K.M., Bottcher, R., and Forstemann, K. (2012). A small RNA response at DNA ends in Drosophila. Nucleic Acids Res 40, 9596-9603.

Moabbi, A.M., Agarwal, N., El Kaderi, B., and Ansari, A. (2012). Role for gene looping in intron-mediated enhancement of transcription. Proc Natl Acad Sci U S A 109, 8505-8510.

Neil, H., Malabat, C., d'Aubenton-Carafa, Y., Xu, Z., Steinmetz, L.M., and Jacquier, A. (2009). Widespread bidirectional promoters are the major source of cryptic transcripts in yeast. Nature 457, 1038-1042.

Ni, J.Q., Markstein, M., Binari, R., Pfeiffer, B., Liu, L.P., Villalta, C., Booker, M., Perkins, L., and Perrimon, N. (2008). Vector and parameters for targeted transgenic RNA interference in Drosophila melanogaster. Nat Methods 5, 49-51.

Okamura, K., Balla, S., Martin, R., Liu, N., and Lai, E.C. (2008a). Two distinct mechanisms generate endogenous siRNAs from bidirectional transcription in Drosophila melanogaster. Nat Struct Mol Biol 15, 581-590.

Okamura, K., Chung, W.J., Ruby, J.G., Guo, H., Bartel, D.P., and Lai, E.C. (2008b). The Drosophila hairpin RNA pathway generates endogenous short interfering RNAs. Nature 453, 803-806.

Okamura, K., Ishizuka, A., Siomi, H., and Siomi, M.C. (2004). Distinct roles for Argonaute proteins in small RNA-directed RNA cleavage pathways. Genes Dev 18, 1655-1666.

Ono, M., Murakami, T., Kudo, A., Isshiki, M., Sawada, H., and Segawa, A. (2001). Quantitative comparison of anti-fading mounting media for confocal laser scanning microscopy. J Histochem Cytochem 49, 305-312.

Pal-Bhadra, M., Bhadra, U., and Birchler, J.A. (1997). Cosuppression in Drosophila: gene silencing of Alcohol dehydrogenase by white-Adh transgenes is Polycomb dependent. Cell 90, 479-490.

Pandey, N.B., Chodchoy, N., Liu, T.J., and Marzluff, W.F. (1990). Introns in histone genes alter the distribution of 3' ends. Nucleic Acids Res 18, 3161-3170.

Pasquinelli, A.E., Reinhart, B.J., Slack, F., Martindale, M.Q., Kuroda, M.I., Maller, B., Hayward, D.C., Ball, E.E., Degnan, B., Muller, P., et al. (2000). Conservation of the sequence and temporal expression of let-7 heterochronic regulatory RNA. Nature 408, 86-89.

Rearick, D., Prakash, A., McSweeny, A., Shepard, S.S., Fedorova, L., and Fedorov, A. (2011). Critical association of ncRNA with introns. Nucleic Acids Res 39, 2357-2366.

Rehwinkel, J., Behm-Ansmant, I., Gatfield, D., and Izaurralde, E. (2005). A crucial role for GW182 and the DCP1:DCP2 decapping complex in miRNA-mediated gene silencing. RNA 11, 1640-1647.

Saito, K., Ishizuka, A., Siomi, H., and Siomi, M.C. (2005). Processing of pre-microRNAs by the Dicer-1-Loquacious complex in Drosophila cells. PLoS Biol 3, e235.

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

Schultz, J., Lorenz, P., Gross, G., Ibrahim, S., and Kunz, M. (2008). MicroRNA let-7b targets important cell cycle molecules in malignant melanoma cells and interferes with anchorage-independent growth. Cell Res 18, 549-557.

Schwarz, D.S., Hutvagner, G., Du, T., Xu, Z., Aronin, N., and Zamore, P.D. (2003). Asymmetry in the assembly of the RNAi enzyme complex. Cell 115, 199-208.

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

Shaner, N.C., Campbell, R.E., Steinbach, P.A., Giepmans, B.N., Palmer, A.E., and Tsien, R.Y. (2004). Improved monomeric red, orange and yellow fluorescent proteins derived from Discosoma sp. red fluorescent protein. Nat Biotechnol 22, 1567-1572.

Sharp, D.J., Rogers, G.C., and Scholey, J.M. (2000). Microtubule motors in mitosis. Nature 407, 41-47.

Siomi, M.C., Tsukumo, H., Ishizuka, A., Nagami, T., and Siomi, H. (2005). A potential link between transgene silencing and poly(A) tails. RNA 11, 1004-1011.

Spradling, A.C., Stern, D.M., Kiss, I., Roote, J., Laverty, T., and Rubin, G.M. (1995). Gene disruptions using P transposable elements: an integral component of the Drosophila genome project. Proc Natl Acad Sci U S A 92, 10824-10830.

Steller, H., and Pirrotta, V. (1985). A transposable P vector that confers selectable G418 resistance to Drosophila larvae. EMBO J 4, 167-171.

Sun, M., Hurst, L.D., Carmichael, G.G., and Chen, J. (2006). Evidence for variation in abundance of antisense transcripts between multicellular animals but no relationship between antisense transcriptionand organismic complexity. Genome Res 16, 922-933.

Tabach, Y., Billi, A.C., Hayes, G.D., Newman, M.A., Zuk, O., Gabel, H., Kamath, R., Yacoby, K., Chapman, B., Garcia, S.M., et al. (2013). Identification of small RNA pathway genes using patterns of phylogenetic conservation and divergence. Nature 493, 694-698.

Takamizawa, J., Konishi, H., Yanagisawa, K., Tomida, S., Osada, H., Endoh, H., Harano, T., Yatabe, Y., Nagino, M., Nimura, Y., et al. (2004). Reduced expression of the let-7 microRNAs in human lung cancers in association with shortened postoperative survival. Cancer Res 64, 3753-3756.

Tam, O.H., Aravin, A.A., Stein, P., Girard, A., Murchison, E.P., Cheloufi, S., Hodges, E., Anger, M., Sachidanandam, R., Schultz, R.M., et al. (2008). Pseudogene-derived small interfering RNAs regulate gene expression in mouse oocytes. Nature 453, 534-538.

Tan-Wong, S.M., Zaugg, J.B., Camblong, J., Xu, Z., Zhang, D.W., Mischo, H.E., Ansari, A.Z., Luscombe, N.M., Steinmetz, L.M., and Proudfoot, N.J. (2012a). Gene Loops Enhance Transcriptional Directionality. Science.

Tan-Wong, S.M., Zaugg, J.B., Camblong, J., Xu, Z., Zhang, D.W., Mischo, H.E., Ansari, A.Z., Luscombe, N.M., Steinmetz, L.M., and Proudfoot, N.J. (2012b). Gene loops enhance transcriptional directionality. Science 338, 671-675.

Tomari, Y., Matranga, C., Haley, B., Martinez, N., and Zamore, P.D. (2004). A protein sensor for siRNA asymmetry. Science 306, 1377-1380.

Townley-Tilson, W.H., Pendergrass, S.A., Marzluff, W.F., and Whitfield, M.L. (2006). Genome-wide analysis of mRNAs bound to the histone stem-loop binding protein. RNA 12, 1853-1867.

van Rooij, E., Sutherland, L.B., Qi, X., Richardson, J.A., Hill, J., and Olson, E.N. (2007). Control of stress-dependent cardiac growth and gene expression by a microRNA. Science 316, 575-579.

Wang, X., Wang, P., Sun, S., Darwiche, S., Idnurm, A., and Heitman, J. (2012). Transgene induced co-suppression during vegetative growth in Cryptococcus neoformans. PLoS Genet 8, e1002885.

Watanabe, T., Totoki, Y., Toyoda, A., Kaneda, M., Kuramochi-Miyagawa, S., Obata, Y., Chiba, H., Kohara, Y., Kono, T., Nakano, T., et al. (2008). Endogenous siRNAs from naturally formed dsRNAs regulate transcripts in mouse oocytes. Nature 453, 539-543.

Wei, W., Ba, Z., Gao, M., Wu, Y., Ma, Y., Amiard, S., White, C.I., Rendtlew Danielsen, J.M., Yang, Y.G., and Qi, Y. (2012). A role for small RNAs in DNA double-strand break repair. Cell 149, 101-112.

Werren, J.H. (2011). Selfish genetic elements, genetic conflict, and evolutionary innovation. Proc Natl Acad Sci U S A 108 Suppl 2, 10863-10870.

Xia, T., SantaLucia, J., Jr., Burkard, M.E., Kierzek, R., Schroeder, S.J., Jiao, X., Cox, C., and Turner, D.H. (1998). Thermodynamic parameters for an expanded nearest-neighbor model for formation of RNA duplexes with Watson-Crick base pairs. Biochemistry 37, 14719-14735.

Yelin, R., Dahary, D., Sorek, R., Levanon, E.Y., Goldstein, O., Shoshan, A., Diber, A., Biton, S., Tamir, Y., Khosravi, R., et al. (2003). Widespread occurrence of antisense transcription in the human genome. Nat Biotechnol 21, 379-386.

Yenerall, P., and Zhou, L. (2012). Identifying the mechanisms of intron gain: progress and trends. Biol Direct 7, 29.

Yu, J.Y., Reynolds, S.H., Hatfield, S.D., Shcherbata, H.R., Fischer, K.A., Ward, E.J., Long, D., Ding, Y., and Ruohola-Baker, H. (2009). Dicer-1-dependent Dacapo suppression acts downstream of Insulin receptor in regulating cell division of Drosophila germline stem cells. Development 136, 1497-1507.

Zambon, R.A., Vakharia, V.N., and Wu, L.P. (2006). RNAi is an antiviral immune response against a dsRNA virus in Drosophila melanogaster. Cell Microbiol 8, 880-889.

Zhang, B., Pan, X., Cobb, G.P., and Anderson, T.A. (2007). microRNAs as oncogenes and tumor suppressors. Dev Biol 302, 1-12.

Zhou, R., Czech, B., Brennecke, J., Sachidanandam, R., Wohlschlegel, J.A., Perrimon, N., and Hannon, G.J. (2009). Processing of Drosophila endo-siRNAs depends on a specific Loquacious isoform. RNA 15, 1886-1895.

Zhou, R., Hotta, I., Denli, A.M., Hong, P., Perrimon, N., and Hannon, G.J. (2008). Comparative analysis of argonaute-dependent small RNA pathways in Drosophila. Mol Cell 32, 592-599.