Dissertation

Centromere regulation by the RNA-binding proteins Fmr1 and Rump in the germline of *Drosophila melanogaster*

Janina Luitz

2022

Inaugural dissertation

for

obtaining the doctoral degree

of the

Combined Faculty of Mathematics, Engineering and Natural Sciences

of the

Ruprecht - Karls - University

Heidelberg

Presented by
M.Sc. Janina Luitz
born in Schongau, Germany
Oral examination: 22.02.2022

Centromere regulation by the RNA-binding proteins Fmr1 and Rump in the germline of *Drosophila melanogaster*

Referees:
Prof. Dr. Sylvia Erhardt
Prof. Dr. Dr. Georg Stoecklin

This thesis is dedicated to my father
In loving memory

Summary

Every cell division relies on functional centromeres to guarantee correct chromosome segregation. They are particularly important in the germline to ensure successful reproduction and embryonic development. Centromeres are epigenetically defined by the incorporation of the histone H3 variant CENP-A and are embedded into pericentromeric chromatin. Non-coding RNAs generated by transcription from centromeric and pericentromeric regions propagate centromere function and specification. Yet, little is known about how (peri-) centromeric RNAs are regulated and which interaction partners they have.

The primary goal of my doctoral thesis was to identify factors that are associated with and influence the function of the lncRNA SatIII from the (peri-) centromeric regions of the X chromosome of *Drosophila melanogaster*. In this thesis, I identified RNA-binding proteins (RBPs) present at the centromere by re-analyzing a published mass spectrometry dataset that was produced by a crosslinked CENP-A pulldown approach. In doing so, I found out that the RBP Fmr1 is one of the top enriched factors and demonstrated that it binds to SatIII RNA *in vitro*. Additionally, I performed Fmr1 CLIP-Seq and showed that Fmr1 binds to (peri-)centromeric RNAs in the female germline. Moreover, Fmr1 associates with a subset of mRNAs including some that encode proteins which are important for centromere function, and also positively affects their translation. In order to find novel interactors of Fmr1 mass spectrometry was used. Thereby, I was able to identify the RBP Rump, which interacts with Fmr1 in an RNA-dependent manner, indicating their co-binding to common target RNAs. In fact, there is an overlap between Fmr1 and Rump CLIP-Seq data in the non-coding RNAs that both proteins target. Additionally, I produced RNA-Seq data and showed that both proteins regulate the levels of non-coding RNAs that emerge from the centromere and elsewhere in the genome (in the male germline). Finally, Rump was shown to be required in maintaining CENP-A levels in mature sperm.

Overall, this thesis demonstrates an implication of the RBPs Fmr1 and Rump in centromere function and regulation in the germline of *Drosophila melanogaster*.

Zusammenfassung

Jede Zellteilung ist auf funktionierende Zentromere angewiesen, um eine korrekte Chromosomentrennung zu gewährleisten. In der Keimbahn sind sie besonders wichtig, um eine erfolgreiche Reproduktion und Embryonalentwicklung sicherzustellen. Zentromere werden epigenetisch durch den Einbau der Histon H3 Variante CENP-A definiert und sind in perizentromerisches Chromatin eingebettet. Nicht-kodierende RNAs, die durch Transkription aus zentromerischen und perizentromerischen Regionen erzeugt werden, fördern die Funktion und Spezifikation des Zentromers. Dennoch ist wenig darüber bekannt, wie (peri-)zentromerische RNAs reguliert werden und welche Interaktionspartner sie haben.

Das primäre Ziel meiner Doktorarbeit war es, Faktoren zu identifizieren, die mit der lncRNA SatIII aus den (peri-)zentromerischen Regionen des X-Chromosoms von *Drosophila melanogaster* assoziiert sind und deren Funktion beeinflussen. In dieser Dissertation identifizierte ich RNA-bindende Proteine (RBPs), die am Zentromer vorhanden sind, indem ich einen veröffentlichten Massenspektrometrie-Datensatz, der durch einen vernetzten CENP-A Pulldown Ansatz erzeugt wurde, erneut analysierte. Dabei fand ich heraus, dass das RBP Fmr1 einer der am stärksten angereicherten Faktoren ist und zeigte, dass es *in vitro* an SatIII RNA bindet. Zusätzlich habe ich Fmr1 CLIP-Seq durchgeführt und gezeigt, dass Fmr1 an (peri-)zentromerische RNAs in der weiblichen Keimbahn bindet. Darüber hinaus assoziiert Fmr1 mit einer Untergruppe von mRNAs, darunter einige, die für Proteine kodieren, die für die Funktion des Zentromers wichtig sind, und beeinflusst auch deren Translation positiv. Um neue Interaktoren von Fmr1 zu finden, wurde Massenspektrometrie durchgeführt. Auf diese Weise konnte ich das RBP Rump identifizieren, das mit Fmr1 in einer RNA-abhängigen Weise interagiert, was darauf hindeutet, dass sie gemeinsam an RNAs binden. Tatsächlich überschneiden sich die CLIP-Seq Daten von Fmr1 und Rump bei den nicht-kodierenden RNAs, an die beide Proteine binden. Darüber hinaus konnte ich anhand von RNA-Seq Daten zeigen, dass beide Proteine die Menge nicht-kodierender RNAs regulieren, die aus dem Zentromer und an anderen Stellen im Genom (in der männlichen Keimbahn) entstehen. Schließlich wurde gezeigt, dass Rump für

die Aufrechterhaltung der CENP-A Pegel in reifen Spermien erforderlich ist.

Insgesamt zeigt diese Doktorarbeit, dass die RBPs Fmr1 und Rump an der Funktion und Regulierung des Zentromers in der Keimbahn von *Drosophila melanogaster* beteiligt sind.

Contents

Chapter 1

Introduction

1.1 Chromatin organisation and the regulatory epigenome

The term "epigenetics" was introduced in 1942 by Conrad Waddington who described it as "the branch of biology which studies the causal interactions between genes and their products, which bring the phenotype into being" (Waddington, 1942). Broadly speaking, epigenetics involves heritable changes in gene function without altering the underlying DNA sequence (Russo et al., 1996). Over the years the study of epigenetics has branched into a new field of study and many molecular mechanisms with underlying epigenetic phenomena have been uncovered. Epigenetic effects on phenotypic traits may be a result of environmental cues or be part of development itself. For example, all cells of multicellular organisms share the same genetic information and yet can differentiate into various cell types with different functions (Goldberg et al., 2007). Epigenetic studies converge on DNA modifications and the biology of chromatin, including chromatin marks, higher order structure of chromosomes and RNAs involved in chromatin regulation (Bird, 2007).

1.1.1 Chromatin structure

Chromatin, first described by W. Flemming in 1878, is a complex of DNA, RNA and associated protein that organizes the genomic material within the nuclei of all eukaryotic cells (Flemming, 1878; Yadav et al., 2018). The organi-

zation of chromatin follows a hierarchical principle from single nucleosomes to the condensation into higher order chromatin structures (Fig. 1.1). The nucleosome represents the basic unit of chromatin, with ~146 base pairs of DNA wrapped 1.65 times around a histone octamer, which consists of two copies of each histone protein, H2A, H2B, H3 and H4. Nucleosomes are joined together through linker DNA, which is associated with the linker histone (H1), thereby building up the primary chromatin structure often referred to as "beads on a string" (Kornberg, 1977; Luger et al., 1997; Olins and Olins, 1974). Nucleosomes are coiled into 30 nm chromatin fibers, which are further folded for additional compaction with increasing complexity to reach higher-order organization with chromatin interaction domains or so called topologically associated domains (TADs) (Dixon et al., 2012). The highest compaction of DNA is reached during cell division, where chromatin is condensed into chromosomes (Alberts et al., 2002). The organization into chromosomes is crucial to ensure the faithful distribution of genomic material into the daughter cells during mitosis. The fidelity of chromosome segregation is further ensured by a unique chromatin site, the centromere, which attaches to spindle microtubules that pull apart the sister chromatids (Villasante et al., 2007). Not only the centromere, but also telomeres and regulatory elements display distinct architectural features in chromatin composition, which are vital for cell viability and proliferation (Probst et al., 2009).

The degree of genome compaction distinguishes two types of chromatin states: euchromatin and heterochromatin (Grewal and Jia, 2007). Euchromatin is a lightly packed form of chromatin, which is often gene-rich and, therefore, requires to be easily accessible to promote transcription in these regions. In contrast, heterchromatin represents a highly compacted and condensed state of chromatin, which is more inaccessible and has been generally described as transcriptionally inactive (Huisinga et al., 2006). Heterochromatin can be further categorized into constitutive or facultative heterochromatin. The latter type can form at different chromosomal regions, usually containing genes that require to be dynamically silenced or activated depending on cellular signals and developmental cues. In contrast, constitutive heterochromatin occurs at the same genomic region in every cell type, at centromeres and telomeres that bear a high amount of repetitive DNA and transposable elements and are

generally low in genes. Therefore, constitutive heterochromatin is believed to be more static than facultative heterochromatin (Brown, 1966; Grewal and Jia, 2007). However, in the last decades it has become evident that constitutive heterochromatin is transcriptionally active in spite of the highly compacted state, which will be discussed later (Hall et al., 2012; Azzalin and Lingner, 2008).
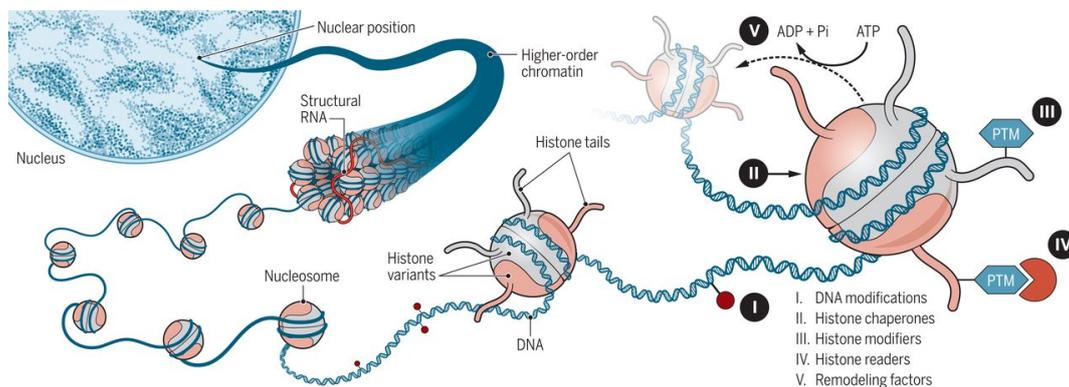
Figure 1.1: Chromatin structure and compaction. The schematic depicts chromatin organization during interphase. DNA is wrapped around a histone octamer to form the basic unit of chromatin - the nucleosome. Arrays of nucleosomes are further organized into higher-order chromatin with RNA. Chromatin is variable depending on DNA or histone modifications, histone variants and chromatin remodeling (Yadav et al., 2018).

1.1.2 The epigenome and epigenetic inheritance

Chromatin not only serves the purpose of packaging DNA, but carries diverse gene regulatory information to enable chromatin dynamics from accessible to inaccessible states and vice versa. Chromatin structure and function can be altered and regulated by currently five known key mechanisms that are propagated through mitotic cell divisions: DNA methylation, post-translational modifications of histone proteins, chromatin remodeling, histone variants and chromatin-associated non-coding RNAs (Allis and Jenuwein, 2016). Chromatin with all these characteristics forms the epigenome and numerous processes such as imprinting, X-inactivation, heterochromatin formation and maintenance and transcriptional regulation are regulated by the above mentioned epigenetic mechanisms (Almouzni and Cedar, 2016).

DNA methylation is a chemical modification of the cytosine bases which can be found in the context of CpG dinucleotides at for instance promoters and CpG islands, and is associated with gene silencing and heterochromatin formation (Greenberg and Bourchis, 2019). Histone modifications add another epigenetic regulatory layer to chromatin and are thought to constitute a "histone code", which is set by histone modifying enzymes thereby defining transcriptional states. Numerous histone modifications have been described including acetylation, phosphorylation and methylation, that mostly occur at the histone tail that protrudes from the nucleosome. Inter-nucleosomal interactions are affected by histone modification, which as a result affect overall chromatin structure (Jenuwein and Allis, 2001). Furthermore, histone modifications recruit chromatin remodeling enzymes, which are able to change the packaging state of chromatin by depositing or ejecting already integrated nucleosomes. Chromatin remodeler can also alter the composition of the nucleosomes by replacing canonical histones with histone variants (Clapier and Cairns, 2009). Contrary to canonical histones, histone variants are incorporated throughout the cell cycle in a replication-independent manner and attribute different structural properties on nucleosomes leading to distinct functions. For example, the incorporation of the histone H3 variant CENP-A results in the formation of centromeric chromatin (Talbert and Henikoff, 2021).

A variety of non-coding RNAs have been shown to be essential in chromatin regulation. RNAs can associate at the site of their synthesis (cis-acting RNAs) or interact with different genomic loci (trans-acting RNAs). Gene expression can be affected through direct RNA-DNA contact or more indirect through RNA-binding proteins (RBPs) that interact with the RNA and localize to chromatin (Khelifi and Hussein, 2020; Li and Fu, 2019). For example, Xist RNA, one of the first discovered lncRNA in the 1990s (Brown et al., 1991), is considered to be the master regulator of X-chromosome inactivation by coating the entire X-chromosome and transforming it into a uniquely organized heterochromatic entity (Loda and Heard, 2019). Furthermore, heterochromatin formation can be promoted by the RNA intereference pathway, in which heterochromatic transcripts guide the transcriptional silencing complex (RISC) to transcription sites leading to the recruitment of epigenetic factors that modify chromatin (Martienssen and Moazed, 2015; Volpe et al., 2002).

1.2 The centromere

Cell division is a fundamental process for life. When a cell divides, the duplicated genome is distributed equally to the daughter cells. The process is tightly controlled to maintain the integrity of the genome. During mitosis, condensed chromosomes are separated through movement along the mitotic spindle. This movement is powered and regulated by the centromere, a specific chromatin domain that is essential to guarantee the fidelity of chromosome segregation. The centromere is present throughout the cell cycle and serves as transient attachment site for kinetochore formation which mediates the association of spindle microtubules during mitosis (Allshire and Karpen, 2008). In general, impaired centromere function results in chromosome segregation defects that are detrimental to the cell, leading to aneuplody, which can cause genetic disorders or cancer (Taylor et al., 2018; Hassold and Hunt, 2001).

1.2.1 The histone H3 variant CENP-A

One of the most unifying characteristic of centromeres is the presence of nucleosomes containing histone H3 variant centromere protein A (CENP-A or CID in *Drosophila melanogaster*). CENP-A is highly conserved from yeast to human and represents one of the key epigenetic marks of the centromere, as its integration is independent of the underlying DNA sequence (Black and Bassett, 2008). The maintenance of genome integrity depends on the proper regulation of CENP-A levels, as both the depletion and overexpression of CENP-A promote mitotic defects and centromeric instability (Maehara et al., 2010; Shrestha et al., 2017). CENP-A was discovered by William Earnshaw in 1985, who examined samples from patients with CREST syndrome of scleroderma and found antibodies specific for CENP-A, B and C appearing in their antisera (Earnshaw and Rothfield, 1985). CENP-A shares 62% identity at the C-terminal histone-fold domain with the canonical histone H3 and has a unique N-terminal tail (Sullivan et al., 1994). The region within the C-terminal histone-fold domain, the CENP-A targeting domain (CATD) is known to be sufficient and necessary to target CENP-A to the centromere (Black et al., 2007). The N-terminal tail varies between species and is required to recruit kinetochore proteins to the centromere in yeast (Chen et al., 2000). Kineto-

chore recruitment during mitosis is one of the major functions of CENP-A (Musacchio and Desai, 2017; Fukagawa and Earnshaw, 2014).

1.2.2 Kinetochore assembly

The kinetochore is a large proteinaceous structure implicated in the attachment of microtubules to chromosomes during mitosis and can be grouped into an inner and outer kinetochore. The inner kinetochore creates an interface with centromeric chromatin, whereas the outer kinetochore contributes to a microtubule-binding interface (Santaguida and Musacchio, 2009). In flies, the inner kinetochore is built through the recruitment of CENP-C to the centromere. CENP-C localisation to the centromere depends on CENP-A and its loading factor Cal1. CENP-A loading is replication-independent and occurs during mitosis in *Drosophila* (Schuh et al., 2007). All three factors are required for centromere localisation and function (Erhardt et al., 2008; Mellone et al., 2011). CENP-C serves as a linker between CENP-A in centromeric chromatin and the outer kinetochore, which in part consists of the KMN network (KNL1/Spc105, the Mis12 complex, and the Ndc80 complex) (Przewloka et al., 2011). Spindle attachment and the alignment of chromosomes on the metaphase plate are monitored by the spindle assembly checkpoint (SAC), which delays the onset of anaphase until errors in spindle attachments are corrected (Musacchio and Salmon, 2007). The transition from metaphase to anaphase is regulated by the anaphase promoting complex/cyclosome (APC/C), a ubiquitin ligase complex that initiates sister-chromatid separation and exit from mitosis by targeting cyclin B and securin for destruction by the 26S proteasome (Peters, 2006). The activation of the APC/C is prevented by the Mitotic Ceckpoint Complex (MCC) with the kinetochore proteins Mad2, BubR1 and Bub3 that sequester Cdc20, a co-activator of the APC/C. Other proteins such as Bub1, Mad1, the Mps1 and the RZZ complex have been also shown to be involved in the SAC (Fig. 1.2) (Musacchio, 2015).
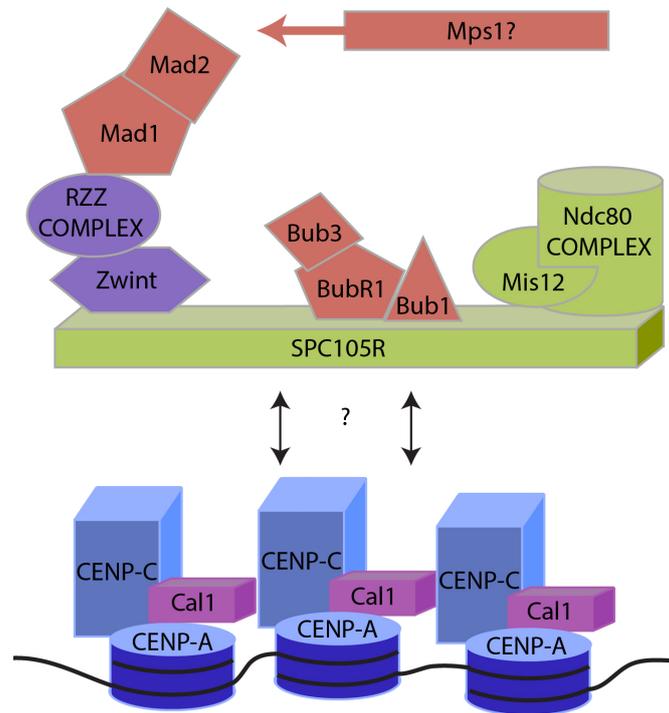
Figure 1.2: The kinetochore in *Drosophila melanogaster*. Schematic of the centromere-kinetochore interface. CENP-A (CID in *Drosophila*) is deposited into centromeric chromatin by the chaperone Cal1, which interacts with CENP-C for centromere recognition (proteins in blue are DNA-binding). Components of the outer kinetochore mediate the spindle attachment (proteins in green) and checkpoint proteins supervise the process (proteins in red). Adapted from Orr et al. (2010).

1.2.3 Centromeric DNA

Although centromeres play a key function during mitosis, centromeric DNA represents one of the most rapidly evolving sequence in the genome (Bayes and Malik, 2008). Furthermore, centromeric DNA can not only differ between species but also between chromosomes within the same species. In contrast, centromeres contain highly conserved proteins. Therefore, the location of centromeres is not determined genetically but rather epigenetically by sequence-independent mechanisms (Sullivan et al., 2001). Despite the lack of sequence identity, centromeres are characterized by repetitive DNA and satellites, often rich in A/T nucleotides and arranged in tandem units in many organisms (Aldrup-MacDonald and Sullivan, 2014). The fact that centromeric repeats have endured evolutionary, implies an integral role of reiterated DNA in supporting centromere formation and function. However, centromeres deprived of satellite repeats exist in chromosomes of horses, zebras and donkeys, guaranteeing robust centromere propagation, which brings the requirement of DNA

repeats at centromeres under scrutiny (Giulotto et al., 2017). Furthermore, neocentromeres that are functionally and structurally similar to endogeneous centromeres can form at non-repetitive sites (Alonso et al., 2010). Therefore, specific properties of centromeric DNA or chromatin environment might be required to establish functional centromeres rather than the sequence itself. Indeed, Heterochromatin-associated Protein 1 (HP1) is recruited to several non-repetitive neocentromeres, indicating that centromere function depends on heterochromatin formation (Aagaard et al., 2000; Henikoff et al., 2001). Other research suggests that secondary DNA structures manifest centromeres and are required for specific DNA-binding factors (Kasinathan and Henikoff, 2018; Gallego et al., 1997). For example, human centromeric alpha satellites, murine centromeric Y-satellites and the *Drosophila* Dodeca repeats that are present at centromeres in the respective species are shown to form dimeric i-motifs (Garavís, Méndez-Lago, Gabelica, Whitehead, González and Villasante, 2015; Garavís, Escaja, Gabelica, Villasante and González, 2015).

The primary DNA sequence of all human centromeres is composed of alpha satellite repeats. Monomers of 171 bp are tandem repeated and organized in arrays building a higher order repeat (HOR). HOR themselves are reiterated numerous times and span megabases (Willard, 1985; Aldrup-MacDonald and Sullivan, 2014). In *Drosophila*, centromeric DNA contains large amounts of repeats, including satellite repeats, as well as transposable elements and tandemly repeated rRNA genes (Le et al., 1995). However, until recently, the repetitive character of *Drosophila* centromeres has been refractory to DNA sequencing and assembly, leading to large gaps in centromeric sequence assembly due to the lack of sequence information at and around centromeres. Chang et al. (2019) have combined long-read sequencing and chromatin immunoprecipitation for CENP-A and as a result assembled a novel *Drosophila* genome, that includes centromeric sequences with islands of complex DNA sequences enriched in retroelements. The islands were shown to be flanked by arrays of satellite repeats. Interestingly, the DNA composition, although quite similar, differs between centromeres and each centromere is distinct in size and arrangement of its DNA elements. The most enriched retroelement was shown to be G2/Jockey-3, which is also shared among all centromeres. Furthermore, CENP-A was enriched with the intergenic spacer of ribosomal genes (IGS)

present in 240 copies on chromosome 3 (Chang et al., 2019; Shatskikh et al., 2020). Other transposable elements, such as Doc and Doc2 were also commonly found at centromeres. It is worth noting that transposable elements enriched at centromeres in *Drosophila* are not unique and can be found elsewhere in the genome. Short tandem repeats, such as (AATAG)n and (AATAT)n, Dodeca satellite and Prodsat are shown to surround the retroelements making up the majority of centromeric tandem repeats (Talbert et al., 2018; Chang et al., 2019). The centromere is flanked by pericentromeric heterochromatin of highly repetitive nature and therefore not assembled yet (Fig. 1.3).
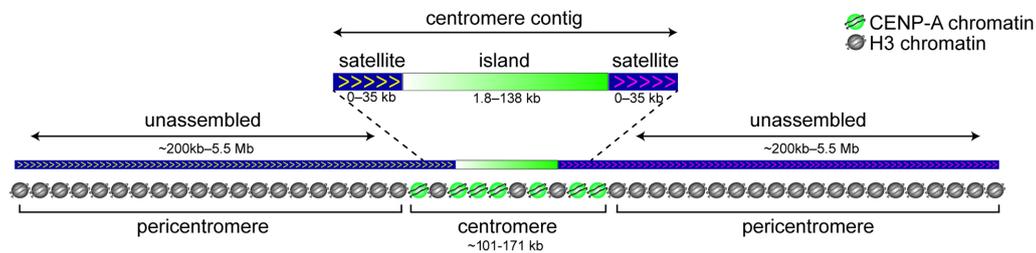
Figure 1.3: Centromere organization in *Drosophila melanogaster* CENP-A chromatin is associated with centromere islands containing complex DNA sequences enriched in retroelements. The size of CENP-A domains ranges between 101 and 171 kb. Islands are flanked by satellites and centromeres are embedded into large pericentromeric regions that span megabases (Chang et al., 2019).

1.2.4 Centromeres are embedded in pericentromeric heterochromatin

In *Drosophila*, pericentromeric DNA differs from centromeric DNA. A major part of pericentromeric chromatin contains complex satellite DNA, including satellites of the 1.688 family and Responder (Rsp). Rsp satellite DNA consists of a dimer of two related 120 bp repeats and is found at the pericentromere of chromosome 2R (Larracuente, 2014). The 1.688 family are a family of related repeats with 356 bp and 353 bp repeats located in the pericentromeric region of chromosome 3L and 260 bp repeats on chromosome 2L. The longest arrays of 1.688 satellite DNA are found in the pericentromeric region of the X-chromosome with 359 bp repeats. There are also shorter arrays of 359 bp in heterochromatic regions of chromosome 2 and 3 (Losada and Villasante, 1996; Abad et al., 2000).

The pericentromere differs from the centromere not only by DNA sequence and organization, but also by a distinct chromatin signature put in place at each region. In fact, centromeric chromatin exhibits a unique histone modification pattern that differs from both euchromatin and heterochromatin, whereas pericentromeric chromatin carries typical heterochromatic histone marks (Sullivan and Karpen, 2004). In more detail, canonical histone H3 that is interspersed in centromeric regions is methylated at lysine 4 and lysine 36 (H3K4me1, H3K4me2, H3K36me2, H3K36me3), producing histone marks commonly found in euchromatin. Conversely, histones H3/H4 lack acetylation at centromeres, which is normally associated with open (active) chromatin (Bergmann et al., 2011). Furthermore, H3K9me3, a histone modification associated with transcriptional repression is also found within centromeric chromatin (Bergmann et al., 2012). This suggests that centromeric chromatin embodies both heterochromatic and euchromatic states (Sullivan and Karpen, 2004). In contrast, pericentromeric regions contain only repressive histone marks such as H3K9me2/me3, H3k27me3 and K4K20me3 and are associated with HP1 to help form and maintain the heterochromatic state and coordinate cohesion and replication activities (Nishibuchi and Déjardin, 2017; Saksouk et al., 2015). Furthermore, non-coding RNAs that are transcribed from heterochromatin and processed into small RNAs have been shown to be important for heterochromatin formation in several organisms (Volpe et al., 2002), as discussed below. Transcription at heterochromatic regions sounds paradox, due to the high compaction and because heterochromatin itself induces transcriptional silencing. Nevertheless, there is active transcription at pericentromeric and centromeric regions, with important roles for heterochromatin establishment and centromere function (Smurova and De Wulf, 2018).

1.2.5 (Peri-)centromeric chromatin is transcriptionally active

Transcription at (peri-)centromeres occurs in many species and has first become evident by the presence of RNA Polymerase II at these sites and by the production of centromere RNA transcripts (Wong et al., 2007; Saffery et al., 2003; Chan et al., 2012). RNA Polymerase II is phosphorylated at Ser2 of the C-terminal domain indicating active transcription elongation, which has

been demonstrated by immunofluoresence assays (Rošić et al., 2014; Molina et al., 2016). In *Drosophila*, the recruitment of RNA polymerase II depends on the CENP-A loading factor Cal1 and the chromatin remodeling activites of the FACT (Facilitates Chromatin Transcription) complex (Chen et al., 2015). Although we have knowledge of RNA polymerases and centromeric transcription, nothing is known about the promoters and only a few transcription factors have been discovered. The identification of promoter sequences has been challenging, due to the high repetitive nature of centromeres, but studies proposed that centromeric transcription is driven by retroviral promoters (Carone et al., 2009; Topp et al., 2004). Despite the gaps in understanding how centromere transcription is carried out, research has advanced to uncover the functional relevance of centromeric RNAs to a certain extent which is outlined below.

1.2.6 (Peri-)centromeric transcripts and function

Two types of transcripts are known to emerge from (peri-)centromeres: small RNAs and long non-coding RNAs (lncRNAs) (Arunkumar and Melters, 2020). Due to the high repetive nature of centromeres and the fact that repeats span megabases, it has been a challenge to identify length and sequence of lncRNAs. Indeed, studies show a high variability in the length of cenRNAs. For example different sizes for satellite RNA have been reported in human cells ranging from 0.3 - 2 kb (McNulty et al., 2017). In *Drosophila* SatIII RNA emerges from arrays of 359 bp repeats from pericentromeric regions and is transcribed in sense and antisense direction. The longest polyadenylated SatIII transcript found by 3' rapid amplification of cDNA ends (RACE) was 1.3 kb in length, corresponding to 4 repeating units (Rošić et al., 2014). However, it remains yet to be determined whether there are longer and/or non-polyadenylated SatIII transcripts. Retrotransposons at centromeres have also been shown to give rise to lncRNAs (Neumann et al., 2007; Topp et al., 2004). Small RNAs originate from centromeric DNA across kingdoms and have been shown to strongly associate with the RNA interference (RNAi) machinery (Castel and Martienssen, 2013).

The role of RNAs in heterochromatin formation

Heterochromatin formation by RNA interference RNAi is a process that serves the repression of transcription through small interfering RNAs (siRNAs) (Fire et al., 1998). Furthermore, in yeast RNAi has a well established role in the formation of constitutive heterochromatin at pericentromeric regions. Thereby, pericentromeric RNA is transcribed bidirectionally to form double stranded RNA (dsRNA) which is processed by Dicer ribonuclease into siRNA (Volpe et al., 2002). siRNAs are loaded into Ago1, the main component of the RNAi transcriptional silencing complex (RITS) which together with other factors localizes to pericentromeric heterochromatin resulting in H3K9 methylation and spreading. In conclusion, siRNAs bind to nascent (peri-)centromeric non-coding RNAs, which means that the (peri-)centromeric RNAs can repress their own transcription (Shimada et al., 2016).

The Piwi-interacting RNA (piRNA) pathway is another form of RNAi that has evolved in the germline of many organisms as a defence system for transposable elements (TEs). The mobilization of TEs is a danger to the genome, as it can trigger mutations, DNA breaks and chromosome rearrangements. Therefore, transposon silencing is particularly important in the germline, as the genome is transmitted to the future generation (Khurana and Theurkauf, 2010). piRNAs are derived from piRNA clusters containing partial or defective nested transposons and satellite sequences, which result in a library of those sequences. The majority of piRNA cluster is present at the pericentromere and telomere (Brennecke et al., 2007). At the core of the piRNA pathway, transcription of piRNAs is induced by the Rhino-Deadlock complex, with Rhino being a HP1 homolog that recognizes heterochromatic marks and Deadlock a recruiter of Moonshiner, a TFIIA-L paralog of RNA polymerase II preinitiation complex (Klattenhoff et al., 2009; Mohn et al., 2014; Andersen et al., 2017). In germline cells, piRNAs undergo two types of processing mechanism; in the first one, piRNAs are generated to target transposons from single-stranded precursor RNAs. In the second one, the sense transcripts of transposons are cleaved to produce sense piRNAs in a process guided by antisense piRNAs. This leads to a further cleavage of cluster transcripts for antisense piRNA production, forming a piRNA amplification loop (ping-pong cycle). The piRNAs

are loaded into Piwi, forming the piRNA-induced silencing complex, resulting in transposon silencing by inducing heterochromatin formation (Iwasaki et al., 2015). Depletion of proteins central to the piRNA pathway, such as Aub and Piwi results in the derepression of pericentric heterochromatin and the accumulation of transposable elements and other piRNA target RNAs (Teo et al., 2018; Gu and Elgin, 2013).

RNAs as regulator of heterochromatin by phase separation Dynamic compartmentalisation occurs in the nucleus with RNA playing a crucial role as architectural factor and phase separation being the driving mechanism (Caudron-Herger and Rippe, 2012; Nozawa and Gilbert, 2019; Erdel and Rippe, 2018). Phase separation describes a physical process in which a supersaturated phase coexists with a dilute phase in a solution, similar to the demixing of oil drops in water. It is driven by multivalent interactions, often between proteins with low complexity intrinsically disordered regions (IDRs) and RNA (Alberti et al., 2019). Heterochromatin formation has been linked to phase separation mechanism by the accumulation of HP1 (Larson et al., 2017; Strom et al., 2017) and RNA derived from major Satellite repeats is required to anchor HP1 at pericentric heterochromatin (Johnson et al., 2017). Another study showed, that SAFB nuclear matrix protein interacts with mouse Satellite RNAs, leading to chromatin condensation and stabilization of heterochromatin by phase separation (Huo et al., 2020). TERRA, a repeat-containing RNA from telomeres, is connected to heterochromatin establishment, which is reminiscent of the above mentioned pericentric Satellite repeats that recruit proteins to maintain silencing (Arnoult et al., 2012). Taken together, repetitive RNAs have the potential to nucleate phase separation processes by locally concentrating proteins involved in H3K9me3-mediated heterochromatin formation (Frank and Rippe, 2020).

The role of RNAs in centromere function

In the last decade it has become more and more evident that centromeric transcription and the resulting transcripts are essential for centromere function. In several organisms, CENP-A loading and maintenance was shown to depend on both the act of transcription and centromeric transcripts. Depletion of cen-

tromeric RNAs in Xenopus eggs and mammalian cell lines results in reduced
CENP-A levels at the centromere (Grenfell et al., 2016; Wong et al., 2007;
Chan et al., 2012). In this line, recent work in *Drosophila* S2 cells showed
that RNA Polymerase II transcription results in the incorporation and stabi-
lization of CENP-A (Bobkov et al., 2018). Furthermore, the inner kinetochore
protein CENP-C is RNA-binding and required for stable spindle association
and centromere maintenance (McNulty et al., 2017; Rošić et al., 2014). The
Erhardt laboratory identified SatIII RNA to be bound by CENP-C in flies
and the depletion of SatIII with LNA probes results in severe mitotic defects
(Rošić et al., 2014). Centromeric chromatin organization also involve RNAs,
which is highlighted by the fact that R-loops (RNA-DNA Triplex) can form at
centromeres and promote correct sister chromatid segregation during mitosis
(Kabeche et al., 2018).
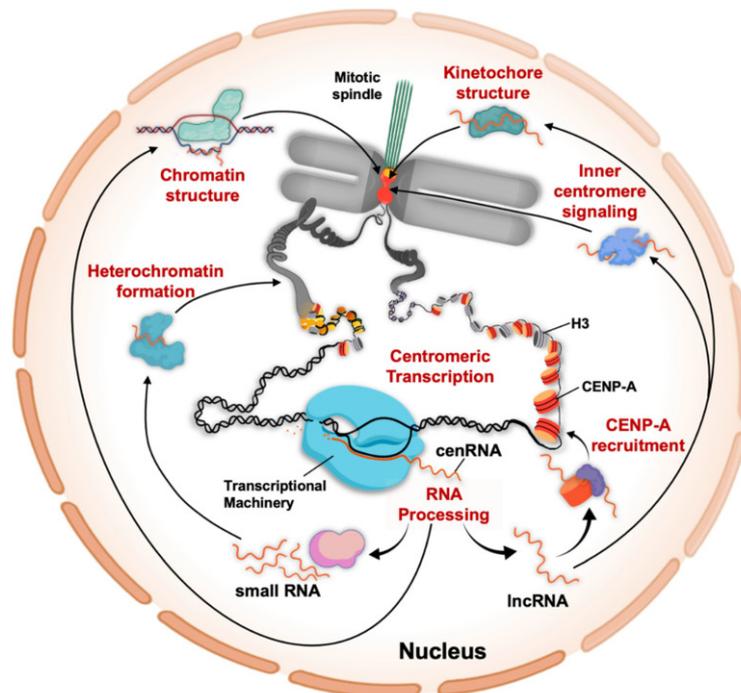
Figure 1.4: The role of centromeric transcripts. Centromeric transcription occurs in many species
and transcripts are found to be either lncRNAs or small RNAs. Pericentric heterochromatin formation is
mediated through small RNAs by the RNAi pathway. R-loops are present at centromeres and modulate the
chromatin structure. CENP-A loading and maintenance depends on both, centromeric transcription and
lncRNAs. Furthermore, the kinetochore protein CENP-C binds to centromeric RNAs, which is required for
kinetochore formation and function. At last, inner centromere signalling requires centromeric transcription
in order to promote faithful chromosome segregation (Arunkumar and Melters, 2020).

In addition to the functions above, centromeric RNAs also play an important role in inner centromere signalling. For example, Shugoshin (Sgo1), a component of the inner-centromere which protects centromeric cohesion from cleavage during prophase, brings RNA Polymerase II to the centromere and also associates with centromeric RNA. This suggests, that mitotic transcription is required to target regulatory factors to highly compacted mitotic chromatin (Liu et al., 2015). Furthermore, Aurora B, the enzymatic member of the chromosomal passenger complex that ensures mitotic completion and cytokinesis, is bound to centromeric RNA which regulates both its activity and localisation (Blower, 2016). The roles of centromeric transcription and transcripts that are described here are illustrated in Fig. 1.4.

1.3   Functions of RNA-binding proteins

Despite our knowledge of centromeric transcripts and what roles they have, we still lack the understanding in how centromeric transcripts precisely perform their functions. Generally, RNAs are rarely unaccompanied. Most RNAs require RNA-binding proteins (RBPs) and form ribonucleoprotein (RNP) complexes in order to propagate their function. For RNA recognition, RNAs usually bear conserved RNA sequence motifs or RNA recognition elements that are targeted by RBPs (Gerstberger et al., 2014). On the other hand, RBPs contain different types of RNA binding motifs that are very often present in multiple copies (Cléry and Allain, 2012). One of the most common and best characterized RNA binding motif is the RNA-recognition motif (RRM), which contains aromatic residues that interact with nucleobases of the RNA (Lunde et al., 2007). The second most prevalent RNA recognition motif is the RGG box (glycine-arginine-rich) and is intrinsically disordered (Ozdilek et al., 2017) with affinity to primary and secondary nucleic acid structures including G-quadruplex (nucleic acid structures formed by guanine tetrads arranged in a planar conformation) (Phan et al., 2011; Hanakahi et al., 1999). Another ubiquitous present RNA binding domain in eukaryotes is the heterogenous nuclear (hn)RNP K-homology domain (KH domain), which achieves RNA recognition by hydrogen bonding and electrostatic interactions instead of aromatic amino acids. The KH domain can bind both single-stranded DNA and RNA (Lewis

et al., 2000; Lunde et al., 2007). RBPs are involved in a myriad of molecular processes including RNA transport, RNA processing, mRNA export, ribosome assembly, translation and RNA decay.

Due to the growing number of centromeric RNAs that are being identified, it is essential to study RBPs in order to get a better understanding of their molecular regulation. The work of this thesis is dedicated to shining light on this aspect. Therefore, we set out to identify novel RBPs with roles in centromere function. The RBPs Fmr1 and Rump are central to this study and some of their known properties and biological functions are outlined below.

1.3.1 The hnRNP M homolog Rumpelstiltskin

Functional relevance of hnRNPs Heterogeneous nuclear ribonucleoproteins (hnRNPs) are a large family of functionally diverse RBPs with nuclear and cytoplasmic functions and are key players in RNA metabolism including splicing, mRNA stabilization and transcriptional and translational regulation (Dreyfuss et al., 1993; Geuens et al., 2016). The first hnRNPs were found by the isolation of the 40S core particle with sucrose density gradients (Beyer et al., 1977). Other hnRNPs were identified later that share general features, but differ in domain composition and functional properties (van Eekelen et al., 1981). In the nucleus, hnRNPs are known to bind to promoter and enhancer sequences, thereby directing transcription together with transcription factors and other RBPs. Thereafter, hnRNPs are stabilizing freshly transcribed pre-mRNAs and promoting splicing and export of transcripts through the nuclear pore complex. In the cytoplasm, hnRNPs regulate translation, depending on the RNP composition at translation initiation sites (Fig. 1.5) (Geuens et al., 2016). It is noteworthy that not every member of the hnRNP family is involved in every nuclear or cytoplasmic function. For example, hnRNP M localises predominantly in the nucleus and is associated with nuclear speckles. Most known functions of hnRNP M are connected to splicing of pre-mRNAs. Indeed, hnRNP M was detected in proteomic analyses with purified spliceosomes in the pre-spliceosomal H-complex and in spliceosome assembly (Rappsilber et al., 2002; Wahl et al., 2009).
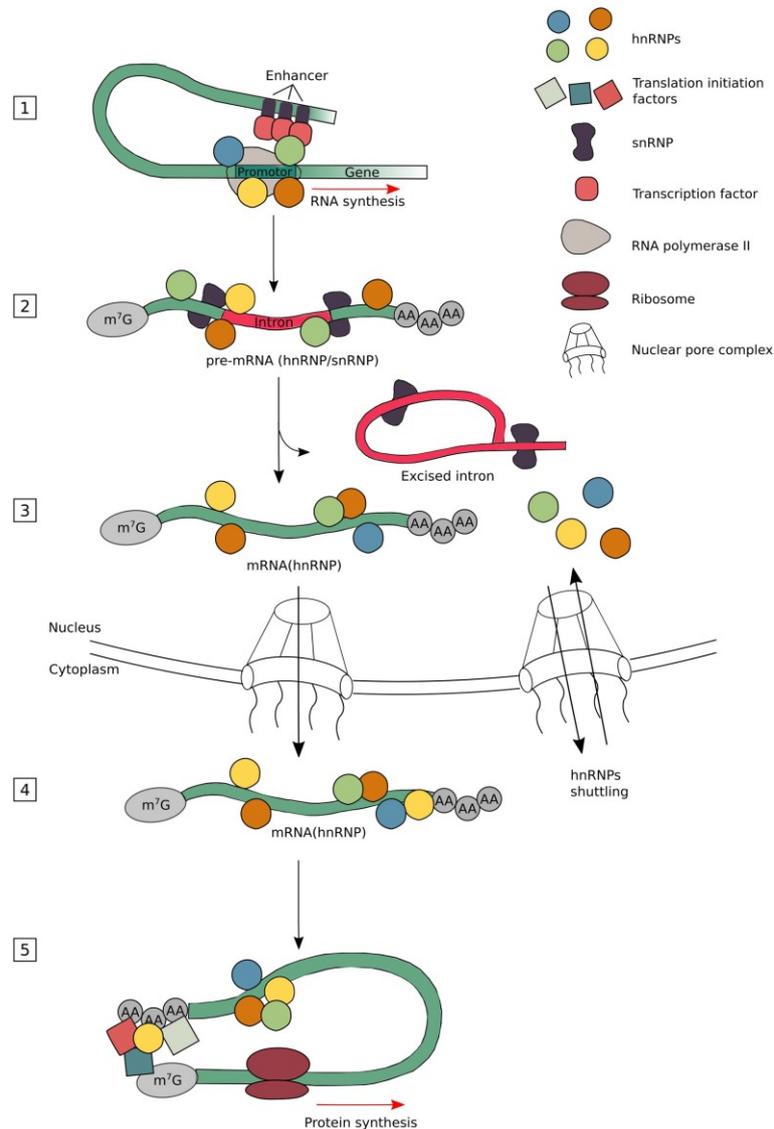
Figure 1.5: Functions of hnRNPs. In the nucleus, hnRNPs are involved in the regulation of transcription by binding to promoter and enhancer sequences (1). The nascent transcript is bound by hnRNPs and stabilized (2). Furthermore, pre-mRNAs are spliced which involves hnRNPs (3) and then exported to the cytoplasm (4), where hnRNPs regulate their translation (5). (Geuens et al., 2016).

The role of Rumpelstiltskin in *Drosophila* The RBP Rumpelstiltskin (Rump and also known as Hrp59 in *Chironomus*), studied in this thesis, is a Drosophila homolog of mammalian hnRNP M and contains three highly conserved RNA recognition motifs (RRM) (King et al., 2014; Jain and Gavis, 2008). Similar to its homolog, the insect version is abundant in the nucleus and affects the splicing of various mRNAs including its own (Kiesler et al., 2005). Several more functions were discovered: Hrp59 has been shown to bind

co-transcriptionally to pre-mRNAs and accompanies the transcripts to the nuclear envelope for nuclear export (Hase et al., 2006; Kiesler et al., 2005). Jain and Gavis (2008) uncovered that, Rump is also involved in mRNA transport and localisation of the mRNAs nanos and oscar, which are distributed to the posterior pole of the developing embryo, thereby regulating anterior-posterior axis patterning. The same study also showed, that homozygous Rump mutant females have reduced number of progeny that complete embryonic development, which could be rescued by a singly copy of a genomic transgene. Similarly, homozygous Rump mutant males are semi-sterile, with only 10% of males being able to fertilize eggs (Jain and Gavis, 2008). Finally, Rump has a role in mediating chromatin function by antagonizing gypsy chromatin insulators in tissues outside the CNS. This results in the expression of downstream genes (King et al., 2014).

1.3.2 The Fragile X mental retardation protein

The second RBP studied in this thesis is Fragile X mental retardation protein (FMRP in mammals, Fmr1 in *Drosophila*). The most common form of inherited mental impairment and autism is Fragile X syndrome (FXS), which is due to the absence or reduction of FMRP (Pieretti et al., 1991). FXS is caused by a CGG trinucleotide repeat expansion in the 5'-untranslated region of the Fmr1 gene which, in case of a full mutation (>230 repeats), leads to hypermethylation of the promoter region and repression of transcription (Fu et al., 1991; Jin and Warren, 2000). Alleles with a repeat number between 60 and 230 are found in premutation carriers, who have normal transcript and protein levels. However, the premutations are extremely unstable and expansion into full mutation can occur by maternal transmission, dependent on the length of the maternal premutation. Therefore, FXS is transmitted as an X-linked dominant trait with a prevalence of 1:7.143 in males and 1:11.111 in females (Hunter et al., 2014). Adult-onset of neurodegenerative disorders, such as Fragile X-associated tremor/ataxia syndrome (FXTAS) and a form of ovarian dysfunction known as Fragile X-associated primary ovarian insufficiency (FXPOI) occur sporadically in individuals carrying premutation alleles (Usdin et al., 2014).

FMRP is a translational regulator

FMRP as well as the *Drosophila* homolog Fmr1 harbor three canonical RNA-binding domains (KH1, KH2 and RGG-Box) and is found to shuttle between cytoplasm and the nucleus due to its Nuclear Localization Signal (NLS) and a Nuclear Export Signal (NES) (Bardoni et al., 1997).

FMRP is involved in various steps of RNA metabolism, but its function as a regulator of translation is the most studied and debated. Initial studies reported that FMRP is a repressor of translation, which was supported by complementary research over the last two decades (Li et al., 2001; Tang et al., 2015; Laggerbauer et al., 2001; Sung et al., 2003). Multiple mechanisms have been proposed for translational repression through FMRP, such as for example repression through polyribosome-stalling (Darnell et al., 2011) or association of FMRP with the RNAi silencing complex (RISC) (Muddashetty et al., 2011; Jin et al., 2004). Another mechanism was proposed by Worpenberg et al. (2021), who found that Fmr1 interacts with Ythdf, both readers of the N6-methyladenosine (m^{6}A) RNA modification and bind together to a common subset of modified targets, leading to translational repression, thereby modulating axonal growth. In contrast, a recent study by Greenblatt and Spradling (2018) showed, that Fmr1 is required to promote the translation of large autism-related genes in the female germline. Additional data in mice and human brains showed that some of FMRPs targets escape translational repression (Brown et al., 2001). To make matters more complicated, ribosome profiling in the presence and absence of FMRP in mouse adult neural stem cells displayed both reduced and increased translation (Liu et al., 2018).

While it is hard to reconcile all the observations it is important to understand how FMRP binds RNA, which RNAs are targeted in which tissue, and during which developmental stage. In this thesis we study the Fmr1 target RNAs in the female germline of *Drosophila*. How RNA is bound by FMRP is largely established by its ability to bind to a sequence motif as well as secondary structures of RNA, such as G-quadruplex (Suhl et al., 2014; Anderson et al., 2016; Maurin et al., 2018). In mammals, the majority of motifs bound by FMRP are linked to its role as a translation repressor and it remains to be

investigated whether this is conserved in *Drosophila*.

FMRP is a multifunctional protein

Besides FMRP's implication in the regulation of translation, the scope of genetic functions continues to expand. Novel functions include RNA-editing via the interaction of FMRP with the adenosine deaminase ADAR in Zebrafish and mouse FXS models, as well as pre-mRNA alternative splicing by interaction with RNA-binding protein 14 (RBM14) (Zhou et al., 2017; Shamay-Ramot et al., 2015; Filippini et al., 2017).

In the nucleus, FMRP has been found to bind chromatin through its tandem Tudor (Agenet) domains when mouse embryonic fibroblasts (MEF) are exposed to DNA damage, thereby regulating the DNA damage response (DDR) (Alpatov et al., 2014). More specifically, FMRP is detected in the vicinity of peri-centromeric domains and co-localized with phosphorylated histone H2A.X, which is a marker for DNA damage (Stiff et al., 2004). During meiosis double strand breaks are generated in prophase in order to promote homologous recombination and crossing over. Interestingly, FMRP is loaded onto chromosomes in male spermatocytes and regulates the placement of phosphorylated histone H2A.X. Moreover, H3K79 methylation, a modification found on sex chromosomes and the centromere in mouse spermatocytes (Ontoso et al., 2014), is required for the recruitment of FMRP to chromatin (Alpatov et al., 2014). Defects during spermatogenesis have been found in the absence of FMRP in *Drosophila* (Zhang et al., 2004), mice (Slegtenhorst-Eegdeman et al., 1998) and human (Johannisson et al., 1987; O'Donnell and Warren, 2002), attributing important germline functions to FMRP.

1.4 The germline of *Drosophila melanogaster*

During metazoan evolution, the germline plays an essential role due to its important function of transmitting genetic and epigenetic information through generations (Johnson et al., 2011). The germline comprises highly specialized haploid gametes, which fuse during sexual reproduction giving rise to a diploid zygote. Gametes are derived from germ cells which initially divide mitotically followed by a meiotic cycle to guarantee constant chromosome numbers in the next generation. The germline of *Drosophila melanogaster* is a well-studied model system and presents a unique opportunity to study novel functions of the RBPs Fmr1 and Rump, both of which impact successful reproduction.

1.4.1 Oogenesis and early development

Female flies have two ovaries, that consist of approximately 18 ovarioles each, in which the oocytes develop. The germarium is situated at the anterior end of the ovariole, containing somatic and germline stem cells. Oocyte maturation occurs along the ovariole in egg chambers containing nurse cells that nurture the growing oocyte. The mature oocyte is found at the posterior end of the ovariole and is competent for fertilization (Bastock and St Johnston, 2008). After fertilisation and pronuclear apposition the zygote is totipotent and early development begins. The first rounds of nuclear divisions are rapid and synchronized, lacking gap phases and the nuclei share a common cytoplasm forming syncytial blastoderm (Farrell and O'Farrell, 2014). Interestingly, the initial divisions occur in the absence of *de novo* transcription from the zygotic genome and instead the early embryo relies on maternal transcripts and proteins, that were deposited into the oocyte prior fertilisation (De Renzis et al., 2007). Zygotic transcription starts gradually at the onset of the maternal-to-zygotic transition (MZT) around cycle 14 (Fig. 1.6) (Pritchard and Schubiger, 1996; Hamm and Harrison, 2018). Maternal transcripts are degraded afterwards (Tadros and Lipshitz, 2009). Post-transcriptional regulation of maternal transcripts is regulated by RBPs, that are involved in mRNA localisation, translation and degradation (Johnstone and Lasko, 2001). Mutations in genes that are required for maternal deposition lead to maternal-effect defects in the early embryo leading to reduced viability (Schupbach and Wi-

eschaus, 1986).
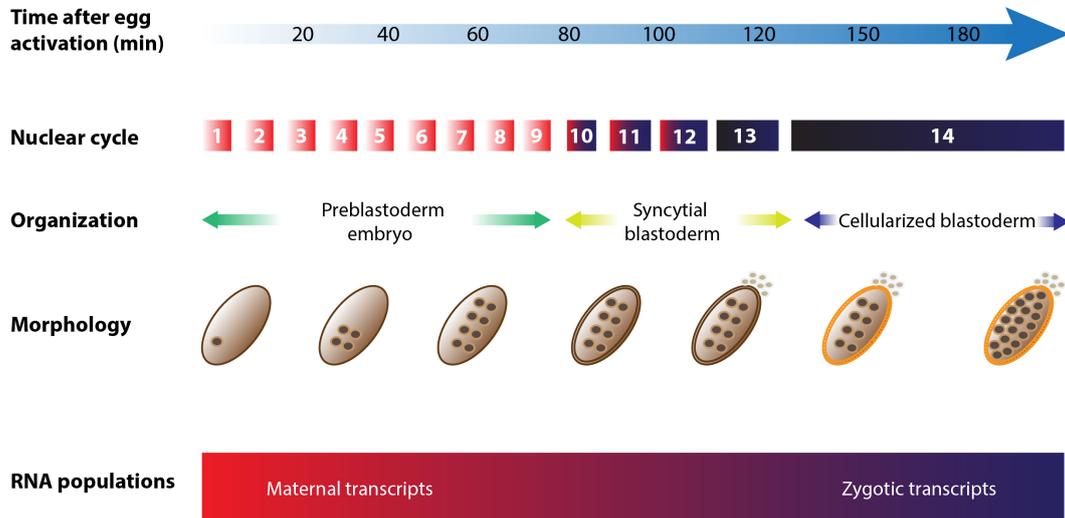
Figure 1.6: Early embryonic development and transcriptomic changes. The schematic depicts the timescale of nuclear divisions and shows the morphology and organization of the developing embryo. The zygotic genome remains silent until zygotic genome activation (ZGA) which begins in late preblastoderm embryos and gradually increases. Until ZGA, the embryo uses maternally-deposited transcripts and proteins to coordinate development and growth. Adapted from (Lefebvre and Lécuyer, 2018).

1.4.2 Spermatogenesis and inheritance of CENP-A

The reproductive organ in male flies consists of a pair of tubular coiled testes. The stem cell niche is located at the apical tip of the testis, where assymetric cell division leads to the production of a cell that maintains stem cell features and the gonialblast, a cell that initiates differentiation. The gonialblast undergoes mitotic divisions, after which cells differentitate into spermatocytes. These cells grow in volume and chromatin organization changes with the separation of the chromosomes and the nucleolus into distinct compartments within the nucleus. The spermatocytes undergo two meiotic divisions, resulting in haploid spermatids (Demarco et al., 2014). Post meiosis, haploid spermatids undergo differentiation stages of sperm development termed spermiogenesis, to become mature sperm that are released into the lumen of seminal vesicles for storage. The process of spermiogenesis is accompanied by three remarkable morphological changes: (i) spermatids change their round shape to a thin and elongated state, (ii) which is accompanied by dramatic chromatin condensation and (iii) finally undergo individualisation to form the needle shaped nuclei

of mature sperm (Fabian and Brill, 2012). Moreover, chromatin of spermatids undergoes a striking reorganization, by which almost all histones are replaced by small basic proteins called protamines (Fig. 1.7). Protamines are evolutionary related to histone H1, but have a higher arginine content, which gives improved condensation properties due to a greater flexibility in the formation of hydrogen bonds with the DNA backbone (Kimmins and Sassone-Corsi, 2005). Histone-to-protamine transition is conserved between mammals and *Drosophila* and the replacement requires transition proteins, such as Tpl^{94D} and is likely facilitated by histone variants, post-translational histone modifications, chromatin-remodeling complexes, as well as transient DNA strand breaks (Rathke et al., 2014).
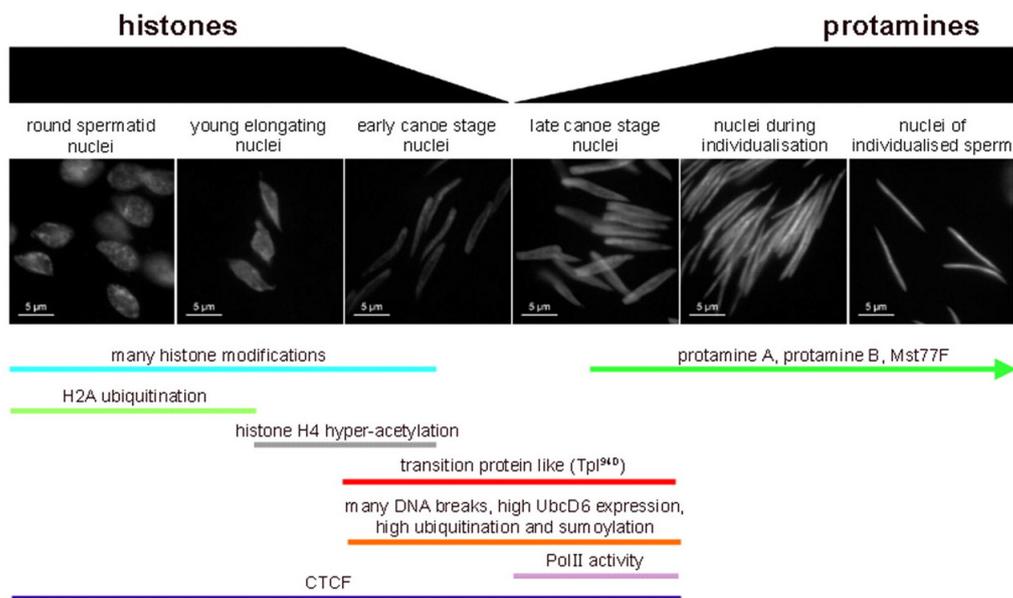
Figure 1.7: Chromatin remodelling during spermiogenesis. Spermatid nuclei undergo morphological changes which are accompanied by chromatin condensation. When spermiogenesis is complete, mature sperm have a characteristic needle-like shape. A key event during sperm maturation is the histone-to-protamine exchange. This process requires histone variants and histone modifications, such as H4 hyper-acetylation and H2A ubiquitination and transition-like proteins. CTCF and active RNA Polymerase II are present during spermiogenesis (Rathke et al., 2007).

In addition, the zinc finger protein CTCF is associated with chromatin post meiosis (Rathke et al., 2007). CTCF is a multifunctional protein that can either activate transcription or act as an insulator protein that can block the actions of cis-acting elements, such as enhancers, and prevent gene activation (Kim et al., 2015). Transcription is very active in early spermatogenesis, but it

has been confirmed that transcription also occurs at later stages (Witt et al., 2019). Therefore, Rathke et al. (2007) suggested that CTCF might be required for chromatin accessibility to RNA Polymerase II, or for the insulation of active genes from inactive ones.

Although CENP-A is a histone protein, it escapes the histone-to-protamine exchange in *Drosophila melanogaster*, *Xenopus laevis* and mammals. Indeed, CENP-A is present throughout spermiogenesis and marks the centromere of mature sperm in these species (Raychaudhuri et al., 2012; Palmer et al., 1990; Milks et al., 2009). Furthermore, CENP-A is present during the formation of the male pronucleus after fertilization and it was shown to be essential for centromere specification and maintenance in the next generation. Fertilization of normal oocytes with sperm lacking CENP-A, leads to the loss of paternal chromosomes during the first mitotic cycles in the early embryo. As a result, embryos are haploid and embryogenesis is halted in late embryogenesis (Edgar et al., 1986; Raychaudhuri et al., 2012). Remarkably, neither CENP-C nor Cal1 are present in mature sperm, suggesting that other factors may be at play to maintain CENP-A at centromeres (Dunleavy et al., 2012). However, the exact mechanism remains elusive and by studying the RBPs Rump and Fmr1 in the germline of *Drosophila melanogaster*, this thesis tries to shed light on this process.

1.5 Aim of the thesis

Centromeres are epigenetically defined by the incorporation of CENP-A into centromeric chromatin. Besides the CENP-A loading factor and important kinetochore proteins, centromere specification and function is propagated by (peri-)centromeric transcripts and the act of transcription itself. In the germline, centromere maintenance is particularly sensible, as functional centromeres are required to guarantee successful reproduction and early embryonic development. However, which factors are at play to regulate centromeric transcripts remains elusive. RNA-binding proteins (RBPs) are key players in the regulation of RNAs and involved in every step of RNA metabolism. Therefore, the aim of this doctoral thesis was to study RBPs that are relevant for centromere function and present in the germline of *Drosophila melanogaster*. I have addressed the two aspects:

1. Identification and characterization of RBPs that interact with centromeric RNAs.

2. Analyze the functional relevance of the RBPs Fmr1 and Rump in the germline of *Drosophila melanogaster*.

Chapter 2

Results

2.1 The RNA-binding protein Fmr1 regulates transcripts important for centromere identity

The first time centromeric transcripts were reported was in the late 1960s but their functionality has only recently been appreciated (Rieder, 1979; Chan et al., 2012; Bulut-Karslioglu et al., 2012). Centromeric RNAs have a critical role in centromere structure and function. However, it is not yet fully understood how centromeric transcripts are regulated. Therefore, I set out to identify RNA-binding proteins from previously published mass spectrometry data that were obtained with a crosslinked CENP-A pulldown approach (Demirdizen et al., 2019). To this end, I performed a GO-term analysis to identify RNA-binding proteins that are enriched at the centromere. Thereby, I obtained a list of enriched RNA-binding proteins (Table A.1) with one of the top enriched factors being Fragile X mental retardation 1 (Fmr1), a well known RNA-binding protein that has been widely studied in human, mouse and *Drosophila melanogaster*. Interestingly, Fmr1 was also present in mass spectrometry data acquired by a SatIII RNA pulldown performed by Saskia Höcker (unpublished data). Therefore, I started with the validation of Fmr1 binding to SatIII RNA.

2.1.1   Fmr1 binds to SatIII RNA *in vitro*

In order to study the binding of Fmr1 protein to SatIII RNA, an electrophoretic
mobility shift assay (EMSA) was performed. Fmr1 full-length protein with a
N-terminally GST-tag was recombinant expressed and purified from *E.coli*.
SatIII Sense RNA, consisting of two 359 bp repeats was *in vitro* transcribed
and fluorescently labelled by the incorporation of 488-UTP nucleotides. Tubu-
lin mRNA was used as RNA control and GST-only as a protein control. In-
creasing concentrations of recombinant purified Fmr1 protein were incubated
with either of the RNAs.
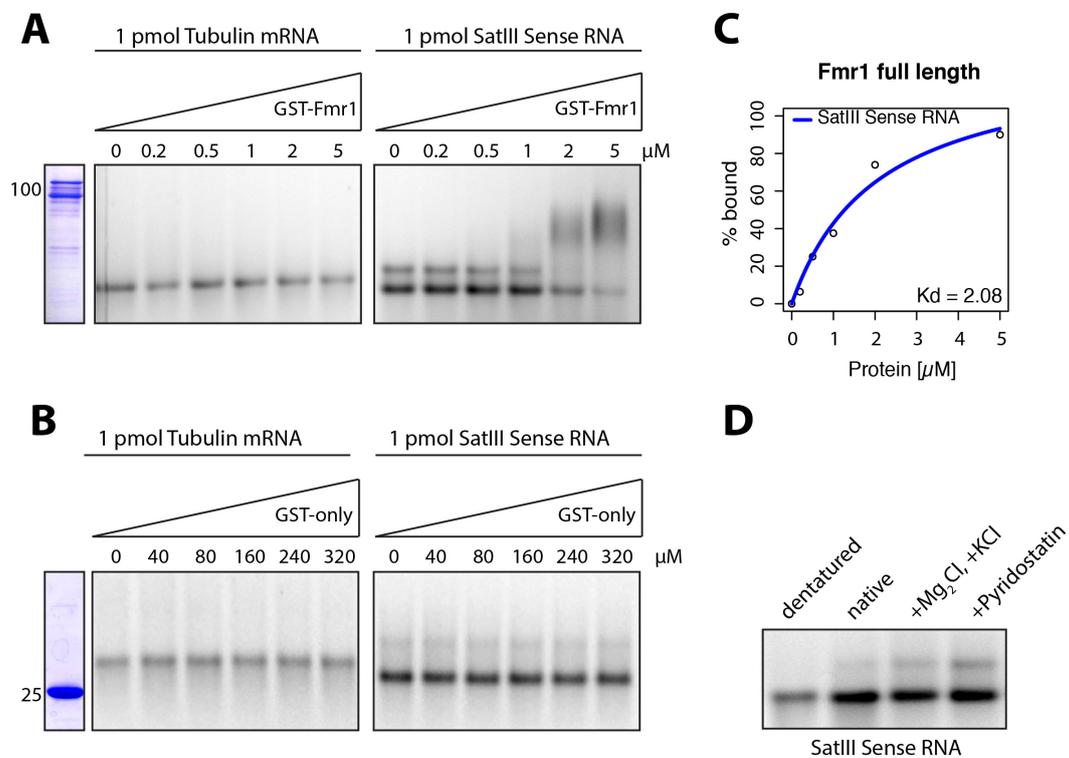
Figure 2.1: Validation of Fmr1 interaction with SatIII Sense RNA. A) GST-Fmr1 was purified
and used for EMSA. GST-Fmr1 was detected at the expected size of 100 kDa. EMSA was performed using
1 pmol RNA and a titration of GST-tagged protein as indicated. A shift was observed with SatIII Sense
RNA but not with Tubulin mRNA. B) GST-only tag was purified and used as negative control in the
EMSA. The protein was detected at the expected size of 25 kDa and did not result in a shift with any of
the RNAs used. C) To determine the fraction of bound RNA by GST-Fmr1 titration, the data were plotted
using non-linear regression. The Kd equals 2.08 μM. D) SatIII Sense RNA was tested for its potential *in
vitro* secondary structure formation under denaturing conditions and native conditions that are in favour
for secondary structure formation. Size of lower band: 719 bp. Upper band 1 kb.

A shift was observed solely with SatIII Sense RNA, thus confirming the inter-
action between SatIII RNA and Fmr1 protein. The binding was also specific to

Fmr1, as GST-only protein did not result in a shift either (Fig. 2.1.A, B). The percentage of bound RNA to protein in μM was plotted by nonlinear regression. The dissociation constant equals the concentration of free RNA at which half of the total molecules of RNA are associated with protein (Fig. 2.1.C). Fmr1 was enriched in both SatIII Sense and Antisense RNA mass spectrometry data from Saskia Höcker and therefore it remains to be tested whether Fmr1 binds to SatIII Antisense RNA.

Interestingly, SatIII Sense RNA seems to have a secondary structure *in vitro*, as two distinct bands can be observed on the Agarose gel. The second upper band that represents a putative secondary structure disappeared when loaded under denaturing conditions and was stabilized in conditions favouring a RNA secondary structure, such as with buffers containing either Mg^{2+}, K^+ ions, or Pyridostatin, a G-quadruplex stabilizing drug (Fig. 2.1.D). This indicates that SatIII is folded *in vitro*.
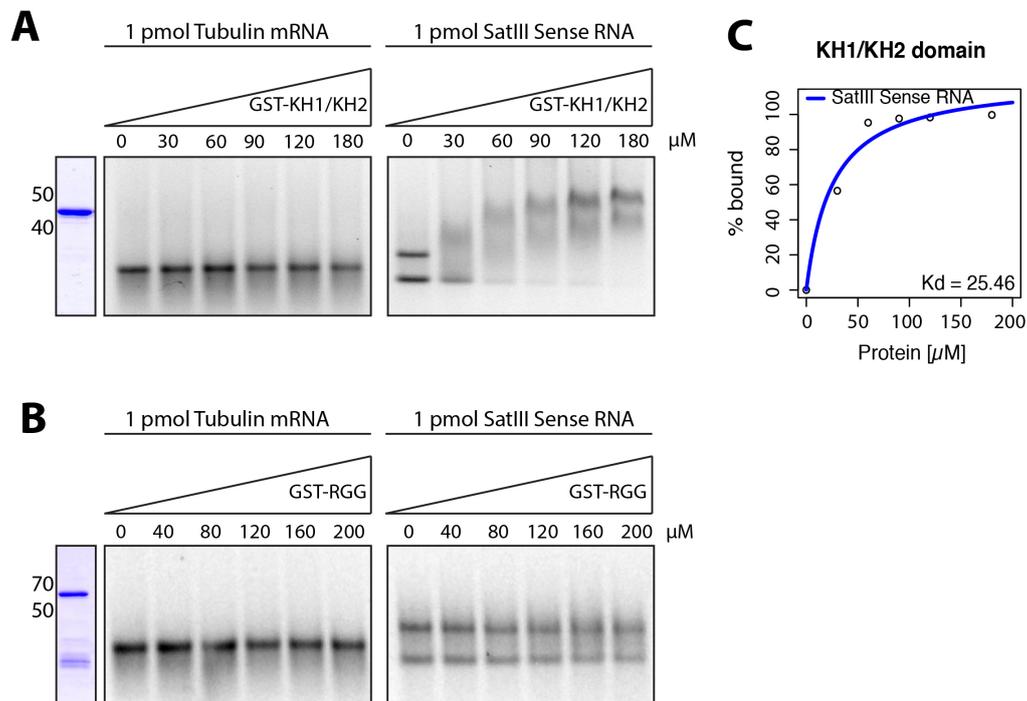
Figure 2.2: Fmr1 binds to SatIII RNA via the KH domain. A) The RNA binding domains KH1/2 were recombinant purified and used for EMSA. GST-KH1/2 was detected at the expected size of 49 kDa by Coomassie stain and was found to specifically bind to SatIII Sense RNA. B) The third RNA binding domain of Fmr1 (RGG domain) was also studied for RNA binding. The purified protein had the expected size of 57 kDa on Coomassie and there was no shift detectable. C) Non-linear regression was used to determine the Kd which equals 25.46 μM.

Fmr1 has three known RNA-binding domains. Two centrally located KH domains and one RGG domain (Fig. 2.3.A) (Siomi et al., 1993). In order to determine, which of the RNA-binding domains is responsible for SatIII binding, the KH domains and the RGG domain were purified and assayed for binding by EMSA as described above. The result shows that SatIII binding is mediated by the KH domain but not the RGG domain (Fig. 2.2.A, B). The KH domain was purified as a trimer (seen by size exclusion chromatography), which might explain the lower affinity for SatIII compared to the full-length protein (Fig. 2.2.C)

2.1.2 Absence of Fmr1 in the germline results in low fertility

To study Fmr1 function *in vivo*, I obtained a Fmr1 mutant fly line that was generated by Lee et al. (2003). The *Fmr1⁴* allele carries a point mutation that changes amino acid 289 to a stop codon (Fig. 2.3.A). *Fmr1⁴* flies are null for Fmr1 and the flies were verified by Western blot using whole fly extract from wt *w¹¹¹⁸* and *Fmr1⁴* flies. No band could be observed for Fmr1 in flies with the *Fmr1⁴* mutation (Fig. 2.3.B). Furthermore, I assessed the fertility of the Fmr1 mutant by crossing female mutants with wt male flies and vice versa: 25 homozygous flies were crossed with 25 wt *w¹¹¹⁸* flies, embryos were collected after 12 h and the hatching rate was monitored after 36 h. The control cross with wt flies resulted in a high viabilty, with over 90% of hatched embryos. Only 6% of offspring obtained from female *Fmr1⁴* flies hatch, suggesting a very low female fertility. Male *Fmr1⁴* flies were shown to be completely sterile (Fig. 2.3.C). Reduced fecundity in Fmr1 mutant flies has been previously shown (Zhang et al., 2004; Greenblatt and Spradling, 2018; Specchia et al., 2017). Male gonads of Fmr1 mutant flies are known to be enlarged and spermatogenesis is arrested specifically in late-stage spermatid differentiation following individualisation. Furthermore, Fmr1 mutants specifically loose the central pair of microtubules in the sperm tail axoneme (Zhang et al., 2004) . The ovaries of *Fmr1⁴* mutant females are smaller in size (Fig. 2.3.D). Ovaries have aberrant egg chambers with fewer or supernumerary germ cells, likely due to both under and over-proliferation scenarios (Epstein et al., 2009).

This study tries to shed light on centromeric transcript regulation by Fmr1, which might add an additional factor to the reduced fertility of *Fmr1⁴* mutant flies.
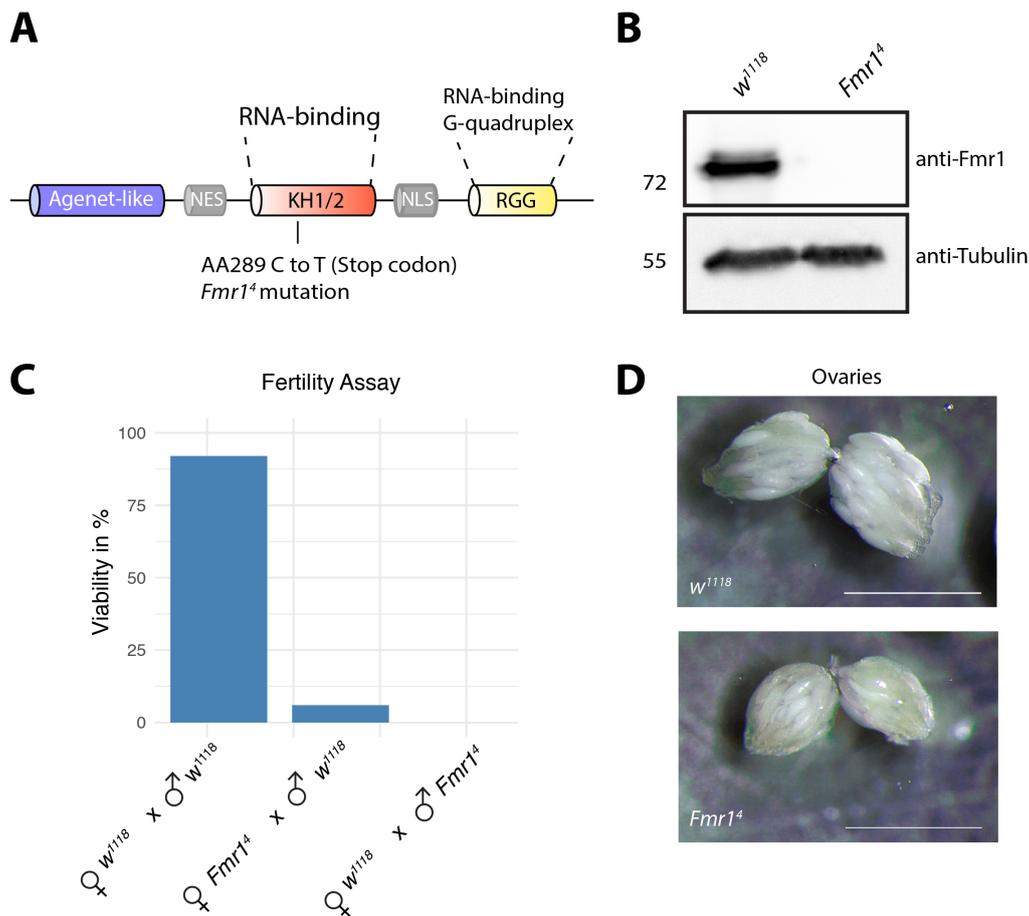
Figure 2.3: Fmr1 mutation causes reduced female fertility and male sterility. A) Schematic depicting the domains of the Fmr1 protein. A point mutation in the KH1 domain results in a nonsense mutation (*Fmr1⁴*). Adapted from (Lee et al., 2003). B) Western blot with wt w^{1118} and *Fmr1⁴* fly extract using Fmr1 and Tubulin antibodies. C) A fertility assay was performed by crossing male flies of *Fmr1⁴* mutant to wt w^{1118} female flies and vice versa. As a control wt flies were used. The embryos that were laid were collected and hatched embryos were counted. D) Bright-field images of adult fly ovaries. Female flies were aged for 3 days before dissection. The ovaries of *Fmr1⁴* mutant flies are smaller. Scale bar 100 μm.

2.1.3 Fmr1 binds to CENP-A containing chromatin

Next, I wanted to verify that Fmr1 is a potential interaction partner of CENP-A as indicated by the mass spectrometry analysis. Therefore, I performed a mononucleosome IP. In this experiment, CENP-A-containing chromatin from

S2 cells was extracted and fragmented by Micrococcal Nuclease (MNase) digestion. Thereafter, recombinant purified proteins bound to Glutathione-Sepharose beads were incubated with mononucleosomes and analysed for interaction on Western blot (Fig. 2.4.A). The preparation of mononucleosomes was monitored by taking a sample in a time series from 0-30 min (Fig. 2.4.B). DNA was isolated and tested for its length by Agarose gel electrophoresis.
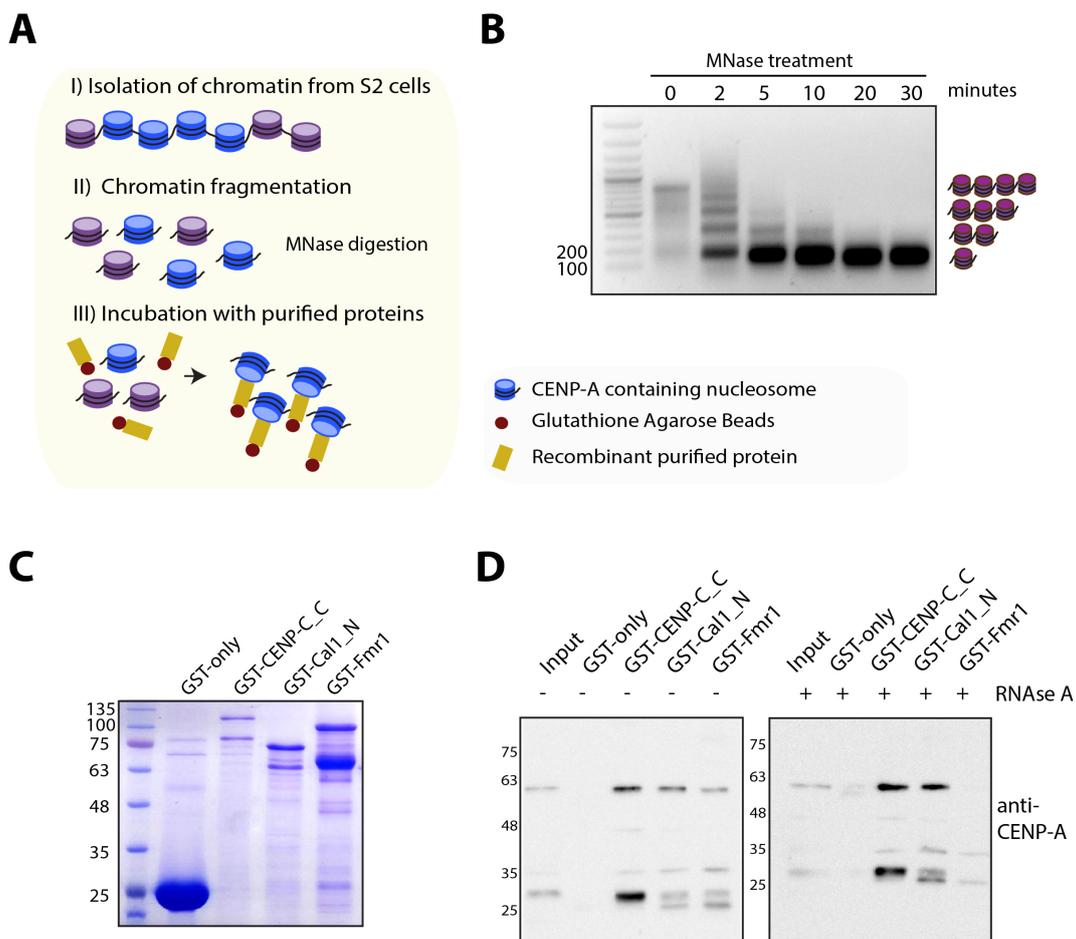
Figure 2.4: **Fmr1 interacts with CENP-A mononucleosomes** *in vitro.* A) Schematic of mononucleosome IP. Mononucleosomes were prepared from S2 cells by fragmenting chromatin with micrococcal nuclease. Proteins of interest were recombinant expressed in bacteria and purified leaving the GST-tag on the protein. Proteins were then incubated with mononucleosomes and analysed for interaction. B) Chromatin from S2 cells was digested at 37 °C and analysed on agarose gel after different time points. Mononucleosomes were obtained after 30 min. C) Indicated GST-tagged proteins were purified and analysed on SDS-PAGE followed by Coomassie staining. Proteins were detected at expected sizes. D) Mononucleosomes were prepared for IP and prior pulldown mononucleosomes were either left untreated (left) or treated with 40 μg of RNAse A (right). Treatment with RNase seems to disrupt the interaction of Fmr1 with ubiquitylated CENP-A (band at 55kDa) as well as endogenous CENP-A (band at 30 kDa). Both positive controls GST-Cenp-C and GST-Cal1 were unaffected by the RNase A treatment.

Around 147 bp of DNA are wrapped around one nucleosome core unit. Hence, mononucleosomes were obtained after 30 min at 37 °C. Di-, tri-, and tetrameric nucleosomes were observed with shorter incubation times. For the IP, GST-tagged proteins were purified from *E.coli*. GST-only was used as a negative control. As positive controls, I purified the C-terminal region of CENP-C and the N-terminal region of Cal1, both known interacting proteins of centromeric chromatin (Fig. 2.4.C) (Erhardt et al., 2008). CENP-C is an inner kineto-chore protein that is strongly associated with CENP-A chromatin and Cal1 is the loading factor of CENP-A (Mellone et al., 2011; Chen et al., 2014). The input of the IP consisted of mononucleosomes. The IP was immunoblotted and analysed using anti-CENP-A antibody. Fig. 2.4.D (left) shows that both positive controls and GST-Fmr1 were able to pull down CENP-A. The band at 55 kDa represents ubiquitylated CENP-A, the band at 30 kDa shows endogeneous CENP-A protein (Bade et al., 2014). The band at 35 kDa is likely unspecific, but needs to be confirmed by a knockdown experiment of CENP-A in the future. Interestingly a second band appeared below the 30 kDa band in the IP's with GST-Cal1_N and GST-Fmr1. It may be interesting to find out if this is due to a phosphoshift. In order to investigate if the interaction of Fmr1 with CENP-A nucleosomes is RNA-dependent, I incubated mononucleosomes with RNaseA for 1h at RT prior to the pulldown (Fig. 2.4.D right). Interestingly, treatment of mononucleosomes with RNaseA seems to be sufficient to abolish the signal of ubiquitylated CENP-A as well as the known CENP-A signal at 30 kDa. This indicates, that Fmr1 interacts with CENP-A chromatin in an RNA-dependent manner.

2.1.4 Fmr1 targets cenRNAs and has regulatory effects on transcripts

Since Fmr1 likely binds to centromeric chromatin in an RNA-dependent manner and is able to bind SatIII RNA *in vitro*, we hypothesized that Fmr1 may be able to bind to cenRNAs. In order to determine whether this is the case, I decided to perform a Fmr1 CLIP experiment followed by RNA-Seq from *Drosophila* ovaries. For that, I dissected 50 ovaries per IP in PBS containing RNase inhibitor. Ovaries were subsequently UV-crosslinked with 0.6 J/cm^2 which resulted in a slight upward shift of Fmr1 due to the crosslink and was

reversible with RNaseA (Fig. 2.5.A). IgG antibody with the same isotype than the Fmr1 antibody was used to control for unspecific binding. RNA fragmentation was performed with sonication and RNAse I and fragments were assayed for their length on an RNA gel. The length of the majority of RNA fragments was below 1000 bp (Fig. 2.5.B). The success of the CLIP was confirmed by Western blot and RNA was extracted from the IP's by cutting out the membrane as indicated in Fig. 2.6.A. The isolated RNA was used for library preparation and deep sequencing. The CLIP RNA-Seq data were analysed in order to find new target RNAs for Fmr1.
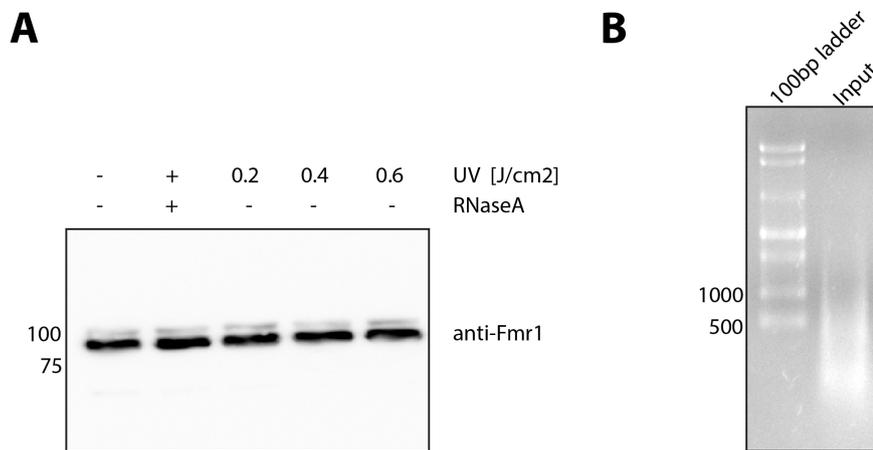
Figure 2.5: Establishment of the CLIP experiment in ovaries. A) UV crosslinking efficiency was tested by exposing ovaries to different energy levels. Subsequently protein was extracted and loaded onto SDS-page followed by Western blot. A slight shift of Fmr1 protein was detected after crosslinking with 0.6 J/cm^2. The shift was reversable with RNaseA. B) RNA gel showing the efficiency of RNA fragmentation after sonication and RnaseI treatment for 3 min at 37 °C. RnaseI was applied in 1:300 dilution. The size of RNA fragments was between 100-750bp.

To this end, I analysed whether Fmr1 is associated with transcripts from centromeric and pericentromeric regions. Therefore, all samples were aligned to the most recent genome assembly that includes all complex satellites and CENP-A enriched sequences in *Drosophila melanogaster* (Chang et al., 2019). The data of two replicates from Fmr1 and IgG CLIP were used for analysis as they showed enough variance between CLIP sample and input (Fig. A.1.A). After mapping CLIP and input reads to the novel genome assembly, I summarized read counts for each type of complex repeats (*e.g.* TEs and complex satellite repeats) using an R script. The Fmr1/input ratio was normalized by the total number of mapped reads. Interestingly, 359 bp satellite repeats

were enriched in the Fmr1 CLIP confirming the EMSA results from Fig. 2.1 and Fig. 2.2 *in vivo*. Furthermore, an enrichment of Fmr1 on dodeca satellite transcripts was observed. Dodeca satellites are 12 bp tandem repeats that are present at the pericentromere of chromosome 3 (Abad et al., 1992). Transcripts of transposable elements that were enriched in the Fmr1 CLIP over Input include for example DOC2 and G which are non-LTR retroelements in the Jockey family especially abundant within centromere islands of chromsomes X, Y and 4. TART elements and DM1731 are found within CENP-A chromatin but outside of centromere islands and are enriched in the Fmr1 CLIP (Chang et al., 2019). All RNAs that are known to originate from (peri-) centromeric chromatin and are enriched in the Fmr1 CLIP are shown in the barplot in Fig. 2.6.B (Table A.8) .
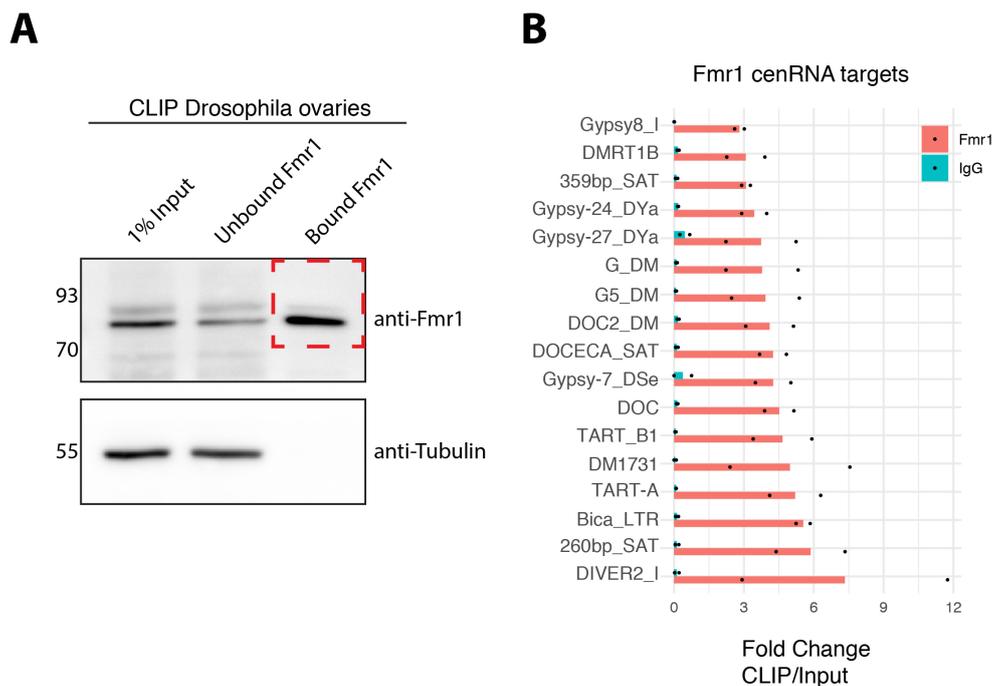
Figure 2.6: Fmr1 is enriched at RNAs with centromeric origin. A) Cross-linking immunoprecipitation of Fmr1 followed by RNA-Seq analysis was performed using fly ovaries. The IP was confirmed by immunoblotting with Fmr1 antibody. Tubulin was not present in the pulldown of Fmr1. B) Sequencing data were aligned to the most recent *Drosophila* genome assembly including all complex satellite repeats and CENP-A enriched sequences in *Drosophila melanogaster* (Chang et al., 2019). The barplot shows the enrichment of Fmr1 compared to the input on centromeric RNAs (cenRNAs red bars) from two replicates. IgG CLIP was used as negative control (turquoise bars).

Since Fmr1 binds to centromeric RNAs, we hypothesized that it might also have a regulatory effect on their expression levels. To test this, I sequenced total RNA obtained from wt and $Fmr1^4$ mutant testes. Three independent RNA isolations were conducted for each genotype and used for library preparation and sequencing. The three replicates clustered together and did show differences to the wt control (Fig. A.3.B,C). Differential expression changes in transposable elements and complex satellite repeats were analysed using the Deseq2 package in R (Love et al., 2014). The count data were transformed to log2 scale using the rlog function which normalizes the counts in respect to the library size (Table A.9). The transformed data were visualized in heatmaps (Fig. 2.7).
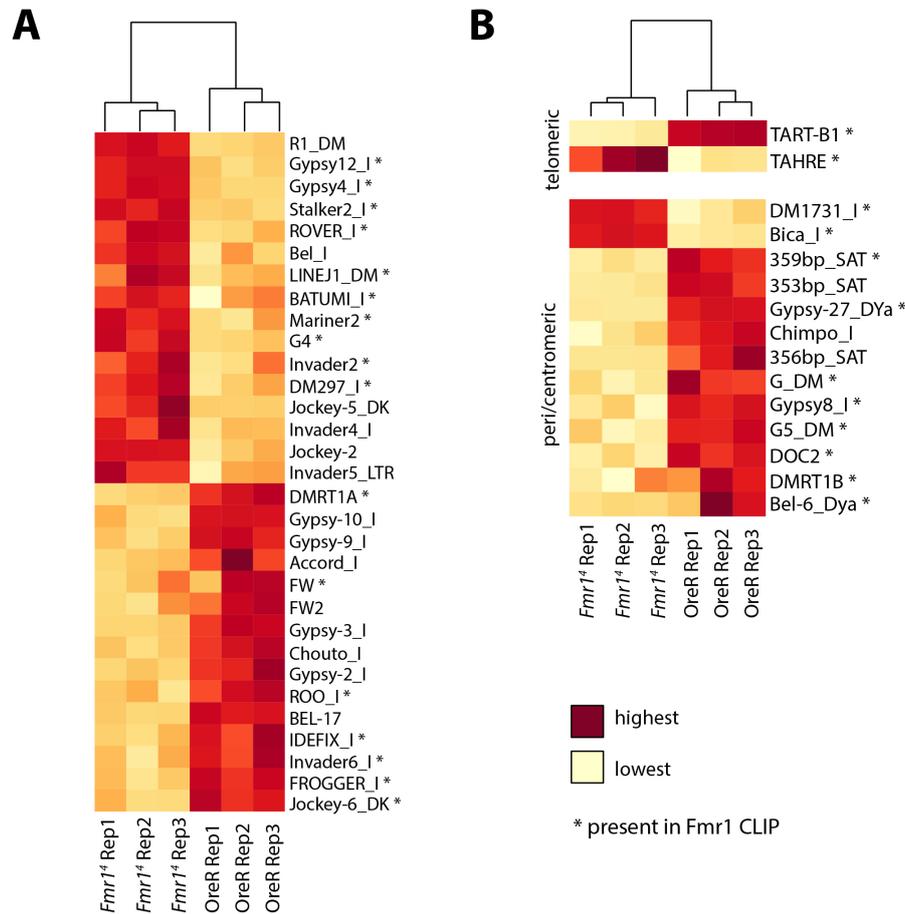
Figure 2.7: Total RNA-Seq data from testes of wt OreR and $Fmr1^4$ mutant flies reveal differential expression changes of transposable elements and complex satellite repeats. A) Heatmap showing significant changes of transposable elements outside of centromeric and telomeric regions. B) Heatmaps displaying the difference in transposable elements and complex satellite repeats that are located at the centromere and telomeres. Transposable elements and repeats that were found enriched in the Fmr1 CLIP are marked with a star.

Only significant hits were plotted and sorted into transposable elements that are present at centromeres, telomeres and outside of these regions. Fig. 2.7.A displays changes in transposable element expression levels that are not present in centromeric islands nor at the telomere. Interestingly, Fmr1 has the potential to upregulate as well as downregulate transposable elements. Most transposable elements that are present at centromeres, however, are downregulated in the *Fmr1⁴* mutant (Fig. 2.7.B), suggesting that Fmr1 is required to activate the expression or is needed for transcript stability. Satellite repeats (353 bp, 356 bp and 359 bp) from the 1.688 satellite family are also misregulated. Most of the differentially expressed transcripts were also found to be enriched in the Fmr1 CLIP, (*e.g.* 359bp satellite repeats, Doc2, DMRT1B) indicating a direct regulatory role on these RNAs by Fmr1.

2.1.5 Fmr1 loss results in mitotic defects and impedes the translation of certain mitotic proteins

Previous work on SatIII RNA from the Erhardt laboratory has shed light on the importance of SatIII transcripts during mitosis. The depletion of SatIII results in severe mitotic defects in S2 cells as well as embryos due to missegregation of all major chromosomes (Rošić et al., 2014). Since SatIII repeats are bound by Fmr1 and the transcript levels are reduced in the germline (as shown above), I decided to look at the *Fmr1⁴* mutant embryos to further investigate centromere regulation. The mutant fly line can only be kept as a heterozygous stock, as *Fmr1⁴* homozygous adults are unable to produce viable fertile offspring. At the embryonic stage, homozygous Fmr1 mutants can be recognized based on their pole cell phenotype, where mutants have a reduced pole cell number and internalised pole cells (Deshpande et al., 2006). I performed immunofluorescence for CENP-A protein in wt *w^1118* and *Fmr1⁴* on 0-4 h old embryos. *Fmr1⁴* mutant embryos show mitotic defects, including lagging chromosomes, chromosome bridges, and chromosome breakage. However, there was no obvious effect on CENP-A signal detectable (Fig. 2.8). The mitotic defect phenotype has been previously described in Deshpande et al. (2006), where they used another Fmr1 mutant (*Fmr1³*) and found that Fmr1 is implicated in correct centrosomal assembly. In this study, however, we focus on the centromeric regulation by Fmr1, which could give new insights to

Fmr1's functionality.
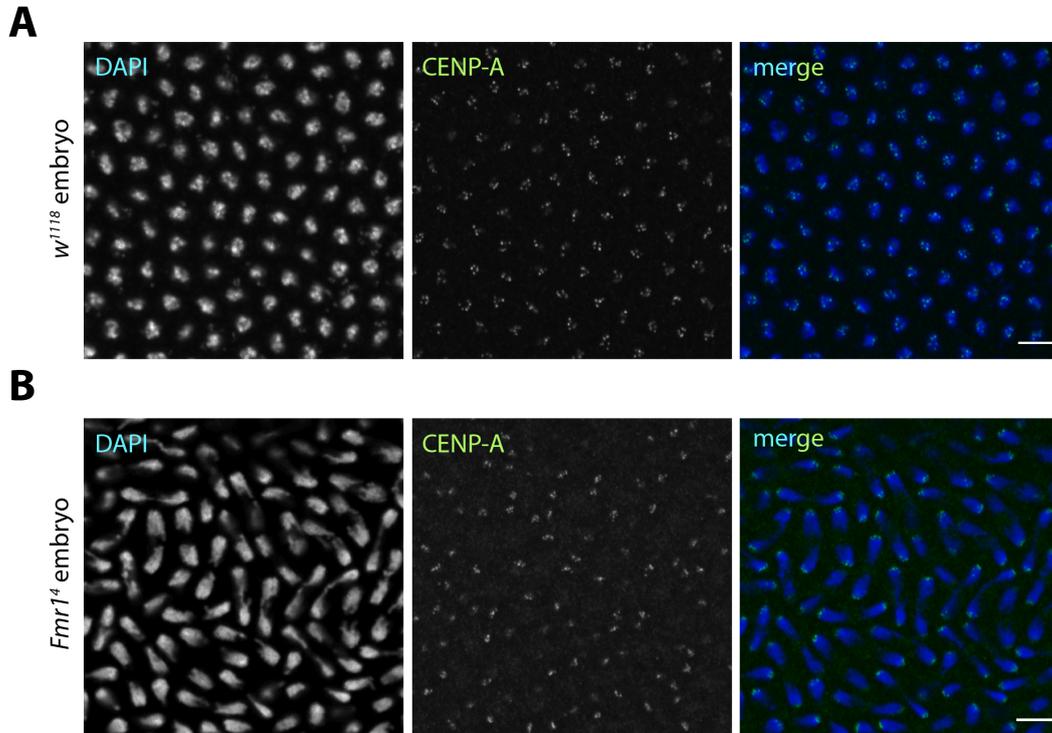
Figure 2.8: Lack of Fmr1 results in mitotic defects in Drosophila embryos. Immunofluorescence for CENP-A was performed on wt w^{1118} (A) and $Fmr1^4$ mutant embryos (B). Severe mitotic defects were ovserved in the absence of Fmr1. DNA was counterstained with DAPI. Scale bar 5 μm.

Since the observed phenotype was detected in early embryos, we hypothesized that proteins and RNAs that are important for mitosis may be maternally deposited, as major zygotic transcription is known to start after cycle 13 in the embryo (Hamm and Harrison, 2018). Interstingly, Greenblatt and Spradling (2018), have found a role of Fmr1 protein in the activation of translation in *Drosophila* ovaries. Therefore, I decided to re-analyse these data in order to look for downregulated proteins that are associated with mitosis and meiosis. The dataset includes data from Ribosome Footprinting and mRNA sequencing using germline Fmr1 RNAi compared to control RNAi. Overall, I found 8% of genes associated to GO terms such as mitosis and meiosis that showed a significant downregulation of translation in Fmr1 depleted ovaries (Table A.2). Proteins important for centromere function are marked in the volcano plot (Fig. 2.9.A). Interestingly, mitotic genes that are translationally reduced by Fmr1 RNAi from Ribosome Footprinting experiments do not show significant changes in mRNA levels (Fig. 2.9.B).
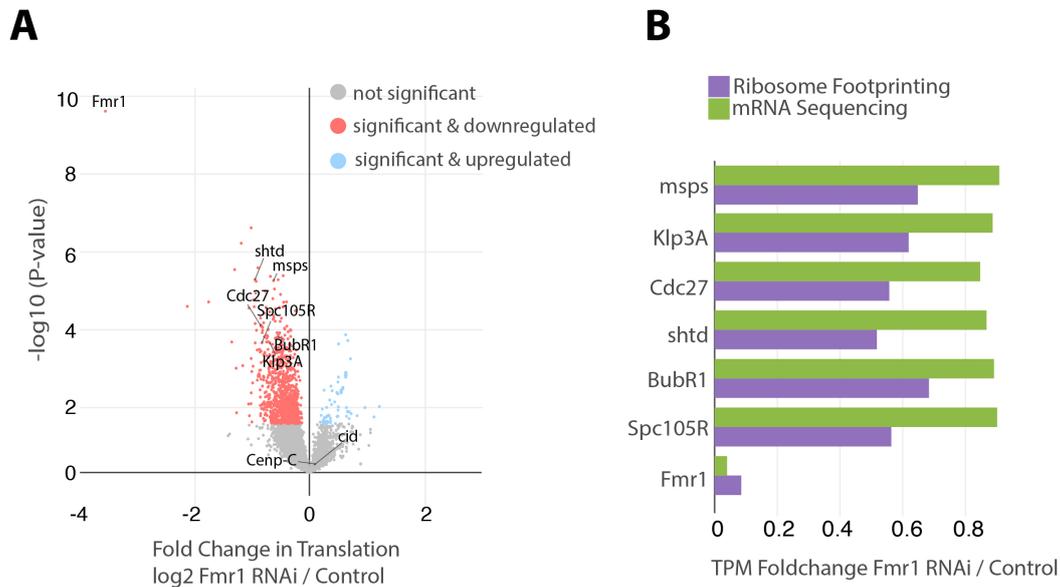
Figure 2.9: Certain mitotic proteins are downregulated in the absence of Fmr1 in ovaries. A) Re-analyses of the Ribosome Profiling dataset from (Greenblatt and Spradling, 2018) shows downregulated proteins in red, upregulated proteins in blue and not significant changes in gray. Certain mitotic hits are annotated. B) Mitotic significant genes that are translationally reduced by Fmr1 RNAi from Ribosome Profiling do not show significant changes at mRNA level. Fold change Fmr1 RNAi to Control was plotted.

Amongst the candidate genes that show a significant downregulation of translation are Spc105R and BubR1. Spc105R encodes an essential component of the kinetochore and is not only important for kinetochore-microtubule attachment, but also required for co-orientation of sister centromeres at meiosis I in oocytes (Radford et al., 2015). BubR1 is required during mitosis for the spindle assembly checkpoint and has been shown to be essential in maintaining sister chromatid cohesion during meiotic progression in both sexes (Malmanche et al., 2007). Other candidates include shtd and cdc27, which are part of the anaphase promoting complex (Tanaka-Matakatsu et al., 2007) and Klp3A, that encodes for a microtubule motor protein involved in chromosome segregation and mitotic spindle morphogenesis (Kwon et al., 2004). Msps is also microtubule-associated and is required for cytoplasmic microtubules in oogenesis (Moon and Hazelrigg, 2004).

I validated the downregulation of Spc105R and BubR1 in the *Fmr1*^{4} mutant ovaries at protein level (Fig. 2.10.A). Furthermore, the total mRNA levels of candidate genes were not significantly changed between *Fmr1*^{4} mutant and control *w*^{1118} ovaries, which is in accordance with the analysis (Fig. 2.10.B). In order to understand the translational regulation of mitotic mRNAs, I per-
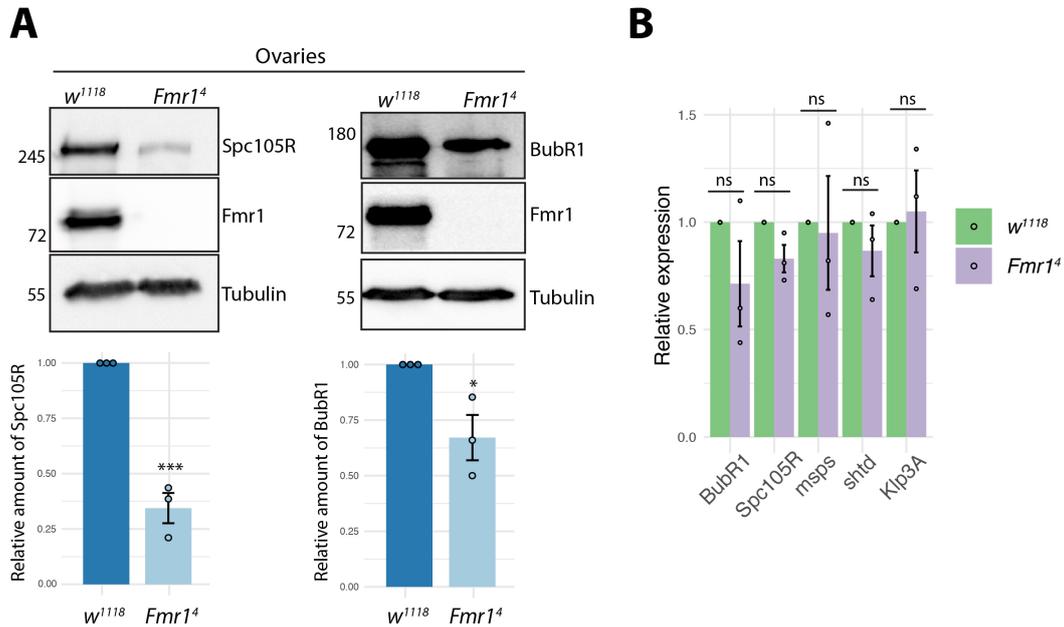
Figure 2.10: The translation of the kinetochore proteins Spc105R and BubR1 is affected in *Fmr1⁴* mutant ovaries. A) Western blot from ovaries of w^{1118} and $Fmr1^4$ for Spc105R, BubR1 and Tubulin as loading control. The band intensity of Spc105R and BubR1 was quantified and normalized to Tubulin. Statistical significance was assessed using unpaired Student's t-test. The error bars represent the standard deviation of pooled data from three independent experiments. Spc105R and BubR1 proteins were significantly reduced in $Fmr1^4$ ovaries (p-value < 0.001 and < 0.05, respectively.) B) The relative expression levels of indicated mRNAs obtained by RT-qPCR from $Fmr1^4$ do not reveal significant changes compared to the wt control.

formed sucrose gradient fractionation using ovary extracts and analysed the distribution profile in each gradient by RT-qPCR. The UV-absorbance profile was overlaid for each genotype. The profiles of the w^{1118} and the $Fmr1^4$ mutant were similar (Fig. 2.11.A), which suggests that the translational downregulation of genes upon Fmr1 depletion does not cause a significant global change in translation. For each fraction that was taken, protein was extracted for Western blot analysis. To confirm the presence of monosomes and polysomes according to the UV-absorbance profile, Rps6 protein was analysed. Rps6 is a ribosomal protein and cosedimented with polyribosomes. All detectable Fmr1 sedimented near the top of the gradient, suggesting that it is not directly bound to polyribosomes in *Drosophila* ovaries (Fig. 2.11.A). A similar result was reported by Monzo et al. (2006), where they could not detect Fmr1 at polyribosomes in *Drosophila* embryos. For each mRNA, data were plotted as the percentage of total amount applied to the gradient detected in each fraction. The data were normalized to Tubulin67C, which is known to remain

translationally active throughout oogenesis (Li et al., 2014). The error bars depict the standard deviation between triplicates during RT-qPCR. When I assessed the mRNA distribution of Spc105R, Bubr1 and cdc27 in the *Fmr1⁴* mutant gradient, I found that compared to the wt gradient there is a reduction in the amount of each mRNAs present in the polysome-containing fractions (missing peak in Fraction 9) (Fig. 2.11.C). There is also a slight shift towards the lighter fractions. Cenp-C mRNA was used as a control mRNA and I could not observe similar differences between wt and *Fmr1⁴* mutant ovaries. However, Cenp-C mRNA was not highly enriched in the polysome fractions, suggesting that it is not highly translated in ovaries.
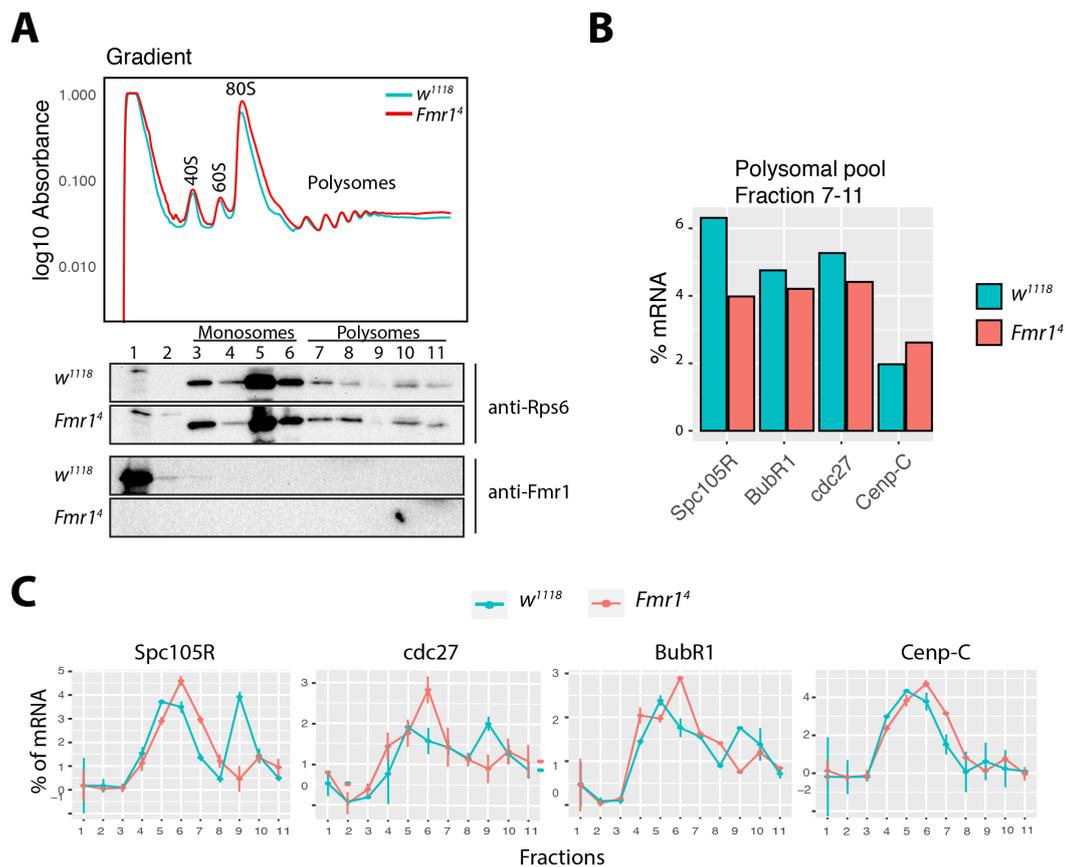
Figure 2.11: Spc105R, cdc27, BubR1 mRNA shows a reduction at polysomes in ovaries absent of Fmr1. A) UV absorbance profile for indicated genotypes obtained by a sucrose gradient from ovary extract. Each fraction was used for RNA and protein extraction. Protein was analysed by Western blot for Rps6 and Fmr1. B) Comparison of the amount indicated mRNAs present in the fast-sedimenting polysome-containing fractions (fractions 7-11 of the polysomal pool). C) Distribution of indicated mRNAs in the wt (turquoise) and Fmr1 mutant (orange) gradients. Plotted was the percentage of total mRNA in each fraction normalized to Tubulin67C. Error bars depict the standard deviation between CT values of triplicates acquired during RT-qPCR. Data are representative of one experiment.

Based on the profile I concluded that the vast majority of polysomes is represented in the fast-sedimenting fractions of the gradient (fractions 7-11). I summed over these fractions and could observe a slight decrease for all mRNAs except for Cenp-C (Fig. 2.11.B).

Next, I wanted to assess whether Fmr1 might bind directly to the downtranslated transcripts thereby promoting translational initiation in association with RNPs or affecting mRNA transport to sites of active initiation. To explore direct binding, I analysed my Fmr1 CLIP dataset for coding genes (Table A.4).
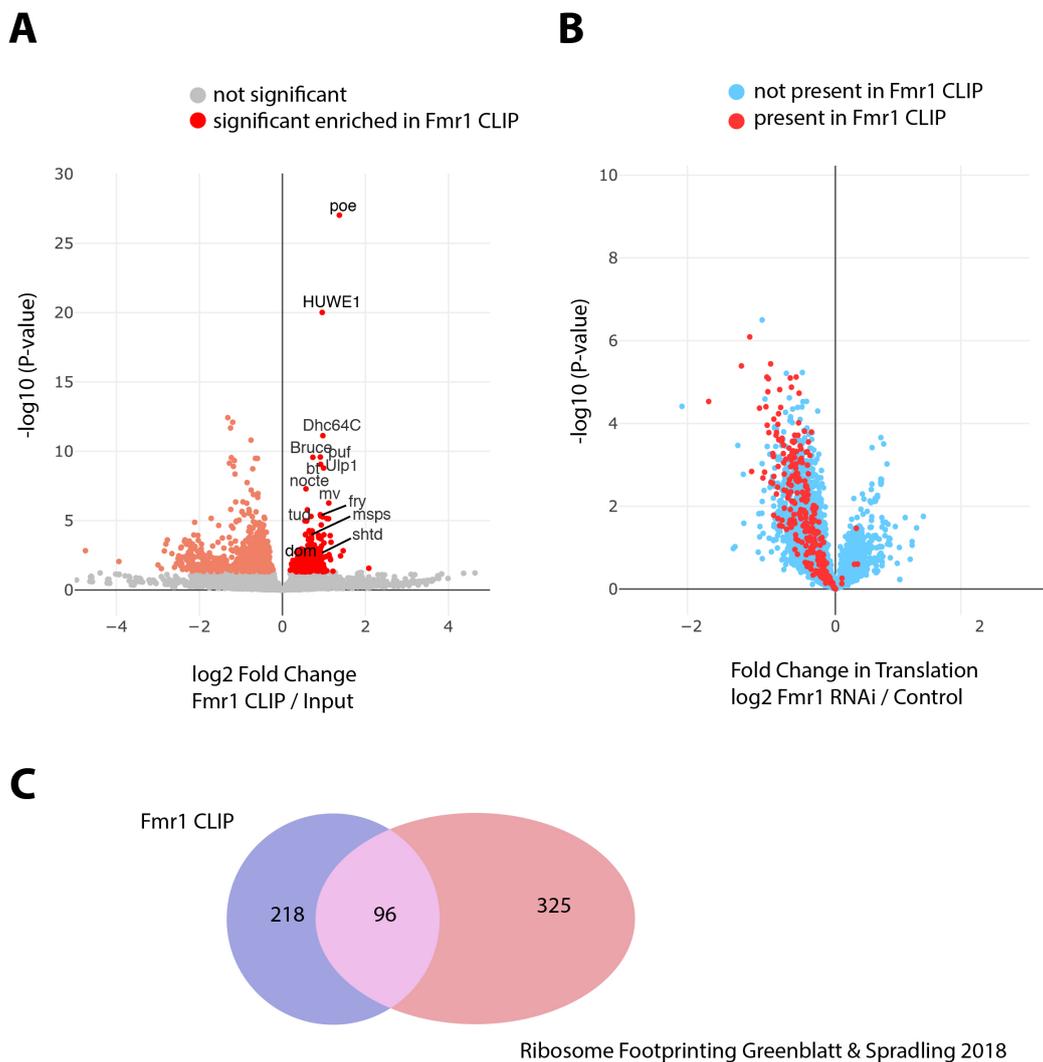
Figure 2.12: RNAs that show a reduced translation efficiency in Fmr1 depleted ovaries are partly present in the Fmr1 CLIP. A) Volcano plot showing significant enrichment for Fmr1 targets (red). Interesting candidates are annotated. B) Data from the Fmr1 CLIP were overlapped with the Ribosome Profiling data from Greenblatt and Spradling (2018). The volcano plot shows a decrease in translation of Fmr1 CLIP targets. C) Venny diagram showing the overlap between both datasets.

The two top hits from this analysis were poe and HUWE1, which were also amongst the most downregulated genes in the Ribosome Footprinting data from Greenblatt and Spradling (2018). Both proteins are E3 ubiquitin ligases and mutations contribute to early onset dementia and intellectual disability in humans (Monies et al., 2017; Bosshard et al., 2017). Interestingly, I found also some mitotic factors, such as msps and shtd (discussed above) and dom enriched in the Fmr1 CLIP that overlap with the Ribosome Footprinting data (Fig. 2.12.A). Furthermore, when I overlapped the Fmr1 CLIP targets with the Ribosome Footprinting dataset, most Fmr1 targets where actually down-translated suggesting a direct role of Fmr1 in this process (Fig. 2.12.B). Around 1/3 of Fmr1 targets found in the CLIP overlap with the Ribosome Footprinting dataset from Greenblatt and Spradling (2018) (Fig. 2.12.C).

2.1.6 Fmr1 and phase separation

It has been previously shown, that purified recombinant human FMRP can phase separate *in vitro* due to its C-terminal low-complexity region containing the RGG motif. Moreover, the low complexity region of FMRP has been shown to phase separate in combination with RNA in a sequence-independent manner (Tsang et al., 2019). Therefore, I wanted to explore whether the *Drosophila* homolog of Fmr1 has similar characteristics. The amino acide sequence of Fmr1 isoform A was fed into the Protein DisOrder prediction System which predicts natively disordered regions of proteins. Indeed the highest disorder probability was found within the C-terminal region of Fmr1, containing the RGG domain (Fig. 2.13.A). Changes in environmental conditions such as temperature, pH and salt has been shown to lower the threshold for phase separation (Alberti et al., 2019). I triggered condensate formation, using 3 mM recombinant purified protein containing only the RNA-binding domains of Fmr1, by changing the salt concentration of the proteins environment from high salt to a lower salt concentration. Droplets formed immediately and could be observed using differential interference contrast (DIC). No droplets were formed with GST-only protein (Fig. 2.13.B). Furthermore, I mixed either Tubulin or SatIII Sense RNA labelled with 488-UTP nucleotides with GST-KH-RGG protein. Droplet formation was observed using DIC and fluorescence microscopy. Both RNAs were able to demix with the protein containing the

low complexity region of Fmr1. As stated in Tsang et al. (2019), protein RNA droplets do not depend on a sequence-specific interaction but rather on the electrostatic interactions between protein and the RNA backbone. I confirmed that the *Drosophila* Fmr1 acts in a similar manner to the human homolog.
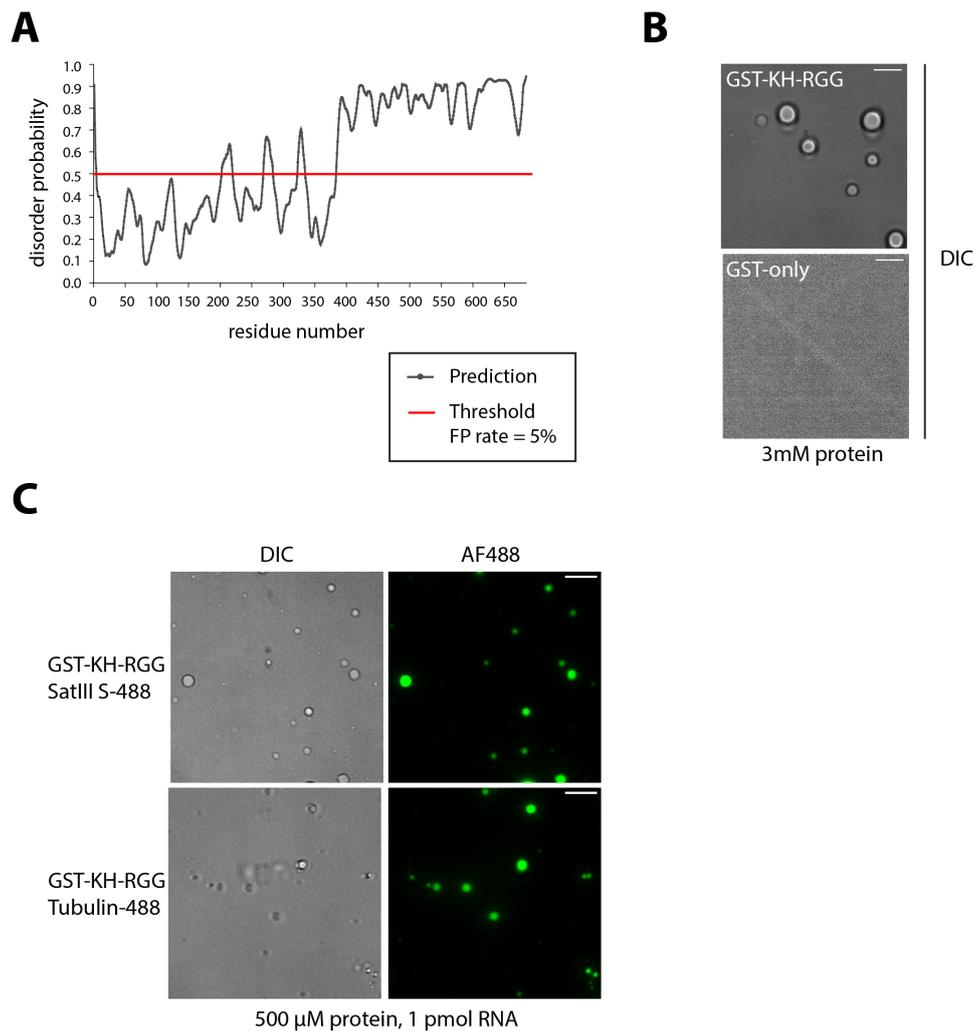
Figure 2.13: Fmr1 is able to form droplet *in-vitro*. A) The Protein DisOrder prediction System showed that Fmr1 contains a disordered region in the C-terminus of the protein which contains the RGG domain. B) Droplet formation could be induced by adding GST-KH-RGG from a high salt buffer into a low salt buffer. C) Addition of RNA fluorescently labelled 488-UTP nucleotides showed the formation of protein-RNA droplets. RNA demixing with GST-KH-RGG is independent of RNA sequence. Scale bar 5 μm.

2.1.7 Fmr1 interacts with Rump in an RNA-dependent manner

RNAs bound by RNA-binding proteins form dynamic ribonucleoprotein (RNP) complexes, which are key to every step of RNA metabolism from synthesis to decay (Beckmann et al., 2016). Therefore, I wanted to find out whether Fmr1 interacts with other RNA-binding proteins through co-binding to Fmr1's target RNAs. In order to do so, I conducted Fmr1 immunoprecipitations from ovaries and analysed the eluates by liquid chromatography-tandem mass spectrometry (LC-MS/MS) (Table A.3).
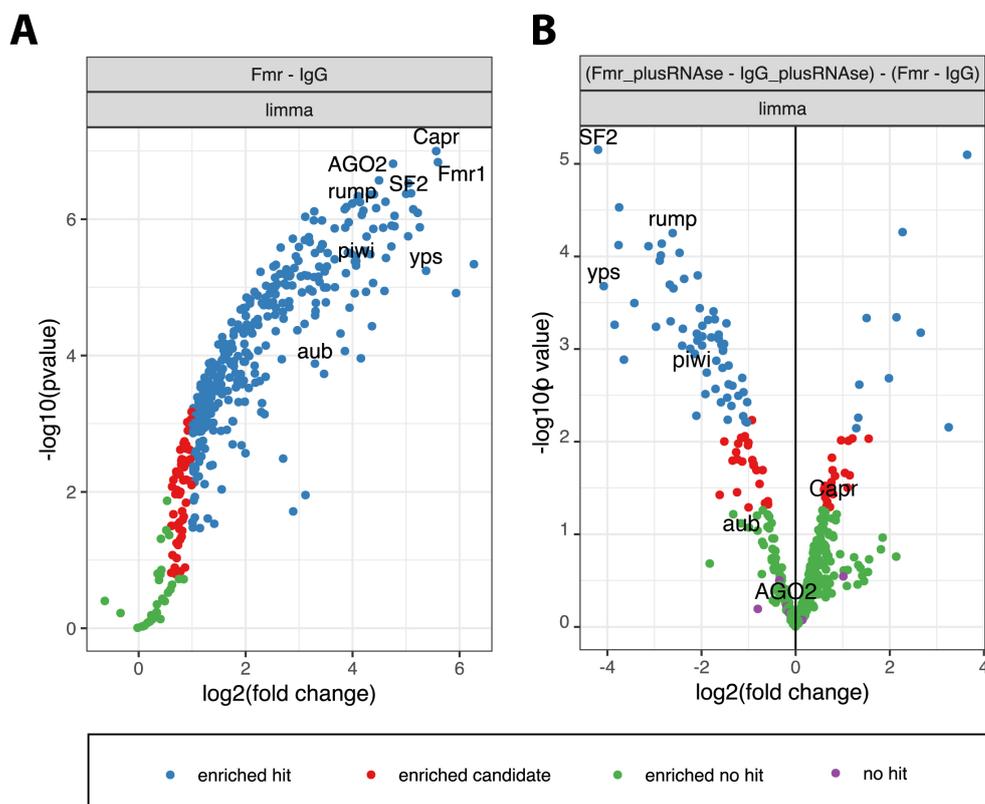
Figure 2.14: Novel and known interaction partners identified by Fmr1 mass spectrometry analysis. A) Interaction candidates of Fmr1 were identified by immunoprecipitation of Fmr1 from wt w^{1118} ovaries followed by quantitative tandem mass spectrometry. Unspecific binding was controlled by IgG which was also used as a reference for limma differential expression analysis. Resulting fold change over IgG values for each identified candidate was plotted on the x-axis and the corresponding p-value on the y-axis. Significant enriched hits marked in blue. B) To differentiate proteins that associate with Fmr1 by protein-protein interactions versus through co-binding to Fmr1's target RNAs, IPs were performed, respectively, in the presence or absence of RNase A. Proteins marked on the left side of the volcano are interactors of Fmr1 that depend on RNA.

Lysates were either left untreated or treated with RNase A in order to remove factors that are only bound to Fmr1 via common RNA targets. IgG was used as a control for unspecific binding and as a reference for differential expression analysis. The experiment was performed in duplicates and only significant interactors in both replicates were analysed. The volcano plot in Fig. 2.14.A shows enriched hits present in Fmr1 compared to IgG. One of the most enriched proteins is Fmr1 itself confirming the success of the immunoprecipitation. The list of enriched hits contained many known interactors of Fmr1, such as AGO2, piwi and aubergine linking Fmr1 to the RNAi and piRNA pathway (Specchia et al., 2017; Ishizuka et al., 2002). Caprin has also previously shown to interact with Fmr1 (Baumgartner et al., 2013). To my knowlegde, SF2, rump and yps are novel interactors of Fmr1 (Fig. 2.14.A). The volcano plot in Fig. 2.14.B compares the enrichment of proteins in the untreated condition (Fmr1 - IgG) to the RNase-treated condition (Fmr1 RNase-treated - IgG RNase-treated). Factors that are strongly enriched in the untreated condition versus treated condition are likely interactors that co-bind to Fmr1 target RNAs. Interestingly, the novel candidate proteins SF2, Rump, yps as well as piwi were found to be RNA-dependent interactors of Fmr1. Caprin is interacting with Fmr1 via protein-protein interaction, since the RNase treatment did not affect the interaction.

We found Rump to be the most interesting candidate as it was also found in the SatIII pulldown mass spectrometry data from Saskia Höcker. We therefore decided to focus our further study on Rump and its involvement in centromere regulation.

2.2 The RNA-binding protein Rump binds centromeric transcripts and controls CENP-A retention in mature sperm

2.2.1 Rump associates with satellite repeats of the 1.688 satellite family

The RNA-binding protein Rump has three RNA-binding domains (Fig. 2.15.A) and has been shown to associate to chromatin co-transcriptionally via premRNA (Kiesler et al., 2005). In order to study Rump's centromere-relevant functionality *in vivo*, I used a fly line with a mutant *rump¹* allele, that was generated by imprecise excision of the P-element *rump^{KG02834}*. This resulted in a deletion that removed the transcription start site, the entire 5'UTR and the first 152 codons of *rump* (Jain and Gavis, 2008). The *rump¹* mutant flies were confirmed by Western blot using whole fly extract. There was no Rump protein detected in the mutant compared to the wt OreR control (Fig. 2.15.B).
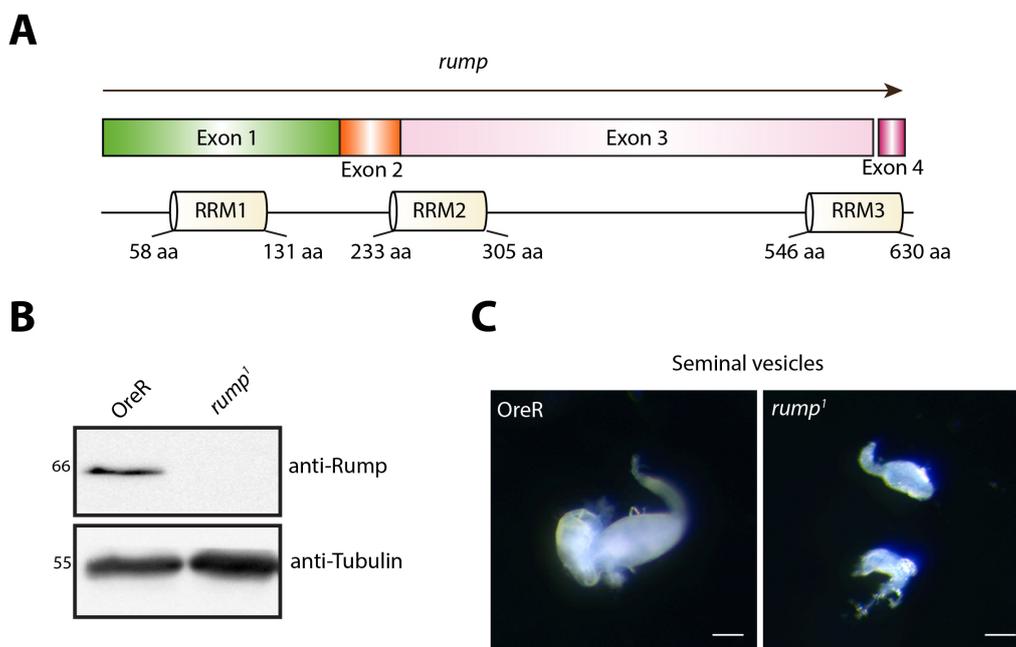
Figure 2.15: Seminal vesicles of wt Oregon-R (OreR) and *rump¹* mutant males. A) Schematic depicting the RNA-binding domains of Rump protein. B) Western blot with wt OreR and *rump¹* fly extract using Rump and Tubulin antibodies. C) Bright-field images of adult fly seminal vesicles. Flies were aged for 3 days before dissection. The seminal vesicles of *rump¹* mutant flies are smaller. Scale bar 100 μm.

Homozygous Rump mutant males derived from heterozygous females have been described to display partial male sterility, with only 10% of males being able to fertilize eggs (Jain and Gavis, 2008). I found that only about 3% of these embryos hatch (n=200), suggesting that despite fertilization, even fewer embryos develop. Indeed, Rump mutant seminal vesicles, in which mature sperm are stored for fertilization, are significantly smaller than wt (Fig. 2.15.C). This indicates defects during spermatogenesis, which may result in a lower sperm count in $rump^1$. However, even if reduced in numbers, motile and mature sperm can be found in almost every Rump mutant seminal vesicle.

Mass spectrometry data from our laboratory show that Rump binds to chromatin in an RNA-dependent manner (MS data Samuel Corless). Therefore, I decided to test whether Rump associates with SatIII transcript. To this end, I conducted an RNA Immunoprecipitation (RIP) followed by RT-qPCR. For this experiment 50 male flies were used for Rump IP and an IgG antibody of the same isotype was used as a control for unspecific binding. The IP was controlled by Western blot for the presence of Rump in the bound fraction of the IP (Fig. 2.16.A).
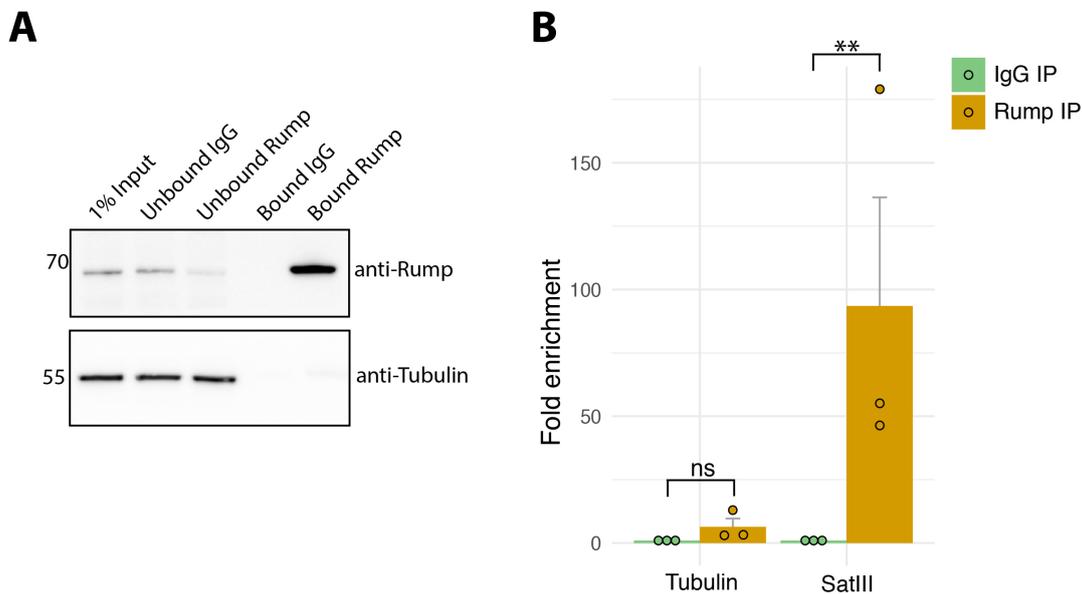
Figure 2.16: Rump binds to SatIII RNA *in vivo*. A) RNA immunoprecipitation was perfomed by immunoprecipitating Rump protein from whole male flies. The IP was verified on Western blot. B) RNA was extracted from Rump pulldowns. IgG was used as a control for unspecific binding. SatIII RNA and Tubulin mRNA were quantified by RT-qPCR and fold enrichment was calculated with the delta-delta Ct method. SatIII RNA was specifically enriched in the Rump pulldown. The standard deviation is the result of three independent experiments. The significance was calculated using a paired Student's t-test (p-value < 0.01).

RNA was extracted from the bound fraction of IgG and Rump IPs and assayed for accumulation of SatIII and Tubulin RNA by RT-qPCR. The fold enrichment between Rump IP and IgG Ip was plotted for both RNAs (Fig. 2.16.B). SatIII RNA was strongly enriched in the Rump IP, which was not the case for Tubulin mRNA. This indicates that Rump associates with SatIII transcripts *in vivo*.

2.2.2 Rump modulates 1.688 satellite repeat levels in testes

Since I found Rump to be bound to SatIII, we hypothesized that in the absence of Rump, SatIII transcript levels could be affected. In order to explore this, I carried out RT-qPCR for SatIII in testes from 3-day old males. I found that SatIII levels in the testes of Rump mutants were significantly reduced (Fig. 2.17.A). However, as satellite repeats from the 1.688 satellite family are very similar, I also tested the specificity of the SatIII qPCR primer pair. I performed PCR using this primer pair on plasmids containing one repeating unit of each subfamily of the 1.688 satellite family.
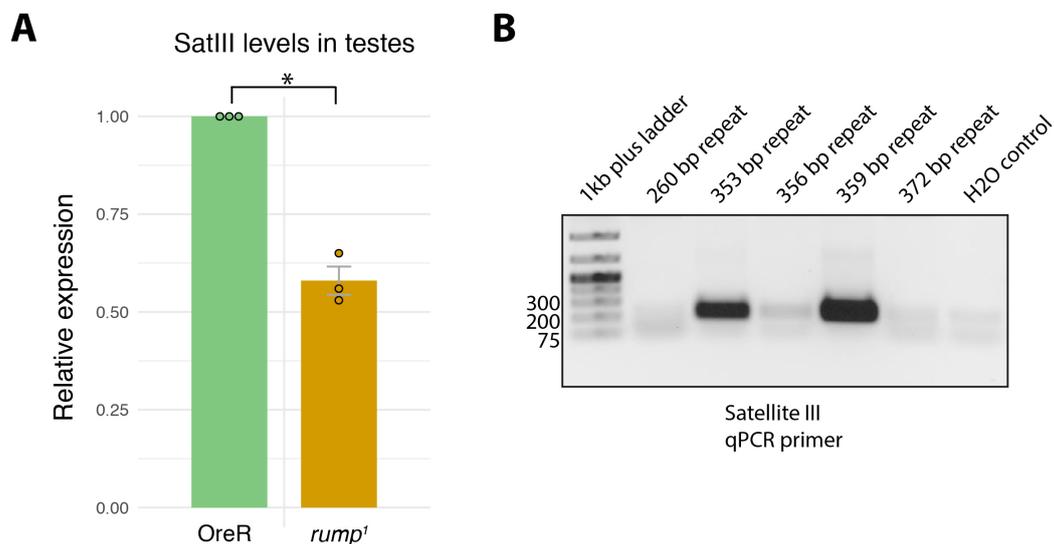
Figure 2.17: SatIII transcript levels in testes of *rump¹* mutant males are significantly reduced. A) RT-qPCR quantification of SatIII in wt OreR and Rump mutant testes. The error bar represents the standard deviation of three independent experiments. The significance was assessed using an unpaired Student's t-test (p-value < 0.05). B) One repeating unit of the indicated satellite repeats of the 1.688 family was cloned into pBluescript SK plasmid. The plasmids were used to test the specificity of the SatIII primer pair by PCR. The PCR products were loaded on Agarose gel and visualized. The primer recognizes the repeating unit of 353bp, 356bp and 359bp satellites.

The primer pair primarily detected 359 bp, but also detected 353 bp repeats and 356 bp to a very low level (Fig. 2.17.B). Hence, the RT-qPCR using the SatIII primer pair can not fully distinguish between the three satellite repeats and Rump might interact with either one or all of them. Nevertheless, the level of satellite repeat transcripts is reduced in *rump¹*.

In order to find out which satellite subfamilies of the 1.688 satellite family are reduced in *rump¹* mutant males, total RNA-Seq was performed using RNA from wt OreR and *rump¹* mutant testes. Therefore, three independent RNA extractions were carried out for each genotype and used for library preparation and sequencing. RNA-Seq analysis was carried out using the data from three independent replicates and reads were aligned to the novel *Drosophila* genome assembly that includes comples satellite repeats as well as CENP-A enriched sequences (Chang et al., 2019). Differential expression analysis with the Deseq2 package (Love et al., 2014) revealed that only the 356bp satellite repeats were downregulated upon Rump protein loss (Table A.10).
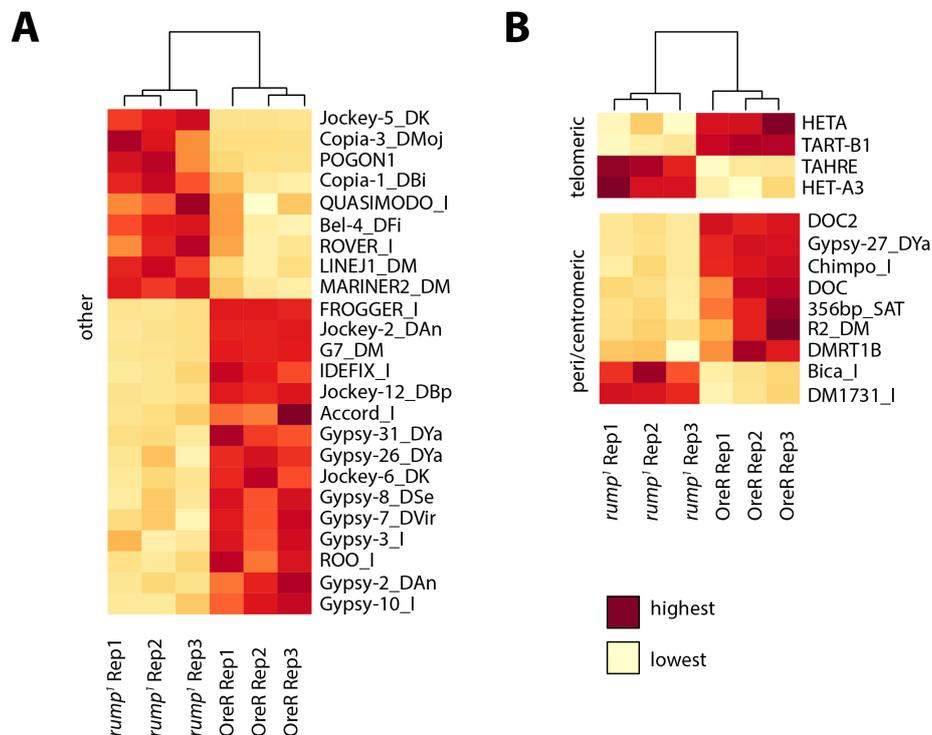
Figure 2.18: Transposable elements are differentially expressed in *rump¹* testes. A) Heatmap showing significant changes of transposable elements outside of centromeric and telomeric regions. B) Significant expression changes of transposable elements emerging from centromeric regions and the 356bp satellite repeat emerging from pericentromeric regions. Telomeric RNAs are depicted in the top panel.

This suggests, that the reduction in SatIII levels observed by RT-qPCR is a result of decreased levels of 356bp satellite repeats, rather than the other two satellite repeats (359bp and 353bp) that the SatIII primer pair recognizes. Interestingly, levels of certain transcripts of transposable elements that originate from the centromeric islands are also reduced, suggesting that Rump is required for transcription or RNA stability (Fig. 2.18.B). Differential expression of transposable elements could also be observed outside centromeric islands (Fig. 2.18.A).

2.2.3 Rump binds to transposable elements and satellite repeats in ovaries

To get a better understanding of how Rump modulates the expression levels of the transposable elements shown above, I wanted to study the target RNAs of Rump. Therefore, I performed Rump CLIP followed by RNA-Seq from ovaries as described for Fmr1 in Fig. 2.6. All three independent replicates of the Rump CLIP showed great distance to the respective input samples and were used for the analysis (Fig. A.1.B). As I did for the Fmr1 CLIP analysis, the samples were aligned to the *Drosophila* genome assembly that contains all complex satellites and CENP-A enriched transposable elements (Chang et al., 2019). Coverage of repeats from the same subfamily were summarized and normalized to the input. The Rump/input ratio was further normalized to the read depth. Enriched transposable elements were further annotated with their respective family name. In order to see the distribution of repeats that Rump preferentially binds to, the percentage of enriched repeat familes was plotted in a pie chart. Interestingly, almost half of all the Rump target repeat RNAs belong to the Gypsy-like family (Fig. 2.19.A). The top ten enriched transposable elements that were bound by Rump are also from the Gypsy clade (with exception of hAT-1N) (Fig. 2.19.B) (Table A.6). A high enrichment could be also found over transposable elements present in centromeric islands as well as satellite RNAs that emerge from pericentromeric regions (Fig. 2.19.C) (Table A.7).
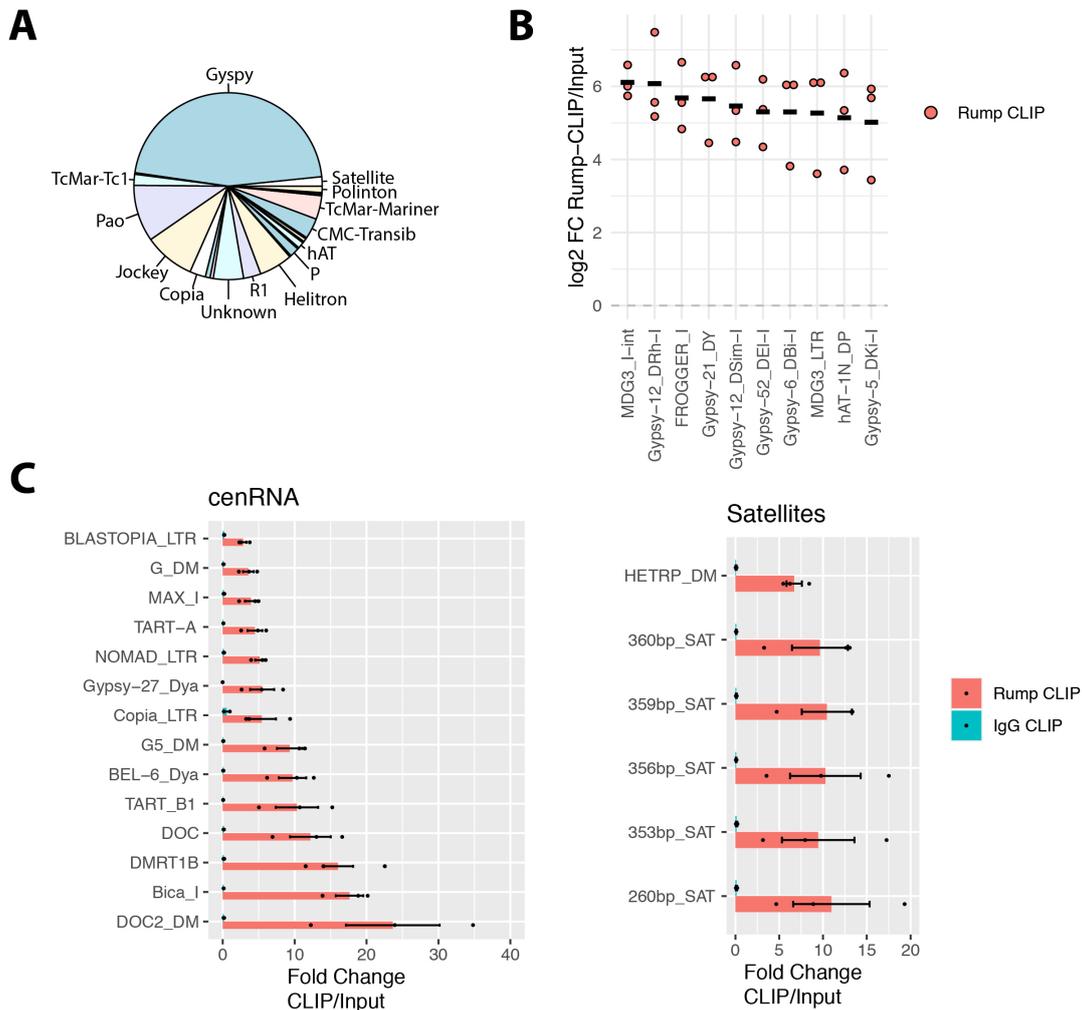
Figure 2.19: Enrichment of Rump over complex repeats. A) Cross-linking immunoprecipitation of Rump followed by RNA-Seq analysis was performed using fly ovaries. Sequencing data were aligned to the assembly containing centromeric contigs and complex satellite repeats from (Chang et al., 2019). Data obtained from three independent replicates were used for analysis. The pie chart depicts the distribution of enriched transposable elements in the Rump-CLIP compared to input. The majority of Rump binds to *gypsy*-like elements (46%) and to *pao* and *jockey*-like elements (10 and 9% respectively). B) Median plot showing the top 10 most enriched transposable elements of Rump. C) Barplot showing the enrichment over centromeric RNAs (cenRNAs) and satellites.

Since Rump was identified via mass spectrometry analysis as an RNA-dependent interactor of Fmr1, I aimed to find target RNAs that are commonly bound by both RNA-binding proteins. Since Fmr1 targets a lot of messenger RNAs (mRNAs), I also researched whether Rump binds to coding RNAs. Therefore, the Rump CLIP data were aligned to the *Drosophila* genome assembly BDGP6, which excludes the centromeric contigs but still contains transposable elements that are present outside of centromeres. Interestingly, by doing

so, I found only very few coding RNAs enriched in the Rump CLIP dataset (*e.g.* fru, Btk29A). In fact, the top enriched candidates found by this analysis, were (long) non-coding RNAs and transposable elements that are annotated in the assembly (Fig. A.2) (Table A.5).
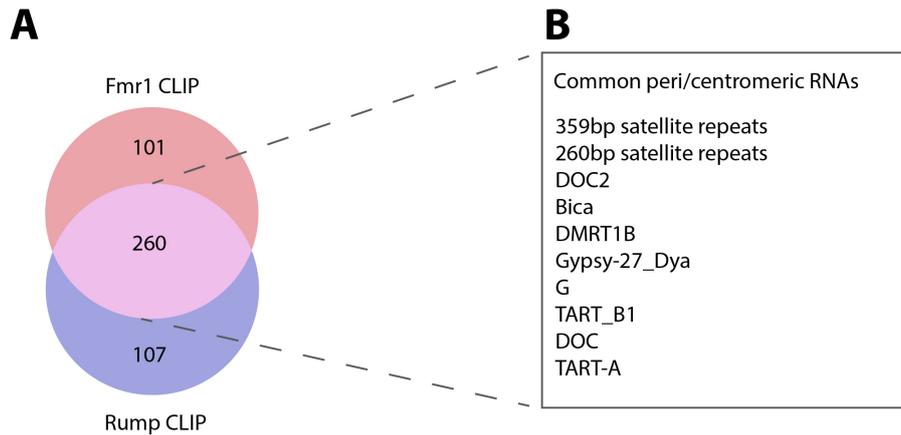
Figure 2.20: Transposable elements and satellite repeats that are common in Fmr1 and Rump CLIP A) Venny diagram depicting the overlap of RNAs that are targeted by Fmr1 and Rump. B) Common peri/centromeric RNAs are listed.

This suggests, that Fmr1 and Rump most likely co-bind to transposable elements and satellite repeats. Indeed, there was a large overlap of transposable elements that are bound by both factors (Fig. 2.20.A), some of which ermerge from centromeric islands or pericentromeric heterochromatin (Fig. 2.20.B).

2.2.4 Rump localizes to the nucleoplasm in spermatocytes of testes

Since most of the data related to Rump were produced from male flies or testes (with exception of the Rump CLIP data from ovaries), I decided to look at Rump's localisation in testes. To mark heterochromatic regions, I made use of a fly line that has a GFP-tag insertion at the N-terminal end of the open reading frame of Su(var)205 (GFP-HP1). During spermatogenesis, the stem cell daughter undergoes four rounds of mitotic divisions resulting in 16 primary spermatocytes that develop within 6 substages (S1-6) into mature spermatocytes. After the growth phase, spermatocytes undergo two meiotic divisions and differentiate into 64 spermatids that further elongate and mature

to form the sperm (Demarco et al., 2014). Interestingly, Rump is expressed throughout all stages of spermatocyte development and most strongly in S6 spermatocytes, which mark the onset of meiosis I (Giansanti and Fuller, 2012).
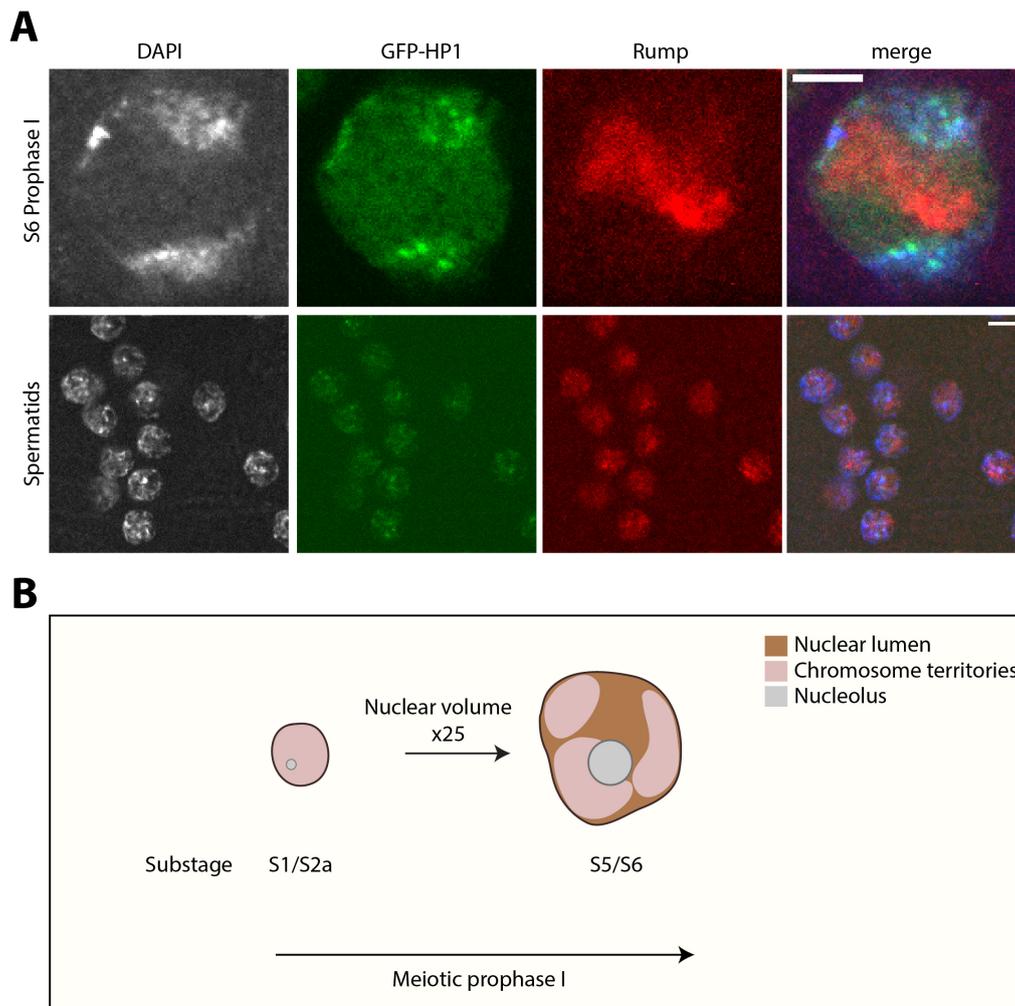
Figure 2.21: Rump does not localize to heterochromatic regions in S6 spermatocytes. A) To determine the localization of Rump in spermatocytes and post-meiotic spermatids, an immunofluorescence for Rump protein was conducted in testes from GFP-HP1 flies. An antibody against GFP was used to co-stain HP1. Rump did not colocalize with HP1. Scale bar 5 μm. B) Schematic of spermatocyte development. Adapted from (Collins et al., 2018).

Rump is not detectable in the meiotic stages that follow but is re-expressed in spermatids post-meiosis (Fig. 2.21.A). Spermatocyte growth leads to a 25-fold increase in volume and the four chromosomes are separated into three

areas with the 2nd and 3rd chromosome forming one large area each and the X-Y chromosomes clustering together with the 4th in another (Fig. 2.21.B) (Collins et al., 2018). Rump localizes to the nucleoplasm which contains very loose euchromatin that can not be stained by DAPI and the nucleolus. I could not confirm a direct co-localisation with HP1, which does, however, not exclude a low presence over these regions that can not be visualised by immunofluorescence.

2.2.5 Rump is required for CENP-A retention in mature sperm

Centromeric transcription has been shown to be essential for stable CENP-A incorporation (Bobkov et al., 2018). Since some of the transposable elements located at the centromere are downregulated in *rump¹* mutant testes, we hypothesized that CENP-A stability and incorporation might be impaired. To find out more about it, CENP-A was analysed in live mature sperm dissected from seminal vesicles of 3-day old wt, *rump¹* and *Zhr¹* males. *Zhr¹* flies were used because they lack most of the 359bp satellite repeats on chromsome X (Sawamura et al., 1993). CENP-A was visualized by Dendra2, which is a photoconvertible tag that was used to endogenously tag CENP-A by the Chen lab (Ranjan et al., 2019). The CENP-A-Dendra2 flies were fully viable and did not show impaired development. CENP-A-Dendra2 flies were crossed with *rump¹* and *Zhr¹* mutant flies and stocks homozygous for CENP-A-Dendra2 and *rump¹* or *Zhr¹* mutation were obtained. I tested whether CENP-A levels change in whole male flies and found that they were unaffected in both mutant stocks (Fig. A.4.A, B). Dendra2 was fully photo converted from green to red by a 10 second exposure to 405 fluorescent light. Interestingly, a significant reduction in CENP-A levels could be observed in mature sperm with *rump¹* and *Zhr¹* mutant background compared to wt flies expressing CENP-A-Dendra2 (Fig. 2.22.A, B). The experiment and data analysis was performed by Dr. Sreemukta Acharya in triplicates.
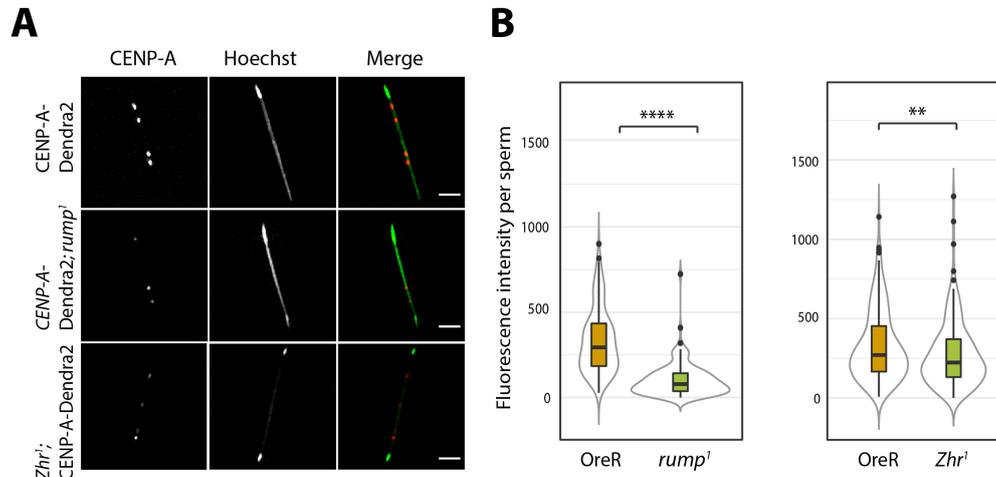
Figure 2.22: CENP-A is reduced in $rump^1$ mutant sperm and Zhr^1 sperm. A) CENP-A-Dendra2 was live imaged after photoconversion in mature sperm from wt, $rump^1$ and Zhr^1 mutant seminal vesicles. B) Total CENP-A levels were significally reduced from mutant sperm ($rump^1$ p-value < 0.0001 and Zhr^1 p-value < 0.01). Statistical significance was assessed using Mann-Whitney test. The error bars represent the standard deviation of pooled data points from three independent experiments. Scale bar 2 μm. Experiment and data analysis carried out by Dr. Sreemukta Acharya.

I further tested whether the absence of Fmr1 interferes with CENP-A retention in mature sperm by following the same technique. However, I could not observe the CENP-A reduction phenotype as CENP-A levels remained unchanged in mature sperm of $Fmr1^4$ mutants (Fig. A.5). This suggests, that Fmr1 has either a more redundant role in CENP-A retention or it is not involved in the pathway.

To test if CENP-A is not properly loaded in $rump^1$ mutant testes, I performed a Western blot to assess whether the protein levels of the loading factor Cal1 are changed (Chen et al., 2014). However, I did not detect a difference between the $rump^1$ and the wt control OreR (Fig. 2.23.A). Furthermore, I also tested a regulation of Cal1, the inner kinetochore protein Cenp-C and CENP-A at mRNA level by RNA-Seq. No significant difference was detected in mRNA levels of all three genes (Fig. 2.23.B). This suggests, that the CENP-A reduction phenotype observed is not due to erroneous CENP-A loading by Cal1, but rather due to misregulated cenRNA, or because cenRNA might be involved in accurate loading of CENP-A as suggested in Rošić et al. (2014).
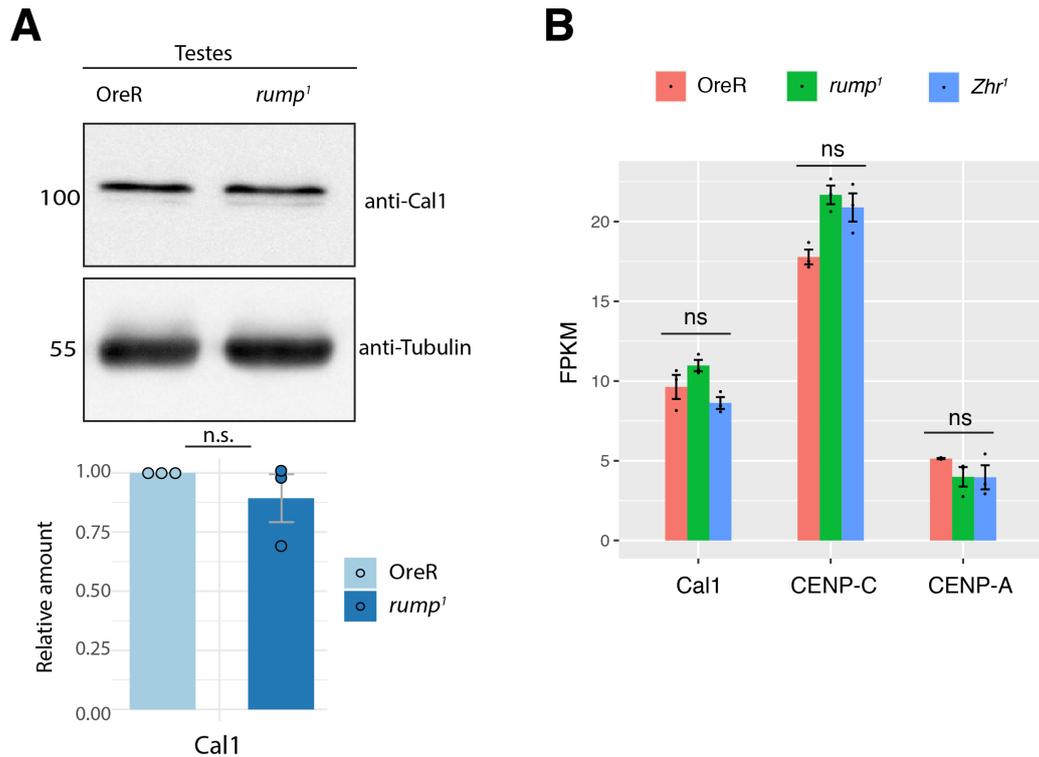
Figure 2.23: The protein levels of the CENP-A loading factor Cal1 are unchanged in $rump^1$ mutant testes. A) To test potential Cal1 protein level changes in $rump^1$ Western blot from wt and mutant testes was performed. The band intensity was assessed using unpaired Student's t-test. The error bars represent the standard deviation of pooled data from three independent experiments. Cal1 levels were unchanged in mutant testes. B) mRNA level changes of Cal1, CENP-C and CENP-A were assessed by analysing the RNA-Seq data of OreR, $rump^1$ and Zhr^1 obtained from testes. There was no significant change detectable.

Since the CENP-A reduction phenotype was also observed in the Zhr^1 mutant fly line, I aimed to characterize these flies better. It is known that the Zhr^1 flies have lower 359bp satellite repeat DNA levels due to a X-Y translocation which deleted large blocks of centromeric heterochromatin (Sawamura and Yamamoto, 1993). However, the details of X-chromosome heterochromatin composition in these flies remains uncharacterised, largely because of the problems to properly annotate the highly repetitive pericentromeric regions. To fully understand which RNAs might be downregulated, I carried out RNA-Seq from Zhr^1 testes. As a result I could see a clear downregulation of 359bp and 353bp satellite repeats (Fig. 2.24.B). Both repeats are found in arrays together in the genome assembly containing the centromeric contigs which was used for the analysis of these data.
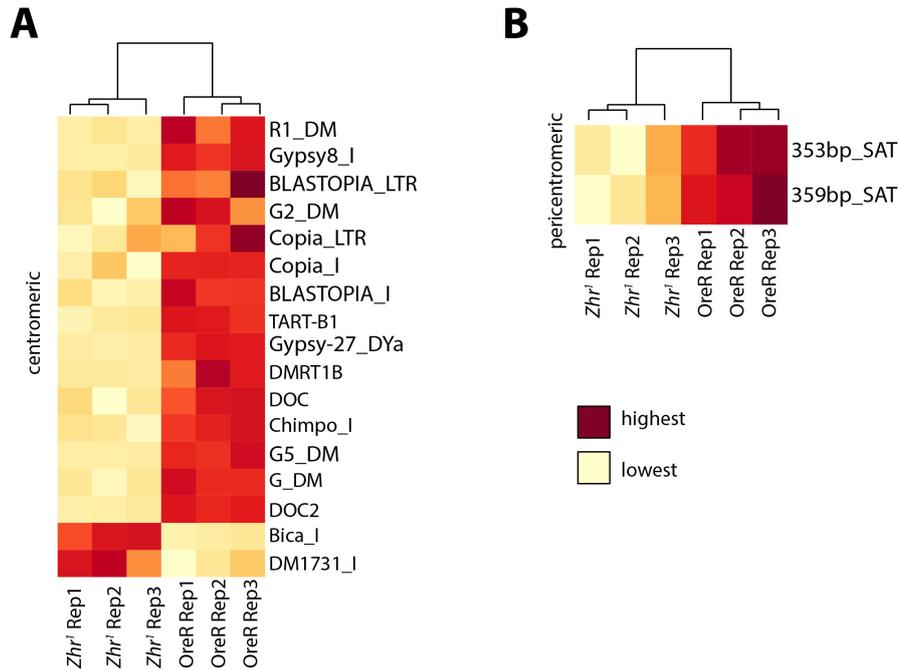
Figure 2.24: Characterization of the Zhr^1 mutant transcriptome. A) The Zhr^1 mutant shows depletion in retroelements emerging from the centromere as well as a strong downregulation of the 353bp and 359bp satellite repeats B).

Interestingly, also several transposable elements that are found in the centromeric islands of chromosome X and Y are downregulated in the Zhr^1 mutant (Fig. 2.24.A) (Table A.11). This suggests that these regions are likely missing or disturbed by the X-Y translocation. Interestingly, the transposable elements Bica and DM1731 are upregulated in Zhr^1. Both of them are moderately enriched with CENP-A CHIP-Seq but not present in centromere islands (Chang et al., 2019). It might be a secondary or compensatory effect due to misregulation of the other transposable elements. Alternatively, either of the RNA-binding proteins Rump or Fmr1 might be misregulated in Zhr^1.

Chapter 3

Discussion

Centromeres are epigenetically marked by the presence of the histone H3 variant CENP-A and are embedded into highly heterochromatic chromatin. The entire region undergoes active transcription in spite of its highly compacted state. Interestingly, centromeric transcription is essential for centromere specificity and function. However, how (peri-)centromeric transcripts precisely carry out their functions remains unknown, which is what I have addressed during the course of my PhD thesis.

RNA-binding proteins are bound to nascent transcripts and are key players in the regulation of every step in RNA metabolism. I therefore focused on identifying RNA-binding proteins (RBPs) relevant to centromeric RNA regulation and centromere function for this doctoral thesis. To this end, I used published mass spectrometry data from our laboratory, which were obtained from a cross-linked CENP-A pulldown (Demirdizen et al., 2019). The cross-linked approach offers the possibility of identifying weak and transient interactors (Klockenbusch and Kast, 2010). This was helpful to the aim of identifying RBPs associated with centromeric chromatin, as they may not actually interact with CENP-A directly, but rather bind to centromeric chromatin via RNA. By re-analysing the data from Demirdizen et al. (2019), I found the RBP Fmr1 as one of the top enriched factors (Table A.1). I further showed that Fmr1 is able to bind to (peri-)centromeric RNAs both *in vitro* and *in vivo* and affects their levels in the germline of *Drosophila melanogaster*. Furthermore, Fmr1 also binds to mRNAs that encode for proteins important for

centromere function and positively affects their translation. Since most RBPs form ribonucleoprotein (RNP) complexes with other RBPs, I set out to identify novel interactors of Fmr1 that may co-bind to target RNAs. In this way, I identified the RBP Rump, which co-binds to centromeric transcripts with Fmr1, suggesting that Rump and Fmr1 are components of a common RNP complex. Additionally, Rump was found to be crucial in maintaining CENP-A levels in mature sperm. In conclusion, this thesis demonstrates that the RBPs Fmr1 and Rump are important novel centromere regulators.

3.1 Fmr1 is an RNA-dependent centromeric factor

Besides, the cross-linked mass spectrometry dataset of Demirdizen et al. (2019), Fmr1 was also present in mass spectrometry data from a SatIII RNA pulldown (performed by Saskia Höcker, unpublished data). SatIII RNA is originating from the pericentromeric 1.688 satellite DNA and the longest transcript that was found to be transcribed in sense direction consists of two repeating units of the 359 bp satellite family (Rošić et al., 2014). Therefore, I studied the interaction of Fmr1 with the SatIII Sense transcript *in vitro* and was able to confirm the association (Fig. 2.1). Additionally, recombinant purified Fmr1 protein was able to pull down CENP-A mononucleosomes, thus confirming the CENP-A mass spectrometry data. CENP-A binding through Fmr1 occured in an RNA-dependent manner, indicating that Fmr1 does not directly associate with CENP-A protein, but binds to centromeric RNAs in the vicinity (Fig. 2.4). However, the experiment did not rule out the possibility of Fmr1 binding to canonical histones or other histone variants. It is possible that Fmr1 binds to chromatin in general in an RNA-dependent manner. As a result, Fmr1 is likely to localize depending on the type of RNA it targets in chromatin. Future experiments are required to determine whether this is the case in *Drosophila*.

The next question I wanted to explore was whether the binding of Fmr1 to SatIII RNA occurs because of the secondary structure the RNA forms, or due to the sequence itself. The SatIII Sense transcript likely forms a secondary structure *in vitro*, since a second band appeared under native conditions (Fig.

2.1.D). I found that the binding of Fmr1 to the SatIII Sense transcript is likely sequence-specific, since only the KH domain (Ashley et al., 1993) was able to associate with the RNA (Fig. 2.2.A). The RGG domain of Fmr1 is known to bind to secondary structures of RNAs, such as G-quadruplex structures but was not sufficient to interact with SatIII on its own (Fig. 2.2.B). It can, however, not be excluded that the binding to the RNA exclusively requires the KH domain, as the RNA-binding could become more stable when the other domain is present as well. Indeed, the affinity of Fmr1 to SatIII was greater when the full length protein was used compared to the KH domain alone. However, this could also be explained by the fact that the KH domain was purified as a trimer, which might have resulted in fewer accessible binding sites. In regards of the sequence properties of SatIII RNA, it is worth noting, that it has a very high AU content (70%), which could lead to the affinity of Fmr1 to the RNA *in vitro*. The reason is, that FXR1, an autosomal paralog of FMRP in human binds to AU-rich elements that are present in mRNA 3'-UTRs (Vasudevan and Steitz, 2007). Therefore, it was essential to confirm these data *in vivo*, in order to see whether the interaction also occurs under more physiological conditions.

3.2 Fmr1 - an important factor during cell division

Fmr1 mutant embryos display severe mitotic defects (Fig. 2.8), similar to *Zhr¹* mutant embryos that have reduced levels of SatIII transcripts (Rošić et al., 2014; Deshpande et al., 2006), indicating that Fmr1 may be required to regulate SatIII RNA and/or other centromeric transcripts. RNAs derived from SatDNAs were suggested to play important roles during early embryogenesis in insects (Wei et al., 2021; Pathak et al., 2013) and therefore, it is likely that the observed phenotype is due to a maternal-effect defect. Furthermore, loss of Fmr1 leads to reduced fertility in both sexes (Fig. 2.3.C) (Zhang et al., 2004), which is why I aimed to study Fmr1's function in the germline. In the germline, transposable elements pose a threat to the genome for which animals have developed responses to safeguard genome integrity (Zamudio and Bourc'his,

2010). Transposable elements are present in islands at the centromere (Chang et al., 2019) and therefore, we were interested in studying the binding of Fmr1 not only to SatIII, but also to other centromeric RNAs, including transposable elements *in vivo*. To this end, I carried out crosslinked immunoprecipitation (CLIP)-Seq from ovaries. CLIP-Seq is the state of the art technique to identify target RNAs of an RBP and up to date Fmr1 CLIP-Seq has been carried out using brains from mice or human cells in order to shed light on Fmr1's neurological functions (Ascano et al., 2012; Sawicka et al., 2019; Darnell et al., 2011; Maurin et al., 2018). All of these studies focused on uniquely mapped reads and ignored repetitive sequences, leaving the extent of Fmr1's binding to these, unexplored. Thus, the Fmr1 CLIP dataset that I generated for this study was analysed for both, multimapped repetitive sequence and unique mappers, such as mRNAs. As a result, I discovered a subset of mRNAs that are targeted by Fmr1 (Fig. 2.12.A). Furthermore, centromeric repeat RNAs as well as SatIII RNA was found to be enriched in the dataset (Fig. 2.6.B).

3.2.1 Fmr1 binding to target mRNA and translation activation in ovaries

Fmr1 has been recently shown to promote the translation of large autism-related mRNAs in ovaries (Greenblatt and Spradling, 2018). Hence, one part of this thesis aimed to determine if translation of mRNAs that encode proteins involved in mitosis and centromeres are similarly affected. Re-analysis of the ribosome footprinting data from Greenblatt and Spradling (2018) showed that several mitotic proteins are down-translated upon Fmr1 RNAi, including kinetochore proteins, microtubule motor proteins and components of the anaphase promoting complex (Fig. 2.9). Although their mRNA levels are comparable between wt and Fmr1-depleted ovaries, Spc105R and BubR1 protein levels are reduced, suggesting translational regulation (Fig. 2.10). The results were further supported by the reduction of mRNA levels at polysomes in ovaries that are depleted of Fmr1 (Fig. 2.11). But how does Fmr1 precisely activate the translation of larger transcripts? To gain a better understanding of this process, the Fmr1 CLIP data were analysed for uniquely mapped-reads in order to find novel target mRNAs of Fmr1 in the female germline. The top enriched mRNAs from this analysis were poe and Huwe1 (Fig. 2.12.A), both of which

are involved in dementia and intellectual disability in humans (Monies et al., 2017; Bosshard et al., 2017). Interestingly, poe and Huwe1 were amongst the top down-regulated genes in the Ribosome Footprinting data from Greenblatt and Spradling (2018), suggesting a direct connection between Fmr1 binding and translation. Some of the mitotic factors that were down-translated due to the absence of Fmr1 were also enriched in the Fmr1 CLIP. The mRNAs of Spc105R and BubR1 were not found to be enriched for Fmr1. In addition to the possibility that these mRNAs are not targets of Fmr1, it should be noted that the ovaries were cross-linked with UV, which may limit the penetration of the radiation due to the thickness of the tissue. As a result, it is likely that some RNAs, particularly weakly associated ones or low abundant RNAs, are not captured in the process, even though they may be targets of Fmr1. Interestingly, overlaying the Fmr1 CLIP data with the published Ribosome Footprinting data resulted in around 1/3 overlap, with most of the Fmr1 targets being down-translated (Fig. 2.12.B, C). This suggests, that Fmr1 activates the translation of mRNAs by a mechanism that involves direct binding. A recent publication reported a novel function of the piwi protein Aubergine (Aub) in translational control of mRNAs in the germline of *Drosophila* (Ramat et al., 2020). The authors showed that Aub physically interacts with the translation initiation factors poly(A)-binding protein (pAbp) and eIF3, leading to translational activation of nos mRNA. Taking this into account, Fmr1 might also be involved in the same pathway, since it is a known interaction partner of Aub. Further, the mass spectrometry data acquired by Fmr1 pull down in this thesis showed an enrichment of the translation initiation factor eIF3, Aub as well as pAbp (Table A.3). Further studies are needed to ascertain whether Fmr1 and Aub act in the same pathway to promote translation in the female germline.

Spc105R is part of the KMN complex, which provides major microtubule binding activity to the kinetochore (Schittenhelm et al., 2009) and BubR1 is a key protein involved in the spindle assembly checkpoint (Musacchio and Salmon, 2007). Reduced protein levels can, therefore, jeopardize the fidelity of chromosome segregation, as observed in the $Fmr1^4$ mutant. Furthermore, both proteins are found to have roles during female meiosis: Spc105R depletion in

Drosophila oocytes results in disrupted kinetochore assembly and unattached microtubules (Radford et al., 2015), BubR1 depletion leads to a similar phenotype in mouse oocytes and impairs meiotic cell cycle progression (Wei et al., 2010). This suggests, that proteins important for centromere function are present in oocytes and likely maternally transmitted into the next generation, where they are required for proper cell cycle control. Cell cycle defects in the Fmr1 mutant embryo have been described before and were attributed to defects in centrosome assembly (Deshpande et al., 2006). Therefore, reduced levels of mitotic proteins as a result of Fmr1 loss may contribute to the observed cell cycle defects. However, further implications of Fmr1 in centromere regulation can not be excluded, as it was also shown that Fmr1 is required for proper HP1 localization in the precursor cells of the germline (pole cells) and assembly of centric heterochromatin (Deshpande et al., 2006). Hence, the main part of this thesis investigated whether Fmr1 regulates centromeric RNAs and to find novel interactors involved in the same pathway.

3.2.2 A potential nuclear function of Fmr1

The results of the Fmr1 CLIP-Seq experiment showed that Fmr1 binds to 359 bp satellite RNA in the female germline, which confirmed the *in vitro* result. Apart from 359 bp repeats, also the 260 bp satellite repeats and Dodeca satellite repeats were enriched in the Fmr1 CLIP (Fig. 2.6.B). It is interesting, that other satellite repeats of the 1.688 family, such as the 356 bp and 353 bp satellite repeats were not enriched, albeit having a high sequence similarity to 359 bp repeats. Additionally, centromeric transcripts from transposable elements within and outside of centromere islands are bound by Fmr1. Some of them have a moderate to strong enrichment of CENP-A at chromatin level (*e.g.* DMRT1B, Gyspy8, DM1731, DOC2, TART-A, Bica) (Chang et al., 2019), which may explain why Fmr1 could pull down CENP-A containing mononucleosomes.

Despite its high cytoplasmic localization, there is evidence that FMR1 shuttles between the nucleus and cytoplasm, suggesting that FMR1 has nuclear functions. Indeed, Fmr1 has been shown to be important for mRNA transport and nuclear export of m^6A-containing mRNAs (Edens et al., 2019). In addition,

FMRP was reported to bind to chromatin during the DNA damage response (Alpatov et al., 2014). SatIII RNA has not been observed outside the nucleus (RNA FISH data from our laboratory) and generally cenRNA transcripts are likely chromatin-bound. The binding of Fmr1 to cenRNA transcripts suggests a novel nuclear function.

As described in the introduction (see Section 1.2.6), transposable elements are usually silenced by the piRNA pathway in the germline of many species, which is thought to be a defence system to safeguard genome integrity (Brennecke et al., 2007). Fmr1 has been shown to play essential roles in the piRNA pathway in both male and female germline, which was demonstrated by its involvement in the *crystal-Stellate* system. The *crystal* locus (also known as *Suppressor of Stellate*) produces piRNAs, which silence the expression of the *Stellate* RNA. Loss of Fmr1 leads to the accumulation of crystalline aggregates in spermatocytes, due to the production of the Stellate protein. Furthermore, Fmr1 loss results in a derepression of transposable elements in testes and ovaries (Bozzetti et al., 2015). Therefore, I wanted to study, whether the binding of Fmr1 to the RNAs of transposable elements leads to the disruption in the piRNA pathway. As a result, I expected an accumulation of transposable element transcripts, since mutations affecting the piRNA pathway induce transposable element up-regulation, leading to deleterious and cumulative effects that ultimately result in animal sterility. To find out about it, an RNA-Seq experiment was carried out from wt and Fmr1 mutant testes and analysed for differential expression of multi-mapped reads. As a result, I found an up-regulation of certain transposable elements due to Fmr1 loss, such as for example R1_DM (Fig. 2.7.A), in accordance with literature (Bozzetti et al., 2015). However, to my surprise, many transposable element transcripts were down-regulated upon Fmr1 loss-of-function. Significant down-regulation was observed with transposable elements that are present at the centromere, with the exception of DM1731 and Bica. Furthermore, transcripts of the 1.688 subfamilies 359 bp, 356 bp and 353 bp showed a severe reduction in their RNA levels (Fig. 2.7.B). In addition, most of the repetitive RNAs that showed a significant change in their levels, were also found to be enriched in the Fmr1 CLIP. This suggests that Fmr1 targets repeating RNAs, *i.e.* satel-

lite repeats and transposable elements, and regulates their levels. Since CLIP was performed using ovaries and RNA-Seq was performed using testes, direct comparison is not possible. As many target RNAs of Fmr1 from the female germline cause differential expression changes in the male germline when Fmr1 is absent, it is however possible that Fmr1 plays a role in both male and female germline pathways. More studies are required to dissect whether this is the case.

3.3 Phase separation - a mechanism by which Fmr1 carries out its functions?

FMRP has an isolated disordered region containing the RGG motif, which has been shown to be sufficient for phase separation (Tsang et al., 2019). In this thesis, I was able to demonstrate that the RGG domain of the *Drosophila* homolog of Fmr1 is sufficient to induce droplet formation *in vitro* (Fig. 2.13.B). Furthermore, RNA, independent of its sequence, was found to associate with the droplets (Fig. 2.13.C). A phase separation dependent mechanism has been proposed to regulate translation by FMRP. In the model of Tsang et al. (2019) phase separation is promoted by phosphorylation of FMRP in neurons, resulting in neuronal granule assembly and translational silencing. Dephosphorylation of FMRP promotes neuronal granule disassembly and activates translation. Post-translational modifications (PTM) are important regulators of biomolecular condensates, as they modulate protein valency and interaction strength. PTMs occur rapidly and are reversible, which opens up the possibility to fine-tune phase separation in response to different cues (Snead and Gladfelter, 2019).

Phase separation has proven to be a regulatory component in many processes. As described in the introduction, heterochromatin formation is linked to phase separation using repetitive RNAs as driving force (Huo et al., 2020; Strom et al., 2017). As an example, the nuclear matrix protein SAFB interacts with the major satellite RNA in mouse cells, thereby driving phase separation required for heterochromatin stabilization and condensation (Huo et al., 2020). It is possible that Fmr1 may be required in a similar scenario to maintain

RNAs on chromatin. This may be a prerequisite for heterochromatin formation, which includes the recruitment of HP1 protein and the spreading of the H3K9me3 mark. This hypothesis is supported by the fact that HP1 is mislocalized in Fmr1 mutant precursor germ line cells (Deshpande et al., 2006). Henceforth, the absence of Fmr1 may lead to RNA instability and degradation at (peri-)centromeric regions.

In other recent studies, phase separation has been proposed to play a role in regulating transcription by RNA Polymerase II. Condensates have been shown to form around transcription factors and serve for the concentration and delivery of proteins needed for transcription initiation (Cramer, 2019). Considering what is known about phase separation and Fmr1 in different contexts, combined with my data, allows for further speculation. Fmr1 may be required for condensate formation at the centromere or pericentromeric heterochromatin, thereby helping attract transcription factors. A mechanism like this may be useful for allowing transcription to occur in an untypical environment characterized by highly compacted chromatin. Additionally, repetitive RNAs may work as a scaffold for phase separation, since their repeated sequences could bind RBPs in a multivalent manner, resulting in a scaffold for condensation. PTM of Fmr1 may help in promoting condensate formation and keeping the chromatin environment dynamic. The recent finding of casein kinase II at centromeres, the kinase that phosphorylates proteins such as Fmr1, further supports this hypothesis (Siomi et al., 2002; Huang et al., 2019). Taken together, phase separation may represent a mechanism by which (peri-)centromeric transcript levels are regulated by Fmr1. Fmr1 may either stabilize RNA on chromatin and contribute to heterochromatin formation, or assist in attracting factors required for transcription.

3.4 Fmr1 potentially interacts with proteins involved in splicing and chromatin binding

There are several different types of granules known to contain FMRP protein, including phase separated structures of RNP complexes, Cajal bodies, P-bodies, stress granules, and transport granules (Dury et al., 2013). The

formation of granules can be regulated by FMRP, and dysregulation of granules that contain FMRP is associated with neurological disorders (Lai et al., 2020). As a result, in order to better understand the role of FMRP, it is vital to study the composition of RNP complexes that include FMRP. Therefore, I sought to investigate the interactome of Fmr1 in ovaries. The enriched proteins included many already known Fmr1 interactors, which confirmed the validity of the experiment, and furthermore, several novel Fmr1 interactors were detected. As expected, several of interacting proteins were found to be RNA-binding (26%), suggesting common roles in RNA metabolism. In order to determine which RBPs interact with Fmr1 through co-binding to Fmr1's target RNAs, I included a condition that was treated with RNase and compared to the untreated condition. This led to the identification of RNA co-binding factors involved in splicing, including RBP Rump (Fig. 2.14). In light of Fmr1's nuclear functions, potential interactors were identified within the context of splicing, DNA repair, chromatin binding, and nucleocytoplasmic shuttling. Known interactors include factors relevant for translation regulation, RNA interference, and the piRNA pathway, among others. Overall, most factors were related to translation and splicing (Table A.3).

As a potential novel interactor, Rump stood out, as it is not only involved in splicing but also a known RNA-dependent chromatin bound protein (Kiesler et al., 2005). Furthermore, Rump was present in SatIII mass spectrometry data from our laboratory (Höcker et al., unpublished), suggesting a (peri-) centromeric function. Homozygous Rump mutant males are partially sterile, and only few embryos develop when Rump mutant males are crossed to wt females. Rump mutant males have smaller seminal vesicles and a lower sperm count (Fig. 2.15.C). Although it may be the cause of impaired fertilization, it may also be caused by defective sperm. Motile and mature sperm are present, and a few embryos are able to develop, which implies that sperm penetrates the egg and issues begin to arise at later stages. Embryos that were produced by rump mutant females were also shown to arrest development prematurely (Jain and Gavis, 2008). Similar to Fmr1, Rump plays an important role in the germline of *Drosophila* and is required for faithful development of the offspring. This indicates, that Fmr1 and Rump might have related functions

at the centromere and led me to characterize Rump further.

3.5 The RBP Rump co-binds cenRNAs with Fmr1 and maintains CENP-A in mature sperm

In order to determine whether Rump and Fmr1 bind to the same target RNAs, I conducted Rump CLIP from ovaries similarly to Fmr1. Surprisingly, the majority of target RNAs of Rump were non-coding RNAs, including transposable elements, satellite and other long non-coding RNAs (lncRNAs) (Fig. 2.19). This indicates that Rump co-binds to these RNAs with Fmr1. Indeed, common centromeric transcripts are found for both proteins (Fig. 2.20). Furthermore, Rump is a strong binding factor of satellite repeats, including 359 bp repeats in both female germline and male flies (Fig. 2.16.B and 2.19.C). This is indicative for a common pathway that includes Rump in both male and female germline.

A RNA-Seq experiment was conducted on Rump mutant testes to see if binding of Rump to non-coding transcripts impacts their levels, comparable to the observations made with Fmr1. As a result, transposable element RNA levels from centromeres were significantly down-regulated in Rump testes, with exception of Bica and DM1731, similar to Fmr1. An up-regulation of certain transposable elements due to the loss of Rump was also observed, suggesting that Rump might act in the piRNA pathway with Fmr1 (Fig. 2.18). In line with this idea, Aub was shown to co-purify with Rump in a complex with nos RNA (Becalska et al., 2011). However, to my knowledge, a role of Rump in the piRNA pathway has not been reported. Overall, Rump and Fmr1 share centromeric target RNAs mostly leading to their down-regulation.

Furthermore, Rump loss led to a reduction in CENP-A in mature sperm (Fig. 2.22). Drosophila mature sperm undergo drastic chromatin organization changes, since histones are exchanged to protamines. However, CENP-A is one of the few histone variants that are retained on the mature sperm and

is transmitted into the next generation. Loss of CENP-A from mature sperm results in defects during the first cell cycle in the offspring and developmental arrest (Raychaudhuri et al., 2012). CENP-A is loaded by its loading factor Cal1, but the levels of Cal1 were unaffected in the Rump mutant testes, suggesting that the reduced CENP-A phenotype is not due to defective CENP-A loading (Fig. 2.23). Interestingly, mature sperm from the Zhr^1 mutant, displayed a similar phenotype of defective CENP-A retention as observed with the Rump mutant (Fig. 2.22). In this thesis, I characterized the Zhr^1 mutation, which results from a X-Y translocation, by RNA-Seq. As expected, transcripts of pericentromeric satellite repeats located on the X-chromosome were down-regulated. Additionally, transcripts of transposable elements from the X and Y-chromosome were down-regulated, too (Fig. 2.24). This suggests, that either the transcription of (peri-)centromeric RNAs, or the transcripts itself are important to control CENP-A levels in mature sperm.

Although Fmr1 did not show a reduction of CENP-A in mature sperm (Fig. A.5), centromeric RNAs were shown to be mis-regulated similar to the Rump mutant. Therefore, it is possible that both proteins might be involved in transcription at the centromere, cenRNA stability or splicing (Fig. 3.1), but that Rump has an additional function that involves the regulation of CENP-A protein levels in sperm. These possibilites are discussed below.

3.5.1 Transcription of cenRNAs

Rump is a highly abundant nuclear protein. In this thesis, I studied the localization of Rump in testes and found that it localizes to the nucleus of spermatocytes and post-meiotic spermatides (Fig. 2.21). In more detail, Rump localizes to euchromatic regions or the nucleoplasm. Transcription is highly active in spermatocytes and Rump has been reported to localize to loci of active transcription by binding co-transcriptionally to a subset of pre-mRNAs (Kiesler et al., 2005). Although Rump was not found at heterochromatic loci with Hp1, it cannot be excluded that a small proportion of Rump localizes to these sites. It is known that hnRNPs direct transcription by binding to promoter or enhancer sequences and recruitment of transcription factors (Geuens et al., 2016). Rump has been shown to antagonize gypsy insulators by di-

rect binding, which results in an enhanced transcription of downstream genes (King et al., 2014). As gypsy elements are present at centromeres, Rump might have a similar function and upregulate transcription of nearby transposable elements. The requirement of centromeric transcription has been shown to be important for the stable integration of CENP-A into chromatin (Bobkov et al., 2018). It has been suggested by Corless et al. (2020) in a recent review, that cenRNA might serve as a guide that recruits proteins to a specific locus in the genome. It is possible to propose that RBPs that recognize and interact with cenRNAs serve as a signal for recruiting additional proteins that are needed for centromere regulation. As mentioned above, Fmr1 and its phase separation properties might be an additional factor involved in transcription of common target RNAs. However, transcription factors were not found in the interactome of Fmr1. This suggests, that Fmr1 may create the environment to attract factors required for transcription, but may not be a direct interactor, or has different roles in cenRNA regulation. It remains to be determined whether Rump is required to recruit transcription factors in future experiments.

3.5.2 RNA stability

RNA undergoes a life cycle that begins with transcription and ends with degradation. It is the balance between newly synthesized RNA and decay of RNA that determines the level of RNA at any given time, and a disturbance of this equilibrium can lead to disease (Weskamp and Barmada, 2018). RNA stability is crucially dependent on RBPs, and while RBPs are well known for regulating mRNA stability and decay, it is less clear how RBPs affect non-coding RNAs. There is evidence that several RBPs are able to interact with lncRNAs, resulting in their stabilization (Jonas et al., 2020). However, as for mRNAs, RBPs regulate both stability and decay of lncRNAs. As a matter of fact, research has shown that the RBP HuR stabilizes some lncRNAs while promoting others' degradation, suggesting that HuR acts context-dependently (Chai et al., 2016; Yoon et al., 2013). Accordingly, it is likely that the RBPs Fmr1 and Rump contribute to the stabilization of centromeric transcripts and may destabilize others. Specific secondary structures of the transcripts or post-transcriptional modifications, such as m^6A modification may be important signals for RBPs to bind and protect from RNA degradation. Since centromeric transcripts are

transposable elements, a fine balance between their transcription and the removal is particularly important, regarding the detrimental effects transposable elements can have on genome integrity and inheritance.

3.5.3   Splicing of cenRNAs

HnRNP M, a human homolog of Rump, is part of the spliceosome complex and associates with nuclear speckles that are enriched in splicing factors (Lleres et al., 2010; Marko et al., 2010). Rump has been reported to bind to splicing enhancers on pre-mRNA targets, implicating it as a splicing factor (Kiesler et al., 2005). Interestingly, apart from Rump, Splicing factor 2 (SF2) was one of the potential interaction partners that was highly enriched in the Fmr1 mass spectrometry data (Fig. 2.14). This indicates that Rump and Fmr1 may regulate common transcripts by splicing. But are (peri-)centromeric RNAs post-transcriptionally processed and why? The results of genome-wide screens that included Drosophila and human cells revealed that splicing factors play a crucial role in cell division (Goshima et al., 2007; Kittler et al., 2007; Somma et al., 2008; Neumann et al., 2010). The role of splicing factors in centromere function was unexpected and the mechanism by which they do so remains largely unknown. Initially it was suggested that splicing is required for pre-mRNAs that encode for mitotic proteins, or that splicing may be required in the production of spindle- or centromere-associated structural RNAs. However, splicing factors were found at mitotic kinetochores and were reported to interact with non-coding centromeric RNAs (Montembault et al., 2007; Nishimura et al., 2019; Mutazono et al., 2017). The reasons for processing cenRNA are as follows; the processing of cenRNAs was stated to be required for CENP-A, CENP-C, and NCD80 to recruit to centromeres and for spindle maintenance (Grenfell et al., 2016). Another study described the necessity of spliced RNA from retrotransposon in centromeric chromatin organization via R-loops and chromatin loops (Liu et al., 2020). If we consider all the known research, it is possible that the RBPs studied here are involved in cenRNA processing. This may lead to the observed reduction in CENP-A levels in mature sperm upon Rump loss, as it remains unknown how active transcription is at this stage. It is unclear why the centromeric transcript levels are reduced rather than elevated in the Fmr1 and Rump mutants, since one would expect

that non-processed transcripts accumulate. The transcript could, however, become unstable if left unprocessed. Further experiments, such as Northern Blots are required to determine the lengths of centromeric transcripts in wt compared to the mutants. This could determine whether an involvement of RBPs in splicing cenRNAs is a possibility.
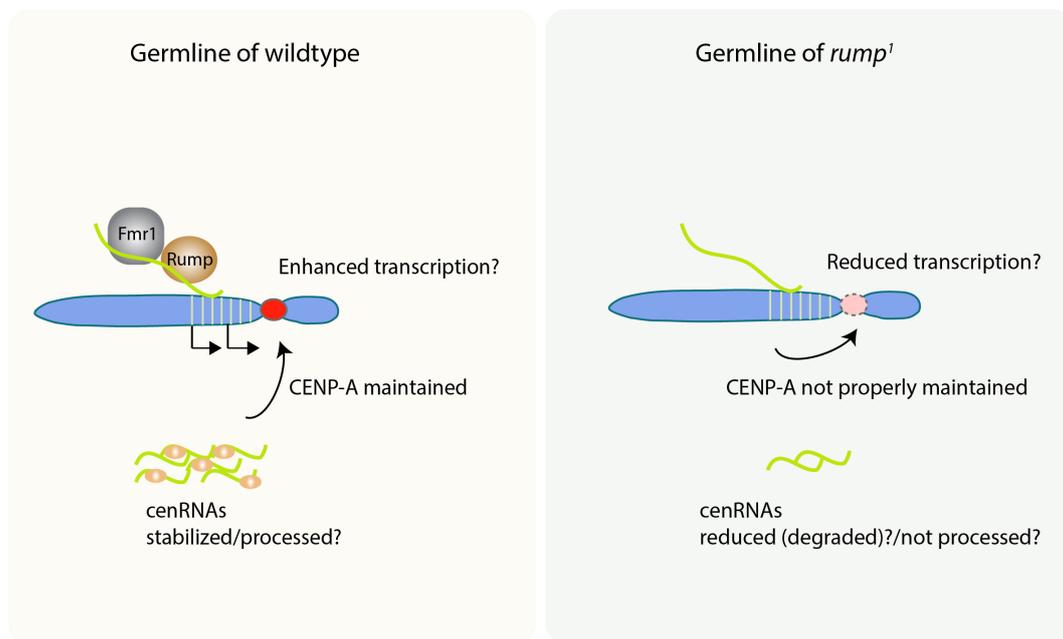
Figure 3.1: Working Model Rump and Fmr1 bind to (peri-)centromeric RNAs and modulate their levels by enhancing their transcription, or by being involved in RNA stability and processing. Lack of Rump causes reduced CENP-A integration in mature sperm, which may be a result of reduced transcription, degradation or unprocessed (peri-)centromeric RNAs.

3.6 Conclusion and Perspectives

In this thesis, the RBPs Fmr1 and Rump are shown to be important for centromere regulation in the germline of *Drosophila melanogaster*. As reported by Greenblatt and Spradling (2018), Fmr1 enhances the translation of large autism-related transcripts. Here, Fmr1 is highlighted to additionally activate the translation of proteins important for centromere/kinetochore function in ovaries, many of which are large proteins and have therefore long transcripts. Fmr1 CLIP carried out in this study revealed that most of the mRNAs that

require Fmr1 for translation are also directly bound by Fmr1. Moreover, Fmr1 loss results in mitotic defects in the early embryo, suggesting that Fmr1 functions are maternally essential. Whether this is true can be determined by fertilizing $Fmr1^4$ mutant females with wt males and studying the offspring for cell cycle defects.

Besides the subset of mRNAs that Fmr1 binds, the lncRNA SatIII and transcripts of transposable elements present at the centromere and elsewhere in the genome are shown to be target RNAs of Fmr1. Furthermore, the study of the Fmr1 interactome identified Rump as a potential new interaction partner, an RBP involved in transcription, mRNA localization, and splicing. Both, Rump and Fmr1 bind to common centromeric and non-centromeric non-coding target RNAs and modulate their levels. This suggests, that both proteins may be found in an RNP complex. The mechanism by which the RBPs regulate common RNAs remains unresolved in this thesis. Potential mechanisms are discussed above (Fig. 3.1), but further experiments are essential to determine the likelihood of the possibilities. For example, Northern Blotting can detect a potential change in RNA length and RNA FISH is useful to study changes in RNA localization.

The histone variant CENP-A is reduced in mature sperm of $rump^1$ and Zhr^1 mutants, suggesting that either the process of transcription or the transcripts itself are important for CENP-A maintenance in sperm. CENP-A is transgenerationally inherited and loss of CENP-A from mature sperm causes defects during the first mitotic cycle in the next generation (Raychaudhuri et al., 2012). To identify the biological consequences of the CENP-A reduction phenotype, it would be interesting to trial the effect of fertilizing normal oocytes with $rump^1$ or Zhr^1 mutant sperm that have decreased CENP-A levels on the offspring's cell cycle.

In summary, this thesis gives insights into centromere regulation by the RBPs Rump and Fmr1 in the germline of *Drosophila melanogaster*. The study expands our understanding of centromere biology and paves the way for further research on centromeric RNAs.

Chapter 4

Materials & Methods

4.1 Materials

All materials used in this work are generally used in the Erhardt laboratory unless otherwise specified.

Chemicals

Table 4.1: Chemicals and reagents

Chemicals	Provider
2-Propanol	Sigma
30% Acrylamid solution	AppliChem
Agarose	Sigma
Albumin (BSA)	AppliChem
Ampicillin sodium salt	AppliChem
Ammonium persulfate (APS)	Thomas Geyer
ATP	Thermo Scientific
Bacto Agar	Difco
Bacto yeast nitrogen without amino acids	Difco
Bluestain protein ladder	Thomas Geyer
b-Mercaptoethanol	AppliChem
Bromphenol blue	AppliChem
Calcium chloride dihydrate	AppliChem
Chloroform	Roth
Complete EDTA-free Protease Inhibitor Cocktail	Roche
Coomassie Brilliant Blue R-250	AppliChem
Cycloheximide	Sigma
DAPI	Sigma
DEPC	AppliChem
Dimethyl Sulfoxide (DMSO)	Sigma
Distilled water, DNase/RNase-free	Gibco
DNA ladder, 1 kb, 100 bp	NEB
DL-Dithiothreitol (DTT)	Sigma
ECL	Thermo Scientific
EDTA disodium dihydrate	Sigma
EGTA	AppliChem
Ethanol	AppliChem
Ethidiumbromide	AppliChem
Formaldehyde solution 37%	Baker
Glutathione reduced	Sigma
Glycerol	Honeywell

Glycine	Sigma
GlycoBlue	Invitrogen
Guanidine HCl	Sigma
Hepes	AppliChem
Heptane	Sigma
Hydrogen Peroxide	Thomas Geyer
IPTG	Roth
LB-Agar (Luria/Miller)	Roth
Magnesium acetate tetrahydrate	Sigma
Magnesium chloride hexahydrate	AppliChem
Methanol	Sigma
Methyl 4-hydrosybenzoate (Nipagin)	Sigma
MOPS	AppliChem
Mounting medium Aqua polymount	Polysciences
Milk powder	AppliChem
Non-fat dry milk powder	AppliChem
Nonidet P-40 (NP-40)	AppliChem
NuPAGE LDS Sample buffer 4x	Thermo Scientific
NuPAGE MOPS SDS Running Buffer (20x)	Thermo Scientific
NuPAGE Transfer Buffer (20x)	Thermo Scientific
Paraformaldehyde	AppliChem
Phenol-Chloroform:IAA (25:24:1)	AppliChem
Phenol-Chlorofrom:IAA (125:25:1)	Sigma
Phenylmethylsulfonyl fluoride (PMSF)	Sigma
Ponceau S-Solution	AppliChem
Potassium Chloride	AppliChem
RNA loading buffer (2x)	NEB
RNaseZAP	Sigma
RNasin Ribonuclease Inhibitor	Promega
SDS	AppliChem
Sodium acetate	Sigma
Sodium azide	AppliChem
Sodium chloride	Sigma
Sodium hydroxyde (NaOH)	AppliChem
ssRNA ladder	NEB
Sucrose	Sigma
TEMED	AppliChem
Tris	AppliChem
Tris-HCl	AppliChem
TRIsure	bioline
Triton X-100	Merck
Tween 20	AppliChem
Yeast tRNA	Ambion
Yeast cube	Rewe

Equipment and consumables

Table 4.2: Equipment and consumables

Equipment	Provider
Agarose gel electrophoresis tank	ZMBH workshop
kta Pure and appendant equipment	GE Healthcare
Balance	Sartorius, Kern EG
Biorupter Plus	Diagenode
Blotting paper	Bio-Rad
Cell counting system, LUNA	Logos Biosystems
Cell culture flasks	TPP
Cell culture incubator	Heraeus
Centrifuge, 5810 R	Eppendorf
Centrifuge, RC 6 Plus	Sorvall
Coplin Staining Jar	Neolab
Coverslips	Neolab
Deltavision microscope	Olympus/GE Healthcare
Dynabeads Protein G and A	Thermo Scientific
Fly cages	ZMBH workshop
Fly vials	Gosslein
Fragment Analyzer 5200	Agilent
Freezer, -20 °C	Liebherr

Freezer, -80 °C	Heraeus
Fridge, 4 °C	Liebherr
Gel Doc XR+ System	Bio-Rad
LAS-4000	Fujifilm Life Science
LightCycler 480 and appendant equipment	Roche
Magnetic stand	GE Healthcare
Microfluidizer, Emulsiflex-C5	Avestin
Micropipettes	Gilson
Microscopy slides	Thermo Scientific
Microwave	Sharp
Mini centrifuge	Nippon Genetics
Mini-PROTEAN Tetra Vertical Electrophoresis Cell	Bio-Rad
Mr. Frosty	Thermo Scientific
Multichannel Pipet	Thermo Scientific
Nanodrop A260	Nanodrop
NextSeq550 Sequencer	Illumina
Nitrocellulose membranes	Amersham Biosciences
Nunc Cryo Tubes	Sigma Aldrich
Parafilm	Bemis
PCR machine	Bio-Rad
pH-meter	Sartorius, Kern AG
Pipette tips	Greiner, Sarstedt
Plastic Pipettes, sterile	VWR
Power Supplies	Bio-Rad
Qubit^{TM} 3 Fluorometer	Invitrogen
Reaction tube, 0.2 50 ml	Sarstedt
Reaction tube, 1.5 ml, protein low binding	Eppendorf
ReadyPrep Mini Grinders	Bio-Rad
Rotation Wheel	Labinco BV
Semi dry blotting machine	ZMBH workshop
Shaker, Mini-100 Orbital-Genie	Scientific Industries
Thermomixer	Eppendorf
Tube Roller	IDL
UV crosslinker	Carl Roth
Vortex	Scientific industries
Waterbath	Memmert, Roth
Zeiss LSM 780	Zeiss

Buffers and solutions

Table 4.3: Buffers and solutions

Buffer	Ingredients
SDS Page and western blotting	
SDS-PAGE separation gel (12%)	0.375 M Tris-HCl pH 8.8
	10.5% acrylamide/bisacrylamide 30:0.8%
	0.1% SDS
	0.05% APS
	0.05% TEMED
Stacking gel	0.375 M Tris-HCl pH 6.8
	4.4% acrylamide/bisacrylamide 30:0.8% stock solution
	0.1% SDS
	0.03% APS
	0.1% TEMED
	in ddH2O
SDS running buffer	25 mM Tris
	190 mM Glycine
	0.1% SDS
	in ddH2O
4x Laemmli sample loading buffer	50 mM Tris-HCl pH=6.8
	10% Glycerol
	2% SDS
	0.5% b-Mercaptoethanol
	0.02% Bromphenolblue
	in ddH2O
Western blocking buer	1x PBS
	0.1% Tween 20

Western washing buer	5% Milk powder
	1x PBS
	0.1% Tween 20
Western transfer buffer	3.03 g Tris
	14.4 g Glycine
	200 ml Methanol
	Adjust to 1l with ddH2O
10x PBS buffer	137 mM NaCl
	2.7 mM KCl
	10 mM Na2HPO4
	1.7 mM KH2PO4
	Adjust to pH 7.5 (HCl)
Blocking buffer	1x PBS buffer
	0.1% Tween20
	5% Milk powder
Ponceau	0.2% Ponceau
	3% TCA
PBS	137 mM NaCl
	2.7 mM KCl
	10 mM Na2HPO4
	1.7 mM KH2PO4
	adjusted to pH 7.5 (HCl)

Gel electrophoresis

50x Tris-acetate-EDTA (TAE)	242 g/l Tris-HCl
	18.6 g/l EDTA
	pH 7.7 adjusted with acetic acid
	in ddH2O

Embryo collection

Embryo collection plate	3.5 g Agar
	150 ml ddH2O
	50 ml Apple juice

Immunofluorescence

Blocking solution	1x PBS buffer
	0.1% Triton X-100
	5% BSA fraction V
Permeabilisation solution	1x PBS
	0.1% Triton X-100
Paraformaldehyde (PFA) 4% in 50 ml PBS	2 g PFA
	75.7 μl 1N KOH
	5 ml 10x PBS
	45 ml H2O

Biochemistry

CLIP Lysis buffer	50 mM Tris/HCl
	150 mM NaCl
	10% glycerol
	1 mM PMSF
	0.1% NP40
	1 mM MgCl2
	0.1% SDS
	1x Protease Inhibitor
	RNasin Inhibitor
Proteinase K buffer	50 mM Tris-HCl (pH8)
	10 mM EDTA
	150 mM NaCl
	1% SDS
IP Lysis buffer	50 mM Tris/HCl
	150 mM NaCl
	10% glycerol
	1 mM PMSF
	0.1% NP40
	1 mM MgCl2
	1x Protease Inhibitor
Polysome buffer	50 mM Tris HCl, pH 7.5
	250 mM NaCl
	25 mM Magnesium acetate
	1 mM DTT
	0.1% Triton
	200 μg/ml cycloheximide
	1x Protease Inhibitor
	RNasin Inhibitor

	DNase (12.5 l/25 ml)
EX-100 buffer	10 mM Hepes
	300 mM NaCl
	1.5 mM MgCl2
	0.5 mM EGTA
	10% glycerol
	10 mM b-glycerophosphate
	1 mM DTT
	1x Protease Inhibitor
Lysis buffer (protein purification)	PBS
	150 NaCl
	1 mM DTT
	1x Protease Inhibitor
Wash buffer (protein purification)	PBS
	350 mM NaCl
	1 mM DTT
	1x Protease Inhibitor
Elution buffer (protein purification)	PBS
	30 mM Glutathione
	1 mM DTT
Storage buffer (proteins)	PBS
	Glycerol added to 10% after buffer exchange

Enzymes

Table 4.4: Enzmyes

Enzyme	Provider
Ambion RNase I	Thermo Scientific
DNase I (RNase-free	NEB
Fast AP	Thermo Scientific
Lysozyme	AppliChem
MNase	Thermo Scientific
PNK	Thermo Scientific
Proteinase K	Thermo Scientific
Q5 Polymerase	NEB
Taq Polymerase	NEB
Turbo DNase	Ambion
Restriction enzymes	NEB
RNase A	Sigma

Kits

Table 4.5: Kits

Kit	Provider
NextSeq 500/550 High Output Kit v2.5 (75 Cycles)	Illumina
NextSeq 500/550 High Output Kit v2.5 (150 Cycles)	Illumina
NucleoSpin Gel and PCR clean up	Machery-Nagel
NucleoSpin Plasmid (Mini)	Machery-Nagel
QuantiTect Reverse Transcription Kit	Qiagen
RNA Clean & Concentrator-5	Zymo Research
RNA mini Quick Spin Columns	Roche
SMARTer smRNA-Seq KIT	Takara
SuperSignal West Pico Chemiluminescent Substrate	Thermo Scientific
Universal Plus Total RNA-Seq with Nuquant KIT	Tecan

Tissue culture reagents

Table 4.6: Tissue culture reagents

Reagent	Provider
Fetal Bovine Serum (FBS)	PAN
Penicillin Streptomycin	Capricorn Scientific
Schneider's Drosophila medium	Gibco

Primary antibodies

Table 4.7: Primary antibodies

Primary antibody	Species	Application	Source
Fmr1 5B6	mouse	IP 5μg	DSHB
Fmr1 5A11	mouse	WB 1:1000	DSHB
IgG1	mouse	IP 5μg	Sigma M5284
Tubulin	mouse	WB 1:5000	Sigma T5168
CENP-A	rabbit	WB 1:2000, IF 1:1000	active motif 39713
GST	goat	WB 1:1000	GE Healthcare
Spc105R	sheep	WB 1:3000	Dr. Marcin Przewloka
BubR1	sheep	WB 1:1000	Dr. Marcin Przewloka
Rps6	rabbit	WB 1:1000	Prof. Aurelio Teleman
Rump 10C3	mouse	WB, IF 1:1000	DSHB
Rump 5G4	mouse	IP 100 μl supernatant	DSHB
Cal1	rabbit	WB 1:3000	Prof. Aaron Straight
H3	rabbit	WB 1:10000	abcam ab1791

Secondary antibodies

Table 4.8: Secondary antibodies

Secondary antibody	Species	Application	Source
IgG True Blot$^{\text{TM}}$ ULTRA	mouse	WB 1:1000	biomol
Alexa Fluor 488 goat IgG	rabbit	IF 1:500	Invitrogen
Alexa Fluor 546 goat IgG	mouse	IF 1:500	Invitrogen
Goat IgG-H&L (HRP)	mouse	1:10000	abcam
Goat IgG-H&L (HRP)	rabbit	1:10000	abcam
Donkey IgG-H&L (HRP)	goat	1:10000	abcam
Rabbit IgG-H&L(HRP)	sheep	1:10000	abcam

Primers

Table 4.9: Primers

Application	Name	Sequence
In vitro transcription	Tubulin84B fw	TAATACGACTCACTATAGGGATGCGTGAATGTATCTCT
In vitro transcription	Tubulin84B rev	TTAGTACTCCTCAGCGCCCTCACCCT
qPCR	Tubulin fw	ACGTTTGTCAAGCCTCATAGC
qPCR	Tubulin rev	TGGATAGAGATACATTCACGCATA
qPCR	Tubulin67C fw	CACCCAATTTTAGTCCAGCAA
qPCR	Tubulin67C rev	CAGGTACAGCTCCCAGCAG
qPCR	Actin5C fw	GATCTGTATGCCAACACCGT
qPCR	Actin5C rev	GCGGGGCAATGATCTTGATC
qPCR	18S rRNA fw	GGCCGTTCTTAGTTCGTGGA
qPCR	18S rRNA rev	CAACAGGTACGGCTCCACTT
qPCR	SatIII fw	AATGGAAATTAAATTTTTTGGCC
qPCR	SatIII rev	GTTTTGAGCAGCTAATTACC
qPCR	cdc27 fw	GCCCCTAGAAAGTCGCACAT
qPCR	cdc27 rev	CGTCGCTCTCCACTTCTGAG
qPCR	shtd fw	TGGCAACGTTCCTACACCAG

qPCR	shtd rev	AGCATCATAGCTCCGTGCTC
qPCR	msps fw	TCGGACGGATGTCAACTACC
qPCR	msps rev	TTCCGTTCAGGCAGAGGTTC
qPCR	Klp3A fw	ACCGCCAGAGAAACAGGATG
qPCR	Klp3A rev	TGCTCTTGCAAACGCAAGTC
qPCR	Cenp-C fw	CCAGAGTTTGAGGAGACCGT
qPCR	Cenp-C rev	TTTCCGGCTCTGGTTCTTGT
qPCR	Spc105R fw	GCAGAAGTTCGCTGCGTAAG
qPCR	Spc105R rev	CTCACGGACTTCTTGCCACT
qPCR	BubR1 fw	AGGAACTTAACGAGCGACGG
qPCR	BubR1 rev	GAAGGCTCGACCTGAACCAC

Plasmids

Table 4.10: Plasmids

Name	Source
pBluescript_MS2_10xbs_SATIII_sense	Dr. Silvana Rosic
pBSK_260 repeat	Dr. Sreemukta Acharya
pBSK_353 repeat	Dr. Sreemukta Acharya
pBSK_356 repeat	Dr. Sreemukta Acharya
pBSK_359 repeat	Dr. Sreemukta Acharya
pBSK_372 repeat	Dr. Sreemukta Acharya
pGex6p1-Fmr1	this study
pGex6p1-CenpC_C	this study
pGex6p1-Cal1_N	this study
pGex6p1-KH-RGG	this study

E.coli strains

Table 4.11: *E.coli* strains

Name	Genotype
DH5	F- Phi80dlacZ DeltaM15 Delta(lacZYA-argF)U169 deoR recA1 endA1 hsdR17(rK-mK+)phoA supE44 lambda- thi-1
BL21(DE3)	F- ompT hsdSB(rB-, mB-) gal dcm (DE3)

Drosophila fly lines

Table 4.12: Fly lines

Name	Source	Genotype
w^{1118}	Erhardt laboratory	
OreR	Erhardt laboratory	
$Fmr1^4$	Dr. Soojin Lee	+; $Fmr1^4$
$rump^1$	Bloomington 57697	st* $rump^1$ e^1
Zhr^1	Bloomington 25140	Zhr^1; +; +;
GFP-HP1	Bloomington 30561	w^{1118}; PGFP-HP13/TM6B, Tb^1

4.2 Methods

The methods listed here are standard protocols used in the Erhardt laboratory unless otherwise specified and reported similarly in different documents of the Erhardt laboratory. Buffers, solutions and materials used for experiments are listed in 4.1.

4.2.1 Molecular biology techniques

Molecular cloning

For protein purification gene constructs were cloned into the pGEX-6p-1 vector (Amersham). Proteins expressed from this vector carry an N-terminal GST-tag and expression in *E.coli* is under the control of a strong IPTG inducible tac promoter. Gibson cloning was performed by amplifying the coding regions of genes from S2 cell or *Drosophila* cDNA with primer pairs overlapping the cDNA of interest and the vector multiple cloning site. BamHI/XhoI restriction sites were part of the primer pairs in order to recover them in the final construct. Primer pairs were designed with the NEBuilder Assembly Tool (https://nebuilder.neb.com/). The Gibson assembly reaction was set up with 1x Gibson Assembly Master Mix, linearized vector and insert (1:2 or 1:3 ratio) and incubated at 50 °C for 15-30 min. The ligated plasmid was then transformed into chemically competent *E.coli* using a standard protocol. Bacterial colonies obtained were used for plasmid recovery using the NucleoSpin Plasmid kit from Machery Nagel. All inserts were confirmed by Sanger sequencing. The 1.688 Satellite repeats were cloned into pBSK vector. Therefore, the repeating unit was ordered as a gBlock from Integrated DNA Technologies. gBlocks were PCR-amplified and ligated with the linearised vector. The cloning of repeats was conducted by Dr. Sreemukta Acharya.

RNA isolation, reverse transcription and RT-qPCR

5 Ovaries or 30-50 testes were used for RNA extraction. The tissue was dissected in DEPC treated PBS+0.1% Tween20 (PBST) and collected in 1 ml PBST in Eppendorf tubes and flash frozen in liquid nitrogen after removal of PBST. Subsequently, tissue was submerged in 200 μl and stored at -80 °C until

further use. To extract RNA, tissue was homogenized using a pestle and total RNA was isolated with TRIsure reagent according to the standard protocol. cDNA for RT-qPCR was prepared using the QuantiTect Reverse Transcription Kit (Qiagen) which includes a genomic DNAse wipeout step. Transcript levels were quantified in triplicate using SYBR Green PCR Master Mix (Roche) and the oligonucleotides indicated in 4.1 on a LightCycler 480 instrument. Actin, Tubulin or 18S rRNA were used as references to normalize expression levels.

gDNA digestion of RNA extractions

Due to the high abundance of satellite repeats at DNA level, a further gDNA digestion step was performed after RNA extraction to quantify SatIII RNA by RT-qPCR. Therefore, up to 10 μg of RNA was digested with Turbo DNase for 30 min at 37 °C in a 50 μl reaction. Subsequently, the RNA was re-precipitated using Phenol-Chloroform. RNA was filled up to 300 μl with RNase-free H_2O and Phenol-Chloroform-Isoamyl was added in a 1:1 ratio. The tubes were mixed vigorously and then centrifuged for 5 min at 16.000 x g at RT. The upper phase was transferred to a new Eppendorf tube and 1 μl Glycoblue was added and precipitated over night at -20 °C. On the next day, the supernatant was removed and the pellet was washed with 1 ml 75% Ethanol. The pellet was air-dried and resuspended in 10 μl Rnase-free H_2O. The RNA was left to resuspend for 30 min on ice and then stored at -80 °C until further use.

In vitro transcription

Transcription of SatIII Sense RNA was carried out by using a pBKS plasmid containing the T7 promoter upstream of two repeating units of SatIII Sense as a template. For Tubulin84B mRNA transcription, a PCR amplicon was used as template. The amplicon was acquired by performing a PCR from cDNA using a primer pair specific to Tubulin84B with the forward primer containing the T7 promoter sequence. Upon addition of the T7 polymerase in the presence of ChromaTide 488-UTP and unlabeled nucleotides, RNA was synthesized. To get a higher yield, the plasmid was linearized with XhoI. The cutting site was immediately downstream of the probe sequence so that transcripts of unique sizes were produced. For transcription reaction the following ingredients were added together before addition of the template DNA:

- 1 μl 10 mM ATP

- 1 μl 10 mM CTP

- 1 μl 10 mM GTP

- 0.75 μl 10 mM UTP

- 2.5 μl 1 mM ChromaTide UTP

- 2.0 μl 10x Transcription buffer

- 1 unit RNase inhibitor

- Rnase-free H_2O to bring the final volume (including the template and enzyme) to 20 μl

  Mix well and add

- 500 ng linearized DNA or 50 ng PCR template

- 2-4 units T7 RNA polymerase

The reaction was mixed well by tapping the tube and centrifuged for a couple of seconds. *In vitro* transcription was carried out at 37 °C over night. On the following day 10 units of RNase-free DNase I were added and incubated for 15 min at 37 °C. The reaction was stopped by addition of 2 μl 0.2 M EDTA, pH 8.0. The RNA was subsequently purified using mini Quick Spin RNA Columns according to the manual (Roche).

Native Agarose gel electrophoresis of RNA

1x loading dye was added to the RNA sample and 1 μg RNA was loaded on 1% agarose gels in TAE buffer. For the EMSA, RNA was heated for 1 min at 95 °C and then let cool down at RT for 30 min to allow folding. The gel was run at 100 V for 5-6 h until the samples have migrated through 2/3 of the gel. For visualization, the gel was transferred to an ethidium bromide bath and incubated for 15 min. The gel was rinsed in H_2O and imaged with the GelDoc system from BioRad. In case the loaded RNA was fluorescently labelled by the incorporation of fluorescent nucleotides into the RNA during *in vitro* transcription, the RNA could be visualized using a blue LED Fluorescence light source (of 460 nm) on the ImageQuant$^{\text{TM}}$ LAS system.

Denaturing Agarose gel electrophoresis of RNA

In case a more precise RNA analysis was required, samples were loaded on denaturing gels. For a 1% gel, 100 g of agarose was melted in 84.5 ml H_2O and 10 ml of 10x MOPS was added. The solution was allowed to cool down to 60 °C and then supplemented with 5.5 ml formaldehyde, mixed and poured into a gel tray. Before loading the samples an equal volume of 2x RNA loading dye (NEB) was added and heated at 70 °C for 10 min. Prior to loading the samples were chilled on ice and spun down.

4.2.2   Cell biology techniques

S2 cell culture

Schneider 2 (S2) cells were grown under sterile conditions in medium supplemented with 10% (v/v) Fetal Bovine Serum and 200 μg/ml penicillin and streptomycin at 25 °C. Cells were split twice a week to a density of approximately 10^6 cells/ml (usually 1:10).

Thawing and freezing of S2 cells

S2 cells were kept in culture for a maximum of 20 passages and then replaced by freshly thawn ones. For thawing, cells were taken from their storage container in liquid nitrogen and quickly thawn in a 37 °C water bath. Cell suspension was added to a fresh flask containing 3 ml of medium and cells were left to settle over night. On the next day the medium was exchanged and the cells were allowed to recover for one or two weeks before experiments were conducted. For freezing, cells were grown in a T75 cell culture flask and harvested by centrifugation for 3 min at 3.000 rpm. Then 1.6 ml of the supernatant was kept (conditional medium) and the rest was discarded. The cell pellet was taken up in 1.6 ml conditional medium, 1.6 ml fresh medium and 400 μl DMSO and distributed into 4 cryotubes. The aliquots were placed in a Mr. Frosty freezing container (Thermo Fisher) with isopropanol and place at -80 °C. For long-term storage, cells were transferred to liquid nitrogen.

4.2.3 *Drosophila* techniques

Fly husbandry and breeding

Fruit flies were kept in precisely controlled conditions as follows. For optimum growth, flies were put at 25 °C with 65% humidity and a day/night cycle of twelve hours. Fly food was exchanged every 2-3 weeks except for fly stocks that were kept at 18 °C and flipped every 3-4 weeks. The food was made of 7.2% (w/v) maize, 2.4% molasses, 7.2% (w/v) malt, 0.72% (w/v) agar, 0.88% (w/v) soya, 1.464% (w/v) yeast and acid mix (1% propionic acid + 0.064% orthophosphoric acid). Female virgins were collected at the pupae stage in which the gender can be differentiated under a light microscope. Female pupae were put into a fresh vial and used for crosses after hatching. Crosses were set up with a 3:1 ratio of virgin females to males and kept at 25 °C for mating.

Grape juice plates

To prepare Agarplates with grape juice for embryo collection, 3.5 g Agar was added in 150 ml H_2O and brought to a boil to dissolve the Agar. 50 ml of grape juice was added and left to cool until 60 °C. Plates that were used for survival assay were further supplemented with 350 mg of Nipagin diluted in 2 ml 100% Ethanol. The solution was then poured into plastic Petri dishes.

Survival assay

To test the survival of Drosophila embryos, a cross of 25 homozygous mutant female flies was set up with 25 wt male flies and vice versa. Embryos were collected after 12 h. As a control, a cross with wt flies only was set up. Embryos were transferred to a new grape juice plate with Nipagin (Sigma) and sorted into groups of 10. Hatching rates were assessed after 36 h.

Testes immunofluorescence

Three pairs of testes were dissected in PBS and each pair was placed into a 5 μl drop on a siliconized coverslip. A glass slide was carefully placed on top and testes were snap-frozen. The coverslip was immediately flicked off upon removal of the liquid nitrogen. The slide was then put into a Coplin jar containing cold 100% Ethanol and left at -20 °C for 10 min. Afterwards, the

slide was left to dry for 3 min and testes were fixed with 4% PFA in PBS for 20 min. Testes were permeabilized in PBS-T (PBS+0.1% TritonX-100) for 30 min before blocking in blocking solution (5% BSA dissolved in PBS-T) for 1 h. Primary antibodies were diluted in blocking solution and added over night at 4 °C. On the following day, slides were washed 3x for 15 min each in PBS-T and then incubated with fluorescent secondary antibodies for 2h at room temperature, protected from light. The testes were washed 3x with PBS-T and then counterstained with DAPI (1:1000) in PBS for 5 min. Slides were rinsed once in PBS, excess PBS was removed and testes were mounted with Aqua Polymount medium. Finally, the slides were left to cure over night at room temperature and stored at 4 °C until imaging.

Embryo stainings

Embryo collection and fixation 2-3 h old staged embryos that were laid on grape juice plates were taken up with H_2O and a brush and transferred to a net. Embryos were dechorionated in 50 % Klorix bleach for 90 sec. After washing them the embryos were put into a scintillation vial containing 2 ml heptane and 2 ml PBS. The embryos accumulated at the interphase and 37% formaldehyde was added to a final concentration of 4%. The embryos were fixed for 20 min with constant shaking. The lower phase was removed and embryos were washed twice with the addition of 2 ml of PBS. For devitellinisation, 2 ml of methanol was added and the vial was vortexed for 30 sec. Embryos that lost their vitelline membrane sank to the bottom and were taken up and transferred to a fresh Eppendorf tube. They were washed twice with methanol and stored in methanol at -20 °C until further processing.

Immunostaining Embryos stored in methanol were rehydrated with three washes in PBS-T (PBS+0.1% Tween20) before blocking them for 1 h in 1 ml of 5% BSA in PBS-T. Primary antibody dilutions were added in PBS-T and embryos were incubated at 4 °C with constant rotation over night. On the next day, embryos were washed 4x for 15 min in PBS-T and fluoroscent secondary antibody dilutions were added in PBS-T and incubated for 2h at room temperature. Following 4x 15 min washes in PBS-T, embryos were stained with DAPI for 5 min, rinsed once with PBS and mounted using Aqua

Polymount mounting medium.

Sperm sample preparation for live imaging

Seminal vesicles from 3 day old male flies were dissected in PBS and added to a drop of 12 μl PBS and 3 μl NucBlue^{TM} Live ReadyProbes^{TM} Reagent (Invitrogen, catalog number: R37605) on a coverslip. A small amount of "Play-doh" was added to each edge of the coverslip to create a small bridge once the slide is put on top. The seminal vesicles were poked open with a fine needle to allow sperm to swim out. A glass slide was put on top carefully and gentle pressure was applied until the drop distributed evenly between the glass surfaces. The edges were sealed with nail-polish and sperm was live imaged within the next 15 min as described in 4.2.7. Sperm sample preparation for $rump^1$ and Zhr^1 live imaging was conducted by Dr. Sreemukta Acharya.

4.2.4 Biochemical techniques

Preparation of protein extracts from whole flies and fly tissue

For fly protein extracts, 4 male and female flies each were anaesthetized with CO_2 and put in an Eppendorf tube. Flies were then shock frozen in liquid nitrogen and either stored at -80 °C until further use or directly taken up in 80 μl 1x Laemmli sample loading buffer. The flies were homogenized with a pestle and then boiled at 95 °C for 5 min. For ovary or testes protein extract, 5 ovaries or 20-30 testes were dissected in PBS and treated similarly than whole fly extracts. The protein extracts could be kept at -20 °C until further analysis by SDS-PAGE and Western blot.

SDS-PAGE and Western blot

Proteins were separated on a 12% SDS-PAGE gel and transferred on a Nitrocellulose membrane using a semi-dry blotting system. The SDS gel was run at 100 V for 10 min followed by an increase of voltage to 180 for 45 min. Blotting was performed at 24 V for 1 h with 0.4 μm nitrocellulose membranes. All PAGE and Western blot equipment used was from Bio-Rad except for the CLIP experiments for which the NuPAGE gel electrophoresis system was used (Thermo Fisher Scientific). After blocking with 5% milk in PBST

(PBS+0.1% Tween 20) for 30 min at room temperature, the membrane was incubated with primary antibody in blocking solution overnight at 4 °C. The membrane was washed 3 times in PBST for 5 min and incubated 1 h at room temperature with a secondary antibody in blocking solution. Detection was done by chemiluminescence using Pierce™ ECL Western Blotting-Substrat (Thermo Fisher).

Sucrose gradient fractionation of ribosomes

For the polysome gradient, 100 female flies were fed with yeast for 3-4 days and then ovaries were dissected in ice-cold PBS and snap frozen. The tissue was stored at -80 °C until further use. The ovaries were homogenized using a pestle in 700 μl of polysome buffer containing 200 μg/ml cycloheximide and homogenates were cleared by centrifugation at 4 °C for 10 min at 16.000 x g. The supernatant was collected and used for ultracentrifugation. Therefore, the ovary extract was layered carefully onto 11 ml 10-50% sucrose gradient containing 250 mM NaCl, 25 mm MgAc and 50 mM Tris/Cl pH 7.5. Ultracentrifugation was carried out in a Beckman SW41 Ti rotor at 35.000 rpm for 4 h at 4 °C. Fractionation of the gradient was performed using a piston gradient fractionator, monitoring profiles continously at 254 nm. Hot acidic Phenol-Chloroform:IAA (125:25:1) was used to extract protein from the organic phase and RNA from the aqueous phase of each fraction. For RNA extraction 700 μl acidic Phenol-Chloroform:IAA (125:25:1) was added to 700 μl fraction. The tubes were incubated at 65 °C for 10 min and then centrifuged at 15.000 rpm at RT for 10 min. The upper phase was taken and another 700 μl acidic Phenol-Chloroform was added. After another centrifugation step the upper phase was taken and 700 μl chloroform was added followed by another centrifugation step. Subsequently 1 ml Isopropanol and 1 μl glycoblue was added to the aqueous phase and put to precipitate overnight at -20 °C. On the following day, RNA was pelleted by centrifugation for 30 min at 15.000 rpm at 4 °C. The pellet was washed once with 75% Ethanol and then left to dry and finally dissolved in RNase-free H₂O. For protein extraction 300 μl 100% Ethanol was added to 700 μl organic phase and spun for 5 min at 2000 g. The supernatant was supplemented with 1 ml Isopropanol and the proteins were left to precipitate overnight at -20 °C. On the next day the suspension was

pelleted by centrifugation for 15 min at 15.000 rpm. The protein pellet was washed with 95% Ethanol containing 0.3 M Guanidine/HCl and centrifuged for 5 min at 15.000 rpm. Pellets are left for drying and then resuspended in 1x Laemmli.

Immunoprecipitation and mass spectrometry

Immunoprecipitation For Fmr1 IP-MS experiments, 50 Ovaries were dissected per IP in PBS and put into low protein binding Eppendorf tubes. The supernatant was removed and ovaries were shock frozen in liquid nitrogen and subsequently stored at -80 °C until use. For the IP, ovaries were homogenized and lysed in 500 μl Lysis buffer with or without 0.35 μg/μl RNase A containing 1x protease inhibitor cOmplete Roche cocktail. Subsequently, the lysate was sonicated using Diagenode Bioruptor with the settings 3 cycles of 30 sec followed by 30 sec pause, "LOW" function toggled. The lysate was then cleared by centrifugation at 4 °C for 15 min. The fat layer on top was sucked off using a vacuum pump and the supernatant was added to 50 μl magentic protein G Dynabeads beads that were bound to 5 μg Fmr1-5B6 antibody and the IP was carried out for 2 h at 4 °C on a rotation wheel. Unspecific binding was controlled for using mouse IgG1 antibody (Sigma M5284) bound to magnetic beads in the same ratio. The antibodies were bound to beads prior IP by washing the beads three times in PBS-T (PBS+0.1% Tween 20) and then incubate them with antibody for 1 h at room temperature. Captured proteins were then washed 3x for 5 min with cold lysis buffer at 4 °C on a rotation wheel. The proteins were eluted by incubation in 50 μl Laemmli buffer at 95 °C for 5 min. To analyse the isolated interaction partners of Fmr1, the samples were subjected to Liquid Chromatography with tandem mass spectrometry.

LC-MS/MS analysis An in-solution tryptic digest was carried out with lysates using a modified version of the Single-Pot Solid-Phase-enhanced Sample Preparation (SP3) protocol (Hughes et al., 2014; Moggridge et al., 2018). Two biological replicates were prepared from Fmr1 and IgG IPs that were prepared from untreated and RNaseA treated lysates (n=2). The lysates were incubated with Sera-Mag Beads (Thermo Scientific) in 10 μl 15% formic acid and 30 μl of ethanol shaking for 15 min at room temperature. To remove

the SDS 4 washes with 200 μl of 70% ethanol were carried out prior to over night digestion of proteins with 0.4 μg of sequencing grade modified trypsin (Promega) in 40 μl Hepes/NaOH, pH 8.4 in the presence of 1.25 mM TCEP and 5 mM chloroacetamide (Sigma-Aldrich). On the following day, the beads were separated, washed with 10 μl of an aqueous solution of 2% DMSO and the combined eluates were dried down. 10 μl of H_2O was added and left to reconstitute peptides for 1 h at room temperature with 80 μg of TMT10plex (Thermo Scientific) (Werner et al., 2014) label reagent dissolved in 4 μl of acetonitrile. Quenching of excess TMT reagent was achieved by the addition of 4 μl of an aqueous 5% hydroxylamine solution (Sigma). Subsequently, 0.1% formic acid was added and combined peptides were purified by a reverse phase clean-up step. Peptides were then analyzed by LC-MS/MS on an Orbitrap Fusion Lumos mass spectrometer (Thermo Scientific). Acquired data were analyzed using IsobarQuant (Franken et al., 2015) and Mascot V2.4 (Matrix Science) using a reverse UniProt FASTA Drosophila melanogaster database (UP000000803) including common contaminants. LC-MS/MS analysis was carried out at the proteomics facility at EMBL in Heidelberg by Dr. Per Haberkant and Dr. Frank Stein.

Data analysis The raw output files of IsobarQuant (protein.txt files) were processed using R. Only proteins that were quantified with at least two unique peptides were considered for the analysis yielding in 430 proteins. Raw TMT reporter ion intensities (signal_sum columns) were first cleaned for batch effects using limma (Ritchie et al., 2015) and further normalized using vsn (variance stabilization normalization - (Huber et al., 2002)). For vsn normalization, different coefficients were estimated for Fmr1 and IgG conditions. Proteins were tested for differential expression using the limma package. A protein was annotated as a hit with a false discovery rate (fdr) smaller 5% and a fold-change of at least 100% and as a candidate with a fdr below 20% and a fold-change of at least 50%. Data analysis was conducted by Dr. Frank Stein from the proteomics facility at EMBL, Heidelberg.

Protein purification

Affinity purification Plasmids containing the GST-tagged protein coding sequences were transformed into BL21 bacteria. Liquid cultures were grown until the density of OD 0.6 before protein expression was induced by the addition of 0.3 mM IPTG for 16-18 h at 18 °C. Induced cultures were centrifuged at 4.000 rpm for 15 min at 4 °C. The bacterial pellet was washed once in PBS and then stored at -80 °C until further use. For purification the pellet was resuspended in 20 ml Lysis buffer and fresh lysozyme (1mg/ml) and 2.5 μl DNase 1 was added. Cells were further disrupted in two cycles with the Avestin Emulsiflex Homogenizer. The lysate was cleared by centrifugation for 30 min at 4 °C at 15.000 rpm. The supernatant was passed through a Glutathione Sepharose prepacked column (GSTrap HP column), washed with 10 ml wash buffer and eluted with a gradient of Glutathione (0-30 mM in PBS, pH = 8.0). All steps were automated using the Unicorn Software of the Äkta Pure System from GE Healthcare. The buffer of the eluted protein was then exchanged to PBS using a PD-10 desalting column (GE Healthcare). The protein was concentrated to 1 mg/ml using a 10K MWCO Ultra Centrifugal Filter (Amicon). Protein purity and stability was determined by Coomassie staining. The final protein product was supplemented with 10% glycerol and aliquoted before snapfreezing it in liquid nitrogen. Proteins were stored at -80 °C.

Gel filtration In case of protein impurity, size exclusion chromatography (SEC) was carried out using a HiLoad 16/600 Superdex 200pg (GE Healthcare) column. Therefore, the column was washed with 2 bed volumes of filtered H_2O and equilibrated in 2 bed volumes of PBS before the protein sample was injected to be loaded onto the column. Separation of molecules was monitored with the UV curve in the Unicorn software and 2 ml fractions were collected. Fractions were subsequently loaded onto SDS PAGE followed by Coomassie staining to verify the purity and presence of the protein of interest. The column was washed with PBS and H_2O in 2 bed volumes each and then equilibrated in degassed 20% Ethanol.

Phase separation experiments

Recombinant GST-KH-RGG and GST-only proteins were purified in *E.coli* as described above. Furthermore, size exlusion was carried out in order to loose impurities and to exchange the proteins into high salt (500 mM NaCl) buffer. The proteins were concentrated to 1 mg/ml and used for droplet formation assays. Therefore, 3 mM or 500 μM of protein was added into 200 μl low salt buffer (150 mM NaCl) in a live cell imaging chamber (Ibidi). For RNA-protein demixing, 1 pmol of *in vitro* transcribed 488-labelled RNAs was mixed with 500 μM of GST-KH-RGG in low salt buffer (150 mM NaCl). Droplet formation was observed using a DeltaVision Core System as described in 4.2.7.

Electrophoretic mobility shift assay (EMSA)

The 488-labelled RNA was heated to 95 °C for 1 min and then let cool down at RT for 30 min to allow folding. The following EMSA reaction was set up:

- 5x EMSA buffer 4 μl

- yeast tRNA 1 μl

- *in vitro* transcribed RNA 1 pmol

- Protein titrated per reaction (*e.g.* 0-300 μM)

- RNase-free H$_2$O to 20 μl

The reaction tubes were mixed and kept at room temperature for 30 min. Then, 4 μl EMSA loading dye was added to samples and mixed. The samples were loaded on a 1% native Agarose gel and run at 100 V until the dye front has migrated through 2/3 of the gel. RNA visualization was carried out using the ImageQuant$^{\text{TM}}$ LAS system.

Preparation of mononucleosomes and pull-down

Mononucleosomes were prepared from *Drosophila* Schneider (S2) cells. There-fore, 4×10^7 cells were resuspended in PBS containing 1x protease inhibitor cOmplete Roche cocktail and 0.3% TritonX-100. The cell suspension was put on a rotation wheel for 10 min and then centrifuged at 4 °C for 10 min at 3.000 g. The pellet was washed in PBS containing protease inhibitor followed

by another centrifugation step at 3.000 g for 5 min at 4 °C. Subsequently, the pellet was resuspended in 500 μl EX-100 buffer containing freshly added 1 mM DTT and protease inhibitor. In oder for the MNase enzyme to be active, the CaCl$_2$ concentration was adjusted to 2 mM and 300 U of MNase (Thermo Fisher) was added and left for incubation at 37 °C for 30 min. The reaction was stopped by the addition of 10 mM EGTA and samples were spun for 30 min at 20.000 rpm at 4 °C. The supernatant contained mononucleosomes which could be verified by Phenol-Chloroform DNA extraction and loading on a 2 % Agarose gel. Mononucleosomes were further used for the pull-down by incubating the lysate with GST-tagged recombinant proteins that were left on Glutathione beads for 2 h at 4 °C. Beads were washed 5x with PBS and proteins with bound mononucleosomes were eluted in 1x Laemmli buffer at 95 °C for 5 min. Western blot was carried out to test a potential interaction between CENP-A containing mononucleosomes and GST-tagged proteins. In order to test whether the interaction depended on RNAs, RNase A was added to the mononucleosome lysate prior to the pull-down.

4.2.5   Next generation sequencing related techniques

Cross-linking immunoprecipitation with high-throughput sequencing (CLIP-Seq)

UV-crosslinking ovaries   For CLIP experiments, 50 Ovaries were dissected in PBS containing RNase Inhibitor (Promega). Ovaries were then spread out on a Petri dish and UV-crosslinked using an UV-crosslinker (Carl Roth) with 0.6 J/cm^2. Ovaries were kept on ice during this process. Afterwards the tissue was transferred back into an Eppendorf tube, shock frozen and stored at -80 °C until use.

CLIP   For the IP, ovaries were lysed in 200 μl Lysis buffer containing freshly added 1x protease inhibitor cOmplete Roche cocktail and RNase Inhibitor (Promega). 0.1% SDS was added to ensure RNA fragmentation during sonication. The lysate was sonicated using the Diagenode Bioruptor with the settings 5 cycles of 30 sec followed by 30 sec at 4 °C, low. After sonication 1 μl Turbo DNase and 10 μl RNase 1 was added (1:300 dilution) and incubated at 37 °C for 3 min. Samples were immediately put on ice and left on ice for 10

min in order to let the SDS inactivate the RNase 1. Thereafter, the samples were spun for 10 min at 4 °C at maximum speed to clear the lysate. The fat layer was sucked off and 400 μl Lysis buffer without SDS was added. Thereby the SDS was diluted to an extent sufficient to prevent antibody loss from the beads during IP. An input sample was taken for Western blotting and for the RNA extraction. For the Western blot sample 30 μl was combined with 10 μl Nupage loading dye and 4 μl 1 M DTT and incubated at 1000 rpm for 10 min at 70 °C before loading on the gel. For the size-matched input (SMI) 30 μl was combined with 5 μl H_2O, 5 μl 10x FastAP buffer and 10 μl FastAP and incubated for 15 min at 37 °C before enzyme inactivation at 80 °C for 5 min. The sample was then placed on ice and further combined with 5 μl 10x PNK buffer, 10 μl ATP (10 mM), 25 μl H_2O and 10 μl PNK and incubated at 37 °C for 15 min. Subsequently, 20 μl corresponding to 1% Input were kept and 6.6 μl Nupage loading dye and 2.5 μl DTT was added. SMI was incubated at 1000 rpm for 10 min at 70 °C prior loading on the gel. The IP was carried out by incubating the lysate for 2 h at 4 °C with antibodies bound to magnetic Dynabeads protein G. 5 μg of Fmr1-5B6, Rump-5G4 and IgG1 antibody were incubated with Dynabeads for 1 h at room temperature prior IP. After the IP an "Unbound" sample was taken for Western blot from each IP. Then the IPs were washed 3x with Lysis buffer using a rotation wheel at 4 °C for 5 min. The beads were then washed twice with TBST (TBST+0.1% Tween), while beads remained attached to the magnet. For the end-repair 100 μl dephosphorylation mix (80 μl H_2O, 10 μl 10x FastAP buffer, 8 μl FastAP and 1 μl RNasin) was added and beads were incubated at 1000 rpm and 37 °C for 15 min. Subsequently, beads were washed 2x again with cold TBST and 100 μl PNK mix (70 μl H_2O, 10 μl 10mM ATP, 10 μl 10x PNK buffer, 8 μl PNK, 1 μl RNasin) was added for another incubation at 37 °C for 10 min with 1000 rpm in a thermoblock. The beads were resuspended in 1 ml TBST and 200 μl was taken for Western blot control. Supernatant was removed from the beads and 15 μl H_2O, 5 μl Nupage loading dye and 2 μl 1M DTT was added. Samples were incubate at 70 °C with 1000 rpm for 10 min. Beads were collected on the magnet and the eluate was transferred to a fresh Eppendorf tube.

NuPAGE gel electrophoresis and Western blot   For the SDS PAGE, 22 μl of the IP eluate and 30 μl of the SMI-control were loaded on a 4-12% BisTris gel (NuPAGE gel) with a protein ladder on either sides and in between. The gels were run at 180 V in 1x NuPAGE MOPS SDS Running Buffer until the dye front was at the bottom. The gel was then equilibrated with 2x NuPAGE Transfer Buffer (without methanol) with 0.02% SDS to ensure a better transfer of large proteins (>100 kDa). RNA-Protein complexes were then transferred onto 0.45 μm Nitrocellulose membrane in 1x NuPAGE Transfer Buffer containing 20% methanol for 1 h at 20 V. Using a fresh scalpel, the membrane area corresponding to the molecular weight of the protein of interest plus the area above (75-150 kDa) was excised. The membrane piece was cut into small slices and transferred to a new Eppendorf tube.

RNA release and purification   Proteins were degraded to release the RNA into solution. Therefore, 200 μl Proteinase K buffer and 50 μl Proteinase K was added to the membrane pieces and incubated for 30 min at 37 °C. Then 150 μl 6M Urea was added in Proteinase K buffer and incubated another 10 min at 37 °C. RNA was subsequently extracted by adding 400 μl acidic Phenol-Chloroform:IAA (125:25:1). The samples were mixed well by shaking and incubated 5 min at 37 °C in a thermocycler with 1200 rpm. After incubating, samples were spun at 15.000 rpm for 10 min to separate the phases. The upper phase was transferred to a new Eppendorf tube and 600 μl Isopropanol and 1 μl glycoblue was added for over night precipitation at -20 °C. On the following day, RNA was pelleted by centrifgation for 30 min at 15.000 rpm at 4 °C. The pellet was washed once with 75% Ethanol and then dissolved in 7 μl RNase-free H$_2$O after letting the pellet dry for 3 min. Acquired RNA was used for library preparation.

Library preparation   Libraries were prepared using the SMARTer smRNA-Seq KIT (Takara) according to the manufacturer's protocol applying 16 cycles of PCR amplification. Purified libraries were quantified on a Qubit Fluorometer using the QUbit dsDNA HS Assay Kit. Afterwards, the size distribution of each library was evaluated on a Fragment analyser. Size selection was carried out in order to get rid of adapter dimers and fragments larger than 1000 bp. To do so Agencourt AMPure XP Beads were used with the ratios 1.2-

0.5 (Left-Right). After size selection, the concentration of each library was assessed again and the expected size distribution of each library was verified with the Fragment analyser. The molarity was calculated with the average fragment size of each library and the measured concentration. Finally, all libraries were pooled to a final concentration of 4 nM in a final volume of 20 μl. 75 bp single end sequencing was performed by the Bukau laboratory at ZMBH in Heidelberg using the Illumina NextSeq 550 Sequencing System.

Total RNA Sequencing from *Drosophila* testes

In this study, total RNA-Seq was performed using RNA from *Drosophila* testes. For that, RNA was extracted with TriSure reagent from 50 pairs of testes as described in 4.2.1. After the RNA extraction genomic DNA was removed with the RNA Clean & Concentrator Kit (Zymo Research) following the manufacturer's instructions. RNA concentration was assessed with the Qubit Fluorometer using the RNA BR Assay Kit. Libraries were prepared using the Universal Plus Total RNA-Seq with Nuquant Kit from Tecan according to the manual with 13 cycles PCR amplification. The kit contained a targeted rRNA depletion with custom AnyDeplete probes for *Drosophila melanogaster*. The molarity of each library was determined with Nuquant Qubit assay. Libraries were pooled to a final concentration of 4 nM in a volume of 20 μl and sequencing was performed on the Illumina NextSeq 550 Sequencing System (ZMBH, Bukau laboratory) with a 150-cycle high output kit.

4.2.6 Computational analyses

All computational analyses were performed using Baden-Württemberg's high performance computing (bwHPC). Further downstream analyses and visualisation of results was carried out in R.

In general, a quality control check was conducted on all raw sequence data from this study using FastQC. Thereafter, samples were either mapped against the most recent Drosophila PacBio genome assembly from Chang et al. (2019), which includes a comprehensive repeat library and centromeric contigs or against the Drosophila assembly BDGP6. The data were mapped to the respective genomes using HISAT2. Samtools view (version 1.3.1) was used with -F 4 (to filter out unmapped reads) and -q 10 (mapping quality >10) or -q

40 to keep high quality reads. Bam files were further sorted and indexed with respect to their position in the reference genome and duplicates were removed with Picard (version 2.20.0). The steps that differ between the analysis of the CLIP-Seq and RNA-Seq data are outlined below.

CLIP-Seq analysis

Before the alignment of the CLIP-Seq data, adaptors were trimmed using cutadapt with -m 15 -u 3 -a AAAAAAAAAA input.fastq >output.fastq (Martin, 2011). The data were then further processed as described above. For the repeat analysis, multimappers were kept and read counts of input and CLIP were summarized for each type of complex repeat (*e.g.* TEs and complex satellite repeats >100) using a self-written R script. The CLIP / input ratio was calculated and further normalized by the number of mapped reads to the genome assembly. The enrichment of complex repeats in CLIP versus input was visualised by barplots or medianplots. The CLIP samples (Fmr1 and Rump) were further compared to the control IgG CLIP. The reads that were generated by mapping to the Drosophila genome BDGP6 were summarized using FeatureCounts which assigns mapped reads to genomic features. The generated count data were then further processed with the Deseq2 package in R to test differential expression analysis (Love et al., 2014). Principal component analysis (PCA) of samples was performed to visualize sample-to-sample distances. Significant enriched RNAs were visualized with a volcano plot using plotly.

Total RNA-Seq analysis

Fastq files were used for mapping with HISAT2 with the parameters –nomixed –no-discordant -q –phred33 –rna-strandness FR. Multi-reads were kept to determine the differential expression of transposable elements or complex repeats. TEtranscripts was used to perform quantification from alignment results in BAM-format with the "multi" mode (Jin et al., 2015). Differential expression analysis was carried out in R with Deseq2 with the count tables obtained from TEtranscripts. Significant changes in transcript expression were visualised with heatmaps using normalized data from Deseq2. Cufflinks (version 2.2.1) was used to compare expression levels of non-repetitive transcripts between groups.

Re-analysis of Ribosome Footprint and mRNA-Seq data

For re-analysis, raw data that were pre-filtered against Drosophila rRNA sequences and mapped to the BDGP release 6 genome were used (Greenblatt and Spradling, 2018). All non-coding genes were filtered out and data were normalized to transcripts per million reads (TPM) with the following formula:

$$TPM = \frac{FPKM}{\sum FPKM} x 10^6$$

Thereafter, follicle cell-specific genes, identified as genes whose expression was reduced by >50% in mRNA-Sequencing experiments of stage 14 de-folliculated oocytes as compared to control ovary samples, were removed from analysis. Unpaired t-test was performed for statistics and fold change was calculated by dividing Fmr1 RNAi sample averages by Control RNAi averages for both footprinting and mRNA-Seq data. Significance versus relative (fold) change was plotted as volcano plot with downregulated proteins (P <0.05 ttest, Fold change <1), upregulated proteins (P <0.05 ttest, Fold change >1) and non-significant hits (P >0.05 ttest).

Illustration

Plots that were created in R were exported as eps or svg file. All figures of this thesis were prepared and designed with Adobe Illustrator. This thesis is written in LATEX.

4.2.7   Microscopy techniques

Light microscope

All dissections and fly sortings were done using a Stereo microscope from Zeiss with an external light source.

Deltavision

A DeltaVision Core Imaging System (GE Healthcare) with softWorx v3.7.1. suite (AppliedPrecision) and a charge-coupled device camera (CoolSNAP HQ2;

Photometrics) were used for microscopy of phase separation experiments and sperm live imaging.

Phase separation imaging Droplets were imaged in a live cell chamber (Ibidi) using differential interference contrast (DIC) and Alexa filter FITC to visualize 488-labelled RNA. Acquisition was done with 100x UPlan-SApochromat (NA 1.4; Olympus). All images represent a single focal plane focused onto the surface of the chamber. Images were processed with ImageJ (NIH).

Sperm Live imaging Images were acquired using 100x UPlan-SApochromat (NA 1.4) and was carried out within 15 min after poking the seminal vesicles open. Each sperm was located and illuminated for 10 sec ith 405 nm wavelength in order to convert Dendra2 excitation wavelength from 490 (green) to 553 (red), followed by immediate imaging. The image size for each image was 512 x 512 pixels and 60 z-stacks of 0.2 μm each were obtained per image. Imaging was carried out by Dr. Sreemukta Acharya with exception of *Fmr1*[4] mutant sperm imaging.

Zeiss LSM 780 confocal microscopy

To image *Drosophila* testes and embryos the Zeiss LSM 780 confocal microscope with the Plan-Apochromat 63x/1.4 Oil objective was used. Representative images were adjusted for brightness and contrast in ImageJ (NIH).

Quantification of CENP-A in mature sperm

Images were deconvolved and projected with maximum intensity in ImageJ. Images were cropped out into an individual file of 150 x 150 pixels and images that were obtained on the same day were concatenated onto a single stack. The concatenated stack was background subtracted with a rolling-ball radius of 5 pixels. Thresholding was carried out using the MaxEntropy algorithm. CENP-A dots were selected using the wand tool and the total intensity was measured. Mutant intensity was normalised to the wt control obtained on the same day. Normalised values from all three days were pooled for the figure. Image quantification for *rump*[1] and *Zhr*[1] sperm was carried out by Dr. Sreemukta Acharya.

Bibliography

Aagaard, L., Schmid, M., Warburton, P. and Jenuwein, T. (2000), 'Mitotic phosphorylation of suv39h1, a novel component of active centromeres, coincides with transient accumulation at mammalian centromeres', *Journal of cell science* **113**(5), 817–829. 8

Abad, J., Agudo, M., Molina, I., Losada, A., Ripoll, P. and Villasante, A. (2000), 'Pericentromeric regions containing 1.688 satellite dna sequences show anti-kinetochore antibody staining in prometaphase chromosomes of drosophila melanogaster', *Molecular Genetics and Genomics* **264**, 371–377. 9

Abad, J. P., Carmena, M., Baars, S., Saunders, R., Glover, D. M., Ludena, P., Sentis, C., Tyler-Smith, C. and Villasante, A. (1992), 'Dodeca satellite: a conserved g+ c-rich satellite from the centromeric heterochromatin of drosophila melanogaster.', *Proceedings of the National Academy of Sciences* **89**(10), 4663–4667. 35

Alberti, S., Gladfelter, A. and Mittag, T. (2019), 'Considerations and challenges in studying liquid-liquid phase separation and biomolecular condensates', *Cell* **176**(3), 419–434. 13, 43

Alberts, B., Johnson, A., Lewis, J., Raff, M., Roberts, K. and Walter, P. (2002), Chromosomal dna and its packaging in the chromatin fiber, *in* 'Molecular Biology of the Cell. 4th edition', Garland Science. 2

Aldrup-MacDonald, M. E. and Sullivan, B. A. (2014), 'The past, present, and future of human centromere genomics', *Genes* **5**(1), 33–50. 7, 8

Allis, C. D. and Jenuwein, T. (2016), 'The molecular hallmarks of epigenetic control', *Nature Reviews Genetics* **17**(8), 487–500. 3

Allshire, R. C. and Karpen, G. H. (2008), 'Epigenetic regulation of centromeric chromatin: old dogs, new tricks?', *Nature Reviews Genetics* **9**(12), 923–937. 5

Almouzni, G. and Cedar, H. (2016), 'Maintenance of epigenetic information', *Cold Spring Harbor Perspectives in Biology* **8**(5), a019372. 3

102

Alonso, A., Hasson, D., Cheung, F. and Warburton, P. E. (2010), 'A paucity of heterochromatin at functional human neocentromeres', *Epigenetics & chromatin* **3**(1), 1–12. 8

Alpatov, R., Lesch, B. J., Nakamoto-Kinoshita, M., Blanco, A., Chen, S., Stützer, A., Armache, K. J., Simon, M. D., Xu, C., Ali, M. et al. (2014), 'A chromatin-dependent role of the fragile x mental retardation protein fmrp in the dna damage response', *Cell* **157**(4), 869–881. 20, 65

Andersen, P. R., Tirian, L., Vunjak, M. and Brennecke, J. (2017), 'A heterochromatin-dependent transcription machinery drives pirna expression', *Nature* **549**(7670), 54–59. 12

Anderson, B. R., Chopra, P., Suhl, J. A., Warren, S. T. and Bassell, G. J. (2016), 'Identification of consensus binding sites clarifies fmrp binding determinants', *Nucleic acids research* **44**(14), 6649–6659. 19

Arnoult, N., Van Beneden, A. and Decottignies, A. (2012), 'Telomere length regulates terra levels through increased trimethylation of telomeric h3k9 and hp1α', *Nature structural & molecular biology* **19**(9), 948–956. 13

Arunkumar, G. and Melters, D. P. (2020), 'Centromeric transcription: a conserved swiss-army knife', *Genes* **11**(8), 911. 11, 14

Ascano, M., Mukherjee, N., Bandaru, P., Miller, J. B., Nusbaum, J. D., Corcoran, D. L., Langlois, C., Munschauer, M., Dewell, S., Hafner, M. et al. (2012), 'Fmrp targets distinct mrna sequence elements to regulate protein expression', *Nature* **492**(7429), 382–386. 62

Ashley, C. T., Wilkinson, K. D., Reines, D. and Warren, S. T. (1993), 'Fmr1 protein: conserved rnp family domains and selective rna binding', *Science* **262**(5133), 563–566. 61

Azzalin, C. M. and Lingner, J. (2008), 'Telomeres: the silence is broken', *Cell cycle* **7**(9), 1161–1165. 3

Bade, D., Pauleau, A.-L., Wendler, A. and Erhardt, S. (2014), 'The e3 ligase cul3/rdx controls centromere maintenance by ubiquitylating and stabilizing cenp-a in a cal1-dependent manner', *Developmental cell* **28**(5), 508–519. 33

Bardoni, B., Sittler, A., Shen, Y. and Mandel, J. (1997), 'Analysis of domains affecting intracellular localization of the fmrp protein', *Neurobiology of disease* **4**(5), 329–336. 19

Bastock, R. and St Johnston, D. (2008), 'Drosophila oogenesis', *Current Biology* **18**(23), R1082–R1087. 21

Baumgartner, R., Stocker, H. and Hafen, E. (2013), 'The rna-binding proteins fmr1, rasputin and caprin act together with the uba protein lingerer to restrict tissue growth in drosophila melanogaster', *PLoS genetics* **9**(7), e1003598. 46

Bayes, J. J. and Malik, H. S. (2008), 'The evolution of centromeric dna sequences', *eLS* . 7

Becalska, A. N., Kim, Y. R., Belletier, N. G., Lerit, D. A., Sinsimer, K. S. and Gavis, E. R. (2011), 'Aubergine is a component of a nanos mrna localization complex', *Developmental biology* **349**(1), 46–52. 69

Beckmann, B. M., Castello, A. and Medenbach, J. (2016), 'The expanding universe of ribonucleoproteins: of novel rna-binding proteins and unconventional interactions', *Pflügers Archiv-European Journal of Physiology* **468**(6), 1029–1040. 45

Bergmann, J. H., Jakubsche, J. N., Martins, N. M., Kagansky, A., Nakano, M., Kimura, H., Kelly, D. A., Turner, B. M., Masumoto, H., Larionov, V. et al. (2012), 'Epigenetic engineering: histone h3k9 acetylation is compatible with kinetochore structure and function', *Journal of cell science* **125**(2), 411–421. 10

Bergmann, J. H., Rodríguez, M. G., Martins, N. M., Kimura, H., Kelly, D. A., Masumoto, H., Larionov, V., Jansen, L. E. and Earnshaw, W. C. (2011), 'Epigenetic engineering shows h3k4me2 is required for hjurp targeting and cenp-a assembly on a synthetic human kinetochore', *The EMBO journal* **30**(2), 328–340. 10

Beyer, A. L., Christensen, M. E., Walker, B. W. and LeStourgeon, W. M. (1977), 'Identification and characterization of the packaging proteins of core 40s hnrnp particles', *Cell* **11**(1), 127–138. 16

Bird, A. (2007), 'Perceptions of epigenetics', *Nature* **447**(7143), 396. 1

Black, B. E. and Bassett, E. A. (2008), 'The histone variant cenp-a and centromere specification', *Current opinion in cell biology* **20**(1), 91–100. 5

Black, B. E., Jansen, L. E., Maddox, P. S., Foltz, D. R., Desai, A. B., Shah, J. V. and Cleveland, D. W. (2007), 'Centromere identity maintained by nucleosomes assembled with histone h3 containing the cenp-a targeting domain', *Molecular cell* **25**(2), 309–322. 5

Blower, M. D. (2016), 'Centromeric transcription regulates aurora-b localization and activation', *Cell reports* **15**(8), 1624–1633. 15

Bobkov, G. O., Gilbert, N. and Heun, P. (2018), 'Centromere transcription allows cenp-a to transit from chromatin association to stable incorporation', *Journal of Cell Biology* **217**(6), 1957–1972. 14, 55, 71

Bosshard, M., Aprigliano, R., Gattiker, C., Palibrk, V., Markkanen, E., Backe, P. H., Pellegrino, S., Raymond, F. L., Froyen, G., Altmeyer, M. et al. (2017), 'Impaired oxidative stress response characterizes huwe1-promoted x-linked intellectual disability', *Scientific reports* **7**(1), 1–11. 43, 63

Bozzetti, M. P., Specchia, V., Cattenoz, P. B., Laneve, P., Geusa, A., Sahin, H. B., Di Tommaso, S., Friscini, A., Massari, S., Diebold, C. et al. (2015), 'The drosophila fragile x mental retardation protein participates in the pirna pathway', *Journal of cell science* **128**(11), 2070–2084. 65

Brennecke, J., Aravin, A. A., Stark, A., Dus, M., Kellis, M., Sachidanandam, R. and Hannon, G. J. (2007), 'Discrete small rna-generating loci as master regulators of transposon activity in drosophila', *Cell* **128**(6), 1089–1103. 12, 65

Brown, C. J., Ballabio, A., Rupert, J. L., Lafreniere, R. G., Grompe, M., Tonlorenzi, R. and Willard, H. F. (1991), 'A gene from the region of the human x inactivation centre is expressed exclusively from the inactive x chromosome', *Nature* **349**(6304), 38–44. 4

Brown, S. W. (1966), 'Heterochromatin: Heterochromatin provides a visible guide to suppression of gene action during development and evolution.', *Science* **151**(3709), 417–425. 3

Brown, V., Jin, P., Ceman, S., Darnell, J. C., O'Donnell, W. T., Tenenbaum, S. A., Jin, X., Feng, Y., Wilkinson, K. D., Keene, J. D. et al. (2001), 'Microarray identification of fmrp-associated brain mrnas and altered mrna translational profiles in fragile x syndrome', *Cell* **107**(4), 477–487. 19

Bulut-Karslioglu, A., Perrera, V., Scaranaro, M., De La Rosa-velazquez, I. A., Van De Nobelen, S., Shukeir, N., Popow, J., Gerle, B., Opravil, S., Pagani, M. et al. (2012), 'A transcription factor–based mechanism for mouse heterochromatin formation', *Nature structural & molecular biology* **19**(10), 1023–1030. 27

Carone, D. M., Longo, M. S., Ferreri, G. C., Hall, L., Harris, M., Shook, N., Bulazel, K. V., Carone, B. R., Obergfell, C., ONeill, M. J. et al. (2009), 'A new class of retroviral and satellite encoded small rnas emanates from mammalian centromeres', *Chromosoma* **118**(1), 113–125. 11

Castel, S. E. and Martienssen, R. A. (2013), 'Rna interference in the nucleus: roles for small rnas in transcription, epigenetics and beyond', *Nature Reviews Genetics* **14**(2), 100–112. 11

Caudron-Herger, M. and Rippe, K. (2012), 'Nuclear architecture by rna', *Current opinion in genetics & development* **22**(2), 179–187. 13

Chai, Y., Liu, J., Zhang, Z. and Liu, L. (2016), 'Hur-regulated lnc rna neat 1 stability in tumorigenesis and progression of ovarian cancer', *Cancer medicine* **5**(7), 1588–1598. 71

Chan, F. L., Marshall, O. J., Saffery, R., Kim, B. W., Earle, E., Choo, K. A. and Wong, L. H. (2012), 'Active transcription and essential role of rna polymerase ii at the centromere during mitosis', *Proceedings of the National Academy of Sciences* **109**(6), 1979–1984. 10, 14, 27

Chang, C.-H., Chavan, A., Palladino, J., Wei, X., Martins, N. M., Santinello, B., Chen, C.-C., Erceg, J., Beliveau, B. J., Wu, C.-T. et al. (2019), 'Islands of retroelements are major components of drosophila centromeres', *PLoS Biology* **17**(5), e3000241. 8, 9, 34, 35, 50, 51, 52, 58, 62, 64, 97

Chen, C.-C., Bowers, S., Lipinszki, Z., Palladino, J., Trusiak, S., Bettini, E., Rosin, L., Przewloka, M. R., Glover, D. M., ONeill, R. J. et al. (2015), 'Establishment of centromeric chromatin by the cenp-a assembly factor cal1 requires fact-mediated transcription', *Developmental cell* **34**(1), 73–84. 11

Chen, C.-C., Dechassa, M. L., Bettini, E., Ledoux, M. B., Belisario, C., Heun, P., Luger, K. and Mellone, B. G. (2014), 'Cal1 is the drosophila cenp-a assembly factor', *Journal of Cell Biology* **204**(3), 313–329. 33, 56

Chen, Y., Baker, R. E., Keith, K. C., Harris, K., Stoler, S. and Fitzgerald-Hayes, M. (2000), 'The n terminus of the centromere h3-like protein cse4p performs an essential function distinct from that of the histone fold domain', *Molecular and cellular biology* **20**(18), 7037–7048. 5

Clapier, C. R. and Cairns, B. R. (2009), 'The biology of chromatin remodeling complexes', *Annual review of biochemistry* **78**, 273–304. 4

Cléry, A. and Allain, F. (2012), 'From structure to function of rna binding domains', *RNA binding proteins* pp. 137–58. 15

Collins, C. M., Malacrida, B., Burke, C., Kiely, P. A. and Dunleavy, E. M. (2018), 'Atp synthase f 1 subunits recruited to centromeres by cenp-a are required for male meiosis', *Nature communications* **9**(1), 1–11. 54, 55

Corless, S., Höcker, S. and Erhardt, S. (2020), 'Centromeric rna and its function at and beyond centromeric chromatin', *Journal of molecular biology* **432**(15), 4257–4269. 71

Cramer, P. (2019), 'Organization and regulation of gene transcription', *Nature* **573**(7772), 45–54. 67

Darnell, J. C., Van Driesche, S. J., Zhang, C., Hung, K. Y. S., Mele, A., Fraser, C. E., Stone, E. F., Chen, C., Fak, J. J., Chi, S. W. et al. (2011), 'Fmrp stalls ribosomal translocation on mrnas linked to synaptic function and autism', *Cell* **146**(2), 247–261. 19, 62

De Renzis, S., Elemento, O., Tavazoie, S. and Wieschaus, E. F. (2007), 'Unmasking activation of the zygotic genome using chromosomal deletions in the drosophila embryo', *PLoS biology* **5**(5), e117. 21

Demarco, R. S., Eikenes, Å. H., Haglund, K. and Jones, D. L. (2014), 'Investigating spermatogenesis in drosophila melanogaster', *Methods* **68**(1), 218–227. 22, 54

Demirdizen, E., Spiller-Becker, M., Förtsch, A., Wilhelm, A., Corless, S., Bade, D., Bergner, A., Hessling, B. and Erhardt, S. (2019), 'Localization of drosophila cenp-a to non-centromeric sites depends on the nurd complex', *Nucleic acids research* **47**(22), 11589–11608. 27, 59, 60, 129, 155

Deshpande, G., Calhoun, G. and Schedl, P. (2006), 'The drosophila fragile x protein dfmr1 is required during early embryogenesis for pole cell formation and rapid nuclear division cycles', *Genetics* **174**(3), 1287–1298. 37, 61, 64, 67

Dixon, J. R., Selvaraj, S., Yue, F., Kim, A., Li, Y., Shen, Y., Hu, M., Liu, J. S. and Ren, B. (2012), 'Topological domains in mammalian genomes identified by analysis of chromatin interactions', *Nature* **485**(7398), 376–380. 2

Dreyfuss, G., Matunis, M. J., Pinol-Roma, S. and Burd, C. G. (1993), 'hn-rnp proteins and the biogenesis of mrna', *Annual review of biochemistry* **62**(1), 289–321. 16

Dunleavy, E. M., Beier, N. L., Gorgescu, W., Tang, J., Costes, S. V. and Karpen, G. H. (2012), 'The cell cycle timing of centromeric chromatin assembly in drosophila meiosis is distinct from mitosis yet requires cal1 and cenp-c', *PLoS biology* **10**(12), e1001460. 24

Dury, A. Y., El Fatimy, R., Tremblay, S., Rose, T. M., Côté, J., De Koninck, P. and Khandjian, E. W. (2013), 'Nuclear fragile x mental retardation protein is localized to cajal bodies', *PLoS genetics* **9**(10), e1003890. 67

Earnshaw, W. C. and Rothfield, N. (1985), 'Identification of a family of human centromere proteins using autoimmune sera from patients with scleroderma', *Chromosoma* **91**(3-4), 313–321. 5

Edens, B. M., Vissers, C., Su, J., Arumugam, S., Xu, Z., Shi, H., Miller, N., Ringeling, F. R., Ming, G.-l., He, C. et al. (2019), 'Fmrp modulates neural differentiation through m6a-dependent mrna nuclear export', *Cell reports* **28**(4), 845–854. 64

Edgar, B. A., Kiehle, C. P. and Schubiger, G. (1986), 'Cell cycle control by the nucleo-cytoplasmic ratio in early drosophila development', *Cell* **44**(2), 365–372. 24

Epstein, A. M., Bauer, C. R., Ho, A., Bosco, G. and Zarnescu, D. C. (2009), 'Drosophila fragile x protein controls cellular proliferation by regulating cbl levels in the ovary', *Developmental biology* **330**(1), 83–92. 30

Erdel, F. and Rippe, K. (2018), 'Formation of chromatin subcompartments by phase separation', *Biophysical journal* **114**(10), 2262–2270. 13

Erhardt, S., Mellone, B. G., Betts, C. M., Zhang, W., Karpen, G. H. and Straight, A. F. (2008), 'Genome-wide analysis reveals a cell cycle–dependent mechanism controlling centromere propagation', *The Journal of cell biology* **183**(5), 805–818. 6, 33

Fabian, L. and Brill, J. A. (2012), 'Drosophila spermiogenesis: big things come from little packages', *Spermatogenesis* **2**(3), 197–212. 23

Farrell, J. A. and O'Farrell, P. H. (2014), 'From egg to gastrula: how the cell cycle is remodeled during the drosophila mid-blastula transition', *Annual review of genetics* **48**, 269–294. 21

Filippini, A., Bonini, D., Lacoux, C., Pacini, L., Zingariello, M., Sancillo, L., Bosisio, D., Salvi, V., Mingardi, J., La Via, L. et al. (2017), 'Absence of the fragile x mental retardation protein results in defects of rna editing of neuronal mrnas in mouse', *RNA biology* **14**(11), 1580–1591. 20

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**(6669), 806–811. 12

Flemming, W. (1878), 'Zur kenntnis der zelle und ihrer teilungs-erscheinungen', *Schriften Naturwiss. Vereins Schl.-Holsk* **3**(1), 26. 1

Frank, L. and Rippe, K. (2020), 'Repetitive rnas as regulators of chromatin-associated subcompartment formation by phase separation', *Journal of molecular biology* **432**(15), 4270–4286. 13

108

Franken, H., Mathieson, T., Childs, D., Sweetman, G. M., Werner, T., Tögel, I., Doce, C., Gade, S., Bantscheff, M., Drewes, G. et al. (2015), 'Thermal proteome profiling for unbiased identification of direct and indirect drug targets using multiplexed quantitative mass spectrometry', *Nature protocols* **10**(10), 1567–1593. 91

Fu, Y.-H., Kuhl, D. P., Pizzuti, A., Pieretti, M., Sutcliffe, J. S., Richards, S., Verkert, A. J., Holden, J. J., Fenwick Jr, R. G., Warren, S. T. et al. (1991), 'Variation of the cgg repeat at the fragile x site results in genetic instability: resolution of the sherman paradox', *Cell* **67**(6), 1047–1058. 18

Fukagawa, T. and Earnshaw, W. C. (2014), 'The centromere: chromatin foundation for the kinetochore machinery', *Developmental cell* **30**(5), 496–508. 6

Gallego, J., Chou, S.-H. and Reid, B. R. (1997), 'Centromeric pyrimidine strands fold into an intercalated motif by forming a double hairpin with a novel t: G: G: T tetrad: solution structure of the d (tcccgtttcca) dimer', *Journal of molecular biology* **273**(4), 840–856. 8

Garavís, M., Escaja, N., Gabelica, V., Villasante, A. and González, C. (2015), 'Centromeric alpha-satellite dna adopts dimeric i-motif structures capped by at hoogsteen base pairs', *Chemistry-A European Journal* **21**(27), 9816–9824. 8

Garavís, M., Méndez-Lago, M., Gabelica, V., Whitehead, S. L., González, C. and Villasante, A. (2015), 'The structure of an endogenous drosophila centromere reveals the prevalence of tandemly repeated sequences able to form i-motifs', *Scientific reports* **5**(1), 1–10. 8

Gerstberger, S., Hafner, M. and Tuschl, T. (2014), 'A census of human rna-binding proteins', *Nature Reviews Genetics* **15**(12), 829–845. 15

Geuens, T., Bouhy, D. and Timmerman, V. (2016), 'The hnrnp family: insights into their role in health and disease', *Human genetics* **135**(8), 851–867. 16, 17, 70

Giansanti, M. G. and Fuller, M. T. (2012), 'What drosophila spermatocytes tell us about the mechanisms underlying cytokinesis', *Cytoskeleton* **69**(11), 869–881. 54

Giulotto, E., Raimondi, E. and Sullivan, K. F. (2017), 'The unique dna sequences underlying equine centromeres', *Centromeres and Kinetochores* pp. 337–354. 8

Goldberg, A. D., Allis, C. D. and Bernstein, E. (2007), 'Epigenetics: a landscape takes shape', *Cell* **128**(4), 635–638. 1

Goshima, G., Wollman, R., Goodwin, S. S., Zhang, N., Scholey, J. M., Vale, R. D. and Stuurman, N. (2007), 'Genes required for mitotic spindle assembly in drosophila s2 cells', *Science* **316**(5823), 417–421. 72

Greenberg, M. V. and Bourchis, D. (2019), 'The diverse roles of dna methylation in mammalian development and disease', *Nature reviews Molecular cell biology* **20**(10), 590–607. 4

Greenblatt, E. J. and Spradling, A. C. (2018), 'Fragile x mental retardation 1 gene enhances the translation of large autism-related proteins', *Science* **361**(6403), 709–712. 19, 30, 38, 39, 42, 43, 62, 63, 73, 99, 129, 155

Grenfell, A. W., Heald, R. and Strzelecka, M. (2016), 'Mitotic noncoding rna processing promotes kinetochore and spindle assembly in xenopus', *Journal of Cell Biology* **214**(2), 133–141. 14, 72

Grewal, S. I. and Jia, S. (2007), 'Heterochromatin revisited', *Nature Reviews Genetics* **8**(1), 35–46. 2, 3

Gu, T. and Elgin, S. C. (2013), 'Maternal depletion of piwi, a component of the rnai system, impacts heterochromatin formation in drosophila', *PLoS genetics* **9**(9), e1003780. 13

Hall, L. E., Mitchell, S. E. and ONeill, R. J. (2012), 'Pericentric and centromeric transcription: a perfect balance required', *Chromosome Research* **20**(5), 535–546. 3

Hamm, D. C. and Harrison, M. M. (2018), 'Regulatory principles governing the maternal-to-zygotic transition: insights from drosophila melanogaster', *Royal Society Open Biology* **8**(12), 180183. 21, 38

Hanakahi, L., Sun, H. and Maizels, N. (1999), 'High affinity interactions of nucleolin with gg-paired rdna', *Journal of Biological Chemistry* **274**(22), 15908–15912. 15

Hase, M. E., Yalamanchili, P. and Visa, N. (2006), 'The drosophila heterogeneous nuclear ribonucleoprotein m protein, hrp59, regulates alternative splicing and controls the production of its own mrna', *Journal of Biological Chemistry* **281**(51), 39135–39141. 18

Hassold, T. and Hunt, P. (2001), 'To err (meiotically) is human: the genesis of human aneuploidy', *Nature Reviews Genetics* **2**(4), 280–291. 5

Henikoff, S., Ahmad, K. and Malik, H. S. (2001), 'The centromere paradox: stable inheritance with rapidly evolving dna', *Science* **293**(5532), 1098–1102. 8

Huang, A., Kremser, L., Schuler, F., Wilflingseder, D., Lindner, H., Geley, S. and Lusser, A. (2019), 'Phosphorylation of drosophila cenp-a on serine 20 regulates protein turn-over and centromere-specific loading', *Nucleic acids research* **47**(20), 10754–10770. 67

Huber, W., Von Heydebreck, A., Sültmann, H., Poustka, A. and Vingron, M. (2002), 'Variance stabilization applied to microarray data calibration and to the quantification of differential expression', *Bioinformatics* **18**(suppl_1), S96–S104. 91

Hughes, C. S., Foehr, S., Garfield, D. A., Furlong, E. E., Steinmetz, L. M. and Krijgsveld, J. (2014), 'Ultrasensitive proteome analysis using paramagnetic bead technology', *Molecular systems biology* **10**(10), 757. 90

Huisinga, K. L., Brower-Toland, B. and Elgin, S. C. (2006), 'The contradictory definitions of heterochromatin: transcription and silencing', *Chromosoma* **115**(2), 110–122. 2

Hunter, J., Rivero-Arias, O., Angelov, A., Kim, E., Fotheringham, I. and Leal, J. (2014), 'Epidemiology of fragile x syndrome: A systematic review and meta-analysis', *American Journal of Medical Genetics Part A* **164**(7), 1648–1658. 18

Huo, X., Ji, L., Zhang, Y., Lv, P., Cao, X., Wang, Q., Yan, Z., Dong, S., Du, D., Zhang, F. et al. (2020), 'The nuclear matrix protein safb cooperates with major satellite rnas to stabilize heterochromatin architecture partially through phase separation', *Molecular cell* **77**(2), 368–383. 13, 66

Ishizuka, A., Siomi, M. C. and Siomi, H. (2002), 'A drosophila fragile x protein interacts with components of rnai and ribosomal proteins', *Genes & development* **16**(19), 2497–2508. 46

Iwasaki, Y. W., Siomi, M. C. and Siomi, H. (2015), 'Piwi-interacting rna: its biogenesis and functions', *Annual review of biochemistry* **84**, 405–433. 13

Jain, R. A. and Gavis, E. R. (2008), 'The drosophila hnrnp m homolog rumpelstiltskin regulates nanos mrna localization'. 17, 18, 47, 48, 68

Jenuwein, T. and Allis, C. D. (2001), 'Translating the histone code', *Science* **293**(5532), 1074–1080. 4

Jin, P. and Warren, S. T. (2000), 'Understanding the molecular basis of fragile x syndrome', *Human molecular genetics* **9**(6), 901–908. 18

Jin, P., Zarnescu, D. C., Ceman, S., Nakamoto, M., Mowrey, J., Jongens, T. A., Nelson, D. L., Moses, K. and Warren, S. T. (2004), 'Biochemical and genetic interaction between the fragile x mental retardation protein and the microrna pathway', *Nature neuroscience* **7**(2), 113–117. 19

Jin, Y., Tam, O. H., Paniagua, E. and Hammell, M. (2015), 'Tetranscripts: a package for including transposable elements in differential expression analysis of rna-seq datasets', *Bioinformatics* **31**(22), 3593–3599. 98

Johannisson, R., Rehder, H., Wendt, V. and Schwinger, E. (1987), 'Spermatogenesis in two patients with the fragile x syndrome', *Human genetics* **76**(2), 141–147. 20

Johnson, A. D., Richardson, E., Bachvarova, R. F. and Crother, B. I. (2011), 'Evolution of the germ line-soma relationship in vertebrate embryos.', *Reproduction (Cambridge, England)* **141**(3), 291–300. 21

Johnson, W. L., Yewdell, W. T., Bell, J. C., McNulty, S. M., Duda, Z., ONeill, R. J., Sullivan, B. A. and Straight, A. F. (2017), 'Rna-dependent stabilization of suv39h1 at constitutive heterochromatin', *Elife* **6**, e25299. 13

Johnstone, O. and Lasko, P. (2001), 'Translational regulation and rna localization in drosophila oocytes and embryos', *Annual review of genetics* **35**(1), 365–406. 21

Jonas, K., Calin, G. A. and Pichler, M. (2020), 'Rna-binding proteins as important regulators of long non-coding rnas in cancer', *International journal of molecular sciences* **21**(8), 2969. 71

Kabeche, L., Nguyen, H. D., Buisson, R. and Zou, L. (2018), 'A mitosis-specific and r loop–driven atr pathway promotes faithful chromosome segregation', *Science* **359**(6371), 108–114. 14

Kasinathan, S. and Henikoff, S. (2018), 'Non-b-form dna is enriched at centromeres', *Molecular biology and evolution* **35**(4), 949–962. 8

Khelifi, G. and Hussein, S. M. (2020), 'A new view of genome organization through rna directed interactions', *Frontiers in Cell and Developmental Biology* **8**, 517. 4

Khurana, J. S. and Theurkauf, W. (2010), 'pirnas, transposon silencing, and drosophila germline development', *Journal of Cell Biology* **191**(5), 905–913. 12

Kiesler, E., Hase, M. E., Brodin, D. and Visa, N. (2005), 'Hrp59, an hnrnp m protein in chironomus and drosophila, binds to exonic splicing enhancers and

112

is required for expression of a subset of mrnas', *The Journal of cell biology* **168**(7), 1013–1025. 17, 18, 47, 68, 70, 72

Kim, S., Yu, N.-K. and Kaang, B.-K. (2015), 'Ctcf as a multifunctional protein in genome regulation and gene expression', *Experimental & molecular medicine* **47**(6), e166–e166. 23

Kimmins, S. and Sassone-Corsi, P. (2005), 'Chromatin remodelling and epigenetic features of germ cells', *Nature* **434**(7033), 583–589. 23

King, M. R., Matzat, L. H., Dale, R. K., Lim, S. J. and Lei, E. P. (2014), 'The rna-binding protein rumpelstiltskin antagonizes gypsy chromatin insulator function in a tissue-specific manner', *Journal of cell science* **127**(13), 2956–2966. 17, 18, 71

Kittler, R., Pelletier, L., Heninger, A.-K., Slabicki, M., Theis, M., Miroslaw, L., Poser, I., Lawo, S., Grabner, H., Kozak, K. et al. (2007), 'Genome-scale rnai profiling of cell division in human tissue culture cells', *Nature cell biology* **9**(12), 1401–1412. 72

Klattenhoff, C., Xi, H., Li, C., Lee, S., Xu, J., Khurana, J. S., Zhang, F., Schultz, N., Koppetsch, B. S., Nowosielska, A. et al. (2009), 'The drosophila hp1 homolog rhino is required for transposon silencing and pirna production by dual-strand clusters', *Cell* **138**(6), 1137–1149. 12

Klockenbusch, C. and Kast, J. (2010), 'Optimization of formaldehyde cross-linking for protein interaction analysis of non-tagged integrin β1', *Journal of Biomedicine and Biotechnology* **2010**. 59

Kornberg, R. D. (1977), 'Structure of chromatin', *Annual review of biochemistry* **46**(1), 931–954. 2

Kwon, M., Morales-Mulia, S., Brust-Mascher, I., Rogers, G. C., Sharp, D. J. and Scholey, J. M. (2004), 'The chromokinesin, klp3a, drives mitotic spindle pole separation during prometaphase and anaphase and facilitates chromatid motility', *Molecular biology of the cell* **15**(1), 219–233. 39

Laggerbauer, B., Ostareck, D., Keidel, E.-M., Ostareck-Lederer, A. and Fischer, U. (2001), 'Evidence that fragile x mental retardation protein is a negative regulator of translation', *Human molecular genetics* **10**(4), 329–338. 19

Lai, A., Valdez-Sinon, A. N. and Bassell, G. J. (2020), 'Regulation of rna granules by fmrp and implications for neurological diseases', *Traffic* **21**(7), 454–462. 68

Larracuente, A. M. (2014), 'The organization and evolution of the responder satellite in species of the drosophila melanogaster group: dynamic evolution of a target of meiotic drive', *BMC evolutionary biology* **14**(1), 1–12. 9

Larson, A. G., Elnatan, D., Keenen, M. M., Trnka, M. J., Johnston, J. B., Burlingame, A. L., Agard, D. A., Redding, S. and Narlikar, G. J. (2017), 'Liquid droplet formation by hp1α suggests a role for phase separation in heterochromatin', *Nature* **547**(7662), 236–240. 13

Le, M.-H., Duricka, D. and Karpen, G. H. (1995), 'Islands of complex dna are widespread in drosophila centric heterochromatin.', *Genetics* **141**(1), 283–303. 8

Lee, A., Li, W., Xu, K., Bogert, B. A., Su, K. and Gao, F.-B. (2003), 'Control of dendritic development by the drosophila fragile x-related gene involves the small gtpase rac1'. 30, 31

Lefebvre, F. A. and Lécuyer, É. (2018), 'Flying the rna nest: Drosophila reveals novel insights into the transcriptome dynamics of early development', *Journal of developmental biology* **6**(1), 5. 22

Lewis, H. A., Musunuru, K., Jensen, K. B., Edo, C., Chen, H., Darnell, R. B. and Burley, S. K. (2000), 'Sequence-specific rna binding by a nova kh domain: implications for paraneoplastic disease and the fragile x syndrome', *Cell* **100**(3), 323–332. 15

Li, W., Klovstad, M. and Schüpbach, T. (2014), 'Repression of gurken translation by a meiotic checkpoint in drosophila oogenesis is suppressed by a reduction in the dose of eif1a', *Development* **141**(20), 3910–3921. 41

Li, X. and Fu, X.-D. (2019), 'Chromatin-associated rnas as facilitators of functional genomic interactions', *Nature Reviews Genetics* **20**(9), 503–519. 4

Li, Z., Zhang, Y., Ku, L., Wilkinson, K. D., Warren, S. T. and Feng, Y. (2001), 'The fragile x mental retardation protein inhibits translation via interacting with mrna', *Nucleic acids research* **29**(11), 2276–2283. 19

Liu, B., Li, Y., Stackpole, E. E., Novak, A., Gao, Y., Zhao, Y., Zhao, X. and Richter, J. D. (2018), 'Regulatory discrimination of mrnas by fmrp controls mouse adult neural stem cell differentiation', *Proceedings of the National Academy of Sciences* **115**(48), E11397–E11405. 19

Liu, H., Qu, Q., Warrington, R., Rice, A., Cheng, N. and Yu, H. (2015), 'Mitotic transcription installs sgo1 at centromeres to coordinate chromosome segregation', *Molecular cell* **59**(3), 426–436. 15

Liu, Y., Su, H., Zhang, J., Liu, Y., Feng, C. and Han, F. (2020), 'Back-spliced rna from retrotransposon binds to centromere and regulates centromeric chromatin loops in maize', *PLoS biology* **18**(1), e3000582. 72

Lleres, D., Denegri, M., Biggiogera, M., Ajuh, P. and Lamond, A. I. (2010), 'Direct interaction between hnrnp-m and cdc5l/plrg1 proteins affects alternative splice site choice', *EMBO reports* **11**(6), 445–451. 72

Loda, A. and Heard, E. (2019), 'Xist rna in action: Past, present, and future', *PLoS genetics* **15**(9), e1008333. 4

Losada, A. and Villasante, A. (1996), 'Autosomal location of a new subtype of 1.688 satellite dna of drosophila melanogaster', *Chromosome Research* **4**(5), 372–383. 9

Love, M. I., Huber, W. and Anders, S. (2014), 'Moderated estimation of fold change and dispersion for rna-seq data with deseq2', *Genome biology* **15**(12), 1–21. 36, 50, 98

Luger, K., Mäder, A. W., Richmond, R. K., Sargent, D. F. and Richmond, T. J. (1997), 'Crystal structure of the nucleosome core particle at 2.8 å resolution', *Nature* **389**(6648), 251–260. 2

Lunde, B. M., Moore, C. and Varani, G. (2007), 'Rna-binding proteins: modular design for efficient function', *Nature reviews Molecular cell biology* **8**(6), 479–490. 15, 16

Maehara, K., Takahashi, K. and Saitoh, S. (2010), 'Cenp-a reduction induces a p53-dependent cellular senescence response to protect cells from executing defective mitoses', *Molecular and cellular biology* **30**(9), 2090–2104. 5

Malmanche, N., Owen, S., Gegick, S., Steffensen, S., Tomkiel, J. E. and Sunkel, C. E. (2007), 'Drosophila bubr1 is essential for meiotic sister-chromatid cohesion and maintenance of synaptonemal complex', *Current Biology* **17**(17), 1489–1497. 39

Marko, M., Leichter, M., Patrinou-Georgoula, M. and Guialis, A. (2010), 'hnrnp m interacts with psf and p54nrb and co-localizes within defined nuclear structures', *Experimental cell research* **316**(3), 390–400. 72

Martienssen, R. and Moazed, D. (2015), 'Rnai and heterochromatin assembly', *Cold Spring Harbor perspectives in biology* **7**(8), a019323. 4

Martin, M. (2011), 'Cutadapt removes adapter sequences from high-throughput sequencing reads', *EMBnet. journal* **17**(1), 10–12. 98

Maurin, T., Lebrigand, K., Castagnola, S., Paquet, A., Jarjat, M., Popa, A., Grossi, M., Rage, F. and Bardoni, B. (2018), 'Hits-clip in various brain areas reveals new targets and new modalities of rna binding by fragile x mental retardation protein', *Nucleic acids research* **46**(12), 6344–6355. 19, 62

McNulty, S. M., Sullivan, L. L. and Sullivan, B. A. (2017), 'Human centromeres produce chromosome-specific and array-specific alpha satellite transcripts that are complexed with cenp-a and cenp-c', *Developmental cell* **42**(3), 226–240. 11, 14

Mellone, B. G., Grive, K. J., Shteyn, V., Bowers, S. R., Oderberg, I. and Karpen, G. H. (2011), 'Assembly of drosophila centromeric chromatin proteins during mitosis', *PLoS genetics* **7**(5), e1002068. 6, 33

Milks, K. J., Moree, B. and Straight, A. F. (2009), 'Dissection of cenp-c–directed centromere and kinetochore assembly', *Molecular biology of the cell* **20**(19), 4246–4255. 24

Moggridge, S., Sorensen, P. H., Morin, G. B. and Hughes, C. S. (2018), 'Extending the compatibility of the sp3 paramagnetic bead processing approach for proteomics', *Journal of proteome research* **17**(4), 1730–1740. 90

Mohn, F., Sienski, G., Handler, D. and Brennecke, J. (2014), 'The rhino-deadlock-cutoff complex licenses noncanonical transcription of dual-strand pirna clusters in drosophila', *Cell* **157**(6), 1364–1379. 12

Molina, O., Vargiu, G., Abad, M. A., Zhiteneva, A., Jeyaprakash, A. A., Masumoto, H., Kouprina, N., Larionov, V. and Earnshaw, W. C. (2016), 'Epigenetic engineering reveals a balance between histone modifications and transcription in kinetochore maintenance', *Nature communications* **7**(1), 1–16. 11

Monies, D., Abouelhoda, M., AlSayed, M., Alhassnan, Z., Alotaibi, M., Kayyali, H., Al-Owain, M., Shah, A., Rahbeeni, Z., Al-Muhaizea, M. A. et al. (2017), 'The landscape of genetic diseases in saudi arabia based on the first 1000 diagnostic panels and exomes', *Human genetics* **136**(8), 921–939. 43, 63

Montembault, E., Dutertre, S., Prigent, C. and Giet, R. (2007), 'Prp4 is a spindle assembly checkpoint protein required for mps1, mad1, and mad2 localization to the kinetochores', *The Journal of cell biology* **179**(4), 601–609. 72

Monzo, K., Papoulas, O., Cantin, G. T., Wang, Y., Yates, J. R. and Sisson, J. C. (2006), 'Fragile x mental retardation protein controls trailer hitch expression and cleavage furrow formation in drosophila embryos', *Proceedings of the National Academy of Sciences* **103**(48), 18160–18165. 40

Moon, W. and Hazelrigg, T. (2004), 'The drosophila microtubule-associated protein mini spindles is required for cytoplasmic microtubules in oogenesis', *Current biology* **14**(21), 1957–1961. 39

Muddashetty, R. S., Nalavadi, V. C., Gross, C., Yao, X., Xing, L., Laur, O., Warren, S. T. and Bassell, G. J. (2011), 'Reversible inhibition of psd-95 mrna translation by mir-125a, fmrp phosphorylation, and mglur signaling', *Molecular cell* **42**(5), 673–688. 19

Musacchio, A. (2015), 'The molecular biology of spindle assembly checkpoint signaling dynamics', *Current biology* **25**(20), R1002–R1018. 6

Musacchio, A. and Desai, A. (2017), 'A molecular view of kinetochore assembly and function', *Biology* **6**(1), 5. 6

Musacchio, A. and Salmon, E. D. (2007), 'The spindle-assembly checkpoint in space and time', *Nature reviews Molecular cell biology* **8**(5), 379–393. 6, 63

Mutazono, M., Morita, M., Tsukahara, C., Chinen, M., Nishioka, S., Yumikake, T., Dohke, K., Sakamoto, M., Ideue, T., Nakayama, J.-i. et al. (2017), 'The intron in centromeric noncoding rna facilitates rnai-mediated formation of heterochromatin', *PLoS genetics* **13**(2), e1006606. 72

Neumann, B., Walter, T., Hériché, J.-K., Bulkescher, J., Erfle, H., Conrad, C., Rogers, P., Poser, I., Held, M., Liebel, U. et al. (2010), 'Phenotypic profiling of the human genome by time-lapse microscopy reveals cell division genes', *Nature* **464**(7289), 721–727. 72

Neumann, P., Yan, H. and Jiang, J. (2007), 'The centromeric retrotransposons of rice are transcribed and differentially processed by rna interference', *Genetics* **176**(2), 749–761. 11

Nishibuchi, G. and Déjardin, J. (2017), 'The molecular basis of the organization of repetitive dna-containing constitutive heterochromatin in mammals', *Chromosome Research* **25**(1), 77–87. 10

Nishimura, K., Cho, Y., Tokunaga, K., Nakao, M., Tani, T. and Ideue, T. (2019), 'Deah box rna helicase dhx38 associates with satellite i noncoding rna involved in chromosome segregation', *Genes to Cells* **24**(8), 585–590. 72

Nozawa, R.-S. and Gilbert, N. (2019), 'Rna: nuclear glue for folding the genome', *Trends in cell biology* **29**(3), 201–211. 13

O'Donnell, W. T. and Warren, S. T. (2002), 'A decade of molecular studies of fragile x syndrome', *Annual review of neuroscience* **25**(1), 315–338. 20

Olins, A. L. and Olins, D. E. (1974), 'Spheroid chromatin units (ν bodies)', *Science* **183**(4122), 330–332. 2

Ontoso, D., Kauppi, L., Keeney, S. and San-Segundo, P. A. (2014), 'Dynamics of dot1l localization and h3k79 methylation during meiotic prophase i in mouse spermatocytes', *Chromosoma* **123**(1-2), 147–164. 20

Orr, B., Afonso, O., Feijão, T. and Sunkel, C. E. (2010), 'Driving chromosome segregation: lessons from the human and drosophila centromere–kinetochore machinery'. 7

Ozdilek, B. A., Thompson, V. F., Ahmed, N. S., White, C. I., Batey, R. T. and Schwartz, J. C. (2017), 'Intrinsically disordered rgg/rg domains mediate degenerate specificity in rna binding', *Nucleic acids research* **45**(13), 7984–7996. 15

Palmer, D. K., O'Day, K. and Margolis, R. L. (1990), 'The centromere specific histone cenp-a is selectively retained in discrete foci in mammalian sperm nuclei', *Chromosoma* **100**(1), 32–36. 24

Pathak, R., Mamillapalli, A., Rangaraj, N., Kumar, R., Vasanthi, D., Mishra, K. and Mishra, R. K. (2013), 'Aagag repeat rna is an essential component of nuclear matrix in drosophila', *RNA biology* **10**(4), 564–571. 61

Peters, J.-M. (2006), 'The anaphase promoting complex/cyclosome: a machine designed to destroy', *Nature reviews Molecular cell biology* **7**(9), 644–656. 6

Phan, A. T., Kuryavyi, V., Darnell, J. C., Serganov, A., Majumdar, A., Ilin, S., Raslin, T., Polonskaia, A., Chen, C., Clain, D. et al. (2011), 'Structure-function studies of fmrp rgg peptide recognition of an rna duplex-quadruplex junction', *Nature structural & molecular biology* **18**(7), 796–804. 15

Pieretti, M., Zhang, F., Fu, Y.-H., Warren, S. T., Oostra, B. A., Caskey, C. T. and Nelson, D. L. (1991), 'Absence of expression of the fmr-1 gene in fragile x syndrome', *Cell* **66**(4), 817–822. 18

Pritchard, D. K. and Schubiger, G. (1996), 'Activation of transcription in drosophila embryos is a gradual process mediated by the nucleocytoplasmic ratio.', *Genes & development* **10**(9), 1131–1142. 21

Probst, A. V., Dunleavy, E. and Almouzni, G. (2009), 'Epigenetic inheritance during the cell cycle', *Nature reviews Molecular cell biology* **10**(3), 192–206. 2

118

Przewloka, M. R., Venkei, Z., Bolanos-Garcia, V. M., Debski, J., Dadlez, M. and Glover, D. M. (2011), 'Cenp-c is a structural platform for kinetochore assembly', *Current Biology* **21**(5), 399–405. 6

Radford, S. J., Hoang, T. L., Głuszek, A. A., Ohkura, H. and McKim, K. S. (2015), 'Lateral and end-on kinetochore attachments are coordinated to achieve bi-orientation in drosophila oocytes', *PLoS genetics* **11**(10), e1005605. 39, 64

Ramat, A., Garcia-Silva, M.-R., Jahan, C., Naït-Saïdi, R., Dufourt, J., Garret, C., Chartier, A., Cremaschi, J., Patel, V., Decourcelle, M. et al. (2020), 'The piwi protein aubergine recruits eif3 to activate translation in the germ plasm', *Cell research* **30**(5), 421–435. 63

Ranjan, R., Snedeker, J. and Chen, X. (2019), 'Asymmetric centromeres differentially coordinate with mitotic machinery to ensure biased sister chromatid segregation in germline stem cells', *Cell stem cell* **25**(5), 666–681. 55

Rappsilber, J., Ryder, U., Lamond, A. I. and Mann, M. (2002), 'Large-scale proteomic analysis of the human spliceosome', *Genome research* **12**(8), 1231–1245. 16

Rathke, C., Baarends, W. M., Awe, S. and Renkawitz-Pohl, R. (2014), 'Chromatin dynamics during spermiogenesis', *Biochimica et Biophysica Acta (BBA)-Gene Regulatory Mechanisms* **1839**(3), 155–168. 23

Rathke, C., Baarends, W. M., Jayaramaiah-Raja, S., Bartkuhn, M., Renkawitz, R. and Renkawitz-Pohl, R. (2007), 'Transition from a nucleosome-based to a protamine-based chromatin configuration during spermiogenesis in drosophila', *Journal of cell science* **120**(9), 1689–1700. 23, 24

Raychaudhuri, N., Dubruille, R., Orsi, G. A., Bagheri, H. C., Loppin, B. and Lehner, C. F. (2012), 'Transgenerational propagation and quantitative maintenance of paternal centromeres depends on cid/cenp-a presence in drosophila sperm', *PLoS biology* **10**(12), e1001434. 24, 70, 74

Rieder, C. L. (1979), 'Ribonucleoprotein staining of centrioles and kinetochores in newt lung cell spindles.', *The Journal of cell biology* **80**(1), 1–9. 27

Ritchie, M. E., Phipson, B., Wu, D., Hu, Y., Law, C. W., Shi, W. and Smyth, G. K. (2015), 'limma powers differential expression analyses for rna-sequencing and microarray studies', *Nucleic acids research* **43**(7), e47–e47. 91

Rošić, S., Köhler, F. and Erhardt, S. (2014), 'Repetitive centromeric satellite rna is essential for kinetochore formation and cell division', *Journal of Cell Biology* **207**(3), 335–349. 11, 14, 37, 56, 60, 61

Russo, V. E., Martienssen, R. A. and Riggs, A. D. (1996), *Epigenetic mechanisms of gene regulation*, Cold Spring Harbor Laboratory Press. 1

Saffery, R., Sumer, H., Hassan, S., Wong, L. H., Craig, J. M., Todokoro, K., Anderson, M., Stafford, A. and Choo, K. A. (2003), 'Transcription within a functional human centromere', *Molecular cell* **12**(2), 509–516. 10

Saksouk, N., Simboeck, E. and Déjardin, J. (2015), 'Constitutive heterochromatin formation and transcription in mammals', *Epigenetics & chromatin* **8**(1), 1–17. 10

Santaguida, S. and Musacchio, A. (2009), 'The life and miracles of kinetochores', *The EMBO journal* **28**(17), 2511–2531. 6

Sawamura, K. and Yamamoto, M.-T. (1993), 'Cytogenetical localization of zygotic hybrid rescue (zhr), a drosophila melanogaster gene that rescues interspecific hybrids from embryonic lethality', *Molecular and General Genetics MGG* **239**(3), 441–449. 57

Sawamura, K., Yamamoto, M. and Watanabe, T. K. (1993), 'Hybrid lethal systems in the drosophila melanogaster species complex. ii. the zygotic hybrid rescue (zhr) gene of d. melanogaster.', *Genetics* **133**(2), 307–313. 55

Sawicka, K., Hale, C. R., Park, C. Y., Fak, J. J., Gresack, J. E., Van Driesche, S. J., Kang, J. J., Darnell, J. C. and Darnell, R. B. (2019), 'Fmrp has a cell-type-specific role in ca1 pyramidal neurons to regulate autism-related transcripts and circadian memory', *Elife* **8**, e46919. 62

Schittenhelm, R. B., Chaleckis, R. and Lehner, C. F. (2009), 'Intrakinetochore localization and essential functional domains of drosophila spc105', *The EMBO journal* **28**(16), 2374–2386. 63

Schuh, M., Lehner, C. F. and Heidmann, S. (2007), 'Incorporation of drosophila cid/cenp-a and cenp-c into centromeres during early embryonic anaphase', *Current Biology* **17**(3), 237–243. 6

Schupbach, T. and Wieschaus, E. (1986), 'Germline autonomy of maternal-effect mutations altering the embryonic body pattern of drosophila', *Developmental biology* **113**(2), 443–448. 21

Shamay-Ramot, A., Khermesh, K., Porath, H. T., Barak, M., Pinto, Y., Wachtel, C., Zilberberg, A., Lerer-Goldshtein, T., Efroni, S., Levanon, E. Y.

120

et al. (2015), 'Fmrp interacts with adar and regulates rna editing, synaptic density and locomotor activity in zebrafish', *PLoS genetics* **11**(12), e1005702. 20

Shatskikh, A. S., Kotov, A. A., Adashev, V. E., Bazylev, S. S. and Olenina, L. V. (2020), 'Functional significance of satellite dnas: Insights from drosophila', *Frontiers in Cell and Developmental Biology* **8**, 312. 9

Shimada, Y., Mohn, F. and Bühler, M. (2016), 'The rna-induced transcriptional silencing complex targets chromatin exclusively via interacting with nascent transcripts', *Genes & development* **30**(23), 2571–2580. 12

Shrestha, R. L., Ahn, G. S., Staples, M. I., Sathyan, K. M., Karpova, T. S., Foltz, D. R. and Basrai, M. A. (2017), 'Mislocalization of centromeric histone h3 variant cenp-a contributes to chromosomal instability (cin) in human cells', *Oncotarget* **8**(29), 46781. 5

Siomi, H., Siomi, M. C., Nussbaum, R. L. and Dreyfuss, G. (1993), 'The protein product of the fragile x gene, fmr1, has characteristics of an rna-binding protein', *Cell* **74**(2), 291–298. 30

Siomi, M. C., Higashijima, K., Ishizuka, A. and Siomi, H. (2002), 'Casein kinase ii phosphorylates the fragile x mental retardation protein and modulates its biological properties', *Molecular and cellular biology* **22**(24), 8438–8447. 67

Slegtenhorst-Eegdeman, K. E., De Rooij, D. G., Verhoef-Post, M., van de Kant, H. J., Bakker, C. E., Oostra, B. A., Grootegoed, J. A. and Themmen, A. P. (1998), 'Macroorchidism in fmr1 knockout mice is caused by increased sertoli cell proliferation during testicular development', *Endocrinology* **139**(1), 156–162. 20

Smurova, K. and De Wulf, P. (2018), 'Centromere and pericentromere transcription: roles and regulation in sickness and in health', *Frontiers in Genetics* **9**, 674. 10

Snead, W. T. and Gladfelter, A. S. (2019), 'The control centers of biomolecular phase separation: how membrane surfaces, ptms, and active processes regulate condensation', *Molecular cell* **76**(2), 295–305. 66

Somma, M. P., Ceprani, F., Bucciarelli, E., Naim, V., De Arcangelis, V., Piergentili, R., Palena, A., Ciapponi, L., Giansanti, M. G., Pellacani, C. et al. (2008), 'Identification of drosophila mitotic genes by combining co-expression analysis and rna interference', *PLoS genetics* **4**(7), e1000126. 72

Specchia, V., DAttis, S., Puricella, A. and Bozzetti, M. P. (2017), 'dfmr1 plays roles in small rna pathways of drosophila melanogaster', *International journal of molecular sciences* **18**(5), 1066. 30, 46

Stiff, T., ODriscoll, M., Rief, N., Iwabuchi, K., Löbrich, M. and Jeggo, P. A. (2004), 'Atm and dna-pk function redundantly to phosphorylate h2ax after exposure to ionizing radiation', *Cancer research* **64**(7), 2390–2396. 20

Strom, A. R., Emelyanov, A. V., Mir, M., Fyodorov, D. V., Darzacq, X. and Karpen, G. H. (2017), 'Phase separation drives heterochromatin domain formation', *Nature* **547**(7662), 241–245. 13, 66

Suhl, J. A., Chopra, P., Anderson, B. R., Bassell, G. J. and Warren, S. T. (2014), 'Analysis of fmrp mrna target datasets reveals highly associated mrnas mediated by g-quadruplex structures formed via clustered wgga sequences', *Human molecular genetics* **23**(20), 5479–5491. 19

Sullivan, B. A., Blower, M. D. and Karpen, G. H. (2001), 'Determining centromere identity: cyclical stories and forking paths', *Nature Reviews Genetics* **2**(8), 584–596. 7

Sullivan, B. A. and Karpen, G. H. (2004), 'Centromeric chromatin exhibits a histone modification pattern that is distinct from both euchromatin and heterochromatin', *Nature structural & molecular biology* **11**(11), 1076–1083. 10

Sullivan, K. F., Hechenberger, M. and Masri, K. (1994), 'Human cenp-a contains a histone h3 related histone fold domain that is required for targeting to the centromere.', *The Journal of cell biology* **127**(3), 581–592. 5

Sung, Y. J., Dolzhanskaya, N., Nolin, S. L., Brown, T., Currie, J. R. and Denman, R. B. (2003), 'The fragile x mental retardation protein fmrp binds elongation factor 1a mrna and negatively regulates its translation in vivo', *Journal of Biological Chemistry* **278**(18), 15669–15678. 19

Tadros, W. and Lipshitz, H. D. (2009), 'The maternal-to-zygotic transition: a play in two acts', *Development* **136**(18), 3033–3042. 21

Talbert, P. B. and Henikoff, S. (2021), 'Histone variants at a glance', *Journal of Cell Science* **134**(6), jcs244749. 4

Talbert, P. B., Kasinathan, S. and Henikoff, S. (2018), 'Simple and complex centromeric satellites in drosophila sibling species', *Genetics* **208**(3), 977–990. 9

Tanaka-Matakatsu, M., Thomas, B. J. and Du, W. (2007), 'Mutation of the apc1 homologue shattered disrupts normal eye development by disrupting g1 cell cycle arrest and progression through mitosis', *Developmental biology* **309**(2), 222–235. 39

Tang, B., Wang, T., Wan, H., Han, L., Qin, X., Zhang, Y., Wang, J., Yu, C., Berton, F., Francesconi, W. et al. (2015), 'Fmr1 deficiency promotes age-dependent alterations in the cortical synaptic proteome', *Proceedings of the National Academy of Sciences* **112**(34), E4697–E4706. 19

Taylor, A. M., Shih, J., Ha, G., Gao, G. F., Zhang, X., Berger, A. C., Schumacher, S. E., Wang, C., Hu, H., Liu, J. et al. (2018), 'Genomic and functional approaches to understanding cancer aneuploidy', *Cancer cell* **33**(4), 676–689. 5

Teo, R. Y. W., Anand, A., Sridhar, V., Okamura, K. and Kai, T. (2018), 'Heterochromatin protein 1a functions for pirna biogenesis predominantly from pericentric and telomeric regions in drosophila', *Nature communications* **9**(1), 1–12. 13

Topp, C. N., Zhong, C. X. and Dawe, R. K. (2004), 'Centromere-encoded rnas are integral components of the maize kinetochore', *Proceedings of the National Academy of Sciences* **101**(45), 15986–15991. 11

Tsang, B., Arsenault, J., Vernon, R. M., Lin, H., Sonenberg, N., Wang, L.-Y., Bah, A. and Forman-Kay, J. D. (2019), 'Phosphoregulated fmrp phase separation models activity-dependent translation through bidirectional control of mrna granule formation', *Proceedings of the National Academy of Sciences* **116**(10), 4218–4227. 43, 44, 66

Usdin, K., Hayward, B. E., Kumari, D., Lokanga, R. A., Sciascia, N. and Zhao, X.-N. (2014), 'Repeat-mediated genetic and epigenetic changes at the fmr1 locus in the fragile x-related disorders', *Frontiers in genetics* **5**, 226. 18

van Eekelen, C. A., Riemen, T. and van Venrooij, W. J. (1981), 'Specificity in the interaction of hnrna and mrna with proteins as revealed by in vivo cross linking', *FEBS letters* **130**(2), 223–226. 16

Vasudevan, S. and Steitz, J. A. (2007), 'Au-rich-element-mediated upregulation of translation by fxr1 and argonaute 2', *Cell* **128**(6), 1105–1118. 61

Villasante, A., Abad, J. P. and Méndez-Lago, M. (2007), 'Centromeres were derived from telomeres during the evolution of the eukaryotic chromosome', *Proceedings of the National Academy of Sciences* **104**(25), 10542–10547. 2

Volpe, T. A., Kidner, C., Hall, I. M., Teng, G., Grewal, S. I. and Martienssen, R. A. (2002), 'Regulation of heterochromatic silencing and histone h3 lysine-9 methylation by rnai', *science* **297**(5588), 1833–1837. 4, 10, 12

Waddington, C. H. (1942), 'The epigenotype', *Endeavour* **1**, 18–20. 1

Wahl, M. C., Will, C. L. and Lührmann, R. (2009), 'The spliceosome: design principles of a dynamic rnp machine', *cell* **136**(4), 701–718. 16

Wei, L., Liang, X.-W., Zhang, Q.-H., Li, M., Yuan, J., Li, S., Sun, S.-C., Ouyang, Y.-C., Schatten, H. and Sun, Q.-Y. (2010), 'Bubr1 is a spindle assembly checkpoint protein regulating meiotic cell cycle progression of mouse oocyte', *Cell cycle* **9**(6), 1112–1121. 64

Wei, X., Eickbush, D. G., Speece, I. and Larracuente, A. M. (2021), 'Heterochromatin-dependent transcription of satellite dnas in the drosophila melanogaster female germline', *bioRxiv* pp. 2020–08. 61

Werner, T., Sweetman, G., Savitski, M. F., Mathieson, T., Bantscheff, M. and Savitski, M. M. (2014), 'Ion coalescence of neutron encoded tmt 10-plex reporter ions', *Analytical chemistry* **86**(7), 3594–3601. 91

Weskamp, K. and Barmada, S. J. (2018), 'Rna degradation in neurodegenerative disease', *RNA Metabolism in Neurodegenerative Diseases* pp. 103–142. 71

Willard, H. F. (1985), 'Chromosome-specific organization of human alpha satellite dna.', *American journal of human genetics* **37**(3), 524. 8

Witt, E., Benjamin, S., Svetec, N. and Zhao, L. (2019), 'Testis single-cell rna-seq reveals the dynamics of de novo gene transcription and germline mutational bias in drosophila', *Elife* **8**, e47138. 24

Wong, L. H., Brettingham-Moore, K. H., Chan, L., Quach, J. M., Anderson, M. A., Northrop, E. L., Hannan, R., Saffery, R., Shaw, M. L., Williams, E. et al. (2007), 'Centromere rna is a key component for the assembly of nucleoproteins at the nucleolus and centromere', *Genome research* **17**(8), 1146–1160. 10, 14

Worpenberg, L., Paolantoni, C., Longhi, S., Mulorz, M. M., Lence, T., Wessels, H.-H., Dassi, E., Aiello, G., Sutandy, F. R., Scheibe, M. et al. (2021), 'Ythdf is a n6-methyladenosine reader that modulates fmr1 target mrna selection and restricts axonal growth in drosophila', *The EMBO journal* **40**(4), e104975. 19

124

Yadav, T., Quivy, J.-P. and Almouzni, G. (2018), 'Chromatin plasticity: a versatile landscape that underlies cell fate and identity', *Science* **361**(6409), 1332–1336. 1, 3

Yoon, J.-H., Abdelmohsen, K., Kim, J., Yang, X., Martindale, J. L., Tominaga-Yamanaka, K., White, E. J., Orjalo, A. V., Rinn, J. L., Kreft, S. G. et al. (2013), 'Scaffold function of long non-coding rna hotair in protein ubiquitination', *Nature communications* **4**(1), 1–14. 71

Zamudio, N. and Bourc'his, D. (2010), 'Transposable elements in the mammalian germline: a comfortable niche or a deadly trap?', *Heredity* **105**(1), 92–104. 61

Zhang, Y. Q., Matthies, H. J., Mancuso, J., Andrews, H. K., Woodruff III, E., Friedman, D. and Broadie, K. (2004), 'The drosophila fragile x-related gene regulates axoneme differentiation during spermatogenesis', *Developmental biology* **270**(2), 290–307. 20, 30, 61

Zhou, L.-T., Ye, S.-H., Yang, H.-X., Zhou, Y.-T., Zhao, Q.-H., Sun, W.-W., Gao, M.-M., Yi, Y.-H. and Long, Y.-S. (2017), 'A novel role of fragile x mental retardation protein in pre-mrna alternative splicing through rna-binding protein 14', *Neuroscience* **349**, 64–75. 20

Supplemental data

Supplemental figures

A

B

C
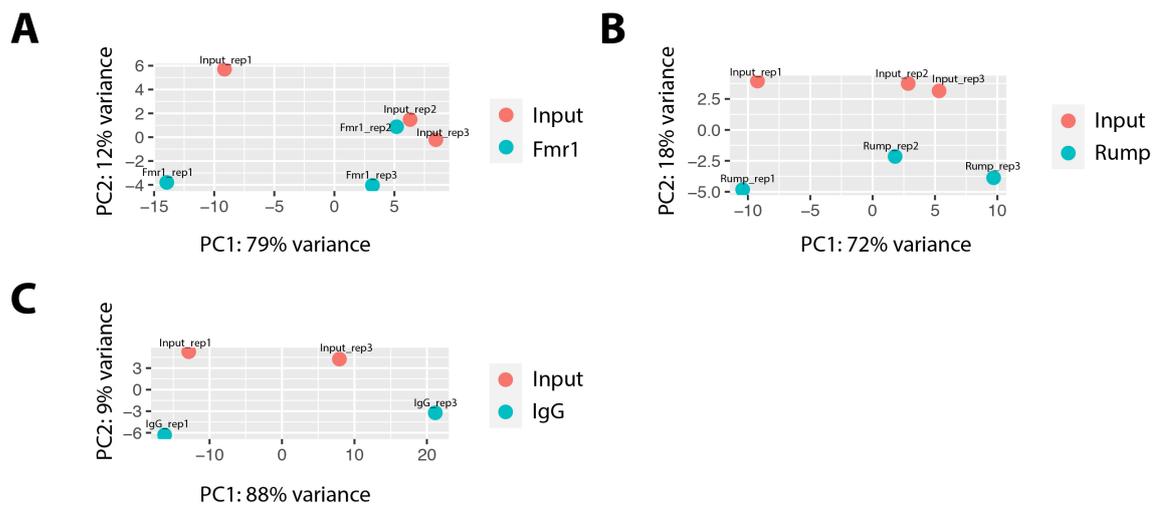
Figure A.1: CLIP experiment followed by RNA-Seq from Drosophila ovaries. A) Principal component plot, showing the samples in 2D plane spanned by their first two principal components. The second replicate of the Fmr1 CLIP experiment was removed from analysis due to low distance between input and Fmr1 IP. B) All three replicates of the Rump CLIP were used for analysis because of high variances between input and IPs. C) PCA plot showing the samples of IgG Clip and input. Both replicates were used to determine specificity of enriched Rump and Fmr1 targets.
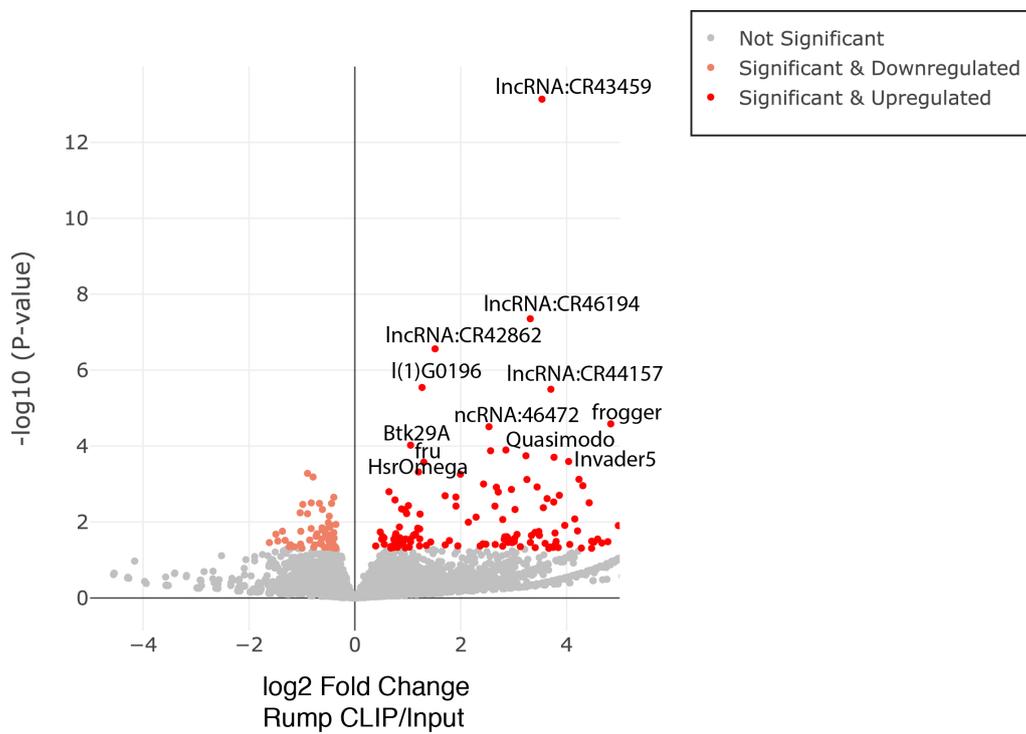
Figure A.2: Volcano plot depicting Rump RNA targets. Rump CLIP data were aligned to the *Drosophila* genome BDGP6 which excludes the centromeric contigs but still contains transposable elements that are present outside of centromeres. Top hits are marked.
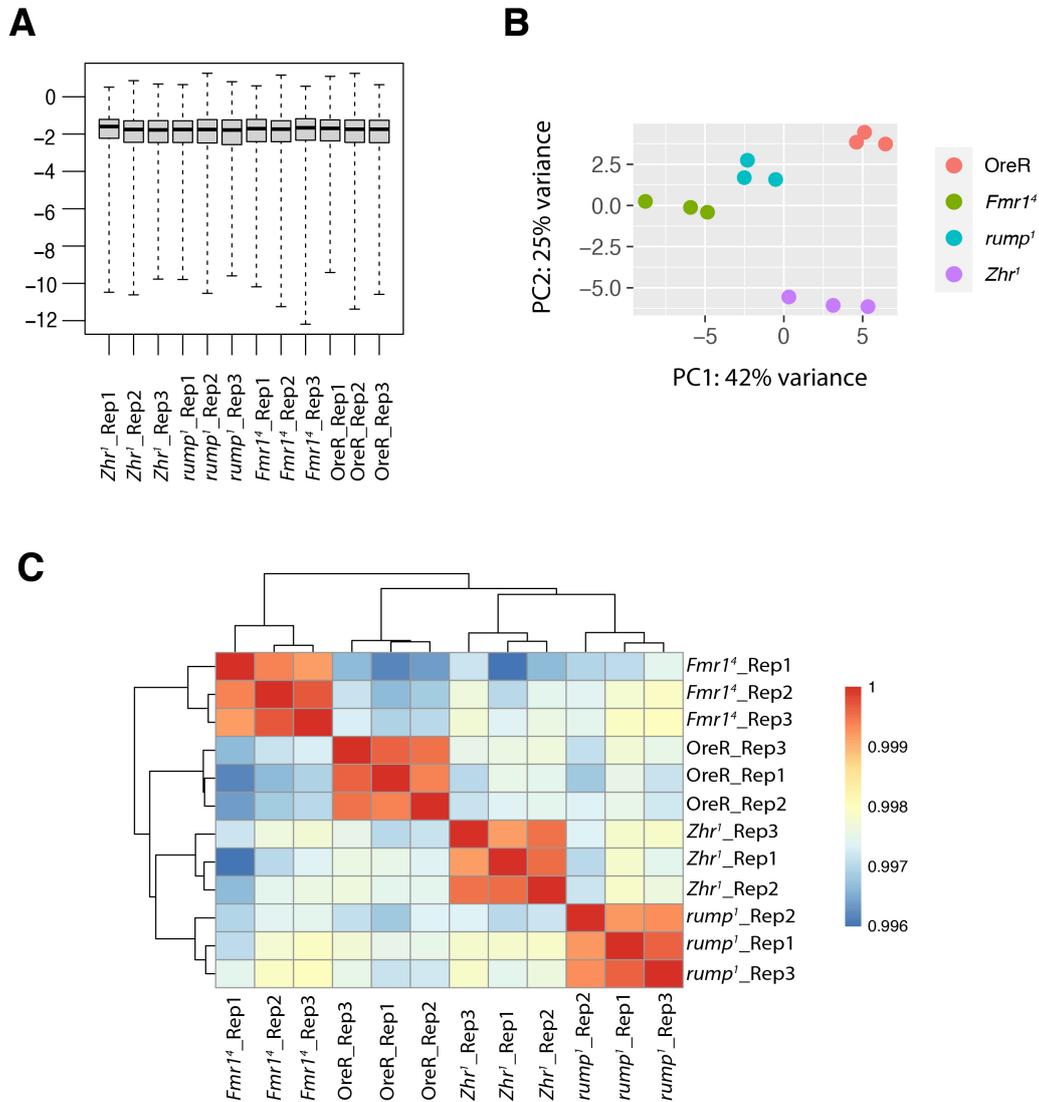
Figure A.3: Total RNA Sequencing from *Drosophila* testes. A) Outliers were assessed for every gene for every sample by Cook's distance. Large values of Cook's distance (y-axes) indicate an outlier count. No sample is significantly higher than another suggesting that the number of outliers is very limited. B) Principal component plot, showing the samples in 2D plane spanned by their first two principal components. PC1 shows the difference between groups, PC2 the variability between replicates within one group. C) Heatmap showing hierarchical clustering based on the sample distances. Replicates of the same groups cluster together, showing their high similarity, while the groups are distant from each other. All quality controls were performed with Deseq2 in R.

128
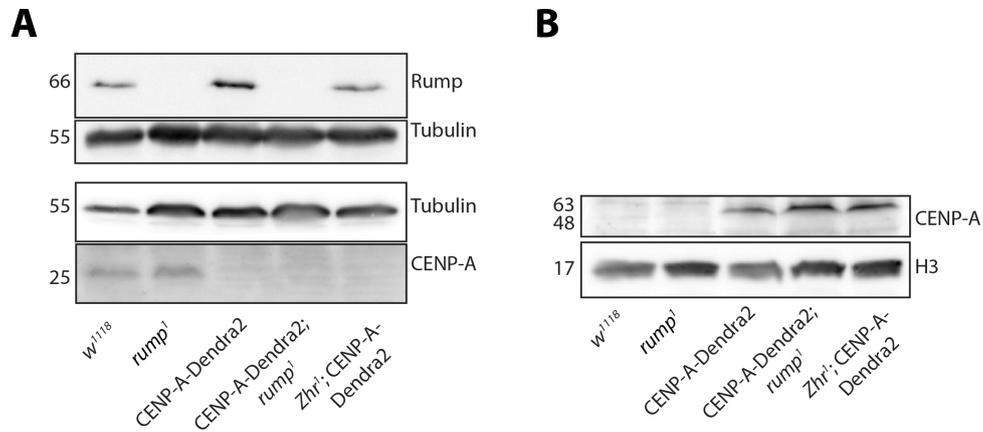
Figure A.4: CENP-A is not reduced in whole male flies. A) Western blot showing endogeneous untagged CENP-A levels in wt and $rump^1$ mutant. Rump null allele was confirmed by immunoblotting with anti-Rump antibody. B) Western blot showing endogenous Dendra2-tagged CENP-A in wt and mutant flylines that were crossed into CENP-A-Dendra2 background.
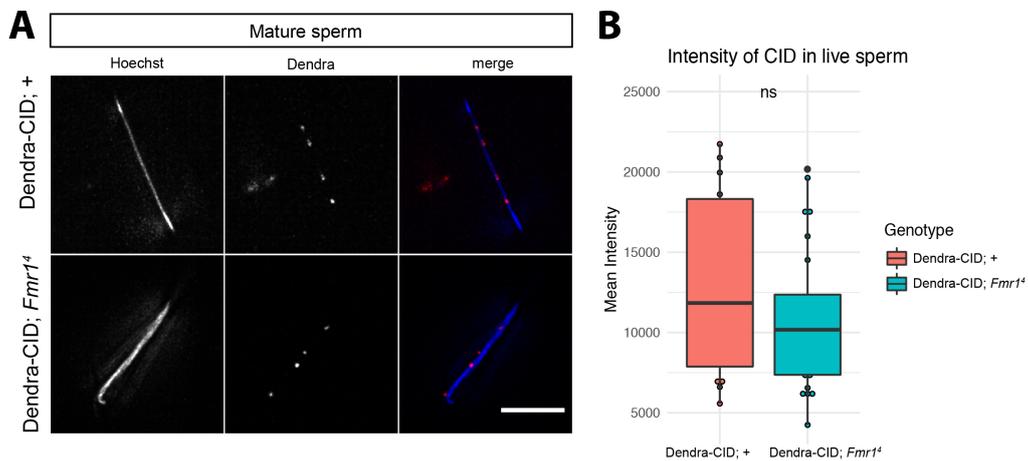
Figure A.5: CENP-A levels are unaffected in $Fmr1^4$ mutant males. A) CENP-A-Dendra2 was live imaged in mature sperm from wt and $Fmr1^4$ mutant sperm. Scale bar 2 μm. B) Total CENP-A levels in sperm were quantified and statistical significance was assessed using Mann-Whitney test. The CENP-A levels were not significantly different in wt and $Fmr1^4$.

Supplemental tables

Table A.1: RNA-binding proteins enriched in CENP-A-eGFP mass spectrometry data (Demirdizen et al., 2019)

Gene name	LQF intensity CENP-A	LQF intensity GFP-only	MS/MS count	Peptides
eIF4G	35081000	10952000	162	58
eIF3-S10	23843000	9585400	112	43
RpL5	80272000	34676000	106	23
pAbp	33402000	11699000	105	29
mask	5925600	1764400	95	47
eIF3-S8	17666000	6394100	89	34
Top2	34501000	2692000	85	43
RpL7	52554000	19468000	83	25
bel	16478000	7389300	74	31
Fmr1	19051000	8375100	71	24
larp	11351000	2693800	63	28
RpS9	47872000	22492000	59	22
AGO2	5515500	1997600	52	19
hyd	8820900	928780	51	38
eIF3-S9	8186400	2835200	50	24
eIF-3p66	14337000	6349500	50	17
RpS11	45038000	13521000	49	19
Rm62	12240000	5414500	48	23
RpS2	28519000	13957000	46	13
RpL9	846930	390000	42	12
RpL8	38191000	17943000	42	13
mod	7762200	2716400	39	14
sqd	7252200	1205200	36	13
Srp72	3700400	694790	34	19
Ssrp	7091700	1427600	30	15
Fib	7026200	2533400	28	10
Trip1	4059200	1556600	26	11
eIF5B	2538000	551040	25	13
eIF3-S6	5711100	2610600	24	12
eIF5	1462700	519570	24	15
Imp	3213400	1119200	22	14
Rbp2	3720500	1857900	22	10
eIF-3p40	4224500	1872900	22	10
CG2263	1279500	560810	20	10
rin	2758300	1336200	19	9
Nsun2	1770800	649460	17	10
bl	1616900	755050	14	8
Pep	2998800	953320	12	6
La	1689100	782840	12	7
Su(var)205	2238800	915380	9	5
Hrb87F	1701500	743720	5	4
eIF4AIII	1226300	547580	5	7
CG4364	533670	264160	5	2
Srp19	846930	390000	4	3
Ref1	710500	352890	4	3

Table A.2: Mitotic genes down-regulated in the Ribo-Seq data from Greenblatt and Spradling (2018)

FBgn	gene	gene length	pvalue	FC
FBgn0004391	shtd	6897	7.58E-06	0.517574281
FBgn0027948	msps	9874	7.98E-06	0.64805735
FBgn0012058	Cdc27	3859	0.000128922	0.557047697
FBgn0026181	Rok	13402	0.000161083	0.794011623
FBgn0264326	DNApol-epsilon255	7035	0.000192375	0.676188483
FBgn0263855	BubR1	5286	0.000246954	0.683674037
FBgn0011606	Klp3A	4894	0.000352321	0.618982592
FBgn0037025	Spc105R	6443	0.000358258	0.563472574
FBgn0025802	Sbf	7655	0.000663668	0.675378841

FBgn0041171	ago	7963	0.000765166	0.65512763
FBgn0050085	Rif1	4838	0.000827728	0.738647123
FBgn0001108	DCTN1-p150	5269	0.000882633	0.727467035
FBgn0010355	Taf1	9159	0.00094899	0.684201581
FBgn0020306	dom	18417	0.001033174	0.659981704
FBgn0040208	Kat60	4920	0.001261416	0.804160314
FBgn0261385	scra	5123	0.001324976	0.687163996
FBgn0003268	rod	7293	0.00147376	0.625700696
FBgn0040290	RecQ4	5338	0.001529957	0.617521087
FBgn0030268	Klp10A	7756	0.001673083	0.712576078
FBgn0025742	mtm	2284	0.002050085	0.869101279
FBgn0045035	tefu	10024	0.0020729	0.498396428
FBgn0033845	mars	5390	0.002520041	0.698627239
FBgn0039680	Cap-D2	4606	0.00329666	0.72439459
FBgn0027518	Wdr24	3120	0.003647323	0.59488763
FBgn0035227	Iml1	10620	0.003931406	0.634597953
FBgn0000140	asp	6664	0.004051427	0.610786263
FBgn0013765	cnn	11161	0.004082846	0.599634552

Table A.3: Mass spectrometry data enriched hits of Fmr1 IP from ovaries

gene	logFC	pvalue	comparison
Capr	5.563581494	1.00E-07	Fmr - IgG
Fmr1	5.595810233	1.46E-07	Fmr - IgG
AGO2	4.756789931	1.54E-07	Fmr - IgG
lig	4.494896865	2.70E-07	Fmr - IgG
SF2	5.04597858	2.98E-07	Fmr - IgG
Cen	5.095290236	4.18E-07	Fmr - IgG
rin	5.002689262	4.26E-07	Fmr - IgG
pcm	4.400294036	4.33E-07	Fmr - IgG
RpS13	4.322329609	4.38E-07	Fmr - IgG
Atx2	4.101792649	4.59E-07	Fmr - IgG
larp	4.092897613	5.16E-07	Fmr - IgG
tral	4.615889413	5.56E-07	Fmr - IgG
CG30122	4.14024022	5.63E-07	Fmr - IgG
rump	3.994311209	5.92E-07	Fmr - IgG
tyf	3.886987056	6.65E-07	Fmr - IgG
Tdrd3	4.440209113	6.90E-07	Fmr - IgG
Nop60B	3.852939371	7.14E-07	Fmr - IgG
CG5787	5.133889109	7.18E-07	Fmr - IgG
RpS15	4.202549811	7.49E-07	Fmr - IgG
RpS19a	3.284248425	7.70E-07	Fmr - IgG
CG5726	5.217882465	8.09E-07	Fmr - IgG
qkr58E-1	4.173040166	8.65E-07	Fmr - IgG
exu	4.784954647	8.95E-07	Fmr - IgG
RpS5	3.115505669	9.22E-07	Fmr - IgG
Msr-110	3.274568022	1.04E-06	Fmr - IgG
stau	3.427914499	1.05E-06	Fmr - IgG
RpS9	3.92851854	1.11E-06	Fmr - IgG
nocte	4.730064688	1.24E-06	Fmr - IgG
lost	4.784638918	1.27E-06	Fmr - IgG
CG8915	5.257038252	1.32E-06	Fmr - IgG
Pep	3.864130706	1.33E-06	Fmr - IgG
nonA	4.571799863	1.34E-06	Fmr - IgG
Rm62	3.665225801	1.37E-06	Fmr - IgG
Imp	4.388233324	1.39E-06	Fmr - IgG
SRPK	5.039268368	1.79E-06	Fmr - IgG
CG30497	4.262819998	1.80E-06	Fmr - IgG
Nop56	3.182898039	1.84E-06	Fmr - IgG
RpL31	2.883148589	1.94E-06	Fmr - IgG
BicC	3.105756429	2.03E-06	Fmr - IgG
RpS28b	3.19985075	2.17E-06	Fmr - IgG
RpS16	3.433530872	2.33E-06	Fmr - IgG
CG13096	3.442257814	2.37E-06	Fmr - IgG
eIF-4E	4.725456156	2.52E-06	Fmr - IgG
lark	2.996718303	2.56E-06	Fmr - IgG
RpS17	2.655391683	2.61E-06	Fmr - IgG

RpS4	3.399309322	2.71E-06	Fmr - IgG
RpS25	4.232329452	2.93E-06	Fmr - IgG
Ref1	3.912935145	3.10E-06	Fmr - IgG
Upf1	3.535080025	3.14E-06	Fmr - IgG
Edc3	3.296968225	3.14E-06	Fmr - IgG
RpS3A	3.966729334	3.18E-06	Fmr - IgG
cup	4.333407179	3.27E-06	Fmr - IgG
osk	4.062830825	3.45E-06	Fmr - IgG
RpS3	2.939372504	3.51E-06	Fmr - IgG
me31B	4.623636303	3.71E-06	Fmr - IgG
La	3.664991458	3.89E-06	Fmr - IgG
nop5	2.986146377	4.07E-06	Fmr - IgG
piwi	4.069001719	4.12E-06	Fmr - IgG
Rbp1	4.043724386	4.17E-06	Fmr - IgG
btz	6.268742612	4.58E-06	Fmr - IgG
RpS20	2.566781591	4.68E-06	Fmr - IgG
DCP1	4.065684761	4.81E-06	Fmr - IgG
Srp72	3.519481439	4.94E-06	Fmr - IgG
jar	2.519990483	4.95E-06	Fmr - IgG
Top3beta	3.376915117	4.98E-06	Fmr - IgG
RpS15Aa	2.913458719	5.22E-06	Fmr - IgG
Rack1	2.763729554	5.44E-06	Fmr - IgG
yps	5.374933774	5.72E-06	Fmr - IgG
RpL11	3.480045154	5.86E-06	Fmr - IgG
ball	2.278760515	5.92E-06	Fmr - IgG
RpS11	3.876622264	6.21E-06	Fmr - IgG
CG4586	2.642816065	6.36E-06	Fmr - IgG
CG7878	2.834535031	6.63E-06	Fmr - IgG
CG17018	3.102990675	6.64E-06	Fmr - IgG
RpS10b	2.763357234	6.71E-06	Fmr - IgG
RpS2	3.159117842	6.77E-06	Fmr - IgG
B52	2.819520975	6.89E-06	Fmr - IgG
RpS6	3.031586791	6.94E-06	Fmr - IgG
Pih1D1	3.361287023	6.97E-06	Fmr - IgG
CG12301	3.281520339	7.09E-06	Fmr - IgG
Ptip	2.550910439	7.19E-06	Fmr - IgG
Fib	3.382350151	8.36E-06	Fmr - IgG
CG6693	4.384775295	8.64E-06	Fmr - IgG
wupA	2.727087089	8.81E-06	Fmr - IgG
zip	2.58984652	8.93E-06	Fmr - IgG
RpL22	2.584979105	9.00E-06	Fmr - IgG
RpS14a	2.764706992	9.35E-06	Fmr - IgG
CG2982	2.873205914	9.41E-06	Fmr - IgG
mod	2.383803709	9.49E-06	Fmr - IgG
CG10754	2.554245849	9.79E-06	Fmr - IgG
ncd	2.50231348	9.81E-06	Fmr - IgG
stc	3.670239584	9.97E-06	Fmr - IgG
RpS18	3.515568852	1.01E-05	Fmr - IgG
eIF-3p66	2.482904666	1.01E-05	Fmr - IgG
aret	3.468717139	1.01E-05	Fmr - IgG
bel	3.194664818	1.07E-05	Fmr - IgG
pAbp	4.597811999	1.13E-05	Fmr - IgG
CG16941	3.250749694	1.15E-05	Fmr - IgG
wisp	4.251554866	1.17E-05	Fmr - IgG
Hsc70-4	2.167153053	1.17E-05	Fmr - IgG
NHP2	3.054059576	1.19E-05	Fmr - IgG
His1	5.935763291	1.21E-05	Fmr - IgG
His3	4.039619382	1.22E-05	Fmr - IgG
RpS12	2.448524583	1.23E-05	Fmr - IgG
up	2.557701827	1.26E-05	Fmr - IgG
Act87E	2.227064419	1.32E-05	Fmr - IgG
didum	2.377175737	1.41E-05	Fmr - IgG
LysRS	2.017777339	1.41E-05	Fmr - IgG
Hsc70-3	2.086202105	1.51E-05	Fmr - IgG
noi	2.813505422	1.60E-05	Fmr - IgG
Tm2	2.511849053	1.65E-05	Fmr - IgG
CG7518	3.499940489	1.66E-05	Fmr - IgG
CG8963	3.314048304	1.70E-05	Fmr - IgG
rig	2.304243346	1.70E-05	Fmr - IgG
Hrb98DE	2.691393355	1.70E-05	Fmr - IgG
Top2	2.081244137	1.73E-05	Fmr - IgG

Mhc	2.470352349	1.73E-05	Fmr - IgG
Prm	2.337051759	1.83E-05	Fmr - IgG
Ddx1	2.434760599	1.84E-05	Fmr - IgG
Patr-1	2.818099155	1.96E-05	Fmr - IgG
Zn72D	3.942564808	1.99E-05	Fmr - IgG
tws	1.999480729	1.99E-05	Fmr - IgG
CG5641	2.5638312	1.99E-05	Fmr - IgG
sta	1.960396003	2.01E-05	Fmr - IgG
Ubi-p63E	1.907748685	2.12E-05	Fmr - IgG
CG4038	3.306644488	2.24E-05	Fmr - IgG
RpL5	2.697844124	2.29E-05	Fmr - IgG
Cp7Fb	2.551019248	2.54E-05	Fmr - IgG
Cbp80	3.301298417	2.58E-05	Fmr - IgG
Aldh	2.00319584	2.63E-05	Fmr - IgG
Tm1	2.411770753	2.67E-05	Fmr - IgG
mahe	2.709237074	2.89E-05	Fmr - IgG
CG3800	2.227244166	2.92E-05	Fmr - IgG
eIF4G	2.174786576	3.20E-05	Fmr - IgG
Mtpalpha	1.737923204	3.25E-05	Fmr - IgG
asp	3.102922119	3.44E-05	Fmr - IgG
Ge-1	2.02332588	3.51E-05	Fmr - IgG
Klp10A	1.858558737	3.52E-05	Fmr - IgG
Srp68	2.234829304	3.57E-05	Fmr - IgG
His2B	4.363102919	3.72E-05	Fmr - IgG
CG11148	2.101295379	3.73E-05	Fmr - IgG
spoon	1.91797467	4.11E-05	Fmr - IgG
His4	2.966917521	4.27E-05	Fmr - IgG
eIF3-S8	1.972051467	4.36E-05	Fmr - IgG
hoip	2.562322203	4.40E-05	Fmr - IgG
Unr	1.78189766	4.40E-05	Fmr - IgG
RpL10Ab	2.109056108	4.54E-05	Fmr - IgG
RpL7A	2.124931701	4.76E-05	Fmr - IgG
CG10077	3.774229355	4.76E-05	Fmr - IgG
dre4	1.964430962	4.76E-05	Fmr - IgG
SkpA	1.775878141	5.06E-05	Fmr - IgG
RpA-70	1.563861364	5.36E-05	Fmr - IgG
eIF5B	1.733912927	6.67E-05	Fmr - IgG
CG5642	1.562549012	6.71E-05	Fmr - IgG
Hsp26	1.97699272	6.76E-05	Fmr - IgG
eIF-2beta	1.629398172	6.80E-05	Fmr - IgG
RpL30	1.492963013	6.88E-05	Fmr - IgG
RpS23	2.376464698	7.15E-05	Fmr - IgG
Droj2	1.613016733	7.60E-05	Fmr - IgG
CkIIbeta	1.756168577	7.89E-05	Fmr - IgG
eIF3-S9	1.588255144	8.37E-05	Fmr - IgG
RpL19	2.236259771	8.48E-05	Fmr - IgG
CG15735	3.856215743	8.61E-05	Fmr - IgG
Hsp27	1.998845065	8.96E-05	Fmr - IgG
vas	1.681651699	9.05E-05	Fmr - IgG
tud	2.021418055	9.12E-05	Fmr - IgG
eIF3-S10	1.645982125	9.42E-05	Fmr - IgG
Prp19	1.607041679	9.84E-05	Fmr - IgG
RpL6	2.055496475	9.94E-05	Fmr - IgG
RpL7	2.067169323	0.00010161	Fmr - IgG
Mlc-c	1.857259308	0.000102217	Fmr - IgG
Dp1	1.604189246	0.000103347	Fmr - IgG
Hrb27C	4.155205114	0.000109947	Fmr - IgG
Usp7	1.639006815	0.000111866	Fmr - IgG
RpL23	2.67616434	0.000113606	Fmr - IgG
Thiolase	1.557228799	0.000117287	Fmr - IgG
CG4747	1.780661174	0.000121491	Fmr - IgG
alt	1.705458163	0.000122875	Fmr - IgG
RpL4	1.986388903	0.000123192	Fmr - IgG
lds	1.966876262	0.000123209	Fmr - IgG
CG4692	1.370722751	0.000125299	Fmr - IgG
Tudor-SN	1.316578276	0.00013029	Fmr - IgG
aub	3.296800256	0.000131742	Fmr - IgG
Prp8	1.728121907	0.00013408	Fmr - IgG
RpLP0	1.455617539	0.000136872	Fmr - IgG
RpL28	2.168694785	0.000141005	Fmr - IgG
Nap1	1.401242345	0.00014108	Fmr - IgG

Gnf1	1.780602986	0.000143057	Fmr - IgG
Rfabg	1.411707089	0.000151673	Fmr - IgG
Nup358	1.574772996	0.000156834	Fmr - IgG
capt	1.348397461	0.000157033	Fmr - IgG
Caf1	1.314982776	0.000171948	Fmr - IgG
14-3-3epsilon	1.543627446	0.000175746	Fmr - IgG
Hsp23	1.810935259	0.000181352	Fmr - IgG
RpL14	2.172696361	0.000183081	Fmr - IgG
CG4806	3.467063952	0.000186055	Fmr - IgG
CG10254	1.908432586	0.000186441	Fmr - IgG
Trip1	1.390171124	0.000192803	Fmr - IgG
Rpn5	1.541098472	0.000200469	Fmr - IgG
RpL10	2.373904951	0.000205236	Fmr - IgG
alphaCOP	1.701848201	0.00021094	Fmr - IgG
Elf	1.231630052	0.000211034	Fmr - IgG
14-3-3zeta	1.462615234	0.000212278	Fmr - IgG
RpL21	1.471823014	0.000227425	Fmr - IgG
RpS7	1.966496645	0.000233634	Fmr - IgG
Mlc2	2.265227506	0.000237173	Fmr - IgG
Mf	1.84707301	0.000237599	Fmr - IgG
ATPsyngamma	1.300250493	0.000239712	Fmr - IgG
eIF3-S6	1.516242352	0.000250096	Fmr - IgG
RpS21	2.182671393	0.000250223	Fmr - IgG
RpL8	1.93361197	0.000263652	Fmr - IgG
OstDelta	1.160305864	0.00026618	Fmr - IgG
ACC	1.545282886	0.00026724	Fmr - IgG
sw	1.37650769	0.000276248	Fmr - IgG
eIF3ga	1.405469046	0.000284997	Fmr - IgG
alphaTub84B	1.27138999	0.000288374	Fmr - IgG
CG3226	1.244844876	0.000291125	Fmr - IgG
betaCOP	1.317295466	0.000291509	Fmr - IgG
Jabba	1.159219868	0.000292089	Fmr - IgG
RpL13	1.819410431	0.000293101	Fmr - IgG
CG12304	1.511628346	0.000293249	Fmr - IgG
Chc	1.456002179	0.000299802	Fmr - IgG
mtrm	2.059537629	0.000299834	Fmr - IgG
CG12163	1.346489385	0.00030517	Fmr - IgG
porin	1.221806307	0.000318193	Fmr - IgG
Parp	1.352541905	0.000323762	Fmr - IgG
fon	1.664122172	0.000330936	Fmr - IgG
eIF-3p40	1.575146246	0.000331032	Fmr - IgG
Drp1	1.445364226	0.000332747	Fmr - IgG
Actn	1.525846491	0.000336537	Fmr - IgG
Ote	1.242227163	0.000340725	Fmr - IgG
SdhA	1.208240245	0.000344932	Fmr - IgG
CG3364	1.557307421	0.00036866	Fmr - IgG
CG1416	1.231162672	0.000370189	Fmr - IgG
IleRS	1.287258428	0.000405172	Fmr - IgG
Karybeta3	1.188344764	0.000405748	Fmr - IgG
nudC	1.198829334	0.000407485	Fmr - IgG
r	1.097552132	0.00040971	Fmr - IgG
Dhc64C	1.357731961	0.000410457	Fmr - IgG
Tango7	1.187528799	0.000422571	Fmr - IgG
CG14309	1.259606999	0.000431612	Fmr - IgG
CG9684	1.679023374	0.000450249	Fmr - IgG
Fs(2)Ket	1.211259487	0.000452095	Fmr - IgG
Lam	1.156295789	0.000454995	Fmr - IgG
Pen	1.259437042	0.000458771	Fmr - IgG
epsilonCOP	1.173399314	0.000475902	Fmr - IgG
faf	1.985184105	0.000477401	Fmr - IgG
Ef1gamma	1.317231243	0.000480371	Fmr - IgG
RpL3	1.866290671	0.000481792	Fmr - IgG
AsnRS	1.124047782	0.000482746	Fmr - IgG
Sec31	1.307958239	0.000487494	Fmr - IgG
beta'COP	1.249837634	0.0004895	Fmr - IgG
RpL23A	2.30648208	0.000499888	Fmr - IgG
Rpn6	1.490867051	0.000500825	Fmr - IgG
CG6439	1.258034581	0.000503644	Fmr - IgG
RpL38	1.344811396	0.000541739	Fmr - IgG
CSN4	1.348097829	0.000544866	Fmr - IgG
Gtp-bp	1.377516398	0.000545847	Fmr - IgG

ThrRS	1.185748515	0.000554732	Fmr - IgG
shi	1.326340514	0.000555467	Fmr - IgG
CG5525	1.137632582	0.000571454	Fmr - IgG
bor	1.287086106	0.000582055	Fmr - IgG
yl	1.421186756	0.00058388	Fmr - IgG
Rpt3	1.022316068	0.000591843	Fmr - IgG
RfC38	1.308821095	0.000611147	Fmr - IgG
CG10407	1.129935719	0.000620732	Fmr - IgG
RPA2	1.035147838	0.0006375	Fmr - IgG
Aats-glupro	1.33702201	0.000639647	Fmr - IgG
Nmt	1.319514765	0.000644035	Fmr - IgG
AP-2alpha	1.346260991	0.000648418	Fmr - IgG
RpL17	2.292610339	0.00067485	Fmr - IgG
Uba1	1.349873768	0.000711806	Fmr - IgG
Vap-33A	1.241136952	0.000725316	Fmr - IgG
RpL24	2.353611511	0.000726199	Fmr - IgG
CHOp24	1.059207223	0.00076133	Fmr - IgG
sesB	1.152838393	0.000792852	Fmr - IgG
Cctgamma	1.177576981	0.000793254	Fmr - IgG
gammaTub37C	1.267505134	0.000796036	Fmr - IgG
FK506-bp1	1.070289418	0.000796053	Fmr - IgG
Phb2	1.131103613	0.000829378	Fmr - IgG
l(1)G0156	1.195071285	0.000838421	Fmr - IgG
Hsp83	1.163468658	0.000853414	Fmr - IgG
RpL9	1.7574608	0.000917765	Fmr - IgG
betaTub56D	1.284132682	0.000922916	Fmr - IgG
sqd	1.765328269	0.00092913	Fmr - IgG
poe	1.078958979	0.000963477	Fmr - IgG
Cam	1.396278026	0.00096867	Fmr - IgG
SERCA	1.010825195	0.000980784	Fmr - IgG
Nup205	1.214000548	0.001023723	Fmr - IgG
CG6089	1.011261917	0.001053439	Fmr - IgG
P32	1.275791539	0.001071236	Fmr - IgG
AP-1-2beta	1.015326727	0.00108091	Fmr - IgG
Rpt6	1.079768902	0.001096272	Fmr - IgG
bai	1.016126181	0.001120879	Fmr - IgG
sea	1.230608447	0.001147859	Fmr - IgG
RpLP1	1.052139202	0.001178228	Fmr - IgG
RpL13A	1.650615373	0.001230913	Fmr - IgG
alphaTub67C	1.395815248	0.001254	Fmr - IgG
Tcp-1zeta	1.023207257	0.001270222	Fmr - IgG
swa	1.539085989	0.00127709	Fmr - IgG
Eb1	1.171688985	0.001295847	Fmr - IgG
eca	1.13490585	0.001308796	Fmr - IgG
Dbp80	1.018938749	0.001372368	Fmr - IgG
RpL18A	1.202671541	0.001780201	Fmr - IgG
Cp1	1.139577263	0.001887457	Fmr - IgG
NAT1	1.758299209	0.002009625	Fmr - IgG
Mlc1	1.923318038	0.002076453	Fmr - IgG
RanGAP	1.144651505	0.002151399	Fmr - IgG
Prosalpha4	1.238346498	0.00232485	Fmr - IgG
msps	1.07111407	0.002531095	Fmr - IgG
Not1	1.084101649	0.002702258	Fmr - IgG
RpL12	1.997126298	0.002714606	Fmr - IgG
Rpn11	1.038858194	0.002967995	Fmr - IgG
CG7033	1.172564293	0.003019537	Fmr - IgG
Act57B	2.701474112	0.00324898	Fmr - IgG
CG13083-RB	1.128705915	0.00363414	Fmr - IgG
Ppn	1.382933707	0.004040295	Fmr - IgG
Ankle2	1.33926853	0.004139317	Fmr - IgG
Rpn3	1.02269676	0.004255403	Fmr - IgG
UQCR-C2	1.06110588	0.004566702	Fmr - IgG
CG12262	1.04062461	0.004653346	Fmr - IgG
blw	1.023167891	0.004739596	Fmr - IgG
CG17514	1.026080625	0.005039044	Fmr - IgG
Rpn1	1.124679172	0.005938308	Fmr - IgG
yellow-g2	1.093821007	0.005971236	Fmr - IgG
26-29-p	1.046939734	0.007230091	Fmr - IgG
yellow-g	1.233480124	0.007703278	Fmr - IgG
trol	1.552049516	0.009201167	Fmr - IgG
orb	3.119446022	0.011141031	Fmr - IgG

FKBP59	1.051593254	0.011891051	Fmr - IgG
Ef1alpha48D	1.029673727	0.013354684	Fmr - IgG
RpS8	2.889358792	0.019357567	Fmr - IgG
eIF-4a	1.051736891	0.023800678	Fmr - IgG
Nup154	1.292506976	0.024583862	Fmr - IgG
vig2	1.003189622	0.027872816	Fmr - IgG
Cct5	1.00813802	0.028253038	Fmr - IgG
CG8142	1.413963469	0.02940282	Fmr - IgG
Nacalpha	1.014348591	0.033404002	Fmr - IgG
Fdh	1.146639878	0.033963479	Fmr - IgG
SF2	-4.198555612	7.05E-06	(Fmr + RNase - IgG + RNase) - (Fmr - IgG)
Parp	3.64373597	8.00E-06	(Fmr + RNase - IgG + RNase) - (Fmr - IgG)
CG30497	-3.750700643	2.96E-05	(Fmr + RNase - IgG + RNase) - (Fmr - IgG)
RPA2	2.271197527	5.47E-05	(Fmr + RNase - IgG + RNase) - (Fmr - IgG)
rump	-2.612184926	5.60E-05	(Fmr + RNase - IgG + RNase) - (Fmr - IgG)
Tdrd3	-2.845755182	7.29E-05	(Fmr + RNase - IgG + RNase) - (Fmr - IgG)
SRPK	-3.762948545	7.55E-05	(Fmr + RNase - IgG + RNase) - (Fmr - IgG)
Imp	-3.12669899	7.76E-05	(Fmr + RNase - IgG + RNase) - (Fmr - IgG)
RpS13	-2.467251947	9.15E-05	(Fmr + RNase - IgG + RNase) - (Fmr - IgG)
Top3beta	-2.865019531	9.75E-05	(Fmr + RNase - IgG + RNase) - (Fmr - IgG)
CG12301	-2.888396076	0.000112202	(Fmr + RNase - IgG + RNase) - (Fmr - IgG)
lark	-2.0817181	0.00016038	(Fmr + RNase - IgG + RNase) - (Fmr - IgG)
Pep	-2.369738877	0.000174741	(Fmr + RNase - IgG + RNase) - (Fmr - IgG)
exu	-2.675361644	0.000201832	(Fmr + RNase - IgG + RNase) - (Fmr - IgG)
yps	-4.076911462	0.000209702	(Fmr + RNase - IgG + RNase) - (Fmr - IgG)
NHP2	-2.597200395	0.000221369	(Fmr + RNase - IgG + RNase) - (Fmr - IgG)
Zn72D	-3.426901096	0.00031936	(Fmr + RNase - IgG + RNase) - (Fmr - IgG)
CG13096	-2.03863656	0.000363008	(Fmr + RNase - IgG + RNase) - (Fmr - IgG)
stau	-1.745255897	0.000392269	(Fmr + RNase - IgG + RNase) - (Fmr - IgG)
Ssrp	2.142559561	0.000454617	(Fmr + RNase - IgG + RNase) - (Fmr - IgG)
RpA-70	1.50661488	0.000463181	(Fmr + RNase - IgG + RNase) - (Fmr - IgG)
Nop56	-1.717515507	0.000477233	(Fmr + RNase - IgG + RNase) - (Fmr - IgG)
pcm	-1.859376514	0.00048547	(Fmr + RNase - IgG + RNase) - (Fmr - IgG)
CG8915	-2.657270076	0.000502895	(Fmr + RNase - IgG + RNase) - (Fmr - IgG)
ball	-1.469991332	0.000527961	(Fmr + RNase - IgG + RNase) - (Fmr - IgG)
btz	-3.846507723	0.000547532	(Fmr + RNase - IgG + RNase) - (Fmr - IgG)
CG2982	-1.982523897	0.000561347	(Fmr + RNase - IgG + RNase) - (Fmr - IgG)
CG6693	-2.967881662	0.000575928	(Fmr + RNase - IgG + RNase) - (Fmr - IgG)
Rbp1	-2.398091912	0.000604735	(Fmr + RNase - IgG + RNase) - (Fmr - IgG)
sqd	2.656797881	0.000666838	(Fmr + RNase - IgG + RNase) - (Fmr - IgG)
La	-2.100102908	0.000682084	(Fmr + RNase - IgG + RNase) - (Fmr - IgG)
Atx2	-1.635711619	0.000698708	(Fmr + RNase - IgG + RNase) - (Fmr - IgG)
Cen	-1.983058465	0.000731568	(Fmr + RNase - IgG + RNase) - (Fmr - IgG)
RpS9	-1.792575726	0.000748698	(Fmr + RNase - IgG + RNase) - (Fmr - IgG)
ncd	-1.636728676	0.000776927	(Fmr + RNase - IgG + RNase) - (Fmr - IgG)
Ref1	-2.088304102	0.000812456	(Fmr + RNase - IgG + RNase) - (Fmr - IgG)
BicC	-1.520185371	0.000881637	(Fmr + RNase - IgG + RNase) - (Fmr - IgG)
Nop60B	-1.578000875	0.000898079	(Fmr + RNase - IgG + RNase) - (Fmr - IgG)
Srp72	-1.987460512	0.000915562	(Fmr + RNase - IgG + RNase) - (Fmr - IgG)
CG4038	-2.410772125	0.000920817	(Fmr + RNase - IgG + RNase) - (Fmr - IgG)
RpS11	-2.239983335	0.000993724	(Fmr + RNase - IgG + RNase) - (Fmr - IgG)
RpS28b	-1.541570806	0.001017183	(Fmr + RNase - IgG + RNase) - (Fmr - IgG)
larp	-1.543393403	0.001044121	(Fmr + RNase - IgG + RNase) - (Fmr - IgG)
piwi	-2.140144746	0.001132821	(Fmr + RNase - IgG + RNase) - (Fmr - IgG)
His1	-3.649133449	0.001303293	(Fmr + RNase - IgG + RNase) - (Fmr - IgG)
RpS6	-1.686423381	0.001338492	(Fmr + RNase - IgG + RNase) - (Fmr - IgG)
tyf	-1.424948598	0.001507054	(Fmr + RNase - IgG + RNase) - (Fmr - IgG)
RpS15	-1.553414683	0.001600033	(Fmr + RNase - IgG + RNase) - (Fmr - IgG)
lost	-1.889465334	0.001798181	(Fmr + RNase - IgG + RNase) - (Fmr - IgG)
RpS5	-1.135643305	0.00205185	(Fmr + RNase - IgG + RNase) - (Fmr - IgG)
swa	1.981556665	0.00206877	(Fmr + RNase - IgG + RNase) - (Fmr - IgG)
Ddx1	-1.425395196	0.00240277	(Fmr + RNase - IgG + RNase) - (Fmr - IgG)
dre4	1.352790288	0.002422578	(Fmr + RNase - IgG + RNase) - (Fmr - IgG)
nop5	-1.344146948	0.002490243	(Fmr + RNase - IgG + RNase) - (Fmr - IgG)
hoip	-1.703524093	0.002697155	(Fmr + RNase - IgG + RNase) - (Fmr - IgG)
RpL31	-1.108574831	0.002926461	(Fmr + RNase - IgG + RNase) - (Fmr - IgG)
CG7518	-1.917483081	0.003067968	(Fmr + RNase - IgG + RNase) - (Fmr - IgG)
CG4586	-1.21934744	0.003198448	(Fmr + RNase - IgG + RNase) - (Fmr - IgG)
rin	-1.452611514	0.003353295	(Fmr + RNase - IgG + RNase) - (Fmr - IgG)
RpS19a	-1.029109944	0.003743369	(Fmr + RNase - IgG + RNase) - (Fmr - IgG)
CG5787	-1.588558787	0.003759707	(Fmr + RNase - IgG + RNase) - (Fmr - IgG)

CG5641	-1.364615746	0.004113258	(Fmr + RNase - IgG + RNase) - (Fmr - IgG)
pAbp	-2.110720076	0.005250323	(Fmr + RNase - IgG + RNase) - (Fmr - IgG)
eIF-3p66	-1.114459465	0.00531354	(Fmr + RNase - IgG + RNase) - (Fmr - IgG)
Cp7Fb	1.330028419	0.005519265	(Fmr + RNase - IgG + RNase) - (Fmr - IgG)
Fib	-1.443189597	0.005813088	(Fmr + RNase - IgG + RNase) - (Fmr - IgG)
spoon	-1.065743409	0.006051562	(Fmr + RNase - IgG + RNase) - (Fmr - IgG)
Prp19	-1.034364567	0.006193956	(Fmr + RNase - IgG + RNase) - (Fmr - IgG)
Act57B	3.250811966	0.006999141	(Fmr + RNase - IgG + RNase) - (Fmr - IgG)
fon	1.288435722	0.007164685	(Fmr + RNase - IgG + RNase) - (Fmr - IgG)

Table A.4: Fmr1 CLIP enriched mRNAs

Flybase identifier	gene	log2FC	p-value	gene length
FBgn0011230	poe	1.373222215	9.39E-28	17302
FBgn0030674	HUWE1	0.959933155	9.70E-21	19307
FBgn0261797	Dhc64C	0.976032427	7.44E-12	18319
FBgn0039214	puf	0.914148431	2.60E-10	14773
FBgn0266717	Bruce	0.734045936	2.71E-10	19314
FBgn0027603	Ulp1	0.921631524	8.72E-10	6084
FBgn0005666	bt	0.993287287	1.62E-09	52074
FBgn0261710	nocte	0.567837515	5.11E-08	12810
FBgn0265988	mv	1.117880618	5.38E-07	13872
FBgn0261836	Msp300	0.601346608	1.67E-06	106125
FBgn0031006	rictor	0.913604992	3.73E-06	9129
FBgn0016081	fry	0.925601691	4.62E-06	47269
FBgn0260003	Dys	0.689000718	4.97E-06	136203
FBgn0086451	l(2)k09022	0.966937769	5.30E-06	6707
FBgn0052602	Muc12Ea	1.070074068	7.03E-06	10426
FBgn0031052	Elys	1.112729697	7.31E-06	7102
FBgn0003891	tud	0.542479008	1.03E-05	9591
FBgn0004649	yl	0.583545314	1.07E-05	6744
FBgn0037541	CG2747	0.935985622	2.04E-05	14580
FBgn0004167	kst	0.643831108	5.28E-05	31273
FBgn0011785	BRWD3	0.716486319	5.46E-05	9619
FBgn0027948	msps	0.704993503	9.40E-05	9874
FBgn0013733	shot	0.564204696	9.50E-05	77873
FBgn0250789	alpha-Spec	0.554002925	0.000103739	16843
FBgn0000052	Pfas	0.999909048	0.000104478	4869
FBgn0015269	Nf1	0.868312536	0.000112796	12724
FBgn0027079	ValRS	1.150643231	0.00012044	3859
FBgn0015477	Rme-8	0.703610506	0.000160607	9498
FBgn0033194	Vps13	0.875392729	0.000170779	13866
FBgn0004598	Fur2	0.634657634	0.000198109	8696
FBgn0266111	ana3	0.929648357	0.000205387	6342
FBgn0023458	Rbcn-3A	0.703974741	0.000219195	14742
FBgn0033688	Prp8	0.625832546	0.000242309	8740
FBgn0261547	Exn	1.171860255	0.000361583	24375
FBgn0027655	htt	1.002073311	0.000404557	38574
FBgn0015589	Apc	0.706370421	0.000451389	12447
FBgn0030266	CG11122	0.768864103	0.000569199	12331
FBgn0010660	Nup214	0.614086728	0.000608898	5805
FBgn0263600	DNApol-delta	0.943170583	0.000674913	3493
FBgn0032479	CG16974	0.643501167	0.00067492	8130
FBgn0028476	Usp1	0.630037795	0.000787571	6141
FBgn0263968	nonC	0.737165662	0.000842972	13626
FBgn0026059	Mhcl	0.660832545	0.000894953	30238
FBgn0030246	CG1582	0.919832602	0.000921918	4555
FBgn0030286	Gapvd1	0.764653877	0.000965999	8686
FBgn0014006	Ask1	0.78282932	0.000990801	6311
FBgn0034964	IntS1	0.712052126	0.0010171	6515
FBgn0022787	Hel89B	0.825126668	0.001050086	10153
FBgn0087008	e(y)3	0.457617497	0.001092023	11821
FBgn0001324	kto	0.830514391	0.001126354	8330
FBgn0005632	faf	0.488152109	0.001195123	13661
FBgn0266696	Svil	0.408578291	0.001317283	121571
FBgn0011286	RyR	1.464203566	0.001485653	27733
FBgn0038108	CG7518	0.58756351	0.001606061	10315

FBgn0008635	betaCOP	0.812610633	0.00168402	3693
FBgn0035713	velo	0.68237981	0.001696197	8326
FBgn0031078	Nup205	0.731489402	0.001793598	7165
FBgn0004391	shtd	0.928370259	0.001849946	6897
FBgn0034878	pita	0.841755967	0.001901241	5745
FBgn0031107	HERC2	0.607690301	0.001923721	17852
FBgn0260794	ctrip	0.343645069	0.001981291	25812
FBgn0261671	tweek	0.452745717	0.002234654	28562
FBgn0261574	kug	0.574063346	0.002244932	19306
FBgn0010355	Taf1	0.665902262	0.002418355	9159
FBgn0029688	lva	0.529932252	0.002473246	9784
FBgn0031266	Sf3b1	0.850679065	0.002605715	4567
FBgn0035267	CG13921	1.124098967	0.002841602	5483
FBgn0026143	Cdc45	0.811820894	0.002906607	2265
FBgn0021796	Tor	0.637120809	0.00305746	8717
FBgn0039727	Vps13B	0.762480745	0.003088448	12229
FBgn0052113	Vps13D	0.738753345	0.003208596	13556
FBgn0034643	CG10321	0.729021473	0.003217037	3720
FBgn0028982	Spt6	0.511289167	0.003278562	7187
FBgn0034400	CG15099	0.553923287	0.003420153	9798
FBgn0037717	CG8301	1.401312254	0.0034896	5342
FBgn0031392	AIF	0.79631943	0.003536997	3635
FBgn0261554	CG42672	0.725626586	0.003586154	14759
FBgn0026239	gukh	0.905839932	0.00361072	38411
FBgn0026598	Apc2	0.89693325	0.003629659	4298
FBgn0026713	l(1)G0007	1.053706642	0.003861779	23370
FBgn0086906	sls	0.497447791	0.003986826	75930
FBgn0039696	Rnb	0.989963382	0.004162019	3657
FBgn0010825	Gug	0.365952925	0.004483853	22747
FBgn0263667	Lpt	0.99480558	0.00454082	5100
FBgn0040233	cana	0.509338387	0.004550597	8066
FBgn0004374	neb	1.142083622	0.004600591	13941
FBgn0033246	ACC	0.441337232	0.004785077	17747
FBgn0085451	htk	0.585628728	0.004805258	12195
FBgn0004647	N	0.457689191	0.004999309	37351
FBgn0016756	Usp47	0.474288971	0.005108311	8913
FBgn0037249	eIF3a	0.421666304	0.005318351	4097
FBgn0036448	mop	0.52835505	0.005473753	6788
FBgn0041171	ago	0.44362851	0.005517809	7963
FBgn0034786	CG13531	0.807825546	0.005865015	5995
FBgn0039581	Moca-cyp	0.682048779	0.005995721	4835
FBgn0053208	Mical	0.625856586	0.006166585	41341
FBgn0264357	SNF4Agamma	0.408897732	0.006187705	73530
FBgn0037371	Sym	0.90397402	0.006315753	4298
FBgn0031390	tho2	0.661523404	0.006316066	7837
FBgn0034618	CG9485	0.605546337	0.006490721	7170
FBgn0003048	pcx	0.494138323	0.00655903	12749
FBgn0283503	Neurl4	1.160224186	0.006574677	6198
FBgn0032395	Wdr81	0.845005062	0.006589116	6656
FBgn0035766	eco	0.682371762	0.007271807	3898
FBgn0043884	mask	0.255636502	0.007314107	19686
FBgn0263603	Zn72D	0.797159791	0.007381455	6032
FBgn0052226	Pex23	0.842328103	0.007462035	5012
FBgn0034853	Ice1	0.677107028	0.007478742	4593
FBgn0023518	trr	0.631455152	0.007559327	8637
FBgn0029912	CG4557	0.936612773	0.007579674	3506
FBgn0250850	rig	0.558152456	0.007925801	6805
FBgn0264326	DNApol-epsilon255	0.78721543	0.00807415	7035
FBgn0260442	rhea	0.526589474	0.008301569	29408
FBgn0266064	GlyS	0.490805666	0.008545001	5688
FBgn0039302	Nup358	0.336868416	0.008918094	9324
FBgn0033757	muskelin	0.52652993	0.008978373	3681
FBgn0038118	timeout	0.912960632	0.009198416	75225
FBgn0021761	Nup154	0.560782681	0.009500449	6101
FBgn0030974	Flacc	0.677886225	0.009895152	4658
FBgn0003189	r	0.657331798	0.01011291	13353
FBgn0039207	CG5789	0.56612182	0.010363001	7710
FBgn0001308	Khc	0.420440744	0.011004193	5011
FBgn0031020	Naa15-16	0.564492629	0.011101033	4659
FBgn0020306	dom	0.254066736	0.011222373	18417
FBgn0002526	LanA	0.588617008	0.011230355	14191

FBgn0086895	pea	0.703649214	0.011338346	3953
FBgn0020647	KrT95D	0.534929833	0.011403294	34289
FBgn0031985	mon2	0.63991809	0.011406816	7045
FBgn0039016	Dcr-1	0.859487357	0.011586733	7084
FBgn0265574	Cdc5	0.69201807	0.01179102	3532
FBgn0082831	pps	0.532068544	0.011810818	7423
FBgn0033766	Nup188	0.819626851	0.011906528	6212
FBgn0284252	Letm1	0.598949573	0.012032647	5812
FBgn0003963	ush	0.617607753	0.012308113	64340
FBgn0025681	CG3558	0.685564208	0.012471896	5856
FBgn0032467	CG9934	0.532783438	0.01269478	6158
FBgn0029763	Usp16-45	0.513027404	0.012757533	5686
FBgn0263490	mld	0.415730606	0.012859488	38246
FBgn0025571	SF1	0.627367406	0.013057501	4727
FBgn0029992	Upf2	0.79456479	0.013074324	5222
FBgn0085436	Not1	0.219623883	0.013187874	12861
FBgn0003464	sol	0.598920574	0.013725182	7632
FBgn0029736	CG4041	0.987624553	0.014031587	3167
FBgn0025802	Sbf	0.675781314	0.014567171	7655
FBgn0024329	Mekk1	0.561028115	0.015211665	10488
FBgn0283531	Duox	0.58141002	0.015267466	14278
FBgn0031990	PAPLA1	0.437444664	0.015433753	25391
FBgn0086655	jing	0.639612488	0.015758858	119786
FBgn0005198	gig	0.795736475	0.015821141	9829
FBgn0053087	LRP1	0.474210368	0.015868544	52273
FBgn0038320	Sra-1	0.524087062	0.015878992	5099
FBgn0030240	CG2202	0.799404816	0.016124135	3570
FBgn0039554	CG5003	0.793207939	0.01659847	2739
FBgn0040236	c11.1	0.575182312	0.016913503	7799
FBgn0000376	crm	0.521388832	0.017229121	4525
FBgn0026181	Rok	0.4139835	0.017284816	13402
FBgn0030858	IntS2	0.847112455	0.01734809	4057
FBgn0064766	CG7600	0.871432239	0.017358148	3311
FBgn0267350	Pi4KIIIalpha	0.449106461	0.017452287	8868
FBgn0000140	asp	0.443321026	0.017615121	6664
FBgn0001108	DCTN1-p150	0.65013413	0.01776188	5269
FBgn0263144	bin3	0.360808545	0.017905289	26034
FBgn0025864	Crag	0.451693133	0.017987448	10005
FBgn0035162	Sf3b3	0.647420642	0.018198178	6665
FBgn0283500	Sac1	0.799955592	0.018683856	2634
FBgn0031023	CG14200	0.559393556	0.018708759	22897
FBgn0019938	RpI1	0.65293726	0.018915941	5605
FBgn0031606	CG15439	0.69540266	0.019063209	3743
FBgn0266557	kis	0.234212297	0.019359674	40088
FBgn0261553	CG42671	0.499941057	0.019779519	16240
FBgn0029861	CG3815	0.7306296	0.020014264	3581
FBgn0031420	Atxn7	0.754304851	0.020036526	4665
FBgn0030141	Gga	0.766027033	0.020096024	3250
FBgn0024277	trio	0.679140617	0.020113556	38893
FBgn0002878	mus101	0.714518907	0.020235649	4867
FBgn0035771	Sec63	0.553747365	0.020397444	3805
FBgn0019968	Khc-73	0.523559624	0.020537476	16481
FBgn0033155	Br140	0.749261194	0.020718952	5826
FBgn0030891	Ada3	0.822346776	0.020784497	2657
FBgn0016641	PTP-ER	0.617578508	0.021063454	5125
FBgn0003575	su(sable)	0.556918535	0.021190554	7870
FBgn0002783	mor	0.469868213	0.021329279	4779
FBgn0250786	Chd1	0.487239033	0.021394129	8161
FBgn0010051	Itpr	0.526051331	0.021593264	22294
FBgn0029088	disp	0.572136116	0.021727076	6614
FBgn0001978	stc	0.582907493	0.021805941	4589
FBgn0039004	Nup133	0.889202934	0.021895067	4165
FBgn0286222	Fum1	0.699990302	0.021960561	3421
FBgn0000562	egl	0.427882305	0.021969131	11925
FBgn0003175	px	0.429182974	0.02224424	67768
FBgn0027588	GCS2alpha	0.461218987	0.022271279	4762
FBgn0028563	sut1	0.668597297	0.022389571	4175
FBgn0038168	omd	0.705815051	0.022566413	7770
FBgn0036913	Usp32	0.5669892	0.022730806	10644
FBgn0027548	nito	0.433107198	0.022810894	7342
FBgn0259734	Nost	0.448216379	0.023496236	24370

FBgn0260990	yata	0.559459664	0.02365407	5309
FBgn0024973	CG2701	1.003094944	0.023843049	2689
FBgn0265045	Strn-Mlck	0.785604856	0.023937946	42565
FBgn0003415	skd	0.303041602	0.024216101	34620
FBgn0034598	CG4266	0.606513324	0.024564855	6946
FBgn0266284	Ns3	0.829956811	0.024571693	2376
FBgn0030530	jub	0.577446596	0.025014735	4276
FBgn0034734	CG4554	0.671433807	0.025410863	8730
FBgn0086694	Bre1	0.647816827	0.025421588	4361
FBgn0286567	gry	0.566873736	0.02543431	12258
FBgn0027783	SMC2	0.582109981	0.02579094	4032
FBgn0045035	tefu	0.560879857	0.025896831	10024
FBgn0039117	tst	0.848335433	0.025988446	4460
FBgn0052529	Hers	0.422183067	0.026092076	41962
FBgn0038968	CG12499	0.451605508	0.026185324	6435
FBgn0038651	Epg5	0.502353092	0.026336379	8437
FBgn0035802	Pura	0.616059033	0.02657105	46482
FBgn0000119	arr	0.603010747	0.026911976	29589
FBgn0259971	CG42481	2.082806974	0.027041477	226
FBgn0041582	tamo	0.553314735	0.027324861	6913
FBgn0039466	CG5521	0.406561495	0.02736913	10948
FBgn0086695	hd	0.716890377	0.027724974	2370
FBgn0264324	spg	0.527396672	0.027855751	25169
FBgn0003977	vir	0.92700879	0.027975158	6163
FBgn0034691	Synj	0.574788396	0.028002083	5591
FBgn0259166	CG42271	0.753581084	0.028244338	6006
FBgn0052479	Usp10	0.299477216	0.028257258	23414
FBgn0043854	slam	0.718537763	0.028709252	4896
FBgn0028538	Sec71	0.611823038	0.028752667	6153
FBgn0023097	bon	0.53663331	0.029001173	19401
FBgn0030420	CG12717	0.589969292	0.029230065	2969
FBgn0261532	cdm	0.625480665	0.02965429	4204
FBgn0036688	Fit2	0.903133496	0.029722507	2789
FBgn0040232	cmet	0.500542343	0.030218223	7976
FBgn0015600	toc	0.527324977	0.030483609	76441
FBgn0263289	scrib	0.454586304	0.030992854	67701
FBgn0035205	Ctr9	0.474254333	0.031001837	4581
FBgn0037109	MED1	0.422061542	0.031156636	6121
FBgn0027087	HisRS	0.646182401	0.031228051	5017
FBgn0050377	CG30377	0.527747514	0.03128696	25007
FBgn0003598	Su(var)3-7	0.403715563	0.031363488	5674
FBgn0263599	l(3)72Ab	0.595210958	0.031378525	7036
FBgn0052685	ZAP3	0.427189374	0.031399262	10634
FBgn0043458	CG12084	0.512447378	0.031479716	5692
FBgn0015371	chn	0.360051232	0.031604957	29547
FBgn0038344	obe	0.584016214	0.031673611	7260
FBgn0030946	CG6659	0.664001138	0.031934845	3894
FBgn0039668	Trc8	0.507605941	0.031951011	4772
FBgn0005674	GluProRS	0.628597066	0.032371905	6255
FBgn0027660	blot	0.438465147	0.03243935	25483
FBgn0262519	Mi-2	0.293763914	0.032537323	28164
FBgn0039508	CG3368	0.747942474	0.033497125	3668
FBgn0050372	Asap	0.577468008	0.033528554	8715
FBgn0037758	CG9467	0.49463632	0.033570174	3257
FBgn0031575	Cep97	0.872130609	0.033700362	3728
FBgn0031048	CG12237	0.731606773	0.033702853	1678
FBgn0286070	cnk	0.555843767	0.033822275	6037
FBgn0010328	woc	0.468066451	0.033831053	7460
FBgn0039633	CG11873	0.346124863	0.034085415	45908
FBgn0030701	CG16952	0.51295942	0.034211744	5424
FBgn0005386	ash1	0.448120828	0.035163524	7982
FBgn0044452	Atg2	0.476324852	0.035232148	6339
FBgn0043796	CG12219	0.714115042	0.035353055	2231
FBgn0050020	CG30020	0.538108148	0.035788408	6381
FBgn0259221	ATP8A	0.650642229	0.035988413	24073
FBgn0034946	CG3065	0.804695541	0.03630582	1928
FBgn0011204	cue	0.441658282	0.036458698	14045
FBgn0000247	ca	0.528766515	0.036613867	8236
FBgn0004368	Ptp4E	0.458847496	0.036698763	16788
FBgn0039241	CG11089	0.501618534	0.036808484	4433
FBgn0263110	CG43367	0.491138898	0.03729087	30010

FBgn0265523	Smr	0.190167768	0.037642479	58415
FBgn0266580	Gp210	0.482950268	0.0377625	13462
FBgn0029840	raptor	0.515627083	0.038678494	7567
FBgn0026430	Grip84	0.535923033	0.038813078	6333
FBgn0000414	Dab	0.537150891	0.039053668	12449
FBgn0261261	plx	0.53785144	0.03911532	26982
FBgn0038816	Lrrk	0.668261202	0.039570338	8799
FBgn0023172	RhoGEF2	0.344957054	0.039867878	17463
FBgn0266083	ocm	0.556591162	0.040540092	8587
FBgn0036451	Helz	0.300325256	0.040547841	11154
FBgn0263705	Myo10A	1.07122134	0.04101724	33206
FBgn0032517	CG7099	0.802306076	0.041268575	6388
FBgn0036574	elg1	0.421279325	0.041535547	4663
FBgn0022764	Sin3A	0.314934454	0.041913249	17677
FBgn0261931	CG42797	0.437038518	0.042110419	14297
FBgn0261535	l(2)34Fd	0.940078723	0.042174843	2491
FBgn0050183	CG30183	0.584196774	0.042416991	7577
FBgn0266917	Sf3a1	0.611688951	0.042823758	2971
FBgn0260936	scny	0.440968037	0.04302962	7269
FBgn0038424	CG17565	0.960759771	0.043104263	1742
FBgn0027866	CG9776	0.376272446	0.04385953	6257
FBgn0004387	Klp98A	0.428992954	0.0441001	11370
FBgn0038889	Fancm	0.862814317	0.044405379	5119
FBgn0260970	Ubr3	0.354118021	0.044509094	17987
FBgn0259736	CG42390	1.218514346	0.044513865	13306
FBgn0039559	NSD	0.499579442	0.044858089	4871
FBgn0034641	mahj	0.571358375	0.044915336	6392
FBgn0086908	egg	0.595163731	0.044978936	4367
FBgn0261985	Ptpmeg	0.485124248	0.045038213	27953
FBgn0030631	CG6227	0.590813301	0.045065959	4852
FBgn0033845	mars	0.652341492	0.045231285	5390
FBgn0035073	CG16896	0.685370392	0.045345258	3505
FBgn0030093	Bap111	0.549223878	0.045518962	2667
FBgn0002872	mu2	0.470471073	0.045983174	4741
FBgn0031655	Marcal1	1.024039763	0.046196551	2783
FBgn0086613	Ino80	0.626250303	0.046198863	34158
FBgn0035878	CG7182	0.49703834	0.046820874	1961
FBgn0034389	Mctp	0.6554448	0.047384565	13066
FBgn0260397	Su(var)3-3	0.542908343	0.047595161	3869
FBgn0260655	l(3)76BDm	0.669138798	0.048112369	5094
FBgn0036372	Abp1	0.520148709	0.048406207	2747
FBgn0030412	Tomosyn	0.450275535	0.048656501	21837
FBgn0263236	SP1029	0.977417141	0.048835802	5490
FBgn0002441	l(3)mbt	0.538660836	0.048954576	6947
FBgn0013995	Calx	0.543465424	0.049046254	37430
FBgn0035237	CG13917	0.33924827	0.049520176	24146
FBgn0260486	Ziz	0.562244181	0.04969814	23794

Table A.5: Rump CLIP enriched RNAs

Flybase identifier	gene	log2FC	p-value
FBgn0263413	lncRNA:CR43459	3.53354796	7.31E-14
FBgn0267913	lncRNA:CR46194	3.312326897	4.43E-08
FBgn0262109	lncRNA:CR42862	1.514421075	2.75E-07
FBgn0027279	l(1)G0196	1.268485753	2.86E-06
FBgn0265040	lncRNA:CR44157	3.701687492	3.18E-06
FBti0019258	frogger558	4.832352024	2.60E-05
FBgn0287603	ncRNA:CR46472	2.532030961	3.09E-05
FBgn0052141	saturn	5.086931898	6.73E-05
FBgn0003502	Btk29A	1.051674323	9.45E-05
FBgn0051538	CG31538	2.85130564	0.00012657
FBgn0033159	Dscam1	2.562696459	0.000132748
FBti0062048	Quasimodo3987	3.230876539	0.000179721
FBgn0266050	CG44815	3.761235264	0.000196655
FBti0059705	invader51644	4.037731011	0.000253924
FBgn0004652	fru	1.302101383	0.000265961
FBgn0001234	lncRNA:Hsromega	1.200320946	0.000478566

FBgn0286827	lncRNA:CR43334	1.992483864	0.000550752
FBgn0265065	asRNA:CR44176	4.232443407	0.000749822
FBgn0287461	lncRNA:CR46450	3.248994822	0.000757448
FBgn0267991	lncRNA:CR46258	2.429674956	0.000996063
FBgn0037506	Mics1	4.306002249	0.001101299
FBti0063070	Cr1a5009	3.44121512	0.001191574
FBgn0267793	lncRNA:CR45232	2.671966201	0.001210699
FBgn0052669	CG32669	2.956497448	0.001384988
FBgn0035720	CG10077	0.644052323	0.001586232
FBgn0264501	lncRNA:CR43900	2.707682895	0.001618862
FBgn0030696	CG8509	3.861832569	0.001974147
FBgn0051038	CG31038	1.70363916	0.002029504
FBgn0267704	lncRNA:flam	1.906941951	0.002192013
FBgn0034327	CG14505	3.630759572	0.002411317
FBgn0053526	PNUTS	0.757849774	0.002611787
FBgn0050323	CG30323	6.606244161	0.002945353
FBti0062182	INE-14121	3.755769635	0.002970634
FBgn0034493	CG8908	4.42579088	0.003110455
FBgn0028577	hfp	1.010509628	0.00368161
FBgn0267320	lncRNA:CR45756	1.90847267	0.003794283
FBgn0031239	CG17075	2.64406225	0.003802198
FBti0062206	jockey4145	3.559175961	0.004182848
FBgn0266308	lncRNA:CR44973	6.314627088	0.004320649
FBgn0051523	CG31523	0.881733183	0.004505762
FBgn0026417	Hus1-like	3.024160198	0.004665721
FBgn0042693	wrd	0.933734571	0.004725876
FBgn0058263	MFS17	0.954016888	0.005066727
FBgn0263290	lncRNA:CR43399	5.429151545	0.005071129
FBgn0032600	BuGZ	0.977428754	0.005992733
FBgn0262976	lawc	1.230549604	0.006160404
FBgn0265042	Irk1	2.288412807	0.007407148
FBgn0266407	lncRNA:CR45047	4.153093414	0.008286598
FBgn0036415	CG7768	2.791992924	0.008573448
FBgn0267656	lncRNA:CR45994	2.142805208	0.010108726
FBgn0035458	ppk27	5.131451257	0.010943773
FBgn0031214	CG11374	3.964073633	0.012245709
FBgn0267781	lncRNA:CR46112	4.976183698	0.012408515
FBgn0032305	CG6700	0.840894472	0.013526085
FBgn0026394	Or24a	5.930770378	0.014209541
FBgn0266918	CG32486	1.187412339	0.014485834
FBti0061469	INE-13408	5.061182864	0.014537261
FBgn0032986	CG3262	1.223422297	0.015157372
FBgn0267601	lncRNA:CR45939	4.205128812	0.017102733
FBgn0267759	lncRNA:CR46090	5.936548104	0.017129429
FBti0060194	INE-12133	3.474700125	0.017691415
FBgn0260934	par-1	0.479645979	0.018289153
FBgn0034700	CG11269	5.857543953	0.018513447
FBti0060191	INE-12130	3.401206454	0.0187541
FBti0063105	Tc1-25044	3.780135407	0.019432245
FBgn0004227	nonA	0.77416567	0.020245571
FBgn0032901	sky	1.146415616	0.02062266
FBgn0003089	pip	3.061588411	0.021169588
FBgn0082582	tmod	1.112701939	0.021821568
FBti0061258	INE-13197	3.484821162	0.022464477
FBti0061812	INE-13751	3.334616188	0.022560094
FBgn0261862	whd	0.765342526	0.022812887
FBgn0032451	spict	1.102159523	0.023089147
FBgn0031374	Wdr62	0.523376127	0.024145256
FBgn0264618	lncRNA:CR43958	2.870213132	0.024275904
FBgn0029838	CG4666	2.830753713	0.024463512
FBgn0263289	scrib	0.849827581	0.025906764
FBgn0032456	MRP	0.549250769	0.026066862
FBgn0265667	lncRNA:CR44474	2.891160065	0.026720151
FBti0063335	INE-15274	3.017147213	0.026884609
FBgn0283500	Sac1	1.224475764	0.02775857
FBgn0260660	Mp	0.959177467	0.0279214
FBgn0040973	CG16824	4.585415968	0.028037417
FBgn0023441	fus	0.507493861	0.028285287
FBgn0011672	Mvl	0.736775451	0.028334971
FBgn0265629	CG44437	2.819802445	0.028455264
FBgn0028563	sut1	1.045947616	0.028459669

FBgn0024320	Npc1a	0.825531664	0.028486657
FBti0061664	Transpac3603	5.572279265	0.030374387
FBgn0041789	Pax	0.837477603	0.030459478
FBgn0263019	lncRNA:CR43314	1.78515404	0.030794267
FBgn0058160	CG40160	0.74747535	0.030966244
FBgn0283741	prage	0.879005766	0.031292734
FBti0063180	invader65119	2.950138944	0.031708159
FBgn0261831	CG42763	4.47247995	0.03194811
FBgn0261267	CG42615	3.801136498	0.032468027
FBti0062608	INE-14547	3.009825249	0.032698986
FBti0060148	INE-12087	4.777724451	0.032921154
FBgn0266414	lncRNA:CR45054	1.052200489	0.033004398
FBgn0039167	CG17786	1.430971708	0.03344953
FBgn0038621	CG10864	4.516721065	0.033597238
FBgn0267336	Glut4EF	0.880033891	0.034035172
FBti0060888	INE-12827	2.895539507	0.034267572
FBti0059683	Cr1a1622	3.314723573	0.034450315
FBti0019900	accord2625	2.969814509	0.034766049
FBti0061227	INE-13166	4.66424141	0.035580751
FBgn0263870	lncRNA:CR43717	3.592247771	0.036395053
FBti0060950	rooA2889	2.418454003	0.037357895
FBgn0050357	CG30357	2.473616062	0.038252578
FBgn0264330	CG43789	3.78776942	0.038315415
FBgn0003256	rl	0.794626583	0.038370661
FBgn0260635	Diap1	0.555575964	0.038473478
FBti0019560	roo41	4.058227006	0.038842279
FBgn0043025	Adgf-A2	2.650379691	0.03898246
FBgn0267376	SelR	0.91413234	0.039408486
FBti0061759	Cr1a3698	5.533008401	0.039738132
FBgn0267073	lncRNA:CR45517	1.705761718	0.039861572
FBgn0052791	DIP-alpha	1.349470184	0.041249886
FBgn0032170	CG4658	0.813471118	0.042325004
FBgn0265316	lncRNA:CR44289	1.94414472	0.042396564
FBgn0027621	Pfrx	1.220327487	0.042523053
FBgn0050354	UQCR-11L	2.364639942	0.04254788
FBgn0020386	Pdk1	0.393444182	0.042596266
FBgn0033074	tomboy40	2.790215013	0.043133977
FBgn0020304	drongo	0.682488977	0.043877595
FBgn0267228	lncRNA:CR45668	3.125220562	0.044375044
FBti0059727	Rt1b1666	5.456723405	0.044905331
FBti0062937	INE-14876	3.69394431	0.045288803
FBgn0259221	ATP8A	1.02789984	0.045628119
FBgn0033482	CG1371	0.803991828	0.046215349
FBgn0264329	CG43788	3.745549081	0.046480445
FBti0060796	diver22735	3.840340934	0.046888284
FBgn0261565	Lmpt	0.724649525	0.046908674
FBgn0261858	Mzt1	3.405201434	0.047088809
FBgn0037203	slif	0.95168228	0.047734815
FBgn0005427	ewg	0.67847198	0.048162858
FBti0060147	INE-12086	4.278301275	0.048452923
FBti0063367	INE-15306	4.470798937	0.049070983
FBgn0259741	CG42395	3.662609773	0.049344984
FBgn0040780	Atg10	5.306568978	0.049350597

Table A.6: Rump CLIP top 20 enriched repeats

Motif	normalized ratio 1	normalized ratio 2	normalized ratio 3	median	Family
Gypsy-21˙DY-I	73.10419841	80.0427521	21.9188379	73.10419841	Gypsy
Gypsy-7˙DVir-I	107.5556023	4.100550825	73.06279301	73.06279301	Gypsy
MDG3˙LTR	65.54169513	12.18449388	71.70977832	65.54169513	Gypsy
Gypsy-6˙DBi-I	64.70141698	14.10589484	66.97422692	64.70141698	Gypsy
MDG3˙I-int	95.89674303	53.36183473	64.02219488	64.02219488	Gypsy
Gypsy-5˙DKi-I	51.46703624	10.82545418	60.88566084	51.46703624	Gypsy
Gypsy-12˙DRh-I	47.23247665	36.14449163	178.5979385	47.23247665	Gypsy
FROGGER˙I-int	28.56945685	46.91030143	100.7995941	46.91030143	Copia
Gypsy-52˙DEl-I	41.4537217	20.33873209	73.06279301	41.4537217	Gypsy
hAT-1N˙DP	82.34725798	13.12176264	40.59044056	40.59044056	hAT

Gypsy-12ˑDSim-I	95.60497978	22.23793458	40.31983762	40.31983762	Gypsy
Gypsy-2ˑDTa-I	38.83952304	20.24500521	34.50187448	34.50187448	Gypsy
BEL-2ˑDSe-I	32.17064889	34.11658286	138.0074979	34.11658286	Pao
Gypsy-24ˑDYa-I	31.46696769	14.72097746	34.07460668	31.46696769	Gypsy
FW3ˑDM	32.35070849	15.9101372	30.78108409	30.78108409	Jockey
Gypsy7ˑI-int	51.08891107	14.30272128	30.10457675	30.10457675	Gypsy
Gypsy5-IˑDya	29.91390188	18.37046769	49.72328969	29.91390188	Gypsy
Invader4ˑLTR	32.65080783	10.98947621	29.28310355	29.28310355	Gypsy
Gypsy11-IˑDya	27.63581447	23.09430224	33.82536713	27.63581447	Gypsy
Gypsy-3ˑDWil-I	33.61112571	10.49741011	27.39854738	27.39854738	Gypsy

Table A.7: Rump CLIP enriched cenRNAs

Motif	normalized ratio 1	normalized ratio 2	normalized ratio 3	median	Family
DOC2ˑDM	34.79373939	12.24697846	23.91030639	23.91030639	Jockey
BicaˑI-int	20.16667542	13.88719879	18.84556169	18.84556169	Gypsy
DMRT1B	22.53746018	11.54334771	13.99773281	13.99773281	R1
DOC	16.61977271	6.921873043	13.04856536	13.04856536	Jockey
TARTˑB1	15.21941984	5.04151933	10.7089673	10.7089673	Jockey
BEL-6ˑDYa-I	10.30018368	6.163252149	12.66421745	10.30018368	Pao
G5ˑDM	10.66787903	5.819216475	11.44650424	10.66787903	Jockey
CopiaˑLTR	3.68121853	3.231841539	9.36198871	3.68121853	Copia
Gypsy-27ˑDYa-I	8.402781426	2.624352528	5.412058741	5.412058741	Gypsy
NOMADˑLTR	5.500002388	3.936528792	5.969182435	5.500002388	Gypsy
TART-A	4.877712145	2.544015206	6.03057974	4.877712145	Jockey
MAXˑI-int	4.509492699	2.271827561	4.974318696	4.509492699	Pao
GˑDM	4.783121735	2.249445024	3.645622902	3.645622902	Jockey
BLASTOPIAˑLTR	2.520834428	2.307620326	3.738593209	2.520834428	Gypsy
260bpˑSAT	19.32639728	4.665515605	8.879158872	8.879158872	Satellite
353bpˑSAT	7.967081649	3.136421314	17.25093724	7.967081649	Satellite
356bpˑSAT	9.747226455	3.561621288	17.50462749	9.747226455	Satellite
359bpˑSAT	13.23806618	4.72383455	13.36421109	13.23806618	Satellite
360bpˑSAT	13.06250567	3.268511785	12.61202975	12.61202975	Satellite
HETRPˑDM	6.242066202	5.4674011	8.40801983	6.242066202	Satellite

Table A.8: Fmr1-CLIP data enriched cenRNAs

Motif	normalized ratio 1	normalized ratio 3	median	Family	Class
DIVER2ˑI-int	11.75157138	2.923757145	7.337664263	Pao	LTR
260bpˑSAT	7.344732112	4.385635718	5.865183915	Satellite	Satellite
BicaˑLTR	5.246237223	5.84751429	5.546875757	Gypsy	LTR
TART-A	6.295484668	4.105193706	5.200339187	Jockey	LINE
TARTˑB1	5.91822716	3.394757536	4.656492348	Jockey	LINE
DOC	5.146057819	3.888903773	4.517480796	Jockey	LINE
Gypsy-7ˑDSe-I	3.497491482	5.012155106	4.254823294	Gypsy	LTR
DODECAˑSAT	3.672366056	4.826519732	4.249442894	Satellite	Satellite
DOC2ˑDM	5.129654174	3.080386992	4.105020583	Jockey	LINE
GˑDM	5.326948565	2.227624492	3.777286529	Jockey	LINE
Gypsy-27ˑDYa-I	5.246237223	2.227624492	3.736930858	Gypsy	LTR
Gypsy-24ˑDYa-I	3.9799041	2.901774009	3.440839055	Gypsy	LTR
359bpˑSAT	3.276597283	2.900114904	3.088356094	Satellite	Satellite
DMRT1B	3.897204794	2.260383675	3.078794235	R1	LINE
Gypsy-8ˑDSim-I	9.79297615	3.75911633	6.77604624	Gypsy	LTR

Table A.9: RNA-Seq data showing differential expression of transposable elements in *Fmr1*[4] mutant testes compared to OreR

Motif	log2FoldChange	padj 3	Family	Class
353bp˙SAT	-1.232902023	4.07E-16	Satellite	Satellite
356bp˙SAT	-3.650029513	3.90E-06	Satellite	Satellite
359bp˙SAT	-1.591045141	8.27E-16	Satellite	Satellite
A-rich	-0.385915468	8.21E-07	Low complexity	Low complexity
ACCORD˙I-int	-2.394759593	7.08E-07	Gypsy	LTR
BATUMI˙I-int	0.474749796	0.023666217	Pao	LTR
BEL˙I-int	0.481387445	0.007254468	Pao	LTR
BEL-17˙DEl-I	-0.552035126	0.00036177	Pao	LTR
BEL-6˙DYa-I	-0.416822059	0.03218075	Pao	LTR
Bica˙I-int	1.500777643	3.45E-12	Gypsy	LTR
Bica˙LTR	2.148099182	0.009505452	Gypsy	LTR
Chimpo˙I-int	-0.869903467	0.00190878	Gypsy	LTR
Chouto˙I-int	-1.261924112	0.003971337	Gypsy	LTR
Copia˙I-int	1.46613196	2.48E-42	Copia	LTR
Copia˙LTR	1.257147416	1.82E-07	Copia	LTR
Copia-3˙DMoj-I	5.255260029	0.003367127	Copia	LTR
DM1731˙I-int	1.350176169	1.16E-13	Copia	LTR
DM297˙I-int	0.653912861	0.004527157	Gypsy	LTR
dmel.rNDA.ITS2	-2.388015174	0.048035594	Unknown	Unknown
DMRT1A	-1.46101963	3.12E-20	R1	LINE
DMRT1B	-0.534678601	0.016417714	R1	LINE
DNAREP2˙Dsim	3.96622004	1.81E-18	Helitron	RC
DOC2˙DM	-0.722685219	5.99E-10	Jockey	LINE
FROGGER˙I-int	-0.453473879	0.016154201	Copia	LTR
FW˙DM	-0.415588735	0.045276397	Jockey	LINE
FW2˙DM	-0.408179563	0.020871749	Jockey	LINE
G˙DM	-0.509682776	0.005035546	Jockey	LINE
G4˙DM	0.520632132	0.00011552	Jockey	LINE
G5˙DM	-0.792949348	7.54E-06	Jockey	LINE
GA-rich	-0.461141422	9.19E-05	Low complexity	Low complexity
Gypsy-10˙DMoj-I	-1.292965897	0.006715933	Gypsy	LTR
Gypsy-10˙DMoj-LTR	-3.121365635	5.05E-05	Gypsy	LTR
Gypsy-11˙DAn-I	-1.311091449	0.001364642	Gypsy	LTR
Gypsy-13˙DEu-I	-1.92052042	0.000833699	Gypsy	LTR
Gypsy-13˙DSim-I	-0.760043467	4.07E-09	Gypsy	LTR
Gypsy-13˙DSim-LTR	-0.413379574	0.004347054	Gypsy	LTR
Gypsy-13˙DVir-I	0.700384301	0.009271097	Gypsy	LTR
Gypsy-15˙DEl-I	-4.108586224	1.24E-10	Gypsy	LTR
Gypsy-16˙DBi-I	1.491725203	0.003297615	Gypsy	LTR
Gypsy-24˙DYa-I	1.461357876	4.73E-09	Gypsy	LTR
Gypsy-25˙DEl-I	-0.407297339	0.006238422	Gypsy	LTR
Gypsy-26˙DBp-I	-1.333405135	0.000593048	Gypsy	LTR
Gypsy-27˙DYa-I	-3.691164994	2.89E-29	Gypsy	LTR
Gypsy-3˙DKi-I	-1.15976147	0.002106366	Gypsy	LTR
Gypsy-31˙DEl-I	-1.158354617	8.32E-11	Gypsy	LTR
Gypsy-31˙DYa-I	-0.706781181	0.004619668	Gypsy	LTR
Gypsy-4˙DSim-I	-0.347809797	0.004581321	Gypsy	LTR
Gypsy-45˙DWil-I	-2.104082872	0.000576154	Gypsy	LTR
Gypsy-6˙DEu-I	-0.662657056	0.042980445	Gypsy	LTR
Gypsy-6˙DSim-I	-0.923471967	2.07E-11	Gypsy	LTR
Gypsy-7˙DBi-I	-1.354393255	4.32E-26	Gypsy	LTR
Gypsy-7˙DSe-I	0.761070098	0.048463688	Gypsy	LTR
Gypsy-7˙DVir-I	-1.167756325	0.034124908	Gypsy	LTR
Gypsy10˙I-int	-0.796815908	3.98E-16	Gypsy	LTR
Gypsy10˙LTR	-0.345367297	0.040197533	Gypsy	LTR
Gypsy12-I˙Dya	0.863590348	2.40E-05	Gypsy	LTR
Gypsy2˙I-int	-0.777184813	0.005157744	Gypsy	LTR
Gypsy3˙I-int	-0.855295854	0.010480423	Gypsy	LTR
Gypsy4˙I-int	1.173038195	9.54E-07	Gypsy	LTR
Gypsy8˙I-int	-0.698584366	0.005143725	Gypsy	LTR
Gypsy9˙I-int	-1.189088685	2.18E-11	Gypsy	LTR
IDEFIX˙I-int	-0.514432908	0.001653405	Gypsy	LTR
Invader2˙I-int	0.487831408	0.042139974	Gypsy	LTR
Invader4˙I-int	1.457375106	0.019728768	Gypsy	LTR
Invader5˙LTR	1.204908793	0.046112482	Gypsy	LTR
Invader6˙I-int	-0.9737108	0.032530678	Gypsy	LTR

Invader6˙LTR	-2.478650073	0.01231718	Gypsy	LTR
Jockey-5˙DK	5.110853024	0.005205933	Jockey	LINE
Jockey-6˙DK	-0.288035094	1.72E-05	Jockey	LINE
Jockey-N1˙Dvi	0.522123049	0.042250392	Jockey	LINE
Jockey2	0.557928543	0.037203824	Jockey	LINE
LINEJ1˙DM	0.894898293	0.022808756	Jockey	LINE
Mariner-1˙DF	-1.24917611	0.004294113	TcMar-Mariner	DNA
Mariner-12˙Dan	1.742762803	0.004145189	TcMar-Mariner	DNA
Mariner2˙DM	0.910308742	0.000819611	TcMar-Mariner	DNA
MICROPIA˙I-int	-0.987119998	0.001242219	Gypsy	LTR
NOMAD˙LTR	1.458950913	0.012503864	Gypsy	LTR
R1˙DM	2.873493281	2.44E-69	R1	LINE
R1-3˙Dya	0.951836845	0.032078624	R1	LINE
R2˙DM	0.905437516	0.04946056	R2	LINE
Rehavkus-1˙DY	-4.182618757	6.72E-05	MULE-NOF	DNA
ROO˙I-int	-0.32937307	0.002408272	Pao	LTR
ROVER-I˙DM	1.717407104	5.10E-07	Gypsy	LTR
Stalker2˙I-int	3.000022191	8.44E-10	Gypsy	LTR
TAHRE	0.964292783	0.001033055	Jocky	LINE
TART˙B1	-1.517016758	8.99E-24	Jockey	LINE
TRANSIB3	-1.84414855	0.00346672	CMC-Transib	DNA

Table A.10: RNA-Seq data showing differential expression of transposable elements in *rump[1]* mutant testes compared to OreR

Motif	log2FoldChange	padj 3	Family	Class
356bp˙SAT	-2.960613757	0.000221432	Satellite	Satellite
ACCORD˙I-int	-1.914880624	0.000119256	Gypsy	LTR
BEL-12˙DEl-I	-1.760267964	0.024836342	Pao	LTR
BEL-4˙DFi-I	1.286924112	0.011067861	Pao	LTR
Bica˙I-int	0.652656389	0.041458935	Gypsy	LTR
Chimpo˙I-int	-1.885599184	2.91E-09	Gypsy	LTR
Copia-1˙DBi-I	1.624430267	0.030939144	Copia	LTR
Copia-2˙DYa-I	-0.380693184	0.044533668	Copia	LTR
Copia-3˙DMoj-I	4.573091982	0.041237432	Copia	LTR
DM1731˙I-int	1.549701734	1.08E-17	Copia	LTR
DMRT1B	-0.602178311	0.004796393	R1	LINE
DMRT1C	-0.928278087	0.022075383	R1	LINE
DOC	-0.236299405	0.027418754	Jockey	LINE
DOC2˙DM	-1.755078763	8.19E-44	Jockey	LINE
FB4˙DM	0.568078295	0.013241707	TcMAr-Tc1	DNA
FROGGER˙I-int	-2.853401741	2.47E-32	Copia	LTR
FW˙DM	-0.735062513	9.35E-05	Jockey	LINE
G5˙DM	-0.932037417	8.41E-07	Jockey	LINE
G7˙DM	-6.240038087	3.38E-14	Jockey	LINE
GA-rich	-0.349193146	0.005138609	Low complexity	Low complexity
Gypsy-1˙DSim-I	-0.57444836	3.98E-05	Gypsy	LTR
Gypsy-13˙DEu-I	-2.667993457	5.25E-05	Gypsy	LTR
Gypsy-2˙DAn-I	-0.637350736	0.013230599	Gypsy	LTR
Gypsy-2˙DTa-I	-2.323732372	0.017700913	Gypsy	LTR
Gypsy-23˙DYa-I	-0.962581199	0.044599559	Gypsy	LTR
Gypsy-26˙DYa-I	-0.433370917	0.001769443	Gypsy	LTR
Gypsy-27˙DYa-I	-3.77145848	1.15E-24	Gypsy	LTR
Gypsy-31˙DEl-I	-0.852700737	8.39E-06	Gypsy	LTR
Gypsy-31˙DYa-I	-0.746480069	0.004432298	Gypsy	LTR
Gypsy-45˙DWil-I	-2.46465388	0.000802317	Gypsy	LTR
Gypsy-6˙DSim-I	-0.364717251	0.033729285	Gypsy	LTR
Gypsy-7˙DBi-I	-1.360780298	2.47E-26	Gypsy	LTR
Gypsy-7˙DVir-I	-1.65684407	0.010852371	Gypsy	LTR
Gypsy-8˙DSe-I	-1.03729507	0.048503862	Gypsy	LTR
Gypsy10˙I-int	-0.331565438	0.009986976	Gypsy	LTR
Gypsy10˙LTR	-0.428946812	0.017717301	Gypsy	LTR
Gypsy12˙I-int	-0.635895099	0.004199026	Gypsy	LTR
Gypsy12-I˙Dya	-0.941003749	0.001853675	Gypsy	LTR
Gypsy3˙I-int	-0.795468981	0.03620516	Gypsy	LTR
Gypsy4˙I-int	0.687691176	0.027095548	Gypsy	LTR
HeT-A3˙Dvi	1.097016134	0.043414482	Unknown	LINE

HETA	-0.39994635	9.71E-05	Jockey	LINE
HOBO	0.970196898	0.024261044	hAT-hobo	DNA
I˙DM	-0.694838497	0.037308355	I	LINE
IDEFIX˙I-int	-0.874470769	1.88E-05	Gypsy	LTR
Jockey-12˙DBp	-3.542749537	5.56E-13	Jockey	LINE
Jockey-2˙Dan	-2.711144764	3.37E-24	Jockey	LINE
Jockey-3˙Dgri	-0.141260778	0.005597711	Jockey	LINE
Jockey-3˙Dper	-0.571687781	0.020050989	Jockey	LINE
Jockey-5˙DK	5.442979687	0.002230803	Jockey	LINE
Jockey-6˙DK	-0.228170448	6.64E-05	Jockey	LINE
LINEJ1˙DM	1.36030889	5.16E-05	Jockey	LINE
Mariner-9˙Dan	-0.479033584	0.000921164	TcMar-Mariner	DNA
Mariner2˙DM	1.382318432	2.35E-08	TcMar-Mariner	DNA
MICROPIA˙I-int	-0.797425988	0.018646918	Gypsy	LTR
P˙DG	2.095133931	0.004638534	P	DNA
POGON1	4.870228689	0.004627194	TcMar-Pogo	DNA
QUASIMODO˙I-int	0.50442811	0.02871762	Gypsy	LTR
QUASIMODO˙LTR	-0.497849081	0.028933552	Gypsy	LTR
R1˙DM	1.167358527	3.27E-07	R1	LINE
R1˙Dmo	-2.542460945	0.021024989	R1	LINE
R2˙DM	-1.958709931	0.002356668	R2	LINE
ROO˙I-int	-0.292514632	0.00609547	Pao	LTR
ROVER-I˙DM	0.956236384	0.044426009	Gypsy	LTR
TAHRE	1.419046756	8.38E-08	Jockey	LINE
TART˙B1	-1.452523011	1.92E-19	Jockey	LINE
ZAM˙I-int	-0.531927747	0.019448845	Gypsy	LTR

Table A.11: RNA-Seq data showing differential expression of transposable elements in Zhr^1 mutant testes compared to OreR

Motif	log2FoldChange	padj 3	Family	Class
353bp˙SAT	-0.928471109	2.23E-07	Satellite	Satellite
359bp˙SAT	-1.245720792	3.72E-08	Satellite	Satellite
ACCORD˙I-int	-1.978028288	5.21E-05	Gypsy	LTR
Baggins1	-0.706616668	0.030552298	LOA	LINE
BEL1-I˙DV	-0.885537825	3.08E-05	Pao	LTR
Bica˙I-int	1.135831382	8.07E-06	Gypsy	LTR
BLASTOPIA˙I-int	-1.310482257	1.74E-06	Gypsy	LTR
BLASTOPIA˙LTR	-1.284603463	0.020734792	Gypsy	LTR
BLOOD˙LTR	3.477386534	0.049936273	Gypsy	LTR
Chimpo˙I-int	-1.341410063	9.00E-06	Gypsy	LTR
CIRCE	-1.01907568	0.046636717	Gypsy	LTR
Copia˙I-int	-0.67861982	9.02E-07	Copia	LTR
Copia˙LTR	-0.886190518	0.018923061	Copia	LTR
Copia-1˙DSe-I	-1.16294472	0.042366869	Copia	LTR
Copia-2˙DKi-I	-4.425706175	0.025632022	Copia	LTR
Copia-2˙DPer-I	-1.539297139	0.000617681	Copia	LTR
Copia-3˙DBp-I	-0.669526811	0.035554846	Copia	LTR
Copia-3˙DMoj-I	6.74457059	5.08E-06	Copia	LTR
DM1731˙I-int	1.05350462	4.50E-06	Copia	LTR
DMCR1A	-0.478568019	0.005546628	CR1	LINE
dmel.rNDA.ITS2	-2.680489535	0.047762477	Unkown	Unknown
DMRT1B	-0.855888807	1.13E-05	R1	LINE
DNAREP1˙DM	-0.716371877	7.44E-12	Helitron	RC
DOC	-0.528802556	3.83E-07	Jockey	LINE
DOC2˙DM	-2.309624326	6.51E-66	Jockey	LINE
FB4˙DM	0.52744096	0.02605729	TcMar-Tc1	DNA
FW˙DM	0.512479258	0.002389435	Jockey	LINE
FW2˙DM	-0.497146865	0.018363364	Jockey	LINE
G˙DM	-1.608903113	6.27E-16	Jockey	LINE
G2˙DM	-0.732768441	0.036302143	Jockey	LINE
G5˙DM	-0.810759986	6.24E-06	Jockey	LINE
G5A˙DM	-0.672285027	2.99E-06	Jockey	LINE
G6˙DM	4.114191282	7.69E-46	Jockey	LINE
GA-rich	-0.433275915	0.000203461	Low complexity	Low complexity
Gypsy-1˙DSim-I	-0.561988722	9.22E-05	Gypsy	LTR
Gypsy-10˙Dmoj	-1.625982397	0.029517422	Gypsy	LTR

Gypsy-12˙DFi-I	-0.906339163	0.028769878	Gypsy	LTR
Gypsy-12˙DSim-I	-0.687392553	0.019063994	Gypsy	LTR
Gypsy-13˙DEu-I	-1.165565043	0.045994397	Gypsy	LTR
Gypsy-13˙DSim-I	-1.744434619	7.08E-37	Gypsy	LTR
Gypsy-13˙Dsim	-1.448778549	1.81E-22	Gypsy	LTR
Gypsy-2˙DAn-I	-1.717430984	1.58E-10	Gypsy	LTR
Gypsy-25˙DEl-I	-0.425062778	0.002715715	Gypsy	LTR
Gypsy-26˙DYa-I	-0.731500386	6.25E-08	Gypsy	LTR
Gypsy-27˙DYa-I	-3.718167148	8.59E-26	Gypsy	LTR
Gypsy-3˙DKi-I	-1.071220377	0.009524607	Gypsy	LTR
Gypsy-31˙DYa-I	-0.763299876	0.003393895	Gypsy	LTR
Gypsy-4˙DSim-I	-0.802396736	1.05E-10	Gypsy	LTR
Gypsy-45˙DWil-I	-3.43856443	1.97E-05	Gypsy	LTR
Gypsy-5˙DSe-I	-0.857963271	0.024090599	Gypsy	LTR
Gypsy-52˙DEl-I	-1.914086307	0.017102404	Gypsy	LTR
Gypsy-6˙DSe-I	-1.030875064	0.008638976	Gypsy	LTR
Gypsy-6˙Dse	2.272722855	0.007957785	Gypsy	LTR
Gypsy-7˙DVir-I	-1.409248757	0.024797921	Gypsy	LTR
Gypsy-8˙DSe-I	-1.964432109	0.000241035	Gypsy	LTR
Gypsy10˙I-int	-0.409414314	0.001333635	Gypsy	LTR
Gypsy12˙I-int	-0.91869641	4.69E-05	Gypsy	LTR
Gypsy12-I˙Dya	-1.039632917	0.000510887	Gypsy	LTR
Gypsy2-I˙DM	-0.490155918	0.019896626	Gypsy	LTR
Gypsy8˙I-int	-0.769174007	0.003216425	Gypsy	LTR
Gypsy8-I˙Dpse	-0.201647227	0.007787895	Gypsy	LTR
Gypsy9˙I-int	-1.406341685	1.88E-14	Gypsy	LTR
Gypsy9˙LTR	-1.457291521	0.012118892	Gypsy	LTR
I˙DM	-0.684882926	0.031499221	I	LINE
IDEFIX˙I-int	-0.386139791	0.032300736	Gypsy	LTR
IVK˙DM	-1.064786656	0.018088356	I	LINE
Jockey-2˙Dan	-0.593455767	0.013882436	Jockey	LINE
Jockey-3˙Dgri	-0.204899078	0.000247844	Jockey	LINE
Jockey-3˙DRh	-1.296553442	0.001856589	Jockey	LINE
Jockey-6˙DK	-0.37039617	6.38E-10	Jockey	LINE
Jockey-N1˙Dvi	-0.831144141	0.011139986	Jockey	LINE
LINEJ1˙DM	0.95002154	0.025988083	Jockey	LINE
Mariner-4˙DBp	-1.554578644	2.95E-07	TcMar-Mariner	DNA
Mariner-9˙Dan	-0.692805765	2.91E-05	TcMar-Mariner	DNA
MDG1˙I-int	-0.793640034	0.001331535	Gypsy	LTR
NINJA˙I-int	-0.6887463	0.000991863	Pao	LTR
NOMAD˙I-int	-0.666657359	5.18E-09	Gypsy	LTR
POGO	1.289663213	0.026987009	TcMar-Pogo	DNA
POGON1	4.93579167	0.003963298	TcMar-Pogo	DNA
QUASIMODO˙I-int	0.603020341	0.004443297	Gypsy	LTR
R1˙DM	-1.07780396	0.002227096	R1	LINE
R1-2˙DM	-1.032783193	0.033855351	R1	LINE
R2˙DM	1.161425164	0.010668149	R2	LINE
ROO˙I-int	-0.442128673	4.51E-05	Pao	LTR
Stalker2˙LTR	4.781259556	0.006269749	Gypsy	LTR
TAHRE	1.510391909	2.53E-09	Jockey	LINE
TART˙B1	-1.010661753	2.19E-11	Jockey	LINE
TRANSIB2	0.587897902	9.00E-06	CMC-Transib	DNA
ZAM˙I-int	-0.746655399	0.001300846	Gypsy	LTR

List of Abbreviations

APC	anaphase promoting complex
Aub	Aubergine
bp	basepairs
Cal1	chromatin alignment defect 1
CATD	CENP-A targeting domain
cDNA	complementary DNA
CENP-A	centromere protein-A
cenRNA	centromeric RNA
CHIP	chromatin immunoprecipitation
CID	centromere identifier
CLIP	cross-linking immunoprecipitation
DDR	DNA damage response
DIC	differential interference contrast
DNA	deoxyribonucleic acid
dsRNA	double-stranded RNA
EMSA	electrophoretic Mobility Shift Assay
FACT	facilitates chromatin transcription
FISH	fluorescence in situ hybridization
Fmr1	fragile X mental retardation 1
FXPOI	fragile X-associated primary ovarian insuciency
FXS	fragile X Syndrome
FXTAS	fragile X-associated tremor/ataxia syndrome
h	hours
hnRNP	heterogeneous nuclear ribonucleoproteins
HOR	higher order repeats
HP1	heterochromatin-associated Protein 1
IDR	intrinsically disordered region
IGS	intergenic spacer

IF	immunofluorescence
IP	immunoprecipitation
kb	kilobase
KH domain	K-homology domain
KMN	KNL1/Spc105, the Mis12 complex and the Ndc80 complex
LC-MS/MS	liquid chromatography-tandem mass spectrometry
LNA	locked nucleic acid
lncRNA	long non-coding RNA
LTR	long terminal repeat
m^6A	N^6-Methyladenosine
MCC	mitotic checkpoint complex
NES	nuclear export signal
NLS	nuclear localization signal
MNase	micrococcal nuclease
mRNA	messenger RNA
MZT	maternal to zygotic transition
nm	nanometer
pAbp	poly(A)-binding protein
PCR	polymerase chain reaction
piRNA	piwi-interacting RNA
pmol	pikomolar
PTM	post-translational modification
RACE	rapid amplication of cDNA ends
RBP	RNA-binding protein
RGG	arginine/glycine-rich
RIP	RNA Immunoprecipitation
RISC	RNA-induced silencing complex
RITS	RNAi transcriptional silencing complex
RNA	ribonucleic acid
RNAi	RNA interference
RNase	ribonuclease
RNA-Seq	RNA-Sequencing
RNP	ribonucleoprotein particle
RT	room temperature
RT-qPCR	real-time quantitative PCR
rRNA	ribosomal RNA

RRM	RNA recognition motif
Rsp	responder
Rump	rumpelstiltskin
S2	*Drosophila* Schneider 2 cell line
SAC	spindle assembly checkpoint
SatIII	satellite III
Sgo1	shugoshin
siRNA	small interfering RNA
TAD	topologically associating domain
TE	transposable element
UTR	untranslated region
UV	ultraviolet
wt	wildtype
μg	mikrogram
μl	mikroliter
μM	mikromolar
μm	mikrometer

List of Figures

154

List of Tables

155

Acknowledgements

After being absorbed by the fantastic world of RNA-binding proteins in the context of centromere regulation, I would like to take a minute to express my deepest gratitude to all the people who have supported and encouraged me over the past years as a PhD student. You all have made big contributions to this thesis, be it scientific or otherwise.

I would like to thank my supervisor Prof. Dr. Sylvia Erhardt for giving me the opportunity to work on this exciting project in an extremely interesting and challenging field. Dear Sylvia, thanks for always being there with an open door, whether in person or via Zoom, for any type of discussion. Above all, for giving me motivation and comforting me when things didn't work out as planned. Thank you also very much for granting me so much freedom to develop my own ideas and helping me pursue them. It was a pleasure to be part of your team and I definitely learned a lot!

I would like to also express my gratitude to the members of the thesis advisory committee, including Prof. Dr. Frank Lyko and Prof. Dr. Aurelio Teleman for fruitful discussions and suggestions. Moreover, I am indebted to Aurelio for allowing me to work in his lab after the ZMBH was inaccessible and giving me the opportunity to continue my PhD without interruption. I would also like to thank all the defense committee members, in particular Prof. Dr. Dr. Georg Stoecklin, for his interest in this thesis and willingness to be my second referee.

I am very grateful to all the Teleman lab members for kindly accommodating me in their lab and helping me out whenever needed.

I would like to acknowledge both the EMBL proteomics facility for carrying out the mass spectrometry analysis, and the ZMBH sequencing unit in the Bukau lab for performing the sequencing of the screens presented here. In