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.) Masroor Ahmad Kahloon

born in: Frankfurt am Main, Germany

Oral examination:

The role of Exon Junction Complex (EJC) and SOLE in *oskar* mRNA transport

Referees:

Prof. Dr. Dr. Georg Stoecklin

Dr. Alexander Aulehla

“In the creation of the heavens and the earth and in the alteration of the night and day there are indeed signs for men of understanding.”

(3:191)

Table of contents

Acknowledgements

The pursuit of knowledge and learning is my driving force to pursue a PhD. Not only is it the duty of every Muslim to increase their knowledge, but it is also a key factor for helping mankind and improving society. I thank God for the opportunity to receive education. In addition to the self-study of literature, knowledge is also gained through mentorship and the experiences of kind individuals who accompany you on the educational path.

First and foremost, I would like to thank my supervisor Dr. Anne Ephrussi for taking me as her student. Her invaluable guidance, encouragement, support and dedication have helped me to develop and refine my ideas, skills and knowledge not only for research but for life itself. I am deeply indebted to her for her mentorship and friendship. I will always hold dear the time I spent in the Ephrussi lab, as well as the support of current and past members of the lab on my academic and personal journey. I especially want to mention my friend and colleague Mainak Bose. He helped me countless times, was available for discussion and feedback and trained me in several ways. The scientific environment at EMBL, along with its various core facilities, provided an incredibly enriching experience. The protein core, genomic core, and protein purification core facilities provided me with significant support during my research.

I am eternally grateful to my wife Tamsila and our two children, Nabeel and Sabiqa, for their love, patience and understanding. Their unwavering support, encouragement and sacrifices have been instrumental in enabling me to pursue my academic and professional goals. I am blessed to have them in my life.

The foundations of my scientific career were laid by my parents, who came to Germany as exiles to practice our faith without restriction. I don't know if I would have the opportunity to pursue a PhD in Pakistan, as even today the place of birth determines the perspective development opportunities. I will be forever grateful to them for sacrificing their lives for me and my siblings. They have instilled in me the values of hard work, perseverance and humility, which have been crucial to my achievements.

I cannot put into words how grateful I am to Allah for His countless blessings, guidance and protection throughout my life. His grace has enabled me to overcome challenges, learn from my mistakes and appreciate the beauty and diversity of His creation. All hardships, every fear and every challenge I have only managed because of Allah. I pray that He continues to guide me on the path of righteousness and success.

Finally, I would like to express my deep appreciation to the Khalif of the Ahmadiyya Movement, Hadhrat Mirza Masroor Ahmad[ABA], for his leadership, wisdom and inspiration. He is my role model encouraging me to grow as a scientist and serve humanity. His prayers lift the weight from my shoulders and ease my path to the future. May God give me further opportunities to gain knowledge, serve God and help humanity with my knowledge and actions.

رَبِّ زِدْنِیْ عِلْمًا

"O my Lord! Increase me in knowledge" (20:115)

Summary

Asymmetric localization of specific RNAs is essential in early embryo development. During *Drosophila melanogaster* oogenesis, *oskar* mRNA is transcribed in nurse cells, transported to the transcriptional silent oocyte and subsequently localized at the posterior pole. The first transport step from the nurse cells into the oocyte is mediated by dynein transport, whereas the second step to the posterior pole is dependent on kinesin. At the posterior pole, *oskar* mRNA is required for the formation of germ plasm and abdominal segments.

The transport of *oskar* mRNA by kinesin requires the exon junction complex (EJC) core components eIF4AIII, Mago nashi, Y14 and Barentsz as well as the spliced *oskar* localization element (SOLE). The exact mechanism by which the EJC and SOLE are involved in kinesin transport has remained unclear. I utilized RNA-protein interaction footprinting to analyze how the protein occupancy on the RNA changes upon disrupting the SOLE secondary structure. Through RNA immunoprecipitation against Mago nashi, I discovered that the presence of the SOLE led to a more stable binding of the EJC to *oskar* mRNA. This demonstrates that the SOLE has a stabilizing effect on EJC binding to *oskar* mRNA. Although the EJC and SOLE were found to be dispensable for kinesin recruitment to *oskar* mRNA, they were essential for *oskar* mRNA transport, suggesting that the EJC-SOLE interactome plays a role in activating kinesin transport to the posterior pole. I optimized a previously described transcript-specific RNA binding protein (RBP) capture protocol and utilized it to analyze the EJC-SOLE interactome. Flies expressing transgenic *lacZ* RNA containing the *oskar* 3'UTR and either wild-type or mutant SOLE were created and used to determine the difference in RBP composition upon disrupting the SOLE secondary structure. I identified RBPs that were specifically bound to the transgenic *lacZ* RNA with the wild-type SOLE, which localizes to the posterior pole and RBPs bound to the *lacZ* RNA with a disrupted SOLE, which does not localize. Further investigations will be carried out to explore which of the RBPs are involved in the kinesin-mediated transport of *oskar* mRNA.

Zusammenfassung

Die asymmetrische Lokalisierung spezifischer RNAs ist für die frühe Embryonalentwicklung wichtig. Während der Oogenese von *Drosophila melanogaster* wird *oskar* mRNA in Nährzellen transkribiert und in die Eizelle transportiert und anschließend am posterior Pol lokalisiert. Der erste Transportschritt von den Nährzellen in die Eizelle wird durch dynein Transport ausgeführt, während der zweite Schritt zum posterior Pol durch kinesin vollzogen wird. Am posterior Pol wird *oskar* mRNA für die Bildung des Keimplasmas und der Abdominalsegmente benötigt.

Für den Transport von *oskar* mRNA durch kinesin sind die Kernkomponenten des Exon Junction Complex (EJC) eIF4AIII, Mago nashi, Y14 und Barentsz sowie das spliced *oskar* localization element (SOLE) erforderlich. Der genaue Mechanismus, wie das EJC und SOLE am kinesin Transport beteiligt sind, blieb bisher unklar. Ich habe RNA-Protein Interaktions-Footprinting verwendet um zu analysieren, wie sich die Proteinbesetzung auf der RNA bei einer Veränderung der SOLE-Sekundärstruktur ändert. Durch RNA-Immunpräzipitation von Mago nashi entdeckte ich, dass die Anwesenheit des SOLE zu einer stabileren Bindung des EJC an *oskar* mRNA führte. Dies demonstriert, dass das SOLE eine stabilisierende Wirkung auf die Bindung des EJC an *oskar* mRNA hat. Obwohl EJC und SOLE für die Rekrutierung von kinesin an *oskar* mRNA entbehrlich sind, sind sie für den Transport von *oskar* mRNA essentiell, was darauf hindeutet, dass das EJC-SOLE-Interaktom eine Rolle bei der Aktivierung von kinesin-Transport zum posterior Pol spielt. Ich optimierte ein zuvor beschriebenes transkript-spezifisches RNA-Bindungsprotein (RBP) Erfassungsprotokoll und setzte es zur Analyse des EJC-SOLE-Interaktoms ein. Fliegen, die transgene *lacZ*-RNA exprimieren, die *oskar* 3'UTR und entweder Wildtyp oder mutiertes SOLE enthalten, wurden erzeugt und verwendet, um den Unterschied in der RBP-Zusammensetzung nach Auflösung der SOLE-Sekundärstruktur zu bestimmen. Ich identifizierte RBPs, die jeweils spezifisch an die transgene *lacZ*-RNA mit dem Wildtyp-SOLE welches am posterior Pol lokalisiert und spezifisch an die *lacZ*-RNA mit einem mutierten SOLE welches nicht lokalisierte banden. In weiteren Untersuchungen soll geklärt werden, welche der RBPs am kinesin vermittelten Transport von *oskar* mRNA beteiligt sind.

List of abbreviations

AP	anteroposterior
APEX	Engineered ascorbate peroxidase
ASO	Antisense oligonucleotide
ATPase	Adenylpyrophosphatase
BicD	Bicaudal-D
CLIP	Crosslinking immunoprecipitation
Co-IP	Complex immunoprecipitation
DL^{Lz}	Distal loop with *lacZ* mutations
DMS-MaPseq	Dimethyl sulfate mutational profiling with sequencing
dsRNA	double-stranded RNA
DTT	Dithiothreitol
Egl	Egalitarian
eIF4AIII	eukaryotic translation initiation factor 4A3
EJC	Exon junction complex
GO	Gene ontology
hid	*head involution defective*
HITS-CLIP	High-throughput sequencing of RNA isolated by crosslinking immunoprecipitation
hs	heat-shock
IL	intronless
IP	Immunoprecipitation
lncRNA	long non-coding RNA
MCP	MS2 coat protein
miRNA	microRNA
mRNA	messenger RNA
MSL^{Lz}	Medial stem loop with *lacZ* mutations
nt	Nucleotide
OES	Oocyte entry signal
osk^{null}	*oskar* RNA-null
PBS	Phosphate Buffered Saline
PCR	Polymerase chain reaction
PS^{Lz}	Proximal stem with *lacZ* mutations
PS^{Lzc}	Proximal stem with complementary *lacZ* mutations to PS^{Lz}
PYM	Partner of Y14 and Mago
RBD	RNA binding domain
RBP	RNA binding protein
RIP	RNA immunoprecipitation
RIP-seq	RNA immunoprecipitation sequencing
RISC	RNA-induced silencing complex

RNAi	RNA interference
RNP	Ribonucleoprotein
RRM	RNA recognition motif
rRNA	robosomal RNA
RT	Room temperature
siRNA	small interfering RNA
SL	Mutations in the SOLE stem-loop
SLc	Mutations in the SOLE stem-loop complementarity to SL
smFISH	Single molecule fluorescence *in situ* hybridization
SOLE	Spliced *oskar* localization element
SRS	Staufen Recognized Structures
ssRNA	single-stranded RNA
TAS	Transport and Anchoring Signal
TDE	2,2' thiodiethanol
TdT	Terminal deoxynucleotidyl transferase
TM1-I/C	Tropomyosin-1 I/C
TMT	Tandem mass tag
TRiP	Transgenic RNAi Project
tRNA	transfer RNA
UTR	Untranslated region
UV	Ultraviolet
WT	wild-type

1. Introduction

The study of molecular processes at various levels, from organelles to the smallest building blocks such as DNA, RNA, proteins and lipids, allows us to gain insights into many biological, cellular and molecular phenomena that are present in all living organisms. Among these building blocks, the nucleic acids DNA and RNA play a crucial role in storing and transmitting genetic information. However, RNA is involved in a variety of other cellular processes beyond just protein translation. Several types of RNA exist such as protein-coding messenger RNA (mRNA), transfer RNA (tRNA), ribosomal RNA (rRNA), long non-coding RNA (lncRNA) and various small RNAs. Each class of RNA has distinct functions in the cell and plays a role in different cellular processes, such as protein synthesis, transcriptional regulation and post-transcriptional regulation. Over the last two decades, many novel functions of RNAs have been discovered. Given the importance and impact of RNA in the cell, they undergo various post-transcriptional modifications that can affect RNA stability, structure and translation. These modifications are crucial for regulating the expression of genes and ensuring that the cell functions properly

1.1. Overview of RNA localization

An important mechanism in post-transcriptional regulation is intracellular RNA localization, for achieving spatial and temporal control of gene expression. RNA localization is conserved from bacteria to higher eukaryotes (Figure 1.1) and plays a critical role in biological processes such as embryonic development, neuronal function and immune responses (Johnstone and Lasko 2001; King et al. 2012; Smith 2004; Lin and Holt 2007; Martin and Zukin 2006; Uehata and Takeuchi 2020; Gadir et al. 2011; Tian and Oktia 2014). In addition, certain RNAs are localized to the perinuclear region or cell periphery, where they regulate cell migration, polarity and division (Gavis and Lehmann 1992; Katz et al. 2012; Groisman et al. 2000). In recent years, RNA localization has gained increasing attention as a key mechanism for regulating gene expression and cellular behavior.

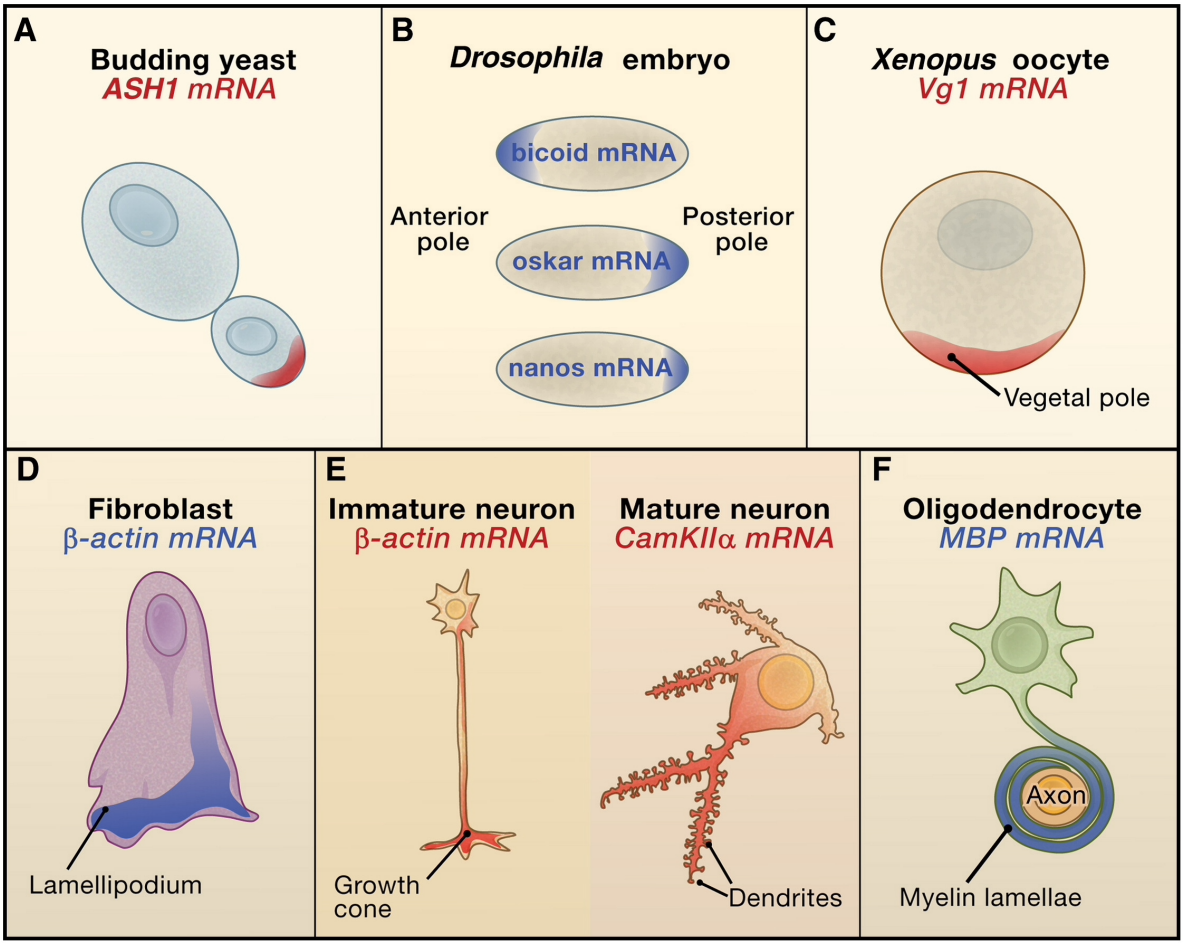

Figure 1.1 RNA localization in different organisms and cells. Specific mRNAs are localized in (A) *Saccharomyces cerevisiae*, (B) *Drosophila melanogaster* embryos, (C) *Xenopus* oocytes, (D) mammalian fibroblasts, (E) mammalian neurons and (F) mammalian oligodendrocytes (Image from Martin & Ephrussi 2009).

RNA localization is important for the functional polarization of cells. In many cell types, such as neurons and epithelial cells, specific RNAs are asymmetrically distributed to different cell regions, contributing to the formation of specialized cellular structures (Mili and Macara 2009; Das et al. 2021). In neurons, RNA localization is essential for the establishment and maintenance of axonal and dendritic compartments, which are critical for proper neuronal function and connectivity. For example, localized RNAs, such as *MAP2* mRNA, are selectively transported to the dendritic compartments and contribute to the formation of dendritic spines, which is essential for synaptic plasticity (Garner et al. 1988; Holt et al. 2019). On the contrary, mRNAs encoding proteins involved in axonal transport and neurotransmitter release are selectively transported to the axonal compartment (Litman et al. 2003; Gervasi et al. 2016). RNA localization is also important in regulating gene expression in response to

environmental cues. In neurons, since the distance between the axon and the nucleus can be extremely large, the mechanism of RNA localization allows rapid translational response independent of the ongoing transcription in the nucleus (Doyle and Kiebler 2011).

RNA localization plays a critical role in the regulation of developmental processes, including the patterning of embryonic tissues and the specification of cell fates. In zebrafish embryos, the specific localization of mRNAs encoding transcription factors and signaling molecules is essential for the specification of different cell types (Lee et al. 2013; Howley and Ho 2000; Holler et al. 2021). Similarly, in *Drosophila melanogaster* embryos, mRNA localization plays a crucial role in establishing the anteroposterior and dorsoventral axes. *bicoid* RNA is localized to the anterior of the embryo, where it specifies the development of the head and thorax while, *oskar* RNA is localized to the posterior pole of the embryo, where it specifies the development of the abdomen (Figure 1.1; Frohnhöfer and Nüsslein-Volhard 1986; Lehmann and Nüsslein-Volhard 1986). A global analysis of RNA localization of the *Drosophila* transcriptome revealed that several thousand RNAs exhibit specific subcellular localization in early embryos indicating the prominent role of RNA localization during organismal development (Lécuyer et al. 2007).

Mislocalization of RNAs can cause severe developmental disorders and can be lethal. Defects in RNA transport are known to underlie several neurodegenerative diseases. Spinal muscular atrophy is caused by mutations in the *SMN* gene, which leads to a reduction in the levels of SMN protein and impaired RNA transport in motor neurons (Rossoll et al. 2002; Fallini et al. 2011). Fragile X syndrome is caused by the absence of FMRP, which regulates the localization of mRNAs involved in synaptic function (Bassell and Warren 2013). Huntington's disease is caused by a mutation in the *HTT* gene, which impairs the transport of mRNAs that are required for neuronal survival (Zuccato & Cattaneo 2009; Savas et al. 2010).

Therefore, RNA localization is a fundamental mechanism for regulating gene expression and cellular behavior. Understanding how RNA molecules are localized and how this process is regulated is critical for deciphering the complex mechanisms underlying different biological processes and developing new strategies for the treatment of various diseases.

1.2. Different modes of RNA localization

The localization of RNA molecules to specific subcellular compartments, such as the cytoplasm, nucleus or organelles is critical for the proper functioning of cells. There are several different modes of RNA localization, each of which is achieved by a distinct set of mechanisms. The most commonly known RNA localization modes are diffusion and anchoring, asymmetric degradation, and active transport along the cytoskeleton (Figure 1.2; Jansen and Niessing 2012).

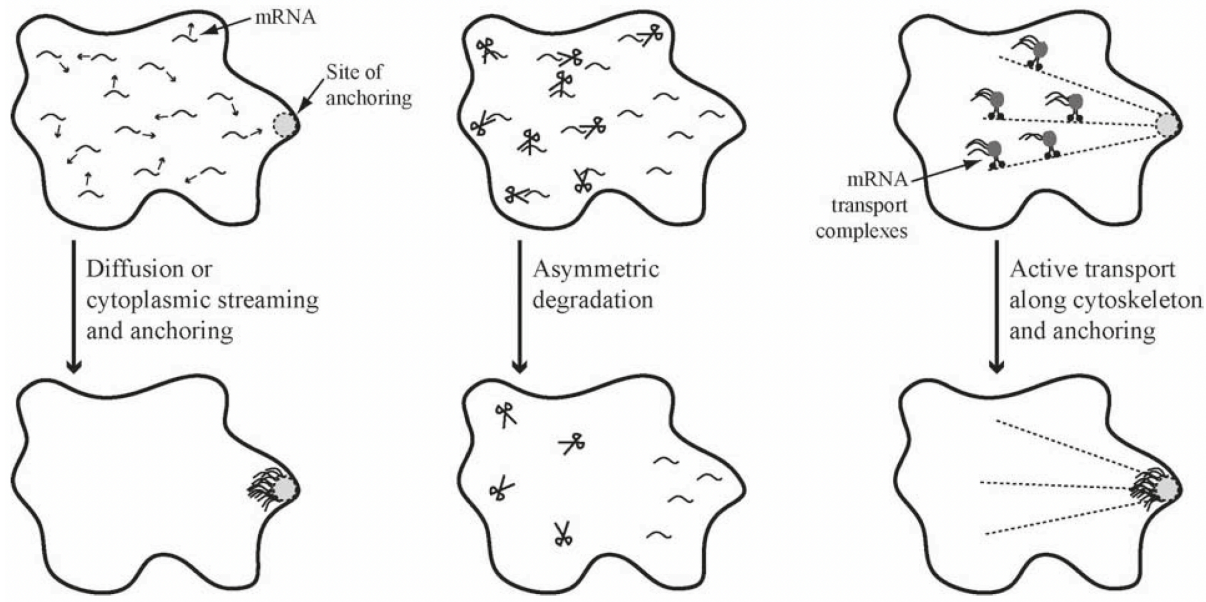

Figure 1.2 Different mechanisms for achieving subcellular RNA localization. Anchoring RNA at a specific site is an important strategy for concentrating RNA in a specific region. This can be coupled to preceding random diffusion or active transport. Asymmetric degradation prevents RNAs from localizing to undesired regions and concentrates them at specific sites. The most commonly studied mechanism is active RNA transport by motor proteins along the cytoskeleton with subsequent anchoring (Image from Jansen and Niessing 2012).

Diffusion-based localization is a simple, passive and low-energy mode of RNA transport. Especially small RNAs, such as microRNAs, can diffuse efficiently in the cell due to their small size (Levine et al. 2007). Diffusion is often coupled to anchoring-based localization, which involves the attachment of RNA molecules to specific subcellular structures or regions. The anchoring is often mediated by RNA-binding proteins that are

already localized subcellularly and recognize specific *cis*-acting sequences or structures within the RNA. This mode of transport is used for RNAs that need to be localized to a specific subcellular region, such as the site of protein synthesis (Vanzo and Ephrussi 2002; Glotzer et al. 1997; Gonzalez et al. 1999; Bloom and Beach 1999). For example, *nanos* mRNA is localized at the posterior pole in the *Drosophila* oocyte through diffusion and anchoring by the germ granules (Forrest and Gavis 2003). Another example of diffusion and local entrapment is the *Xcat2* and *Xdaz1* mRNA in the Balbiani body in the developing *Xenopus* oocyte (Chang et al. 2004; King et al. 2012).

Another mechanism to achieve subcellular RNA localization is through asymmetric RNA degradation. This is achieved when an RNA destabilizing enzyme is not uniformly distributed in the cell or if a degradation-protecting RBP is subcellularly localized. For example, *hsp83* RNA is evenly distributed in *Drosophila* oocytes, but its specific localization to the pole plasm occurs as a result of the asymmetric degradation of *hsp83* RNA during the maternal to zygotic transition. This is achieved by the *hsp83* binding protein Smaug, which recruits a deadenylase complex leading to RNA destabilization. (Ding et al. 1993; Semotok et al. 2005). Another example is *nanos* mRNA which is also bound by Smaug leading to mRNA degradation. In fact, only 4% of *nanos* mRNA is stably localized at the posterior pole whereas the rest is degraded in the cytoplasm (Bergsten & Gavis 1999). Although achieving RNA localization through degradation mechanisms comes at an energetic cost, it is an efficient method for localizing RNAs to specific regions.

An important and most studied mode of RNA localization is by active transport along polarized cytoskeletal elements such as microtubules and actin filaments. As such, this mode of RNA localization was characterized by perturbing the cytoskeleton using drugs. The current notion is that long and short-range active RNA transport use different mechanisms. Short-range transport is facilitated by actin filaments while long-range transport occurs usually along microtubules (Bloom and Beach 1999; Engel et al. 2020). Active transport is often used for the localization of large mRNAs or RNPs that are too large to diffuse rapidly in the cytoplasm. The direction and speed of transport are often regulated by *cis*-acting sequences within the RNA that serve as docking sites for *trans*-acting factors called adaptor proteins that link the RNA to cytoskeletal motor proteins. The motor protein can be, depending

on the site of localization and the proteins bound to the RNA, either myosin when transporting along actin filaments or dynein or kinesin for microtubule-dependent localization (Figure 1.3; Wilhelm and Vale 1993; Jansen and Niessing 2012). These interactions allow for the active transport of multiple RNAs simultaneously. In some instances, several mRNAs are packaged in the same RNP transport granule, as in the case of *CaMKIIα*, *Neurogranin*, and *Arc* mRNAs transport to dendrites (Gao et al. 2008; Mitsumori et al. 2007). The composition and dynamics of cytoskeleton-associated RNPs influence the direction and speed of RNA transport. For example, in budding *Saccharomyces cerevisiae ASH1* mRNA is bound by the *trans*-acting factors She2p and She3p that facilitate myosin V binding and transport along actin filaments to the bud tip (Bookwalter et al. 2009). Similarly, during *Drosophila* oogenesis several mRNAs such as, *gurken*, *bicoid* and *oskar* have *cis*-acting elements that enable directed dynein or kinesin-dependent transport along microtubules to specific regions (MacDougall et al. 2003; Weil et al. 2006; Brendza et al. 2000; Palacios & St Johnston 2002). The exact mechanism is not always understood, since the RNP composition and complex dynamic interplay of different factors render quite complex the analysis of the molecular mechanisms underlying RNA localization.

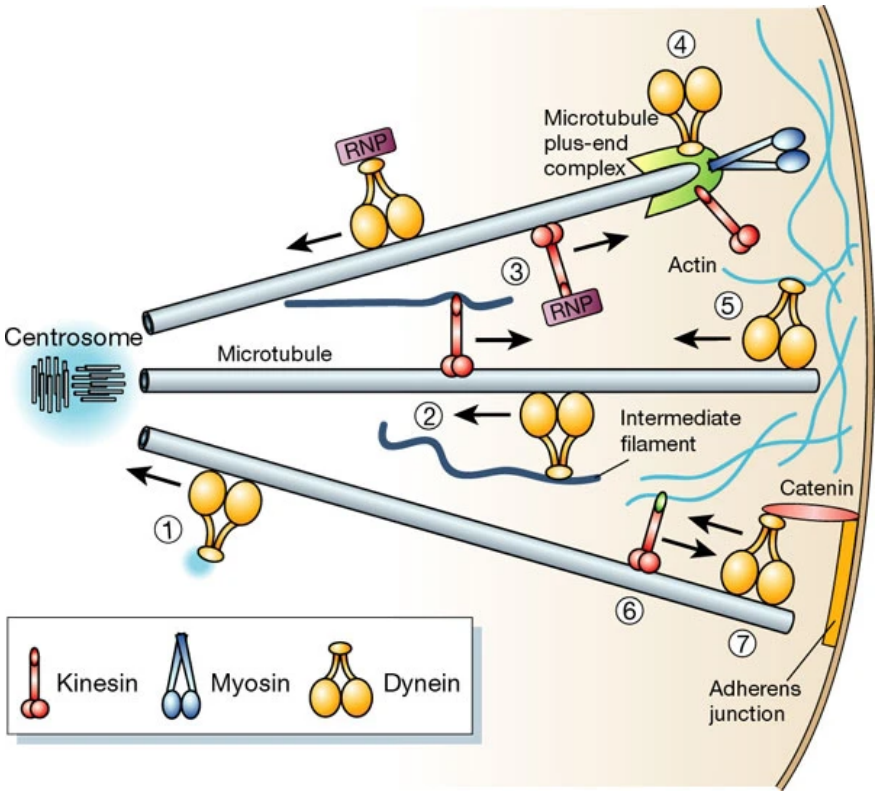

Figure 1.3 Schematic view of motor protein mediated transport along the cytoskeleton. Dynein transports cargo to the minus-end of microtubules, whereas kinesin transports cargo to the plus-end of microtubules. Myosin transports cargo along actin filaments (Image from Schliwa & Woehlke 2003)

1.3. *Drosophila* oogenesis for studying RNA localization

Thousands of RNAs within the *Drosophila* transcriptome have been shown to exhibit specific subcellular localization in early embryos (Lécuyer et al. 2007). Numerous RNAs that exhibit specific subcellular localization have been identified and studied in *Drosophila*. During *Drosophila* oogenesis, multiple mRNAs are transported to specific subcellular regions using different modes of RNA localization (Lasko 2012). This made the *Drosophila melanogaster* a popular model organism for investigating RNA localization.

The *Drosophila* ovary is composed of multiple ovarioles, each of which contains a chain of developing egg chambers. The process of oogenesis can be divided into 14 morphogenetically distinct stages which can be broadly classified into three phases: early, mid and late oogenesis (Figure 1.4). During early oogenesis, germ cells undergo mitotic divisions with incomplete cytokinesis to generate cysts of interconnected cells called cystocytes. One germ cell is specified to be the oocyte, while the others become nurse cells that support oocyte development (Bastock and St Johnston 2008). Meiosis is initiated shortly after the oocyte is specified, but stalls at prophase I and enters a transcriptionally silent state. The germline is surrounded by somatic follicle cells. Mid-oogenesis is characterized by the growth and differentiation of the oocyte, while late-phase oogenesis involves the formation of the eggshell and progression of meiosis (McLaughlin & Bratu 2015). After maturation, the oocyte travels from the ovary to the oviduct and then to the uterus, where it undergoes egg activation. Fertilization occurs while the egg is still in the uterus, after which it is released into the external environment, where embryogenesis takes place (Laver et al. 2015)

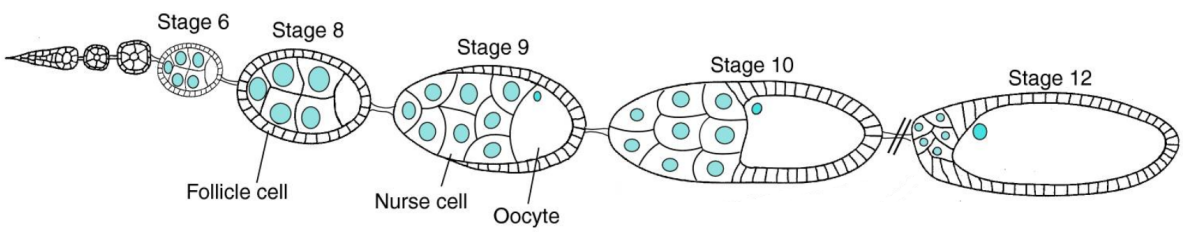

Figure 1.4 During *Drosophila* oogenesis egg chambers develop into mature eggs. The egg chambers consist of 16 germ cells comprised of 15 nurse cells and one posterior localized oocyte surrounded by follicle cells. The development is divided into 14 stages (Image from Becalska & Gavis 2009).

During the process of oocyte maturation, the oocyte is in a transcriptionally silent state, while the nurse cells are responsible for supplying nutrients and other essential molecules required for oocyte development. The nurse cells are connected to the oocyte through cytoplasmic bridges called ring canals, which allow the transfer of molecules between the two cells. One critical process during oogenesis is the specific localization of different maternal RNAs such as *nanos*, *bicoid*, and *oskar*, which specify the anterior-posterior axis of the future embryo. (Berleth et al. 1988; Wang & Lehmann 1991; Kim-Ha et al. 1991; Riechmann & Ephrussi 2001). Other genes, such as *gurken*, are involved in the formation of the dorsal-ventral axis (Figure 1.5; Neuman-Silberberg & Schüpbach 1993; González-Reyes et al. 1995). RNA localization is crucial for the proper polarization of the oocyte and is achieved through a variety of mechanisms, including active transport, anchoring and translational repression.

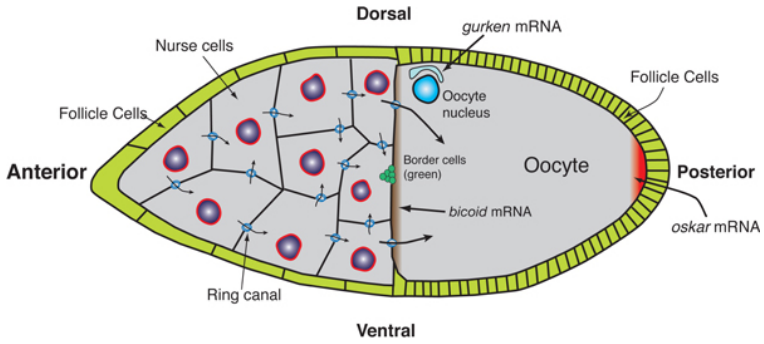

Figure 1.5 Localization of *oskar*, *gurken* and *bicoid* RNA in the *Drosophila* egg chamber during mid-oogenesis. *oskar* mRNA is transported to the posterior pole, *gurken* mRNA is localized at the dorsal-anterior corner and *bicoid* mRNA is translocated to the anterior side of the oocyte (Image from Gonsalvez & Long 2012).

RNA localization is required for establishment of the *Drosophila* body axes. Microtubules nucleate from the microtubule organization center at the oocyte posterior. The minus ends of the germline microtubules are organized in the oocyte and the plus ends extend through the ring canals into the nurse cells (Theurkauf et al. 1992; Grieder et al. 2000). During early oogenesis, *gurken* mRNA is transported by dynein from the nurse cells into the oocyte. The transport is achieved by binding two Egalitarian (Egl) proteins to *gurken* mRNA, which creates a complex that interacts with Bicaudal-D (BicD), leading to a conformational change of BicD. This change enables BicD to bind to dynactin and dynein, facilitating dynein motor-

dependent localization to the microtubule minus-end at the posterior pole (Bullock & Ish-Horowicz 2001; Navarro et al. 2004; Clark et al. 2007; Dienstbier et al. 2009; Goldman et al. 2019). The translation of *gurken* mRNA at the posterior pole initiates the specification of the anteroposterior axis and leads to the repolarization of the microtubules in the oocyte. The microtubules nucleate around the oocyte cortex with their plus-ends focused toward the posterior pole, leading to the migration of the oocyte nucleus and the transport of *gurken* mRNA towards the dorsoanterior corner of the oocyte (González-Reyes & Johnston 1998; Clark et al. 1997; Clark et al. 1994; MacDougall et al. 2003).

Similar to *gurken*, *bicoid* mRNA is transported during mid-oogenesis by a dynein machinery from the nurse cells into the oocyte. The *bicoid* mRNA is the localized anterior determinant (Berleth et al. 1988). During egg chamber stages 8 and 9 of oogenesis, *bicoid* mRNA is concentrated around the anterior-lateral margin of the oocyte, in a ring-like distribution. In stage 10 of oogenesis, the *bicoid* mRNA is relocalized from the anterior-lateral cortex and forms a disc shape at the center of the anterior cortex of the oocyte (St. Johnston et al. 1989; Trovisco et al. 2016). In addition, the posterior determinant Oskar is localized at the posterior pole of the oocyte. *oskar* mRNA is transported by dynein from the nurse cells into the oocyte and subsequently by kinesin to the posterior pole where it is translated and anchored (Ephrussi et al. 1991; Navarro et al. 2004; Brendza et al. 2000). Nanos, the abdominal determinant, localizes at the posterior pole through anchoring, which depends on Oskar (Wang and Lehmann 1991; Forrest and Gavis 2003).

1.4. *oskar* RNA localization during early oogenesis

The correct localization of *oskar* mRNA to the posterior pole of the oocyte is critical for the formation of germ plasm and embryonic patterning (Ephrussi et al. 1991). Similar to other cytoplasmic determinants, *oskar* mRNA is produced in the nurse cells of the *Drosophila* germline syncytium and must be transported to the oocyte. During early oogenesis, microtubules are nucleated in the oocyte and extend into the nurse cells. As a result, the dynein machinery is transporting *oskar* mRNA into the oocyte. The *oskar* 3'UTR alone can localize to the oocyte, which indicates that the coding region is not required for the dynein transport

(Hachet and Ephrussi 2004). Deletion analysis of the *oskar* 3'UTR revealed that a specific sequence named the oocyte entry signal (OES) which is a 67-nucleotide stem-loop forming *cis*-acting factor, is required for dynein-based transport of the RNA (Jambor et al. 2014). The OES is thought to recruit the Egl-BicD-dynein transport machinery (Figure 1.6; Bullock and Ish-Horowicz, 2001; Navarro et al., 2004; Dienstbier et al., 2009; McClintock et al., 2018). The stem structure, rather than the sequence, is essential for OES function, as demonstrated by mutational analysis. (Jambor et al. 2014). After *oskar* mRNA is transported into the oocyte by dynein, it localizes at the anterior of the oocyte during stages 7-8 (Ephrussi et al. 1991; Kim-Ha et al. 1991). During mid oogenesis, the polarity of the microtubule changes resulting in minus-ends enriched at the anterior and plus-ends at the posterior (Theurkauf et al. 1992; Grieder et al. 2000). Contrary to *bicoid* and *gurken* mRNA, which are transported by dynein to the anterior, *oskar* mRNA is transported by a kinesin-dependent mechanism to the posterior pole of the oocyte (Figure 1.7; Brendza et al. 2000).

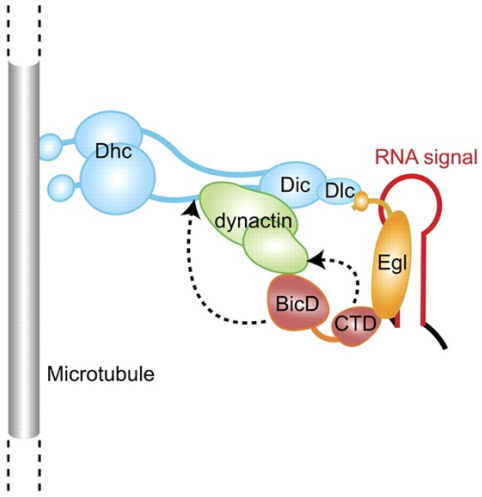

Figure 1.6 Simplistic dynein transport complex visualization. The double stranded RNA binding protein Egl binds the RNA and interacts with the dynein motor adapters BicD and dynactin. The Egl-BicD-dynein complex facilitates microtubule minus-end directed transport (Image from Dienstbier et al. 2009).

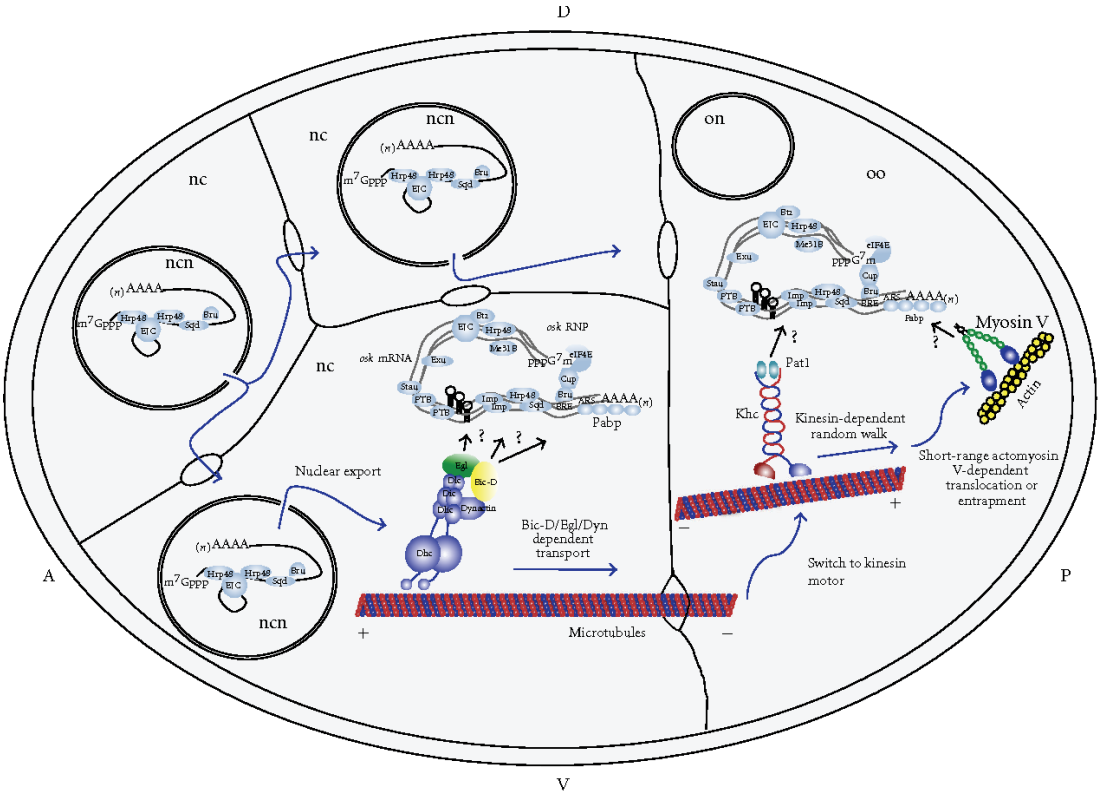

Figure 1.7 2-step transport mechanism of *oskar* mRNA from the nurse cells to the oocyte posterior pole. *oskar* mRNA is transported by a Egl-BicD-dynein transport mechanism from the nurse cells into the oocyte. In the oocyte, the dynein transport is inhibited and kinesin transport is activated. (Image from Vazquez-Pianzola and Suter 2012).

Interestingly kinesin is already recruited to the RNA in the nurse cells and is transported along with the *oskar* RNP complex by dynein into the oocyte (Gáspár et al. 2017). There are previous reports of RNA bound by opposing motor proteins in other cell types and organisms (Kanai et al. 2004; Messitt et al. 2008). A unique isoform of tropomyosin-1 (TM1-I/C) is able to inhibit kinesin activity and link it to the dynein-transported *oskar* RNP. This prevents any tug-of-war between the opposing motors and facilitates the co-transportation of inactive kinesin on *oskar* mRNA by dynein from nurse cells into the oocyte (Heber et al. 2022). During dynein transport of *oskar* mRNA, inactive kinesin remains bound and is only activated once the RNA is in the oocyte during mid-oogenesis (Sanghavi et al. 2013; Gaspar et al. 2021; Gaspar et al. 2023). In the oocyte, the double-stranded RNA binding protein Staufen binds the *oskar* 3'UTR and is required for the kinesin transport and translation of *oskar* mRNA. One of the Staufen Recognized Structures (SRS) in the *oskar* 3'UTR partially overlaps with the Egalitarian binding site termed the Transport and Anchoring Signal (TAS). Staufen prevents

dynein-mediated transport of *oskar* mRNA by displacing Egl from *oskar* RNPs after reaching the oocyte, which allows kinesin to transport the mRNA along with dynein to the posterior pole of the oocyte. (Gaspar et al, 2021). The details of how *oskar*-associated kinesin is activated in the oocyte are not yet fully understood.

Oskar protein is translated from stage 9 onwards when *oskar* mRNA is accumulating at the posterior pole. *oskar* mRNA and protein continue to accumulate at the posterior pole and remain localized until early embryogenesis (Markussen et al. 1995; Kim-Ha et al. 1995; Rongo et al. 1995). At the posterior pole, Oskar protein anchors *oskar* mRNA, which is especially important at stage 10B during ooplasmic streaming when the nurse cells dump their content into the oocyte (Rongo et al. 1995; Vanzo and Ephrussi 2002).

1.5. Exon junction complex in *oskar* mRNA transport

Over time, numerous factors involved in the kinesin-dependent transportation of *oskar* mRNA have been identified. Mutations in the proteins eIF4AIII, Mago nashi, Y14, and Barentsz were revealed to affected *oskar* mRNA localization, but only the second kinesin-based localization step was impacted, as *oskar* mRNA was still enriched in the oocyte but failed to accumulate at the posterior pole (Hachet and Ephrussi 2001; Mohr et al. 2001; Newmark & Boswell 1994; van Eeden 2001; Palacios et al. 2004). These proteins were shown to localize along with *oskar* mRNA to the posterior pole, which suggested that they are part of the *oskar* RNP complex. The four proteins were characterized as the core components of the exon junction complex (EJC). The EJC is a multiprotein complex that forms on mRNA concomitant with the splicing process in the nucleus (Le Hir et al. 2000A; Le Hir et al. 2000B; Kataoka et al. 2001; Shibuya et al. 2004; Palacios et al. 2004; Degot et al. 2004). The EJC acts as a binding platform for various proteins and thus modulates the composition of RBPs (Le Hir et al. 2001). The EJC is engaged in a number of post-transcriptional regulatory processes, such as RNA localization, translation and degradation (Tange et al. 2004).

The core components of the EJC in *Drosophila melanogaster* consist of the proteins eIF4AIII, Mago nashi, Y14, and Barentsz (Figure 1.8). During mRNA splicing the DEAD-

box RNA helicase eIF4AIII is first deposited on the RNA by the spliceosome component CWC22 (Barbosa et al. 2012). The RNA helicase eIF4AIII associates with the RNA 20 to 24 nucleotides upstream of the exon-exon junction site in a sequence-independent manner. However, the binding of eIF4AIII alone is not stable (Le Hir & Andersen 2008). The Y14-MAGO heterodimer bind to eIF4AIII, inhibiting its ATPase activity and stabilizing it on the RNA (Ballut et al. 2005; Andersen et al. 2006; Xiol et al. 2014). It is thought that a pre-EJC complex, consisting of eIF4AIII, Y14, and MAGO, is then exported to the cytoplasm, where Barentsz and other peripheral proteins form the mature EJC (Palacios et al. 2004; Gehring et al. 2009). The complete EJC is stably bound to the mRNA and is only removed later in the RNA life e.g. by ribosome-associated PYM during the first round of translation or RNA degradation (Bono et al. 2004; Gehring et al. 2009).

The core components of the EJC are conserved in *Drosophila* and their deposition was thought to occur at all exon-exon junctions, as suggested by *in vitro* studies. However, it has been shown that the EJC is not present at all exon-exon junctions, suggesting that some sort of stabilization or regulatory effect may be involved (Saulière et al. 2010; Saulière et al. 2012; Singh et al. 2012). The presence of secondary structures at EJC binding sites can disrupt EJC deposition and may even cause the EJC be deposited up- or downstream of its canonical binding site (Mishler et al. 2008). A transcriptome-wide analysis of EJC binding sites in *Drosophila melanogaster* revealed that the majority of EJCs are deposited on internal exons (Obrdlik et al. 2019). Several factors including long introns, strong splice sites and CG-rich hexamers influence EJC stability. Theoretical base-pairing capability analysis of regions adjacent to the EJC site suggests that the EJC is preferentially bound in close proximity to secondary structures (Obrdlik et al. 2019). However, it is not clear how these factors contribute to regulating or stabilizing EJC binding. Additionally, it has been shown that the EJC can bind to non-canonical sites on mRNA, independent of splicing (Saulière et al. 2012; Singh et al. 2012). Compared to mammalian cells, in *Drosophila melanogaster* the number of EJC at non-canonical sites is relatively low (Obrdlik et al. 2019).

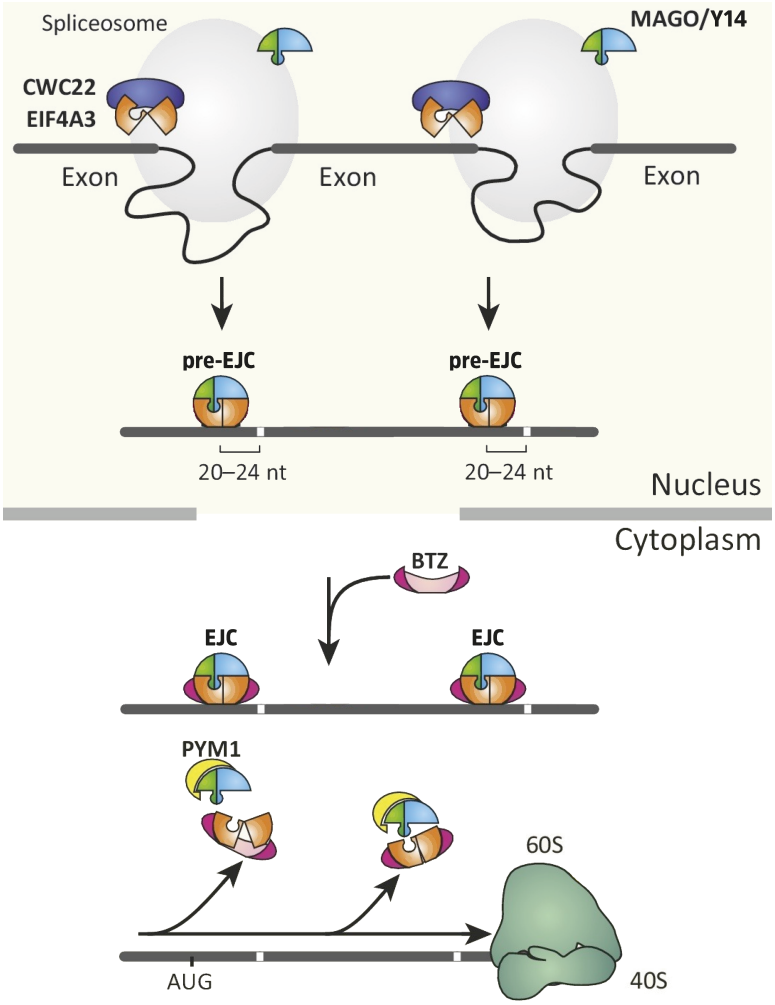

Figure 1.8 EJC formation and dissociation on a processed mRNA. eIF4A3 is recruited by CWC22 to the RNA splice site. It binds 20 to 24 nucleotides upstream of the exon-exon junction site. Mago and Y14 as a heterodimer associate with eIF4A3 to form the pre-EJC. Once exported to the cytoplasm Barentsz completes the core EJC. The EJC is removed from the RNA by PYM e.g. during RNA translation (adapted from Boehm & Gehring 2016).

1.6. Spliced *oskar* localization element

The finding that the EJC is required for *oskar* mRNA localization seemed contradictory to previous findings showing that the *oskar* 3'UTR is sufficient to localize the intronless, *E. coli lacZ* RNA to the oocyte posterior pole (Kim-Ha et al. 1993), and raised the question whether splicing is necessary for *oskar* mRNA localization. To test this, the ability of an *oskar* mRNA derived from transgenic *oskar* cDNA to localize was examined, in oocytes lacking endogenous *oskar* RNA expression (*oskar* "RNA-null" background). The cDNA derived

RNA failed to localize in absence of endogenous *oskar* mRNA, revealing that in previous studies the intronless *oskar* 3'UTR-containing transcript localized by "hitch-hiking" with endogenous *oskar*, and that splicing of *oskar* mRNA or at least *oskar* introns were required for proper *oskar* RNA localization (Hachet and Ephrussi 2004; Jambor et al. 2011). Further experiments showed that splicing of the second or third *oskar* intron is not required for localization, whereas splicing of the first intron is necessary. Replacing the first intron with the third intron, or even with an intron of an unrelated gene, did not affect *oskar* RNA localization, which demonstrated that the position of the first intron, rather than its sequence, is essential.

The requirement for the EJC and splicing at the position of the first intron suggested that EJC deposition at the first exon-exon junction site is crucial for the kinesin transport (Hachet and Ephrussi 2004). To confirm this, the *oskar* coding sequence was replaced by a *lacZ* coding sequence of similar length, which failed to localize to the posterior pole (Ghosh et al. 2012). This indicated that in addition to splicing at the first intron and EJC deposition, additional exonic sequences of *oskar* mRNA are necessary. A series of substitutions of the *oskar* coding region with *lacZ* sequence fragments demonstrated that the last 18 nucleotides of the first exon and the first 10 nucleotides of the second exon are required, while the remaining *oskar* coding sequence could be replaced by *lacZ* and still localize to the posterior pole (Ghosh et al. 2012). This bipartite 28-nucleotide sequence was named spliced *oskar* localization element (Figure 1.9; SOLE).

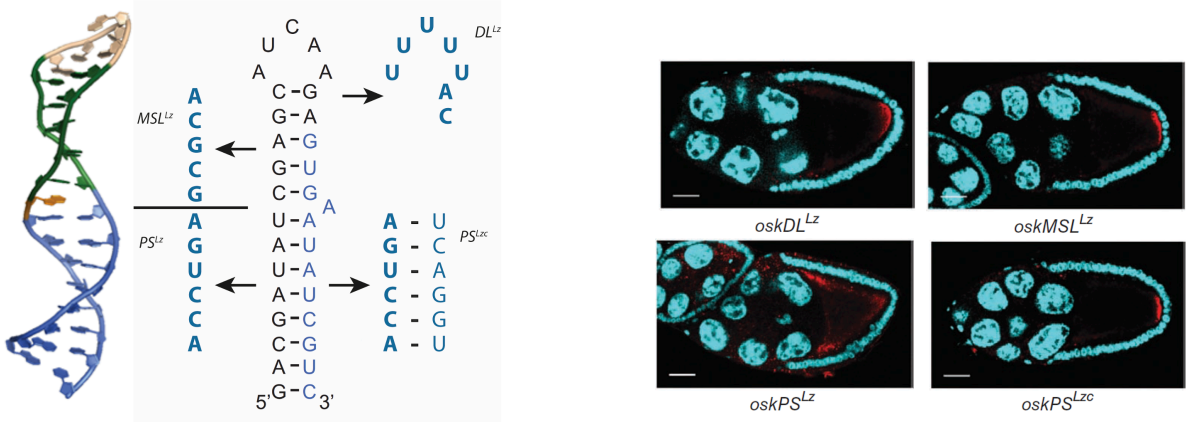

Figure 1.9 SOLE secondary structure is relevant for *oskar* mRNA localization. While mutations in the loop (DL^{Lz}) and terminal stem (MSL^{Lz}) do not affect RNA localization, mutations in the proximal stem (PS^{Lz}) disrupt RNA transport. RNA localization can be rescued upon restoring the SOLE secondary stem-loop structure (PS^{Lzc}) (adapted from Ghosh et al. 2012 and Simon et al. 2015).

The SOLE contains a small loop consisting of 5 nucleotides and an elongated stem, which is formed by non-Watson-Crick base pairing (Simon et al. 2015). Mutations in the SOLE sequence were used to analyze the contribution of SOLE in *oskar* kinesin transport. However, the loop (DL^{Lz}) or the terminal stem sequence (MSL^{Lz}) has no effect on *oskar* mRNA localization (Ghosh et al. 2012). The replacement in the terminal part of the stem still allowed the formation of a stem-loop structure due to non-Watson-Crick base pairing (Simon et al. 2015). On the other hand, mutating the proximal stem (PS^{Lz}) and significantly disrupting the stem-loop structure abolished *oskar* mRNA transport to the posterior pole. The localization was rescued by introducing a compensatory sequence (PS^{Lzc}) at the second exon that is complementary to the mutation in the proximal stem. These findings suggest that the stem-loop structure, rather than the sequence, is crucial for the kinesin-based transport of *oskar* mRNA. Ultimately, structural analysis by NMR confirmed that, at least *in vitro*, the SOLE forms a stem-loop structure (Figure 1.9).

In summary, splicing at the position of the first intron, EJC deposition, and the SOLE structure promote kinesin-based transport of *oskar* mRNA to the posterior pole (Figure 1.10). However, the exact role and molecular mechanism whereby the EJC and SOLE achieve this function remain to be understood.

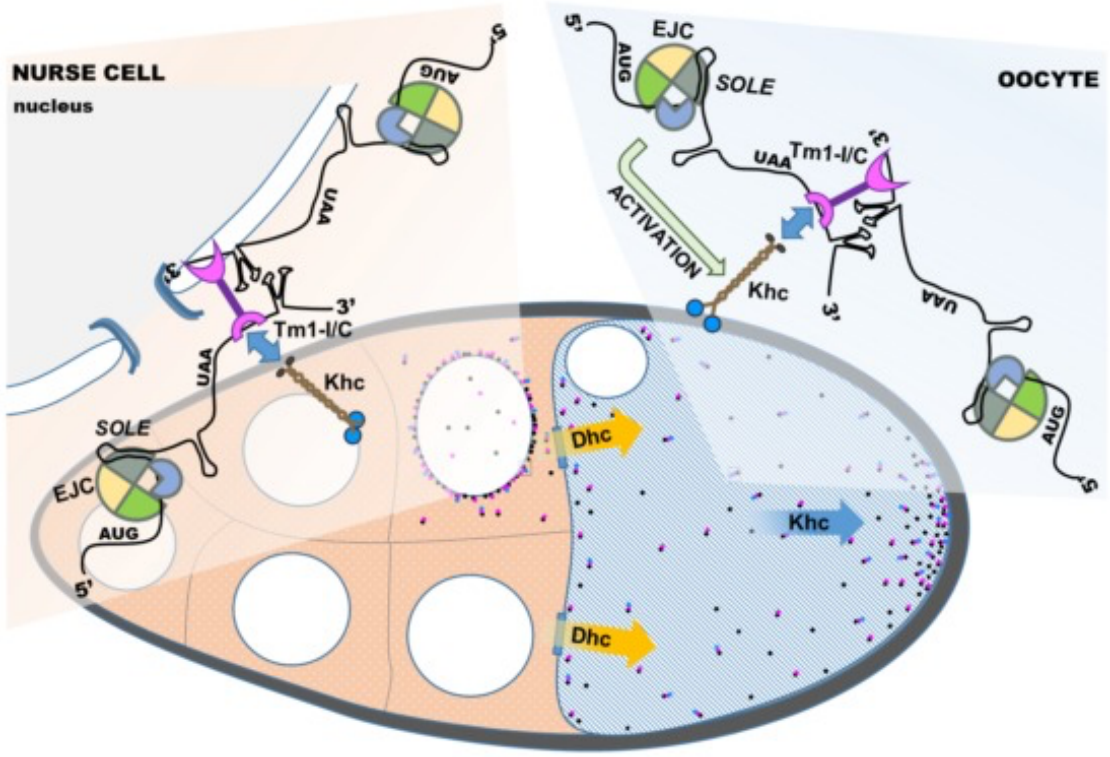

Figure 1.10 2-step transport mechanism of *oskar* mRNA. *oskar* mRNA is transported along kinesin and the motor adapter TM-1/C from the nurse cells into the oocyte by dynein mediated transport. In the oocyte the kinesin transport is activated and translocates *oskar* mRNA to the posterior pole (Image from Gáspár et al. 2017).

1.7. Identification of RNA binding proteins

RNA binding proteins (RBPs) determine the fates of RNA. RBPs are essential for various biological processes including the correct localization of maternal mRNAs during *Drosophila* egg development. *oskar* mRNA transport requires the EJC which is associated with several different peripheral proteins and is used as a platform for numerous proteins to interact with. The identification of RBPs and their target RNA can help understand specific biological or molecular processes. Therefore, several methods for RBP identification have been developed over the last decades. The following are some of the methods used for RBP identification.

To verify that a particular protein has the ability to bind RNA, immunoprecipitation (IP) can be utilized to isolate the protein-RNA complex and analyze the bound RNA. IP was one

of the first methods used to biochemically verify the RNA binding capacity of RBPs. Proteins that were expected to bind RNA, for example, due to their protein structure containing RNA binding domains (RBDs) or RNA recognition motifs (RRMs) were verified through IP (Dreyfuss et al. 1984; Adam et al. 1986; Blencowe et al. 1998; Mayeda et al. 1999; McGarvey et al. 2000; Le Hir et al. 2000; Kataoka et al. 2000). The bound RNA was analyzed by reverse transcription and PCR to confirm the RNA binding capacity of the specific protein. IP can be performed either in native conditions or after UV-crosslinking (CLIP), which stabilizes protein-RNA interactions (Ule et al. 2005). Advances in sequencing led to the development of high-throughput sequencing of RNA isolated by CLIP (HITS-CLIP; Licatalosi et al. 2008) and many more variations on the CLIP approach to identify transcriptome-wide RBP targets or RBP binding sites (Hafner et al. 2010; König et al. 2010; Huppertz et al. 2014; Zarnegar et al. 2016; Zhao et al. 2019; Strittmatter et al. 2020). This approach has been recently used and revealed that FMR1 shares the same binding site as Bruno and likely compete with each other for binding to the *oskar* 3'UTR (Vaishali 2022).

To determine the RBP composition, proteins or a complex that is predominantly binding to RNAs such as the EJC or the RNA-induced silencing complex (RISC) can be targeted by complex immunoprecipitation (Co-IP). Protein-RNA binding is stabilized by UV-crosslinking and protein-protein binding by formaldehyde treatment (Singh et al. 2013; Patton et al. 2020). RISC is a multiprotein complex that binds to small non-coding RNAs, such as microRNAs (miRNAs) and small interfering RNAs (siRNAs) and targets mRNA for degradation or translational repression (Iwakawa and Tomari 2022). By targeting one component of the RISC or the EJC by Co-IP the associated RNAs and bound proteins can be purified and identified by RNA sequencing and mass spectrometry, respectively. By targeting Ago2, the protein eIF1A was identified as part of the RISC (Yi et al. 2015). Similarly, this method was used to analyze the RNA-EJC interactome by targeting the EJC via immunoprecipitation (Wang et al. 2018; Mabin et al. 2018).

To determine the RNA binding proteome of all mRNAs, RNA affinity chromatography using oligo(dT) was performed in different cells and organisms and analyzed by mass spectrometry (Baltz et al. 2012; Castello et al. 2012; Beckmann et al. 2015; Wessels et al. 2016; Sysoev et al. 2016). Many proteins were identified to bind mRNA in which a large

proportion was shown to lack known RBDs (Castello et al. 2016). The drawback of this method is the focus on the poly(A) tail which can create a bias and underrepresents pre-mRNA and RNAs with a shortened poly(A) tail. The detected RBPs need to be validated by other methods to exclude false positives (Vaishali et al. 2021).

1.8. RNA-centric RBP analysis

Contrary to protein-centric methods, RNA-centric RBP analysis enables the identification of all RNA-specific RBPs. A transcript-specific RBP analysis is able to determine how sequence or structural changes in an RNA changes the RBP composition. An RNA-centric RBP analysis approach can be achieved by using biotinylated RNAs. Biotinylation throughout the RNA sequence or at the 3'end enables RNP pulldown by streptavidin beads. *In vitro* transcribing RNA in presence of a biotin-nucleotide such as biotin-UTP or biotin-CTP generates RNA with several biotin modifications (Panda et al. 2014; Panda et al. 2016). The biotinylated RNA is incubated in cell lyses to facilitate protein binding of potential RNA-specific binding proteins and their identification via western blot or mass spectrometry. A systematic analysis of RNA truncation can lead to the identification of specific *cis*-elements required for protein binding (Panda et al. 2014). However, biotinylation of the entire RNA sequence has the potential to cause adverse binding to streptavidin beads, obstruct RBP docking sites or have an unknown impact on the RNA secondary structure. One way to minimize the potential side effects is by using only 3'biotinylated RNA generated by attaching a 3′-biotinylated nucleotide to an *in vitro* transcribed RNA (Richardson and Gumport 1983; Cole et al. 2004). While the *in vitro* approach offers a simple method for identifying RBPs, it does not involve *in vivo* mechanisms that recruit proteins to the RNA, either directly or indirectly.

An *in vivo* RNA-centric method to purify RNP complexes is to attach an RNA aptamer tag to the RNA of interest, enabling the capture of the tagged RNA using a ligand with a high binding affinity. The aptamers form a stem-loop structure that specifically interacts with a ligand protein. The MS2 coat protein (MCP) from the bacteriophage MS2 binds with high affinity to MS2 aptamers and has been used to identify RNA-specific RBPs (Slobodin and

Gerst et al. 2010). Other aptamer-based systems have been developed that involve tagging the bacterial streptavidin-binding S1 aptamer or the smaller optimized S1m aptamer to RNA and using them with streptavidin beads to identify RNA-specific RBPs (Srisawat & Engelke 2001; Leppek & Stoecklin 2014; Dix et al. 2013). RNA aptamers provide a sturdy and adaptable approach for analyzing RNP complexes *in vitro* and *in vivo*. However, the presence of an aptamer tagged to an RNA or proteins bound to the aptamer may change the global structure of the RNA (Gerber 2021).

Proximity-dependent biotin labeling methods have recently emerged for proteomic profiling of organelles, subcellular domains, protein interactomes, but also proteins in close proximity to an RNA of interest (Hung et al. 2016; Benhalevy et al. 2018; Ramanathan et al. 2018; Fazal et al. 2019; Mukherjee et al. 2019; Zhang et al. 2020). Contrary to other approaches the advantage of a proximity labeling approach is the identification of local interactomes and transient interacting factors. Typically, proximity labeling is reliant on enzymes such as BirA ligase or engineered ascorbate peroxidase (APEX) that transforms a substrate into a reactive radical, which then covalently tags neighboring proteins with biotin (Roux et al. 2012; Rhee et al. 2013; Jan et al. 2014). For analyzing the RBP composition of a specific RNA, BirA or APEX can be used in conjunction with the MCP-MS2 system. BirA or APEX is conjugated to MCP which leads to the biotinylation of proteins in close proximity to the transgenic MS2-RNA and thus the identification of RNA-specific RBPs. Once proximity labeling is carried out in living cells, the cells are lysed and biotinylated proteins are collected via streptavidin beads and subsequently analyzed by mass spectrometry. This method was successfully used to analyze the interactome of the *β-actin* mRNA in mouse embryonic fibroblasts (Mukherjee et al. 2019).

To determine RBPs of a specific endogenously expressed RNA without introducing any transgene, antisense oligonucleotides can be utilized. Unlike protein-centric RBP capture, targeting specific RNAs with antisense oligonucleotides is fairly inefficient, which limits their sensitivity despite their potential. Therefore, this was originally developed to analyze protein complexes on highly expressed RNAs such as rRNA, telomerase RNAs and U4/U6 small nuclear RNPs (Yehle et al. 1987; Blencowe et al. 1989; Lingner and Cech 1996). For many years this method was neglected due to its inability to efficiently target less abundant RNAs,

including lncRNAs and specific mRNAs. The improved sensitivity of mass spectrometry analysis has brought the antisense oligo-based method back into focus.

1.9. Transcript-specific RBP capture of *oskar* RNP

The localized *oskar* mRNA is bound by different RBPs throughout its transport from the nurse cells to the oocyte posterior pole. Recently, a transcript-specific RBP capture approach using antisense oligonucleotides was developed to analyze the RBP composition of RNAs in the *Drosophila* ovary in a quantitative manner (Figure 1.11; Wippich and Ephrussi 2020). Due to the inefficiency of RNA-centric pulldown methods, a large amount of *oskar* mRNA from *Drosophila* ovaries was required. The mass isolation of *Drosophila* ovaries was achieved by mechanical grinding flies in grinding mills and size separation via sieving (Jambor et al. 2016). The dynamic RNA-protein interactions were stabilized by physical and chemical crosslinking via UV and formaldehyde, respectively (Urdaneta and Beckmann 2020). Both crosslinking methods were required to capture not only direct RBPs but also peripheral proteins of the *oskar* RNP composition. 3'-biotinylated DNA oligonucleotide probes antisense to numerous sites to the *oskar* mRNA were used to capture *oskar* RNPs by streptavidin beads (Wippich and Ephrussi 2020). This led to the identification of many known *oskar* RBPs and some previously unknown *oskar* RBPs such as FMR1. This method was shown to determine the RBPs of a specific RNA in *Drosophila* oocytes and could potentially be utilized to analyze posttranscriptional mechanisms during *Drosophila* egg development.

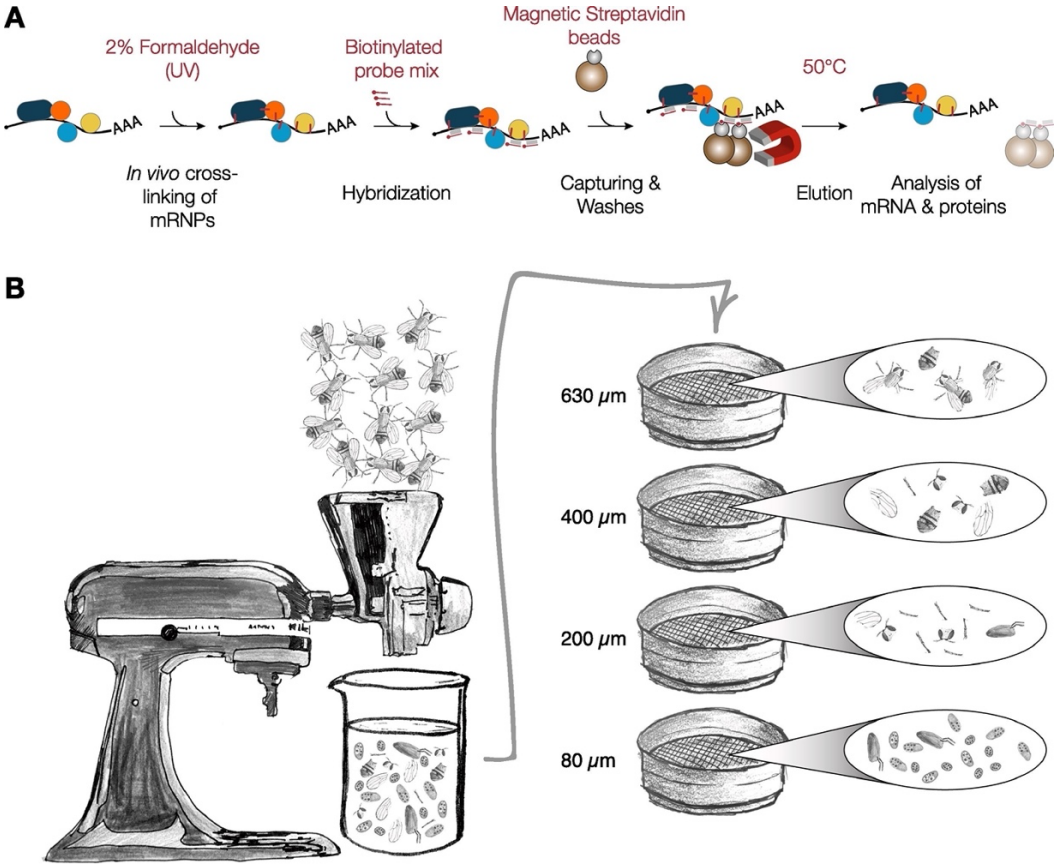

Figure 1.11 Transcript-specific RBP capture in *Drosophila* ovaries. (A) *Drosophila* ovaries are crosslinked and incubated with short 3'-biotinilated DNA probes that hybridize to the target RNA. Streptavidin beads are used to capture the RNA and its RBP composition. (B) Mass isolation of *Drosophila* ovaries is achieved through mechanical grinding by a kitchen grinder and subsequent sieving (Image from Wippich and Ephrussi 2020).

Aim

The aim of this study was to investigate the role of the EJC and SOLE in *oskar* mRNA transport. *oskar* mRNA is transcribed in the nurse cells and transported by dynein to the oocyte. During mid-oogenesis, kinesin transports *oskar* mRNA to the posterior pole of the oocyte. While the EJC and SOLE are not necessary for dynein transport, they are required for the kinesin transport of *oskar* mRNA. Since the EJC is binding in close proximity (2 to 6 nucleotides) to the SOLE, it was hypothesized that the EJC and SOLE interact with each other. Therefore, the study aimed to determine whether the SOLE stabilizes the EJC on the RNA by introducing mutations in the SOLE regions to disrupt the secondary stem-loop structure and observe its effect on EJC stability. I aimed to investigate whether the EJC and SOLE recruit

unknown proteins necessary for the kinesin transport of *oskar* mRNA. To address this, I generated transgenes with mutations in the SOLE and utilized transcript-specific RBP capture to compare the RBP composition of a transgenic RNA that localizes to the posterior pole versus an RNA that lacks the kinesin transport. Identifying the proteins involved in the kinesin-mediated transport of *oskar* mRNA to the posterior pole could elucidate the transport mechanism.

2. Materials and Methods

2.1. Materials

2.1.1. Enzymes

Name	Application	Source
DpnI	Cloning, generation of transgenic fly lines	Thermo
Proteinase K	RNA-protein footprinting, RNA analysis after transcript-specific RBP capture	Thermo
RNase A	RNA-protein footprinting, RNA analysis after transcript-specific RBP capture	Thermo
RNase T1	DNA probe binding analysis	Thermo
Terminal Deoxynucleotidyl Transferase (TdT)	smFISH probe labelling	Thermo

2.1.2. Antibodies

Name	Type, Application and Dilution	Source
Rabbit anti-Actin	Primary, Western Blot, 1:1500	Sigma
Rabbit anti-Bruno	Primary, Western Blot, 1:1000	Made in-house
Rabbit anti-eIF4A3	Primary, Western Blot, 1:3000	Isabel Palacios
Rabbit anti-GFP	Primary, Western Blot, 1:3000	Torres Pines Biolabs
Rabbit anti-Staufen	Primary, Western Blot, 1:1000	Made in-house
Rabbit ECL anti-Rabbit Ig, HRP linked whole antibody	Secondary, Western Blot, 1:10000	Sigma

2.1.3. Chemicals and reagents

Name	Source
10x Casein blocking buffer	Sigma
2,2'-thiodiethanol	Sigma
Acidic Phenol:Chloroform pH 4,5	Thermo
Amino-11-ddUTP	Lumiprobe
Atto-565	ATTO-TEC
Atto-633	ATTO-TEC
Calcium chloride ($CaCl_2$)	Merck
Carboxyl modified magnetic beads	New England Biolabs
Chloroform	Merck
ChromoTek Binding Control Magnetic Agarose Beads	ChromoTek
ChromoTek GFP-Trap® Agarose beads	ChromoTek
cOmplete® mini EDTA-free protease inhibitor	Roche
DAPI	Thermo
Dithiothreitol (DTT)	Biomol
Dynabeads™ MyOne™ Streptavidin C1	Invitrogen
EDTA	Merck
Ethanol	Merck
Ethylene carbonate	Sigma
Formaldehyde	Merck
Glycine	Merck
GlycoBlue™	Thermo
Heparin	Sigma
Immobilon western HRP Substrate	Merck
Isopropanol	Merck
Linear acrylamide	Thermo
Magnesium chloride ($MgCl_2$)	Sigma
NP-40	USB
NuPAGE™ LDS Sample Buffer (4X)	Thermo

Pierce® Avidin Agarose	Thermo
PMSF	Sigma
Potassium Chloride (KCl)	Merck
Potassium dihydrogen phosphate dihydrate (KH_2PO_4)	Merck
RiboLock RNase Inhibitor	Thermo
SDS	AppliChem
Sodium Acetate	Sigma
Sodium Hydrogencarbonate ($NaHCO_3$)	Sigma
Sodium Chloride (NaCl)	Merck
Sodium Citrate	Sigma
Sodium Deoxycholate	Sigma
Sodium Phosphate Dibasic (Na_2HPO_4)	Sigma
TdT Buffer	Thermo
Tris-Cl	Sigma
Tris-HCl	Sigma
Triton X-100	Sigma
Tween® 20	Sigma

2.1.4. Consumables

Name	**Source**
Container for plant tissue culture (fly bottle)	Greiner Bio-One
Cover slips	VWR
DNA LoBind® Tubes	Eppendorf
Drosophila vials	Dominique Dutscher
Immobilon-P nitrocellulose membrane	Merck
Novex™ TBE-Urea Gels, 6%, 10 well	Thermo
Nunc™ Square BioAssay Dish (tissue culture dish)	Thermo
NuPAGE™ 4 to 12%, Bis-Tris gel	Thermo
Protein LoBind® Tubes	Eppendorf

Saf-instant dry yeast	Lesaffre
Whatman paper	GE HealthCare

2.1.5. Buffers, solutions and mixes

Name	Composition	Application
Phosphate Buffered Saline (PBS)	137 mM NaCl, 2.7 mM KCl, 10 mM Na_2HPO_4, 1.8 mM KH_2PO_4 (pH 7.4)	Ovary dissection, Rapid Ovary Mass-Isolation, RIP, transcript-specific RBP capture, RNA-protein footprinting
Lysis Buffer	50 mM Tris-Cl (pH 7.5), 10 mM EDTA, 1% (v/v) SDS, 0.05% Sodium Deoxycholate, 1 mM PMSF, 1 tablet/25 ml cOmplete® mini EDTA-free protease inhibitor, 1:2000 RiboLock RNase Inhibitor	Transcript-specific RBP capture, RNA-protein footprinting
Hybridization Buffer	750 mM NaCl, 1% (v/v) SDS, 50 mM Tris-Cl (pH 7.5), 1 mM EDTA, 15% (v/v) Ethylene carbonate, 1 mM PMSF, 1 tablet/25 ml cOmplete® mini EDTA-free protease inhibitor, 1:2000 RiboLock RNase Inhibitor	Transcript-specific RBP capture, RNA-protein footprinting
Low Salt Wash Buffer	300 mM NaCl, 30 mM Sodium citrate (pH 7.0), 0.5% (v/v) SDS, 1 mM PMSF, 1 tablet/25 ml cOmplete® mini EDTA-free protease inhibitor, 1:2000 RiboLock RNase Inhibitor	Transcript-specific RBP capture, RNA-protein footprinting

High Salt Wash Buffer	750 mM NaCl, 30 mM Sodium citrate (pH 7.0), 0.5% (v/v) SDS, 1 mM PMSF, 1 tablet/25 ml cOmplete® mini EDTA-free protease inhibitor, 1:2000 RiboLock RNase Inhibitor	Transcript-specific RBP capture, RNA-protein footprinting
Elution Buffer	10 mM Tris-Cl (pH 7.5)	Transcript-specific RBP capture, RNA-protein footprinting
5x Proteinase K buffer	50 mM Tris-Cl (pH 7.5), 750 mM NaCl, 1% (v/v) SDS, 50 mM EDTA, 2.5 mM DTT, 25 mM $CaCl_2$	Transcript-specific RBP capture, RNA-protein footprinting
10x RNase A Buffer	100 mM Tris-Cl (pH 7.5), 1.5 M NaCl, 0.5% (v/v) NP-40, 5 mM DTT	Transcript-specific RBP capture, RNA footprinting
10x Cross-link blocking solution	1.25 M Glycine in PBS	Ovary harvest, Transcript-specific RBP capture, RNA-protein footprinting
Apple Agar-plate	11.5 l water, 5 l apple juice, 0.75 l beet syrup, 500 g agar, 0.42 l 10% Nipagin	Fly maintenance and harvest
Fly food	3 l water, 36 g agar, 54 g dry yeast, 30 g soya flour, 66 g beet sirup, 240 g malt extract, 240 g cornmeal, 18.6 ml propionic acid, 7.2 g Nipagin	Fly maintenance and harvest
PBT	137 mM NaCl, 2.7 mM KCl, 10 mM Na_2HPO_4, 1.8 mM KH_2PO_4 (pH 7.4), 0.1% Triton X100	Western blot, smFISH

HYBEC buffer	0.3 M sodium chloride, 0.03 M sodium citrate, 15% ethylene carbonate, 50 µg/mL heparin, 1% Triton X-100, 1 mM EDTA	smFISH
20x SSC buffer	3.0 M sodium chloride, 0.3 M sodium citrate (pH 7.0)	RIP, transcript-specific RBP capture, RNA-protein footprinting
10x RIPA buffer	20 mM Tris-Cl (pH 7.5), 150 mM KCl, 0.5 mM EDTA, 1% Triton X-100, 1% deoxycholate, 0.1% SDS, 5 mM DTT, 1 tablet/10 ml cOmplete® mini EDTA-free protease inhibitor, 1 mM PMSF, 1:2000 RiboLock RNase Inhibitor	RIP
RIP Low Salt Wash Buffer	20 mM Tris-Cl (pH 7.5), 150 mM KCl, 0.5 mM EDTA, 1 tablet/10 ml cOmplete® mini EDTA-free protease inhibitor, 1 mM PMSF, 1:2000 RiboLock RNase Inhibitor	RIP
RIP High Salt Wash Buffer	20 mM Tris-Cl (pH 7.5), 500 nM NaCl, 1 mM EDTA, 0.1% SDS, 0.5 mM DTT, 1 tablet/10 ml cOmplete® mini EDTA-free protease inhibitor, 1 mM PMSF, 1:2000 RiboLock RNase Inhibitor	RIP
Casein blocking buffer	50 µg/mL Heparin, 10x casein blocking buffer	RIP
Towbin buffer	25 mM tris base, 192 mM glycine, 20% (v/v) methanol	Western blot

2.1.6. Kits & Mixes

Name	Application	Source
MEGAscript™ T7 Transcription Kit	*In vitro* transcription	Thermo
NEXTFLEX® Small RNA-seq v3 Kit	RNA sequencing	Perkin Elmer
Phusion Flash High-Fidelity PCR Master Mix	Cloning, PCR	Thermo
SuperScript™ III First-Strand Synthesis SuperMix	cDNA synthesis for RT-qPCR	Thermo
SYBR™ Green PCR Master Mix	RT-qPCR	Thermo
TMT10plex™ Isobaric Label Reagents and Kits	Mass spectrometry analysis	Thermo
Zymo Quick-RNA Microprep Kit	RNA isolation	Zymo Research

2.1.7. Primers

Name	Sequence (5' to 3')	Purpose
MK1 F-pUASP *lacZ* SL	TTGACCTGACGAGCATCAAGAGTGAGTAGC	Generate *lacZ* SL, forward primer
MK2 R-pUASP *lacZ* SL	TCGTCAGGTCAAATTCAGACGGCAAACGAC	Generate *lacZ* SL, reverse primer
MK3 F-pUASP *lacZ* SL1c	GTGATCAGGTAAACCGCCTCGCGGTGATG	Generate *lacZ* SLc, forward primer
MK4 R-pUASP *lacZ* SL1c	GGTTTACCTGATCACCTGGCGAATTTCAATTATAGAATTATTTAAGG	Generate *lacZ* SLc, reverse primer
MK5 F-pUASP *lacZ* IL	CATCAAGAGTGAATATCGAAACCGCCTCGC	Generate *lacZ* IL, forward primer

MK6 R-pUASP *lacZ* IL	GATATTCACTCTTGATGCTCGATATCGCAAATTCAGAC	Generate *lacZ* IL, reverse primer
MK7 F-*lacZ* for sequencing	ATCCGACGGGTTGTTACTC	Sanger sequencing, forward primer
MK10 F-Gal4 for sequencing	CAAGGGTCGAGTCGATAG	Sanger sequencing, forward primer
MK21 F-cut tr. intron+coding	CAGAACTTTCCTCCAAGCGATGACCATG	Remove *P-element* intron and coding sequence, forward primer
MK22 R-cut tr. intron+coding	GAGGAAAGTTCTGTTTGTGTACTCCCACTG	Remove *P-element* intron and coding sequence, reverse primer
MK37 F-*lacZ* 2 sequencing	AAGCCGTTGCTGATTCGAGG	Sanger sequencing, forward primer
MK38 F-*lacZ* 3 sequencing	ACGATTTACGCTGATGGATCG	Sanger sequencing, forward primer
MK39 F-*lacZ* 4 sequencing	ATGTATTGATGGTGATCACG	Sanger sequencing, forward primer
MK40 F-mini *white* end start	ACTGCACTGGATATCATTG	Sanger sequencing, forward primer
MK41 F-mini *white* end mid 1	TTCGCAGAGCTGCATTAACCAGG	Sanger sequencing, forward primer
MK42 F-mini *white* end mid 2	ATCCTTCTGATGGCCGAGGG	Sanger sequencing, forward primer
MK43 F-mini *white* end mid 3	AGGCCCGAAGTCGACTTTATCG	Sanger sequencing, forward primer
MK44 F-mini *white* end	AGTTGCTCTTTCGCTGTCTCC	Sanger sequencing, forward primer
MK45 F-after insert 1	AGCAGATTACGCGCAGAAAAAAAGG	Sanger sequencing, forward primer
MK46 F-after insert 2	AGCAGATTACGCGCAGAAAAAAAGG	Sanger sequencing, forward primer

MK47 F-plasmid end	TCTGGGTGAGCAAAAACAGG	Sanger sequencing, forward primer
MK63 F-qPCR *lacZ* over intron	ACTCGGCGTTTCATCTGTGG	RT-qPCR, forward primer
MK64 R-qPCR *lacZ* over intron	ATCCTGATCTTCCAGATAACTGC	RT-qPCR, reverse primer
MK123 F-T7 promotor2	CATTAATGCAGGTTAACCTGG	*In vitro* transcription of *lacZ* WT, forward primer
MK124 R-S1m2	GGTGACACTATAGAACCAGATCT	*In vitro* transcription of *lacZ* WT, reverse primer

2.1.8. Plasmids

Name	**Source**
pUASp-*lacZ*(IIb,III)osk	Sanjay Ghosh
pUASp-*lacZ*-WT	This study
pUASp-attB-ΔK10	Ephrussi lab
pUASp-attB-ΔK10-*lacZ*-WT	This study
pUASp-attB-ΔK10-*lacZ*-SL	This study
pUASp-attB-ΔK10-*lacZ*-SLc	This study
pUASp-attB-ΔK10-*lacZ*-IL	This study
pSP73-4xS1m-*lacZ*-WT	This study

2.1.9. Fly stocks

Genotype	**Description**	**Source**
w^{1118}	wild-type	Ephrussi lab

oskGAL4/CyO; lacZ-WT, osk^{attP,3P3GFP}/TM3, P{w[+mC]=hs-hid}14, ry[*] Sb[1]	Stable stock for fly maintenance and amplification. Expression of *lacZ* WT under UAS promotor driven by oskGAL4.	This study
oskGAL4/CyO; lacZ-SL, osk^{attP,3P3GFP}/TM3, P{w[+mC]=hs-hid}14, ry[*] Sb[1]	Stable stock for fly maintenance and amplification. Expression of *lacZ* SL under UAS promotor driven by oskGAL4.	This study
oskGAL4/CyO; lacZ-SLc, osk^{attP,3P3GFP}/TM3, P{w[+mC]=hs-hid}14, ry[*] Sb[1]	Stable stock for fly maintenance and amplification. Expression of *lacZ* SLc under UAS promotor driven by oskGAL4.	This study
oskGAL4/CyO; lacZ-IL, osk^{attP,3P3GFP}/TM3, P{w[+mC]=hs-hid}14, ry[*] Sb[1]	Stable stock for fly maintenance and amplification. Expression of *lacZ* IL under UAS promotor driven by oskGAL4.	This study
oskGAL4/CyO; lacZ-WT, osk^{attP,3P3GFP}/lacZ-WT, osk^{attP,3P3GFP}	Expression of *lacZ* WT under UAS promotor driven by oskGAL4 in absence of endogenous *oskar* RNA. The insertion in the *oskar* allele abolishes *oskar* RNA expression.	This study
oskGAL4/CyO; lacZ-SL, osk^{attP,3P3GFP}/lacZ-SL, osk^{attP,3P3GFP}	Expression of *lacZ* SL under UAS promotor driven by oskGAL4 in absence of endogenous *oskar* RNA. The insertion in the *oskar* allele abolishes *oskar* RNA expression.	This study
oskGAL4/CyO; lacZ-SLc, osk^{attP,3P3GFP}/lacZ-SLc, osk^{attP,3P3GFP}	Expression of *lacZ* SLc under UAS promotor driven by oskGAL4 in absence of endogenous *oskar* RNA. The insertion in the *oskar* allele abolishes *oskar* RNA expression.	This study

oskGAL4/CyO; lacZ-IL, $osk^{attP,3P3GFP}$/lacZ-IL, $os^{kattP,3P3GFP}$	Expression of *lacZ* IL under UAS promotor driven by oskGAL4 in absence of endogenous *oskar* RNA. The insertion in the *oskar* allele abolishes *oskar* RNA expression.	This study
w-, oskGAL4/oskGAL4	Fly line for Gal4 expression under *oskar* promoter.	Ephrussi lab, Imre Gaspar
GFP-Mago; FLAG-HA-eIF4A3-1/FLAG-HA-eIF4A3-1; Myc-Y14, tubGAL4:VP16/TM3 Ser	Fly line with tagged EJC core components. Used to introduce GFP-Mago in other fly lines.	Ephrussi lab, Ales Obrdlik
GFP-Mago; oskGAL4/CyO; lacZ-WT, $osk^{attP,3P3GFP}$/TM3 Ser	Stable stock for fly maintenance and amplification. Expression of *lacZ* WT under UAS promotor driven by oskGAL4 and GFP-Mago expression.	This study
GFP-Mago; oskGAL4/CyO; lacZ-SL, $osk^{attP,3P3GFP}$/TM3 Ser	Stable stock for fly maintenance and amplification. Expression of *lacZ* SL under UAS promotor driven by oskGAL4 and GFP-Mago expression.	This study
GFP-Mago; oskGAL4/CyO; lacZ-WT, $osk^{attP,3P3GFP}$/lacZ-WT, $osk^{attP,3P3GFP}$	Expression of *lacZ* WT under UAS promotor driven by oskGAL4 in absence of endogenous *oskar* RNA. The insertion in the *oskar* allele abolishes *oskar* RNA expression. GFP-Mago expression.	This study
GFP-Mago; oskGAL4/CyO; lacZ-SL, $osk^{attP,3P3GFP}$/lacZ-SL, $os^{kattP,3P3GFP}$	Expression of *lacZ* SL under UAS promotor driven by oskGAL4 in absence of endogenous *oskar* RNA. The insertion in the *oskar* allele abolishes *oskar* RNA expression. GFP-Mago expression.	This study
if/CyO; sb/TM3 Ser	Double balancer fly line	Ephrussi lab

2.1.10. DNA probes

DNA probes used in smFISH and transcript-specific RBP capture are listed in Appendix.

2.1.11. Equipment

Name	Application	Source
ChemiDoc™ Touch Imaging System	Western blot, SDS-PAGE	Bio-Rad
Dounce tissue grinder (tight / B)	RNA-protein footprinting, Transcript-specific RBP capture	Kimble Chase
Illumina MiSeq	RNA sequencing	Illumina
Kitchen Aid® with grain mill	RNA-protein footprinting, Transcript-specific RBP capture	Kitchen Aid
Leica SP8 confocal microscope	Imaging	Leica
Magnetic Rack (15 ml)	RNA-protein footprinting, Transcript-specific RBP capture, RIP	NEB
Magnetic Rack (1.5 ml) (DynaMag®)	RNA-protein footprinting, Transcript-specific RBP capture, RIP	Thermo
Mechanical homogenizer	Ovary lysis, RIP	VWR
NanoDrop spectrophotometer	Measurement of RNA, DNA or protein amount	Thermo
Sieves (630, 400, 200, 125 & 80 µm)	Ovary harvest, RIP, RNA-protein footprinting, Transcript-specific RBP capture	ATECHNIK
StepOnePlus Real-Time PCR Systems	RT-qPCR	Thermo
UV Crosslinker	RIP, RNA-protein footprinting, Transcript-specific RBP capture	Stratalinker®

2.1.12. Software

Name	Source
Biorender	biorender.com
Fiji CortAnalysis plugin	Gaspar et al. 2014
Fiji software	Schindelin et al. 2012
Galaxy	Afgan et al. 2018
Integrative Genomics Viewer	Robinson et al. 2011
LAS X software	Leica
Prism 8.0	GraphPad
R studio	R Studio Team

2.2. Methods

2.2.1. Generation of transgenic *lacZ* fly lines

The primers MK21 and MK22 were used in a deletion-based restriction-free polymerase chain reaction (PCR) cloning approach to remove the 58 nucleotide *P-element* intron (Rørth 1998) and the 83 nucleotide *P-element* transposase coding region upstream of the *oskar* 5'UTR in the pUASp-*lacZ*(IIb,III)^{osk} plasmid (Ghosh et al. 2012). The primers MK21 and MK22 were designed to contain two parts, where one anneals to the *oskar* 5'UTR and the other upstream of the *P-element* intron. A two-step PCR with the Phusion Flash High-Fidelity PCR Master Mix was used to remove the site of interest and generate a plasmid in a restriction-free manner. The template plasmid was denatured by DpnI and the PCR mix containing the *lacZ* WT transgene in absence of the *P-element* intron (pUASp-*lacZ*-WT plasmid) was transformed into competent bacterial cells for plasmid amplification. Site-directed mutagenesis cloning was used to generate mutations in the SOLE site for *lacZ* SL and *lacZ* SLc transgenes, while restriction-free cloning was used to remove the first *oskar* intron for *lacZ* IL. Primers MK1 and MK2 were used for *lacZ* SL, primers MK3 and MK4 for *lacZ* SLc and primers MK5 and MK6 for *lacZ* IL generation. The primers MK1 to MK4 had a pentanucleotide mutation to generate the mutations in the SOLE sequence for *lacZ* SL and *lacZ* SLc. The primer MK5 and MK6 were designed for a two-step PCR so that one part of the primer would anneal to the first exon and another part to the second exon. Cloning success was verified by Sanger sequencing. The *lacZ* sequences were cloned into a pUASp-attB-ΔK10 plasmid and microinjected in *Drosophila* embryos for site-specific integration by the PhiC31 integrase to facilitate recombination using the attB and attP sites (Bateman et al. 2006). The injection service was carried out by Alessandra Reversi (EMBL). The transgenes *lacZ* WT, *lacZ* SL, *lacZ* SLc and *lacZ* IL were integrated at the same position in the *Drosophila* chromosome (III) to ensure similar expression levels.

To create flies that do not express any endogenous *oskar* RNA, the fly line *osk*^{attP,3P3GFP} was utilized. This fly line harbors a loxP cassette with a 3xP3-EGFP marker inserted after the first 26 nucleotides of the *oskar* transcription start site to prevent endogenous *oskar* RNA expression (Gáspár et al., 2017). Recombination was utilized to produce flies that have the

lacZ transgene and the *osk*^{attP,3P3GFP} insertion on the same chromosome to enable *lacZ* RNA expression in the absence of endogenous *oskar* RNA.

2.2.2. Fly maintenance and ovary harvest

All *Drosophila* fly lines were maintained either at 18°C, room temperature (RT) or at 25°C and kept in vials or bottles containing basic food made of soy flour, cornmeal and malt extract. One day prior to experiments, the food was supplemented with yeast and fly lines were kept at 25°C overnight. Large amounts of wild-type flies (w^{1118}) were kept in cages and constantly fed by apple agar plates supplemented with yeast paste, which was kindly managed by the EMBL fly kitchen ladies, Anna Cyrklaff and Matthew Benton. However, the transgenic fly lines *lacZ* WT, *lacZ* SL, *lacZ* SLc and *lacZ* IL were kept in bottles even after amplification because they were not able to be maintained in cages due to their poor health and low fecundity. Each transgenic fly line was amplified to at least 500 bottles, which served as a base to generate more flies for ovary harvesting.

For almost all experiments, the *lacZ* flies were required to be in an *oskar* RNA-null (*osk*^{null}) background, which could be achieved by having *osk*^{attP,3P3GFP} in homozygous form. Since *osk*^{null} results in sterile flies, the stable *lacZ* fly stocks were kept in heterozygous form. One chromosome (III) of the *lacZ* fly lines contained the *lacZ* transgene and *osk*^{attP,3P3GFP} insertion, while the other chromosome (III) had an apoptosis-inducing *hid* gene under the heat-shock (hs) promoter. To achieve homozygous form, the flies in stock (500 bottles) were transferred to new bottles with yeast-supplemented food and kept at 25°C. After three to five days, bottles with pupa and larva were heat-shocked in a water bath at 37°C for 90 mins on two consecutive days to induce apoptosis by the *hs-hid* gene in all heterozygous (III chromosome) larvae and pupae.

After the heat-shock treatment, pupae were given 2 days at 25°C or 3 days at room temperature for hatching and then collected and transferred to a cage containing apple agar plates supplemented with yeast paste at 25°C. On the following day, *Drosophila* ovaries were harvested by utilizing the kitchen grinder and sieves of different sizes (Jambor et al., 2016).

Sieves of sizes 630 µm, 400 µm, 200 µm, and 125 µm were used to filter-out body parts and eggs. Egg chambers were collected using an 80 µm sieve, transferred into a tube, washed twice with phosphate buffered saline (PBS) and fixed with 2% formaldehyde for 10 min at RT while moderate shaking. Crosslinking was stopped by adding 10x Cross-link blocking solution (1:10) and incubating for 5 min. Egg chambers were washed two more times with PBS and centrifugation at 600 rpm for 10 s at RT and snap-frozen in liquid nitrogen. The required amount of starting material (egg chambers) varied depending on the experiment, thus the quantity of flies processed using the kitchen grinder was adapted accordingly. Typically, grinding 100 ml of flies produced 1 to 1.5 ml of ovaries.

2.2.3. Probe design and labeling for smFISH

The antisense oligonucleotides DNA probes used for smFISH and transcript-specific RBP capture targeting the *lacZ* sequence were designed using the R script smFISHprobe_finder.R, created by Imre Gaspar (Gaspar et al. 2018). These probes were carefully designed to meet specific criteria, including a melting temperature of approximately 70°C, a GC content ranging between 45-60%, a gap between each probe binding site of at least 10 nucleotides and a primer length between 18 to 21 nucleotides. In total, a set of 45 probes were designed against the *lacZ* coding region for smFISH and transcript-specific RBP capture.

The DNA probes used for smFISH were labeled with ddUTP-Atto565 and ddUTP-Atto633 according to the protocol described in Gaspar et al. 2017. A reaction mix containing Amino-11-ddUTP, Atto-NHS ester (1:2 ratio) and 0.1 M $NaHCO_3$ was incubated at RT in the dark for 2-3 hours for labeling. The reaction was quenched by adding 1 M Tris-HCl. The Atto565 or Atto633 conjugated ddUTPs were subsequently incubated in a 15 µl reaction volume with 1000 pmol DNA probe mix, 24 units of terminal deoxynucleotidyl transferase (TdT) and TdT buffer at 37°C in the dark for 16 to 18 hours. After increasing the reaction volume to 200 µl containing 0.3 M sodium acetate and 0.5 µl of 5gm/ml linear acrylamide, the mixture was subjected to incubation with 800 µl of ice-cold ethanol for 20-60 min at -20°C. The precipitated DNA probes were pelleted by centrifuging for 30 min at 16,000 g

at 4°C, washed twice with ethanol, air dried and resuspended in 25 µl of DNase and RNase-free water. The concentration and labeling efficiency of the DNA probes were calculated by measuring the UV absorbance at 260 and 634 nm with a NanoDrop spectrophotometer, as described in Gaspar et al. 2017.

2.2.4. Single molecule fluorescence *in situ* hybridization (smFISH)

To prepare the female flies for imaging, they were first anesthetized with CO_2 and their ovaries were carefully hand-dissected in PBS at RT using tweezers. The ovaries were fixed in 2% formaldehyde in PBT (0.1% Triton X100 in PBS) for 20 minutes at RT while shaking. After fixation, the ovaries were washed twice with 1 ml PBT for 10 minutes each, followed by incubation in 200 µl of HYBEC buffer at 42°C for 15 minutes for pre-hybridization. For hybridization, 50 µl of HYBEC buffer supplemented with 2 nM per probe concentration was added to the ovaries and incubated for 2 hours at 42°C (for 45 DNA probes the final concentration of probes in 250 µl HYBEC buffer was 90 nM). After hybridization, the ovaries were washed twice with HYBEC buffer, to remove unbound probes, for 15 min each at 42°C on a nutator. Ovaries were then incubated in PBT supplemented with DAPI (final concentration of 2 µg/ml) for 5 minutes at RT followed by a PBT wash for 10 min on a nutator. The ovaries were subsequently transferred to mounting media consisting of 80% 2,2'-thiodiethanol (TDE) in PBS and incubated overnight at 4°C. Individual ovarioles were pipetted on a glass slide and carefully separated using tungsten needles. A cover glass was placed over the sample and sealed for imaging. The samples were imaged using a Leica SP8 confocal microscope.

2.2.5. RNA distribution analysis from smFISH

The Leica SP8 laser scanning confocal microscope with LAS X software and a 63x (1.3 NA) glycerol immersion objective were used for imaging *Drosophila* egg chambers. Samples were imaged at 565 nm and 633 nm depending on whether the smFISH probes were labeled with Atto-565 (*oskar* 3'UTR) or Atto-633 (*lacZ* coding region), respectively. To

measure the mean fluorescence values of RNA labeled with smFISH in the oocytes, Fiji software and the CortAnalysis plugin were used, according to the protocol outlined in Gaspar et al. 2014. The oocyte cell boundaries and the anteroposterior (AP) axis were manually defined and the CortAnalysis plugin was used to create profiles of signal distribution for labeled RNAs with smFISH. A 100x100 matrix was generated for each imaged oocyte, in which the fluorescence intensities were rearranged while maintaining both the signal intensity and the positional data. The matrices from all imaged oocytes were combined, and the average center of mass of the signal intensity distribution was determined relative to the geometric center of the oocyte. The center of mass was compared using Mann-Whitney U test in a pairwise manner.

2.2.6. RNA-protein footprinting

Drosophila ovaries were harvested using a kitchen grinder and washed twice with PBS and centrifugation at 600 rpm for 10 s before crosslinking in 2% formaldehyde in PBS. *lacZ* RNAs were enriched using a biotinylated antisense oligonucleotide (ASO) based RBP capture method. The RNA-protein footprinting analysis involved resuspending 1 ml of *Drosophila* ovaries in 3 ml of lysis buffer, crosslinking with 0.8 J/cm^2 at 254 nm UV-light using a stratalinker and homogenizing using a dounce tissue grinder with pestle "tight" or "B". The lysate was cleared by centrifuging twice at 140 g for 5 min and added to 2 volumes of hybridization buffer. The sample was precleared using 1:50 (v/v) avidin-agarose beads for 30 min at RT while rotating on a nutator followed by removal of beads by centrifugation at 140 g for 5 min. This step was repeated to ensure that all avidin-agarose beads were removed, which could potentially bind to biotinylated ASOs and reduce RNA capture. The sample was then incubated with 0.25 µg biotinylated DNA probes per 1 ml lysate, which corresponds to 0.2 µl from a 200 µM stock, for 1 hour at 37°C. An optimized DNA probe set (21 DNA probes) was utilized that was evaluated for its efficiency against *lacZ* RNA (Figure 2.10). Magnetic streptavidin beads were added and incubated for 2 hours at 37°C. After incubation, the supernatant was removed and the beads were retained using a magnetic rack.

To perform protein-protein complex protected RNA-protein footprinting, a mixture of 200 µl of lysis buffer and hybridization buffer (1:2) along with 20 µl of 10x RNase A buffer and 20 µg of RNase A were added to the streptavidin beads and incubated for 1 hour at 37°C. The supernatant was then transferred to a new tube while the remaining beads were heated in 100 µl of 10 mM Tris-Cl buffer at 90°C for 10 minutes and then transferred to the supernatant.

For the direct "zero distance" RNA-protein footprinting, the beads were washed twice in 1 ml of low salt wash buffer for 5 min at 37°C, then incubated in 100 µl elution buffer for 10 minutes at 90°C. The supernatant was transferred using a magnetic rack and supplemented with 10x RNase A buffer (1:10) and 20 µg of RNase A, followed by incubation for 1 hour at 37°C.

In both cases, the protein-protein complex and “zero-distance” RNA-protein footprinting, RNase A was inactivated with 400 U RiboLock RNase inhibitor and proteins were degraded by adding 25 µg of Proteinase K (in Proteinase K buffer) and incubating for 1 hour at 50°C. The RNA was then purified by phenol-chloroform extraction using 600 µL of Phenol:chloroform:isoamyl alcohol (pH 4.8; 125:24:1), rigorous vortexing and centrifuging at 16,000 rpm for 10 minutes at RT. The aqueous phase was transferred to a new tube and mixed with an equal volume of chloroform. After rigorous vortexing, the mixture was centrifuged at 16,000 rpm for 10 minutes at RT. The resulting aqueous phase was then pipetted to a new tube and supplemented with 0.1 volume of 3 M sodium acetate, 1 µl of GlycoBlue, 0.01 M $MgCl_2$ and 3 volumes of ethanol for RNA precipitation. The mixture was incubated at -20°C overnight, allowing the RNA to precipitate. The RNA was pelleted by centrifugation at 16,000 rpm for at least 60 minutes at 4°C. The RNA pellet was washed twice with ice-cold ethanol, air-dried and resuspended in 10 µl of RNase and DNase-free water. The RNA was sequenced to generate a *lacZ* gene coverage profile.

2.2.7. RNA immunoprecipitation (RIP)

Flies, in which GFP-Mago and either *lacZ* WT or *lacZ* SL were expressed (without any endogenous *oskar* RNA expression) were used to perform the RIP analysis. Approximately

100 ovary pairs from each fly line were hand-dissected in PBS, followed by crosslinking in 2% formaldehyde for 10 minutes at RT. After washing twice with PBS, the ovaries were snap-frozen in liquid nitrogen. The ovary pellets were resuspended in 100 µl of 10x RIPA buffer and pipetted onto a dish and crosslinked using 1 J/cm^2 at 254 nm UV light. The ovaries were collected in a tube and the volume of the sample was adjusted to 200 µl with 10x RIPA buffer. The sample was ground with a mechanical homogenizer and centrifuged for 1 minute at 500 g at 4°C to clear cell debris. Protein concentration was measured by a NanoDrop spectrophotometer, adjusted to 10 mg/ml, and 9 volume units of RIP low salt wash buffer was added to the supernatant. 15 µl of GFP-Trap beads and agarose beads (negative control) were blocked with Casein blocking buffer for 60 minutes at RT. Beads were washed once in RIP low salt wash buffer and then incubated in sample lysate for 90 minutes at 4°C on a rotating nutator. The beads were washed six times with RIP high salt wash buffer for 5 minutes at moderate shaking (900 rpm). After washing, 5% of the beads were analyzed by western blot to confirm the successful immunoprecipitation of GFP-Mago and quantify the amount of captured protein. The remaining beads were used to isolate RNA bound to GFP-Mago using the Zymo QuickRNA microprep kit, following the manufacturer's instructions. The eluted RNA was reverse transcribed with the SuperScript™ III First-Strand Synthesis SuperMix kit according to the manufacturer's instructions. The resulting cDNA was used for RT-qPCR analysis to quantify the RNA of interest captured by GFP-Mago. The cDNA was prepared in a reaction mix with SYBR green qPCR mix and primers specific to the *lacZ* coding region and analyzed by the StepOnePlus Real-Time PCR Systems from Applied Biosystems. The primers MK63 and MK64 (10 µM each) were used to quantify *lacZ* RNA.

2.2.8. Total RNA isolation and analysis

Total RNA was extracted from *Drosophila* ovaries immediately after dissection in PBS. 3 to 5 ovary pairs were lysed by a mechanical homogenizer in RNA lysis buffer from the Zymo QuickRNA microprep kit. After centrifugation at 13,000 rpm for 1 min, the supernatant was used for RNA purification using the Zymo QuickRNA microprep kit as per the manufacturer's instructions. First-strand cDNA was synthesized from the purified RNA with the SuperScript™ III First-Strand Synthesis SuperMix kit according to the manufacturer's

instructions with the slight alteration that all reaction components were used in half. One-tenth of the resulting cDNA, 10 µM MK63 and MK64 primers and the SYBR green qPCR mix were used for RT-qPCR analysis.

2.2.9. Western blotting

Western blot analysis was performed to analyze specific protein levels after RIP and transcript-specific RBP capture. Protein samples were first denatured in NuPAGE™ LDS Sample buffer and 10 mM DTT for 10 min at 95°C and then separated by SDS-PAGE using a pre-cast NuPAGE™ 4-12% Bis-Tris gel according to the manufacturer's instructions. The proteins were transferred onto a nitrocellulose membrane using a tank blotting system, sandwiching the membrane and SDS-gel between Whatman paper. The transfer occurred in towbin buffer for 2.5 hours at constant 0.25 A at 4°C. The membrane was then blocked in 5% milk in TBST with moderate shaking for 60 min at room temperature. The primary antibody was added and the blot was incubated overnight at 4°C, washed three times in PBT buffer for 10 min each at RT and then incubated for 2 hours with the secondary antibody conjugated with horseradish peroxidase at RT. All antibodies were diluted in 5% milk in PBT. Finally, the membrane was washed three times with PBT for 10 min each and developed using Immobilon western HRP Substrate (ECL) and the ChemiDoc™ Touch Imaging System.

2.2.10. Transcript-specific RBP capture

The transcript-specific RBP capture required a significant amount of input material. To obtain 5 to 10 ml of ovaries, sequential harvests were performed from homozygous *lacZ-osk*^{null} flies. Approximately 50 to 100 ml of flies were ground in each round, yielding between 0.5 and 1.5 ml of *Drosophila* ovaries. The ovaries were washed twice with PBS, centrifuged at 600 rpm for 10 s and crosslinked with 2% formaldehyde for 10 minutes at RT. After two PBS washes, the samples were snap-frozen in liquid nitrogen. Harvesting was repeated until enough ovaries were obtained from a localizing (*lacZ* WT or *lacZ* SLc) and non-localizing fly line (*lacZ* SL or *lacZ* IL). Transcript-specific RBP capture was then performed on two fly

lines, one localizing and one non-localizing, simultaneously. The ovaries were thawed in 2-3 volumes of lysis buffer, transferred to a plastic tissue culture dish and UV-crosslinked using 0.8 Joule/cm^2 at 254 nm wavelength. The sample was then homogenized using dounce tissue grinder (pestle "tight" or "B") and cleared three times by centrifugation at 140 g for 5 minutes. 2 volumes of hybridization buffer were added to the supernatant and precleared by prewashed avidin agarose beads (1:50) for 30 minutes at RT while rotating on a nutator. The avidin agarose beads were removed from the sample by centrifuging twice at 140 g for 5 minutes. 0.1% lysate was reserved for analyzing the RNA and protein levels in the input sample using RT-qPCR and western blot, respectively. Biotinylated *lacZ*-DNA probes or scrambled probes (negative control) at a concentration of 0.25 µg per 1 ml lysate were added to the sample, which corresponded to 0.2 µl of a 200 µM stock. The sample was incubated for 2 hours at 37°C while rotating to facilitate hybridization. Prewashed streptavidin beads (3.75 µl per 1 ml lysate) were added to the sample and the mixture was incubated for an additional hour at 37°C. The beads were then collected utilizing a magnetic rack and the supernatant was removed. The beads were resuspended in 1 ml low salt wash buffer, transferred to a 1.5 ml low binding tube and washed for 5 minutes at 37°C with 900 rpm on a thermomixer. This wash step was repeated twice, after which the beads were transferred to a new tube and washed twice with 1 ml high salt wash buffer under the same conditions as before. The beads were transferred to a new tube and subjected to three additional washes with low salt wash buffer. Particular attention was paid to removing as much supernatant as possible during the final wash step. The remaining beads were resuspended in 0.75 µl elution buffer per 1 µl beads, transferred to a new tube and incubated at 70°C for 10 minutes on a thermomixer while shaking at 900 rpm. The supernatant was collected and snap-frozen in liquid nitrogen for mass spectrometry analysis. For RNA analysis, 5% of the elution was separated and supplemented with 5x Proteinase K buffer and 10 µg Proteinase K. The sample was incubated at 50°C for 45 minutes and stored in the RNA lysis buffer from the Zymo QuickRNA microprep kit at -20°C. RNA purification utilizing the Zymo QuickRNA microprep kit, cDNA synthesis by the SuperScript™ III First-Strand Synthesis SuperMix kit and RT-qPCR analysis with the SYBR green qPCR mix were done as advised by each manufacturer. For protein analysis, 5% of the elution was separated and incubated in RNase buffer and 10 µg RNase A. The sample was incubated at 37°C for 45 minutes and stored at -20°C for western blot analysis.

2.2.11. *In vitro* RNA transcription

DNA templates for *in vitro* transcription were generated by PCR. The primers MK123 and MK124 were used for amplifying the *lacZ* WT sequence with the T7 promotor from the pSP73-4xS1m-*lacZ*-WT plasmid. The MEGAscript™ T7 Transcription Kit was used to transcribe the *lacZ* WT RNA *in vitro* following the manufacturer's instructions. The *in vitro* transcribed RNA was purified using phenol:chloroform extraction. This involved mixing the sample with phenol:chloroform and chloroform, separating both times the aqueous phase and then precipitating the RNA with isopropanol overnight at -20°C. The RNA was pelleted by centrifuging at 16,000 rpm for 60 min at 4°C and resuspended in RNase and DNase-free water. The quality of the RNA was assessed by RNA PAGE using 6% Urea-TBE gels as described by the manufacturer. The RNA was quantified using a NanoDrop spectrophotometer.

2.2.12. DNA probe binding analysis

Fixed *lacZ* ovaries (using formaldehyde) were resuspended in 3 volume lysis buffer and UV-crosslinked with 0.8 J/cm^2 at 254-nm. The chemically and physically crosslinked ovaries were then homogenized with a dounce tissue grinder. The resulting lysate was cleared by centrifuging two times at 140 g for 5 min and supplemented with 2 volumes of hybridization buffer and precleared with avidin agarose beads (1:50 v/v) for 30 min at RT while rotating. The avidin agarose beads were removed by centrifugung twice at 140 g for 5 min. The supernatant was transferred to a new 15 ml falcon tube and supplemented with 0.25 µg biotinylated *lacZ* DNA probes and incubated for 2 hours at 37°C while shaking. After incubation, 1000 U RNase T1 was added to the sample and incubated for 1 hour, followed by adding 3.75 µl streptavidin beads per 1 ml lysate and incubating for an additional 1 hour at 37°C. The beads were collected utilizing a magnetic rack and the supernatant was removed. The beads were washed three times with low salt wash buffer while shaking for 10 min at 900 rpm at 37°C, transferred to a new tube and washed two times in high salt wash buffer. The beads were transferred to a new tube and washed three times with low salt wash buffer. The RNA was eluted from the beads in 100 µl elution buffer by heating the sample at 70°C for 15 min. The RNA was purified using phenol:chloroform extraction and precipitated with

ice-cold isopropanol and 0.3 M sodium acetate overnight. The RNA was pelleted by centrifuging at 16,000 rpm for 60 min, washed twice with ethanol and resuspended in 20 µl RNase and DNase-free water. The RNA was stored at -20°C until sequencing.

2.2.13. RNA sequencing

RNA sequencing was performed by the EMBL Genomics Core Facility. Following RNA-protein footprinting or DNA probe binding analysis, the RNA samples were prepared for sequencing using the NEXTFLEX® Small RNA-seq v3 Kit. The RNA was prepared according to the manufacturer's instructions, starting with 50 ng of RNA and using a 1:5 adaptor dilution and 20 PCR cycles. Clean-up was performed after reverse transcription. Multiple RNA samples were combined and sequenced on an Illumina MiSeq in 60-basepair single-end read mode. Charles Girardot (EMBL) assisted in analyzing the raw results from RNA sequencing using the EMBL galaxy platform. The barcode sequences introduced during multiplexing were utilized to allocate the reads to their respective samples, followed by reads trimming and mapping to the *Drosophila* genome. Any reads that were either unmapped to the genome or mapped to the *oskar* RNA were stored and subsequently mapped to the transgenic *lacZ* WT RNA. The results were then displayed using Integrative Genomics Viewer.

2.2.14. Mass spectrometry analysis

The protein samples after transcript-specific RBP capture were analyzed using mass spectrometry (performed by the EMBL Protein Core Facility). Cysteine-containing proteins were reduced via dithiothreitol to break disulfide bridges and treated with 2-chloroacetamide to alkylate the reduced cysteines. Samples were processed by an SP3 clean-up protocol (Hughes et al. 2014) and Trypsin digestion in a 1:50 ratio overnight at 37°C. Peptides were collected in HEPES, washed and labeled with TMT6plex Isobaric Label Reagent according to the manufacturer’s instructions, allowing for quantitative analysis of different samples. The high pH reverse phase fractionation was carried out prior to the mass spectrometry analysis.

2.2.15. Gene ontology (GO) analysis

The GO analysis of the proteins identified in the mass spectrometry analysis was performed using the programming language R (performed by the EMBL Protein Core Facility). The R function "bitr" was utilized to map the gene symbols of detected proteins to their corresponding Entrez IDs. The "enrichGO" function was used to conduct a gene ontology analysis based on either cellular compartment, molecular function or biological process. The odds-ratio of each GO term was calculated and visualized with the R function “calculateFE”.

3. Results

3.1. Mutations in SOLE affect *lacZ* RNA localization to the posterior pole

3.1.1. The presence of the *P-element* intron leads to an extra splicing event

Previous studies have demonstrated that the splicing event at the first intron, the SOLE and the exon junction complex (EJC) are necessary for the localization of *oskar* mRNA at the posterior pole of the oocyte (Ghosh et al. 2012). Mutational analysis revealed that the secondary structure of SOLE is important for *oskar* RNA localization, whereas the sequence of the SOLE is not essential. At the time, to exclusively study splicing at the SOLE position, the *oskar* transgene contained only the first *oskar* intron. Those studies were conducted *in vivo* by expressing transgenic *oskar* RNA using the Gal4-UASp system. The UASp vector contained a UAS regulatory sequence that is bound by the transcription factor GAL4 from *Saccharomyces cerevisiae* (Giniger et al. 1985). This characteristic was used to create a two-component expression system enabling precise regulation of transgenic *oskar* RNA expression in both time and space. For increased transcription efficiency, a *P-element* intron was inserted between the promoter and the 5' end of the gene to be expressed (O'Hare & Rubin 1983; Rørth 1998), in this instance the *oskar* gene (Figure 1.9; Figure 2.1A). Considering that the EJC is deposited upon splicing of RNA and additional splicing might increase the number of EJCs on the RNA (Figure 2.1A), it was unclear if the presence of the *P-element* intron affected the transport of the previously studied transgenic *oskar* mRNAs (Ghosh et al. 2012).

To test if the *P-element* intron has any effect on the localization of *oskar* RNA and to specifically investigate the EJC's function in *oskar* RNA transport, it was necessary to remove the *P-element* intron. I therefore generated a new *lacZ* WT transgene comprising the 5'UTR, 3'UTR, SOLE and first intron of *oskar* mRNA which would only be spliced at a single position, and therefore should bear only a single EJC per transcript, positioned immediately upstream of the SOLE (Figure 2.1B-C).

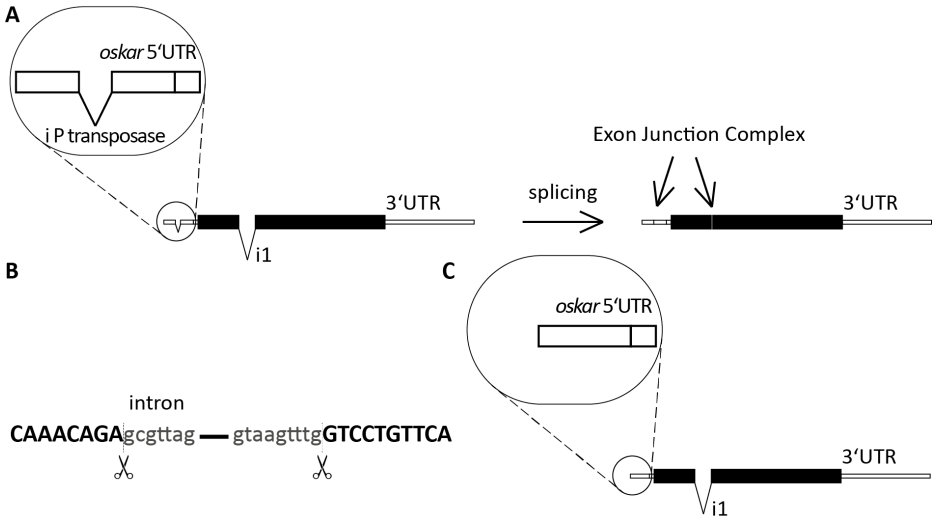

Figure 2.1 The presence of the *P-element* intron introduces an additional splice junction and resulting EJC deposition. (A) Transgenes used in previous studies were cloned in UASp vector that bears an intron 5' to the inserted RNA sequence. A *lacZ* RNA with an *oskar* intron would be spliced at two positions resulting in two possible EJC depositions. (B) The sequence of the *P-element* intron and its neighboring *P-element* coding region. (C) *lacZ* WT RNA after removal of the *P-element* intron.

3.1.2. *lacZ* constructs facilitate the analysis of the role of SOLE and EJC in *oskar* mRNA transport

In previous work (Ghosh et al. 2012), *oskar* transgenes with mutations in the SOLE sequence were created to investigate its role in posterior localization. Mutations in the proximal stem of SOLE (SL) that disrupt secondary structure were found to hinder *oskar* mRNA localization (Figure 2.2A), while introducing further mutations that restored complementarity (SLc) rescued RNA localization to the posterior pole.

In order to investigate the roles of the EJC and SOLE in *oskar* mRNA localization through both genetic and biochemical approaches, I generated four transgenic fly lines that express a *lacZ* sequence of a length similar to that of *oskar* RNA and contain the *oskar* 5'UTR, 3'UTR and the first intron. The *lacZ* WT transgene additionally contains the bipartite 28-nt SOLE sequence flanking the intron, which forms a stem-loop structure upon splicing. The *lacZ* SL transgene contains a mutated SOLE in which the stem-loop structure is disrupted, while the *lacZ* SLc transgene bears a compensatory sequence that restores base pairing and

thus the stem-loop structure. Furthermore, an intronless variant (*lacZ* IL), which lacks the intron and therefore cannot be spliced and should be devoid of an EJC, was also generated. As EJCs are typically displaced from RNA during translation (Dostie and Dreyfuss 2002; Lejeune et al. 2002; Gehring et al. 2009), I introduced a stop codon in all *lacZ* constructs to prevent further translation and displacement of the EJC by the ribosomes.

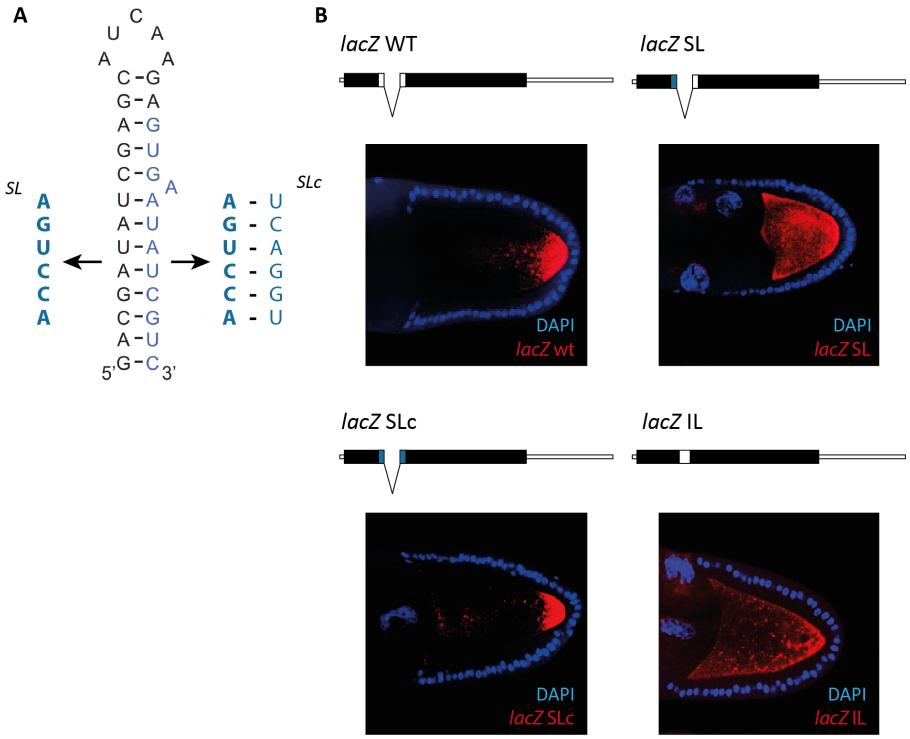

Figure 2.2 The stem-loop structure of SOLE and splicing are essential for the posterior localization of *lacZ* RNAs. (A) SL and SLc mutation introduced in the *lacZ* RNA. (B) FISH staining against the *lacZ* sequence (red) and DNA by DAPI (blue) in the four fly lines expressing transgenic *lacZ* RNA. *lacZ* WT and *lacZ* SLc RNA localize to the posterior pole as their SOLE sequence forms a stem-loop structure. Disrupting the stem-loop structure in *lacZ* SL abolishes posterior localization similar to the *lacZ* IL RNA, which is not spliced and therefore carries no EJC.

I next analyzed the localization of the transgenic RNAs in absence of endogenous *oskar* mRNA, by single molecule fluorescent *in situ* hybridization (smFISH). The *lacZ* WT RNA with the SOLE sequence localized to the posterior pole, mimicking endogenous *oskar* mRNA (Figure 2.2B). In contrast, the *lacZ* SL construct containing the disrupted stem-loop structure did not localize to the posterior pole. The compensatory rescue sequence *lacZ* SLc localized

to the posterior pole, whereas the intronless *lacZ* IL did not (Figure 2.2B). These experiments, in which the *P-element* intron was removed from the vector and only the first *oskar* intron was present reveal that splicing at the SOLE site is sufficient for RNA localization at the posterior pole. They also suggest that splicing of the *P-element* intron did not have any, or any major influence on the localization of the *oskar* transcripts in the previous study (Ghosh et al. 2012).

3.1.3. Analysis of *lacZ* RNA variant localization during oocyte development

The proper development of the *Drosophila* embryo is contingent upon the localization and translation of *oskar* mRNA at the posterior pole of the oocyte (Ephrussi et al. 1991; Kim-Ha et al. 1991). In the early stages of oogenesis development, *oskar* mRNA is transcribed in the nurse cells and exported by dynein into the oocyte, where it is enriched (Clark et al, 2007). During mid-oogenesis, the RNA is transported by kinesin to the posterior pole where it is exclusively translated (Brendza et al, 2000). Oskar protein is required to anchor *oskar* mRNA resulting in a gradual accumulation of *oskar* mRNA and protein at the posterior pole (Vanzo and Ephrussi 2002; Eichler 2022). This is especially important after egg chamber stage 10B when the nurse cells release their contents into the oocyte. During the ensuing cytoplasmic streaming, Oskar protein is required to maintain *oskar* mRNA to the posterior pole (Glotzer et al. 1997; Palacios and St Johnston 2002; Vanzo and Ephrussi 2002). In addition to Oskar protein-mediated anchoring, *oskar* mRNA is capable of co-transporting transgenic RNAs bearing the *oskar* 3'UTR to the posterior pole through a process called "hitch-hiking" (Hachet and Ephrussi 2004; Jambor et al. 2011). To prevent both hitch-hiking and anchoring, the kinesin transport of *lacZ* RNA in transgenic flies was analyzed in mutant flies in which endogenous *oskar* mRNA and protein are absent. Previous studies on *oskar* transgenes were also performed in such an "*oskar* RNA-null" background (Hachet and Ephrussi 2004, Ghosh et al. 2012)

During stage 8 of oogenesis, all four *lacZ* RNAs (WT, SL, SLc and IL) were enriched in the oocyte (Figure 2.3 and 2.4). At stage 9, the *lacZ* WT and *lacZ* SLc RNAs localized to the posterior pole comparable to endogenous *oskar* mRNA in wild-type flies (w^{1118} flies,

which are wild-type with respect to *oskar*). In contrast, the *lacZ* SL and *lacZ* IL RNAs failed to localize and instead were spread throughout the oocyte. During stage 10A, the *lacZ* WT and *lacZ* SLc RNA were not concentrated tightly at the posterior pole due to the absence of Oskar protein and anchoring. With the onset of cytoplasmic streaming in stage 10B the amount of *lacZ* WT and *lacZ* SL RNA localized at the posterior pole reduced. For *lacZ* SL and *lacZ* IL RNA, no evidence of specific localization at the posterior pole was detected (Figure 2.3 and 2.4).

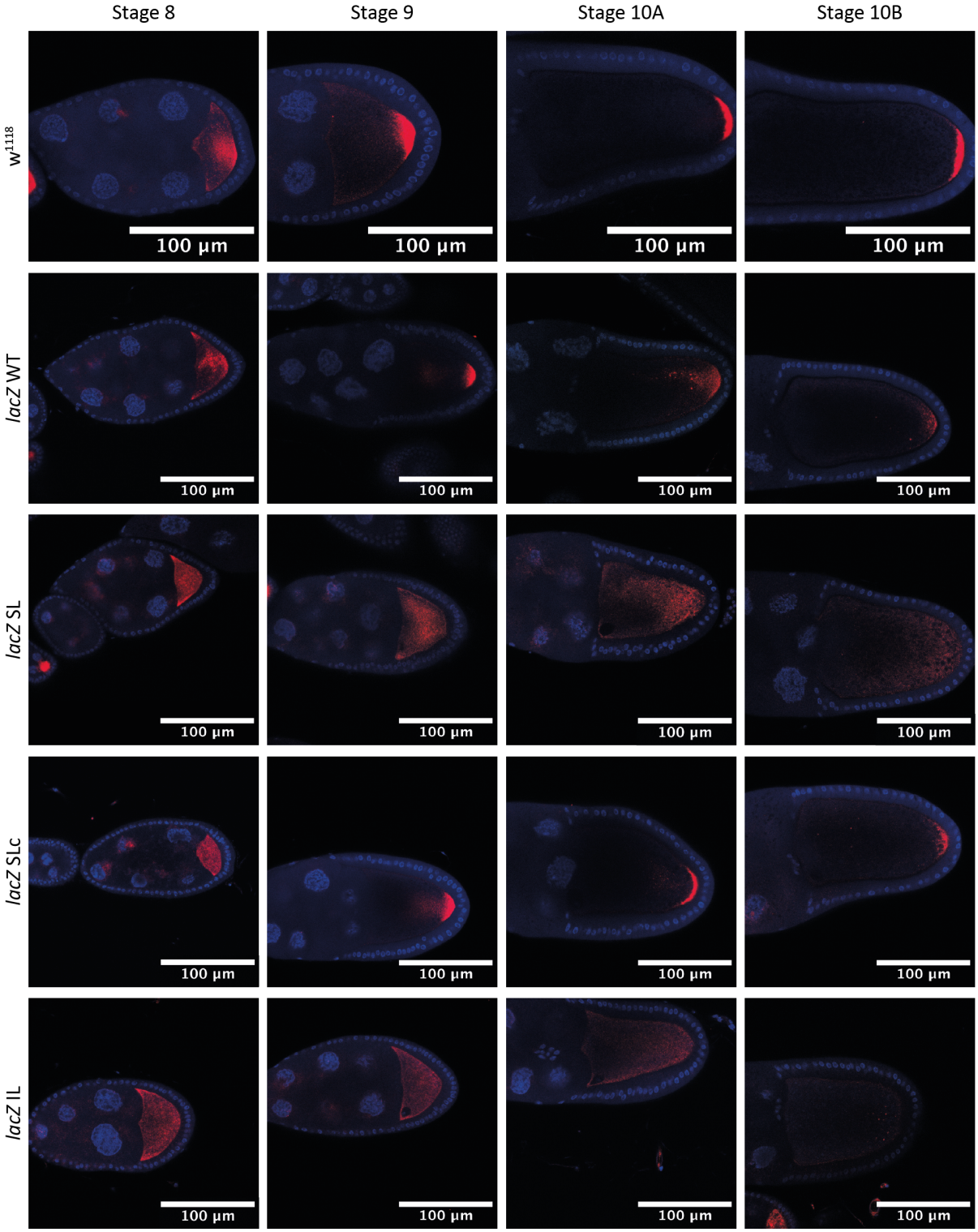

Figure 2.3 RNA localization to the oocyte posterior pole requires splicing-dependent assembly of a stem-loop forming SOLE. The transgenic *lacZ* RNAs were visualized by FISH (red) and DNA by DAPI staining (blue). *oskar*, *lacZ* WT and *lacZ* SLc RNA form a stem-loop structure upon splicing and localize to the posterior pole from stage 9 onwards. *lacZ* SL and *lacZ* IL RNAs lack the SOLE secondary structure and presumably the EJC, respectively. As a result, although both RNAs are transported into the oocyte, they fail to accumulate at the posterior pole.

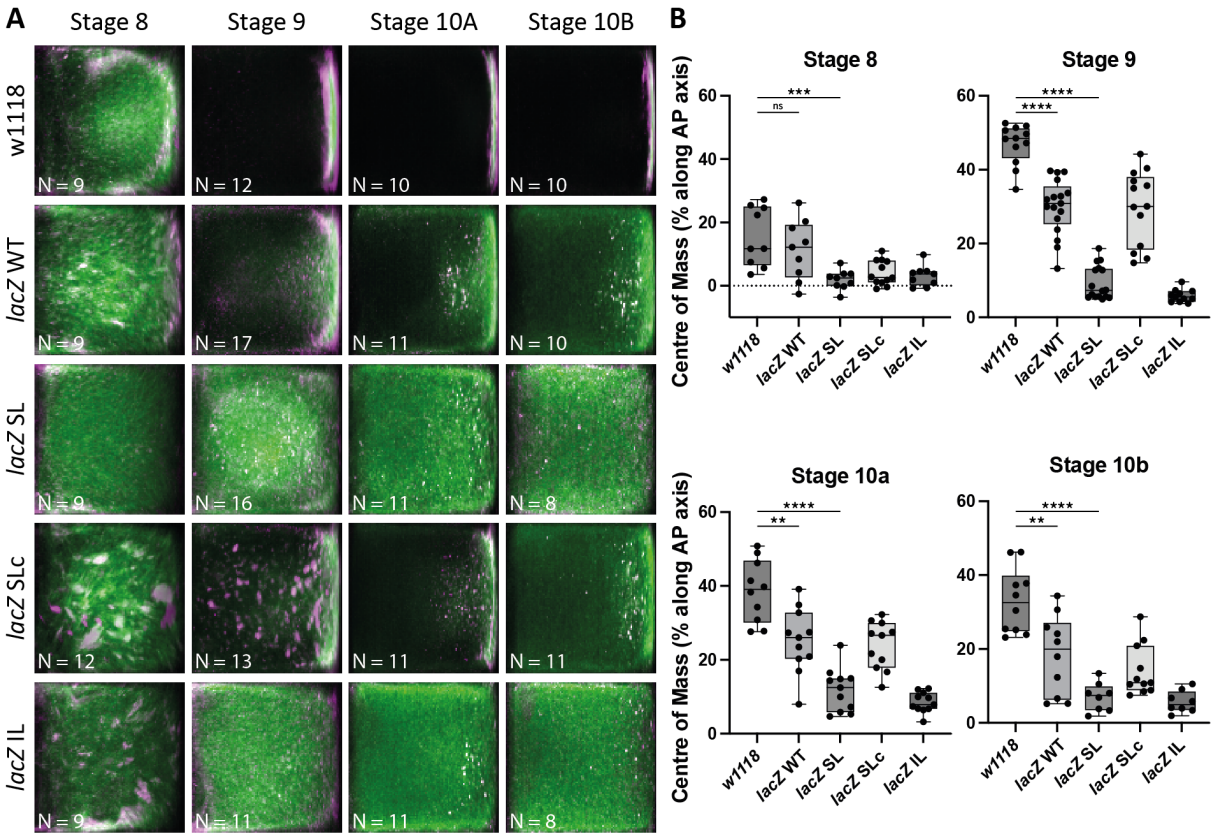

Figure 2.4 *oskar* 3'UTR and splicing to form a stem-loop SOLE are sufficient to localize the *lacZ* RNA to the posterior. (A) Analysis of RNA distribution by displaying the average distribution of RNA (green) of N oocytes visualized by smFISH. The purple colour represents a high concentration of RNA. The anterior is at the left and posterior at the right of the panels. N = number of oocytes imaged. (B) The RNA centre of mass distribution is visualized as a box plot with dots representing individual data points. The distribution along the anterior-posterior axis is relative to the oocyte centre, which is dedicated as 0 in the box plot. Unpaired Mann–Whitney U-tests for p-value calculation: ns = not significant, ** <0.01, *** <0.001, **** <0.0001

3.2. A stem-loop forming SOLE increases EJC stability on the RNA

The transport of *oskar* RNA to the posterior pole by kinesin requires splicing of the first intron, the EJC, and the presence of a stem-loop SOLE structure at the first exon-exon junction (Hachet and Ephrussi 2004, Palacios et al. 2004; Ghosh et al. 2012). However, the direct relationship between the EJC and the SOLE is unclear. Why the EJC is selectively bound at the first, and not at the second or third exon-exon junction in *oskar* RNA is not known (Obrdlik et al. 2019). The EJC can affect several downstream RNA processing events, such as RNA

degradation, translation and also localization by serving as a platform for recruiting peripheral proteins to the RNA (Le Hir et al. 2001; Tange et al. 2004).

In order to investigate whether presence of the SOLE stem-loop structure affects EJC stability on the RNA, I first analyzed protein occupancy on the *lacZ* WT (localizing) and *lacZ* SL (non-localizing) RNAs by performing a RNA-protein footprinting analysis (Figure 2.5A). I conducted the RNA-protein footprinting of *lacZ* WT and *lacZ* SL RNA *ex vivo*. Transgenic fly ovaries were crosslinked with UV and formaldehyde to stabilize protein binding after cell lysis. UV crosslinking stabilizes RNA-protein interactions (Hockensmith et al. 1986), whereas formaldehyde crosslinking stabilizes protein-protein interactions (Niranjanakumari et al. 2002). The *lacZ* RNAs were specifically enriched by utilizing biotinylated antisense DNA probes against the *lacZ* sequence and streptavidin beads. The *lacZ* RNAs were treated with RNase A, resulting in selective degradation of regions not bound by protein. Only the protein-bound regions, which the RNase could not access, remained intact and were sequenced to generate a gene coverage profile of the *lacZ* WT and *lacZ* SL RNA. The gene coverage profiles of the *lacZ* WT and *lacZ* SL RNAs were similar (Figure 2.5B). However, I observed that the *lacZ* WT RNA produced more reads and therefore a higher peak was present at the position of the first exon-exon junction. This could be a result of direct RNA-protein interactions at that site, or be due to the RNP structure, which might hinder RNA accessibility by the enzyme and obstruct RNA degradation, independent of direct protein binding.

To determine whether the variation in the reads at the exon-exon junction is due to direct RNA-protein binding or the RNP structure, I conducted an RNA-protein footprinting analysis specifically to detect direct RNA-protein interactions (Figure 2.5C). To accomplish this, I captured the *lacZ* RNAs from transgenic fly ovary lysates as previously described and in addition incubated the enriched *lacZ* RNAs at 70°C to reverse the protein-protein crosslinking that was mediated by formaldehyde, thereby leaving only the direct UV-crosslinked RNA-protein bonds intact. The footprinting analysis of direct RNA-protein binding of *lacZ* WT RNA (Figure 2.5C) produced a significant amount of reads at the EJC binding site (-20-24 upstream of the exon-exon junction). This demonstrated that proteins bind directly and in a stable manner at the EJC binding site of the *lacZ* WT RNA. The gene coverage profile of direct RNA-protein interactions of the *lacZ* SL RNA (Figure 2.5C) exhibited different RBP

binding pattern compared to the protein-protein complex protected RNA-protein footprinting with UV and formaldehyde crosslinking (Figure 2.5A). According to the *lacZ* SL gene coverage profile, the 3'UTR exhibited a higher protein binding occupancy than the *lacZ* coding region. Notably, no binding was detected at the EJC binding site in *lacZ* SL RNA.

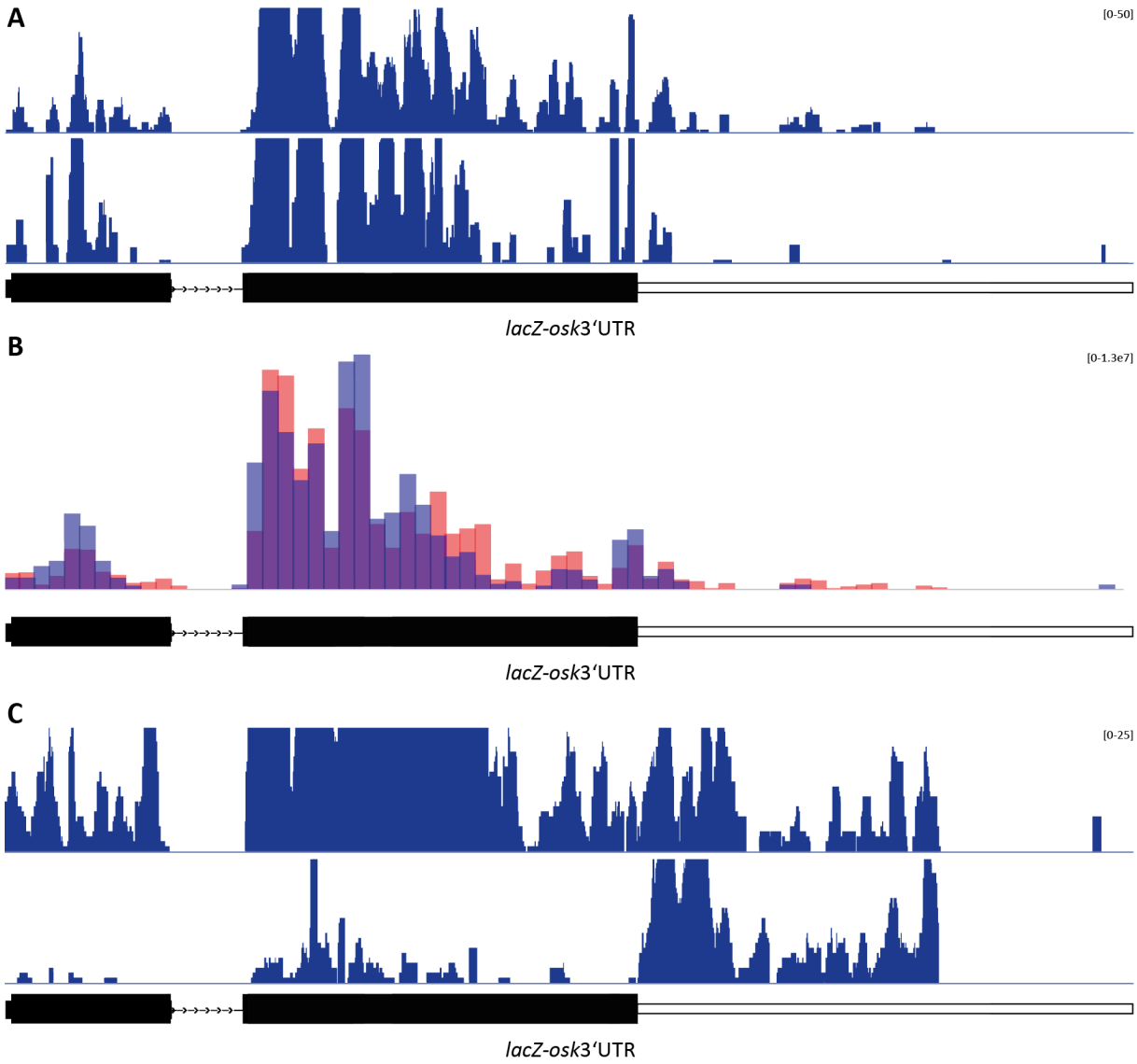

Figure 2.5 Gene coverage profiles after RNA sequencing of protein protected regions in *lacZ* WT and *lacZ* SL RNA. (A) RNA sequencing of protein-protein complex protected region in the *lacZ* RNAs. White smaller bar corresponds to the *oskar* 3'UTR and small line with arrows the first *oskar* intron. (B) Overlay of the gene coverage profiles visualizes the difference in protein-protein complex protected RNA regions. The EJC binding site is protected stronger in the *lacZ* WT (red) than *lacZ* SL RNA (blue). (C) Gene coverage profile of direct RNA-protein footprinting on the *lacZ* WT and *lacZ* SL RNA.

In order to investigate whether the stem-loop structure of the SOLE directly affects EJC stability on the RNA, I performed an RNA immunoprecipitation (RIP) experiment to determine if EJC is differentially stabilized on the *lacZ* WT and *lacZ* SL RNA (Figure 2.6A). To do so, I generated flies expressing GFP-Mago, a core component of the EJC, and either *lacZ* WT or *lacZ* SL RNA.

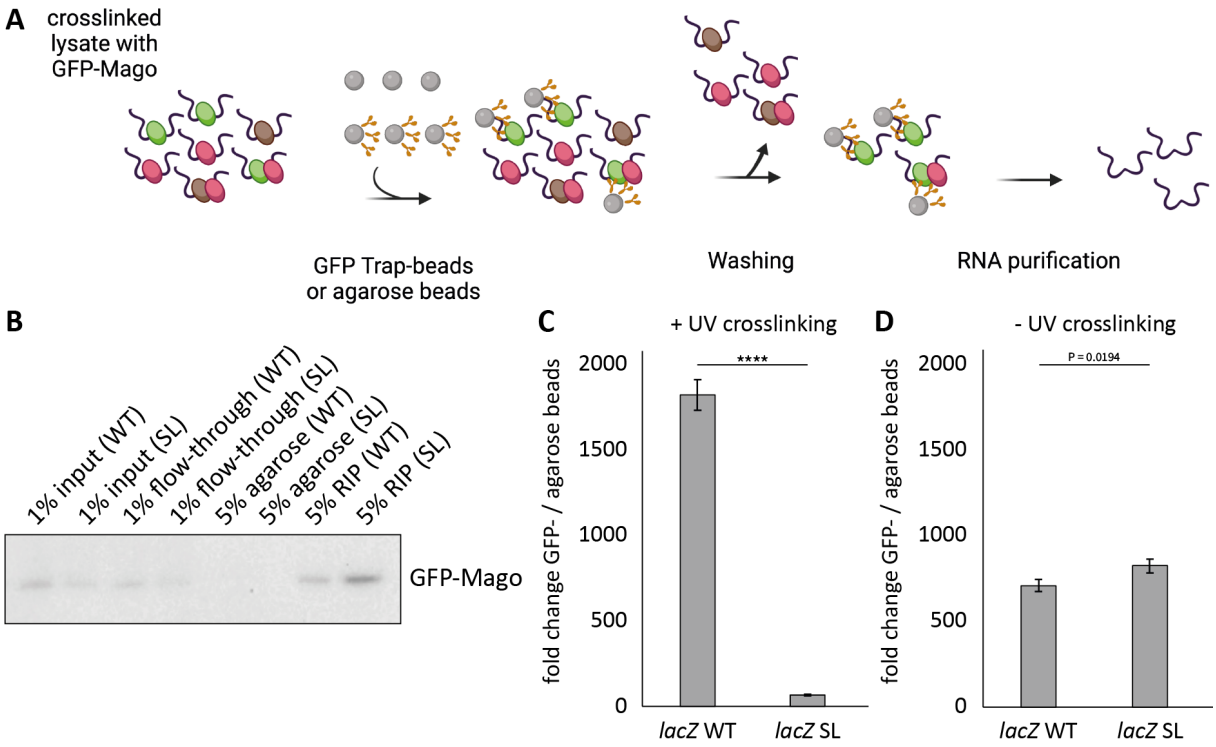

Figure 2.6 RNA immunoprecipitation against GFP-Mago reveals that upon UV crosslinking the EJC is stabilized on *lacZ* WT more than *lacZ* SL RNA. (A) Schematic view of the RIP protocol. Lysate is crosslinked and incubated with either GFP-TRAP beads or agarose beads and the RNA is analyzed after purification. (B) The western blot shows GFP-Mago detected by antibody staining against GFP. (C & D) RT-qPCR analysis of captured RNA by immunoprecipitating GFP-Mago using GFP-Trap beads.

The ovaries of the transgenic flies were collected and RNA-protein interactions were fixed through crosslinking with UV and formaldehyde. After immunoprecipitation, the efficiency of GFP-Mago capture was analyzed by western blotting (Figure 2.6B). The copurified RNA was quantified by RT-qPCR (Figure 2.6C), which revealed that the *lacZ* WT RNA was significantly more enriched by GFP-Mago RIP than the *lacZ* SL RNA, indicating that the EJC is more stably bound to the *lacZ* WT RNA (Figure 2.6C). Executing the RIP

experiment without crosslinking and RNA-protein stabilization did not demonstrate any difference between the captured *lacZ* WT and *lacZ* SL RNA (Figure 2.6D), indicating that the enrichment of *lacZ* WT RNA is due to direct EJC to RNA binding. These RIP experiments displayed that the SOLE stem-loop structure increased the stability of EJC binding to the *lacZ* WT RNA. This may be one explanation for why the EJC is selectively enriched at the first exon-exon junction site and why only splicing of the first *oskar* intron at the SOLE site can promote *oskar* mRNA localization.

3.3. Transcript-specific mRNP capture of transgenic *lacZ* RNAs

3.3.1. Kinesin is recruited to the *oskar* RNA independent of EJC and SOLE

During early oogenesis *oskar* mRNA is transported by dynein from the nurse cells into the oocyte (Clark et al. 2007) and during mid-oogenesis kinesin transports *oskar* RNA within the oocyte to the posterior pole (Brendza et al. 2000; Zimyanin et al. 2008). The EJC and SOLE are dispensable for dynein-dependent transport of the RNA from nurse cells to the oocyte. However, both the EJC and the SOLE are required for kinesin-dependent transport of the RNA within the oocyte (Hachet and Ephrussi 2004). Therefore, one hypothesis was that the EJC and SOLE are required for kinesin recruitment. In order to investigate this, Imre Gaspar created transgenic fly lines that expressed kinesin-mKate2, a fluorescently labeled protein, along with either full-length *oskar* RNA or only the *oskar* 3'UTR. Single molecular fluorescence *in situ* hybridization (smFISH) was utilized to analyze the colocalization of kinesin-mKate2 with *oskar* full-length RNA or *oskar* 3'UTR in the *Drosophila* oocyte (Figure 2.7A). This revealed that there was no significant difference between the degree of colocalization of kinesin-mKate2 molecules with full-length *oskar* RNA and *oskar* 3'UTR. Since the *oskar* 3'UTR lacks the SOLE and an EJC, this demonstrated that kinesin is recruited to *oskar* RNA independent of splicing, the SOLE, EJC and the coding region. This raised the question what the exact role of the EJC and SOLE is in the mechanism of *oskar* transport by kinesin. As it is not required for kinesin recruitment, it is likely that EJC/SOLE has a kinesin activating effect, e.g. by recruiting or stabilizing additional factors to the *oskar* RNP.

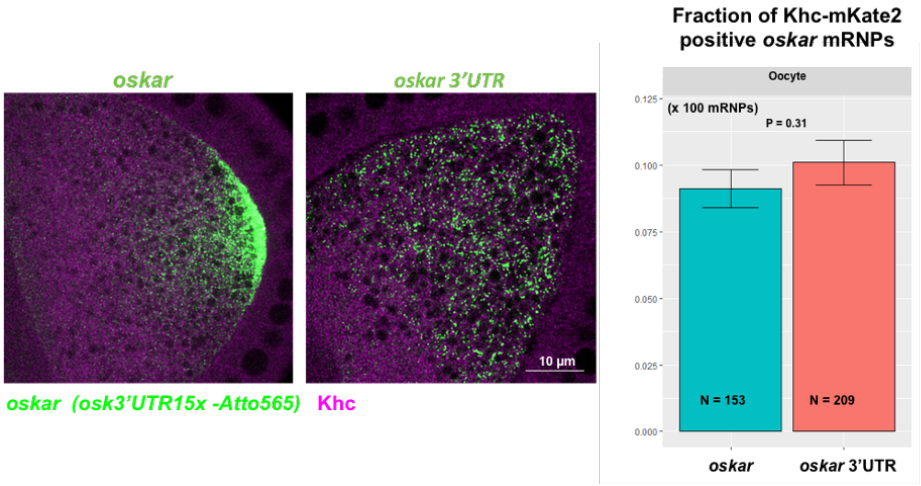

Figure 2.7 The *oskar* 3'UTR is sufficient to recruit kinesin to the RNA. (A) FISH staining of *oskar* full-length RNA (left, green) and *oskar* 3'UTR (right, green) and kinesin-mKate2 (purple). (B) Analysis of fractions of kinesin-mKate2 positive *oskar* mRNPs reveals no significant difference (Data and images from Imre Gáspár).

3.3.2. *oskar*-RNPs are copurified with *lacZ*-RNPs

Identifying the interactome of the EJC and SOLE could shed light on their role in *oskar* RNA transport and explain how they facilitate kinesin transport. Especially the comparison of the EJC-SOLE interactome of a localizing RNA (*lacZ* WT) and non-localizing RNA (*lacZ* SL) could identify proteins that are specifically enriched on the *lacZ* WT RNA required to activate kinesin transport. To identify the protein factors associated with the EJC and SOLE on *oskar* RNA I aimed to perform a transcript-specific mRNP capture using biotinylated antisense DNA oligonucleotides (ASOs) against the *lacZ* RNA. This method has been utilized in analyzing the composition of *oskar* RNPs (Appendix Figure S1, Wippich and Ephrussi 2020).

Typically for RNA-centric RBP capture and identification methods, a significant amount of starting material (RNA of interest) is required. The previously performed *oskar*-specific RBP capture required 10-15 ml ovaries per replicate, harvested from approximately 600 ml of flies (Wippich and Ephrussi 2020). As mentioned earlier, endogenous *oskar* mRNA is able to promote RNA localization of transgenic RNAs bearing the *oskar* 3'UTR to the posterior pole through hitch-hiking and Oskar protein facilitates anchoring and accumulation

of *oskar* RNA at the posterior pole (Hachet and Ephrussi 2004; Vanzo and Ephrussi 2002). Systematic analysis of transgenic RNA transport to the oocyte posterior pole requires the absence of endogenous *oskar* RNA, since it could confound the analysis. To prevent endogenous *oskar* RNA expression, a loxP cassette containing a 3xP3-EGFP marker was inserted after the first 26 nucleotides of the *oskar* transcription start site in an *oskar* allele ($osk^{attP,3P3GFP}$, Gáspár et al. 2017). This *oskar* allele was then recombined with the *lacZ* transgene containing chromosome. For experiments such as smFISH, RIP and RNA-protein footprinting transgenic *lacZ* flies only homozygous for osk^{3P3GFP} were used by crossing out the *oskar* RNA-producing allele.

Maintaining a stable stock of *oskar* "RNA-null" flies (osk^{null}) over generations is not possible, as the *oskar* 3'UTR is required for oogenesis progression and localized production of Oskar protein for *Drosophila* embryo development (Jenny et al. 2006). As a result, *oskar* RNA-null flies produce no progeny. Therefore, all stable *Drosophila* lines must express at least one wild-type *oskar* transcribing allele, either from a wild-type chromosome III (on which the *oskar* gene is located), or from a functional *oskar* transgene. Given that the *lacZ* RNA sequence contains the *oskar* 3'UTR, *Drosophila* oocyte development is rescued in the absence of endogenous *oskar* RNA, but embryo development is impaired, which results in no progeny being produced.

Generating enough material for transcript-specific RBP capture from all four transgenic flies (*lacZ* WT, SL, SLc and IL) in an *oskar* RNA-null background in triplicate is not feasible due to the large amount of starting material required (approximately 600 ml of flies to generate 10 to 15 ml of ovaries). Therefore, the option to use *lacZ* flies with endogenous *oskar* RNA, which can be maintained as a stable stock, needed to be explored. A potential issue for the transcript-specific RBP analysis in *oskar* mRNA expressing flies is the co-purification of *oskar* RNPs due to hitch-hiking, which could result in enrichment of factors involved in kinesin transport under all conditions, independently of SOLE integrity or EJC deposition on the *lacZ* RNA. It was therefore important to determine if the *lacZ* transcript-specific RBP capture of transgenic *lacZ* RNA in presence of endogenous *oskar* RNA exclusively captures *lacZ* RNA. To do so, I performed a transcript-specific RBP capture of *lacZ* WT RNA in flies, where endogenous *oskar* mRNA was present. Ovaries from the transgenic *lacZ* WT fly line

were crosslinked *ex vivo*, lysed and incubated with 45 different DNA oligonucleotide probes antisense to the *lacZ* RNA sequence and bearing a Biotin-TEG at the 3'end. Scrambled probes with random sequences were used as a negative control for the transcript-specific RBP capture (Figure 2.8A).

The RBP capture of *lacZ* WT RNA using specific *lacZ* probes successfully enriched *lacZ* WT RNA, but enriched endogenous *oskar* RNA as well. To determine whether the DNA probes are binding to *oskar* RNA or if *oskar* RNA is copurified with *lacZ* hitch-hiking with endogenous *oskar* RNA, I conducted a *lacZ*-specific mRNP capture in wild-type flies (w^{1118} flies), which do not express any *lacZ* RNA. In wild-type flies, neither *lacZ* nor *oskar* RNA was enriched (Figure 2.8B). This indicates that in ovaries from transgenic *lacZ* flies, the *oskar* RNA was copurified with the *lacZ* RNA owing to *oskar* 3'UTR-dependent co-packaging of transcripts in the RNP granules (Jambor et al. 2011). This presented a challenge for our intended transcript-specific RBP analysis, as the co-purification of *oskar* RNP would compromise the EJC interactome analysis. Therefore, to study the difference in the EJC-SOLE interactome between localizing (*lacZ* WT and *lacZ* SLc) and non-localizing RNAs (*lacZ* SL and *lacZ* IL), it was necessary to harvest ovaries from flies lacking endogenous *oskar* RNA.

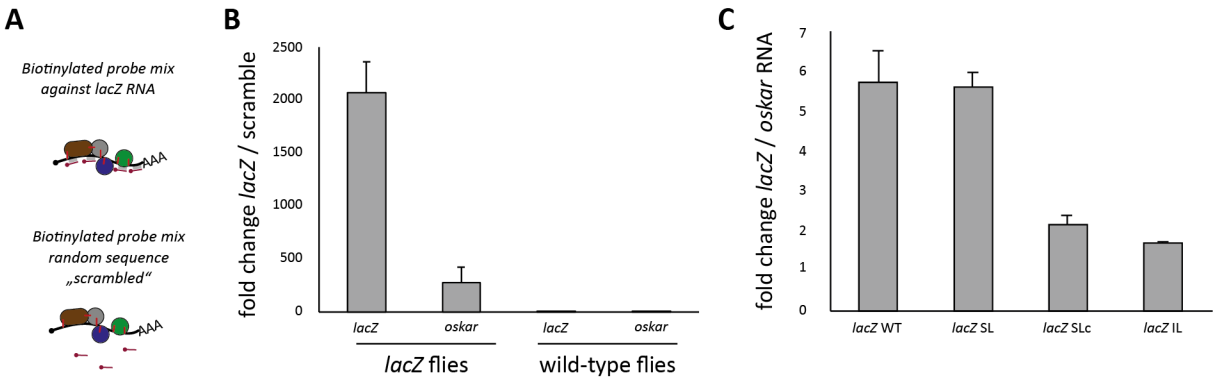

Figure 2.8 *oskar* RNP and *lacZ* WT RNP are pulled together during the transcript-specific RBP capture. (A) 45 different biotinylated DNA oligomers are used to capture the *lacZ* RNAs. As a negative control and to assess background information randomly designed DNA oligomers named "scrambled" are used. (B) RT-qPCR analysis of RNAs captured after the transcript-specific RBP capture in *lacZ* WT flies and wild-type flies. (C) Difference of *lacZ* and *oskar* RNA expression in presence of one allele with the *lacZ* gene and one allele with the *oskar* gene analyzed by RT-qPCR.

3.3.3. Optimizing sample generation

Since manually sorting approximately 600 ml of flies corresponding to 10 to 15 ml of ovaries in *oskar* RNA-null background (homozygous for $osk^{attP,3P3GFP}$) was not feasible, I used the *hid* gene under control of the heat-shock (hs) promoter (Starz-Gaiano et al. 2001). The hs-*hid* gene induces apoptosis upon heat activation, enabling the harvest of flies in an *oskar* RNA-null background. To achieve this, I generated flies bearing a hs-*hid* gene on the *oskar* RNA-expressing third chromosome (Figure 2.9A). After a heat-shock at 37°C, only larvae and pupae in an *oskar* RNA-null background (homozygous for $osk^{attP,3P3GFP}$) survive and hatch, which can then be used to harvest ovaries for the transcript-specific RBP capture.

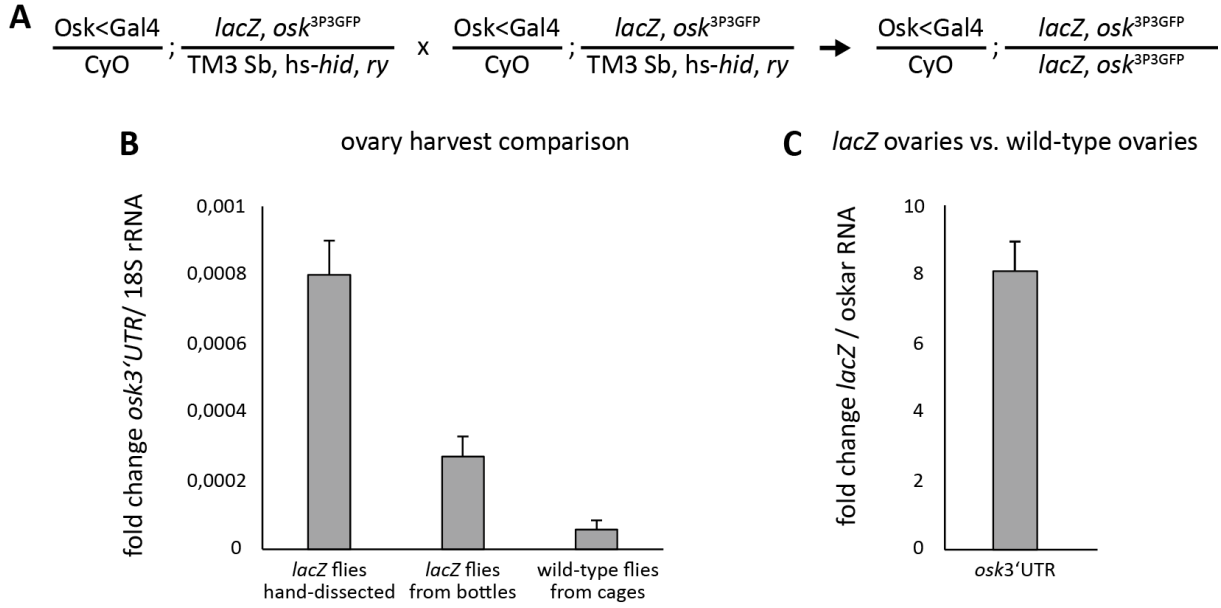

Figure 2.9 The quality of harvested *lacZ* WT ovaries is sufficient for transcript-specific RBP capture. (A) RT-qPCR analysis of *lacZ* WT and ribosomal *18S* RNA is performed to analyze the ovary-to-background ratio. (B) RT-qPCR analysis of the amount of *lacZ* WT and endogenous *oskar* RNA of ovaries from similar volume generated from mechanically grinding transgenic *lacZ* and wild-type flies, respectively.

Smaller numbers of flies can be maintained in vials or bottles, while larger numbers are typically kept in large cages. However, I found that my *lacZ* mutant flies could not be successfully amplified or maintained in cages, as the number of flies decreased over each generation, likely due to the *lacZ* mutant flies being less healthy than wild-type flies.

Additionally, the *lacZ* mutant flies have one *osk*^{attP,3P3,GFP} chromosome and one bearing an *oskar* allele. This means that approximately half of the *lacZ* mutant offspring are homozygous for *osk*^{attP,3P3,GFP} and thus lack Oskar protein, leading to the death of the embryos (Ephrussi et al. 1991; Kim-Ha et al. 1991). Therefore, I amplified and harvested flies from bottles rather than cages. Large amounts of ovaries were harvested by mechanically grinding flies and size sorting the egg-chambers using sieves of different mesh sizes. The method was essentially as described in previous studies (Jambor et al. 2016; Wippich and Ephrussi 2020), however additionally the crushed sample was cleared using a 120-micron sieve to remove egg chambers older than stage 12. As my study was focused on the role of the EJC in *oskar* transport by kinesin, which occurs during mid-oogenesis (Brendza et al. 2000; Zimyanin et al. 2008), I sought to exclude later-stage egg chambers, which are much larger and would otherwise constitute the bulk of the material. Therefore, the additional size selection step should reduce background noise in the downstream mass spectrometry analysis.

Having modified the strategy for fly amplification and ovary harvest of *lacZ* flies, I assessed the quality of the *lacZ* ovary samples by measuring the ratio of transgenic *oskar* 3'UTR to *18S* ribosomal RNA (rRNA). As a reference, I compared the ratio of endogenous *oskar* 3'UTR to *18S* rRNA in hand-dissected wild-type ovaries with that in ovaries of flies from a wild-type population cage that were harvested by mechanical grinding. Directly dissected wild-type ovaries yielded the highest endogenous *oskar*:*18S* ratio, no doubt due to the absence of contamination from body parts (Figure 2.9B). However, the *lacZ* RNA sample had a lower *oskar*:*18S* ratio than the hand-dissected wild-type ovary samples, but a higher ratio than the wild-type ovary samples obtained from flies in cages. The quantities of *lacZ* and *oskar* RNA were measured from *lacZ* and wild-type ovaries, respectively. The results showed that *lacZ* RNA from *lacZ* ovaries was more highly enriched than *oskar* RNA in wild-type ovaries (Figure 2.9C). Since the wild-type ovaries from population cages were used in a previous *oskar*-specific RBP capture (Wippich and Ephrussi 2020), the results demonstrated that the adjusted harvesting protocol was able to generate an ovary lysate of sufficient quality to proceed to transcript-specific RBP capture.

3.3.4. DNA probe optimization

The DNA probes against *lacZ* RNA for transcript-specific RBP capture were designed to have similar features (melting temperature, GC content, length). Since they were targeted to random sites in the *lacZ* coding region, the hybridization efficiency of the DNA probes for *lacZ* RNA capture needed to be tested. Amplifying and harvesting *lacZ* mutant flies in an *oskar* RNA-null background is easier when using the hs-*hid* selection strategy, but the amplification of bottles rather than cages impedes *lacZ* ovary harvest. To save *lacZ* ovary material, I validated the efficacy of the 45 DNA probes designed for capturing *lacZ* RNA and associated RBPs in wild-type ovary lysate supplemented with *in vitro* transcribed *lacZ* WT RNA.

The *in vitro* transcribed *lacZ* WT RNA capture was successful, indicating that the probes are functional (Figure 2.10A). To confirm the enrichment of *lacZ* WT specific RBPs, I measured the enrichment of Staufen, a bona fide *oskar* 3'UTR binding protein. However, the capture yielded a lower enrichment of Staufen than previously reported RBP captures of endogenous *oskar* RNA in wild-type ovaries (Wippich and Ephrussi 2020). Therefore, it seemed plausible that a significant amount of my DNA probes did not efficiently bind the *lacZ* WT RNA. To improve the capture efficiency, I designed a second set of DNA probes targeted to different sites on the *lacZ* RNA, but this did not result in any significant increase in the total amount of *lacZ* RNA captured (Figure 2.10B).

Given that the probes were designed to target random sites in the *lacZ* RNA sequence, it is possible that not all probes could bind effectively due to RNA secondary structure or inaccessibility as a result of protein binding to some RNA sequences. Biotinylated DNA probes that do not efficiently hybridize to the *lacZ* RNA are competing with *lacZ* RNA:DNA hybrids for binding to streptavidin beads. This competition could reduce the binding capacity of streptavidin beads, which could reduce the efficiency of the capture.

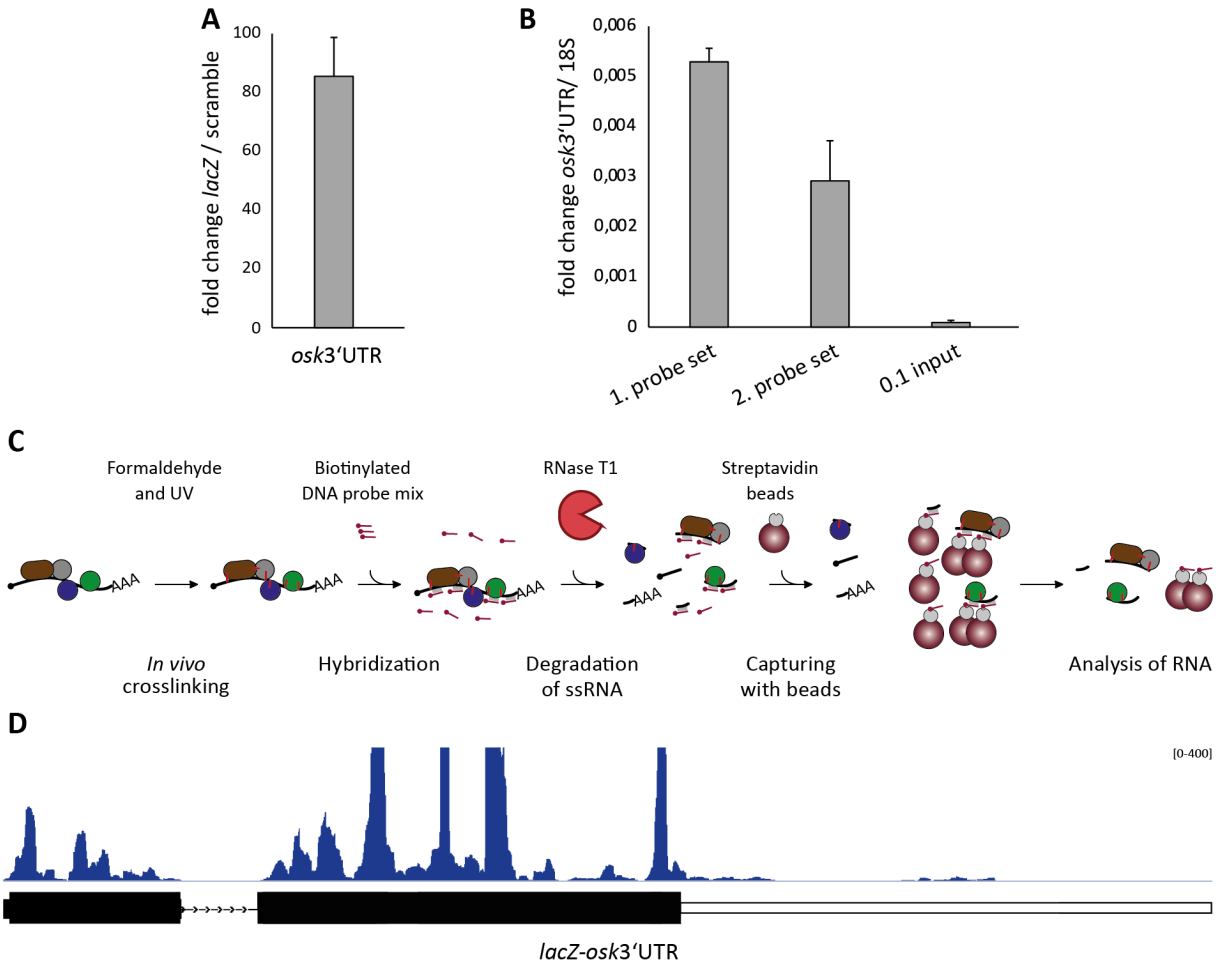

Figure 2.10 The evaluation of probe binding efficiency facilitates the optimization of the DNA probe set for effective capture of the desired RNA. (A) RT-qPCR analysis after a transcript-specific enrichment of *in vitro* transcribed *lacZ* WT RNA in wild-type lysate. The initially designed DNA probe set of 45 oligomers is able to enrich the RNA of interest. (B) RT-qPCR analysis of a transcript-specific pulldown using two different DNA probe sets, with the input sample demonstrating successful enrichment. However, there is so significant difference between the two probe sets. (C) Schematic overview of analyzing DNA probe binding efficiency. Subsequent to DNA:RNA hybridization the RNA is cleaved into smaller fragments by RNase T1. Magnetic streptavidin beads are then used to capture the DNA:RNA hybrids. The RNA fragments are then sequenced. (D) Sequenced RNA fragments are used to create a gene coverage profile. Higher peaks correspond to more reads and thus to a higher number of captured RNA fragments.

To identify the probes that actually bind the RNA, I modified the RBP capture protocol by first sequencing RNA fragments that hybridized with the DNA probes. After incubating the 45 DNA probes from the first probe set in *lacZ* ovary lysate, the total RNA was degraded by RNase T1. RNase T1 cleaves single-stranded RNA after a guanine position, thus only double-stranded RNA, RNA fragments protected by RBPs and RNA fragments hybridized to

DNA probes remain intact (Pace et al. 1991). The RNA:DNA hybrids were then specifically enriched using streptavidin beads and subjected to RNA sequencing. This allowed me to generate a *lacZ* gene coverage profile to identify the probes that bound effectively (Figure 2.10C & D). The binding efficiency of each DNA probe was analyzed by mapping the 45-oligomer probe set to the *lacZ* gene coverage profile. I selected 21 DNA probes that mapped to the highest peaks in the *lacZ* gene coverage profile and trimmed the probe set to include only these. Through the removal of inefficient DNA probes, I was able to decrease the proportion of unbound biotinylated probes competing for binding to the streptavidin beads, thereby increasing the specificity of the capture process.

3.3.5. Adjustment of beads to probe ratio

The initial probe set, which consisted of 45 DNA probes, was reduced to 21 DNA probes using the preceding DNA probe optimization step. I validated the functionality of the trimmed DNA probe set (21 probes) by performing transcript-specific RBP capture using ovaries from *lacZ* flies that do not express endogenous *oskar* mRNA (transgene in *oskar* RNA-null background). Comparing the total RNA captured using the optimized 21 DNA probes set with that of the 45 DNA probes showed a reduction in the total amount of captured RNA (Figure 2.11A). Although I successfully reduced the background, as was shown by the lower amount of ribosomal *18S* RNA captured, I was unable to further enrich *lacZ* RNA. One possible explanation for the lack of enrichment in captured RNA could be the streptavidin beads being the limiting factor. The transcript-specific RBP capture used streptavidin beads with a binding capacity of approximately half of the total amount of biotinylated DNA probes present in one pull-down. Thus, in principle the beads should be able to bind a maximum of half of the probes in a given reaction. This ratio was chosen in a previous study to minimize background corresponding to unspecific RNA and protein (Wippich and Ephrussi 2020). I increased the amount of beads, which resulted in an increase of captured *lacZ* WT RNA (Figure 2.11C), with approximately 15% of input RNA being captured (Figure 2.11D). Unfortunately, this also increased the *18S* rRNA, which is an indication of unspecific RNA capture. Analyzing the levels of captured Staufen protein confirmed that, besides RNA, RBPs were being enriched as well (Figure 2.11E).

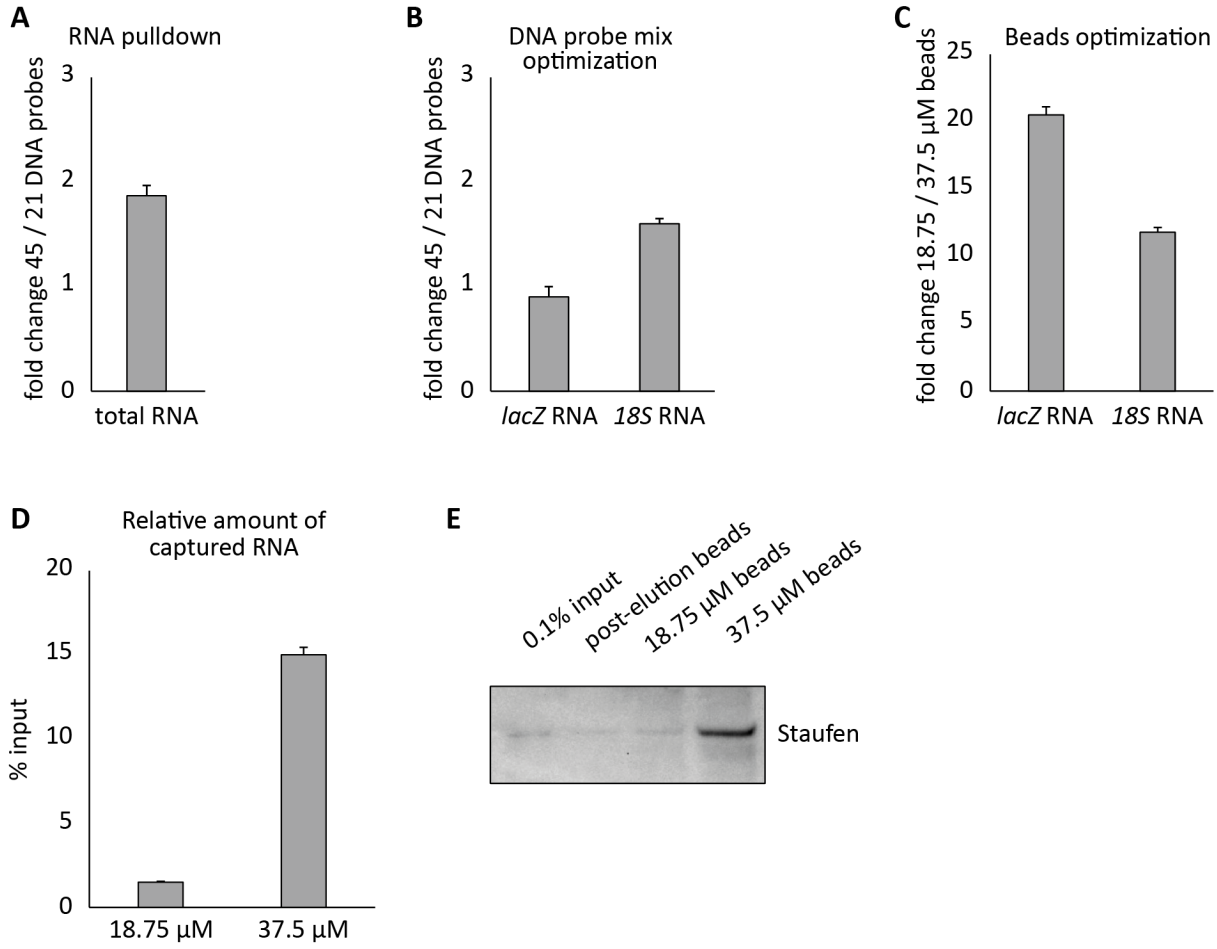

Figure 2.11 Increasing the number of beads and the resulting biotin-binding capacity improves the amount of RNA and protein captured. (A) Total RNA captured from a transcript-specific pulldown decreases to almost half after using the optimized DNA probe mix. (B) RT-qPCR analysis of captured RNA using both DNA probe sets reveals that the *lacZ* WT RNA amount does not significantly change after optimizing the biotinylated oligomers. However, the amount of ribosomal RNA (*18S*) is reduced when using the optimized DNA pool. (C) RT-qPCR analysis of *lacZ* WT and *18S* RNA after increasing the number of streptavidin beads from a binding capacity of 18.75 to 37.5 µM. The amount of *lacZ* WT and *18S* RNA is increased, whereas the first is increased more than the background RNA. (D) RT-qPCR analysis of the relative *lacZ* WT RNA captured from the input. Increasing the number of beads significantly increases the percentage of captured *lacZ* WT RNA. (E) Western blot analysis demonstrates that not only the RNA but also its RBPs are enriched by analyzing the known *oskar3*'UTR binding protein Staufen.

3.4. Change in RBP composition upon manipulating the SOLE sequence

3.4.1. Optimized RBP and RNA capture

Identifying the RBPs on a specific RNA by utilizing biotinylated DNA probes requires a significant amount of input sample, since pulling on an RNA is generally inefficient (Gerber 2021). Due to the inability to maintain the *lacZ* WT flies in an *oskar* RNA-null background and establish a stable stock, the sample generation step became the bottleneck in performing the RBP capture. This increased the necessity to improve the efficiency of the protocol at several steps including sample generation, DNA probes, and streptavidin beads to probe ratio.

I aimed to identify the RBPs of all transgenic *lacZ* RNAs and identify differences in RNP composition between localizing (*lacZ* WT and *lacZ* SLc) and non-localizing RNA (*lacZ* SL and lacZ IL) using mass spectrometry analysis. To assess the RBP capture quality, I analyzed the relative amount of captured RNA by RT-qPCR, as well as the enrichment of known *oskar* 3'UTR binding proteins by western blot (Figure 2.12). Staufen, a bona fide *oskar* 3'UTR binding partner, was significantly enriched over the negative control and actin, serving as a marker for non-specific proteins, was successfully washed away in all four transcript-specific RBP captures. The enrichment of *lacZ* RNA in all mass spectrometry samples suggested that the proteins enriched are putative *lacZ* or *oskar* 3'UTR binding partners. With a successful RBP capture for *lacZ* established, the next step was to perform mass spectrometry to analyze the RBP composition of all four transgenic *lacZ* RNAs and identify differences between their respective EJC-SOLE interactomes.

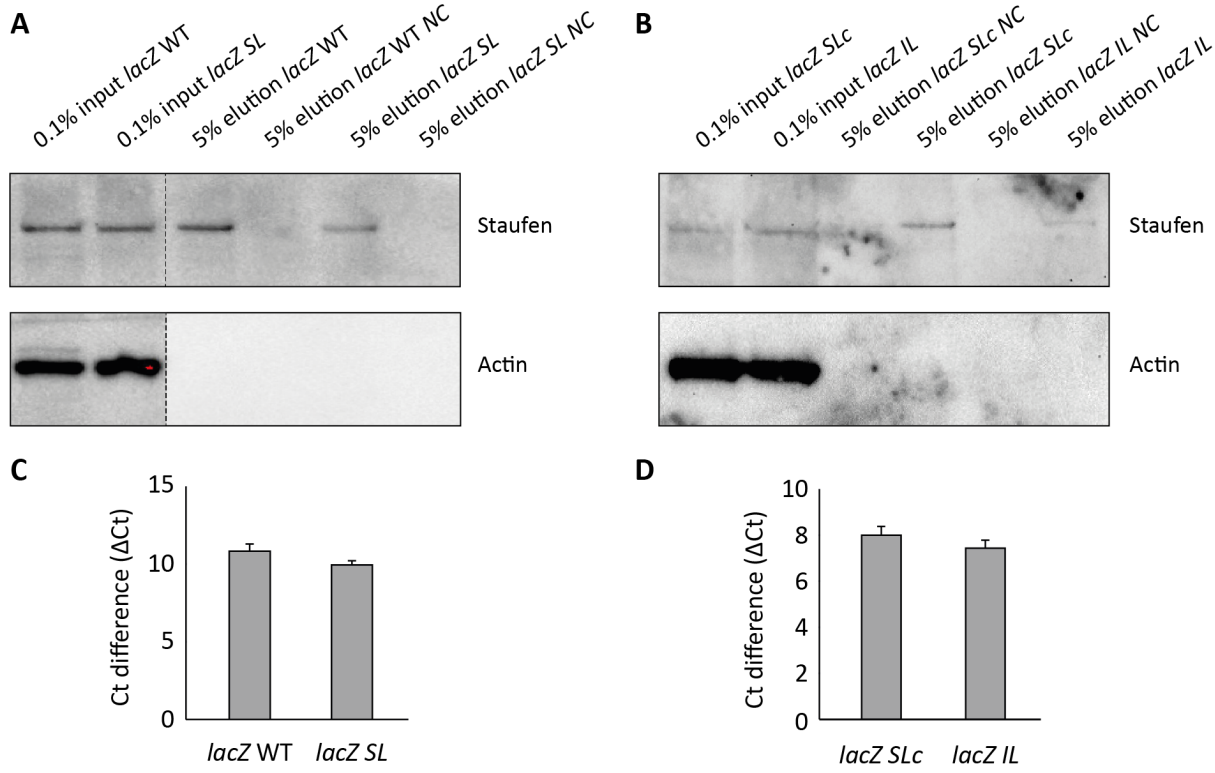

Figure 2.12 Quality assessment on a protein and RNA level demonstrates successful transcript-specific RBP capture. (A, B) Western blot analysis of the Staufen protein level in the pulldown. Staufen serves as a control as it is a known *oskar* 3'UTR binding protein. Its presence in the elution and absence in the scrambled sample (negative control) reveals that the RBPs are significantly enriched. Analyzing Actin protein shows that unspecific proteins are successfully washed away. (C, D) RT-qPCR analysis of eluted *lacZ* RNA shows a significant enrichment over the negative control.

3.4.2. The difference in RBPs between localizing and non-localizing *lacZ* RNA

Transcript-specific RBPs were successfully enriched and used for an unbiased quantitative protein identification using mass spectrometry by utilizing in-solution digestion, SP3 clean-up (Hughes et al. 2014), and isobaric tandem mass tag (TMT) labelling for multiplexed proteomics quantification. The mass spectrometry data of each sample was compared to its negative control to identify which proteins are *lacZ* transcript-specific. The enriched proteins in the *lacZ* WT sample were plotted against those in the *lacZ* SL sample (Figure 2.13). The first mass spectrometry analysis revealed a number of proteins enriched in

the *lacZ* WT over the *lacZ* SL mutant sample. The colocalization analysis of kinesin with either full-length *oskar* RNA or *oskar* 3'UTR revealed that there is no significant difference in kinesin recruitment (Figure 2.7). However, kinesin remained in an inactive state since no transport to the posterior pole occurred. Therefore, it would seem more likely that an activating factor is missing in a non-localizing RNA such as *lacZ* SL and *lacZ* IL, rather than a factor that is blocking the transport being present.

A significant number of proteins were enriched in the *lacZ* WT condition, but more hits were present in the *lacZ* SL sample. By inspection of the enriched hits in the *lacZ* SL condition, I detected proteins such as Bruno (*aret*) and Cup, which act as translational repressors and are known to bind to the *oskar* 3'UTR. These proteins likely dissociate from the RNA upon *oskar* RNP remodeling at the posterior pole, resulting in the alleviation of translational repression. The fact that the *lacZ* SL RNA does not localize to the posterior pole could explain why several *oskar* 3'UTR binding proteins involved in translational repression are slightly more enriched in the *lacZ* SL sample. After repeating the transcript-specific RBP capture and normalizing all runs, the number of significant differences between the *lacZ* WT and *lacZ* SL samples decreased. But the tendency of many known *oskar* 3'UTR binding proteins being slightly more enriched in the *lacZ* SL RNA persisted. In contrast, many significant hits in the *lacZ* WT RNA during the first run did not remain after multiple replicates. In fact, displaying the normalized mass spectrometry results in a fold change correlation plot demonstrates that the RNP composition of *lacZ* WT and *lacZ* SL RNA are mostly similar (Figure 2.13B), despite the different localization patterns of *lacZ* WT and *lacZ* SL RNA in the *Drosophila* oocyte.

Small differences in the RNP composition may not be detected by mass spectrometry analysis. However, even the RNP composition of *lacZ* SLc and *lacZ* IL are shown to be similar (Figure 2.14), despite the absence of introns and thus splicing and EJC deposition in *lacZ* IL RNA. Notably, mass spectrometry analysis did not detect any differences in the EJC core components in any of the *lacZ* samples. Even, the RBP composition of *lacZ* IL and *lacZ* SLc did not show any differences in the EJC core components, despite the *lacZ* IL RNA lacking splicing and EJC deposition. I incubated *lacZ* ovary lysate with streptavidin and carboxy beads and followed the transcript-specific RBP capture protocol, without using any DNA

probes. The EJC core component eIF4A3 was shown to bind to the bead matrix independently of DNA probes, *lacZ* RNA enrichment, biotin or streptavidin (Appendix Figure S2). This explains why the EJC was detected in all *lacZ* samples and negative controls, which makes it challenging to distinguish any differences. Barentsz was the only EJC core component that was slightly differently enriched after transcript-specific RBP capture of transgenic *lacZ* RNAs. In the *lacZ* WT and *lacZ* SL analysis, Barentsz was slightly more enriched in the *lacZ* WT RNA, which is in line with the RIP experiments that showed that the EJC is more stably bound to the *lacZ* WT RNA due to the SOLE stem-loop structure. In the RNP analysis of *lacZ* SLc and *lacZ* IL, Barentsz was slightly more enriched in the *lacZ* SLc sample, which is expected due to the absence of splicing and EJC deposition in *lacZ* IL RNA.

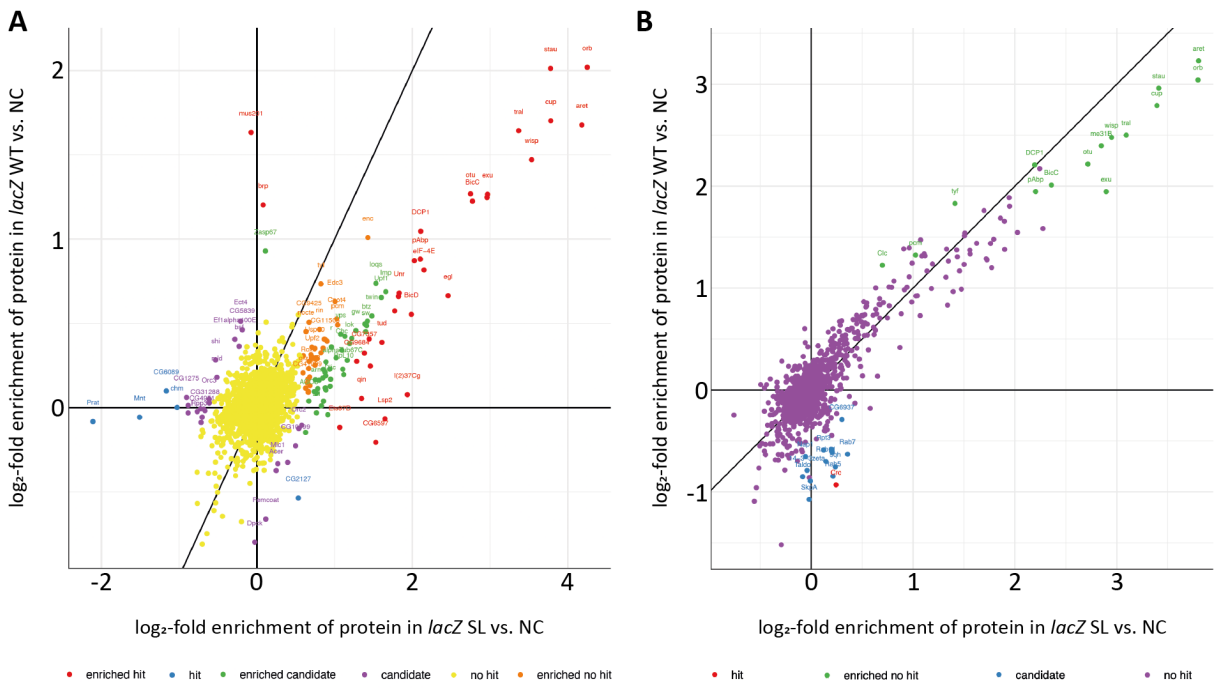

Figure 2.13 Difference of RBP composition of *lacZ* WT and *lacZ* SL RNA detected via transcript-specific RBP capture and mass spectrometry analysis. (A, B) Fold enrichment correlation of enriched proteins detected in *lacZ* WT and *lacZ* SL samples. The closer the hits near the x or y axe, the more prominent they are present only in a single condition. Proteins close to the diagonal are similarly enriched in both conditions. (A) First RBP capture and mass spectrometry analysis demonstrated several hits in each condition, however (B) normalizing after several replicates revealed that the RBP composition is mostly similar. Enlarged images are in Appendix.

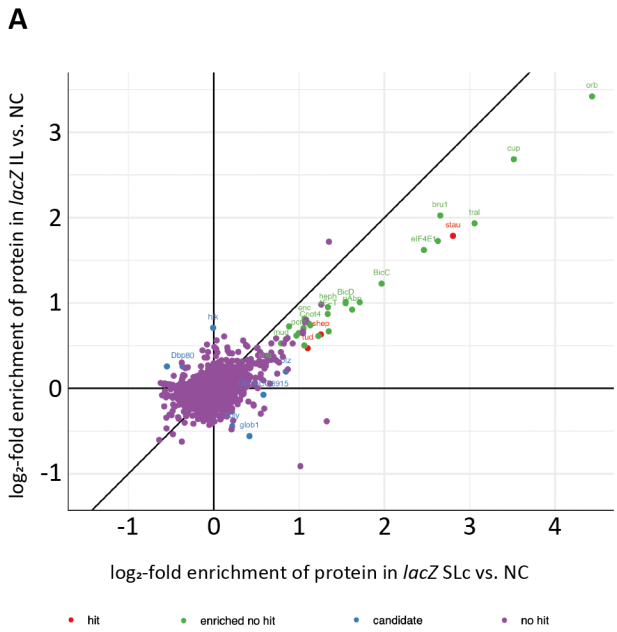

Figure 2.14 RBP composition difference in *lacZ* IL and *lacZ* SLc. (A) Fold enrichment correlation of RBPs in *lacZ* IL and *lacZ* SLc sample. Mass spectrometry analyses were performed in triplicates. Enlarged image is in Appendix.

3.4.3. Proteins with a putative role in *oskar* RNA transport

The majority of the detected RBPs were known to either directly bind the *oskar* 3'UTR or be a part of the *oskar* RNP complex. To analyze the detected RBPs by the *lacZ*-specific RBP capture, I performed a gene ontology (GO) enrichment analysis for biological processes, which revealed that mainly RNA processing proteins were enriched (Figure 2.15). This confirms the high sensitivity of the transcript-specific RBP capture for identifying *lacZ* RNA-specific RBPs, given that proteins involved in RNA processing and many known *oskar* RNP components were among the enriched proteins. However, the identification of differences in RBP composition between a localizing and non-localizing *lacZ* RNA was not entirely conclusive.

For the RBP capture approach, I harvested *Drosophila* ovaries by mechanically grinding and sieving flies. The resulting ovary sample was heterogeneous, containing egg chambers of different stages, but predominantly consisting of early- and mid-stage chambers. Kinesin transport of *oskar* RNA is occurring in mid-oogenesis, but in early oogenesis the dynein

transport is dominant. Because the ovary sample for the *lacZ*-specific RBP capture includes both early and mid-stage egg chambers, *lacZ* RNA and its RBPs during dynein (early oogenesis) and kinesin (mid oogenesis) transport are both captured. Mutations in the SOLE that disrupt the stem-loop secondary structure in *lacZ* SL do not impede the dynein transport. Therefore, the RNP complex in early oogenesis might not significantly vary between *lacZ* WT and *lacZ* SL RNA, which could result in a significant overlap in the RNPs captured in both samples. Consequently, differences in RBP composition that are due to either active or inactive kinesin transport might be entirely or partially masked.

Therefore, it is possible that even proteins that show only slight enrichment in the localizing condition could be involved in kinesin-dependent *oskar* RNA transport. To investigate this, I selected a set of candidate proteins (Figure 2.16). I primarily chose proteins that were higher enriched in the localizing than in the non-localizing condition. In addition, I selected some proteins that were selectively enriched in only one condition, even if to a lesser degree, in order to validate the mass spectrometry data. Proteins were chosen depending if they have an RNA binding site, are known to interact with motor proteins, the cytoskeleton or *oskar* RNA. Unannotated proteins significantly enriched were also included in the list. I will carry out RNAi screens of the selected proteins by using flies from the Transgenic RNAi Project (TRiP) to conditionally knock down the gene of interest in the germline and investigate if *oskar* localization is impaired.

Overall, the mass spectrometry results of this study provide valuable insights into the RBP composition of the *lacZ* RNP complex. The identification of candidate proteins through the *lacZ*-specific RBP capture approach and subsequent knock-down experiments will provide a more comprehensive understanding of the molecular mechanisms involved in *oskar* mRNA localization to the posterior pole, which is critical for proper embryonic development in *Drosophila*.

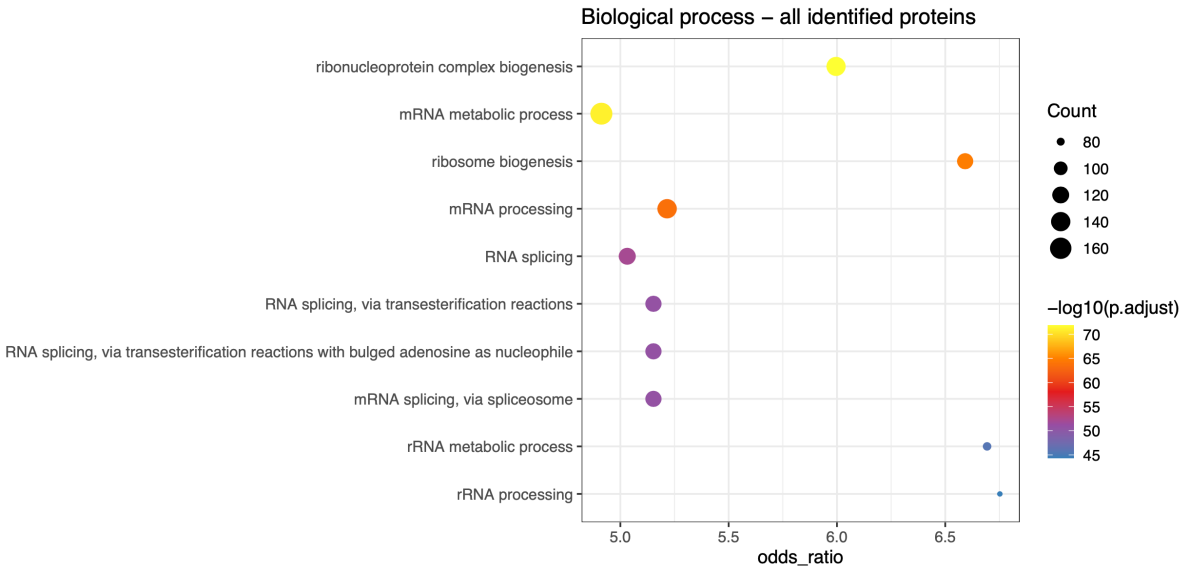

Figure 2.15 Biological process gene ontology analysis of all identified proteins in the *lacZ* WT and *lacZ* SL RBP capture. (A) All detected proteins are classified according to their GO biological process which reveals that most RBPs are involved in RNA processes. This emphasizes that the detected proteins have a high true positive rate as most proteins are known to either interact with *oskar* RNA or RNA in general. The size of the dot correlates to the number of proteins classified for that specific biological process. The color displays the p-value.

Investigating following hits:		
Acsl	*ArfGAP3*	*bel*
brp	*bsf*	*CG10077*
CG41099	*CG5839*	*CG6089*
chm	*Clc*	*Dlic*
ect4/sarm	*enc*	*ens*
lig	*lok*	*lost*
Lsp2	*mld*	*Mnt*
mus201	*nocte*	*Prat*
rin	*shep*	*shi*
Tango1	*TBCB*	*tyf*
Vap-33A		

Figure 2.16 Potential hits for future investigation. The involvement of the hits in kinesin transport of *oskar* mRNA will be investigated by TRiP-line based RNAi systems.

4. Discussion

4.1. The effect of SOLE mutations in *oskar* mRNA localization

The localization of *oskar* RNA in the *Drosophila melanogaster* oocyte offers an excellent system to investigate the different RNA transport mechanisms, including active transport and diffusion and entrapment mediated localization. The switch from dynein to kinesin during the *oskar* mRNA transport is intriguing and involves various *cis*- and *trans*-factors for each motor protein. While the mechanism of mRNA transport by dynein has been extensively studied *in vivo* and *in vitro* (Bullock and Ish-Horowicz 2001; Dienstbier et al. 2009; McClintock et al. 2018; Sladewski et al. 2018), the process of kinesin-mediated transport to the posterior pole of the oocyte is less clearly understood. Previous studies have shown that splicing at the first intron, creation of a SOLE stem-loop and EJC deposition are required for proper kinesin transport (Ghosh et al. 2012). Indeed, incorporating the SOLE, an intron, and the *oskar* 3'UTR is sufficient to enable the localization of *lacZ* RNA to the posterior pole, thus mimicking *oskar* mRNA localization. The requirement of the first *P-element* transposase intron for boosting transgenic RNA expression should be reinvestigated since the *lacZ* RNAs were expressed at high levels even in the absence of the *P-element* intron, as measured by RT-qPCR.

The localization of *lacZ* WT and *lacZ* SLc RNA to the posterior pole was not affected by the removal of the *P-element* intron and the accompanying EJC. Previous research showing that splicing and EJC deposition only at the SOLE site is required for *oskar* mRNA localization was conducted in presence of the *P-element* intron (Hachet and Ephrussi 2004; Ghosh et al. 2012). My findings complement these studies by demonstrating that the additional EJC that is presumably deposited upon splicing of the *P-element* intron had no effect on mRNA localization to the posterior pole. Thus, the data support the notion that *oskar* mRNA requires only one splicing event at the first intron position for its localization to the posterior pole of the oocyte.

The *lacZ* WT and *lacZ* SLc RNA, which contain the stem-loop SOLE, localize to the posterior pole during stage 9 of egg chamber development similar to *oskar* mRNA. The mode

of function of SOLE in the *lacZ* WT RNA is similar to in endogenous *oskar* mRNA, since mutations in the SOLE have the same effect on localization of both RNAs (Ghosh et al. 2012). However, unlike *oskar* mRNA, which is translated at the posterior pole and leads to Oskar protein mediated RNA anchoring (Vanzo and Ephrussi 2002), the *lacZ* RNAs in an *oskar* RNA-null background had no such mechanism to anchor the RNA at the posterior pole. In absence of successful anchoring at the posterior cortex, the *lacZ* WT RNA delocalizes after the kinesin transport, which results in the *lacZ* WT RNA being less tightly localized to the posterior pole compared to endogenous *oskar* mRNA. This effect is further enhanced in stage 10a and 10b egg chambers, where the *lacZ* WT RNA center of mass is shifted more towards the anterior. The *lacZ* SL and the *lacZ* IL RNA, which lacks the essential splice event and EJC, are unable to localize to the posterior pole and are distributed throughout the oocyte. The observation regarding the impact of mutations in the SOLE on *lacZ* RNA localization is consistent with previous studies on *oskar* transgenes (Hachet and Ephrussi 2004; Ghosh et al. 2012). However, analogous mutations in *oskar* mRNA still resulted in reduced localization at the posterior pole during late oogenesis, which was achieved due to ooplasmic streaming and local anchoring by Oskar produced by the endogenous *oskar* mRNA (Glotzer et al. 1997; Hachet and Ephrussi 2004; Eichler et al. 2022).

4.2. Stabilization of EJC binding in presence of a RNA stem-loop structure

The exact relationship between the EJC and SOLE has been unclear. Splicing of only the first *oskar* intron is required for posterior localization. At this site, the SOLE forms upon splicing and the EJC is deposited 20 to 24 nucleotides upstream of the exon-exon junction (Ghosh et al. 2012; Obrdlik et al. 2019). The first exon contributes 18 nucleotides to the 28 nucleotide SOLE stem-loop. Thus, the 144 kDa EJC is only two to six nucleotides away from the SOLE, making it very likely that the EJC and SOLE interact.

The RNA-protein footprinting experiment revealed the RNA regions bound by proteins and thus protected from RNase degradation. Two *lacZ* RNA-protein footprints were generated, one analyzing the direct binding of proteins to the RNA and the other at the RNP

level including peripheral proteins indirectly binding the *lacZ* RNA. The direct RNA-protein footprinting showed only the RNA regions protected by proteins that bind directly to the RNA at "zero-distance". In contrast, the protein complex-protected RNA-protein footprint at the RNP level included RNA regions protected due to the RNP structure and protein complexes bound to the RNA. Thus, protection was provided not only by proteins that interact directly with the RNA but also by protein complexes including indirectly bound proteins that block RNase access.

The protein complex-protected RNA footprint indicated that more proteins are bound to the *lacZ* coding region than the *oskar* 3'UTR, irrespective of the SOLE structure. The generated RNA-protein footprints of *lacZ* WT and *lacZ* SL RNA were similar. However, the EJC binding site of the *lacZ* WT RNA, which has an intact SOLE secondary structure, was found to be more protected than the EJC binding site of the *lacZ* SL RNA. Even the *lacZ* RNA footprint from direct RNA-protein interactions revealed that the EJC site is more protected in *lacZ* WT than *lacZ* SL RNA. This showed that in the presence of a stem-loop SOLE, there was a higher protein occupancy at the EJC binding site. Moreover, the *lacZ* RNA footprint from direct RNA-protein interactions showed that in a localizing RNA (*lacZ* WT) the coding region is more protected relative to the *oskar* 3'UTR. The RNA footprint of the *lacZ* SL RNA displayed a higher protein occupancy in the 3'UTR relative to its *lacZ* coding region. The mass spectrometry data demonstrated that known *oskar* 3'UTR binding proteins such as Bruno, Orb, Tral and Cup were slightly higher enriched in *lacZ* SL than *lacZ* WT RNA. This could explain the difference in RNA-protein footprints.

Earlier studies of EJC occupancy after splicing *oskar* mRNA *in vitro* showed that the EJC is bound at all three exon-exon junctions. *In vitro*, the SOLE mutations did not affect the EJC deposition, as was shown by RNase H protection (Ghosh et al. 2012). An *in vivo* approach using RNA immunoprecipitation sequencing (RIP-seq) on the other hand demonstrated that the EJC is selectively enriched at the first exon-exon junction site (Obrdlik et al. 2019). The majority of EJCs in *Drosophila melanogaster* are deposited on internal exons. Several factors including long introns, strong splice sites and CG-rich hexamers influence EJC stability on RNA. Bioinformatic analysis also suggested that the EJC binds most efficiently or stably to single-stranded regions in RNA (Obrdlik et al. 2019). To identify if the stem-loop SOLE

increases the EJC stability on the RNA, a RIP experiment on *lacZ* WT and *lacZ* SL flies was performed. The RIP experiment against GFP-Mago in *lacZ* WT and *lacZ* SL flies confirmed that it is the SOLE stem-loop structure that is primarily affecting the EJC stability at the first exon, as disrupting the stem-loop decreased EJC stability. The SOLE stem-loop structure is thus stabilizing the EJC on the *oskar* mRNA, which results in selective enrichment of the EJC at the first exon-exon junction. This is consistent with a transcriptome-wide EJC binding study, which analyzed the base-pairing probability close to the EJC binding sites and concluded that adjacent dsRNA structure may affect EJC binding in *Drosophila melanogaster* (Obrdlik et al. 2019). A similar tendency of EJC occupancy near RNA secondary structure was not found in mammals (Hauer et al. 2016). The analysis I performed is the first to show that an RNA stem-loop is able to increase EJC stability in *Drosophila melanogaster*.

Double-stranded RNA at the EJC binding position negatively impacts EJC deposition (Mishler et al. 2008). A transcriptome-wide RNA structural analysis in mouse embryonic stem cells demonstrated that stem-loops upstream of the EJC binding site are more common than downstream (Saha et al. 2020). These stem-loops close to the EJC binding site are thought to affect splicing efficiency. The precedence of SOLE stabilizing the EJC may include that the stem-loops not solely affect RNA splicing but also stabilize the EJC. Since the EJC has been shown to be required for downstream splicing (Hayashi et al. 2014), the EJC adjacent stem-loop may have a dual function by directly affecting splicing and indirectly by EJC stabilization.

The precise mechanism by which the SOLE stabilizes the EJC remains unclear. Although the SOLE is located in close proximity to the EJC and suggests direct interaction, it could potentially recruit *trans*-acting factors to indirectly stabilize the EJC. In mammalian cells, the EJC can dissociate during translation through the action of Pym (Bono et al. 2004; Dostie and Dreyfuss 2002; Lejeune et al. 2002; Gehring et al. 2009), while in *Drosophila*, Pym can dissociate the EJC through a translation-independent interaction with the Y14-Mago heterodimer (Ghosh et al. 2014). In my system, neither *lacZ* WT nor *lacZ* SL RNA were translated, indicating that EJC dissociation is also occurring translation-independent. The presence of SOLE could lead to direct EJC stabilization, indirect stabilization by recruiting RBPs or by blocking Pym association to the RNA and Y14-Mago. Since Y14-Mago can

inhibit the ATPase activity of eIF4A3, it is possible that the SOLE stabilizes the association of Y14-Mago with eIF4A3. It would be interesting to determine at which stage EJC dissociates from the *lacZ* SL RNA and to investigate its underlying mechanism and address whether Pym or another unknown mechanism is involved. Since the EJC is not deposited on all exon-exon junction sites (Saulière et al. 2010), the possibility exists that the EJC is not stably bound during splicing on *lacZ* SL RNA at all.

However, the possibility that in addition to the SOLE stabilizing the EJC, the EJC having a reciprocal stabilizing effect on the SOLE structure and thus function as a unit remains to be tested. Several stem-loop forming *cis*-acting RNA sequences required for protein binding and RNA localization have been identified previously. *bicoid* and *oskar* mRNA contain such sequences which form stem-loop structures, recruit RBPs and facilitate dynein transport (Ferrandon et al. 1997; Macdonald and Kerr 1997; Jambor et al. 2014). The SOLE could have the potential to recruit an RBP to the RNA that might have an effect on activating kinesin transport, either directly or indirectly through interaction with the EJC or its interactome. It is possible that disruption of the SOLE secondary structure may not only result in decreased stability of the EJC but also hinder the binding of important RBPs to the SOLE.

Structural analysis that is planned for the future should be able to answer if the presence of the EJC affects the SOLE stem-loop structure. Dimethyl sulfate (DMS) mutational profiling with sequencing (DMS-MaPseq) based *in vivo* structural analysis has been performed on *oskar* mRNA in a previous study (Zubradt et al. 2016). In DMS-MaPseq and similar methods, the RNA is treated with a chemical compound that is specifically modifying single-stranded RNAs (ssRNAs). Mutations are introduced in the cDNA sequence during reverse transcription of chemically modified ssRNA, which can be sequenced to analyze the RNA secondary structure (Smola et al. 2015; Smola & Weeks 2018). The method previously confirmed the *in vivo* configuration of the stem-loop OES within the *oskar* 3'UTR (Zubradt et al. 2016), suggesting its potential usefulness for *in vivo* RNA structure analysis in *Drosophila* ovaries. The secondary structure of the SOLE has also been demonstrated through NMR analysis *in vitro* (Simon et al. 2015). Investigating how the SOLE structure is affected by the mutations in the *lacZ* SL, the lack of splicing in *lacZ* IL, or the absence of EJC will complement the current understanding of the role of EJC and SOLE in *oskar* mRNA transport.

4.3. Increasing the efficiency of the transcript-specific RBP capture

Unlike protein-centric RBP capture methods, the transcript-specific RBP capture approach is constrained by the low abundance of RNA of interest relative to the total RNA, low RNA expression levels and inefficient RNA pulldown by antisense oligonucleotides (AOS; Gerber 2021). However, contrary to aptamer-based RNA pulldown, the use of antisense DNA probes does not introduce an additional RNA sequence that can form a secondary structure, which could potentially compromise the integrity of the global RNA structure. The recently developed BioID methods, while suitable for capturing local RBPs at a site of interest, might not efficiently capture larger RNAs due to the biotinylation radius being only 10 nm (Kim et al. 2014). This limitation highlights the potential of the antisense oligonucleotide-based approach, which targets unmodified endogenous RNA. However, the challenge of capturing the RBP composition of RNAs with low expression levels and inefficient pulldowns remains. As a result, the RBP composition of only a handful of RNAs, mostly highly expressed ones, has been studied so far (Blencowe et al. 1989; Lingner & Cech 1996; McHugh et al. 2015; Rogell et al. 2017; Wippich and Ephrussi 2020). Three independent ASO-based transcript-specific RBP captures of the lncRNA XIST displayed only a minor overlap of detected RBPs (Minajigi et al. 2015; Chu et al. 2015; McHugh et al. 2015), suggesting the need for optimized transcript-specific RBP capture. To address this, I developed a general optimization protocol for transcript-specific RBP capture, making the ASO approach more attractive.

The transcript-specific RBP capture was used before to analyze the RBP composition of *oskar* RNP (Wippich and Ephrussi 2020; Appendix Figure S1). Initially, the same DNA probes that were effective in smFISH were utilized to target RNA in the transcript-specific RBP capture method. In ASO-based RBP capture, probes were generated randomly or by analyzing RNA structure through mfold (Zuker 2003; Rogell et al. 2017). A previous study has employed a similar strategy of RNase degradation followed by sequencing (West et al. 2014). However, in their RNase H mapping approach for the lncRNAs MALAT1 and NEAT1, accessible single-stranded RNA fragments were generally identified instead of directly detecting hybridized DNA probes. My approach complements previous techniques for identifying efficient AOS, thus improving the efficiency of ASO-based transcript-specific

RBP capture. Since smFISH uses the same principle in ASO-based hybridization to a target RNA (Femino et al. 1998), probe optimization can also be used for optimizing smFISH DNA probes, when the signal-to-noise ratio is critical.

A critical point in RNA-centric RBP capture is to keep background to a minimum. While capturing more beads can lead to more RNA and RBPs being captured, it can also increase the amount of non-specific RNA and protein being captured. Thus, there is a tradeoff between increasing the capture of RNA of interest and a high background, or capturing less RNA and reducing the background, which needs to be chosen on a case-by-case basis for each target transcript. In my case, it was necessary to increase the number of beads and background to obtain enough RNA-bound proteins for mass spectrometry analysis. The relative amount of RNA capture after my optimization is similar to other RNA-centric RBP studies (Rogell et al. 2017; Minajigi et al. 2015; Wippich and Ephrussi 2020). Previous RBP capture studies considered a 2-fold protein enrichment over the negative control as significant (Minajigi et al. 2015; Rogell et al. 2017; Spiniello et al. 2018). The previously performed transcript-specific RBP capture (Figure 1.11) on *oskar* mRNA was less stringent, as proteins known to bind *oskar* mRNA were enriched only 20-30% over the negative control (Wippich and Ephrussi 2020; Appendix Figure S1). Staufen, a bona fide *oskar* 3'UTR binding protein was approximately 2.5-fold enriched, whereas in the *lacZ*-specific RBP capture, Staufen was 8-fold enriched over the negative control using scrambled DNA probes (Figure 2.13B). Overall, I developed a protocol that optimizes the previously reported transcript-specific RBP capture, which can be used in principle for studying RBPs associating with any RNA.

4.4. RBP composition analysis of *lacZ* RNAs

Imaging-based colocalization analysis of the *oskar* full-length RNA and the *oskar* 3'UTR performed by Imre Gaspar in our lab revealed that kinesin is recruited to the RNA independent of the coding sequence, splicing, EJC and SOLE. But the inability of *oskar* 3'UTR and kinesin to localize to the posterior pole suggests that kinesin is in an inactive state. Therefore, it was anticipated that the missing EJC-SOLE interactome involves yet unidentified factors required for kinesin activation to promote localization to the posterior

pole. To identify novel factors involved in *oskar* mRNA transport I aimed to analyze and compare the EJC-SOLE interactome of an RNA that localized to the posterior pole and an RNA with a mutant SOLE that did not localize. The RIP data demonstrate that disrupting the stem-loop structure affects EJC occupancy, and thus comparing the RBP composition of different RNAs localizing and not localizing should give insight into the EJC-SOLE interactome of a posterior pole localizing RNA.

Since the translation of the RNA displaces the bound EJC (Dostie and Dreyfuss 2002), the designed *lacZ* RNA contained a stop codon and therefore remained untranslated. The resulting untranslatable *lacZ* transgenes provide a good method for analyzing the kinesin transport-relevant RBP composition of *oskar* mRNA transport. To ensure that the analysis would be focused on the EJC flanking the SOLE, the *P-element* intron in the transgenesis vector (Rørth 1998) was specifically removed to generate a clean system. Additionally, to prevent the formation of an RNP complex containing both the *oskar* and the *lacZ* RNA as a result of "hitch-hiking" and presumably co-packaging into RNPs by interaction of the dimerization loop in the 3'UTRs of endogenous *oskar* with that in the *lacZ* transegne (Hachet and Ephrussi 2004; Jambor et al. 2011), the *lacZ* transgenes were expressed in flies lacking endogenous *oskar* RNA. Contrary to the previous *oskar* RNP analysis (Frank Wippich, Appendix Figure S1), Vasa was not enriched in my experiments (Figure 2.13). This is presumably due to the absence of endogenous *oskar* mRNA since the lack of Oskar protein has been shown to recruit Vasa to the posterior pole (Breitwieser et al. 1996). The generated transgenic *lacZ* fly lines facilitate analyzing the impact of SOLE structures and sequences on the RBP composition *in vivo*.

According to the GO analysis, the majority of proteins detected in the *lacZ* RNP are involved in RNA binding and processing. Mass spectrometry analysis of the *lacZ*-specific RBP capture successfully identified many known *oskar* 3'UTR binding proteins. In fact, the most enriched proteins in all transgenic *lacZ* RNA pulldowns were mostly known *oskar* 3'UTR binding proteins and/or components of the *oskar* RNP complex such as Staufen, Orb, Bruno, Cup, Tral, Wisp and Me31B (Breitwieser et al. 1996; Chang et al. 1999; Kim-Ha et al. 1995; Nakamura et al. 2001; Wilhelm et al. 2003; Besse et al. 2009; Cui et al. 2013).

The first RBP capture detected several proteins specifically enriched in each *lacZ* RNA sample. Proteins such as Mus201, Brp and Zasp67 were strongly enriched in the *lacZ* WT sample, whereas they were not found to associate with the *lacZ* SL RNA. Ect4, CG5839, Ef1alpha100E, Bsf and Shi were only moderately enriched in the *lacZ* WT sample. *mus201* is involved in UV DNA damage repair (Boyd et al. 1982), *brp* encodes a cytoskeletal protein required for structural integrity and vesicle fusion (Wagh et al. 2006) and Zasp67 is able to bind actin filaments and participates in muscle development and myofibril assembly (Katzemich et al. 2013). Ect4, CG5839, Ef1alpha100E, Bsf and Shi are not known to interact with *oskar* mRNA or be involved in active kinesin transport. Bsf has been reported to affect RNA stability and Shi (Dynamin) affects the production of microtubule bundles (Mancebo et al. 2001; Shpetner and Vallee 1991). No direct association can be made between highly enriched proteins in *lacZ* WT and RNA localization or *oskar* mRNA binding.

After multiple replicates the transcript-specific RBP capture demonstrated similar RNP composition of *lacZ* WT and *lacZ* SL RNA, even despite differences in RNA localization. Proteins that could be involved in kinesin transport of *oskar* mRNA such as Pat1 (Loiseau et al. 2010) or UNC-76 (Gindhart et al. 2003) were not majorly enriched in one RNA over the other in the mass spectrometry dataset. It is possible that factors involved in kinesin transport are only slightly more enriched in *lacZ* WT than *lacZ* SL RNA, therefore differences are not detectable via mass spectrometry. Surprisingly proteins involved in dynein transport such as BicD, Dhc64C, Short wing, DCTN2-p50 and Dlic were among the top enriched proteins detected in all conditions, whereas kinesin was not enriched over the negative control. As the fly ovary is a complex tissue containing egg chambers from 14 morphogenetically distinct stages of development, and egg chambers of all the different stages are collected during the ovary harvesting process (Jambor et al. 2016), the transcript-specific RBP capture enriches distinct *lacZ* RNPs at different stages. The RBP composition is likely different during each transport step, and the localization pattern suggests that the SOLE and EJC do not affect dynein transport (Ghosh et al. 2012). The RBP composition of *lacZ* WT and *lacZ* SL RNA is possibly similar during dynein transport, resulting in a significant overlap in the RNP captured during the RBP capture. This unavoidable heterogeneity in the input material might explain the increased similarity in RBP composition between *lacZ* WT and SL RNA after additional biological replicates.

This study sheds new light on the role of the EJC and SOLE in *oskar* mRNA transport. Although RIP showed that the EJC was more stable on *lacZ* WT RNA than on *lacZ* SL RNA, this was not confirmed by the mass spectrometry data since the EJC was similarly captured in all *lacZ* samples, including the negative controls. Incubating carboxy agarose beads in *Drosophila* ovary lysate revealed EJC binding to the beads (Appendix Figure S2), which could explain why the EJC was not enriched in the *lacZ* WT transcript-specific RBP capture. It is unclear how the EJC is interacting with the beads, but it is possible that the EJC core proteins and associated proteins somehow bind to the bead matrix. Therefore, factors involved in kinesin transport that associate with the EJC-SOLE interactome could potentially bind to the beads matrix and thus be present in the negative control and consequently not enriched in the transcript-specific RBP capture. It is important to take this into account when conducting future bead-based pull-downs.

Kinesin was shown to be associated with transgenic *oskar* 3'UTR, yet the RNA did not undergo transport by kinesin (Figure 2.7). Since the EJC and SOLE are required for the kinesin transport, a hypothesis was proposed that some proteins of the EJC-SOLE interactome activate kinesin transport. The RIP experiment showed that the EJC is more stable on the SOLE-containing RNA, suggesting that the EJC and its interactome are the factors that promote kinesin activity and RNA localization. In addition to its involvement in *oskar* mRNA localization, the EJC was recently shown to be required for *NIN* RNA transport in human retinal pigment epithelial-1 cells (Kwon et al. 2021). This may indicate that EJC-dependent RNA localization is not unique to *Drosophila*. Although *in vitro* experiments in our lab have failed to detect a direct interaction between kinesin and the EJC, recent studies in neuroblastoma cells suggest that KIF1C directly interacts with the EJC (Nagel et al. 2022). In a functionally inactive mutant of KIF1C, which has been known to cause hereditary spastic paraplegia (Oteyza et al. 2014), the EJC and RNA have been shown to be mislocalized. The interaction between KIF1C and EJC was found to be RNA-dependent. The EJC could be used as a marker to detect processed and untranslated RNA for kinesin transport. The stabilization of EJC by SOLE on the *oskar* mRNA could be sufficient to promote kinesin transport to the posterior pole. An immunoprecipitation experiment of eIF4A3 from a previous study showed no direct protein-protein interaction between kinesin and eIF4A3 in *Drosophila* ovaries (Bansal et al. 2020). However, it is yet to be determined whether kinesin can interact with the

EJC in an RNA-dependent manner in *Drosophila*. Single-molecule reconstitution assays (McClintock et al. 2018; Heber et al. 2022) may be useful in determining whether the EJC directly activates kinesin, or whether secondary factors interacting with the EJC-SOLE are required for transport.

4.5. Transcript-specific RBP capture for future applications

The data generated by the transcript-specific RBP capture experiments which I performed underline the effectiveness of the method for general RBP analysis. Known *oskar* 3'UTR binding proteins were significantly enriched in each condition. The transcript-specific RBP capture approach shows that careful optimization before choosing a target RNA is beneficial and aids to reduce background RNA and protein. I achieved similar RNA capture efficiency as previous studies using different RNA-centric methods (Rogell et al. 2017; Baltz et al. 2012; Castello et al. 2012; Wippich et al. 2020). The abundance of the target RNA and the amount of required *Drosophila* ovaries are the biggest bottlenecks for transcript-specific RBP capture in *Drosophila*. An additional limitation is the length of the target RNA, as shorter RNAs afford fewer sites for probe binding, which thus reduces the efficiency of the capture.

The data obtained from mass spectrometry indicate that the identification of RNA binding proteins for a specific RNA is robust. Multiple known *oskar* 3'UTR binding proteins were consistently detected across all replicates. The fold change relative to the negative control was significantly higher than in previous RBP capture attempts (Minajigi et al. 2015; Rogell et al. 2017; Spiniello et al. 2018). The mass spectrometry method may not be sensitive enough to detect slight differences or short-lived changes in RBP composition. In the case of my *lacZ* RNAs, there are significant differences in RNA localization that suggest a potentially significant difference in the RBP composition. However, the mass spectrometry data only showed slight differences in RBP composition. Additionally, the EJC, which has been shown to recruit various other proteins to the RNA (Hayashi et al. 2014), was not detected, despite the RIP experiment indicating that the EJC is more stable in the presence of a wild-type SOLE. Proteins captured due to interaction with the bead matrix will not be significantly enriched and therefore fall through the identification of RBPs. The transcript-specific RBP capture

method can analyze and identify specific RBPs, but it is important to keep in mind that not all RBPs may be detected due to technical limitations. In addition, the detected RBPs need to be validated by other methods to exclude false positives (Vaishali et al. 2021). Nevertheless, this approach allows for the identification of many novel RBPs for a given RNA.

5. Conclusion & Outlook

The localization of RNA has been extensively researched during *Drosophila* embryogenesis because asymmetric localization is crucial for proper development. In the developing oocyte, *oskar* mRNA utilizes both dynein and kinesin-based RNA transport mechanisms to reach the posterior pole, making it interesting to examine how a single RNA exploits both of these highly conserved motors for subcellular localization. The EJC is deposited on the RNA in the nucleus, and once the *oskar* RNP reaches the oocyte through dynein-mediated transport, the *oskar* RNP activates kinesin transport. In addition to the EJC, proper *oskar* RNA localization requires the SOLE. A transgenic RNA requires only the SOLE, *oskar* 3'UTR and splicing/EJC to localize to the posterior pole.

It has been more than a decade since the SOLE was identified. My data support the hypothesis that the SOLE stabilizes the EJC on the RNA. However, it is unclear whether the SOLE has additional roles beyond stabilizing the EJC. It is possible that the SOLE and EJC function as a unit, with the SOLE stabilizing the EJC and the EJC stabilizing the SOLE stem-loop structure. Further structural analysis will facilitate the investigation of how the EJC interacts with the SOLE and the impact of the EJC's absence on the secondary structure of the SOLE. This analysis can enhance our understanding of the roles of the EJC and SOLE in transporting *oskar* mRNA.

The technique of transcript-specific RBP capture was utilized to identify new factors associated with kinesin transport in an unbiased manner. Ideally, I would have preferred to select proteins for further investigation based on their significance and enrichment on the localizing relative to the non-localizing RNA. However, many proteins that were captured appeared to be similarly enriched, resulting in no clear candidates to select for testing *in vivo*. Nevertheless, to test if a knockdown of some of the detected proteins affects *oskar* RNA localization, *Drosophila* RNAi lines from public stocks were ordered. Should this be the case, additional approaches to determine the interaction between the proteins and *oskar* RNA, kinesin and EJC will be investigated. Single-molecule *in vitro* reconstitution assays could be instrumental in verifying the effect of putative hits. Ultimately, it would be interesting to investigate the mechanism behind kinesin transport activation of *oskar* mRNA in the oocyte.

My findings offer a greater understanding of the role of the SOLE and EJC in kinesin transport and provide data and an approach that can be further refined and used to fully elucidate the mechanism and the proteins involved in kinesin transport. After I optimized the RBP capture technique, further investigation of other EJC and SOLE mutants should be easier and will complement the data and help unravel the localization mechanism of *oskar* mRNA to the posterior pole.

Bibliography

Adam, S. A.; Nakagawa, T.; Swanson, M. S.; Woodruff, T. K.; Dreyfuss, G., 1986: mRNA polyadenylate-binding protein: gene isolation and sequencing and identification of a ribonucleoprotein consensus sequence. *Molecular and Cellular Biology*., **6**, 2932–2943.

Afgan, E.; Baker, D.; Batut, B.; van den Beek, M.; Bouvier, D.; Cech, M.; Chilton, J.; Clements, D.; Coraor, N.; Grüning, B. A.; Guerler, A.; Hillman-Jackson, J.; Hiltemann, S.; Jalili, V.; Rasche, H.; Soranzo, N.; Goecks, J.; Taylor, J.; Nekrutenko, A. et al., 2018: The Galaxy platform for accessible, reproducible and collaborative biomedical analyses: 2018 update. *Nucleic Acids Research*., **46**, W537–W544.

Andersen, C. B. F.; Ballut, L.; Johansen, J. S.; Chamieh, H.; Nielsen, K. H.; Oliveira, C. L. P.; Pedersen, J. S.; Séraphin, B.; Le Hir, H.; Andersen, G. R., 2006: Structure of the exon junction core complex with a trapped DEAD-box ATPase bound to RNA. *Science*., **313**, 1968–1972.

Ballut, L.; Marchadier, B.; Baguet, A.; Tomasetto, C.; Séraphin, B.; Le Hir, H., 2005: The exon junction core complex is locked onto RNA by inhibition of eIF4AIII ATPase activity. *Nature Structural & Molecular Biology*., **12**, 861–869.

Baltz, A. G.; Munschauer, M.; Schwanhäusser, B.; Vasile, A.; Murakawa, Y.; Schueler, M.; Youngs, N.; Penfold-Brown, D.; Drew, K.; Milek, M.; Wyler, E.; Bonneau, R.; Selbach, M.; Dieterich, C.; Landthaler, M., 2012: The mRNA-bound proteome and its global occupancy profile on protein-coding transcripts. *Molecular Cell*., **46**, 674–690.

Bansal, P.; Madlung, J.; Schaaf, K.; Macek, B.; Bono, F., 2020: An Interaction Network of RNA-Binding Proteins Involved in Drosophila Oogenesis. *Molecular & Cellular Proteomics*., **19**, 1485–1502.

Barbosa, I.; Haque, N.; Fiorini, F.; Barrandon, C.; Tomasetto, C.; Blanchette, M.; Le Hir, H., 2012: Human CWC22 escorts the helicase eIF4AIII to spliceosomes and promotes exon junction complex assembly. *Nature Structural & Molecular Biology*., **19**, 983–990.

Bassell, G. J.; Warren, S. T., 2008: Fragile X syndrome: loss of local mRNA regulation alters synaptic development and function. *Neuron*., **60**, 201–214.

Bastock, R.; St Johnston, D., 2008: Drosophila oogenesis. *Current Biology*., **18**, R1082-7.

Bateman, J. R.; Lee, A. M.; Wu, C. ting, 2006: Site-specific transformation of Drosophila via phiC31 integrase-mediated cassette exchange. *Genetics*., **173**, 769–777.

Becalska, A. N.; Gavis, E. R., 2009: Lighting up mRNA localization in Drosophila oogenesis. *Development*., **136**, 2493–2503.

Beckmann, B. M.; Horos, R.; Fischer, B.; Castello, A.; Eichelbaum, K.; Alleaume, A. M.; Schwarzl, T.; Curk, T.; Foehr, S.; Huber, W.; Krijgsveld, J.; Hentze, M. W., 2015: The RNA-binding proteomes from yeast to man harbour conserved enigmRBPs. *Nature Communications*., **6**, 10127.

Benhalevy, D.; Anastasakis, D. G.; Hafner, M., 2018: Proximity-CLIP provides a snapshot of protein-occupied RNA elements in subcellular compartments. *Nature Methods*., **15**, 1074–1082.

Bergsten, S. E.; Gavis, E. R., 1999: Role for mRNA localization in translational activation but not spatial restriction of nanos RNA. *Development*., **126**, 659–669.

Berleth, T.; Burri, M.; Thoma, G.; Bopp, D.; Richstein, S.; Frigerio, G.; Noll, M.; Nüsslein-Volhard, C., 1988: The role of localization of bicoid RNA in organizing the anterior pattern of the Drosophila embryo. *The EMBO Journal*., **7**, 1749–1756.

Besse, F.; López de Quinto, S.; Marchand, V.; Trucco, A.; Ephrussi, A., 2009: Drosophila PTB promotes formation of high-order RNP particles and represses oskar translation. *Genes & Development*., **23**, 195–207.

Blencowe, B. J.; Sproat, B. S.; Ryder, U.; Barabino, S.; Lamond, A. I., 1989: Antisense probing of the human U4/U6 snRNP with biotinylated 2'-OMe RNA oligonucleotides. *Cell*., **59**, 531–539.

Blencowe, B. J.; Issner, R.; Nickerson, J. A.; Sharp, P. A., 1998: A coactivator of pre-mRNA splicing. *Genes & Development*., **12**, 996–1009.

Bloom, K.; Beach, D. L., 1999: mRNA localization: motile RNA, asymmetric anchors. *Current Opinion in Microbiology*., **2**, 604–609.

Boehm, V.; Gehring, N. H., 2016: Exon junction complexes: supervising the gene expression assembly line. *Trends in Genetics*., **32**, 724–735.

Bono, F.; Ebert, J.; Unterholzner, L.; Güttler, T.; Izaurralde, E.; Conti, E., 2004: Molecular insights into the interaction of PYM with the Mago-Y14 core of the exon junction complex. *EMBO Reports*., **5**, 304–310.

Bookwalter, C. S.; Lord, M.; Trybus, K. M., 2009: Essential features of the class V myosin from budding yeast for ASH1 mRNA transport. *Molecular Biology of the Cell*., **20**, 3414–3421.

Boyd, J. B.; Snyder, R. D.; Harris, P. V.; Presley, J. M.; Boyd, S. F.; Smith, P. D., 1982: Identification of a second locus in Drosophila melanogaster required for excision repair. *Genetics*., **100**, 239–257.

Breitwieser, W.; Markussen, F. H.; Horstmann, H.; Ephrussi, A., 1996: Oskar protein interaction with Vasa represents an essential step in polar granule assembly. *Genes & Development*., **10**, 2179–2188.

Brendza, R. P.; Serbus, L. R.; Duffy, J. B.; Saxton, W. M., 2000: A function for kinesin I in the posterior transport of oskar mRNA and Staufen protein. *Science*., **289**, 2120–2122.

Bullock, S. L.; Ish-Horowicz, D., 2001: Conserved signals and machinery for RNA transport in Drosophila oogenesis and embryogenesis. *Nature*., **414**, 611–616.

Bullock, S. L.; Ringel, I.; Ish-Horowicz, D.; Lukavsky, P. J., 2010: A'-form RNA helices are required for cytoplasmic mRNA transport in Drosophila. *Nature Structural & Molecular Biology*., **17**, 703–709.

Caballero Oteyza, A.; Battaloğlu, E.; Ocek, L.; Lindig, T.; Reichbauer, J.; Rebelo, A. P.; Gonzalez, M. A.; Zorlu, Y.; Ozes, B.; Timmann, D.; Bender, B.; Woehlke, G.; Züchner, S.; Schöls, L.; Schüle, R., 2014: Motor protein mutations cause a new form of hereditary spastic paraplegia. *Neurology*., **82**, 2007–2016.

Cassella, L.; Ephrussi, A., 2022: Subcellular spatial transcriptomics identifies three mechanistically different classes of localizing RNAs. *Nature Communications*., **13**, 6355.

Castello, A.; Fischer, B.; Eichelbaum, K.; Horos, R.; Beckmann, B. M.; Strein, C.; Davey, N. E.; Humphreys, D. T.; Preiss, T.; Steinmetz, L. M.; Krijgsveld, J.; Hentze, M. W., 2012: Insights into RNA biology from an atlas of mammalian mRNA-binding proteins. *Cell*., **149**, 1393–1406.

Castello, A.; Fischer, B.; Frese, C. K.; Horos, R.; Alleaume, A. M.; Foehr, S.; Curk, T.; Krijgsveld, J.; Hentze, M. W., 2016: Comprehensive Identification of RNA-Binding Domains in Human Cells. *Molecular Cell*., **63**, 696–710.

Chang, J. S.; Tan, L.; Schedl, P., 1999: The Drosophila CPEB homolog, orb, is required for oskar protein expression in oocytes. *Developmental Biology*., **215**, 91–106.

Chang, P.; Torres, J.; Lewis, R. A.; Mowry, K. L.; Houliston, E.; King, M. L., 2004: Localization of RNAs to the mitochondrial cloud in Xenopus oocytes through entrapment and association with endoplasmic reticulum. *Molecular Biology of the Cell*., **15**, 4669–4681.

Chu, C.; Zhang, Q. C.; da Rocha, S. T.; Flynn, R. A.; Bharadwaj, M.; Calabrese, J. M.; Magnuson, T.; Heard, E.; Chang, H. Y., 2015: Systematic discovery of Xist RNA binding proteins. *Cell*., **161**, 404–416.

Clark, A.; Meignin, C.; Davis, I., 2007: A Dynein-dependent shortcut rapidly delivers axis determination transcripts into the Drosophila oocyte. *Development*., **134**, 1955–1965.

Clark, I.; Giniger, E.; Ruohola-Baker, H.; Jan, L. Y.; Jan, Y. N., 1994: Transient posterior localization of a kinesin fusion protein reflects anteroposterior polarity of the Drosophila oocyte. *Current Biology*., **4**, 289–300.

Clark, I. E.; Jan, L. Y.; Jan, Y. N., 1997: Reciprocal localization of Nod and kinesin fusion proteins indicates microtubule polarity in the Drosophila oocyte, epithelium, neuron and muscle. *Development*., **124**, 461–470.

Cohen, R. S.; Zhang, S.; Dollar, G. L., 2005: The positional, structural, and sequence requirements of the Drosophila TLS RNA localization element. *RNA (New York)*., **11**, 1017–1029.

Cole, K.; Truong, V.; Barone, D.; McGall, G., 2004: Direct labeling of RNA with multiple biotins allows sensitive expression profiling of acute leukemia class predictor genes. *Nucleic Acids Research*., **32**, e86.

Cui, J.; Sartain, C. V.; Pleiss, J. A.; Wolfner, M. F., 2013: Cytoplasmic polyadenylation is a major mRNA regulator during oogenesis and egg activation in Drosophila. *Developmental Biology*., **383**, 121–131.

Das, S.; Vera, M.; Gandin, V.; Singer, R. H.; Tutucci, E., 2021: Intracellular mRNA transport and localized translation. *Nature Reviews. Molecular Cell Biology*., **22**, 483–504.

Degot, S.; Le Hir, H.; Alpy, F.; Kedinger, V.; Stoll, I.; Wendling, C.; Seraphin, B.; Rio, M. C.; Tomasetto, C., 2004: Association of the breast cancer protein MLN51 with the exon junction complex via its speckle localizer and RNA binding module. *The Journal of Biological Chemistry*., **279**, 33702–33715.

Dienstbier, M.; Boehl, F.; Li, X.; Bullock, S. L., 2009: Egalitarian is a selective RNA-binding protein linking mRNA localization signals to the dynein motor. *Genes & Development*., **23**, 1546–1558.

Ding, D.; Parkhurst, S. M.; Halsell, S. R.; Lipshitz, H. D., 1993: Dynamic Hsp83 RNA localization during Drosophila oogenesis and embryogenesis. *Molecular and Cellular Biology*., **13**, 3773–3781.

Dix, C. I.; Soundararajan, H. C.; Dzhindzhev, N. S.; Begum, F.; Suter, B.; Ohkura, H.; Stephens, E.; Bullock, S. L., 2013: Lissencephaly-1 promotes the recruitment of dynein and dynactin to transported mRNAs. *The Journal of Cell Biology*., **202**, 479–494.

Dostie, J.; Dreyfuss, G., 2002: Translation is required to remove Y14 from mRNAs in the cytoplasm. *Current Biology*., **12**, 1060–1067.

Doyle, M.; Kiebler, M. A., 2011: Mechanisms of dendritic mRNA transport and its role in synaptic tagging. *The EMBO Journal*., **30**, 3540–3552.

Dreyfuss, G.; Choi, Y. D.; Adam, S. A., 1984: Characterization of heterogeneous nuclear RNA-protein complexes in vivo with monoclonal antibodies. *Molecular and Cellular Biology*., **4**, 1104–1114.

Eichler, C. E.; Grunberg, M. E.; Gavis, E. R., 2022: Overlapping and distinct cis-acting requirements for *oskar* mRNA localization pathways. *BioRxiv*.

Engel, K. L.; Arora, A.; Goering, R.; Lo, H. Y. G.; Taliaferro, J. M., 2020: Mechanisms and consequences of subcellular RNA localization across diverse cell types. *Traffic*., **21**, 404–418.

Ephrussi, A.; Dickinson, L. K.; Lehmann, R., 1991: Oskar organizes the germ plasm and directs localization of the posterior determinant nanos. *Cell*., **66**, 37–50.

Fallini, C.; Zhang, H.; Su, Y.; Silani, V.; Singer, R. H.; Rossoll, W.; Bassell, G. J., 2011: The survival of motor neuron (SMN) protein interacts with the mRNA-binding protein HuD and regulates localization of poly(A) mRNA in primary motor neuron axons. *The Journal of Neuroscience*., **31**, 3914–3925.

Fazal, F. M.; Han, S.; Parker, K. R.; Kaewsapsak, P.; Xu, J.; Boettiger, A. N.; Chang, H. Y.; Ting, A. Y., 2019: Atlas of Subcellular RNA Localization Revealed by APEX-Seq. *Cell*., **178**, 473-490.e26.

Femino, A. M.; Fay, F. S.; Fogarty, K.; Singer, R. H., 1998: Visualization of single RNA transcripts in situ. *Science*., **280**, 585–590.

Ferrandon, D.; Koch, I.; Westhof, E.; Nüsslein-Volhard, C., 1997: RNA-RNA interaction is required for the formation of specific bicoid mRNA 3' UTR-STAUFEN ribonucleoprotein particles. *The EMBO Journal*., **16**, 1751–1758.

Forrest, K. M.; Gavis, E. R., 2003: Live imaging of endogenous RNA reveals a diffusion and entrapment mechanism for nanos mRNA localization in Drosophila. *Current Biology*., **13**, 1159–1168.

Frohnhöfer, H. G.; Nüsslein-Volhard, C., 1986: Organization of anterior pattern in the Drosophila embryo by the maternal gene bicoid. *Nature*., **324**, 120–125.

Gadir, N.; Haim-Vilmovsky, L.; Kraut-Cohen, J.; Gerst, J. E., 2011: Localization of mRNAs coding for mitochondrial proteins in the yeast Saccharomyces cerevisiae. *RNA (New York)*., **17**, 1551–1565.

Gao, Y.; Tatavarty, V.; Korza, G.; Levin, M. K.; Carson, J. H., 2008: Multiplexed dendritic targeting of alpha calcium calmodulin-dependent protein kinase II, neurogranin, and activity-regulated cytoskeleton-associated protein RNAs by the A2 pathway. *Molecular Biology of the Cell*., **19**, 2311–2327.

Garner, C. C.; Tucker, R. P.; Matus, A., 1988: Selective localization of messenger RNA for cytoskeletal protein MAP2 in dendrites. *Nature*., **336**, 674–677.

Gaspar, I.; Yu, Y. V.; Cotton, S. L.; Kim, D. H.; Ephrussi, A.; Welte, M. A., 2014: Klar ensures thermal robustness of oskar localization by restraining RNP motility. *The Journal of Cell Biology*., **206**, 199–215.

Gáspár, I.; Sysoev, V.; Komissarov, A.; Ephrussi, A., 2017: An RNA-binding atypical tropomyosin recruits kinesin-1 dynamically to oskar mRNPs. *The EMBO Journal*., **36**, 319–333.

Gaspar, I.; Wippich, F.; Ephrussi, A., 2018: Terminal Deoxynucleotidyl Transferase Mediated Production of Labeled Probes for Single-molecule FISH or RNA Capture. *Bio-protocol*., **8**, e2750.

Gáspár, I.; Phea, L. J.; McClintock, M. A.; Heber, S.; Bullock, S. L.; Ephrussi, A., 2021: An RNA-based feed-forward mechanism ensures motor switching in oskar mRNA transport. *BioRxiv*.

Gavis, E. R.; Lehmann, R., 1992: Localization of nanos RNA controls embryonic polarity. *Cell*., **71**, 301–313.

Gehring, N. H.; Lamprinaki, S.; Kulozik, A. E.; Hentze, M. W., 2009a: Disassembly of exon junction complexes by PYM. *Cell*., **137**, 536–548.

Gehring, N. H.; Lamprinaki, S.; Hentze, M. W.; Kulozik, A. E., 2009b: The hierarchy of exon-junction complex assembly by the spliceosome explains key features of mammalian nonsense-mediated mRNA decay. *PLoS Biology*., **7**, e1000120.

Gerber, A. P., 2021: RNA-Centric Approaches to Profile the RNA-Protein Interaction Landscape on Selected RNAs. *Non-coding RNA*., **7**.

Gervasi, N. M.; Scott, S. S.; Aschrafi, A.; Gale, J.; Vohra, S. N.; MacGibeny, M. A.; Kar, A. N.; Gioio, A. E.; Kaplan, B. B., 2016: The local expression and trafficking of tyrosine hydroxylase mRNA in the axons of sympathetic neurons. *RNA (New York)*., **22**, 883–895.

Ghosh, S.; Marchand, V.; Gáspár, I.; Ephrussi, A., 2012: Control of RNP motility and localization by a splicing-dependent structure in oskar mRNA. *Nature Structural & Molecular Biology*., **19**, 441–449.

Ghosh, S.; Obrdlik, A.; Marchand, V.; Ephrussi, A., 2014: The EJC binding and dissociating activity of PYM is regulated in Drosophila. *PLoS Genetics*., **10**, e1004455.

Gindhart, J. G.; Chen, J.; Faulkner, M.; Gandhi, R.; Doerner, K.; Wisniewski, T.; Nandlestadt, A., 2003: The kinesin-associated protein UNC-76 is required for axonal transport in the Drosophila nervous system. *Molecular Biology of the Cell*., **14**, 3356–3365.

Giniger, E.; Varnum, S. M.; Ptashne, M., 1985: Specific DNA binding of GAL4, a positive regulatory protein of yeast. *Cell*., **40**, 767–774.

Glotzer, J. B.; Saffrich, R.; Glotzer, M.; Ephrussi, A., 1997: Cytoplasmic flows localize injected oskar RNA in Drosophila oocytes. *Current Biology*., **7**, 326–337.

Goldman, C. H.; Neiswender, H.; Veeranan-Karmegam, R.; Gonsalvez, G. B., 2019: The Egalitarian binding partners Dynein light chain and Bicaudal-D act sequentially to link mRNA to the Dynein motor. *Development*., **146**.

Gonsalvez, G. B.; Long, R. M., 2012a: Spatial regulation of translation through RNA localization. *F1000 Biology Reports*., **4**, 16.

Gonsalvez, G. B.; Long, R. M., 2012b: Faculty Opinions. *F1000Prime Rep*.

Gonzalez, I.; Buonomo, S. B.; Nasmyth, K.; von Ahsen, U., 1999: ASH1 mRNA localization in yeast involves multiple secondary structural elements and Ash1 protein translation. *Current Biology*., **9**, 337–340.

González-Reyes, A.; Elliott, H.; St Johnston, D., 1995: Polarization of both major body axes in Drosophila by gurken-torpedo signalling. *Nature*., **375**, 654–658.

González-Reyes, A.; St Johnston, D., 1998: Patterning of the follicle cell epithelium along the anterior-posterior axis during Drosophila oogenesis. *Development*., **125**, 2837–2846.

Grieder, N. C.; de Cuevas, M.; Spradling, A. C., 2000: The fusome organizes the microtubule network during oocyte differentiation in Drosophila. *Development*., **127**, 4253–4264.

Groisman, I.; Huang, Y. S.; Mendez, R.; Cao, Q.; Theurkauf, W.; Richter, J. D., 2000: CPEB, maskin, and cyclin B1 mRNA at the mitotic apparatus: implications for local translational control of cell division. *Cell*., **103**, 435–447.

Hachet, O.; Ephrussi, A., 2001: Drosophila Y14 shuttles to the posterior of the oocyte and is required for oskar mRNA transport. *Current Biology*., **11**, 1666–1674.

Hachet, O.; Ephrussi, A., 2004: Splicing of oskar RNA in the nucleus is coupled to its cytoplasmic localization. *Nature*., **428**, 959–963.

Hafner, M.; Landthaler, M.; Burger, L.; Khorshid, M.; Hausser, J.; Berninger, P.; Rothballer, A.; Ascano, M.; Jungkamp, A. C.; Munschauer, M.; Ulrich, A.; Wardle, G. S.; Dewell, S.; Zavolan, M.; Tuschl, T., 2010: Transcriptome-wide identification of RNA-binding protein and microRNA target sites by PAR-CLIP. *Cell*., **141**, 129–141.

Hauer, C.; Sieber, J.; Schwarzl, T.; Hollerer, I.; Curk, T.; Alleaume, A. M.; Hentze, M. W.; Kulozik, A. E., 2016: Exon Junction Complexes Show a Distributional Bias toward Alternatively Spliced mRNAs and against mRNAs Coding for Ribosomal Proteins. *Cell reports*., **16**, 1588–1603.

Hayashi, R.; Handler, D.; Ish-Horowicz, D.; Brennecke, J., 2014: The exon junction complex is required for definition and excision of neighboring introns in Drosophila. *Genes & Development*., **28**, 1772–1785.

Heber, S.; McClintock, M. A.; Simon, B.; Hennig, J.; Bullock, S. L.; Ephrussi, A., 2022: Tropomyosin 1-I/C co-ordinates kinesin-1 and dynein motors during *oskar* mRNA transport. *BioRxiv*.

Hockensmith, J. W.; Kubasek, W. L.; Vorachek, W. R.; von Hippel, P. H., 1986: Laser cross-linking of nucleic acids to proteins. Methodology and first applications to the phage T4 DNA replication system. *The Journal of Biological Chemistry*., **261**, 3512–3518.

Holt, C. E.; Martin, K. C.; Schuman, E. M., 2019: Local translation in neurons: visualization and function. *Nature Structural & Molecular Biology*., **26**, 557–566.

Howley, C.; Ho, R. K., 2000: mRNA localization patterns in zebrafish oocytes. *Mechanisms of Development*., **92**, 305–309.

Hughes, C. S.; Foehr, S.; Garfield, D. A.; Furlong, E. E.; Steinmetz, L. M.; Krijgsveld, J., 2014: Ultrasensitive proteome analysis using paramagnetic bead technology. *Molecular Systems Biology*., **10**, 757.

Hung, V.; Udeshi, N. D.; Lam, S. S.; Loh, K. H.; Cox, K. J.; Pedram, K.; Carr, S. A.; Ting, A. Y., 2016: Spatially resolved proteomic mapping in living cells with the engineered peroxidase APEX2. *Nature Protocols*., **11**, 456–475.

Huppertz, I.; Attig, J.; D'Ambrogio, A.; Easton, L. E.; Sibley, C. R.; Sugimoto, Y.; Tajnik, M.; König, J.; Ule, J., 2014: iCLIP: protein-RNA interactions at nucleotide resolution. *Methods*., **65**, 274–287.

Iwakawa, H. O.; Tomari, Y., 2022: Life of RISC: Formation, action, and degradation of RNA-induced silencing complex. *Molecular Cell*., **82**, 30–43.

Jambor, H.; Brunel, C.; Ephrussi, A., 2011: Dimerization of oskar 3' UTRs promotes hitchhiking for RNA localization in the Drosophila oocyte. *RNA (New York)*., **17**, 2049–2057.

Jambor, H.; Mueller, S.; Bullock, S. L.; Ephrussi, A., 2014: A stem-loop structure directs oskar mRNA to microtubule minus ends. *RNA (New York)*., **20**, 429–439.

Jambor, H.; Mejstrik, P.; Tomancak, P., 2016: Rapid Ovary Mass-Isolation (ROMi) to Obtain Large Quantities of Drosophila Egg Chambers for Fluorescent In Situ Hybridization. *Methods in Molecular Biology*., **1478**, 253–262.

Jan, C. H.; Williams, C. C.; Weissman, J. S., 2014: Principles of ER cotranslational translocation revealed by proximity-specific ribosome profiling. *Science*., **346**, 1257521.

Jansen, R. P.; Niessing, D., 2012: Assembly of mRNA-protein complexes for directional mRNA transport in eukaryotes--an overview. *Current Protein & Peptide Science*., **13**, 284–293.

Jenny, A.; Hachet, O.; Závorszky, P.; Cyrklaff, A.; Weston, M. D. J.; Johnston, D. S.; Erdélyi, M.; Ephrussi, A., 2006: A translation-independent role of oskar RNA in early Drosophila oogenesis. *Development*., **133**, 2827–2833.

Johnstone, O.; Lasko, P., 2001: Translational regulation and RNA localization in Drosophila oocytes and embryos. *Annual Review of Genetics*., **35**, 365–406.

Kanai, Y.; Dohmae, N.; Hirokawa, N., 2004: Kinesin transports RNA: isolation and characterization of an RNA-transporting granule. *Neuron*., **43**, 513–525.

Kanke, M.; Jambor, H.; Reich, J.; Marches, B.; Gstir, R.; Ryu, Y. H.; Ephrussi, A.; Macdonald, P. M., 2015: oskar RNA plays multiple noncoding roles to support oogenesis and maintain integrity of the germline/soma distinction. *RNA (New York)*., **21**, 1096–1109.

Kataoka, N.; Yong, J.; Kim, V. N.; Velazquez, F.; Perkinson, R. A.; Wang, F.; Dreyfuss, G., 2000: Pre-mRNA splicing imprints mRNA in the nucleus with a novel RNA-binding protein that persists in the cytoplasm. *Molecular Cell*., **6**, 673–682.

Kataoka, N.; Diem, M. D.; Kim, V. N.; Yong, J.; Dreyfuss, G., 2001: Magoh, a human homolog of Drosophila mago nashi protein, is a component of the splicing-dependent exon-exon junction complex. *The EMBO Journal*., **20**, 6424–6433.

Katz, Z. B.; Wells, A. L.; Park, H. Y.; Wu, B.; Shenoy, S. M.; Singer, R. H., 2012: β-Actin mRNA compartmentalization enhances focal adhesion stability and directs cell migration. *Genes & Development*., **26**, 1885–1890.

Katzemich, A.; Liao, K. A.; Czerniecki, S.; Schöck, F., 2013: Alp/Enigma family proteins cooperate in Z-disc formation and myofibril assembly. *PLoS Genetics*., **9**, e1003342.

Kaufman, P. D.; Rio, D. C., 1991: Drosophila P-element transposase is a transcriptional repressor in vitro. *Proceedings of the National Academy of Sciences of the United States of America*., **88**, 2613–2617.

Kim, D. I.; Birendra, K. C.; Zhu, W.; Motamedchaboki, K.; Doye, V.; Roux, K. J., 2014: Probing nuclear pore complex architecture with proximity-dependent biotinylation. *Proceedings of the National Academy of Sciences of the United States of America*., **111**, E2453-61.

Kim-Ha, J.; Smith, J. L.; Macdonald, P. M., 1991: oskar mRNA is localized to the posterior pole of the Drosophila oocyte. *Cell*., **66**, 23–35.

Kim-Ha, J.; Webster, P. J.; Smith, J. L.; Macdonald, P. M., 1993: Multiple RNA regulatory elements mediate distinct steps in localization of oskar mRNA. *Development*., **119**, 169–178.

Kim-Ha, J.; Kerr, K.; Macdonald, P. M., 1995: Translational regulation of oskar mRNA by bruno, an ovarian RNA-binding protein, is essential. *Cell*., **81**, 403–412.

King, M. L.; Messitt, T. J.; Mowry, K. L., 2005: Putting RNAs in the right place at the right time: RNA localization in the frog oocyte. *Biology of the Cell*., **97**, 19–33.

König, J.; Zarnack, K.; Rot, G.; Curk, T.; Kayikci, M.; Zupan, B.; Turner, D. J.; Luscombe, N. M.; Ule, J., 2010: iCLIP reveals the function of hnRNP particles in splicing at individual nucleotide resolution. *Nature Structural & Molecular Biology*., **17**, 909–915.

Kwon, O. S.; Mishra, R.; Safieddine, A.; Coleno, E.; Alasseur, Q.; Faucourt, M.; Barbosa, I.; Bertrand, E.; Spassky, N.; Le Hir, H., 2021: Exon junction complex dependent mRNA localization is linked to centrosome organization during ciliogenesis. *Nature Communications*., **12**, 1351.

Lasko, P., 2012: mRNA localization and translational control in Drosophila oogenesis. *Cold Spring Harbor Perspectives in Biology*., **4**.

Laver, J. D.; Marsolais, A. J.; Smibert, C. A.; Lipshitz, H. D., 2015: Regulation and Function of Maternal Gene Products During the Maternal-to-Zygotic Transition in Drosophila. *Current Topics in Developmental Biology*., **113**, 43–84.

Le Hir, H.; Moore, M. J.; Maquat, L. E., 2000a: Pre-mRNA splicing alters mRNP composition: evidence for stable association of proteins at exon-exon junctions. *Genes & Development*., **14**, 1098–1108.

Le Hir, H.; Izaurralde, E.; Maquat, L. E.; Moore, M. J., 2000b: The spliceosome deposits multiple proteins 20-24 nucleotides upstream of mRNA exon-exon junctions. *The EMBO Journal*., **19**, 6860–6869.

Le Hir, H.; Gatfield, D.; Izaurralde, E.; Moore, M. J., 2001: The exon-exon junction complex provides a binding platform for factors involved in mRNA export and nonsense-mediated mRNA decay. *The EMBO Journal*., **20**, 4987–4997.

Le Hir, H.; Andersen, G. R., 2008: Structural insights into the exon junction complex. *Current Opinion in Structural Biology*., **18**, 112–119.

Lécuyer, E.; Yoshida, H.; Parthasarathy, N.; Alm, C.; Babak, T.; Cerovina, T.; Hughes, T. R.; Tomancak, P.; Krause, H. M., 2007: Global analysis of mRNA localization reveals a prominent role in organizing cellular architecture and function. *Cell*., **131**, 174–187.

Lee, M. T.; Bonneau, A. R.; Takacs, C. M.; Bazzini, A. A.; DiVito, K. R.; Fleming, E. S.; Giraldez, A. J., 2013: Nanog, Pou5f1 and SoxB1 activate zygotic gene expression during the maternal-to-zygotic transition. *Nature*., **503**, 360–364.

Lehmann, R.; Nüsslein-Volhard, C., 1986: Abdominal segmentation, pole cell formation, and embryonic polarity require the localized activity of oskar, a maternal gene in Drosophila. *Cell*., **47**, 141–152.

Lejeune, F.; Ishigaki, Y.; Li, X.; Maquat, L. E., 2002: The exon junction complex is detected on CBP80-bound but not eIF4E-bound mRNA in mammalian cells: dynamics of mRNP remodeling. *The EMBO Journal*., **21**, 3536–3545.

Leppek, K.; Stoecklin, G., 2014: An optimized streptavidin-binding RNA aptamer for purification of ribonucleoprotein complexes identifies novel ARE-binding proteins. *Nucleic Acids Research*., **42**, e13.

Levine, E.; McHale, P.; Levine, H., 2007: Small regulatory RNAs may sharpen spatial expression patterns. *PLoS Computational Biology*., **3**, e233.

Licatalosi, D. D.; Mele, A.; Fak, J. J.; Ule, J.; Kayikci, M.; Chi, S. W.; Clark, T. A.; Schweitzer, A. C.; Blume, J. E.; Wang, X.; Darnell, J. C.; Darnell, R. B., 2008: HITS-CLIP yields genome-wide insights into brain alternative RNA processing. *Nature*., **456**, 464–469.

Lingner, J.; Cech, T. R., 1996: Purification of telomerase from Euplotes aediculatus: requirement of a primer 3' overhang. *Proceedings of the National Academy of Sciences of the United States of America*., **93**, 10712–10717.

Litman, P.; Barg, J.; Rindzoonski, L.; Ginzburg, I., 1993: Subcellular localization of tau mRNA in differentiating neuronal cell culture: implications for neuronal polarity. *Neuron*., **10**, 627–638.

Little, S. C.; Sinsimer, K. S.; Lee, J. J.; Wieschaus, E. F.; Gavis, E. R., 2015: Independent and coordinate trafficking of single Drosophila germ plasm mRNAs. *Nature Cell Biology*., **17**, 558–568.

Loiseau, P.; Davies, T.; Williams, L. S.; Mishima, M.; Palacios, I. M., 2010: Drosophila PAT1 is required for Kinesin-1 to transport cargo and to maximize its motility. *Development*., **137**, 2763–2772.

Mabin, J. W.; Woodward, L. A.; Patton, R. D.; Yi, Z.; Jia, M.; Wysocki, V. H.; Bundschuh, R.; Singh, G., 2018: The Exon Junction Complex Undergoes a Compositional Switch that Alters mRNP Structure and Nonsense-Mediated mRNA Decay Activity. *Cell reports*., **25**, 2431-2446.e7.

Macdonald, P. M.; Kerr, K., 1997: Redundant RNA recognition events in bicoid mRNA localization. *RNA (New York)*., **3**, 1413–1420.

MacDougall, N.; Clark, A.; MacDougall, E.; Davis, I., 2003: Drosophila gurken (TGFalpha) mRNA localizes as particles that move within the oocyte in two dynein-dependent steps. *Developmental Cell*., **4**, 307–319.

Markussen, F. H.; Michon, A. M.; Breitwieser, W.; Ephrussi, A., 1995: Translational control of oskar generates short OSK, the isoform that induces pole plasma assembly. *Development*., **121**, 3723–3732.

Martin, K. C.; Ephrussi, A., 2009: mRNA localization: gene expression in the spatial dimension. *Cell*., **136**, 719–730.

Mayeda, A.; Badolato, J.; Kobayashi, R.; Zhang, M. Q.; Gardiner, E. M.; Krainer, A. R., 1999: Purification and characterization of human RNPS1: a general activator of pre-mRNA splicing. *The EMBO Journal*., **18**, 4560–4570.

McClintock, M. A.; Dix, C. I.; Johnson, C. M.; McLaughlin, S. H.; Maizels, R. J.; Hoang, H. T.; Bullock, S. L., 2018: RNA-directed activation of cytoplasmic dynein-1 in reconstituted transport RNPs. *eLife*., **7**.

McGarvey, T.; Rosonina, E.; McCracken, S.; Li, Q.; Arnaout, R.; Mientjes, E.; Nickerson, J. A.; Awrey, D.; Greenblatt, J.; Grosveld, G.; Blencowe, B. J., 2000: The acute myeloid leukemia-associated protein, DEK, forms a splicing-dependent interaction with exon-product complexes. *The Journal of Cell Biology*., **150**, 309–320.

McHugh, C. A.; Chen, C. K.; Chow, A.; Surka, C. F.; Tran, C.; McDonel, P.; Pandya-Jones, A.; Blanco, M.; Burghard, C.; Moradian, A.; Sweredoski, M. J.; Shishkin, A. A.; Su, J.; Lander, E. S.; Hess, S.; Plath, K.; Guttman, M., 2015: The Xist lncRNA interacts directly with SHARP to silence transcription through HDAC3. *Nature*., **521**, 232–236.

McLaughlin, J. M.; Bratu, D. P., 2015: Drosophila melanogaster Oogenesis: An Overview. *Methods in Molecular Biology*., **1328**, 1–20.

Messitt, T. J.; Gagnon, J. A.; Kreiling, J. A.; Pratt, C. A.; Yoon, Y. J.; Mowry, K. L., 2008: Multiple kinesin motors coordinate cytoplasmic RNA transport on a subpopulation of microtubules in Xenopus oocytes. *Developmental Cell*., **15**, 426–436.

Mili, S.; Macara, I. G., 2009: RNA localization and polarity: from A(PC) to Z(BP). *Trends in Cell Biology*., **19**, 156–164.

Minajigi, A.; Froberg, J.; Wei, C.; Sunwoo, H.; Kesner, B.; Colognori, D.; Lessing, D.; Payer, B.; Boukhali, M.; Haas, W.; Lee, J. T., 2015: Chromosomes. A comprehensive Xist interactome reveals cohesin repulsion and an RNA-directed chromosome conformation. *Science*., **349**.

Mishler, D. M.; Christ, A. B.; Steitz, J. A., 2008: Flexibility in the site of exon junction complex deposition revealed by functional group and RNA secondary structure alterations in the splicing substrate. *RNA (New York)*., **14**, 2657–2670.

Mitsumori, K.; Takei, Y.; Hirokawa, N., 2017: Components of RNA granules affect their localization and dynamics in neuronal dendrites. *Molecular Biology of the Cell*., **28**, 1412–1417.

Mohr, S. E.; Dillon, S. T.; Boswell, R. E., 2001: The RNA-binding protein Tsunagi interacts with Mago Nashi to establish polarity and localize oskar mRNA during Drosophila oogenesis. *Genes & Development*., **15**, 2886–2899.

Mukherjee, J.; Hermesh, O.; Eliscovich, C.; Nalpas, N.; Franz-Wachtel, M.; Maček, B.; Jansen, R. P., 2019: β-Actin mRNA interactome mapping by proximity biotinylation. *Proceedings of the National Academy of Sciences of the United States of America*., **116**, 12863–12872.

Nagel, M.; Noss, M.; Xu, J.; Horn, N.; Ueffing, M.; Boldt, K.; Schuele, R., 2022: The kinesin motor KIF1C is a putative transporter of the exon junction complex in neuronal cells. *RNA (New York)*., **29**, 55–68.

Nakamura, A.; Amikura, R.; Hanyu, K.; Kobayashi, S., 2001: Me31B silences translation of oocyte-localizing RNAs through the formation of cytoplasmic RNP complex during Drosophila oogenesis. *Development*., **128**, 3233–3242.

Navarro, C.; Puthalakath, H.; Adams, J. M.; Strasser, A.; Lehmann, R., 2004: Egalitarian binds dynein light chain to establish oocyte polarity and maintain oocyte fate. *Nature Cell Biology*., **6**, 427–435.

Neuman-Silberberg, F. S.; Schüpbach, T., 1993: The Drosophila dorsoventral patterning gene gurken produces a dorsally localized RNA and encodes a TGF alpha-like protein. *Cell*., **75**, 165–174.

Newmark, P. A.; Boswell, R. E., 1994: The mago nashi locus encodes an essential product required for germ plasm assembly in Drosophila. *Development*., **120**, 1303–1313.

Niranjanakumari, S.; Lasda, E.; Brazas, R.; Garcia-Blanco, M. A., 2002: Reversible cross-linking combined with immunoprecipitation to study RNA-protein interactions in vivo. *Methods*., **26**, 182–190.

Obrdlik, A.; Lin, G.; Haberman, N.; Ule, J.; Ephrussi, A., 2019: The Transcriptome-wide Landscape and Modalities of EJC Binding in Adult Drosophila. *Cell reports*., **28**, 1219-1236.e11.

O'Hare, K.; Rubin, G. M., 1983: Structures of P transposable elements and their sites of insertion and excision in the Drosophila melanogaster genome. *Cell*., **34**, 25–35.

Pace, C. N.; Heinemann, U.; Hahn, U.; Saenger, W., 1991: Ribonuclease T1: structure, function, and stability. *Angewandte Chemie International Edition*., **30**, 343–360.

Palacios, I. M.; St Johnston, D., 2002: Kinesin light chain-independent function of the Kinesin heavy chain in cytoplasmic streaming and posterior localisation in the Drosophila oocyte. *Development*., **129**, 5473–5485.

Palacios, I. M.; Gatfield, D.; St Johnston, D.; Izaurralde, E., 2004: An eIF4AIII-containing complex required for mRNA localization and nonsense-mediated mRNA decay. *Nature*., **427**, 753–757.

Panda, A. C.; Abdelmohsen, K.; Yoon, J. H.; Martindale, J. L.; Yang, X.; Curtis, J.; Mercken, E. M.; Chenette, D. M.; Zhang, Y.; Schneider, R. J.; Becker, K. G.; de Cabo, R.; Gorospe, M., 2014: RNA-binding protein AUF1 promotes myogenesis by regulating MEF2C expression levels. *Molecular and Cellular Biology*., **34**, 3106–3119.

Panda, A. C.; Abdelmohsen, K.; Martindale, J. L.; Di Germanio, C.; Yang, X.; Grammatikakis, I.; Noh, J. H.; Zhang, Y.; Lehrmann, E.; Dudekula, D. B.; De, S.; Becker, K. G.; White, E. J.; Wilson, G. M.; de Cabo, R.; Gorospe, M., 2016: Novel RNA-binding activity of MYF5 enhances Ccnd1/Cyclin D1 mRNA translation during myogenesis. *Nucleic Acids Research*., **44**, 2393–2408.

Patton, R. D.; Sanjeev, M.; Woodward, L. A.; Mabin, J. W.; Bundschuh, R.; Singh, G., 2020: Chemical crosslinking enhances RNA immunoprecipitation for efficient identification of binding sites of proteins that photo-crosslink poorly with RNA. *RNA (New York)*., **26**, 1216–1233.

Perkins, L. A.; Holderbaum, L.; Tao, R.; Hu, Y.; Sopko, R.; McCall, K.; Yang-Zhou, D.; Flockhart, I.; Binari, R.; Shim, H. S.; Miller, A.; Housden, A.; Foos, M.; Randkelv, S.; Kelley, C.; Namgyal, P.; Villalta, C.; Liu, L. P.; Jiang, X. et al., 2015: The transgenic rnai project at harvard medical school: resources and validation. *Genetics*., **201**, 843–852.

Ramanathan, M.; Majzoub, K.; Rao, D. S.; Neela, P. H.; Zarnegar, B. J.; Mondal, S.; Roth, J. G.; Gai, H.; Kovalski, J. R.; Siprashvili, Z.; Palmer, T. D.; Carette, J. E.; Khavari, P. A., 2018: RNA-protein interaction detection in living cells. *Nature Methods*., **15**, 207–212.

Rhee, H. W.; Zou, P.; Udeshi, N. D.; Martell, J. D.; Mootha, V. K.; Carr, S. A.; Ting, A. Y., 2013: Proteomic mapping of mitochondria in living cells via spatially restricted enzymatic tagging. *Science*., **339**, 1328–1331.

Richardson, R. W.; Gumport, R. I., 1983: Biotin and fluorescent labeling of RNA using T4 RNA ligase. *Nucleic Acids Research*., **11**, 6167–6184.

Riechmann, V.; Ephrussi, A., 2001: Axis formation during Drosophila oogenesis. *Current Opinion in Genetics & Development*., **11**, 374–383.

Robinson, J. T.; Thorvaldsdóttir, H.; Winckler, W.; Guttman, M.; Lander, E. S.; Getz, G.; Mesirov, J. P., 2011: Integrative genomics viewer. *Nature Biotechnology*., **29**, 24–26.

Rogell, B.; Fischer, B.; Rettel, M.; Krijgsveld, J.; Castello, A.; Hentze, M. W., 2017: Specific RNP capture with antisense LNA/DNA mixmers. *RNA (New York)*., **23**, 1290–1302.

Rongo, C.; Gavis, E. R.; Lehmann, R., 1995: Localization of oskar RNA regulates oskar translation and requires Oskar protein. *Development*., **121**, 2737–2746.

Rørth, P., 1998: Gal4 in the Drosophila female germline. *Mechanisms of Development*., **78**, 113–118.

Rossoll, W.; Kröning, A. K.; Ohndorf, U. M.; Steegborn, C.; Jablonka, S.; Sendtner, M., 2002: Specific interaction of Smn, the spinal muscular atrophy determining gene product, with hnRNP-R and gry-rbp/hnRNP-Q: a role for Smn in RNA processing in motor axons? *Human Molecular Genetics*., **11**, 93–105.

Roux, K. J.; Kim, D. I.; Raida, M.; Burke, B., 2012: A promiscuous biotin ligase fusion protein identifies proximal and interacting proteins in mammalian cells. *The Journal of Cell Biology*., **196**, 801–810.

Saha, K.; England, W.; Fernandez, M. M.; Biswas, T.; Spitale, R. C.; Ghosh, G., 2020: Structural disruption of exonic stem-loops immediately upstream of the intron regulates mammalian splicing. *Nucleic Acids Research*., **48**, 6294–6309.

Sanghavi, P.; Laxani, S.; Li, X.; Bullock, S. L.; Gonsalvez, G. B., 2013: Dynein associates with oskar mRNPs and is required for their efficient net plus-end localization in Drosophila oocytes. *Plos One*., **8**, e80605.

Saulière, J.; Haque, N.; Harms, S.; Barbosa, I.; Blanchette, M.; Le Hir, H., 2010: The exon junction complex differentially marks spliced junctions. *Nature Structural & Molecular Biology*., **17**, 1269–1271.

Saulière, J.; Murigneux, V.; Wang, Z.; Marquenet, E.; Barbosa, I.; Le Tonquèze, O.; Audic, Y.; Paillard, L.; Roest Crollius, H.; Le Hir, H., 2012: CLIP-seq of eIF4AIII reveals transcriptome-wide mapping of the human exon junction complex. *Nature Structural & Molecular Biology*., **19**, 1124–1131.

Savas, J. N.; Ma, B.; Deinhardt, K.; Culver, B. P.; Restituito, S.; Wu, L.; Belasco, J. G.; Chao, M. V.; Tanese, N., 2010: A role for huntington disease protein in dendritic RNA granules. *The Journal of Biological Chemistry*., **285**, 13142–13153.

Schindelin, J.; Arganda-Carreras, I.; Frise, E.; Kaynig, V.; Longair, M.; Pietzsch, T.; Preibisch, S.; Rueden, C.; Saalfeld, S.; Schmid, B.; Tinevez, J. Y.; White, D. J.; Hartenstein, V.; Eliceiri, K.; Tomancak, P.; Cardona, A., 2012: Fiji: an open-source platform for biological-image analysis. *Nature Methods*., **9**, 676–682.

Schliwa, M.; Woehlke, G., 2003: Molecular motors. *Nature*., **422**, 759–765.

Semotok, J. L.; Cooperstock, R. L.; Pinder, B. D.; Vari, H. K.; Lipshitz, H. D.; Smibert, C. A., 2005: Smaug recruits the CCR4/POP2/NOT deadenylase complex to trigger maternal transcript localization in the early Drosophila embryo. *Current Biology*., **15**, 284–294.

Sharma, S.; Wongpalee, S. P.; Vashisht, A.; Wohlschlegel, J. A.; Black, D. L., 2014: Stem-loop 4 of U1 snRNA is essential for splicing and interacts with the U2 snRNP-specific SF3A1 protein during spliceosome assembly. *Genes & Development*., **28**, 2518–2531.

Shibuya, T.; Tange, T. Ø.; Sonenberg, N.; Moore, M. J., 2004: eIF4AIII binds spliced mRNA in the exon junction complex and is essential for nonsense-mediated decay. *Nature Structural & Molecular Biology*., **11**, 346–351.

Shpetner, H. S.; Vallee, R. B., 1991: Purification and characterization of dynamin. *Methods in Enzymology*., **196**, 192–201.

Simon, B.; Masiewicz, P.; Ephrussi, A.; Carlomagno, T., 2015: The structure of the SOLE element of oskar mRNA. *RNA (New York)*., **21**, 1444–1453.

Singh, G.; Kucukural, A.; Cenik, C.; Leszyk, J. D.; Shaffer, S. A.; Weng, Z.; Moore, M. J., 2012: The cellular EJC interactome reveals higher-order mRNP structure and an EJC-SR protein nexus. *Cell*., **151**, 750–764.

Singh, G.; Ricci, E. P.; Moore, M. J., 2014: RIPiT-Seq: a high-throughput approach for footprinting RNA:protein complexes. *Methods*., **65**, 320–332.

Sladewski, T. E.; Billington, N.; Ali, M. Y.; Bookwalter, C. S.; Lu, H.; Krementsova, E. B.; Schroer, T. A.; Trybus, K. M., 2018: Recruitment of two dyneins to an mRNA-dependent Bicaudal D transport complex. *eLife*., **7**.

Slobodin, B.; Gerst, J. E., 2010: A novel mRNA affinity purification technique for the identification of interacting proteins and transcripts in ribonucleoprotein complexes. *RNA (New York)*., **16**, 2277–2290.

Smith, R., 2004: Moving molecules: mRNA trafficking in Mammalian oligodendrocytes and neurons. *The Neuroscientist*., **10**, 495–500.

Smola, M. J.; Rice, G. M.; Busan, S.; Siegfried, N. A.; Weeks, K. M., 2015: Selective 2'-hydroxyl acylation analyzed by primer extension and mutational profiling (SHAPE-MaP) for direct, versatile and accurate RNA structure analysis. *Nature Protocols*., **10**, 1643–1669.

Smola, M. J.; Weeks, K. M., 2018: In-cell RNA structure probing with SHAPE-MaP. *Nature Protocols*., **13**, 1181–1195.

Srisawat, C.; Engelke, D. R., 2001: Streptavidin aptamers: affinity tags for the study of RNAs and ribonucleoproteins. *RNA (New York)*., **7**, 632–641.

St Johnston, D.; Driever, W.; Berleth, T.; Richstein, S.; Nüsslein-Volhard, C., 1989: Multiple steps in the localization of bicoid RNA to the anterior pole of the Drosophila oocyte. *Development*., **107 Suppl**, 13–19.

Starz-Gaiano, M.; Cho, N. K.; Forbes, A.; Lehmann, R., 2001: Spatially restricted activity of a Drosophila lipid phosphatase guides migrating germ cells. *Development*., **128**, 983–991.

Strittmatter, L. M.; Capitanchik, C.; Newman, A. J.; Hallegger, M.; Norman, C. M.; Fica, S. M.; Oubridge, C.; Luscombe, N. M.; Ule, J.; Nagai, K., 2020: psiCLIP reveals dynamic RNA binding by DEAH-box helicases before and after exon ligation. *BioRxiv*.

Sysoev, V. O.; Fischer, B.; Frese, C. K.; Gupta, I.; Krijgsveld, J.; Hentze, M. W.; Castello, A.; Ephrussi, A., 2016: Global changes of the RNA-bound proteome during the maternal-to-zygotic transition in Drosophila. *Nature Communications*., **7**, 12128.

Tange, T. Ø.; Nott, A.; Moore, M. J., 2004: The ever-increasing complexities of the exon junction complex. *Current Opinion in Cell Biology*., **16**, 279–284.

Theurkauf, W. E.; Smiley, S.; Wong, M. L.; Alberts, B. M., 1992: Reorganization of the cytoskeleton during Drosophila oogenesis: implications for axis specification and intercellular transport. *Development*., **115**, 923–936.

Tian, L.; Okita, T. W., 2014: mRNA-based protein targeting to the endoplasmic reticulum and chloroplasts in plant cells. *Current Opinion in Plant Biology*., **22**, 77–85.

Tiruchinapalli, D. M.; Oleynikov, Y.; Kelic, S.; Shenoy, S. M.; Hartley, A.; Stanton, P. K.; Singer, R. H.; Bassell, G. J., 2003: Activity-dependent trafficking and dynamic localization of zipcode binding protein 1 and beta-actin mRNA in dendrites and spines of hippocampal neurons. *The Journal of Neuroscience*., **23**, 3251–3261.

Trovisco, V.; Belaya, K.; Nashchekin, D.; Irion, U.; Sirinakis, G.; Butler, R.; Lee, J. J.; Gavis, E. R.; St Johnston, D., 2016: bicoid mRNA localises to the Drosophila oocyte anterior by random Dynein-mediated transport and anchoring. *eLife*., **5**.

Uehata, T.; Takeuchi, O., 2020: RNA Recognition and Immunity-Innate Immune Sensing and Its Posttranscriptional Regulation Mechanisms. *Cells*., **9**.

Ule, J.; Jensen, K.; Mele, A.; Darnell, R. B., 2005: CLIP: a method for identifying protein-RNA interaction sites in living cells. *Methods*., **37**, 376–386.

Urdaneta, E. C.; Beckmann, B. M., 2020: Fast and unbiased purification of RNA-protein complexes after UV cross-linking. *Methods*., **178**, 72–82.

Vaishali; Dimitrova-Paternoga, L.; Haubrich, K.; Sun, M.; Ephrussi, A.; Hennig, J., 2021: Validation and classification of RNA binding proteins identified by mRNA interactome capture. *RNA (New York)*., **27**, 1173–1185.

Vaishali, V., 2022: Characterisation of RNA binding proteins and their roles in the Drosophila germline. *Heidelberg University Library*.

van Eeden, F. J.; Palacios, I. M.; Petronczki, M.; Weston, M. J.; St Johnston, D., 2001: Barentsz is essential for the posterior localization of oskar mRNA and colocalizes with it to the posterior pole. *The Journal of Cell Biology*., **154**, 511–523.

Vanzo, N. F.; Ephrussi, A., 2002: Oskar anchoring restricts pole plasm formation to the posterior of the Drosophila oocyte. *Development*., **129**, 3705–3714.

Vazquez-Pianzola, P.; Suter, B., 2012: Conservation of the RNA Transport Machineries and Their Coupling to Translation Control across Eukaryotes. *Comparative and functional genomics*., **2012**, 287852.

Wagh, D. A.; Rasse, T. M.; Asan, E.; Hofbauer, A.; Schwenkert, I.; Dürrbeck, H.; Buchner, S.; Dabauvalle, M. C.; Schmidt, M.; Qin, G.; Wichmann, C.; Kittel, R.; Sigrist, S. J.; Buchner, E., 2006: Bruchpilot, a protein with homology to ELKS/CAST, is required for structural integrity and function of synaptic active zones in Drosophila. *Neuron*., **49**, 833–844.

Wang, C.; Lehmann, R., 1991: Nanos is the localized posterior determinant in Drosophila. *Cell*., **66**, 637–647.

Wang, Z.; Ballut, L.; Barbosa, I.; Le Hir, H., 2018: Exon Junction Complexes can have distinct functional flavours to regulate specific splicing events. *Scientific Reports*., **8**, 9509.

Wessels, H. H.; Imami, K.; Baltz, A. G.; Kolinski, M.; Beldovskaya, A.; Selbach, M.; Small, S.; Ohler, U.; Landthaler, M., 2016: The mRNA-bound proteome of the early fly embryo. *Genome Research*., **26**, 1000–1009.

West, J. A.; Davis, C. P.; Sunwoo, H.; Simon, M. D.; Sadreyev, R. I.; Wang, P. I.; Tolstorukov, M. Y.; Kingston, R. E., 2014: The long noncoding RNAs NEAT1 and MALAT1 bind active chromatin sites. *Molecular Cell*., **55**, 791–802.

Wilhelm, J. E.; Vale, R. D., 1993: RNA on the move: the mRNA localization pathway. *The Journal of Cell Biology*., **123**, 269–274.

Wilhelm, J. E.; Hilton, M.; Amos, Q.; Henzel, W. J., 2003: Cup is an eIF4E binding protein required for both the translational repression of oskar and the recruitment of Barentsz. *The Journal of Cell Biology*., **163**, 1197–1204.

Wippich, F.; Ephrussi, A., 2020: Transcript specific mRNP capture from Drosophila egg-chambers for proteomic analysis. *Methods*., **178**, 83–88.

Xiol, J.; Spinelli, P.; Laussmann, M. A.; Homolka, D.; Yang, Z.; Cora, E.; Couté, Y.; Conn, S.; Kadlec, J.; Sachidanandam, R.; Kaksonen, M.; Cusack, S.; Ephrussi, A.; Pillai, R. S., 2014: RNA clamping by Vasa assembles a piRNA amplifier complex on transposon transcripts. *Cell*., **157**, 1698–1711.

Yehle, C. O.; Patterson, W. L.; Boguslawski, S. J.; Albarella, J. P.; Yip, K. F.; Carrico, R. J., 1987: A solution hybridization assay for ribosomal RNA from bacteria using biotinylated DNA probes and enzyme-labeled antibody to DNA:RNA. *Molecular and Cellular Probes*., **1**, 177–193.

Yi, T.; Arthanari, H.; Akabayov, B.; Song, H.; Papadopoulos, E.; Qi, H. H.; Jedrychowski, M.; Güttler, T.; Guo, C.; Luna, R. E.; Gygi, S. P.; Huang, S. A.; Wagner, G., 2015: eIF1A augments Ago2-mediated Dicer-independent miRNA biogenesis and RNA interference. *Nature Communications*., **6**, 7194.

Zarnegar, B. J.; Flynn, R. A.; Shen, Y.; Do, B. T.; Chang, H. Y.; Khavari, P. A., 2016: irCLIP platform for efficient characterization of protein-RNA interactions. *Nature Methods*., **13**, 489–492.

Zhang, Z.; Sun, W.; Shi, T.; Lu, P.; Zhuang, M.; Liu, J. L., 2020: Capturing RNA-protein interaction via CRUIS. *Nucleic Acids Research*., **48**, e52.

Zhao, Y.; Zhang, Y.; Teng, Y.; Liu, K.; Liu, Y.; Li, W.; Wu, L., 2019: SpyCLIP: an easy-to-use and high-throughput compatible CLIP platform for the characterization of protein-RNA interactions with high accuracy. *Nucleic Acids Research*., **47**, e33.

Zimyanin, V. L.; Belaya, K.; Pecreaux, J.; Gilchrist, M. J.; Clark, A.; Davis, I.; St Johnston, D., 2008: In vivo imaging of oskar mRNA transport reveals the mechanism of posterior localization. *Cell*., **134**, 843–853.

Zubradt, M.; Gupta, P.; Persad, S.; Lambowitz, A. M.; Weissman, J. S.; Rouskin, S., 2017: DMS-MaPseq for genome-wide or targeted RNA structure probing in vivo. *Nature Methods*., **14**, 75–82.

Zuccato, C.; Cattaneo, E., 2009: Brain-derived neurotrophic factor in neurodegenerative diseases. *Nature Reviews. Neurology*., **5**, 311–322.

Zuker, M., 2003: Mfold web server for nucleic acid folding and hybridization prediction. *Nucleic Acids Research*., **31**, 3406–3415.

List of figures

Appendix

A. Supplementary figures

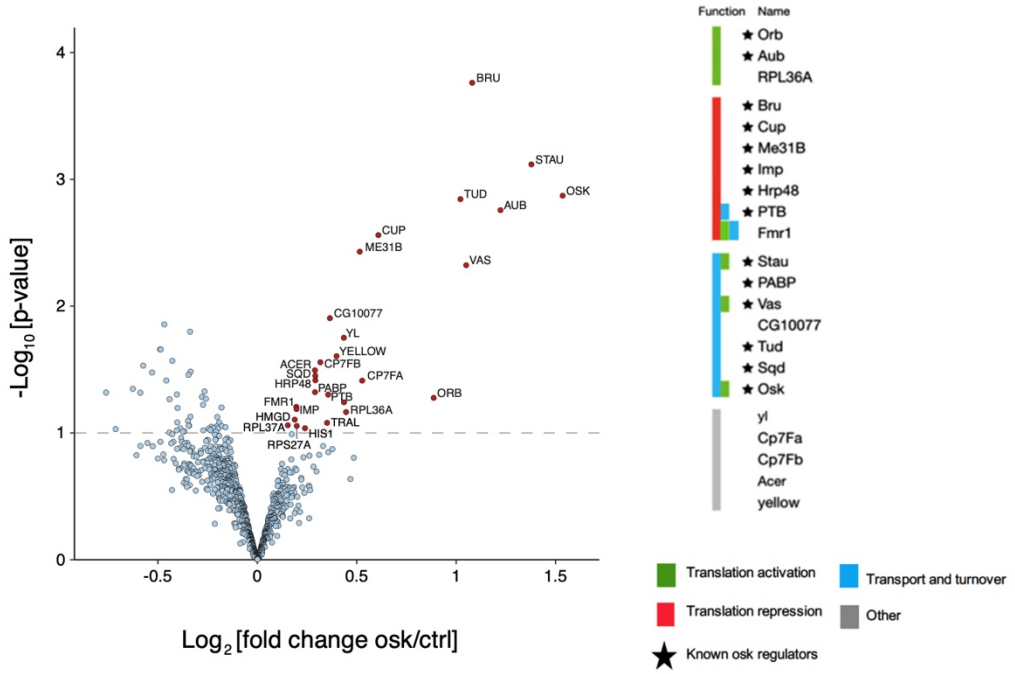

Figure S1 Transcript-specific RBP capture of endogenous *oskar* mRNA. The volcano plot displays the RBP identified by mass spectrometry. Known protein functions of identified proteins are listed on the right (Data and image from Frank Wippich).

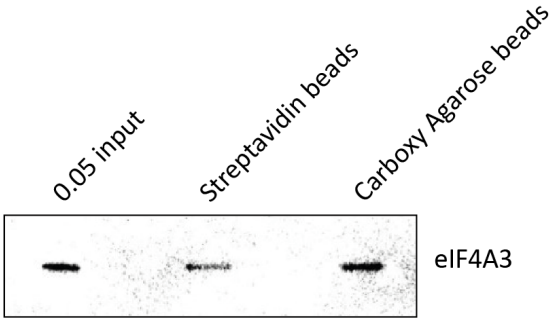

Figure S2 Western blot analysis of *Drosophila* ovary lysate incubated with naked beads. The EJC core component eIF4A3 is binding to the beads matrix independent of RNA, biotin or streptavidin.

Figure S3 Enlarged image of Figure 2.13 (A). Mass spectrometry results after the first transcript-specific RBP capture of *lacZ* WT and *lacZ* SL RNA.

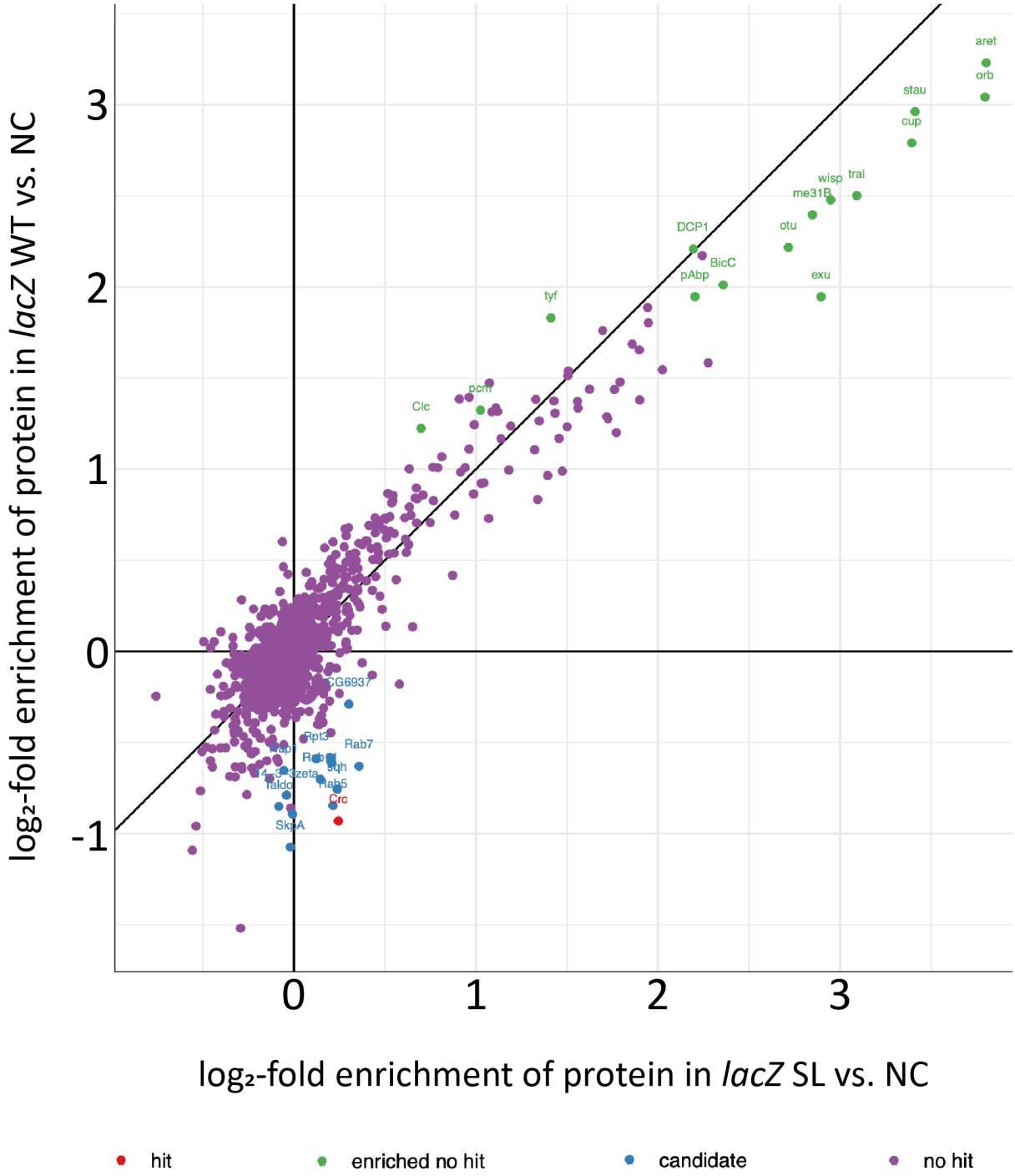

Figure S4 Enlarged image of Figure 2.13 (B). Mass spectrometry results normalized after multiple transcript-specific RBP capture of *lacZ* WT and *lacZ* SL RNA.

Figure S5 Enlarged image of Figure 2.14. Mass spectrometry results normalized after multiple transcript-specific RBP capture of *lacZ* SLc and *lacZ* IL RNA.

B. Antisense DNA oligonucleotides for smFISH

Name	Sequence (5' to 3')
lacZ_1	TAGCGAAACCGCCAAGAC
lacZ_2	GGGCGTATTCGCAAAGGA
lacZ_3	TCCAGGTAGCGAAAGCCA
lacZ_4	GGATCATCGGTCAGACGA
lacZ_5	GCAGGATATCCTGCACCA
lacZ_6	ACCTGACCATGCAGAGGA
lacZ_7	AGTTCAACCACCGCACGA
lacZ_8	CGACGTTCAGACGTAGTG
lacZ_9	GCATAACCACCACGCTCA
lacZ_10	CGTAGGTAGTCACGCAAC
lacZ_11	CTGAACTTCAGCCTCCAG
lacZ_12	CGGGTTGCCGTTTTCATCA
lacZ_13	GCGAGTGGCAACATGGAAA
lacZ_14	CGTTCATACAGAACTGGCGA
lacZ_15	CGCGTACATCGGGCAAATAA
lacZ_16	CCGTGGGTTTCAATATTGGC
lacZ_17	GACCATTTTCAATCCGCACC
lacZ_18	GTTTATGCAGCAACGAGACG
lacZ_19	TCGCCAAAATCACCGCCG
lacZ_20	ATCCAGCGATACAGCGCG
lacZ_21	CATTCCCCAGCGACCAGA
lacZ_22	GGGTGATTACGATCGCGC
lacZ_23	AGCGTGTACCACAGCGGA
lacZ_24	CAGCGCACGGCGTTAAAG
lacZ_25	CCTCGAATCAGCAACGGC

Name	Sequence (5' to 3')
lacZ_26	GACATCGCAGGCTTCTGC
lacZ_27	GATTTCGGCGCTCCACAG
lacZ_28	CTGCGTTTCACCCTGCCA
lacZ_29	CCGCTCATCCGCCACATA
lacZ_30	CACTCCAACGCAGCACCA
lacZ_31	CACAGATGAAACGCCGAG
lacZ_32	TGAGCGAGTAACAACCCG
lacZ_33	GGATAGGTCACGTTGGTG
lacZ_34	TAACCGTGCATCTGCCAG
lacZ_35	GTATCGGCCTCAGGAAGA
lacZ_36	AACTGTTGGGAAGGGCGA
lacZ_37	CTTCGCTATTACGCCAGC
lacZ_38	GATGTGCTGCAAGGCGAT
lacZ_39	ATTCGCGTCTGGCCTTCC
lacZ_40	GTGGGAACAAACGGCGGA
lacZ_41	GCAAAGCGCCATTCGCCA
lacZ_42	ACGCCAGGGTTTTCCCAG
lacZ_43	CAAATTCAGACGGCAAACGAC
lacZ_44	AGCTTTCCGGCACCGCTTC
lacZ_45	ACGACGGCCAGTGAATCCG

Name	Sequence (5' to 3')
osk3'UTR-1	ATCGCGCAAATGCTTCAC
osk3'UTR-2	TTAAGGGCAAGTGGCAGG
osk3'UTR-3	ACGTGATCACCATCAATACA
osk3'UTR-4	AGCTGTAAATTACGCCAGAA
osk3'UTR-5	TGCTACAAACAAGCGCTTAG
osk3'UTR-6	TCTGCAGCAGAGTGTAAGCA
osk3'UTR-7	AATTTGCTTGAGCACATCAA
osk3'UTR-8	GTTGATTTTGTGCAAGCGAA

Name	Sequence (5' to 3')
osk3'UTR-9	TTCCAAGTAAAGCAGTGCA
osk3'UTR-10	TTACGGCCAAAATGCAGCA
osk3'UTR-11	TGTATACGTACCACGCCCAC
osk3'UTR-12	TGATACAGGAGCATGCCGAA
osk3'UTR-13	GCGGAAAAGTTTGAAGAGAAG
osk3'UTR-14	CTGCTTGCGCTTATTTTGCA
osk3'UTR-15	CGAATTCCGTAAAAGCCGA

C. Antisense DNA oligonucleotides for transcript-specific RBP capture

Name	Sequence (5' to 3')
Lz_1_1	TAGCGAAACCGCCAAGAC[BtnTg]
Lz_1_2	GGGCGTATTCGCAAAGGA[BtnTg]
Lz_1_3	TCCAGGTAGCGAAAGCCA[BtnTg]
Lz_1_4	GGATCATCGGTCAGACGA[BtnTg]
Lz_1_5	GCAGGATATCCTGCACCA[BtnTg]
Lz_1_6	ACCTGACCATGCAGAGGA[BtnTg]
Lz_1_7	AGTTCAACCACCGCACGA[BtnTg]
Lz_1_8	CGACGTTCAGACGTAGTG[BtnTg]
Lz_1_9	GCATAACCACCACGCTCA[BtnTg]
Lz_1_10	CGTAGGTAGTCACGCAAC[BtnTg]
Lz_1_11	CTGAACTTCAGCCTCCAG[BtnTg]
Lz_1_12	CGGGTTGCCGTTTTCATCA[BtnTg]
Lz_1_13	GCGAGTGGCAACATGGAAA[BtnTg]
Lz_1_14	CGTTCATACAGAACTGGCGA[BtnTg]
Lz_1_15	CGCGTACATCGGGCAAATAA[BtnTg]
Lz_1_16	CCGTGGGTTTCAATATTGGC[BtnTg]
Lz_1_17	GACCATTTTCAATCCGCACC[BtnTg]
Lz_1_18	GTTTATGCAGCAACGAGACG[BtnTg]
Lz_1_19	TCGCCAAAATCACCGCCG[BtnTg]
Lz_1_20	ATCCAGCGATACAGCGCG[BtnTg]
Lz_1_21	CATTCCCCAGCGACCAGA[BtnTg]
Lz_1_22	GGGTGATTACGATCGCGC[BtnTg]
Lz_1_23	AGCGTGTACCACAGCGGA[BtnTg]
Lz_1_24	CAGCGCACGGCGTTAAAG[BtnTg]
Lz_1_25	CCTCGAATCAGCAACGGC[BtnTg]

Name	Sequence (5' to 3')
Lz_1_26	GACATCGCAGGCTTCTGC[BtnTg]
Lz_1_27	GATTTCGGCGCTCCACAG[BtnTg]
Lz_1_28	CTGCGTTTCACCCTGCCA[BtnTg]
Lz_1_29	CCGCTCATCCGCCACATA[BtnTg]
Lz_1_30	CACTCCAACGCAGCACCA[BtnTg]
Lz_1_31	CACAGATGAAACGCCGAG[BtnTg]
Lz_1_32	TGAGCGAGTAACAACCCG[BtnTg]
Lz_1_33	GGATAGGTCACGTTGGTG[BtnTg]
Lz_1_34	TAACCGTGCATCTGCCAG[BtnTg]
Lz_1_35	GTATCGGCCTCAGGAAGA[BtnTg]
Lz_1_36	AACTGTTGGGAAGGGCGA[BtnTg]
Lz_1_37	CTTCGCTATTACGCCAGC[BtnTg]
Lz_1_38	GATGTGCTGCAAGGCGAT[BtnTg]
Lz_1_39	ATTCGCGTCTGGCCTTCC[BtnTg]
Lz_1_40	GTGGGAACAAACGGCGGA[BtnTg]
Lz_1_41	GCAAAGCGCCATTCGCCA[BtnTg]
Lz_1_42	ACGCCAGGGTTTTCCCAG[BtnTg]
Lz_1_43	CAAATTCAGACGGCAAACGAC[BtnTg]
Lz_1_44	AGCTTTCCGGCACCGCTTC[BtnTg]
Lz_1_45	ACGACGGCCAGTGAATCCG[BtnTg]

Name	Sequence (5' to 3')
Lz_2_1	CCAGTTCGTGCCAAGAATG[BtnTg]
Lz_2_2	GGTAGAGCTTCGATAGCAG[BtnTg]
Lz_2_3	CGTATCGCCAAAATCACCG[BtnTg]
Lz_2_4	AAGACTGTTACCCATCGCG[BtnTg]
Lz_2_5	TCTCTCCAGGTAGCGAAAG[BtnTg]
Lz_2_6	ATGGACCATTTCGGCACAG[BtnTg]
Lz_2_7	TACAGCGCGTCGTGATTAG[BtnTg]
Lz_2_8	ATAATGCGAACAGCGCACG[BtnTg]
Lz_2_9	GTTCTGCTTCATCAGCAGG[BtnTg]
Lz_2_10	CATCGTCTGCTCATCCATG[BtnTg]
Lz_2_11	TCGACGTTCAGACGTAGTG[BtnTg]
Lz_2_12	ATAACCACCACGCTCATCG[BtnTg]
Lz_2_13	TAGGTAGTCACGCAACTCG[BtnTg]
Lz_2_14	ACTTCAGCCTCCAGTACAG[BtnTg]
Lz_2_15	GTCGGTTTATGCAGCAACG[BtnTg]
Lz_2_16	AAAGGGGATGTGCTGCAAG[BtnTg]
Lz_2_17	AAGCCAATTTGGCGGGAAC[BtnTg]
Lz_2_18	AAACAAGCGCTTAGTGCGC[BtnTg]
Lz_2_19	TCTGCAGCAGAGTGTAAGC[BtnTg]
Lz_2_20	GTGGCAGGTTGTGTGATAC[BtnTg]
Lz_2_21	CAGCTTTTGGGATAGCAGC[BtnTg]
Lz_2_22	TGGACTAGGACAAGTGCAC[BtnTg]
Lz_2_23	ATTTTGTCGGTTCCGGTGC[BtnTg]
Lz_2_24	CCGTTCATACAGAACTGGC[BtnTg]
Lz_2_25	AATCAGCGACTGATCCACC[BtnTg]

Name	Sequence (5' to 3')
Lz_2_26	ATACTGACGAAACGCCTGC[BtnTg]
Lz_2_27	ACCATGCAGAGGATGATGC[BtnTg]
Lz_2_28	ACGGTTAACGCCTCGAATC[BtnTg]
Lz_2_29	TTTTCAATCCGCACCTCGC[BtnTg]
Lz_2_30	AATTTCACCGCCGAAAGGC[BtnTg]
Lz_2_31	TCCGCCACATATCCTGATC[BtnTg]
Lz_2_32	CAGATAACTGCCGTCACTC[BtnTg]
Lz_2_33	CGAGGCGGTTTCGATATTC[BtnTg]
Lz_2_34	CCTGTAGCCAGCTTTCATC[BtnTg]
Lz_2_35	TGTGAGCGAGTAACAACCC[BtnTg]
Lz_2_36	GACCGTAATGGGATAGGTC[BtnTg]
Lz_2_37	TATCGGCCTCAGGAAGATC[BtnTg]
Lz_2_38	AATAAATCAGCCGCGGCGT[BtnTg]
Lz_2_39	TATACGTACCACGCCCACT[BtnTg]
Lz_2_40	GTTTTCAGCTGCTTGCGCT[BtnTg]
Lz_2_41	ACCATTCGCGTTACGCGTT[BtnTg]
Lz_2_42	GGATCATCGGTCAGACGAT[BtnTg]
Lz_2_43	AGTTCAACCACCGCACGAT[BtnTg]
Lz_2_44	AGTACAGCGCGGCTGAAAT[BtnTg]
Lz_2_45	CACAGATGAAACGCCGAGT[BtnTg]

Name	Sequence (5' to 3')
lzseq_1	GGATCATCGGTCAGACGA [BtnTg]
lzseq_2	ACCTGACCATGCAGAGGA[BtnTg]
lzseq_3	AGTTCAACCACCGCACGA[BtnTg]
lzseq_4	CGACGTTCAGACGTAGTG[BtnTg]
lzseq_5	GCATAACCACCACGCTCA[BtnTg]
lzseq_6	CGTAGGTAGTCACGCAAC[BtnTg]
lzseq_7	CTGAACTTCAGCCTCCAG[BtnTg]
lzseq_8	CGGGTTGCCGTTTTCATCA[BtnTg]
lzseq_9	GCGAGTGGCAACATGGAAA[BtnTg]
lzseq_10	CGTTCATACAGAACTGGCGA[BtnTg]
lzseq_11	CCGTGGGTTTCAATATTGGC[BtnTg]

Name	Sequence (5' to 3')
lzseq_12	GTTTATGCAGCAACGAGACG [BtnTg]
lzseq_13	TCGCCAAAATCACCGCCG[BtnTg]
lzseq_14	ATCCAGCGATACAGCGCG[BtnTg]
lzseq_15	CCTCGAATCAGCAACGGC[BtnTg]
lzseq_16	GATTTCGGCGCTCCACAG[BtnTg]
lzseq_17	GGATAGGTCACGTTGGTG[BtnTg]
lzseq_18	GTATCGGCCTCAGGAAGA[BtnTg]
lzseq_19	GATGTGCTGCAAGGCGAT[BtnTg]
lzseq_20	ACGCCAGGGTTTTCCCAG[BtnTg]
lzseq_21	AGCTTTCCGGCACCGCTTC[BtnTg]