# Functional conservation of germ plasm organizers Bucky ball in zebrafish and *Drosophila* Oskar

Doctoral Thesis

Dissertation for the award of the degree

"Doctor rerum naturalium (Dr. rer. nat)"

in the GGNB program: Genes and Development

at the Georg August University Göttingen

Faculty of Biology

Submitted by

Pritesh Krishnakumar

born in Chennai, India

Göttingen, September 2017

# Members of the Thesis Committee

Supervisor:

Dr. Roland Dosch (Reviewer)

Department of Developmental Biochemistry, Georg-August-University Göttingen

Second member of the thesis committee:

Prof. Dr. Michael Kessel (Reviewer)

Department of Developmental Biology, Max Planck Institute for Biophysical Chemistry (MPI-BPC), Göttingen

Third member of the thesis committee:

Prof. Dr. Dirk Görlich

Department of Cellular Logistics, Max Planck Institute for Biophysical Chemistry (MPI-BPC), Göttingen

Date of the oral examination:

# Affidavit

Herewith I declare that I prepared the PhD thesis "Functional conservation of germ plasm organizers Bucky ball in zebrafish and *Drosophila* Oskar" on my own and with no other sources and aids than quoted.


30.09.2017

_____
Pritesh Krishnakumar

# Acknowledgements

# Table of Contents

# Abbreviations List

| | |
|---|---|
| °C | Degrees Celsius |
| A | Adenine |
| aa | Amino acids |
| bp | Base pairs |
| BSA | Bovine serum albumin |
| C | Cytosine |
| C | Cysteine |
| *C. elegans* | *Caenorhabditis elegans* |
| cDNA | Complementary DNA |
| Co-IP | Co-immunoprecipitation |
| DAPI | 4',6-diamidino-2-phenylindole |
| DAZ | Deleted in azoospermia |
| Dm | *Drosophila melanogaster* |
| DNA | Deoxyribonucleic acid |
| dNTP | Deoxynucleotide triphosphate |
| *E. coli* | *Escherichia coli* |
| *e.g.* | *Exempli gratia* |
| EDTA | Ethylenediamine tetraacetic acid |
| eGFP | Enhanced green fluorescent protein |
| *et al.* | *Et alii* |
| fwd | Forward |
| G | Guanine |
| GFP | Green fluorescent protein |
| hpf | Hours post fertilization |
| HMM | Hidden Markov models |
| IDP | Intrinsically Disordered Proteins |
| IDR | Intrinsically Disordered Region |
| K | Lysine |
| kb | Kilo base pairs |
| Kd | Kilodalton |
| M | Molar |
| miRNA | Micro RNA |
| ml | Millilitre |
| mM | Millimolar |
| MPI-BPC | Max Planck Institute for Biophysical Chemistry |
| mRNA | Messenger RNA |

| | |
|---|---|
| n | Number |
| ng | Nanogram |
| nl | Nanolitre |
| pg | picogram |
| PAGE | Polyacrylamide gel electrophoresis |
| PBS | Phosphate-buffered saline |
| PBT | Phosphate-buffered saline Triton X-100 |
| PCR | Polymerase chain reaction |
| PGC | Primordial germ cell |
| pH | Potentium hydrogenium |
| piRNA | Piwi-interacting RNA |
| R.F. | Relative Fluorescence |
| RNA | Ribonucleic acid |
| RNase | Ribonuclease |
| RT-PCR | Reverse transcription PCR |
| rev | Reverse |
| s | Second |
| SDS | Sodium dodecyl sulfate |
| S | Serine |
| T | Thymine |
| T | Threonine |
| Taq | Thermus aquaticus |
| TBE | Tris-Borate-EDTA |
| Tris | Tris(hydroxymethyl)aminomethane |
| U | Units |
| UMG | University Medical Center Göttingen |
| UTR | Untranslated region |
| UV | Ultraviolet |
| V | Volt |
| Y | Tyrosine |
| Zf | Zebrafish |
| µg | Microgram |
| µl | Microlitre |
| µm | Micrometer |
| µM | Micromolar |

# Abstract

The reproductive cells of many animals are specified by a maternal determinant termed germ plasm. Cells inheriting germ plasm during embryogenesis will develop into primordial germ cells (PGCs), the precursors to eggs and sperm. Zebrafish specify their germ cells by deposition of germ plasm in the egg. To this end, only two proteins are known in the entire animal kingdom that program germ cell development *in vivo*: Oskar, which is specific to insects and Bucky ball (Buc), which our lab discovered in the zebrafish. Based on this unique activity, these two proteins are termed germ plasm organizers.

In my study, I show that the two germ plasm organizers transform a somatic cell into PGCs in zebrafish. This result indicates that both proteins act through a conserved biochemical network to specify germ cell fate. A comparison of the amino acid sequence of *Drosophila* Oskar and zebrafish Bucky Ball, showed no conserved domains to explain their similar function. Buc is a novel protein, whose protein sequence shows no homology to other known functional domains and hence, its biochemical activity is currently unknown. My analysis of the biochemical interaction network of Bucky ball revealed that it interacts with similar germ plasm factors as Oskar. These results suggest that germ plasm organizers form a common core complex to specify germ cells.

Nonetheless, Oskar and Bucky ball represent a deviation from the generally accepted sequence-structure-function paradigm. I therefore analyzed their biophysical properties and discovered that both encode intrinsically disordered proteins (IDPs). IDPs are known to quickly change their sequence during evolution, because their function is based on structural features. IDPs are known to assemble into hydrogels by liquid-liquid phase separation. Pharmacologically disrupting hydrogels fragmented Buc-complexes in zebrafish indicating their critical role during aggregation of germ plasm. These data suggest that Buc and Oskar exert their conserved role through similar biophysical and biochemical properties.

It has been frequently postulated in the literature that unrelated protein sequences perform the same function. To my knowledge, my data represents the first example of two proteins which were previously classified as "novel" or "species-specific", but are indeed functionally conserved across distant species.

# 1. Introduction

Biological systems differ from chemical and physics based setups in their capacity to reproduce. This unique ability allows organisms to propagate their genetic material to the next generation and thus continue their lineage.

Two different types of reproduction have evolved: asexual and sexual reproduction. In asexual reproduction, the future offspring arise from a single parent. The genetic material is of one organism and is transmitted to the next generation, thus making it genetically identical to its parent. This is seen in organism such as archaebacteria, protists, and eubacteria. The energy requirements are less compared to other organisms, because no cells are set aside to make gametes and no meiosis takes place (Crow, 1994).

Sexual reproduction is the main form of reproduction in almost all animals and plants. Gametes arise from germ cells, which are set aside during embryogenesis. The cells undergo meiosis to differentiate into gametes, which are then stored in somatic gonads. During reproduction, male and female gametes fuse giving rise to an offspring. The developing progeny are a genetic combination of the female and the male parent. Even though this method is more costly, when it comes to energy requirements, it has certain advantages. For instance, the recombination of genetic material from both parents usually results in a strong hybrid, sometimes certain favorable mutations are inserted into the genome, like the adaption to environmental changes and the elimination of harmful mutations (Crow, 1994).

In sexually reproducing organisms, the development of the germline development is essential for their propagation. The germline contains the genetic information which will be passed on to the next generation(Weismann, 1893). Hereditary information can move only from germline to somatic (body) cells and never in reverse (Weismann, 1893). During embryonic development, the cells set aside to make the germline will differentiate into primordial germ cells; these cells will migrate to the future gonadal region. This decision which demarcates the germ line from the soma forms the underlying basis of a fundamental concept, in which the somatic cells die with each generation, whereas the germ line is considered to be 'immortal', passes information from one generation to the next. This separation between the germ line and the soma is apparently never breached, which is known as the Weismann barrier. Hence, the whole germline development has to be regulated and controlled to ensure the development of a fertile adult organism. Any mis-regulation or mutation in the pathway would lead to no offspring being formed and eventually the end of that lineage of the organism.

The initial step of specifying the cells is the most critical period during the germline development and is also very little known about it. Understanding the mechanism of specifying a group of cells to pursue a certain fate will answer how other cell specification processes occur during development of multicellular organisms.

Compared to somatic tissue, very little is known about the critical period of PGC specification. Understanding the biochemical activity of all germ plasm components could help us to grasp, how germ cells get specified. Furthermore, it could identify how "stemness" is achieved at the molecular level.

## 1.1 Mechanisms of germ cell specification
Two different modes of germ cell specification have been described.
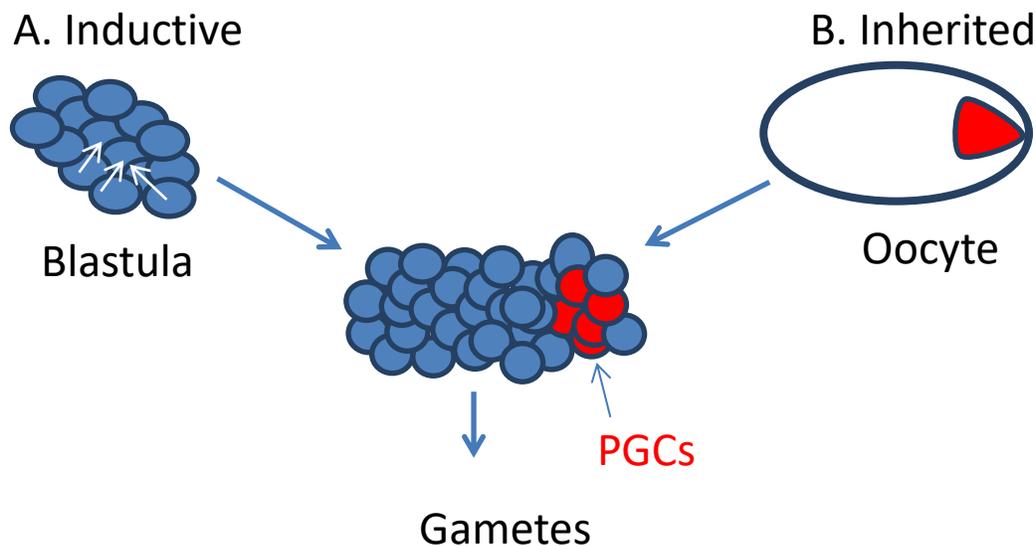


**Figure 1: Inherited vs Inductive mode.**

(A) Inductive mode. Somatic cells induce germ cells (white arrows) within the blastula to express germline factors and differentiate into PGCs (red). (B) Inherited mode. Maternal RNP granules or germ plasm (red) are asymmetrically localized in the oocyte and are inherited by a subset of blastomeres, which specifies PGCs Modified from (Seervai and Wessel, 2013).

A.      Inductive mode

Germ cell specification by induction is often described as the ancestral or more prevalent mode (Fig.1A)(Extavour and Akam, 2003). In the induction mode, germ cell fate is specified through external signals from developing embryonic cells. Induction was described in some invertebrates and in some vertebrates like mammals (Ullmann *et al.*, 1997; Noce, Okamoto-Ito and Tsunekawa, 2001). The most studied example is the mouse(Tam and Zhou, 1996; Toyooka *et al.*, 2000; Costa *et al.*, 2006). One of the signals inducing germ cells is BMP4 (Lawson *et al.*, 1999). However, it is currently not clear how conserved this signal is during germ cell specification in other species.

Regardless whether PGCs are specified by induction or inheritance, they show several commonalities at the molecular level. In most species, numerous proteins and mRNAs like Vasa, Piwi, and Nanos are conserved(Tsuda, 2003; Seydoux and Braun, 2006). In spite of two different modes of specification, they activate common downstream components.

PGCs adopt different lineages, if transplanted to different parts of the embryo(Wylie, 1999). In the mouse, which uses the induction mode, transplanted PGCs later on colocalize with neural plate and surface ectoderm cells(Tam and Zhou, 1996). In Xenopus, which uses the inherited mode, transplanted PGCs generate lineages of the three germ layers(Tada *et al.*, 2012). These results suggest that despite different specification modes, both types of PGCs still require signaling from extrinsic sources to maintain their fate as fully determined PGCs(Bertocchini and Chuva de Sousa Lopes, 2016). Hence even though the two mechanisms seem starkly different, there may be a common underlying signaling mechanism.

The key to understanding the specification of PGCs is to separate species-specific adaptations from a core program of germ cell formation. As information about the initial phase of germ cell specification is still quite fragmentary in different organisms, the core program of germ cell specification is unclear. For instance, the molecule that acts as a master or 'kick starter' for the germ plasm or PGC program appears to be different in each organism.


B.      Inherited mode

Inheritance of cytoplasmic determinants represents the second mode, by which germ cells are specified (Fig 1B). This mechanism of germ cell specification is described amongst others in dipteran insects (*e.g. Drosophila*), nematodes (*e.g. C. elegans*) anuran amphibians (*e.g. Xenopus*), zebrafish, and birds(Tsunekawa *et al.*, 2000; Extavour and Akam, 2003). The molecular mechanisms of germ cell specification are probably better understood at the molecular-genetic and biochemical level than induction, because forward genetics identified most of the known maternal key factors (Schupbach and Wieschaus, 1986a; Hekimi, Boutis and Lakowski, 1995; Pelegri *et al.*, 1999; Dosch *et al.*, 2004). The best studied examples are probably *Drosophila* and C. elegans (Strome and Updike, 2015; Lehmann, 2016).

Table 1: Summarizes selected historical highlights in the context of germ plasm.

| Publication | Finding/Hypothesis |
| --- | --- |
| Weismann (1893) | Inheritance depends on germ cells. Postulates that germ plasm localizes to the nucleus. |
| Hegner (1911), Boveri (1910) | Germline determinants (germ plasm) localize to the cytoplasm. Germ plasm is necessary (Hegner) and sufficient (Boveri) for germline development. |
| Bounoure (1934) | Germ plasm for the first time visualized in a vertebrate egg. |

| | |
|---|---|
| Smith (1966) | UV-irradiation of *Drosophila* eggs reduces the number of PGCs. The UV-wavelength suggest that nucleic acids are critical for germline development. |
| Illmensee and Mahowald (1974) | Ectopic germ plasm is sufficient for PGC formation. |
| Heasman (1984) | The Balbiani body of *Xenopus* contains the germ plasm. |
| Ephrussi and Lehman (1992) | Ectopic expression of a single protein termed Oskar gives rise to functional PGCs in *Drosophila*. |
| Hashimoto (2004) | Ablation of germ plasm in zebrafish reduces PGCs. |
| Bontems (2009) | Ectopic Expression of a single protein termed Bucky ball induces PGCs in zebrafish. |
| Brangwynne (2009) | Biophysical studies on embryonic germ plasm reveal a liquid-like hydrogel in *C*. elegans. |
| Tada (2012) | Germ plasm transplantation in Xenopus induces ectopic germ cells. |
| Boke (2016) | The *Xenopus* Balbiani body forms amyloid aggregates. |

(Weismann, 1893; Hegner, 1911; Boveri, 1910; Bounoure, 1934; Smith, 1966; Illmensee and Mahowald, 1974; Heasman, Quarmby and Wylie, 1984; Ephrussi and Lehmann, 1992; Hashimoto *et al.*, 2004; Bontems *et al.*, 2009; Brangwynne *et al.*, 2009; Tada *et al.*, 2012; Boke *et al.*, 2016)

## 1.2 Germ Plasm

Composition

RNA/protein (RNP) assemblies form the basis of a lot of structures like germ granules, processing bodies, stress granules, and nucleoli. These are key players in the regulation of gene expression(Anderson and Kedersha, 2006). These RNP granules, have roles in RNA processing like splicing, degradation, and translational repression of mRNA. They lack a membrane and appear as puncta in both the cytoplasm and the nucleus under the microscope. The germ plasm is an RNP granule. The germ plasm consists of maternally provided RNAs, proteins, and organelles like mitochondria and endoplasmic reticulum (ER). The entire assembly forms a cytoplasmic structure in the oocyte named Balbiani body(Kloc *et al.*, 2014). Sometimes it is also referred to as the mitochondrial cloud in *Xenopus*(Heasman, Quarmby and Wylie, 1984). The Balbiani body was discovered in spiders and it seems to be omnipresent in oocytes of most species of invertebrates (*e.g.* spiders, insects and molluscs) and vertebrates (*e.g.* frogs, birds, teleosts and mammals)(Guraya, 1979; Kloc *et al.*, 2014).

Organelles

Studies in Xenopus and *Drosophila* suggest that the Balbiani body accumulates a subset of mitochondria. These mitochondria are designated to be delivered to the germ plasm and ultimately to the next generation via primordial germ cells(Marinos and Billett, 1981; Pepling

and Spradling, 1998; Cox and Spradling, 2003)). Interesting experiments in *Drosophila* proposed that germ plasm selects a healthy set of mitochondria by their level of ATP production (Pepling and Spradling, 1998; Cox and Spradling, 2003) The mitochondria in oocytes show high levels of mitochondrial inner membrane potential(Wilding *et al.*, 2001; Tworzydlo *et al.*, 2016). Perhaps this mechanism provides germ cells and by extension gametes with the fittest organelles. This might ensure that the healthiest mitochondria and its descendants are passed on to the next generation.

Function

Inheritance of the germ plasm by cells specifies those cells as PGCs. Germ plasm is an assembly of numerous proteins, mRNAs and other types of small RNAs. Germ plasm components are believed to convey longevity and totipotency. Many components of germ plasm like Vasa and Nanos have important roles in the function of germ cell specification. Piwi another germ plasm component functions in maintaining both germline and somatic stem cells in *Drosophila* (González-Morales *et al.*, 2015). Loss of germ plasm leads to a decrease or no germ cells (Smith, 1966) whereas in gain of function experiments more germ plasm leads to more germ cells in *Drosophila* (Illmensee and Mahowald, 1974)(table1). Loss of the germ plasm component Ziwi (the Piwi counterpart in zebrafish), which is involved in the piRNA pathway, leads to loss of germ cells (Houwing *et al.*, 2007). Hence the function of the germ plasm is attributed to its individual components, which are well regulated.

**1.3 Two molecular components of germ plasm Nanos and Vasa.**

**Vasa, a RNA helicase which is essential for germ cell development.**

Vasa seems to be the most widely used molecular marker to identify germ cells(Lasko and Ashburner, 1988; Raz, 2000; Yajima and Wessel, 2015). The sequence of Vasa is well conserved and required for germline development. Vasa is a member of the DEAD-box protein family of RNA helicase suggesting that it resolves duplex RNA or RNA-protein hybrids. Mutations in Vasa show defects in posterior patterning and in germ cell specification in the *Drosophila* embryo (Lasko and Ashburner, 1988; Tomancak *et al.*, 1998) Vasa mutant zebrafish do not form gametes and develop as sterile males (Hartung, Forbes and Marlow, 2014). Vasa-null male mice are infertile because their germ cell do not proliferate and differentiate (Ikenishi and Tanaka, 2000).The VASA-like gene DBY in humans also appears to be required for male fertility (Foresta, Ferlin and Moro, 2000). In gain of function experiments, ectopic Vasa expression in chicken embryonic stem cells induces expression of specific germline and meiotic genes(Lavial *et al.*, 2009). When these cells are transplanted into chick embryos, they migrate to the gonad anlagen and differentiate into gametes. Vasa is also involved in piRNA processing. The amount

of Vasa in the germ plasm therefore prevents the degradation of the germ cell genome by transposon activity, but piRNAs could also play an undiscovered role in early germ cells(Pek, Patil and Kai, 2012). Aubergine, a well-known component of the piRNA pathway is needed for Osk translation, which also needs Vasa to localize. This could indicate a feedback mechanism ensuring all the downstream germ plasm members are expressed(Harris and Macdonald, 2001). Overall these results support the theory that Vasa has a central role in establishing germ line identity and function, though the precise signaling pathways it initiates is mechanism still not known.

Vasa RNA or protein expression is frequently used to label PGCs in animals. As at least one homolog seems to be present in all metazoans, Vasa is also an easily accessible marker across the animal kingdom(Van Doren *et al.*, 1998). However, the restriction of Vasa at the blastula stage to the germ plasm and prospective PGCs varies across species. In some species like the zebrafish, Vasa protein is ubiquitous and later gets restricted into PGCs(Knaut *et al.*, 2000), which raised concerns about the role of Vasa during germ cell specification.

**Nanos** (Nos)**, a RNA-binding protein and repressor which is needed for primordial germ cell maintenance.**

Nanos codes for an RNA-binding protein which contains a zinc finger domain. This is highly conserved among metazoans (Mosquera *et al.*, 1993; Curtis, Apfeld and Lehmann, 1995). Nanos protein has also been shown to be involved in translational repression in early embryonic patterning in *Drosophila*(Gavis and Lehmann, 1994). Studies in *Drosophila*, *C. elegans* and mice indicate that *nanos* genes play a role in survival, migration and chromatin remodeling of primordial germ cells(Nakamura and Seydoux, 2008). Germ cells, which show a decrease in Nanos activity, show a precocious expression of the other germ cell genes and this leads to migration defects and loss of functional germ cells(Kobayashi *et al.*, 1996). In zebrafish knockdown of the *nanos* homolog nanos3 by morpholino, showed a loss of proper migration and survival of PGCs (Koprunner *et al.*, 2001) suggesting that Nanos is crucial for migration and survival of germ cells.

Nanos3 mRNA is specifically localized to the germ plasm in zebrafish (Koprunner et al., 2001) .The signal for the specific localization of *nanos3* mRNA is present in its 3'UTR. The nanos3 3'UTR has been shown to be sufficient to protect gfp mRNA from degradation in primordial germ cells, whereas it is degraded in the soma (Koprunner et al., 2001). Even in *Drosophila*, localization of *nanos* mRNA is controlled by sequences within its own 3'UTR in *Drosophila* (Gavis and Lehmann, 1994). This signal in the 3'UTR of *nanos 3* has been used as a tool mark primordial germ cells, as well as target injected mRNA into the germ cells. Maternally provided nanos3 mRNA is localized similarly to vasa transcripts, which indicates that nanos3 mRNA is indeed a component of the germ plasm (Koprunner *et al.*, 2001; Kosaka *et al.*, 2007).

## 1.4 Oskar, a germ plasm organizer in *Drosophila*

Oskar protein acts as a master regulator of germ plasm assembly(Ephrussi and Lehmann, 1992). In *Drosophila*, germ plasm is localized to the posterior pole during late oogenesis and hence, also known as pole plasm (Fig. 2). *oskar* was isolated in mutagenesis screens for maternal-effect genes required for embryonic patterning (Lehmann and Nuesslein-Volhard, 1986). Oskar mutants showed posterior patterning defects and no germ cells. *osk* RNA localizes to the posterior pole, where the protein gets translated and starts the assembly of germ plasm(Ephrussi, Dickinson and Lehmann, 1991; Kim-Ha, Smith and Macdonald, 1991). This shows that Oskar is the initiator protein to form germ plasm and by extension germ cells.

| A | *oskar* mutant | B | wildtype | C | *oskar* overexpression |
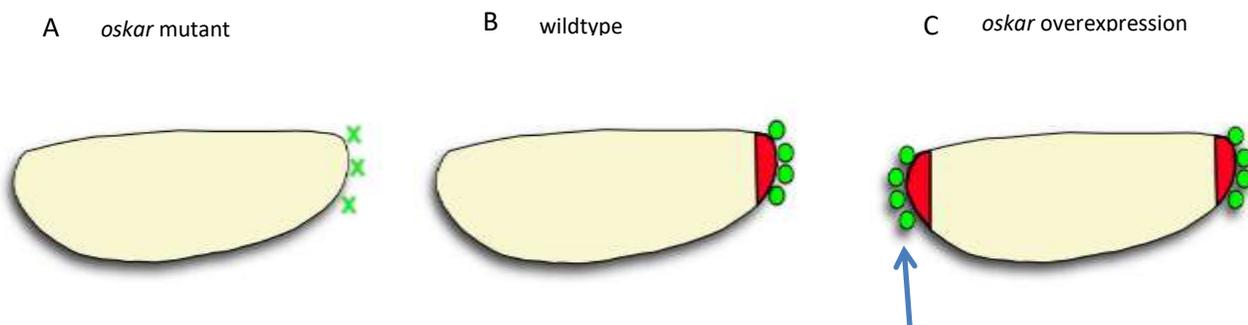


**Figure 2: The role of Oskar in loss- and gain-of-function experiments**

*Drosophila* embryo at stage 5, lateral view, anterior to the left (A) *oskar* mutant showing loss of germ cells (green x). (B) Wildtype embryo showing germ cells in green. (C) Oskar overexpressed at the anterior end showing ectopic germ cells (arrow blue).(modified from (Ephrussi and Lehmann, 1992)).

Mislocalization of Oskar protein at the anterior end of the embryo leads to ectopic germ cells as seen in Fig 2 (Ephrussi & Lehmann, 1992). Oskar was the first protein, which is both necessary and sufficient to assemble germ plasm. Increasing the amount of Oskar protein in the fly embryo causes increase in activity of Nos protein. Thus, the amount of Osk protein and the level of Nos protein accumulation are related. Mutations in *oskar* affect the enrichment of other RNAs and proteins at the posterior pole, which are present in the germ plasm. Mutations in *nos*, infact result in an increase in germ cells (Kim-Ha, Smith and Macdonald, 1991). Possibly the heightened expression of Nos represses the somatic cell fate pushing the cell to a germ cell lineage(Smith, Wilson and Macdonald, 1992). Using a dosage of 6x *osk*, it shows the presence of germ cells in the background of mutations in maternal genes like *staufen* and *cappuccino*. However, the high dosage of Oskar could not rescue germ cell formation in *vasa* mutants(Schupbach and Wieschaus, 1986b). This indicates that Oskar acts through Vasa for specifying germ cells. Such an activity supports the role of Oskar as a master regulator of PGC specification in invertebrates.

*osk* mRNA is translated into two protein isoforms by alternative translation initiation (Markussen *et al.*, 1995; Rongo, Gavis and Lehmann, 1995). Long Osk (lOsk) is translated at the first start codon and encodes a protein of 606 amino acids. LOsk mainly anchors germ plasm at the posterior end(Suyama *et al.*, 2009). Long Oskar also traps and accumulates mitochondria at the posterior at the site of PGC formation. Mutating this long oskar form strongly decreases the number of mtDNA molecules inherited by PGCs (Hurd *et al.*, 2016). Short Oskar (sOsk) starts at Methionine 139 and encodes a protein of 467 amino acids (Markussen *et al.*, 1995; Rongo, Gavis and Lehmann, 1995). sOsk assembles germ plasm and thereby plays a critical role to specify PGCs.

A

| lOsk | NTE | E-LOTUS | | OSK | 606 aa |

| sOsk | E-LOTUS | | OSK | 467 aa |
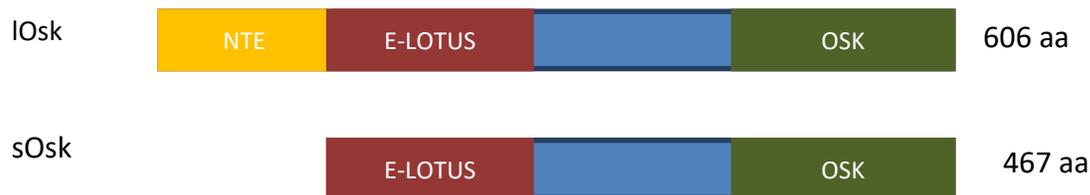
B

**Table 2: Stating the differences between long and short Oskar**

**Stating the differences between long and short Oskar.**

| Long Oskar | Short Oskar |
|---|---|
| 606 amino acids long | 467 amino acids long |
| Anchoring germ plasm | Assembling germ plasm |
| Associated with endosomes | Associated with RNA granules |
| Interacts with Lasp to be tethered to posterior pole | Interacts with Lasp to be tethered to posterior pole |
| Not essential for patterning and germ cell formation | Necessary for germ cell formation and posterior patterning |

(Tanaka, Kato, Matsuda, Hanyu-Nakamura and Nakamura 2011, Suyama et al., 2009, Markussen et al., 1995; Rongo et al., 1995; Vanzo & Ephrussi, 2002 Tanaka & Nakamura, 2008)

**Figure 3: Oskar isoforms in Drosophila**

(A) Comparison of long (lOsk) and short (sOsk) Oskar proteins. The NTE domain in lOs inhibits Vasa-interaction and RNA-binding(Rongo *et al.*, 1997; Vanzo and Ephrussi, 2002; Lehmann, 2016).The eLOTUS (extended LOTUS) domain consists of the minimal LOTUS domain along with a short disordered region of 18aa, which are essential to bind Vasa. The OSK domain binds to RNA(Jeske *et al.*, 2015; Yang *et al.*, 2015) (B) A table listing the differences between long and short oskar.

## 1.5 Germ cell specification by Oskar

Fascinating insight into sOsk function was recently gathered by crystallizing two of its domains. A domain at the N-terminus of sOsk (139–240aa) termed LOTUS domain was previously predicted to be involved in RNA-binding. The second structure described the C-terminal "OSK" domain, which resembles a SGNH hydrolase (Jeske *et al.*, 2015; Yang *et al.*, 2015) (Fig. 3). However, looking carefully at the biochemical interactions and crystallizing sOsk with these binding partners revealed some unexpected information.

sOsk directly interacts with Vasa(Breitwieser *et al.*, 1996), which is an ATP-dependent helicase(Jeske *et al.*, 2015; Jeske, Müller and Ephrussi, 2017). Interesting biochemical and biophysical studies show that the eLOTUS domain of Oskar does not interact with RNA, but in fact binds to the RNA helicase Vasa. Surprisingly the extension of the LOTUS domain (eLOTUS) encodes an intrinsically disordered motif, which forms a structured domain upon Vasa binding. This stretch of 18 amino acids outside of the LOTUS domain is essential for the Vasa interaction. Moreover, binding the eLOTUS domain increases the ATPase activity of Vasa. This is the first time an instructive role was assigned to Oskar, which was previously regarded as a scaffold protein aggregating germ plasm components within the Drosophila oocyte(Jeske, Müller and Ephrussi, 2017). Taken together, Osk is a Vasa activator.

The OSK domain shows a lot of similarity to a SGNH hydrolase, but lacks three of the four residues of the SGNH motif, as well as the serine triad to be an active hydrolase(Jeske *et al.*, 2015). The C-terminal OSK-domain forms a globular structure, which carries several basic, positively charged residues at its surface suggesting it could interact with nucleic acids. Indeed, this domain binds in *in vitro* experiments mRNAs like the osk and nos 3'UTRs(Yang *et al.*, 2015). When the basic residues of the OSK domain are mutated, binding to RNA is disrupted (Yang et al., 2015). In vivo pull-down experiments after UV-crosslinking suggest that Osk interacts with nos, pgc and gcl mRNA in vivo(Jeske *et al.*, 2015). All three RNAs are known to be localized to the germ plasm. Again, these exciting discoveries identify sOsk as a novel RNA-binding protein and suggest a more instructive role of in germline development than previously anticipated.

Taken the interaction data of sOsk together, a modified picture of germ cell specification emerges. sOsk initiates the assembly of germ plasm by binding to Vasa and mRNA. This interaction activates Vasa and might sterically bring it in proximity with specific RNA(s). This could regulate translation or stability of the RNA(s) involved in specifying PGCs (Smith, Wilson and Macdonald, 1992). Hence Vasa and Osk seem to act in a cooperative manner to specify germ cells. Figuring out the biochemical process, which is initiated by sOsk/Vasa is probably key to understand the molecular mechanism of the germ cell specification program.

## 1.6 Germ cell specification in zebrafish

Compared to invertebrates such as *Drosophila* and C. elegans, much less is known about the molecular processes occurring in the germ plasm of vertebrates. In Xenopus, germ plasm research is mostly focused on processes during oogenesis(Machado *et al.*, 2005; Kosaka *et al.*, 2007; Nijjar and Woodland, 2013). Among vertebrates that specify their germ cells through inheritance of germ plasm, there are a number of studies on the zebrafish. Zebrafish combines a number of features for early developmental studies. Among these features the following are helpful for germline specification research. Embryos and oocytes are easily accessible and available in high numbers, which is beneficial for biochemical experiments. The embryo develops outside the mother, which is advantageous because the cellular movements occurring during early and late embryogenesis can be studied in detail and imaged. Moreover, its transparent embryos enable tracing of fluorescently-tagged proteins in vivo and allow detection of endogenous proteins by immunostaining. The genome is completely sequenced, and genomic manipulations via CRISPR/Cas9 are easy. Therefore, zebrafish as a vertebrate model is very well suited for the analysis of germ cell specification.

## 1.7 Germline development in Zebrafish.
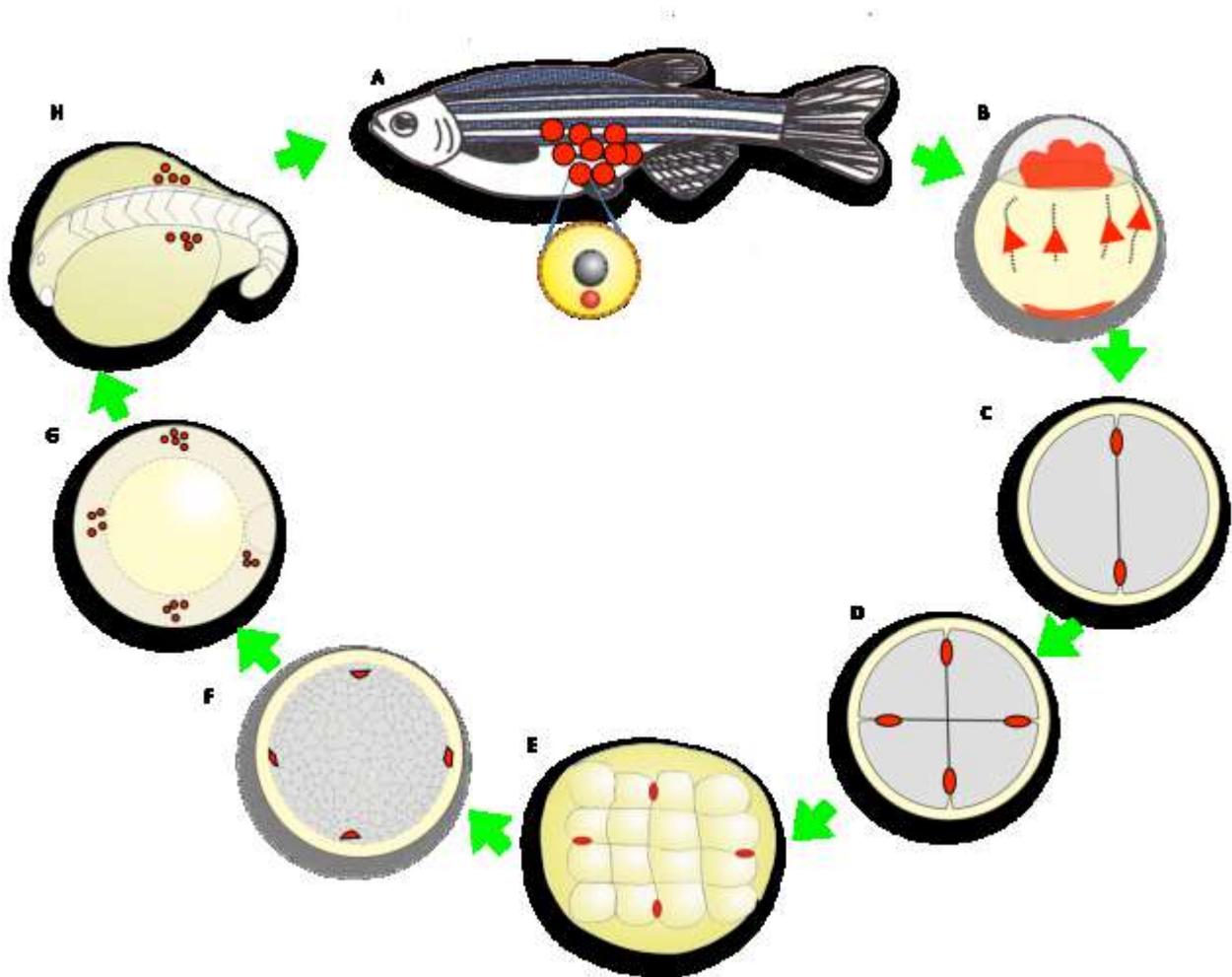


**Figure 4: Key stages of germline development in zebrafish.**

(A) Zebrafish female containing oocytes (red circles). Magnification of Stage I oocyte with the Balbiani body containing germ plasm (red) vegetal to the nucleus. (B) Wildtype zebrafish embryo after fertilization. Germ plasm components stream (red triangles) into the animal blastodisc. (C) At the two-cell stage (animal view) embryos localize germ plasm to two spots on either side of the first cleavage furrow. (D) At the four-cell stage, germ plasm localizes to the four cleavage furrows and no further germ plasm aggregates are formed in the later stages. (E) 16 cell stage embryo (1.5h post-fertilization (hpf)) which is the first stage with a clear separation between prospective somatic cells and prospective germ cells. (F) In the 3hpf stage embryo, four populations of germ cells get specified. (G) PGC clusters have formed, which then start to migrate. (H) At the 15-somite stage (16.5 hpf), PGCs arrived at the prospective gonad, where they will differentiate during sexual maturation into sperm in males or oocytes in females (A). (Dosch, 2015).

Germ cell specification in the fish is with the inherited strategy. Germ plasm put in by the mother into her oocyte in Fig 4A gets streamed up into the embryo after fertilization into the animal cell layer (4B). As embryogenesis occurs four clusters of germ plasm form which get inherited into cells thus becoming primordial germ cells (PGCs) (4C-4G). Once specified, they migrate to the

future gonadal region and differentiate into gametes (4G and 4H). If the sex is female, the gametes develop into oocytes once again forming germ plasm in their vegetal pole.

## 1.8  Bucky ball, germ plasm organizer in Zebrafish

Maternal-effect inheritance is the kind of heritability in which the maternal genotype controls the embryonic phenotype. It was first studied in the molecular regulation of embryonic development of snails. In the snail Limnaea, maternal factors provided by the mother controlled the sinistral or dextral turning of the shell(Sturtevant, 1923; Diver, Boycott and Garstang, 1925).

To identify maternal factors controlling early vertebrate development, maternal-effect mutant screens were carried out in zebrafish (Pelegri *et al.*, 1999; Bauer and Goetz, 2001; Dosch *et al.*, 2004; Kishimoto *et al.*, 2004; Pelegri and Mullins, 2016). Among the mutant mothers, one of the mutants gave embryos in which it showed a defect in processes prior to midblastula transition (MBT). After fertilization it showed a radial segregation of cytoplasm. By contrast, in the wild type cytoplasm streamed to the animal pole (Fig 4B and 5A). In addition, the fertilized mutant embryo did not show cellular cleavages and hence does not develop beyond the 1-cell stage. As the mutant embryo lacks polarity similar to Buckminsterfullerenes, it was referred to as *bucky ball* (*buc*) as seen in Fig 5B (Dosch *et al.*, 2004).
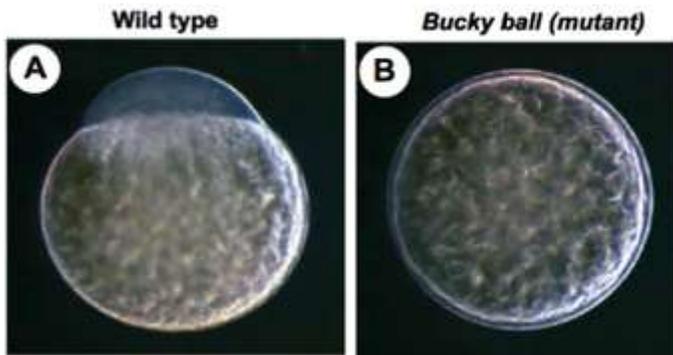


**Figure 5: Phenotype of Bucky ball mutant.**

Phenotype of Bucky ball mutant. One-cell stage zebrafish embryo in lateral view, animal to the top (A) Wildtype (B) Buc mutant shows loss of polarity and uniform radial distribution of cytoplasm. Development of *buc* mutant stops after the 1-cell stage and shows no cell cleavage (Dosch et al 2004).

In the oocyte, Buc mutants did not assemble germ plasm into a Balbiani body (Bb) (Fig 6b and 7A). Instead, germ plasm components like nanos and vasa mRNA are no longer localized to the Balbiani body, but rather distributed ubiquitously in the ooplasm (Bontems *et al.*, 2009). This result described the first gene in vertebrates required for the formation of the Balbiani body and the localization of germ plasm components in the oocyte. Moreover, if the cDNA of Buc is ectopically expressed from a transgene, ectopic Bbs are seen (Fig 7C)(Heim *et al.*, 2014). This leads to the conclusion that Buc, similar to sOsk in *Drosophila*, is necessary and sufficient for germ plasm assembly.
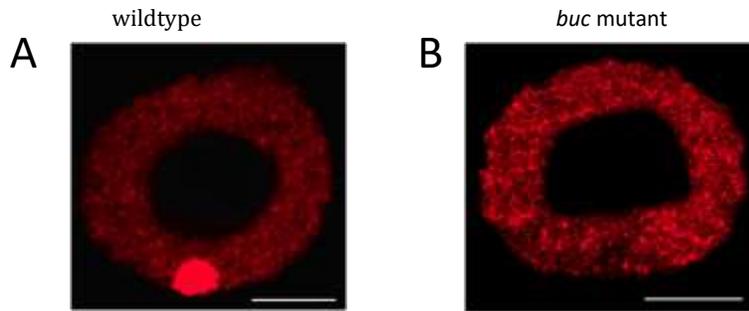
**Figure 6: Buc is required for Balbiani body formation.**

Stage I oocyte lateral view. Fluorescent in-situ hybridization showed in red for *nanos* a germ plasm marker. (A) The wild type oocyte showing an assembled germ plasm, whereas in (B) the *buc* mutant the germ plasm is disrupted(Bontems *et al.*, 2009). Scale bar is 500 μm.
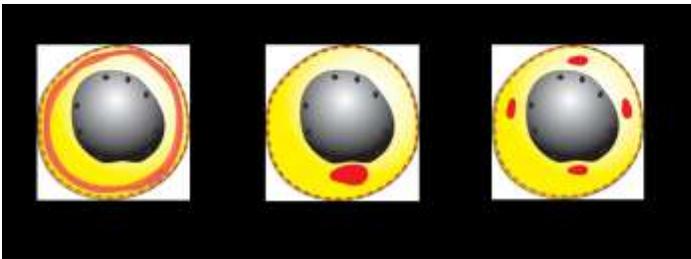


**Figure 7: Buc is necessary and sufficient for germ plasm assembly during zebrafish oogenesis.**

Cartoon of stage I oocytes. (A) In *buc* mutant oocytes *(buc-)*, germ plasm assembly is disrupted, and Balbiani body components are ubiquitously distributed in the oocyte (red haze) (Bontems et al 2009). (B) Wild-type stage I zebrafish oocyte, the central nucleus (germinal vesicle; grey), germ plasm/Balbiani body (red). (C) A transgene with the Buc cDNA is over-expressed, which leads to the ectopic formation of multiple Bbs (red) (Bontems *et al.*, 2009; Heim *et al.*, 2014)

Interestingly, in the embryo *buc* also mimics germ plasm activity Fig. 8 *i.e.* Overexpression reprograms somatic cells into germ cells(Bontems *et al.*, 2009). These results identify Buc as a crucial regulator of germ cell formation in vertebrates.
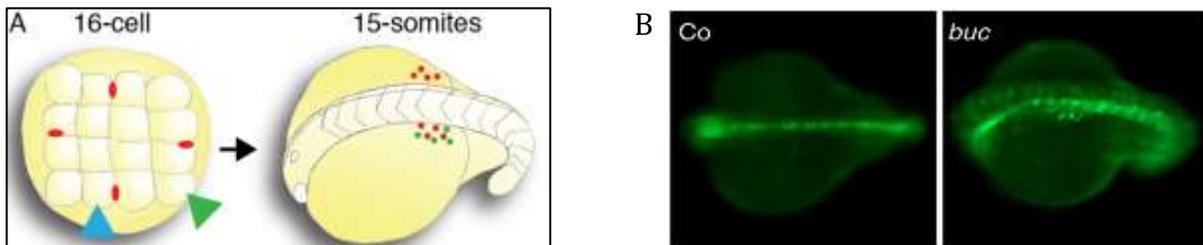


**Figure 8: Scheme and results of germ cell induction assay (16 cell assay).**

Scheme and results of germ cell induction assay (16 cell assay) (A) Scheme of germ cell induction assay.  Left:  16-cell embryo, animal view injected with a PGC-reporter (*GFP-nos3'UTR*) into a prospective germ cell (middle) (blue arrowhead) containing endogenous germ plasm (red dots) or a somatic cell (corner blastomere; green arrowhead). Right: Dorso-lateral view of a 15-somite embryo (18 hours post fertilization, hpf), anterior to the left. After injection, a fraction of the germ cells is fluorescent (green dots), if the reporter was targeted to a germ cell or transformed a somatic into a germ cell.  Otherwise only the neural tube shows some background fluorescence. The injection into the middle and corner blastomere is with the germ cell report eGFP3'nos UTR which is stabilized only in germ cells. As the corner blastomere is not fated to be a germ cell, there is low fluorescence and endogenous PGCs are not visible (red cells in A). Somatic cells are reprogrammed into ectopic germ cells upon Buc co-injected with the reporter. (B) Left panel shows an embryo injected with the germ cell reporter, right panel shows ectopic fluorescent germ cell on co-injection of Buc and the reporter. (Bontems et al 2009)

## 1.9 The conservation of Buc across the vertebrate kingdom.

Buc is present vertebrates but even across its homologs, the sequence changes quite rapidly (Bontems *et al.*, 2009). Zebrafish has two paralogs of Buc in its genome, whereas the salmon has three (Škugor *et al.*, 2016). Currently, the function of the other paralogs is not clear. In Fig 9A the Buc protein in the zebrafish is very dissimilar to the Buc protein in mammals. However, the synteny seemed to be conserved (Bontems *et al.*, 2009)(data not shown). The Xenopus Buc homolog Xvelo exists in two splice forms, long Xvelo and short Xvelo. Both seem to play redundant roles in maintaining germ plasm assembly (Nijjar and Woodland, 2013). In humans, two genetic loci show homology to Buc protein (Gene ID EU128483, EU128484) (Bontems *et al.*, 2009), but the sequence is interrupted by STOP-codons and hence, does not encode an open reading frame. Human ovaries show RNA expression from these loci, but their function is not known (Lyautey et al., unpublished). BUC might act as a noncoding RNA or encode a short peptide(Bontems *et al.*, 2009; Chew, Pauli and Schier, 2016). Currently, Buc is a novel protein whose protein sequence shows no homology to other known functional domains and hence, its biochemical activity is currently unknown (Fig 9B). Nevertheless, the N-terminal 100 amino acids, termed BUVE-motif, are conserved among Buc vertebrate homologs (Fig 9B). By aligning Buc with its homologs across other vertebrate species certain domains could be seen emerging. However no information about their function could be gathered from it.

**A**

**B**

Conc. N-Terminus
BUVE-motif
Highly-cons. Domain
Positive charged domain

Bucky Ball, 639 aa
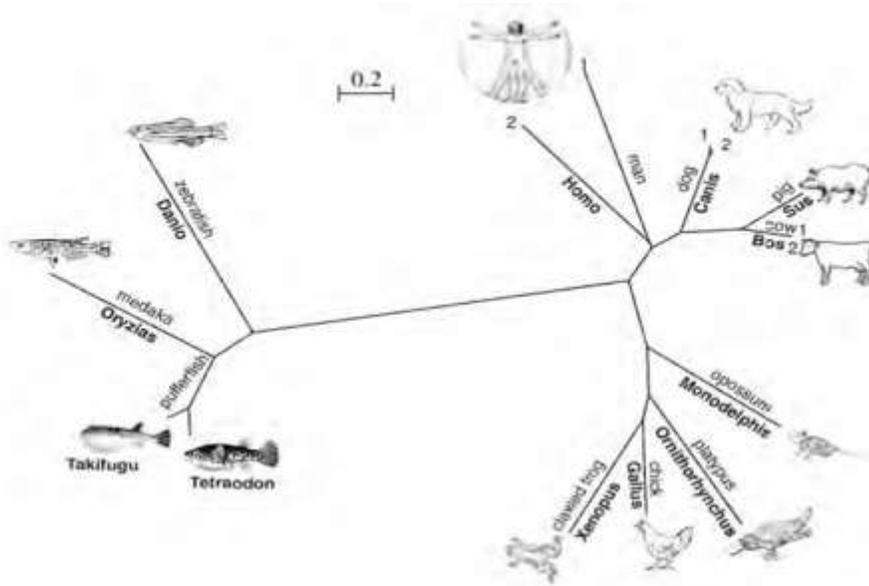
**Figure 9: Sequence conservation in Buc protein across vertebrates.**

(A) Unrooted star diagram displaying the phylogenetic conservation of Bucky ball proteins among vertebrates, The scale indicates the number of substitutions per amino acid residue. (Bontems et al 2009). The scale indicates the number of substitutions per amino acid residue. (B) The predicted domains from comparing Buc across species of the animal kingdom (Riemer, 2014).

## 1.10 *buc* and its mutant alleles

Two mutant alleles were isolated in the matenal-effect mutagenesis screens in zebrafish. The alleles buc[p106] and buc[p43] contain nonsense mutations that result in premature stop codons see Fig 10(Bontems *et al.*, 2009). Recent studies indicated that buc[p106] as well as buc[p43] mutant oocytes showed no localization of Buc. However, Vasa was still localized to the perinuclear region. The immunostainings for Buc also showed the loss of Buc protein expression in the mutant oocytes(Riemer *et al.*, 2015). This showed that the two mutant alleles are probably null alleles.
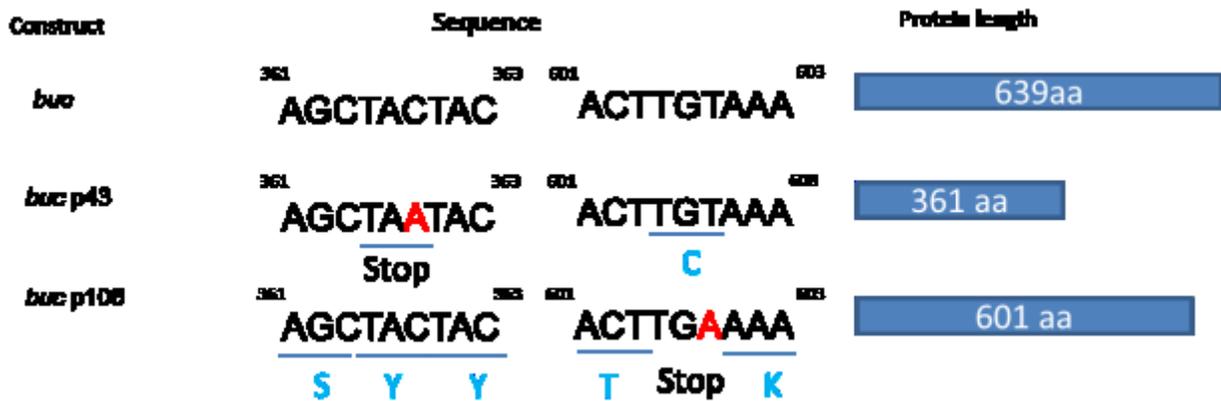


**Figure 10: Buc and its mutant alleles.**

Numbers indicate nucleotide positions. Red nucleotides indicate position of termination. Light blue letters indicate amino acids.

## 1.11 Transgenic BucGFP continuously localizes with germ plasm.

In order to study the localization of Buc in the zebrafish embryo, a transgenic zebrafish *buc-gfp* line was created which expressed one copy of *buc-gfp* in the *bucp106* homozygous mutant background Fig 11A(Bontems *et al.*, 2009). This line was further crossed to make a stable line with two copies of *buc-gfp*. Hence, in this transgenic line, *buc-gfp* rescues the mutant phenotype, showing that the transgene indeed mirrors the activity of endogenous Buc.
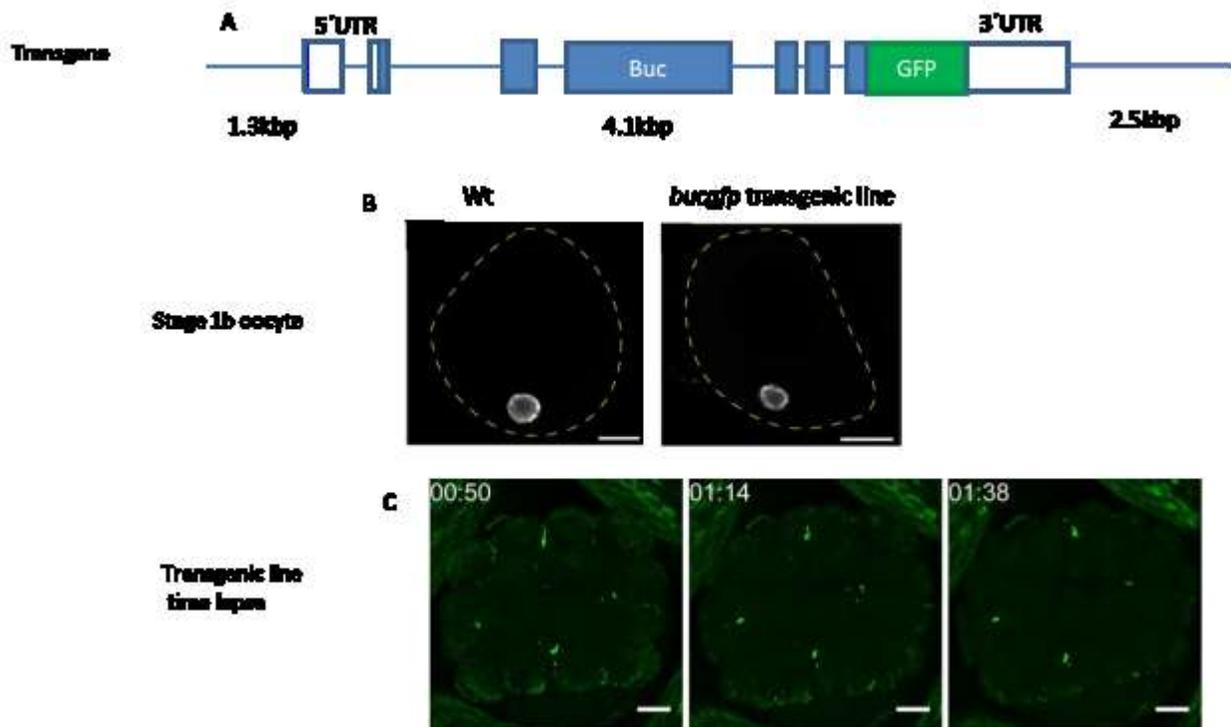
**Figure 11: Construction and characterization of BucGFP transgenic line**

(A) The transgene inserted into the transgenic line. (B) (Below) Stage 1b oocyte Wildtype stained with Buc antibody,(right) GFP showing Buc in stage matched oocyte, scale bar is 10µm. (C) Pictures of a time lapse movie showing a transgenic buc-gfp embryo. Scale bar 100µm. from ((Riemer, 2014; Riemer *et al.*, 2015).

In wild type stage IB oocytes and in stage matched transgenic Buc-GFP, Buc was localized to the Balbiani body near the vegetal pole of the oocyte (Fig 11B)(Riemer *et al.*, 2015). This line was further imaged during early embryonic development (Fig 11C). During early and late embryonic development, the Buc GFP protein localized with the germ plasm. Hence, demonstrating this transgenic *buc-gfp* line is the first transgenic line that marks the germ plasm *in vivo* throughout embryogenesis (Riemer *et al.*, 2015).

## 1.12 Low complexity Proteins

Both Buc and sOsk have been suggested to have low complexity regions (Toretsky and Wright, 2014; Jeske *et al.*, 2015; Boke *et al.*, 2016)). Low complexity (LC) proteins are of two types, Amyloid and intrinsically disordered proteins (IDPs) (Dyson and Wright, 2005; Oldfield and Dunker, 2014; Toretsky and Wright, 2014).

Table 3: Compares the differences between the two types of LC-proteins.

| Properties | Amyloid | IDP |
|---|---|---|
| Structure | Low complexity regions form beta sheets | Very low complexity with FG or FXXG repeats, in most cases with no secondary structure formation. |
| Chemical | Aggregates are resistant to SDS and high salt concentrations. | Aggregates are dissolved by SDS or high salt concentrations. |
| Aggregation | Aggregates are resistant to 1,6 hexanediol | 1,6-hexanediol dissolves hydrogels formed by IDPs. |
| Staining | Stain positively with Thioflavin S and T. | No accumulation of Thioflavin |
| | Examples: Amyloid plaques, Balbiani body Xvelo protein in Xenopus(Boke et al., 2016) | Examples: Nuclear pores (Nucleoporins(Frey, Richter and Gorlich, 2006)), germ plasm in C. elegans |

 (Frey, Richter and Gorlich, 2006; Brangwynne *et al.*, 2009; Coletta *et al.*, 2010; Kroschwald *et al.*, 2015; Boke *et al.*, 2016)

Both Buc and sOsk have been suggested to have low complexity regions (Toretsky and Wright, 2014; Jeske *et al.*, 2015; Boke *et al.*, 2016). Indeed, it was shown that sOsk contains an intrinsically disordered region critical for Vasa binding. By contrast, in Buc and Velo1, it was shown that parts of the conserved BUVE-motif form prions or amyloid-like aggregates.

IDPs are also known to act as hubs for supra-molecular complexes and also more prevalent in RNA-binding proteins. Moreover, IDPs form liquid-liquid phase separations such as RNA-granules, which was also described for the germ plasm in C. elegans (Brangwynne *et al.*, 2009).

Some evidence was provided by in vivo imaging of germ plasm in zebrafish (Riemer *et al.*, 2015) and *Drosophila* (Sinsimer *et al.*, 2013) that germ plasm is liquid. Nonetheless, the level of intrinsic disorder of germ plasm organizers and the liquid properties of germ plasm in fly and fish are still not clear. It is presently unknown how the protein components like Oskar, Vasa, assemble into a germ granule aggregate. RNA-binding proteins have been shown to undergo phase transitions from a soluble to viscous state (Brangwynne, 2013; Lin *et al.*, 2015; Zhang, Cao and Liu, 2015). Thus, RNAs may be trapped by germ plasm aggregates, which become a granule and thereby facilitate more RNA–RNA and RNA-protein interactions. Oskar has been suggested to contain disordered regions, which link the domains that were crystallized. These regions could push for the propensity to form aggregates as well.
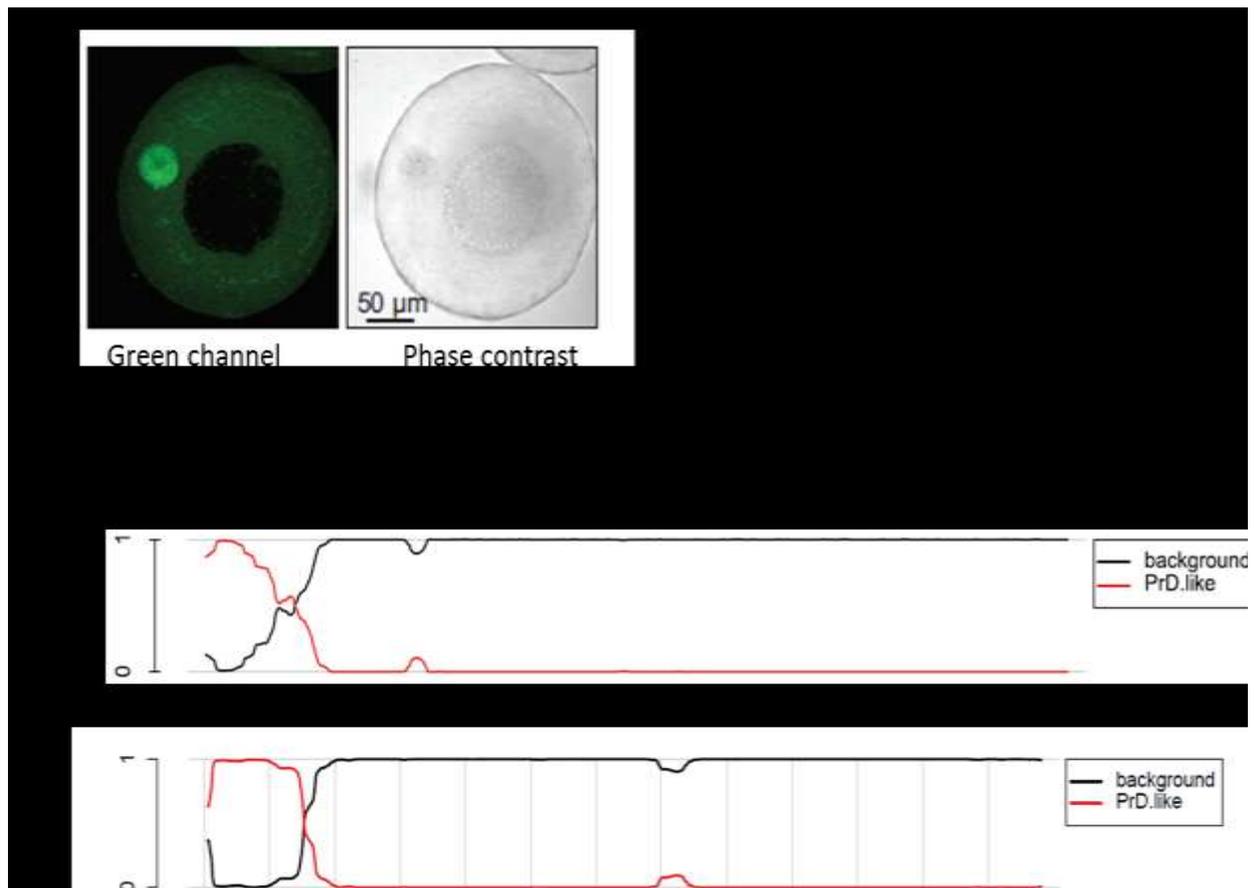


**Figure 12: Amyloid nature of Xvelo and Buc.**

(A) Stage I *Xenopus laevis* oocytes stained with Thioflavin T which stains the prion (amyloid) domain in Xvelo, the Buc homolog in frogs. (B) Predicted prion (amyloid) domain using prion domain predictor software PLAAC in Xvelo and Buc protein. Red peaks indicate presence of prion domain (PrD like) in the N terminal region of both proteins.(from (Boke *et al.*, 2016)). The x axis indicates the region of the protein.

Interestingly, Buc has been discussed to have both amyloid (Fig 12B) and IDP regions. In Xenopus, the Buc homolog Velo1 aggregates into an amyloid like assembly forming the Balbiani

body (Boke *et al.*, 2016) (Fig 12A). Both Buc and Xvelo were predicted to have prion like domains in the N terminal region (Fig 10B). Prion like domains indicated the presence of amyloid regions. Staining of Buc protein with Thioflavin T, which was expressed in the frog oocytes, confirmed the presence of amyloid regions within it (Boke *et al* 2016). By contrast, BucGFP molecules showed hydrogel or liquid droplet-like behavior in the early zebrafish similar to the P-granules in C. elegans (Brangwynne *et al.*, 2009; Riemer *et al.*, 2015). This suggests a controlled transition from an amyloid plaque to a soluble hydrogel at the end of oogenesis. If Buc indeed does form a hydrogel, it would mirror the biophysical similarity to Osk which is discussed to be an IDP.

# Aims and Rational

As mentioned before, the Buc protein sequence shows no homology to other known functional domains, hence its biochemical activity is unknown. Therefore, the molecular process, initiated by Buc to specify germ cell development, is an unanswered question. The typical approach to analyze the biochemical function of a novel protein is to dissect its functional domains in a biological assay and then search for interactors. In contrast to this blind, but safe approach, I planned to take advantage of knowledge about PGCs specification in other organisms, especially in Drosophila.

Based on the Buc overexpression experiment in the zebrafish, which results in ectopic germ cells it bears a highly reminiscent effect of an experiment in *Drosophila*, where Oskar acts as a induces extra germ cells in the fly (Ephrussi and Lehman., 1992). Thus, Oskar shows the same unique activity in invertebrates as Bucky ball in the zebrafish.

Therefore by studying and comparing the Osk protein with Buc it would be possible:

1) To test **the *Drosophila* Oskar function in the Zebrafish** and look for ectopic germ cells. Injecting Oskar into the zebrafish would provide the answer whether Buc and Osk have the same biological role.

2) Study the nature by which Osk and Buc perform the same biochemical function by looking at sequence \biophysical similarities and common interactions partners.
-Looking for common sequences and biophysical **similarities between Osk\Buc**.

3) **Comparing known interactors** of Osk and experimentally determined ones for Buc.
Both are complementary approaches and would answer the question whether there is a functional conservation between the two proteins.

# 2. Results

## 1. Scheme and rationale behind the germ cell induction assay.

The germ cell induction assay, which exploits that the first somatic cells in zebrafish segregate from the germline at the 16-cell stage (Fig. 13) was used to test for germ cell induction in the zebrafish

Injecting the gfp-nos3'UTR reporter (Koprunner *et al.*, 2001) into the middle blastomere (Fig. 13A) labels some endogenous germ cells (Fig. 13B). This reporter highlights the germ cells arising from the lineage of that blastomere as shown in Fig 13A´ (green germ cells among red germ cells). Injecting the reporter into the corner blastomere, which gives rise to somatic cells, leads to reporter degradation and no fluorescence is seen (Fig 13B&B´) (presence of only red germ cells). Injecting the mRNA of interest together with the reporter into the corner blastomere was to test the ability of the encoded protein to induce germ cells. If the mRNA is sufficient to give rise to ectopic germ cells it will be highlighted by expression of the reporter, which labels ectopic green germ cells among invisible, endogenous germ cells (red) (Bontems *et al.*, 2009)
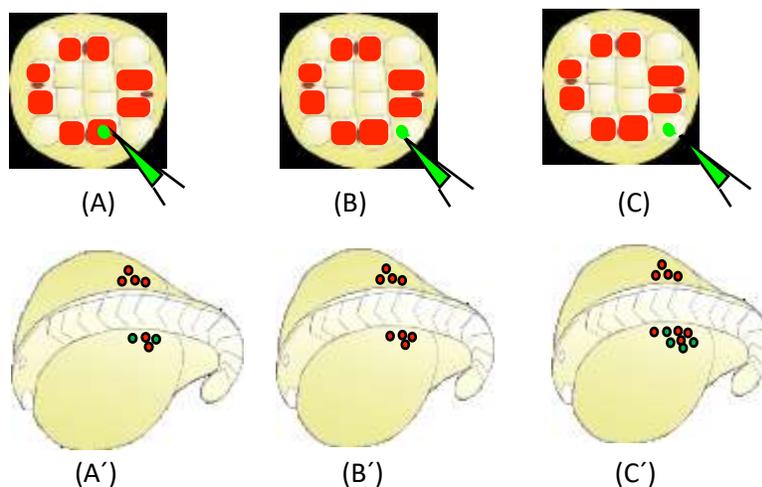


**Figure 13: Scheme of germ cell induction assay.**

Left panel:  Animal view of a 16-cell embryo injected with PGC-reporter (green injection) into a middle blastomere (A) containing endogenous germ plasm (red dots) or a somatic cell (corner blastomere B & C).Oblique, dorsal view of a 15-somite stage embryo (18 hours post fertilization, hpf), anterior to the left.  Fluorescent germ cells (green dots) emerge by targeting the reporter to a PGC (13A´)or transforming a somatic into a PGC (13C´). (Modified from (Bontems *et al.*, 2009))

## 2.1 Buc protein and not its RNA induce germ cells in the Zebrafish.
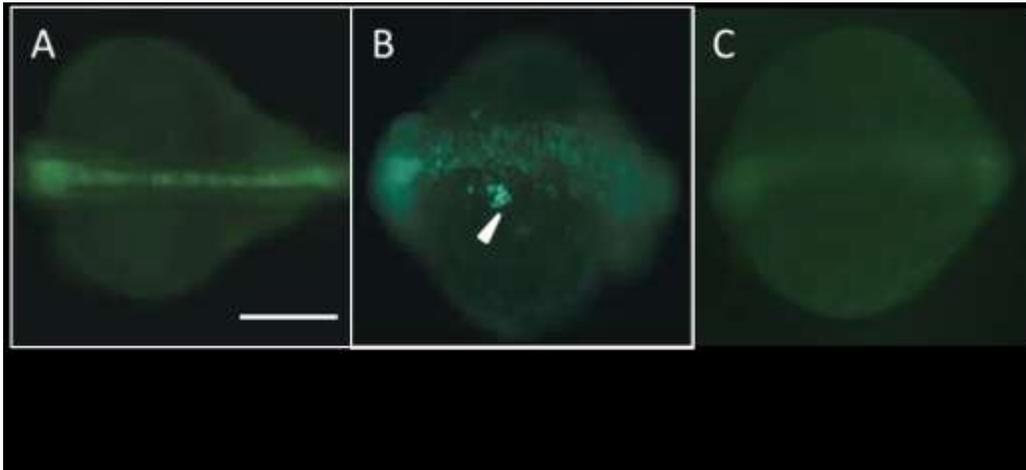


**Figure 14:Buc but not its mutant allele induces ectopic germ cells.**

Fluorescent germ cells (white arrowhead) emerge by targeting the reporter to a PGC or transforming a somatic cell into a PGC. (A) Live 18 hpf embryo after injection of PGC-reporter into a corner blastomere (A; 31±4.5%; n=70;). (B) Wild-type buc overexpression in the corner blastomere reprograms the blastomere to the germline (B; wt= 78±2.6%; n=71). (C) Mutant Buc (bucp43 mRNA) sequence is identical to wt, apart from a point mutation (Y362STOP) (Bontems F, *et al*. (2009)) (C; mut=35±1.6%; n=68). Scale bar= 200μm.

It has been shown that *bucky ball* has the ability to induce extra germ cells in the Zebrafish (Fig 14B, (Bontems *et al.*, 2009). In order to see whether the germ cell inducing activity is present in the RNA or protein, the known mutant allele buc$^{p43}$ (aa 1-363) was injected along with the reporter into the corner blastomere. No effect (i.e no extra germ cells) was observed (Fig 14 C). The RNA of the mutant has identical sequences to its wild-type counterpart besides a point mutation, generating a premature STOP-codon (Bontems *et al.* 2009). Thus, overexpression of *buc* mRNA *per se* is not sufficient to induce germ cells. This data shows that Buc protein is sufficient to induce germ cells and not its RNA or mutant allele.

## 2.1.1 Xvelo, homolog of Buc induces germ cells.

Xvelo is a homolog of Buc in frogs. When *xvelo* mRNA was injected into the corner blastomere along with the reporter it also showed it could reprogram the fate from somatic to PGC (Fig 15B). When *ziwi* RNA (homolog of *piwi (Drosophila)* in the zebrafish), a member of germ plasm was injected into the

corner blastomere it could not induce germ cells (data not shown). This data suggests that these two germ plasm organizers across species have the ability to interact with the same pathway to specify germ cells. It also shows not all germ plasm members have the ability to induce germ cells.
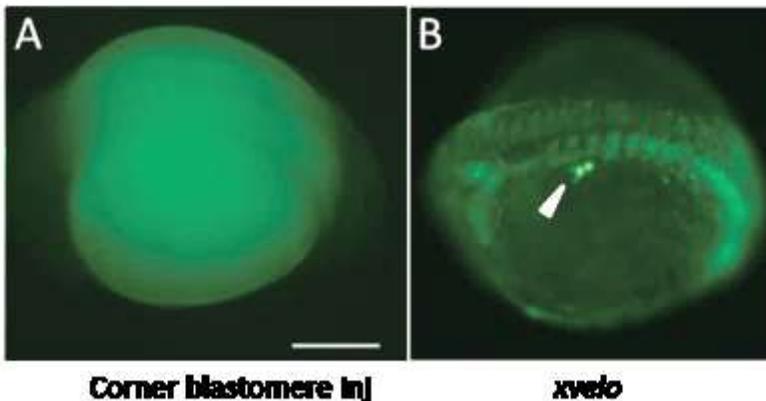


**Figure 15: Frog germ plasm organizer induces germ cells in the fish.**

(A) Live 18 hpf embryo after injection of PGC-reporter into a corner blastomere (A; 31±4.5%; n=70;).
(B) *Xenopus velo1* acts as a functional homolog of Buc. It induces fluorescent germ cells (white arrowhead) (B; 61±3.5%; n=41). Scale bar: 200 µm

## 2.1.2 Osk induces germ cells in the Zebrafish.

Oskar shows the same activity in invertebrates as Bucky ball in the zebrafish (Ephrussi and Lehmann, 1992; Bontems *et al.*, 2009). Upon injection of *Drosophila* Oskar into the fish embryo, it induces extra germ cell formation (Fig. 16A). This indicates that it is highly likely Oskar uses the same molecular pathway as Buc to induce germ cells.In order to see whether the germ cell inducing activity is present upon injection of mutant sOsk RNA (aa 139-254) was injected. There was no induction of germ cells observed (Fig 16B). The $osk^{84}$ RNA also had identical sequences to its wild-type counterpart besides a point mutation generating a premature STOP-codon (Kim-Ha, Smith and Macdonald, 1991).Thus, overexpression of osk mRNA *per se* is not sufficient to induce germ cells.

Quantification indicated a similar number of induced PGCs by Osk and Buc (Fig 16C). These results suggest that the germ plasm organizers sOsk, Buc and their homologs share the unique capacity to induce germ cells.
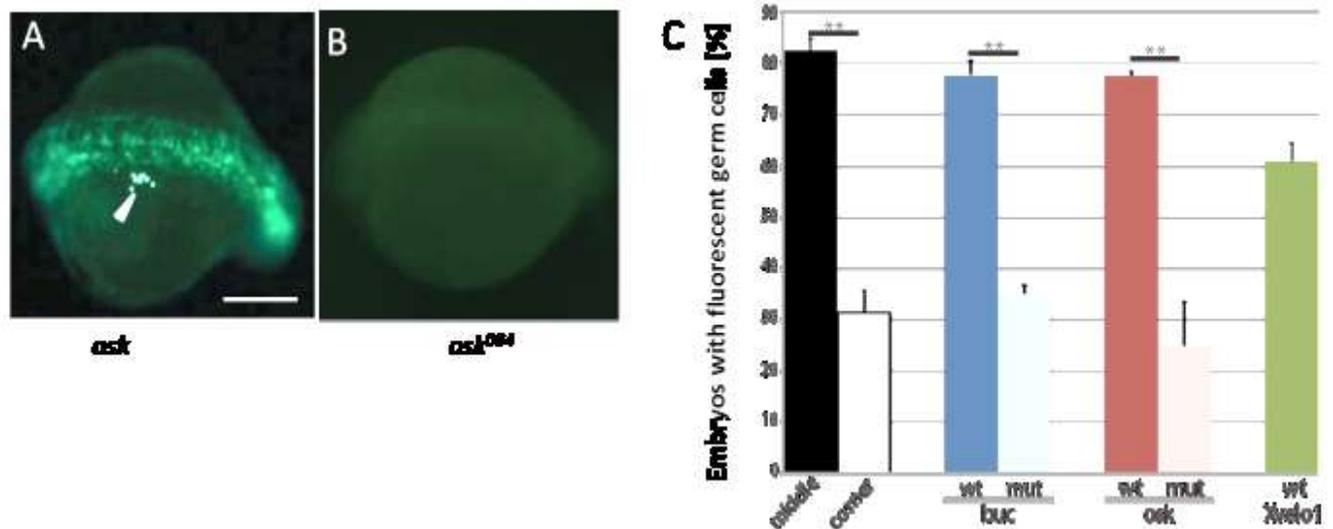
**Figure 16: Germ cell induction by Oskar.**

(A)*Drosophila sosk* induces germ cell formation (A; wt= 78±1.1%; n=81).(B)Mutant s*osk^084^* (aa139-254) (B; mut=25±8.7%; n=62). (C) Quantification of injection results (three independent experiments for each RNA).  Error bars represent standard deviation of the mean. Scale bar: 200 µm Student's t-test; P-value: **<0.01.

Therefore, the results (2.1- 2.1.2) suggest that the germ plasm organizers sOsk, Buc and their homologs share the unique capacity to induce germ cells.

## 2.2 Osk and Buc show no conserved sequence motif.

According to the sequence-structure-function paradigm, proteins with the same activity contain homologous sequence motifs to interact with similar binding partners (Doolittle, 1981). Conserved amino acid sequences were previously described for *Xenopus* Velo and zebrafish Buc, but not between *Drosophila* Osk and Buc (Bontems *et al*., 2009; Srouji and Extavour, 2010).  Therefore a stepwise strategy was pursued for their direct, *in silico* comparison starting with pairwise alignments (Fig 17). Only a 11.5% similarity between both proteins. The longOsk (lOsk) isoform, which is inactive in germ cell induction in *Drosophila*(Markussen *et al*., 1995), reduced similarity to Buc even further down to 10% (Fig. 17B).  A comparison of zebrafish Buc with *Drosophila* Vasa as an unrelated sequence showed 18.5% similarity, while Vasa homologs in zebrafish and *Drosophila* were 60.8% similar (Fig. 17B and Appendix Table 1). Studies done previously showed that when orthologs from different species are aligned, conserved domains and hidden similarities are revealed(Liu *et al*., 2008). The sequences of

14 vertebrate Buc orthologs were aligned, thus discovering two conserved motifs (aa 24-84 and 114-128) within the previously described BUVE-sequence (aa 23-136) (Bontems F, *et al*. 2009)and a novel motif in the center of Buc (aa 372-394) (Fig. 17C). The same approach was used with Osk and detected published motifs: the Lotus domain (aa 154-236 lOsk) (Anantharaman, Zhang and Aravind, 2010; Callebaut and Mornon, 2010; Ahuja and Extavour, 2014), the Lasp binding region (aa 290-369) (Suyama *et al.*, 2009)and the OSK domain (aa 452-490) (Jeske *et al.*, 2015; Yang *et al.*, 2015)(Figure 17C). A hidden Markov models (HMM) profile was generated of both proteins, but it did not detect significant hits by searching the Drosophila genome for sequence similarities with the Buc-HMM. The Buc-HMM consensus sequence, however, showed 43% identity of aa 56-81 in Buc to the DAZ motif in zebrafish Dazl (Fig 17D; Appendix Table 2)(Maegawa *et al.*, 2002). Searching with Osk-HMM identified Tdrd5 and -7 in zebrafish and Tejas in Drosophila, which all contain LOTUS-motifs, but no similarity to Buc (Fig. 17D; Appendix Table 2). Comparing the HMM-models of sOsk and Buc to each other also did not discover conserved motifs (Appendix Table 2). Taken together, the extensive bioinformatic analysis did not detect hidden sequence similarities between the two germ plasm organizers Osk and Buc, thus unable to explain their similar activity. Hidden Markov models (HMM) of both proteins, did not detect significant hits by searching the *Drosophila* genome for sequence similarities with the Buc-HMM. The results for this Fig 17 were performed by Dr. S. Riemer and Dr. T. Lingner (Riemer, 2014).
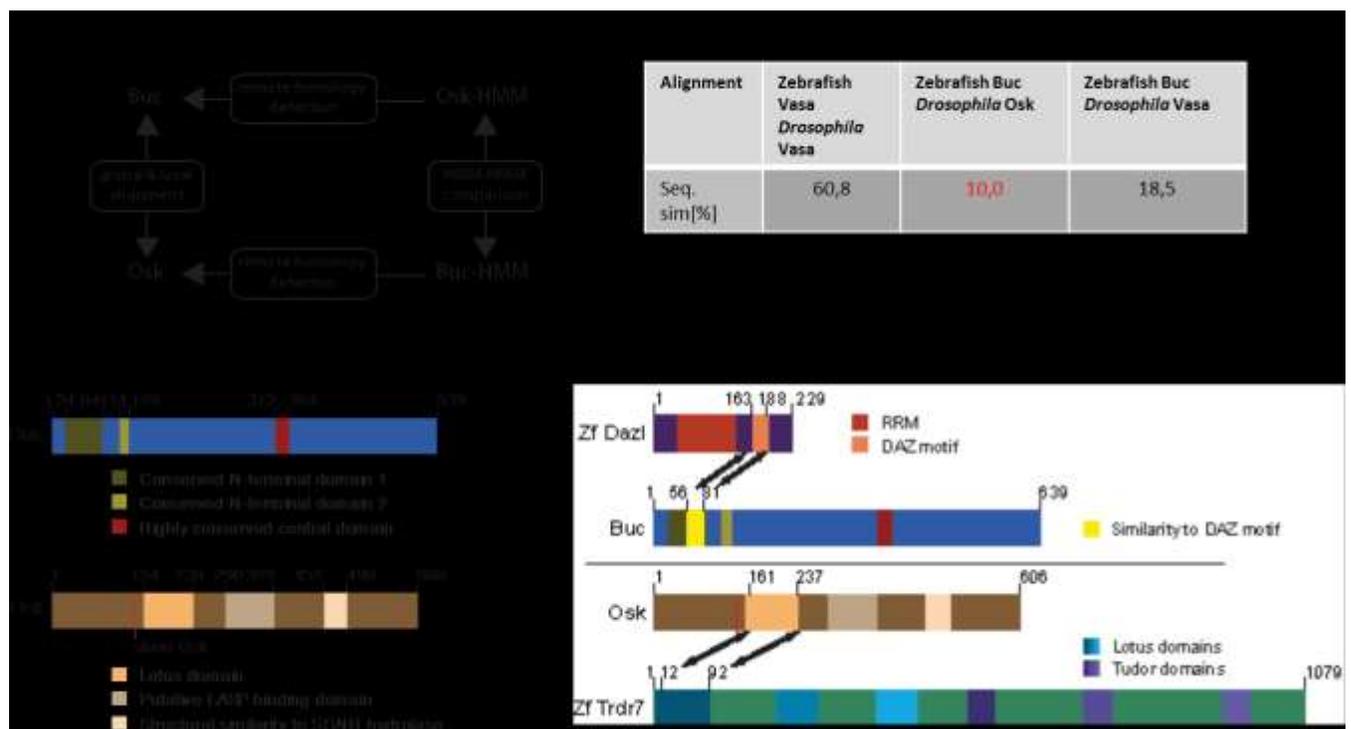


**Figure 17: Drosophila Osk and zebrafish Buc display unrelated protein sequences.**

(A) Strategy for Osk-Buc sequence comparison using first global and local alignment algorithms, then hidden Markov models (HMM) of both proteins to detect remote homologies, and eventually a comparison of HMM to each other. (B) Graph comparing sequence similarity of Buc with short (aa139-606) and long Osk (aa1-606) isoforms based on global pairwise sequence alignments. Alignment of Buc and Drosophila Vasa (Dm Vasa) serves as negative, whereas zebrafish and

Drosophila Vasa as positive control.  (C) Scheme of conserved domains in Buc identified by alignments of 14 protein sequences.  Conserved motifs in Osk protein.  Red line indicates alternative translation initiation of short Osk at Met139. (D) Scheme indicating significant remote sequence similarity based on HMMER analysis for Buc with zebrafish Dazl and for Osk with the Lotus domain of zebrafish Tdrd7. Numbers indicate amino acid positions.(Riemer, 2014)

## 2.2.1 No conservation of LOTUS and RNA binding domain between Oskar and Buc

Recently Oskar was shown to have a LOTUS domain which is required for multimerization and helps in the interaction with Vasa (Jeske *et al.*, 2015). Buc does not have a visible LOTUS domain (Fig 17). I aligned the delineated LOTUS domain of Osk (aa 139-240) with Buc in order to gain insight into the domain of Buc that could bind Vasa.

Lotus domain Homology between Buc and Osk

```
Oskar     ------------------------------------------------------------
Buc       MEGINNNSQPMGVGQPHHPVNHTRPFFYVQPPSQPYFMYQWPMNPYGHYGFPGPALHFGR


Oskar     ------------------MTIIESNYISVREEYPDIDSEVRAILLSHAQNGITIS-----
Buc       PYMAPYQFMQYPGYVIPHAPMQPIDYRRINPHYPSVASYDLRVRHHFQNAGMHRETACSE
                         :    :*  :. .**.: *      :   . : *:  .

Oskar     ------------------------------------------------------------
Buc       VQTDPSDSVNKLIDKIESLKACELGSDKGPNNVVSSTPDVVQGEKLTRLNEDSNLEVATK


Oskar     ------------------------------------------------------------
Buc       ECKEDPVTRPTTYSDSAYDAESSQGRLDECVFSDVLPLDSSSVHEEEEEEEKDVNEEDEP


Oskar     ------------------------------------------------------------
Buc       QTVADEICSQNEMSASTTSNVFCSGVQSIADPTECHDLEKLGDEQKQDIPSADAAAVIEP


Oskar     -----------------------------------------------------SIKSEY--
Buc       LISLSEDFDLPYQILRLPCNKTTTGLSLEREIDPLVYFDSPSTLLPPQNYLSSIGSAYSY
                                                               **  *  *

Oskar     ---------------------------------RKLTGNPFPL-------HDNVTDFL
Buc       SYYPQVTQERQSVLSPSIDELSSRDEMFSTDVEDLEVVPGHVYVGGGRLAEASDMPVRSR
                                           . : *. :          *  .

Oskar     LTIPNVTAECSESGKRIFNLKASL------------------------------------
Buc       KELPSVDKTCSVCQKTCACCGSTLQDEVGMCKMAEHSHPERDEMSDQDCDYDLEAEVRSN
             :*.*   ** . *        ::*

Oskar     ------------------------------------------------------------
Buc       CESPRVSKRKCCSRHALPSCGHHCAKHRHRKLLCEGQESCDLREQARVHPKGECCEEYGA
```

```
Oskar     --------KNGHLLDMVLNQKERTS-----------------------------------
Buc       LAKADKRIQKGALCRPCIEQQWREGVVSDQENWASCGAKPRSWRQVTGPQDQGRTPLRRS
                   ::*  *      ::*: *  .
```

Alignment 1: No conserved LOTUS region seen when Buc is aligned with Oskar. The LOTUS domain of Oskar highlighted in red aligns incongruously with Buc. (Identity: 24.2%; aligned using EMBOSS WATER)

No conserved region was detected in Buc confirming that Buc does not have a LOTUS domain.

Buc also has no motif with a known homology to any RNA binding domain. However, the OSK RNA-binding domain was also not described previously in other proteins and many RNA binding motifs do not show conserved domains(Castello *et al.*, 2016).  Even when the OSK domain (aa 401-606) is aligned with Buc it results in low sequence identity.

```
Osk       ------------------------------------------------------------
Buc       MEGINNNSQPMGVGQPHHPVNHTRPFFYVQPPSQPYFMYQWPMNPYGHYGFPGPALHFGR


Osk       ------------------------------------------------------------
Buc       PYMAPYQFMQYPGYVIPHAPMQPIDYRRINPHYPSVASYDLRVRHHFQNAGMHRETACSE


Osk       ------------------------------------------------------------
Buc       VQTDPSDSVNKLIDKIESLKACELGSDKGPNNVVSSTPDVVQGEKLTRLNEDSNLEVATK


Osk       --------TSGTYNDSLLTINS---------------------------------------
Buc       ECKEDPVTRPTTYSDSAYDAESSQGRLDECVFSDVLPLDSSSVHEEEEEEEKDVNEEDEP
                   **.**     :*

Osk       --------------------------------------------------DYDAYLLD
Buc       QTVADEICSQNEMSASTTSNVFCSGVQSIADPTECHDLEKLGDEQKQDIPSADAAAVIEP
                                                               *   *  :

Osk       FPLMGDDFMLYLARMELKCRFRRHE-----------------------------------
Buc       LISLSEDFDLPYQILRLPCNKTTTGLSLEREIDPLVYFDSPSTLLPPQNYLSSIGSAYSY
          :  :.:** *      :.* *.

Osk       ----RV------------------------------LQSGLCVSGLTINGARNR--LKR
Buc       SYYPQVTQERQSVLSPSIDELSSRDEMFSTDVEDLEVVPGHVYVGGGRLAEASDMPVRSR
              :*                              :  . : *.*  :  * :     .*

Osk       VQLPE---------------GTQIIVNIGSVDIMR-----GKPLVQIEHDFRLLIKEMHN
Buc       KELPSVDKTCSVCQKTCACCGSTLQDEVGMCKMAEHSHPERDEMSDQDCDYDLEAEVRSN
           :**.                *: :  ::*  .:.      . : : : *: *  :    *

Osk       MR-------LVPILTNLAPLGNYC----HDKVLC------DKIYRFNKFIRSECCHLKVI
Buc       CESPRVSKRKCCSRHALPSCGHHCAKHRHRKLLCEGQESCDLREQARVHPKGECCEEYGA
           .       *  *.:*    * *:**     *   :  . . :.***.
Osk       DIHSCLINERGVVRFDCFQASPRQVTGSKEPYLFWNKIGRQR----------------VL
Buc       LAKADKRIQKGALCRPCIEQQWRE—GVVSDQENWASCGAKPRSWRQVTGPQDQGRTPLR
          ::     ::*.:   *::  .  *: *  .    *  . * :                :
```
```

38
```

```
Osk     Q----------------VIETSLEY--------------
Buc     RSTCKSIHQQRPRSEYNDYDETEFTYCQRGRGSMKKRGSRY
        :               ..* .:.*
```

Alignment 2: No conserved OSK domain seen when Buc is aligned with Oskar. The OSK domain of Oskar highlighted in blue aligns incongruously with Buc. (Identity: 19%; aligned using EMBOSS WATER)

Overall this would suggest that the structure or biophysical nature of both proteins might be similar in order to accomplish the same activity.

## 2.2.2 Osk and Buc encode low-complexity proteins.

The *Xenopus* homolog Velo1 was shown to contain a low-complexity motif within the BUVE domain, which forms insoluble amyloids(Boke *et al.*, 2016). In addition, Osk and Buc were proposed to encode intrinsically disordered proteins (IDP) (Toretsky and Wright, 2014; Jeske *et al.*, 2015). I investigated to see whether, the functional conservation of Osk and Buc could be attributed to biophysical similarities of both proteins. As the intrinsic disorder of Osk and Buc was previously not shown, I analyzed the intrinsic disorder prediction of Osk and Buc using the PONDR-FIT algorithm. PONDR-FIT is a metapredictor, which conservatively combines the results of six prediction algorithms (Xue *et al.*, 2010) to show the regions of disorder within a protein. Both protein sequences displayed large disordered regions (Figs. 18A and 18B). Interestingly, the previously identified prion-like domain in the N-terminus (aa 1-150) (Boke *et al.*, 2016) represents the largest ordered sequence in Buc suggesting that almost the entire protein consists of low complexity sequences.
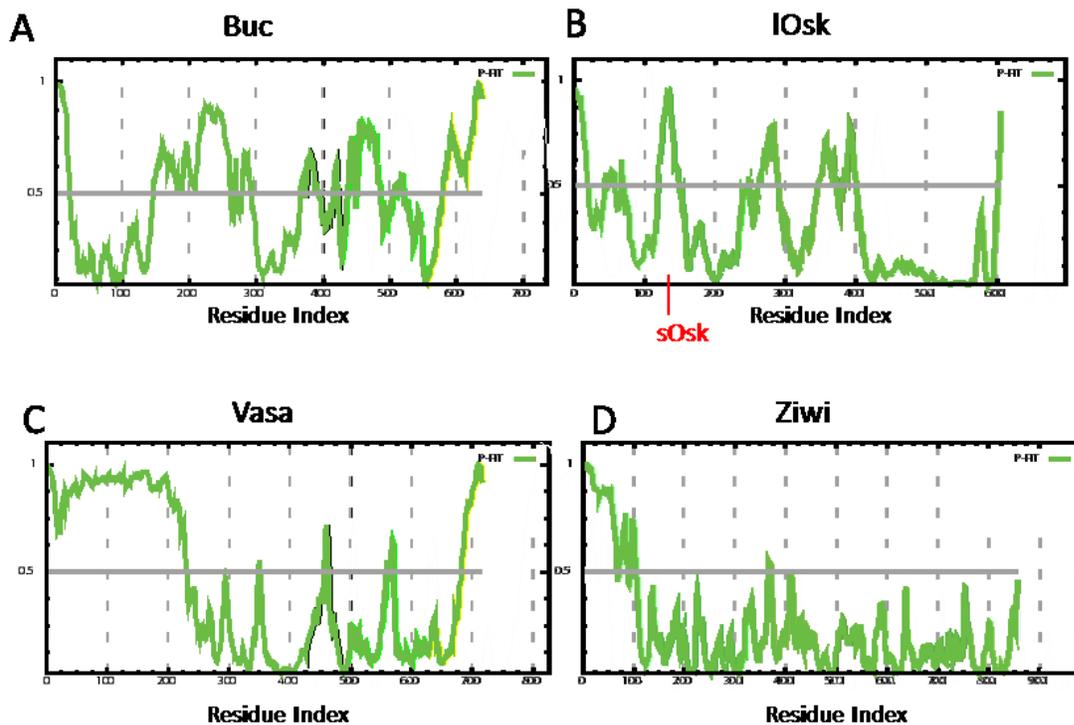
**Figure 18:Buc and Osk contain intrinsically disordered regions.**

Predicted protein disorder in (A) Buc, (B) Osk, (C), Vasa, and (D) Ziwi. Disorder disposition (y-axis) plotted against the amino acid residue index (x-axis). Values above the 0.5 threshold (grey bar) show the propensity for disordered regions. The red line at aa 139 in Osk indicates the alternative translation initiation site for short Oskar.

As a positive control for an IDP, I analyzed zebrafish Vasa (Fig. 19C), which showed the known unstructured domain of about 200 aa at the N-terminus(Yoon, Kawakami and Hopkins, 1997; Nott *et al.*, 2015). By contrast, the germ plasm component Ziwi was largely structured (Fig. 19D). This data suggests that Osk and Buc are indeed disordered.

## 2.2.3 Osk and Buc form aggregates in HEK cells.

Many IDPs with RNA binding function show the formation of granule formation upon overexpression. Indeed, Osk was recently shown to form aggregates in insect S2-cells supporting its function as an IDP (Jeske, Müller and Ephrussi, 2017). To investigate, whether Buc forms similar protein aggregates, I transfected HEK293 cells with plasmids encoding GFP-fusions. Buc and Osk formed protein aggregates, whereas the GFP control was uniformly distributed (Figures 19A-C).
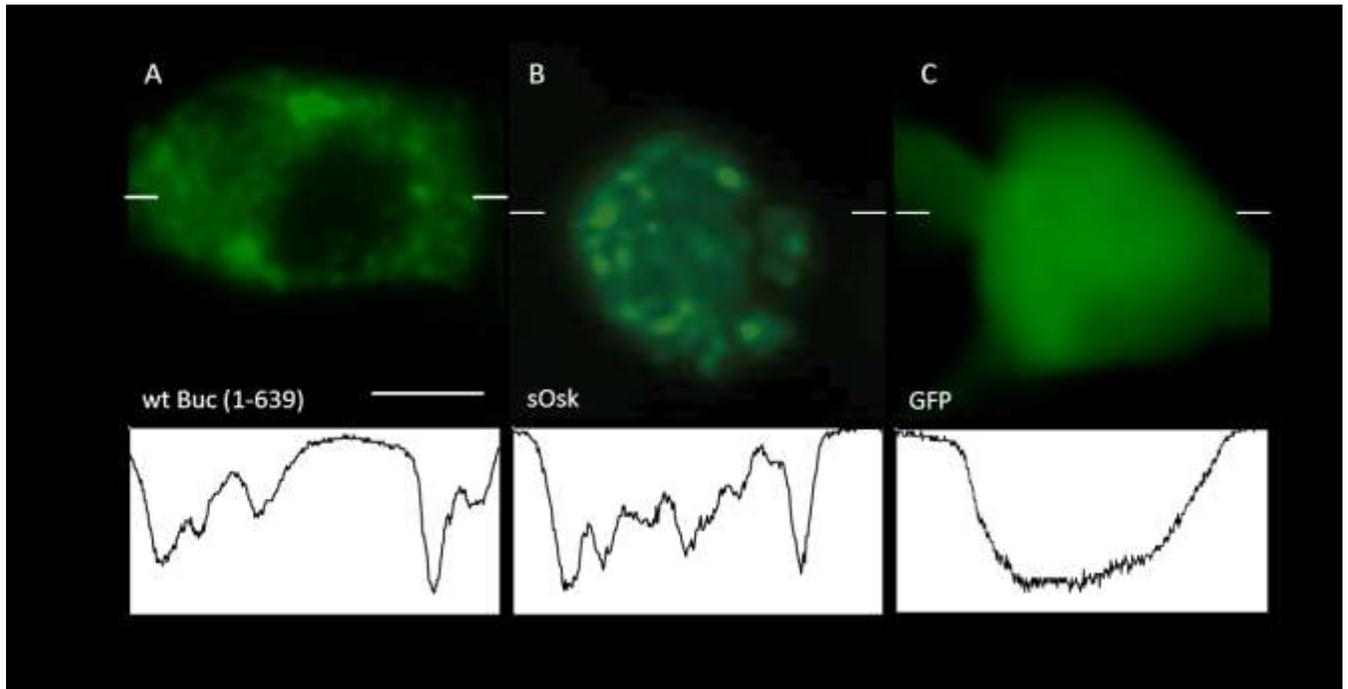
**Figure 19: Germ plasm organizer form cytoplasmic aggregates in vitro.**

Protein aggregates upon transfection of HEK cells with (A) Buc-GFP (99.3 ±1.15%; n=111), (B) sOsk-GFP (83.17± 8.18%; n= 90) and (C) GFP (0%; n= 81). The profiles below the pictures show levels of fluorescent intensity along the line indicated by white dashes. The peaks in the profiles of Buc and Osk transfected cells indicate presence of aggregates, whereas GFP show a uniform distribution within the cytoplasm. Scale bar: 10µm

These data indicate that Osk and Buc encode IDPs with a propensity to form cellular protein aggregates *in vitro*.

## 2.2.4 Buc forms hydrogels in zebrafish embryos.

The aliphatic solvent 1,6-hexanediol dissolves hydrogels formed by IDPs, but not amyloid-like aggregates as described for Velo1 in *Xenopus* oocytes (Boke et al., 2016; Kroschwald et al., 2015). Therefore this drug will allow me to distinguish whether the aggregates formed by Buc are amyloid-like similar to the Xenopus oocyte or more hydrogel-like similar to the C. elegans germ plasm. To address this question in vivo, I took advantage of the Buc-GFP line, which was previously shown to report Buc activity in vivo. After hexanediol-treatment of transgenic embryos (Riemer et al., 2015), Buc-GFP was fragmented into smaller aggregates whereas control-treated embryos showed no change (Figs 20A-C). Fig. 20D shows the number of embyros displaying fragmentation of the Buc-eGFP aggregates in the control and the hexanediol treated. These results indicate that zebrafish germ plasm forms an intracellular hydrogel, whose aggregation is mediated by IDPs like Buc.
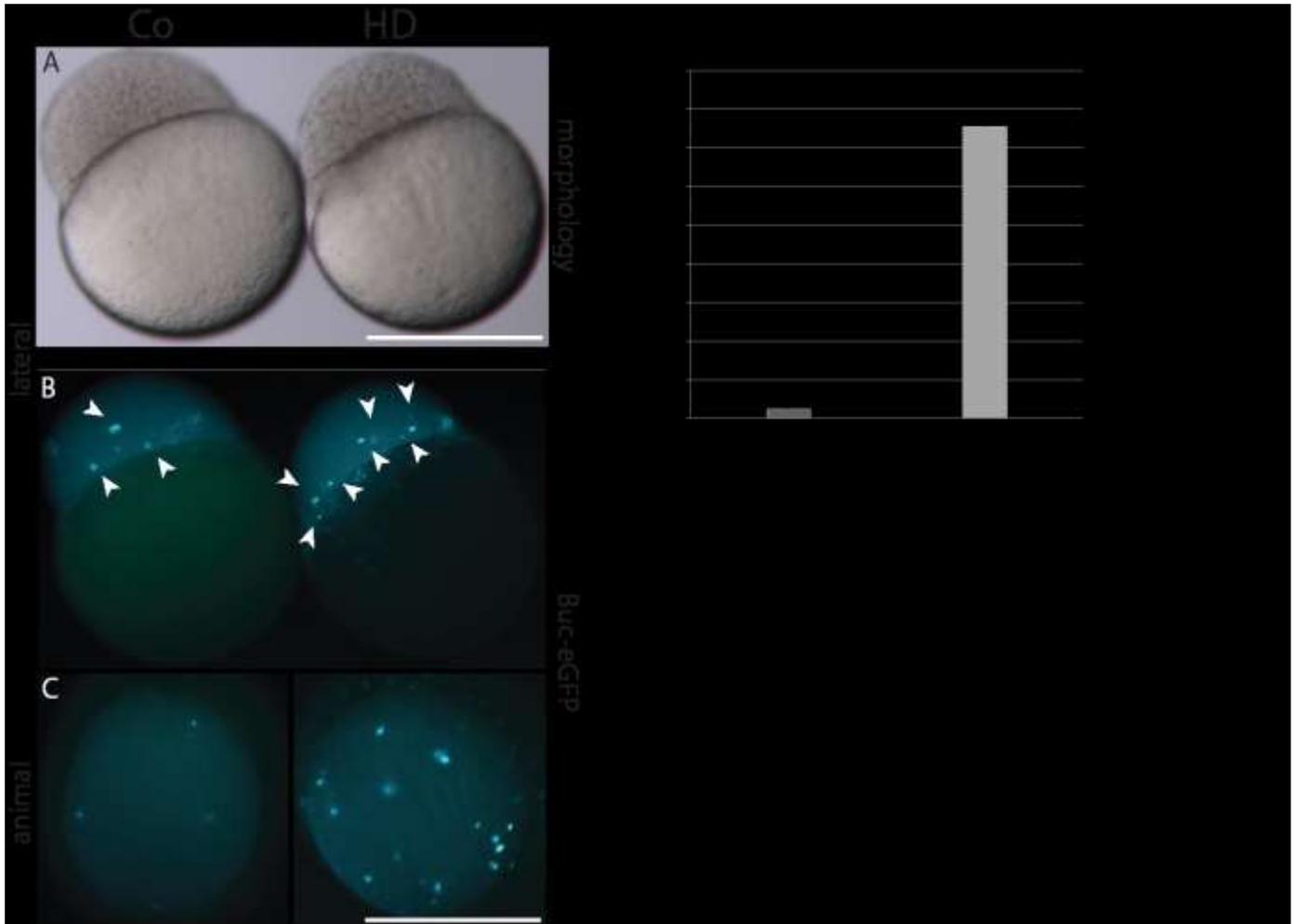
**Figure 20: Pharmacological disruption of IDR-interactions leads to fragmented Buc-GFP aggregates.**

(A-C) The hydrogel disruptor 1,6-hexanediol (HD) fragments aggregates in 3 hpf embryos transgenic for Buc-eGFP. The morphology after (Co) and HD-treatment is normal (A), whereas fluorescent Buc-eGFP aggregates are fragmented with HD (right embryo in B; white arrowheads, lateral view, animal pole to the top-left) (C, animal view).  Scale bar: 500 μm.  (D) Quantification of Buc-eGFP embryos displaying fragmentation control treated vs HD treated Student's t-test; P-value: ***<0.005

## 2.2.5 Intrinsically disordered regions of Buc are not sufficient to induce germ cells.

To investigate whether Buc aggregation is sufficient for germ cell specification, I used the mutant *p106* allele of Buc discovered in the mutagenesis screen (Fig.10). This allele encodes amino acids 1-601 of Buc and hence, lacks 38 C-terminal amino acids. This mutant, Buc (aa1-601) injected into 16-cell embryos in the PGC induction assay (16 –cell assay) did not express the PGC reporter (Figures 21A-C). These results suggest that the IDRs of Buc are not sufficient to specify germ cells.
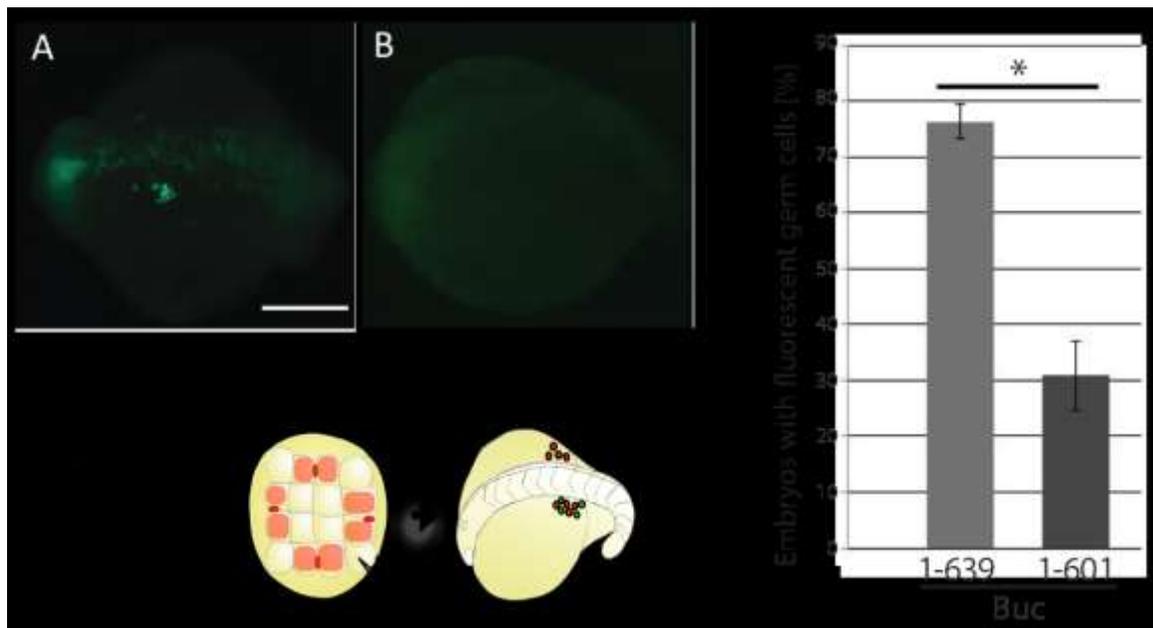


**Figure 21: IDR regions of Buc insufficient to induce germ cells.**

IDR regions of Buc insufficient to induce germ cells. Embryos form germ cells (white arrowheads) after injection with wt Buc mRNA (aa 1-639) (A; 76.3±4.0%; n=60), but not with mutant Buc (B; aa1-601) containing most IDRs (D; 30.8±8.7%; n=60; p=0.04). Scale bar: 200 µm. (C) Quantification of injection results (three independent experiments for each RNA). Error bars represent standard deviation of the mean. Student's t-test; P-value: *<0.05

## 2.2.5 Intrinsically disordered regions of Buc are sufficient to form aggregates.

To confirm that the retained IDRs of Buc are still able to form aggregates, I transfected their GFP tagged versions into HEK cells. Buc (1-601) still forms protein aggregates in HEK 293 cells compared to a GFP control (Fig. 22A). Using the p43 allele of Buc, which reduces the protein further to aa 1-363,

still leads to protein aggregation (Fig. 22 B), but this form did also not induce germ cells (Fig. 14C). These results demonstrate that aggregation of Buc is not sufficient for PGC formation. Moreover, the IDR of germplasm organizers are consistent with the rapid evolution of their protein sequence, but other biochemical interactions are critical for germ cell specification.
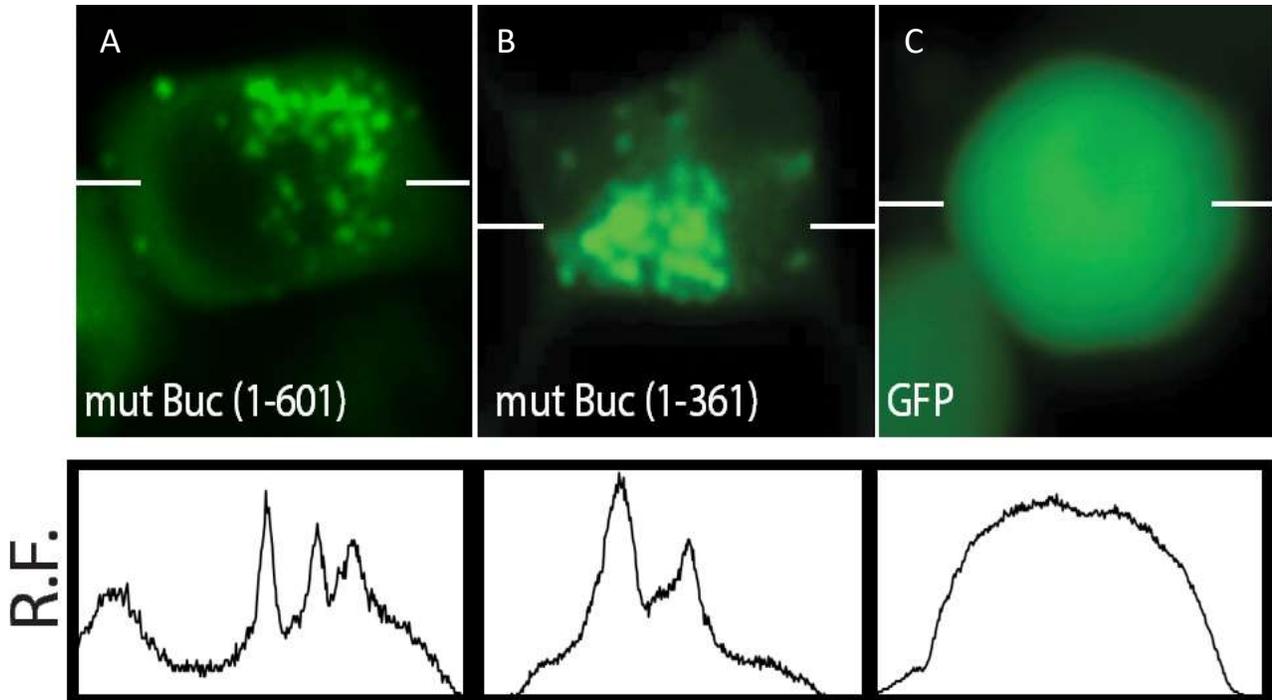


**Figure 22: Protein aggregates formed upon transfection of HEK cells**

HEK cells transfected with (A) Buc(1-601)-GFP (50.32±2.95%; n=70), (B) Buc(1-363)-GFP (77.9±8.8%; n= 89) and (C) GFP (0%; n= 81). Scale bar: 10µm. The peaks in the profiles of the transfected cells indicate presence of aggregates, whereas in GFP show a uniform ditribution within the cytoplasm.

Therefore take together results (2.2 -2.25) it shows that Buc and Osk have no sequence similarity; however they show the propensity and characteristics of disordered proteins. These shared characteristics are what give them biophysical similarities.

## 2.3 Buc and Osk bind similar proteins.

The similar function of Osk and Buc postulates that they most likely perform similar biochemical interactions, which then initiate the PGC-specification program. Osk binds directly to Smaug, Valois, and Vasa protein (Anne, 2010; Breitwieser et al., 1996; Dahanukar et al., 1999). The homologs of these Drosophila proteins are Samd4b, Mep 50 and Ddx4 also known as zebrafish Vasa. To test whether these proteins are conserved in the Buc interactome, the Buc-GFP line was exploited the again. The rescue of Buc mutants shows that Buc functions correctly and provides at the same time a tagged version of Buc protein for immuno-precipitation of endogenous protein. At the same time, the transgenic line avoids unspecific interactions, which would occur after overexpression of Buc-GFP by mRNA overexpression, because the transgenic fish express Buc under control of its own promotor (Riemer et al., 2015). Buc-eGFP was immunoprecipitated from 3hpf zebrafish embryos, when PGC specification occurs (Fig 23A). Binding partners were identified by mass-spectrometry and searched for the zebrafish homologs of the three Osk binding partners in the Buc interactome (Fig 23C). As a control, eGFP was injected into zebrafish embryos and pulldown at the 3hpf stage as well. Proteins that bound to it were identified by mass spectrometry as well. I determined the enrichment of the homologs of the three binding partners of Oskar. Interestingly, MACF1 was also found to be highly enriched in the Buc interactome. Zebrafish mutants in *macf1* and *buc* show defects in embryonic polarity and Balbiani body localization (Dosch *et al.*, 2004; Gupta *et al.*, 2010; Bontems *et al.*, 2011; Escobar-Aguirre *et al.*, 2017) supporting the specificity of the biochemical interaction.



| Protein | Acc-number | Drosophila Homolog | Enrichment [Buc-GFP/GFP] |
|---------|------------|---------------------|--------------------------|
| Ddx4 | gi\|528491414 | Vasa | 3.4 |
| Wdr77/Mep50 | gi\|41054711 | Valois | 1.8 |
| Samd4b | gi\|326674545 | Smaug | - |
| Macf1 | gi\|528510265 | Short stop | High enrichment |
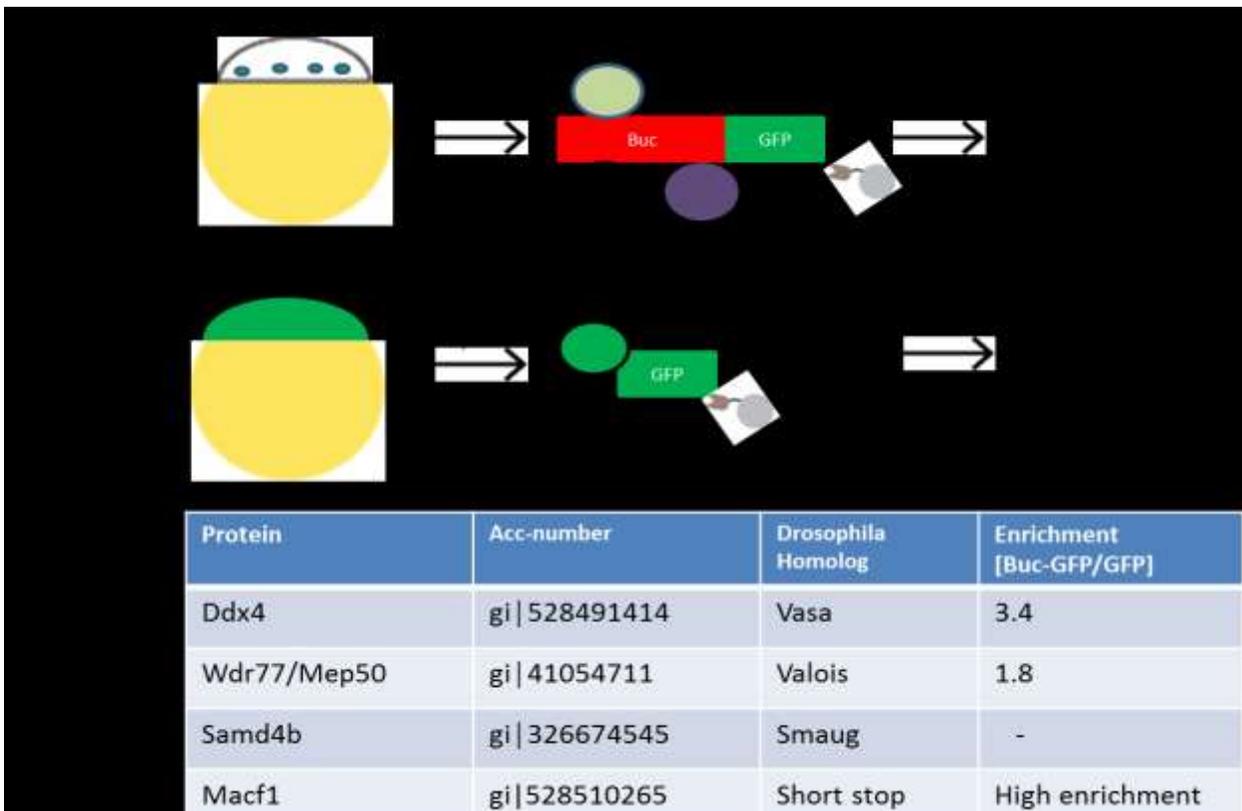
**Figure 23: Identification of interactors of Buc.**

(A) Embryos of the transgenic *buc-gfp* line were lysed at 3hpf. (B) *egfp* RNA injected embryos at the 1 cell stage were used as negative control. Subsequent to lysis an IP against the GFP-tag was carried out. Interacting proteins were identified by mass spectrometry (Core Facility of Proteome Analysis, UMG, Goettingen). (Performed by Dr. S.Riemer ). (C) Table showing zebrafish homologs of known Oskar binding proteins in the Buc-interactome detected Vasa (Ddx4) and Valois (Wdr77/Mep50), but not Smaug (Samd4b).  Enrichment indicates the ratio of unique peptide counts after Buc-GFP pulldown to GFP-control samples. Macf1 is shown as a control for the specificity of the pulldown as a known interactor of Buc.

From the table the homologs of Vasa, Valois seemed to be enriched in the pulldown, but not Smaug. This data suggests that Osk and Buc might indeed utilize some interactors conserved across *Drosophila* and zebrafish.

### 2.3.1 Buc binds zebrafish Vasa.

As shown in the table in Fig. 23C Vasa seemed to be more strongly enriched than the other Oskar interactors. Therefore, I selected Vasa for further analysis, since its stronger enrichment suggested a greater probability to interact with Buc. To verify that Buc interacts with Vasa during the period of germ cell specification, I pulled down Buc-GFP from embryonic extracts of transgenic embryos at 3 hpf (Fig 24). As a control another maternal GFP transgenic line H2AGFP was used, as it also shows GFP fluorescence during 3hpf. Even though H2AGFP is localized to the nucleus and Buc to the cytoplasm it was used as a negative control, due to the paucity of alternative maternally expressed GFP-fusions in the zebrafish field.
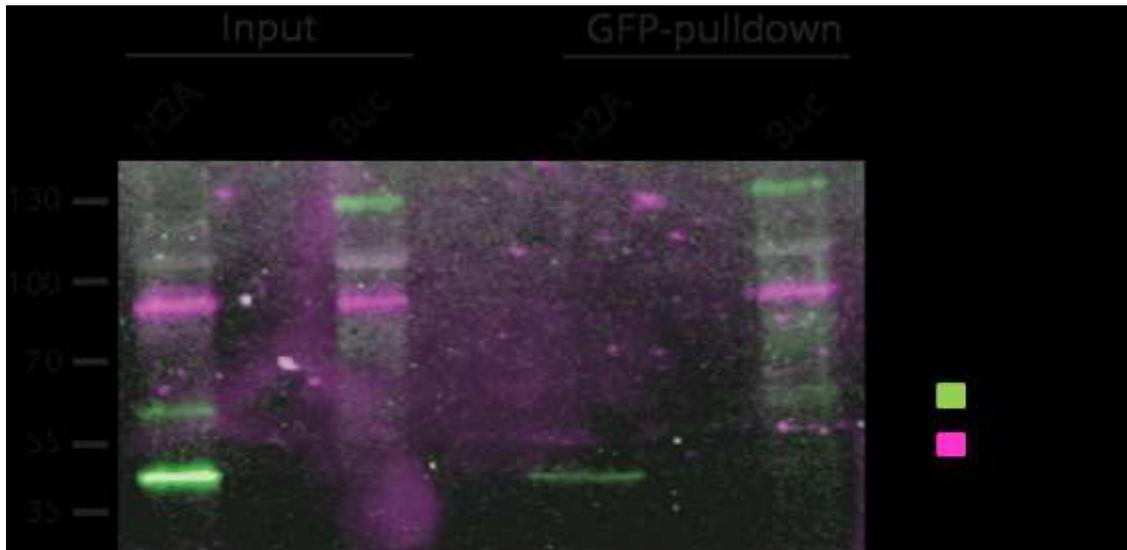


**Figure 24: Buc binds to Vasa in vivo during germ cell specification.**

Immunoprecipitations from 3 hpf H2A-GFP (42 kD) or Buc-GFP (130 kD) transgenic embryos blotted with GFP (green) and Vasa (magenta) (input = 20% of pulldown). H2A-GFP embryos were used as a control.

This data shows that Vasa interacts with Buc in vivo during germ cell specification and confirms the mass spectrometry data.

## 2.3.2 Buc and Vasa co-localize during germ cell specification.

During the mechanical disruption of the embryos to generate the protein extracts for the immuno-precipitation, proteins from different compartments are mixed. It is therefore critical to determine whether two potential interaction partners co-localize during embryogenesis. To determine whether Buc and Vasa co-localize during germ cell specification, I labeled zebrafish embryos by antibody staining. Vasa is ubiquitous at the 16-cell stage and at 3 hpf as previously described (Figs. 25A-F) (Braat *et al.*, 2000; Knaut *et al.*, 2000). Buc localization overlaps with Vasa only in the germ plasm, which further supports the in vivo interaction of Buc and Vasa (Fig 24).
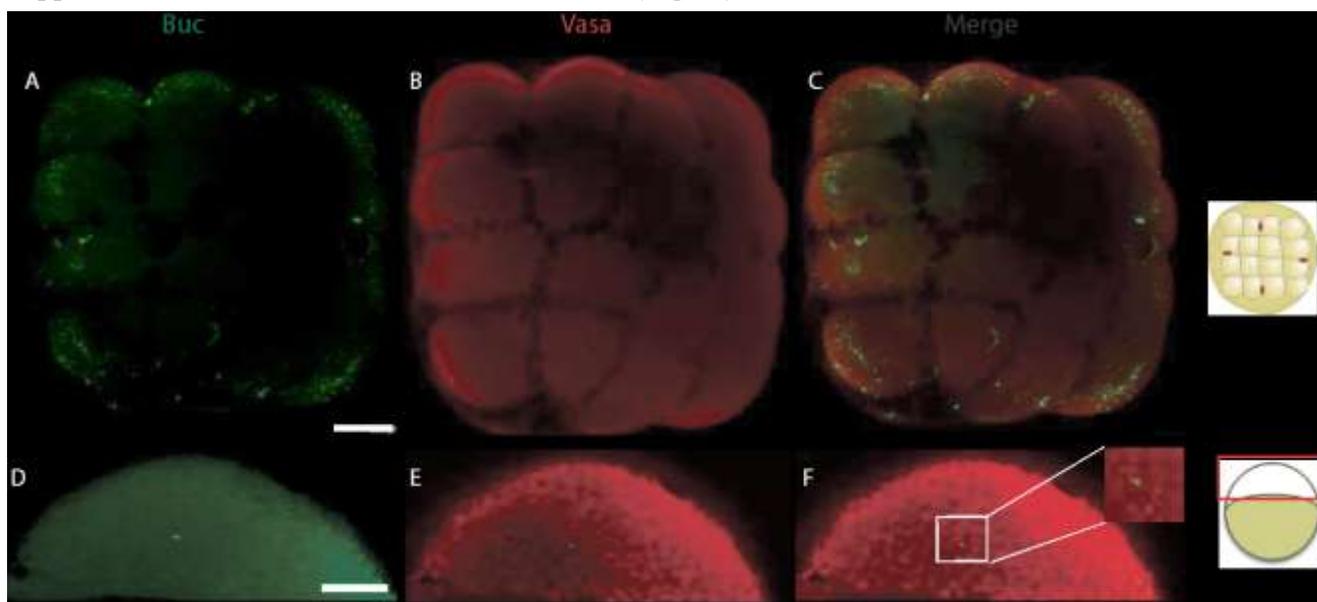


**Figure 25: Co-localization of Buc and Vasa.**

A-F) Immunostaining of 16-cell stage (A-C) or 3 hpf (D-F) embryo as indicated by cartoon on the right showing co-localization of Buc (green) and Vasa (red), inset shows a 10x magnification of boxed area. Scale bar: 200µm

Hence this data suggests that Buc and Vasa co-localize in prospective germ cells and this initiates the next step to specify them as primordial germ cells.

### 2.3.3 Vasa is sufficient to induce germ cells in the zebrafish.

Previous studies in chicken showed that Vasa overexpression reprograms embryonic stem cells to a germline fate (Lavial *et al.*, 2009). Furthermore, studies in *Drosophila* showed that Osk enhances Vasa activity suggesting that Vasa has a key role during germline specification(Jeske, Müller and Ephrussi, 2017). Therefore the role of Vasa was analyzed with the zebrafish germ cell induction assay. Fascinatingly Vasa also induced ectopic germ cells, whereas another known Buc binding protein Hermes(Nijjar and Woodland, 2013; Heim *et al.*, 2014) showed no activity (Fig 26 A-C).
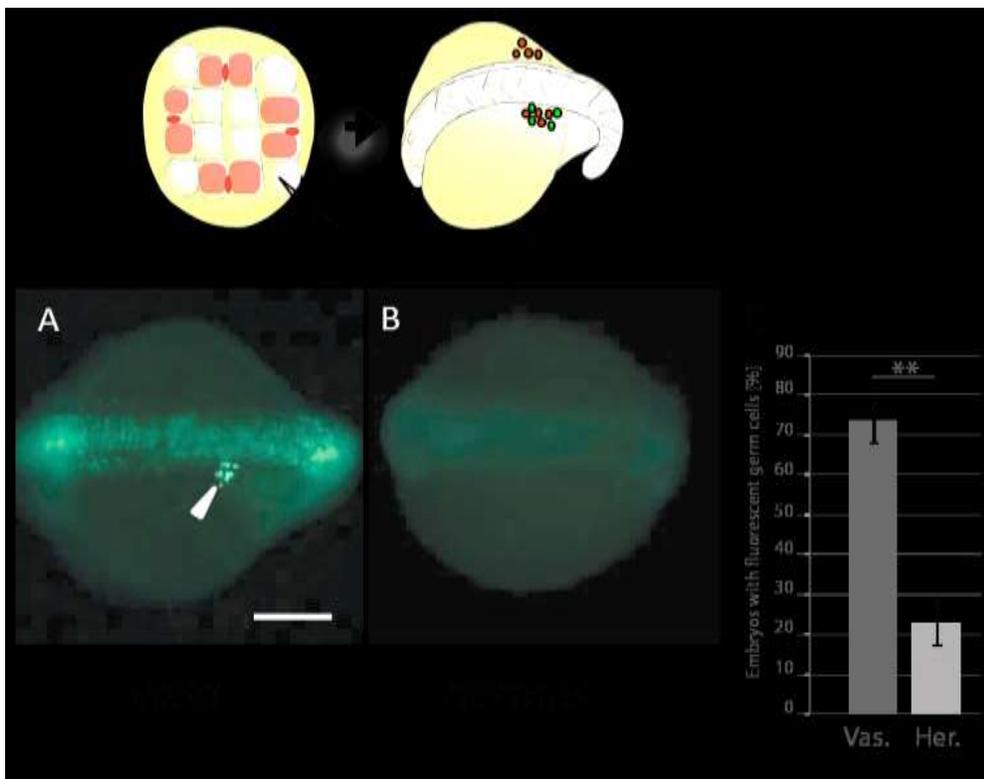


**Figure 26:Vasa induces germ cells in the zebrafish.**

Cartoon of germ cell induction assay above below results of the assay showing germ cell formation (white arrowhead) after injection with Vasa mRNA (A; 73.9±5.3%; n=60; p=0.01) but not with Hermes (B; 22.9±4.8%; n=60).  (C) Quantification of injection results (three independent experiments for each RNA).  Error bars represent standard deviation of the mean. Student's t-test; P-value: **<0.01

This result suggests that Vasa performs a central activity during germ cell specification.

## 2.3.4 Buc and Osk both bind zebrafish Vasa.

Osk activates the activity of *Drosophila* Vasa (Jeske, Müller and Ephrussi, 2017). I therefore speculated that Osk might interact with zebrafish Vasa and this triggers germ cell formation in zebrafish. I investigated, whether *in vitro* translated Osk-GFP binds to zebrafish Vasa using in vitro translated GFP as a negative control. Indeed, Osk pulled down zebrafish Vasa (Fig 27A), supporting the hypothesis that Osk and Buc share conserved interactions. The Buc (1-363) mutant does not induce PGCs and we therefore analyzed its interaction with Vasa. Buc (1-363) bound Vasa just like wt Buc (Fig 27A).

This result suggests that Vasa interaction is not sufficient for PGC specification.

Based on the previous result Buc (aa 1-363) binds Vasa just like wt Buc. This implies that the mutant Buc protein could indeed compete with the endogenous Buc and bind Vasa. To test whether Buc (1-363) interferes with PGC formation, I injected transgenic Buc-eGFP embryos at the one cell stage, and scored the number of PGCs at 18hpf. However, as seen in Fig 27B there is no decrease in number of germ cells. This suggests that either the mutant Buc protein could not interfere with an already pre-bound Buc/Vasa complex or the mutant Buc bound the excess ubiquitous Vasa and did not interfere with germ cell specification.
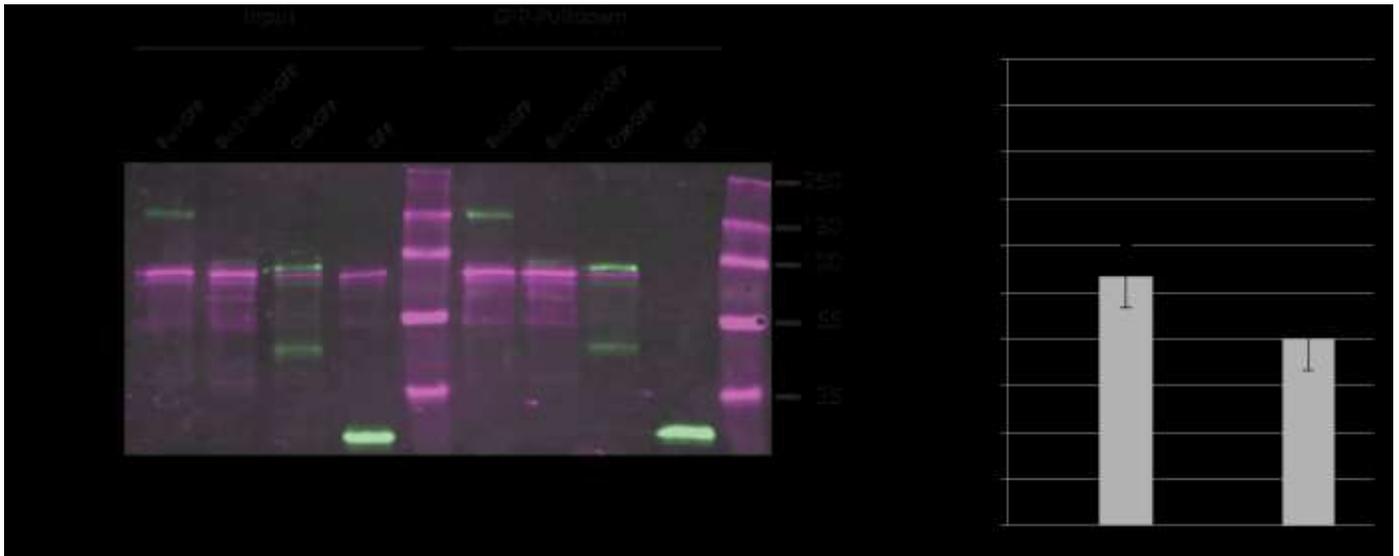


**Figure 27: : Buc, Osk bind Vasa and Buc p43 does not seem to compete for Vasa binding.**

(A) Western blot of Buc-GFP, mutant Buc-GFP (1-363), Osk-GFP, and GFP-control (green) together with Vasa (magenta) after in vitro translation (input = 40% of pulldown) and after GFP-pulldown. Vas (magenta) interacts with Buc, Buc (1-361) and Osk, but not GFP controls (green ).(B) Comparison of number of germ cells in Buc-eGFP embryo at 24hpf uninjected (20.6± 4.16) and injected with mutant Buc (aa 1-363) ( 18± 3.05) encoding mRNA (300ng). No significant decrease in number of germ cells seen. Student's t-test; P-value: 0.4, n= 30.

## 2.3.5 Buc and Osk interacts with *nanos3* RNA

Mutant Buc (aa 1-363) binds to Vasa, but does not induce germ cells suggesting that this protein lacks an additonal interaction. From two recent publications, it was shown that the OSK domain of Osk protein binds RNA (Jeske *et al.*, 2015; Yang *et al.*, 2015). Many IDPs are also RNA-binding proteins (reviewed in(Järvelin *et al.*, 2016)). Aggregates of the Xenopus homolog VeloI were previously shown recruit nanos mRNA(Boke *et al.*, 2016). I therefore investigated whether Buc, could also bind RNA (Fig 28A). Buc has a positively charged C terminal end which could be involved in RNA binding. Also this could suggest that full-length Buc performs additional interactions with RNA. To address whether Buc interacts with the 3´UTR of zebrafish *nanos3* mRNA (Köprunner *et al.*, 2001), I co-expressed its GFP-tagged versions in HEK293 cells (Fig 28B). After immunoprecipitation of Buc, I detected zebrafish *nanos3 3´UTR* by RT-PCR, but not an abundant, endogenous control (*18S rRNA*) (Fig 28C). Similarly, Osk-GFP bound to zebrafish *nanos3* mRNA consistent with its conserved activity during germ cell specification in zebrafish. These results show that Buc can bind RNA and further, both Osk and Buc can bind *nanos3* RNA.
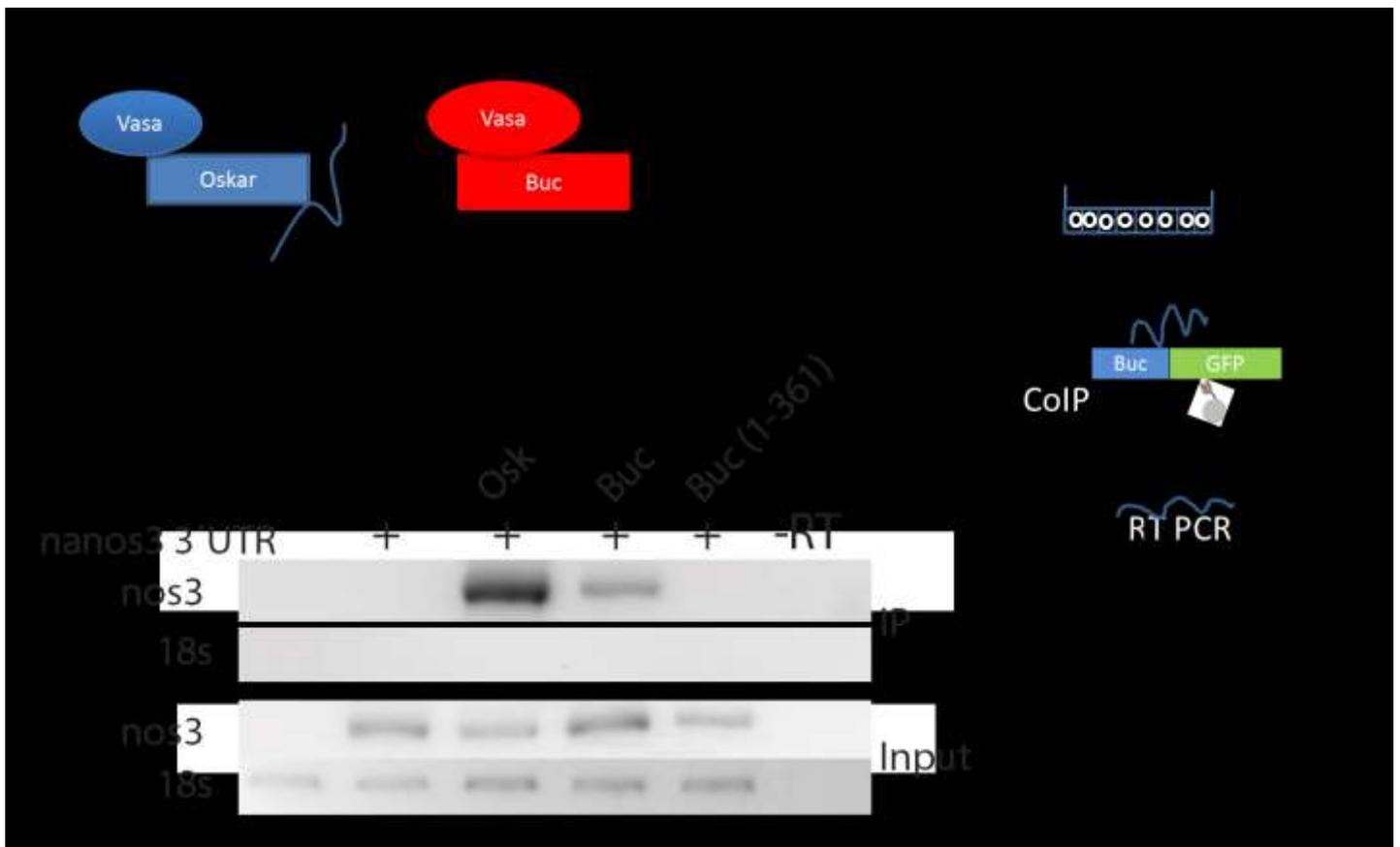


**Figure 28 : Buc and Osk bind zebrafish nanos3 3´UTR**

(A) Comparison of known interactors of Osk and Buc. (B) Scheme of experiment to check for RNA binding in Buc. (C) Buc interacts with RNA. RT-PCR of zebrafish nanos3 and human 18S rRNA after GFP pulldown (IP) of HEK293 cells either untransfected (-) or transfected with Cherry-nos-3'UTR (+), alone or co-transfected with Osk-GFP, Buc-GFP or Buc(1-363)-GFP.

## 2.3.6  Is the binding between Buc and Vasa mediated by RNA?

As both Buc and Vasa interact with RNA, their interaction might be mediated indirectly via RNA. Apart from Buc binding RNA, Vasa is an RNA helicase, it could independently bind RNA and complex it together with Buc in Fig 29A or Buc might bind the RNA independently from Vasa. Therefore two experimental setups were performed to identify the type of interaction (Fig 29B). RNase treatment (with RNAse A and T1) did not inhibit Buc-Vasa binding in both setups (i) (Fig 29C) and (ii) (Appendix Fig 2), showing that the complex was held possibly together by protein–protein interactions, protein interactions, or was protected from the RNAse by being the bridge between the two proteins, thereby shielded from the activity. But specifically looking at Fig 29C, it suggests that the Buc-Vasa interaction is probably not mediated by RNA.
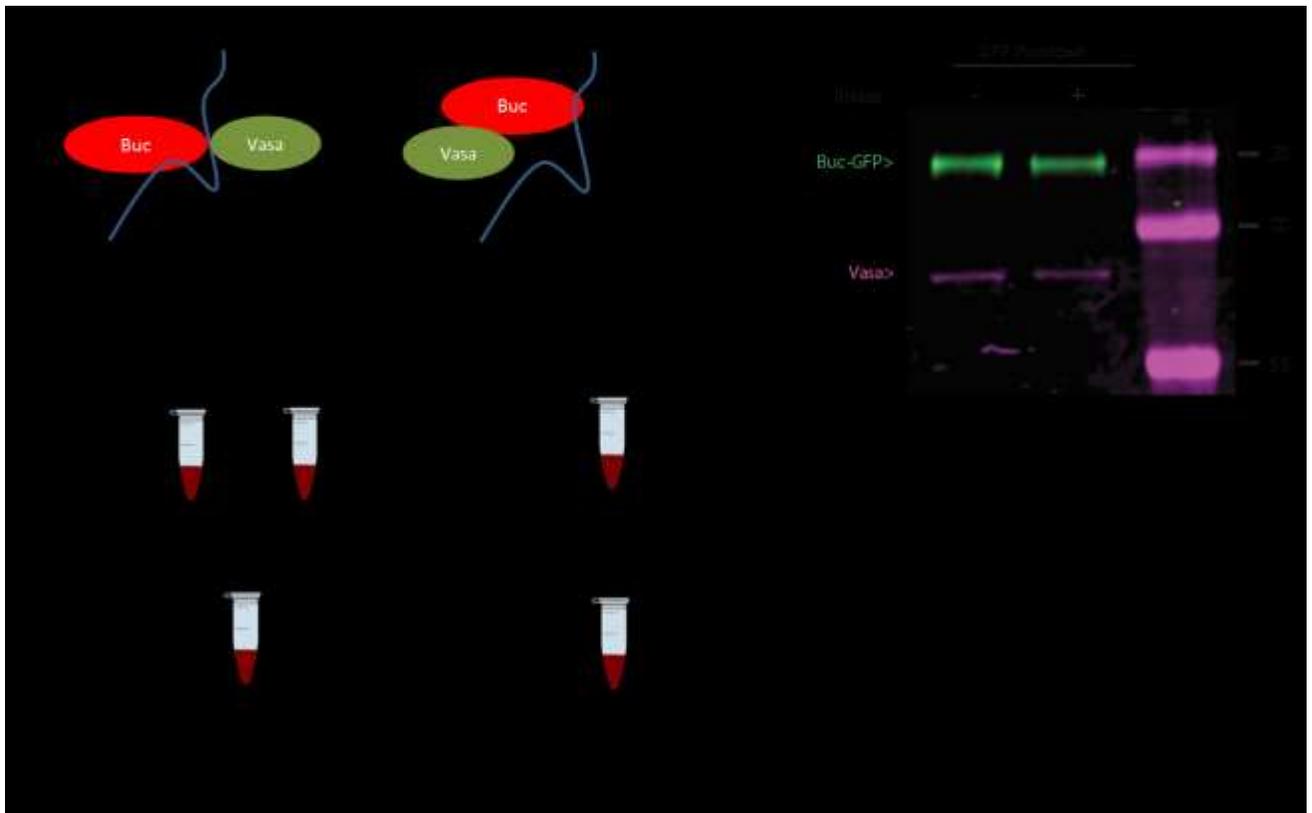


**Figure 29:RNA does not seem to mediate Buc –Vasa interaction**

(A) Two possible scenarios of Buc and Vasa interaction mediated by RNA . (B) Two experimental setups that were performed to check whether RNA mediates the Buc-Vasa binding. (C) Western blot showing Buc Vasa pull down from setup

(i) RNA without reverse transcriptase control(-),  RNase treatment (+) prior to GFP pulldown of *in vitro.* Translated protein from setup (ii) does not disrupt Buc-Vasa interaction either.  Buc labelled in green, Vasa in magenta.

Therefore results (2.3-2.3.6) shows that Buc and Osk indeed display a conserved interactome.

In summary of results (2 - 2.3.6), the biophysical similarities of Buc and Osk allow for self-aggregation and complex formation in the embryo. The concerted role of Buc, Vasa and Nanos 3 mRNA interacting complex within this aggregation plays a central role for germ cell specification in zebrafish.

# 3. Discussion

My results show that Buc and Osk, the two germ plasm organizers have a number of similarities. Both have the ability to induce germ cells in the zebrafish. These two proteins have no common sequences to encode for common domains. Also the two proteins show regions of intrinsic disorder and Buc was shown to form a liquid hydrogel aggregate. Perhaps since both are intrinsically disordered, this is what gives them the ability to have the same interactome. Both interact with Vasa protein and *nanos3* RNA. This shows that Buc and Osk are functionally conserved.

## 3.1 The 16-cell or germ cell induction assay and its advantages.

Analyzing whether a particular factor positively influences or contributes to germ cell specification in the zebrafish was made possible by this powerful assay. Injecting the mRNA encoding for the protein of interest in a blastomere can tell us the properties of that RNA or protein with regard to germ cell specification.

There are two main advantages that were considered when selecting this procedure for analyzing Bucky ball's function in germline determination. First, it can distinguish whether a factor is merely bound to Buc and for example could be involved in its localization or it has a genuine role in germ cell specification. Secondly, it leads to the overexpression of the factor of interest in a single blastomere, hence it will have a concentrated action in that blastomere and its future daughter cells. One drawback is that, if a factor needs to be activated for germ cell specification it might not be able to self-activate in all cases and give rise to germ cells. Overall this assay has worked well in identifying factors responsible for specifying germ cells.

## 3.2 Functionality of the induced ectopic germ cells in the zebrafish

The ectopic germ cells induced by Oskar in *Drosophila* were shown to be functional(Ephrussi and Lehmann, 1992). Buc induces germ cells in the zebrafish(Bontems *et al.*, 2009).The germ cells induced in the 16 cell assay show expression of *gfpnos 3´UTR,* express *vasa* and show targeted migration to the gonadal ridge. These characteristics are seen in endogenous germ cells as well. This suggests that the ectopic germ cells behave exactly like PGCs.

However whether these germ cells are functional remains an open question. This can be resolved through the following experiment (Scheme in Fig 30).

*buc*, *cherrynos 3´UTR (a variant of gfpnos 3´UTR)* will be injected into a Buc-GFP (Riemer *et al.*, 2015)16-cell stage embryo. The BucGFP fish strain has horizontal stripes across its body. At 24 hours post fertilization (hpf) the injected embryos with both Buc and *cherrynos 3´UTR* mRNAs will form

extra germ cells. These germ cells will be disctinct from the endogenous ones because of the combined expression of BucGFP and Cherry rather than only BucGFP.

Germ cells from these injected embryos will be isolated. Once the cells are separated from the body of the embryo they will be isolated via FACS (Fluorescence-activated cell sorting).

These sorted cells will be transplanted into the high stage embryos (3hpf) of another zebrafish line with spots on the skin and bred for offspring. If one detects the Buc-GFP genotype or stripes in some of the offspring it will show that the transplanted and hence ectopic germ cells are functional . Hence also showing that Buc is sufficient for the specification of functional germ cells in zebrafish. Alternatively, one can eliminate Buc after fertilization to creat a protein knock-out as described in the following section 3.3.
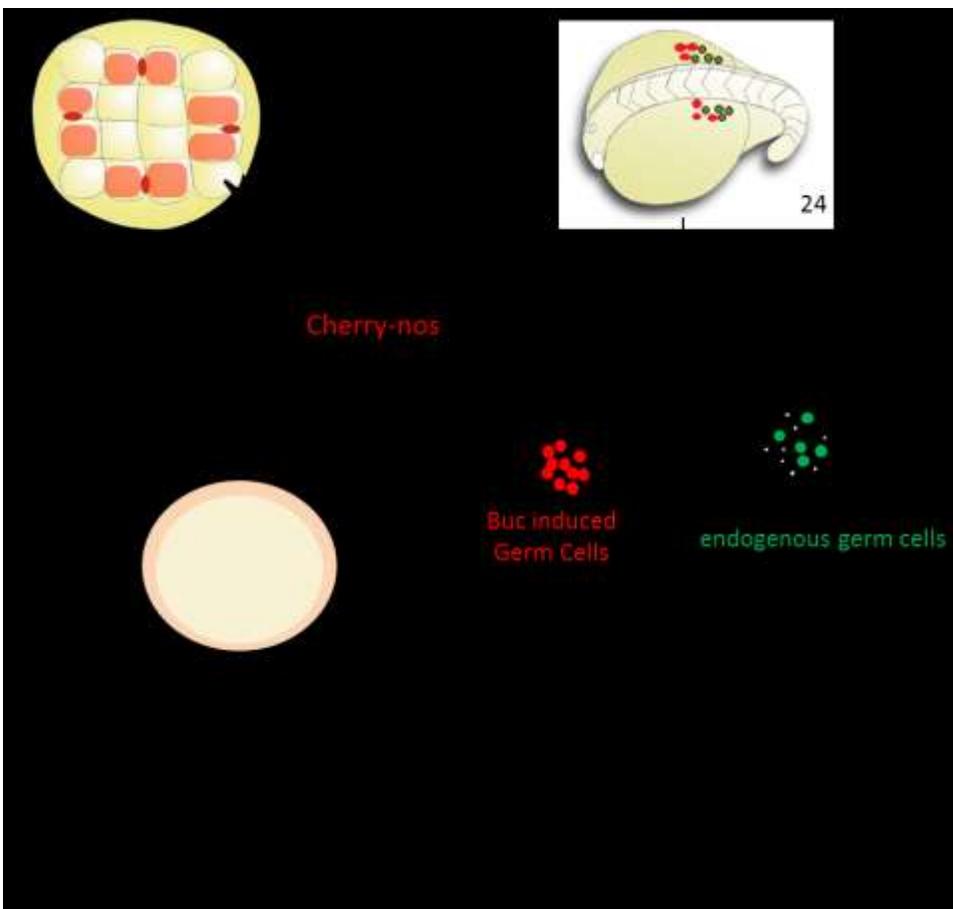


**Figure 30: Scheme outlining the experiment to mark the ectopic cells and transfer them**

Scheme outlining the experiment to mark the ectopic cells in a striped fish strain, isolate them and transplant it into a spotted fish strain.

### 3.3 Deciphering the endogenous role of Buc.

*osk* mutants have been shown to lack germ cells (Ephrussi and Lehmann, 1992) which shows that Osk is necessary for germ cell formation. *buc* mutants arrest at the one cell stage(Dosch *et al.*, 2004), making it unfeasible to analyze a potential role in germ cell formation. One solution would be to do a temporal knockdown of the Buc protein by using deGradFP(Caussinus, Kanca and Affolter, 2011), which actively depletes BucGFP after the one-cell stage (Fig 30) thus creating a protein knockout of Buc.

The BucGFP line as described in the introduction is a functional copy of BucGFP rescuing the mutant line. Hence the only functional copy of Buc is tagged with a GFP moiety. This is where the deGradFP tool will be useful. deGradFP is a system developed in *Drosophila* (Caussinus, Kanca and Affolter, 2011)and has been adapted in the zebrafish by our lab (unpublished data). This tool can deplete maternally tagged GFP proteins in the zebrafish. deGradFP contains a Ubquitin binding element and a GFP nanobody. When the nanobody binds the GFP tagged protein the entire complex gets signaled for degradation.
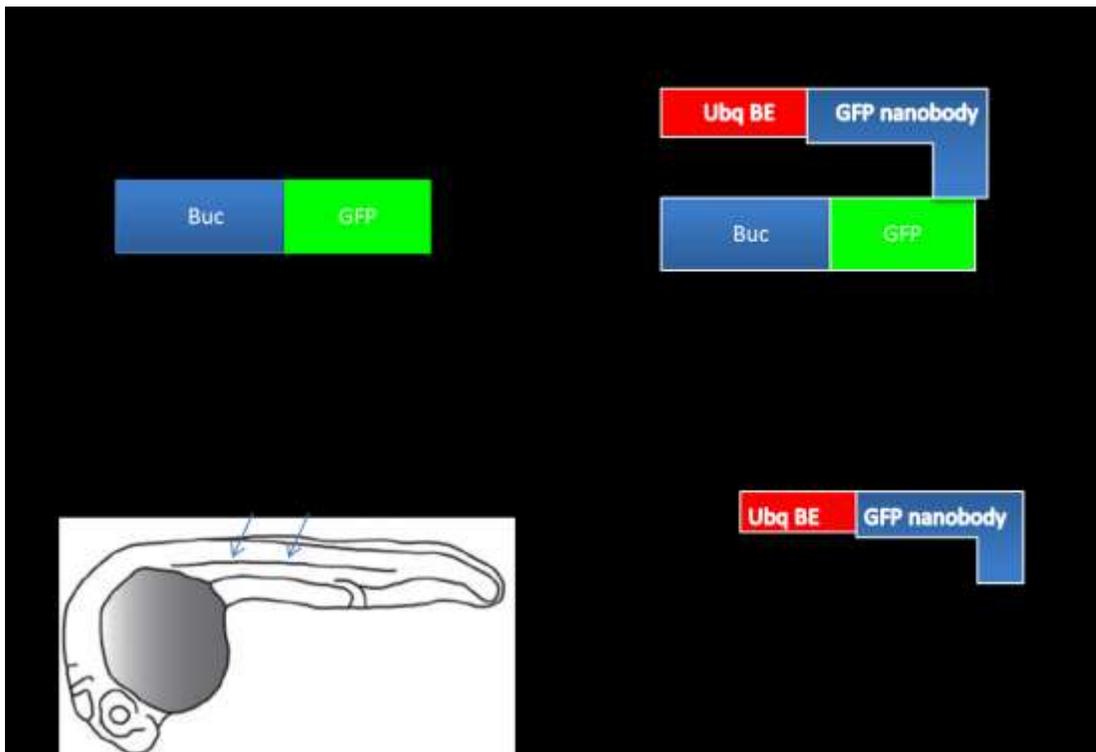


**Figure 31: Mode of Action of deGradFP**

(1) BucGFP embryo is injected at the one cell stage. (2) Injected mRNA codes for a Ubquitin binding element and a GFP nanobody which are the individual parts of the deGradFP. (3) This protein degrades BucGFP leaving only the mutant Buc p106 copy. (4) Hence once the functional protein is depleted the embryo will not have germ cells.

Using this tool, one can deplete the active BucGFP protein and check for germ cell formation in the 24hpf embryo (staining for Vasa and other germ cell markers). Since Buc is sufficient for germ cell formation the idea would be that it could be also necessary for it as well. Hence, the resulting embryos

would potentially be sterile and would give rise to sterile fish. This would conclusively prove that Buc is necessary for germ cell formation like Osk in *Drosophila* and also the endogenous role of Buc.

The advantages of the deGradFP tool would be:

1. It could deplete GFP fusion proteins, in transgenic maternal-GFP lines which would result in a phenocopy of loss-of function mutants. Depletion could be visualized easily by loss of fluorescence.
2. It can be applied in a temporal manner and hence after the critical lethality period of maternal protein mutants.
3. deGradFP could be also integrated into the genome using the tol 2 system (Kawakami, 2007). The deGradFP construct could be fused downstream of a tissue or stage-specific promotor to the gene of interest promoter. Hence, when the gene of interest switches on, deGradFP becomes active to degrade the GFP tagged product of the gene of interest.

If the degradation of the BucGFP proves to be tenuous one can try **either** one of the below listed changes.

1) Inject the deGradFP protein instead of the RNA. This could directly and act early enough on BucGFP instead of the delay, because of its need for translation.
2) Inject the RNA into the BucGFP oocyte and fertilize it (method described in(Nair, Lindeman and Pelegri, 2013)) if the protein proves to be difficult to purify. Injecting it early enough gives it time for translation.
3) deGradFP injection along with a mild treatment of Hexanediol could partially disassemble the aggregates and deplete the BucGFP.
4) A co-injection of the Buc morpholino and deGradFP that would both knockdown RNA and protein in the embryo, hence no continuous supply of protein would be generated.

**3.4 Comparison between Oskar and Buc/Buc homologs.**

Buc and sOsk show striking similarities at the genetic level regarding germ plasm formation. Mutant embryos for both factors show a defect in polarity and a failure of germ plasm aggregation(Lehmann and Nuesslein-Volhard, 1986; Ephrussi, Dickinson and Lehmann, 1991; Marlow and Mullins, 2008; Bontems *et al.*, 2009). Remarkably, ectopic overexpression of sOsk and Buc induces the formation of germ cells (Bontems *et al.*, 2009; Ephrussi and Lehmann, 1992). To this current date, no other proteins have been described, which can induce PGC formation in an organism.

Looking at the two germ plasm organizers the following comparison can be made:

Table 4: Comparing the similarities and differences of Oskar and Buc.

| Oskar | Buc and its Homologs |
|---|---|
| Induction of ectopic germ cells. (Ephrussi and Lehman 1992) | Induction of ectopic germ cells. (Bontems et al 2009) |
| It is RNA binding.(Jeske et al 2015, Yang et al 2015) | Xvelo has been shown to bind RNA (Boke et al 2016). Buc, in my study has been shown to bind RNA. |
| It binds Vasa (Jeske et al 2015 and 2017) and Zf Vasa (Fig 27A) | I have shown that Buc binds ZfVasa. (Fig 24,27A) |
| Contains IDRs (Jeske 2015 and 2017) and prediction in Fig. 18 and 19. | I have shown it contains IDRs and forms aggregates. (Fig 18,19,22.) |
| Has two isoforms long and short Oskar due to alternative translation. | Xvelo has two splice variants, long and short (Nijjar and Woodland 2013). |
| Oskar has been shown to multimerize and bind itself (Jeske 2015). | Xvelo has been shown to multimerize with itself (Nijjar and Woodland 2013). |
| *osk* mRNA localizes to the germ plasm in early embryos (Kim Ha et al 1991). | *buc* mRNA localization to the germ plasm is no longer present during late oogenesis (Bontems et al 2009). |
| Oskar is bound by Lasp (actin-binding protein) which localizes it at the posterior pole of the embryo (Suyama et al 2009). | Buc binds Kinesin-1(Campbell et al 2015) and non-muscle myosin II (NMII) (Riemer Thesis 2014) which recruits it to the cleavage furrow. Both are motor proteins, however only NMII binds actin. |

(Kim-Ha, Smith and Macdonald, 1991; Ephrussi and Lehmann, 1992; Bontems *et al.*, 2009; Suyama *et al.*, 2009; Nijjar and Woodland, 2013; Riemer, 2014; Jeske, Müller and Ephrussi, 2017)

**3.5 The domain of the Buc responsible for interaction with Vasa.**

Osk has been shown to directly bind Vasa (Jeske *et al.*, 2015). Unlike Osk, Buc does not have a LOTUS domain, which is involved in Vasa binding. However my results show that Buc and Vasa interact with each other. This interaction could be direct or indirect. One concurrent method to show that the binding is direct would be to purify both proteins from a bacterial system and check if they interact in pull-down assays. However Buc is very difficult express in bacteria and purify it.

To solve the problem of insoluble Buc, it would be interesting to precisely identify the domain of Buc responsible for binding to Vasa. A smaller domain has a higher chance to be soluble and hence, purifying that domain should be easier. I showed that Buc (1-361) can bind Vasa (Fig 27A). This part would be an excellent starting point to generate shorter variants, which are then tested in a cell free lysate system for binding to Vasa. Even if the binding is not direct like in Osk, Buc might have evolved to activate Vasa activity.

## 3.6 Buc can cause activation of Vasa

Osk has recently been shown to activate Vasa activity (Jeske, Müller and Ephrussi, 2017). Vasa is an RNA helicase with ATPase activity and in the presence of Osk it shows increased ATP hydrolysis (Jeske, Müller and Ephrussi, 2017). The same could be tested for Buc in a cell free system with Vasa, dsRNA and ATP; the presence or absence of Buc would indicate the rate of ATP hydrolysis showing Buc activates Vasa. This would suggest Buc like Osk can increase Vasa activity.

Vasa is ubiquitous at the 3hpf stage (stage during germ cell specification) and Buc is present only in selected cells (germ cells). Vasa might have a different activity in somatic cells since Buc is not present to modulate or activate it. However, in the germ cells Buc is present to ´switch on´ a different or an accelerated function of Vasa which specifies a different cell fate. This would also explain my data of the antibody staining of Buc, Vasa (Fig 25) and why Buc and Vasa need not co-localize in every cell.

Probably artificially localize/restricting Vasa in the 16 cell assay in high amounts indicates that it might have ´kicks started´ its own activity without the presence of Buc. This may be the reason of ectopic germ cells (Fig 26).

The protein interaction of Buc and Vasa seems to be mRNA independent. Treatment with RNAase did not affect the pulldown (Fig 29C). Nevertheless, it could be that if the RNA is present, it could be shielded by the two proteins binding to neighboring sites in the mRNA.

## 3.7 Hierarchy of Buc, Vasa in the germ cell specification pathway.

In Drosophila Osk localizes Vasa and activates it (Breitwieser *et al.*, 1996; Jeske, Müller and Ephrussi, 2017) and *vas$^{4617}$* embryos do not show germ cell formation (Johnstone, 2004). This stands to reason that Osk is higher on the hierarchy in activating germ cell formation and Vasa is downstream. Probably this is the case in the zebrafish too, where Buc is upsteam of Vasa.

To study this hierarchy in the zebrafish after fertilization is technically very challenging. One solution might be to use the deGradFP, CRISPR-Cas9 system and results from the experiment described in **3.3**. Using CRISPR-Cas9 it is possible to insert a GFP coding region into the endogenous Vasa coding locus in the zebrafish(Irion, Krauss and Nusslein-Volhard, 2014). Hence all Vasa protein molecules would have a GFP moiety. Then at the one cell stage when the mother has already deposited the Vasa-GFP molecules, the deGradFP mRNA can be injected and deplete the Vasa protein (Fig 32). If the fish larvae at 24hpf have no germ cells and the result from **3.3** does show germ cells, it would suggest that Vasa is downstream of Buc.
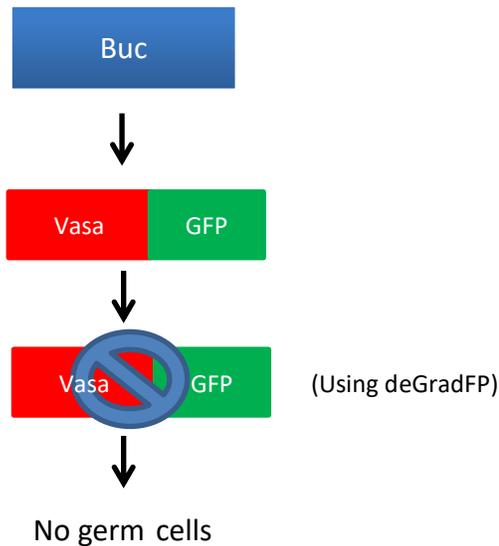
**Figure 32: Hierarchy of Buc , Vasa in specifying germ cells**

If Buc is upstream of Vasa, depleting Vasa tagged GFP with deGradFP will not show germ cells in the developing fish.

### 3.8 Osk binds Dm *nanos* and both Osk, Buc bind *nanos3*

Osk has been shown to directly bind Drosophila *nanos3-3'UTR* (Yang *et al.*, 2015). In an *in vitro* system Xvelo was shown to bind the *nanos3* homolog in the frog (Boke *et al.*, 2016). I discovered that both Osk and Buc are able to interact with the zebrafish counterpart *nanos3 3ÚTR*. Currently it is not clear whether the binding of Buc and *nanos* is direct. However, this can be tested by UV- crosslinking experiments coupled with immunoprecipitation against the protein and *nanos* in an expression system (embryo or in cell culture). Hence, if the protein yields *nanos* RNA a direct interaction can be determined. To calculate the exact stoichiometric of the interaction Isothermal titration calorimetry (ITC) can be performed to quantify the number of protein molecules binding RNA, however this would also need the full Buc purified protein which is insoluble. To get around the problem one would need to identify the domain of Buc binding *nanos3* 3´UTR. I have shown that Buc (aa 1-361) does not bind *nanos3* 3´UTR, therefore the responsible domain is the between (aa 361-639). Using my experimental setup for protein-RNA precipitation the domain responsible in Buc can be delineated and used in ITC experiments. Thus this results would add to the expanding list of protein domains that can bind RNA.

### 3.9 Possible function of Buc bind the *nanos3* 3´UTR

The 3´UTR of RNA has an important function in the regulation and stability of the RNAs(Iwai *et al.*, 1991). Micro RNAs bind to partial complementary sequences in the 3′UTR regions of mRNAs to

decrease their expression by inhibiting translation and occasionally the decay of that mRNA(Bagga *et al.*, 2005; Lim *et al.*, 2005; Pillai, 2005).A protein called Dead end competes with miRNA for a binding site in the *3´UTR* of *nanos3* and stabilizes it in germ cells(Kedde *et al.*, 2007). The *3´UTR* of the Xenopus *nanos* is responsible for localizing it correctly(Kosaka *et al.*, 2007). Hence binding of the two germ plasm organizers to the 3´UTR must provide a vital function in stabilizing the *nanos3* RNA or increasing its expression. It could also be involved in bridging Vasa and Nanos for Vasa protein to act upon it, in an RNA unwinding or editing activity.

## 3.10 Minimum domains of Buc sufficient to specify germ cells.

Narrowing down and identifying the domains in Buc responsible for binding Vasa and Nanos will also point to the regions sufficient in Buc to specify germ cells (Fig 33).
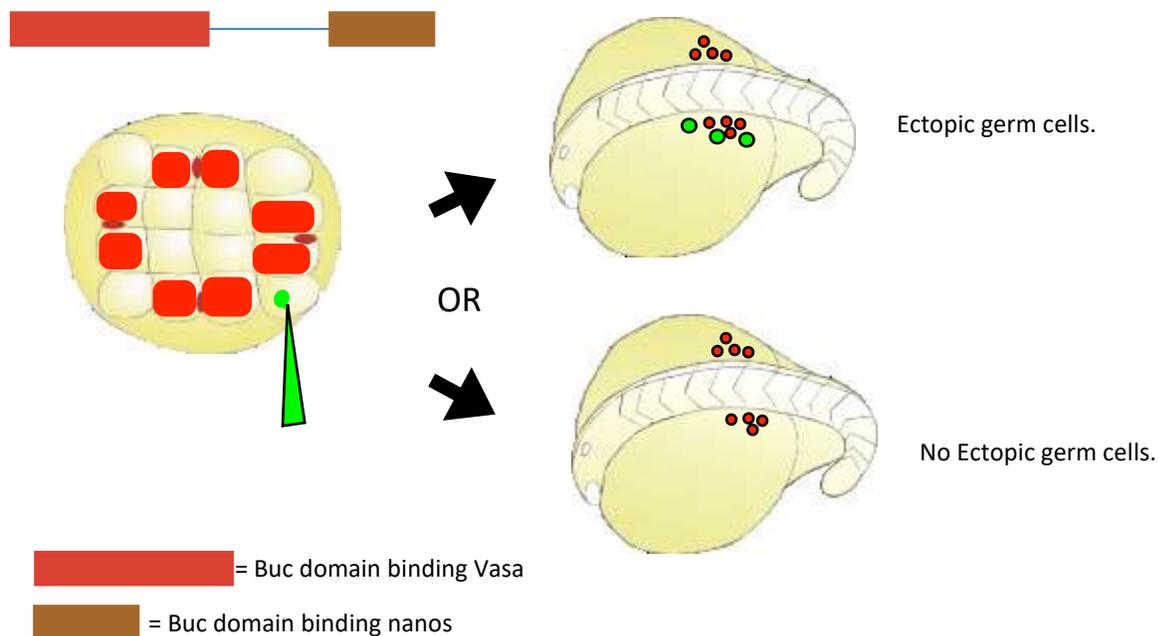


**Figure 33: Sufficiency of Buc domains to specify a germ cell**

To check if the Vasa binding domain of Buc (vermillion) and *nanos3* binding domain of Buc (brown), injected into the corner blastomere is sufficient to induce germ cells via the 16 cell assay.

In summary, this study has identified a conserved ´core complex´, which is required for germ cell specification.  This complex includes the conserved germline components Vasa protein (Raz E  2000 and Lasko P 2013) and *nanos* mRNA as well as a germ plasm organizer like Osk or Buc.  These molecules are probably not the only components of the complex and might contain additional proteins or RNAs, since numerous, canonical germ plasm components are conserved in metazoan genomes(Ewen-Campen, Schwager and Extavour, 2010; Juliano, Swartz and Wessel, 2010).

In summary, from the data regarding the conservation of the Osk and Buc interactome, one can come to the conclusion of the following scheme (Scheme below).

## Biochemical interactions common between Osk and Buc



sOsk interacts with Dm Vasa

Buc interacts with Zf Vasa

sOsk activates Vasa ↔ Osk binds *nanos 3´UTR*

Buc activates Vasa? ↔ Buc binds *nanos3 3´UTR*

Germ cells specified
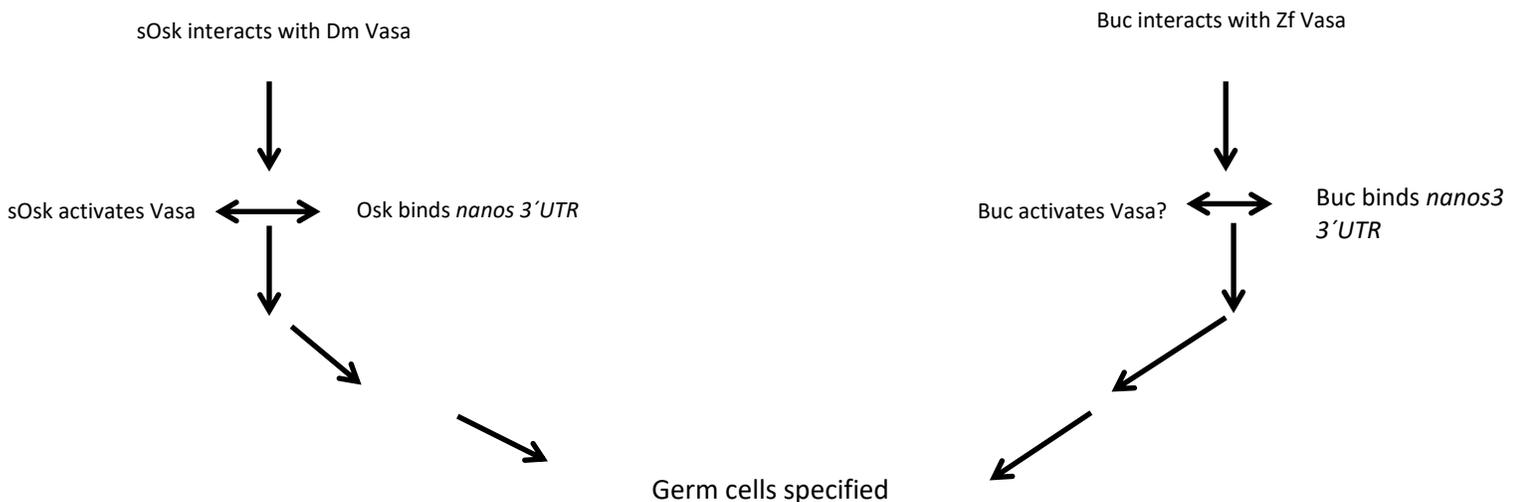
**Figure 34: Mode of action of Osk and Buc**

Osk and Buc have an overlap in their biochemical network, which they use to form germ plasm and specify germ cells.

**3.11 Buc and Oskar have IDR domains within their protein which helps them function.**

As shown in Fig 18. Buc and Osk have IDRs in their proteins. These IDRs seem to be responsible in obscuring sequence similarity between the two proteins which would explain the problem to align the two proteins. As they are fast evolving, the similarity of binding domains will be abolished. This makes it a problem when performing bioinformatics analysis to identify similar proteins; on the other hand it relaxes the constraint on proteins in order to bind the same partners.

When comparing germ plasm members across various species, the aggregation of proteins seems to emerge as a central theme when initiating germ cell specification. Vasa in *Drosophila* and zebrafish, which has intrinsically disordered (IDRs), shows aggregating and phase transition (Nott *et al.*, 2015)(Nott et al., 2015). P-granules which contain MEG1 and MEG 3 proteins in *C.elegans* demonstrate aggregating properties(Wang *et al.*, 2014). Buc and Osk tend to self- aggregate. When reaching a certain threshold all these aggregators undergo phase transitions and assemble a hydrogel. That's probably the

main key to have a separate compartment with all the proteins and RNAs. This compartment probably protects the RNAs and regulates their expression thus specifying the germ cell fate.

### 3.12 1,6-Hexanediol assay as a tool to study Buc-GFP aggregation dynamics.

To disrupt BucGFP, the hydrogel disrupter 1,6-hexanediol was used to separate the aggregation. In this case, treatment followed immediately after dechorionation, which usually take about 20-30 min. Most embryos were still in 1-cell stage at the beginning of treatment. This means that hexanediol will inhibit the aggregation throughout the first and maybe the second cell cleavage. This already resulted in fragmented BucGFP aggregates. This would explain why there are more numerous and larger aggregates in treatment than in the control (Fig 20C). Starting treatment at a later time point in development, like at 4-cell stage, could show fragmentation as well as smaller aggregates like shown in Fig 35.
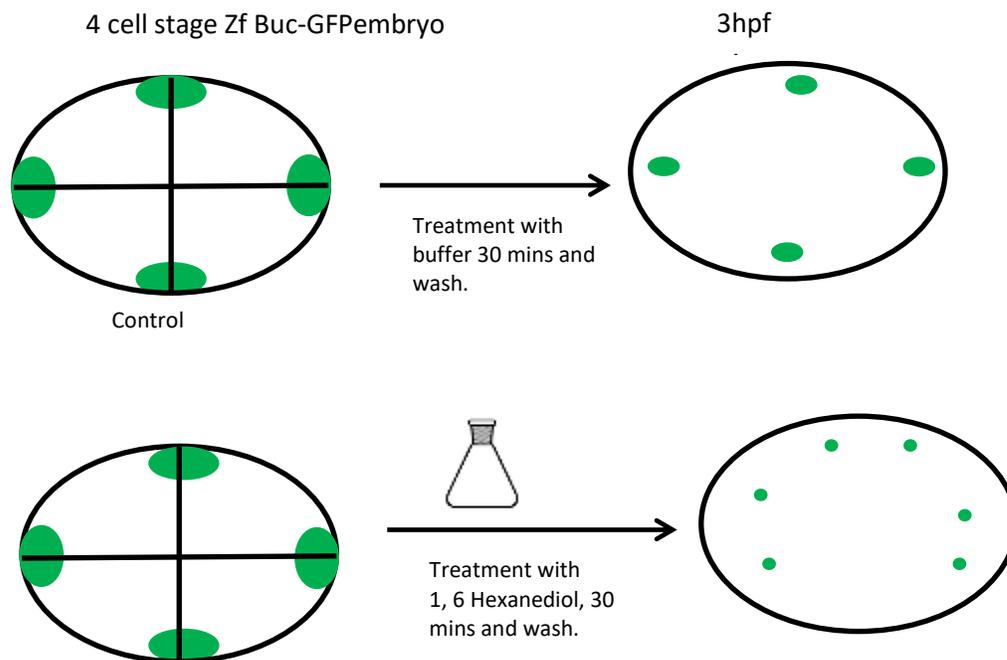
**Figure 35: Hexandiol treatment of Buc-eGFP embryos**

At the 4 cell stage in the zebrafish Buc-GFP embryo which is at 1 hpf give the respective treatments of buffer (for the control) and Hexanediol. Wash after 30mins and image embryo at 3hpf. Possibly the treated embryo, would show smaller aggregates than its control.

One other reason for the extra aggregates could be because of symmetrical inheritance. Recently an experiment in yeast, showed that aggregates tend to re-assemble after their dissolution by

hexanediol(Bolognesi *et al.*, 2016). Infact, the aggregates get symmetrically inherited by daughter cells, whereas the untreated control aggregates are only inherited to one daughter cell(Bolognesi *et al.*, 2016). A similar phenomenon might be occurring in the fish embryo which would lead to the increased number of aggregates (Fig 20 C).

Another experimental possibility would be to raise embryos treated with 1,6-hexanediol to maturity to see whether the disruption (and therefore dilution) of the aggregate components (i.e. germ plasm) has an influence on germ line development. In untreated embryos, only four cells will inherit the bulk of germ plasm present in the embryo. Disruption and subsequent dilution of germ plasm component, including BucGFP, could have one of two outcomes: either more PGCs would be specified, since more cells inherit germ plasm. Contrary, dilution of germ plasm could lead to fewer PGCs, since the critical threshold (i.e. concentration of germ plasm) necessary for germ cell specification would not be reached. The later outcome, however, would give no evidence for the necessity of *buc* to specify germ cells, but would rather show that there is a critical threshold or concentration for germ plasm to be able to specify a given cell as a germ cell. However, though hexanediol is the best available drug to interfere with aggregation, it is unspecific and disrupts all hydrogels in the cell (*eg.* nuclear pore protein complexes). The embryos might all die during later development making it difficult to check number of germ cells in older embryos. Therefore a mild and short treatment could be optimized to ensure embryo survival.

**3.13 Potential assays to identify germ plasm organizers / similar proteins.**

Looking at the two organizers Buc and Osk, as shown in this study both self-aggregate, bind Vasa and bind *nanos*. These three must be the minimum characteristics to possess to qualify as a germ plasm organizer. Other molecules which are identified in the future as potential germ plasm organizers could be subjected to tests to check whether they fulfill these criteria. For one I) they could be injected in the 16-cell assay and check for germ cell induction. II) Next if they bind Vasa/*nanos* indirectly or directly with a pulldown. Third, III) do they have the tendency to self-aggregate when transfected as cDNA into cells or even injected as mRNA or protein in an *in vivo* system. I have stated the above criteria because germ plasm organizers cannot be identified by current bioinformatics based simply on protein sequence. Only a database comparing genetic phenotypes with gain and loss-of-function experiments would detect these functional similarities.

## 3.14 Model or recipe to make germ cells.
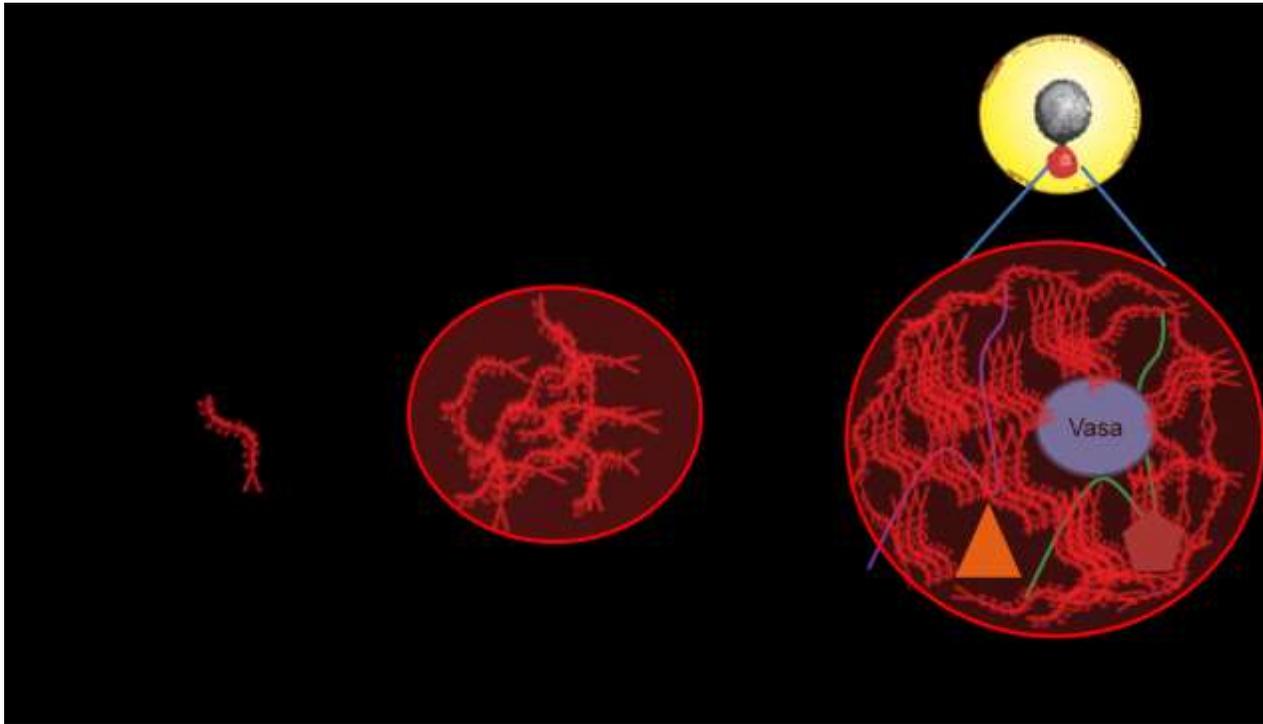


**Figure 36: Model**

Model for germ plasm formation. Single monomer molecules of germ plasm organizer (red) like Buc and Osk aggregate through weak interactions of their intrinsically disordered regions (hooks and loops), until a threshold concentration is reached. This leads to a liquid-liquid phase separation (red haze) to form hydrogel-like germ plasm. The aggregate then selectively recruits proteins (geometric shapes) and mRNAs like *nanos3* (lines). This gets packed into germ plasm *e.g.* as shown above in the Balbiani body of the oocyte. The same principle would apply when this germ plasm gets inherited into a cell, specifying that cell as a germ cell.

If Osk and Buc have diverged from a common ancestor, their precursor would have been an ancient protein of low complexity, which induces germ cell formation. Both proteins probably have unrelated sequences as consequence of their role as intrinsically disordered scaffolds. This structural role releases the constraints to maintain a defined protein structure as described for other IDPs(Forman-Kay and Mittag, 2013). This divergence probably hides conserved motifs, which bind to a similar interactome such as Vasa, Valois and probably other common mRNA binding partners. Finding interaction partners and mapping the interaction motifs like for the sOsk-Vasa interaction will determine, to which level interaction motifs are conserved between sOsk and Buc.

In the Fig. 36, the organizer proteins like Buc and Osk form a scaffold probably via self-aggregation or upon binding with their interactors, which form during polymerization. During this process, germ plasm assembles and thereby integrates RNA like *nanos* and proteins into this 3D liquid lattice. The assembly also initiates Vasa's activity to start the downstream program e.g. to protect RNAs and proteins from degradation (Aguero *et al.*, 2017). When inherited into a cell, the germ plasm probably releases some proteins whose translation and stability is tightly controlled. Once these factors are unleashed from the plasm spheres, they change the transcriptional program to specify the maturation of a PGC to a gamete.

## 3.15 Origin of life: A germ plasm perspective

When animals started to become multi-cellular, they needed to set the germline apart from the soma for their reproduction (Grosberg and Strathmann, 2007). For this task, they had to evolve proteins, which served as master switches for germ cell specification. Any changes to the function of these proteins could have deleterious consequences on the propagation of that species. However if these proteins were IDPs, they could still perform their function, despite of rapid (localized or random) changes. These changes could have roles in speciation or better coordinated control of specification. Whichever the case, if they still aggregated and setup the ´core´ complex, a germ cell would have still formed.

This hypothesis is supported by looking at unicellular organisms. Ciliates form a cytoplasmic aggregate called the conjusome(Janetopoulos *et al.*, 1999). This structure is present only during sexual reproduction. Similar to the Balbiani body in Xenopus and the P-granules in C. elegans, the conjusome is made up of fibrous, electron dense material (Janetopoulos *et al*., 1999). It also contains a Piwi related protein TWI, which protects the integrity of the genome as in metazoans (Mochizuki *et al.*, 2002). These commonalities with germ plasm are very striking and suggest that a conjusome-like structure might be the ancestral form of germ plasm. Hence, the organizer protein of ciliates probably displays a very different amino acid sequence from Osk and Buc. However, the Ciliate organizer might have similar characteristics like Osk and Buc like forming the protein-RNA core or even induce germ cells in zebrafish. If indeed a germ plasm like structure existed in unicellular organisms, germ cell specification by induction could only have emerged after the transition to multicellularity, because signaling requires a multicellular environment. This would challenge the current dogma that induction is the ancestral form of germ cell formation. It will therefore be quite interesting to find out to which level germ plasm in metazoans and structures like the conjusome in unicellular organisms are conserved.

Why should germ cell specification be conserved? Reproduction is a conserved feature of all biological systems and must have been therefore be present in the first metazoans, which did not use other cell types like neurons, muscle or a vascular system. Germ cell specification was therefore present before the formation of an eye or even a nervous system. Nonetheless, the conservation of the master regulator Pax6/Eyeless showed that light sensing organs were already present at the base of metazoan evolution

(Gehring, 2005). Although this hallmark finding is currently accepted in the scientific literature, the insect compound eye and the vertebrate camera-eye were regarded as a paradigm for convergent adaptations. Therefore one can speculate that germ cell formation as the more ancient tissue compared to eyes, would use an even more conserved molecular regulation than Pax6/Eyeless.

If the beginning of life was an RNA world (Cech, 2010) and formation of a cell was needed to protect the genetic material, it would have been easier to have a hydrogel aggregate of slime or protein lock the RNA into an RNA granule like a primitive germ plasm than to establish a lipid bilayer with an internal framework. Indeed if that was the case, this structure at the origin of life would have been more similar to the germ plasm that we see today than to a membrane-bound cell.

# 4. **Conclusions**

In summary, a fascinating finding in this study is that Osk and Buc share many common features such as conserved biochemical interactions despite the lack of recognizable sequence homologies. Two scenarios could explain this functional conservation. In the first scenario, the evolution of both proteins happened in parallel. At a certain time they converged by recruiting a similar interactome during evolution. Such a model is more likely for somatic tissues, where the loss of an organ might not give an evolutionary dead-end. However if drastic errors where added within the evolving germ plasm organizer, it is likely that it would result in a failure to reproduce. Eventually the entire species would be extinct Hence, the independently evolving germ plasm organizers showing convergence becomes increasingly complex to explain.

In the second scenario, Osk and Buc have diverged from a common ancestor. They probably have highly unrelated sequences, because their role as IDPs scaffolds relaxes their constraints to maintain a defined protein structure (Forman-Kay and Mittag, 2013). This divergence probably obscures conserved motifs, which bind to a similar interactome such as Vasa and nanos mRNA. Osk and Buc could only be termed true homologs, if the identity of a common ancestor is known. Moreover the structural identity of the Vasa and the nanos mRNA binding domain would be similar. However, without this information, the conservation of their activity is very exciting and important addition to the germ cell specification field. In a more global perspective, Osk and Buc could represent a widespread phenomenon, in which probably additional protein-pairs exist, which are currently listed as novel or species-specific in genome databases, but could be functional homologs across distant species.

# 5. Material and Methods

## 5.1 Zebrafish

Fish were maintained as described (Westerfield, 2000) in accordance with regulations of the Georg-August University Goettingen, Germany. Injections were performed into wild-type embryos (hybrid of *ABxTLF) and *buc-egfp* embryos.

## 5.2 Microinjection of embryos

A Needle puller (model PN-30; Narishige, Japan) was used to make glass needles for injection (GB100F-8P; Science Products, Hofheim). The synthesized capped sense RNA (Section 5.8.1) was diluted with 0.1 M KCl and 0.05% phenol red (Sigma-Aldrich, Hannover). The one cell injection was done to check expression of the *buc* mutant RNAs (100pg/nl each) was injected in the wildtype embryo. $buc^{p43}$ was injected into the buc-egfp embryos to test if it could disrupt germ cell formation. 2nl was injected into the embryo for the one cell injection.

16-cell stage zebrafish embryos were injected with 0.5 nl capped sense RNA. 16-cell embryos were injected as previously described (Bontems, et al. 2009). At least 20 embryos were sorted per injection and for biological replicates independent egg lays were used. One blastomere was injected with 0.5 nl RNA-solution containing 100 pg/nl of PGC-reporter (*gfp nos 3′UTR* ) plus 100 pg/nl mRNA encoding a germ plasm component. *buc* and short *osk* mRNAs contained their 5′ and 3′UTR sequences, respectively. $buc^{p43}$ and $buc^{p106}$ were identical to wt mRNA except for a premature stop codon in 361 aa and 601 aa, respectively (Bontems, et al. 2009). Short *osk* and $osk^{084}$ mRNA were identical except for the premature stop codon in $osk^{084}$ (Kim-Ha, et al. 1991).

## 5.3 Dechorionation of embryos

For the 16 cell assay or to prepare embryo lysates an entire clutch of ~ 200 of embryos were enzymatically dechorionated using pronase (30 mg/ml; Roche, Mannheim). The embryos were incubated 3-5 min in pronase solution (3 mg/ml in 1x E3 medium which contains 5 mM NaCl, 0.17 mM KCl, 0.33 mM CaCl2, 0.33 mM MgSO4, 0.00001 % methylene blue). Later the embryos along with the dish were immersed into a beaker of 300ml E3 medium. The embryos sank to the bottom while the empty chorions floated on top. The E3 was discarded (leaving some behind just enough to cover the embryos) and refilled again till the 300ml mark (Washing the embryos). This washing was done three times. A fire polished Pasteur pipette was used to transfer the embryos into an agar coated dish so as to

not damage the embryos. For imaging the embryos were mechanically dechorionated with sharp forceps.

## 5.4 Deyolking embryos

The embryos (3hpf stage) were deyolked by the following chemical procedure to remove most of the yolk (Link, Shevchenko and Heisenberg, 2006). The following buffer 1/2 Ginzburg Fish Ringer with Calcium (55 mM NaCl, 2.7 mM CaCl2, 1.8 mM KCl, 1.25 mM NaHCO3) was used to lyze the dechorionated embryos. After pelleting the lysate it was washed with deyolking buffer 3 times (110 mM NaCl, 3.5 mM KCl, 2.7 mM CaCl2, 10 mM Tris pH 8.5). The washed cell pellet was directly used for further experiments.

## 5.5 Preparation of embryo lysates

The embryo lysates were were resuspended in 2x SDS loading buffer (100 mM Tris pH 6.8, 20 % glycerol, 4 % SDS, 200 mM β-mercaptoethanol, 0.02 % bromophenol blue). Then it was boiled at 95 °C for 10 min. The sample was then cooled on ice for 5 mins, then stored at -20°C.

## 5.6 1,6-hexandiol treatment

Dechorionated embryos Buc GFP transgenic embryos at the one cell stage were treated with 1,6-hexandiol (5% w/v in E3-medium) for 30 min, whereas control treated embryos were exposed to E3-medium.  Both were incubated at 28.5 °C in glass dishes for 30min and then transferred into fresh E3-medium in agarose coated dishes at 28.5 °C until 3hpf.

### 5.6.1 Live-Embryo Imaging

Transgenic *buc-gfp* line and embryos from the 16 cell injection assay were imaged with the stereo microscope SteREO Lumar.V12 (Carl Zeiss Microscopy, Jena). Images analysis was done using the Axio Vision Rel. 4.8 software (Carl Zeiss Microscopy, Jena).

## 5.7  Protein biochemistry

### 5.7.1  Proteomics of *buc-egfp* embryos and *egfp* injected embryos, Mass spectrometry and Mass spectrometry Data analysis

Performed as described in (Riemer, 2014).

### 5.7.2  Co-immunoprecipitation

The embryo cell pellet from section 4.4 was used. The cell pellet were homogenized on ice in lysis buffer (10 mM Tris pH 7.5, 150 mM NaCl, 0.5 mM EDTA, 0.5% NP-40, 1x complete protease inhibitor cocktail (Roche Mannheim) and centrifuged. After centrifugation the cell debris were discarded. The supernatant was incubated with GFP-Trap magnetic beads (ChromoTek) for 3hrs at 4°C.  Beads were then washed 4 times with dilution buffer (10 mM Tris pH 7.5, 150 mM NaCl, 0.5 mM EDTA). Beads were then resuspended in SDS-loading buffer, boiled at 95 °C for 10 min and stored at -20°C for gel electrophoresis.

### 5.7.3  Protein in vitro analysis
The Promega TnT® SP6 Quick Coupled Transcription/Translation System was used to synthesize proteins.  Products were diluted in YSS buffer (50 mM Tris pH 8, 75 mM NaCl, 1mM MgCl$_2$, 100mM Sucrose 5 mM EDTA, 0.05% NP-40, 1mMDTT and ¼ x tablet of complete protease inhibitor cocktail (Roche Mannheim)). For pull-downs GFP Trap beads were used as described above and loaded later on a western blot.  In case of the RNAse (A and T1) treatment experiment it was done as shown in Fig.29B. The pulldown was done after the treatment.

### 5.7.4  SDS - polyacrylamide gel electrophoresis (PAGE)
A 10% SDS PAGE gel was used to separate proteins under denaturating conditions. Depending on the ongoing experiment samples from section 4.5/ 4.7.1/4.7.2 were loaded into a precleared gel well. The Page Ruler prestained plus protein ladder (Thermo Scientific, USA) was also loaded into a separate well to determine the molecular weight of the proteins (Kd). The gels were run vertically in 1x Laemmli

buffer (25 mM Tris, 250 mM glycine, 0.01 % SDS) at a constant voltage of 70 V. When the running front reached the resolving gel, the voltage was raised to 130 V.

### 5.7.5 Protein Transfer

This was done by via a semi dry transfer method using the Biometra B44 fastblot instrument. The nitrocellulose membrane and Whatmann papers were presoaked in blotting buffer (50 mM glycine, 50 mM Tris, 0.037 % SDS, 20 % methanol) and the gel was sandwiched between the stack with the membrane facing towards the positive electrode. The semi-dry blotting for carried out for 70 min at 25V.

### 5.7.6 Western blot

Blots were blocked in 5% TBS/milk for 1 hr, incubated overnight with primary antibody at 4°C, washed with 3 times with TBST (20 mM Tris pH 8.0, 150 mM NaCl, 0.05 % Tween20), incubated with Li-Cor IRDye secondary antibody for 1 hr (protected from light). Washed again with TBST twice. The Li-Cor Odyssey CLx Infrared Imaging system was used to detect bands on the blot (Li-Cor, Lincoln, USA). Antibodies: rabbit anti-Vasa (1:500; Genetex and gift from Vorbrueggen, Univ of Goettingen), guinea-pig anti-Buc (1:5,000; BioGenes, Berlin, Riemer et al 2015), goat anti-rabbit (1:20,000; IRDye, Li-Cor), goat anti-guinea pig 800CW (1:20,000; IRDye, Li-Cor).

### 5.7.7 Protein aggregation assay

HEK cells ($10^4$/per well) in an eight-chambered slide (Sarstedt) were grown in Dulbeccos modified Eagle Medium (DMEM) (Biochrom). 24 hrs later they were transfected using ScreenFect®A reagent with 100ng of the plasmids of interest. Cells were imaged after 48hrs using 10X objective with a 10X digital zoom with an LSM780 confocal microscope and cell profiles were analyzed with ZEN2011 software (Carl Zeiss Microscopy, Göttingen).

### 5.7.8 Fixation of zebrafish embryos for immunostaining

Embryos were collected at the desired stage. Dechorionation by pronase treatment according to section 4.3 and fixed. Fixation was done at 4 °C overnight in 4 % paraformaldehyde in PBS (10 mM Na2HPO4, 1.76 mM KH2PO4  pH 7.4, 137 mM NaCl, 2.7 mM KCl,) in 5 ml glass vials. The embryos were washed three times with PBS and dehydrated in a methanol dilution series from (50% to 100% methanol). The embryos were then stored in 100 % methanol at -20 °C in the glass vials. Before the embryos were immunostained, they were rehydrated in PBT (10 mM Na2HPO4, 1.76 mM KH2PO4 pH 7.4, 137 mM NaCl, 2.7 mM KCl, Triton X-100, 0.2% BSA) by replacing the methanol.

### 5.7.9  Immunostaining of zebrafish embryos and Imaging

Immunostaining was performed as described in Riemer *et al* 2015(Riemer *et al.*, 2015). For imaging, the embryos were transferred to a imaging dish (Fluorodish 35 mm; WPI, Sarasota, USA).The methanol was removed and replaced with Murray's clearing medium (2/3 benzyl benzoate, 1/3 benzyl alcohol) that made the yolk transparent. Embryos were imaged using a LSM780 confocal microscope (Carl Zeiss Microscopy, Jena). The Z-stacks generated images were analyzed using the ZEN 2011 software (Carl Zeiss Microscopy, Jena). Antibodies: rabbit anti-Vasa (1:500; Genetex and gift from Vorbrueggen, Univ of Goettingen), guinea-pig anti-Buc (1:5,000; BioGenes, Berlin, Riemer et al 2015)

### 5.8 Molecular biological methods

### 5.8.1  *In vitro* transcription

The microinjection into zebrafish as described in section 5.2 used capped sense RNA. This was synthesized *in vitro* using the SP6 mMessage mMachine kit (Life Technologies, USA).

RNA synthesized was purified using Illustra Probe Quant G-50 columns (GE Healthcare, UK).

### 5.8.2  RNA-Immunoprecipitation

HEK cells ($0.5 \times 10^6$) per well were co-transfected for protein and RNA expression with the indicated combinations of plasmids (table 6). Cells were incubated for 48hrs and wells were visibly screened for expression of GFP and Cherry fluorescence using an inverted fluorescence microscope (Carl Zeiss Microscopy, Jena). Cells were then homogenized in (1 ml) YSS buffer and centrifuged. The supernatant was taken, 20 µl were kept aside as the input fraction and the rest was incubated with GFP nanotrap beads (Chromotek) for 3hrs at 4°C. Beads were washed (YSS buffer) and the bound fraction was released from the beads with 10% SDS. RNA was isolated using phenol/chloroform/isoamylalcohol and precipitated (0.3 M sodium acetate pH 4.5 in ethanol). Precipitated RNA was used for cDNA synthesis in section 5.8.3.

### 5.8.3 RT-PCR (Polymerase Chain Reaction)

RNA was reverse transcribed for first strand synthesis using random hexamers and MuLV RTase according to the NEB ProtoScript. The random hexamers annealed to the RNA for 10 min at 20 °C. The cDNA was synthesized at 42 °C for 60 min and then terminated for 5 min at 95 °C. The cDNA obtained was stored at -20 °C. The required gene from the cDNA was amplified by primers in table 7 .

### 5.8.4 Polymerase chain reaction (PCR)

DNA was amplified using the standard PCR reaction (Mullis *et al.*, 1986). 50-100 ng of DNA was used as template for the reaction. A 50 µl reaction containing 1 Unit of Phusion polymerase (5 U/µl, New England BioLabs, USA), 1x Phusion buffer (New England BioLabs, USA), 0.4 µM of each primer (Table ) (Sigma Aldrich, Hannover), 0.5 mM of each dNTP (Thermo Scientific, USA) . UNOII thermocycler (Biometric, Geneva) was used to run the PCR according to the steps below. Depending on the primers, the temperature for annealing was calculated by the program in the Intergrated DNA technologies server and confirmed with a gradient PCR .35-39 cycles was used depending on the experiment. For cloning 39 cycles was used. For the RNA Immunoprecipitation in 5.8.2, 35 cycles was used to amplify the input and pulldown (control and experimental samples).

Step1: 95°C, 2 min denaturation

Step 2: 95°C, 45 sec denaturation

Step 3: Based on server or gradient(each well with different temperature) PCR reaction, 30 secs annealing

Step 4: 72°C, (0.5 min per Kb length) min elongation, cycle back to step 2

Step 5: 72°C, 10 mins

### 5.8.5 Colony-PCR

A colony-PCR reaction was performed to confirm the length of sequence in bacterial clones after chemical transformation. The bacteria colony was used as template in a 10 µl reaction. It also had 0.4 µM of each primer (Table), 0.2 mM of each dNTP (Thermo Scientific, USA), 1 U of Taq polymerase

(40 U/ml) in 1x PCR buffer (100 mM Tris pH 8.4, 500 mM KCl, 15 mM MgCl2, 1 mg/ml BSA).  The standard PCR program from section 5.8.4 was used for 25 cycles. The UNOII thermocycler was used.

### 5.8.6  Sequence analysis

The Sanger method of sequencing (Dye-termination) was used (Sanger, Nicklen and Coulson, 1977)here. The Big DyeTM Terminator Kit (Applied Biosystems) was used according to manufacturer instructions.

### 5.8.7  Agarose gel electrophoresis

DNA or RNA was separated in an agarose gel. An electrical field was applied (Sharp, Sugden and Sambrook, 1973). Agarose gels were prepared in TBE buffer (90 mM Tris pH 8.0, 90 mM boric acid, 2 mM EDTA) and the electrophoresis took place in the same buffer as well at 100V. The gels contained 0.5 μg/ml of ethidium bromine to detect the nucleic acids. The loading of the samples was done in DNA loading dye (6x, Ambion). A1 kb Plus DNA ladder (Life Technologies, USA) was used to determine the size (Kb). The gels were scanned with the gel documentation system (Bio-Rad, Munich).

### 5.8.8  Chemical transformation of plasmid into bacterial cells

This was performed in bacterial strain DH5 α (Life Technologies, USA). 100ng of plasmid DNA was mixed with one aliquot of 100μl thawed bacteria. This was kept on ice for 30 min and heat shocked for 45 s at 42 °C in a water bath. Then followed by a quick incubation for 2 min on ice. 900 μl LB medium was added and the cells incubated for 80 min at 37 °C and 210 rpm. The 1ml culture was plated onto 1.5 % LB-agar plates (containing the selection antibiotic) and incubated at 37 °C overnight.

### 5.8.9  Plasmid DNA preparation

The NucleoBond Xtra Midi Kit (Macherey and Nagel, Dueren) was used to isolate plasmid DNA, according to the manufacturer's instructions. The DNA was quantified with the NanoDrop 1000c spectrophotometer (Thermo Scientific).

### 5.8.10 Gateway cloning

The Gateway cloning technology (Invitrogen, USA) was used according to the manufacturer's instructions. Reagents provided in the kit were used. Vectors pDONR221 and pCSDEST was used to sequentially transfer the cloned PCR product into the DONR vector then the Destination vector. The Gateway system uses site specific recombination to embed the PCR amplified region of interest into the vector for SP6 driven transcription of RNA.

### 5.8.11 In-fusion cloning

In fusion cloning was used to insert a DNA fragment into the PCS2+ vector. Protocol followed according to manufacturer's instructions (Clonetech, Takara, USA).

### 5.9 Plasmids

Table 5 Plasmids used

| Name | Backbone | Insert | Cloning strategy |
|---|---|---|---|
| pCS2$^+$ buc-eGFP | pCS2$^+$ | *buc-egfp* | (Bontems *et al*., 2009) |
| pCS2$^+$ buc$^{p43}$-eGFP | pCS2$^+$ | *buc$^{p43}$-egfp* | (Riemer, 2014) |
| pCS2$^+$ buc$^{p106}$eGFP | pCS2$^+$ egfp | *buc$^{p106}$-egfp* | *buc$^{p106}$* sequence was amplified from pCS2$^+$buc-eGFP, using the primers BamHI_bucp106_for, Buc-p106_XbaI_rev. The PCR product was inserted into precut (BamHI and XbaI) pCS2$^+$ eGFP by Infusion cloning (gift from C. Niehrs, IMB Mainz). |
| pCS2$^+$ osk-eGFP | pCS2$^+$ | *osk-egfp* | (Riemer, 2014) |
| pCS2$^+$ osk$^{084}$-eGFP | pCS2$^+$ | *osk$^{084}$-egfp* | (Riemer, 2014) |
| pCSDestEx hermes | pCSDEST | *hermes* | *hermes* was amplified from pCS2$^+$GFP-hermes with attB1 and attB2 sites on the forward and reverse primers, respectively. Using the Gateway system it was recombined into pDONR221 vector and then later into pCSDEST vector |
| pCSDestEx vasa | pCSDEST | *vasa* | *vasa* was amplified from pBS vasa with attB1 and attB2 sites on the forward and reverse primers, respectively. Using the Gateway system it was recombined into pDONR221 vector and then later into pCSDEST vector. |

Other plasmids were pCS2$^+$XveloI (gift from M. Claußen & T. Pieler, GZMB Göttingen), pBSvasa, pSP64gfp-nos-3'UTR and pSP64cherry-nos-3´UTR (gift from E. Raz, ZMBE, Münster), pCS2$^+$GFP-hermes (gift from K. Inoue, Japan). All plasmids were verified by sequencing.

## 5.10 Primers

Table 6

| Name | Sequence (5' to 3' direction) | Purpose |
|---|---|---|
| ClaI_buc_fwd | GGGATCGATAATGTGGATCTCTGGAAACAG | pCS2$^+$ plasmids |
| Buc-p43_XbaI_rev | GCTCTAGAGCTGTAGGAATAAGCACTGCC | pCS2$^+$ plasmids |
| Osk_BamHI_fwd | GGGGGATTCATGACCATCATCGAGAGCAAC | pCS2$^+$ plasmids |
| Osk_XbaI_rev | GGGTCTAGAATACTCCAGACTCGTTTCAAT | pCS2$^+$ plasmids |
| 18SrRNA fwd | AACTGAGGCCATGATTAA | 18SrRNA |
| 18SrRNA rev | GGAACTACGACGGTATCTGA | detection |
| attB1 hermes | GGGGACAAGTTTGTACAAAAAAGCAGGCTTAATGAGTGTCAAGTCCGAC | *hermes* cloning |
| attb2 hermes | GGGGACCACTTTGTACAAGAAAGCTGGGTATTAACAGAACTGTCGGGA | |
| attB1 vasa | GGGGACAAGTTTGTACAAAAAAGCAGGCTTAATGGATGACTGGGAGGAA | *vasa* cloning |
| attB2 vasa | GGGGACCACTTTGTACAAGAAAGCTGGGTATTATTCCCATTCCTCATC | |
| SP6 | ATTTAGGTGACACTATAGAA | Sequencing |
| T7 (Pcs2+) | GTAATACGACTCACTATAG | |
| T3 | TATTAACCCTCACTAAAGGG | |

## 5.11 Bioinformatics methods

### 5.11.1 Multiple sequence alignments

As described in Riemer et al 2014, Multiple sequence alignments of Buc and Osk homologs have been accomplished in collaboration with Dr. Thomas Lingner (Department of Bioinformatics, Georg August University Göttingen) with the T-Coffee software of the EMBL-EBI (McWilliam et al., 2013).

Following Buc vertebrate homologs (gi number is given) were used: 292610748, 47225100, 148230857, 301615136, 118086206, 513169733, 73976581, 327275069, 642119256, 410909482, 432930267.

### 5.11.2  Hidden Markov models analysis

As described in Riemer et al 2015, Hidden Markov models analysis was performed in collaboration with Dr. Thomas Lingner by HMMER (http://hmmer.janelia.org/)(Finn, Clements and Eddy, 2011). Using this method, remote homologs can be detected more accurately in comparison to BLAST.

### 5.11.3  Pairwise alignments
This was performed using the EMBL EMBOSS water software (McWilliam *et al.*, 2013).

### 5.12  Statistics
Error bars indicate the standard deviation of the average (at least three independent experiments).  The statistical significance (P-value) of two groups of values was calculated using a two-tailed, two-sample unequal variance t-test.

# List of Tables

# List of Figures

# Bibliography

Aguero, T. *et al.* (2017) *Vertebrate Development*. doi: 10.1007/978-3-319-46095-6.

Ahuja, A. and Extavour, C. G. (2014) 'Patterns of molecular evolution of the germ line specification gene oskar suggest that a novel domain may contribute to functional divergence in Drosophila', *Development Genes and Evolution*, 224(2), pp. 65–77. doi: 10.1007/s00427-013-0463-7.

Anantharaman, V., Zhang, D. and Aravind, L. (2010) 'OST-HTH: a novel predicted RNA-binding domain', *Biology Direct*, 5(1), p. 13. doi: 10.1186/1745-6150-5-13.

Anderson, P. and Kedersha, N. (2006) 'RNA granules', *Journal of Cell Biology*, pp. 803–808. doi: 10.1083/jcb.200512082.

Bagga, S. *et al.* (2005) 'Regulation by let-7 and lin-4 miRNAs results in target mRNA degradation', *Cell*, 122(4), pp. 553–563. doi: 10.1016/j.cell.2005.07.031.

Bauer, M. P. and Goetz, F. W. (2001) 'Isolation of Gonadal Mutations in Adult Zebrafish from a Chemical Mutagenesis Screen 1', *Biology of reproduction*, 64(2), pp. 548–554. doi: 10.1095/biolreprod64.2.548.

Bertocchini, F. and Chuva de Sousa Lopes, S. M. (2016) 'Germline development in amniotes: A paradigm shift in primordial germ cell specification', *BioEssays*, 38(8), pp. 791–800. doi: 10.1002/bies.201600025.

Boke, E. *et al.* (2016) 'Amyloid-like Self-Assembly of a Cellular Compartment', *Cell*, 166(3), pp. 637–650. doi: 10.1016/j.cell.2016.06.051.

Bolognesi, B. *et al.* (2016) 'A concentration-dependent liquid phase separation can cause toxicity upon increased protein expression', *Cell Reports*, 16(1), pp. 222–232. doi: 10.1016/j.celrep.2016.05.076.

Bontems, F. *et al.* (2009) 'Bucky Ball Organizes Germ Plasm Assembly in Zebrafish', *Current Biology*. Elsevier Ltd, 19(5), pp. 414–422. doi: 10.1016/j.cub.2009.01.038.

Bontems, F. *et al.* (2011) 'Efficient mutation identification in zebrafish by microarray capturing and next generation sequencing', *Biochemical and Biophysical Research Communications*, 405(3), pp. 373–376. doi: 10.1016/j.bbrc.2011.01.024.

Bounoure (1934) 'Recherches sur la ligne´e germinale chez la grenouille rousse aux premiers stades du de´veloppement.', *Ann. Sci. Nat., 67-248*.

Boveri (1910) 'Die Potenzen der Ascaris-Blastomeren bei abgeanderter Furchung. Festschrift fur R. Hertwig'.

Braat, A. K. *et al.* (2000) 'Vasa protein expression and localization in the zebrafish', *Mechanisms of Development*, 95(1–2), pp. 271–274. doi: 10.1016/S0925-4773(00)00344-0.

Brangwynne, C. P. *et al.* (2009) 'Germline P Granules Are Liquid Droplets That Localize by Controlled Dissolution/Condensation', *Science*, 324(5935), pp. 1729–1732. doi: 10.1126/science.1172046.

Brangwynne, C. P. (2013) 'Phase transitions and size scaling of membrane-less organelles.', *The Journal of cell biology*, 203(6), pp. 875–881. doi: 10.1083/jcb.201308087.

Breitwieser, W. *et al.* (1996) 'Oskar protein interaction with vasa represents an essential step in polar granule assembly', *Genes and Development*, 10(17), pp. 2179–2188. doi: 10.1101/gad.10.17.2179.

Callebaut, I. and Mornon, J. P. (2010) 'LOTUS, a new domain associated with small RNA pathways in the germline', *Bioinformatics*, 26(9), pp. 1140–1144. doi: 10.1093/bioinformatics/btq122.

Castello, A. *et al.* (2016) 'Comprehensive Identification of RNA-Binding Domains in Human Cells', *Molecular Cell*. The Author(s), 63(4), pp. 696–710. doi: 10.1016/j.molcel.2016.06.029.

Caussinus, E., Kanca, O. and Affolter, M. (2011) 'Fluorescent fusion protein knockout mediated by anti-GFP nanobody', *Nature Structural & Molecular Biology*. Nature Publishing Group, 19(1), pp. 117–121. doi: 10.1038/nsmb.2180.

Cech, T. R. (2010) 'The RNA world in context', *RNA Worlds*, pp. 9–13. doi: 10.1101/cshperspect.a006742.

Chew, G.-L., Pauli, A. and Schier, A. F. (2016) 'Conservation of uORF repressiveness and sequence features in mouse, human and zebrafish.', *Nature communications*. Nature Publishing Group, 7(May), p. 11663. doi: 10.1038/ncomms11663.

Coletta, A. *et al.* (2010) 'Low-complexity regions within protein sequences have position-dependent roles.', *BMC systems biology*, 4(1), p. 43. doi: 10.1186/1752-0509-4-43.

Costa, Y. *et al.* (2006) 'Mouse MAELSTROM: The link between meiotic silencing of unsynapsed chromatin and microRNA pathway?', *Human Molecular Genetics*, 15(15), pp. 2324–2334. doi: 10.1093/hmg/ddl158.

Cox, R. T. and Spradling, A. C. (2003) 'A Balbiani body and the fusome mediate mitochondrial inheritance during Drosophila oogenesis.', *Development (Cambridge, England)*, 130(8), pp. 1579–1590. doi: 10.1242/dev.00365.

Crow, J. F. (1994) 'Advantages of sexual reproduction', *Developmental Genetics*, 15(3), pp. 205–213. doi: 10.1002/dvg.1020150303.

Curtis, D., Apfeld, J. and Lehmann, R. (1995) 'Nanos Is an Evolutionarily Conserved Organizer of Anterior-Posterior Polarity.', *Development (Cambridge, England)*, 121, pp. 1899–1910.

Diver, C., Boycott, A. E. and Garstang, S. (1925) 'The inheritance of inverse symmetry in Limnaea peregra', *Journal of Genetics*, 15(2), pp. 113–200. doi: 10.1007/BF02983106.

Doolittle, R. F. (1981) 'Similar amino acid sequences: chance or common ancestry?', *Science (New York, N.Y.)*, 214(4517), pp. 149–159. doi: 10.1126/science.7280687.

Van Doren, M. *et al.* (1998) 'HMG-CoA reductase guides migrating primordial germ cells.', *Nature*, 396(6710), pp. 466–469. doi: 10.1038/24871.

Dosch, R. *et al.* (2004) 'Maternal control of vertebrate development before the midblastula transition: Mutants from the zebrafish I', *Developmental Cell*, 6(6), pp. 771–780. doi:

10.1016/j.devcel.2004.05.002.

Dosch, R. (2015) 'Next generation mothers: Maternal control of germline development in zebrafish', *Critical Reviews in Biochemistry and Molecular Biology*. Informa Healthcare USA, Inc, 50(1), pp. 54–68. doi: 10.3109/10409238.2014.985816.

Dyson, H. J. and Wright, P. E. (2005) 'Intrinsically unstructured proteins and their functions.', *Nature reviews. Molecular cell biology*, 6(3), pp. 197–208. doi: 10.1038/nrm1589.

Ephrussi, A., Dickinson, L. K. and Lehmann, R. (1991) 'Oskar Organizes the Germ Plasm and Directs Localization of the Posterior Determinant Nanos', *Cell*, 66(1), pp. 37–50. doi: 10.1016/0092-8674(91)90137-N.

Ephrussi, A. and Lehmann, R. (1992) 'Induction of germ cell formation by oskar', *Nature*, 358(6385), pp. 387–392. doi: 10.1038/358387a0.

Escobar-Aguirre, M. *et al.* (2017) 'Microtubule-actin crosslinking factor 1 (Macf1) domain function in Balbiani body dissociation and nuclear positioning', *PLOS Genetics*. Public Library of Science, 13(9), p. e1006983. Available at: https://doi.org/10.1371/journal.pgen.1006983.

Ewen-Campen, B., Schwager, E. E. and Extavour, C. G. M. (2010) 'The molecular machinery of germ line specification', *Molecular Reproduction and Development*, 77(1), pp. 3–18. doi: 10.1002/mrd.21091.

Extavour, C. G. and Akam, M. (2003) 'Mechanisms of germ cell specification across the metazoans: epigenesis and preformation.', *Development (Cambridge, England)*, 130(24), pp. 5869–5884. doi: 10.1242/dev.00804.

Finn, R. D., Clements, J. and Eddy, S. R. (2011) 'HMMER web server: Interactive sequence similarity searching', *Nucleic Acids Research*, 39(SUPPL. 2), pp. 29–37. doi: 10.1093/nar/gkr367.

Foresta, C., Ferlin, A. and Moro, E. (2000) 'Deletion and expression analysis of AZFa genes on the human Y chromosome revealed a major role for DBY in male infertility.', *Human molecular genetics*. England, 9(8), pp. 1161–1169.

Forman-Kay, J. D. and Mittag, T. (2013) 'From sequence and forces to structure, function, and evolution of intrinsically disordered proteins', *Structure*. Elsevier Ltd, 21(9), pp. 1492–1499. doi: 10.1016/j.str.2013.08.001.

Frey, S., Richter, R. P. and Gorlich, D. (2006) 'FG-Rich Repeats of Nuclear Pore Proteins Form a Three-Dimensional Meshwork with Hydrogel-Like Properties', *Science*, 314(5800), pp. 815–817. doi: 10.1126/science.1132516.

Gavis, E. R. and Lehmann, R. (1994) 'Translational regulation of nanos by RNA localization', *Nature*, 369, pp. 315–318. doi: 10.1038/369315a0.

Gehring, W. J. (2005) 'New perspectives on eye development and the evolution of eyes and photoreceptors', *Journal of Heredity*, 96(3), pp. 171–184. doi: 10.1093/jhered/esi027.

González-Morales, N. *et al.* (2015) 'The Atypical Cadherin Dachsous Controls Left-Right Asymmetry in Drosophila', *Developmental Cell*, 33(6), pp. 675–689. doi: 10.1016/j.devcel.2015.04.026.

Grosberg, R. K. and Strathmann, R. R. (2007) 'The Evolution of Multicellularity: A Minor Major Transition?', *Annu. Rev. Ecol. Evol. Syst*, 38(August), pp. 621–54. doi: 10.1146/annurev.ecolsys.36.102403.114735.

Gupta, T. *et al.* (2010) 'Microtubule actin crosslinking factor 1 regulates the balbiani body and animal-vegetal polarity of the zebrafish oocyte', *PLoS Genetics*, 6(8), pp. 14–18. doi: 10.1371/journal.pgen.1001073.

Guraya, S. S. (1979) 'Recent advances in the morphology, cytochemistry, and function of Balbiani's vitelline body in animal oocytes.', *International review of cytology*. United States, 59, pp. 249–321.

Harris,  a N. and Macdonald, P. M. (2001) 'Aubergine encodes a Drosophila polar granule component required for pole cell formation and related to eIF2C.', *Development (Cambridge, England)*, 128(14), pp. 2823–2832.

Hartung, O., Forbes, M. M. and Marlow, F. L. (2014) 'Zebrafish vasa is required for germ-cell differentiation and maintenance', *Molecular Reproduction and Development*, 81(10), pp. 946–961. doi: 10.1002/mrd.22414.

Hashimoto, Y. *et al.* (2004) 'Localized maternal factors are required for zebrafish germ cell formation', *Developmental Biology*, 268(1), pp. 152–161. doi: 10.1016/j.ydbio.2003.12.013.

Heasman, J., Quarmby, J. and Wylie, C. C. (1984) 'The mitochondrial cloud of Xenopus oocytes: The source of germinal granule material', *Developmental Biology*, 105(2), pp. 458–469. doi: 10.1016/0012-1606(84)90303-8.

Hegner, R. W. (1911) The Effects of Reinoving the Gernm Cell Determinants from the Eggs of Some Chnysoinelid Beetles III

Heim, A. E. *et al.* (2014) 'Oocyte polarity requires a Bucky ball-dependent feedback amplification loop', *Development*, 141(4), pp. 842–854. doi: 10.1242/dev.090449.

Hekimi, S., Boutis, P. and Lakowski, B. (1995) 'Viable maternal-effect mutations that affect the development of the nematode Caenorhabditis elegans', *Genetics*, 141(4), pp. 1351–1364.

Houwing, S. *et al.* (2007) 'A Role for Piwi and piRNAs in Germ Cell Maintenance and Transposon Silencing in Zebrafish', *Cell*, 129(1), pp. 69–82. doi: 10.1016/j.cell.2007.03.026.

Hurd, T. R. *et al.* (2016) 'Long Oskar Controls Mitochondrial Inheritance in Drosophila melanogaster', *Developmental Cell*. Elsevier Inc., 39(5), pp. 560–571. doi: 10.1016/j.devcel.2016.11.004.

Ikenishi, K. and Tanaka, T. S. (2000) 'Spatio-temporal expression of Xenopus vasa homolog, XVLG1, in oocytes and embryos: The presence of XVLG1 RNA in somatic cells as well as germline cells', *Development Growth and Differentiation*, 42(2), pp. 95–103. doi: 10.1046/j.1440-169X.2000.00493.x.

Illmensee, K. and Mahowald, A. P. (1974) 'Transplantation of posterior polar plasm in Drosophila. Induction of germ cells at the anterior pole of the egg', *Proceedings of the National Academy of Sciences of the United States of America*, 71(4), pp. 1016–1020. doi: 10.1073/pnas.71.4.1016.

Irion, U., Krauss, J. and Nusslein-Volhard, C. (2014) 'Precise and efficient genome editing in zebrafish using the CRISPR/Cas9 system', *Development*, 141(24), pp. 4827–4830. doi: 10.1242/dev.115584.

Iwai, Y. *et al.* (1991) 'Identification of sequences within the murine granulocyte-macrophage colony-stimulating factor mRNA 3′-untranslated region that mediate mRNA stabilization induced by mitogen treatment of EL-4 thymoma cells', *Journal of Biological Chemistry*, 266(27), pp. 17959–17965.

Janetopoulos, C. *et al.* (1999) 'The conjusome : a novel structure in Tetrahymena found only during sexual reorganization', 1011, pp. 1003–1011.

Järvelin, A. I. *et al.* (2016) 'The new (dis)order in RNA regulation', *Cell Communication and Signaling*, 14(1), p. 9. doi: 10.1186/s12964-016-0132-3.

Jeske, M. *et al.* (2015) 'The crystal structure of the Drosophila germline inducer Oskar identifies two domains with distinct Vasa Helicase- and RNA-binding activities', *Cell Reports*, 12(4), pp. 587–598. doi: 10.1016/j.celrep.2015.06.055.

Jeske, M., Müller, C. W. and Ephrussi, A. (2017) 'The LOTUS domain is a conserved DEAD-box RNA helicase regulator and essential for recruitment of Vasa to germ plasm (Manuscript submitted to Molecular Cell)', pp. 939–952. doi: 10.1101/gad.297051.117.

Johnstone, O. (2004) 'Interaction with eIF5B is essential for Vasa function during development', *Development*, 131(17), pp. 4167–4178. doi: 10.1242/dev.01286.

Juliano, C. E., Swartz, S. Z. and Wessel, G. M. (2010) 'A conserved germline multipotency program', *Development*, 137(24), pp. 4113–4126. doi: 10.1242/dev.047969.

Kawakami, K. (2007) 'Tol2: a versatile gene transfer vector in vertebrates', *Genome Biology*, 8(Suppl 1), p. S7. doi: 10.1186/gb-2007-8-s1-s7.

Kedde, M. *et al.* (2007) 'RNA-Binding Protein Dnd1 Inhibits MicroRNA Access to Target mRNA', *Cell*, 131(7), pp. 1273–1286. doi: 10.1016/j.cell.2007.11.034.

Kim-Ha, J., Smith, J. L. and Macdonald, P. M. (1991) 'oskar mRNA is localized to the posterior pole of the Drosophila oocyte', *Cell*, 66(1), pp. 23–35. doi: 10.1016/0092-8674(91)90136-M.

Kishimoto, Y. *et al.* (2004) 'Zebrafish maternal-effect mutations causing cytokinesis defect without affecting mitosis or equatorial vasa deposition', *Mechanisms of Development*, 121(1), pp. 79–89. doi: 10.1016/j.mod.2003.10.001.

Kloc, M. *et al.* (2014) 'Balbiani body, nuage and sponge bodies - The germ plasm pathway players', *Arthropod Structure and Development*. Elsevier Ltd, 43(4), pp. 341–348. doi: 10.1016/j.asd.2013.12.003.

Knaut, H. *et al.* (2000) 'Zebrafish vasa RNA but not its protein is a component of the germ plasm and segregates asymmetrically before germline specification', *Journal of Cell Biology*, 149(4), pp. 875–888. doi: 10.1083/jcb.149.4.875.

Kobayashi, S. *et al.* (1996) 'Essential role of the posterior morphogen nanos for germline development in Drosophila', *Nature*, 380(6576), pp. 708–711. doi: 10.1038/380708a0.

Koprunner, M. *et al.* (2001) 'A zebrafish nanos -related gene is essential for the development of primordial germ cells', *Genes & development*, 15, pp. 2877–2885. doi: 10.1101/gad.212401.showed.

Kosaka, K. *et al.* (2007) 'Spatiotemporal localization of germ plasm RNAs during zebrafish oogenesis', *Mechanisms of Development*, 124(4), pp. 279–289. doi: 10.1016/j.mod.2007.01.003.

Kroschwald, S. *et al.* (2015) 'Promiscuous interactions and protein disaggregases determine the material state of stress-inducible RNP granules', *eLife*, 4(AUGUST2015), pp. 1–32. doi: 10.7554/eLife.06807.

Lasko, P. F. and Ashburner, M. (1988) 'The product of the Drosophila gene vasa is very similar to eukaryotic initiation factor-4A.', *Nature*, pp. 611–617. doi: 10.1038/335611a0.

Lavial, F. *et al.* (2009) 'Ectopic expression of Cvh (Chicken Vasa homologue) mediates the reprogramming of chicken embryonic stem cells to a germ cell fate', *Developmental Biology*, 330(1), pp. 73–82. doi: 10.1016/j.ydbio.2009.03.012.

Lawson, K. A. *et al.* (1999) 'Bmp4 is required for the generation of primordial germ cells in the mouse embryo.', *Genes & development*, 13(4), pp. 424–36. doi: 10.1101/gad.13.4.424.

Lehmann, R. (2016) *Germ Plasm Biogenesis-An Oskar-Centric Perspective*. 1st edn, *Current Topics in Developmental Biology*. 1st edn. Elsevier Inc. doi: 10.1016/bs.ctdb.2015.11.024.

Lehmann, R. and Nuesslein-Volhard, C. (1986) 'Abdominal segmentation, pole cell formation, and embryonic polarity require the localized activity of oskar, a maternal gene in drosophila', *Cell*, 47(1), pp. 141–152. doi: 10.1016/0092-8674(86)90375-2.

Lim, L. P. *et al.* (2005) 'Microarray analysis shows that some microRNAs downregulate large numbers of-target mRNAs', *Nature*, 433(7027), pp. 769–773. doi: 10.1038/nature03315.

Lin, Y. *et al.* (2015) 'Formation and Maturation of Phase-Separated Liquid Droplets by RNA-Binding Proteins', *Molecular Cell*. Elsevier Inc., 60(2), pp. 208–219. doi: 10.1016/j.molcel.2015.08.018.

Link, V., Shevchenko, A. and Heisenberg, C.-P. (2006) 'Proteomics of early zebrafish embryos.', *BMC developmental biology*, 6, p. 1. doi: 10.1186/1471-213X-6-1.

Liu, J. *et al.* (2008) 'The C. elegans SYS-1 Protein Is a Bona Fide β-Catenin', *Developmental Cell*, 14(5), pp. 751–761. doi: 10.1016/j.devcel.2008.02.015.

Machado, R. J. *et al.* (2005) 'Xenopus Xpat protein is a major component of germ plasm and may function in its organisation and positioning', *Developmental Biology*, 287(2), pp. 289–300. doi: 10.1016/j.ydbio.2005.08.044.

Maegawa, S. *et al.* (2002) 'Zebrafish DAZ-like protein controls translation via the sequence "GUUC"', *Genes to Cells*, 7(9), pp. 971–984. doi: 10.1046/j.1365-2443.2002.00576.x.

Marinos, E. and Billett, F. S. (1981) 'Mitochondrial number, cytochrome oxidase and succinic dehydrogenase activity in Xenopus laevis oocytes', *J Embryol Exp Morphol*, 62(609), pp. 395–409. Available at: http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=6268730.

Markussen, F.-H. *et al.* (1995) 'Translational control of oskar generates Short OSK, the isoform that induces pole plasm assembly', *Development* , 3732(121), pp. 3723–3732.

Marlow, F. L. and Mullins, M. C. (2008) 'Bucky ball functions in Balbiani body assembly and animal-vegetal polarity in the oocyte and follicle cell layer in zebrafish', *Developmental Biology*, 321(1), pp. 40–50. doi: 10.1016/j.ydbio.2008.05.557.

McWilliam, H. *et al.* (2013) 'Analysis Tool Web Services from the EMBL-EBI.', *Nucleic acids research*, 41(Web Server issue), pp. 597–600. doi: 10.1093/nar/gkt376.

Mochizuki, K. *et al.* (2002) 'Analysis of a piwi -Related Gene Implicates Small RNAs in Genome Rearrangement in Tetrahymena', 110, pp. 689–699.

Mosquera, L. *et al.* (1993) 'A mRNA localized to the vegetal cortex of Xenopus oocytes encodes a protein with a nanos-like zinc finger domain.', *Development (Cambridge, England)*, 117(1), pp. 377–386. doi: 10.1016/0168-9525(93)90157-D.

Mullis, K. *et al.* (1986) 'Specific enzymatic amplification of DNA in vitro: The polymerase chain reaction', *Cold Spring Harbor Symposia on Quantitative Biology*, 51(1), pp. 263–273. doi: 10.1101/SQB.1986.051.01.032.

Nair, S., Lindeman, R. E. and Pelegri, F. (2013) 'In vitro oocyte culture-based manipulation of zebrafish maternal genes', *Developmental Dynamics*, 242(1), pp. 44–52. doi: 10.1002/dvdy.23894.

Nakamura, A. and Seydoux, G. (2008) 'Less is more: specification of the germline by transcriptional repression', *Development*, 135(23), pp. 3817–3827. doi: 10.1242/dev.022434.

Nijjar, S. and Woodland, H. R. (2013) 'Protein interactions in Xenopus germ plasm RNP particles', *PLoS ONE*, 8(11). doi: 10.1371/journal.pone.0080077.

Noce, T., Okamoto-Ito, S. and Tsunekawa, N. (2001) 'Vasa Homolog Genes in Mammalian Germ Cell Development', *Cell Structure and Function*, 26(3), pp. 131–136. doi: 10.1247/csf.26.131.

Nott, T. J. *et al.* (2015) 'Phase Transition of a Disordered Nuage Protein Generates Environmentally Responsive Membraneless Organelles', *Molecular Cell*, 57(5). doi: 10.1016/j.molcel.2015.01.013.

Oldfield, C. J. and Dunker,  a K. (2014) 'Intrinsically disordered proteins and intrinsically disordered protein regions.', *Annual review of biochemistry*, 83, pp. 553–84. doi: 10.1146/annurev-biochem-072711-164947.

Pek, J. W., Patil, V. S. and Kai, T. (2012) 'piRNA pathway and the potential processing site, the nuage, in the Drosophila germline.', *Development, growth & differentiation*. Japan, 54(1), pp. 66–77.

Pelegri, F. *et al.* (1999) 'A mutation in the zebrafish maternal-effect gene nebel affects furrow formation and vasa RNA localization', *Current Biology*, 9(24), pp. 1431–1440. doi: 10.1016/S0960-9822(00)80112-8.

Pelegri, F. and Mullins, M. C. (2016) 'Genetic screens for mutations affecting adult traits and parental-effect genes', *Methods in Cell Biology*, 135, pp. 39–87. doi: 10.1016/bs.mcb.2016.05.006.

Pepling, M. E. and Spradling,  a C. (1998) 'Female mouse germ cells form synchronously dividing cysts.', *Development*, 125(17), pp. 3323–3328.

Pillai, R. S. (2005) 'Inhibition of Translational Initiation by Let-7 MicroRNA in Human Cells', *Science*,

309(5740), pp. 1573–1576. doi: 10.1126/science.1115079.

Raz, E. (2000) 'The function and regulation of vasa -like genes in germ-cell development Erez Raz', *Genome Biology*, 1(3), pp. 1–6. doi: 10.1186/gb-2000-1-3-reviews1017.

Riemer, S. (2014) 'Analyzing the molecular mechanism of Bucky ball localization during germ cell specification in zebrafish', *Dissertation*, (October 2014).

Riemer, S. *et al.* (2015) 'A functional Bucky ball-GFP transgene visualizes germ plasm in living zebrafish', *Gene Expression Patterns*. Elsevier B.V., 18(1–2), pp. 44–52. doi: 10.1016/j.gep.2015.05.003.

Rongo, C. *et al.* (1997) 'Germ plasm assembly and germ cell migration in Drosophila.', *Cold Spring Harbor symposia on quantitative biology*. United States, 62, pp. 1–11.

Rongo, C., Gavis, E. R. and Lehmann, R. (1995) 'Localization of oskar RNA regulates oskar translation and requires Oskar protein.', *Development (Cambridge, England)*, 121(9), pp. 2737–46. Available at: http://www.ncbi.nlm.nih.gov/pubmed/7555702.

Sanger, F., Nicklen, S. and Coulson, A. R. (1977) 'DNA sequencing with chain-terminating inhibitors', *Proceedings of the National Academy of Sciences*, 74(12), pp. 5463–5467. doi: 10.1073/pnas.74.12.5463.

Schupbach, T. and Wieschaus, E. (1986a) 'Germline autonomy of maternal-effect mutations altering the embryonic body pattern of Drosophila', *Developmental Biology*, 113(2), pp. 443–448. doi: 10.1016/0012-1606(86)90179-X.

Schupbach, T. and Wieschaus, E. (1986b) 'Maternal-effect mutations altering the anterior-posterior pattern of the Drosophila embryo.', *Roux's archives of developmental biology : the official organ of the EDBO*. Germany, 195(5), pp. 302–317. doi: 10.1007/BF00376063.

Seervai, R. N. H. and Wessel, G. M. (2013) 'Lessons for inductive germline determination', *Molecular Reproduction and Development*, 80(8). doi: 10.1002/mrd.22151.

Seydoux, G. and Braun, R. E. (2006) 'Pathway to Totipotency: Lessons from Germ Cells', *Cell*, 127(5), pp. 891–904. doi: 10.1016/j.cell.2006.11.016.

Sharp, P. A., Sugden, B. and Sambrook, J. (1973) 'Detection of Two Restriction Endonuclease Activities in Haemophilus parainfluenzae Using Analytical Agarose-Ethidium Bromide Electrophoresis', *Biochemistry*, 12(16), pp. 3055–3063. doi: 10.1021/bi00740a018.

Sinsimer, K. S. *et al.* (2013) 'Germ plasm anchoring is a dynamic state that requires persistent trafficking', *Cell Reports*. The Authors, 5(5), pp. 1169–1177. doi: 10.1016/j.celrep.2013.10.045.

Škugor, A. *et al.* (2016) 'Multiplicity of Buc copies in Atlantic salmon contrasts with loss of the germ cell determinant in primates, rodents and axolotl', *BMC Evolutionary Biology*, 16(1), p. 232. doi: 10.1186/s12862-016-0809-7.

Smith, J. L., Wilson, J. E. and Macdonald, P. M. (1992) 'Overexpression of oskar directs ectopic activation of nanos and presumptive pole cell formation in Drosophila embryos', *Cell*, 70(5), pp. 849–859. doi: 10.1016/0092-8674(92)90318-7.

Smith, L. D. (1966) 'The role of a "germinal plasm" in the formation of primordial germ cells in Rana pipiens', *Dev Biol*, 14(2), pp. 330–347. doi: 10.1016/0012-1606(66)90019-4.

Srouji, J. R. and Extavour, C. Chapter 16, pp. 360-397 in Key Transitions in Animal Evolution, R. DeSalle and B. Schierwater, eds. CRC Press, Taylor & Francis

Strome, S. and Updike, D. (2015) 'Specifying and protecting germ cell fate', *Nature Reviews Molecular Cell Biology*. Nature Publishing Group, 16(7), pp. 406–416. doi: 10.1038/nrm4009.

Sturtevant, A. H. (1923) 'Inheritance of coiling direction in Limnaea', *Science*, 58, pp. 269–270.

Suyama, R. *et al.* (2009) 'The actin-binding protein Lasp promotes Oskar accumulation at the posterior pole of the Drosophila embryo.', *Development (Cambridge, England)*, 136(1), pp. 95–105. doi: 10.1242/dev.027698.

Tada, H. *et al.* (2012) 'Ectopic formation of primordial germ cells by transplantation of the germ plasm: Direct evidence for germ cell determinant in Xenopus', *Developmental Biology*. Elsevier, 371(1), pp. 86–93. doi: 10.1016/j.ydbio.2012.08.014.

Tam, P. P. and Zhou, S. X. (1996) 'The allocation of epiblast cells to ectodermal and germ-line lineages is influenced by the position of the cells in the gastrulating mouse embryo.', *Developmental biology*, 178(1), pp. 124–132. doi: 10.1006/dbio.1996.0203.

Tomancak, P. *et al.* (1998) 'Oocyte polarity depends on regulation of gurken by Vasa.', *Development (Cambridge, England)*, 125, pp. 1723–1732.

Toretsky, J. A. and Wright, P. E. (2014) 'Assemblages: Functional units formed by cellular phase separation', *Journal of Cell Biology*, 206(5), pp. 579–588. doi: 10.1083/jcb.201404124.

Toyooka, Y. *et al.* (2000) 'Expression and intracellular localization of mouse Vasa-homologue protein during germ cell development', *Mechanisms of Development*, 93(1–2), pp. 139–149. doi: 10.1016/S0925-4773(00)00283-5.

Tsuda, M. (2003) 'Conserved Role of nanos Proteins in Germ Cell Development', *Science*, 301(5637), pp. 1239–1241. doi: 10.1126/science.1085222.

Tsunekawa, N. *et al.* (2000) 'Isolation of chicken vasa homolog gene and tracing the origin of primordial germ cells.', *Development (Cambridge, England)*, 127(12), pp. 2741–2750.

Tworzydlo, W. *et al.* (2016) 'Exclusion of dysfunctional mitochondria from Balbiani body during early oogenesis of Thermobia', *Cell and Tissue Research*. Cell and Tissue Research, 366(1), pp. 191–201. doi: 10.1007/s00441-016-2414-x.

Ullmann, S. L. *et al.* (1997) 'Migration of primordial germ cells to the developing gonadal ridges in the tammar wallaby Macropus eugenii', *Journal of reproduction and fertility*, 110, pp. 135–143.

Vanzo, N. F. and Ephrussi, A. (2002) 'Oskar anchoring restricts pole plasm formation to the posterior of the Drosophila oocyte.', *Development (Cambridge, England)*, 129, pp. 3705–3714.

Wagner, D. S. *et al.* (2004) 'Maternal Control of Development at the Midblastula Transition and beyondMutants from the Zebrafish II', *Developmental Cell*, 6(6), pp. 781–790. doi:

10.1016/j.devcel.2004.04.001.

Wang, J. T. *et al.* (2014) 'Regulation of RNA granule dynamics by phosphorylation of serine-rich, intrinsically disordered proteins in C. elegans', *eLife*, 3, pp. 1–23. doi: 10.7554/eLife.04591.

Weismann (1893) *The Germ-plasm: A Theory of Heredity.*

Westerfield, M. (2000) *The Zebrafish Book*, *Eugene*.

Wilding, M. *et al.* (2001) 'Mitochondrial aggregation patterns and activity in human oocytes and preimplantation embryos.', *Human reproduction (Oxford, England)*. England, 16(5), pp. 909–917.

Wylie, C. (1999) 'Germ cells.', *Cell*, 96(2), pp. 165–174. doi: 10.1016/s0092-8674(00)80557-7.

Xue, B. *et al.* (2010) 'PONDR-FIT: A meta-predictor of intrinsically disordered amino acids', *Biochimica et Biophysica Acta - Proteins and Proteomics*, 1804(4), pp. 996–1010. doi: 10.1016/j.bbapap.2010.01.011.

Yajima, M. and Wessel, G. M. (2015) 'Essential elements for translation: the germline factor Vasa functions broadly in somatic cells', *Development*, 142, pp. 1960–1970. doi: 10.1242/dev.118448.

Yang, N. *et al.* (2015) 'Structure of *Drosophila* Oskar reveals a novel RNA binding protein', *Proceedings of the National Academy of Sciences*, 112(37), pp. 11541–11546. doi: 10.1073/pnas.1515568112.

Yoon, C., Kawakami, K. and Hopkins, N. (1997) 'Zebrafish vasa homologue RNA is localized to the cleavage planes of 2- and 4-cell-stage embryos and is expressed in the primordial germ cells', 3166, pp. 3157–3165.

Zhang, Y., Cao, H. and Liu, Z. (2015) 'Binding cavities and druggability of intrinsically disordered proteins', *Protein Science*, 24(5), pp. 688–705. doi: 10.1002/pro.2641.

# Appendix

## Table 1: Bucky ball and Oskar do not share sequence homology.
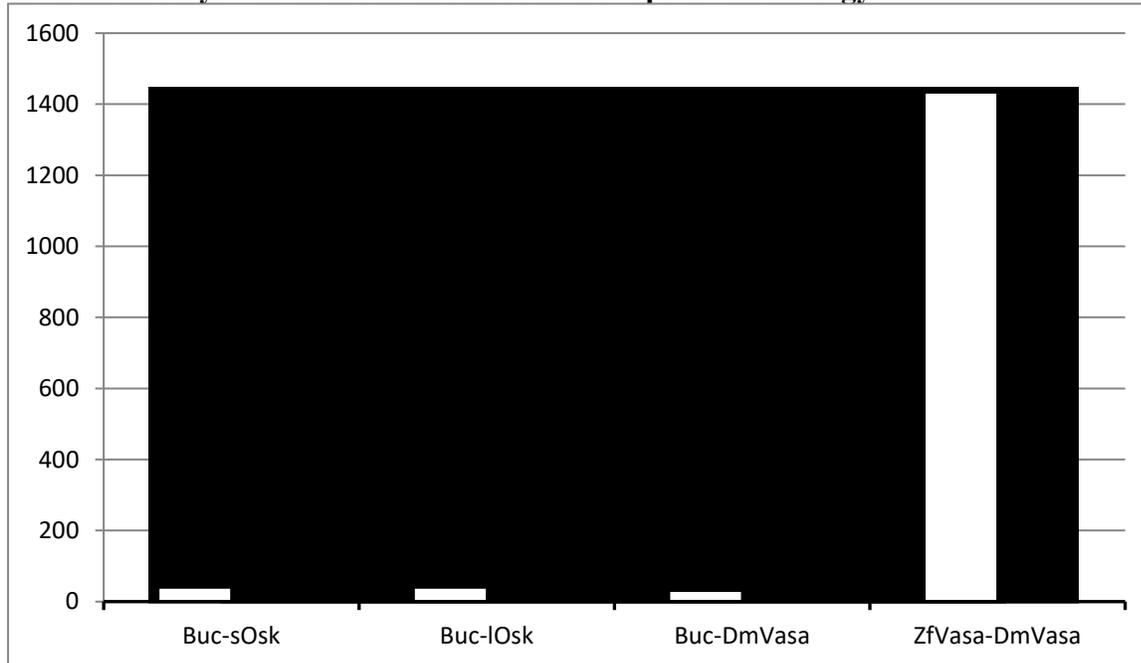


Graph summarizing scores of global (white bar) and local alignments (black bar). Buc-Osk alignments are equally low as the negative control (Buc-Dm Vasa), whereas ZfVasa and DmVasa show a characteristic score of two homologous sequences. From (S.Riemer 2014)
White bar  is the  Needleman-Wunsch algorithm.
Black bar is the Smith-Waterman.

A) Global alignment with Needleman-Wunsch

| Alignment | Bucky ball short Oskar (Isoform C) | Bucky ball long Oskar (Isoform A) | Zf Bucky ball Dm Vasa | Zf Vasa Dm Vasa |
|---|---|---|---|---|
| Seq. similarity [%] | 11.5% | 10.0% | 18.5% | 59.4% |
| Seq. identity [%] | 6.3% | 5.5% | 10.8% | 45.2% |
| score | 39.0 | 39.0 | 30.5 | 1433 |

B) Local alignment with Smith-Waterman

| Alignment | Bucky ball short Oskar (Isoform C) | Bucky ball long Oskar (Isoform A) | Zf Bucky ball Dm Vasa | Zf Vasa Dm Vasa |
|---|---|---|---|---|
| Seq. similarity [%] | 30.6 | 35.9 | 29.9 | 59.8 |
| Seq. identity [%] | 17.1 | 17.7 | 18.4 | 45.7 |
| score | 43.0 | 59.5 | 46.5 | 1434 |

Analysis of protein sequences with global pairwise alignments using the Needleman-Wunsch algorithm (A; http://www.ebi.ac.uk/Tools/psa/emboss_needle/; standard settings) or with local pairwise alignments using the Smith-Waterman algorithm (B; http://www.ebi.ac.uk/Tools/psa/emboss_water/; standard settings). Depicted are the percentages of similar and identical amino acids of two aligned protein. From (S. Riemer, 2014)

**Appendix Table 2: Comparison of Buc and Osk with Hidden-Markov-Models.**

| Used Hidden-Markov-Model | Searched NCBI organism database | Significant hits | E-Value |
|---|---|---|---|
| Buc-HMM | *Danio rerio* | Bucky ball | 9.4e-223 |
| | | zDazl | 0.022 |
| | *Drosophila melanogaster* | - | - |
| Osk-HMM | *Danio rerio* | Trd7A | 1.4e-11 |
| | | Trd7A, isoform 1 | 3.2e-11 |
| | | Tdrd5 | 2.4e-07 |
| | | Trd7B | 1e-06 |
| | | Her13 | 0.017 |
| | *Drosophila melanogaster* | Oskar; isoform A | 2.8e-223 |
| | | Oskar; isoform C | 1.1e-205 |
| | | RE24380p | 8.8e-198 |
| | | CG8920, isoform D /E | 1.9e-07 |
| | | CG8920, isoform B / C | 2.8e-07 |
| | | CG34007 | 0.00015 |
| | | Tejas / FI02030p | 0.0048 |
| | | IP20666p | 0.016 |

Homology search with conserved domains using Hidden-Markov-Models (www.HMMer.org) of the respective proteins did not reveal any conserved domains between Oskar and Bucky ball. Hits of the used HMM in the NCBI databases are shown with their corresponding E-value. From (S. Riemer, 2014)

**Buc p43 and Buc p106 are expressed in the Zebrafish**

The mRNAs of *bucp⁴³* and *bucp¹⁰⁶* were injected into the one cell embryo to check for their expression. At 3hpf stage the embryos were lysed and check for the mutant protein expression via western blot using the Buc antibody(Riemer *et al.*, 2015). This shows both mutant proteins are expressed in the zebrafish.
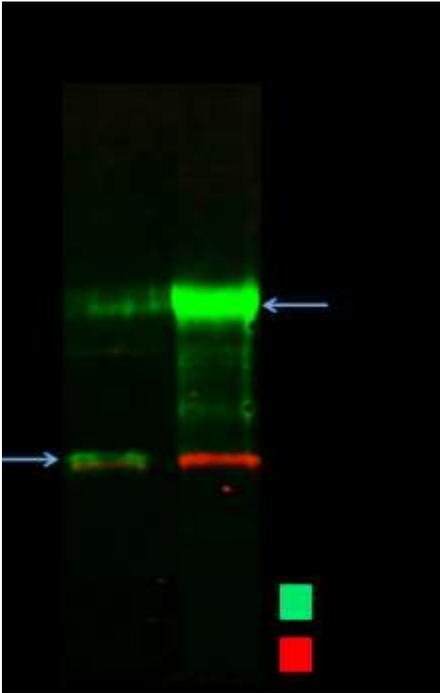


Appendix Fig.1: Buc p43 and Buc p106 are expressed in the Zebrafish embryo. Blue arrows indicate the respective protein band. Tubulin as the loading control.

**Buc- Vasa interaction is RNA independent.**

The blot shown below, shows the result of scenario (ii) where the proteins Buc and Vasa were co-translated together and treated with RNAse (A and T1). This shows that Buc-Vasa interaction is not RNA dependant.
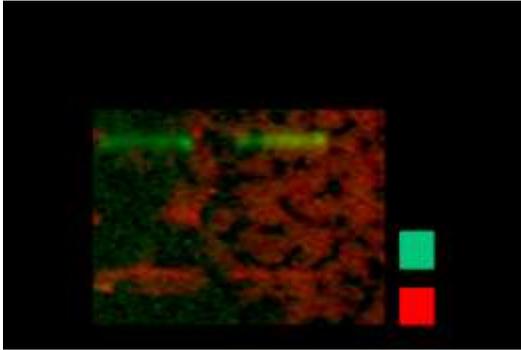


Appendix Fig.2: Buc- Vasa interaction is not dependent on RNA.  Blot showing two lanes of Buc- Vasa pulldowns.  In the first lane the sample was not treated with RNAses (A and T1) before pulldown. In the second lane the samples were treated with RNAses (A and T1) before pulldown.

# Curriculum Vitae

**Name:** Pritesh Krishnakumar
**Date of Birth:** 08.03.1985
**Place of Birth:** Chennai, India
**Address:** Am Vogelsang, Apt 312. 37075, Göttingen, Germany
**Email:** Pritesh.Krishnakumar@med.uni-goettingen.de

## Education and Research:

| Position/Institution | Guide | Project | Year |
|---|---|---|---|
| PhD Student University of Goettingen, Germany | Dr. R.Dosch | Functional Conservation of germ plasm organizers Bucky ball in zebrafish and *Drosophila* Oskar. | 2013-2017 |
| Senior Research Fellow Advance Centre for Treatment Research and Education in Cancer (ACTREC), Navi Mumbai,India | Prof. I. Mittra | Circulating DNA as a causative agent for cancer. | 2012-2013 |
| Research Assistant Virginia Polytechnic and State University Virginia, USA | Dr. D. Banerjee | Role of Lin-42 in stem cell fate maintenance in *C.elegans.* | 2009-2011 |
| Junior Research Fellow Tata Institute of Fundamental Research, Mumbai, India | Dr. M. Narasimha | Role of cell death in dorsal closure in *Drosophila melanogaster.* | 2008-2009 |
| Masters in Biochemistry Maharaja Sayajirao University ,Vadodara, India | Dr. P. Robin | Plant defense mechanism in *Brassica juncea seed extracts.* | 2006-2008 |
| Bachelors in Biochemistry and Life Sciences St. Xavier´s College Mumbai, India | . | | 2003-2006 |

**Awards and Stipends:**

1. GGNB bridging fund (2017-present)

2. DAAD fellowship for PhD program (2013-2017)

3. Fellowship from Council of Scientific & Industrial Research (India) (2008-2009).