Cell fate specification in the *Drosophila* embryo: Maternal Roles for Dorsal, Caspar and Groucho

विद्या वाचस्पति की
उपाधि की अपेक्षाओं की आंशिक पूर्ति में प्रस्तुत शोध प्रबंध

A thesis Submitted in partial fulfillment

of the requirements of the degree Doctor of Philosophy

द्वारा / By

सुभ्रदीप दास/Subhradip Das

20193641

डॉ. गिरीश रत्नपारखी/Dr. Girish Ratnaparkhi

भारतीय विज्ञान शिक्षा एवं अनुसंधान संस्थान पुणे
Indian Institute of Science Education and Research, Pune

2025

Certificate

I certify that the original work presented in this thesis titled **'Cell fate specification in the *Drosophila* embryo: Maternal Roles for Dorsal, Caspar and Groucho'** was carried out by Mr. Subhradip Das at IISER, Pune, under my guidance, from August 2019 to July 2022. The work presented here or a part of it has not been included in any other thesis submitted previously for the award of any degree or diploma from any other University or institution.

Girish Ratnaparkhi

Thesis supervisor

Declaration

I declare that this written submission represents my ideas in my own words, and where others' ideas have been included, I have adequately cited and referenced the original sources. I also declare that I have adhered to all academic honesty and integrity principles and have not misrepresented, fabricated, or falsified any idea/data/fact/source in my submission. I understand that violation of the above will cause disciplinary action by the Institute and can also evoke penal action from the sources, which have thus been appropriately cited, or from whom proper permission has not been taken when needed.

Subhradip Das.

Subhradip Das

20193641

In my capacity as supervisor of the candidate's thesis, I certify that the above statements are true to the best of my knowledge.

G S Ratnaparkhi

Dr. Girish Ratnaparkhi

To my Mom and Dad, for your
unwavering love, support and belief
in me. Thank you for always being
there for me – I owe it all to you

Acknowledgments

My PhD journey was a symphony of beauty, chaos, and discovery—a delicate dance between the known and the unknown. It was as if I were chasing fireflies in the dark; each moment of illumination brought wonder, yet the shadows that followed whispered mysteries yet to be uncovered. The path was anything but linear, winding through triumphs that took my breath away and setbacks that demanded my patience and resolve. The unpredictability was its greatest gift, teaching me to find solace in uncertainty and marvel at the unexpected. It wasn't just about the research but the quiet nights spent pondering the infinite; the camaraderie forged through shared struggles, and the subtle transformation of my spirit. As Rumi so beautifully put it, *"Be like a tree and let the dead leaves drop. Let yourself be silently drawn by the strange pull of what you really love. It will not lead you astray."* My PhD became a journey of falling and rising, of becoming—an extraordinary mosaic of moments that will forever remain etched in my soul.

First and foremost, I would like to express my heartfelt gratitude to my supervisor, **Dr. Girish Ratnaparkhi,** for always being there for me and helping me in every way possible. His unwavering support, guidance, and patience have been nothing short of remarkable. He has been a true mentor, friend, and advocate, always encouraging me to pursue science in my own way. Through all our fights and open discussions, I have grown immensely as a scientist and person.

I would also like to thank my Research Advisory Committee members **(Dr. Richa Rikhy, Dr. Nishad Matagne, and Dr. Mahendra Sonawane)** and the Thesis Review Committee members for their valuable feedback, constructive criticism, and insightful suggestions. I would also like to thank **Dr. Girish Deshpande** for his help in developing and structuring Caspar's story.

To my colleagues and friends, you have been the constellation in the night sky of my PhD journey, illuminating my path when darkness threatened to envelop it. To my lab mates, your shared curiosity and unyielding passion turned the sterile hum of the lab into a symphony of discovery. Through endless experiments, quiet frustrations, and moments of triumph, you became not just collaborators but kindred spirits in this quest for understanding. I am grateful to my past lab members **Prajna** and **Shweta** for mentoring me initially and teaching me fly genetics and molecular biology, and present lab members **Namrata, Lovleen, Kundan,**

Pulkit, Vidhyadesh, Adheena, Amrita, Sanhita, and Kanika for all the chai sessions and discussions. To my friends beyond the lab, you reminded me that life exists in the spaces between—spaces filled with laughter, warmth, and the gentle art of being present. **Souradeep** and **Debasmita** for being a constant source of joy in my life. Batchmates and friends **Arnab, Atreyi, Shivani, Gauri,** and **Mukundan** for all the game and fun sessions. Finally, I would like to thank my family members and my Mom and Dad for believing in me and encouraging me to always follow my passion and dreams.

I would also like to thank all the staff and faculty of IISER, Pune, whose dedication to academic excellence and support for graduate students have been instrumental in making this journey possible. I would also like to thank CSIR for the funding, NCBS Fly Facility for generating CRISPR flies and injections, and BDSC for reagents.

Table of Contents

Chapter III

3. Phosphorylation of the Dorsal/NFκB dimer plays a critical role in embryonic dorsoventral patterning

Chapter IV

4. Caspar specifies primordial germ cell count and identity in *Drosophila melanogaster*

Synopsis

Early embryonic development in *Drosophila* is controlled by maternal factors deposited in the egg during oogenesis. Two major events during this time include the specification of embryonic polarity and primordial germ cells. Embryonic polarity, which later specifies the future body axes, results from the asymmetric distribution of gene products during oogenesis and early embryogenesis. After fertilization, the *Drosophila* embryo undergoes 13 rapid nuclear divisions without cytoplasmic divisions, resulting in a syncytial blastoderm. Primordial germ cells (pole cells), which later form future gonads of the fly, are specified around nuclear cycle 9 (NC9) when a few nuclei from the center of the embryo move towards the posteriorly anchored specialized cytoplasm called germplasm (containing germ cell determinant) and assumes pole cell fate. Cell fate specification is a crucial event during this time as it lays out the map for the future development of the embryo. This is regulated by activity gradients of maternally deposited transcription factors and signaling molecules within the embryo, specifying the anterior-posterior (AP) and dorsoventral (DV) axes. Specific threshold concentrations of these maternal factors create multiple discrete domain s of gene expression along these axes, leading to finer subdivision of the embryo. Four gene networks define the polarity of the *Drosophila* embryo and establish the developmental axes. Three of these, the anterior, posterior, and terminal gene networks, work in conjunction to develop the AP axis, while the specification of DV polarity and DV axis is the responsibility of the DV gene network. Each gene network comprises a specific set of maternal factors that are actively regulated with the embryo through maternal to zygotic transition (MZT).

The specification of the anteroposterior axis during *Drosophila* development is laid out during oogenesis, much before fertilization and egg laying. *Bicoid* and *hunchback* are two maternal effect genes that are found in the anterior part of the embryo and are predominantly responsible for anterior patterning. In contrast, *caudal* and *nanos* are found towards the embryo's posterior pole and are predominantly concerned with abdominal patterning. The DV axis is specified by the location of the oocyte nucleus much before fertilization. The anterior dorsal localization of the oocyte nucleus results in the secretion of Gurken (Grk), a TGFα-like ligand that activates Torpedo (Top), the *Drosophila* Epidermal Growth Factor Receptor (EGFR) on the surface of the follicle cells lying in proximity giving them a dorsal cell fate. Without Grk, EGFR is not activated, and Pipe is expressed in the ventral follicle cells. Pipe activation results in a multi-signaling cascade, ultimately activating Toll Signaling. Dorsal (DL), the effector molecule of

the Toll Signaling pathway, translocates inside the nucleus and acts as a transcription factor by activating ventral-specific genes while repressing dorsal determinant genes. At the posterior end of the embryo, the terminal gene network comprising Groucho and Torso regulates terminal patterning.

Somatic sexual identity is regulated by the absolute concentration of zygotically transcribed, X-chromosome linked elements or XCEs (e.g., sisterless-a, scute or sisterless-b, runt) that work in conjunction with the uniformly distributed maternal regulators such as Daughterless. While these events happen around NDC11, germline/soma distinction is established even earlier, when nuclei dividing in the center of the embryo start their journey towards the nuclear periphery. A few nuclei precociously migrate into the posteriorly localized and anchored specialized cytoplasm or 'germ plasm' that contains RNA and protein determinants required for germ cell formation and specification. The recruitment and stepwise assembly of germ plasm components depends on Oskar, the critical player shown to be necessary and sufficient for the formation of functional PGCs.

Cell fate specification and embryonic patterning are also regulated by post-translational modifications (PTMs). Over the past two decades, SUMOylation has been identified as one of the most important post-translational modifiers to thousands of proteins in the mammalian system. SUMO (Small Ubiquitin-like Modifier) is a reversible post-translational modifier that shows great versatility in modulating the function of its target proteins. After conjugating with its target, SUMO can modulate its function in various ways, including its activity, localization, and stability (half-life). Current research has implicated the role of SUMO in the cell cycle, DNA repair, morphogenesis, and development, as well as some neurodegenerative diseases like ALS, Alzheimer's, and Huntington's diseases.

For my PhD project, I have worked on cell fate specification in the early embryo. I have looked at three genes that are crucial in early embryonic patterning. These three genes are **Dorsal**, the master regulator for DV patterning; **Groucho**, which is involved in terminal patterning; and **Caspar**, which is essential for pole cell formation and identity.

In **Chapter 1,** I discuss *Drosophila* early embryonic patterning and cell fate specification. During early embryogenesis, the zygotic genome remains largely inactive, and development is driven by the maternally deposited protein and mRNAs. As the embryo progresses, the maternal-to-zygotic transition (MZT) occurs, marked by maternal clearance and zygotic genome activation (ZGA), which reshapes gene expression and establishes zygotic control over

development. Key events during this period include embryonic patterning and cell fate specification. Embryonic polarity, which defines body axes, arises from asymmetrically distributed maternal gene products. Primordial germ cells (pole cells), which later form future gonads of the fly, are specified around nuclear cycle 9 (NC9) when a few nuclei from the center of the embryo move towards the posteriorly anchored specialized cytoplasm called germplasm (containing germ cell determinant) and assumes pole cell fate. Post-translational modifications, particularly SUMOylation, significantly influence embryonic patterning and cell fate specification. SUMO and its enzymes are abundant, underscoring their critical role. In this chapter, I aim to discuss how different gene networks govern embryonic patterning and specify cell fate in the early *Drosophila* embryo. I will also highlight how SUMOylation influences these events.

In **Chapter 2**, I explore the importance of SUMO conjugation to Groucho in the early development of *Drosophila melanogaster*. Here, I aim to study the SUMOylation of Groucho, a drosophila embryo terminal group Gene Regulatory Network member. Groucho is both SUMO conjugated and contains a putative SUMO binding site - a SUMO-interaction motif (SIM). SUMO Conjugation resistant (SCR) mutants for Groucho (K167R and K177R) were successfully generated using the CRISPR/Cas9 strategy. The double mutant Groucho line shows no apparent developmental defects, and the mutant Groucho protein was expressed in the correct positions, similar to the wild type. However, pull-down experiments suggest that the SCR Gro^{2M} line is defective in interacting with Dorsal, a key transcription factor. The preliminary results indicate that SUMOylation of Groucho is responsible for Dorsal-Groucho interaction and might prove a vital link in uncovering the role of SUMO in the dorsoventral and terminal patterning of the embryo. Although SIM sites in Groucho appear not to be involved in cis-SUMOylation, I find that it is essential for its stability and function.

In **Chapter 3,** I explored the role of SUMO Interacting Motif (SIM) in regulating Dorsal function and stability. Post-translational modifications such as phosphorylation and SUMOylation tightly regulate DL's activity. Phosphorylation at S312 stabilizes DL in the cytoplasm, while S317 phosphorylation regulates its nuclear import. Additionally, SUMOylation at K382 enhances DL's transcriptional activity, with a SUMO-resistant form of DL (dl^{SCR}) acting as a more effective transcription factor. I mutated two predicted SUMO-interacting motifs (SIMs) on DL in this study. While SIMα(VVVV) turned out to be redundant, SIMβ (IILL) was critical for dimerization. Mutation of SIMβ destabilized DL reduced its

SUMOylation, and triggered protein degradation. Interestingly, this degradation occurs through a novel pathway independent of proteasomal or autophagy mechanisms. To validate these findings, I generated CRISPR mutants with the SIMβ mutation, which resulted in embryonic lethality and dorsalized embryos. Although the DL^{β} protein was absent, its mRNA was stable, suggesting post-translational degradation. Molecular modeling showed that SIMβ forms a hydrophobic pocket essential for dimerization. When this pocket is disrupted, DL cannot dimerize, becomes unstable as a monomer, and is degraded. Through this work, I discovered that DL relies on dimerization for its stability and function. Phosphorylation can only occur in dimeric DL, and its absence leads to the degradation of DL and its partner protein, Cactus. I identified SIMβ as a novel site critical for DL's dimerization and stability, with mutations at this site rendering DL functionally null.

In Chapter 4, I worked on Caspar, the *Drosophila* orthologue of mammalian Fas Associated Factor 1 (FAF1). While Casp has a well-established role in the immune response, modENCODE RNAseq and proteomics data suggest that *casp* is also highly expressed in the 0-3 hour old embryo. Analysis of a hypomorphic allele of *casp* demonstrated that roughly half of the embryos maternally compromised in *casp*, fail to undergo gastrulation. Furthermore, such embryos display developmental defects, including aberrant cytoskeletal networks starting from early blastoderm stages. Interestingly, Casp is expressed strongly in primordial germ cells (PGCs). Consistent with the enrichment, maternal reduction of *casp* significantly affects the total number of pole cells. Here, I present an analysis of *casp* function during early embryonic development and its role in the formation and/or specification of PGCs. We show that Casp activity regulates Oskar levels and centrosome function, two critical determinants of PGC fate in *Drosophila* embryos. Upon loss of *casp*, the total amount of Oskar and Smaug changed reciprocally to influence the PGC count. Ubiquitin-based protein degradation is critically involved during early embryonic events, including the maternal-zygotic transition. We present a model explaining the involvement of Casp and its protein partner, TER94, during germ cell development, considering their influence on the clearance of Smaug, a critical regulator of maternal to zygotic transition (MZT).

Publications

- **Das S**, Hegde S, Wagh N, Sudhakaran J, Roy AE, Deshpande G, Ratnaparkhi GS. Caspar specifies primordial germ cell count and identity in *Drosophila melanogaster*. Elife. 2024 Dec 13;13:RP98584. doi: 10.7554/eLife.98584. PMID: 39671304; PMCID: PMC11643641.

- Deshpande G, **Das S**, Roy AE, Ratnaparkhi GS. A face-off between Smaug and Caspar modulates primordial germ cell count and identity in *Drosophila* embryos. Fly (Austin). 2025 Dec;19(1):2438473. doi: 10.1080/19336934.2024.2438473. Epub 2024 Dec 24. PMID: 39718186.

- Kushwaha S, Mallik B, Bisht A, Mushtaq Z, Pippadpally S, Chandra N, **Das Subhradip,** Ratnaparkhi G, Kumar V. dAsap regulates cellular protrusions via an Arf6-dependent actin regulatory pathway in S2R+ cells. FEBS Lett. 2024 Jun; 598(12):1491-1505. doi: 10.1002/1873-3468.14954. Epub 2024 Jun 11. PMID: 38862211.

- **Das Subhradip**, Roy Amrita, Mukundan S, Madhusudhan MS, Ratnaparkhi Girish (2025) Phosphorylation of the Dorsal/ NFκB dimer plays a critical role in embryonic dorsoventral patterning. (In preparation

Chapter 1: An introduction to Early embryonic patterning and cell fate specification in *Drosophila melanogaster*

1.1 Abstract

Embryogenesis in all animals begins with a single zygotic nucleus formed through parental pronuclear fusion. The zygote, a single cell containing maternal cytoplasm and a nuclear genome from both parents, undergoes a remarkable transformation into a complex organism with distinct structures and functions. This process is initially governed by maternally deposited mRNA and proteins in the egg, which regulate biosynthetic processes, mitotic divisions, and primary axis specification. During early embryogenesis, the zygotic genome remains largely inactive, and development is driven by these maternal factors. As the embryo progresses, the maternal-to-zygotic transition (MZT) occurs, marked by maternal clearance and zygotic genome activation (ZGA), which reshapes gene expression and establishes zygotic control over development.

Key events during this period include embryonic polarity and primordial germ cell specification. Embryonic polarity, which defines body axes, arises from asymmetrically distributed maternal gene products. Rapid nuclear divisions without cytoplasmic cleavage form the syncytial blastoderm, where primordial germ cells are specified by nuclei migrating to germplasm at nuclear cycle 9. Maternal transcription factor gradients specify anterior-posterior (AP) and dorsal-ventral (DV) axes by creating distinct domains of gene expression, further refining embryonic subdivisions. Four gene networks establish these axes: the anterior, posterior, and terminal systems define AP polarity, while the DV system specifies DV polarity.

Post-translational modifications, particularly SUMOylation, significantly influence embryonic patterning and cell fate specification. SUMO (Small Ubiquitin-like Modifier) is a versatile, reversible modifier regulating target protein activity, localization, and stability. SUMOylation has been linked to processes including the cell cycle, DNA repair, morphogenesis, and development, as well as neurodegenerative diseases like ALS and Alzheimer's. During early *Drosophila* embryogenesis, maternally deposited SUMO and its enzymes are abundant, underscoring their critical role. Proteomic studies have identified over 140 SUMOylation targets in early embryos, implicating this modification in cell cycle regulation, RAS/MAPK

signaling, and embryonic patterning. Maternal depletion of SUMO (*smt3*) results in high embryonic lethality, and loss of E1 enzymes (*Aos1* and *Uba2*) causes severe developmental defects and lethality at the larval-pupal transition. SUMOylation is essential for anterior-posterior and dorsal-ventral axis formation and overall embryonic development.

Keywords

Early development, *Drosophila melanogaster*, Embryo patterning, Cell fate specification, SUMOylation, Pole cells

1.2 Introduction

In all animals, a single zygotic nucleus formed by parental pronuclear fusion begins the process of embryogenesis. The embryo is a single cell composed of maternal cytoplasm and nuclear genome from both parents. What has long baffled scientists is how a zygote, which is a single cell finally forms a fully grown organism with such complex structure and function. The initial stages of development in *Drosophila* is controlled by maternally supplied mRNA and proteins loaded into the egg during oogenesis. These maternally deposited products are mainly involved in regulating the basic biosynthetic processes in the early embryo, including the first mitotic division and specifying the primary axis of development, cell fate, and embryo patterning. Most of the zygotic genome remains in a transcriptionally quiescent state during the early stages of embryogenesis, and as the development progresses, the zygotic genome takes over gradually. This *maternal to-zygotic transition* (MZT) is characterized by two major developmental events: first, the maternal clearance, whereby the maternally supplied products (mRNA and proteins) are cleared from the system, and second, the *zygotic genome activation* (ZGA), the transcriptional activation of the zygotic genome and establishing its control over the developmental processes **(Vastenhouw et al. 2019)(Hamm and Harrison 2018).** Together, these two events remodel the gene expression pattern in embryonic landscape and cell fates, a process that will be carried on throughout the entire lifetime of the organism.

Two significant events during this time include the specification of embryonic polarity and primordial germ cells. Embryonic polarity, which later specifies the future body axes, results from the asymmetric distribution of gene products during oogenesis and early embryogenesis. After fertilization, the embryo undergoes 13 rapid nuclear divisions without cytoplasmic divisions, resulting in a syncytial blastoderm **(Campos-Ortega J et al., 2013).** Primordial germ cells (pole cells), which later form future gonads of the fly, are specified around nuclear cycle 9 (NC9) when a few nuclei from the center of the embryo move towards the posteriorly anchored specialized cytoplasm called germplasm (containing germ cell determinant) and assumes pole cell fate **(Santos and Lehmann 2004).** Cell fate specification is a crucial event during this time as it lays out the map for the future development of the embryo. This is regulated by activity gradients of maternally deposited transcription factors and signaling molecules within the embryo, specifying the anterior-posterior (AP) and dorsoventral (DV) axes (**Figure 1)**. Specific threshold concentrations of these maternal factors create multiple discrete domains of gene expression along these axes, leading to finer subdivision of the embryo . Four gene networks define the polarity of the *Drosophila* embryo and establish the

developmental axes. Three of these, the anterior, posterior, and terminal gene networks, work in conjunction to develop the AP axis, while the specification of DV polarity and DV axis is the responsibility of the DV gene network **(Driever and Nüsslein-Volhard 1988)(Huang et al. 1997) (Stathopoulos and Levine 2002).**

Cell fate specification and embryonic patterning are also regulated by post-translational modifications (PTMs). Over the past two decades, SUMOylation has been identified as one of the most important post-translational modifiers to thousands of proteins in the mammalian system. SUMO (Small Ubiquitin-like Modifier) is a reversible post-translational modifier that shows great versatility in modulating the function of its target proteins. After conjugating with its target, SUMO can modulate its function in various ways, including its activity, localization, and stability (half-life). Current research has implicated the role of SUMO in the cell cycle, DNA repair, morphogenesis, and development, as well as some neurodegenerative diseases like ALS, Alzheimer's, and Huntington's diseases.

SUMO is expressed throughout the *Drosophila* development. However, during early embryogenesis, the amount of maternal SUMO and its enzymes remains highly abundant, suggesting a critical role during *Drosophila's* early development. Recently, two-step purification followed by mass spectrometric analysis of SUMO proteome in early *Drosophila* embryos identified over 140 direct SUMOylation targets **(Nie et al. 2009)**. Among these are players involved in key embryonic development processes, including cell cycle control, RAS/MAPK signaling, embryonic patterning, etc. This suggests a critical role of SUMO in early metazoan development. Removal of maternal *smt3* in *Drosophila* (using maternal driver) causes embryonic lethality (70% fail to hatch), and those that do die in their first instar larvae. However, if the zygotic *smt3* is knocked down, the larvae survive past their first instar but die in their second. *When mutated, Aos1 and Uba2, the Drosophila E1 enzymes* result in severe developmental defects like abnormal imaginal disc development and lethality during larval to pupal transition **(Kanakousaki and Gibson 2012)**. One of the most important events occurring during the early stages of embryogenesis is the formation of developmental axes (anterior-posterior and dorsoventral axis). SUMOylation plays a significant role in embryo patterning in *Drosophila* **(Johnston and Nüsslein-Volhard 1992)(Van Eeden and St Johnston 1999).**

In this chapter, I aim to discuss how different gene networks govern embryonic patterning and specify cell fate in the early *Drosophila* embryo. I will also highlight how SUMOylation influences these events.

1.3 The anterior-posterior gene network

The establishment of the anterior-posterior axis in *Drosophila* is one of the most well understood and extensively researched developmental processes. This was formerly laid by the groundbreaking work of Christiane Nüsslein-Volhard, Eric Wieschaus, and their colleagues, who conducted comprehensive forward genetic screens to identify mutations impacting this process. By the early 1990s, results from these screens and the free availability of mutants enabled scientists to understand RNA-level gene regulatory mechanisms governing the AP axis, thereby deepening our knowledge of this process. These early characterizations helped us to research further, and even after 30 years, new findings continue to emerge from this model system, advancing our understanding of how physical processes regulate genes involved in cellular and developmental switches. Nonetheless, significant questions remain unresolved.

The derived nature of anterior-posterior (AP) patterning in *Drosophila* relies on the protein Bicoid (Bcd). The ***bcd*** gene, encoding a homeodomain transcription factor, is found exclusively in higher dipterans. ***Bcd*** mRNA is maternally deposited during *Drosophila* oogenesis and is localized to the anterior pole of the egg. Following fertilization, ***bcd*** mRNA is translated, and the resulting protein diffuses from the anterior pole, forming a morphogenetic gradient with its highest concentration at the anterior pole of the embryo. This gradient enables Bcd to regulate gene transcription in a concentration-dependent manner **(Lynch and Desplan 2003)(McGregor 2005)(Stauber et al. 1999)(St. Johnston et al. 1989)(Berleth et al. 1988).**

Early nuclear divisions in the *Drosophila* embryo result in a syncytium, where the nuclei share a common cytoplasm, allowing Bcd and other factors to diffuse freely throughout the embryo. Bcd's transcriptional activity is closely linked with Hunchback (Hb). Maternal ***hb*** mRNA, unlike ***bcd***, is distributed uniformly within the egg but is regulated post-transcriptionally. During zygotic genome activation, Bcd activates zygotic ***hb*** transcription while maternal ***hb*** mRNA is translationally repressed by the posterior-specific protein Nanos (Nos). This interplay creates a higher concentration of Hb levels at the anterior of the embryo and low levels at the posterior. Together, the Bcd and Hb gradients activate anterior-specific genes necessary for forming anterior body structures **(Ma et al. 1996)(Simpson-Brose et al. 1994)(Driever and Nüsslein-Volhard 1989)(Lehmannt and Akam 1989).**

Loss of function mutations of ***bcd*** lead to severe anterior patterning defects, including the absence of head, thoracic, and certain abdominal segments. Conversely, overexpression of Bcd

can induce ectopic anterior structures, demonstrating that Bcd is both necessary and sufficient for anterior specification **(Frohnhofer and Niisslein-volhard 1986)**. Additionally, the loss of maternal and zygotic Hb disrupts all Bcd-regulated processes, suggesting that Hb is a co-factor for Bcd. The concentration gradients established by **Bcd** and **Hb** morphogens produce distinct effects on the transcription of downstream gap genes, which subsequently refine AP body patterning. These gradients are central to establishing the precise molecular framework required for anterior specification and subsequent embryonic development **(Simpson-Brose et al. 1994).**

Similar to anterior specification, *Drosophila* embryos rely on maternal factors to establish posterior specification. These anterior and posterior determinants interact to produce a graded anterior–posterior (AP) response across the embryo. This interplay is particularly evident in the absence of Bicoid (Bcd), where regions that should develop as anterior are misspecified as posterior due to the posteriorizing influence of Caudal (Cad) **(Frohnhofer and Niisslein-volhard 1986)(PhiL et al. 1982).**

cad mRNA is uniformly deposited throughout the developing oocyte, but its translation is repressed in the anterior by Bcd binding to the **cad** 3′ untranslated region (UTR). This repression establishes a gradient of Cad protein, with the highest levels at the posterior end of the embryo. Cad, a homeobox transcription factor, regulates posterior-specific gene expression in *Drosophila*, as it does in many other bilaterians. Another key posterior determinant is Nanos (Nos), whose mRNA is maternally deposited and localized to the posterior pole of the egg. After fertilization, **nos** mRNA is translated, creating a Nos protein gradient with its peak at the posterior end. Nos functions as a translational repressor, specifically inhibiting the translation of Hunchback (Hb) **(Lehmann and Nüsslein-Volhard 1991) (Gavis and Lehmann 1992)(Wang et al. 1994).** This repression ensures that Hb, an anterior-specific transcription factor, is excluded from the posterior regions of the embryo.

The first evidence that SUMOylation is required for embryo patterning came from the analysis of *ubc9* loss-of-function mutation, which led to the loss of anterior segments, a phenotype similar to that observed in *hunchback* mutation. Mutation of *Drosophila semushi (*semi) gene, which encodes the E2 enzyme (Ubc9), causes cuticle defects that range from the absence of just one thoracic segment (T1) to the absence of all the anterior segments (T1-A5) **(Epps and Tanda 1998)**. The accumulation of *bicoid* towards the anterior pole of the syncytial blastoderm determines the pattern in which zygotic segmentation genes are laid. The rapid nuclear division

during this phase requires quick translocation of *bicoid* into the nucleus after each mitotic cycle, which is regulated by RanGAP1. RanGAP1 is Sumoylated, and knocking down *Drosophila Ubc9* causes the mis-localization of *bicoid.* In a separate proteomic screen by Minghua et al., it was showed that both *hunchback* and *bicoid* are targets of SUMOylation, and disrupting the gene encoding SUMO causes a cascade of anteroposterior defects **(Nie et al. 2009).**

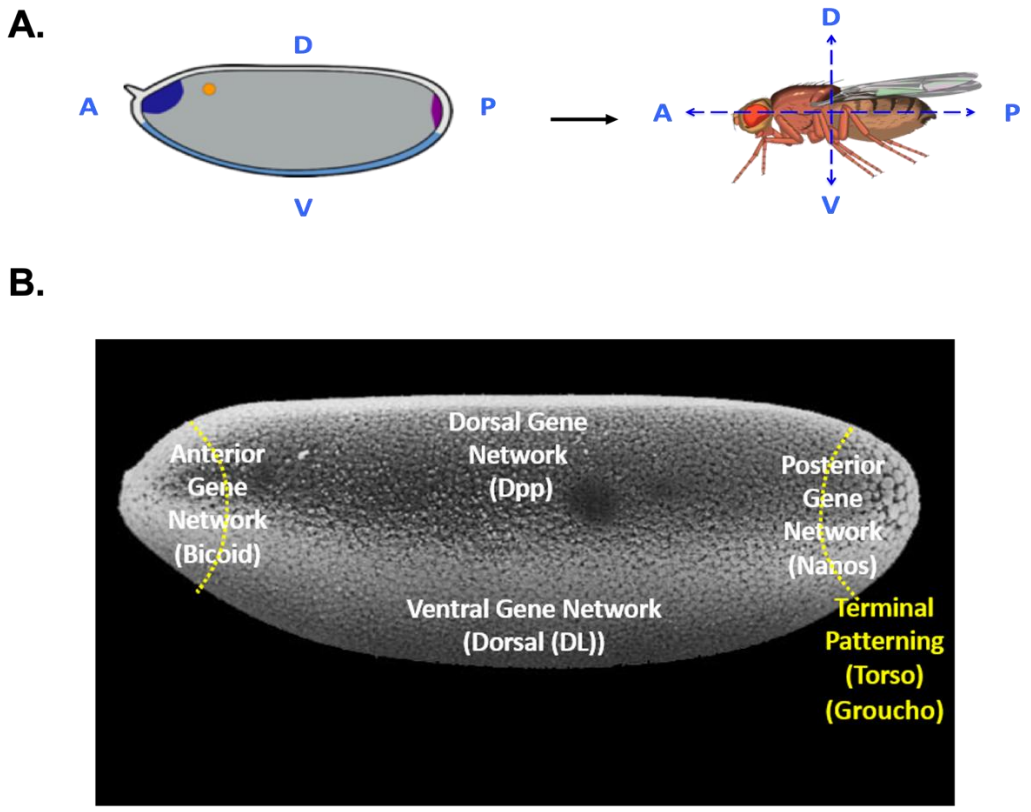


Figure 1.1 Developmental axis formation in early *Drosophila* embryo

(A) Axis specification from embryo to adult. A, P, D, and V correspond to anterior, aosterior, dorsal and ventral axis of the fly. For the fly to develop into a mature adult from the embryo, the anterior-posterior and dorso-ventral axis must be laid correctly before the gastrulation begins. **(B)** A scanning Electron Micrograph (SEM) image of *Drosophila* embryo at 2.5 hrs (stage 4-5). Embryonic polarity is mainly determined by maternal genes. While *nanos* and *bicoid* determine the anterior-posterior axis, *dorsal* and *dpp* governs the dorso-ventral axis. The terminal patterning by *torso* is independent of AP-DV axis.

Together, the activities of Bcd and Hb in the anterior and Cad and Nos in the posterior in conjunction to their fine-tuning by post translational modifications like SUMOylation establish a robust AP patterning system across the embryo. These maternal factors regulate downstream genes that further refine AP patterning and generate the precise segmental organization of the embryo **(Nüsslein-volhard et al. 1987).**

1.4 The dorsoventral gene network

The shape of the *Drosophila* egg reflects its dorsal-ventral polarity from the time it is laid. The eggshell is curved on the ventral side, which corresponds to the sembryo's prospective ventral region, while the flattened dorsal side is marked by a pair of respiratory appendages at the anterior. Around three hours post-egg-laying, following cellularization, distinct cell types along the dorsal-ventral axis begin to emerge. These differences are evident in changes in cell shape and division patterns during gastrulation and germ band elongation. Ultimately, the blastoderm cells differentiate into four distinct regions: cells at the ventral midline form mesodermal derivatives such as muscles, ventrolateral cells develop into the ventral epidermis and the ventral nerve cord, dorsolateral cells produce dorsal epidermis, and cells at the dorsal midline become the extraembryonic amnioserosa **(Stein and Stevens 2014).**

There are mainly three signaling pathways whose sequential crosstalk patterns the DV axis of the fly **(Morisato and Anderson 1995).**

1. **Oogenesis and Gurken/Torpedo Signaling**: The DV polarity is already established even before the egg is laid **(Roth and Schüpbach 1994)**. The germline-derived oocyte is surrounded by somatically derived follicle cells that secrete components of the eggshell. The anterior dorsal localization of the oocyte nucleus results in the secretion of Gurken (Grk), a TGFα-like ligand that activates Torpedo (Top), the *Drosophila* Epidermal Growth Factor Receptor (EGFR) on the surface of the follicle cells lying in proximity giving them a dorsal cell fate **(Ruohola-Baker et al. 1994)(Schüpbach 1987).**
2. **Ventral Signaling and Dorsal/Cactus Pathway**: After fertilization, the second pathway involves eleven dorsal group genes and **cactus**, which generate a ventral signal. Without gurken, EGFR is not activated, and Pipe is expressed in the distal follicle cells. Pipe activation results in a multi signaling cascade, ultimately activating Toll Signaling. Dorsal (DL), the effector molecule of the Toll Signaling pathway. This pathway relies on **spätzle** (the ligand) and **Toll** (the receptor), with spatial restriction achieved through localized proteolytic activation of the Toll ligand, regulated by a cascade originating in the follicle cells. Activation of Toll triggers the nuclear translocation of the Dorsal protein, a transcription factor, leading to a graded ventral-to-dorsal distribution of Dorsal. This gradient governs the transcriptional activation or repression of a set of zygotic genes **(Stein and Stevens 2014) (Morisato and Anderson 1995)(Govind and Steward 1991).**

3. **Dorsal Patterning and Decapentaplegic/TGF-β Signaling**: The third pathway establishes dorsal patterning through the ligand **decapentaplegic (dpp)**, a member of the TGF-β family. The response to Dpp is mediated by a receptor complex composed of **punt** (type II receptor) and **thick veins** or **saxophone** (type I receptors). A dorsal-to-ventral gradient of Dpp activity arises through post-translational enhancement on the dorsal side and inhibition on the ventral side.

The role of SUMOylation in the dorsoventral patterning of the embryo came from the proteomic screen, which identified *dorsal* as a target for SUMOylation in early embryos **(Nie et al. 2009)**. However, the mechanism behind how SUMOylation regulates *dorsal*-dependent transcription during early embryonic patterning is still unclear

A better example of how SUMOylation regulates dorsoventral patterning comes from its role in Decapentaplegic (*Dpp*) signaling **(Miles et al. 2008)**. Decapentaplegic is a BMP morphogen critical for dorsal ectoderm patterning through the SMAD family of transcription factors, like the *Drosophila* Mad and Medea (*Med)* proteins. Dpp first phosphorylates Mad, which translocates to the nucleus and activates Med. Mad-Med heterodimer then functions as a transcription factor activating Dpp target genes **(Affolter et al. 2001; Miles et al. 2008)**. When the SUMO pathway genes are disrupted, an expanded expression of Dpp targets suggests a possible mechanism of how SUMO may regulate Dpp signaling by restricting its range of action. Med, a target for SUMOylation when mutated at all three potential SUMO conjugation sites, shows a similar expression profile of Dpp target genes. Supporting this fact, overexpression of fused Med-SUMO protein leads to partial retraction of Dpp target gene expression, suggesting that the SUMOylation of Med acts as a negative regulator of Dpp target genes **(Long et al. 2004).**

Although there is no hard evidence suggesting how SUMOylation of Med negatively regulates Dpp signaling, several experiments suggest reduced motility of unmodified Med. When GFP-Med fusion proteins were photobleached, the rate of fluorescence recovery in the nucleus (due to translocation of unbleached cytoplasmic GFP-Med to the nucleus) was found to be significantly reduced in triple mutant Med background. Thus, the possibility that SUMOylation increases Med mobility suggests that the cyclic conjugation and deconjugation of Med during nuclear shuttling tightly regulates the expression of Dpp target genes in a spatio-temporal manner.

These signaling pathways collectively create the robust dorsal-ventral polarity essential for proper embryonic development, with each pathway sequentially refining the axis and contributing to the precise spatial organization of cell fates.

1.5 The terminal gene network

The fourth gene network that patterns the terminal region of the embryo involves mainly Torso, Groucho and other related proteins. Several morphogenetic effects of the anterior-posterior (AP) and dorsal-ventral (DV) patterning systems, such as segmentation and mesoderm formation, are suppressed at the terminal regions of the *Drosophila* embryo. This suppression is mediated by the terminal patterning system, which was identified through genetic screens for patterning mutants. In these studies, larval cuticle preparations from mutants showed normal segmentation patterns but lacked head and tail structures, including the larval head skeleton, filzkorper, and posterior spiracles. Conversely, some mutations caused expansion of head and tail structures at the expense of the segmented trunk. These phenotypes were linked to gain- or loss-of-function mutations in the receptor tyrosine kinase (RTK) Torso, revealing its role in terminal patterning **(Goyal et al. 2018).** Further research demonstrated that Torso signaling activates the canonical Ras/ERK cascade, which acts as a maternal cue specifying head- and tail-specific structures while suppressing mesodermal and trunk fates. This signaling also ensures proper progression of morphogenetic processes initiated by AP and DV systems, such as germband extension, which is disrupted in the absence of Torso activity **(Irvine and Wieschaus 1994)**.

Studies on the terminal patterning system have evolved from analyzing larval cuticle preparations to highly quantitative investigations of molecular and cellular dynamics in live embryos. With advances in genome editing for fluorescent tagging of terminal components and optogenetic manipulation of Ras/ERK signaling, researchers have achieved unprecedented spatiotemporal resolution. Unlike AP and DV systems, which utilize factors like Bicoid and Toll exclusively during early embryogenesis, RTK signaling in the terminal system is reused throughout multiple developmental stages in *Drosophila* and other animals. This mechanism provides precise spatial and temporal transcriptional control, as seen in diverse developmental contexts such as insect compound eye formation and mammalian neocortex patterning.

RTK signaling outcomes are highly sensitive to variations in the spatial distribution and temporal duration of activation. This sensitivity has been highlighted in genetic studies of

model organisms and in human developmental disorders linked to deregulated RTK signaling. To fully understand RTK signaling, it is essential to quantify the spatiotemporal limits of activation that are critical for avoiding developmental defects. Terminal patterning involves a gradient of activated RTK signaling, with the highest signal at the poles gradually decreasing toward the center of the embryo. This graded signal elicits multiple transcriptional and cell fate responses, consistent with the classical concept of a morphogen. Interestingly, the gradient is shaped by nuclear trapping, wherein nuclear import and export of ERK modulate its diffusion profile. Understanding how this mechanism establishes positional information in the anterior and posterior regions remains an open question **(Jindal et al. 2015)**.

The terminal patterning system has been instrumental in identifying conserved genes with relevance to human diseases. For instance, the **corkscrew (csw)** gene, named for its twisted loss-of-function phenotype, encodes a non-receptor tyrosine phosphatase essential for RTK signaling. Deregulation of **csw** is implicated in developmental syndromes and cancers. Similarly, **capicua (cic)**, which was identified for its loss-of-function phenotype in early embryos, has been associated with neurodegenerative diseases and neural cancers. Advances in live imaging and quantitative techniques have positioned the terminal system as one of the best-characterized RTK-dependent pattern formation processes. This has propelled our understanding of signaling and morphogenetic mechanisms and highlighted key unresolved questions, including how ERK gradients define positional information and the broader implications of RTK signaling in development and disease **(Perkins et al. 1992)(Jiménez et al. 2000)(Simón-Carrasco et al. 2018).**

1.6 Germ cell fate specification

Somatic sexual identity is regulated by the absolute concentration of zygotically transcribed, X-chromosome linked elements or XCEs (e.g., sisterless-a, scute or sisterless-b, runt) that work in conjunction with the uniformly distributed maternal regulators such as Daughterless **(Cline and Meyer 1996).** While these events happen around NDC11, germline/soma distinction is established even earlier, when nuclei dividing in the center of the embryo start their journey towards the nuclear periphery. A few nuclei precociously migrate into the posteriorly localized and anchored specialized cytoplasm or ‘germ plasm’ that contains RNA and protein determinants required for germ cell formation and specification. The recruitment and stepwise

assembly of germ plasm components depends on Oskar, the critical player shown to be necessary and sufficient for the formation of functional PGCs.

At the posterior pole, **osk** serves as the key maternal determinant for germ cell specification and the establishment of posterior somatic structures **(Ephrussi et al. 1991)(Kim-Ha et al. 1991)(Lehmann and Nüsslein-Volhard 1986). osk** mRNA localizes to the posterior pole through a series of mechanisms and is subsequently translated. The resulting Osk protein organizes germ plasm and polar granules in the oocyte and early embryo **(Ephrussi et al., 1991; Smith et al., 1992).** Polar granules, containing key components like Vasa (a DEAD-box RNA helicase), Tudor, and Aubergine (a Piwi-family protein interacting with piRNAs), are essential for primordial germ cell (PGC) specification. Additional mRNAs, such as **nanos (nos)**, **germ cell-less (gcl)**, and **polar granule component (pgc)**, are recruited to polar granules and play roles in posterior patterning and PGC development **(Jongens et al. 1992)(Kobayashi and Lasko 2016).**

While **nos** is the primary determinant of posterior patterning, it is not required for PGC specification but is crucial for their subsequent development **(Asaoka-Taguchi et al. 1999)(Forbes and Lehmann 1998).** After fertilization, as zygotic nuclei migrate to the posterior pole, Osk protein also drives the formation of nuclear germ granules, specialized ribonucleoprotein particles distinct from polar granules. These granules, or potentially Osk itself, promote pole cell division through mechanisms yet to be fully understood **(Kistler et al. 2018).**

1.7 References

Affolter M, Marty T, Vigano MA, Jaźwińska A. (2001). Nuclear interpretation of Dpp signaling in Drosophila. EMBO J. 20(13):3298–3305. doi:10.1093/emboj/20.13.3298.

Asaoka-Taguchi M, Yamada M, Nakamura A, Hanyu K, Kobayashi S. (1999). Maternal Pumilio acts together with Nanos in germline development in Drosophila embryos. Nat Cell Biol. 1(7):431–437. doi:10.1038/15666.

Berleth T, Burri M, Thoma G, Bopp D, Richstein S, Frigerio G, Noll M, Nüsslein-Volhard C. (1988). The role of localization of bicoid RNA in organizing the anterior pattern of the Drosophila embryo. EMBO J. 7(6):1749–1756. doi:10.1002/j.1460-2075.1988.tb03004.x.

Cline TW, Meyer BJ. (1996). Vive la différence: Males vs Females in Flies vs Worms. Annu Rev Genet. 30(31):637–702. doi:10.1146/annurev.genet.30.1.637.

Driever W, Nüsslein-Volhard C. (1988). A gradient of bicoid protein in Drosophila embryos. Cell. 54(1):83–93. doi:10.1016/0092-8674(88)90182-1.

Driever W, Nüsslein-Volhard C. (1989). The bicoid protein is a positive regulator of hunchback transcription in the early Drosophila embryo. Nature. 337(6203):138–143. doi:10.1038/337138a0.

Van Eeden F, St Johnston D. (1999). The polarisation of the anterior-posterior and dorsal-ventral axes during Drosophila oogenesis. Curr Opin Genet Dev. 9(4):396–404. doi:10.1016/S0959-437X(99)80060-4.

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

Epps JL, Tanda S. 1998. The Drosophila semushi mutation blocks nuclear import of bicoid during embryogenesis. Curr Biol. 8(23):1277–1280. doi:10.1016/s0960-9822(07)00538-6.

Forbes A, Lehmann R. (1998). Nanos and Pumilio have critical roles in the development and function of Drosophila germline stem cells. Development. 125(4):679–690. doi:10.1242/dev.125.4.679.

Frohnhofer HG, Niisslein-volhard C. (1986). 324120a0. (V).

Gavis ER, Lehmann R. (1992). Localization of nanos RNA controls embryonic polarity. Cell. 71(2):301–313. doi:10.1016/0092-8674(92)90358-J.

Govind S, Steward R. (1991). Dorsoventral pattern formation in Drosophila: signal transduction and nuclear targeting. Trends Genet. 7(4):119–125. doi:10.1016/0168-9525(91)90456-Z.

Goyal Y, Schüpbach T, Shvartsman SY. (2018). A quantitative model of developmental RTK signaling. Dev Biol. 442(1):80–86. doi:10.1016/j.ydbio.2018.07.012.

Hamm DC, Harrison MM. (2018). Regulatory principles governing the maternal-to-zygotic transition: insights from Drosophila melanogaster. Open Biol. 8(12):180183. doi:10.1098/rsob.180183. https://royalsocietypublishing.org/doi/10.1098/rsob.180183.

Huang AM, Rusch J, Levine M. (1997). An anteroposterior dorsal gradient in the Drosophila embryo. Genes Dev. 11(15):1963–1973. doi:10.1101/gad.11.15.1963.

Irvine KD, Wieschaus E. 1994. Irvine (1994)_CellIntercalation. 841:827–841.

Jiménez G, Guichet A, Ephrussi A, Casanova J. (2000). Relief of gene repression by Torso RTK signaling: Role of capicua in Drosophila terminal and dorsoventral patterning. Genes Dev. 14(2):224–231. doi:10.1101/gad.14.2.224.

Jindal GA, Goyal Y, Burdine RD, Rauen KA, Shvartsman SY. (2015). Erratum: RASopathies: unraveling mechanisms with animal models (DMM Disease Models and Mechanisms 8 (769-782)). DMM Dis Model Mech. 8(9):1167. doi:10.1242/dmm.022442.

St. Johnston D, Driever W, Berleth T, Richstein S, Nusslein-Volhard C. (1989). Multiple steps in the localization of bicoid RNA to the anterior pole of the Drosophila oocyte. Development. 107(SUPPL.):13–19. doi:10.1242/dev.107.supplement.13.

Johnston DS, Nüsslein-Volhard C. (1992). The origin of pattern and polarity in the Drosophila embryo. Cell. 68(2):201–219. doi:10.1016/0092-8674(92)90466-P.

Jongens TA, Hay B, Jan LY, Jan YN. (1992). The germ cell-less gene product: A posteriorly localized component necessary for germ cell development in Drosophila. Cell. 70(4):569–584. doi:10.1016/0092-8674(92)90427-E.

Kanakousaki K, Gibson MC. (2012). A differential requirement for SUMOylation in proliferating and non-proliferating cells during Drosophila development. Development.

139(15):2751–2762. doi:10.1242/dev.082974.

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

Kistler KE, Trcek T, Hurd TR, Chen R, Liang FX, Sall J, Kato M, Lehmann R. (2018). Phase transitioned nuclear oskar promotes cell division of drosophila primordial germ cells. Elife. 7:1–35. doi:10.7554/eLife.37949.

Kobayashi S, Lasko PF. (2016). Requirement for a Noncoding RNA in Drosophila Polar Granules for Germ Cell Establishment Author (s): Akira Nakamura , Reiko Amikura , Masanori Mukai , Satoru Kobayashi and Paul F . Lasko Published by : American Association for the Advancement of Science Stable URL : http://www.jstor.org/stable/2891360 Requirement for a Noncoding RNA in Drosophila Polar Granules for Germ Cell Establishment. 274(5295):2075–2079.

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

Lehmann R, Nüsslein-Volhard C. (1991). The maternal gene nanos has a central role in posterior pattern formation of the Drosophila embryo. Development. 112(3):679–691. doi:10.1242/dev.112.3.679.

Lehmannt R, Akam M. (1989). activity. 338(April):646–648.

Long J, Wang G, Dongming HE, Liu F. (2004). Repression of Smad4 transcriptional activity by SUMO modification. Biochem J. 379(1):23–29. doi:10.1042/BJ20031867.

Lynch J, Desplan C. (2003). Evolution of development: Beyond bicoid. Curr Biol. 13(14):557–559. doi:10.1016/S0960-9822(03)00472-X.

Ma X, Yuan D, Diepold K, Scarborough T, Ma J. (1996). The Drosophila morphogenetic protein bicoid binds DNA cooperatively. Development. 122(4):1195–1206. doi:10.1242/dev.122.4.1195.

McGregor AP. (2005). How to get ahead: The origin, evolution and function of bicoid. BioEssays. 27(9):904–913. doi:10.1002/bies.20285.

Miles WO, Jaffray E, Campbell SG, Takeda S, Bayston LJ, Basu SP, Li M, Raftery LA,

Ashe MP, Hay RT, et al. (2008). Medea SUMOylation restricts the signaling range of the Dpp morphogen in the Drosophila embryo. Genes Dev. 22(18):2578–2590. doi:10.1101/gad.494808.

Morisato D, Anderson K V. (1995). Signaling pathways that establish the dorsal-ventral pattern of the Drosophila embryo. Annu Rev Genet. 29:371–399. doi:10.1146/annurev.genet.29.1.371.

Nie M, Xie Y, Loo JA, Courey AJ. (2009). Genetic and proteomic evidence for roles of Drosophila SUMO in cell cycle control, Ras signaling, and early pattern formation. PLoS One. 4(6). doi:10.1371/journal.pone.0005905.

Nüsslein-volhard AC, Frohnhöfer HG, Lehmann R, Ntysslein-volhard C, Frohnhofer HG, Lehmann R. (1987). Linked references are available on JSTOR for this article : Determination of Anteroposterior Polarity in Drosophila. 238(4834):1675–1681.

Perkins LA, Larsen I, Perrimon N. (1992). Corkscrew Encodes a Putative Protein Tyrosine Phosphatase That Functions To Transduce the Terminal Signal From the Receptor Tyrosine Kinase Torso. Cell. 70(2):225–236. doi:10.1016/0092-8674(92)90098-W.

PhiL AH, Macdonald PM, Struhf G. (1982). 18. Vittorio, N. & Silk. 42 Ostriker, J Vishniac, E Astrophys J. 307(9):1655–1657.

Roth S, Schüpbach T. (1994). The relationship between ovarian and embryonic dorsoventral patterning in Drosophila. Development. 120(8):2245–2257. doi:10.1242/dev.120.8.2245.

Ruohola-Baker H, Jan LY, Jan YN. (1994). The role of gene cassettes in axis formation during Drosophila oogenesis. Trends Genet. 10(3):89–94. doi:10.1016/0168-9525(94)90231-3.

Santos AC, Lehmann R. (2004). Germ cell specification and migration in Drosophila and beyond. Curr Biol. 14(14):578–589. doi:10.1016/j.cub.2004.07.018.

Schüpbach T. (1987). Germ line and soma cooperate during oogenesis to establish the dorsoventral pattern of egg shell and embryo in Drosophila melanogaster. Cell. 49(5):699–707. doi:10.1016/0092-8674(87)90546-0.

Simón-Carrasco L, Jiménez G, Barbacid M, Drosten M. (2018). The Capicua tumor suppressor: a gatekeeper of Ras signaling in development and cancer. Cell Cycle.

17(6):702–711. doi:10.1080/15384101.2018.1450029.

Simpson-Brose M, Treisman J, Desplan C. (1994). Synergy between the hunchback and bicoid morphogens is required for anterior patterning in Drosophila. Cell. 78(5):855–865. doi:10.1016/S0092-8674(94)90622-X.

Stathopoulos A, Levine M. (2002). Dorsal gradient networks in the Drosophila embryo. Dev Biol. 246(1):57–67. doi:10.1006/dbio.2002.0652.

Stauber M, Jäckle H, Schmidt-Ott U. (1999). The anterior determinant bicoid of Drosophila is a derived Hox class 3 gene. Proc Natl Acad Sci U S A. 96(7):3786–3789. doi:10.1073/pnas.96.7.3786.

Stein DS, Stevens LM. (2014). Maternal control of the Drosophila dorsal-ventral body axis. Wiley Interdiscip Rev Dev Biol. 3(5):301–330. doi:10.1002/wdev.138.

Vastenhouw NL, Cao WX, Lipshitz HD. (2019). The maternal-to-zygotic transition revisited. Development. 146(11):dev161471. doi:10.1242/dev.161471. https://doi.org/10.1242/dev.161471.

Wang C, Dickinson LK, Lehmann R. (1994). Genetics of nanos localization in Drosophila. Dev Dyn. 199(2):103–115. doi:10.1002/aja.1001990204.

Chapter 2: Understanding the functional significance of SUMO conjugation to Groucho in terminal patterning of *Drosophila* embryo

2.1 Abstract

A hallmark of early embryonic development is the establishment of the body axes: the anterior-posterior, the dorsoventral, and the left-right axis. Maternal proteins work towards establishing these axes, even before fertilization. A significant fraction of these proteins are post-translationally modified, but very little is known about the modulation of protein function by post-translational modifiers (PTMs) in early embryonic development. The small ubiquitin-like protein modifier SUMO, has been identified as a major regulator of cellular processes. SUMO, as well as the enzymes that help conjugate SUMO to target proteins, are maternally deposited.

Previous studies, including those from our laboratory, have demonstrated that reducing maternal SUMO conjugation in *Drosophila melanogaster* disrupts the developmental program, leading to embryonic lethality and axis patterning defects. Many of the biological functions of SUMO depend on SIM (Sumo Interacting Motif). Although very little is known about SIM, it is becoming clear that a large fraction, perhaps most of the functions regulated by SUMO relies on the non-covalent interaction of covalently conjugated SUMO with SIMs of other interacting proteins.

Here, we aim to study the SUMOylation of Groucho, a member of the *drosophila* embryo terminal group Gene Regulatory Network. Groucho is both SUMO conjugated and contains a putative SUMO binding site - a SUMO-interaction motif (SIM). SUMO Conjugation Resistant mutants for Groucho (K167R and K177R) were successfully generated using the CRISPR/Cas9 strategy and validated by sequencing. The double mutant Groucho (Gro^{2M}) line shows no apparent developmental defects, and the mutant Groucho protein was expressed in the correct positions, similar to the wild type. However, pull-down experiments suggest that the SCR Gro^{2M} line is defective in its interaction with Dorsal, a key transcription factor.

The preliminary results that seem to suggest that SUMOylation of groucho is responsible for Dorsal-Groucho interaction might prove a vital link in uncovering the role of SUMO in dorsoventral and terminal patterning of the embryo. Although SIM sites in Dorsal and Groucho appear not to be involved in cis-SUMOylation, we hypothesize that these motifs may mediate

Dorsal-Groucho interaction or interactions with other proteins critical for embryonic patterning.

Our findings highlight a potentially novel role for SUMOylation in early embryonic development, providing insights into the complex regulatory mechanisms underlying axis specification and the establishment of embryonic patterning.

Keywords: Immunity, PTM, signal transduction, SIM, Ubiquitin, Activation

2.2 Introduction

The *Drosophila groucho (gro)* gene was first discovered when the flies developed a thick tuft of sensory bristles over their eye, resembling the bushy eyebrows of Groucho Marx, as a result of a weak hypomorphic allele (gro^1) **(Lindsley and Grell, 1968)**. Groucho (gro) is the founding member of the Gro/TLE family of metazoan co-repressors. Gro mediates repression through physical interaction with a wide array of DNA-binding repressors to direct gene silencing and is thereby involved in a myriad of developmental processes, including neurogenesis, axis patterning, and sex determination. Groucho family members are essential regulators of critical signaling pathways such as Wingless/Wnt, Notch, and Dpp/TGFß. Groucho carries out its co-repressor role through multiple molecular mechanisms, depending on the regulatory context **(Buscarlet & Stifani, 2007)**. It is recruited by a specific set of transcriptional repressors via a short conserved sequence, which includes the WRPW or Eh1 motif. Groucho domain analysis shows that the N terminal Q domain is responsible for its tetramerization and transcriptional repression, while the CcN and SP domains are essential in nuclear localization and signal reception. The C-terminal WD domain functions as a site for repressor binding **(Nuthall et al., 2004) (Kim et al., 2005)** ***(Figure 2.1A)*****.** Groucho forms high molecular weight complexes with its repressor and represses transcription of genes by recruiting HDAC1, which binds to deacetylated histones and directs chromatin condensation **(Kim et al., 1999)(Yochum & Ayer, 2001).**

Groucho co-repressor was found to be covalently conjugated to SUMO-1 at multiple lysine residues in COS-7 cells. Deletion analysis on Groucho determined four distinct lysines within the GP and SP domains, which are SUMOylated. These lysines are K167, K177, K272 and K279. SUMOylation of Groucho was shown to increase its co-repressor activity by luciferase reporter assay, whereas Groucho SUMO conjugation resistant (SCR) mutant showed loss of co-repressor activity. This is probably done by efficiently recruiting HDAC1 through its SUMO Interacting Motif (SIM). It was shown that wild-type HDAC1 binds more strongly to SUMO-Groucho fusion protein than wild-type Groucho, whereas HDAC1 SIM mutant did not **(Ahn et al., 2009)**. Protein-protein interaction shows that Dorsal and Groucho can bind to each other directly.

Furthermore, Dorsal can activate but not repress transcription in embryos lacking Groucho co-repressor, which strongly suggests that Dorsal can recruit Groucho to a template where it can repress transcription. Thus, Groucho plays a predominant role in Dorsal-mediated dorsoventral patterning of the embryo **(Dubnicoff et al., 1997)**. Here, we aim to test the importance of

Groucho SUMOylation in the patterning and development of *Drosophila melanogaster*. CRISPR-Cas9 technology was used to generate SUMO conjugation-resistant Groucho flies. We have also cloned the SIM mutant *groucho* construct and expressed it in *Drosophila* 529Su cells to test its effect on Groucho sumoylation.

2.3 Results

2.3.1 Generation of SUMO conjugation resistant (SCR) mutants for Groucho

In recent years, the advent of CRISPR-Cas9 genome editing technology has allowed the generation of point mutations in a straightforward and site-directed manner **(Bassett & Liu, 2014; Bier & Nizet, 2021; Gratz et al., 2014).** Groucho has four distinct lysine residues (K167, K177, K272, and K279) subjected to SUMO conjugation. We targeted these lysines using CRISPR-Cas9 technology to generate a 4M Groucho (K167R, K177R, K272R and K279R). The genome-edited lines are a cleaner way to look at the effects of SUMOylation, as the level of mutant transcript/protein will be driven by its natural promoter. These lines are generated with the help of Dr. Deepti Trivedi, NCBS. The Groucho SCR lines were made using the Ds-Red Scarless editing strategy. In this strategy, two homology arms containing the desired lysine to arginine mutations were inserted into a pHD-Scarless DsRed vector, and the final donor was injected into embryos expressing cas9 in the *vasa* domain. The Groucho homology arm 1 (HR1) contained two mutations (K167R and K177R), and the homology arm 2 (HR2) cont ained two mutations (K272R and K279R). pHD-Scarless DsRed is a donor vector that has a 3xP3-DsRed marker with PBac transposon ends flanking both sides. This vector can be used to generate targeted insertions, deletions, SNPs, or any other modification with minimal locus disruption. Following DsRed-mediated identification of transgenic lines, the marker cassette can easily be removed by crossing the fly with a PBac transposase ***(Figure 2.1B)*.**

We received seven lines that showed a positive expression of the Ds-Red marker in the eye. These seven Groucho lines (named L1-L7) were taken forward to validate the mutation's presence through sequencing. Single-fly genomic DNA prep was made from each of these

lines, and the entire Groucho cassette containing the Ds-Red marker was amplified by PCR. Each single PCR band was then gel extracted, and purified DNA was sent for sequencing.

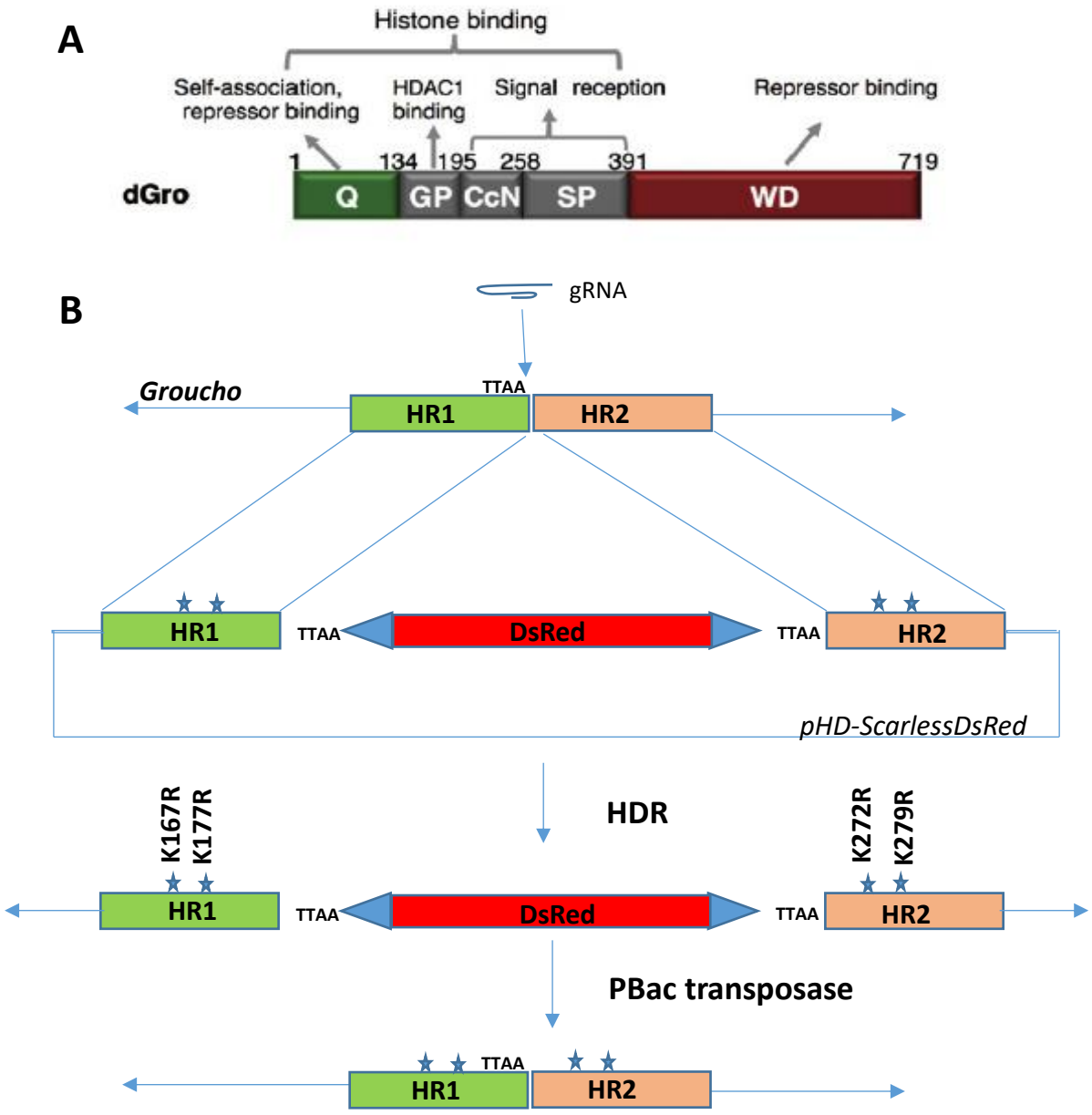


Figure 2.1 Design of the Groucho CRISPR mutant.

(A) Domain organization in the Gro/TLE family. Most Gro/TLE family members share a conserved structural organization, which consists of a glutamine-rich (Q) domain, followed by a glycine/proline-rich (GP) domain, a so-called CcN domain, a serine/proline-rich (SP) domain, and a WD-repeat domain responsible for repressor binding. (Wiam Turki-Judeh and Albert J. Courey, 2012)**. (B)** The Groucho SCR lines were made using the Ds-Red Scarless editing strategy. In this strategy, two homology arms containing K-R mutations were inserted into a pHD-Scarless DsRed vector, and the final donor was injected into cas9 flies. The Groucho homology arm 1 (HR1) contained two mutations (K167R and K177R), and the homology arm 2 (HR2) contained two mutations (K272R and K279R).

The sequencing result was aligned against the template, and it was found that only one line (L7) out of the seven lines contained two out of four mutations. From here on, we will call it Groucho^{2M}. The two mutations K167R and K177R lie in the Groucho homology arm 1 (HR1). The rest of the lines had no mutations in them. L2 had multiple random mutations in the sequence; hence, this line was no longer used for further experiments.

2.3.2 Groucho^{2M} does not show any defects in lifespan/ embryonic lethality or protein expression

We first checked if the mutant Groucho line (L7) showed any obvious visible phenotype compared to the wild type. Since Groucho plays an essential role during the early embryo patterning (dorsoventral and terminal patterning) and SUMOylation of Groucho regulates its co-repressor activity, we decided to check if the mutant line shows any embryonic defects.

First, we checked the embryonic lethality of the Groucho CRISPR lines. Around 80 females and 20-30 males are put together in a cage. The flies were fed with yeast paste to stimulate egg production. The flies were allowed to lay eggs for 12 hours, after which the cages were changed. The eggs collected on the sugar agar plate were arranged on another plate with 200 embryos per plate. The embryos were incubated at 25°C, and after 24 h, the number of hatched embryos were counted. This exact procedure was repeated after 48 h. The ratio of unhatched to total embryos plated for each genotype gives us the lethality of the embryos. Three replicates of the experiment were done. The Groucho two mutant (Gro^{2M}) line (L7) embryos showed no significant lethal defects compared to the control ***(Figure 2.2B)*.**

Next, we did a life span assay to check if the Gro^{2M} L7 line shows a normal life span compared to the control. The entire assay was done at 29°C, i.e., stress. Since Groucho shows a variable amount of expression in males and females, we checked the life span assay separately for males and females of each genotype. The number of males and females assayed for each genotype was 60. The CRISPR lines (L1-L7 except L2) survived for 30-34 days under the specified conditions. There was no significant difference between the males and females of the CRISPR lines. The control set (w^{1118}) survived for 40-45 days (both males and females) at 29°C. The lower life span for the CRISPR lines compared to the w^{1118} control was probably due to the gene disruption and insertion of the Ds-Red cassette ***(Figure 2.2C, D)*.**

Further, we checked for proper Groucho expression in these lines using a western blot. Since Groucho expression is the highest in early embryos, we collected 0-4 hour embryos, crushed them in RIPA buffer (supplemented with protease inhibitor complex), and loaded an equal amount of protein after protein estimation by BCA assay. All 6 CRISPR lines (L1-L7 except L2) expressed Groucho at the correct position (~80 kDa) compared to control (w^{1118}). The two mutant L7 also showed proper Groucho expression. This data confirms that there exists no problem with Groucho translation, and the said protein is expressed properly in the mutant lines like its wild-type counterpart ***(Figure 2.2A)*.**

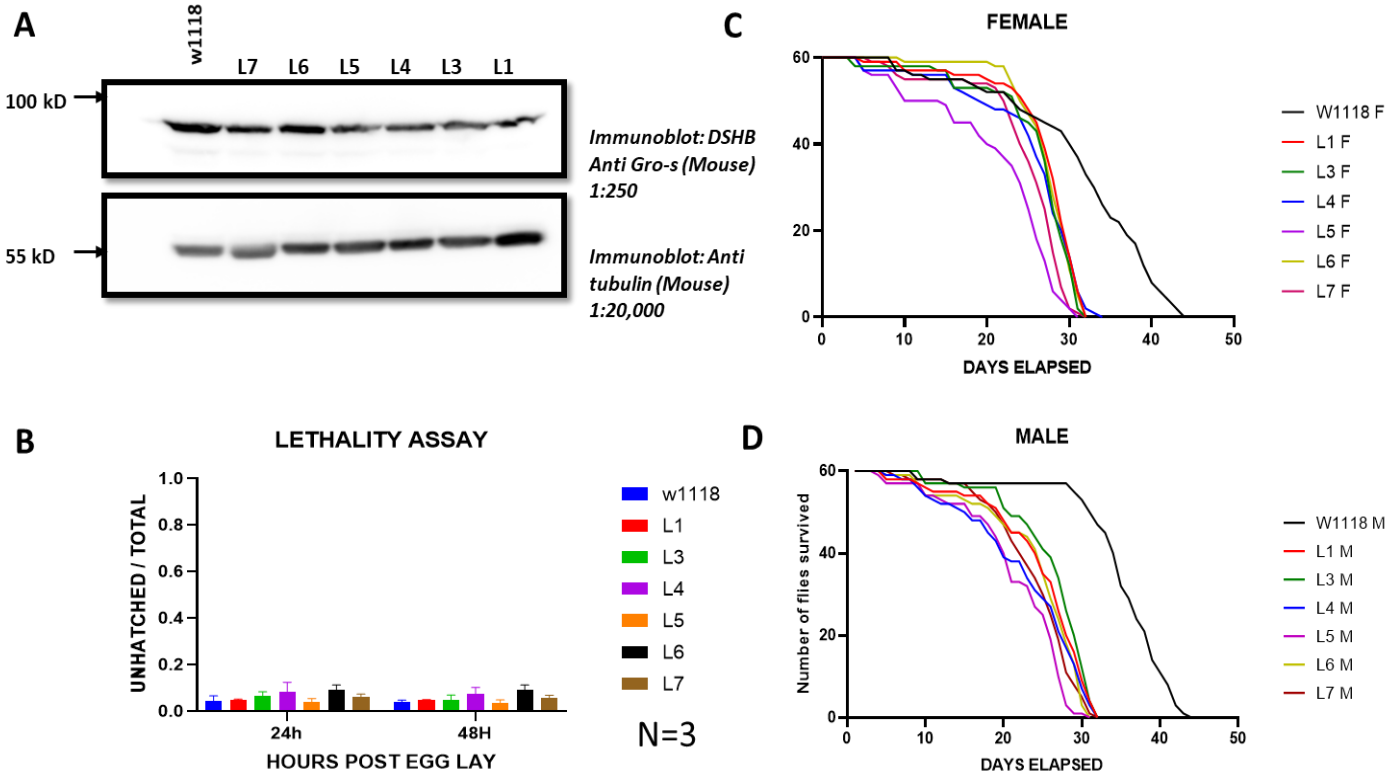


Figure 2.2 Phenotypic analysis of Gro^{2M} line.

(A) The two mutant L7 and the other CRISPR lines showed expression of Groucho at the correct molecular weight compared to wild type control (w1118). Alpha tubulin was used as loading control**. (B)** Lethality assay using the CRISPR lines showed no significant lethal defect in the Gro^{2M} line compared to CRISPR control and W^{1118}. **(C, D)** Life span assay for the Gro^{2M} line is similar to the CRISPR controls in both males and females. These CRISPR mutants survive 30-34 days, whereas wild-type w1118 flies survive for around 42-44 days. The assay was done at 29°C.

2.3.3 Gro^{2M} fails to interact with Dorsal

Dorsal and Groucho are both targets of SUMO, and SUMOylation of a target protein has been shown to modulate its interaction with its partner. In the context of Gro^{2M} L7, it will be interesting to see how the conjugation of SUMO to Groucho regulates its interaction with Dorsal and its effect on the downstream processes. $Dorsal^{K382R}$ has been shown to yield about 5-10 fold higher levels of reporter gene activity than the wild-type counterpart. This difference was not due to difference in protein expression since each protein was expressed with similar efficiency. Hence, conjugation of SUMO with Dorsal seems to apply a break to Dorsal activity and fine-tunes it depending on the context **(Bhaskar V. et al., 2002).** This break in Dorsal activity by SUMO could be done by interaction with Groucho directly or indirectly since Groucho is a well-known co-repressor which interacts with Dorsal. We have already generated $Dorsal^{K382R}$ and Gro^{2M}.

Native pull-down experiments with Gro^{2M} was done with *CS* flies as control and L1 as CRISPR control and visualized using western blot. Since Groucho expression is the highest in early embryos, we collected 0-5 hour embryos, crushed them in RIPA buffer (supplemented with protease inhibitor complex and 20mM N-ethyl maleimide), and loaded an equal amount of protein after protein estimation by BCA assay.

Groucho (input) was shown to be expressing in all three fly lines, with the CRISPR control showing a slightly lesser intensity of the Groucho band. Tubulin was used as a loading control ***(Figure 2.3A)*****.** Groucho was successfully pulled down using an anti-rabbit Groucho antibody***(Figure 2.3B)*****.** Interestingly, when the blot was probed with an antibody against Dorsal, it was seen that the Groucho2M did not pull down any Dorsal protein with it while the control CS and CRISPR control L1 does so ***(Figure 2.3C)*****.** Furthermore, when we tried to pull down Groucho using an antibody against *Drosophila* Sumo (Smt3), we saw a very faint band corresponding to the Groucho protein in the Gro^{2M} line ***(Figure 2.3D)***. Notably, the molecular weight at which the Groucho band was seen was precisely the same as the unmodified Groucho. We believe that this Groucho band corresponds to an unmodified groucho interacting with some other sumoylated protein. We were not successful in pulling down SUMOylated form of Groucho. However, it was shown that the mutant Groucho fails to interact with Dorsal.

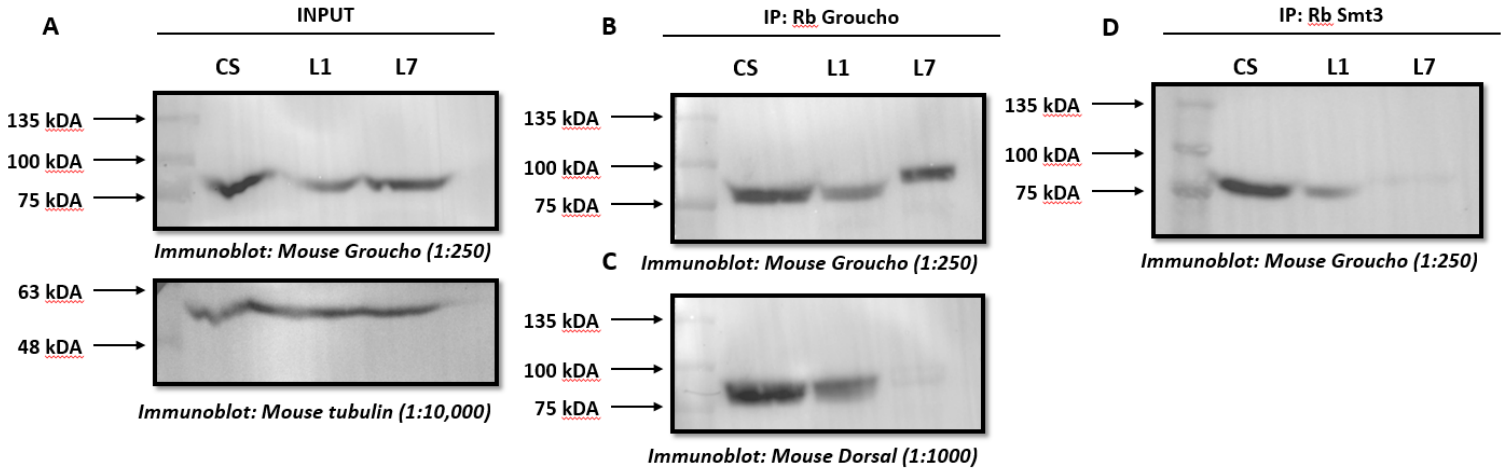


Figure 2.3 Dorsal-Groucho interaction is affected in Gro^{2M} line.

(A) Western blot showing bands for inputs of CS(control), L1(CRISPR control) and L7(mutant: Gro^{2M}). Tubulin has been used as loading control. **(B)** Native pull down experiments were carried out using antibody against Groucho and probed with antibodies against the same. The presence of three distinct Groucho bands show that pull down was successful. **(C)** Native pull down experiments were carried out using antibody against Groucho and probed with antibodies against Dorsal. Dorsal was seen to be interacting with CS as well as L1, whereas L7 (partial SCR mutant Groucho line) does not seem to interact with Dorsal at all. **(D)** Native pull down using antibody against drosophila SUMO (Smt3) and probed with antibodies against Groucho. A very faint band of Groucho in the L7 line was seen.

2.3.4 Understanding the role of SIM sites in Groucho and its importance in mediating its co-repressor activity

Groucho co-repressor was found to be covalently conjugated to SUMO-1 at multiple lysine residues in COS-7 cells. Deletion analysis on Groucho determined four distinct lysines within the GP and SP domains which are SUMOylated. These lysines are K167, K177, K272 and K279. SUMOylation of Groucho was shown to increase its co-repressor activity by luciferase reporter assay whereas Groucho SUMO conjugation resistant (SCR) mutant showed loss of co-repressor activity. This is probably done by efficient recruitment of HDAC1 by its SUMO Interacting Motif (SIM). It was shown that wild type HDAC1 binds more strongly to SUMO-Groucho fusion protein than wild type Groucho, whereas HDAC1 SIM mutant did not. Protein-protein interaction shows Dorsal and Groucho can bind to each other directly **(Dubnicoff T et al., 1997).** SUMO-SIM mediated protein-protein interaction is a well-known phenomenon in the biological system. Groucho contains multiple SIM sites in its key repressor binding

domains. Hence it would be interesting to check the role of these SIM sites in Groucho co-repressor activity.

SIM site prediction by JASSA shows that Groucho has three (VVDV 236-239, VEVL 630-633 and VLSL 650-653). 2 out of 3 SIM sites in Groucho lies within the WD domain (required for repressor binding). The binding of the Groucho co-repressor to Dorsal converts the otherwise active transcription factor into a repressor. We hypothesize that the conjugation of SUMO to Dorsal has a role to play in this. SUMOylated Dorsal recruits Groucho co-repressor in a context-dependent manner by SUMO-SIM interaction and thereby functions as a repressor. To test this hypothesis, we aimed to create the SIM mutants for Groucho (Gro Δ630-633 and Gro Δ650-653) and check the role of SIM sites in Dorsal Groucho interaction in wild type as well as SCR background.

We first amplified the Groucho cds region and introduced the required mutations in the SIM sites (VVDV to AVDA, VEVL to AEAL and VLSL to ALSA). The wild-type and mutant *gro* were tagged in their C-terminus with an HA-tag or HIS-tag in their N-terminus to aid in subsequent purification and pull-down experiments.

2.3.5 Mutation of Groucho SIM sites leads to enhanced degradation of the mutant protein

Both the wild-type *gro* and mutant *gro* were tagged in their C-terminus with an HA-tag to aid in subsequent purification and pull-down experiments. The construct was then introduced in the vector pET45b(+) through SLiCE-mediated ligation and cloning to generate pET45b(+)-gro^{wt}-HA and pET45b(+)-gro^{SIM}-HA. After verification through sequencing, the final construct was transformed into BL21DE3 to check for their successful expression. The successfully expressing clones were selected and co-transformed into BL21 cells (Novagen) with either the Q^{SUMO} or the control $Q^{\Delta GG}$ expression vectors **(Nie et al., 2009).** Several transformed bacterial colonies, selected for kanamycin and ampicillin resistance, were picked to inoculate 5 ml cultures of a non-inducing medium and shaken at 37°C overnight. This culture was then used to inoculate 10 ml secondary culture. When the bacterial culture reached an OD of 0.4 to 0.5, it was induced using 0.5 mM IPTG and grown overnight at 18°C (180 rpm) to induce the expression of all five proteins. 1 ml culture was taken and boiled in SDS dye and 20 µl was loaded in a 10% denaturing SDS gel and analyzed using western blot. The blots were then

probed with antibodies against the HA-tag and against the protein to confirm the experiment's credibility.

Groucho wild-type protein, when co-transformed with Q^{SUMO} showed four slow migrating bands (three clearly visible) corresponding to the sumoylated form of dorsal, whereas when transformed with $Q^{\Delta GG,}$ the slow migrating band disappears completely. SIM mutation does not seem to affect the SUMOylation status of the protein in any way **(*Figure 2.4A)*.** We further tried to confirm these results by transfecting the dorsal sim mutant in *Drosophila* 529Su cells.

529Su is a modified form of the *Drosophila* S2 cell line. It is stably transfected with expression vectors for FLAG-tagged Smt3 and HA-tagged Ubc9 (lwr), both under the control of the Cu2+ inducible metallothionein promoter. Both the wild-type groucho and mutant groucho cds were tagged in their N-terminus with an HIS-tag to aid in subsequent purification and pull-down experiments. The construct was then introduced in the vector pRM-HIS through SLiCE-mediated ligation and cloning to generate pRM-HIS-gro^{wt} and pRM-HIS-gro^{SIM}. After verification through sequencing, the final construct was transfected into 529Su to check for their successful expression. The cells were seeded in a six well tissue culture plate prior to start of transfection. Mirus TransIT-Insect was used as a transfection reagent. Induction was done using 0.5mM CuSO4 post-transfection, and cells were harvested 24h post-transfection. Collected cells were washed once in 1x PBS, boiled in SDS dye for 10 mins, loaded in a 10% SDS polyacrylamide gel, and analyzed using western blot.

$Groucho^{WT}$ protein when transfected into 529Su cells, did not show a clear slow migrating band corresponding to the sumoylated form of Groucho. Giving a heat shock (37°C for 30 mins) to the transiently transfected 529Su cells did not show any significant alteration to the SUMOylation status of Groucho **(*Figure 2.4B, right panel)*.** However, we see a clear reduction in the intensity of $Groucho^{SIM}$ upon heat shock. As_2O_3 is a known chemical agent used to increase the global SUMOylation within a cell. When As_2O_3 was used alone or in combination with heat shock, the decrease in $Groucho^{SIM}$ band intensity was severe (almost absent).

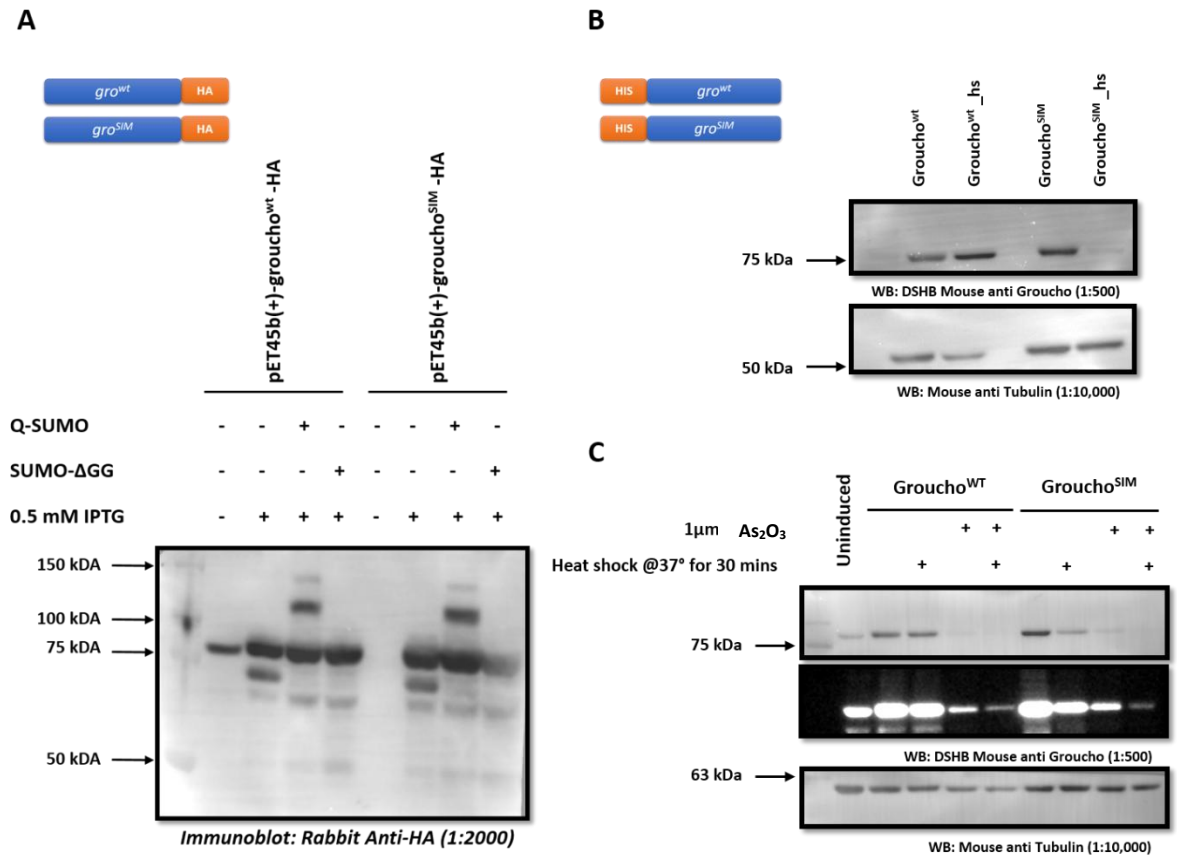


Figure 2.4 Groucho SIM sites are important for protein stability under stress conditions

(A) In-bacto SUMOylation assay of C-terminus HA-tagged Groucho wild type and Groucho SIM mutant. Groucho full-length cds with or without mutations were cloned in pET45b(+), co-transformed in BL21DE3 cells (with either the Q^{SUMO} or the control $Q^{\Delta GG}$ expression vectors following protocol in Nie et al., 2009) and induced with 0.5mM IPTG overnight with shaking. Lysates were loaded onto 10% SDS gel and analyzed using a western blot. Groucho wild type protein, when co-transformed with Q^{SUMO} showed four slow migrating band (three clearly visible) corresponding to the sumoylated form of Groucho, whereas when transformed with $Q^{\Delta GG}$ the slow migrating band disappears completely. **(B)** N-terminus HIS-tagged full-length Groucho cds with or without mutations were cloned in a pRM vector, induced with 0.5mM CuSO4, and harvested after 24h of induction. Cells were given heat shock for 30 mins at 37°C and lysates were loaded on a 10% SDS gel and analyzed using western blot. Slow migrating sumoylated forms of proteins bands were very faint and hence had to be overexposed to visulaize them clearly. HIS-tagged Caspar was used as a positive control of successful transfection. (C) 529Su cells were transiently transfected with wild type and mutant *groucho* and given heat shock or treated with 1 µM of As2O3 to upregulate global SUMOylation. Cells were lysed in RIPA supplemented with 0.04M N-ethyl maleimide. Equal protein was loaded on a 10% SDS gel and analysed using a western blot. Tubulin was used as a loading control.

2.5 Materials and methods

2.5.1 Fly husbandry and stocks

Flies were raised on standard cornmeal agar at 25°C unless stated otherwise. Groucho CRISPR lines were generated in collaboration with NCBS transgenic facility.

2.5.2 Generation of transgenic CRISPR lines

The gRNA was designed using the Target Finder tool (http://targetfinder.flycrispr.neuro.brown.edu/). A gRNA sequence was selected based on its lack of off-target sites and suitability for our experimental design. This sequence was then cloned into the pBFv-U6.2B vector. To generate homology regions, the genomic locus of *gro* was amplified from a single w^{1118} fly. The 5′ and 3′ homology regions (HR) of *gro* were prepared separately, with the K167R and K177R mutations introduced into the 5′HR, and the K272R and K279R mutations incorporated into the 3′HR. The pHD-scarless-DsRed vector was amplified in two fragments using specific primers. These fragments, along with the 5′HR and 3′HR of *gro*, were assembled using Gibson assembly (NEB) and verified by sequencing.

To screen for the gro^{2M} mutant after excising the DsRed cassette, genomic DNA was extracted by homogenizing single flies in 50 µL of DNA extraction buffer containing 10 mM Tris, 1 mM EDTA, and 20 µM Proteinase K. The homogenate was incubated at 37°C for 30 minutes, followed by heat inactivation at 85°C for 5 minutes. The *gro* genomic region was amplified and sequenced for each line. Among the screened lines, only line 7 (L7) carried two of the four desired mutations (K272R and K279R) ***(Supplementary Figure S1)***. Fly lines were routinely sequenced to confirm the genotype. A list of all primers used is provided in Figure (***Supplementary Table S2).***

2.5.3 Embryonic lethality

0-3 hr embryos were collected, transferred to a fresh sugar-agar plate, and unhatched embryos were scored after 48 hours to determine viability.

2.5.4 Lifespan and survival analysis

Survival assays were performed on the *Gro*2M CRISPR line and its derivative genotypes. Approximately ~60 male and female flies of the desired genotype were collected for each experiment, with no more than 15 age-matched flies per vial. Flies were transferred to fresh vials every four days, and the number of surviving individuals in each vial was recorded daily

until all flies, including those in control groups, had died. The resulting survival data were analyzed and plotted.

2.5.5 Cloning and generation of constructs for overexpression

The *Groucho* (*gro*, FBgn0001139) gene was amplified from the *Drosophila* Gold Collection Library (https://www.fruitfly.org/EST/gold_collection.shtml) using specific primers (listed in ***Figure Supplementary Table S2***). The resulting amplicons were independently cloned into the pET45b+ vector for bacterial SUMOylation assays and the pRM-HA3 vector for transfection into 529Su cells, employing a modified Seamless Ligation Cloning Extract (SLiCE) protocol (Zhang Y et al., 2012). Site-directed mutagenesis was performed using specific primers to modify the SUMO-interacting motif (SIM) sites on Groucho, introducing the following substitutions: VVDV to AVDA, VEVL to AEAL, and VLSL to ALSA. All constructs were verified by sequencing and subsequently utilized in downstream experiments.

2.5.6 S2 cell culture, transfections, cycloheximide treatment, LPS induction and immunostaining

S2 cells stably transfected with Flag-SUMO and HA-Ubc9, referred to as 529SU cells, were generously provided by Prof. Albert Courey. The cells were cultured and maintained in Gibco Schneider's *Drosophila* Medium (Thermo Fisher Scientific, #21720024) supplemented with 10% heat-inactivated fetal bovine serum (FBS) (Thermo Fisher Scientific, #10082147) at 23°C. Transfections were performed using 1 µg of plasmid DNA per 1 mL of cells, following the manufacturer's protocol for the TransIT-Insect Transfection Reagent (Mirus, #6100). Induction of transfected cells was carried out using 0.5 M $CuSO_4$, and cells were harvested 24 or 48 hours post-induction, depending on the experimental requirements. For heat shock treatment, 529SU cells were harvested and subjected to a 30-minute heat shock at 37°C in a water bath. Protein extracts were subsequently analyzed via western blotting.

2.5.7 Bacterial SUMOylation assay

This is a modified in-vitro SUMOylation assay that was previously described (Nie M et al., 2009). The quartet vector comprising of the *Drosophila* SUMO machinery components was co-transformed with HA tagged *dl*. The bacterial culture was induced with 1mM of Isopropyl β-D-1-thiogalactopyranoside (IPTG) for 6 hours at 25°C. 10ml of the bacterial culture was harvested in 1ml 50mM *tris* aminomethane (TRIS) buffer containing 150mM NaCl, 1mM Dithiothreitol (DTT), 1ug/ml lysozyme, and 1mM phenylmethylsulfonyl fluoride (PMSF). The cells were lysed using a VibraCell probe sonicator with 2/3sec ON/OFF cycle for 2 min at 60% amplitude.

2.5.8 Western blot analysis

For in-bacto SUMOylation assay, several transformed bacterial colonies, selected for kanamycin and ampicillin resistance, were picked to inoculate 5 ml cultures of a non-inducing medium and shaken at 37°C overnight. This culture was then used to inoculate 10 ml secondary culture. When the bacterial culture reached an OD of 0.4 to 0.5, it was induced using 0.5 mM IPTG and grown overnight at 18°C (180 rpm) to induce the expression of all five proteins. 1 ml culture was taken, boiled in SDS dye, and 20 µl was loaded in a 10% denaturing SDS gel and analyzed using western blot.

For cell-culture experiments, the cells were seeded in a six well tissue culture plate prior to start of transfection. Mirus TransIT-Insect was used as a transfection reagent. Induction was done using 0.5mM CuSO4 post-transfection and cells were harvested 24h post-transfection. Collected cells were washed once in 1x PBS, boiled in SDS dye, and equal amount of protein was loaded in a 10% denaturing SDS gel and analyzed using western blot.

For flies, 0-3h embryos were collected and were lysed in RIPA buffer (50 mM Tris-Cl, 150 mM NaCl, 0.1% SDS, 0.01% Sodium azide, 0.5% sodium deoxycholate, 1 mM EDTA, 1% Triton X-100, 1X PIC) with a pellet pestle (Kontes). Lysates were cleared by centrifugation at 21,000g at 4 °C for 30 minutes. Protein concentration was estimated using a BCA assay (Pierce) and 80 µg of total protein was loaded onto the gel after boiling in 1X Laemmli Sample Buffer. Proteins separated by 10% SDS-PAGE were transferred onto a PVDF membrane (Immobilon-E, Merck) and blocked in 5% milk in Tris-Buffer Saline (TBS) with 0.1% Tween 20 (TBS-T) for an hour. Blots were then incubated overnight with primary antibody diluted in 5% milk in TBS-T at 4 °C. Following three washes with TBS-T, blots were incubated with

secondary antibodies diluted in 5% milk in TBS-T for 1 hour at room temperature. Blots were washed thrice with TBS-T and visualized on a LAS4000 Fuji imaging system after incubating with Immobilon Western Chemiluminescent HRP substrate (Merck). The following antibodies were used: Mouse anti-α-Tubulin, 1:10000 (T6074, Sigma), Rabbit anti-HA, 1:2000 (04-902 DW-2, Sigma-Aldrich), Mouse anti-Groucho (DSHB, 1:250), Mouse anti- Dorsal 7A4 (DSHB, 1:1000), Goat anti-rabbit HRP and Goat anti-mouse HRP secondary antibodies, each at 1:10000 (Jackson ImmunoResearch).

2.5.9 Immunoprecipitation

0–3-hour embryos were lysed in Co-IP Lysis Buffer (20mM Tris pH 8.0, 137mM NaCl, 1% IGEPAL, 2mM EDTA, 1X PIC) using a Dounce homogenizer, and centrifuged at 21,000 g for 30 minutes. 3 mg of total lysate was incubated with 5 µg of primary antibody (Rb anti-Gro and Rb anti-Smt3) and 5 µg of Normal Rabbit IgG overnight at 4 °C. Antigen-antibody complexes were captured using 50 µL of BioRad SureBeads Protein A (1614013) at 4 °C for 4 hours. Beads were washed six times with Co-IP Lysis Buffer and protein complexes eluted by boiling in 1X Laemmli Sample Buffer. Eluted proteins were resolved on a 10% polyacrylamide gel.

2.6 Acknowledgements

We thank Fly facility at the National Centre for Biological Sciences (NCBS), Bangalore, for generation of CRISPR mutants.

2.7 References

Lindsley, D.L. and Grell, E.H. (1968) Genetic Variation of Drosophila melanogaster. Carn. Inst. Wash. Publ., Washington, 627.

Ahn, J. W., Lee, Y. A., Ahn, J. H., & Choi, C. Y. (2009). Covalent conjugation of Groucho with SUMO-1 modulates its co-repressor activity. *Biochemical and Biophysical Research Communications*, *379*(1), 160–165. https://doi.org/10.1016/j.bbrc.2008.12.043

Bassett, A. R., & Liu, J. L. (2014). CRISPR/Cas9 and Genome Editing in Drosophila. *Journal of Genetics and Genomics*, *41*(1), 7–19. https://doi.org/10.1016/j.jgg.2013.12.004

Bier, E., & Nizet, V. (2021). Driving to Safety: CRISPR-Based Genetic Approaches to Reducing Antibiotic Resistance. *Trends in Genetics*, *37*(8), 745–757. https://doi.org/10.1016/j.tig.2021.02.007

Buscarlet, M., & Stifani, S. (2007). The "Marx" of Groucho on development and disease. *Trends in Cell Biology*, *17*(7), 353–361. https://doi.org/10.1016/j.tcb.2007.07.002

Dubnicoff, T., Valentine, S. A., Chen, G., Shi, T., Lengyel, J. A., Paroush, Z., & Courey, A. J. (1997). Conversion of Dorsal from an activator to a repressor by the global co-repressor Groucho. *Genes and Development*, *11*(22), 2952–2957. https://doi.org/10.1101/gad.11.22.2952

Gratz, S. J., Ukken, F. P., Rubinstein, C. D., Thiede, G., Donohue, L. K., Cummings, A. M., & Oconnor-Giles, K. M. (2014). Highly specific and efficient CRISPR/Cas9-catalyzed homology-directed repair in Drosophila. *Genetics*, *196*(4), 961–971. https://doi.org/10.1534/genetics.113.160713

Kim, Y. H., Choi, C. Y., & Kim, Y. (1999). Covalent modification of the homeodomain-interacting protein kinase 2 (HIPK2) by the ubiquitin-like protein SUMO-1. *Proceedings of the National Academy of Sciences of the United States of America*, *96*(22), 12350–12355. https://doi.org/10.1073/pnas.96.22.12350

Kim, Y. H., Sung, K. S., Lee, S. J., Kim, Y. O., Choi, C. Y., & Kim, Y. (2005). Desumoylation of homeodomain-interacting protein kinase 2 (HIPK2) through the cytoplasmic-nuclear shuttling of the SUMO-specific protease SENP1. *FEBS Letters*, *579*(27), 6272–6278. https://doi.org/10.1016/j.febslet.2005.10.010

Nie M, Xie Y, Loo JA, Courey AJ. Genetic and proteomic evidence for roles of Drosophila

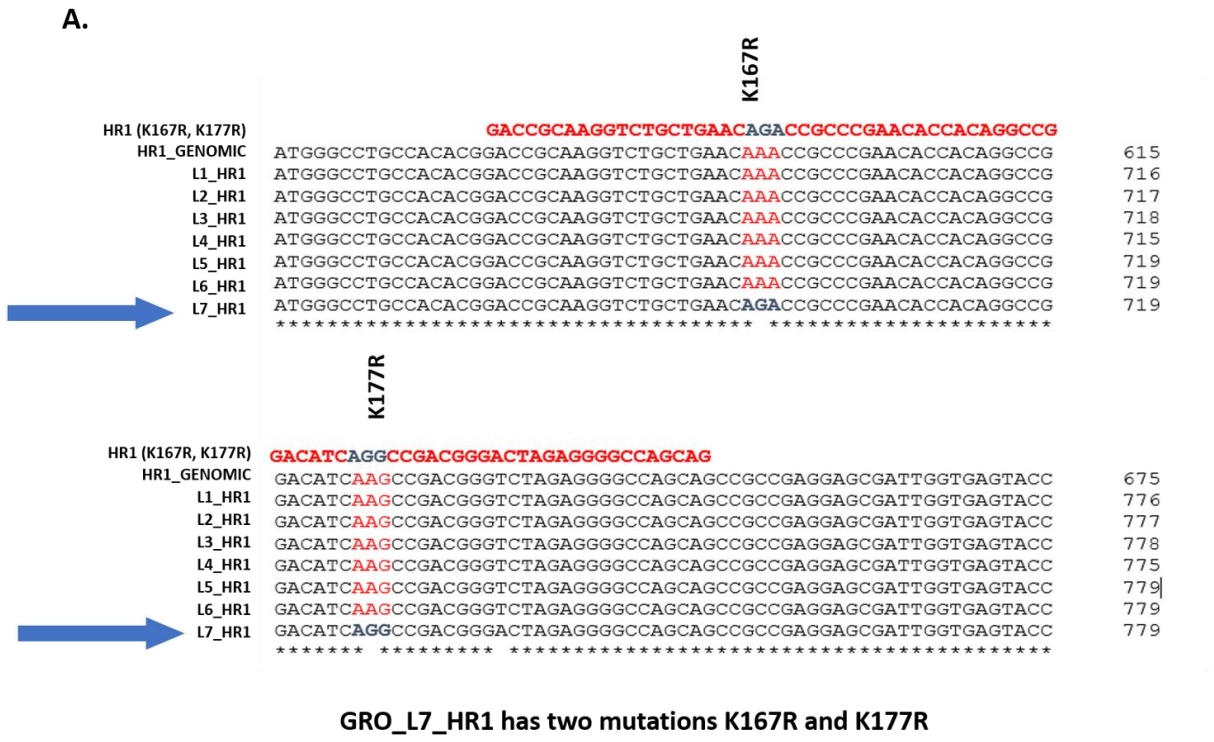
SUMO in cell cycle control, Ras signaling, and early pattern formation. PLoS One. 2009 Jun 16;4(6):e5905. doi: 10.1371/journal.pone.0005905. PMID: 19529778; PMCID: PMC2692000.

Nuthall, H. N., Joachim, K., & Stifani, S. (2004). Phosphorylation of Serine 239 of Groucho/TLE1 by Protein Kinase CK2 Is Important for Inhibition of Neuronal Differentiation. *Molecular and Cellular Biology*, *24*(19), 8395–8407. https://doi.org/10.1128/mcb.24.19.8395-8407.2004

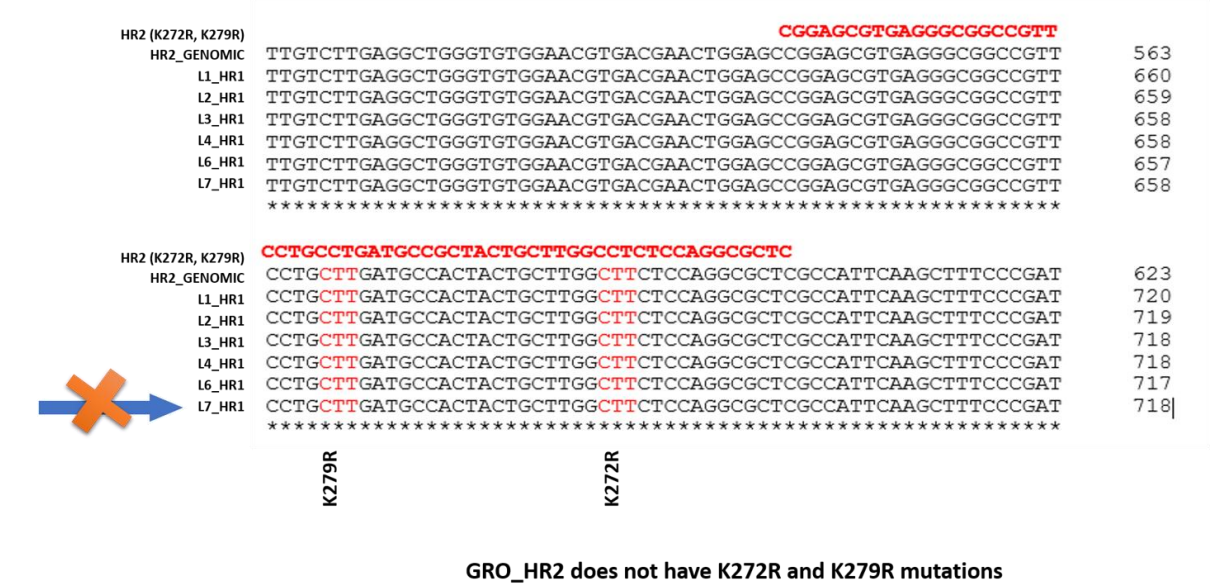
Yochum, G. S., & Ayer, D. E. (2001). Pf1, a Novel PHD Zinc Finger Protein That Links the TLE Co-repressor to the mSin3A-Histone Deacetylase Complex. *Molecular and Cellular Biology*, *21*(13), 4110–4118. https://doi.org/10.1128/mcb.21.13.4110-4118.2001

Zhang Y, Werling U, Edelmann W. SLiCE: a novel bacterial cell extract-based DNA cloning method. Nucleic Acids Res. 2012 Apr;40(8):e55. doi: 10.1093/nar/gkr1288. Epub 2012 Jan 12. PMID: 22241772; PMCID: PMC3333860.

2.8 Supplementary Figures



GRO_L7_HR1 has two mutations K167R and K177R



GRO_HR2 does not have K272R and K279R mutations

Figure S1: Multiple Sequence Alignment of *Drosophila gro^{wt} and gro^{SCR}* lines

(A) The genomic regions of homology region 1 (HR1) of *gro^{SCR}* lines (L1, L2, L3, L4, L5, L6, and L7) generated using CRISPR-Cas9 were aligned against their wild type genomic regions. L7 has two mutations as confirmed by sequencing (K167R and K177R). The mutated sequence is shown in red on the top denoted as HR1 (K167R and K177R). Mutated sequence is denoted in Blue.

(B) The genomic regions of homology region 2 (HR2) of *gro^{SCR}* lines (L1, L2, L3, L4, L6, and L7) generated using CRISPR-Cas9 were aligned against their wild type genomic regions. None of the lines have the two mutations as confirmed by sequencing (K272R and K279R). The mutated sequence is shown in red on the top denoted as HR2 (K272R and K279R). Mutated sequence is denoted in Blue.

Tabl2 S2: List of Groucho primers used during the course of experiments

PRIMER NAME	*SEQUENCE (5' - 3')*
sg2 white groucho-hr-F	*AGCGACACATACCGGCGCCCAGGGTGCAGCATCAGCAGCAACA*
seq groucho 5'F	*GCTTCAAATGCCGCCAATTG*
seq2 groucho 5'F	*GCACGATGTTCATTATTCCG*
groucho hr sdm K167R 3'R	*TGTGGTGTTCGGGCGGTCTGTTCAGCAGACCTTGCGGTC*
groucho hr sdm K177R 5'F	*GGCCGGACATCAGGCCGACGGGACTAGAGGGGCCAGCAG*
groucho hr sdm K272R-K279R 5'F	*GAGCGCCTGGAGAgGCCAAGCAGTAGCGGCATCAgGCAGGAAC GGCCGCCCTCACGCTCCG*
groucho hr sdm K272R-K279R 3'R	*GCCATTCAAGCTTTCCCGATC*
seq2 groucho 3'R	*TGTACCCGGCTTTTCCATCT*
seq groucho 3'R	*TGGAATGCCATTGGTTCGCA*
sg2 white groucho-hr-R	*AGCGACACATACCGGCGCCCAGGACTCCTATCTGTGCCGATTC*
groucho-hr1-F	*CCTATAGGAGACGTATATGGTCTTCTTTTCCCGAGGGTGCAGCA TCAGCAGCAACATGG*
gro hr1 FP (sd)	*CCCCAGATGGTCTTTGCTTATGC*
groucho-hr1-r	*CAATATGATTATCTTTCTAGGGTTAAACCAACTACTAGAAATTGAA AACCGG*
ts_gro hr1_RP	*ATACAGACCGATAAAACACATGCGTCA*
dsred hr1 RP (sd)	*ATCTCGAACTCGTGGCCGTTCA*
ds red hr2 FP (sd)	*TGCCCGGCTACTACTACGTGGA*
ts_gro hr2_FP	*CAGTGACACTTACCGCATTGACAAG*
groucho-hr2-F	*CGCAGACTATCTTTCTAGGGTTAAGGTGTATCTCAACTTATAGAA TTTTGTTTC*
gro hr2 RP	*GTGTACCCGGCTTTTCCATCT*

gro cds pet45b FP	*GAAGGAGATATACCATGGCAATGTATCCCTCACCGGTGCG*
pet45b+ RP	*TGCCATGGTATATCTCCTTCTTAAAGTTAAACAAAA*
gro cds sim 1 FP	*GCGATCAAGATTTAGCAGTAGATGCCGCAAATGAAATGGAATCC CACTCACC*
gro cds sim1 RP	*GGTGAGTGGGATTCCATTTCATTTGCGGCATCTACTGCTAAATCT TGATCGC*
gro cds sim seq RP	*GAGGGATGTGGAATGCCATTGG*
t7 eurofins FP	*TAATACGACTCACTATAGGG*
gen gro hr2 RP	*TTGTAGACTACCCTCGCCGTTCATAT*
gro cds sim2 FP	*GGAGAACTCGCATGCCGAGGCCCTGCACGCATCGAAAC*
gro cds sim2 RP	*GTTTCGATGCGTGCAGGGCCTCGGCATGCGAGTTCTCC*
gro cds sim3 FP	*CACGAGAGCTGCGCCCTGTCGGCCCGCTTTGCCGCC*
gro cds sim3 RP	*GGCGGCAAAGCGGGCCGACAGGGCGCAGCTCTCGTG*
ha pet45b+ FP	*TACCCATACGATGTTCCAGATTACGCTtaaCCGCACTCGAGTCTG GTAAAG*
gro-ha pet45b RP	*TCTGGAACATCGTATGGGTATCCTCCATAAATAACTTCGTAGACA GTAGCCTTCTTATCG*
t7 eurofins RP	*CTAGTTATTGCTCAGCGGT*
prm seq FP	*AGCATCTGGCCAATGTGC*
prm his RP	*TCCGTGGTGGTGGTGGTGGTGCATGGTACCGAGCTCGAATTCC C*
prm his gro FP	*ATGCACCACCACCACCACCACGGAGGATATCCCTCACCGGTGC GCCA*
prm FP	*GGATCCTCTAGAGTCGACCTGC*
prm gro RP	*GCAGGTCGACTCTAGAGGATCCTTAATAAATAACTTCGTAGACAG TAGCCTTCTTATC*

Chapter 3: Phosphorylation of the Dorsal/NFκB dimer plays a critical role in embryonic dorsoventral patterning

3.1 Abstract

The maternal factor Dorsal (DL) is the master regulator of embryonic dorsoventral (DV) patterning. In the early embryo, asymmetric Toll signalling leads to the phosphorylation of DL (Drier et al.,1999), which is important for its stability and nuclear import. Once in the nucleus, DL is SUMOylated and regulates the transcription of target genes (Hedge et al., 2022). SUMOylated proteins often interact with partner proteins at their SUMO-Interacting Motif (SIM) sites.

In this study, we have generated mutations (VVVV>AVVA & IILL>AILA) in *dl* that would disrupt both the predicted SIM site(s), SIMα ($DL^{98\text{-}VVVV\text{-}103}$) and SIMβ ($DL^{242\text{-}IILL\text{-}247}$). In S2 cells in culture, we find that this disruption does not affect DL SUMO conjugation, but intriguingly disrupts DL phosphorylation with enhanced degradation of the mutant protein. Embryos from homozygous dl^{β} mothers, generated by CRISPR/Cas9 genome editing, were completely dorsalized (D0), indicating that the dl^{β} is a null allele. Our data points to DL^{β} being unstable, lacking phosphorylation and prone to degradation. Further, molecular modelling of the DL dimer suggests that mutating the SIMβ site leads to the weakening of the DL:DL dimeric interface, which directly or indirectly affects its phosphorylation. The haploinsufficiency of DL^{β} at both 25°C and 29°C indicates that the DL: DL^{β} dimer is stable, and capable of efficient phosphorylation and nuclear import. A bioinformatics analysis involving structural modeling of Dorsal (DL) reveals that the SIMβ site is located within the dimerization interface of the protein, forming a hydrophobic pocket. Mutations at this site disrupt the integrity of the dimerization interface, leading to destabilization of the dimer. This structural instability promotes enhanced protein degradation, suggesting a critical role for the SIMβ site in maintaining Dorsal's structural and functional stability.

In summary, disruption of a predicted SIM2 site, which is an extended beta-sheet at the DL dimer interface, appears not to affect SUMOylation but rather the phosphorylation of DL. The consequence of this mutation is the inability of the DL^{β} variant to pattern the DV axis.

Keywords

Dorsal, Dorso-ventral Pattering, SUMO Interacting Motif, Protein degradation, Phosphorylation, Early embryogenesis, Maternal factor

3.2 Introduction

The dorsoventral (DV) axis of *Drosophila* embryos is established in the syncytial blastoderm by a nuclear concentration gradient of the Dorsal (DL) protein. DL is a member of the Rel family of NF-kappaB transcription factors **(Govind and Steward 1991) (Rushlow and Warrior 1992)**. Once translocated to the nucleus, it can bind to DNA to activate or repress genes in the DV gene regulatory network (Morisato and Anderson 1995). DL is an effector for the Toll signaling cascade, whose activation is set up in the maturing egg chamber before fertilization. Cactus, the *Drosophila* orthologue of mam malian IκB, plays an important regulatory role by binding and sequestering DL in the cytoplasm, blocking DL from performing its role as a master regulator of DV patterning. Rel proteins feature a conserved amino-terminal domain, the Rel homology region (RHR), encompassing approximately 300 amino acids. The C termini of Rel proteins exhibit significant divergence and are implicated in transcriptional activation. Investigations into vertebrate Rel proteins have established that crucial elements necessary for their regulated nuclear import are contained within the Rel homology region (RHR).

Furthermore, the RHR is responsible for mediating DNA binding, facilitating both homodimer and heterodimer formation among family members, and engaging in interactions with the ankyrin repeats of IkB proteins **(Ghosh et al. 1995), (Gilmore and Morin 1993) (Kidd 1992) (Müller et al. 1995)**. The hallmark of Rel/NF-KB proteins is their rapid translocation from the cytoplasm to the nucleus in response to extracellular signaling (Verma et al. 1995). In the absence of stimulation, these proteins remain inactive in the cytoplasm, sequestered by IkappaB **(Baeuerle and Baltimore 2018).** In a newly laid *Drosophila* egg, the product of the *dl* gene, DL protein, is uniformly distributed throughout the embryo's cytoplasm. Upon asymmetric Toll signaling, DL undergoes selective nuclear import in a graded manner; its concentration is maximal in the most ventral nuclei, diminishing progressively in the lateral nuclei, while it is excluded from the dorsally positioned nuclei **(Roth et al. 1989) (Rushlow et al. 1989) (Steward 1989).**

The vertebrate Rel family comprises several members, including the two NF-kB subunits p50 and p65/RelA, the p105 precursor of p50, p52 and its precursor p100, c-Rel, and RelB **(Ghosh et al. 1990)(Liou and Baltimore 1993)(Nolan et al. 1991) (Ruben et al. 2016)(Ryseck et al. 1992)(Sieben et al. 2024)**. These Rel proteins can form various interchangeable homodimers and heterodimers that play roles in the vertebrate immune system. In contrast, DL operates as a homodimer during *Drosophila* embryogenesis, unlike the NF-kB heterodimers. Genetic and

proteomic evidence shows that the nuclear translocation of DL protein is regulated at two levels, both requiring dissociation of the Dorsal-Cactus complex **(Govind et al. 1993)(Roth et al. 1991)**. The first basal level of nuclear translocation is achieved by free DL, which occurs in lateral cells and is independent of signaling. It is observed in embryos from mothers that are simultaneously mutant for both the signal and Cactus. The second level of regulation is signal-dependent and necessitates additional processes to reach the elevated concentrations of nuclear DL typically found in ventral cells.

The nuclear import and proper functioning of DL are also regulated by post-translational modifications in response to signaling. DL is an embryonic phosphoprotein. DL undergoes signal-dependent phosphorylation (limited to specific serine residues) while bound to Cactus. After being released from Cactus, DL is stabilized against immediate degradation in the cytoplasm by phosphorylation on its S312 residue. While DL^{S312} retains much of its function, it is less stable in a *cactusA2, dl^{null}* background, resulting in dorsalized embryos. This suggests that while S312 may not be critical for nuclear import, it plays a significant role in DL's stability. DL is translocated inside the nucleus upon activation of the Toll pathway. Drier et al. found that phosphorylation of S317 is necessary for its target to the nucleus. The study finds that DL^{S317A} significantly alters DL's phosphoform distribution and results in reduced nuclear localization in a *dl^{null}* background. Additionally, in a cactusA2, *dl^{null}* background, this mutation leads to lateralized rather than ventralized embryos, demonstrating a failure to establish the typical nuclear gradient. While the DL^{S317A} mutation does not entirely eliminate Dorsal's nuclear import, it prevents the protein from responding effectively to Toll signaling. The study proposes that hyperphosphorylation of S317 is crucial for high levels of nuclear targeting in ventral cells. Furthermore, the basal level of nuclear import in the S317A mutant also requires some form of phosphorylation **(Drier et al. 1999).**

DL has also been shown to be a target for another post-translational modification, SUMOylation (K382) **(Bhaskar et al. 2000)(Bhaskar et al. 2002)**. SUMO (Small Ubiquitin Like Modifier) is a reversible post-translational modifier that shows great versatility in modulating the functions of its target proteins. SUMO proteins are small (~10 kDa) and share some structural resemblance with the ubiquitin protein family. Despite this similarity, there are notable differences between SUMO and ubiquitin, particularly in sequence homology (approximately 20% similarity) and surface charge distribution, which set them apart in their biological roles and interactions **(Bayer et al. 1998) (Mossessova and Lima 2000) (Bernier-villamor et al. 2002).** Like ubiquitin, SUMO attaches to its target proteins via an isopeptide

bond, formed between the C-terminal glycine residue of SUMO and the ε-amino group of a lysine residue on the substrate protein (most often part of a Ψ-K-X-E/D consensus motif) **(Johnson 2004)(Hay 2005)(Varejão et al. 2020)**. Some SUMO substrates possess a SUMO-Interacting Motif (SIM), which enhances their SUMOylation. The best-characterized class of SIMs features a core of hydrophobic residues, typically following the pattern [V/I]-x-[V/I]-[V/I], flanked by acidic, negatively charged amino acids. The hydrophobic core of the SIM binds to a complementary hydrophobic surface on SUMO (when complexed with UBC9), while the surrounding acidic residues contribute to binding affinity and orientation. This interaction effectively increases the local concentration of the substrate, thereby facilitating the transfer of SUMO (Wang et al. 2009).

During early development, where alterations in the Dorsal (DL) gradient can disrupt DV patterning, it was found that SUMOylation of the DL protein has no significant effect on DL nuclear import and, therefore, the DL gradient. Surprisingly, it was demonstrated that the SUMO Conjugation Resistant form of DL (DLK^{382R} or DL^{SCR}) is a better transcriptional enhancer in the embryo under conditions of haploinsufficiency, where, counterintuitively, DL target genes are activated at levels comparable to wild-type **(Hegde et al. 2022)**.

DL has two conserved SIM sites (predicted) in its Rel homology domain. In this study, we have generated mutations (VVVV>AVVA & IILL>AILA) in DL that would disrupt both the predicted SIM site(s), SIM1 ($DL^{98\text{-}VVVV\text{-}103}$) and SIM2 ($^{DL242\text{-}IILL\text{-}247}$). In S2 cells in culture, this disruption does not affect DL SUMO conjugation but intriguingly disrupts DL phosphorylation. Embryos from homozygous $dl^{242\text{-}AILA\text{-}247}$ mothers, generated by CRISPR/Cas9 genome editing, were completely dorsalized (D0), indicating that the dl^{AILA} behaves as a null allele. Our data points to DL^{AILA} as unstable, lacking phosphorylation, and prone to degradation. Further, molecular modeling of the DL dimer suggests that mutating the SIM2 site weakens the DL:DL dimeric interface, which directly or indirectly affects its phosphorylation. The haploinsufficiency of DL^{AILA} at both 25°C and 29°C indicates that the DL: DL^{AILA} dimer is stable and capable of efficient phosphorylation and nuclear import. In summary, disruption of a predicted SIM2 site, which is an extended beta-sheet at the DL dimer interface, appears not to affect SUMOylation but rather the phosphorylation of DL. The consequence of this mutation is the inability of the DL^{AILA} variant to pattern the DV axis.

3.3 Results

3.3.1 Mutation of SIM sites in DL modulates its SUMOylation status

Building on our initial studies on DL SUMOylation **(Hegde et al. 2022)**, we aimed to investigate the functional role of SUMO-interacting motif (SIM) sites within the DL protein. SIM sites are known to modulate both SUMOylation status and protein-protein interactions, potentially impacting protein function. We used the Joined Advanced SUMOylation Site and Sim Analyser (JASSA) **(Varejão et al. 2020)** to predict two putative well-conserved SIM sites in DL, which we refer to as SIMα (amino acid residues 99-102, VVVV) and SIMβ (amino acid residues 243-246, IILL) **(Figure 3.1A, C).** Both SIMα and SIMβ lie within the RHD **(Figure 1B).** In order to investigate the biological roles of these SIM sites in DL, we generated point mutations in both sites. We examined their effect on the SUMOylation of DL using the *in-bacto* SUMOylation assay **(Nie et al. 2009) (Figure S1 A).** Specifically, we PCR-amplified the DL coding sequence and introduced mutations (VVVV to AVVA for SIMα and IILL to AILA for SIMβ). In previous studies, similar amino-acid substitutions led to the disruption of the SIM sites **(Lin et al. 2006; Pidugu et al. 2024)**. The wild type and mutant *dl* were tagged in their C-terminus with a DNA sequence corresponding to an HA-tag to aid in subsequent purification and pull-down experiments.

The constructs were then cloned into the *pET45b(+)* vector through SLiCE-mediated ligation to generate *pET45b(+)-dl^{wt}-HA* and *pET45b(+)-dl$^{\alpha\beta}$-HA*. After verification through sequencing, the final construct was transformed into BL21DE3 to check for their successful expression. Clones expressing the protein were then co-transformed with the control Q^{SUMO} or $Q^{\Delta GG}$ expression vectors (Nie et al., 2009). Q^{SUMO} vector expresses mature 6XHis-tagged SUMO (SUMO-GG), while $Q^{\Delta GG}$ expressed n immature/dead 6XHis-tagged

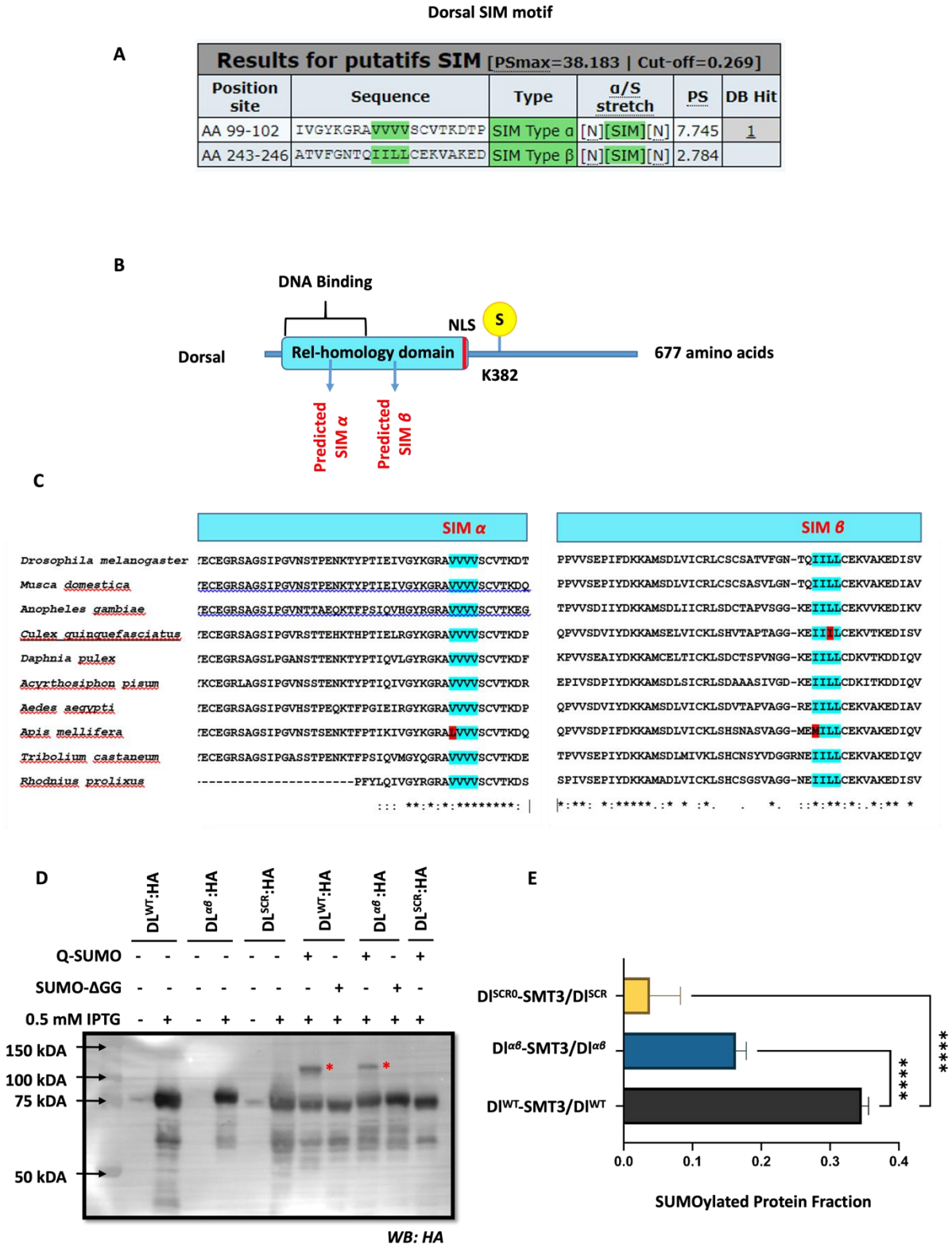


Fig. 3.1 In-bacto SUMOylation assay of Dorsal

(A) Table showing SIM site prediction for Dorsal by JASSA showing two probable sites VVVV (99-102) and IILL (243-246) (B) Schematic diagram showing location of SIM and SUMO conjugation sites

in Dorsal protein. (C) Homology sequences showing conservation of SIM sites among Dipterans (D) Bacterial sumoylation assay was used to validate the effect of SIM sites on the SUMOylation status of Dorsal. HA-tagged Dorsal was expressed in BL21 cells co-transformed with Q^{SUMO} or $Q^{\Delta GG}$ vectors, and immunoblotted using antibodies against HA and Dorsal. Red asterisk point to the bands representing sumoylated proteins. (E) Western blots were quantified to determine percentage of the total protein SUMOylated using ImageJ and plotted as bar graph. N=4, One way ANOVA, (***) P <0.001, (*) P <0.05

SUMO (SUMO-ΔGG). Protein expression was confirmed by SDS-PAGE and western blot analysis, with *pET-dl^{SCR}-HA*, a SUMO conjugation-resistant (SCR) form of DL, used as a negative control. The blots were then probed with antibodies against the HA tag and against the protein to re-confirm the experimental results **(Figure 3.1D, Supplementary S1 B)**. DL expression was found to be similar across all constructs. A slow-migrating band was observed in the samples co-expressing Q^{SUMO}, indicative of a SUMOylated DL fraction (shown with red asterisk). A single SUMO band confirms previous findings that DL has only one SUMO conjugation site. Interestingly, there was a decrease in the intensity of the SUMOylated fraction of $DL^{\alpha\beta}$, but the SUMOylation status is maintained. To determine the effect of SIM sites on cis-SUMOylation of the protein, we calculated the fraction of the total protein that is SUMOylated (SUMOyalted DL/(SUMOylated DL+DL). Quantifying SUMOylated vs. total protein, we found that SUMOylation in $DL^{\alpha\beta}$ mutants was reduced by approximately 50% compared to the wild-type protein, while Dl^{SCR} exhibited minimal or no SUMOylation **(Figure 3.1E, Supplementary S1C).** DL co-expressed with immature $Q^{\Delta GG}$ did not show a SUMO-conjugated species, validating the fact that SUMO requires the di-Gly residues to form an isopeptide bond with its substrate. Our findings indicate that the conserved SIMα and SIMβ sites contribute to efficient SUMOylation of DL but are not essential. The mutations substantially reduce SUMOylated protein levels, potentially impacting DL's function. The experimental examination of SUMOylation in-bacto is straightforward and on par with in-vitro experiments. These results, however, may not reflect conditions inside a cell. In the following section, we tested SUMO conjugation of DL SIM mutants in *Drosophila* cell lines.

3.3.2 $DL^{\alpha\beta}$ shows loss of phospho-isoforms and enhanced degradation upon heat shock

To investigate the role of the SUMO-interacting motif (SIM) in the SUMOylation of DL protein, we expressed DL in *Drosophila* 529Su cell lines, a derivative of S2 cells that are stably transfected with vectors expressing FLAG-tagged Smt3 and HA-tagged Ubc9 (*lwr*) under the control of a Cu^{2+}-inducible metallothionein promoter **(Figure 3.2A).** The wild-type *dl* and

mutant *dl* cds were tagged in their N-terminus with a 6X-HIS-tag to aid in subsequent purification and pull-down experiments. The construct was then introduced in the vector pRM-HIS through SLiCE-mediated ligation and cloning to generate pRM-HIS-dl^{wt} , pRM-HIS-dl^{α} , pRM-HIS-dl^{β} and pRM-HIS-$dl^{\alpha\beta}$. After verification through sequencing **(Figue S2),** the final construct was transfected into 529Su cells to check for their successful expression. pRM-his-dl^{SCR}, a SUMO conjugation-resistant form of DL, was used as a negative control for the SUMOylation assay.

Wild-type DL migrates with multiple slow migrating bands corresponding to its phospho-isoforms, whereas DL^{β} and $DL^{\alpha\beta}$ show reduced expression and absence of one or more of its phosphorylated isoforms. Interestingly, DL^{α} expression mirrored that of the wild type, retaining all phospho-isoforms. Surprisingly, DL^{SCR} showed expression similar to DL^{wt} with a much higher proportion of phospho-isoforms **(Figure 3.2B).** In Fig 2B red asterisk denote unphosphorylated native DL protein whereas yellow asterisk denote its phospho-isoforms. Previous studies have shown that DL needs to be phosphorylated to enter the nucleus, and a hyperphosphorylated form of DL is considered to be the active form. We have previously shown that DL^{SCR} is a better transcriptional activator (hegde et al). A higher abundance of phospho-isoforms in DL^{SCR} hints to a mechanism that in absence of SUMOylation, DL is phosphorylated and acts as a better transcriptional activator.

We next checked for the role of SIM in the cis-SUMOylation of DL. Heat shock is known to induce SUMOylation in general and hence is used to visualize the SUMOylated form of DL. N-ethylmaleimide (a SUMO deconjugase inhibitor) was used while preparing the sample to prevent the loss of the SUMOylated fraction. When transfected into 529Su cells, DL wild-type protein shows a slow migrating band corresponding to the SUMOylated form of DL. Notably, this band was absent in cells transfected with dl^{SCR}, validating the identity of the SUMOylated band. $DL^{\alpha\beta}$ shows significantly less or absence of its phospho-isoforms and an overall decrease in total protein. The total SUMOylated fraction of $DL^{\alpha\beta}$ is much lesser than the wild type, which may result from less total mutant protein **(Figure 3.2C).** Heat shock did not enhance DL SUMOylation; however, it resulted in the degradation of $Dl^{\alpha\beta}$ (as observed after quantification), indicating that this mutant has a lower half-life and heightened susceptibility to degradation under stress conditions **(Figure 3.2D).** These findings suggest that $Dl^{\alpha\beta}$ is both phosphorylation-deficient and less stable than wild-type Dorsal, supporting the hypothesis that this mutant form is more readily targeted for degradation upon heat shock.

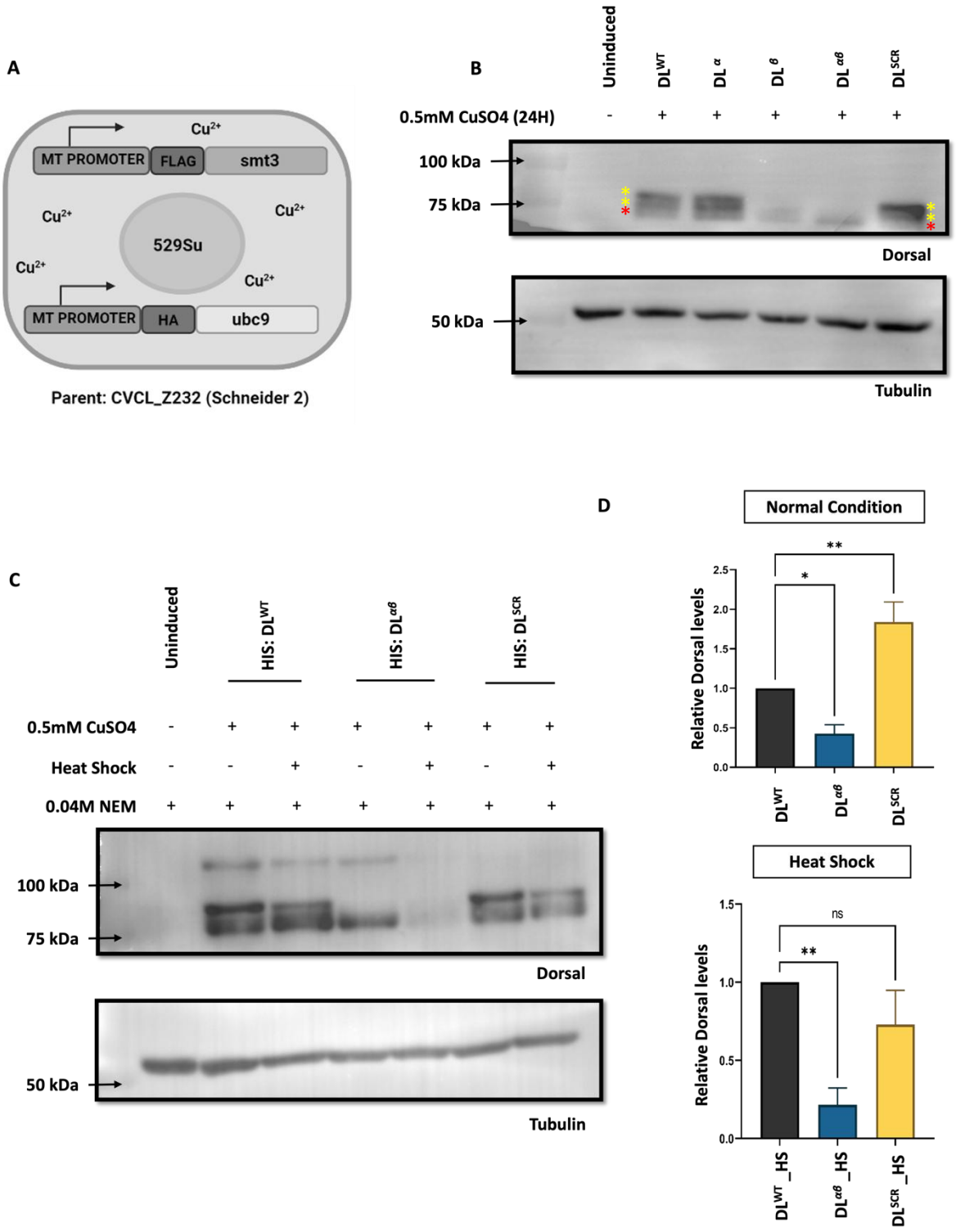


Fig. 3.2 Dl$^{\alpha\beta}$ shows reduced phosphoisoforms and enhanced degradation upon heat shock

(A) A Cu^{2+} induced S2 SUMOylation assay. The cds of FLAG-tagged *smt3* (*Drosophila* SUMO) and HA- tagged *ubc9* (*Drosophila* SUMO conjugase) were cloned in pRM vector under the metallothionein promoter and stably transfected inside *Drosophila* Schneider cells (S2). The stably transfected cell lines were called 529Su and were used for SUMOylation assay of Dorsal. (B) Site directed mutagenesis was carried out to make mutants for *dorsal* and cloned inside pRM vector. Expression levels of the mutants were checked by transfecting HIS-tagged dl^{wt} , dl^{α} , dl^{β}, $dl^{\alpha\beta}$ and dl^{SCR} in 529Su cells. After induction,

cells were lysed and run on SDS PAGE and probed with mouse anti-Dorsal antibody (1:1000). Tubulin was used as a loading control. (C) Effect of SIM sites on SUMOylation of Dorsal was carried out by transfecting 529Su cells with the wild type or mutant form of dorsal. After induction with 0.5 mM CuSO4 for the indicated amount of time, cells were lysed and the supernatant was run on an SDS PAGE and immunoblotted with antibodies against Dorsal. Red asterisk point to the bands representing sumoylated proteins. Tubulin was used as a loading control. (D) Western blots(C) were quantified to determine the total protein using ImageJ and plotted as bar graph. N=3-5, mean ±SD One way ANOVA, (**) P <0.01, (*) P <0.05

3.3.3 Mutation of predicted SIMα and SIMβ disrupts DL's ability to be phosphorylated

DL function is regulated through post-translational modifications, with phosphorylation and SUMOylation being critical for its stability and activity **(Gillespie and Wasserman 1994).** There are multiple phosphorylation sites on DL, of which S312 and S317 have been highlighted in literature with significant biological roles. S312 phosphorylation has been hypothesized to stabilize DL in the cytoplasm **(Drier et al. 1999)** after it is released from Cactus, in response to Toll signaling. Phosphorylation of S317, on the other hand, has been shown to be crucial for DL nuclear import. The classical dl^1 allele is a phosphomutant (S317N), which behaves as a genetic null **(Hegde et al. 2022).** In recent papers **(Mcgehee and Stathopoulos 2024)**, phosphorylation near DL's SUMOylation site (K382) is thought to affect its interactions with chromatin and other co-regulators, highlighting a complex interplay of modifications that fine-tune its function. This predicted phosphorylation site (S389) was identified through a phosphoproteomic screen **(Pasteur et al. 2019)** and is part of a bipartite PDSM (phosphorylation-dependent SUMOylation motif) consisting of a SUMO consensus site adjacent to a proline-directed phosphorylation site (ΨKxExxSP). Both the S312 and S317 phosphorylation sites lie on the Rel homology domain, whereas S389 lies very close to DL SUMOylation site K382 **(Figure 3.3A).** Investigating DL phosphorylation is of interest because it provides insights into how multiple post-translational modifications can integrate environmental signals to control precise developmental outcomes.

Our goal was to check if the phosphomutant forms of DL behave in a manner similar to the $DL^{\alpha\beta}$ Both the wild-type and mutant dorsal cDNA sequences were tagged in their N-terminus with a 6X-HIS-tag to aid in subsequent purification and pull-down experiments. The construct was then introduced in the vector pRM-HA through SLiCE-mediated ligation and cloning to generate *pRM-6XHis-*dl^{wt}, *pRM-6XHis-*dl^{S312A}, *pRM-6XHis-*dl^{S317A}, *pRM-6XHis -*$dl^{S312A,\ S317A}$, and *pRM-6XHis -*dl^{S389A}. After verification through sequencing **(Figure S3)**, the final

constructs were transfected into 529SU cells and evaluated for their protein expression. We find that DL^{S312A}, DL^{S317A} and $DL^{S312A, S317A}$ showed significantly lower expression levels and the absence of phospho-isoforms similar to $DL^{\alpha\beta}$ **(Figure 3.3B).** The non-phosphorylated form of DL was also seen with lesser intensity, highlighting the importance of phosphorylation, in general, for the stability of the DL protein. Concomitant to an overall decrease in DL levels, there was a decrease in the SUMOylated fraction of DL (DL^{SUMO}). The expression pattern of DL^{S389A} is similar to its wild-type counterpart, with a visible slow migrating high molecular weight band corresponding to its SUMOylated fraction **(Figure 3.3C).** This would suggest that the phosphorylation state of S389 is not a requirement for the SUMOylation of DL. Our data further suggests that SIM mutation-induced degradation may stem from an inability to undergo phosphorylation, leading to instability and degradation. Alternatively, the mutation might alter protein stability as a primary cause, leading to degradation. The lack of phosphorylated bands is because a degraded protein cannot be phospho-conjugated.

To test the two hypotheses, we decided to stabilize the fraction of DL phosphorylated states by blocking phosphatase activity in cellular lysates. When $Dl^{\alpha\beta}$ cellular lysates were treated with sodium vanadate, the intensity of phospho-isoforms of DL^{wt} was significantly higher. However, we did not see any phosphorylated bands for $DL^{\alpha\beta}$ **(Figure 3.3D)** on western blots, which would indicate remnant phosphorylation state(s). From our data, we conclude that it is more likely that DL SIM mutants ($DL^{\alpha,}$ $DL^{\beta,}$ $DL^{\alpha\beta}$) are not amenable to phosphorylation by a kinase, and this, in turn, leads to DL instability and accelerated degradation, rather than the phospho-isoforms of DL getting degraded owing to the mutation in the predicted SIM sites.

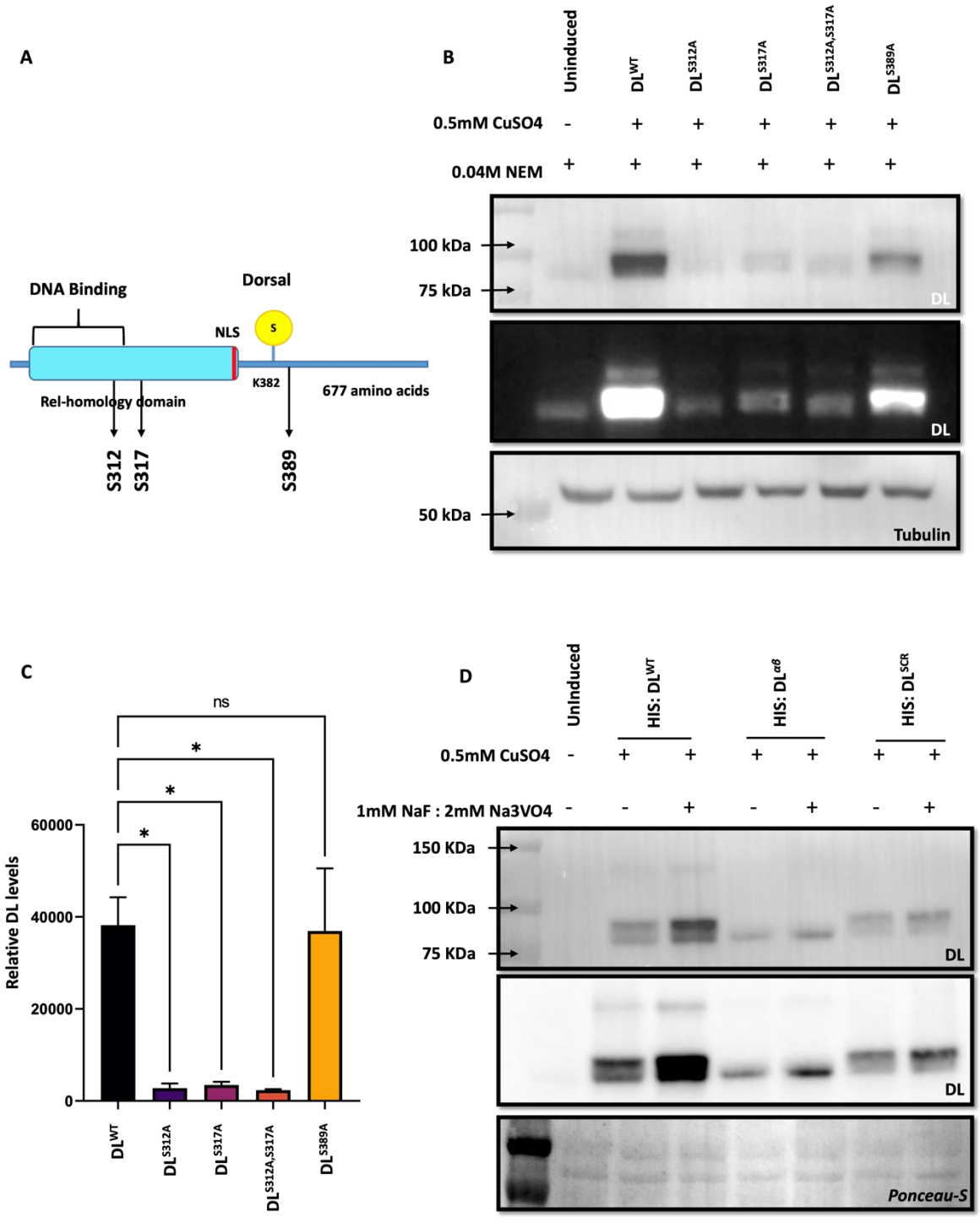


Fig 3.3 Dl$^{\alpha\beta}$ shows reduced phosphoisoforms similar to Dl phosphomutants and is not rescued by phosphatase inhibitors

(A) Schematic showing known phosphorylation sites on Dorsal (S312 and S317) and a PDSM site identified by JASSA (S389) in the Dorsal protein. (B) Dl phosphomutants were analyzed for their presence or absence of phosphoisoforms by running them on SDS PAGE and probed with mouse anti-Dorsal antibody (1:1000). Tubulin was used as a loading control. The middle panel is a more contrasted version of the top to show clearly the SUMOylated fraction. (C) Western blots in (B) were quantified

to determine the total protein using ImageJ and plotted as bar graph. N=3, mean ±SEM One way ANOVA, (*) P <0.1 (D) 529Su cells expressing dl^{wt} , $dl^{\alpha\beta}$ and dl^{SCR} were treated with phosphatase inhibitors (1mM NaF : 2mM Na3VO4) and run on an SDS PAGE to check the expression pattern of Dorsal. Ponceau-s was used as a loading control. The middle panel is a more contrasted version of the top to show clearly the SUMOylated fraction. N=3

To figure out how DL^{β} is degraded we explored two possible mechanisms, the 26S proteasomal machinery and autophagy. Treatment of 529Su cells with 25 μM MG132 for 8h (to block proteasomal degradation pathway) or 2mM 3-methyladenosine for 24h (to block autophagy) did not show any significant stabilization of the mutant protein. Surprisingly, addition of MG132 or 3-MA leads to enhanced degradation of the protein. Applying heat shock (37°C) for 30 mins was shown to enhance degradation of DL^{β} and $DL^{\alpha\beta}$, however DL^{wt} was found to be stable. Here, we see degradation of the wild type protein when heat shock is given in combination with MG132 or 3-MA. DL^{SCR} which was used as an internal control also showed similar degradation pattern like wild tuype protein. Our data indicates an unknown mechanism different from 26S proteasomal degradation or autophagy that is responsible for DL degradation **(Supplementary Figure S4)**.

Mammalian VCP/p97 is an essential chaperone for proteostasis that modulates several ubiquitin- associated processes. We decided to knockdown the *Drosophila* counterpart Ter94 using ds-RNA against it; thereby disrupted the 26S proteasomal pathway and then checked DL levels in 529Su cells. Ter94 knockdown using ds-RNA was found to be highly efficient with around 80-90% reduction in Ter94 reduction. Our data indicates that there is no change in the DL levels before and after ds-RNA knockdown, suggesting that DL degradation does not go through Ter94 dependent 26S proteasomal degradation. Cactus, the negative regulator of Toll pathway is known to sequester DL in the cytoplasm until Toll is activated. Following activation of Toll pathway , Cactus degrades releasing DL in the cytoplasm when then enters the nucleus and acts as a transcription factor. We predicted that $DL^{\alpha\beta}$ is unable to release itself from the sequestration of Cact and hence is not phosphorylated and translocated inside the nucleus. Hence a knockdown of Cact will result in an abundance of free DL in the cytoplasm which then can enter nucleus. Knocking down of Cact using ds-RNA against it seemed to be very robust, however we do not see any significant difference in the expression pattern of DL. $DL^{\alpha\beta}$ did not show recovery of the phosphoisoforms and degraded faster than wild type. Cact is known to degrade in the absence of DL. We see reduced Cact when cells are transfected with $DL^{\alpha\beta}$, which shows a highly reduced expression compared to the wild-type. Our data hints

towards an unknown degradation pathway other than 26S proteasome and autophagy that clears DL from the cell **(Supplementary Figure S4)**.

3.3.4 Generation of *dl$^{\alpha\beta}$* by CRISPR-Cas9 genome editing

In recent years, genome editing using CRISPR-Cas9 has become the go-to tool for generating mutations in a straightforward and site-directed manner **(Bassett and Liu 2014) (Bier et al. 2018).** After mapping the functional landscape of *dl$^{\alpha\beta}$* activity in cells, we focused on exploring the effect of predicted SIM mutation in the organism. We used CRISPR/Cas9 technology to generate a *dl$^{\alpha\beta}$* animal to achieve this. One of its key advantages is its targeted approach, which minimizes off-target modifications and ensures that the mutant allele is expressed under the control of its native promoter. Our approach specifically employed the Scarless system to achieve these outcomes **(Gratz et al. 2014)**.

The process of generating the *dl$^{\alpha\beta}$* mutant involved several carefully executed steps. First, a guide RNA (gRNA) targeting the third exon of the *dorsal* locus was designed and cloned into the pBFv-U6.2B vector. A transgenic fly line was then created by integrating this vector into the attP40 site on the second chromosome through embryonic microinjection **(Fig. 3.4A)**. The resulting single-guide RNA line (*sg3-dorsal*) was balanced and validated for further experiments.

To incorporate the two SIM mutations into the *dorsal* genomic locus, we used the pHD-Scarless-DsRed vector ((Gratz et al. 2014); Fig. 4B). This vector contains a DsRed cassette flanked by PBac inverted repeats (IR) targeting a TTAA site, with the IRs further flanked by homology regions (HR) specific to the target genomic region. The DsRed cassette, driven by the P3 promoter, serves as a visual marker for successful insertion, as its expression is restricted to the eye. The desired VVVV>AVVA and IILL>AILA mutations were incorporated into the 5′HR region, and the pHD-Scarless-dl$^{\alpha\beta}$ vector was constructed using Gibson assembly to assemble the mutated *dl* fragments with the vector components (Fig. 4B). DNA sequencing confirmed the accuracy of the construct.

Next, pHD-Scarless-dl$^{\alpha\beta}$ was injected into 360 embryos expressing Cas9 and *sg3-dorsal* (Fig. 4C, 4D). Emerging F1 flies were balanced using a second chromosome balancer, and successful integrations of the pHD-Scarless-dl$^{\alpha\beta}$ cassette were identified through the presence of DsRed fluorescence in the fly eyes. Homology-directed repair (HDR) proved to be

successful, resulting in 2 lines with successful integrations. These were then crossed with a transposase line to excise the DsRed cassette. This excision restored the duplicated TTAA site to a single genomic TTAA site, leaving a "scarless" modification.

From each of these lines, a single fly genomic DNA prep was made, and the entire Dl cassette containing the Ds-Red marker was amplified by PCR. Each single PCR band was then gel extracted, and purified DNA was sent for sequencing. The sequencing result was aligned against the template, and it was found that both lines contained only one mutation out of two (IILL-AILA). Interestingly, the IILL-AILA mutation alone was sufficient to produce the mutant phenotype in cells, while the VVVV-AVVA mutation was associated with a wild-type phenotype. From here on, we call these lines $dl^{\#28}$ and $dl^{\#29}$. These genome-edited lines allow us to study the effects of the SIM mutation on DL function with greater precision, as in the animal, the mutant transcript and protein levels are driven by an endogenous promoter, eliminating overexpression artifacts.

3.3.5 dl^{β} is a null allele

We checked mRNA levels for $dl^{\#28}$ and $dl^{\#29}$ to check if the mutant gene was transcribed correctly. dl^{1} (the classical Dl null allele) and dl^{K382R} (the SCR form of Dl) were used as controls along with w^{1118}. RT-PCR results showed that all the alleles showed mRNA levels comparable Next, we assessed the protein expression level of these mutant alleles. We took 0-3h embryos from homozygous mothers and visualized the expression of DL protein. Embryos from $dl^{\#28}$ and $dl^{\#29}$ mothers had no DL protein similar to $dl^{1,}$ whereas embryos from dl^{SCR} showed expression similar to the wild type **(Figure 3.5B1, B2).** Cactus, the negative regulator of DL, degrades in the absence of DL, and this was found to be true in our experiments. We checked Cactus protein levels in the DL mutants and found that it is absent in embryos from mothers homozygous for $dl^{\#28}$, $dl^{\#29,}$ and dl^{1} **(Figure 3.5B3).** Our results indicate that the mutant *dl* is transcribed correctly; however, the transcribed mRNA is either not translated, or the translated protein undergoes rapid degradation. This data directly correlates to our *Drosophila* 529SU cell culture data, which shows rapid degradation of the protein phospho-isoforms and total protein. Unlike cells in culture with basal DL expression, a maternally null female lays eggs without any wild-type DL/dl.

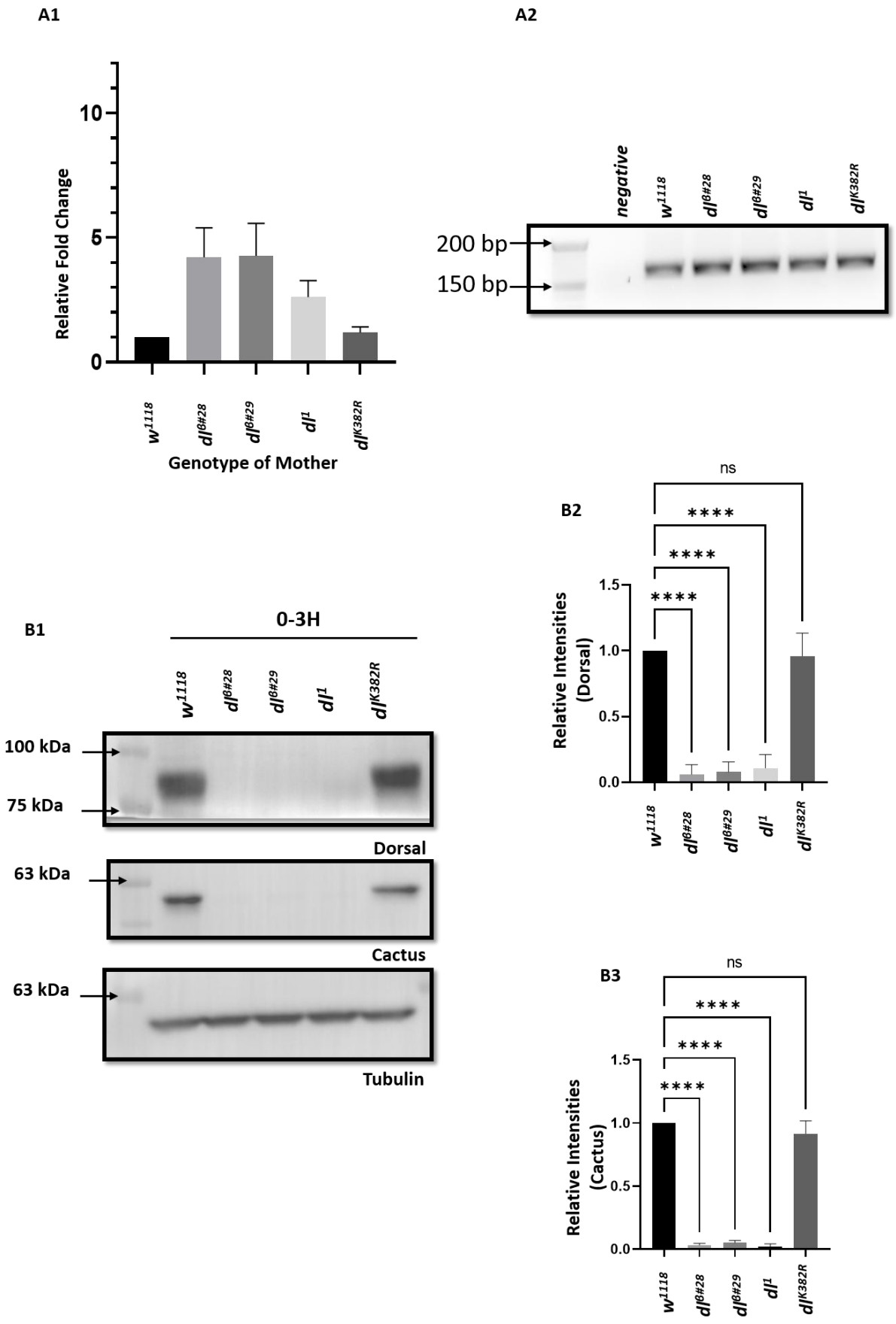

A1
Relative Fold Change
10
5
0
w^{1118}
$dl^{\beta\#28}$
$dl^{\beta\#29}$
dl^{1}
dl^{K382R}
Genotype of Mother
A2
negative
w^{1118}
$dl^{\beta\#28}$
$dl^{\beta\#29}$
dl^{1}
dl^{K382R}
200 bp
150 bp
B1
0-3H
w^{1118}
$dl^{\beta\#28}$
$dl^{\beta\#29}$
dl^{1}
dl^{K382R}
100 kDa
75 kDa
Dorsal
63 kDa
Cactus
63 kDa
Tubulin
B2
ns
****
****
****
Relative Intensities (Dorsal)
1.5
1.0
0.5
0.0
w^{1118}
$dl^{\beta\#28}$
$dl^{\beta\#29}$
dl^{1}
dl^{K382R}
B3
ns
****
****
****
Relative Intensities (Cactus)
1.5
1.0
0.5
0.0
w^{1118}
$dl^{\beta\#28}$
$dl^{\beta\#29}$
dl^{1}
dl^{K382R}

Figure 3.5 DL$^{\beta}$ is degraded post translationally

(A1-A2) Graph showing relative transcript levels for w^{1118} , dl^{β} *(L28),* dl^{β} *(L29),* dl^{1} and dl^{SCR} assayed by qRT-PCR from 0-3h old embryos. N=4, mean ±SEM. PCR product of the indicated genotype was loaded on an agarose gel to check the expression level. (B1) 0-3h embryonic lysates from w^{1118} , dl^{β}/ dl^{β}*(L28),* dl^{β} / dl^{β}*(L29),* dl^{1}/dl^{1} and dl^{SCR}/ dl^{SCR} were loaded on an SDS-PAGE gel and probed with rabbit anti-Dorsal (1:10,000) and mouse anti-Cactus (1:100) antibodies. Tubulin was used as a loading control. (B2) Western blots were quantified to determine the total protein using ImageJ and plotted as bar graph. N=4, mean ±SD One way ANOVA, (****) *P* <0.0001

3.3.6 dl^{β} does not support DV patterning

Results from the previous section demonstrated that the expression of DL protein from embryos derived from $dl^{\#28}$ and $dl^{\#29}$ mothers (from here on denoted as dl^{β}) mirrors that of the classical null allele dl^{1}. Next, we examined embryos derived from mothers homozygous for dl^{β}. Embryos from mothers homozygous for dl^{β} were embryonic lethal. To test if the lethality is because of any off-target effects, we used a dl null allele dl^{1} and also a deficiency (J4) line along with dl^{β} and found that the embryonic lethality persists **(Figure 3.6 A1-A3),** suggesting that the dlB locus was a dl null.

The dl^{1} allele is haplosufficient in terms of embryonic lethality at 25°C. Eggs laid by dl^{1}/ dl^{WT} develop in to normal adults at 25°C without any significant developmental defects. Its lethality could be rescued when a wild-type copy of *dl* is introduced maternally **(Fig 3.5A).** dl^{1} is haploinsufficient at 29°C, with 50-60% of embryos laid by dl^{1}/ dl^{WT} mothers dying before hatching. We tested the dl^{β} allele and found that it differs from dl^{1} in being haplosufficient at both 25°C and 29°C.

One of the well-established indicators of patterning defect is abnormal cuticle phenotype. The absence of maternal DL in the embryo correlates with the dorsalized cuticular phenotype. We sought to determine if our mutant allele shows a similar phenotype with respect to cuticular defects. Embryos laid by homozygous dl^{β} mothers show a severe dorsalized phenotype (D0), in accord with the 100% embryonic hatching defect. The cuticle of wild-type embryos shows thick denticle bands on the ventral side and fine hair on the dorsal side. In contrast, dl^{β} progeny displayed completely dorsalized cuticles lacking all ventral structures, corresponding to the observed high embryonic lethality. This result is comparable to cuticle preparations of progeny from dl^{1}, dl^{β}/ dl^{1}, and dl^{β} /J4 mothers **(Figure 3.6 B1, C1, D1, E1 and F1).** We immunostained

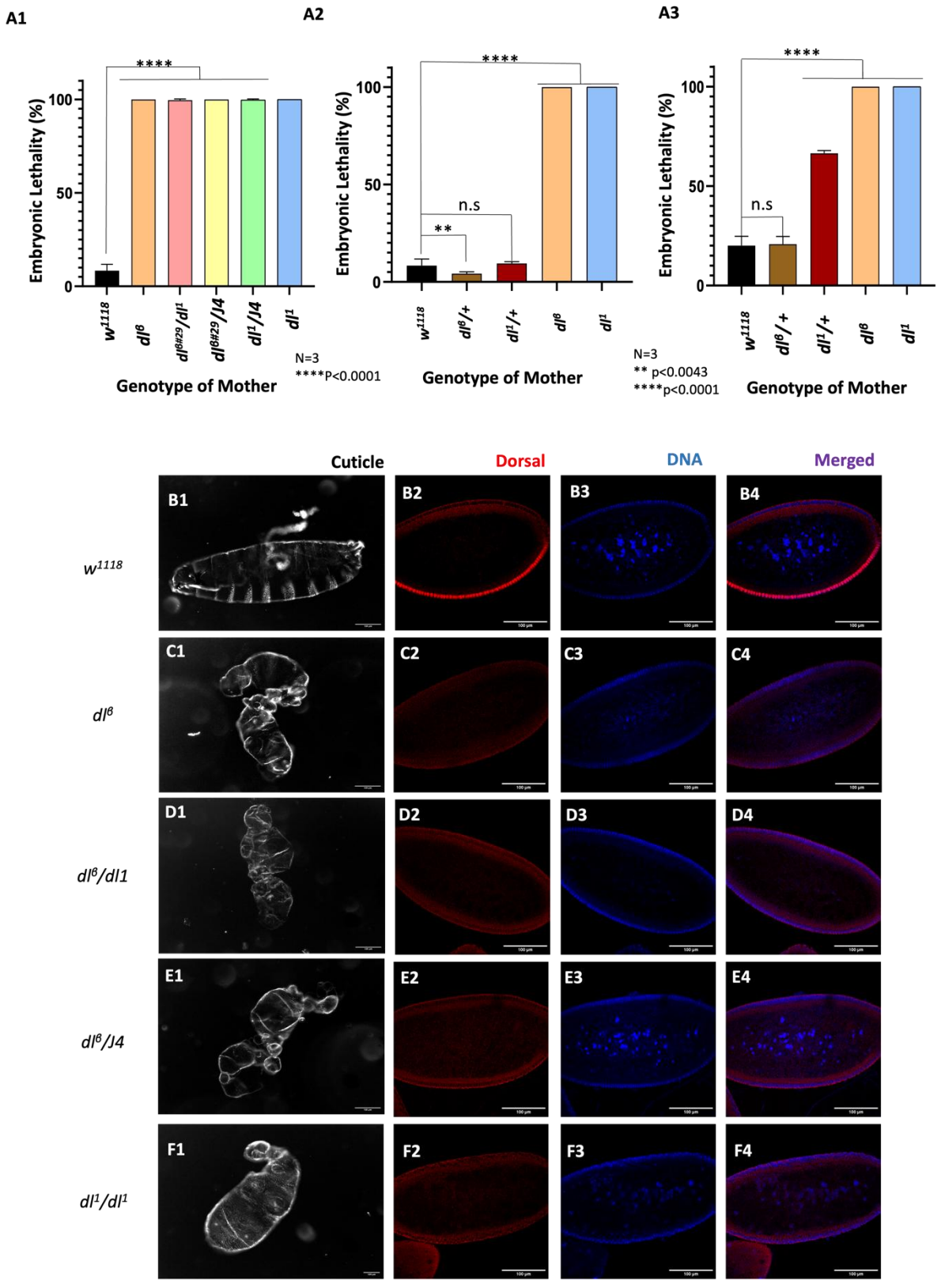


Figure 3.6 dl^{β} does not support dorsoventral patterning:

(A1) Embryos from indicated maternal genotype were scored for percentage lethality. dl^{β} mothers show high percentage lethality when compared to embryos from w1118. Similar lethality is also observed in embryos from dl^{1} (classical mutant), dl^{β} / dl^{1} and dl^{β} /J4 mothers. (A2) Percentage lethality of embryos from dl^{β} /+ and dl^{1} /+ mothers are comparable to that of w1118 embryos at 25°C. (A3) At 29°C about

half the embryos from dl^1/+ mother are non-viable in contrast to embryos from dl^β/+.N=3, mean ±SEM, 2- way ANOVA, (n.s) p<, (**) p<0.0043, (****)p<0.0001. (B1-F1) Depict cuticle preparations. Cuticle of wildtype embryo show thick denticle bands on the ventral side and fine hair on the dorsal side. Cuticles of progeny of dl^β are completely dorsalized and lack any ventral structures, which explains the high embryonic lethality observed. This results is comparable to cuticle preparations of progeny from dl^1, dl^β / dl^1 and dl^β /J4 mothers. (B2-F2) Depicts nuclear localization of DL on ventral side of the embryo, (B3-F3) Depicts 13-14th nuclear division in embryo and (B4-F4) depicts merged images.

0-3h embryos with DL antibody to further validate the absence of protein in embryos. Wild-type embryos show DL protein localized inside the nucleus of the ventral side of the embryo, whereas cytoplasmic DL was visible on the dorsal side. Embryos laid by mothers homozygous for dl^β show a complete absence of protein on the embryo's ventral and dorsal side, similar to the dl^1 null allele **(Figure 3.6 B2-B4, C2-C4, D2-D4, and E2-E4).** Our results show that dl^β is a strong hypomorph and a protein null. This mutant allele causes embryonic lethality and severe dorsalized cuticular phenotype in homozygous conditions.

3.3.7 Model of Dorsal-Cactus and Dorsal-Dorsal dimers

To understand the structural aspects of Dorsal and Cactus and their interactions, we have modelled the structure of DL homodimer and Dorsal-Cactus heterodimer using a combination of AlphaFold, MODELLER, and CLICK. Structures of the IkappaBalpha/NF-kappaB complex (PDB id:1NF1) **(Jacobs and Harrison 1998)** were used as the template to model dorsal cactus interaction. The structure of NF-kappaB p52 homodimer (PDB id:1A3Q) **(Cramer et al. 1997)** was used as the template for the DL homodimer. Individual domains predicted by AlphaFold **(Jumper et al. 2021; Abramson et al. 2024)** were superimposed on the respective templates and refined using MODELLER**(Šali and Blundell 1993)** and OpenMM **(Eastman et al. 2017)**.

The structure of DL-Cactus heterodimer shows a large binding interface consisting of 72 residues (36 from Cactus and 36 from DL). Surface charges representation shows that the negatively charged C terminus (PEST domain) of the Cactus binds to a positively charged cleft in the DL. This region of Dorsal, in its nuclear DNA binding pose, likely binds to DNA. Interestingly, the PEST domain of Cactus also contains multiple phosphorylation sites for Drosophila casein kinase II **(Liu et al. 1997)**. These phosphorylation sites can be mapped to the interface of the heterodimer.

The structure of the DL homo dimer shows interactions along the dimerization domain as well as along the C terminus alpha-helical region. We have identified a possible disulfide bond by

residue 231 between the two chains as well as histidine-mediated hydrogen bonds by residues 340 and 281 in the interface. This suggests that Dorsal dimerization could be affected by its cellular environment.

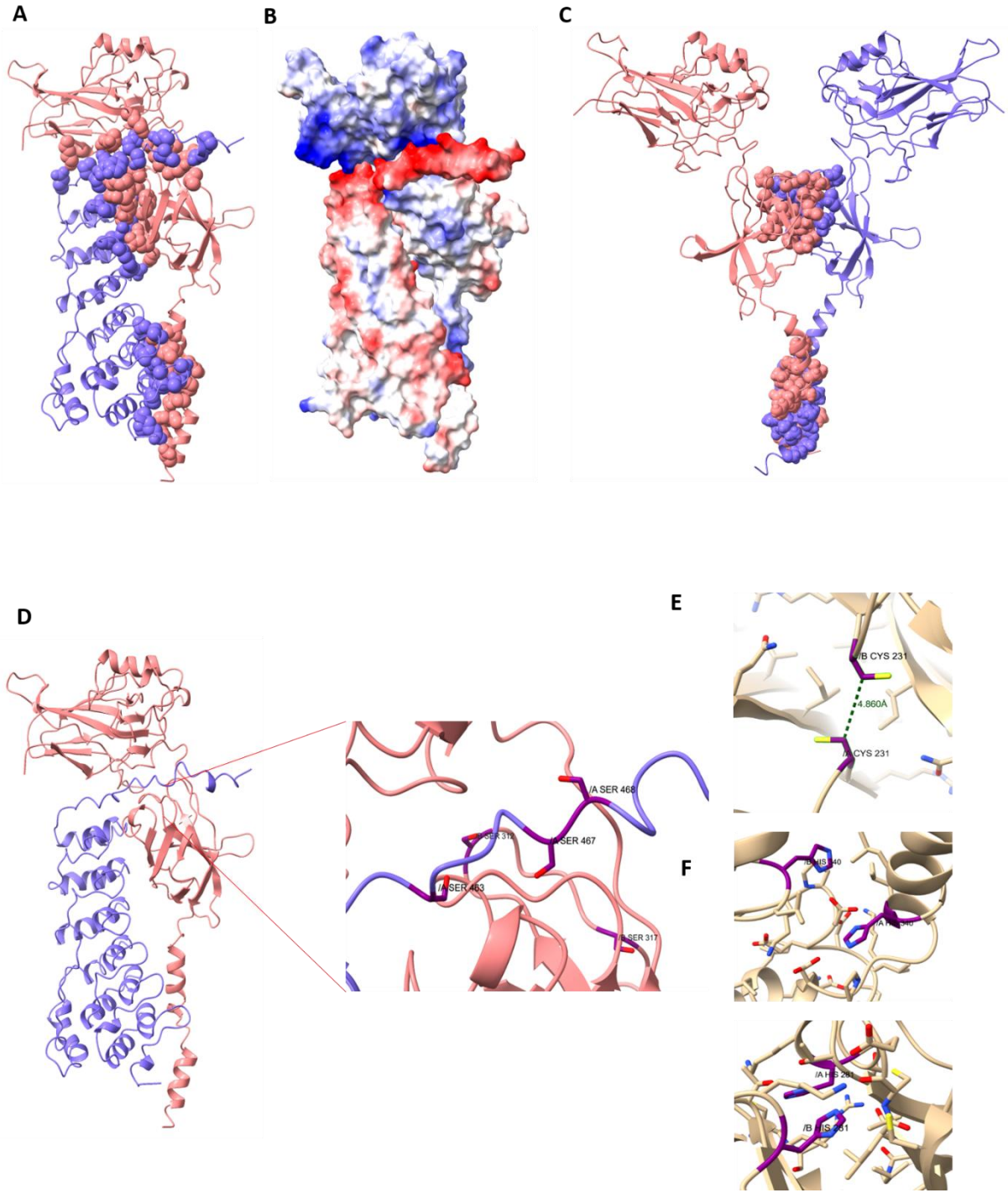


Figure 3.7 Modelling of DL:Cact dimer

(A) Dorsal cactus heterodimer with interface residue shown in sphere representation. Red: Dorsal 44-360 Blue: cactus 217-481 **(B)** Surface charges of Dorsal-Cactus dimer **(C)** Ribbon representation of Dorsal homodimer with the interface residues shown in sphere representation **(D)** Ribbon representation of modelled Dorsal-Cactus Dimer. Inset shows cactus phosphorylation sites in the PEST

domain and dorsal phosphorylation site 317 that is in proximity. **(E)** Candidate for disulfide bond in between monomers of dorsal dimer **(F)** Histidine mediated interactions in the binding interface

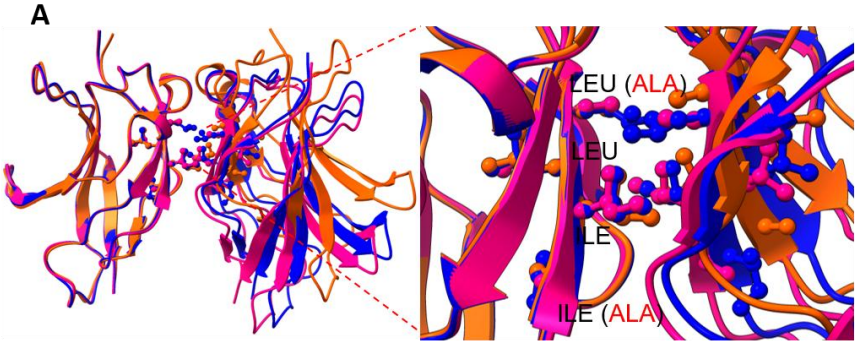
3.3.7 IILL to AILA mutation likely destabilizes Dorsal homodimer

In this section, we explore the effect of mutation of dorsal motif IILL (residue 243-246) to AILA on the stability of Dorsal homodimer. We first modeled the dimers (for both IILL and AILA) formed by the dimerization domain of dorsal using AlphaFold, MODELLER as well as MODELLER models subjected to a 200ns MD simulation. We selected the AlphaFold models of the dimers as they showed the most significant number of interactions. These models along with the homologous region from 1A3Q (a positive control), were subjected to molecular dynamics simulations for 200ns (3 replicates each). The interfaces formed by the AlphaFold model of wildtype and mutant and the MODELLER model of the mutant is show in the figure. The IILL/AILA motifs lie at the core of the interface.

The analysis of the trajectories of these simulations shows that the centroid distance between the monomers is noticeably lower for all replicates of the mutant dimers with AILA. The X-ray crystal structure 1A3Q showed the most consistent trajectory. The wildtype dimer switches between the centroid distance shown by 1A3Q and an intermediate state. This suggests a difference in the binding mechanism between the mutant and the wild type.

We then tracked hydrogen bonds between the two chains of the dimer in each frame of the trajectory. We found that the number of consistent hydrogen bonds in the mutant was lower compared to the wildtype (apart from one replicate), with values reaching as low as 1 in one of the mutant replicates. This suggests that the dimers with AILA in the interface have lower stability than the wild type due to missing hydrogen bonds **(Rose 2021).**

A



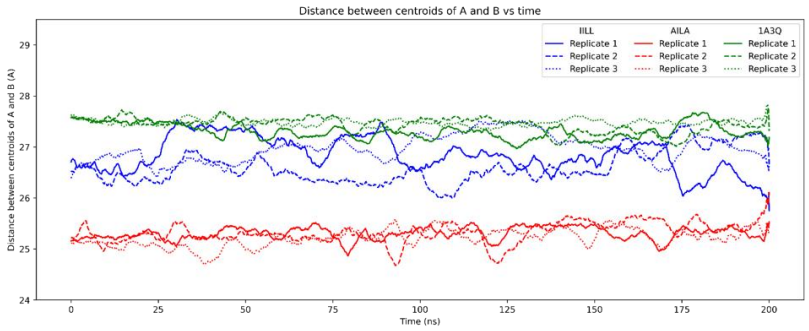
B

Structure 1	Structure 2	RMSD
IILL MODELLER	IILL Alphafold	1.68 Å
IILL MD last frame	IILL Alphafold	2.46 Å
AILA MODELLER	AILA Alphafold	1.38 Å
AILA MD last frame	AILA Alphafold	1.6 Å
IILL Alphafold	AILA Alphafold	1.6 Å

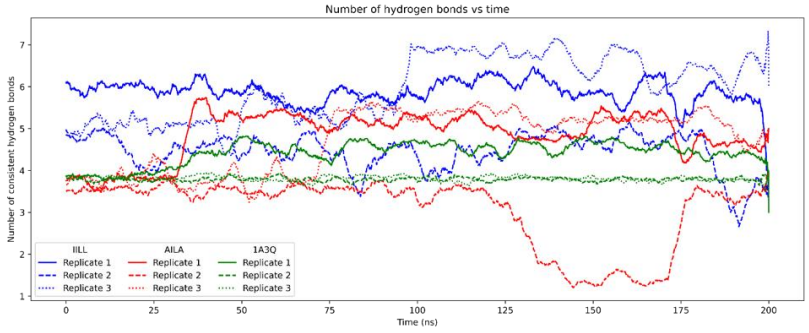
C

Structure	No. of H bonds	Salt bridges
IILL MODELLER	6	1
AILA MODELLER	5	0
IILL MD last frame	8	1
AILA MD last frame	8	2
IILL Alphafold	16	0
AILA Alphafold	14	0

D



E



F

Dimer	Interacting residues pairs	Raw score	Normalized raw score	Z score	Prediction
IILL-IILL (AlphaFold)	32	-124.702	-3.897	2.738	BINDER
IILL-AILA (AlphaFold)	30	-136.933	-4.564	2.726	BINDER
AILA-AILA (AlphaFold)	26	-715.225	-27.509	2.329	BINDER
IILL-IILL (MODELLER)	25	-2198.423	-87.937	1.285	NON-BINDER
AILA-AILA (MODELLER)	21	-2609.830	-124.278	0.656	NON-BINDER
IILL-IILL MD last frame (MODELLER)	31	-800.714	-25.829	2.358	BINDER
AILA-AILA MD last frame (MODELLER)	27	-1399.067	-51.817	1.909	BINDER

Figure 3.8 IILL to AILA mutation likely destabilizes Dorsal homodimer.

(A) Dimers of Dorsal dimerization domains superimposed by one chain showing different binding poses in models predicted by different methods. The regions with motif IILL (residue 243-246) is shown in ball and stick. An inset of the same is shown on the side with the residue types labeled (mutation in red). Blue: Model of Wildtype dimer with IILL motif by AlphaFold Pink: Model with mutant dimer with AILA motif by AlphaFold Orange: Model with mutant dimer with AILA motif by MODELLER. **(B)** RMSD of between models with IILL and AILA motifs, modelled by MODELLER, AlphaFold and the last frame from a 200ns MD simulation of the MODELLER model **(C)** Number of Hydrogen bonds and salt bridges between the interfaces among models from different methods. **(D)** Distance between the centroids of chains A and B across 200ns of MD for 3 replicates of IILL (blue), AILA(red) and 1A3Q(green) **(E)** Number of consistent hydrogen bonds across 200ns of MD for 3 replicates of IILL (blue), AILA(red) and 1A3Q(green) (F) PIZSA scores of interfaces formed by combinations of mutant and wildtype dimerization domains of Dorsal.

To evaluate the stability of the mutant and wildtype interface, we used Protein Interaction Z-Score Assessment (PIZSA) **(Dhawanjewar et al. 2019 Feb; Roy et al. 2019)**, a statical potential that evaluates the binding interfaces of proteins. PIZSA scores of the dimers are highest for the wild type (IILL-IILL), followed by the chimeric complex (IILL-AILA), and lowest for the mutant dimer (AILA-AILA). The number of interacting residues and Z score follows the same order. Nevertheless, PIZSA predicts all three dimers are 'BINDERS.'

3.4 Discussion

The *Drosophila* NFkappaB protein DL is well studied in the context of dorsoventral patterning in the early embryo as well as a key player in the toll-mediated immune response pathway. The activity of DL is highly regulated by post-translational modifications. Phosphorylation of DL has been shown to regulate its stability and nuclear import. In this context, S312 and S317 deserves special mention. Phosphorylation of S312 stabilizes the protein in the cytoplasm after it has been released from Cactus, post Toll activation. DL is a transcriptional factor that translocates into the nucleus after Toll pathway activation. Phosphorylation of S317 on DL regulates this nuclear import. Mutation of this Ser317 leads to a failure in its nuclear import resulting in degradation of the protein. The classical *dl* null allele (dl^{1}) is a S317N mutant. Apart from phosphorylation, SUMOylation plays an important role in regulating its transcriptional activity. DL is SUMOylated at K382. A SUMO Conjugation Resistant form of DL (dl^{SCR}) was found to act as a better transcriptional factor than the wild-type DL. The mammalian NFkappaB proteins are known to homo and hetero-dimerize and this is important for their activity. Although DL has been studied extensively for over 50 years, there still

remains a few unanswered questions. We are not able to specifically pinpoint the kinase that phosphorylates the protein or the machinery that degrades the protein.

Here, we revisit DL and try to answer some of these questions in the context of early embryonic development. We started off by mutating two predicted SUMO-Interacting Motifs (SIM) and looked at its effect on SUMOylation. One of the predicted SIM sites (SIMα VVVV) turned out to be completely redundant having no effect upon mutation. While the second predicted SIM site (SIMβ IILL) turned out to be a dimeric interface for DL. Mutation of this site showed a decrease in protein stability as well as loss of its phosphoisoforms. The total SUMOylated fraction of the protein was also decreased in the mutant (in-bacto SUMOylation assay). However, the decrease in SUMO fraction in S2 cells can be attributed to the degradation of the protein itself. This is similar to the expression pattern seen in DL phosphomutants (S312A and S317A). In all these cases, the protein is degraded and there are no phosphoisoforms. In order to investigate, how the protein is degraded we blocked the 26s proteasomal pathway with MG132 as well as autophagy with 3-MA and checked for the expression pattern of the mutant protein. None of these two well known precesses seem to contribute to the degradation of the protein and it seems that the protein might be degraded by a novel degradation machinery.

To further validate the results observed in *Drosophila* S2 cells, we generated CRISPR mutants harbouring the mutation (IILL – AILA). It was found that embryos laid by homozygous mutant mothers are embryonic lethal and produce cuticles that are completely dorsalized. Moreover, DL^{β} protein is completely absent as determined by both immunoblotting and immunostaining, but the mRNA is stable and transcribed properly as determined by real time PCR. dl^{β} allele was found to be haplosufficient, since a single wild type cope of *dl* was able to rescue the embryonic lethality. MD simulation and molecular modelling shows that the SIMβ site lies in the dimeric interface of DL. It creates a hydrophobic pocket that is essential for protein protein interaction. Mutation of this site disrupts this pocket which probably leads to failure in forming dimers, hence destabilizing the protein, triggering its degradation.

Here, we report that DL functions as a dimer and in its monomeric state it is unstable and degraded through an unknown mechanism. The degradation of the protein precedes its phosphorylation. Conversely, the kinase can only bind and phosphorylate the protein in its dimeric state. In the absence of a dimer, DL is degraded and so does Cactus in the absence of DL. DL^{β} is however able to dimerize in the presence of a wild type protein and can behave as

wild type. We discover a novel site on the DL protein which when mutated makes it a protein null.

3.5 Materials and methods

3.5.1 Fly husbandry and stocks

Flies were raised on standard cornmeal agar at 25°C unless stated otherwise. The following fly stock Was procured from the Bloomington Drosophila Stock Centre: dl^{1} /CyO (3236). The dl deficiency allele w-, y; J4/CyO containing a precise deletion of the dl and dif loci, was a kind gift from the Govind laboratory, City University of New York (CUNY), NY.

3.5.2 Generation of transgenic CRISPR lines

The gRNA was designed using the Target Finder tool (http://targetfinder.flycrispr.neuro.brown.edu/). A gRNA sequence was selected based on its lack of off-target sites and suitability for our experimental design. This sequence was then cloned into the pBFv-U6.2B vector. To generate homology regions, the genomic locus of *dl* was amplified from a single w^{1118} fly. The 5′ and 3′ homology regions (HR) of *dl* were prepared separately, with the IILL>AILA mutations introduced into the 5′HR. The pHD-scarless-DsRed vector was amplified in two fragments using specific primers. These fragments, along with the 5′HR and 3′HR of *dl*, were assembled using Gibson assembly (NEB) and verified by sequencing.

To screen for the $dl^{IILL>AILA}$ (dl^{β}) mutant after excising the DsRed cassette, genomic DNA was extracted by homogenizing single flies in 50 µL of DNA extraction buffer containing 10 mM Tris, 1 mM EDTA, and 20 µM Proteinase K. The homogenate was incubated at 37°C for 30 minutes, followed by heat inactivation at 85°C for 5 minutes. The dl^{β} and dl^{WT} fly lines were regularly sequenced to confirm the desired genotype. All the primers used are listed in ***(Supplementary Table S5).***

3.5.3 Embryonic lethality

0-3 hr embryos were collected, transferred to a fresh sugar-agar plate, and unhatched larvae were scored after 48 hours to determine viability.

3.5.4 ds-RNA synthesis and ds-RNA mediated knockdown

To prepare the dsRNA, we designed specific ~ 20-mer oligos to amplify a PCR product of approximately 400 bp from cDNA. These oligos were designed with a 50 T7 RNA polymerase binding site ***(Supplementary Figure S6).*** The PCR products were purified and used as the template for the preparation of dsRNA in vitro. The dsRNA was synthesized using the MEGAshortscript T7 kit (#AM1354; ThermoFisher Scientific, Waltham, MA, USA), following the manufacturer's guidelines. RNAeasy columns from QIAGEN (Hilden, Germany) (#74116) were used to purify the synthesized dsRNA following the kit instructions and eluted in the desired volume. Five micrograms of dsRNA was used for each knockdown experiment. The efficiency of the knockdown of genes was assessed by western blotting. Following dsRNA treatment, 529Su cells were harvested and lysed, and total protein was quantified. An equal amount of protein was loaded for SDS/PAGE, and western blotting was performed using appropriate antibodies to assess the extent of knockdown with dsRNA treatment.

3.5.5 Cloning and generation of constructs for overexpression

Dl (FBgn0260632) was amplified from the *Drosophila* gold collection library (https://www.fruitfly.org/EST/gold_collection.shtml) using specific primers ***(Supplementary Table S5).***These amplicons were independently cloned into the pET45b+ vector for bacterial SUMOylation assay and pRM-HA3 vector for transfection into 529Su cells using a modified Seamless Ligation Cloning Extract (SLiCE) protocol (Zhang Y et al., 2012). The site-directed mutagenesis approach with specific primers was used to modify the predicted SIMα (VVVV>AVVA) and SIMβ site (IILL>AILA). All the clones were confirmed by sequencing and used for downstream experiments.

3.5.6 S2 cell culture, transfections, cycloheximide treatment, LPS induction and immunostaining

S2 cells that were stably transfected with Flag-SUMO (referred to as 529SU cells) were a kind gift from Prof. Albert Courey. The cells were grown and maintained in Gibco Schneider's *Drosophila* Medium (Thermo Fischer Scientific, #21720024) supplemented with

10% heat-inactivated fetal bovine serum (FBS) (Thermo Fischer Scientific, #10082147) at 23°C. 1ug of plasmid was transfected per 1ml of cells using TransIT-Insect Transfection Reagent (Mirus, #6100) as per the manufacturer's protocol. The cells were induced with 0.5M CuSO4 and harvested after 24/48h of induction (as required). For the MG132 and 3-MA treatment, 529SU cells were seeded appropriately, transfected, and induced as described above. MG132 and 3-MA were added at a final concentration of 25μM and 2mM and harvested after 8h and 24h respectively. The protein was analyzed using a western blot.

3.5.7 Bacterial SUMOylation assay

This is a modified in-vitro SUMOylation assay that was previously described (Nie M et al., 2009). The quartet vector comprising of the *Drosophila* SUMO machinery components was co-transformed with HA tagged *dl*. The bacterial culture was induced with 1mM of Isopropyl β-D-1-thiogalactopyranoside (IPTG) for 6 hours at 25°C. 10ml of the bacterial culture was harvested in 1ml 50mM *tris* aminomethane (TRIS) buffer containing 150mM NaCl, 1mM Dithiothreitol (DTT), 1ug/ml lysozyme, and 1mM phenylmethylsulfonyl fluoride (PMSF). The cells were lysed using a VibraCell probe sonicator with 2/3sec ON/OFF cycle for 2 min at 60% amplitude.

3.5.8 Cuticle preparation

Embryos were collected for 3 hr and aged for 22 hr at 25 or 29C, depending on the nature of the experiment. They were dechorionated in a 4% sodium hypochlorite solution for 90s. Dechorionated embryos were washed thoroughly under running tap water and transferred to a scintillation vial containing 1:1 methanol: heptane. The vial was shaken vigorously for a few minutes, and de-vitellinized embryos in the lower methanol phase were transferred to a new vial with fresh methanol. Embryos were transferred onto a slide, mounted in 85% lactic acid, and incubated overnight at 55C on a slide warmer. Cuticles were imaged on a Zeiss Axio Imager Z1 microscope, using dark field illumination, with a 10 objective.

3.5.9 Embryo staining

The 0–3-hr embryos were dechorionated in 4% sodium hypochlorite for 2 min. Embryos were rinsed and fixed in a 1:1 solution of 4% formaldehyde in 1x phosphate-buffered saline

(PBS):heptane for 20 minutes. The aqueous phase containing formaldehyde was removed, and embryos were devitellinized by adding an equal volume of ice-cold methanol followed by vigorous shaking. Devitellinized embryos were washed thrice in methanol. Embryos were re-hydrated and permeabilized by giving six 15- min washes in 1 PBS containing 0.3% Triton X-100 (0.3% PBS-T). After blocking with 2% bovine serum albumin (BSA) in 0.3% PBST, embryos were incubated overnight at 4C with the primary antibody. Following four 15-min washes with 0.3% PBS-T, embryos were incubated with the secondary antibody for an hour at room temperature. Embryos were washed thrice in 0.3% PBS-T, and Hoechst was added in the penultimate wash. Embryos were mounted in SlowFade Gold mountant (Invitrogen) and imaged on a Zeiss Anisotropy Copnfocal microscope 20x oil-immersion objective. The following antibodies were used: Mouse anti-Dorsal, 1:1,000 (DSHB 7A4-c) and goat anti mouse Alexa 647 secondary antibody, 1:1,000 (Invitrogen). Hoechst was used to stain the nuclei.

3.5.10 Western blot analysis

For in-bacto SUMOylation assay, several transformed bacterial colonies, selected for kanamycin and ampicillin resistance, were picked to inoculate 5 ml cultures of a non-inducing medium and shaken at 37°C overnight. This culture was then used to inoculate 10 ml secondary culture. When the bacterial culture reached an OD of 0.4 to 0.5, it was induced using 0.5 mM IPTG and grown overnight at 18°C (180 rpm) to induce the expression of all five proteins. 1 ml culture was taken, boiled in SDS dye, and 20 µl was loaded in a 10% denaturing SDS gel and analyzed using western blot.

For cell-culture experiments, the cells were seeded in a six well tissue culture plate prior to start of transfection. Mirus TransIT-Insect was used as a transfection reagent. Induction was done using 0.5mM $CuSO_4$ post transfection and cells were harvested 24h post transfection. Collected cells were washed once in 1x PBS, boiled in SDS dye, and equal amount of protein was loaded in a 10% denaturing SDS gel and analyzed using western blot.

For flies, 0-3h embryos collected and were lysed in RIPA buffer (50 mM Tris-Cl, 150 mM NaCl, 0.1% SDS, 0.01% Sodium azide, 0.5% sodium deoxycholate, 1 mM EDTA, 1% Triton X-100, 1X PIC) with a pellet pestle (Kontes). Lysates were cleared by centrifugation at 21,000g at 4 °C for 30 minutes. Protein concentration was estimated using a BCA assay (Pierce) and 80 µg of total protein was loaded onto the gel after boiling in 1X Laemmli Sample Buffer. Proteins separated by 10% SDS-PAGE were transferred onto a PVDF membrane (Immobilon-E,

Merck) and blocked in 5% milk in Tris-Buffer Saline (TBS) with 0.1% Tween 20 (TBS-T) for an hour. Blots were then incubated overnight with primary antibody diluted in 5% milk in TBS-T, at 4 °C. Following three washes with TBS-T, blots were incubated with secondary antibodies diluted in 5% milk in TBS-T, for 1 hour at room temperature. Blots were washed thrice with TBS-T and visualized on a LAS4000 Fuji imaging system after incubating with Immobilon Western Chemiluminescent HRP substrate (Merck). The following antibodies were used: Mouse anti-α-Tubulin, 1:10000 (T6074, Sigma), Rabbit anti-HA 1:2000 (04-902 DW-2, Sigma-Aldrich), Rabbit anti-VCP, 1:2000, Mouse anti-Dorsal 7A4 (DSHB, 1:1000), Rabbit anti- Dorsal (1:10,000) (A kind gift from Courey Lab, Mouse anti Cactus (DSHB, 1:50). All secondary antibodies Goat anti-rabbit HRP and Goat anti-mouse HRP secondary antibodies, each at 1:10000 (Jackson ImmunoResearch).

3.5.11 Quantitative PCR

RNA was extracted from appropriately staged embryos (0-3H post fertilization) using the RNeasy Plus Universal mini kit (Qiagen) according to the manufacturer's instructions. 1 μg of total RNA was used to generate cDNA using the High-Capacity cDNA Reverse Transcription kit (Thermo Fisher Scientific). The qPCR reaction was performed on a qTOWER3 real-time thermal cycler (Analytik Jena) with KAPA SYBR FAST master mix (SigmaAldrich). Gene expression was monitored using gene-specific primers. Transcript levels were calculated using the comparative Ct method to obtain fold change values. Relative mRNA levels were calculated using the delta Ct values. Rp49 was used as a reference gene. The following primer pairs were used (Forward primer, F and reverse primer, R):

rp49 F: GACGCTTCAAGGGACAGTATC,

rp49 R: AAACGCGGTTCTGCATGAG;

dl F: ATCCGTGTGGATCCGTTTAA,

dl R: AATCGCACCGAATTCAGATC;

3.5.12 Modelling Dorsal-Cactus and Dorsal-Dorsal dimer

The structures of Dorsal and Cactus monomers were modelled using AlphaFold **(Abramson et al., 2024; Jumper et al., 2021).** The low pLDDT regions corresponding to the unstructured parts were removed resulting in a Dorsal segment ranging from residues 45-360 and a Cactus segment ranging from residues 217-481. The two domains of Dorsal modelled by AlphaFold are in its nuclear, DNA binding pose. Since we aimed to also model the Dorsal-Cactus dimer

in the cytoplasmic state, we segmented Dorsal (45-360) to Dorsal-A (45-222) and Dorsal-B (222-360). These domains were then superimposed onto the structure 1NFI **(Jacobs and Harrison, 1998)** using the CLICK (**Nguyen et al., n.d.; Nguyen and Madhusudhan, 2011)**. MODELLER **(Šali and Blundell, 1993)** was used to energy minimize the resultant structure and reattach the Dorsal-A and Dorsal-B. The remaining clashes were removed and further minimized using, molecular dynamics using OpenMM **(Eastman et al., 2017)** energy minimization. Dorsal-Dorsal dimer was modelled using AlphaFold as well as the above method using 1A3Q **(Cramer et al., 1997)** as the template.

3.5.13 Evaluating the binding interface of Dorsal-Dorsal dimer

Molecular dynamics simulations were performed for the dimerization domain of Dorsal (222-360) using GROMACS **(Pronk et al., 2013)** for 200ns in triplicates with AMBER99SB-ILDN **(Lindorff-Larsen et al., 2010)** force field. The trajectory was sampled every 0.1ns to generate a trajectory with 2000 frames in PDB format **(Berman et al., 2000).** A python3 implementation of Kearsley's algorithm **(Kearsley, 1989)** was used to superimpose models with each other. These files were then analysed using python3 scripts to evaluate RMSD, RMSF, centroid distances and hydrogen bonds. Hydrogen bond donors and acceptors within a range of 2.5-3.5 Å Euclidian distance and donor-acceptor-antecedent angle above 90º were considered as hydrogen bonded. Hydrogen bonds that are found in at least half the frames were deemed as consistent hydrogen bonds. Protein Interaction Z-Score Assessment (PIZSA) was used to score the interfaces via the webserver (http://cospi.iiserpune.ac.in/pizsa/) **(Dhawanjewar et al., 2019; Roy et al., 2019).**

3.6 Contributions

Amrita helped with characterizing the CRISPR mutant fly. Mukundan and Madhusudan did the structure dimer analysis and all related bioinformatics.

3.7 Acknowledgements

Stocks obtained from the Bloomington Drosophila Stock Center (NIH P40OD018537) were used in this study: NCBS fly facility and Deepti Trivedi for CRISPR injections. Snehal Patil and Yashwant Pawar are responsible for fly media and stock maintenance; IISER Microscopy facility, Vijay Vittal, is responsible for training and maintenance.

3.8 References

Abramson J, Adler J, Dunger J, Evans R, Green T, Pritzel A, Ronneberger O, Willmore L, Ballard AJ, Bambrick J, et al. (2024). Accurate structure prediction of biomolecular interactions with AlphaFold 3. Nature. 630(8016):493–500. doi:10.1038/s41586-024-07487-w.

Baeuerle PA, Baltimore D. (2018). Linked references are available on JSTOR for this article : 1K B : A Specific Inhibitor of the NE-KB Transcription Factor. 242(4878):540–546.

Bassett A, Liu J. (2014). CRISPR / Cas9 mediated genome engineering in Drosophila. Methods. 69(2):128–136. doi:10.1016/j.ymeth.2014.02.019. http://dx.doi.org/10.1016/j.ymeth.2014.02.019.

Bayer P, Arndt A, Metzger S, Mahajan R, Melchior F, Jaenicke R, Becker È. (1998). Structure Determination of the Small Ubiquitin-related Modifier SUMO-1.

Bernier-villamor V, Sampson DA, Matunis MJ, Lima CD. (2002). Structural Basis for E2-Mediated SUMO Conjugation Revealed by a Complex between Ubiquitin-Conjugating Enzyme Ubc9 and RanGAP1. 108:345–356.

Bhaskar V, Smith M, Courey AJ. (2002). Conjugation of Smt3 to Dorsal May Potentiate the Drosophila Immune Response. 22(2):492–504. doi:10.1128/MCB.22.2.492.

Bhaskar V, Valentine SA, Courey AJ. (2000). A Functional Interaction between Dorsal and Components of the Smt3 Conjugation Machinery *. J Biol Chem. 275(6):4033–4040. doi:10.1074/jbc.275.6.4033. http://dx.doi.org/10.1074/jbc.275.6.4033.

Bier E, Harrison MM, O'connor-Giles KM, Wildonger J. (2018). Advances in engineering the fly genome with the CRISPR-Cas system. Genetics. 208(1):1–18. doi:10.1534/genetics.117.1113.

Cramer P, Larson CJ, Verdine GL, Müller CW. (1997). Structure of the human NF-κB p52 homodimer-DNA complex at 2.1 Å resolution. EMBO J. 16(23):7078–7090. doi:https://doi.org/10.1093/emboj/16.23.7078.

Dhawanjewar AS, Roy AA, Madhusudhan MS. (2019) Feb. A knowledge-based scoring function to assess the stability of quaternary protein assemblies. bioRxiv.:562520. doi:10.1101/562520.

Drier EA, Huang LH, Steward R. (1999). Nuclear import of the Drosophila Rel protein Dorsal is regulated by phosphorylation. :556–568.

Eastman P, Swails J, Chodera JD, McGibbon RT, Zhao Y, Beauchamp KA, Wang L-P, Simmonett AC, Harrigan MP, Stern CD, et al. (2017). OpenMM 7: Rapid development of high performance algorithms for molecular dynamics. PLOS Comput Biol. 13(7):e1005659.

Ghosh G, Van Duyne G, Ghosh S, Sigler PB.(1995). Structure of nf-κb p50 homodimer bound to a κb site. Nature. 373(6512):303–310. doi:10.1038/373303a0.

Ghosh S, Gifford AM, Riviere LR, Tempst P, Nolan GP, Baltimore D. (1990). Cloning of the p50 DNA binding subunit of NF-κB: Homology to rel and dorsal. Cell. 62(5):1019–1029. doi:10.1016/0092-8674(90)90276-K.

Gillespie SKH, Wasserman SA. (1994). dorsal , a Drosophila Rel-Like Protein , Is Phosphorylated upon Activation of the Transmembrane Protein Toll. 14(6):3559–3568.

Gilmore TD, Morin PJ. (1993). The IκB proteins: members of a multifunctional family. Trends Genet. 9(12):427–433. doi:10.1016/0168-9525(93)90106-R.

Govind S, Brennan L, Steward R. (1993). Homeostatic balance between dorsal and cactus proteins in the Drosophila embryo. Development. 117(1):135–148. doi:10.1242/dev.117.1.135.

Govind S, Steward R. (1991). Dorsoventral pattern formation in Drosophila: signal transduction and nuclear targeting. Trends Genet. 7(4):119–125. doi:10.1016/0168-9525(91)90456-Z.

Gratz SJ, Ukken FP, Rubinstein CD, Thiede G, Donohue LK, Cummings AM, Oconnor-Giles KM. 2014. Highly specific and efficient CRISPR/Cas9-catalyzed homology-directed repair in Drosophila. Genetics. 196(4):961–971. doi:10.1534/genetics.113.160713.

Hay RT. (2005). SUMO: A history of modification. Mol Cell. 18(1):1–12. doi:10.1016/j.molcel.2005.03.012.

Hegde S, Sreejan A, Gadgil CJ, Ratnaparkhi GS. (2022). SUMOylation of Dorsal attenuates Toll / NF- j B signaling. 221(May).

Jacobs MD, Harrison SC. (1998). Structure of an IκBα/NF-κB Complex. Cell. 95(6):749–758. doi:https://doi.org/10.1016/S0092-8674(00)81698-0.

Johnson ES. (2004). Rotein odification. doi:10.1146/annurev.biochem.73.011303.074118.

Jumper J, Evans R, Pritzel A, Green T, Figurnov M, Ronneberger O, Tunyasuvunakool K, Bates R, Žídek A, Potapenko A, et al. (2021). Highly accurate protein structure prediction with AlphaFold. Nature. 596(7873):583–589. doi:10.1038/s41586-021-03819-2.

Kidd S. (1992). Characterization of the Drosophila cactus locus and analysis of interactions between cactus and dorsal proteins. Cell. 71(4):623–635. doi:10.1016/0092-8674(92)90596-5.

Lin DY, Huang YS, Jeng JC, Kuo HY, Chang CC, Chao TT, Ho CC, Chen YC, Lin TP, Fang HI, et al. (2006). Role of SUMO-Interacting Motif in Daxx SUMO Modification, Subnuclear Localization, and Repression of Sumoylated Transcription Factors. Mol Cell. 24(3):341–354. doi:10.1016/j.molcel.2006.10.019.

Liou HC, Baltimore D. (1993). Regulation of the NF-ηB/rel transcription factor and IηB inhibitor system. Curr Opin Cell Biol. 5(3):477–487. doi:10.1016/0955-0674(93)90014-H.

Liu ZP, Galindo RL, Wasserman SA. (1997). A role for CKII phosphorylation of the cactus PEST domain in dorsoventral patterning of the Drosophila embryo. Genes Dev. 11(24):3413–3422. doi:10.1101/gad.11.24.3413.

Mcgehee J, Stathopoulos A. (2024). Mechanisms for controlling Dorsal nuclear levels. 5(August):1–6. doi:10.3389/fcell.2024.1436369.

Morisato D, Anderson K V. (1995). Signaling pathways that establish the dorsal-ventral pattern of the Drosophila embryo. Annu Rev Genet. 29:371–399. doi:10.1146/annurev.genet.29.1.371.

Mossessova E, Lima CD. (2000). Ulp1-SUMO Crystal Structure and Genetic Analysis Reveal Conserved Interactions and a Regulatory Element Essential for Cell Growth in Yeast. 5:865–876.

Müller CW, Van Duyne FA, Sodeoka M, Verdine GL, Harrison SC. (1995). Structure of the nf-κb p50 homodimer bound to dna. Nature. 373(6512):311–317. doi:10.1038/373311a0.

Nie M, Xie Y, Loo JA, Courey AJ. (2009). Genetic and proteomic evidence for roles of Drosophila SUMO in cell cycle control, Ras signaling, and early pattern formation. PLoS One. 4(6). doi:10.1371/journal.pone.0005905.

Nolan GP, Ghosh S, Liou HC, Tempst P, Baltimore D. (1991). DNA binding and IκB inhibition of the cloned p65 subunit of NF-κB, a rel-related polypeptide. Cell. 64(5):961–969.

doi:10.1016/0092-8674(91)90320-X.

Pasteur AL, Rnai D, Hu Y, Sopko R, Chung V, Foos M, Studer RA, Landry SD, Liu D, Rabinow L, et al. (2019). iProteinDB : An Integrative Database of Drosophila Post-translational Modi fi cations. 9(January):1–11. doi:10.1534/g3.118.200637.

Pidugu LS, Servius HW, Espinosa KB, Cook ME, Varney KM, Drohat AC. (2024). Sumoylation of thymine DNA glycosylase impairs productive binding to substrate sites in DNA. J Biol Chem. 300(11):107902. doi:10.1016/j.jbc.2024.107902. https://doi.org/10.1016/j.jbc.2024.107902.

Rose GD. (2021). Protein folding - seeing is deceiving. Protein Sci. 30(8):1606–1616. doi:https://doi.org/10.1002/pro.4096.

Roth S, Hiromi Y, Godt D, Nüsslein-Volhard C. (1991). Cactus, a maternal gene required for proper formation of the dorsoventral morphogen gradient in Drosophila embryos. Development. 112(2):371–388. doi:10.1242/dev.112.2.371.

Roth S, Stein D, Nüsslein-Volhard C. (1989). A gradient of nuclear localization of the dorsal protein determines dorsoventral pattern in the Drosophila embryo. Cell. 59(6):1189–1202. doi:10.1016/0092-8674(89)90774-5.

Roy AA, Dhawanjewar AS, Sharma P, Singh G, Madhusudhan MS. (2019). Protein Interaction Z Score Assessment (PIZSA): an empirical scoring scheme for evaluation of protein-protein interactions. Nucleic Acids Res. 47(W1):W331–W337. doi:10.1093/nar/gkz368.

Ruben SM, Dillon PJ, et al. (2016). Isolation of a rel-Related Human cDNA That Potentially Encodes the 65-kD Subunit of NF- κB Published by : American Association for the Advancement of Science Stable URL : http://www.jstor.org/stable/2875831 JSTOR is a not-for-profit service that helps sch. 251(5000):1490–1493.

Rushlow C, Warrior R. (1992). The rel family of proteins. BioEssays. 14(2):89–95. doi:10.1002/BIES.950140204. [accessed (2024) Nov 12]. https://onlinelibrary.wiley.com/doi/full/10.1002/bies.950140204.

Rushlow CA, Han K, Manley JL, Levine M. (1989). The graded distribution of the dorsal morphogen is initiated by selective nuclear transport in Drosophila. Cell. 59(6):1165–1177. doi:10.1016/0092-8674(89)90772-1.

Ryseck R-P, Bull P, Takamiya M, Bours V, Siebenlist U, Dobrzanski P, Bravo R. (1992). RelB,

a New Rel Family Transcription Activator That Can Interact with p50-NF-κB. Mol Cell Biol. 12(2):674–684. doi:10.1128/mcb.12.2.674-684.1992.

Šali A, Blundell TL. (1993). Comparative Protein Modelling by Satisfaction of Spatial Restraints. J Mol Biol. 234(3):779–815. doi:10.1006/jmbi.1993.1626.

Sieben U, Franzoso G, Brown K. (2024). Structure , regulation and function of nf-1d3.

Steward R. (1989). Relocalization of the dorsal protein from the cytoplasm to the nucleus correlates with its function. Cell. 59(6):1179–1188. doi:10.1016/0092-8674(89)90773-3.

Varejão N, Lascorz J, Li Y, Reverter D. (2020). Molecular mechanisms in SUMO conjugation. Biochem Soc Trans. 48(1):123–135. doi:10.1042/BST20190357. [accessed 2024 Nov 12]. /biochemsoctrans/article/48/1/123/221692/Molecular-mechanisms-in-SUMO-conjugation.

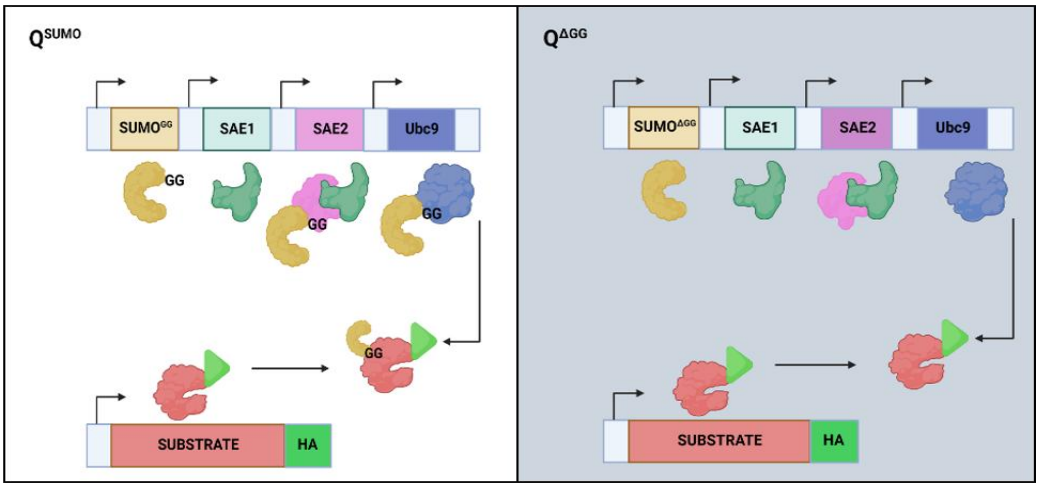
Verma IM, Stevenson JK, Schwarz EM, Antwerp D Van. (1995). of Association and Dissociation. :2723–2735.

Wang Y, Dasso M, Wang Y, Dasso M. (2009). SUMOylation and deSUMOylation at a glance SUMOylation and DeSUMOylation at a Glance. 1:4249–4252. doi:10.1242/jcs.050542.

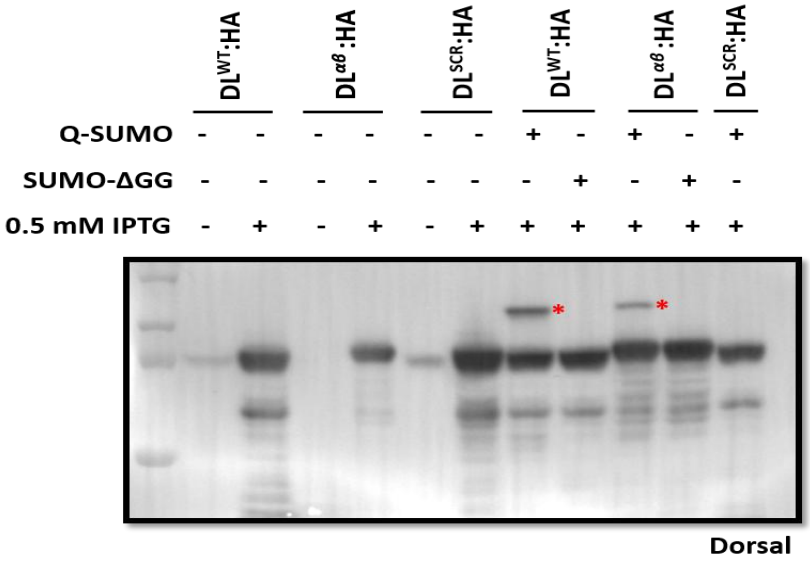
Zhang Y, Werling U, Edelmann W. SLiCE: a novel bacterial cell extract-based DNA cloning method. Nucleic Acids Res. (2012) Apr;40(8):e55. doi: 10.1093/nar/gkr1288. Epub 2012 Jan 12. PMID: 22241772; PMCID: PMC3333860.

3.9 Supplementary materials

A



B



C

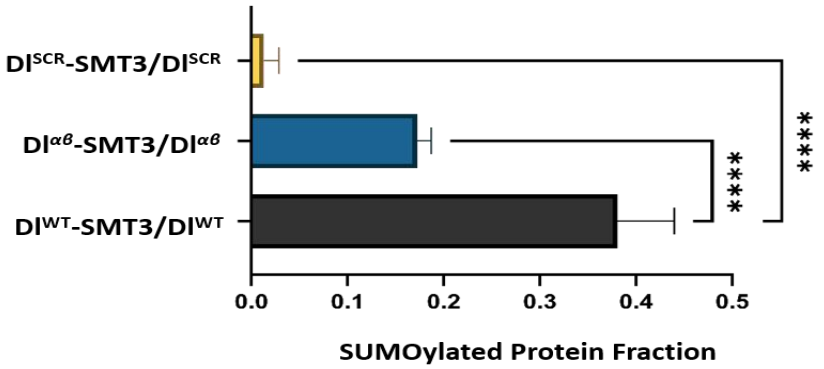
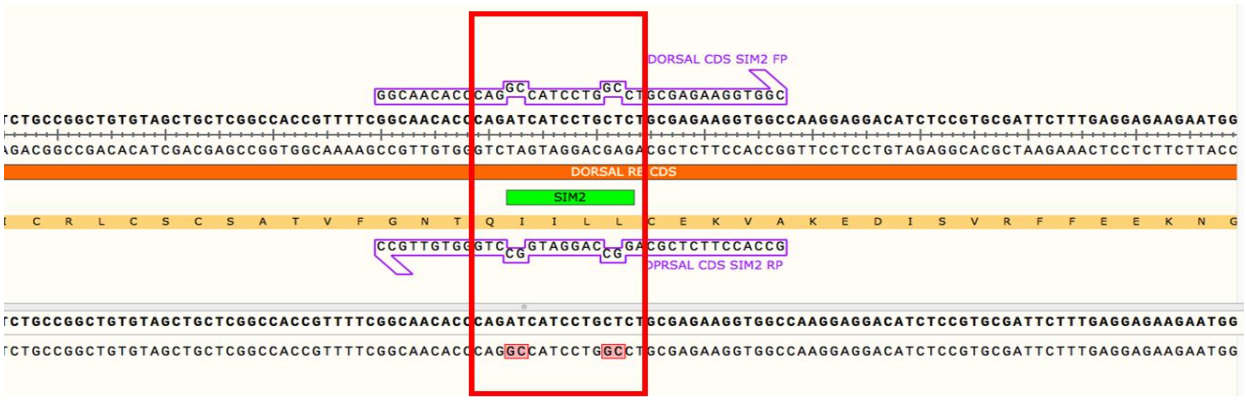


Fig. S1. ***In-bacto SUMOylation assay of Dorsal shows SIM sites are important***

(A) A bacterial sumoylation assay. The Q^{SUMO} vector, which encodes the mature form of SUMO ($SUMO^{GG}$) along with SAE1, SAE2, and Ubc9 expressed from separate *T7/lac* promoters, was used in combination with a vector expressing a HA-tagged candidate substrate. As a negative control, $Q^{\Delta GG}$, which expresses a conjugation defective form of SUMO ($SUMO^{\Delta GG}$), was used in place of Q^{SUMO}. (B) Bacterial sumoylation assay was used to validate the effect of SIM sites on the SUMOylation status of Dorsal. HA-tagged Dorsal was expressed in BL21 cells co-transformed with Q^{SUMO} or $Q^{\Delta GG}$ vectors, and immunoblotted using antibodies against HA and Dorsal. Red asterisk point to the bands representing sumoylated proteins. (C) Western blots were quantified to determine percentage of the total protein SUMOylated using ImageJ and plotted as bar graph. N=4, One way ANOVA, (***) $P < 0.001$, (*) $P < 0.05$

SIM1: VVVV - AVVA

SIM1: IILL - AILA



SCR: K382R

Figure S2: Multiple Sequence Alignment of *Drosophila* $dl^{\alpha\beta}$ *and* dl^{SCR} lines

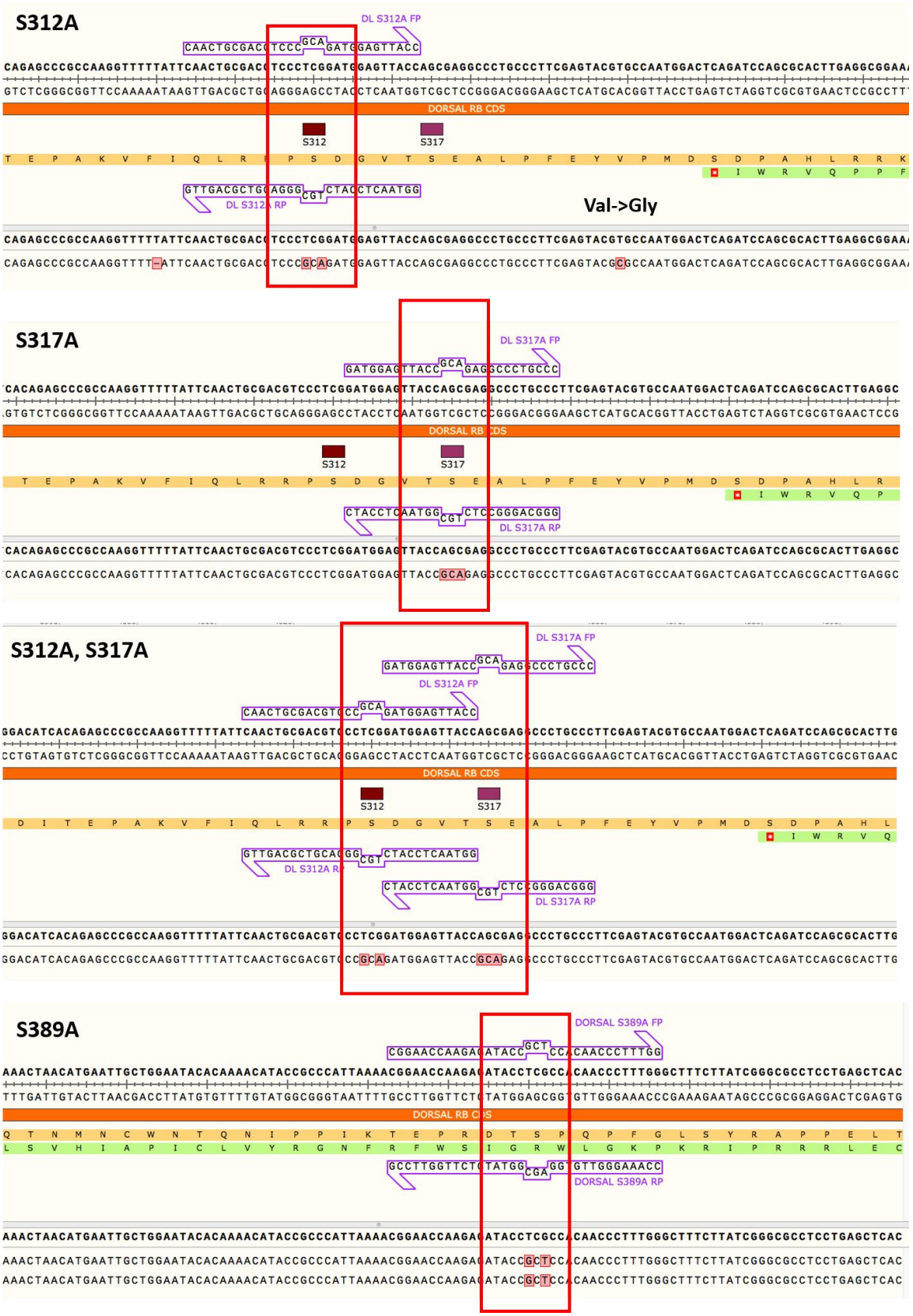


Figure S3: Multiple Sequence Alignment of *Drosophila* dl^{S312A} , dl^{S317A} , $dl^{S312A,S317A}$ and dl^{S389A} lines

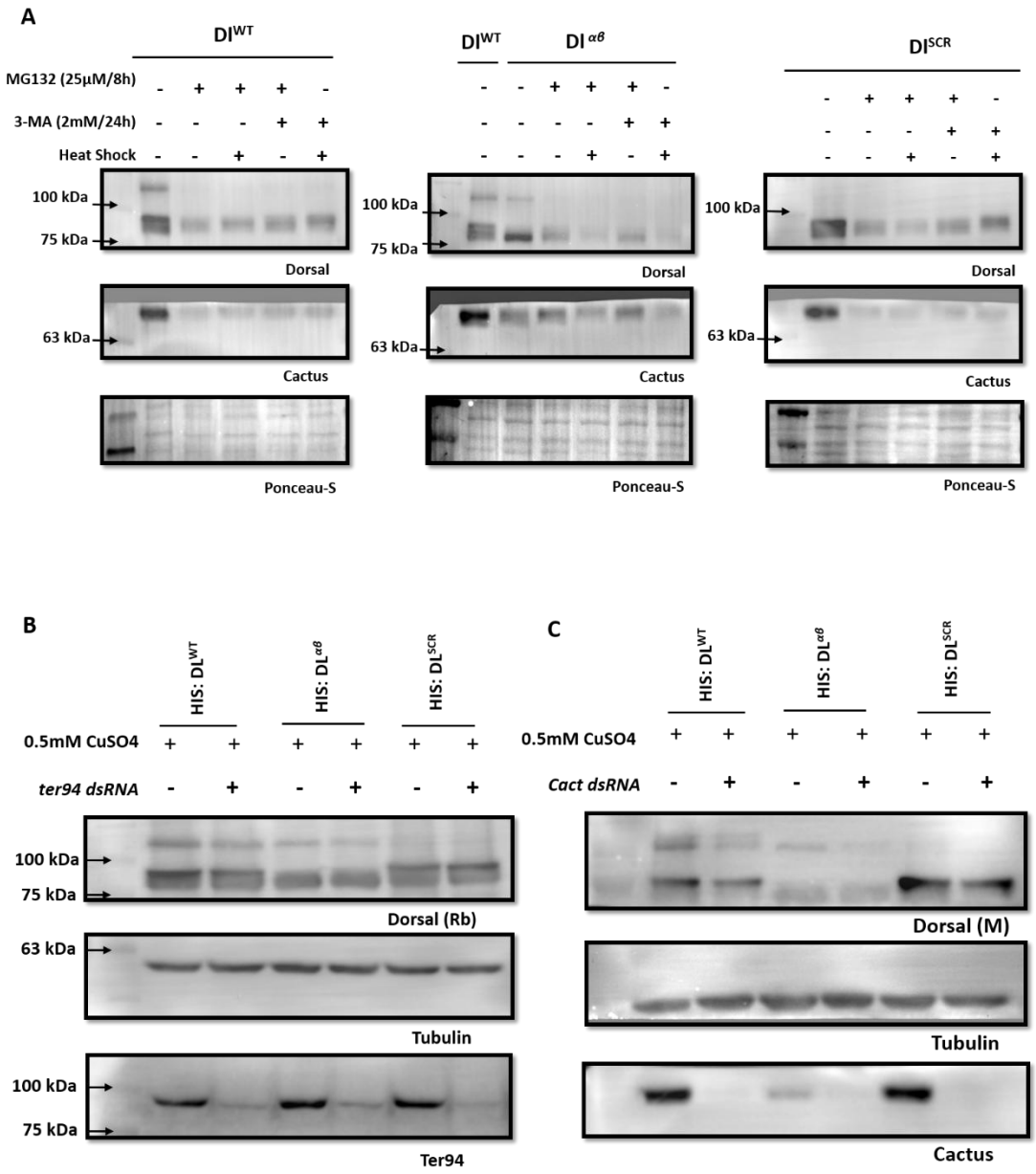


Fig. S4. Dl$^{\alpha\beta}$ does not undergo degradation through 26s Proteasome.

(A) Immunoblot showing effect of MG132, 3-MA and heat shock on DL^{WT}, DL$^{\alpha\beta}$ and DL^{SCR}. Cactus is used as an internal control. Equal protein loading is shown using Ponceau staining. **(B)** Immunoblot showing effect of Ter94 dsRNA treatment on DL. Tubulin is used as a loading control. Lower panel shows efficacy of Ter94 dsRNA knockdown in 529SU cells. **(C)** Immunoblot showing effect of Cactus dsRNA treatment on DL. Tubulin is used as a loading control. Lower panel shows efficacy of Cactus dsRNA knockdown in 529SU cells.

Table S5: List of Dorsal primers used during the course of experiments

PRIMER NAME	*SEQUENCE (5' - 3')*
dorsal-hr-F	*GCTCACGATTAACGAAGACC*
dorsal-ga-5'F	*CCTATAGGAGACGTATATGGTCTTCTTTTCCCGAGGCAGCAGTTGATGGCCAACAGAG*
seq screen dorsal hr F	*GATCACCGAACAACCGGCAG*
dorsal hr sdm avva 3'R	*CCACAATTTCGATTGTCGGATAG*
dorsal hr sdm avva 5'F	*GCTACAAGGGACGCGCCGCAGTTGTTGCCTCCTGCGTCACAAAGG*
dorsal hr sdm aila 5'F	*GTTTTCGGCAACACCCAGGCCATCCTGGCCTGCGAGAAGGTGGCCAAG*
dorsal hr sdm aila 3'R	*GGTGGCCGAGCAGCTACACAG*
seq screen dorsal hr R	*GAAGTCACCAAAGGCCTCCC*
dorsal-ga-5'R	*CAATATGATTATCTTTCTAGGGTTAAAGGTAATGGCAGTCTGCTTGTG*
dorsal-ga3'F	*CGCAGACTATCTTTCTAGGGTTAAGACGCCGCGCTATCATACC*
dorsal-hr-R	*CCTCGCTGGTAACTCCATC*
dorsal-ga-3'R	*GTATAGGAGACCTATAGTGTCTTCGGGGCCGAGGCCTCTGAACTTTGAAGTCCGTG*
t7 eurofins FP	*TAATACGACTCACTATAGGG*
dorsal pet45b cds FP	*GAAGGAGATATACCATGGCAATGTTTCCGAACCAGAACAATGGAG*
pet45b+ RP	*TGCCATGGTATATCTCCTTCTTAAAGTTAAACAAAA*
dorsal cds sim1 FP	*CAAGGGACGCGCCGCCGTTGTTGCGTCCTGCGTCAC*
dorsal cds sim1 RP	*GTGACGCAGGACGCAACAACGGCGGCGCGTCCCTTG*
rt dl1 fp dl1 seq FP	*ATCCGTGTGGATCCGTTTAA*
dorsal cds sim2 FP	*GGCAACACCCAGGCCATCCTGGCCTGCGAGAAGGTGGC*
dprsal cds sim2 RP	*GCCACCTTCTCGCAGGCCAGGATGGCCTGGGTGTTGCC*

dorsal k382r FP	*CAAAACATACCGCCCATTCGAACGGAACCAAGAGATAC*
dorsal k382r RP	*GTATCTCTTGGTTCCGTTCGAATGGGCGGTATGTTTTG*
dl cds sim seq RP	*GGCGTGCTGTTGTGGTTGTAG*
ha pet45b+ FP	*TACCCATACGATGTTCCAGATTACGCTtaaCCGCACTCGAGTCTGGTAAAG*
dorsal-ha pet45b RP	*TCTGGAACATCGTATGGGTATCCTCCCGTGGATATGGACAGGTTCGATATCTG*
prm seq FP	*AGCATCTGGCCAATGTGC*
prm his RP	*TCCGTGGTGGTGGTGGTGGTGCATGGTACCGAGCTCGAATTCCC*
prm his dorsal FP	*ATGCACCACCACCACCACCACGGAGGAATGTTTCCGAACCAGAACAATGGAG*
dorsal s389a FP	*CGGAACCAAGAGATACCGCTCCACAACCCTTTGG*
dorsal s389a RP	*CCAAAGGGTTGTGGAGCGGTATCTCTTGGTTCCG*
dorsal prm RP	*AGGTCGACTCTAGAGGATCCTTACGTGGATATGGACAGGTTCGATATCT*
prm FP	*GGATCCTCTAGAGTCGACCTGC*
prm seq RP	*ATTTTTCCATATTTTTCATGGAAAGTTTTA*
dl s312a FP	*CAACTGCGACGTCCCGCAGATGGAGTTACCAGC*
dl s312a s317a FP	*CAACTGCGACGTCCCGCAGATGGAGTTACCGCAGAGGCCCTGCCCTTC*
dl s317a FP	*CTCGGATGGAGTTACCGCAGAGGCCCTGCCCTTC*
dl s312a RP	*CTGGTAACTCCATCTGCGGGACGTCGCAGTTGAATAAAAAC*
dl s312a s317a RP	*GAAGGGCAGGGCCTCTGCGGTAACTCCATCTGCGGGACGTCGCAGTTGAATAAAAAC*
dl s317a RP	*GAAGGGCAGGGCCTCTGCGGTAACTCCATCCGAG*

Table S6: List of primers used for making dsRNA against *cact* and *ter94*

PRIMER NAME	***SEQUENCE (5' - 3')***
cact ds RNA FP	*TAATACGACTCACTATAGGGGAGAACGCTGTGTGCATTTG*
cact ds RNA RP	*TAATACGACTCACTATAGGGGCTTCTCCAGGATGTTCTGC*
ter94 ds RNA FP	*TAATACGACTCACTATAGGGTTCGGTGTCCGATAATGACA*
ter94 ds RNA RP	*TAATACGACTCACTATAGGGGTAAAGATCATCGTCGCCGT*

Chapter 4: Caspar specifies primordial germ cell count and identity in *Drosophila melanogaster*

4.1 Abstract

Repurposing of pleiotropic factors during execution of diverse cellular processes has emerged as a regulatory paradigm. Embryonic development in metazoans is controlled by maternal factors deposited in the egg during oogenesis. Here, we explore maternal role(s) of Caspar (Casp), the *Drosophila* orthologue of human Fas-associated factor-1 (FAF1) originally implicated in host-defense as a negative regulator of NF-κB signaling. Maternal loss of either Casp or it's protein partner, Transitional endoplasmic reticulum 94 (TER94) leads to partial embryonic lethality correlated with aberrant centrosome behavior, cytoskeletal abnormalities, and defective gastrulation. Although ubiquitously distributed, both proteins are enriched in the primordial germ cells (PGCs), and in keeping with the centrosome problems, mutant embryos display a significant reduction in the PGC count. Moreover, the total number of pole buds is directly proportional to the level of Casp. Consistently, it's 'loss' and 'gain' results in respective reduction and increase in the Oskar protein levels, the master determinant of PGC fate. To elucidate this regulatory loop, we analyzed several known components of mid-blastula transition and identify the translational repressor Smaug, a zygotic regulator of germ cell specification, as a potential critical target. We present a detailed structure-function analysis of Casp aimed at understanding its novel involvement during PGC development.

Keywords

Pole cells, Poleplasm , Ter94, Smaug, Oskar, Centrosome

4.2 Introduction

As a model organism, *Drosophila melanogaster*, has been instrumental in establishing and advancing several developmental paradigms underlying embryonic development. *Drosophila* has also emerged as a relatively simple yet tremendously useful model, to analyze the underpinnings of the immune response **(Medzhitov 2001)(Lemaitre and Hoffmann 2007)**. Critically, despite the small size and modest cellular complexity, the insect immune system shares many fundamental traits with the higher vertebrates, including the humoral and innate arms and the dialogue between the two **(Hultmark 1993)(Adams et al. 2000)**. Furthermore, several factors that contribute to the proper functioning of the insect immune system are highly conserved and perform analogous functions across evolution **(Hoffmann and Reichhart 2002)(Lemaitre and Hoffmann 2007).**

Interestingly, almost all the protein components of the insect immune system are highly pleiotropic and involve diverse activities during developmental progression. For example, the Toll class of proteins belonging to a larger family of pattern recognition receptors are essential while mounting robust immune response against the microorganismal invasion **(Brennan and Anderson 2004)(Wang and Ligoxygakis 2006)**. The same proteins are also deployed earlier during embryonic patterning and morphogenesis **(Govind 1999)(Hoffmann and Reichhart 2002)(Valanne et al. 2011).** Evidently, in several instances, different pathway components or modules used during early development are repurposed to mediate immunity both in the insect and the mammalian context **(Belvin and Anderson 1996)(Govind 1999)(Roth 2023)**.

The involvement of early embryonic morphogen, Dorsal, during humoral response in flies is a canonical example of the context-specific and diverse functions of immune system components **(Govind 1999)(Roth 2023)**. Maternal loss of function of genes involved in Toll/Dorsal signaling affect Dorso-ventral patterning **(Belvin and Anderson 1996)**. A nuclear-cytoplasmic gradient of Dorsal, set up by asymmetric Toll signalling establishes cell fate across the Dorsal-ventral axis in the syncytial embryo **(Nusslein-Volhard 2022)**. Pathogenic invasion of insects induces both the humoral and cellular immune response **(Belvin and Anderson 1996)(Williams 2007).** The humoral response consists of the production of antimicrobial peptides by the fat bodies, which serve as a first line of defence **(Belvin and Anderson 1996)(Imler and Hoffmann 2000)**. Two Rel family member proteins, Dorsal and Dif, homologous to the mammalian NF-kappa B, induce the expression of defense peptides **(Belvin and Anderson 1996)(Govind 1999)(Buchon et al. 2014)**. NF-kappa B is essential for differentiating lymphocytes, which engineer the acute-phase response. Altogether, undertaking

functional analysis of proteins, in a temporally distinct developmental context, has proven to be highly informative and insightful. Notably, however, thus far, such analysis has focused on only activator proteins. Here we have investigated the embryonic function of Caspar (Casp), a protein involved in inhibiting the immune response **(Kim et al. 2006).**

Casp is an intracellular negative regulator discovered in a genetic screen to identify suppressors of antibacterial immunity **(Kim et al. 2006).** Flies mutant for casp were identified due to their high rates of melanization, an innate immune response that leads to encapsulation of pathogens in the gut and fat body. Interestingly, casp mutants isolated in this study were resistant to Gram-negative bacterial infections due to elevated expression of the antimicrobial peptide (AMP) diptericin; consequently, infected flies survived longer than their wild-type counterparts. Strikingly, casp overexpression inhibited nuclear localization of Rel in response to infection in the fat body. Excess levels of Casp result in the cytoplasmic retention of Rel in its uncleaved, inactive form, presumably due to inhibition of the protease, Dredd **(Kim et al. 2006).**

Intriguingly, sequence analysis of Casp indicated a high degree of similarity to the mammalian Fas-associated factor 1 (FAF1) protein **(Kim *et al*. 2006)(Tendulkar *et al*. 2022)** (**Fig. 4.1A**). FAF1 is evolutionarily conserved and was initially discovered as an interactor of Fas, a pro-apoptotic member of the tumor necrosis factor receptor family **(Chu *et al*. 1995)**. FAF1 also interacts with the components of the death-inducing signaling complex (DISC) such as the Fas-associated death domain (FADD) and Caspase-8 proteins **(Ryu *et al*. 2003)**. These interactions are mediated by the death effector domains (DED) in FADD and Caspase-8 and the DED-interacting domain (DEDID) in FAF1 **(Ryu *et al*. 2003)**. Both 'loss' and 'gain' of function experiments indicated that FAF1 is crucial in promoting cell death via transduction of the apoptotic signal **(Ryu and Kim 2001)(De Zio *et al*. 2008).** FAF1, like its *Drosophila* ortholog, is a negative regulator of NFκB signaling **(Min-Young *et al*. 2004)(Park *et al*. 2007)**.

FAF1's myriad cellular functions can be attributed to its multiple protein-interaction domains that allow its participation in ubiquitin-related processes such as protein degradation (Song *et al*. 2005; Menges *et al*. 2009; Zhang *et al*. 2011). Consistently, FAF1 harbors a Ubiquitin-associated (UBA) domain, a UAS domain and a Ubiquitin-like regulatory X (UBX) domain (**Fig. 4.1A**). Furthermore, the N-terminal UBA domain recruits polyubiquitinated proteins, leading to their accumulation **(Song *et al*. 2009)**. The C-terminal UBX domain, on the other hand, interacts with the molecular chaperone, valosin-containing protein (VCP/p97) bound to the Npl4-Ufd1 heterodimeric complex **(Schuberth and Buchberger 2008)(Kloppsteck *et al*.**

2012)(Ewens *et al.* 2014). FAF1 complexed with VCP-Npl4-Ufd1 and polyubiquitinated proteins is known to assist endoplasmic reticulum-associated degradation (ERAD) **(Lee *et al.* 2013)**. The UAS domain, a domain of unknown function, is involved in interaction with long-chain fatty acids, which is thought to promote the polymerization of FAF1 **(Kim *et al.* 2013).**

While Casp has a well-established role in the immune response, modENCODE RNAseq and proteomics data suggest that *casp* is also highly expressed in the 0-3 hour old embryo **(Brown *et al.* 2014)(Casas-Vila *et al.* 2017)**. Consistently, a snapshot of *casp* staining in a high-throughput RNA *in-situ* experiment indicates ubiquitous expression of maternally deposited *casp* in the *Drosophila* embryo **(Weiszmann *et al.* 2009).** We thus wondered if Casp expression in early embryos is functionally relevant for proper developmental progression. To explore the possible developmental function of Casp, we first assessed if Casp function is needed for viability. Analysis of a hypomorphic allele of *casp* demonstrated that Casp is indeed required maternally for embryonic development. Consequently, roughly half of the embryos maternally compromised in *casp*, fail to undergo gastrulation. Furthermore, such embryos display developmental defects including aberrant cytoskeletal network starting from early blastoderm stages. Interestingly, Casp is expressed strongly in primordial germ cells (PGCs). Consistent with the enrichment, maternal reduction of *casp* significantly affects the total number of pole cells. Here, we present an analysis of *casp* function during early embryonic development and its role in the formation and/or specification of PGCs. We show that Casp activity regulates Oskar levels and centrosome function, two critical determinants of PGC fate in *Drosophila* embryo. Upon loss of *casp*, the total amount of Oskar and Smaug changed reciprocally to influence the PGC count. Ubiquitin-based protein degradation is critically involved during early embryonic events, including the maternal-zygotic transition. We present a model explaining the involvement of Casp and its protein partner, TER94, during germ cell development, considering their influence on the clearance of Smaug, a critical regulator of maternal to zygotic transition (MZT).

4.3 Results

4.3.1 Caspc^{04227} is a loss of function allele of casp

To better understand embryonic function of Casp we decided to first characterize a *casp* allele, *w^{1118}; pBac(PB)caspc^{04227}* (Bloomington stock number:11373) (Supplementary Fig. 1A; labelled as S1A). This allele is induced by a piggyBac insertion in the 5' regulatory region of *CG8400/FBgn0034068/Casp* locus which is situated at 2R, region 52D14-15 **(Thibault *et al.***

2004). As summarized in **Fig. 4.1** , our data indicate that *caspc*^{04227} is a strong hypomorphic allele and is referred to as a 'loss of function' allele (*casp*^{lof}) here onwards.

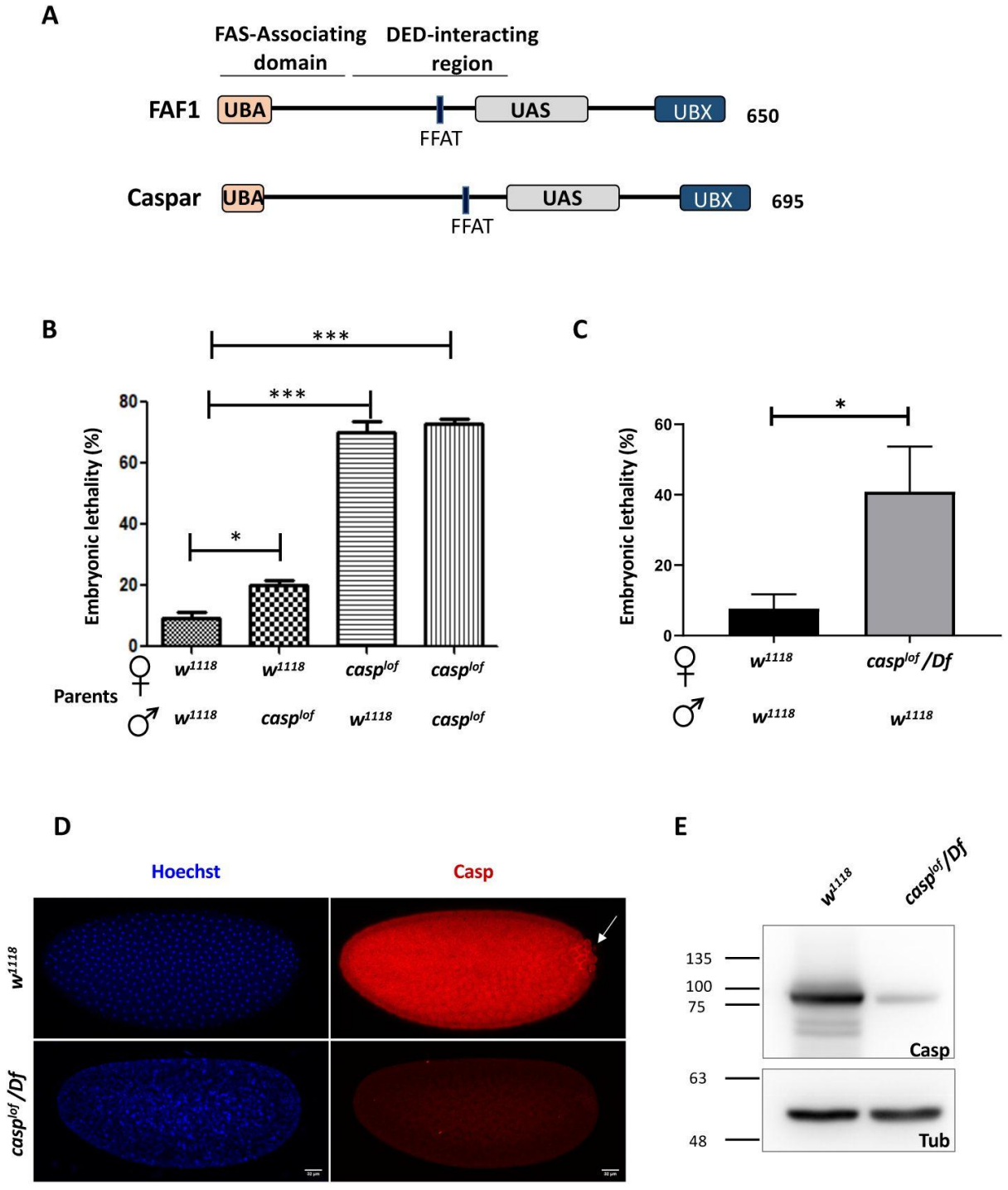


Figure 4.1 *casp* is a maternal effect gene.

(A) Comparison between human FAF1 and Casp shows conserved protein domains, which are described in the text. (**B**) Embryos laid by homozygous $casp^{lof}$ females show ~70% lethality, irrespective of the paternal genotype (w^{1118} or $casp^{lof}$), suggesting a strictly maternal function of *casp* (**C**) The use of a Deficiency in the *casp* locus validates the lethal phenotype associated with the mutation, which drops to ~40%. In panels B and C, the parental genotype is listed on the X-axis, with percent embryonic lethality plotted as a bar graph. N = 3, ordinary one-way ANOVA/ unpaired t-test, (***) P <0.001, (*) P <0.05. (**D**) Immunofluorescence images of 0-3 h embryos derived from w^{1118} and $casp^{lof}/Df$ females, stained with Hoechst and Casp (**E**) Casp protein levels were assessed in 0-3 embryos laid by w^{1118} and $Casp^{lof}/Df$ animals, evaluated via western blotting. Tubulin is used as a loading control.

~70% embryos laid by the homozygous $casp^{lof}$ females failed to hatch into larvae (**Fig. 4.1B**). Moreover, the number of embryos that failed to hatch did not change substantially when $casp^{lof}$ virgin females were mated either with w^{1118} males or $casp^{lof}$ males. That the extent of viability was independent of the paternal genotype established that *casp* is a maternal effect gene and its activity is required for viability (**Fig. 4.1B**). To rule out the possibility that the lethality was induced by non-specific background mutations, we procured multiple deficiency (Df) lines spanning the 52D locus and confirmed that $casp^{lof}/Df$ mothers also displayed significant embryonic lethality (40%) (**Fig. 4.1C**). Lastly, as in the case of $casp^{lof}$ females, $casp^{lof}$ males are also semi-fertile which was assessed by mating the mutant males with w^{1118} virgin females (**Fig. 4.1B**).

To better understand the function of Casp, we first decided to analyze how *casp* RNA and Casp protein are expressed during embryogenesis. Consistent with the maternal effect lethality, modENCODE **(Muers 2011)(Chen *et al.* 2014)** data suggests that *casp* mRNA is deposited maternally, being highly expressed in the 0-2 hours embryo, with a drop in expression 2-4 hours post fertilization (**Fig. S1B**)**(Brown *et al.* 2014)**. Publicly available *In-Situ* databases such as FlyAtlas, Fly-FISH and BDGP **(Weiszmann *et al.* 2009) (Tomancak *et al.* 2002)(Lecuyer *et al.* 2007)** also confirmed maternal deposition of *casp* mRNA, with ubiquitous expression seen in stage 1-3 embryos. Fly-FISH data **(Lecuyer *et al.* 2007)(Wilk *et al.* 2016)** further suggests that expression in the pole cells persists while maternally deposited somatic transcripts are selectively degraded from stage 4 onwards. We immunostained 0-3 hours old embryos derived from w^{1118} mothers with the anti-Casp antibody **(Tendulkar *et al.* 2022)** and found that Casp protein was predominantly cytoplasmic, and as in the case of RNA, exhibited a relatively ubiquitous distribution in the somatic compartment (**Fig. 4.1D**). Again, consistent with the *in-situ* hybridization pattern, Casp-specific antibody staining appeared to be enriched in the posteriorly localized primordial germ cells (PGCs) compared to the

surrounding soma (**Fig. 4.1D** , arrow). To establish the specificity of the anti-Casp antibodies in the embryonic context, we stained the embryos laid by both the $casp^{lof}$/ $casp^{lof}$ and $casp^{lof}$/*Df* mothers. As expected, Casp protein was nearly absent in embryos of both these genotypes suggesting that Casp protein is significantly reduced in the 0-3 hours old $casp^{lof}$ mutant embryos. Analysis of the embryonic lysates using Western blotting, suggested that trace amounts of Casp protein (estimated 5-10% of w^{1118}), were present in embryos laid by $casp^{lof}$/$casp^{lof}$ and $casp^{lof}$/*Df* mothers, supporting the conclusion that $casp^{lof}$ is a strong hypomorphic allele and not a null (**Fig. 4.1E**). Interestingly, mature egg chambers of wild-type animals had Casp expression in somatically derived follicle cells, but Casp protein could not be readily detected in the egg itself (**Fig. S1C**). This observation suggested that, unlike the RNA, only trace amounts of Casp protein may be maternally deposited. Thus, the protein detected in the young embryos is likely generated post-fertilization, by the translation of maternally deposited *casp* mRNA.

4.3.2 Casp function is required for early embryonic development

Larval cuticle patterns are an excellent readout for major patterning defects in *Drosophila (*Nusslein-Volhard and Wieschaus 1980*)*. To better understand the function of Casp protein during embryonic development, we analyzed the cuticles of *casp* mutant embryos. 0-3 hours old embryos laid by $casp^{lof}$/*Df* mothers were collected and aged for 22 hours to prepare cuticles (see Material and Methods). The cuticular patterns were visualized under a dark field microscope. While ~60% of larvae displayed cuticular patterns comparable to wild type larvae, development in the remaining 40% embryos appeared to have stalled before they could deposit cuticle. These data were consistent with the extent of maternal effect lethality.

Next, to define the stage of developmental arrest, we performed time-lapse live Bright-field imaging **(Cavey and Lecuit 2008)**. These data (Movies, Fig. S2A) confirmed that 40% of embryos did not proceed to gastrulation and showed developmental arrest before germ band extension (around stage 6; **Fig. S2** , panels C4, D4). Such embryos displayed irregular, uncoordinated morphogenetic movements with blebbing of the plasma membrane possibly inhibiting both cephalic furrow formation, and germband elongation (Fig. S2 panels C2-C3, D2-D3).

4.3.3 Compromising casp activity leads to centrosomal abnormalities in early embryos

To trace the defects during mid-embryogenesis including failure of gastrulation, we sought to visualize *casp*^{lof} embryos at blastoderm stage. We labelled the 0-3 hours old *casp*^{lof} and control embryos with the nuclear dye Hoechst and the cytoskeletal F-actin marker phalloidin (**Fig. 4.2A,B**). *casp*^{lof} embryos displayed significant structural abnormalities that are reflected in an irregular actin network that lacks stereotypical organization, which is partially interrupted in several places (**Fig. 4.2B** 1, compared to **Fig. 4.2A** 1). Similarly, unlike the age-matched control embryos, regular nuclear spacing is disrupted, and nuclei are unevenly distributed across the embryo with occasional instances of nuclear fusion and possibly mitotic asynchrony (**Fig. 4.2B** vs **Fig. 4.2A**).

The nuclear and cytoskeletal defects observed in *casp* embryos prompted us to analyze the centrosomes in embryos deficient in *casp* function. Centrosomes function as the microtubule organizing centers, and actomyosin-based cytoskeletal defects have been correlated with aberrant centrosome activity **((Wu and Akhmanova 2017)(Blake-Hedges and Megraw 2019)** and references therein). In wild-type nuclei, centrosome duplication occurs simultaneously with the initiation of the nuclear division cycle. After completion of duplication, centrosomes separate and migrate along the nuclear membrane to reach the opposite sides of the nucleus (reviewed by **(Lattao *et al*. 2017)(Wu and Akhmanova 2017)(Blake-Hedges and Megraw 2019)**). In the embryos maternally compromised for *casp*, several characteristic aspects of centrosome behaviour during mitotic divisions are altered (Compare **Fig. 4.2B** 3 to **Fig. 4.2A** 3). At times, centrosomes appeared to divide even without a nucleus (or DNA), and many 'orphan' centrosomes devoid of nuclear DNA were observed (**Fig. 4.2B** 3, inset). By contrast, in some instances, centrosomes were duplicated but remained in proximity indicating failed or incomplete migration (see arrows in **Figure 4.2B** 3, inset). Quantitation of the centrosomal abnormalities in *casp*^{lof} embryos (**Fig. 4.2C, D**) further revealed that the behaviour of the centrosomes, especially their ability to separate and migrate to the opposite poles correctly deteriorated progressively (Compare **Fig. 4.2D** to **Fig. 4.2C**), as nuclear division cycles advanced.

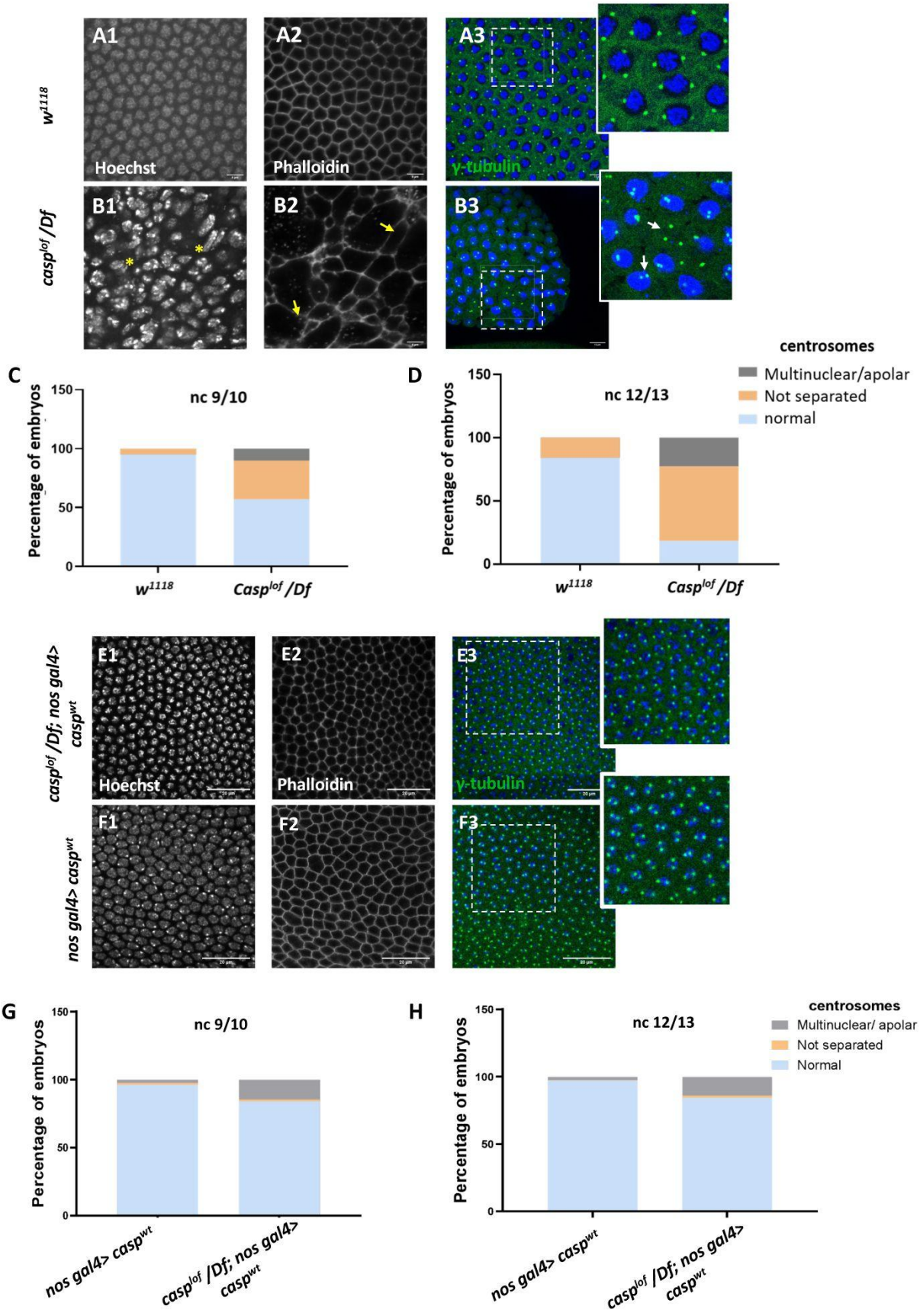

A1
A2
A3
Hoechst
Phalloidin
γ-tubulin
w1118
B1
B2
B3
casplof /Df
C
nc 9/10
Percentage of embryos
150
100
50
0
w1118
Casplof /Df
D
nc 12/13
centrosomes
Multinuclear/apolar
Not separated
normal
E1
E2
E3
F1
F2
F3
casplof /Df; nos gal4> caspwt
nos gal4> caspwt
G
nc 9/10
nos gal4> caspwt
casplof /Df; nos gal4> caspwt
H
nc 12/13
centrosomes
Multinuclear/ apolar
Not separated
Normal

Fig. 4.2 A Significant proportion of embryos (~45%) laid by *casp^{lof}/Df* mothers display nuclear division and cytoskeletal defects.

(**A-B**) Confocal images of nuclear cycle 13/14 embryos (single sections) of the indicated genotypes stained with Hoechst, phalloidin, and gamma-tubulin. Both the regular arrangement and uniform density of nuclei are disrupted in the mutants, as indicated by a yellow asterisk. F-actin, marked with phalloidin, shows a regular, hexagonal compartmentalization in *w^{1118}* (panel A2), while disorganized F-actin (white arrows; panel B3) is observed in the mutant. Defective centrosomes marked with gamma-tubulin can also be observed in the *Casp* mutant (inset of B3, compare to inset A3), indicated by yellow arrows. The extent of defects is quantified in nuclear cycle(nc) 9/10 (**C**) and nc 12/13 (**D**). (**E-F**) Confocal images of sections of nc 13/14 embryos of the indicated genotypes stained with Hoechst, phalloidin, and gamma-tubulin. The nuclear, cytoskeletal, and centrosomal defects are rescued (panel E) when wild-type *casp* is expressed in the *casp^{lof}/Df* background (compare panels **G- H** to C-D). Overexpression of *casp* on its own (panel F) does not appear to affect nuclear or cytoskeletal architecture. The bar charts represent the percentage of defective centrosomes in nuclear cycle 9/10 and 12/13 embryos. Number of embryos imaged ~15.

To confirm the specificity of the phenotypic consequences induced by the maternal loss of *casp*, we overexpressed *UAS-casp^{wt}* transgene in *casp^{lof}/Df* embryos using *nos Gal4*, a maternal Gal4 driver. Embryos derived from such mothers were stained using nuclear dye Hoechst, phalloidin and anti-Gamma tubulin antibodies that mark the centrosomes. As shown in **Fig. 4.2E** , maternal overexpression of *casp* substantially rescues centrosomal and cytoskeletal abnormalities seen in *casp^{lof}* (compare panels E1-E3 to A1-A3). This rescue is quantitated in **Fig. 4.2 G, H** (compare with **Fig. 4.2C, D**). The *casp^{lof}/Df; nos Gal4/UASp-casp^{wt}* showed a rescue of lethality of 20-25% as compared to 40-45% for *casp^{lof}/Df*.

Taken together, these data showed that Casp plays an essential role in early embryonic development in *Drosophila*, and loss of *casp* results in conspicuous nuclear and cytoskeletal defects that correlate with incomplete gastrulation movements and developmental arrest, for ~45% of embryos. Moreover, such defects do not appear to be region-specific or localized within an embryo arguing in support of ubiquitous function across the embryo for Casp protein.

4.3.4 Maternal Casp protein is enriched in pole cells and controls total pole cell count in blastoderm embryos

Primordial germ cells (PGCs) are precursors of the gametes—sperm and eggs. Their specification is essential for the development of the germline, which in turn passes genetic information to the next generation. During early embryonic development, PGCs are specified and segregated from somatic cells. In *Drosophila*, the PGCs are formed in the posterior end of the embryo. As the early embryonic syncytial nuclear division cycles progress, a few nuclei and centrosomes associated with them invade posteriorly localized and anchored pole plasm

ahead of the rest of the somatic nuclei **(Raff and Glover 1989)**. The precocious entry of the centrosomes results in the release and microtubule-dependent transport of pole plasm, which is sequestered in newly cellularized PGCs **(Raff and Glover 1989)(Lerit *et al*. 2017).** Thus, the PGCs or the germline stem cell precursors are set aside during early embryogenesis. Moreover, centrosome behaviour and dynamics are crucial for proper cellularization and mitotic cell divisions of early germ cells **(Lerit and Gavis 2011)(Lerit *et al*. 2017).** Centrosomes, the microtubule (MT)-organizing centers, ensure the faithful segregation of germ plasm, a reservoir of germ cell determinants, into PGCs. Taken together with the observation that Casp protein is readily detectable in early PGCs, we wondered if it could influence the total pole cell count. To examine this possibility, stage 5 embryos derived from *cs* and $casp^{lof}$/ *Df* mothers were immunostained for the pole cell marker Vasa and visualized under a confocal fluorescence microscope. Interestingly, while *cs* and *nos-Gal4* embryos had an average of ∼30 pole cells (**Fig 4.3** , panels A1-B3, for quantitation see panel F) $casp^{lof}$/ *Df* showed a drastic reduction in total PGC count to ∼10 (**Fig. 4.3** panels C1-3, F). By contrast, stage 5 embryos derived from *nosGal4>UAS-casp* mothers, where *casp* was overexpressed (**Fig. 4.3G**), displayed considerably elevated total number of PGCs (**Fig. 4.3** , panels E1-E3, F). To confirm that the loss of *casp* is specifically responsible for the reduction in the total number of PGCs, we simultaneously overexpressed *casp* in the $casp^{lof}$/*Df* mothers in a germline-specific manner, which resulted in a significant rescue (∼26 pole cells; **Fig. 4.3** , panels D1-D3, F). Taken together, these data suggested that the total pole cell number in late syncytial /cellular blastoderm embryos specifically depends on *casp* function. Furthermore, increasing *casp* levels maternally, is sufficient to elevate the total pole cell number substantially.

4.3.5 Casp interacts with TER94 in the early embryo

Data presented in the previous section demonstrated that both the 'loss' and 'gain' of *casp* activity exert reciprocal influence on total PGC numbers in early embryos. As very few germ plasm components have been shown to display this trait **(Jongens *et al*. 1994)**, we decided to explore the phenomenon further. We have previously reported that Casp protein physically associates with both transitional endoplasmic reticulum ATPase (TER94; also called valosin-containing protein, VCP or p97) and vesicle-associated membrane protein-associated protein B (VAPB/VAP33A) **(Tendulkar *et al*. 2022).** Furthermore, based on interaction studies and biochemical analysis, we proposed that Casp may act as an adapter that either directly or indirectly mediates the physical association between VAPB and TER94 **(Tendulkar *et al*.**

2022). To arrive at this conclusion, we had primarily relied on S2 cell lysates as well as adult and fly head total protein extracts. To specifically identify the major protein partners of Casp in the early embryos, we performed similar immunoprecipitations using anti-Casp antiserum in early (0-3 hour) embryonic lysates and interactors were identified via mass spectrometry (**Fig. S3A**). Peptides from 122 proteins were recovered in the Casp IP but not the IgG IP and thus, were considered significant interactors. The top 23 interactors are presented in **Fig. S3A** . Interestingly, as in the case of brain and S2 cell lysates, TER94 and VAPB were the top hits, with germ cell determinants also identified as interactors (**Fig. S3C** ; Discussed in a subsequent section). These observations suggested that a functional protein complex between Casp, TER94 and VAPB likely exists in many different tissue/cellular contexts, including early embryos.

4.3.6 Maternal requirement of TER94

Mammalian VCP/p97 is an essential chaperone for proteostasis that modulates several ubiquitin- associated processes **(Peters *et al*. 1990)(Dai *et al*. 1998)(Meyer *et al*. 2000)(Meyer 2005)(Ye 2006)(Jentsch and Rumpf 2007)(Meyer *et al*. 2012).** Its *Drosophila* counterpart TER94 has been studied mostly in the context of neurodegeneration **(Griciuc *et al*. 2010)(Chang *et al*. 2011)(Azuma *et al*. 2014)(Kushimura *et al*. 2018)(Tendulkar *et al*. 2022)(Thulasidharan *et al*. 2024).** Loss of TER94 was shown to ameliorate polyQ-induced eye degeneration. Moreover, the overexpression of TER94 promoted the apoptosis of neuronal cells **(Higashiyama *et al*. 2002).** TER94 has been uncovered in a genetic screen for maternal proteins that are regulated by phosphorylation **(Zhang *et al*. 2018**) and have roles in oogenesis and early development.

TER94 is one of the significant interactors of Casp protein in the embryonic context as suggested by the Casp interactome (**Fig. S3A**) and has been implicated in ER-associated degradation **(Chang *et al*. 2011)(Tendulkar *et al*. 2022).** Thus, we wondered if TER94 also has a maternal function and whether embryos maternally compromised for *TER94* display similar phenotypes as *casp*. To achieve this, we employed VALIUM 20/22 maternal RNAi lines (*UAS-TER94i*). First, we used a *mat-α4tubulin:VP16-Gal4* (*Mat-αtubGal4*) driver to deplete *TER94* activity in the late stages of oogenesis (**Fig. 4.4**).

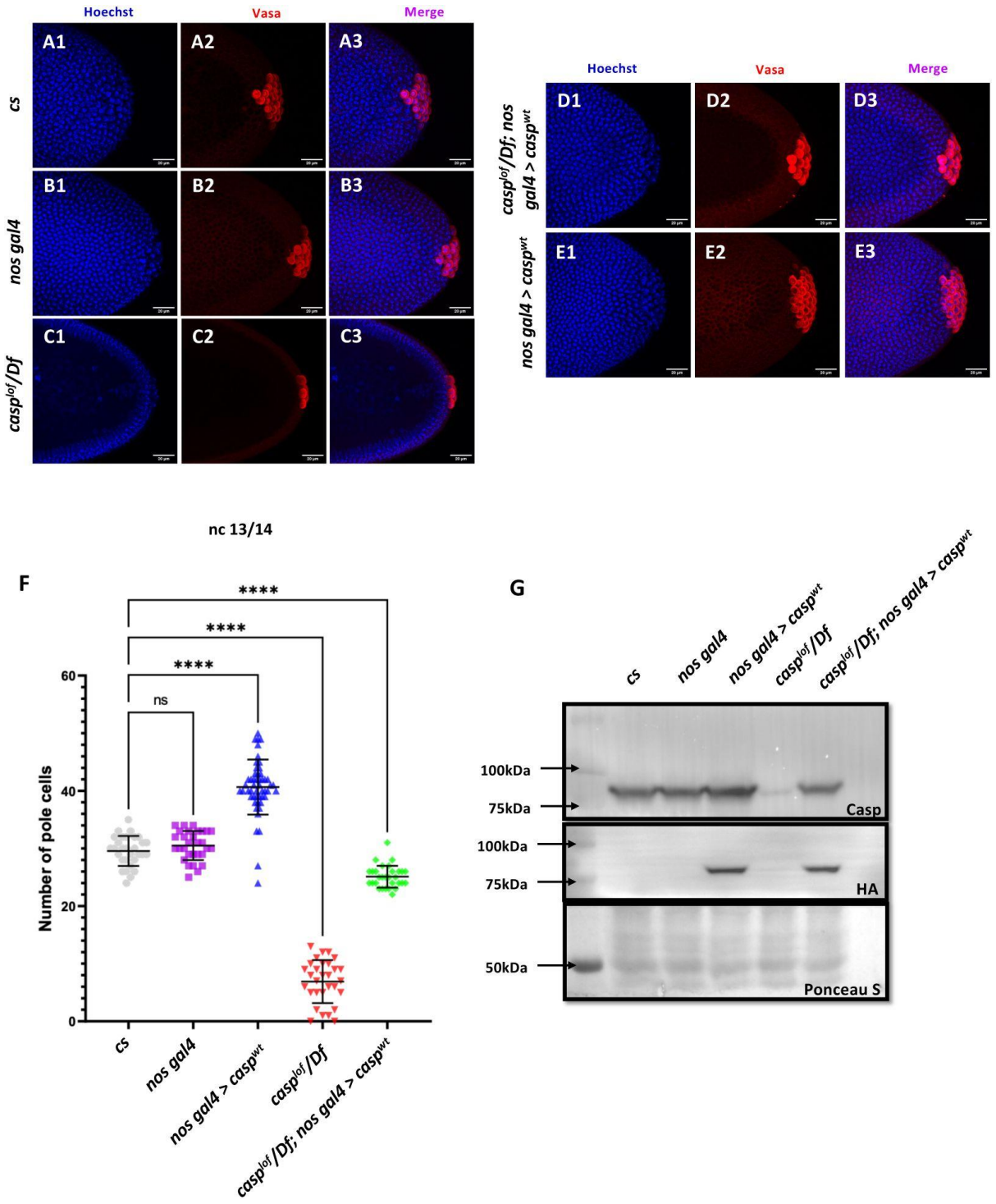


Fig. 4.3 Casp influences total PGC count.

(**A-E**): Shown are the Confocal images of the posterior terminii of nuclear cycle 13/14 embryos of the following genotypes: *cs* (A1-A3), *nos gal4* (B1-B3), *casp^{lof}/Df* (C1-C3), casp^{lof}/*Df*; *nos gal4>casp^{wt}* (D1-D3) and *nos gal4> casp^{wt}* (E1-E3). Embryos were immunostained with Vasa (1:50) antibody. Hoechst marks the nuclei. (**F**) The total number of germ cells marked with Vasa were quantified and plotted as bar graphs. N (embryos)= 30. Ordinary one-way ANOVA, (****) P <0.0001. (**G**) Casp protein levels were assessed in 0-3 embryos via western blotting. Rabbit anti-Casp (1:10,000) and rabbit anti-HA (1:2000) antibodies were used to probe the blot. Ponceau is used as a loading control. N = 3.

We sought to observe the consequence of *TER94* knockdown both in the adult females and early embryos derived from such females. In the females compromised for *TER94*, egg-laying behavior remained largely unaffected. Interestingly, however, the viability of such embryos was severely impaired, with >95% of embryos failing to hatch (**Fig 4.4A**), in agreement with earlier studies **(Zhang *et al*. 2018).**

A western blot of lysates derived from *Mat-GAL4>UAS-TER94i* mothers probed with anti-TER94 antiserum indicated a robust knockdown of TER94 as compared to control (**Fig. 4.4B**). Phenotypic analysis of embryos derived from *Mat-GAL4> TER94i* females revealed that >70% of *Mat-GAL4/ TER94i* embryos failed to progress to the syncytial blastoderm stage, with even fewer reaching Stage 5. The late-stage syncytial/ cellular blastoderm embryos were further assessed for cell cycle defects. As in the case of $casp^{lof}/Df$, *Mat-GAL4>TER94i* embryos also displayed defects ranging from irregular nuclear distribution to perturbed F-actin localization, as assessed by HOECHST and phalloidin staining, respectively (Compare **Fig. 4.4D1-D3** with **Fig. 4.4C1-C3**).

We also examined if maternal loss of *TER94* activity results in centrosome aberrations like those observed in $casp^{lof}$ embryos. Indeed, embryos deficient in *TER94* function showed both the characteristic phenotypes observed in the *casp* embryos including a) inefficient separation of duplicated centrosomes and b) multiple instances of 'orphan' centrosomes devoid of nuclear DNA (**Fig. 4.4E, F**). An earlier study **(Zhang *et al*. 2018)** had found that TER94 RNAi leads to multi- polar spindles with supernumerary centrosomes.

Taken together, these data demonstrated that embryos maternally compromised for either Casp or TER94 functions share several phenotypes, suggesting that both the proteins likely perform essential and likely related functions during syncytial nuclear cycles in *Drosophila* embryos. Furthermore, the phenotypic consequences observed upon maternally compromising *casp* and *TER94* overlap but are not identical, with TER94 embryos showing higher penetrance in terms of phenotypes and embryonic lethality. Future experiments will be necessary to resolve the functional distinction between the two.

4.3.7 TER94, a known component of pole plasm, is detectible in PGCs

Maternal loss of TER94 function mimicked both the cytoskeletal and corresponding centrosome aberrations observed upon similar loss of *casp* activity. Both centrosomes and germ

plasm, are essential for the proper specification and formation of PGCs in *Drosophila* embryos **(Lerit and Gavis 2011)(Lerit *et al*. 2017).** This prompted us to examine the Casp interactome for the possible enrichment of germ plasm components and proteins that may regulate centrosome dynamics and/or behaviour. **Fig. S3B** lists Casp interactors that are important constituents of the germplasm, including eIF4A, me31B, and TER94, known constituents of the polar granules **(Thomson *et al*. 2008).**

Among the different components of the germ plasm, Oskar serves as the principal determinant of the PGC fate **(Kim-Ha *et al*. 1991)(Snee and Macdonald 2004)(Vanzo *et al*. 2007)(Lehmann 2016**). Supporting the conclusion, the loss and gain of function of *oskar* leads to reciprocal phenotypes. Compromising *oskar* activity maternally, leads to reduction in total number of PGCs whereas anterior ectopic expression of *oskar* using the *bicoid* mRNA localization signal induces pole cell formation at the anterior **(Ephrussi and Lehmann 1992).** The germ plasm is supplemented with mitochondria and polar granules, which consist of ribonucleoprotein complexes. Downstream of *oskar*, *vasa* and *tudor* are two genes that are essential for the assembly of pole plasm **(Breitwieser *et al*. 1996)(Arkov *et al*. 2006)(Jones and Macdonald 2007).** Biochemical proteomic analysis of Vasa (VAS) and Tudor (TUD) containing polar granule complexes identified eIF4A, me31B, and TER94. This indicated that the germ plasm consists of the components of translational machinery and the endoplasmic reticulum assembly **(Thomson *et al*. 2008).**

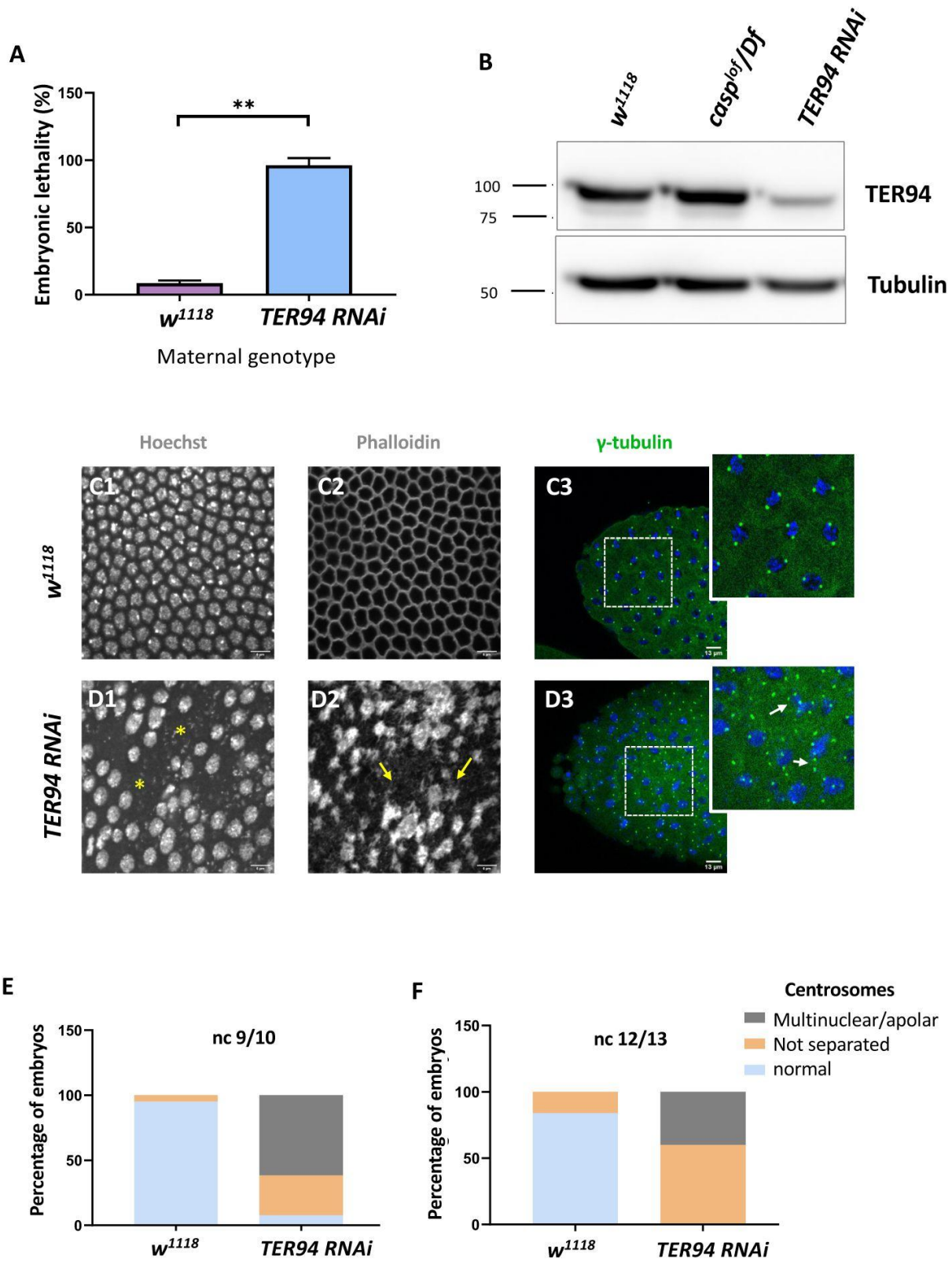


Fig. 4.4 Maternal requirement of *TER94* during early embryogenesis.

(**A**) Viability of embryos derived from w^{1118} and *Mat-αtubGAL4 > TER94i* (referred to as *TER94 RNAi* henceforth) mothers is represented as a bar graph. N = 3, ordinary one-way ANOVA, (***) P <0.001. (**B**) Western blot analysis indicates the knockdown efficiency of TER94 in the 0-3 h embryo, estimated to be ~90%., levels remain unaffected in $casp^{lof}/Df$ embryos. Tubulin was used as a loading control. N = 3. (**C, D**) Confocal images of sections of nuclear cycle 13/14 embryos of the indicated genotypes

stained with Hoechst, phalloidin, and gamma-tubulin. Nuclei and F-actin are visualized with Hoechst and phalloidin respectively. Nuclear disruption (D, asterisks) and F-actin aggregates (D2, arrows) are observed in the mutant. Defective centrosomes (D3, arrows, inset) can be observed in the *TER94 RNAi* embryos. (**E, F**) The percentage of defective centrosomes in nuclear cycles 9/10 and 12/13 are represented as bar charts. n = 10.

TER94 is a component of pole plasm and is also physically associated with Casp. Furthermore, proteomic analysis indicated that Casp could be part of a complex comprising several germ plasm proteins including Tudor and Me31B. To confirm that, as in the case of Casp, TER94 is also detected in PGCs. We used antibodies generated against Casp (**Fig. S3D**) and TER94 (**Fig. S3E**) to label wild-type embryos. Embryos were also co-immunostained for germ cell specific marker, Vasa. As can be seen, both Casp and TER94 proteins are found in PGCs, colocalized with Vasa.

4.3.8 Casp and TER94 regulate embryonic germ cell formation

As we were specifically interested in investigating possible similarities and distinctions between the respective functions of TER94 and Casp during early embryonic germ cell development, we decided to examine their possible functions during PGC formation (also referred to as pole cell budding). This seemed especially pertinent as centrosome behavior and dynamics are crucial for proper formation and cellularization of PGCs. Evidently, a germ plasm component, Germ-cell-less **(Cinalli and Lehmann 2013)** was shown to be necessary for proper separation of daughter centrosomes in the dividing pole buds **(Lerit *et al*. 2017).** Similarly, loss of centrosome components such as Centrosomin results in partial loss of early embryonic PGCs. As maternal loss of both *casp* and *TER94* results in aberrant centrosomes in the surrounding somatic nuclei in early blastoderm embryos, we sought to investigate if posteriorly positioned centrosomes in the vicinity of pole plasm also display similar problems. Indeed, maternally compromising *casp* led to significant loss of pole buds and corresponding defective centrosome behaviour (**Fig. 4.5**).

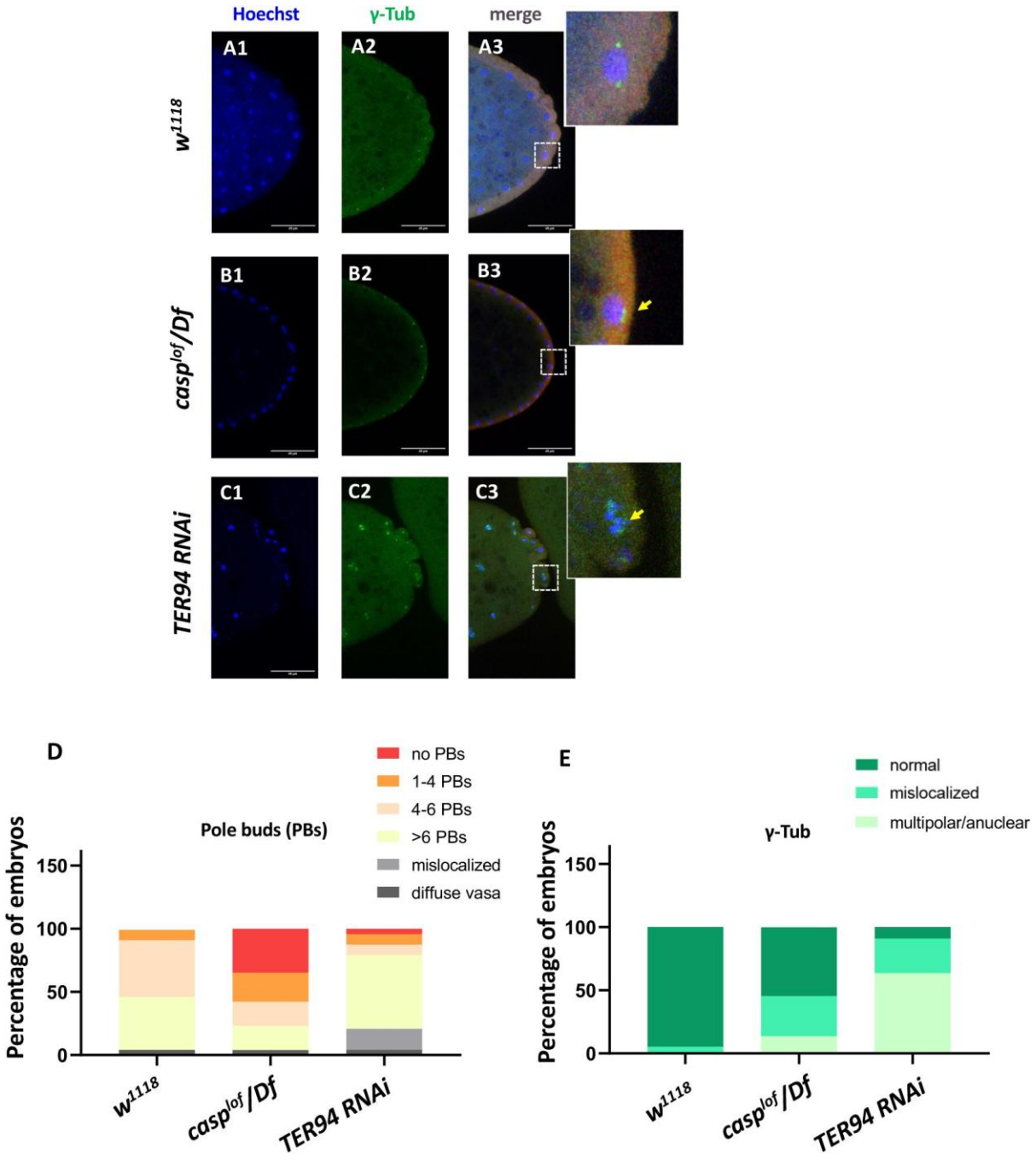


Fig. 4.5 *casp and TER94* regulate pole bud formation.

(A-C) Confocal sections of the posterior termini of nuclear cycle 9/10 embryos derived from w^{1118} (A1-A3), $casp^{lof}/Df$ (B1-B3), and *TER94 RNAi* (C1-C3) mothers immunostained with gamma-tubulin. ψ-tubulin labels centrosomes of pole buds, identified based on Vasa staining. Nuclei are stained with Hoechst. The inset(s) for B3 and C3 highlight the centrosomal defects (arrow) when compared with A3. Defects were quantified and plotted as bar graphs in (D) and (E) respectively. n = 20. oocyte, the defects in pole cell budding including their ectopic positioning could be a consequence of inappropriate segregation of *oskar*. Moreover, the early nuclear division cycles and nuclear migration defects seen in are defective in *TER94i* embryos could also contribute to PGC formation/ cellularization. Notably, although the pole bud count is elevated in *TER94i* embryos, presumably, these pole buds don't survive through subsequent nuclear cycles, and consequently, late blastoderm *TER94i* embryos almost completely lack PGCs.

Especially these pole buds display defective centrosome separation and/or orphan centrosomes (**Fig. 4.5B1-B2,D,E**), as observed in the somatic nuclei. Intriguingly, while *casp*^{lof} embryos show discernible loss of pole buds, *TER94i* embryos displayed an increased number of round-shaped, 'bud-like' cells at the posterior of the stage 3 embryos (**Fig 4.5 C1-C2**). Typically, such ectopically localized pole buds also showed inadequate separation of centrosomes and could be identified as pole buds due to the enrichment of Vasa. However, Vasa distribution and accumulation in *TER94i* buds was variable and non-uniform. Since TER94 is required for *Osk* mRNA localization in the oocyte, the defects in pole cell budding including their ectopic positioning could be a consequence of inappropriate segregation of *oskar*. Moreover, the early nuclear division cycles and nuclear migration defects seen in *TER94i* embryos could also contribute to PGC formation/ cellularization defects. Notably, although the pole bud count is elevated in *TER94i* embryos, presumably, these pole buds don't survive through subsequent nuclear cycles, and consequently, late blastoderm *TER94i* embryos almost completely lack PGCs.

4.3.9 Casp activity is needed for the accumulation of Oskar protein at the embryonic posterior pole

Taken together, our data demonstrate that change in Casp levels leads to corresponding alteration in the total PGC count. Moreover, PGC formation is affected by maternal loss of Casp, although the precise nature of its involvement in this process remains to be determined. As Oskar is the master determinant of germ cell fate, we sought to determine if Oskar levels (Kim-Ha *et al*. 1991 ; Ephrussi and Lehmann 1992) are correspondingly altered upon change in Casp. To assess this possibility, we stained wild-type control embryos and embryos maternally compromised for *casp* using anti-Oskar antibodies (Compare **Fig 4.6B to 4.6A**). Simultaneously, we stained embryos derived from *nosGal4>UASp-casp* mothers (**Fig 4.6C1-C3**) that display significantly elevated number of PGCs, presumably due to excess levels of Casp protein in the germline.

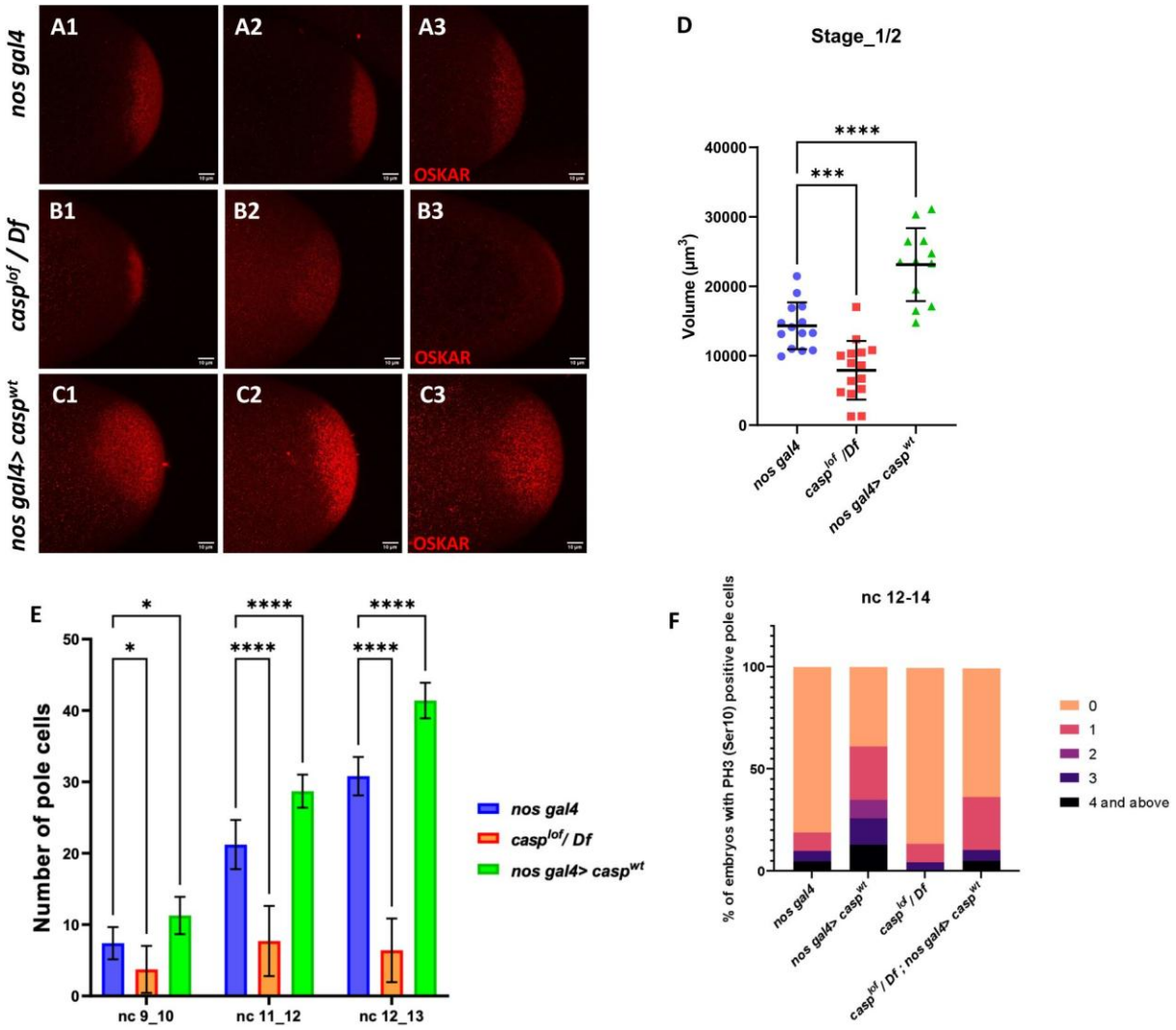


Fig. 4.6 Casp regulates Oskar levels and modulates cell division

(A-C) Confocal microscopy images of stage 1/2 embryo laid by w^{1118}, $casp^{lof}/Df$, and *nos gal4> casp*^{wt} females and immunostained with Oskar antibodies. Replicate (1-3) are shown to highlight variable reduction in Oskar levels. In addition, as compared to controls (w^{1118}) embryos (A1-A3), spread of Oskar appears to be restricted in $casp^{lof}/Df$ (B1-B3) whereas it is expanded in *nos gal4> casp*^{wt} (C1-C3). (**D**) Volume occupied by Oskar, per embryo is measured and plotted for all three genotypes (w^{1118}, $casp^{lof}/Df$, and *nos gal4> casp*^{wt}), n = 8. Ordinary one-way ANOVA, (**) P <0.01, (*) P <0.05. (**E**)Number of pole buds (nc 9/10) and pole cells (nc 11/12, nc 12/13) in 0-3h old *nos gal4*, $casp^{lof}/Df$ *and nos gal4> casp*^{wt} embryos at different nuclear cycles. Two-way ANOVA, (*) P <0.05, (****) P <0.0001. n = 10 (**F**) Graph showing the distribution of actively dividing pole cells as indicated by the presence of phospho-histone 3 (Ser10) antibody. Quantitative analysis was performed using nuclear cycle 12-14 in 0-3h old, *nos gal4*, $casp^{lof}/Df$, *nos gal4> casp*^{WT} and $casp^{lof}/Df$; *nos gal4> casp*^{WT} embryos.

Satisfyingly, the change in Oskar protein levels is consistent with the alteration in the Casp levels in the maternal germline (**Fig 4.6D**). Embryos derived from *casp* mutant mothers show

considerably diminished levels of Oskar as compared to the control whereas overexpression of Casp in the maternal germline results in elevation in the Oskar protein amount (**Fig 4.6D**). Furthermore, in all instances, Oskar protein appears to be anchored to the posterior pole as the wild-type control embryos and the only discernible change is observed in its levels.

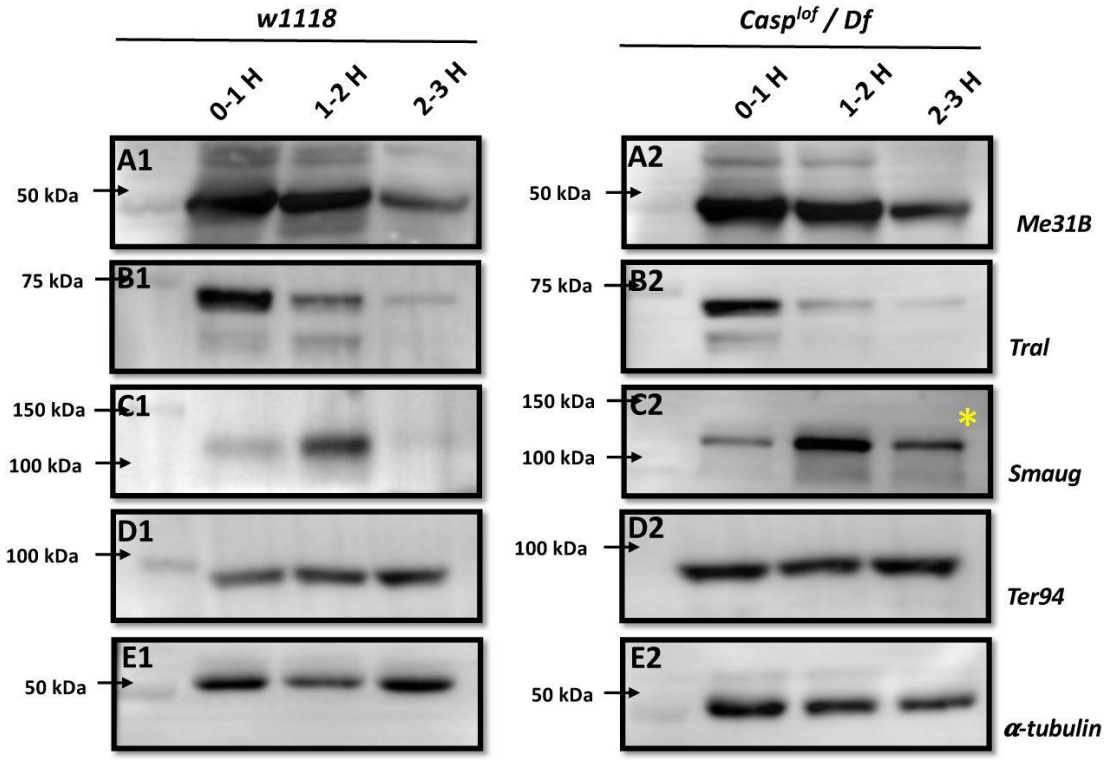
4.3.10 Casp levels influence total number of phospho- histone3 (pH3) positive pole buds and PGCs

Changes in Casp protein levels influence the accumulation of Oskar protein. Oskar is necessary and sufficient to assemble pole plasm at the posterior pole which controls the total number of pole buds in an embryo. To directly evaluate the influence of Casp on pole bud formation, total number of pole buds and pole cells were quantitated upon maternal loss and gain of Casp against the control embryos at the same stage (**Fig. 4.6E**). Consistent with the data presented in the previous sections, total number of pole buds (nc 9 and 10) were reduced in *casp*^{lof} embryos whereas the number was elevated in the presence of excess Casp (**Fig 4.6E**).

Similar quantitation was performed in older embryos in two stages (nc11-12 and NC12-13), and a progressive increase was observed in the total number of PGCs upon the gain of Casp function (**Fig. 4.6E**). By contrast, the total PGC count for *casp*^{lof}*/Df* embryos did not increase appreciably at nc11-12 and nc12-13. Primordial germ cell number is determined by limited mitotic cell divisions that each pole bud undergoes in a stochastic and asynchronous manner. As these are mitotic divisions, the dividing PGCs can be identified using antibodies against the miotic marker phosphoHistone3 (pH3). As the total number of pole buds and PGCs change per Casp levels, we wanted to examine if this is also reflected in the total number of pH3-positive germ cells simultaneously labelled with the anti-Vasa antibodies. We decided to focus on syncytial blastoderm embryos between NC12-14. As anticipated, maternal overexpression of Casp led to a corresponding increase. In control (*nos gal4*) embryos, roughly 20% PGCs display either 1 or >1 pH3 positive PGCs whereas, 70% of *nos-gal4>Casp*^{WT} embryos show 1 or >1 pH3 positive PGCs (**Fig. 4.6F**). Conversely, less than 10% of *casp*^{lof} embryos showed either 1 or >1 pH3 positive PGCs and pole cells with 3 or >3 pH3 positive cells were completely absent in this background as opposed to 25% present in the control embryos (**Fig. 4.6F**). Importantly, maternal expression of the rescue construct in the *casp*^{lof} embryos ameliorated the loss of pH3 positive PGCs seen in just the mutant (*casp*^{lof}) PGCs.

4.3.11 Does Casp function affect canonical MBT regulators?

Our analysis thus far has revealed two related yet distinct phenotypes associated with maternal loss of *casp*. First, it can influence proper PGC formation and specification during early embryonic development via its effect on the accumulation of germ cell determinant, Oskar. Subsequently, it helps orchestrate cellular movements leading up to gastrulation during mid-embryogenesis. While the first activity is likely germ cell autonomous, the second relates to its role in the somatic cells/nuclei. Moreover, both the somatic and germline compartments of $casp^{lof}$ embryos share centrosomal as well as cytoskeletal aberrations. We thus wondered whether the two activities are mechanistically connected, and if the possible connection relates to the maternal-to-zygotic transition (MZT). One aspect of the MZT that has been highlighted in recent years is the active degradation of maternal proteins **(Cao *et al*. 2020)**. These proteins comprise 2% of the maternal proteome and are degraded abruptly, at the end of the MZT. In *Drosophila*, ubiquitin-proteasome- based degradation of three repressor proteins, namely Smaug (Smaug), Trailer hitch (Tral), Maternal expression at 31B (Me31B), marks the MZT **(Cao *et al*. 2020).** We thus decided to probe, using antibodies generated against these proteins (**Fig. 4.7**) if the degradation of any of these proteins is affected due to maternal loss of Casp/*casp* during the initial hours of embryonic development (0-1, 1-2 & 2-3 hours). TER94 protein is unchanged (Fig. D1, D2) as is the α-tubulin loading control (**Fig. 4.7E1** , **E2**). Furthermore, pattern of degradation of Me31B and TRAL proteins in the $casp^{lof}$ embryos is relatively unaffected, when compared to the w^{1118} control samples.


w1118
Casp^lof / Df
0-1 H
1-2 H
2-3 H
A1
A2
50 kDa
Me31B
B1
B2
75 kDa
Tral
C1
C2
150 kDa
100 kDa
*
Smaug
D1
D2
Ter94
E1
E2
α-tubulin

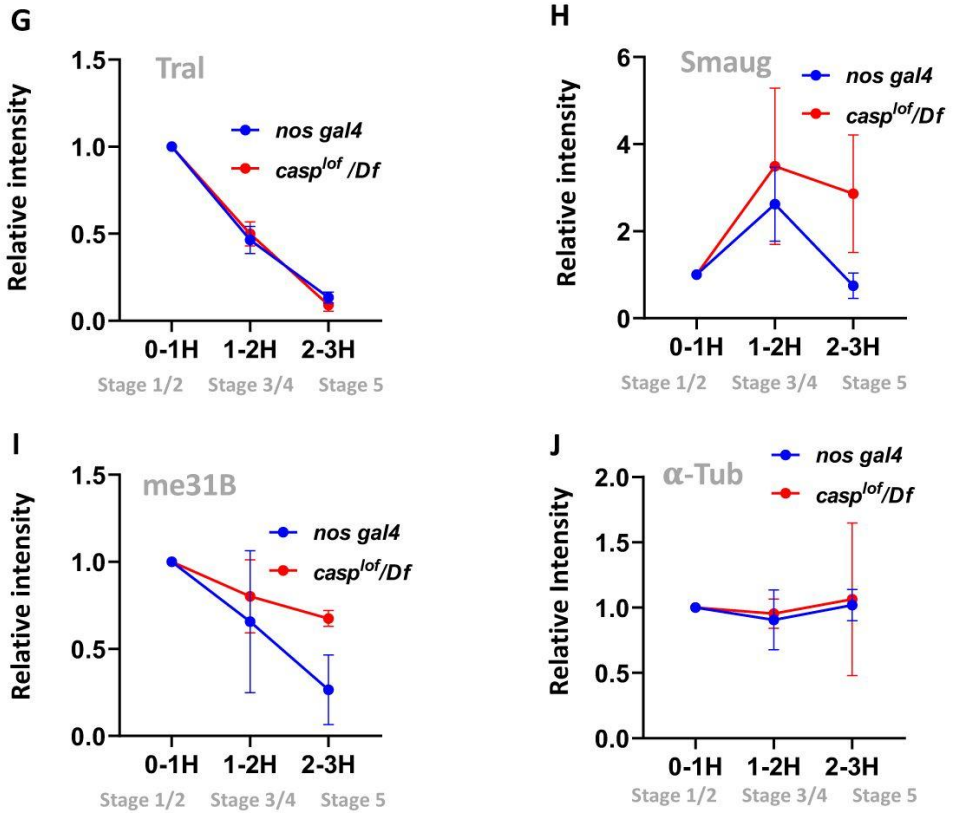

G
Tral
Relative intensity
nos gal4
casp^lof /Df
0-1H
1-2H
2-3H
Stage 1/2
Stage 3/4
Stage 5
H
Smaug
I
me31B
J
α-Tub
Relative Intensity

Fig. 4.7 Reduction in Casp activity specifically affects Smaug degradation during the MZT.

At the MZT, a few maternal proteins are actively degraded (C_{AO} *et al*. 2020). Embryos from mothers with genotype w^{1118} **(A1-E1)** and $casp^{lof}/Df$ **(A2-E2)** were collected using three time intervals (0-1, 1-2, 2-3 Hrs), and embryonic lysates were separated on SDS-PAGE gels. The Western blots were probed with antibodies against ME31B (A1-A2), Tral (B1-B2) and Smaug (C1-C2) to assess the extent of protein degradation, as part of the MZT. α-tubulin (E1-E2) was used as a loading control, while TER94 (D1-D2) was used as a negative control as it is a maternal protein that does not undergo degradation at MZT. (**G-J**) Quantitation of the intensity of change of protein band (normalized to 1), for Tral (G), Smaug (H), ME31B (I) and tubulin (J). Each data point is an average of over 5 western blots. The intensities at each time point for *nos gal4* vs $casp^{lof}/Df$ are statistically not significant.

4.3.12. Germ cell specific Smaug levels are influenced by casp activity

Since Smaug levels appear to be elevated in embryonic lysates in the MZT, we decided to test if the Smaug levels are elevated specifically in $casp^{lof}/Df$ pole cells. To this end, pre-syncytial as well as early syncytial blastoderm embryos of both the genotypes (control and $casp^{lof}/Df$) were stained with anti-Smaug and anti-Vasa antibodies. Aligned with the Western blot data presented earlier, levels of Smaug protein are significantly elevated in pre-syncytial embryos (**Fig. 4.8** , compare panel B3 with A3). Also, in $casp^{lof}/Df$ embryos, Smaug protein seems to accumulate in discernible large puncta which are barely visible at this stage in control embryos (**Fig. 4.8** , B3 vs A3). At the syncytial blastoderm stage (Stage 4), however, the Smaug positive puncta appear larger and more numerous in the few surviving PGCs from the $casp^{lof}/Df$ embryos (**Fig. 4.8** , D3) when compared to controls (C3).

Taken together, our data suggest that Smaug protein levels are appreciably elevated in pole cells from $casp^{lof}/Df$ embryos. The specific increase in Smaug levels may, in part, be due to inappropriate accumulation and/or stabilization of Smaug. Smaug protein is necessary for translational control of the posterior determinant *nanos*. Early reports indicated that unlocalized *nos* RNA is translationally repressed by Smaug which binds to Smaug Response Elements (SREs) within 3'UTR of *nos* RNA. In addition to *nos*, it also regulates *hsp83* translation.

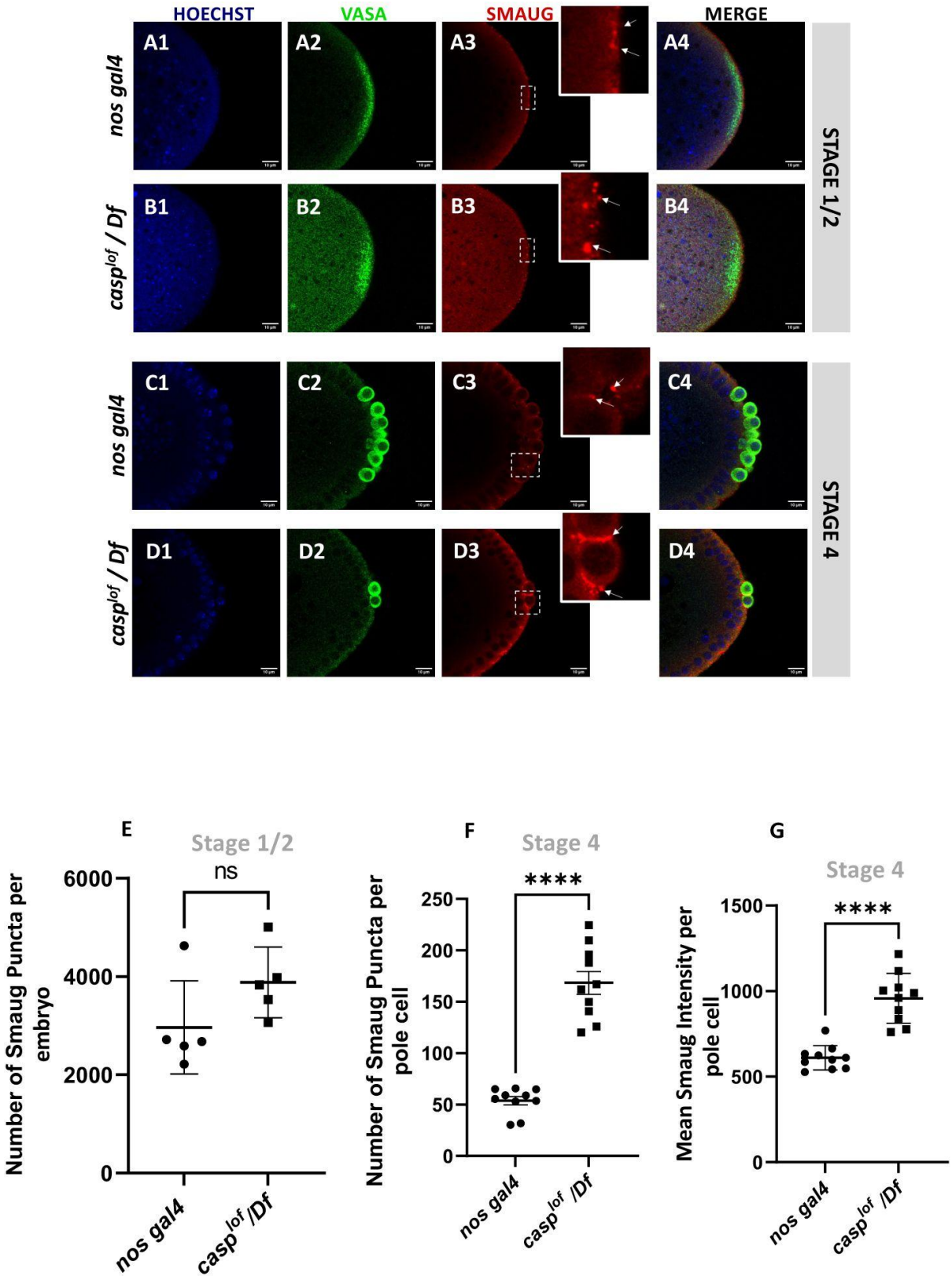


Fig. 4.8 Decrease in Casp correlates with an increase in Smaug levels in the pole buds and PGCs.

Confocal microscopy images of the posterior termini of stage 1-2, and 4 embryos derived from *nos gal4* and *casp^{lof}/Df* females. Embryos were immunostained with Vasa (green) and Smaug (red) antibodies. Hoechst marks the nuclei (A-D). Smaug expression was quantified across the two genotypes and plotted as bar graphs (E-G). N=5 embryos (Stage 2) and 10 embryos (Stage 4).

Smaug participates in multiple, overlapping mechanisms including interaction with the components of the translation machinery as well as deadenylation to regulate translation/localization of the target RNAs **(Dahanukar *et al*. 1999)(Zaessinger *et al*. 2006).**

By contrast, Smaug **(Fig 4.8, C1-C2)** is an interesting exception. In control, 0–1hr old embryos, low levels of Smaug protein were observed, which increased in the 1-2hr old embryo extract but decreased the 2-3 hr time-window (**Fig. 4.7C**), data that agrees with previously published work **(Cao *et al*. 2020).** However, in the $casp^{lof}/Df$ embryos levels of Smaug protein are modestly elevated from the start (0-1 and 1-2 hr old extract samples respectively; **Fig. 4.7C** 2). Critically, unlike control embryos, Smaug protein persists in 2-3hr old $casp^{lof}/Df$ embryonic extracts (**Fig 4.7C** 2, *). Altogether, these data argue that Smaug degradation is specifically adversely influenced in the $casp^{lof}/Df$ embryos at the end of the MZT.

Intriguingly, recent data from Lipshitz lab has, in fact, shown that Smaug protein accumulates in germ granules **(Siddiqui *et al*. 2023)**. Moreover, in addition to its canonical role in regulating *nos* translation, it can also repress *oskar* at the translational level (see Discussion). As the embryonic PGC count depends upon Oskar levels, *Smaug* mutant embryos show an increase in the total number of PGCs. These data fit nicely with our observations and are entirely consistent with a model incorporating the downregulation of Smaug mediated by Casp.

4.3.13 Functional analysis of different protein domains within Casp

To decipher the specific functions of the individual protein domains of Casp during pole cell formation and division we followed an experimental strategy involving rescue of the loss of function phenotype. We compared the extent of rescue observed in the presence of individual deletion constructs to the full length wild-type control. To this end, fly lines expressing domain deletions of Casp under the *pUASp* promoter were generated (**Fig. 4.9A**). The $pUASp\text{-}casp^{\Delta UBA}$,

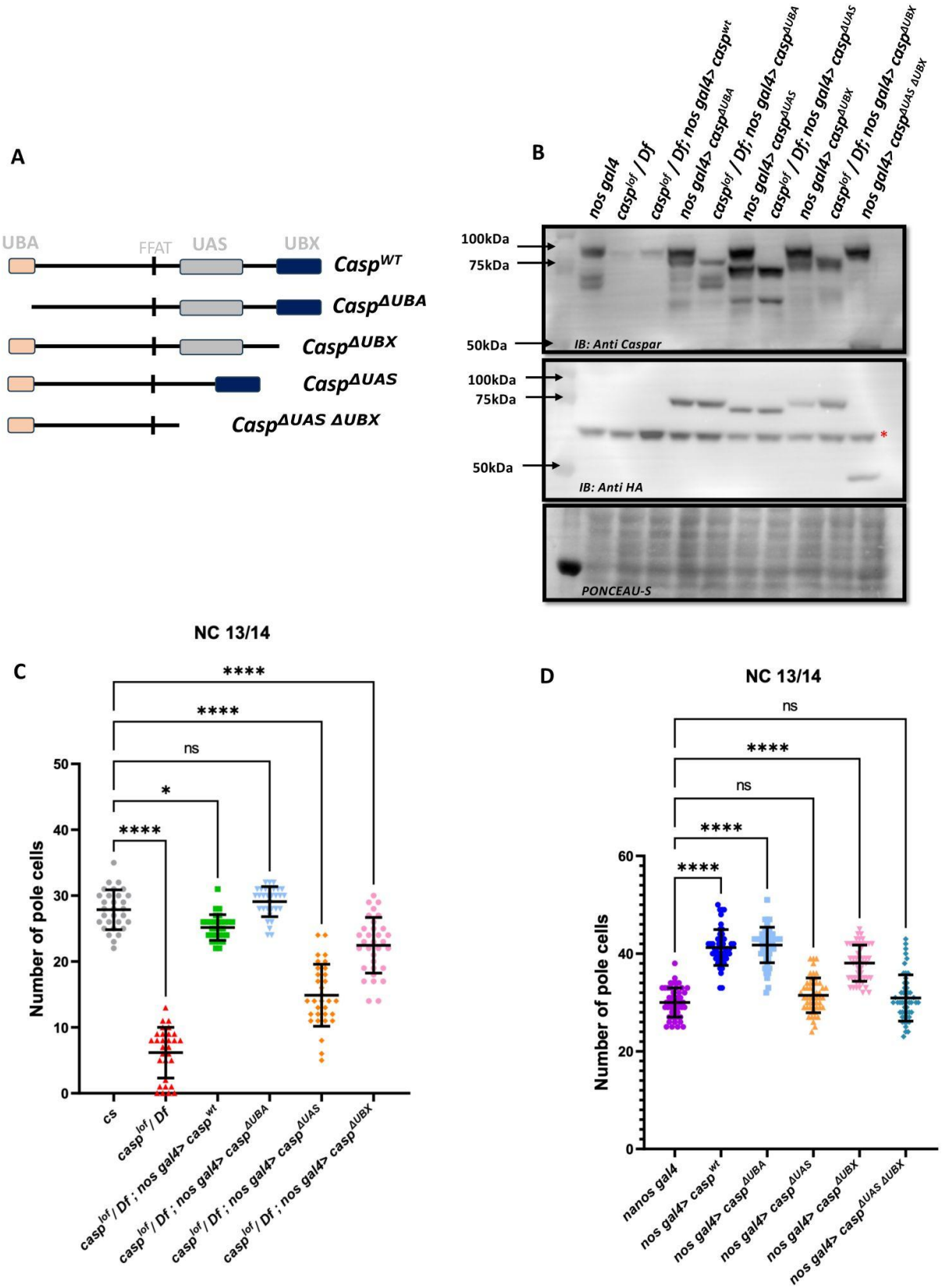

A
UBA
FFAT
UAS
UBX
Casp^WT
Casp^ΔUBA
Casp^ΔUBX
Casp^ΔUAS
Casp^ΔUAS ΔUBX
B
nos gal4
casp^lof / Df
casp^lof / Df; nos gal4> casp^wt
casp^lof / Df; nos gal4> casp^ΔUBA
casp^lof / Df; nos gal4> casp^ΔUAS
casp^lof / Df; nos gal4> casp^ΔUBX
casp^lof / Df; nos gal4> casp^ΔUBX
nos gal4> casp^ΔUAS ΔUBX
100kDa
75kDa
50kDa
IB: Anti Caspar
IB: Anti HA
PONCEAU-S
C
NC 13/14
Number of pole cells
cs
casp^lof / Df
casp^lof / Df ; nos gal4> casp^wt
casp^lof / Df ; nos gal4> casp^ΔUBA
casp^lof / Df ; nos gal4> casp^ΔUAS
casp^lof / Df ; nos gal4> casp^ΔUBX
D
NC 13/14
Number of pole cells
nanos gal4
nos gal4> casp^wt
nos gal4> casp^ΔUBA
nos gal4> casp^ΔUAS
nos gal4> casp^ΔUBX
nos gal4> casp^ΔUAS ΔUBX

Fig. 4.9 Structure-function analysis of Casp protein domains in regulating pole cell number.

(A) Schematic representation of different domain deletion variants of Casp. As shown in the schematic, the WT Casp protein consists of different functional domains including UBA, UAS, UBX and the FFAT-like motif. **(B)** Western blot of the truncated Casp proteins, deficient in the domains indicated, probed with rabbit anti-casp (1:10,000) and rabbit anti-HA (1:2000) antibody. *nos gal4* served as a control. Ponceau staining was used as a loading control. Asterisk (*) denotes non-specific antibody binding. **(C)** The total number of germ cells from nc 13/14 embryos derived from mothers expressing different *casp* domain deletion constructs in the *casplof/Df* background were marked with Vasa antibodies, quantified and plotted as a bar graph. n = 30, Ordinary one-way ANOVA, (****) P <0.0001, (*) $P<0.05$. **(D)** The total number of germ cells from nc 13/14, derived from mothers expressing casp domain deletion, in an otherwise *wild-type* background, were stained with anti-Vasa antibodies. Vasa positive cells (i.e. PGCs) were quantified and plotted as a bar graph. n = 30, Ordinary one-way ANOVA, (****) P <0.0001.

pUASp-casp$^{\Delta UBX}$ and *pUASp-casp*$^{\Delta UAS}$ construct represent deletions of the UBA, UBX, and UAS domains, respectively. Additionally, we also generated fly lines that express the *pUASp-casp*$^{\Delta UAS\Delta UBX}$ which deletes both the N and C-terminal domains simultaneously. A full-length coding sequence of Casp was also cloned, and this *pUASp-casp*^{WT} line was used as a positive control carrying the 'rescue' construct. A western blot confirmed the expression of these constructs in adult animals with different domain deletions (Δ), *UAS-casp*$^{\Delta}$ driven by *nos-GAL4* driver (**Fig. 9B**). To assess the domain-specific functions of Casp in isolation without the confounding effects of endogenous Casp, the fly lines were balanced with a *casp*^{lof} allele on the second chromosome. These balanced lines were crossed to a fly line with the maternal driver *nos- GAL4* on the third chromosome and an allele with a deficiency for *casp (Df)*. Embryos laid by mothers of the genotype *casp*^{lof}*/Df; nos-GAL4/pUASp-casp*$^{\Delta}$ were used to determine maternal effects of the domain deletions (**Fig. 9C**).

Embryos were collected from females of different genotypes and stained using anti-Vasa antibodies. As can be seen by the comparison shown in the bar graph (**Fig. 9C**), in terms of the rescue of PGC numbers, deletion of the UBA domain does not affect the rescuing activity of the Casp protein, the UBX domain deletion can rescue, but is weaker than the UBA domain deletion, whereas deletion of the UAS domain is the least effective with regards to the rescue.

To extend these observations, we overexpressed individual Casp deletion mutants in an otherwise wild-type background and counted the total number of PGCs. Neither the maternal overexpression of *pUASp-casp*^{wt} nor any of the *pUASp-casp*$^{\Delta}$ increased lethality in terms of decrease in hatching of embryos, in all cases the hatching was in the range 90-95%. Again, as seen before, overexpression of *casp*^{wt} under these conditions led to a significant increase in the total number of PGCs in blastoderm embryos, with average PGC's at 40. The ability to enhance

the PGC division is retained in the absence of the UBA domain as, in this particular background, the mean count of the total number of germ cells was similar to embryos that expressed full-length protein. By contrast, deletion of the UAS domain is unable to support the (increased) germ cell division, and the total PGC count dropped to the same level as *nos Gal4* embryos.

Interestingly, deletion of UBX domain, yet again, displayed intermediate activity, with the *ΔUBX- ΔUBA* domain deletion also being unable to drive the increase in PGCs. Proteomic analysis of the embryonic extracts suggested that in addition to TER94 protein, Casp is also associated with VAP (**Fig. S4A** ; **(Tendulkar *et al*. 2022)**). We thus sought to test if physical interaction between Casp and VAP is relevant for pole cell formation.

To assess this possibility, we used a deletion variant of Casp, $Casp^{\Delta FFAT}$, which cannot interact with VAPB **(Tendulkar *et al*. 2022)**. Expression of $casp^{\Delta FFAT}$ under the control of maternal driver, in a $casp^{lof}$/Df mother showed hatching (~60%) frequencies at par with wild type (**Fig. S4B**), with expression levels equal to other *casp* constructs (**Fig. S4C**). Importantly, *nos-Gal4-* dependent maternal overexpression *of UAS-*$casp^{\Delta FFAT}$ led to an increase in the total number of PGCs in syncytial blastoderm embryos, comparable to full-length *UAS-*$casp^{wt}$(average PGC ~23). Lastly, expression of $casp^{\Delta FFAT}$ could rescue the PGC loss upon maternal loss of *casp* (compare **Fig. S4** F1-3 to **Fig. 4.9C**). These data suggested that interaction between Casp and VAP is either not essential or partially redundant for Casp function in the context of determining PGC. Altogether, the functional analysis of different protein domains within Casp underscores the importance of UAS and UBX domains, especially in embryonic contexts, including PGC development.

4.4 Discussion

Proper development of biological systems in an organismal context involves complex interactions between various individual pathways. The qualitative nature of the interaction(s) between pathway components ultimately determines how different pathways intersect in a context-specific manner. Curiously however, there are only a defined number of signaling pathways/circuits that have been elucidated thus far. Consistently, the entire molecular cassettes that constitute a given pathway or a select number of pathway components, are reiteratively used in a variety of biological contexts even within a lifecycle of the same organism. Moreover, while such repurposing is relatively frequent, the corresponding

biological outcomes are diverse in nature and insightful in ways more than one. Our data detailing the functional involvement of Casp and TER94 in the PGCs is a case in point. Casp function was initially characterized in the context of *Drosophila* immune response **(Kim *et al*. 2006)**. However, the original study did not investigate the developmental roles of Casp and its interactors. Here, we report novel activities of *Drosophila melanogaster* immune components Casp and TER94 during early embryonic development with a specific focus on germline.

Early embryonic patterning is a dynamic process in metazoans which is primarily regulated by the deposition of maternal gene products including RNAs and proteins. Such maternal determinants are localized in a spatially restricted manner **(Schier 2007)**. On many instances, the unique pattern of localization of specific factors underlies their activities during embryo patterning. To execute their functions properly, patterning determinants heavily rely upon the members of the housekeeping machinery that carefully calibrate the synthesis, stability, transport, and degradation of diverse regulatory components. Here, we have explored novel embryonic functions for the FAF1 ortholog, Casp in *Drosophila*. Initially, we uncovered that maternal loss of *casp* activity results in partially penetrant embryonic lethality. The *casp* allele was determined to be a strong loss of function, but not a null. The partial penetrance could also be due to redundancy. Consistent with either possibility, quantitation of the individual phenotypes yielded significant but variable penetrance.

Our data demonstrate Casp's involvement during cellular movements that lead to gastrulation, a likely cause underlying the lethality. Surprisingly, Casp protein is enriched in the PGCs, which are specified at the posterior pole under the control of the master germ cell determinant, Oskar. PGC formation and specification in a young *Drosophila* embryo depends on the posteriorly anchored specialized cytoplasm (or pole plasm) enriched in RNA and protein components essential to determine germ cell identity and behavior. Two important traits distinguishing early pole cells from the surrounding somatic nuclei include precocious budding and limited mitotic self-renewal. Consistent with its functional involvement in both these processes, maternal loss of either Casp or its protein partner TER94 resulted in a reduced number of buds and subsequent loss of PGCs. While qualitatively similar, the severity of phenotypic consequence due to loss of Casp and TER94 regarding PGC loss, is not identical. Interestingly, however, both proteins appear to influence centrosome behavior and dynamics that are of paramount significance in forming pole cells in *Drosophila* embryos.

Thus far, Germ cell-less (Gcl) is the only protein shown to control PGC formation in a similar manner **(Jongens *et al*. 1992)(Cinalli and Lehmann 2013)**. Importantly Gcl activity depends on its ability to influence proper separation of centrosomes and elaboration of astral microtubules in dividing pole buds **(Lerit *et al*. 2017)**. Aberrant centrosome behavior in *gcl* mutant embryos adversely affects PGC budding and equitable distribution of germ plasm among daughter cells. Intriguingly, these phenotypic traits can be recapitulated by simply engineering defective centrosome separation **(Lerit *et al*. 2017).** It is thus noteworthy that maternal loss of Casp and TER94 leads to similar defects and future experiments will reveal mechanistic details underlying roles of these two proteins in regulating centrosome dynamics in PGCs and their possible interaction with Gcl and its protein partners. Of note, a germplasm interactome **(Thomson *et al*. 2008)** included TER94 along with other important germ plasm proteins including Vasa, Tudor, and others.

TER94, an ER protein, is a major Casp interactor. Our data also indicate that in addition to PGC formation, TER94 and Casp regulate early nuclear division cycles in syncytial blastoderm embryos and subsequent cell divisions in the somatic compartment cell cycle processes in the gastrulating embryos. (Casp and TER94 are distinct compared to Gcl in this regard which has a strictly germ cell specific function). Taken together, these observations suggest that Casp and TER94 contribute to critical early developmental functions leading to mid-blastula transition which precedes gastrulation and germband extension. In the mammalian context, the FAF1-VCP interaction is mediated by the UBX domain. Thus, it is unsurprising that deletion of the protein domain crucial for association between Casp and TER94 resulted in a somewhat diminished function as compared to a full-length version. Curiously however, deleting UAS fragment, a protein domain of unknown function, implicated in self-association, resulted in substantially compromised activity as compared to native protein. Future experiments will focus on the specific molecular interactions this (and other) individual domain(s) are involved in. It will be also important to determine how protein degradation especially engineered via ubiquitin modification contributes to Casp stability and function.

Intriguingly our data also argue that regulation mediated by *casp* is likely critical for maintenance of PGC fate via maintenance of Oskar levels. Importantly, several independent observations suggest that this influence is likely zygotic. First, barely detectable amount of Casp protein is deposited in the egg and bulk of the Casp protein is generated by translation post-fertilization. Second, Casp protein controls overall levels of Smaug at the post-transcriptional level including in the early embryonic PGCs. It was recently reported that

Smaug protein accumulates in the germ granules where it controls Oskar (and Bruno) translation negatively by binding to Smaug response elements present in the untranslated regions within the respective RNAs **(Siddiqui *et al*. 2023)**. Consistently loss of function mutations in *Smaug* result in a modestly elevated PGC count in stage 4-5 embryos whereas a reciprocal phenotype is observed in $casp^{lof}$ embryos. Taken together with the increase in Smaug levels in $casp^{lof}$ embryos, it would be reasonable to propose that PGC- specific phenotypes observed upon loss of *casp* are, in part, mediated by excess accumulation of Smaug. Moreover, these authors also suggest that Smaug dependent regulation of *oskar* RNA likely has a significant zygotic i.e. embryonic component which aligns well with our data. Future experiments will be necessary to elucidate the mechanism underlying regulation of Smaug levels by Casp. It will also be interesting to examine if Smaug and Casp regulate centrosome behaviour reciprocally to control the final PGC count.

Degradation of maternally deposited RNAs and proteins constitutes an important transition during embryonic development. Zygotic Genome Activation (ZGA) and turning over of maternal determinants (Both RNAs and proteins) happen almost simultaneously. Together these two events constitute MZT which is delayed in the germ cell compartment as opposed to soma. Nonetheless, recent data have suggested that the two events possibly occur in a coordinated manner possibly via shared components **(Colonnetta *et al*. 2023)**. Many of the embryonic phenotypes that $casp^{lof}$ embryos display (aberrant nuclear migration, cellularization defects, defective gastrulation etc.), are shared by mutations in gene products either directly or indirectly involved in one of these processes. It will be important to elucidate new interactors of Casp as its function likely impacts protein degradation and/or stability of the target proteins. Recent reports have indicated potential involvement of ZGA regulators such as Zelda and CLAMP in germline/soma distinction. Critically this function of the components of ZGA depends on proper anchoring, release and transmission of posteriorly anchored pole plasm that involves centrosome function. Maternally compromising ZGA components resulted in inappropriate release and transmission of pole plasm RNAs, a phenotype that is partially recapitulated in $casp^{lof}$ and *TER94i* embryos.

While our data have clearly established the involvement of Casp and TER94 during PGC formation and specification, further studies will be necessary to elucidate their precise molecular function(s) underlying this activity. Several observations are noteworthy in this regard and will guide the course of future investigation. The first set of results points to the

possible participation of both Casp and TER94 during ubiquitin-dependent protein degradation.

A regulatory network incorporating Bru1, Cup, Oskar, and Smaug are key to PGC specification. The Skp Cullin and F-box (SCF) containing complex marks proteins such as Smaug for degradation by functioning as an E3 ubiquitin ligase **((Cao *et al*. 2020 ; Cao *et al*. 2022)**; **Fig. 4.10A**). The ubiquitinated Smaug is then degraded by the proteasome. In the *mammalian* context, the Casp/TER94 orthologs FAF1/VCP (in *green* font; **Fig. 4.10A**) have been found to interact with the SCF complex. Our data supports a similar scenario in *Drosophila*.

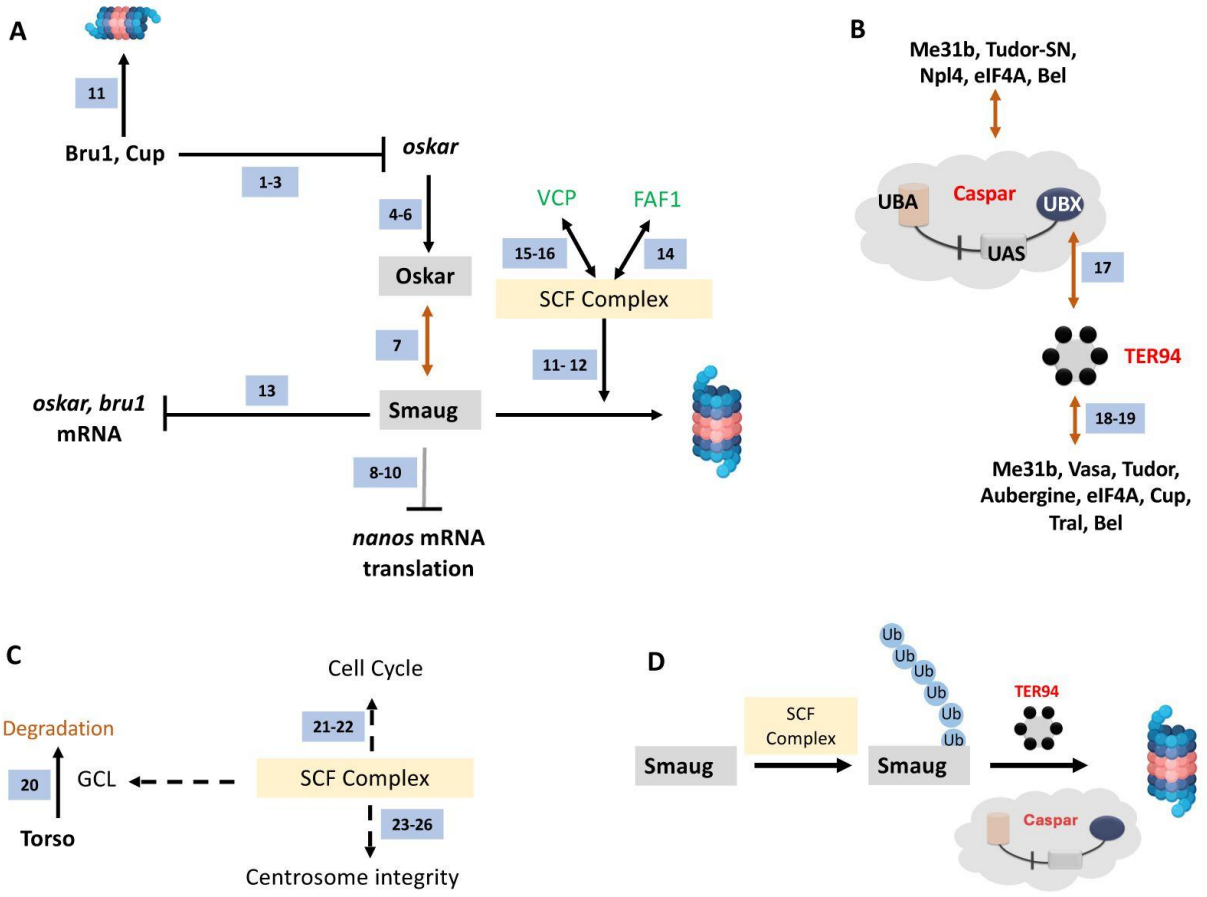


Figure 4.10 A model for Casp function in PGCs.

(**A**) Smaug is marked for degradation by the SCF complex. Specific pole plasm components including Cup, Oskar and Smaug are actively degraded. *oskar* is regulated by Bru1 and Cup (*ref 1-3*; (NAKAMURA *et al*. 2004; KIM *et al*. 2015; BAYER *et al*. 2023)). Oskar (*ref 4-6*; (MAHOWALD 2001; HUYNH AND ST JOHNSTON 2004; LEHMANN 2016)) is the master determinant of PGC fate, and hence the stability of the Oskar:Smaug complex (*ref 7*;(KUBIKOVA et al. 2023)) is a key to PGC determination and proper specification. Oskar also regulates posterior cell fate by regulating *nos* translation which is modulated by Smaug (*ref 8-10;* (DAHANUKAR et al. 1999; ZAESSINGER et al. 2006; JESKE et al. 2011)). The SCF complex is a multi-protein E3 ubiquitin ligase complex, and Smaug is one of its prominent targets (*ref 11-12;* (CAO et al. 2020; CAO et al. 2022)). Smaug appears to

repress the translation of *oskar* and *Bruno 1* mRNA (*ref 13*; (SIDDIQUI *et al*. 2023)). In mammals (marked with green font), FAF1 modulates the SCF complex (*ref 14*; (MORAIS-DE-SA *et al*. 2013)) with VCP/p97 assisting in the degradation of ubiquitinated Smaug (*ref 15-16;* (LI et al. 2014; REIM et al. 2014)), suggesting potential conservation of the activities in the fly orthologs, Casp and TER94.

(**B**) Casp/TER94 interact with an overlapping set of proteins. TER94 has earlier been demonstrated to be associated with Oskar (RUDEN *et al*. 2000). Our data points to an association between *Drosophila* Casp and TER94 (this study and *ref 17;* (TENDULKAR et al. 2022)) that participates in protein degradation. In a proteomic analysis of germ cell components, TER94 was identified (this study and *ref 18*; (THOMSON *et al*. 2008)), along with other bona-fide germ cell constituents including Cup, Tral, Bel, eIF4A, Tud and Vasa (*ref 18* and *19* (THOMSON *et al*. 2008; DEHAAN *et al*. 2017)). Both, Casp and TER94 are thus enriched in the pole cells and interact with proteins that specify PGC fate.

(**C**). The SCF Complex, Gcl and Centrosome integrity. The SCF complex is localized to the centrosome. Gcl interacts with Cullin 3 to degrade Torso receptor to promote PGC fate (*ref 20*; (PAE *et al*. 2017)). The SCF complex regulates the cell cycle (*ref 21-22*; (MARGOTTIN-GOGUET *et al*. 2003; ROGERS *et al*. 2009)) and is a known regulator of centrosomal integrity (*ref 23-26*;(WOJCIK *et al*. 2000; MURPHY 2003; PHUONG THAO *et al*. 2006; CUNHA-FERREIRA *et al*. 2009)), thereby influencing PGC specification (LERIT *et al*. 2017).

(**D**) Casp and TER94 assist in the degradation of Smaug. The SCF complex works as a ubiquitin E3 ligase to mark Smaug for Degradation and PolyUb-Smaug is degraded by the action of Casp and TER94 during the late-MZT.

It would be reasonable to propose that Casp recruits TER94 via Casp's UBX domain (**Fig. 4.10B**). Independent proteomic studies **((Thomson *et al.* 2008)(Dehaan *et al.* 2017)(Tendulkar *et al.* 2022)**; **Fig 4.10 B**) suggest multiple overlaps between Casp/TER94 and germ-cell specific protein complexes (**Fig. 4.10B**), again suggesting functional relationships between Casp/TER94 and germ cell determinants. In addition, the SCF complex contributes to cell cycle progression and integrity of centrosomes (**(Wojcik *et al.* 2000)(Murphy 2003)(Phuong Thao *et al.* 2006)**; **Fig 4.10C**). Total PGC count in early embryos depends on centrosome function and efficient mitotic divisions. So SCF complex based degradation may have multiple targets that participate during germ cell formation and specification (**Fig. 4.10C**).

Casp likely functions as an adapter protein that works in conjunction with the SCF complex (**(Cao *et al.* 2020)(Cao *et al.* 2022)**; **Fig. 4.10D**) and TER94 (**Fig. 4.10D**). The binary complex between TER94 and Casp subsequently targets ubiquitinated proteins such as Smaug which are present in a complexed form with other proteins. Homeostatic maintenance of Smaug levels via Casp-mediated proteolytic degradation could explain many of our observations. Consistently, increased Smaug levels, due to diminished Casp function leads to loss of PGCs whereas, *smaug* mutant embryos show the opposite phenotype **(Siddiqui *et al.* 2023)**. Intriguingly, even though Smaug and Oskar are members of a complex, the reduction in Casp

activity results in the stabilization of Smaug. By contrast, Oskar levels are diminished resulting in loss of PGCs. The mechanism underlying the substrate-dependent divergent activities is not known at this point.

The similarity between the phenotypic consequences due to loss of ZGA components and Casp/TER94 also needs a special mention. The only study implicating TER94 during early embryogenesis **(Zeng *et al*. 2014)** suggested that TER94 can potentiate BMP signaling. Interestingly, *decapentaplegic* (*dpp*), a BMP ligand which is one of the important targets of ZGA regulators including *Zelda*, can also influence the specification of embryonic PGCs **(Colonnetta *et al*. 2023).** PGCs need BMP signals (Dpp) to maintain their identity.

Furthermore, the exposure to BMP signals needs careful calibration as excess BMP signaling in the surrounding soma leads to PGC loss. Importantly, components of protein degradation machinery including those involved in ubiquitination (Smurf, a Ubiquitin ligase) and sumoylation (Ubc9) appear to be involved in fine-tuning the signaling **(Deshpande *et al*. 2014)**. Thus, it will be of interest to determine if the aberrant germ cell specification observed due to loss of *casp* and the components of BMP signaling pathway is mechanistically connected. This seems especially relevant in the light of their respective dependence on the components of the protein degradation machinery. It will be of considerable interest to investigate whether and how, zygotic activity of Smaug, one of important canonical MBT regulators, fits into this picture. In sum, while the specific details await detailed examination, it is apparent that possible recapitulation of maternal regulation in the zygotic context may be a recurrent theme rather than an isolated anomaly.

4.5 Materials and Methods

4.5.1 Fly husbandry and stocks

Flies were raised on standard cornmeal agar at 25 °C unless stated otherwise. *Casp*^{lof} (11373), *Casp Df* (23691), *nos Gal4* (4937), *Mat α4-tubulinGal4:VP16* (7063) and *TER94 RNAi* (32869) lines were procured from the Bloomington *Drosophila* Stock Centre, with numbers in brackets indicated Stock numbers.

4.5.2 Cloning of *casp* deletion constructs and generation of transgenic flies

UASp-casp$^{\Delta}$ constructs (*pUASp casp$^{\Delta UBA}$*, *pUASp casp$^{\Delta UAS}$*, *pUASp casp$^{\Delta UBX}$*, *pUASp casp$^{\Delta UAS\Delta UBX}$*) were generated, starting from a *pUASp-attB-Casp* construct (Tendulkar *et al*. 2022), with the design including an N-terminal HA tag. The constructs were injected into a *w1118; Attp2* embryo in the NCBS-CAMP transgenic injection facility (*https://www.ncbs.res.in/research-facilities/drosophila- services*), and stable lines were generated by balancing against a *w^{1118}; TM3Sb/TM6Tb* animal.

For cloning and amplification, a functional N-terminal HA tag was introduced with the 5'-forward primers and a 3'-homology arm was introduced using the 3'-reverse primers. The exception was the *pUASp Casp$^{\Delta UAS}$* construct where two fragments, one upstream and one downstream of the UAS domain were amplified separately with 22 nucleotide homologous overhang between them to mediate homologous recombination. The upstream fragment contained the HA sequence while the downstream fragment contained the 3'-homology arm.

The sequences of the primers used are as follows. *TGTTCCAGATTACGCTGGCGGC* was used as the 5'-forward primer for the *pUASp casp*^{wt}, *pUASp casp$^{\Delta UBX}$*, *pUASp casp$^{\Delta UAS\Delta UBX}$* and the upstream fragment of the *pUASp casp$^{\Delta UAS}$* constructs.

TGTTCCAGATTACGCTGGCGGCCCCCATCCTATCCTGGTGCC was used as the 5'-forward primer for the *pUASp casp$^{\Delta UBA}$* construct.

CGGATGATGAGATAAGTGGCTCCACGGAAACATGCGAAATGTTTGAGGAGCAG was used as forward primer for amplifying downstream fragment of the *pUASp casp$^{\Delta UAS}$* construct.

ACCATGGGTTTAGGTATAATGTTATCAAGCTCC was the reverse primer for *pUASp casp^{wt}*, *pUASp casp$^{\Delta UBA}$*, and downstream fragment of *pUASp casp$^{\Delta UAS}$*. *TTAGGTATAATGTTATCAAGCTCCTCATTCGGACGGCTCCTGAGGTAG* was the reverse primer for *pUASp casp$^{\Delta UBX}$*.

TTAGGTATAATGTTATCAAGCTCCTCACGTGGAGCCACTTATCTCATCATCCG was used as the reverse primer for *pUASp casp$^{\Delta UAS\Delta UBX}$*.

CGTGGAGCCACTTATCTCATCATCCG was the reverse primer for upstream fragment of *pUASp casp$^{\Delta UAS}$*. PCR products were further amplified with a common 5'-primer *ATAGGCCACTAGTGGATCTGATGTACCCATACGATGTTCCAGATTACGCTGGCGGC* and their respective reverse primers to introduce the 5'-homology arm. Inserts were recombined into *pUASp AttB* vector linearised at the BamH1 site, using a variation of the SLiCE cloning

method (Zhang *et al*. 2012 ; Zhang *et al*. 2014). In short, E.coli DH10B expressing the optimised λ-prophage red recombinase system (PPY cells) was cultured in the presence of arabinose to induce recombinase expression and subsequently used for preparing competent cells. DNA fragments containing homologous overhangs were co-transformed into the competent bacteria and resultant colonies were screened through PCR to identify proper recombinants and subsequently confirmed through sequencing. The pUASp *casp*$^{\Delta FFAT}$ was PCR amplified from *pRM-HA:casp*$^{\Delta FFAT}$ **(Tendulkar *et al*. 2022)**, cloned into pUASp-AttB, sequenced for validation and injected into a *w*^{1118}*; Attp2* animal.

4.5.3 Embryonic lethality

0-3 hr embryos were collected, transferred to a fresh sugar-agar plate, and unhatched larvae were scored after 48 hours to determine viability.

4.5.4 Immunoprecipitation

0–3-hour embryos were lysed in Co-IP Lysis Buffer (20mM Tris pH 8.0, 137mM NaCl, 1% IGEPAL, 2mM EDTA, 1X PIC) using a Dounce homogenizer, and centrifuged at 21,000 g for 30 minutes. 3 mg of total lysate was incubated with 5 µg of primary antibody (Rb anti-Casp) and 5 µg of Normal Rabbit IgG overnight at 4 °C. Antigen-antibody complexes were captured using 50 µL of BioRad SureBeads Protein A (1614013) at 4 °C for 4 hours. Beads were washed six times with Co-IP Lysis Buffer and protein complexes eluted by boiling in 1X Laemmli Sample Buffer. Eluted proteins were resolved on a 10% polyacrylamide gel followed by western blotting or in-gel trypsin digestion, described in the following sections.

4.5.5 Western blot analysis

Embryos collected at varied time points (0-3, 0-1, 1-2, 2-3 hours) were lysed in RIPA buffer (50 mM Tris-Cl, 150 mM NaCl, 0.1% SDS, 0.01% Sodium azide, 0.5% sodium deoxycholate, 1 mM EDTA, 1% Triton X-100, 1X PIC) with a pellet pestle (Kontes). Lysates were cleared by centrifugation at 21,000 g at 4 °C for 30 minutes. Protein concentration was estimated using a BCA assay (Pierce) and 30-40 µg of total protein was loaded onto the gel after boiling in 1X Laemmli Sample Buffer. Proteins separated by 10% SDS-PAGE were transferred onto a PVDF membrane (Immobilon-E, Merck) and blocked in 5% milk in Tris-Buffer Saline (TBS) with

0.1% Tween 20 (TBS-T) for an hour. Blots were then incubated overnight with primary antibody diluted in 5% milk in TBS-T, at 4 °C. Following three washes with TBS-T, blots were incubated with secondary antibodies diluted in 5% milk in TBS-T, for 1 hour at room temperature. Blots were washed thrice with TBS-T and visualized on a LAS4000 Fuji imaging system after incubating with Immobilon Western Chemiluminescent HRP substrate (Merck). The following antibodies were used: Rabbit anti-VAP, 1:10000 (Tendulkar *et al*. 2022), Mouse anti-α-Tubulin, 1:10000 (T6074, Sigma), Mouse anti-Ubiquitin, 1:1000 (P4D1, Santa Cruz Biotechnology), Rabbit anti-HA, 1:2000 (04-902 DW-2, Sigma-Aldrich), Rabbit anti-Casp, 1:10000 (Tendulkar *et al*. 2022), Goat anti-rabbit HRP and Goat anti-mouse HRP secondary antibodies, each at 1:10000 (Jackson ImmunoResearch). Rabbit anti-me31B, 1:1000, Rabbit anti-Smaug, 1:500, Rat anti-Tral, 1:1000 were a kind gift from Elmar Whale (Cao *et al*. 2020).

4.5.6 In-gel Trypsin Digestion and LC-MS/MS Analysis

Before in-gel trypsin digestion of the Co-IP eluate, the antibody was crosslinked to the SureBeads using DMP (Sigma) according to the NEB crosslinking protocol to avoid elution of the antibody. After crosslinking 10µg Casp antibody, Co-IP was performed as described above. In-gel trypsin digestion was carried out as previously described (SHEVCHENKO et al. 2006). Briefly, Coomassie- stained bands on the gel were excised and cut into 1 mm cubes. Gel pieces were transferred to a clean microcentrifuge tube and destained with buffer containing 50% acetonitrile in 50 mM Ammonium bicarbonate.

Reduction and alkylation were carried out on the destained gel pieces by incubating with 10mM dithiothreitol (DTT) followed by incubating with 20mM iodoacetamide. Gel pieces were saturated with sequencing grade Trypsin (Promega) at a concentration of 10 ng/µL and incubated overnight at 37 °C. Peptides were extracted by sequential addition of 100 µl of 0.4% Trifluoroacetic acid (TFA) in 10% ACN, 100 µl of 0.4% TFA in 60% ACN and 100 µl of ACN. The pooled extract was dried in a vacuum centrifuge and reconstituted with 50 µl of 0.1% TFA. The peptides in TFA were purified using the StageTip protocol (RAPPSILBER et al. 2007). LC– MS/MS analysis was performed on the Sciex TripleTOF6600 mass spectrometer interfaced with an Eksigent nano-LC 425. Tryptic peptides (1 µg) were loaded onto an Eksigent C18 trap (5 µg capacity) and subsequently eluted with a linear acetonitrile gradient on an Eksigent C18 analytical column (15 cm × 75-µm internal diameter). A typical LC run lasted 2 h post loading onto the trap at a constant flow rate of 300 nL/min with solvent A consisting of

water + 0.1% formic acid and solvent B consisting of acetonitrile. The gradient schedule for the LC run was 5% (vol/vol) B for 10 min, a linear gradient of B from 0% to 80% (vol/vol) over 80 min, 80% (vol/vol) B for 15 min and equilibration with 5% (vol/vol) B for 15 min. Data was acquired in an information-dependent acquisition (IDA) mode over a mass range of 300–2,000 m/z. Each full MS survey scan was followed by MS/MS of the 15 most intense peptides. Dynamic exclusion was enabled for all experiments (repeat count 1; exclusion duration 6 s). Peptides were identified and quantified using the SCIEX ProteinPilot software at a false discovery rate (FDR) of 1%. A RefSeq *Drosophila* protein database (release 6) was used for peptide identification. Proteins that were identified in two or more replicates were tabulated.

4.5.7 Fixation, immunostaining, and imaging of embryos

0-3 hr embryos were collected in a sieve and dechorionated in 4% sodium hypochlorite for 90 seconds. After thorough washes with distilled water, embryos were fixed in a 1:1 heptane:4% PFA solution for 20 minutes. The PFA layer was removed, and embryos in heptane were re-constituted with an equal volume of methanol. Embryos were devitellinized by vigorous shaking in the 1:1 heptane:methanol mixture. The heptane layer and the interphase containing non-devitellinized embryos were carefully removed. Devitellinized embryos in the bottom methanol phase were washed twice with methanol and stored at -20 °C till they were ready to be imunostained. For phalloidin and Smaug staining, embryos fixed in heptane: 4% PFA were hand de-vitellinized, after which the standard immunostaining procedure was followed. For immunostaining, embryos were rehydrated by washing thrice with 0.3% PBS-TritonX 100 (PBS-T) for 15 minutes each. Embryos were blocked in 2% BSA in 0.3% PBS-T for 1 hour at room temperature (RT). Embryos were incubated at 4 °C overnight with primary antibodies diluted in 2% BSA in 0.3% PBS-T at the appropriate dilutions. Following three 15-minute washes with 0.3% PBS-T, embryos were incubated in the appropriate secondary antibodies for 1 hour at room temperature. The following antibodies were used: Rabbit anti-Casp, 1:1000 (Tendulkar *et al*. 2022); Rat anti-α-vasa; 1:50 (DSHB); Rabbit anti-VCP, 1:200 (#2648, Cell Signaling Technology); Rabbit anti-Oskar, 1:1000 (generated by Mandy Jeske, Anne Ephrussi lab); Rabbit anti-Smaug, 1:50; Mouse anti-γ-tubulin, 1:1000 (T6557, Sigma), Rabbit P-Histone H3 (S10) 1:300 (#9701, Cell Signalling); Alexa Fluor 568 Phalloidin, 1:1000 (Invitrogen) The following secondary antibodies were used, goat anti-mouse Alexa488/ goat anti-rabbit Alexa568/ goat anti-rat Alexa647/ goat anti-rabbit Alexa647/ goat anti- rabbit Alexa480, 1:1000 (DSHB). Embryos were washed thrice with 0.3% PBS-T, and DAPI/Hoechst (1:500)

was added in the penultimate wash. Embryos were mounted in 70% glycerol and observed under a Leica sp8 confocal microscope with a 20X objective.

4.5.8 Live imaging of embryos

Embryos at the appropriate stage were washed and dechorionated as described previously. A 2- well NuncTM Lab-TekTM II Chamber SlideTM System (ThermoFisher Scientific) was affixed with a 3M Scotch double-sided tape, and dechorionated embryos with intact vitelline membranes were mounted on the tape under a dissecting microscope. Halocarbon oil 200 was used to cover the embryos to prevent dehydration during imaging. Time-lapse imaging of embryos was performed on an inverted LSM confocal system (Zeiss multiphoton 710) at 20X for ~8 hours. Embryo images were acquired at 5 min intervals.

4.5.9 Quantification of western blots

Western blots were quantified using ImageJ. Each protein was normalized to the tubulin loading control, and the highest signal for each was set to 1 and all other tubulin band intensities were normalized against it. For MZT experiments, western blots were used to determine protein degradation dynamics. 3-5 biological replicates, representing 3-5 independent embryo collection were assayed. Band intensities were quantified using ImageJ and normalized to α-tubulin as loading control. For each replicate intensities were normalized to the first time point (0-1 hour). Significance thresholds are presented in the figures and figure captions.

4.5.10 Quantification of the Smaug Immunostaining

All quantification of Smaug stained images were done on ImageJ. ROIs were made for a single pole cell through the z-stacks, thereby only analysing one pole cell at a time. For each embryo five such pole cells were measured. To quantify intensity, mean intensity was measured for each ROI per stack and added to get the final intensity. Each data point on the graph represents an average of all five cell intensities per embryo. For Smaug punctae measurement, a combined ROI for each cell was taken and duplicated. Each of the stack was then analysed using 3D objects counter, where thresholding for all images was kept within the range of 80-90 with a minimum size filter kept at zero. The number of objects counted was taken as the total number of Smaug punctae per cell. Five such pole cells were analysed per embryo and the average was

plotted as a single data point on the graph. For quantification of poleplasm volume in the embryos, embryos were stained with Oskar. An ROI (kept same for all images) encompassing the posterior part of the embryo was drawn. Gaussian blur 3D (X sigma = 2, Y sigma = 2 and Z sigma = 2) was applied to each image. The volume of pole plasm was analysed by using 3D objects counter after appropriate thresholding.

4.6 Contributions

SD performed most of the experiments, with contributions from SH, NW, JS, and AR. GR and GD conceptualised the project with input from SD and SH. SD, GR, and GD analyzed and interpreted the data. GD, GR and SD wrote the manuscript. GR and GD supervised the project and acquired funding.

4.7 Acknowledgements

Stocks obtained from the Bloomington Drosophila Stock Center (NIH P40OD018537) were used in this study; Elmar Whale and Christiane Rammelt for the kind gift of anti-Tral, anti-Me31B, and anti-Smaug antibodies; Snehal Patil and Yashwant Pawar for fly media and stock maintenance; IISER Microscopy facility, Dr Santosh Podder, and Vijay Vittal, for training and maintenance.

4.8 References

Adams M. D., Celniker S. E., Holt R. A., Evans C. A., Gocayne J. D., et al. (2000) The genome sequence of Drosophila melanogaster *Science* 287:2185–2195

Arkov A. L., Wang J. Y., Ramos A., Lehmann R. (2006) The role of Tudor domains in germline development and polar granule architecture *Development* 133:4053–4062

Azuma Y., Tokuda T., Shimamura M., Kyotani A., Sasayama H., et al. (2014) Identification of ter94, Drosophila VCP, as a strong modulator of motor neuron degeneration induced by knockdown of Caz, Drosophila FUS *Hum Mol Genet* 23:3467–3480

Bayer L. V., Milano S., Formel S. K., Kaur H., Ravichandran R., et al. (2023) Cup is essential for oskar mRNA translational repression during early Drosophila oogenesis *RNA Biol* 20:573–587

Belvin M. P., Anderson K. V. (1996) A conserved signaling pathway: the Drosophila toll-dorsal pathway *Annu Rev Cell Dev Biol* 12:393–416

Blake-Hedges C., Megraw T. L. (2019) Coordination of Embryogenesis by the Centrosome in Drosophila melanogaster *Results Probl Cell Differ* 67:277–321

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

Brennan C. A., Anderson K. V. (2004) Drosophila: the genetics of innate immune recognition and response *Annu Rev Immunol* 22:457–483

Brown J. B., Boley N., Eisman R., May G. E., Stoiber M. H., et al. (2014) Diversity and dynamics of the Drosophila transcriptome *Nature* 512:393–399

Buchon N., Silverman N., Cherry S. (2014) Immunity in Drosophila melanogaster--from microbial recognition to whole-organism physiology *Nat Rev Immunol* 14:796–810

Cao W. X., Kabelitz S., Gupta M., Yeung E., Lin S., et al. (2020) Precise Temporal Regulation of Post-transcriptional Repressors Is Required for an Orderly Drosophila Maternal-to- Zygotic Transition *Cell Rep* 31

Cao W. X., Karaiskakis A., Lin S., Angers S., Lipshitz H. D. (2022) The F-box protein Bard (CG14317) targets the Smaug RNA-binding protein for destruction during the Drosophila maternal-to-zygotic transition *Genetics* 220

Casas-Vila N., Bluhm A., Sayols S., Dinges N., Dejung M., et al. (2017) The developmental proteome of Drosophila melanogaster *Genome Res* 27:1273–1285

Cavey M., Lecuit T. (2008) Imaging cellular and molecular dynamics in live embryos using fluorescent proteins *Methods Mol Biol* 420:219–238

Chang Y. C., Hung W. T., Chang Y. C., Chang H. C., Wu C. L., et al. (2011) Pathogenic VCP/TER94 alleles are dominant actives and contribute to neurodegeneration by altering cellular ATP level in a Drosophila IBMPFD model *PLoS Genet* 7

Chen Z. X., Sturgill D., Qu J., Jiang H., Park S., et al. (2014) Comparative validation of the D. melanogaster modENCODE transcriptome annotation *Genome Res* 24:1209–1223

Chu K. T., Niu X. H., Williams L. T. (1995) A Fas-Associated Protein Factor, Faf1, Potentiates Fas-Mediated Apoptosis *Proceedings of the National Academy of Sciences of the United States of America* 92:11894–11898

Cinalli R. M., Lehmann R. (2013) A spindle-independent cleavage pathway controls germ cell formation in Drosophila *Nat Cell Biol* 15:839–845

Colonnetta M. M., Schedl P., Deshpande G. (2023) Germline/soma distinction in Drosophila embryos requires regulators of zygotic genome activation *Elife* 12

Cunha-Ferreira I., Rodrigues-Martins A., Bento I., Riparbelli M., Zhang W., et al. (2009) The SCF/Slimb ubiquitin ligase limits centrosome amplification through degradation of SAK/PLK4 *Curr Biol* 19:43–49

Dahanukar A., Walker J. A., Wharton R. P. (1999) Smaug, a novel RNA-binding protein that operates a translational switch in Drosophila *Mol Cell* 4:209–218

Dai R. M., Chen E., Longo D. L., Gorbea C. M., Li C. C. (1998) Involvement of valosin-containing protein, an ATPase Co-purified with IkappaBalpha and 26 S proteasome, in ubiquitin-proteasome-mediated degradation of IkappaBalpha *J Biol Chem* 273:3562–3573

De Zio D., Ferraro E., D'Amelio M., Simoni V., Bordi M., et al. (2008) Faf1 is expressed during neurodevelopment and is involved in Apaf1-dependent caspase-3 activation in proneural cells *Cell Mol Life Sci* 65:1780–1790

DeHaan H., McCambridge A., Armstrong B., Cruse C., Solanki D., et al. (2017) An in vivo proteomic analysis of the Me31B interactome in Drosophila germ granules *FEBS Lett* 591:3536–3547

Deshpande G., Willis E., Chatterjee S., Fernandez R., Dias K., et al. (2014) BMP signaling and the maintenance of primordial germ cell identity in Drosophila embryos *PLoS One* 9

Ephrussi A., Lehmann R. (1992) Induction of germ cell formation by oskar *Nature* 358:387–392

Ewens C. A., Panico S., Kloppsteck P., McKeown C., Ebong I. O., et al. (2014) The p97-FAF1 protein complex reveals a common mode of p97 adaptor binding *J Biol Chem* 289:12077–12084

Govind S. (1999) Control of development and immunity by rel transcription factors in Drosophila *Oncogene* 18:6875–6887

Griciuc A., Aron L., Roux M. J., Klein R., Giangrande A., et al. (2010) Inactivation of VCP/ter94 suppresses retinal pathology caused by misfolded rhodopsin in Drosophila *PLoS Genet* 6

Higashiyama H., Hirose F., Yamaguchi M., Inoue Y. H., Fujikake N., et al. (2002) Identification of ter94, Drosophila VCP, as a modulator of polyglutamine-induced neurodegeneration *Cell Death Differ* 9:264–273

Hoffmann J. A., Reichhart J. M. (2002) Drosophila innate immunity: an evolutionary perspective *Nat Immunol* 3:121–126

Hultmark D. (1993) Immune reactions in Drosophila and other insects: a model for innate immunity *Trends Genet* 9:178–183

Huynh J. R., St Johnston D. (2004) The origin of asymmetry: early polarisation of the Drosophila germline cyst and oocyte *Curr Biol* 14:R438–449

Imler J. L., Hoffmann J. A. (2000) Signaling mechanisms in the antimicrobial host defense of Drosophila *Curr Opin Microbiol* 3:16–22

Jentsch S., Rumpf S. (2007) Cdc48 (p97): a “molecular gearbox” in the ubiquitin pathway? *Trends Biochem Sci* 32:6–11

Jeske M., Moritz B., Anders A., Wahle E. (2011) Smaug assembles an ATP-dependent stable complex repressing nanos mRNA translation at multiple levels *EMBO J* 30:90–103

Jones J. R., Macdonald P. M. (2007) Oskar controls morphology of polar granules and nuclear bodies in Drosophila *Development* 134:233–236

Jongens T. A., Ackerman L. D., Swedlow J. R., Jan L. Y., Jan Y. N. (1994) Germ cell-less encodes a cell type-specific nuclear pore-associated protein and functions early in the germ-cell specification pathway of Drosophila *Genes Dev* 8:2123–2136

Jongens T. A., Hay B., Jan L. Y., Jan Y. N. (1992) The germ cell-less gene product: a posteriorly localized component necessary for germ cell development in Drosophila *Cell* 70:569–584

Kim G., Pai C. I., Sato K., Person M. D., Nakamura A., et al. (2015) Region-specific activation of oskar mRNA translation by inhibition of Bruno-mediated repression *PLoS Genet* 11

Kim H., Zhang H., Meng D., Russell G., Lee J. N., et al. (2013) UAS domain of Ubxd8 and FAF1 polymerizes upon interaction with long-chain unsaturated fatty acids *J Lipid Res* 54:2144– 2152

Kim M., Lee J. H., Lee S. Y., Kim E., Chung J. (2006) Caspar, a suppressor of antibacterial immunity in Drosophila *Proceedings of the National Academy of Sciences of the United States of America* 103:16358–16363

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

Kloppsteck P., Ewens C. A., Forster A., Zhang X., Freemont P. S. (2012) Regulation of p97 in the ubiquitin-proteasome system by the UBX protein-family *Biochim Biophys Acta* 1823:125–129

Kubikova J., Ubartaite G., Metz J., Jeske M. (2023) Structural basis for binding of Drosophila Smaug to the GPCR Smoothened and to the germline inducer Oskar *Proc Natl Acad Sci U S A* 120

Kushimura Y., Tokuda T., Azuma Y., Yamamoto I., Mizuta I., et al. (2018) Overexpression of ter94, Drosophila VCP, improves motor neuron degeneration induced by knockdown of TBPH, Drosophila TDP-43 *Am J Neurodegener Dis* 7:11–31

Lattao R., Kovacs L., Glover D. M. (2017) The Centrioles, Centrosomes, Basal Bodies, and Cilia of Drosophila melanogaster *Genetics* 206:33–53

Lecuyer E., Yoshida H., Parthasarathy N., Alm C., Babak T., et al. (2007) Global analysis of mRNA localization reveals a prominent role in organizing cellular architecture and function *Cell* 131:174–187

Lee J. J., Park J. K., Jeong J., Jeon H., Yoon J. B., et al. (2013) Complex of Fas-associated factor 1 (FAF1) with valosin-containing protein (VCP)-Npl4-Ufd1 and polyubiquitinated proteins promotes endoplasmic reticulum-associated degradation (ERAD) *J Biol Chem* 288:6998– 7011

Lehmann R. (2016) Germ Plasm Biogenesis--An Oskar-Centric Perspective *Curr Top Dev Biol* 116:679–707

Lemaitre B., Hoffmann J. (2007) The host defense of Drosophila melanogaster *Annu Rev Immunol* 25:697–743

Lerit D. A., Gavis E. R. (2011) Transport of germ plasm on astral microtubules directs germ cell development in Drosophila *Curr Biol* 21:439–448

Lerit D. A., Shebelut C. W., Lawlor K. J., Rusan N. M., Gavis E. R., et al. (2017) Germ Cell-less Promotes Centrosome Segregation to Induce Germ Cell Formation *Cell Rep* 18:831–839

Li J. M., Wu H., Zhang W., Blackburn M. R., Jin J. (2014) The p97-UFD1L-NPL4 protein complex mediates cytokine-induced IkappaBalpha proteolysis *Mol Cell Biol* 34:335–347

Mahowald A. P. (2001) Assembly of the Drosophila germ plasm *Int Rev Cytol* 203:187–213

Margottin-Goguet F., Hsu J. Y., Loktev A., Hsieh H. M., Reimann J. D., et al. (2003) Prophase destruction of Emi1 by the SCF(betaTrCP/Slimb) ubiquitin ligase activates the anaphase promoting complex to allow progression beyond prometaphase *Dev Cell* 4:813–826

Medzhitov R. (2001) Toll-like receptors and innate immunity *Nat Rev Immunol* 1:135–145 Menges C. W., Altomare D. A., Testa J. R. (2009) FAS-associated factor 1 (FAF1) Diverse

functions and implications for oncogenesis *Cell Cycle* 8:2528–2534

Meyer H., Bug M., Bremer S. (2012) Emerging functions of the VCP/p97 AAA-ATPase in the ubiquitin system *Nat Cell Biol* 14:117–123

Meyer H. H. (2005) Golgi reassembly after mitosis: the AAA family meets the ubiquitin family *Biochim Biophys Acta* 1744:108–119

Meyer H. H., Shorter J. G., Seemann J., Pappin D., Warren G. (2000) A complex of mammalian ufd1 and npl4 links the AAA-ATPase, p97, to ubiquitin and nuclear transport pathways *EMBO J* 19:2181–2192

Min-Young P., Ji-Hyun M., Eunhee K. (2004) Fas associated factor 1, FAF1, impairs NF-kappaB activation by interacting with the IkappaB kinase complex *Molecular Biology of the Cell* 15:2544–2549

Morais-de-Sa E., Vega-Rioja A., Trovisco V., St Johnston D. (2013) Oskar is targeted for degradation by the sequential action of Par-1, GSK-3, and the SCF(-)Slimb ubiquitin ligase *Dev Cell* 26:303–314

Muers M. (2011) Functional genomics: the modENCODE guide to the genome *Nat Rev Genet* 12

Murphy T. D. (2003) Drosophila skpA, a component of SCF ubiquitin ligases, regulates centrosome duplication independently of cyclin E accumulation *J Cell Sci* 116:2321–2332

Nakamura A., Sato K., Hanyu-Nakamura K. (2004) Drosophila cup is an eIF4E binding protein that associates with Bruno and regulates oskar mRNA translation in oogenesis *Dev Cell* 6:69–78

Nusslein-Volhard C. (2022) The Toll gene in Drosophila pattern formation *Trends Genet* 38:231–245

Nusslein-Volhard C., Wieschaus E. (1980) Mutations affecting segment number and polarity in Drosophila *Nature* 287:795–801

Pae J., Cinalli R. M., Marzio A., Pagano M., Lehmann R. (2017) GCL and CUL3 Control the Switch between Cell Lineages by Mediating Localized Degradation of an RTK *Dev Cell* 42:130–142

Park M. Y., Moon J. H., Lee K. S., Choi H. I., Chung J., et al. (2007) FAF1 suppresses I kappa B kinase (IKK) activation by disrupting the IKK complex assembly *Journal of Biological Chemistry* 282:27572–27577

Peters J. M., Walsh M. J., Franke W. W. (1990) An abundant and ubiquitous homo-oligomeric ring-shaped ATPase particle related to the putative vesicle fusion proteins Sec18p and NSF *EMBO J* 9:1757–1767

Phuong Thao D. T., Ida H., Yoshida H., Yamaguchi M. (2006) Identification of the Drosophila skpA gene as a novel target of the transcription factor DREF *Exp Cell Res* 312:3641–3650

Raff J. W., Glover D. M. (1989) Centrosomes, and not nuclei, initiate pole cell formation in Drosophila embryos *Cell* 57:611–619

Reim G., Hruzova M., Goetze S., Basler K. (2014) Protection of armadillo/beta-Catenin by armless, a novel positive regulator of wingless signaling *PLoS Biol* 12

Rogers G. C., Rusan N. M., Roberts D. M., Peifer M., Rogers S. L. (2009) The SCF Slimb ubiquitin ligase regulates Plk4/Sak levels to block centriole reduplication *J Cell Biol* 184:225–239

Roth S. (2023) Neofunctionalization of Toll Signaling in Insects: From Immunity to Dorsoventral Patterning *Annu Rev Cell Dev Biol* 39:1–22

Ruden D. M., Sollars V., Wang X., Mori D., Alterman M., et al. (2000) Membrane fusion proteins are required for oskar mRNA localization in the Drosophila egg chamber *Dev Biol* 218:314– 325

Ryu S. W., Kim E. (2001) Apoptosis induced by human Fas-associated factor 1, hFAF1, requires its ubiquitin homologous domain, but not the Fas-binding domain *Biochem Biophys Res Commun* 286:1027–1032

Ryu S. W., Lee S. J., Park M. Y., Jun J., Jung Y. K., et al. (2003) Fas-associated factor 1, FAF1, is a member of Fas death-inducing signaling complex *Journal of Biological Chemistry* 278:24003– 24010

Schier A. F. (2007) The maternal-zygotic transition: death and birth of RNAs *Science* 316:406–407

Schuberth C., Buchberger A. (2008) UBX domain proteins: major regulators of the AAA ATPase Cdc48/p97 *Cell Mol Life Sci* 65:2360–2371

Siddiqui N. U., Karaiskakis A., Goldman A. L., Eagle W. V. I., Smibert C. A., et al. (2023) Smaug regulates germ plasm synthesis and primordial germ cell number in Drosophila embryos by repressing the oskar and bruno 1 mRNAs *bioRxiv*

Snee M. J., Macdonald P. M. (2004) Live imaging of nuage and polar granules: evidence against a precursor-product relationship and a novel role for Oskar in stabilization of polar granule components *J Cell Sci* 117:2109–2120

Song E. J., Yim S. H., Kim E., Kim N. S., Lee K. J. (2005) Human Fas-associated factor 1, interacting with ubiquitinated proteins and valosin-containing protein, is involved in the ubiquitin-proteasome pathway *Mol Cell Biol* 25:2511–2524

Song J., Park J. K., Lee J. J., Choi Y. S., Ryu K. S., et al. (2009) Structure and interaction of ubiquitin-associated domain of human Fas-associated factor 1 *Protein Sci* 18:2265–2276

Tendulkar S., Hegde S., Garg L., Thulasidharan A., Kaduskar B., et al. (2022) Caspar, an adapter for VAPB and TER94, modulates the progression of ALS8 by regulating IMD/NFkappaB- mediated glial inflammation in a Drosophila model of human disease *Hum Mol Genet* 31:2857–2875

Thibault S. T., Singer M. A., Miyazaki W. Y., Milash B., Dompe N. A., et al. (2004) A complementary transposon tool kit for Drosophila melanogaster using P and piggyBac *Nat Genet* 36:283–287

Thomson T., Liu N., Arkov A., Lehmann R., Lasko P. (2008) Isolation of new polar granule components in Drosophila reveals P body and ER associated proteins *Mech Dev* 125:865– 873

Thulasidharan A., Garg L., Tendulkar S., Ratnaparkhi G. S. (2024) Age-dependent dynamics of neuronal VAPB(ALS) inclusions in the adult brain *Neurobiol Dis* 196

Tomancak P., Beaton A., Weiszmann R., Kwan E., Shu S., et al. (2002) Systematic determination of patterns of gene expression during Drosophila embryogenesis *Genome Biol* 3

Valanne S., Wang J. H., Ramet M. (2011) The Drosophila Toll signaling pathway *J Immunol* 186:649–656

Vanzo N., Oprins A., Xanthakis D., Ephrussi A., Rabouille C. (2007) Stimulation of endocytosis and actin dynamics by Oskar polarizes the Drosophila oocyte *Dev Cell* 12:543–555

Wang L., Ligoxygakis P. (2006) Pathogen recognition and signalling in the Drosophila innate immune response *Immunobiology* 211:251–261

Weiszmann R., Hammonds A. S., Celniker S. E. (2009) Determination of gene expression patterns using high-throughput RNA in situ hybridization to whole-mount Drosophila embryos *Nat Protoc* 4:605–618

Wilk R., Hu J., Blotsky D., Krause H. M. (2016) Diverse and pervasive subcellular distributions for both coding and long noncoding RNAs *Genes Dev* 30:594–609

Williams M. J. (2007) Drosophila hemopoiesis and cellular immunity *J Immunol* 178:4711–4716

Wojcik E. J., Glover D. M., Hays T. S. (2000) The SCF ubiquitin ligase protein slimb regulates centrosome duplication in Drosophila *Curr Biol* 10:1131–1134

Wu J., Akhmanova A. (2017) Microtubule-Organizing Centers *Annu Rev Cell Dev Biol* 33:51–75 Ye Y. (2006) Diverse functions with a common regulator: ubiquitin takes command of an

AAA ATPase *J Struct Biol* 156:29–40

Zaessinger S., Busseau I., Simonelig M. (2006) Oskar allows nanos mRNA translation in Drosophila embryos by preventing its deadenylation by Smaug/CCR4 *Development* 133:4573–4583

Zeng Z., de Gorter D. J., Kowalski M., ten Dijke P., Shimmi O. (2014) Ter94/VCP is a novel component involved in BMP signaling *PLoS One* 9

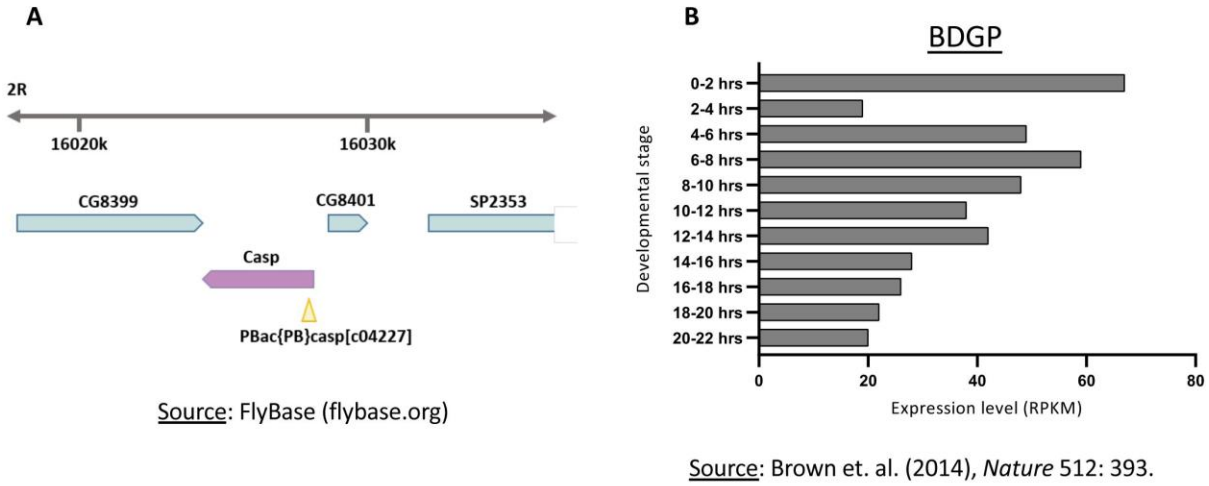
Zhang L., Zhou F., van Laar T., Zhang J., van Dam H., et al. (2011) Fas-associated factor 1 antagonizes Wnt signaling by promoting beta-catenin degradation *Mol Biol Cell* 22:1617–1624

Zhang Y., Werling U., Edelmann W. (2012) SLiCE: a novel bacterial cell extract-based DNA cloning method *Nucleic Acids Res* 40

Zhang Y., Werling U., Edelmann W. (2014) Seamless Ligation Cloning Extract (SLiCE) cloning method *Methods Mol Biol* 1116:235–244

Zhang Z., Krauchunas A. R., Huang S., Wolfner M. F. (2018) Maternal Proteins That Are Phosphoregulated upon Egg Activation Include Crucial Factors for Oogenesis, Egg Activation and Embryogenesis in Drosophila melanogaster *G3 (Bethesda)* 8:3005–3018

4.9 Supplementary materials



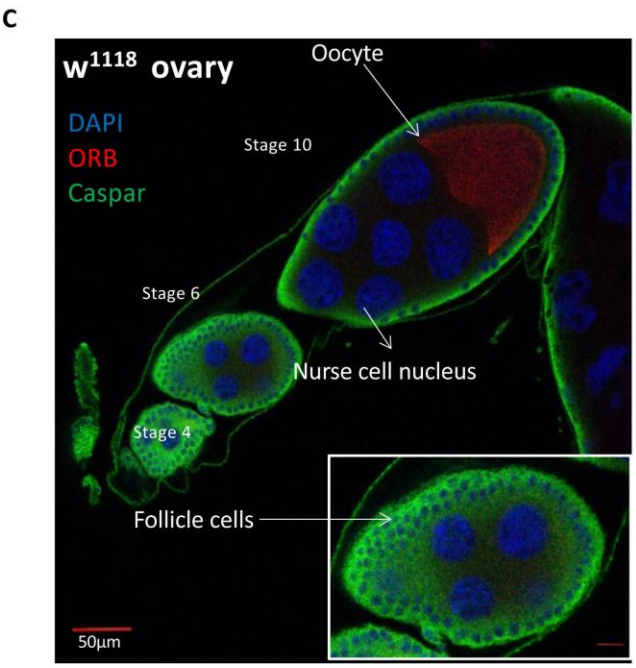


Fig. S1.

(A) Schematic of the the chromosomal location of the *casp* locus and the associated transposon insertion that generates the *caspc04227* allele. This allele is referred to as *casplof* hereafter. **(B)** RNA expression data from the modENCODE database, plotted as a bar chart **(C)**. DAPI marks the nuclei, orb marks the oocyte. Casp marks the follicle cells, but is not significantly enriched in the germline (egg, nurse cells).

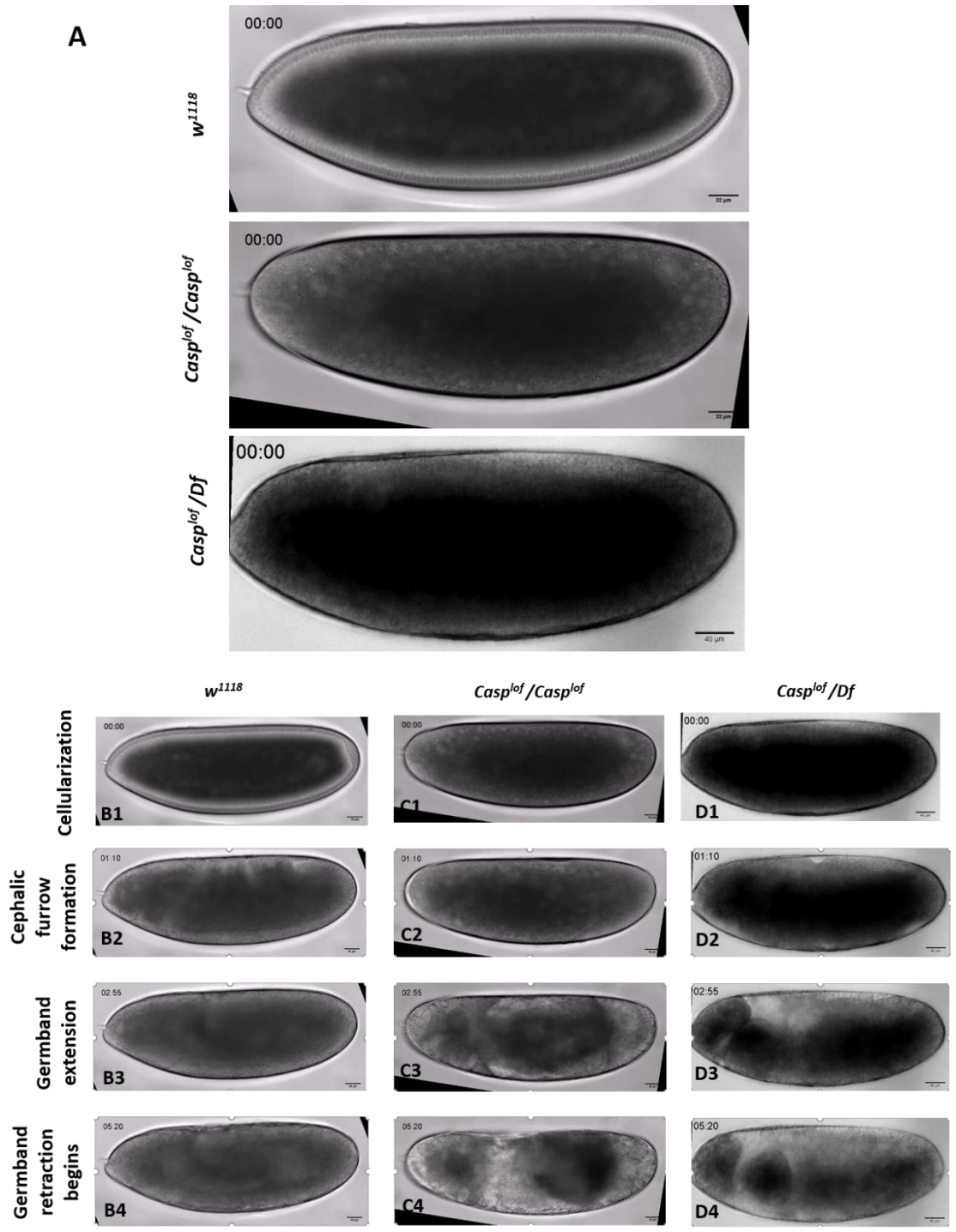


Fig. S2. *Casp* assists in progression to gastrulation.

(A) Representative movies for the early development of the indicated genotypes are shown in D.

(B-D) Time-lapse images are presented for the wild-type and mutants in B1-D4. Out of 96 time points imaged, four are shown for each genotype. Images were acquired from a single, constant plane unique to each embryo across 5-minute intervals. The developmental trajectory for w^{1118} can be followed from the cellularization stage to the beginning of germ band retraction, for the time points indicated. Crucial developmental milestones are highlighted (B1-B4). *casp* mutants did not exhibit characteristic stages of morphogenesis at comparable time points (C1-C4, D1-D4), n = 10.

A

Protein	Peptide count
casp	73.5
TER94	18
Vap33	2.5
Tudor-SN	2.25
alphaTub85E	2
eIF4A	2
Npl4	2
SERCA	1.75
bel	1.75
CCT2	1.5
mRpS14	1.5
CCT6	1.5
Hacl	1.25
mAcon1	1.25
ATPsyngamma	1.25
mRpL27	1.25
alphaTub67C	1.25
eIF3b	1
ScsbetaA	1
26-29-p	1
lig	1
me31B	1
nocte	1
Ufd1-like	1

C

Protein	Known function in the germline	Reference
TER94	polar granule component; associated with vasa and Tudor	Thomson et al., 2008
Tudor-SN	regulates spermatogenesis	Ku et al., 2016
eIF4A	polar granule component; associated with vasa and Tudor	Thomson et al., 2008
belle	required for female fertility and maintenance of male germline stem cells	Johnstone et al., 2005
CCT6	expressed in primordial germ cells	Fisher et al., 2012
α-Tubulin at 67C	regulation of male germline stem cells, centrosome positioning and segregation	Tao et al., 2021
Lingerer	interacts with germline component vasa	Durdevic and Ephrussi, 2019
me31B	polar granule component; associated with vasa and Tudor	Thomson et al., 2008

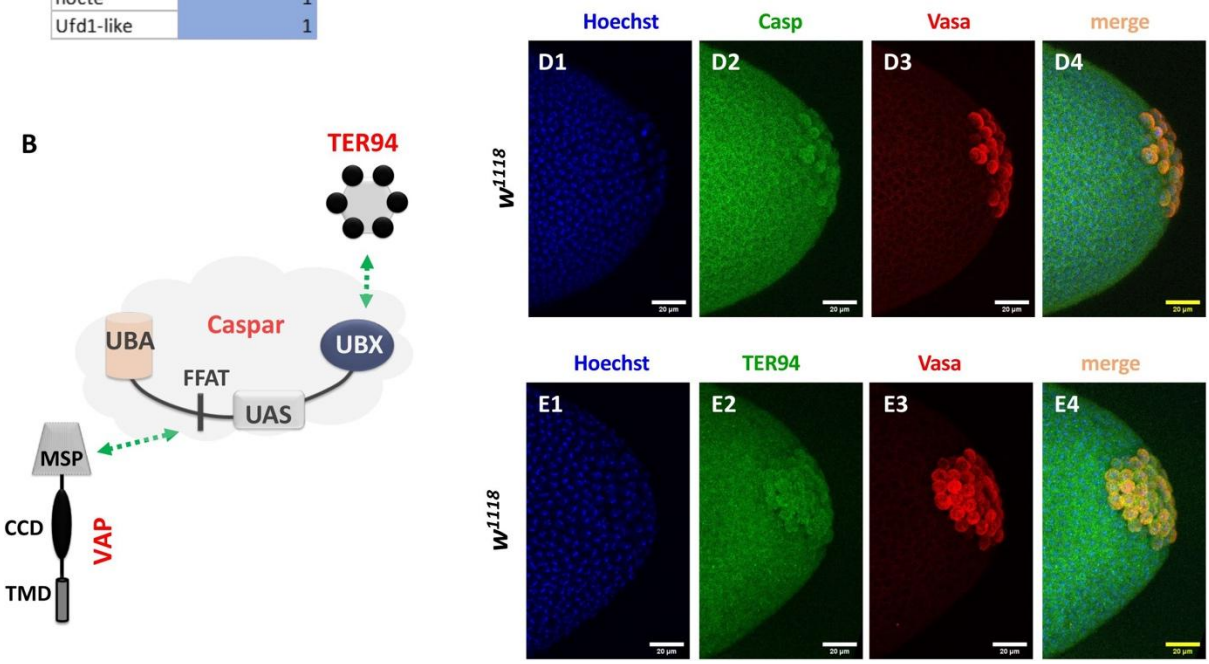


Fig. S3. TER94 and germline components are major interactors of Casp.

(**A**) shows a list of the 24 proteins enriched after a Casp IP, followed by mass spectrometry, listed in the order of their peptide counts. The peptide counts are averaged from 4 biological replicates. A subset of interactors with known functions or expression in the germline are tabulated in (**B**) A schematic of the interaction of Casp with both TER94 and VAPB (Tendulkar et al., 2022). TER94 interacts with the Casp UBX domain, while VAPB with the FFAT motif (**C**) List of proteins that are associated with Casp and are previously identified as pole-cell components (**D**-**E**) Immunostaining of wild-type nuclear cycle 13/14 embryos with Casp (D2) and TER94-specific (E2) antibodies indicates expression in the

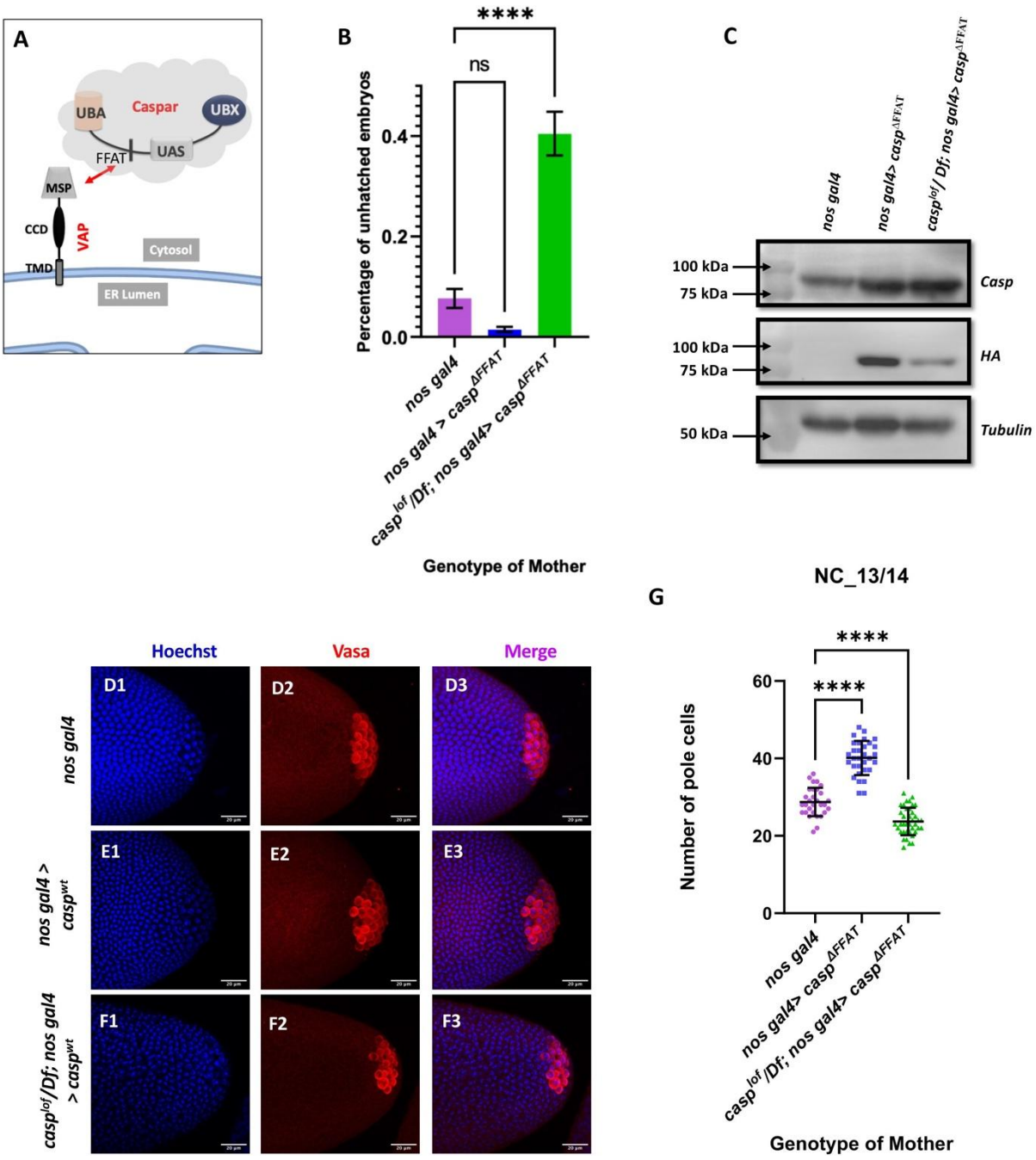


Fig S4. Casp:VAP interaction is not required for defining pole cell number.

(**A**) Casp interacts with VAPB via its FFAT-like motif (**B**) represents embryonic lethality at 25° C plotted as a bar graph. The genotype of the mother is listed on the X-axis. N = 3. n = 200, ordinary one-way ANOVA, (****) P <0.0001. (**C**) Casp protein levels were assessed in 0-3h embryos via western blotting. Rabbit anti Casp (1:10,000) and rabbit anti HA (1:2000) antibodies were used to probe the blot. Mouse anti Tubulin (1:10,000)is used as a loading control. (**D-F**) Confocal microscopy images of the posterior of nuclear cycle 13/14 embryos from mated *nos gal4* (D1-D3), *nos gal4> caspΔFFAT* (E1-E3), and *casplof/Df; nos gal4> caspΔFFAT* (F1-F3) females immunostained with vasa antibody (1:50). Hoechst marks the nuclei. (**G**) The germ cells at the posterior marked by vasa quantified and plotted as bar graphs. n = 30-40.

Chapter 4: Conclusions and future directions

Early embryonic development in *Drosophila melanogaster* is a tightly regulated process orchestrated by maternally deposited factors during oogenesis, which set the stage for body axis formation, germ cell specification, and subsequent cellular differentiation. Before fertilization even occurs, maternal RNAs and proteins establish the blueprint for the anterior-posterior (AP) and dorsoventral (DV) axes of the embryo. Key maternal effect genes—*bicoid* and *hunchback* for anterior patterning, *caudal* and *nanos* for posterior development—are asymmetrically localized within the egg, creating concentration gradients that later drive the spatial expression of zygotic genes. The DV axis, by contrast, is defined by the position of the oocyte nucleus, which dictates the dorsal expression of *gurken* and subsequent activation of the EGF receptor *torpedo* in adjacent follicle cells. In the absence of this signal, *pipe* is expressed, initiating a cascade that activates the Toll pathway. Its effector, Dorsal (DL), translocates into nuclei along a gradient and activates ventral-specific genes while repressing dorsal fate genes.

This spatial patterning of the embryo is facilitated by a transition in developmental control from maternal gene products to zygotic gene expression, known as the maternal-to-zygotic transition (MZT). It is during this phase that somatic sex identity and the segregation of germline cells from somatic lineages are also determined. Germ cell precursors, or pole cells, are specified around nuclear cycle 9 when a few nuclei migrate into the posterior germ plasm, a specialized cytoplasm enriched with determinants such as Oskar. This migration marks the beginning of germ-soma segregation and is essential for the formation of future gonads.

Overlaying this transcriptional and translational regulation is a sophisticated layer of post-translational control, particularly via SUMOylation—the covalent attachment of the Small Ubiquitin-like Modifier (SUMO) protein to target substrates. SUMOylation fine-tunes the activity, localization, and stability of key developmental regulators during early embryogenesis. The presence of enriched maternal SUMO and its conjugation machinery in the early *Drosophila* embryo underscores the significance of this reversible modification during this critical window of development. Functional studies have shown that disruption of SUMOylation leads to a range of developmental defects, highlighting its essential role in coordinating gene regulatory networks.

Among the key players subject to SUMOylation are Groucho and Dorsal. Groucho, a transcriptional co-repressor involved in terminal patterning, is SUMOylated and contains a SUMO-interacting motif (SIM). Although mutants resistant to SUMO conjugation show no gross defects in development, biochemical studies reveal that SUMOylation is required for Groucho to interact with Dorsal, suggesting that SUMOylation facilitates critical protein-protein interactions that influence axis patterning. Despite multiple approaches, we were unable to detect SUMOylation of Groucho either in cell culture or in vivo. Future studies could benefit from enriching SUMOylated protein fractions via immunoprecipitation using FLAG-tagged constructs in 529SU cells, followed by detailed expression analysis. Interestingly, mutation of Groucho's SUMO-interacting motifs (SIMs) leads to a marked reduction in protein half-life, though the underlying mechanism remains unclear. Groucho functions as a well-established transcriptional co-repressor, acting in concert with HDACs to repress various transcription factors, including Dorsal. Further research should aim to dissect how SUMOylation and SIM-mediated interactions influence Groucho's transcriptional regulatory activity and its downstream effectors, particularly during early embryonic development. Structural modeling of Groucho, especially in the context of SIM mutations, may reveal whether these alterations contribute to protein destabilization and functional impairment.

Similarly, Dorsal's stability and activity are closely tied to both phosphorylation and SUMOylation. SUMOylation at lysine 382 enhances its transcriptional potency, while dimerization—enabled by the SIMβ motif—is essential for its nuclear stability and function. Mutations in SIMβ prevent Dorsal dimerization, leading to its destabilization and degradation via a novel pathway, ultimately resulting in embryonic lethality and disrupted DV axis formation. Our experiments consistently revealed that Dorsal (DL) exists in more than two phosphorylated isoforms, indicating the presence of additional phosphorylation sites beyond the known residues S312 and S317. We ruled out S389, identified in a phosphoproteomic screen, as a phosphorylation site. Previous studies (Drier et al., 1999) had already eliminated S70, S79, S103, and S312 as potential targets. Another candidate, S665, was highlighted in a separate phosphoproteomic study (Hilger et al., 2009) and warrants further investigation as a potential phosphorylation site. The mechanism underlying DL degradation remains unclear, particularly whether it proceeds via the 26S proteasome or autophagy. Furthermore, the identity of the kinase(s) responsible for DL phosphorylation is still unknown. The possible crosstalk between SUMOylation and phosphorylation of DL also remains to be explored. Interestingly, our data show that the DLSCR variant is highly active and exhibits elevated phosphorylation

levels, though the mechanism driving this remains to be elucidated. Future studies should aim to identify the regulatory pathways and modifications responsible for DL activity and turnover.

In addition to these regulators, the protein Caspar, a *Drosophila* homolog of mammalian FAF1, emerges as a novel component influencing early development. Caspar is highly expressed in early embryos and in primordial germ cells, and its maternal depletion results in severe gastrulation defects and a reduced number of pole cells. Mechanistically, Caspar modulates the levels of key germ cell determinants such as Oskar and Smaug, possibly through a partnership with TER94 and a role in ubiquitin-mediated protein degradation during MZT. This regulatory role connects Caspar to both cytoskeletal integrity and germ cell fate specification, underlining the multifaceted nature of post-translational control in early development. Our findings clearly establish the involvement of Caspar (Casp) and TER94 in the formation and specification of primordial germ cells (PGCs). However, further investigation is required to define their precise molecular functions in this process. Notably, Casp remains the only protein reported to regulate germ cell fate and number by modulating both somatic signals and pole plasm components. The phenotypic parallels observed between the loss of ZGA components and Casp/TER94 suggest convergence on a common regulatory axis involving BMP signaling and protein degradation pathways. Future studies should aim to elucidate the mechanistic interplay between Casp/TER94, BMP signaling, and proteostasis. Additionally, exploring the role of zygotic factors such as Smaug may provide critical insights into how maternal and zygotic regulatory networks intersect during early embryogenesis.

In summary, early embryogenesis in *Drosophila* is not solely a result of gene expression patterns but is profoundly shaped by post-translational modifications such as SUMOylation. These modifications ensure precise temporal and spatial control of protein function, stability, and interaction. The interplay between maternal effect genes, signaling pathways, and SUMO-regulated protein networks governs not only the formation of body axes and cell identities but also the critical transition from maternal to zygotic control. This study underscores the importance of SUMOylation in embryonic patterning and germline specification, offering insights that extend beyond *Drosophila* to broader principles of developmental biology.