

Identification and Functional Analysis of Genes Controlling Germ Cell Migration and Gonad Formation

Dissertation

der Mathematisch-Naturwissenschaftlichen Fakultät
der Eberhard Karls Universität Tübingen
zur Erlangung des Grades eines
Doktors der Naturwissenschaften
(Dr. rer. nat.)

vorgelegt von
Ratna Tripathy
aus Cuttack

Tübingen
2013

Tag der mündlichen Qualifikation:

13.06.2013

Dekan:

Prof. Dr. Wolfgang Rosenstiel

1. Berichterstatter:

Dr. Andrew D. Renault

2. Berichterstatter:

Prof. Dr. Rolf Reuter

to Ved...

... you are always there with me

Acknowledgements

The successful journey through my doctoral thesis would not have been possible without the direct and indirect support of many people. Although I deeply thank all of them, unfortunately I can only mention a few of them here.

I owe my deepest gratitude to Dr. Andrew Renault for not only providing me with this project but also for his guidance, continuous support, and inspiration, enabling the completion of this work. I would also like to thank him for helping me develop a deep understanding of the subject. I am further indebted to Prof. Dr. Rolf Reuter for his extraordinary support as my university supervisor. Moreover, I would like to express my gratitude to the members of my thesis advisory committee, Prof. Dr. Gerd Jürgens, Prof. Dr. Rolf Reuter and Dr. Bernard Moussian, for their invaluable advice and for their scientific guidance. I also owe my thanks to Dr. Uwe Irion for discussions and helpful suggestions.

I would like to deeply thank all members of the Renault lab, past and present, for their insightful discussions, their help in troubleshooting everyday experiments and for a great lab atmosphere. In particular, I thank Dr. Kristina Ile and Amrita Mukherjee for patient and careful proofreading of the thesis. I thank Tina Bresser and Matthias Schneider for spending their time and effort in the translation of the abstract. I am deeply grateful to the two technicians, past and present, Michaela Schuppe and Jennifer Bergmann, for so efficiently handling the lab organization.

I thank Dr. Christa Lanz for her help in the library preparation and Illumina whole genome sequencing. I also thank Dagmar Sigurðardóttir and Simon Perathoner for helping in all the technical aspects of thesis writing.

My heartfelt gratitude goes out to all my friends here in Tübingen, in Europe and in India, who have been with me and made my journey through the past few years more enjoyable. I would also like to thank Christian Wiesner, not only for his help in the formatting of the thesis, but also for his encouragement and unquestionable support through my PhD and the writing process.

In the end, I would like to thank my family, particularly my parents and my sister who have continuously supported me, encouraged me, and have been there for me through all the circumstances.

Summary

Gonad formation in *Drosophila* involves the intricate movements of two cell types, the germ cells and the somatic gonadal precursor cells (SGPs), to generate a simple three-dimensional structure. The germ cells must firstly migrate to SGP clusters, which in turn fuse and ensheath the germ cells and the two cell types ultimately coalesce. Underlying this seemingly simple process is an array of genes required to guide germ cells to the SGPs. Additional complex transcriptional networks in SGPs are essential for their specification, maintenance and to initiate their ensheathment behavior. This system represents a good model to study organogenesis using essentially two cell types. In this thesis I have examined aspects of both germ cell migration and SGP behavior.

Firstly I examined the role of the *Drosophila* lipid phosphate phosphatases, Wun and Wun2 (referred to collectively as Wunens), in guiding the migration of germ cells. These multi-pass transmembrane enzymes, expressed both in the germ cells and the embryonic somatic cells, are thought to create an extracellular gradient of lipid phosphate, which is required for both migration and survival of the germ cells. However in order to form this extracellular lipid gradient, Wunens would be expected to act as ecto-enzymes. To investigate the site of requirement of Wun2 activity, I examined the localization of endogenous and different tagged versions of Wun2. Although both the GFP and the Myc tagged Wun2 were functional, the GFP version replicated the endogenous Wun2 plasma membrane localization.

In parallel I made a construct in which a TEV protease cleavage site was inserted in a predicted extracellular loop of Wun2. I demonstrated that secretion of TEV protease by neighboring cells reduced the ability of this construct, to rescue *wunen* loss of function in germ cells. Taken together these data support the notion that the cell surface localization and the ecto-phosphatase activity of Wunens is critical for germ cell migration and survival.

Secondly I have characterized the role of three mutants in which gonad formation is defective. In these mutants, although early germ cell migration is unaffected, in later embryos germ cells are scattered and SGP clusters do not fuse or compact. Whole genome sequencing coupled to deficiency complementation was performed to determine the causative mutations in two of the mutants. I identified mutations in two transcription factors, Midline (Mid) and Longitudinals lacking (Lola) and showed that both genes are expressed by SGPs and are required cell autonomously.

The mutation in *lola* affects only one (*lola-R*) of the 30 predicted *lola* isoforms. Whilst the SGP defects resulting from mutation of *Lola-R* could be rescued by expression of this *Lola* isoform, another functional *Lola* isoform (*Lola-BC*) was unable to do so. Therefore, this study identifies the first *lola-R* specific allele, and describes a role in gonad formation that cannot be substituted by other *Lola* isoforms. Additionally, I have shown that robust gonad expression of another transcription factor Traffic jam is lost in *mid* mutants, placing Mid into a SGP transcriptional cascade. Lastly, *mid* and *lola* interact genetically suggesting that they could have one or more overlapping downstream targets.

Zusammenfassung

Die Morphogenese der Gonaden in *Drosophila* beinhaltet die verworrene Bewegung zweier Zelltypen, der Keimzellen und der somatischen Gonadenvorläuferzellen (*somatic gonadal precursor cells*, SGPs), um schließlich eine einfache dreidimensionale Struktur zu bilden. Die Keimzellen müssen als erstes zu den SGP-Clustern migrieren, die wiederum die Keimzellen umschließen und sich mit ihnen zu einer Gonade verdichten. Diesem scheinbar einfachen Prozess liegen eine Reihe von Genen zugrunde, die notwendig sind, um die Keimzellen zu den SGPs zu leiten. Außerdem sind komplexe Transkriptionsnetzwerke in den SGPs notwendig, um deren Spezifizierung und ihre Erhaltung zu gewährleisten, sowie um die Umhüllung der Keimzellen zu initiieren. Dieses System stellt ein gutes Model dar, um die Organogenese mit nur zwei Zelltypen zu untersuchen. In dieser Arbeit habe ich sowohl Aspekte der Keimzellwanderung, als auch das Verhalten der SGPs untersucht.

Als erstes habe ich die Rolle der *Drosophila* Lipid-Phosphat-Phosphatasen Wun und Wun2 (zusammen bezeichnet als Wunen) bei der Steuerung der Keimzellmigration untersucht. Diese Transmembranenzyme, welche die Membran mehrfach durchspannenden, die sowohl in den Keimzellen, als auch in den somatischen Zellen des Embryos exprimiert werden, wird angenommen, dass sie einen extrazellulären Lipid-Phosphat-Gradienten aufbauen, der sowohl für die Migration, als auch für das Überleben der Keimzellen notwendig ist. Um jedoch diesen extrazellulären Gradienten zu erzeugen, würde erwartet werden, dass Wunen als Ektoenzyme agieren. Um zu ermitteln, an welchem Ort die Wun2-Aktivität benötigt wird, habe ich die Lokalisation von endogenen und unterschiedlich markierten Versionen von Wun2 untersucht. Obwohl sowohl GFP- als auch Myc-markiertes Wun2 funktional waren, hat nur die GFP-Version die endogene Wun2 Plasmamembran-Lokalisation gezeigt.

Außerdem habe ich ein Konstrukt hergestellt, in welches eine von der TEV-Protease erkannte Schnittstelle in eine prognostizierte extrazelluläre Schleife von Wun2 eingefügt wurde. Ich habe dabei gezeigt, dass die Sekretion von TEV-Protease durch benachbarte Zellen, die Fähigkeit dieses Konstrukts den *wunen* Funktionsverlust in den Keimzellen zu retten, verringert. Zusammengefasst unterstützen diese Daten die Auffassung, dass die Lokalisation an der Zelloberfläche und die Ekto-Phosphataseaktivität von Wunen entscheidend für die Keimzellmigration und deren Überleben sind.

Als zweites habe ich die Rolle von drei Mutanten charakterisiert, in welchen die Gonadenentwicklung fehlerhaft ist. Obwohl in diesen Mutanten die frühe Keimzellmigration unbeeinträchtigt ist, sind die Keimzellen in älteren Embryonen verstreut und die SGP-Cluster fusionieren oder komprimieren sich nicht. Mittels *whole-genome-sequencing* und Defizienz-Komplementation wurden die zugrundeliegenden Mutationen in zwei der Mutanten untersucht. Dabei habe ich Mutationen in zwei Transkriptionsfaktoren, Midline (Mid) und Longitudinals lacking (Lola), identifiziert und gezeigt, dass beide Gene von SGPs exprimiert werden und zellautonom erforderlich sind.

Die Mutation in *lola* beeinflusst lediglich eine (*lola-R*) der 30 prognostizierten *lola*-Isoformen. Während die SGP-Defekte, die aus der Mutation von *Lola-R* resultieren, durch die Expression dieser *Lola*-Isoform gerettet werden konnten, war eine andere funktionale

Lola-Isoform (Lola-BC) nicht dazu in der Lage. Demzufolge identifiziert diese Studie das erste *lola-R* spezifische Allel und beschreibt eine Rolle in der Gonadenentwicklung, die durch andere Lola-Isoformen nicht ersetzt werden kann. Darüberhinaus habe ich gezeigt, dass die stabile gonadische Expression eines anderen Transkriptionsfaktors, Traffic jam, in *mid* Mutanten verloren geht, was demzufolge Mid in eine SGP-spezifischen Transkriptionskaskade einfügt. Zuletzt interagieren *mid* und *lola* genetisch, was suggeriert, dass sie ein oder mehrere gemeinsame stromabwärts -liegende Ziele haben könnten.

Abbreviations

°C.....	degree Celsius
2R.....	right arm of chromosome 2
2L.....	left arm of chromosome 2
ABC	adenosine binding cassette
AEL.....	after egg laying
BAC.....	Bacterial artificial chromosome
BSA	bovine serum albumin
BTB.....	Broad complex, Tramtrack, Bric-a-Brac
C1P	ceramide-1-Phosphate
CDS.....	coding sequence
CHO.....	Chinese hamster ovary
CNS	central nervous system
CXCR	chemokine Receptor
DAB	diaminobenzidine
DAPI.....	4',6-diamidino-2-phenylindole
DE-cadherin	<i>Drosophila melanogaster</i> E-cadherin
EB.....	elution buffer
EMS	ethane methyl sulfonate
FACS.....	fluorescent activated cell sorter
FasIII.....	Fascilin III
FLP	Flippase
FP	forward primer
FRT	Flippase recognition target
GAL4	Galactose 4 transcription factor
GG.....	geranyl-geranylation
GLC.....	germline clone
GPCR	G protein coupled receptor
GSC.....	germline stem cell
JNK.....	Jun N-terminal kinase
LPP	lipid phosphate phosphatases
LPT	lipid phosphotransferases
LPA	lysophosphatidic acid
LB	luria Broth
msSGP.....	male specific somatic gonadal precursor cells
NA.....	numerical apperture
PA	phosphatidic acid
PBS	phosphate buffered saline
PBST	PBS with 0,1% Tween
PCR	polymerase chain reaction
PLD.....	Phospholipase D

PS..... parasegment
 RP.....reverse primer
 RT.....room temperature
 S1Psphingosine 1 phosphate
 SDFstromal derived factor
 SGPsomatic gonadal precursor
 SK1.....Sphingosine kinase 1
 SMSsphingomyelin synthases
 SNP.....single nucleotide polymorphism
 SP.....signal peptide
 SPPsphingosine phosphate phosphatases
 Tbx.....T-box
 TEV.....Tobacco etch virus
 UASupstream activating sequence
 VNC.....ventral nerve cord
 Wunens.....Wun and Wun2 collectively

Drosophila genes

abdA.....*abdominal A*
abdB.....*abdominal B*
arm.....*armadillo*
cli.....*clift*
dl.....*delta*
Dsix4.....*Drosophila six4*
ena.....*enabled*
eya.....*eyes absent*
foi.....*fear of intimacy*
fpps*farnesyl pyrophosphate synthase*
fra.....*frazzled*
ftz*fushitarazu*
ggt1.....*geranylgeranyl transferase type 1*
hid*head involution defective*
hmgcr*Hmg-CoA-Reductase*
lola.....*longitudinals lacking*
Ly*lyra*
mdr49.....*multi drug resistance 49*
mid.....*midline*
N.....*Notch*
nos*nanos*
pim.....*pimples*

qm*quemao*
robo*roundabout*
Sb*stubble*
shg.....*shotgun*
sli.....*slit*
spir.....*spir*
thr.....*three rows*
tin.....*tinman*
tj.....*traffic jam*
tre1*trapped in endoderm 1*
twi.....*twist*
w*white*
wg.....*wingless*
wun*wunen*
wun2*wunen2*
zfh1.....*zinc finger homeodomain 1*

Contents

1. Introduction	1
1.1 Germ cells	2
1.2 Cell migration	2
1.2.1 <i>Drosophila</i> germ cell migration	3
1.2.2 Genes involved in germ cell migration.....	5
1.3 LPPs in <i>Drosophila</i> germ cell migration	8
1.3.1 Structure and function of LPPs.....	8
1.3.2 <i>Drosophila</i> LPPs: <i>wunen</i> and <i>wunen 2</i>	9
1.3.3 <i>wunens</i> affect germ cell migration and survival.....	10
1.3.4 Model: Competition between somatic and germ line <i>Wunens</i> for a common substrate.....	12
1.3.5 Differential localization of mammalian LPPs	13
1.4 Gonad Morphogenesis	15
1.4.1 Genes involved in SGP specification.....	17
1.4.2 Genes involved in SGP cluster fusion.....	18
1.4.3 Genes involved in SGP ensheathment	19
1.4.4 Genes involved in gonad compaction and coalescence	19
2. Results	23
2.1 Investigation of the site of <i>Wunen2</i> activity required in the germ cells for their survival	24
2.1.1 Endogenous <i>Wun2</i> localizes to the plasma membrane of pericardial cells.....	24
2.1.2 <i>Wun2</i> protein shows tag-specific localization in germ cells	25
2.1.3 Using TEV protease to reveal the site of <i>Wun2</i> activity	28
2.1.4 The region of <i>tev</i> -site insertion is relatively unstructured.....	29
2.1.5 <i>Wun2</i> with a <i>tev</i> -cleavage site insertion is functional	31
2.1.6 Using the <i>nanos</i> promoter to ensure expression of constructs at physiological levels	32
2.1.7 Germ cells do not show high levels of endocytosis	33
2.1.8 Mesodermal expression of secreted TEV protease reduces the ability of <i>Wun2tev</i> to rescue the death of <i>wunen</i> null germ cells	35
2.2 Analysis of EMS mutants showing defective gonad morphogenesis	39
2.2.1 Gonad formation mutants showing normal early germ cell migration	39
2.2.2 Mutants display normal SGP specification.....	40
2.2.3 At later stages mutant gonads show SGP-related defects	42
2.2.4 Mesoderm derived tissues are fusion-competent in the mutants	44
2.2.5 Complementation screen approach to determine the locus responsible for the <i>A44</i> , <i>B23</i> , and <i>C28</i> mutant phenotypes.....	45
2.2.6 <i>Df(2R)BSC595</i> uncovers the <i>C28</i> causative locus	46
2.2.7 <i>Df(2R)BSC595</i> uncovers the <i>B23</i> causative locus	49
2.2.8 Whole genome sequencing of <i>A44</i> , <i>B23</i> and <i>C28</i> mutants to identify the causative mutations	50
2.2.9 Combining complementation data with the whole genome sequencing SNP results.....	56
2.2.10 The <i>C28</i> mutation is an allele of <i>longitudinals lacking</i>	58
2.2.11 The <i>B23</i> mutation is an allele of <i>midline</i>	60
2.2.12 Sanger sequencing reaffirms the <i>mid[B23]</i> and <i>lola[C28]</i> causative SNPs.....	61
2.2.13 Mutation in <i>mid[B23]</i> leads to a truncation in the T-box domain of the protein 63	

2.2.14 <i>lola</i> [C28] mutation affects the R-isoform of Lola	65
2.2.15 <i>lola</i> -R displays a unique pattern of expression compared to the remaining <i>lola</i> isoforms	67
2.2.16 <i>mid</i> [B23] but not <i>lola</i> [C28] shows defective VNC formation	69
2.2.17 <i>lola</i> -R but not <i>lola</i> -BC expression is observed in the gonad	70
2.2.18 Midline and Lola protein can be detected in the SGPs	72
2.2.19 <i>twiGal4</i> drives expression in the SGPs	74
2.2.20 <i>lola</i> -R but not <i>lola</i> -BC can rescue the <i>lola</i> [C28] mutant gonad defects	75
2.2.21 Expression of <i>midline</i> in the mesoderm rescues the gonad phenotypes in <i>mid</i> [B23] mutants.....	78
2.2.22 Stage specific loss of Traffic Jam in <i>mid</i> [B23] is rescued by mesodermal Midline expression.....	80
2.2.23 DE-cadherin, Armadillo and Robo are expressed in <i>mid</i> [B23] and <i>lola</i> [C28] mutant gonads	82
2.2.24 Tinman is expressed in SGPs but remains unaffected in <i>mid</i> [B23] mutant gonads	86
2.2.25 <i>mid</i> [B23] and <i>lola</i> [C28] genetically interact	88
3. Discussion.....	91
3.1 Requirement of cell surface activity of Wun2 for germ cell migration and survival	92
3.1.1 Differential localization of the tagged-versions of Wun2	92
3.1.2 Cleavage of Wun2 at the inserted tev-site	94
3.1.3 Implication of cell surface Wun2 activity in germ cells.....	95
3.2 Transcription factors regulating several steps of gonad morphogenesis	96
3.2.1 Normal mesodermal patterning and SGP defects in mutants.....	97
3.2.2 Severity of the EMS induced mutation in <i>mid</i> [B23] and <i>lola</i> [C28].....	98
3.2.3 Severity of phenotype in <i>lola</i> [C28] homozygotes vs. transheterozygotes.....	99
3.2.4 Identification of an isoform specific role for Lola	100
3.2.4 Stage specific loss of Traffic Jam expression in <i>mid</i> [B23] mutant gonad.....	101
3.2.5 Interaction between <i>midline</i> and <i>lola</i>	102
3.2.6 Genes unaffected in gonads of <i>mid</i> [B23] and <i>lola</i> [C28]	102
3.2.7 Other possible players in the gonad acting downstream of Mid and Lola	105
3.2.8 Implication on vertebrate gonad morphogenesis.....	106
4. Materials and Methods.....	109
4.1 Materials	110
4.1.1 Fly husbandry	110
4.1.2 Apple juice plate recipe	110
4.1.3 <i>Drosophila</i> strains	110
4.1.4 Antibodies.....	113
4.1.5 Oligonucleotides	114
4.2 Methods in molecular biology	116
4.2.1 Polymerase chain reaction	116
4.2.2 TOPO® cloning	116
4.2.3 Transformation of chemically competent cells	117
4.2.4 Sequencing	117
4.2.5 Gateway reaction.....	118
4.3 Methods using <i>Drosophila</i> embryos.....	119
4.3.1 Methanol fixation	119
4.3.2 Antibody staining of <i>Drosophila</i> embryos	120

4.3.3 Genomic DNA isolation for PCR	120
4.3.4 RNA Isolation	121
4.3.5 cDNA synthesis.....	121
4.3.6 Synthesis of riboprobes.....	122
4.3.7 <i>In-situ</i> hybridisation in <i>Drosophila</i> embryos.....	124
4.3.8 Making Transgenic Flies.....	125
4.4 Insertion of <i>tev</i>-site in <i>wun2</i> gene	126
4.5 Whole genome sequencing	130
4.5.1 Genomic DNA isolation from mutants.....	130
4.5.2 Library Preparation.....	130
4.5.3 Validation of Library	132
4.5.4 Cluster preparation on flowcell and SBS Sequencing by Synthesis sequencing.	133
4.5.5 SHORE analysis of obtained reads	134
4.5.6 Analysis of sequences using CLC Genomic workbench	134
4.5.7 Manual curation	137
4.6 Microscopy	138
4.6.1 Confocal Microscopy	138
4.6.2 Light Microscopy	138
Contributions	139
Appendix A.....	141
The Bloomington deficiency kit for chromosomal arm 2L.....	141
The Bloomington deficiency kit for chromosomal arm 2R	144
Appendix B.....	147
Fly cross schematic for rescue of <i>wunen</i> loss function in germ cells by <i>nos>wun2con/tev</i> in the presence of the TEV protease	147
Fly cross schematic for rescue of <i>lola</i> mutant gonad phenotype with Lola-R-GFP expression.....	148
Fly cross schematic for rescue of <i>mid</i> mutant gonad phenotype with Mid expression.....	149
Bibliography.....	151

Chapter 1

Introduction

1.1 Germ cells

In most multicellular organisms, germ cells are cells with the unique capability of giving rise to new individuals of the same species. These cells are formed and set aside from the rest of the soma very early in embryonic development. At this stage, they also differ greatly from the somatic cells with respect to their cellular content, migratory behavior and transcriptional repression (Seydoux and Braun, 2006). The quiescent transcription of germ cells, achieved through the action of various maternally deposited mRNAs, prevents their trans-differentiation into the somatic lineage (Van Doren et al., 1998c). Hence, germ cells are the ultimate stem cells, with an ever-lasting lineage. Because of these unique properties, the embryonic germ cells of *Drosophila melanogaster* provide a great system for studying two basic cell biological processes:

- Single cell migration
- Organ formation

In *Drosophila*, as in many other organisms, germ cells after being specified have to migrate along a stereotyped path, meet the somatic gonadal precursor cells (SGPs) and together form the embryonic gonads. Migratory cues are necessary and often central to the process of germ cell migration. Attractive cues, such as those provided by SDF1 in zebrafish (Doitsidou et al., 2002); (Knaut et al., 2003) and mouse (Ara et al., 2003); (Molyneaux et al., 2003), or by Hmgcr in flies (Van Doren et al., 1998b), seem to play an important role in guiding the germ cells to the gonads.

1.2 Cell migration

Cell migration is a basic biological process required by many cell types during their development. Migration is needed in early developing embryos during gastrulation to lay out the future body plan of the organism. Therefore cells may not require migration for their inherent function but instead to come together in the

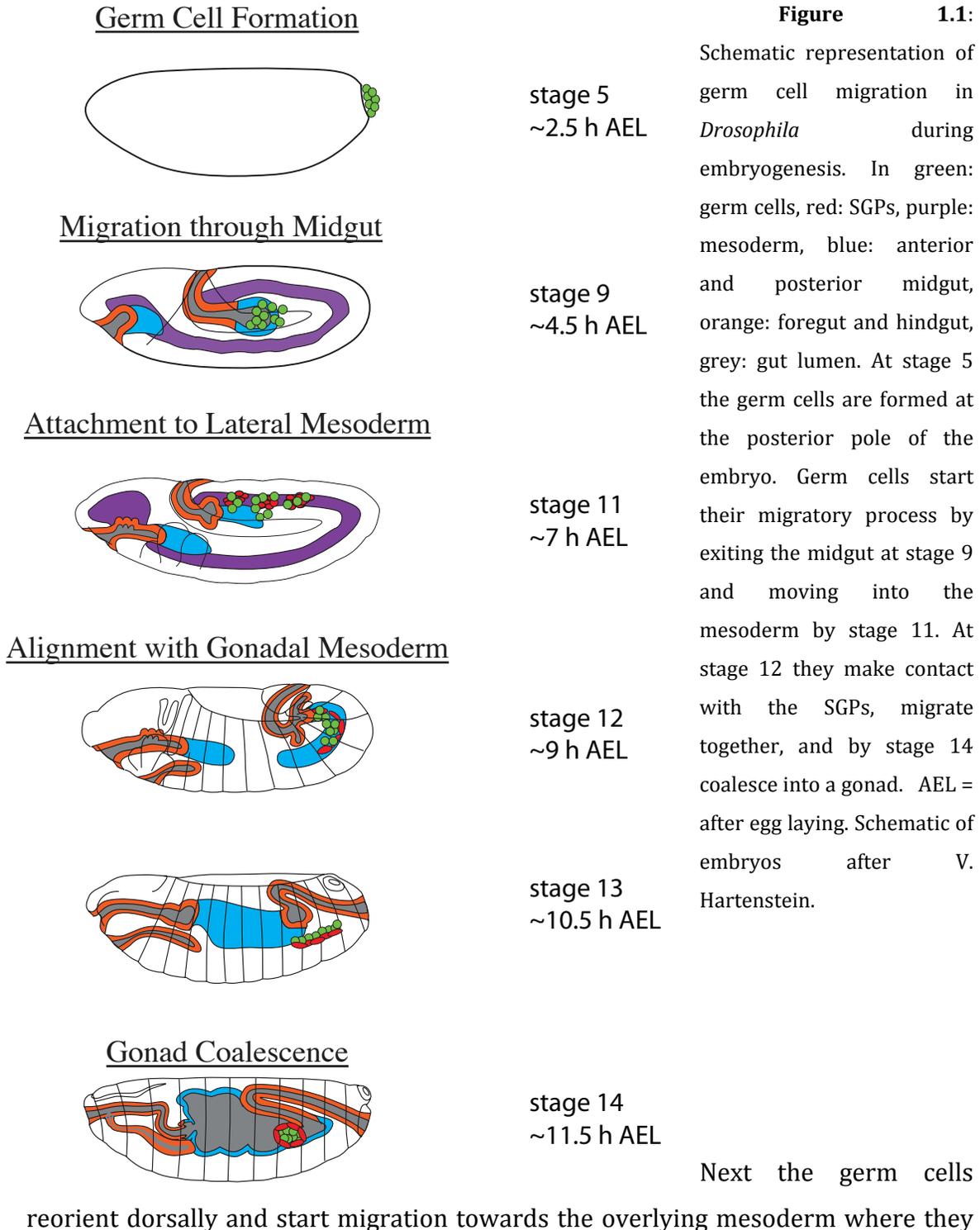
right place and form a functional organ or tissue. Moreover, this process is equally important in a mature organism for its role in immune defense, wound healing and maintenance of gut and skin epithelia. For these cells, migration is necessary for their inherent activities.

However, besides these innate processes, migration also plays a key role in various diseases and pathological conditions. This includes metastasis of cancer cells, infiltration of immune cells in asthma and migration of vascular smooth muscles cells in atherosclerosis. Therefore, understanding the various aspects and mechanisms of migration is vital for understanding both basic developmental processes and pathological conditions.

1.2.1 *Drosophila* germ cell migration

In *Drosophila* the germ cells are the first cells to form during embryogenesis (Fig. 1.1) (Starz-Gaiano and Lehmann, 2001). They form at the posterior tip of the developing embryo, which is rich in many maternally deposited mRNAs and proteins (Mahowald et al., 1976). The germ cells arise by the cellularization of the nuclei at the posterior pole taking in the germ plasm. At this stage, the germ cells appear to show migratory behavior with visible cytoplasmic protrusions (Jaglarz and Howard, 1995). These cells then divide to give rise to approximately 40 germ cells. Once gastrulation begins, they are passively swept into the embryo along with the midgut primordium to which they are attached. During this movement, the germ cells cease their cytoplasmic protrusions and appear spherical in shape, suggesting their passive movement into the midgut pocket (Kunwar et al., 2008). Subsequently, around late stage 9, the germ cells start their active migratory process. This is also the stage when the germ cells become transcriptionally active (Van Doren et al., 1998d). The clustered germ cells start individualizing, detaching from the midgut primordia as well as each other, to undergo trans-epithelial migration (Kunwar et al., 2008). They become highly polarized with their leading edge oriented towards the midgut epithelium and migrate through the midgut pocket (Kunwar et al., 2008).

However, the polarized germ cells are dependent on the endoderm remodeling occurring at this stage, and cannot pass through intact midgut epithelia (Seifert and Lehmann, 2012).



split into two bilateral clusters on either side of the midline (Sano et al., 2005). This step ensures that both of the gonads that will form at the either side of the embryo will be populated. The two clusters of germ cells continue their migration along with somatic gonadal precursor cells (SGPs). By stage 13 they stop their migration and align with the SGPs. At the end of stage 14 the two cell types coalesce to give rise to the embryonic gonads (Boyle and DiNardo, 1995). There the germ cells proliferate to give rise to the germ line stem cells.

1.2.2 Genes involved in germ cell migration

A number of genes have been identified as being involved in germ cell migration, eliciting diverse functions to regulate this complicated process. They include those that are expressed in the germ cells or in the soma. Some of the encoded proteins are required for making and sensing of attractive and repulsive cues, to aid the migrating germ cells along the correct migratory route. Some of the key genes are discussed below.

➤ *tre-1* in transepithelial migration

The onset of germ cell polarization and subsequent migration at stage 9 is controlled by *tre-1*, or *trapped in endoderm 1* (Kunwar et al., 2003) (Kunwar et al., 2008). This gene encodes for a G-protein coupled receptor (GPCR) expressed by the germ cells (Fig 1.2), whose transcript is maternally provided (Kunwar et al., 2003). In embryos laid by females that are mutant for this gene, the germ cells are unable to individualize and undergo transepithelial migration to exit the posterior midgut (Kunwar et al., 2003). Although the ligand for this GPCR is still unknown, it is believed that signaling through Tre-1 leads to the redistribution of a G-protein β subunit, the GTPase Rho1, DE-cadherin (DE-cad) and other adherens junction components to the lagging edge of the germ cells. DE-cad, encoded by the *shotgun* gene (*shg*), is a major component of the adherens-junctions, which are the primary junctions required for cell-cell adhesion and interaction. Therefore, its down regulation helps the germ cells polarize and separate into individual cells (Kunwar

et al., 2008). Recent studies indicate that Tre-1 activated germ cells still need normal epithelial remodeling to allow their migration through the midgut (Seifert and Lehmann, 2012).

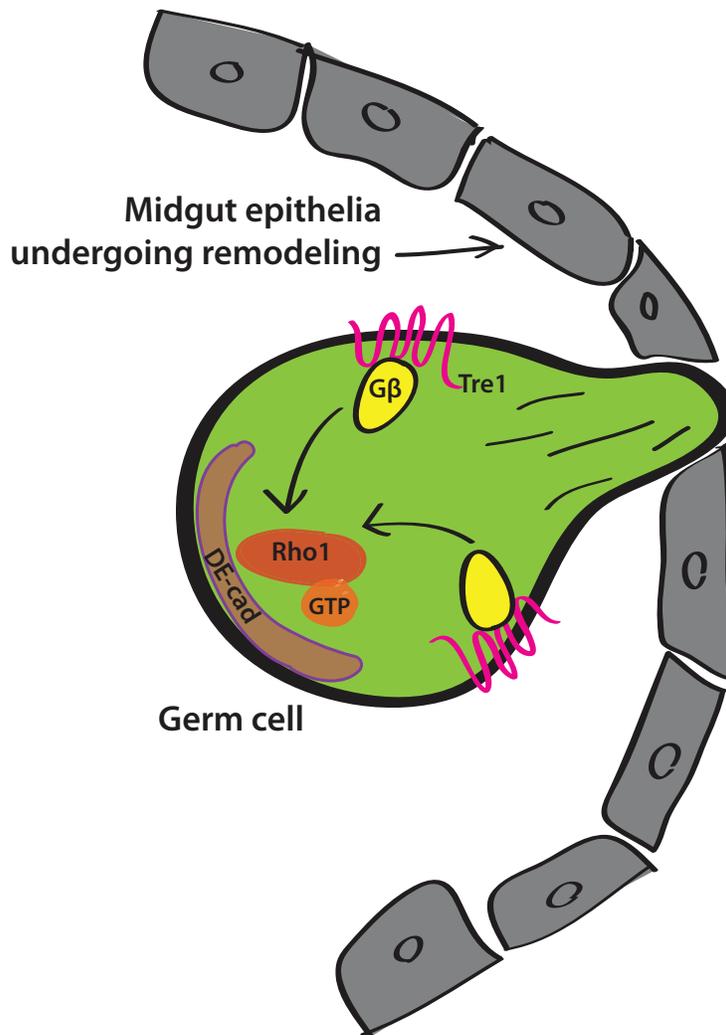


Figure 1.2: Tre1, a G-protein coupled receptor drives germ cell transepithelial migration through the midgut pocket. Tre1 (pink) signals through the G-protein β -subunit (yellow) and Rho1 GTPase (orange), and leads to the redistribution of adherens junction protein like DE-cadherin (brown) to the lagging edge of the germ cell (green). This leads to cytoskeletal changes in the germ cells required for migration. The polarized germ cell then orients towards the midgut epithelia (gray), which is undergoing remodeling during this stage, and exits the midgut pocket. Adapted from (Richardson and Lehmann, 2010)

➤ **wunen in mesodermal migration**

Following exit from the midgut pocket, the germ cells need to be guided into their next target tissue, the mesoderm. This guidance cue comes from the Wunens, and their mechanism of action is discussed in detail in sections 1.3.2-1.3.4 (Zhang et al., 1997); (Starz-Gaiano et al., 2001).

➤ **hmgcr in association with SGPs**

Hydroxymethylglutaryl-CoA Reductase (Hmgcr) is an important and rate limiting enzyme of the cholesterol biosynthesis pathway (Goldstein and Brown, 1990). Zygotic mutants for *hmgcr* show germ cell migration defects (Van Doren et al., 1998b). The germ cells fail to associate with the somatic gonadal precursor cells, instead scattering in the posterior of the embryo. *In situ* hybridization shows enrichment of *hmgcr* expression the SGPs. Moreover, ectopic expression of *hmgcr* in the ectoderm or the CNS is sufficient to attract germ cells to those tissues, implying that expression of this enzyme is sufficient for germ cell attraction (Van Doren et al., 1998b).

Many of the enzymes downstream of Hmgcr, like farnesyl-diphosphate synthase (Fpps), geranyl-geranyl-diphosphate synthase (Quemao) and geranylgeranyl transferase type I (GGT1), which are involved in attaching geranyl-geranyl (GG) groups onto proteins (Fig 1.3), also show germ cell migration phenotypes (Santos and Lehmann, 2004). It is speculated that the attractive cue is a geranyl-geranylated peptide that is released by the SGPs through the ABC transporter Mdr49 expressed by SGPs (Ricardo and Lehmann, 2009).

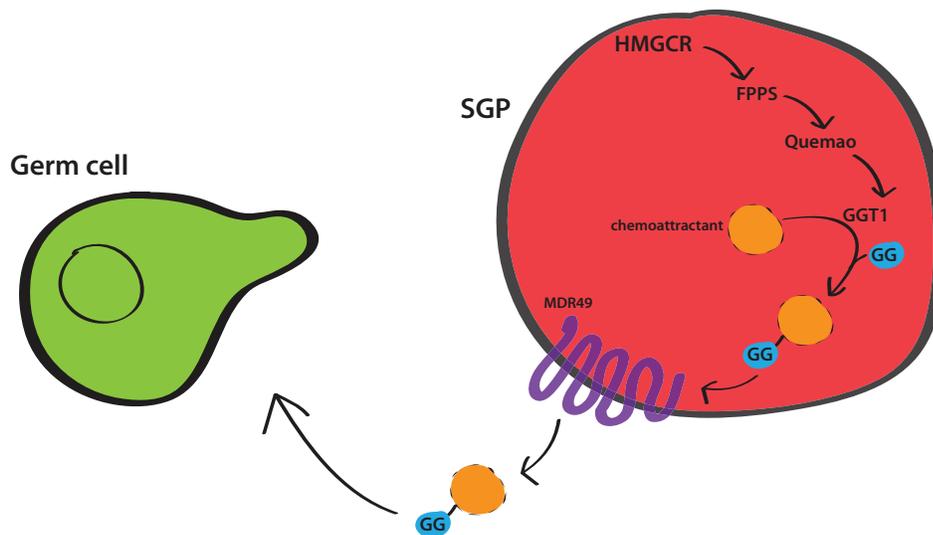


Figure 1.3:

The HMGCR pathway is required for attracting germ cells to the SGPs. HMGCR and the downstream enzymes like FPPS, Quemao and GGT1 are thought to be required in the SGPs (red) for the production of a

geranyl-geranylated (blue) chemoattractant (orange) in the SGPs, which is then secreted by the ABC transporter MDR49 (purple) for the germ cells (green) to sense. Adapted from (Richardson and Lehmann, 2010)

1.3 LPPs in *Drosophila* germ cell migration

However, repulsive mechanisms are also needed to work in conjunction with the attractive cues, to prevent the germ cell from drifting into tissues that are not the target tissue. For example, CXCR7B, expressed in the soma of zebrafish, causes internalization and destruction of the SDF1 germ cell chemoattractant. This leads to the formation of a gradient of SDF1, and prevents mis-migration of germ cells. In *Drosophila* germ cell migration, a similar destruction of migratory cue is performed by a lipid phosphate phosphatase (LPP).

1.3.1 Structure and function of LPPs

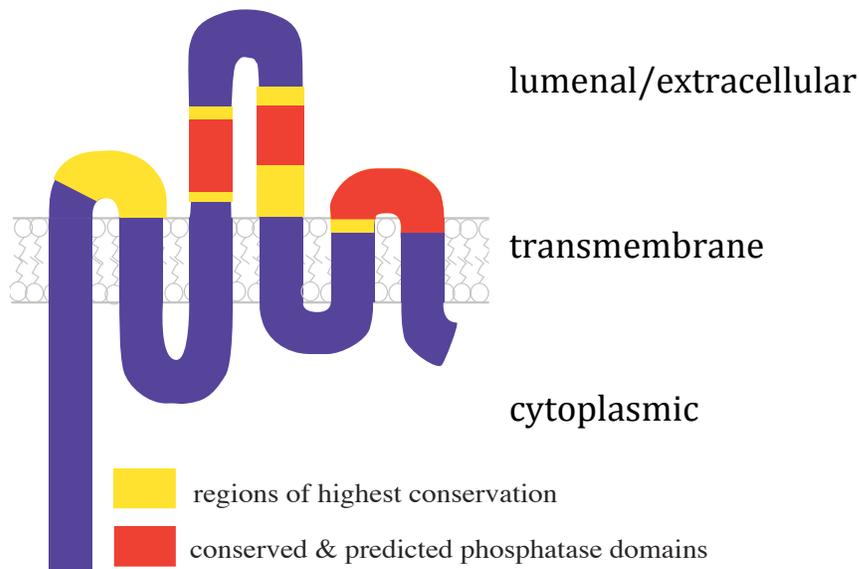
LPPs belong to the family of LPTs (lipid phosphatases/phosphotransferases) along with SPPs (sphingosine phosphate phosphatases), SMSs (sphingomyelin synthases) and other enzymes involved in lipid phosphate metabolism (reviewed in (Sigal et al., 2005)). These proteins have similar predicted structures (Fig. 1.4), with six transmembrane spanning alpha helices connected by five extramembrane loops (Barila et al., 1996). These membrane proteins have flanking N- and C- terminal regions of varied lengths and sequence composition. There are three catalytic motifs in LPPs (Fig. 1.4): two contained in the third extramembrane loop and the last in the fifth extramembrane loop (Stukey and Carman, 1997). These catalytic motifs interact to catalyze the dephosphorylation of lipid phosphates. Some of the *in vitro* lipid substrates of LPPs are lysophosphatidic acid (LPA), phosphatidic acid (PA), sphingosine-1-Phosphate (S1P), and ceramide-1-Phosphate (C1P) (Roberts et al., 1998); (Waggoner et al., 1996).

There are 3 mammalian LPPs; LPP1, LPP2, and LPP3 encoded by the genes Ppap2a, Ppap2c, and Ppap2b, respectively (Roberts et al., 1998). Although these LPPs show overlapping activities and substrate specificity *in vitro*, their knockout phenotype in mice indicate non-redundant functions (Zhang et al., 2000; Escalante-Alcalde et al., 2003; Tomsig et al., 2009).

1.3.2 *Drosophila* LPPs: *wunen* and *wunen2*

The *Drosophila* genome encodes for eight LPPs, two of which, *wunen* (*wun*) and *wunen2* (*wun2*) (collectively referred to as *wunens*) show embryonic expression. These two genes are present next to each other on chromosome 2R, cytological position 45D, and are transcribed in opposite directions from a common promoter region leading to similar early embryonic expression patterns (Starz-Gaiano et al., 2001).

Like other LPPs, Wunens are also six transmembrane spanning proteins (Fig. 1.4). Expression of *wun2* in insect cell lines leads to the accumulation of fluorescently labeled lipid substrate PA within the cells (Renault et al., 2004b). This indicates uptake of the lipid by the cell following its dephosphorylation by Wun2. *In vitro*, the LPP substrates, PA, LPA and C1P, are also dephosphorylated by Wunens (Burnett and Howard, 2003; Renault et al., 2004b). However, their *in vivo* substrate is still unknown.



and unconserved residues in blue.

Figure 1.4: Wunens belong to the family of LPPs. The six transmembrane domain topology of the protein is shown here with region of highest conservation shown in yellow, predicted catalytic phosphatase domains (extracellular) in red

1.3.3 *wunens* affect germ cell migration and survival

Wun and Wun2 were shown to play a role in germ cell migration through a deficiency screen (Zhang et al., 1996), and a misexpression screen (Starz-Gaiano et al., 2001), respectively. Loss of both genes is required to affect migration, implying their redundancy in this process.

In embryos deficient for both *wun* and *wun2* the germ cells fail to reach the mesoderm and scatter over the posterior of the embryo (Zhang et al., 1996). Conversely, ectopic expression of either *wun* or *wun2* in the surrounding somatic tissues leads to a dramatic reduction in germ cell number (Starz-Gaiano et al., 2001). Therefore somatic *wunen* expression affects both germ cell migration (in case of deficiency) and negatively regulates survival (in case of somatic overexpression).

Germ cells avoid areas of *wunen* expression during their migration suggesting that Wunens act via repulsion (Zhang et al., 1996; Zhang et al., 1997; Starz-Gaiano et al., 2001). It is proposed that the stage 10 posterior midgut expression orients the germ cells upon exit of the midgut to allow them to move into the overlying mesoderm, while the ectodermal CNS expression leads to the bilateral splitting of the germ cell population (Starz-Gaiano et al., 2001; Sano et al., 2005).

Wunens are also expressed in germ cells and this expression is necessary for germ cell survival. *wun* and *wun2* mRNA is maternally provided and for *wun2* this maternal transcript is clearly retained in the germ cells (Fig. 1.5) (Renault et al., 2002b).



Figure 1.5: In situ hybridization using a probe against *wun2* in a *Drosophila* stage 5 embryo (Renault et al., 2002b). The arrow points towards the germ cells at the posterior end of the embryo, which stain positively for *wun2* mRNA.

In maternal *wun2* mutant embryos, half of the germ cells survive and migrate normally to the embryonic gonads (Hanyu-Nakamura et al., 2004; Renault et al.,

2004b) whereas depletion of both maternal *wun* and *wun2* leads to near complete loss of germ cells during migration (Renault et al., 2010). Expressing a wild type version of Wun2 in otherwise *wun wun2* null germ cells fully rescues the germ cell death phenotype whereas a catalytically dead version of Wun2 is unable to do so (Renault et al. 2010). These experiments indicate a requirement for LPP phosphatase activity in germ cells for their survival (Renault et al., 2004b).

In the absence of both somatic and maternal *wun* and *wun2* expression, germ cells survive but mismigrate (Renault et al. 2010). Hence, loss of somatic *wunens* can rescue germ cell death caused by lack of germ cells *wunens*, but not the mismigration phenotype (summarized in Table 1.1). In addition, blocking germ cell migration also rescues the death of *wunen*-lacking germ cells. Therefore, the requirement of Wunens in germ cell survival can be eliminated by inhibiting germ cell migration (Renault et al., 2010).

Table 1.1: Effect of Wunen dosage on germ cell survival and migration. - indicates loss of function, + indicates wild type levels and +++ indicates overexpression.

Germ cells	Somatic cells	Phenotype
+	+	Proper migration and survival
+	-	Mismigration
+	+++	Death
-	+	Death
-	-	Mismigration

1.3.4 Model: Competition between somatic and germ line Wunens for a common substrate

The complex interplay between the somatic and germ cell *wunens* has been summarized into a model (Fig. 1.6) of how these *Drosophila* LPPs may be regulating germ cell migration and survival:

Both germ cells and somatic cells express *wunens*. Once the germ cells begin their migration, they move away from the tissues expressing somatic *wunens*. This is because the germ cell *wunens* and the somatic *wunens* compete with each other for a common lipid phosphate substrate. For germ cells, internalization of this lipid following its dephosphorylation by *wunens* is required as a survival signal. Somatic *wunens* act to deplete this survival factor hence making a lipid phosphate gradient in the embryo according to the somatic *wunens* expression pattern. Since the germ cells move towards higher lipid phosphate levels, they move away from these somatic tissues (Renault et al., 2004b).

When germ cells lack *Wunens*, they cannot receive the survival factor and they die. Similarly, on overexpression of *Wunens* in the soma, the germ cells are unable to compete with the somatic *Wunens* for the substrate, again leading to germ cell death. However, when the soma is depleted of *wunens* while the germ cell *wunens* is intact, the substrate is present for internalization by the germ cells but there is no longer a gradient of the lipid phosphate, causing the germ cells to mismigrate into other tissues. Moreover, there is also competition for the substrate between the germ cells as well, helping them to individualize and remain separate (Renault et al., 2010).

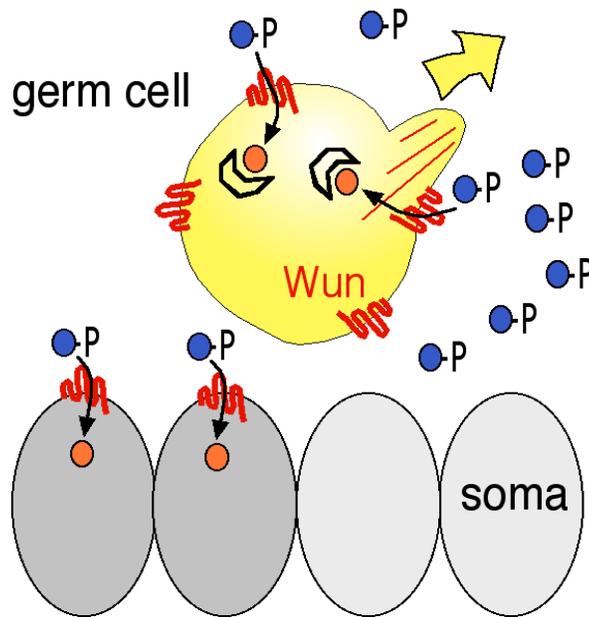


Figure 1.6: Model showing the role of *wunens* in germ cell migration and survival (Renault et al., 2004b)

Although this model provides a compelling explanation of the germ cell migration and survival phenotypes associated with changes in *wunen* levels in both germ and somatic cells, there are still some points left unexplained. First, what is this lipid phosphate substrate? Though Wunen can cleave a number of lipid phosphates *in vitro*, its *in vivo* substrate is still unknown. It is also not known which tissues produce and release this substrate helping to make the gradient. The mechanism of germ cell death, due to the lack of survival factor, is also not understood. It has been shown that wunen related germ cell death is *reaper*, *hid* or *grim* (pro-apoptotic *Drosophila* genes) independent (Renault et al., 2004b) and that these cells don't stain positively for active caspase-3 (Hanyu-Nakamura et al., 2004). Therefore, germ cell death occurs through an apoptosis-independent mechanism.

1.3.5 Differential localization of mammalian LPPs

LPPs are transmembrane proteins which can localize to various cellular compartments within the cell, including plasma membrane, and intracellular membranes like the endoplasmic reticulum and Golgi apparatus (Sigal et al., 2005).

Many studies (a few described below) show LPP localization to be dynamic and cell-type specific.

➤ **Ecto-activity of LPPs**

A key example of this is the endogenously expressed LPP1 in cultured human platelets. Exposure to LPA, leads to plasma membrane localization of LPP1 in these cells. Degradation of extracellular LPA by LPP1 leads to decreased signaling through the LPA-receptors, preventing platelet aggregation (Smyth et al., 2003).

3T3F442A preadipocytes express high levels of LPP1 and LPP3, and require their ecto-activity to regulate differentiation. Similar to the platelets, the LPP activity is required to control extracellular levels of LPA to regulate preadipocyte proliferation (Simon et al., 2002). Interestingly, on differentiation into adipocytes, these cells reduce their LPP expression by 80% (Simon et al., 2002). Therefore demonstrating the regulated ecto-phosphatase activity during adipose tissue development.

In another cell-type, human endothelial cells, ecto-LPP1 activity dephosphorylates extracellular S1P, to aid uptake of sphingosine by these cells. This is then reconverted to S1P by intracellular sphingosine kinase 1 (SK1) to activate internal signaling cascades (Barzik et al., 2005).

➤ **Intracellular LPP activity**

LPPs acting inside the cells can regulate cell signaling by altering the balance between a lipid and its phosphorylated form. The activity of LPP2 on phospholipase-D (PLD)-derived phosphatidic acid (PA), and LPP3 on S1P can regulate the balance between cell survival and apoptosis in Chinese hamster ovary (CHO) cells (Riechmann et al., 1998). Endogenous LPP3, seen during rest phase in cytoplasmic vesicles in CHO cells, translocates to the perinuclear compartment upon agonists induced activation of PLD (Riechmann et al., 1998).

In cultured migrating fibroblasts, it has been demonstrated that LPA induced LPA-receptor activity is required for LPP1 control of PA formation to inhibit fibroblast migration. Overexpression of LPP1 in these cells prevents LPA-induced

fibroblast migration but does not decrease extracellular LPA levels, indicating an intracellular requirement of LPP1 activity (Pilquill et al., 2006).

Therefore the subcellular localization of LPPs is key to their physiological roles. In *Drosophila* it is not known where the Wunen elicit their phosphatase activity to aid germ cell migration. Although the model of Wunen activity is based on its ecto-catalytic activity, there is no conclusive proof of it being required at the cell surface.

1.4 Gonad Morphogenesis

For the germ cells to reach their ultimate fate as gametes, they need supporting somatic cells. These cells, called the somatic gonadal precursor (SGP) cells, are formed bilaterally in the mesoderm as three separate clusters in parasegments (PS) 10, 11 and 12 (Brookman et al., 1992b), in embryonic stage 10. A fourth 'male-specific' SGP (msSGP) cluster is specified at this stage in PS13 in both sexes, but is eventually lost in the females and maintained only in males (DeFalco et al., 2003).

After splitting into two bilateral clusters the germ cells make contact with the SGPs (Boyle and DiNardo, 1995). Following contact formation, at stage 13 the SGPs extend cytoplasmic processes to surround and individualize the germ cells (Fig 1.7) in a process called 'ensheathment' (Jenkins et al., 2003). Simultaneously, the three SGP clusters along with the ensheathed germ cells fuse to form an elongated gonad (Boyle and DiNardo, 1995; Jenkins et al., 2003). Subsequently, the gonad commences its process of 'compaction' and begins to round up. At stage 14 the msSGP cluster joins the compacting gonad, and by stage 15 the embryonic gonad is formed in PS10 and appears as a round, tight ball-like structure (Boyle and DiNardo, 1995).

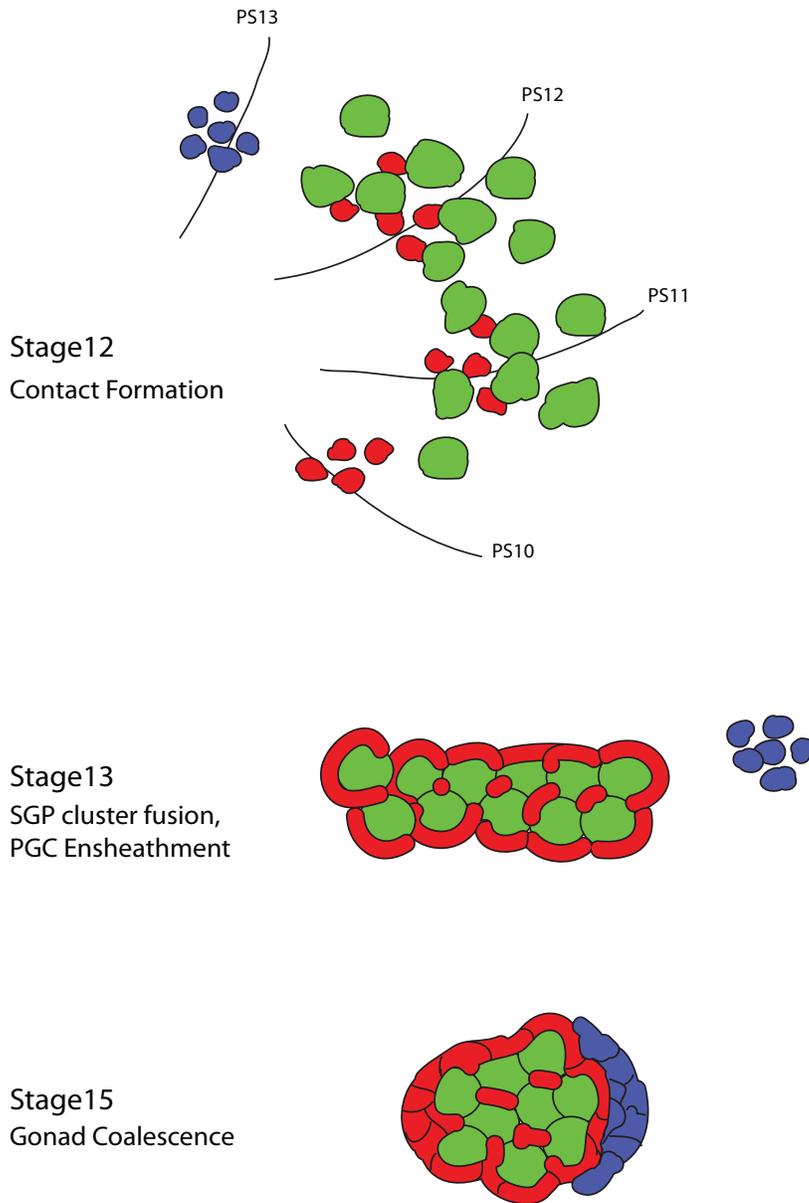


Figure 1.7: Schematic representation of gonad formation. In red: SGPs, purple: msSGPs, green: germ cells, the lines drawn denote PS10-13. The SGPs, which are specified in 3 separate clusters in PS10-12 and an additional male-specific cluster (msSGPs) in PS13, make contact with the germ cells at stage 12. By stage 13, these separate clusters of SGPs fuse to form a contiguous tissue. Simultaneously, the SGPs at this stage send out cytoplasmic protrusions to ensheath the germ cells and by stage 15 the two cell types have

come together to form the embryonic gonad. Only following compaction do the msSGPs fuse at the posterior of gonad.

By this time, the gonads are already sexually dimorphic. The msSGPs, which go on to form the terminal epithelium in the adult testis (Tokuyasu, 1974), are present only in the male gonads. Another sex-specific cell type is the pigment cells, also found in the male gonads first around late embryonic stage 17 (DeFalco et al., 2008), and thought to be required for testis morphogenesis in pupal stages (Nanda et al.,

2009). The anterior-most SGPs in the male embryonic gonad form the 'hub cells', which in turn signal to the anterior most germ cells in the gonad to form the germline stem cells (GSCs) (Le Bras and Van Doren, 2006). While in males, the first signs of spermatogenesis begins by the first larval stage (Aboïm 1945)(Le Bras and Van Doren, 2006), whereas by contrast the female germline niche formation doesn't begin until the late third instar larval stage (King, 1970).

1.4.1 Genes involved in SGP specification

SGPs arise from the dorsolateral mesoderm (Brookman et al., 1992b; Mainieri et al., 2006). Correct SGP specification therefore requires proper mesodermal specification. The mesoderm starts differentiating into different derivatives at stage 10. This can be observed with the exclusive dorso-lateral expression of a combination of genes including *tinman*, *bagpipe* (Azpiazu and Frasch, 1993), and *zinc finger homeodomain-1 (zfh1)* (Broihier et al., 1998), from where the SGPs will arise.

Although *zfh1* expression can be seen earlier in the mesoderm, it is upregulated in the gonadal mesoderm from stage 10 (Broihier et al., 1998). *tinman* (*tin*) expression is also restricted to the dorsal mesoderm in this stage (Azpiazu and Frasch, 1993). Mutants lacking both *tin* and *zfh1* have no SGPs, however single mutants of either of these genes still have some SGPs (Broihier et al., 1998).

Following proper dorsolateral mesoderm specification, the SGPs also need to be specified in clusters within this region. Two homeotic genes, *abdA* and *abdB*, work together in specifying the 4 clusters of SGPs. While *abdA* is expressed in a gradient, its highest level of expression is in the PS10 and 11 (Karch et al., 1990). This expression then leads to specification of the anterior SGPs (Brookman et al., 1992b). *abdB*, on the other hand, is expressed highest in PS13-14 (Delorenzi and Bienz, 1990), hence specifying the posterior SGPs (Brookman et al., 1992b). Consistent with the expression patterns, in *abdA* mutants there are no anterior SGPs formed and the 4th cluster of SGP derived from PS13 dies in the absence of the gonad (DeFalco et al., 2008), and *abdB* mutants lack posterior SGPs. Moreover, ectopic

expression of these genes can lead to generation of SGPs at the expense of fat body, indicating their instructive nature in SGP specification (Boyle and DiNardo, 1995; Moore et al., 1998b; Riechmann et al., 1998).

Another gene, *clift (cli)/eyes absent (eya)*, is expressed strongly in the SGP clusters by stage 11, and is a useful maker for SGPs (Boyle and DiNardo, 1995). While *eya* is not required for SGP specification, its expression is needed to maintain SGP fate. In *eya* mutants, the SGPs are lost soon after they are specified, and the germ cells scatter at the posterior of the embryo (Boyle et al., 1997). Yet another early expressing marker in the SGPs is the retrotransposon 412 (Brookman et al., 1992b). It is expressed at high levels in PS10-12 at stage 11, and its expression is maintained in the SGPs through gonad formation. Although expression of transposable elements is not uncommon in the gonad, any functional significance is unclear. Nevertheless, the 412 retrotransposon has proved to be a valuable reagent aiding the study of SGPs.

1.4.2 Genes involved in SGP cluster fusion

At stage 13, the three separate clusters of SGPs must fuse together to form one contiguous tissue. The failure of this process will lead to a several SGP-germ cell clumps, but no functional gonad and hence sterility.

In spite of the importance of SGP cluster fusion, few genes are known that are required for or regulate this process. A recent screen performed by (Weyers et al., 2011), identified several mutants which show a defect at this stage of gonad formation. Mutations in two genes, *three rows (thr)* and *pimples (pim)*, whose molecular functions are not known, showed what was described as fragmented gonads, with loosely structured SGPs, possibly due to loss of cell adhesion. The role of the Slit/Robo pathway was also illustrated in this screen. Slit is a secreted ligand for the Robo-family receptors, discovered first in its role in axonal path finding (Brose et al., 1999). There are 3 Robo receptors in *Drosophila*, Robo, Robo2 (or Leak) and Robo3, and loss of function of any of these receptors led to SGP fusion and compaction defects (Weyers et al., 2011). As expected, *slit* mutants also display

similar defects. Robo is localized at SGP-PGC junctions, while Robo2 is seen at both SGP-germ cell and SGP-SGP junctions (Weyers et al., 2011). However, no Slit protein was detected in the vicinity of the SGPs. Therefore the source of the Slit ligand remains unknown and whether it provides directional or permissive signals to the Robo-receptors in the gonads is unclear.

1.4.3 Genes involved in SGP ensheathment

Ensheathment of the germ cells is achieved through the cytoplasmic processes extended by the SGPs around the germ cells, to individualize them (Jenkins et al., 2003). A transcription factor, Traffic Jam (TJ), is expressed in the SGPs around stage 12-13, and is required for proper ensheathment to occur (Li et al., 2003). In *tj* mutants the SGPs are specified and undergo cluster fusion, however, they do not individualize the germ cells. This leads to reduced germ cell proliferation in later stages, and hence adult sterility (Li et al., 2003). The ovarian follicle cells of adult female *tj* mutants show increased DE-cad levels (Li et al., 2003), and it is not known where this is the case in the embryonic gonads of these mutants.

Another gene, *raw*, was also recently shown to be required in the process of germ cell ensheathment (Weyers et al., 2011; Jemc et al., 2012). Similar to *tj* mutants, *raw* mutants also show SGP ensheathment defects. Raw, whose molecular function is not unknown, works by negatively regulating JNK signaling, downstream of which cell-adhesion molecules like DE-cad and Arm (β -catenin) are regulated. In *raw* mutants, Arm, is mislocalized away from the cell surface (Jemc et al., 2012). Moreover, over-expression of DE-cad in the *Drosophila* gonad is sufficient to rescue the ensheathment phenotype (Jemc et al., 2012). Therefore, cell-adhesion molecules appear to be central to proper gonad formation.

1.4.4 Genes involved in gonad compaction and coalescence

On successful completion of the initial stages of germ cell-SGP interaction and cluster fusion, the two cell types involved must compact themselves into a round,

tight structure. The role of cell adhesion molecules is of great importance here. In the gonads, DE-Cad is expressed both by the germ cells and the SGP, and localized to their site of contact (Tepass et al., 1996; Jenkins et al., 2003). Mutants in *shg* show compaction and ensheathment defects. Moreover, in these mutants the msSGPs fail to join the gonad in later stages (Jenkins et al., 2003). A gene encoding a zinc ion transporter, *fear of intimacy (foi)*, exhibits similar gonadal defects (Van Doren et al., 2003). It was demonstrated that *Foi* functions upstream of DE-cad, to post-transcriptionally regulate its levels in germ cells and SGP but not msSGPs (Mathews et al., 2006). Recently, a role of Enabled (*Ena*), an actin regulator, was revealed in gonad coalescence, also through the regulation of DE-cad localization (Sano et al., 2012).

Table 1.2: Genes required in the various steps of gonad formation, with their molecular function.

Gene	Molecular Function	Role in Gonad Formation
<i>zinc finger homeodomain1 (zfh1)</i>	Transcription factor	SGP specification
<i>tinman (tin)</i>	Transcription factor	SGP specification
<i>abdominal A, B (abdA, abdB)</i>	Transcription factor	SGP specification
<i>clift (cli)/eyes absent (eya)</i>	Transcription factor	SGP specification
<i>three rows (thr)</i>	Unknown	SGP cluster fusion
<i>pimples (pim)</i>	Unknown	SGP cluster fusion
<i>roundabout, roundabout 2, 3 (robo, robo2, robo3)</i>	Cell surface receptor	SGP cluster fusion, Compaction
<i>slit (sli)</i>	Robo ligand	SGP cluster fusion, Compaction
<i>Traffic jam (tj)</i>	Transcription factor	SGP ensheathment
<i>raw</i>	Unknown	SGP ensheathment
<i>armadillo (arm)</i>	β -catenin	SGP ensheathment and coalescence
<i>Drosophila E-cadherin (DE-cad)/shotgun (shg)</i>	Adhesion molecule	SGP ensheathment and coalescence
<i>fear of intimacy (foi)</i>	Zinc ion transporter	SGP coalescence
<i>enabled (ena)</i>	Actin regulator	SGP coalescence

Although a number of genes are known to play a role in the process of gonad morphogenesis (summarized in Table 1.2), it is clear that many genes remain to be

identified. The regulation of DE-cad has been studied in some depth, but the mechanism of action of many of the remaining genes is unresolved. Understanding how this genetic network is regulated and interacts to effect gonad formation remains a major challenge.

Chapter 2

Results

2.1 Investigation of the site of Wunen2 activity required in the germ cells for their survival

Overexpressed Wunens in somatic cells have been reported to be largely cell surface localized in insect cell lines and *Drosophila* mesodermal cells (Starz-Gaiano et al., 2001; Burnett and Howard, 2003). These observations have prompted the assumption that in the current model of the role of Wunens in germ cell migration and survival (Fig 1.8), Wunens are acting at the cell surface. However, presence of cytoplasmic Wunens can be clearly seen in these cell types. Therefore, it is important to study the localization of Wunens in germ cells, which has not been examined so far.

2.1.1 Endogenous Wun2 localizes to the plasma membrane of pericardial cells

To look at the endogenous subcellular localization of Wun2, a Wun2 antibody was used to stain wild type and *wun2* null embryos. The most prominent expression in wild type embryos was in the heart. The *Drosophila* heart is a linear organ comprising of two rows of cardioblasts surrounded on either side by a row of pericardial cells. In particular, Wun2 expression was seen in the pericardial cells, but not the cardioblasts (Fig 2.1). Two lines of evidence suggested that this staining represented genuine Wun2 expression: 1) because no such staining was observed in *wun2^Δ* embryos, a RNA null mutant (Hanyu-Nakamura et al., 2004). 2) *wun2* mRNA is highly expressed in pericardial cells (Renault et al., 2002a). Wun2 was localized primarily on the cell surface, co-localizing with alpha-spectrin, which marks the plasma membrane (Fig 2.1 A', yellow arrow heads).

However, germ cells were not strongly stained using this antibody in spite of high levels of *wun2* mRNA present in the germ cells at stage 5 (Renault et al., 2002a) which endures until stage 10 (A. Mukherjee, unpublished results). This lack of

staining is possibly due to lower levels of Wun2 expression or the deeper position of the germ cells within the embryo, as compared to the heart, making staining more difficult.

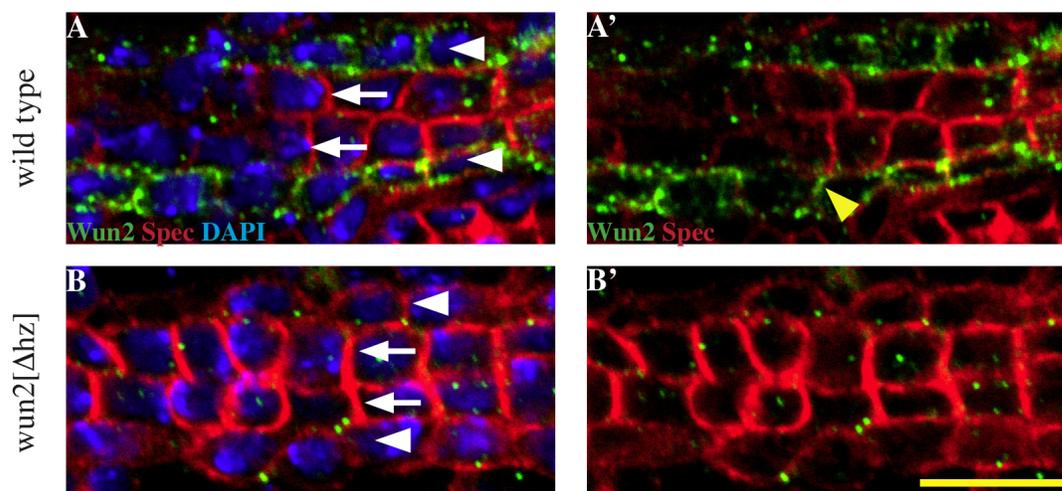


Figure 2.1: Endogenous Wun2 signal in the embryonic heart using a Wun2 antibody. The heart can be visualized as two rows of cardioblasts (white arrows) and two rows of pericardial cells (white arrow heads) on either side. In wild type hearts (A, A'), Wun2 (green) expression can be detected in the pericardial cells (B', yellow arrow head) co-localizing with the cell membrane marker alpha-spectrin (red). In *wun2^Δ* mutants (B, B') however, no Wun2 signal could be detected. Scale bar: 10 μ m.

2.1.2 Wun2 protein shows tag-specific localization in germ cells

To address the question of the localization of Wun2 in germ cells, tagged versions of the protein were used. While a functional Myc-tagged *wun2* construct was already available, a GFP-tagged form was synthesized and expressed in flies for this study. Both *UASwun2-myc* and *UASwun2-GFP*, were expressed specifically in the germ cells using the *nosGal4* driver. Antibodies against the respective tags were then used to visualize their localization in germ cells of stage 10-11 migrating germ cells. This is the stage where Wun2 activity is required for germ cell migration and survival.

Compared to the ubiquitous intracellular presence of the control GFP protein (Fig 2.2 A, A'), expression of Wun2GFP was observed as a ring surrounding the

cytoplasmic Vasa protein (Fig 2.2 B, B'). This demonstrates a plasma membrane localization of the GFP-tagged Wun2.

The localization of the Myc-tagged protein was, in contrast, observed to be mostly intracellular. Wun2myc largely co-localized with Vasa, however, some protein could be detected on the surface of the germ cells (Fig 2.2 C, C').

The GFP and Myc tags gave conflicting results as to the localization of the Wun2 protein. To determine if GFP tagged Wun2 was functional, this construct was used to rescue the death of *wun wun2* null germ cells. (Fig 2.2 E). While in wild type, an average of 20 germ cells were observed per embryo at stage 13, this number decreased drastically in embryos lacking maternal *wunens*. On re-expressing either Wun2GFP or Wun2myc in the germ cells of these mutants using *nosGal4*, germ cell number increased back to wild-type levels. Therefore GFP tagged Wun2 is equally functional as the Myc tagged version, in spite of their disparate patterns of localization. Therefore these data do not pinpoint the site where Wun2 activity is required.

To avoid possible tag related artifacts, I overexpressed un-tagged Wun2 in germ cells using the *nosGal4* driver and line *EP2650*, a UAS containing P-element insertion upstream of *wun2*. On using the Wun2 antibody, the overexpressed untagged Wun2 was detected at the plasma membrane of germ cells (Fig 2.2 D, D'). In addition, intracellular concentrations of Wun2 were often noted in the form of puncta in the germ cells (Fig 2.2 D', yellow arrowheads). Overall therefore, the GFP tagged construct behaves most similarly to the untagged protein with both showing the cell surface as the prominent site of accumulation, which likely represents the site of action. However, given that Wunens must traffic to the cell surface, and the observed intracellular puncta, it cannot be excluded that Wunens are active in intracellular vesicles.

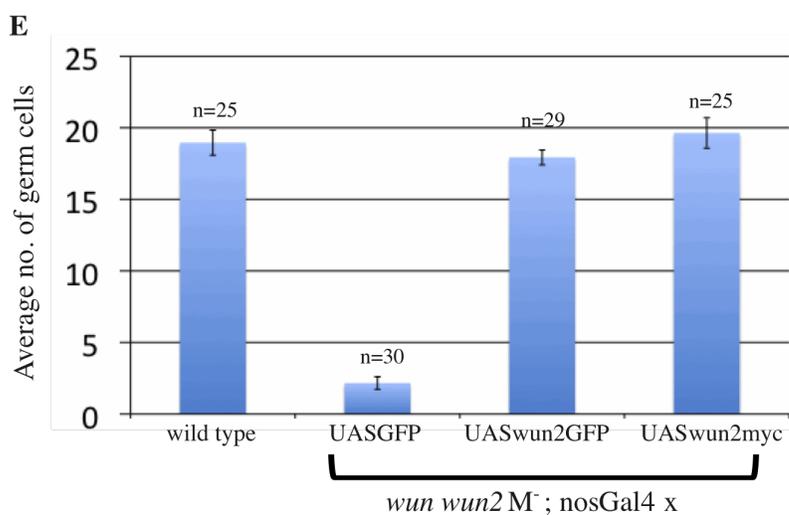
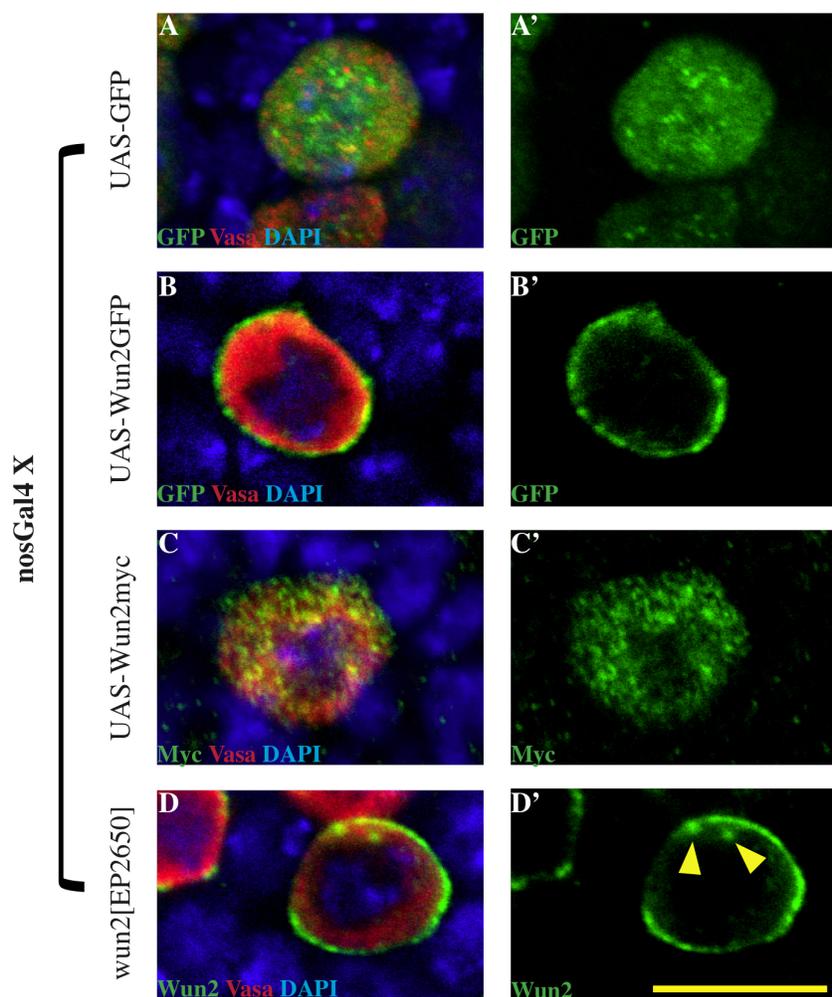


Figure 2.2: Localization and functionality of tagged Wun2. A-D: germ cells of *nosGal4* females crossed to males carrying the specified UAS constructs. A germ cell-specific marker, Vasa, is in red and the antibodies used against the tags or Wun2 are in green. Compared to the ubiquitous presence of GFP in *UAS-GFP* expressing germ cells (A, A'), Wun2GFP (B, B') localizes to the cell membrane of germ cells. Wun2myc (C, C') on the other hand is largely cytoplasmic with some cell surface localization. Untagged Wun2 expressed from the *EP2650* line (D, D'), is largely cell surface localized, with a few intense cytoplasmic puncta (yellow arrow heads). Scale bar: 10µm. The graph (E) shows the average germ cell number at stage 13 in embryos from *wun wun2*

M⁻ mutants expressing various UAS rescue constructs. In spite of the differential localization of Wun2GFP and Wun2myc, both are equally capable of rescuing germ cell death in these mutants. Error bars show the standard error, and 'n' represents number of embryos scored.

2.1.3 Using TEV protease to reveal the site of Wun2 activity

To distinguish between the cell surface and intracellular compartments as the sites critical for Wunen activity, a different approach was taken. The TEV (tobacco etch virus) protease has been used previously to cleave proteins intracellularly in *Drosophila* among other systems (Harder et al., 2008; Pauli et al., 2008; Svendsen et al., 2009; Formaz-Preston et al., 2012). The TEV protease recognizes a 7 amino acid motif (Glu-Asn-Leu-Tyr-Phe-Gln-Gly), hereafter referred to simply as a tev-site, and cleaves between Gln and Gly (Greig and Akam, 1995).

Therefore a TEV protease cleavable form of Wun2 was synthesized. A tev-site was inserted in the 3rd loop of Wun2, facing the luminal or extracellular domain (Fig 2.3). The tev-site was placed between the two catalytic domains (Fig 2.3, C₁ and C₂) contained within this loop. Exposure of the cleavable form of Wun2 (Wun2tev) to the TEV protease should lead to cleavage at the tev-site. This would presumably lead to the separation of the two catalytic domains, rendering the Wun2 enzyme catalytically inactive. Given that Wun2 catalytic activity is required for germ cell survival (Starz-Gaiano et al., 2001), the TEV cleaved enzyme would not be expected to be functional and should not be able to rescue the germ cell death of germ cells lacking wunens.

In order to be able to test the cell surface activity of Wun2, a secretable form of the TEV protease was kindly provided by Reinhard Schuh (MPI for Biophysical Chemistry, Göttingen). This UAS-promoter driven transgene contained a wingless-signal peptide (wgSP) to enable secretion of the protease outside the cell, and a mCherry-tag to verify expression and localization.

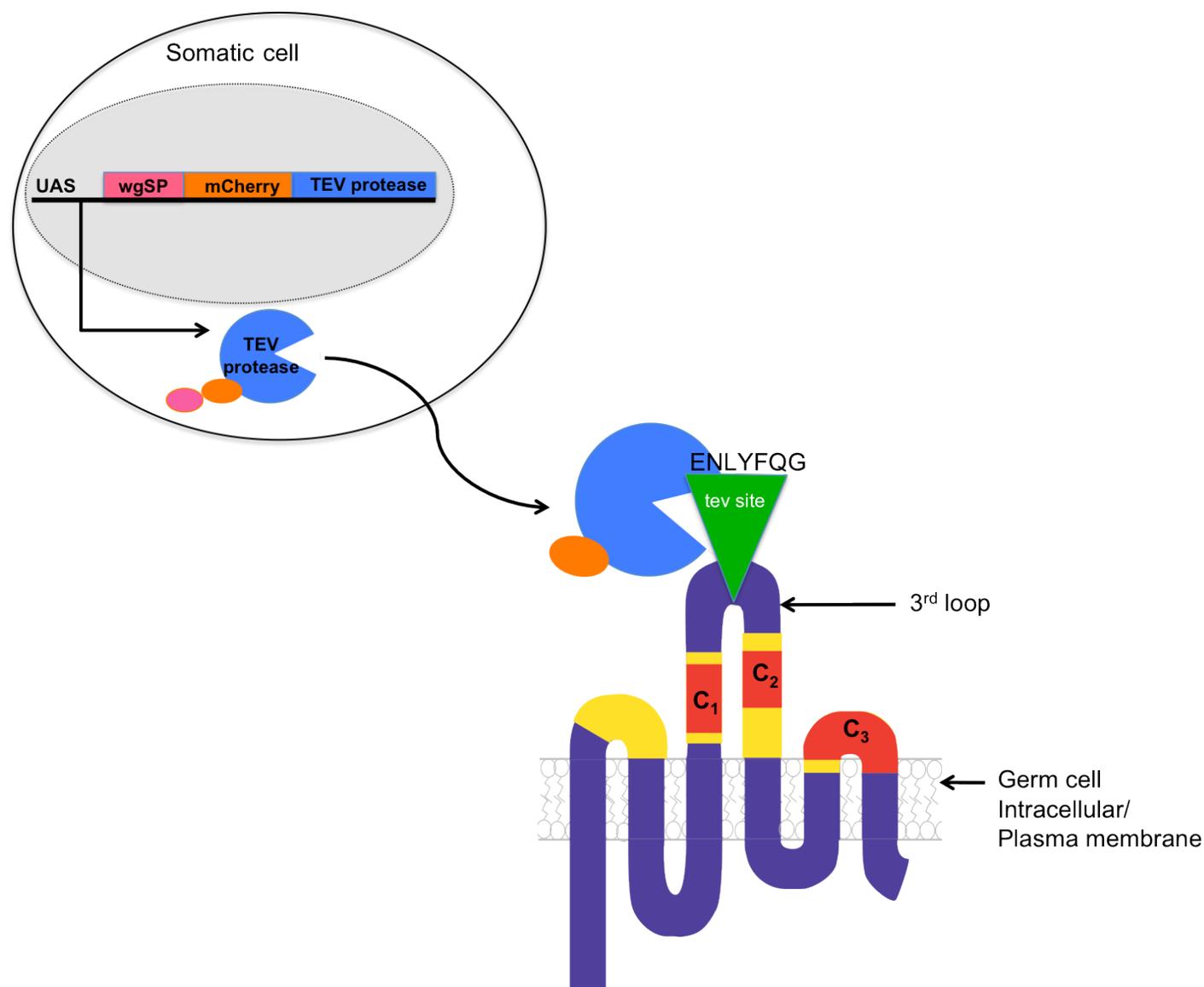


Figure 2.3: Schematic representation of TEV protease cleavage of Wun2tev. The inducible UAS promoter leads to the expression of the TEV protease (blue), which is tagged with mCherry (orange) and contains a wingless-signal peptide (wgSP, pink). The wgSP enables the secretion of the TEV protease outside of the cell, where it can cleave at the 7-amino acid (ENLYFQG) recognition site (green) in the 3rd loop of Wun2 between its two catalytic domains, C1 and C2 (red). This should render the enzyme inactive and block its ability to rescue the death of *wunen* null germ cells.

2.1.4 The region of tev-site insertion is relatively unstructured

The Myc-tagged form of Wun2, active and displaying both intracellular and plasma membrane localization, provided a good system to test the intracellular

two catalytic domains in this loop are highlighted in red and the region between the first two catalytic domains is highlighted in purple. The arrow indicates the point of insertion of the *tev*-site. It can be seen that the amino acids around these regions are predicted with a high confidence value to be involved in a coil structure.

2.1.5 Wun2 with a *tev*-cleavage site insertion is functional

Following these bioinformatic predictions, the *tev*-site was inserted into *wun2myc*, forming *wun2-tev-myc*, using overlap extension PCR. In parallel, a control without the *tev* site (*wun2-con-myc*), was synthesized in the same manner as *wun2-tev-myc*. The two constructs were then cloned into the pTW Gateway vector containing a *UAS* promoter upstream of the insertion site. Sequencing the Gateway clones revealed that the integrity of the *wun2* gene was maintained through the cloning process and the *tev*-site was successfully inserted in the desired location (Fig. 2.5).

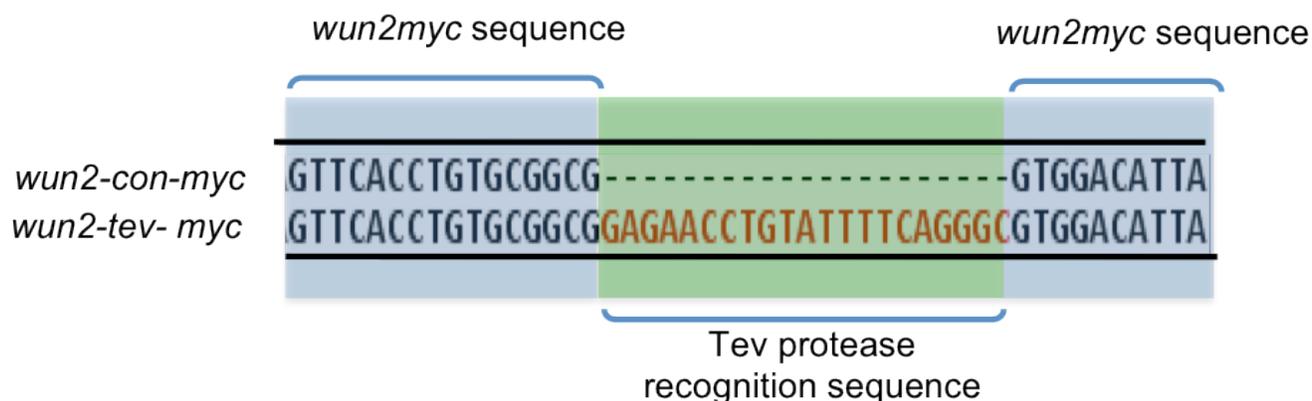


Figure 2.5: Sequencing data from the wild type (upper line) and *tev* site containing (lower line) Wun2 constructs. The blue highlighted sequences are those of *wun2* whereas the green highlighted sequences indicate the inserted *tev* site. The sequencing also showed that the reading frame was preserved.

To test if the *tev*-site containing Wun2 retained its functionality *in vivo*, *nosGal4* was used to express either the *UASwun2-con-myc* or the *UASwun2-tev-myc* transgene in *wun wun2* maternal mutants. Both constructs were equally capable of rescuing the germ cell death in these embryos indicating that Wun2tevmyc was functional and therefore retained catalytic activity (Fig. 2.6).

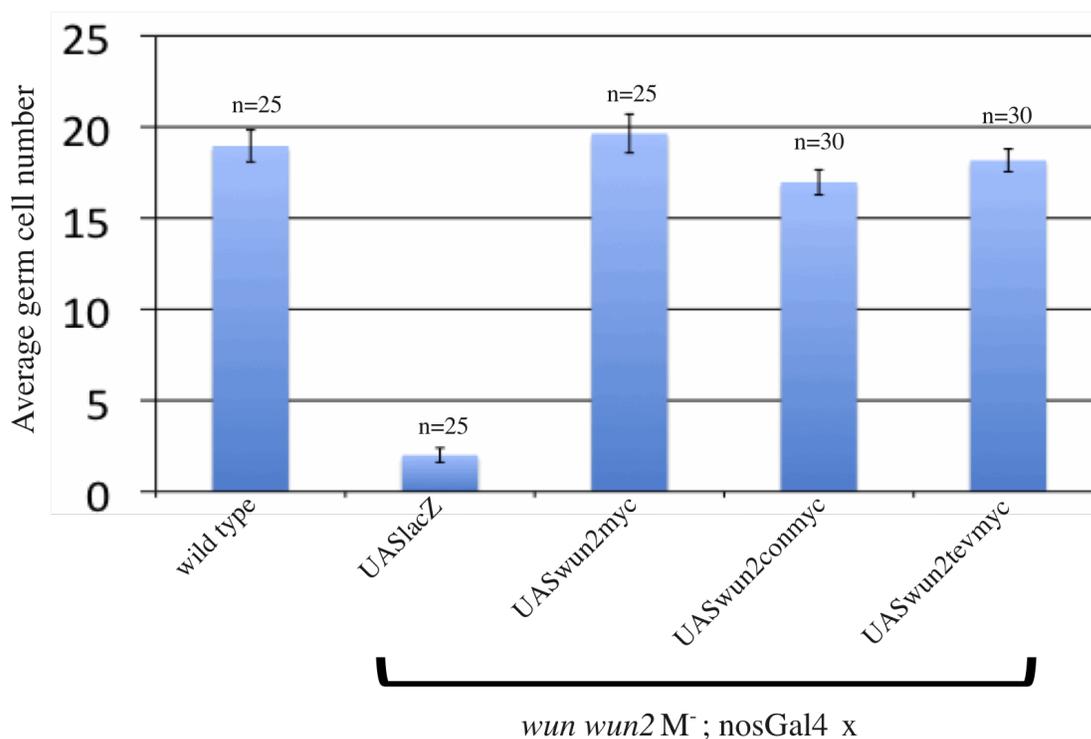


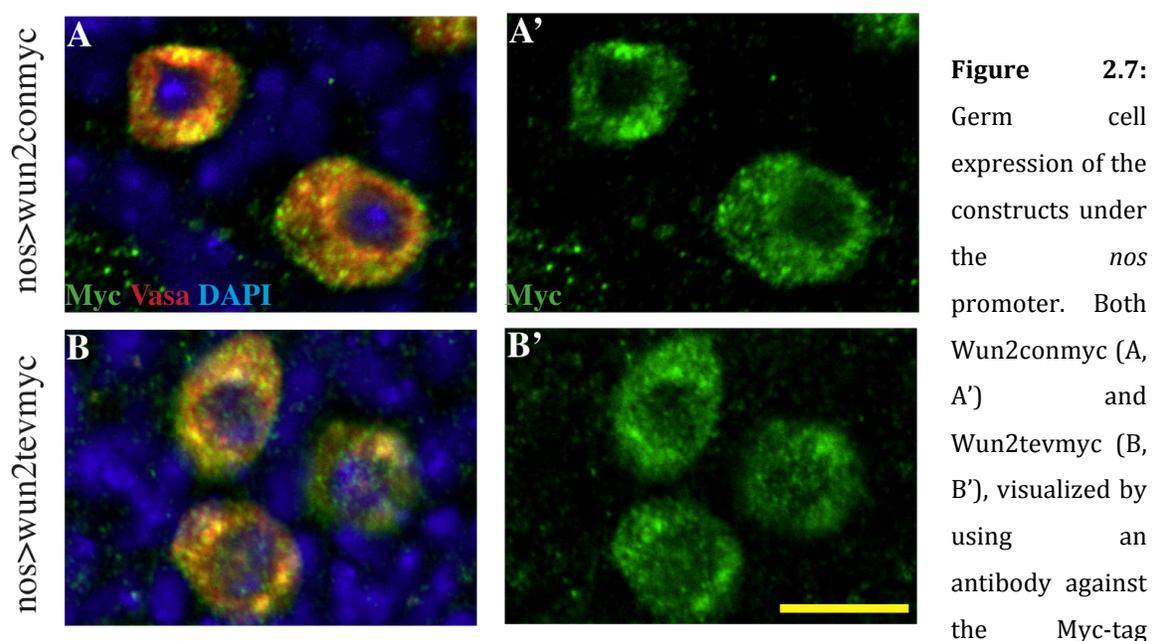
Figure 2.6: Graph depicting the average germ cell number at stage 13 in embryos from *wun wun2 M* mutants expressing various UAS rescue constructs. Driving *UASwun2-tev-myc* in the germ cells using the *nosGal4* driver rescues germ cell death due to lack of *wunens*, compared to those expressing a control construct, *UASlacZ*. Moreover, *Wun2* with a *tev*-site rescues to the same extent as the control *Wun2* without the *tev*-site. Error bars show the standard error, and ‘n’ represents number of embryos scored.

2.1.6 Using the *nanos* promoter to ensure expression of constructs at physiological levels

Although the *nosGal4* driven expression is also restricted only to the germ cells, the UAS-Gal4 is an overexpression system, which would lead to the production of overwhelming amounts of the protein. Therefore, to ensure the germ cell specific expression of the *wun2tevmyc* at physiological levels, the *wun2* constructs were transferred to a *nos* expression vector. This vector contains a *nos* promoter, a gateway cloning site and *nos* 3'UTR, the latter being sufficient to localize transcripts to the posterior of the embryo and hence be incorporated into germ cells. This shift

to the *nos* expression cassette also allows for the usage of the UAS-Gal4 system to later express the TEV protease.

Transgenic flies expressing *wun2(con/tev)myc* under the *nos* promoter (*nos>wun2(con/tev)myc*) were checked for detectable transgene expression. On using an antibody against the Myc-tag, both Wun2conmyc and Wun2tevmyc could be detected in the migrating germ cells in stage 10-11 embryos. The expression of the transgene was restricted to the germ cells (Fig 2.7, A, A' B, B'). Moreover, the subcellular localization of these Myc tagged proteins were similar to the pattern using a *nosGal4* driver (Fig 2.2 C, C'): largely cytoplasmic and some plasma membranous.



(green), co-localize primarily with cytoplasmic Vasa (red). Some plasma membranous presence can also be observed beyond the Vasa signal. Scale bar: 10 μ m.

2.1.7 Germ cells do not show high levels of endocytosis

The germ cells spend the majority of their migratory route in the mesoderm. Therefore to maximize the exposure time of the germ cell Wun2tevmyc to the protease, the latter was expressed (from a *UASwgmcherryTEV construct*) using the pan-mesodermal driver *twistGal4* (*twiGAL4*). Due to the wgSP at its N-terminus, the

TEV protease would then be secreted out of the mesodermal cells and into extracellular spaces, and could further be visualized due to its mCherry-tag.

However, it could be possible that the migrating germ cells internalize the secreted protease, hence making it difficult to differentiate between cleavage of Wun2tevmyc present at the germ cell surface and of that present within the cell. To address this, we imaged germ cells migrating through the mesoderm of embryos expressing secretable mCherry-tagged TEV protease. The germ cells were imaged for the presence of mCherry inside the germ cells. As a control, migrating germ cells were imaged in embryos expressing a non-secretable mCherry (*UASmcherrymoesin*) in the mesoderm.

As expected, no mCherry signal could be detected in the germ cells in the mCherryMoesin expressing embryos (Fig 2.8 A, A'). Moreover, similar to the control, although the germ cells were surrounded by mCherryTEV expressing cells, there was no detectable protease in the cytoplasm of the germ cells (Fig 2.8 B, B'). Hence it is unlikely that the germ cells internalize the secreted protease.

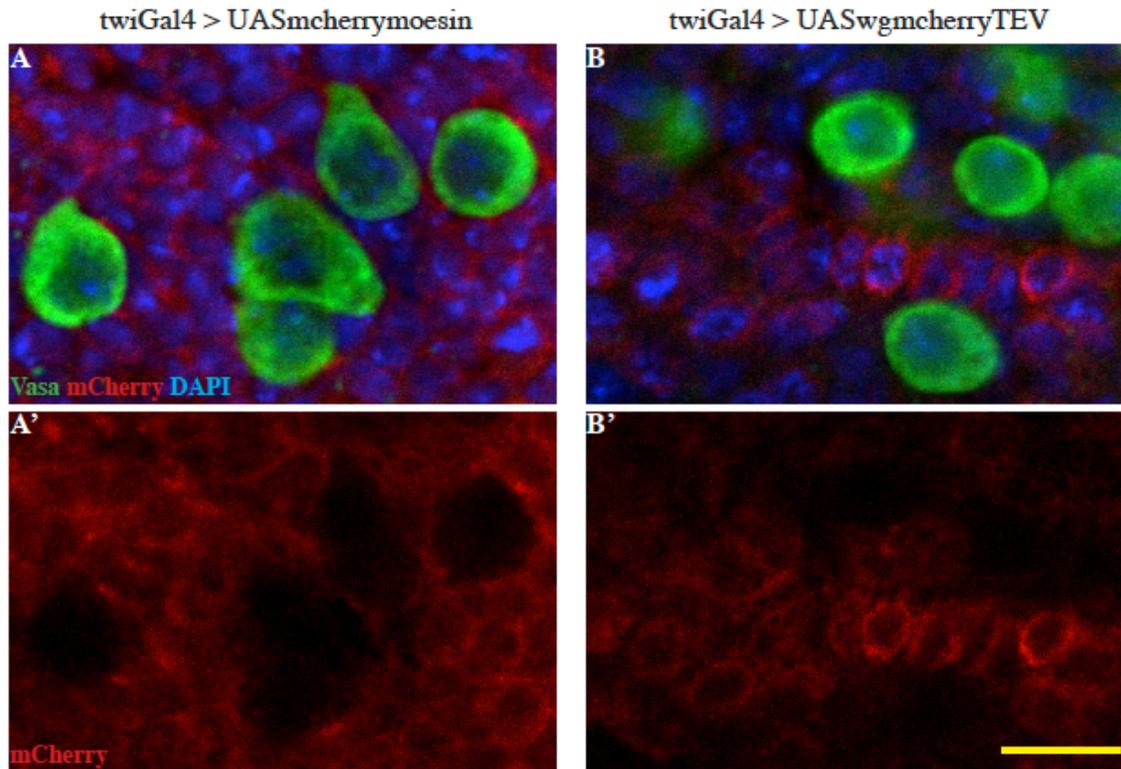


Figure 2.8: Germ cells do not uptake secreted TEV protease: Migrating germ cells (green) in *twiGal4>UASmcherrymoesin* (A, A'), or *twiGal4>UASwgmcherryTEV* embryos (B, B') have no detectable mCherry signal (red) in their cytoplasm. Scale bar: 10µm.

2.1.8 Mesodermal expression of secreted TEV protease reduces the ability of Wun2tev to rescue the death of *wunen* null germ cells

In order to test the site of Wun2 activity in germ cells, it was important to have the germ cells depend entirely on the cleavable form of Wun2. To achieve this, the *nos>wun2tevmyc* construct was introduced into mothers that lay embryos lacking endogenous *wunens*. Therefore the germ cells in such embryos would receive only Wun2tevmyc but the soma however, would express *wunen* zygotically. In such embryos I also expressed the secretable TEV protease specifically in the mesoderm (Fig 2.9).

I reasoned that if Wun2 was required on the surface of the germ cells, then the cleavable Wun2 would be cleaved and rendered inactive by the TEV protease

present in the extracellular spaces. Due to this, the germ cells would be unable to dephosphorylate the lipid survival factor, leading to germ cell death. On the other hand, if Wun2 is required intracellularly, the TEV protease would not have access to the cleavable Wun2 inside the cell, hence lipid dephosphorylation would still occur and germ cells would survive (Fig 2.9).

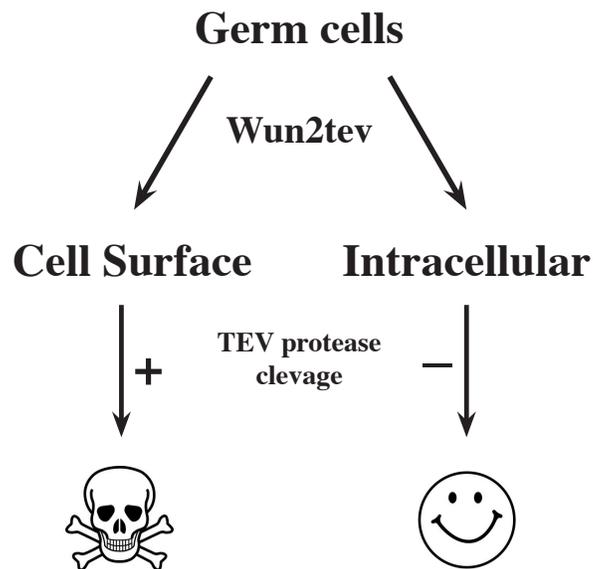
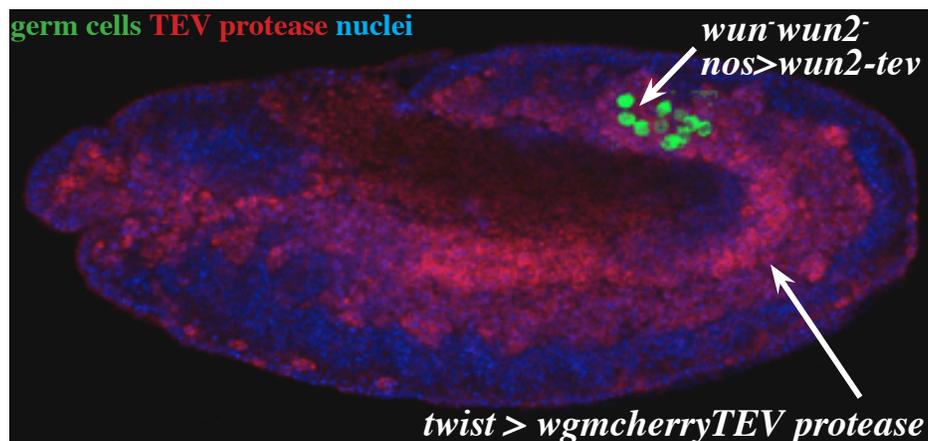


Figure 2.9: The *nos* promoter drives the expression of the cleavable form of *wun2* (*nos>wun2tev*) in germ cells (green) lacking endogenous *wunens* (*wun wun2 M*). The soma expresses the secretable TEV protease (red) in the mesoderm (*twiGal4>UASwgmCherryTEV*). The secreted protease can act on the cell surface Wun2tev (+) but the intracellular Wun2tev is protected from TEV protease cleavage (-) If Wun2 is required at the cell surface, then the inactivated Wun2tev at the cell surface would lead to germ cell death. But germ cells would survive if they require only intracellular Wun2 activity.

Germ cell numbers were scored in *wun wun2* M- embryos from *nos>wun2tevmyc* and *nos>wun2conmyc* containing mothers. Several independent insertions of each transgene were tested. Firstly, although the *nos>wun2tevmyc* and *nos>wun2conmyc* did rescue germ cell death, unlike on overexpression with the *nosGal4* (Fig 2.6), the rescue was partial and wild type number of germ cells were not observed. Moreover, the two *nos>wun2conmyc* lines, line#1 and line#2, showed germ cell numbers of 12 and 8, respectively, indicating that there are insertion line specific differences (Fig 2.10), most likely due to differing expression levels.

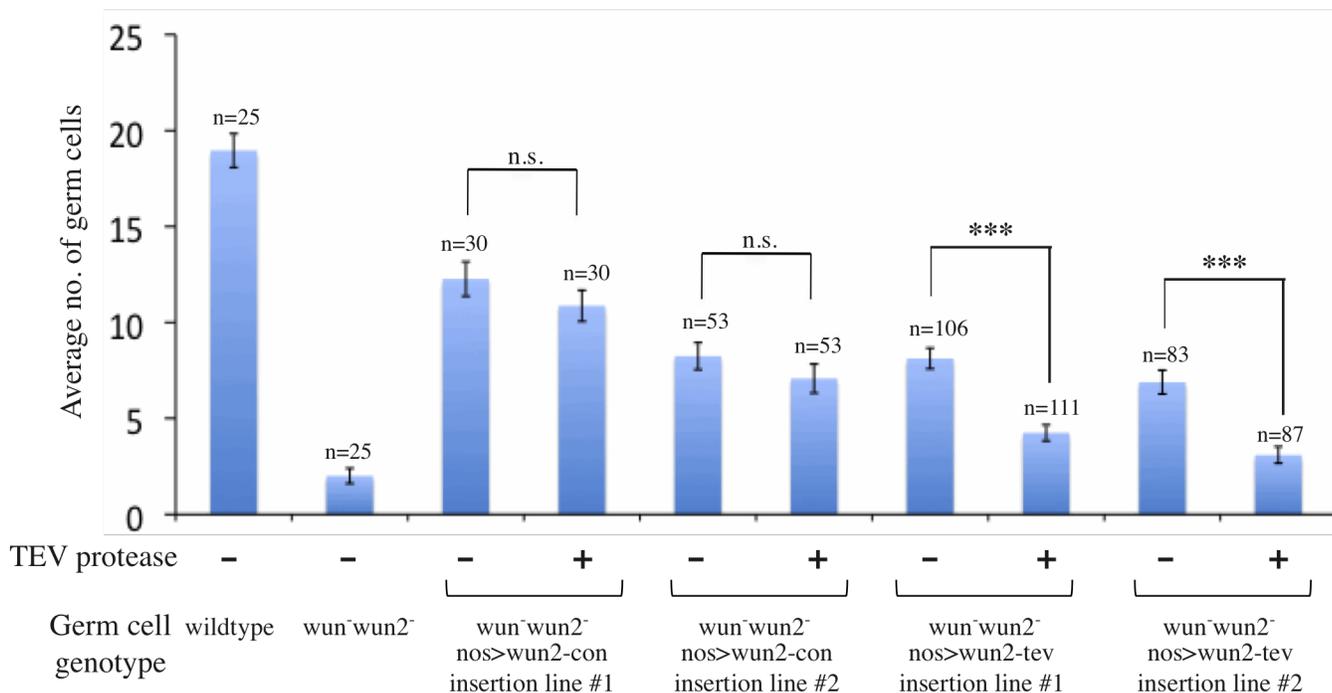


Figure 2.10: Graph depicting the average germ cell number at stage 13 in embryos from *wun wun2* M- mutants expressing various *nos* driven rescue constructs. On the X-axis, the expression of the TEV protease is denoted as ‘-’ for absent and ‘+’ for present. Further, the different germ cell genotype in each case is denoted, keeping in mind that in all cases somatic *wunen* expression is normal. The *nos* driven *wun-con* as well as *wun2-tev* insertion lines rescue germ cell death seen on complete lack of *wunen* in germ cells, but not back to wild type levels. However, on expressing the TEV protease, there is no significant (n.s.) difference in rescue of the two control lines when compared to absence of the protease (p-value of 0.25 for *wun2con* line#1 and p-value of 0.25 for *wun2con* line#2 by Student T-test). The *wun2-tev* insertion lines, in contrast, show a highly significant difference (p-value of 3.6 E-08 for *wun2tev* line#1 and p-value of 1.06 E-06 for *wun2tev* line#2) between the rescue levels with or without the protease. Error bars show the standard error, and ‘n’ represents number of embryos scored.

Similar to the control lines, the two *nos>wun2tevmyc* lines #1 and #2, showed germ cell numbers of 8 and 7, respectively (Fig 2.10). Mesodermal expression of the secreted TEV protease resulted in significantly reduced germ cell numbers (by Student T-test) to 4 and 3 in lines #1 and #2, respectively (Fig 2.10). The control lines on the other hand showed negligible differences in germ cell numbers upon expression of the protease (11 and 7 for lines #1 and #2 respectively).

Therefore, the absence or the presence of the TEV protease did not significantly affect the ability of the control Wun2 to rescue germ cell death. However, in case of the cleavable form of Wun2, the expression of the protease lead to a substantial decrease in germ cell number with an average reduction of 4 germ cells per embryo. This highlights the difference in the rescue ability of Wun2-tev but not Wun2-con upon expression of the TEV protease, and is consistent with a cell surface requirement of Wun2 catalytic activity.

2.2 Analysis of EMS mutants showing defective gonad morphogenesis

Gonad morphogenesis couples the end of germ cell migration with SGP specification, and the following interplay between the two cell types to form the gonad. The genetic dissection of this complex process can be performed using the powerful tool of forward genetic screens.

2.2.1 Gonad formation mutants showing normal early germ cell migration

In a mutagenesis screen using ethyl methanesulphonate (EMS) (Leal et al., 2009), three mutants were identified, *A44*, *B23* and *C28*, all exhibiting a defect in gonad formation. The screen was designed to uncover both maternal and zygotic effect mutations. The *A44* mutant displays a maternal effect phenotype, i.e., the phenotype is seen in the offspring only when the mothers are homozygous for the mutation. However, since the *A44* line is homozygous lethal, germ line clone (GLC) females have to be used in order to see the phenotype. The other two, *B23* and *C28*, are zygotic mutants.

Interestingly, all three mutants display normal germ cell migration: at stage 11 the germ cells move into the mesoderm (Fig. 2.11, A1, B1, C1) and split into two bilateral clusters (data not shown), at stage 12 they continue their migration in the mesoderm (Fig. 2.11, A2, B2, C2). However, in the following stages, the germ cells display irregular behavior; at stage 13, the germ cells don't align in a row, but instead appear scattered (Fig. 2.11, A3, B3, C3). This irregularity is amplified in later stages, where by stage 15 the germ cells don't always come together to form a gonad, but are instead scattered at the at the posterior of the embryo, clumped together in several clusters, or form an uncompact gonad (Fig. 2.11, A4, B4, C4). In spite of their defects in gonad formation, all three mutants show no apparent gross morphological/patterning defects as judged by gut patterning. However, they are all

homozygous lethal, and survive only until late embryonic to early larval stages, with only a small percentage of A44 mutant larvae developing into pupae.

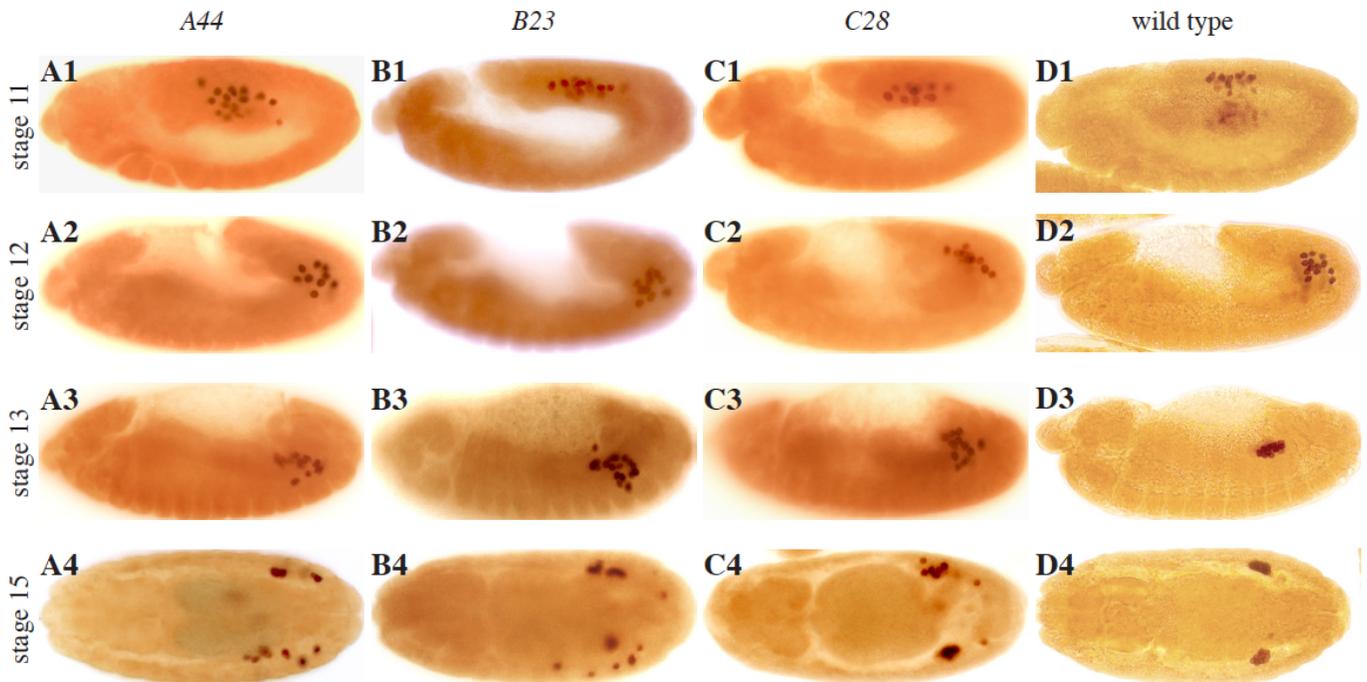


Figure 2.11: Vasa-labeled germ cells in three mutants displaying gonad formation defects. Early germ cell migration at stage 11 (A1-D1), and stage 12 (A2-D2) in maternal *A44* (A1-A4) mutant, *B23* (B1-B4) and *C28* (C1-C4) zygotic mutants, appear normal and similar to wild type (D1-D4). At stage 13 (A3-D3), the germ cells appear more scattered, and stage 15 (A4-D4) display clear defects in gonad formation in the mutants.

2.2.2 Mutants display normal SGP specification

Since staining of the germ cells suggests that gonad formation is defective in the three isolated mutants, it was necessary to investigate if the SGPs were specified. An early marker of SGP specification is the *412* retrotransposon. It is expressed in broad mesodermal stripes early in embryogenesis, but by stage 11 it is maintained mostly in dorso-lateral clusters in the mesoderm (Brookman et al., 1992a). The embryos of the three mutants were subjected to fluorescent *in situ* hybridization using a *412* probe, and the Vasa antibody was used at the same time to visualize the germ cells. In wild type stage 12 embryos, three clusters of SGPs could be observed

(Fig. 2.12, A, white arrow heads) making contact with the germ cells. Similar to wild type, in the mutants, 3 distinct clusters of SGPs were observed (Fig. 2.12, B, C, D, white arrow heads). Moreover, the SGP clusters in the mutants appeared to establish contact and interact with the migrating germ cells. This implied that the early steps of gonad formation, i.e., SGP specification and contact formation, were normal in the *A44*, *B23*, and *C28* mutants. Additionally, normal germ cell migration and SGP specification suggested that the mesoderm in these mutants has undergone proper patterning.

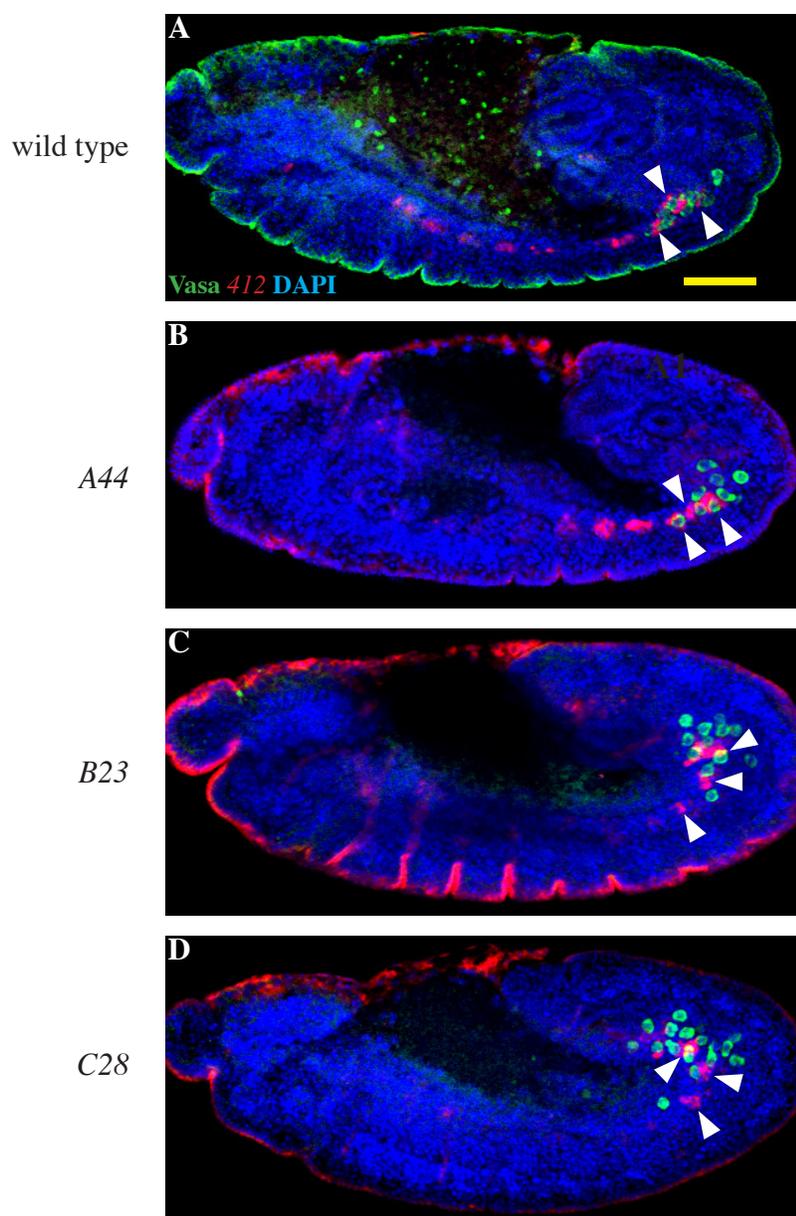


Figure 2.12: SGP specification in stage 12 mutant embryos. SGPs were labeled using the 412 probe (red), germ cells were labeled using the Vasa antibody (green), and the nuclei were stained with DAPI (blue). Similar to wild type (A), where three clusters of SGPs are observed (white arrow heads), the mutants *A44* (B), *B23* (C) and *C28* (D) also showed 3 specified clusters of SGPs (white arrow heads) making contact with the germ cells. During this stage, the 412 labeling can also be seen in other mesodermal clusters (A, B) besides PS10-13. Scale bar: 50 μ m

2.2.3 At later stages mutant gonads show SGP-related defects

In the mutants, the early stage SGPs showed competence in initiating contact formation with the germ cells. However, since the germ cells at stage 13 appeared dispersed instead of being aligned in a line (Fig 2.11, A3, B3, C3), this suggested a defect in SGP-germ cell interaction at this stage. Therefore, the SGPs in the three mutants were examined at the later stages of gonad formation. In wild type stage 13 embryos, the three clusters of SGPs fuse into one contiguous tissue (Fig 2.13, A1). Simultaneously, the cytoplasmic protrusions emanating from the SGPs ensheath the germ cells and individualize them (Fig 2.13, A1). In the mutants, *A44*, *B23*, and *C28*, however, both of the steps of SGP-cluster fusion, and germ cell ensheathment by the SGPs at stage 13 were impaired. In most occasions, in the three mutants, minimally one cluster (typically, but not necessarily, the anterior cluster) was disjoined from the other two clusters (Fig 2.13, B1, C1, D1 white arrow heads). Moreover, many germ cells at stage 13 were scattered and found to be not associated with the SGPs (Fig 2.13, B1, C1, D1 yellow arrow heads).

At stage 15, while the wild type gonad coalesced into a tight and round gonad (Fig 2.13, A2), the mutant gonads appeared abnormal. The earlier lack of SGP cluster fusion or ensheathment was not overcome with time (Fig 2.13, B2, C2, D2), indicating that these processes were defective in the mutants, and not caused simply by a delay in development. Moreover, SGP clusters that were occupied with germ cells did not appear compacted. Therefore these mutants appear to display defects in many of the SGP-driven processes required for gonad formation.

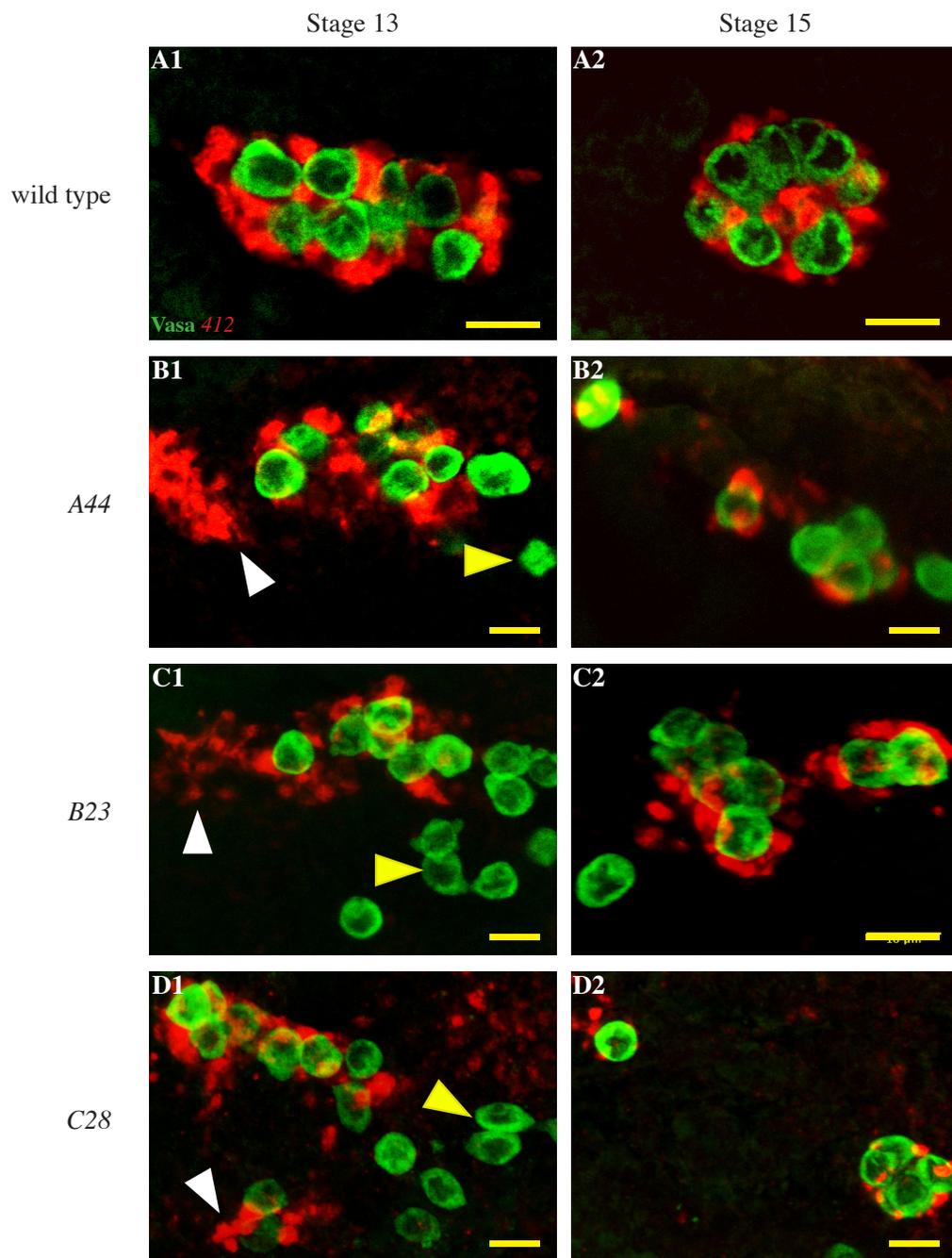


Figure 2.13: Mutant SGPs in later stages of gonad formation. SGPs were labeled using the *412* probe (red), germ cells were labeled using the Vasa antibody (green). Wild type stage 13 gonad (A1) display fused clusters of SGPs, ensheathing the germ cells. In the mutants *A44* (B1), *B23* (C1) and *C28* (D1) stage 13 embryos however, at least one separate cluster of SGP can be seen (white arrow head) along with further scattered germ cells (yellow arrow head). Moreover, while the wild type stage 15 gonads (A2) appear round and compact, the mutant gonads at this stage (B2, C2, D2) fail to compact with scattered SGP-germ cell clumps. Scale bar: 10 μ m

2.2.4 Mesoderm derived tissues are fusion-competent in the mutants

Although SGP specification indicates that mesodermal patterning is not grossly disrupted, the observed SGP cluster fusion defect in the mutants could potentially arise due to lack of fusion-competence in mesoderm-derived tissues in general. To test this, a Fasciclin III (FasIII) antibody was used to visualize the visceral mesoderm, which, like the SGPs, is also specified in clusters in the early mesoderm. These clusters must fuse over the course of development, to form a single contiguous tissue, to eventually surround and separate the gut from the rest of the embryo (Azpiazu and Frasch, 1993). In wild type embryos, at stage 13 the FasIII-labeled visceral mesoderm appeared as a single continuous tissue (Fig 2.14, A1 white arrow), and by stage 15 it could be observed as a thin sheet enveloping the developing gut (Fig 2.14, A2 white arrow). In *A44*, *B23*, and *C28* mutants, the visceral mesoderm at stage 13 (the stage when SGP cluster fusion would normally occur) formed a single continuous belt of tissue and was indistinguishable from wild type (Fig 2.14, B1, C1, D1). The visceral mesoderm around the gut continued to be intact in the stage 15 mutant embryos (Fig 2.14, B2, C2, D2). These observations support the notion that the mesoderm is correctly patterned in the mutants, but also show that other mesoderm-derived tissues are still fusion competent.

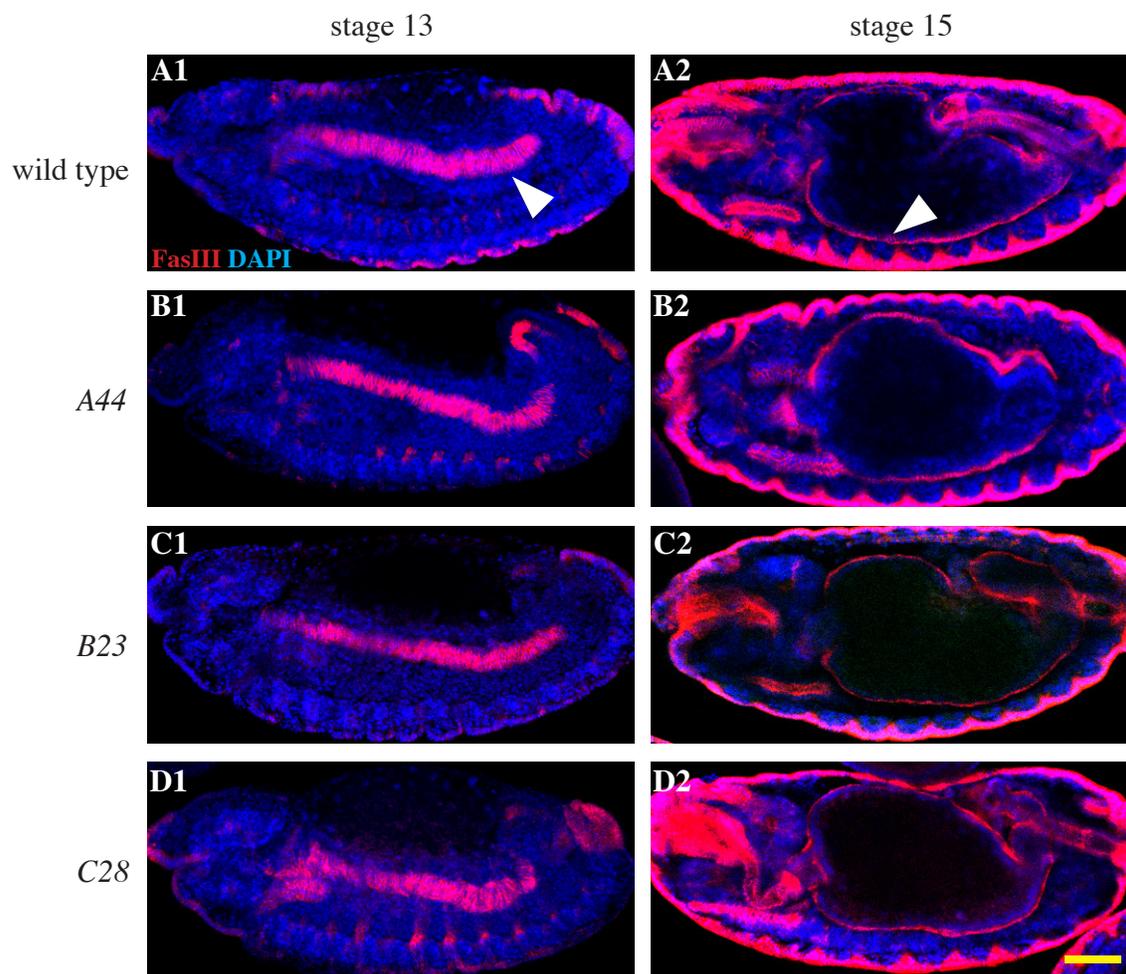


Figure 2.14: Visceral mesoderm in wild type and mutant embryos. Wild type stage 13 embryos (A1) display a coherent and continuous FasIII-labeled (red) visceral mesoderm (white arrow head), which at later stages (A2) form a thin layer around the gut (white arrow head). The visceral mesoderm in *A44* (B1), *B23* (C1) and *C28* (D1) stage 13 embryos also appears smooth and contiguous and similarly surrounds the gut in stage 15 embryos (B2, C2, D2), as observed in wild type. Scale bar: 50 μ m

2.2.5 Complementation screen approach to determine the locus responsible for the *A44*, *B23*, and *C28* mutant phenotypes

Identifying the causative genes in the three mutants could potentially reveal a new understanding of the aspects of SGP behavior required to undertake the various steps involved in normal gonad morphogenesis. In order to do this, a complementation screen was performed using the Bloomington deficiency kit. Since

the EMS mutagenesis screen, from which these mutants were isolated, was designed in a manner to isolate mutantions only on the second chromosome, the deficiency kit for this chromosome was used for the lethality screen. The *A44* mutant phenotype was observed with 100% penetrance with GLC females having the Flip Recombinase target (FRT) site at 42D, making only the right arm of the 2nd chromosome (2R) homozygous in these females. Therefore, the *A44* causative mutation must lie on chromosome 2R, whereas, the *B23* and *C28* mutation could be either on 2L or 2R chromosome. Assuming that the mutation causing the phenotype is lethal, each deficiency line (deleting a small part of the second chromosome), was crossed to each mutant line (homozygous lethal), and the offspring were scored for lethality.

Two deficiency lines from chromosome 2R, *Df(2R)BSC595* and *Df(2R)BSC136* uncovering the cytological region 47A3-47F1 and 59F5-60B6 respectively, were found to be lethal with the *C28* mutant. The lethality screen also revealed two lethal deficiencies with the *B23* mutant on the left arm of the 2nd chromosome (2L), i.e., *Df(2L)ED4651* and *Df(2L)Exel6012* uncovering 22E1-22F3 and 25D5-25E6, respectively.

Unfortunately, none of the deficiencies proved to be lethal with the *A44* line, hence preventing narrowing down of the causative locus in this mutant. I conclude that either the causative mutation on chromosome arm 2R is not lethal but there is a second mutation on chromosome arm 2L which results in lethality or that the causative mutation is lethal but is in a location not covered by the deficiencies screened.

2.2.6 *Df(2R)BSC595* uncovers the *C28* causative locus

The locus causing lethality need not be coupled with the locus leading to the mutant gonad phenotype. Therefore, each of the lethal deficiencies was tested for non-complementation of the gonad formation defects, in a trans-heterozygous situation with the respective mutant lines.

To quantify the phenotype and characterize its pleiotropy, the various observed phenotypes were first categorized. The pleiotropy seen in these mutants are common to mutants affecting gonad formation (Weyers et al., 2011). The categorization was performed on Vasa-labeled germ cells in stage 15 and older embryos. Although, I have demonstrated that defects in these mutants are essentially SGP-related, the germ cells act as a good read out of the condition of the gonad, and allows for a faster way of screening. The first category is the 'wild type' category where the gonads appear normal, defined as two round gonads with 6 or fewer germs scattered outside, as a few germs per embryo can be found lost in many genetic backgrounds. The second is the ' ≥ 7 germ cells (gc) outside gonad', where the observed phenotype is some scattered germ cells, with visible intact gonads on each bilateral side. This high threshold enables quantification of truly affected gonads. Next is the ' ≥ 2 germ cells (gc) clusters/gonad', where the germ cells are not necessarily scattered, but instead are present as 2 or more neighboring clumps, possibly representing the unfused SGP-germ cell clusters. The fifth category 'no real gonad', the most severe category, is scored when there occur many separate clusters/clumps containing only a few germ cells, indicating a complete gonad loss. The last category was that of 'uncompacted or elongated gonad' where the gonads are visible, but remain elongated at late stages, indicating a failure of gonadal coalescence.

Compared to the heterozygous *C28* phenotype containing approximately 90% wild type gonads, *C28* homozygous embryos displayed 90% mutant gonads covering the whole array of phenotypes (Fig. 2.15, A, E). Nearly 50% of the homozygous embryos showed the most severe class of phenotype, 'no real gonad'. Two deficiency lines that are lethal with *C28* (*Df(2R)BSC136* and *Df(2R)BSC610*) were crossed with *C28* and assessed for the gonad phenotype. We found that *C28* complemented the mutant phenotype when trans-heterozygous (data not shown). Therefore the lethal mutation uncovered by these deficiencies is not the cause of the gonad phenotype.

The region uncovered by the second deficiency line *Df(2R)BSC595* was also divided into smaller fractions with two other deficiency lines, *Df(2R)BSC336* and *Df(2R)ED2098*. Of the two, only *Df(2R)ED2098* proved to be lethal with *C28*. On testing for a mutant gonad phenotype, both deficiencies *Df(2R)BSC595* (Fig. 2.15, B, E) and *Df(2R)ED2098* (Fig. 2.15, C, E), showed the observed Vasa-labeled gonadal defects with penetrance of 80% and 70%, respectively, similar to the homozygous *C28* mutant embryos (Fig. 2.15, A). However, the highest represented category had shifted from the severe, 'no real gonad' to the less severe '≥7 gc outside gonad' category. As expected, the smaller non-lethal deficiency *Df(2R)BSC336* (Fig. 2.15, D), showed normal embryonic gonads in trans with *C28*.

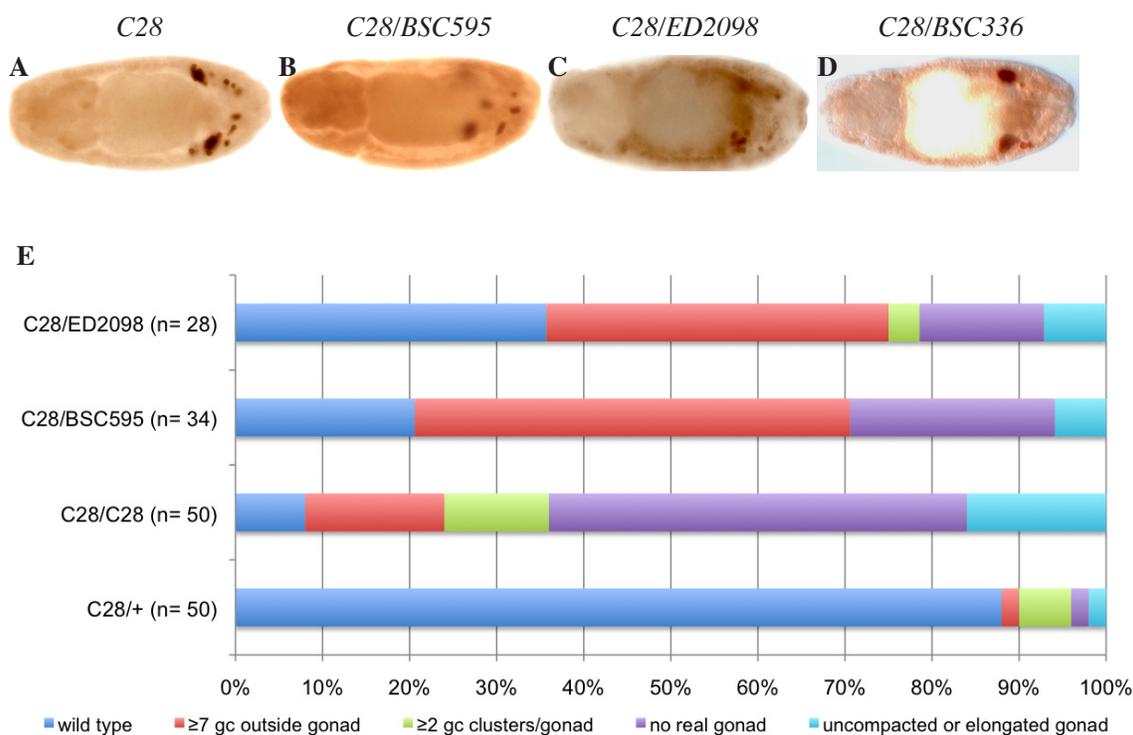


Figure 2.15: Germ cell migration phenotypes and their quantification in the *C28* mutant and in trans to several deficiency lines. Complementation tests of deficiency lines *BSC595* (B, E) and *ED2098* (C, E), with *C28* display similar Vasa-labeled germ cell phenotypes in stage 15 embryos as *C28* homozygotes (A, E). Moreover, the *BSC336* non-lethal deficiency complements the gonad phenotype of *C28* (D). The phenotype is further quantified (E) using the indicated categories (for explanation of the categorization, please see text, section 2.2.6). The *C28* mutant (the first indicated genotype is

always derived from the female) crossed to the deficiencies show a high level of penetrance of the germ cell migration phenotype.

Therefore the causative mutation in *C28* is lethal and lies within the overlap of the regions deleted in *Df(2R)BSC595* and *Df(2R)ED2098*, namely 47A7-47C6.

2.2.7 *Df(2L)Exel6012* uncovers the *B23* causative locus

A similar phenotypic complementation test was performed with the deficiencies found to be lethal with the *B23* mutant and the phenotypes were further quantified (Fig. 2.16, D). When in single copy the *B23* mutation already caused some scattering of germ cells in a small percentage of embryos. However, on being homozygous, almost 100% of embryos displayed a mutant phenotype with approximately 90% of the mutant embryos having no discernable gonads (Fig. 2.16, A, D). While the *Df(2L)ED4651* line was found to complement the *B23* gonad defects (data not shown), the *Df(2L)Exel6012* line was unable to phenotypically complement *B23* when in trans (Fig. 2.16, B, D). Similar to the *B23* homozygotes, the majority of *B23/Df(2L)Exel6012* embryos displayed the most severe category of no discernable gonads. This implicates the cytological region 25D5-25E6, as defined by the breakpoints of the deficiency *Df(2L)Exel6012*, as the causative locus for the *B23* mutant gonad phenotype.

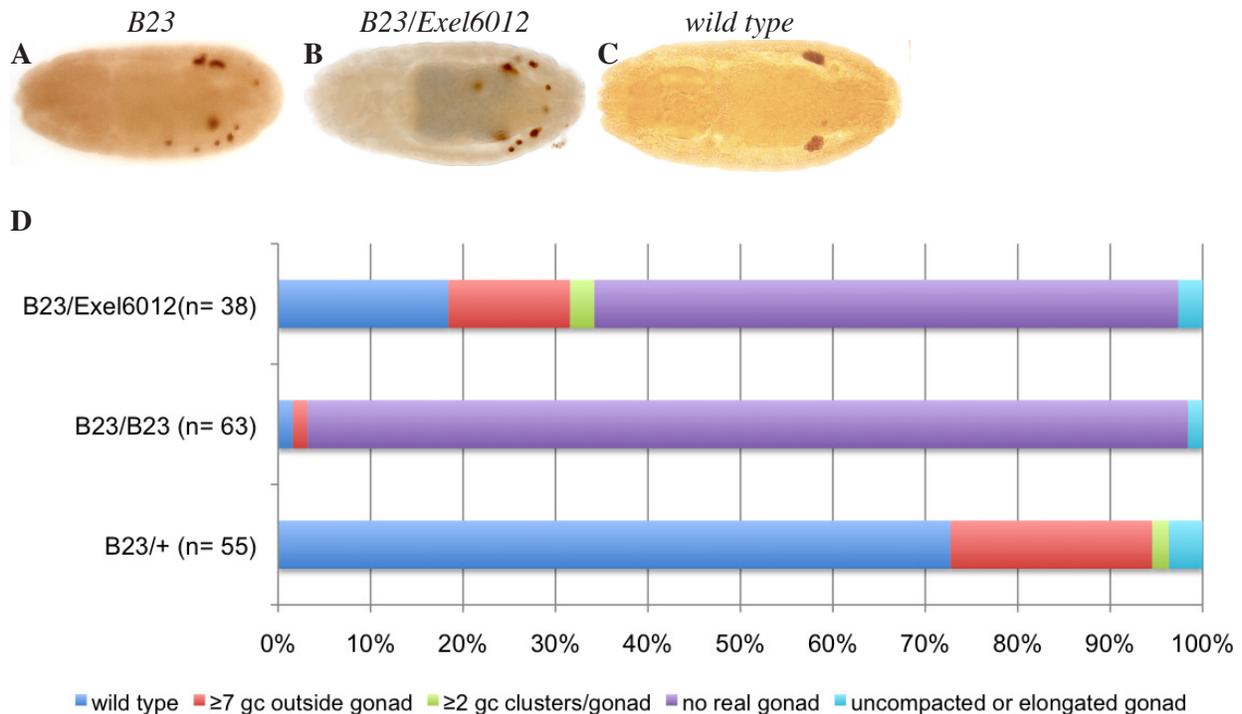


Figure 2.16: Germ cell migration phenotypes and their quantification in the *B23* mutant and in trans to deficiency line. Complementation test of the deficiency line *Exel6012* (B, D) with *B23*, displays similar Vasa-labeled germ cell phenotypes in stage 15 embryos as *B23* homozygotes (A, D). The phenotype is further quantified (D) using the indicated categories (for explanation of the categorization, please see text, section 2.2.6). The *B23* mutant (the first indicated genotype is always derived from the female) crossed to the deficiencies show a high level of penetrance of the germ cell migration phenotype.

2.2.8 Whole genome sequencing of *A44*, *B23* and *C28* mutants to identify the causative mutations

In a parallel approach, I sought to identify the causative mutation using whole genome sequencing. *A44*, *B23* and *C28* mutant lines were isogenized and homozygous embryos were collected to isolate a minimum of 3µg of genomic DNA from each line. The genomic DNA was sheared into 500bp fragments and paired-end libraries were prepared. These libraries were then subjected to 100bp paired end-Illumina sequencing. The obtained reads were then analyzed using the CLC Genomics Workbench software, versions 4.7 - 5.5. Using the Genomic Gateway

plugin tool (later integrated into the software as the 'NGScore tools'), the reads from the three mutants were mapped onto the chromosome 2L complete sequence (GenBank NT_033779 Release 5.30) as well as chromosome 2R complete sequence (GenBank NT_033778 Release 5.30).

The reads from mutants *A44* and *C28*, each covered 99% of chromosome 2R (since the causative gene in the two mutants lies on this arm) with a coverage of 16x and 13x respectively (Table 2.1). The reads from the *B23* line on the other hand covered 100% of the chromosome 2L (where the causative gene for this line lies) with a coverage of 19x (Table 2.1).

Table 2.1: Table showing the coverage statistics for each mutant on the corresponding chromosomal arm.

	A44 (2R)	C28 (2R)	B23 (2L)
Total reference length	21,146,708	21,146,708	23,011,544
Total read count	3,657,497	2,901,079	4,721,864
Fraction of reference covered by reads	0.99	0.99	1.00
Average coverage	16X	13X	19X

SNPs (Single Nucleotide Polymorphisms) between the mapped mutant reads and the reference genomic sequence were identified using CLC Genomics Workbench. Table 2.2 denotes the number of SNPs obtained in the case of each mutant for their respective chromosomal arm. Since the three mutants were derived from EMS mutagenesis of flies of the same isogenized genetic background, the SNPs that were common between all three mutants were identified as tolerated naturally occurring polymorphisms that were likely present in the premutagenesis chromosome and removed. This drastically reduced the number of candidate SNPs remaining for each case (Table 2.2). Subsequently, non-synonymous SNPs, i.e. SNPs that lead to a change in amino acid in the protein coding regions, and SNPs at splice sites were determined. The mutations at these sites were verified by manual inspection of the sequencing reads to examine the quality of the reads. In addition, a selected number of SNPs were verified by PCR and Sanger sequencing and in 11 of 11 cases the SNPs were confirmed.

Table 2.2: Table showing the number of SNPs detected after each step in the three mutants for the corresponding chromosomal arm.

	A44 (2R)	C28 (2R)	B23 (2L)
SNPs compared to the reference genome	114,066	106,639	124,425
Unique SNPs when compared to the other two mutants	2,988	2,517	3,401
SNPs leading to amino acid changes	118	89	81
SNPs leading to splice site changes	1	0	2
SNPs after manual curation	22	29	17

This led to a final refined list of SNPs that could possibly account for the mutant gonad phenotype for the *A44* mutant (22 SNPs on chromosome 2R, Table 2.3), *B23* mutant (17 SNPs on chromosome 2L, Table 2.4) and the *C28* mutant (29 SNPs on chromosome 2R, Table 2.5).

Table 2.3: Table showing the list of SNPs on chromosome 2R in the *A44* mutant, after manual curation. The chromosomal position (Release 5.30) and the reference residue at that position is indicated along with the SNP in the *A44* mutant, with the % frequency that the SNP was seen in the no. of reads (coverage). The amino acid change caused by each SNP and the corresponding gene affected in also displayed.

Position 2R	Reference Residue	A44 Residue	Frequency	Coverage	Amino Acid Change	Gene
1642151	G	A	100	5	Glu11Lys	CG12792
2566116	G	A	100	8	Ser194Phe	Tdc1
3721670	A	T	100	13	Asn95Tyr	CG12821
3992926	G	A	88,9	9	Thr121Met	Cul-4
6001140	G	A	100	5	Cys246Tyr	Pfk
6170036	G	A	100	4	Arg1064His	CAP
8224687	G	A	100	8	Glu66Lys	CG34021
8293781	G	A	100	7	Val72Ile	Cpr49Af
8383622	G	A	100	5	Gly94Ser	s-cup
10089183	G	A	100	9	Gln764*	Sox15
10257153	G	A	100	10	Asp93Asn	CG34444
10470656	G	A	100	3	Ala479Thr	ttv
12829777	G	A	100	6	Gly527Glu	Pkc53E
12888027	A	T	100	10	Leu397*	ste24a
13313770	G	A	100	3	Ala366Val	CG6550
13594526	G	A	100	5	Ala340Val	CG6424
16888682	G	A	100	3	Gly69Arg	CG10543
17228514	G	A	100	7	Gly3160Asp	dom
18114188	G	A	100	10	Pro87Ser	Oatp58DC
19893140	G	A	100	7	Pro1567Ser	Unc-89
19949758	G	A	85,7	7	Pro24Ser	RpL39

Table 2.4: Table showing the list of SNPs on chromosome 2R in the *B23* mutant, after manual curation. The chromosomal position (Release 5.30) and the reference residue at that position is indicated along with the SNP in the *B23* mutant, with the % frequency that the SNP was seen in the no. of reads (coverage). The amino acid change caused by each SNP and the corresponding gene affected in also displayed.

Position 2L	Reference Residue	<i>B23</i> Residue	Frequency	Coverage	Amino Acid Change	Gene
421966	A	T	100	9	Asn56Ile	CG13692
3060268	G	A	100	12	Asp198Asn	CG3523
3336647	G	A	83,3	6	Arg367Trp	E23
5131102	G	A	100	23	Glu3513Lys	Msp-300
7613525	G	A	100	9	Asp41Asn	CG6739
7682125	G	A	93,8	16	Glu476Lys	Myo28B1
9769652	A	T	100	9	Glu190Val	CG17633
11108106	G	A	100	10	Ala319Thr	l(2)gd1
11525165	G	A	100	7	Pro168Ser	CG31705
13909954	G	A	100	7	Gly87Arg	Ance-2
13981373	G	A	100	5	Gly328Ser	nimC2
16452421	G	A	100	6	Gly300Ser	ldgf3
17972401	G	A	100	4	Ala160Thr	CG5674
19512611	G	A	100	3	Ser821Leu	CG10186
20644865	G	A	100	10	Pro156Ser	phr6-4
5465893	G	A	100	9	Splice site disruption (AG acceptor of 3rd and final intron changed to AA)	midline
10461342	G	A	100	9	Splice site disruption (AG acceptor of penultimate intron changed to AA)	CG6206

Table 2.5: Table showing the list of SNPs on chromosome 2R in the *C28* mutant, after manual curation. The chromosomal position (Release 5.30) and the reference residue at that position is indicated along with the SNP in the *C28* mutant, with the % frequency that the SNP was seen in the no. of reads (coverage). The amino acid change caused by each SNP and the corresponding gene affected in also displayed.

Position 2R	Reference Residue	<i>C28</i> Residue	Frequency	Coverage	Amino Acid Change	Gene
1863038	C	T	80	5	Gly823Asp	mle
2865254	C	T	100	4	Glu286Lys	vimar
3242797	T	A	100	7	Arg281*	dscam
3382870	C	T	100	5	Gly278Asp	CG1620
3389110	C	T	87,5	8	Thr242Ile	didum
5191818	C	T	100	8	Ala399Thr	Myd88
5712796	C	T	100	8	Gly547Glu	CG1625
6370617	T	A	87,5	8	Asn835Ile	lola
6878347	C	T	100	5	Asp129Asn	luna
7120774	C	T	100	5	Gly146Arg	CG13218
7472790	C	T	75	4	Ala1751Thr	CG8839
8736071	C	T	100	8	Ser117Asn	CG30486
8809089	C	T	100	4	Ser451Phe	NAT1
8825540	C	T	100	11	Arg79Gln	CG30487
9842953	T	A	100	6	Val11Glu	CG33155
10147737	C	T	100	4	Ala296Val	β4GalNAcTA
12028713	C	T	100	6	Ser289Leu	Got1
12232345	C	T	100	3	Pro87Leu	Nup62
12544807	C	T	100	10	Glu81Lys	Picot
13554256	C	T	100	5	Pro208Leu	eIF3-S8
14425758	C	T	100	5	Ala204Val	Pepck
17012350	T	A	100	4	Gln685Leu	CG4050
17408683	C	T	100	7	Asp148Asn	CG30289
17662141	T	A	100	7	Tyr286*	CG4021
18071524	C	T	100	4	Ser270Leu	MED16
19740401	C	T	100	7	Pro375Leu	gbb
19854112	C	T	87,5	8	Pro521Ser	Alas
19944935	C	T	100	3	His591Tyr	CG13563
19985069	C	T	100	9	Trp36*	gek

2.2.9 Combining complementation data with the whole genome sequencing SNP results

The large number of unique SNPs in each of the mutant lines coupled with the fact that there is only a single allele available in each case prevented us from identifying the causative mutation by sequencing alone. However by combining the sequencing with the deficiency data we were able to drastically reduce the number of candidate genes. The SNPs obtained from the *B23* and the *C28* mutants were mapped onto the deficiency regions that failed to complement both by lethality and phenotype.

In case of the *B23* mutant, only one SNP identified from its whole genome sequencing corresponded to the non-complementing locus 25D5-25E6, uncovered by the deficiency *Df(2L)Exel6012*. This was the G to A mutation at chromosomal position 5465893 in the *midline* gene (Figure 2.17, A), which was predicted to convert the AG splice site acceptor in the 3rd and final intron to AA (Table 2.4), possibly disrupting the splicing in of the last 819bp exon.

On the other hand, SNPs in three different genes, *lola*, *luna* and *CG13218*, were contained within the deficiency *Df(2R)BSC595* line which failed to complement the *C28* mutant (Figure 2.17, B). However, the genes *luna* and *CG13218* are uncovered also by the smaller non-interacting deficiency *Df(2R)BSC336*, whereas the gene *lola* is deleted by the second smaller interacting deficiency *Df(2R)ED2098*. In the *C28* mutant, the SNP A to T change at position 6370617 leads to an amino acid change of Asn835Ile in the protein encoded by the *lola* locus.

Since no data was available from the lethality tests for the maternal effect mutant *A44*, all of the 22 SNPs occurring on the 2R arm of the mutant were considered. For this reason the SNP-list was further refined by ruling out genes where mutations were not likely to change gene function. Two approaches were used to determine whether an amino acid change was likely to be important. First, in some of the affected genes, the missense mutation caused the amino acid in the encoded protein to be replaced by another similar amino acid, for example,

Arg1064His in CAP, Val72Ile in Cpr49Af, Ala366Val in CG6550, etc., and these changes are not likely to affect gene function. Second, we checked the degree of conservation of the affected amino acids using the online available alignment tools Blast and ConSurf. For the alignment, the available annotated proteins from several *Drosophila* species were used. This analysis revealed that many of the affected amino acids were not conserved, and the observed mutation occurred naturally in other species.

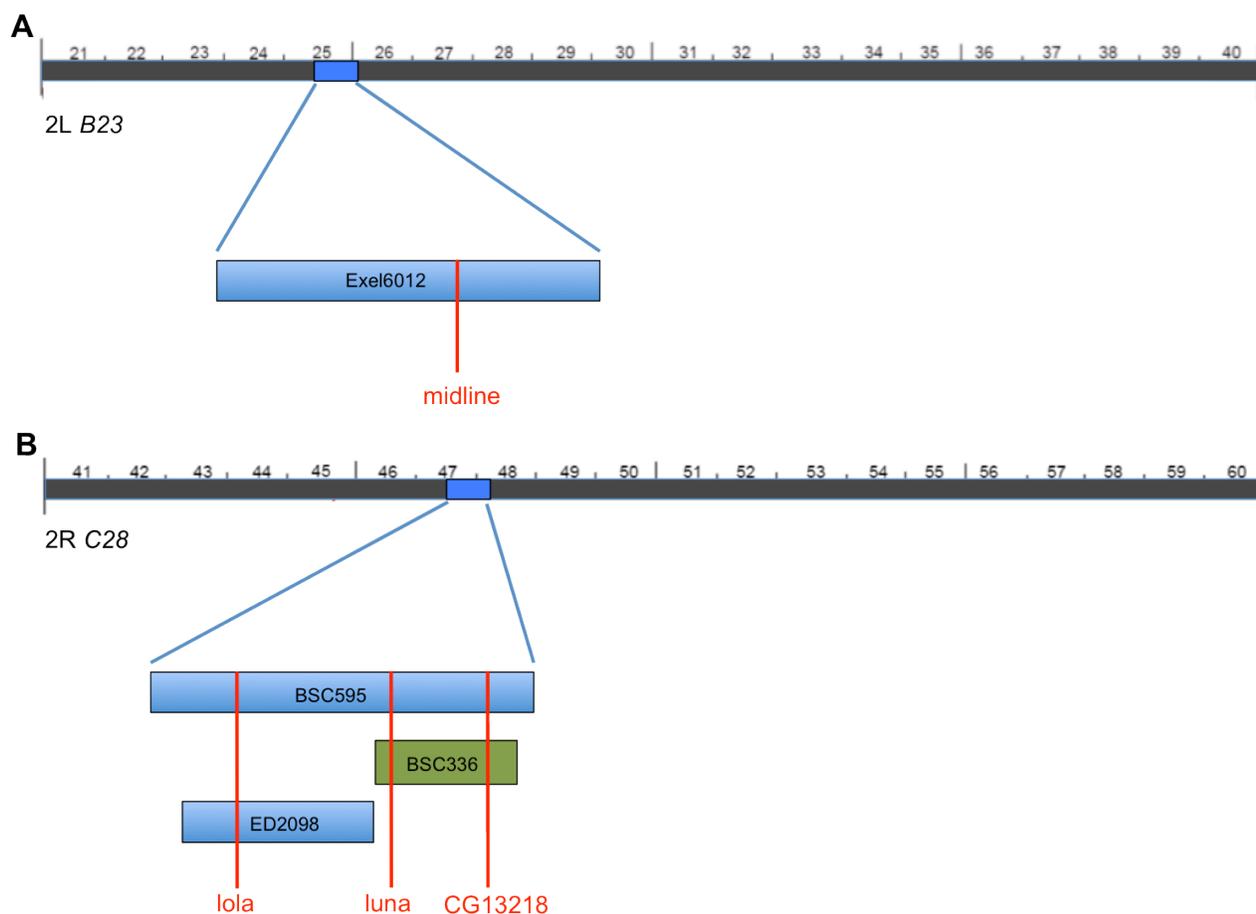


Figure 2.17: Schematic representation of the SNPs uncovered by non-complementing deficiencies for *B23* and *C28*. . A: the 2L chromosomal arm of *B23* is shown, indicating the locus depleted by the *Exel6012* deficiency (blue box). From the list of genes containing SNPs in *B23*, the only gene that maps to the deficiency locus is *midline* (red). B: From the 2R *C28* chromosome, the non-complementing deficiency *BSC595* (blue box), along with the smaller deficiencies in the region, *ED2098* (blue box) and the complementing *BSC336* (green box) are shown. *C28* contains SNPs in three genes, *lola* *luna* and *CG13218* (red) which map to these deficiency regions.

Taking all of these factors into account, 11 of the 22 affected genes were thought to be promising candidates for future studies. Deficiencies that deleted these genes were first crossed to the *A44* line and from the offspring, females containing the deficiency chromosome in trans with *A44* chromosome were selected. These females were then crossed to wild type males, and the resulting embryos were analyzed for the *A44* maternal mutant gonad phenotype, using the Vasa-stained germ cells as the read out. Of the 10 deficiencies checked, all of them complemented the *A44* gonadal defects (data not shown). Therefore, this ruled out the 11 genes: *CG12792*, *tdc1*, *sox15*, *CG34444*, *ttv*, *Pkc53E*, *ste24a*, *CG6424*, *CG10543*, *dom*, *unc-89*.

Although unlikely, one of the remaining 10 SNPs discarded from the list during refinement based on conservation, could prove to be in the causative gene. Moreover, the gene *rpL39*, coding for a ribosomal protein, is not depleted by any available deficiency, and could therefore not be tested. Therefore, due to the inability of cloning the *A44* mutant gene, the rest of the study will focus on the *B23* and *C28* mutant lines.

2.2.10 The *C28* mutation is an allele of *longitudinals* lacking

The results obtained on combining the complementation and the sequencing data, led to *lola* being the prime candidate as the causative gene in the *C28* mutant. To verify this, an available mutant of *lola*, *lola[e76]*, having an early stop in the *lola* locus and considered a protein null (Madden et al., 1999), was used to check for complementation both by lethality and phenotype. *C28* in trans to *lola[e76]* flies were lethal. Further, on observing the Vasa-labeled germ cells, these embryos showed similar gonad defects (Fig 2.18, B) as that seen in *C28* homozygotes (Fig 2.18, A). This was quantified according to the classification used earlier. As was the case with the non-complementing deficiencies, the severity of *C28* over the *lola[e76]* allele was seen to shift to the scattered germ cells and uncompact gonad category, but the overall penetrance of the mutant phenotypes continued to be high (70-80%) (Fig 2.18, C).

Due to the reduced severity of the phenotype in the transheterozygote gonads as compared to the homozygotes, deficiencies deleting 12 other SNP-containing genes were also tested by phenotype, to check for a second possible locus that contributes to the *C28* phenotype. However, all of the tested deficiencies complemented the *C28* phenotype.

Therefore, the *C28* mutation is established as an allele of the *longitudinals lacking (lola)* gene, hereafter referred to as *lola*[*C28*].

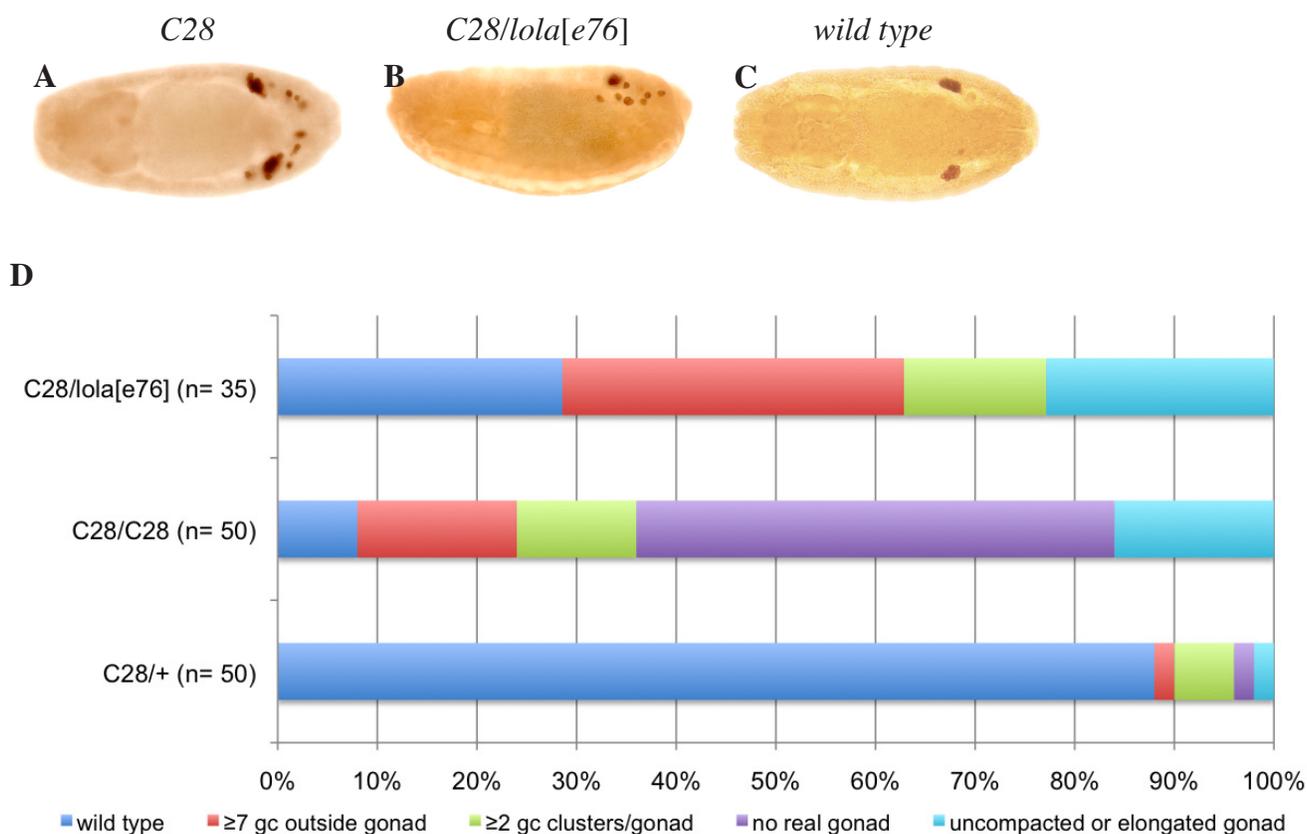


Figure 2.18: Germ cell migration phenotypes and their quantification in *C28* mutant trans to *lola* null allele. Complementation test of *lola*[*e76*] (B, C) with *C28* displays similar Vasa-labeled germ cell phenotypes in stage 15 embryos as *C28* homozygotes (A, C). The phenotype is further quantified (C) using the indicated categories (for explanation of the categorization, please see text, section 2.2.6). The *C28* mutant (the first indicated genotype is always derived from the female) crossed to the *lola* null allele show a high level of penetrance of the germ cell migration phenotype.

2.2.11 The *B23* mutation is an allele of *midline*

midline (*mid*) was the only gene containing a SNP which could affect the coding sequence, within the chromosomal region uncovered by the non-complementing deficiency *Df(2L)Exel6012*. To verify this as the causative mutation, two further *mid* mutant alleles were tested for gonad formation defects. *mid[1]* and *mid[2]* have early stop codon mutations and are null alleles (Liu et al., 2009). Both *mid[1]* and *mid[2]* were lethal with the *B23* mutant line. On analyzing the gonad phenotype in Vasa-labeled *B23/mid[1]* (Fig 2.19, B) and *B23/mid[2]* (Fig 2.19, C) embryos, both alleles showed gonad formation defects with near identical penetrance and severity, as observed in *B23* homozygotes (Fig 2.19, D).

These results clearly demonstrate that *B23* is an allele of *mid*, and hence hereafter will be referred to as *mid[B23]*. Furthermore the mutation in *mid* is causative for the gonad formation defect in this line.

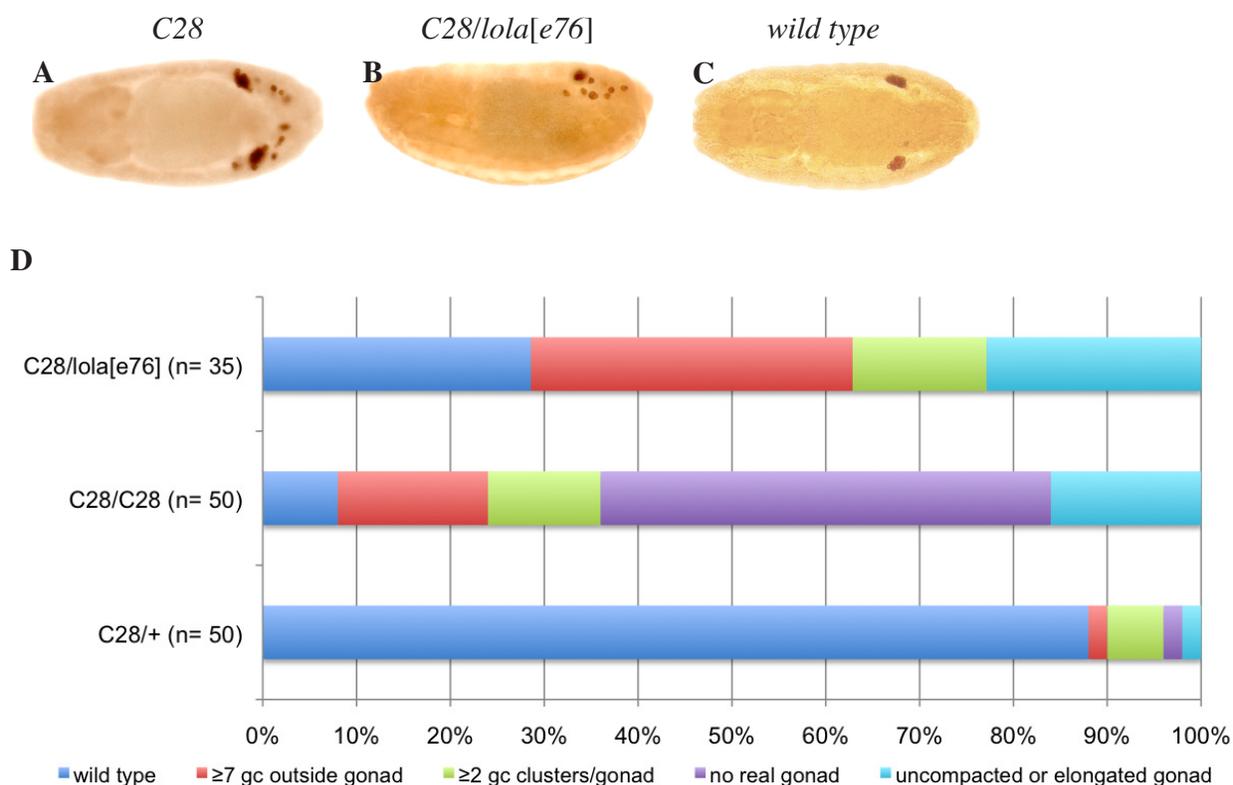


Figure 2.19: Germ cell migration phenotypes and their quantification in *B23* mutant trans to *mid* null alleles. Embryos transheterozygous for *mid[1]* (B, D) and *mid[2]* (C, D) with *B23*, display similar Vasa-labeled germ cell phenotypes in stage 15 embryos as *B23* homozygotes (A, D). The phenotype is

further quantified (D) using the indicated categories (for explanation of the categorization, please see text, section 2.2.6). The *B23* mutant (the first indicated genotype is always derived from the female) crossed to the *mid* alleles shows a high level of penetrance of the germ cell migration phenotype.

2.2.12 Sanger sequencing reaffirms the *mid*[*B23*] and *lola*[*C28*] causative SNPs

Before continuing with further analysis on the two cloned mutants, I confirmed the causative SNP data obtained from the whole genome sequencing. For this, isolated genomic DNA (gDNA) of the two mutants, obtained from homozygous mutant embryos, was used to amplify the region surrounding the site of mutation for both the *midline* and *lola* genes. The obtained amplicons were then subjected to Sanger sequencing, and the sequences were compared.

The amplified *mid* intron 3 region from *mid*[*B23*] and from *lola*[*C28*] were aligned to the available Flybase sequence. It was observed that while the last nucleotide of *mid* intron 3 in the *lola*[*C28*] sequence matched that of the annotated one, but there was a change in the nucleotide (G to A) in the *mid*[*B23*] sequence (Fig. 2.20, A). This confirmed the previously observed SNP at this location, along with reaffirming the quality of the reads obtained from the whole genome sequencing.

In a similar manner the amplified *lola* region from the two mutants were also compared with the annotated *lola* genomic and protein sequence. The *mid*[*B23*] amplicon matched the annotated sequence, while the *lola*[*C28*] amplicon revealed a nucleotide change in the second base of the Asparagine (AAT) codon converting it into an Isoleucine (ATT) codon (Fig. 2.20, B). This confirms the data obtained by whole genome sequencing.

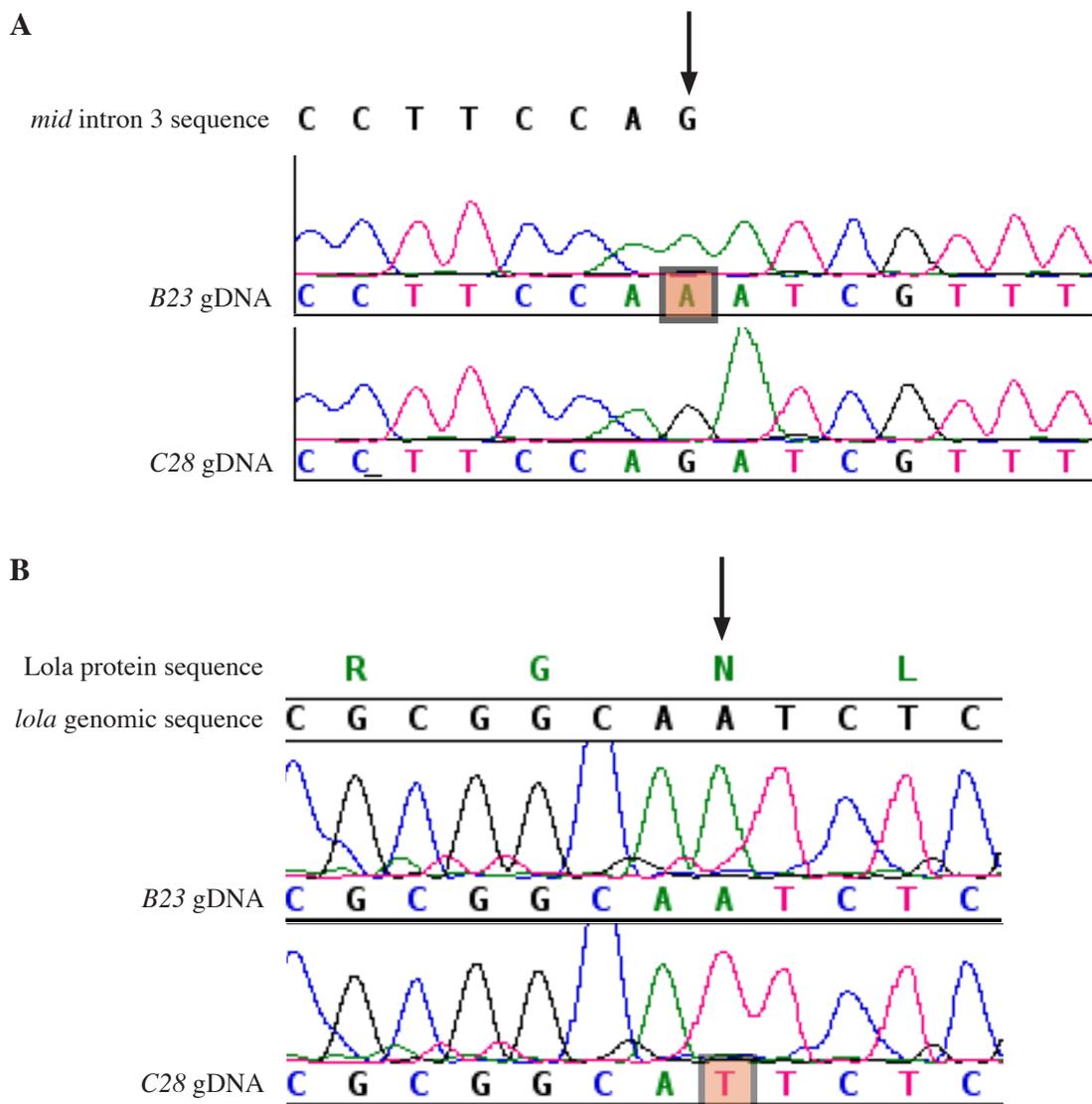


Figure 2.20: Sanger sequencing of the causative SNPs in *B23* and *C28*. The gDNA amplified *midline* intron 3 region (A) and *lola* region (B) from *B23* and *C28* is aligned with the respective Flybase annotated sequence. Sanger sequencing shows that the G base (arrow) in the splice acceptor site on the intron is mutated to A (orange box) in *B23*, but unchanged in *C28* (A). In the *lola* locus the N amino acid residue (arrow) is affected by a mutation from A to T (orange box) in *C28* (B).

2.2.13 Mutation in *mid*[B23] leads to a truncation in the T-box domain of the protein

midline (also called *neuromancer2*) was first identified in the classical screen for mutants affecting the larval cuticle (Nüsslein-Volhard et al., 1984). *midline* mutants have narrower denticle bands at the ventral midline. It encodes a transcription factor belonging to the Tbx20 (T-box 20) family (Nusslein-Volhard et al., 1984), and has been implicated in various aspects of development including heart formation (Miskolczi-McCallum et al., 2005; Qian et al., 2005), neurogenesis (Buescher et al., 2006) and limb formation (Buescher et al., 2004a). The T-box in *Mid* is the domain responsible for sequence specific DNA-binding in the promoters of the target genes (Lecuyer et al., 2008).

The *mid*[B23] mutant contains a mutation in the splice acceptor site, present at the 3' end of the last intron-exon boundary (Fig. 2.21, A), mutating the AG acceptor into a presumable non-functional AA site. Although this mutation could potentially cause the deletion of the last and largest exon, the presence of possible cryptic acceptor sites could lead to splicing within the last exon. To verify the effect of the mutation in *mid*[B23] on splicing, RNA was isolated from the *mid*[B23] mutant and wild type embryos. The synthesized cDNAs were then amplified for the *midline* transcript, and sequenced. On aligning the obtained wild type and mutant cDNA sequences with the annotated *midline* sequence, it was observed that indeed an internal cryptic acceptor site had been used. This led to a 30 base pair, or 10 amino acid deletion, in the last exon of *mid* (Fig. 2.21, B). The usage of this cryptic acceptor site preserved the reading frame, and would allow the rest of the protein to be correctly translated.

This deletion lies with the T-box domain, required for DNA-binding to target genes. To verify if the residues deleted were conserved and therefore likely to be important for function, the *D. melanogaster* *Midline* protein sequence was aligned with T-box containing proteins from zebrafish (*Danio rerio*, Tbx1), frog (*Xenopus laevis*, TBX1A), mouse (*Mus musculus*, Tbx1) and human (*Homo sapiens*, TBX1). The 10 amino acids lacking in the *mid*[B23] mutant are indeed highly conserved in the

different organisms (Fig. 2.21, C) thus we expect the mutant protein to be non-functional.

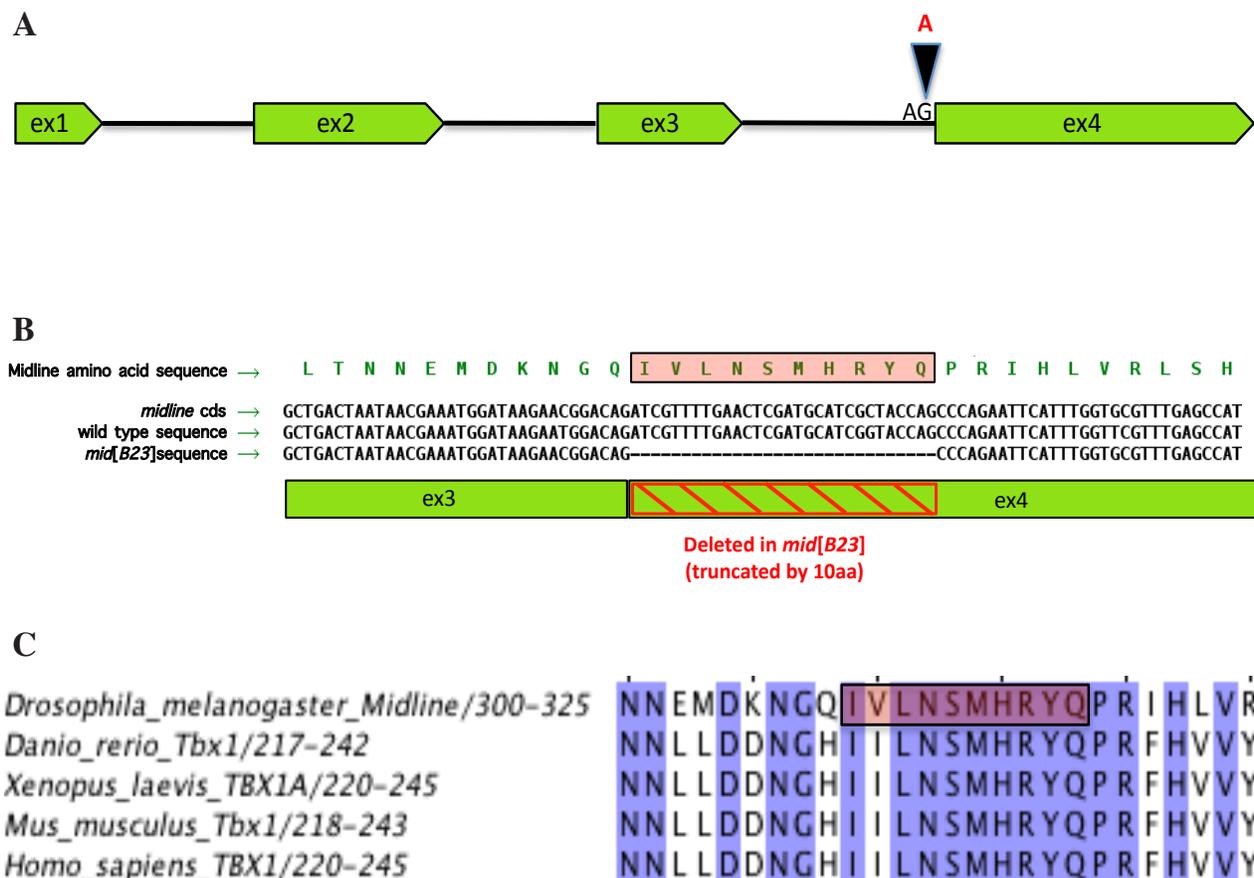


Figure 2.21: Sequencing and alignment of the deleted region in Mid[B23]. A shows the intron (black line) and exon (ex1-ex4, green boxes) arrangement in *midline*. In *mid[B23]* the AG splice acceptor site is mutated to AA (black arrow head). The cDNA obtained from *mid[B23]* along with a control cDNA were sequenced and aligned with the annotated *midline* CDS. It is seen that 30 nucleotides (shown as ----) or 10 amino acids (red box) are deleted at the ex3-ex4 (green boxes) boundary in the mutant (B). C shows the ClustalW protein alignment of the deletion region (red box) in Mid with other indicated T-box proteins from zebrafish, frog, mouse and human. Amino acids showing 100% identity are highlighted in purple.

2.2.14 *lola*[C28] mutation affects the R-isoform of Lola

The *lola* locus encodes for more than 30 different predicted isoforms of which at least 20 have been experimentally verified (Goeke et al., 2003). All of the isoforms share a common N-terminal domain containing a Broad complex, Tramtrack, Bric-a-Brac (BTB) protein dimerization domain, whereas alternate splicing leads to the varying C-termini (Barbosa et al., 2007). The isoform-specific C-termini typically contain either one or two zinc-finger motifs of the typical C2H2 or the atypical C2HC class (Goeke et al., 2003), making them nuclear-localized transcription factors (Giniger et al., 1994).

An allele of *lola*, *lola*[22.05], was recently isolated in a screen performed in search of gonad formation mutants (Weyers et al., 2011). Sequencing this allele revealed a mutation, converting the 97th amino acid of the protein from Gln to a stop codon. This premature stop early within the common region of Lola would lead to loss of all Lola isoforms.

However, the mutation in *lola*[C28] is predicted to affect only two isoforms, *lola*-G and *lola*-R (according to Flybase annotation), also referred to collectively as *lola*4.7 (Giniger et al., 1994), or *lola*-T (Goeke et al., 2003). *lola*-R and *lola*-G differ in their pre-mRNAs, however, their mature transcripts encode for the same CDS and therefore will hereafter be referred to as *lola*-R. Along with the N-terminal BTB dimerization domain, Lola-R contains a unique C-terminus with 2 Zn-fingers, one C2HC- and the other C2H2-type (Fig. 2.22, A). The Asn835Ile mutation occurs in the within the second Zn-finger motif.

To see if the mutated Asn is a conserved and likely important residue, the *D. melanogaster* Lola-R protein sequence was first used to BLAST against the proteome of other *Drosophila* and insect species. The protein from each species showing the highest % identity with Lola-R were chosen for multiple sequence alignment using ClustalW. In addition, the *D. melanogaster* Tramtrack protein, which is also a Zn-finger containing transcription factor and one of the first identified members of the BTB domain containing proteins, was added to the alignment. On focusing on the region of the second Zn-finger motif, it was observed that the two cysteines and the

two histidines forming the finger were completely conserved through the species. More interestingly, the mutated Asn in *lola*[C28] also showed 100% conservation (Fig. 2.22, B, red arrow), indicating it is very likely important for functionality.

This is supported by the structure of the Tramtrack protein, which was crystallized in complex with an 18bp oligonucleotide. The corresponding Asn residue in Tramtrack contacts directly with the oligonucleotide (Fairall et al., 1993). These studies together, indicate that the Asn835Ile mutation in *lola*[C28] leads to an isoform specific functionally defective Lola-R protein.

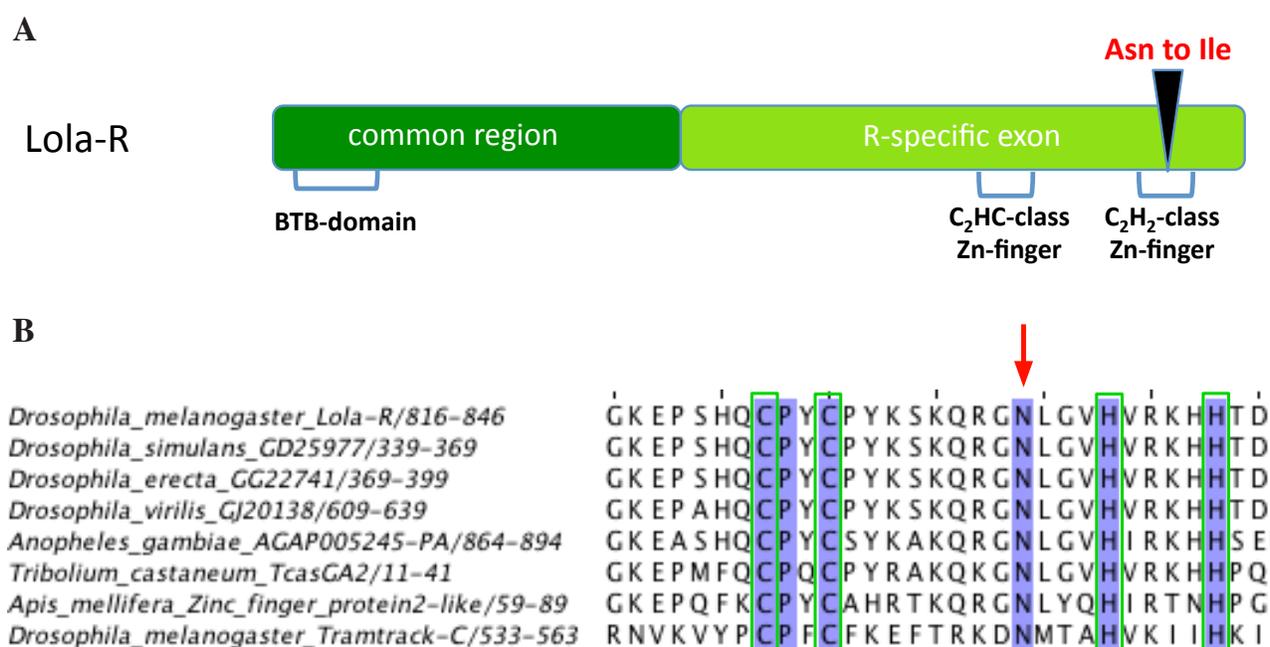


Figure 2.22: Protein domain representation and alignment of the Lola-R isoform. A shows a schematic representation of the Lola-R protein isoform with the common N-terminal region shared by all isoforms (dark green) containing the BTB dimerization domain. The C-terminal, encoded by the R-isoform specific exon (light green), contains the two Zn-finger domains. The Asn835Ile mutation (black arrow head) in *lola*[C28] occurs in the second Zn-finger domain. B shows the ClustalW protein alignment 2nd Zn finger domain in Lola-R with similar proteins from other insect species, along with the *Drosophila* Tramtrack protein. Amino acids showing 100% identity are highlighted in purple. The two cysteines and histidines (green box), along with the asparagine mutated in C28 (red arrow) are fully conserved.

2.2.15 *lola-R* displays a unique pattern of expression compared to the remaining *lola* isoforms

The different isoforms of *lola* show tissue specific expression (Goeke et al., 2003), which is highly suggestive of isoform-specific functions. To identify the tissues where the R-isoform is expressed, an antisense probe was generated against the R-specific exon (*lola-R* probe), and whole mount *in situ* hybridization was performed on embryos. In parallel, a probe against the common region shared by all *lola* isoforms was also used (*lola-c* probe). Both probes revealed maternally deposited transcripts in the preblastoderm stage 5 embryos, although the *lola-c* probe showed a much higher signal (Fig 2.23, A1, A2). In the early stages of embryogenesis, *lola-R* and *lola-c* are expressed in similar domains such as the posterior midgut (Fig 2.23, B1, C2) and the mesoderm (Fig 2.23, C1, B2) in stages 8 through 10. However, unlike *lola-R*, *lola-c* expression can be detected homogeneously in the ectoderm early on (Fig 2.23, B1) and becomes more segmental by stage 10 (Fig 2.23, C1). Moreover, while a strong expression of *lola-c* is seen in the brain lobe from stage 10 onwards (Fig 2.23, C1-F1), a low level expression of *lola-R* is seen in this tissue only from stage 13 (Fig 2.23, D2). In the following stages, the staining patterns of the probes become more restricted. While the central nervous system (CNS) and brain becomes the most prominent site of *lola-c* expression in the later stages (Fig 2.23, D1-F1), *lola-R* expression becomes restricted to the salivary glands at these stages (Fig 2.23, D2-F2).

Therefore, *lola-R* shows a very distinct pattern of expression as compared to the other *lola* isoforms. Although no prominent gonad expression could be seen in these embryos, the colorimetric Alkaline Phosphatase-based detection method used in this experiment limits the resolution.

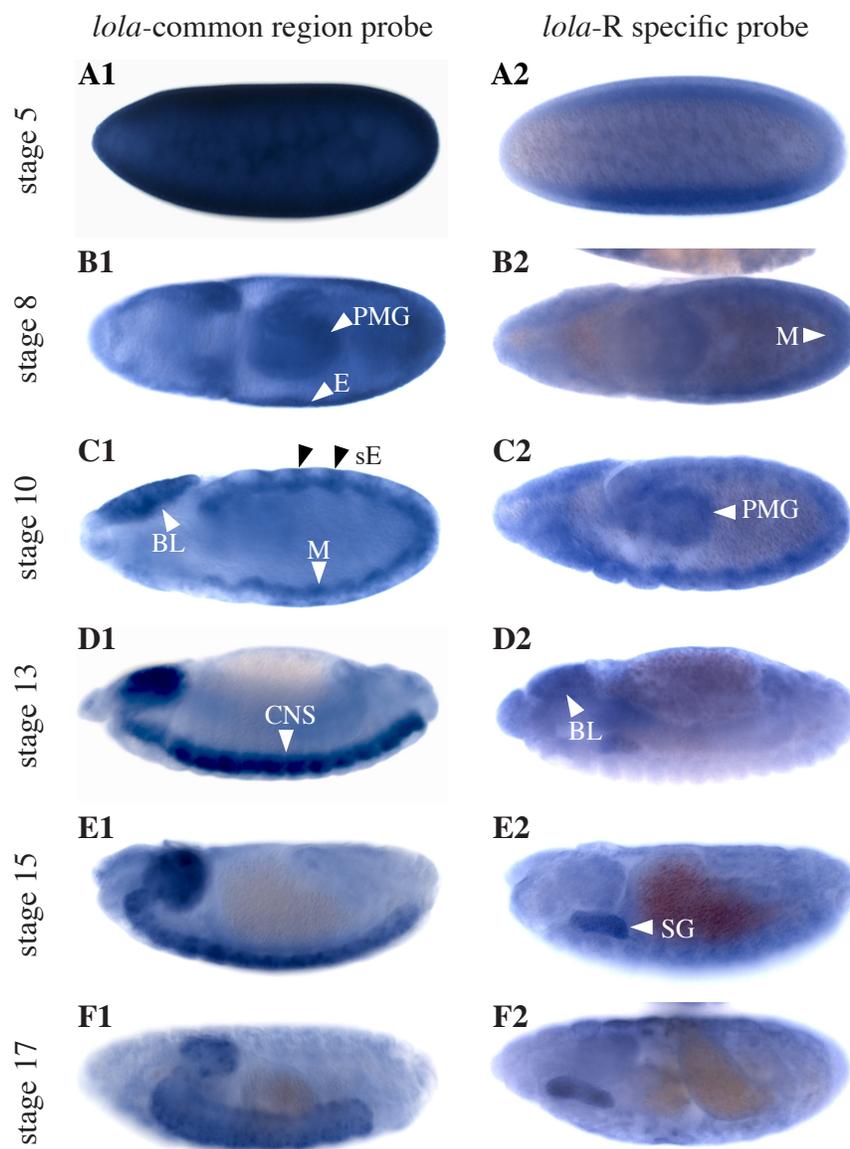


Figure 2.23: Expression pattern of all *lola* transcripts compared to the *lola-R* transcript. *In situ* hybridization using a probe against the *lola-R* specific exon (A2-F2) and a probe against the *lola*-common region (A1-F1) shared by all *lola* isoforms on whole mount embryos is shown. Both probes indicate the presence of maternal transcripts (A1, A2). Early embryonic expression of *lola-c* (B1, C1,) occurs in the posterior midgut rudiment (PMG), ectoderm (E), which become segmented (sE) in later stages, mesoderm (M) and brain lobe (BL). *lola-R* (B2, C2, D2) in these stages is also expressed in similar domains (mesoderm, primordial midgut and brain lobe), except the ectoderm. In the later stages, while *lola-c* expression (D1-F1) becomes limited to the central nervous system (CNS), high levels of *lola-R* (E2-F2) can be detected mostly in the salivary glands (SG).

2.2.16 *mid*[B23] but not *lola*[C28] shows defective VNC formation

Lack of functional Midline leads to defects in the ventral nerve cord (VNC) (Liu et al., 2009). Similarly, in agreement with its observed expression pattern, Lola is also implicated in neurogenesis (Giniger et al., 1994; Madden et al., 1999; Crowner et al., 2002). Null mutants of *lola*, containing lesions in the initial common exons also show defects in the VNC, with loss of longitudinal axons. Since both *midline* and *lola* are required for proper nervous system development, the VNC in *mid*[B23] and *lola*[C28] mutant embryos was examined. In addition, the *lola*[22.05] null allele, which also showed gonad defects, was also used.

Thereafter, the VNC of the three mutants *mid*[B23], *lola*[C28] and *lola*[22.05] was labeled using an antibody BP102 to reveal the axonal scaffold. As has been described for the *mid*[1] allele, *mid*[B23] displayed severe interruptions in the longitudinal axonal tracts (Fig. 2.24, B, white arrow). The VNC in *lola*[22.05] also showed disruptions in the longitudinal tracts (Fig. 2.24, C, white arrow), confirming the phenotype seen previously in other *lola* null mutants (Giniger et al., 1994). Interestingly, the *lola*[C28] mutant did not display any defects in the VNC (Fig. 2.24, D, white arrow), with intact axonal tracts. This correlates with *lola*-R, unlike other *lola* isoforms, not being expressed in the VNC.

This experiment confirms firstly that the deletion of 10 amino acids in the T-box domain of Midline that occurs in the *mid*[B23] mutant does affect the functionality of the protein. It also uncovers an isoform-specific allele of *lola* that separates the roles of Lola in nervous system and gonad development.

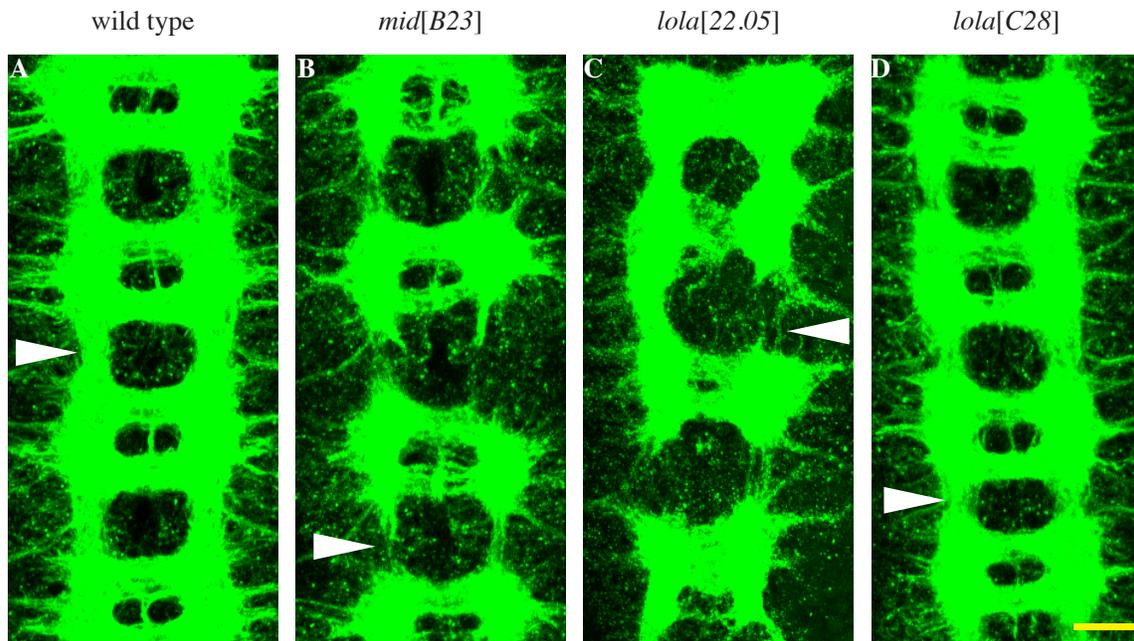


Figure 2.24: Axonal staining in *midline* and *lola* mutants. The antibody BP102 (green) is used mark the axonal scaffold in stage 16 embryos. While the longitudinal tracts (white arrow head) in the VNC of wild type (A) embryos are intact. In *mid[B23]* (B) and *lola[22.05]* (C) severe disruptions in these tracts can be observed. In *lola[C28]* (D) mutants however, the axonal tracts appear unaffected. Scale bar: 10 μ m

2.2.17 *lola*-R but not *lola*-BC expression is observed in the gonad

The previously described Alkaline Phosphatase-based *in situ* hybridization established the tissues expressing high levels of *lola*-R. However, due to background staining and the inability to independently label the gonad, it did not demonstrate whether *lola*-R is expressed in this tissue. To investigate this, fluorescent *in situ* hybridization coupled with antibody based SGP labeling was performed on whole mount embryos and imaged using confocal microscopy to obtain a better spatial resolution.

SGPs were labeled with an antibody against the nuclear protein Traffic jam (Tj) and germ cells were labeled with an antibody against Vasa. In stage 15 gonads *lola*-R probe signal could be observed surrounding the nuclear SGP labeling (Fig. 2.25, A, A'', yellow arrows) demonstrating the presence of *lola*-R transcripts in the

cytoplasm of the SGPs. These transcripts could also be visualized in the germ cells (Fig. 2.25, A, A', pink arrows), opening up the possibility of an additional role of Lola-R in germ cells.

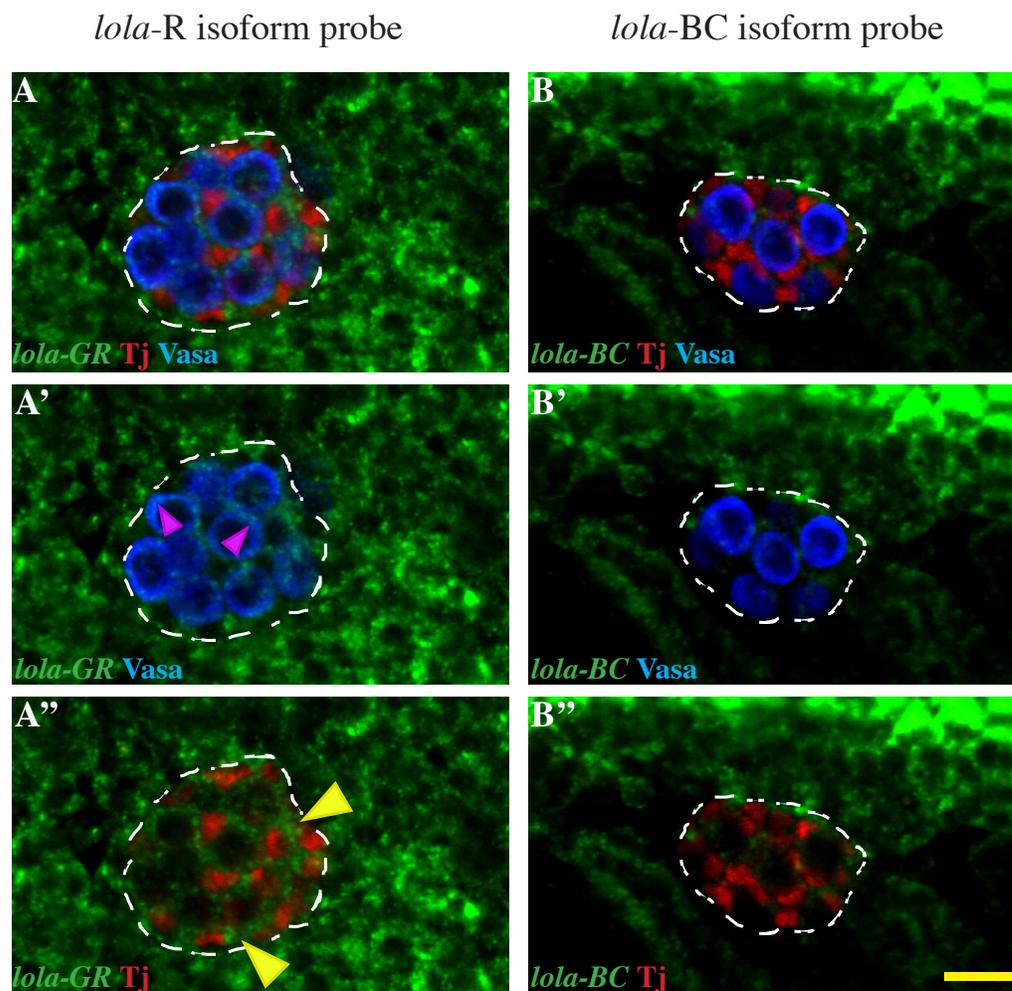


Figure 2.25: Comparative expression of *lola* isoforms in the gonad. Fluorescent *in situ* hybridization using the *lola*-R isoform specific probe (green) in a stage 15 gonad (A-A'') reveals expression, both in germ cells (stained with anti Vasa antibody) (A, A' pink arrow heads) and in SGPs (stained with anti TJ antibody) (A, A'' yellow arrow heads). The *lola*-BC probe (green) (B-B'') indicates the lack of expression of this isoform in both germ cells (B, B') and SGPs (B, B''). Scale bar: 10 μ m

As expected, the *lola*-c probe also showed a similar pattern of expression in and around the gonad (data not shown). However, to verify the specificity of the labeling observed for *lola*-R, a probe against another isoform of *lola* was used. An isoform

specific mutant for *lola*-BC (also known as *lola*-L), *lola*[*ORE119*], displayed no gonad formation defects (data not shown). Therefore, embryos labeled with a probe against the *lola*-BC specific exon were examined. These embryos did not show any SGP or germ cell specific labeling (Fig. 2.25, B, B', B'') demonstrating that the observed gonad expression of *lola*-R is specific.

2.2.18 Midline and Lola protein can be detected in the SGPs

By *in situ* hybridization, *midline* expression was detected in ectodermal segmental stripes, in the heart, and can be observed in region of the SGPs (Brook and Cohen, 1996; Buescher et al., 2004b). Mid protein expression has been reported in the heart and many neurons in the CNS (Moore et al., 1998b). However Mid protein expression has not been reported in SGPs and the importance of the gonadal *midline* expression has not been previously investigated.

To check if the SGP *mid* RNA expression correlated with the protein expression, an available antibody against the Midline protein was utilized. The Mid antibody showed a clear labeling of the heart cells as expected. Moreover, on inspecting the SGPs, counter-labeled with Tj, in stage 13 embryos, Mid protein could be observed co-localizing with the Tj signal (Fig. 2.26, A, A' white arrows). The nuclear localization of both Tj and Mid is expected as they both are transcription factors. Thus Mid is expressed in SGPs.

To detect the Lola protein an antibody made against the common region of Lola was used (Giniger et al., 1994). This should detect all possible Lola isoforms. The labeling pattern with the Lola antibody was very similar to the observed *lola*-c *in situ* hybridization pattern, with high expression in the CNS at later stages. In stage 13 embryos, nuclear Lola protein could be detected colocalizing with the Eya SGP marker (Fig. 2.26, B, B' white arrows). Therefore Lola is expressed in SGPs. In addition, Lola could also be detected in germ cell nuclei (Fig. 2.26, B, B' pink arrows).

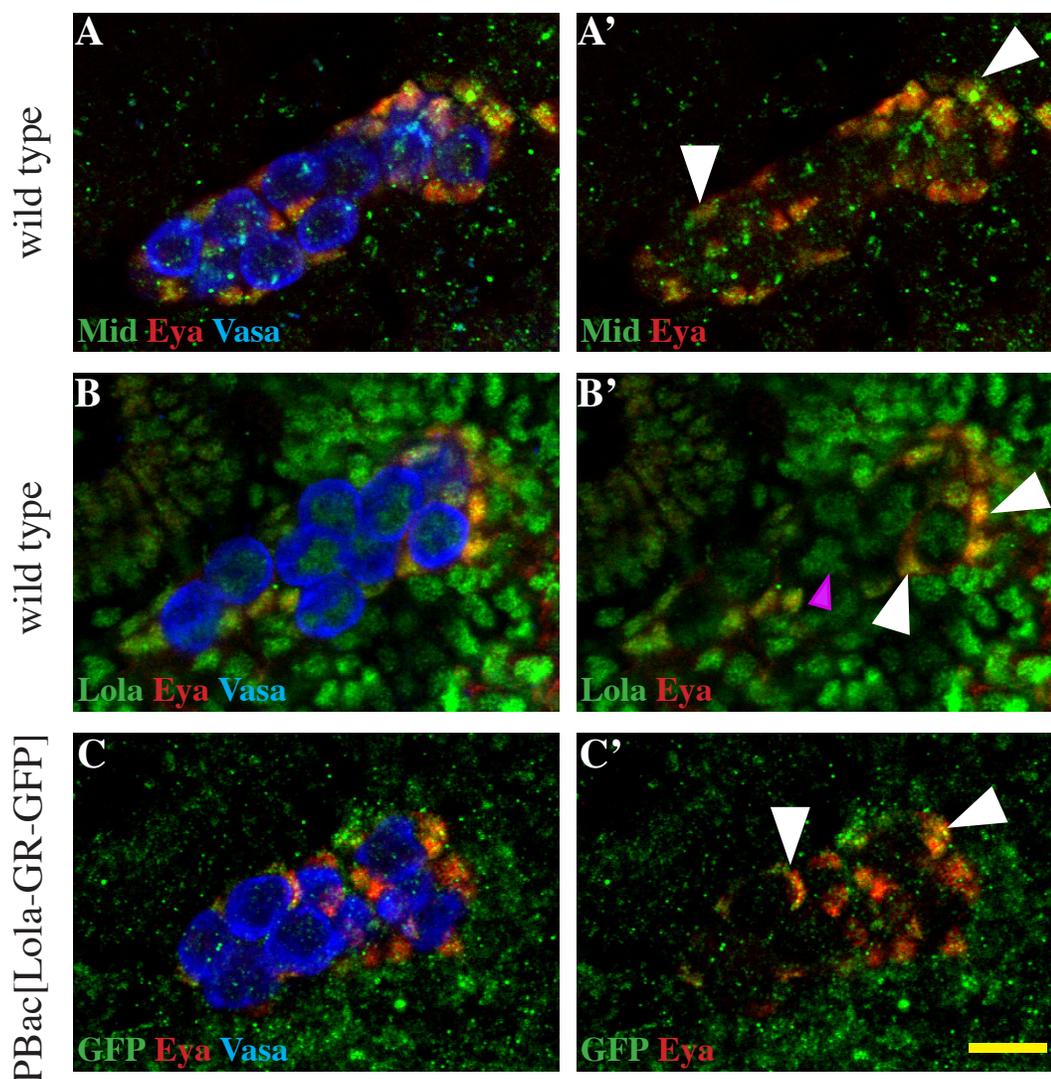


Figure 2.26: Midline and Lola protein expression in the gonad. In wild type stage 13 embryos, Mid protein (green) is seen in the SGPs (red) counter-stained with Eya antibody (A, A' white arrow heads), but no specific staining is observed in the germ cell (blue) nuclei. On using the pan-Lola antibody, Lola protein localizes to SGP nuclei (B, B' white arrow heads) and germ cell nuclei (B, B' pink arrow head). In stage 13 gonads of embryos containing a BAC insertion with GFP inserted 3' of the *lola-R* isoform. GFP expression (green) is seen in the SGPs (C, C' white arrow heads), but not in the germ cells. Scale bar: 10 μ m

To ascertain if the Lola-R isoform was expressed in the gonad, GFP expression in embryos containing a BAC (bacterial artificial chromosome)-insertion of the genomic piece of the *lola* locus, in the third chromosome at 65B2 cytological position, with a GFP-tag 3' of the R-specific exon of *lola* was examined. To avoid

potential maternal Lola-R GFP contribution, male flies containing the BAC-insertion line were crossed to wild type females. GFP was observed in mesoderm of early embryos and the brain lobe and the salivary glands in later stage embryos, which is in agreement with the *lola-R* RNA expression pattern (Fig 2.23). Lola-R-GFP could additionally, be observed in the Eya co-labeled SGPs (Fig. 2.26, C, C' white arrows). This was in agreement with the observed *lola-R* fluorescent probe labeling of the gonad. However, the germ cell expression seen by *lola-R* *in situ* hybridization and the Lola antibody labeling was not observed with the Lola-R-GFP BAC-insertion line. The germ cell *lola-R* signal is therefore likely due to maternally deposited transcripts. Unfortunately, the relevance of expression of *lola-R* in germ cells is difficult to study due to the poorly developed embryos layed by *lola*[C28] germ-line clone females.

2.2.19 *twiGal4* drives expression in the SGPs

To verify that Mid and Lola are both required autonomously in the SGPs for gonad formation I performed tissue specific rescue experiments using the UAS-Gal4 system. *Dsix4Gal4* (*Drosophila* Six4, a transcription factor expressed in the SGPs) is a SGP-specific driver. However, it turns on expression of the transgene in later stages (starting from approximately stage 13, data not shown), when the SGP phenotype in *mid*[B23] and *lola*[C28] embryos have already manifested.

Since SGPs are mesodermally derived cells, an early pan-mesodermal driver such as *twiGal4* could also be used. To test if the *twiGal4* drove expression of a transgene early in the SGPs, the driver line was crossed to a *UASlola-R-GFP* line. In addition to the *twiGal4*, this line also contained the 68-77 lacZ enhancer trap line, which marks the SGPs with lacZ.

GFP expression could be detected throughout the mesoderm of early stage embryos (Fig. 2.27, A). Lola-R-GFP was observed in the nucleus of the LacZ-labeled SGPs at stage 12 (Fig. 2.27, A', A''). Although the expression of the transgene in the whole embryo became more restricted (Fig. 2.27, B), SGP-specific expression of GFP could still be seen at stage 13 (Fig. 2.27, B', B''). But towards the end of gonad

morphogenesis, around stage 15, the *twiGal4*-induced expression of the transgene was no longer detected in the embryo including the gonad (Fig. 2.27, C, C', C''). However, since it appears that the Lola-R and Mid proteins are expressed and required from earlier stages, this loss of transgene expression at later stages should not affect our ability to rescue the gonad formation defects. Therefore *twiGal4* is a suitable driver for SGP rescue experiments.

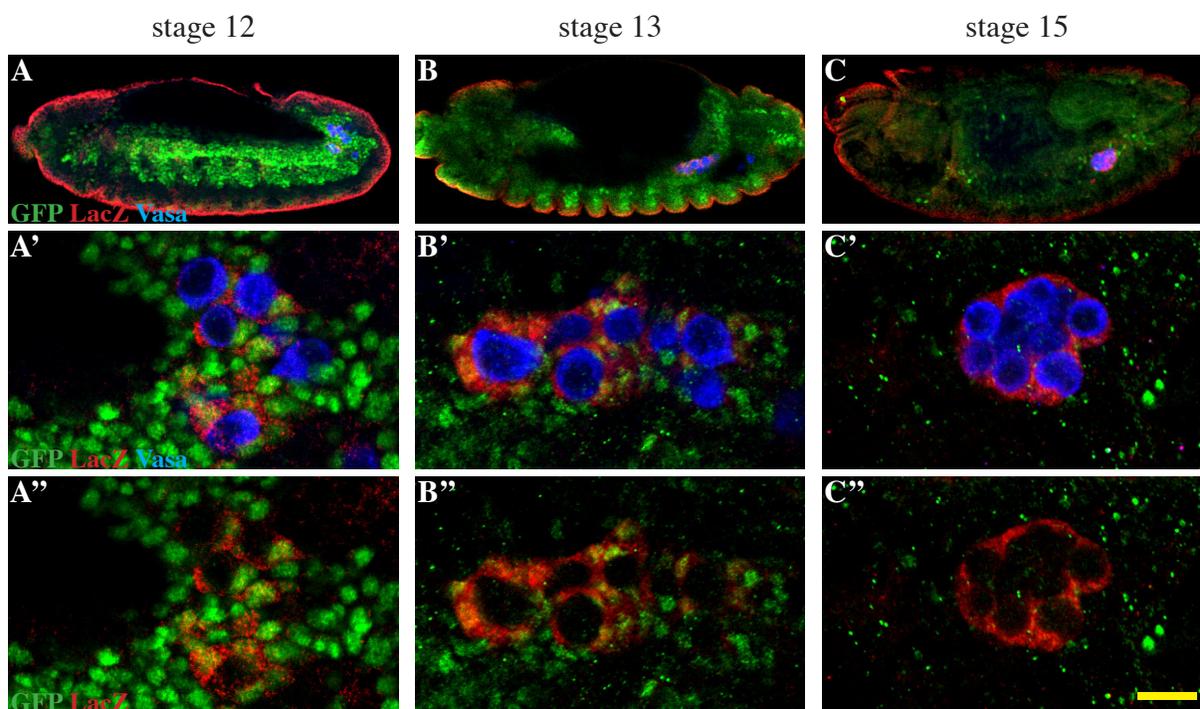


Figure 2.27: *twiGal4* driver expression of *UASlola-R-GFP* transgene in the mesoderm. Transgene expression (green, using GFP antibody) is observed in the SGPs (red, LacZ expression from the 68-77 enhancer trap), but not in germ cells at stage 12 (A-A''), using the *twiGal4* driver. The driver-induced expression becomes restricted at stage 13 in the embryo but is still present in the SGPs (B-B''), but by stage 15, transgene expression cannot be detected in the gonad (C-C''). Scale bar: 10 μ m

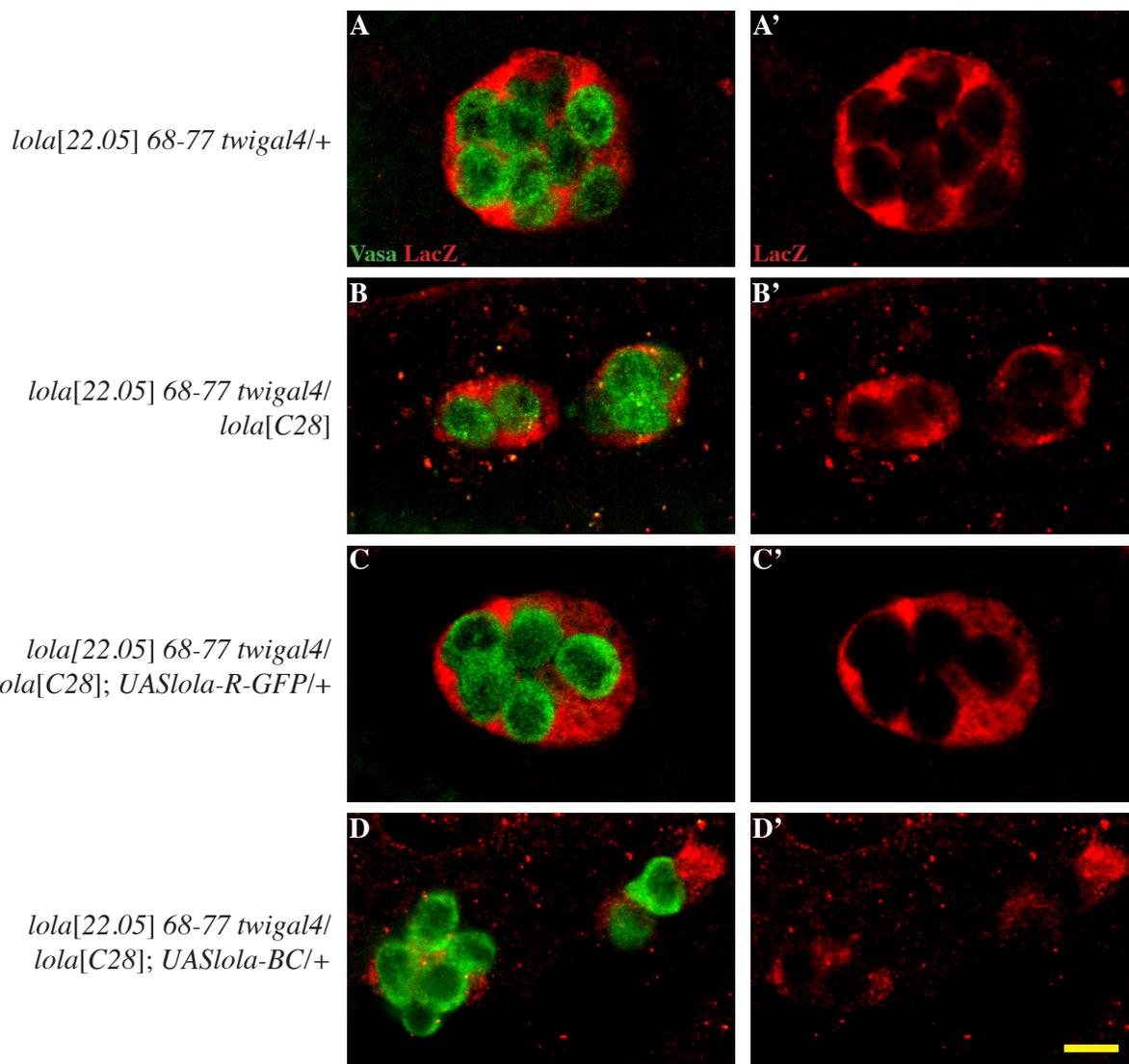
2.2.20 *lola-R* but not *lola-BC* can rescue the *lola[C28]* mutant gonad defects

To ascertain if the expression of the Lola-R or other Lola isoforms in the SGPs can rescue the gonad defects of *lola[C28]* mutants, the *lola[22.05]* 68-77 line was recombined with the *twiGal4* driver and the resulting recombinants checked for gonad defects. The *lola[22.05]* allele contains a mutation converting the 97th amino

acid of the protein from Gln to a stop codon (data not shown) leading to loss of all Lola isoforms and therefore this allele is a null.

As expected, no defects were observed with the *lola*[22.05] 68-77 *twiGal4* line when heterozygous (Fig. 2.28 A, A'). Using this line in trans with *lola*[C28] (Fig. 2.28 B, B'), however lead to a high penetrance of gonad defects similar to that observed in the case of homozygous *lola*[C28] mutants (Fig. 2.13 D1, D2). The rescue of the defects in the transheterozygotes was attempted using both the *UASlola-R-GFP* transgene and a *UASlola-BC* transgene, which encodes a functional Lola isoform containing a BTB dimerization domain and two Zn-finger DNA-binding domains (Goeke et al., 2003).

Embryos expressing *lola-R-GFP* under the *twiGal4* driver were able to rescue the gonad defects of *lola*[22.05]/*lola*[C28] transheterozygotes, leading to round and compact stage 15 gonads and few lost germ cells (Fig. 2.28 C, C'). The *UASlola-BC* transgene on the other hand, was unable rescue *lola*[22.05]/*lola*[C28] transheterozygotes, the gonads of which displayed the full spectrum of phenotypes with high penetrance (Fig. 2.28 D, D'). Importantly, expression of the Lola-BC in the mesoderm in a wild-type background doesn't disrupt gonad formation (data not shown). For quantification (Fig. 2.28 E) of the phenotype in these genotypes, it was possible to make use of the lacZ enhancer trap labeling the SGPs, along with the usual labeling of the germ cells using Vasa. This led to the replacement of the '≥2 gc clusters/gonad' and the 'no real gonad' categories with '≥2 SGP clusters/gonad', as the state of SGPs could directly be used to assess the phenotype. While the *lola*[22.05] 68-77 *twiGal4*/+ embryos demonstrated 100% wild type gonads, the transheterozygous *lola*[22.05] 68-77 *twiGal4*/ *lola*[C28] embryos had nearly 90% mutant gonads with nearly half of them having 2 or more unfused clusters of SGPs. Further, these defects can be rescued in nearly 80% of the embryos on using the *UASlola-R-GFP* transgene. The Lola-BC isoform on the other hand didn't show any rescue and continued to display 90% defective gonads. This proves that the functions performed by the R-specific isoform of Lola in the formation of gonads cannot be replaced by the BC isoform and suggests that the downstream targets of the different isoforms are not identical.



E

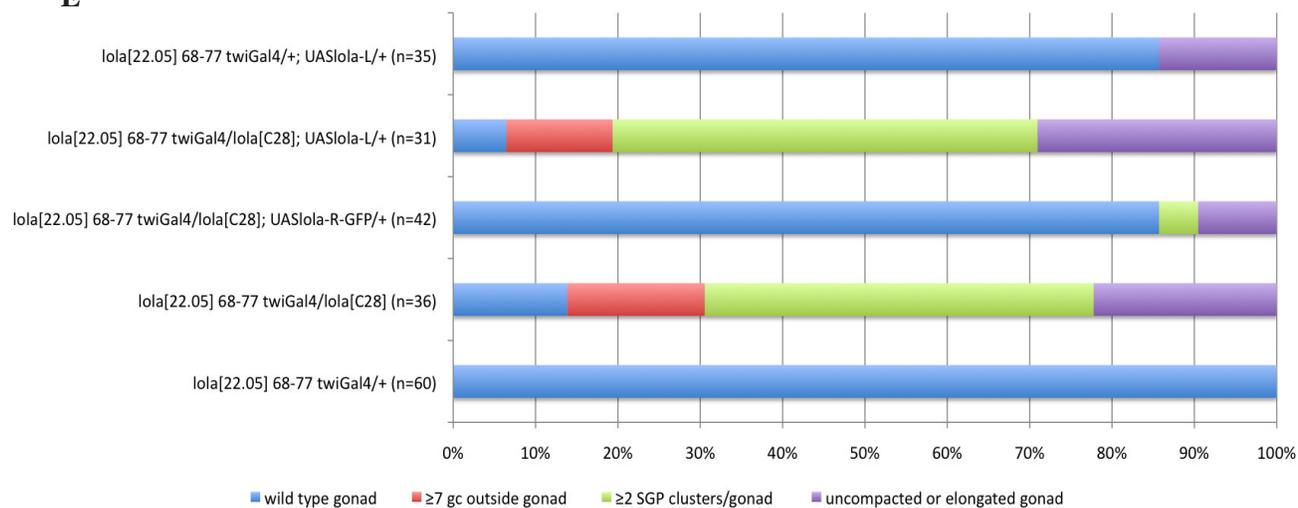


Figure 2.28: Rescue of *lola*[C28] gonad defects using different *lola* isoforms. The recombined *lola*[22.05] 68-77 *twiGal4* line when heterozygous (A, A') has normal stage 15 gonads, however when in trans with *lola*[C28] (B, B'), shows defective gonads. While these defects in gonad formation can be rescued by expression of the Lola-R-GFP (C, C') in the mesoderm, the *lola-BC* isoform (D, D') is unable to do so. The graph (E) quantifies the phenotype using fluorescent germ cell and SGP labeling according to the indicated categories. Only expression of the *lola-R* transgene in the mutants rescues the mutant gonad phenotypes with 80-90% of gonads rescued to wild type status. Scale Bar: 10µm

2.2.21 Expression of *midline* in the mesoderm rescues the gonad phenotypes in *mid*[B23] mutants

As has been established thus far, lack of functional *midline* leads to defects in several steps for gonad formation during embryonic development. To prove that Midline is required autonomously in the SGPs, a rescue experiment was attempted in the *mid* mutants. The *twiGal4* driver was used to express the inducible *UASmid* transgene in the mesoderm of the *mid*[1]/*mid*[B23] transheterozygotes.

As expected, the recombined *mid*[1] *twiGal4* line, when heterozygous (Fig. 2.29 A, A') did not show any gonad defects. *mid*[1] *twiGal4*/*mid*[B23] transheterozygous embryos displayed severe gonad formation defects, similar to that in *mid*[B23] (Fig. 2.13 C1, C2, Fig. 2.29 B, B'). On expression of the *UASmid* transgene in the mesoderm, including the SGPs, in such a *midline* mutant background, the SGP-related defects, including the lack of fusion of the SGP clusters and the ensheathment of the germ cells by the SGPs, were rescued (Fig. 2.29 C, C'). Moreover the scattered germ cells observed at this stage were also rescued.

The rate of the observed rescue was further quantified (Fig. 2.29 D) using Vasa-labeled germ cells. While *mid*[1] *twiGal4* /+ embryos showed 80% wild type gonads, similar to *mid*[B23] embryos, the transheterozygous *mid*[1] *twiGal4*/*mid*[B23] displayed a 100% penetrance for mutant gonads, with the highest severity category being most represented. The *mid*[1] *twiGal4*/*mid*[B23]; *UASmid* embryos, on the other hand showed an approximately 75% wild type gonads at stage 15.

This clearly demonstrates that Midline is required in the SGPs for normal gonad morphogenesis.

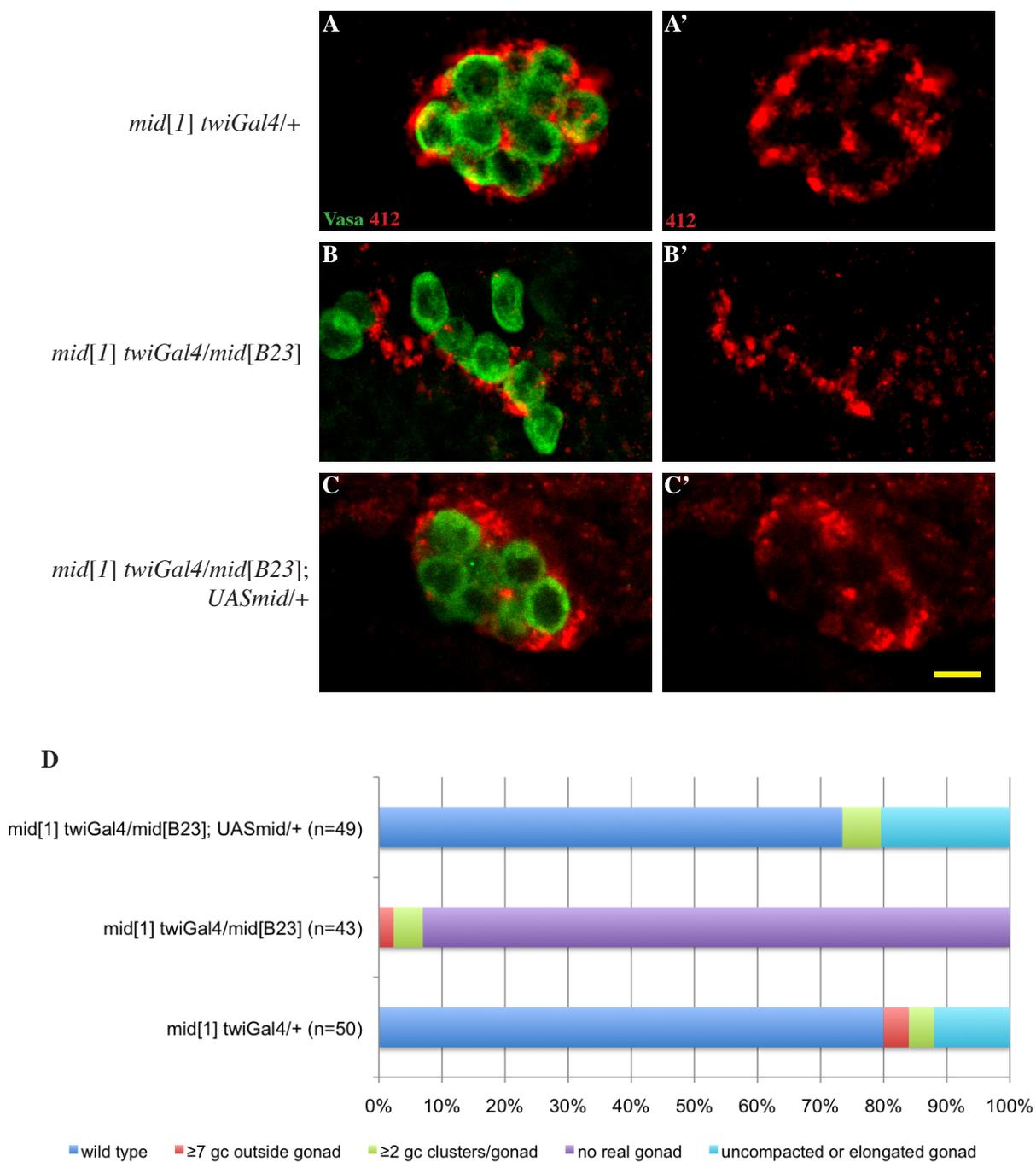


Figure 2.29: Rescue of *mid[B23]* gonad defects using *midline* overexpression. The recombinant *mid[B23] twiGal4* line when heterozygous (A, A') has normal stage 15 gonads, however when in trans with *mid[B23]* (B, B'), shows defective gonads. These defects in gonad formation can be rescued by expression of the *UASmid* (C, C') transgene in the mesoderm. The graph (E) further quantifies the phenotype using the indicated categories (for explanation of the categorization, please see text, section 2.2.6). Expression of the *UASmid* transgene in the mutants rescues the mutant gonad phenotypes with 75% of gonads rescued to wild type status. Scale Bar: 10 μ m

2.2.22 Stage specific loss of Traffic Jam in *mid[B23]* is rescued by mesodermal Midline expression

Traffic jam (Tj) is a transcription factor expressed by SGPs and required for the process of germ cell ensheathment, (Li et al., 2003). One of the defects observed in *mid[B23]* mutant embryos is the complete lack of germ cell ensheathment (section 2.2.3; Fig. 2.13 C1, C2), in spite of early SGP-germ cell contact initiation. Therefore I tested whether Tj was downstream of Mid using an available antibody against Tj.

In addition to the VNC expression, Tj could be observed in the nucleus of SGPs, co-labeled against the *412* RNA in stage 13 (Fig. 2.30 A1, A1') and stage 15 (Fig. 2.30 A2, A2') *midline* heterozygous embryos. However, in stage 13 *mid[1] twiGal4/mid[B23]* transheterozygotes (Fig. 2.30 B1, B1'), in spite of the presence of 412-labeled SGPs, no nuclear or cytoplasmic Tj expression could be observed. Interestingly Tj expression was seen in the nucleus of a few stage 15 mutant gonads (Fig. 2.30 B2, B2'). However, this late Tj expression was not consistent in all embryos or SGPs. Moreover, the VNC Tj was unaffected no matter the stage (data not shown).

If Mid is required autonomously in SGPs for robust Tj expression, then Tj expression should be restored in *mid* mutant embryos rescued with mesodermally driven Mid. Indeed, the gonads in the *mid[1] twiGal4/mid[B23]; UASmid* rescue embryos showed Tj expression at stage 13 (Fig. 2.30 C1, C1'). Furthermore, the germ cells were ensheathed by the SGPs. Similarly, the stage 15 rescued gonads showed wild type Tj levels, with all the SGPs showing expression (Fig. 2.30 C2, C2').

However, neither *A44* (data not shown) nor *lola[C28]* (Fig. 2.30 D1, D2) mutants displayed any deficiency in SGP Tj expression or localization. This indicates that gonad-specific Tj expression is indeed regulated by Mid, but is not affected in other mutants having gonadal defects.

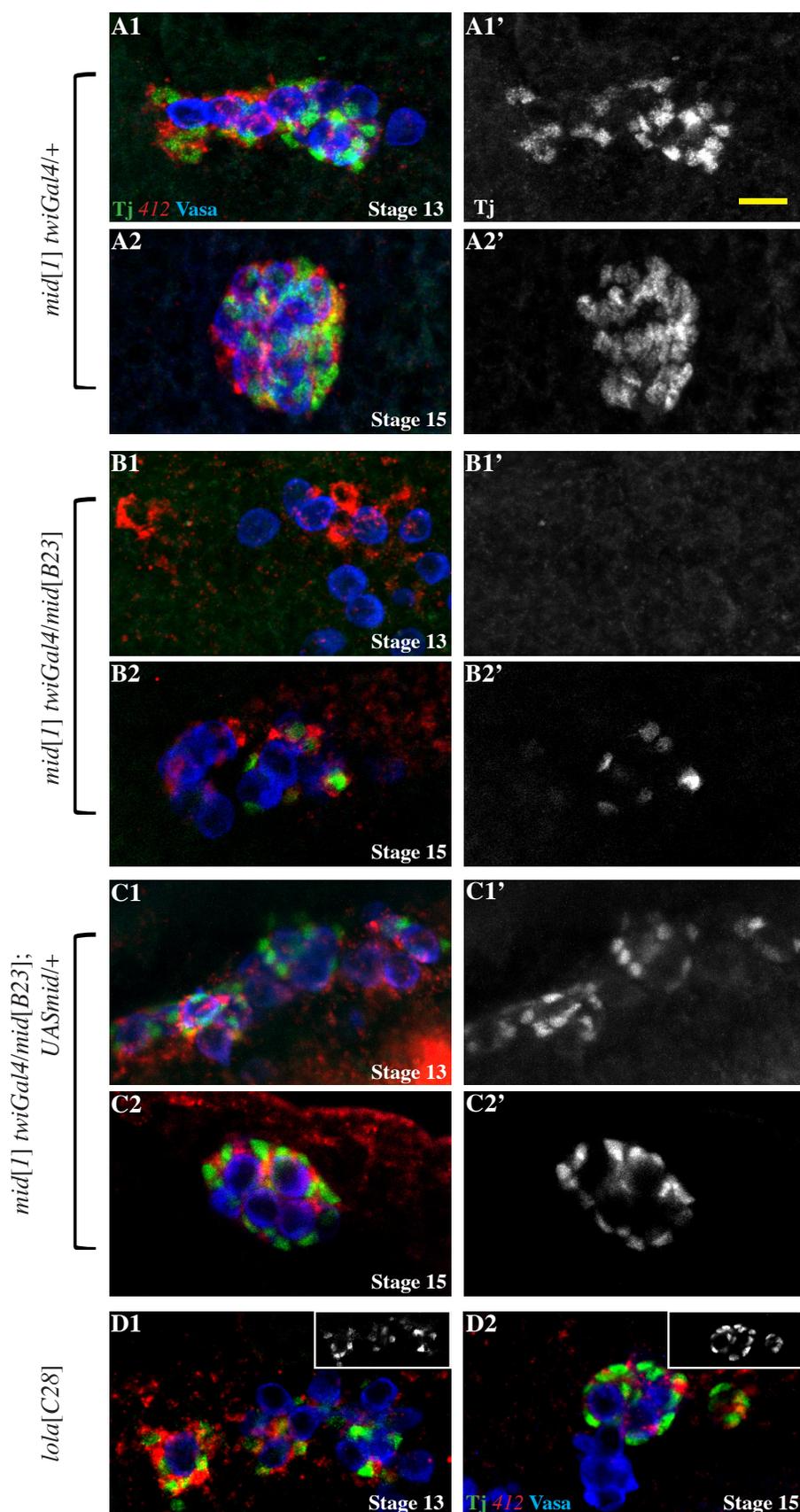


Figure 2.30: Rescue of Traffic jam expression in *mid[B23]* gonads. Vasa-labeled germ cells are in blue, 412 labeled SGPs in red, and Traffic jam in green or grey. Heterozygous *mid[1] twiGal4/+* gonads at stage 13 (A1, A1') and stage 15 (A2, A2') display normal Tj levels in SGP. *mid[1] twiGal4/mid[B23]* embryos show loss of Tj expression in the 412 labeled SGP of stage 13 gonads (B1, B2). Patchy Tj expression is observed in mutant stage 15 gonads (B2, B2'). This loss/reduction of Tj protein in the gonad is rescued on expressing *UASmid* transgene (C1, C1', C2, C2') in the mesoderm of the mutants. *lola[C28]* mutants show normal levels of Tj expression in stage 13 (D1, inset) and stage 15 (D2, inset) gonads. Scale Bar: 10 μ m

2.2.23 DE-cadherin, Armadillo and Robo are expressed in *mid[B23]* and *lola[C28]* mutant gonads

Since both Mid and Lola are transcription factors, they could potentially regulate numerous downstream genes. A number of genes are already known to be involved in gonad morphogenesis (Table 1.2), and I wanted to examine if some of them are regulated by Mid or Lola.

The adhesion proteins, DE-cadherin and Arm, are required in various steps of gonad formation, especially gonadal compaction at stage 15 (Jenkins et al., 2003; Mathews et al., 2006) (Warrior, 1994). Therefore, antibodies against DE-cad and Arm were used to examine the localization and levels of these proteins in wild type and in *mid[B23]* and *lola[C28]* mutant gonads. In wild type embryos, DE-cad (Fig. 2.31 A-A'') and Arm (Fig. 2.32 A, A') were observed both in the germ cells (pink arrow head) and in the SGPs (yellow arrow head), and localized to SGP-germ cell junctions. In the mutants *mid[B23]* (Fig. 2.31 B-B'') and *lola[C28]* (Fig. 2.31 C-C'') DE-cad was expressed by the germ cells and the SGPs although I cannot exclude that there is a small decrease in the labeling intensity in *mid[B23]* stage 15 gonads. Moreover both cell types localized DE-cad to the sites of cell contact. Similar observations were made in the case of Arm-labeling. In the mutant gonads (Fig. 2.32 B, B', C, C'), both SGPs and germ cells expressed and localized Arm similar to that observed in wild type gonads. Additionally, other tissues, like the gut and the trachea expressed normal levels of the two adhesion proteins in the mutants (data not shown, bright DE-cad staining in the gut can be observed in all panels of Fig 2.31). Therefore, no gross differences could be observed in the DE-cad and Arm expression between wild type and mutants.

Another protein implicated in gonad formation at the step of SGP cluster fusion and gonadal compaction, is the cell surface receptor Robo (Weyers et al., 2011). Moreover, both Mid (Liu et al., 2009) and Lola (Crownier et al., 2002) have been demonstrated to regulate Robo levels in the CNS, making it a likely target in the gonads. In wild type stage 15 embryos (Fig. 2.33 A, A'), Robo localized to the cell

surface of both SGPs (yellow arrow head) and germ cells (pink arrow head). In *mid[B23]* (Fig. 2.33 B, B') and *lola[C28]* (Fig. 2.33 C, C') mutants the germ cells could be seen with Robo signal at their cell surface. Furthermore, the SGPs of mutants also expressed Robo at levels not distinguishable from wild type gonads.

Hence three factors, DE-cad, Arm and Robo, implicated in gonad formation appear unaffected in the *mid[B23]* and *lola[C28]* mutants.

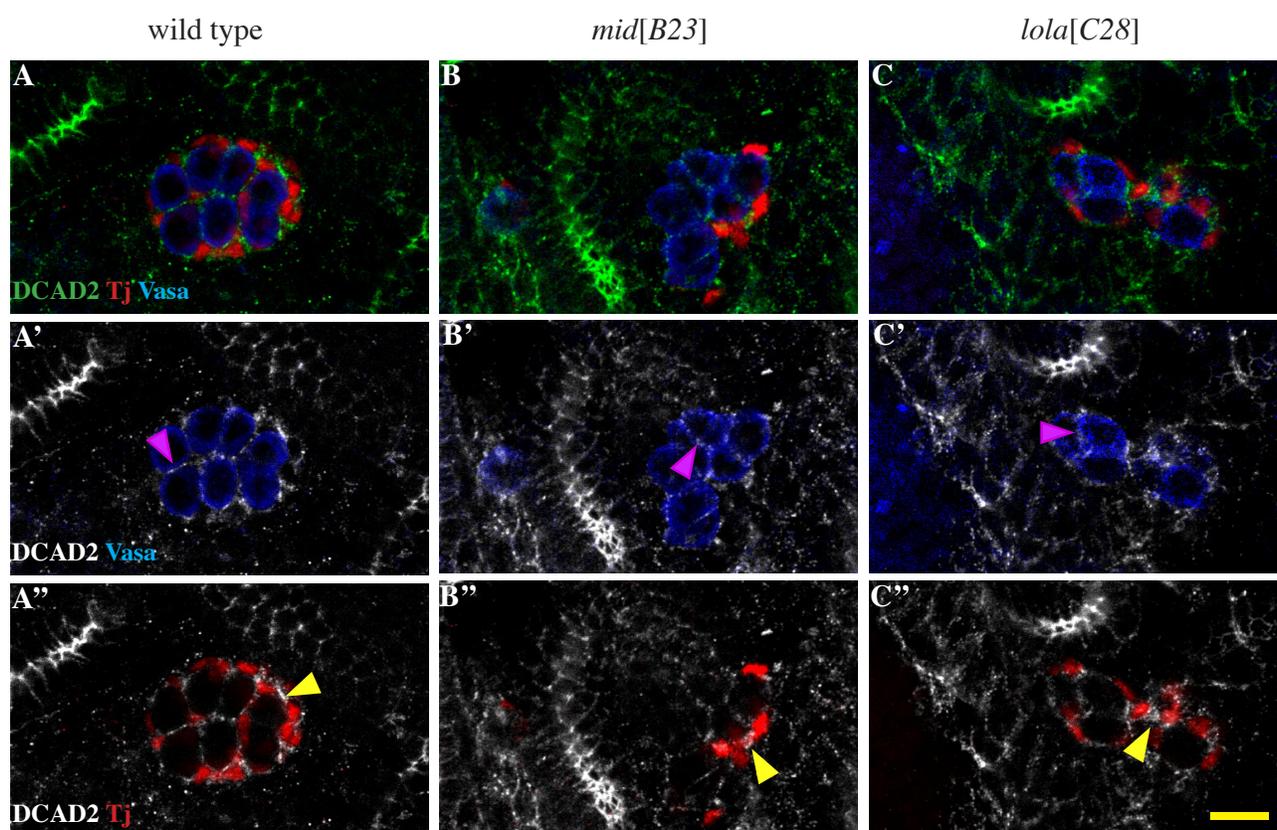


Figure 2.31: DE-cadherin levels in stage 15 *mid[B23]* and *lola[C28]* mutant gonads. In wild type (A-A'') DE-cad (green, gray) expression is observed in both Vasa-labeled (blue) germ cells (A' pink arrow head) and Tj labeled (red) SGPs (A'' yellow arrow head). Similar to wild type, *mid[B23]* (B-B'') and *lola[C28]* (C-C'') mutant gonads also display germ cell (B', C' pink arrow head) and SGP (B'', C'' yellow arrow head) specific DE-cad expression. Scale Bar: 10 μ m

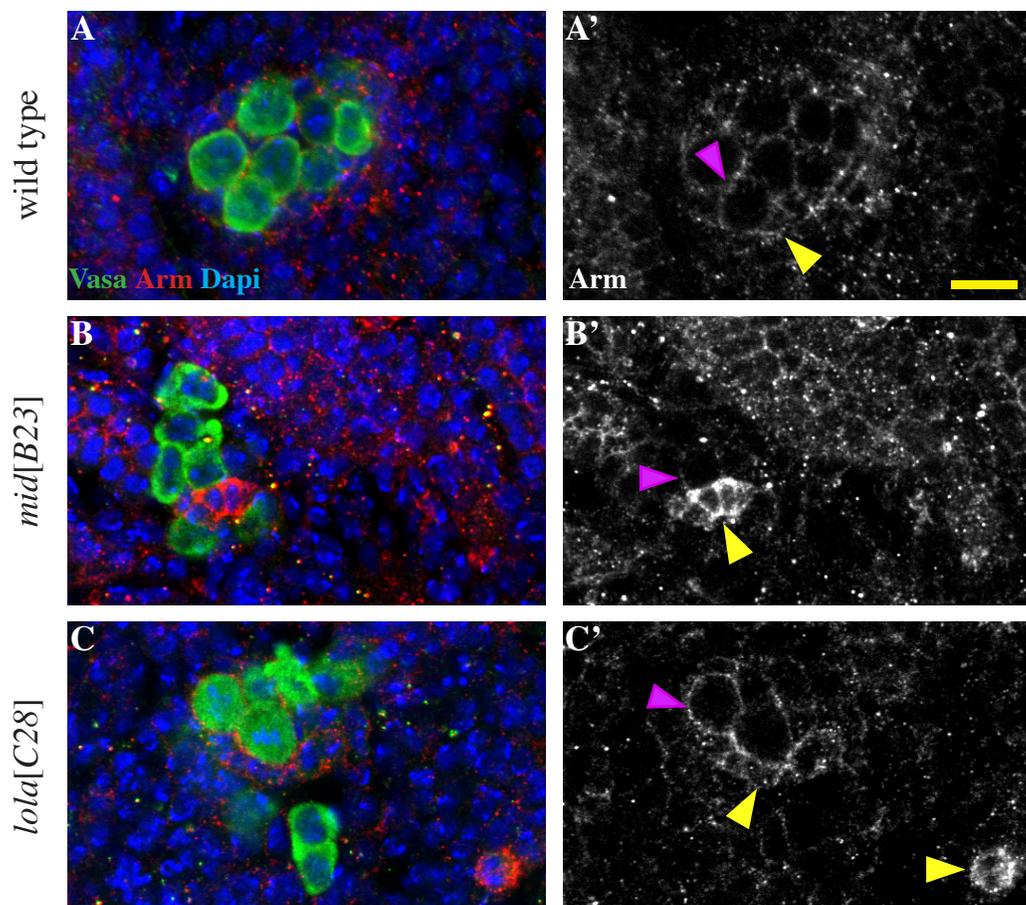


Figure 2.32: Armadillo levels in stage 15 *mid[B23]* and *lola[C28]* mutant gonads. In wild type (A, A') Arm (red, gray) expression is observed in both Vasa-labeled (green) germ cells (A' pink arrow head) and closely associated cells, SGPs (A' yellow arrow head). Similar to wild type, *mid[B23]* (B, B') and *lola[C28]* (C, C') mutant gonads also display germ cell (B', C' pink arrow head) and SGP (B', C' yellow arrow head) specific Arm expression. Scale Bar: 10 μ m

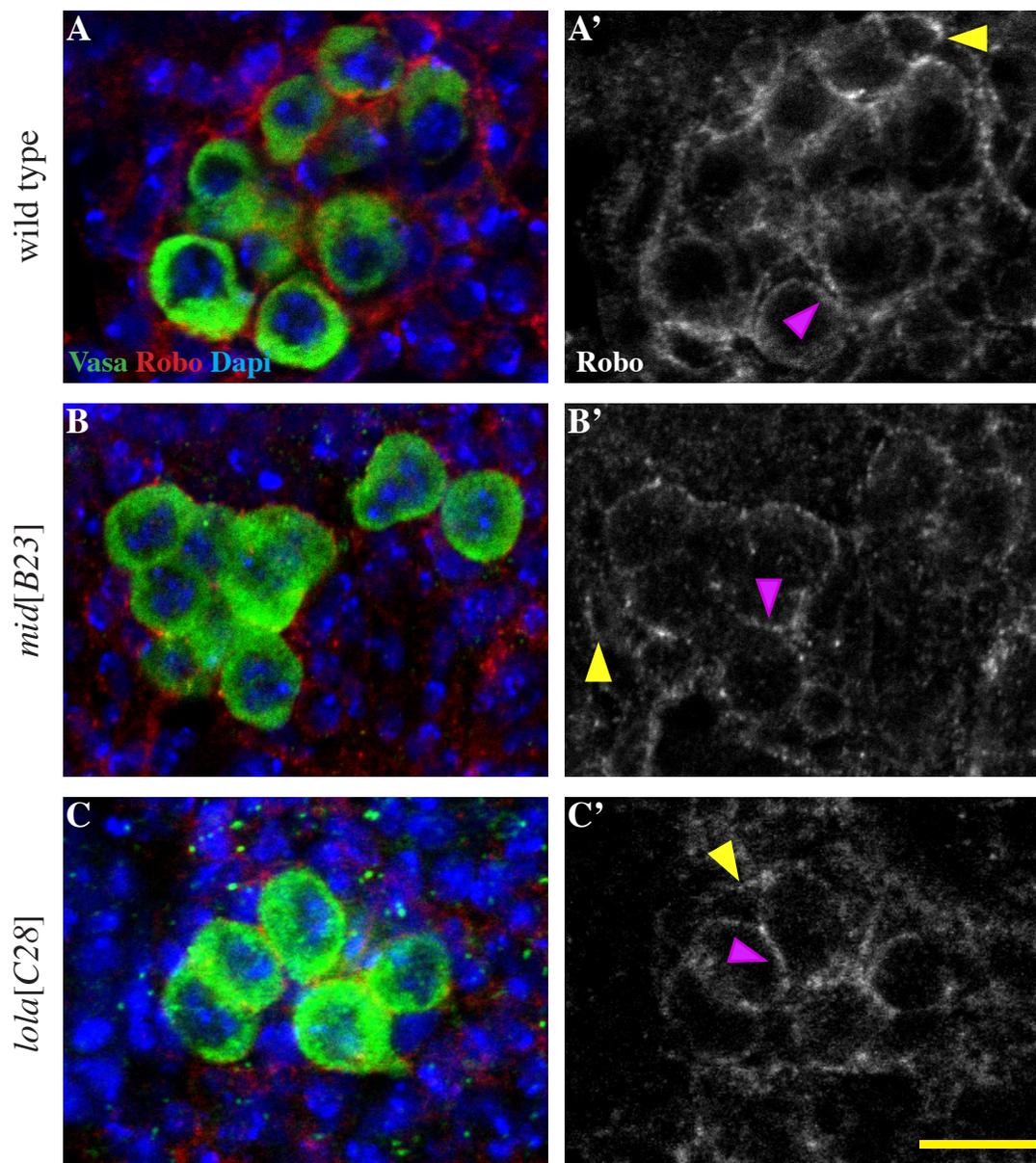


Figure 2.33: Robo levels in stage 15 *mid[B23]* and *lola[C28]* mutant gonads. In wild type (A, A') Robo (red, gray) expression is observed in both Vasa-labeled (green) germ cells (A' pink arrow head) and closely associated cells, SGPs (A' yellow arrow head). Similar to wild type, *mid[B23]* (B, B') and *lola[C28]* (C, C') mutant gonads also display germ cell (B', C' pink arrow head) and SGP (B', C' yellow arrow head) specific Robo expression. Scale Bar: 10 μ m

2.2.24 Tinman is expressed in SGPs but remains unaffected in *mid[B23]* mutant gonads

Tinman (Tin) is an early expressing mesodermal transcription factor required for proper patterning of many tissues derived from the mesoderm, including the heart, the muscles and the gonad (Azpiazu and Frasch, 1993; Bodmer, 1993). In case of gonad development, *tin* expression in the dorsolateral mesoderm is required for SGP specification (Boyle et al., 1997). In *tin* mutants there is a reduction in SGP number, manifestation of the germ cell phenotype is observed scattered in clumps in posterior of late stage embryos (Moore et al., 1998a). Moreover, it has been reported that although *tin* is required early for SGP specification, it is not expressed later in the SGPs (Jemc, 2011).

In the heart, there is a molecular connection between Mid and Tin: Tin expression in the heart is required to directly activate *mid* expression (Brook and Cohen, 1996; Ryu et al., 2011). Interestingly, Mid is later required to maintain Tin expression in the 4 out of the 6 cardioblasts in one hemisegment of the heart (Reim et al., 2005).

To test the possibility that Tin is downstream of Mid in the SGPs, wild type gonads were analyzed for Tin expression using a Tin antibody. Surprisingly, Tin expression could be detected in all the SGPs, counter-labeled with Tj, in stage 13 embryos (Fig. 2.34 A1, A1'). Interestingly, this expression became more restricted in a stage 15 gonad, where the anterior most SGPs showed the highest expression levels of Tin (Fig. 2.34 A1, A1' yellow arrow head) with a decreasing gradient of expression towards the posterior.

On analyzing the *mid[B23]* mutants, although SGP-specific Tj expression was lost in the stage 13 gonads, expression of Tin could still be observed (Fig. 2.34 B1, B1'). Similarly, in the mutant stage 15 gonad, Tin expression could be observed in cells making contact with the germ cells (Fig. 2.34 B2, B2'). However, no particular differences in the level of Tin expression could be seen in the stage 15 SGPs, indicating a possible absence of anterior and posterior SGP identity in the mutants.

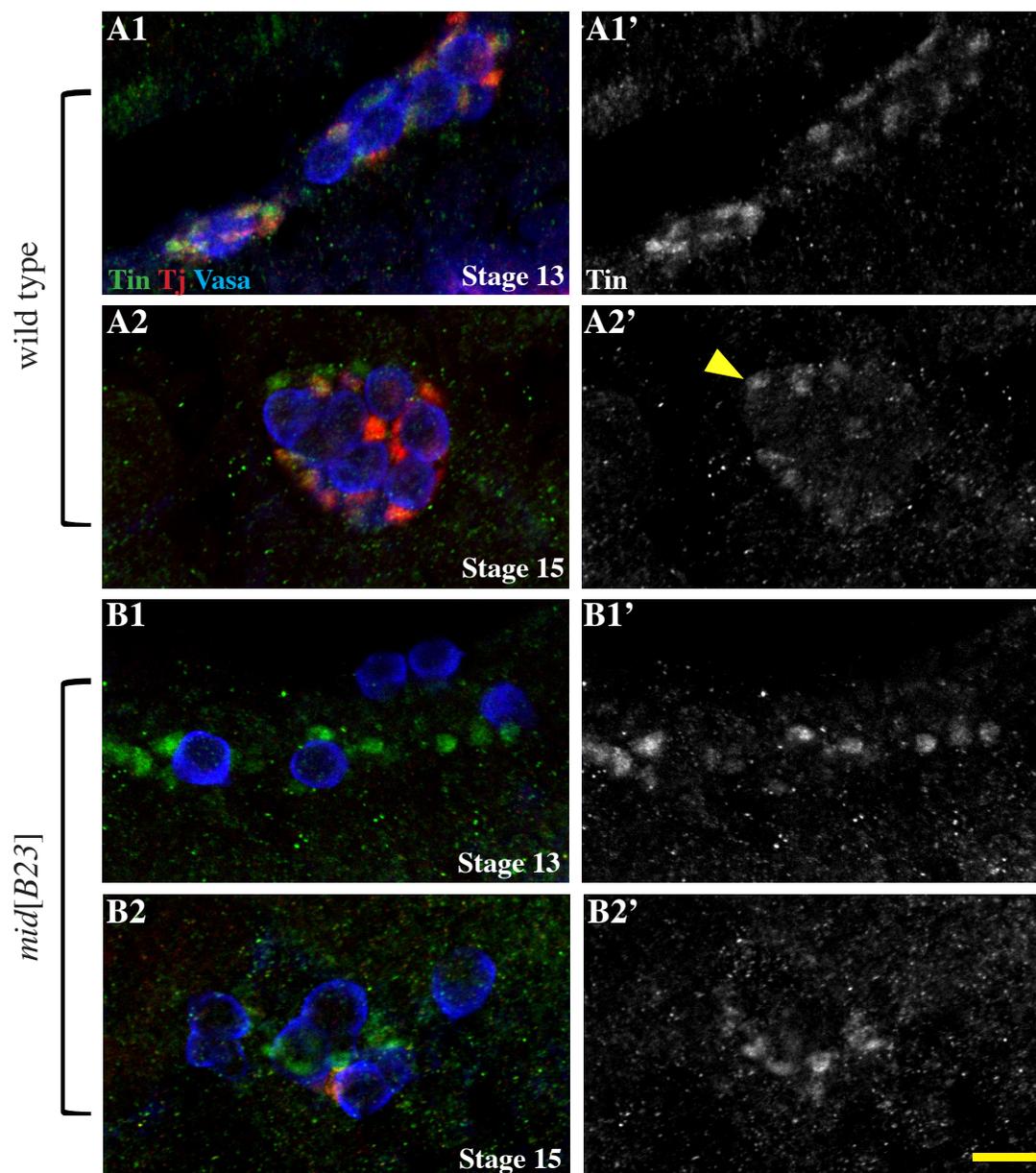


Figure 2.34: Tinman expression in wild type and *mid[B23]* gonads. In wild type stage 13 gonad (A1, A1'), Tin (red, gray) expression is observed to co-localize with all nuclear Tj-labeled (red) SGPs. At stage 15 (A2, A2'), Tin is observed in the SGPs in a gradient fashion with the anterior SGPs (A2' yellow arrow head) showing highest levels of expression. Similar to wild type, *mid[B23]* stage 13 gonad (B1, B1') and stage 15 gonad (B2, B2'), in spite of absent and reduced Tj expression respectively, still express Tin in the mutant gonads. Vasa-labeled germ cells are seen in blue. Scale Bar: 10 μ m

2.2.25 *mid*[B23] and *lola*[C28] genetically interact

The two mutants described in this study show similar defects in the gonad formation suggesting that some of the downstream genes may be regulated by both Mid and Lola. Although given that *mid*[B23] affects Tj expression, while *lola*[C28] does not, at least some of the downstream genes must be unique.

To check if *mid* acts downstream of *lola*, the *UASmid* transgene was expressed in the mesoderm of *lola* mutants. As expected, embryos expressing *UASmid* in the *lola*[22.05] 68-77 *twiGal4/+* background, showed normal, compact stage 15 gonads (Fig. 2.35, A, A'). Surprisingly however, in approximately 50% of stage 15 gonads of *lola*[22.05] 68-77 *twiGal4/lola*[C28] embryos expressing the *UASmid* transgene in the mesoderm, yet another phenotype was seen, observed in neither *lola*[C28] nor *mid*[B23] mutants. These gonads appeared 'super-elongated', containing increased number of germ cells and SGPs (Fig. 2.35, B, B'). Moreover instead of being positioned at PS10, these 'super-elongated' gonads appeared to span more than one parasegment. However, the SGPs appear to interact and ensheath the germ cells through the length of the 'super-elongated' gonad.

Given these findings, I wanted to test for a functional relationship between *lola* and *mid* by testing for a single copy genetic interaction between the mutant alleles (Fig. 2.36). As shown in previous quantifications, *lola*[C28] shows only 10% mutant gonads when heterozygous, while in heterozygous *mid*[B23], approximately 30% of the gonads are mutant. However, on testing embryos transheterozygous for the two mutants, there was a marked increase in mutant gonad phenotypes compared to the single mutant heterozygous situation. On crossing *lola*[C28] females with *mid*[B23] males, the resulting embryos showed a 40% penetrance in mutant phenotype was observed compared to 10% in *lola*[C28]/+ embryos. In the embryos from the reciprocal cross, using *mid*[23] females, an even higher penetrance of 70% mutant gonads was observed in comparison to the *mid*[23] females crossed to wild type males (*mid*[B23]/+).

To test if the interaction seen with *mid*[B23] is specific to the *lola* locus, or if its due to some other mutation in the *lola*[C28] background, *mid*[B23] was further tested for interaction using the *lola* deficiency Df(2R)BSC595 and the *lola* null allele *lola*[22.05]. In both scenarios, a 50-60% penetrance of mutant gonad phenotypes could be observed (Fig. 2.36), implying a *lola*-specific interaction.

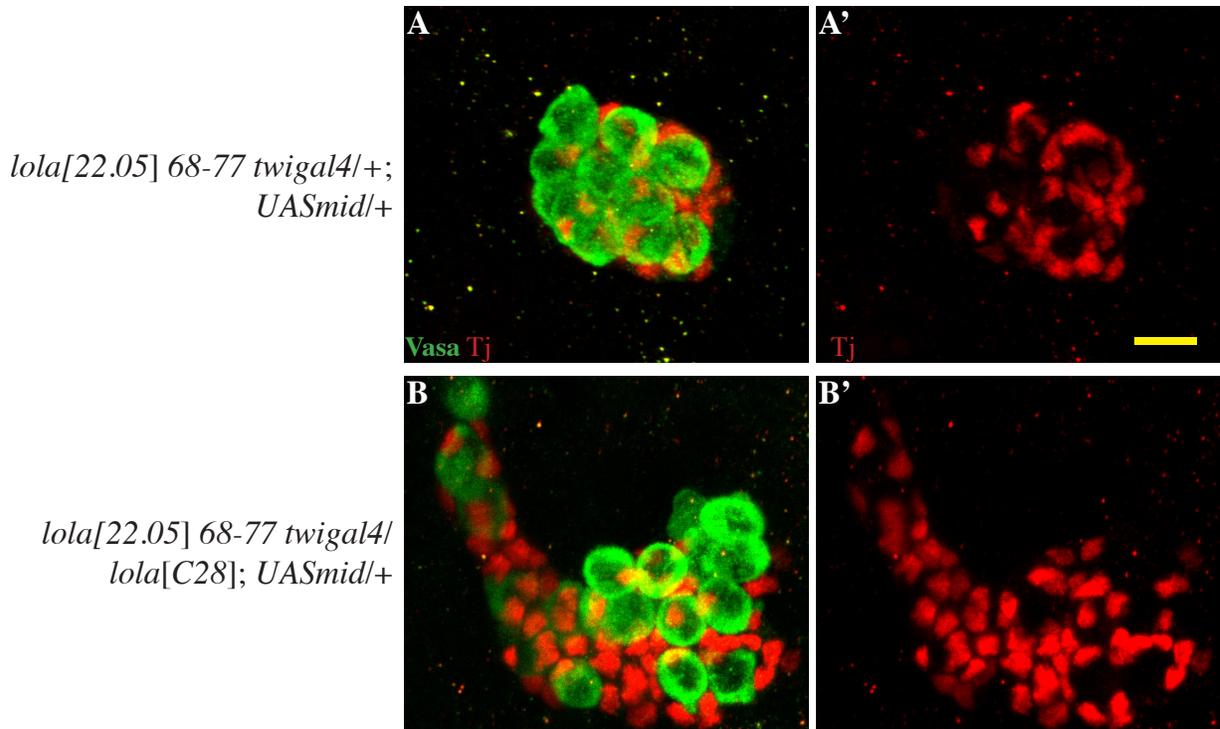


Figure 2.35: ‘Super-elongated’ gonads in *UASmid* ‘rescued’ *lola*[C28] mutants. While overexpression of *UASmid* in the mesoderm of *lola* heterozygous embryos (A, A’) leads to wild type gonad morphology and Tj expression (red), such overexpression in *lola* mutants (B, B’) leads to the formation of abnormal ‘super-elongated’ stage 15 gonads. Such a gonad typical spans more than one parasegment, and displays a greater number of germ cells and SGPs. Vasa-labeled germ cells are seen in green. Scale bar: 10µm

The strong genetic interaction observed between the two mutant alleles when one functional copy of each gene is still available appears to be synergistic, and indicates that the two genes, *midline* and *lola*, interact mechanistically at an unknown level. However, expression of Lola-R does not rescue *mid*[B23] mutant gonad phenotype (data not shown). Additionally, Mid over-expression in *lola*

mutants leads to the formation of the ‘super-elongated’ gonads. Therefore, the possibility of the two transcription factors acting in parallel pathways, with one or more overlapping targets could explain the observed interactions. Studying these pathways and their regulation could reveal many novel mechanisms involved in SGP behavior and the process of gonad morphogenesis.

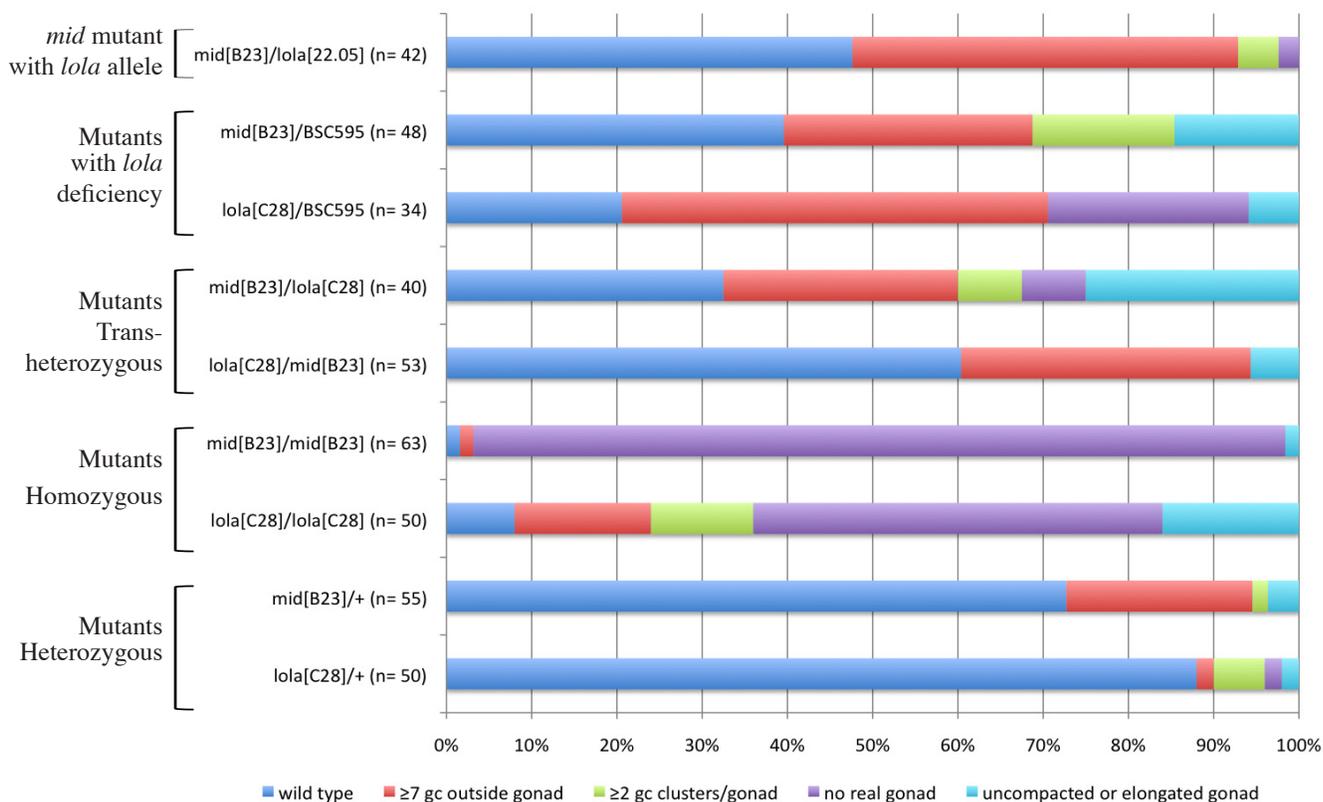


Figure 2.36: Quantification of the genetic interaction bet *midline* and *lola*. The graph quantifies the phenotype using the indicated categories (for explanation of the categorization, please see text, section 2.2.6). The bottom two bars show the penetrance in heterozygous *lola*[C28] and *mid*[B23] mutants, while the next two bars display the penetrance in homozygous mutants. The two bars above display the single copy genetic interaction between the two mutants resulting from reciprocal crosses. Following this, the interaction of *mid*[B23] with the *lola* deficiency *BSC595* (compared to *lola*[C28]/*BSC595* non-complementation), and *lola* null allele *lola*[22.05], showing approximately 60% and 50% interaction is displayed.

Chapter 3

Discussion

3.1 Requirement of cell surface activity of Wun2 for germ cell migration and survival

In this study, I focused on addressing one of the crucial assumptions on which the model of Wunen's role in germ cell migration and survival is based. I used two approaches to ascertain the site of Wun2 activity in germ cells.

In the first approach, I overexpressed tagged and untagged versions of Wun2 in the germ cells. The untagged and GFP tagged versions localized primarily to the plasma membrane whilst the Myc tagged version was more intracellular. All 3 versions showed identical functionality in rescue experiments. I conclude that the GFP version more accurately reflects the native protein and that the cell surface represents the site of accumulation of Wun2 in germ cells.

In a parallel approach, I synthesized a cleavable form of Wun2 by inserting a TEV protease cleavage site (tev-site) in the third loop of Wun2 between the first two catalytic domains. Cleavage at this site, by extracellular TEV protease, would separate the catalytic domains, and presumably render the enzyme inactive. Expression of the secretable TEV protease in the surrounding tissues had no effect on control Wun2 without a tev-site to rescue germ cell loss in *wun wun2* M-embryos, but significantly reduced the ability of Wun2tev to do so. This was confirmed by using several insertion lines of both Wun2con and Wun2tev, all of which showed the same pattern of germ cell survival with/without protease expression.

These results strongly implicate that Wun2 acts on the surface of the germ cells, where it dephosphorylates extracellular lipid phosphates, to aid germ cell migration and survival.

3.1.1 Differential localization of the tagged-versions of Wun2

Wunen localization has been studied previously in somatic cells. Localization

of the tagged forms of Wun, WunGFP and Wunmyc, is reported on the plasma membrane of both S2 cells (using electron microscopy and anti-GFP or anti-myc antibodies followed by detection using gold particles) (Burnett et al., 2004) and mesodermal cells (fluorescent staining) (Burnett and Howard, 2003). Although similar cell surface localization of Wun2myc is claimed in the ectodermal cells of *Drosophila* embryos (Starz-Gaiano et al., 2001), closer inspection of their figures indicate more intracellular localization. However, the localization of Wun and Wun2 in germ cells had not been looked into.

To examine where Wunens localize in migrating germ cells I used two tagged forms of Wun2, and compared them to overexpressed untagged Wun2. Interestingly, the localization of overexpressed untagged Wun2 (visualized by using an antibody against Wun2) was very similar to Wun2-GFP, both being largely cell surface, but highly dissimilar to the largely cytoplasmic presence of Wun2-myc (both visualized by using an antibody against the tags). Although these results, suggest that Wun2 localizes to the cell surface, several observations leave open the possibility that Wunens act inside the cell. Firstly intense cytoplasmic punctae of Wun2 were observed when untagged-Wun2 was overexpressed in germ cells. Secondly, from live imaging work performed earlier in the lab, fast moving cytoplasmic 'dots' of Wun-GFP were seen in migrating and stationary germ cells (J. Baumbach, Diploma thesis). Finally, in spite of a starkly different pattern of localization, Wun2-myc was equally capable of rescuing germ cell loss, as Wun2GFP, in maternal *wunen* lacking embryos.

One line of reasoning to explain the above observations could be that the tag affects the localization pattern of the protein. It is possible that the GFP tag on Wun2 induces and/or stabilizes the localization of the protein at the plasma membrane. Alternatively, the Myc tag could hinder its proper cell surface localization, by causing the intracellular trapping of the protein. This hypothesis is supported by the observation that WunGFP localized to the cell surface in insect S2 cell lines, while Wun with a 3Xmyc tag was observed in 'sub-cortical vesicles', along with cell surface localization (Burnett et al., 2004). Lastly it also possible, that the C-terminal Myc tag is cleaved off, and what we observe by Myc antibody staining in the cytoplasm, is

only the tag.

Alternatively, the cytoplasmic presence of Wun2 and Wun2-myc could be explained by the inability of all the overexpressed protein to localize to the plasma membrane. Furthermore, the cytoplasmic punctae could be explained as the recycling or the transport of the protein between the interior of the cell and the cell surface.

Taken together, I conclude that Wunens localize to the cell surface, however I cannot exclude they are functionally required also in some intracellular site.

3.1.2 Cleavage of Wun2 at the inserted tev-site

The insertion of tev-sites for cleavage of intracellular proteins has been performed previously in *Drosophila* cells (Harder et al., 2008; Pauli et al., 2008). A 3x-tev-site was inserted in the Rad21 protein, separating the N- and C-terminus, to investigate the role of Cohesin in post-mitotic embryonic cells. The cleavage of Rad21 on heat shock-induced TEV protease expression in the same cells could be observed in a Western blot (Pauli et al., 2008). In a similar study, co-expression of the TEV protease with a cleavable form of cytoplasmic Tao protein led to the nuclear accumulation of the cleaved C-terminus. However, in contrast to the previously mentioned study, only one tev-site was inserted in the Tao protein (Harder et al., 2008).

A single cleavage site was used in Wun2 as it was already shown to be sufficient for recognition by the TEV protease, and additionally to avoid altering the protein structure more than necessary. Moreover, unlike the earlier works, this study expressed the TEV protease and the cleavable Wun2tev substrate in separate cell types, increasing the complexity of the experiment.

In spite of using several experimental set ups, I was unable to demonstrate the cleavage of Wun2tev *in vitro* by the TEV protease. I used Schneider2 (S2) insect cell lines, to express Wun2con and Wun2tev, and attempted TEV cleavage using exogenous addition of TEV protease to intact cells and total protein extracts. However, several bands were detected on Western blots using an anti-Myc antibody

even without protease addition, occluding the predicted cleavage band. This large number of observed bands likely represents several degradation products of the overexpressed Wun2tev protein in the S2 cells.

However, the experiment in this study demonstrating the significant reduction in the ability of Wun2tev to rescue germ cell loss in the presence of the protease can only be explained by recognition of the tev-site by the protease.

3.1.3 Implication of cell surface Wun2 activity in germ cells

Wild type germ cells can sense a lipid phosphate gradient up to a distance of 33 μ m (Mukherjee, Neher et al. unpublished data). However, it is not known how germ cells sense this extracellular lipid phosphate gradient and whether the sensing is directly or indirectly performed by the expression of Wunens in germ cells. Furthermore, overexpressing Wunens in insect cell lines results in the accumulation of fluorescent lipid analogs within the cell (Renault et al., 2004a) but the mechanism of entry of the lipid is still unknown.

It is possible that the extracellular lipid phosphates are received by a cell surface receptor (possibly the G-protein coupled receptor Tre-1), which is internalized and the substrates are subsequently dephosphorylated by Wunens, hence leading to indirect sensing. However, to date, no lipid-specific receptors have been found in *Drosophila*, which makes the direct sensing by Wunen expression in germ cells the more favorable model of Wunen activity (Renault et al., 2004a). For this it is important that Wunen can act as an 'ecto-enzyme' in order to directly dephosphorylate the extracellular lipid phosphates and cause their subsequent uptake.

This work strongly suggests the role of Wun2 as an 'ecto-enzyme' in the germ cells, hence supporting the model of Wunens in germ cells acting as a direct sensor of the extracellular lipid cue. Germ cells could possibly sense the rate of Wunen activity by detecting the rate of accumulation of the dephosphorylated form of the lipid inside the cell. Hence germ cells close to an area of somatic wunen expressing

cells, where lipid phosphate levels are lower, would migrate away from these regions because of decreased internalization of the lipid.

The redundancy of Wun and Wun2 in germ cell migration and survival is known (Starz-Gaiano et al., 2001). Wunens are also required in other morphogenetic processes like tracheal formation and heart formation. Interestingly, defects in trachea can occur by the loss of only Wun, however, both Wun2myc and Wun2GFP can rescue these defects (Kispert and Herrmann, 1993). Apart from proving a redundancy in the required activity of Wun and Wun2, it also suggests a cell surface activity of Wun similar to Wun2.

3.2 Transcription factors regulating several steps of gonad morphogenesis

In this study, I phenotypically characterized three gonad formation mutants *A44*, *B23*, *C28* obtained from an EMS mutagenesis screen. The mutants displayed normal early germ cell migration, with the migration through the midgut into the mesoderm and the bilateral sorting of the germ cells being unaffected. Analyzing the stage 13 and stage 15 mutant gonads revealed defects in several steps of gonad formation, including SGP-cluster fusion, germ cell ensheathment by the SGPs, and gonad compaction.

I used two independent approaches, whole genome sequencing and complementation by deficiencies, to clone the causative mutations. I established *B23* as an allele of *midline* and *C28* as an allele of *longitudinals lacking*. The observed gonadal defects could be rescued by expression of the respective transgenes in the mesoderm of *mid*[*B23*] and the *lola*[*C28*] mutant embryos.

Both *mid* and *lola* encode transcriptional factors, and were seen to be expressed in the SGPs. The missense mutation in *lola*[*C28*] affects only one isoform

of *lola*, *lola-R*. Another isoform of *lola*, *lola-BC*, could not be detected in the gonad, nor could expression of this isoform rescue the *lola*[C28] mutant gonadal phenotype.

Expression of the adhesion proteins DE-cadherin and Armadillo, and the cell surface receptor Robo was not affected in the mutants. However, expression of another transcription factor Traffic Jam, was lost in stage 13 and patchy in stage 15 *mid*[B23] mutant SGPs. Moreover, Tj levels were normal in the other two gonad formation mutants, *A44* and *lola*[C28]. Additionally, I found that Tinman, earlier thought to be required only for SGP specification, is also expressed in the SGPs at later stages. This expression was however, unperturbed in *mid*[B23] mutants.

Lastly, I observed and quantified a single copy genetic interaction between *mid*[B23], *lola*[C28] and other *lola* deficiencies and null alleles. The presence of an interaction between the two genes was further demonstrated by the observed 'super-elongated' gonad phenotype on expression of Midline in the mesoderm of *lola*[C28] mutants.

Therefore, this study has revealed two new factors essential to the process of gonad morphogenesis and required in several stages of this process. The two implicated genes *mid* and *lola*, both encode for transcription factors that could regulate gonadal expression of overlapping or unique sets of effector genes.

3.2.1 Normal mesodermal patterning and SGP defects in mutants

The specification of SGPs occurs in the dorsolateral mesoderm (Brookman et al., 1992b). Therefore, normal mesoderm specification and its proper patterning is a prerequisite for normal SGP formation. In the gonad formation mutants characterized in this study, three lines of evidence pointed towards a correctly patterned mesoderm: i) Normal germ cell migration, which requires normal mesodermal patterning (Warrior, 1994), was observed. ii) Expression of early SGP markers such as the *412* retrotransposon, and the transcription factor Eyes absent (*Eya*), could be detected. Additionally, these markers also established that the specified SGPs were capable of initiating contact formation with germ cells.

iii) Lastly, another marker FasIII, labeling the visceral mesoderm, demonstrated a normal fused and contiguous mesodermally derived tissue in the mutants. These lines of evidence show that the compromised SGP cluster fusion occurring in stage 13 mutants is specifically due to SGP-related defects and not faulty mesoderm patterning.

Apart from unfused and separate SGP-germ cell clusters, germ cell ensheathment by the cytoplasmic processes extended from the SGPs in stage 13 mutants was lacking. This loss of ensheathment was more penetrant in *A44* and *mid[B23]* compared to *lola[C28]*, where the larger SGP-germ cell cluster often appeared ensheathed.

The few gonads that succeeded in SGP-cluster fusion, however, failed in the following steps of gonadal compaction. These uncompact/elongated gonads appeared similar to wild type stage 13 gonads. Along with the compromised SGP-cluster fusion and germ cell ensheathment, several germ cells in the mutants were also found scattered at the posterior of the embryo. Such a scattering could result from the loss of the initially established SGP-germ cell contact, possibly also due to lack of ensheathment of the germ cells. Live imaging of the behavior of the two cell-types with respect to each other could tell us more about the dynamics of their interaction in the mutant situation.

3.2.2 Severity of the EMS induced mutation in *mid[B23]* and *lola[C28]*

The conserved nature of the affected locus in the two mutants was demonstrated by the performed multiple sequence alignment using homologues in other species. Due to a mutation in the splice site acceptor residue, The Mid protein in the *mid[B23]* mutants lacks 10 amino acids in the conserved T-box domain. Using an antibody synthesized against C-terminus of the protein, Mid expression could be observed in nucleus of the heart cells in *mid[B23]* mutants (data not shown), indicating normal translation and localization of the mutant protein. The axonal tracts, however, demonstrated a non-functional Mid protein in *mid[B23]* where the observed VNC phenotype was similar in severity to that of the published protein

null *mid* allele, *mid*[1] (Liu et al., 2009). Furthermore, as reported with the *mid*[1] allele, the hearts of the *mid*[B23] mutants showed abnormal Tin expression (data not shown). These data together demonstrate that *mid*[B23] is also a functional null allele.

The Asn835Ile mutation in *lola*[C28] on the other hand, unlike the null *lola*[22.05] allele, did not lead to an axonal tract phenotype. The study of the crystal structure of Tramtrack protein, which is one of the founding members of BTB domain containing proteins, provides evidence of the possible effect of the missense mutation in *lola*[C28]. It demonstrated the importance of the Asn residue in question, to aid the Zn-fingers in making contact with the target DNA (Fairall et al., 1993). Therefore, the Asn835Ile mutation, affecting the *lola*-R isoform, could potentially prevent its binding to the target DNA in the gonad. Therefore it is likely that *lola*[C28] would result in a non-functional *lola*-R isoform. Furthermore, the lack of VNC phenotype in this mutant could be explained by the expression of the other *Lola* isoforms in this tissue.

3.2.3 Severity of phenotype in *lola*[C28] homozygote versus transheterozygote

The quantification of the phenotypes in *lola*[C28] transheterozygous with *lola* null alleles always indicated a decrease in the severity of the gonad defects compared to the *lola*[C28] homozygous embryos, as determined by germ cell-based scoring.

Hence the possibility of a second locus contribution to the *lola*[C28] phenotype was looked into. Several deficiency lines uncovering the SNPs identified on chromosome 2R were tested, but none of them showed non-complementation with *lola*[C28] by phenotype. However, the EMS mutagenesis induced possible second locus could potentially lie on the 2L arm, or affect a non-coding region anywhere in the second chromosome.

Therefore, to establish if the mutated *lola*-R locus is singly responsible for the gonad phenotype observed in *lola*[C28] and to allow the rescue of the 'severe'

phenotype of *lola*[C28] with Lola-R isoform, recombinants were made between the FRT (42B) containing *lola*[C28] (47A) chromosome and the *twiGal4* (55D) chromosome. The lines were defined as recombinant if they contained either i) both the mutant *lola* locus and the *twiGal4* insertion, or ii) the FRT locus and the *twiGal4* insertion, or lastly iii) none of the markers. Of the 21 were recombinants isolated, 8 were lethal with *lola*[22.05], hence containing the mutant *lola* locus, and were also positive for the *twiGal4* driver. All of the 8 *lola twiGal4* recombinant lines showed a phenotype with *lola*[C28], while the remaining 13 recombinants that were not lethal with *lola*[22.05], did not. Two lines *lola*[C28] *twiGal4* #14 and *lola*[C28] *twiGal4* #39 were further quantified both for severity and rescue. Both recombinant lines in trans with the original *lola*[C28] chromosome, when compared to the *lola* null alleles, showed an increased penetrance in the more severe categories of ‘≥2 germ cell clusters/gonad’ and ‘no real gonad’. The observed penetrance in the transheterozygotes was very similar to the *lola*[C28] homozygotes and higher than the *lola*[C28]/*lola*[null] embryos. Moreover, expression of Lola-R isoform in these transheterozygotes, fully rescued all the mutant phenotypes.

These recombinants demonstrate that the region between the FRT and *twiGal4* insertion, namely, 42B-55D, does not contain a second locus responsible for the *lola*[C28] phenotype. Although the 2L chromosome did not contain any markers for recombination, it is quite likely that in the recombinants that did not show a phenotype, some recombination events would have occurred on the left arm. Therefore, a contribution from the left arm is unlikely. Putting all these lines of evidence together, strongly indicates that the severity of gonad defects observed in the *lola*[C28] can be attributed solely to the mutation at the *lola*-R locus.

3.2.4 Identification of an isoform specific role for Lola

Recent work has implicated the role of different Lola isoforms in the germline of both male and female flies. Lola-I (aka Lola-K) is required in the ovaries for the program cell death (PCD) of late stage nurse cells. Germ line clones of *lola*[629], containing a lesion in the I-specific exon, exhibit defective nuclear

organization, chromatin condensation and DNA fragmentation required for PCD in the nurse cells (Paige Bass et al., 2007). Furthermore, using available isoform-specific mutants and inducible RNAi in the adult testis, isoforms Lola-I and Lola-BC (Lola-L) were implicated in cell autonomous germline stem cell (GSC) maintenance and survival (Davies et al., 2013). Lack of either isoform led to differentiation and subsequent loss of GSC fate. Also, as mentioned before, *lola[22.05]*, a null allele for all *lola* isoforms, was identified in a screen, along with several other genes, to be responsible for gonadal defects (Weyers et al., 2011).

This study adds to the limited knowledge of the isoform specific functions of Lola in embryogenesis. Here another isoform-specific allele of Lola, Lola-R (Lola-T/Lola-4.7) is isolated, and its role in a stage prior to germ cell differentiation, i.e., gonad formation, and in the somatic cells of the gonad, is described.

3.2.4 Stage specific loss of Traffic Jam expression in *mid[B23]* mutant gonad

Expression of Traffic Jam, which is a MAF family transcription factor, is first seen in late stage 12 or early stage 13 SGPs. Mutants of *tj* show defective germ cell ensheathment, where the SGPs peripherally surround the germ cells, but fail to send out cytoplasmic projects to individualize them (Li et al., 2003). In *mid[B23]*, the larger SGP-germ cell clusters appear very similar to the *tj* mutant gonads, consistent with the observed lack of stage 13 SGP Tj expression. Interestingly, although a few SGPs regain the ability of Tj expression in later stages, ensheathment defects are not restored in the mutants. Therefore, it appears that firstly Tj is required in early stages for correct ensheathment, and secondly there is a Mid-dependant regulation of Tj expression at stage 13, but other factors also contribute to its expression at later stages. It would be interesting to study if Mid directly binds the Tj promoter, and what other transcription factors regulate Tj expression during development.

3.2.5 Interaction between *midline* and *lola*

Since Lola and Midline are both transcription factors, they could potentially regulate the same downstream targets. The attempted rescue of *mid*[B23] mutant phenotype with expression of *lola*-R-GFP did not function to improve or worsen the condition of the gonads (data not shown). This indicates that *lola* is not downstream of *mid* in the gonad. However, the appearance of a neomorphic phenotype in the reciprocal rescue attempt argues for a genetic interaction between the two genes. Overexpression of Mid in a wild type embryo, however, does not affect the gonads. This could be due to presence of a functional Lola protein in the wild type, which downregulates Mid activity.

Interestingly, a similar 'super-elongated' gonad phenotype has been observed on overexpression of *abdominal A* (*abdA*), under the *heat shock* (*hs*) promoter, during early embryonic expression (Boyle and DiNardo, 1995). Although *abdA* is required for SGP specification, while *mid* and *lola* are not, it would be intriguing to see if the three genes interact to cause such an unusual phenotype in later stages.

Furthermore, the synergistic interaction observed with the loss of just one copy of both genes also indicates a functional relationship between the two genes. This observed synergy, could also indicate the convergence of the two separate pathways involving Lola and Mid at a node. Double mutants of *mid* and *lola* could be investigated in the future to check for increased penetrance and/or severity of mutant gonad phenotypes.

3.2.6 Genes unaffected in gonads of *mid*[B23] and *lola*[C28]

The levels of three proteins with roles in gonad formation, DE-cad, Arm, and Robo, were tested in *mid*[B23] and *lola*[C28], using the appropriate antibodies. Furthermore, other genes implicated to function downstream of either *lola* or *mid* in other tissues, were also tested for expression/interaction.

➤ **Adhesion proteins**

Mutation in several genes, including *raw*, *enabled* and *fear of intimacy*, that show defects in gonad formation similar to that seen in *lola* and *mid* mutants, display defective expression or localization of adhesion proteins (Van Doren et al., 2003; Mathews et al., 2006; Sano et al., 2012). However, unlike in these mutants, expression of the adhesion proteins, DE-cad and Arm, could be detected in *mid[B23]* and *lola[C28]*, and moreover, they localized properly to the cell junctions in the mutant gonads. The fact that *mid[B23]* and *lola[C28]* do display gonadal defects without detectable changes in DE-cad or Arm levels/localization suggests that disruption in other downstream targets can lead to similar gonad phenotypes.

➤ **Robo**

Similar to the adhesion proteins, the expression of Robo, a cell surface receptor, also appeared to be normal in the mutant gonads. Previous reports stated that a reduction of Robo levels was observed in the CNS of *mid* mutants (Liu et al., 2009). They also identified a consensus binding site of the Mid transcription factor, and demonstrated Mid binding in the promoter region of Robo. However, I found no observable differences in CNS Robo levels in *mid[B23]* or *mid[1]* mutants (data not shown). Moreover, a recently published study demonstrated the previously reported binding site to be incorrect and showed a novel consensus site (Najand et al., 2012). This puts into question the published effect of Mid on Robo levels, although it remains clear that although Robo is required for gonad formation, the gonads of the *mid[B23]* and *lola[C28]* mutants have normal gonad Robo expression.

➤ **Frazzled**

Another receptor, Frazzled (Fra), is reported to be governed downstream of both Lola (Gates et al., 2011) and Mid (Liu et al., 2009) in the CNS. Although Fra expression has never been reported in the gonads, I performed several experiments to investigate a possible connection between *fra* and the two transcription factors in the gonad. Firstly, embryos lacking *fra* were analyzed for a gonad phenotype. Vasa-labeled germ cells in homozygous *fra* mutants displayed a slight dispersed and/or clumping phenotype, however, with a low penetrance (data not shown). Further, attempting a single copy genetic interaction with both *lola* and *mid* did not yield a

significant difference in perturbation of gonad formation. However, an effect would be expected only if the levels of available protein were crucial. Lastly, a *UASfra* transgene driven by the *twiGal4* driver was unable to rescue the *mid[B23]* and *lola[C28]* gonad phenotypes. However, the lack of rescue by the *UASfra* transgene could also be explained by the requirement of expression of other genes in the gonad of these mutants. Therefore, although these experiments failed to provide a link, the possibility that *fra* is downstream of either *mid* and/or *lola* can not be completely ruled out.

➤ **Spire**

Lola negatively regulates *spir*, which encodes an actin nucleation factor, required for proper axon guidance (Gates et al., 2011). To test for a role of Spire in gonad formation, both *spir* mutants as well as *spir* overexpression embryos were analyzed. However, neither the lack nor the overexpression of *spir* affected gonad morphology (data not shown).

➤ **Tinman**

In spite of the earlier belief that Tinman expression is required only in the mesoderm for SGP specification, and not expressed by the SGPs themselves (Azpiazu and Frasch, 1993), this work shows that late stage SGPs do express Tin. Its expression becomes restricted to the anterior SGPs upon gonadal coalescence (Fig. 2.34). This scenario is similar to the heart, where, early Tin expression is required for cardioblast specification, but late stage Tin expression is observed only in a subset of the cardioblasts (Bodmer, 1993). This late stage and restrictive maintenance of Tin expression is performed by Mid protein in the heart cells (Brook and Cohen, 1996; Reim et al., 2005). In *mid* mutants, Tin expression is lost in the cardioblasts, and the expression domain of another transcription factor, Drosocross, is expanded (Reim et al., 2005). However, the unaffected expression of Tin in the SGPs of *mid[B23]* and *mid[1]* (data not shown), indicates a Mid-independent mechanism of maintenance of Tin expression.

3.2.7 Other possible players in the gonad acting downstream of Mid and Lola

Besides the genes mentioned so far, only a few other pathways are known to interact with Midline and/or Lola. One such interactor is the Notch (N) receptor and its ligand Delta (Dl). N-Dl signaling is required to restrict early Mid expression to two cardioblasts per hemisegment during heart development (Brook and Cohen, 1996). In the eye, Lola functions to attenuate the N-Dl induced gene expression in some photoreceptors to allow one developmental fate over the other (Zheng and Carthew, 2008). Moreover, Notch signaling in the SGPs is shown to induce the formation of hub cells after gonadal coalescence in the male embryonic gonads. Therefore, a similar hierarchical interplay between N-Dl, Lola, and Mid is a plausible mechanism of regulating expression of the two transcription factors and/or their target genes in the process of gonad formation.

Another pathway that could be involved is JNK (Jun amino (N)-terminal Kinase) signaling. Raw negatively regulates JNK signaling in the gonad (Jemc et al., 2012). *raw* mutants display a germ cell ensheathment defect also observed in *mid* and *lola* mutants. It is therefore possible that JNK signaling is required downstream of the transcription factors to regulate germ cell ensheathment or other phenotypes observed in the mutants. An enhancer trap reporter, puckered-lacZ, could be used to investigate if the levels of JNK signaling in the *mid* and *lola* mutant backgrounds is affected.

Recently published data, implicates an actin regulator, Enabled (Ena), in gonad formation. *ena* mutants display several phenotypes similar to that observed in the *mid*[B23] and *lola*[C28], including unfused SGP clusters, elongated and split gonads (Sano et al., 2012). Therefore, a commercially available Ena antibody could be used to check if Ena is downstream of either Mid or Lola.

In order to get a more global picture of the genes downstream of Mid and *lola*, and to find novel gonad specific targets, transcriptional profiling of the SGPs in the two mutants could prove fruitful. This could be performed by FACS (fluorescent activated cell sorting) mediated isolation of labeled SGPs and the subsequent

sequencing of their isolated RNAs. The difficulty of this approach is the lack of an available driver that expresses exclusively in the SGPs. *Dsix4Gal4*, often used as an SGP driver, also expresses early on in several other tissues in the mesoderm. Only by late stages of embryogenesis, are the highest expression levels seen in the SGPs. While it is possible to use this driver for the FACS of SGPs of late stages, i) the presence of non-SGP cells would be difficult to exclude, and ii) transcriptional profiling of late stage SGPs might not be ideal because the phenotypes manifest earlier in development.

An approach to circumvent these challenges could be to perform transcriptional profiling of FACS isolated gonads (Shigenobu et al., 2006), using labeled germ cells with reduced trypsinization and homogenization, enabling whole gonad isolation. Although this would cause the merging of germ cell and SGP profiles, comparison to germ cell only datasets could aid their dissection. Obtaining the mutant gonads, however, could be a challenge due to gonad fragmentation, but could be circumvented by scaling up the collections compared to wild type.

Alternatively, an enhancer/suppressor EMS or deficiency screen in the mutant backgrounds could be undertaken to look for possible interactors. Although the severity *mid[B23]* gonad phenotype might be too high for screening for enhancers, looking for suppressors in this background could still be plausible. The *lola[C28]* phenotype on the hand would be equally amenable for screening of both enhancement and suppression.

3.2.8 Implication on vertebrate gonad morphogenesis

Homologues of Mid and Lola in other organisms are required in several important developmental processes. The Mid homologues, hrT in zebrafish and Tbx20 in mouse, are implicated in heart formation (Griffin et al., 2000; Stennard et al., 2003). Furthermore, other T-box proteins are also known for their involvement in mesodermal specification, differentiation and migration in zebrafish, mouse, and frogs (Brook et al., 1996; Brookes et al., 1996; Russ et al., 2000). Moreover, BTB containing Zn-finger transcription factors, similar to Lola, are implicated in

vertebrate embryonic development and hematopoiesis (Lecuit et al., 1996; Shaknovich et al., 1998).

Studies on the process of gonad formation in other model organisms are very limited. SGP expression of homing ligands, like SDF1A in zebrafish (Bernstein et al., 1996; Knaut et al., 2003) and SDF1 in mouse (Ara et al., 2003), is required to direct germ cell migration. Additionally, the vertebrate Hox gene, *Hox8*, homologue of *abda* and *abdB* *Drosophila* *hox* genes, are known to express in the gonads of humans (Redline et al., 1992).

Hmgcr, the rate-limiting enzyme of cholesterol biosynthesis, was first shown to be required in the *Drosophila* SGPs to attract germ cells (Van Doren et al., 1998a). Following this, studies using inhibitors of HMGCR have also shown the role of this enzyme in primordial germ cell migration in zebrafish (Thorpe et al., 2004). Deficient prenylation of the G γ subunit of many G-proteins lead to disrupted GPCR signaling, causing the zebrafish germ cells to mismigrate. Some of these G γ subunits (G γ 2 and G γ 3) are expressed in tissues where the germ cells migrate (Mulligan et al., 2010). Furthermore, migration of germ cells in mouse is also dependent on HMGCR expression. Recent studies have shown that the genital ridges (developing gonads), which are populated by the germ cells, have accumulated cholesterol. Reducing cholesterol levels by inhibiting Hmgcr leads to germ cell migration defects (Ding et al., 2008).

These studies taken together demonstrate that genes shown to play a role in *Drosophila* gonad formation can be extrapolated to other vertebrate systems. Furthermore, this process, involving only two cell types, proves to be a good model system for studying changes in cellular behaviors and cell-cell interactions, required for organogenesis. Therefore, investigating the role Midline and Lola could provide useful insights into the various mechanisms that govern *Drosophila* gonad development.

Chapter 4

Materials and Methods

4.1 Materials

4.1.1 Fly husbandry

Flies were cultured in vials filled with approximately 10 ml standard cornmeal medium. Large- scale breeding can be performed in mid-size and large vials filled with approximately 25 ml and 50 ml medium, respectively. Stocks were kept at room temperature (RT). Crosses were performed at 25°C. For collection of embryos 30 to 40 flies were put into a laying cage covered with a yeasted apple juice plate at 25°C overnight.

4.1.2 Apple juice plate recipe

4 g agar was added to 740 ml bidest water and autoclaved. Subsequently 25 g sucrose diluted in 250 ml apple juice and 10 ml Nipagin were added.

4.1.3 *Drosophila* strains

The fly lines used in this study are listed below. The mutants which were homozygous lethal, were maintained over the *Cyo P{ry+ ftz>lacZ}* or *TM3 Sb p{ry+ ftz>lacZ}* second and third balancer chromosomes, respectively. These balancers expressed the β -galactosidase (*lacZ*) in a 7 ectodermal stipes under the *fushitarazu* (*ftz*) promoter. The homozygous null embryos could then be recognized as lacking the *lacZ* stripes.

white (*w*) flies are used as a wild type control in this study.

Germ line clones of *wun wun2* (to generate *wun wun2* M⁻ embryos) were made using the FRT42B *wun*[49] *wun2*[EP2650ex34] chromosome and the ovo^D dominant female sterile technique (Chou and Perrimon, 1996). For germ cell specific rescue experiments with the germ line clones, the females also contained a *nosGal4* driver.

Table 4.1: List of the mutants or deficiencies used in this study, and their source/reference. The BL number indicates the Bloomington stock number.

Genotype	Source/Reference
<i>white-</i>	Bloomington stock center
<i>wun2[Δ]</i>	(Hanyu-Nakamura et al., 2004)
<i>w ; FRTw+42B wun[49]</i> <i>wun2[EP2650ex34]</i>	(Renault et al., 2004a)
<i>w P{w+ faf>lacZ} ; FRTw+42B</i> <i>A44</i>	(Leal et al., 2009)
<i>w P{w+ faf>lacZ} ; FRTw+42B</i> <i>B23</i>	(Leal et al., 2009)
<i>w P{w+ faf>lacZ} ; FRTw+42B</i> <i>C28</i>	(Leal et al., 2009)
<i>Df(2R)ED2098</i>	BL 9277
<i>Df(2R)BSC336</i>	BL 24360
<i>w[*]; lola[e76]</i>	BL 28283/(Madden et al., 1999)
<i>mid[1] cn[1] bw[1] sp[1]</i>	BL 3086/(Nuesslein-Volhard et al., 1984; Liu et al., 2009)
<i>mid[2] cn[1] bw[1] sp[1]</i>	BL 18/(Nuesslein-Volhard et al., 1984; Liu et al., 2009)
<i>w; al lola[22.05] 68-77 sp</i>	Gift of Mark VanDoren/(Weyers et al., 2011)
<i>w[1118]; PBac{ lola.GR-GFP.FLAG}VK00033</i>	BL 38661
<i>P{FRT(w[hs])G13 fra[3]</i>	BL 8813/(Kolodziej et al., 1996)
<i>spir[1] cn[1] bw[1]</i>	BL 5113/(Kitadate and Kobayashi, 2010)

Table 4.2: List of transgenic lines used for expression in the embryos, and their source/reference. The BL number indicates the Bloomington stock number.

Genotype	Source/Reference
<i>nosGal4</i>	(Crowner et al., 2002)
<i>twiGal4</i>	Bloomington stock center
<i>UASGFP</i>	Bloomington stock center
<i>UASwun2myc</i>	(Starz-Gaiano et al., 2001)
<i>wun2[EP2650]</i>	(Yin et al., 1997)
<i>UASwun2GFP</i>	Made in this study
<i>UASwgmcherryTEV</i>	Gift of Reinhard Schuh
<i>UASwun2conmyc</i>	Made in this study
<i>UASwun2tevmyc</i>	Made in this study
<i>nos>wun2conmyc</i>	Made in this study
<i>nos>wun2tevmyc</i>	Made in this study
<i>UASlacZ</i>	(Reim and Frasch, 2005)
<i>UASmcherrymoesin</i>	Gift of Tom Millard
<i>UASlola-R-GFP</i>	Made in this study
<i>UASmid</i>	Gift of William Brook
<i>UASlola-BC</i>	BL 28829/(Spletter et al., 2007)

4.1.4 Antibodies

Primary antibodies used in this work are listed in table 4.3. For visualization of the primary antibody-staining pattern, secondary antibodies conjugated to the fluorophore Alexa488 (Invitrogen), Cy3, Cy5 and Biotin (Jackson ImmunoResearch) were used at 1:500. DAPI was used as a nuclear stain at 1:1000.

Table 4.3: Showing the primary and secondary antibodies with the dilution used.

Primary Antibody	Dilution used	Reference
anti-Vasa (rabbit)	1/10,000	Gift from Ruth Lehmann
anti-Vasa (chicken)	1/10,000	Gift from Ruth Lehmann
anti- β -galactoside (rabbit)	1/10,000	Cappel
anti- β -galactoside (mouse)	1/1000	Promega Z378a
anti-Wun2 (rabbit)	1/200	Gift from Andrew Renault
anti-Myc (rabbit)	1/1000	Abcam ab9106
anti-dsRED/mCherry (rabbit)	1/500	Clontech
anti-GFP (rabbit)	1/1000	Abcam ab290
anti-GFP (mouse)	1/1000	Clotech JL-8
anti-GFP (chicken)	1/1000	Abcam ab13970
anti-Digoxigenin conjugated to POD (sheep)	1/250	Roche
anti-Digoxigenin conjugated to AP (sheep)	1/2000	Roche
anti-BP102 (mouse)	1/1000	Abcam ab12455
anti-Nmr2 (rabbit)	1/1000	Gift from Sandra Leal
anti-Lola (rabbit)	1/50	Gift from Edward Giniger
anti-Traffic jam (guinea pig)	1/10,000	Gift from Dorothea Godt
anti-Tinman (rabbit)	1/1000	Gift from Manfred Frasch

anti-Spectrin (mouse)	1/10	DSHB 3A9
anti-Armadillo (mouse)	1/50	DSHB N27A1
anti-DE-cadherin (mouse)	1/25	DSHB DCAD2
anti-Eyes Absent (mouse)	1/12	DSHB 10H6
anti-FasciclinIII (mouse)	1/50	DSHB 7G10
anti-Robo (mouse)	1/10	DSHB 13C9

4.1.5 Oligonucleotides

Several oligonucleotides have been used in this study for cloning, sequencing and riboprobe manufacture.

Table 4.4: List of oligonucleotides, their sequences and usage.

Primer name	5'-3' sequence	Annealing temperature	Used for
wun2tev FP	CACCATGAGCACCCCTGCGACCCGT	58°C	Overlap-extension pcr for <i>wun2tev</i>
wun2tev(2) RP	GCCCTGAAAATACAGGTTCTCCGC CGCACAGGTGAACTCCTC		
wun2tev(2) FP	GAGAACCTGTATTTTCAGGGCGTG GACATTACATCGAAGCAG		
wun2tev RP	TTAGGTGAGGTCGCCCAAGCTCTC		
wun2tev FP	CACCATGAGCACCCCTGCGACCCGT	55°C	Cloning of <i>wun2GFP</i>
wun2-nostop RP	CATAGCTTTAAATCGATGGGATCTCC		
lola-ex5begin	ATTTGAGCAGGAAGGAGAACACCG	55°C	Sanger sequencing of <i>C28</i> mutation, Synthesis of <i>lola</i> -R riboprobe
lola-ex5end	TTATTTGGTTTCAAGCTCTCCCTTCC		

mid int3-4 genomic FP	AAGCACTTTCCTAATTCCTCCAGCTATG	58°C	Sanger sequencing of <i>B23</i> mutation
mid-cds1713 RP	CGGGGCTGCGTGAGGCCTTTAC		
mid cds-begin	ATGCTGGTTGGATCTCATCCG	58°C	Sequencing of <i>midline</i> cds
mid cds-end	GTGGGCGGGGCTGCG		
lola common FP	ATGGATGACGATCAGCAGTTTTG	55°C	Synthesis of <i>lola-c</i> riboprobe
lola common RP	CTTGGGGATCCCGTTGCTG		
lola-exonBC FP	CTGAGTATTACGCGTATAGCGGGACTC	58°C	Synthesis of <i>lola-BC</i> riboprobe
lola-exonBC RP	CGAGAAGTGGGGATGCAACTCC		
lola-R cds FP	CACCATGGATGACGATCAGCAGTTTTG	55°C	Cloning of <i>lola-R</i>
lola-R cds nostop RP	TTTTGTTTTCAAGCTCTCCCTTCCC		
T7 FP	TAATACGACTCACTATAGGG	55°C	Synthesis of <i>412</i> riboprobe
T3 RP	AATTAACCCTCACTAAAGGG		
Sp6	CATTTAGGTGACACTATA	-	Synthesis of riboprobes
GW1	GTTGCAACAAATTGATGAGCAATGC	-	Sequencing forward primer
GW2	GTTGCAACAAATTGATGAGCAATTA	-	Sequencing reverse primer
M13 Forward	GTAAAACGACGGCCAG	-	Sequencing forward primer
M13 Reverse	CAGGAAACAGCTATGAC	-	Sequencing reverse primer
HSPf	TATAAATAGAGGCGCTTCGT	-	Sequencing forward primer

4.2 Methods in molecular biology

4.2.1 Polymerase chain reaction

The following protocol was used:

10X Pfx buffer	2.5 μ l
Pfx enzyme (2.5 units/ μ l)	0.25 μ l
50mM MgSO ₄	0.5 μ l
10mM dNTPs	0.5 μ l
DNA	2-10ng
Forward primer (100pmol/ μ l)	0.5 μ l
Reverse primer (100pmol/ μ l)	0.5 μ l
MilliQ water	q.s.
Total	25 μ l

The PCR was performed in a DNA Engine thermal cycler (Biorad) with the following program. The annealing temperature (Table 4.) was specific to the melting temperature (T_m) of the primers.

1. Denature	94°C	5 min
2. Denature	94°C	30 sec
3. Anneal	$T_m-5^\circ\text{C}$	1 min
4. Elongate	72°C	2 mins
5. Repeat steps 2 to 4, 29 times		
6. Final extension	72°C	10 min
7. Halt reaction	4°C	forever

PCR products were visualized on a 1% Agarose-Gel stained with SYBRSafe DNA gel stain.

4.2.2 TOPO® cloning

DNA was cloned into entry vectors for the Gateway® System including the pCR®8/GW/TOPO® TA and pENTR™/D-TOPO®, and into dual promoter vectors including pCR™II-TOPO®. (see appendix for the pCR8 vector map).

The following modified TOPO reaction was used:

DNA	5-10ng
Vector	1 μ l
Salt solution	1 μ l
Water	q.s.
Total	6 μ l

This reaction was kept at RT for 30 minutes. 3 μ l of the reaction was used for transforming competent *E. coli* cells.

4.2.3 Transformation of chemically competent cells

Transformation was performed using 50 μ l of One Shot[®] TOP10 Chemically Competent *E. coli* cells (Invitrogen). The cells were thawed on ice and the required plasmid was added and left on ice for 30 minutes. Then the cells were incubated in a waterbath at 42°C for one minute, and immediately put back on ice. The cells were then allowed to recover in 1ml of LB medium for one hour at 37°C and then plated on agar plates containing the appropriate antibiotic selection. Positive colonies could also be selected for plasmids containing the *beta-galactosidase* gene at the cloning site. For this 40 μ l X-Gal (5-bromo-4-chloro-3-indolyl b-D-galactopyranoside) was plated on the LB-agar plate, 30 minutes prior to cell plating. The next day, blue-white-screening could then identify colonies carrying the vector with the insert.

For plasmid isolation 2ml liquid LB-medium containing the appropriate antibiotic was inoculated with a single colony and grown overnight at 37°C. The next day plasmid mini preparations were done using the Qaigen QIAprep Spin Miniprep Kit following the recommended protocol.

4.2.4 Sequencing

Sequencing was carried out by mixing 0.5 μ l of BDT RRMix, 1.9 μ l of 5x Buffer, 150 – 300ng of Template DNA, 4pmol of primer and bidest water was added to make up the volume to 10 μ l. Thermal cycling was performed with the following programme.

1. Denaturation 96°C 2 min
2. Denaturation 96°C 20 sec
3. Annealing 50°C 10 sec
4. Elongation 60°C 4 min
5. Repeat 2. to 4. 29 times
6. Halt reaction 8°C forever

Sequencing was performed in the Max Planck Institute for Developmental Biology by the in house sequencing service unit on a Abi3730XL (Applied Biosystems) sequencer.

4.2.5 Gateway reaction

The Gateway® System allows the cloning, combining and transferring of DNA fragments from one vector to another. This is done by site-specific recombination mediated at the complementary *att*-sites present on the donor and the recipient vector, so as to preserve the orientation and reading frame of the insert during the transfer (Fig. 4.1). One can choose a range of recipient vectors with different promoters, 5' and 3' tags.

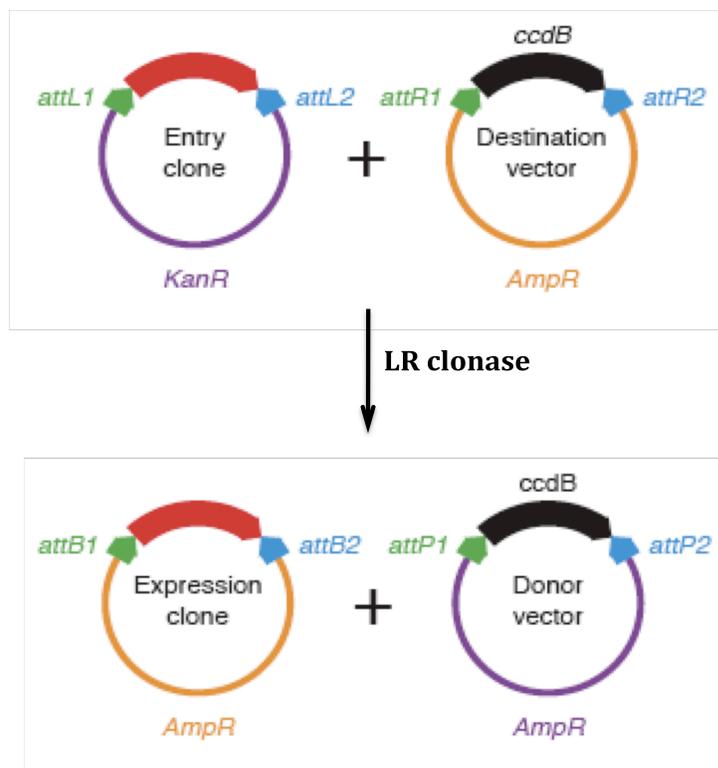


Figure 4.1: Schematic representation of the Gateway reaction. It depicts the transfer of the insert (red) from the entry clone (purple) into the destination vector (orange), catalyzed by the LR clonase enzyme. The *ccdB* gene (shown in black) encodes for a toxin that prevents recovery of the donor vector. (Adapted from the Invitrogen Gateway® manual)

The gateway reaction was carried out as follows:

Entry clone	50-100ng
Destination vector	150ng
TE Buffer (pH 8.0)	5 μ l
LR clonase	2 μ l
Total	10 μ l

The reaction was incubated at RT for 1-2 hours, after which 1 μ l of Proteinase K was added and kept at 37°C for 10 minutes to stop the reaction. 1 μ l of the reaction was then used to transform One Shot® TOP10 Chemically Competent *E. coli* cells.

4.3 Methods using *Drosophila* embryos

4.3.1 Methanol fixation

Embryos were dechorionated by incubating them with 50% bleach for three minutes. Afterwards the embryos were washed in a sieve, and then transferred into a scintillation vial containing 1.75 ml PBS, 8 ml heptane and 0.25 37% formaldehyde. The embryos stayed at the heptane/PBS interface in the vial. They were shaken on an orbital shaker at 200 rpm for 20 min. Fixation was stopped by removing the PBS phase and adding 7 ml methanol and shaken vigorously for one minute to remove the vitelline membrane. Properly fixed embryos sank into the lower methanol phase. The embryos were transferred into an eppendorf tube and washed with methanol several times. Fixed embryos were stored in 100% methanol at -20°C. For staining using the DCAD2 (DHSB) antibody, the fixation was performed adding 2 μ l of 1M CaCl₂ in the fixing solution, and the methanol was replaced by ethanol at all steps. These embryos however, were not stored, but instead used immediately for antibody staining.

4.3.2 Antibody staining of *Drosophila* embryos

Antibody staining was carried out using fixed embryos from over-night laying cages, including embryos of all stages. After fixation the embryos were rehydrated by washing them in PBST. Afterwards they were incubated with 0.1% BSA in PBST for one hour on a nutator in order to reduce unspecific binding of the antibody. The embryos were incubated over night with the primary antibodies of appropriate dilutions in 0.1 % BSA in PBST. The next morning embryos were rinsed with 0.1% BSA in PBST four times, for one minute, 5 min, 15 min and 30 minutes (for fluorescent antibody staining, DAPI was added in a 1:1000 dilution during the 15 min wash). Following the washes, the secondary antibodies were added in a 1:500 dilution in 0.1% BSA in PBST. After 2 hours of incubation the embryos were rinsed four times with PBST for one minute, 5 min, 15 min and 30 minutes. For fluorescent imaging, the embryos were mounted in Aquamount (Polyscience Inc.).

For the detection of biotin conjugated secondary antibodies, during the last wash vectastain solution (VECTOR laboratories) was prepared and then added for 30 minutes. Afterwards the embryos were washed four times and then stained with DAB solution (1 μ l H₂O₂ and 10 μ l DAB (20mg/ml) per ml of PBST). Staining was completed after approximately five to ten minutes, then DAB was removed and the embryos were washed with PBST. The embryos were either directly mounted in 70 % glycerol in PBS or dehydrated by washing them in 30%, 50%, 70% and 100% ethanol, then 100% acetone and then mounted in epon resin which was polymerized by baking overnight at 65° C.

4.3.3 Genomic DNA isolation for PCR

Embryos were dechorionated as in section 4.3.1. 50-100 wild type or homozygous mutant embryos (sorted by lack of fluorescent balancer) were collected in an eppendorf tube and frozen at -80°C. The embryos were homogenized in 60 μ l of Buffer A (100mM Tris-HCl pH 7.5, 100mM EDTA, 100mM NaCl, 0.5% SDS) with a disposable tissue homogenizer. An additional 60 μ l of Buffer A was added with continued homogenization until only the cuticles remained. The tube was incubated

at 65°C for 30 minutes to aid lysis. 240µl of LiCl/KAc (1x 5M KAc: 2.5x 6M LiCl) solution was added, the mixture incubated on ice for at least 10 minutes, and centrifuged for 15 minutes at RT. 1 ml of the supernatant was transferred into a new tube, and 180µl isopropanol added, mixed, and centrifuged again for 15 minutes at RT. The obtained pellet was washed with 70% ethanol, dried, resuspended in 20-30µl Tris-EDTA and stored at -20°C.

4.3.4 RNA Isolation

An overnight collection, consisting of embryos of all stages, was dechorionated as in section 4.3.1. To the collected embryos, 500µl of Trizol® was added and homogenized using an eppendorf pestle. Following incubation of the homogenized embryos for 15 minutes at RT, 100µl of chloroform was added, the mixture vortexed and kept at RT again for 15 minutes. The tube was then centrifuged for 15 mins at 13,000rpm. The upper phase of the mixture was then transferred into a new tube and 250µl of isopropanol was added and briefly vortexed. This solution was then left at RT for 5-10 minutes and centrifuged again at 4°C to pellet the RNA. The obtained pellet was washed by centrifuging for 5 minutes at 4°C with 70% ethanol. The pellet was dried at 65°C for 30 seconds, resuspended with 50µl of double distilled water and stored at -80°C until further use.

4.3.5 cDNA synthesis

1-5µg of the isolated RNA was used for making cDNA, using AMV Reverse transcriptase. The following components were added:

RNA	1-5µg
dNTP (Roche)	1µl
Oligo dT primers (Promega)	1µl
MiliQ water	q.s.
Total	14 µl

The mixture was heated to 65°C and snap-chilled in ice. To this the following components were added:

Buffer	4 μ l
0.1M DTT	1 μ l
RNase Inhibitor	1 μ l
AMV Reverse transcriptase (Invitrogen)	1 μ l

The tube with all the components was kept at 45°C for 1 hour for cDNA synthesis. The reaction was stopped by heat inactivation for 20 minutes at 80°C, and the cDNA immediately stored at -20°C until further use.

4.3.6 Synthesis of riboprobes

In this study, riboprobes against the 412 reterotransposon, *lola*-common region, *lola*-R specific exon and *lola*-BC specific exon were synthesized.

➤ PCR amplification from cDNA

The 412 reterotransposon was already available cloned into the pBluescript vector (gift from Akira Nakamura). To clone the other three sequences, PCR amplification was performed from wild type cDNA, using the JumpStart™ REDTaq® ReadyMix™ RaactionMix (Sigma-Aldrich, including buffer, *Taq* polymerase, dNTPs, loading buffer for the agarose gel)

JumpStart™ REDTaq® ReadyMix™ RaactionMix	7.5 μ L
cDNA	1 μ L
Forward primer (100 pmol/ μ L)	1 μ L
Reverse primer (100 pmol/ μ L)	1 μ L
MiliQ water	4.5 μ L
Total volume	15 μ L

The amplification was performed according to the parameters mentioned in section using a primer specific annealing temperature (section).

➤ Cloning into dual-promoter vector

The PCR product was then cloned into a dual promoter vector using the pCR™II-TOPO® Cloning® Kit (see appendix for the pCR8 vector map). The *pCR™II-TOPO®* vector consists of the *SP6* and the *T7* promoter sequences 5' and 3' of the cloning site, respectively. The TOPO reaction was setup according to the protocol mentioned in section 4.2.2, and was subsequently transformed into the One Shot® TOP10 Chemically Competent *E. coli* cells (section 4.2.3). The positives clones were selected

by blue white screening, and tested for presence of insert by restriction enzyme digestion. The clones containing the insert were then sequenced using the M13 forward and reverse primers, to ascertain the orientation of insertion and hence the promoter to use for antisense probe transcription.

➤ PCR amplification from clones

Once the direction of insertion was determined, the insert was amplified from the plasmid using a primer binding the promoter closest to the 3' end of the insert, and a primer binding to the other end of the insert. The PCR was performed as mentioned in section 4.2.1 with annealing temperature of 55°C.

➤ In vitro transcription

The following components from the SP6/T7 Dig RNA labeling kit (Roche) were added:

Nuclease free water	7µl
PCR product	6µl
DIG labeled NTPs	2µl
Transcription Buffer	2µl
RNase inhibitor	1µl
Sp6 Polymerase	2µl
Total	20µl

The mixture was incubated in a 37°C water bath for 2 hours for transcription. The reaction was stopped by adding 1µl 0.5M EDTA, and run on a 1% agarose gel stained with SYBR Safe DNA gel stain (Invitrogen) in order to check whether riboprobes of the expected size had been successfully synthesized.

➤ Carbonating the riboprobes

Afterwards the riboprobes were carbonate chopped by adding 40 µL water and 50 µL 2x carbonate buffer (120 mM Na₂CO₃, 80 mM NaHCO₃, set pH to 10.2). After an incubation at 60°C for 40 minutes, the reaction was stopped by adding 10 µL 0.2 M Na-Acetate (0.41 g Na-Acetate, 25 mL H₂O_{dd}, 45 uL HCl [5M], pH 6, RNase free).

➤ Precipitation of riboprobes

The riboprobes were precipitated by adding 20 µL 4 M LiCl, 10 µL 20 mg/mL tRNA and 600 µL ethanol (RNase free) and incubating at -20°C for a few hours or at -80°C for 10 minutes. Afterwards the sample was centrifuged at 4°C for 15 minutes and the supernatant was discarded. The obtained pellet was washed with 70% ethanol,

dried and resuspended in 100-150 μ L hybridisation solution [50mM] (50% Formamide, 5x SSC (SSC 20x: 175 g NaCl [1M], 88 g Na₃Citrate, 1000 mL H₂O_{dd}, set pH to 7 with HCl), 100 μ g/mL tRNA, 50 μ g/ml Heparin, 0.1%Tween) and incubated at room temperature for 15 minutes. The riboprobes were stored at -20°C.

4.3.7 *In-situ* hybridisation in *Drosophila* embryos

The fixed embryos were rinsed in methanol (-20°C). Then they were incubated for 25 minutes in 1:1 methanol: 5% formaldehyde in PBST solution for 5 minutes followed by 25 minutes in PBST + 5% formaldehyde. After 5 rinses with PBST, Proteinase K (4 mg/mL in PBST) was added for exactly 3 minutes. Afterwards the embryos were rinsed three times in PBST and then postfixed in PBST + 5% formaldehyde for 25 minutes on the nutator. The embryos were washed 5 times in PBST and then incubated 5 minutes in 1:1 PBST:hybridisation solution. The embryos were rinsed in hybridisation solution and prehybridised in hybridisation solution at 55°C. After 1 - 2 hours incubation time the solution was replaced with 90 - 100 μ L hybridisation solution + 1 - 10 μ L probe and the embryos were incubated at 55°C over night. Following hybridization the probe was removed and stored at -20°C for reuse (up to 3 times). The embryos were washed several times with prewarmed hybridisation solution at 55°C. Afterwards the embryos were incubated for 20 minutes in 1:1 PBST:hybridisation solution while shaking and then washed 4 times, 15 minutes each wash, in PBST.

➤ For chemical staining using Alkaline Phosphatase

Following the washes, anti-Dig conjugated to Alkaline Phosphatase (1:2000 in PBST) was added to the embryos. After incubation of 1 hour at RT (or at 4°C over night), the embryos were washed thrice in PBST for 20 minutes each. Then the embryos were rinsed in AP staining buffer (100mM NaCl, 50mM MgCl₂, 100mM Tris pH 9.5, 0.1%Tween), following which 9 μ L NBT (100 mg/mL) and 7 μ L BCIP (50 mg/mL) in 1 mL AP staining buffer was added. The embryos were transferred to a glass-staining dish to monitor the staining. Once a visible patterning was observed,

replacing the staining buffer with PBST stopped the reaction. The stained embryos were washed in PBST and mounted in 75% glycerol. The slides were stored at 4°C.

➤ For fluorescent staining using Tyramide labeling

After the PBST washes, the embryos were incubated in a 1% BSA blocking solution (1% BSA in PBST). Following this a 1:250 dilution of anti-Dig-Peroxidase in blocking solution was added to the embryos and kept for 4 hours at RT or overnight at 4°C (for combining *in situ* hybridization with antibody staining, a primary antibody was added here). Next, the embryos were washed 4 times for 15 minutes each using 0.1% BSA in PBST. The amplification buffer/H₂O₂ stock solution was prepared by adding 1 μL of H₂O₂ to 200 μL amplification buffer (Invitrogen Molecular Probes™ Tyramide Signal Amplification Kit). Next, a working solution was made by adding 3 μL of the stock solution to 297 μL of the amplification buffer. The 300 μL of amplification buffer working solution was then supplemented with 3 μL of labeled Tyramide solution and added it to the embryos for 60–90 minutes with constant rocking in the dark. Then the embryos were washed 3 times with PBST for 2mins each and unless an antibody staining was being performed the embryos were mounted in Aquamount. If an antibody staining was also performed, then a further blocking for 1 hour was required. Then the appropriate secondary antibody was added, and incubated for two hours at RT. Lastly, the embryos were washed 4 times (10 minutes each) and mounted in Aquamount.

4.3.8 Making Transgenic Flies

For making transgenic lines containing UASwun2GFP, and UASlola-R-GFP, *wun2* and *lola-R* were cloned into the vector pUAS-*attB*-WG (PhD thesis Khalid Al-Jubran), using the Gateway® reaction. The UAS/Gal4 system is often used in *Drosophila* for expression of transgenes in a tissue specific manner. In addition this vector encodes GFP 3' to the cloning site to generate C-terminal GFP fusion proteins.

The bacteriophage PhiC31 *integrase* (Bischof et al., 2007) catalyses the recombination of a specific *attB* sequence (present on the plasmid) with an *attP*

sequence (in the fly genome). Therefore, several fly lines with *attP* sites integrated at different but known chromosomal positions (available from the Basler lab) and germline expressed *integrase* can be used to make site-specific transgenic lines.

DNA for injection was prepared from 50ml of LB culture using the QiaGen QIAprep Spin Midiprep Kit, according to the instructor's manual. The eluted plasmid was precipitated again using isopropanol, and dissolved in 15-20 μ l of MilliQ water. For injection the plasmid concentration was diluted down to 500ng-1 μ g/ μ l.

Flies containing the PhiC31 *attP* site on the third chromosome (lines VK00020) were placed into laying cages at RT. Egg laying was synchronized by repeated changing of apple juice agar-plates every two hours for 6-7 hours. After synchronization, the eggs laid were dechorionated and lined on an agar-piece with the posterior of the egg placed towards the edge. The eggs were then transferred onto a coverslip and dried for 10-12 minutes. Immediately 10S halocarbon oil was applied onto the eggs to prevent further drying. Embryos were injected with DNA into their posterior using a FemtoJet injector (Eppendorf).

Post injection, the embryos were allowed to recover, and the hatched larvae were collected after 2-3 days, and transferred to a food vial kept at 25°C. The hatched flies were then crossed to *white* flies to screen for transgenic lines expressing the mini-white gene in the eyes of the next generation.

4.4 Insertion of *tev*-site in *wun2* gene

To determine the site of requirement of the Wun2 lipid dephosphorylation activity, a cleavage site recognized by the TEV (tomato etch virus) protease was introduced in the gene. The *tev*-site was introduced in the 3rd loop of Wun2, separating the two catalytic domains in the same loop. In order to achieve, this several steps of PCR and cloning were undertaken.

➤ Overlap-Extension PCR

This technique provides a means of introducing mutations in the middle of the PCR fragment. Here I have used this technique to introduce a 21-nucleotide (ntd) protease cleavage site (Fig. 4.2).

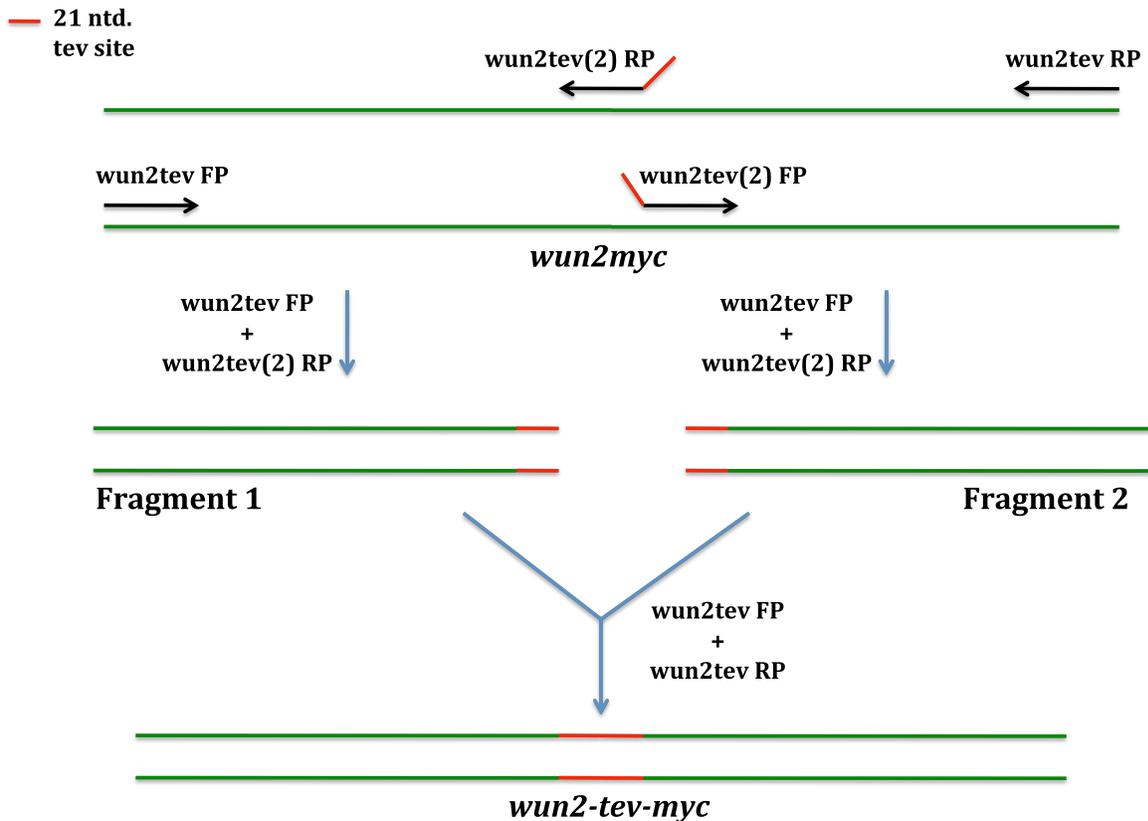


Figure 4.2: Schematic representation highlighting the steps of overlap-extension PCR. The *wun2myc* gene was first divided into two fragments using 2 separate pairs of primers, where the internal primers contained a *tev*-site. The two fragments, now overlapping at the *tev*-site, were then fused back together and amplified with the two external forward and reverse primers, producing *wun2-tev-myc*.

In this technique the gene is divided into two fragments using two different pairs of primers (Fig. 4.2). The first primer pair consists one flanking forward primer (*wun2tev FP*) starting at the 5' end of the gene, and a reverse internal primer [*wun2tev(2) RP*]. Similarly, for the second primer pair the forward primer is an internal primer [*wun2tev(2) FP*] and the reverse primer (*wun2tev RP*) is a flanking primer ending at the 3' end of the gene. The two internal primers however have an

overlapping complementary sequence, the 21ntd *tev*-site.

The two sets of primers were used in two different PCR reactions, resulting in two fragments of the gene, with a small region of complementarity corresponding to the *tev*-site. The PCR protocol mentioned in section was used for amplification from the pBS-wun2myc vector in a DNA Engine thermal cycler from Biorad, with the annealing temperature of 58°C.

This protocol was also used to make *wun2-con-myc* construct, without the *tev*-site, using the two flanking wun2tev FP and wun2tev RP primers.

➤ Touchdown PCR

This technique is used to enhance sensitivity and product specificity. Here instead of one annealing temperature, a range of annealing temperatures is employed in a single PCR reaction. The initial temperature is above the calculated T_m of the primers, and with each progressing cycle the temperature falls down by 1°C till the lowest specified temperature.

The overlap-extension PCR produced two large fragments of the first and the second half of *wun2* each containing the *tev*-site at the 3' and the 5' end, respectively. In this study, touchdown PCR method was adopted to find a suitable temperature for the annealing of the, produced from previous PCR reaction. In addition, touchdown PCR also provided suitable temperature for annealing of the flanking primers, to produce large number of copies of the final product with the inserted *tev* site.

The following protocol was used:

- | | | |
|-------------|---------|--------|
| 1. Denature | 94°C | 5 mins |
| 2. Denature | 94°C | 30 sec |
| 3. Anneal | 60-50°C | 1 min |
| 4. Elongate | 72°C | 2 mins |

This was repeated for 10 cycles (decreasing the annealing temp by 1°C each cycle)

- | | | |
|-------------|------|--------|
| 5. Denature | 94°C | 30 sec |
| 6. Anneal | 50°C | 30 sec |
| 7. Elongate | 72°C | 2 mins |

This was repeated for 30 cycles to give the product.

- | | | |
|--------------------|------|---------|
| 8. Final Extension | 72°C | 10 mins |
| 9. Halt reaction | 4°C | forever |

PCR products were visualized on a 1% Agarose gel stained with SYBRSafe DNA gel stain and bands of the correct size were cut out and eluted using the QuiaGen QIAquick® Gel Extraction Kit, using the recommended protocol.

➤ **TOPO-cloning into the entry vector**

The eluted DNA was then cloned into an entry vector for the Gateway® System using the pCR®8/GW/TOPO® TA Cloning® Kit. Plasmids were screened for insertions using EcoRI digestion. The positive clones were then sequenced (section 4.2.4) using the GW1 and GW2 sequencing primers, to verify the presence of desired *tev* site.

➤ **Gateway reaction**

The 2 inserts of interest, *wun2-con-myc* and *wun2-tev-myc*, cloned into the entry vector pCR8, (containing *att*-sites for recombination), could now be cloned into different expression vectors from the Gateway® System.

The expression vector pTW (The Drosophila Gateway™ Vector Collection, Murphy lab) has a UAS promoter sequence upstream of the insertion site, and no tags at the 5' or 3' ends (a tag was not required as the insert already contained a 3' myc-tag). This vector was suitable to check the functionality of the *wun2-tev-myc* construct by expression in *wunen* maternal mutants, using a *Gal4* driver.

In addition, the two constructs were also cloned into the pNOS vector (gift from Matt DeGenero) using the gateway reaction. This vector contains the *nanos* promoter, and the *nanos 3'UTR*. This would allow localization and expression of the constructs specifically in the germ cells, at physiological levels. The vector contains P-element inverted repeat ends for transposase mediated transformation which was carried out by Best Gene.

4.5 Whole genome sequencing

To identify the causative mutations in the *A44*, *B23*, and *C28* mutants, whole genome sequencing was performed in collaboration with the in house genome sequencing facility in Max Planck Institute of Developmental Biology, with the Illumina sequencing method and using the Genome Analyzer II.

4.5.1 Genomic DNA isolation from mutants

In order to prepare 3µg high quality genomic DNA for the purpose of whole genome sequencing of the three EMS mutants displaying gonad formation defects, high quality DNA was prepared. Minimally 200 homozygous mutant embryos were homogenized in 120-150µl Buffer A. The homogenized embryos were then treated with 6µl of RNases A/T1 mix (Fermentas) for 50-10 minutes at RT. Following treatment with RNase, 6µl of 4mg/ml Proteinase K was used and incubated for 30-45 minutes at 65°C. Next, 16µl of 7.5M Ammonium acetate was added to the mixture and kept on ice for 10 minutes. The solution was then treated as in the genomic DNA isolation described earlier (section 4.3.3), with centrifugation at RT for 15 minutes. The supernatant was transferred into another tube, where the genomic DNA was precipitated using isopropanol, washed with 70% ethanol and resuspended in sequencing grade water for further processing.

4.5.2 Library Preparation

The genomic DNA was used to prepare a library for 100bp pair-end sequencing as follows:

1. Fragmentation of DNA: The isolated 3µg of genomic DNA was sheared into 500bp fragments using the S2 Covaris machine, according to the manufacturer's instructions. This generates dsDNA fragments with 3' to 5'

overhangs. The sheared DNA was purified using the Qiagen Qiaquick PCR-purification kit, and eluted in 31µl of elution buffer (EB). The concentration of the DNA was then quantified and 1µg of DNA was used for subsequent steps.

2. End repairing of the fragmented DNA: The NEB Next End Repair Enzyme Mix with accompanying buffer was added to 1µg of DNA and incubated at 20°C for 30 minutes for end repairing. This was done to convert the overhangs, resulting from the fragmentation process, into blunt ends. The DNA was purified again using the Qiagen Qiaquick PCR-purification kit, and eluted in 42µl of EB.
3. dA-Tailing: The 3' end of the fragments were attached with extra dA residues to allow efficient adaptor ligation. Further, this process prevents the blunt-ended fragments from ligating to one another. The NEB kit containing the dA tailing buffer and Klenow exo, were used at 37°C for 30 minutes, following which the tailed-DNA was purified and eluted in 10µl EB.
4. Adaptor ligation: This process is required to allow the fragments to hybridize onto the flow cell. The adaptor also provides the identity to a DNA sample, and allows multiplexing of samples. 2µl of custom-made adaptors were used per µg of DNA. The adaptor sequence for the *A44* DNA was ACGTAGCT, for *B23* DNA was TACGTCAG, and for *C28* DNA was CGTCGATA. The following reaction was set up for adaptor ligation:

DNA 10µl

Ligation Buffer 5x (NEB) 10µl

Adaptor DNA 2.5µl

T4 DNA ligase (NEB) 5µl

MilliQ water 22.5

Total 50µl

The reaction was incubated for at 30°C for 10 minutes. After ligation, the DNA was again purified and eluted in 30µl EB.

5. Size selection: The eluted DNA was then run on 2% Agarose gel at 120V for 90 minutes. A scalpel was then used to excise selectively the DNA band of

approximately 450-500bp using the marker as a guide. The excised piece of gel was then purified using the Qiagen gel-extraction kit, and eluted in 20µl of EB. This purified the ligation product and removed the unligated adaptors.

6. Enrichment by PCR: The eluted DNA was then PCR amplified to selectively enrich those DNA fragments that have adaptor molecules on both ends. It was done using reagents from the TruSeq DNA sample prep kit with the following reaction mix:

DNA 20µl

Primer cocktail (Illumina) 5µl

Mastermix 25µl

Total 50µl

The reaction mix was then amplified using the following parameters:

Denaturation 98°C 30 sec

Denaturation 98°C 10 sec

Annealing 60°C 30 sec

Elongation 72°C 30 sec

Repeat 2nd to 4th steps 12 times

Elongation 72°C 5 min

Halt reaction 8°C forever

The amplified fragments were then purified and eluted with 50µl of EB.

4.5.3 Validation of Library

Before continuing with the library for Illumina sequencing, a quality control analysis of the library and quantification of the DNA was performed. Validation of the libraries was performed using the Agilent Bioanalyzer 2100, by running 1 µl of sample on a DNA 1000 chip (Agilent Technologies, Germany). This verified the size of the PCR enriched fragments and their size distribution.

The average size of the library for *A44* was 449bp at a concentration of 11.39ng/µl, for *B23* was 448bp at a concentration of 7.51ng/µl and for *C28* was 457bp at a concentration of 11.21ng/µl.

Using the data obtained from the Bioanalyzer, the DNA was then diluted and stored at -20°C as a 10 nM stock in EB buffer (Qiagen) supplemented with 0.1% Tween-20, until cluster generation.

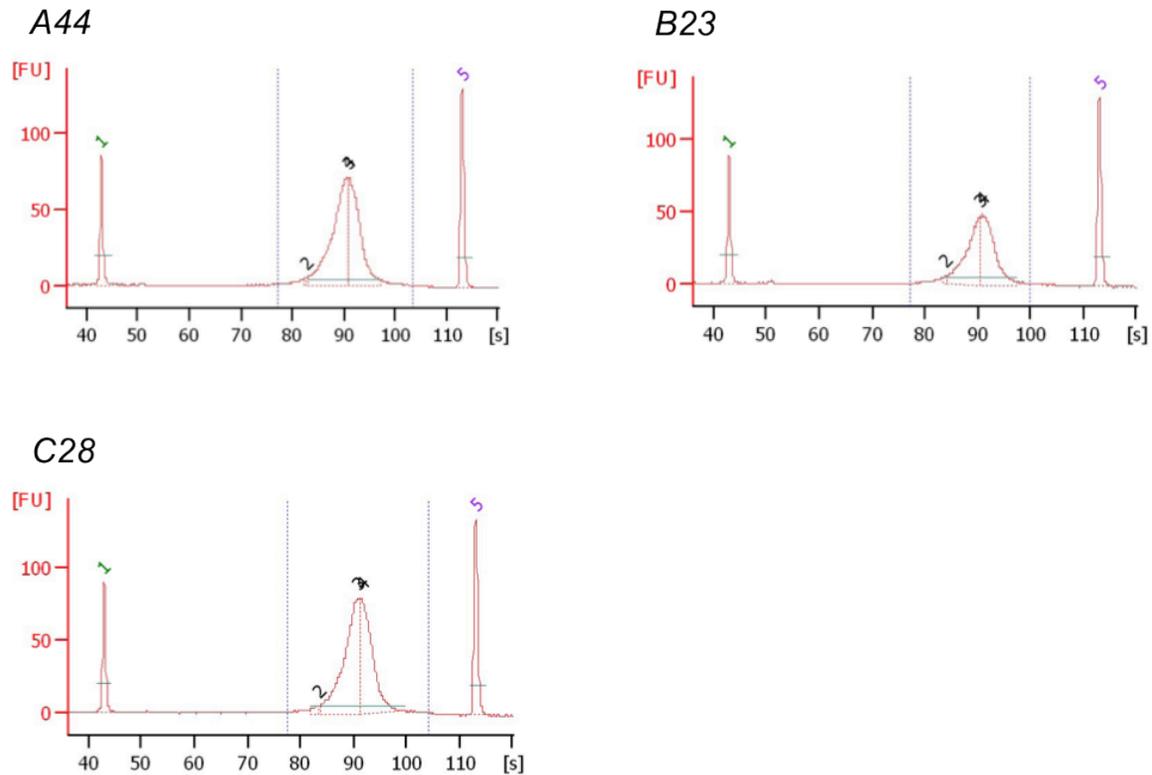


Figure 4.3: Electropherogram produced by the Bioanalyzer, showing the size distribution of the DNA fragments in the three mutant libraries. The x-axis [s] depicts the sedimentation coefficient, a measure of size, and the y-axis [FU] depicts the fluorescence unit, a measure of quantity. The numbers 1 and 5, show the markers at 15bp and 1,500bp respectively. The number 2 and 3 in each graph, denote the base and the height of the peak of the library, in order to calculate the area under the curve.

4.5.4 Cluster preparation on flowcell and SBS Sequencing by Synthesis sequencing

Cluster generation was performed on a Cluster Station (Illumina) with the recipe “PE_Amplification_2P_R1prep_Linearization_combinedBlocking_Priimerhyb_v7.0” and reagents from Cluster Generation Kits (Illumina) according to the manufacturer’s instructions. With a 10 pM solution of denatured DNA, ~575 000

raw clusters per mm² were generated on the flow-cell surface as assessed at the sequencing step (see below). Sequencing was performed on a Genome Analyzer GAI (Illumina) using the SBS TruSEQ™ 36-cycle Sequencing Kits (Illumina) according to the manufacturer's instructions; the .xml file for recipe "GA2-PEM2x-2x101cyclev8.3px.xml" was used to generate 101 bp reads.

4.5.5 SHORE analysis of obtained reads

Image analysis of the output from Genome Analyser GAI runs was performed with GAPipeline version 1.0, with default parameters, resulting in 3.6×10^6 and 2.9×10^6 and 4.7×10^6 read pairs for the *A44*, *B23* and *C28* mutants, respectively. Unfiltered reads were used as input for the SHORE pipeline (Ossowski et al., 2008), version 0.5beta (<http://1001genomes.org/downloads>). Reads were subjected to quality filtering, trimming, adaptor clipping and demultiplexing, leaving reads of an average length of 97bp for each mutant.

4.5.6 Analysis of sequences using CLC Genomic workbench

The CLC Genomic Workbench is a desktop application with a graphical user-interface, commonly used for analyzing and visualizing next generation sequencing data. The CLC Genomics Workbench versions 4.7 - 5.5 were used in this study, and the Genomic Gateway plugin tool (later integrated into the software as the 'NGScore tools') was used for mapping of reads onto the reference genome.

➤ Importing reads

The paired reads from the three mutants, in a FASTQ format were first imported into the CLC genomic software version 4.7.2, using the following parameters:

Discard sequence names = Yes

Discard quality scores = No

Paired orientation = Paired reads (forward-reverse)

Minimum distance = 180

Maximum distance = 250

➤ Mapping reads

The imported reads were then mapped onto the chromosome 2L complete sequence (GenBank NT_033779 Release 5.30) as well as chromosome 2R complete sequence (GenBank NT_033778 Release 5.30). The following parameters were used for the mapping of the reads:

Mismatch cost = 2

Insertion cost = 3

Deletion cost = 3

Length fraction = 0.5

Similarity = 0.8

Paired reads mode = Forward/Reverse

Min distance = 180

Max distance = 250

Alignment mode = Local

This achieved an average genome read coverage of about 16x, 13x and 19x for the *A44*, *B23* and *C28* mutants, respectively.

➤ SNP calling

The aligned reads could then be used for detecting SNPs when compared to the reference genome. This was performed using the SNP detection tool with the following parameters:

Window length = 11

Maximum gap and mismatch count = 2

Minimum central quality = 20

Minimum average quality = 15

Minimum coverage = 3

Minimum variant frequency (%) = 70.0

Maximum expected variations (ploidy) = 2

➤ Overlap filtering

As the three mutants *A44*, *B23* and *C28*, originated from the same isogenized genetic background, the SNPs common between the three mutants were then identified and removed from the SNP-list of each mutant. For example, this was executed for *A44* by using an ‘overlap filter’ with the following parameters:

Refiner = Overlap filter

Track parameter = reads *B23_1* (paired) SNPs

Keep overlapping = No

Refiner = Overlap filter

Track parameter = reads *C28_1* (paired) SNPs

Keep overlapping = No

➤ Non-synonymous SNPs

From the unique SNP-list (Table 2.), non-synonymous SNPs were identified using the ‘amino acid changes filter’. This was done by removing SNPs that did not cause a change in amino acid sequence compared to the reference. The following parameters were used:

Refiner = Amino acid changes annotation/filter

Track parameter (CDS track) = CDS

Filter synonymous changes = Yes

➤ Splice site SNPs

In addition, the SNPs that led to a possible change in the splicing of the mRNA hence a defective protein, were also identified. The filter ‘splice site effect prediction’ was used with the following parameters:

Refiner = Splice site effect prediction

Track parameter (mRNA track) = mRNA

4.5.7 Manual curation

Manual verification of the SNPs called by the software was performed to obtain the final working SNP-list (Tables 2.3, 2.4, 2.5). A number of the SNPs identified by the software were not true SNPs. This means that although, according to set parameters, more than three reads contained the nucleotide change, the minimum variant frequency of 70% was not always observed. Moreover, some SNPs showed the same base change in the other two mutants at that location, although at a lower frequency than 70%. Therefore manual verification significantly reduced the number of SNPs to approximately 20% of the previous list. For *B23* and *C28* mutants, where non-complementing deficiencies had already been identified through the deficiency screen, the final working SNP-list contained 15 and 29 SNPs respectively.

In case of the *A44* mutant, since no non-complementing deficiency could be found, further curation was performed on the verified 21 SNPs, based on the level of conservation of the mutated amino acid. SNPs causing changes to similar amino acids, for example a change of Ala to Val in CG6550 and CG6424, or Val to Ile in Cpr49Af, were not considered for further analysis. Additionally, two separate tools, BLAST (Ohsako et al., 2003) and ConSurf (Ashkenazy et al., 2010) were used to look at the conservation of the affected amino acid within other insect species. SNPs were discarded if the site of the mutation was variable in other insect species. As a result of the selection based on conservation, 11 of the 21 SNPs found on the 2R chromosome of the *A44* mutant were then tested as possible causative SNPs.

4.6 Microscopy

4.6.1 Confocal Microscopy

Fixed tissue imaging was done using the laser scanning Olympus FluoView FV1000 confocal microscope with FluoView Software. The samples were prepared as mentioned in section 2.5.2, and mounted in either aquamount or 1:2 benzylalcohol: benzylbenzoate (this mountant dissolves lipids thereby reducing light scattering).

An overview image of the embryo was taken using the 20X dry objective (UPlanSApo 20x, 0.75NA). To visualize the cell at higher resolution, the 60X water objective (UPlanSApo 60x W, 1.2NA) was used along with a 4x digital zoom. The lasers were used at wavelengths of 405nm (for DAPI) 488nm (for AlexaFluor 488), 559nm (for Cy3) and 635nm (for Cy5). In all cases, sequential scanning was performed to avoid bleed through and Kalman Filter used to decrease background noise. The images were analyzed using the Imaris 6.1.5 software (Bitplane) and an appropriate slice taken for presentation.

4.6.2 Light Microscopy

DAB and BCIP stained embryos were visualised on an Imager Z1 microscope (Zeiss), using a 20X objective (Plan-Apochromat 20x, 0.8NA) and images captured using an AxioCam HRc camera and AxioVision 4.6 software (Zeiss).

Contributions

The work described in this thesis was performed at the Max Planck Institute for Developmental Biology in Tübingen under the supervision of Dr. Andrew Renault. The mutants *A44*, *B23*, and *C28* were isolated by A. Renault, H. Sano, and P. Kunwar, in the lab of Prof. Ruth Lehmann (Skirball Institute of Biomolecular Medicine and Howard Hughes Medical Institute, New York University Medical Center). The library preparation, library validation, and cluster preparation on flowcell for whole genome sequencing was performed by Dr. Christa Lanz in the sequencing facility of Max Planck Institute for Developmental Biology. The obtained raw sequences were demultiplexed and filtered using the SHORE pipeline by Jörg Hagmann, Detlef Weigel department.

Appendix A

The Bloomington deficiency kit for chromosomal arm 2L

used for the lethality screening of the gonad formation mutants.

BL number	Deficiency	Cytological deletion coordinates	Sequence deletion coordinates
24626	Df(2L)ED50001	21A1;21B1	2L:-204333;72671
9353	Df(2L)ED5878	21B1;21B3	2L:67365;161120
8901	Df(2L)ED19	21B3;21B7	2L:159063;285763
24958	Df(2L)BSC454	21B7;21B8	2L:271351;307085
8672	Df(2L)BSC106	21B7;21C2	2L:291728--291846;417947
3548	Df(2L)al	21B8--C1;21C8--D1	
8673	Df(2L)BSC107	21C2;21E2	2L:431096;574741
3084	Df(2L)ast2	21D1--2;22B2--3	
24118	Df(2L)ED105	21E2;22A1	2L:852854;1420528
8908	Df(2L)ED94	21E2;21E3	2L:568095;1036969
7492	Df(2L)Exel6005	22A3;22B1	2L:1555098;1737249
26540	Df(2L)BSC688	22B1;22D6	2L:1736964;2273384--2273572
24959	Df(2L)BSC455	22D5;22E1	2L:2242285;2374023
7783	Df(2L)Exel7011	22E1;22F3	2L:2362917;2492447
90	Df(2L)C144	22F4;23C3	2L:2517598--2551864;2955279--2961962
9176	Df(2L)ED136	22F4;23A3	2L:2492935;2753125
7744	Df(2L)Exel6277	23A2;23B1	2L:2677694;2808100
26544	Df(2L)BSC692	23B3;23B7	2L:2830265--2830267;2868633
9610	Df(2L)BSC180	23B7;23C3	2L:2868633;2958118
8904	Df(2L)ED4651	23B8;23F3	2L:2873954;3478331
6507	Df(2L)drm-P2	23F3--4;24A1--2	
23677	Df(2L)BSC292	23F6;24A2	2L:3515462;3632008
24123	Df(2L)ED247	24A2;24C3	2L:3632218;3771177
5330	Df(2L)ed1	24A2;24D4	
7495	Df(2L)Exel6009	24C3;24C8	2L:3771368;3888977
9600	Df(2L)BSC165	24D4;24D8	2L:4031318;4195308
23680	Df(2L)BSC295	24D4;24F3	2L:4031318;4455780

744	Df(2L)M24F-B	24E1--2;24F6--7	
9270	Df(2L)ED250	24F4;25A7	2L:4477085;4821294
24124	Df(2L)ED7853	25A3;25B10	2L:4701129;5000402
1164	Df(2L)tkv3	25A4--5;25D5	
9605	Df(2L)BSC172	25B10;25C1	2L:5000838;5037253
8835	Df(2L)BSC110	25C1;25C4	2L:5029595;5064620
8674	Df(2L)BSC109	25C4;25C8	2L:5073453;5145500
7497	Df(2L)Exel6011	25C8;25D5	2L:5147258;5305646
7498	Df(2L)Exel6012	25D5;25E6	2L:5305646;5555049
9560	Df(2L)BSC169	25E5;25F3	2L:5524375--5524385;5716224
9343	Df(2L)ED334	25F2;26B2	2L:5658629;5999667
9341	Df(2L)ED385	26B1;26D7	2L:5980272;6465772
6338	Df(2L)BSC6	26D3--E1;26F4--7	
24378	Df(2L)BSC354	26D7;26E3	2L:6465706;6557463
9615	Df(2L)BSC188	26F1;27A2	2L:6612189;6742726
23676	Df(2L)BSC291	27D6;27F2	2L:7042642;7366119
9708	Df(2L)BSC233	27F3;28D2	2L:7388188--7388190;7887564
9502	Df(2L)BSC142	28C3;28D3	2L:7774037;8012787
7807	Df(2L)Exel7034	28E1;28F1	2L:8071311;8205166
9704	Df(2L)BSC227	28E8;29B1	2L:8155863;8346414
8836	Df(2L)BSC111	28F5;29B1	2L:8240266;8362842
24132	Df(2L)ED629	29B4;29E4	2L:8382851;8700124
9631	Df(2L)BSC204	29D5;29F8	2L:8529124;9025734
8906	Df(2L)ED678	29F5;30B12	2L:8958155;9581740
24133	Df(2L)ED690	30B3;30E4	2L:9437469;9918174
6478	Df(2L)BSC17	30C3--5;30F1	
9715	Df(2L)BSC240	30C7;30F2	2L:9744077;9960577--9960586
8469	Df(2L)BSC50	30F4--5;31B1--4	
26541	Df(2L)BSC689	30F5;31B1	2L:9984563;10198946--10198992
9503	Df(2L)BSC143	31B1;31D9	2L:10209408;10333704
1469	Df(2L)J39	31C--D;32D--E	
9635	Df(2L)BSC208	31D7;31D11	2L:10321809;10381214
9637	Df(2L)BSC209	31D7;31E1	2L:10321809;10421268
24135	Df(2L)ED8142	31E1;32A4	2L:10413461;10684679
9642	Df(2L)BSC214	31F5;32B4	2L:10516675;10913342
9641	Df(2L)BSC213	32B1;32C1	2L:10809118;11001451
9505	Df(2L)BSC145	32C1;32C1	2L:10967405;11001945--11001966
9716	Df(2L)BSC241	32C1;32F2	2L:11006679;11445740
9718	Df(2L)BSC244	32F2;33B6	2L:11445733;12002748

24109	Df(2L)ED761	33A2;33E5	2L:11808835;12436439
8907	Df(2L)ED775	33B8;34A3	2L:12010010;12975028
23662	Df(2L)BSC277	34A1;34B2	2L:12832803;13257491
32256	Df(2L)BSC892	34A5;34B9	2L:13060152;13382714
9594	Df(2L)BSC159	34B4;34C4	2L:13290762;13536086
27383	Df(2L)BSC812	34B11;34E1	2L:13421556;13878659
23152	Df(2L)BSC252	34D1;34F1	2L:13819589;14013641
9061	Df(2L)ED793	34E4;35B4	2L:13934848;14689337
6963	Df(2L)ED3	35B2;35D1	2L:14490575;15333760
8946	Df(2L)ED1050	35B8;35D4	2L:14997553;15762784
1491	Df(2L)r10	35D1;36A6--7	
26542	Df(2L)BSC690	35D4;35D4	2L:15745455;15821840
7521	Df(2L)Exel6038	35D6;35E2	2L:15912343;16042754
23663	Df(2L)BSC278	35E1;35F1	2L:16025369;16289284
27353	Df(2L)BSC781	35F1;36A1	2L:16325113;16417726
24113	Df(2L)ED1102	35F12;36A10	2L:16350236;16684883
3180	Df(2L)H20	36A8--9;36E1--2	
24114	Df(2L)ED1161	36A10;36C9	2L:16685279;17473293
9507	Df(2L)BSC148	36C8;36E3	2L:17428248;17965422
7839	Df(2L)Exel7070	36E2;36E6	2L:17903087--17903187;18161791
23156	Df(2L)BSC256	36E3;36F2	2L:18061946;18320008
7840	Df(2L)Exel8038	36E5;36F5	2L:18123514;18455586
9508	Df(2L)BSC149	36F5;36F10	2L:18444727;18673286
8935	Df(2L)ED1203	36F7;37C5	2L:18617225;19158447
24116	Df(2L)ED1272	37C5;38A2	2L:19158440;19753324
8679	Df(2L)ED1303	37E5;38C6	2L:19517610;20382385
9269	Df(2L)ED1315	38B4;38F5	2L:20085397;20917519
9682	Df(2L)ED1378	38F1;39D2	2L:20823195;21397328
9266	Df(2L)ED1473	39B4;40A5	2L:21250892;21828548
9510	Df(2L)BSC151	40A5;40E5	2L:21828581;22139023
26782	Df(2L)lt109	h35;h35	2L:22817519;22885080
4959	Df(2L)C'	h35;h38L	

The Bloomington deficiency kit for chromosomal arm 2R

used for the lethality screening of the gonad formation mutants.

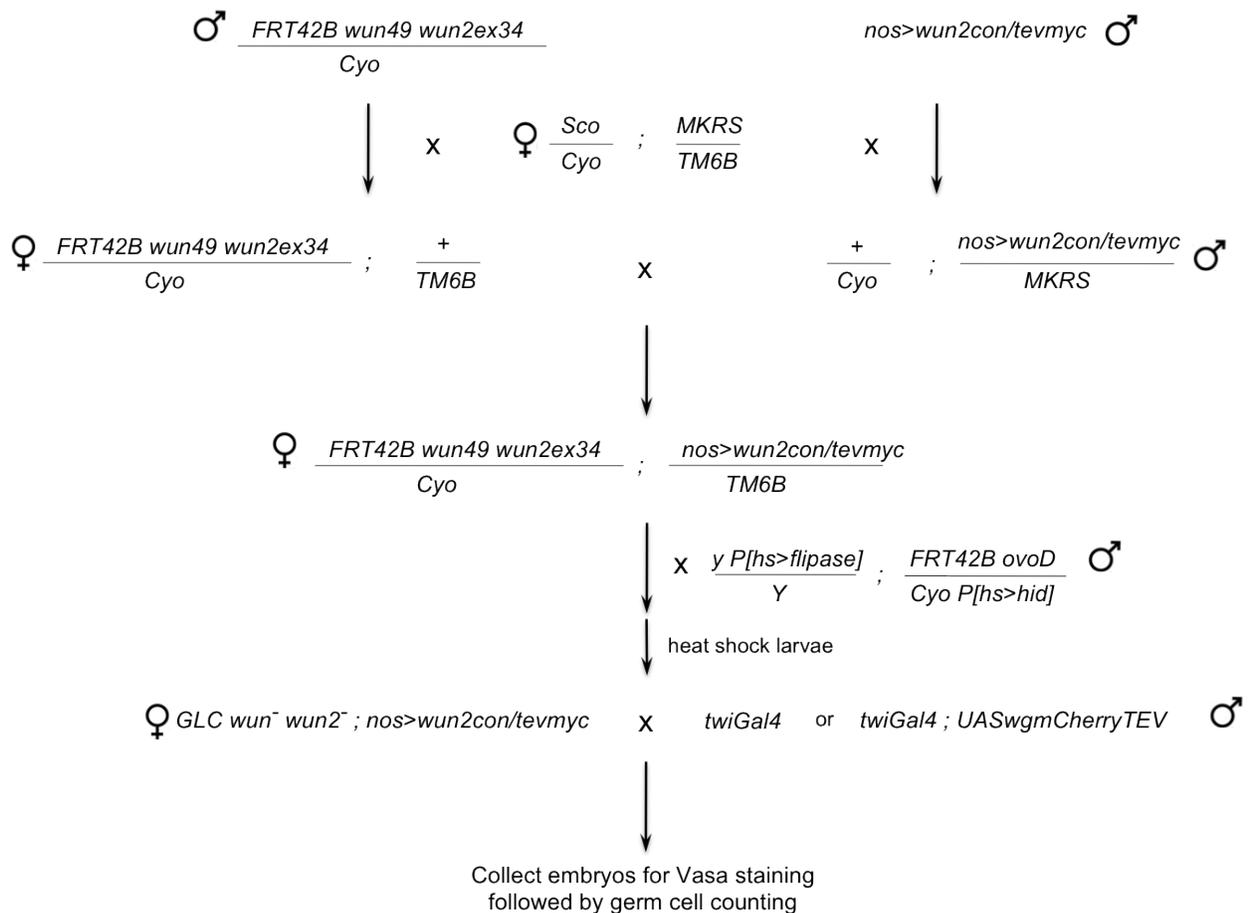
BL number	Deficiency	Cytological deletion coordinates	Sequence deletion coordinates
741	Df(2R)M41A10	41A;41A	
25705	Df(2R)BSC630	41D3;41F11	2R:1015322;1669918
32253	Df(2R)BSC889	41F11;42A13	2R:1669744-- 1669934;2069250
9683	Df(2R)ED1484	42A2;42A14	2R:1781142;2132933
8045	Df(2R)ED1612	42A13;42E6	2R:2108037;2937177
9062	Df(2R)ED1673	42E1;43D3	2R:2873307;3421058
8931	Df(2R)ED1715	43A4;43F1	2R:3214456;3804428
8941	Df(2R)ED1725	43E4;44B5	2R:3501429;4043550
24335	Df(2R)BSC267	44A4;44C4	2R:3970399;4281766
9276	Df(2R)ED1742	44B8;44E3	2R:4061673;4611634
9157	Df(2R)ED1770	44D5;45B4	2R:4543134;5095046
9063	Df(2R)ED1791	44F7;45F1	2R:4810235;5440757
23665	Df(2R)BSC280	45C4;45F4	2R:5180164;5466121
9410	Df(2R)BSC132	45F6;46B4	2R:5482319-- 5482429;5748332
23682	Df(2R)BSC298	46B2;46C7	2R:5621779;5845625
1702	Df(2R)X1	46C;47A1	
9539	Df(2R)BSC152	46C1;46D6	2R:5762817;5918044
23686	Df(2R)BSC303	46E1;46F3	2R:5965798;6035409
23666	Df(2R)BSC281	46F1;47A9	2R:6012734;6350379
25428	Df(2R)BSC595	47A3;47F1	2R:6273472;7176083
8910	Df(2R)ED2219	47D6;48B6	2R:7084917;7552896
8912	Df(2R)ED2247	48A3;48D5	2R:7487611;7876225
9626	Df(2R)BSC199	48C5;48E4	2R:7779605;8059989
26551	Df(2R)BSC699	48D7;48E6	2R:7907386;8070095
24929	Df(2R)BSC425	48F1;49A1	2R:8146790;8268309
7543	Df(2R)Exel6061	48F1;49A6	2R:8149005;8324950

23688	Df(2R)BSC305	49A4;49A10	2R:8313105;8392638
30585	Df(2R)BSC880	49A9;49E1	2R:8384335;8781611-- 8781621
24989	Df(2R)BSC485	49B10;49E6	2R:8520257;8868687
442	Df(2R)CX1	49C1;50D3-5	2R:8587926-- 8622457;9950134--9978645
7544	Df(2R)Exel6062	49E6;49F1	2R:8868689;8922684
7871	Df(2R)Exel8057	49F1;49F10	2R:8922352;9106852-- 9106854
23169	Df(2R)BSC273	49F4;50A13	2R:9047084;9389655
23170	Df(2R)BSC274	50A7;50B4	2R:9317969;9480777
23690	Df(2R)BSC307	50B6;50C18	2R:9510513;9849106
24385	Df(2R)BSC361	50C3;50F1	2R:9648768-- 9648817;10179666
24407	Df(2R)BSC383	50C6;50D2	2R:9726984;9912384
7875	Df(2R)Exel7130	50D4;50E4	2R:9960585;10100288
7876	Df(2R)Exel7131	50E4;50F6	2R:10118170-- 10118172;10247930
8913	Df(2R)ED2354	50E6;51B1	2R:10146550;10462062
7749	Df(2R)Exel6284	51B1;51C2	2R:10462255;10653073-- 10653275
24933	Df(2R)BSC429	51C2;51D1	2R:10657714;10761429
25741	Df(2R)BSC651	51C5;51E2	2R:10740461;11022806
9064	Df(2R)ED2426	51E2;52B1	2R:11016313;11498329
8915	Df(2R)ED2457	52D11;52E7	2R:11887814;12017662
29661	Df(2R)ED2487	52E6;53C4	2R:12012439;12273917
25078	Df(2R)BSC550	53C1;53C6	2R:12224286;12382237
7888	Df(2R)Exel7144	53C8;53D2	2R:12459438;12578579
7546	Df(2R)Exel6064	53C11;53D11	2R:12499138;12716579
9278	Df(2R)ED2747	53D11;53F8	2R:12716578;12984808
24356	Df(2R)BSC331	53D14;54A1	2R:12756835;13027428
7548	Df(2R)Exel6066	53F8;54B6	2R:12985074;13282147
9596	Df(2R)BSC161	54B2;54B17	2R:13192288;13372333
24379	Df(2R)BSC355	54B16;54C3	2R:13349852;13424178
7890	Df(2R)Exel7149	54C10;54D5	2R:13469001;13579292
24371	Df(2R)BSC347	54D2;54E9	2R:13522004;13680154
6780	Df(2R)14H10W- 35	54E5--7;55B5--7	
9066	Df(2R)ED3610	54F1;55C8	2R:13738410;14299538
8918	Df(2R)ED3683	55C2;56C4	2R:14176374;15116496
7551	Df(2R)Exel6069	56B5;56C11	2R:15029454;15213128

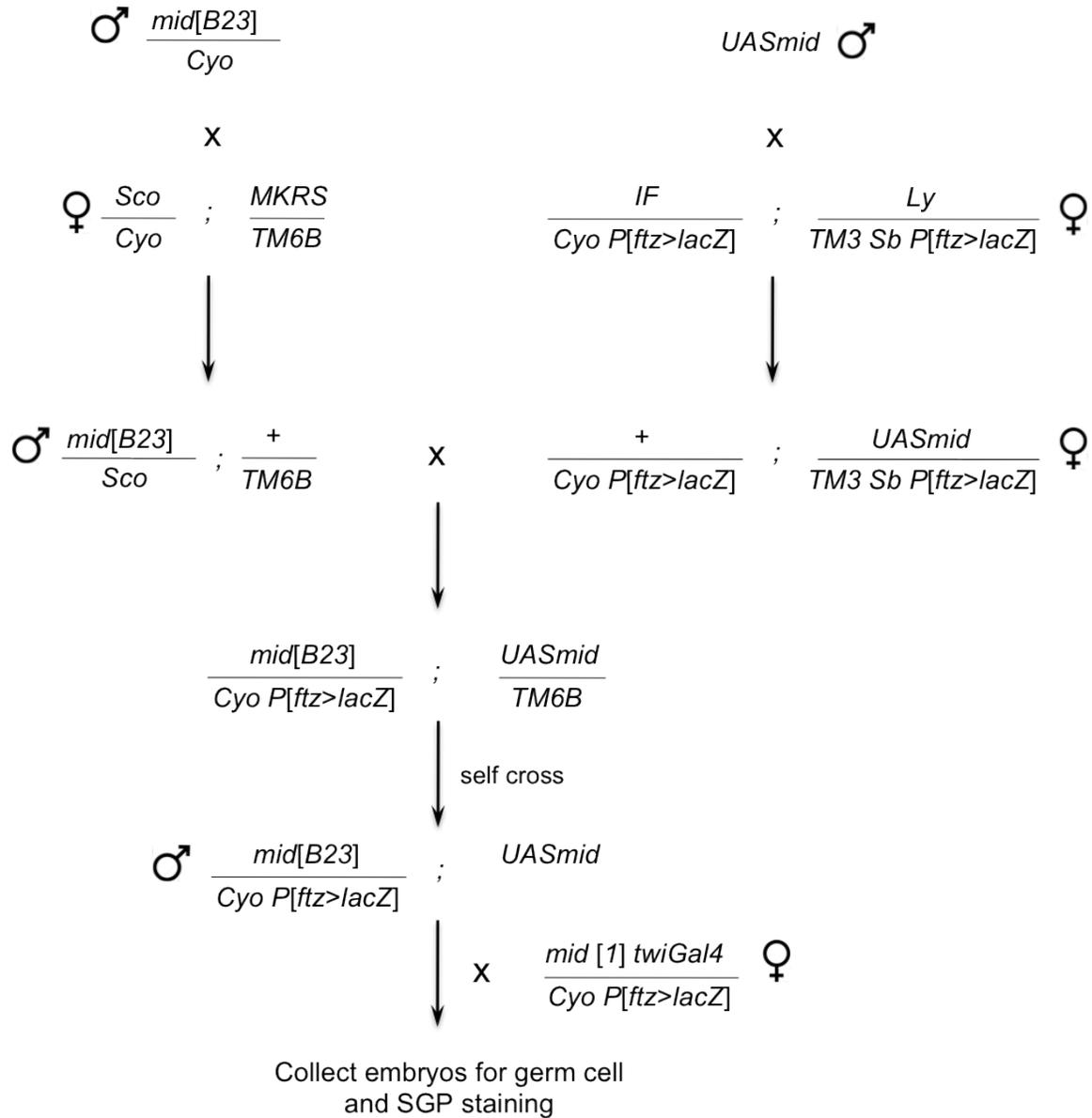
9423	Df(2R)BSC135	56C11;56D5	2R:15215243;15311082
27354	Df(2R)BSC782	56D8;56D14	2R:15338532;15389309
9067	Df(2R)ED3728	56D10;56E2	2R:15349955;15614252
30588	Df(2R)BSC883	56E1;56F11	2R:15519525;16130288
7896	Df(2R)Exel7162	56F11;56F16	2R:16132691-- 16132995;16201140
6609	Df(2R)BSC19	56F12--14;57A4	
26553	Df(2R)BSC701	56F15;57A9	2R:16166339;16554778
26554	Df(2R)BSC702	57A2;57B3	2R:16311622;16758360
30590	Df(2R)BSC885	57D2;57D10	2R:17102728;17189303
27582	Df(2R)BSC821	57D10;57E6	2R:17189303;17384714
26516	Df(2R)BSC664	57D12;58A3	2R:17229152;17759533
25430	Df(2R)BSC597	58A2;58F1	2R:17693855;18480501
282	Df(2R)X58-12	58D1--2;59A	
25431	Df(2R)BSC598	58F3;59A1	2R:18529284;18566186
27359	Df(2R)BSC787	58F4;59B1	2R:18549332;18685826
29988	Df(2R)BSC865	59A4;59B7	2R:18652440;18828551
25432	Df(2R)BSC599	59B1;59B3	2R:18685826;18741702
27356	Df(2R)BSC784	59B4;59B6	2R:18774154;18824283
26866	Df(2R)BSC769	59B7;59D9	2R:18828454;19272990
26513	Df(2R)BSC661	59D8;59F5	2R:19241797;19546662
9424	Df(2R)BSC136	59F5;60B6	2R:19554482;19944987
24380	Df(2R)BSC356	60B8;60C4	2R:19955762;20145427
27352	Df(2R)BSC780	60C2;60D14	2R:20090739;20572714
25437	Df(2R)BSC604	60D4;60E11	2R:20419696;20855955
2471	Df(2R)M60E	60E2--3;60E11--12	
25441	Df(2R)BSC608	60E11;60F2	2R:20864281;20939897
4961	Df(2R)Kr10	60F1;60F5	

Appendix B

- Fly cross schematic for rescue of *wunen* loss function in germ cells by *nos>wun2con/tev* in the presence of the TEV protease



➤ **Fly cross schematic for rescue of *mid* mutant gonad phenotype with Mid expression**



Bibliography

Ara, T., Nakamura, Y., Egawa, T., Sugiyama, T., Abe, K., Kishimoto, T., Matsui, Y. and Nagasawa, T. (2003) 'Impaired colonization of the gonads by primordial germ cells in mice lacking a chemokine, stromal cell-derived factor-1 (SDF-1)', *Proc Natl Acad Sci U S A* 100(9): 5319-23.

Ashkenazy, H., Erez, E., Martz, E., Pupko, T. and Ben-Tal, N. (2010) 'ConSurf 2010: calculating evolutionary conservation in sequence and structure of proteins and nucleic acids', *Nucleic Acids Res* 38(Web Server issue): W529-33.

Azpiazu, N. and Frasch, M. (1993) 'tinman and bagpipe: two homeo box genes that determine cell fates in the dorsal mesoderm of *Drosophila*', *Genes & Development* 7(7b): 1325-1340.

Barbosa, V., Kimm, N. and Lehmann, R. (2007) 'A maternal screen for genes regulating *Drosophila* oocyte polarity uncovers new steps in meiotic progression', *Genetics* 176(4): 1967-77.

Barila, D., Plateroti, M., Nobili, F., Muda, A. O., Xie, Y. H., Morimoto, T. and Perozzi, G. (1996) 'The Dri 42 gene, whose expression is up-regulated during epithelial differentiation, encodes a novel endoplasmic reticulum resident transmembrane protein', *Journal of Biological Chemistry* 271(47): 29928-29936.

Barzik, M., Kotova, T. I., Higgs, H. N., Hazelwood, L., Hanein, D., Gertler, F. B. and Schafer, D. A. (2005) 'Ena/VASP proteins enhance actin polymerization in the presence of barbed end capping proteins', *J Biol Chem* 280(31): 28653-62.

Bernstein, D. P., Cohen, P., Skodol, A., Bezirgianian, S. and Brook, J. S. (1996) 'Childhood antecedents of adolescent personality disorders', *The American journal of psychiatry* 153(7): 907-13.

Bodmer, R. (1993) 'The Gene Tinman Is Required for Specification of the Heart and Visceral Muscles in *Drosophila*', *Development* 118(3): 719-729.

Boyle, M., Bonini, N. and DiNardo, S. (1997) 'Expression and function of clift in the development of somatic gonadal precursors within the *Drosophila* mesoderm', *Development (Cambridge, England)* 124(5): 971-982.

Boyle, M. and DiNardo, S. (1995) 'Specification, migration and assembly of the somatic cells of the *Drosophila* gonad', *Development* 121(6): 1815-25.

Broihier, H., Moore, L., Van Doren, M., Newman, S. and Lehmann, R. (1998) 'zfh-1 is required for germ cell migration and gonadal mesoderm development in *Drosophila*', *Development (Cambridge, England)* 125(4): 655-666.

Brook, W. J. and Cohen, S. M. (1996) 'Antagonistic interactions between wingless and decapentaplegic responsible for dorsal-ventral pattern in the *Drosophila* Leg', *Science* 273(5280): 1373-7.

- Brook, W. J., Diaz-Benjumea, F. J. and Cohen, S. M. (1996) 'Organizing spatial pattern in limb development', *Annual review of cell and developmental biology* 12: 161-80.
- Brookes, S. M., Cohen, S. B., Price, E. J., Webb, L. M., Feldmann, M., Maini, R. N. and Venables, P. J. (1996) 'T cell clones from a Sjogren's syndrome salivary gland biopsy produce high levels of IL-10', *Clinical and experimental immunology* 103(2): 268-72.
- Brookman, J. J., Toosy, A. T., Shashidhara, L. S. and White, R. A. H. (1992b) 'The 412 Retrotransposon and the Development of Gonadal Mesoderm in *Drosophila*', *Development* 116(4): 1185-1192.
- Brose, K., Bland, K. S., Wang, K. H., Arnott, D., Henzel, W., Goodman, C. S., Tessier-Lavigne, M. and Kidd, T. (1999) 'Slit proteins bind robo receptors and have an evolutionarily conserved role in repulsive axon guidance', *Cell* 96(6): 795-806.
- Buescher, M., Svendsen, P. C., Tio, M., Miskolczi-McCallum, C., Tear, G., Brook, W. J. and Chia, W. (2004a) '*Drosophila* T box proteins break the symmetry of hedgehog-dependent activation of wingless', *Current biology : CB* 14(19): 1694-702.
- Buescher, M., Tio, M., Tear, G., Overton, P., Brook, W. and Chia, W. (2006) 'Functions of the segment polarity genes midline and H15 in *Drosophila melanogaster* neurogenesis', *Developmental Biology* 292(2): 418-429.
- Burnett, C. and Howard, K. (2003) 'Fly and mammalian lipid phosphate phosphatase isoforms differ in activity both in vitro and in vivo', *EMBO Rep* 4(8): 793-9.
- Burnett, C., Makridou, P., Hewlett, L. and Howard, K. (2004) 'Lipid phosphate phosphatases dimerise, but this interaction is not required for in vivo activity', *BMC Biochem* 5: 2.
- Crowner, D., Madden, K., Goetze, S. and Giniger, E. (2002) 'Lola regulates midline crossing of CNS axons in *Drosophila*', *Development* 129(6): 1317-25.
- Davies, E., Lim, J., Joo, W., Tam, C. and Fuller, M. (2013) 'The transcriptional regulator lola is required for stem cell maintenance and germ cell differentiation in the *Drosophila* testis', *Developmental Biology* 373(2): 310-321.
- DeFalco, T., Camara, N., Le Bras, S. and Van Doren, M. (2008) 'Nonautonomous sex determination controls sexually dimorphic development of the *Drosophila* gonad', *Dev Cell* 14(2): 275-86.
- DeFalco, T. J., Verney, G., Jenkins, A. B., McCaffery, J. M., Russell, S. and Van Doren, M. (2003) 'Sex-specific apoptosis regulates sexual dimorphism in the *Drosophila* embryonic gonad', *Dev Cell* 5(2): 205-16.
- Delorenzi, M. and Bienz, M. (1990) 'Expression of Abdominal-B homeoproteins in *Drosophila* embryos', *Development* 108(2): 323-9.
- Ding, J., Jiang, D., Kurczyk, M., Nalepka, J., Dudley, B., Merkel, E. I., Porter, F. D., Ewing, A. G., Winograd, N., Burgess, J. et al. (2008) 'Inhibition of HMG CoA reductase reveals an unexpected role for cholesterol during PGC migration in the mouse', *BMC Dev Biol* 8: 120.
- Doitsidou, M., Reichman-Fried, M., Stebler, J., Kopranner, M., Dorries, J., Meyer, D., Esguerra, C. V., Leung, T. and Raz, E. (2002) 'Guidance of primordial germ cell migration by the chemokine SDF-1', *Cell* 111(5): 647-59.

- Escalante-Alcalde, D., Hernandez, L., Le Stunff, H., Maeda, R., Lee, H. S., Jr Gang, C., Sciorra, V. A., Daar, I., Spiegel, S., Morris, A. J. et al. (2003) 'The lipid phosphatase LPP3 regulates extra-embryonic vasculogenesis and axis patterning', *Development* 130(19): 4623-37.
- Fairall, L., Schwabe, J. W., Chapman, L., Finch, J. T. and Rhodes, D. (1993) 'The crystal structure of a two zinc-finger peptide reveals an extension to the rules for zinc-finger/DNA recognition', *Nature* 366(6454): 483-7.
- Formaz-Preston, A., Ryu, J. R., Svendsen, P. C. and Brook, W. J. (2012) 'The Tbx20 homolog Midline represses wingless in conjunction with Groucho during the maintenance of segment polarity', *Developmental Biology* 369(2): 319-29.
- Gates, M., Kannan, R. and Giniger, E. (2011) 'A genome-wide analysis reveals that the Drosophila transcription factor Lola promotes axon growth in part by suppressing expression of the actin nucleation factor Spire', *Neural Development* 6: 37.
- Giniger, E., Tietje, K., Jan, L. and Jan, Y. (1994) 'lola encodes a putative transcription factor required for axon growth and guidance in Drosophila', *Development (Cambridge, England)* 120(6): 1385-1398.
- Goeke, S., Greene, E., Grant, P., Gates, M., Crowner, D., Aigaki, T. and Giniger, E. (2003) 'Alternative splicing of lola generates 19 transcription factors controlling axon guidance in Drosophila', *Nature neuroscience* 6(9): 917-924.
- Goldstein, J. L. and Brown, M. S. (1990) 'Regulation of the mevalonate pathway', *Nature* 343(6257): 425-30.
- Greig, S. and Akam, M. (1995) 'The role of homeotic genes in the specification of the Drosophila gonad', *Current biology : CB* 5(9): 1057-62.
- Griffin, K. J., Stoller, J., Gibson, M., Chen, S., Yelon, D., Stainier, D. Y. and Kimelman, D. (2000) 'A conserved role for H15-related T-box transcription factors in zebrafish and Drosophila heart formation', *Dev Biol* 218(2): 235-47.
- Hanyu-Nakamura, K., Kobayashi, S. and Nakamura, A. (2004) 'Germ cell-autonomous Wunen2 is required for germline development in Drosophila embryos', *Development* 131(18): 4545-53.
- Harder, B., Schomburg, A., Pflanz, R., Kustner, K., Gerlach, N. and Schuh, R. (2008) 'TEV protease-mediated cleavage in Drosophila as a tool to analyze protein functions in living organisms', *Biotechniques* 44(6): 765-72.
- Jaglarz, M. K. and Howard, K. R. (1995) 'The active migration of Drosophila primordial germ cells', *Development* 121(11): 3495-503.
- Jemc, J. (2011) 'Somatic gonadal cells: the supporting cast for the germline', *Genesis (New York, N.Y. : 2000)* 49(10): 753-775.
- Jemc, J., Milutinovich, A., Weyers, J., Takeda, Y. and Van Doren, M. (2012) 'raw Functions through JNK signaling and cadherin-based adhesion to regulate Drosophila gonad morphogenesis', *Developmental Biology* 367(2): 114-125.
- Jenkins, A. B., McCaffery, J. M. and Van Doren, M. (2003) 'Drosophila E-cadherin is essential for proper germ cell-soma interaction during gonad morphogenesis', *Development* 130(18): 4417-4426.

- Karch, F., Bender, W. and Weiffenbach, B. (1990) 'abdA expression in Drosophila embryos', *Genes Dev* 4(9): 1573-87.
- King, R. C. (1970) 'The meiotic behavior of the Drosophila oocyte', *Int Rev Cytol* 28: 125-68.
- Kispert, A. and Herrmann, B. G. (1993) 'The Brachyury gene encodes a novel DNA binding protein', *The EMBO journal* 12(8): 3211-20.
- Kitadate, Y. and Kobayashi, S. (2010) 'Notch and Egfr signaling act antagonistically to regulate germline stem cell niche formation in Drosophila male embryonic gonads', *Proceedings of the National Academy of Sciences of the United States of America* 107(32): 14241-6.
- Knaut, H., Werz, C., Geisler, R. and Nusslein-Volhard, C. (2003) 'A zebrafish homologue of the chemokine receptor Cxcr4 is a germ-cell guidance receptor', *Nature* 421(6920): 279-82.
- Kolodziej, P., Timpe, L., Mitchell, K., Fried, S., Goodman, C., Jan, L. and Jan, Y. (1996) 'frazzled encodes a Drosophila member of the DCC immunoglobulin subfamily and is required for CNS and motor axon guidance', *Cell* 87(2): 197-204.
- Kunwar, P. S., Sano, H., Renault, A. D., Barbosa, V., Fuse, N. and Lehmann, R. (2008) 'Tre1 GPCR initiates germ cell transepithelial migration by regulating Drosophila melanogaster E-cadherin', *Journal of Cell Biology* 183(1): 157-68.
- Kunwar, P. S., Starz-Gaiano, M., Bainton, R. J., Heberlein, U. and Lehmann, R. (2003) 'Tre1, a G protein-coupled receptor, directs transepithelial migration of Drosophila germ cells', *PLoS Biol* 1(3): E80.
- Le Bras, S. and Van Doren, M. (2006) 'Development of the male germline stem cell niche in Drosophila', *Dev Biol* 294(1): 92-103.
- Leal, S. M., Qian, L., Lacin, H., Bodmer, R. and Skeath, J. B. (2009) 'Neuromancer1 and Neuromancer2 regulate cell fate specification in the developing embryonic CNS of Drosophila melanogaster', *Developmental Biology* 325(1): 138-50.
- Lecuit, T., Brook, W. J., Ng, M., Calleja, M., Sun, H. and Cohen, S. M. (1996) 'Two distinct mechanisms for long-range patterning by Decapentaplegic in the Drosophila wing', *Nature* 381(6581): 387-93.
- Lecuyer, E., Necakov, A. S., Caceres, L. and Krause, H. M. (2008) 'High-resolution fluorescent in situ hybridization of Drosophila embryos and tissues', *CSH protocols* 2008: pdb prot5019.
- Li, M., Alls, J., Avancini, R., Koo, K. and Godt, D. (2003) 'The large Maf factor Traffic Jam controls gonad morphogenesis in Drosophila', *Nature cell biology* 5(11): 994-1000.
- Liu, Q.-X., Hiramoto, M., Ueda, H., Gojobori, T., Hiromi, Y. and Hirose, S. (2009) 'Midline governs axon pathfinding by coordinating expression of two major guidance systems', *Genes & Development* 23(10): 1165-1170.
- Madden, K., Crowner, D. and Giniger, E. (1999) 'LOLA has the properties of a master regulator of axon-target interaction for SNb motor axons of Drosophila', *Developmental Biology* 213(2): 301-313.
- Mahowald, A. P., Illmensee, K. and Turner, F. R. (1976) 'Interspecific Transplantation of Polar Plasm between Drosophila Embryos', *Journal of Cell Biology* 70(2): 358-373.

- Mainieri, D., Summermatter, S., Seydoux, J., Montani, J. P., Rusconi, S., Russell, A. P., Boss, O., Buchala, A. J. and Dullo, A. G. (2006) 'A role for skeletal muscle stearoyl-CoA desaturase 1 in control of thermogenesis', *FASEB J* 20(10): 1751-3.
- Mathews, W., Ong, D., Milutinovich, A. and Van Doren, M. (2006) 'Zinc transport activity of Fear of Intimacy is essential for proper gonad morphogenesis and DE-cadherin expression', *Development (Cambridge, England)* 133(6): 1143-1153.
- Miskolczi-McCallum, C., Scavetta, R., Svendsen, P., Soanes, K. and Brook, W. (2005) 'The *Drosophila melanogaster* T-box genes *midline* and *H15* are conserved regulators of heart development', *Developmental Biology* 278(2): 459-472.
- Molyneaux, K. A., Zinszner, H., Kunwar, P. S., Schaible, K., Stebler, J., Sunshine, M. J., O'Brien, W., Raz, E., Littman, D., Wylie, C. et al. (2003) 'The chemokine SDF1/CXCL12 and its receptor CXCR4 regulate mouse germ cell migration and survival', *Development* 130(18): 4279-86.
- Moore, L., Broihier, H., Van Doren, M., Lunsford, L. and Lehmann, R. (1998a) 'Identification of genes controlling germ cell migration and embryonic gonad formation in *Drosophila*', *Development (Cambridge, England)* 125(4): 667-678.
- Moore, L. A., Broihier, H. T., Van Doren, M. and Lehmann, R. (1998b) 'Gonadal mesoderm and fat body initially follow a common developmental path in *Drosophila*', *Development* 125(5): 837-44.
- Mulligan, T., Blaser, H., Raz, E. and Farber, S. A. (2010) 'Prenylation-deficient G protein gamma subunits disrupt GPCR signaling in the zebrafish', *Cell Signal* 22(2): 221-33.
- Najand, N., Ryu, J.-R. and Brook, W. (2012) 'In Vitro Site Selection of a Consensus Binding Site for the *Drosophila melanogaster* Tbx20 Homolog *Midline*', *PLoS ONE* 7(10).
- Nanda, S., DeFalco, T. J., Loh, S. H. Y., Phochanukul, N., Camara, N., Van Doren, M. and Russell, S. (2009) 'Sox100B, a *Drosophila* Group E Sox-domain Gene, Is Required for Somatic Testis Differentiation', *Sexual Development* 3(1): 26-37.
- Nüsslein-Volhard, C., Wieschaus, E. and Kluding, H. (1984) 'Mutations affecting the pattern of the larval cuticle in *Drosophila melanogaster*', *Roux's Archives of Developmental Biology* 193: 267-282.
- Nüsslein-Volhard, C., Wieschaus, E. and Kluding, H. (1984) 'Mutations Affecting the Pattern of the Larval Cuticle in *Drosophila-Melanogaster*. 1. Zygotic Loci on the 2nd Chromosome', *Wilhelm Roux Archives of Developmental Biology* 193(5): 267-282.
- Ohsako, T., Horiuchi, T., Matsuo, T., Komaya, S. and Aigaki, T. (2003) '*Drosophila lola* encodes a family of BTB-transcription regulators with highly variable C-terminal domains containing zinc finger motifs', *Gene* 311: 59-69.
- Ossowski, S., Schneeberger, K., Clark, R. M., Lanz, C., Warthmann, N. and Weigel, D. (2008) 'Sequencing of natural strains of *Arabidopsis thaliana* with short reads', *Genome Res* 18(12): 2024-33.
- Paige Bass, B., Cullen, K. and McCall, K. (2007) 'The axon guidance gene *lola* is required for programmed cell death in the *Drosophila* ovary', *Developmental Biology* 304(2): 771-785.
- Pauli, A., Althoff, F., Oliveira, R. A., Heidmann, S., Schuldiner, O., Lehner, C. F., Dickson, B. J. and Nasmyth, K. (2008) 'Cell-type-specific TEV protease cleavage reveals cohesin functions in *Drosophila* neurons', *Dev Cell* 14(2): 239-51.

- Pilquil, C., Dewald, J., Cherney, A., Gorshkova, I., Tigyi, G., English, D., Natarajan, V. and Brindley, D. N. (2006) 'Lipid phosphate phosphatase-1 regulates lysophosphatidate-induced fibroblast migration by controlling phospholipase D2-dependent phosphatidate generation', *J Biol Chem* 281(50): 38418-29.
- Qian, L., Liu, J. and Bodmer, R. (2005) 'Neuromancer Tbx20-related genes (H15/midline) promote cell fate specification and morphogenesis of the Drosophila heart', *Developmental Biology* 279(2): 509-524.
- Redline, R. W., Williams, A. J., Patterson, P. and Collins, T. (1992) 'Human HOX4E: a gene strongly expressed in the adult male and female urogenital tracts', *Genomics* 13(2): 425-30.
- Reim, I. and Frasch, M. (2005) 'The Dorsocross T-box genes are key components of the regulatory network controlling early cardiogenesis in Drosophila', *Development* 132(22): 4911-25.
- Reim, I., Mohler, J. and Frasch, M. (2005) 'Tbx20-related genes, mid and H15, are required for tinman expression, proper patterning, and normal differentiation of cardioblasts in Drosophila', *Mechanisms of development* 122(9): 1056-1069.
- Renault, A., Kunwar, P. and Lehmann, R. (2010) 'Lipid phosphate phosphatase activity regulates dispersal and bilateral sorting of embryonic germ cells in Drosophila', *Development* 137(11): 1815-1823.
- Renault, A., Sigal, Y., Morris, A. and Lehmann, R. (2004) 'Soma-germ line competition for lipid phosphate uptake regulates germ cell migration and survival', *Science (New York, N.Y.)* 305(5692): 1963-1966.
- Renault, A., Starz-Gaiano, M. and Lehmann, R. (2002) 'Metabolism of sphingosine 1-phosphate and lysophosphatidic acid: a genome wide analysis of gene expression in Drosophila', *Mechanisms of development* 119 Suppl 1: 301.
- Ricardo, S. and Lehmann, R. (2009) 'An ABC transporter controls export of a Drosophila germ cell attractant', *Science* 323(5916): 943-6.
- Richardson, B. and Lehmann, R. (2010) 'Mechanisms guiding primordial germ cell migration: strategies from different organisms', *Nature reviews. Molecular cell biology* 11(1): 37-49.
- Riechmann, V., Rehorn, K. P., Reuter, R. and Leptin, M. (1998) 'The genetic control of the distinction between fat body and gonadal mesoderm in Drosophila', *Development* 125(4): 713-23.
- Roberts, R., Sciorra, V. A. and Morris, A. J. (1998) 'Human type 2 phosphatidic acid phosphohydrolases. Substrate specificity of the type 2a, 2b, and 2c enzymes and cell surface activity of the 2a isoform', *J Biol Chem* 273(34): 22059-67.
- Russ, A. P., Wattler, S., Colledge, W. H., Aparicio, S. A. J. R., Carlton, M. B. L., Pearce, J. J., Barton, S. C., Surani, M. A., Ryan, K., Nehls, M. C. et al. (2000) 'Eomesodermin is required for mouse trophoblast development and mesoderm formation', *Nature* 404(6773): 95-99.
- Ryu, J.-R., Najand, N. and Brook, W. (2011) 'Tinman is a direct activator of midline in the Drosophila dorsal vessel', *Developmental dynamics : an official publication of the American Association of Anatomists* 240(1): 86-95.
- Sano, H., Kunwar, P. S., Renault, A. D., Barbosa, V., Clark, I. B., Ishihara, S., Sugimura, K. and Lehmann, R. (2012) 'The Drosophila Actin Regulator ENABLED Regulates Cell Shape and Orientation during Gonad Morphogenesis', *PLoS ONE* 7(12): e52649.

- Sano, H., Renault, A. and Lehmann, R. (2005) 'Control of lateral migration and germ cell elimination by the *Drosophila melanogaster* lipid phosphate phosphatases Wunen and Wunen 2', *The Journal of cell biology* 171(4): 675-683.
- Santos, A. C. and Lehmann, R. (2004) 'Isoprenoids control germ cell migration downstream of HMGCoA reductase', *Dev Cell* 6(2): 283-93.
- Seifert, J. and Lehmann, R. (2012) 'Drosophila primordial germ cell migration requires epithelial remodeling of the endoderm', *Development (Cambridge, England)* 139(12): 2101-2106.
- Seydoux, G. and Braun, R. (2006) 'Pathway to totipotency: lessons from germ cells', *Cell* 127(5): 891-904.
- Shaknovich, R., Yeyati, P. L., Ivins, S., Melnick, A., Lempert, C., Waxman, S., Zelent, A. and Licht, J. D. (1998) 'The promyelocytic leukemia zinc finger protein affects myeloid cell growth, differentiation, and apoptosis', *Mol Cell Biol* 18(9): 5533-45.
- Shigenobu, S., Kitadate, Y., Noda, C. and Kobayashi, S. (2006) 'Molecular characterization of embryonic gonads by gene expression profiling in *Drosophila melanogaster*', *Proc Natl Acad Sci U S A* 103(37): 13728-33.
- Sigal, Y. J., McDermott, M. I. and Morris, A. J. (2005) 'Integral membrane lipid phosphatases/phosphotransferases: common structure and diverse functions', *Biochem J* 387(Pt 2): 281-93.
- Simon, M. F., Rey, A., Castan-Laurel, I., Gres, S., Sibrac, D., Valet, P. and Saulnier-Blache, J. S. (2002) 'Expression of ectolipid phosphate phosphohydrolases in 3T3F442A preadipocytes and adipocytes. Involvement in the control of lysophosphatidic acid production', *J Biol Chem* 277(26): 23131-6.
- Smyth, S., Sciorra, V., Sigal, Y., Pamuklar, Z., Wang, Z., Xu, Y., Prestwich, G. and Morris, A. (2003) 'Lipid phosphate phosphatases regulate lysophosphatidic acid production and signaling in platelets: studies using chemical inhibitors of lipid phosphate phosphatase activity', *The Journal of biological chemistry* 278(44): 43214-43223.
- Spletter, M., Liu, J., Liu, J., Su, H., Giniger, E., Komiyama, T., Quake, S. and Luo, L. (2007) 'Lola regulates *Drosophila* olfactory projection neuron identity and targeting specificity', *Neural Development* 2(1): 14.
- Starz-Gaiano, M., Cho, N. K., Forbes, A. and Lehmann, R. (2001) 'Spatially restricted activity of a *Drosophila* lipid phosphatase guides migrating germ cells', *Development* 128(6): 983-91.
- Starz-Gaiano, M. and Lehmann, R. (2001) 'Moving towards the next generation', *Mech Dev* 105(1-2): 5-18.
- Stennard, F., Costa, M., Elliott, D., Rankin, S., Haast, S., Lai, D., McDonald, L., Niederreither, K., Dolle, P., Bruneau, B. et al. (2003) 'Cardiac T-box factor Tbx20 directly interacts with Nkx2-5, GATA4, and GATA5 in regulation of gene expression in the developing heart', *Developmental Biology* 262(2): 206-224.
- Stukey, J. and Carman, G. M. (1997) 'Identification of a novel phosphatase sequence motif', *Protein Sci* 6(2): 469-72.
- Svendsen, P. C., Formaz-Preston, A., Leal, S. M. and Brook, W. J. (2009) 'The Tbx20 homologs midline and H15 specify ventral fate in the *Drosophila melanogaster* leg', *Development* 136(16): 2689-93.

- Tepass, U., GruszynskiDeFeo, E., Haag, T. A., Omatyar, L., Torok, T. and Hartenstein, V. (1996) 'Shotgun encodes Drosophila E-cadherin and is preferentially required during cell rearrangement in the neurectoderm and other morphogenetically active epithelia', *Genes & Development* 10(6): 672-685.
- Thorpe, J. L., Doitsidou, M., Ho, S. Y., Raz, E. and Farber, S. A. (2004) 'Germ cell migration in zebrafish is dependent on HMGCoA reductase activity and prenylation', *Developmental Cell* 6(2): 295-302.
- Tokuyasu, K. T. (1974) 'Dynamics of Spermiogenesis in Drosophila-Melanogaster .3. Relation between Axoneme and Mitochondrial Derivatives', *Experimental Cell Research* 84(1-2): 239-250.
- Tomsig, J. L., Snyder, A. H., Berdyshev, E. V., Skobeleva, A., Mataya, C., Natarajan, V., Brindley, D. N. and Lynch, K. R. (2009) 'Lipid phosphate phosphohydrolase type 1 (LPP1) degrades extracellular lysophosphatidic acid in vivo', *Biochem J* 419(3): 611-8.
- Van Doren, M., Broihier, H., Moore, L. and Lehmann, R. (1998a) 'HMG-CoA reductase guides migrating primordial germ cells', *Nature* 396(6710): 466-469.
- Van Doren, M., Mathews, W., Samuels, M., Moore, L., Broihier, H. and Lehmann, R. (2003) 'fear of intimacy encodes a novel transmembrane protein required for gonad morphogenesis in Drosophila', *Development (Cambridge, England)* 130(11): 2355-2364.
- Van Doren, M., Williamson, A. and Lehmann, R. (1998c) 'Regulation of zygotic gene expression in Drosophila primordial germ cells', *Current biology : CB* 8(4): 243-246.
- Waggoner, D. W., Gomez-Munoz, A., Dewald, J. and Brindley, D. N. (1996) 'Phosphatidate phosphohydrolase catalyzes the hydrolysis of ceramide 1-phosphate, lysophosphatidate, and sphingosine 1-phosphate', *J Biol Chem* 271(28): 16506-9.
- Warrior, R. (1994) 'Primordial germ cell migration and the assembly of the Drosophila embryonic gonad', *Developmental Biology*.
- Weyers, J., Milutinovich, A., Takeda, Y., Jemc, J. and Van Doren, M. (2011) 'A genetic screen for mutations affecting gonad formation in Drosophila reveals a role for the slit/robo pathway', *Developmental Biology* 353(2): 217-228.
- Yin, Z., Xu, X. L. and Frasch, M. (1997) 'Regulation of the twist target gene tinman by modular cis-regulatory elements during early mesoderm development', *Development* 124(24): 4971-82.
- Zhang, N., Sundberg, J. P. and Gridley, T. (2000) 'Mice mutant for Ppap2c, a homolog of the germ cell migration regulator wunen, are viable and fertile', *Genesis* 27(4): 137-40.
- Zhang, N., Zhang, J., Cheng, Y. and Howard, K. (1996) 'Identification and genetic analysis of wunen, a gene guiding Drosophila melanogaster germ cell migration', *Genetics* 143(3): 1231-41.
- Zhang, N., Zhang, J., Purcell, K. J., Cheng, Y. and Howard, K. (1997) 'The Drosophila protein Wunen repels migrating germ cells', *Nature* 385(6611): 64-7.
- Zheng, L. and Carthew, R. (2008) 'Lola regulates cell fate by antagonizing Notch induction in the Drosophila eye', *Mechanisms of development* 125(1-2): 18-29.