

AN INTERDISCIPLINARY APPROACH TO STUDY THE DYNAMIC PATTERNING OF
FOLLICLE CELLS

By

NICOLE THERESA REVAITIS

A dissertation submitted to the

Graduate School-Camden

Rutgers, The State University of New Jersey

In partial fulfillment of the requirements

For the degree of

Doctor of Philosophy

Graduate Program in Computational and Integrative Biology

Written under the direction of

Dr. Nir Yakoby

And approved by

Nir Yakoby

Benedetto Piccoli

Jongmin Nam

Gertrud Schüpbach

Camden, New Jersey

May 2019

ABSTRACT OF THE DISSERTATION

An Interdisciplinary Approach to Study the Dynamic Patterning of Follicle Cells

By:

NICOLE THERESA REVAITIS

Dissertation Director:

Dr. Nir Yakoby

Organogenesis requires the spatiotemporal coordination of numerous cell signaling pathways. While many pathways have been investigated, the mechanism for ligand dispersal and quantitative analysis of signaling induction is largely unknown. The epidermal growth factor receptor (EGFR) signaling pathway is activated multiple times in tissues throughout development. In oogenesis, the TGF- α -like ligand, Gurken (GRK), is secreted from around the oocyte nucleus where it activates the EGFR in the overlying follicle cells. In addition to the multiple components of the pathway, the compartments of the egg chamber are changing along with the position of GRK. To study how signaling is regulated in time and space, we developed a mathematical model to recapitulate the dynamics of EGFR signaling in oogenesis. While some parameters were acquired from the literature, others were obtained through genetic perturbations and quantitative analysis. For example, we used CRISPR/Cas9 to generate a homozygous viable EGFR tagged with GFP. Using this fly, we followed the dynamic localization of GRK and EGFR. Importantly, using ELISA, we quantified the number of EGFR molecules per cell during stages 8-14 of oogenesis and stages 2 and 5 of embryogenesis.

While patterning of the follicular epithelium has been extensively studied, information about their regulation is mostly unknown. To better understand this regulation, we took advantage of the FlyLight collection that contains over 7,000 intergenic and intronic DNA fragments that can potentially drive the transcription factor GAL4. Cross listing the 84 genes known to be expressed during oogenesis with the 1200 genes in the FlyLight collection found 22 common genes that are represented by 281 fly lines. Of these lines, 61 show expression patterns in the follicle cells when crossed to a UAS-GFP reporter. Of the 61 lines, 19 recapitulate the full or partial pattern of the endogenous gene pattern. Mapping the distribution of all 61 lines, we found a significant enrichment of enhancers in the first intron in comparison to the 5' proximal or distal regions of the gene model. Since all lines drive a GAL4 transcription factor, thus offering valuable resource for

genetic manipulations. Our screen provides further evidence that complex gene-patterns are regulated combinatorially by enhancers controlling expression in simple domains.

DEDICATION

This work is dedicated to my loving husband, Paul, and our two children, Ella and Michael.

Thank you for motivating me to keep working, stay focused, and make the best of life.

ACKNOWLEDGEMENTS

I am thankful to my advisor, Dr. Nir Yakoby, for without his mentoring and friendship; this work would not be possible. I appreciate his guidance and support in helping me to improve my experimental and analytical skills. Lastly, I admire his creativity and evolving perspectives that have allowed me to gain new insights in scientific understanding. I would like to thank my committee members, Dr. Trudi Schüpbach, Dr. Benedetto Piccoli, and Dr. Jongmin Nam. Their advice and support have been an invaluable resource to me as they have substantially improved the quality of my work, and for that, I am thankful!

A special thank you to Dr. Nastassia Pouradier-Duteil for her innovations to mathematical modeling and willingness to apply them to *Drosophila*. Her collaboration has been a wonderful research opportunity for me to learn more about applied mathematics. I would also like to thank Dr. Robert Marmion for his work generating the tagged EGFR and methodical support. I would like to thank Dr. Matthew Niepielko for his advice and support. I am thankful for the assistance from Dr. Eric Klein with automating the image analysis. I am also grateful for the friendship and many fruitful discussions with my lab mates: Dr. Vikrant Singh, Cody Stevens, Alessio Russomanno, Christopher Sottolano, Nicholas Gattone, and Abrar Abuthaqilah.

My research has greatly benefitted from the generosity of the *Drosophila* community. I would like to thank Dr. Stanislav Shvartsman for graciously being a reliable resource of reagents and fly stocks. I would also like to thank Dr. M. Gonzalez-Gaitan, Dr. C. Berg, Dr. D. Vasilias, Dr. A. Michelson, and Dr. T. Schüpbach, for providing fly stocks and antibodies. I am also thankful for Primmibiotec, the VDRC, DSHB, Hybridoma Bank, and the UC *Drosophila* Stock Center for *Drosophila* fly species, stocks, and reagents. I appreciate Rainbow Transgenics for helping us to produce many transgenic fly stocks.

This work was supported by grants and I appreciate the support of the Center for Computational and Integrative Biology, The Rutgers Dean's Office, National Science Foundation (IOS-1149144, awarded to NY), and National Institute of Health R15 (GM-101597, awarded to NY).

Lastly, I would like to thank my friends and family for their emotional support and encouragement, which none of this would have been possible without.

TABLE OF CONTENTS

TITLE PAGE.....	i
ABSTRACT.....	ii
DEDICATION.....	iv
ACKNOWLEDGEMENTS.....	v
LIST OF FIGURES.....	x

CHAPTERS

1. Introduction.....	1
1.1 <i>Drosophila</i> oogenesis.....	2
1.2 The <i>Drosophila</i> eggshell.....	3
1.3 Signaling pathways.....	4
1.4 Cell signaling and eggshell diversity.....	6
2. Specific Aims.....	8
2.1 Develop a mathematical model that defines parameters and their values contributing to different distributions of GRK/EGFR activation.....	8
2.1.1 Mathematical reasoning.....	8
2.1.2 Building a dynamic model from experimental reasoning.....	9
2.1.3 Modelling the egg chamber growth and source.....	11
2.1.4 Determining parameters.....	13
2.1.5 Validating parameter selection.....	15
2.1.6 The effects of mechanistic changes.....	17
2.2 Quantitative analysis of the EGFR in the follicle cells over developmental stages.....	18
2.2.1 Quantification of the EGFR.....	19
2.2.2 A GFP tagged EGFR construct.....	19
2.2.3 EGFR quantification throughout oogenesis and embryogenesis.....	23
2.2.4 Intracellular dynamics of the EGFR.....	25
2.2.5 EGFR expression in imaginal discs.....	27
2.3 Screening for <i>cis</i> regulation of <i>Drosophila</i> genes.....	29

2.3.1	Screening for regulatory domains.....	29
2.3.2	The FlyLight lines are new resource tool for gene perturbations in oogenesis.....	36
2.3.3	Mapping the distribution of CRMs in the gene model.....	40
2.3.4	The FlyLight lines control expression in multiple tissues.....	44
3.	Conclusions.....	47
3.1	A comprehensive model to study parameters affecting the distribution of GRK/dpERK.....	47
3.2	A quantitative analysis of the EGFR in <i>Drosophila</i> oogenesis.....	48
3.3	Analysis of CRMs in <i>Drosophila</i> oogenesis.....	48
4.	Materials and Methods.....	51
4.1	Fly stocks and strains.....	51
4.2	Immunohistochemistry.....	52
4.3	EGFR protein quantification.....	53
4.4	Microscopy and co-localization analysis.....	54
4.5	RNA-seq analysis.....	55
4.6	Statistical analysis of DNA fragments distribution.....	56
4.7	Pattern annotation and matrix formation.....	56
4.8	Construction of the mathematical model.....	57
4.8.1	Diffusion of ligand in the perivitteline space.....	58
4.8.2	Reactions between ligand, receptors, complexes, inhibitors and signal.....	58
4.8.3	Ligand-Receptor Complex recycling.....	60
4.8.4	Movement of the oocyte nucleus.....	61
4.8.5	Growth of the egg chamber.....	61
4.8.6	Shift of the overlying follicle cells.....	62
4.8.7	Numerical simulations.....	63

4.8.8 Parameterization of the perivitteline space.....	64
4.8.9 Numerical discretization of the domain.....	64
4.8.10 Numerical approximation of the Laplace-Beltrani operator.....	66
4.8.11 Extrapolation of parameter values from intensity measurements.....	68
5. Supplemental Data.....	70
Supplement 1: Measurements of the egg chamber.....	70
Supplement 2: Modelling the PVS: Construction of the cubed spheroidal mesh.....	71
Supplement 3: Table of parameters.....	72
Supplement 4: Fitting parameters selection to intensity data.....	73
Supplement 5: EGFR sfGFP.....	74
6. Bibliography.....	75
7. Curriculum Vitae.....	85

LIST OF FIGURES

Figure 1: Cartoon of <i>Drosophila</i> oogenesis.....	3
Figure 2: <i>Drosophila</i> eggshells.....	4
Figure 3: The EGFR signaling pathway.....	6
Figure 4: Differences in EGFR signaling activation and eggshells among <i>Drosophila</i> species.....	7
Figure 5: Schematic of the mechanisms used for defining parameters of the model.....	10
Figure 6: Measurements of the source.....	13
Figure 7: Simulations of GRK/dpERK over developmental stages.....	14
Figure 8: Validating parameter selections using genetic perturbations.....	15
Figure 9: Exhibiting the mechanistic changes.....	18
Figure 10: Using CRISPR/Cas9 to generate an EGFR tagged with GFP.....	20
Figure 11: Egg chambers show co-localization of EGFR GFP and GRK.....	22
Figure 12: Cartoon of GRK locus targeted by CRISPR/Cas9.....	23
Figure 13: Quantification of EGFR during oogenesis and embryogenesis.....	25
Figure 14: Intracellular localization of the EGFR.....	26
Figure 15: Expression of the EGFR protein detected by GFP in imaginal discs.....	28
Figure 16: Screening for expression domains during oogenesis.....	30
Figure 17: Expression domains of several FlyLight lines.....	32
Figure 17A: Expression domains <i>daughters against dpp</i> (<i>dad</i>).....	32
Figure 17B: Expression domains of <i>broad</i> (<i>br</i>).....	33
Figure 17C: Expression domains of <i>pointed</i> (<i>pnt</i>).....	34
Figure 17D: A binary matrix representing all gene expression patterns and FlyLight GFP positive lines	35
Figure 18: Genetic perturbations using <i>dad</i> ^{44C10} and <i>dpp</i> ^{18E05} FlyLight lines.....	38
Figure 19: Mapping distribution of FlyLight gene fragments.....	42

Figure 20: Spatial and temporal distribution of GFP-positive Flylight lines.....	45
Supplement 1: Dynamic measurements of the egg chamber from S7 to S10A.....	70
Supplement 2: Modelling the PVS: Construction of the cubed spheroidal mesh.....	71
Supplement 3: Table of parameters.....	72
Supplement 4: Fitting parameters selection to intensity data.....	73
Supplement 5: An EGFR tagged with sfGFP.....	74

Chapter 1:

Introduction

The epidermal growth factor receptor (EGFR) signaling pathway is a conserved developmental pathway controlling many cellular processes, including cell proliferation, differentiation, adhesion, and cell migration (Arteaga, 2002; Cavaliere et al., 2008; Dobens and Raftery, 2000). Malregulation of this pathway is associated with tissue pathology, including many types of cancers (Corkery et al., 2009; Matikas et al., 2015; Yewale et al., 2013). During *D. melanogaster*'s oogenesis, EGFR activation, by the EGF-like ligand Gurken (GRK), defines the anterior posterior and dorsal ventral axes of the fly (Gonzalez-Reyes et al., 1995; Neuman-Silberberg and Schupbach, 1993; Schupbach, 1987). GRK is secreted from near the oocyte nucleus to the perivitelline space and signals through EGFR in the overlaying follicle cells (FCs) (Neuman-Silberberg and Schupbach, 1993; Schupbach, 1987). *Drosophila* oogenesis is comprised of 14 morphologically distinct developmental stages of egg chambers, the precursor of the mature egg (Spradling, 1993). The position of the oocyte nucleus is dynamic, where it initially signals to the posterior follicle cells (PFCs) until stage 8. At stage 8, the PFCs signal to the oocyte causing the microtubules to reorganize and nucleus to anchor to the dorsal anterior corner of the oocyte, a process known as repolarization (Gonzalez-Reyes et al., 1995; Roth et al., 1995; Wittes and Schupbach, 2019). The position of the nucleus at S10A, and thus GRK, defines the dorsal side of the egg chamber and all subsequent stages in the *Drosophila* life cycle (Schupbach and Roth, 1994). The dynamic location of the oocyte nucleus sets the flies' axes by controlling GRK secretion.

At these levels of complexity, computational models are useful tools to study and understand the mechanism of GRK distribution and the dynamic activation of EGFR during oogenesis. In this model (Aim 1), we define several parameters that shape the distribution of GRK. While some parameters were already described in the literature, others were measured directly in this study. For example, we developed an EGFP tagged EGFR (Aim 2) to measure the amounts and dynamics of EGFR in the FCs to be directly applied to our model. The model was developed based on the

research done in *D. melanogaster*, however, the long term goal is to use this model to study the differences observed in the distribution of GRK and dpERK in other *Drosophila* species that lead to diversity of eggshell morphologies among species.

While the model is useful for understanding *trans*-acting elements that regulate eggshell formation, we also looked at *cis*-acting elements and how they pattern the FCs of the egg chamber to regulate gene expression (Aim 3). To better understand gene regulation, we utilized the FlyLight collection of drivers expressing the GAL4 transcription factor under the regulation of sequences from intronic or intergenic DNA origin. We cross-listed this list of over 7000 genes with the 83 genes known to be expressed in the FCs during oogenesis. We found that 53 of these constructs have expression patterns in the FCs, 18 of which reflected that of the endogenous gene.

1.1 *Drosophila* Oogenesis

Development occurs in all animals. This process originates from a single-celled egg that is fertilized and divides into a functional multicellular organism. To study the process of development, we use *Drosophila* oogenesis as a model system (Figure 1). Oogenesis occurs in the female fruit fly. Each female has two ovaries containing 14-16 ovarioles (King, 1970). The ovariole is a strand of developing egg chambers, the precursor to the mature egg. Each egg chamber undergoes fourteen morphologically defined stages throughout oogenesis (Ables, Jan. 1, 2015; Spradling, 1993). Since all developmental stages are present within the ovary at any given time, this is an ideal model system for studying changes that occur throughout the developmental process. The cells of the egg chamber are compartmentalized into different cell types. The nurse cells (NCs) at the anterior of the egg chamber supply the growing oocyte with cytoplasmic contents such as RNA and proteins throughout oogenesis in a process known as dumping (Horne-Badovinac and Bilder, 2005). The oocyte is surrounded by a monolayer of epithelial cells known as follicle cells (FCs) that secrete proteins essential for the future eggshell (Berg, 2005; Duhart et al., 2017). The entire process of oogenesis occurs within 3-5 days (Spradling, 1993). At the end of this process, the egg goes through the oviduct for fertilization (Krauchunas and Wolfner, 2013).

Figure 1: Cartoon of *Drosophila* oogenesis (Ables, 2015). A. Ovaries of a *Drosophila* female. B. Stages of oogenesis are defined by morphologies. Different cell types of the egg chamber are: oo, oocyte; nc, nurse cells; fc, follicle cells.

1.2 The *Drosophila* eggshell

The eggshell of *D. melanogaster* features several structures on the dorsal side of the egg. The most prominent of these are the two respiratory filaments known as dorsal appendages (Figure 2A). The size and number of dorsal appendages is different among *Drosophila* species, where some species like *D. guttifera* (Figure 2B) have three dorsal appendages and other species, like *D. cardini* (Figure 2D) have four. Other structures present on the eggshell are the operculum, where the larvae hatches from the egg; and the micropyle, where sperm enters the egg (Figure 2A). An additional structure, a lumen-like dorsal ridge (DR), is present on the eggshells of other species (Figure 2C-D); *D. melanogaster* eggshells lack a DR (Margaritis et al., 1983; Niepielko and Yakoby, 2014). The size and shape of the DR varies among species. In *D. willistoni*, the DR extends over 75% of the eggshell to the posterior (Figure 2C). In *D. cardini*, the DR extends over the entire dorsal side of the eggshell to the posterior end (Figure 2D).

Figure 2: *Drosophila* eggshells. A. Eggshell of *Drosophila melanogaster*. Structures labeled are the dorsal appendages (DAs), operculum (op), and micropyle. B. Eggshell of *D. guttifer* (third DA highlighted in yellow). C. Eggshell of *D. willistoni* (dorsal ridge highlighted in yellow). D. Eggshell of *D. cardini* (dorsal ridge highlighted in yellow).

1.3 Signaling Pathways

Morphogens are chemicals or proteins that change the morphological fate of the cell. Chemical gradients were initially described in a free diffusion model that was later redefined throughout tissues in the context of their position to a relative morphogen source (Turing, 1952; Wolpert, 1989). Many of the signaling pathways present in humans are conserved across animal species. *Drosophila* oogenesis is no exception as many of the pathways present throughout the developmental stages are apparent in humans. For this reason, *Drosophila* oogenesis is a useful system for understanding developmental processes.

Two major signaling pathways that cooperate to pattern the eggshell in *Drosophila* oogenesis are the epidermal growth factor receptor (EGFR) and bone morphogenetic protein (BMP) signaling pathways (Berg, 2005; Deng and Bownes, 1997; Dobens and Raftery, 2000; Pyrowolakis et al., 2017; Wasserman and Freeman, 1998; Yakoby et al., 2008b). BMP signaling is activated by the anteriorly emanating Decapentaplegic (DPP) ligand through the type I receptor Thickveins and type II receptors, Punt and Wishful thinking (Chen and Schupbach, 2006; Dobens

and Raftery, 2000; Lembong et al., 2008; Marmion et al., 2013; Twombly et al., 1996). DPP binding allows the type II receptor to phosphorylate the type I receptor kinase domain (Parker et al., 2003). Then, the type I receptor phosphorylates the R-Smad, Mothers against *dpp* (MAD). Activated R-Smads recruit the common Smad Medea (MED) to form a transcription factor complex (Parker et al., 2003).

The major ligand of the EGFR signaling pathway during mid-oogenesis (S7-10A) is the TGF- α like ligand, Gurken (GRK). Gurken possesses an EGF-like domain that activates the EGFR (Marmion and Yakoby, 2018; Queenan et al., 1997; Schnepp et al., 1998; Schupbach, 1987; Schupbach and Roth, 1994; Shilo, 2005; Van Buskirk and Schupbach, 1999). This activation allows the receptor to recruit a second receptor to dimerize and initiate signaling through Ras/Raf/MEK/ERK phosphorylation cascade (Figure 3A). The kinase, ERK, is di-phosphorylated (dpERK) and regulates the activities of transcriptional regulators. Gurken is secreted from around the oocyte nucleus to the perivitelline space and activates the EGFR in the overlaying FCs (Figure 3B,C). The oocyte nucleus position, and therefore GRK and EGFR activation, is spatially changing over time; initially at the posterior end of the oocyte and later anchored to anterior cortex of the oocyte. The activation of EGFR is regulating several *trans*-acting elements (Figure 3D) (Cavaliere et al., 2008; Dobens and Raftery, 2000; Yakoby et al., 2008b). Another level of EGFR signaling regulation is executed by two main inhibitors, Kekk1 (Kek1) and Sprouty (Sty) (Figure 3E). The transmembrane protein Kek1 directly interacts with the EGF receptor to inhibit ligand-receptor interactions (Ghiglione et al., 2003; Ghiglione et al., 1999). Sty acts directly on Ras/MAPK to inhibit dpERK activation (Figure 3E) (Peri et al., 1999; Shilo, 2005). While the two inhibitors work in a classic negative feedback fashion, the mechanism of signaling regulation is different.

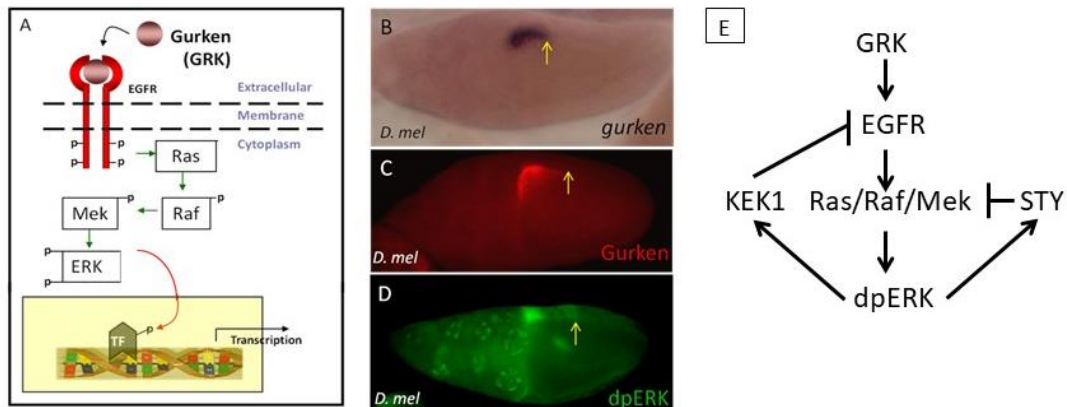


Figure 3: A. The EGFR signaling pathway. B. Distribution of *gurken* in a *D. melanogaster* egg chamber. C. Distribution of GRK in a *D. melanogaster* egg chamber. D. Distribution of dpERK in the FCs of a *D. melanogaster* egg chamber. E. Simplified network of the EGFR signaling pathway featuring negative feedback inhibitors, *Kek1* and *STY*.

1.4 Cell signaling and eggshell diversity

The *Drosophila* eggshells present remarkable diversity in dorsal structures. The formation of dorsal structures is regulated by EGFR signaling, hence suggesting the existence of spatiotemporal changes in EGFR signaling pathway among *Drosophila* species (previously shown in (Dworkin et al., 2003)). A study looking at patterning among three *Drosophila* species found that dorsal ridge formation is downstream of the EGFR signaling pathway (Niepielko et al., 2014). Differences in the distribution of EGFR activation were attributed to a posterior elongation of GRK in species with a dorsal ridge on their eggshells, in comparison to that of *D. melanogaster* (Figure 4) (Niepielko and Yakoby, 2014). The oocyte nucleus association of the *grk* mRNA could not account for changes in EGFR activation patterns among species (Figure 4). Lastly, EGFR activation (detected by antibodies against dpERK) is restricted to the oocyte nucleus in *D. guttifera*, distributed 50% to the posterior on *D. melanogaster*, distributed to the full posterior in *D. cardini*, and 75% to the posterior in *D. willistoni*. The corresponding patterns of GRK were found by using species-specific anti GRK antibodies (Figure 4).

Figure 4: Differences in EGFR activation and eggshells among Drosophila species. Eggshells of D. guttifera, D. melanogaster, D. willistoni, and D. cardini with corresponding staining of dpERK, grk mRNA, and GRK. Anterior is to the left. Eggshells, dpERK, GRK are dorsal views, and grkimages are sagittal views. (D.mel, D.wil., and D. car images from Niepielko and Yakoby, 2014)

Chapter 2: Specific Aims

2.1 Develop a mathematical model that defines parameters and their values contributing to different distributions of GRK/EGFR activation.

This work was performed in collaboration with the Piccoli Mathematics lab. I was responsible for the intensity analysis of the genetic perturbations, and experimental validation of the model. In addition, I discussed the development of the model, prepared figures, and wrote portions of the manuscript that will be submitted for publication.

The distribution of GRK at S10A in *D. melanogaster* extends to 50% to the posterior of the egg chamber on the dorsal side (Niepielko and Yakoby, 2014). The distribution of GRK at the dorsal-ventral axis was initially modeled where 10% of total GRK is present at the ventral side of the egg chamber (Goentoro et al., 2006). This was later validated experimentally using a HRP-tagged GRK construct (Chang et al., 2008).

The formation of GRK gradient is regulated by the rate of ligand secretion, diffusion, receptor binding, and internalization that leads to ligand degradation (Goentoro et al., 2006; VanBuskirk and Schupbach, 1999; Wasserman and Freeman, 1998). The distribution of GRK can be determined by a single dimensionless parameter, the Thiele modulus (ϕ) (Goentoro et al., 2006). It was shown that the Thiele modulus reflects the shape of the GRK gradient and is used to quantify morphogen gradients (Goentoro et al., 2006). While this model can successfully capture the spatial profile of GRK, the spatiotemporal dynamics of GRK distributions remains unaccounted for.

2.1.1 Mathematical Reasoning

The model is designed to simulate EGFR activation dynamics of stages 7 to 10A of egg chamber development. Current models characterized steady state solutions over fixed conditions, i.e. size of egg chamber, rate of diffusion, rate of internalization of receptor (Goentoro et al., 2006; Lembong et al., 2008; Zartman et al., 2011). Our aim is to develop a spatiotemporal dynamic model able to capture the complexity of the system. The model takes into consideration that GRK is secreted initially at the posterior end of the egg chamber and later, as the nucleus anchors to the anterior

cortex of the oocyte, to the overlaying dorsal FCs. Thus, we intend to simulate the dynamics of nucleus/GRK location during these stages when the size of the oocyte changes considerably.

2.1.2 Building a Dynamic Model from Experimental Data

The ligand, Gurken, is secreted from around the oocyte nucleus into the surrounding perivitelline space (PVS). At stages 8 to 10A, the nucleus remains anchored to the dorsal anterior while the FCs shift over it from cuboidal to columnar FCs (Duhart et al., 2017). A previous study found there to be roughly 855 columnar FCs over the oocyte (White et al., 2009). The cells at the anterior transition to a layer of squamous FCs, known as stretched cells. Here we consider three mechanical changes that will affect the distribution of the ligand: 1) the dynamic position of the oocyte nucleus; 2) the growth of the oocyte; and 3) movement of the FCs over the oocyte nucleus or source (Figure 5). By integrating these three factors into our model, we can reflect the dynamics of the pathway by observing the output (dpERK), and its relative changes based on variations in the parameters defined.

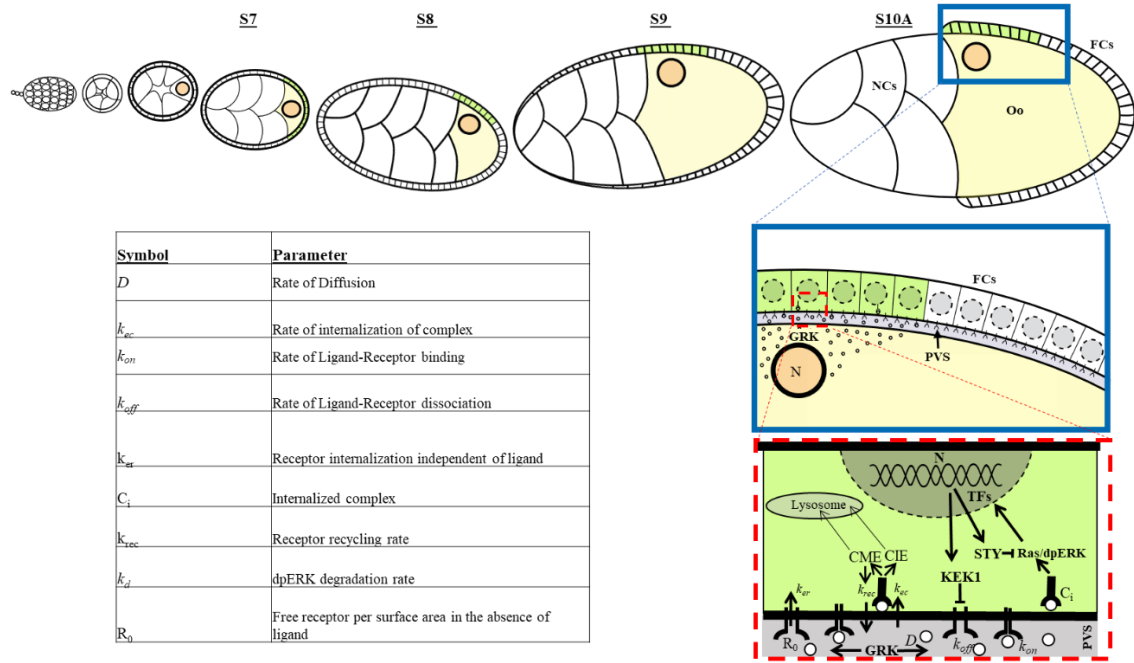


Figure 5: Schematic of the mechanisms used for defining the parameters of the model. The oocyte nucleus has a dynamic localization relative to the FCs during Stages 7 to 10A of oogenesis (top). GRK is localized around the oocyte nucleus, diffuses in the perivitelline space and binds to the EGFR located on the surface of the overlying FCs (blue inset). Internalization of GRK sets off the RAS/RAF/MEK signaling cascade as well as the production of inhibitors that act as negative feedback (red inset).

The parameters considered in our model consist of three movements during stages 7-10A of oogenesis. These movements include: 1) growth, 2) the transient source of GRK in the developing oocyte, 3) the rearrangement of the FCs from cuboidal to columnar and stretched. Several other parameters considered in the model are defined in the table above.

The perivitelline space is a space between the oocyte and FCs (Figure 5, blue inset), where the ligand can diffuse and bind to the EGFR in the overlying FCs with the binding rate k_{on} and dissociation rate, k_{off} (Figure 5, red inset). The ligand-receptor complex is then internalized at the rate k_{ec} , and subsequently activates the EGFR signaling pathway in the overlying FCs. The output of EGFR activation, dpERK, regulates the transcription of genes, or of particular interest to our model, regulators of the pathway.

The EGFR inhibitors, including Argos (Aos), Kekk1 (Kek1), and Sprouty (Sty), regulate eggshells' patterning (Boisclair-Lachance et al., 2009; Ghiglione et al., 1999; Wasserman and

Freeman, 1998; Yakoby et al., 2005; Zartman et al., 2009). *aos* is expressed at later stages of oogenesis (Wasserman and Freeman, 1998), thus it is not included in the model. Our model accounts for two inhibitors that are active during oogenesis at stages 7-10A. The transmembrane protein Kek1 targets the EGF receptor to inhibit dimerization of receptor (Ghiglione et al., 1999). On the other hand, Sty targets directly the Ras/MAPK to inhibit dpERK activation (Peri et al., 1999; Shilo, 2005). Each inhibitor is considered independently based on its inhibitory mechanism.

2.1.3 Modeling the egg chamber growth and source

As we model the growing egg chamber, we have three moving parts to consider. The first is the dynamic position of the oocyte nucleus at the dorsal anterior corner of the oocyte as it grows. The second is the overall growth of the egg chamber, specifically the oocyte. The third is the changing of the FCs as they become columnar and the stretched cells flatten. The respective time-dependent rates of movement, growth and shift were extrapolated from measurements of the egg chamber at each stage (Supplement 1).

The dynamic positioning of the oocyte nucleus at each developmental stage was measured from the anterior oocyte boundary to the posterior most point of the FCs (L_{Oo}). A line was then fit to the measurements to account for the nuclear speed. To account for the size of the overall growth per developmental stage, measurements were taken from the anterior of the egg chamber to the posterior (L_{AP}). To account for the posterior movement of the FCs, measurements were taken from DAPI images for developmental stages 9, early and late, and 10A (L_{FC}), when FC movement takes place. These measurements were then applied to the cubed spheroid coordinate system (see details in Materials and Methods, Supplement 1, Supplement 2)

The model shows the level of detectable GRK secreted from the oocyte into the PVS and is affected by the outlined parameters. The size and the shape of the source of GRK, however, change throughout the developmental stages. As we fine-tuned the model, the shape of the ligand source

is considered an input. Hence, we calibrate this time and space-dependent source directly with experimental measurements (Figure 6). At Stage 7, the oocyte nucleus encompasses the entire size of the oocyte, thus the size of the source is equivalent to the size of the egg chamber. As the oocyte grows, however, the size of the source is refined. Measurements of the source were taken from inside of the oocyte from three different vantage points (Figure 6A, B). At stages 9 early, 9 late, and 10A, direct measurements were taken from the anterior of the oocyte along the dorsal surface using a cross-section and a sagittal view (Figure 6D). The shape of the source is then modeled according to the measurements, and the time-varying flux $V(t, \eta, \theta)$ is designed such that a constant total quantity of ligand is secreted from the source while its size changes over time (Figure 6C). Denoting by $\sigma(t)$ the total surface area of the source at time t , and by V_0 the initial flux of ligand, the total quantity of ligand released at time $t = 0$ is $\sigma(0)V_0$. We set:

$$\begin{cases} V(t, \eta, \theta) = \frac{\sigma(0)V_0}{\sigma(t)} & \text{if the point } (\eta, \theta) \text{ belongs to the source at time } t \\ V(t, \eta, \theta) = 0 & \text{if the point } (\eta, \theta) \text{ is outside the source at time } t. \end{cases}$$

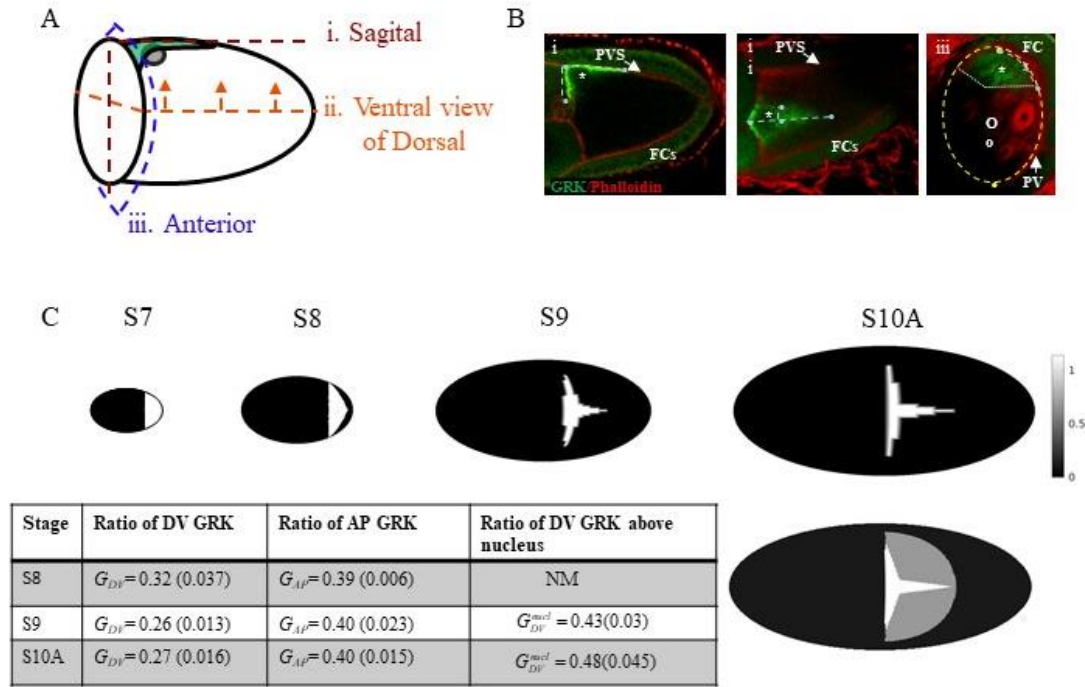


Figure 6: Measurements of the source. A. A cartoon depicting the vantage points from where source measurements were taken. B. Immunohistochemistry staining depicting the vantage points of the egg chamber at S10A: i. Sagittal ii. Dorsal measurement from a ventral view iii. Anterior boundary of oocyte. C. Simulations of the source and table of measurements at specific stages. At S10A we see the current source measured superimposed on what was previously studied (Goentoro et al., 2006).

2.1.4 Determining Parameters

The model parameters were taken either from the literature or directly from experimental measurements (Supplement 3). However, there remain parameters (namely the strengths of the inhibitors $Kek1$ and Sty) that are not found in the literature or measurable from experiments. The strengths of inhibitors γ_{sty} and γ_{kek} were then determined using data obtained from intensity plot profiles from immunostainings of GRK and dpERK at each of the modeled stages, excluding Stage 7 (Figure 7). Intensities were taken in three different genetic backgrounds that represent copy numbers of GRK: 1x, 2x (wild type), and 4x GRK. In the model, this corresponds to setting the initial flux of ligand to $\frac{1}{2}V_0$ for 1x GRK, to V_0 for 2x GRK and to $2V_0$ for 4x GRK. Intensities at the AP and DV axes were compared to simulations containing different parameter values (Supplement

4). The simulations showed the best match with experimental curves for the choice of parameters $\gamma_{\text{Sty}} = 2000 \text{ AU}$ and $\gamma_{\text{kek}} = 5000 \text{ AU}$, as shown in Figure 7A.

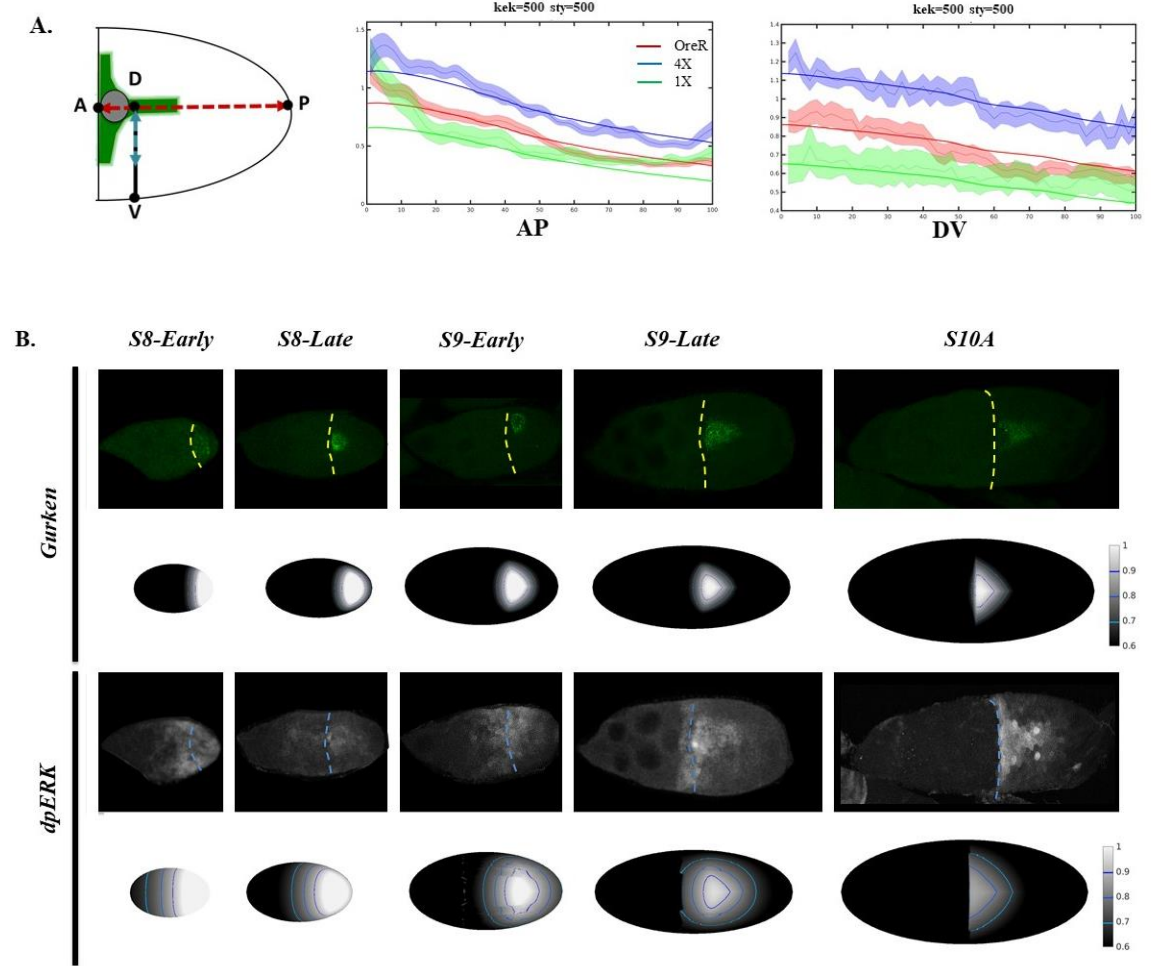


Figure 7: Simulations of GRK/dpERK over developmental stages A. Images from Wild type (2x GRK, red), 4X GRK (blue), and 1x GRK (green) were measured for pixel intensities at the anterior-posterior and dorsal-ventral region of the egg chamber directly after the oocyte nucleus. These values were averaged and then plotted against simulations with many combinations of parameter values. Parameter values were selected for the best fit plots for AP and DV over the five stages considered in this model. B. Immunohistochemistry and model predictions for GRK and dpERK at stages 8-early to stage 10A.

Figure 7B shows the full distributions of dpERK and GRK, obtained by numerical simulations with the chosen parameters, at various stages of development. A threshold was applied to each of the simulations, where all concentrations below 60% of the maximum concentration are equal to zero. The model successfully captures the distribution of GRK and dpERK on a growing manifold with a moving morphogen source.

2.1.5 Validating Parameter Selection

To test the predictability of the model, we used RNA interference perturbations in *sty* and *Egfr*, and looked at the signaling output (dpERK). In the *sprouty* RNAi background, the intensity of GRK is unchanged (Figure 8A). Since the used GAL4 driver in this experiment, CY2-GAL4, is uniformly expressed from stage 8, we gave the system time to adjust, and considered changes only from stage 9.

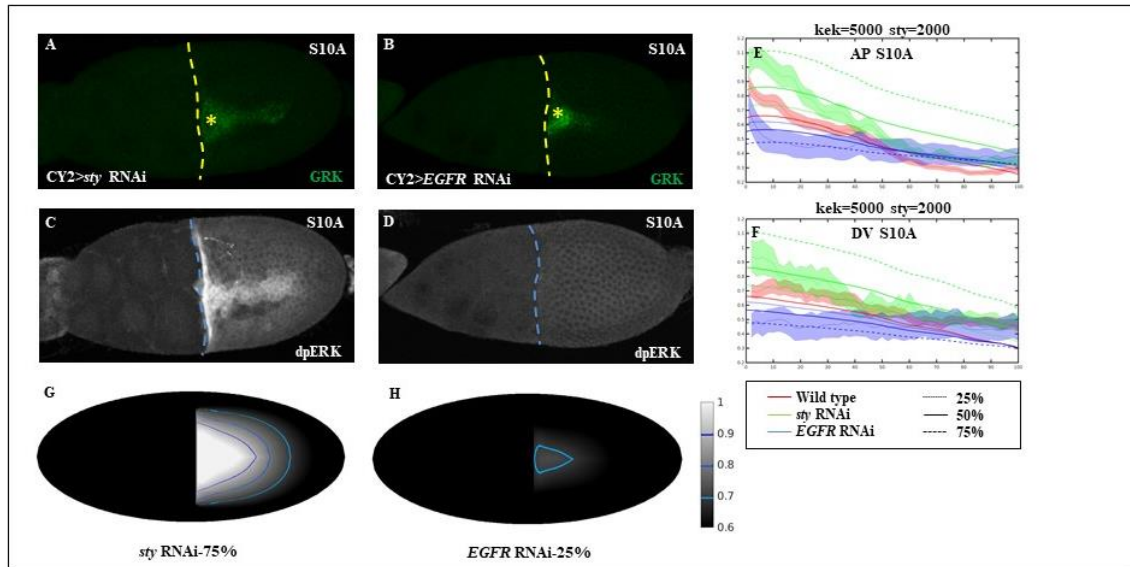


Figure 8: Validating parameter selections using genetic perturbations. A-D. Immunohistochemistry staining at S10A for A. GRK in Cy2>EGFR RNAi ($n=10$), B. GRK in sty RNAi ($n=11$), C. dpERK in Cy2>EGFR RNAi, D. GRK in CY2>sty RNAi. E-F. AP and DV, respectively, intensity profiles of wild-type (red), sty RNAi (green) and EGFR RNAi (blue) and corresponding intensity profiles of simulations where RNAi targets were reduced by half. The wild-type simulation corresponds to the parameters defined in Tables 1 and 2. The sty RNAi perturbation corresponds to a modified $\gamma_{sty}(t) = 0$ for all $t > 6$, to cancel the effect of Sty on the dynamics after Stage 7.

As expected, the effect of this perturbation on the signal, however, shows a significant increase in dpERK in RNAi targeting *sty* and a decrease in signal when EGFR is depleted (Figure 8C, F). The concentration of dpERK along the AP axis at stage 10A ($t = 21h$) is shown in three different simulations (Figure 8E-F).

The *sty* RNAi experimental intensity plot reflects the increase in signal observed in this genetic background compared to the wild type at 21 hours. Since the total amount of *sty* at this stage is unknown, as is the strength of the RNAi, several simulations were run to predict the amount of reduced *sty*. The experimental intensity plots in the AP overlap with the posterior decay predicted by the simulation at a reduction of 25%, or 75% of the endogenous level (Figure 8E, F). However, the DV profile shows slower degradation of dpERK in the plot profile. Surprisingly, when a qPCR analysis of *sty* in egg chambers expressing *sty*-RNAi at S9-10 was performed and compared to the wild type, the qPCR detected a fold change at 78%, or a 22% decrease, in *sty* compared to the wild type which is consistent with the model prediction.

The *EGFR* RNAi perturbation corresponds to a reduced level of available receptors at the surface of the FCs. The experimental intensity plots show a severe decrease of dpERK (Figure 8D) where GRK remains unaffected (Figure 8B). Several simulations were run to estimate the reduction of receptor in the *EGFR* RNAi perturbation. Comparisons of intensity measurements between wild-type and *EGFR* RNAi show a decrease of 75% of *Egfr* in the *Egfr* RNAi perturbation compared to wild-type (Figure 8E, F, H). This corresponds to a depletion of 75% of the receptors in our model, i.e. setting the initial amount of receptors to $R(0, \eta, \theta) = \frac{R_0}{2}$ and their production rate to $\frac{Q_r}{2}$.

To summarize, we recall the three categories of parameters, classified according to the method by which we set their values.

- Parameters taken from previous literature: diffusion rate (D), reaction rates (k_{ec} , k_{on} , k_{off} , k_{er} , k_{rec} , k_{deg} , k_d , α_{rec} , α_{deg} , Q_r), initial concentration of receptors (R_0), initial flux of ligand (V_0)
- Parameters measured experimentally: physical dimensions of the egg chamber (L_{AP} , L_{DV} , L_{OO} , L_{FC} , H), dimensions of the source of ligand

- Parameters calibrated by comparing simulations and immunostainings of GRK and dpERK: strengths of inhibitors (γ_{Sty} , γ_{Kek})

2.1.6 The Effects of Mechanical Changes

We used our model to test the role of mechanical transformations of the egg chamber on the distribution of dpERK at all stages. To study the impact of each transformation, we perturbed individually in numerical simulations the following: 1) movement of the oocyte nucleus, 2) the shift of the FCs, and 3) the growth of the egg chamber. Qualitative simulation results are shown in Figure 9. The second row of Figure 9 shows the distribution of dpERK at Stages 8, 9 and 10A when the movement of the nucleus is stopped at late Stage 8 ($t = 10.5h$) (Supplemental Figure 1). As expected, the signal is perturbed and dpERK is found only in the posterior of the egg chamber compared to the wild type (top row). The signal is lost at the dorsal anterior. This simulation agrees with perturbing microtubule-dependent nuclear anchoring by treating flies with colchicine (Niepielko et al., 2014; Roth et al., 1995).

As a second perturbation, we removed the shift of the FCs, as shown in the third row of Figure 9. This has two effects on the signal. At Stage 9, one can see that the signal is less elongated than in wild-type. At Stage 10A, signal is fainter at the dorsal anterior. Since the same relative shape is observed for both stages suggests that the changing of the FCs is not necessary to create the elongation in the distribution of the signal. However, we did observe that the anterior FCs act to maintain high signal levels by acting as a physical boundary that prevent ligand dispersal anteriorly.

Lastly, we used the model to perturb the growth of the egg chamber. The fourth row of Figure 9 shows the dpERK distribution at Stages 8, 9 and 10A when growth is stopped at late Stage 8 ($t = 10.5h$). This mechanic change has two consequences. First, it affects the diffusion of the signal as the growth affects the curvature of the perivitteline space, where the changes in coordinates affect

the Laplace-Beltrami operator. At earlier stages, the egg chamber's A/P and D/V dimensions are comparable, so the egg chamber is almost spherical. As a consequence, diffusion at earlier stages is almost isotropic. At later stages, wild-type egg chambers are more elongated, so this earlier symmetry is broken and diffusion becomes anisotropic. This biased diffusion of GRK affects the shape of the signal. When growth is perturbed, we observe that the signal profile is less elongated. Second, the perturbation of growth also affects the signal intensity, i.e. as the egg chamber remains of small size, GRK remains more concentrated and creates higher levels of dpERK.

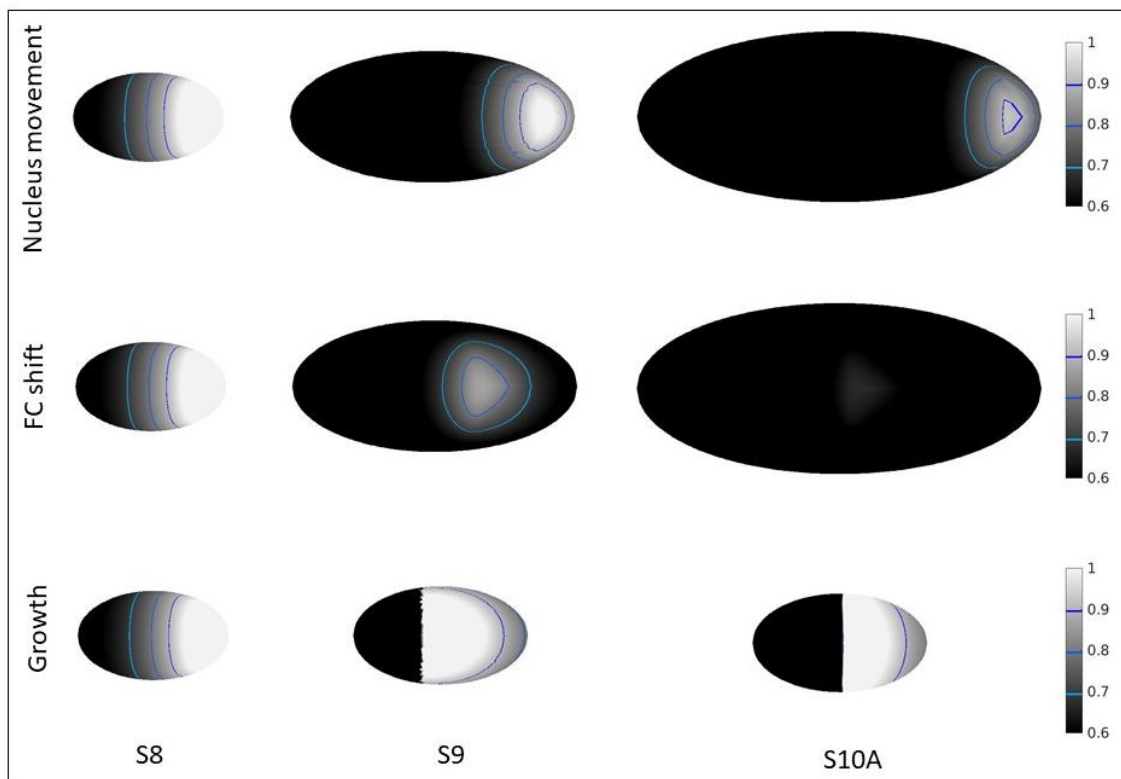


Figure 9: Exhibiting the role of *mechanical* changes. The model was used to predict the effect of *mechanical* perturbations on the distribution of dpERK. Each row shows dpERK at stages 8, 9 and 10A for a different *mechanic* perturbation. First row: the nucleus movement was stopped at late S8 ($t=10.5h$). Second row: the follicle cell shift was removed. Third row: growth was stopped at lates S8 ($t=10.5h$).

Chapter 2.2: Quantitative analysis of the EGFR in the follicle cells over developmental stages

For this project I was responsible for the quantitative analysis of the EGFR and innovation of experimental design. The Egfr-egfp CRISPR/Cas9 was made by Robert Marmion. I generated the EGFR- sfgfp, designed and performed the experiments, produced the figures, and wrote parts of the manuscript.

It was previously shown that the levels of distribution of EGFR in the FCs impacts the distribution and levels of GRK and EGFR activation patterns (Goentoro et al., 2006). The rate at which GRK is internalized effects the overall distribution of signal. While there are known distribution patterns of GRK and dpERK in perturbations of the EGFR, there is little understanding of receptor dynamics throughout oogenesis. It was previously reported that the dynamics of the EGFR change in the presence of high ligand during S10A of oogenesis as a dorsal clearing is present (Sapir et al., 1998). In addition, the mechanisms of EGFR trafficking once it is internalized are unclear, whether the receptor returns to the plasma membrane or be trafficked to the lysosome for degradation (Sigismund et al., 2008). It was shown that the receptor-ligand complex continues to signal after internalization (Vieira, 1996), and its fate after internalization remains unclear.

2.2.1 Quantification of the EGFR

While several reports show the quantification of the EGFR in culture cells, there is little information available to other systems, including tissues (Pribyl et al., 2003b; Zhang et al., 2015). Since the distribution of the EGFR is variable by cell type, it is important to quantify the amount of available receptor during different stages of oogenesis to determine the involvement of the EGFR in shaping the distribution of GRK, and consequently dpERK. To quantify and follow EGFR distribution in the FCs, an endogenously tagged EGFR with EGFP was generated in our lab using CRISPR-Cas9. There are alternative exons at the 5' end of the transcript due to the EGFR isoforms (Schejter et al., 1986). To circumvent this issue and include all annotated transcripts, an EGFP tag was inserted between the last exon and the 3' UTR on the C terminus of the protein (Figure 10A). Flies that were positive for the dsRed selection marker were verified by PCR and sequencing.

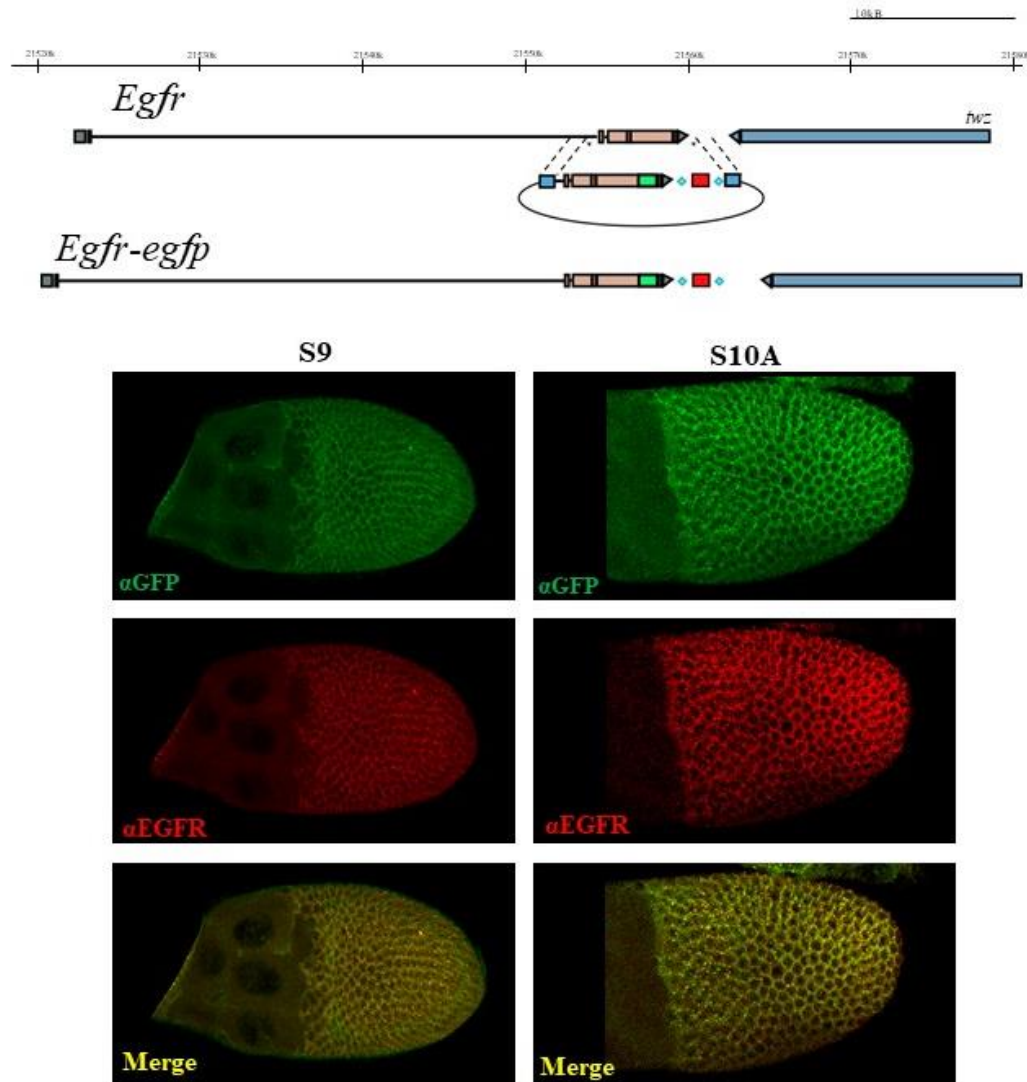


Figure 10: Using CRISPR-Cas9 to generate an EGFR tagged with GFP. A. The EGFR locus was used to generate an EGFR tagged with EGFP. Two guides were designed in the region flanking the four exons at the 3' end of the EGFR locus. A donor construct was generated with the replacement four exons, an EGFP tag, and the 3' UTR. The construct used two homology arms flanking the gRNAs to allow integration of the donor construct. B. Immunohistochemistry for GFP and EGFR in EGFR-EGFP fly line. The EGFR is present in all FCs of the egg chamber at all stages.

2.2.2 A GFP tagged EGFR construct

In order to visualize and quantify EGFR, we endogenously tagged the EGFR with EGFP (Figure 10). Adult homozygous flies were phenotypically wild type. However, 31% of the eggshells (n=371) had a phenotype where appendages were closer together, indicating on a reduction of EGFR signaling in this background. At the same time, the parent flies used for this injection showed

the same phenotype (15%, n=49). It was previously reported by the Shilo Lab that EGFR is expressed uniformly in the FCs (Sapir et al., 1998). Repeating this experiment, we stained *D. melanogaster* egg chambers with EGFR antibodies and found similar results (Figure 10B'-C'). Staining the same egg chambers with anti-GFP antibodies providing a similar pattern (Figure 10B-C). While the overlap is clear for both images, the EGFR antibody stain had increased non-specific staining making localization harder to detect (Figure 10B''-C'').

Interestingly, at stage 8, the immunostaining of EGFP and EGFR show high concentration of bright puncta localized at the apical side of the FCs, which coincides with the localization of GRK (Figure 11A-A''). The puncta were more much elevated in the posterior compared to the anterior of the egg chamber. These bright puncta were not detected at early S9 (Figure 11B-B'') where the EGFP was mainly observed at the lateral surface of the plasma membrane. At this stage, the co-localization of EGFR and GRK is 19% at the anterior and 5% at the posterior (Figure 11D) (n=38). At S10A, the apical localization is lost and the basolateral distribution of EGFR-EGFP continues in all FCs (Figure 11C-C'). At this stage, the reduction in EGFR levels overlapped the high levels of GRK (22%) at the dorsal anterior (Figure 11E, n=27), which is in agreement with a previous study (Sapir et al., 1998). The posterior end of the egg chamber had very little co-localization (Figure 11E).

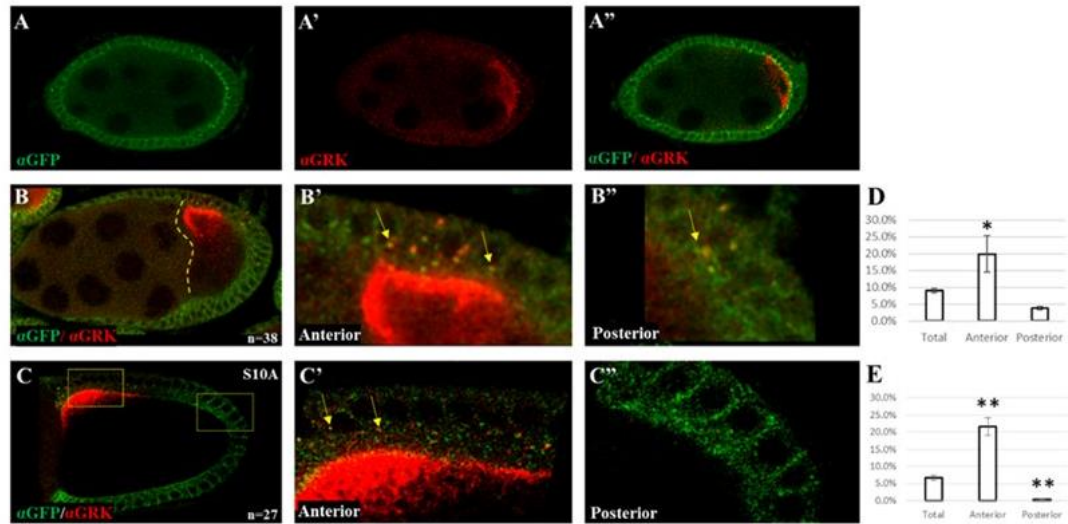


Figure 11: Egg chambers show co-localization of EGFR GFP and GRK. A. S8 egg chamber stained for GFP A'. GRK A'' Merge. B. Egg chamber at S9 co-stained for GFP and GRK B'. Anterior S9 B''. Posterior C. Egg chamber at S10A co-stained for GFP and GRK C'. Anterior S9 C''. Posterior D. Bar graph of co-localization percentages between the total dorsal side, dorsal anterior, and dorsal posterior at S9 E. S10A (* $P < 0.01$, ** $P < 0.001$)

To determine whether the EGFR localization to the apical side at S8 depends on GRK, we used CRISPR/Cas9 to generate a deletion of the *grk* locus (Figure 12A). While most of these flies remain heterozygous, a few sterile homozygous females were recovered. Eggs laid by these females do not possess operculum or dorsal appendages, which are distinct from the wild type eggshells (Figure 12B,C) and consistent with the previously reported null alleles (Schupbach, 1987). Immunohistochemistry using EGFR specific antibodies shows the loss of apical localization of the EGFR at stage 8 (Figure 12E) in comparison to the wild type (Figure 12D). We conclude, apical localization of EGFR at early stages of oogenesis depends on GRK. Due to the absence of a dorsal-ventral axis in a *grk* null background, we could not follow dorsal-anterior changes in EGFR at stage 10.

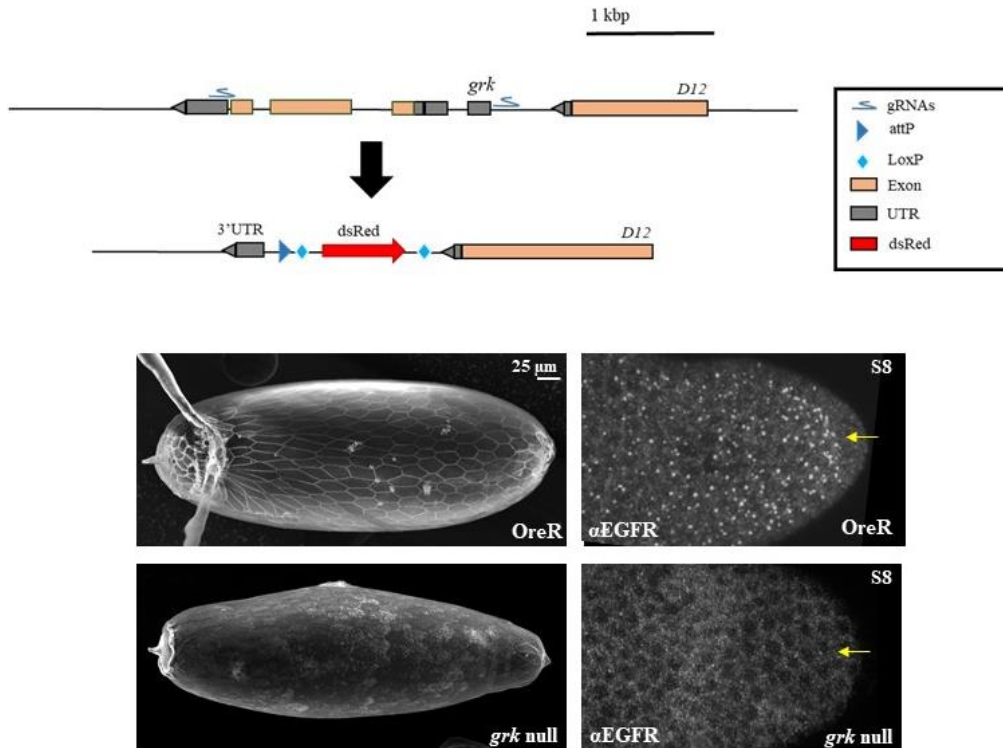


Figure 12 A. Cartoon representation of the GRK locus targeted by CRISPR/Cas9. B. Eggshell of *OreR* C. *grk null*. D. Egg chambers at S8 stained for EGFR in *OreR* (D) and *grk null* (E).

2.2.3 EGFR quantification throughout oogenesis and embryogenesis

The differences in the spatial distribution of the EGFR allow us to question if dynamics of the pathway lead to any quantitative differences observed throughout oogenesis. Furthermore, since only low levels of maternal RNA were detected in the early embryo (Lev et al., 1985; Mohammed and Lai, 2016), we also aimed to quantify EGFR levels in early embryogenesis at pre-cellularization (0-1 hr, or stage 2) and cellularization (2.5 hr, or stage 5) stages. To quantify the levels of EGFR, we collected a known amount of egg chambers and embryos from EGFR-GFP and *yw*- (negative control) flies. We used an ELISA kit to detect GFP (GFP detection sensitivity ranges 30pg-2ng, see details in methods). Since the amount of detectable EGFR in a total protein sample was initially unknown; we used whole ovary in samples ranging from 200-1.25 μ g of total protein to determine the amount of total protein to use per well (See details in methods). To avoid

saturating the wells with signal and obtain enough signal to detect differences among developmental stages, we used 75 μg total protein in each well to quantify EGFR for all samples. We used the total protein collected from the whole ovary to approximate the number of receptors for each FC for all developmental stages. To do this, total protein was collected from 10 flies (twenty ovaries). It can be approximated that each ovary contained 16 ovarioles with 7 egg chambers at different developmental stages per ovary. Each egg chamber contains on average 855 FCs (White et al., 2009). Using these numbers, we calculated the total weight of GFP per cell. If we convert the weight of GFP (27 kDa, or g/mol) to the weight of a single molecule (4.48×10^{12} pg/receptor), we can obtain the number of GFP molecules/per cell, which is equivalent to the number of receptors per cell. The sample containing all developmental stages of oogenesis contains approximately 4000 receptors per cell.

Once established that we have detectable levels of EGFR in our samples, we next collected 100 egg chambers in three groups at stages 8 and 9, stage 10A-B, and stages 11 and greater (Figure 13 A). We included 60 embryos at 0-1 hr and 2.5 hours (Figure 13B). All collections were done in triplicates. Looking at the ELISA results, we found that EGFR is dynamically present in the FCs. Early (S8-9), we calculated around 4000 receptors/cell and a similar value was obtained for egg chambers at >S11. However, this number significantly increased at S10A-B where approximately 7000 receptors/cell were detected (Figure 13C). The amount of detected receptors were in the order of magnitude of 1000-fold less than what has been reported in A431 epidermoid carcinoma cells that were measured to have 1.2×10^6 receptors per cell (Zhang et al., 2015). The amount of receptors in embryos showed no detection at stage 2, and about 200 receptors/cell in the ~6000 cells of the embryo blastocyst stage (Figure 13C).

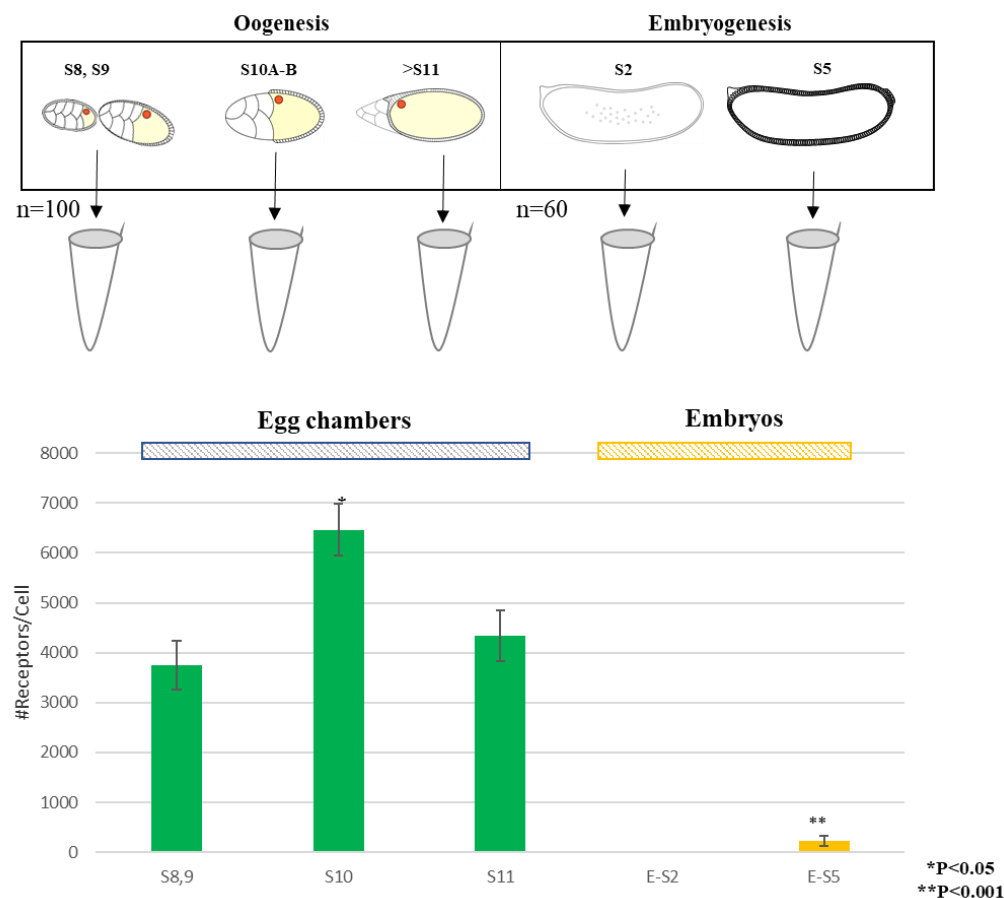


Figure 13: Quantification of the EGFR during oogenesis and embryogenesis A. A cartoon depicting egg chambers collected for stages of oogenesis and embryogenesis. B. A bar graph of receptors per cell in each of the stage specific groups of oogenesis and embryogenesis.

2.2.4 Intracellular dynamics of the EGFR

The trafficking of the EGFR has been widely studied in many cell types (Burke et al., 2001; Fortian and Sorkin, 2013; Yang et al., 2017). However, clear methods of detection are not well established for visualizing trafficking of the EGFR during oogenesis. Previous reports show that the ligand/receptor complex is internalized and sorted through endosomal compartments (Chang et al., 2008). In order to determine the cellular compartment of EGFR, we generated a super folded GFP (sfGFP) tagged EGFR. We expect that the potential delayed folding of EGFP, as well as the auto-dimerization potential of EGFP that was shown as to have functional defects on tagged constructs (Carrell et al., 2017), are eliminated with this construct (Supplement 5). The sfGFP tagged EGFR was used to determine the co-localization of EGFR with the early and late endosomal markers,

Rab5 and Rab7, respectively (Stenmark, 2009). It was previously shown along the dorsal ventral axis that lysosomal degradation was more efficient in regions of the FCs where the concentrations of GRK are high (Chang et al., 2008). We aim to determine if the same intracellular dynamics can be applied to the anterior posterior boundary as well.

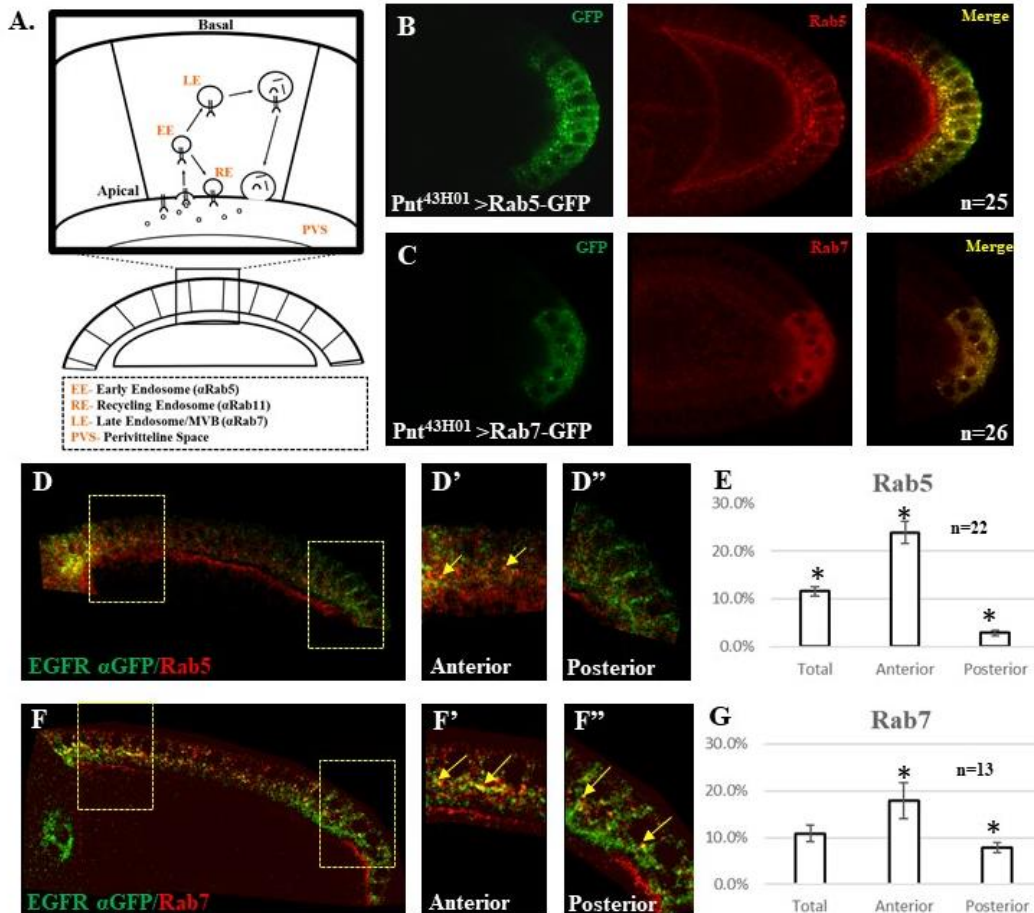


Figure 14. Intracellular localization of the EGFR A. A cartoon depicting the endocytic pathway. Posterior driver (*pnt^{43H04}-GAL4*, Revaitis et al., 2017) expressing B. *UAS-Rab5* and C. *UAS-Rab7*. D. Localization of the EGFR with Rab5 along the dorsal side of the egg chamber at S10A. Insets represent co-localization at the D'. Anterior and D''. Posterior. E. Bar graph representing co-localization with Rab5 in different domains. F. Localization of the EGFR with Rab7 along the dorsal side of the egg chamber at S10A. Insets represent co-localization at the F'. Anterior and F''. Posterior. G. Bar graph representing co-localization with Rab7 in different domains. Yellow arrow points to regions where co-localization detected. (* $P < 0.01$)

The early endosome, or sorting endosome, is necessary for determining whether receptor is recycled to the plasma membrane or trafficked to the multivesicular body (MVB) or late endosome for degradation (Figure 14A) (Bakker et al., 2017). To determine if Rab antibodies are sensitive enough to detect the intracellular compartments and be used to determine co-localization, Rab5 and

Rab7-GFP constructs were expressed with a driver active in the posterior, *pnt*^{43H01} (Revaitis et al., 2017) (Figure 14B-C). We detected co-localization of 87% for Rab5-GFP/Rab5 and 83% for Rab7-GFP/Rab7. Looking directly at the at the Rab5 antibody with the EGFR-GFP, at S10A the overall co-localization of EGFR with Rab5 shows low presence in 11.5% of all puncta (Figure 14D, E). Higher levels of Rab 5/EGFR puncta co-localization are detected in the dorsal side (23.8%) where the levels of GRK and EGFR signaling are high (Figure 14D', E). This is greater than the amount of receptor puncta co-localized with the early endosomal marker at the posterior (2.9%, Figure 14D'', E). The low co-localization could be residual of earlier stages, where the source of ligand was once at the posterior end. To better understand the intracellular distribution of receptor puncta in the late endosome, we next observed that the overall detection of EGFR co-localized with Rab7 at the dorsal side FCs was present in 10.8% of all puncta (Figure 14F). The amount of receptor co-localized with Rab7 was detected in 17.8% of all puncta at the dorsal anterior, which is significantly greater than the co-localization along the dorsal side ($P < 0.01$). The 7.8% that was detected at the dorsal posterior was significantly greater than the 2.9% of EGFR co-localized at the posterior with Rab5 ($P < 0.01$, Figure 14C'-C'').

2.2.5 EGFR expression in imaginal discs

The EGFR is a single receptor in *D. melanogaster*. Since EGFR signaling is used multiple times during development, the EGFR-sfGFP line is not a restricted tool for the FCs. Since the GFP construct is homozygous viable, hence it is expected that the detection of GFP will represent the endogenous patterns of EGFR in all tissues. Imaginal discs of third instar larvae were analyzed for expression of sfGFP. In the wing disc, the EGFR is ubiquitously expressed. High levels of EGFR detection are shown in the wing pouch except in the AP and DV boundaries, which is consistent with detection of transcript by *in situ* hybridization (Guichard et al., 1999; Pallavi et al., 2006) (Figure 15A). The EGFR is ubiquitously detected in the eye disc, and higher levels are observed in the region directly posterior to the morphogenetic furrow (Figure 15B). This observation is consistent with the pattern of EGFR in this tissue (Courgeon et al., 2018; Zak and Shilo, 1992). The EGFR

is uniformly detected in the haltere with elevated levels in the region surrounding the dorsal compartment boundary similar to previously shown by *in situ* hybridization of the haltere (Figure 15C) (Pallavi et al., 2006). The three leg discs have uniform pattern of EGFR, with higher levels at the lower compartment of disc T2 and T3 (Figure 15 D-F). In summary, while the corresponding patterns of EGFR could be confirmed in the wing, haltere, and eye discs, no data is available for expression pattern of EGFR in the leg discs. Based on the correct patterns of the former, we suggest that the patterns found in the leg discs reflect the endogenous pattern of EGFR expression in them.

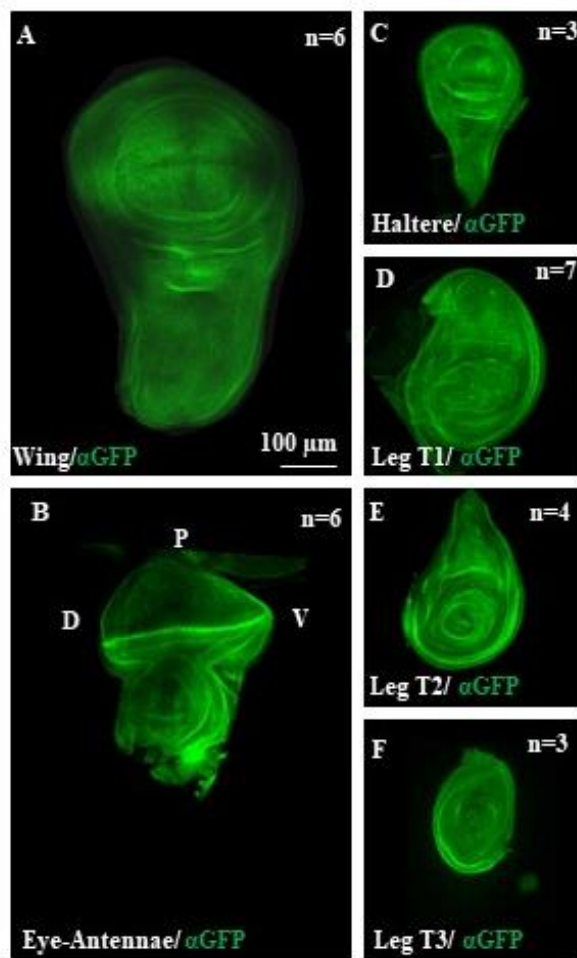


Figure 15: Expression of the EGFR protein detected by GFP in imaginal discs A. Wing B. Eye-Antenna (P:posterior, D:dorsal, V:ventral) C. Haltere D-F. (Leg thoracic discs (T1-T3). D. Leg T1 E. Leg T2 F. Leg T3

Chapter 2.3 Screening for *cis* regulation of *Drosophila* genes (Revaitis *et al.*, G3, 2017)

For this project, I, with the aid of five undergraduate students that I supervised, was responsible for screening the 281 FlyLight lines, performed the RNA-seq analysis of the distribution of the fragments on the gene's model, and performed the analysis on the genetic perturbations. I generated the images and organized them into the supplemental atlas that characterized each of the 54 positive lines, produced the figures, and wrote parts of the manuscript. This chapter was published in G3: Genes, Genomes, Genetics.

The spatiotemporal control of gene expression is a fundamental requirement for animal development (Davidson and Erwin, 2006; Levine, 2010). Research in *Drosophila melanogaster* has provided insight into the complex process of tissue patterning and cell fate determination during animal development (e.g. (Konikoff *et al.*, 2012; Lecuyer *et al.*, 2007; Tomancak *et al.*, 2002)). Large-scale screens for *cis*-regulatory modules (CRMs), which control spatiotemporal expression of genes, provided compelling examples of gene patterning in embryo, central nervous system (CNS), and imaginal disc development (Jenett *et al.*, 2012; Jory *et al.*, 2012; Li *et al.*, 2014; Manning *et al.*, 2012; Pfeiffer *et al.*, 2008). Despite comprehensive screens to systematically search for CRMs in *Drosophila*, our understanding of how genes are regulated in time and space is still limited (Arnold *et al.*, 2013; Kvon *et al.*, 2014; Manning *et al.*, 2012; Pfeiffer *et al.*, 2008). Furthermore, analysis of gene regulation during *Drosophila* oogenesis still remains underexplored.

2.3.1 Screening for regulatory domains

During oogenesis, the egg chamber, the precursor of the mature egg, is extensively patterned through 14 morphologically distinct stages (Figure 1,5,16A-C) (Berg, 2005; Jordan *et al.*, 2005; Spradling, 1993; Yakoby *et al.*, 2008a). Previously, we the expression pattern of >80 genes was characterized in the FCs, a layer of epithelial cells surrounding the oocyte (Fregoso Lomas *et al.*, 2013; Jordan *et al.*, 2005; Niepielko *et al.*, 2014; Yakoby *et al.*, 2008a). The Yakoby lab established

genes that are expressed dynamically in distinct domains that can be combinatorially assembled into more complex patterns (Yakoby et al., 2008, Niepielko et al., 2014). The initial pattern annotations were described as domains that used Boolean operations to make complex patterns combinatorially from the primitive domains anterior (A), dorsal (D), midline (M), Roof (R), Floor (F), and Uniform (U) (Yakoby et al., 2008a). However, to use these domains in a comparative analysis among *Drosophila* species, they were later redefined into mutually exclusive domains which added two new domains, posterior (P), and dorsal ridge (DR, not relevant to this study), and replaced anterior with anterior dorsal (AD) and anterior ventral (AV). In this study, we added to the follicle cell expression domains gene/reporter expression in additional domains, including the germarium (G), stalk cells (StC), border cells (BC), and stretched cells (SC) (Figure 15A-C). These domains serve as a platform for the spatiotemporal patterning analysis.

Figure 16: Screening for expression domains during oogenesis. A) Early stages egg chamber (stages 2-8). Germarium (G) and stalk cells (StC). Egg chamber at stage 10B in a sagittal B) and dorsal C) views. Different groups/domains of cells are marked, including border cells (BC), stretched cells (SC), nurse cells (NC), centripetally migrating follicle cells (CMFCs), oocyte associated follicle cells (FC), oocyte (Oo), oocyte nucleus (N). Cellular domains are marked, including the anterior (A), posterior (P), midline (M), floor (F), and roof (R). D) Summary of the screen for cis-regulatory modules (CRMs) during oogenesis.

To identify the CRMs of genes controlling tissue patterning, we took advantage of the FlyLight collection of flies, which consists of ~7000 fly lines containing intronic and intergenic DNA fragments, representing potential regulatory regions of ~1200 genes (Jenett et al., 2012; Pfeiffer et

al., 2008). Our screen focused on the 84 genes known to pattern the FCs during oogenesis (e.g. (Deng and Bownes, 1997; Dequier et al., 2001; Fregoso Lomas et al., 2013; Jordan et al., 2005; Ruohola-Baker et al., 1993; Yakoby et al., 2008a)). Cross-listing the 84 genes with the FlyLight list yielded 22 common genes (Figure 16D). These genes are associated with a total of 281 fly lines containing CRMs that are potentially active during oogenesis. All DNA fragments in FlyLight collection are upstream to the transcription factor GAL4. Crossing these lines to a UAS-pStinger-GFP fly yielded 54 GFP positive lines. Of importance, 16 of the 54 fly lines recapitulated the full or partial endogenous pattern of the corresponding genes (Figure 16D and Supplemental Material, (Revaitis et al., 2017)).

The BMP inhibitor, *daughters against dpp* (*dad*) (Inoue et al., 1998; Tsuneizumi et al., 1997), is expressed in the stretched cells and the anterior FCs (Jordan et al., 2005; Muzzopappa and Wappner, 2005; Yakoby et al., 2008a) (Figure 17Aa). Three of the six associated FlyLight lines express GFP (Figure 17Ab-k). The *dad*^{H4C10} line recapitulates the *dad* endogenous expression pattern (Figure 17Ac-e). The *dad*^{H4C10} line is expressed in the centripetally migrating cells and in the stretched cells from stage 8 (Supplemental Material, Figure S2Dc, Revaitis et al., 2017). Interestingly, the *dad*^{H3H04} line is restricted to the stretched cells (Figure 17Af-h). Previously, we referred to the anterior domain as the centripetally migrating cells, which include the anterior oocyte-associated FCs (Figure 17B, C) (Niepielko et al., 2014; Yakoby et al., 2008a). Here, we found that the anterior pattern is comprised of two patterns, one is restricted to the stretched cells and another includes both the stretched cells and centripetally migrating FCs (Figure 17Ad, g). The

dad^{45C11} line is expressed in 1-2 border cells (we cannot distinguish whether these are border cells or polar cells) at stages 9-10B (Figure 17Ai-k).

Figure 17: Expression domains of several FlyLight lines. .A. Expression domains of daughters against *dpp* (*dad*) **Aa** A cartoon describing the *dad* expression pattern in the stretched cells (SC) and anterior (A) domains. **Ab** The gene model for *dad* and the associate FlyLight fragments screened during oogenesis. The GFP positive lines are marked in orange. Asterisk (*) indicates a line with expression not seen in endogenous gene patterns. Empty arrowhead denotes the transcription start site and the direction of the gene in the locus. **Ac-e** GFP expression driven by the *dad*^{44C10} during stages 9-10B in the stretched cells (SC) and centripetally migrating FCs, denoted as the anterior (A) domain. Broad (BR) is used as a spatial marker. "n" is the number of images represented by this image. Arrowhead denotes the dorsal midline. **Af-h** GFP expression driven by the *dad*^{43H04} during stages 9-10B in the SC. **Ai-k** GFP expression driven by the *dad*^{45C11} during stages 9-10B in the border cells (BC). **Ai'-k'** Insets of i-k (white arrow denotes the BC).

The zinc finger transcription factor *broad* (*br*) is expressed in a dynamic pattern during oogenesis. At early developmental stages, *br* is uniformly expressed in all FCs. Later, it is expressed in two dorsolateral patches on either side of the dorsal midline (Figure 17Ba, b and (Deng and Bownes, 1997; Yakoby et al., 2008b)). Two lines, *br*^{69B10} and *br*^{69B08}, express GFP (Figure 17Bc-i). The former is expressed in a uniform pattern, similar to the early expression pattern of *br* CRM (*brE*) (Figure 17Bd-f). However, unlike the early pattern of *br*, it does not clear from the dorsal domain at stage 10A (Cheung et al., 2013; Fuchs et al., 2012). The other

line, *br*^{69B08}, is expressed in the roof domain, like the late pattern of the *br* CRM (*brL*) (Figure 17Bg-i). Interestingly, unlike the *br* gene and the published *brL*, this line is also expressed in the floor domain (Figure 17Bg'-i'). We further discuss these CRMs later (Figure 17Bj).

Figure 17B: cont'd. Expression domains of broad (br) Ba, b) A cartoon describing the expression patterns of early and late broad. Bc) The gene model for br and the associate FlyLight fragments screened during oogenesis. The GFP positive lines are marked in orange. Bd-f) GFP expression driven by the br^{69B10} during stages 9-10B is uniform in all FCs. Bg-i) GFP expression driven by the br^{69B08} during stages 10B-12 in the roof (R) and floor (F) domains (brRF). Bg'-i'). Insets of g-i. Bj) The position of the different br fragments. brE, brL, and brS (Charbonnier et al., 2015), and br^{69B08} (brRF - this screen). The br^{69B08} is shorter 250bp and 53bp on the left and right ends, respectively, of the brS fragment.

The ETS transcription factor *pointed* (*pnt*) is necessary for proper development of numerous tissues, including the eye and eggshell (Deng and Bownes, 1997; Freeman, 1994; Morimoto et al., 1996; Zartman et al., 2009). Two isoforms are expressed in the FCs during oogenesis, *pnt-P1* and *pnt-P2* (Figure 17Ca,b) (Morimoto et al., 1996). The *pnt-P1* isoform is expressed in the posterior domain from stage 6 to 9. At stages 10A and 10B, it is expressed in the dorsal midline (Figure 17Ca, b) (Boisclair Lachance et al., 2014; Morimoto et al., 1996). Later, at stage 11, *pnt-P1* isoform is expressed in the floor and posterior domains (Yakoby et al., 2008a). Two overlapping FlyLight lines show a similar pattern of GFP expression (Figure 17Cc-i). The *pnt*^{45D11} and *pnt*^{43H01} lines express GFP in the posterior and border cells (Figure 17Cd-i). In addition, *pnt*^{43H01} is broadly

expressed in the stretched cells (Figure 17Cd-f). None of the screened lines associated with the *pnt* gene were found to contain the information for the midline expression pattern of *pnt-P1*. The midline pattern of the *pnt-P1* transcript could be visualized by the GFP tagging of the endogenous *pnt* gene (Boisclair Lachance et al., 2014). In addition, none of these lines recapitulate the pattern of *pnt-P2*, which is expressed in the midline (at stage 10A) and roof (at stage 10B) domains (Supplemental Material, Figure S2Rb, Revaitis et al., 2017) (Morimoto et al., 1996).

Figure 17C: cont'd Expression domains of pointed (pnt). Ca, b) A cartoon describing the expression patterns of pointed-P1 (pnt-P1) during stages 6-8 in the posterior (P) domain, and at the floor (F) and P domains at stage 11. Cc) The gene model for pnt isoforms and the associate FlyLight fragments screened during oogenesis. The GFP positive lines are marked in orange. Cd-f) GFP expression driven by the pnt^{43H01} during stages 9-10B in the SC, border cells (BC), and P domains. Cg-i) GFP expression driven by the pnt^{45D11} during stages 9-10B in the BC and P domains. Additional stages can be found in Supplemental Material, Figure S2R.

To understand the overlap between the patterns of the GFP positive lines and the endogenous gene, all patterns were annotated as previously described (Figure 17D) (Niepielko et al., 2014). The annotation system is based on simple domains of gene expression that are induced by cell signaling pathways, including BMP (AD, AV, SC) and EGFR (M, D, P), and domains of future dorsal appendages (R, F) (Niepielko et al., 2014; Nilson and Schupbach, 1999; Peri and Roth, 2000; Twombly et al., 1996; Yakoby et al., 2008b). This system was developed to annotate FC patterning as a binary matrix, which allows the addition of domains found in our screen, including germarium (G), stalk cells (StC), border cells (BC), and polar cells (PC) (Figure 16A-C). The overlay of the GFP expression patterns and the endogenous gene patterns revealed that the majority of

recapitulated patterns are within the anterior (AD, AV), stretched cells (SC), posterior (P), and uniform (U) domains (Figure 17D).

Figure 17 D) A binary matrix representing all gene expression patterns (red) and FlyLight GFP positive lines (green). The overlap between the two data sets are denoted in yellow. The matrix is based on assigning mutually exclusive domains to patterns (Figure 30, and Supplemental Material, S2i,ii, Revaitis et al., 2017). Domains include, germarium (G), splitting the anterior domain to anterior dorsal (AD) and anterior ventral (AV), midline (M), roof (R), floor (F), dorsal (D), posterior (P), stretched cells (SC), stalk cells (StC), polar cells (PC), and uniform (U). Additional domains are included as not (/) for domain exclusions. The complete description of these domains can be found in Supplemental Material, Figure S2i and ii, Revaitis et al., 2017). On the Y axis are the gene names at a specified developmental stage. % recapitulation (% Recap.) represents the percent of GFP patterns that overlap with the endogenous pattern in each domain.

Numerous genes are uniformly expressed in the FCs during early oogenesis (Figure 17D). At the same time, the uniform “inducer” is still unknown. Several reporter lines are expressed in the border

cells (Figure 17D). With the exception of one line, none of the known associated genes were reported to be expressed in these cells. Since the border cells travel through the nurse cells, which turn dark during most *in situ* hybridization procedures, it is possible that gene expression in the border cells is masked by the dark nurse cells (Yakoby et al., 2008a). The roof and floor domains (R and F) are regulated jointly by multiple signaling pathways, including EGFR and BMP (Deng and Bownes, 1997; Ward and Berg, 2005; Ward et al., 2006). It is possible that the single floor and roof patterns found is due to the complex regulation of these domains that may require more enhancers working together (Figure 17D). The midline and dorsal domains were not found in our screen; these CRMs may reside in neighboring genes, or require longer DNA fragments, or are present in the gene locus but not covered by the screened fragments.

2.3.2 The FlyLight lines are new resource tool for gene perturbations in oogenesis

The GAL4-UAS system has been a valuable method to manipulate genes in *D. melanogaster* (Brand et al., 1994; Duffy, 2002). To increase the perturbation efficiency, it is necessary to refine and restrict the affected domain. The GFP positive lines present an opportunity to manipulate genes in a domain-specific manner. As far as we know, none of the previously published GAL4 lines are expressed only in the anterior domain, including the centripetally migrating FCs and stretched cells. Here, we used two of the anterior lines, *dad*^{44C10} and *dpp*^{18E05}, to determine their function in genetic perturbations. A limitation of the GAL4-UAS system is the undesired expression of some drivers in multiple tissues, which, in many cases, leads to lethality. Indeed, a complete lethality was observed when these lines were crossed directly to UAS lines of perturbations in EGFR signaling (not shown). Thus, we used a GAL80^{ts} to circumvent the problem.

The regulation of *dpp* during oogenesis is not fully understood. While an earlier study mapped numerous regulatory elements downstream of the 3' end of the *dpp* transcription unit, it did not report expression during oogenesis (Blackman et al., 1991). The posterior repression of *dpp* requires the activation of EGFR signaling (Peri and Roth, 2000; Twombly et al., 1996). Unlike *dpp*,

dad is a known target of BMP signaling (Marmion et al., 2013; Weiss et al., 2010). The expression patterns of *dad*^{44C10} and *dpp*^{18E05} are nearly identical (Supplemental Material, Figure S2D, F, Revaitis et al., 2017). However, if the two CRMs are regulated by different mechanisms, perturbations in cell signaling pathways may impact their activities in a different manner. In order to test this idea, we used the corresponding fragments of DNA in the *dad*^{44C10} and *dpp*^{18E05} (Figs. 17A, 19B, and Supplemental Material, Figure S1 and S2D, F, Revaitis et al., 2017) to generate *lacZ* reporter lines. As expected, the two reporter lines are expressed in a similar anterior pattern (Figure 18A, B). To test whether these reporter lines are regulated by BMP signaling, we crossed these lines to a fly expressing *dpp* in the posterior end of the egg chamber (E4>*dpp*). Interestingly, we detected ectopic posterior expression of β -Galactosidase in the *dad-lacZ* line, but not in the *dpp-lacZ* background (Figure 18C, D). Based on the β -Galactosidase results, we conclude that the two drivers are regulated differently, and thus, perturbations, in addition to affecting the tissue, may have a positive or negative impact on the drivers themselves. The results are not surprising as a *dpp* reporter was previously described to be regulated by EGFR and JAK/STAT signaling pathways (Xi et al., 2003).

Figure 18: Genetic perturbations using *dad*^{44C10} and *dpp*^{18E05} FlyLight lines. **A, B)** β -Galactosidase expression patterns of *dad*^{44C10} and *dpp*^{18E05} lines in the anterior and stretched cells domains (*dad-lacZ* and *dpp-lacZ*). **C, D)** Expression of *dpp* in the posterior end (*E4>dpp*) induces ectopic expression of β -Galactosidase expression in the posterior domain in *dad-lacZ* but not in the *dpp-lacZ* (denoted by a white arrow). BR staining is used as a spatial marker. Arrowhead denotes the dorsal midline. Broken yellow line denotes the anterior boundary of the oocyte. **E-H)** OreR **E)** Eggshell, **F)** pMad (green), **G)** BR (red), **H)** merge. **I-L)** *dad*^{44C10} driving the expression of a dominant negative EGFR (*dnEGFR*). **I)** Eggshell, **J)** pMad,

K) BR, **L)** merge. **M-P)** *dpp*^{18E05} driving the expression of a *dnEGFR*. **M)** Eggshell, **N)** pMad, **O)** BR, **P)** merge. **Q-T)** *dad*^{44C10} driving the expression of a constitutively active EGFR (*caEGFR*). **Q)** Eggshell, **R)** pMad, **S)** BR, **T)** merge. **U-X)** *dpp*^{18E05} driving the expression of *caEGFR*, **U)** no eggshell, **V)** pMad (white arrow denoted the anterior boundary of the future oocyte-associated FCs, also in W), **W)** Merged image of pMad and BR (a separate BR image is not shown). **X)** For comparison, we included the wild type (*OreR*) merged BR/pMad image at S9. We note that oogenesis stopped at stage 9 in the *dpp*^{18E05}>*caEGFR* background. No pMad is present in egg chambers. Egg chambers' developmental stages are denoted. All images are a dorsal view and anterior is to the left.

Eggshell structures are highly-sensitive to changes in EGFR signaling (Neuman-Silberberg and Schupbach, 1993; Queenan et al., 1997; Yakoby et al., 2005). Therefore, we aimed to demonstrate the use of the two drivers to disrupt EGFR signaling and monitor the impact on eggshell structures and egg chamber patterning. Each driver was crossed to a dominant negative EGFR (*dnEGFR*) and a constitutively activated EGFR (*caEGFR*). We looked at patterning of (Broad - BR), BMP signaling (pMad), and eggshell structures. *D. melanogaster* eggshell has two long dorsal appendages. At stage 10B, pMad appears in three rows of cells in the anterior domain, while BR is expressed mostly in two dorsolateral patches on either side of the dorsal midline (Figure 18E-H) (Deng and Bownes, 1997). The eggshell of *dad*^{44C10}>*dnEGFR* has an elongated narrow operculum and two shortened dorsal appendages (Figure 18I). Interestingly, the pattern of pMad and BR remained in one and two rows, respectively, of cells in the anterior domain (Figure 18J-L). The *dpp*^{18E05}>*dnEGFR* generated a short eggshell with a large and wide operculum and two short dorsal appendages (Figure 18M). Unlike the anterior domain of pMad and BR in Figure 18L, this perturbation led to ectopic pMad in the anterior domain but not BR (Figure 18N-P). An activation of EGFR in the posterior domain represses *dpp* expression (Twombly et al., 1996). Following the same logic, overexpression of *dnEGFR* alleviates the anterior repression of *dpp*, and consequently increases BMP signaling and the operculum size (Dobens and Raftery, 1998).

Over-activation of EGFR signaling in the anterior domain with the two drivers generated different phenotypes. In *dad^{44C10}>caEGFR*, the eggshell was short with reduced operculum that extends to the ventral domain (Figure 18Q), which is expected for an increase in anterior EGFR activation (Queenan et al., 1997). The pMad pattern was shifted anteriorly over the stretched cells. Also, BR was shifted anteriorly (Figure 18R-T). These phenotypes indicate that in addition to the increase in EGFR signaling, there is also a decrease in BMP signaling (Yakoby et al., 2008b). Interestingly, *dpp^{18E05}>caEGFR* did not produce any eggs (Figure 18U). The ovarioles and egg chambers of flies grown at 30°C appeared deformed (not shown). Reducing the temperature to 28°C allowed the egg chambers to develop up to stage 9. However, pMad could not be detected (Figure 18V,W), in comparison to the corresponding pMad pattern in the wild type at this developmental stage (Figure 18X). This cross was repeated five times and all egg chambers ceased to develop at stage 9. These results are consistent with the previously published decrease in BMP signaling: medium decrease generated short eggshells, while a strong decrease stopped egg chamber development at stage 9 (Twombly et al., 1996). These results further support the negative regulation of *dpp* signaling by EGFR activation (Xi et al., 2003).

2.3.3 Mapping the distribution of CRMs in the gene model

To date, the computational prediction of CRMs has not been straightforward. Since the FlyLight fragments cover the entire length of the gene, we aimed to determine whether certain locations of the gene locus are more likely to contain CRMs. We binned the distributions of all GFP positive FlyLight lines into three groups based on their relative position to the first exon of the gene model (Figure 19A). All DNA fragments that are upstream of the first exon were classified as *Proximal*. All fragments that are downstream of the first exon and within the first intron are categorized as *Intron 1*, and all other downstream fragments are classified *Distal*. One problem with this analysis is that several genes have multiple isoforms with different locations of the first exon. For example, the *ana* gene has three isoforms, two have the same first exon (*ana*-RA and *ana*-RB) and the third

(*ana*-RC) has a different first exon (Figure 19B). Since no information is available on the oogenesis-specific isoform(s), we carried out an RNA-seq analysis of egg chambers at three developmental groups. Specifically, egg chambers were collected at early (stages ≤ 9), middle (stages 10A-B), and late (stages ≥ 11) stages of oogenesis. We found that the *ana* gene has only two isoforms (*ana*-RA and *ana*-RB) that could be expressed during oogenesis, both have the same transcription start site (TSS) (Figure. 19B). The RNA-seq analysis eliminated discrepancies among isoform transcripts for nine additional genes (Revaitis et al., 2017 Supplemental Material, Figure S2, Revaitis et al., 2017).

*Figure 19: Mapping distribution of FlyLight gene fragments A) A cartoon representation of the gene-fragments' binning that is based on the relative position of a fragment in the gene model. The Proximal includes all fragments that are upstream to the first exon. The Intron 1, includes all fragments that are in the first intron. The Distal includes all fragments downstream of the second exon. B) An example of one of the genes screened, ana, and its three isoforms. The gene has two "first" exons. Using RNA-seq data, we demonstrate in the coverage plot that the ana-RA and/or ana-RA isoforms (marked by *) are expressed during all stages of oogenesis, whereas ana-RC is not expressed. The RNA-seq data is divided into three developmental groups (stage 9 and younger, stages 10A and 10B, and stages 11 and older egg chambers). Peaks indicate the number of reads per base. Gray peaks indicate matched base pairs and colored peaks*

indicate mismatches (see M&M for details). The fragments mapped below the model were screened. The analysis shows that *ana*^{23E11} fragment (orange) is in the *Distal* bin. **C)** A Chi-square test shows that the GFP-expressing FlyLight fragments are distributed significantly different ($p=0.034$, $df=2$) from the expected distribution.

Next, we tested the distribution of the GFP-positive FlyLight lines in the three categories (*Proximal*, *Intron 1*, and *Distal*). The null hypothesis is that the frequency of CRMs among the categories is identical. In this case, the observed frequency of positive CRMs is equal to the expected frequency of positive CRMs for the total number of DNA fragments for each of the three categories. The expected distribution is calculated as the percent of the number of GFP positive lines (54) out of the total number of lines (281), which is 19%. The *Proximal* category includes 140 lines. The expected number of lines expressing GFP in this category is 27. The observed number of GFP expressing lines is 23, which is 15% less than the expected value (Figure 19C). The *Intron 1* category includes 72 lines, thus the expected number of lines expressing GFP is 14. The observed number of GFP expressing lines is 22, which is 57% more than the expected value (Figure 19C). The *Distal* category includes 69 lines, and the expected number of lines expressing GFP is 13. The observed number of GFP expressing lines is 9, which is 31% less than the expected value (Figure 19C). Using a Chi-square test, we determined that the distribution of the observed values is significantly different from the expected values ($p=0.034$, $df=2$).

To determine whether the distribution of CRMs is significantly different than the expected value for each category, we used a binomial test, which checks the significance of deviation between two results. As stated above, the calculated success rate (positive GFP expression) is 19%. Based on this rate, we employed a one-tailed binomial test for each category. The probability that 23 or less out of the 140 fragments in the *Proximal* category will drive GFP expression is 0.25. The probability that 22 or more out of the 72 fragments in the *Intron 1* category will drive GFP expression is 0.012. The probability that 9 or less out of the 69 fragments in the *Distal* category will drive GFP expression is 0.13. Based on the binomial test, we conclude that the number of

CRMs in the *Intron 1* category is significantly greater ($p=0.012$) than the expected number. We note that the average DNA fragment-size for the *Proximal* is 3.17 ± 0.06 kb, for the *Intron 1* is 3.06 ± 0.1 kb, and for the *distal* is 2.56 ± 0.13 kb (fragment size \pm SE kb). While the average size of the *Distal* fragments is significantly shorter than the fragments in the other two categories ($p<0.01$), we do not consider this difference to be the cause for the CRM-enrichment in *Intron 1*, since shorter fragments were found to be more active (Nam et al., 2010). Also, the average fragment size is not significantly different between *Proximal* and *Intron 1* categories.

2.3.4 The FlyLight lines control expression in multiple tissues

On average, the FlyLight lines contain ~3kb fragments of DNA. We were interested to understand whether these fragments contain one or more CRMs regulating expression during oogenesis. Assuming that each domain is regulated by a CRM, we aimed to determine if each fly line controls one or more domains during oogenesis. Each GFP expression domain was counted once for each line per developmental stage for a total of 339 ‘line patterns’. Interestingly, only 35 ‘line patterns’ (10.3%) changed to a different pattern in the next developmental stage (30 ‘line patterns’) or at a later stage that is not the next stage (5 ‘line patterns’) (Figure 20A). Most (34 of the 35) ‘line patterns’ changes are found after stage 8, which is the transition stage from anterior-posterior axis to a dorsal-ventral axis determination as a result of changes in the position of EGFR signaling (Neuman-Silberberg and Schupbach, 1993). Interestingly, 20 ‘line patterns’ change at stage 14, the stage of ovulation. Hence, most domains, once expressed, maintain the same pattern over multiple developmental stages (Figure 20B). CRMs are 500-1000 bp (Ivan et al., 2008; Levine and Tjian, 2003). Under the assumption that all fragments contain CRMs, it is expected that each fragment has 3 or more CRMs. Our results suggest that the ~3kb DNA fragments mostly contain one (48%) or two (33%) CRMs that control the expression of a simple pattern during oogenesis (Figure 20C). Many CRMs are tissue-specific, therefore we wanted to determine whether the same fragments control expression in other tissues (Jenett et al., 2012; Jory et al., 2012; Li et al., 2014; Manning et

al., 2012). Only 8 out of the 281 FlyLight lines are not expressed in any of the examined tissues (Figure 20D). A primary portion of the lines (~50%) are expressed in four tissues (Figure 20D). These analyses cannot distinguish whether the same CRM is expressed in multiple tissues or if each fragment contains numerous CRMs that are expressed in different tissues. Of the FlyLight lines positive for GFP, 75-85% are expressed in other tissues, including the brain, ventral notochord (VNC), larva central nervous system (CNS), and embryo (Figure 20E). Only ~20% of the lines are also expressed in imaginal discs (Figure 20E, and Supplemental Material, Figure S3, Revaitis et al., 2017). These results support our observation that lethality of the tested lines in perturbations without GAL80^{ts} is likely related to the expression of these lines in multiple tissues.

Figure 20: Spatial and temporal distribution of GFP-positive Flylight lines. A) A summary of the temporal distribution of GFP-positive FlyLight patterns throughout oogenesis. Each box represents a developmental stage and the number of lines expressing GFP. Horizontal arrows represent the number of lines with the same spatial pattern in the next developmental stage (refers in the text as 'line patterns'). Diagonal bottom arrows represent the number of lines not expressed in the next developmental stage. The diagonal top arrows represent the new lines expressed in this developmental stage. Horizontal broken arrows represent the number of lines expressed in the next developmental stage that change their spatial pattern. Diagonal broken arrows represent lines expressed in an early developmental stage, and now are expressed in a later developmental stages that is not the proximal stage (the pattern is spatially

different from the earlier pattern). **B)** For each domain, the average number of developmental stages it is expressed in and the standard deviation were calculated. **C)** For each of the GFP positive FlyLight line, the number of expression domains was calculated. The data is presented as % of the total lines expressing GFP for all developmental stages. **D)** Based on the available data for the Flylight expression patterns, we determined the frequency of expression of each of the 281 lines in the five FlyLight tissues and oogenesis (total of six tissues). **E)** Of the 281 lines screened, 84% are expressed in the brain, 78% in the ventral nerve cord, 85% in the larval CNS, 77% in the embryo, 20% in the third instar larvae imaginal discs, and 19% in the ovary.

Chapter 3 Conclusions

3.1 A comprehensive model to study parameters affecting the distribution of GRK/dpERK

In section 2.1, our comprehensive model provides new insights into the mechanisms that contribute to the distribution of GRK and EGFR signaling. First, we were able to develop experimental approaches to determine values for parameters that have so far only been measurable in culture cells (Pribyl et al., 2003a). These parameter values were then applied to a highly dynamic model that considers the mechanical evolution of the egg chamber: its overall growth, the shifting epithelial cells, and a moving morphogen source. Published literature models the distributions of GRK and dpERK at a steady state in fixed coordinates (Goentoro et al., 2006; Zartman et al., 2011). Our model took a new approach by building on the dynamic outputs from earlier stages to achieve an overall distribution at S10A. Using a multistage approach allowed us to elude to the mechanisms behind the posterior extension of GRK after the morphogen source has moved anteriorly.

To establish our dynamic model, we had to establish measurements of the oocyte, egg chamber, and the source of GRK. Observations in EGFR signaling dynamics and overall growth made it necessary to piece apart stages 8 and 9 into additional steps, early and late. Furthermore, we were able to discretize values in current literature and build on previous models for the dynamics of the EGFR, diffusion, internalization of ligand and receptor, and FC counts (Chu et al., 1996; Goentoro et al., 2006; Pribyl et al., 2003a; Sigismund et al., 2008; White et al., 2009; Zartman et al., 2011). The unknown parameters were approximated by running numerous simulations against 1-D intensity plots. After establishing parameters based on 1-D experimental plots we were able to simulate the distribution of dpERK on a growing manifold.

Since the model accounts for the physical movement and growth of the egg chambers' compartments, we could answer questions related to the mechanics of tissue morphogenesis. Specifically, we tested the impacts on signaling in cases when the morphological changes will be perturbed. An increase in localized signal is predicted in both the absence of oocyte nucleus movement and in the absence of oocyte growth. This is unlike the reduced signaling in case of no

FCs movement. We reason that the movement of “new” FCs with high EGFR could maintain the signal close to the source in the first two perturbations, unlike the overuse of “old” FCs in the latter perturbation. This is consistent with our findings in EGFR dynamics (Aim 2).

The model development for the spatio-temporal evolution of GRK and dpERK in a highly dynamic egg chamber has led us to identify the need for a suitable mathematical framework for reaction-diffusion equations in developing organisms. In this work, the growth of the egg chamber is a morphological change that affects the distribution of Gurken through the time-evolving Laplace-Beltrami operator. The rigorous development of the coupling of growth and diffusion was done in a parallel work, in which we give a more general setting: the diffusing quantity (*i.e.* the morphogen) also affects the growth of the organism (Pouradier-Duteil et al., 2015; Rossi et al., 2016). This new framework for the two-way coupling between diffusion and growth has many potential applications in models of development, and for this reason we named the new class of equations *Developmental Partial Differential Equations*.

3.2 A quantitative analysis of the EGFR in *Drosophila* oogenesis

The levels of EGFR receptors in tissues is an important aspect of appropriate ligand distribution, cell signaling, tissue patterning and morphogenesis. The uniform pattern of EGFR in oogenesis was reported by the Shilo group (Sapir et al., 1998), however, these are qualitative description of the pattern. Most of the quantitative measurements of EGFR have been reported out of cancer research, where mis-regulation of EGFR signaling is of great importance (Zhang et al., 2015). In addition, a recent study in mice reported the patterning of a EGFR-EGFP construct in many tissues (Yang et al., 2017). However, like in *Drosophila*, this is a beautiful but qualitative study. Here, we developed a tool to close the current gap between qualitative and quantitative studies in tissues. This tool allowed us to better visualize the localization of the EGFR in the tissues of *D. melanogaster*, as well as to determine the levels and dynamics of EGFR in the tissue.

We found that during early oogenesis, the EGFR is apically localized in the posterior side of the FCs. The loss of this restriction in a *grk* null background suggests that the default distribution is

basolateral. This mechanism is consistent with what was reported in *C. elegans*, the receptor is moving away from the ligand source to lateral domains of the cells (Haag et al., 2014). The actual mechanism in which the EGFR is transcytosed throughout the cell is still not known. Interestingly, the apical localization disappears as the amount of receptor increases.

It is known that upon EGFR activation, the receptor enters the endocytic cycle at the apical side of the FCs (Chang et al., 2008). However, receptor dynamics have been shown to change depending on ligand (Zhang et al., 2015), therefore, how the intracellular processes respond to GRK is not clearly outlined. Our results suggest that the level of EGFR increases after the ligand source leaves the posterior where EGFR activation was once high. The significance in the increase of the receptor amount is contrary to a dorsal anterior decrease in the spatial distribution of the receptor at S10A (Sapir et al., 1998). This decrease is restricted to the dorsal anterior in a small number of the egg chambers. It is also not quantifiable along the AP when we re-examined the spatial distribution of the receptor, regardless of the visual appearance. A possible reason is that most of the EGFR is localized to compartments inside of the cell which may give the qualitative appearance of a restriction of EGFR at the dorsal anterior.

The co-localization of the EGFR to the endosomal compartments elucidates the intracellular dynamics of the EGFR in the absence and presence of ligand. Auto-activation of the EGFR in the absence of ligand is possible and even visible in tiny punctae that were detected in the *grk* null EGFR. However, most auto-activated receptor was shown to recycle back to the plasma membrane (Tanaka et al., 2017). Our analysis detected a relative higher co-localization of the EGFR with Rab7 at the posterior, suggesting this to be residual from previous stages. This further suggests that the degradative process is slow in low GRK domains. The potential uses of the EGFR-GFP line are extensive. We demonstrate that EGFR is detected in the correct patterns in imaginal discs. It will be interesting to look at the pattern of EGFR during embryogenesis. Hence, the quantitative analysis done in the FCs can be repeated in other tissues, providing quantitative measurements for the levels (and dynamics) of EGFR during development. This tool can provide powerful

information about the EGFR and the different ligands' interactions throughout development. It can also provide a valuable quantitative information about the cellular compartments of the receptor and ligands co-localization during development.

3.3 Analysis of CRMs in *Drosophila* oogenesis

Lastly, we developed a comprehensive screen using the FlyLight collection allowed us to search for CRMs during oogenesis (Revaitis et al., 2017). This collection contained a wealth of regulatory information that could be mapped spatially and temporally to better understand CRM activity on a broader scale. We were able to take advantage of the placement of the FlyLight fragments and their comparable size to determine that CRM activity is greater in the first intron during oogenesis. The CRM activity could be mapped temporally so that we can determine the average length of time for an active CRM during development.

Interestingly, genes that are co-expressed in the same compartment at the same developmental stage are considered to share a regulatory mechanism. Our FlyLight screen produced CRMs showing activity in the same cellular compartments. We demonstrated, using perturbations of the BMP and EGFR signaling pathways, that the regulation of these enhancers could be under different regulatory mechanisms. In particular, the example of the genes *dad* and *dpp* showed different responses. These responses were also different when these constructs were used to perturb EGFR signaling in the FCs. The gene, *dad*, is an inhibitor of the BMP signaling pathway, as opposed to *dpp*, whose regulation is unknown. Perturbations using a UAS-*dpp* construct only affected *dad*^{44C10}, but perturbations involving a constitutively active receptor at the anterior proved to be lethal in the *dpp*^{18E05} construct. To add to this complexity, our lab found that the type II receptor, wishful thinking, which is expressed in the same pattern, is regulated by alleviation of BRK repression (Marmion and Yakoby, 2018). Hence, genes that are expressed in the anterior domains in near identical patterns are regulated by BMP (via PMAD or BRK) or by other pathways (EGFR or JAK/STAT) (Xi et al., 2003).

Chapter 4 Materials and Methods

4.1 Fly Stocks and strains

Flies used in this study w, nos-phiC31 double balanced, GMR43H01 (*pnt*)-Gal4 (BSID 47931) (Pfeiffer et al., 2008, Revaitis et al., 2017), UAS-Rab5-GFP, UAS-RAb7-GFP (A generous gift from Gonzalez-Gaitan), GRK null, yw; FRT82B Δ Ras, Cy2-GAL4 (Queenan *et al.*, 1997), UAS-styRNAi (DRSC, BSID Stock #36709(Flockhart et al., 2012)), UAS-EGFR RNAi (DRSC, BSID#36770 (Flockhart et al., 2012)), 4PX, w- *D. melanogaster grk null* [2b]b, grk null [2E12]b (gifts from Trudi Schüpbach, Howard Hughes Medical Institute, Princeton University, Princeton, NJ, USA), X7;28.20 (Neuman-Silberberg and Schüpbach, 1993). All crosses and fly stocks were grown on cornmeal agar and maintained at 23°C.

The Flylight lines (Pfeiffer et al., 2008; Pfeiffer et al., 2010) were obtained through Bloomington *Drosophila* stock center, Indiana University. All tested FlyLight stocks are listed in Supplemental Material, Figure S1 in Revaitis et al., 2017. FlyLight lines (males) were crossed to P[UAS-Stinger]GFP:NLS (Barolo et al., 2000) virgin females. In order to overcome lethality associated with genetic perturbations, FlyLight lines were first crossed to a temperature sensitive GAL80, P[tubP-GAL80^{ts}]10 (Bloomington ID# 7108). The *dad-lacZ* and *dpp-lacZ* reporters (see below) were crossed to E4-GAL4 (Queenan et al., 1997), and a UAS-*dpp* (a gift from Trudi Schüpbach). EGFR signaling was upregulated by a UAS- λ top-4.2 (caEGFR (Queenan et al., 1997)) and downregulated by a UAS-dnEGFR (a gift from Alan Michelson). Progeny were heat shocked at 28°C for three days to alleviate repression by GAL80^{ts}.

The *dad-lacZ* and *dpp-lacZ* reporters were constructed based on the *dad*^{44C10} and *dpp*^{18E05} DNA fragments. The coordinates for these DNA fragments were taken from <http://flweb.janelia.org/cgi-bin/flew.cgi>. These fragments were amplified from OreR using phusion polymerase (NEB), A-tailed with Taq, and cloned into a PCR8/GW/TOPO vector by TOPO cloning (Invitrogen). The fragments were then Gateway cloned into a pattBGWhZn (Marmion et al., 2013). Both reporter

constructs were injected into the attP2 line (Stock# R8622, Rainbow Transgenic Flies, Inc) and integrated into the 68A4 chromosomal position by PhiC31/attB mediated integration (Groth et al., 2004).

To generate the EGFP construct, an EGFP tag was inserted between the last exon and the 3' UTR on the C terminus of the protein in a pHd-dsRed (Gratz et al., 2014) donor plasmid. This construct was generated containing two 1 kb homologous regions directly adjacent to the target sites for gRNAs. Guides were designed in the first intron adjacent to second exon and downstream of 3'UTR in a pU6-BbsI-chiRNA plasmid (Gratz et al., 2013). These homologous regions flank a replacement region of the EGFR with an EGFP tag placed between the last exon and the 3' UTR. To generate the sfGFP construct, a single guide was designed directly upstream of the 3' UTR of the EGFR. Two 1 kb homology arms flanking the gRNA were cloned into a pHd-sfGFP-scarless vector (Addgene, #80811). Flies were injected with both plasmids(Rainbow Transgenic Flies, Inc). To generate the *grk* null construct, a pHd-dsRed donor plasmid was generated containing two 1 kb homologous regions directly adjacent to the target sites for gRNAs. Guides were designed in the 3'UTR and 5' proximal region in the intergenic space. Flies were injected with both plasmids(Rainbow Transgenic Flies, Inc).

For all CRISPR constructs, adult flies surviving injection were crossed to a *w*- strain and screened for the selective marker dsRed. Flies that were positive for the dsRed selection marker were verified by PCR and sequenced.

4.2 Immunohistochemistry

Immunoassays were performed as previously described (Yakoby et al., 2008b). In short, flies 3-7 days old were put on yeast and dissected in ice cold Grace's insect medium, fixed in 4% paraformaldehyde, washed three times, permeabilized (PBS and 1% Triton X-100), and blocked for 1 hour (PBS and 0.2% Triton X-100 and 1% BSA). Ovaries were then incubated overnight at 4°C with primary antibody. After washing three times with PBST (0.2% Triton X-100), ovaries

were incubated in secondary antibodies for 1 hour at 23°C. Then, ovaries were washed three times and mounted in Flouromount-G (Southern Biotech). Immunoassays for dpERK were performed as previously described (Zartman *et al.*, 2009). Primary antibodies used were sheep anti-GFP (1:5000, Serotec), Gurken anti-mouse (1:10, DSHB) MapK (dpERK) anti-rabbit (1:100, Cell Signaling Technologies), Rab5 anti-Rabbit (1:50, a gift from Gonzalez-Gaitan), Rab7 anti-mouse (1:40, DSHB), rabbit anti-beta Galactosidase (1:1000, Invitrogen) (Yakoby *et al.*, 2008b), mouse anti-Broad (BR) (1:400, stock #25E9.D7, Hybridoma Bank), and rabbit anti-phosphorylated-Smad1/5/8 (1:3600, a gift from D. Vasilias, S. Morton, T. Jessell and E. Laufer) (Yakoby *et al.*, 2008b). Secondary antibodies used were Alexa Fluor 488 (anti-mouse), Alexa Fluor 488 (anti-sheep), Alexa Fluor 568 (anti-mouse), Alexa Fluor 647 (anti-mouse), and Alexa Fluor 568 (anti-rabbit) (1:2000, Molecular Probes). Antibodies against Gurken and Rab5 were preabsorbed for 1 hour at room temperature. Nuclear staining was performed using DAPI (84 ng/ml). All immunofluorescent images were captured with a Leica SP8 confocal microscope (Rutgers University Camden, confocal core facility).

In the CRM screen, the pattern of BR was used as a spatial reference to characterize the dorsal side of the egg chamber. Two ovaries from an internal positive control, *rho*^{38A01}, were added in each immunoassay due to the unique expression pattern in the border cells (In Revaitis *et al.*, 2017 see Supplemental Material, Figure S2Tc). For SEM imaging, eggshells were mounted on double sided carbon tape and sputter coated with gold palladium for sixty seconds. Images were taken using a LEO 1450EP.

4.3 EGFR protein quantification

To quantify the amount of GFP in the ovary and embryo, a GFP ELISA kit (Abcam AB171581) was used and the manual was followed regarding antibody dilutions and generating GFP standard curve. For egg chamber protein quantification, ovaries from *yw*; EGFR-EGFP and *yw* (control) were dissected into ice cold Grace's medium. After teasing and separating, 100 stage specific egg

chambers at stages 8,9, stage 10A and 10B, and stages >11 were collected in three biological replicates and added to 150 μ L cold NP-40 lysis buffer treated with 1X protease inhibitor cocktail (ThermoFisher). Egg chambers were sonicated in lysis buffer and agitated for one hour at 4C. The protein quantification of the embryo was performed at 1 and 2.5 hrs after fertilization. One hundred embryos at each developmental time were collected and added to 150 μ L of NP-40 lysis buffer with 1x protease inhibitor cocktail. Embryos were homogenized using a dounce homogenizer and agitated at 4C for one hour.

Both egg chamber and embryo samples were then centrifuged and 150 μ L of supernatant was transferred to a new Eppendorf tube. Total protein was quantified using a BCA assay. An ELISA kit was obtained using precoated wells (AbCam) and 9.375 μ g of total protein in 50 μ L lysis buffer was added to each well. Each stage was run in biological triplicates and two technical replicates. Concentrations were calculated using online software (elisaanalysis.com). Since the amount of GFP in our sample is unknown, we first ran kit with total protein extracted from whole ovaries at different weights determined from a BCA assay kit was used to determine appropriate amount of total protein. Amounts total protein were run at different dilutions and amount of total receptor was calculated using 75ug total protein.

4.4 Microscopy and Co-localization Analysis

Images were acquired using a Leica SP8 confocal microscope. Image acquisition was taken in between frames. Confocal stacks were separated into individual images (z-distance between images used was 1.38 μ m) using For Its Just ImageJ (FIJI) software where five to eight images per stack were analyzed and values for co-localization were averaged. Individual .tif files containing two channels at the same Z position were imported into Matlab. Pixel intensity values within images were multiplied to one another and a binary threshold was applied for pixel value, n , where $n = 1$ if greater than 96% maximum intensity, otherwise $n=0$. Spots were filtered by size spot size, s , in pixels where $s = 2 < s < 50$. Select regions of interest were analyzed in sagittal views of the

FCs along the entire dorsal side in images referred to as dorsal, in the five FCs overlying the anterior-most region of the oocyte for images referencing the anterior, or five posterior-most FCs on the dorsal side of the oocyte for images referencing the posterior. The number of spots co-localized were divided by the total number of spots detected for the GFP.

To validate the Rab5/7 antibodies, UAS constructs for Rab5/7 tagged with GFP were expressed using a posterior driver (*pnr^{43H01}*-GAL4). Images were acquired at S10A and processed as described above.

4.5 RNA-seq analysis

The specific isoform expressed during oogenesis for each of the 22 genes was identified using RNA-sequencing (RNA-seq) analysis (Supplemental Material, Figure S2, Revaitis et al., 2017). Egg chambers were analyzed at three developmental stages: 1) egg chambers at stages 9 or earlier, 2) egg chambers at stages 10A and 10B, 3) egg chambers at stages S11 or greater. Egg chambers' isolation was done manually as previously described (Yakoby et al., 2008a). All RNA samples (~200 egg chambers from each developmental group) were extracted using RNeasy Mini Kit (QIAGEN, Valencia, CA). One microgram of total RNA from each sample was subject to poly-A containing RNA enrichment by oligo-dT bead and then converted to RNA-seq library using the automated Apollo 324TM NGS Library Prep System and associated kits (Wafergen, CA), according to the manufacturer's protocol, utilizing different DNA barcodes in each library. The libraries were examined on Bioanalyzer (Agilent, CA) DNA HS chips for size distribution, and quantified by Qubit fluorometer (Invitrogen, CA). The set of 3 RNA-seq libraries was pooled together at equal amount and sequenced on Illumina HiSeq 2500 in Rapid mode as one lane of single-end 65nt reads following the standard protocol. Raw sequencing reads were filtered by Illumina HiSeq Control Software and only the Pass-Filter (PF) reads were used for further analysis. PF Reads were de-multiplexed using the Barcode Splitter in FASTX-toolkit. Then the reads from each sample were mapped to dm3 reference genome with gene annotation from FlyBase using TopHat 1.5.0 software.

Expression level was further summarized at the gene level using htseq-count 0.3 software, including only the uniquely mapped reads. Data were viewed using IGV software (Robinson et al., 2011; Thorvaldsdottir et al., 2013). The RNA-seq alignments show the coverage plot of each of the screened genes aligned to the reference genome gene track(s) (Figure 4, Supplemental Material, Figure S2Aa-Va). The peaks in the coverage plot represent the number of reads per base pair. The color code represents miscalls or SNPs to the reference genome. In these cases, red, blue, orange, and green represent cytosine, thymine, guanine, and adenosine mismatches, respectively, and gray represents a match. The RNAseq data are available here <http://dx.doi.org/doi:10.7282/T3ZS300V>.

4.6 Statistical analysis of DNA fragments distribution

Fragments of DNA were divided into three bins: those upstream of the transcription start site (TSS) were categorized as *Proximal*, fragments within the first intron were categorized as *Intron 1*, and those downstream of the second exon were categorized as *Distal*. The TSS was assigned by selecting the longest isoform expressed, which was extrapolated from the RNA-seq data, unless otherwise noted from references in the literature. Introns shorter than 300 bp were not included in the FlyLight collection (Pfeiffer et al., 2008). Statistical analysis was performed using a Chi-square test ($n=3$, thus $df=2$). The null hypothesis is that the frequency of the CRMs among categories is identical. Pairwise testing for enrichment of specific categories was determined using a one-tailed binomial test, where the number of expected GFP expressing lines was 22%, and the size for each of the categories was 140, 72, and 69 for *Proximal*, *Intron 1*, and *Distal*, respectively. The observed numbers of GFP expressing fragments were 23, 22, and 9 for *Proximal*, *Intron 1*, and *Distal*, respectively. The p -values were calculated in MatLab using the myBinomTest (s,n,p,Sided) script available at MathWorks.com.

4.7 Pattern annotation and matrix formation.

We adopted the previously developed annotation system for FCs' patterning (Niepielko et al., 2014; Yakoby et al., 2008a). Briefly, the annotation of gene-patterning is based on simple domains,

primitives, which repeat across different expression patterns. The assembly of primitives provides a tool for the description of complex gene expression patterns in the FCs. Each domain is coded into a binary matrix as 0 (no expression) or 1 (expression), which allows us to simply add new domains into the matrix. In our screen, in addition to different domains in the FCs, other domains were added, including stretched cells, border cells, polar cells, and the germarium (Figure 30A-C). In addition, we added two new domains, the dorsal appendages and operculum, for stage 14, that are used for the calculations in Figure 19. These domains are presented in Supplemental Material, Figure 2 in Revaitis et al., 2017. The annotations of the endogenous gene patterns and the patterns of GFP positive FlyLight lines were performed by three independent researchers. Each pattern was annotated into an excel spreadsheet using a binary system for each domain. The annotations for individual lines were collapsed to represent one input for that gene at stages where the endogenous pattern is detected. To determine the overlap between the FlyLight GFP expression pattern and the endogenous expression pattern of the gene, the matrix of GFP positive domains was compared to the *in situ* hybridization matrix (the sources of these expression patterns are included in the captions of Supplemental Material, Figure S2, Revaitis et al., 2017). Matrices' overlay was done in MatLab using *imagesc*.

4.8 Construction of the mathematical model

Our mathematical model accounts for five main mechanisms whose complex interplay generates the spatiotemporal distribution of dpERK during stages 7 to 10A of oogenesis. Hence, we model simultaneously:

- the diffusion of ligand in the perivitteline space
- the reactions between ligand, receptors, complexes, inhibitors and signal
- the movement of the oocyte nucleus, considered to be the source of ligand
- the growth of the egg chamber
- the shift of the overlying FCs

In order to explain the model clearly despite its complexity, we construct it by exposing each of these mechanisms one by one.

4.8.1 Diffusion of ligand in the perivitteline space

The main assumption of our model is that the width of the perivitteline space is negligible compared to the other dimensions of the system (such as the anterior-posterior or dorso-ventral dimensions of the egg-chamber). As a consequence, we consider that Gurken diffuses on a curved surface, which approximates the perivitteline space. The diffusion of the ligand L is then given by the following partial differential equation:

$$\frac{\partial L}{\partial t} = D \Delta_{\text{surf}} L$$

where Δ_{surf} represents the Laplace-Beltrami operator of the surface, and D is the diffusion rate.

4.8.2 Reactions between ligand, receptors, complexes, inhibitors and signal

Many parameters are involved at the interface of the perivitteline space and FCs. Previous models have considered the concentrations of ligand and of ligand-receptor complexes (Pribyl et al., 2003a) and variable concentrations of cell receptors (Pribyl et al., 2003b). Also, in tissue culture, there exist detailed models involving intact ligand, degraded ligand, surface receptors, internalized receptors, surface ligand-receptor complexes and internalized ligand-receptor complexes (Chu et al., 1996). We position ourselves at an intermediate level of complexity, and choose to model the concentrations of ligand, cell receptors, surface ligand-receptor complexes, internalized ligand-receptor complexes, and signal. This choice allows us to take into account feedback mechanisms through the action of inhibitors, as well as the dynamic recycling and degradation of internalized complexes.

All variables are function of time t and of two space coordinates (η, θ) that parameterize the surface. The system of reaction-diffusion equations can be written as:

$$\begin{cases} \frac{\partial L}{\partial t} = D \Delta_{\text{surf}} L + V - \frac{1}{H} \bar{k}_{\text{on}} RL + k_{\text{off}} C \\ \frac{\partial C}{\partial t} = \frac{1}{H} \bar{k}_{\text{on}} RL - (k_{\text{off}} + k_{\text{ec}}) C + \alpha_{\text{rec}} k_{\text{rec}} C_i \\ \frac{\partial C_i}{\partial t} = k_{\text{ec}} C - \alpha_{\text{deg}} k_{\text{deg}} C_i - \alpha_{\text{rec}} k_{\text{rec}} C_i \\ \frac{\partial R}{\partial t} = -\bar{k}_{\text{on}} RL + k_{\text{off}} C - k_{\text{er}} R + Q_r \\ \frac{\partial S}{\partial t} = \bar{k}_S C_i - k_d S \end{cases} \quad (1)$$

where $L = L(t, \eta, \theta)$ represents the concentration of ligand (GRK), $C = C(t, \eta, \theta)$ is the concentration of surface ligand-receptor complexes, $C_i = C_i(t, \eta, \theta)$ is the concentration of internalized ligand-receptor complexes, $R = R(t, \eta, \theta)$ is the concentration of surface receptors (EGFR) and $S = S(t, \eta, \theta)$ is the concentration of signal (dpERK).

All parameters of the model are either taken from literature or measured experimentally, see Table 1. The flux of ligand $V(t, \eta, \theta)$ is a function of time and space and determined by the observed shape and position of the source of ligand.

At this level of model construction, we have a system of ODEs-PDE (respectively *ordinary differential equations* and *partial differential equation*). Indeed the evolution of the surface complexes, internalized complexes, receptors and signal are so far described by pointwise dynamics. On the other hand, the evolution of ligand is modeled by a reaction-diffusion equation.

Inhibitors:

In addition to the dynamics given in the system of equations (1), we model the action of two EGFR inhibitors: Kekkon1 (Kek1) and Sprouty (Sty). Kek1 targets EGFR dimerization, thus reducing GRK uptake and leaving higher levels of free ligand. To account for this effect, we modify the receptor-ligand binding rate via Michaelis-Menten kinetics, as follows:

$$\tilde{k}_{\text{on}}(t, \eta, \theta) = \frac{k_{\text{on}}}{1 + \gamma_{\text{Kek}} K \text{Kek}(t, \eta, \theta)}$$

The parameter γ_{Kek} is the strength of inhibitory feedback of Kek1, and the constant K is defined by: $K = V k_S / (k_d k_{\text{deg}})$.

The binding rate \tilde{k}_{on} is a space-time dependent variable, affected by Kek1. In the absence of Kek1, $\tilde{k}_{\text{on}} \equiv k_{\text{on}}$ and we recover the constant binding rate previously defined.

Sty acts on the intracellular components, affecting signal propagation (Figure 5). We include Sty's negative feedback network motif in the ODE for dpERK, also via Michaelis-Menten kinetics, with:

$$\bar{k}_s(t, \eta, \theta) = \frac{k_s}{1 + \gamma_{\text{Sty}} K \text{Sty}(t, \eta, \theta)}$$

where γ_{Sty} is the strength of inhibitory feedback.

In turn, as targets of the pathway, Sty and Kek1 depend on the concentration of dpERK and are modeled by standard linear kinetics:

$$\begin{cases} \frac{d \text{Sty}}{dt} = k^{\text{Sty}} S - k_d^{\text{Sty}} \text{Sty} \\ \frac{d \text{Kek}}{dt} = k^{\text{Kek}} S - k_d^{\text{Kek}} \text{Kek} \end{cases}$$

where k^{Sty} , k_d^{Sty} , k^{Kek} and k_d^{Kek} are the production rates and degradation rates of Sty and Kek1.

4.8.3 Ligand-Receptor Complex recycling

As represented in the system of reactions, a fraction α_{deg} of the internalized complexes goes to degradation, while the fraction $\alpha_{\text{rec}} = 1 - \alpha_{\text{deg}}$ is recycled and goes back to the membranes of the FCs to be reused. To model the amount of available receptor we consider a previous study that reported two different pathways for the EGFR internalization: clathrin-regulated endocytosis

(CME) and non-clathrin-mediated endocytosis (NCE) (Sigismund et al., 2008). This study found in culture cells that while 70% of the receptors internalized through the CME pathway are recycled, only 15% of the receptors undergoing NCE are recycled (Sigismund et al., 2008). Importantly, at low level of ligand, almost all receptors undergo clathrin-mediated endocytosis. At high level of ligand, 60% of EGFR undergo CME and 40% undergo NCE. We incorporate this data in our model, considering that the fractions of degraded and recycled receptors depend on the level of ligand. At low ligand, $\alpha_{\text{deg}} = 0.3$ and at high ligand, $\alpha_{\text{deg}} = 0.55$. At intermediate levels of ligand, we interpolate linearly as follows:

$$\alpha_{\text{deg}}(t, \eta, \theta) = 0.3 + 0.25 \frac{L(t, \eta, \theta) - L_{\min}(t)}{L_{\max}(t) - L_{\min}(t)}; \quad \alpha_{\text{rec}}(t, \eta, \theta) = 1 - \alpha_{\text{deg}}(t, \eta, \theta)$$

where $L_{\min}(t)$ and $L_{\max}(t)$ are respectively the minimum and maximum values of L at time t .

4.8.4 Movement of the oocyte nucleus

The oocyte nucleus, which acts as a source of ligand, moves from the posterior to the dorsal anterior of the oocyte during stage 7 of oogenesis. We take this movement into account in the function $V(t, \eta, \theta)$ which represents the time and space-dependent flux of ligand.

4.8.5 Growth of the egg chamber

Egg chamber measurements during oogenesis are categorized into six different time points, considering that each of the stages 7 to 10A lasts approximately 6 hours (Spradling, 1993). In addition, since egg chambers at stages 8 and 9 change drastically within this six hour time frame, we take measurements of S8 and S9 at two separate time points, early (here onward denoted by (S8E or S9E) and late (S8L or S9L), that are defined by the size of the egg chamber. To summarize, we group the measurements of the egg chamber dimensions into six time-points: 3hr (corresponding to S7), 7.5hr (S8E), 10.5hr (S8L), 13.5hr (S9E), 16.5hr (S9L), 21hr (S10A).

Experimental measurements show that as oogenesis progresses from Stage 7 to Stage 10A, the

anterior-posterior dimensions of the egg-chamber increase by a factor of 4 and the dorso-ventral ones by a factor of 3 (Supplement 1). This hints that growth may play a fundamental role in shaping the distributions of ligand and signal. We take growth into account in the model by adding a transport term to the equations. This transport is done by the flow of a vector field v that transforms the prolate spheroid representing the oocyte according to the experimental measurements. As another effect of growth, the diffusion operator becomes time-dependent, as at each instant, it depends on the new geometry of the surface. To highlight this time dependence, from here onwards we denote the Laplace-Beltrami operator of the surface at time t by Δ_t .

4.8.6 Shift of the overlying follicle cells

The FCs overlaying the perivitteline space are known to gradually shift from cuboidal to squamous (anterior) and columnar (posterior) of the egg chamber (Supplement 1.) (Spradling, 1993). Since the receptors, complexes, signal and inhibitors are located inside or on the membrane of the FCs, this shift affects them. This phenomenon can be transcribed mathematically by adding another transport term to the equations of these variables. We introduce a time-dependent vector field w tangent to the surface of the prolate spheroid. The full set of equations including the phenomena of growth and shift of the FCs can be rewritten as:

$$\begin{cases} \frac{\partial L}{\partial t} = \nabla \cdot (vL) + D \Delta_t L + V - \frac{1}{H} \bar{k}_{on} RL + k_{off} C \\ \frac{\partial C}{\partial t} = \nabla \cdot (vC) + \nabla \cdot (wC) + \frac{1}{H} \bar{k}_{on} RL - (k_{off} + k_{ec}) C + \alpha_{rec} k_{rec} C_i \\ \frac{\partial C_i}{\partial t} = \nabla \cdot (vC_i) + \nabla \cdot (wC_i) + k_{ec} C - \alpha_{deg} k_{deg} C_i - \alpha_{rec} k_{rec} C_i \\ \frac{\partial R}{\partial t} = \nabla \cdot (vR) + \nabla \cdot (wR) - \bar{k}_{on} RL + k_{off} C - k_{er} R + Q_r \\ \frac{\partial S}{\partial t} = \nabla \cdot (vS) + \nabla \cdot (wS) + \bar{k}_S C_i - k_d S. \end{cases} \quad (2)$$

While the previous system (1) contained a reaction-diffusion PDE coupled with reaction ODEs, this new system (2) includes transport terms due to growth and cell shift, and we now have a system of coupled PDEs.

This set of equations is supplemented by initial conditions for each of the concentrations. All concentrations are initially considered to be zero, except for the initial concentration of receptors, R_0 :

$$\begin{aligned} L(0, \eta, \theta) = 0, \quad C(0, \eta, \theta) = 0, \quad C_i(0, \eta, \theta) = 0, \quad R(0, \eta, \theta) = R_0, \quad S(0, \eta, \theta) = 0, \\ \text{Kek}(0, \eta, \theta) = 0, \quad \text{Sty}(0, \eta, \theta) = 0. \end{aligned}$$

4.8.7 Numerical simulations

The complete model that we are studying is composed of several components that pose numerical challenges in specific ways. Solving the system of partial differential equations numerically requires finding a suitable spatial discretization of the domain, in this case a two-dimensional prolate spheroid. The most natural parameterization of such a symmetric surface is done with the prolate spheroidal coordinates as done previously (Goentoro et al., 2006). However, this system of coordinates is degenerate at the two poles of the spheroid. In consequence, the corresponding mesh constructed with prolate spheroidal coordinates is ill-suited for the numerical approximation of diffusion. As an alternative to prolate spheroidal coordinates, we used cubed spheroidal coordinates, adapted from the cubed sphere coordinates developed in (Guo et al.) and (Nair et al., 2005). More specifically, the quarter prolate spheroid is divided into 4 zones, each parameterized by a separate set of coordinates (x_i, y_i) , with $i \in [0, \dots, 3]$. Matching boundary conditions are set at the interfaces between zones. Neumann boundary conditions, provided by symmetry considerations and zero-flux conditions are set at the boundary of the quarter prolate spheroid. The other quarter prolate spheroid is obtained by symmetry.

The diffusion operator is approximated with finite differences in each zone. The ordinary differential equations are solved using the forward Euler method. All computations and simulations were done in MATLAB.

4.8.8 Parameterization of the perivitteline space

We represent the perivitteline space by a time-evolving prolate spheroid, parameterized at each time t by the coordinates (η, θ) .

At time t , the perivitteline space is the set of points $\mathcal{S}(t) = \{P(t, \eta, \theta), \eta \in [0, \pi], \theta \in [0, 2\pi]\}$ where $P(t, \eta, \theta)$ is given by :

$$P(t, \eta, \theta) = \begin{cases} x(t, \eta, \theta) = L_{DV}(t) \sin \eta \cos \theta \\ y(t, \eta, \theta) = L_{DV}(t) \sin \eta \sin \theta, \\ z(t, \eta, \theta) = L_{AP}(t) \cos \eta \end{cases}$$

The parameters L_{AP} and L_{DV} denote the half lengths of the big and small axes, respectively, and correspond to the anterior-posterior and dorso-ventral half lengths of the egg chamber, which change dynamically with time. Their values are taken by interpolating the measurements done at each stage of oogenesis (see Supplement 1). With these coordinates, the dorsal side of the egg chamber corresponds to $y = \theta = 0$ (see Supplement 2).

4.8.9 Numerical discretization of the domain

Notice that the spheroidal parameterization is not a diffeomorphism from $[0, \pi] \times [0, 2\pi]$ to $\mathcal{S}(t)$. Indeed, for all $\theta \in [0, 2\pi]$, $(x, y, z)(t, 0, \theta) = (0, 0, L_{AP}(t))$ and $(x, y, z)(t, \pi, \theta) = (0, 0, -L_{AP}(t))$.

Because of this, the spheroidal mesh built from these coordinates contains two singularities, or overlapping points, at $(0, 0, L_{AP})$ and $(0, 0, -L_{AP})$. As a consequence, the discretization points close to the poles $z = L_{AP}$ and $z = -L_{AP}$ are much closer than those towards the equator $z = 0$. This characteristic implies that the spheroidal mesh has irregular cell sizes, which makes it ill-suited for the finite-differences approximation of the diffusion operator. For this reason, we present another

system of coordinates that provides a more regular discretization of the domain.

A way to construct a more regular discretization of the spheroid is to divide it into several subdomains, each endowed with their own coordinate system. We developed the cubed spheroid coordinate system by extending the “cubed sphere” approach described in [Nair et al] in the context of the discontinuous Galerkin numerical scheme (Supplement 2).

In order to extend the cubed sphere parameterization, we define a homeomorphism between each point of the prolate spheroid P of small axis length L_{DV} and big axis length L_{AP} and the sphere S_{AP} of radius L_{AP} given by the equation $x^2 + y^2 + z^2 = L_{AP}^2$.

Let Φ_S be defined by $\Phi_S: S_{AP} \rightarrow P$, with

$$\Phi_S: (x_s, y_s, z_s) \mapsto (x, y, z) = \left(\frac{L_{DV}}{L_{AP}} x_s, \frac{L_{DV}}{L_{AP}} y_s, z_s \right).$$

The homeomorphism Φ_S^{-1} transforms each point P of the prolate spheroid P to a point P_s of the sphere S_{AP} by projection along the direction $\overrightarrow{P_z P}$, where P_z has coordinates $(0, 0, z)$.

The division of the sphere into “cubed” subdomains was introduced in [Nair et al]. We recall it here.

Let C_a be the cube of side $2a$ inscribed in S_{AP} , oriented such that the 3D Cartesian axes are orthogonal to its faces (see Supplement Figure S2-C). By definition, $a = \frac{1}{\sqrt{3}} L_{AP}$. We define the mapping $\Phi_C: P_s \in S_{AP} \mapsto P_s \in C_a$ by projection along the direction $\overrightarrow{OP_s}$. Then the point P_s can be parametrized by some coordinates (x_C, y_C) and the parameterization depends on the face of C_a .

Let F_0 be the face of C_a belonging to the plane $z = a$.

Let $(x_C, y_C) \in [-a, a] \times [-a, a]$ be the local coordinates on F_0 (see Supplement 2C).

Geometrically, we have the following relation between (x_C, y_C) and (x_s, y_s, z_s) :

$$\begin{cases} x_C = \frac{a x_s}{z_s} \\ y_C = \frac{a y_s}{z_s} \end{cases}.$$

Then the mapping $\Phi_C^0: F_0 \rightarrow S_{AP}$ mapping the cube's side F_0 to the prolate spheroid is defines by

$$\Phi_C^0: (x_C, y_C) \mapsto (x_s, y_s, z_s) = \left(\frac{x_C L_{AP}}{\sqrt{a^2 + x_C^2 + y_C^2}}, \frac{y_C L_{AP}}{\sqrt{a^2 + x_C^2 + y_C^2}}, \frac{a L_{AP}}{\sqrt{a^2 + x_C^2 + y_C^2}} \right).$$

Similar parametrizations can be given by defining local coordinate systems on the other faces of the cube. Composing Φ_C^0 and Φ_S gives a parameterization of the prolate spheroid that we name *cubed spheroid parameterization*. The respective images of the faces F_i of C_a by $\Phi_S \circ \Phi_C^i$ (for $i \in \{0, \dots, 5\}$) divide the spheroid into 6 domains F_i . For instance, the local coordinates on F_i define a subdomain of \mathcal{S} that we denote by D_0 and that can be parameterized combining the two maps as follows:

$$(x, y, z) = \Phi_S \circ \Phi_C^0(x_C, y_C) = \left(\frac{x_C L_{DV}}{\sqrt{a^2 + x_C^2 + y_C^2}}, \frac{y_C L_{DV}}{\sqrt{a^2 + x_C^2 + y_C^2}}, \frac{a L_{AP}}{\sqrt{a^2 + x_C^2 + y_C^2}} \right).$$

In practice, we restrict ourselves to a quarter prolate spheroid. In this case, only four faces of the cube are needed to parameterize it, which divides the spheroid into four subdomains, denoted by D_0, D_1, D_2 and D_3 (Supplemental 2). The coordinates on the total prolate spheroid can be obtained by symmetry. We create a cubed spheroidal mesh by discretizing these coordinates.

4.8.10 Numerical approximation of the Laplace-Beltrami operator

Given local coordinates $(x_i)_{i \in \{1, \dots, d\}}$ and a metric tensor $(g_{ij})_{(i,j) \in \{1, \dots, d\}^2}$, the Laplace-Beltrami operator is given by the explicit expression:

$$\Delta_{LB} f = \frac{1}{\sqrt{|g|}} \sum_{i=1}^d \partial_{x_i} (\sqrt{|g|} g^{ij} \partial_{x_j} f).$$

Using the cubed spheroidal parameterization, one can compute the metric tensor of \mathcal{S} in each subdomain.

Let $r = (x, y, z)$. Then in D_0 , the metric tensor G_0 is given by:

$$G_0 = \begin{pmatrix} g_0^{11} & g_0^{12} \\ g_0^{12} & g_0^{22} \end{pmatrix}$$

where

$$\begin{cases} g_0^{11} = |\partial_{x_c} r|^2 = \frac{1}{(a^2 + x_c^2 + y_c^2)^3} (L_{DV}^2 (a^2 + y_c^2)^2 + L_{DV}^2 x_c^2 y_c^2 + L_{AP}^2 a^2 x_c^2) \\ g_0^{11} = |\partial_{y_c} r|^2 = \frac{1}{(a^2 + x_c^2 + y_c^2)^3} (L_{DV}^2 (a^2 + x_c^2)^2 + L_{DV}^2 x_c^2 y_c^2 + L_{AP}^2 a^2 y_c^2) \\ g_0^{12} = \langle \partial_{x_c} r, \partial_{y_c} r \rangle = \frac{1}{(a^2 + x_c^2 + y_c^2)^3} (-L_{DV}^2 (2a^2 + x_c^2 + y_c^2)^2 + L_{AP}^2 a^2 x_c y_c). \end{cases}$$

The expression of the metric tensor allows to compute the Laplace-Beltrami operator in the cubed spheroid coordinates in each domain.

Using cubed spheroidal coordinates requires subdividing the prolate spheroid into several subdomains, and treating the interfaces between domains with appropriate boundary conditions. Notice that due to the symmetry of the source function with respect to the (X0Z) plane, the full system is symmetric with respect to the (X0Z) plane. It is then sufficient to solve numerically the system of equations on a half prolate spheroid, and to recover the solution on the full domain by symmetry with respect to (X0Z). The boundaries between domains are treated with Dirichlet boundary conditions. The boundaries of the quarter spheroid inscribed in the (X0Z) plane are

treated with Neumann boundary conditions, for reasons of symmetry (Supplement 2C).

4.8.11 Extrapolation of parameter values from intensity measurements:

Since they cannot be evaluated by direct measurements, the inhibitors' strengths γ_{Kek} and γ_{Sty} were determined by comparing simulation results for different values of these two parameters with intensity measurements of dpERK immunostainings.

Intensity measurements for dpERK were done at early and late Stages 9 and mid Stage 10A along the anterior-posterior axis and the dorso-ventral axis for wild-type, 1px, and 4px flies.

Simulations were run with our model for wild-type, 1px, and 4px, for four different values of γ_{Kek} and four different values of γ_{Sty} . This provided us with 16 AP and 16 DV intensity plots for each of the three stages and each of the three perturbations. These simulation results are presented in the Supplement Figure 4A. Supplement Figure 4A shows the AP plots for the 20 different combinations of γ_{Kek} and γ_{Sty} at S10A. In each of the 20 plots, we represented the three experimental intensity curves for wild-type(2x GRK), 1x GRK, and 4x GRK, and in each of the 20 plots, we normalized the three simulation curves by the same factor K_i (for $i \in \{1, \dots, 16\}$) so that the experimental and simulated wild-type AP curves intersect at 20% of the total length at S10A.

The problem remained of using this wealth of information to select the combination of parameters γ_{Kek} and γ_{Sty} that provide the closest results to the experimental measurements.

This was done by computing a cost function for each of the 20 combinations of parameters. Let us denote by $u_{AP}^i, v_{AP}^i, w_{AP}^i, u_{DV}^i, v_{DV}^i$, and w_{DV}^i respectively the 6 AP and DV curves corresponding to wild-type, 1px and 4px, for the i -th combination of parameters, after normalization by the constant K_i . The corresponding experimental curves are respectively denoted by $U_{AP}^i, V_{AP}^i, W_{AP}^i, U_{DV}^i, V_{DV}^i$, and W_{DV}^i . Then the cost Γ^i of the i -th combination of parameters is:

$$\begin{aligned} \Gamma^i = & l_1(u_{AP}^i - U_{AP}^i) + l_1(v_{AP}^i - V_{AP}^i) + l_1(w_{AP}^i - W_{AP}^i) + l_1(u_{DV}^i - U_{DV}^i) \\ & + l_1(v_{DV}^i - V_{DV}^i) + l_1(w_{DV}^i - W_{DV}^i), \end{aligned}$$

where l_1 denotes the L^1 norm in 1D:

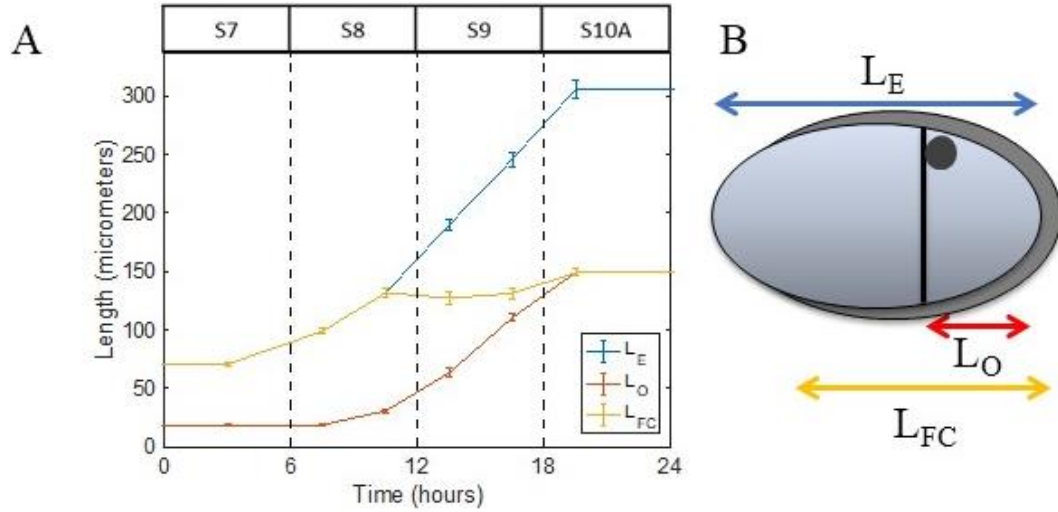
$$l_1(u_{AP}^i - U_{AP}^i) := \int_0^{100} |u_{AP}^i(\xi) - U_{AP}^i(\xi)| d\xi.$$

Notice that both experimental and simulation curves are horizontally normalized so that $\xi \in [0,100]$ represents the percentage of the length of the measured experimental data.

Supplement 5B presents the 16 costs Γ^i in a heat map. The minimum is attained for the combination of parameters $\gamma_{\text{Kek}} = \gamma_{\text{Sty}} = 500$ AU.

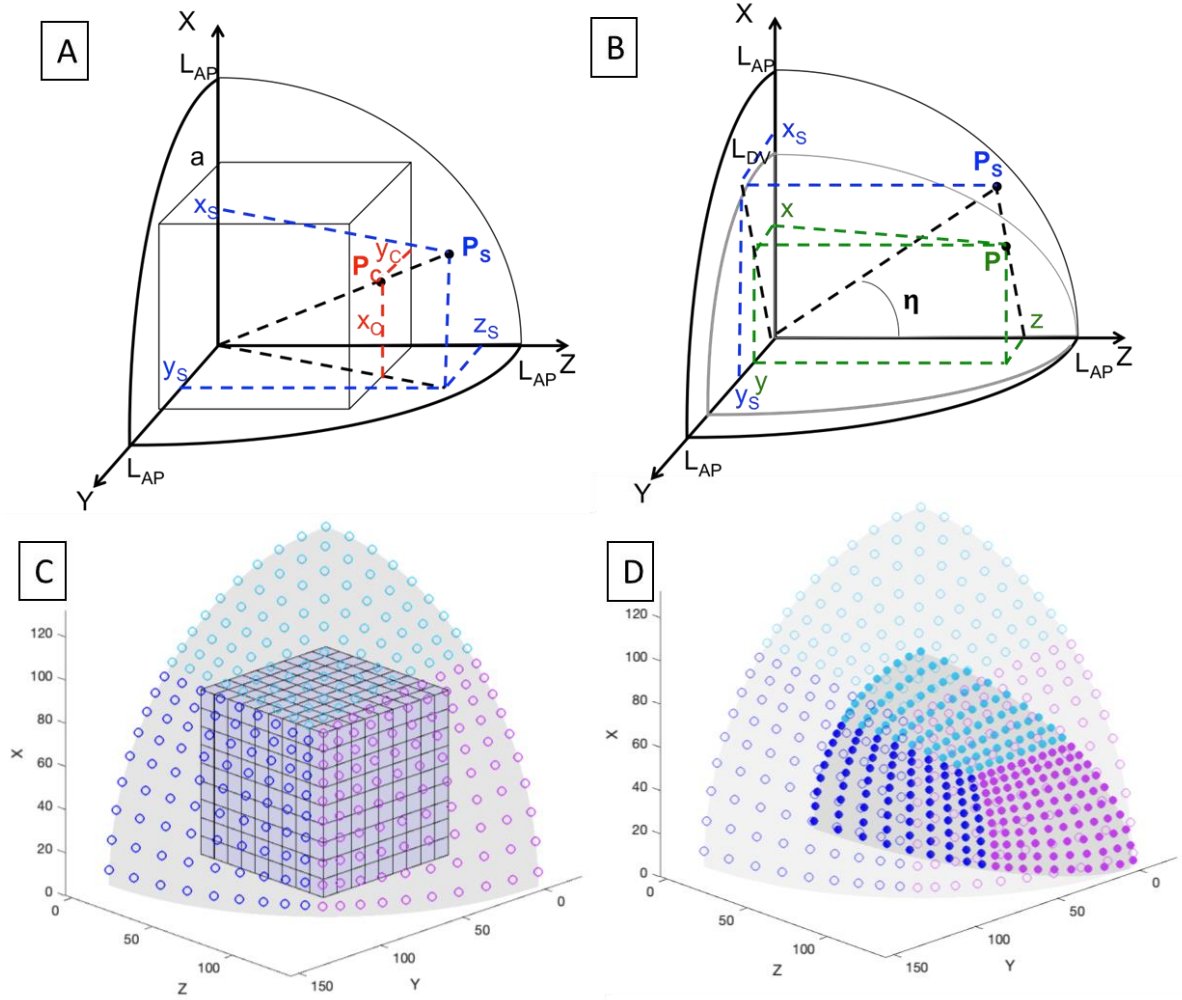
Chapter 5 Supplemental data

Supplement 1: Dynamic measurements of the egg chamber from S7 to S10A



Supplement 1: Dynamic measurements of the egg chamber dimensions from S7 to S10A. Data was collected at 6 consecutive time-points: 3hr (S7), 7.5hr (S8E), 10.5hr (S8L), 13.5hr (S9E), 16.5hr (S9L), 19.5hr (S10AE). A. Three measurements were taken in the AP direction: the total length of the egg chamber L_E , the length of the oocyte L_O , and the length of the FCs L_{FC} . B. A cartoon schematic showing positions of measurements along the AP.

Supplement 2: Modelling the PVS: Construction of the cubed spheroidal mesh.



Supplement 2: Construction of the cubed spheroidal mesh.

The cubed spheroidal mesh is obtained by a two-step transformation, from the cube C_a to the sphere S_{AP} , and from the sphere S_{AP} to the spheroid P .

Consider the sphere of radius S_{AP} centered at 0 and let C_a be its inscribed cube of side $2a = 2 \frac{L_{AP}}{\sqrt{3}}$, also centered at 0. Each side of C_a is discretized by a regular orthonormal mesh (Figure S2-D). Then the cubed spherical mesh of S_{AP} is obtained by taking the radial projection of the mesh of C_a onto S_{AP} (Supplement 2C,A): each vertex $P_C(x_C, y_C, z_C) \in C_a$ of the mesh is projected radially onto S_{AP} , giving the point $P_S(x_S, y_S, z_S) \in S_{AP}$. By this transformation, the sphere S_{AP} is meshed by the cubed sphere projection.

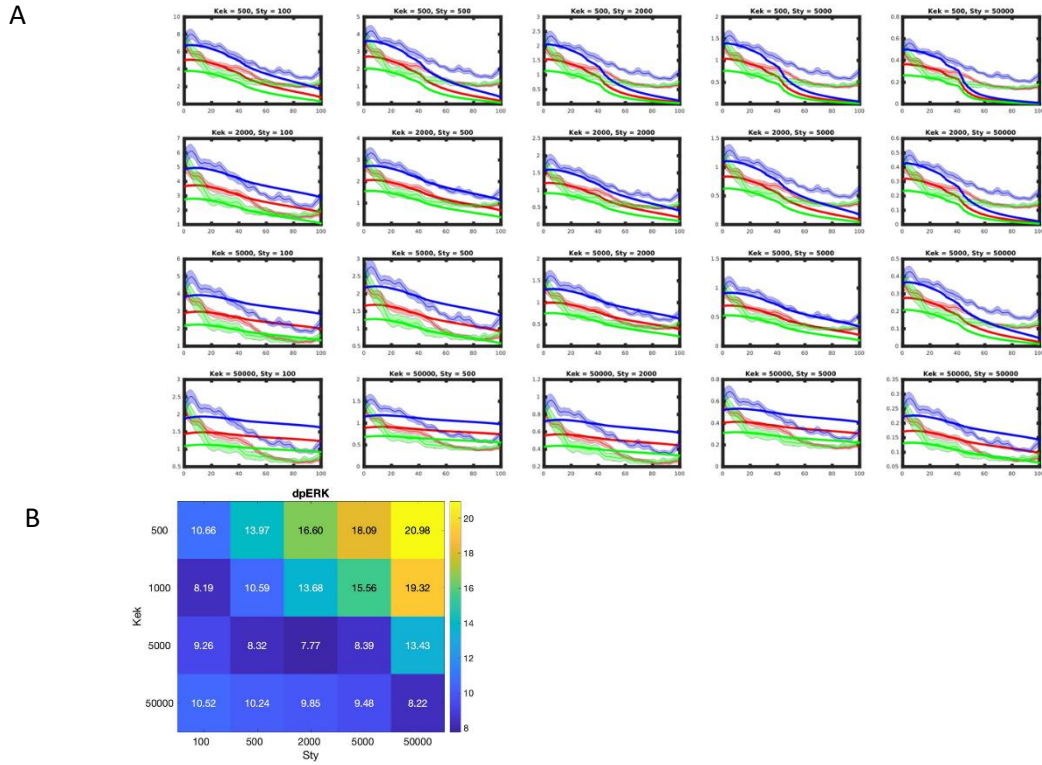
Secondly, each vertex $P_S \in S_{AP}$ of the cubed spherical mesh is projected onto the prolate spheroid P , orthogonally to the z -axis (Supplement 2B). This transformation defines the cubed spheroidal mesh (Supplement 2D).

Supplement 3: Table of parameters

Symbol	Parameter	Value or Range	Source
L_{FC}	Length columnar FCs	Variable (See Fig. S1)	This study
L_{AP}	Length egg chamber	Variable (See Fig. S1)	This study
L_{Oo}	Length oocyte	Variable (See Fig. S1)	This study
L_{DV}	Width Oocyte	Variable (See Fig. S1)	This study
H	Perivitelline space thickness	0.5 μm	This study
D	Rate of Diffusion	3,600-360,000 $\mu\text{m}^2 \text{hr}^{-1}$ 36-360,000 $\mu\text{m}^2 \text{hr}^{-1}$	Pribyl et al., 2003
k_{ec}	Rate of internalization of complex	6 hr^{-1}	Pribyl et al., 2003
k_{on}	Rate of Ligand-Receptor binding	$6 \times 10^{22} - 6 \times 10^{25} \text{mol}^{-1} \mu\text{m}^3 \text{hr}^{-1}$	Pribyl et al., 2003
k_{off}	Rate of Ligand-Receptor dissociation	6 hr^{-1}	Pribyl et al., 2003
R_0	Free receptor per surface area in the absence	$6.7 \times 10^{22} \text{mol} \mu\text{m}^{-2}$	Pribyl et al., 2003
k_{er}	Receptor internalization independent of ligand	0.6-6 hr^{-1}	Pribyl et al., 2003
α_{rec}	Fraction of recycled receptor	0.45-0.7	Sigismund et al., 2008
α_{deg}	Fraction of degraded receptor	0.3-0.55	Sigismund et al., 2008
k_{rec}	Receptor recycling rate	2.3 hr^{-1}	Sigismund et al., 2008
k_{deg}	Receptor degradation rate	2.5 hr^{-1}	Sigismund et al., 2008
k_d	dpERK degradation rate	2.3 hr^{-1}	Pinilla-Macua et al., 2016
V_0	Initial flux of ligand from the oocyte	$2 \times 10^{-21} \text{mol} \mu\text{m}^{-2} \text{hr}^{-1}$	Pribyl et al., 2003
Q_r	Receptor production rate	$4 \times 10^{-22} \text{mol} \mu\text{m}^{-2} \text{hr}^{-1}$	Pribyl et al., 2003
k_s	dpERK production rate	1 hr^{-1}	Arbitrary

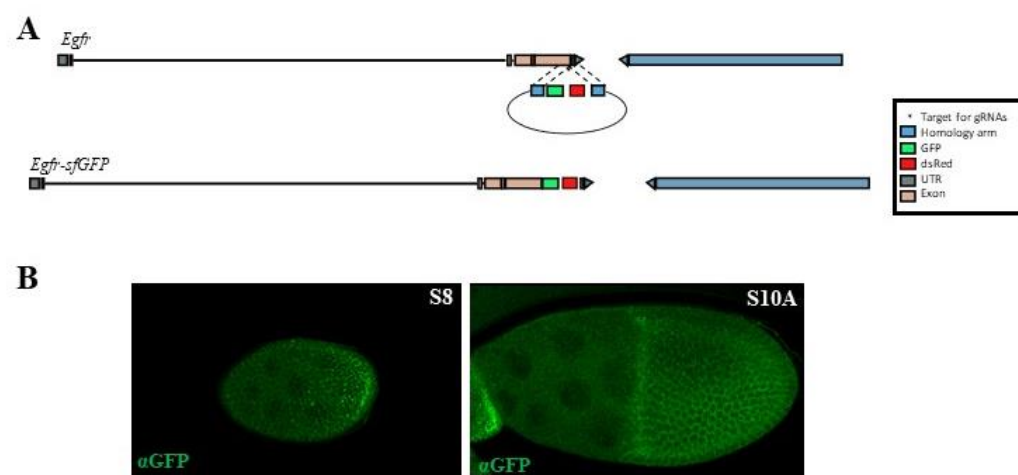
Supplement 3: Parameters of the model

Supplement 4: Fitting parameters selection to intensity data



Supplement 4: Fitting parameters selection to intensity data **A.** AP profiles of experimental intensity measurements for wild type (red), 4x grk (blue), and 1x grk (green) and of the respective simulations at Stage 10A (arbitrary units). This shows that values of Kek and Sty around 2000-5000 give a good fit. **B.** Heat map comparing the L2 norms of the dpERK AP profiles of experimental measurements and of the simulations at Stage 10A. The lower the value, the better the fit. This shows that values of Kek and Sty around 2000-5000 give a good fit.

Supplement 5: An EGFR tagged with sfGFP



Supplement 5: Generating the EGFR-sfGFP A. A cartoon schematic of the EGFR locus, the donor plasmid, and sfGFP tagged EGFR B. Immunohistochemistry images of EGFR using GFP antibodies at stages 8 and 10A.

Chapter 6: Bibliography

1. Ables, E.T., Jan. 1, 2015. *Drosophila* Oocytes as a Model for Understanding Meiosis: An Educational Primer to Accompany "Corolla Is a Novel Protein That Contributes to the Architecture of the Synaptonemal Complex of *Drosophila*". *Genetics* 199, 17-23.
2. Arnold, D.C., 2nd, Schroeder, G., Smith, J.C., Wahjudi, I.N., Heldt, J.P., Richards, G.D., Agarwal, G., Brisbane, W.G., Farley, D.V., Baldwin, D.D., 2013. Comparing radiation exposure between ablative therapies for small renal masses. *Journal of endourology / Endourological Society* 27, 1435-1439.
3. Arteaga, C.L., 2002. Overview of epidermal growth factor receptor biology and its role as a therapeutic target in human neoplasia. *Semin. Oncol.* 23, 3-9.
4. Bakker, J., Spits, M., Neefjes, J., Berlin, I., 2017. The EGFR odyssey – from activation to destruction in space and time. *Journal of Cell Science* 130, 4087-4096.
5. Barolo, S., Carver, L.A., Posakony, J.W., 2000. GFP and beta-galactosidase transformation vectors for promoter/enhancer analysis in *Drosophila*. *BioTechniques* 29, 726, 728, 730, 732.
6. Berg, C.A., 2005. The *Drosophila* shell game: patterning genes and morphological change. *Trends Genet* 21, 346-355.
7. Blackman, R.K., Sanicola, M., Raftery, L.A., Gillevet, T., Gelbart, W.M., 1991. An extensive 3' cis-regulatory region directs the imaginal disk expression of decapentaplegic, a member of the TGF-beta family in *Drosophila*. *Development* 111, 657-666.
8. Boisclair-Lachance, J.F., Fregoso-Lomas, M., Eleiche, A., Bouchard Kerr, P., Nilson, L.A., 2009. Graded Egfr activity patterns the *Drosophila* eggshell independently of autocrine feedback. *Development* 136, 2893-2902.
9. Boisclair Lachance, J.F., Pelaez, N., Cassidy, J.J., Webber, J.L., Rebay, I., Carthew, R.W., 2014. A comparative study of Pointed and Yan expression reveals new complexity to the transcriptional networks downstream of receptor tyrosine kinase signaling. *Dev Biol* 385, 263-278.
10. Brand, A.H., Manoukian, A.S., Perrimon, N., 1994. Ectopic expression in *Drosophila*. *Methods Cell Biol* 44, 635-654.
11. Burke, P., Schooler, K., Wiley, H.S., 2001. Regulation of epidermal growth factor receptor signaling by endocytosis and intraellular trafficking. *Moll Cell Biol* 12.

12. Carrell, S., O'Connell, M., Jacobsen, T., Allen, A.A., Smith, S.M., Reeves, G.T., 2017. A facilitated diffusion mechanism establishes the *Drosophila* Dorsal gradient. *Development*.
13. Cavaliere, V., Bernardi, F., Romani, P., Duchi, S., Gargiulo, G., 2008. Building Up the *Drosophila* Eggshell: First of All the Eggshell Genes Must Be Transcribed. *Dev. Dyn.* 237.
14. Chang, W.L., Liou, W., Pen, H.C., Chou, H.Y., Chang, Y.W., Li, W.H., Chiang, W., Pai, L.M., 2008. The gradient of Gurken, a long-range morphogen, is directly regulated by Cbl-mediated endocytosis. *Development* 135, 1923-1933.
15. Charbonnier, E., Fuchs, A., Cheung, L.S., Chayengia, M., Veikkolainen, V., Seyfferth, J., Shvartsman, S.Y., Pyrowolakis, G., 2015. BMP-dependent gene repression cascade in *Drosophila* eggshell patterning. *Dev Biol* 400, 258-265.
16. Chen, Y., Schupbach, T., 2006. The role of brinker in eggshell patterning. *Mech Dev* 123, 395-406.
17. Cheung, L.S., Simakov, D.S., Fuchs, A., Pyrowolakis, G., Shvartsman, S.Y., 2013. Dynamic model for the coordination of two enhancers of broad by EGFR signaling. *Proc Natl Acad Sci U S A* 110, 17939-17944.
18. Chu, L., Wiley, H.S., Lauffenburger, D.A., 1996. Endocytic Relay as a Potential Means for Enhancing Ligand Transport through Cellular Tissue Matrices: Analysis and Possible Implications for Drug Delivery. *Tissue engineering* 2, 17-38.
19. Corkery, B., Crown, J., Clynes, M., O'Donovan, N., 2009. Epidermal growth factor receptor as a potential therapeutic target in triple-negative breast cancer. *Ann Oncol.* 862-867.
20. Courgeon, M., He, D.Q., Liu, H.H., Legent, K., Treisman, J.E., 2018. The *Drosophila* Epidermal Growth Factor Receptor does not act in the nucleus. *Journal of Cell Science* 131, 1-8.
21. Davidson, E.H., Erwin, D.H., 2006. Gene regulatory networks and the evolution of animal body plans. *Science* 311, 796-800.
22. Deng, W.M., Bownes, M., 1997. Two signalling pathways specify localised expression of the Broad-Complex in *Drosophila* eggshell patterning and morphogenesis. *Development* 124, 4639-4647.
23. Dequier, E., Soud, S., Pal, M., Maroy, P., Lepesant, J.A., Yanicostas, C., 2001. Top-DER- and Dpp-dependent requirements for the *Drosophila* fos/kayak gene in follicular epithelium morphogenesis. *Mech Dev* 106, 47-60.
24. Dobens, L.L., Raftery, L.A., 1998. *Drosophila* oogenesis: a model system to understand TGF-beta/Dpp directed cell morphogenesis. *Ann N Y Acad Sci* 857, 245-247.

25. Dobens, L.L., Raftery, L.A., 2000. Integration of epithelial patterning and morphogenesis in *Drosophila* ovarian follicle cells. *Dev Dyn* 218, 80-93.
26. Duffy, J.B., 2002. GAL4 system in *Drosophila*: a fly geneticist's Swiss army knife. *Genesis* 34, 1-15.
27. Duhart, J.C., Parsons, T.T., Raftery, L.A., 2017. The repertoire of epithelial morphogenesis on display: Progressive elaboration of *Drosophila* egg structure. *Mech Dev* 148, 18-39.
28. Dworkin, I., Palsson, A., Birdsall, K., Gibson, G., 2003. Evidence that *Egfr* contributes to cryptic genetic variation for photoreceptor determination in natural populations of *Drosophila melanogaster*. *Curr Biol* 13, 1888-1893.
29. Flockhart, I.T., Booker, M., Yanhui, H., McElvany, B., Gilly, Q., Mathey-Prevot, B., Perrimon, N., Mohr, S.M., 2012. "FlyRNAi.org--the database of the *Drosophila* RNAi screening center: 2012 update." *Nucleic Acids Res* 40, Pp. D715-719.
30. Fortian, A., Sorkin, A., 2013. Live-cell fluorescence imaging reveals high stoichiometry of Grb2 binding to the EGF receptor sustained during endocytosis. *J Cell Sci* 127, 432-444.
31. Freeman, M., 1994. The *spitz* gene is required for photoreceptor determination in the *Drosophila* eye where it interacts with the EGF receptor. *Mech Dev* 48, 25-33.
32. Fregoso Lomas, M., Hails, F., Lachance, J.F., Nilson, L.A., 2013. Response to the dorsal anterior gradient of EGFR signaling in *Drosophila* oogenesis is prepatterned by earlier posterior EGFR activation. *Cell Rep* 4, 791-802.
33. Fuchs, A., Cheung, L.S., Charbonnier, E., Shvartsman, S.Y., Pyrowolakis, G., 2012. Transcriptional interpretation of the EGF receptor signaling gradient. *Proc Natl Acad Sci U S A* 109, 1572-1577.
34. Ghiglione, C., Amundadottir, L., Andresdottir, M., Bilder, D., Diamonti, J.A., Noselli, S., Perrimon, N., Carraway, I.K., 2003. Mechanism of inhibition of the *Drosophila* and mammalian EGF receptors by the transmembrane protein Kekk1. *Development* 130, 4483-4493.
35. Ghiglione, C., Carraway, K.L., 3rd, Amundadottir, L.T., Boswell, R.E., Perrimon, N., Duffy, J.B., 1999. The transmembrane molecule kekk1 acts in a feedback loop to negatively regulate the activity of the *Drosophila* EGF receptor during oogenesis. *Cell* 96, 847-856.
36. Goentoro, L.A., Reeves, G.T., Kowal, C.P., Martinelli, L., Schupbach, T., Shvartsman, S.Y., 2006. Quantifying the Gurken morphogen gradient in *Drosophila* oogenesis. *Dev Cell* 11, 263-272.
37. Gonzalez-Reyes, A., Elliot, H., St Johnston, D., 1995. Polarization of both major body axes in *Drosophila* by gurken-torpedo signalling. *Nature* 375, 654-658.

38. Gratz, S.J., Cummings, A.M., Nguyen, J.N., Hamm, D.C., Donohue, L.K., Harrison, M.M., Wildonger, J., O'Connor-Giles, K.M., 2013. Genome engineering of *Drosophila* with the CRISPR RNA-guided Cas9 nuclease. *Genetics* 194, 1029-1035.
39. Gratz, S.J., Ukken, F., Rubinstein, C., Thiede, G., Donohue, L., Cummings, A.M., O'Connor-Giles, K.M., 2014. Highly Specific and efficient CRISPR/Cas9-catalyzed homology-directed repair in *Drosophila*. *Genetics* 196, 961-971.
40. Groth, A.C., Fish, M., Nusse, R., Calos, M.P., 2004. Construction of transgenic *Drosophila* by using the site-specific integrase from phage ϕ C31. *Genetics* 166, 1775-1782.
41. Guichard, A., Biehs, B., Sturtevant, M.A., Wickline, L., Chacko, J., Howard, K., Bier, E., 1999. rhomboid and Star interact synergistically to promote EGFR/MAPK signaling during *Drosophila* wing vein development *Development* 126, 2663-2676.
42. Guo, W., Nair, R.D., Qiu, J.-M., A conservative semi-Lagrangian discontinuous Galerkin scheme on the cubed sphere. *Monthly Weather Review* 142, 457-475.
43. Haag, A., Gutierrez, P., Buhler, A., Walser, M., Yang, Q., Langouet, M., Kradolfer, D., Frohli, E., Herrmann, C.J., Hajnal, A., Escobar-Restrepo, J.M., 2014. An In Vivo EGF Receptor Localization Screen in *C. elegans* Identifies the Ezrin Homolog ERM-1 as a Temporal Regulator of Signaling. *PLoS Genet* 10.
44. Horne-Badovinac, S., Bilder, D., 2005. Mass transit: epithelial morphogenesis in the *Drosophila* egg chamber. *Dev Dyn* 232, 559-574.
45. Inoue, H., Imamura, T., Ishidou, Y., Takase, M., Udagawa, Y., Oka, Y., Tsuneizumi, K., Tabata, T., Miyazono, K., Kawabata, M., 1998. Interplay of signal mediators of decapentaplegic (Dpp): molecular characterization of mothers against dpp, Medea, and daughters against dpp. *Molecular biology of the cell* 9, 2145-2156.
46. Ivan, A., Halfon, M.S., Sinha, S., 2008. Computational discovery of cis-regulatory modules in *Drosophila* without prior knowledge of motifs. *Genome Biol* 9, R22.
47. Jenett, A., Rubin, G.M., Ngo, T.T., Shepherd, D., Murphy, C., Dionne, H., Pfeiffer, B.D., Cavallaro, A., Hall, D., Jeter, J., Iyer, N., Fetter, D., Hausenfluck, J.H., Peng, H., Trautman, E.T., Svirskas, R.R., Myers, E.W., Iwinski, Z.R., Aso, Y., DePasquale, G.M., Enos, A., Hulamm, P., Lam, S.C., Li, H.H., Lavery, T.R., Long, F., Qu, L., Murphy, S.D., Rokicki, K., Safford, T., Shaw, K., Simpson, J.H., Sowell, A., Tae, S., Yu, Y., Zugates, C.T., 2012. A GAL4-driver line resource for *Drosophila* neurobiology. *Cell Rep* 2, 991-1001.
48. Jordan, K.C., Hatfield, S.D., Tworoger, M., Ward, E.J., Fischer, K.A., Bowers, S., Ruohola-Baker, H., 2005. Genome wide analysis of transcript levels after perturbation of the EGFR pathway in the *Drosophila* ovary. *Dev Dyn* 232, 709-724.

49. Jory, A., Estella, C., Giorgianni, M.W., Slattey, M., Laverty, T.R., Rubin, G.M., Mann, R.S., 2012. A survey of 6,300 genomic fragments for cis-regulatory activity in the imaginal discs of *Drosophila melanogaster*. *Cell Rep* 2, 1014-1024.
50. King, C.R., 1970. Ovarian Development in *Drosophila melanogaster*. Academic Press, London.
51. Konikoff, C.E., Karr, T.L., McCutchan, M., Newfeld, S.J., Kumar, S., 2012. Comparison of embryonic expression within multigene families using the FlyExpress discovery platform reveals more spatial than temporal divergence. *Dev Dyn* 241, 150-160.
52. Krauchunas, A.R., Wolfner, M.F., 2013. Molecular changes during egg activation. *Curr Top Dev Biol* 102, 267-292.
53. Kvon, E.Z., Kazmar, T., Stampfel, G., Yanez-Cuna, J.O., Pagani, M., Schernhuber, K., Dickson, B.J., Stark, A., 2014. Genome-scale functional characterization of *Drosophila* developmental enhancers in vivo. *Nature* 512, 91-95.
54. Lecuyer, E., Yoshida, H., Parthasarathy, N., Alm, C., Babak, T., Cerovina, T., Hughes, T.R., Tomancak, P., Krause, H.M., 2007. Global analysis of mRNA localization reveals a prominent role in organizing cellular architecture and function. *Cell* 131, 174-187.
55. Lembong, J., Yakoby, N., Shvartsman, S.Y., 2008. Spatial regulation of BMP signaling by patterned receptor expression. *Tissue Eng Part A* 14, 1469-1477.
56. Lev, Z., Shilo, B.Z., Kimchie, Z., 1985. Developmental Changes in Expression of the *Drosophila melaogaster* Epidermal Growth Factor Receptor Gene *Dev Biol* 110, 499-502.
57. Levine, M., 2010. Transcriptional enhancers in animal development and evolution. *Curr Biol* 20, R754-763.
58. Levine, M., Tjian, R., 2003. Transcription regulation and animal diversity. *Nature* 424, 147-151.
59. Li, H.H., Kroll, J.R., Lennox, S.M., Ogundeyi, O., Jeter, J., Depasquale, G., Truman, J.W., 2014. A GAL4 driver resource for developmental and behavioral studies on the larval CNS of *Drosophila*. *Cell reports* 8, 897-908.
60. Manning, L., Heckscher, E.S., Purice, M.D., Roberts, J., Bennett, A.L., Kroll, J.R., Pollard, J.L., Strader, M.E., Lupton, J.R., Dyukareva, A.V., Doan, P.N., Bauer, D.M., Wilbur, A.N., Tanner, S., Kelly, J.J., Lai, S.L., Tran, K.D., Kohwi, M., Laverty, T.R., Pearson, J.C., Crews, S.T., Rubin, G.M., Doe, C.Q., 2012. A resource for manipulating gene expression and analyzing cis-regulatory modules in the *Drosophila* CNS. *Cell Rep* 2, 1002-1013.
61. Margaritis, L.H., Dellas, K., Kalantzi, M.C., Kambysellis, M.P., 1983. The eggshell of Hawaiian *Drosophila*: structural and biochemical studies in *D. grimshawi* and comparison to *D. melanogaster*. *Roux's Archives of Developmental Biology*, 303-316.

62. Marmion, R.A., Jevtic, M., Springhorn, A., Pyrowolakis, G., Yakoby, N., 2013. The *Drosophila* BMPRII, wishful thinking, is required for eggshell patterning. *Dev Biol* 375, 45-53.
63. Marmion, R.A., Yakoby, N., 2018. In locus analysis of patterning evolution in the BMPRII Wishful thinking. *Development* 145.
64. Matikas, A., Mistriotis, D., Georgoulas, V., Kotsakis, A., 2015. Current and Future Approaches in the Management of Non-Small-Cell Lung Cancer Patients With Resistance to EGFR TKIs. *Clin Lung Cancer*.
65. Mohammed, J., Lai, E.C., 2016. Lai lab miRNA expression and conservation data for *Drosophila*.
66. Morimoto, A.M., Jordan, K.C., Tietze, K., Britton, J.S., O'Neill, E.M., Ruohola-Baker, H., 1996. Pointed, an ETS domain transcription factor, negatively regulates the EGF receptor pathway in *Drosophila* oogenesis. *Development* 122, 3745-3754.
67. Muzzopappa, M., Wappner, P., 2005. Multiple roles of the F-box protein Slimb in *Drosophila* egg chamber development. *Development* 132, 2561-2571.
68. Nair, R.D., Thomas, S.J., Loft, R.D., 2005. A discontinuous Galerkin transport scheme on the cubed sphere. *Monthly Weather Review* 133, 814-828.
69. Nam, J., Dong, P., Tarpine, R., Istrail, S., Davidson, E.H., 2010. Functional cis-regulatory genomics for systems biology. *Proceedings of the National Academy of Sciences of the United States of America* 107, 3930-3935.
70. Neuman-Silberberg, F.S., Schupbach, T., 1993. The *Drosophila* dorsoventral patterning gene *gurken* produces a dorsally localized RNA and encodes a TGF alpha-like protein. *Cell* 75, 165-174.
71. Niepielko, M.G., Marmion, R.A., Kim, K., Luor, D., Ray, C., Yakoby, N., 2014. Chorion patterning: a window into gene regulation and *Drosophila* species' relatedness. *Mol Biol Evol* 31, 154-164.
72. Niepielko, M.G., Yakoby, N., 2014. Evolutionary changes in TGFalpha distribution underlie morphological diversity in eggshells from *Drosophila* species. *Development* 141, 4710-4715.
73. Nilson, L.A., Schupbach, T., 1999. EGF receptor signaling in *Drosophila* oogenesis. *Curr Top Dev Biol* 44, 203-243.
74. Pallavi, S., Kannan, R., Shashidhara, L., 2006. Negative regulation of Egfr/Ras pathway by Ultrabithorax during haltere development in *Drosophila*. *Dev Biol* 296, 340-352.
75. Parker, L., Stathakis, D.G., Arora, K., 2003. Regulation of BMP and Activin Signaling in *Drosophila*. *Prog Mol Subcell Biol* 34, 73-101.

76. Peri, F., Bokel, C., Roth, S., 1999. Local Gurken signaling and dynamic MAPK activation during *Drosophila* oogenesis. *Mech Dev* 81, 75-88.
77. Peri, F., Roth, S., 2000. Combined activities of Gurken and decapentaplegic specify dorsal chorion structures of the *Drosophila* egg. *Development* 127, 841-850.
78. Pfeiffer, B.D., Jenett, A., Hammonds, A.S., Ngo, T.T., Misra, S., Murphy, C., Scully, A., Carlson, J.W., Wan, K.H., Lavery, T.R., Mungall, C., Svirskas, R., Kadonaga, J.T., Doe, C.Q., Eisen, M.B., Celniker, S.E., Rubin, G.M., 2008. Tools for neuroanatomy and neurogenetics in *Drosophila*. *Proceedings of the National Academy of Sciences of the United States of America* 105, 9715-9720.
79. Pfeiffer, B.D., Ngo, T.T., Hibbard, K.L., Murphy, C., Jenett, A., Truman, J.W., Rubin, G.M., 2010. Refinement of tools for targeted gene expression in *Drosophila*. *Genetics* 186, 735-755.
80. Pouradier-Duteil, N., Rossi, F., Boscain, U., Piccoli, B., 2015. Developmental partial differential equations. *IEEE 54th Conference on Decision and Control (CDC)*, 3181-3186.
81. Pribyl, M., Muratov, C.B., Shvartsman, S.Y., 2003a. Discrete models of autocrine cell communication in epithelial layers. *Biophys J* 84, 3624-3635.
82. Pribyl, M., Muratov, C.B., Shvartsman, S.Y., 2003b. Long-range signal transmission in autocrine relays. *Biophys J* 84, 883-896.
83. Pyrowolakis, G., Veikkolainen, V., Yakoby, N., Shvartsman, S.Y., 2017. Gene regulation during *Drosophila* eggshell patterning. *PNAS* 114, 5808-5813.
84. Queenan, A.M., Ghabrial, A., Schupbach, T., 1997. Ectopic activation of torpedo/Egfr, a *Drosophila* receptor tyrosine kinase, dorsalizes both the eggshell and the embryo. *Development* 124, 3871-3880.
85. Revaitis, N.T., Marmion, R.A., Farhat, M., Ekiz, V., Wang, W., Yakoby, N., 2017. Simple expression domains are regulated by discrete CRMs during *Drosophila* oogenesis. *G3* 7, 2705-2718.
86. Robinson, J.T., Thorvaldsdottir, H., Winckler, W., Guttman, M., Lander, E.S., Getz, G., Mesirov, J.P., 2011. Integrative genomics viewer. *Nature biotechnology* 29, 24-26.
87. Rossi, F., Pouradier-Duteil, N., Yakoby, N., Piccoli, B., 2016. Control of reaction-diffusion equations on time-evolving manifolds. *IEEE 55th Conference on Decision and Control (CDC)* 1614-1619.
88. Roth, S., Neuman-Silberberg, F.S., Barcelo, G., Schupbach, T., 1995. cornichon and the EGF receptor signaling process are necessary for both anterior-posterior and dorsal-ventral pattern formation in *Drosophila*. *Cell* 81, 967-978.

89. Ruohola-Baker, H., Grell, E., Chou, T.B., Baker, D., Jan, L.Y., Jan, Y.N., 1993. Spatially localized rhomboid is required for establishment of the dorsal-ventral axis in *Drosophila* oogenesis. *Cell* 73, 953-965.
90. Sapir, A., Schweitzer, R., Shilo, B.Z., 1998. Sequential activation of the EGF receptor pathway during *Drosophila* oogenesis establishes the dorsoventral axis. *Development* 125, 191-200.
91. Schejter, E., Segal, D., Glazer, L., Shilo, B.Z., 1986. Alternative 5' exons and tissue-specific expression of the *Drosophila* EGF receptor homolog transcripts. 46, 1091-1101.
92. Schnepf, B., Donaldson, T., Grumblin, G., Ostrowski, S., Schweitzer, R., Shilo, B.Z., Simcox, A., 1998. EGF domain swap converts a *Drosophila* EGF receptor activator into an inhibitor. *Genes Dev* 12, 908-913.
93. Schupbach, T., 1987. Germ line and soma cooperate during oogenesis to establish the dorsoventral pattern of egg shell and embryo in *Drosophila melanogaster*. . *Cell* 49, 699-707.
94. Schupbach, T., Roth, S., 1994. Dorsoventral patterning in *Drosophila* oogenesis. *Current opinion in genetics & development* 4, 502-507.
95. Shilo, B.Z., 2005. Regulating the dynamics of EGF receptor signaling in space and time. *Development* 132, 4017-4027.
96. Sigismund, S., Argenzio, E., Tosoni, D., Cavallaro, E., Polo, S., Di Fiore, P.P., 2008. Clathrin-mediated internalization is essential for sustained EGFR signaling but dispensable for degradation. *Dev Cell* 15, 209-219.
97. Spradling, A.C., 1993. Developmental genetics of oogenesis. In: *The Development of Drosophila melanogaster*. Plainview: Cold Spring Harbor Laboratory Press.
98. Stenmark, H., 2009. Rab GTPases as coordinators of vesicle traffic. *Molecular Cell Biology* 10, 513-525.
99. Tanaka, T., Zhou, Y., Ozawa, T., Okizono, R., Banba, A., Yamamura, T., Muraguchi, A., Sakurai, H., 2017. Ligand-activated epidermal growth factor receptor (EGFR) signaling governs endocytic trafficking of unliganded receptor monomers by non-canonical phosphorylation. *Journal of Biological Chemistry* 293, 2288-2301.
100. Thorvaldsdottir, H., Robinson, J.T., Mesirov, J.P., 2013. Integrative Genomics Viewer (IGV): high-performance genomics data visualization and exploration. *Briefings in bioinformatics* 14, 178-192.
101. Tomancak, P., Beaton, A., Weizmann, R., Kwan, E., Shu, S., Lewis, S.E., Richards, S., Ashburner, M., Hartenstein, V., Celniker, S.E., Rubin, G.M., 2002. Systematic determination of patterns of gene expression during *Drosophila* embryogenesis. *Genome biology* 3, RESEARCH0088.

102. Tsuneizumi, K., Nakayama, T., Kamoshida, Y., Kornberg, T.B., Christian, J.L., Tabata, T., 1997. Daughters against dpp modulates dpp organizing activity in *Drosophila* wing development. *Nature* 389, 627-631.
103. Turing, A.M., 1952. The chemical basis of morphogenesis. *Phil. Trans. R. Soc. Lond.* 237, 37-72.
104. Twombly, V., Blackman, R.K., Jin, H., Graff, J.M., Padgett, R.W., Gelbart, W.M., 1996. The TGF-beta signaling pathway is essential for *Drosophila* oogenesis. *Development* 122, 1555-1565.
105. Van Buskirk, C., Schupbach, T., 1999. Versatility in signalling: multiple responses to EGF receptor activation during *Drosophila* oogenesis. *Trends Cell Biol* 9, 1-4.
106. Vieira, A.V., 1996. Control of EGF receptor signaling by clathrin-mediated endocytosis. *Science* 274, 2086-2089.
107. Ward, E.J., Berg, C.A., 2005. Juxtaposition between two cell types is necessary for dorsal appendage tube formation. *Mech Dev* 122, 241-255.
108. Ward, E.J., Zhou, X., Riddiford, L.M., Berg, C.A., Ruohola-Baker, H., 2006. Border of Notch activity establishes a boundary between the two dorsal appendage tube cell types. *Dev Biol* 297, 461-470.
109. Wasserman, J.D., Freeman, M., 1998. An autoregulatory cascade of EGF receptor signaling patterns the *Drosophila* egg. *Cell* 95, 355-364.
110. Weiss, A., Charbonnier, E., Ellertsdottir, E., Tsirigos, A., Wolf, C., Schuh, R., Pyrowolakis, G., Affolter, M., 2010. A conserved activation element in BMP signaling during *Drosophila* development. *Nat Struct Mol Biol* 17, 69-76.
111. White, P.F., Shreter, D.M., Kolahi, K.S., Classen, A., Bilder, D., Mofrad, M.R.K., 2009. Quantitative Analysis of Epithelial Morphogenesis in *Drosophila* Oogenesis: New Insights Based on Morphometric Analysis and Mechanical Modeling. *Dev Biol* 331, 129-139.
112. Wittes, J., Schupbach, T., 2019. A gene expression screen in *Drosophila melanogaster* identifies novel JAK/STAT and EGFR targets during oogenesis. *G3* 9, 47-60.
- 113.

114. Wolpert, L., 1989. Positional information revisited. *Development* 107, 3-12.
115. Xi, R., McGregor, J.R., Harrison, D.A., 2003. A gradient of JAK pathway activity patterns the anterior-posterior axis of the follicular epithelium. 2003 *Dev Cell*.
116. Yakoby, N., Bristow, C.A., Gong, D., Schafer, X., Lembong, J., Zartman, J.J., Halfon, M.S., Schupbach, T., Shvartsman, S.Y., 2008a. A combinatorial code for pattern formation in *Drosophila* oogenesis. *Developmental cell* 15, 725-737.
117. Yakoby, N., Bristow, C.A., Gouzman, I., Rossi, M.P., Gogotsi, Y., Schupbach, T., Shvartsman, S.Y., 2005. Systems-level questions in *Drosophila* oogenesis. *Syst Biol (Stevenage)* 152, 276-284.
118. Yakoby, N., Lembong, J., Schupbach, T., Shvartsman, S.Y., 2008b. *Drosophila* eggshell is patterned by sequential action of feedforward and feedback loops. *Development* 135, 343-351.
119. Yang, Y.-P., Ma, H., Starchenko, A., Huh, W.J., Li, W., Hickman, F.E., Zhang, Q., Franklin, J.L., Mortlock, D.P., Fuhrman, S., Carter, B.D., Ihrie, R.A., Coffey, R.J., 2017. A Chimeric Egr Protein Reporter Mouse Reveals Egr Localization and Trafficking In Vivo. *Cell Reports* 19, 1257-1267.
120. Yewale, C., Baradia, D., Vhora, I., Patil, S., Misra, A., 2013. Epidermal growth factor receptor targeting in cancer. *Biomaterials* 34.
121. Zak, N.B., Shilo, B.Z., 1992. Localization of DER and the Pattern of Cell Divisions in Wild-Type and *Ellipse* Eye Imaginal Discs. *Developmental Biology* 149, 448-456.
122. Zartman, J.J., Cheung, L.S., Niepielko, M.G., Bonini, C., Haley, B., Yakoby, N., Shvartsman, S.Y., 2011. Pattern formation by a moving morphogen source. *Phys Biol* 8, 045003.
123. Zartman, J.J., Kanodia, J.S., Cheung, L.S., Shvartsman, S.Y., 2009. Feedback control of the EGFR signaling gradient: superposition of domain-splitting events in *Drosophila* oogenesis. *Development* 136, 2903-2911.
124. Zhang, F., Wang, S., Yin, L., Yang, Y., Guan, Y., Wang, W., Xu, H., 2015. Quantification of Epidermal Growth Factor Receptor Expression
125. Level and Binding Kinetics on Cell Surfaces by Surface Plasmon Resonance Imaging. *Analytical Chemistry* 87, 9960-9965.

Chapter 7: Curriculum Vitae

Nicole T. Revaitis

11 Harvard Rd., Atco, NJ 08004

Ph:609-742-6640 Email: npope321@gmail.com

Academic Background

- **2019 Ph. D. in Computational and Integrative Biology**
Rutgers University, Camden, NJ
- **2013 M.S. in Biology**
Rutgers University, Camden, NJ
- **2008 B.S in Biology**
The Richard Stockton College of NJ

Professional Experience

Graduate Assistant, Rutgers University 2017-2019

Teaching Assistant, Rutgers University 2012- 2017

Administrative Assistant, Brite Idea Energy 2010-2011

Administrative Assistant, U.S. Census Bureau 2010

Teachers Assistant, Camden County Educational Services Commission 2009

Publications

N. Revaitis, R.A. Marmion, M. Farhat, V. Ekiz, W. Wang, N. Yakoby. . 2017. G3 7:2705-2718; PMID: 28634244

Presentations

A quantitative analysis of EGFR dynamics during *Drosophila* oogenesis* GSA Annual *Drosophila* Research Conference (Dallas, Tx, Mar. 2019).

Modelling the spatiotemporal dynamics of EGFR activation in the follicular epithelium.[^] Mid-Atlantic Bioinformatics Conference, The Children's Hospital of Philadelphia (Philadelphia, PA Oct. 2018)

The dynamics of the EGFR signaling activation in the follicular epithelium.* GSA Annual *Drosophila* Research Conference (Philadelphia, PA, Apr. 2018).

Parameter selection for modelling EGFR signaling dynamics during *Drosophila* oogenesis.* Annual Graduate Student Symposium, Rutgers University (New Brunswick, NJ Mar. 2018)

Discrete regulatory regions control the expression of simple patterns during *Drosophila* oogenesis.[^] GSA Annual *Drosophila* Research Conference (San Diego, CA Mar. 2017)

Analysis of cis-regulatory modules in *Drosophila* oogenesis: patterning, distributions, and perturbation tools.[^] Society of Developmental Biology Annual Conference (Boston, MA Aug. 2016)

Developing a mathematical model to analyze the evolution of morphologies mediated by the TGF-alpha-like ligand Gurken.[^] GSA Annual *Drosophila* Research Conference (Chicago, IL Mar. 2015)

Gene regulatory domains controlling eggshell patterning are enriched within the first intron.[^] GSA Annual *Drosophila* Research Conference (San Diego, CA Mar. 2014)

Screening for regulatory sequences that pattern the *Drosophila* eggshell.[^] GSA Annual
Drosophila Research Conference (Washington, D.C. Mar. 2013)

**Invited talks ^Poster*

Awards

2018 **Best Student Paper Award** | Rutgers University-Camden

2017 **Dissertation Fellowship Award** | Rutgers University-Camden

2015-2018 **Professional Development Fund** | Rutgers University-Camden

2011-2016 **Dean's Office Research Fund** | Rutgers University-Camden

2006-2008 **Phi Theta Kappa Scholarship** | The Richard Stockton College of NJ

2006 **Environmental Science Scholarship** | Municipal Utilities Authorities Association