Next Generation Sequencing Guided SNP Mapping

BY

John Hasper

A thesis submitted in partial fulfillment of the requirements needed for completing the Degree of Master in Biology from the Department of Biology at The State University of New York at Fredonia, Fredonia, NY

[redacted]
Research Advisor
Department of Biology

[redacted]
Thesis Committee Member
Department of Biology

[redacted]
Dr. Matthew Fountain
Thesis Committee Member
Department of Chemistry & Biochemistry

[redacted]
Dr. Theodore Lee
Thesis Committee Member
Department of Biology

[redacted]
Department of Biology

Dedication

For my Parents,

For always being there and encouraging me

Acknowledgements

First and foremost, I would like to thank my adviser, mentor, teacher, and friend, Dr. Scott Ferguson. Dr. Ferguson was the reason I came to SUNY Fredonia. His passion for genetics was evident from our first meeting so it goes without saying that it was his enthusiasm that hooked me. Prior to becoming an official member of Ferg Lab, I finagled my way in, aiding the student researchers in their experiments. Presumably, after seeing how much of a valuable asset I was, Dr. Ferguson finally relented and accepted me into his lab. I've learned so much from being in his lab over the years and only some of which I would like to forget (phenotyping over 53,000 fruit fly eggs). There has definitely been some great times, like when I dissected more ovaries in an hour than he and Ramses, or the three hundred separate occasions in which we were sure we had found the correct causative mutation of *CA231*. With his advice and friendship I was able to persevere through every one of those candidate rejections, balance my teaching load, research, classes, and personal life. He has my thanks and respect for always being there for me.

Very special thanks to Breanna Myers. Thank you for enduring the tedious task of making all of the recombinant flies and for never missing a virgin collection. Thank you for supporting, encouraging and pushing me. Words cannot convey how much I appreciate everything you have done for me.

Ramses Rodriguez, I don't know if I could have made it through the program without you. Even though you only had a small area in my office (sorry about that), you were always there for me, encouraging me and being excited for me. Taking classes with you helped me learn, not only from the studying together, but by providing competition. You have been a brother and a great friend. Thank you for going through this program with me and putting up with me all of this time. I will try and make great use of all that you have taught me, I know that I will at least have the writing equipment to do so.

I would like to thank the rest of my committee, Dr. Nicholas Quintyne, Dr. Theodore Lee, and Dr. Matthew Fountain for helping me in many ways. I knew I could always get sound advice on experimental design, teaching, grading, and a host of student dilemmas. Dr. Quintyne, thank you for never turning me

away all of those times I would randomly come by your office. At least sixty percent of those times I had academic questions, the other forty percent always ended in laughter.

Special thanks to Ed McCarrick, for constantly bringing joy into work, and for all of the times I ran away expecting something to fly out after me.

Thank you to all of the past Ferg Lab members who have worked on *CA231*, and those who have helped train me or have worked with me. Phil, you'll do wonders at Berkley! To the present Ferg Lab members, keep at it, and saver the small victories. Remember as you are working in Ferg Lab, that the sundews need to be watered- bottom watering is best!

To the rest of the Biology Department, it has been a pleasure and an honor to take classes from you, to learn from you, and to work beside you. Thank you for helping me any time I needed it.

I'd like to extend a big thank you to my family, all of whom have helped me immensely and have always cheered me on. Thank you for being there for me and being understanding when I became busy and didn't call as often as I should have.

Lastly, and certainly least, I'd like to acknowledge Julie Hens for saying I wouldn't do well in Genetics. Thank you!

Abstract

Oogenesis is dependent on precise translational control and localization of numerous morphogens within the oocyte to achieve faithful patterning. Gurken, (Grk) is one such protein and is responsible for specification of the dorsal/ ventral axis. Mutations in the *spindle-B gene* results in inefficient *gurken* translation due to activation of a meiotic DNA damage checkpoint. This checkpoint activation inhibits the Vasa RNA helicase, an essential *grk* translation factor. Without proper Gurken levels, the egg chambers develop defects, the most severe being complete ventralization. A 2004 forward genetic screen targeting the 3rd chromosome identified thirty nine unique mutants in a *spn-B*^{BU} mutant background. Two of these lines had already been mapped, the other lines were screened for their ability to suppress the ventralized *spn-B*^{BU} phenotype and therefore stimulate *grk* translation. Eggs laid by homozygotes from each of the isogenized lines were scored for their dorsal/ventral polarity and compared to those of the control group of *spn-B*^{BU} homozygotes. We have taken advantage of a next-generation sequencing approach to identify candidate mutations in 10 independent lines from a forward genetic screen for regulators of dorsal ventral patterning during *Drosophila* oogenesis. Through a partnership with Roswell Park Cancer Institute in Buffalo, NY, the best suppressor lines were subject to whole-genome re-sequencing using the Illumina HiSeq 2000 platform. A large scale mapping experiment was started, creating recombinant flies for six of the lines. After multiple universal markers were developed to distinguish these chromosomes from the mapping line, a focus was placed on one of the suppressor lines, *CA231*. A previous mapping experiment on this line placed the mutation toward the end of the right arm. A higher density map was made for this area. The screen was limited by the number of recombinants that showed variation in this

area. While the causative mutation has yet to be found, the pool of candidate mutations has been vastly diminished. Furthermore, additional focused mapping projects have been started from the recombinants made in this experiment, using a subset of the markers that are shared with *CA231* as a starting point.

Table of Contents

Dedication ii
Acknowledgements iii
Abstract v
Table of Contents vii
List of tables and figures ix
Oogenesis & Spindle Class Mutations 1
Oogenesis 1
Defects in Dorsoventral Patterning 6
Mapping Techniques 9
Methods 12
Isogenization 12
Genomic DNA (gDNA) Preparation for Routine PCR 13
Primer Design 14
Amplification of the Target Region 14
Taq Polymerase PCR (Single Reaction) 15
Beaded PCR 16
Restriction Digest (20 µL Reactions) & Visualization on an Agarose Gel 17
Analyzing Egg Shell Morphology 19
Determining Dominant Suppression 21
Apple Juice Media for Condos (0.5L Batch, 16-20 Plates) 21
Nextera Tagmentation Full Genome Sequencing 22
Characterization of Sequencing Data 22
Making Recombinant Fly Lines 24
Initial Screening of the Recombinant Lines 25
Creating Universal and Line-Specific SNP Markers 26
Haplotyping Recombinant Flies 27
Candidate Testing 28
Application of Mapping Techniques to the *CA231* Suppressor of *spn-B*^{BU} 29
Identifying Ideal Novel *spn-B*^{BU} Suppression Lines 29
Next Generation Sequencing 38
Characterizing the Haplotype of *CA231* Recombinant Chromosomes 41
Constructing the Haplotype Chromosomes 54

Specific Mutation Testing 59
Conclusions and Future Directions 62
Bibliography 64

List of tables and figures

Figure and Table Title	Page Numbers
Figure 1: Ovary Anatomy	2
Figure 2: Germline cyst Diagram	2
Figure 3: Gurken Signaling Cascade	5
Figure 4: Egg Shell Classifications	6
Figure 5: *grk* Translation Inhibited by $spnB^{BU}$	7
Table 1: Isogenization Crossing Scheme	12
Table 2: Taq PCR Reaction list	15
Table 3: Bead PCR	17
Table 4: Restriction Enzyme Digests	18
Figure 6: Condo Design	20
Table 5: Dominant Suppression Crossing Scheme	21
Table 6: Sequencing Result Format	23
Figure 7: Recombinant Fly Crossing Scheme	25
Table 7: Novel Fly Lines	29
Figure 8: spn-BBU Assessment	30
Figure 9: Isogenizing cross	31
Figure 10: Primary Mutation Retest	31
Figure 11: Additional BU retests	33
Table 8: Egg Shell Analysis	34
Figure 12: Visualization of Table 8	35
Table 9: Statuses of Novel Suppressor Lines After Preliminary Tests	37
Table 10: NGS Preparation	38
Table 11: NGS Total Differences	39
Table 12: Recombinant Lines of CA231	42
Table 13: Recombinant Suppression Analysis	43
Figure 13: Recombinant Suppression Analysis	43
Table 14: Complementation Test of Suppressors	44
Figure 14: Complementation Test of Suppressors	44

Figure 15: Chromosome Map with Tested Loci 46
Figure 16: SNP C 47
Figure 17: SNP E 47
Figure 18: SNP 6 48
Figure 19: SNP F 48
Figure 20: SNP B 49
Figure 21: SNP 2 49
Figure 22: SNP 3 50
Figure 23: tsl 50
Figure 24: btz 51
Figure 25: SNP 4 51
Table 15: Haplotypes Characterized by Sequencing 53
Table 16: Haplotypes Characterized by Phenotypic Markers 53
Figure 26: 3L Haplotype Maps 56
Figure 27: 3R Haplotype Maps 58
Table 17: Knock Out Allele Condo Data 60
Table 18: MBJD & Recombinant 36 Complementation Test 61
Figure 28: MBJD & Recombinant 36 Complementation Test 61

Oogenesis & Spindle Class Mutations

Oogenesis

Developmental processes are complex and if a disturbance occurs, development can be seriously disrupted. This is true for all species and at all stages of development. By narrowing the focus to *Drosophila melanogaster* and oogenesis, the number of interacting cell types can be minimized to improve the quality of cause-and-effect relationships. *Drosophila oogenesis* is a superb model for characterizing interactions related to RNA localization, cytoskeletal rearrangement, translational control, cell migration and more.

Oogenesis is a complex process that occurs in a precise and symmetric way. Within each ovary there are, usually, sixteen ovarioles. Ovarioles run parallel to each other, with their thinnest section being at their anterior and then broadening out posteriorly. The broadening is due to the progressive development of the eggs, which start anteriorly and then exit posteriorly. (Figure 1, Left) The germarium is at the anterior-most end of the ovary and at the tip are one or two stem cells (Figure 1, Right). As the germ-line stem cells divide, they do so asymmetrically, producing a daughter stem cell and a cytoblast. The cytoblasts are pushed posteriorly by other stem cell divisions and by muscular contraction around the ovariole. As the cytoblasts travel, they divide four times. These divisions do not complete cytokinesis, and ring canals are formed connecting the daughter cells. The sixteen-celled cyst has two cells that both have four connections (Figure 2). As all sixteen cells enter a long pre-meiotic S phase, one of these two cells will become the oocyte. The other fifteen cells do not complete meiosis but instead differentiate into nurse cells. These nurse cells have quick growth, do not divide, and develop polytene chromosomes. These nurse cells will provide the oocyte with proteins, mRNAs and other necessary molecules. For additional information see Spradling, 1993.

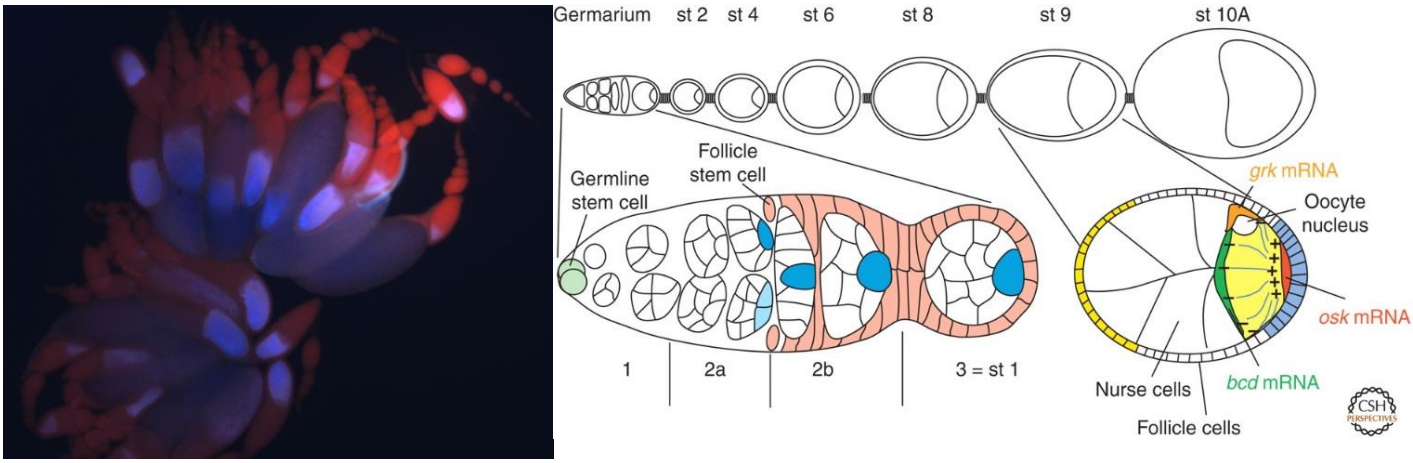


Figure 1: Ovary Anatomy (Left): A dissected ovary. Each string of cells is an ovariole. Vitellogenin, shown in blue, highlights the oocyte, Red Stain is DRAQ5 and marks nuclear proteins (Picture is courtesy of Dr. S. Ferguson.)

Ovariole Anatomy (Right): Top: diagram of an ovariole, the germarium is at the anterior tip and sequential egg chambers are of increasing age. Bottom: Magnified view of germarium and stage 9 egg chamber. In the germarium dark blue marks the oocyte, pink the follicle cells. The Stage 9 egg chamber depicts the three main cell types (names in black) and three mRNAs necessary during dorsal ventral patterning. (González-Reyes, Elliot, and St. Johnston, 1997)

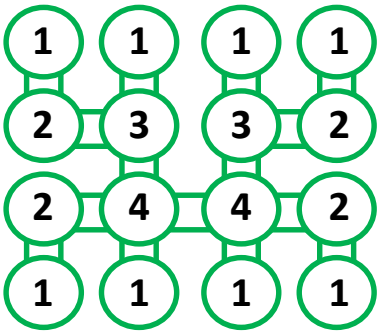


Figure 2: Germline cyst Diagram: A developing, sixteen celled cyst will have interconnected cells in the shown orientation. Eight cells have only one connected cell. Four cells have two connections. Two cells have three connections. Two cells have four connections and it is from one of these cells that the oocyte will develop.

While the germline cyst is forming and the oocyte cell is being determined, there are somatic follicle stem cells that are continually dividing, creating a thin barrier in the middle of the germarium (Figure 1 Right). As the cyst continues to develop it is pushed into this barrier and it continues to move posteriorly. As few as sixteen follicle cells will initially contact the cyst and begin to divide to create a protective epithelial sheet encapsulating the cyst. The space in the posterior end of the germarium is compact as multiple cysts are surrounded by their own follicle cells. As a result, the follicle cells of neighboring cysts can contact one another. Once this encapsulated cyst starts to bud off from the germarium, it is considered to be in its first developmental stage of an egg chamber.

There are a total of fourteen developmental stages an egg chamber must progress through before exiting the ovariole as a mature egg. As the egg chamber traverses the ovariole, development of the nurse cells and the oocyte continue to diverge in their respective roles. By Stage 3, the oocyte represses all transcription capabilities and relies on nurse cells for the supply of ribosomes, mRNA and necessary proteins. This supply flows from the nurse cells to the oocyte via the ring canals. This transcription repression is facilitated by the chromosomes condensing into a karyosome, a single tightly wound mass that transcription factors cannot breach. As the oocyte is becoming transcriptionally quiescent, the nurse cells are becoming polyploidy and by stage four, they have lost all of their nuclear organization (As Reviewed in Spradling, 1993).

Near the end of Stage 5, there are around a thousand follicle cells. The follicle cells cease dividing at Stage 5, but to continue to maintain a protective monolayer around the egg chamber and become polyploid. Follicle cells can undergo as many as five rounds of endoreplication (S phase and mitosis, but no cytokinesis). These rounds of endoreplication keeps the follicle cells large enough to remain a monolayer. These cells will eventually differentiate further due to molecular interactions between them and the nurse cells and oocyte.

During Stages 7 and 8 the anterior / posterior patterning of the oocyte is determined. During Stages 9 and 10 the dorsal / ventral patterning is established. Both of these patterning events require communication between the follicle cells, the nurse cells and the oocyte. These patterning events not only lay a foundation for the organization of the oocyte, they are conserved and built upon during embryogenesis. One of the visual consequences of this patterning event is the formation of the dorsal appendages. A pair of modified follicle cells are shaped into evenly spaced tubes that will aide in gas exchange if the egg becomes submerged in the food it was laid on (As Reviewed in Spradling, 1993).

Dorsal / ventral patterning hinges on the proper localization and translation of *gurken* (*grk*) mRNA. By Stage 10, the nucleus has moved to the future dorsal anterior of the oocyte. Gurken is brought into

the oocyte as translationally quiescent mRNA and is localized around the nucleus (Neuman-Silberberg, F.S. and Schüpbach, T., 1993). Another factor needed for proper patterning is Orb, which has been shown to bind mRNAs including *gurken* (Christerson and McKearin, 1994). Orb aids in the proper localization of *grk* to the dorsal anterior of the oocyte where it can interact with its translation factors (Davidson et al. 2016). One of the proteins expressed on the membrane of the follicle cells is the Epidermal Growth Factor Receptor (EGFR) and Gurken is one of its ligands. All of the follicle cells express EGFR, which is the reason why the localization of Gurken is vital. When Gurken binds to EGFR, it sets off a MAP-kinase dependent mitotic cascade (Ray and Schüpbach, 1996). Since Gurken is localized around the nucleus, only a small set of follicle cells participate in this signaling event. ERK causes rhomboid to be translated, which will cleave Star (Figure 3). Spitz is one of the parts of Star and is shuttled out of the follicle cell. Being another ligand for EGFR, Spitz will then bind nearby EGFRs and initiate the cascade again. This positive feedback loop progresses, expanding the region in which ERK has a high concentration. . When the concentration of ERK reaches a critical concentration, Argos is produced and shuttled out of the cell. Since Argos production is dependent on ERK concentration, Argos will first be produced at the peak of EGFR activation. Argos will bind any free-floating Gurken or Spitz, stopping the positive feedback loop. Where the positive feedback loop produced a single peak of ERK, cutting off the positive feedback loop produced twin peaks of ERK concentration. These peaks align with only a subset of the follicle cells that were initially part of this signaling event. This refined area marks the dorsal side of the oocyte. The follicle cells that are in this area subsequently progress through other signaling cascades causing them to develop into the dorsal appendages (Wasserman et al. 1998, Peri et al. 1999, Sapir et al. 1998). Other proteins that are involved in additional signaling cascades that aid in defining the dorsoventral patterning include decapentaplegic (dpp) and vein, another ligand for the EGFR. Defining where the dorsal appendages develop along the anterior / posterior axis is done in part through dpp's concentration and signaling cascades (Twombly et al. 1996, Deng et al. 1997, Queenan et al. 1997, Peri et al. 2000, Dequier et al. 2001).

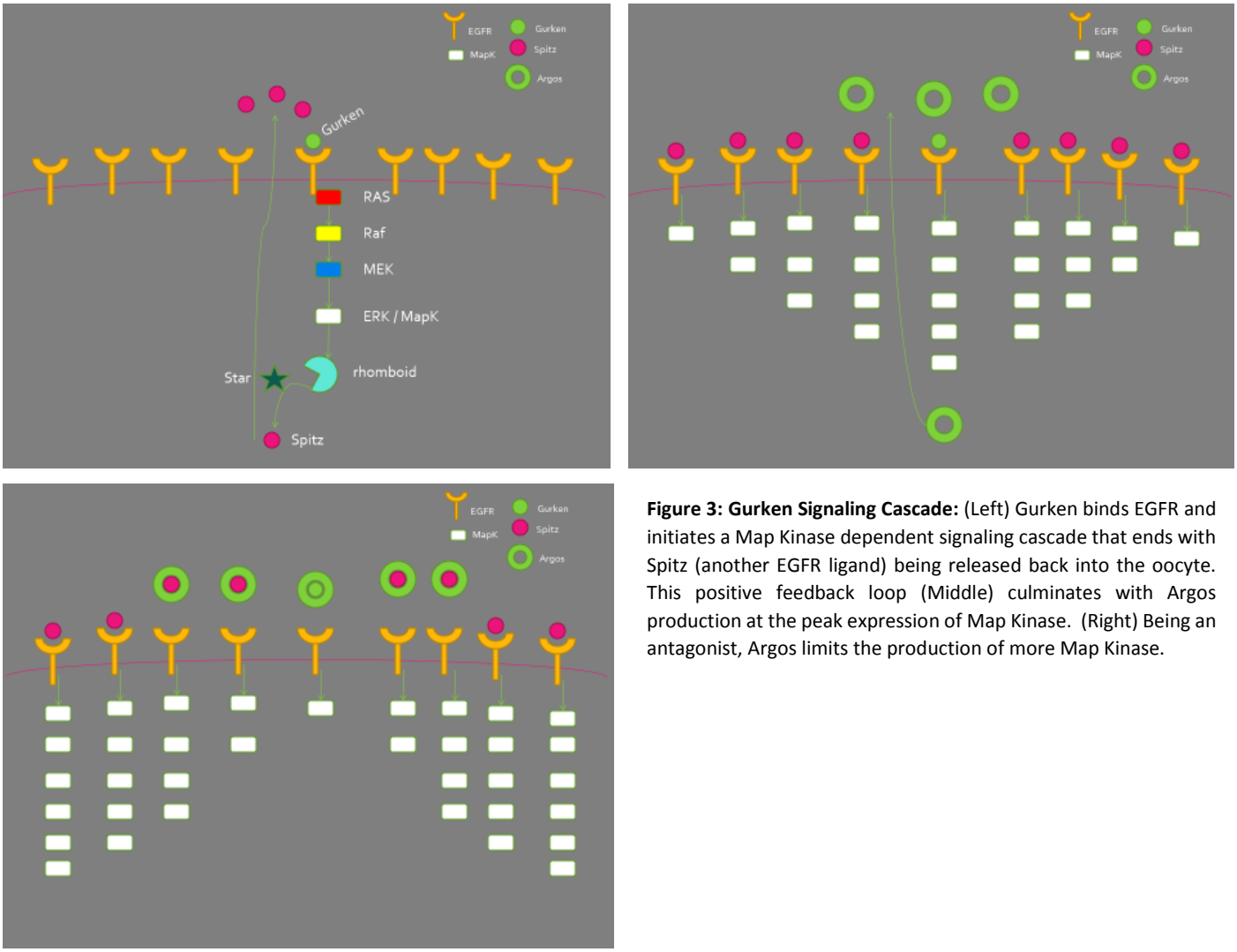


Figure 3: Gurken Signaling Cascade: (Left) Gurken binds EGFR and initiates a Map Kinase dependent signaling cascade that ends with Spitz (another EGFR ligand) being released back into the oocyte. This positive feedback loop (Middle) culminates with Argos production at the peak expression of Map Kinase. (Right) Being an antagonist, Argos limits the production of more Map Kinase.

The concentration of Gurken directly affects the dorsal / ventral patterning, which can be seen by alterations to the dorsal appendages. Proper localization and concentration of Gurken will result in evenly spaced dorsal appendages. At a molecular level, these dorsal appendages correlate to where the peaks of MAP kinase signaling are located. When there is not enough Gurken translated, the initial strength of signaling cascade is diminished, resulting in the twin peaks being closer together. Since there is a direct connection with the MAP kinase peaks and dorsal appendages, the dorsal appendages develop closer together. A very minimal Gurken concentration would not progress to the peak splitting step and would

macroscopically be seen as fused appendages or even a single appendage. If the Gurken concentration is too low or absent, the signaling cascade might not even activate and the dorsal appendages may not form at all. Figure 4 shows four classes of dorsal appendage phenotype (Ghabriel *et al*. 1998). The class 1, wild type, appearance is shown in the first scanning electron micrograph, the appendages separated nicely and paddles forming at the tips. Class 2, the appendages get closer and can even fuse at the base. If the appendages are fused at a location beyond the end of the egg, it is classified as Class 3. The fourth class is an egg devoid of not only its dorsal appendages, but also its whole dorsal side. These eggs only develop ventral aspects and are called ventralized eggs (Figure 4, right panel).

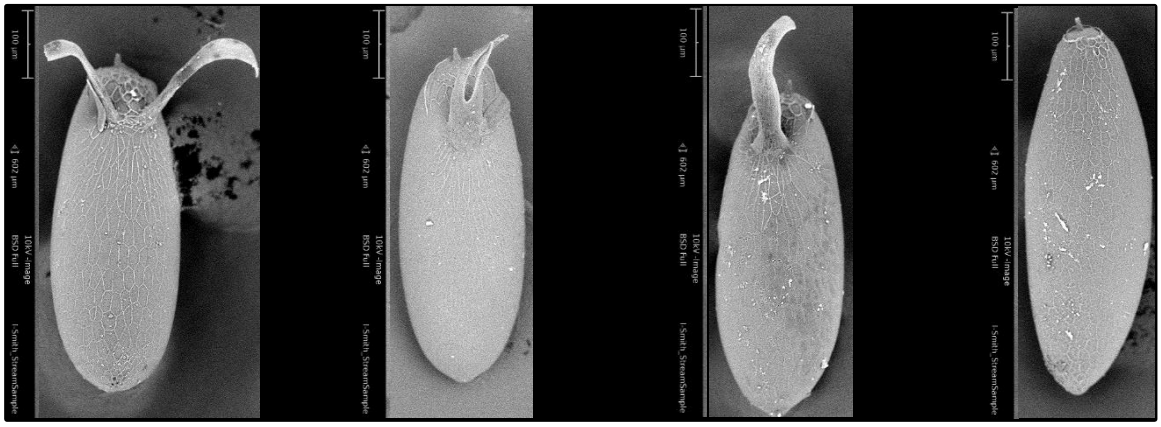


Figure 4: Egg Shell Classifications. From left to right, four phenotypic classes are shown from a wild type egg being Class 1. Class 2 eggs are differentiated by the bases of the dorsal appendages being abnormally close together or even fused. Once this fusion extends passed the main body, the egg is Class 3. Class 4 eggs show no dorsal appendages and are completely ventralized.

Defects in Dorsoventral Patterning

The dorsal ventral patterning event is pivotal and defects here can be lethal for the embryo. To ensure that proper patterning occurs, *gurken* must be translated to provide a proper concentration. Mutations in spindle class DNA repair enzymes, like *spindle B (spnB), spindle D (spnD),* and *okra (okr),* have been shown to have an effect dorsoventral patterning. These effects cause a slight mislocalization of *gurken* mRNA, but more importantly there is a dramatic reduction in Gurken translation. Since Orb and

other oocyte specific proteins are still translated properly, this effect is directed at Gurken (Ghabrial and Schupbach 1999). The entire pathway has yet to be elucidated, therefore finding these missing interactions has been the focus of multiple projects. Figure 5 highlights the main portions of this pathway and indicates the areas that are not known.

Mechanistically, a null mutation in one of the spindle class genes, specifically in *spindle B,* inhibits the repair of the double stranded breaks (DSBs) created during meiosis. The number of DSBs will increase during meiosis and this accumulation activates meiotic checkpoint. *meiotic-41, mei-41,* is a member of the ATM/ATR family (Sekelsky et al. 2000). One of the downstream effects of *mei-41* is the activation of DmChk2, the *Drosophila* Chk2 homolog. DmCk2 then has indirect downstream effects upon Dwee1 and Vasa (Abdu et al. 2002). The effect on vasa is phosphorylation by an unknown kinase and this inhibits its

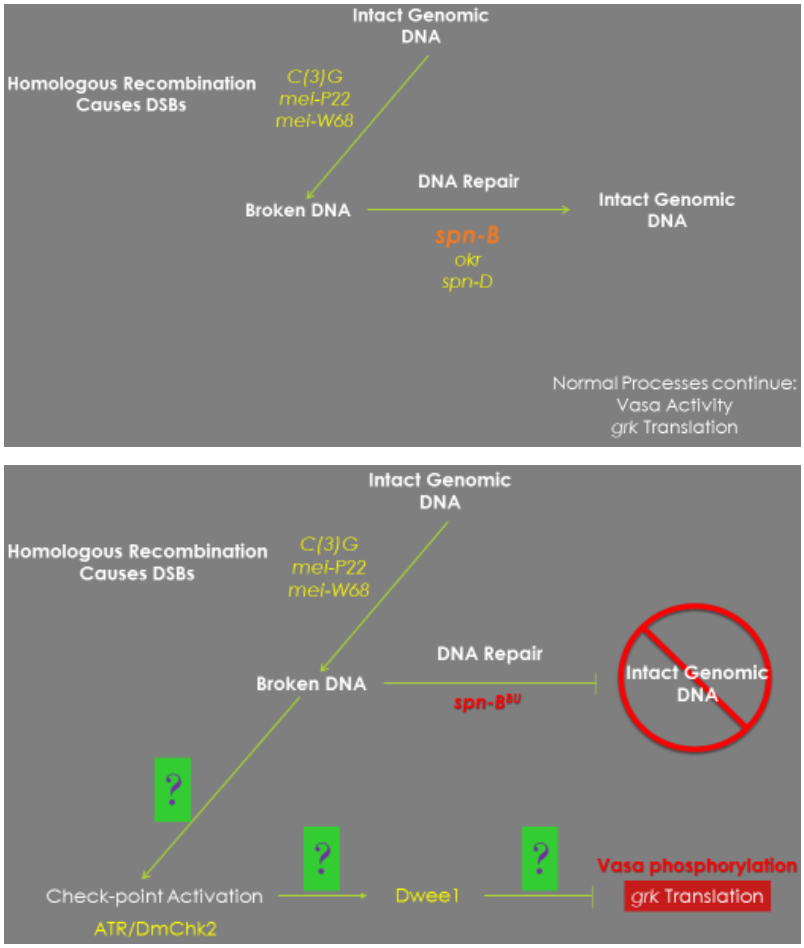


Figure 5: *grk* Translation Inhibited by *spnB*^{BU}. (Top) Homologous recombination causes double stranded DNA breaks, DSBs. C*(3)G, meiotic P22 (mei-P22), and meiotic W68 (meiW68)* are three of the proteins needed for DSB formation. Gurken is produced normally when the genomic DNA is repaired. (Bottom) *gurken* translation is inhibited by an unknown cascade starting from a buildup of double stranded DNA breaks (DSBs) when *spn-B* has a null mutation (BU). The question marks denote areas where not all of the interacting proteins are known.

activity. As vasa is one of Gurken's essential translational factors, the phosphorylation blocks Gurken translation (Abdu et al. 2002, Ghabrial and Schupbach, 1999).

A forward genetic screen was performed in 2004 to identify the unknown proteins involved in this pathway. The unknown proteins interact between DSB accumulation and *mei-41* /*DmCk2* activation and then other proteins are activated to interact with vasa. This forward genetic screen was done at Princeton by Martha Klovstad and Uri Abdu in Trudi Schüpbach's Lab. The general process was that males with a null mutation of *spindle B, spnB*^{BU}, were fed ethyl methanesulfonate (EMS), which randomly induces G to A mutations during mitosis. In a very elegant crossing scheme, third chromosomes carrying the *spn-B*^{BU} allele and additional novel EMS mutations were captured by a third chromosome balancer and established in their own unique stocks. From these stocks, females that were homozygous for these novel third chromosomes were screened for changes in their eggshell phenotypes, including suppression and enhancement of the *spn-B*^{BU} ventralization defect. Thirty-nine lines were isolated that showed suppression of the patterning defects. Since EMS is a random mutagen, the causative mutation needed to be mapped for each of the lines.

Mapping Techniques

The traditional methods of mapping novel mutations include using deficiency chromosomes, or making recombinant progeny mapping fly lines. Deficiency chromosomes have wild type genes but with known sections of chromosome deleted. Multiple deficiency chromosomes are separately crossed to the novel fly line. Mutations in the spindle class genes were first mapped using deficiencies (Schüpbach and Wieschaus 1991). All of the spindle class genes are haplosufficient and therefore mapping these mutations can be done with deficiencies. The only time a spindle mutant line would produce ventralized eggs would be when the deficiency overlapped with the mutation. In other words, when the deficiency chromosome's lack of a locus aligns with the locus the mutation resides at, the mutant allele would be the only available allele making the organism hemizygous and it will result in the mutant phenotype. When the deficiency doesn't overlap with the mutation, the deficiency provides a wild type allele, masking the mutant phenotype. At first, large deficiencies are used, and then smaller deficiencies reduce the possible areas. Deficiencies with overlapping areas can help create limiting sections. If both overlapping deficiencies produce the mutant phenotype, the mutation is in the overlapping area. Conversely, if only one deficiency causes the mutant phenotype, then the mutation is not in the overlapping area. When the area is refined to a small span containing a handful of genes, rescue experiments are done. This method could be used to find the novel suppressors of *spnB*^{BU}, but additional crossing schemes are necessary due to the fact that the BU mutation must remain homozygous. Therefore, before the deficiency chromosomes can be used the BU mutation must be recombined into each of the deficiency lines. When those balanced stocks are made, then all of the deficiency, BU double mutants will then be crossed with the suppressor lines. The use of deficiencies may not be ideal since some of the suppressors of the BU allele have dominant suppression, such as the line used for this project, *CA231*. False positives would be common, visually seeing a phenotype when the fly is heterozygous would normally indicate that the

deficiency is overlapping with the suppressive mutation, but for dominant suppression, the heterozygote could be phenotypically indistinguishable from the hemizygote.

Another way to map novel mutations is to recombine it with a mapping line. A mapping line has multiple phenotypes across its chromosome. The progeny between the novel fly line and the mapping line would retain only a subset of these phenotypic markers. Excluding progeny that didn't inherit the BU mutation, the suppression ability of the recombinants can be determined as well as which phenotypic markers they possess, effectively linking phenotypic markers to suppression ability. By assessing the recombinants, linkage between the phenotypic markers and the causative mutation can be determined. This way helps narrow down the location of the novel mutation to between two phenotypic markers. At this point target deficiency chromosomes or rescue constructs can be used to test multiple gene spans narrowing the possible area further.

Both of these processes, while very effective, can be very time consuming and difficult considering the nature of a dominant suppressor. With new technologies becoming more cost effective, full genome sequencing can enhance the ability to distinguish the locus of a causative mutation, not between two phenotypic markers, but within a few hundred kilobases or even within a single gene (Berger *et al*. 2001). In the Berger et al. study (2001), 8,615 polymorphisms were identified between their mutant line and their mapping line. The area of coverage was averaged at one marker every 114kb. Once recombinant lines were produced and screened for the mutant phenotype, the screening process took less than three days to map each of the mutations.

In this study we begin the mapping process for multiple suppressors that came out of the Schüpbach lab's 2004 Forward Genetic Screen. There have already been fruitful findings from this study with only two of these lines being mapped. One of these lines was mapped, by positional cloning, to the *mei-P22* locus. *mei-P22* is active as the synaptonemal complex is forming, binding to the chromosome and thereby marking the locations to which *me-W68* will later create double stranded breaks (McKim et

al. 2002). By preventing double stranded breaks from occurring in the first place, this mutation rendered the non-functional Spn-B enzyme unnecessary because there were no breaks to repair. The other line from the screen was also mapped by positional cloning, this one linked to the *Lnk* locus, which is a positive regulator of the insulin-like pathway (Ferguson *et al.* 2012). This finding added another layer of complexity to dorsoventral patterning, not only in the scope of the developing egg chamber, but also to the environmental conditions experienced by the adult female as the egg chamber is developing. Normally, mutations in *spindle B* activate *DmChk2* which, through an unknown kinase, phosphorylates Vasa, a necessary translation factor for *gurken* mRNA. Ferguson et al. 2012 proposed a second mode of *gurken* translation through an Internal Ribosomal Entry Site, IRES. The mutation in *Lnk* blocks TOR activation. With an inactive TOR, eIF4E-Binding Protein is able to bind eIF4E. Sequestering eIF4E effectively shuts down cap-dependent translation. Since this mutation is causing Gurken to be produced at near normal levels, a second mode of translation must be occurring (Ferguson *et al.* 2012).

In this study, preliminary groundwork has been done to map the remaining suppressive mutations. These novel fly lines were all scored for their suppressive phenotypes and those that still showed clear suppression were sent in for Next Generation Sequencing. A focus was then placed on six lines that exhibited a high degree of suppression. Recombinants were made from these lines and a preliminary set of single nucleotide polymorphisms (SNP) were determined to distinguish between the haplotype of the suppressor line and the mapping line. The focus shifted to mapping one specific line, *CA231*. This line is still being mapped, but the number of possible causative mutations has been drastically reduced.

Methods

Isogenization

When first given the novel mutants, it was important to ensure that the fly lines were not heterogeneous, and therefore isogenization was paramount. Normally a novel mutation is captured on a chromosome with its homolog being a balancer chromosome. This was the case and so a simple crossing scheme was done using another balanced stock, but with a visual phenotype on its homolog. The *ruPrica / TM6C, Sb* fly line was used for this process. The *ruPrica* homolog has many marker mutations on it, *Prickly* being one of them. This line is always seen heterozygously with the balancer chromosome *TM6C, Sb* because *ruPrica* and *TM6C, Sb* are both have recessive lethals. Crossing virgin female *ruPrica / TM6C, Sb* with heterozygous novel males, will lead to a quarter of the progeny having the novel chromosome and the *TM6C, Sb* balancer (Table 1). Males and virgin females, of this phenotype, were then used to create an isogenized stock of the novel mutation.

Table 1: Isogenization Crossing Scheme. The *ruPrica* line has *Prickly* (*Pr*) on its third chromosome and the balancer *TM6C* for its homolog. Both *Pr* and *Sb* (*Stubble*) are dominant markers. Crossing a balanced fly with a novel mutation with a *ruPrica* fly will lead to only one living type of fly that isn't *ruPrica*. These flies are then used to make a new balanced stock, all with the same 3rd chromosome (Highlighted in yellow). Homozygosity for the balancer chromosome is lethal. *Prickly* is epistatic to *Stubble*.

	Novel Mutation	*TM6C, Sb*
ruPrica	Novel Mutation *ruPrica*	*ruPrica* *TM6C, Sb*
TM6C, Sb	Novel Mutation *TM6C, Sb*	*TM6C, Sb* *TM6C, Sb*

Genomic DNA (gDNA) Preparation for Routine PCR

DNA needed to be isolated from the novel mutant line to use as a template for PCR. The best way to do this was with a single fly genomic prep on a homozygous male. In a balanced stock, heterozygous flies carry a balancer chromosome, thus two different alleles could be amplified. Therefore homozygous flies were used wherever possible. The yolk in the ovaries of female flies could also interfere with amplification so male homozygous flies are the ideal source of gDNA. Single flies were collected and physically crushed to increase the surface area. Proteinase K was used to degrade the proteins and release the gDNA. The proteinase K was then heat inactivated to prevent the degradation of DNA polymerase. At the end of the protocol the supernatant containing the gDNA was transferred to a new tube, diluted 1:5 with Milli-Q water and stored at four degrees Celsius.

Homogenization Buffer (10mL):	**Master Mix**
Tris: 10mM, pH 8.3 (100µL of a 1M stock) EDTA: 1mM, pH 8.0 (20µL of a 0.5M stock) NaCl: 25mM (50µL of a 5M stock) Nuclease Free H_2O: to 10 mL	Homogenization Buffer (Sample number +0.5) x 50µL Proteinase K (PK) (Sample number +0.5) x 1µL

- ☐ Collect single homozygous male flies (one per genotype), putting them in separate PCR tubes.
- ☐ After all of the flies are collected, put them on ice.
- ☐ After the flies settle down, add 50µL of the Master Mix into each tube.
- ☐ Mash the fly (use clean square round toothpicks or the pestles from Kontes)
- ☐ Incubate at 37°C for 1 hour
- ☐ Incubate at 100°C for 2 minutes.
- ☐ Quick spin down.
- ☐ Transfer supernatant to a clean labeled PCR tube.
- ☐ Dilute 1:5 with Milli-Q water
- ☐ Store at 4°C

Primer Design

Custom designed primers are necessary for PCR and Sanger DNA sequencing. The nucleotide sequence of the target locus was downloaded from Flybase.org. The target sequences were single nucleotide polymorphisms, small changes of a handful of nucleotides or were large insertions or deletions of nucleotides. The target nucleotide was found and primers were searched for on both sides, at least one hundred bases away from the target nucleotide. Primers were designed to have equal CG:AT content, have a CC, GG, CG, or GC clamp on the 3' end. Stretches of five or six ATs or GCs were also avoided. The sequence from fly base is given in a 5' to 3' direction. Designing forward primers prior to the target nucleotide can use the sequence as provided. Reverse primers are designed to bind downstream of the target region. Since the sequence is downloaded in the 5' to 3' direction only, for a primer to bind and extension to occur in the direction of the target region, the reverse complement was used. Note that this primer must still have the GC cap on its 3' side, which is the 5' end of the reverse complement. If the amplicon was being sent for sequencing, a third, nested, primer was designed. Primer designs were purchased through Integrated DNA Technologies (IDT) and were received dehydrated. Primers were resuspended to a concentration of 50µM, unless otherwise stated.

Amplification of the Target Region

Amplification was performed using Bead PCR from the GE Healthcare Illustra Taq Bead Kit, or from NEB 2x Taq mastermix. NEB's 2x Taq mastermix was cost effective especially when performing a lot of reactions. GE's Bead Kit is more robust for difficult amplicons. The details for these two types of PCR are described below.

Taq Polymerase PCR (Single Reaction)

The protocols in Table 2 were used when using Taq Polymerase. When there were multiple samples being amplified, the Multiple Reactions volumes were scaled up to the number of reactions (N) plus a half reaction to make a Master Mix. When a single reaction was needed, a dilution of the primers was done to have amounts that could more easily be pipetted.

Table 2: Taq PCR Reaction list. The protocol for Taq PCR is shown below. When preparing an "N" number of reactions, the protocol was scaled by N and an additional half reaction (N.5), to account for any pipetting errors. "*" denotes the components added together to make a master mix for multiple reactions. "**" both protocols have 9µL of components prior to the addition of gDNA. When a single reaction was needed, the other protocol was used to have more accurate pipetting with larger volumes.

		10 µL (Multiple Reactions)	**10 µL (Single Reaction)**
	Nuclease Free H_2O	(3.8 x N.5) µL	2.0 µL
*	Forward Primer (50µM Stock)	(0.1 x N.5) µL	1.0 µL (1:10 dilution)
	Reverse Primer (50µM Stock)	(0.1 x N.5) µL	1.0 µL (1:10 dilution)
	Taq Master Mix with Buffer	(5.0 x N.5) µL	5 µL
**		9.0 µL	9.0 µL
	Template DNA	1.0 µL (1:5 dilution)	1.0 µL (1:5 dilution)

For Taq PCR, the thermocycler was set to the following parameters:

1. 94°C, 30 seconds (Pre-cycling denaturation)
2. 94°C, 30 seconds (In-cycle denaturation)
3. 55°C, 30 seconds (Primer annealing)
4. 68°C, 40 seconds (Extension)
5. Go to #2, 29 times
6. 4°C, hold

Occasionally a primer would not amplify well, because primers can vary slightly on which annealing temperature is more efficient. When attempting to design primers, choices for primer locations were sometimes limited making it necessary to have slightly uneven AT:GC content or large stretches of either pair. This altered the annealing temperature for the primer. To identify what the most efficient annealing temperature was, a temperature gradient was set up in the thermocycler for the annealing

step, and multiple trials were tested. The highest concentration of DNA after the thirty cycles was then used when repeating the amplification step for that locus. If the temperature gradient didn't solve the problem, varying the primer concentration was tried. Lowering the primer concentration could help solve problems with nonspecific annealing.

Beaded PCR

Occasionally, the primers for a given locus would not amplify well at any annealing temperature or primer concentration for Taq polymerase. When this happened, bead PCR was used. The GE Healthcare, illustra™ PuReTaq™ Ready-To-Go™ PCR beads kit was used. Annealing temperatures did vary for some primers, and were assessed the same way as for Taq PCR. The general protocol for Bead PCR is shown below.

1. 94°C, 30 seconds (Pre-cycling denaturation)
2. 94°C, 30 seconds (In-cycle denaturation)
3. 54°C, 45 seconds (Primer Annealing)
4. 72°C, 30 seconds (Extension)
5. Go to #2, 29 times
6. 4°C, hold

This kit has a dehydrated bead containing of all of the buffers and Taq polymerase necessary for a single 25 µl reaction. The other components needed for each reaction were water, primers and gDNA (Table 3). The following protocol was used for single reactions, when multiple reactions were needed, everything for the reaction except for the gDNA was added together plus an additional half reaction. Then twenty-four microliters of this mixture was added to each tube. Even though this kit helped streamline the preparation time, it is important to know that the dehydrated PCR bead is very prone to static electricity, and it was a good idea to attempt to keep the ball at the bottom of the tube, either by flicking it and/or centrifugation. Also, since the ball is highly concentrated salts, buffers and enzymes, when

adding anything to the tube, it was vital to not touch the pipette tip to the ball or the foam that the ball creates.

	Single Reaction (25.0 µL)	
	Nuclease Free H_2O	23.5 µL
*	Forward Primer	0.25 µL
	Reverse Primer	0.25 µL
**		24.0 µL
	gDNA	1.0 µL

Table 3: Bead PCR. The protocol for Taq PCR is shown below. Scaling the single reaction protocol can be done easily by multiplying by the number of reactions. It is encouraged to add a half reaction to account for pipetting errors. "*" denotes the components added together to make a master mix for multiple reactions. "**" indicated the aliquot from the master mix needed for each sample.

Restriction Digest (20 µL Reactions) & Visualization on an Agarose Gel

After amplifying a section of DNA, visualizing the difference between a normal wild type sequence (the same as flybase) and a sequence that contains a mutation was done in a variety of ways. InDels could be assessed directly via agarose gel electrophoresis. Loci that were RFLPs first needed to undergo a restriction digest before being visualized by agarose gel electrophoresis. Enzymes needed for each RFLP site had variations on their needs, what buffers to use, what incubation temperature was more efficient, etc. Enzymes were also held at various concentrations. Each restriction digest followed the same basic protocol (Table 4), and was modified with the enzyme's unique needs. After the restriction digest, samples were then visualized via agarose gel electrophoresis. The agarose gel percentage required for each RFLP is indicated in the appendix.

Table 4: Restriction Enzyme Digests. Enzymes can have varying needs for optimizing agents. "*"Each Enzyme will have its own optimal buffer, review the neb website to find which buffer is needed. "**"Some enzymes require additional salts and cofactors to have optimized function. Please review the enzyme on the NEB if additional buffers are necessary. "***"Enzymes can be purchased in different concentrations and each aliquot of enzyme product will have so many units of enzyme per microliter. Calculate how much volume is needed for two units of enzyme. "****"While nuclease free water is the last amount calculated, it should be the first substance in the tube, followed by the enzyme buffer(s) and the enzyme. The PCR product is usually added last because a reaction master mix is usually made.

	Final Volume	**20 µL**
*	Enzyme Buffer	2.0 µL
**	Additional Enzyme Buffer	H µL
***	2 Units of Enzyme	K µL
	PCR Product	5.0 µL
****	Nuclease Free Water	Remaining Volume

Analyzing Egg Shell Morphology

When investigating oogenesis, egg shell morphology reflects the level of EGFR signaling occurring during development. In this study, the primary mutation is *spnB*^{BU}, which leads to the ventralization of a large portion of the eggs laid by homozygous mothers. After the forward genetic screen, novel mutations were captured that corrected this ventralization, but with varying degrees. To quantify how much these novel fly lines could correct ventralization, homozygous females and any males were placed in a plastic chamber with a fruit media lining the bottom. The fruit media plates were made to provide a desirable location for females to lay eggs. The fruit media also had a spot of yeast on it to provide a high protein diet. The flies ate the yeast and the females laid eggs mostly on the fruit media and some in the yeast spot. A high throughput apparatus was designed to hold and keep multiple fly lines separate. This apparatus, a "condo", is a nineteen-chambered plastic containment device that is both flat and open on one end (Figure 6). The condo and the fruit plate were lightly held together with rubber bands. The seal needed to be tight enough to leave no opening for flies to escape, but loose enough that the inside chambers didn't become humid due to compression of the agar plate. The tension on the rubber bands was adjusted by tying multiple bands together.

In each condo experiment, flies were put into their designated condo chamber, or lane. When possible, the condo experiments had control lanes, one with homozygous *spnB*^{BU} females and one with a wild type line. The ideal number of flies in each chamber was five females and three males. Any more flies made the lane too crowded, reducing the quantity or quality of eggs produced. Fresh yeast was made prior to each condo experiment as a 1:3 weight ratio of distilled water to Fleischmann's active dry yeast. Condo plates were changed twice a day for seven days. The first six plates (three days) were not counted to ensure that the females are well fed and that nutritional restriction was not affecting growing egg chambers. For each lane, all eggs were classified by hand as one of the four classes of ventralization (Figure 4). Some eggs needed to be manually rolled, or removed from the yeast, to identify their class. If

multiple lanes were filled with the same genotype then after all of the eggs were counted and classified, the data were added.

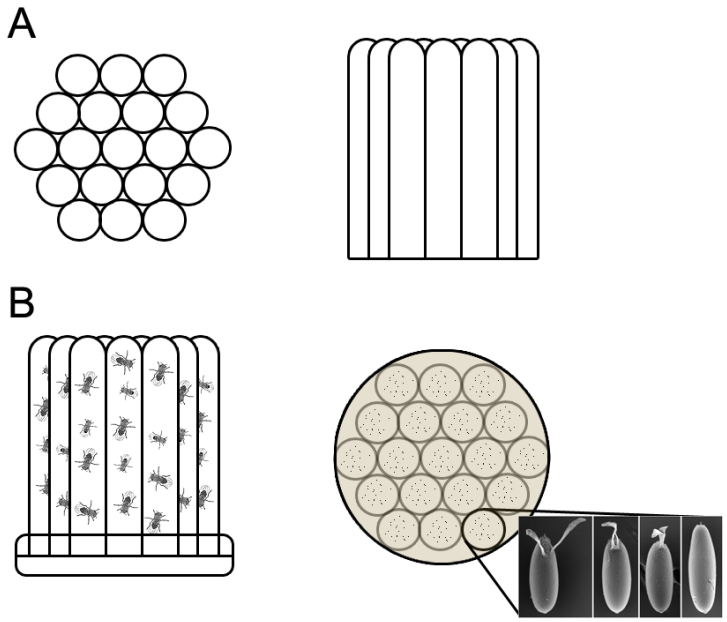


Figure 6: Condo Design. A) Bottom view of the condo, this side is open and flat (Left). A portrait view of the condo, only the bottom is open. The rounded tops have small perforations in them to allow gas exchange but too small for flies to escape through (Right). B) Portrait view of the condo resting on a fruit media plate (Left). A bottom view of the condo with a fruit media plate. Females will lay their eggs on the media (Right). Figure courtesy of Malachi Blundon (Blundon 2012).

The number of Class 1, Class 2, Class 3, and Class 4 eggs, for each individual line, were summed across the last four days of the condo experiment. A way to compare the suppression ability of different fly lines was to compare their Suppression Index. This term was calculated by summing the Class 1 and Class 2 egg counts and dividing by the sum of Class 3 and Class 4. The overall number of eggs was also counted and percentages for the four classes were graphically displayed. The homozygous *spn-B*^{BU} line was compared to all of the other lines being tested. Since this line has no suppressor, its suppression Index is between 0 and 1. If a line has a significantly larger Suppression Index than the homozygous *spn-B*^{BU} line, then that line is suppressing the BU phenotype.

Determining Dominant Suppression

These suppressive secondary mutations can either be dominantly suppressing or recessively suppressing. To differentiate between the two, a cross between the novel fly line and the reference *spn-B^{BU}* line was performed (Table 5). Females that were not stubble were collected because these flies were homozygous for *spnB^{BU}* and heterozygous for their suppressive mutation. These females were used in a condo experiment. If the fly line contained a similar suppression index, whether it is heterozygous or homozygous, then the mutation that is causing the suppression has a dominant phenotype.

Table 5: Dominant Suppression Crossing Scheme. The novel suppressing fly lines are crossed back to a *spnB^{BU}* line to capture flies that are homozygous for *spnB^{BU}* but are heterozygous for their suppressive mutation. "*" denotes the unknown suppressive mutation. Non-balanced flies are collected and used for the condo experiment.

	*spnB^{BU}, **	*TM6C, Sb*
spnB^{BU}	$\frac{spnB^{BU}, *}{spnB^{BU}}$	$\frac{spnB^{BU}}{TM6C, Sb}$
TM6C, Sb	$\frac{spnB^{BU}, *}{TM6C, Sb}$	$\frac{TM6C, Sb}{TM6C, Sb}$

Apple Juice Media for Condos (0.5L Batch, 16-20 Plates)

Erlenmeyer Flask #1 (1 Liter Volume)

- dH_2O (**375mL**)
- Bacto Agar (**12.5g**)
- A Stir Bar
- Cap the flask and apply autoclave tape

Erlenmeyer Flask #2 (250 milliliter Volume)

- Sucrose (**12.5g**)
- Mott's Apple Juice (**125mL**)
- p-Hydroxybenzoic Acid (**0.75g**)
- A Stir Bar

- Sterilize Erlenmeyer Flask #1 and all of its contents in the autoclave on a liquid cycle for 15 minutes. Be mindful of the rules for your specific machine.
- Heat Erlenmeyer Flask #2 on a stirring hot plate, removing from heat at the onset of boiling.
    - For best results, time the end of the sterilization cycle for Flask #1 with the boiling of Flask #2.
- Using sterile technique, pour Flask #2 into Flask #1
    - If possible, avoid pouring the stir bar by careful pouring or placing a magnet on the outside of Flask #2
- Place Flask #1 back on a stir plate allowing the contents to mix and cool.
- Once the media has cooled to a temperature that is comfortable to place your hand on the flask, the plates are poured.
- Pour the plates using sterile technique paying particular care to ensure that they are very flat with no lumps.

Nextera Tagmentation Full Genome Sequencing

To fully characterize the differences between the suppressor lines and the *spnB*^{BU} line, all lines were sequenced. Homozygous males were collected and put into containers with an aqueous solution of sucrose (2:1) soaking filter paper and were left for twenty four hours. This allowed time for the flies to clean themselves and digest or pass any yeast from their bodies. This prevents contamination of the genomic prep with yeast DNA. Up to twenty-five milligrams of tissue was obtained from each fly line. The flies collected were all homozygous so that the balancer chromosome wasn't sequenced as well. DNA extraction was performed by the QIAGEN DNeasy Blood & Tissue Kit where two elutions were obtained for each sample. Their concentrations were determined by Qubit Fluorometry. These elutions were sent to Roswell Park Cancer Institute in Buffalo NY for Next Generation Sequencing.

Prior to sequencing, each of the samples was fragmented and had linkers added to the fragments using the Nextera Tagmentation kit. Full Genome Sequencing was conducted on the Illumina HiSeq 2000 platform using 100 base pair paired end reads. The samples were multiplexed at a density of 7 genomes per flow cell lane. This achieved a mean read depth of 30x across the entire genome. Multiplexing was achieved by the addition of a unique 8 bp bar code sequence during sample preparation of each genotype. After sequencing, only the differences on the 3rd chromosome were kept since this is where the Forward Genetic screen captured these novel mutations

Characterization of Sequencing Data

After sequencing each fly line had its own Excel file where any variations between it and the reference fly line was recorded. Since multiple fly lines were individually sequenced at the same time and characterized at the same time, the same variations among different fly lines were also noted (Table 6).

Table 6: Sequencing Result Format. The results from the Next Gen. Sequencing were returned and categorized in the following format. Which chromosome and which arm, the position on that arm, the reference genome's sequence, the alternative sequence from the flyline in question, and the next columns is either filled with a 0, that that mutant fly line doesn't have the same mutation, or a 1, meaning that it does. The count column at the end, represents how many fly lines share that specific mutation.

CHROM	POS	REF	ALT	CE149	CE121	CA23	CB429	CA121	CA231	CR324	CD485	CR549	CA1301	count

The first step in characterizing the sequencing results was to see how close the mutations are to one another. This was done by the addition of a new column and taking the difference between the second mutation's nucleotide position and the first's. This calculation was done for all mutations. If any stretches of deviations between the novel line and the mutant were very close (a handful all being within a hundred base pairs), then this track exhibits a classic gene conversion track. Gene conversion can occur when a chromosome of a stem cell attempts to undergo recombination with a balancer chromosome during meiosis. While the full recombination will fail, a small region can be converted to the balancer allele prior to the failure. This suggests that it is a gene conversion. This can be validated by sequencing the balancer chromosome. These possible gene conversion regions were noted in another column for each of the nucleotide changes in the region. Since these possible gene conversions tracks cannot be shown to be actual gene conversion, they were given a second note as well. This note was identified the same as any non-gene conversion mutation. Each location was input into flybase.org and the location where the mutation occurred was described. This location was described in a variety of ways, having occurred outside of gene spans, inside of an intron of a gene, in the five prime or three prime untranslated region (UTR), or within a gene's exons. Some differences in the exon did not produce an amino acid change, these mutations were classified as exon synonymous and were ruled out as candidates. Changes in an exon that did produce a protein change were the exon non-synonymous and were the most likely candidates. Special notes were given to any mutation that occurred in a gene if its functions had anything to do with the oogenesis, DNA repair, translation, or other processes germane to *grk* translation.

Making Recombinant Fly Lines

The amount of information obtained when having the entirety of the third chromosome sequenced can make characterizing and testing each of these loci arduous. To focus the scope of possible candidates, a mapping experiment was done. The first step was making recombinant flies. Virgin *Oregon R* flies were crossed with heterozygous suppressor flies. The female progeny with a 3rd chromosome from the suppressor and one from *Oregon R* were the non- stubble and thus non-balanced flies. This allowed the two homologs to undergo recombination in their ovaries. This allowed these females to pass genetically mosaic recombinant chromosomes to the next generation. These chromosomes were caught and preserved by crossing these virgin females with male *ruPrica* flies. In the next generation, males that obtained the balancer chromosome also had to have received one of these recombinant chromosomes. These males were then individually crossed to virgin *ruPrica* flies. Non-*ruPrica* male and virgin female flies were gathered from each cross and used to start their respective balanced stocks (Figure 7).

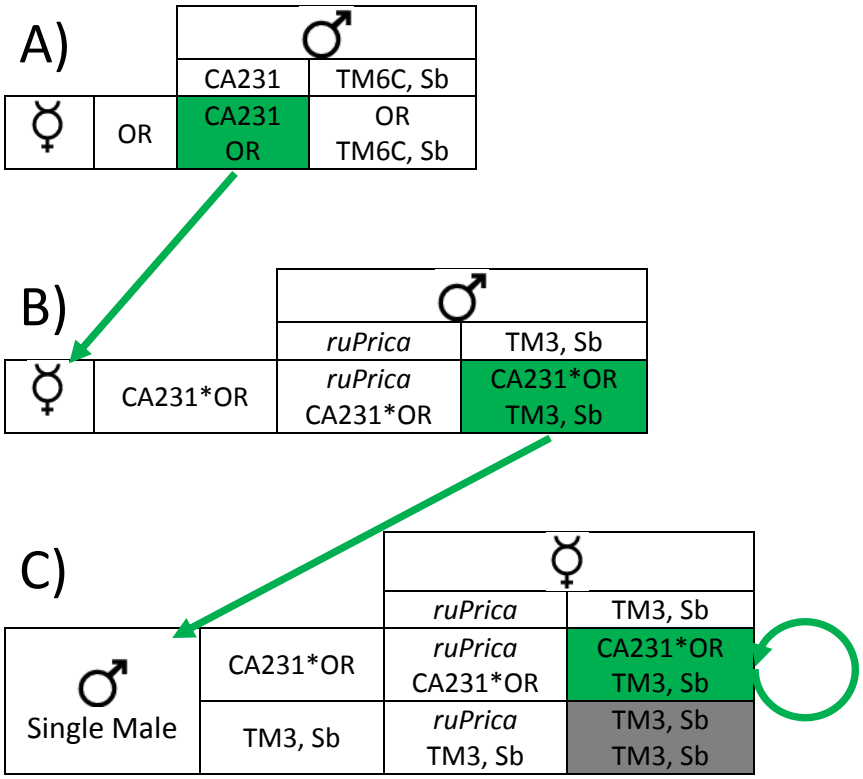


Figure 7: Recombinant Fly Crossing Scheme. In the first cross (**A**), Oregon R (OR) virgin females are crossed with heterozygous male *CA231* flies. From this cross, virgin females that are not *Stubble* were collected. These females, being unbalanced, will have recombination occur during meiosis and will pass along hybrid chromosomes, some parts randomly from *CA231* and other parts from OR (*CA231*OR*). These hybrid chromosomes are then rebalanced with the second cross (**B**). Single male flies that are *Stubble* and not *ruPrica* are then used to found their own lineage with their unique recombined chromosome (**C**). These single males are crossed with virgin *ruPrica* flies and non-ruPrica flies, males and virgin females, are used to make a balanced stock. Green depicts the flies used in the next cross. The grayed box are dead from recessive lethality of the balancer chromosome.

Initial Screening of the Recombinant Lines

Since inheriting the *spnB*^{BU} allele follows Mendelian genetics, it was expected that half of the progeny would not have this primary mutation. Therefore it was imperative that the new recombinant fly lines be tested to see if they had inherited it. This was achieved by PCR RFLP analysis for the *spn-B*^{BU} allele. Any fly lines that did not have the *spnB*^{BU} allele were discarded. To speed up the selection process, the founding males were collected and screened once larvae could be seen in their vials. After the

recombinant flies were screened for the BU allele, they were tested to see if they had retained the secondary, suppressive, mutation; this was done by performing another condo experiment.

Creating Universal and Line-Specific SNP Markers

As a result of sequencing multiple lines and comparing each of the novel lines not only to the reference but also to the other novel lines, non-unique mutations were found that all the novel lines have. These loci that were common to all of the suppressors were investigated to see if they had restriction fragment length polymorphisms (RFLP) differences or were insertion / deletions (InDels). For each of the locations, that position was found in "GBrowse" on Flybase.org, expanded so that 300 – 400 nucleotides flanked the region and then the sequence was downloaded. A duplicate sequence was made, but was edited to reflect the changes between the BU line and the suppressor line. For InDel mutations, primers were then made in the flanking region, scaling the amplicon appropriately to the size of the InDel. Markers amplifying the area around an InDel could then be easily tested by gel electrophoresis on the amplicon. If the locus was only a change in a single nucleotide, the area was investigated slightly differently. After the two sequence files were made, one for the *BU* allele and one for the suppressor, both sequences were put into pDraw and possible restriction sites were shown within thirty base pairs of the site in question. If there was any change in enzyme restriction site, this locus could be used as an RLFP. Almost all of the RFLP loci used in this experiment used enzymes already on hand in the lab.

As the sequence changes allowed, RFLPs and InDels developed from these common universal, mutations were separated by around 1 million base pairs. These markers could then be used to characterize the general haplotype of the recombinants. To augment these universal markers ability to characterize the chromosomes of the recombinants, mutations that were unique to an individual line were used to create additional RFLP and InDel markers. When the RFLP and InDel loci were exhausted, the loci of interest were characterized by direct sequencing. This was achieved by designing three primers to flank it. The two outside primers were used to amplify the region by PCR and the third, middle, primer

was sent with the amplicon for Sanger sequencing by GeneWiz. During marker development, it became advantageous to develop markers from the protein changing mutations and any other mutations of interest such that complete linkage is present between the marker used for screening and these mutations.

Haplotyping Recombinant Flies

Each recombinant fly had their chromosomes characterized with the developed markers, first with the Universal and Line-Specific RFLPs and InDels, then with the sequencing markers. Identifying the regions of recombination where these recombinants lost some of the original suppressive chromosome helped build an accurate haplotype for that recombinant. Recombinants that still suppress have retained the suppressive mutation and thus the areas that share the same haplotype as the suppressive chromosome is where the causative mutation lies. Recombinants that have lost the ability to suppress have lost the suppressive mutation and thus the areas that do not share the same haplotype as the suppressive chromosome is where the causative mutation lies. For all recombinants of a single novel line, they will all share the same causative mutation. The area of interest is developed from the overlay of all of the locations of each recombinant where the causative mutation has been either been kept or lost, depending on being able to suppress or not, respectively. When evaluating the information gained from each of the recombinant's haplotype chromosome, the recombinants with the most variation were evaluated at more of the markers. The variations allowed limiting chromosomes to be identified, ones with the least amount of the original suppressive chromosome yet still suppress, and those with the most amount of the original suppressive chromosome but do not suppress. These limiting chromosomes provided most of the information for the area of interest. Any other recombinants that provided more clarity in the area of interest helped with narrowing it further.

Candidate Testing

Once an area of interest had been defined, more molecular markers were designed to either narrow it further, or to directly test the chromosome at candidate mutations. Literature reading about the genes containing mutations in the area of interest dictated what mutations became candidate mutations. Once a candidate mutation was identified, the haplotype of each of the limiting recombinants was identified. For a candidate to pass the first test, each suppressing recombinant must still maintain the suppressor haplotype and each non-suppressing recombinant must have the haplotype of the mapping line. If this is not the case, then the area of interest is refined and another candidate is chosen.

Once a candidate mutation has passed the first test, then other experiments can be performed. These include creating knock out of that gene in a $spnB^{BU}$ background, recreating the candidate mutation via CRISPR-Cas9 mutagenesis in a $spnB^{BU}$ background, and rescuing the wild type phenotype by creating a transgenic fly with a wild type allele of the candidate gene into $spnB^{BU}$, *suppressor* background. Unfortunately the mapping of *CA231* has not yet progressed to this point.

Application of Mapping Techniques to the *CA231* Suppressor of *spn-B*^{BU}

Identifying Ideal Novel *spn-B*^{BU} Suppression Lines

The fly lines studied in this project had been maintained since 2004, when the original forward genetic screen was performed. There were originally thirty-nine novel mutant lines from the forward genetic screen, thirty-seven of which hadn't been characterized (Table 7). The decade of fly maintenance meant that each of these novel lines went through well over two hundred generations. With this many generations it was quite possible that random mutations arose in the fly lines. These mutations result in more variation from the reference genome, or their ability to suppress the *spn-B*^{BU} phenotype could have been altered. Having a heterogeneous population of flies can interfere with producing high-quality sequencing results and introduce ambiguity due to acquired polymorphisms.

Table 7: Novel Fly Lines. From the *spn-B*^{BU} Forward Genetic Screen at Princeton University in 2004, these 39 lines had been maintained as balanced stocks until 2013 or 2015 when they were sequenced. *CA231*, having one of the better suppression abilities, has been the focus of the majority of Dr. Ferguson's attention since 2009.

BG 88	*CA 1147*	*CA 121*	*CA 1215*	*CA 1301*	*CA 23*	*CA 231*	*CA 290*	*CA 911*	*CB 100*
CB 23	*CB 34*	*CB 429*	*CB 641*	*CB 649*	*CC 562*	*CC 687*	*CC 701*	*CD 100*	*CD 485*
CD542	*CD545*	*CD740*	*CD755*	*CD801*	*CE121*	*CE149*	*CE261*	*CR139*	*CR31*
CR324	*CR549*	*CR55*	*CR590*	*CR642*	*CR855*	*LB575*	*LC114*	*LC32*	

The suppresser lines being tested must still be homozygous for the *spnB*^{BU} mutation in order for the suppressive mutation to manifest its phenotype. False positives can arise if the BU mutation has been selected against in these populations. Therefore multiple homozygous flies were assessed at their *spindle B* locus by way of a PCR Restriction Fragment Length Polymorphism (PCR-RFLP) that the *BU* allele causes. The full-length amplicon is 454 bps and for a *spnB*^{BU} fly the amplicon is digested by *Hpy188III* producing two parts (189, 265 bps). All of the flies were originally made so that their unique 3rd chromosomes were balanced over *TM6C*, *Sb,* blocking recombination. Since all of the fly lines tested didn't have *Stubble*, they should have been homozygous for *spn-B*^{BU}. Some of the flies that were tested show heterogeneity at this

locus (Figure 8). To be tested accurately, the *spindle B* locus needs to be homozygous for the BU allele, so that the effect of the suppressive mutation can be assessed.

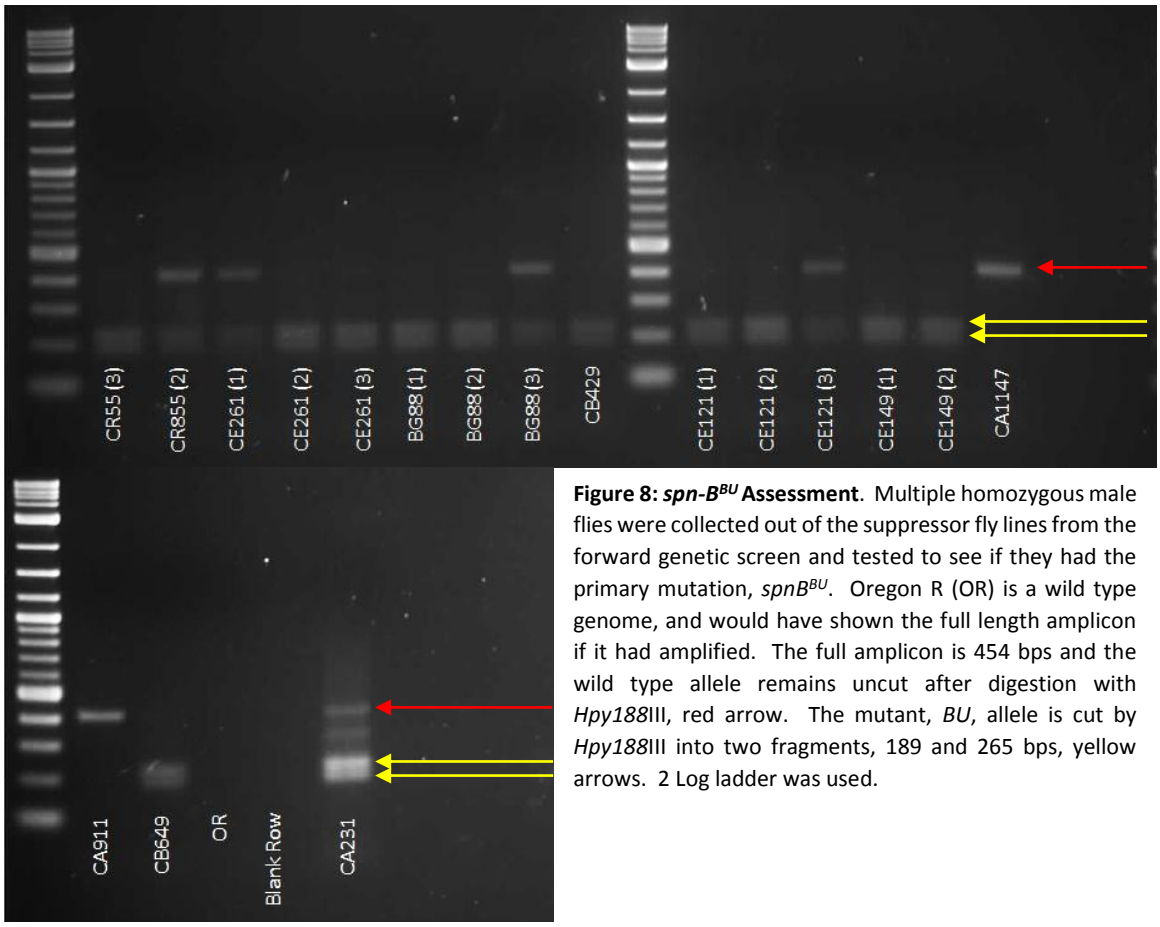


Figure 8: *spn-B*^{BU} Assessment. Multiple homozygous male flies were collected out of the suppressor fly lines from the forward genetic screen and tested to see if they had the primary mutation, *spnB*^{BU}. Oregon R (OR) is a wild type genome, and would have shown the full length amplicon if it had amplified. The full amplicon is 454 bps and the wild type allele remains uncut after digestion with *Hpy188*III, red arrow. The mutant, *BU*, allele is cut by *Hpy188*III into two fragments, 189 and 265 bps, yellow arrows. 2 Log ladder was used.

Since these suppressor lines were heterozygous at the *spindle B* locus, the first step was isogenizing all of the lines. The fly lines that had retained the BU mutation were also isogenized so that they too would have a working stock of a homogenous population. Each fly line underwent the following cross in triplicates to have better success in capturing the suppressive mutation (Figure 9). These triplicate isogenized lines were differentiated by an α, β, or γ succeeding their original names. A few of the suppressor lines were isogenized by previous students (differentiated by an A, B, or C succeeding the original fly line's name), to be thorough, these were also isogenized.

$$\male \frac{\text{Novel Chromosome}}{\text{TM3,Sb}} \times \female \frac{ruPrica}{\text{TM3,Sb}}$$

	ruPrica	TM3,Sb
Novel Chromosome	$\frac{ruPrica}{\text{Novel Chromosome}}$	$\frac{\text{Novel Chromosome}}{\text{TM3, Sb}}$
TM3,Sb	$\frac{ruPrica}{\text{TM3, Sb}}$	$\frac{\text{TM3, Sb}}{\text{TM3, Sb}}$

Figure 9: Isogenizing cross. Single heterozygous males were crossed with multiple virgin *ruPrica* flies (Top). Each male founded its own isogenized line. The Punett square for this cross shows that one third of the progeny are the genotype that is desired (highlighted in yellow), the other two living genotypes have the *Prickly* phenotype. The grayed out box shows a lethal combination of chromosomes.

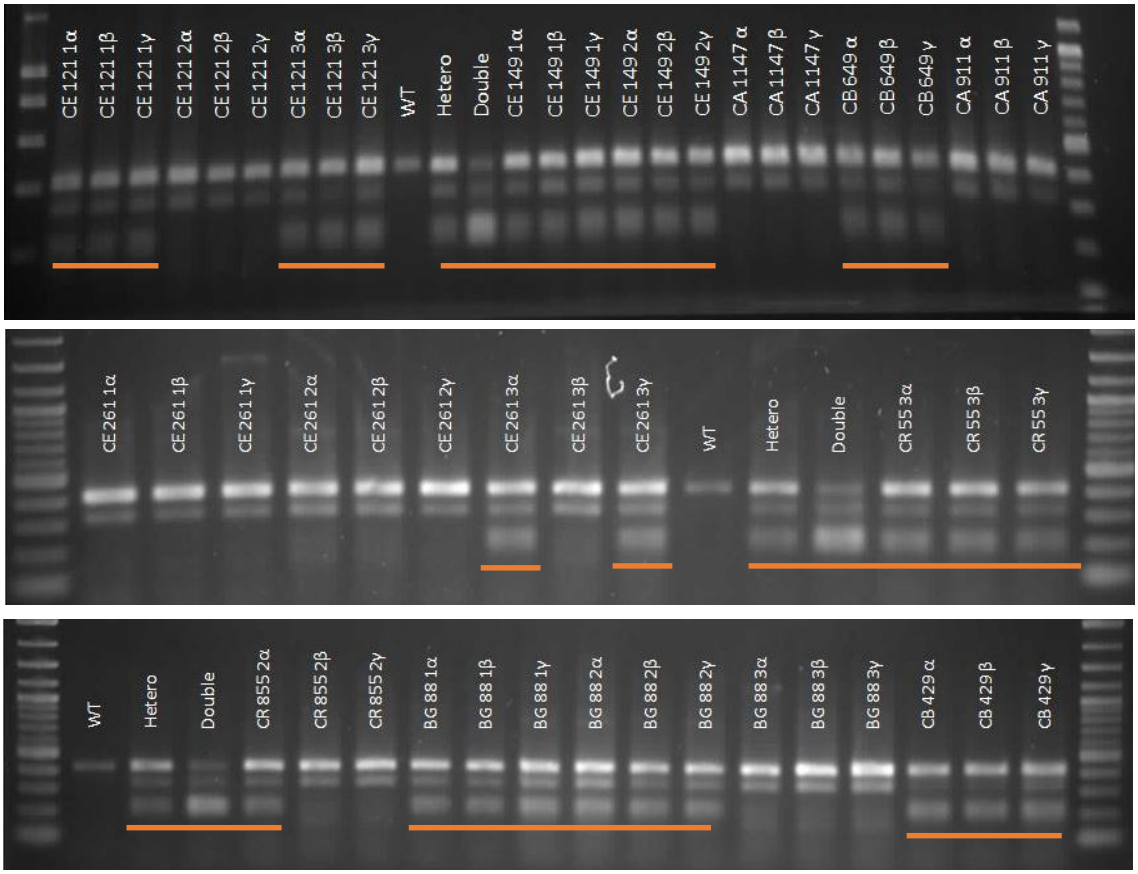
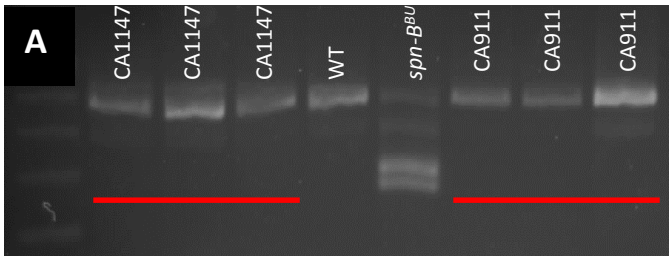
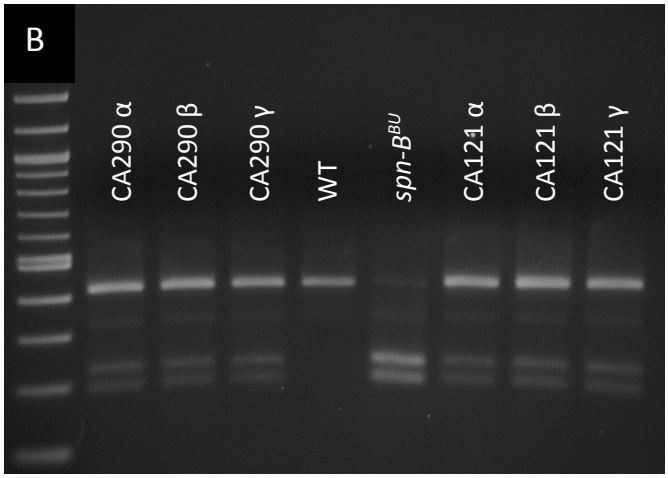


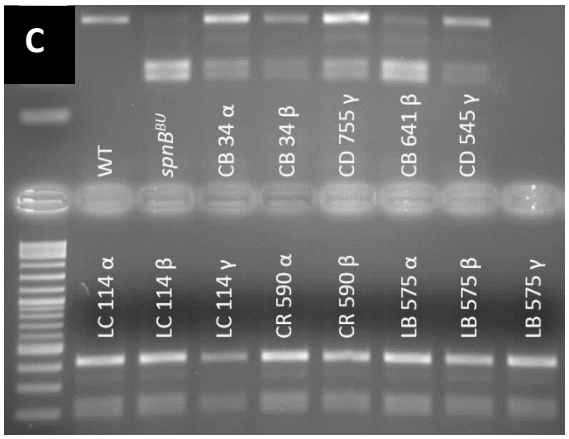
Figure 10: Primary Mutation Retest. (Top) Isogenized mutants were retested for having the primary mutation, *spnB*^{BU}. The founding, heterozygous, males of each of the isogenized lines were tested to see if they contained the BU allele. The mutation is visualized by enzymatic digestion with *Hpy188III* to go from a 454 bp amplicon to a 189 and 265 doublet. Doublets are underlined in orange. 2 Log ladder was used.

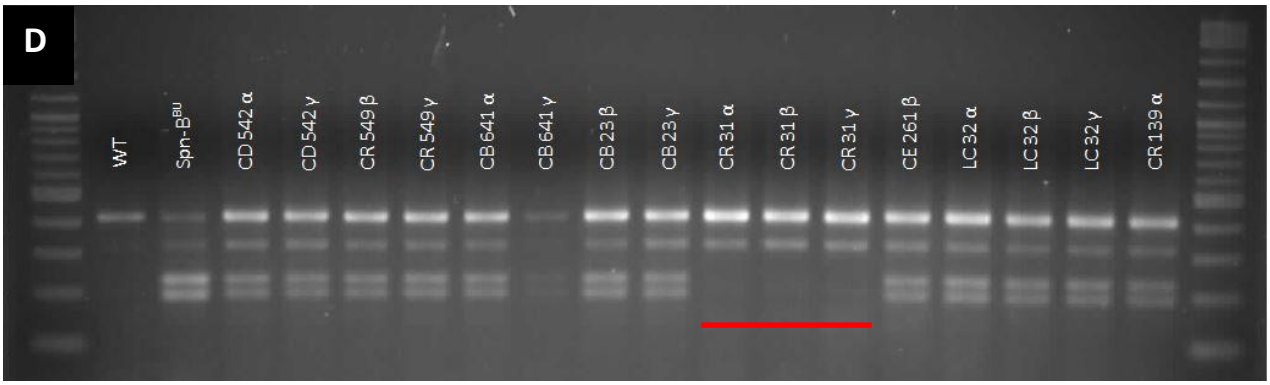
Figure 10 shows the results of the rescreening. To be more efficient, the flies tested were the founding heterozygous males of the balanced stock cross. Since his third chromosome was being captured in the crossing scheme, we tested the whole line, just by testing him. In the top gel in Figure R3, the smallest band seen on the left hand side is the 200bp marker, the 100bp marker can be seen on the right. Since all of the flies tested, except the controls, were heterozygous, we should have seen a doublet and a full-length amplicon. The sample called "Double" refers to a homozygous *spnB*^{BU} fly line. This sample should have only contained a doublet because it is homozygous; a full-length amplicon is seen though, which is the result of incomplete digestion. Looking at the other fly lines tested, some have the BU allele and some do not. Only those lines containing the *spn-B*^{BU} allele were kept.

Neither CA911 nor CA1147 had the BU mutation in any of the three isogenized lines. To better assess the full non-isogenized stock, a large genomic prep was done with eight male homozygotes from CA911 and nine from CA1147. From these large genomic preps, the *spindle B* locus was tested in triplicates. These results from this analysis are shown in Figure 11. Every fly used in these multi-fly genomic preps for CA911 and CA1147 does not contain the BU mutation in the *spindle B* gene (Figure 9A). As a result, these fly lines were discarded. Figure 11B, C, D, and E show additional BU screens, only the CR31 isolates didn't have the BU mutation. This line and its isogenized lines were also discarded. Shown in blue is the β isogenized *CA231* line that is the main focus of the paper.









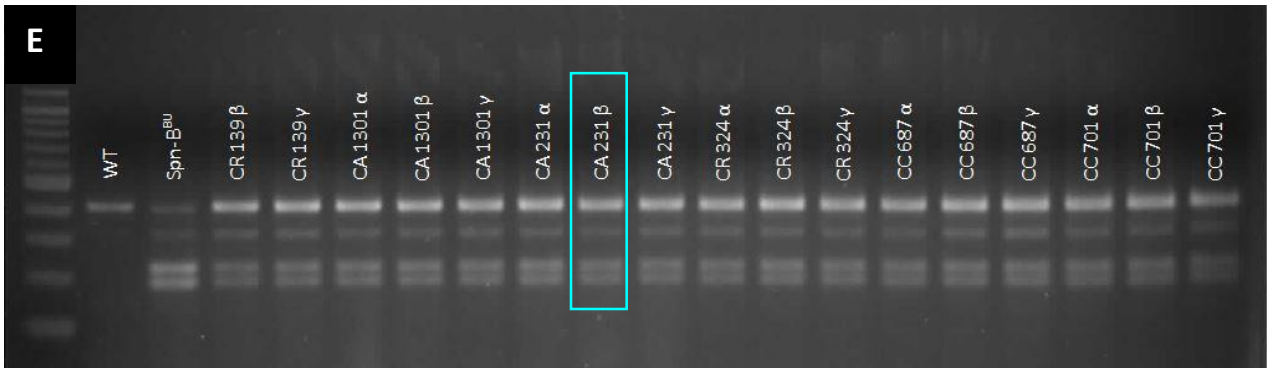


Figure 11: Additional BU retests. The *spindle B* locus was amplified in each fly line shown. Afterwards, the amplicon was digested with *Hpy188*III which cleaves the amplicon only when the BU mutation is present. Fly lines that do not contain the BU allele are underlined in red. **A)** A multi-fly genomic prep was done for each line. Neither CA1147 nor CA911 have the BU allele. **B)** All six isogenized lines have the BU mutation. **C)** All novel lines have the BU mutation. **D)** Out of all of the lines tested, only the three isogenized lines of CR31 didn't possess the BU mutation. **E)** All lines tested had the BU mutation. The CA231β isolate is highlighted in blue. 2 Log ladder was used.

After the isogenized lines had their BU mutation verified, the next step was to assess their ability to suppress the ventralized eggshell phenotype seen in *spn-B*^{BU} flies. Multiple condos were set up to test the suppression ability of the mutant lines. Table 8 and Figure 12 show the data from one condo Experiment. Table 8 gives the number of eggs counted with their Class breakdown. The table also gives the Suppression Index, a quantification of how ventralized the eggs are (defined in the method section). *spn-B*^{BU} flies should have the most ventralized eggs. Any fly line that has a similar Suppression index to the *spn-B*^{BU} line does not suppress. The larger the suppression index, the greater the fly line's ability to suppress is. Shown in blue is the *CA231*β line, which shows good suppression capabilities. Appendix B gives the Condo Experiment's data Tables and accompanying Figures for the other suppressor lines.

Table 8: Egg Shell Analysis. The isogenized variants of multiple fly lines were tested for their phenotypic suppression of ventralized eggs by means of a Condo experiment. The number of eggs and the class break down are shown. The Suppression index is the percent of class 1 and 2 eggs divided by the percent of class 3 and 4 eggs. One of the best suppressors, *CA231*β is shown highlighted in blue.

Genotype	Total Eggs	Class 1	Class 2	Class 3	Class 4	Suppression Index
spnB^{BU} Homozygotes	1669	2%	3%	4%	91%	0.05
CA 231 α	538	41%	43%	16%	1%	5.11
CA 231 β	812	56%	38%	6%	1%	15.24
CA 231 γ	985	32%	53%	14%	1%	5.48
LC 32 α	156	42%	42%	13%	3%	5.24
LC 32 β	18	39%	56%	0%	6%	17.00
LC 32 γ	11	45%	18%	36%	0%	1.75
CE 261 3β	264	16%	25%	12%	47%	0.69
CE 261 3γ	1294	22%	24%	11%	44%	0.83
CD 485 A	61	11%	30%	31%	28%	0.69
CD 485 B	14	7%	36%	29%	29%	0.75
CD 485 C	139	22%	43%	27%	8%	1.90
CR 549 α	45	56%	33%	7%	4%	8.00
CR 549 β	362	51%	42%	6%	1%	14.74
CR 549 γ	54	50%	46%	4%	0%	26.00
CA 290 β	187	16%	39%	14%	32%	1.17
CA 290 γ	137	25%	47%	12%	16%	2.61
CD 542 γ	359	44%	44%	7%	5%	7.35
CD 542 α	187	12%	49%	36%	3%	1.60

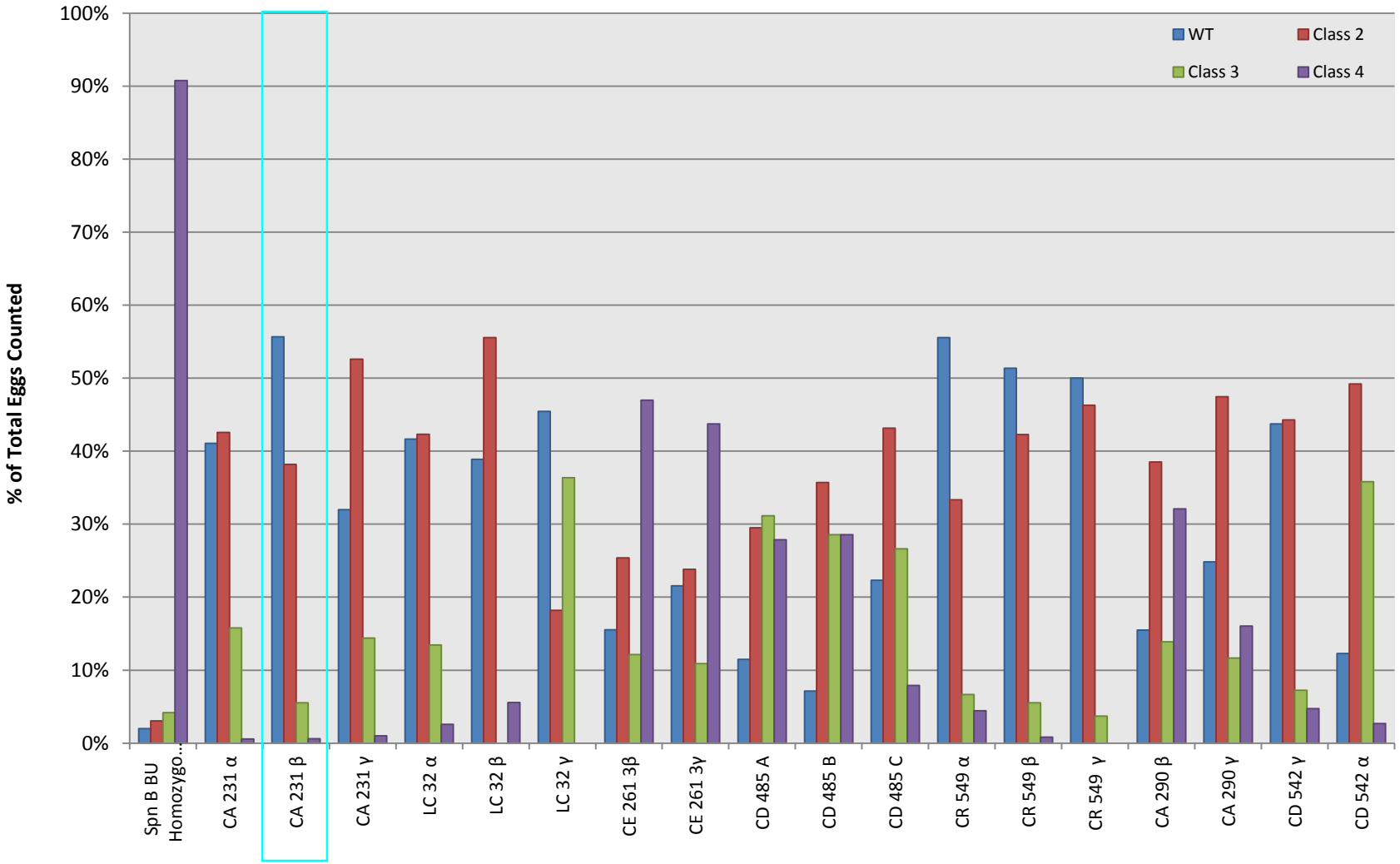


Figure 12: Visualization of Table R2. The percent of each phenotypic class from the isogenized suppressor lines are shown. The class break down for *spnB*^{BU} homozygous is what all of the other lines are being compared to. The more eggs in Class 1 or Class 2 indicates a suppression of the *spnB*^{BU} phenotype. One of the best suppressors, CA231β is shown outlined in blue.

Now working with thirty-six novel lines, there were two variables to consider in attempting to rank them: their suppression ability and the number of eggs laid. Reviewing Tables R2 and those in Appendix B, some of the lines had a dramatic reduction in the number of eggs laid. These lines were ranked lower because it was harder to discern suppression. This phenotype of reduced number of eggs, could have a connection to the way the lines are suppressing the *spnB*^{BU} phenotype, however this has not yet been determined. While these lines were not used for further experiments in this paper, they may provide information about a novel mode of suppression and should be investigated at a later time. The final group ranking of the novel lines was done with the top-ranking group having the greater suppression index and egg laying ability (Table 9). The ten genotypes that are shown in bright blue were the ones with the best suppression ability. These lines were then selected and prepared for Next Generation Sequencing (NGS). The green-highlighted lines were also good suppressors and were sent for NGS in a second batch. The three lines in black either never had the BU mutation, or it was crossed out of the lines through negative selection over many generations. The four red genotypes were also removed, these for their poor suppression ability. Dr. Scott Ferguson mapped *CA1215* to the *mei-P22* locus and *CR642* to the *lnk* locus (Ferguson *et al.* 2012).

Genotype	Suppression Index	Egg Count	Status
CA 23 - C	0.602	660	Sequenced, Initial Batch
CA 121 - γ	1.310	1014	Sequenced, Initial Batch
CA 231 - β	15.240	812	Sequenced, Initial Batch
CA 1301 - α	1.252	635	Sequenced, Initial Batch
CB 429 - α	8.833	118	Sequenced, Initial Batch
CD 485 - C	1.896	139	Sequenced, Initial Batch
CE121 – 1β	1.516	161	Sequenced, Initial Batch
CE149 – 1γ	5.207	360	Sequenced, Initial Batch
CR324 - β	22.068	1015	Sequenced, Initial Batch
CR549 - β	14.739	362	Sequenced, Initial Batch
CA 290 - γ	2.605	137	Sequenced, Second Batch
CB 34 - α	2.565	82	Sequenced, Second Batch
CD 100 - β	1.085	711	Sequenced, Second Batch
CD542 - γ	7.349	359	Sequenced, Second Batch
LB575 - α	8.910	357	Sequenced, Second Batch
LC32 - α	5.240	156	Sequenced, Second Batch
LC114 - α	4.236	555	Sequenced, Second Batch
CR590	8.377	497	Further Data Needed
CB 23 - γ	0.932	228	Further Data Needed
CC 701 - β	2.640	659	Further Data Needed
CD801 - A	6.000	7	Further Data Needed
CA 1215	-	-	Mapped
CR642	-	-	Mapped
BG 88	-	-	Potentially Lethal
CB 100	-	-	Potentially Lethal
CB 649	-	-	Potentially Lethal
CC 562	-	-	Potentially Lethal
CC 687	-	-	Potentially Lethal
CD545	-	-	Potentially Lethal
CD740	-	-	Potentially Lethal
CD755	-	-	Potentially Lethal
CR139	-	-	Potentially Lethal
CB 641	-	-	Removed
CE261	-	-	Removed
CR55	-	-	Removed
CR855	-	-	Removed
CA 911	-	-	Removed

Table 9: Statuses of Novel Suppressor Lines After Preliminary Tests. All thirty nine suppressors underwent the original *BU* screen, and attempts at isogenization. The fly lines, shown in blue, show clear suppression. gDNA from each line was sent for NGS in either the first or second batch. The green lines have mild suppression abilities and need further analysis. The yellow lines had previously been mapped. The purple lines were unable to be isogenized, any attempts resulted in the capture of a recessive lethal mutation on the third chromosome. The black lines were removed either because they no longer could suppress or they did not contain the *BU* allele.

Next Generation Sequencing

The ten novel fly lines chosen for sequencing (Table 10) and the reference line *spnB*^{BU} had their genomic DNA prepped for NGS as described in the methods section. Table 10 shows the breakdown of how many flies were used and what final mass was achieved. The flies collected were all homozygous so that the balancer chromosome wasn't sequenced as well. DNA was isolated using the QIAGEN DNeasy Blood & Tissue Kit. After the gDNA was isolated, its concentration was determined by Qubit Fluorometry. These elutions were sent to Roswell Park Cancer Institute for Next Generation Sequencing.

Table 10: NGS Preparation. The top suppressors and the reference line, each had their gDNA isolated for Nex Gen Sequencing. The number of flies used and the concentrations obtained from the first two elutions were recorded. Concentration was obtained by Qubit Flourometry.

Genotype	Number of Flies	gDNA Concentration 1st Elution	gDNA Concentration 2nd Elution
spnB^{BU}	30	22.9 µL/mL	5.60 µL/mL
CE149 1γ	32	17.9 µL/mL	4.88 µL/mL
CE121 - γ	33	23.9 µL/mL	5.15 µL/mL
CA23 - C	34	19.4 µL/mL	5.15 µL/mL
CB429 - α	28	13.2 µL/mL	3.80 µL/mL
CA231 - β	14	15.6 µL/mL	2.98 µL/mL
CR324 - β	9	9.93 µL/mL	2.49 µL/mL
CR549 - β	8	7.13 µL/mL	1.79 µL/mL
CD485 - C	15	16.0 µL/mL	3.07 µL/mL
CA121 - γ	9	10.9 µL/mL	3.03 µL/mL
CA1301 - α	12	20.1 µL/mL	3.29 µL/mL

The *spn-B*^{BU} line was vital to the NGS results, the genetic sequence of this line would be compared to the suppressors. The changes that make each suppressor line unique from *spnB*^{BU} were what we were after. Therefore we had to determine the genotype of *spnB*^{BU} with the utmost precision. This line was sequenced to an average coverage of 136.5x on the right arm and 139.2x on the left arm of the third chromosome. This meant that the nucleotide sequence of this line was assessed over one hundred and thirty times. This allowed for a very accurate read with low ambiguity. Each of the 10 suppressors was sequenced to an average coverage of 34.5x.

The number of differences between each of the suppressor lines to the reference, *spn-B*^{BU}, were roughly the same, except for the *CE121* line (Table 11). This line had a tremendous number of changes to it. All of the mutations in *CA231*, *CA121*, and *CR549* were characterized and *CE149* has some initial characterization. The *CA231* line was chosen to be investigated for multiple reasons. It has a great suppression ability, the quantity of eggs laid is not diminished, the sequencing results had a manageable amount of differences, and previous work had already been done to map the causative mutation. Malachi Blundon had previous mapped the mutation to the right arm of the chromosome, possibly to a gene called *pyk.* The sequencing results showed no mutation within a few million nucleotides of *pyk,* thereby refuting this hypothesis. The complete list of differences between *CA231*β and *spnB*^{BU} is shown in Appendix C. All of the mutations were characterized individually, as described in the methods.

The plethora of differences, 379, between *CA231* and *spnB*^{BU} included several strong candidates but there are still a lot of mutations spread across both arms of the third chromosome. To begin to narrow down the list of candidate mutations, another mapping experiment was done, but this one guided by the Next Generation Sequencing results.

Table 11: NGS Total Differences. The total number of differences between each suppressor line and *spnB*^{BU} are shown below.

Suppressor Line	Number of Mutations
CA231	379
CA121	409
CR549	336
CA1301	474
CA23	360
CB429	437
CD485	427
CR324	353
CE149	375
CE121	>10,000

Multiple mapping lines were available to use in creating these recombinants. The most vital aspect of these mapping lines was to have clear differences between the suppressor line and the mapping line. To test potential mapping lines, universal SNP markers were generated from the differences that arose in all ten of the suppressor lines. These universal molecular markers, primers, chromosomal locations and distinguishing characteristics are shown in Appendix A. The 2014 Molecular Genetics Class aided in creating some of the molecular markers. The markers in Appendix A that have the alphabetical names (SNPA, SNPB, etc) were developed from this class. Even though the markers were made to identify the differences between the reference line and the suppressor line, there was no guarantee that the mapping line would provide unambiguous results to distinguish itself from the suppressor lines. Each of the mapping lines tested shared multiple loci with *CA231* and therefore were not used. Instead of a regular mapping line, the wild type line, *Oregon R*, was used to make recombinants. With the help of Breanna Myers, over 50 recombinants each were made for *CR549, LB575, CA121, LC114, CA231* and *CR324*.

Characterizing the Haplotype of *CA231* Recombinant Chromosomes

First, all of the recombinants were screened for *spn-B*^{BU}. Out of the fifty-two recombinant stocks made from *CA231*, only nineteen of them retained the *spn-B*^{BU} allele (Table 12). The other recombinants were discarded. A condo experiment was performed with the remaining recombinants. It was expected that the recombinants would have differing suppression capabilities based on which ones obtained the suppressive mutation. The results were indeed mixed, but the suppression abilities of most of the recombinants were not easily discernable as suppressing the BU phenotype or not (Table 13 & Figure 13). Rec 34 and Rec 36 clearly suppressed the BU phenotype. Rec 38 clearly did not suppress the BU phenotype. Additional condo experiments were done to obtain less ambiguous results, but those results were still ambiguous for the remaining recombinants (Data not shown). Since *CA231* had previously been shown to dominantly suppress the BU phenotype, a condo experiment was done with the recombinants, testing them heterozygously but with a *spn-B*^{BU} homolog. This would make their causative mutation heterozygous while leaving the *spindle B* locus homozygous for the BU allele. These results were also ambiguous and didn't distinguish any more of the recombinants (Data not shown). A previous mapping experiment on *CA231* produced a strongly suppressing recombinant, MBJD, and this recombinant was investigated alongside the newly made ones. This recombinant was made by recombining with the *17620* (Ferguson *et al.* 2012) mapping line. This previous mapping experiment mapped the causative mutation to the end of the right arm between ebony and claret. Further analysis seemed to point to the *pyk* gene (Blundon 2011). This area became the location of interest for this study.

Multiple suppressor lines, mainly *CA231*, *CR324* and *CA1301*, share many differences with the *spn-B*^{BU} line. One possibility was that the causative mutation is shared between each of the suppressor lines. Therefore a complementation cross was set up to assess whether or not these suppressor lines were in the same complementation group or not (Table 14 & Figure 14). Among the homozygous and the

trans-heterozygous flies tested, it was quite clear that the eggs laid by the trans-heterozygotes do not suppress the BU phenotype as well as when their respective homozygotes. Therefore it was concluded that each is in its own complementation group and that none of the shared mutations between these three fly lines are the causative mutation.

Table 12: Recombinant Lines of *CA231*. These nineteen lines retained the *spnB*^{BU} allele. Their haplotypes were further characterized at the various developed molecular markers.

Rec 03	Rec 08	Rec 11	Rec 12	Rec 16	Rec 18	Rec 20
Rec 25	Rec 30	Rec 34	Rec 36	Rec 38	Rec 39	Rec 40
Rec 41	Rec 42	Rec 46	Rec 48	Rec 49		

Table 13: Recombinant Suppression Analysis. The recombinants were put on apple juice plates with a yeast diet. The plates were changed twice a day for 7 days. Eggs from the last four days of plates were counted and sorted into the four phenotypic classes.

Genotype	Total Eggs	Class 1	Class 2	Class 3	Class 4	Suppression Index
spn-B^{BU}	667	18%	26%	28%	28%	0.77
CA231	338	39%	46%	14%	1%	5.50
MBJD	242	33%	63%	3%	1%	25.89
Rec 3	606	9%	46%	43%	1%	1.24
Rec 11	415	14%	60%	23%	4%	2.77
Rec 30	335	26%	47%	25%	2%	2.72
Rec 34	583	61%	38%	1%	1%	57.30
Rec 36	742	66%	20%	12%	2%	6.49
Rec 38	281	8%	24%	60%	9%	0.46
Rec 39	223	7%	45%	44%	4%	1.08
Rec 40	245	21%	41%	30%	8%	1.63

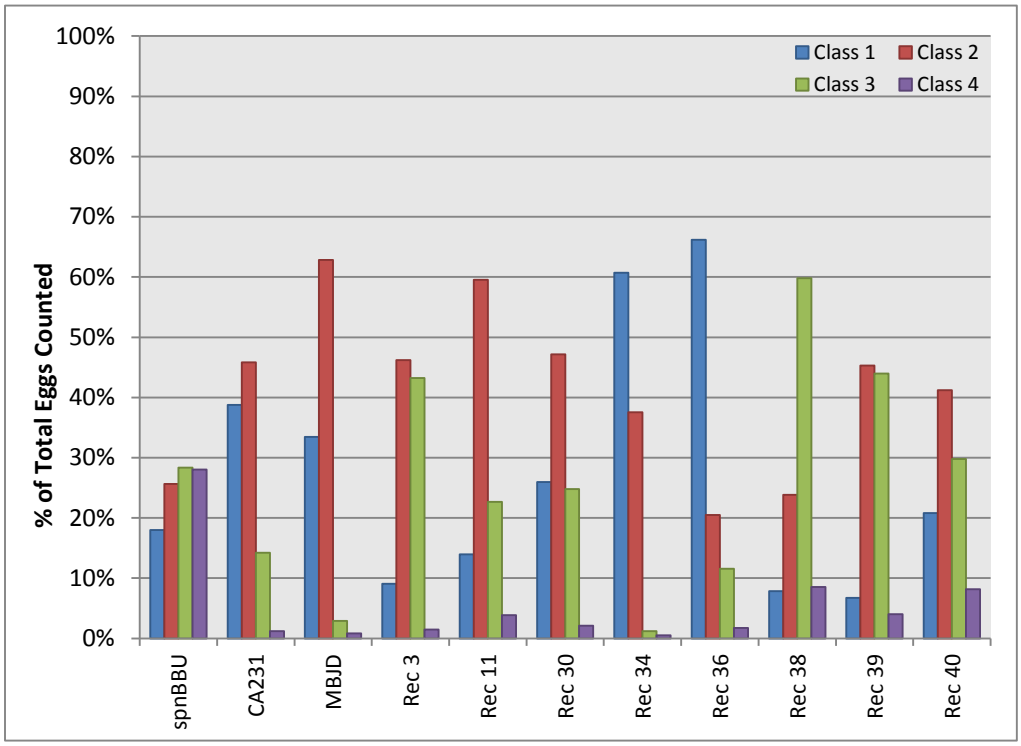


Figure 13: Recombinant Suppression Analysis. The homozygous recombinants were put on apple juice plates with a yeast diet. The plates were changed twice a day for 7 days. Eggs from the last four days of plates were counted and sorted into the four phenotypic classes.

Table 14: Complementation Test of Suppressors. Multiple mutations are shared between *CA231*, CR324 and CA1301. Condo data is shown from the homozygous lines and flies that are trans-heterozygous for two different suppressor chromosomes. None of the three suppressor lines are in the same complementation group.

Genotype	Total Eggs	Class 1	Class 2	Class 3	Class 4	Suppression Index
spnB^{BU}	456	5%	4%	17%	75%	0.09
CA1301 α	291	37%	18%	36%	9%	1.22
CR324 β	255	89%	5%	5%	1%	16.00
CA231 β	374	44%	20%	30%	6%	1.73
CA1301 α x CR324 β	1244	31%	18%	32%	20%	0.93
CA231 β x CA1301 α	1097	21%	18%	50%	11%	0.62
CR324 β x *CA231* β	1441	22%	21%	50%	7%	0.74

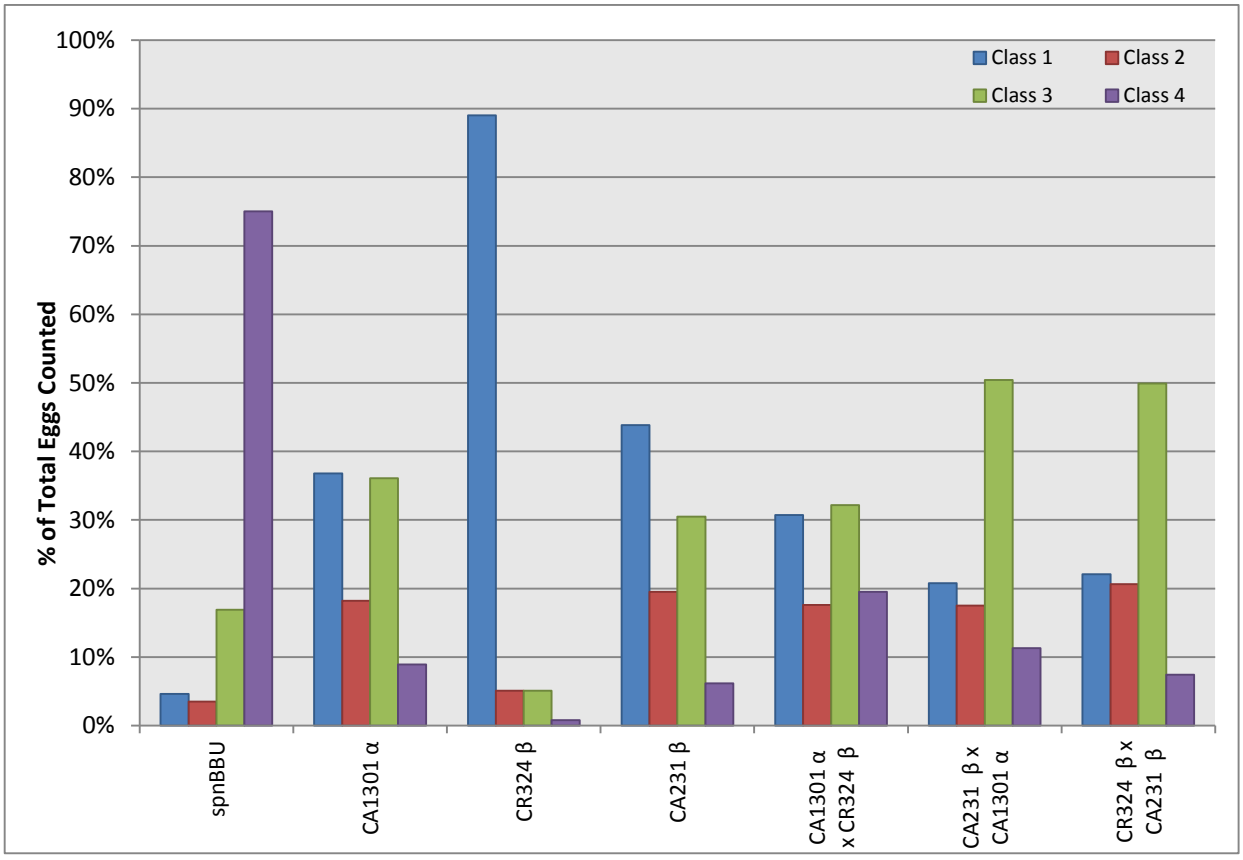


Figure 14: Complementation Test of Suppressors. Multiple mutations are shared between *CA231*, CR324 and CA1301. Condo data is shown from the homozygous lines and flies that are trans-heterozygous for two different suppressor chromosomes. None of the three suppressor lines are in the same complementation group.

Since the complementation test indicated that the causative mutation is not among the shared mutations, the list of possible causative mutations decreased from 379 to 214. These mutations are still spread across both sides of the chromosome. Individual haplotype maps were constructed by analyzing multiple polymorphic sites. (Figure 15). The Universal RFLPs and InDels were tested first (Figure 15 purple) then the *CA231* specific RFLPs (pink, Figure R8). Haplotype data from multiple loci could not be used because *Oregon R* had the same haplotype as *CA231*. SNP 1, SNP 7, and SNP 8 are three loci where this was the case.

Analyzing the loci that did produce data, starting at the beginning of the Left arm, SNP C is the first locus (Figure 16). SNPC has variation among the recombinants with some exhibiting the original *CA231* haplotype and others having that of *Oregon R*. While the SNP D loci did indicated that the *Oregon R* and *CA231* haplotypes were different, the recombinants' haplotypes weren't pursued due to data that implicated a different area of the chromosome. There was not much variation at the SNP E locus, Figure 17. Recombinants 3, 36, and 39 were the only ones to have recombination occur in this area. The SNP 6 locus has a similar variation with the recombinants 3, 36, and 39, and MBJD all having *Oregon R*'s haplotype (Figure 18). At the SNP F locus only Recombinants 3 and 39 had *Oregon R*'s haplotype (Figure 19). Since the SNPB locus is so close to the *spindle B* locus, it is of no surprise that all of the recombinants have the suppressor haplotype (Figure 20). Likewise, SNP2 is just to the right of the *spindle B* locus and again, all of the recombinants have the *CA231* haplotype (Figure 21). Further away from the *spindle B* locus, SNP3 has some variation with recombinants 36, 38, and 40 showing the *Oregon R* haplotype (Figure 22). The *tsl, btz,* and SNP 4 loci were also investigating showing variation among the haplotypes of Rec. 34, 36, 38 and MBJD (Figures 23-25).

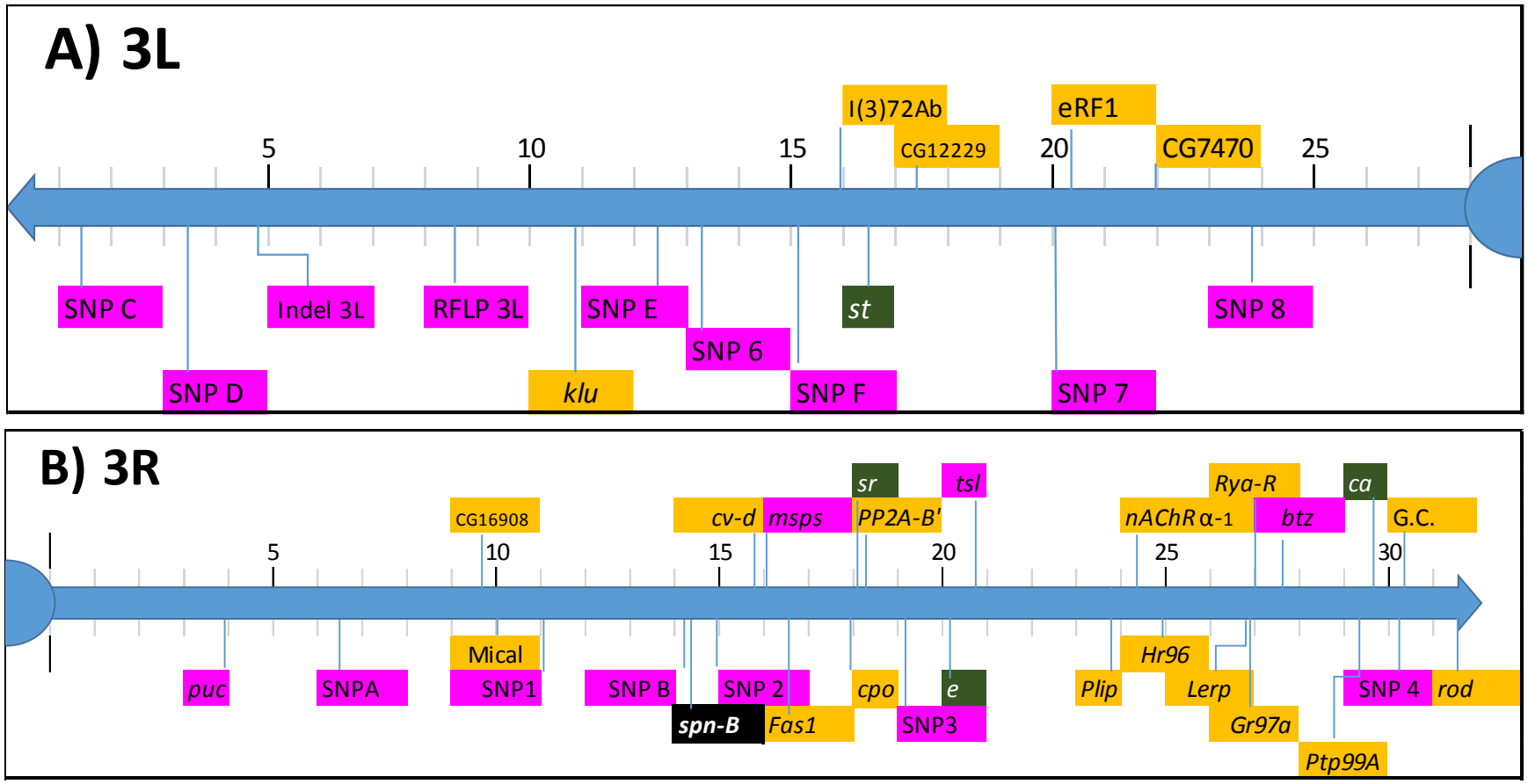


Figure 15: Chromosome Map with Tested Loci. The 3rd chromosome is split into the left arm (**A**) and Right arm (**B**). Universal RFLPs and InDels (purple), *CA231* specific RFLPs and InDels (pink), sequenced loci (orange), phenotypic loci (green), and the *spindle B* loci (black) are shown to scale. The locations are written in millions of nucleotides.

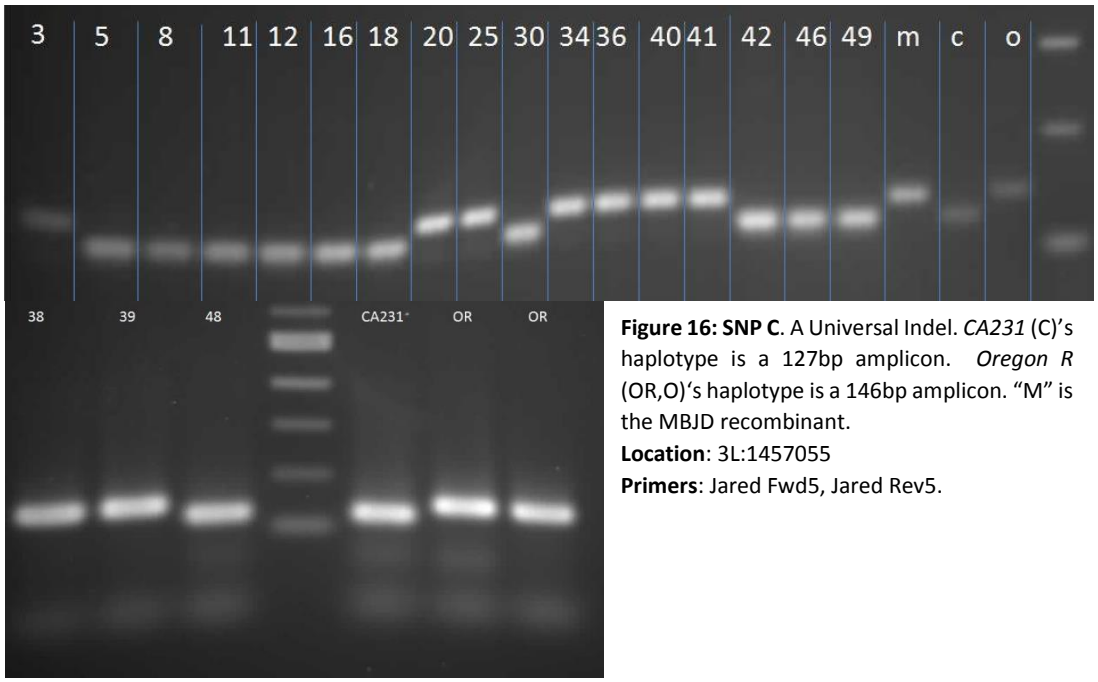


Figure 16: SNP C. A Universal Indel. *CA231* (C)'s haplotype is a 127bp amplicon. *Oregon R* (OR,O)'s haplotype is a 146bp amplicon. "M" is the MBJD recombinant.
Location: 3L:1457055
Primers: Jared Fwd5, Jared Rev5.

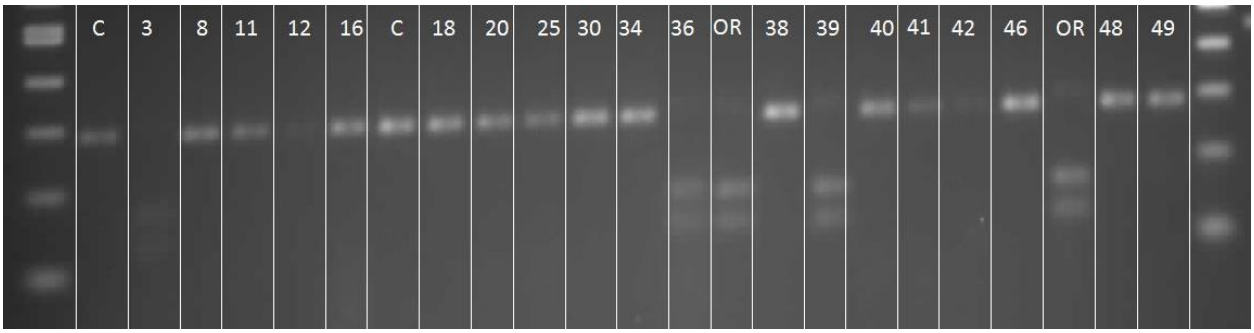


Figure 17: SNP E. A Universal RFLP. *CA231* (C)'s haplotype is a 293bp amplicon. *Oregon R* (OR)'s haplotype is an amplicon cut by HpaII (169, 124bp).
Location: 3L: 12,456,662
Primers: Anoud Fwd3, Anoud Rev3.

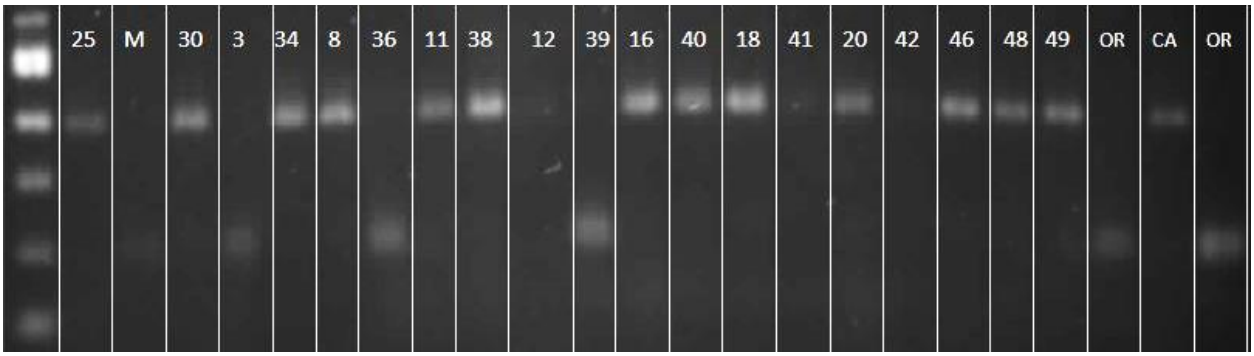


Figure 18: SNP 6. A Universal RFLP. *CA231* (CA)'s haplotype is a 388bp amplicon. *Oregon R* (OR)'s haplotype is an amplicon cut by AcuI (202, 186bp). M is the MBJD recombinant.
Location: 3L: 13,243,020.
Primers: JH11, JH12.

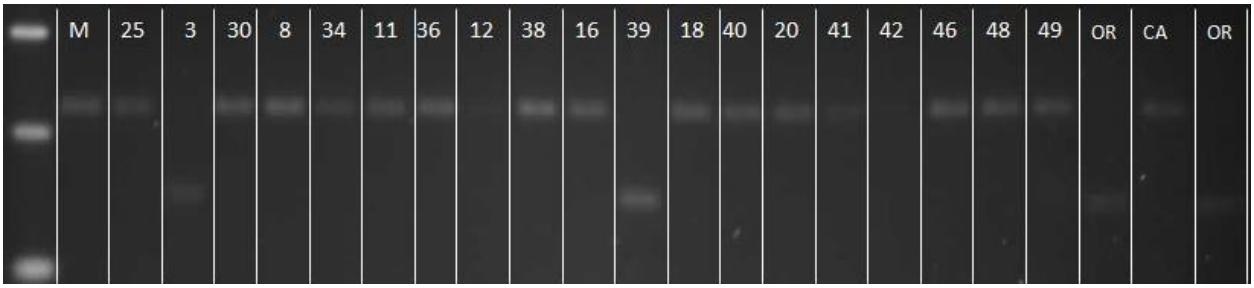


Figure 19: SNP F. A Universal RFLP. *CA231* (CA)'s haplotype is a 217bp amplicon. *Oregon R* (OR)'s haplotype is an amplicon cut by HinP1I (147, 70 bp). M is the MBJD recombinant.
Location: 3L:15,166,730.
Primers: Alissa Fwd4, AlissaRev4.

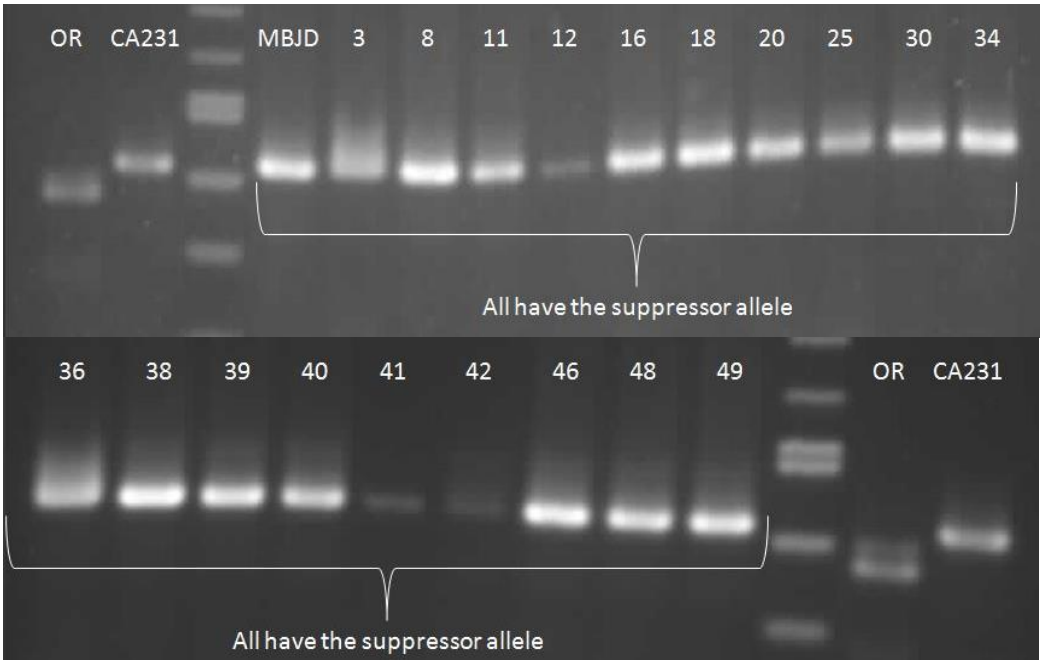


Figure 20: SNP B. A Universal Indel. *CA231*'s haplotype is a 414bp amplicon. *Oregon R* (OR)'s haplotype is a 368bp amplicon.
Location: 3R:14,100,888.
Primers: Kyle Fwd3, Kyle Rev3.

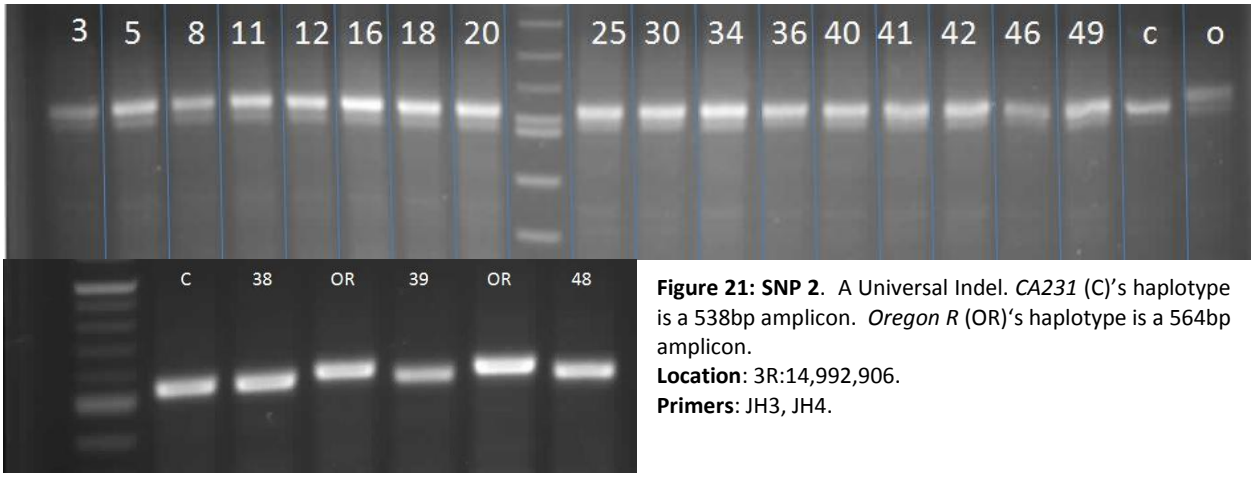


Figure 21: SNP 2. A Universal Indel. *CA231* (C)'s haplotype is a 538bp amplicon. *Oregon R* (OR)'s haplotype is a 564bp amplicon.
Location: 3R:14,992,906.
Primers: JH3, JH4.

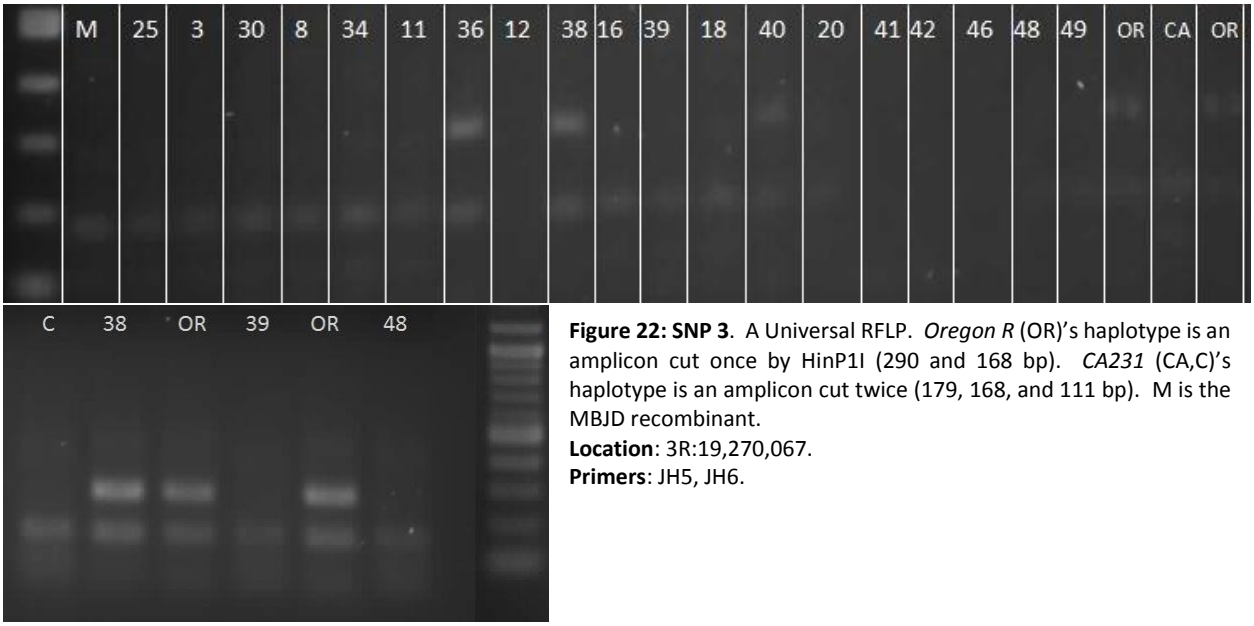


Figure 22: SNP 3. A Universal RFLP. *Oregon R* (OR)'s haplotype is an amplicon cut once by HinP1I (290 and 168 bp). *CA231* (CA,C)'s haplotype is an amplicon cut twice (179, 168, and 111 bp). M is the MBJD recombinant.
Location: 3R:19,270,067.
Primers: JH5, JH6.

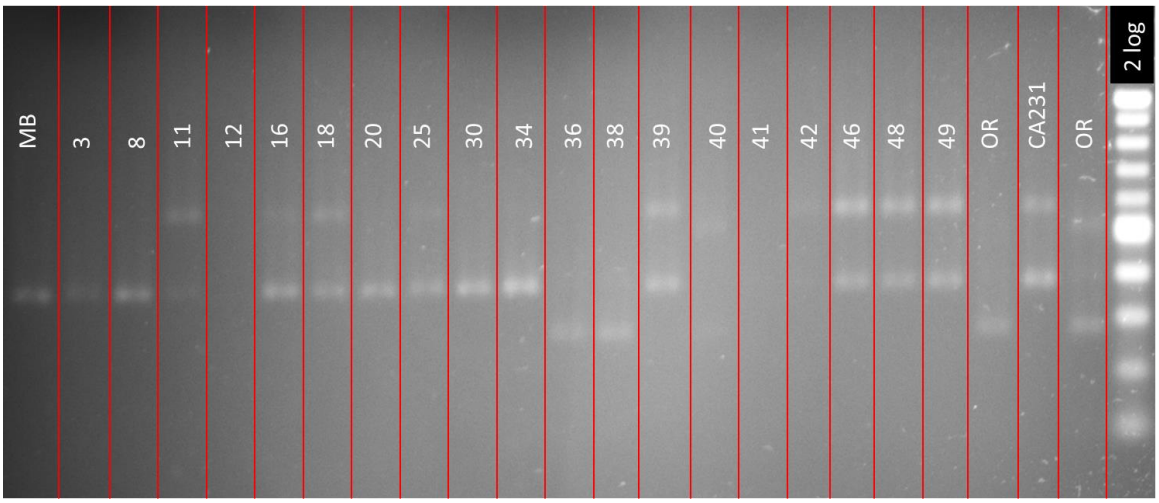


Figure 23: *tsl*. A *CA231* specific RFLP. *CA231*'s haplotype is a 372bp amplicon. *Oregon R* (OR)'s haplotype is an amplicon cut by *SfaNI* (273, 93bp). MB is the MBJD recombinant.
Location: 3R:21,778,495.
Primers: SF631, SF632

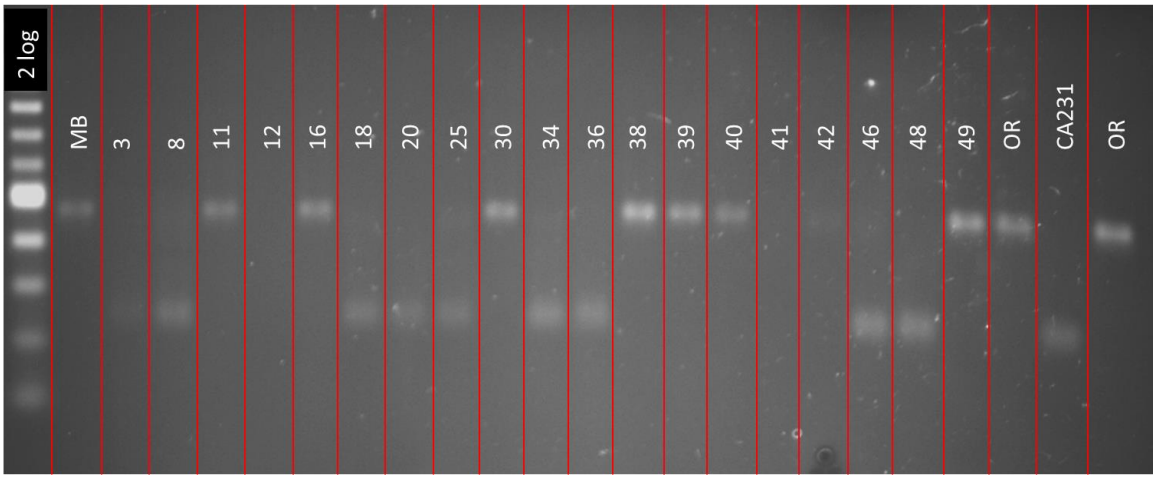
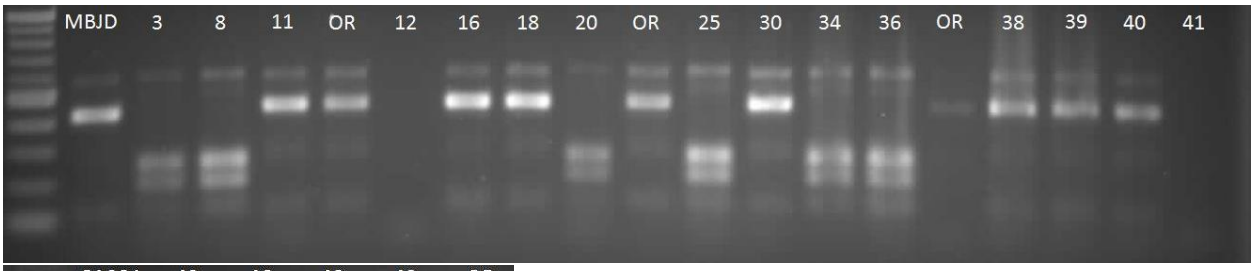


Figure 24: *btz.* A *CA231* specific RFLP. *Oregon R* (OR)'s haplotype is a 475bp amplicon. *CA231*'s haplotype is an amplicon cut by MluCI (248 and 229bp). MB is the MBJD recombinant.
Location: 3R:27,701,050.
Primers: SF627, SF628.



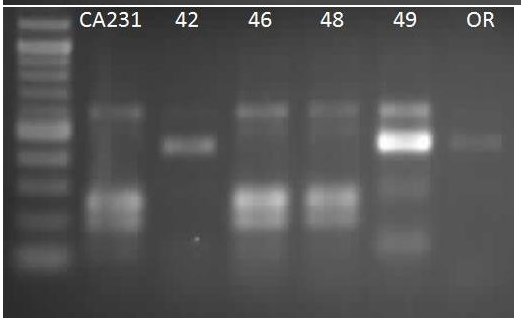


Figure 25: SNP 4. A Universal RFLP. *Oregon R* (OR)'s haplotype is a full length amplicon (442bps). *CA231*'s haplotype is an amplicon cut by BsrI (251 and 192 bp).
Location: 3R:30,348,884.
Primers: JH7, JH8.

As the testing progressed, it became clear that Recombinants 34, 36, and 38 generated the cleanest data so they, along with MBJD, were assessed at additional loci. These additional loci were mostly protein changes between *CA231* and *spn-B*^{BU}. Recall that the recombinants from the experiment were crossed with *Oregon R*, and the MBJD line was made by recombination with the mapping line *17620*. Since the haplotypes of the *17620* line is unknown at these loci, a map of the MBJD haplotypes couldn't be constructed with certainty. Regardless of the haplotype of the mapping chromosomes, the MBJD line is able to suppress the BU phenotype. The MBJD line was used as a positive control line. When investigating each candidate locus, the suppressing recombinants (34, 36 and MBJD) all had to have the mutation while Recombinant 38 must not. Loci that did not follow this pattern could not be the causative mutation.

Twenty addition regions were tested for Recombinants 34, 36, 38 and MBJD (orange, Figure 15). The haplotypes for these recombinants are displayed in Table 15. Recombinant 38 has extensive *Oregon R* haplotypes at the end of the right arm. Recombinant 34 on the other hand, contains the *CA231* haplotype for almost all of its characterized loci.

Along with these additional loci, phenotypic markers were assessed to give further haplotypes. These phenotypic markers are shown in Table 16 in green. The *ebony* marker was the only one that was identifiable in all four recombinants. MBJD's haplotype for *scarlet*, and Recombinant 38's haplotype for *stripe* were also not discernable.

Table 15: **Haplotypes Characterized by Sequencing**. These loci were characterized by amplifying the area and sending the product to GeneWiz for sequencing. Recombinants that share the *Oregon R* (OR) haplotype are shown in red, those that share the *CA231* haplotype in green, and any region that couldn't be identified are shown in yellow.

Name		Location	MBJD	34	36	38
klu	*klumpfuss*	3L:10,984,445	OR	?	OR	*CA231*
l(3)72Ab	*lethal (3) 72Ab*	3L:15,966,999	*CA231*	*CA231*	*CA231*	*CA231*
CG12229		3L:17,435,941	?	*CA231*	*CA231*	*CA231*
eRF1	*eukaryotic release factor 1*	3L:20,350,702	?	?	?	*CA231*
CG7470		3L:21,982,217	*CA231*	*CA231*	*CA231*	*CA231*
CG16908		3R:9,753,267	*CA231*	*CA231*	*CA231*	*CA231*
Mical		3R:10,015,100	*CA231*	*CA231*	*CA231*	*CA231*
cv-d	*crossveinless d*	3R:15,844,615	*CA231*	*CA231*	*CA231*	*CA231*
Fas1	*Fasciclin 1*	3R:16,638,094	*CA231*	*CA231*	*CA231*	*CA231*
cpo	*couch potato*	3R:17,997,601	?	?	?	?
PP2A-B'		3R:18,172,727	*CA231*	*CA231*	*CA231*	*CA231*
Plip	*PTEN-like phosphatase*	3R:23,758,772	OR	?	OR	?
nAChRα1	*nicotinic Acetylcholine Receptor α1*	3R:24,405,397	OR	*CA231*	OR	OR
Hr96	*Hormone receptor-like in 96*	3R:25,026,721	OR	*CA231*	OR	OR
Lerp	*Lysosomal enzyme receptor protein*	3R:26,861,311	OR	*CA231*	*CA231*	OR
Gr97a	*Gustatory receptor 97a*	3R:26,965,308	OR	*CA231*	*CA231*	OR
RYa-R	*RYamide receptor*	3R:27,002,187	OR	*CA231*	*CA231*	OR
Ptp99A	*Protein tyrosine phosphatase 99A*	3R:29,401,231	OR	*CA231*	*CA231*	OR
Gene conversion track in CG2217 CG42740		3R:30,419,147	OR*	*CA231*	*CA231*	OR
rod	*rough deal*	3R:31,581,796	OR	*CA231*	?	OR

Table 16: Haplotypes Characterized by Phenotypic Markers. On the original spn-BBU line there were three recessive mutations that caused phenotypic changes to the adult fly. These were *scarlet, stripe* and *ebony.* A visual test for each of these was performed. Some recombinants couldn't be determined visually (yellow). The others either showed the *CA231* haplotype (green) or didn't (*Oregon R*'s haplotype, red).

Name	Location	MBJD	34	36	38
scarlet (st)	3L: 16,497,651	?	*CA231*	*CA231*	*CA231*
stripe (sr)	3R: 18,090,803	*CA231*	*CA231*	*CA231*	?
ebony	3R:21,229,839	*CA231*	*CA231*	OR	OR

Constructing the Haplotype Chromosomes

Figure 15 shows all of the loci investigated in this study; this figure was used as a template. Individually, for each of the four recombinants, their haplotype at each locus was determined. By visualizing their haplotypes across the third chromosome, we were able to get a sense of the areas that have been recombined away (Figures 26 & 27). Since the MBJD line was made from recombination with a different line, the haplotype map cannot be fully trusted, especially at the Universal SNP areas where the change might be from after the Forward Genetic Screen. Looking at MBJD's 3L haplotype map, Figure 26A, from the RFLP 3L loci to SNP F, it goes from OR's haplotype, to *CA231*'s, back to OR's and finally back to *CA231*'s. If MBJD was made by recombining with OR, then this would mean that three recombination loops were established in a seven million base pair range. While this is theoretically possible, it is not very likely. A more plausible explanation would be that at the SNPE locus, the *17620* mapping line already had the same haplotype as *CA231* and so even though this area is from the mapping line, it still shares the same haplotype. Besides this region, the haplotype of 3L from the MBJD line has a stretch of *CA231* haplotypes from SNPF to CG7470. The 3L haplotype map of recombinant 34, (Figure 26B), shows a large section being all from *CA231*, with flanking regions that have yet to be characterized. Since the SNPC region, at the very tip of the chromosome, is of the *OR* haplotype, recombination had to have happened somewhere in the 1.5 million - 12.5 million base range. The 3L haplotype map of recombinant 36, (Figure 26C), looks like it may be split in half between SNP6 and SNPF. The 3L haplotype map of recombinant 38 looks almost identical to that of recombinant 34, except the haplotype of recombinant 38 has additional haplotype information matching *CA231*.

Looking at the haplotype maps of recombinant 36 and recombinant 38, while focusing on the fact that the former suppresses the BU allele and the latter doesn't, a large section of the left arm is ruled out from containing the causative mutation. Since recombinant 38 cannot suppress the BU allele, any section

that contains the *CA231* haplotype must therefore not contain the causative mutation. Conversely, the causative mutation must still be within the *CA231* haplotype range for recombinant 36, which could rule out the first fifteen million base pairs of the left arm, assuming that double recombination hasn't occurred. If recombinant 38 has no areas of recombination passed the *CG7470* locus, the whole left arm of the chromosome may not contain the causative mutation.

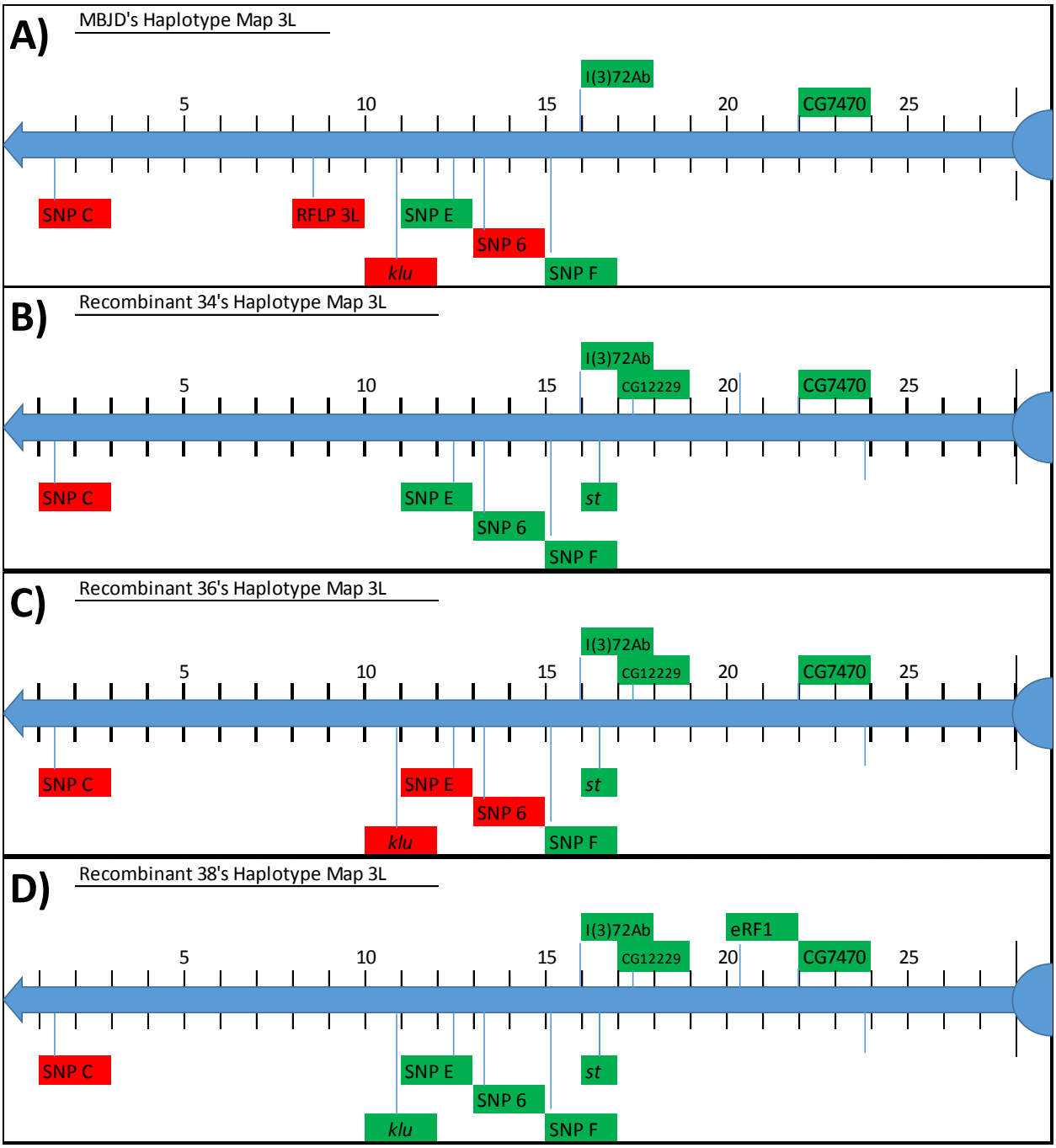


Figure 26: 3L Haplotype Maps. The left arm of the third chromosome for MBJD (A), recombinant 34 (B), recombinant 36 (C), and recombinant 38 (D), are filled in with their respective haplotypes at various locations. The *CA231* haplotype is shown in green, *Oregon R*'s haplotype in red and unknown haplotypes are omitted.

Repeating the process with the right arm, the haplotype map for MBJD appears to be mostly that of the original *CA231* haplotype with the last ten million base pairs being replaced by the *OR* haplotype (Figure 27A). Recombinant 34 appears to have all *CA231* haplotypes on 3R. This recombinant gives us no information on limiting the area where the causative mutation lies (Figure 27B). Recombinant 36 contains some variation with a double cross over occurring between the *PP2A-B'* and *Lerp* loci (Figure 27C). With this area being of OR's haplotype, and recombinant 36 still suppressing, the causative mutation must lay outside this area. Lastly, the 3R haplotype map of recombinant 38 provides additional insight on the areas of the chromosome necessary for suppression (Figure 27D). Assuming that no other areas of recombination have occurred in this recombinant, the causative mutation could be on the far edge of the right arm, anywhere beyond *PP2A-B'*. Looking at the haplotype maps of recombinants 36 and 38, the only area to which recombinant 36 retains the *CA231* haplotype while recombinant 38 has the *OR* haplotype, would be from just after the *Hr96* loci, to the end of the chromosome. Coupled with the information from the left arm, the causative mutation is most likely in this span. MBJD's haplotype reduces the possible candidates in this area. While MBJD's haplotypes cannot give us certainty on which areas of the chromosome are from *CA231* and which are from *17620*, it must contain the causative mutation because it can still suppress the BU phenotype. The number of candidates in this region is reduced down to 42. Five of these are intronic mutations, six don't occur in a gene span, two are in 3' UTRs and 29 of them are most likely gene conversion mutations.

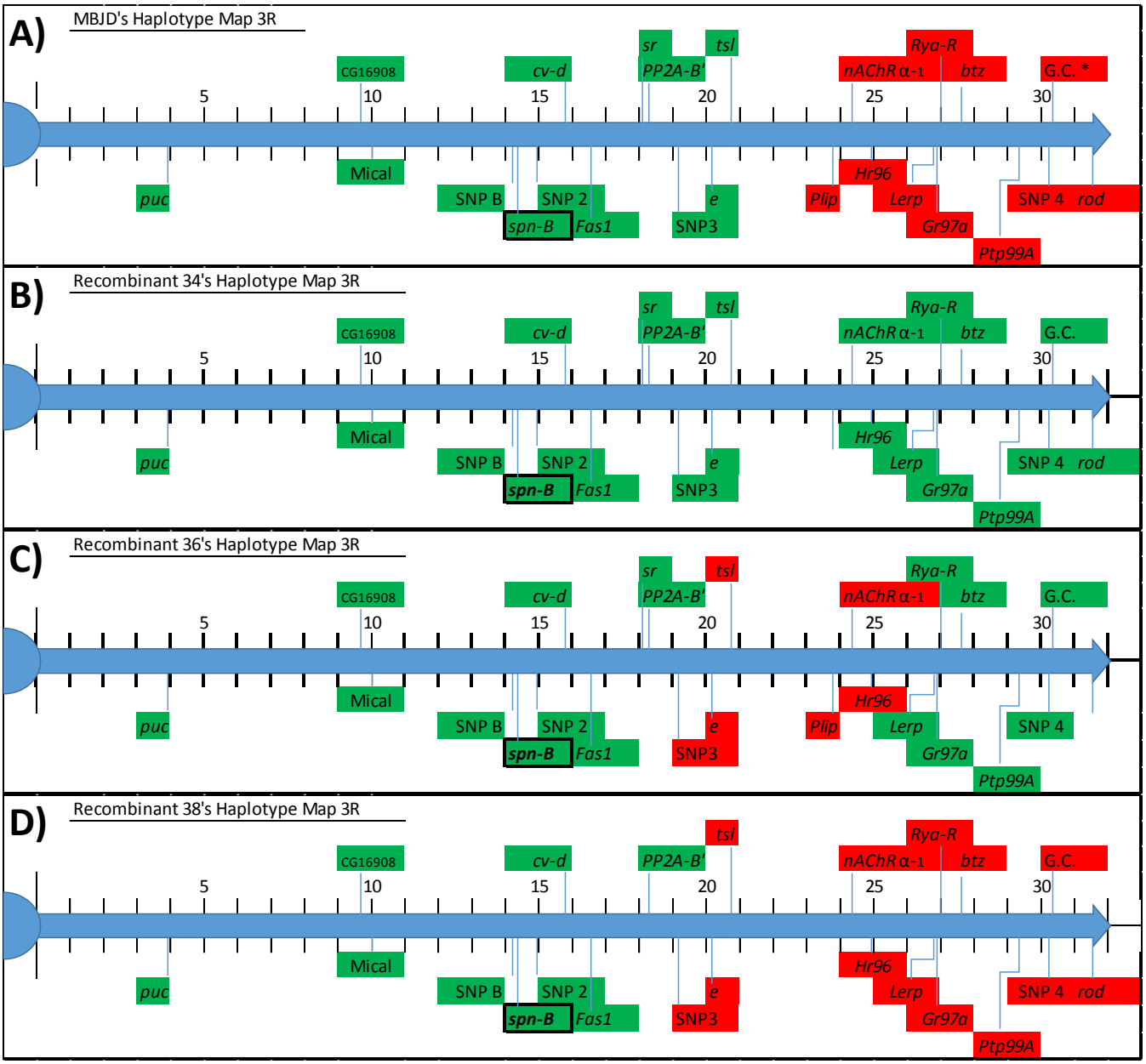


Figure 27: 3R Haplotype Maps. The right arm of the third chromosome for MBJD (A), recombinant 34 (B), recombinant 36 (C), and recombinant 38 (D), are filled in with their respective haplotypes at various locations. The *CA231* haplotype is shown in green, *Oregon R*'s haplotype in red and unknown haplotypes are omitted.

Specific Mutation Testing

During the Fall 2014 Genetics Class at the State University of New York at Fredonia, the amino acid altering mutations were investigated by incorporating knockout alleles of these genes into a *spnB*^{BU} background. Among these knock out alleles were *klu, tsl,* and another amino acid changing mutation, *MsR2.* A condo experiment was performed with homozygous female double mutants (*spnB*^{BU}, and one of the knock out alleles) were tested to see if suppression occurred, females homozygous for the BU allele but heterozygous for the knock out tested for dominant suppression, and a complementation analysis was done between the new double mutant and *CA231*. *MsR2,* located on the tip of the left arm, did not suppress the BU phenotype. *klu* and *tsl* on the other hand, showed dominant suppression and the homozygous *klu* double mutants showed very few defects in the dorsoventral egg patterning (Table 17). Investigating *klu* in the literature, *klumpfuss* has negative regulation of the Ras signaling cascade (Brachmann and Cagan, 2003). With a decrease in Gurken translation, there would be only a small initial signal to the Ras signal cascade, but by stopping any negative regulator, this small signal would grow, possibly to wild type levels, or beyond.

tsl has been shown to have an effect on anterior-posterior development of an embryo (Baek KH, Lee KY, 1999). This developmental patterning is based off of the signaling events that occur in a Stage 7 egg chamber during oogenesis. Gurken is also one of the signals needed for anterior-posterior patterning. Before delving into possible causes and effects for proper *gurken* translation and localization, reviewing the SNP analysis on both of these loci rules them out as being the causative suppressor in *CA231.* Neither the MBJD recombinant nor the recombinant 36 have the CA231 haplotype for *klu,* yet they are suppressing (Table 15). Similarly, recombinant 36 has the *Oregon R* haplotype at the *tsl* locus (Figure 21).

Table 17: Knock Out Allele Condo Data. Two null alleles were crossed into a *spnB*^{BU} background and then assessed for dominant suppression, complementation with *CA231,* and homozygous suppression. **A)** A *klu* knock out allele is assessed and the reference homozygous *spnB*^{BU} data is also present. The students who did this work were Emily Bystrak, Joe Grieco, Rebecca Hartling, Taylor Eddington, Emily Wojcieszek. **B)** A *tsl* knock out allele is assessed for dominant suppression and complementation with *CA231,* but not for homozygous suppression. The students who work to this mutation were Zac Bunge, Zack Eklum, Molly Kalasinski, & Jordan Lamb.

A)	*spn-B*^{BU} / *spn-B*^{BU}		*spn-B*^{BU} / *spn-B*^{BU}, *klu*$^{-}$		*spn-B*^{BU}, *CA231* / *spn-B*^{BU}, *klu*$^{-}$		*spn-B*^{BU}, *klu*$^{-}$ / *spn-B*^{BU}, *klu*$^{-}$	
Class	WT	Ventralized	WT	Ventralized	WT	Ventralized	WT	Ventralized
Relative Frequency	0.041	0.959	0.981	0.019	0.984	0.016	0.966	0.034

B)	*spn-B*^{BU} / *spn-B*^{BU}		*spn-B*^{BU} / *spn-B*^{BU}, *tsl*$^{-}$		*spn-B*^{BU}, *CA231* / *spn-B*^{BU}, *tsl*$^{-}$	
Class	WT	Ventralized	WT	Ventralized	WT	Ventralized
Relative Frequency	0.118	0.882	0.321	0.679	0.629	0.371

In light of these data (Table 13), there is a possibility that there are two suppressing mutations on the 3rd chromosome; the MBJD line may have inherited one and recombinant 36 may have the other mutation. The data is inconclusive as to if there are one mutation or two, and this is a result of dominant suppression (Table 18 & Figure 28). Each of the heterozygotes has well defined suppression and any of the complementation crosses between *CA231,* MBJD and Recombinant 36 have very few pattern defective eggs. One possibility is that there is only one mutation and we are seeing a stronger phenotype as a homozygote. The other possibility is that there are two suppressive mutations and they are both dominant and are additive in the way they suppress. This second possibility is not as plausible for two reasons, the first being that EMS, the random mutagen, would have to mutate two genes that have an effect on the BU phenotype, in the same fly. The chances are quite low for that happening since out of over 10,000 flies screened, only 39 of them could originally suppress the BU phenotype. The second reason why this is unlikely is that the suppression ability of both MBJD and Rec 36, together, have a large Suppression Index and similarly when they are homozygous (Table 18). Even discounting normal variation between condo experiments, CA231, if it contains two dominantly suppressing mutations, should have a higher ability to suppress, but it doesn't.

Table 18: MBJD & recombinant 36 Complementation Test. Crosses were done to determine if the MBJD line carries a different suppressive mutation than recombinant 36. The suppression index of MBJD and recombinant 36 is quite high, either indicating that they have the same mutation and being homozygous gives them a *CA231*-like suppression. Or that they both have dominant suppressors and that they are additive in their suppression ability.

Genotype	Total Eggs	Class 1	Class 2	Class 3	Class 4	Suppression Index
BU Homo	90	0%	3%	12%	84%	0.034
CA231 Homo	124	60%	31%	9%	0%	10.27
BU x CA231	1244	43%	28%	20%	9%	2.45
BU x MBJD	2062	54%	28%	16%	3%	4.43
BU x 36	803	62%	17%	20%	0%	3.87
MBJD x CA231	1573	76%	22%	2%	0%	49.74
CA231 x 36	1277	79%	18%	3%	0%	28.70
36 x MBJD	1631	79%	19%	2%	0%	45.60

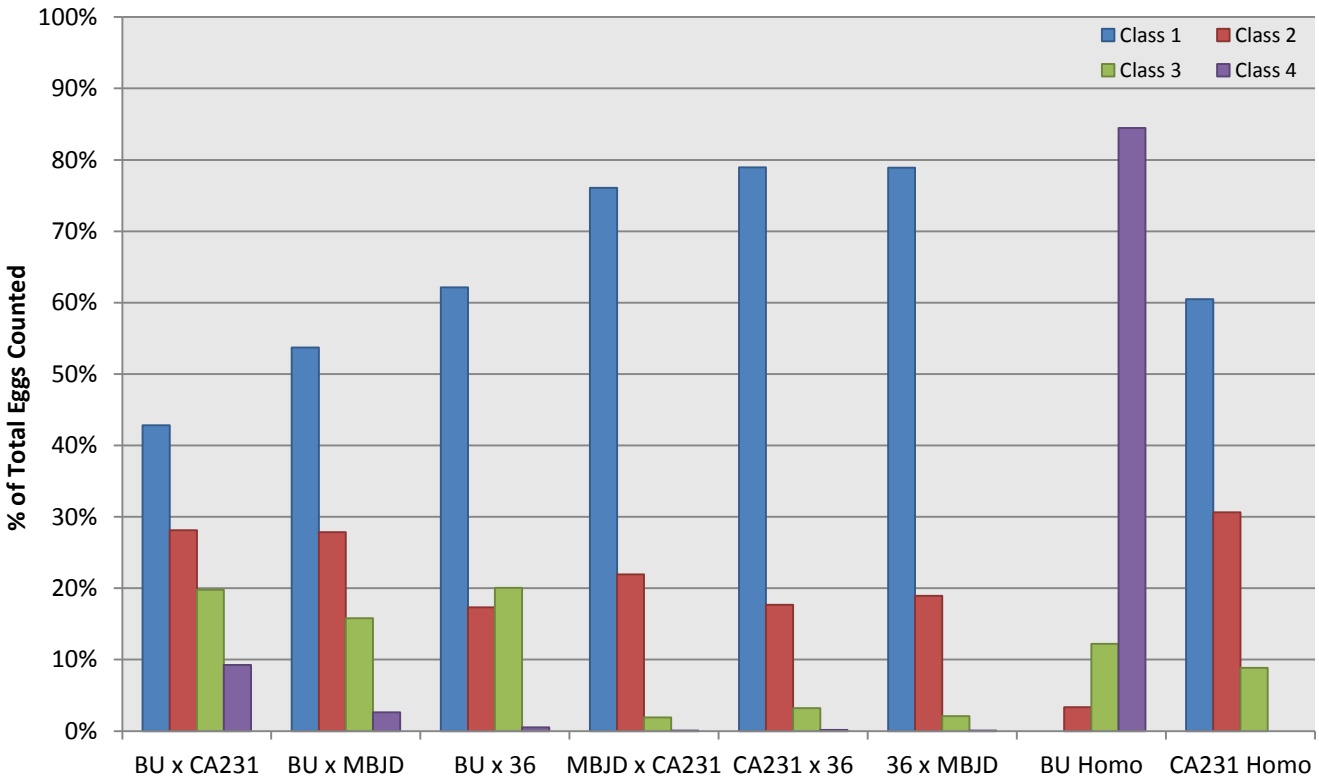


Figure 28: MBJD & recombinant 36 Complementation Test. Crosses were done to determine if the MBJD line carries a different suppressive mutation than recombinant 36. The suppression index of MBJD and recombinant 36 is quite high, either indicating that they have the same mutation and being homozygous gives them a *CA231*-like suppression. Or that they both have dominant suppressors and that they are additive in their suppression ability.

Conclusions and Future Directions

The causative mutation for *CA231* has yet to be found, but the list of candidate genes that are causing the suppression of the BU phenotype has been vastly reduced. Previous researchers, while unable to elucidate the mechanisms of how *CA231* is suppressing, have found unrelated novel suppressors of *spnB*^{BU}, *pyk* being one (Blundon, 2011). In this study, mutations in *klu* and *tsl* have shown that they are novel suppressors of *CA231*. While a rescue construct was attempted for the *klu* locus, the focus remained on identifying the novel mutation in *CA231*.

Multiple loci could not be used as mapping markers because the *CA231* haplotype matched that of *Oregon R*, even though these loci contained clear differences between all of the suppressor lines and *spnB*^{BU}. One possible cause for this was the length of time since the suppressor screen. The suppressor lines were not the only ones being maintained generation after generation accumulating random mutations, the *spn-B*^{BU} experienced the same conditions. It is very possible that after the suppressor lines were already created the *spn-B*^{BU} line was randomly mutated at multiple loci. Since the *spn-B*^{BU} line originally contained all wild type alleles except for three mutated genes, *spindle B*, *stripe* and *ebony*, it makes sense that the suppressor lines would still share the original wild type alleles. When the Next Generation Sequencing data was pooled, these differences were on the list.

A way forward would be to expand the number of recombinants by using recombinant 36 as a parent line. The *Oregon R* line could be used again to maintain consistency with the already established loci. Conversely, mapping lines could provide additional information on limiting the area of the causative mutation. Either way, more recombinants would provide more variation, which would help assess the candidate areas of the chromosome. Collecting unambiguous suppression ability of the recombinants was difficult. Having more recombinants with clear suppression phenotypes would greatly increase the

ability to map the mutation. For the new recombinants, Instead of characterizing all of the loci individually, a low coverage Next Generation Sequencing experiment can be done. This would save time and money in identifying the areas of the chromosome that came from CA231. Any ambiguous readings from the sequencing can be assessed by the loci described above, or additional loci can be developed.

Bibliography

- Abdu, Uri. Brodsky, Michael. and Schüpbach, Trudi. (2002). Activation of a Meiotic Checkpoint during Drosophila Oogenesis Regulates the Translation of Gurken through Chk2/Mnk. Current Biology. 12, 1645-1651.
- Baek KH, Lee KY. (1999). Signal transduction pathway for anterior-posterior development in Drosophila. J Biomed Sci; 6(5): 314-9
- Brachmann, Carrie Baker. Cagan, Ross. (2003). Patterning the fly eye: the role of apoptosis. TRENDS in Genetics; 19(2):91-96
- Christerson, Lori B., and Dennis M. McKearin. "Orb Is Required for Anteroposterior and Dorsoventral Patterning during Drosophita Oogenesis." *GENES & DEVELOPMENT* 8 (1994): 614-62. Web.
- Davidson, A. et al. (2016). Localized Translation of gurken/TGF-α mRNA during Axis Specification Is Controlled by Access to Orb/CPEB on Processing Bodies. Cell Reports; 14(10):2451-62
- Deng, W.M. and Bownes, M. (1997) Two signaling pathways specify localized expression of the Broad-Complex in Drosophila eggshell patterning and morphogenesis. Development 124, 4639–4647
- Dequier, E. et al. (2001) top-DER- and dpp-dependent requirements for the Drosophila fos/kayak gene in follicular epithelium morphogenesis. Mech. Dev. 106, 47–60
- Ghabrial, A., Ray, R. P. & Schüpbach, T. (1998). okra and spindle-B encode components of the RAD52 DNA repair pathway and affect meiosis and patterning in Drosophila oogenesis. Genes Dev. 12, 2711–2723.
- Ghabrial, A., and Schüpbach, T. (1999). Activation of a meiotic checkpoint regulates translation of Gurken during Drosophila oogenesis. Nat. Cell Biol. 1, 354–357.
- Neuman-Silberberg, F.S. and Schüpbach, T. (1993) The Drosophila dorsoventral patterning gene gurken produces a dorsally localized RNA and encodes a TGFa-like protein. Cell 75, 165–174
- Ray, R. and Schüpbach, T. (1996) Intercellular signaling and the polarization of the body axes during Drosophila oogenesis. Genes Dev. 10, 1711–1723
- Peri, F. and Roth, S. (2000) Combined activities of Gurken and Decapentaplegic specify dorsal chorion structures of the Drosophila egg. Development 127, 841–850
- Peri, F. et al. (1999) Local Gurken signaling and dynamic MAPK activation during Drosophila oogenesis. Mech. Dev. 81, 75–88
- Queenan, A.M. et al. (1997) Ectopic activation of Torpedo/EGFR, a Drosophila receptor tyrosine kinase, dorsalizes both the eggshell and the embryo. Development 124, 3871–3880
- Sapir, A. et al. (1998) Sequential activation of the EGF receptor pathway during Drosophila oogenesis establishes the dorsoventral axis. Development 125, 191–200
- Sekelsky, J.J., Brodsky, M.H., and Burtis, K.C. (2000). DNA repair in Drosophila: insights from the Drosophila genome sequence. J. Cell Biol. 150, F31–F36.
- Schupbach, Trudi. Wieschaus, Eric. (1991). Female Sterile Mutations on the Second Chromosome of Drosophila melanogaster. II. Mutations Blocking Oogenesis or Altering Egg Morphology. Genetics 129:1119-1136
- Spradling, A.C. (1993) Developmental genetics of oogenesis. In The Development of Drosophila melanogaster (Bate, M. and Martinez Arias, A., eds), pp. 1–70, Cold Spring Harbor Laboratory Press
- Twombly, V. et al. (1996) The TGF-b signaling pathway is essential for Drosophila oogenesis. Development 122, 1555–1565

- Wasserman, J.D. and Freeman, M. (1998) An autoregulatory cascade of EGF receptor signaling patterns the Drosophila egg. Cell 95, 355–364

Appendix A: Primer & Restriction Enzyme List

The primers that were used for the experiment are shown below. Primers that are used together, differ only by their name and their oligo sequence. Their Oligo Sequences are shown in the 5' to 3' direction. The chromosomal location for each of the primer groups refer to the location of the mutation to which the primers were designed around. If the primer group was given another name, that is shown first in the Amplicon Data section, followed by the type of mutation it is. The greyed out row shows a SNP that was designed but was unable to be used.

Primer Type

Name	Oligo Sequence	Chr. Location	Amplicon Data	Amplification Method	Annealing Temp. °C	Analysis Method

Primary Mutation

spn-B^{BU} Fwd	CAATCGTGCGAAAATGCACTCC	3R:14328056	*spn-B^{BU}*: A RFLP. The wild type amplicon is the full length (454 bp). The *spn-B^{BU}* amplicon is cut by Hpy188III (189, 265 bp).	Taq	64	3% Gel
spn-B^{BU} Rev	AGAAAATTCAGCTCCATTTCGG					

Universal SNPs

JH1	TTTTGCCTGCCAGTCGAAGC	3R:11054150	SNP 1: Oregon R and CA231 shared the same haplotype at this locus.	Beads	54	3%
JH2	AACTTCTCCACCTCGTTGGG					
JH3	ATGCTAAGCCTGCAACCTGC	3R:14992906	SNP 2: An INDEL. CA231's haplotype is a 538bp amplicon. Oregon R's haplotype is a 564bp amplicon.	Beads	64	2%
JH4	GCACCAGAAACAGCAGCACC					
JH5	TCCTCTAGCCCATGACGACG	3R:19270067	SNP 3: A RFLP. Oregon R's haplotype is an amplicon cut once by HinP1I (290, 168 bp). CA231's haplotype is an amplicon cut twice (179, 168, 111 bp).	Beads	64	2%
JH6	ACATCATCACACACACACCC					
JH7	GTGTCTGACAACAACTCTCG	3R:30348884	SNP 4: A RFLP. Oregon R's haplotype is a full length amplicon (442bps). CA231's haplotype is an amplicon cut by BsrI (251, 192 bp).	Beads	54	2%
JH8	ACGACCAGGACAAATCGAGG					
JH9	CGATGAGGAGCAAGTGGTCC		SNP 5: Oregon R and CA231 shared the same haplotype at this locus.			
JH10	GTTGACTCTACAGCAGCTGC					
JH11	CGATGCTCGACCGCAATTCC	3L:13243020	SNP 6: A RFLP. CA231's haplotype is a 388bp amplicon. Oregon R's haplotype is an amplicon cut by AcuI (202, 186bp).	Taq	54	2%
JH12	CATGACTGGACAACAGTGGC					
JH13	TTCGCATGTCAGTGGCGTGG	3R:20070396	SNP 7: Oregon R and CA231 shared the same haplotype at this locus.	Beads	64	2%
JH14	CTGCATGTGTGTGAGTACGG					
JH15	CAAGCCTGAAGGTAATCTCG	3L:23964636	SNP 8: Oregon R and CA231 shared the same haplotype at this locus.	Beads	64	4%
JH16	AAGGCTCGATAGTATCCAGC					
Mike Fwd1	ACCAAGATTAAACTGTCGCC	3R:6485328	SNP A: A RFLP. The reference amplicon is the full length (213 bp). The mutant amplicon is cut by MfeI (170 and 43 bp).	Taq	64	3%
Mike Rev1	GAAAGAGATGTGTGAATCCG					
Kyle Fwd3	CTGCGTTTGACTGGAATTGG	3R:14100888	SNP B: An INDEL. CA231's haplotype is a 414bp amplicon. Oregon R's haplotype is a 368bp amplicon.	Taq	64	2%
Kyle Rev3	CGGTGATGACCTATGTCTCC					

Universal SNPs Cont.

Jared Fwd5	GGTGCAACAACTGCAACTGC	3L:1457055	SNP C. An InDel. CA231's haplotype is a 127bp amplicon. Oregon R's haplotype is a 146bp amplicon.	Taq	64	4%
Jared Rev5	CGCAATGGAAGCCGTTGACC					
Rachel Fwd3	ACTGTAAGCACAATTTAAAACGACC	3L:3433922	SNP D: A RFLP. Oregon R's haplotype is a full length amplicon (159 bp). CA231's haplotype is an amplicon cut by MwoI (105, 54 bp).	Taq	64	3%
Rachel Rev3	TGTTTTCGTAGTCAACTTGTTTGG					
Anoud Fwd3	GTTCGGTGAGCATTATCAGC	3L:12456662	SNP E: A RFLP. CA231's haplotype is a 293bp amplicon. Oregon R's haplotype is an amplicon cut by HpaII (169, 124bp).	Taq	64	3%
Anoud Rev3	CTTAGAGATTGTGGCCATGC					
Alissa Fwd4	CTAAACTTTTCTGCCATCGG	3L:15166730	SNP F: A RFLP. CA231's haplotype is a 217bp amplicon. Oregon R's haplotype is an amplicon cut by HinP1I (147, 70 bp).	Taq	64	3%
Alissa Rev4	TATTCAGCCAACGGATGTGG					

CA231 Specific SNPs

SF627	ACGACTCCGAGTACGACACG	3R:27701050	*btz*: A CA231 specific RFLP. Oregon R's haplotype is a 475bp amplicon. CA231's haplotype is an amplicon cut by MluCI (248, 229bp).	Taq	64	2%
SF628	TCGTGAGACCAGCGATCACC					
SF629	GGTCCGTGCTGACATCATGG	3R:8,108,333	*puc*: A CA231 specific RFLP. *spn-B*^{BU}'s haplotype is a full length amplicon (452 bp). CA231's haplotype is an amplicon cut by FokI (231, 221 bp).	Taq	64	2%
SF630	TGCTCTACGTTTGCATGTGG					
SF631	TGTTGTCGAGAACCTCGTGG	3R:21778495	*tsl*: A CA231 specific RFLP. CA231's haplotype is a 372bp amplicon. Oregon R's haplotype is an amplicon cut by SfaNI (273, 93bp).	Taq	64	2%
SF632	AGGTGTTTGTGTACAGCAGG					
SF653	CTTGGTCCATTTCCTGTTCG	3L:10984445	*klu*: A CA231 specific mutation. Haplotype is discerned by sequencing. SF654 is the sequencing primer	Bead	54	Seq
SF654	TCGCTCTTGATGTACTTGCC					
SF655	CAATCTCACTCCCACTACGG					
SF691	TTTATGCTTTTGTGCCTGGCC	3L:4890926	Oregon R and CA231 shared the same haplotype at this locus.	Taq	54	4%
SF692	GGATATTGATGGCGTTCTCG					
SF693	ATTCATTGTCGCGTGTTGCG	3L:8657304	Oregon R and CA231 shared the same haplotype at this locus.	Taq	54	
SF694	AGTCTGTCCGTTTCTGTTGC					
SF695	CTACATCGGAGGAATTCTGC	3R:15844615	*cv-d:* A CA231 Specific SNP. The haplotype was discerned by sequencing. SF697 is the sequencing primer.	Taq	54	Seq
SF696	AATTGTGGAGTGTACGCAGC					
SF697	ACAGGAGTTGGAGTGTTTGC					
SF698	AAAGGTGAACGCGTTGAACC	3R:16052929	msps: A CA231 Specific SNP. Oregon R's haplotype is an amplicon cut twice with SfaNI (120, 112, 18). CA231's is an amplicon cut once (130, 120).	Taq	54	4% Gel
SF699	TCGATTTCTTCAACGCCAGC					
SF700	ACACACTCTCATACTCGAGC	3R:16638094	A CA231 Specific SNP. The haplotype was discerned by sequencing. SF702 is the sequencing primer	Taq	54	Seq
SF701	GTAACTAAGTGCAGGATGGC					
SF702	CATTCCCCCTATTCCTATCC					
SF703	TTTAGCCAACAACAGCAGCG	3R:17997601	A CA231 Specific SNP. The haplotype was discerned by sequencing. SF704 is the sequencing primer	Taq	54	Seq
SF704	ACACTGGCAAACACTTCAGC					
SF705	GCTAAAACGGGTTTGACAGC					
SF706	TAATTGCTTGGGCGGTTTCG	3R:18172727	A CA231 Specific SNP. The haplotype was discerned by sequencing. SF707 is the sequencing primer	Taq	54	Seq
SF707	CGGTTTCGACTCACAAATCG					
SF708	GCAGCAGTCGAAGCAACAAA					
SF709	AGTTACAGCTGTTGCGTTCG	3R:23758772	*plip*: A CA231 Specific mutation. The haplotype was discerned by sequencing. SF710 is the sequencing primer.	Taq	54	Seq
SF710	TCAGGACATTGTACAGCAGG					
SF711	ATCACATGCTCATCGATGCG					
SF712	AGGGTATGATCAGGTTGACC	3R:24405397	*nAChRalpha1*: A CA231 Specific mutation. The haplotype was discerned by sequencing. SF713 is the sequencing primer.	Taq	54	Seq
SF713	TGTAGAACTT CTCGTTCCGC					
SF714	AAAGAAAGACCAGCTGGTCC					

CA231 Specific SNPs Cont.

SF715	AGTTCATGAATACCCCAGCG	3R:25026721	*Hr96*: A CA231 Specific mutation. The haplotype was discerned by sequencing. SF716 is the sequencing primer.	Taq	54	Seq
SF716	GTTGGCTATGTAAGGCAAGG					
SF717	AGGTCTTCAGATCGAAGTCC					
SF718	ATGAGAATCAGCAGGCTTCC	3R:26965308	*Gr97a*: A CA231 Specific mutation. The haplotype was discerned by sequencing. SF719 is the sequencing primer.	Taq	54	Seq
SF719	TACCTTATTGACCAGCGAGG					
SF720	ACGACTATTTGCAGACCAGG					
SF721	ACTACTACTGAGTCCACTCC	3R:27002187	*Rya-R*: A CA231 Specific mutation. The haplotype was discerned by sequencing. SF722 is the sequencing primer.	Taq	54	Seq
SF722	GAAATGTGTCGATGGAGAGC					
SF723	CCCATTTATGTCCCTGATGC					
SF724	TCACAGCTCATTTGTCGAGG	3R:26861311	*Lerp*: A CA231 Specific mutation. The haplotype was discerned by sequencing. SF725 is the sequencing primer.	Taq	54	Seq
SF725	GATCATTACAGGCCAAAGGC					
SF726	ATACGAGCAATCGTCAAGGC					
SF728	ATCGGAGGGAAAGGAAAACC	3R:29401231	Ptp99A: A CA231 Specific mutation. The haplotype was discerned by sequencing. SF730 is the sequencing primer.	Taq	54	Seq
SF729	TTCCCCTGCTTTTATGACCC					
SF730	CTCAAGGCTG GATCCAAAGC					
SF731	ATGAGACCTGCTTGGAAACG	3R:31581796	*rod*: A CA231 Specific mutation. The haplotype was discerned by sequencing. SF732 is the sequencing primer.	Taq	54	Seq
SF732	ATTCGTGTCACGTCATTGGC					
SF733	GAAGATCTGCTTTAGGCTGC					
SF734	GGAAAGCCAGTGAGTTATGC	3R:30419147	A CA231 Specific gene conversion track. The haplotype was discerned by sequencing. SF735 is the sequencing primer.	Taq	54	Seq
SF735	ATTCTGCTGGCTTCTAGTCG					
SF736	AATAGCAACGCCGATTCAGC					

Enzyme List: Details about the enzymes used, including their Order Number through NEB. For each digest, two Units of enzymes were used per sample digested. Reactions are done in ten microliters with the buffer being one tenth the final volume. The optimal buffer and any additional additive are shown. The enzymes optimal functioning temperatures are also shown.

Name	Enzyme Order Number	Units / mL	Volume for 2 Units (μL)	Buffer	Additional Buffer(s)	Optimal Working Temperature
MluCI	R0538S	10000	0.2	cutsmart	-	37
Hpy188III	R0622L	5000	0.4	cutsmart	-	37
FokI	R0109S	5000	0.4	cutsmart	-	37
BccI	R0704S	10000	0.2	cutsmart	-	37
HinP1I	R0124S	10000	0.2	cutsmart	-	37
BsrI	R0527S	10000	0.2	3.1	-	65
MboII	R0148S	5000	0.4	cutsmart	-	37
AcuI	R0641S	5000	0.4	cutsmart	40 uM of SAM	37
HpyAV	R0621S	2000	1	cutsmart	-	37
MfeI	R3589S	20000	0.1	cutsmart	-	37
MwoI	R0573S	5000	0.4	cutsmart	-	60
HpaII	R0171S	10000	0.2	cutsmart	-	37
HinP1I	R0124S	10000	0.2	cutsmart	-	37
SfaNI	R0172S	2000	1	3.1	-	37

Appendix B: Additional Suppression Data

Table C1: Condo 1 Suppression Data. Multiple variants of the many fly lines are tested to identify any line that still contains the suppressive mutation captured in the forward genetic screen. The *spnB*^{BU} heterozygous line was used as a control for wild type eggs. Some lines failed to produce many eggs, some none at all, and for those lines, data wasn't available (-).

Genotype	Total Eggs	Class 1	Class 2	Class 3	Class 4	Suppression Index
spnB^{BU} homo	552	1%	10%	7%	82%	0.1242
spnB^{BU} hetero	275	100%	0%	0%	0%	-
CE261 3γ	169	13%	47%	18%	22%	1.4853
CR55 3α	0	-	-	-	-	-
CR55 3γ	0	-	-	-	-	-
CR55 3γ	19	0%	11%	0%	89%	0.1176
CR855 2α	2	0%	50%	0%	50%	1
CB429 α	7	71%	29%	0%	0%	-
CB 429 β	2	100%	0%	0%	0%	-
CB 429 β	6	67%	33%	0%	0%	-
CB 429 γ	8	63%	38%	0%	0%	-
CE 121 1α	116	0%	49%	47%	3%	0.9661
CE 121 1β	172	13%	58%	24%	5%	2.3725
CE 121 1β	383	19%	63%	18%	0%	4.4714
CE 121 3α	155	3%	63%	34%	0%	1.9245
CE 121 3β	83	0%	17%	80%	4%	0.2029
CE 149 1α	54	6%	72%	20%	2%	3.5
CE 149 2α	141	23%	62%	15%	0%	5.7143
CE 149 2β	4	50%	50%	0%	0%	-

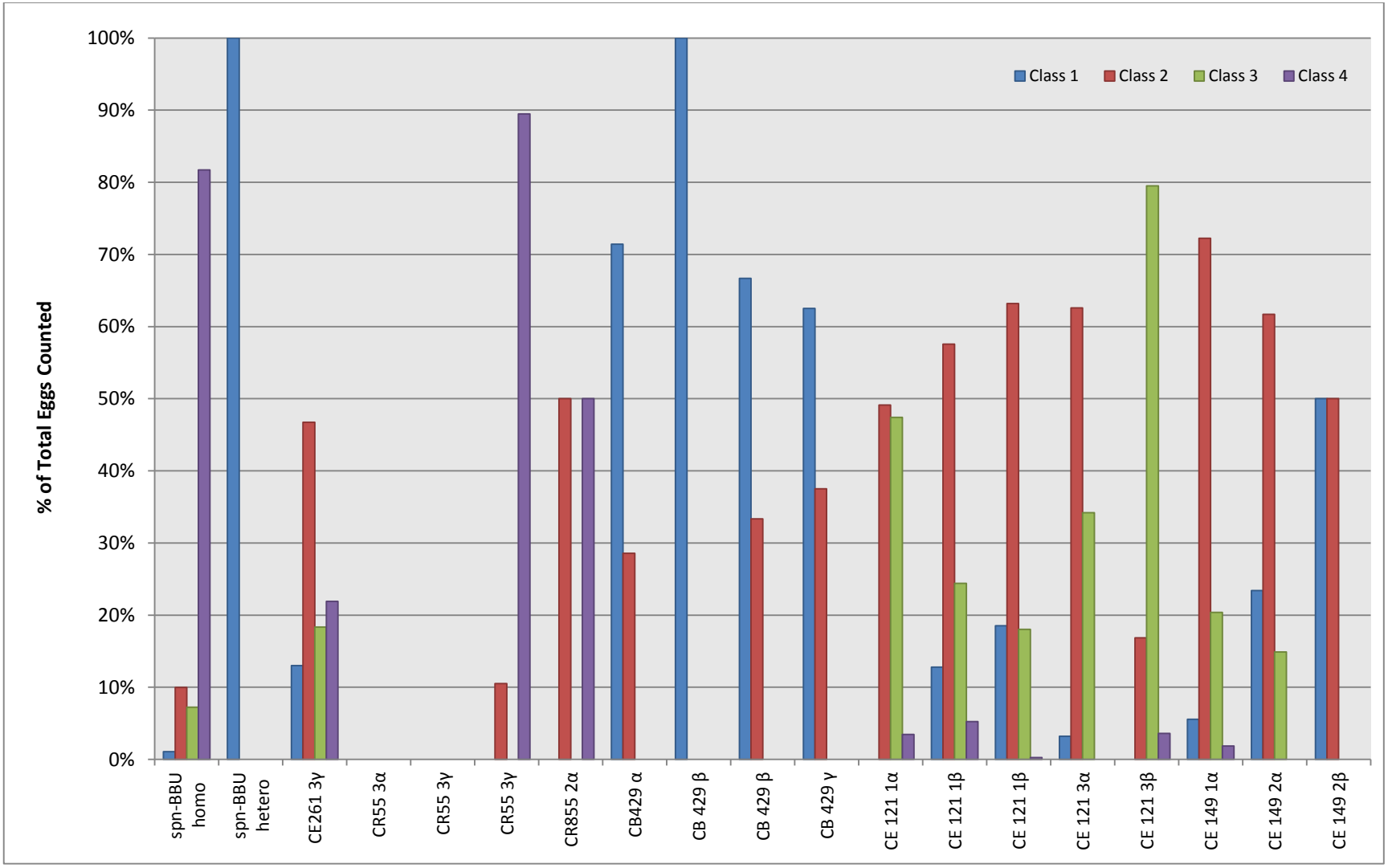


Figure C1: Condo 1 Egg Shell Classifications. Based on the information from Table C1, the percent of each egg shell type from each novel line tested. The *spnB*^{BU} homo (homozygous) line shows the high percent of Class Four eggs. The *spnB*^{BU} hetero (heterozygous) shows only wild type eggs (Class 1). Mutants that have good suppression capabilities should look more like the wild type, *spnB*^{BU} heteros than the *spnB*^{BU} homos.

Table C2: Condo 2 Suppression Data. Multiple variants of the many fly lines are tested to identify any line that still contains the suppressive mutation captured in the forward genetic screen. Some lines failed to produce many eggs, some none at all, and for those lines, data wasn't available (-). The green highlighted mutant lines are the isogenized isolates chosen to represent the parent line.

Genotype	Total Eggs	% WT	% V2	%V3	% V4	Suppression Index
WT	-	-	-	-	-	-
spnB^{BU} Homo	574	6%	9%	13%	72%	0.176229508
CE 121 1α	0	-	-	-	-	-
CE 121 1β	161	13%	47%	39%	1%	1.515625
CE 121 3α	50	0%	12%	88%	0%	0.136363636
CE 121 3γ	164	2%	18%	71%	9%	0.251908397
CB 429 α	118	56%	34%	8%	2%	8.833333333
CB 429 β	63	19%	48%	33%	0%	2
CB 429 γ	29	45%	41%	10%	3%	6.25
CE 149 1α	120	13%	67%	13%	8%	3.8
CE 149 1β	130	2%	39%	46%	12%	0.710526316
CE 149 1γ	360	17%	67%	12%	4%	5.206896552
CE 149 2β	155	9%	88%	3%	1%	30

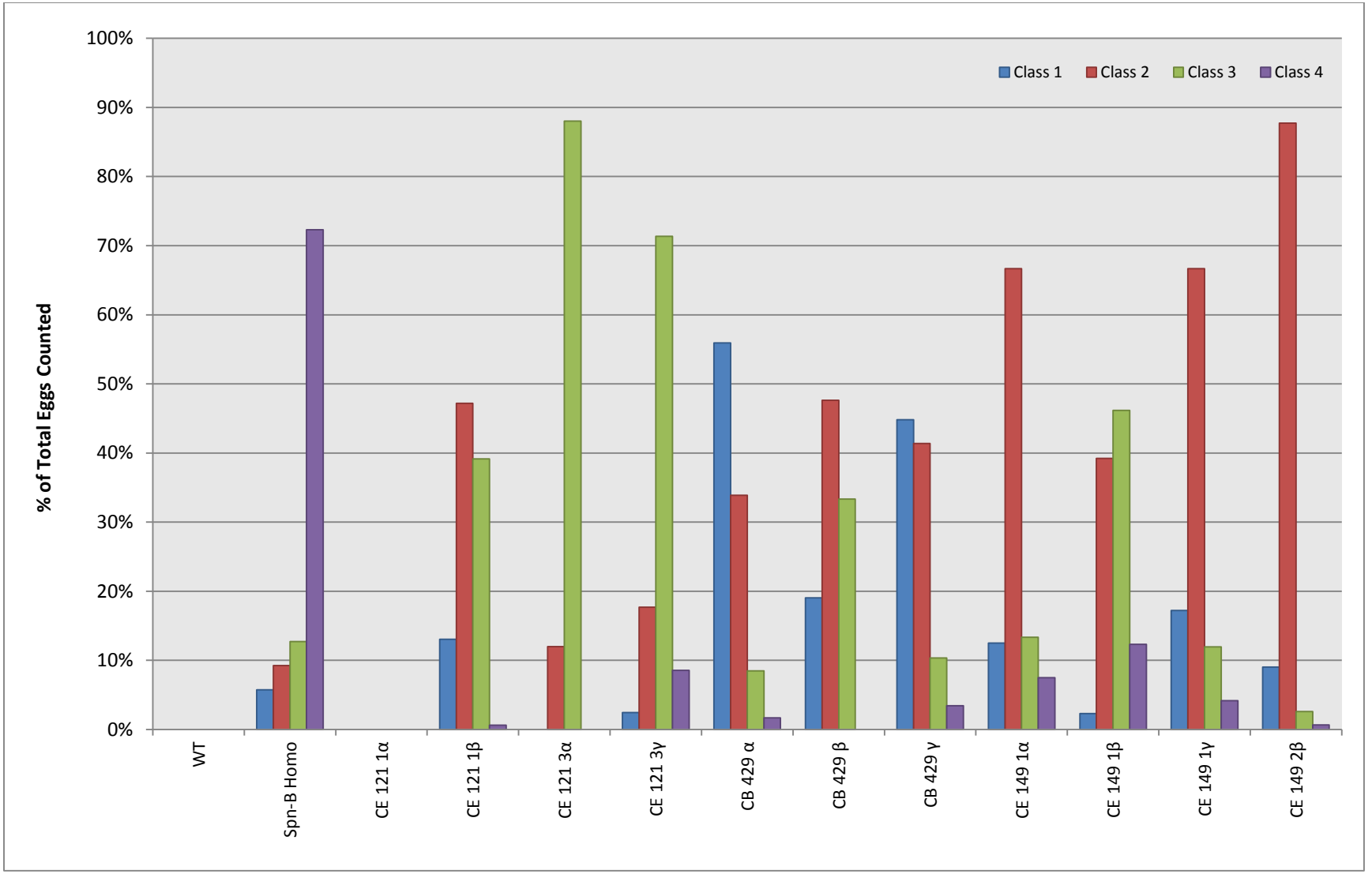


Figure C2: Condo 2 Egg Shell Classifications. Based on the information from Table C2, the percent of each egg shell type from each novel line tested. The *spnB*^{BU} homo (homozygous) line shows the high percent of Class Four eggs. The Wild Type (WT) line was used as a reference to what Class 1 eggs look like. Mutants that have good suppression capabilities should have more Class 1 and Class 2 eggs.

Table C3: Condo 3 Suppression Data. Multiple variants of the many fly lines are tested to identify any line that still contains the suppressive mutation captured in the forward genetic screen. The WT line did produce eggs, but they were used as reference for identifying Class 1 eggs. CR 855 2α failed to produce any eggs (-). The orange highlighted mutant line is the isogenized isolate chosen to represent the parent line (CA23). It is orange and not green, to reflect the lower suppression ability.

Genotype	Total Eggs	Class 1	Class 2	Class 3	Class 4	Suppression Index
WT	-	-	-	-	-	-
spnB^{BU} Homo	718	2%	9%	6%	82%	0.130708661
CA 23 A	877	4%	29%	23%	45%	0.486440678
CA 23 B	1085	5%	23%	19%	53%	0.383928571
CA 23 C	660	6%	31%	35%	27%	0.601941748
CR 55 3α	17	29%	35%	24%	12%	1.833333333
CR 55 3β	159	0%	2%	34%	64%	0.019230769
CR 55 3γ	155	3%	17%	19%	61%	0.25
CR 855 2α	0	-	-	-	-	-

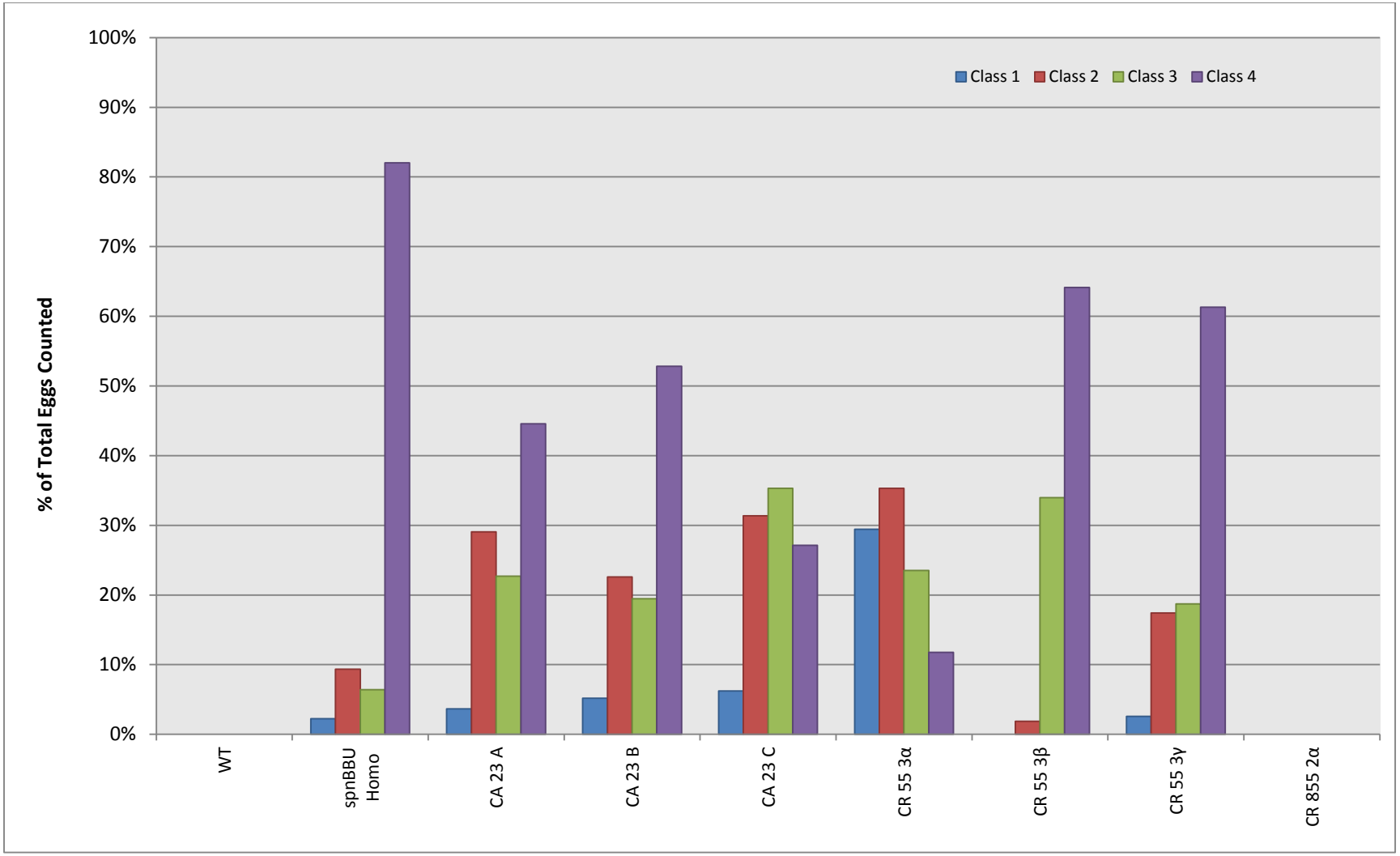


Figure C3: Condo 3 Egg Shell Classifications. Based on the information from Table C3, the percent of each egg shell type from each novel line tested. The *spnB*^{BU} homo (homozygous) line shows the high percent of Class Four eggs. The Wild Type (WT) line was used as a reference to what Class 1 eggs look like. Mutants that have good suppression capabilities will have more Class 1 and Class 2 eggs.

Table C4: Condo 4 Suppression Data. Multiple variants of the many fly lines are tested to identify any line that still contains the suppressive mutation captured in the forward genetic screen. The WT line did produce eggs, but they were used as reference for identifying Class 1 eggs. The green highlighted mutant lines are the isogenized isolate chosen to represent their parent lines.

Genotype	Total Eggs	Class 1	Class 2	Class 3	Class 4	Suppression Index
spn-B^{BU} Homo	411	1%	8%	9%	82%	0.098930481
WT	-	-	-	-	-	-
CB 34 α	82	39%	33%	11%	17%	2.565217391
CB 34 γ	95	5%	15%	11%	69%	0.25
CE 261 β	400	6%	8%	1%	85%	0.166180758
CE 261 3γ	360	8%	17%	14%	61%	0.338289963
CC 701 β	582	19%	33%	26%	22%	1.086021505
CB 641 α	158	2%	1%	1%	97%	0.025974026
CB 641 β	3	0%	0%	0%	100%	0
CB 641 γ	198	1%	0%	0%	99%	0.010204082
CD 100 α	590	13%	17%	5%	65%	0.418269231
CD 100 β	711	22%	30%	3%	45%	1.085043988

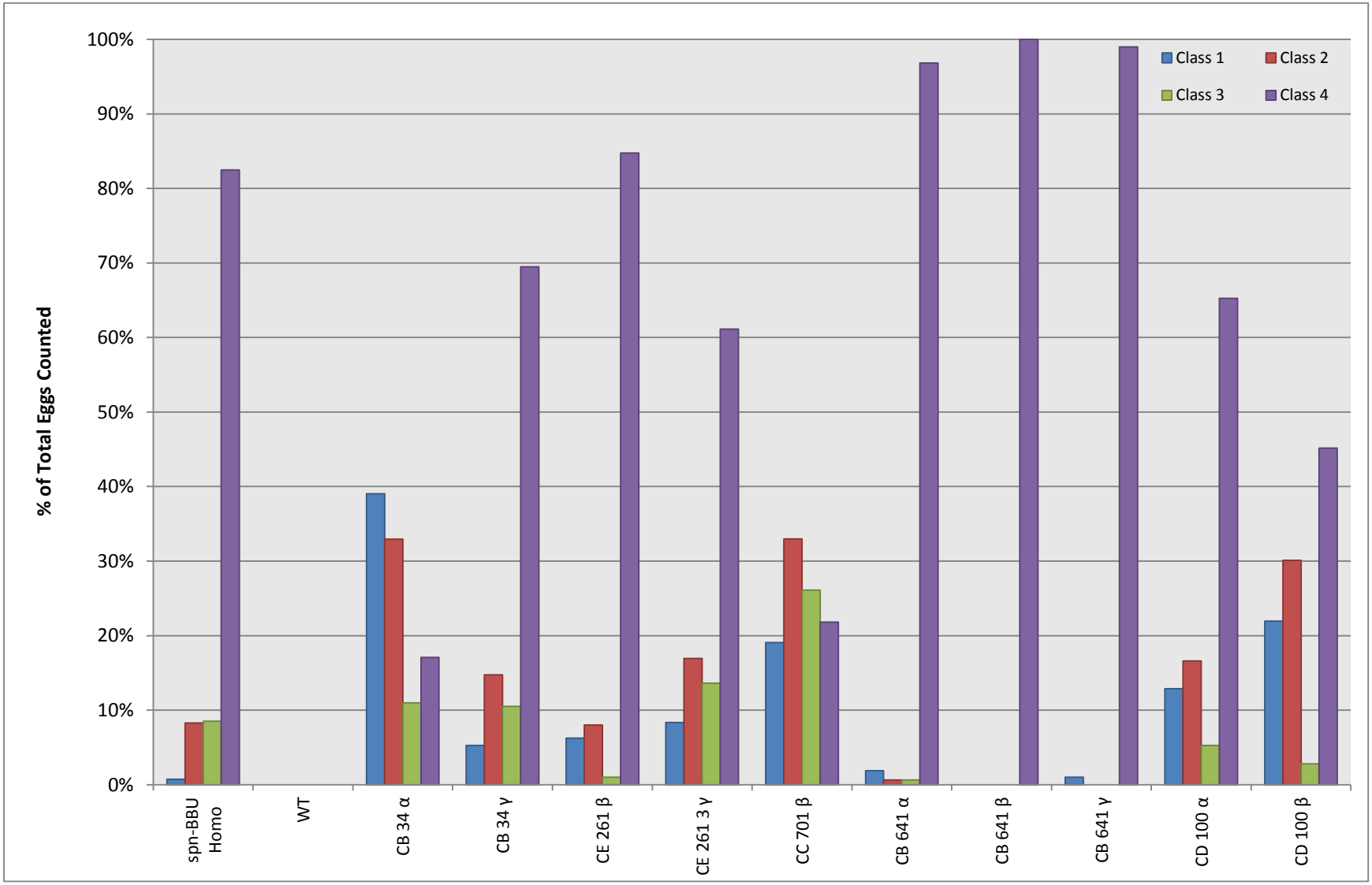


Figure C4: Condo 4 Egg Shell Classifications. Based on the information from Table C4, the percent of each egg shell type from each novel line tested. The *spnB*^{BU} homo (homozygous) line shows the high percent of Class Four eggs. The Wild Type (WT) line was used as a reference to what Class 1 eggs look like. Mutants that have good suppression capabilities will have more Class 1 and Class 2 eggs.

Table C5: Condo 5 Suppression Data. Directed Testing of a few variants of Novel fly lines to identify any line that can suppress the *spnB*^{BU} mutation. The green highlighted mutant line is the isogenized isolate chosen to represent their parent lines. The orange isolates were also chosen to represent their lines, but their suppression ability isn't as good.

Genotype	Total Eggs	Class 1	Class 2	Class 3	Class 4	Suppression Index
WT Control	707	100%	0%	0%	0%	-
spn-B^{BU} Control	278	2%	11%	0%	87%	0.153526971
CR590	246	41%	58%	1%	0%	60.5
CB23α	15	7%	27%	0%	67%	0.5
CB23γ	225	11%	17%	0%	72%	0.388888889
CD801α	7	14%	71%	0%	14%	6
CD801β	11	18%	27%	0%	55%	0.833333333

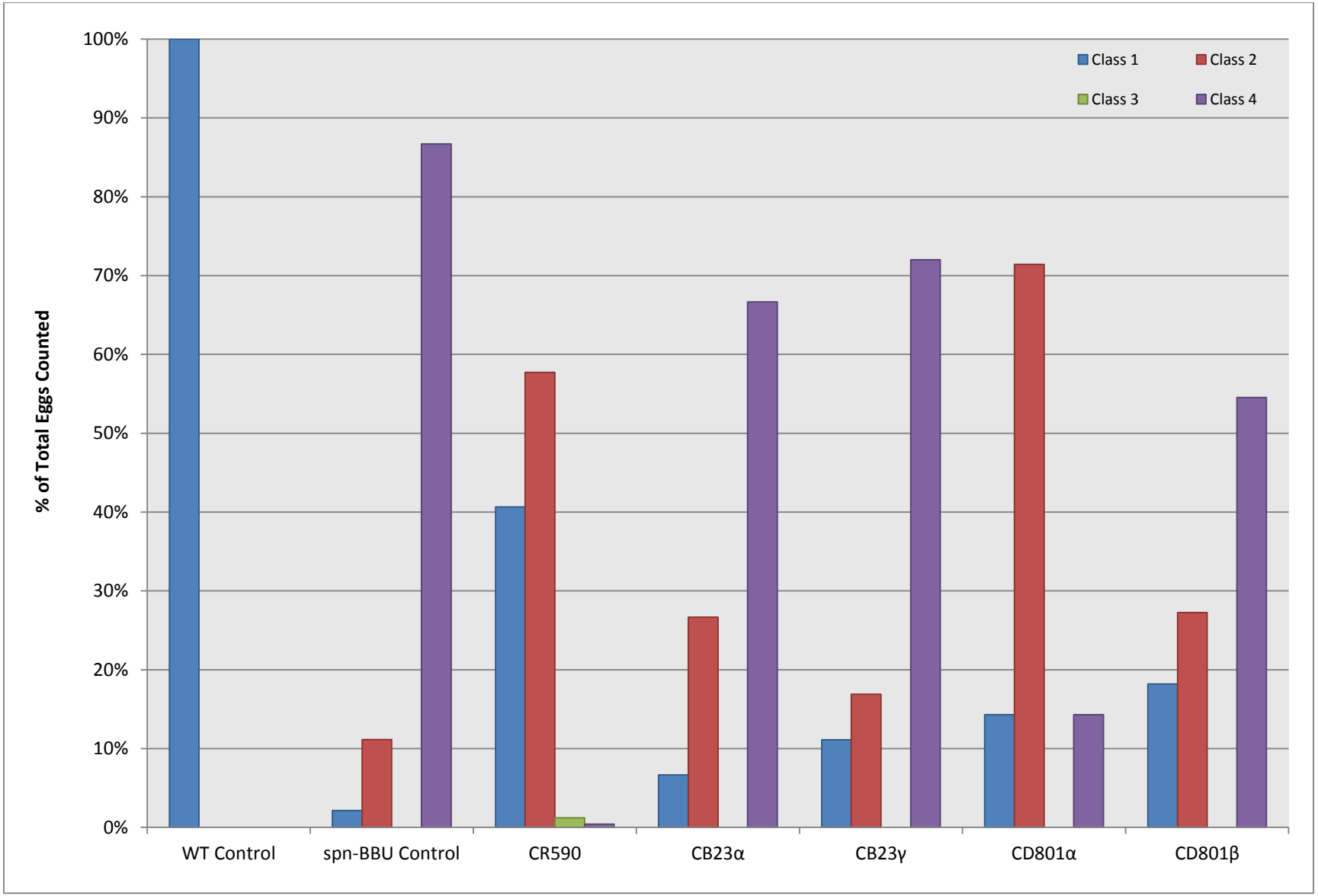


Figure C5: Condo 5 Egg Shell Classifications. Based on the information from Table C5, the percent of each egg shell type from each novel line tested. The *spnB*^{BU} homo (homozygous) line shows the high percent of Class Four eggs. The Wild Type (WT) line was used as a reference to what Class 1 eggs look like. Mutants that have good suppression capabilities will have more Class 1 and Class 2 eggs.

Appendix C: Annotated Next Generation Sequencing (NGS) Results

The annotated and simplified results from the NGS are shown below. The chromosomal positions are based on the 6th edition of Flybase. Changes between the reference sequence, homozygous *spnB*^{BU}, and the alternative sequence, CA231, are shown in columns two and three, respectively. Purple fill indicates that those differences are EMS like mutations. The number of fly lines that have the same mutation are shown in the "count" column. The count column has a color gradient scaled to easily identify those mutations that are solely unique (no fill) from mutations that are in all ten of the sequencing lines (red fill).

The last column has characteristics of the mutation. The mutations that do not occur in a gene do not have a white fill. Ones that may be part of a gene conversion track have red lettering. Introns are shown in blue. Synonymous mutations are shown in pink, while yellow shows the nonsynonymous mutations. Brown shows mutations that occur in either UTRs or is from a non-protein coding gene. Mutations that are light green represent any mutation that is of interest. Any row that is colored hunter green, represents phenotypic marker mutations like *spindle B*, *scarlet, etc* that give additional information to locations on the third chromosome.

Position based on Flybase V6	*spnB*^{BU} REF Seq	CA231 ALT Seq	count	Ferguson / Hasper Annotations
3R:4,691,036	T	TAATAATA	1	Intron of *CG43427*
3R:4,691,038	T	TTAAACGAA	1	Intron of *CG43427*
3R:4,759,550	C	T	1	Intron of *atms* and *CG44908*
3R:5,057,215	C	A	1	No Gene
3R:5,143,135	C	T	3	Present in CA231, CR324, and CA1301
3R:5,345,962	C	T	1	Intron of *mtd*
3R:5,577,470	C	T	1	Intron of *CRMP*
3R:5,708,775	C	T	1	Synonymous mutation in *plx*
3R:6,282,120	G	A	1	Gene Conversion
3R:6,282,223	G	A	1	Gene Conversion
3R:6,282,318	T	C	1	Gene Conversion
3R:6,485,328	A	T	10	Present in all suppressors - Likely mutation in parental line
3R:6,485,364	A	T	10	Present in all suppressors - Likely mutation in parental line
3R:6,828,399	C	T	1	Intron of *Scr*
3R:7,161,033	G	A	1	Gene Conversion
3R:7,161,188	C	T	1	Gene Conversion
3R:7,161,667	A	G	1	Gene Conversion
3R:7,161,672	A	G	1	Gene Conversion
3R:7,161,761	C	G	1	Gene Conversion
3R:7,441,451	C	T	1	Synonymous mutation in *CG14598*
3R:7,474,644	GTTAC	G	3	Present in CA231, CR324, and CA1301
3R:7,533,850	C	T	1	3' UTR of *alpha-Est5*
3R:7,792,529	C	T	1	Synonymous mutation in *CG10086*
3R:8,108,333	C	T	1	R to W mutation in *puckered* - Ovarian follicle cell development.
3R:8,364,143	C	T	1	R to W mutation in *CG7963-RB* - Beyond the 3' end of the gene.
3R:8,507,113	A	C	1	No Gene
3R:9,025,684	C	T	1	Intron of *neur* - ubiquitin ligase

3R:9,567,815	C	T	1	Synonymous mutation in *CG16789*
3R:9,753,267	C	T	1	D to N mutation in *CG16908*
3R:10,015,100	C	A	1	G to C mutation in *Mical*
3R:10,259,601	A	G	10	Present in all suppressors - Likely mutation in parental line
3R:11,054,150	GTGA	G	10	Present in all suppressors - Likely mutation in parental line
3R:11,272,437	C	A	1	Intron of ncRNA of *CR31386*
3R:11,335,365	G	A	7	Intron in two isoforms / 5' UTR of one of *pros - prospero*
3R:11,438,175	C	T	1	Intron of *CG17230*, E to K mutation in *sals*
3R:11,798,169	A	T	1	Intron of *ClC-a*
3R:11,847,192	C	T	1	Intron of *CG6959*
3R:11,849,384	G	A	1	Intron of *CG6959*
3R:11,973,273	C	T	1	Synonymous mutation in *Tango9*
3R:12,080,430	C	T	1	Intron of *CG14741*
3R:12,349,644	C	T	1	Synonymous mutation in *CG10038*
3R:12,686,130	C	T	1	Intron of *kar*
3R:12,887,914	C	T	1	No Gene
3R:13,441,876	G	A	3	Present in CA231, CR324, and CA1301
3R:13,442,157	G	C	3	Present in CA231, CR324, and CA1301
3R:13,442,234	A	G	3	Present in CA231, CR324, and CA1301
3R:13,442,425	T	TTGAGAG	3	Present in CA231, CR324, and CA1301
3R:13,442,431	G	A	3	Present in CA231, CR324, and CA1301
3R:13,443,499	A	C	3	Present in CA231, CR324, and CA1301
3R:13,515,002	C	T	1	No Gene
3R:13,568,798	AG	A	10	Present in all suppressors - Likely mutation in parental line
3R:13,671,298	G	T	1	Intron of *Task6*
3R:13,671,311	A	G	1	Intron of *Task6*
3R:14,074,510	G	C	1	Intron of *foxo*
3R:14,100,694	C	T	10	*Pk1-R* - Pyrokinin 1 receptor
3R:14,100,888	C	T	10	*Pk1-R* - Pyrokinin 1 receptor
3R:14,100,909	A	G	10	Present in all suppressors - Likely mutation in parental line
3R:14,100,953	A	G	10	Present in all suppressors - Likely mutation in parental line
3R:14,100,982	G	T	10	Present in all suppressors - Likely mutation in parental line
3R:14,100,983	T	A	10	Present in all suppressors - Likely mutation in parental line
3R:14,101,025	T	TC	10	Present in all suppressors - Likely mutation in parental line
3R:14,101,026	T	TTCATCTGC	10	Present in all suppressors - Likely mutation in parental line
3R:14,101,029	T	TGATG	10	Present in all suppressors - Likely mutation in parental line
3R:14,101,031	A	ATTCGAGTTATTCG	10	Present in all suppressors - Likely mutation in parental line
3R:14,101,034	A	T	10	Present in all suppressors - Likely mutation in parental line
3R:14,101,037	T	TGCGCTGTCTCT TGAAA	10	Present in all suppressors - Likely mutation in parental line
3R:14,101,041	A	ACATT	10	Present in all suppressors - Likely mutation in parental line
3R:14,101,079	C	T	10	Intron of *Pk1-R* - Pyrokinin 1 receptor
3R:14,101,091	C	A	10	Present in all suppressors - Likely mutation in parental line
3R:14,326,051				*spn-B*
3R:14,730,036	T	G	10	Present in all suppressors - Likely mutation in parental line
3R:14,992,906	G	T	10	Present in all suppressors - Likely mutation in parental line
3R:14,992,934	C	A	10	Present in all suppressors - Likely mutation in parental line
3R:14,992,935	AGCTTCACTTG	A	10	Present in all suppressors - Likely mutation in parental line
3R:14,992,964	ATTGCC	A	10	Present in all suppressors - Likely mutation in parental line
3R:14,992,973	C	T	10	No Gene
3R:14,992,989	A	G	10	Present in all suppressors - Likely mutation in parental line
3R:14,993,015	C	CT	10	Present in all suppressors - Likely mutation in parental line
3R:14,993,017	GATCTTTTTCTAC	G	10	Present in all suppressors - Likely mutation in parental line

3R:14,993,088	C	G	10	Present in all suppressors - Likely mutation in parental line
3R:14,993,095	C	A	10	Present in all suppressors - Likely mutation in parental line
3R:14,993,099	T	G	10	Present in all suppressors - Likely mutation in parental line
3R:15,470,526	TGC	T	10	Present in all suppressors - Likely mutation in parental line
3R:15,470,529	C	A	10	Present in all suppressors - Likely mutation in parental line
3R:15,470,599	C	T	10	Intron of *AdamTS-A*
3R:15,470,620	A	G	10	Present in all suppressors - Likely mutation in parental line
3R:15,470,622	T	C	10	Present in all suppressors - Likely mutation in parental line
3R:15,538,821	C	T	1	Synonymous mutation in *CG18519*
3R:15,844,615	C	T	1	R to H mutation in *cv-d*
3R:15,878,072	C	T	1	No Gene
3R:16,052,929	G	A	3	*mini spindles* (*msps*) Present in CA231, CR324, and CA1301
3R:16,229,375	C	G	1	Intron of *tara*
3R:16,572,931	C	T	1	Synonymous mutation in *CG14891*
3R:16,638,094	C	T	1	P to S mutation in *Fas1*
3R:17,997,601	C	T	1	Intron of *cpo*
3R:18,044,488	C	A	10	Present in all suppressors - Likely mutation in parental line
3R:18,044,536	C	G	10	Present in all suppressors - Likely mutation in parental line
3R:18,044,636	C	T	10	No Gene
3R:18,143,785	GCTC	G	1	No Gene
3R:18,146,031	G	A	1	No Gene
3R:18,172,727	C	T	1	Intron of *PP2A-B'*
3R:19,212,139	G	A	1	No Gene
3R:19,270,067	A	G	10	Present in all suppressors - Likely mutation in parental line
3R:19,270,318	C	T	10	Present in all suppressors - Likely mutation in parental line
3R:19,272,216	C	T	1	No Gene
3R:19,601,954	C	T	1	Synonymous mutation in *CG7333*
3R:20,016,761	C	T	1	Intron of *CG4836*
3R:21,121,508	C	T	1	Intron of *slmb - supernumerary limbs*
3R:21,229,839				ebony
3R:21,245,059	C	T	1	P to S mutation in *CG5892*
3R:21,661,727	G	A	1	3' UTR of *CG6353*
3R:21,723,081	C	T	1	Non-protein coding gene *CR44096*
3R:21,761,867	T	TA	1	Intron of *CG42335*
3R:21,778,495	C	T	1	D to N mutation in *tsl - torso-like*
3R:22,126,601	CAA	C	1	Intron of *CG42390*
3R:22,126,655	C	T	1	Intron of *CG42390*
3R:22,365,536				*pyk*
3R:23,338,231	GATCAATCC	G	1	Intron of *pnt - pointed*
3R:23,624,853	C	A	1	Intron of *CG31145*
3R:23,758,772	C	A	1	Change just 5' to the splice acceptor site in *Plip PTEN-like phosphatase*
3R:23,976,830	C	T	1	P to S mutation in CG31140
3R:24,064,912	GT	G	10	Present in all suppressors - Likely mutation in parental line
3R:24,074,990	AAAC	A,AAACAACAAC	8	Not unique
3R:24,143,789	G	A	10	Intron of *Kal1 - Kallmann syndrome 1* ortholog
3R:24,289,235	C	T	1	Intron of *CG34290*
3R:24,405,397	G	A	1	Gene Conversion
3R:24,405,476	G	C	1	Gene Conversion
3R:24,405,480	A	T	1	Gene Conversion
3R:24,406,491	T	C	1	Gene Conversion
3R:24,406,495	G	A	1	Gene Conversion
3R:24,429,080	C	T	1	Intron of *nAChRalpha1 - nicotinic Acetylcholine Receptor alpha 1*
3R:25,026,721	C	T	1	P to S mutation in *Hr96 - Hormone receptor-like in 96*

3R:25,636,244	C	T	3	Present in CA231, CR324, and CA1301
3R:25,636,576	G	T	2	No Gene - Present in both CA231 and CR324
3R:25,636,972	C	A	3	Present in CA231, CR324, and CA1301
3R:25,637,013	T	C	3	Present in CA231, CR324, and CA1301
3R:25,637,096	A	G	3	Present in CA231, CR324, and CA1301
3R:25,637,107	T	C	3	Present in CA231, CR324, and CA1301
3R:25,640,463	C	T	1	Intron of *jigr1 - jing interacting gene regulatory 1*
3R:25,654,871	C	T	1	Intron of *Tnks - tankyrase*
3R:26,065,868	A	T	1	No Gene
3R:26,077,363	T	A	1	3' UTR of the C isoform of *dys - dysfusion*, dorsal branch primordium, lateral trunk primordium
3R:26,146,548	C	T	1	No Gene
3R:26,331,111	C	T	3	Present in CA231, CR324, and CA1301
3R:26,646,617	T	A	1	3' UTR of *beat-VII*
3R:26,861,311	C	T	1	V to M mutation in *Lerp - lysosomal enzyme receptor protein*
3R:26,965,308	C	T	1	V to M mutation in *Gr97a - Gustatory receptor 97a*
3R:27,002,187	C	T	1	P to L mutation in *RYa-R - RYamide receptor* and *CG31077*
3R:27,701,050	C	T	1	P to S mutation in *btz - barentsz*: ovary expression
3R:27,720,385	C	T	1	No Gene
3R:28,942,071	C	T	1	Intron of *CG14521*
3R:28,984,878	G	A	3	Present in CA231, CR324, and CA1301
3R:28,984,981	G	A	3	Present in CA231, CR324, and CA1301
3R:28,985,044	CTACAA	C	3	Present in CA231, CR324, and CA1301
3R:28,985,059	A	G	3	Present in CA231, CR324, and CA1301
3R:28,985,073	T	TTA	3	Present in CA231, CR324, and CA1301
3R:28,985,142	T	G	3	Present in CA231, CR324, and CA1301
3R:28,985,160	C	G	3	Present in CA231, CR324, and CA1301
3R:29,211,562	C	T	1	No Gene
3R:29,313,388	A	T	10	Present in all suppressors - Likely mutation in parental line
3R:29,401,231	C	T	1	Intron of *Ptp99A - Protein tyrosine phosphatase 99A*
3R:29,461,090	T	TTGTTGGCTTCC CTGTC	6	Not unique
3R:29,497,481	C	T	1	Synonymous mutation in *CG2010*
3R:29,502,692	G	GTGAATCCTTTCCCA	10	Present in all suppressors - Likely mutation in parental line
3R:29,502,780	G	A	10	*Trc8*
3R:29,806,811				*claret*
3R:29,949,597	T	A	1	No Gene
3R:30,130,432	A	G	3	Present in CA231, CR324, and CA1301
3R:30,154,694	G	C	1	Intron of *CG11498*
3R:30,348,884	G	C	10	Present in all suppressors - Likely mutation in parental line
3R:30,418,395	A	T	1	Gene Conversion
3R:30,418,536	C	A	1	Gene Conversion
3R:30,418,563	CT	C	1	Gene Conversion
3R:30,418,657	G	A	1	Gene Conversion
3R:30,418,659	A	C	1	Gene Conversion
3R:30,418,718	A	C	1	Gene Conversion
3R:30,418,773	G	A	1	Gene Conversion
3R:30,418,783	G	T	1	Gene Conversion
3R:30,418,827	T	A	1	Gene Conversion
3R:30,418,830	G	C	1	Gene Conversion
3R:30,418,894	A	C	1	Gene Conversion
3R:30,418,934	A	G	1	Gene Conversion
3R:30,418,974	A	C	1	Gene Conversion
3R:30,418,978	G	T	1	Gene Conversion

3R:30,418,993	C	A	1	Gene Conversion
3R:30,418,996	A	G	1	Gene Conversion
3R:30,419,052	A	T	1	Gene Conversion
3R:30,419,147	C	T	1	Gene Conversion
3R:30,419,224	G	A	1	Gene Conversion
3R:30,419,254	C	A	1	Gene Conversion
3R:30,419,298	G	A	1	Gene Conversion
3R:30,419,339	T	A	1	Gene Conversion
3R:30,419,352	A	T	1	Gene Conversion
3R:30,419,379	G	A	1	Gene Conversion
3R:30,419,520	G	GAGTTTACT	1	Gene Conversion
3R:30,419,536	A	C	1	Gene Conversion
3R:30,419,538	A	T	1	Gene Conversion
3R:30,419,585	T	G	1	Gene Conversion
3R:30,419,591	G	T	1	Gene Conversion
3R:30,419,596	A	T	1	Gene Conversion
3R:30,419,618	CGCAGTATT	C	1	Gene Conversion
3R:30,772,825	C	T	1	Intron of *zfh1 - Zn finger homeodomain 1*
3R:30,821,141	G	C	3	Present in CA231, CR324, and CA1301
3R:31,581,796	C	T	1	A to V mutation in *rod - rough deal*
3R:32,009,958	C	T	1	No Gene
3L:389,226	T	A	1	Intron of *trh - trachealess*
3L:651,815	A	G	1	Gene Conversion
3L:652,018	G	T	1	Gene Conversion
3L:652,196	G	A	1	Gene Conversion
3L:732,600	AACACACACAC	A	3	Present in CA231, CR324, and CA1301
3L:732,815	G	A	3	Present in CA231, CR324, and CA1301
3L:842,822	C	T	1	Synonymous mutation in *Gr61a - Gustatory receptor 61a*
3L:856,955	C	G	1	5' UTR of *dpr20*
3L:1,111,216	C	T	1	No Gene
3L:1,124,870	C	T	1	No Gene
3L:1,301,103	C	G	1	3' UTR of *Rac1*
3L:1,356,499	C	T	1	Intron of *Ptp61F - Protein tyrosine phophatase 61F* - eye disc
3L:1,377,208	C	T	1	Intron of both *Ptp61F* and *ru - roughoid*
3L:1,457,055	AGCGGTACCGC CGAGTACAT	A	10	Present in all suppressors - Likely mutation in parental line
3L:1,457,295	A	G	10	Present in all suppressors - Likely mutation in parental line
3L:1,457,347	C	T	10	Intron of *Ptp61F - Protein tyrosine phosphatase 61F*
3L:1,457,368	T	A	10	Intron of *Ptp61F - Protein tyrosine phosphatase 61F*
3L:1,457,393	T	A	10	Intron of *Ptp61F - Protein tyrosine phosphatase 61F*
3L:1,457,426	TAGGG	T	9	Not unique
3L:1,457,431	T	TCGA	9	Not unique
3L:1,457,448	A	T	9	Not unique
3L:1,457,450	A	C	9	Not unique
3L:1,457,461	G	A	9	Intron of *Ptp61F - Protein tyrosine phosphatase 61F*
3L:1,457,470	G	T	9	Not unique
3L:1,457,491	C	T	9	Intron of *Ptp61F - Protein tyrosine phosphatase 61F*
3L:1,457,496	G	C	9	Not unique
3L:1,457,647	G	A	10	Intron of *Ptp61F - Protein tyrosine phosphatase 61F*
3L:1,457,672	GGAGGACA	G	10	Present in all suppressors - Likely mutation in parental line
3L:1,814,981	C	T	1	Intron of *GV1*
3L:1,865,031	C	T	1	Synonymous mutation in *HBS1*
3L:2,014,614	C	T	1	No Gene
3L:2,298,471	G	A	1	W to Stop mutation in *MsR2 - Myosuppressin receptor 2*

3L:2,444,707	T	A	1	Synonymous mutation in *CG42669*
3L:2,605,434	T	G	1	Intron of *Pxn - Peroxidasin*
3L:3,386,561	C	A	1	No Gene
3L:3,386,580	G	GGTGGAT	1	No Gene
3L:3,433,922	A	C	10	Present in all suppressors - Likely mutation in parental line
3L:3,435,023	A	G	10	Present in all suppressors - Likely mutation in parental line
3L:3,435,111	C	T	10	intron of *eIF5B*
3L:3,435,114	TG	T	10	intron of *eIF5B*
3L:3,435,123	A	T	10	intron of *eIF5B*
3L:3,435,213	C	T	10	intron of *eIF5B*
3L:3,554,226	C	T	1	Intron of *Eip63E - Ecdysone-induced protein 63E*
3L:4,890,769	C	T	1	Gene Conversion
3L:4,890,780	A	C	1	Gene Conversion
3L:4,890,837	C	CA	1	Gene Conversion
3L:4,890,913	C	T	1	Gene Conversion
3L:4,890,922	C	T	1	Gene Conversion
3L:4,890,954	T	A	1	Gene Conversion
3L:4,891,001	AT	A	1	Gene Conversion
3L:4,891,023	C	CCA	1	Gene Conversion
3L:4,891,026	T	TAGATATGTCTT TTGAACTTAC	1	Gene Conversion
3L:4,920,263	G	C	1	No Gene
3L:5,445,464	C	T	1	No Gene
3L:5,766,599	C	T	1	S to F mutation in *Txl - Thioredoxin-like*
3L:5,830,263	C	T	1	Intron of *vn - vein*
3L:5,924,738	A	T	1	Intron of *CG33523*
3L:5,975,467	C	T	1	Intron of *CG10479*
3L:6,126,563	G	A	1	Intron of *l(3)mbn - lethal (3) malignant blood neoplasm*
3L:6,126,571	T	A	1	Intron of *l(3)mbn - lethal (3) malignant blood neoplasm*
3L:6,127,729	C	CAA	1	Intron of *l(3)mbn - lethal (3) malignant blood neoplasm*
3L:6,485,899	C	T	1	3' UTR of *CG42747*
3L:6,842,147	C	T	1	No Gene
3L:7,042,349	T	A	9	Not unique
3L:7,042,352	G	C	9	Not unique
3L:7,285,872	C	T	1	Intron of *unc-13-4A*
3L:8,100,880	C	T	1	Intron of *Ect4 - Ectoderm-expressed 4*
3L:8,101,582	C	T	1	Intron of *Ect4 - Ectoderm-expressed 4*
3L:8,156,858	G	A	3	Present in CA231, CR324, and CA1301
3L:8,657,594	T	A	9	Not unique
3L:8,657,595	A	C	9	Not unique
3L:8,657,598	T	C	9	Not unique
3L:8,657,603	T	C	9	Not unique
3L:8,657,612	A	T	9	Not unique
3L:8,657,615	C	G	9	Not unique
3L:8,657,645	C	A	9	Not unique
3L:8,657,670	G	A	9	No Gene
3L:8,657,676	G	C	9	No Gene
3L:8,657,771	C	T	9	No Gene
3L:8,764,575	AT	A	9	Not unique
3L:8,765,054	G	A	9	Intron of *Fhos - Formin homology 2 domain containing ortholog*
3L:8,765,055	G	A	9	Intron of *Fhos - Formin homology 2 domain containing ortholog*
3L:8,765,287	C	T	9	Intron of *Fhos - Formin homology 2 domain containing ortholog*
3L:8,765,435	C	G	9	Not unique
3L:8,805,397	C	T	1	Synonymous mutation in *CG34427*

3L:8,915,252	C	T	1	Intron of *Tsp66E - Tetraspanin 66E /// ncRNA of CR43970*
3L:8,915,875	C	T	1	Intron of *Tsp66E - Tetraspanin 66E*
3L:9,940,273	G	T	9	Not unique
3L:9,940,333	C	A	9	Not unique
3L:9,940,657	A	T	9	Not unique
3L:10,090,955	C	A	1	Intron of *dpr6* /// E to D mutation in *CG32053*
3L:10,097,916	C	T	1	Intron of *dpr6* /// G to E in variant 2 *CG42825*
3L:10,644,304	G	A	10	Present in all suppressors - Likely mutation in parental line
3L:10,644,389	C	G	10	Present in all suppressors - Likely mutation in parental line
3L:10,644,390	A	C	10	Present in all suppressors - Likely mutation in parental line
3L:10,644,419	T	A	10	Present in all suppressors - Likely mutation in parental line
3L:10,743,289	A	G	3	Present in CA231, CR324, and CA1301
3L:10,743,380	TAC	T	3	Present in CA231, CR324, and CA1301
3L:10,743,614	T	TTACA	3	Present in CA231, CR324, and CA1301
3L:10,952,299	A	AATAT	3	Present in CA231, CR324, and CA1301
3L:10,952,554	G	T	3	Present in CA231, CR324, and CA1301
3L:10,984,445	C	T	1	D to N mutation in *klu - klumpfuss*
3L:11,474,764	T	A	3	Present in CA231, CR324, and CA1301
3L:11,474,782	A	G	3	Present in CA231, CR324, and CA1301
3L:11,474,799	A	T	3	Present in CA231, CR324, and CA1301
3L:11,474,807	A	AAAT	3	Present in CA231, CR324, and CA1301
3L:11,474,828	A	C	3	Present in CA231, CR324, and CA1301
3L:11,474,867	A	T	3	Present in CA231, CR324, and CA1301
3L:11,474,878	T	A	3	Present in CA231, CR324, and CA1301
3L:11,474,969	T	A	3	Present in CA231, CR324, and CA1301
3L:11,474,997	A	G	3	Present in CA231, CR324, and CA1301
3L:11,814,309	C	T	1	Synonymous mutation in *CG42255*
3L:12,028,552	C	T	1	Intron of *rols - rolling pebbles*: late extended germ band embryo
3L:12,071,267	C	T	1	Intron of *Sema-5c - Semaphorin-5c*
3L:12,456,190	ATTATTACATT	A	9	Not unique
3L:12,456,630	GGA	G	10	Present in all suppressors - Likely mutation in parental line
3L:12,456,633	CTGGGG	C	10	Present in all suppressors - Likely mutation in parental line
3L:12,456,662	C	T	10	No Gene
3L:12,456,710	GTGA	G	10	Present in all suppressors - Likely mutation in parental line
3L:12,457,136	A	AAGCCG	10	Present in all suppressors - Likely mutation in parental line
3L:13,039,402	C	T	1	T to I mutation in intron 3 bps away from splice spot in *SRm160*
3L:13,242,945	A	AG	10	Present in all suppressors - Likely mutation in parental line
3L:13,243,020	C	G	10	Present in all suppressors - Likely mutation in parental line
3L:13,243,290	A	ATTT	10	Present in all suppressors - Likely mutation in parental line
3L:13,243,367	A	G	10	Present in all suppressors - Likely mutation in parental line
3L:13,243,386	T	TTTTTTTG	10	Present in all suppressors - Likely mutation in parental line
3L:13,243,400	T	A	10	Present in all suppressors - Likely mutation in parental line
3L:13,243,672	C	T	10	Intron of caps - capricious
3L:13,290,027	T	A	3	Present in CA231, CR324, and CA1301
3L:13,781,477	C	T	1	Intron of *bru-3 - bruno-3* - negative regulation of translation
3L:13,828,367	C	T	1	No Gene
3L:13,865,050	C	A	1	Intron of *CG8745*
3L:14,179,044	C	T	1	No Gene
3L:15,014,905	AAG	A	1	3' UTR of *Sytbeta - Synaptotagmin beta*
3L:15,014,909	A	G	1	3' UTR of *Sytbeta - Synaptotagmin beta*
3L:15,165,691	G	C	10	Present in all suppressors - Likely mutation in parental line
3L:15,166,607	GTTTGGGGGTTT	G	10	Present in all suppressors - Likely mutation in parental line
3L:15,166,730	G	C	10	Present in all suppressors - Likely mutation in parental line
3L:15,166,956	A	C	10	Present in all suppressors - Likely mutation in parental line

3L:15,166,965	T	A	10	Present in all suppressors - Likely mutation in parental line
3L:15,167,047	C	G	10	Present in all suppressors - Likely mutation in parental line
3L:15,685,630	C	T	1	Synonymous mutation in *pgant8 - polypeptide GaINAc transferase 8*
3L:15,902,455	C	T	1	3' UTR of *sff - sugar-free frosting*
3L:15,966,999	C	T	1	D to N mutation in *l(3)72Ab - lethal (3)72Ab - helicase activity*
3L:16,624,385	C	T	1	Synonymous mutation in *Abl - Abl tyrosine kinase*
3L:16,962,482	C	T	1	Synonymous mutation in *Nc73EF - Neural conserved at 73EF*
3L:17,105,479	C	T	1	Intron of *Rbp6 - RNA-binding protein 6* - embryonic
3L:17,306,940	C	T	1	No Gene
3L:17,435,941	C	T	1	Intron of *CG32176* /// V to M mutation in *CG12229*
3L:17,533,202	T	A	1	Intron of *Nedd4*
3L:18,654,245	C	T	1	Intron of *MYPT-75D*
3L:18,988,776	G	A	10	3' UTR of *CG34256*
3L:19,782,370	C	T	1	Intron of *Gyc76C* /// Intron of *CG42637*
3L:19,902,335	C	T	1	Intron of *Su(Tpl)* /// Intron of *Mi-2*
3L:20,070,396	G	A	10	Intron of *sNPF-R - short neuropeptide F receptor*
3L:20,350,702	C	T	1	L to F mutation in *eRF1 - eukaryotic release factor 1*
3L:20,402,032	A	G	3	Present in CA231, CR324, and CA1301
3L:20,645,321	A	G	9	Not unique
3L:20,647,359	A	G	10	Present in all suppressors - Likely mutation in parental line
3L:20,647,364	A	G	10	Present in all suppressors - Likely mutation in parental line
3L:20,749,086	C	T	1	No Gene
3L:20,972,994	C	T	1	Synonymous mutation in *CG33286*
3L:21,187,182	T	C	1	Intron of *CG42337*
3L:21,982,217	C	T	1	P to L mutation in *CG7470*
3L:22,934,610	C	T	1	No Gene
3L:22,985,792	GT	G	1	No Gene
3L:23,228,447	C	T	1	Intron of *nACHRalpha4 - nicotinic Acetylcholine Receptor alpha 4*
3L:23,511,783	C	T	1	No Gene - repeat region
3L:23,964,636	G	C	10	Present in all suppressors - Likely mutation in parental line
3L:23,964,931	A	C	10	Present in all suppressors - Likely mutation in parental line
3L:23,964,972	T	A	10	Present in all suppressors - Likely mutation in parental line
3L:23,965,219	T	A	10	Present in all suppressors - Likely mutation in parental line
3L:23,965,332	A	G	10	Present in all suppressors - Likely mutation in parental line
3L:23,965,594	CAA	C	10	Present in all suppressors - Likely mutation in parental line
3L:23,966,116	C	T	10	Intron of *CG40470*
3L:24,070,026	T	TG	1	No Gene - repeat region
3L:24,070,028	C	CA	1	No Gene - repeat region
3L:24,079,670	C	T	1	Intron of *Snap25 - Synapse protein 25*