Spring 2018

Genetic Basis of Larval Crystal Cell Quantity Variation in the Drosophila Genetic Reference Panel (DGRP)

Brian Tang
Bernard M Baruch College & CUNYBA Program

Brian Tang

Genetic Basis of Larval Crystal Cell Quantity Variation in the *Drosophila* Genetic Reference Panel (DGRP)

April 30th, 2018

Submitted to the Committee on Undergraduate Honors at Baruch College of the City University of New York in partial fulfillment of the requirements for the degree of Bachelor of Arts from Baruch & CUNY Baccalaureate for Unique and Interdisciplinary Studies in Behavioral Neuroscience and Biochemistry.

Signature of Faculty Sponsor

Rebecca F Spokony

Dr. Rebecca F. Spokony
Assistant Professor of Biology
Baruch College Natural Science Department

Signatures of full-time Baruch faculty members who have read and approved this thesis:

Helene Barbour

Dr. Helene Eisenman Barbour
Assistant Professor of Biology
Baruch College Natural Science Department

Krista C Dobi

Dr. Krista Dobi
Assistant Professor of Biology
Baruch College Natural Science Department

Table of Contents

Abstract

Crystal cells are one of three requisite hemocytes that take part in fighting infection and wound healing in *Drosophila melanogaster* (common fruit flies). The developmental genetics of crystal cell formation is only beginning to be discovered. To address this question, we performed a Genome-Wide Association Study (GWAS) on larval crystal cell number from 78 isolines of the *Drosophila* Genetic Reference Panel (DGRP) collection. The DGRP consists of naturally caught fruit flies that are inbred to near homozygosity with completely sequenced genomes. By placing the wandering third instar larvae under heatshock, a process that induces the melanization of crystal cells, it was possible to manually score the number of black crystal cells throughout the larvae. We discovered that the average number of crystal cells in each DGRP line (10 larvae per line) varied, ranging from 0 to 730. The average of all the crystal cells counted (all DGRP lines collectively) was calculated to be 220. Also, of the 78 DGRP lines, 75 of them had more crystal cells when compared to the Oregon-R control line (with a mean of 20 crystal cells per larvae). GWAS of the DGRP crystal cell count data found 128 polymorphisms ($p<10^{-5}$) that may be associated with differences in crystal cell number between the lines. From the list of polymorphisms, we chose to test 10 genes (with smaller p-values) mapped to the polymorphisms. However, overexpression was done for 4 of the 10 genes. Overexpressing 3 of the 4 genes led to altered crystal cell number. In total, we have found 3 new genes (*domino*, *extra-extra*, and *hemese*) involved in crystal cell development.

Objectives

The *Drosophila* innate immune system consists of different myeloid-like cells that are similar to human monocytes, macrophages, neutrophils, basophils, eosinophils, erythrocytes, and megakaryocytes [1]. These myeloid-like cells are present throughout the fruit flies' development. Of the different myeloid-like cells, three major hemocytes known as plasmatocytes, lamellocytes and crystal cells contribute to the processes of phagocytosis, encapsulation, and melanization, respectively, within *Drosophila* [2]. Crystal cells in particular, melanize upon injury, microbial infections, and parasitization.

In our study, one of the focuses was to determine if larvae of different genotypes vary in the number of crystal cells they develop. The fruit fly immune system is very much similar to that of the human immune system, which has a spectrum ranging from a poor immune system to an overly active immune system. The extreme ends of the spectrum cause to an imbalance in immunity, inflammation, and tissue regeneration (related to macrophages), which has been determined to be one of many causes of tumorigenesis in various organisms [3]. As a myeloid-like cell is similar to that of human macrophages, crystal cell production imbalance may also be one of many causes of tumorigenesis in fruit flies.

Due to flies being easy to maintain, reproducing large number of offspring, and having many analogous genes to humans, they serve as good models to conduct experiments. By using flies from lines that belong to the DGRP fly collection, larval crystal cell count data was obtained. Next, we wanted to determine whether there is any association between genetic variation between the DGRP flies and crystal cell variation. We then wanted to know whether these genes increase or decrease crystal cell number. Finally, we developed a model of where these genes may act within the established crystal cell developmental pathway.

Introduction

Hematopoiesis in *Drosophila melanogaster*

Unlike the immune system of the human body, *Drosophila melanogaster*, also known as the common fruit fly, has only an innate immune system but no adaptive immune system [4]. The *Drosophila's* innate immune system fulfills the role of aiding the survival of the organism against infection and invaders. The innate immune system consists of three types of hemocytes that contribute to their immunity throughout *Drosophila* development and further into adulthood. These hemocyte cells differentiate from cells known as prohemocytes, which originate from the mesoderm [5].

Drosophila prohemocytes, a collection of hemocyte precursors, are stem cells that are capable of giving rise to different types of cells, therefore *Drosophila* prohemocytes are considered to be pluripotent [5, 6]. The three hemocytes are plasmatocytes, crystal cells, and lamellocytes. The process in which prohemocytes differentiate into hemocytes is called hematopoiesis, which takes place in the lymph gland and sessile pockets throughout larval development [6, 7]. A gene, *Serpent* (*Srp*), codes for GATA Serpent transcription factors (Srp) that are required for production of prohemocytes [8]. It has also been shown to contribute to the formation of fat bodies and insect organs that are similar to the human liver, in *Drosophila* [9].

Hematopoiesis occurs in two phases. The first phase of hematopoiesis occurs in the early embryonic stage (indicated by purple box in Figure 1) where prohemocytes originate from the mesoderm on the head and differentiate into two cell types known as crystal cells and plasmatocytes. As the first phase comes to a close, a specialized organ known as the lymph gland begins to develop along the dorsal vessel (represented by the orange structure in Figure 1). The second phase of hematopoiesis occurs in the larval stage of development in the lymph gland,

which gives rise to adult hemocytes (indicated by yellow box Figure 1) [7]. The lymph gland becomes fully matured during larval development. Even though hematopoiesis occurs in two phases, both hemocytes that are produced as embryos and those produced as larvae exist in adult flies (stage 5 in Figure 1) [5]. The pupae and adults, however, do not produce hemocytes.

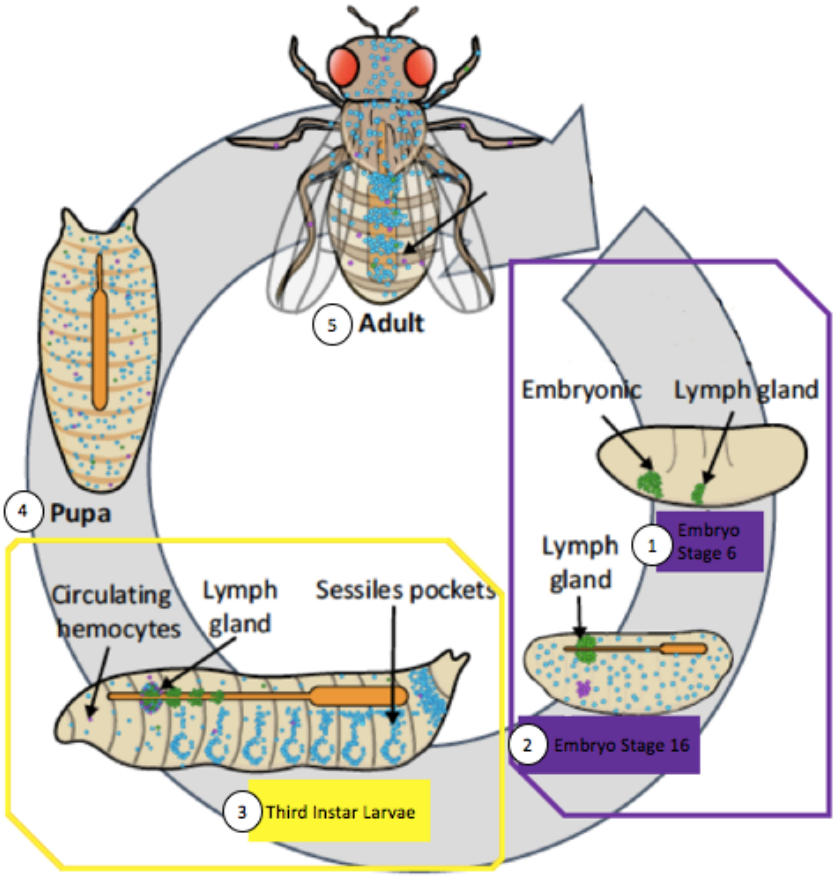

***Figure 1 Drosophila* Life Cycle** The diagram above illustrates the life cycle of a fruit fly, *Drosophila melanogaster*. Numbers 1 and 2 (purple box) on the diagram represent stages 6 and 16 respectively, during embryonic development. Number 3 (yellow box) shows the 3rd instar larva stage of development. Number 4 represents the pupa stage of development. Last but not least, number 5 represents flies at adulthood. The colors and what it symbolizes are as follows: green dots (embryonic lymph glands), purple dots (plasmatocytes), blue dots (crystal cells). The orange is the cardiac/tube dorsal vessel. *Modified version from Letourneau et al., 2016* [2].

Hemocytes: Plasmatocytes, Crystal Cells, and Lamellocytes

Plasmatocytes (green), crystal cells (blue), and lamellocytes (orange) make up the hemocyte population shown in Figure 2. The hemocytes each play similar and different roles in

the organism such as assisting in wound healing, protection against bacterial infection, and engulfment of parasitoid eggs [2]. Both plasmatocytes and crystal cells are usually present in the *Drosophila* hemolymph, but lamellocytes differentiate from either prohemocytes or plasmatocytes only when the immune system is challenged. Under normal conditions without infections, 95% of hemocytes exist as plasmatocytes and only 5% exist as crystal cells. Crystal cells majorly contribute to encapsulation and melanization during immune reactions and wound healing processes at the site of infection and/or injury [10]. Crystal cells particularly are named for the paracystalline inclusions, structures found in the cell's cytoplasm that contain large amounts of molecules involved in a melanization [3].

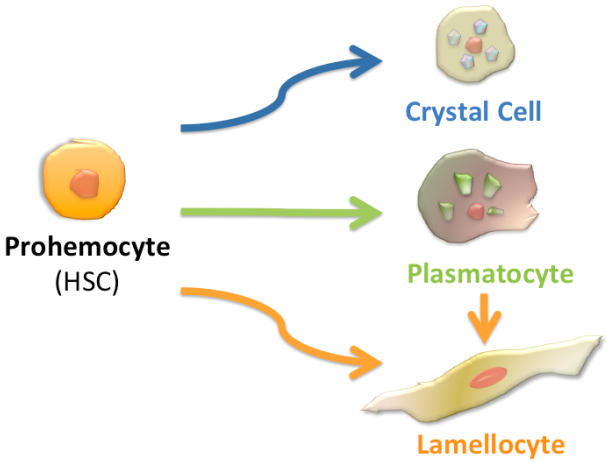

***Figure 2 Drosophila* Hemocytes** The *Drosophila* has no adaptive immune system, but do have an innate immune system. The immune system consists of different myeloid-like cells: lamellocytes (orange), plasmatocytes (green), and crystal cells (blue). The three hemocytes differentiate from prohemocytes, but lamellocytes are not constantly present in the hemolymph. Lamellocytes only differentiate from either prohemocytes or plasmatocytes immune system is challenged. Lamellocytes contribute to encapsulation, plasmatocytes to phagocytosis, and crystal cells to both melanization and encapsulation.

Molecular Pathway: Locking Prohemocytes into Crystal Cell Fate

Hematopoiesis is the general process in which prohemocytes differentiate into either of the following hemocytes: plasmatocytes, lamellocytes, or crystal cells. However, specific

molecular pathways are required to induce the differentiation of prohemocytes into a specific hemocyte. A summary of the known molecular pathway involving Serpent, Serrate (Jagged-1 in mammals) Notch Ligand, Notch, RUNX Lozenge, U-shaped protein, SerpentNC, Serpin 27A, Serine protease, and PPO1/PPO2 to induce the differentiation of prohemocytes into crystal cells is portrayed in Figure 3 below.

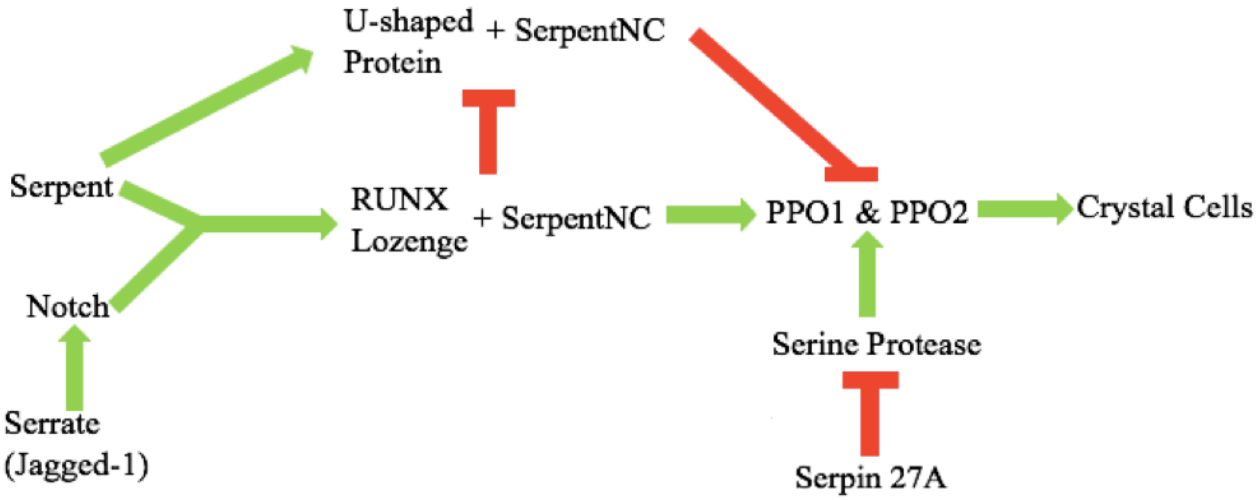

Figure 3 **Summary of Known Molecular Crystal Cell Pathway** The known molecular pathway involved in crystal cell formation consists of Serpent, Serrate (Jagged-1) Notch Ligand, Notch, RUNX Lozenge, U-shaped protein, SerpentNC, Serpin 27A, Serine protease, and PPO1/PPO2 [8-17]. This molecular pathways was compiled using information found from previous literature. The green arrows indicate processes that upregulate crystal cell formation. The red T-shaped structures indicate processes that inhibit and downregulate crystal cell production.

Crystal Cell Differentiation: Serpent, Notch Signaling, and Lozenge

Serpent (Srp) and its isoform SerpentNC (SrpNC), that consists of zinc ions (zinc fingers) present at and stabilizes both its N and C terminus, contribute to crystal cell differentiation from prohemocytes [8, 11, 12]. Increase in crystal cell differentiation is upregulated by interaction of Serpent with the Notch pathway. Though two ligands exist for Notch signaling, one being Delta and the other being Serrate, Serrate is the specific ligand that contributes to crystal cell

formation. The interaction of Serpent with the Notch pathway is important for crystal cell fate due to the binding of two proteins that their interaction activates. Once RUNX Lozenge and SrpNC binding is activated by the interaction between Serpent and Notch, prohemocytes are locked into crystal cell fate [13]. However, without the Notch pathway, Serpent alone would actually inhibit crystal cell formation. In fact, dysfunctional Notch proteins lead to a significant decrease in crystal cell quantity and an overexpression of Notch proteins lead to an increase in crystal cell quantity [14].

Crystal Cell Differentiation: Serpent, SerpentNC, Lozenge, and U-shaped Protein

SrpNC plays both the role of upregulation and downregulation of crystal cell production depending on whether it binds to RUNX Lozenge factor or GATA U-shaped protein. Interaction between Serpent and Notch leads to binding of SrpNC to Lozenge, activating crystal cell production that occurs downstream of the Notch molecular pathway [15]. The direct downstream effect of Lozenge and SrpNC complex is decreased availability of U-shaped Protein and inability of it to bind to SrpNC. On the contrary, without Notch, Serpent alone would promote the binding of SrpNC to U-shaped protein, and when SerpentNC binds to the GATA U-shaped protein, crystal cell production is inhibited [15]. SrpNC factors that cannot bind to U-shaped proteins are not able to inhibit the production of crystal cells, once again indicating that the Serrate-induced Notch is particularly required for crystal cell production [8].

Crystal Cell Differentiation: Serine Protease and Prophenoloxidases (PPOs) 1 & 2

Crystal cells have been found to produce two prophenoloxidases, PPO1 and PPO2, both of which are released from ruptured crystal cells into the hemolymph (blood in invertebrates)

upon injury [16]. However, the two prophenoloxidases are not always active and are regulated to prevent excessive amounts of melanization [17]. The PPOs are created as zymogens, which are inactive enzymes until activated by another enzyme. In the *Drosophila* hemolymph, the PPOs are activated by serine proteases to become their active form known as phenoloxidases (POs) [16]. In turn, phenoloxidases catalyze the production of quinones, which are precursors for the production of melanin polymers [16]. Simultaneously, the phenoloxidase produces highly unstable and reactive oxygen species that have been speculated to kill off infections [16, 18]. The process of melanization is not limited to fruit flies. Interestingly, prophenoloxidase-activating enzymes from different insect species and arthropods have been found to exist, but are activated by serine proteases that differ from that of the ones in *Drosophila* [17]. To simulate the reaction of crystal cells to infection and injury, previous research has shown that under heatshock conditions, melanization of crystal cells occurs similarly to when caused by injury or infection [19].

Drosophila with PPO1 and/or PPO2 deletions have shown that both prophenoloxidases contribute to the role that crystal cells play in the immune system in different ways. Another prophenoloxidase, PPO3, is not found in crystal cells but is produced in lamellocytes and plays a role in capsule formation [16]. A previous study on PPO1 and PPO2 mutants found that PPO2, not PPO1, are absolutely required for the melanization of crystal cells [16]. However, both PPOs are required for normal melanization of crystal cells as seen in wild type larvae. Larvae of PPO2 mutants and PPO1,PPO2 double mutants showed no crystal cell melanization after heating. Larvae of PPO1 mutants portrayed slower melanization of crystal cells but was similar to that of the wild type larvae [16]

Crystal Cell Differentiation: Serpin 27A, PPO1, and PPO2

Though melanization provides protection against infections and parasites, regulation of PPO1 and PPO2 are essential to prevent unnecessary melanization by crystal cells. PPO1 and PPO2 are released as zymogens and are activated upon the hydrolysis of their peptide bonds by serine proteases, to become their active form of phenoloxidases (POs). Serpin-27A, a serine protease inhibitor protein, regulates the activity of PPO1 and PPO2 within the Toll/Imd pathway [17, 20]. As a negative regulator for phenoloxidases, Serpin-27A significantly decreases the ability of *Drosophila* larvae to produce melanotic capsules [20]. Unlike PPO1 and PPO2 mutations, double mutations of *Serpin-27A* that produced no Serpin-27A proteins cause lethality in homozygous *Drosophila* larvae [17]. However, mutation of the Serpin-27A gene that codes for dysfunctional Serpin-27A protein, produced flies that are only deficient (not completely lack) of Serpin-27A protein [17]. These animals exhibited spontaneous melanization of crystals around the larval body and an overall decreased survivability as adult flies [17].

PPO Mutants and Serpin 27A Mutants Against Infections

Previous research has tested the susceptibility of PPO1,PPO2 double mutant adult flies and their wild types to *Erwinia carotovora*, *Salmonella typhimurium*, *Listeria monocytogenesis*, *Enterococcus faecalis, Staphylococcus aureus*, and *Candida albicans* infection [16]. Due to PPO1,PPO2 double mutant's inability to produce melanization, their ability to prevent the colony growth of these infections after exposure was expected to be compromised. However, the expected outcome was not fully obtained. Over the course of four days post-infection, the colony forming unit (CFU) data of each type of infection per fly in both mutant and wild type groups was collected daily. Among the CFU collected for each infection, only the CFU difference between wild types and mutants for *S. aureus* significantly differed after four days of exposure

[16]. The CFU differences over the course of four days for *E. carotovora*, *S. typhimurium*, *L. monocytogenesis*, *E. faecalis*, and *C. albicans* were determined to be insignificant [16]. This is indicative of mechanisms other than PPOs that may have contributed the variation in vulnerability of PPO1,PPO2 double mutants adult flies to the different types of infection.

Additionally, susceptibility of null Serpin-27A mutant adult flies to *Escherichia coli*, *Micrococcus Luteus*, *Aspergillus fumigatus*, and *Beauveria bassiana* infections was obtained in terms of adult fly survival rate. Both null Serpin-27A mutants and wild types were infected and the percentage of adult flies that survived eight days after infection were calculated [17]. All null Serpin-27A mutant adults had lower survival rates in comparison to their wild type. Though these mutants were expected to have increased adult survival rate against infection due to upregulation of crystal cell formation, the lower adult survival rates can be attributed to the fixed amount of hemocytes the organism has. Without functional Serpin-27A, majority of the crystal cells are used in the larval stages of development, leading to a depleted supply of crystal cells for adult flies to fight against infection. As for the comparison of null Serpin-27A mutant adults' survival rate in relation to different types of infection, the results are as follows from least to greatest survival rate: *B. bassiana, M. Luteus, A. fumigatus,* and *E. coli* [17]. The wild type adults, though, had a different rate of survival pattern to each infection. Once again, there is indication of other mechanisms that may have contributed to the differences in adult survival rate and immunity within both null Serpin-27A mutant and wild type groups.

Introduction to GWAS

Genome-Wide Association Study (GWAS) is a statistical analysis of a genome-wide set of genetic variants in different individuals to see if any variant is associated with a trait. This

method involves searching of an organisms' genome for polymorphisms, small variations in the genome, that appear in higher frequency within the genome of individuals who have a particular trait [21]. For humans, GWAS can be used to discover genetic variants that correlate with diseases. When associations between genes and diseases have been made, scientists can utilize the information to invent more efficient methods to determine risks of developing the gene-related disease, as well as methods of treatment. For this study, the organism studied was *Drosophila melanogaster* from the DGRP collection (flies with fully sequenced genomes) and the trait of study is the quantity of crystal cells in third instar larvae.

Introduction to DGRP Immunity Studies

The flies which are used in this study are flies from the *Drosophila* Genetic Reference Panel (DGRP) collection. Several studies of *Drosophila* immunity have been done with the DGRP collection [22, 23]. However, these immunity studies utilized adult flies rather than larvae. Variation in mean survival times was detected after 188 DGRP lines were exposed to *Metarhizium anisopliae* (a fungal pathogen) or *Pseudomonas aeruginosa* (bacterium). The variation was attributed to genetic differences impacting the ability of each DGRP line to adapt to pathogen exposures [22]. Interestingly, in addition to observing difference in immunity of each DGRP line, comparison of immunity based on sex of the flies within each line showed that males demonstrated higher resistance and survival rate than females for most lines. 13 candidate genes (with polymorphisms found through GWAS analysis) that were associated with immunity of the flies to *M. anisopliae* and *P. aeruginosa* were functionally tested to observe if mutations caused by p-element insertions into those genes led to significant differences in resistance and

survival rate of the adult flies [22]. The genes were: *S*, *msn*, *Shn*, *CG33172*, *tai*, *sik3*, *Rdl*, *f*, *CG9990*, *CG32066*, *CG33111*, *puf*, and *FOXO*.

Exposure of DGRP adult flies from 172 lines to bacterium *Providencia rettgeri* investigated the ability of flies to limit the negative consequences of infection (also known as tolerance) and how it impact their survival at a given level of infection intensity [23]. Researchers looked to identify genes that contribute to tolerance of adult DGRP flies to *P. rettgeri* infection. Through a GWAS analysis, they identified gene associated polymorphisms that contributed to variation in tolerance between the DGRP lines. Using the GWAS results, they conducted functional validation experiements for 10 candidate genes (RNAi knock down) to confirm that the candidate genes impacted variation in tolerance [23]. The genes were: *Rbp9*, *mspo*, *fhos*, *CG4174*, *gus*, *beat-IIIc*, *u-shaped (ush)*, *grainyhead (grh)*, *debris buster (dsb)*, and *CG30098*. The variation in pathogen susceptibility may be related to differences in hemocyte composition between the DGRP lines.

Materials and Methods

Drosophila melanogaster stocks

Seventy-eight isolines of the *Drosophila* Genetic Reference Panel (DGRP), mutants, cDNA, and RNAi flies were purchased from the Bloomington Stock Center to use in this study (Table 1). The *Drosophila* Genetic Reference Panel (DGRP) collection consists of naturally caught fruit flies that are inbred to near homozygosity [24]. Their fully sequenced genomes allow for the identification of genes/polymorphisms, through the Genome-Wide Association Study (GWAS), that are relevant to the phenotype of crystal cell quantity being studied. Using the polymorphisms provided by the GWAS analysis, mutant flies, flies with the cDNA constructs (overexpression), and flies with RNAi (inhibition/under expression) of the polymorphism were studied. All flies were quarantined before use in this study.

Table 1A

Bloomington Stock #	DGRP Genotypes	Bloomington Stock #	DGRP Genotypes	Bloomington Stock #	DGRP Genotypes	Bloomington Stock #	DGRP Genotypes
25174	208	28144	142	28213	589	28249	850
25175	301	28145	149	28215	595	28250	853
25186	360	28146	153	28217	646	28251	855
25187	362	28148	161	28218	703	28255	882
25190	380	28154	217	28219	716	28258	892
25191	391	28160	237	28220	721	28260	897
25193	427	28161	239	28223	738	28261	900
25201	712	28165	287	28224	748	28262	907
25203	732	28166	309	28227	761	28263	908
25206	786	28167	317	28229	776	28265	913
25208	820	28173	338	28230	783	28278	409

25210	859	28176	350	28231	787	29652	57
25211	Oregon-R	28178	356	28232	790	29658	439
25745	714	28179	359	28233	796	55014	31
28122	21	28183	371	28235	802	55019	348
28123	26	28189	382	28237	805	55022	395
28128	45	28191	385	28239	810	55023	397
28129	59	28204	502	28240	812	55030	627
28138	101	28208	535	28244	822	55031	630
28140	109	28212	584	28247	843	28249	850

Table 1B

Bloomington Stock #	**Mutant Genotypes**
367	**pk**[1] cn[1]
5806	**prd**[8]/CyO
5370	w[*]; **htl**[AB42]/TM3, P{ry[+t7.2]=ftz/lacC}SC1, ry[RK] Sb[1] Ser[1]
9260	y[1] w[*]; **dom**[3]/SM6a
36980	y[1] w[*]; Mi{y[+mDint2]=MIC}**CG4390**[MI03759]/TM3, Sb[1] Ser[1]
34149	y[1] w[*]; Mi{y[+mDint2]=MIC}Indy-2[MI01115] **CG17193**[MI01115] CG33934[MI01115]
9930	w[*]; **exex**[KK30] e[s]/TM3, P{ry[+t7.2]=ftz-lacZ.ry[+]}TM3, Sb[1] ry[*]
11707	P{ry[+t7.2]=PZ}**msn**[06946] ry[506]/TM3, ry[RK] Sb[1] Ser[1]

Table 1C	
Bloomington Stock #	**cDNA Genotypes**
5419	*heartless*
5867	*PPO1* (chromosome 2)
5868	*PPO1* (chromosome 3)
8697	*hemese*
9929	*extra-extra*
64261	*domino*

Table 1D	
Bloomington Stock #	**RNAi Genotypes**
38243	*echinoid*
58134	*PPO1*
58289	*heartless*

***Table 1A, 1B, 1C, 1D* Bloomington Stock Number of Flies Used** The genes of all of the flies used in this study and their corresponding Bloomington stock numbers are shown. The tables consists of Bloomington stock numbers for DGRP lines (Table 1A), mutant lines (Table 1B), and cDNA lines (Table 1C), RNAi lines (Table 1D).

Scoring of Crystal Cells for DGRP and Mutant Flies

Scoring for both DGRP and Mutant Flies

Ten late third instar larvae (five male and five female) from each of the genotypes/isolines were collected from food vials and placed into micro centrifuge tubes labeled with the proper genotype and sex of the larvae (Table 1A and 1B). The microcentrifuge tubes were then placed into a thermal cycler that has been set at a steady temperature of 70^{o}C for the heat shocking process. Timers were set for ten minutes to ensure consistent amount of heat shock in a constant amount of time for all genotypes. After ten minutes, micro centrifuge tubes were removed from the thermal cycler and the male larvae within, were placed onto a petri dish with wet filter paper (repeated process for female larvae of the same genotype). Utilizing microscopes, the number of crystal cells were observed at higher visual quality and counted for each individual larvae (hand tally counters were used to maintain count of crystal cells). Cameras attached and synced with the microscopes documented photos of individual larvae that were scored for crystal cell number.

Genome-Wide Association Study

Genome-Wide Association Study (GWAS) is a statistical analysis of a genome-wide set of genetic variants in different individuals to see if any variant is associated with a trait. It utilizes the experimental data collected and analyzes it by identifying possible polymorphisms that may be associated with a trait/phenotype. Upon collection of crystal cell count (trait) data from 78 DGRP isolines, the Mackay lab GWAS pipeline was utilized [24]. GWAS analysis provided polymorphisms that may be correlated with variation in crystal cell count between isolines. Selected polymorphism, from the list of possible polymorphism, were used for the purchase of mutant, cDNA, and RNAi flies for functional testing.

Scoring of Crystal Cells for cDNA Flies

Scoring of UAS-cDNA and UAS/Gal4-cDNA

The crystal cell counting procedure for UAS-cDNA genotypes was the same as that for DGRP Isolines & Mutants. However, two differences in procedure were that: (1) more than ten third instar larvae were collected to have their crystal cell number counted for both groups and (2) the UAS-cDNA required crosses with a Gal4 driver genotype line to exhibit their genetic characteristics as UAS/Gal4-cDNA genotypes.

Preparation of cDNA Flies for Crystal Cell Scoring

In the GAL4-UAS system, GAL4 transcription activator proteins bind to UAS enhancers to activate transcription of the cDNA downstream (Figure 4A) [25]. The cDNA genotypes with UAS enhancers (*heartless, PPO1* (chromosome 2)*, PPO1* (chromosome 3)*, hemese, extra-extra, and domino*) and their corresponding Bloomington Stock number are shown in Table 1C. The cDNA genotype flies were crossed with the driver genotype line, Bloomington Stock number

8700, to increase transcription of those UAS-genes in hemocytes [26]. Expression of GFP in the dissected lymph gland of wandering third instar larvae was observed underneath a microscope (Figure 4B). Hemese transmembrane protein was found to regulate the activation and recruitment of hemocytes and are specifically expressed in *Drosophila* hemocytes and lymph gland. Zettervall et. al in 2004 fused the *Hemese* promoter to the yeast GAL4 gene to generate a *Hemese*-Gal4 driver *Drosophila* line through a third chromosome insertion of *P{Hemese-GAL4}85*, expressing strong GFP in hemocytes.

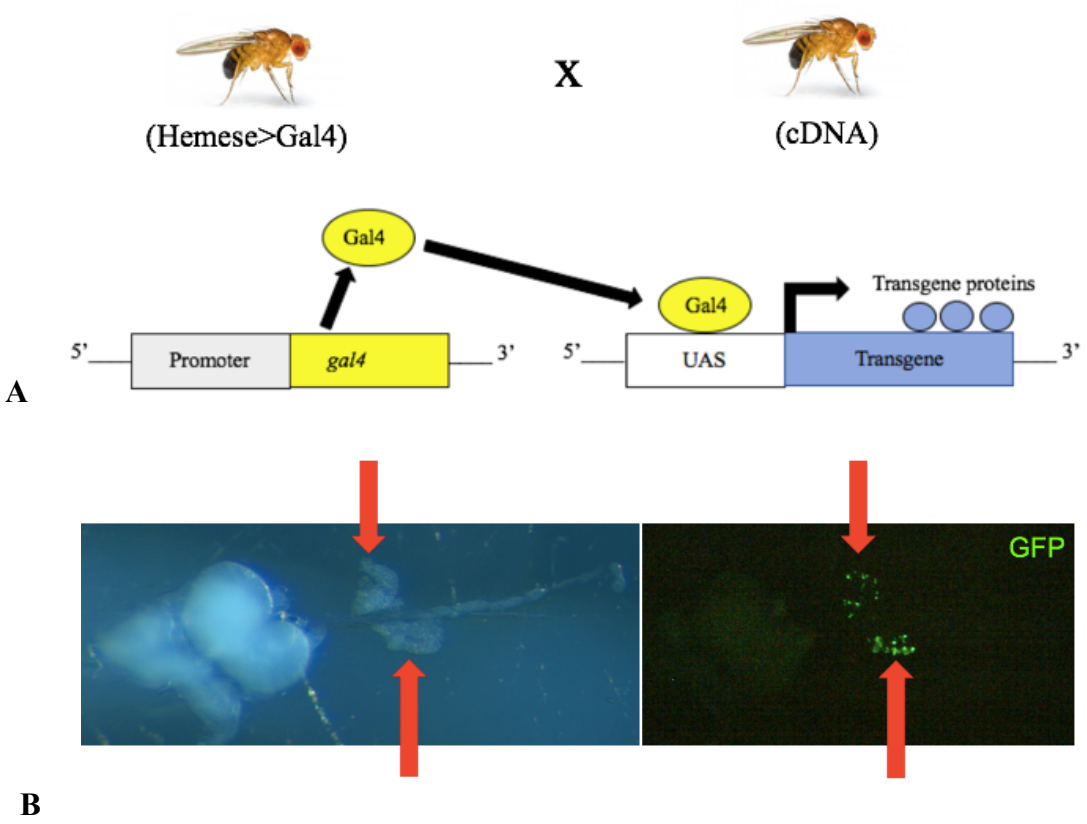

Figure 4 **GAL4-UAS System in 8700 line (A)** GAL4-UAS System is a biochemical method used to study gene expression in organisms such as *Drosophila melanogaster.* GAL4 transcription activator proteins will bind to UAS enhancers to activate transcription of the transgene downstream. *Adapted from Brand and Perrimon, 1993* [25]. **(B)** Expression of GFP in the dissected lymph gland of the 8700 driver line wandering third instar larvae (indicated by the red arrows).

Scoring of Crystal Cells for RNAi Flies

Preparation and Scoring of RNAi Flies

The crystal cell counting procedure for RNAi genotypes was exactly the same as that for the UAS-cDNA genotypes. However, the only difference was that some RNAi lines required multiple crosses to produce a stable stock (Table 1D). Then the flies from the stable stock were crossed with the driver genotype line 8700 that expresses GAL4 in hemocytes. However, the collection of the crystal cell data for the RNAi lines are still in progress.

Results

Crystal Cell Number Variation Across the DGRP

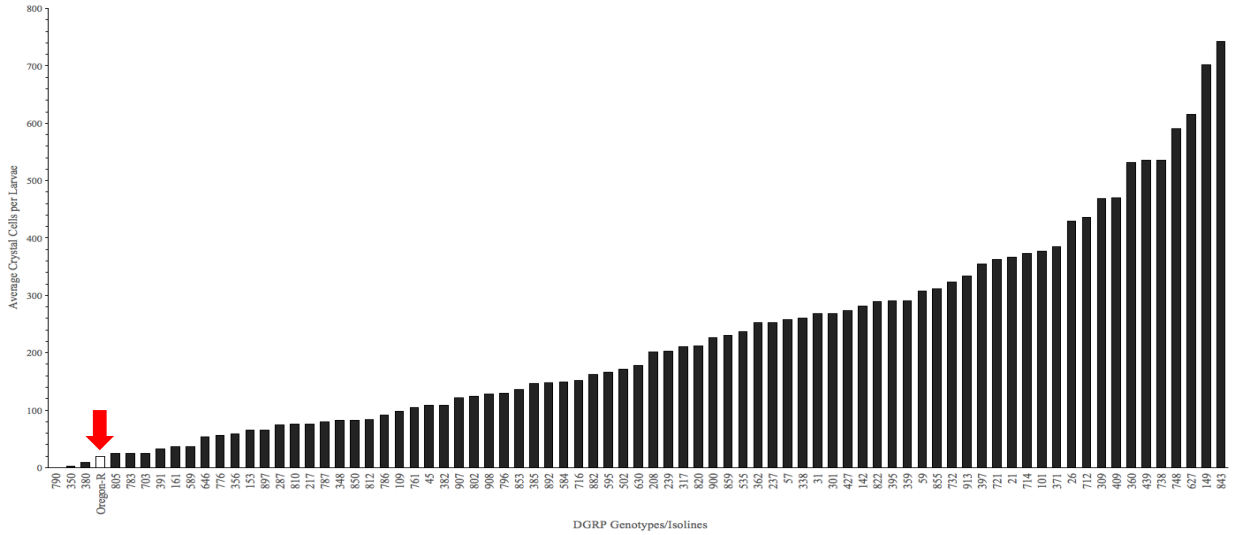

Figure 5 **Crystal Cell Variation Across DGRP Lines** Average number of crystal cells of 10 larvae per genotype from the DGRP were counted. The DGRP genotypes/lines on the x-axis are organized by crystal cell number. 78 DRGP isolines (black bars) and an Oregon-R line (white bar marked with red arrow) are shown for comparison.

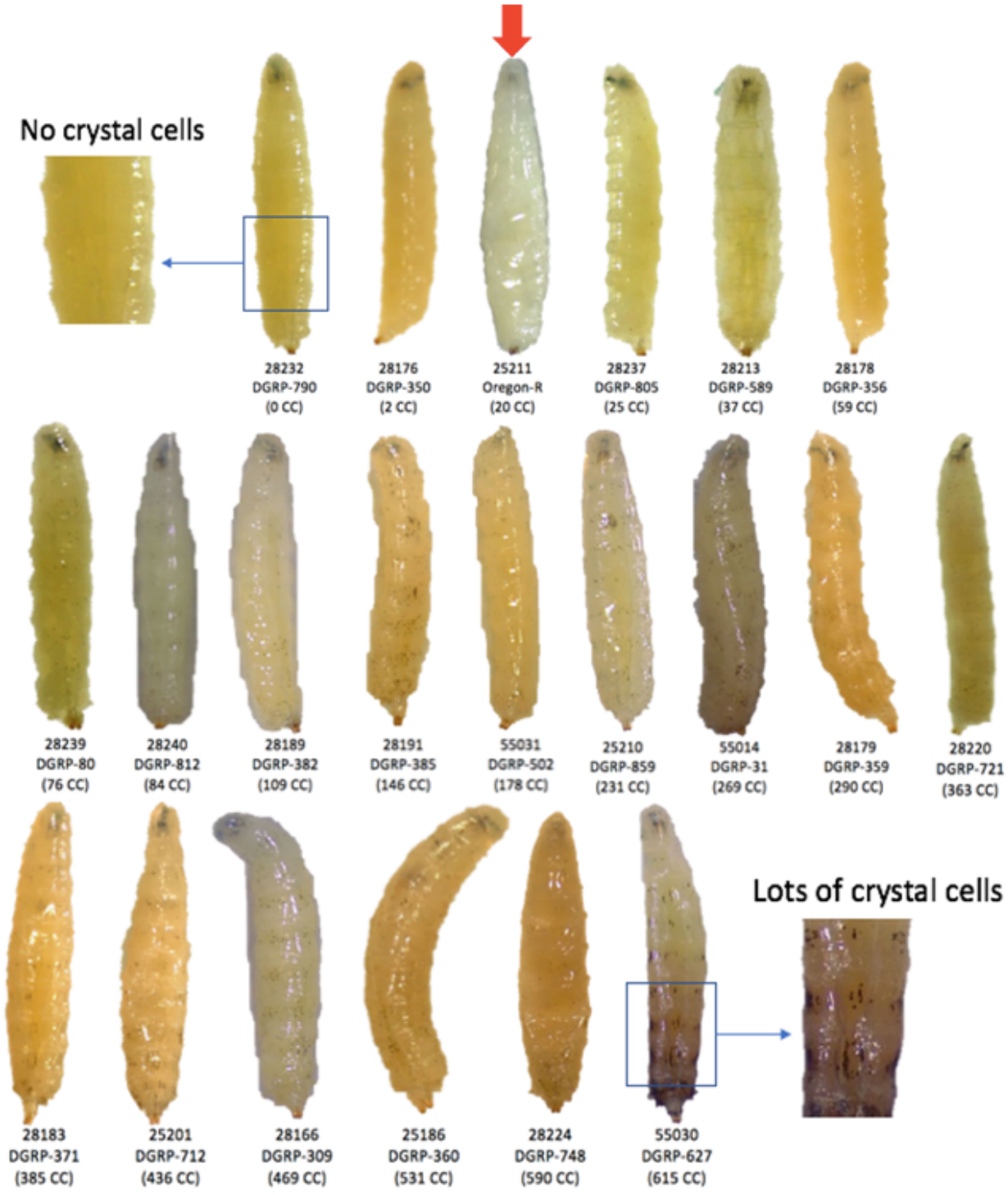

Figure 6 **Larvae of DGRP Isolines After Heatshock** The larvae of 21 of 78 randomly selected DGRP genotypes are shown with the average crystal cells per genotype. Larvae of Oregon-R line (marked with red arrow) with 20 crystal cells per larvae, is shown for comparison. Larvae are shown with various quantities of crystal cells per larvae ranging from 0 crystal cells to 615 crystal cells.

Of the 200 isolines in the DGRP collection, 78 isolines were randomly selected from the collection for crystal cell quantification. For each isoline, 10 larvae were placed into the thermal cycler for heatshock to visualize crystal cells after melanization. The crystal cell count per larvae in each isoline was calculated by taking an average of all 10 larvae, resulting in the data displayed in Figure 5**.** For the first 35 lines, sex of the larvae was not determined. For the remaining 43 lines, 5 males and 5 females were scored for each isoline. Future analysis will be needed to determine whether or not the average crystal cells between males and females of the same line are significantly different.

As the DGRP collection consists of naturally caught fruit flies that are inbred to near homozygosity, differences between the crystal cell count of DGRP collection and wild type flies were compared. The Oregon-R line, which is not part of the DGRP collection, but rather a wild type fly, was used as a reference for comparison (white bar in Figure 5). Whereas 20 crystal cell count per larvae were observed in the Oregon-R line, there was a wide range of crystal cell count per larvae across the DGRP isolines, with a range of 0 crystal cells to 730 crystal cells (black bars in Figure 5). Of the 78 DGRP lines, 75 of them had more crystal cells when compared to the Oregon-R control line. Pictures documented with camera synced microscopes of larvae representing their isolines (not all 78 lines) are shown in Figure 6. The small black dots are crystal cells that have melanized after the larvae have been exposed to heatshock. The mean crystal cell count of all the isolines' crystal cell count per larvae was determined to be 220.

GWAS: Polymorphisms of Genes Associated with Hematopoiesis and Immune Response

After collection of crystal cell count (trait) data from 78 DGRP isolines, a GWAS analysis was conducted, which located polymorphisms associated with variation in crystal cell

count between isolines (Table 2). There were 5 million possible polymorphisms, therefore, only polymorphisms with an association supported by a p-value less than 10^{-5} were reported as the "top hits." Genes (highlighted in yellow in Table 2) were selected for functional testing and for the purchase of mutant, cDNA, and RNAi flies to verify that those genes had an impact on crystal cell count variation. The genes selected from the list included: *CG4390*, *CG17193*, *domino*, *prickle*, *extra-extra*, and *heartless*.

Polymorphism Location	Gene Annotation	Gene	P-value
2R_13764187_SNP	SiteClass[FBgn0011589]	elk	1.43E-08
3R_15882599_SNP	SiteClass[FBgn0038771]	CG4390	7.40E-07
3R_15879811_SNP	SiteClass[FBgn0040571]	CG17193	1.42E-07
2R_17228425_SNP	SiteClass[FBgn0020306]	dom	3.07E-07
3R_15881036_SNP	SiteClass[FBgn0040571]	CG17193	5.52E-07
2R_3109796_SNP	SiteClass[FBgn0003090]	pk	3.76E-07
2L_6159789_SNP	SiteClass[FBgn0031786]	CG13989	6.30E-07
2R_13764180_SNP	SiteClass[FBgn0011589]	elk	3.71E-07
X_10098602_SNP	SiteClass[FBgn0052690]	CR32690	4.91E-07
3R_15877377_SNP	SiteClass[FBgn0040571]	CG17193	1.12E-06
2R_19330815_SNP	SiteClass[FBgn0264339]	CG43795	7.59E-07
2L_2277282_SNP	SiteClass[FBgn0264084]	CR43753	4.00E-06
2L_22631044_SNP	SiteClass[FBgn0058006]	CG40006	1.46E-06
3R_15123660_SNP	SiteClass[FBgn0265063]	CG44174	3.34E-06
2L_16707980_SNP	SiteClass[FBgn0062978]	CG31808	3.09E-06
2R_13764149_SNP	SiteClass[FBgn0011589]	elk	6.59E-07
2R_13764156_SNP	SiteClass[FBgn0011589]	elk	8.29E-07
3R_15876868_SNP	SiteClass[FBgn0040571]	CG17193	1.92E-06
3R_15878120_SNP	SiteClass[FBgn0040571]	CG17193	2.54E-06
2L_3130164_SNP	SiteClass[FBgn0015600]	toc	1.63E-06
2L_3130168_SNP	SiteClass[FBgn0015600]	toc	1.63E-06
2R_12139947_SNP	SiteClass[FBgn0261612]	Cng	2.70E-06
2R_1655688_SNP	SiteClass[FBgn0033039]	gp210	1.49E-06
2R_17190075_SNP	SiteClass[FBgn0261554]	CG42672	1.63E-06
X_1055492_SNP	SiteClass[FBgn0040368]	eIF4E-7	2.20E-06
3L_7941395_SNP	SiteClass[FBgn0041156]	exex	1.60E-06
2L_4961585_SNP	SiteClass[FBgn0083960]	CG34124	2.83E-06
2L_4961615_SNP	SiteClass[FBgn0083960]	CG34124	2.83E-06
3R_15878285_SNP	SiteClass[FBgn0040571]	CG17193	3.83E-06
3L_16317914_SNP	SiteClass[FBgn0036608]	CG13040	2.87E-06

X_1070108_SNP	SiteClass[FBgn0040363]	CG11384	2.01E-06
X_14229128_SNP	SiteClass[FBgn0052600]	dpr8	1.86E-06
2R_18533793_SNP	SiteClass[FBgn0034753]	CG2852	2.00E-06
X_1055470_SNP	SiteClass[FBgn0040368]	eIF4E-7	3.03E-06
3L_7941820_SNP	SiteClass[FBgn0041156]	exex	1.45E-06
3R_15878847_SNP	SiteClass[FBgn0040571]	CG17193	3.09E-06
3R_15874505_SNP	SiteClass[FBgn0040571]	CG17193	7.03E-06
2L_12919914_SNP	SiteClass[FBgn0040509]	ACXB	2.37E-06
2L_13679639_SNP	SiteClass[FBgn0051814]	CG31814	4.00E-06
3L_7944257_SNP	SiteClass[FBgn0041156]	exex	3.27E-06
2L_12923576_SNP	SiteClass[FBgn0040510]	ACXA	2.69E-06
2L_12925525_SNP	SiteClass[FBgn0040510]	ACXA	2.69E-06
3L_7942672_SNP	SiteClass[FBgn0041156]	exex	3.62E-06
2L_12919936_SNP	SiteClass[FBgn0040509]	ACXB	2.35E-06
3L_16349980_SNP	SiteClass[FBgn0036616]	CG4893	3.82E-06
3L_16152983_SNP	SiteClass[FBgn0036576]	CG5151	2.90E-06
3L_9542042_DEL	SiteClass[FBgn0261555]	CG42673	6.73E-06
2R_1654551_SNP	SiteClass[FBgn0033039]	gp210	4.16E-06
3R_24508379_INS	SiteClass[FBgn0039594]	CG9990	7.59E-06
3L_4484411_SNP	SiteClass[FBgn0035552]	CG11350	4.78E-06
3L_7942502_SNP	SiteClass[FBgn0041156]	exex	3.27E-06
3L_7942529_SNP	SiteClass[FBgn0041156]	exex	3.27E-06
3R_13874873_SNP	SiteClass[FBgn0010389]	htl	3.21E-06
3R_15123673_SNP	SiteClass[FBgn0265063]	CG44174	1.09E-05
3R_15884874_SNP	SiteClass[FBgn0038771]	CG4390	1.12E-05
2L_4847449_SNP	SiteClass[FBgn0031637]	CG2950	8.02E-06
3R_15884449_SNP	SiteClass[FBgn0038771]	CG4390	9.36E-06
3L_7942511_INS	SiteClass[FBgn0041156]	exex	3.87E-06
2R_9929453_SNP	SiteClass[FBgn0002643]	mam	6.57E-06
2R_17048116_SNP	SiteClass[FBgn0050389]	CG30389	4.96E-06
2R_1662444_SNP	SiteClass[FBgn0025693]	CG11163	3.08E-06
2R_1666225_DEL	SiteClass[FBgn0025693]	CG11163	6.30E-06
2R_8992352_SNP	SiteClass[FBgn0265429]	CG44341	7.63E-06
X_15980370_INS	SiteClass[FBgn0028397]	Tob	6.14E-06
2R_1656388_SNP	SiteClass[FBgn0033039]	gp210	6.43E-06
3L_7942587_SNP	SiteClass[FBgn0041156]	exex	3.76E-06
X_14727829_SNP	SiteClass[FBgn0030582]	CG14411	4.41E-06
2L_10701132_SNP	SiteClass[FBgn0023496]	Lip1	1.10E-05
3L_5911346_SNP	SiteClass[FBgn0035649]	CG10483	3.73E-06
X_15980346_SNP	SiteClass[FBgn0028397]	Tob	8.23E-06
X_15980354_SNP	SiteClass[FBgn0028397]	Tob	7.21E-06
X_15980389_SNP	SiteClass[FBgn0028397]	Tob	7.21E-06
2L_4709708_DEL	SiteClass[FBgn0085380]	CG34351	9.85E-06
2R_1661643_SNP	SiteClass[FBgn0033039]	gp210	7.45E-06
2L_3899200_SNP	SiteClass[FBgn0000256]	capu	2.24E-05

2R_1683839_SNP	SiteClass[FBgn0033042]	Tsp42A	6.25E-06
3R_15876056_SNP	SiteClass[FBgn0040571]	CG17193	1.16E-05
2L_16852760_SNP	SiteClass[FBgn0051810]	CG31810	2.56E-05
2L_5300905_INS	SiteClass[FBgn0016076]	vri	9.28E-06
2L_10700912_SNP	SiteClass[FBgn0023496]	Lip1	1.33E-05
2R_18427641_SNP	SiteClass[FBgn0003175]	px	6.03E-06
2R_18427642_SNP	SiteClass[FBgn0003175]	px	6.03E-06
2R_1669353_SNP	SiteClass[FBgn0025693]	CG11163	9.81E-06
2L_3899247_SNP	SiteClass[FBgn0000256]	capu	2.20E-05
2R_1668155_SNP	SiteClass[FBgn0025693]	CG11163	8.68E-06
2R_1669755_INS	SiteClass[FBgn0025693]	CG11163	8.68E-06
2L_8272531_SNP	SiteClass[FBgn0003502]	Btk29A	8.40E-06
2R_1669376_SNP	SiteClass[FBgn0025693]	CG11163	1.14E-05
3R_24841783_SNP	SiteClass[FBgn0039620]	CG1443	9.18E-06
2R_1668154_SNP	SiteClass[FBgn0025693]	CG11163	8.78E-06
3R_15877860_SNP	SiteClass[FBgn0040571]	CG17193	1.11E-05
3R_15877862_SNP	SiteClass[FBgn0040571]	CG17193	1.11E-05
X_15008903_SNP	SiteClass[FBgn0030615]	Cyp4s3	1.06E-05
3R_15864341_DEL	SiteClass[FBgn0040571]	CG17193	1.09E-05
2R_1656665_SNP	SiteClass[FBgn0033039]	gp210	7.34E-06
2R_1671078_SNP	SiteClass[FBgn0025693]	CG11163	9.88E-06
2L_3899225_INS	SiteClass[FBgn0000256]	capu	2.58E-05
3L_5912625_SNP	SiteClass[FBgn0035649]	CG10483	5.89E-06
2R_1671039_SNP	SiteClass[FBgn0025693]	CG11163	8.93E-06
2R_1682240_SNP	SiteClass[FBgn0033042]	Tsp42A	8.93E-06
2R_1682699_SNP	SiteClass[FBgn0033042]	Tsp42A	8.93E-06
2R_1664687_SNP	SiteClass[FBgn0025693]	CG11163	1.01E-05
2R_14009929_SNP	SiteClass[FBgn0003520]	stau	1.44E-05
3R_25605938_SNP	SiteClass[FBgn0039694]	fig	1.32E-05
2L_5902481_SNP	SiteClass[FBgn0031747]	CG9021	6.39E-06
2R_1669317_INS	SiteClass[FBgn0025693]	CG11163	1.23E-05
3R_24538539_SNP	SiteClass[FBgn0027655]	htt	2.46E-05
3R_13432919_SNP	SiteClass[FBgn0038524]	sll	1.52E-05
3L_258643_SNP	SiteClass[FBgn0035120]	wac	2.23E-05
2R_1683819_SNP	SiteClass[FBgn0033042]	Tsp42A	1.24E-05
X_19300601_SNP	SiteClass[FBgn0031016]	kek5	2.42E-05
X_9746079_SNP	SiteClass[FBgn0052698]	CG32698	1.11E-05
2R_1670058_SNP	SiteClass[FBgn0025693]	CG11163	1.17E-05
2R_18180862_SNP	SiteClass[FBgn0034717]	CG5819	1.64E-05
2L_15083679_SNP	SiteClass[FBgn0028879]	CG15270	2.25E-05
2R_10155171_SNP	SiteClass[FBgn0033919]	CG8547	7.07E-06
3R_15877846_SNP	SiteClass[FBgn0040571]	CG17193	1.72E-05
3L_8149841_SNP	SiteClass[FBgn0259916]	CG42445	2.63E-05
2R_16208268_SNP	SiteClass[FBgn0027529]	CG8920	9.29E-06
3L_9878342_SNP	SiteClass[FBgn0011836]	Taf2	1.22E-05

3L_9878344_SNP	SiteClass[FBgn0011836]	Taf2	1.22E-05
3L_5920084_SNP	SiteClass[FBgn0053523]	CG33523	9.92E-06
3R_15876621_SNP	SiteClass[FBgn0040571]	CG17193	8.12E-06
X_15434283_SNP	SiteClass[FBgn0022710]	Ac13E	8.86E-06
2R_12139641_SNP	SiteClass[FBgn0261612]	Cng	9.59E-06
2R_16592412_INS	SiteClass[FBgn0040726]	dpr	2.80E-06
2L_7242492_SNP	SiteClass[FBgn0259111]	Ndae1	5.63E-06
2R_17248386_SNP	SiteClass[FBgn0000395]	cv-2	9.03E-06

Table 2 **GWAS Analysis Results** After collecting crystal cell count data from 78 DGRP lines, a GWAS analysis was done. Over 100 polymorphisms found to be associated with variation in crystal cell count between DGRP lines, $p<10^{-5}$. Highlighted in yellow are genes and polymorphisms that were selected for functional testing.

Heartless codes for a transmembrane tyrosine kinase receptor for fibroblast growth factors (FGF) and plays a role in *Drosophila* embryo mesoderm migration that leads to differentiation of the mesoderm into different cell types. Studies have found that *heartless* mutants have normal formation of the mesoderm layer, but lack mesodermal invagination [27]. Without the migration of the mesoderm, the heart and other muscles are not formed. In non-mutants, a transforming growth factor known as Decapentaplegic (Dpp) induces the formation of the heart and muscle organs in the embryo [27].

Extra-extra codes for a homeodomain transcription factor that can be mainly found in motor neurons within fruit flies. It was found to regulate the differentiation of motor neurons that travel to and control ventral body wall muscles [28]. *Exex* does so by limiting the expression of Lim3 and Even-skipped (Eve), which are two other homeodomain proteins required for development of neurons other than motor neurons [29]. Exex protein is dependent on Groucho corepressor protein in the Notch pathway to have an inhibitory effect on Eve. More specifically, the ventral body wall muscles that *Exex* motor neurons innervate are important for eclosion. Removal of those motor neurons produced defects in adult eclosion [28].

Prickle codes for a REST/NRSF-interacting LIM domain protein that helps with the organization of microtubule polarity in axons of neurons. As such, the gene plays a role in vesicle transportation (back and forth) between the soma (cell body of the neuron) and the terminal button by passing through the axon [30]. Mutations of *prickle* were found to be associated with seizures in flies as well as humans. Whereas heterozygous *prickle* flies showed improvement after being administered anti-epileptic medication, homozygous *prickle* fly mutants had major brain defects [31].

Domino codes for a protein of the SWI2/SNF2 family that contributes to DNA related processes such as transcription, replication, and repair [32]. Mutations of the *domino* gene lead to genomic mutations occurring at a greater frequency, leading to the conclusion that it is necessary for cell viability and proliferation [33]. Additionally, studies have found that domino contributes to the process of hematopoiesis in fruit flies. Homozygous *domino* mutant larvae have been found to have good survival rates as larvae, but usually die as prepupae [34]. The study also found that hemolymph collected from homozygous third instar *domino* mutant larvae had absence of circulating hemocytes. On the contrary, wild type larvae have been found to have intense melanization when exposed to fungal infections [17, 34].

Lastly, *CG4390* and *CG17193* genes are lesser known in comparison to the aforementioned genes. However, what are known about these two genes are what proteins they code for. Whereas *CG4390* codes for S-formylglutathione hydrolase protein, *CG17193* codes for a protein called GEO09915p1 [35, 36].

Increased Crystal Cell Formation in Mutant Genotypes

Wild type Gene	Bloomington Stock #	Average Crystal Cell Per Larvae
Oregon-R-modENCODE	25211	20
Mutant Genes		
pk[1] cn[1]	367	187 *+
prd[8]/CyO	5806	169 *+
w[*]; **htl**[AB42]/TM3, P{ry[+t7.2]=ftz/lacC}SC1, ry[RK] Sb[1] Ser[1]	5370	695 *+
y[1] w[*]; **dom**[3]/SM6a	9260	115 *+
y[1] w[*]; Mi{y[+mDint2]=MIC}**CG4390**[MI03759]/TM3, Sb[1] Ser[1]	36980	160 *+
y[1] w[*]; Mi{y[+mDint2]=MIC}Indy-2[MI01115] **CG17193**[MI01115] CG33934[MI01115]	34149	130 *+
w[*]; **exex**[KK30] e[s]/TM3, P{ry[+t7.2]=ftz-lacZ.ry[+]}TM3, Sb[1] ry[*]	9930	546 *+
P{ry[+t7.2]=PZ}**msn**[06946] ry[506]/TM3, ry[RK] Sb[1] Ser[1]	11707	755 *+

Note: *+ indicates significantly more with p-value < 0.05 in comparison of mutants to 25211 Oregon-R; bolded words were associated genes found via GWAS analysis

Table 3 **Mutant Genotypes** *Drosophila* mutants of the some of the associated genes found through GWAS to crystal cell count number variation was obtained. Larvae with mutations of the above genes were heatshocked and their crystal cells counted. Analyzed using a t-test, all of the mutant genotype crystal cell counts were significantly greater than the Oregon-R line (p-value<0.05). However, Oregon-R may not be the appropriate control.

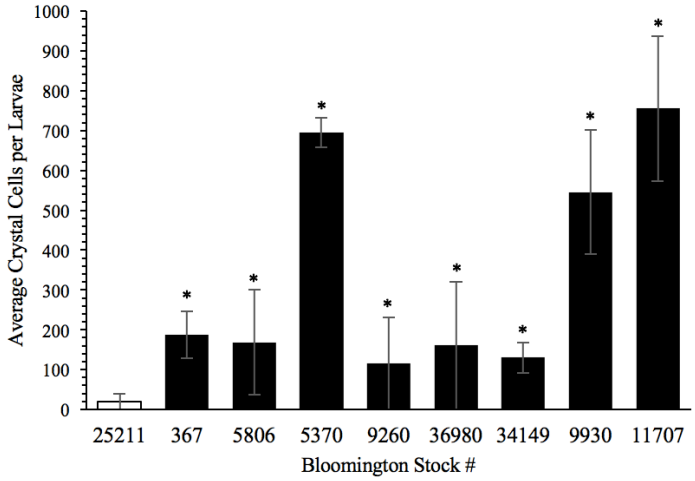

Figure 7 **Graph of Data Represented in Table 3 (above)** The white bar represents Oregon-R line (25211) and the black bars represent the mutant genotypes. The error bars indicate standard deviations.
* indicates significance with p-value < 0.05 in comparison of mutants to 25211 Oregon-R (white bar)

After the GWAS analysis, genes assigned to polymorphisms associated with variation in crystal cell count between DGRP isolines were selected for functional assessment. The purpose of the functional test was to determine if mutations of these genes will lead to significant differences (more or less) in crystal cell count per larvae. Mutant genotypes of those genes were obtained through the Bloomington Stock center and utilized to quantify crystal cells. Due to Oregon-R line (white bar) being a line with a genotype that is close to that of a wild type fly, its crystal cell count data was used to compare with that of the mutant type flies (black bars) shown in Table 3 and Figure 7. Using a t-test to compare the average Oregon-R crystal cell count of 20 cells per larvae to all the mutant genotype larvae, we found that each of the larvae from all the mutant genotypes had significantly greater crystal cell count than Oregon-R larvae (p-value<0.05). According to the results, it indicates that the selected genes do play a role in and have an impact on crystal cell differentiation. However, Oregon-R may not be a good control to be compared to the mutants because the mutations of the selected genes are not the only mutations that are found in the genotypes of flies that were purchased. Instead, there are also

marker mutations that were present in the background. Future experiments will examine crystal cell number in lines with the marker mutations alone.

Formation of More Crystal Cells Through Overexpression of *domino*, *extra-extra*, and *PPO1* and Fewer Crystal Cells Through Overexpression of *hemese*

cDNA Transgenes	UAS-cDNA Bloomington #	UAS-cDNA: Average Crystal Cell Per Larvae	UAS/Gal4-cDNA	UAS/Gal4-cDNA: Average Crystal Cell Per Larvae
hemese	8697	460	8700 x 8697	179 *-
domino	64261	145	8700 x 64261	327 *+
heartless	5419	255	8700 x 5419	317
extra-extra	9929	258	8700 x 5867	431 *+
PPO1 (chromosome 2)	5867	298	8700 x 9929	605 *+
PPO1 (chromosome 3)	5868	77	8700 x 5868	110 *+

Note: *+ indicates significantly more with p-value < 0.05 in comparison of UAS-cDNA to UAS/Gal4-cDNA; *- indicates significantly less with p-value < 0.05 in comparison of UAS-cDNA to UAS/Gal4-cDNA

***Table 4* Overexpression of Transgenes** The transgenes that correspond with the Bloomington Stock number shown in Figure 8 (below) are listed. Crystal cell count data of larvae from no-driver (UAS-cDNA) and driver crossed lines (UAS/Gal4-cDNA) were compared through a t-test.

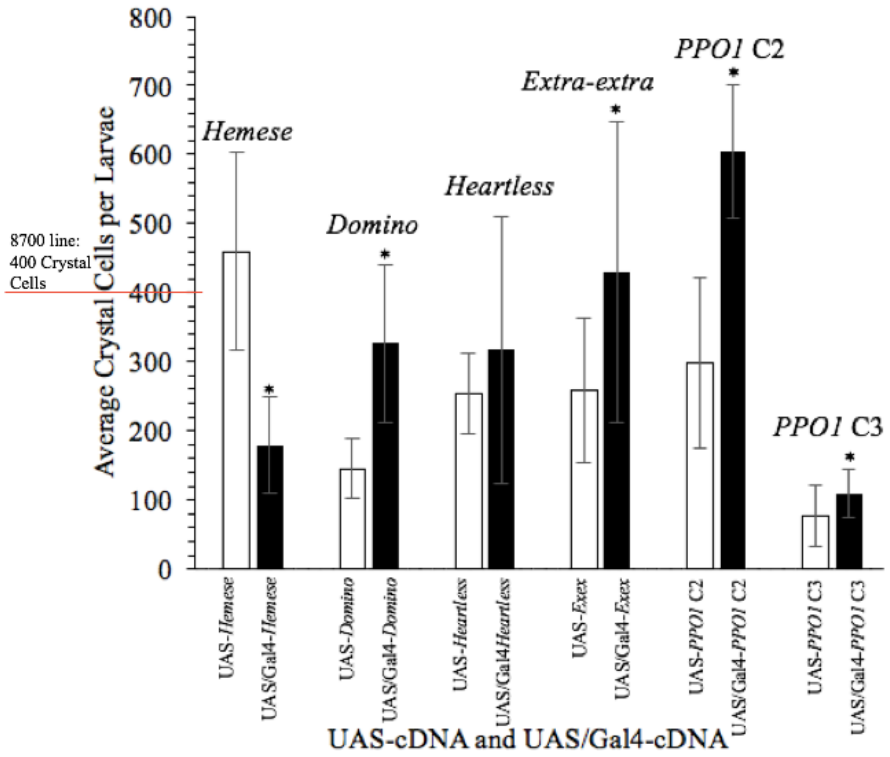

Figure 8 **Graph of Data Represented in Table 4 (above)** The white bars are crystal cell count per larvae from no-driver (UAS-Transgene). The black bars are crystal cell count per larvae of driver crossed lines (UAS/Gal4-Transgene). The error bars indicate standard deviations. The red line next to the y-axis indicates the number of crystal cells for the Gal4 driver line alone (average of 400 crystal cells).

* indicates significant difference with p-value < 0.05 in comparison of UAS-cDNA to UAS/Gal4-cDNA

In the GAL4-UAS system, GAL4 transcription activator proteins bind to UAS enhancers to activate transcription of the transgene downstream (Figure 4). The UAS enhancer cDNA genotypes for UAS-genes (*heartless, PPO1* (chromosome 2)*, PPO1* (chromosome 3)*, hemese, extra-extra, and domino*) were crossed with the driver genotype line 8700 that expresses GAL4 in hemocytes (Table 4 and Figure 8). Four out of five significant differences (obtained through t-test) were of more crystal cell formation through overexpression (p-value<0.05). The other one out of the five was of less number of crystal cell formation was observed through overexpression of the gene. Significantly more crystal cell formation was observed through the overexpression of *domino*, *extra-extra*, and *PPO1* (both chromosome 2 and 3) genes. However, significantly less

number of crystal cell formation was observed through overexpression of *hemese* and no significant difference for overexpression of *heartless* were observed. When looking at the data in terms of males vs females in the same UAS/Gal4-cDNA lines, both groups showed similar patterns in significance in comparison to their UAS-cDNA lines. In other words, males and females both had a significant increase in crystal cell count in overexpressed *domino* and *extra-extra* lines; both males and females had a significant decrease in crystal cell count in overexpressed *hemese* line.

Looking at the 8700 driver line alone, the number of average crystal cells per larvae was observed to be 400. This indicates that the driver may have some impact on increase in crystal cell number for some of the UAS-cDNA lines. As a result, a follow-up experiment has to be done with a different driver to determine if similar results are achieved.

Discussion

From DGRP to GWAS to Functional Validation Assessments

78 isolines were randomly selected from the DGRP collection for crystal cell quantification. When compared with the crystal cell quantity of Oregon-R line, a line with genotype that is close to that of a wild type fly, it was evident that flies of the DGRP collection differ from the wild type flies. Simultaneously, the wide range in variation of crystal cell counts across the DGRP lines indicate that the DGRP lines not only differs from Oregon-R, but also from each other as well. Using the data from the DGRP crystal cell scoring experiment, a GWAS analysis was performed which provided a list of over 100 gene polymorphisms that may be responsible for variation in crystal cell count between the DGRP isolines. Out of the 128 gene polymorphisms, a few genes were selected for functional assessment to verify that those genes have an impact on crystal cell count variation.

Functional assessments of these genes included the use of mutants as well as cDNA (UAS/Gal4 system) genotypes. Of the genes functionally tested, *domino*, *extra-extra*, and *hemese* flies showed significant differences in crystal cell numbers. *Domino* and *extra-extra* mutant larvae showed significantly higher crystal cell number per larvae as compared to that of the Oregon-R line ($p\text{-value}<0.05$). Surprisingly, the *domino* mutants in this study increased crystal cell count number contrary to previous literature. As for the overexpression of *domino* and *extra-extra* transgenes using the UAS/Gal4 system, both genes significantly increased crystal cell count per larvae. Interestingly, the overexpression of the *domino* gene aligns with previous literature regarding its involvement in quantity of circulating hemocytes in third instar larvae. However, overexpression of *hemese* lead to a significantly decreased crystal cell count per larvae.

New Additions to Known Crystal Cell Pathway

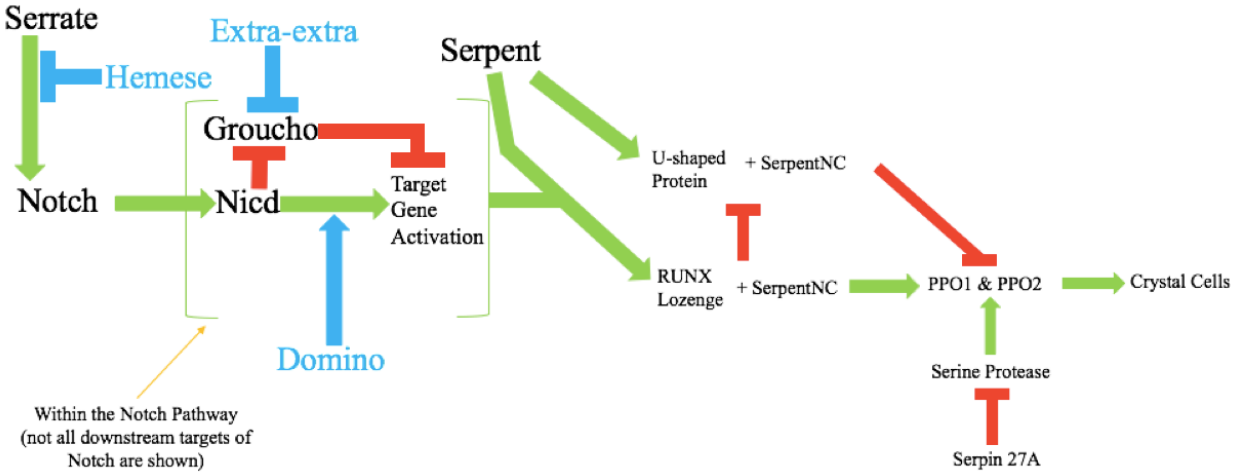

Figure 9 **New Hypothetical Pathway Based on Study** New hypothetical pathway from this study shows in blue incorporates domino, extra-extra, and hemese into the pathway shown in Figure 3.

Utilizing the data obtained from this experiment and information from previous literature, it was possible to create a hypothetical pathway that incorporates *domino*, *extra-extra*, and *hemese*. The pathway (Figure 3) was created through information obtained from previous literature and was used as a foundation to incorporate the three new genes. However, unlike the pathway shown in Figure 3, extra molecules downstream of Notch (but within the Notch pathway) were added to provide more specific locations in which *domino* and *extra-extra* contribute to upregulating the known crystal cell differentiation pathway. The molecules added were Notch intracellular domain (Nicd) and Groucho [37].

Extra-extra and Domino: Notch Pathway Activator

To fit extra-extra into the pathway, general knowledge has to be known about the Notch pathway. When Serrate ligands bind to Notch receptors, a part of the Notch receptor that lies

within the membrane known as Nicd is cleaved [37]. The cleavage of Nicd leads to target gene activation downstream of the signaling pathway. On the other hand, Groucho is one of four protein subunits that form a complex that represses targets downstream of Nicd in the absence of Nicd [37]. Thus, when Notch is not activated, Groucho plays the role of repressing unintentional Notch signaling from occurring. It is previously known that *extra-extra* codes for proteins that represses Groucho proteins. With the experimental overexpression of *extra-extra* leading to a significantly increase in crystal cell count, it can be assumed that having large amounts of extra-extra induce the effects similar to that of a Notch pathway activator. This may be attributed to the overwhelming number of Groucho protein repression by extra-extra proteins that bypasses the requirement of Serrate binding and Nicd cleavage to activate the Notch pathway.

Similar to extra-extra, domino is involved within the Notch pathway. Instead of repressing a repressor as extra-extra, domino has been known to be recruited after Nicd cleavage to contribute to target gene activation [37]. Experimental overexpression of domino led to a significant increase in larval crystal cell count, which may indicate that overexpressing *domino* increased Notch signaling. Though domino is recruited after Nicd cleavage, overabundance of domino proteins may have increased Notch signaling due to its protein concentration within the cell. Alternatively, overabundance of domino bypassed the requirement of Nicd cleavage to activate the Notch pathway. As a result, domino was placed into the pathway after Nicd cleavage.

Hemese: Notch Pathway Inhibitor

Last but not least, *hemese* gene that codes for only hemese transmembrane proteins was placed to inhibit the start of the Notch signaling pathway. Experimentally driving the

overexpression of *hemese* resulted in an overabundance amount of hemese transmembrane proteins to be found on the surface of hemocyte cell membrane. As Notch signaling requires ligand-receptor crosstalk across the cell membrane, overexpression of hemese transmembrane protein may compete with the amount of receptors available for the ligand-receptor crosstalk. In other words, overexpression of *hemese* may oversaturate the cell membrane with hemese transmembrane protein and ultimately interfere with the binding of Serrate ligand to the Notch receptors.

Conclusion

Having collected larval crystal cell quantity data for 78 of more than 200 DGRP lines and submitted it for GWAS analysis, it was possible to identify 128 polymorphisms mapped to genes that contribute to differentiation of prohemocytes into specifically crystal cells. Of the mapped genes, we tested four new genes which were selected for their small p-values. Three of the four genes (*domino*, *extra-extra*, and *hemese*) significantly altered crystal cell count in the functional validation experiments and one of the four (*heartless*) did not significantly alter crystal cell count.

The next steps that need to be taken are to count more DGRP lines and collect more data for male vs female larvae of the same line. This is because the DGRP collection consists of 200+ isolines and only 78 of them have been experimented on. Furthermore, we would like to validate our DGRP larval crystal cell data by treating lines with antibiotics to ensure that the differences in larval crystal cell count aren't due to any form of infection. As infections can contribute to increase in crystal cell production, we hope to use antibiotics against these infections (if any) and recount crystal cells in some of the lines. Simultaneously, we hope to continue to functionally validate the other genes that are mapped to the 128 polymorphisms. Once again, we will be testing these genes starting with the ones that have the smaller p-values. As the number of average crystal cells per larvae for the 8700 driver line was observed to be 400, another functional validation experiment has to be conducted utilizing a different driver to determine if similar results are to be achieved.

Lastly, though currently in progress, the next step is to complete RNAi experiments for *domino*, *extra-extra*, and *hemese*. As overexpression of *domino*, *extra-extra*, and *hemese* altered crystal cell number, it is expected that knockdown (using RNAi constructs) of the same gene

would cause the opposite effect. Ultimately, as we progress through these future experiments, we hope to continue adding new genes into the already established crystal cell differentiation pathway. As flies serve as model organisms that have many analogous genes to that of humans, understanding the function of the new candidate genes and how they may contribute to crystal cell production may provide us with some insight into human genes that contribute to immune system imbalance.

References

1. Iwasaki, Hiromi, and Koichi Akashi. “Myeloid Lineage Commitment from the Hematopoietic Stem Cell.” *Immunity*, vol. 26, no. 6, 22 June 2007, pp. 726–740., doi:10.1016/j.immuni.2007.06.004.
2. Letourneau, Manon, Francois Lapraz, Anurag Sharma, Nathalie Vanzo, Lucas Waltzer, and Michèle Crozatier. "Drosophila hematopoiesis under normal conditions and in response to immune stress." FEBS Letters 590.22 (2016): 4034-051.
3. Wang, Lihui, et al. “Drosophila as a Model to Study the Role of Blood Cells in Inflammation, Innate Immunity and Cancer.” *Frontiers in Cellular and Infection Microbiology*, vol. 3, Jan. 2014, doi:10.3389/fcimb.2013.00113.
4. Leulier, François, et al. “The Drosophila immune system detects bacteria through specific peptidoglycan recognition.” *Nature Immunology*, vol. 4, no. 5, 2003, pp. 478–484., www.nature.com/articles/ni922.
5. Holz, Anne, Barbara Bossinger, Thomas Strasser, Wilfried Janning, and Robert Klapper. "The two origins of hemocytes in Drosophila." Development 130 (2003): 4955-962.
6. Minakhina, Svetlana, and Ruth Steward. "Hematopoietic stem cells in Drosophila." Development 137.1 (2010): 27-31.
7. Letourneau, Manon, Francois Lapraz, Anurag Sharma, Nathalie Vanzo, Lucas Waltzer, and Michèle Crozatier. "Drosophila hematopoiesis under normal conditions and in response to immune stress." FEBS Letters 590.22 (2016): 4034-051.
8. Fossett, N., et al. “Combinatorial interactions of Serpent, Lozenge, and U-Shaped regulate crystal cell lineage commitment during Drosophila hematopoiesis.” *Proceedings of the National Academy of Sciences*, vol. 100, no. 20, 2003, pp. 11451–11456., doi:10.1073/pnas.1635050100.
9. Rehorn, K P, et al. “A molecular aspect of hematopoiesis and endoderm development common to vertebrates and Drosophila.” *Development (Cambridge, England).*, U.S. National Library of Medicine, Dec. 1996, www.ncbi.nlm.nih.gov/pubmed/9012522.
10. Dudzic, Jan P., et al. “Drosophila innate immunity: regional and functional specialization of prophenoloxidases.” BMC Biology, vol. 13, no. 1, Jan. 2015
11. Muratoglu, Selen, et al. “The GATA factor Serpent cross-Regulates lozenge and u-Shaped expression during Drosophila blood cell development.” *Developmental Biology*, vol. 311, no. 2, 2007, pp. 636–649., doi:10.1016/j.ydbio.2007.08.015.
12. Klug, A., and D. Rhodes. “Zinc Fingers: A Novel Protein Fold for Nucleic Acid Recognition.” *Cold Spring Harbor Symposia on Quantitative Biology*, vol. 52, Jan. 1987, pp. 473–482., doi:10.1101/sqb.1987.052.01.054.
13. Terriente-Felix, A., et al. “Notch cooperates with Lozenge/Runx to lock haemocytes into a differentiation programme” *Development*, vol. 140, no. 4, 2013, pp. 926–937., doi:10.1242/dev.086785.
14. Duvic, Bernard, et al. “Notch Signaling Controls Lineage Specification during Drosophila Larval Hematopoiesis.” *Current Biology*, vol. 12, no. 22, 2002, pp. 1923–1927., doi:10.1016/s0960-9822(02)01297-6.
15. Fossett, Nancy G. “Genetic Regulation of Drosophila Hematopoiesis.” *Grantome*, NIH, 1 Sept. 2006, grantome.com/grant/NIH/R01-DK072229-02.

16. Binggeli, Olivier, et al. “Prophenoloxidase Activation Is Required for Survival to Microbial Infections in Drosophila.” *PLoS Pathogens*, vol. 10, no. 5, Jan. 2014, doi:10.1371/journal.ppat.1004067.
17. Gregorio, Ennio De, et al. “An Immune-Responsive Serpin Regulates the Melanization Cascade in Drosophila.” *Developmental Cell*, vol. 3, no. 4, 2002, pp. 581–592., doi:10.1016/s1534-5807(02)00267-8.
18. Cerenius, Lage, et al. “The proPO-System: pros and cons for its role in invertebrate immunity.” *Trends in Immunology*, vol. 29, no. 6, 2008, pp. 263–271., doi:10.1016/j.it.2008.02.009.
19. Lanot, René, et al. “Postembryonic Hematopoiesis in Drosophila.” *Developmental Biology*, vol. 230, no. 2, 2001, pp. 243–257., doi:10.1006/dbio.2000.0123.
20. Nappi, A.j., et al. “Drosophila serpin 27A is a likely target for immune suppression of the blood cell-Mediated melanotic encapsulation response.” *Journal of Insect Physiology*, vol. 51, no. 2, 2005, pp. 197–205., doi:10.1016/j.jinsphys.2004.10.013.
21. Bush, William S., and Jason H. Moore. “Chapter 11: Genome-Wide Association Studies.” *PLoS Computational Biology*, Public Library of Science, Dec. 2012, www.ncbi.nlm.nih.gov/pmc/articles/PMC3531285/.
22. Lu, Jonathan B. WangHsiao-Ling, and Raymond J. St. Leger. “The Genetic Basis for Variation in Resistance to Infection in the Drosophila Melanogaster Genetic Reference Panel.” *PLOS Pathogens*, Public Library of Science, 3 Mar. 2017, journals.plos.org/plospathogens/article?id=10.1371/journal.ppat.1006260.
23. Howick, V M, and B P Lazzaro. “The Genetic Architecture of Defence as Resistance to and Tolerance of Bacterial Infection in Drosophila Melanogaster.” *Molecular Ecology.*, U.S. National Library of Medicine, 16 Feb. 2017, www.ncbi.nlm.nih.gov/pubmed/28099780.
24. Mackay, Trudy et al. "The Drosophila Melanogaster Genetic Reference Panel." Nature 482, no. 7384 (2012): 173-78. doi:10.1038/nature10811.
25. Brand, A H, and N Perrimon. “Targeted Gene Expression as a Means of Altering Cell Fates and Generating Dominant Phenotypes.” Development (Cambridge, England)., U.S. National Library of Medicine, June 1993, www.ncbi.nlm.nih.gov/pubmed/8223268.
26. Zettervall, Carl-Johan, et al. “A Directed Screen for Genes Involved in Drosophila Blood Cell Activation.” *Proceedings of the National Academy of Sciences of the United States of America*, National Academy of Sciences, 28 Sept. 2004, www.ncbi.nlm.nih.gov/pmc/articles/PMC521135/.
27. Shishido, E, et al. “Requirements of DFR1/Heartless, a Mesoderm-Specific Drosophila FGF-Receptor, for the Formation of Heart, Visceral and Somatic Muscles, and Ensheathing of Longitudinal Axon Tracts in CNS.” *Development (Cambridge, England).*, U.S. National Library of Medicine, June 1997, www.ncbi.nlm.nih.gov/pubmed/9187139.
28. Banerjee, S, et al. “dHb9 Expressing Larval Motor Neurons Persist through Metamorphosis to Innervate Adult-Specific Muscle Targets and Function in Drosophila Eclosion.” *Developmental Neurobiology.*, U.S. National Library of Medicine, Dec. 2016, www.ncbi.nlm.nih.gov/pubmed/27168166.
29. Broihier, H. T., and J. B. Skeath. “Drosophila Homeodomain Protein dHb9 Directs Neuronal Fate via Crossrepressive and Cell-Nonautonomous Mechanisms.” *Neuron*, Cell Press, 17 July 2002, www.sciencedirect.com/science/article/pii/S0896627302007432.

30. Ehaideb, S N, et al. “Prickle Modulates Microtubule Polarity and Axonal Transport to Ameliorate Seizures in Flies.” *Proceedings of the National Academy of Sciences of the United States of America.*, U.S. National Library of Medicine, 29 July 2014, www.ncbi.nlm.nih.gov/pubmed/25024231.
31. Tao, H, et al. “Mutations in Prickle Orthologs Cause Seizures in Flies, Mice, and Humans.” *American Journal of Human Genetics.*, U.S. National Library of Medicine, 11 Feb. 2011, www.ncbi.nlm.nih.gov/pubmed/21276947.
32. Ellis, Kaitlyn, et al. “Drosophila Domino Exhibits Genetic Interactions with a Wide Spectrum of Chromatin Protein-Encoding Loci.” *PLOS ONE*, Public Library of Science, journals.plos.org/plosone/article?id=10.1371/journal.pone.0142635.
33. Ruhf, M L, et al. “The Domino Gene of Drosophila Encodes Novel Members of the SWI2/SNF2 Family of DNA-Dependent ATPases, Which Contribute to the Silencing of Homeotic Genes.” *Development (Cambridge, England).*, U.S. National Library of Medicine, Apr. 2001, www.ncbi.nlm.nih.gov/pubmed/11262242.
34. Braun, Anne, et al. “Analysis of the Drosophila Host Defense in Domino Mutant Larvae, Which Are Devoid of Hemocytes.” *PNAS*, National Academy of Sciences, 24 Nov. 1998, www.pnas.org/content/95/24/14337.full.
35. European Bioinformatics InstituteProtein Information ResourceSIB Swiss Institute of Bioinformatics. “S-Formylglutathione Hydrolase.” *Dmel\CG4390 - S-Formylglutathione Hydrolase - Drosophila Melanogaster (Fruit Fly) - Dmel\CG4390 Gene & Protein*, 28 Mar. 2018, www.uniprot.org/uniprot/Q9VDP9.
36. European Bioinformatics InstituteProtein Information ResourceSIB Swiss Institute of Bioinformatics. “GEO09915p1.” *BcDNA:RE41733 - GEO09915p1 - Drosophila Melanogaster (Fruit Fly) - BcDNA:RE41733 Gene & Protein*, 28 Mar. 2018, www.uniprot.org/uniprot/Q8SXH7.
37. Bray, Sarah J. “Notch Signaling: A Simple Pathway Becomes Complex.” Nature Reviews Molecular Cell Biology, vol. 7, 1 Sept. 2006, pp. 1–1., doi:10.1002/9783527678679.dg08522.