



January 2015

Molecular Mechanisms Underlying Wing Color Pattern Development In *Vanessa Cardui* And Evolution Of The Polycomb Repressive Complex 2

Heidi Connahs

Follow this and additional works at: <https://commons.und.edu/theses>

Recommended Citation

Connahs, Heidi, "Molecular Mechanisms Underlying Wing Color Pattern Development In *Vanessa Cardui* And Evolution Of The Polycomb Repressive Complex 2" (2015). *Theses and Dissertations*. 1758.
<https://commons.und.edu/theses/1758>

This Dissertation is brought to you for free and open access by the Theses, Dissertations, and Senior Projects at UND Scholarly Commons. It has been accepted for inclusion in Theses and Dissertations by an authorized administrator of UND Scholarly Commons. For more information, please contact zeinebyousif@library.und.edu.

MOLECULAR MECHANISMS UNDERLYING WING COLOR PATTERN
DEVELOPMENT IN *VANESSA CARDUI* AND EVOLUTION OF THE POLYCOMB
REPRESSIVE COMPLEX 2

by

Heidi Connahs
Bachelor of Science, Brighton University, East Sussex, UK, 2002
Master of Science, Tulane University, New Orleans, Louisiana, 2009

A Dissertation
Submitted to the Graduate Faculty

of the

University of North Dakota

In partial fulfillment of the requirements

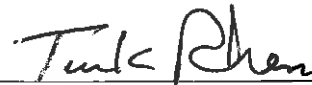
For the degree of

Doctor of Philosophy

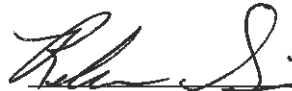
Grand Forks, North Dakota

May
2015

This dissertation, submitted by Heidi Connahs in partial fulfillment of the requirements for the Degree of Doctor of Philosophy from the University of North Dakota, has been read by the Faculty Advisory Committee under whom the work has been done and is hereby approved.



Turk E. Rhen, Co-advisor



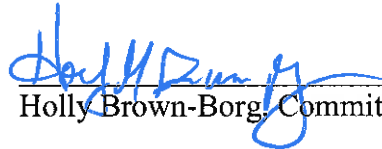
Rebecca B. Simmons, Co-advisor



Brian Darby, Committee Member

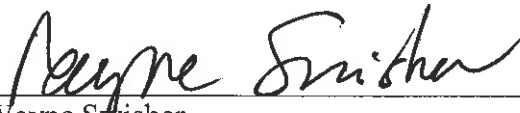


Diane Darland, Committee Member



Holly Brown-Borg, Committee Member

This dissertation is being submitted by the appointed advisory committee as having met all of the requirements of the School of Graduate Studies at the University of North Dakota and is hereby approved.



Wayne Swisher
Dean of the School of Graduate Studies

april 09, 2015

Date

PERMISSION

Title Molecular mechanisms underlying wing color pattern development in
Vanessa cardui and evolution of the polycomb repressive complex 2

Department Biology

Degree Doctor of Philosophy

In presenting this dissertation in partial fulfillment of the requirements for a graduate degree from the University of North Dakota, I agree that the library of this University shall make it freely available for inspection. I further agree that permission for extensive copying for scholarly purposes may be granted by the professors who supervised my dissertation work, or in their absences, by the Chairperson of the department or the dean of the School of Graduate Studies. It is understood that any copying or publication or other use of this dissertation or part thereof for financial gain shall not be allowed without my written permission. It is also understood that due recognition shall be given to me and to the University of North Dakota in any scholarly use which may be made of any material in my dissertation.

Heidi Connahs
May 7, 2015

TABLE OF CONTENTS

LIST OF FIGURES	x
LIST OF TABLES	xiii
ACKNOWLEDGMENTS	xv
ABSTRACT	xvii
CHAPTER	
I. WING COLOR PATTERNING, PHENOTYPIC PLASTICITY AND EPIGENETICS.....	1
Lepidoptera As A Model System For Understanding Color Pattern Development.....	2
The Development of Butterfly Wings and Color Pattern Elements.....	4
Phenotypic Plasticity of Wing Patterns.....	7
Why is Plasticity So Important in Evolution?.....	8
Environmental Effects on Trait Integration and Modularity	10
Molecular Mechanisms Underlying Phenotypic Plasticity.....	12
What is Epigenetics?.....	12
Research Objectives.....	15
II. TRANSCRIPTOME ANALYSIS OF THE GENE REGULATORY NETWORK UNDERLYING WING PATTERN DEVELOPMENT AND PIGMENTATION IN THE PAINTED LADY BUTTERFLY <i>VANESSA CARDUI</i>	20
Abstract.....	20

Introduction.....	21
Study System	27
Methods.....	28
Tissue Collection	28
cDNA Library Construction and Illumina Sequencing	29
Transcriptome Assembly and Sequence Annotation	29
Quantitative PCR Validation	30
Statistical Analyses	31
Results.....	32
Correlation of qPCR and RNA-seq Data	33
Wing GRN Patterning Genes.....	33
Ommochrome Pigmentation	38
Melanin Pigmentation.....	39
Discussion.....	42
Temporal Expression Patterns of the Wing GRN.....	42
Melanin Genes Up-regulated During Late Pupation	45
Ommochrome Gene Expression Mirrors the Wing GRN.....	46
Conservation and Divergence of Pigment Gene Expression Across Butterflies.....	47
Conclusions.....	48
III. PHENOTYPIC PLASTICITY REVEALS MODULARITY OF EYESPOT DEVELOPMENT IN THE PAINTED LADY BUTTERFLY <i>VANESSA CARDUI</i>	49
Abstract.....	49

Introduction.....	50
Methods.....	54
Butterfly Rearing and Experimental Setup.....	54
Eyespot Color Analysis.....	55
Statistical Analyses.....	56
Results.....	58
Eyespot Size.....	59
Black Border (Outer Ring).....	60
Yellow Pigment.....	61
Orange Pigment.....	63
Blue Scales.....	64
Black Focus.....	65
Tests for Modularity and Integration.....	67
Discussion.....	71
Eyespot Size Plasticity Varies Across Wing Segments.....	72
Inner Pigment Rings Are More Sensitive to Perturbations Than the Eye Spot Border.....	74
Do Gradient Models Adequately Explain How The Same Pigment Develops in Different Rings?.....	76
Temperature Shock Alters Patterns of Integration and Modularity.....	78
Conclusions.....	79
 IV. TEMPERATURE SHOCK AND HEPARIN ALTER EXPRESSION OF GENES INVOLVED IN EYESPOT DEVELOPMENT.....	 81
Abstract.....	81

Introduction.....	82
Methods.....	86
Transcriptome Analysis	86
Temperature and Heparin Experiment: Butterfly Rearing and Eyespot Dissections.....	87
RNA Isolation and Quantitative Real Time PCR	88
Statistical Analyses	90
Results.....	91
Expression of Polycomb Repressive Complex Genes During Wing Development.....	91
Comparison of Pupal Development Time and Pigmentation for Eyespot Dissections	92
Differential Expression of Patterning and Pigment Genes Across Eyespots.....	93
Effects of Temperature and Heparin on Pigment, Patterning and Polycomb Genes	95
Discussion.....	96
Expression Patterns of Pigment and Patterning Genes Across Hindwing Eyespots.....	97
Heparin Increases Expression of Patterning Genes, but Has No Effect on Genes Promoting Melanization.....	98
<i>Engrailed</i> , <i>sal</i> and <i>tan</i> are Thermosensitive Genes	101
Polycomb Repressive Complex Genes are Expressed During Wing Development	102
Conclusions.....	104

V.	MOLECULAR EVOLUTION OF AN EPIGENETIC SILENCER: POLYCOMB REPRESSIVE COMPLEX 2.....	105
	Abstract.....	105
	Introduction.....	106
	Function of the PRC2	108
	Origin and Evolution of the PRC2.....	110
	Research Questions.....	111
	Methods.....	112
	Mapping of Domains	114
	Gene Tree Comparisons.....	114
	Results and Discussion	114
	Nematodes Do Not Reconstruct The Monophyletic Relationship of Animals	116
	SANT1 Domain is Poorly Conserved in Primitive Taxa and Nematodes.....	122
	Low Conservation of the <i>Drosophila</i> EZ-ESC Binding Region.....	124
	Identification of Suz12 Homologs in Nematodes.....	124
	Evolutionary Conservation and Divergence in β-Blades and Loops of ESC.....	126
	Why Are Polycomb Genes So Different in Nematodes Compared to All Other Animals?.....	128
	Conclusions.....	132
VI.	EPILOGUE	134
	<i>Drosophila</i> Wing Gene Regulatory Network is Conserved in <i>V. cardui</i> and Peaks in Expression During the Larval and Early Pupal Stages	135

Eyespots Reveal Complex Patterns of Trait Integration and Modularity That Are Disrupted Following Plastic Responses to Environmental Perturbation	136
Environmental Perturbation Alters Expression of Patterning and Pigment Genes While Polycomb Genes Remain Unaffected.....	137
The Evolution of the Polycomb Repressive Complex 2 Shows Conservation Across Diverse Metazoans and Significant Divergence in Nematodes	138
APPENDICES	141
REFERENCES	205

LIST OF FIGURES

Figure	Page
1.1. Nymphalid groundplan described by Nijhout (2001)	3
1.2. Live pupal wing of <i>Vanessa cardui</i> at 3 days post-pupation, no pigmentation is visible at this stage	5
1.3. Patterning gene expression in butterfly eyespots	6
1.4. Dorsal and ventral wings of <i>Vanessa cardui</i> . The arrows point to the dorsal and ventral eyespots respectively	18
2.1. Bivariate analysis of fold change expression for RNA-seq and qPCR	34
2.2. RNA-seq and qPCR data for patterning and pigment genes showing fold change expression.	35
2.3. Wing gene regulatory network	37
2.4. RNA-seq expression patterns for the different functional groups of the wing GRN.	38
2.5. RNA-seq temporal expression patterns for ommochrome genes.	40
2.6. RNA-seq temporal expression patterns for melanin genes	41
3.1. Images of eyespots from the different treatment groups illustrating representative phenotypes	55
3.2. Color composition of eyespot in <i>Vanessa cardui</i>	56
3.3 Appendix. Bivariate analysis of the sum of all pigments in each eyespot relative to the total eyespot area	154
3.4. Eyespot size across all four eyespots in the different treatments	59
3.5. Area of the black border (outer ring) across all four eyespots in the different treatments.	60

3.6. Area of yellow pigment across all four eyespots in the different treatments.....	62
3.7. Area of orange pigment across all four eyespots in the different treatments.....	64
3.8. Area of blue pigment across all four eyespots in the different treatments.....	65
3.9. Area of the black focus across all four eyespots in the different treatments.	66
3.10. Hypotheses of integration and modularity for control eyespots based on overall phenotype data.	67
3.11 Appendix. Partial correlation matrices for eyespot size and proportion of each color ring.	156
3.12 Appendix. Edge exclusion deviance matrices based on partial correlations and calculated using the EED formula described in the methods.	157
3.13. Comparison of phenotypic correlations within and between control and temperature shock groups for eyespot size and pigment rings.	68
4.1. Appendix. Representative images of pupal hindwings at 6 days post-pupation prior to eyespot dissections.	160
4.2. Adult hindwing of <i>Vanessa cardui</i> with eyespots labelled.....	89
4.3. Expression of the polycomb repressive complex during larval and pupal wing development in <i>V. cardui</i>	92
4.4. Appendix Gene expression patterns across all three hindwing eyespots (Control only) for the patterning genes and melanin genes.	161
4.5. Expression of genes across eyespots for each treatment group for <i>V. cardui</i>	96
5.1. Major components of the PRC2 in invertebrates, showing the interaction between Ez and Suz12 and ESC, adapted from O’Meara & Simon, (2012).	108
5.2 Appendix. Reconstruction of the animal phylogeny based on the Tree of Life website	181
5.3. Phylogenetic tree of Enhancer of zeste.	117
5.4. Phylogenetic tree of ESC.	118

5.5. Phylogenetic tree of Suz12.	119
5.6. Multiple sequence alignment of the enhancer of zeste (EZ) SET domain.	121
5.7. Appendix. Amino acid alignment of the CXC domain from enhancer of zeste.	182
5.8. Appendix. Amino acid alignment of the SET domain from EZ.	183
5.9. Multiple sequence alignment of the enhancer of zeste (EZ) SANT domains (SANT1 top two panels, SANT2 bottom panel) for representatives across major animal groups.	123
5.10. Appendix. Amino acid alignment of the SANT1 domain from EZ.	186
5.11. Appendix. Amino acid alignment of the SANT2 domain from enhancer of zeste.	191
5.12. Appendix. ESC-EZ binding site identified in <i>Drosophila melanogaster</i> is poorly conserved across all taxa.	192
5.13. Multiple sequence alignment of the Suz12 VEFS box.	125
5.14. Appendix. Amino acid alignment of the VEFS domain from Suz12.	193
5.15. Appendix. Amino acid alignment of the Zinc finger domain from Suz12.	197
5.16. WD-40 repeats of the ESC protein for select taxa across major animal groups...	129
5.17. Appendix. Amino acid alignment of the WD40 repeats from Esc.	198

LIST OF TABLES

Table	Page
2.1. Appendix. NCBI Blastn and Blastp results for <i>Vanessa cardui</i> transcripts following transcriptome assembly with CLC Genomics Workbench and multiblast (Blastx) against the <i>Drosophila</i> peptide database (FlyBase).	142
2.2. Appendix. Primers used for qPCR validation.....	149
2.3. Appendix. Summary of de-novo transcriptome assembly performed using CLC Genomics.....	150
2.4. Appendix. References for classification of functional groups in the wing GRN. ...	151
3.1. Appendix. Dunn-Šidák corrections for multiple comparisons following 2-way ANOVA for treatment x eyespot interactions comparing changes in pigment area (cm ²).	153
3.2. Percent difference in area (cm ²) for eyespot size and pigmentation in response to temperature shock (37°C, 48 hrs.) relative to the control.....	70
3.3. Total number of edges across all traits between each pair of eyespots for the Control and temperature shock (37°C, 48 hrs.) groups.	71
4.1. Age of pupae at time of eyespot dissection and total duration of pupation for butterflies reared to adults across treatment groups.	88
4.2. Appendix. List of primer pairs used for quantitative real time PCR.	158
4.3. Results from two-way ANCOVA of gene expression. Main effects include treatment (control, temperature, heparin), eyespot number (1, 2, 3) and the interaction between treatment and eyespot.	94
5.1A. Appendix. Taxon list for Enhancer of zeste (Ez) with accession numbers for NCBI and Uniprot.	162
5.1B. Appendix. Taxon list for Suppressor of zeste (Suz12) with accession numbers for NCBI and Uniprot. NF indicates no record found in that database. ...	166

5.1C. Appendix. Taxon list for Embryonic Sex Combs (ESC) with accession numbers for NCBI and Uniprot.	169
5.2A. Appendix. Domain sequence divergence for Enhancer of zeste.....	173
5.2B. Appendix. Domain sequence divergence for Suz12.	176
5.2C. Appendix. Domain sequence divergence for ESC.	177

ACKNOWLEDGMENTS

This PhD has been a very long journey and I have many people to thank who have provided continual support and encouragement along the way. First and foremost I wish to express my sincere gratitude and appreciation to my PhD advisors, Dr. Rebecca Simmons and Dr. Turk Rhen. I thank Dr. Simmons for her warm encouragement and providing a cheerful, fun and supportive environment to work in. I especially thank her for training me in phylogenetics, assisting with rearing hundreds of butterflies including working with me during weekends. I also thank Dr. Rhen for his extraordinary patience, the many long hours he committed to training me in transcriptomics, qPCR and statistical analyses, without which I would not have been able to conduct this research. I also thank him for introducing me to the exciting field of epigenetics and for the many interesting conversations we had. I am extremely grateful to both advisors for always being available whenever I needed help. I would also like to thank my committee members Dr. Diane Darland and Dr. Brian Darby for their insightful comments and guidance and their generosity in allowing me to borrow whatever equipment I needed from their labs! I also thank Dr. Holly Brown-Borg for agreeing to be part of my advisory committee. I would like to thank the graduate students, in particular Katherine Hernandez, Justin Burum, Riley McGlynn and Stephanie Snyder for the many laughs we had, even their persistent teasing about my ‘love’ of the polar vortex and for helping me to survive in North

Dakota! I would also like to thank other friends who have been supportive over the years, including Christina Gomez-Mira, Kate Scheurer, Genoveva Rodriguez, Mark Tobler and also Rebecca Forkner who has always believed in me and has been an inspiring mentor. I am also very grateful to Ken and Maureen Drees for providing me a place to live when I first arrived and for their continued support. I am also very grateful to Archana Dhasarathy for being a wonderful teacher, and for her support in helping me land my dream post-doc! I am also thankful to Annette Aiello who has been an important mentor to me over the last 12 years and has provided many letters of recommendation that have helped me get to where I am today. I also wish to thank those who provided funding for my research including Graduate Women in Science (GWIS), Adele Lewis Grant Fellowship, Dr. Diana Wheeler, Vice President of Research office at UND, NSF Experimental Program to Stimulate Competitive Research (EPSCoR) Doctoral Dissertation Award, the Biology Department Academic Programs and Student Awards Committee (APSAC), and additionally, the School of Graduate Studies for the Summer Research Fellowships. I am also grateful to Dr. Ike Schlosser and Dr. Jeffery Carmichel for their support during my time here. I also want to thank all of the undergraduate students who assisted with my research including Joslin Seidel, Kasey Chelemedos, Kiara Pochardt, Jace Kussler, Boma Afonya and Whitney Redman. I also cannot finish without thanking Kathleen Peterson, who has helped me through the most difficult times and has been a guiding light during my time here. Finally, I wish to express my deepest thanks to my wonderful family who have been an enormous support to me throughout this long journey. I hope this will make them very proud.

This dissertation is dedicated to the memory of my beautiful mother
(1956 – 2011)

ABSTRACT

In this dissertation, I present four papers. Three explore different aspects of wing color pattern development in the painted lady butterfly, *Vanessa cardui*; while the fourth examines the evolution of an epigenetic silencer complex across invertebrate animals. In the first paper, I used transcriptomics to identify patterning genes from the *Drosophila* wing gene regulatory network (GRN) in larval and pupal wings of *V. cardui* and to examine how temporal expression dynamics of this gene network correspond to expression of ommochrome and melanin pigment genes. This study identified key developmental periods of gene upregulation and highlights the temporal separation between peak expression of patterning and melanin pigment genes. In the second paper, I present evidence that hind wing eyespots of *V. cardui* exhibit phenotypic plasticity. Using morphometrics, I quantified how temperature shock and heparin modify eyespot size and pigment ring composition. This information is used to examine whether eyespot plasticity was a function of trait integration or modularity. In the third paper, I used qPCR to explore the role that epigenetic mechanisms may play in phenotypic plasticity of *V. cardui* eyespots. I examined expression of an epigenetic silencer, the polycomb repressive complex (PRC) in developing wings and in modified eyespots at 6 days post-pupation. I present evidence that the PRC is expressed during butterfly wing development and exhibits a similar pattern of expression to the wing GRN. Polycomb genes were not differentially expressed in eyespots modified by temperature shock and heparin sulfate;

however, expression of several patterning genes was altered by these treatments. In the final paper, I present a comprehensive phylogenetic analysis of the PRC2 across non-vertebrate animals. This analysis revealed that the evolutionary history of the PRC2 does not reconstruct the known phylogeny of animals, due to significant sequence divergence in the nematode lineage. Thus, PRC2 has undergone significant evolutionary changes in nematodes that may be a consequence of Hox gene depletion and re-organization in this lineage.

CHAPTER I
WING COLOR PATTERNING, PHENOTYPIC PLASTICITY AND
EPIGENETICS

Introduction

An important challenge in evo-devo is to understand the evolution of morphological diversity by identifying the molecular and environmental factors that promote the development of novel phenotypes. The extraordinary diversity of animal color patterns is perhaps one of the most dramatic examples of morphological diversity and represents an ideal system to address such questions. Colors and patterns mediate how animals interact with their environment by serving as visual communication signals to attract mates and deter predators. Some animals including insects, fish and amphibians use aposematic coloration to warn of their toxicity (Prudic et al., 2006), or alternatively to allure mates (Maan & Cummings, 2009). Camouflage enables leaf-mimic insects to blend perfectly with their surroundings (Vallin et al., 2006). In addition to ecological benefits, color patterns confer physiological benefits such as protection from ultra-violet light (Protas & Patel, 2008). Thus, a variety of selective forces from the environment help shape the evolution of color pattern variation. Although the ecological function of animal pigmentation has been well documented across a broad range of taxa, precisely how color patterns develop at the molecular level and how the environment influences these

processes remains poorly understood (Wittkopp et al., 2002; Miyazawa, et al. 2010; Werner et al. 2010).

Lepidoptera As A Model System for Understanding Color Pattern Development

Lepidoptera are an excellent model for studying the evolution and development of animal color patterning. Lepidoptera are an extremely diverse order with approximately 20,000 described species of butterflies and between 150,000 -250,000 described species of moths (Brock & Kaufman, 2006). The enormous diversity observed in this group is characterized by a stunning array of wing color patterns. These patterns are generated by colored scales that are modified sensory bristles; a synapomorphy that gave the order its scientific name; scaled wings (Knüttel & Fiedler, 2001; Mcmillan et al. 2002). Coloration can be the result of pigmentation that is synthesized *de novo* during scale development in the pupa, or it can develop as structural modifications that result in iridescent scales.

In most plants and animals, color patterns are arranged randomly and are quite variable (Nijhout, 2001; Ohno & Otaki, 2012). This is especially the case in vertebrates; for example, the markings of leopards are highly variable within a species, and even between the left and right sides of an individual (Nijhout, 2001; Miyazawa et al., 2010). In contrast, wing color patterning in Lepidoptera is highly organized and essentially represents two-dimensional anatomical structures (Ohno and Otaki 2012). Color patterns are generally fixed and can be used as characters to distinguish different species, which have subtle modifications of pattern elements. Thus, wing color patterns are likely to follow distinct developmental pathways with individual pattern elements developing autonomously from adjacent pattern elements. This allows an uncoupling of changes across the different elements, increasing the opportunity for diversification of color

patterns. These specific characteristics of butterfly wing patterns make them an ideal model system for dissecting the processes underlying morphological diversification.

Wing color patterns are highly modular and their diversity can be thought of as modifications to the Nymphalid groundplan, a system that has been developed to describe the general symmetrical color patterns of Nymphalid butterflies (Nijhout, 2001).

Although this system was originally based on Nymphalids, its utility has been extended to many other Lepidoptera, including moths. Wing color patterns are composed of at least three symmetry systems laying side by side in parallel form from the anterior costal margin to the posterior hind margin; 1. the border symmetry system, 2. the central symmetry system and 3. the basal symmetry system near the proximal base site of the wing (Figure 1.1). Each symmetry system is a collection of elements. These elements appear to be regulated by eyespot focal organizers as well as the wing margin, which also acts as an organizing center (Nijhout 1991).

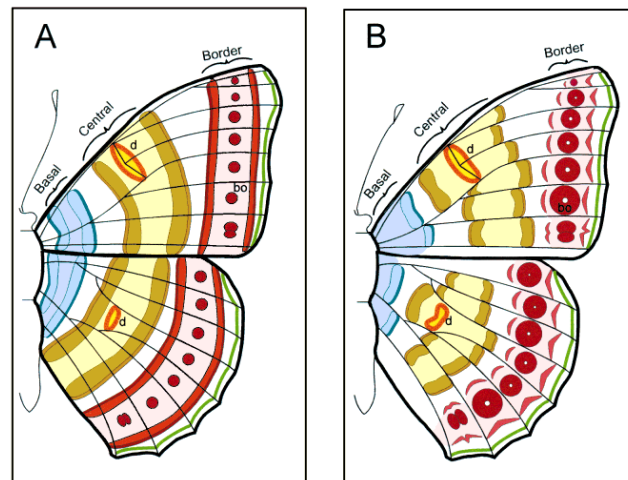


Figure 1.1 Nymphalid groundplan described by Nijhout (2001). Panel A emphasizes the vertical pattern elements. Panel B emphasizes the wing veins that break up the symmetry system.

The Development of Butterfly Wings and Color Pattern Elements

The development of butterfly wings begins inside the immature caterpillar where imaginal discs within the meso and meta-thorax will differentiate into fore and hindwings respectively. These imaginal discs develop as an invagination of the body wall and form an undifferentiated epithelial bilayer representing the dorsal and ventral wing surfaces (Macdonald et al., 2010; Cho & Nijhout, 2013; Iwata et al., 2014). During the early larval period, the imaginal discs grow continuously until the late larval stage where they rapidly expand in size (Kremen & Nijhout, 1998; Nijhout et al., 2007). Following pupation, the wings are extruded through the body wall to assume their final adult position (Cho & Nijhout, 2013). During early pupation, the epithelial wing cells divide and differentiate into scale and socket cells which are regularly arranged in parallel rows (Figure 1.2). Array formation of scale cells has been shown to be complete by approximately two days post-pupation (Iwata et al., 2014). Each individual scale develops as a flattened projection of a scale-building cell that will become colored as a single pigment during later stages of pupation. Why each individual scale develops one specific pigment out of all the possible range of pigments present on the wing is one of the most fascinating yet poorly understood aspects of wing color pattern development. The specification of a single pigment suggests that molecular mechanisms must be employed that repress alternative pigmentation pathways.

The development of wing color patterns is initiated during the late larval and early pupal stages, when scale cell fate determination and differentiation takes place (Otaki et al. 2005). Although pigments are not visible during these early stages, gene expression

patterns have already established the template of wing color patterns (Brunetti et al., 2001) (Figure 1.3).

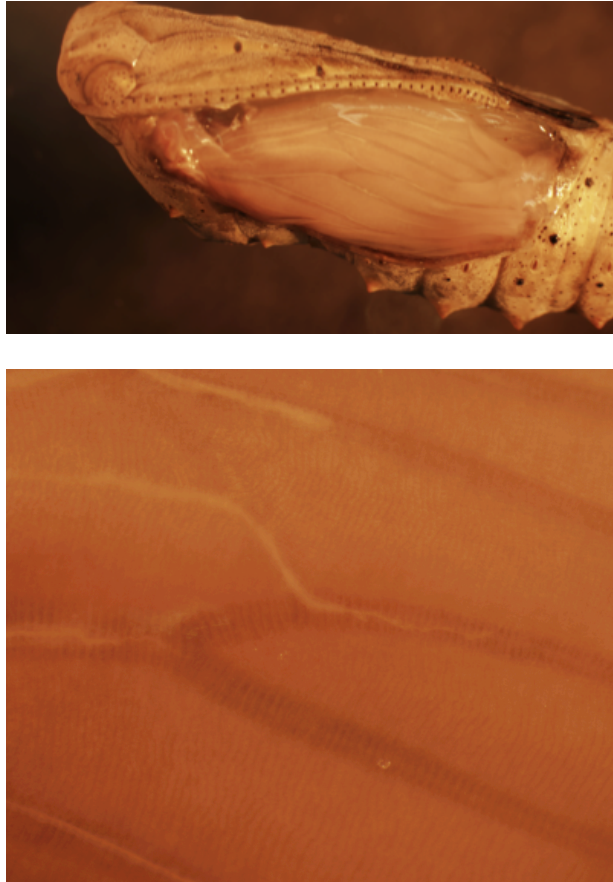


Figure 1.2 Live pupal wing of *Vanessa cardui* at 3 days post-pupation, no pigmentation is visible at this stage. Bottom image shows a close-up of the wing revealing thousands of tiny parallel rows of developing scale cells that form perpendicular to the proximal-distal axis of the wing. The branching lines represent developing wing veins also shown clearly in the top image.

Pattern elements are formed in response to positional information from a putative morphogen signal that is emitted from organizing centers through passive diffusion (Nijhout 1991). Color pattern formation can be divided into four sequential steps: 1. Signaling, 2. Reception, 3. Interpretation, and 4. Expression (Otaki et al., 2005a; Otaki, 2008c). The signaling step is a process of morphogen production and emission from an

organizing center, such as an eyespot center, resulting in the establishment of a stable morphogen gradient. These signals are received and interpreted by neighboring scale cells through an array of receptors and signal transducers. This process ultimately results in pigment biosynthesis and color pattern development. Many aspects of this signaling process remain a mystery, including the link between morphogen signaling and expression of patterning genes, and how these early pupal events trigger expression of genes involved in pigment formation at subsequent developmental stages.

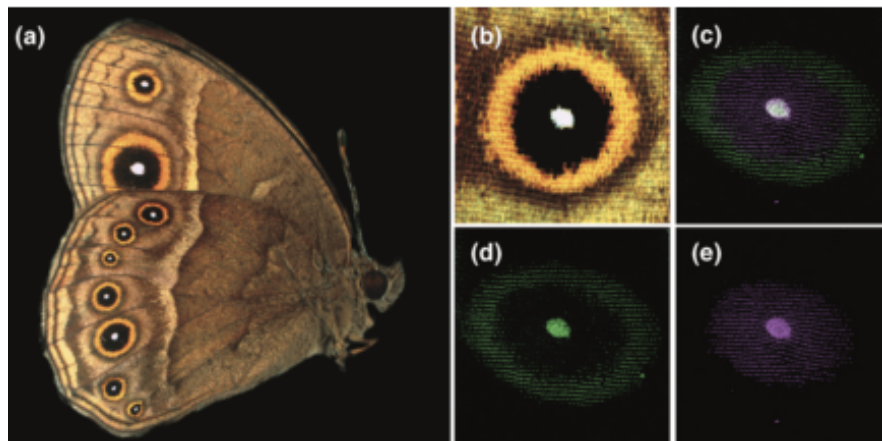


Figure 1.3 Patterning gene expression in butterfly eyespots. Extracted from Brunetti et al. (2001) showing immunolabelling of patterning genes (En/Inv - green) and Spalt - purple) in the eyespots of the African butterfly, *Bicyclus anynana* during the first 24 hours of pupation.

Butterfly eyespots are the most widely studied color pattern elements. Eyespots are an evolutionary novelty in Lepidoptera, functioning as a visual signal for mating and for predator deterrence. Intriguingly, some of the patterning genes identified in butterfly eyespots are the same as those expressed in the regulatory network for wing development in the fruit fly (Carroll et al., 1994; Brunetti et al., 2001; Carroll et al., 2001). The finding that genes can be re-used and co-opted for novel functions has been one of the most

fascinating discoveries in the field of Evo-Devo (Carroll et al., 2001; Monteiro, 2012). In the case of butterfly wings, it appears that wing color patterns may have evolved via co-option of the wing gene regulatory network. Thus, ectopic expression of wing patterning genes outside of their normal expression domains may have produced the diversity of wing color patterns that we observe today. Currently, little is known about this wing gene regulatory network in butterfly wings, other than the spatial expression of individual genes in pattern elements such as eyespots (Brunetti et al., 2001). Therefore, we do not have a clear understanding of the composition of this network, the temporal expression dynamics, or correspondence with pigment gene expression during wing color pattern development.

Phenotypic Plasticity of Wing Patterns

Although wing color patterns are largely fixed and can be used to distinguish different species, many butterflies exhibit phenotypic plasticity in response to environmental conditions. The well-studied African butterfly *Bicyclus anynana* displays seasonal variation in wing color patterns (Roskam & Brakefield, 1999). The wet season butterflies have wings with conspicuous eyespots to deter active summer predators, whereas wings of dry season butterflies display cryptic coloration that matches the background of dead leaves (Lyytinen et al. 2004). Thus, plasticity enables these butterflies to express the optimal phenotype for particular environmental conditions (de Jong et al. 2010) by revealing hidden phenotypic and genetic variability. Environmental changes may alter the timing or intensity of the initial signaling step, resulting in heterochronic uncoupling of receptors and transducers. This alteration, in turn, could lead to changes in gene expression or even expression of novel genes that modify wing color patterns.

It is important to emphasize that wing pattern modifications occur without any changes to the underlying genetic code. These environmentally induced changes in gene expression may arise simply as a byproduct of a generalized stress response with no immediate ecological benefit, or may lead to the evolution of an adaptive phenotype, as in the case of *B. anynana*. Regardless of whether the phenotype is adaptive or not, plasticity plays a fundamental role in promoting morphological diversification in butterfly wings.

Why is Plasticity So Important in Evolution?

One of the most interesting ideas to emerge about phenotypic plasticity is the potential role it may play in facilitating speciation (West-Eberhard, 1989; Pigliucci et al. 2006; Pfennig et al., 2010; Fitzpatrick, 2012; Minelli & Fusco, 2012). Speciation can occur when selection on phenotypic or behavioral traits drives a population into genetically isolated groups resulting in assortative mating. Given that selection acts on phenotypes, rather than genotypes, any phenotypic variation has the potential to promote evolutionary change, whether it is due to mutation or environmentally induced plasticity (West-Eberhard, 2005). It has been proposed that when plastic traits are adaptive and persist in a population they may eventually become constitutively expressed even in the absence of the environmental factor that induced the plasticity (West-Eberhard, 2005). The mechanism by which plastic traits become canalized is known as genetic assimilation. Genetic assimilation is a special case of a more general phenomenon known as genetic accommodation which refers to the evolution of either plastic or canalized traits (Braendle & Flatt, 2006; Pfennig & Ehrenreich 2014).

These ideas, introduced by Waddington (1942) and West-Eberhard (2005), have remained highly controversial due in part to limited empirical evidence in natural populations. Further, at the core of these ideas is the notion that plasticity is heritable and can promote the evolution of adaptive phenotypes (Braendle & Flatt, 2006; Ghalambor et al. 2007; Pfennig et al., 2010; Wund, 2012). In addition, adaptive plasticity is typically assumed to dampen divergent selection, because different genotypes may converge on the same phenotype. Convergent phenotypes would reduce the power of selection to filter out particular genotypes as more or less fit (Crispo, 2008; Pfennig et al., 2010; Fitzpatrick, 2012). In this sense, adaptive plasticity appears to act as a buffer against selection by inducing phenotypes that can survive in a range of conditions, thus constraining evolutionary processes by shielding genetic variation (Ghalambor et al. 2007; Wund, 2012). It has been argued that adaptive plasticity can accelerate evolution compared to random mutations because an environmental factor can affect numerous individuals simultaneously, increasing the chance that selection can act on variation in regulatory genes controlling favorable plastic traits (West-Eberhard, 2005; Pfennig et al. 2010; Shaw et al., 2014). In this way, phenotypic plasticity influences which particular phenotypes are exposed to selection. If the induced phenotype is adaptive, and there is genetic variation for plasticity, these environmentally induced regulatory networks will be selected and may eventually become genetically assimilated (Schlichting & Smith, 2002; Pigliucci et al., 2006).

The butterfly genus, *Vanessa* (Nymphalidae), seems to exemplify the idea that phenotypic plasticity can cause individuals in one species to resemble related species, occupying different environments (Price et al. 2003; Pfennig et. al 2010; Wund, 2012).

Vanessa butterflies are characterized by several orange pattern elements that form a large patch in the middle of the forewings, the size of which varies across different species (Otaki 2008; Hiyama et al. 2012). It is possible to modify the size of this orange patch by exposing butterfly pupae to temperature shock treatments. Interestingly, the color patterns of modified individuals can be arranged in a progressively linear series that resemble the natural color patterns of other, related species (Otaki & Yamamoto, 2004). Otaki (2008a) speculated that the ancestral species of *Vanessa* was isolated to mountainous regions and therefore exposed to fluctuating temperatures resulting in wing color pattern modifications. Otaki (2008a) proposed that these color pattern modifications represent ancestral expressions of phenotypic plasticity that have since become genetically assimilated in derived species. These color pattern modifications are induced by different degrees of cold shock and are thought to be a physiological byproduct of a cold shock hormone (Otaki, 2008a). Selection for increased cold shock resistance would also lead to modified wing patterns due to trait integration (Otaki, 2008a). Thus, phenotypic plasticity of an ancestral species may have facilitated divergence during the course of speciation in *Vanessa* (Otaki 2008).

Environmental Effects on Trait Integration and Modularity

Though phenotypic plasticity has been widely studied, most research has focused on examining plasticity of a single trait. In contrast, an organisms' phenotype comprises a multitude of traits that may even be correlated with each other (phenotypic integration). Exposure to novel environments may affect plasticity of multiple traits, particularly if traits are integrated (Price et al. 2003; Pfennig et al. 2010; Montague et al. 2012). Therefore, to fully understand the role of phenotypic plasticity in generating novel

phenotypes, it is also important to consider how the environment influences suites of developmental traits, and whether correlations among traits vary based on environmental conditions. Studies in plants and birds have shown that variation in plasticity among different traits can alter patterns of integration and modularity (Schlichting, 1989; Montague et al., 2012), which could have important evolutionary consequences by changing the outcome of selection (Schlichting, 1989; Schlichting & Smith, 2002; Montague et al., 2012). Despite the common occurrence of phenotypic plasticity in butterflies, there are no studies examining the impact of environmental conditions on patterns of modularity and integration. Studies that have examined modularity and integration in *B. anynana* reveal complex patterns of both concerted and independent evolution across different eyespots (Beldade et al. 2002; Allen, 2008). Identifying integrated versus modular traits can be instructive for understanding the degree of morphological flexibility within a phenotype and potential evolvability. Modular traits are presumed to evolve independently allowing selection to optimize these traits. Integrated traits are more likely to be constrained by pleiotropy, whereas, modularity has been proposed as a mechanism that promotes plasticity by facilitating the independent evolution of novel traits and alternative phenotypes (Snell-Rood et al. 2010). In contrast, research on clonal plants has demonstrated integration can reduce plasticity, but it can also enhance plasticity, by increasing the availability of resources (Alpert, 1999). Underlying processes promoting plasticity seem to be related to the specific nuances of the environment and the genetic background of the organism. Plasticity may vary across different levels of morphological organization and environmental conditions. Thus, it is

important to understand the potential roles of modularity and integration in promoting/constraining phenotypic plasticity, and ultimately, morphological diversity.

Molecular Mechanisms Underlying Phenotypic Plasticity

The idea of phenotypic plasticity has become an important concept in understanding how organisms cope with environmental variation (West-Eberhard, 1989). Until recently, these expressions of plasticity were thought to be ecologically important, but not heritable. Further, phenotypic plasticity was once considered a nuisance in evolutionary biology (Pigliucci, 2005; Forsman, 2014); however, its significance to evolution has since been argued by numerous authors including: West-Eberhard (1989, 2003, 2005), Schlichting, (1986) and Pigliucci, (2001). This discussion resulted in wide acceptance of the importance of phenotypic plasticity as an ecological phenomenon, though its evolutionary importance still remains controversial. Although phenotypic plasticity has been documented across all domains of life, we still do not have a clear understanding how alternative phenotypes are generated at the molecular level (Wray et al., 2003; Pfennig & Ehrenreich, 2014). Because the genetic make-up of an organism, does not change, epigenetic regulation is a plausible explanation for the molecular basis of phenotypic plasticity (Jablonka & Lamb, 2002; Champagne, 2013; Geng et al., 2013).

What is Epigenetics?

The term epigenetics was first coined by Waddington (1942) to explain the interactions among genes and differential gene expression during development. Today, epigenetics refers to the process by which heritable changes in gene expression alter an organisms' phenotype without any changes to the DNA coding sequence. Gene expression is regulated by transcription factors, cis-regulatory elements, non-coding RNAs, DNA

methylation and histone modifications. It has become evident that the structure of chromatin contributes to transcriptional regulation. The chromatin structure of DNA is dynamic and can be tightened or unwound by a suite of enzymes that are involved in the addition and removal of various chemical moieties to histones (e.g., methyl or acetyl groups, ubiquitin). Generally speaking, DNA methylation of cytosines is involved in silencing genes, while histone acetylation is involved in unraveling the chromatin, enabling the activation of genes (Feil & Fraga, 2011). This dynamic process of unwinding and rewinding the chromatin appears to be sensitive to endogenous and exogenous signals (Feil & Fraga, 2011). One of the most well studied histone modifiers is polycomb repressive complex (PRC2), which catalyzes the trimethylation of lysine 27 on histone 3 resulting in chromatin compaction and gene silencing (Margueron & Reinberg, 2011; Jeffrey et al., 2013). The PRC2 has received significant attention because it is known to regulate expression domains of Hox genes, which play a critical role in animal development and morphogenesis (Müller et al., 2002). Additionally, the PRC2 is involved in stem cell biology and cancer (Müller et al., 2002).

Evidence is accumulating that nutrition, environmental chemicals, stressors, and pharmaceutical agents can affect the epigenetic state (Feil & Fraga, 2011; Crews & Gore, 2012). For example, maternal diet has been shown to influence the epigenetic state of a transposable element in the Agouti gene of A_{vy} mice as well as their offspring (Waterland & Jirtle, 2003; Dolinoy, 2008). The agouti gene regulates the production of pigment in individual hair follicles. If this gene is not properly methylated the mice develop yellow instead of brown fur and are obese and prone to diabetes and cancer. Providing a methyl-enriched diet to female mice before and during pregnancy permanently increases DNA

methylation of the agouti gene in offspring and largely reverses the deleterious phenotype (Waterland, & Jirtle, 2003). These results show that environmental factors can influence animal coloration (= phenotypic plasticity) via epigenetic mechanisms. In contrast, most research investigating molecular and developmental mechanisms of animal color patterns have examined this question by looking at the genetic basis of these traits. For example, research on *Drosophila* and Lepidoptera have revealed the genetic basis for wing patterning including a variety of structural and regulatory genes that control pigmentation (Wittkopp et al., 2002; Reed et al., 2011; Tong et al. 2012). Similarly research on mice and zebra fish have also revealed a suite of genes involved in coloration (Bennett & Lamoreux, 2003; Quigley et al., 2004). Despite this work, many colors and patterns displayed by animals are induced by environmental conditions. Some animals change color instantaneously to match their surroundings (flounder and flatfish; Protas & Patel, 2008), while other animals produce distinctive color morphs in response to environmental stressors (polyphenism in moths and butterflies; Rountree & Nijhout, 1995; Nijhout, 2003). These observations suggest that the molecular mechanisms regulating these changes are flexible and sensitive to the environment. Given that environmental changes are known to alter color patterns, I decided to examine whether epigenetic mechanisms explain phenotypic plasticity in animal coloration.

Although epigenetic regulation of color patterning has not been investigated, research suggests that epigenetic mechanisms may underlie developmental plasticity. DNA methylation has been shown to be critical for caste development in honeybees (Lyko et al., 2010; Weiner & Toth, 2012). Honeybee larvae are genetically identical and are fed a diet of royal jelly. Larvae that are fed more royal jelly develop into queens while

those fed less royal jelly develop into workers. When researchers down-regulated DNMT3, an important DNA methyltransferase, in honeybee larvae, all the treated larvae developed into queens (Lyko et al., 2010). A component in royal jelly may function by repressing DNMT3 preventing methylation of genes involved in reproduction (Kucharski et al., 2008). Epigenetic mechanisms may also be involved in regulating the switch between winged and wingless morphs in pea aphids (Srinivasan & Brisson, 2012). These examples provide tantalizing clues of the potential importance of epigenetic mechanisms in shaping morphological diversity in insects. All of these studies have focused on the role of DNA methylation in regulating plasticity; however, chromatin remodelers such as the PRC2 may also be involved. Studies of the PRC have been focused primarily in vertebrate systems with an emphasis on biomedical applications (Cao & Zhang, 2004; Sparmann & van Lohuizen, 2006; Willert & Nusse, 2012; Zeng et al., 2012). With the exception of *Drosophila* little is known about this complex in other invertebrates including its evolutionary history, patterns of expression during development, and its potential role in regulating phenotypic plasticity. Further insight into the PRC in non-model organisms will help diversify current knowledge on this important epigenetic regulator.

Research Objectives

The introduction outlined above provides an overview to four major areas that I explore in this dissertation, 1) the gene regulatory network underlying butterfly wing color pattern development, 2) phenotypic plasticity of wing color patterns, 3) patterns of PRC expression in butterfly wings and 4) evolution of the polycomb repressive complex 2 in invertebrates. The study system for this research is the cosmopolitan butterfly, *Vanessa*

cardui (Linnaeus 1758) (Nymphalidae) commonly known as the Painted Lady. *Vanessa cardui* is a long distance migrant and is one of the most widespread of all butterflies, occurring on every continent with the exception of Antarctica and South America (Brock & Kaufman, 2006; Wahlberg & Rubinoff, 2011). *Vanessa cardui* is a popular and well-known butterfly particularly among children largely because it is colorful, relatively cheap to purchase and easy to rear in captivity. For this reason, it has served as an important educational tool for learning about lepidopteran lifecycles and is also an ideal system for developmental studies. Interestingly, NASA selected *V. cardui* along with the Monarch (*Danaus plexippus*) as the first butterflies to travel into space as part of an educational outreach experiment to observe their development in microgravity (www.nasa.gov/mission_pages/station/expeditions/expedition22/butterflies). Thus, *V. cardui* has served as a valuable educational tool to inspire student interest in science.

Most scientific research on *V. cardui* has focused on their migration ability (Stefanescu et al., 2007; Brattström et al., 2008), ecology (Bowers, 1998; O'Neill et al., 2010) and morphology of the dorsal wing patterns (Otaki & Yamamoto, 2004; Otaki, 2007). There are very few studies examining the molecular basis of wing color pattern development. Most work investigating the molecular basis of wing color patterning has been conducted in *B. anynana*, *Junonia coenia* and *Heliconius* butterflies (Carroll et al., 1994; Monteiro et al., 1994; Beldade, et al., 2006; Joron et al. 2006). Studies from *B. anynana* and *J. coenia* have revealed that some genes from the regulatory network for *Drosophila* wing development are also expressed in butterfly eyespots and other wing regions. While spatial expression patterns for some of these genes have been described in butterflies, there has been no investigation on the temporal dynamics of this network in

butterfly wings. The genes in this network are known as patterning genes, as they define specific regions on the wings of *Drosophila* and also appear to define aspects of pattern elements on butterfly wings, such as the colored rings of the eyespots (Brunetti et al. 2001). Research suggests that these patterning genes must be involved in regulating the expression of downstream pigment genes (Beldade & Brakefield, 2002). Thus, it would be valuable to determine the timeline for expression of patterning and pigment networks, and examine whether these networks overlap during wing color pattern development. If so, this would suggest that patterning genes directly regulate pigment genes. If they are temporally separated, this would suggest regulation is indirect. In Chapter 2, I utilize a transcriptomic approach to investigate temporal dynamics of the *Drosophila* wing gene regulatory network during wing color pattern development in *V. cardui*. To accomplish this, I sampled imaginal discs at multiple time points during larval and pupal development. In addition, I also examined expression patterns of two different pigmentation pathways. The melanin pathway is involved in producing brown and black pigments while the ommochrome pathway is involved in the production of red, yellow and orange pigments. This research provides the first expression timeline comparing the wing GRN and pigment pathways during wing pattern development in butterflies. These results also provide molecular resources on the entire wing transcriptome for *V. cardui*.

Previous work reveals *V. cardui*, like other species of *Vanessa*, exhibit phenotypic plasticity of dorsal wing patterns when exposed to temperature shock and pharmacological agents during early pupation (Otaki 2008). Less attention has been paid to the ventral hindwings, where these butterflies display a series of beautiful, complex eyespots composed of different colored rings (Figure 1.4).



Figure 1.4 Dorsal and ventral wings of *Vanessa cardui*. The arrows point to the dorsal and ventral eyespots respectively. Image extracted from www.jardinsauvage.fr

Nijhout (1984) demonstrated that these eyespots exhibit phenotypic plasticity when exposed to cold shock; however, there has been no detailed quantitative analysis of the response of these eyespots to environmental perturbation. It is also not known if the different eyespots exhibit similar responses; for example, if eyespot size or color composition is altered or whether particular pigments or eyespots are sensitive to perturbations. Furthermore, while studies have examined patterns of integration and modularity of eyespots, they have only been conducted in a single butterfly species, *B. anynana*. It is also unclear if these patterns are disrupted when butterflies are exposed to novel environments. In Chapter 3, I conduct a quantitative analysis on the response of *V. cardui* hindwing eyespots to environmental perturbation including heat shock and injection of heparin sulfate, which has been shown to mimic the effects of cold shock in butterflies (Serfas & Carroll, 2005). This study also compares patterns of integration and modularity of eyespot size and coloration following pupal exposure to heat shock. This information provides interesting insights into patterns of plasticity, integration and modularity across different eyespots and between the colored rings. These changes could influence the response of wing patterns to selection when exposed to novel environments.

In Chapter 4, I expand on this research to examine the molecular basis underlying eyespot plasticity in *V. cardui* following heat shock and heparin injection. I use quantitative PCR to examine expression of a select group of genes from the wing gene regulatory network and genes involved in the melanin pigmentation pathway. Furthermore, I examine expression of polycomb repressive complex 2 during wing color pattern development and across eyespots exposed to environmental perturbation to explore whether epigenetic mechanisms are involved in regulating wing color patterns. This research provides information on newly identified genes expressed in butterfly eyespots including expression patterns of an epigenetic silencer, the PRC2.

Finally in Chapter 5, I take a broad approach in examining processes regulating morphological diversity by investigating the evolution of the PRC2 across a wide diversity of invertebrate animals. I examine how this epigenetic silencer has evolved from the earliest extant metazoans, including sponges and comb jellies, through to the pre-vertebrate lineages, including tunicates and cephalochordates. This analysis encompasses three of the core subunits of the PRC2 to examine if the evolution of these closely interacting subunits exhibit similar evolutionary histories and if their evolution mirrors the known phylogeny of animals. Lastly, I explore the evolution of the major domains that comprise each of these subunits and examine whether these functional units have remained highly conserved or diverged significantly across this morphologically diverse group of animals.

CHAPTER II

**TRANSCRIPTOME ANALYSIS OF THE GENE REGULATORY
NETWORK UNDERLYING WING DEVELOPMENT AND
PIGMENTATION IN THE PAINTED LADY BUTTERFLY, *VANESSA
CARDUI*.**

Abstract

Introduction: Butterfly wing color patterns are an important model system for understanding the evolution and development of morphological diversity and animal pigmentation. Color patterns develop from a complex network composed of highly conserved patterning genes and pigmentation pathways. Patterning genes are involved in regulating pigment synthesis however the temporal expression dynamics of these interacting networks is poorly understood. Here, we employ next generation sequencing to describe expression patterns of the wing gene regulatory network (GRN) and look for evidence of correlated expression with genes involved in pigmentation.

Results: Homologs of genes involved in wing development in *Drosophila* are expressed in the developing wings of *Vanessa cardui*. Most of these genes exhibit peak levels in expression during the late larval and early pupal stages and then decline throughout pupal development. The most highly expressed genes included the Hox cofactor *extradenticle*, the selector gene *vestigial* and the signaling gene *serrate*. *Serum response factor* and

achaete-scute, which are involved with wing vein positioning and scale building, were expressed at the lowest levels across all developmental stages. Ommochrome pigment genes exhibit a similar expression pattern to the wing GRN, with the exception of *kynurenine formamidase*, which increased one day prior to butterfly eclosion. In contrast, expression of genes involved in melanin synthesis increase from larval to pupal development with the highest levels occurring one day prior to eclosion.

Conclusions: Here we provide a detailed expression timeline for all major genes involved in wing patterning, melanin and ommochrome pigmentation. Our results reveal that patterning genes display coordinated expression patterns with other members of the network despite significant differences in function, and exhibit a developmental peak that corresponds with ommochrome but not melanin gene expression.

Introduction

A fundamental question in biology centers on understanding the origin and evolution of morphological diversity and its regulation at the genome level. Arguably, among the most striking examples of morphological variation are the stunning array of colors and patterns that decorate the wings of butterflies. The spectacular diversity of butterfly wing patterns has been shaped by natural selection to serve a variety of adaptive functions, ranging from mate recognition and courtship to predator avoidance and deterrence (Brakefield & French, 1999; Brunetti et al., 2001a; Beldade & Brakefield, 2002). Although many of the ecological processes shaping color patterns are well documented, the underlying molecular and developmental program generating these patterns still remains largely unknown (Beldade & Brakefield, 2002; Werner et al. 2010; Martin et al., 2012). The emergence of the exciting field of Evo-devo positioned butterfly

wings as a valuable model system for studying morphological diversity and adaptation (Nijhout, 2001; Beldade & Brakefield, 2002; Monteiro & Prudic, 2010; French-Constant, 2012) and stimulated an increased focus on the discovery of genes and regulatory mechanisms underlying color pattern development (Carroll et al., 1994; Martin & Reed, 2010; Reed et al., 2011; Stoehr, et al., 2013).

Over the past two decades, research has revealed that genes involved in wing color pattern development also belong to an ancient gene regulatory network (GRN) for wing construction (Brakefield & French, 1999; Keys et al., 1999; Saenko et al., 2008). This network has been proposed to serve as a pre-patterning template for downstream pigment genes (McMillan et al., 2002; Beldade & Brakefield, 2002; French-Constant, 2012). Studies on wing development in *Drosophila melanogaster*, ants and aphids have characterized expression patterns of this gene regulatory network (Abouheif & Wray, 2002, Brisson et al. 2010); however, no comprehensive analysis has been conducted in butterfly wings. Thus, we do not have a clear understanding of the expression dynamics of this network during butterfly wing development and how it may correspond to temporal changes in pigment gene expression.

The wing GRN is comprised of at least 18 major developmental genes representing selector genes, morphogens and a suite of transcription factors that co-operate in wing development (Abouheif & Wray, 2002). Selector genes encode a unique class of transcription factors that act as master switches, controlling genes that regulate the development of specific cells, tissues and organs (Carroll et al., 2001; Mann & Carroll, 2002; Wolpert, 2003). Selector genes include the Hox genes, which function as regional selector genes and specify segment identity along the anterior/posterior axis; for example,

ultrabithorax (ubx) which regulates butterfly hindwing identity (Weatherbee et al., 1999; Krupp et al. 2005). At a finer scale, field-specific selector genes control growth of entire fields of cells and structures whereas compartment specific selector genes regulate development of dorsal/ventral or anterior/posterior identity (Carroll et al., 2001; Halder & Carroll, 2001). Many of these genes are also pleiotropic and have important developmental roles outside of the wing (Monteiro & Podlaha, 2009).

Development of the imaginal disc to the adult wing is governed by an intricate network of patterning genes that appear to have been conserved in holometabolous insects for over 300 million years (Abouheif & Wray, 2002). Wing development is initiated when morphogen and selector genes map out a coordinate system by dissecting the wing into functionally distinct compartments (anterior/posterior [A/P], dorsal/ventral [D/V] and proximal/distal [P/D]). Early in development, the compartment selector genes, *apterous (ap)* and *engrailed/invented (en/inv)* subdivide the wing disc into D/V and A/P regions, respectively (Carroll et al., 1994; Keys et al., 1999; O’Keefe & Thomas, 2001). Activity of these selector genes initiates a signal transduction cascade, triggering expression of other genes that regulate wing development.

Expression of *en/inv* induces signaling of the short-range molecule *hedgehog (hh)* from posterior to anterior regions (Wolpert, 2003). Diffusion of *hh* creates the A/P boundary, which induces expression of the long-range morphogen, *decapentaplegic (dpp)* in a thin stripe of anterior cells along the A/P boundary, promoting outgrowth of the wing blade (Posakony et al. 1990). *Dpp* activity also induces expression of other developmental genes, including the transcription factor *spalt (sal)*, which plays a key role in wing vein development during pupation (de Celis & Barrio, 2000).

In dorsal compartments, *ap* induces the expression of *serrate* (*ser*) and *delta*, two important ligands for the receptor *notch* (Neumann & Cohen, 1996). The signaling of these ligands induces the diffusion of another long-range morphogen, *wingless* (*wg*), which coordinates the dorsal/ventral boundary and the wing margin, the edges of which are marked by the transcription factor *cut* and the field-specific selector *distal-less* (*dll*) (Carroll et al., 1994; Krupp et al., 2005; Macdonald et al., 2010; Iwata et al., 2014). Later in development, a feedback loop between *wg* signaling and ligands *ser* and *delta* maintain *wg* and *cut* activity at the D/V boundary, sculpting the final wing shape (Milán & Cohen, 2000, Macdonald et al., 2010). *Wingless* also induces expression of the field selector genes, *vestigial* (*vg*) and *scalloped* (*sd*) that together promote wing differentiation (Bray, 1999; Carroll et al., 2001).

In addition to regulating wing development, many of these selector genes and morphogens appear to have been redeployed in novel developmental contexts to specify wing color patterns (Beldade & Brakefield, 2002; Martin & Reed, 2010; Monteiro & Podlaha, 2009; Oliver et al., 2012). Eyespots are the most well studied wing color pattern elements and studies have revealed that at least 12 genes are expressed in the focus and concentric colored rings (Brunetti et al., 2001; Monteiro et al. 2003; Otaki, 2011; Saenko et al. 2011; Oliver et al., 2012; Oliver et al. 2013). Intriguingly, many of these genes are the same developmental genes involved in the wing GRN indicating a potential co-option event (Monteiro, 2012). In nymphalid butterflies, the focus of the eyespot is associated with expression of *antennepedia*, *en*, *sal*, *dll* and *notch* (Brunetti et al., 2001; Oliver et al., 2012). The morphogens *wg* and *dpp* may also play a role in eyespot positioning and appear to function as a signal that activates *dll* and *sal* expression (Beldade & Brakefield,

2002; Monteiro et al. 2006; Held, 2012). Many of these same wing developmental genes are also expressed in other pattern elements. For example, *wg* is involved in the development of stripes on the wings of moths and butterflies and expression of *env/inv* has been correlated with the development of the disc spots in saturniid moths (Monteiro et al., 2006; Martin & Reed, 2010).

These studies reveal a remarkably diverse role of the wing GRN during development in controlling both wing size and shape and in generating novel wing color patterns. However, developmental genes are not the whole story; wing color patterns are ultimately the product of pigment synthesis pathways in individual scales (Ffrench-Constant, 2012). Coloration can be the result of pigmentation that is synthesized *de novo* during scale development in the pupa, or it can develop as structural modifications that result in iridescent scales (Knüttel & Fiedler, 2001).

Wing color patterns are determined when each scale cell specifies a particular color pigment and pattern diversity arises with variation in the color, size and arrangement of these pigments. Studies on *Heliconius* butterflies have identified specific pigment genes associated with particular color pattern elements (Ferguson & Jiggins, 2009; Ferguson et al. 2011; Reed et al., 2011). Major pigment pathways include pteridins (white), the ommochromes (red, yellow and orange-- found only in nymphalids), and the melanins (black, brown and tan) (Reed & Nagy, 2005; Ferguson & Jiggins, 2009; Martin & Reed, 2010; Ferguson, Maroja, et al., 2011). The ommochrome pathway is characterized by the enzymes tryptophan oxidase (*vermillion*), kynurenine 3-hydroxylase (*cinnabar*) and *kynurenine formamidase (kf)* (Reed & Nagy, 2005; Ferguson & Jiggins, 2009). These enzymes convert the precursor tryptophan into a variety of ommochrome

pigments. In general, ommochrome pigments appear earlier in pupal wing development than melanin pigments (Koch et al. 2000; Ferguson & Jiggins, 2009; Iwata et al., 2014;). The melanin pigmentation pathway has also been well characterized in *Drosophila* and Lepidoptera and is comprised of the enzymes, tyrosine hydroxylase (*pale*), Dopa decarboxylase (*DDC*), NBAD hydrolyase (*tan*), NBAD synthetase (*ebony*) and *yellow* (Wittkopp & Beldade, 2009). *Ebony* has been shown to be up-regulated in non-melanic tissues, while *tan* is associated with melanin pigmentation (Ferguson & Jiggins, 2009).

While many of the genes involved in pigmentation are well characterized, the connection between the developmental genes in the wing GRN and pigmentation pathways remains unclear (McMillan et al., 2002; Hines et al., 2012). A link has been established between developmental genes and specific pigments; for example, *en* has been mapped to the ring of gold scales around the eyespots of *Bicyclus* (Brunetti et al., 2001; Monteiro et al., 2006; Oliver et al., 2012). Melanin pigmentation has also been shown to be associated with *sal* expression in pierid butterflies (Stoehr et al., 2013) and *wg* signaling in *Heliconius* butterflies (Gibert et al. 2007; Martin et al. 2012). These examples clearly illustrate a role for patterning genes in regulating downstream pigment genes, yet an important challenge remains understanding the underlying mechanisms linking these interacting networks.

Next generation sequencing has become a valuable tool for surveying the transcriptome of non-model organisms (Ekblom & Galindo, 2011). Although Lepidoptera are a diverse order of insects, there are still relatively few well annotated genomic resources (Ferguson et al., 2010). Advances in our understanding of the transcriptome during wing color pattern development has come primarily from microarray studies on

the temporal patterns of gene expression during pupation in *Heliconius erato* (Hines et al., 2012), EST data from larval and pupal wings from *B. anynana* (Beldade et al. 2006), 454 sequencing of pooled larval and pupal wings in *Heliconius melpomene* (Ferguson et al., 2010) and Illumina sequencing of microRNA expression of pooled larval and pupal wings in *Heliconius melpomene* (SurrIDGE et al., 2011). Here, we use Illumina sequencing to examine temporal expression patterns of the wing GRN and pigment genes during larval and pupal development. This work serves as a valuable contribution towards our understanding of the wing transcriptome during color pattern development.

Study System

Our current understanding of the genes involved in wing color pattern development is based on a small selection of species, primarily *Junonia coenia*, *Bicyclus anynana* and members of *Heliconius* (Oliver et al., 2013; Supple et al., 2013). A greater diversity of species should be examined to better understand how wing color patterning has evolved in butterflies. We used the painted lady butterfly, *Vanessa cardui* Linnaeus (Nymphalidae) as our study organism. *Vanessa cardui* is a long distance migrant and is one of the most widespread of all butterflies occurring worldwide with the exception of South America (Brock and Kauffman 2006). Similar to *Bicyclus* and *Heliconius* butterflies, the wing color patterns of *V. cardui* have been well studied; however, in contrast to *Bicyclus* and *Heliconius*, fewer molecular resources are available (but see Brunetti et al., 2001; Macdonald et al., 2010; Oliver et al., 2012). Furthermore, *V. cardui* is closely related to well-studied models, which will facilitate comparative analysis of gene expression patterns across different genera within the same family.

Methods

Tissue Collection

Vanessa cardui caterpillars and artificial diet were purchased from Carolina Biological Supply Company (Burlington, NC). The caterpillars were reared individually at ambient temperature (~28°C). Wing discs were dissected from caterpillars at two developmental time points in the final instar (2 days and 4 days post-molt), and at three time points in the pupal stage (2 days, 5 days and 8 days post-pupation). Prior to harvest, larvae were weighed. The thorax, including the first abdominal segment, was harvested and placed immediately in RNAlater[®] (Ambion) and stored at 4°C for at least 48 hours prior to dissection. Pupal wings were dissected from live pupa using a Zeiss Stemi-2000 Microscope and placed immediately in RNAlater and stored at 4°C. Imaginal wing discs were carefully dissected from the larva and placed in RNazol[®] RT (Molecular Research Center Inc.) for RNA isolation. For pupal wing samples, samples were placed in RNazol for RNA isolation. All tissues were weighed and processed using an electric homogenizer followed by RNA isolation using isopropanol. Concentration of RNA was measured using a ND-1000 spectrophotometer (NanoDrop products, Wilmington, DE) (A260/A280 >1.8) and integrity was assessed using electrophoresis on a formaldehyde-agarose gel. The RNA samples were diluted in water to a concentration of 25ng/μl in 50μl. For larval wing samples, 5 individual larvae were pooled for each developmental time point for a total of 10 individuals (one biological replicate per time point). For the 2 and 5-day pupal wings, 4 individuals were pooled, and 3 were pooled for the 8-day time point. Two biological replicates were prepared for each pupal time point. The samples were shipped on dry ice over night to the Utah Microarray and Genomic Analysis Facility (University

of Utah, Salt Lake City, UT) for library preparation and sequencing on an Illumina HiSeq 2000 sequencer.

cDNA Library Construction and Illumina Sequencing

Library construction was performed using the Illumina TruSeq RNA Sample Preparation Kit v2.. Briefly, total RNA (100 ng to 4 ug) was poly-A selected using poly-T oligo-attached magnetic beads. The Poly-A RNA was eluted from magnetic beads, fragmented and primed with random hexamers. First strand cDNA synthesis was performed using Superscript II Reverse Transcriptase (Invitrogen) and then converted to blunt-ended fragments with an A-base following second strand synthesis. Adapters containing a T-base overhang were ligated to the A-tailed DNA fragments. The ligated fragments were PCR-amplified (12 cycles) and the amplified library purified by Agencourt AMPure XP beads (Beckman Coulter Genomics). Concentration of the amplified library was measured with a NanoDrop spectrophotometer. To determine the size distribution of the sequencing library an aliquot was resolved on an Agilent 2200 Tape Station. Quantitative PCR (KapaBiosystems Kapa Library Quant Kit) was used to calculate the molarity of adapter ligated library molecules and the concentration of the libraries was adjusted to a concentration of 10 nM. Library concentration was further adjusted in preparation for Illumina sequence analysis.

Transcriptome Assembly and Sequence Annotation

De-novo transcriptome assembly was conducted using CLC Genomics Workbench 6.5.1 with a word size of 40. The parameters were modified throughout the assembly and mapping process to optimize similarity and length fraction required for robust mapping and assembly. Mismatch, insertion and deletion costs were set at 2, 3 and 3 respectively.

Following completion of the assembly, a multiblast (tblastn) was conducted through CLC Genomics using the *Drosophila* peptide database downloaded from FlyBase. Results from the multiblast were used to mine the transcriptome for genes from the wing gene regulatory network and genes involved in ommochrome and melanin pigmentation. Genes of interest that were not retrieved from the *Drosophila* multiblast were obtained from NCBI and used to build a list of proteins from related insect species for a separate tblastn multiblast against the *V. cardui* transcriptome. Using these two approaches we were able to identify most genes of interest from our transcriptome. Top hits of *V. cardui* transcripts from the multiblast were used in a reciprocal blastn search against the non-redundant database at NCBI to validate the identity of each gene and also to annotate start and stop codons (Appendix Table 2.1). All assembled contigs (including annotated and non-annotated) were used for RNA-seq analysis, which was also performed using CLC genomics. Expression values were normalized to reads per kilobase of exon per million reads mapped (RPKM). Gene expression analyses for target genes were performed in JMP version 10.0.2 SAS Institute Inc. Data were log or square root transformed where necessary to meet assumptions of normality and equal variance.

Quantitative PCR Validation

An independent experiment was designed to validate the transcriptome results. Wing discs and pupal wings were dissected at the same developmental stages as the transcriptome study with seven biological replicates per stage. RNA isolation was performed as described above. RNA quality was checked for degradation on a formaldehyde-agarose gel and examined for genomic DNA contamination using qPCR, with primers for the glutamate receptor, which also served as the housekeeping gene.

cDNA synthesis was performed with an iScript kit (BioRad) in a single run for all samples using 1 µg of input RNA (20 µl reaction). An aliquot of cDNA was diluted to the equivalent of 2 ng total RNA input/µl for qPCR. Primers were designed for the following genes: *wg*, *sal*, *en*, *dll*, *ddc*, *pale*, *ebony*, *tan*, *vermillion*, *kf*, and *cinnabar* (Appendix table 2.2). cDNA (2 ng/µl) was amplified from wing samples using PRC Accuzyme™ 2x reaction mix (Bioline). *Glutamate receptor* was selected as a housekeeping gene based on results from the transcriptome data. The PCR was checked for a single band (75 bp) on a 1% agarose gel, purified using a Thermo Scientific purification kit and quantified using Nanodrop. Standard curves were generated using an initial concentration of 2 picograms of PCR product and serial 10-fold dilutions. qPCR was performed using 2 µl of cDNA template with Evagreen Supermix (BIO-RAD) (10 µl reaction/well), and run on a CFX384 Real time system (Bio-rad C1000 Thermocycler) with the following conditions 95°C 30s, 95°C 5s, 60°C 5s for 40 cycles.

Statistical Analyses

RNA-seq analysis was performed on RPKMs to compare the expression of genes from the GRN and genes involved in melanin and ommochrome pigmentation. We also examined expression of two putative receptors associated with the ommochrome pathway, *scarlet* and *white*. Twelve genes were also selected for quantitative PCR to validate expression patterns observed with the transcriptome analysis (Table 2.2). Two-way ANOVA was conducted for all analyses in JMP Version 11.0.0 (SAS Institute Inc., Cary, NC) using the *glutamate receptor* as an internal control. For qPCR validation we selected four genes from the GRN, all three ommochrome genes and all melanin genes with the exception of *yellow* of which there were multiple paralogs. Due to the single

biological replicate for the larval stages, statistical analyses were performed only on pupal stages for the RNA-seq analysis. All developmental stages were analyzed from the qPCR experiment.

Results

A total of 17 libraries of raw reads (50 bp) were used to assemble the wing transcriptome for *V. cardui*. Four libraries were from the early 4th instar and eight libraries were from the late 4th instar, with 15-18 million reads per library. The pupal stages were each represented by two libraries. The total number of reads obtained from the larval libraries was 265,105,531 reads and 181,176,198 reads from pupal libraries. A greater number of libraries were obtained for the larval stages as these included treatment groups for a separate experiment that is not part of the current study. Only data from the control groups (2 libraries one each from early and late 4th instar) were used for downstream expression analyses for comparison across developmental stages. Appendix Table 2.3 presents a full summary of the transcriptome assembly. A total of 446,282,529 raw reads were used to construct the transcriptome. The final transcriptome comprised 89,069 contigs with a mean length of 779.8 bp and N50 of 1266 bp after removal of short sequences <200 bp. Mapping of the raw reads back to the transcriptome revealed that 91% of the reads mapped to the final assembly. When larval and pupal libraries were mapped separately, 94% of reads from the larval libraries and 87% of reads from the pupal libraries mapped to the assembled transcriptome.

Of the 89,069 contigs, only a small handful of putative genes were identified that exhibited constant levels of expression across all developmental stages. Following blast searches we identified one of these genes as a putative *glutamate receptor*. The remaining

contigs did not match any known sequences on NCBI and are likely non-coding RNA. Quantitative PCR confirmed that expression of the *glutamate receptor* did not vary across developmental stages ($p>0.05$, $F=0.6$). *Actin* and *GAPDH* were also examined as potential internal controls however expression was variable across developmental stages, which was verified by qPCR.

Correlation of qPCR and RNA-seq Data

A bivariate analysis of fold change in expression relative to the early 4th larval stage for all twelve genes (including the *glutamate receptor*) revealed that the RNA-seq and qPCR results are largely consistent with each other (Figure 2.1). Examination of fold change of each gene individually reveals very similar expression patterns and a high correlation between the RNA-seq and qPCR analysis for most genes (Figure 2.2). Weaker correlations were found for genes expressed at very low levels e.g. *dll* and *en*.

Wing GRN Patterning Genes

Both the transcriptome and qPCR analyses revealed a dynamic pattern of expression during wing development from larval to pupal stages. For the wing patterning genes, the transcriptome analysis revealed that 13 out of the 17 genes had relatively low expression, less than 50 RPKM across all developmental stages. There was no expression of *Abd-A* in the wing transcriptome. For the analyses of the transcriptome pupal data, we categorized genes into the following three functional groups: 1. selector genes, 2. signaling molecules and 3. general transcription factors (Figure 2.3 and see Appendix Table 2.4 for literature references). For most patterning genes, peak expression occurred during the late larval and early pupal stages followed by a sharp decline at 5 or 8 days post-pupation (Figure 2).

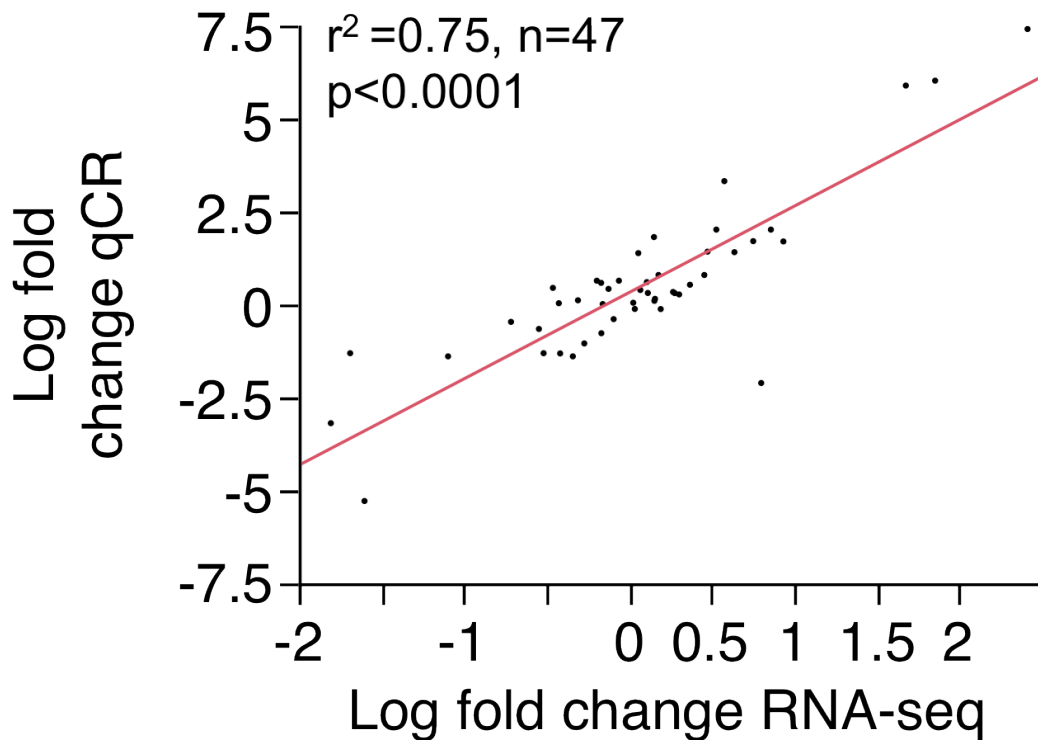


Figure 2.1 Bivariate analysis of fold change expression for RNA-seq and qPCR. Fold change is relative to the early 4th larval wing for all developmental stages across all genes. The regression shows a strong correlation for results obtained using these two different methods. Data points for RNA-seq are the result of one pooled (5 individuals) biological replicate for larval stages and two pooled (3-5 individuals) biological replicates for pupal stages. Data points for qPCR are based on 7 biological replicates. Correlation coefficient, p value for the hypothesis $r = 0$, and sample size for gene expression data for 12 genes are also presented (*dll*, *en*, *wg*, *sal*, *vermillion*, *kf*, *cinnabar*, *pale*, *ddc*, *ebony*, *tan* and *glutamate receptor*)

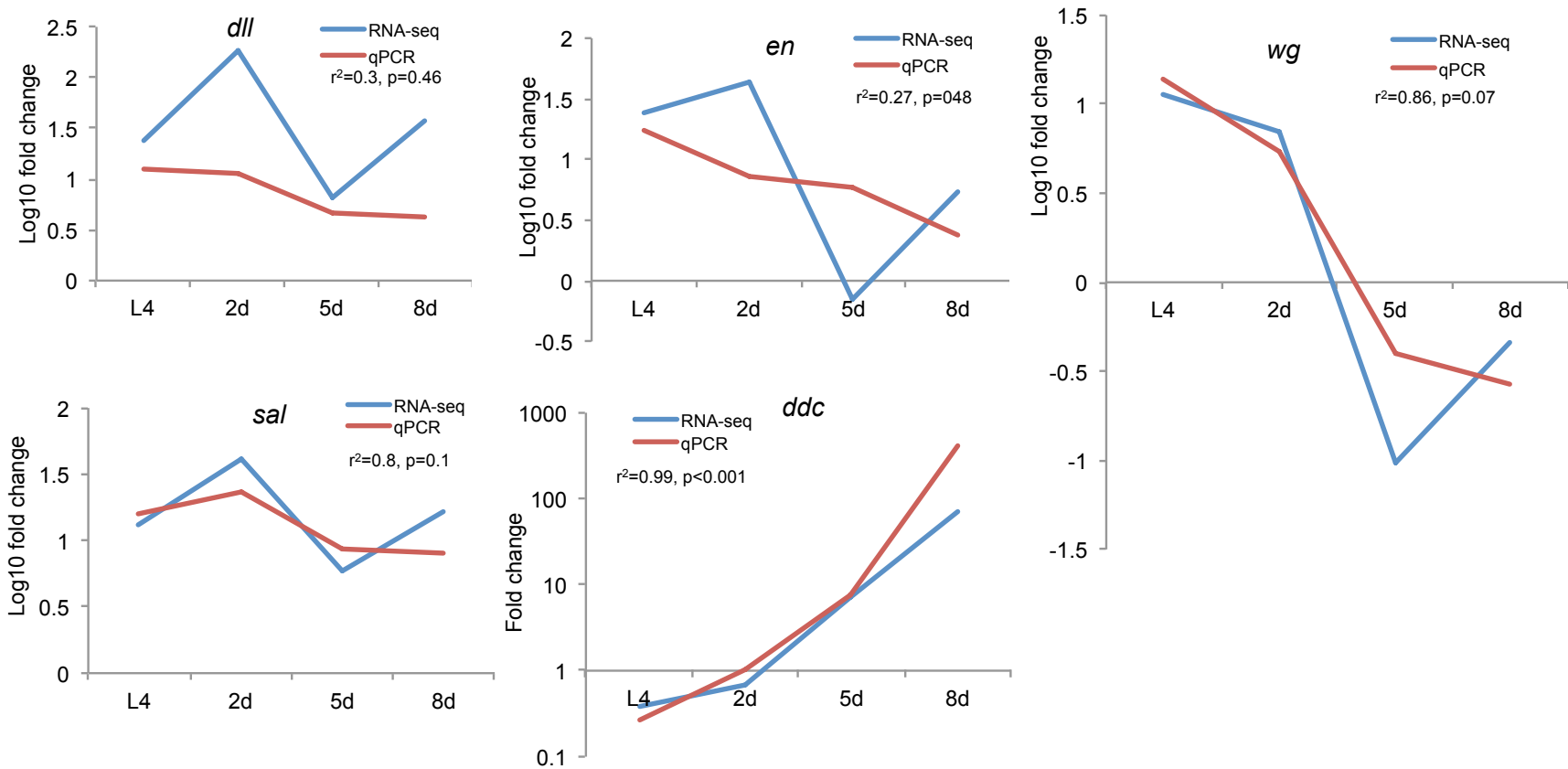


Figure 2.2 RNA-seq and qPCR data showing fold change expression for patterning and pigment genes. Fold change is calculated for individual genes at each developmental stage relative to early 4th instar. Correlation coefficient, p value for the hypothesis $r = 0$ are also presented.

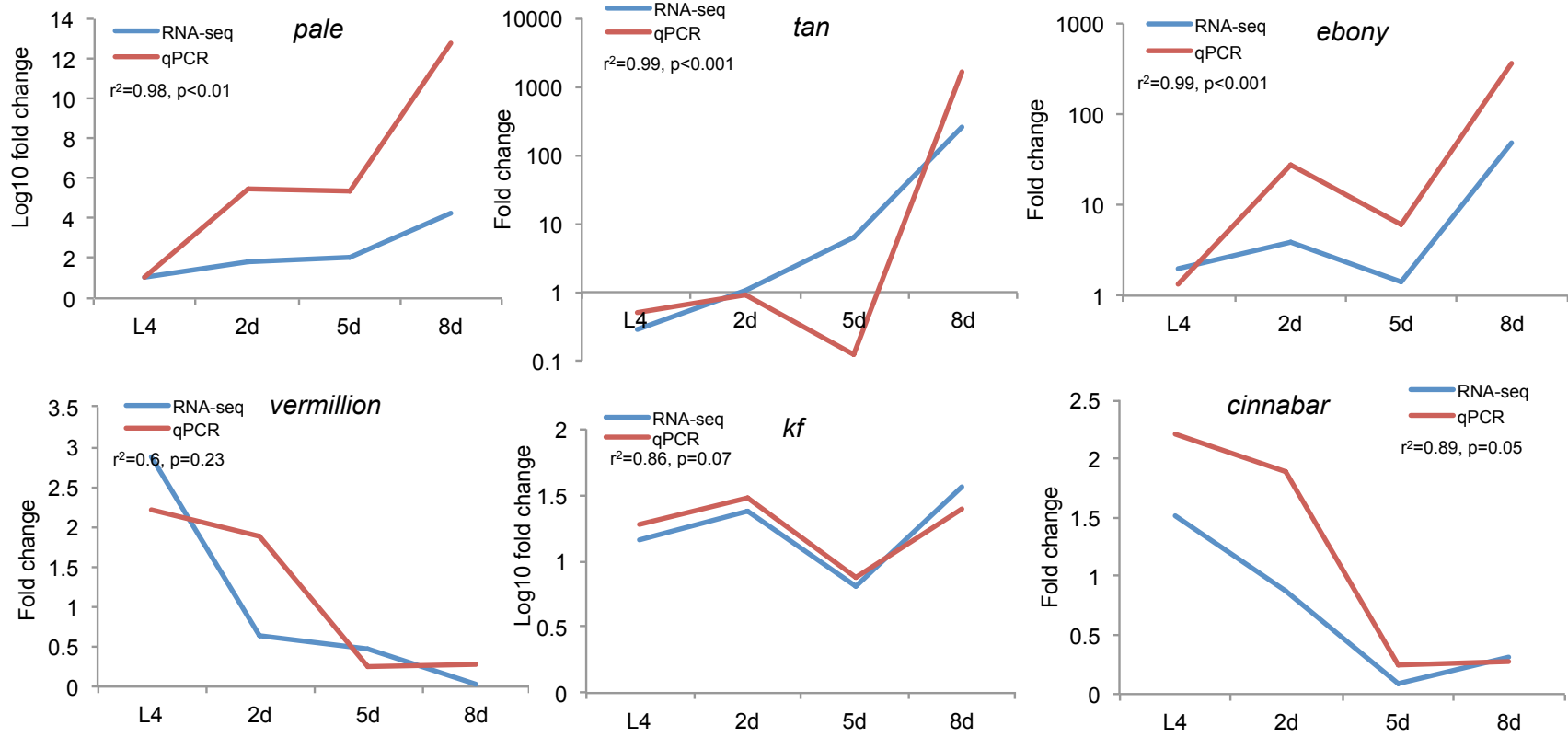


Figure 2.2 cont. RNA-seq and qPCR data for pigment genes showing fold change expression. Fold change is calculated for individual genes at each developmental stage relative to early 4th instar. Correlation coefficient, p value for the hypothesis $r = 0$ are also presented.

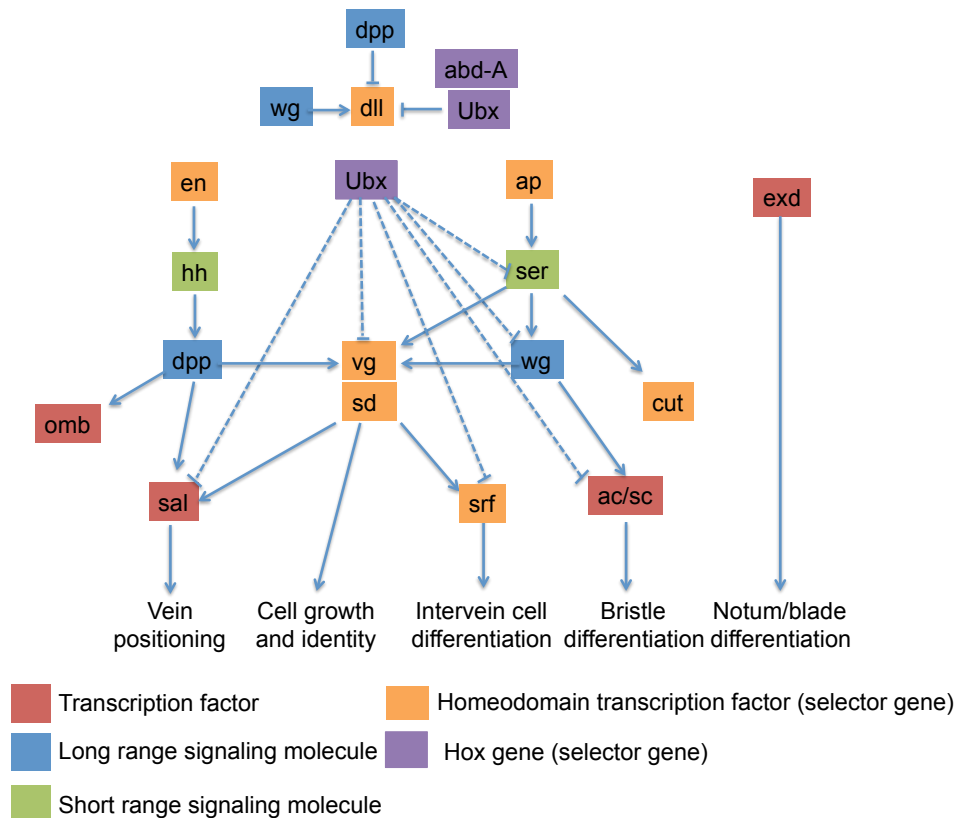


Figure 2.3 Wing gene regulatory network. Model of the gene regulatory network for wing development characterized from *Drosophila melanogaster* and adapted from Abouheif and Wray (2002). The network depicts the hierarchy of patterning genes involved in the establishment of the imaginal disc and developmental of wings during the larval stages. The different functional groups are color-coded to highlight their role and placement within the network.

Within each functional group, there was a significant interaction between gene and developmental stage (day post-pupation); selector genes: $F_{(14, 23)} = 7.8$ $p < 0.0001$), signaling molecules, ($F_{(6, 11)} = 17.85$, $p < 0.0001$) and general transcription factors ($F_{(6, 10)} = 53.24$, $p < 0.0001$). Within the selector genes, *vg* was the most highly expressed gene across all pupal stages, while *cut* was expressed at the lowest levels at days 5 and 8. For the signaling molecules, *dpp* and *ser* were expressed at significantly higher levels than *hh*

and *wg* across all pupal stages. *Extradenticle* was the most highly expressed transcription factor and *aschaete scute ac/sc* exhibited the lowest expression levels while expression of *serum response factor (srf)* was barely detectable.

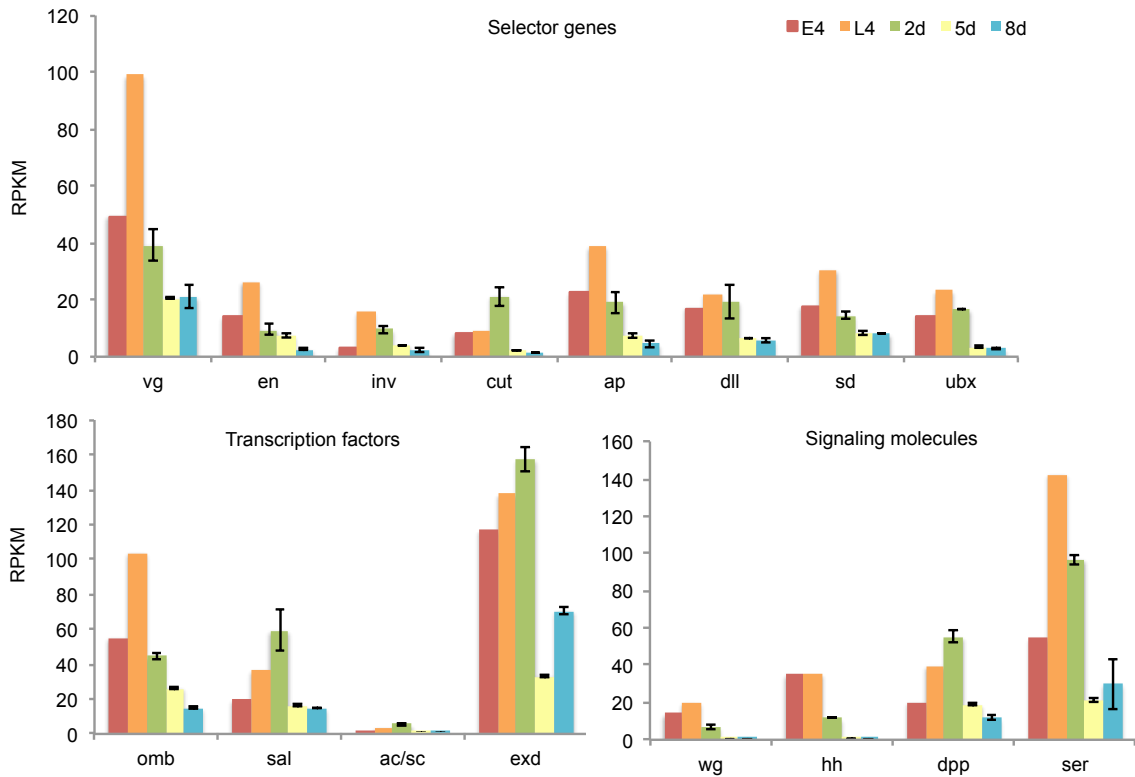


Figure 2.4. RNA-seq expression patterns for the different functional groups of the wing GRN. Larval stages represent one pooled biological replicate; pupal stages represent two pooled biological replicates. Error bars represent 1 SD from the mean.

Ommochrome Pigmentation

In *V. cardui*, white pigmentation is deposited at 5 days and red pigmentation is deposited around 6 days post-pupation (Figure 2.5). The transcriptome and qPCR analysis revealed that expression of pigment genes involved in ommochrome synthesis is upregulated during the larval wing stages with most genes exhibiting a peak in expression during the late larval stages and 2 days post-pupation. Expression patterns of the three pigment

genes, *vermillion*, *kf* and *cinnabar* varied significantly during pupal development with a significant interaction between gene and developmental stage ($F_{(4, 8)} = 32.2$, $p < 0.0001$). The transcriptome analysis revealed that only *cinnabar* and *kf* exhibited a peak in expression at 2 days post-pupation and both declined significantly at 5 days post-pupation.

Interestingly, *kf*, exhibited a dramatic increase in expression from 5 – 8 days post-pupation, while expression levels of *cinnabar* did not vary between these two stages. The transcriptome data showed that *vermillion* exhibited a gradual decline in expression during pupation with significant decline between days 5 and 8 post-pupation. The qPCR results also produced the same overall trend, although it suggests that *vermillion* does exhibit a significant peak in expression at 2 days post-pupation. Overall, *kf* was the most highly expressed ommochrome gene. *Scarlet* and *white* ommochrome receptors declined in expression across pupal development. A significant interaction was observed between gene and day ($F_{(2, 5)} = 44.87$, $p < 0.001$) (Figure 2.5). Both genes exhibited a peak at 2 days post-pupation and the white receptor was expressed at significantly higher levels across all pupal stages.

Melanin Pigmentation

The qPCR and RNA-seq analysis revealed that genes involved in the melanin synthesis pathway exhibit a very different pattern from the wing GRN and ommochrome genes. Most melanin genes increased in expression during pupation and a significant interaction was observed between gene and day ($F_{(18, 29)} = 88.6$, $p < 0.001$) (Figure 2.6). Expression of most melanin genes peaked at 8 days post-pupation, which is one day prior to butterfly emergence.

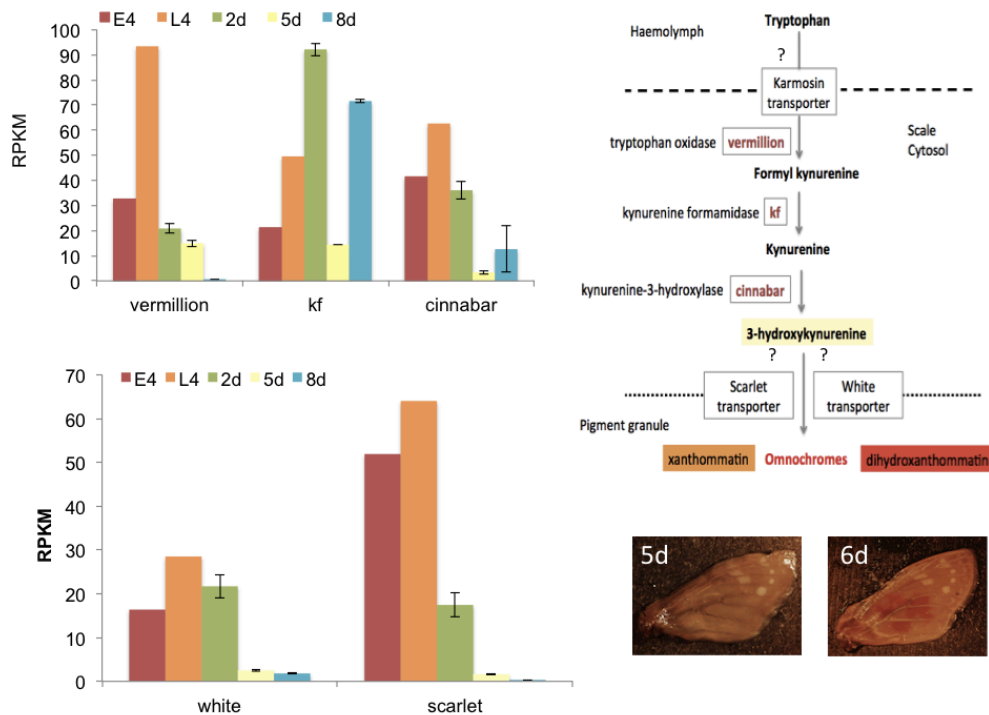


Figure 2.5 RNA-seq temporal expression patterns for ommochrome genes. Hypothesized ommochrome pigment biosynthesis pathway in butterfly wings based on the pathway in *Drosophila melanogaster* ommatidia (Ryall & Howells, 1974) and adapted from (Reed & Nagy, 2005). Tryptophan enters the scale cell from the hemolymph through the putative *karmoisin* transporter where it is converted to ommochrome precursors via oxidation of three major enzymes encoded by *vermillion*, *kf* and *cinnabar* (Reed & Nagy, 2005; Reed et al., 2008). The final conversion of 3-OHK to ommochromes (orange xanthommatin and red dihydro- xanthommatin) occurs in pigment granules. Uptake of ommochrome precursors into pigment granules is thought to occur through the scarlet and white transporters as in *D. melanogaster* (Warren et al. 1996). Wings from *Vanessa cardui* at 5 days post-pupation and 6 days at onset of ommochrome pigmentation.

We also identified multiple paralogs of the *yellow* gene and annotated the full coding sequence for six paralogs, *yellow*, *yellow-b*, *yellow-c*, *yellow-d*, *yellow-f3* and *yellow-x*. Although most melanin genes increased in expression during pupation, some of the *yellow* paralogs exhibited contrasting patterns of expression. In particular, *yellow* exhibits a dramatic peak at 5 days post-pupation, while *yellow-f3* and *yellow-d* exhibit a sharp

decline at this pupal stage, as does *ebony*. *Yellow* and *yellow-f3* are also the only melanin genes, which do not show a peak in expression at 8 days post-pupation. Similar to the ommochrome genes, expression of melanin genes was also evident in the larval stages, albeit at much lower levels, long before black pigmentation is visible (around 7 days post-pupation). Overall, *pale* (TH) was the most highly expressed melanin gene and *ebony* and *yellow-f3* were expressed at the lowest levels.

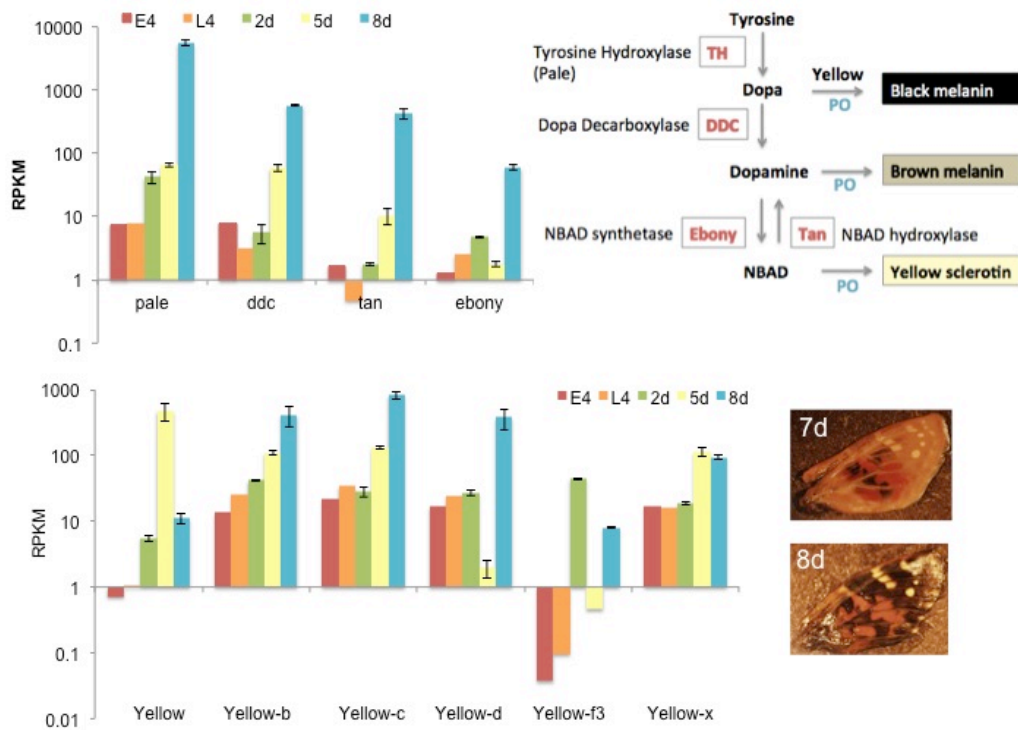


Figure 2.6. RNA-seq temporal expression patterns for melanin genes. The melanin biosynthesis pathway adapted from (Wittkopp & Beldade, 2009). Tyrosine hydroxylase (TH) encoded by the *pale* gene catalyzes the conversion of the precursor tyrosine to dopa, which is further catalyzed to dopamine by dopa decarboxylase (*ddc*). Dopa and dopamine are converted into black melanin and brown melanin respectively through the action of phenol oxidases. *Yellow* is required for the production of black melanin, though its specific biochemical function is unknown (Wittkopp, et al. 2002). Dopamine is converted to NBAD by *ebony*, (N- β -alanyl dopamine synthetase, NBAD synthetase) which produces tan pigmentation. NBAD is oxidized to produce yellow sclerotin or converted back to dopamine by the enzyme *tan* (NBAD hydrolyase) to produce tan pigment (Wittkopp et al. 2002). Wings from *Vanessa cardui* at 7 and 8 days post-pupation illustrate development of melanin pigmentation.

Discussion

Here, we describe the first transcriptome analysis of larval and pupal wing development in the painted lady butterfly, *Vanessa cardui*. Using next generation sequencing with Illumina Hiseq2000 we successfully assembled the transcriptome using short reads (50bp), and obtained a final assembly of 85,065 contigs, over which 50% (40,000), were greater than 200bp. The total number of contigs, mean contig length and N50 are within the ranges recently reported for Illumina sequencing in other arthropods (Van Belleghem et al., 2012 Croucher et al. 2013; Li et al., 2013). A major goal of this study was to identify and quantify expression patterns of genes from the wing GRN previously characterized in *Drosophila* and compare temporal expression patterns with genes involved in pigmentation. We identified all the major genes that are depicted in the wing GRN with the exception of *Abd-A* and obtained full coding sequences for 10 of the 17 genes examined. We also identified all of the major enzymes involved in the ommochrome and melanin pigmentation pathways that have been characterized in *Drosophila* and other butterfly species. We identified several paralogs of *yellow-y*, some of which were assembled into very short sequences (*yellow-e*, *yellow-f4*, *yellow-h2* and *yellow-h3*).

Temporal Expression Patterns of the Wing GRN

It has been suggested that genes involved in the wing GRN may play a role in pre-patterning the downstream pigment genes (Mcmillan et al., 2002; Reed & Nagy, 2005). Therefore, we would expect patterning genes to be upregulated during late larval or early pupal stages, and pigment genes to be upregulated later in pupal development. Results from the RNA-seq analysis revealed that the different members of the wing GRN exhibit

very similar temporal patterns of expression. Most genes are upregulated during the larval stages and early pupal stages and then decline significantly during pupation. *Extradenticle*, *ser*, *omb* and *vg* were the most highly expressed genes across all developmental stages, with the exception of 2 days post-pupation. At 2 days post pupation, *sal* and *dpp* were more highly expressed than *vg* and *omb*.

Overall, *exd* was the most highly expressed gene throughout wing development (32-157 RPKMs) and interestingly exhibited increased expression shortly before eclosion. *Extradenticle* is a homeobox transcription factor, which acts as a cofactor for Hox genes in the specification of segmental identity (Rauskolb et al. 1995). Interactions of Hox genes with cofactors are thought to be critical in the selective regulation of Hox target genes such as *ubx*, due to their low DNA binding specificities. Hox genes have also been shown to function independently of cofactors, particularly in the case of distal appendages in arthropods and vertebrates (Galant et al., 2002). Transcript levels of *exd* are uniform throughout the imaginal wing disc in *Drosophila* (Rauskolb et al., 1995). However, the protein is mostly expressed in the cytoplasm with nuclear distribution restricted only to the notum and wing hinge (Aspland & White, 1997). Thus, the spatial distribution of *exd* suggests it is involved in regulating proximodistal polarity in addition to segmental identity. In *Drosophila*, *exd* interacts with the selector gene *en* to mediate repression of *en* target genes, indicating that *exd* has functions in addition to its role as a Hox cofactor (González-Crespo & Morata, 1995, Kobayashi, 2003). Whether *exd* exhibits a similar functional role and spatial patterning during butterfly wing development is currently unknown.

We detected extremely low transcript abundance for two genes located at the base of the wing GRN hierarchy, *ac/sc* and barely detectable levels of *srf*. *Serum response factor* is a MADS box transcription factor involved in wing vein formation and differentiation of intervein wing tissue in both *Drosophila* and butterflies (Montagne et al., 1996; Galant et al., 1998). *Achaete scute* encodes a basic helix-loop-helix (bHLH) transcription factor involved in bristle development in *Drosophila*. Patterns of *ac/sc* expression in wing discs of *Precis coenia* are similar to that of *Drosophila*; this gene regulates development of innervated sensory scales (Galant et al., 1998). In pupal wings, expression of *ac/sc* is also observed in scale precursor cells that later differentiate into the socket and scale cells (Galant et al., 1998). Transcript levels of *srf* were barely detectable across all developmental stages but peak expression of *ac/sc* was observed during the larval wing stages and 2 days post-pupation, which then dropped to barely detectable levels at 5 and 8 days. These results indicate that very low levels of these proteins are required in the regulation of vein and scale development and that the functional role of *ac/sc* is likely restricted to larval and early pupal stages.

For the selector genes, *vg* emerged as the most highly expressed gene across all developmental stages. Overall, the range of expression of most genes across developmental stages was very similar, which is consistent with the observation that developmental genes are very tightly regulated (Macneil & Walhout, 2011). Of the signaling molecules, *dpp* and *ser* were consistently expressed at the highest levels with *wg* and *hh* exhibiting the lowest levels of expression. Expression of *hh* and *wg* declined to barely detectable levels at 5 and 8 days post-pupation, suggesting that they play more

important functional roles during larval and early pupal stages. These results also show no differences in expression levels between short versus long range signaling molecules.

Recent detailed work on live pupal imaging in the butterfly *Junonia orithya* has provided a window into the physiological processes during wing development from 0h post-pupation through butterfly eclosure, thus capturing temporal patterns of pigment deposition (Iwata et al., 2014). This imaging study revealed a potential organizing center for the marginal band system and a prospective eyespot focus immediately following pupation with the eyespot focus becoming apparent within the first 48 hours. Some of the genes from the wing GRN involved in eyespot development are already expressed in the larval wing disc (Keys et al., 1999; Martin & Reed, 2010; Monteiro et al., 2013). Fluorescent imaging has revealed expression of *en*, *sal* and *dll* within the first 24h of pupal wing development in *V. cardui* (Brunetti et al., 2001). The results reported here support the hypothesis that late larval and early pupal stages are a critical and potentially sensitive period for establishment of wing patterning.

Melanin Genes Up-regulated During Late Pupation

In contrast to the wing GRN, genes involved in the melanin pathway exhibited a pattern of increasing expression during wing color pattern development with the highest levels occurring at 8 days post-pupation. These expression patterns mirror the patterns of melanin pigment deposition on the wing which appear late in pupal development (Figure 2.6). We found that tyrosine hydroxylase (*pale*), displayed a sharp increase from the early to mid pupal stages with extremely high expression at 8 days followed by increased expression of *ddc* and *tan*. The qPCR data revealed a strong downregulation of *tan* and *ebony* during upregulation of *pale* and *ddc* (day 5 post-pupation). These results suggest

that down-regulation of *tan* and *ebony* during this developmental period may be required for appropriate melanization.

The *yellow* genes showed a complex pattern of expression, indicating multiple functions of these enzyme products. Among the melanin genes *tan*, *ebony*, *yellow-d* and *yellow-f3* all exhibit similar expression profiles during pupation with a significant decline at day 5. In *Drosophila*, *yellow-y* and *ebony* have opposite effects on pigmentation with *yellow-y* promoting melanization and *ebony* involved in melanin repression (Wittkopp et al., 2002). However, contrasting roles of *yellow* genes have also been observed, for example *yellow-d* is associated with melanized regions in *Bombyx mori* (Xia et al., 2006) and unmelanized regions in *Heliconius* species (Ferguson et al. 2011; Hines et al., 2012). The functional role of the *yellow* genes is not well understood, although different *yellow* genes appear to exhibit variable patterns of expression during development (Han et al., 2002, Ferguson et al. 2011; Hines et al., 2012). The diversity of patterns of *yellow* genes suggests that they may have different functional roles in regulating the intensity of melanin pigmentation in *Vanessa* and other insects.

Ommochrome Gene Expression Mirrors the Wing GRN

Expression patterns of the ommochrome pigment genes, particularly *vermillion* and *cinnabar*, were similar to those of the wing GRN. In *V. cardui*, ommochrome pigmentation becomes visible at 6 days post-pupation, one day prior to melanin deposition. Interestingly, genes involved in ommochrome pigmentation are upregulated long before the pigments are visible on the wing compared to melanin genes, which are upregulated within a few days prior to melanin pigmentation. Similarly, genes that are involved in generating pattern templates that are destined to develop black pigment are

upregulated more than a week prior to melanization. In fact strong upregulation of melanin pigment genes occurs during the developmental period when patterning genes are being downregulated. These results show that timing of gene upregulation for patterning genes corresponds with upregulation of the ommochrome genes, but not the melanin genes. Either melanin regulation of patterning genes occurs indirectly through some as yet unknown developmental pathway or low levels of patterning genes are required for melanization.

It has been suggested that developmental patterning genes may alter the morphology or rate of wing scale development which then ultimately influences pigmentation (Koch et al., 2000; ffrench-Constant, 2012). Our results suggest that if scale maturation is linked to competency in responding to pigment precursors then scales maturing earlier will be exposed to a high availability of ommochrome precursors and those maturing later to melanin precursors. A number of studies have shown that exposure to temperature shock and pharmacological treatments during early pupation can alter patterns of melanin pigmentation (Nijhout, 1984; Serfas & Carroll, 2005; Otaki, 2007). If patterning genes influence scale morphology or development rate, this could explain why perturbations that occur during this developmental window alter patterns of melanin pigmentation later in pupation.

Conservation and Divergence of Pigment Gene Expression Across Butterflies

Surprisingly, comparison of our results with those of Hines et al. (2012), revealed that *pale*, *ddc*, *tan*, *ebony* and *vermillion* show very similar temporal expression patterns between *V. cardui* and *Heliconius* species despite these butterflies exhibiting widely divergent wing patterns. Similar to the Hines et al. (2012) study we also found that the

putative ommochrome transporters *scarlet* and *white* were expressed at relatively low levels particularly during late pupation. Not all genes however were comparable across species. Hines et al. (2012) found slightly different expression patterns of *yellow-y* and *yellow-d* compared to *V. cardui*. Differences were also observed for the ommochrome genes with opposing patterns of expression for *cinnabar* and *kf*. Our study has revealed that several of the pigment genes exhibit a sharp decline at 5 days post-pupation, just one day prior to ommochrome pigmentation. Whether this decline in expression is related to ommochrome pigmentation or due to the white pigments visible on the wing at this developmental stage remains unclear.

Conclusions

In sum, we found a strong correlation between the RNA-seq and qPCR results. Our results highlight that genes in the wing GRN generally exhibit similar expression patterns despite significant differences in function. *Extradenticle* emerged as the most highly expressed gene in the network while *ac/sc* and *srf* were expressed at the lowest levels. Whether *exd* plays any role in regulating wing color patterning or if its function is restricted to controlling other aspects of wing development is currently unknown. This work highlights the differences in temporal expression patterns between the wing GRN and genes involved in melanin and ommochrome pigmentation. We have shown that expression of ommochrome genes is correlated with upregulation of the wing GRN in the larval and early pupal stages while melanin genes are upregulated much later in pupal development. These results raise questions about the molecular mechanisms in which patterning genes regulate the expression of different pigments whether by directly interacting with pigment pathways or indirectly by altering scale development.

CHAPTER III

PHENOTYPIC PLASTICITY REVEALS MODULARITY OF EYESPOT DEVELOPMENT IN THE PAINTED LADY BUTTERFLY, *VANESSA CARDUI*.

Abstract

Homology among eyespots suggests they share a common developmental basis and may function as an integrated unit in response to selection. Despite strong evidence of genetic integration, eyespots also exhibit phenotypic plasticity indicating an underlying flexibility in pattern development. These observations call into question whether eyespots are developmentally integrated or if they function as independent modules. Modularity in eyespot development could facilitate phenotypic plasticity by allowing uncoupling of traits both within and between different eyespots, promoting pattern diversification. We conducted a morphometric analysis to examine whether eyespots of *Vanessa cardui* exhibit phenotypic plasticity and identify which eyespots and eyespot features are most sensitive to perturbation by temperature shock and injection of heparin sulfate. In both treatments, the two central eyespots exhibited the highest levels of plasticity and the inner pigment rings were more strongly affected than the outer ring. We observed changes in both the strength of phenotypic correlations and patterns of integration and modularity. Phenotypic plasticity of eyespots was associated with a loss of integration among eyespots and a doubling of independent modules, suggesting that plasticity promotes

modularity. The results of our study suggest that phenotypic plasticity is associated with altered patterns of phenotypic correlations that may have consequences for selection in different environments.

Introduction

Eyespots are one of the most striking and diverse features displayed on butterfly wings. These colorful pattern elements are composed of concentric rings that can vary widely in size, number and color composition even on the same wing surface. The contrasting colors of concentric rings create a bold, conspicuous pattern, which may have evolved as a visual signal to intimidate or deflect predators (Nijhout, 1996; Stevens et al., 2008). In some butterfly species, eyespots appear to have evolved specialized functions; dorsal eyespots are employed for courtship display while ventral eyespots are used for predator deterrence (Prudic et al., 2011; Oliver et al., 2014). The complexity and diversity of butterfly eyespots has drawn the attention of evolutionary biologists to understand not only their functional role but also the underlying developmental program that generates these patterns (Carroll et al., 1994; Brakefield et al., 1996; Keys et al., 1999; Koch et al., 2003; Beldade et al. 2008; Monteiro et al., 2013).

Several models have been proposed to explain eyespot formation based on a series of elegant studies that manipulated eyespot development and identified genes involved in their initial establishment (Monteiro et al., 1997; Brunetti et al., 2001; Dilão & Sainhas, 2004; Beldade et al., 2008; Otaki, 2011). Eyespot specification begins in the late larval wing discs where a group of organizing cells, the focus, form the presumptive eyespot center (French & Brakefield, 1995; Beldade et al. 2002). Gradient models propose that organizing cells emit one or more putative long-range morphogens that

diffuse radially through gap junctions forming a concentration gradient (Monteiro et al., 1997; Nijhout, 1980). During early pupation, the surrounding epidermal scale cells are thought to respond to positional information specified by the concentration gradient. These scale cells trigger an unknown series of molecular events that lead to the synthesis of different colored pigments. Although the identity of the focal signaling molecule(s) is currently unknown, a number of transcription factors and morphogens have been implicated in regulating eyespot development (Brunetti et al., 2001). Interestingly, the genes identified in eyespot development are the same as those involved in wing development, thus co-option of the wing gene regulatory network may explain how eyespots originated (Oliver et al., 2012). Experimental data suggests that modifications of these developmental networks may have generated eyespot diversity by altering properties of the focal signal and/or response thresholds of scale cells (Brakefield et al., 1996).

In many butterflies, eyespots develop as a series of homologous pattern elements along the wing margin known as the border ocelli system (Monteiro et al., 2003). Each eyespot develops within a wing cell created by a border of wing veins that specify different wing compartments. Artificial selection for eyespot size, color and shape, have revealed correlations among serially repeated eyespots providing strong evidence of developmental integration (Monteiro et al., 1997; Monteiro et al., 1997; Beldade & Brakefield, 2003). Genetic coupling among eyespots, perhaps due to linkage or pleiotropy, prompted the idea that the entire border ocelli functions as a discrete integrated unit separate from other pattern elements on the wing (Brakefield & French, 1999; Brakefield, 2001). Despite evidence of developmental integration among eyespots,

many studies have also documented modularity (individuality) between wing compartments including selection experiments uncoupling the size of eyespots on the same wing surface (Brakefield et al., 1996; Monteiro et al., 1997; Beldade et al., 2002). These results suggest that eyespots in different wing cells are regulated independently either by different networks, sub-networks or differences in network sensitivity. Most of these studies have been conducted in the African butterfly *Bicyclus anynana* (but see Breuker et al., 2007) and reveal a complex picture where different levels of integration and independence regulate eyespot development.

Modularity is a common theme in organismal development, enabling spatial partitioning of semi- autonomous developmental units that are then free to diversify in function or morphology (Magwene, 2001; Allen, 2008; Klingenberg, 2008). In this way, modularity can promote flexibility during development (Breuker et al. 2007). Modularity may facilitate rapid responses to environmental heterogeneity through phenotypic plasticity by permitting independent networks or sub-networks to be induced by environmental cues (Breuker et al., 2006; Moczek, 2010; Snell-Rood et al., 2010). This idea is pertinent to eyespot development as many butterflies exhibit phenotypic plasticity in response to changing environmental conditions with some eyespots exhibiting greater sensitivity than others (Brakefield et al., 1996; Gibbs & Breuker, 2006; Mateus et al., 2014). Variability in plasticity across eyespots implies differences in how the underlying developmental network integrates and responds to environmental cues resulting in altered pattern development.

To better understand the evolution of eyespot diversity researchers have used a variety of experimental approaches to modify eyespot development. The most common

approach is to conduct perturbation experiments (cautery, temperature shock, injection of hormones and pharmacological agents) and examine which aspects of pattern development are modified (Nijhout, 1984; Takayama & Yoshida, 1997; Serfas & Carroll, 2005; Otaki et al., 2005; Mateus et al., 2014). Although modifications are not always representative of phenotypic plasticity in wild populations, some of these modifications mimic patterns found in related species or resemble aberrant wing patterns that are occasionally observed in nature (Nijhout, 1984; Otaki, 2007). These studies provide a way to dissect the underlying organization of eyespot development in the absence of transgenic tools (Serfas & Carroll, 2005; Otaki, 2008). Studying which eyespots or eyespot features are more susceptible to modification may reveal developmental biases or constraints that have influenced the evolution of eyespot diversity (Nijhout, 1984; Brakefield, 2001).

Here we use heat shock and heparin injections on the painted lady butterfly *Vanessa cardui* (Nymphalidae) to explore phenotypic plasticity and look for evidence of integration and modularity in eyespot development. Heparin is an extracellular proteoglycan that modifies morphogen gradients by a variety of mechanisms including controlling diffusion, signaling, and intracellular trafficking (Yan & Lin, 2009). Heparin is thought to influence secretion of a cold shock hormone and has been shown to modify wing color patterns in other species of butterflies, although the precise mechanism by which this occurs remains unknown (Serfas & Carroll, 2005; Martin et al., 2012). A number of studies have also shown that *V. cardui* is sensitive to heat and cold shock resulting in modified wing patterns (Otaki & Yamamoto, 2004; Otaki, 2007). Using these treatments as a tool, we examine whether eyespots or eyespot traits (color and size) in *V.*

cardui are equally sensitive to perturbations or whether responses to these treatments are eyespot or trait specific. Finally we examine whether phenotypic plasticity is associated with altered patterns of integration and modularity for different eyespot traits.

Methods

Butterfly Rearing and Experimental Setup

Vanessa cardui caterpillars and artificial diet were purchased from Carolina Biological Supply Company (carolina.com). Caterpillars arrived as 2nd instar larvae and were randomly assigned to the following treatment groups, temperature shock, heparin injection and control. Caterpillars were reared in individual containers in ambient conditions (23°C) under a 12-hour light : dark cycle and fed 3 grams of artificial diet every other day until pupation at approximately 7 days. Within 12 hours of pupation caterpillars assigned to the temperature treatment were transferred to an incubator set at 37°C for 48 hours. A stock solution of heparin (Sigma Aldrich) (5µg/ul) was prepared by diluting 0.005g in 1 ml of UV treated water. Caterpillars assigned to the heparin group were injected with 2ul (10µg) of heparin using a Hamilton needle (2mm) within 12 hours of pupation. Injections were performed at the margin of the left wing and the needle was cleaned with 70% ethanol between each injection. Sham injections were also performed with and without water for a small group of pupae to ensure there was no effect of the needle or water injection on the wing phenotype. Following the treatments, all pupae were returned to the same rearing conditions as the control group. All pupae were placed in clean containers with paper tissue to act as support material to ensure successful eclosion of butterflies. Throughout the experiment, containers containing caterpillars and pupae were arranged with alternating treatment groups to reduce potential variation in environmental conditions.

Eyespot Color Analysis

Following eclosion, butterflies were immediately placed at -20°C for storage until wing pattern analysis. Butterflies were spread on a pinning board for a minimum of 3 days. Hind wings of the butterflies were carefully removed and images captured using a Canon digital camera (DP70) mounted on an Olympus SZX12 microscope. Images of the ventral surface included a scale bar. Morphometric measurements were made using ImageJ (NIH) and the number of pixels was calibrated to one cm. Unmanipulated *V. cardui* have a series of marginal eyespots composed of five major concentric rings; an outer black border, a yellow ring, an orange ring, blue coloration, and finally a black center (Figure 3.1 + 3.2). The area of the wing was measured along with eyespot area, and area of the individual colored rings.

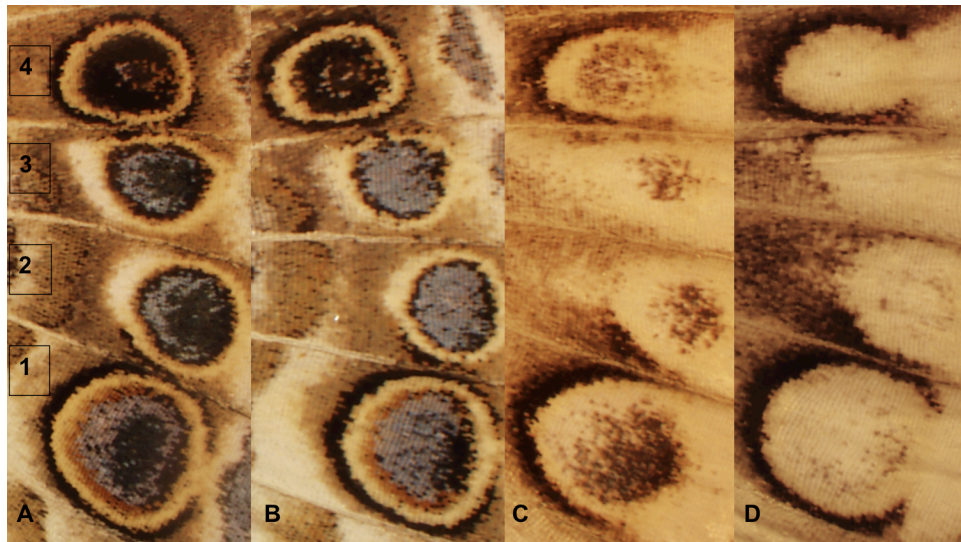


Figure 3.1. Images of eyespots from the different treatment groups illustrating representative phenotypes. **Panel A** Control, **Panel B** Temperature shock, **Panel C** Heparin, **Panel D** Heparin extreme phenotype with complete loss of central eyespots. (1-4 = Eyespots 1-4).

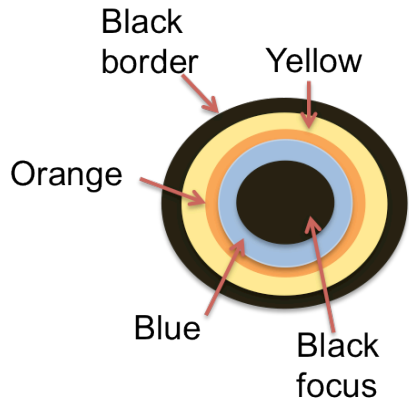


Figure 3.2. Color composition of eyespot in *Vanessa cardui*. Each of the four eyespots represents a combination of these different color elements that were measured for the color pattern analysis.

Total eyespot area and sum of the area of the different colored pigments were compared throughout the process to ensure measurements were consistent. This entire experiment was replicated three times (May 2013, June and September 2014). The heparin treatment was only included in the September 2014 experiment as a single study. Heparin eyespots were measured twice and the average of these measurements was used in the final analyses.

Statistical Analyses

Measurements of total eyespot area and sum of eyespot pigments were assessed using bivariate correlations to ensure high repeatability between the two sets of measurements for each eyespot. A two-way ANOVA was performed using Type III sums of squares to compare differences in eyespot size and area of the different pigments both within the control group and also between the control and treatment groups (Control n = 45, Heparin n = 20, Temperature shock n = 46). The model consisted of treatment and eyespot as main effects, overall wing size as a covariate for eyespot size, date as a blocking factor for

replicated experiments and interaction terms for treatment x eyespot, wing x treatment and date x treatment where appropriate. All data were examined for normality of residuals (Shapiro Wilk $p > 0.05$) and equal variance (Levene Test $p > 0.05$). A Dunn-Šidák correction for multiple comparisons was used when the interaction term was significant. All analyses were performed in JMP version 11 (SAS, Institute, Cary, NC). For analyses of color proportions relative to eyespot size, data were analyzed with a one-way ANOVA or non-parametric tests including Welch ANOVA and Kruskal-Wallis. The proportion of each color was also analyzed to examine whether changes in pigmentation were simply related to eyespot size or whether there were specific changes in the relative size of each ring. Eyespots were analyzed individually due to the dramatic effects of heparin on eyespot size resulting in a large number of outliers for the whole model. Heparin was excluded from some of these analyses due to the presence of many zero measurements.

Associations among eyespots and eyespot traits were assessed using measures of conditional independence (modularity) and graphical modeling as described in Magwene (2001). Graphical Modeling was conducted for eyespots from the control and temperature shock treatment only. Partial correlations between all eyespots for eyespot size and % area of each pigment ring were obtained for the control and temperature treatments separately using multiple regression after checking for normality and equal variances between all associations (Control $n = 45$, Temperature shock $n = 46$). Wing size was included as a covariate. Partial correlations represent the association between two variables (size or color) after accounting for correlations among all other variables. Edge Exclusion Deviance (EED) is a theoretical measure of whether a particular edge can be

eliminated from a saturated model of complete integration Magwene (2001). EED was calculated using the following formula $-N \ln. [1-p_{ij}^2]$ where N represents sample size and p_{ij} is the partial correlation between two variables. The strength of the edge (correlation) was calculated using $-(0.5) \ln. [1-p_{ij}^2]$. The value of each EED is tested against the χ^2 distribution with one degree of freedom. Values less than 3.84 ($p < 0.05$) are rejected as having an edge i.e. the traits are not significantly integrated thus inferring modularity (Allen, 2008). Values greater than 3.84 indicate the traits are developmentally integrated (not conditionally independent). The matrix of EED values was used to construct a graphical model illustrating patterns of integration and modularity among all eyespots. Eyespots that are not connected i.e. those with no edge are inferred to be conditionally independent (modular). Integrated eyespots are those with significant edges that are correlated independent of their associations with the other eyespots.

Results

All results for multiple comparisons using the Dunn-Šidák post hoc test are presented in Appendix Table 3.1. Bivariate plots for regression analyses of total eyespot area and sum of pigment area revealed a high correlation between these two sets of measurements for all eyespots across each treatment ($r^2 = 0.95 - 0.99$) (Appendix Figure 3.3). A one-way ANOVA was conducted to compare wing area across the treatment groups using date as a covariate. There was no date by treatment interaction for the replicated experiments ($F_{(2, 313)} = 1.9, p < 0.14$) and no effect of either treatment on wing area ($F_{(2, 92)} = 2.8, p = 0.07$), however date was significant ($F_{(2, 92)} = 7.7, p = 0.001$) with larger wings observed in the experiment conducted in June. As wing area was not significantly different between

treatments it was used as a covariate in a 2 way ANOVA to examine differences in eyespot size. The sham injections had no effect on any trait examined (data not shown).

Eyespot Size

For eyespot area, there was no date ($F_{(2, 302)} = 0.05, p = 0.9$), date by treatment interaction ($F_{(2, 94)} = 1.6, p = 0.2$) or date by wing interaction ($F_{(2, 380)} = 1.73, p = 0.17$); however, there was a treatment by eyespot interaction ($F_{(2, 380)} = 27.2, p < 0.0001$). Within the control group, eyespot 1 was significantly larger than eyespots 2, 3 and 4. Eyespot 2 was larger than eyespot 3 although there was no difference in size between eyespots 2 and 3 compared to eyespot 4. The treatments had variable effects on eyespot size; heparin dramatically reduced the size of eyespots 2 and 3 often eliminating them entirely while it had no effect on the size of eyespots 1 and 4 (Figure 3.4). Temperature shock significantly reduced the size of all eyespots with the exception of eyespot 4.

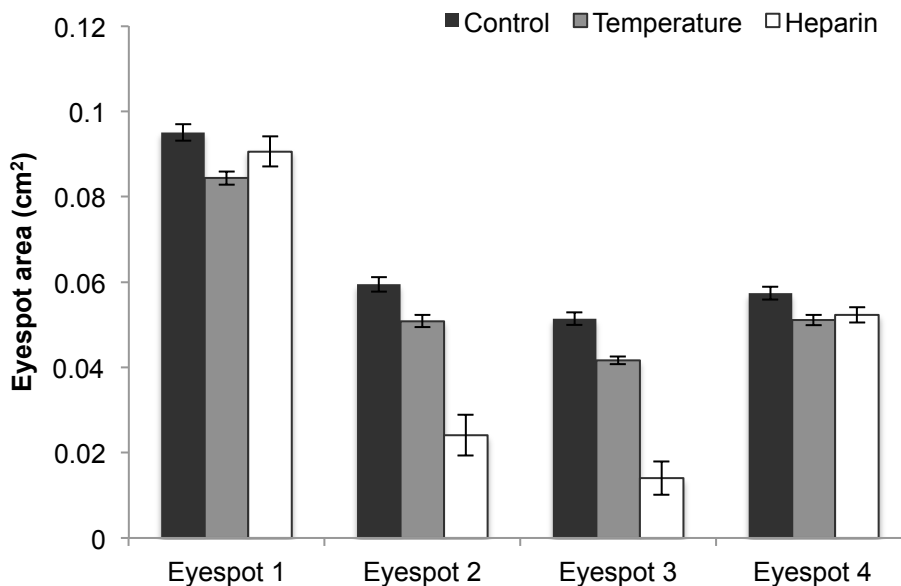


Figure 3.4 Eyespot size across all four eyespots in the different treatments. Data represent the area in cm² with error bars representing 1 SE from the mean. The 2-way ANOVA was performed on square-root transformed data.

Black Border (Outer Ring)

There was no date by treatment interaction for the replicated experiments ($F_{(2, 348)} = 2.0$, $p = 0.12$) although date was significant ($F_{(2, 424)} = 10.45$, $p < 0.0001$); butterflies reared in June had a larger black margin compared to those reared in May and September. Within the control group, the area of the black border varies across all eyespots. Eyespot 1 exhibits the largest border, followed by eyespot 4 with the two central eyespots possessing the smallest black border. A significant treatment by eyespot interaction was found ($F_{(6, 424)} = 4.62$, $p < 0.0001$). However, overall neither treatment had an effect on the area of this eyespot ring for most eyespots. The exception is that butterflies treated with heparin had a significantly smaller black border in eyespot 3 compared to control groups (Figure 3.5).

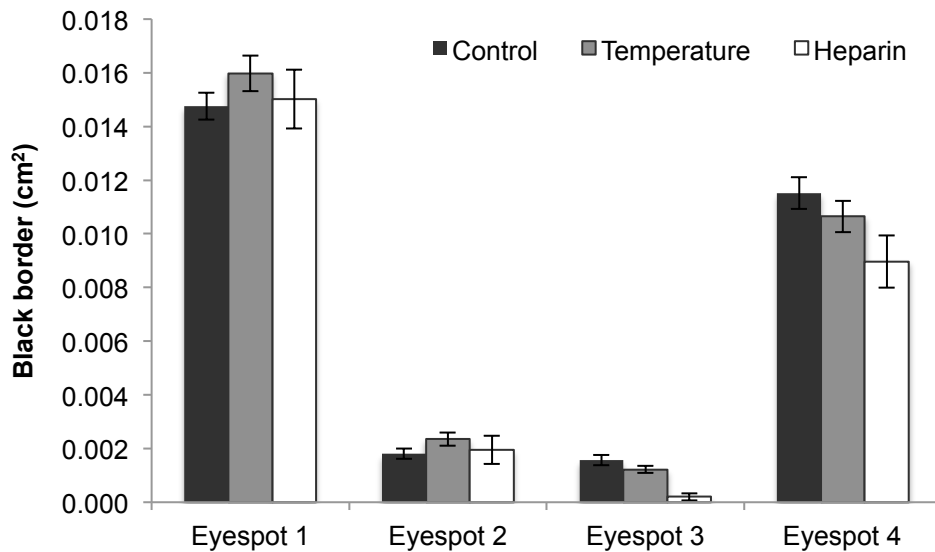


Figure 3.5 Area of the black border (outer ring) across all four eyespots in the different treatments. Data represent the area in cm^2 with error bars representing 1 SE from the mean. The 2-way ANOVA was performed on square-root transformed data.

Differences were observed in the proportion of the black border relative to eyespot size. Butterflies exposed to temperature shock had a proportionally larger black border in eyespot 1 ($F_{(2, 105)} = 7.6, p = 0.001$) and eyespot 2 ($F_{(1, 88)} = 6.24, p < 0.05$). No differences were observed for the other eyespots or the heparin treatment.

Yellow Pigment

Due to the strong effects of heparin on eyespot pigmentation, the analysis of yellow pigment was conducted first by examining the effects of temperature on all eyespots, and then examining the effect of both treatments on eyespots 1 and 4. No date by treatment interaction was found for the replicated experiments ($F_{(2, 348)} = 0.99, p = 0.3$), although butterflies reared in May had overall smaller area of yellow ($F_{(2, 348)} = 9.8, p < 0.0001$). There was a significant treatment by eyespot interaction ($F_{(6, 348)} = 5.15, p = 0.0017$). In the control group, the amount of yellow pigmentation exhibited a gradual decline in area over the four eyespots, with the largest amount of yellow pigment observed in eyespot 1 and the smallest amount in eyespot 4. A similar pattern was also observed in the temperature treatment. Temperature shock had a marginally significant effect on reducing the area of yellow pigmentation in eyespot 1 although there was no effect on the overall proportion of this pigment compared to the control (Figure 3.6). There was a significant reduction in the area of yellow pigment in eyespots 2 and 3.

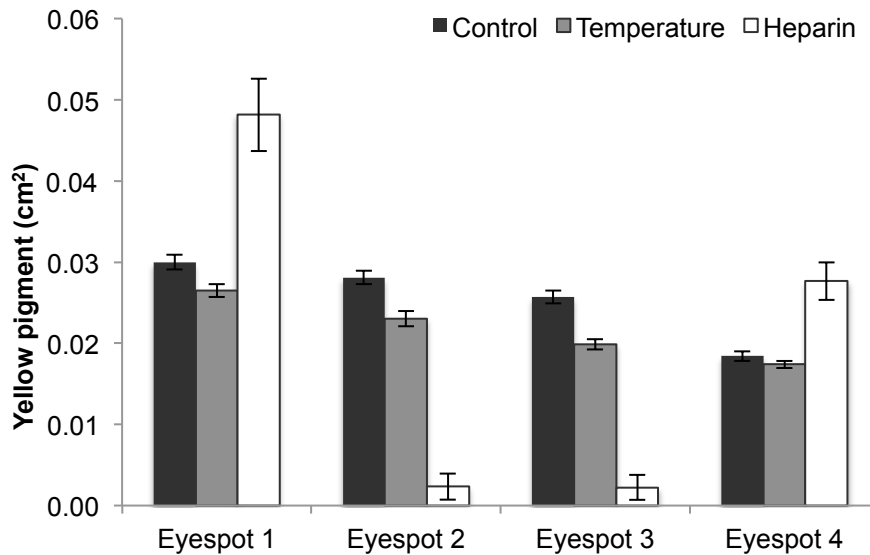


Figure 3.6 Area of yellow pigment across all four eyespots in the different treatments. Data represent the area in cm^2 with error bars representing 1 SE from the mean. The 2-way ANOVA was performed on \log_{10} -transformed data.

When eyespot size was considered, the proportion of yellow pigment was reduced in eyespot 2 ($F_{(1, 87)} = 7.3$, $p = 0.01$) but no reduction was observed in eyespot 3 ($F_{(1, 86)} = 3.5$, $p = 0.28$). There was no effect of temperature shock on the area or proportion of yellow pigment in eyespot 4.

A separate analysis including heparin treatment was done for eyespots 1 and 4 only because heparin abolished eyespots 2 and 3 in many butterflies. For this comparison, there was no treatment x eyespot interaction for yellow pigmentation in eyespots 1 and 4 across all treatments ($F_{(2, 212)} = 0.95$, $p = 0.3885$), however, treatment and eyespot were significant with more yellow pigment observed in eyespot 1 compared to eyespot 4. Heparin significantly increased the area of yellow pigment in both eyespots ($F_{(2, 212)} = 37.73$, $p < 0.0001$) and the proportion of yellow pigment in eyespot 1 ($F_{(2, 102)} = 14.3$, $p = 0.0001$), and eyespot 4 ($F = 13.9$, $p = 0.001$ Welch ANOVA) compared to the control (Figure 3.6).

Orange Pigment

The analysis revealed no date ($F_{(2, 344)} = 0.9, p = 0.4$) or date by treatment interaction ($F_{(2, 344)} = 1.6, p = 0.2$). However, there was a significant treatment by eyespot interaction ($F_{(6, 387)} = 10.8, p < 0.0001$). Within the control group, the area of orange pigmentation was significantly larger in eyespot 1 compared to eyespots 2, 3 and 4 while no difference was observed among the latter eyespots. Temperature shock had no effect on the area of orange pigment in any eyespot; however, the proportion of orange in eyespot 1 was affected by both treatments ($F_{(2, 104)} = 34.43, p = 0.001$) with the highest proportion found in the temperature treatment (Figure 3.7). The proportion of orange pigment in temperature-shocked butterflies was similar to control butterflies for eyespot 2 ($F = 7.6, p = 0.12$, Welch ANOVA, $\chi^2 = 2.8, p = 0.09$) and eyespot 4 ($F_{(1, 82)} = 1.9, p = 0.12$); however, a significantly higher proportion of orange was found in eyespot 3 ($F = 5.8, p = 0.02$ Welch ANOVA, $\chi^2 = 6.0, p = 0.02$).

Heparin significantly reduced the area of orange pigment in eyespot 1 but increased this color in eyespots 2, 3 and 4. The proportion of orange pigment was also significantly reduced in eyespot 1. It was not possible to examine the proportion of orange in the remaining eyespots due to the strong effects of heparin on eyespot size.

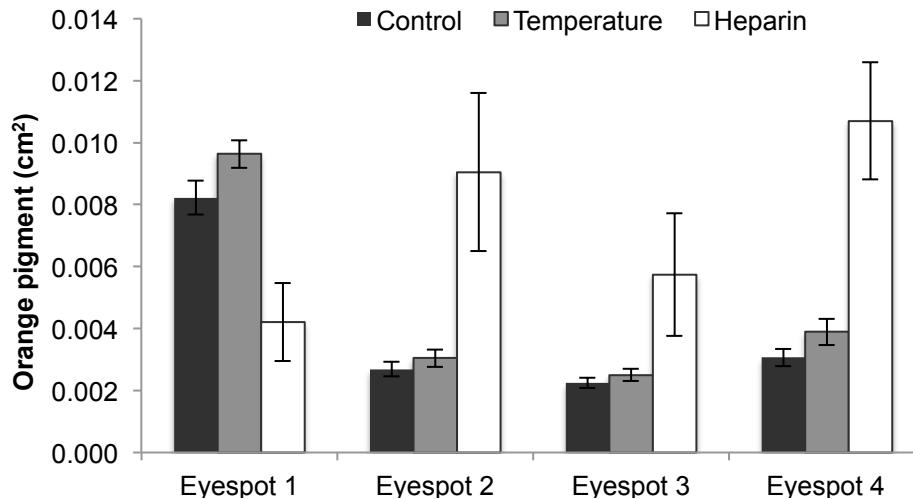


Figure 3.7 Area of orange pigment across all four eyespots in the different treatments. Data represent the area in cm² with error bars representing 1 SE from the mean. The 2-way ANOVA was performed on log₁₀-transformed data.

Blue Scales

Heparin virtually eliminated blue scales in all eyespots, thus only the temperature and control groups were compared for eyespots 1-3 as this structural color is rarely observed in eyespot 4. There was no date ($F_{(2, 261)} = 1.3, p = 0.4$) or date by treatment interaction ($F_{(2, 261)} = 0.9, p = 0.4$). However, there was a significant treatment by eyespot interaction ($F_{(2, 264)} = 15.08, p < 0.0001$). Within the control group, the area of blue was significantly higher in eyespot 1 versus eyespots 2 and 3, however no difference was observed between eyespots 2 and 3. Temperature shock had no effect on the area of blue pigment in eyespot 1 however the treatment significantly increased the amount of blue pigment in eyespots 2 and 3, with the strongest effect observed in eyespot 2 (Figure 3.8). There was no effect of temperature on the proportion of blue in eyespot 1 ($F_{(1, 88)} = 2.3, p = 0.1$). However, the proportion of blue was significantly higher in eyespots 2 ($F_{(1, 86)} = 100.7, p < 0.0001$) and 3 ($F_{(1, 83)} = 55.8, p = 0.0001$) compared to the control.

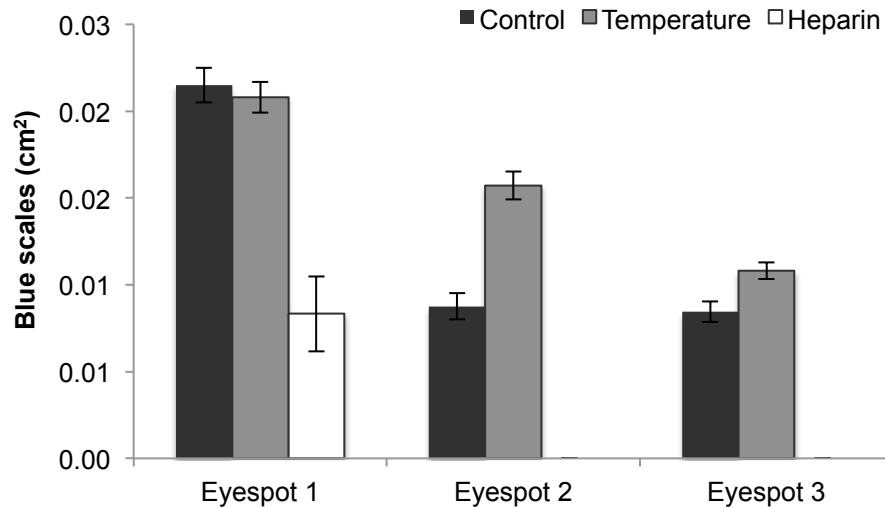


Figure 3.8 Area of blue scales across all four eyespots in the different treatments. Data represent the area in cm^2 with error bars representing 1 SE from the mean. The 2-way ANOVA was performed on square-root transformed data.

Black Focus

For the black focus, there was no date x treatment interaction ($F_{(2, 349)} = 0.7, p = 0.5$) although date was significant. Butterflies reared in September had a smaller black focus compared to those reared in May and June ($F_{(2, 425)} = 15.11, p < 0.0001$). There was a significant treatment by eyespot interaction ($F_{(2, 425)} = 13.77, p < 0.0001$). In the control butterflies, the area of the black focus varied significantly across eyespots. There was no difference in the area of the black focus between eyespot 1 and 2 however comparisons between all other eyespots were significantly different with the smallest focus observed in eyespot 3 and the largest in eyespot 4. Both treatments significantly reduced the size of the black focus in most of the eyespots (Figure 3.9). Heparin did not affect the size of the black focus in eyespot 1; however this treatment significantly reduced the size of the black focus in the three other eyespots with a dramatic reduction in eyespot 4.

Temperature shock significantly reduced the size of the black focus in all eyespots however the strongest effect was observed in eyespot 2.

Both treatments significantly reduced the proportion of the black focus in eyespot 1 ($F_{(2, 103)} = 24.4, p < 0.0001$) and also had significant effects in eyespot 2 ($F_{(2, 103)} = 2.14, p < 0.0001$) and eyespot 3 ($F_{(2, 95)} = 50.12, p < 0.0001$). Heparin significantly increased the proportion of the black focus in the central eyespots, largely because the black focus was all that remained. In contrast, temperature shock reduced the proportion of the black focus in the two central eyespots and also in eyespot 4 ($F_{(1, 82)} = 18.2, p < 0.0001$). Heparin also dramatically reduced the proportion of the black focus in eyespot 4 by eliminating it completely in many individuals.

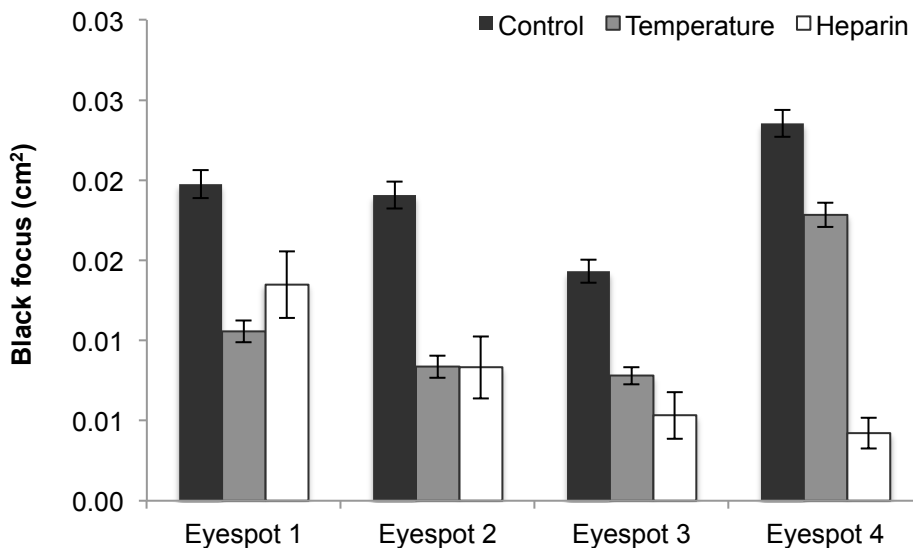


Figure 3.9 Area of the black focus across all four eyespots in the different treatments. Data represent the area in cm^2 with error bars representing 1 SE from the mean. The 2-way ANOVA was performed on square-root transformed data. Sample sizes are shown in the bars.

Tests for Modularity and Integration

Based on results from the phenotype data, we tested the prediction that eyespots 2 and 3 form an independent module that may exhibit weak integration between eyespots 1 and 2. Eyespot 4 would represent a further independent module that also may exhibit weak integration with eyespot 1 for the black border (Figure 3.10). Matrices for partial correlations and EED values are presented in Appendix Figures 3.11 and 3.12 respectively. The graphical model for control eyespots reveals that eyespot size in *V. cardui* is highly integrated across eyespots (Figure 3.13). The strength of integration is highly variable with eyespots 2 and 4 exhibiting the weakest edge strength. Overall, the yellow ring and black focus display similar integration patterns across eyespots with a lack of integration between eyespots 2 and 4. Only orange pigment and the black border showed evidence of integration between these two eyespots, and overall exhibited the lowest levels of integration.

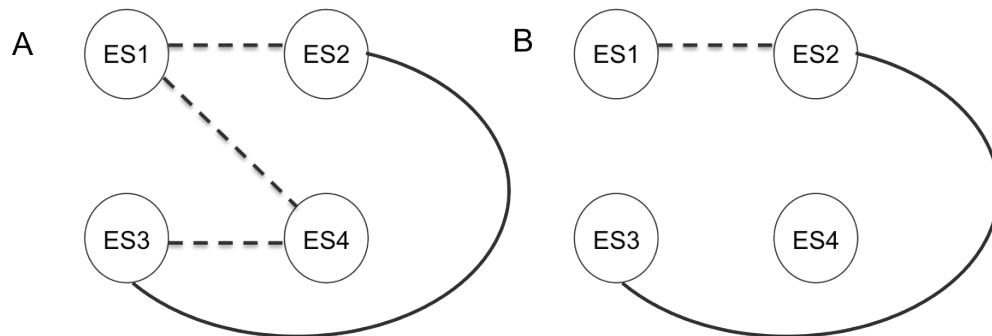


Figure 3.10 Hypotheses of integration and modularity for control eyespots based on overall phenotype data. **A** represents the most intergrated model inferring weak integration between neighboring eyespots and between eyespots 1 and 4 (for the black border). **B** represents the most conservative model showing weak integration only between eyespots 1 and 2.

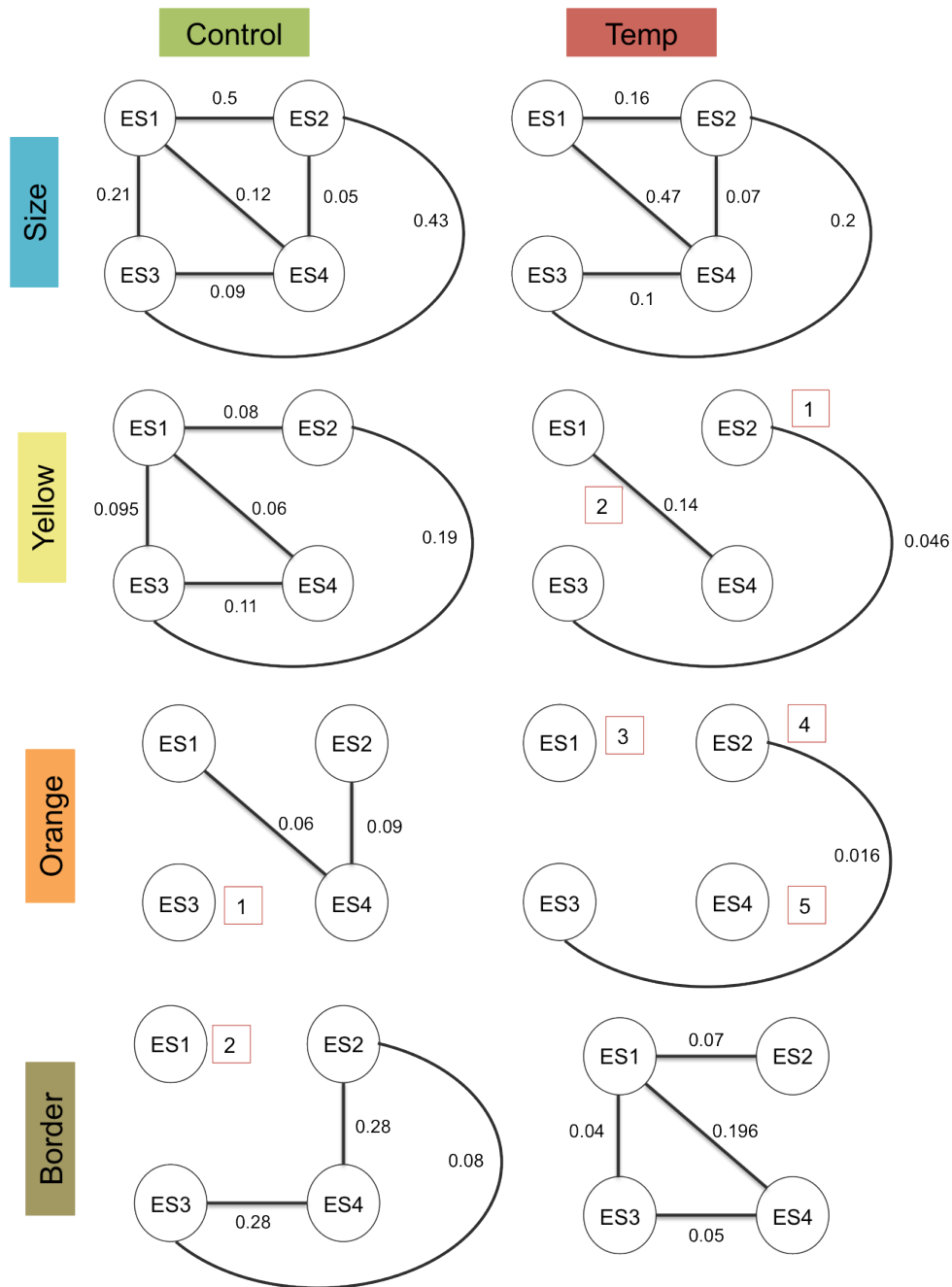


Figure 3.13 Comparison of phenotypic correlations within and between control and temperature shock groups for eyespot size and pigment rings. Edges infer integration among eyespots based on edge exclusion deviance values above the critical value from the χ^2 distribution. Eyespots without edges are conditionally independent from all other eyespots, $\chi^2 < 3.84$, $p < 0.05$. Edge strength values represent the strength of integration between eyespots. Temperature shock results in a doubling of independent modules (shown by red numbered boxes).

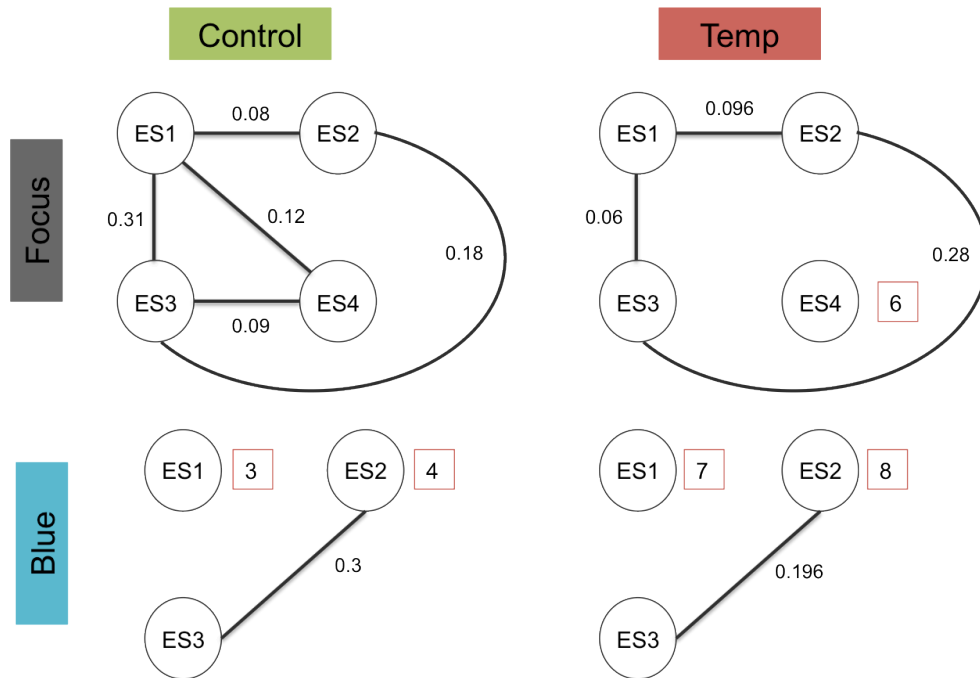


Figure 3.13 cont. Comparison of phenotypic correlations within and between control and temperature shock groups for eyespot size and pigment rings. Edges infer integration among eyespots based on edge exclusion deviance values above the critical value from the χ^2 distribution. Eyespots without edges are conditionally independent from all other eyespots, $\chi^2 < 3.84$, $p < 0.05$. Edge strength values represent the strength of integration between eyespots. Temperature shock results in a doubling of independent modules (shown by red numbered boxes).

These results suggest that for many traits the eyespots are largely integrated with some eyespots demonstrating modularity for certain pigments. Following temperature shock, patterns of integration were largely intact for eyespot size, with only one edge removed between eyespots 1 and 3. Edge strengths weakened for some eyespots but increased for others. The pigment genes displayed very different patterns of integration compared to controls with an overall loss of integration among eyespots resulting in a doubling of independent modules. The yellow ring and black focus were reduced to two independent modules and the orange pigment was reduced to three. Examination of the blue ring revealed an independent module composed of eyespots 2 and 3 in both the control and

temperature group. These modules reflect changes in the proportions of color for temperature shock relative to the control (Table 3.2).

Table 3.2. Percent difference in area (cm²) for eyespot size and pigmentation in response to temperature shock (37°C, 48 hrs.) relative to the control. Module number represents groups identified using graphical modeling (Fig. 9) that corresponds to percent changes in pigmentation.

	Eyespot no.	Temperature	Module
Spot size	1	-11.2	0
	2	-14.7	0
	3	-19.0	0
	4	-10.9	0
Black focus	1	-46.5	1
	2	-56.1	1
	3	-45.5	1
	4	-24.2	2
Blue	1	-3.2	1
	2	44.3	2
	3	21.8	2
Orange	1	14.6	1
	2	11.7	2
	3	10.0	2
	4	21.1	3
Yellow	1	-11.7	1
	2	-18.1	2
	3	-22.8	2
	4	-5.6	1
Black border	1	7.6	0
	2	23.3	0
	3	-22.4	0
	4	-9.4	0

The black border exhibited an increase in patterns of integration between eyespots and a new edge was created between eyespots 2 and 3 for the orange pigment. Overall, in both the control and treatment groups, the two central eyespots retained the highest number of edges between each other across all traits (Table 3.3) and displayed some of the highest edge strength values. We also observed a general trend of concerted changes in response

to the treatment. Yellow pigment rings and the black focus declined in response to temperature shock across all eyespots while the orange pigment ring and blue region increased.

Table. 3.3 Total number of edges across all traits between each pair of eyespots for the Control and temperature shock (37°C, 48 hrs.) groups. The two central eyespots (ES2 and ES3) are highlighted as showing the highest number of edges and no change in edge number following temperature shock.

	Control	Temperature shock	Difference
ES1 vs. ES2	3	3	0
ES1 vs. ES3	3	2	-1
ES1 vs. ES4	4	3	-1
ES2 vs. ES3	5	5	0
ES2 vs. ES4	4	2	-2
ES3 vs. ES4	3	1	-1

Discussion

We have found that hind wing eyespots of the butterfly *Vanessa cardui* exhibit phenotypic plasticity in response to temperature shock and injection of heparin sulfate. Temperature shock had subtle but significant effects on eyespot development, while eyespot morphology was dramatically altered by heparin sulfate. Despite the use of two very different treatments we observed common effects on eyespot development indicating that some underlying property of the eyespots influence pattern modification. We found that eyespots in *V. cardui* exhibit both concerted and individual responses to environmental perturbation providing evidence for developmental integration for eyespot size, but also modularity for individual eyespots and specific eyespot elements. Our

results suggest that particular eyespots and eye spot elements are sensitive to environmental perturbations.

Eyespot Size Plasticity Varies Across Wing Segments

Eyespots often vary in size across the wings of butterflies with some eyespots appearing more conspicuous than others. In *V. cardui* the posterior eyespot (eyespot 1) was found to be significantly larger than the other three eyespots, which were all similar in size.

Wings with large eyespots or a series of three small eyespots are known to confer a protective benefit by intimidating predators (Stevens et al., 2008). Thus, predation may have shaped this particular combination of eyespot sizes in *V. cardui*. Given that eyespot size is a plastic trait in many butterfly species (Brakefield et al., 1996; Lyytinen et al. 2004), we examined whether eyespot size in *V. cardui* is also phenotypically plastic and whether plasticity varies across eyespots.

Our experiments revealed that both treatments had significant and variable effects on size of different eyespots independent of wing area. Temperature shock significantly reduced the size of eyespots 1- 3 while heparin significantly reduced the size of eyespots 2 and 3. While similar reductions in eyespot size have been reported following heparin injections in another Nymphalid, *Junonia coenia* (Serfas & Carroll, 2005) temperature shifts have been shown to have opposite effects in *B. anynana* (Brakefield et al., 1998; Oliver, et al., 2013). Thus, temperature seems to have complex effects on butterfly eyespot development, increasing eyespot size in some species while decreasing it in others.

Grafting experiments have revealed that eyespot size is regulated by some property of the focal organizing cells such as a morphogen signal (Monteiro, et al., 1994;

French & Brakefield, 1995; Beldade et al., 2008). Thus, warmer temperatures may increase the concentration of this signal in *B. anynana* while weakening the focal signal in *V. cardui*. The precise mechanisms by which temperature and heparin alter eyespot size remain unknown (Serfas & Carroll, 2005; Oliver et al., 2013) although many hypotheses have been proposed to explain plasticity of eyespot size (Brakefield et al., 1998; Oliver et al., 2013). Research by Oliver et al. (2013) in *B. anynana* has shown that while changes in temperature do not alter the temporal order of gene expression for eyespot associated genes; cooler temperatures lead to earlier onset of gene expression. Thus, temperature may induce heterochronic shifts in expression of eyespot-associated genes (Oliver et al., 2013). In addition to temperature-induced variation in expression of eyespot genes, hormone titers have also been implicated in regulating eyespot size plasticity (Brakefield, 1996). These treatments may have reduced eyespot size by altering the dynamics of hormone signaling and/or selector genes, or by modifying the competency of scale cells to respond appropriately to morphogen signals (Brakefield, 2001). Whatever the mechanism that resulted in modified eyespots, in both treatments the two central eyespots were most affected.

Interestingly, neither temperature shock nor heparin affected the size of eyespot 4 suggesting it is less plastic. These results demonstrate that eyespot plasticity varies across the wing in *V. cardui*. Centrally positioned eyespots appear more vulnerable to environmental perturbations. Although eyespots are formed from a common developmental program, eyespot development in *V. cardui* may fall into three independent developmental modules: **1.** eyespot 1, **2.** eyespots 2 and 3, and **3.** eyespot 4) that exhibit different levels of phenotypic plasticity (Figure 3.8). These developmental

modules may have evolved due to differences in the focal signal, or variation in the timing of eyespot development driven by anterior-posterior gradients in transcription factors across the wing (Keys et al., 1999; Monteiro, 2014). Whether eyespot size plasticity has evolved in response to selection for mate attraction or predator avoidance in *V. cardui* is currently unknown.

Inner Pigment Rings Are More Sensitive to Perturbations Than the Eyespot Border

Eyespots are composed of a series of concentric rings of different colors that also exhibit phenotypic plasticity. Variation in eyespot rings in response to environmental conditions may provide opportunities for the generation of novel eyespot patterns if certain eyespot rings are easily modified. In *V. cardui*, exposure to heparin and temperature shock produces an overall simplification of the eyespot pattern. Simplification is particularly evident in the heparin treatment, which virtually eliminated eyespots 2 and 3 and strongly impacted the inner rings of eyespots 1 and 4. The bleaching effects of heparin were strikingly similar to those observed in *V. cardui* following cold shock (Nijhout, 1984) and sodium tungstate (Otaki & Yamamoto, 2004). Similar to previous work (Serfas & Carroll, 2005), we found individual variation in the response to heparin, with some butterflies exhibiting stronger effects on eyespot development than others (Figure 3.1 C + D). This pattern is also observed in butterflies exposed to cold shock with a range of aberrant forms produced. However, heparin does not have similar effects in all butterfly species. In contrast to a bleaching effect, Martin et al. (2012) found that heparin increases melanization in *Heliconius* butterflies. Heparin is thought to influence wingless signaling which may be involved in promoting melanin synthesis in *Heliconius*, but suppressing its production in other butterflies.

Interestingly, both cold shock (Nijhout 2001) and heparin show similar effects where the black border or outer ring is more resistant to modification and in many individuals is the only remaining eyespot ring. Serfas and Carroll (2005) also found similar effects of heparin on the black outer ring in eyespots of *Junonia coenia*. In our experiments, neither temperature shock nor heparin had any effect on the overall area of the black outer ring, although the proportion was increased by temperature shock when controlling for size in eyespots 1 and 2. Interestingly, heparin altered the color of the dorsal hind wing eyespots changing their color from black to white, a phenotype that closely resembles aberrant forms of the sister species *V. kershawi* (Otaki, 2007). In fact many wing pattern modifications in response to cold shock, heparin and sodium tungstate resemble the range of phenotypes observed across different species within *Vanessa* including occasional wild-caught aberrant individuals of *V. cardui* (Otaki & Yamamoto, 2004). These observations suggest developmental constraints in eyespot formation resulting in a limited range of phenotypic possibilities.

Black pigmentation is also found in the focal region of the eyespot. In heparin treated butterflies, the two central eyespots were mostly eliminated leaving just partial remnants of the black focus, some of which developed as orange pigment. Temperature shock also significantly reduced the black focus in all eyespots with the strongest effects occurring in the two central eyespots. Although temperature shock did not produce any bleaching effect, it did increase the amount, both area and proportion, of the structural blue color in the focal region of the two central eyespots, particularly for eyespot 2. Thus, in the two central eyespots the black focus appears to have been partially replaced or masked by expansion of blue. The blue region may also have expanded at the expense of

the yellow ring, which also decreased in response to temperature shock for all eyespots with the exception of eyespot 4.

Replacement and expansion of pigment rings have also been observed in other butterflies. In the Goldeneye mutant of *B. anynana*, the outer ring of gold scales replaced the inner ring of black scales due to changes in the expression domains of transcription factors. In these mutants, Spalt expression was replaced by Engrailed/Invected in scales that originally developed black pigment (Brunetti et al., 2001). Expansion of the blue ring in response to temperature shock observed here may have occurred via a similar mechanism. Spalt also corresponds to black pigmentation in *V. cardui*; thus, temperature shock may repress Spalt expression and/or increase expression of Engrailed/Invected. Immunolabeling experiments would be required to test this hypothesis.

Do Gradient Models Adequately Explain How the Same Pigment Develops in Different Rings?

These results and those of others (Nijhout, 1984; French & Brakefield, 1995) suggest that the outer ring of eyespots is less vulnerable to perturbations than the inner pigment rings. These observations indicate that expression of morphogens or transcription factors are less stable in the focal region. For *V. cardui* the most sensitive eyespot ring appears to be the black focus, which was significantly modified in all eyespots by both treatments. Several studies have shown that temperature shock can reduce melanin pigmentation (Gibert et al. 2007; Otaki, 2007); however, in the case of *V. cardui* eyespots it is not melanin *per se* that is affected, but where this pigment is localized that influences its vulnerability to perturbation. These results raise two important questions regarding the development of concentric colored rings. First: how can a concentration gradient produce the same color in both the focus and the outer most ring? Second: how do environmental

perturbations significantly modify the inner pigment rings while not affecting the outermost ring?

According to current models, the eyespot develops from an organizing focus at the center of the eyespot, which contains the highest concentration of a putative morphogen signal. This morphogen diffuses radially through the wing epidermis to produce concentric rings of colored scales (Nijhout, 1978; French & Brakefield, 1995; Brakefield & French, 1999). Thus, each ring represents a different threshold response to the diffusing morphogen resulting in a signal transduction cascade leading to the development of different colored pigments (Monteiro, 2014).

According to this model both low and high concentrations of the putative morphogen leads to the same colored pigment, suggesting that these two eyespot regions are producing melanin via alternative mechanisms. Spatial differences in interactions between hormones and transcription factors could influence how epidermal cells interpret the signal, leading to the synthesis of the same pigment. Immunolabeling however does not show evidence of similar expression patterns in the focus and outer rings during the first 24 hours of pupation. Fluorescent labeling of eyespot 4 in *V. cardui* shows that different genes are expressed in the inner and outer rings with co-expression of Distal-less and Spalt in the focus and Engrailed in the outer ring (Brunetti et al., 2001). Work by Brunetti et al. (2001) suggests that similar pigments produced at the ends of the concentration gradient are due to expression of different eyespot genes. These different expression profiles may also explain why the outer ring is less vulnerable to perturbation if certain expression domains are more resistant to environmental changes. It also suggests that different genes are involved in producing the same pigment.

It is important to note that expression patterns vary substantially during wing color pattern development (Reed et al. 2007) and we do not yet have a detailed time series of expression patterns during eyespot development but rather snapshots at particular time points. Clearly, variation in threshold responses to a concentration gradient cannot explain the same pigment developing in different rings. Further work is required to develop models that adequately explain the diversity of different eyespot pigment patterns including those composed of different rings of the same pigment.

Temperature Shock Alters Patterns of Integration and Modularity

Studies investigating integration and modularity of butterfly eyespots have produced conflicting results, suggesting that developmental processes influencing multi-trait correlations are complex and labile. We investigated patterns of integration and modularity in *V. cardui* and examined whether phenotypic plasticity is associated with changes in phenotypic correlations among eyespot traits. We tested the hypothesis that patterns of integration and modularity would follow those described in Figure 3.8 and that these patterns would explain variation in phenotypic plasticity. We did not find support for our hypothesis; the graphical models revealed that most traits were largely integrated across eyespots in control butterflies with some evidence of modularity in eyespot 1 and 3. Contrary to studies documenting strong correlations between neighboring eyespots, we observed that eyespots 1 and 2 exhibited conditional independence for several pigments although they were highly correlated for size. The two central eyespots, however, do appear to be more tightly co-regulated as they displayed the highest number of edges and among the strongest edge strength values. Fewer edges were observed for the orange and blue ring and black border, indicating stronger

integration for eyespot size than pigmentation suggesting greater independence between eyespots during pigment synthesis.

A growing number of studies have demonstrated that phenotypic correlations among traits are not necessarily static and are modified in different environmental conditions (Schlichting, 1989; Urren et al., 2002; Plaistow & Collin, 2004; Montague et al., 2012). We also found that temperature shock altered patterns of integration and modularity in *V. cardui*. Overall, there was a loss of integration among eyespots with a corresponding doubling of independent modules. Thus, eyespot modularity did not predict plasticity: (e.g. the two central eyespots were integrated with the other eyespots in the control butterflies), but phenotypic plasticity generated novel independent modules. The two central eyespots retained their connections to each other for all traits, with exception of the black border, providing further evidence that they are strongly integrated. Patterns observed for the yellow, orange and blue ring and black focus reveal distinct modules that closely correspond to changes in the proportion of these different pigments. In addition to an increase in modularity, new correlation patterns emerged that were not present in the controls. Temperature -induced plasticity generates novel patterns of integration and modularity, which may be due to variation in plasticity between different traits. These results lend support to the idea that phenotypic correlations are dependent on environmental conditions (Schlichting, 1989).

Conclusions

Overall, we found that the central eyespots and inner pigment rings were highly sensitive to modification. In contrast, eyespot 4 and the black border were more resistant to perturbation. This variation between eyespots has also been observed in other species.

Gibbs & Breuker, (2006) found that one of the hind wing eyespots (HW-OC4) in *Precis aegeria*, was more sensitive to resource shortage during development compared to the other eyespots. We observed that the central eyespots appear to have linked responses to stimuli; this is also observed in other species. A number of studies have also revealed that neighboring eyespots appear to be more highly integrated compared to distant eyespots (Beldade et al., 2002; Breuker et al. 2007; Allen, 2008). Monteiro et al. (2003) discovered a mutant produced by x-rays with reduced eyespots 3 and 4, yet the remaining hind wing eyespots were unaffected. Why certain eyespots or eyespot rings are more sensitive to perturbation remains unclear. Hormones or morphogens that influence pigment pathways seem to be subject to alteration in specific regions of the wing. It is clear however that changes in environmental conditions disrupt patterns of integration and promote modularity leading to variation in phenotypic plasticity and novel eyespot patterns. Whether environmentally induced phenotypic correlations influence selective regimes and ultimately the evolution of novel phenotypes still remains a topic of intense investigation (Schlichting & Smith, 2002; Snell-rood et al., 2010; Fitzpatrick, 2012; Montague et al., 2012).

CHAPTER IV

TEMPERATURE SHOCK AND HEPARIN ALTER EXPRESSION OF GENES INVOLVED IN EYESPOT DEVELOPMENT

Abstract

Phenotypic plasticity in wing color patterns is commonly observed in butterflies; however, underlying changes in gene expression that cause plasticity are poorly understood. Environmentally induced changes in wing color patterns may be caused by changes in patterning/pigmentation genes or upstream regulators such as epigenetic modifiers. Here we investigated whether melanin suppression in eyespots following temperature shock and heparin injections was caused by upregulation of an epigenetic silencer, the polycomb repressive complex (PRC) and associated down-regulation of patterning and pigment genes. We first investigated whether the PRC is expressed during butterfly wing development. We found that all major genes from PRC1 and the PRC2 are dynamically expressed and peak in expression during late larval and early pupal stages. Despite evidence of PRC expression in butterfly wing tissue we did not observe treatment effects on transcription of two PRC genes, *polycomb* and *enhancer of zeste* in modified eyespots though *polycomb* was expressed at significantly higher levels than *enhancer of zeste*. Temperature shock significantly increased expression of *tan*, *spalt* and *engrailed*,

while heparin increased expression of *distal-less* and *engrailed*. Expression of *engrailed* was altered in both treatments, suggesting it is a highly sensitive gene. This work expands current knowledge on genes expressed in butterfly eyespots and provides a foundation for future studies investigating epigenetic regulation of wing color patterns.

Introduction

Many organisms respond to changes in the environment by modifying their phenotype, a phenomenon referred to as phenotypic plasticity (West-erberhard, 1989; Schlichting & Smith, 2002; Pigliucci, 2005). These developmental adjustments to the environment may be driven by a generalized stress response or an adaptive phenotypic change that leads to enhanced fitness. Precisely how environmental factors shape the developmental trajectory of an organism is poorly understood, but likely involves complex interactions across multiple levels including genetic and epigenetic regulation. Thus, a key step towards understanding the molecular basis of phenotypic plasticity will require identifying environmentally sensitive genes or networks that generate alternative phenotypes.

Butterfly wings are an excellent model system to investigate the molecular basis of phenotypic plasticity. Many butterfly species exhibit distinct seasonal morphs or plasticity in wing color patterns ranging from expansion of melanin pigment across the wing to local changes affecting specific color pattern elements (Roskam & Brakefield, 1999; Otaki, 2008; Simpson et al., 2011). Butterfly eyespots have been widely studied as a highly plastic trait. These spots can vary in size, number and color composition in response to temperature shifts and pharmacological treatments (Nijhout, 1984; Serfas & Carroll, 2005; Dhungel & Otaki, 2009; Monteiro et al., 2013). The focus on butterfly eyespots has resulted in the identification of a number of patterning genes that are

expressed during eyespot development, including selector genes (*distal-less*, *spalt*, *engrailed*), Hox genes (*ultrabithorax*, *antennepedia*), morphogens (*wingless* and *hedgehog*) and receptors (*notch* and *ecdysone receptor*). These patterning genes are part of the general toolkit for animal development; thus, any disruption or modification in their expression has the potential to induce significant changes in morphology.

Temperature, for example is known to alter expression of some of these patterning genes leading to dramatic changes in eyespot size. In the polyphenic butterfly, *Bicyclus anynana*, the conspicuous eyespots of wet season morphs exhibit larger spatial expression of the selector gene *distal-less* compared to the smaller eyespots of dry season morphs.

Variation in temporal expression of *notch* and *engrailed* has also been linked to temperature-induced plasticity in eyespot size, although it is not precisely clear how temperature causes these shifts in gene expression.

Expression of several patterning genes have also been mapped to eyespot rings providing strong evidence of their involvement in pigmentation (Brakefield et al., 1996). *Spalt* and *distal-less* correspond to melanic scales in several butterfly species, suggesting their expression is involved in regulatory switches that trigger upregulation of melanin genes (*pale*, *ddc*, *tan* and *yellow*). Similar to eyespot size, scale color can be modified in response to environmental perturbations as evidenced by seasonal polyphenisms, implying that the regulatory system underlying pigmentation is also flexible.

The precise molecular mechanisms regulating phenotypic plasticity are not well understood although epigenetic mechanisms provide a plausible explanation. Epigenetics involves changes in gene expression without changing the underlying DNA sequence (Feil & Fraga, 2011). Epigenetic mechanisms include DNA methylation and histone

modifications that regulate gene silencing and activation. Polycomb repressive complex 1 (PRC1) and polycomb repressive complex 2 (PRC2) are well-studied epigenetic regulators that respond to internal and external stimuli by catalyzing mono-ubiquitination of lysine 119 of histone 2A and trimethylation of lysine 27 on histone 3, respectively (Margueron & Reinberg, 2011). These histone modifications result in chromatin compaction and silencing of Hox and thousands of other development genes (Feil & Fraga, 2011), including those identified in butterfly eyespots. Early studies of PRC revealed that mutations within the complex lead to ectopic expression of developmental genes outside of their normal domains resulting in dramatic homeotic transformations (Lewis, 1978; Jürgens, 1985; Struhl & Akam, 1985). These studies highlight the crucial role of the PRC in regulating cell fate and differentiation and raise the possibility that environmentally induced changes in PRC expression may mediate phenotypic outcomes.

In insects, PRC1 contains 4 major proteins, Polycomb (*Pc*), Polyhomeotic (*Ph*), Sex combs extra (*Sce/Ring*) and Posterior sex combs (*Psc*), while PRC2 contains three major proteins, Enhancer of zeste (*Ez*), Suppressor of zeste (*Suz12*) and Extra sex combs (*Esc*). Enhancer of zeste catalyzes the trimethylation of lysine 27 on histone 3 (H3K27me3), which leads to propagation of this mark on neighboring histones (Cao et al., 2002; Panning, 2010; van der Velden et al., 2012). Polycomb recognizes this mark through its chromodomain and initiates binding of the PRC1 via catalyzing H2A-K119ub1 (Margueron & Reinberg, 2011), and together these two complexes cooperate to regulate chromatin compaction and gene silencing.

In recent years there has been a surge of interest in the role of epigenetic regulation in phenotypic plasticity in insects. These studies have focused primarily on

DNA methylation in Hymenoptera, revealing that epigenetic control plays an important role in regulating diet-induced plasticity in caste determination (Foret, & Maleszka, 2008; Kucharski et al., 2008; Chittka & Chittka, 2010; Lyko et al., 2010; Weiner & Toth, 2012; Hunt et al., 2013a, 2013b). Epigenetic regulation, mostly in the form of DNA methylation has also been investigated in a number of other insects including the pea aphid, locusts, beetles, and stick insects (Krauss et al, 2009; Hunt et al., 2010; Boerjan et al., 2011; Feliciello et al., 2013). In contrast, there are very few studies investigating epigenetic regulation in Lepidoptera. Although, evidence of DNA methylation has been observed in the silkworm, *Bombyx mori* (Xiang et al., 2010) and the cabbage moth (Mandrioli & Volpi, 2003), *Mamestra brassica* and expression of both PRC complexes has been observed in *B. mori* (Li et al., 2012). However, no published studies have examined epigenetic regulation in butterflies or wing tissue of Lepidoptera.

We have previously demonstrated phenotypic plasticity in the painted lady butterfly, *Vanessa cardui*, in response to temperature shock and heparin injection. Exposure of pupae to temperature shock reduced eyespot size and melanin pigmentation; heparin dramatically altered eyespot formation and pigmentation. These observations led to the hypothesis that alterations in pigmentation may be due to the PRC repressing genes involved in patterning and pigmentation. Here we test this hypothesis by examining 1) whether the PRC is expressed in butterfly wings and 2) whether reduced melanin pigmentation is associated with a down-regulation of patterning/pigment genes and upregulation of the PRC. We also examine how expression of patterning and pigment genes varies across eyespots that differ in size and color composition.

Methods

Transcriptome Analysis

As part of a larger effort to explore the transcriptome during wing color patterning in *V. cardui*, we examined whether the PRC is expressed during wing pattern development, particularly during initial stages of eyespot establishment (larval and early pupal stages). *Vanessa cardui* caterpillars and artificial diet were purchased from Carolina Biological Supply Company (NC). Caterpillars were reared individually at ambient conditions (28°C). Wing discs were dissected from caterpillars at two stages in the fourth instar, the final stage prior to pupation. Prior to dissection, caterpillars were cut in half after the first abdominal segment and the thorax was placed immediately in RNAlater[®] (Ambion). The samples were stored at 4°C for 2 days. Both fore and hind wing discs were carefully dissected from the thorax and placed in RNAlater[®] in preparation for RNA isolation. Pupal wings were dissected from live pupa at 2, 5 and 8 days post-pupation. Following dissections, both fore and hind wings per individual were placed immediately in RNAlater[®] (Ambion) and stored at 4°C. RNazol was used to extract RNA from larval and pupal wing samples. RNA concentration was assessed on a spectrophotometer (A260/A280 >1.8) and integrity was checked on a 1% formaldehyde-agarose gel.

The RNA from each sample was diluted to 25ng/μl in RNase free water. RNA from wing discs (both fore and hindwings) was then pooled into a single biological replicate for each stage. Thus, each pooled sample contained wing discs from 5 individuals for the early 4th and late 4th instar. There were two biological replicates for each pupal stage representing pooled fore and hindwings from 3-5 individuals. The RNA samples were shipped overnight on dry ice to the Huntsman Cancer Institute at the

University of Utah Microarray and Genomic Analysis Facility for library preparation and sequencing using illumina Hiseq 2000 sequencer. Transcriptome assembly and RNA-seq analysis was performed using CLC Genomics 6.5.1. Details for cDNA library preparation, illumina sequencing and transcriptome assembly are provided elsewhere (Chapter 2).

Temperature and Heparin Experiment: Butterfly Rearing and Eyespot Dissections

Vanessa cardui caterpillars and artificial diet were purchased from Carolina Biological Supply Company (NC). Caterpillars were reared individually in a growth chamber at 28°C 16L: 16D cycle. Caterpillars were randomly assigned to treatment groups: 1. Control, 2. Heparin injection and 3. Temperature shock at 37°C. Heparin injections (10µg in 2µl water) were performed using a Hamilton syringe within 12 hours of pupation on the left side of pupa close to the wing. The needle was cleaned with 70% ethanol between injections. For the temperature treatment, pupae were moved to an incubator set at 37°C for 48 hrs. Following treatments, all pupae were returned to the growth chamber. Pupae were harvested for dissections at a time point when pigments are visible on the cuticle (6 days post-pupation) (Appendix Figure 4.1). Pigments were used as a landmark for estimating when pupae reached a similar developmental stage. Images were taken of the pupae to document cuticle pigmentation prior to dissections. Wings were dissected from live pupae (6 days post-pupation) under Zeiss Stemi 2000-C microscope. One of the dissected hind wings was immediately placed in RNAlater® (Life Technologies) and the other hind wing was photographed using an Olympus DP70 camera mounted on an Olympus SZ12 microscope. Images of the hind wings and pupae were carefully examined to select butterflies with similar levels of pigmentation (Appendix Figure 4.1).

The duration of pupation until dissection time was also calculated to compare pupal age in the control and treatment groups (Table 4.1). Butterflies with similar hind wing phenotypes and age were selected for quantitative real time PCR. A subset of butterflies from each group was reared to adults for morphometric analysis of eyespots (Chapter 3).

Table 4.1. Age of pupae at time of eyespot dissection and total duration of pupation for butterflies reared to adults across treatment groups. SE represents 1 standard error from the mean and n = sample size.

	Pupal age at dissection (days)	SE	Pupal age at dissection (hrs.)	SE	N	Age at eclosure (days)	SE	n
Control	6.88	0.13	165.8	2.6	8	7.81	0.16	21
Heparin	6.80	0.13	167.0	1.6	10	7.74	0.14	19
Temperature	6.00	0.0	144.8	1.2	8	6.69	0.18	26

RNA Isolation and Quantitative Real Time PCR

The first, second and third eyespots (Figure 4.2) were carefully dissected from select hind wings and placed immediately in RNAlater. Micro dissections of all three eyespots were conducted on the same wing simultaneously and were performed by alternating between the control and each treatment group to prevent any systematic bias in dissections. Each dissected eyespot was transferred to 100µl of RNazol® RT (Molecular Research Center Inc.) and processed using an electric homogenizer for 1 minute each. Eyespots for all treatment groups were processed at the same time, alternating between treatments and eyespots to reduce any bias in processing. RNA was quantified using a ND-1000 spectrophotometer (NanoDrop products, Wilmington, DE). All samples were treated with Dnase 1 followed by ethanol precipitation. RNA was diluted to 2ng/µl and qRT-PCR was conducted using primers for β-actin to confirm absence of genomic DNA contamination.

Total RNA was reverse transcribed using ABI High Capacity Reverse Transcription kit (Applied Biosystems) using 100ng input RNA. All samples were converted to cDNA in the same run using an ICycler thermocycler (Bio-rad) with the following conditions: 95°C 30s, 95°C 5s, 60°C 5s for 40 cycles. cDNA was diluted with RNAase free water to 2ng/μl for quantitative PCR. Each 10μl reaction consisted of 2μl cDNA (4ng), 5 μl Evagreen Supermix (BioRad), 0.3 μl of forward and reverse primer (10μM), and 2.4μl of RNAase free water. Reactions were performed on a CFX384 Real time system (Bio-rad C1000 Thermocycler) with an initial incubation at 95°C for 30s, 95°C and 60°C for 5s each over 45 cycles.

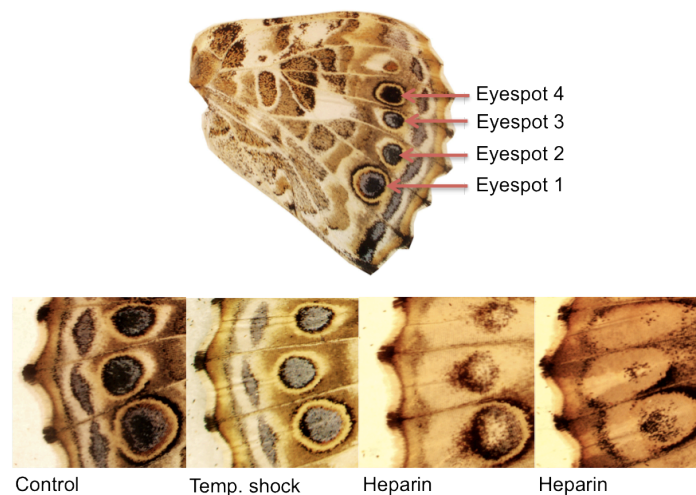


Figure 4.2 Adult hindwing of *Vanessa cardui* with eyespots labelled. Only eyespots 1-3 were dissected for qRT-PCR. The bottom panel shows close up of representative images of eyespots 1-3 for each treatment group. Two images are shown for heparin to illustrate individual variation in response to treatment.

Genes were selected based on studies demonstrating expression in developing eyespots: *engrailed (en)*, *spalt (sal)*, *distal-less (dll)* and *ultra-bithorax (ubx)* (Brunetti et al., 2001; Tong et al., 2014). Genes involved in the melanin synthesis were also examined: *tan*, *ebony*, *ddc* and *pale* (Wittkopp et al., 2003). *Enhancer of zeste (Ez)* was chosen from the

PRC2, as it is involved in the trimethylation of lysine 27 on histone 3. The *polycomb (Pc)* gene of the PRC1 was selected as it recognizes the trimethylation mark and initiates assembly of the PRC1. Primers from the following genes (Appendix Table 4.2) were designed using the software Primer Express (Life Technologies). Primers were based on sequences obtained from the wing transcriptome for *Vanessa cardui* (Chapter 2). Gene identity was confirmed in a tblastx using the *Drosophila* peptide database and a reciprocal blastn in NCBI. All genes retrieved an E value = 0.0 and mapped to homologous sequences in related butterflies or *Bombyx mori*. Primers were used to amplify cDNA to generate a standard curve for qRT-PCR in the following manner: 25ng of cDNA was amplified in a 50µl reaction using the 2x Acuzyme kit (Bioline) using the following conditions (94°C, 60°C, and 68°C for 30s each over 40 cycles). The PCR product was purified using GeneJET PCR purification kit (Thermoscientific) and run on a 1% agarose gel to confirm the presence of a single band. A series of ten fold dilutions was generated using a starting concentration of 2 picograms/ml.

Statistical Analyses

All analyses were conducted in JMP version 10.0 (SAS). Pupal gene expression was analyzed using only RPKM (reads per kilobase exon per million reads). A two-way ANCOVA was employed with day and gene as main effects and log expression as the dependent variable. The *glutamate receptor* was identified as a suitable internal control based on the transcriptome analysis, which revealed consistent levels of expression during all stages of pupal wing development. For the temperature and heparin experiment, gene expression was analyzed using a two-way ANCOVA with treatment and eyespot as main effects, treatment by eyespot as an interaction term and *ash1* as a

covariate. *Ash1* was selected as an internal control as it did not vary in expression with treatment or eyespot. All data were square root transformed and the residuals checked for normality. Homoskedasticity was checked visually by plotting residuals and also by conducting a Levene test for equal variances.

Results

Expression of Polycomb Repressive Complex Genes During Wing Development

The transcriptome analysis revealed that all major genes of PRC 1 and PRC2 are expressed during development of larval wing discs and the pupal wing. Genes from both complexes exhibited dynamic changes during wing development with higher levels of expression during the late larval and early pupal stages with a dramatic decline during late pupation (Figure 4.3). Expression of PRC1 genes varied significantly among pupal stages with the highest levels observed at day 2 followed by day 5 then day 8 ($F_{(2, 11)} = 108.3$, $p < 0.0001$). There was a significant gene x day interaction with *Psc* showing an increase in expression from day 5 to day 8 ($F_{(10, 11)} = 5.3$, $p < 0.001$). Overall, *Ring* was the most highly expressed gene and *Pc* was expressed at the lowest levels.

Pupal expression levels varied among the PRC2 genes with *Suz12* exhibiting significantly higher expression than *esc* ($F_{(2, 8)} = 7.5$, $p = 0.015$). Although no gene by day interaction was observed ($F_{(4, 8)} = 2.5$, $p = 0.13$). All genes exhibited a significant peak in expression at 2 days post-pupation ($F_{(2, 8)} = 335.8$, $p < 0.0001$) and no difference in expression was observed between days 5 and 8 ($F_{(1, 5)} = 3.7$, $p = 0.13$). We have previously confirmed that patterns of gene expression are highly correlated ($R^2 = 0.75$) between RNA-seq and qPCR data (Chapter 2).

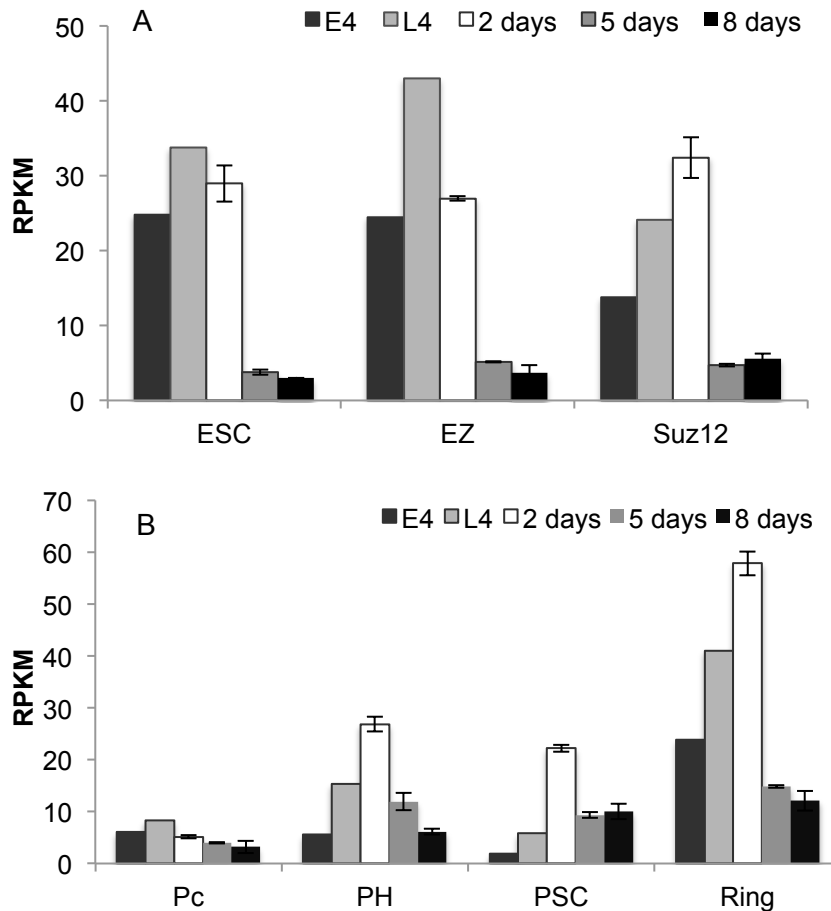


Figure 4.3 Expression of the polycomb repressive complex during larval and pupal wing development in *V. cardui*. RPKM represents reads per kilobase of exon per million reads mapped. Panel A shows expression levels for members of polycomb repressive complex 2. Panel B shows expression levels for members of the polycomb repressive complex 2. E4 and L4 represent early and late 4th instar, 2, 5 and 8 days represent number of days following pupation. Error bars are $1 \pm SE$ from the mean. No error bars are present for the larval stages as these represent one biological replicate of 5 pooled samples. Pupal stages represent the mean RPKM of 2 biological replicates of 3-5 pooled samples.

Comparison of Pupal Development Time and Pigmentation for Eyespot Dissections

We found that butterflies exposed to 37°C for 48 hours within 12 hours of pupation developed faster than control butterflies, eclosing at 7 days post-pupation compared to 8 days for controls (Table 4.1). Heparin injected butterflies eclosed at the same time as control butterflies. Pigmentation appeared 22 hours earlier in pupae exposed to temperature shock compared to the control and heparin pupae. Due to the advanced

development in temperature treated butterflies, eyespots were dissected one day earlier than control and heparin groups to ensure that all pupae were at a similar developmental stage. No significant difference was observed in pupal age at dissection time between the heparin and control groups (Table 4.1).

At 6 days post-pupation, butterfly eyespots are not fully formed; the melanin border that forms around eyespot 1 and 4 is not completely developed however all eyespots exhibit some degree of melanin and blue pigmentation. Temperature effects on eyespot pigmentation are difficult to distinguish from the controls at this developmental stage. In contrast, heparin effects are visible, showing dramatic modifications in eyespot formation and overall pigment patterns on the hind wing (Appendix Figure 4.1). Heparin injections have variable effects on butterfly wing patterns, although eyespots 2 and 3 exhibit dramatic reductions in black and blue pigmentation. Eyespots were greatly reduced in size or almost eliminated in many heparin treated butterflies. The melanin border, however, appeared less susceptible to the modifying effects of heparin compared to the other pigment elements.

Differential Expression of Patterning and Pigment Genes Across Eyespots.

A total of 16 genes were tested for differences in gene expression in response to the treatments and across the different eyespots. *Actin* and *glutamate receptor* were tested as potential internal controls, but showed significant treatment effects (Table 4.3).

Expression of *ash1* did not vary with treatments or across eyespots and was selected as a covariate for analysis of all other genes (Table 4.3).

Table 4.3 Results from two-way ANCOVA of gene expression. Main effects include treatment (control, temperature, heparin), eyespot number (1, 2, 3) and the interaction between treatment and eyespot. *Ash1* did not vary with treatment, eyespot or treatment \times eyespot interaction and was used as an internal control for the analysis of all genes. Significant results ($p \leq 0.05$) are indicated with an asterisk.

Gene	Treatment		Eyespot			Interaction		
	F	p	F	p	df	F	p	df
<i>Ash1</i>	0.03	0.97	1.86	0.16	2, 67	0.51	0.72	4, 67
<i>Actin</i>	3.65	0.03*	2.9	0.06	2, 65	3.2	0.02*	4, 65
<i>Glutamate</i>	11.96	<0.0001*	0.62	0.54	2, 66	0.37	0.83	4, 66
<i>Ddc</i>	2.32	0.10	2.71	0.07	2, 66	0.96	0.43	4, 66
<i>Ebony</i>	2.93	0.06	1.37	0.26	2, 66	0.18	0.95	4, 66
<i>Tan</i>	8.64	<0.001*	4.27	0.02*	2, 66	0.60	0.66	4, 66
<i>Pale</i>	2.77	0.07	3.34	0.04*	2, 65	1.13	0.35	4, 66
<i>Ubx</i>	0.98	0.34	3.07	0.05*	2, 66	1.37	0.27	4, 66
<i>Dll</i>	3.08	0.05*	3.20	<0.05*	2, 66	0.38	0.82	4, 66
<i>En</i>	4.9	<0.01*	5.20	<0.001*	2, 66	1.63	0.18	4, 66
<i>Sal</i>	4.48	0.02*	2.68	0.07	2, 66	0.33	0.86	4, 66
<i>Ez</i>	0.97	0.38	0.86	0.43	2, 65	1.08	0.37	2, 65
<i>Pc</i>	0.39	0.68	0.48	0.62	2, 66	0.97	0.43	4, 66

Across all eyespots and treatment groups, *sal* was the most highly expressed gene followed by *en*, *ubx* and *dll* (Appendix Figure 4.24). *Wingless* was also tested but expression was barely detectable at 6 days post-pupation. *Engrailed*, *ubx* and *dll* varied in expression across eyespots with eyespot 1 exhibiting higher expression than eyespots 2 and 3 (Appendix Figure 4.4). Overall, the most highly expressed pigment gene in eyespots was *pale* followed by *ddc*, *tan* and *ebony* and expression of *pale* and *tan* was significantly higher in eyespot 1 compared to eyespots 2 and 3 (Appendix Figure 4.4).

Effects of Temperature and Heparin on Pigment, Patterning and Polycomb Genes.

We did not observe any treatment effects or treatment by eyespot interaction on expression of *pale* or *ddc* (Table 4.3). Temperature shock significantly increased expression of *tan*. Although the interaction term was not significant, the effects were most pronounced in eyespots 2 and 3 (Figure 4.5, Table 4.3). Heparin had a marginally significant effect on expression of *ebony* ($p = 0.06$). Both temperature shock and heparin significantly increased expression of *en* and *sal* (Table 4.3). Although a treatment by eyespot interaction was not observed, the treatments appear to have more strongly affected expression of genes in eyespot 2 (Figure 4.5). Heparin significantly increased expression of *dll* expression particularly in eyespot 1 (Figure 4.5). There were no treatment effects on *Ez* or *Pc* at this developmental stage (6 days post-pupation) and no variation in expression across eyespots (Table 4.3). However, expression of *Pc* was measured at significantly higher levels than *Ez* across all eyespots ($F_{(1, 45)} = 75.58$, $p < 0.0001$).

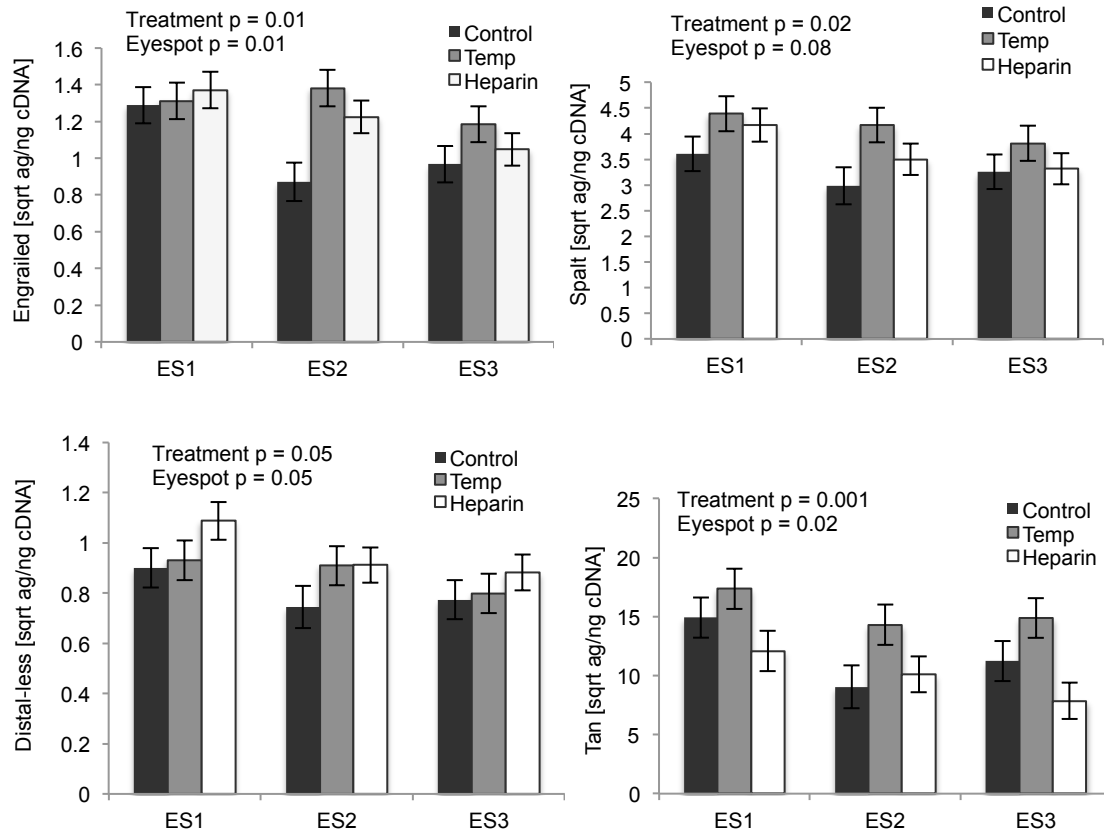


Figure 4.5 Expression of genes across eyespots for each treatment group for *V. cardui*. Data represent the least square means for square root transformed data. Errors bars are $1 \pm$ SE of the mean. Expression values of each gene are in attagrams (ag) of mRNA per nanogram (ng) of total RNA. Sample sizes: Control n=8, Temp n=8, Heparin n = 10.

Discussion

Phenotypic plasticity in butterfly wing patterns is a well-documented phenomenon, yet few studies have examined how environmental perturbations alter gene expression. Another unexamined issue is the possibility that epigenetic regulation of patterning and pigmentation genes is the source of this plasticity. We have previously shown that *Vanessa cardui* exhibits phenotypic plasticity in response to temperature shock and heparin injection with significant effects on size and pigmentation across three hindwing eyespots. In this study, we quantified expression of pigment, patterning and polycomb

genes in these hindwing eyespots and identified the most highly expressed genes and those varying in expression across eyespots. We also identified which genes are sensitive to environmental perturbation and associated with phenotypic plasticity. Our results suggest that although genes involved in epigenetic regulation are expressed in *V. cardui* wings, temperature and heparin treatments had no effect on their expression in modified eyespots, however expression of some patterning and pigment genes was significantly affected.

Expression Patterns of Pigment and Patterning Genes Across Hindwing Eyespots

Studies on butterfly eyespots have focused primarily on examining spatial expression of patterning genes in larval wing discs and early pupal stages (Carroll et al., 1994; Brunetti et al., 2001; Reed et al., 2007; Oliver et al., 2012; Oliver et al., 2013). Brunetti et al. (2001) have shown that in *V. cardui*, *sal*, *en* and *dll* are expressed in spatial patterns in eyespot 4 that correspond with yellow and black pigment rings. We also observed expression of these genes in the other hindwing eyespots at 6 days post-pupation along with the Hox gene *ubx*. Expression of Hox genes has also been observed in the eyespots of other butterfly species. *Ultrabithorax* has been identified in *Precis coenia*, and both *ubx* and *antennapedia* have been observed in *B. anynana* (Weatherbee et al., 1999; Saenko, et al. 2011; Tong, et al., 2014). These studies suggest that in addition to their role in controlling segmental identity along the anterior-posterior axis (Gellon & McGinnis, 1998), Hox genes may have evolved a functional role in eyespot development.

Across all eyespots, in control and treatment groups *sal* was the most highly expressed gene and also the only patterning gene that did not vary in expression across the three eyespots. In contrast, *en*, *dll* and *ubx* exhibited significantly higher expression in

eyespot 1 relative to eyespots 2 and 3, which is significantly smaller in size. These results indicate that expression levels of these genes vary across the wing between large and small eyespots. A number of studies have implicated a role for these genes in regulating eyespot size. Recent transgenic experiments in *B. anynana*, have shown that both RNAi reduction of *distal-less* and overexpression of *ubx* resulted in butterflies developing smaller eyespots (Monteiro et al., 2013; Tong et al., 2014). A similar trend was also observed in a homeotic *ubx* mutant of *P. coenia* where increased expression of *dll* led to the development of larger eyespots (Weatherbee et al., 1999). Thus, *ubx* and *dll* appear to have conserved roles in regulating eyespot size in different butterfly species, and may also perform a similar function in *V. cardui*.

Expression of two genes from the melanin synthesis pathway also varied across eyespots in control and treatment groups. *Pale* and *tan* were more highly expressed in the larger eyespot, which has significantly more black pigment than the two central eyespots. Interestingly, *pale* and *ddc* were more highly expressed than *tan* and *ebony* in all eyespots, which is the same order of expression levels, observed for the entire wing (*pale*>*ddc*>*tan*>*ebony*) (Chapter 2). These results are in contrast to expression levels observed for *Heliconius* butterflies, where *tan* and *ebony* exhibit higher expression levels than *ddc* (Hines et al., 2012). Thus, expression levels of these biosynthetic enzymes appear to have diverged in different butterfly species.

Heparin Increases Expression of Patterning Genes but Has No Effect on Genes Promoting Melanization.

Heparin is an extracellular proteoglycan that is known to modify the activities of morphogens such as *wingless*, *decapentaplegic* and *hedgehog* (Yan & Lin, 2009). It is

widely proposed that eyespot development is regulated by diffusion of a putative morphogen that interacts with transcription factors such as *sal*, *dll* and *en* to determine the differentiation of concentric colored rings (Brunetti et al., 2001; Mcmillan et al., 2002). Variation in expression of this putative morphogen could affect the concentration gradient required to set up the concentric rings and this may alter expression of these downstream transcription factors. Heparin severely disturbs eyespot development in *V. cardui* and also *Junonia coenia* (Serfas & Carroll, 2005), resulting in an overall loss or bleaching of the eyespots; however, it is not clear whether the effects of heparin on morphogen diffusion results in an increase or decrease in expression of patterning genes. We could not examine effects of heparin on *wingless* because this morphogen is barely detectable at 6 days post pupation. However, we found that heparin injection within 12 hours of pupation results in increased expression of *en* and *dll* during the developmental period when pigments are deposited in the wing. Serfas and Carroll (2005) found that injection of heparin significantly reduced the spatial expression of *dll* in the hind wing eyespot of *Junonia coenia* at 2 days post-pupation, suggesting that heparin decreases transcription of key eyespot genes. Because we did not measure expression during early pupation, we do not know whether heparin influences gene expression dynamics during development. However, the results from our study and those of Serfas and Carroll (2005) suggest that heparin may have distinct effects on eyespot patterning genes at different stages of pupal development.

Although heparin significantly reduced melanin pigmentation in all eyespots, the treatment had no effect on decreasing expression of enzymes required for melanization (*pale*, *ddc* and *tan*). We expected that *ebony* might exhibit increased expression because

it inhibits formation of melanin pigmentation (Wittkopp et al., 2002). Though not significant, we did observe a trend for higher expression of *ebony* in all eyespots, which may have contributed to the suppression of melanin pigment in these eyespots. These results suggest that the dramatic effects of heparin on pigmentation are not due to altered expression levels of enzymes that promote melaninization. Although heparin had no significant effect on transcript levels of these melanin genes, the treatment may have perturbed their spatial expression by modifying expression domains of patterning genes prior to melanization. Alternatively, heparin may have affected downstream cascades by binding to pigment precursors, or altering levels of precursors that are transported into scale cells.

Our results do show that during the period of pigmentation, heparin elevated expression of *dll* and *en*, which was also associated with increased expression of *ebony*. Whether patterning genes directly or indirectly influence expression of pigment genes remains unknown, although in *Drosophila* the Hox gene, *Abd-B* has been shown to directly regulate expression of melanin pigment gene *yellow* (Jeong et al., 2006). Serfas and Carroll (2005) propose that heparin may alter the secretion of a cold shock hormone that may be present in the hemolymph, influencing activity of an intracellular regulator involved in pigmentation. Butterflies exposed to cold shock exhibit similar phenotypes to those injected with heparin (Nijhout, 1984). Thus, the effect of heparin on pigmentation may not be due to direct effects of patterning genes on pigment genes but is potentially mediated by a cold shock sensitive regulator. Regardless of the mechanism, these results suggest that altered expression of *dll* and *en* and potentially *ebony* are involved in the dramatic phenotypic changes in eyespot pigmentation in response to heparin.

Engrailed, Sal and Tan Are Thermosensitive Genes.

Similar to the heparin treatment, temperature shock had no detectable effect on the early acting enzymes (i.e., *pale* and *ddc*) in the melanin pathway. There was also no effect on *ebony* expression. However, *tan* expression was significantly increased in the two central eyespots. This observation was surprising because the two central eyespots show a reduction of melanin pigmentation and an increase in blue scales. The structural blue color is produced not by pigments but changes in scale microstructure that interfere with reflectance of incident light (Ghiradella, 1991). Structural colors however are backed by melanin pigmentation which intensifies the color by absorbing excess light (Ghiradella, 1991), which may explain upregulation of *tan*. Thus, melanin pigmentation may have increased in these eyespots but was simply masked by changes in scale structure leaving only the blue color visible. In addition to melanin pigmentation *tan* may also regulate scale microstructure or morphology. Many pigment genes including *tan* are pleiotropic and are known perform multiple functions during development (Wittkopp & Beldade, 2009).

In addition to changes in *tan* expression, we also observed increased expression of *sal* and *en*, particularly in eyespot 2, which exhibits the largest increase in blue coloration. It has been proposed that patterning genes may also act pleiotropically to regulate two separate developmental pathways involving pigmentation and scale ultrastructure (Janssen et al. 2001). If patterning genes influence the development of scale microstructure, then temperature-induced upregulation of *sal* and *en* may explain the development of blue scales. The molecular basis of structural color determination is

unknown; therefore we can only speculate how a change from a pigmented to a structural color may have occurred in response to temperature shock.

Finally, we cannot rule out the possibility that temperature shock reduced expression of the *yellow* gene, which is also involved in melanin production (Wittkopp et al., 2002; Ferguson, et al. 2011). Changes in expression of this gene could have played an important role in the final color fate of these scales. In *V. cardui*, as in other butterflies, there are at least 10 paralogs of the *yellow* gene (Ferguson, et al. 2011); however, their individual function in melanin pigmentation remains unclear. Further work is required to determine the function of these different paralogs and examine whether any members of the *yellow* gene family are also thermosensitive.

Polycomb Repressive Complex Genes Are Expressed During Wing Development

To explore the potential role of epigenetic regulation of plasticity in wing color patterns, we examined whether genes within the PRC are expressed during wing color pattern development. If the complex were involved in regulating expression of eyespot genes it would likely have an effect during the early stages of eyespot development (i.e., during the late larval and first 2 days of pupation). It is not possible to dissect butterfly eyespots at these stages; instead, we examined whether the PRC1 and PRC2 were expressed in whole wing tissue during these larval and pupal stages. We show that all major genes that comprise PRC1 and PRC2 are expressed during wing development in a dynamic and stage-specific pattern. Expression peaks during early pupation when wing patterns are being established and declines significantly during late pupation in a similar manner to patterning genes (Chapter 2). Genes within the PRC2 exhibited relatively similar patterns of expression although overall expression of *Suz12* was significantly higher than *Esc*.

Within the PRC1, transcript levels among the genes were more variable; *Pc* was expressed at the lowest levels and *Ring* were expressed at the highest levels throughout pupation.

The PRC has been shown to vary in expression and composition across different tissues, cell types and developmental stages (Gunster et al., 2001). Our results suggest that although expression levels of the PRC vary during wing development the composition of the major components does not change. Surprisingly, we found that within the PRC2, *Esc* was transcribed at similar levels to *Suz12* and *Ez*. Expression of *Esc* is thought to occur primarily during oogenesis and early embryogenesis compared to *Ez* and *Suz12* which are expressed throughout development (Tie, et al., 1998; Ng, et al., 2000). Our observations of *Esc* expression in *V. cardui* during pupation suggest that this gene may have evolved additional functions in wing development.

To examine whether changes in expression of patterning and pigment genes were also associated with altered expression of the PRC, we focused on two genes from each complex, *Pc* (PRC1) and *Ez* (PRC2). Both genes were expressed at consistent levels across eyespots with no effects of either treatment on their expression. Although we did not observe any treatment effects on expression for either gene at this developmental stage, we did find that expression of *Pc* was significantly higher than *Ez* in all three eyespots. Whether these differences in expression have functional consequences for regulating genes involved in pattern development would be an interesting avenue for future studies, as would investigation of expression of the PRC at different stages of eyespot development.

Conclusion

We identified a suite of genes that were sensitive to environmental perturbation during eyespot development. *Engrailed* was shown to be a highly sensitive gene as it was the only gene that was affected by both treatments. Interestingly, *en* expression is known to exhibit increased variability during later stages of eyespot pattern formation in larval wing discs compared to *notch* and *sal* (Reed et al., 2007). This variability in expression may contribute to its sensitivity in responding to environmental perturbation suggesting it may be involved in regulating eyespot plasticity in *V. cardui*. Genes involved in the polycomb complex were not sensitive to developmental perturbations; however, we only sampled at one time point during late stages of eyespot development. Thus, we cannot rule out the possibility that polycomb, patterning or pigment genes show significant responses at earlier or later stages during pupation. Because expression of these genes varies dramatically during wing development, a fine-scale time series analysis is essential to obtain a robust picture of gene expression dynamics. Our results do suggest that perturbations during early pupation can alter gene expression much later in development during pigmentation and these changes may have consequences for scale color fate.

CHAPTER V

MOLECULAR EVOLUTION OF AN EPIGENETIC SILENCER: POLYCOMB REPRESSIVE COMPLEX 2

Abstract

The polycomb repressive complex comprises a group of interacting proteins that play a crucial role in maintaining silencing of genes involved in organismal development. The evolution of this complex has received some attention in vertebrates and plants, which have experienced multiple duplication events during their evolution. Little is known about the evolutionary history of these proteins and their domains in invertebrate animals, other than a lack of gene duplication in these taxa. Here, we conducted a large-scale phylogenetic analysis of three core members of polycomb repressive complex 2 (Enhancer of zeste, Suz12 and Extra sex combs) to examine the degree of conservation and divergence across invertebrate animals. We found that the gene trees do not reconstruct the known phylogeny of these animals, nor do they share similar evolutionary histories. Nematodes display significant sequence divergence in all functional domains, which does not reconstruct a monophyletic Animalia. We also identify residues and domains that are highly conserved versus those displaying significant sequence divergence during the evolution of early metazoans. These findings demonstrate that polycomb genes are generally highly conserved across lineages that exhibit dramatically different modes of development.

Introduction

Animal development is coordinated by a complex interplay of genetic and epigenetic regulators that control spatial and temporal patterns of gene expression.

During early development, maternal factors and segmentation genes initiate the developmental program that establishes which genes are switched on or off (Schroeder et al., 2004; Levine, 2008). These initiating factors are expressed along a concentration gradient and determine the appropriate expression domains of homeotic (Hox) genes that define regional identity along the anterior-posterior axis (Gellon & McGinnis, 1998; Carroll et al. 2001). While maternal factors and segmentation genes set up the initial cascade of expression patterns, their transient nature means additional mechanisms are required to maintain cell identity throughout the lifetime of the organism (Ringrose & Paro, 2007; Schwartz & Pirrotta, 2007).

The cellular memory system that functions in heritable gene silencing is propagated by a group of proteins called the polycomb repressive complex (PRC) (Margueron & Reinberg, 2011; Pirrotta, 2011). The PRC plays a critical role in determining cell identity via negative regulation of Hox clusters and many other developmental genes (Breiling, et al. 2007; Margueron & Reinberg, 2011). Genome wide mapping in human embryonic fibroblasts has identified thousands of polycomb gene targets in addition to Hox genes including transcription factors, morphogens, receptors and signaling proteins that are important in developmental processes and cell fate decisions (Bracken et al. 2006, O'Meara & Simon, 2012). In each cell only the specific genes required for a particular development pathway are active while alternative genetic programs are silenced by the PRC (Schwartz & Pirrotta, 2007; Prezioso & Orlando, 2011).

The two major PRC complexes that have been most intensively studied are PRC1 and PRC2. PRC1 is composed of 4 major core components, Polycomb (Pc), Sex combs extra (dRing1/Scf), Polyhomeotic (Ph), and Posterior sex combs (Psc) proteins (Kerppola, 2009). PRC was originally named based on mutation screens that identified *polycomb* gene misexpression as a source of homeotic transformations in *Drosophila* (Lewis, 1978; Jürgens, 1985; Struhl & Akam, 1985). Mutations in the *polycomb* (Pc) gene in *Drosophila* transformed posterior legs into anterior legs that contain comb like bristles, and mutations in other polycomb group genes resulted in similar phenotypes due to altered expression of Hox genes. The *polycomb* gene contains a chromodomain that binds tri-methylated lysines on histone 3, a mark that is propagated by PRC2. The canonical model proposes that PRC1 is recruited to H3K27me3 following methylation activity of PRC2 and that interaction between both complexes is necessary for transcriptional repression (Schwartz & Pirrotta, 2007; Kerppola, 2009; Klose et al. 2013).

In invertebrates, PRC2 is made up of three major core subunits. The first protein, enhancer of zester (EZ) is composed of four major domains, SANT1, SANT2, CXC and the catalytic SET domain. Extra sex combs (ESC) is composed primarily of WD40 repeats and suppressor of zeste (Suz12) is characterized by two domains, the zinc finger and VEF box (Cao and Zhang 2004, Schuettenbruger et al 2007, Margueron & Reinberg, 2011) (Figure 5.1). Despite the important role of PRC silencing in organism development and disease, the link between H3K27me3 and subsequent polycomb mediated chromatin compaction are poorly understood. This link may involve interference of transcriptional machinery or other chromatin remodelers such as SWI/SNF (Dellino et al., 2004; Wilson et al., 2010; Simon & Kingston, 2013).

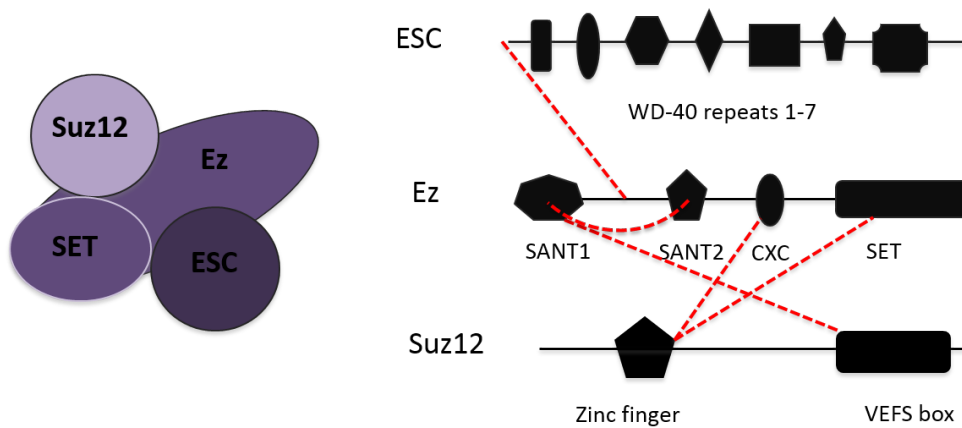


Figure 5.1 Major components of the PRC2 in invertebrates, showing the interaction between Ez and Suz12 and Ez and ESC, adapted from O’Meara & Simon, (2012). Schematic (right) illustrates the major domains present in each polycomb gene and some of the interactions that have been identified by Ciferri et al. (2012). The N terminal region of ESC interacts with a region downstream from the SANT1 domain of EZ. The SANT1 domain also interacts with the VEFS domain, while the SET domain interacts with the zinc finger of Suz12 and the CXC domain binds the zinc finger of Suz12. SANT domains are known to couple histone tail-binding to enzymatic activity (Shaver et al. 2010). These interactions likely facilitate coupling of histone 3 to the SET domain, and mediate its methyltransferase activity through the interactions with VEFS domain and active chromatin marks (Ciferri et al. 2012).

Function of the PRC2

Initially, PRC2 was thought to function in the permanent repression of genes based on early developmental decisions. However, it has become apparent that the activity of the PRC2 is highly dynamic and dependent on particular developmental stages, cell cycle and interactions with activator complexes (Pirrotta, 2011). In addition to body patterning, PRC2 also plays an important role in a variety of other biological processes including differentiation, cell cycle regulation, apoptosis, cell identity maintenance, proliferation, pluripotency and stem cell self-renewal (Whitcomb et al. 2007, Pirrotta, 2011; Margueron & Reinberg, 2011).

Recent work has begun to unravel the molecular architecture of the polycomb subunits and their interacting domains (Ciferri et al., 2012). Despite this work, details of

their interactions are still poorly understood due to a lack of structural information regarding the active site of EZ (O'Meara & Simon, 2012). The catalytic SET domain of EZ has methyltransferase activity, which is critically dependent on its interaction with both the WD40 domains of ESC and the VEFS domain of SUZ12 (Cao & Zhang, 2004; Han et al., 2007; O'Meara & Simon, 2012; Ciferri et al., 2012). The assembly of PRC2 onto chromatin is initiated by ESC, by its recognition of the H3K27 tri-methylation mark through its WD40 domain (Hansen & Helin, 2009). This interaction promotes the spread of H3K27 tri-methylation by the complex. The conserved WD domain folds into a seven bladed β -propeller that functions as a scaffold for protein interactions (Han et al., 2007).

The top portion of the β -propeller features an aromatic cage that can specifically bind repressive chromatin marks such as H3K27me₃ and H3K9me₃ (O'Meara & Simon, 2012), while the bottom of the propeller (C terminus) binds an N-terminal motif on EZ (Jones et al., 1998; Han et al., 2007; Ciferri et al., 2012). Thus, ESC presents the H3 substrate to the catalytic site of EZ facilitating the propagation of H3K27me₃ (Han et al. 2007). Methylation then occurs when AdoMet (the methyl donor) and lysine₂₇ are bound to the catalytic binding site of the SET domain. Following de-protonation of the lysine residue a methyl group is transferred from AdoMet to the lysine side chain (Hamamoto et al. 2015). Spreading of H3K27me₃ will only proceed if neighboring nucleosomes are in contact with Suz12, which appears to be able to 'sense' the local chromatin landscape (Ciferri et al., 2012; Herz et al., 2013). The critical role of Suz12 has been demonstrated in knockdown experiments, resulting in cell growth defects and genome-wide decreases of H3K27 tri-methylation in addition to upregulation of certain Hox genes (Cao & Zhang, 2004). In addition to stimulating the methyltransferase activity of PRC2, Suz12 can also

inhibit its activity via its interaction with H3K4me3 at the C-terminal of the VEFS domain (Schmitges et al. 2011). In this way, the PRC2 can ‘sense’ marks associated with active transcription and adjust its own methyltransferase activity. Similarly, repressed chromatin can stimulate PRC2 methyltransferase activity creating a positive feedback loop (Margueron & Reinberg, 2011; Schmitges et al., 2011; O’Meara & Simon, 2012). In addition to the VEFS domain, Suz12 also possess a C2H2 zinc finger, the function of which is currently unknown (Schmitges et al. 2011). Thus, PRC subunits function as molecular cogs that facilitate fine-tuning of the PRC2 methyltransferase activity through their ability to detect the local chromatin landscape (O’Meara and Simon 2012).

Origin and Evolution of the PRC2

The PRC2 is an ancient epigenetic regulator that arose very early in the evolution of eukaryotes. All three major PRC2 subunits are present in unicellular algae (Shaver et al. 2010); in contrast, PRC1 is less conserved, indicating that it is more recently derived than PRC2 (Shaver et al., 2010; Derkacheva & Hennig, 2013). It has been proposed that PRC2 may have originally evolved as a defense mechanism against genomic parasites, later evolving specialized functions for developmental regulation (Shaver et al. 2010, Dekacheva and Hennig 2013). PRC2 appears to be largely conserved across algae, fungi, insects and mammals, while it has been lost independently in some lineages (e.g. yeast) (Shaver et al., 2010; Margueron & Reinberg, 2011). Its function has also diversified in some lineages; for example, the PRC2 is essential for embryo development in animals while in plants it is required for developmental transitions (embryo to seedling, vernalization, flowering) (Derkacheva and Hennig 2013).

Specialization in lineages may be driven in part by gene duplication and the

gain/loss of functional domains. Both PRC1 and PRC2 have experienced multiple duplication events in vertebrates (Whitcomb et al., 2007; Senthilkumar & Mishra, 2009) and plants (Hennig & Derkacheva, 2009). Furthermore, PRC1 and PRC2 paralogs have evolved specialized expression patterns and functional roles due to the evolution of novel domains (Gunster et al., 2001; Whitcomb et al., 2007). Vertebrate Hox genes also display duplication events; thus, Hox genes and PRC may have co-evolved, driving the evolution of specialization during vertebrate development (Whitcomb et al., 2007). In contrast, no subunit duplication within PRC has been observed in invertebrates with the exception of Psc in PRC1. These observations raise the question why polycomb gene expansion did not occur in invertebrate animals, which are also morphologically diverse. Diversification in invertebrates may have occurred within the polycomb genes themselves by changes in domain architecture or binding motifs, rather than through gene duplication.

Research Questions

Although the evolution of polycomb genes and their paralogs have been examined in vertebrates and plants, less is known about the evolution of polycomb genes and their domains in invertebrates. Here, we conducted a detailed examination of PRC2 evolution across invertebrates. Specifically, we examine 1) whether the evolutionary histories of EZ, Suz12 and ESC are consistent or have diverged from the known phylogeny of animals 2) whether the polycomb genes share a similar or divergent evolutionary history and 3) whether it is possible to identify regions of conservation/divergence in sequence features of each major domain. Understanding the evolutionary history of these proteins will provide insights into PRC2 function in animal development and the role of these domains and residues in the evolution of early metazoans.

Methods

Sequences were retrieved as mRNA from both NCBI non-redundant database and UniprotKB and were recorded as complete or partial based on its database annotation. A variety of approaches were used to obtain sequences for the majority of invertebrate taxa for which data were available including BlastN, BlastP, tBlastN and position specific iterative PSI-Blast using default settings. Sequences were retrieved for the three major subunits of the PRC2: EZ, Suz12 and ESC. Sequences from *Drosophila melanogaster* were used as a seed for retrieval of invertebrate sequences for initial searches. Multiple taxa from these initial results were then used as seeds for subsequent searches to recover a broader range of taxa, particularly taxa at the base of the animal tree. Sequences for *Mnemiopsis leidyi* were obtained from the draft genome at the *Mnemiopsis* Genome Project portal (research.nhgri.nih.gov/Mnemiopsis, accessed January 2014) using the Cnidarian, *Nematostella vectensis*, as a seed. Sequences were verified using SMART (smart.embl-heidelberg.de) (Letunic et al., 2014) for identification of specific domains (e.g. VEFS box). Nematode sequences were obtained from WormBase (www.wormbase.org/), Uniprot and NCBI. For outgroups, sequences from the Fungi were collected from both NCBI and UniprotKB and for choanoflagellates which were obtained from the Broad Institute (www.broadinstitute.org). Accession numbers for all taxa included in the study are listed in Appendix Table 5.1A (EZ), 5.1B (Suz12) and 5.1C (ESC).

All sequences were subject to quality control by checking several criteria.

Sequences were translated in Expasy (web.expasy.org/translate/) using the compact setting to check for open reading frames and to verify the absence of internal stop codons.

Translated sequences were immediately examined in SMART using Pfam and outlier homologs (schnipsel database) to identify conserved domains. Domain position and amino acid length was recorded. The following domain sequences were also collected: SANT 1 and 2 (EZ), CXC (EZ), the Set domain (EZ), Zinc Finger (Suz12) and VEFS box (Suz12), and all WD40 repeats (ESC). Sequences without any recovered domains were discarded. Nucleotide sequences for phylogenetic analysis of the entire protein-coding region were translated to protein sequences in Seaview (Gouy et al. 2010). Taxa that was only available as protein sequences were incorporated into this final dataset after checking for the presence of appropriate domains in SMART.

Protein sequences were aligned in Seaview using MUSCLE. Ambiguous regions were removed using Gblocks (relaxed settings) (Talavera & Castresana, 2007). Following removal of ambiguous regions, the final alignments were reduced to 261, 264 and 392 amino acids for Suz12, ESC and EZ respectively. Phylogenetic trees were constructed using MrBayes version 3.1.2 with a fixed Blosum62 model (Henikoff & Henikoff, 1992). Two simultaneous analyses of four Markov Chain Monte Carlo chains (one heated, three cold) were conducted; cold chains were sampled every 100 runs. The analysis was stopped when runs converged on results that had <0.01 split frequencies. Posterior probabilities were used to assess the strength of relationships found via these methods. All trees were edited in PAUP (Swofford, 1993). Fungi and choanoflagellates were designated as outgroups. Taxon sampling for each gene region is given in Appendix Table 5.1.

Mapping of Domains

The presence of domains in each gene (Ez, Suz12 and ESC) was recorded from SMART (Pfam and schnipsel). Pairwise comparisons were also conducted using NCBI BLAST using *Drosophila melanogaster* to obtain percent similarity scores for the entire protein and individual domains for all taxa. A combined approach using SMART, BLAST and multiple sequence alignments identified both conserved domains and divergent sequences. Domains that were highly conserved and considered homologous (e-value <0.0001) across all taxa were mapped onto the known phylogeny of animals with the taxon sampling from phylogenetic analyses described above. Jalview was used for multiple sequence alignment and annotation of binding sites and percent similarity.

Gene Tree Comparisons

A two-sided Kashino-Hasegawa test was also conducted in PAUP to examine if gene trees topologies were significantly different from each other and from the known animal phylogeny. The animal phylogeny was reconstructed using classification information obtained from the Tree of Life website (tolweb.org/tree/) (Appendix Figure 5.2). Prior to the pairwise comparisons each gene and species tree was pruned in MacClade (Maddison & Maddison, 2000) to ensure each dataset contained identical taxa.

Results and Discussion

Overall, the gene trees correctly reconstructed relationships within clades, particularly for insects. Relationships at deeper nodes were less well resolved despite strong branch support (Figures 5.3-5.5). The Suz12 gene tree was poorly resolved, with a large polytomy at the base of the main animal clade (Figure 5.5). This polytomy may be due to the a reduction in amino acid data used to construct the Suz12 gene tree; after removing

ambiguously aligned data the Suz12 alignment was reduced to 34% of the total gene length. Two thirds of Suz12 was highly variable; the remaining third was extremely conserved lacking phylogenetic signal to resolve the relationships of distantly related species. Similar to the EZ and ESC tree, Suz12 resolved within clade relationships between molluscs, plathelminthes, nematodes and insects with the exception of Lepidoptera (*Danaus plexippus* and *Bombyx mori*) and the pea aphid *Acyrthosiphon pisum*.

The evolutionary history of the three core subunits of PRC2 has diverged significantly not only from the known phylogeny of animals but also from each other ($p < 0.0001$ for all KH pairwise comparisons between gene trees). The lack of similarity between gene trees is surprising, given the close interactions between the proteins in the PRC2 (Jones et al., 1998; Cao & Zhang, 2004; Han et al., 2007; Ciferri et al., 2012). Although KH comparisons reveal that the different genes trees did not reconstruct the same relationships among taxa, it is still possible that coevolution has occurred at individual residues or functional domains. It may be that changes in binding sites for one protein may influence substitutions in the complimentary binding site in a different protein. Three-dimensional modeling would be one approach for further examination of fine-scale co-evolution between major protein domains within PRC2.

The observation that the individual gene trees are not consistent with the known animal phylogeny was less surprising. Gene trees often do not reconstruct species trees due to gene duplication, extinction or deep coalescences that occur if related species do not share paralogs (Nichols, 2001; Maddison, 2008). Although deep coalescence is often associated with closely related species it can also occur at deeper nodes in the phylogeny

leading to violation of monophyletic relationships (Degnan & Rosenberg, 2009). One of the main drivers for the mismatch of gene trees with the animal tree appears to be the extremely divergent protein sequences found in the nematodes, as discussed below.

Nematodes Do Not Reconstruct the Monophyletic Relationship of Animals

We assumed that polycomb genes would be most divergent in outgroups and early lineages (i.e. sponges, comb jellies and placozoans) compared to those animals with more complex body plans. As expected, polycomb genes in fungi and choanoflagellates varied in amino acid composition in multiple domains compared to metazoans, reflecting distant common ancestry and developmental functional differences. Contrary to our expectation, we found that polycomb genes were remarkably conserved from comb jellies to arthropods despite significant differences in morphology and development.

The most surprising result involved the phylogenetic position of nematodes; all three genes placed nematodes as sister to fungi, exclusive of the remainder of animals (Figures 5.3-5.5). These results did not appear to be due to long-branch attraction in this lineage, based on a comparison of overall branch length in each gene tree. Shaver et al. (2010) and Luo et al., (2009) also report similar findings despite different taxon sampling, alignment software and phylogenetic approaches.

To further explore the unusual placement of Nematoda, we investigated whether its position in gene trees was due to domain loss or sequence divergence. Our results show significant sequence divergence in nematodes compared to other animals. Amino acid changes can result from conservative or non-conservative substitutions with the latter potentially altering the physiochemical properties of the residue and ultimately protein function (Betts & Russell, 2007).

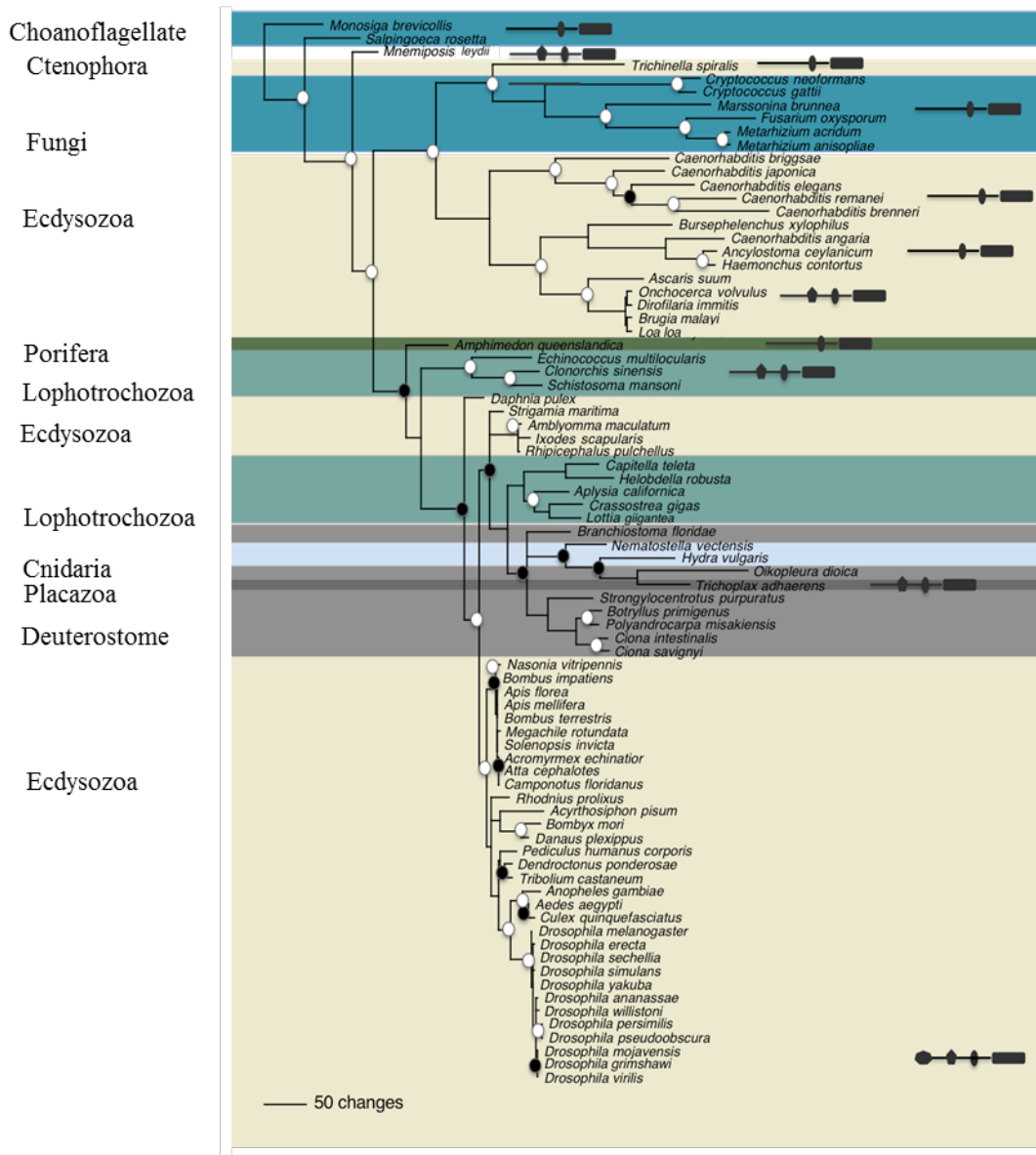


Figure 5.3 Phylogenetic tree of *Enhancer of zeste*. Open circles represent posterior probability = 1, closed circles represent posterior probability > 0.9. Taxa are color coded by major groups. Bottom right shows schematic of domains (SANT1, SANT2, CXC and SET). Unlabeled taxa possess all 4 domains. Labeled taxa show conserved domains whereas missing domains are highly divergent.

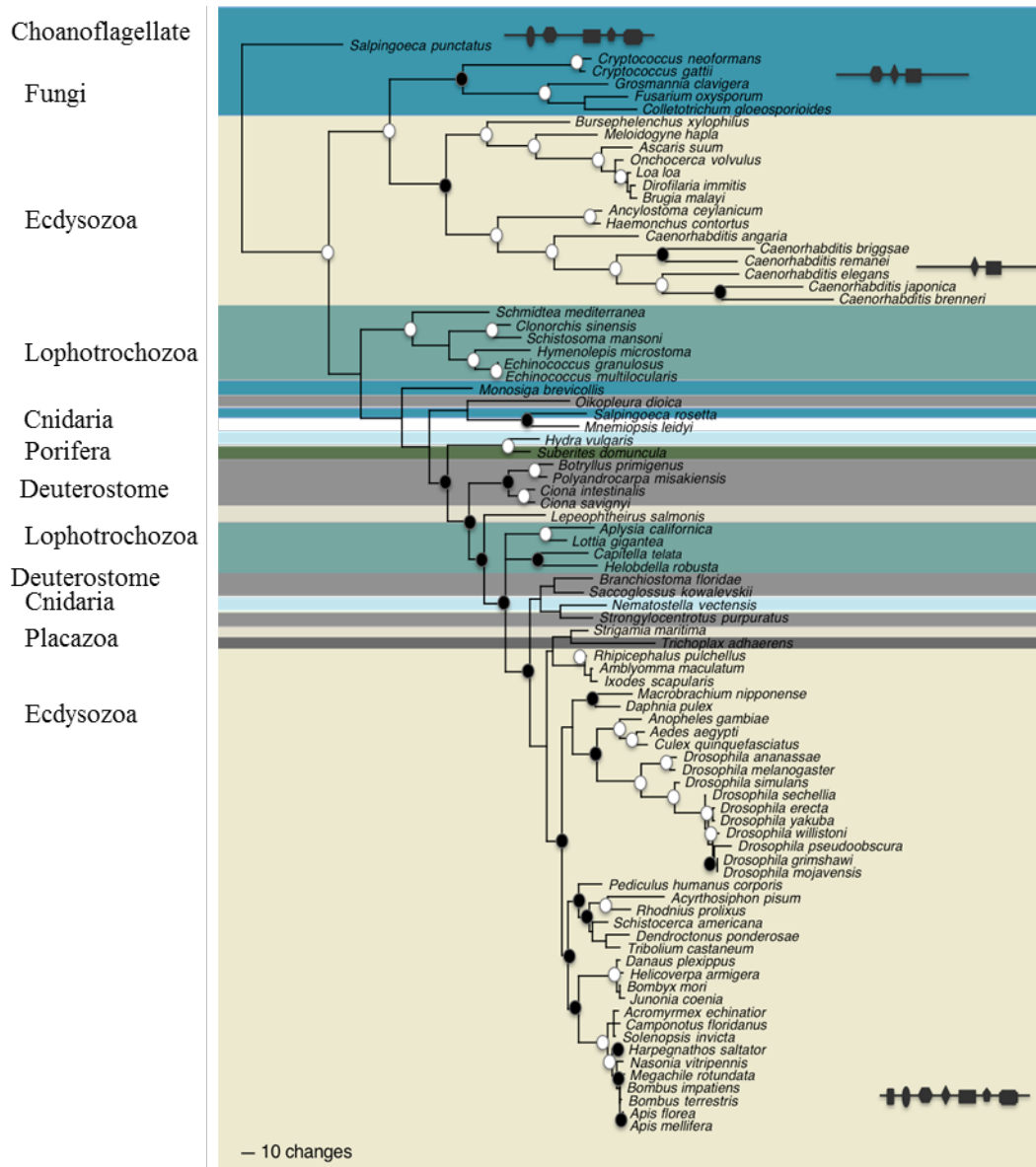


Figure 5.4 Phylogenetic tree of ESC. Open circles represent posterior probability = 1, closed circles represent posterior probability > 0.9. Taxa are color coded by major groups (see methods). Bottom right shows schematic of domains (Seven WD40 repeats). Unlabeled taxa possess all 7 repeats. Labeled taxa show the conserved repeats and missing repeats are divergent.

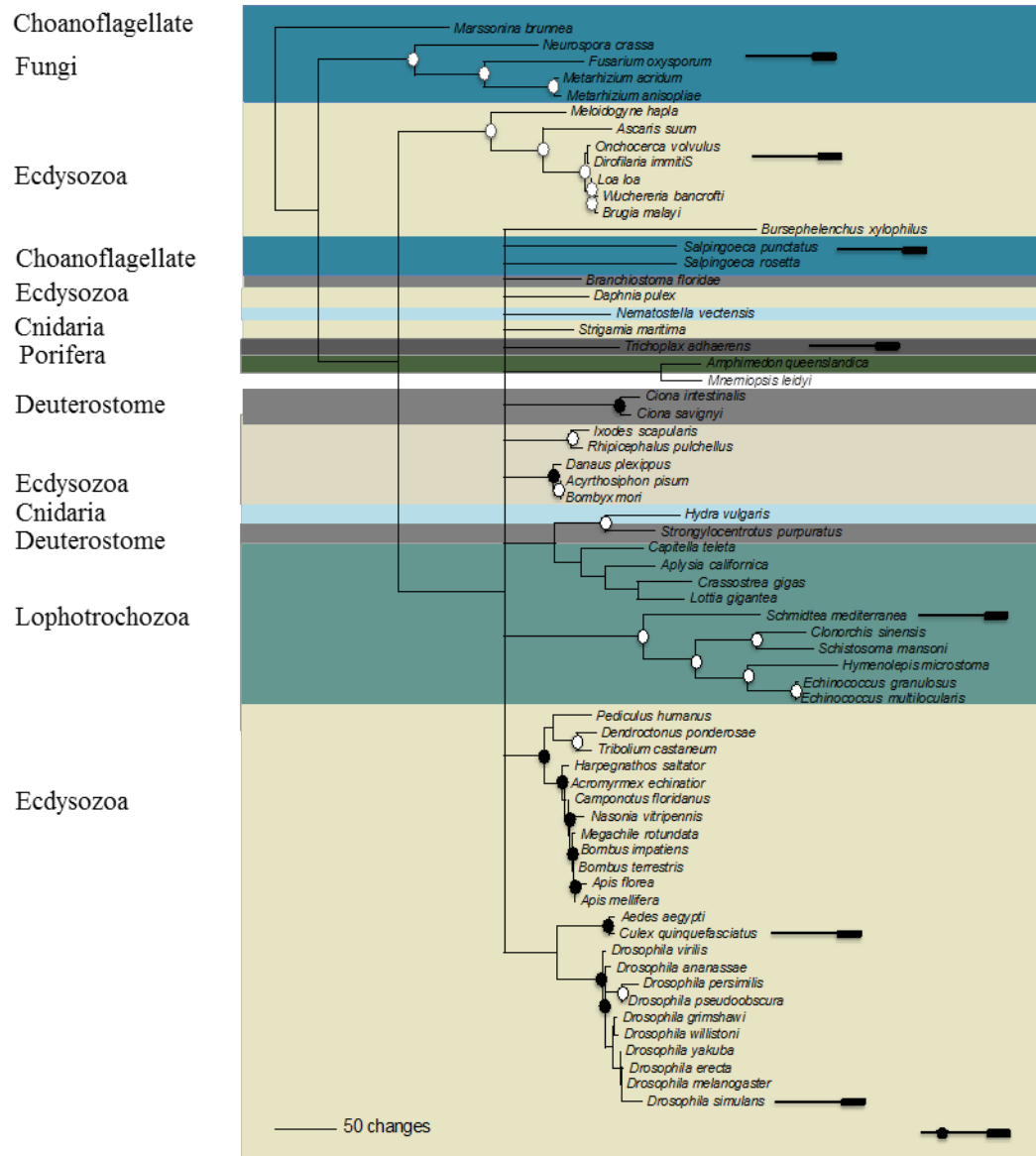


Figure 5.5 Phylogenetic tree of Suz12. Open circles represent posterior probability = 1, closed circles represent posterior probability > 0.9. Taxa are color coded by major groups (see methods). Bottom right shows schematic of domains (zinc finger and VEFS box). Unlabeled taxa possess both domains. Labeled taxa show the conserved VEFS box and a missing zinc finger that is highly divergent.

We compared multiple sequence alignments for each domain of EZ, Suz12 and ESC to examine whether nematodes displayed non-conservative substitutions at conserved sites. Overall, we found that many of the amino acid changes involved non-conservative substitutions, which we also observed in fungi and choanoflagellates. The substitutions that occurred in nematodes were generally not the same as those observed in fungi and choanoflagellates, indicating that these different lineages have not converged on substitutions with similar physiochemical properties.

Mutation studies in *Drosophila melanogaster* have identified residues with important functional roles. Remarkably, many of these studies have revealed that even a single non-conservative substitution can dramatically alter PRC2 assembly and enzymatic activity (Jones et al., 1998; Tie et al., 1998; Ketel et al., 2005; Han et al., 2007; Joshi et al., 2008; Rai et al., 2013). Some important functional motifs for histone methylation have been identified in the catalytic SET domain of EZ. These include FLF, a hydrophobic motif that contacts the lysine substrate, YCG a motif involved with the catalytic site, and GWG, a motif involved in methyl donor binding (Dillon et al. 2005; Joshi et al., 2008). Another highly conserved region is the pseudoknot, which brings together two highly conserved motifs (RFANHS and EELFFDY) to form an active site adjacent to the methyl donor (Ado/Met) binding region (Dillon et al., 2005). We found that these motifs are extremely well conserved even across distantly related taxa; however, nematodes and fungi have accumulated non-conservative mutations at these sites (Figure 5.6). The mollusk *Aplysia californica* also displayed a non-conservative mutation in the Ado/Met binding region; deletions at this site are found in echinoderms (*Strongylocentrotus purpuratus*) and in one species of nematode (*Trichinella spiralis*),

suggesting that these animals may have evolved different methyl donor binding sites. We also observed a similar pattern in the highly conserved CXC domain with nematodes and fungi displaying the most divergent sequences. Despite divergence in fungi and nematodes, cysteine residues, which help to stabilize PRC2, were largely conserved (Appendix Figure 5.7). The overall high conservation of these regions across phylogenetically distant taxa suggests that some functional constraints have been relaxed in certain taxa.

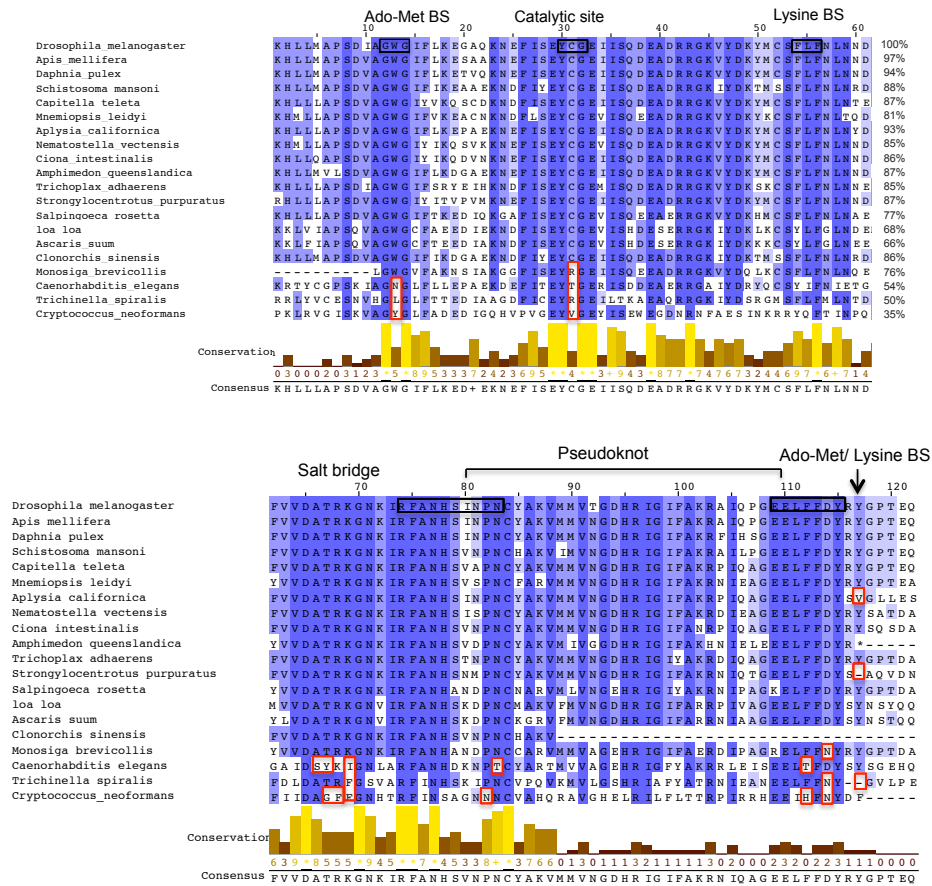


Figure 5.6. Multiple sequence alignment of the enhancer of zeste (EZ) SET domain. Functional binding sites (BS) and structural regions identified in the literature are annotated above the sequence. Taxa are organized by pairwise identity (ID) and e-value scores compared to *Drosophila melanogaster*. Red boxes indicate non-conservative amino acid changes/deletions in functional motifs. Full taxon alignment in Appendix Figure 5.8. Percentage identity to the consensus sequence is color coded from dark blue (>80%) to light blue (>40%) and white (<40%).

SANT1 Domain is Poorly Conserved in Primitive Taxa and Nematodes.

Enhancer of zeste is also composed of two SANT domains, which were less conserved than the SET and CXC domains even between closely related taxa (Figure 5.9).

Identification of SANT1 via SMART was not possible for many basal taxa, including fungi, comb jelly (*Mnemiopsis leiydi*), sponge (*Amphemidon queenslandica*), placozoan (*Trichoplax adhaerens*), nematodes and platyhelminths. Sequences from these species also had poor similarity scores in a Blast pairwise alignment with *D. melanogaster* (Appendix: Table 5.2). Multiple sequence alignment highlighted several conserved residues (Gln, Glu and Leu) in nematodes and platyhelminthes in the N-terminus, suggesting these lineages possess a divergent SANT1 compared with other animals.

Sponge and comb jelly had only small fragments that were identified in the SANT1 region; it is unclear whether these basal taxa have lost this domain entirely or if it is so divergent that it is no longer recognizable or even functional. *Trichoplax adhaerens* does share many conserved residues with distantly related taxa (i.e. *Drosophila melanogaster*) despite having a poor Blast score, indicating that SANT1 may have an ancient origin.

Fungi and choanoflagellates also share some residues with other taxa, although it is difficult to determine if this region is homologous to SANT1 or whether this domain is a metazoan duplication of SANT2 because of the low similarities to both regions.

Compared with SANT1, the SANT2 domain appears to be more conserved across the phylogeny, though sequence conservation declines within fungi, nematodes and certain basal taxa compared to the remainder of species sampled.

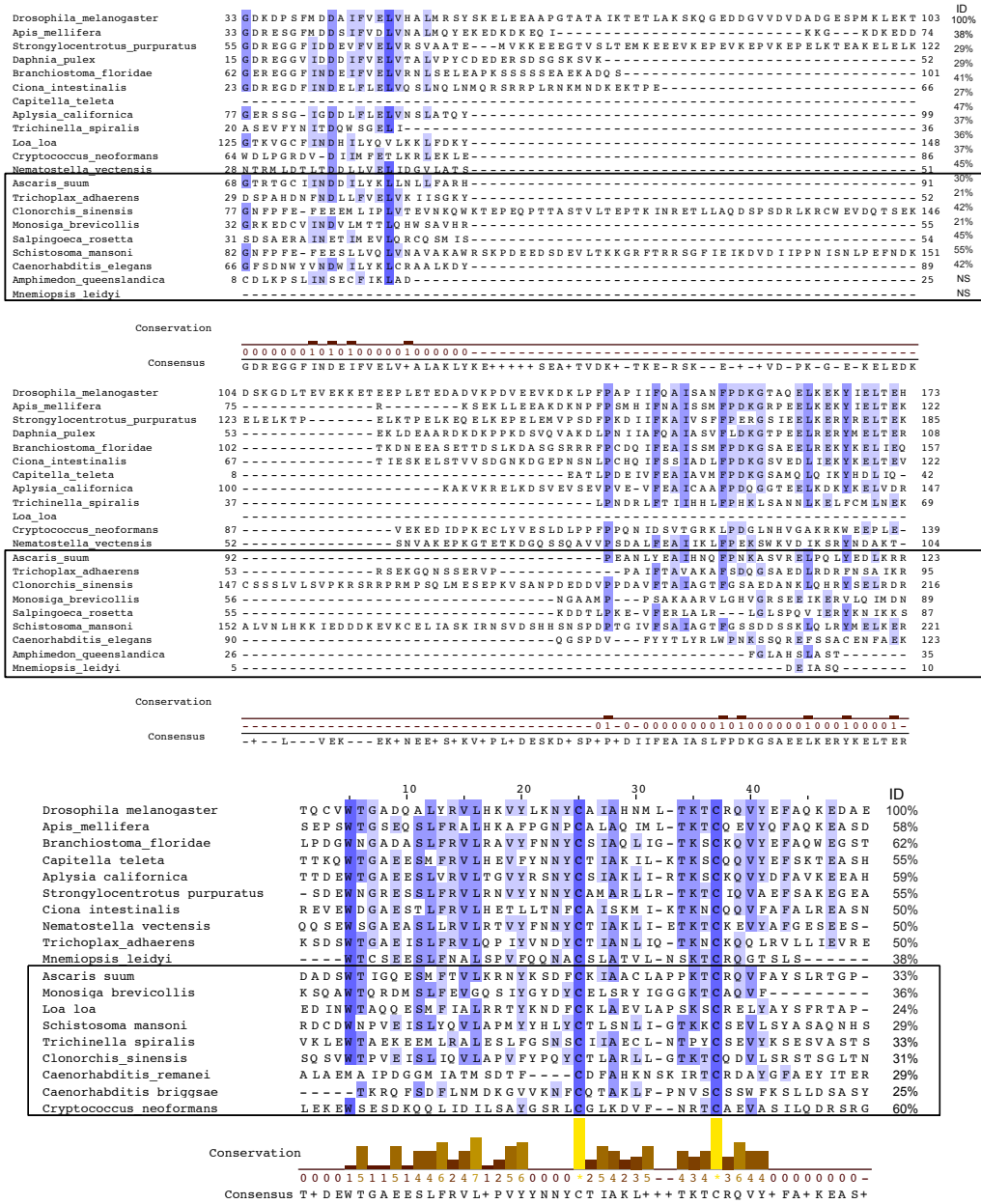


Figure 5.9 Multiple sequence alignment of the enhancer of zeste (EZ) SANT domains (SANT1 top two panels, SANT2 bottom panel) for representatives across major animal groups. Taxa are organized by pairwise identity (ID) and e-value scores compared to *Drosophila melanogaster*. Nematodes are placed within the grey box. Full taxon alignment in Appendix Figures 5.10 and 5.11. Percentage identity to the consensus sequence is color coded from dark blue (>80%) to light blue (>40%) and white (<40%).

Low Conservation of the *Drosophila* EZ-ESC Binding Region.

In addition to the four major domains described above, EZ possesses an N-terminal motif of 33 amino acids, known as EBD. This motif has been identified in *Drosophila* as an important binding region for ESC (Jones et al., 1998; Han et al., 2007). Similar to Jones et al. (1998) we found that this motif was poorly conserved across taxa (Appendix Figure 5.12). Despite the low similarity of this region, Jones et al. (1998) demonstrated that asparagine at position 40 was crucial for ESC binding and that this particular residue was highly conserved across fly, mouse and human. We also found high conservation of this residue with the exception of *Hydra vulgaris*, nematodes and fungi. Four residues, W-1, N-4 in the N-terminus and K-44, W-47 in the C-terminus were highly conserved within insects and some non-insect arthropods (Appendix Figure 5.12). Outside of arthropods, there is low conservation suggesting that this ESC-EZ binding region has undergone significant evolutionary changes and may have evolved species or lineage-specific motifs for binding ESC if residues other than asparagine are required.

Identification of Suz12 Homologs in Nematodes

Suz12 is the least studied protein of the three core subunits. Though the functional role of Suz12 is still not fully understood, the VEFS box is known to be an important binding site for EZ (Yamamoto et al. 2004). Using SMART we were able to identify the VEFS box in all taxa with Suz12 homologs, which is supported by significant BLAST scores in pairwise alignments with *D. melanogaster* (E-value >4E-04). As in other studies, we found no evidence of Suz12 homologs in the *Caenorhabditis* lineage (Ketel et al., 2005). Homologs were identified for other nematode species, indicating that this

polycomb gene has not been entirely lost in nematodes as previously assumed (Birve et al., 2001) (Figure 5.13).

Although Suz12 homologs were identified in nematodes, we found that the DSE-E-D motif was highly divergent. Mutations in this motif, particularly the first aspartic acid (D), have been shown to reduce histone methyltransferase activity by at least 10-fold in *Drosophila* (Ketel et al., 2005). Other highly conserved residues downstream from DSE-E-D motif have been shown to play important roles in catalytic efficiency (EK) and EZ-Suz12 assembly (W, N) (Rai et al., 2013) and these also displayed non-conservative mutations in fungi, choanoflagellates and nematodes. It is unknown whether these non-conservative mutations have similar negative consequences to the alanine substitutions used in the mutagenic experiments (Simon et al. 1995; Ng et al. 2000).

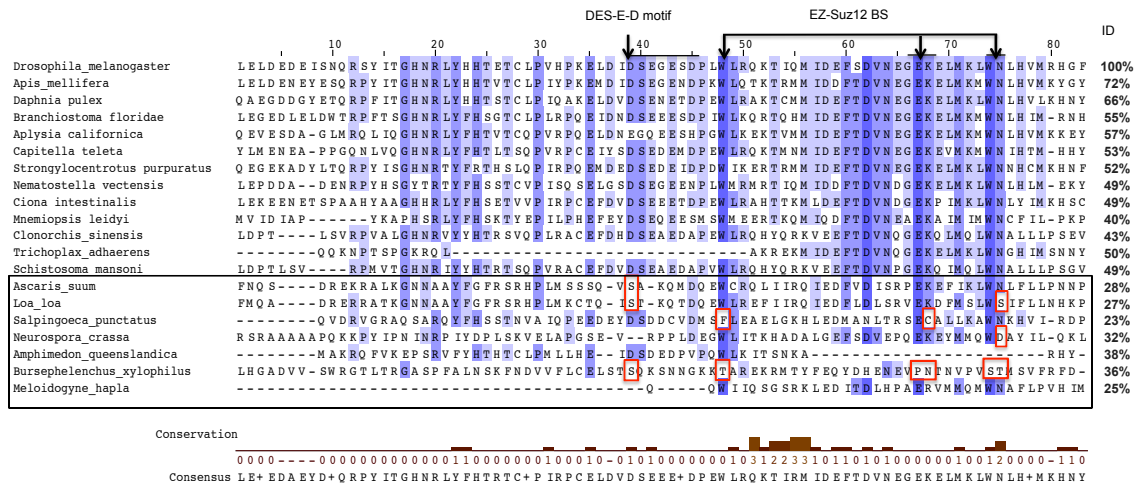


Figure 5.13 Multiple sequence alignment of the Suz12 VEFS box. Functional motifs and binding sites (BS) are annotated above the sequence. Red boxes indicate non-conservative residues in known functional regions. Taxa are organized by pairwise identity (ID) and e-value scores compared to *Drosophila melanogaster*. Nematodes, sponge, fungi and choanoflagellates are placed within the grey box. Full taxon alignment in Appendix Figure 5.14. Percentage identity to the consensus sequence is color coded from dark blue (>80%) to light blue (>40%) and white (<40%).

The zinc finger is another region that has been shown to be critical for EZ-Suz12 binding and PRC2 assembly (Birve et al., 2001; Rai et al., 2013). The Cis₂-His₂ zinc finger was less conserved than the VEFS box and was not recognized by SMART for nematodes and some platyhelminthes. These sequences also produced low BLAST scores (Appendix Table 5.2) when compared to *Drosophila melanogaster*. No record of this domain was found for either *T. adhaerens* or *Drosophila simulans*, which possessed fragmented residues in this region suggesting loss of this domain in these taxa. In contrast, this domain was identified in fungi by SMART; however, Blast scores of this domain were very poor (Appendix Table 5.2B). Thus, results from SMART and Blast are not always universally successful in identifying homologs.

In some fungi, choanoflagellates and nematodes the highly conserved terminal H and C-H residues were replaced by non-conservative substitutions (Appendix Figure 5.15). Recent work has shown that while the zinc finger is dispensable for PRC2 assembly and enzymatic activity, it is critical for Suz12 binding at chromatin targets (Rai et al., 2013). These results suggest that the non-conservative mutations observed in fungi and nematodes may have evolved as alternative binding sites or that other residues in the protein perform this function.

Evolutionary Conservation and Divergence in β -blades and Loops of ESC

In order to assemble onto chromatin targets, Suz12 and EZ must bind in a complex with ESC. The ESC protein is largely composed of WD40 repeats which adopt a circular β -blade propeller structure and function as a scaffold for protein-protein interactions. External loops also protrude from the top and bottom surface of this structure, serving as additional binding sites for protein interactions (Ng, et al. 1997; Ng et al. 2000). We

found that most taxa possessed all seven repeats; all of which are required for ESC function due to the structural constraints of the β -propeller (Simon et al. 1995; Tie et al. 1998). Taxa missing repeats were most likely due to the presence of partial sequences in the database. Interestingly, our data suggest that functional importance may vary across repeats. In pairwise comparisons with *D. melanogaster*, we found very high conservation for WD40-5 (75%) followed by WD40-4 (67%) and WD40-3 (63%) with a similarity of 53-59% for the remaining repeats. Some of these repeats, particularly WD40-1, 2, 6 and 7, were difficult to identify in fungi and nematodes due to high sequence divergence; these could only be identified in SMART using the outlier homologs database or in sequence alignments with related taxa.

We annotated the multiple alignments with the β -strands and exposed loops to examine the degree of conservation in these protein-binding regions. Each repeat folds into four stacked β -strands (a, b, d and d). On top of the β -propeller structure are two exposed loops per repeat that connect the specific β -strands, which include the d-a and b-c loop and on the bottom surface, the a-b and c-d loop. The external loops and the d β -strand are the most exposed parts of ESC serving as potentially important binding sites for protein interactions and therefore should be subject to strong evolutionary constraints (Ng et al., 1997; Jones et al., 1998). Overall, we found that the a-b loop exhibited very poor conservation compared to the d-a loop and the b-c loop, which were more highly conserved across taxa particularly in WD3, 4 and 5 (Figure 5.16). Furthermore, a region of the WD4 d-a loop (RDE-216 and M-263), important for efficient binding of EZ, was also divergent in nematodes, plathelminthes and the outgroups (Tie et al., 1998; Ng et al., 2000). These residues were highly conserved across all other taxa.

As for the other polycomb genes, we found that residues shown to be important for ESC function in *D. melanogaster* (Jones et al., 1998; Ng et al., 2000, 1997) were replaced by substitutions that alter amino acid properties in fungi, choanoflagellates and nematodes. Additionally the results suggest that the d-a and b-c loops are under strong purifying selection in particular repeats, while in others, selection appears to be relaxed. Thus, WD3, 4 and 5 may be the primary binding sites for assembly of PRC2, while loops in other repeats may be involved in lineage specific protein interactions.

Why Are Polycomb Genes So Different in Nematodes Compared to All Other Animals?

The results of our study raise the question as to why polycomb genes have been able to tolerate so many non-conservative substitutions in nematodes compared to all other animals. The PRC is an important regulator of the Hox gene cluster, which has been dramatically reduced, dispersed, and reorganized throughout the genome in nematodes (Aboobacker & Blaxter, 2003). Divergence in Hox gene sequence evolution may provide some explanation for the unusual placement of nematodes in the polycomb phylogeny, indicating potential co-evolutionary changes between the Hox and polycomb group genes. Intriguingly, although nematodes appear to have PRC2 homologs, a functional PRC1 has not been identified, suggesting that nematodes have evolved alternative strategies for gene repression, possibly reflected in the divergent sequences observed in this study.

Figure 5.16. WD-40 repeats of the ESC protein for select taxa across major animal groups. External loops are annotated above the sequence alignment and outlined as red boxes. Each β -blade is indicated as arrows below. Non-conservative substitutions in functional motifs in WD40 are shown in small red boxes. All taxa are organized by pairwise identity (ID) and e-value to *Drosophila melanogaster*. Full taxon alignments in Appendix Figure 5.17. Percentage identity to the consensus sequence is color coded from dark blue (>80%) to light blue (>40%) and white (<40%).

WD40-1

	d-a loop	a-b loop	b-c loop	ID
<i>Drosophila_melanogaster</i>	K E N H G A N I F G V A F N T L L G K D E P	O V F A T A G S N R R V T V Y E		100%
<i>Mnemiopsis_leidy</i>	K E D H K E P I F G V A F N P Y N H P D H P	V V F A T V G S N R R V T I Y E		65%
<i>Saccoglossus_kowalevskii</i>	K E D H G Q P L F G V Q F N T H C Q E G D A O	I F A T V G S S R R V T V Y E		61%
<i>Daphnia_pulex</i>	K E D H G Q P L F G V Q F N H L L R D G Q	L V F A T V G S H R I S V Y E		58%
<i>Capitella_teleta</i>	K E D H G Q P V F G V Q F N Y H T K D G D P	V L F A T V G S N R R V T V Y E		56%
<i>Apis_mellifera</i>	K E D H G Q P L F G V Q F N H H L K E G E F	L I F A S V G S N R R V S I Y E		55%
<i>Strongylocentrotus_purpuratus</i>	K E D H G Q P I F G V I F N P Y R K E S D P	N V F C S V G S N R R V S I Y E		54%
<i>Hydra_vulgaris</i>	K E D H K Q P I F G V Q F O O L I G E D D E	L I F G T V G S N R R V S V Y K		53%
<i>Ciona_intestinalis</i>	K E D H G Q P L F G V S F C H Q T S K D E Y	P M F A S V G S N R R I A V Y E		53%
<i>Loa_loa</i>	Y E G H K K T I Y G V A F N P Y L - I A N F H	- P A T V G E N R R V S I Y S		51%
<i>Branchiostoma_floridae</i>	K E D H G Q P L F G V Q I C F Y Y K E S Q A I	I F A T V G S N R R V T I Y E		50%
<i>Lottia_gigantea</i>	Q E D H S Q P I F G L Q I N L N A P E T D P	L T F A T V G N N R R V S V Y E		50%
<i>Clonorchis_sinensis</i>	R E S H G R S V F G V A F S I R S R I S D P	L L F A T V A G N F V T I Y Q		45%
<i>Schistosoma_mansoni</i>	R E T H S Q S V F G V A F S V R S Q P T D P	L L F A T V A S H Y V T V Y Q		43%
<i>Trichoplax_adhaerens</i>	L E D Q K K A I Y G C A F N Q Y A G I D E H O	A V A T V G G S F L H M Y S		43%
<i>Caenorhabditis_elegans</i>	L E D Q K K A I Y G C A F N Q Y A G I D E H O	A V A T V G G S F L H M Y S		41%
<i>Meloidogyne_hapla</i>	V E S H K T T V Y A I A F N T F T P Q E E T	S Y F A T A G K N K V S V Y S		40%
<i>Ancylostoma_ceilanicum</i>	Y Q Q H R O P I Y A C A F N P Y Q P E G C V	P V L A T A A K N M I T I Y E		39%
<i>Caenorhabditis_japonica</i>	G E Q H D K L Y N C D E N P Y I G W E Q T O	V L A T V G G T K R L V H E		37%
<i>Cryptococcus_neoformans</i>	E A T H S P E S Y R R D A N S Q R W T S H P Y	V I I H Q G D S L T N Y D		35%
<i>Salpingoeca_punctatus</i>	R O P V N D K L Y D V R A N L F T G V - - -	E F A V V G C G V S I W T		33%



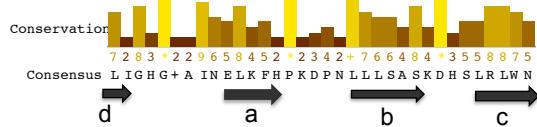
WD40-2

	d-a loop	a-b loop	b-c loop	ID
<i>Drosophila_melanogaster</i>	D P P F D E V F Y T C A W	S Y D L K T S S P L L A A A G Y R G V	I R V I D	100%
<i>Daphnia_pulex</i>	D P D L E E N F Y T C A W	S Y D E E T G K P I L A A A G S R G I V R I I S	64%	
<i>Apis_mellifera</i>	D P P F E E N F Y T C T W	T Y D - D S G K P L L A V A G S R G V V R V I S	64%	
<i>Trichoplax_adhaerens</i>	D P P F E D F Y T C A W	S V L H N T S E L I L A I A G A G R V I R I I N	62%	
<i>Strongylocentrotus_purpuratus</i>	D A P S D E N F Y T C A W	T W E E T T G I P L L A V A G S R G V I R I I S	61%	
<i>Saccoglossus_kowalevskii</i>	D P P A D E N Y Y T C A W	T I E E N T G A P L L A V A G S R G I R I R L I S	58%	
<i>Capitella_teleta</i>	D A P A D E S F Y T C A W	T Y D D V S H E P L L V A A G A R G I R F L S	55%	
<i>Branchiostoma_floridae</i>	D A N M E E N F Y T C A W	T M D E V A R O P L I A V A G L R G V I R I I S	55%	
<i>Lottia_gigantea</i>	D P S N E E N F Y C C A W	S H D D I T K Q P L V A V A G V R G I V R L I S	52%	
<i>Ciona_intestinalis</i>	D P S E E E N F Y S C T W	T V D S T S G H P L L A V A G S R G I R V R L N	49%	
<i>Hydra_vulgaris</i>	D S P E E S F Y A C S W	T Y D P D N R N P L F C F A G A K G I H I L N	47%	
<i>Meloidogyne_hapla</i>	D O P A K E C F Y A I T W	A Y N L D T S L H V L V V G G H R G I R V I S	46%	
<i>Cryptococcus_neoformans</i>	N O P E D D T L Y T L A W	T W H P F T C H P L I A V A G A N A L Y I I D	45%	
<i>Caenorhabditis_japonica</i>	T P N Q E E D L V A V T W	A L D T Y Q N A H R I V T G G L H G Q L Y V I N	43%	
<i>Mnemiopsis_leidy</i>	D P C T E E E I F Y T V C W	V V E N E K I E T M V A I A G L R G L I R V V S	39%	
<i>Loa_loa</i>	D S A K T E W E F S V C W	A Y D T E N D V H V V I A G G N R G I R V I D	38%	
<i>Ancylostoma_ceilanicum</i>	D P S P D M D I F T L T W	C Y D I T D K A H R I A F G G Y S G L I R L V D	37%	
<i>Salpingoeca_punctatus</i>	- P P - - - D V F C D W	L Y D E K E A K C H I A A G G S D G F L M V F D	37%	
<i>Clonorchis_sinensis</i>	D P A D D E E F Y C C A W	S R D T S G N Q Q L V A A A G K R G V I R I L C	30%	
<i>Schistosoma_mansoni</i>	D P A D K E E F Y C C A W	S R D T S G N Q Q V V A A A G K R G V I R I L C	30%	
<i>Caenorhabditis_elegans</i>	K V R R E E S L F T V T W	C Y D T Y R N P F K V V T G G T L G H I Y V I D	NS	



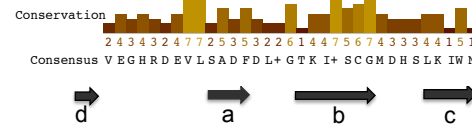
WD40-3

	d-a loop	a-b loop	b-c loop	ID
<i>Drosophila_melanogaster</i>	Y I S H G Q A I N E L K F H P H K L Q	L L L S G S K D H A I R L W N		100%
<i>Saccoglossus_kowalevskii</i>	Y V S H G N A I N E L K F H P H D Q N L L S V	S K D H S L R L W N		69%
<i>Daphnia_pulex</i>	Y V S H G H A I N E L K F H P S D P N L L S V	S K D H A L R L W N		68%
<i>Apis_mellifera</i>	Y I S H G H A I N E L K I H P K D P N	I L L S A S K D H A L R L W N		68%
<i>Capitella_teleta</i>	F I S H G Q S V N E L K F H P K D P N	I L M S V S K D H A L R L W N		65%
<i>Schistosoma_mansoni</i>	L V S H G S S I N E L R F H P R D P A L L F S	F S K D Y T I R L W N		64%
<i>Clonorchis_sinensis</i>	L V S H G A A I N E L R F H P R D P A L L F S	F S K D Y T V R L W N		64%
<i>Ciona_intestinalis</i>	Y I S H G N A V N E L K F H P Q M Q P	I L L S A S K D H S L R V W N		62%
<i>Lottia_gigantea</i>	Y V S H G N A I N E L K F H P K D S N L L S V	S K D H S M R I W N		62%
<i>Trichoplax_adhaerens</i>	Y P S H G N A I N E L K F H P L D P N	I L A S V G K D H I I H L W N		61%
<i>Hydra_vulgaris</i>	L Q S H G S A I N E L K T H P I E P L I I L S A	S K D H T I R M W N		59%
<i>Strongylocentrotus_purpuratus</i>	F I A H G N A V N E L K T H P H D S N L L S V	S K D H S V R L W N		58%
<i>Mnemiopsis_leidy</i>	L M S C G D S I N E V K L H P K D N N L L S A	S K D N S L R L W N		56%
<i>Branchiostoma_floridae</i>	Y T S H G H S V N E L K F H P S K P S I M L S V	S K D H S L R L W N		46%
<i>Ancylostoma_ceilanicum</i>	M Y S H G D H V N E M R T D P N N S M I F A S V	S K D T T I R L W N		45%
<i>Meloidogyne_hapla</i>	L L S H G D S I N E L R T S P T H P M I V A S A	S K D F T A R I W N		44%
<i>Cryptococcus_neoformans</i>	L K S H G D E I L C L A F A P L N P H I A S T S	S S D R S T R I W N		42%
<i>Salpingoeca_punctatus</i>	L H S H G S H V N D I R T H P K D E L L F A T A	S C D L S A R L W N		38%
<i>Loa_loa</i>	L I E H G D A V N D V R V F P N D S M I A S A	S K D F T A R I W N		38%
<i>Caenorhabditis_japonica</i>	L Q S C G G A I N D I R T S P A N S N L V A V A	S A D Q T V R I F H		32%
<i>Caenorhabditis_elegans</i>	L R S V G W E I N D I R T C P A N S N L I V C A S	S D Q S I R I H H	NS	



WD40-4

	d-a loop	a-b loop	b-c loop	ID
<i>Drosophila_melanogaster</i>	V E G H R D E V L S A D F L E G R R V I S	S G M D H S L K L W N	C	100%
<i>Daphnia_pulex</i>	V E G H R D E V L S A D F L E G R R V I S	S C G M D H S L K L W R		77%
<i>Apis_mellifera</i>	V E G H R D E V L S A D F D M K G E R I I	S C G M D H A L K L S		76%
<i>Branchiostoma_floridae</i>	V E G H R D E V L S A D F N A E G T R V V	S C G M D H S L K I W N		76%
<i>Strongylocentrotus_purpuratus</i>	V E G H R D E V L S G D F D I D G L R I A	S C G M D H S L K I W N		74%
<i>Trichoplax_adhaerens</i>	I D G H R D E V L S V D F D I L G K K I I	S G M D H S I K M W T		69%
<i>Hydra_vulgaris</i>	V D G H R D E V L G I D F D V L G T K I V	S C G M D H S L M P W S		69%
<i>Saccoglossus_kowalevskii</i>	I E G H R D E V L S A D F E D L D G K K I	I S C G M D H S L K I W N		67%
<i>Lottia_gigantea</i>	V D G H R D E L L S A D E N L E G T K I I	S C G M D H S L K I W N		67%
<i>Capitella_teleta</i>	V D G H R D E V L S G D I N L E G T M I V	S C G M D H S L K I W R		67%
<i>Ciona_intestinalis</i>	V E G H R D E V L S C D F N I F G T K I I	S C G M D H S L K I W N		62%
<i>Mnemiopsis_leidy</i>	L E G H R D E V L S C D E F L N A S Y V L	S C G M D H S V K M W S		62%
<i>Ancylostoma_ceilanicum</i>	Y E G H K D Q I L S L D W S L D S K Y I V	S C S M D H S I R L W Y		56%
<i>Schistosoma_mansoni</i>	A E G H R A E V L H G D I S L T G D L L L	S A G M D H C V K I H Y		55%
<i>Loa_loa</i>	V E G H L D Q V I S V D F D A E S E Y L A	S A S M D H T P K L H Y		55%
<i>Meloidogyne_hapla</i>	V Q G H R D Q V I S L D F E A T S H F L A T A	S M D H A V K L W H		54%
<i>Clonorchis_sinensis</i>	V E G H R A E I L H G D I S L T G D L L L	T A G M D H C I K I W R		52%
<i>Caenorhabditis_japonica</i>	L N C H R D Q V I S L D W D R D G N F L V	S C G M D H L S M R W		49%
<i>Caenorhabditis_elegans</i>	L E C H A G T I L S V D W S T D G D F I L	S C G F D H Q L M E W D		40%
<i>Cryptococcus_neoformans</i>	E G K G G H R A Y V V S C A F H P T K R A	I A T C S M D I T A K I		37%
<i>Salpingoeca_punctatus</i>	A T E F Q R M G I L S L A W H H T G K K L L V G	E K D G I V R L W	NS	



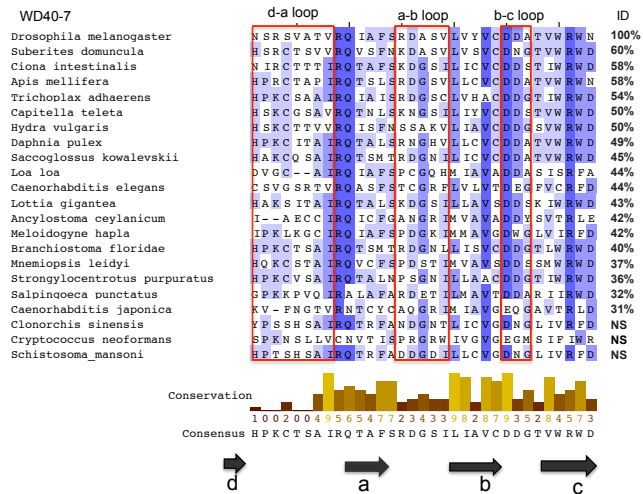
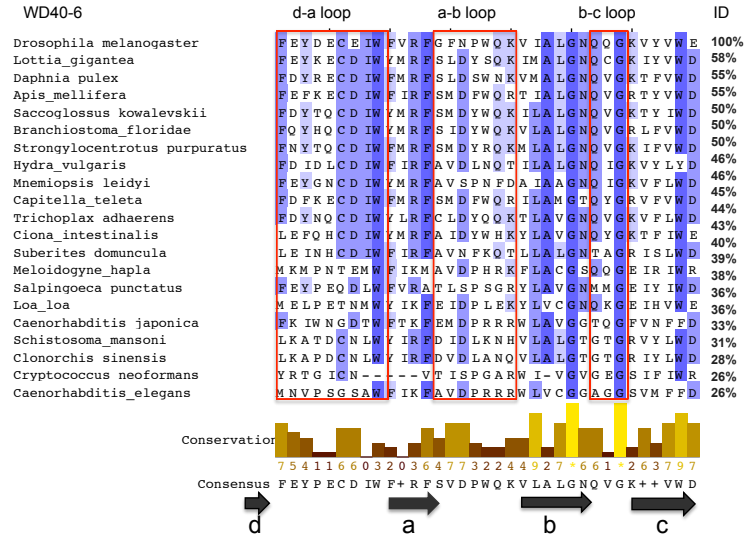
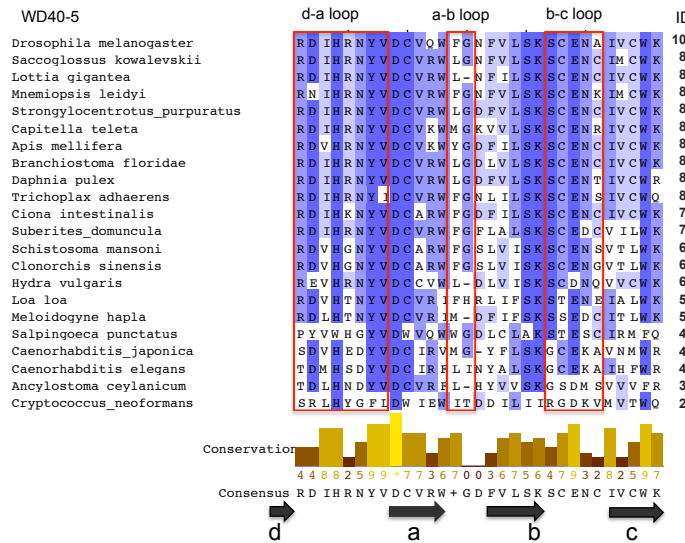


Figure 5.16 cont. WD-40 repeats of the ESC protein for select taxa across major animal groups. External loops are annotated above the sequence alignment and outlined as red boxes. Each β -blade is indicated as arrows below. Non-conservative substitutions in functional motifs in WD40 are shown in small red boxes. All taxa are organized by pairwise identity (ID) and e-value to *Drosophila melanogaster*. Full taxon alignments in Appendix Figure 5.9. Percentage identity to the consensus sequence is color coded from dark blue (>80%) to light blue (>40%) and white (<40%).

The lack of a Suz12 homolog in *Caenorhabditis* also supports this view. Within Nematoda, Hox sequence evolution has been more rapid in *Caenorhabditis* (Aboobaker & Blaxter, 2003); we also found a similar trend with the greatest sequence divergence in polycomb genes for this lineage. An unusual rapid rate of sequence evolution may be a hallmark of the Nematoda, as evidenced by the more rapid evolution of orthologs in *C. elegans* versus other metazoans (Mushegian et al. 1998; Coghlan & Wolfe, 2002). Nematodes have very short generation times, increasing the opportunity for accumulation of chromosomal rearrangements and mutations over evolutionary time. This rapid rate of protein evolution may be in response to selection for reduced genome size, increased development rates and the staggering range of specialized lifestyles exhibited by nematodes (Kortschak et al. 2003). Based on these findings we propose that the evolutionary tinkering of polycomb genes in nematodes may have been a consequence of relaxed selection due to the dispersion, rapid evolution and loss of Hox genes and/or selection of novel binding sites for alternative interacting partners/cofactors.

Conclusions

The evolution of the polycomb repressive complex in invertebrate animals has remained largely conserved even across distantly related taxa from sponges and comb jellies to early deuterostomes. Despite overall strong conservation in residues of known functional importance, certain domains, for example the Suz12 zinc finger, the SANT1 domain from Enhancer of zeste and some of the ESC WD40 repeats, appear to be subject to relaxed selection. These regions were more divergent and sometimes difficult to identify with current databases. These divergent sequences were mostly found in primitive taxa but surprisingly also in nematodes, which may reflect differences in Hox gene regulation.

Strong conservation is likely due to structural constraints and important binding sites between the different polycomb subunits, while divergent regions may have evolved for lineage-specific protein interactions. This work provides insight into the evolution of PRC2 and reveals that the subunits of this complex have distinct evolutionary histories. The β -propeller structure of Extra sex combs emerges as the most conserved element, while Enhancer of zeste SANT1 domain is the most divergent.

CHAPTER VI

EPILOGUE

My journey into the field of Evo-Devo was initially inspired by the work of Dr. Sean B. Carroll; my introduction to his work was through his Howard Hughes Medical Institute lectures on science and his book, *From DNA to Diversity: Molecular Genetics and the Evolution of Animal Design*. Together, these resources introduced me to the fascinating insight that, despite the extraordinary diversity of morphological forms, animal bodies are built from a common genetic toolkit. Furthermore, diversity can arise from differential gene expression of conserved genes and co-option of regulatory networks. This research led me to the exciting work on the evolution and development of butterfly wing color patterning by researchers such as Antonia Montiero, Paul Brakefield, Patricia Beldade, Fred Nijhout, Robert Reed, and Chris Jiggins. The work of these authors has played a major influential role in the development of my research and also my enthusiasm for this field.

Studying Evo-Devo raised my awareness of the current debate regarding the potential limitations of the Modern Synthesis. Many biologists including Sean Carroll, Massimo Pigliucci, Carl Schlichting and Mary West-Eberhard are calling for an extension of the Modern Synthesis to incorporate gene regulatory networks, phenotypic plasticity and epigenetics in generating heritable, novel phenotypes. Below, I highlight some of the

main conclusions from this dissertation that focuses on each of these areas. In doing so, I identify possible future avenues of research that would greatly improve understanding of butterfly wing pattern development.

1. *Drosophila* Wing Gene Regulatory Network is Conserved in *V. cardui* and Peaks in Expression During Late Larval and Early Pupal Stages.

Prior to this dissertation, there was little information available on gene expression patterns in the developing wings of *V. cardui*. There have also been no studies examining the temporal dynamics of the wing gene regulatory network for any butterfly species and the correspondence of these networks to temporal dynamics of the downstream pigment genes. Many genes have been highly conserved across major animal phyla and are redeployed in novel developmental contexts to produce a diversity of animal forms. The research conducted in Chapter 2 reveals that the regulatory genes characterized in the network for wing development in *Drosophila* are also expressed in the developing butterfly wing. These results indicate high conservation of this regulatory network between insects separated by ~200 million years, despite significant differences in wing morphology. In *V. cardui*, and probably other butterflies, expression of genes in this regulatory network expression peaks as wing color patterns are being established. Expression then declines dramatically during late pupal development, as melanin gene expression is at peak levels. In contrast, the peak in expression of ommochrome genes overlaps with expression of the patterning genes.

Taken together, these results demonstrate that patterns of expression appear to be highly coordinated during wing development. With the exception of the Hox cofactor *extradenticle*, expression of all genes declined to low levels as pupal development progressed. One important observation was that genes involved in melanin synthesis have

peak expression during late pupation in conjunction with increased expression of *extradenticle*. Whether this Hox cofactor has been co-opted to perform a novel developmental function in wing color patterning or whether it has simply retained its role in notum development (as observed in *Drosophila*) is completely unknown. If patterning genes directly regulate melanin pigment genes, then either low expression levels are sufficient or regulation is via an indirect and as yet unidentified pathway.

Immunohistochemistry work examining spatial expression patterns of Extradenticle would help elucidate its function in butterfly wings.

2. Eyespots Reveal Complex Patterns of Trait Integration and Modularity That Are Disrupted Following Plastic Responses to Environmental Perturbation.

The development of butterfly wing colors patterns seems to involve gene regulatory networks that have been co-opted in specific regions of the wing. Environmental stimuli also alter the expression of these networks. Chapters 3 and 4 demonstrate that hindwing eyespots exhibit phenotypic plasticity in response to temperature shock and pupal injection of heparin sulfate. Chapter 3 revealed a number of interesting observations regarding eyespot plasticity. First, plasticity varies significantly across different eyespots, with the centrally located eyespots exhibiting the highest levels of plasticity for both treatments. These results strongly imply that some intrinsic property of these eyespots, perhaps morphogen or hormone levels along the central wing axis enhances their susceptibility to environmental perturbation. Whether this sensitivity is due to variation in morphogen gradients during early pupation remains an intriguing question for future investigations.

Not only does plasticity vary across different eyespots, but also between the colored rings that comprise an individual eyespot. The outermost ring exhibits the lowest degree

of plasticity compared to the inner pigment rings, suggesting this ring is highly buffered to environmental perturbations. These results reveal that patterns of integration and modularity are highly complex, with integration of some traits (size) and modularity of others (pigment). I also found that these associations can be disrupted by environmental perturbations and may potentially alter the outcome of selection.

Finally, the results of this study raised queries regarding the effectiveness of concentration gradient models to explain the formation of the same pigment in different rings. Both the focal ring and the outer ring develop black pigment so it is not clear how a declining morphogen gradient from the interior to the outer ring can trigger the expression of the same pigment. These results call for a revised model to explain how rings located at different places within an eyespot display the same pigment.

3. Environmental Perturbation Alters Expression of Patterning and Pigment Genes While Polycomb Genes Remain Unaffected.

Although phenotypic plasticity is commonly observed in butterflies, few studies have tested whether environmental factors alter expression of genes involved in pigment synthesis or epigenetic modifications. Chapter 4 presents evidence that the epigenetic silencer, the polycomb repressive complex 1 and 2, is expressed during butterfly wing development and shows similar developmental trends and expression levels to those observed for the wing gene regulatory network.

Although it is thought that ESC is expressed primarily during embryogenesis, I observed similar expression of levels for ESC with EZ and SUZ12. This observation indicates that ESC may have additional functions during post-embryonic development in butterflies. Upregulation of this complex during the late larval and early pupal stages suggests these genes may play a role in the establishment of wing color patterns. Any role

that the PRC2 may play in eyespot development must occur prior to six days post-pupation; I did not observe differential expression of PRC2 genes in modified eyespots. Eyespot plasticity was associated with altered expression of *Engrailed*, *Spalt* and *Distal-less*, highlighting these as environmentally sensitive patterning genes. Only a single pigment gene (*tan*) was affected despite significant changes in eyespot pigmentation.

This is the first study to quantify expression patterns of pigment and patterning genes in butterfly eyespots. A handful of genes have previously been identified in butterfly eyespots; this study reveals novel genes expressed in this pattern element. Finally, because this study represents a single snapshot of gene expression, more studies are needed to explore the temporal dynamics of these genes during eyespot development. In particular future work should examine whether epigenetic mechanisms (DNA methylation and histone modifications) are involved in the regulation of patterning and pigment genes in butterflies.

4. The Evolution of Polycomb Repressive Complex 2 Shows Conservation Across Diverse Metazoans and Significant Divergence in Nematodes.

The PRC2 is a well-known epigenetic silencer, that regulates Hox genes and other developmental genes that influence animal morphogenesis. Most studies of the PRC focus on vertebrates and the PRC1, because it has experienced many duplication events. Chapter 5 broadens the focus from butterflies to a comprehensive phylogenetic analysis. This work examines the evolutionary history of the PRC2, an epigenetic silencer, across most invertebrates, starting from the earliest extant animals.

The phylogenetic analysis revealed that the evolutionary history of the major components of the PRC2 has diverged significantly from the known phylogeny of animals. Gene trees of each component of PRC2 also differ from each other, despite their

protein interactions. These results suggest that the three major core subunits of the PRC2 have experienced different selection pressures relative to each other and for neutrally evolving genes used for constructing animal phylogenies. These results do not exclude the possibility of coevolution between individual residues, which represent important binding sites.

An important observation from this study was that, despite the relatively high conservation of the core subunits across morphologically diverse taxa, PRC2 appears to have diverged significantly in the Nematoda. This result is particularly interesting considering that Hox genes also exhibit significant divergence in this lineage. It is possible that the divergence of PRC2 may have been influenced by changes in Hox organization in the nematode genome. Perhaps, in nematodes, PRC2 is more redundant due to the evolution of other regulatory mechanisms. These observations are intriguing but purely correlational. Future studies exploring PRC regulation of Hox genes in nematodes may shed light into the significant divergence that has occurred in these two important and conserved complexes that regulate animal morphogenesis.

Animals have evolved all manner of shapes, sizes, colors and patterns. Unraveling how such diversity arises from a common developmental toolkit has been an important and exciting challenge in evolutionary biology. Moving forward will require identifying the role of gene and network co-option, phenotypic plasticity and epigenetics in driving the evolution of morphological novelties. Thus, far, research on butterfly wing patterns has focused on the genetic aspects of pattern development including the genetic underpinnings of phenotypic plasticity. Here, I have attempted to broaden the scope of understanding wing pattern plasticity by incorporating the potential role of epigenetics. I

hope future investigations will also embrace the possibility that non-genetic mechanisms also play an important role in generating novel wing patterns. Finally, my work on the evolution of polycomb repressive complex 2 demonstrates high conservation of this epigenetic silencer across morphologically diverse animals. This analysis supports the widely observed conservation of genes across the animal phylogeny emphasizing that gene regulation rather than diversity in coding genes underlies the diversification of metazoans. My future work will continue to examine underlying mechanisms for the extraordinary diversity of animals beyond what genetic diversity can explain.

APPENDICES

Appendix Table 2.1

NCBI Blastn and Blastp results for *Vanessa cardui* transcripts following transcriptome assembly with CLC Genomics Workbench and multiblast (Blastx) against the *Drosophila* peptide database (FlyBase). Length of *V. cardui* transcript provided in amino acids (aa) and basepairs (bp). Status of the *V. cardui* coding sequence is provided as partial or complete. Complete is the length of the coding sequence.

Gene	Length	Coding sequence	Blastp best hit sp./Accession no.	Query cover	Identity	E-value
<i>Ultrabithorax</i>	253 aa	Complete	<i>Junonia coenia</i> Q8T940.1	100%	100%	0.0
<i>Ultrabithorax</i>	762 bp	Complete	<i>Junonia coenia</i> AY074760.1	100%	91%	0.0
<i>Wingless</i>	392 aa	Complete	<i>Helicoverpa armigera</i> AHN95659.1	100%	93%	0.0
<i>Wingless</i>	1179 bp	Complete	<i>Helicoverpa armigera</i> KJ206240.1	100%	82%	0.0
<i>Decapentaplegic</i>	379 aa	Complete	<i>Danaus plexippus</i> EHJ78539.1	100%	90%	0.0
<i>Decapentaplegic</i>	1110 bp	Complete	<i>Bombyx mori</i> NH_001145329.1	100%	72%	7e-159
<i>Spalt</i>	1058 aa	Partial	<i>Danaus plexippus</i> EHJ67088.1	100%	76%	0.0
<i>Spalt</i>	4558 bp	Partial	<i>Bombyx mori</i> XM_004931039.1	64%	78%	0.0

Appendix Table 2.1 cont.

Gene	Length	Coding sequence	Blastp best hit sp./Accession no.	Query cover	Identity	E-value
<i>Cut</i>	1200 aa	Partial	<i>Danaus plexippus</i> EHJ1235.1	78%	91%	0.0
<i>Cut</i>	4333 bp	Partial	<i>Bombyx mori</i> XM_004922717.1	77%	75%	0.0
<i>Invected</i>	141aa	Complete	<i>Junonia coenia</i> AAB46364.1	100%	97%	1e-93
<i>Invected</i>	426 bp	Complete	<i>Precis coenia</i> L41929.1	100%	84%	3e-126
<i>Vestigial</i>	695 aa	Partial	<i>Bombyx mori</i> XP_0049252918.1	46%	68%	5e-108
<i>Vestigial</i>	2134 bp	Partial	<i>Bombyx mori</i> XM_004925861.1	37%	80%	2e-169
<i>Distal-less</i>	357 aa	Complete	<i>Bicyclus anynana</i> AAL69325.1	100%	97%	0.0
<i>Distal-less</i>	1074 bp	Complete	<i>Junonia coenia</i> AF404110.1	98%	86%	0.0
<i>Engrailed</i>	362 aa	Complete	<i>Bombyx mori</i> XP_004933319.1	100%	76%	0.0
<i>Engrailed</i>	1089 bp	Complete	<i>Bombyx mori</i> XM_004933262.1	100%	73%	0.0

Appendix Table 2.1 cont.

Gene	Length	Coding sequence	Blastp best hit sp./Accession no.	Query cover	Identity	E-value
<i>Extradenticle</i>	367 aa	Complete	<i>Bombyx mori</i> XP_004928672.1	100%	96%	0.0
<i>Extradenticle</i>	1104bp	Complete	<i>Bombyx mori</i> XM_004928615.1	93%	82%	0.0
<i>Hedgehog</i>	382 aa	Complete	<i>Bombyx mori</i> XP_004925213.1	100%	84%	0.0
<i>Hedgehog</i>	1149 bp	Complete	<i>Bombyx mori</i> XM_004925156.1	99%	75%	0.0
<i>Scalloped TEF-1</i>	317 aa	Complete	<i>Danaus plexippus</i> EHJ63836.1	100%	86%	0.0
<i>Scalloped TEF-1</i>	954 bp	Complete	<i>Bombyx mori</i> AK38837.1	100%	83%	0.0
<i>Aschaete-scute</i>	106 aa	Partial	<i>Junonia coenia</i> AAC24714.1	100%	96%	1e-65
<i>Aschaete-scute</i>	319 bp	Partial	<i>Junonia coenia</i> AF071498.1	100%	87%	4e-105
<i>Serrate</i>	875 aa	Partial	<i>Danaus plexippus</i> EHJ63211.1	99%	85%	0.0
<i>Serrate</i>	2625 bp	Partial	<i>Bombyx mori</i> XM_004925686.1	65%	81%	0.0

Appendix Table 2.1 cont.

Gene	Length	Coding sequence	Blastp best hit sp./Accession no.	Query cover	Identity	E-value
<i>Optometer blind</i>	603 aa	Partial	<i>Bombyx mori</i> XP_004929044.1	96%	84%	0.0
<i>Optometer blind</i>	2110 bp	Partial	<i>Bombyx mori</i> XM_004928987.1	87%	85%	0.0
<i>ApterosusA</i>	399 aa	Complete	<i>Bombyx mori</i> BAK19079.1	100%	90%	0.0
<i>ApterosusA</i>	1200 bp	Complete	<i>Bombyx mori</i> AB587301.1	100%	78%	0.0
<i>Tan</i>	390 aa	Partial	<i>Heliconius melpomene</i> ADU32897.1	98%	86%	0.0
<i>Tan</i>	1175 bp	Partial	<i>Heliconius melpomene</i> GU386341.1	99%	76%	0.0
<i>Ebony</i>	574 aa	Partial	<i>Heliconius melpomene malleti</i>	100%	84%	0.0
<i>Ebony</i>	1724 bp	Partial	<i>Heliconius melpomene malleti</i> GU386340.1	99%	72%	0.0
<i>Pale</i>	558 aa	Complete	<i>Heliconius melpomene malleti</i> ADU32895.1	100%	93%	0.0
<i>Pale</i>	1677 bp	Complete	<i>Heliconius melpomene malleti</i> GU386339.1	100%	86%	0.0

Appendix Table 2.1 cont.

Gene	Length	Coding sequence	Blastp best hit sp./Accession no.	Query cover	Identity	E-value
<i>Ddc</i>	404 aa	Partial	<i>Danaus plexippus</i> EHJ63554.1	100%	92%	0.0
<i>Ddc</i>	1212 bp	Partial	<i>Vanessa carye</i> JQ786196.1	92%	92%	0.0
<i>Yellow</i>	494 aa	Complete	<i>Heliconius erato</i> ADX87341.1	99%	80%	0.0
<i>Yellow</i>	1485 bp	Complete	<i>Heliconius melpomene</i> GU063822.1	86%	77%	0.0
<i>Yellow-b</i>	454 aa	Complete	<i>Heliconius melpomene</i> ADX87345.1	99%	85%	0.0
<i>Yellow-b</i>	1365 bp	Complete	<i>Heliconius melpomene</i> GU063825.1	94%	76%	0.0
<i>Yellow-c</i>	408 aa	Complete	<i>Heliconius erato</i> ADX87347.1	99%	88%	0.0
<i>Yellow-c</i>	1227 bp	Complete	<i>Heliconius erato</i> GU063827.1	100%	78%	0.0
<i>Yellow-d</i>	896 aa	Complete	<i>Danaus plexippus</i> EHJ69631.1	100%	79%	0.0
<i>Yellow-d</i>	1341 bp	Complete	<i>Heliconius melpomene</i> GU063831.1	92%	72%	0.0

Appendix Table 2.1 cont.

Gene	Length	Coding sequence	Blastp best hit sp./Accession no.	Query cover	Identity	E-value
<i>Yellow-e</i>	268 aa	Partial	<i>Danaus plexippus</i> EHJ78055.1	84%	90%	7e-149
<i>Yellow-e</i>	806 bp	Partial	<i>Heliconius numata</i> GU063835.1	78%	78%	2e-144
<i>Yellow-f3</i>	467 aa	Complete	<i>Papilio xuthus</i> NBAM18870.1	95%	68%	0.0
<i>Yellow-f3</i>	1401 bp	Complete	<i>Papilio xuthus</i> AK402248.1	73%	70%	3e-126
<i>Vermillion</i>	410 aa	Complete	<i>Danaus plexippus</i> EHJ70119.1	100%	87%	0.0
<i>Vermillion</i>	1233 bp	Complete	<i>Vanessa cardui</i> partial DQ005628.1	54%	99%	0.0
<i>Cinnabar</i>	1347 bp	Complete	<i>Bombyx mandarina</i> EF210332.1	93%	73%	0.0
<i>Cinnabar</i>	448 aa	Complete	<i>Bombyx mori</i> ABM68366.1	95%	78%	0.0
<i>Kf</i>	297 aa	Complete	<i>Heliconius melpomene</i> ACS66705.1	100%	90%	0.0
<i>Kf</i>	894 bp	Complete	<i>Heliconius melpomene</i> GQ183897.1	100%	77%	0.0

Appendix Table 2.1 cont.

Gene	Length	Coding sequence	Blastp best hit sp./Accession no.	Query cover	Identity	E-value
<i>Scarlet</i>	671 aa	Complete	<i>Danaus plexippus</i> EHJ70567.1	100%	76%	0.0
<i>Scarlet</i>	20146 bp	Complete	<i>Bombyx mori</i> NM_001256993.2	91%	72%	0.0
<i>White</i>	686 aa	Complete	<i>Bombyx mori</i> BAH03523.1	100%	88%	0.0
<i>White</i>	2061 bp	Complete	<i>Bombyx mori</i> NM_001043569.1	99%	75%	0.0
<i>Serum response factor</i>	318 aa	Partial	<i>Bombyx mori</i> XP_004933723.1	41%	98%	1E-87
<i>Serum response factor</i>	2202 bp	Partial	<i>Juonia coenia</i> AF120007.1	16%	92%	1E-149

*Top hit for serrate nucleotide sequence was Bm8 interacting protein

Appendix Table 2.2. Primers used for qPCR validation.

Gene	Primer sequences
<i>Glutamate receptor</i>	Forward -TGGTATCGTCGCCATATTCG Reverse - GGAGAATATCAGCGCACCGA
<i>Wingless</i>	Forward - AAAAGCTGGCGAACCAAACA Reverse - GGTTGCGTGAACCTCCTGGAT
<i>Spalt</i>	Forward - GAAAACGATGGAGGGCAAGA Reverse - AGTCCATGCTGCAGTCGTCA
<i>Engrailed</i>	Forward - GTACACCTGCACCACCATCG Reverse - CGGTGAGTTCGGTTGGACTTT
<i>Distal-less</i>	Forward - GGCTTGGGATGTAAAGGTTGG Reverse - TGGTGGCTTCACGTCACAA
<i>Ddc</i>	Forward - ACGACATCGAGCGCGTTATA Reverse - GCTGTGCGGAAATAGGCGT
<i>Tan</i>	Forward - ATCCCCACGCAAGAAGACAG Reverse - GCAAGTGACCCGCATAGCA
<i>Pale</i>	Forward - CTCGTAGATGACGCCCGCT Reverse - GTGCACGAGCCTCTTCAAGC
<i>Ebony</i>	Forward - CATCCTGACTTTGGCCGTCT Reverse - TGCCAGCGAACAAGATGAGA
<i>Kf</i>	Forward - GCATGTGGTCGACGAGGTTT Reverse - TCGCTTGCTGTGGTAACGAA
<i>Vermillion</i>	Forward - AATGCGTGAACCCAACGAAG Reverse - GGATCGGTATATTTCCCGCC
<i>Cinnabar</i>	Forward - ACGAGGACATCGAGTGTCCT Reverse – AATGGAACGCCCTCGTACAT

Appendix Table 2.3

Summary of de-novo transcriptome assembly performed using CLC Genomics.

Descriptive Statistic	Value
Total size of transcriptome	31,689,449 bp
Total number of reads	446, 282,529
Mean no. reads for early 4th larval libraries	25,699,084.3
Mean no. reads for early 4th larval libraries	27,632,760
Mean no. reads for 2 day pupal libraries	23,187,027.7
Mean no. reads for 5 day pupal libraries	28,375,390.5
Mean no. reads for 8 day pupal libraries	34,585,348.5
Total number of contigs (unigenes)	89,065
Mean contig length	779.8 bp
Median contig length	446 bp
Max contig length	15,506 bp
N50	1,266 bp

* contigs >200bp

Appendix Table 2.4. References for classification of functional groups in the wing GRN.

Gene	Abbreviation	Function	Reference
<i>Cut</i>	<i>cut</i>	Selector gene	Brewster et al. (2001)
<i>Apterous</i>	<i>ap</i>	Selector gene	O’Keefe & Thomas (2001)
<i>Ultrabithorax</i>	<i>ubx</i>	Selector gene	Mann & Hogness (1990)
<i>Distal-less</i>	<i>dll</i>	Selector gene	Gebelein et al. (2002)
<i>Scalloped</i>	<i>sd</i>	Selector gene	Halder & Carroll (2001)
<i>Vestigial</i>	<i>vg</i>	Selector gene	Mann & Carroll (2002)
<i>Engrailed</i>	<i>en</i>	Selector gene	Zecca et al. (1995)
<i>Invected</i>	<i>inv</i>	Selector gene	Blair & Ralston (1997)
<i>Decapentaplegic</i>	<i>dpp</i>	Signaling molecule	Shen & Dahmann (2005)
<i>Wingless</i>	<i>wg</i>	Signaling molecule	Werner et al. (2010)
<i>Hedgehog</i>	<i>hh</i>	Signaling molecule	Ingham & McMahon (2001)
<i>Serrate</i>	<i>ser</i>	Signaling molecule	Walters et al. (2005)
<i>Optomotor blind</i>	<i>omb</i>	Transcription factor	Umemori et al. (2007)
<i>Extradenticle</i>	<i>exd</i>	Transcription factor	Rauskolb et al. (1993)
<i>Spalt</i>	<i>sal</i>	Transcription factor	Organista & De Celis (2013)

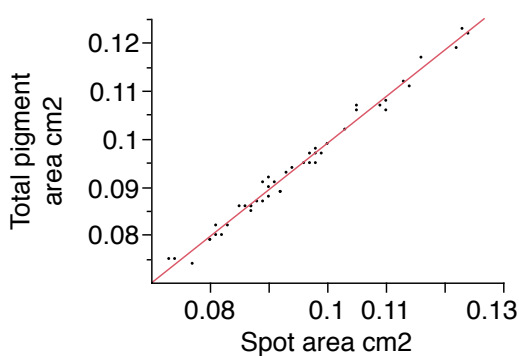
Appendix Table 2.4 cont.

Gene	Abbreviation	Function	Reference
<i>Ashaete-scute</i>	<i>ac/sc</i>	Transcription factor	García-Bellido & De Celis (2009)
<i>Serum response factor</i>	<i>srf</i>	Transcription factor	Chai & Tarnawski (2002)

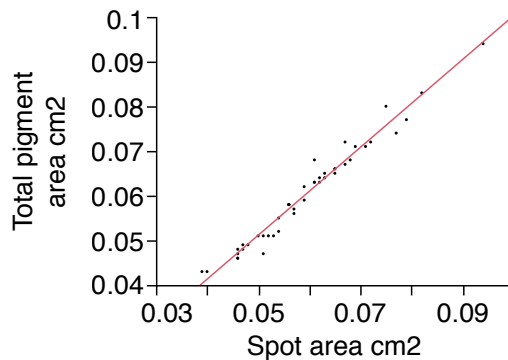
Appendix Table 3.1. Dunn-Šidák corrections for multiple comparisons following 2-way ANOVA for treatment x eyespot interactions comparing changes in pigment area (cm²). Significance indicated when p<0.006 for all 14 comparisons and p<0.009 for 6 comparisons.

Eyespot multiple comparisons	Eyespot	Black				
	Area	border	Yellow	Orange	Blue	Black focus
Control E1 vs E2	p<0.0001	p<0.0001	NS	p<0.0001	p<0.0001	NS
Control E1 vs E3	p<0.0001	p<0.0001	p<0.0002	p<0.0001	p<0.0001	p<0.0001
Control E1 vs E4	p<0.0001	p<0.0001	p<0.0001	p<0.0001		p<0.001
Control E2 vs E3	p<0.0001	NS	p<0.05	NS	NS	p<0.0001
Control E2 vs E4	NS	p<0.0001	p<0.0001	NS		p<0.0001
Control E3 vs E4	NS	p<0.0001	p<0.0001	p<0.008		p<0.0001
Control E1 vs Heparin E1	NS	NS	NS	NS		p<0.005
Control E1 vs Temp E1	p<0.001	NS	p<0.006	NS	NS	p<0.0001
Control E2 vs Heparin E2	p<0.0001	NS	NS	p<0.0001		p<0.0001
Control E2 vs Temp E2	p<0.01	p<0.05	p<0.0001	NS	p<0.0001	p<0.0001
Control E3 vs Heparin E3	p<0.0001	p<0.0001	NS	p<0.0001		p<0.0001
Control E3 vs Temp E3	p<0.002	NS	p<0.0001	NS	p<0.005	p<0.0001
Control E4 vs Heparin E4	NS	p<0.05	p>0.05	p<0.0001		p<0.0001
Control E4 vs Temp E4	NS	NS	NS	NS		p<0.0001

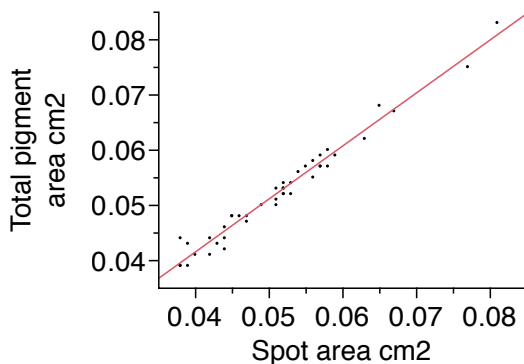
*Only 6 comparisons shown for blue pigment due to absence of blue in eyespots of butterflies treated with heparin and the lack of blue pigment in eyespot 4 for all treatment groups. p>0.006 NS



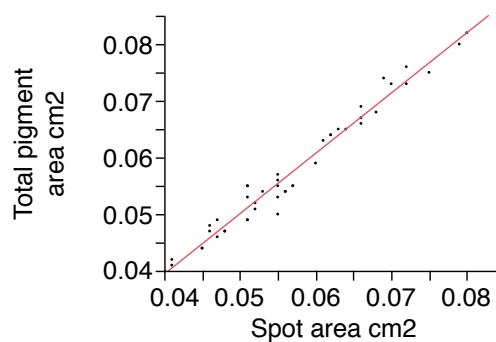
Eyespot 1: Control $r^2=0.99$ $n=45$, $p<0.0001$



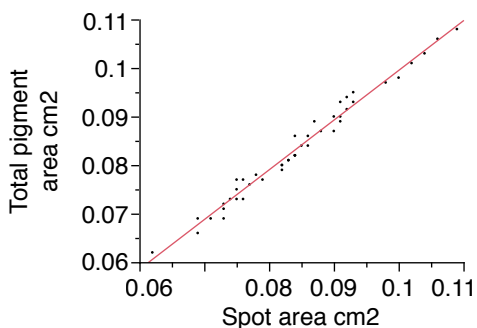
Eyespot 2: $r^2=0.97$ Control $n=45$, $p<0.0001$



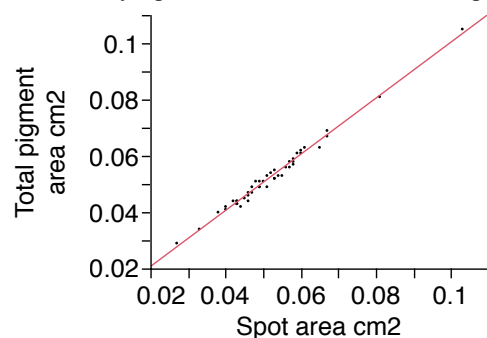
Eyespot 3: Control $r^2=0.97$ $n=44$, $p<0.0001$



Eyespot 4: Control $r^2=0.97$ $n=44$, $p<0.0001$

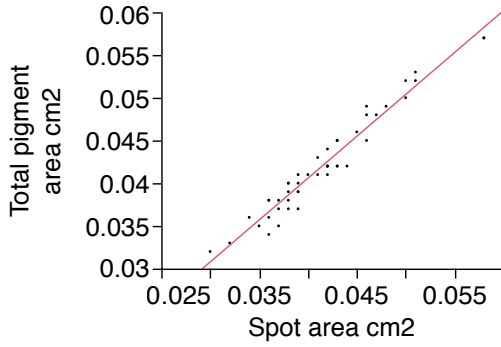


Eyespot 1: 37°C $r^2=0.98$ $n=45$, $p<0.0001$

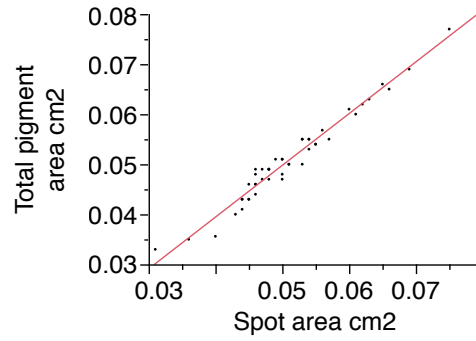


Eyespot 2: 37°C $r^2=0.99$ $n=46$, $p<0.0001$

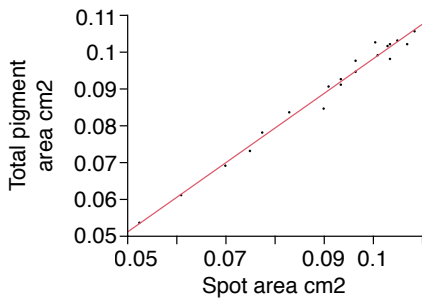
Appendix figure 3.3. Bivariate analysis of the sum of all pigments in each eyespot relative to the total eyespot area. Plots reveal a strong correlation in the measurements of the pigment area to the total eyespot area indicating area of the individual pigments were measured with high precision.



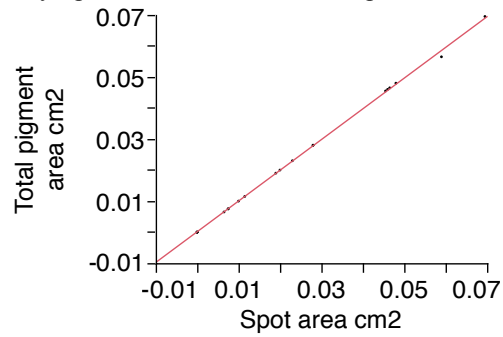
Eyespot 3: 37°C $r^2 = 0.95$ $n=46$, $p<0.0001$



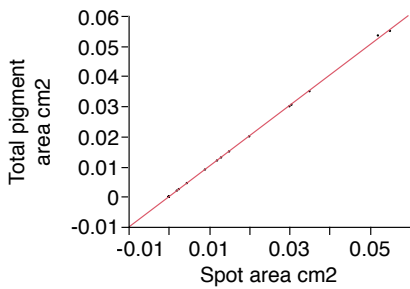
Eyespot 4: 37°C $r^2 = 0.96$ $n=46$, $p<0.0001$



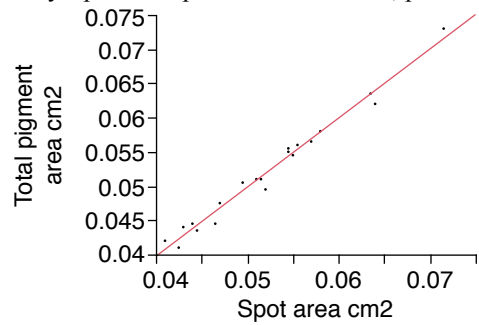
Eyespot 1: Heparin $r^2 = 0.98$ $n=20$, $p<0.0001$



Eyespot 2: Heparin $r^2 = 0.99$ $n=20$, $p<0.0001$



Eyespot 3: Heparin $r^2 = 0.99$ $n=20$, $p<0.0001$



Eyespot 4: Heparin $r^2 = 0.98$ $n=20$, $p<0.0001$

Appendix figure 3.3 cont. Bivariate analysis of the sum of all pigments in each eyespot relative to the total eyespot area. Plots reveal a strong correlation in the measurements of the pigment area to the total eyespot area indicating area of the individual pigments were measured with high precision.

Size	ES1	ES2	ES3	ES4
ES1	*	0.523	0.2	0.47
ES2	0.795	*	0.574	0.367
ES3	0.589	0.761	*	0.443
ES4	0.453	0.411	0.315	*

Border	ES1	ES2	ES3	ES4
ES1	*	0.35	0.29	0.57
ES2	0.01	*	0.16	0.28
ES3	0.156	0.387	*	0.32
ES4	0.125	0.29	0.653	*

Yellow	ES1	ES2	ES3	ES4
ES1	*	0.083	0.215	0.5
ES2	0.387	*	0.298	-0.15
ES3	0.416	0.567	*	0.1
ES4	0.34	0.151	0.446	*

Orange	ES1	ES2	ES3	ES4
ES1	*	-0.0122	-0.12	-0.2
ES2	0.15	*	0.52	-0.026
ES3	-0.0767	-0.17	*	-0.18
ES4	0.34	0.4	-0.026	*

Blue	ES1	ES2	ES3	ES4
ES1	*	0.16	0.2	*
ES2	0.18	*	0.57	*
ES3	0.1589	0.676	*	*
ES4	*	*	*	*

Black focus	ES1	ES2	ES3	ES4
ES1	*	0.418	0.322	-0.113
ES2	0.379	*	0.651	0.051
ES3	0.683	0.549	*	0.182
ES4	0.47	0.004	0.397	*

Appendix figure 3.11. Partial correlation matrices for eyespot size and proportion of each color ring. Values below the diagonal represent control eyespots and those above represent the temperature shock treatment. Wing area was used as a covariate. Partial correlations are used to measure the association (or integration) between pairs of traits, independent of associations with all other measured traits (Magwene, 2001, Allen 2008).

Size	ES1	ES2	ES3	ES4
ES1	*	14.7	1.88	11.48
ES2	44.988	*	18.4	6.7
ES3	19.17	38.94	*	10.05
ES4	10.34	8.33	4.7	*

Border	ES1	ES2	ES3	ES4
ES1	*	5.88	3.95	17.68
ES2	0.004	*	1.167	3.67
ES3	1.1	7.3	*	4.86
ES4	0.7	3.95	25	*

Yellow	ES1	ES2	ES3	ES4
ES1	*	0.32	2.12	13.23
ES2	7.3	*	4.27	1.05
ES3	8.55	17.45	*	0.46
ES4	5.53	1.04	9.98	*

Orange	ES1	ES2	ES3	ES4
ES1	*	0.006	0.67	1.88
ES2	1.02	*	14.5	0.03
ES3	0.268	1.32	*	1.51
ES4	5.53	7.85	0.03	*

Blue	ES1	ES2	ES3	ES4
ES1	1	1.167	1.84	*
ES2	1.48	1	17.68	*
ES3	1.152	27.48	1	*
ES4	*	*	*	*

Black focus	ES1	ES2	ES3	ES4
ES1	*	8.833	4.93	0.578
ES2	6.98	*	24.8	0.119
ES3	28.3	16.14	*	1.55
ES4	11.23	0	7.72	*

Appendix figure 3.12 Edge exclusion deviance matrices based on partial correlations and calculated using the EED formula described in the methods. EED uses partial correlations to test for conditional independence. Values highlighted in bold (<3.82) suggest conditional independence and values > 3.82 indicate eyespots are integrated. Values below the diagonal represent control eyespots and those above represent the temperature shock treatment.

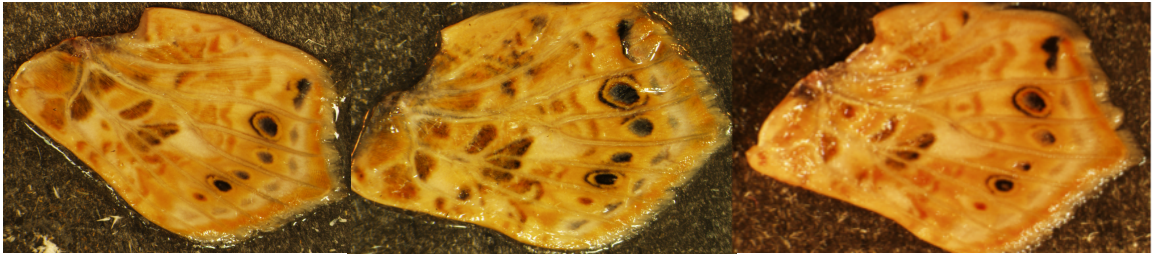
Appendix Table 4.2. List of primer pairs used for quantitative real time PCR.

Gene	Primer sequences
B-Actin	Forward -TGGTATCGTCGCCATATTCG Reverse - GGAGAATATCAGCGCACCGA
Ash1	Forward -AAACAGAATTCCCACCGGGT Reverse - AAATATTGTGGTGTACGGCGC
Glutamate receptor	Forward -TGGTATCGTCGCCATATTCG Reverse - GGAGAATATCAGCGCACCGA
Spalt (sal)	Forward - GAAAACGATGGAGGGCAAGA Reverse - AGTCCATGCTGCAGTCGTCA
Engrailed (En)	Forward - GTACACCTGCACCACCATCG Reverse - CGGTGAGTTCGGTTGGACTTT
Distal-less (Dll)	Forward - GGCTTGGGATGTAAAGGTTGG Reverse - TGGTGGCTTCACGTCACAA
Ultrabithorax (Ubx)	Forward - AGGCCTCAGGACTCCCCATA Reverse - TTGCGTATTGCTGCTCTCCC
Dopa decarboxylase (DDC)	Forward - ACGACATCGAGCGCGTTATA Reverse - GCTGTCGGGAAATAGGCGT
Tan	Forward - ATCCCCACGCAAGAAGACAG Reverse - GCAAGTGACCCGCATAGCA
Tyrosine hydroxylase (TH)	Forward - CTCGTAGATGACGCCCGCT Reverse - GTGCACGAGCCTCTTCAAGC
Ebony	Forward - CATCCTGACTTTGGCCGTCT Reverse – TGCCAGCGAACAAGATGAGA

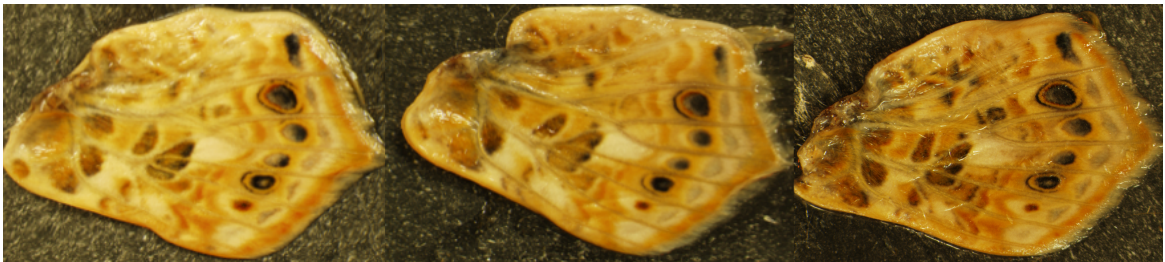
Appendix Table 4.2. cont.

Gene	Primer sequences
Polycomb (Pc)	Forward -ACCGAAACACAACACCTGGG
	Reverse - TTCGCCTCGTTCGTAGCTCT
Enhancer of zeste (Ez)	Forward - AGCTGAAGAAGGACTCCGCC
	Reverse - TTGTCGCAGGGCTGGTTAG

Control



Temperature shock



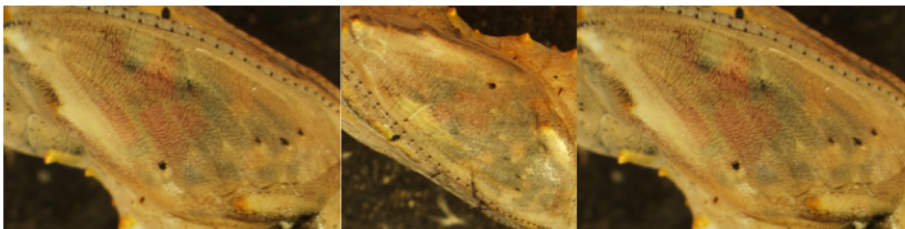
Heparin



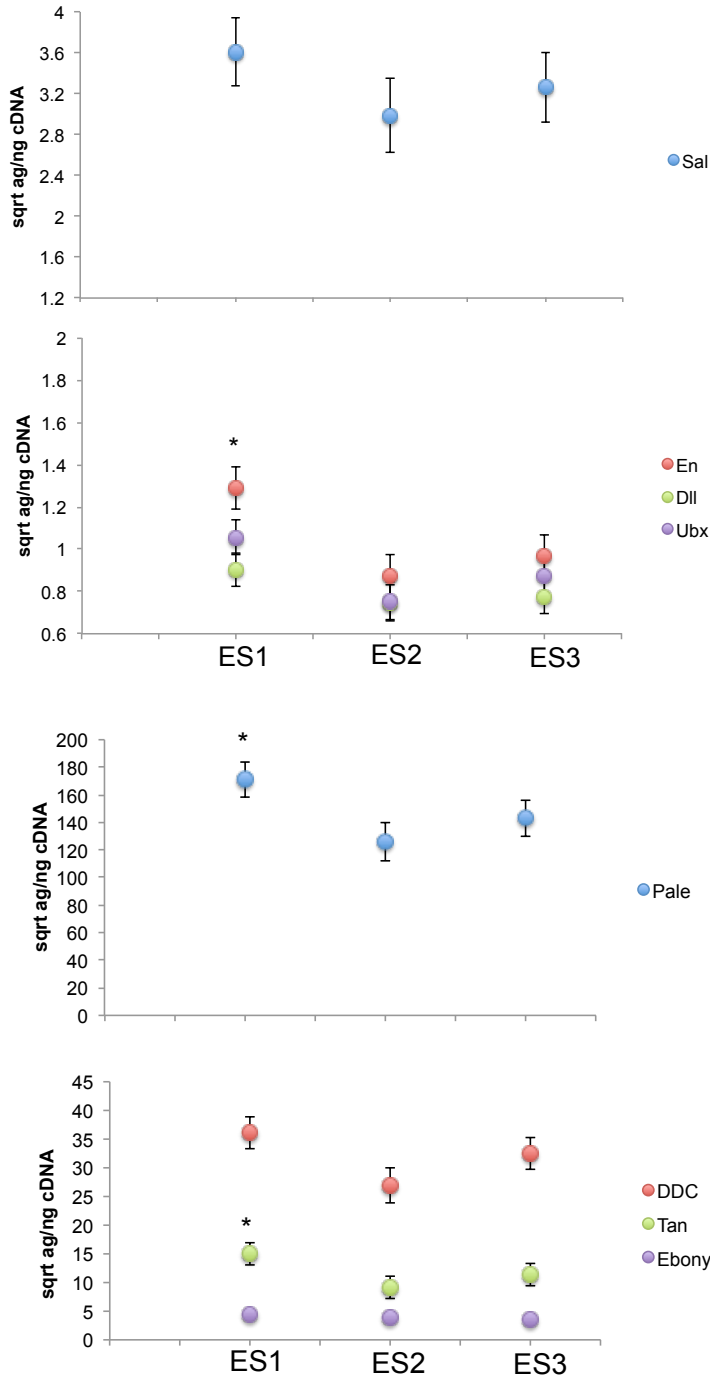
Control

Temp. shock

Heparin



Appendix Figure 4.1. Representative images of pupal hindwings at 6 days post-pupation prior to eyespot dissections. Bottom panel shows black and red pigmentation visible on the pupal cuticle used as an aid for timing of wing dissection.



Appendix Figure 4.4 Gene expression patterns across all three hindwing eyespots (Control only) for the patterning genes and melanin genes. Data represent the least square means square root transformed and errors bars are $1 \pm SE$ from the mean.

Appendix Table 5.1A. Taxon list for Enhancer of zeste (Ez) with accession numbers for NCBI and Uniprot. NF indicates no record was found in that database. Sequences for taxa in bold were obtained from WormBase (Nematoda), Mnemiopsis Genome Project (Ctenophora) or The Broad Institute (Choanoflagellate). Asterisks indicate sequences that were listed as partial in both NCBI and Uniprot. Amino acid length = aa.

Taxa	Common name	Phylum	NCBI reference	Uniprot reference	aa
<i>Acromyrmex echinator</i>	Panamanian leafcutter ant	Arthropoda	EGI68054.1	NF	761
<i>Acyrtosiphon pisum</i>	Pea aphid	Arthropoda	XP_003240462.1	J9K871	745
<i>Aedes aegypti</i>	Yellow fever mosquito	Arthropoda	XP_001663394.1	Q16JU6	752
<i>Amblyomma maculatum</i>	Gulf coast tick	Arthropoda	NF	G3MM78	715
<i>Amphimedon queenslandica</i>	Demosponge	Porifera	XP_003390508.1	I1EWD7	345
<i>Ancylostoma ceylanicum</i>	Hookworm	Nematoda	NF	NF	944
<i>Anopheles gambiae</i>	African malaria mosquito	Arthropoda	XP_307419.2	Q7PTY9	742
<i>Apis florea</i>	Common Eastern Bumble Bee	Arthropoda	XP_003690343.1	NF	746
<i>Apis mellifera</i>	Dwarf honey bee	Arthropoda	XP_003249917.1	H9KEQ7	746
<i>Aplysia californica</i>	California sea hare	Arthropoda	XP_005096266.1	NF	788
<i>Ascaris suum</i>	Pig roundworm	Arthropoda	NF	F1KYX6	676
<i>Atta cephalotes</i>	Leaf cutter ant	Arthropoda	NF	H9HT54	757
<i>Bombus impatiens</i>	Common Eastern Bumble Bee	Arthropoda	XM_003485728.1	NF	754
<i>Bombus terrestris</i>	Buff-tailed Bumblebee	Arthropoda	XM_003397352.1	NF	762
<i>Bombyx mori</i>	Silkworm	Arthropoda	XP_003397400.1	H9JY58	442
<i>Botryllus primigenius*</i>	Sea squirt	Urochordata	AB852574.1	T2HSG9	555
<i>Branchiostoma floridae</i>	Florida lancelet	Cephalochordata	XM_002605849.1	C3YCV4	625
<i>Brugia malayi</i>	Roundworm	Nematoda	XM_001902435.1	A8QGD6	652
<i>Bursephelenchus xylophilus</i>	Roundworm	Nematoda	NF	NF	745
<i>Caenorhabditis angaria</i>	Roundworm	Nematoda	NF	NF	324
<i>Caenorhabditis brenneri</i>	Roundworm	Nematoda	EGT54479	G0N6Y4	764
<i>Caenorhabditis elegans</i>	Roundworm	Nematoda	NP_496992.3	O17514	773

Appendix Table 5.1A continued

Taxa	Common name	Phylum	NCBI reference	Uniprot reference	aa
<i>Caenorhabditis japonica</i>	Roundworm	Nematoda	NF	K7GZH2	288
<i>Caenorhabditis remanei</i>	Roundworm	Nematoda	XP_003104313	E3MI00	841
<i>Caenorhabditis briggsiae</i>	Roundworm	Nematoda	CAP29342.2	A8X9M7	516
<i>Camponotus floridanus</i> *	Florida carpenter ant	Arthropoda	EFN68978.1	E2AC62	755
<i>Capitella teleta</i>	Polychaete worm	Annelida	ELT87938.1	R7T4Q6	527
<i>Clonorchis sinensis</i> *	Chinese liver fluke	Platyhelminthes	GAA55462.1	G7YR82	940
<i>Ciona intestinalis</i>	Transparent sea squirt	Urochordata	XP_002126205.2	H2YDH7	692
<i>Ciona savignyi</i>	Transparent sea squirt	Urochordata	NF	H2YDH6	734
<i>Crassostrea gigas</i>	Pacific oyster	Mollusca	EKC36964.1	K1R0K5	807
<i>Cryptococcus gattii</i>	Fungi		XM_003196759.1	XM_003196759.1	731
<i>Cryptococcus neoformans</i>	Fungi		XP_567801.1	XP_567801.1	718
<i>Culex quinquefasciatus</i>	Southern house mosquito	Arthropoda	XP_001848357	B0WI23	763
<i>Danaus plexippus</i>	Monarch butterfly	Arthropoda	EHJ78862	G6CIT8	733
<i>Daphnia pulex</i>	Common water flea	Arthropoda	EFX90346	E9FSB6	790
<i>Dendroctonus ponderosae</i> *	Mountain pine beetle	Arthropoda	ENN78407	N6UID1	742
<i>Dirofilaria immitis</i>	Heart worm	Arthropoda	NF	NF	721
<i>Drosophila ananassae</i>	Fruit fly	Arthropoda	XP_001957822	B3M5C3	751
<i>Drosophila erecta</i>	Fruit fly	Arthropoda	XM_001972254.1	B3NCL9	761
<i>Drosophila grimshawi</i>	Fruit fly	Arthropoda	XM_001984539.1	B4J1K3	762
<i>Drosophila melanogaster</i>	Fruit fly	Arthropoda	NM_079297.3	P42124	760
<i>Drosophila mojavensis</i>	Fruit fly	Arthropoda	XM_002009014	B4K VX5	741
<i>Drosophila persimilis</i>	Fruit fly	Arthropoda	XM_002026486.1	B4H6P3	749
<i>Drosophila pseudoobscura</i>	Fruit fly	Arthropoda	XM_001353744.2	Q2LZJ3	749
<i>Drosophila sechellia</i>	Fruit fly	Arthropoda	XM_002029970.1	B4HLW0	753

Appendix Table 5.1A continued

Taxa	Common name	Phylum	NCBI reference	Uniprot reference	aa
<i>Drosophila simulans</i>	Fruit fly	Arthropoda	XP_002084393	B4QP80	675
<i>Drosophila virilis</i>	Fruit fly	Arthropoda	XM_002047818.1	B4LET9	741
<i>Drosophila willistoni</i>	Fruit fly	Arthropoda	XM_002069020.1	B4N6Q3	768
<i>Drosophila yakuba</i>	Fruit fly	Arthropoda	XM_002094276.1	B4PEF8	760
<i>Echinococcus multilocularis</i>	Tapeworm	Nematoda	CDI99875.1	A0A068XY53	930
<i>Fusarium oxysporum cubense</i>	Fungi		EMT61537.1	N1RBD1	1026
<i>Haemonchus contortus</i>	Barbers pole worm	Nematoda	CDJ88240.1	U6P2S3	999
<i>Helobdella robusta</i>	Californian leech	Mollusca	ESO04666	T1G9T5	624
<i>Hydra vulgaris</i>	Fresh water polyp	Cnidarian	XP_004207451.1	T2MD94	724
<i>Ixodes scapularis</i> *	Black legged tick	Arthropoda	XM_002408909.1	B7Q167	737
<i>Loa loa</i>	Eye worm	Nematoda	EFO28287	E1FH06	732
<i>Lottia gigantea</i>	Owl limpet	Mollusca	ESP00855	V4CFQ6	783
<i>Marssonina brunnea</i>	Fungi		XM_007292563.1	K1XWQ2	1073
<i>Megachile rotundata</i>	Alfalfa leafcutter bee	Arthropoda	XM_003707504.1	Not found	758
<i>Metarhizium acridum</i>	Fungi		XM_007815540.1	E9EBZ3	1139
<i>Metarhizium anisopliae</i>	Fungi		XM_007818675.1	E9EMH0	1148
<i>Mnemiopsis leidyi</i>	Sea walnut	Ctenophora	NF	NF	304
<i>Monosiga brevicollis</i>	Choanoflagellate		XP_001742056	A9UNS2	2169
<i>Nasonia vitripennis</i>	Jewel wasp	Arthropoda	XM_001599009.2	K7JBX9	781
<i>Nematostella vectensis</i>	Starlet sea anenome	Cnidarian	XM_001622372.1	A7T142	688
<i>Oikopleura dioica</i>	Sea squirt	Urochordata	CBY12204	E4XQW8	692
<i>Onchocerca volvulus</i>	River blindness nematode	Nematoda	NF	NF	696
<i>Pediculus humanus corporis</i>	Human body louse	Arthropoda	XM_002427044.1	E0VLU5	729
<i>Polyandrocarpa misakiensis</i> *	Sea squirt	Urochordata	AB671227.1	G1UK06	566
<i>Rhipicephalus pulchellus</i>	Ivory ornamented tick	Arthropoda	NF	L7LTF5	715

Appendix Table 5.1A continued

Taxa	Common name	Phylum	NCBI reference	Uniprot reference	aa
<i>Rhodnius prolixus</i> *	Assassin bug	Arthropoda	NF	T1HYG5	747
<i>Salpingoeca rosetta</i>	Choanoflagellate		XP_004992157	F2UEQ4	508
<i>Schistosoma mansoni</i>	Blood fluke	Platyhelminthes	XM_002578972.1	C4QIH3	1026
<i>Solenopsis invicta</i> *	Red fire ant	Arthropoda	EFZ16931	E9IR73	639
<i>Strigamia maritima</i>	Coastal centipede	Arthropoda	NF	T1IUN1	345
<i>Strongylocentrotus purpuratus</i>	Purple sea urchin	Echinodermata	XP_790741	H3JA00	794
<i>Tribolium castaneum</i>	Red flour beetle	Arthropoda	XP_001811652.1	D6WFD9	721
<i>Trichinella spiralis</i>	Trichina worm	Nematoda	XP_003377384.1	E5SCV7	633
<i>Trichoplax adhaerens</i>	Flat animal	Placozoa	XM_002110946.1	B3RS40	682

Appendix Table 5.1B. Taxon list for Suppressor of zeste (Suz12) with accession numbers for NCBI and Uniprot. NF indicates no record found in that database. Sequences for taxa in bold were obtained from WormBase (Nematoda), Mnemiopsis Genome Project (Ctenophora) or The Broad Institute (Choanoflagellate). Asterisks indicate sequences that were listed as partial in both NCBI and Uniprot. Amino acid length = aa.

Taxa	Common name	Phylum	NCBI reference	Uniprot	aa
<i>Acromyrmex echinator</i>	Panamanian leafcutter ant	Arthropoda	EGI70249.1	F4W6B4	735
<i>Acyrtosiphon pisum</i>	Pea aphid	Arthropoda	XR_045851.2	NF	747
<i>Aedes aegypti</i>	Yellow fever mosquito	Arthropoda	XP_001653313.1	Q16YD4	835
<i>Amphimedon queenslandica</i>	Demosponge	Porifera	Not found	I1FL37	670
<i>Apis florea</i>	Dwarf honey bee	Arthropoda	XM_003692795.1	NF	690
<i>Apis mellifera</i>	European honey bee	Arthropoda	XM_006569644.1	H9KLG2	651
<i>Aplysia californica</i>	California Sea Hare	Mollusca	XM_005102037.1	NF	634
<i>Ascaris suum</i>	Pig roundworm	Nematoda	ERG85993.1	U1MP83	1122
<i>Bombus impatiens</i>	Common Eastern Bumble Bee	Arthropoda	XM_003484368.1	NF	747
<i>Bombus terrestris</i>	Buff-tailed Bumblebee	Arthropoda	XM_003403279.1	NF	745
<i>Bombyx mori</i>	Silk Moth	Arthropoda	XM_004931509.1	H9JQN6	747
<i>Branchiostoma floridae</i>	Florida lancelet	Cephalochordata	XM_002613537.1	C3XQZ7	581
<i>Brugia malayi</i>	Roundworm	Nematoda	XP_001899529.1	A8Q0U5	1208
<i>Bursephelenchus xylophilus</i>	Roundworm	Nematoda	NF	NF	710
<i>Camponotus floridanus</i>	Florida carpenter ant	Arthropoda	EFN60214.1	E2B288	848
<i>Capitella teleta</i>	Polychaete worm	Annelida	ELU03450.1	R7UAU1	717
<i>Ciona intestinalis</i>	Transparent sea squirt	Urochordata	XM_002129088.2	F6UZQ2	741
<i>Ciona savignyi</i>	Transparent sea squirt	Urochordata	NF	H2YZ91	600
<i>Clonorchis sinensis</i>	Chinese liver fluke	Platyhelminthes	GAA50274.1	G7YBE0	1086
<i>Crassostrea gigas</i>	Pacific oyster	Mollusca	EKC34399.1	K1QT27	607
<i>Culex quinquefasciatus</i>	Southern house mosquito	Arthropoda	XP_002613583.1	B0X2T6	294
<i>Danaus plexippus</i>	Monarch butterfly	Arthropoda	XP_001653313.1	G6DEL2	747
<i>Daphnia pulex</i>	Common water flea	Arthropoda	EFX88992.1	E9FW32	662

Appendix Table 5.1B continued

Taxa	Common name	Phylum	NCBI reference	Uniprot	aa
<i>Dendroctonus ponderosae</i>	Mountain pine beetle	Arthropoda	ENN76057.1	N6UBL7	689
<i>Dirofilaria immitis</i>	Heartworm	Nematoda	NF	NF	1203
<i>Drosophila ananassae</i>	Fruit fly	Arthropoda	XM_001958239.1	B3M8T0	936
<i>Drosophila erecta</i>	Fruit fly	Arthropoda	XP_001973404.1	B3NE06	940
<i>Drosophila grimshawi</i>	Fruit fly	Arthropoda	XP_001984226.1	B4IYC4	913
<i>Drosophila melanogaster</i>	Fruit fly	Arthropoda	NP_730465.1	Q9NJD9	900
<i>Drosophila persimilis</i>	Fruit fly	Arthropoda	XP_002024211.1	B4H032	911
<i>Drosophila pseudoobscura</i>	Fruit fly	Arthropoda	XP_001354015.2	Q2LYV8	958
<i>Drosophila simulans</i>	Fruit fly	Arthropoda	XP_002043635.1	B4IIS5	942
<i>Drosophila virilis</i>	Fruit fly	Arthropoda	XP_002085559.1	B4QRD1	783
<i>Drosophila willistoni</i>	Fruit fly	Arthropoda	XP_002061847.1	B4MLH7	1043
<i>Drosophila yakuba</i>	Fruit fly	Arthropoda	XP_002095547.1	B4PFW2	894
<i>Echinococcus granulosus</i>	Dog tapeworm	Platyhelminthes	CDJ24181.1	U6JHB6	877
<i>Echinococcus multilocularis</i>	Tape worm	Platyhelminthes	NF	U6HQR2	896
<i>Fusarium oxysporum</i>	Fungi		EMT62392.1	N1RDD5	791
<i>Harpegnathos saltator</i>	Jerdon's jumping ant	Arthropoda	EFN85407.1	E2BG05	882
<i>Hydra vulgaris</i> *	Fresh-water polyp	Cnidaria	XP_002153958.2	T2MCM8	539
<i>Hymenolepis microstoma</i>	Rodent tapeworm	Platyhelminthes	CDJ15069.1	U6IY71	823
<i>Ixodes scapularis</i> *	Black-legged tick	Arthropoda	XP_002416184.1	B7QLX9	635
<i>Loa loa</i>	African eye worm	Nematoda	EJD75529.1	J0DP84	1212
<i>Lottia gigantea</i>	Owl limpet	Mollusca	ESP03845.1	V4CNA0	605
<i>Marssonina brunnea</i>	Fungi		XP_007293392.1	K1XU93	1852
<i>Megachile rotundata</i>	Alfalfa leafcutter bee	Arthropoda	XP_003706889.1	NF	857
<i>Metarhizium acridum</i>	Fungi		EFY85899.1	E9EDU0	743
<i>Metarhizium anisopliae</i>	Fungi		EFY99265.1	E9EZ74	741
<i>Mnemiopsis leidyi</i>	Sea walnut	Ctenophora	MLRB05513 MGP	NF	457
<i>Nasonia vitripennis</i>	Jewel Wasp	Arthropoda	XP_001605309.1	K7IM36	770

Appendix Table 5.1B continued

Taxa	Common name	Phylum	NCBI reference	Uniprot	aa
<i>Nematostella vectensis</i> *	Starlet sea anemone	Cnidaria	XP_001634713.1	A7S0W2	618
<i>Neurospora crassa</i>	Fungi		CAD11320.1	Q96U06	898
<i>Onchocerca volvulus</i>	River blindness nematode	Nematoda	NF	NF	1213
<i>Pediculus humanus corporis</i>	Human body louse	Arthropoda	XP_002425246.1	E0VGK2	695
<i>Rhipicephalus pulchellus</i>	Ivory-ornamented tick	Arthropoda	NF	L7MHD0	663
<i>Salpingoeca punctatus</i>	Choanoflagellate		NF	NF	380
<i>Salpingoeca rosetta</i>	Choanoflagellate		XP_004994745.1	F2U818	605
<i>Schistosoma mansoni</i>	Blood fluke	Platyhelminthes	XP_002576086.1	G4LWA0	1140
<i>Schmidtea mediterranea</i>	Planarian	Platyhelminthes	AFD29606.1	H9CXT5	673
<i>Strigamia maritima</i>	Coastal centipede	Arthropoda	NF	T1JJ78	770
<i>Strongylocentrotus purpuratus</i>	Purple Sea Urchin	Echinodermata	XP_788076.2	H3IGD6	780
<i>Tribolium castaneum</i>	Red flour beetle	Arthropoda	XM_970065.1	D6WHD3	673
<i>Trichoplax adhaerens</i>	Flat animal	Placozoa	XP_002109732.1	B3RMS3	444
<i>Wuchereria bancrofti</i>	Parasitic roundworm	Nematoda	EJW87074.1	J9FID3	1194

Appendix Table 5.1C. Taxon list for Embryonic Sex Combs (ESC) with accession numbers for NCBI and Uniprot. NF indicates no record was found in that database. Sequences for taxa in bold were obtained from WormBase (Nematoda), Mnemiopsis Genome Project (Ctenophora) or The Broad Institute (Choanoflagellate). Asterisks indicate sequences that were listed as partial in both NCBI and Uniprot. Amino acid length = aa.

Taxa	Common name	Phylum	NCBI reference	Uniprot	aa
<i>Acromyrmex echinatior</i>	Panamanian leafcutter ant	Arthropoda	EGI64071.1	F4WP90	425
<i>Acyrtosiphon pisum</i>	Pea aphid	Arthropoda	XM_001949733.2	J9K6S7	409
<i>Aedes aegypti</i>	Yellow fever mosquito	Arthropoda	XM_001648915.1	Q16G31	425
<i>Amblyomma maculatum</i>	Gulf coast tick	Arthropoda	NF	G3MT61	318
<i>Ancylostoma ceylanicum</i>	Hookworm	Nematoda	EYB97523.1	U6NPK7	466
<i>Anopheles gambiae</i> *	African malaria mosquito	Arthropoda	XM_557691.3	Q5TSA2	322
<i>Apis florea</i>	Dwarf honey bee	Arthropoda	XM_003691321.1	No record	427
<i>Apis mellifera</i>	European honey bee	Arthropoda	XM_623805.3	H9KFU0	427
<i>Aplysia californica</i>	California Sea Hare	Mollusca	XM_005111287.1	NF	310
<i>Ascaris suum</i>	Pig roundworm	Nematoda	NF	F1LG42	187
<i>Bombus impatiens</i>	Common Eastern Bumble Bee	Arthropoda	XM_003490736.1	NF	427
<i>Bombus terrestris</i>	Buff-tailed Bumblebee	Arthropoda	XM_003393707.1	NF	427
<i>Bombyx mori</i>	Silk Moth	Arthropoda	NM_001201437.1	E5RWX8	412
<i>Botryllus primigenius</i> *	Sea squirt	Urochordata	AB852575.1	T2HUX0	292
<i>Branchiostoma floridae</i>	Florida lancelet	Cephalochordata	XM_002599094.1	C3YXD4	439
<i>Brugia malayi</i>	Roundworm	Nematoda	XM_001894369.1	A8P2A2	374
<i>Bursephelenchus xylophilus</i>	Roundworm	Nematoda	NF	NF	393
<i>Caenorhabditis angaria</i>	Roundworm	Nematoda	NF	NF	480
<i>Caenorhabditis brenneri</i>	Roundworm	Nematoda	EGT34195.1	G0NMP5	429
<i>Caenorhabditis briggsae</i>	Roundworm	Nematoda	CAP33746.2	A8XM71	484
<i>Caenorhabditis elegans</i>	Roundworm	Nematoda	NP_001021320.1	Q9GYS1	459
<i>Caenorhabditis japonica</i>	Roundworm	Nematoda	NF	H2VY07	465
<i>Caenorhabditis remanei</i>	Roundworm	Nematoda	XP_003093535.1	E3NDJ1	470

Appendix Table 5.1C continued

Taxa	Common name	Phylum	NCBI reference	Uniprot	aa
<i>Camponotus floridanus</i>	Florida carpenter ant	Arthropoda	EFN68408.1	E2ADP6	425
<i>Capitella teleta</i>	Polychaete worm	Annelida	ELU03609.1	R7UC89	376
<i>Ciona intestinalis</i>	Transparent sea squirt	Urochordata	XM_002128576.2	F6UZU2	424
<i>Ciona savignyi</i>	Transparent sea squirt	Urochordata	NF	H2YT04	388
<i>Clonorchis sinensis</i>	Chinese liver fluke	Platyhelminthes	GAA50728.1	G7YCP4	1170
<i>Colletotrichum gloeosporioides</i>	Fungi		XP_007286258.1	L2FE45	484
<i>Cryptococcus gattii</i>	Fungi		XP_003194233.1	E6R6V5	572
<i>Cryptococcus neoformans</i>	Fungi		AFR95484.1	J9VSB1	571
<i>Culex quinquefasciatus</i>	Southern house mosquito	Arthropoda	XM_001842037.1	B0W055	422
<i>Danaus plexippus</i>	Monarch butterfly	Arthropoda	EHJ72379.1	G6D2A5	412
<i>Daphnia pulex</i>	Common water flea	Arthropoda	EFX73235.1	E9H4Z8	426
<i>Dendroctonus ponderosae</i>	Mountain pine beetle	Arthropoda	AEE62083.1	J3JV55	427
<i>Dirofilaria immitis</i>	Heartworm	Nematoda	NF	NF	405
<i>Drosophila ananassae</i>	Fruit fly	Arthropoda	XP_001965085.1	B3MMX9	466
<i>Drosophila erecta</i>	Fruit fly	Arthropoda	XM_001969774.1	B3N4F7	688
<i>Drosophila grimshawi</i>	Fruit fly	Arthropoda	XM_001993400.1	B4JR31	425
<i>Drosophila melanogaster</i>	Fruit fly	Arthropoda	NM_058083.4	Q9VKD5	425
<i>Drosophila mojavensis</i>	Fruit fly	Arthropoda	XM_002002983.1	B4KIV5	426
<i>Drosophila pseudoobscura</i>	Fruit fly	Arthropoda	XM_001356924.2	Q29LL9	424
<i>Drosophila sechellia</i>	Fruit fly	Arthropoda	XM_002041944.1	B4IE20	425
<i>Drosophila simulans</i>	Fruit fly	Arthropoda	XM_002079110.1	B4Q3A0	425
<i>Drosophila willistoni</i>	Fruit fly	Arthropoda	XM_002064837.1	B4MVY9	418
<i>Drosophila yakuba</i>	Fruit fly	Arthropoda	XM_002088360.1	B4P1K1	675
<i>Echinococcus granulosus</i>	Fox tapeworm	Platyhelminthes	CDS18218	A0A068WDP7	467
<i>Echinococcus multilocularis</i>	Fox tapeworm	Platyhelminthes	CDJ01264.1	A0A068Y217	467
<i>Fusarium oxysporum</i>	Fungi		EMT66423.1	N1RIL3	527
<i>Grosmannia clavigera</i>	Fungi		EFW98560.1	F0XUH9	513

Appendix Table 5.1C continued

Taxa	Common name	Phylum	NCBI reference	Uniprot	aa
<i>Haemonchus contortus</i>	Barbers pole worm	Nematoda	CDJ82446.1	U6NPK7	446
<i>Harpegnathos saltator</i>	Jerdon's jumping ant	Arthropoda	EFN80548.1	E2BUS5	428
<i>Helicoverpa armigera</i>	Corn earworm	Arthropoda	JQ744271.1	R4IT83	413
<i>Helobdella robusta</i>	Californian leech	Mollusca	ESO12851.1	T1EDW3	427
<i>Hydra vulgaris</i>	Fresh-water polyp	Cnidarian	AAR06604.1	Q69DT2	421
<i>Hymenolepis microstoma</i>	Rodent tapeworm	Nematoda	CDJ14311.1	U6IP89	467
<i>Ixodes scapularis</i>	Black legged tick	Arthropoda	XM_002413658.1	B7QDT9	444
<i>Junonia coenia</i>	Buckeye	Arthropoda	AAC05331.1	O16021	412
<i>Lepeophtheirus salmonis</i>	Salmon louse	Arthropoda	BT078622.1	C1BVJ3	428
<i>Loa loa</i>	Eye worm	Nematoda	EFO21763.2	E1G0P4	405
<i>Lottia gigantea</i>	Owl limpet	Mollusca	ESO89106.1	V4A747	429
<i>Macrobrachium nipponense</i>	Freshwater shrimp	Arthropoda	AGI50961.1	NF	355
<i>Megachile rotundata</i>	Alfalfa leafcutter bee	Arthropoda	XM_003700925.1	NF	427
Meloidogyne hapla	Root knot nematode	Nematoda	NF	NF	385
Mnemiopsis leidyi	Sea walnut	Ctenophora	NF	NF	527
<i>Monosiga brevicollis</i>	Choanoflagellate		XP_001746355.1	A9V144	304
<i>Nasonia vitripennis</i>	Jewel wasp	Arthropoda	XM_003424048.1	K7IRS7	427
<i>Nematostella vectensis</i> *	Starlet sea anemone	Cnidarian	XM_001634033.1	A7S2L1	299
<i>Oikopleura dioica</i>	Sea squirt	Urochordata	CBY32762.1	E4YB25	537
Onchocerca volvulus	African river blindness nematode	Nematoda	NF	NF	358
<i>Pediculus humanus corporis</i>	Human body louse	Arthropoda	XM_002427528.1	E0VN79	437
<i>Polyandrocarpa misakiensis</i> *	Sea squirt	Urochordata	AB617630.1	BAJ78350	276
<i>Rhipicephalus pulchellus</i>	Ivory ornamented tick	Arthropoda	NF	L7LZK8	431
<i>Rhodnius prolixus</i>	Assassin bug	Arthropoda	NF	T1HXR8	422
<i>Saccoglossus kowalevskii</i>	Acorn worm	Hemichordata	XM_006824055.1	NF	451
Salpingoeca punctatus	Choanoflagellate		NF	NF	914
<i>Salpingoeca rosetta</i>	Choanoflagellate		XP_004996458.1	F2U3T1	253

Appendix Table 5.1C continued

Taxa	Common name	Phylum	NCBI reference	Uniprot	aa
<i>Schistocerca americana</i>	American grasshopper	Arthropoda	AF003604.1	O16022	437
<i>Schistosoma mansoni</i>	Trematode flatworm	Platyhelminthes	XM_002579057.1	G4VT41	507
<i>Schmidtea mediterranea</i>	Freshwater planarian	Platyhelminthes	JQ425136.1	H9CXT3	466
<i>Solenopsis invicta</i> *	Red fire ant	Arthropoda	EFZ16293.1	E9ISN8	425
<i>Strigamia maritima</i>	Coastal centipede	Arthropoda	NF	T1J4Y1	434
<i>Strongylocentrotus purpuratus</i>	Purple sea urchin	Echinodermata	XM_781252.3	H3JHM5	461
<i>Suberites domuncula</i> *	Sponge	Porifera	AM084418.1	Q0KHA0	344
<i>Tribolium castaneum</i>	Red flour beetle	Arthropoda	XM_968687.2	D6WCC2	423
<i>Trichoplax adhaerens</i> *	Flat animal	Placazoa	XM_002116910.1	B3S9R9	353

**All hymenopteran sequences were identified in BLAST as ESC-L, which is functionally similar to ESC but exhibits temporal differences in expression (Rai et al., 2013).

Appendix Table 5.2A. Domain sequence divergence for Enhancer of zeste. Taxa recovered with E-value > 1E-4 in BLAST pairwise alignments with SANT1 and SANT2 domains of *Drosophila melanogaster*. NS indicates that no sequence similarity was found.

SANT1	Description	Identity	Query cover	E value	Phylum
	<i>Oikopleura dioica</i>	39%	34%	7.00E-04	Urochordate
	<i>Trichinella spiralis</i>	36%	42%	0.004	Nematoda
	<i>Loa loa</i>	37%	76%	0.006	Nematoda
	<i>Onchocerca volvulus</i>	37%	89%	0.007	Nematoda
	<i>Dirofilaria immitis</i>	37%	68%	0.007	Nematoda
	<i>Brugia malayi</i>	37%	63%	0.009	Nematoda
	<i>Cryptococcus gattii</i>	45%	14%	0.019	Fungi
	<i>Cryptococcus neoformans</i>	45%	39%	0.024	Fungi
	<i>Nematostella vectensis</i>	30%	75%	0.038	Cnidarian
	<i>Ascaris suum</i>	21%	59%	0.1	Nematoda
	<i>Trichoplax adhaerens</i>	42%	13%	0.11	Nematoda
	<i>Ancylostoma ceylanicum</i>	35%	72%	0.16	Nematoda
	<i>Clonorchis sinensis</i>	21%	95%	0.25	Nematoda
	<i>Haemonchus contortus</i>	35%	51%	0.26	Nematoda
	<i>Monosiga brevicollis</i>	45%	37%	0.3	Choanoflagellate
	<i>Echinococcus multilocularis</i>	28%	22%	0.33	Platyhelminthes
	<i>Caenorhabditis remanei</i>	26%	65%	0.49	Nematoda
	<i>Salpingoeca rosetta</i>	28%	42%	0.57	Choanoflagellate
	<i>Helobdella robusta</i>	21%	56%	0.92	Mollusca
	<i>Caenorhabditis briggsae</i>	86%	21%	0.93	Nematoda
	<i>Caenorhabditis angaria</i>	30%	16%	0.97	Nematoda
	<i>Caenorhabditis japonica</i>	29%	19%	1.1	Nematoda

Appendix Table 5.2A continued

SANT1	Description	Identity	Query cover	E value	Phylum
	<i>Schistosoma mansoni</i>	55%	18%	1.8	Platyhelminthe
	<i>Caenorhabditis elegans</i>	42%	14%	2.4	Nematoda
	<i>Metarhizium anisopliae</i>	38%	39%	2.8	Fungi
	<i>Metarhizium acridum</i>	38%	20%	2.8	Fungi
	<i>Fusarium oxysporum</i>	26%	27%	4.5	Fungi
	<i>Marssonina brunnea</i>	27%	18%	5.9	Fungi
	<i>Amphimedon queenslandica</i>	NS			Porifera
	<i>Bombyx mori</i>	NS			Arthropoda
	<i>Bursephelenchus xylophilus</i>	NS			Nematoda
	<i>Caenorhabditis brenneri</i>	NS			Nematoda
	<i>Mnemiopsis leidyi</i>	NS			Ctenophora
SANT2	<i>Caenorhabditis angaria</i>	32%	42%	0.002	Nematoda
	<i>Ancylostoma ceylanicum</i>	33%	71%	0.097	Nematoda
	<i>Metarhizium anisopliae</i>	29%	73%	0.21	Nematoda
	<i>Caenorhabditis remanei</i>	29%	38%	0.23	Nematoda
	<i>Caenorhabditis briggsae</i>	25%	48%	0.41	Nematoda
	<i>Cryptococcus gattii</i>	60%	10%	0.64	Fungi
	<i>Cryptococcus neoformans</i>	60%	10%	0.65	Fungi
	<i>Metarhizium acridum</i>	26%	63%	0.99	Fungi
	<i>Fusarium oxysporum</i>	38%	42%	1.1	Fungi
	<i>Caenorhabditis brenneri</i>	44%	32%	1.3	Nematoda
	<i>Caenorhabditis japonica</i>	29%	42%	1.6	Nematoda
	<i>Amphimedon queenslandica</i>	NS			Porifera

Appendix Table 5A continued

SANT2	Description	Identity	Query cover	E value	Phylum
	<i>Caenorhabditis elegans</i>	NS			Nematoda
	<i>Haemonchus contortus</i>	NS			Nematoda
	<i>Marssonina brunnea</i>	NS			Nematoda
	<i>Salpingoeca rosetta</i>	NS			Choanoflagellate
	<i>Strigamia maritima</i>	NS			Arthropoda

* CXC and SET domain were highly conserved (E-value >2E-5 and 2E-15 respectively) with only *Strigamia maritima* recovered as having no sequence similarity with *Drosophila melanogaster* due to a partial sequence.

Appendix Table 5.2B. Domain sequence divergence for Suz12. Taxa recovered with E-value > 1E-4 in BLAST pairwise alignments with the Zinc finger domain of *Drosophila melanogaster*. NS indicates that no sequence similarity was found.

Taxon list	Identity	Query cover	E value	Taxonomic ID
<i>Schmidtea mediterranea</i>	43%	87%	3E-04	Platyhelminthe
<i>Marssonina brunnea</i> **	44%	87%	0.009	Fungi
<i>Ascaris suum</i>	63%	41%	0.029	Nematoda
<i>Neurospora crassa</i>	100%	16%	0.16	Fungi
<i>Culex quinquefasciatus</i>	83%	25%	0.19	Arthropoda
<i>Salpingoeca punctatus</i>	100%	16%	0.25	Choanoflagellate
<i>Fusarium oxysporum</i> **	50%	41%	0.71	Fungi
<i>Meloidogyne hapla</i>	100%	16%	5.7	Nematoda
<i>Metarhizium acridum</i> **	67%	20%	12	Fungi
<i>Bursaphelenchus xylophilus</i>	100%	12%	16	Nematoda
<i>Metarhizium anisopliae</i> **	100%	8%	16	Fungi
<i>Brugia malayi</i>	57%	37%	19	Nematoda
<i>Dirofilaria immitis</i>	50%	33%	19	Nematoda
<i>Loa loa</i>	50%	37%	19	Nematoda
<i>Onchocerca volvulus</i>	50%	33%	19	Nematoda
<i>Wuchereria bancrofti</i>	50%	37%	19	Nematoda
<i>Drosophila simulans</i>	NS			Arthropoda
<i>Trichoplax adhaerens</i>	NS			Placozoa

* No Suz12 homolog was found in the genus *Caenorhabditis*. Homologs for the VEFS box were identified in SMART and pairwise alignments with *D. melanogaster* for all 65 taxa. **Taxa identified with a Zinc finger in SMART.

Appendix Table 5.2C. Domain sequence divergence for ESC. Taxa recovered with E-value > 1E-4 in BLAST pairwise alignments with the Zinc finger domain of *Drosophila melanogaster*. NS indicates that no sequence similarity was found. *Potential partial sequences in these repeat regions.

Repeat	Taxa	Identity	Query cover	E value	Taxonomic ID
WD1	<i>Cryptococcus gattii</i>	33%	84%	7.00E-04	Fungi
	<i>Cryptococcus neoformans</i>	33%	84%	0.001	Fungi
	<i>Caenorhabditis briggsae</i>	31%	71%	0.007	Nematoda
	<i>Grosmannia clavigera</i>	31%	82%	0.014	Fungi
	<i>Fusarium oxysporum</i>	22%	82%	0.32	Fungi
	<i>Colletotrichum gloeosporioides</i>	33%	61%	0.79	Fungi
	<i>Nematostella vectensis</i> *	40%	25%	6.8	Cnidarian
	<i>Amblyomma maculatum</i> *	NS			Mollusca
	<i>Ascaris suum</i> *	NS			Nematoda
	<i>Botryllus primigenius</i> *	NS			Urochordata
	<i>Caenorhabditis remanei</i>	NS			Nematoda
	<i>Monosiga brevicollis</i> *	NS			Choanoflagellate
	<i>Polyandrocarpa misakiensis</i> *	NS			Urochordata
	<i>Salpingoeca punctatus</i>	NS			Choanoflagellate
<i>Suberites domuncula</i> *	NS			Porifera	
WD2	<i>Grosmannia clavigera</i>	26%	91%	4.00E-04	Fungi
	<i>Fusarium oxysporum</i>	25%	95%	6.00E-04	Fungi
	<i>Caenorhabditis briggsae</i>	37%	82%	0.001	Nematoda
	<i>Polyandrocarpa misakiensis</i> *	70%	34%	0.012	Urochordata
	<i>Caenorhabditis japonica</i> *	43%	30%	0.69	Nematoda
	<i>Ascaris suum</i> *	NS			Nematoda

Appendix Table 5.2C continued

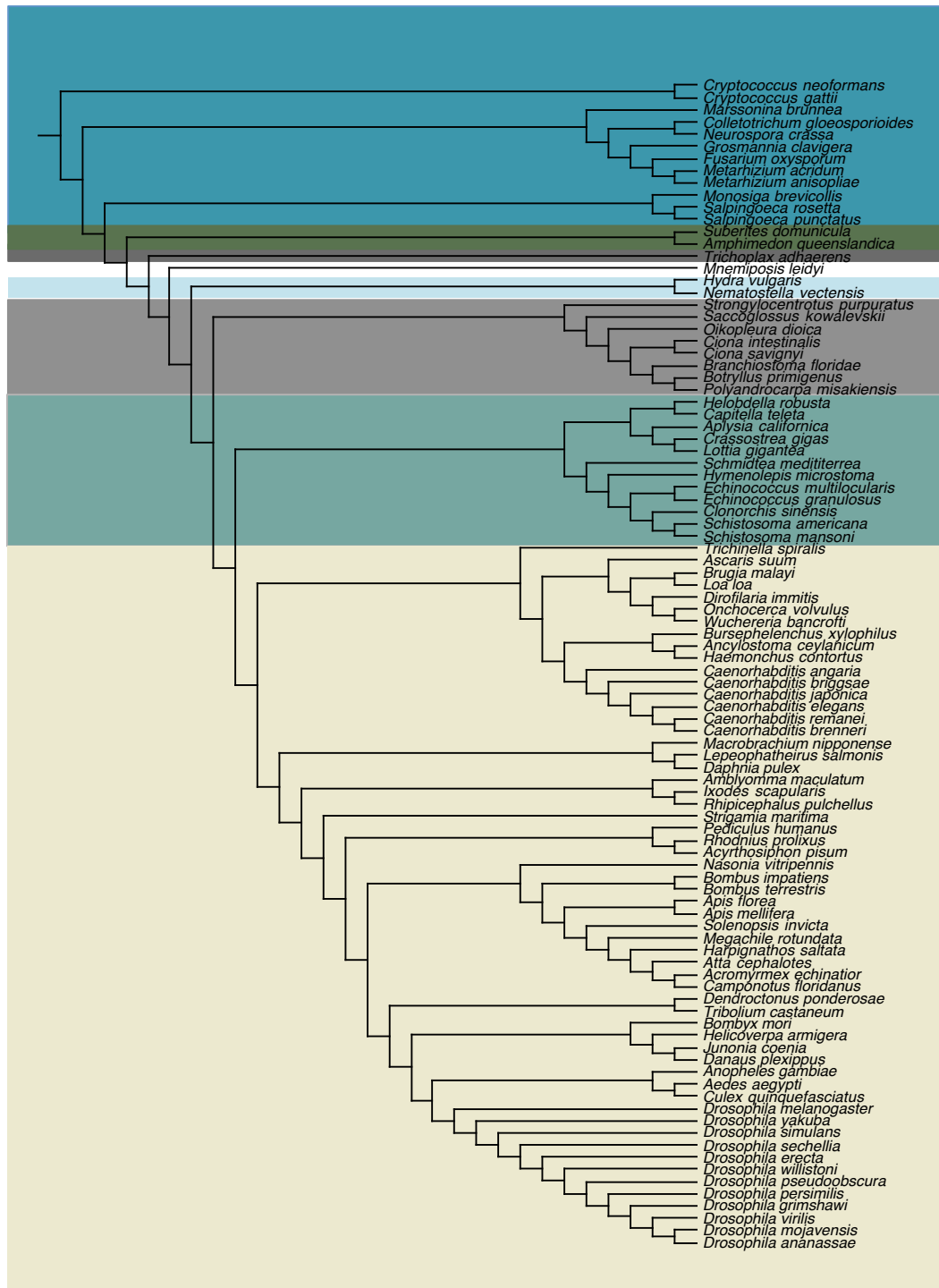
Repeat	Taxa	Identity	Query cover	E value	Taxonomic ID
WD2	<i>Botryllus primigenius</i> *	NS			Urochordata
	<i>Caenorhabditis angaria</i>	NS			Nematoda
	<i>Caenorhabditis brenneri</i>	NS			Nematoda
	<i>Caenorhabditis elegans</i>	NS			Nematoda
	<i>Caenorhabditis remanei</i>	NS			Nematoda
	<i>Monosiga brevicollis</i> *	NS			Choanoflagellate
	<i>Nematostella vectensis</i> *	NS			Cnidaria
	<i>Suberites domuncula</i> *	NS			Porifera
WD3	<i>Caenorhabditis elegans</i>	32%	68%	4.00E-04	Nematoda
	<i>Caenorhabditis briggsae</i>	31%	75%	4.00E-04	Nematoda
	<i>Monosiga brevicollis</i> *	75%	21%	0.002	Choanoflagellate
	<i>Ascaris suum</i> *	57%	17%	0.008	Nematoda
	<i>Botryllus primigenius</i> *	31%	31%	0.008	Urochordata
	<i>Caenorhabditis brenneri</i>	26%	82%	0.022	Nematoda
	<i>Salpingoeca rosetta</i>	31%	39%	0.069	Choanoflagellate
	<i>Suberites domuncula</i> *	39%	43%	0.1	Porifera
	<i>Caenorhabditis remanei</i>	28%	60%	0.22	Nematoda
	<i>Colletotrichum gloeosporioides</i>	NS			Fungi
WD4	<i>Salpingoeca punctatus</i>	24%	95%	7.00E-04	Choanoflagellate
	<i>Ascaris suum</i> *	75%	90%	0.002	Nematoda
	<i>Salpingoeca rosetta</i>	31%	30%	0.46	Choanoflagellate
	<i>Suberites domuncula</i> *	NS			Porifera
WD5	<i>Caenorhabditis briggsae</i>	35%	92%	8.00E-04	Nematoda

Appendix Table 5.2C continued

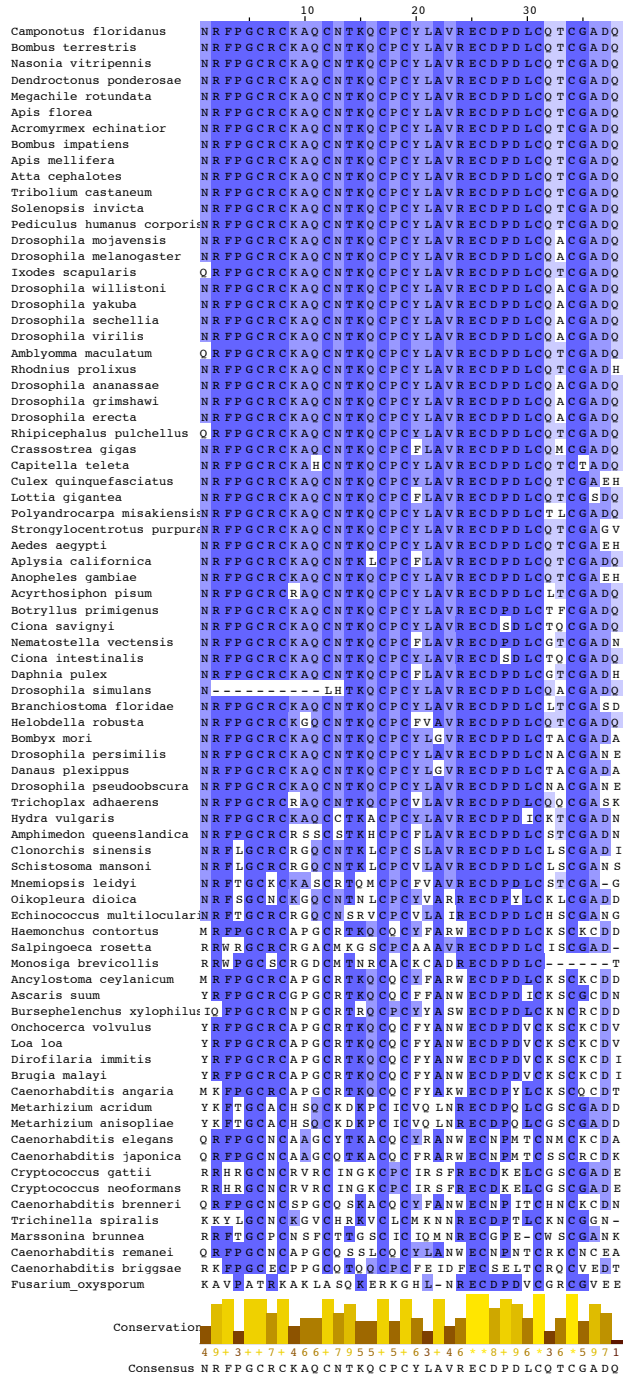
Repeat	Taxa	Identity	Query cover	E value	Taxonomic ID
WD5	<i>Salpingoeca rosetta</i>	NS			Choanoflagellate
WD6	<i>Caenorhabditis elegans</i>	26%	65%	3.00E-04	Nematoda
	<i>Colletotrichum gloeosporioides</i>	36%	97%	5.00E-04	Fungi
	<i>Grosmannia clavigera</i>	33%	100%	0.005	Fungi
	<i>Cryptococcus neoformans</i>	28%	92%	0.01	Fungi
	<i>Cryptococcus gattii</i>	28%	78%	0.01	Fungi
	<i>Fusarium oxysporum</i>	29%	95%	0.43	Fungi
	<i>Aplysia californica</i> *	38%	48%	0.64	Mollusca
	<i>Anopheles gambiae</i> *	NS			Arthropoda
	<i>Caenorhabditis briggsae</i>	NS			Nematoda
	<i>Caenorhabditis remanei</i>	NS			Nematoda
WD7	<i>Haemonchus contortus</i>	46%	61%	1.00E-04	Nematoda
	<i>Caenorhabditis elegans</i>	44%	64%	1.00E-04	Nematoda
	<i>Meloidogyne hapla</i>	42%	61%	1.00E-04	Nematoda
	<i>Ascaris suum</i>	41%	64%	3.00E-04	Nematoda
	<i>Caenorhabditis remanei</i>	36%	59%	0.005	Nematoda
	<i>Caenorhabditis japonica</i>	31%	61%	0.057	Nematoda
	<i>Grosmannia clavigera</i>	34%	66%	0.12	Fungi
	<i>Cryptococcus gattii</i>	29%	50%	0.27	Fungi
	<i>Colletotrichum gloeosporioides</i>	32%	52%	0.68	Fungi
	<i>Monosiga brevicollis</i>	38%	38%	0.76	Choanoflagellate
	<i>Hymenolepis microstoma</i>	50%	35%	0.83	Nematoda
	<i>Caenorhabditis briggsae</i>	19%	61%	1.3	Nematoda

Appendix Table 5.2C cont.

Repeat	Taxa	Identity	Query cover	E value	Taxonomic ID
WD7	<i>Anopheles gambiae</i> *	NS			Arthropoda
	<i>Aplysia californica</i> *	NS			Mollusca
	<i>Caenorhabditis angaria</i>	NS			Nematoda
	<i>Caenorhabditis brenneri</i>	NS			Nematoda
	<i>Clonorchis sinensis</i>	NS			Nematoda
	<i>Cryptococcus neoformans</i>	NS			Fungi
	<i>Fusarium oxysporum</i>	NS			Fungi
	<i>Macrobrachium nipponense</i> *	NS			Arthropoda
<i>Schistosoma mansoni</i>	NS			Platyhelminthe	



Appendix Figure 5.2 Reconstruction of the animal phylogeny based on the Tree of Life website. Major animal groups are color coded from top to bottom: Outgroups, Porifera, Placozoa, Ctenophora, Cnidaria, Deuterostomes (Echinodermata, Hemichordata, Urochordata, Cephalochordata, Lophotrochozoans and Ecydsozoans).

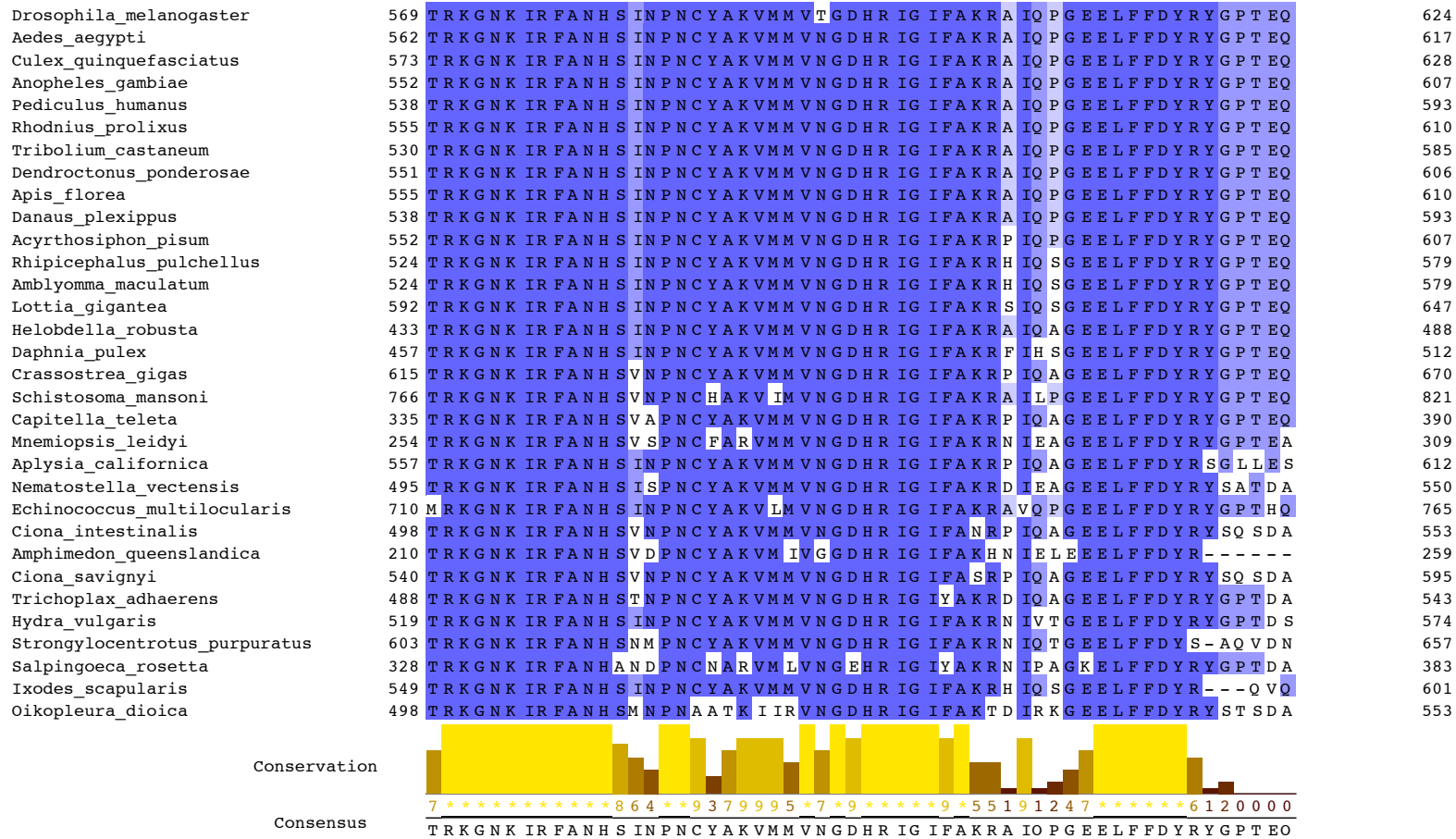


Appendix Figure 5.7 Amino acid alignment of the CXC domain from enhancer of zeste. Sequences were aligned in Seaview using Muscle and manually edited in Jalview. Percentage identity to the consensus sequence is color coded from dark blue (>80%) to light blue (>40%) and white (<40%). Taxa are sorted by pairwise comparisons to each other.

<i>Drosophila_melanogaster</i>	505	KHLLM	APSD	IAGWG	IFLK	ES	EAQ	KNEF	ISEYCGESQ	DEADRRGK	VYDKYM	C	SFLFN	LNN	D	FV	VDA	568																															
<i>Aedes_aegypti</i>	498	KHLLM	APSDVAGWG	IFLK	ES	EAQ	KNEF	ISEYCGESQ	DEADRRGK	VYDKYM	C	SFLFN	LNN	D	FV	VDA	561																																
<i>Culex_quinquefasciatus</i>	509	KHLLM	APSDVAGWG	IFLK	ES	EAQ	KNEF	ISEYCGESQ	DEADRRGK	VYDKYM	C	SFLFN	LNN	D	FV	VDA	572																																
<i>Anopheles_gambiae</i>	488	KHLLM	APSDVAGWG	IFLK	ES	EAQ	KNEF	ISEYCGESQ	DEADRRGK	VYDKYM	C	SFLFN	LNN	D	FV	VDA	551																																
<i>Pediculus_humanus</i>	474	KHLLM	APSDVAGWG	IFLK	ES	EAQ	KNEF	ISEYCGESQ	DEADRRGK	VYDKYM	C	SFLFN	LNN	D	FV	VDA	537																																
<i>Rhodnius_prolixus</i>	491	KHLLM	APSDVAGWG	IFLK	ES	EAQ	KNEF	ISEYCGESQ	DEADRRGK	VYDKYM	C	SFLFN	LNN	D	FV	VDA	554																																
<i>Tribolium_castaneum</i>	466	KHLLM	APSDVAGWG	IFLK	ES	EAQ	KNEF	ISEYCGESQ	DEADRRGK	VYDKYM	C	SFLFN	LNN	D	FV	VDA	529																																
<i>Dendroctonus_ponderosae</i>	487	KHLLM	APSDVAGWG	IFLK	ES	EAQ	KNEF	ISEYCGESQ	DEADRRGK	VYDKYM	C	SFLFN	LNN	D	FV	VDA	550																																
<i>Apis_florea</i>	491	KHLLM	APSDVAGWG	IFLK	ES	AAKNEF	ISEYCGESQ	DEADRRGK	VYDKYM	C	SFLFN	LNN	D	FV	VDA	554																																	
<i>Danaus_plexippus</i>	474	KHLLL	APSDVAGWG	IFLK	EA	AH	KNEF	ISEYCGESQ	DEADRRGK	VYDKYM	C	SFLFN	LNN	D	FV	VDA	537																																
<i>Acyrtosiphon_pisum</i>	488	KHLLM	APSDVAGWG	IFLK	ES	AQ	KNEF	ISEYCGE	TD	DEADRRGK	VYDKYM	C	SFLFN	LNH	D	FV	VDA	551																															
<i>Rhipicephalus_pulchellus</i>	460	KHLLL	APSDVAGWG	IFLK	ET	AQ	KNEF	ISEYCGESQ	DEADRRGK	VYDKYM	C	SFLFN	LNN	D	FV	VDA	523																																
<i>Amblyomma_maculatum</i>	460	KHLLL	APSDVAGWG	IFLK	DT	AQ	KNEF	ISEYCGESQ	DEADRRGK	VYDKYM	C	SFLFN	LNN	D	FV	VDA	523																																
<i>Lottia_gigantea</i>	528	KHLLL	APSDVAGWG	IFLK	EP	A	EKNEF	ISEYCGESQ	DEADRRGK	VYDKYM	C	SFLFN	LNN	D	FV	VDA	591																																
<i>Helobdella_robusta</i>	369	KHLLL	APSDVAGWG	IFLK	ES	EKNEF	ISEYCGE	TD	DEADRRGK	VYDK	RYM	C	SFLFN	LNS	D	FV	VDA	432																															
<i>Daphnia_pulex</i>	393	KHLLM	APSDVAGWG	IFLK	ET	VQ	KNEF	ISEYCGESQ	DEADRRGK	VYDKYM	C	SFLFN	LNN	D	FV	VDA	456																																
<i>Crassostrea_gigas</i>	551	KHLLL	APSD	IAGWG	IFLK	VP	AEKNEF	ISEYCGESQ	DEADRRGK	VYDKYM	C	SFLFN	LNN	D	FV	VDA	614																																
<i>Schistosoma_mansoni</i>	702	KHLLM	APSDVAGWG	IF	IK	EAA	EKND	F	IY	EYCGESQ	DEADRRGK	IYDK	TM	SS	SFLFN	LNR	D	FV	VDA	765																													
<i>Capitella_teleta</i>	271	KHLLL	APSDVAGWG	IY	VK	Q	SCDKND	F	ISEYCGESQ	DEADRRGK	VYDKYM	C	SFLFN	LNT	E	FV	VDA	334																															
<i>Mnemiopsis_leidy</i>	190	KHMLL	APSDVAGWG	IF	VK	EAC	NKND	F	L	SEYCGESQ	E	EA	DRRGK	VYDKY	K	C	SFLFN	LNTQ	DY	VVDA	253																												
<i>Aplysia_californica</i>	493	KHLLL	APSDVAGWG	IFLK	EP	A	EKNEF	ISEYCGESQ	DEADRRGK	VYDKYM	C	SFLFN	LNY	D	FV	VDA	556																																
<i>Nematostella_vectensis</i>	431	KHMLL	APSDVAGWG	IY	IK	Q	SVKNEF	ISEYCGESQ	DEADRRGK	VYDKYM	C	SFLFN	LNN	D	FV	VDA	494																																
<i>Echinococcus_multilocularis</i>	646	KHLLM	AS	SDVAGWG	IF	M	KDGA	EKNEF	IY	EYCGESQ	DEA	ERRGK	IYDR	TM	C	SFLFN	LNRQ	FA	VDA	709																													
<i>Ciona_intestinalis</i>	434	KHLLQ	APSDVAGWG	IY	IK	Q	DVNKNEF	ISEYCGESQ	DEADRRGK	VYDKYM	C	SFLFN	LNN	D	FV	VDA	497																																
<i>Amphimedon_queenslandica</i>	146	KHLLM	V	LSDVAGWG	IF	LK	DGA	EKNEF	ISEYCGESQ	DEADRRGK	VYDKYM	C	SFLFN	LNN	D	Y	VVDA	209																															
<i>Ciona_savignyi</i>	476	KHLLQ	APSDVAGWG	IY	VK	EDVN	KNEF	ISEYCGESQ	DEADRRGK	VYDKYM	C	SFLFN	LNN	D	FV	VDA	539																																
<i>Trichoplax_adhaerens</i>	424	KHLLL	APSD	IAGWG	IF	SR	YE	IHKND	F	ISEYCGESQ	DEADRRGK	VYDK	SK	C	SFLFN	LNN	E	FV	VDA	487																													
<i>Hydra_vulgaris</i>	455	MNLLL	APSD	IAGWG	IY	LK	NDVT	KNTL	I	SEYCGESQ	DEA	ERRGK	VYDK	TM	C	SFLFN	LNH	E	FV	VDA	518																												
<i>Strongylocentrotus_purpuratus</i>	539	RHLLL	APSDVAGWG	IY	IT	V	P	M	KNEF	ISEYCGESQ	DEADRRGK	VYDKYM	C	SFLFN	LNN	D	FV	VDA	602																														
<i>Salpingoeca_rosetta</i>	264	KHLLL	APSDVAGWG	IF	T	K	ED	IQ	KGA	F	ISEYCGESQ	E	EA	ERRGK	VYDK	H	M	C	SFLFN	LNAEY	VVDA	327																											
<i>Ixodes_scapularis</i>	500	K	-----	-----	-----	V	SETAQ	KNEF	ISEYCGESQ	DEADRRGK	VYDKYM	C	SFLFN	LNN	D	FV	VDA	548																															
<i>Oikopleura_dioica</i>	434	H	LLM	APSD	IAGWG	C	YAK	ND	IK	KNDY	I	S	D	Y	C	G	E	S	Q	E	E	A	D	R	R	G	K	V	Y	D	K	H	K	C	S	F	L	F	D	L	N	S	S	F	C	I	D	A	497



Appendix Figure 5.8 Amino acid alignment of the SET domain from EZ. Sequences were aligned in Seaview using Muscle and manually edited in Jalview. Percentage identity to the consensus sequence is color coded from dark blue (>80%) to light blue (>40%) and white (<40%). Taxa are sorted by pairwise comparisons to each other.

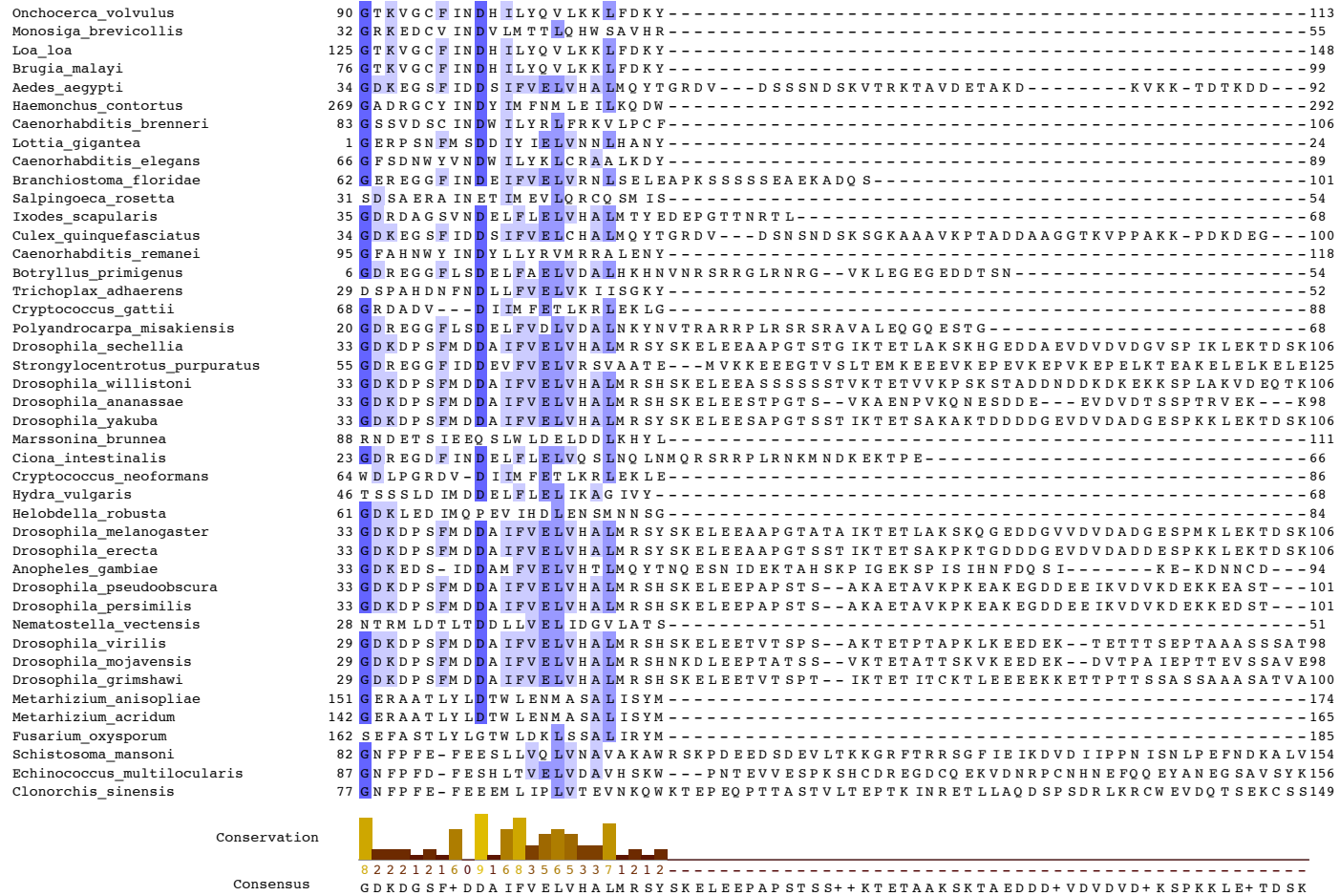


Appendix Figure 5.8 continued Amino acid alignment of the SET domain from EZ. Sequences were aligned in Seaview using Muscle and manually edited in Jalview. Percentage identity to the consensus sequence is color coded from dark blue (>80%) to light blue (>40%) and white (<40%). Taxa are sorted by pairwise comparisons to each other.

Brugia_malay	488	V	IR	FANH	SKDPNCMAKVF	VNGDHRIGIFARRP	IVAGEELFFDY	SYNSYQQ	538								
Dirofilaria_immitis	557	V	IR	FANH	SKDPNCMAKVF	VNGDHRIGIFARRP	IVAGEELFFDY	SYNSYQQ	607								
Onchocerca_volvulus	533	V	IR	FANH	SKDPNCMAKVF	VNGDHRIGIFARRP	IVAGEELFFDY	SYNSYQQ	583								
Loa_loa	568	V	IR	FANH	SKDPNCMAKVF	VNGDHRIGIFARRP	IVAGEELFFDY	SYNSYQQ	618								
Ascaris_suum	499	V	IR	FANH	SKDPNCKGRVFM	VNGDHRIGIFARRN	IAAGEELFFDY	SYNSTQQ	549								
Clonorchis_sinensis	764	K	IR	FANH	SVNPNCHAKV	-----	-----	-----	780								
Bursephelenchus_xylophilus	542	L	IR	FANHS	SNPNCYAKVVV	VNTDHRIGIFASRF	IEKGEELFFDY	AYSKNHQ	592								
Monosiga_brevicollis	471	K	IR	FANHAND	PNCARVMM	VAGEHRIGIFAPERD	IPAGREELFFNY	RYGPTDA	521								
Caenorhabditis_japonica	183	LAR	FANH	DKNPSLY	ARTMV	VAGEHRIGIFYAKRR	LEPND	ELTFDY	SYGEHQE	233							
Botryllus_primigenus	417	K	IR	FA	-----	-----	-----	-----	421								
Polyandrocarpa_misakiensis	438	K	IR	-----	-----	-----	-----	-----	440								
Bombyx_mori	248	K	IR	FANHS	INPNCYAKVMM	VNGDHRIGIFAKRA	IQPG EELFFDY	RYGPT EQ	298								
Caenorhabditis_elegans	561	LAR	FANH	SKNPTCY	ARTMV	VAGEHRIGIFYAKRR	LEISEELTFDY	SYSGEHQ	611								
Ancylostoma_ceyLANicum	227	L	IR	FANHNNN	ANCSSSEIK	IVNGEHRIGVYASRH	ILCGEELTFDY	NYGQ TWN	277								
Haemonchus_contortus	271	L	IR	FANHNNN	ANCSSSEIK	IVNGEHRIGVYASRH	ILFGEELTFDY	NYGQ TWN	321								
Caenorhabditis_brenneri	560	ISR	FANH	KKHPTVY	AKTIV	VAGELRIGFFAKRQ	LSPGDEELTFDY	SYNAIRQ	610								
Caenorhabditis_remanei	591	LSR	FANHE	INPTVNA	AKTMV	VNGEHRIGFYARRELK	ANTELTFDY	GYEKEHK	641								
Marssonina_brunnea	261	KTR	F	INHSS	ATNCEAKIM	LVNGEHR	IKFFALRD	IDAGEELLFN	YGEGLNKT	311							
Solenopsis_invicta	-----	-----	-----	-----	-----	-----	-----	-----	-----								
Trichinella_spiralis	466	V	AR	F	INH	SKIPNCV	PQVKM	VLGSHR	IAFYATR	NIEANEELFFNY	--GVLPE	514					
Metarhizium_anisopliae	200	LSRY	INH	A	SEHG	CNPRILY	VNGEYR	IKFTAMRD	IEAGEELFFNY	GENLTKK	250						
Metarhizium_acridum	191	LSRY	INH	A	SEHG	CNPRILY	VNGEYR	IKFTAMRD	IEAGEELFFNY	GENLTKK	241						
Caenorhabditis_briggsae	297	VLR	FLNN	S	DEPN	CIKYKY	VKGD	LRIGFY	TLKAMKT	QELFIN	RYRPEAA	347					
Fusarium_oxysporum	206	LSRY	INH	A	SESG	CNPRILY	VNGEYR	IKFTAMRD	IKAGEELFFNY	GENLTKK	256						
Caenorhabditis_angaria	163	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	197					
Cryptococcus_neoformans	527	HTR	F	IN	SQGN	N	CV	AHQRA	VGH	ELRILFL	TTRP	IRRH	EEIH	FN	YGD	-----	572
Cryptococcus_gattii	562	HTR	F	IN	SQEN	N	CV	AHQRA	VGH	ELRILFL	TTRP	IKRH	EEIH	FN	YGD	-----	607



Appendix Figure 5.10 Amino acid alignment of the SANT1 domain from EZ. Sequences were aligned in Seaview using Muscle and manually edited in Jalview. Percentage identity to the consensus sequence is color coded from dark blue (>80%) to light blue (>40%) and white (<40%). Taxa are sorted by pairwise comparisons to each other.



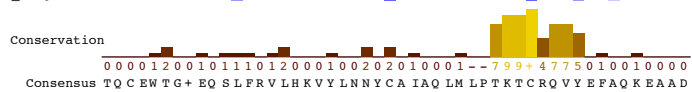
Appendix Figure 5.10 continued Amino acid alignment of the SANT1 domain from EZ. Sequences were aligned in Seaview using Muscle and manually edited in Jalview. Percentage identity to the consensus sequence is color coded from dark blue (>80%) to light blue (>40%) and white (<40%). Taxa are sorted by pairwise comparisons to each other.

Onchocerca volvulus	114	-----SD--IADQV	IYRAVYEQ	FNKASVQQ	LPFLFED	LKRR	148
Monosiga brevicollis	56	-----NGAAME	---PSAKAARV	LGHVGRSE	EIKERV	LQIMDN	89
Loa loa	149	-----SDIADRV	IYKAVYEQ	FNKASVQQ	LPFLFED	LKRR	183
Brugia malayi	100	-----SDVG IADQM	IYRAVYEQ	FNKASVQQ	LP--FED	LKRR	134
Aedes aegypti	93	-----NATD TAKEEK	LEPFVIVK	KDEKPF	PPPII	FOAISSQ	FPDHG
Haemonchus contortus	293	-----EDRSN	SEKIIYYA	IFKLFN	KLSHRQ	LV	TARGDLEER
Caenorhabditis brenneri	107	-----NDT	PDAA--	FYVAVY	CLWPN	KL	LSQRQLSFAWCDLYAE
Lottia gigantea	25	-----PD	IDDNGFKAK	LNFL	DDDDPK	P	TAVRDIIFKSI
Caenorhabditis elegans	90	-----QSPDV	---FYTL	LYRLW	PNKSSQ	REFSSAC	ENFAEK
Branchiostoma floridae	102	-----TKDNEEA	SETTDSL	KDASG	SRRRFP	CDQIF	EAISSMFPDKGSAEELREKYKELIEQ
Salpingoeca rosetta	55	-----KDDT	LFKE-V	FERL	ALR---	LGLSPQ	VIERYKNIKKS
Ixodes scapularis	69	-----TEANY	YDICARR	PLRN	L	FP	EKAPSPSDFIFAAICSVFPDKRTPBELKERYRELM EK
Culex quinquefasciatus	101	-----TV	IAAKDD	LEPFV	L	K	KDEKPFSPPII
Caenorhabditis remanei	119	-----DGN	IDV---	FY	TYCLW	PNK	FSQRQLSYIYCARYAE
Botryllus primigenus	55	-----ASTA	ETNSV	ISGD	NNE--	EL	KKTPEADEIFDAIADYFPDKGSGEDLKDKYRELTTEL
Trichoplax adhaerens	53	-----RSEK	GQNSS	ERP	V-----	PA	I
Cryptococcus gattii	89	-----IK	EEDID	PKDC	FV	ES	LDMPPPP---LAPRNINPVEGV--KLPDSSKRVVGR
Polyandrocampa misakiensis	69	-----D	SSDDAD	SDLDP	DSVK	IT	PKPTFADEIFEA
Drosophila sechellia	107	VD	LT	V	E	K	K
Strongylocentrotus purpuratus	126	L	K	T	P	---	EL
Drosophila willistoni	107	A	M	P	V	A	V
Drosophila ananassae	99	A	E	S	V	E	V
Drosophila yakuba	107	V	D	L	T	V	I
Marssonina brunnea	112	---	L	R	D	V	G
Ciona intestinalis	67	---	T	I	E	S	K
Cryptococcus neoformans	87	---	V	E	K	E	D
Hydra vulgaris	69	---	Q	E	Y	R	S
Helobdella robusta	85	---	G	N	K	K	N
Drosophila melanogaster	107	G	L	T	E	V	E
Drosophila erecta	107	V	D	P	T	V	E
Anopheles gambiae	95	---	E	N	T	C	N
Drosophila pseudoobscura	102	---	D	E	K	K	E
Drosophila persimilis	102	---	D	E	K	K	E
Nematostella vectensis	52	---	S	N	V	A	E
Drosophila virilis	99	---	E	D	K	E	A
Drosophila mojavensis	99	---	D	D	K	E	E
Drosophila grimshawi	101	A	D	D	N	D	V
Metarhizium anisopliae	175	---	A	A	R	E	P
Metarhizium acridum	166	---	A	A	R	E	P
Fusarium oxysporum	186	---	A	S	R	E	P
Schistosoma mansoni	155	N	L	H	K	I	E
Echinococcus multilocularis	157	S	K	L	S	S	---
Clonorchis sinensis	150	S	L	V	L	S	V



Appendix Figure 5.10 continued Amino acid alignment of the SANT1 domain from EZ. Sequences were aligned in Seaview using Muscle and manually edited in Jalview. Percentage identity to the consensus sequence is color coded from dark blue (>80%) to light blue (>40%) and white (<40%). Taxa are sorted by pairwise comparisons to each other.

Camponotus_floridanus S E F S W T G S E Q S L F R A L H K A F P G N P C A L A Q I M L T K T C Q E V Y E F A Q K E A S D
 Strigamia_maritima - - - - -
 Acromyrmex_echinaiator S E F S W T G S E Q S L F R A L H K A F P G N P C A L A Q I M L T K T C Q E V Y E F A Q K E A S D
 Atta_cephalotes S E F S W T G S E Q S L F R A L H K A F P G N P C A L A Q I M L T K T C Q E V Y E F A Q K E A S D
 Bombus_terrestris S E F S W T G S E Q S L F R A L H K A F P G N P C A L A Q I M L T K T C Q E V Y Q F A Q K E A S D
 Apis_florea S E F S W T G S E Q S L F R A L H K A F P G N P C A L A Q I M L T K T C Q E V Y Q F A Q K E A S D
 Apis_mellifera S E F S W T G S E Q S L F R A L H K A F P G N P C A L A Q I M L T K T C Q E V Y Q F A Q K E A S D
 Bombus_impatiens S E F S W T G S E Q S L F R A L H K A F P G N P C A L A Q I M L T K T C Q E V Y Q F A Q K E A S D
 Solenopsis_invicta S E Y S W T G S E Q S L F R A L H K A L P G N P C A L A Q I M L T K T C Q E V Y E F A Q K E A S D
 Megachile_rotundata N E F S W T G S E Q S L F R A L H K A F P G N P C A L A Q I M L T K T C Q O V Y Q F A Q K E A S D
 Nasonia_vitripennis S E L S W T G S E Q S L F R A L H K A F P G N P C A L A Q I M L T K T C K E V Y E F S L K E A S D
 Salpingoeca_rosetta - - - - - S A V H S R - - - - -
 Caenorhabditis_japonica - - - - - H E K L - - - - - R R Q A R D K G - - - - -
 Amphimedon_queenslandica - - - - - L A E M I - - - - - A S R P L Y - - - - - P G S S T
 Caenorhabditis_angaria - - - - - V G N N - - - - - Q K S R O L Y T Y S R K P S E I
 Tribolium_castaneum D H K E W T G S D E S L F R G L H K I F L N N Y C A I A Q I M L T K T C Q O V Y E F A Q K E A D
 Fusarium_oxysporum - - - - - S I L P R I S D T L - - - - - K R R R R L P Y - - - - -
 Dendroctonus_ponderosae D H N E W T G S D E S L F R A L H K V F L N N Y C A I A Q I I L T K T C Q O V Y Q F A K K E A A D
 Rhodnius_prolixus E S K E W T G S D Q S L F R A L Y K V F L R N Y C A I S Q I M L T K T C Q O V Y H F A Q K D P S V
 Pediculus_humanus_corporis Q A E E W T G S D Q S L F R A L H K I F L N N Y C A I K D C M L T K T C R O I Y E F A Q K E A E
 Danaus_plexippus I E S E W T G S D Q S L F R A L H K V F P S N Y C A I A O L M L T S K T C Q O V Y T Y W I R T G Q E
 Drosophila_grimshawi T Q C V W T G A D Q A L F R V L H K V Y L K N Y C A V A H N M L T K T C R O V Y E F A Q K E A E
 Drosophila_ananassae S Q C V W T G A D Q A L F R V L H K V Y L K N Y C A I A H N M L T K T C R O V Y E F A Q K E A E
 Drosophila_willistoni T Q C V W T G A D Q A L F R V L H K V Y L K N Y C A I A H N M L T K T C R O V Y E F A Q K E A E
 Drosophila_mojavensis T Q C V W T G A D Q A L F R V L H K V Y L K N Y C A V A H N M L T K T C R O V Y E F A Q K E A E
 Drosophila_virilis T Q C V W T G A D Q A L F R V L H K V Y L K N Y C A I A H N M L T K T C R O V Y E F A Q K E A E
 Amblyomma_maculatum L Q E E W T G A E Q S L F R V L W R V F Y G N Y C A L A T L I L T K T C A Q V Y A F A Q R E A D
 Memiopsis_leidy - - - - - W T C S E S E S L F N A L S P V F Q Q N A C S L A T V L N S K T C R O G T S L S - - - - -
 Aedes_aegypti A D E E W N G S D K S F P R T L H K V Y L N N Y C A I A E A M L M K T C Q O V Y M F A Q K E A D
 Acyrthosiphon_pisum D K Q V W T G S D Q S I F R A L R R T F L N N Y C V I A O M M L T K T C Q O V Y E F A Q N E N D E
 Drosophila_persimilis T Q C V W T G A D Q A L F R V L H K V Y L R N Y C A I A H N M L T K T C R O V Y E F A Q K D D A E
 Drosophila_melanogaster T Q C V W T G A D Q A L Y R V L H K V Y L K N Y C A I A H N M L T K T C R O V Y E F A Q K E A D
 Daphnia_pulex V Q T W T P S E Q T L F R V V H P I F L N N Y C A I A Q T I L S K T C K O V Y R F A Q Q E A D
 Drosophila_erecta T Q C V W T G A D Q A L Y R V L H K V Y L K N Y C A I A H N M L T K T C R O V Y E F A Q K E A D
 Rhipicephalus_pulchellus L Q E E W S G A E Q S L F R V L W R V F Y G N Y C A L A T L I L S K T C A Q V Y A F A Q R E A D
 Drosophila_simulans T Q C V W T G A D Q A L Y R V L H K V Y L K N Y C A I A H N M L T K T C R O V Y E F A Q K E A E
 Drosophila_yakuba T Q C V W T G A D Q A L Y R V L H K V Y L K N Y C A I A H N M L T K T C R O V Y E F A Q K E A E
 Drosophila_pseudoobscura T Q C V W T G A D Q A L F R V L H K V Y L R N Y C A I A H N M L T K T C R O V Y E F A Q K D D A E
 Drosophila_sechellia T Q C V W T G A D Q A L Y R V L H K V Y L K N Y C A I A H N M L T K T C R O V Y E F A Q K E A E
 Capitella_teleta T T K Q W T G A E E S M F R V L H E V F Y N N Y C T I A K I L K T K S C O O V Y E F S K T E A S H
 Anopheles_gambiae K D T E W N G S D K S F P R T L Q K T F L N N Y C A I A E A M L T K T C Q O V Y R F V Q Q E A A A
 Bombyx_mori G A A E W T G S D V S L H R A L H K V F P H N Y C A Q A Q L M L S K T C Q O V W R R R Q R G G Q
 Culex_quinquefasciatus A D E E W N G S D K S F P R S L Q T I Y L N N Y C A I A E A M L M K T C Q O V Y Q F A Q K E A D
 Ciona_intestinalis R E V E W D G A E S T L F R V L H E T L L T N F C A I S K M I K T K N Q O V F A F A L R E A S N
 Branchiostoma_floridae L P D G W N G A D A S L F R V L R A V Y F N N Y C S I A Q L I G T K S Q K O V Y E F A Q W E S T
 Botryllus_primigenus V T C E W D G A E S T L F R V L Q T C Y P N N F C V I A K L M K T K N Q S E V Y A F A L Q E A T R
 Lottia_gigantea T D P V W N G A E E S L F R V V H D M Y R N N Y C A I S K I I G S K N K O V Y E F A I K E E A H
 Ixodes_scapularis L Q E E W S G A E Q S L F R V L W R V F Y G N Y C A L A T L I L T K T C A Q S C M O R S S I A D
 Strongylocentrotus_purpurus S D E W N G R E S S L F R V L R N V Y N N Y C A M A R L L R T K T C I Q V A E F S A K E G E A
 Nematostella_vectensis Q Q S E W S G A E A S L L R V L R T V Y F N N Y C T I A K L I E T K C K E V Y A F G E S E S -
 Helobdella_robusta T T V D W S N A E K S M F R A L V N V L Y G N Y C A I S R V L Q T R S C L O V Y E F A K L E A D F
 Aplysia_californica T T D E W T G A E E S L R V L T G V Y R S N Y C S I A K L I R T K S C K O V Y D F A V K E A A H
 Crassostrea_gigas R I D D W M P G E E S L F R V I Q D M H R T N Y C A M A K L I K T K T C O V O Y S F A L K E A A H
 Ciona_savignyi R K V E W D G A E S T L F R V L H E S L I T N F C O I S K L I K T K T C Q O V F A F A L R E T S S
 Monosiga_brevicollis K S Q A W T Q R D M S L F E V G Q S I Y G Y D Y C E L S R Y I G G K T C A O V F - - - - -
 Hydra_vulgaris N D E P W T S S D L S M F R V L I K N F P N N Y C T I A O L L Y S H T C K O I Y R H A M L E P R D
 Polyandrocampa_misakiensis L D I H R D G A E S T L Y R V L R T S Y P N N F C A I S K L I K T K S C K D V Y A F A L Q E A T R
 Metarhizium_acridum - - - - - P L E D D E R I V L R S I H S N Y T G D P C L A A E F L - - - - - N R H C D E V Y E F K S M G I S L
 Metarhizium_anisopliae - - - - - P L E D D E R I V L R S I H S T Y M G D P C L A A E F L - - - - - N R H C D E V Y E F K S M G I S L
 Clonorchis_sinensis S Q S V W T P V E I S L I Q V L A P V F Y P Q Y C T L A R L L G T K T C Q D V L S R S T S G L T N
 Echinococcus_multilocularis K E P E W I S E Q T L F E V L A P I Y F Q N Y C V L S G F I N T R T C K E V L R Y C E C K I S K
 Trichoplax_adhaerens K S D S W T G A E I S L F R V L Q P I Y V N D Y C T I A N L I Q T K N K Q O L R V L L I E V R E
 Caenorhabditis_briggsae - - - - - T K R O F S D F L N M D K G V V K N F Q T A K L F P N V S S S M F K S L L D S A S Y
 Loa_loa E D I N W T A Q Q E S M F I A L R R T Y K N D F C K L A E V L A P S K S R E L Y A Y S F R T A P -
 Dirofilaria_immitis E D I S W T P Q Q E A M F I A L R R T Y K N D F C K L A E V L A P P K S R E L Y A Y S F R T A P -
 Trichinella_spiralis V K L E W T A E K E E M L A L E S L F G S N S C I I A E C L N T P Y C S E Y Y K S S V A S T S
 Haemonchus_contortus K V S P F - - - - - N P T L M N I L V A L L A G E Q C L I T A Q M V F S K T C S E I H R L A S O L A - -
 Caenorhabditis_brenneri D M L R M S H E E G G S I V S A N L P F - - - - - D P A Q L A A E F K R C S G I Y E K C Q K L S E Q
 Schistosoma_mansoni R D C D W N P V E I S L Y Q V L A P M Y Y H L Y C T L S N L I G T K R C S E W L S Y A S A Q N H S
 Oikopleura_dioica R R K K W T Q S E S V L M R T L V E I Y Q G N L C T I S S C M G - G R S C K S I F A P D D S V S S
 Onchocerca_volvulus E N I N W T P Q Q E T M F I A L R R T Y K N D F C K L A E V L A P P K S R E L Y A Y S F R T A P -
 Brugia_malayi E D I N W T A Q Q E S M F I A L R R T Y K N D F C K L S E V L A P S K S R E L Y A Y S F R T A P -
 Ancylostoma_ceilanicum R V A P F S - - - - - N P T L M N I L V S L L A G E K C I I T A Q M A Q P R T C S E I Y R L A S Q L A Q A
 Bursaphelenchus_xylophilus K Q V S W P H E E S M M A L I K S V G I T D C C L I A R A L Q Q P K T C N D V Y E Y F R R Q S P -
 Ascaris_suum D A D S W T I G Q E S M F T V L K R N Y K S D F C K I A C L A P P K T C R O V F A Y S L R T G P -
 Cryptococcus_neoformans L E K E S E S D K Q Q L I D I L S A Y G S R L G L K D V F - - - - - N R T C A E V A S I Q D R S R G
 Caenorhabditis_remanei A L A E M A I P D G G M I A T M S D T F - - - - - D F A H K N S K I R T C R D A Y G F A E Y I T E R
 Marssonina_brunnea E N L A Y I P Q M L L V E K - - - - - D P V Y R A - - - - - C L I A F A L - - - - - D I P C W T Y Y K L E R E G L P -
 Cryptococcus_gattii E K R G W S E S D K Q Q L I D I L S A Y G S G L G L K A V F - - - - - S R T C A E I A T I Q D R S R G
 Caenorhabditis_elegans R I A K M P I E D G A L I V N I Y I P F - - - - - C E F V K K Y S K I R S C R D A Y S M A E N V S A R



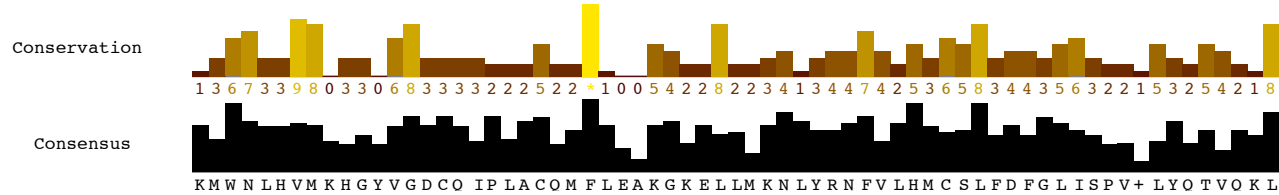
Appendix Figure 5.11 Amino acid alignment of the SANT2 domain from enhancer of zeste. Sequences were aligned in Seaview using Muscle and manually edited in Jalview. Percentage identity to the consensus sequence is color coded from dark blue (>80%) to light blue (>40%) and white (<40%). Taxa are sorted by pairwise comparisons to each other.

Mnemiopsis_leidy	10	MV	ID	IAP	----	YK	AP	HS	SR	LY	FH	SK	TY	EP	IL	PH	EFE	YD	SE	EQ	EE	SM	SW	ME	ERT	KQM	IQ	DF	TD	VNE	EAK	IM	74																																									
Drosophila_erecta	3	LE	LD	ED	E	IS	NQ	RS	Y	IT	GH	NR	LY	HHT	ET	CL	PV	HP	KE	LD	ID	SE	GE	SD	PL	WR	QR	KT	IQ	ME	IF	ES	DV	NEG	EK	EL	M	73																																				
Drosophila_yakuba	3	LE	LD	ED	E	IS	NQ	RS	Y	IT	GH	NR	LY	HHT	ET	CL	PV	HP	KE	LD	ID	SE	GE	SD	PL	WR	QR	KT	IQ	ME	IF	ES	DV	NEG	EK	EL	M	73																																				
Drosophila_melanogaster	3	LE	LD	ED	E	IS	NQ	RS	Y	IT	GH	NR	LY	HHT	ET	CL	PV	HP	KE	LD	ID	SE	GE	SD	PL	WR	QR	KT	IQ	ME	IF	ES	DV	NEG	EK	EL	M	73																																				
Drosophila_virilis	3	LE	LD	ED	D	IS	NQ	RS	Y	IT	GH	NR	LY	HHT	ET	CL	PV	HP	KE	LD	ID	SE	GE	SD	PL	WR	QR	KT	IQ	ME	IF	ES	DV	NEG	EK	EL	M	73																																				
Megachile_rotundata	2	LE	LD	ENE	E	YS	QR	RP	Y	IT	GH	NR	LY	HHT	VT	CL	P	Y	PK	EM	ID	SE	GE	ND	PK	W	LQ	TK	T	M	M	IDD	FT	DN	EG	EK	EL	M	72																																			
Drosophila_persimilis	3	LE	LD	ED	D	IS	NQ	RS	Y	IT	GH	NR	LY	HHT	ET	CL	PV	HP	KE	LD	ID	SE	GE	SD	PL	WR	QR	KT	IQ	ME	IF	ES	DV	NEG	EK	EL	M	73																																				
Loa_loa	98	FM	QA	---	---	DR	ER	RAT	K	G	N	N	A	A	Y	F	G	F	R	S	R	H	P	L	M	K	CT	Q	-	IST	-	KQ	T	DQ	EW	LR	EF	I	IR	Q	IED	FL	D	LS	RV	EK	DF	M	162																									
Aplysia_californica	2	Q	EV	ES	DA	-	G	LM	RQ	L	I	Q	G	H	N	R	L	Y	F	H	T	V	T	C	Q	P	V	R	P	Q	E	L	D	NE	G	Q	EE	SH	P	G	W	L	K	E	K	T	V	M	IED	FT	DN	EG	EK	EL	M	71																		
Dirofilaria_immitis	87	FM	QA	---	---	DR	ER	RAT	K	G	N	N	A	A	Y	F	G	F	R	S	R	H	P	L	M	K	CT	Q	-	IST	-	KQ	T	DQ	EW	LR	EF	I	IR	Q	IED	FL	D	LS	RV	EK	DF	M	151																									
Acromyrmex_echinatior	2	LE	LD	ENE	F	ES	QR	RP	Y	IT	GH	NR	LY	HHT	VT	CL	P	Y	PK	EM	ID	SE	GE	ND	PK	W	LQ	TK	T	M	M	IDD	FT	DN	EG	EK	EL	M	72																																			
Drosophila_pseudoobscura	3	LE	LD	ED	D	IS	NQ	RS	Y	IT	GH	NR	LY	HHT	ET	CL	PV	HP	KE	LD	ID	SE	GE	SD	PL	WR	QR	KT	IQ	ME	IF	ES	DV	NEG	EK	EL	M	73																																				
Brugia_malayi	80	FM	QA	---	---	DR	ER	RAT	K	G	N	N	A	A	Y	F	G	F	R	S	R	H	P	L	M	K	CT	Q	-	IST	-	KQ	T	DQ	EW	LR	EF	I	IR	A	IED	FL	D	LS	RV	EK	DF	M	144																									
Onchocerca_volvulus	85	FM	QA	---	---	DR	ER	RAT	K	G	N	N	A	A	Y	F	G	F	R	S	R	H	P	L	M	K	CT	Q	-	IST	-	KQ	T	DQ	EW	LR	EF	I	IR	Q	IED	FL	D	LS	RV	EK	DF	M	149																									
Apis_florea	15	LE	LD	ENE	E	YS	QR	RP	Y	IT	GH	NR	LY	HHT	VT	CL	P	Y	PK	EM	ID	SE	GE	ND	PK	W	LQ	TK	T	R	M	IDD	FT	DN	EG	EK	EL	M	85																																			
Apis_mellifera	2	LE	LD	ENE	E	YS	QR	RP	Y	IT	GH	NR	LY	HHT	VT	CL	P	Y	PK	EM	ID	SE	GE	ND	PK	W	LQ	TK	T	R	M	IDD	FT	DN	EG	EK	EL	M	72																																			
Acyrtosiphon_pisum	2	LE	LD	D	AD	L	DA	Q	R	P	Y	V	T	G	H	N	R	L	Y	H	H	T	I	T	C	L	P	V	Y	P	N	E	L	D	IS	E	S	E	T	D	P	L	W	L	Q	K	T	M	M	IED	FT	DN	EG	EK	EL	M	72																	
Camponotus_floridanus	2	LE	LD	ENE	F	ES	QR	RP	Y	IT	GH	NR	LY	HHT	VT	CL	P	Y	PK	EM	ID	SE	GE	ND	PK	W	LQ	TK	T	M	M	IDD	FT	DN	EG	EK	EL	M	72																																			
Bombyx_mori	2	LE	LD	D	AD	L	DA	Q	R	P	Y	V	T	G	H	N	R	L	Y	H	H	T	I	T	C	L	P	V	Y	P	N	E	L	D	IS	E	S	E	T	D	P	L	W	L	Q	K	T	M	M	IED	FT	DN	EG	EK	EL	M	72																	
Bombus_terrestris	2	LE	LD	ENE	F	ES	QR	RP	Y	IT	GH	NR	LY	HHT	VT	CL	P	Y	PK	EM	ID	SE	GE	ND	PK	W	LQ	TK	T	M	M	IDD	FT	DN	EG	EK	EL	M	72																																			
Harpegnathos_saltator	2	LE	LD	ENE	F	ES	QR	RP	Y	IT	GH	NR	LY	HHT	VT	CL	P	Y	PK	EM	ID	SE	GE	ND	PK	W	LQ	TK	T	M	M	IDD	FT	DN	EG	EK	EL	M	72																																			
Danaus_plexippus	2	LE	LD	D	D	N	D	V	D	A	Q	R	P	Y	L	T	G	H	N	R	L	Y	H	H	T	I	T	C	L	P	V	Y	P	N	E	L	D	IS	E	S	E	T	D	P	L	W	L	Q	K	T	M	M	IED	FT	DN	EG	EK	EL	M	72														
Bombus_impatiens	2	LE	LD	ENE	E	YS	QR	RP	Y	IT	GH	NR	LY	HHT	VT	CL	P	Y	PK	EM	ID	SE	GE	ND	PK	W	LQ	TK	T	M	M	IDD	FT	DN	EG	EK	EL	M	72																																			
Dendroctonus_ponderosae	2	LE	LD	D	C	E	Y	D	G	Q	R	P	F	I	T	G	H	N	R	L	Y	H	H	T	T	C	L	P	Y	PK	EM	D	IS	E	GE	ND	P	E	W	L	Q	S	K	T	M	M	IED	FT	DN	EG	EK	EL	M	72																				
Tribolium_castaneum	2	LE	LD	D	C	E	Y	D	G	Q	R	P	F	I	T	G	H	N	R	L	Y	H	H	T	T	C	L	P	Y	PK	EM	D	IS	E	GE	ND	P	E	W	L	R	N	K	T	M	M	IED	FT	DN	EG	EK	EL	M	72																				
Hymenolepis_microstoma	7	L	A	K	E	T	I	S	A	K	E	D	A	T	K	M	V	K	H	R	R	K	F	Y	H	S	S	T	L	Q	I	T	A	G	D	N	-	Y	D	S	E	E	E	S	A	K	W	L	R	D	Q	Y	R	R	V	Q	E	I	T	D	V	N	K	G	E	K	EL	M	76					
Salpingoeca_rosetta	21	A	S	T	S	A	R	G	E	L	K	A	R	L	N	V	Q	Q	R	R	-	Y	Y	S	S	P	F	C	V	E	Q	P	G	I	G	D	D	D	D	D	D	D	V	D	E	S	W	L	T	H	S	S	N	A	L	L	D	E	F	N	D	V	N	S	G	E	K	A	F	M	90			
Neurospora_crassa	23	R	S	R	A	A	A	A	A	A	P	Q	K	K	P	Y	I	P	N	I	N	R	P	I	Y	D	P	L	S	K	V	E	L	A	P	G	S	E	-	V	-	-	R	P	P	L	D	E	G	W	L	I	T	K	H	A	D	A	L	G	E	F	S	D	V	E	P	Q	E	K	E	Y	M	90
Bursephelenchus_xylophilus	2	L	H	G	A	D	V	V	-	S	W	R	G	T	L	T	R	G	A	S	P	F	A	L	N	S	K	F	N	D	V	V	F	L	C	E	L	S	T	S	Q	K	S	N	N	G	K	K	T	A	R	E	K	R	M	T	Y	F	E	Q	Y	D	H	E	N	E	V	P	N	T	N	V	71	
Metarhizium_acridum	15	P	L	R	R	P	V	K	A	S	V	S	K	V	L	V	P	N	I	P	Q	P	L	F	H	P	I	S	K	A	K	L	K	P	G	Q	E	-	V	P	Q	-	N	T	P	D	N	T	W	L	I	Q	K	H	R	E	S	I	A	D	F	S	D	V	T	A	A	E	K	E	Y	I	83	
Metarhizium_anisopliae	15	P	L	R	R	P	V	K	A	S	A	S	K	V	L	V	P	N	I	P	Q	P	L	F	H	P	I	S	K	A	K	L	K	P	G	Q	D	-	V	P	Q	-	N	T	P	D	I	T	W	L	I	Q	K	H	R	E	S	I	A	D	F	S	D	V	T	A	A	E	K	E	Y	I	83	
Fusarium_oxysporum	20	P	A	A	V	Q	A	V	I	E	K	K	K	V	I	I	P	E	S	C	P	L	F	D	P	V	S	K	A	R	L	K	P	G	E	E	-	L	P	K	-	P	V	V	D	N	A	W	L	L	Q	K	H	R	E	D	I	G	E	F	S	D	V	S	K	E	E	Y	I	88				

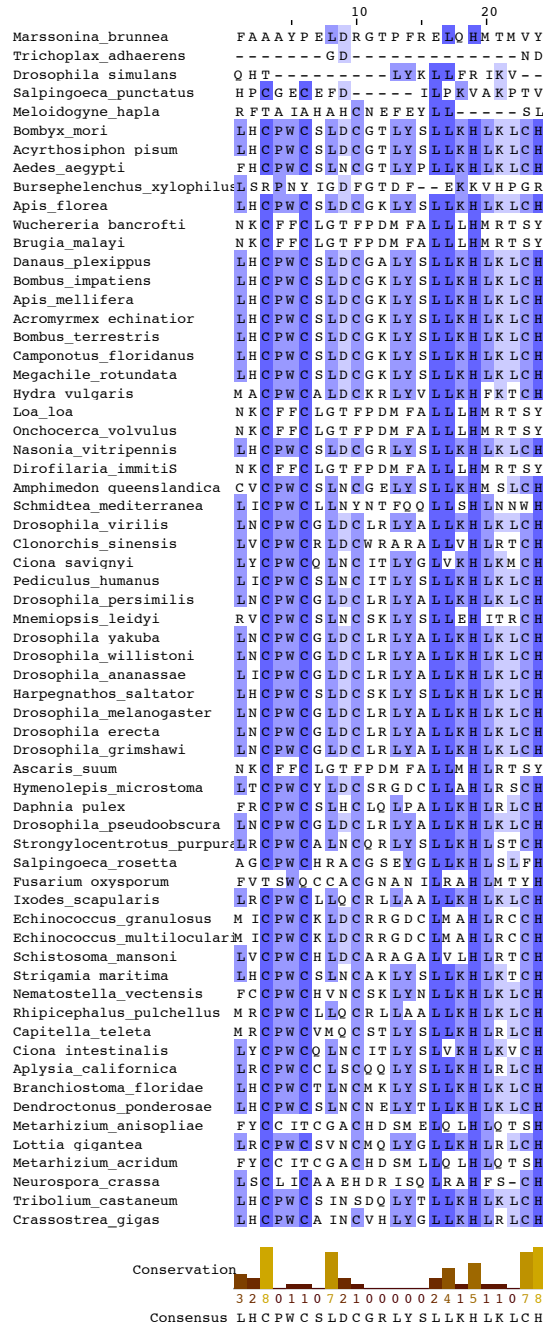


Appendix Figure 5.14 continued Amino acid alignment of the VEFs domain from Suz12. Sequences were aligned in Seaview using Muscle and manually edited in Jalview. Percentage identity to the consensus sequence is color coded from dark blue (>80%) to light blue (>40%) and white (<40%). Taxa are sorted by pairwise comparisons to each other.

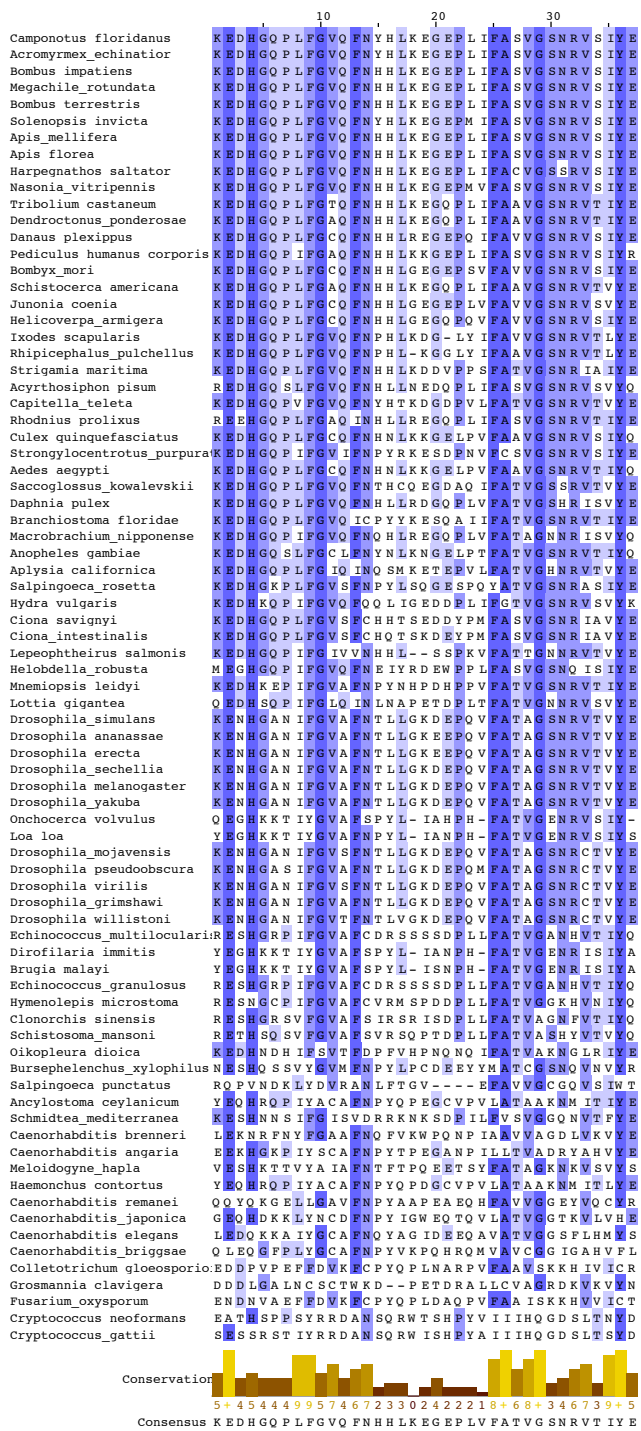
Mnemiopsis_leidy	75	IMWNCFIL- PKPLADHQV PGLARE	FV	EQ	KG	GE	IA	QG	LV	KN	IL	HL	LT	NL	HS	FG	LL	KP	EE	LE	SI	MV	GF	140																					
Drosophila_erecta	74	KLWN LHVM RHG FVGDCQ LP IAC EM	FL	DA	KG	TE	IV	RK	NY	RN	FI	LH	MC	SL	FD	YGL	IA	AE	TV	YK	TV	QK	KL	140																					
Drosophila_yakuba	74	KLWN LHVM RHG FVGDCQ LP IAC EM	FL	DA	KG	TE	IV	RK	NY	RN	FI	LH	MC	SL	FD	YGL	IA	AE	TV	YK	TV	QK	KL	140																					
Drosophila_melanogaster	74	KLWN LHVM RHG FVGDCQ LP IAC EM	FL	DA	KG	TE	IV	RK	NY	RN	FI	LH	MC	SL	FD	YGL	IA	AE	TV	YK	TV	QK	KL	140																					
Drosophila_virilis	74	KLWN LHVM RHG FVGDCQ LP IAC EM	FL	DA	KG	HE	IV	RK	LY	RN	FI	LH	MC	SL	FD	YGL	IA	AE	TV	YK	TV	QK	KL	140																					
Megachile_rotundata	73	KMWN LHVM KHGY VGDCQ IPLACHM	FL	ET	KG	KE	EL	LM	KN	LY	RN	FV	LH	MC	SL	FD	FGL	IS	PV	VV	LY	QT	IQ	KL	139																				
Drosophila_persimilis	74	KIWN LHVM RHG FVGDCQ LP IAC EM	FL	DA	KG	HE	IV	RK	LY	RN	FI	LH	MC	SL	FD	YGL	IS	NE	HV	YK	TV	QK	KL	140																					
Loa_loa	163	SLWS IFLLNHHKPLGRCHTYRTRCRL	FL	QK	HR	QD	IL	QN	FQ	NA	VV	FH	LT	AL	YE	KE	EA	LD	TE	EV	YD	LT	MR	229																					
Aplysia_californica	72	KMWN LHVM KK EY IGDCM VPRACHT	FV	EE	QA	VV	QK	IL	CR	NF	LL	HL	LV	NL	ND	FG	LIR	PE	IV	QNT	FAM	L	138																						
Dirofilaria_immitis	152	SLWS IFLLNHHKPLGRCHTYRTRCRL	FL	QK	HR	QD	IL	QN	FQ	NA	VV	FH	LT	AL	HE	KE	EA	LD	TE	EV	YD	LT	MR	218																					
Acromyrmex_echinaior	73	KMWN LHVM KHGY VGDCQ IPLACQM	FL	EN	KG	KE	EL	LM	KN	LY	RN	FV	LH	MC	SL	FD	FGL	IS	PV	IL	YQ	TIQ	KL	139																					
Drosophila_pseudoobscura	74	KIWN LHVM RHG FVGDCQ LP IAC EM	FL	DA	KG	HE	IV	RK	LY	RN	FI	LH	MC	SL	FD	YGL	IS	NE	HV	YK	TV	QK	KL	140																					
Brugia_malayi	145	SLWS IFLLNHHKPLGRCHTYRTRCRL	FL	QK	HR	QD	IL	QN	FQ	NA	VV	FH	LT	AL	HE	KE	EA	LD	TE	EV	YD	LT	MR	211																					
Onchocerca_volvulus	150	SLWS IFLLNHHKPLGRCHTYRTRCRL	FL	QK	HR	QD	IL	QN	FQ	NA	VV	FH	LT	AL	HE	KE	EA	LD	TE	EV	YD	LT	MR	216																					
Apis_florea	86	KMWN LHVM KYGY VGDCQ IPLACQM	FL	ET	KG	KE	EL	LM	KN	LY	RN	FV	LH	MC	SL	FD	FGL	IS	PV	VV	LY	QT	IQ	KL	152																				
Apis_mellifera	73	KMWN LHVM KYGY VGDCQ IPLACQM	FL	ET	KG	KE	EL	LM	KN	LY	RN	FV	LH	MC	SL	FD	FGL	IS	PV	VV	LY	QT	IQ	KL	139																				
Acyrtosiphon_pisum	73	KMWN LHVM KYNY VGDCQ IPLACQM	FL	QMR	KG	KE	EL	LE	EKN	LY	RN	FI	LH	MC	SL	HD	FGL	IS	AV	AL	YQ	TV	QM	139																					
Camponotus_floridanus	73	KMWN LHVM KYGY VGDCQ IPLACQM	FL	ES	KG	KE	EL	LM	KN	LY	RN	FV	LH	MC	SL	FD	FGL	IS	PV	IL	YQ	TIQ	KL	139																					
Bombyx_mori	73	KMWN LHVM KYNY VGDCQ IPLACQM	FL	QMR	KG	KE	EL	LE	EKN	LY	RN	FI	LH	MC	SL	HD	FGL	IS	AV	AL	YQ	TV	QM	139																					
Bombus_terrestris	73	KMWN LHVM KYGY VGDCQ IPLACQM	FL	ET	KG	KE	EL	LM	KN	LY	RN	FV	LH	MC	SL	FD	FGL	IS	PV	VV	LY	QT	IQ	KL	139																				
Harpegnathos_saltator	73	KMWN LHVM KYGY VGDCQ IP IACQM	FL	DN	KG	KE	EL	LM	KN	LY	RN	FV	LH	MC	SL	FD	FGL	IS	PV	IL	YQ	NI	QK	KL	139																				
Danaus_plexippus	73	KMWN LHVM KYNY VGDCQ IPLACQM	FL	QM	KG	KE	EL	LE	EKN	LY	RN	FI	LH	MC	SL	HD	FGL	IS	PA	AL	YQ	TV	QM	139																					
Bombus_impatiens	73	KMWN LHVM KYGY VGDCQ IPLACQM	FL	ET	KG	KE	EL	LM	KN	LY	RN	FV	LH	MC	SL	FD	FGL	IS	PV	VV	LY	QT	IQ	KL	139																				
Dendroctonus_ponderosae	73	KMWN LHVM KYG VGDCQ IPLACQM	FV	QK	GE	EL	LL	KN	LY	RN	FV	LH	LT	SL	FD	FGL	IS	AV	CL	Y	TS	SL	QK	139																					
Tribolium_castaneum	73	KMWN LHVM KYG FVGDCQ IPLACQM	FV	QK	GG	KE	EL	LL	KN	LY	KN	FV	LH	MG	SL	FD	FGL	IS	AV	CL	Y	TT	IQ	KL	139																				
Hymenolepis_microstoma	77	QIWN AY ILPRDV VADCN IDQ LLMG	FV	K	SY	GVR	MNN	Q	N	LR	NA	FI	IL	HL	LT	NY	DY	GI	LS	PT	SM	HR	CV	LA	143																				
Salpingoeca_rosetta	91	KLWN R SVMN FSVM ADRRVPD L LRS	FA	CR	HR	TA	IR	EG	RV	GN	EK	VH	MA	LL	VQ	FG	LIT	AK	IT	D	IT	LD	IL	157																					
Neurospora_crassa	91	MQW DAY IL-Q KLC SEQ YLPREFRN	FV	RE	KA	TW	LLE	K	RS	RA	EL	GK	H	MA	V	L	L	A	R	R	V	D	D	A	T	V	M	A	V	T	K	E	L	156											
Bursephelenchus_xylophilus	72	PVSTM SVFRFD-MGNR IDNTK IRIF	F	-	-	-	T	GA	EM	D	CL	N	E	K	GE	F	I	E	L	K	T	Q	F	N	Q	I	G	V	Y	K	E	M	K	W	L	Q	S	F	L	134					
Metarhizium_acridum	84	WEWDGY ILRQN ITSAAYFPRAWLN	FV	EE	KA	GW	L	A	QA	H	RM	E	FG	K	H	S	SV	L	L	A	R	D	V	L	D	D	E	M	Q	Q	A	F	R	Y	I	150									
Metarhizium_anisopliae	84	WEWDGY ILRQN ITSAAYFPRAWLN	FV	EE	KA	GW	L	A	QA	H	RM	E	FG	K	H	S	SV	L	L	A	R	D	V	L	D	D	E	M	Q	Q	A	F	R	Y	I	150									
Fusarium_oxysporum	89	RRWD A F I L - Q E V T S G Q Y F R R A W V K	FV	K	E	N	A	S	W	L	V	G	A	N	Y	R	M	E	Y	G	K	H	L	C	V	L	M	A	R	D	V	L	D	N	A	S	V	E	K	A	G	K	L	I	154



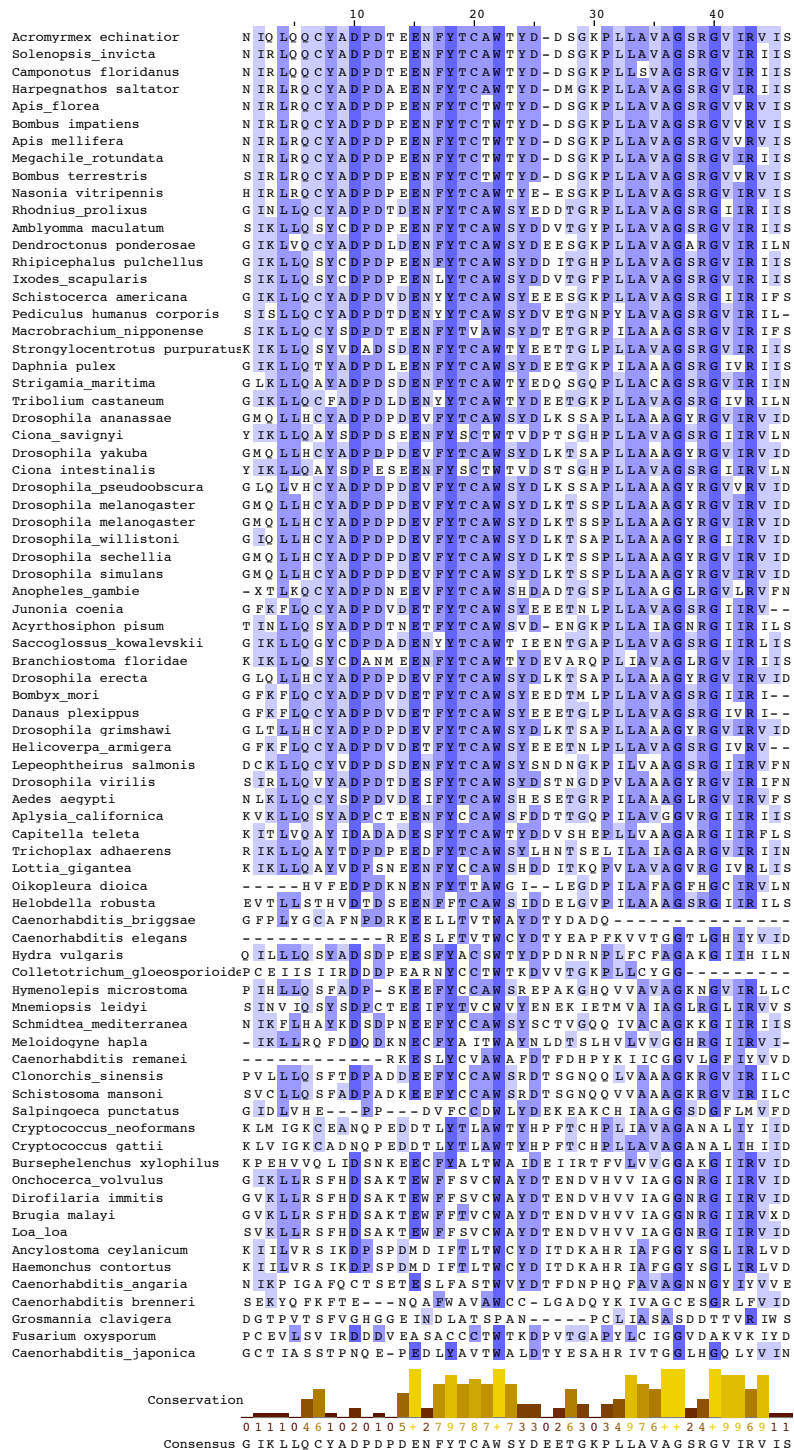
Appendix Figure 5.14 continued Amino acid alignment of the VEFS domain from Suz12. Sequences were aligned in Seaview using Muscle and manually edited in Jalview. Percentage identity to the consensus sequence is color coded from dark blue (>80%) to light blue (>40%) and white (<40%). Taxa are sorted by pairwise comparisons to each other.



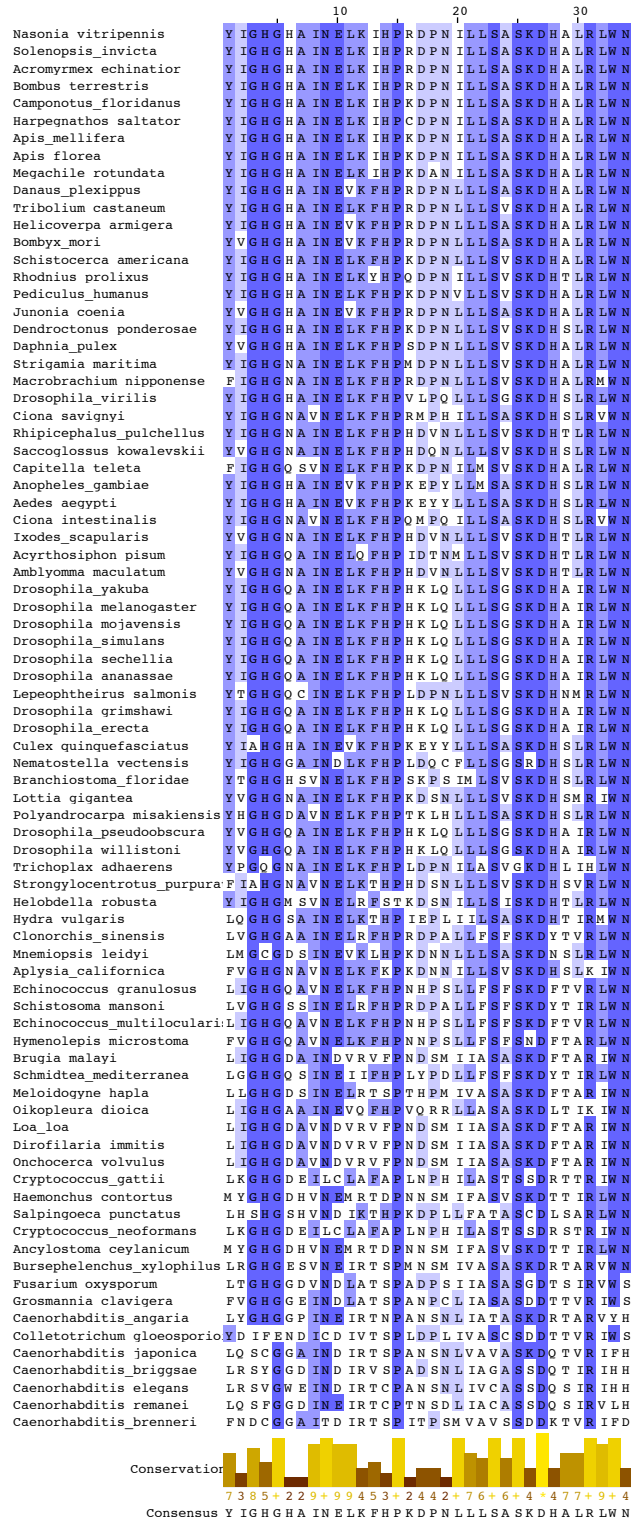
Appendix Figure 5.15 Amino acid alignment of the Zinc finger domain from Suz12. Sequences were aligned in Seaview using Muscle and manually edited in Jalview. Percentage identity to the consensus sequence is color coded from dark blue (>80%) to light blue (>40%) and white (<40%). Taxa are sorted by pairwise identity in Jalview.



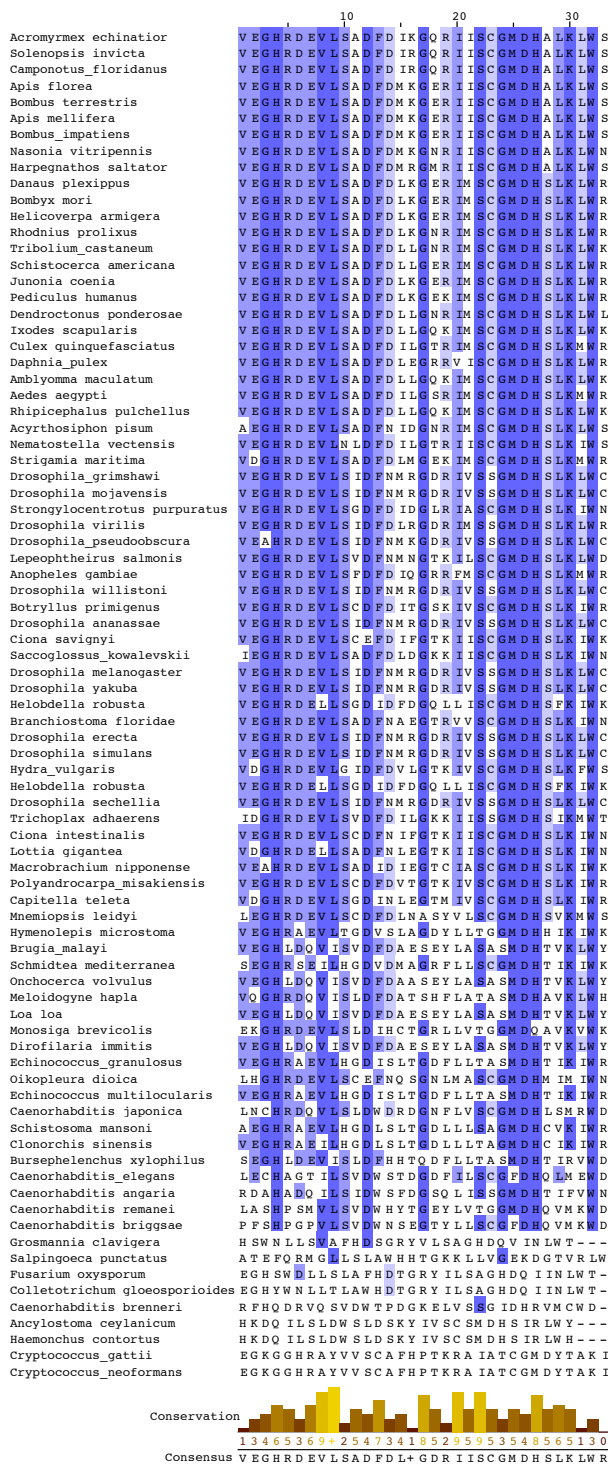
Appendix Figure 5.17 Amino acid alignment of the WD40-1 repeat from Esc. Sequences were aligned in Seaview using Muscle and manually edited in Jalview. Percentage identity to the consensus sequence is color coded from dark blue (>80%) to light blue (>40%) and white (<40%). Taxa are sorted by pairwise identity in Jalview.



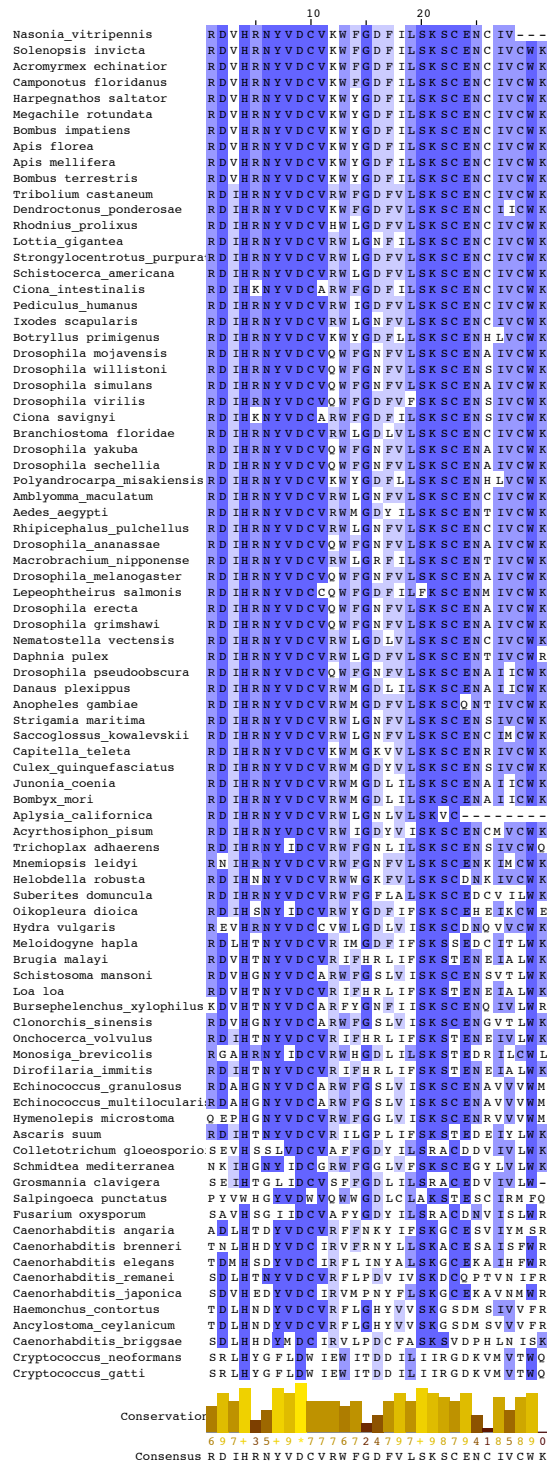
Appendix Figure 5.17 continued Amino acid alignment of the WD40-2 repeat from *Esc.* Sequences were aligned in Seaview using Muscle and manually edited in Jalview. Percentage identity to the consensus sequence is color coded from dark blue (>80%) to light blue (>40%) and white (<40%). Taxa are sorted by pairwise identity in Jalview.



Appendix Figure 5.17 continued Amino acid alignment of the WD40-3 repeat from Esc. Sequences were aligned in Seaview using Muscle and manually edited in Jalview. Percentage identity to the consensus sequence is color coded from dark blue (>80%) to light blue (>40%) and white (<40%). Taxa are sorted by pairwise identity in Jalview.



Appendix Fig. 5.17 continued Amino acid alignment of the WD40-4 repeat from Esc. Sequences were aligned in Seaview using Muscle and manually edited in Jalview. Percentage identity to the consensus sequence is color coded from dark blue (>80%) to light blue (>40%) and white (<40%). Taxa are sorted by pairwise identity in Jalview.

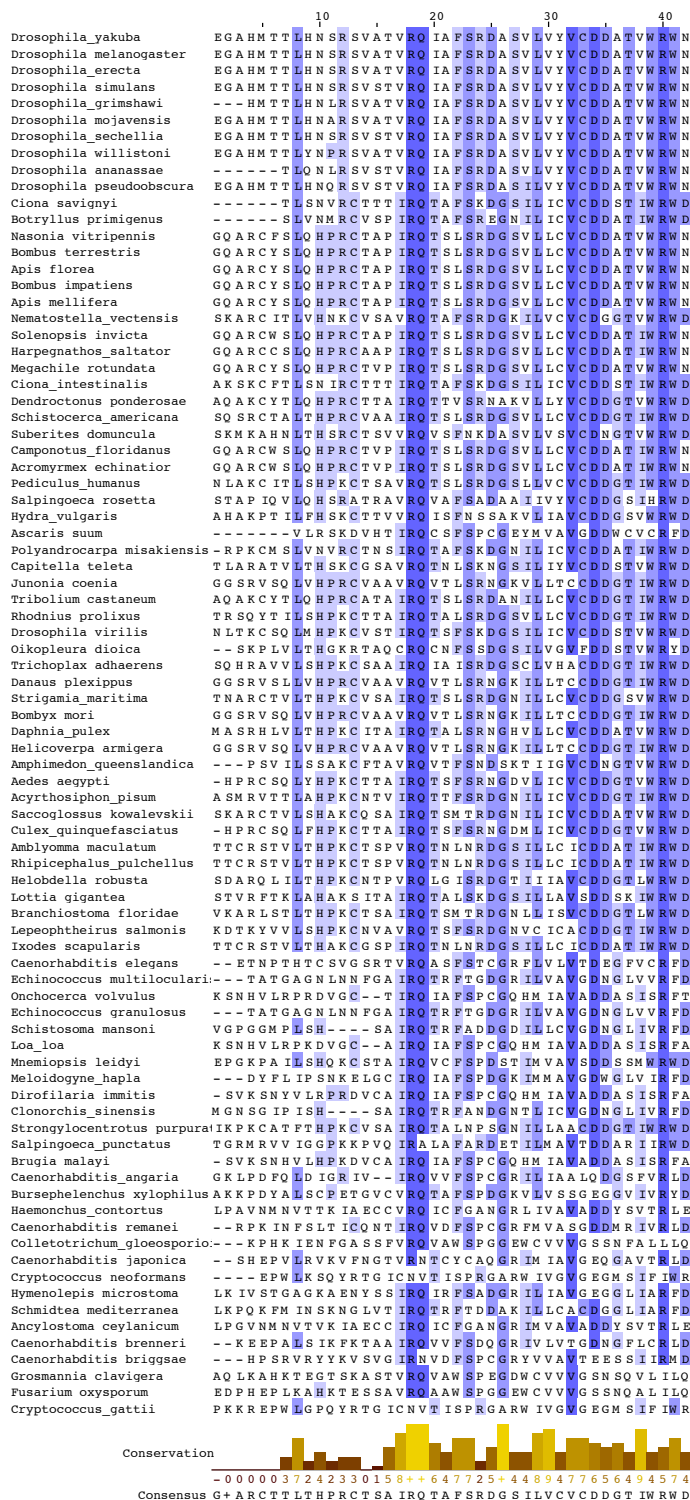


Appendix Figure 5.17 continued Amino acid alignment of the WD40-5 repeat from *Esc.* Sequences were aligned in Seaview using Muscle and manually edited in Jalview. Percentage identity to the consensus sequence is color coded from dark blue (>80%) to light blue (>40%) and white (<40%). Taxa are sorted by pairwise identity in Jalview.

	10	20	30
Camponotus floridanus	F E F K E C D I W F I R F S M D F W Q R T I A L G N O V G R T Y V W D		
Megachile rotundata	F E F K E C D I W F I R F S M D F W Q R T I A L G N O V G R T Y V W D		
Acromyrmex echinatior	F E F K E C D I W F I R F S M D F W Q R T I A L G N O V G R T Y V W D		
Apis mellifera	F E F K E C D I W F I R F S M D F W Q R T I A L G N O V G R T Y V W D		
Apis florea	F E F K E C D I W F I R F S M D F W Q R T I A L G N O V G R T Y V W D		
Solenopsis invicta	F E F K E C D I W F I R F S M D F W Q R T I A L G N O V G R T Y V W D		
Bombus impatiens	F E F K E C D I W F I R F S M D F C O R T I A L G N O V G R T Y V W D		
Nasonia vitripennis	F E F K E C D I W F I R F S M D F W Q R T I A L G N O V G R T Y V W D		
Harpegnathos saltator	F E F K E C D I W F I R F S M D F W Q R T I A M G N O V G R T Y V W D		
Bombus terrestris	F E F K E C D I W F I R F S M D F C O R T I A L G N O V G R T Y V W D		
Macrobrachium nipponense	F D Y K E C E I W F M R F A L D F W Q K - - - - -		
Pediculus humanus corporis	F E Y K E C E I W F V R F S M D F W Q R I L L A L G N O A G R T F V W D		
Ixodes scapularis	F E Y K E C N I W F M R F S M D F E Q R I L L A L G N O V G K T Y V W N		
Capitella teleta	F D F K E C D I W F M R F S M D F W Q R I L L A M G P D Y G R V F V W D		
Amblyomma maculatum	F E Y R E C N I W F M R F S M D F E Q R I L L A L G N O V G K T Y V W D		
Rhipicephalus pulchellus	F E Y R E C N I W F M R F S M D F E Q R I L L A L G N O V G K T Y V W D		
Strigamia maritima	F D Y K E C E I W F M R F S I D F W Q R I M A L G N O V G K T F V W D		
Schistocerca americana	F E Y R E C E I W F V R F A M D F W Q K I L L A L G N O V G K T F V W D		
Rhodnius prolixus	F D F K E C E I W F V R F A L D Y W Q K I L L A L G N O V G R T F V W D		
Dendroctonus ponderosae	F E Y K E C E I W F V R F A M D F W Q K I L L A L G N O T G K V F V W D		
Tribolium castaneum	F E Y K E C E I W F V R F A M D F W Q K I L L A L G N O T G K I F V W D		
Aedes aegypti	L E Y K E C D I W F I R F S L D Y W Q K Y L L A L G N O N G K T Y L W E		
Bombyx mori	F D Y K E C E I W F I R F A V D Y S Q R V I A L G N O C G K T M V W E		
Daphnia pulex	F D Y R E C D I W F M R F S L D S W N K V M A L G N O V G K T F V W D		
Saccolossus kowalevskii	F D Y T Q C D I W Y M R F S M D Y W Q K I L L A L G N O V G K T Y I W D		
Junonia coenia	F D Y K E C E I W F I R F A V D Y S Q R V I A L G N O C G K T M V W E		
Acyrtosiphon pisum	Y D F K D C D V W F I R F S M D F S Q K I L L A L G N T I G K I Y V W D		
Culex quinquefasciatus	L E Y K D C D I W F I R F S L D Y W Q K Y L L A L G N O I G K T Y I W E		
Lottia gigantea	F E Y K E C D I W Y M R F S L D Y S Q R I M A L G N O C G K I Y V W D		
Danaus plexippus	F D Y K E C E I W F I R F A V D Y S Q R V I A L G N O C G K T M V W E		
Helicoverpa armigera	F D Y K E C E I W F I R F A V D Y S Q R V I A L G N O C G K T M V W E		
Strongylocentrotus purpuratus	F N Y T Q C D I W F M R F S M D F W Q R M L L A L G N O V G K I F V W D		
Branchiostoma floridae	F O Y H Q C D I W Y M R F S I D Y W Q K V L L A L G N O V G A L F V W D		
Drosophila virilis	F E Y D E C E I W F V R F G F N P W Q K V I A L G N O Q G K V Y V W E		
Drosophila grimshawi	F E Y D E C E I W F V R F G F N P W Q K V I A L G N O Q G K V Y V W E		
Drosophila ananassae	F E Y D E C E I W F V R F G F N P W Q K V I A L G N O Q G K V Y V W E		
Drosophila erecta	F E Y D E C E I W F V R F G F N P W Q K V I A L G N O Q G K V Y V W E		
Drosophila sechellia	F E Y D E C E I W F V R F G F N P W Q K V I A L G N O Q G K V Y V W E		
Drosophila simulans	F E Y D E C E I W F V R F G F N P W Q K V I A L G N O Q G K V Y V W E		
Drosophila melanoqaster	F E Y D E C E I W F V R F G F N P W Q K V I A L G N O Q G K V Y V W E		
Drosophila mojavensis	F E Y D E C E I W F V R F G F N P W Q K V I A L G N O Q G K V Y V W E		
Drosophila yakuba	F E Y D E C E I W F V R F G F N P W Q K V I A L G N O Q G K V Y V W E		
Drosophila willistoni	F E Y D E C E I W F V R F G F N P W Q K V I A L G N O Q G K V Y V W E		
Nematostella vectensis	E D F S Q C E I W Y M R F S L D F E Q R L V A L G N O C G K V F V W D		
Polyandrocarpa misakiensis	L Q Y H C E I W Y M R F S M D L R Q R F L A L G N O V G K T F V W D		
Lepeophtheirus salmonis	L N Y K D N E I W F I R F A L D K G Q K L L A L G N O M G R T Y I W D		
Hydra vulgaris	F d I D L C D I W F I R F A V D L N Q T I L L A L G N O I G K V Y L Y D		
Drosophila pseudoobscura	F S Y D E C E I W F V R F G F N P W H K V I A L G N O H G K V Y V W E		
Suberites domuncula	L E I H C D I W F I R F A V N F K O T L L A L G N T A G R I S L W D		
Trichoplax adhaerens	F D Y N C D I W Y L R F C L D F Q Q R T L A L G N O V G K V F L W D		
Amphimedon queenslandica	F E Y P W C E I W F I R F A M D R K M K Y L L A L G M Q I E I H I W D		
Helobdella robusta	H D P S A C E I W F M R F S L D Y D Q K I L L A V G S O T G K I F V W D		
Ciona intestinalis	L E P Q H C D I W Y M R F A I D Y W H K Y L A V G N O V G K T F I E		
Schmidtea mediterranea	F E L D D C D I W Y V R F D I D V K R G L L A L G N E L H I V W N		
Botryllus primigenus	L O Y Q H C E I W Y M R F A M D M K Q R F L A L G N O V G K V F L W D		
Ciona savignyi	L D F O H C D I W Y M R F A I D Y W H K Y L A V G N O V G K T F I E		
Mnemioopsis leidyi	F E Y G N C D I W Y M R F A V S P N F D A I A A G M O I G K V F L W D		
Echinococcus multilocularis	L R L P D C D L W Y V R F D L H L Q R L L A L G T G V G R I F L W D		
Fusarium oxysporum	- - - - - P Q F F M R F K L H F O H P V L A F C N A G G N V F F W D		
Echinococcus granulosus	L R L P D C D L W Y V R F D L H L Q R L L A L G T G V G R I F L W D		
Salpingoeca punctatus	F E Y P E Q D L W E V R A T L S P S G R Y L L A V G N M M G E I Y I W D		
Hymenolepis microstoma	L Q M D D T E L W Y I R F D L H L Q R L L A L G T G A G R I Y L W D		
Schistosoma mansoni	L K A T D C D L W Y I R F D I D L K N H V L A L G T G T G R V Y L W D		
Colletotrichum gloeosporioides	- - - - - G P Q F M R F K L H F O H P V L A F C N A N K I F F W D		
Grosmanina clavigera	- - - - - S Q F F T R F G L F N H H P T L A F C N T Y A K I F F W D		
Ancylostoma ceylanicum	M D L P D S D I W F E I K F D I D P L N R W I V S G N K M G Q L C F W D		
Clonorchis sinensis	L K A P D C D L W Y I R F D V D L A N Q V L L A L G T G T G R I Y L W D		
Haemonchus contortus	M D L P D S D I W F I K F D I D P L N R W I V S G N K M G Q L C M W D		
Ascaris suum	L S W P E T N M W Y I K F E I D P A K Q Y L V C G N O K E I H I W D		
Caenorhabditis briggsae	H A I G E G K R F K F S I D P K R R W I G G D E S I M F P D		
Loa loa	M E L P E T N M W Y I K F E I D P L E K Y L V C G N O K E I H V W E		
Oikopleura dioica	I S L P Y S P N W Y V R F G L D R Y L Q Y M A G N L N G M Y V W D		
Salpingoeca rosetta	L P L D N C D I W E V K E D V E A T F T F L A A G N O A G K V F L W D		
Cryptococcus neoformans	Y R T G I G N - - - - - V T I S P G A R W I - V G V E S I F I R R		
Caenorhabditis japonica	F K I W N G D T W F T K F E M D P R R R W L A V G T O G F V N F F D		
Bursaphelenchus xylophilus	W E V P K S S I W F I K F D I D P D N K Y L A V G N E E T V K L V D		
Caenorhabditis angaria	F D Y V D C E T W E V K F D L D P L N R W I T C G N N R G D V F F S		
Brugia malayi	M E L P E T N M W Y I K F E I D P L E K Y L V C G N O K E I H I W E		
Onchocerca volvulus	M Q L P E T N M W Y I K F E I D P L E K Y L V C G N O K E I H V W E		
Dirofilaria immitis	M E L P D T N M W Y I K F E I D P L E K Y L V C G N O K E I H I W E		
Cryptococcus gattii	F D Y S D S D M W F R H H Q M D P S K H W Y A T G S A H N S L L V R R		
Meloidogyne hapla	- K M P N T E M W F I K M A V D P H R R F L A C C S D O C E I R I R R		
Caenorhabditis elegans	M N V P S G S A W F I K F A V D P R R R W L V C G G A G G S V M F F D		
Caenorhabditis brenneri	I G A K S L K E W F C K G G V D P L R K Y I G V G R G G L O P H M E		
Monosiga brevicollis	L D V P R C N M W Y I R F A Y A A Q V P F V G I G N T T E I F L Y N		
Caenorhabditis remanei	V T N D N G E V W F T K A I D P R R R W L V C G C T R A I V N F I D		



Appendix Figure 5.17 continued Amino acid alignment of the WD40-6 repeat from Esc. Sequences were aligned in Seaview using Muscle and manually edited in Jalview. Percentage identity to the consensus sequence is color coded from dark blue (>80%) to light blue (>40%) and white (<40%). Taxa are sorted by pairwise identity in Jalview.



Appendix Figure 5.17 continued Amino acid alignment of the WD40-7 repeat from Esc. Sequences were aligned in Seaview using Muscle and manually edited in Jalview. Percentage identity to the consensus sequence is color coded from dark blue (>80%) to light blue (>40%) and white (<40%). Taxa are sorted by pairwise identity in Jalview.

REFERENCES

- Aboobacker, A., & Blaxter, M. (2003). Hox gene evolution in nematodes : novelty conserved. *Current Opinion in Genetics & Development*, 13(6), 593–8.
- Aboobaker, A., & Blaxter, M. L. (2003). Hox gene loss during dynamic evolution of the nematode cluster. *Current Biology*, 13(02), 37–40.
- Abouheif, E., & Wray, G. A. (2002). Evolution of the gene network underlying wing polyphenism in ants. *Science*, 297(5579), 249–52.
- Alpert, P. (1999). Effects of clonal integration on plant plasticity in *Fragaria chiloensis*. *Plant Ecology*, 141(1), 99–106.
- Aspland, S. E., & White, R. A. (1997). Nucleocytoplasmic localisation of extradenticle protein is spatially regulated throughout development in *Drosophila*. *Development*, 124(3), 741–7.
- Beldade, P., & Brakefield, P. M. (2002). The genetics and evo-devo of butterfly wing patterns. *Nature Reviews. Genetics*, 3(6), 442–52.
- Beldade, P., & Brakefield, P. M. (2003). Concerted evolution and developmental integration in modular butterfly wing patterns, 179, 169–179.
- Beldade, P., Brakefield, P. M., & Long, A. D. (2002). Contribution of *Distal-less* to quantitative variation in butterfly eyespots. *Nature*, 415, 315–318.
- Beldade, P., Koops, K., & Brakefield, P. M. (2002). Modularity, individuality, and evo-devo in butterfly wings. *Proceedings of the National Academy of Sciences*, 99(22), 14262–14267.
- Beldade, P., Rudd, S., Gruber, J. D., & Long, A. D. (2006). A wing expressed sequence tag resource for *Bicyclus anynana* butterflies, an evo-devo model. *BMC Genomics*, 7, 130.
- Bennett, D. C., & Lamoreux, M. L. (2003). Review – Pigment Gene Focus The Color Loci of Mice – A Genetic Century. *Pigment Cell Research / Sponsored by the European Society for Pigment Cell Research and the International Pigment Cell Society*, 16, 333–344.

- Betts, M. J., & Russell, R. B. (2007). Amino-Acid Properties and Consequences of Substitutions. *Bioinformatics for Geneticists: A Bioinformatics Primer for the Analysis of Genetic Data: Second Edition*, 4, 311–342.
- Birve, A, Sengupta, A. K., Beuchle, D., Larsson, J., Kennison, J. A, Rasmuson-Lestander A, & Müller, J. (2001). Su(z)12, a novel Drosophila Polycomb group gene that is conserved in vertebrates and plants. *Development*, 128, 3371–3379.
- Blair, S. S., & Ralston, A. (1997). Smoothed-mediated Hedgehog signalling is required for the maintenance of the anterior-posterior lineage restriction in the developing wing of Drosophila. *Development*, 124, 4053–4063.
- Boerjan, B., Sas, F., Ernst, U. R., Tobback, J., Lemièrre, F., Vandegehuchte, M. B., ... De Loof, A. (2011). Locust phase polyphenism: Does epigenetic precede endocrine regulation? *General and Comparative Endocrinology*, 173(1), 120–8.
- Bowers, M. D. (1998). Effects of hostplant species and artificial diet. *Journal of the Lepidopterists Society*, 52(1), 73–83.
- Bracken, A. P., Dietrich, N., Pasini, D., Hansen, K. H., & Helin, K. (2006). Genome-wide mapping of Polycomb target genes unravels their roles in cell fate transitions. *Genes & Development*, 20, 1123–1136.
- Braendle, C., & Flatt, T. (2006). A role for genetic accommodation in evolution? *BioEssays*, 28, 868–873.
- Brakefield, P., Gates, J., Keys, D., Kesbeke, F., Wijngaarden, P., Monteiro, A., ... Carroll, S. (1996). Development, plasticity and evolution of butterfly eyespot patterns. *Nature*, 384(6606), 236–242.
- Brakefield, P. M. (1996). Seasonal polyphenism in butterflies and natural selection. *Trends in Ecology & Evolution*, 11(7), 275–277.
- Brakefield, P. M. (2001). Structure of a Character and the Evolution of Butterfly Eyespot Patterns. *Journal of Experimental Zoology. Part B, Molecular and Developmental Evolution*, 104, 93–104.
- Brakefield, P. M. (2006). The power of evo-devo to explore evolutionary constraints : experiments with butterfly eyespots . *Zoology*, 106(2003), 283–290.
- Brakefield, P. M., & French, V. (1999). Butterfly wings: the evolution of development of colour patterns. *BioEssays*, 21(5), 391–401.

- Brakefield, P. M., Kesbeke, F., & Koch, P. B. (1998). The regulation of phenotypic plasticity of eyespots in the butterfly *Bicyclus anynana*. *The American Naturalist*, *152*(6), 853–60.
- Brattström, O., Kjellén, N., Alerstam, T., & Åkesson, S. (2008). Effects of wind and weather on red admiral, *Vanessa atalanta*, migration at a coastal site in southern Sweden. *Animal Behaviour*, *76*(2), 335–344.
- Bray, S. (1999). *Drosophila* development: Scalloped and Vestigial take wing. *Current Biology : CB*, *9*(7), R245–7.
- Breiling, A., Sessa, L., & Orlando, V. (2007). Biology of polycomb and trithorax group proteins. *International Review of Cytology*, *258*(07), 83–136.
- Breuker, C. J., Gibbs, M., Dyck, H. V. A. N., Brakefield, P. M., & Klingenberg, C. P. (2007). Integration of Wings and Their Eyespots in the Speckled Wood Butterfly *Pararge aegeria*. *Journal of Experimental Zoology. Part B, Molecular and Developmental Evolution*, *463*, 454–463.
- Brewster, R., Hardiman, K., Deo, M., Khan, S., & Bodmer, R. (2001). The selector gene cut represses a neural cell fate that is specified independently of the Achaete-Scute-Complex and atonal. *Mechanisms of Development*, *105*, 57–68.
- Brisson, J. A., Ishikawa, A., & Miura, T. (2010). Wing development genes of the pea aphid and differential gene expression between winged and unwinged morphs. *Insect Molecular Biology*, *19 Suppl 2*, 63–73.
- Brock, Jim, P., & Kaufman, K. (2006). *Kaufman Field Guide to Butterflies of North America*. Houghton Mifflin Harcourt.
- Brunetti, C. R., Selegue, J. E., Monteiro, a, French, V., Brakefield, P. M., & Carroll, S. B. (2001). The generation and diversification of butterfly eyespot color patterns. *Current Biology : CB*, *11*(20), 1578–85.
- Cao, R., Wang, L., Wang, H., Xia, L., Erdjument-Bromage, H., Tempst, P., ... Zhang, Y. (2002). Role of histone H3 lysine 27 methylation in Polycomb-group silencing. *Science*, *298*(5595), 1039–43.
- Cao, R., & Zhang, Y. (2004a). SUZ12 is required for both the histone methyltransferase activity and the silencing function of the EED-EZH2 complex. *Molecular Cell*, *15*(1), 57–67.
- Cao, R., & Zhang, Y. (2004b). The functions of E(Z)/EZH2-mediated methylation of lysine 27 in histone H3. *Current Opinion in Genetics & Development*, *14*(2), 155–64.

- Carroll, S. B., Gates, J., Keys, D. N., Paddock, S. W., Panganiban, G. E., Selegue, J. E., & Williams, J. a. (1994). Pattern formation and eyespot determination in butterfly wings. *Science (New York, N.Y.)*, 265(5168), 109–14.
- Carroll, S., Grenier, J., & Weatherbee, S. (2001). Wiley: From DNA to Diversity: Molecular Genetics and the Evolution of Animal Design, 2nd Edition - Sean B. Carroll, Jennifer K. Grenier, Scott D. Weatherbee. *Blackwell Science*.
- Cedar, H., & Bergman, Y. (2009). Linking DNA methylation and histone modification: patterns and paradigms. *Nature Reviews. Genetics*, 10(5), 295–304.
- Cerisse E Allen. (2008). The “Eyespot Module ” and Eyespots as Modules : Development , Evolution , and Integration of a Complex Phenotype. *Journal of Experimental Biology*, 190, 179–190.
- Chai, J., & Tarnawski, A. S. (2002). Serum response factor: Discovery, biochemistry, biological roles and implications for tissue injury healing. *Journal of Physiology and Pharmacology*, 53, 147–157.
- Champagne, F. A. (2013). Epigenetics and developmental plasticity across species. *Developmental Psychobiology*, 55(1), 33–41.
- Cho, E. H., & Nijhout, H. F. (2013). Development of polyploidy of scale-building cells in the wings of *Manduca sexta*. *Arthropod Structure & Development*, 42(1), 37–46.
- Ciferri, C., Lander, G. C., Maiolica, A., Herzog, F., Aebersold, R., & Nogales, E. (2012). Molecular architecture of human polycomb repressive complex 2. *eLife*, 1, e00005.
- Coghlan, A., & Wolfe, K. H. (2002). Fourfold faster rate of genome rearrangement in nematodes than in *Drosophila*. *Genome Research*, 12, 857–867.
- Crews, D., & Gore, A. C. (2012). Epigenetic synthesis: a need for a new paradigm for evolution in a contaminated world. *F1000 Biology Reports*, 4(18), 1–6.
- Crispo, E. (2008). Modifying effects of phenotypic plasticity on interactions among natural selection, adaptation and gene flow. *Journal of Evolutionary Biology*, 21, 1460–1469.
- Croucher, P. J. P., Brewer, M. S., Winchell, C. J., Oxford, G. S., & Gillespie, R. G. (2013). De novo characterization of the gene-rich transcriptomes of two color-polymorphic spiders, *Theridion grallator* and *T. californicum* (Araneae: Theridiidae), with special reference to pigment genes. *BMC Genomics*, 14, 862.
- De Celis, J. F., & Barrio, R. (2000). Function of the spalt/spalt-related gene complex in positioning the veins in the *Drosophila* wing. *Mechanisms of Development*, 91(1-2), 31–41.

- Degnan, J. H., & Rosenberg, N. A. (2009). Gene tree discordance, phylogenetic inference and the multispecies coalescent. *Trends in Ecology and Evolution*, 24(March), 332–340.
- Dellino, G. I., Schwartz, Y. B., Farkas, G., McCabe, D., Elgin, S. C. R., & Pirrotta, V. (2004). Polycomb silencing blocks transcription initiation. *Molecular Cell*, 13, 887–893.
- Derkacheva, M., & Hennig, L. (2013). Variations on a theme: Polycomb group proteins in plants. *Journal of Experimental Botany*.
- Dilão, R., & Sainhas, J. (2004). Modelling butterfly wing eyespot patterns. *Proceedings. Biological Sciences / The Royal Society*, 271(1548), 1565–9.
- Dillon, S. C., Zhang, X., Trievel, R. C., & Cheng, X. (2005). The SET-domain protein superfamily: protein lysine methyltransferases. *Genome Biology*, 6, 227.
- Dolinoy, D. C. (2008). The agouti mouse model: an epigenetic biosensor for nutritional and environmental alterations on the fetal epigenome. *Nutrition Reviews*, 66 Suppl 1, S7–11.
- Eklom, R., & Galindo, J. (2011). Applications of next generation sequencing in molecular ecology of non-model organisms. *Heredity*, 107(1), 1–15.
- Feil, R., & Fraga, M. F. (2011). Epigenetics and the environment: emerging patterns and implications. *Nature Reviews. Genetics*, 13(2), 97–109.
- Feliciello, I., Parazajder, J., Akrap, I., & Ugarković, D. (2013). First evidence of DNA methylation in insect *Tribolium castaneum*: environmental regulation of DNA methylation within heterochromatin. *Epigenetics : Official Journal of the DNA Methylation Society*, 8(5), 534–41.
- Ferguson, L. C., Green, J., Surridge, A., & Jiggins, C. D. (2011). Evolution of the insect yellow gene family. *Molecular Biology and Evolution*, 28(1), 257–72.
- Ferguson, L. C., & Jiggins, C. D. (2009). Shared and divergent expression domains on mimetic *Heliconius* wings. *Evolution & Development*, 11(5), 498–512.
- Ferguson, L. C., Maroja, L., & Jiggins, C. D. (2011). Convergent, modular expression of ebony and tan in the mimetic wing patterns of *Heliconius* butterflies. *Development Genes and Evolution*, 221(5-6), 297–308.
- Ferguson, L., Lee, S. F., Chamberlain, N., Nadeau, N., Joron, M., Baxter, S., ... Jiggins, C. (2010). Characterization of a hotspot for mimicry: assembly of a butterfly wing transcriptome to genomic sequence at the HmYb/Sb locus. *Molecular Ecology*, 19 Suppl 1, 240–54.

- French-Constant, R. H. (2012). Butterfly wing colours are driven by the evolution of developmental heterochrony. Butterfly wing colours and patterning by numbers. *Heredity*, 108(6), 592–3.
- Fitzpatrick, B. M. (2012). Underappreciated Consequences of Phenotypic Plasticity for Ecological Speciation. *International Journal of Ecology*, 2012, 1–12.
- Forsman, A. (2014). Rethinking phenotypic plasticity and its consequences for individuals, populations and species. *Heredity*, (August), 1–9.
- French, V., & Brakefield, P. M. (1995). Eyespot development on butterfly wing- the focal signal.pdf. *Developmental Biology*, 168, 112–123.
- Galant, R., Walsh, C. M., & Carroll, S. B. (2002). Hox repression of a target gene: extradenticle-independent, additive action through multiple monomer binding sites. *Development*, 129(13), 3115–26.
- García-Bellido, A., & De Celis, J. F. (2009). The complex tale of the achaete-scute complex: A paradigmatic case in the analysis of gene organization and function during development. *Genetics*, 182, 631–639.
- Gebelein, B., Culi, J., Ryoo, H. D., Zhang, W., & Mann, R. S. (2002). Specificity of Distalless repression and limb primordia development by abdominal Hox proteins. *Developmental Cell*, 3, 487–498.
- Gellon, G., & McGinnis, W. (1998). Shaping animal body plans in development and evolution by modulation of Hox expression patterns. *BioEssays : News and Reviews in Molecular, Cellular and Developmental Biology*, 20(2), 116–25.
- Geng, Y., Gao, L., & Yang, J. (2013). Epigenetic flexibility underlying phenotypic plasticity. *Progress in Botany*, 74, 153–163.
- Ghalambor, C. K., McKay, J. K., Carroll, S. P., & Reznick, D. N. (2007). Adaptive versus non-adaptive phenotypic plasticity and the potential for contemporary adaptation in new environments. *Functional Ecology*, 21, 394–407.
- Gibbs, M., & Breuker, C. J. (2006). Effect of larval-rearing density on adult life-history traits and developmental stability of the dorsal eyespot pattern in the speckled wood butterfly, *Pararge aegeria*. *Entomologia Experimentalis et Applicata*, 118, 41–47.
- Gibert, J.-M., Peronnet, F., & Schlötterer, C. (2007). Phenotypic plasticity in *Drosophila* pigmentation caused by temperature sensitivity of a chromatin regulator network. *PLoS Genetics*, 3(2), e30.

- Gouy, M., Guindon, S., & Gascuel, O. (2010). SeaView version 4: A multiplatform graphical user interface for sequence alignment and phylogenetic tree building. *Molecular Biology and Evolution*, *27*(2), 221–224.
- Gunster, M. J., Raaphorst, F. M., Hamer, K. M., den Blaauwen, J. L., Fieret, E., Meijer, C. J. L. M., & Otte, A. P. (2001). Differential expression of human Polycomb group proteins in various tissues and cell types. *Journal of Cellular Biochemistry*, *81*(S36), 129–143.
- Halder, G., & Carroll, S. B. (2001). Binding of the Vestigial co-factor switches the DNA-target selectivity of the Scalloped selector protein. *Development*, *128*(17), 3295–305.
- Hamamoto, R., Saloura, V., & Nakamura, Y. (2015). Critical roles of non-histone protein lysine methylation in human tumorigenesis. *Nature Reviews Cancer*, *15*(2), 110–124.
- Han, Q., Fang, J., Ding, H., Johnson, J., Bruce, C., & Li, J. (2002). Identification of *Drosophila melanogaster* yellow-f and yellow-f2 proteins as dopachrome-conversion enzymes. *Biochem. J.*, *340*, 333–340.
- Han, Z., Xing, X., Hu, M., Zhang, Y., Liu, P., & Chai, J. (2007). Structural basis of EZH2 recognition by EED. *Structure (London, England : 1993)*, *15*(10), 1306–15.
- Hansen, K. H., & Helin, K. (2009). Epigenetic inheritance through self-recruitment of the polycomb repressive complex 2. *Epigenetics : Official Journal of the DNA Methylation Society*, *4*(3), 133–8.
- Held, L. I. (2012). Rethinking Butterfly Eyespots. *Evolutionary Biology*, *40*(1), 158–168.
- Henikoff, S., & Henikoff, J. G. (1992). Amino acid substitution matrices from protein blocks. *Proceedings of the National Academy of Sciences of the United States of America*, *89*(22), 10915–9.
- Hennig, L., & Derkacheva, M. (2009). Diversity of Polycomb group complexes in plants: same rules, different players? *Trends in Genetics*, *25*, 414–423.
- Herz, H.-M., Garruss, A., & Shilatifard, A. (2013). SET for life: biochemical activities and biological functions of SET domain-containing proteins. *Trends in Biochemical Sciences*, *38*(12), 621–39.
- Hines, H. M., Papa, R., Ruiz, M., Papanicolaou, A., Wang, C., Nijhout, H. F., ... Reed, R. D. (2012). Transcriptome analysis reveals novel patterning and pigmentation genes underlying *Heliconius* butterfly wing pattern variation. *BMC Genomics*, *13*(1), 288.

- Hiyama, A., Taira, W., & Otaki, J. M. (2012). Color-pattern evolution in response to environmental stress in butterflies. *Frontiers in Genetics*, 3(February), 15.
- Hunt, B. G., Brisson, J. A., Yi, S. V., & Goodisman, M. A. D. (2010). Functional conservation of DNA methylation in the pea aphid and the honeybee. *Genome Biology and Evolution*, 2, 719–28.
- Ingham, P. W., & McMahon, A. P. (2001). Hedgehog signaling in animal development : paradigms and principles Hedgehog signaling in animal development : paradigms and principles. *Genes and Development*, 15, 3059–3087.
- Iwata, M., Ohno, Y., & Otaki, J. M. (2014). Real-time in vivo imaging of butterfly wing development: revealing the cellular dynamics of the pupal wing tissue. *PloS One*, 9(2), e89500.
- Jablonka, E., & Lamb, M. J. (2002). The changing concept of epigenetics. *Annals of the New York Academy of Sciences*, 981, 82–96.
- Janssen, J. M., Monteiro, A., & Brakefield, P. M. (2001). Correlations between scale structure and pigmentation in butterfly wings. *Evolution and Development*, 3(6), 415–423.
- Jeong, S., Rokas, A., & Carroll, S. B. (2006). Regulation of body pigmentation by the Abdominal-B Hox protein and its gain and loss in *Drosophila* evolution. *Cell*, 125(7), 1387–99.
- Jones, C. A., Ng, J., Peterson, A. J., Morgan, K., Simon, J., & Jones, R. S. (1998). The *Drosophila* *esc* and *E(z)* Proteins Are Direct Partners in Polycomb Group-Mediated Repression The *Drosophila* *esc* and *E(z)* Proteins Are Direct Partners in Polycomb Group-Mediated Repression. *Molecular and Cellular Biology*.
- Joron, M., Jiggins, C. D., Papanicolaou, a, & McMillan, W. O. (2006). Heliconius wing patterns: an evo-devo model for understanding phenotypic diversity. *Heredity*, 97(3), 157–67.
- Joshi, P., Carrington, E. A., Wang, L., Ketel, C. S., Miller, E. L., Jones, R. S., & Simon, J. a. (2008). Dominant alleles identify SET domain residues required for histone methyltransferase of polycomb repressive complex 2. *Journal of Biological Chemistry*, 283(41), 27757–27766.
- Jürgens, G. (1985). A group of genes controlling the spatial expression of the bithorax complex in *Drosophila*. *Nature*, 316(6024), 153–155.
- Kerppola, T. K. (2009). Polycomb group complexes--many combinations, many functions. *Trends in Cell Biology*, 19(12), 692–704. d

- Ketel, C. S., Andersen, E. F., Vargas, M. L., Suh, J., Strome, S., & Simon, J. A. (2005). Subunit Contributions to Histone Methyltransferase Activities of Fly and Worm Polycomb Group Complexes, *25*(16), 6857–6868.
- Keys, D. N., Lewis, D. L., Selegue, J. E., Pearson, B. J., Goodrich, L. V, Johnson, R. L., ... Carroll, S. B. (1999). Recruitment of a hedgehog regulatory circuit in butterfly eyespot evolution. *Science (New York, N.Y.)*, *283*(5401), 532–4.
- Kim, D.-H., & Sung, S. (2014). Genetic and epigenetic mechanisms underlying vernalization. *The Arabidopsis Book / American Society of Plant Biologists*, *12*, e0171.
- Klingenberg, C. P. (2008). Morphological Integration and Developmental Modularity. *Annual Review of Ecology, Evolution and Systematics*, *39*, 115–132.
- Klose, R. J., Cooper, S., Farcas, A. M., Blackledge, N. P., & Brockdorff, N. (2013). Chromatin sampling--an emerging perspective on targeting polycomb repressor proteins. *PLoS Genetics*, *9*(8), e1003717.
- Knüttel, H., & Fiedler, K. (2001). Host-plant-derived variation in ultraviolet wing patterns influences mate selection by male butterflies. *The Journal of Experimental Biology*, *204*(Pt 14), 2447–59.
- Kobayashi, M. (2003). Engrailed cooperates with extradenticle and homothorax to repress target genes in *Drosophila*. *Development*, *130*(4), 741–751.
- Koch, P. B., Lorenz, U., Brakefield, P. M., & ffrench-Constant, R. H. (2000). Butterfly wing pattern mutants: developmental heterochrony and co-ordinately regulated phenotypes. *Development Genes and Evolution*, *210*(11), 536–44.
- Koch, P. B., Merk, R., Reinhardt, R., & Weber, P. (2003). Localization of ecdysone receptor protein during colour pattern formation in wings of the butterfly *Precis coenia* (Lepidoptera: Nymphalidae) and co-expression with Distal-less protein. *Development Genes and Evolution*, *212*(12), 571–84.
- Kortschak, R. D., Samuel, G., Saint, R., & Miller, D. J. (2003). EST Analysis of the Cnidarian *Acropora millepora* Reveals Extensive Gene Loss and Rapid Sequence Divergence in the Model Invertebrates. *Current Biology*, *13*, 2190–2195.
- Krauss, V., Eisenhardt, C., & Unger, T. (2009). The genome of the stick insect *Medauroidea extradentata* is strongly methylated within genes and repetitive DNA. *PloS One*, *4*(9), e7223. d
- Kremen, C., & Nijhout, H. F. (1998). Control of pupal commitment in the imaginal disks of *Precis coenia* (Lepidoptera: Nymphalidae). *Journal of Insect Physiology*.

- Krupp, J. J., Yaich, L. E., Wessells, R. J., & Bodmer, R. (2005). Identification of genetic loci that interact with cut during *Drosophila* wing-margin development. *Genetics*, *170*(4), 1775–95.
- Kucharski, R., Maleszka, J., Foret, S., & Maleszka, R. (2008). Nutritional control of reproductive Status in honeybees via DNA methylation. *Science*, *319*(March), 1827–1830.
- Letunic, I., Doerks, T., & Bork, P. (2014). SMART: recent updates, new developments and status in 2015. *Nucleic Acids Research*, *43*(6), D257–D260.
- Levine, M. (2008). A systems view of *Drosophila* segmentation. *Genome Biology*, *9*, 207.
- Lewis, E. B. (1978). A gene complex controlling segmentation in *Drosophila*. *Nature*, *276*(5688), 565–570.
- Li, Z., Tatsuke, T., Sakashita, K., Zhu, L., Xu, J., Mon, H., ... Kusakabe, T. (2012). Identification and characterization of Polycomb group genes in the silkworm, *Bombyx mori*. *Molecular Biology Reports*, *39*(5), 5575–88.
- Li, Z.-Q., Zhang, S., Ma, Y., Luo, J.-Y., Wang, C.-Y., Lv, L.-M., ... Cui, J.-J. (2013). First Transcriptome and Digital Gene Expression Analysis in Neuroptera with an Emphasis on Chemoreception Genes in *Chrysopa pallens* (Rambur). *PloS One*, *8*(6), e67151.
- Luo, M., Platten, D., Chaudhury, A., Peacock, W. J., & Dennis, E. S. (2009). Expression, imprinting, and evolution of rice homologs of the polycomb group genes. *Molecular Plant*, *2*(4), 711–723.
- Lyko, F., Foret, S., Kucharski, R., Wolf, S., Falckenhayn, C., & Maleszka, R. (2010). The honey bee epigenomes: differential methylation of brain DNA in queens and workers. *PLoS Biology*, *8*(11), e1000506.
- Lyytinen, A., Brakefield, P. M., Lindström, L., & Mappes, J. (2004). Does predation maintain eyespot plasticity in *Bicyclus anynana*? *Proceedings. Biological Sciences / The Royal Society*, *271*(1536), 279–83.
- Maan, M. E., & Cummings, M. E. (2009). Sexual dimorphism and directional sexual selection on aposematic signals in a poison frog. *Proceedings of the National Academy of Sciences of the United States of America*, *106*(45), 19072–7.
- Macdonald, W. P., Martin, A., & Reed, R. D. (2010). Butterfly wings shaped by a molecular cookie cutter: evolutionary radiation of lepidopteran wing shapes associated with a derived Cut/wingless wing margin boundary system. *Evolution & Development*, *12*(3), 296–304.

- Macneil, L. T., & Walhout, A. J. M. (2011). Gene regulatory networks and the role of robustness and stochasticity in the control of gene expression. *Genome Research*, *21*(5), 645–57.
- Maddison, D., & Maddison, W. (2000). MacClade 4. *Manual*, 1–492.
- Maddison, W. P. (2008). Gene Trees in Species Trees. *Systematic Biology*, *46*(3), 523–536.
- Magwene, Paul, M. (2001). New tools for studying integration and modularity. *Evolution*, *55*(9), 1734–1745.
- Mahdi, S. H. A., Yamasaki, H., & Otaki, J. M. (2011). Heat-shock-induced color-pattern changes of the blue pansy butterfly *Junonia orithya*: Physiological and evolutionary implications. *Journal of Thermal Biology*, *36*(6), 312–321.
- Mandrioli, M., & Volpi, N. (2003). The genome of the lepidopteran *Mamestra brassicae* has a vertebrate-like content of methyl-cytosine. *Genetica*, *119*(2), 187–91.
- Mann, R. S., & Carroll, S. B. (2002). Molecular mechanisms of selector gene function and evolution. *Current Opinion in Genetics & Development*, *12*(5), 592–600.
- Mann, R. S., & Hogness, D. S. (1990). Functional Dissection in *D. melanogaster* of Ultrabithorax Proteins. *Cell*, *60*, 597–610.
- Margueron, R., & Reinberg, D. (2011). The Polycomb complex PRC2 and its mark in life. *Nature*, *469*(7330), 343–9.
- Martin, A., Papa, R., Nadeau, N. J., Hill, R. I., Counterman, B. A., & Halder, G. (2012). Diversification of complex butterfly wing patterns by repeated regulatory evolution of a Wnt ligand. *Proceedings of the National Academy of Sciences*, *109*(31), 12632–12637.
- Martin, A., Papa, R., Nadeau, N. J., Hill, R. I., Counterman, B. A., Halder, G., ... Reed, R. D. (2012). Diversification of complex butterfly wing patterns by repeated regulatory evolution of a Wnt ligand. *Proceedings of the National Academy of Sciences of the United States of America*, *109*(31), 12632–7.
- Martin, A., & Reed, R. D. (2010). Wingless and aristaless2 define a developmental ground plan for moth and butterfly wing pattern evolution. *Molecular Biology and Evolution*, *27*(12), 2864–78.
- Mateus, A. R. A., Marques-pita, M., Oostra, V., Lafuente, E., Brakefield, P. M., & Zwaan, B. J. (2014). Adaptive developmental plasticity : Compartmentalized responses to environmental cues and to corresponding internal signals provide phenotypic flexibility. *BMC Biology*, *12*(97), 1–15.

- McMillan, W. O., Monteiro, A., & Kapan, D. D. (2002). Development and evolution on the wing. *Trends in Ecology and Evolution*, *17*(02), 125–133.
- Milán, M., & Cohen, S. M. (2000). Temporal regulation of apterous activity during development of the *Drosophila* wing. *Development*, *127*(14), 3069–78.
- Minelli, A., & Fusco, G. (2012). On the evolutionary developmental biology of Speciation. *Evolutionary Biology*, *39*, 242–254.
- Miyazawa, S., Okamoto, M., & Kondo, S. (2010). Blending of animal colour patterns by hybridization. *Nature Communications*, *1*, 66.
- Moczek, A. P. (2010). Phenotypic plasticity and diversity in insects. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *365*(1540), 593–603.
- Montagne, J., Groppe, J., Guillemin, K., Krasnow, M. A., Gehring, W. J., & Affolter, M. (1996). The *Drosophila* Serum Response Factor gene is required for the formation of intervein tissue of the wing and is allelic to blistered. *Development*, *122*(9), 2589–97.
- Montague, M. J., Danek-gontard, M., & Kunc, H. P. (2012). Phenotypic plasticity affects the response of a sexually selected trait to anthropogenic noise. *Behavioral Ecology*, *342–348*.
- Monteiro, A. (2012). Gene regulatory networks reused to build novel traits: co-option of an eye-related gene regulatory network in eye-like organs and red wing patches on insect wings is suggested by optix expression. *BioEssays : News and Reviews in Molecular, Cellular and Developmental Biology*, *34*(3), 181–6.
- Monteiro, A. (2014). Origin, Development, and Evolution of Butterfly Eyespots. *Annual Review of Entomology*, *60*, 253–71.
- Monteiro, A., Brakefield, P. M., & French, V. (1994). The evolutionary genetics and developmental basis of wing pattern variation in *Bicyclus anynana*. *Evolution*, *48*(4), 1147–1157.
- Monteiro, A., Brakefield, P. M., & French, V. (1997). Butterfly Eyespots : The Genetics and Development of the Color Rings. *Evolution*, *51*(4), 1207–1216.
- Monteiro, A., Brakefield, P. M., & French, V. (1997). The relationship between eyespot shape and wing shape in the butterfly *Bicyclus anynana* : A genetic and morphometrical approach. *Journal of Evolutionary Biology*, *10*, 787–802.
- Monteiro, A., Chen, B., Ramos, D. M., Oliver, J. C., Tong, X., Guo, M., ... Kamal, F. (2013). Distal-less regulates eyespot patterns and melanization in *Bicyclus*

- butterflies. *Journal of Experimental Zoology. Part B, Molecular and Developmental Evolution*, 320(5), 321–31.
- Monteiro, A., Glaser, G., Stockslager, S., Glansdorp, N., & Ramos, D. (2006). Comparative insights into questions of lepidopteran wing pattern homology, 13, 1–13.
- Monteiro, A., & Podlaha, O. (2009). Wings, horns, and butterfly eyespots: how do complex traits evolve? *PLoS Biology*, 7(2), e37.
- Monteiro, A., Prijs, J., Bax, M., Hakkaart, T., & Brakefield, P. M. (2003). Mutants highlight the modular control of butterfly eyespot patterns. *Evolution & Development*, 5(2), 180–7.
- Morange, M. (2009). How phenotypic plasticity made its way into molecular biology. *Journal of Biosciences*, 34(4), 495–501.
- Müller, J., Hart, C. M., Francis, N. J., Vargas, M. L., Sengupta, A., Wild, B., ... Simon, J. a. (2002). Histone methyltransferase activity of a *Drosophila* Polycomb group repressor complex. *Cell*, 111(2), 197–208.
- Murren, C. J., Pendleton, N., & Pigliucci, M. (2002). Evolution of phenotypic integration in *Brassica* (Brassicaceae). *American Journal of Botany*, 89(4), 655–663.
- Mushegian, A. R., Garey, J. R., Martin, J., & Liu, L. X. (1998). Large-scale taxonomic profiling of eukaryotic model organisms: A comparison of orthologous proteins encoded by the human, fly, nematode, and yeast genomes. *Genome Research*, 8, 590–598.
- Neumann, C. J., & Cohen, S. M. (1996). A hierarchy of cross-regulation involving Notch, wingless, vestigial and cut organizes the dorsal/ventral axis of the *Drosophila* wing. *Development*, 122(11), 3477–85.
- Ng, J., Hart, C. M., Morgan, K., & Simon, J. a. (2000). A *Drosophila* ESC-E(Z) protein complex is distinct from other polycomb group complexes and contains covalently modified ESC. *Molecular and Cellular Biology*, 20(9), 3069–3078.
- Ng, J., Li, R., Morgan, K., & Simon, J. (1997). Evolutionary conservation and predicted structure of the *Drosophila* extra sex combs repressor protein. *Molecular and Cellular Biology*, 17(11), 6663–6672.
- Nichols, R. (2001). Gene trees and species trees are not the same. *Trends in Ecology & Evolution*, 16(01), 358–364.
- Nijhout, H. F. (1980). Pattern formation on lepidopteran wings: Determination of an eyespot. *Developmental Biology*, 80(2), 267–274.

- Nijhout, H. F. (1984). Colour pattern modification by coldshock in Lepidoptera. *Journal of Embryology and Experimental Morphology*, 81, 287–305.
- Nijhout, H. F. (1996). Focus on butterfly eyespot development. *Nature*, 384, 209–210.
- Nijhout, H. F. (2001). Elements of butterfly wing patterns. *The Journal of Experimental Zoology*, 291(3), 213–25.
- Nijhout, H. F. (2003). Development and evolution of adaptive polyphenisms. *Evolution & Development*, 5(1), 9–18.
- Nijhout, H. F., Smith, W. A., Schachar, I., Subramanian, S., Tobler, A., & Grunert, L. W. (2007). The control of growth and differentiation of the wing imaginal disks of *Manduca sexta*. *Developmental Biology*.
- O’Keefe, D. D., & Thomas, J. B. (2001). *Drosophila* wing development in the absence of dorsal identity. *Development*, 128(5), 703–10.
- O’Meara, M. M., & Simon, J. A. (2012). Inner workings and regulatory inputs that control Polycomb repressive complex 2. *Chromosoma*, 121(3), 221–34.
- O’Neill, B. F., Zangerl, A. R., Delucia, E. H., & Berenbaum, M. R. (2010). Olfactory preferences of *Popillia japonica*, *Vanessa cardui*, and *Aphis glycines* for *Glycine max* grown under elevated CO₂. *Environmental Entomology*, 39(4), 1291–301.
- Ohno, Y., & Otaki, J. M. (2012). Eyespot colour pattern determination by serial induction in fish: Mechanistic convergence with butterfly eyespots. *Scientific Reports*, 2, 290.
- Oliver, J. C., Beaulieu, J. M., Gall, L. F., Piel, W. H., Monteiro, A., & B, P. R. S. (2014). Nymphalid eyespot serial homologues originate as a few individualized modules. *Proceedings of the Royal Society. B*, 281, 20133262.
- Oliver, J. C., Ramos, D., Prudic, K. L., & Monteiro, A. (2013). Temporal gene expression variation associated with eyespot size plasticity in *Bicyclus anynana*. *PloS One*, 8(6), e65830.
- Oliver, J. C., Tong, X.-L., Gall, L. F., Piel, W. H., & Monteiro, A. (2012). A single origin for nymphalid butterfly eyespots followed by widespread loss of associated gene expression. *PLoS Genetics*, 8(8), e1002893.
- Oostra, V., Brakefield, P. M., Hiltemann, Y., Zwaan, B. J., & Brattström, O. (2014). On the fate of seasonally plastic traits in a rainforest butterfly under relaxed selection. *Ecology and Evolution*, 4(13), 2654–67.

- Organista, M. F., & De Celis, J. F. (2013). The Spalt transcription factors regulate cell proliferation, survival and epithelial integrity downstream of the Decapentaplegic signalling pathway. *Biology Open*, 2, 37–48.
- Otaki, J. M. (2007). Stress-Induced Color-Pattern Modifications and Evolution of the Painted Lady Butterflies *Vanessa cardui* and *Vanessa kershawi*. *Zoological Science*, 24(8), 811–819.
- Otaki, J. M. (2008a). Phenotypic plasticity of wing color patterns revealed by temperature and chemical applications in a nymphalid butterfly *Vanessa indica*. *Journal of Thermal Biology*, 33(2), 128–139.
- Otaki, J. M. (2008b). Physiological side-effect model for diversification of non-functional or neutral traits: a possible evolutionary history of *Vanessa* butterflies (Lepidoptera, Nymphalidae). *Trans. Lepid. Soc. Japan*, 59(2), 87–102.
- Otaki, J. M. (2008c). Physiologically induced color-pattern changes in butterfly wings: mechanistic and evolutionary implications. *Journal of Insect Physiology*, 54(7), 1099–112.
- Otaki, J. M. (2011). Color-pattern analysis of eyespots in butterfly wings: a critical examination of morphogen gradient models. *Zoological Science*, 28(6), 403–13.
- Otaki, J. M., Ogasawara, T., & Yamamoto, H. (2005a). Morphological comparison of pupal wing cuticle patterns in butterflies. *Zoological Science*, 22(1), 21–34.
- Otaki, J. M., Ogasawara, T., & Yamamoto, H. (2005b). Tungstate-induced color-pattern modifications of butterfly wings are independent of stress response and ecdysteroid effect. *Zoological Science*, 22(6), 635–44.
- Otaki, J. M., & Yamamoto, H. (2004). Color-pattern Modifications and Speciation in Butterflies of the Genus *Vanessa* and its Related Genera *Cynthia* and *Bassaris*. Color-pattern Modifications and Speciation in Butterflies of the Genus *Vanessa* and its Related Genera *Cynthia* and *Bassaris*. *Zoological Science*, 21(9), 967–976.
- Beldade, P., French, V., & Brakefield, P. M. (2008). Developmental and Genetic Mechanisms for Evolutionary Diversification of Serial Repeats : Eyespot Size in *Bicyclus anynana* Butterflies. *Journal of Experimental Zoology. Part B, Molecular and Developmental Evolution*, 201(April 2007), 191–201.
- Panning, B. (2010). Fine-tuning silencing. *Cell Stem Cell*, 6(1), 3–4.
- Pfennig, D. W., & Ehrenreich, Ian, M. (2014). Towards a gene regulatory network perspective on phenotypic plasticity , genetic accommodation and genetic assimilation. *Molecular Ecology*, 23(August), 4438–4440.

- Pfennig, D. W., Wund, M. A., Snell-Rood, E. C., Cruickshank, T., Schlichting, C. D., & Moczek, A. P. (2010). Phenotypic plasticity's impacts on diversification and speciation. *Trends in Ecology & Evolution*, 25(8), 459–67.
- Pigliucci, M. (2001). *Phenotypic Plasticity: Beyond Nature and Nurture*. JHU Press.
- Pigliucci, M. (2005). Evolution of phenotypic plasticity : where are we going now? *Trends in Ecology & Evolution*, 20(9), 481–6.
- Pigliucci, M., Murren, C. J., & Schlichting, C. D. (2006). Phenotypic plasticity and evolution by genetic assimilation. *The Journal of Experimental Biology*, 209(Pt 12), 2362–7.
- Pirrotta, V. (2011). Polycomb Mechanisms and Epigenetic Control of Gene Activity. In T. Tollefsbol (Ed.), *Handbook of Epigenetics. The New Molecular and Medical Genetics*.
- Plaistow, S. J., & Collin, H. (2004). Phenotypic integration plasticity in *Daphnia magna* : an integral facet of G x E interactions. *Journal of Evolutionary Biology*, 27, 1913–1920.
- Posakony, L. G., Raftery, L. A., & Gelbart, W. M. (1990). Wing formation in *Drosophila melanogaster* requires decapentaplegic gene function along the anterior-posterior compartment boundary. *Mechanisms of Development*, 33(1), 69–82.
- Prezioso, C., & Orlando, V. (2011). Polycomb proteins in mammalian cell differentiation and plasticity. *FEBS Letters*, 585(13), 2067–77.
- Price, T. D., Qvarnström, A., & Irwin, D. E. (2003). The role of phenotypic plasticity in driving genetic evolution. *Proceedings. Biological Sciences / The Royal Society*, 270(1523), 1433–40.
- Protas, M. E., & Patel, N. H. (2008). Evolution of coloration patterns. *Annual Review of Cell and Developmental Biology*, 24, 425–46.
- Prudic, K. L., Jeon, C., Cao, H., & Monteiro, A. (2011). Developmental plasticity in sexual roles of butterfly species drives mutual sexual ornamentation. *Science (New York, N.Y.)*, 331(6013), 73–5.
- Prudic, K. L., Skemp, A. K., & Papaj, D. R. (2006). Aposematic coloration, luminance contrast, and the benefits of conspicuousness. *Behavioral Ecology*, 18(1), 41–46.
- Quigley, I. K., Turner, J. M., Nuckels, R. J., Manuel, J. L., Budi, E. H., MacDonald, E. L., & Parichy, D. M. (2004). Pigment pattern evolution by differential deployment of neural crest and post-embryonic melanophore lineages in Danio fishes. *Development*, 131(24), 6053–69.

- Rai, A. N., Vargas, M. L., Wang, L., Andersen, E. F., Miller, E. L., & Simon, J. A. (2013). Elements of the polycomb repressor SU(Z)12 needed for histone H3-K27 methylation, the interface with E(Z), and in vivo function. *Molecular and Cellular Biology*, 33(24), 4844–56.
- Rauskolb, C., Peifer, M., & Wieschaus, E. (1993). extradenticle, a regulator of homeotic gene activity, is a homolog of the homeobox-containing human proto-oncogene pbx1. *Cell*, 74(6), 1101–1112.
- Rauskolb, C., Smith, K. M., Peifer, M., & Wieschaus, E. (1995). Extradenticle determines segmental identities throughout Drosophila development. *Development*, 121(11), 3663–73.
- Reed, R. D., Chen, P.-H., & Frederik Nijhout, H. (2007). Cryptic variation in butterfly eyespot development: the importance of sample size in gene expression studies. *Evolution & Development*, 9(1), 2–9.
- Reed, R. D., McMillan, W. O., & Nagy, L. M. (2008). Gene expression underlying adaptive variation in Heliconius wing patterns: non-modular regulation of overlapping cinnabar and vermilion prepatterns. *Proceedings. Biological Sciences / The Royal Society*, 275(1630), 37–45.
- Reed, R. D., & Nagy, L. M. (2005). Evolutionary redeployment of a biosynthetic module: expression of eye pigment genes vermilion, cinnabar, and white in butterfly wing development. *Evolution & Development*, 7(4), 301–11.
- Reed, R. D., Papa, R., Martin, A., Hines, H. M., Counterman, B. A., Pardo-Diaz, C., ... McMillan, W. O. (2011). Optix Drives the Repeated Convergent Evolution of Butterfly Wing Pattern Mimicry. *Science*, 333(6046), 1137–41.
- Ringrose, L., & Paro, R. (2007). Polycomb/Trithorax response elements and epigenetic memory of cell identity. *Development*, 134(2), 223–32.
- Roskam, J. C., & Brakefield, P. M. (1999). Seasonal polyphenism in Bicyclus (Lepidoptera : Satyridae) butterflies different climates need different cues. *Biological Journal Of the Linnean Society*, 66, 345–356.
- Rountree, D. B., & Nijhout, H. F. (1995). Hormonal control of a seasonal polyphenism in *Precis coenia* (Lepidoptera: Nymphalidae). *Journal of Insect Physiology*, 41(11), 987–992.
- Ryall, R. L., & Howells, A. J. (1974). Ommochrome biosynthetic pathway of *Drosophila melanogaster*: Variations in levels of enzyme activities and intermediates during adult development. *Insect Biochemistry*, 4(1), 47–61.

- Saenko, S. V., French, V., Brakefield, P. M., & Beldade, P. (2008). Conserved developmental processes and the formation of evolutionary novelties: examples from butterfly wings. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, 363(1496), 1549–55.
- Saenko, S. V., Marialva, M. S., & Beldade, P. (2011). Involvement of the conserved Hox gene *Antennapedia* in the development and evolution of a novel trait. *EvoDevo*, 2(1), 9.
- Schlichting, C. D. (1986). The Evolution of Phenotypic Plasticity in Plants. *Annual Review of Ecology and Systematics*, 17(1), 667–693.
- Schlichting, C. D. (1989). Phenotypic plasticity in *Phlox* II . Plasticity of character correlations. *Oecologia*, 78, 496–501.
- Schlichting, C. D., & Smith, H. (2002). Phenotypic plasticity : linking molecular mechanisms with evolutionary outcomes. *Evolutionary Ecology*, 16, 189–211.
- Schmitges, F. W., Prusty, A. B., Faty, M., Stützer, A., Lingaraju, G. M., Aiwazian, J., ... Thomä, N. H. (2011). Histone methylation by PRC2 is inhibited by active chromatin marks. *Molecular Cell*, 42(3), 330–41.
- Schroeder, M. D., Pearce, M., Fak, J., Fan, H., Unnerstall, U., Emberly, E., ... Gaul, U. (2004). Transcriptional control in the segmentation gene network of *Drosophila*. *PLoS Biology*, 2(9).
- Schwartz, Y. B., & Pirrotta, V. (2007). Polycomb silencing mechanisms and the management of genomic programmes. *Nature Reviews. Genetics*, 8(1), 9–22.
- Senthilkumar, R., & Mishra, R. K. (2009). Novel motifs distinguish multiple homologues of Polycomb in vertebrates: expansion and diversification of the epigenetic toolkit. *BMC Genomics*, 10, 549.
- Serfas, M. S., & Carroll, S. B. (2005). Pharmacologic approaches to butterfly wing patterning: sulfated polysaccharides mimic or antagonize cold shock and alter the interpretation of gradients of positional information. *Developmental Biology*, 287(2), 416–24.
- Shaver, S., Casas-Mollano, J. A., Cerny, R. L., & Cerutti, H. (2010). Origin of the polycomb repressive complex 2 and gene silencing by an E(z) homolog in the unicellular alga *Chlamydomonas*. *Epigenetics : Official Journal of the DNA Methylation Society*, 5(4), 301–12.
- Shaw, J. R., Hampton, T. H., King, B. L., Whitehead, A., Galvez, F., Gross, R. H., ... Stanton, B. A. (2014). Natural selection canalizes expression variation of

- environmentally induced plasticity-enabling genes. *Molecular Biology and Evolution*, 31(11), 3002–3015.
- Shen, J., & Dahmann, C. (2005). The role of Dpp signaling in maintaining the *Drosophila* anteroposterior compartment boundary. *Developmental Biology*, 279, 31–43.
- Simon, J. A., & Kingston, R. E. (2013). Occupying chromatin: Polycomb mechanisms for getting to genomic targets, stopping transcriptional traffic, and staying put. *Molecular Cell*, 49(5), 808–24.
- Simon, J., Bornemann, D., Lunde, K., & Schwartz, C. (1995). The extra sex combs product contains WD40 repeats and its time of action implies a role distinct from other Polycomb group products. *Mechanisms of Development*, 53, 197–208.
- Simpson, S. J., Sword, G. A., & Lo, N. (2011). Polyphenism in insects. *Current Biology* : *CB*, 21(18), R738–49.
- Snell-rod, E. C., Dyken, J. D. Van, Cruickshank, T., Wade, M. J., & Moczek, A. P. (2010). Toward a population genetic framework of developmental evolution: the costs, limits, and consequences of phenotypic plasticity. *BioEssays*, 32(1), 71–81.
- Sparmann, A., & van Lohuizen, M. (2006). Polycomb silencers control cell fate, development and cancer. *Nature Reviews. Cancer*, 6(11), 846–56.
- Srinivasan, D. G., & Brisson, J. A. (2012). Aphids: a model for polyphenism and epigenetics. *Genetics Research International*, 2012, 431531.
- Stefanescu, C., Alarcón, M., & Avila, A. (2007). Migration of the painted lady butterfly, *Vanessa cardui*, to north-eastern Spain is aided by African wind currents. *The Journal of Animal Ecology*, 76(5), 888–98.
- Stern, S., Fridmann-Sirkis, Y., Braun, E., & Soen, Y. (2012). Epigenetically heritable alteration of fly development in response to toxic challenge. *Cell Reports*, 1(5), 528–42.
- Stevens, M., Hardman, C. J., & Stubbins, C. L. (2008). Conspicuousness , not eye mimicry , makes ““ eyespots ”” effective antipredator signals. *Behavioral Ecology*, 525–531.
- Stoehr, A. M., Walker, J. F., & Monteiro, A. (2013). Spalt expression and the development of melanic color patterns in pierid butterflies. *EvoDevo*, 4(1), 6.
- Struhl, G., & Akam, M. (1985). Altered distributions of Ultrabithorax transcripts in extra sex combs mutant embryos of *Drosophila*. *The EMBO Journal*, 4(12), 3259–64.

- Surridge, A. K., Lopez-Gomollon, S., Moxon, S., Maroja, L. S., Rathjen, T., Nadeau, N. J., ... Jiggins, C. D. (2011). Characterisation and expression of microRNAs in developing wings of the neotropical butterfly *Heliconius melpomene*. *BMC Genomics*, *12*(1), 62.
- Swofford, D. L. (1993). PAUP*: Phylogenetic Analysis Using Parsimony.
- Takayama, E., & Yoshida, A. (1997). Color pattern formation on the wing of a butterfly, *Pieris rapae*. 1. Cautery induced alteration and scale color and delay of arrangement formation. *Development, Growth, Differentiation*, *39*, 23–31.
- Talavera, G., & Castresana, J. (2007). Improvement of phylogenies after removing divergent and ambiguously aligned blocks from protein sequence alignments. *Systematic Biology*, *56*(4), 564–77.
- Tie, F., Furuyama, T., & Harte, P. J. (1998). The *Drosophila* Polycomb Group proteins ESC and E(Z) bind directly to each other and co-localize at multiple chromosomal sites. *Development*, *125*, 3483–3496.
- Tong, X., Hrycaj, S., Podlaha, O., Popadic, A., & Monteiro, A. (2014). Over-expression of Ultrabithorax alters embryonic body plan and wing patterns in the butterfly *Bicyclus anynana*. *Developmental Biology*, *394*(2), 357–66.
- Tong, X., Lindemann, A., & Monteiro, A. (2012). Differential Involvement of Hedgehog Signaling in Butterfly Wing and Eyespot Development. *PLoS ONE*, *7*(12), e51087.
- Umemori, M., Takemura, M., Maeda, K., Ohba, K., & Adachi-Yamada, T. (2007). *Drosophila* T-box transcription factor Optomotor-blind prevents pathological folding and local overgrowth in wing epithelium through confining Hh signal. *Developmental Biology*.
- Vallin, A., Jakobsson, S., Lind, J., & Wiklund, C. (2006). Crypsis versus intimidation - Anti-predation defence in three closely related butterflies. *Behavioral Ecology and Sociobiology*, *59*, 455–459.
- Van Belleghem, S. M., Roelofs, D., Van Houdt, J., & Hendrickx, F. (2012). De novo transcriptome assembly and SNP discovery in the wing polymorphic salt marsh beetle *Pogonus chalceus* (Coleoptera, Carabidae). *PloS One*, *7*(8), e42605.
- Van der Velden, Y. U., Wang, L., van Lohuizen, M., & Haramis, A.-P. G. (2012). The Polycomb group protein Ring1b is essential for pectoral fin development. *Development*, *139*(12), 2210–20.
- Waddington, C. H. (1942). Canalization of Development and the Inheritance of Acquired Characters. *Nature*, *150*, 563–565.

- Wahlberg, N., & Rubinoff, D. (2011). Vagility across Vanessa (Lepidoptera: Nymphalidae): Mobility in butterfly species does not inhibit the formation and persistence of isolated sister taxa. *Systematic Entomology*, 36, 362–370.
- Walters, J. W., Muñoz, C., Paaby, A. B., & DiNardo, S. (2005). Serrate-Notch signaling defines the scope of the initial denticle field by modulating EGFR activation. *Developmental Biology*, 286, 415–426.
- Warren, William, D., Palmer, S., & Howells, Anthony, J. (1996). Molecular characterization of the cinnabar region of *Drosophila melanogaster* : Identification of the cinnabar transcription unit. *Genetica*, 98, 249–262.
- Waterland, R. A., & Jirtle, R. L. (2003). Transposable Elements : Targets for Early Nutritional Effects on Epigenetic Gene Regulation Transposable Elements : Targets for Early Nutritional Effects on Epigenetic Gene Regulation. *Molecular and Cellular Biology*, 23(15), 5293–5300.
- Weatherbee, S. D., Nijhout, H. F., Grunert, L. W., Halder, G., Galant, R., Selegue, J., & Carroll, S. (1999). Ultrabithorax function in butterfly wings and the evolution of insect wing patterns. *Current Biology : CB*, 9(3), 109–115.
- Weiner, S. A., & Toth, A. L. (2012). Epigenetics in social insects: a new direction for understanding the evolution of castes. *Genetics Research International*, 2012, 609810.
- Werner, T., Koshikawa, S., Williams, T. M., & Carroll, S. B. (2010). Generation of a novel wing colour pattern by the Wingless morphogen. *Nature*, 464(7292), 1143–8.
- West-Eberhard, M. J. (1989). Phenotypic plasticity and the origins of diversity. *Annual Review of Ecology and Systematics*, 20, 249–78.
- West-Eberhard, M. J. (2005). Developmental plasticity and the origin of species differences. *Proceedings of the National Academy of Sciences of the United States of America*, 102 Suppl (2), 6543–6549.
- West-eberhard, M. J. (2005). Phenotypic accomodation: adaptive innovation due to developmental plasticity. *Journal of Experimental Zoology*, 304B(September), 610–618.
- Whitcomb, S. J., Basu, A., Allis, C. D., & Bernstein, E. (2007). Polycomb Group proteins: an evolutionary perspective. *Trends in Genetics : TIG*, 23(10), 494–502.
- Willert, K., & Nusse, R. (2012). Wnt proteins. *Cold Spring Harbor Perspectives in Biology*, 4(9), a007864.

- Williams, J. A., Paddock, S. W., & Carroll, S. B. (1993). Pattern formation in a secondary field : a hierarchy of regulatory genes subdivides the developing *Drosophila* wing disc into discrete subregions, *584*, 571–584.
- Wilson, B. G., Wang, X., Shen, X., McKenna, E. S., Lemieux, M. E., Cho, Y. J., ... Roberts, C. W. M. (2010). Epigenetic antagonism between polycomb and SWI/SNF complexes during oncogenic transformation. *Cancer Cell*, *18*(4), 316–328.
- Wittkopp, P. J., & Beldade, P. (2009). Development and evolution of insect pigmentation: genetic mechanisms and the potential consequences of pleiotropy. *Seminars in Cell & Developmental Biology*, *20*(1), 65–71.
- Wittkopp, P. J., Carroll, S. B., & Kopp, A. (2003). Evolution in black and white: genetic control of pigment patterns in *Drosophila*. *Trends in Genetics : TIG*, *19*(9), 495–504.
- Wittkopp, P. J., True, J. R., & Carroll, S. B. (2002). Reciprocal functions of the *Drosophila* yellow and ebony proteins in the development and evolution of pigment patterns. *Development*, *129*(8), 1849–58.
- Wittkopp, P. J., Vaccaro, K., & Carroll, S. B. (2002). Evolution of yellow gene regulation and pigmentation in *Drosophila*. *Current Biology : CB*, *12*(18), 1547–56.
- Wolpert, L. (2003). Cell boundaries: knowing who to mix with and what to shout or whisper. *Development (Cambridge, England)*, *130*(19), 4497–500.
- Wray, G. A., Hahn, M. W., Abouheif, E., Balhoff, J. P., Pizer, M., Rockman, M. V., & Romano, L. a. (2003). The evolution of transcriptional regulation in eukaryotes. *Molecular Biology and Evolution*, *20*(9), 1377–419.
- Wund, M. A. (2012). Assessing the impacts of phenotypic plasticity on evolution. *Integrative and Comparative Biology*, *52*(1), 5–15.
- Xia, A.-H., Zhou, Q.-X., Yu, L.-L., Li, W.-G., Yi, Y.-Z., Zhang, Y.-Z., & Zhang, Z.-F. (2006). Identification and analysis of YELLOW protein family genes in the silkworm, *Bombyx mori*. *BMC Genomics*, *7*, 195.
- Xiang, H., Zhu, J., Chen, Q., Dai, F., Li, X., Li, M., ... Wang, J. (2010). Single base-resolution methylome of the silkworm reveals a sparse epigenomic map. *Nature Biotechnology*, *28*(5), 516–20.
- Yamamoto, K., Sonoda, M., Inokuchi, J., Shirasawa, S., & Sasazuki, T. (2004). Polycomb Group Suppressor of Zeste 12 Links Heterochromatin Protein 1 α and Enhancer of Zeste 2. *Journal of Biological Chemistry*, *279*(1), 401–406.
- Yan, D., & Lin, X. (2009). Shaping morphogen gradients by proteoglycans. *Cold Spring Harbor Perspectives in Biology*, *1*(3), a002493.

- Zecca, M., Basler, K., & Struhl, G. (1995). Sequential organizing activities of engrailed, hedgehog and decapentaplegic in the *Drosophila* wing. *Development*, *121*, 2265–2278.
- Zeng, J., Kirk, B. D., Gou, Y., Wang, Q., & Ma, J. (2012). Genome-wide polycomb target gene prediction in *Drosophila melanogaster*. *Nucleic Acids Research*, *40*(13), 5848–63.