Elucidating the Genetic Basis of Pigmentation Differences between *Drosophila* Species

by

Lisa L. Sramkoski

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Molecular, Cellular, and Developmental Biology)
In the University of Michigan
2012

Doctoral Committee:

Associate Professor Patricia J. Wittkopp, Chair
Associate Professor Scott E. Barolo
Associate Professor Kenneth M. Cadigan
Associate Professor Gyorgyi Csankovszki
Associate Professor Catherine E. Krull

To: My husband & parents who have encouraged & supported
this and so many other life adventures, and
to my daughter, may she explore the world
by following her own unique path of her choosing.

Acknowledgements

There are many people who should be acknowledged for their support and assistance that made the research included herein possible. First and foremost, I owe a tremendous amount of gratitude to my advisor, Trisha Wittkopp. Over the last several years, she has challenged and supported me an enormous amount both professionally and personally. She has been an incredible mentor in so many aspects, I cannot begin to even describe them all, so simply put for *everything*, thank you!

The members of the Wittkopp lab have created a scientifically rich environment that lends itself well to fostering problem-solving and theoretical discussions with a variety of expertise represented. These discussions and assistance were instrumental in being able to successfully complete the experiments described in the subsequent chapters. Specific thanks is given to Wesley McLaughlin, Arielle Cooley, Kraig Stevenson, Dave Yuan, Emma Stewart, Laura Shefner, Elizabeth Thompson, Elizabeth Walker, and Adam Neidert for their contributions to the projects included in this thesis.

I was also fortunate enough to have a very talented and knowledgable cohort member that has assisted on innumerable occasions troubleshooting experiments, lending reagents and equipment as well as discussing ideas: the elusive Gish Fish. She not only provided professional support, but has been a

tremendous friend, providing a listening ear, thoughtful advice, as well as countless (and necessary) laughs. You have been essential to the process.

During my graduate career, two faculty members really helped to inspire and encourage my educational ambitions, Jo Kurdziel and Laura Olsen. Working with Jo as a GSI really sparked my desire for teaching. I was able to learn an enormous amount from her expertise during our time working together. My enthusiasm for the classroom was then stirred by the Biology Educational Journal Club organized by both Jo and Laura that provided a foundation for knowledge in biology educational research. I owe my involvement in the Master's program for post-secondary science education to Laura who first told me about the program and its potential. I cannot express enough gratitude for these wonderful opportunities as they have shaped my career goals and aspirations.

I also am indebted to my family. My parents have lent a great deal of support in so many ways while completing this degree. It certainly would not have been possible without them. They have always encouraged me to follow my own path, no matter where it has led. The path has certainly been long and curvy all the way from the Creek, but also an exceptional adventure. Thank you for such unlimited opportunities.

Additionally, my extraordinary husband and daughter have been fantastically loving, patient, supportive and helpful during my time completing this degree. I am appreciative beyond words for everything over the years. You have made an excellent "home" team and I feel honored to be part of The Sram Fam!

Lastly, I would like to thank all other family and friends not mentioned explicitly. So many have offered a helpful and supportive word or gesture over the last several years. No matter the size, big or small, every one of them made a difference to me in one way or another.

Table of Contents

List of Figures

Figure

List of Tables

Table

ABSTRACT

Elucidating the Genetic Basis of Pigmentation Differences between *Drosophila* Species

by

Lisa L. Sramkoski

Chair: Patricia J. Wittkopp

Understanding how phenotypes change has been a long-standing question in evolutionary biology. The research presented here investigates the genetic basis of the pigmentation difference exhibited between the dark pigmented *Drosophila americana* and lightly pigmented *Drosophila novamexicana.* It shows that the pigmentation genes *ebony* and *tan* contribute to this pigmentation divergence between these species and that *tan* mRNA is more highly expressed in *D. americana*. Variation in pigmentation within *D. americana* is also documented and is shown to be due in part to the same genes: *ebony* and *tan.* The contribution of *tan* to the pigmentation differences between species is further examined, and the 3' region of *tan* intron 1 is determined to likely contain some of the causative nucleotides through functional transgenic analyses. The differences exhibited among *D. americana* strains was quantified and found to vary longitudinally with the darkest flies found in the eastern U.S. and lighter

variants found in western populations of the U.S. This variation appears to be a consequence of local adaptation among *D. americana* populations. Desiccation resistance was investigated as a potential ultimate cause for the observed pigmentation differences, but found not to be associated with pigmentation. Two populations of *D. americana* were identified to contain *D. novamexicana*-like alleles of either *tan* or *ebony.* A survey of 51 populations was conducted to examine the frequency of this phenomena. No additional populations were found to contain *D. novamexicana*-like alleles, although genetic heterogeneity is identified among *D. americana* strains. This research presents a case of how phenotypic evolution has occurred by identifying genetic contributors and specific DNA sequence associated with this change in pigmentation both between species and within species.

Research on the effectiveness of a specific pedagogical technique, online quizzes, in a large genetics course is also included in this thesis that resulted from participation in the Post-Secondary Science Education Master's degree program. Although quizzes were not found to be statistically associated with increased student performance, students and instructor expressed positive sentiments towards the quizzes. This illustrates how classroom techniques can improve student attitude even if student performance is not directly impacted.

Chapter 1

Introduction:

Understanding Evolutionary Mechanisms and How Phenotypes are Formed

Evolution can act in all populations of living organisms from microscopic bacteria, plants, insects such as fruit flies and butterflies to mammals such as humans and mice. It is the process of evolution that has created the wide array of fantastic diversity found around the globe. Understanding how the phenotypes of organisms change has long been a curiosity of evolutionary biologists.

Phenotypic evolution refers to how physical traits of an organism change over evolutionary time. Understanding how traits evolve can be studied at two levels, either by examining ultimate or proximate causes of phenotypic changes.

An ultimate cause of phenotypic change refers to the natural force for why a phenotype might change (or the evolutionary forces at work) (HOFREITER and SCHÖNEBERG 2010). For example, in the classic example of the peppered moth, moth coloration became quite dark among populations to match the darkened tree bark (KETTLEWELL 1961). The ultimate cause in this instance being selection on individuals that were concealed visually via camouflage to increase moth survival from predators.

The proximate cause of phenotypic change refers to the underlying molecular mechanisms for an organism's phenotypic modification (HOFREITER

and SCHÖNEBERG 2010). This could be many various possibilities including a mutation in the DNA sequence that results in an amino acid change in a protein (coding change) or sequence important for regulation of a gene important for a particular phenotype (regulatory change). Continuing with the peppered moth as an example, the underlying genetic change between the wild type light-colored morph and derived dark-colored morph has been largely unknown until recently. A 200 kb region has been identified to likely contain the mutation(s) responsible for this redesign of pigmentation in United Kingdom populations of the peppered moth (although the specific gene(s) have yet to be determined) (VAN'T HOF *et al.* 2011). Understanding these proximate causes of how phenotypes change can provide insights into how evolutionary mechanisms act at the sequence level.

There is much to be learned from investigations on the ultimate and proximate causes of phenotypic change. Although each can provide insights into evolutionary processes, a complete picture is only obtained when both aspects are investigated (NACHMAN *et al.* 2003; PROTAS *et al.* 2006; HOEKSTRA *et al.* 2006). Again with the peppered moth, not only establishing the environmental cause for why the dark-colored morphs rose in frequency among populations, but also elucidating the underlying genetics has led to a deeper understanding of how melanin pattern formation occurs in lepidopteran species.

Understanding how novel phenotypes arise in individual species or across multiple species can provide valuable insight into how evolutionary mechanisms can create so many varieties of organisms. There are several traits that have undergone study such as trichome patterning and bristle length in fruit flies,

morphological changes in stickleback fish, branch length in plants, among many others (CARROLL 2005; KOPP 2009), but one phenotype that has received particular attention and holds great potential to answer many evolutionary questions is pigmentation patterns of organisms.

Pigmentation studies can enhance our understanding of phenotypic evolution

Pigmentation presents an ideal phenotypic trait to study in a variety of organisms. Many species (and closely related species) hold a spectacular array of natural pigmentation variation. In fact, pigmentation is one of the most variable traits among organisms, providing innumerable examples to study how phenotypic change occurs. Many studies have been conducted examining the pigmentation patterns in vertebrates such as mammals and fish (QUIGLEY and PARICHY 2002; SUGIE *et al.* 2004; NACHMAN 2005; CARO 2005; PROTAS *et al.* 2006; HOEKSTRA *et al.* 2006; PARICHY 2006; STEINER *et al.* 2007; PROTAS and PATEL 2008; GREENWOOD *et al.* 2011), within plants (MARTIN and GERATS 1993; HOLTON and CORNISH 1995; WINKEL-SHIRLEY 2001; SPELT *et al.* 2002; KOES *et al.* 2005; COOLEY and WILLIS 2009) , and invertebrates such as *Drosophila* (KOPP and DUNCAN 2002; GOMPEL *et al.* 2005; PRUD'HOMME *et al.* 2006; POOL and AQUADRO 2007; JEONG *et al.* 2008; WITTKOPP and BELDADE 2009; REBEIZ *et al.* 2009a).

There are multiple roles for pigmentation to play that fundamentally either increase chances of survival or reproduction for individuals. Survival mechanisms

of crypsis or camoflauge allow an organism to blend in with its background and remain hidden from predators, as demonstrated by the well-studied peppered moth that adapted to the darkened backgrounds on lichened tree trunks, rock, or dead wood resulting from the industrial revolution with a darker melanic form (KETTLEWELL 1961; 1965; LEES 1968; COOK *et al.* 2012). Additionally, an organism's coloration can be an aposematic warning to predators of distasteful or toxic biochemicals. Populations of the strawberry poison frog of Panama, *Dendrobates pumilio*, have been surveyed to have color patterns in a strong association with toxicity (MAAN 2012). Pigmentation can also serve as a means to regulate body temperature of some organisms. For example, the two-spotted ladybird, exhibits a melanic form that has been found to have a higher body temperature than its non-melanic form (JONG *et al.* 1996). Sexual selection can also drive diverse pigmentation patterns that lead to reproductive success among individuals. For instance, female choice has helped to shape the diverse array of pigmentation patterns seen in butterflies, particularly in *Pieris occidentalis*, where more marginal forewing melanization in males leads to increased mating success (WIERNASZ 1989). Other butterflies, such as the *Bicyclus* species show differing rates of evolution for characters (eyespots and other pigmentation patterns) on the dorsal forewing between sexes versus other areas of the wing (OLIVER *et al.* 2009). These are to name just a few of the roles pigmentation can fulfill. There are still other functions of pigmentation not mentioned here, but reviewed elsewhere (TRUE 2003; WITTKOPP *et al.* 2003a; CARO 2005; KRONFORST *et al.* 2012).

Pigmentation is clearly an important trait with such a variety of selection pressures available. Coloration patterns are easily visible on all organisms that exhibit them. Additionally, there is a multitude of pigmentation diversity among not only distantly related species, but also closely related species in the natural world. This allows phenotypic evolution to be studied at many divergent levels. The numerous functions of pigmentation among organisms, its ease of visibility and measurement, and the great array of pigmentation diversity makes pigmentation an ideal phenotype for study.

A few groups that exemplify this pigmentation diversity, such as mammals, fish, plants and flies, are reviewed below. Pigmentation is also being researched in other groups that display a variety of pigmentation patterns including moths, butterflies, birds and a more developing system in lizards (KETTLEWELL 1961; GRANT *et al.* 1996; GRANT 2004; ROSENBLUM 2006; JORON *et al.* 2006; NADEAU *et al.* 2007; PAPA *et al.* 2008; BAXTER *et al.* 2008; STUART-FOX and MOUSSALLI 2009; ROSENBLUM *et al.* 2010; KRONFORST *et al.* 2012).

In order to understand how different pigmentation patterns form across different species, there are essentially two main factors to consider: what pigments are available to use and how those pigments are distributed. Research pertaining to these elements and the evolutionary forces that shape populations is reviewed here.

A detailed review of pigmentation in mammals, fish, plants and flies is provided below. Although each system discussed in this review illustrates a varying level of known environmental and genetic components, every one

demonstrates the multiple ways in which color patterns can be formed across many taxa and the wealth of knowledge to be gained on how evolutionary, genetic and cellular processes interact. As the research in the reminder of this thesis is specific to elucidating the genetic mechanisms in flies, more details are provided for the fly system.

Pigmentation in mammals

Mammalian species display a tremendous amount of coat color variation. This diversity in pigmentation has interested humans as early as the eighteenth century with selective mouse breeding (PROTAS and PATEL 2008) and continues to be both a captivating and tractable system to examine pigmentation changes across species.

Among mammals, there are a few adaptive mechanisms for pigmentation, including concealment, communication and regulation of physiological processes (PROTAS and PATEL 2008). Concealment, which includes crypsis, has been argued to be the most prominent selective pressure for mammal pigmentation, although in some instances it is argued whether thermoregulation versus crypsis is truly the selective pressure at work (CARO 2005). Further research across multiple species is needed for a broader understanding of the adaptive significance of mammalian pigmentation.

Over the course of the last century, more than one hundred genes (with only ~60 being cloned) have been identified as playing a role in mammalian pigmentation (BENNETT and LAMOREUX 2003). Due to the complexity of such a

system, it is not surprising that research still has much to discover about the regulation of mammalian pigmentation. Despite this challenge however, scientists have obtained a substantial understanding of the pigmentation processes, with recent research focusing on two particular genes (*MC1R* and *Agouti*-see below). This research holds great value as it has not only provided explanations for many of the colorations and patterns found among mammalian species and underlying evolutionary processes, but it has also led to insights on the genetic basis the phenotypic variation in human skin, hair and eye color (SULEM *et al.* 2007), as well as providing valuable knowledge used to combat human diseases such as melanoma (KABBARAH and CHIN 2006).

Mammalian (Mouse) pigmentation biosynthesis

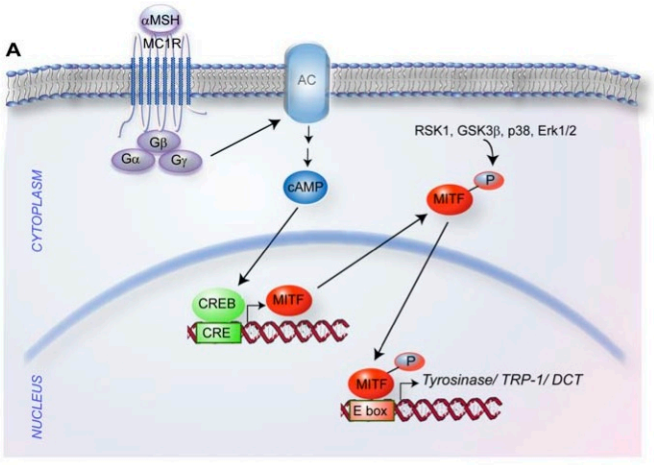
Mammalian pigmentation biosynthesis has primarily been elucidated using the model system of mice; therefore, the following discussion is grounded in this research.

In mice, pigment is formed from specialized cells called melanocytes. These cells (except those found in the pigmented retinal epithelium or PRE) are formed in and migrate from the neural crest through the dermis and then move into the epidermis where they enter the developing hair bulb (JACKSON 1994). Melanocytes secrete pigment granules into growing hairs or neighboring keratinocytes. Two different types of melanin pigments are produced from melanocytes in the hair follicle: eumelanin, which is dark melanin (brown or black

in color) and pheomelanin, which is a lighter (yellow or red in color). (Figure 1.1) (JACKSON 1994; JACKSON *et al.* 1994; NACHMAN *et al.* 2003)

Coat color is determined primarily by the interaction of two main proteins, Melanocortin-1-receptor (MC1R) and Agouti (NACHMAN *et al.* 2003; HOEKSTRA and NACHMAN 2003). MC1R is typically activated by the peptide hormone, α-Melanocyte Stimulating Hormone (α-MSH), which then leads to increased eumelanin production (NACHMAN *et al.* 2003; HOEKSTRA and NACHMAN 2003). MC1R signaling acts to produce the dark eumelanin, where as when MC1R is inhibited, the yellow pheomelanin is produced. The protein encoded by *agouti* acts as an antagonist of eumelanin production by blocking the action of α-MSH, thereby inhibiting MC1R activation and leads to pheomelanin development (JACKSON *et al.* 1994; NACHMAN *et al.* 2003; HOEKSTRA and NACHMAN 2003).

Figure 1.1: MC1R signaling pathway & melanin synthesis



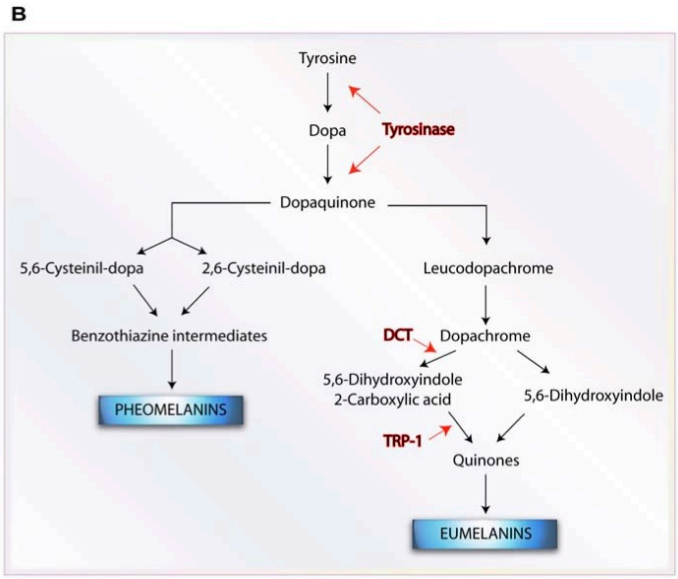


Figure 1.1: (A.) The MC1R pathway. Binding of α-MSH to MC1R activates adenylate cyclase (AC) through a heterotrimeric G-protein complex. AC catalyzes the production of cAMP, leading to the phosphorylation of members of the CREB family of transcription factors, leading to transcriptional activation of multiple genes, including *MITF*, which is a critical CREB target gene in melanocytes. Activated Mitf regulates other genes important for pigmentation and differentiation. (B) The biochemical pathways leading to pheomelanin and eumelanin synthesis. *Figure adapted from (LOMAS *et al.* 2008).

MC1R signaling ultimately leads to the transcription of Tyrosinase and other proteins involved with the production of melanin (Figure 1.1) (LOMAS *et al.* 2008). Tyrosinase has long been known to play a critical role in melanin production in mammals and was found to be encoded by the *albino* locus wherein a homozygous mutant mouse at this locus is incapable of any melanin synthesis and is white with pink eyes (JACKSON 1994).

MC1R has been implicated in many natural mouse populations displaying pigmentation differences. One example includes rock pocket mice, which are

found in the deserts of the southwestern United States and northern Mexico. Many populations live among the sandy desert rocks and have a light tan color, whereas other populations have adapted to live on rocky outcrops of harden molten lava flows and are more melanic providing concealment for protection from predators (Figure 1.2A). (NACHMAN *et al.* 2003) Coding mutations in *MC1R* were identified as causative sites for the adaptive color change in one Arizona population of rock pocket mice (NACHMAN *et al.* 2003). However, when separate dark-colored populations from New Mexico were surveyed, there was no association with the same *MC1R* mutations, but new polymorphisms in *MC1R* were identified, although had no significant association with coat color (NACHMAN *et al.* 2003; HOEKSTRA and NACHMAN 2003). These data show that this dark coloration has evolved independently across multiple populations via different genes.

Figure 1.2: Pigmentation of rock pocket mice & *Peromyscus polionotus* mice

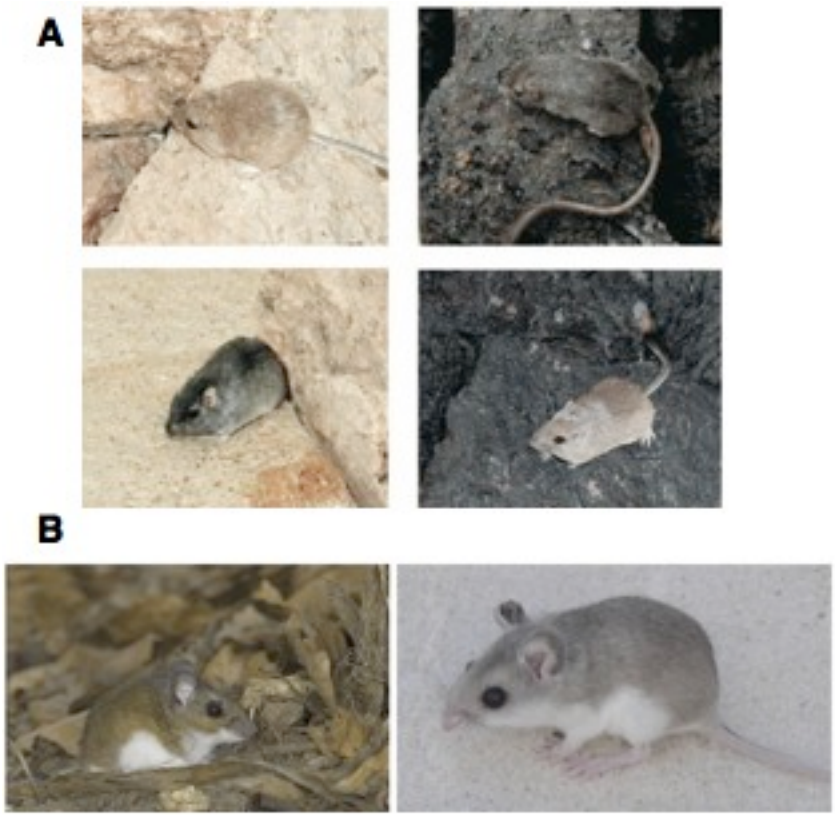


Figure 1.2: (A) Light and dark *C. intermedius* (rock pocket mice) from the Pinacate locality on light and dark rocks. *Adapted from (NACHMAN *et al.* 2003). (B) Color polymorphism in *Peromyscus poliotiotus* (mainland mouse and beach mouse) . *Adapted from (ARENDT and REZNICK 2008).

Natural populations of beach mice (*Peromyscus polionotus leucocephalus)*, have also displayed an association of coat color with *MC1R*. Compared with the mainland mice (*Peromyscus polionotus subgriseus),* some beach mouse populations display a distinct pigmentation difference (Figure 1.2B). These light-colored populations were found to have a single coding mutation in *MC1R* that reduces ligand binding and G protein coupling, leading to

a decrease in cAMP formation and thus, less of the dark eumelanin (HOEKSTRA *et al.* 2006). However, this *MC1R* mutation has only been shown to impact pigmentation in a genetic background also containing a derived *Agouti* allele, which increases *Agouti* mRNA expression (STEINER *et al.* 2007). As an antagonist of MC1R activity, *Agouti* leads to the formation of the light colored pheomelanin (JACKSON 1994). Consequently, the color pattern demonstrated in the light beach mouse populations is a result of the interplay between the changes in protein expression of MC1R and the regulatory modifications of *Agouti* (STEINER *et al.* 2007), illustrating that pigmentation production can be a fine balance among all possible molecular players and the underlying mechanisms may not be as easy as switching one gene on or off to create a single coloration pattern.

Studies have shown that distinct mutations in *MC1R* and *Agouti* have led to pigmentation differences in many mammalian species, illustrating that a similar genetic toolkit has been used multiple times to create new phenotypes. Ongoing research will help identify additional loci involved in melanin production that may be associated with coloration change, such as the *K* locus, which codes for an alternative ligand of MC1R, identified in wolf coat color (ANDERSON *et al.* 2009). Determining other loci important for pigmentation can help as research further investigates natural populations and ultimate causes for phenotypic change.

Pigmentation in fish

Some of the most vibrant colors and patterns can be found among teleost fish. Pigmentation patterns in fish have been long studied in the zebrafish (*Danio rerio)* and its relatives. *D. rerio* displays prominent horizontal stripes, whereas other relatives exhibit spots, vertical stripes, a mixture of stripes and spots or no pattern at all (Figure 1.3A) (PARICHY 2006). Similarly, other species such as cichlids also exhibit an extraordinary array of bright pigmentation and patterns that have also been investigated (Figure 1.4) (WEBBER *et al.* 1973; SUGIE *et al.* 2004; PROTAS and PATEL 2008; GARCÍA 2011).

Pigment patterns in teleost fish have been associated with a variety of evolutionary pressures, including shoaling behavior, camouflage and warning coloration and mate recognition and mate choice (PARICHY 2006). In guppies, *Poecilia reticulata,* male pigmentations have been shaped from a balance between sexual selection (female choice) and crypsis. Males have a number of strategies to remain inconspicuous to predators while being strikingly distinct to females among particular backgrounds. (ENDLER 1983) Male guppies display such a variety of pigmentation, in fact, that no two males are deemed to be the same (ENDLER 1983). The notable pigmentation diversity among cichlids is also important for crypsis, species-recoginition, and mate-choice (SUGIE *et al.* 2004).

Figure 1.3: Pigmentation diversity in *Danio* fish & mutant phenotypes

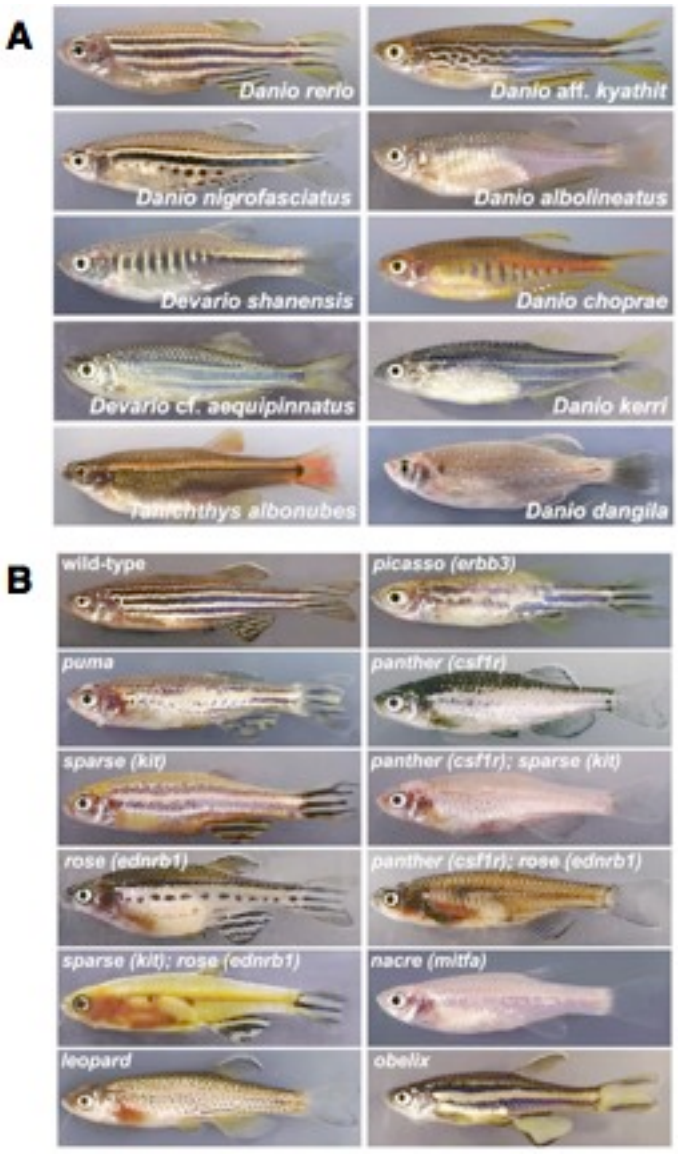


Figure 1.3: (A.) *Danio* pigment pattern diversity. Columns show phenotypes of several species within *Danio*, the closely related *Devario*, and more distant *Tanichthys*. The group encompasses a range of stripes, bars, uniform, and more complex pigment patterns. (B) Pigment pattern mutants within *D. rerio*. Mutant names are shown along with gene identities in parentheses when known. *Figures adapted from (PARICHY 2006).

Figure 1.4: Pigmentation diversity exhibited in African cichlid species

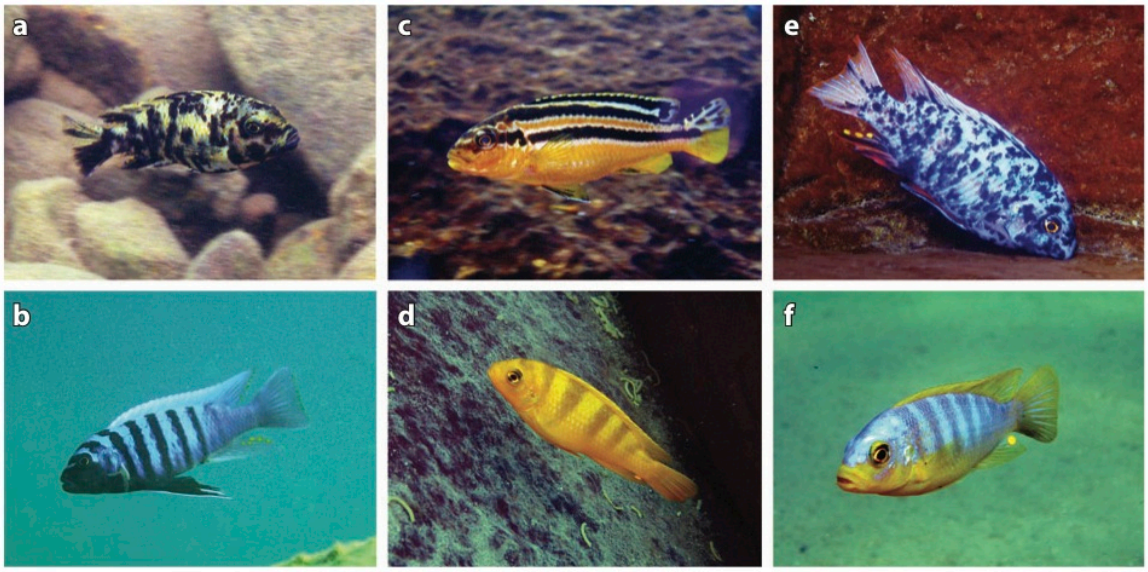


Figure 1.4: Coloration differences in cichlids. In the Great Lakes of East Africa, there are almost 2000 species of cichlids that have evolved relatively recently with a wide variety of color patterns. (A) Orange blotch phenotype in a female *Metriaclima zebra*. (B) Blue-with- black-bar phenotype in a male *M. zebra*. (C) *Metriaclima aurora*. (D) *Metriaclima lombardoi*. (E) *Labeotropheus fuelleborni*. (F) *Metriaclima auratus*. *Figure adapted from (PROTAS and PATEL 2008).

Formation of pigment patterns in teleost fish

In contrast to other vertebrates such as mammals that have one type of pigment cell, teleost fish have several classes of pigment cells or chromatophores. These include the black melanophores (containing melanin), yellow or orange xanthophores (containing pteridines and carnotenoids) and silvery iridophores (containing guanine-rich reflective platelets) (PARICHY 2003). There are also red erythrophores, but these are only present in some *Danio* species, but not others (PARICHY 2006). All of these classes of chromatophores however, are derived from neural crest cells, similarly to mammals. The pigments

of these chromatophores remain packaged in these cells in special organelles, whereas, melanocytes of mammals transfer pigments to hairs. The coloration patterns exhibited by teleost fish are therefore the spatial organization of the chromatophores themselves. (PARICHY 2006)

The interactions of these chromatophores can be quite important for patterning development. In fact, the initial formation and maintenance of the stripes in *D. rerio* has been shown to be dependent upon an interaction between xanthophores and melanophores; in mutants with decreased or no xanthophores, stripes fail to form, but are rescued when xanthophores are restored (PARICHY 2006). The molecular mechanisms at play during these cellular interactions have yet to be identified.

Fish can have varying pigmentation patterns between larval and adult developmental stages. Often these patterns shift during the metamorphosis phase and develop into the adult coloration (PARICHY 2003; KELSH 2004; PARICHY 2006; KONDO *et al.* 2009). Mutations identified in *D. rerio* have been instrumental in determining the genes and mechanisms involved in pigment patterning (Figure 1.3B) (PARICHY 2006; GREENWOOD *et al.* 2011). In *D. rerio,* adult melanophores that are used to form the stripes differentiate during the metamorphosis stage from latent precursors, whereas *D. nigrofasciatus* re-utilizes embryonic and early larval melanophores to form the stripes (PARICHY 2006). This developmental divergence is thought to be caused by a gene(s) in the *puma* pathway (although molecular identification of *puma* is required before this can be tested) (PARICHY 2006). Most *Danio* species, however, exhibit the patterning process of *D. rerio,*

where adult melanophores are primarily formed during metamorphosis , suggesting the re-employment of embryonic and larval melanophores in *D. nigrofasciatus* is a derived developmental mechanism (PARICHY 2006).

Development of melanophores is critical for pigmentation patterning. *mitf (nacre)* mutants lack pigment and were found to produce no melanophores throughout any developmental stage (LISTER *et al.* 1999). Among the African Great Lake cichlid species, *mitf* was identified to have undergone accelerated evolution that may have coincided with the pigment pattern diversification (SUGIE *et al.* 2004). The ocular and cutaneous albinism-2 (*Oca2*) gene may potentially also play some kind of role in melanophore formation. *Oca2* locus is the most commonly mutated gene in human albinism (BENNETT and LAMOREUX 2003) and has been shown to cause complete pigmentation loss in three different populations of the Mexican tetra cavefish, *Astyanax (A. mexicanus* and *A. fasciatus*) (PROTAS *et al.* 2006). Additionally, melanophores have been associated with divergent pigmentation patterning between the marine and freshwater three-spine sticklebacks. Marine sticklebacks are silvery with a high number iridophores, whereas the freshwater sticklebacks have prominent vertical bars across their flank with substantially less iridophores (Figure 1.5) (GREENWOOD *et al.* 2011). QTL mapping identified two loci associated with two separate features of the barred pattern: melanophore number and the degree of melanization within the melanophore (GREENWOOD *et al.* 2011). The number and distribution of melanophores are clearly important across multiple teleost species and can be used to create different pigment patterns.

Figure 1.5: Pigmentation variation observed in three-spine stickleback

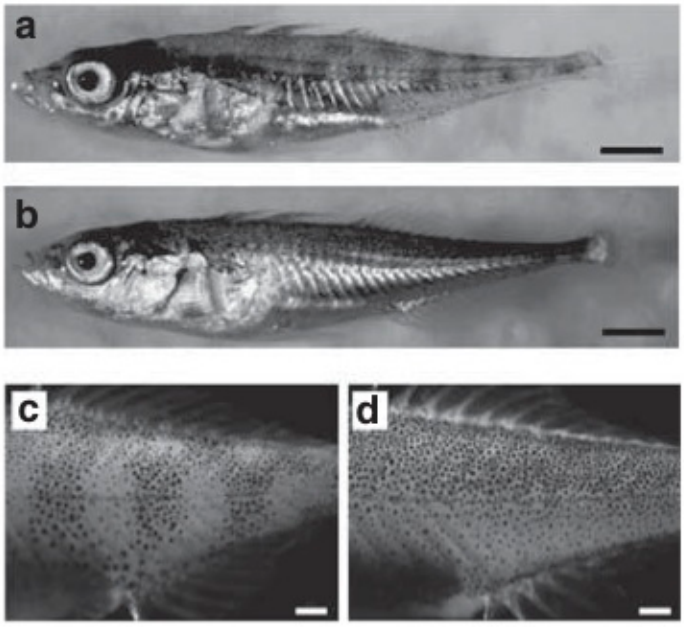


Figure 1.5: Pigment patterns of juvenile freshwater and marine sticklebacks. Photographs of 6-week-old sticklebacks from (a) a freshwater population from Hotel Lake (HL) and (b) a marine population from Little Campbell River (LCM). Scale bar 1⁄4 2 mm. (c, d) Photomicrographs of melanophore distribution across the flank of fish from (panel c) HL and (panel d) LCM after immersion in formalin to reduce iridophore pigmentation. Scale bar 1⁄4 0.5 mm. *Figure adapted from (GREENWOOD *et al.* 2011).

Although less is known about the details of the developmental and cellular mechanisms that shape pigmentation and patterns in fish compared to other organisms, much is being still being discovered. Research using natural populations of fish such as the Mexican tetra cavefish, African cichlids and the three-spine stickleback are showing that pigmentation patterns have arisen using a similar genetic toolkit in response to diverse environments and other selective pressures. Mexican cavefish illustrate the importance of genes also known to

function in humans (PROTAS *et al.* 2006). Further studies identifying molecular mechanisms combined with surveys of natural populations will likely uncover valuable insights into evolutionary and genetic processes.

Pigmentation in plants

Plant pigmentation has long been recognized as being an important tool for genetics as early as Mendel's pea plant experiments in the late nineteenth century. It has not been until recently, in the late twentieth century though that research on plant pigmentation has really gained momentum with the recent technological advances and the ability to investigate the genetics of pigment patterns in a variety of plant genera such as *Arabidopsis, Petunia*, *Antirrhinum* (snapdragons), *Ipomoea* (morning glory), *Aquilegia* (columbine), and *Mimulus* (monkeyflower) (KOPP 2009).

Flower coloration varies dramatically both within and between plant species. This diversity can have significant adaptive impacts. Flower color is most typically used to attract pollinators, so changes in pigmentation can lead to pollinator shifts (BRADSHAW and SCHEMSKE 2003; WHIBLEY *et al.* 2006; HOBALLAH *et al.* 2007; RAUSHER 2008), although other (unknown) factors can also influence flower pigmentation (COOLEY *et al.* 2008).

Pigmentation biosynthesis in plants

Pigment in flowers results from the accumulation of flavenoids (which include anthocyanins) and carotenoids (MOL *et al.* 1998). The anthocyanins produce red, purple and blue colors, where as carotenoids produce yellow or orange colors (COOLEY and WILLIS 2009). Although in addition to the accumulation of anthocyanins and carotenoids, flower pigments can also be impacted by colorless co-pigments, vacuolar pH, and cell shape (MOL *et al.* 1998).

Anthocynanins are formed in the epidermal cells of flowers (MARTIN and GERATS 1993) and accumulate in the cell's vacuole (KOES *et al.* 2005). The anthocynanin synthesis pathway has been well characterized (Figure 1.6) (KOES *et al.* 2005; KOPP 2009). This pathway notably contains serval branches for non-anthocyanin metabolites that are important for several physiological functions including pollen fertility, heat stress tolerance, UV resistance, pathogen and herbivore defense (MOL *et al.* 1998).

Much of this pathway is regulated by a transcriptional complex made up of bHLH and MYB-domain transcription factors and a WD40-repeat scaffolding protein (KOPP 2009). These enzymes and the regulatory genes that control their expression are generally conserved among flowering plants (WINKEL-SHIRLEY 2001; KOES *et al.* 2005; KOPP 2009). However, the molecular mechanism(s) employed by these enzymes to regulate most of this pathway is still unknown (KOES *et al.* 2005).

Figure 1.6: Anthocynanin synthesis pathway

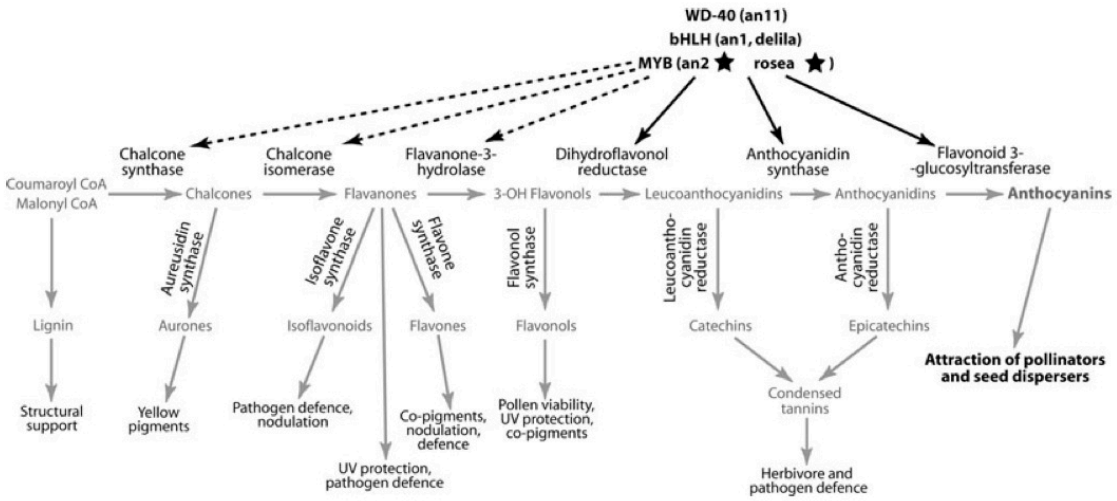


Figure 1.6: A generalized scheme of the anthocyanin synthesis pathway. Genes are shown in black and metabolites in gray. Transcriptional interactions (direct or indirect) are indicated by black arrows, and chemical reactions by gray arrows. The three upstream enzymes are co-regulated by the MYB/bHLH/WD-40 complex only in some plant lineages (dashed arrows).*Figure adapted from (KOPP 2009).

Among the studied plants, anthocyanin loss is a frequent occurrence between species, which leads to white or yellow flowers (RAUSHER 2008; KOPP 2009). For instance, the difference between the reddish-purple flower of *Petunia integrifolia* and the white flower of *Petunia axillaris* has been associated with a coding sequence mutation in the *an2* gene, which encodes a MYB-domain transcription factor (HOBALLAH *et al.* 2007). Additionally, *Antirrhinum* species that display a reduction or loss of anthocyanin pigment contain changes in the *rosea* and *venosa* genes (MYB-domain genes), and not bHLH genes (SCHWINN *et al.* 2006; WHIBLEY *et al.* 2006).

The emphasis on coloration changes stemming from changes in MYB genes has not gone unnoticed. It has been noted that MYB proteins have less pleiotropic effects than bHLH or WD40 regulators (MOL *et al.* 1998; KOES *et al.* 2005; KOPP 2009). MYBs are expressed throughout many tissues, meaning there are specialized MYB proteins for various tissues and physiological functions that can work with common bHLH and WD40 components (KOES *et al.* 2005). This is exemplified in the fact that in loss-of-function mutations in bHLH or WD40 genes (*an1* and *an11)* result in completely white flowers whereas mutations in two MYB genes (*an2* and *an11*) have more spatially confined impacts that are tissue-specific (DE VETTEN *et al.* 1997; SPELT *et al.* 2002; KOES *et al.* 2005; KOPP 2009).

Although plant pigmentation presents a complex pathway with multiple possibilities for interactions with other physiological functions, research is pressing forward to unravel and investigate the underlying molecular and genetic mechanisms, as well as, the evolutionary processes that have led to the great coloration diversity displayed among plants.

Pigmentation in flies (*Drosophila*)

Fruit flies or *Drosophila* display a wide variety of pigmentation patterns for both body and wing coloration (Figures 1.7, 1.8). This great diversity provides excellent examples of phenotypic evolution, as well as, an amenable system for research. *Drosophila* has been studied since T.H. Morgan first started his genetics research in the early twentieth century. Since those initial experiments, many experimental techniques have been developed for researchers to

investigate both how and why the multitude of distinct pigmentation patterns have arisen.

Since *Drosophila* has been used for experimental analyses for over a century, it is one of the most widely studied systems. This has led to a multitude of tools and techniques among researchers in the fly community such as genome resources for several species, procedures for genetic manipulations such as targeted gene expression, and developmental techniques like fluorescent *in situ* hybridization and live imaging of cellular process among numerous others. These many tools make *Drosophila* a tractable system in addition to its short generation time, large progeny sizes and ease to work with in the laboratory.

Pigmentation has been well-characterized in *Drosophila.* Due to their obvious phenotypes, some of the first mutants identified in *D. melanogaster* were pigmentation mutants leading the understanding of the pigmentation synthesis pathway (Figure 1.9) in *Drosophila.* Also, pigmentation varies greatly both within and between *Drosophila* species. Many of the pigmentation differences have evolved repeatedly in different lineages (WITTKOPP et al. 2003a), providing an excellent model to study the genetic mechanisms of phenotypic evolution.

There have also been several environmental conditions associated with pigmentation changes in *Drosophila.* As pigmentation development is linked with cuticle formation, desiccation resistance has been associated with various levels of pigmentation. In *D. polymorpha,* it was found that darker flies survived better under desiccating conditions as compared to lighter individuals (BRISSON *et al.* 2005). Pigmentation intensity has also been correlated with thermoregulation in

that some populations of darker flies are found in cooler areas (GIBERT *et al.* 1999), but other species display darker phenotypes closer to the equator (HOLLOCHER *et al.* 2000).

Pigmentation patterns in *Drosophila* may also be shaped by other processes. Several species exhibit sexual dimorphic patterning, suggesting a role in mate choice. A few *Drosophila* species have developed spots on their wings that correlate with the wing display in courtship behavior (KOPP and TRUE 2002b; PRUD'HOMME *et al.* 2006). Other dimorphic pigmentation patterns among species have little evidence to support a role in mating success (KRONFORST *et al.* 2012). Additionally, given that many pigmentation genes have pleiotrophic effects, pigmentation changes may also be a consequence of selective pressures on other physiological traits linked with pigmentation.

It is obvious from the research that there no general trends for the ecological and evolutionary processes that shape *Drosophila* pigmentation, but rather each species is subject to its unique geography and selective pressures.

Drosophila pigmentation patterning

Typically, *Drosophila* have a mixture of both light and dark pigments that are arrayed in some form of spatial pattern. A frequently found color design includes a dark black or brown on a lighter tan or yellow background. Among the multitude of *Drosophila* patterns, there are shared molecular elements that contribute to this diversity.

The most typical body pattern found among *Drosophila* is a band of dark pigment at the posterior edge of each dorsal cuticular plate (also called a "tergite"). These posterior stripes can also vary along the dorso-ventral axis, with some species having a wider pigmented area at the dorsal midline, while others can have wider pigmentation at the lateral edges of the tergites. Still others have much more complex patterns where posterior stripes are broken up into spots or are widened at multiple places. Some species also display sex-specific pigmentation (Compare Figure 1.7A and 1.7B). Yet other species, have solid pigmentation throughout the entire tergite or have very little dark pigment present at all (Figure 1.7F and 1.7I).

Wings also exhibit a spectacular array of pigmentation patterns in *Drosophila.* Some can have no noticeable pigmentation patterning, where others can have a single spot or numerous spots throughout the wing in either simple or highly intricate patterns (Figure 1.7E, 1.7G-H, 1.8) . Although much more is known about body pigmentation development in flies, research is uncovering more about the mechanisms and genes involved in *Drosophila* wing patterning. For example, Prud'homme and colleagues have unraveled that male-specific wing spot patterning has been gained and lost multiple times within a particular *Drosophila* clade (PRUD'HOMME *et al.* 2006), while others have identified a *cis*-regulatory sequence at the *yellow* locus that contributes to the wing spot in *D. biarmipes,* as well as as one of the transcription factors (Engrailed) that acts as a repressor (GOMPEL *et al.* 2005).

Figure 1.7: *Drosophila* pigmentation diversity

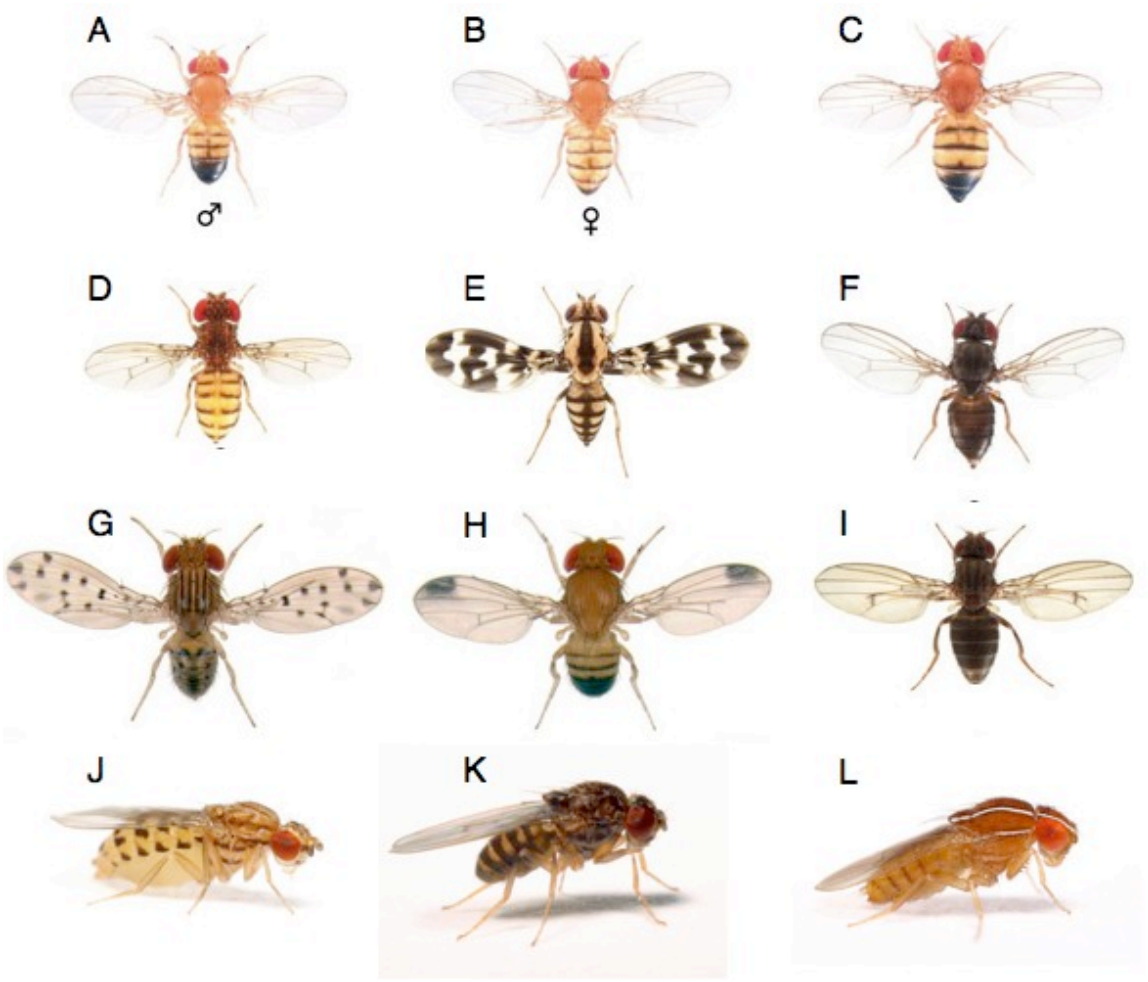


Figure 1.7: Pigmentation patterns of *Drosophila*. (A) male *D. melanogaster* (B) female *D. melanogaster* (C) *D. yakuba* (D) *D. mojavensis* (E) *D. grimshawi* (F) *D. pseudobscura* (G) *D. guttifera* (H) *D. biarmipes* (I) *D. virilis* (J) *D. busckii* (K) *D. hydei* (L) *zaprionus bagyi* *Photos compiled from Prud'home & Gompel lab website (http://www.ibdml.univ-mrs.fr/equipes/BP_NG/Illustrations/Images%201.html) and Obbard lab website (http://www.biology.ed.ac.uk/research/groups/obbard/photos.html).

Figure 1.8: *Drosophila* wing diversity

Figure 1.8: Various wing patterns from several *Drosophila* species. *Photo from Prud'homme & Gompel (http://www.ibdml.univ-mrs.fr/equipes/BP_NG/Illustrations/Images%201.html).

Transcriptional regulators can also impact abdominal pigment patterning. The spatial patterning of pigmentation has been primarily studied in the abdomen of *D. melanogaster.* The posterior stripe of dark pigment of each abdominal tergite is regulated by *optomotor-blind* (*omb*), which is controlled by the Hedgehog (Hh) signaling pathway (KOPP *et al.* 1997; KOPP and DUNCAN 2002; WITTKOPP *et al.* 2003a). Pigmentation stripes can be wider at the dorsal midline

and then gradually taper off at the lateral edges of the tergite. This tapering has been found to be regulated by the Decapentaplegic (Dpp) signaling pathway, which works together with Wingless (Wg) and epidermal growth factor receptor (EGF) signaling to control the dorso-ventral patterning of the abdominal segments (KOPP *et al.* 1999). Additionally, for the sex-specific pigmentation pattern in *D. melanogaster* males, where the last two tergites are completely pigmented, the Hox protein, Abdominal-B (Abd-B), has been identified to directly regulate the expression of a pigmentation gene (*yellow)* (JEONG *et al.* 2006). Abd-B also represses *bric a brac* (*bab*) expression in these pigmented tergites in males to produce the dimorphic pattern. In *D. melanogaster* females, Abd-B, *bab* and *doublesex* (*dsx*) all interact to inhibit male-specific pigmentation (KOPP *et al.* 2000).

Drosophila pigmentation biosynthesis

Pigmentation synthesis in *Drosophila* has been very well-characterized from studies in *D. melanogaster*. Some of the first mutants identified in *Drosophila* were pigmentation mutants due to their obvious phenotypes (KRONFORST *et al.* 2012). Consequently, many pigmentation genes were determined very early on and their enzymatic functions were established through genetic and biochemical studies (WRIGHT 1987).

In *Drosophila,* there are no specialized cell types or structures for pigments. Cuticular pigments are produced and secreted by epidermal cells and then polymerized in the overlying cuticle. These pigment precursors can only

move a few cell diameters; therefore, pigmentation is determined by the specific combination of pigment precursors in any given area. In this way, pigmentation is more or less cell autonomous (WITTKOPP *et al.* 2003a; KRONFORST *et al.* 2012).

Cuticular pigments include catecholamine polymers, which include dark melanins (black or brown), and light sclerotins (yellow, tan or colorless). All these pigments in *Drosophila* are synthesized from the derivatives dopa (dihydroxyphenylalanine) and dopamine (dihydroxyphenylethylamine) in a branched biosynthesis pathway stemming from the amino acid, tyrosine (Figure 1.9). Tyrosine is converted into the dopa by Tyrosine hydroxylase (encoded by the *pale* gene). Dopa melanization requires the *yellow* gene, which encodes an extracellular protein of unknown function (WRIGHT 1987; WITTKOPP *et al.* 2002a; TRUE 2003; WITTKOPP *et al.* 2003a). Dopa decarboxylase (DDC) converts dopa into dopamine, which then serves as a precursor to brown melanin and two types of sclerotin. Yellowish-tan pigmentation is formed when dopamine is used to form N-β-alanyl-dopamine (NBAD) through the actions of NBAD synthetase (encoded by *ebony)*. NBAD can then be used to produce NBAD sclerotin. (WRIGHT 1987) Alternatively, NBAD can be converted back into dopamine through a NBAD hydrolase encoded by the *tan* gene, and can lead to the formation of dopamine melanin that is brown (TRUE *et al.* 2005). N-acetyl-dopamine (NADA) can also be derived from dopamine, which is the primary constituent of unpigmented NADA sclerotin (WRIGHT 1987; WITTKOPP *et al.* 2003a). There are several other enzymes (encoded by additional genes) that play a role in pigment biosynthesis

in *Drosophila*, such as *black* and *silver.* (See (WRIGHT 1987) for a more detailed description of the process of melanization and sclerotization in *Drosophila*.)

Figure 1.9: *Drosophila* pigmentation biosynthesis pathway

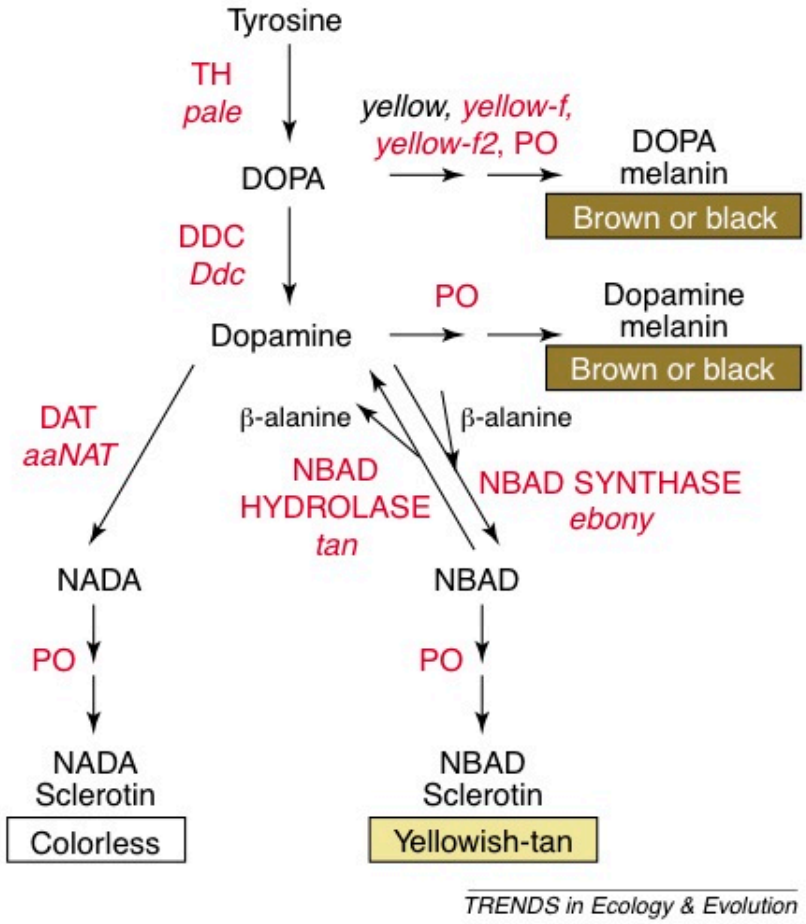


Figure 1.9: Simplified pigmentation biosynthesis pathway in *Drosophila*. Enzymes are indicated in red type. Genes in *D. melanogaster* known to encode these enzymes or corresponding to known pathway steps are in italics. Pathways shown with two arrows comprise several enzymatic steps and are still not characterized fully. Final pigmentation states are given in boxes indicating their color at the end of each pathway. Abbreviations: DAT, dopamine acetyltransferase; DDC, DOPA decarboxylase; DOPA, dihydroxyphenylalanine; NADA, N-acetyl dopamine; NBAD, N-b-alanyl dopamine; PO, phenol oxidases; TH, tyrosine hydroxylase. *Figure adapted from (TRUE 2003).

The genes, *yellow, ebony,* and *tan* have been the focus for many pigmentation studies in *Drosophila.* The enzymes encoded by these genes function at major points in the biosynthetic pathway (Figure 1.9), making them critical components for pigmentation.

yellow has been widely studied among *Drosophila* species. The expression of *yellow* leads to the formation of black melanin. Several analyses have found *yellow* expression to correspond with the pigmentation pattern of black melanin in either the body or wing (WITTKOPP *et al.* 2002a; KOPP and TRUE 2002a; WITTKOPP *et al.* 2002b; GOMPEL *et al.* 2005; PRUD'HOMME *et al.* 2006; WERNER *et al.* 2010; KALAY and WITTKOPP 2010). *yellow* has also been implicated in divergent pigmentations between species (JEONG *et al.* 2008) and is involved with the sex-specific dimorphic pigmentation pattern in *D. melanogaster* (JEONG *et al.* 2006).

ebony has also been highly investigated due to its role in pigmentation. Given that *ebony* functions in forming lighter yellow sclerotin, it has been found to act reciprocally of *yellow* (WITTKOPP *et al.* 2002a; GOMPEL *et al.* 2005). *ebony* has also been observed to be associated with pigmentation differences both within and between species. Among Ugandan populations of *D. melanogaster, ebony* was not only determined as contributing to the pigmentation variation (POOL and AQUADRO 2007), but several causative mutations in *ebony*'s 5' upstream region (presumably a region containing regulatory sequences/enhancers) were identified (REBEIZ *et al.* 2009a). Additionally, *ebony* has also been found to play a

role in the pigmentation divergence between two sister species, *D. americana* and *D. novamexicana* (WITTKOPP *et al.* 2009).

tan has been more recently investigated (TRUE *et al.* 2005) compared to *yellow* and *ebony,* but has been found to be equally important for pigmentation in many species. Both males and females of *D. yakuba* exhibit dark abdominal pigmentation in the posterior tergites, whereas *D. santomea* flies lack all abdominal pigmentation. *tan* was identified in conjunction with *yellow* to contribute to this pigmentation divergence (JEONG *et al.* 2008). Inactivation of a *cis*-regulatory element for *tan* was identified as the cause for decreased *tan* expression in the more lightly pigmented *D. santomea* (JEONG *et al.* 2008; REBEIZ *et al.* 2009b). *tan* has also been found to contribute to the pigmentation difference between *D. americana* and *D. novamexicana* (WITTKOPP *et al.* 2009).

These genes, *yellow, ebony* and *tan,* do not function solely for pigmentation purposes. All three have other functions in *Drosophila,* meaning they have pleiotropic affects. *yellow* functions in courtship behavior (DRAPEAU *et al.* 2003; 2006) and is also associated with the immune response. When injured, *Drosophila* will melanize the injured area, creating a barrier to help prevent infection by forming a scab over the exposed region (LAVINE and STRAND 2002; GALKO and KRASNOW 2004). Given that *yellow* is critical for dark melanin production, its role here would not be surprising. *ebony* and *tan* both function in histidine metabolism in the eye (BORYCZ *et al.* 2002; TRUE *et al.* 2005; WAGNER *et al.* 2007). Additionally, *ebony* also plays a role in courtship behavior in *Drosophila* (JACOBS 1978; HOVEMANN *et al.* 1998; KOHN and WITTKOPP 2007).

To gain further insight into how and why pigmentation changes (phenotypic evolution) have occurred in natural populations, further detailed analyses are needed that can provide a complete picture of the underlying genetic and molecular mechanisms by identifying and confirming functionality of causative mutations. Few studies have achieved such a comprehensive examination of phenotypic evolution (COLOSIMO *et al.* 2005; HOEKSTRA *et al.* 2006; TISHKOFF *et al.* 2007; MCGREGOR *et al.* 2007; JEONG *et al.* 2008). The experiments described in the subsequent chapters add to this collection by identifying how and why natural pigmentation changes have arisen between two *Drosophila* species, as well as investigating the pigmentation variation within one of these species.

Drosophila americana *and* Drosophila novamexicana *as a model for pigmentation divergence*

To study how pigmentation changes can occur in *Drosophila*, two closely related sister species from the *virilis* group, *D. americana* and *D. novamexicana,* were used as a model throughout the work presented here. These are interfertile species that shared a common ancestor with *D. virilis* approximately 4 million years ago (Figure 1.10A) (WITTKOPP *et al.* 2003b) and diverged from each other about 300,000-500,000 years ago (CALETKA and MCALLISTER 2004; MORALES-HOJAS *et al.* 2008). Although these two species only recently diverged, they exhibit very distinct pigmentation. *D. americana* is highly melanized (dark coloration) (Figure 1.10B) and *D. novamexicana* has relatively little melanization

and has a quite light pigmentation (Figure 1.10B). Except for *D. novamexicana,* all other members of the *virilis* group display the melanized dark phenotype.

Figure 1.10: *D. novamexicana* displays a novel pigmentation

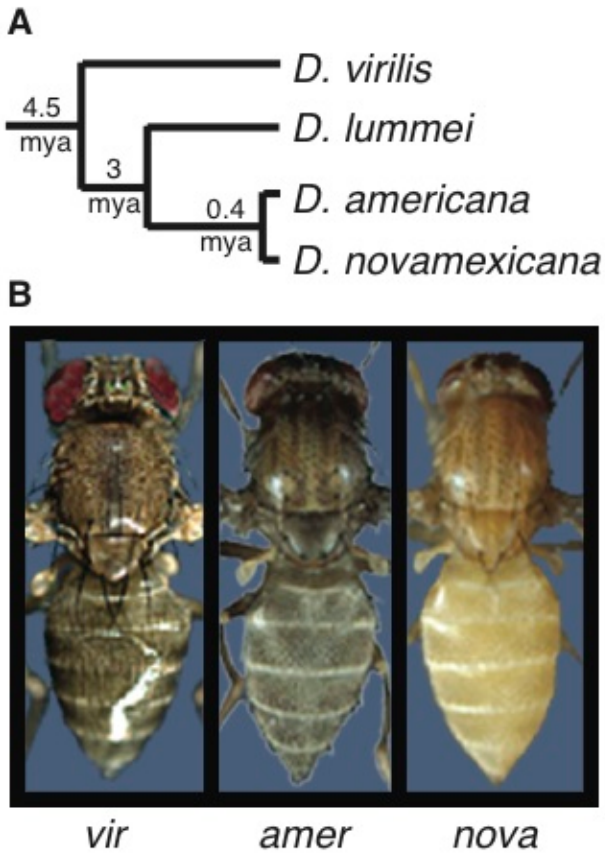


Figure 1.10: (A) Phylogenetic relationships among members of the virilis phylad within the virilis group of Drosophila are shown with estimated divergence times (CALETKA and MCALLISTER 2004) at each node (mya=millions of years ago). (B) Dorsal body pigmentation is shown for *D. virilis (vir), D. americana (amer),* and *D. novamexicana (nova). D. lummei* (not shown) has pigmentation similar to that of *D. virilis* and *D. americana* (THROCKMORTON 1982). Figure adapted from (WITTKOPP *et al.* 2009).

Although *D. americana* and *D. novamexicana* can mate successfully under lab conditions, in nature the two are distributed into allopatric populations, meaning that the natural populations of these species are geographically isolated from each other. *D. novamexicana* is found in the southwestern United States (Arizona, New Mexico, southern Utah, etc), whereas *D. americana* is found east of the Rocky Mountains across the United States ranging from Montana to eastern states such as Vermont, South Carolina, and Florida. The Rocky Mountain Range creates a physical barrier between the two populations and there are no known natural hybrid zones.

D. americana also presents a unique model in that there is pigmentation variation within this species that corresponds with location. Flies that are found in more westerly locations (such as Montana and Nebraska) exhibit relatively lighter phenotypes (although these flies are still much darker than any *D. novamexicana* individuals) and the more eastern locations display darker coloration.

The pigmentation difference between these species has been investigated previously (WITTKOPP *et al.* 2003b). *ebony* was found to contribute to this pigmentation divergence and the lighter colored *D. novamexicana* was shown to have higher Ebony protein levels as compared to the darker *D. americana* (WITTKOPP *et al.* 2003b). This intuitively makes sense with what is known about *ebony*'s role in pigmentation biosynthesis, in that it promotes the formation of light-colored (yellow) sclerotin. Since *D. novamexicana* is very light and yellowish in color, it makes sense for there to be high levels of Ebony expression.

Additionally, *yellow* was found to be not associated with the pigmentation difference between *D. americana* and *D. novamexicana* (WITTKOPP *et al.* 2003b).

These species, *D. americana* and *D. novamexicana,* offer a model system of relatively recent divergence, with a striking phenotypic difference that we can use to begin to examine the genetic and molecular basis of pigmentation divergence between species, as well as the genetic and molecular basis of pigmentation differences within a species. And since these are *Drosophila* species, many of the techniques and tools developed for other *Drosophila* species are applicable to *D. americana* and *D. novamexicana*. These species also present a system to investigate the ultimate cause of this phenotypic change; addressing why this novel pigmentation in *D. novamexicana* may have occurred.

Thesis Overview

In the following chapters, the genetic basis of pigmentation differences between *D. americana* and *D. novamexicana,* as well as the pigmentation variation within *D. americana* is investigated. In Chapter 2, *ebony* and *tan* alleles are shown to contribute to both the interspecific and intraspecific pigmentation differences. Additionally, *tan* is shown to be differentially expressed in the darker *D. americana* and a region within *tan* is identified to be important for pigmentation. In Chapter 3, this previously determined region of *tan* is further investigated using transgenic analysis to isolate causative sequence(s). The 3' half of *tan* intron 1 is discovered to likely harbor some of the pertinent nucleotides

associated with the pigmentation difference between *D. americana* and *D. novamexicana.* In Chapter 4, the pigmentation variation in *D. americana* is measured, analyzed and found to be a consequence of local adaptation among *D. americana* populations. Desiccation resistance is also explored as a potential selective pressure for pigmentation in *D. americana*. In Chapter 5, the pigmentation differences among *D. americana* are further studied by testing the function of alleles of *ebony, tan* and *yellow* among multiple populations to assess their association with phenotypic variation within this species.

My dissertation is unique in that I have conducted research in both Molecular, Cellular, and Developmental Biology, and Post-Secondary Science Education. This latter work is described in Chapter 6, which focuses on testing the usefulness of a specific pedagogical technique, online quizzes, in a large genetics class. Although quizzes were found to be statistically insignificant overall, students and instructor both expressed positive sentiments towards the quizzes.

References:

ANDERSON T. M., VONHOLDT B. M., CANDILLE S. I., MUSIANI M., GRECO C., STAHLER D. R., SMITH D. W., PADHUKASAHASRAM B., RANDI E., LEONARD J. A., BUSTAMANTE C. D., OSTRANDER E. A., TANG H., WAYNE R. K., BARSH G. S., 2009 Molecular and evolutionary history of melanism in North American gray wolves. Science **323**: 1339–1343.

ARENDT J., REZNICK D., 2008 Convergence and parallelism reconsidered: what have we learned about the genetics of adaptation? Trends Ecol Evol **23**: 26–32.

BAXTER S. W., PAPA R., CHAMBERLAIN N., HUMPHRAY S. J., JORON M., MORRISON C., FFRENCH-CONSTANT R. H., MCMILLAN W. O., JIGGINS C. D., 2008 Convergent evolution in the genetic basis of Müllerian mimicry in heliconius butterflies. Genetics **180**: 1567–1577.

BENNETT D. C., LAMOREUX M. L., 2003 The color loci of mice--a genetic century. Pigment Cell Res **16**: 333–344.

BORYCZ J., BORYCZ J. A., LOUBANI M., MEINERTZHAGEN I. A., 2002 tan and ebony genes regulate a novel pathway for transmitter metabolism at fly photoreceptor terminals. J Neurosci **22**: 10549–10557.

BRADSHAW H. D., SCHEMSKE D. W., 2003 Allele substitution at a flower colour locus produces a pollinator shift in monkeyflowers. Nature **426**: 176–178.

BRISSON J., DETONI D., DUNCAN I., ARTEMPLETON, 2005 Abdominal Pigmentation Variation in Drosophila Polymorpha: Geographic Variation in the Trait and Underlying Phylogeography. Evolution **59**: 1046–1059.

CALETKA B. C., MCALLISTER B. F., 2004 A genealogical view of chromosomal evolution and species delimitation in the Drosophila virilis species subgroup. Mol Phylogenet Evol **33**: 664–670.

CARO T., 2005 The Adaptive Significance of Coloration in Mammals. BioScience **55**: 125–136.

CARROLL S. B., 2005 Evolution at two levels: on genes and form. PLoS Biol **3**: e245.

COLOSIMO P. F., HOSEMANN K. E., BALABHADRA S., VILLARREAL G., DICKSON M., GRIMWOOD J., SCHMUTZ J., MYERS R. M., SCHLUTER D., KINGSLEY D. M., 2005 Widespread parallel evolution in sticklebacks by repeated fixation of Ectodysplasin alleles. Science **307**: 1928–1933.

COOK L. M., GRANT B. S., SACCHERI I. J., MALLET J., 2012 Selective bird predation on the peppered moth: the last experiment of Michael Majerus. Biol Lett. [Epub ahead of print]

COOLEY A. M., WILLIS J. H., 2009 Genetic divergence causes parallel evolution of flower color in Chilean Mimulus. New Phytol. **183**: 729–739.

COOLEY A. M., CARVALLO G., WILLIS J. H., 2008 Is floral diversification associated with pollinator divergence? Flower shape, flower colour and pollinator preference in Chilean Mimulus. Ann. Bot. **101**: 641–650.

DE VETTEN N., QUATTROCCHIO F., MOL J., KOES R., 1997 The an11 locus controlling flower pigmentation in petunia encodes a novel WD-repeat protein conserved in yeast, plants, and animals. Genes Dev **11**: 1422–1434.

DRAPEAU M. D., CYRAN S. A., VIERING M. M., GEYER P. K., LONG A. D., 2006 A cis-regulatory sequence within the yellow locus of Drosophila melanogaster required for normal male mating success. Genetics **172**: 1009–1030.

DRAPEAU M. D., RADOVIC A., WITTKOPP P. J., LONG A. D., 2003 A gene necessary for normal male courtship, yellow, acts downstream of fruitless in the Drosophila melanogaster larval brain. J Neurobiol **55**: 53–72.

ENDLER J. A., 1983 Natural and sexual selection on color patterns in poeciliid fishes. Environ Biol Fish **9**: 173–190.

GALKO M. J., KRASNOW M. A., 2004 Cellular and genetic analysis of wound healing in Drosophila larvae. PLoS Biol **2**: E239.

GARCÍA A., 2011 Exploring Female Preference for Male Melanic Pigmentation Patterns in the Malawian Cichlid Metriaclima Zebra. In: Bowling Green State University, pp. 1–41.

GIBERT P., MORETEAU B., MUNJAL A., DAVID J. R., 1999 Phenotypic plasticity of abdominal pigmentation in Drosophila kikkawai: multiple interactions between a major gene, sex, abdomen segment and growth temperature. Genetica **105**: 165–176.

GOMPEL N., PRUD'HOMME B., WITTKOPP P. J., KASSNER V. A., CARROLL S. B., 2005 Chance caught on the wing: cis-regulatory evolution and the origin of pigment patterns in Drosophila. Nature **433**: 481–487.

GRANT B. S., 2004 Allelic melanism in American and British peppered moths. J. Hered. **95**: 97–102.

GRANT B. S., OWEN D. F., CLARKE C. A., 1996 Parallel Rise and Fall of Melanic Peppered Moths in America and Britain. J. Hered. **87**: 351–357.

GREENWOOD A. K., JONES F. C., CHAN Y. F., BRADY S. D., ABSHER D. M., GRIMWOOD J., SCHMUTZ J., MYERS R. M., KINGSLEY D. M., PEICHEL C. L., 2011 The genetic basis of divergent pigment patterns in juvenile threespine sticklebacks. Heredity **107**: 155–166.

HOBALLAH M. E., GÜBITZ T., STUURMAN J., BROGER L., BARONE M., MANDEL T., DELL'OLIVO A., ARNOLD M., KUHLEMEIER C., 2007 Single gene-mediated shift in pollinator attraction in Petunia. The Plant Cell **19**: 779–790.

HOEKSTRA H. E., NACHMAN M. W., 2003 Different genes underlie adaptive melanism in different populations of rock pocket mice. Molecular Ecology **12**: 1185–1194.

HOEKSTRA H. E., HIRSCHMANN R. J., BUNDEY R. A., INSEL P. A., CROSSLAND J. P., 2006 A single amino acid mutation contributes to adaptive beach mouse color pattern. Science **313**: 101–104.

HOFREITER M., SCHÖNEBERG T., 2010 The genetic and evolutionary basis of colour variation in vertebrates. Cell Mol Life Sci **67**: 2591–2603.

HOLLOCHER H., HATCHER J. L., DYRESON E. G., 2000 Genetic and developmental analysis of abdominal pigmentation differences across species in the Drosophila dunni subgroup. Evolution **54**: 2057–2071.

HOLTON T. A., CORNISH E. C., 1995 Genetics and Biochemistry of Anthocyanin Biosynthesis. The Plant Cell **7**: 1071.

HOVEMANN B. T., RYSECK R. P., WALLDORF U., STÖRTKUHL K. F., DIETZEL I. D., DESSEN E., 1998 The Drosophila ebony gene is closely related to microbial peptide synthetases and shows specific cuticle and nervous system expression. Gene **221**: 1–9.

JACKSON I. J., 1994 Molecular and developmental genetics of mouse coat color. Annu. Rev. Genet. **28**: 189–217.

JACKSON I. J., BUDD P., HORN J. M., JOHNSON R., RAYMOND S., STEEL K., 1994 Genetics and Molecular Biology of Mouse Pigmentation. Pigment Cell Research **7**: 73–80.

JACOBS M. E., 1978 Influence of beta-alanine on mating and territorialism in Drosophila melanogaster. Behav Genet **8**: 487–502.

JEONG S., REBEIZ M., ANDOLFATTO P., WERNER T., TRUE J., CARROLL S. B., 2008 The evolution of gene regulation underlies a morphological difference between two Drosophila sister species. Cell **132**: 783–793.

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**: 1387–1399.

JONG P., GUSSEKLOO S., BRAKEFIELD P., 1996 Differences in thermal balance, body temperature and activity between non-melanic and melanic two-spot ladybird beetles (Adalia bipunctata) under controlled conditions. Journal of Experimental Biology **199**: 2655–2666.

JORON M., PAPA R., BELTRÁN M., CHAMBERLAIN N., MAVÁREZ J., BAXTER S., ABANTO M., BERMINGHAM E., HUMPHRAY S. J., ROGERS J., BEASLEY H., BARLOW K., FFRENCH-CONSTANT R. H., MALLET J., MCMILLAN W. O., JIGGINS C. D., 2006 A conserved supergene locus controls colour pattern diversity in Heliconius butterflies. PLoS Biol **4**: e303.

KABBARAH O., CHIN L., 2006 Advances in malignant melanoma: genetic insights from mouse and man. Front. Biosci. **11**: 928–942.

KALAY G., WITTKOPP P. J., 2010 Nomadic Enhancers: Tissue-Specific cis-Regulatory Elements of yellow Have Divergent Genomic Positions among Drosophila Species. PLoS Genet **6**: e1001222.

KELSH R. N., 2004 Genetics and Evolution of Pigment Patterns in Fish. Pigment Cell Research **17**: 326–336.

KETTLEWELL H. B. D., 1961 The Phenomenon of Industrial Melanism in Lepidoptera. Annu. Rev. Entomol. **6**: 245–262.

KETTLEWELL H. B. D., 1965 Insect Survival and Selection for Pattern: Most camouflage and survival mechanisms, though highly perfected, can be adapted to changing environments. Science **148**: 1290–1296.

KOES R., VERWEIJ W., QUATTROCCHIO F., 2005 Flavonoids: a colorful model for the regulation and evolution of biochemical pathways. Trends in Plant Science **10**: 236–242.

KOHN M. H., WITTKOPP P. J., 2007 Annotating ebony on the fly. Molecular Ecology **16**: 2831–2833.

KONDO S., IWASHITA M., YAMAGUCHI M., 2009 How animals get their skin patterns: fish pigment pattern as a live Turing wave. Int. J. Dev. Biol. **53**: 851–856.

KOPP A., 2009 Metamodels and Phylogenetic Replication: A Systematic Approach to the Evolution of Developmental Pathways. Evolution **63**: 2771–2789.

KOPP A., DUNCAN I., 2002 Anteroposterior patterning in adult abdominal segments of Drosophila. Dev Biol **242**: 15–30.

KOPP A., TRUE J. R., 2002a Evolution of male sexual characters in the oriental Drosophila melanogaster species group. Evol Dev **4**: 278–291.

KOPP A., TRUE J. R., 2002b Phylogeny of the Oriental Drosophila melanogaster species group: a multilocus reconstruction. Syst Biol **51**: 786–805.

KOPP A., BLACKMAN R. K., DUNCAN I., 1999 Wingless, decapentaplegic and EGF receptor signaling pathways interact to specify dorso-ventral pattern in the adult abdomen of Drosophila. Development **126**: 3495–3507.

KOPP A., DUNCAN I., GODT D., CARROLL S. B., 2000 Genetic control and evolution of sexually dimorphic characters in Drosophila. Nature **408**: 553–559.

KOPP A., MUSKAVITCH M. A., DUNCAN I., 1997 The roles of hedgehog and engrailed in patterning adult abdominal segments of Drosophila. Development **124**: 3703–3714.

KRONFORST M. R., BARSH G. S., KOPP A., MALLET J., MONTEIRO A., MULLEN S. P., PROTAS M., ROSENBLUM E. B., SCHNEIDER C. J., HOEKSTRA H. E., 2012 Unraveling the thread of nature's tapestry: the genetics of diversity and convergence in animal pigmentation. Pigment Cell & Melanoma Research **25**: 411–433.

LAVINE M. D., STRAND M. R., 2002 Insect hemocytes and their role in immunity. Insect Biochem Mol Biol **32**: 1295–1309.

LEES D. R., 1968 Genetic control of the melanic form Insularia of the peppered moth Biston betularia (L.). Nature **220**: 1249–1250.

LISTER J. A., ROBERTSON C. P., LEPAGE T., JOHNSON S. L., RAIBLE D. W., 1999 nacre encodes a zebrafish microphthalmia-related protein that regulates neural-crest-derived pigment cell fate. Development **126**: 3757–3767.

LOMAS J., MARTIN-DUQUE P., PONS M., 2008 The genetics of malignant melanoma. [Front Biosci. 2008] - PubMed - NCBI. Front. Biosci. **13**: 5071–5093.

MAAN M., 2012 Poison Frog Colors are Honest Signals of Toxicity Particularly for Bird Predators. Am. Nat. **179**: E1–E14.

MARTIN C., GERATS T., 1993 Control of Pigment Biosynthesis Genes during Petal Development. The Plant Cell **5**: 1253.

MCGREGOR A. P., ORGOGOZO V., DELON I., ZANET J., SRINIVASAN D. G., PAYRE F., STERN D. L., 2007 Morphological evolution through multiple cis-regulatory mutations at a single gene. Nature **448**: 587–590.

MOL J., GROTEWOLD E., KOES R., 1998 How genes paint flowers and seeds. Trends in Plant Science **3**: 212–217.

MORALES-HOJAS R., VIERIRA C. P., VIEIRA J., 2008 Inferring the evolutionary history of Drosophila americana and Drosophila novamexicana using a multilocus approach and the influence of chromosomal rearrangements in single gene analyses. Molecular Ecology **17**: 2910–2926.

NACHMAN M. W., 2005 The genetic basis of adaptation: lessons from concealing coloration in pocket mice. Genetica **123**: 125–136.

NACHMAN M. W., HOEKSTRA H. E., D'AGOSTINO S. L., 2003 The genetic basis of adaptive melanism in pocket mice. PNAS **100**: 5268–5273.

NADEAU N. J., BURKE T., MUNDY N. I., 2007 Evolution of an avian pigmentation gene correlates with a measure of sexual selection. Proc. R. Soc. B **274**: 1807–1813.

OBBARD D.J., 2012 Drosophila in the lab from Obbard Lab Website. University of Edinburgh, Edinburgh, UK. http://www.biology.ed.ac.uk/research/groups/obbard/photos.html (accessed June 26, 2012)

OLIVER J. C., ROBERTSON K. A., MONTEIRO A., 2009 Accommodating natural and sexual selection in butterfly wing pattern evolution. Proceedings of the Royal Society B: Biological Sciences **276**: 2369–2375.

PAPA R., MORRISON C. M., WALTERS J. R., COUNTERMAN B. A., CHEN R., HALDER G., FERGUSON L., CHAMBERLAIN N., FFRENCH-CONSTANT R., KAPAN D. D., JIGGINS C. D., REED R. D., MCMILLAN W. O., 2008 Highly conserved gene order and numerous novel repetitive elements in genomic regions linked to wing pattern variation in Heliconius butterflies. BMC Genomics **9**: 345.

PARICHY D. M., 2003 Pigment patterns: fish in stripes and spots. Current Biology **13**: R947–50.

PARICHY D. M., 2006 Evolution of danio pigment pattern development. Heredity **97**: 200–210.

POOL J. E., AQUADRO C. F., 2007 The genetic basis of adaptive pigmentation variation in Drosophila melanogaster. Molecular Ecology **16**: 2844–2851.

PROTAS M. E., PATEL N. H., 2008 Evolution of coloration patterns. Annu Rev Cell Dev Biol **24**: 425–446.

PROTAS M. E., HERSEY C., KOCHANEK D., ZHOU Y., WILKENS H., JEFFERY W. R., ZON L. I., BOROWSKY R., TABIN C. J., 2006 Genetic analysis of cavefish reveals molecular convergence in the evolution of albinism. Nat Genet **38**: 107–111.

PRUD'HOMME B., GOMPEL N., 2012 Images from the Prud'homme & Gompel Lab Website, Institut de Biologie du Développement de Marseille-Luminy, Marseille, France. http://www.ibdml.univ-mrs.fr/equipes/BP_NG/Illustrations/Images%201.html (accessed June 26, 2012)

PRUD'HOMME B., GOMPEL N., ROKAS A., KASSNER V. A., WILLIAMS T. M., YEH S.-D., TRUE J. R., CARROLL S. B., 2006 Repeated morphological evolution through cis-regulatory changes in a pleiotropic gene. Nature **440**: 1050–1053.

QUIGLEY I. K., PARICHY D. M., 2002 Pigment pattern formation in zebrafish: a model for developmental genetics and the evolution of form. Microsc. Res. Tech. **58**: 442–455.

RAUSHER M. D., 2008 Evolutionary Transitions in Floral Color. Int. J Plant Sci. **169**: 7–21.

REBEIZ M., POOL J. E., KASSNER V. A., AQUADRO C. F., CARROLL S. B., 2009a Stepwise modification of a modular enhancer underlies adaptation in a Drosophila population. Science **326**: 1663–1667.

REBEIZ M., RAMOS-WOMACK M., JEONG S., ANDOLFATTO P., WERNER T., TRUE J., STERN D. L., CARROLL S. B., 2009b Evolution of the tan Locus Contributed to Pigment Loss in Drosophila santomea: A Response to Matute et al. Cell **139**: 1189–1196.

ROSENBLUM E. B., 2006 Convergent evolution and divergent selection: lizards at the White Sands ecotone. Am. Nat. **167**: 1–15.

ROSENBLUM E. B., RÖMPLER H., SCHÖNEBERG T., HOEKSTRA H. E., 2010 Molecular and functional basis of phenotypic convergence in white lizards at White Sands. Proceedings of the National Academy of Sciences **107**: 2113–2117.

SCHWINN K., VENAIL J., SHANG Y., MACKAY S., ALM V., BUTELLI E., OYAMA R., BAILEY P., DAVIES K., MARTIN C., 2006 A Small Family of MYB-Regulatory Genes Controls Floral Pigmentation Intensity and Patterning in the Genus Antirrhinum. The Plant Cell **18**: 831–851.

SPELT C., QUATTROCCHIO F., MOL J., KOES R., 2002 ANTHOCYANIN1 of petunia controls pigment synthesis, vacuolar pH, and seed coat development by genetically distinct mechanisms. The Plant Cell **14**: 2121–2135.

STEINER C., WEBER J., HOEKSTRA H., 2007 Adaptive Variation in Beach Mice Produced by Two Interacting Pigmentation Genes. PLoS Biol **5**: e219.

STUART-FOX D., MOUSSALLI A., 2009 Camouflage, communication and thermoregulation: lessons from colour changing organisms. Phil. Trans. R. Soc. B. **364**: 463–470.

SUGIE A., TERAI Y., OTA R., OKADA N., 2004 The evolution of genes for pigmentation in African cichlid fishes. Gene **343**: 337–346.

SULEM P., GUDBJARTSSON D. F., STACEY S. N., HELGASON A., RAFNAR T., MAGNUSSON K. P., MANOLESCU A., KARASON A., PALSSON A., THORLEIFSSON G., JAKOBSDOTTIR M., STEINBERG S., PÁLSSON S., JONASSON F., SIGURGEIRSSON B., THORISDOTTIR K., RAGNARSSON R., BENEDIKTSDOTTIR K. R., ABEN K. K., KIEMENEY L. A., OLAFSSON J. H., GULCHER J., KONG A., THORSTEINSDOTTIR U., STEFANSSON K., 2007 Genetic determinants of hair, eye and skin pigmentation in Europeans. Nat Genet **39**: 1443–1452.

THROCKMORTON L.H., 1982 in The Genetics and Biology of Drosophila. Ashburner M., Carson H.L., Thompson J.N., Eds (Academic Press, London), vol 3b, 227-296

TISHKOFF S. A., REED F. A., RANCIARO A., VOIGHT B. F., BABBITT C. C., SILVERMAN J. S., POWELL K., MORTENSEN H. M., HIRBO J. B., OSMAN M., IBRAHIM M., OMAR S. A., LEMA G., NYAMBO T. B., GHORI J., BUMPSTEAD S., PRITCHARD J. K., WRAY G. A., DELOUKAS P., 2007 Convergent adaptation of human lactase persistence in Africa and Europe. Nat Genet **39**: 31–40.

TRUE J. R., 2003 Insect melanism: the molecules matter. Trends Ecol Evol **18**: 640–647.

TRUE J. R., YEH S.-D., HOVEMANN B. T., KEMME T., MEINERTZHAGEN I. A., EDWARDS T. N., LIOU S.-R., HAN Q., LI J., 2005 Drosophila tan encodes a

novel hydrolase required in pigmentation and vision. PLoS Genet **1**: e63.

VAN'T HOF A. E., EDMONDS N., DALÍKOVÁ M., MAREC F., SACCHERI I. J., 2011 Industrial melanism in British peppered moths has a singular and recent mutational origin. Science **332**: 958–960.

WAGNER S., HESEDING C., SZLACHTA K., TRUE J., PRINZ H., HOVEMANN B., 2007 Drosophila Photoreceptors Express Cysteine Peptidase Tan. J. Comp. Neurol. **500**: 601–611.

WEBBER R., WEBBER BARLOW G., BRUSH A. H., 1973 Pigments of a color polymorphism in a cichlid fish. Comparative Biochemistry and Physiology Part B: Comparative Biochemistry **44**: 1127–1135.

WERNER T., KOSHIKAWA S., WILLIAMS T. M., CARROLL S. B., 2010 Generation of a novel wing colour pattern by the Wingless morphogen. Nature **464**: 1143–1148.

WHIBLEY A. C., LANGLADE N. B., ANDALO C., HANNA A. I., BANGHAM A., THÉBAUD C., COEN E., 2006 Evolutionary paths underlying flower color variation in Antirrhinum. Science **313**: 963–966.

WIERNASZ D., 1989 Female Choice and Sexual Selection of Male Wing Melanin Pattern in Pieris Occidentalis (Lepidoptera). Evolution **43**: 1672–1682.

WINKEL-SHIRLEY B., 2001 It takes a garden. How work on diverse plant species has contributed to an understanding of flavonoid metabolism. Plant Physiol. **127**: 1399–1404.

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**: 65–71.

WITTKOPP P. J., CARROLL S. B., KOPP A., 2003a Evolution in black and white: genetic control of pigment patterns in Drosophila. Trends Genet **19**: 495–504.

WITTKOPP P. J., STEWART E. E., ARNOLD L. L., NEIDERT A. H., HAERUM B. K., THOMPSON E. M., AKHRAS S., SMITH-WINBERRY G., SHEFNER L., 2009 Intraspecific polymorphism to interspecific divergence: genetics of pigmentation in Drosophila. Science **326**: 540–544.

WITTKOPP P. J., TRUE J. R., CARROLL S. B., 2002a Reciprocal functions of the Drosophila yellow and ebony proteins in the development and evolution of pigment patterns. Development **129**: 1849–1858.

WITTKOPP P. J., VACCARO K., CARROLL S. B., 2002b Evolution of yellow gene regulation and pigmentation in Drosophila. Current Biology **12**: 1547–1556.

WITTKOPP P. J., WILLIAMS B. L., SELEGUE J. E., CARROLL S. B., 2003b Drosophila pigmentation evolution: divergent genotypes underlying convergent phenotypes. PNAS **100**: 1808–1813.

WRIGHT T. R. F., 1987 The Genetics of Biogenic Amine Metabolism, Sclerotization, and Melanization in Drosophila melanogaster. Advances in Genetics **24**: 127–222.

Chapter 2

Genetic Basis of Pigmentation between *D. americana* and *D. novamexicana*[1]

Abstract

Phenotypes that differ both within and between species can be caused by similar or dissimilar genetic mechanisms. To address this, individual genes and divergent sites within those genes must be identified and compared within and between species. Here, we show that the same alleles of *ebony* and *tan* contribute to the pigmentation difference seen between *D. americana* and *D. novamexicana,* as well as the pigmentation variation observed within *D. americana.* Additionally, *tan* mRNA is shown to be more highly expressed in the darker *D. americana* species and divergent *tan* sequence is identified that shows an association with pigmentation. The functions of *ebony* and *tan* in pigmentation and their respective protein and mRNA expression patterns in these species lead to a model of pigmentation that illustrates how natural populations may have utilized changes in these alleles to create intraspecific polymorphism and interspecific divergence in *Drosophila*.

---

[1]This chapter is a compilation of data from several individuals published as Wittkopp, P.J., Stewart, E.E., Arnold, L.L., Neidert, A.H., Haerum, B. K., Thompson, E.M., Akhras, S., Smith-Winterberry, G. & Shefner, L. (2009) Intraspecific Polymorphism to Interspecific Divergence: Genetics of Pigmentation in *Drosophila*. *Science*. 326: 540-544. My contributions include data for the following figures: Figure 2.5, Figure 2.8, and Figure 2.9. Our combined efforts have shown that the same alleles underlying interspecific divergence, also contribute to intraspecific polymorphism (within *D. americana)* as well as identifying candidate sequences that may play a role in divergent pigmentation.

Introduction

Similar phenotypes that vary within and between species may or may not be caused by similar genetic mechanisms. Quantitative trait mapping shows that loci contributing to polymorphism and divergence of a single character map to the same region of the genome approximately half of the time (Table 2.1). These overlapping quantitative trait loci (QTLs) may or may not result from changes in the same genes, and most studies lack the power to distinguish between these possibilities. To determine whether the same genes (and potentially even the same alleles of these genes) contribute to phenotypic diversity within and between species, intra- and interspecific QTLs must be resolved to individual genes, functionally divergent sites must be localized within these genes, and specific alleles must be compared within and between species.

Results and Discussion

ebony *and* tan *QTLs contribute to pigmentation divergence*

To investigate the relationship between intraspecific polymorphism and interspecific divergence, we examined the genetic basis of pigmentation differences within and between a pair of closely related *Drosophila* species. *D. americana* and *D. novamexicana* are sister taxa within the *Drosophila virilis* species group that diverged approximately 300,000 – 500,000 years ago (CALETKA and MCALLISTER 2004; MORALES-HOJAS *et al.* 2008, Figure 2.1A). D. novamexicana has a derived light yellow body color, while other members of this

group (including *D. americana*) retain an ancestral dark brown body color (THROCKMORTON 1982, Figure 2.1B). In the laboratory, these species can mate and produce fertile offspring. Genetic mapping showed that a region of the second chromosome containing the *ebony* gene contributes to pigmentation divergence between *D. novamexicana* and *D. americana* (WITTKOPP *et al.* 2003b). This gene is required for pigmentation in *D. melanogaster* (BRIDGES AND MORGAN 1923). Three other autosomal regions, as well as an unidentified region of the X-chromosome, also contribute to pigmentation divergence, although none of these regions were linked to other pigmentation genes tested (i.e., *yellow*, *dopa-decarboxylase*, *optomotor blind*, and *bric-a-brac*).

Figure 2.1: *D. novamexicana yellow* body color is derived

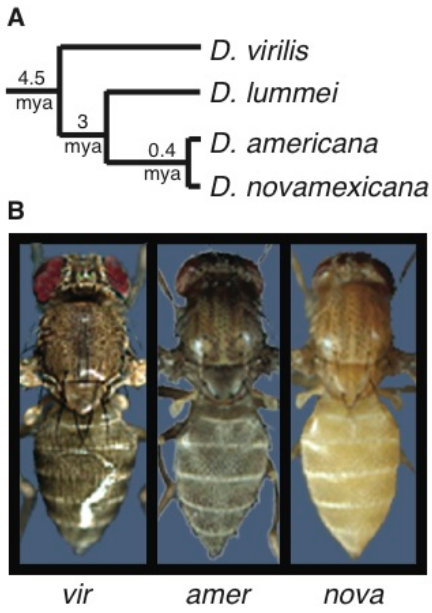


Figure 2.1: (A) Phylogenetic relationships among members of the virilis phylad within the virilis group of Drosophila are shown with estimated divergence times (CALETKA and MCALLISTER 2004) at each node (mya=millions of years ago). (B) Dorsal body pigmentation is shown for *D. virilis*

(vir), D. americana (amer), and *D. novamexicana (nova). D. lummei* (not shown) has pigmentation similar to that of *D. virilis* and *D. americana* (THROCKMORTON 1982).

Recently, the X-linked pigmentation gene *tan* was cloned in D. melanogaster (TRUE *et al.* 2005). To test whether this gene might contribute to pigmentation differences between *D. americana* and *D. novamexicana*, we crossed *D. americana* females to *D. novamexicana* males, backcrossed F_1 hybrid females to *D. novamexicana* males, and scored 495 backcross progeny for body color (Figure 2.3). All of the lightest male offspring (n = 10) inherited the *D. novamexicana* allele of *tan*, while all of the darkest male offspring (n = 24) inherited the *D. americana* allele of this marker. These data show that sequences linked to *tan* contribute to pigmentation divergence ($P_{\text{Fisher's exact test}} = 8\text{x}10^{-9}$). The previously described pigmentation QTL linked to ebony and the lack of a pigmentation QTL linked to yellow (WITTKOPP *et al.* 2003b) were also reconfirmed in this population ($P_{\text{Fisher's exact test}} = 3\text{x}10^{-8}$ and 0.7, respectively).

To determine the phenotypic effects of QTLs linked to ebony and tan, we created lines of *D. novamexicana* in which genomic regions containing these genes were replaced with orthologous sequences from *D. americana*. These genotypes were constructed by marker-assisted introgression, moving *ebony* and *tan* alleles from *D. americana* into *D. novamexicana*. F_1 hybrid females were backcrossed to *D. novamexicana* males, and a single female inheriting the *D. americana tan* (or *ebony*) allele was randomly selected and backcrossed to *D. novamexicana* males again. This process was repeated for ten generations

(Figure 2.4), with females carrying the *D. americana ebony* or *tan* allele selected randomly in each generation without regard to pigmentation. Introgressed *D. americana* sequences linked to either *tan* or *ebony* darkened pigmentation relative to wild-type *D. novamexicana* (Figure 2.2 A, B, C), with sequences linked to ebony (Figure 2.2C) causing darker pigmentation than sequences linked to *tan* (Figure 2.2B). Digital quantification of pigmentation showed that, when combined, the introgressed *tan* and *ebony* regions recapitulated 87% of the pigmentation difference between species (Figure 2.2A, D, E).

Figure 2.2: QTLs linked to *tan* and *ebony* account for the majority of pigmentation divergence between species

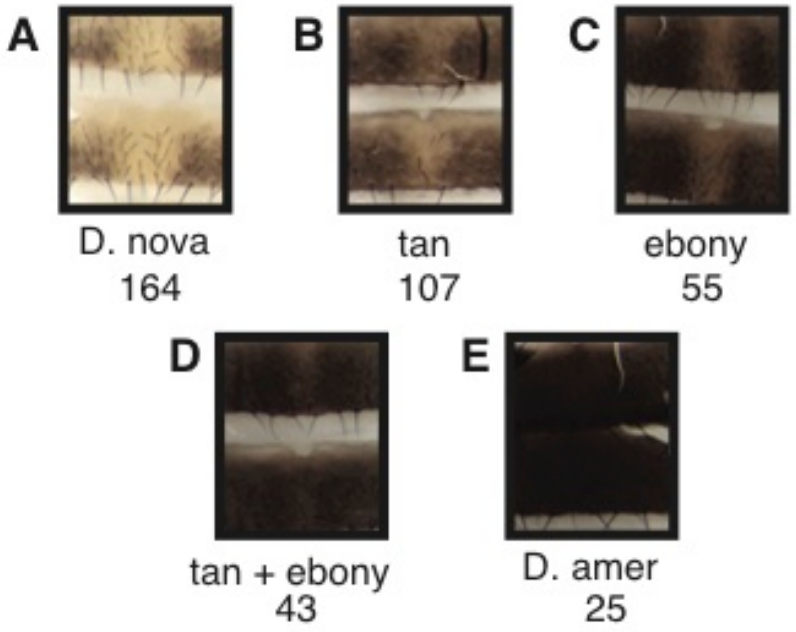


Figure 2.2: Dorsal abdominal cuticle is shown from segments A4 and A5 of 7-10 day old adult females. Compared to *D. novmexicana* (A), introgression of alleles linked to *tan* (B) or *ebony* (C)

darkened pigmentation. Together, introgressed regions produced even darker pigmentation (D), although these flies were not as dark as wild-type *D. americana* (E). Numbers indicate intensity of gray-scale images, where 0=black and 255=white. Panels B, C, and D are all heterozygous for the introgressed region(s).

Figure 2.3: Interspecific backcross progeny are distributed among five pigmentation classes

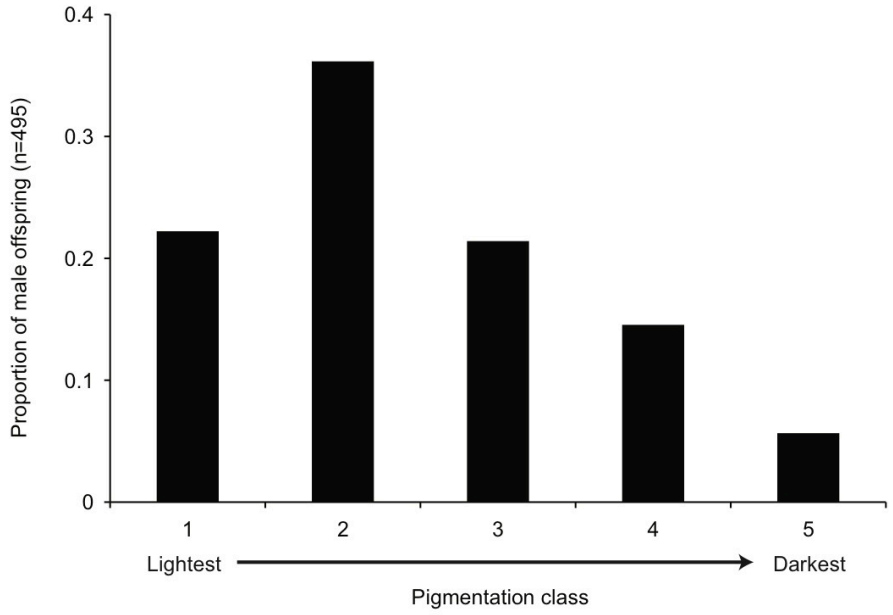


Figure 2.3: F1 hybrid females, produced by crossing *D. americana* (A00) females to *D. novamexicana* (N14) males, were backcrossed to *D. novamexicana* (N14) males. Body color was scored by eye for 495 male progeny. Pigmentation phenotypes were not continuous, but rather fell into five distinct classes. The phenotypic distribution observed is similar to that reported in (WITTKOPP *et al.* 2003b). Light and dark flies genotyped came from classes 1 and 5, respectively.

Figure 2.4: Crossing scheme used to introgress *D. americana* alleles linked to *tan* into *D. novamexicana*

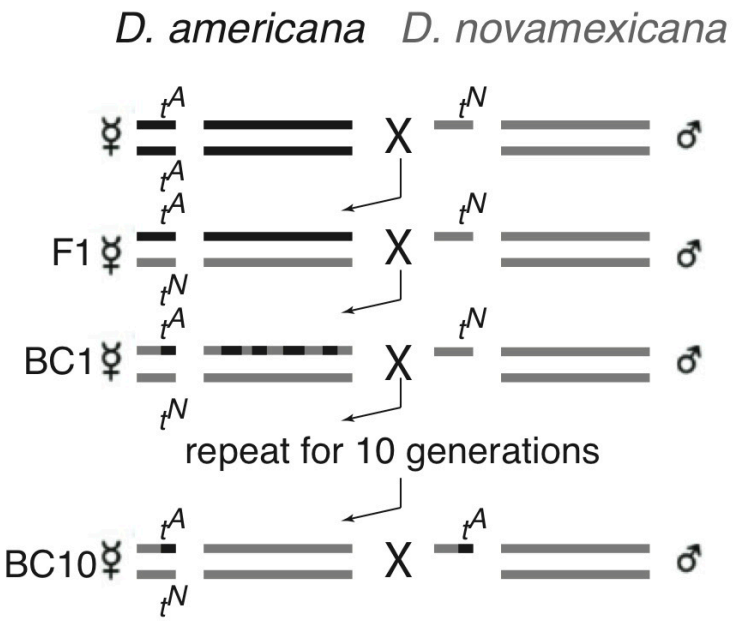


Figure 2.4: Short bars represent X chromosomes and long bars represent all five autosomes for *D. americana* (black) or *D. novamexicana* (gray). Alleles of *tan* from *D. americana* (*tA*) and *D. novamexicana* (*tN*) are indicated. An analogous crossing scheme was used to introgress autosomal alleles linked to *ebony* from *D. americana* into *D. novamexicana*.

ebony *and* tan *affect pigmentation development*

Studies of pigmentation in *D. melanogaster* suggest that *ebony* and *tan* may themselves be responsible for these interspecific QTLs. Loss-of-function mutations in *ebony* darken pigmentation (Bridges and Morgan 1923), while loss of- function mutations in *tan* lighten it (MCEWEN 1918). Biochemically, Ebony catalyzes the conversion of dopamine into N-beta-alanyl-dopamine (NBAD),

which is a precursor for (yellow) sclerotin, and Tan catalyzes the reverse reaction, converting NBAD back into dopamine, which is a precursor for (brown) melanin (reviewed in (WITTKOPP *et al.* 2003a, Figure 2.5A). Ectopic expression of Ebony induces yellow pigmentation (WITTKOPP *et al.* 2002, Figure 2.5D), while ectopic expression of Tan induces brown pigmentation (TRUE *et al.* 2005, Figure 2.5E). Ectopic expression of both proteins simultaneously results in pigmentation intermediate to that caused by ectopic expression of either protein alone (Figure 2.5F), showing that the balance between Ebony and Tan enzymatic activity affects pigmentation. Genetic and biochemical pathways controlling pigment synthesis are highly conserved among insects (WITTKOPP and BELDADE 2009), suggesting that the *D. americana* and *D. novamexicana tan* and *ebony* genes function similarly to their *D. melanogaster* orthologs. Consistent with this prediction, the Ebony protein is more abundant in epidermal cells of the yellowish *D. novamexicana* during late pupal stages than in the darker *D. americana* (WITTKOPP *et al.* 2003b).

Figure 2.5: Ebony and Tan have reciprocal effects on pigmentation development

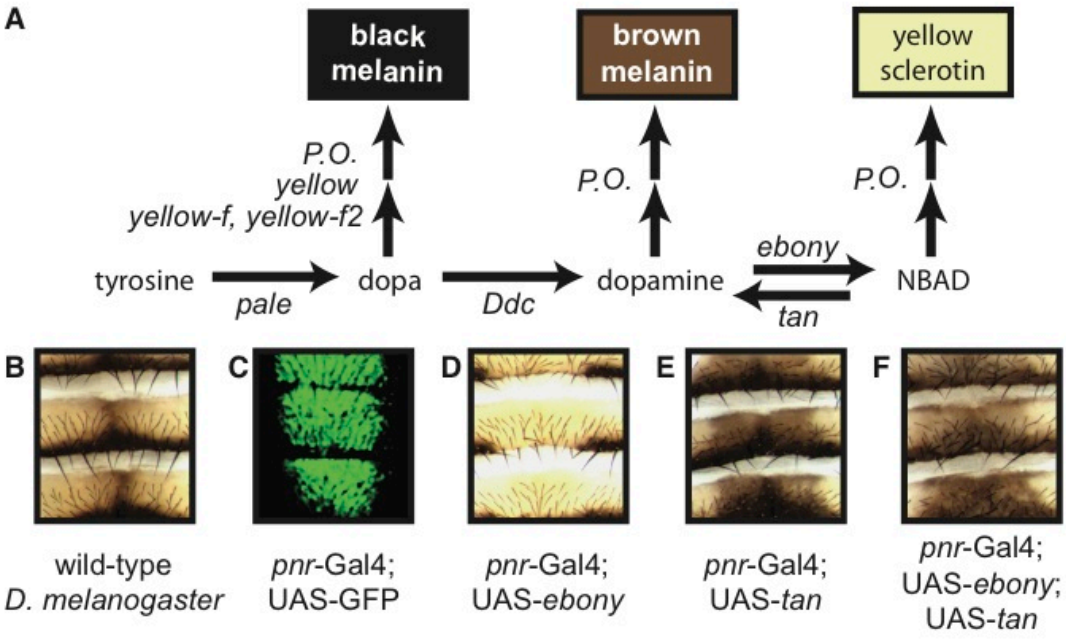


Figure 2.5: (A) A simplified melanin biosynthesis pathway is shown. For a more complete pathway, see (TRUE *et al.* 2005). The gene(s) controlling each enzymatic step are shown in italics. *P.O.* Indicates genes encoding phenol oxidase proteins. Branches with two consecutive arrows include multiple enzymatic steps that are not well defined. (B-F) Dorsal abdominal cuticle (segments A3-A5) is shown for various *D. melanogaster* genotypes. (B) Canton-S, a wild-type strain of *D. melanogaster*, shows the striped dorsal abdominal pigment pattern typical of this species. (C) Expression of UAS-GFP (green) shows that *pnr*-Gal4, the driver used to ectopically express Ebony and Tan in panels D-F, activates gene expression in a stripe along the dorsal midline during late pupal development. (D) Ectopic expression of UAS-Ebony caused increased yellow pigmentation; (E) Ectopic expression of UAS-Tan caused increased brown pigmentation; and (F) the simultaneous expression of both UAS-Tan and UAS-Ebony resulted in an intermediate phenotype. Cuticle is from 3-5 day old females in all panels except (C), in which cuticle is from a female pupa just prior to eclosion (stage P15).

Non-coding changes in tan *contribute to pigmentation divergence*

The above results are consistent with changes in *ebony* and *tan* contributing to pigmentation divergence, but cannot distinguish divergence affecting these genes from divergence affecting linked loci. This is particularly concerning for *ebony* because it is located in a part of the genome that is inverted between species (WITTKOPP *et al.* 2003b, HSU, 1951). Inversions effectively suppress recombination, precluding genetic dissection of the region. Nonetheless, differences in Ebony protein expression between *D. americana* and *D. novamexicana* (WITTKOPP *et al.* 2003b) strongly suggest that this gene is involved in pigmentation divergence.

Unlike *ebony*, *tan* is in a freely recombining region of the genome. This allowed us to use fine-scale genetic mapping to separate the effects of *tan* from neighboring genes and to determine whether *tan* contributes to the altered pigmentation observed in the *tan* introgression line (Figure 2.2B). A 2.7 kb region of *tan* was identified that contributes to pigmentation divergence (Figure 2.6) and contains 57 single nucleotide differences and 19 insertions or deletions (indels) (Figure 2.7). All of these changes affect non-coding sequences, and the region includes the entire first intron (Figure 2.6). Differences located 3' of this region must also affect pigmentation, however, because the recombinant fly inheriting *D. americana tan* sequence only in this region was not as dark as flies inheriting *D. americana* sequence for the full *tan* gene (Figure 2.6). Within *tan*, this 3’ region

includes many non-coding differences as well as two non-synonymous differences that affect amino acids 190 and 267.

Figure 2.6: A 2.7kb region of *tan* contributes to pigmentation divergence

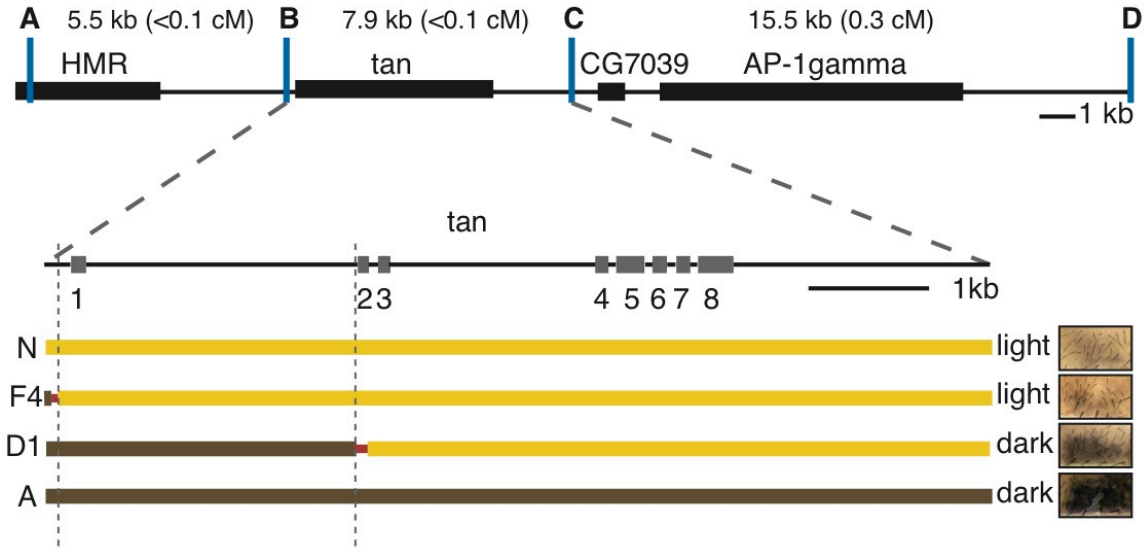


Figure 2.6: Genomic structure of *tan* and flanking genes is shown to scale. Genetic distances indicated between molecular markers (labeled A-D) were determined empirically. Below, a more detailed schematic of *tan* is shown with vertical dotted lines delineating the 2.7 kb region containing sites inferred to affect pigmentation, and gray boxes representing exons. Recombination breakpoints occurred between positions 689 and 752 in F4 and between positions 3500 and 3658 in D1. Representations of recombinant genotypes, their corresponding phenotypic classifications (light or dark), and pictures of dorsal abdominal cuticle recovered after DNA extraction (see Materials and Methods) are also shown. Yellow represents *D. novamexicana* sequence and brown represents *D. americana* sequence. Red bars show regions that do not differ between species. Genotypes A and N show flies with *D. americana* or *D. novamexicana* alleles at molecular markers B and C, respectively.

Figure 2.7: Sequence alignment reveals candidate sites for functional divergence of *tan* alleles

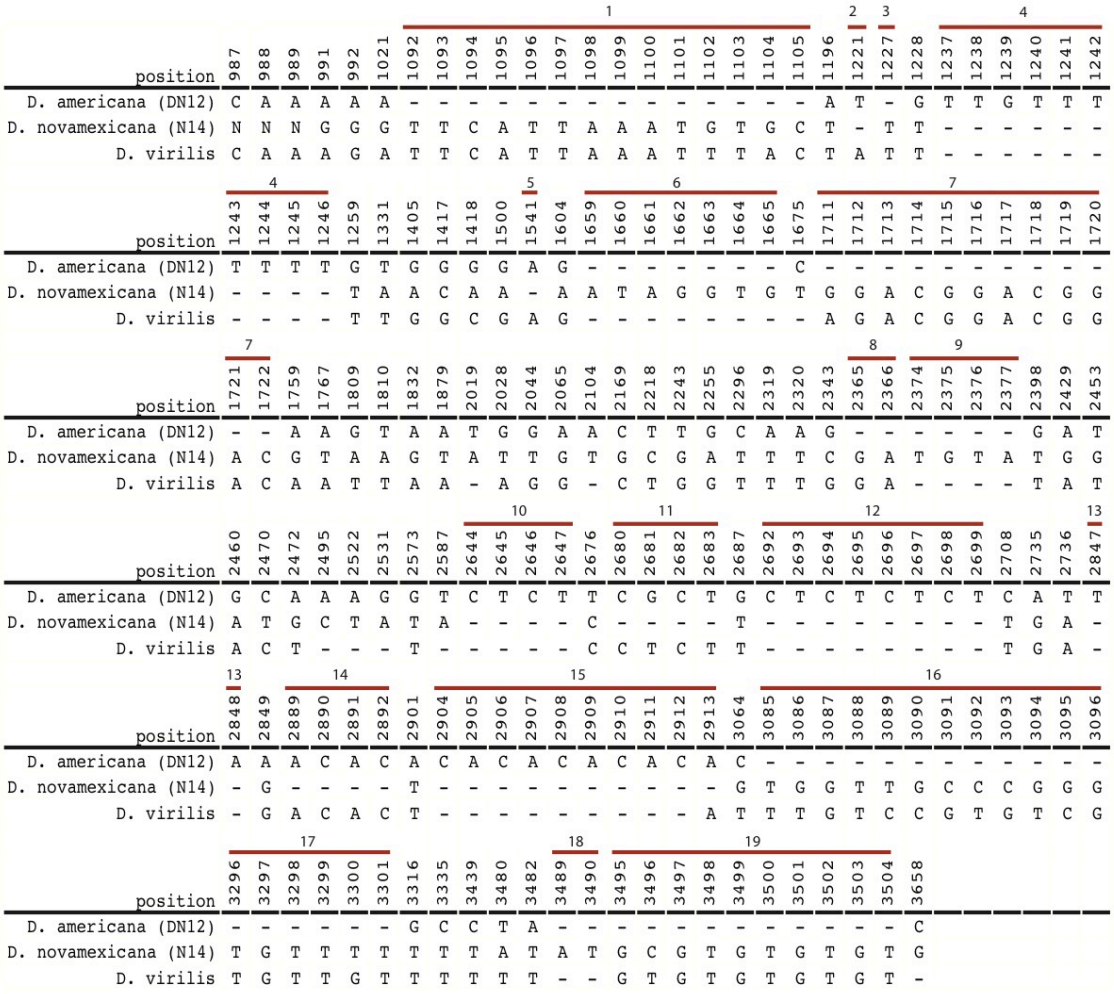


Figure 2.7: Fine- scale genetic mapping identified a 2.7 kb region of *tan* that is functionally divergent between *D. americana* and *D. novamexicana* (Figure 2.6). All divergent sites between the DN12 allele of *D. americana* and the N14 allele of *D. novamexicana* are shown. *D. virilis* alleles at these sites are also shown. Site positions refer to PopSet alignment of GQ457336-GQ457353. The 2.7 kb candidate region begins at position 752 and extends through position 3658 in this alignment. The red lines numbered 1 to 19 indicate indels.

tan *expression correlates with pigmentation differences*

Given the absence of coding changes in the 2.7 kb mapped region of *tan*, we expect that divergent sites in this region affect pigmentation by altering *tan* expression. Because of its darker pigmentation, we hypothesized that *D. americana* has higher levels of *tan* expression than *D. novamexicana*. *in situ* hybridization showed that *tan* is expressed throughout each dorsal abdominal segment ("tergite") in both species during the P14 and P15 pupal stages (Ashburner, 1989) when pigmentation develops (Figure 2.8A, B). This expression pattern correlates with the distribution of pigments in adult *D. americana* and *D. novamexicana* tergites, and is distinct from the patterns of *tan* expression in Drosophila species with other pigment patterns (JEONG *et al.* 2008). Differences in *tan* expression detected with *in situ* hybridization correlate with pigmentation divergence in these other species (JEONG *et al.* 2008), yet we saw no obvious expression differences between *D. americana* and *D. novamexicana* during the same developmental stages with this technique (Figure 2.8A, B).

Figure 2.8: *tan* mRNA is more abundant in *D. americana* than in *D. novamexicana*

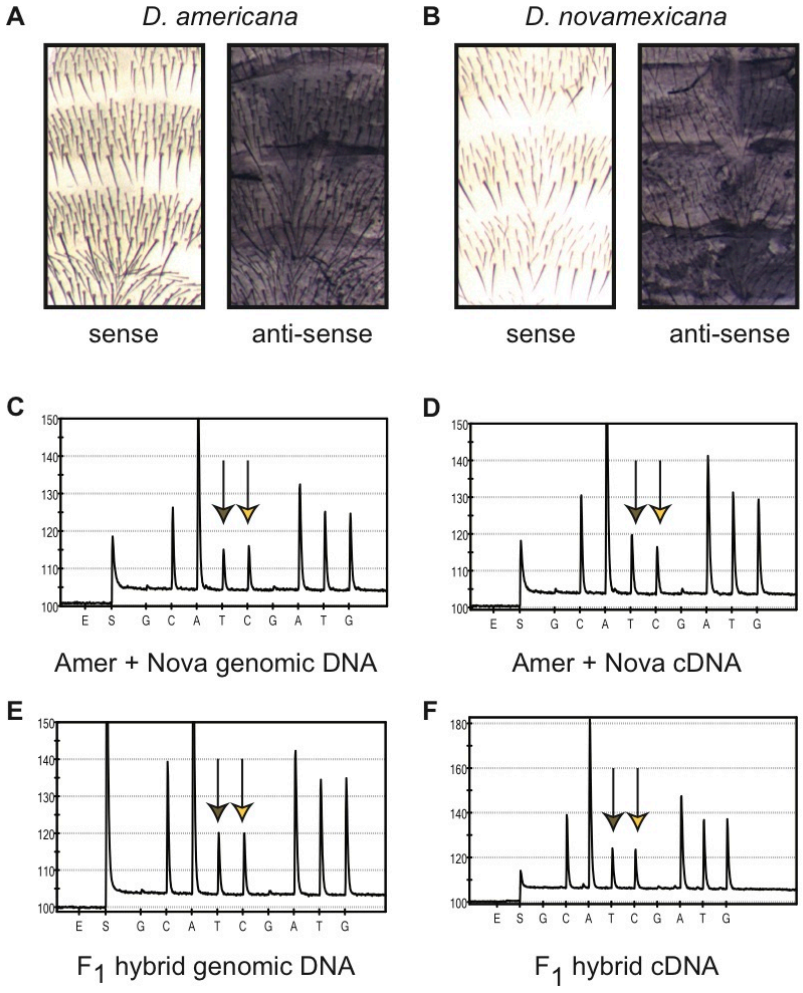


Figure 2.8: *in situ* hybridization with an oligonucleotide probe complementary to *tan* mRNA (anti-sense) showed similar expression of *tan* in *D. americana* (A) and *D. novamexicana* (B). Control probes composed of sequences identical to the *tan* mRNA (sense) showed no hybridization signal in either species (A, B). The intensity of staining was variable from cuticle to cuticle and experiment to experiment, and no systematic differences in hybridization signal were apparent between species. To more quantitatively compare *tan* expression between species, we used Pyrosequencing (WITTKOPP *et al.* 2004). Sample pyrograms are shown for (C) genomic DNA extracted from a pool of *D. americana* and *D. novamexicana* pupae, (D) cDNA synthesized from

RNA extracted from the same pool of pupae, (E) F1 hybrid genomic DNA, and (F) cDNA synthesized from F1 hybrid pupal RNA. Letters below each pyrogram refer to the sequential addition of enzymes (E), substrate (S), and single nucleotides (A, G, C, or T) to the reaction. (C-F) *D. americana*-specific peaks are indicated by brown arrowheads and *D. novamexicana*-specific peaks are indicated by yellow arrowheads.

To quantitatively compare levels of *tan* expression, we measured the relative abundance of *tan* transcripts in stage P14 and P15 pupae of each species with Pyrosequencing (AHMADIAN 2000). We observed an average of 34% more *tan* transcripts in *D. americana* females than in *D. novamexicana* females ($n = 4$ samples, each containing 6 flies; $t = 3.7$, $P_{t\text{-test}} = 0.03$; Figure 2.8C, D), consistent with the darker pigmentation of *D. americana*. To determine whether this expression difference results from *cis*-regulatory divergence of *tan*, we compared transcript abundance of *D. americana* and *D. novamexicana tan* alleles in F_1 hybrid females during the same pupal stages with Pyrosequencing. Surprisingly, no significant differences in allele-specific expression were observed ($n = 5$ samples, each containing 6 flies, $t = 0.72$, $P_{t\text{-test}} = 0.51$; Figure 2.8E, F). Divergent expression levels may therefore be caused by differences in *trans*-regulatory factors and/or differences in the number of *tan* expressing cells between species (WITTKOPP *et al.* 2004). The non-coding differences we identified by fine-scale genetic mapping may alter fine-scale temporal control (*c.f.,* CONG *et al.* 2002) and/or posttranscriptional regulation (*c.f.,*LAURIE and STAM 1994) of *tan*. It is also possible that these non-coding differences may affect transcriptional regulation of a neighboring gene that is also involved in

pigmentation (*c.f.,*JEONG *et al.* 2008) or a cryptic, small, non-coding RNA encoded by the *tan* intron (*c.f.,*STARK *et al.* 2007).

Phenotypic consequences of tan *divergence revealed in transgenic flies*

To determine whether evolutionary changes in the *tan* gene itself are sufficient to affect pigmentation, we inserted transgenes carrying the *D. americana* and *D. novamexicana tan* alleles into the *D. melanogaster* genome (see Materials and Methods). Both transgenes were integrated at the same site, allowing us to compare pigmentation of flies whose genomes differed only for divergent sites within the transgenes. Both the *D. americana* and *D. novamexicana* alleles of *tan* rescued pigmentation in a *D. melanogaster tan* null mutant (Figure 2.9A-D), indicating that the transgenes were expressed in *D. melanogaster* and that Tan protein function is (at least largely) conserved. Flies carrying the *D. americana tan* allele had darker pigmentation than flies carrying the *D. novamexicana tan* allele (Figure 2.9C, D; $F = 26.94$, $P < 0.0001$ for abdominal segments A3 and A4, and $F = 6.51$, $P = 0.03$ for the darker A5 segment). This is consistent with the darker pigmentation of *D. americana* relative to *D. novamexicana*. We also compared the phenotypic effects of *D. americana* and *D. novamexicana tan* alleles in *D. americana* and *D. novamexicana* themselves by randomly inserting both *tan* transgenes into the genomes of both species. Two independent insertions were recovered for each transgene in each species. In *D. americana*, we were unable to detect a difference in pigmentation between transformed and untransformed flies,

presumably because of the already dark pigmentation of this species (see Figure 2.1B). In *D. novamexicana*, however, transformant flies carrying the *D. americana tan* transgene (Figure 2.9F) were visibly darker than flies carrying the *D. novamexicana tan* transgene (Figure 2.9E).

Figure 2.9: The *D. americana* allele of *tan* causes darker pigmentation than the *D. novamexicana* allele of *tan*

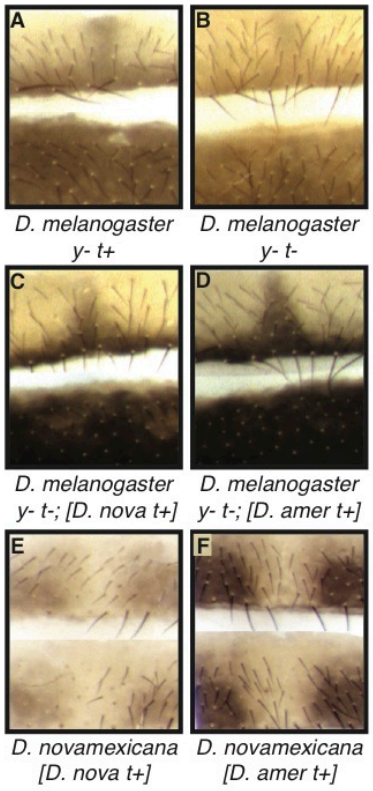


Figure 2.9: Transgenes containing *tan* alleles from *D. americana* and *D. novamexicana* were transformed into *D. melanogaster* and *D. novamexicana*. In *D. melanogaster*, transgenes were crossed into a genetic background homozygous for null mutations in *yellow* (*y*-) and *tan* (*t*-). The *yellow* mutation was used to lighten pigmentation, making the effects of *tan* transgenes easier to see. (A) *D. melanogaster yellow* (*y*-) mutant, which is wild-type for *tan*. (B) *D. melanogaster*

yellow, *tan* (*y-, t-*) double mutant. (C) *D. melanogaster yellow, tan* mutant carrying the *D. novamexicana tan* transgene ([*D. nova t+*]). (D) *D. melanogaster yellow, tan* mutant carrying the *D. americana tan* transgene ([*D. amer t+*]). (E) Wild-type *D. novamexicana* carrying the *D. novamexicana tan* transgene. (F) Wild-type *D. novamexicana* carrying the *D. americana tan* transgene.

ebony *and* tan *QTLs underlie variable pigmentation within* D. americana

Pigmentation of *D. americana* is always distinct from that of *D. novamexicana* (THROCKMORTON, 1982), but the intensity of dark pigmentation varies within *D. americana*. This variation is geographically structured, with *D. americana* captured in the eastern United States visibly darker than those captured from the western part of the species range (THROCKMORTON, 1982). These pigmentation differences remain visible after rearing flies under common environmental conditions in the laboratory (Figure 2.11A), indicating a genetic basis for the pigmentation cline. Sequence variation within *D. americana* at putatively neutral loci shows no population structure (MCALLISTER 2003; VIEIRA *et al.* 2003; MASIDE *et al.* 2004; SCHAFER *et al.* 2006; MCALLISTER and EVANS 2006; MORALES-HOJAS *et al.* 2008), suggesting that this cline is due to local adaptation.

To determine whether sites linked to *ebony* and/or *tan* contribute to this intraspecific polymorphism, we used *D. novamexicana* alleles as a reference to compare the phenotypic effects of *ebony* and *tan* QTL alleles among lines of *D. americana.* Genetic mapping was performed using four isofemale lines of *D. americana* (A01, DN2, DN4, and DN12) with lighter pigmentation than the line of *D. americana* (A00) used previously. *D. americana* females from each strain were

crossed to *D. novamexicana* males; F_1 hybrid females were backcrossed to *D. novamexicana* males; and molecular markers in *ebony* and *tan* were genotyped in flies from the lightest and darkest pigmentation classes (Figure 2.10). Despite their lighter pigmentation, the DN12 and DN4 lines of *D. americana* produced similar mapping results to A00: both *ebony* and *tan* showed highly significant linkages to loci affecting pigmentation (Figure 2.11B). Genotyping 101 males from the DN4 backcross population and fitting their genotypes and phenotypes to a linear model (see Materials and Methods) showed that *ebony* (E) and *tan* (T) both have significant additive effects on pigmentation ($F_E = 132.98$, $P_E < 0.0001$; $F_T = 160.85$, $P_T < 0.0001$) with no significant epistatic interaction between them ($F = 3.02$, $P = 0.09$). These additive effects explained 76% of the pigmentation variance in the DN4 backcross population. For DN2, sites linked to *tan* contributed to pigmentation differences between species, but sites linked to *ebony* did not (Figure 2.11B). The converse was true for A01 — sites linked to *ebony* contributed to pigmentation differences between species, while sites linked to *tan* did not (Figure 2.11B).

Taken together, these data reveal three distinct genotypes among *D. americana* lines with a light pigmentation phenotype. DN2 has alleles linked to *ebony* that appear to be functionally equivalent to those found in *D. novamexicana*; A01 has alleles linked to *tan* that appear to be functionally equivalent to those found in *D. novamexicana*; and DN4 and DN12 have alleles linked to *tan* and *ebony* that appear to be functionally distinct from those found in *D. novamexicana*. It remains to be seen whether the DN12 and DN4 alleles of

these QTLs have the same effect on pigmentation as each other or as the alleles from the darker A00 line of *D. americana*. The three *D. americana* lines starting with “DN” were collected from Duncan, Nebraska, with DN2 and DN4 collected in the same year and DN12 collected the following year (Table 2.2), suggesting that genetic heterogeneity for pigmentation exists within this local population. This heterogeneity may be caused by gene flow among populations and/or balancing selection within the population.

Figure 2.10: Different phenotypic distributions were observed among backcross progeny from different strains of *D. americana*

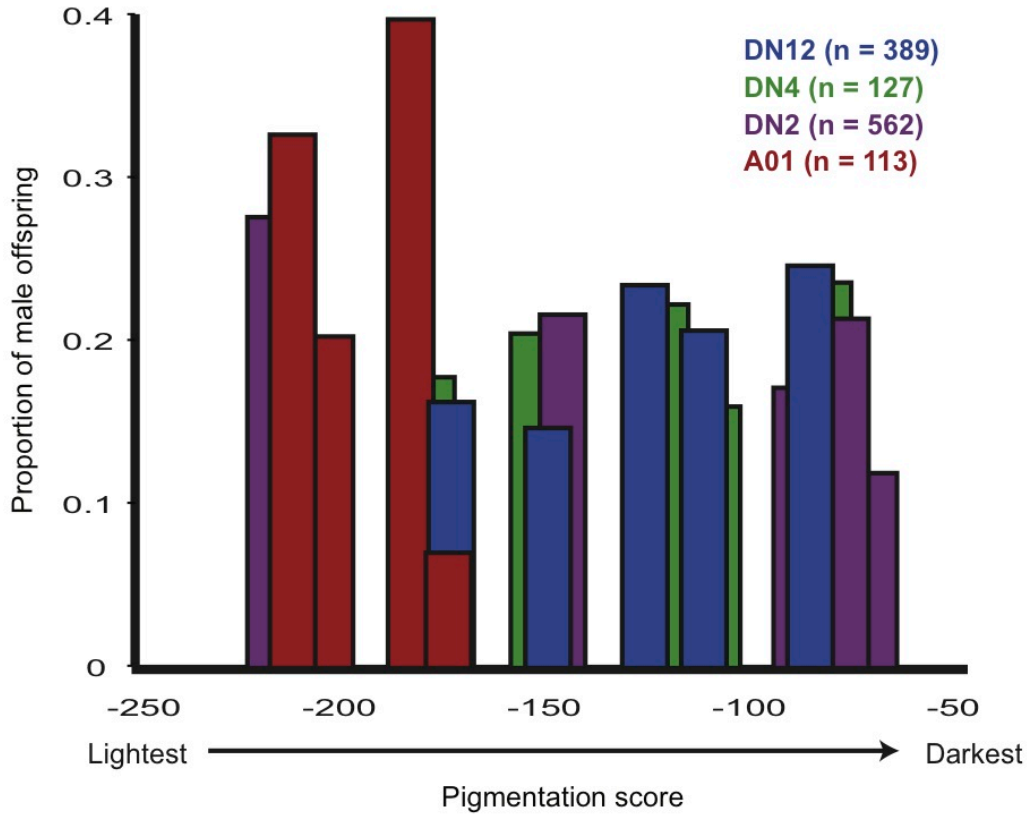


Figure 2.10: Females from the DN12, DN4, DN2 and A01 strains of *D. americana* were each mated to *D. novamexicana* (N14) males; F1 females were backcrossed to *D. novamexicana* (N14) males; and male progeny were classified by eye into as many distinguishable pigmentation classes as possible. A histogram of pigmentation phenotypes is shown, with the sample size (n) for each backcross population listed in the key. Pigmentation classes are plotted on the X-axis on the basis of the average pigmentation intensity of five randomly chosen flies from that class. DN12 (blue) and DN4 (green) backcrosses produced similar phenotypic distributions in terms of both the frequency and pigmentation of each phenotypic class. These distributions were distinct from that observed with A00 (Figure 2.3). The DN2 backcross also produced a phenotypic distribution distinct from A00 (Figure 2.3), but this population was also distinct (in both pigmentation phenotypes and frequency) from the DN12 and DN4 backcross distributions. Crosses with the A01 line of *D. americana* produced only four recognizable classes of backcross progeny, all of which were lighter than most backcross progeny produced with DN2, DN4, or DN12.

Figure 2.11: *ebony* and *tan* QTLs also contribute to polymorphism

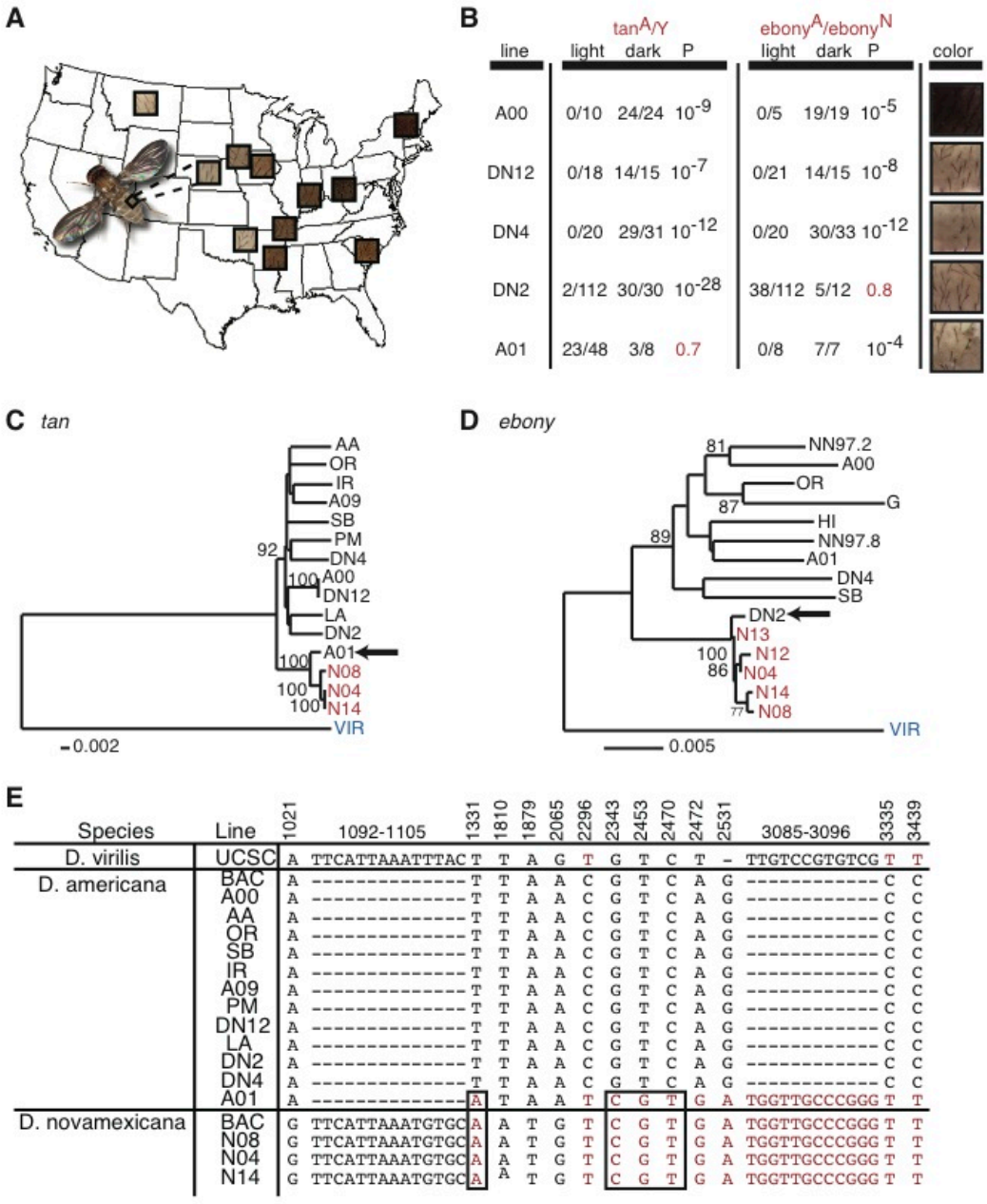


Figure 2.11: (A) Dorsal abdominal cuticle from *D. americana* isofemale lines (Table 2.2) is shown. Eastern populations are darker than western populations. (B) Phenotypes, genotypes, and statistical significance for interspecific QTL mapping experiments are shown. Dorsal abdominal pigmentation of each of isofemale line is shown in the color column, the middle columns show the

proportion of male backcross progeny genotyped from the lightest (light) and darkest (dark) pigmentation classes with *D. americana tan* (tan^A) and *ebony* ($ebony^A$) alleles. P-values (P) are from 2x2 Fisher's exact tests of the genotype count data. (C, D) Neighbor-joining trees of *tan* (C, 7625 bp) and *ebony* (D, 1136 bp) from *D. americana* (black), *D. novamexicana* (red), and *D. virilis* (blue) are shown with bootstrap values >75% (n=1000). Branch lengths are to scale. (E) Fixed differences within the 2.7kb candidate region of *tan* are shown. Sites 889-981 and 3521-3616 are exons 1 and 2, respectively. The *D. virilis* allele is from the 2005 assembly of the *D. virilis* genome sequence (DROSOPHILA 12 GENOMES CONSORTIUM *et al.* 2007). Positions refer to alignment of GQ457336–GQ457353. Alleles shared between *D. novamexicana* and A01 are red; the derived subset, relative to *D. virilis,* is boxed.

Shared pigmentation alleles contribute to polymorphism and divergence

The functional similarity observed for the A01 and *D. novamexicana* alleles linked to *tan*, as well as for the DN2 and *D. novamexicana* alleles linked to *ebony,* may result from shared ancestry (i.e., alleles that are identical-by-descent) or from convergent evolution. Sequences of *tan* and *ebony* from multiple lines of *D. americana* and *D. novamexicana* show that the functional similarity most likely reflects shared ancestry as the A01 *tan* sequence is more similar to *D. novamexicana* alleles than to other *D. americana* alleles (Figure 2.11C) and the DN2 *ebony* sequence is more similar to *D. novamexicana* alleles than to other *D. americana* alleles (Figure 2.11D). The DN2 line of *D. americana* also has the same arrangement of the *ebony*-containing inversion ("In(2)b" in Hsu, 1951) as *D. novamexicana* (Figure 2.12, MENA AND MCALLISTER), further suggesting that pigmentation alleles linked to *ebony* in DN2 and *D. novamexicana* have a common origin.

Figure 2.12: The genomic region containing *ebony* is inverted between *D. novamexicana* and DN12 line of *D. americana*, but not the DN2 line of *D. americana*

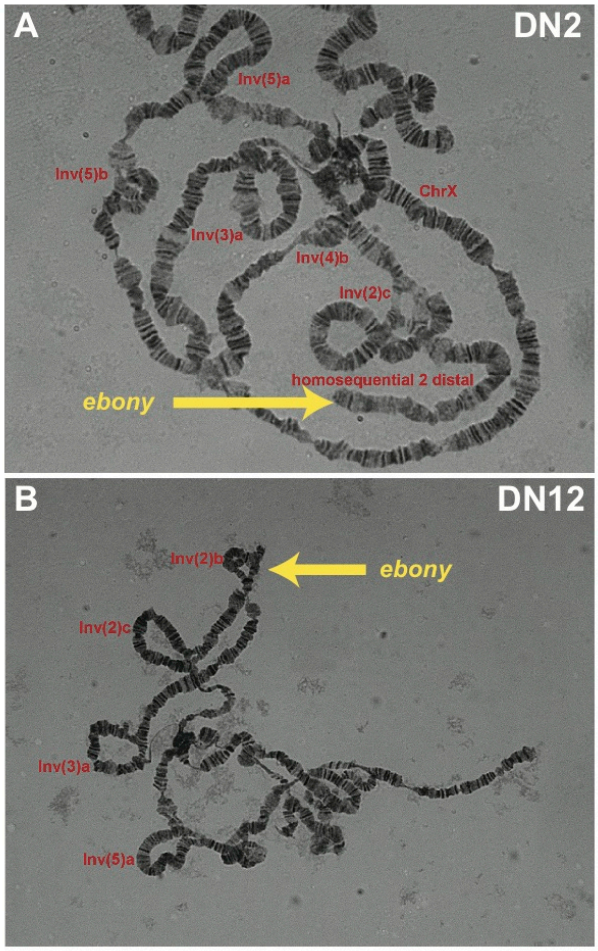
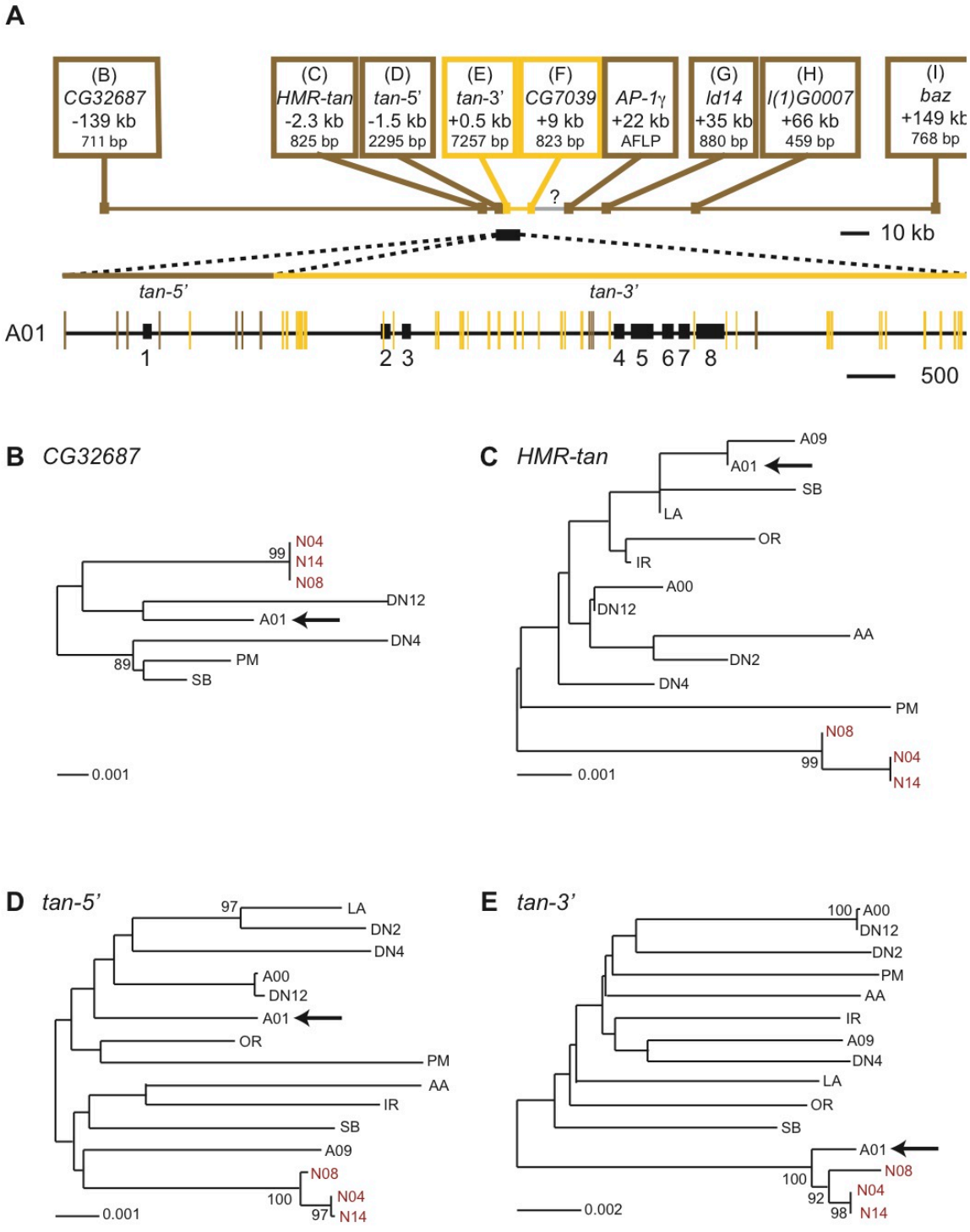


Figure 2.12: Polytene chromosomes from interspecific F1 hybrid males show the location of inversions between *D. novamexicana* and two strains of *D. americana*: DN2 (A) and DN12 (B). The distal end of the 2nd chromosome, which contains *ebony*, is indicated with a yellow arrow in each panel. Note the presence of the *ebony*-containing In(2)b inversion in DN12 (B) but not DN2 (A). Other visible inversions are also indicated. This experiment was performed and analyzed by Paulina Mena and Bryant McAllister (University of Iowa); images are reproduced here with permission.

Sequence variation identifies candidate sites for divergent pigmentation

Sequence similarity between A01 and *D. novamexicana* was found to be highest beginning in the first intron of *tan* and extending 3' of *tan* (Figure 2.13). Within the 2.7 kb region identified by fine-scale mapping (Figure 2.6), we observed thirteen fixed single nucleotide differences and two fixed indels between *D. americana* (excluding A01) and *D. novamexicana* (Figure 2.11E). The A01 allele of *D. americana* contains the same sequence as *D. novamexicana* at nine of these thirteen divergent sites and shares one of the two indels (Figure 2.11E, red). Only four of the shared substitutions are derived changes relative to *D. virilis* (Figure 2.11E, boxed). Because *D. virilis* has pigmentation similar to *D. americana* (Figure 2.1B), we consider these four non-coding changes to be the best candidates for divergent function in this region. Derived changes outside of this region that are also unique to A01 and *D. novamexicana tan* may contribute to pigmentation divergence as well. The two non-synonymous differences between alleles used for fine-scale mapping are both polymorphic and thus unlikely to contribute to fixed differences between species.

Figure 2.13: Haplotype sharing between *D. novamexicana* and the A01 line of *D. americana* is limited to sequences within and near *tan*



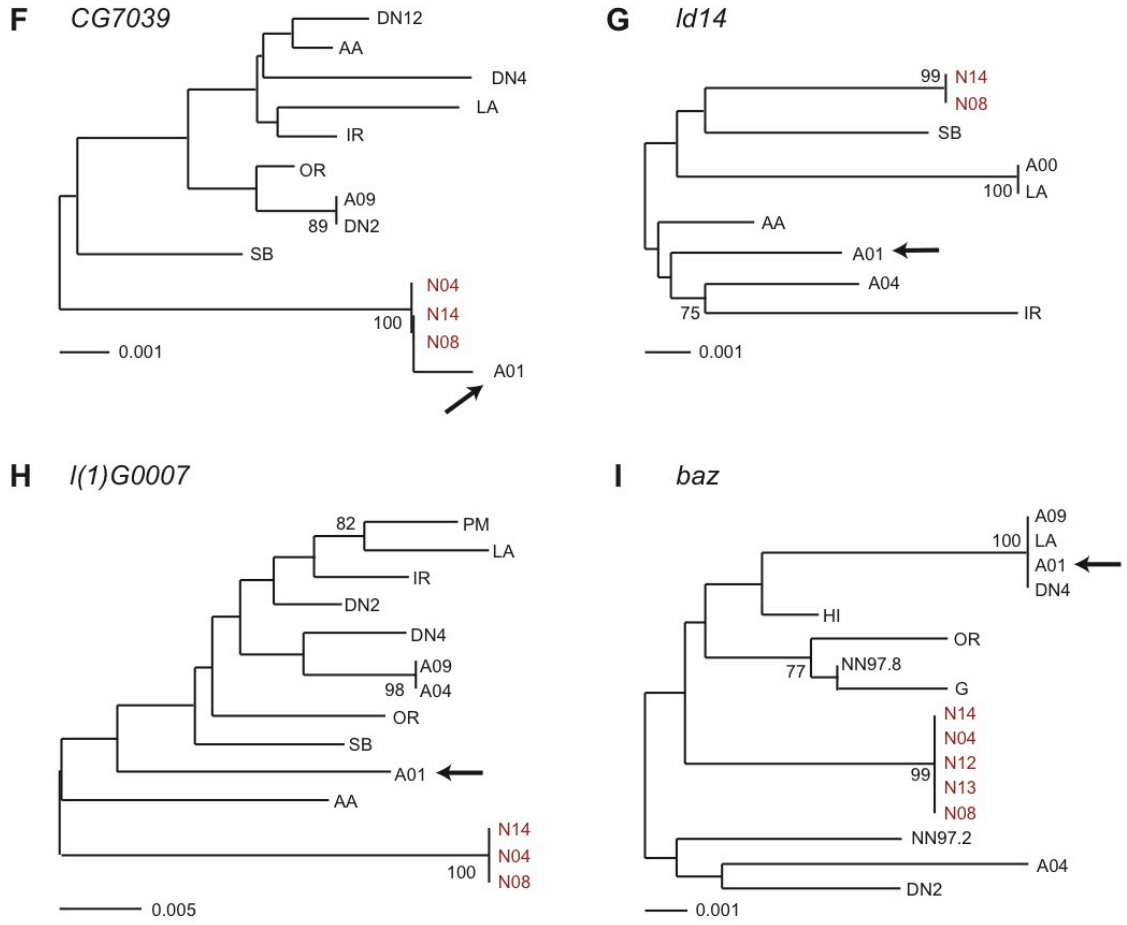


Figure 2.13: To determine the extent of haplotype sharing, we sequenced multiple loci in and around *tan* (A) from multiple lines of *D. americana* and *D. novamexicana* and constructed a neighbor joining tree for each locus (B - I). Nodes appearing in >75% of 1000 replicate bootstrap trees are indicated by their exact frequency, and branch lengths are drawn to scale. Arrows indicate the location of A01 within each tree, and all *D. novamexicana* alleles are shown in red. (A) A summary of these neighbor joining trees is shown. Loci for which the A01 allele was most similar to other alleles of *D. americana* are shown in brown, and loci for which the A01 allele was most similar to *D. novamexicana* alleles are shown in yellow. The gene containing (or adjacent to) each locus is indicated with its position, in kilobases (kb), relative to the start of the *tan* gene (i.e., *D. virilis*, scaffold_12932: 1,852,445). The number of aligned basepairs (bp) analyzed for each locus (excluding gaps) is also indicated. The one exception is the *AP-1 Gamma* (*AP-1*) locus,

which was genotyped with an amplified fragment length polymorphism. 840 bp of sequence (GQ457370-GQ457381) from the region shown in gray (indicated by the question mark) was also analyzed, but did not contain enough fixed differences to distinguish the *D. americana* and *D. novamexicana* alleles with bootstrap support greater than 75%. A larger schematic of the *tan* region, including exons 1 – 8 (black boxes), is also shown in panel A, with vertical lines indicating the positions of fixed differences between species GQ457336-GQ457353. Sequence of the A01 allele is indicated at each of these divergent sites, with brown lines representing *D. americana* alleles and yellow lines representing *D. novamexicana* alleles. On the basis of the pattern of haplotype sharing, this region was divided into two sections (*tan*-5' and *tan*-3') for phylogenetic analysis.

A model of pigmentation evolution

Our data reveal the relationship between intraspecific polymorphism and interspecific divergence by showing that the same alleles contribute to pigmentation differences within and between species (Figure 2.14). These alleles may have been present in the common ancestor of *D. americana* and *D. novamexicana* or may have arisen in *D. novamexicana* and subsequently introgressed into *D. americana* following hybridization. Distinguishing between these two scenarios is notoriously difficult (HOLDER *et al.* 2001; NOOR and FEDER 2006; DEGNAN and ROSENBERG 2009), although haplotype sharing between *D. americana* and *D. novamexicana* has been postulated to be due to shared ancestral variation (MORALES-HOJAS *et al.* 2008). Our data are consistent with this interpretation: sequences of *D. americana* alleles that appear to have the same function as *D. novamexicana* alleles are basal to these *D. novamexicana* alleles in gene trees (Figure 2.11C, D); there is evidence of recombination (or

gene conversion) within the *D. americana* A01 *tan* haplotype (Figure 2.8A) which argues against a recent introgression event; and *D. novamexicana* is thought to have evolved from a peripheral population of the common ancestor shared with *D. americana* (MORALES-HOJAS *et al.* 2008). Therefore, we propose that light pigmentation alleles segregating in this common ancestor became fixed in *D. novamexicana,* contributing to its yellow body color, and continue segregating in *D. americana*, contributing to clinal variation. Additional *D. novamexicana*-like alleles of *D. americana* are needed to further evaluate this model.

Figure 2.14: A genetic model of pigmentation polymorphism and divergence

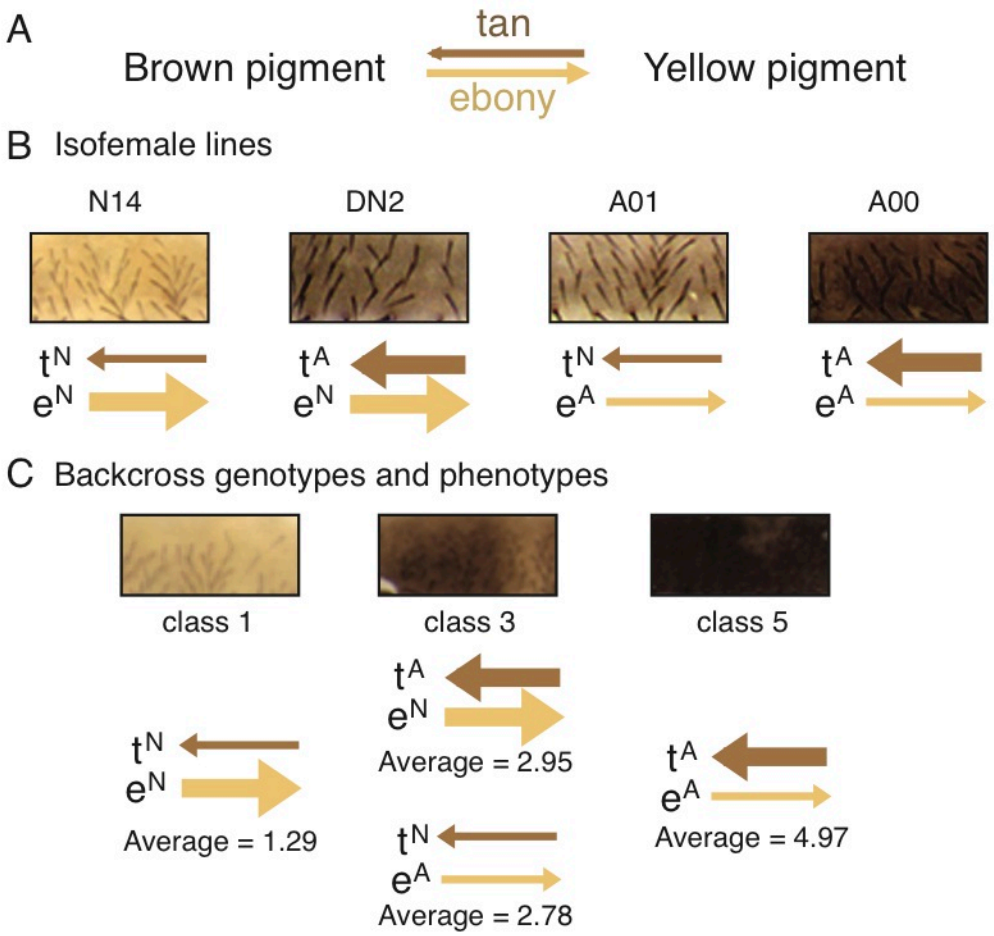


Figure 2.14: (A) *ebony* and *tan* control the production of yellow and brown pigments, respectively (Figure 2.5A), and the relative expression of these proteins determines adult pigmentation (Figure 2.5F). Tan is expressed at higher levels in *D. americana* (Figure 2.6D), and Ebony is expressed at higher levels in *D. novamexicana* (WITTKOPP *et al.* 2003b). To reconcile the genotypes and phenotypes observed in this study, we propose a model of pigmentation evolution in which Tan and Ebony expression differences between species are caused by changes in the *tan* and *ebony* genes themselves. Specifically, we assume that the *D. americana tan* allele (*tA*) causes greater Tan protein expression than the *D. novamexicana* allele (*tN*), and that the *D. novamexicana ebony* allele (*eN*) causes greater Ebony protein expression than the *D. americana* allele (*eA*). For simplicity, we ignore the contribution of other genes, which in the DN4 backcross population explained up to 24% of the pigmentation variance. (B) *ebony* and *tan* genotypes are shown for isofemale lines of *D. americana* (DN2, A01, and A00) and *D. novamexicana* (N14). The thicker yellow arrow represents greater activity of *eN* relative to *eA*, and the thicker brown arrow represents greater activity of *tA* relative to *tN*. In N14, which carries *tN* and *eN*, our model predicts a net production of yellow pigment. In A00, which carries *tA* and *eA*, our model predicts a net production of brown pigment. In DN2 and A01, both of which carry one allele from each species, our model predicts pigmentation intermediate between A00 and N14. Dorsal abdominal cuticle from the A4 segment is shown for 7-10 day old males. (C) The four possible *tan* and *ebony* genotypes are shown along with the average pigmentation class for each genotype in the DN4 backcross population. Dorsal abdominal cuticle from the A4 segment of male flies categorized as class 1, class 3, and class 5 in the DN4 backcross are also shown. Flies inheriting *eN* and *tN* had light pigmentation, presumably because of excess *ebony* activity, while flies inheriting *eA* and *tA* had dark pigmentation, presumably because of excess *tan* activity. Flies inheriting the *D. americana* allele of only *ebony* or *tan* had intermediate pigmentation, presumably because of more balanced *tan* and *ebony* activities. The distribution of phenotypes in the DN4 backcross population is consistent with a simple two-locus Mendelian model assuming that class 1 contains

tN;eN flies, classes 2, 3, and 4 contain *tN;eA* and *tA;eN* flies, and class 5 contains *tA;eA* flies ($\chi 2 = 4.7$, df = 2, P = 0.1).

Materials and Methods:

Sequencing D. americana *and* D. novamexicana tan *and* ebony

Bacterial artificial chromosomes (BACs) containing *tan* and *ebony* were identified in *D. americana* and *D. novamexicana* genomic libraries by screening filters from the Arizona Genomics Institute. Radioactively labeled probes used for screening were produced with PCR products amplified by the following primers: *HMR*, located 5' of *tan* (5'-CATCTCGTCCAACTCCAGGT-3' and 5'-GCGCTATAAATATCAGCGTCA-3'); *CG7039* located 3' of *tan* (5'-CATTGCTGCACGGCTTTTAC-3' and 5'-CTCCACCAGCCATTTGATG-3'); *ETHR* located 5' of *ebony* (5'- GGCTGTCGCTGCTGTTATTT-3' and 5'-CCAAGCCGCAAATAAGTTTC-3'); and *CG5874* located 3' of *ebony*, (5'-GCCTGCACCTGCACCA-3' and 5'- CCACGCTAATTCCAACCAAC-3'). These primers were designed on the basis of sequence from the August 2005 *D. virilis* genome assembly (DROSOPHILA 12 GENOMES CONSORTIUM *et al.* 2007). We ordered six *D. americana* and six *D. novamexicana* clones positive for both genes flanking *tan* and five *D. americana* and ten *D. novamexicana* clones positive for both genes flanking *ebony* from the Arizona Genomics Institute. Each clone was evaluated by testing for amplification of PCR products from *tan*, *ebony*, and flanking genes. Ultimately, the DA_ABa0020L7 (*D. americana*) and DN_Ba0024C15 (*D. novamexicana*) clones containing *tan* as well as the

DA_ABa0029H3 (*D. americana*) and DN_Ba0007J18 (*D. novamexicana*) clones containing *ebony* were selected for sequencing. With a combination of primers designed on the basis of the *D. virilis* sequence and primer walking, we sequenced ~14 kb from each *tan* BAC clone and ~4kb from each *ebony* BAC clone.

Fly strains, rearing, and imaging

The following lines of *D. melanogaster* were used for this work: *pnr*-Gal4 (G. Morata); *w-* ;P[*w+*, UAS-Tan] (J. True); *w-*;P[*w+*, UAS-Ebony] (J. True); *yellow1*, *tan5* (J. True), and CantonS. A *w-*;*CyO/Sb*; *TM2/TM6* balancer line was used to construct genotypes described in Figure 2.5. *D. americana* and *D. novamexicana* lines and sources are shown in Table 2.2. All flies were reared on standard yeast-glucose media, with *D. americana* and *D. novamexicana* reared at 20°C, and *D. melanogaster* reared at 25°C. Pigmentation of individual flies was documented by placing age-controlled adults in a solution of 10% glycerol in ethanol, storing at room temperature for 1-30 days, dissecting the dorsal abdominal cuticle, mounting the cuticle in Hoyer's solution, and imaging with a Scion 1394 (Frederick, MD) digital camera. All images presented within a figure (or within a panel for Figure 2.11) are from age-matched flies with images captured at the same time, under the same lighting conditions. Colors were adjusted in Photoshop CS2 (Adobe, San Jose, CA) to best reproduce visual observations, with an identical color transformation applied to all images shown within the same figure (or panel for Figure 2.11). Pigmentation of a fly was quantified by

measuring the intensity of gray-scale images with ImageJ (NIH, Bethesda, MD) in five to ten non-overlapping regions and averaging the median intensity from each region. Measurements range from 0 (black) to 255 (white).

Quantitative trait locus (QTL) mapping

All QTL mapping experiments were performed with interspecific backcross populations, as described in the main text. Within each population, progeny were visually classified into distinct pigmentation classes. All phenotypic scoring was performed by E.E.S. under controlled lighting conditions and pigmentation of each fly was verified by P.J.W prior to DNA extraction. Single fly DNA preparations were performed with the protocol described in (GLOOR *et al.* 1993). Genotypes of *tan*, *ebony*, and *yellow* were determined for backcross progeny with either DNA sequencing or amplified fragment length polymorphisms resolved with agarose gel electrophoresis. For *yellow,* genotypes were determined by directly sequencing a 632 bp band amplified with the following PCR primers: 5'-CTAAACATGCCTGAAAATCAATCACGGA-3' and 5'-CGTTGGTAAACGAAAGTCCAATTGG-3'. For *tan*, the primers 5'-CGAGTTTTTATTCCCACTGAATTAT-3' and 5'-GGGTTCGTCTTATCCACGAT-3' amplified a 99 bp band in *D. americana* and a 64 bp band in *D. novamexicana*. For *ebony*, the primers 5'- GTTGTGCCAAACTGAAAGATCC-3' and 5'-CACATTCACACTTTGTGCACTTG-3' amplified a 162 bp band in *D. novamexicana* and a 244 bp band in all *D. americana* lines except DN2, which amplified a 162 bp band identical to *D. novamexicana*. For the DN2 backcross

population, *ebony* genotypes were determined by examining heterozygous bases in chromatograms from directly sequenced 162 bp PCR products.

To test for a significant association between genotypes and phenotypes, two-sided Fisher exact tests, evaluated with "fisher.test()" in R Cocoa GUI 1.12 (http://www.r-project.org/), were used to compare the number of *D. americana* alleles observed among the lightest and darkest flies. For the DN4 backcross population, 101 out of 127 flies were successfully genotyped for both *ebony* and *tan*. The proportion of pigmentation variance explained by *ebony* and *tan* genotypes in the DN4 backcross was determined by fitting the following model with PROC MIXED in SAS v.9.1 (Cary, NC) with Type III sums-of-squares: $Y_{ijk} = E_i + T_j + ET_{ij} + !_{ijk}$, where *Y* is the pigmentation score of each fly, *E* is the *ebony* genotype (i = *D. americana* or *D. novamexicana*), *T* is the *tan* genotype (j = *D. americana* or *D. novamexicana*), *ET* is the interaction between *tan* and *ebony* genotypes, and ! is a random error term. *E*, *T*, and *ET*, were treated as fixed effects in the model. A second model, lacking the interaction term, was used to quantify the amount of variance explained by additive effects of *ebony* and *tan* genotypes.

Constructing introgression lines

D. americana alleles of *ebony* and *tan* from the DN12 line were introgressed into the N14 line of *D. novamexicana*. This line of *D. americana* was used after three failed attempts to introgress alleles from the line of *D. americana* (A00) used for the initial genetic mapping. Note that the *tan* sequences from A00 and

DN12 are very similar (see Figure 2.11C) and differ only by 2 bp in the 2.7 kb region identified by fine-scale genetic mapping (GQ457339 and GQ457347). DN12 and A00 also share the same arrangement of the In(2)b inversion containing *ebony* (MENA AND MCALLISTER).

As shown in Figure 2.4, *D. americana* virgin females were crossed to *D. novamexicana* males, and F_1 hybrid virgin females were backcrossed to *D. novamexicana* males. Virgin females were collected from this first back-cross (BC1), and then we set up twenty matings, each containing one virgin female from BC1 and one *D. novamexicana* male. When third instar larvae were visible (two to three weeks after mating), the female parent was removed from each vial, DNA was extracted, and genotypes at *tan* and *ebony* were determined with DNA sequencing and PCR-based genotyping as described above for QTL mapping. One vial containing larval progeny from a mother heterozygous for *tan* and/or *ebony* was randomly selected, and twenty virgin females were collected from this brood after eclosion. To begin the next backcross generation (BC2), twenty pair-matings were set-up by crossing each of these females to one *D. novamexicana* male. This process was repeated for ten generations.

On average, after ten generations of backcrossing, an introgressed region extends 10 cM to either side of the selected locus (i.e., *ebony* and *tan*) (FALCONER AND MACKAY, 1996). The precise breakpoints of introgressed *tan* and *ebony* regions remain unknown, although DNA sequencing showed that *D. americana* alleles remained at loci ~500 kb from *tan* in both directions. Polytene chromosomes squashes of the *tan* introgression line showed the loss of *D.*

americana alleles for all chromosomal regions inverted between *D. americana* and *D. novamexicana*, and squashes of the *ebony* introgression lines showed the loss of inversions other than the *ebony* containing In(2)b. Importantly, mendelian inheritance of pigmentation was observed for both of the final introgression lines, showing that any remaining *D. americana* alleles not linked to *tan* or *ebony* do not visibly affect pigmentation.

Fine-scale genetic mapping

Virgin females heterozygous for the introgressed *D. americana tan* allele were crossed to *D. novamexicana* males. 5048 male offspring were visually scored for pigmentation by A.H.N. under constant light conditions, with each classified as either light or dark. Light flies had pigmentation most similar to wild-type *D. novamexicana* and dark flies had pigmentation most similar to the *tan* introgression line. DNA was extracted from each fly with the protocol described in (GLOOR *et al.* 1993), except that flies were homogenized by placing a single fly into a well of a 96-well PCR plate, adding a single glass bead, and shaking on a Mixer Mill MM301 (Retsch, Inc., Haan, Germany) for 10 seconds at a frequency of 25 Hz. This shaking condition was found to homogenize the flies sufficiently for DNA extraction without completely destroying the abdominal cuticle.
Genotypes at molecular makers A, B, C, and D in Figure 2.6 were determined with the following primer pairs: (A) 5'-TTATATCGCCGGGTATCAGC-3' and 5'-CGTCTGATGCTTTCTGACGA-3'; (B) 5'-CGAGTTTTTATTCCCACTGAATTAT-3' and 5'-GGGTTCGTCTTATCCACGAT-3'; (C) 5'-

GGAGTCCATGTGGCCTAAGAAC -3' and 5'-GCCTTATCTTAATAGAAGTTTAATATGC-3'; and (D) 5'-TCGAACATGTTTGGCCTTGTCAC-3' and 5'-GTTTATAGCCAGCAGTTGCTG-3'. PCR products from B, C, and D differed in length between the N14 line of *D. novamexicana* and the DN12 line of *D. americana* (i.e., the lines used for fine-scale genetic mapping). The PCR product amplified from locus A was cut with HaeIII, producing different sized fragments for the two alleles. Two flies were found that inherited the *D. americana* allele at one of these loci and the *D. novamexicana* allele at the other. The location of the recombination breakpoint in each of these flies was determined by re-sequencing the *tan* gene. The recombination breakpoint was located between positions 689 and 752 in F4 and between positions 3500 and 3658 in D1. These positions refer to the PopSet alignment of GQ457336-GQ457353. Flies were also genotyped for molecular markers outside of this region (i.e., A and D in Figure 2.6) to estimate genetic distances between A and B as well as between C and D.

Analyzing tan *mRNA expression*

in situ hybridizations were performed as described in (JEONG *et al.* 2008). Briefly, *D. americana* (A00) and *D. novamexicana* (N14) pupae were collected one to four hours prior to eclosion and heated to 100°C for one minute. Dorsal abdominal cuticle was dissected, fixed in 4% paraformaldehyde, dehydrated, and stored at -20°C. After re-hydrating, cuticles were fixed again in 4% paraformaldehyde, treated with proteinase K, fixed a third time, and incubated

overnight at 65°C with a Digoxigenin (DIG)-labeled RNA probe. Sense and anti-sense RNA probes were synthesized with T7 polymerase (Promega, Madison, WI) to transcribe PCR products containing ~300 bp of sequence from *D. novamexicana tan* exon 8. *D. americana* and *D. novamexicana tan* alleles have identical sequence in this region. After washing, samples were incubated overnight with anti-DIG AP Fab fragments (Roche, Basel, Switzerland), washed again, and incubated with a solution containing NBT/BCIP (Promega, Madison, WI) for 20 minutes or until sufficient colorametric signal was obtained. Three to twenty cuticles of *D. americana* and *D. novamexicana* were processed in parallel, and a titration series of probe and antibody conditions was examined. This experiment was repeated more than ten times.

To quantify standing levels of *tan* mRNA, stage P14 and P15 pupae from *D. americana* (A00), *D. novamexicana* (N14) and F_1 hybrids were collected and stored at -80. Prior to freezing, pupae were dissected from their cases, and their heads were removed to eliminate *tan* transcripts associated with the visual system (TRUE *et al.* 2005). Four samples, each containing three *D. americana* and three *D. novamexicana* pupal bodies, as well as five samples, each containing six F_1 hybrid bodies, were homogenized and used for sequential RNA and genomic DNA extractions, as described in (WITTKOPP *et al.* 2004). cDNA was synthesized from each RNA sample with a polyT primer and Superscript 2 reverse transcriptase (Invitrogen, Carlsbad, CA).

Pyrosequencing was performed as described in (WITTKOPP *et al.* 2008) with the following pair of primers: 5'- GATGCTGAAGTCCAGCGTGTC-3' and 5'-BIO-

CAGCCGCCAGTGACATCA-3', where "BIO" indicates the addition of a biotin molecule. A Pyrosequencing primer (5'-CGAGCACGATGTCCG-3') was used to analyze the sequence CAAYATG, in which the *D. americana* allele contains a thymine (T) and the *D. novamexicana* allele contains a cytosine (C) at the variable position. Pyrosequencing reactions were performed for each cDNA and genomic DNA sample, with a minimum average peak height of twelve for the conserved C, T, and G positions required for quality control.

The relative abundance of *D. americana* and *D. novamexicana tan* alleles was calculated as the ratio between the polymorphic T and C peaks (i.e., *D. americana*/*D. novamexicana* = T/C). These ratios were $\log_2$ transformed to make them normally distributed (WITTKOPP *et al.* 2008), and the $\log_2$ genomic DNA ratio was subtracted from the corresponding $\log_2$ cDNA ratio to correct for any bias between alleles in PCR- amplification and/or nucleic acid extraction in mixed species pools (WITTKOPP *et al.* 2004). We tested for expression differences between species and for allele-specific expression differences in F_1 hybrids using "t.test()" in R Cocoa GUI 1.12 (H_0: $\mu = 0$). The average percent difference in *tan* expression between species was calculated as $(1-2x) * 100$, where x = the mean value of $\log_2(T/C)$ for the four biological replicate samples.

Transgene construction and transformation

piggyBac transgenes containing the *D. americana* and *D. novamexicana tan* alleles were constructed by recombineering with the general protocols and strains described at http://recombineering.ncifcrf.gov/. Targeting plasmids for *D.*

americana and *D. novamexicana tan* were produced by amplifying ~500 bp "homology arms" from HMR and CG7039 with PCR primers, using PCR sewing to connect them with an Xho I site in between, and inserting them into the AscI site of a piggyBac plasmid (HORN and WIMMER 2000) with AscI sites flanking the homology arms that were introduced during the initial PCR amplification. The piggyBac plasmid, already containing a 3xP3-EGFP transformation marker (HORN and WIMMER 2000), was modified by inserting the attB sequence (GROTH *et al.* 2004) into the Xba I site located in the pUC18 backbone. Separate targeting vectors were made for *D. americana* and *D. novamexicana tan*, and the sequence of these constructs was confirmed by DNA sequencing.

Each 7 kb piggyBac targeting vector was linearized with Xho I, gel purified, treated with calf intestine alkaline phosphotase, and electroporated into SW102 cells carrying the appropriate BAC clone: DA_ABa0020L7 for *D. americana* and DN_Ba0024C15 for *D. novamexicana*. SW012 cells contain all of the genetic resources needed for recombineering (WARMING *et al.* 2005). Following induction of the recombinase and ampicillin selection for circularized piggyBac plasmids, individual colonies were screened by PCR, mini-prepped, and subject to diagnostic restriction digests. One positive clone from each species was re-sequenced for the entire 14kb *tan* transgene. In both cases, recombineering was found to have produced an exact replica of the BAC sequence.

piggyBac plasmids carrying the *D. americana* and *D. novamexicana tan* transgenes were injected into *D. melanogaster white* mutants carrying the attP16 site on chromosome two (MARKSTEIN *et al.* 2008), and inserted into the genome

with the phiC31 integrase (GROTH *et al.* 2004). Each piggyBac plasmid was also co-injected with a piggyBac transposase source (HORN and WIMMER 2000) into wild-type *D. novamexicana* (N14) and *white* mutant *D. americana* (provided by B. McAllister). Genetic Services, Inc (Sudbury, MA) performed all embryo injections and screening.

Sequence polymorphism discovery and analysis

DNA was extracted from males of each isofemale line listed in Table 2.2. *tan* was sequenced in twelve lines of *D. americana* and three lines of *D. novamexicana* with primers developed during the initial sequencing of *tan* alleles from BAC clones. Additional loci flanking *tan*, described in Figure 2.13A, were also sequenced in multiple lines of *D. americana* and *D. novamexicana*. Three regions of *ebony* were amplified and sequenced from ten lines of *D. americana* and five lines of *D. novamexicana* (see Figure 2.11D). These regions were concatenated for phylogenetic analysis. Primers used to collect these sequences are summarized in Table 2.3. In lieu of sequencing, AP-1 Gamma was genotyped with the primers 5'- TCGAACATGTTTGGCCTTGTCAC-3' and 5' TTTATAGCCAGCAGTTGCTG-3, which amplified a 158 bp product in *D. novamexicana*, and a 100bp product in *D. americana* (including the A01 allele). Sequencing reactions were performed by the University of Michigan Sequencing Core Facility and raw sequence data was analyzed with Sequencher 4.6 (Gene Codes Corporation, Ann Arbor, MI) and Codon Code Aligner 2.0.6 (CodonCode Corporation, Dedham, MA). Sequences for each allele were aligned with

MUSCLE (EDGAR 2004) and then manually inspected by P.J.W. Phylogentic trees were built with MEGA 4.0.2 (KUMAR *et al.* 2004). DNAsp 4.90.1 (ROZAS *et al.* 2003) and SITES (HEY and WAKELEY 1997) were used to identify fixed differences and polymorphic sites.

Table 2.1: Prior studies comparing intra- and interspecific QTL

	Intraspecific			Interspecific			
Trait	**Species**	**#QTLs**	**Ref.**	**Species**	**#QTLs**	**Ref.**	**Overlap[1]**
sex comb tooth number	*Drosophila melanogaster*	2	NUZHDIN and REIWITCH 2000	*D. simulans* and *D. mauritiana*	>= 5	COYNE 1985 [2]	1
male courtship song	*Drosophila melanogaster*	3	GLEASON *et al.* 2002	*D. simulans* and *D. sechellia*	6	GLEASON and RITCHIE 2004	0
abdominal pigmentation	*Drosophila melanogaster*	3	KOPP *et al.* 2003	*D. melanogaster* and *D. willistoni*	N/A[3]	WILLIAMS *et al.* 2008	1
various[4]	*Helianthus petiolaris*	28 [5]	LEXER *et al.* 2005	*H. annuus* and *H. petiolaris*	72	LEXER *et al.* 2005	28
sex comb tooth number	*Drosophila simulans*	7	TATSUTA and TAKANO-SHIMIZU 2006	*D. simulans* and *D. mauritiana*	2	TRUE *et al.* 1997	2
floral traits	*Mimulus guttatus*	16	HALL and BASTEN 2006	*M. guttatus* and *M. nasutus*	24	FISHMAN *et al.* 2002	11

[1] The number of QTLs that map to the same genomic region within and between species is shown.

[2] Interspecific QTLs were identified on each chromosome in this low-resolution study, providing a minimum estimate of 5 QTLs. True et al. (TRUE *et al.* 1997) refined this work, but only identified two QTLs, both on the third chromosome, and neither of which overlap the intraspecific QTL identified in (NUZHDIN and REIWITCH 2000). The inconsistency between studies may also reflect intraspecific variation within *D. simulans* and/or *D. mauritiana*.

[3] Interspecific divergence of *bric-a-brac* (located within one of the intra-specific QTL) was demonstrated using transgenic assays.

[4] This study includes 40 different morphological, physiological, and life history traits.

[5] Intraspecific QTL were not mapped genome wide; rather, 72 interspecific QTL were directly tested for the presence of intraspecific QTL.

Table 2.2: Lines of *D. americana* and *D. novamexicana* used in this work

Species	Line	Full-ID	Source	Collection Site	Collection Year
D. americana	A00	15010-0951.00	Drosophila Species Stock Center	Anderson, IN	unknown
D. americana	A01	15010-0951.01	Drosophila Species Stock Center	Poplar, MT	1947
D. americana	A04	15010-0951.04	Drosophila Species Stock Center	Lake Champlain, VT	1948
D. americana	A09	15010-0951.09	Drosophila Species Stock Center	Myrtle Beach, SC	1961
D. americana	AA	AA 99.6	B. McAllister[6]	Augusta, AR	1999
D. americana	DN12	DN 01.2	B. McAllister[7]	Duncan, NE	2001
D. americana	DN2	DN 00.2	B. McAllister[7]	Duncan, NE	2000
D. americana	DN4	DN 00.4	B. McAllister[7]	Duncan, NE	2000
D. americana	G	G96.21	B. McAllister[7]	Gary, IN	1996
D. americana	IR	IR 03.10	B. McAllister[8]	Iowa River, IA	2003
D. americana	LA	LA 99.48	B. McAllister[6]	Lake Ashbaugh, AR	1999
D. americana	NN97.2	NN 97.2	B. McAllister[7]	Niobrara, NE	1997
D. americana	NN97.8	NN97.8	B. McAllister[7]	Niobrara, NE	1997
D. americana	OR	OR 01.52	B. McAllister[9]	Ottawa, OH	2001
D. americana	PM	PM 99.28	B. McAllister[6]	Puxico, MO	1999
D. americana	SB	SB 02.06	B. McAllister[8]	Saulsbury, IA	2002
D. novamexicana	N04	15010-1031.04	Drosophila Species Stock Center	Moab, UT	1949
D. novamexicana	N08	15010-1031.08	Drosophila Species Stock Center	San Antonio, NM	1947
D. novamexicana	N12	15010-1031.12	Drosophila Species Stock Center	Antlers, CO	1949
D. novamexicana	N13	15010-1031.13	Drosophila Species Stock Center	Patagonia, AZ	1953
D. novamexicana	N14	15010-1031.14	Drosophila Species Stock Center	Moab, UT	1949

[6](McAllister 2002)

[7](McAllister 2003)

[8](McAllister and Evans 2006)

[9](McAllister 2001)

Table 2.3: Primers used for phylogenetic analysis

Gene name	Primer name	Primer sequence	Accession # containing resulting sequence
tan	39	CGAACCGCAACTGATATTGA	GQ457336-GQ457353
tan	91	TAGTGAGTGCCACGTGTATAGAGAACG	
tan	204	AAGCTAGGCAAACGGCATGC	
tan	205	ACAATTTCGAATTCGATGAGC	
tan	226	GGCGCTCTTCAATGAGCCAAACAA	
tan	94	TCAGTTTGAATTCTGCCTTCAAGCGCT	
tan	40	GTTGTTGTTGTTGGGGGTTC	
tan	90	CGTTCTCTATACACGTGGCACTCACTA	
tan	204	AAGCTAGGCAAACGGCATGC	
tan	205	ACAATTTCGAATTCGATGAGC	
tan	191	CCCTTACCCACTTTCTATGG	
tan	229	TCACGACTGATCGACAGGGCAAAC	
tan	227	CCCGCGCACATAATTAACAAGCTG	
tan	89	GCGACTTTGGCCTTAGCTTC	
tan	69	GCCGAGGTGGAGTTCCA	
tan	81	GGTATTCAATCTTCGGCGTGCCAAA	
tan	71	CACACCTTTCAAAAGAT	
tan	85	CGCTGACCAGAATTTCAAATTTAATTGCC	
tan	87	GACCATTTAATGGTGCTCAAAATATGG	
tan	206	GGAATGCCTTTTACTGCACATAATG	
tan	230	CCGGCATAGCAGAGCGACATGAA	
tan	86	CCATATTTTGAGCACCATTAAATGGTC	
tan	228	GCCTTGACGGAGACGCTTATTCAT	
tan	41	TACTTTGTGGTTGCGCACAT	
tan	42	ATTTGTAAAGCAGGGGCAAC	
tan	180	CGCATACACTTGGACCAGGCC	
tan	231	TTGGCTTCGTTCTAACGGGCATCA	
tan	246	CATTAAGCTTTGTAGCTGACTATG	
tan	193	CACACACGCGAATTAGGCAAAAG	
ld14	215	GCGTGCAGTGTCTGTTAGCAG	GQ457425-GQ457433
ld14	216	AGGCCACGCCCACTAACTAAC	
l(1)G0007	217	GCGCGTCTCCGATGAGATGG	GQ457434-GQ457447
l(1)G0007	218	AATACCATCGGTCATGTACTTGATGAC	
CG32687	234	CCAGCGAAAGCACATGCAG	GQ457401-GQ457408
CG32687	235	GGATATCTGGAAGATGCAAAG	
baz	61	CGGTGGCATACATCAGAATG	GQ457409-GQ457424
baz	62	CGACTGCAAAAGCGTACAAA	
CG7039-AP1gamma	238	CAAATGGCTGGTGGAGGC	GQ457388-GQ457400
CG7039-AP1gamma	239	CCTGCTCATAATACGGATAC	

ebony	45	AATTACCCAACTGCGACTGG	GQ457354-GQ457369
ebony	46	CGCCCTCCATCTTCAGATAC	
ebony	130	CGCTCCCTGCTCATGTATCT	
ebony	131	GGCGACGTTCTTCTCAACCT	
ebony	e4f	CGTGTGCGCCCAATTAGC	
ebony	e4r	TGCTTAGATTTCACCTCATCAACAGAA	
HMR[1]	142	CATCTCGTCCAACTCCAGGT	GQ457448-GQ457453
HMR[1]	144	GCGCTATAAATATCAGCGTCA	
CG7039[1]	145	CATTGCTGCACGGCTTTTAC	GQ457382-GQ457387
CG7039[1]	146	CTCCACCAGCCATTTGATG	
AP1gamma[1]	213	GGAAATGTGCGAGAACAGTTCGG	GQ457370-GQ457381
AP1gamma[1]	214	GATTTTCACCTGCAGGAACGGATC	

References:

AHMADIAN A., 2000 Single-Nucleotide Polymorphism Analysis by Pyrosequencing. Analytical Biochemistry **280**: 103–110.

BRIDGES C.B., MORGAN T.H. 1923 The third-chromosome group of mutant characters of *Drosophila melanogaster*. Publ. Carnegie Inst. **327**: 251.

CALETKA B. C., MCALLISTER B. F., 2004 A genealogical view of chromosomal evolution and species delimitation in the Drosophila virilis species subgroup. Mol Phylogenet Evol **33**: 664–670.

CONG B., LIU J., TANKSLEY S. D., 2002 Natural alleles at a tomato fruit size quantitative trait locus differ by heterochronic regulatory mutations. PNAS **99**: 13606–13611.

COYNE J. A., 1985 Genetic studies of three sibling species of Drosophila with relationship to theories of speciation. Genet. Res. **46**: 169–192.

DEGNAN J. H., ROSENBERG N. A., 2009 Gene tree discordance, phylogentic inference and the multispecies coalescent. Trends Ecol Evol **24**: 332–340.

DROSOPHILA 12 GENOMES CONSORTIUM, *ET AL*., 2007 Evolution of genes and genomes on the Drosophila phylogeny. Nature **450**: 203–218.

EDGAR R. C., 2004 MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Research **32**: 1792–1797.

FALCONER D.S., MACKAY, T.F.C., 1996 Introduction to Quantitative Genetics. Longman Group Ltd., Essex, UK, ed 4th.

FISHMAN L., KELLY A. J., WILLIS J. H., 2002 Minor quantitative trait loci underlie floral traits associated with mating system divergence in Mimulus. Evolution **56**: 2138–2155.

GLEASON J. M., NUZHDIN S. V., RITCHIE M. G., 2002 Quantitative trait loci affecting a courtship signal in Drosophila melanogaster. Heredity **89**: 1–6.

GLEASON J. M., RITCHIE M. G., 2004 Do quantitative trait loci (QTL) for a courtship song difference between Drosophila simulans and D. sechellia coincide with candidate genes and intraspecific QTL? Genetics **166**: 1303–1311.

GLOOR G. B., PRESTON C. R., JOHNSON-SCHLITZ D., NANASSIF, PHILLIS R., BENZ W., ROBERTSON H., ENGELS W., 1993 Type 1 Repressors of P Element Mobility. Genetics **135**: 80–95.

GROTH A. C., FISH M., NUSSE R., CALOS M. P., 2004 Construction of transgenic Drosophila by using the site-specific integrase from phage phiC31. Genetics **166**: 1775–1782.

HALL M., BASTEN C., 2006 Pleiotropic Quantitative Trait Loci Contribute to Population Divergence in Traits Associated With Life-History Variation in Mimulus guttatus. Genetics **172**: 1829–1844.

HEY J., WAKELEY J., 1997 A coalescent estimator of the population recombination rate. Genetics **145**: 833–846.

HOLDER M. T., ANDERSON J. A., HOLLOWAY A. K., 2001 Difficulties in Detecting Hybridization. Syst Biol **50**: 978–982.

HORN C., WIMMER E. A., 2000 A versatile vector set for animal transgenesis. Dev Genes Evol **210**: 630–637.

HSU, T.C., 1951 Chromosomal variation and evolution in the *virilis* group of *Drosophila.* Thesis, University of Texas, Austin, TX.

JEONG S., REBEIZ M., ANDOLFATTO P., WERNER T., TRUE J., CARROLL S. B., 2008 The evolution of gene regulation underlies a morphological difference between two Drosophila sister species. Cell **132**: 783–793.

KOPP A., GRAZE R. M., XU S., CARROLL S. B., NUZHDIN S. V., 2003 Quantitative trait loci responsible for variation in sexually dimorphic traits in Drosophila melanogaster. Genetics **163**: 771–787.

KUMAR S., TAMURA K., NEI M., 2004 MEGA3: Integrated software for Molecular Evolutionary Genetics Analysis and sequence alignment. Brief. Bioinformatics **5**: 150–163.

LAURIE C. C., STAM L. F., 1994 The Effect of an Intronic Polymorphism on Alcohol Dehydrogenase Expression in Drosophila melanogaster. Genetics **138**: 378–385.

LEXER C., ROSENTHAL D. M., RAYMOND O., DONOVAN L. A., RIESEBERG L. H., 2005 Genetics of species differences in the wild annual sunflowers, Helianthus annuus and H. petiolaris. Genetics **169**: 2225–2239.

MARKSTEIN M., PITSOULI C., VILLALTA C., CELNIKER S., 2008 Exploiting position effects and the gypsy retrovirus insulator to engineer precisely expressed transgenes. Nat Genet **40**: 476–483.

MASIDE X., WEISHAN LEE A., CHARLESWORTH B., 2004 Selection on Codon Usage in Drosophila americana. Current Biology **14**: 150–154.

MCALLISTER B. F., 2001 Genetic Analysis of Sex-Chromosome Arrangement in Drosophila americana: a laboratory exercise for undergraduate or advanced placement students. Dros. Inf. Serv. **84**: 227–234.

MCALLISTER B. F., 2002 Chromosomal and allelic variation in Drosophila americana: selective maintenance of a chromosomal cline. Genome **45**: 13–21.

MCALLISTER B. F., 2003 Sequence Differentiation Associated With an Inversion on the Neo-X Chromosome of Drosophila americana. Genetics **165**: 1317–1328.

MCALLISTER B. F., EVANS A. L., 2006 Increased Nucleotide Diversity with Transient Y Linkage in Drosophila americana. Plos One **1**: e112.

MCEWEN R., 1918 The Reactions to Light and to Gravity in Drosophila and its Mutants. Journal of Experimental Zoology **25**: 49-106.

MENA P., MCALLISTER B.F. Personal Communication.

MORALES-HOJAS R., VIERIRA C. P., VIEIRA J., 2008 Inferring the evolutionary history of Drosophila americana and Drosophila novamexicana using a multilocus approach and the influence of chromosomal rearrangements in single gene analyses. Molecular Ecology **17**: 2910–2926.

NOOR M. A. F., FEDER J. L., 2006 Speciation genetics: evolving approaches. Nat Rev Genet **7**: 851–861.

NUZHDIN S. V., REIWITCH S. G., 2000 Are the same genes responsible for intra- and interspecific variability for sex comb tooth number in Drosophila? Heredity **84 (Pt 1)**: 97–102.

ROZAS J., SÁNCHEZ-DELBARRIO J. C., MESSEGUER X., ROZAS R., 2003 DnaSP, DNA polymorphism analyses by the coalescent and other methods. Bioinformatics **19**: 2496–2497.

SCHAFER M., ORSINI L., MCALLISTER B., SCHOTTERER C., 2006 Patterns of microsatellite variation through a transition zone of a chromosomal cline in Drosophila americana. Heredity **97**: 291–295.

STARK A., KHERADPOUR P., PARTS L., BRENNECK J., HODGES E., HANNON G. J., KELLIS M., 2007 Systematic discovery and characterization of fly microRNAs using 12 Drosophila genomes. Genome Research **17**: 1865–1879.

TATSUTA H., TAKANO-SHIMIZU T., 2006 Genetic architecture of variation in sex-comb tooth number in Drosophila simulans. Genet. Res. **87**: 93–107.

THROCKMORTON L.H., 1982 in The Genetics and Biology of Drosophila. Ashburner M., Carson H.L., Thompson J.N., Eds (Academic Press, London), vol 3b, 227-296.

TRUE J., LIU J., STAM L., ZENG Z., 1997 Quantitative Genetic Analysis of Divergence in Male Secondary Sexual Traits Between Drosophila Simulans and Drosophila Mauritiana. Evolution **51**: 816–832.

TRUE J. R., YEH S.-D., HOVEMANN B. T., KEMME T., MEINERTZHAGEN I. A., EDWARDS T. N., LIOU S.-R., HAN Q., LI J., 2005 Drosophila tan encodes a novel hydrolase required in pigmentation and vision. PLoS Genetics **1**: e63.

TRUE J., LIU J., STAM L., ZENG Z., 1997 Quantitative Genetic Analysis of Divergence in Male Secondary Sexual Traits Between Drosophila Simulans and Drosophila Mauritiana. Evolution **51**: 816–832.

VIEIRA C. P., COELHO P. A., VIEIRA J., 2003 Inferences on the Evolutionary History of the Drosophila americana Polymorphic X/4 Fusion From Patterns of Polymorphism at the X-Linked paralytic and elav Genes. Genetics **164**: 1459–1469.

WARMING S., COSTANTINO N., COURT D. L., JENKINS N. A., COPELAND N. G., 2005 Simple and highly efficient BAC recombineering using galK selection. Nucleic Acids Research **33**: e36.

WILLIAMS T., SELEGUE J., WERNER T., GOMPEL N., 2008 The Regulation and Evolution of a Genetic Switch Controlling Sexually Dimorphic Traits in Drosophila. Cell **134**: 610–623.

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**: 65–71.

WITTKOPP P. J., CARROLL S. B., KOPP A., 2003a Evolution in black and white: genetic control of pigment patterns in Drosophila. Trends Genet **19**: 495–504.

WITTKOPP P. J., HAERUM B. K., CLARK A. G., 2004 Evolutionary changes in cis and trans gene regulation. Nature **430**: 85–88.

WITTKOPP P. J., HAERUM B. K., CLARK A. G., 2008 Regulatory changes underlying expression differences within and between Drosophila species. Nat Genet **40**: 346–350.

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**: 1849–1858.

WITTKOPP P. J., WILLIAMS B. L., SELEGUE J. E., CARROLL S. B., 2003b Drosophila pigmentation evolution: divergent genotypes underlying convergent phenotypes. PNAS **100**: 1808–1813.

Chapter 3

Identifying Causative Regions in *tan* for Pigmentation between *D.americana* and *D. novamexicana*[2]

Abstract

In order to better understand how phenotypic evolution can occur, the underlying genes must first be identified, as well as the individual causative nucleotides. Studies of this resolution provide valuable insight into the mechanisms of how new traits are formed. Despite its importance in evolutionary biology, few studies have achieved such level of detail. A case study is presented here comparing two *Drosophila* species, *D. americana* and *D. novamexicana,* that have distinct pigmentations. A previously identified region of the pigmentation gene, *tan*, contributing to pigmentation divergence between *D. americana* and *D. novamexicana* is further dissected to identify specific regions associated with the pigmentation difference between the two species using functional transgenic analyses. These data suggest that the 3' half of the first intron of *tan* is likely to contain some of the divergent sequence(s) contributing to the pigmentation difference between *D. americana* and *D. novamexicana.* Future analyses will further narrow this region to single nucleotides or a group of

[2] The research presented here was done in collaboration with postdoctoral fellow, Arielle Cooley and assistance from undergraduate researchers Bradley Lankowsky and Wesley McLaughlin with fly rearing, dissections and imaging.

nucleotides necessary for the pigmentation difference between these species, allowing the details of the genetic and molecular basis of this phenotypic change between species to be deciphered.

Introduction

Understanding the mechanisms of phenotypic evolution has been a longstanding challenge for evolutionary biology. By elucidating the genetic and molecular processes underlying morphological traits, we can begin to gain a greater insight into how evolution can shape phenotypes. Identifying specific nucleotides responsible for changes in gene or protein function can allow us to discern how developmental and cellular processes are altered to result in new phenotypes.

Despite being such an important evolutionary topic, few studies have been able to provide a detailed analysis that not only determine the gene(s) responsible for phenotypic change, but also the specific nucleotide(s) and the mechanisms by which those genetic changes lead to phenotypic divergence (COLOSIMO *et al.* 2005; HOEKSTRA *et al.* 2006; TISHKOFF *et al.* 2007; MCGREGOR *et al.* 2007; JEONG *et al.* 2008). Many more investigations that can supply this level of resolution are needed to truly understand how genetic and developmental processes can lead to evolutionary change.

Here, we use two *Drosophila* species with divergent pigmentation as a model to investigate how genetic differences can contribute to phenotypic change. *D. americana* and *D. novamexicana* are sister species within the *virilis*

group that diverged about 300,000-500,000 years ago (CALETKA and MCALLISTER 2004; MORALES-HOJAS *et al.* 2008). *D. americana* exhibits a very dark melanic pigmentation, whereas *D. novamexicana* has a light yellowish pigmentation (Figures 1.10 and 2.1).

Previously, it was found that the pigmentation genes *ebony* and *tan* are strongly associated with the pigmentation difference in these species. *ebony* and *tan* act reciprocally of each other in the biosynthesis of light and dark pigment in *Drosophila* (WRIGHT 1987; TRUE 2003; WITTKOPP *et al.* 2003a; WITTKOPP *et al.* 2009). *ebony*, which promotes the formation of light pigment, has been found to have higher protein expression in the lightly pigmented *D. novamexicana* (WITTKOPP *et al.* 2003b). In comparison, *tan,* which promotes dark pigment production, has been found to have higher mRNA expression in the darker *D. americana* (WITTKOPP *et al.* 2009; COOLEY *et al.* 2012).

Additionally, a region in *tan* has previously been shown to contribute to the functional divergence between these alleles (WITTKOPP *et al.* 2009). This sequence of *tan* is comprised of primarily the first intron, the small exon 1 (~90 bp) and a small portion of the 5' UTR. In total, the region is about 2.7 kb and contains primarily non-coding sequence in intron 1 (WITTKOPP *et al.* 2009). The fact that this region contains mostly non-coding sequence combined with the differential mRNA expression in *D. americana* suggests a regulatory change in *tan* between species. Surprisingly, expression differences for *tan* have been shown to be due largely to *trans* effects with small *cis*-acting effects (COOLEY *et al.* 2012; WITTKOPP *et al.* 2009), although the exact details have yet to be worked

out. By elucidating the specific sequences of *tan* that are important for the pigmentation difference between *D. americana* and *D. novamexicana,* the mechanisms by which *tan* functions or is regulated can be better understood.

This presents an excellent system to identify and dissect how genetic differences in *tan* between species leads to such dramatic phenotypes. The following analysis ultimately aims to identify specific nucleotides or regions that are important for *tan*'s impact on the pigmentation difference between *D. americana* and *D. novamexicana* through transgenic analyses. As a step towards this goal, the data presented here suggests that some of the causative sequence is likely located in the 3' half of the first intron of *tan.*

Materials and Methods

Transgene construction and transformation

D. americana and *D. novamexicana tan* sequence was modified in a bacterial artificial chromosome (BAC) for each species. *tan* sequence was modified using a two-step recombineering protocol as described in (WARMING *et al.* 2005), which recombines a selectable marker (*galK*) into a specified location, and is then replaced with the desired sequence. Homology arms are designed to have ~300-800 bp overhang for recombination. Chimeric *tan* intron 1 constructs were created using this method.

Prior to creating these chimeric intron constructs, one amino acid was changed in each species' *tan* allele. These changes replaced a rare polymorphism with the most common allele found in each species. The effects of

these rare polymorphisms were examined by over-expressing Tan proteins containing these alleles using the GAL4-UAS system and were found to have no visible effect on pigmentation intensity (Figure 3.1). All chimeric constructs were created in a genetic background with the rare polymorphisms replaced with the most common allele found in both species. The pure *D. americana tan* allele is not included in the results because it has not yet been successfully transformed into flies.

Two chimeric *tan* constructs were created wherein, the 5' half of intron 1 is replaced with sequence from the opposing species (*Da tan + Dn* 5' half intron 1 and *Dn tan + Da* 5' half intron 1-See Figure 3.2) using the recombineering protocol described above.

Figure 3.1: Rare *tan* polymorphisms do not impact Tan function

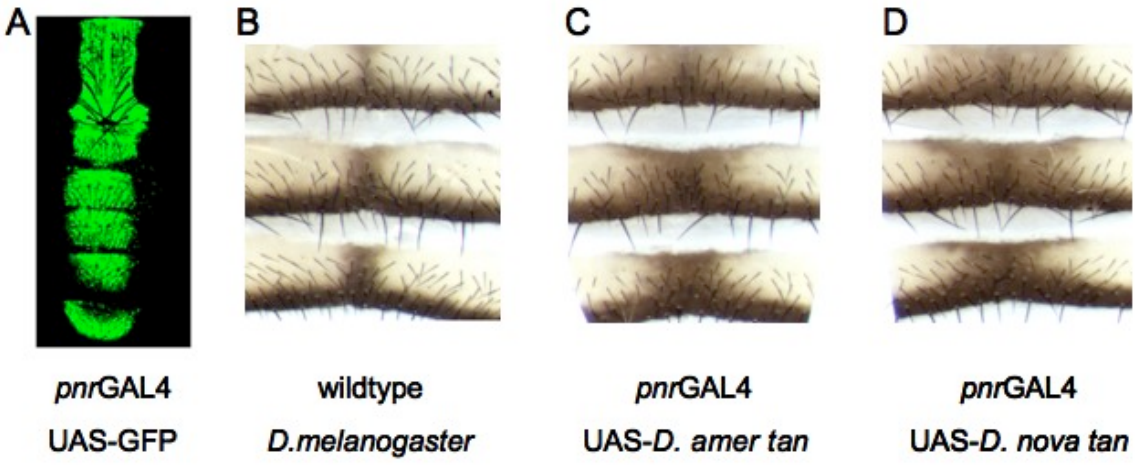


Figure 3.1: Over-expression of *D. americana* and *D. novamexicana tan* alleles containing rare polymorphism using the GAL4-UAS system. Dorsal abdominal cuticles (abdominal segments 3-5) are shown from 7-10 day old female flies. The *pannier*-GAL4 (*pnr*GAL4) is used to drive expression through the dorsal midline of the fly (A). Both *D. americana (D. amer)* (C) and *D.*

novamexicana (D. nova) (D) *tan* alleles are over-expressed compared to the wildtype (B), but do not show any visual difference in pigmentation intensity between species' alleles.

Figure 3.2 : Constructs used for transgenics

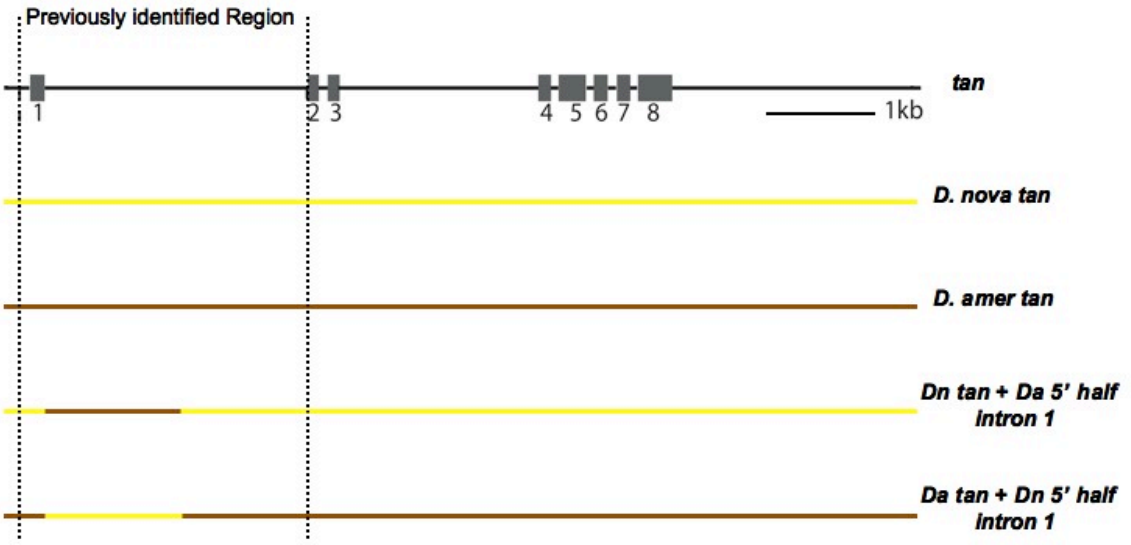


Figure 3.2: Constructs used for transgenic analysis. A schematic representing *tan* is shown at the top with all eight exons (gray boxes) and some flanking sequence. The previously identified region is denoted by vertical dashed lines. Yellow lines represent *D. novamexicana* sequence, and brown lines represent *D. americana* sequence. No successful transformants have been obtained for the pure *D. americana tan* allele; therefore, it is not included in subsequent analyses. Abbreviations: *Dn: D. novamexicana, Da: D. americana*. All schematics drawn approximately to scale.

Chimeric oligonucleotides used for recombination (5' arm + target sequence + 3' arm) were constructed using a PCR fusion protocol (LAB 2006). Each homology arm and target sequence were amplified individually using the following standard PCR conditions, 94°C for 30s, 56°C for 30s and 72°C for 1min/1kb sequence amplified for 30 cycles. The three individual components (5' arm + target sequence + 3' arm) were then used as templates (equal amounts of

each) to construct the complete construct via PCR fusion as described in (LAB 2006). Primers used for each PCR are listed in Table 3.1.

Constructed chimeric oligonucleotides from PCR fusion were used for recombineering as described in (WARMING *et al.* 2005) and subsequent colonies were screened for positive transformants accordingly using PCR with the same primers described for chimeric oligonucleotide synthesis. Positive clones were then mini-prepped and subject to diagnostic restriction digests and sequenced to ensure the intended sequence changes had been successful. Modified *tan* sequences were then moved into an appropriate injection vector (piggyBac) via a second recombineering protocol.

The same piggyBac vector and recombineering protocol was used as described in (WITTKOPP *et al.* 2009).This targeting vector contains a species-specific (*D. americana* or *D. novamexicana*) pair of homology arms, 3xP3-EGFP transformation marker and attB sequence. The piggyBac vector was linearized by using the introduced XhoI restriction site located between the two homology arms and electroporated into SW102 cells (as described in WITTKOPP *et al.* 2009) carrying the appropriate modified BAC clone for each construct. Electroporation was performed using Eppendorf Electroporator 2510 at 1250 Volts. Following electroporation, SW102 cells were incubated in 450ul LB at 30^{o}C rotating for 1-2 hours, then diluted (100%, 10% and 1%) and spread on LB agar plates supplemented with ampicillin (50 ug/ml) and grown overnight at 30^{o}C to select for circularized piggyBac plasmid containing cells. Individual colonies were screened by PCR, mini-prepped and subject to diagnostic restriction digests. Positive

clones were also sequenced across key species-specific sequence transitions (i.e. where *D. novamexicana* sequence was introduced into an otherwise *D. americana tan* gene).

The piggyBac plasmids containing the *D. americana* and *D. novamexicana* modified *tan* transgenes were injected into *D. melanogaster* white mutants carrying the attP sequence located on chromosome 2 (inserted using phiC31 integrase (GROTH *et al.* 2004)), and screened for GFP activity. All injections were performed by Genetic Services, Inc (Sudbury, MA).

Fly strains, crossing schemes and imaging

Transformant *D. melanogaster* flies for each construct were crossed into *yellow; tan; white* mutant background using the balancer stock *w-; I(2)/CyO; D¹/ TM6b, Tb+*. Transgenes must be maintained in a *white* mutant background in order to visualize GFP expression in the eye (as the transformation marker). A *yellow; tan* genetic background was chosen to eliminate all endogenous dark pigment in fly abdomens. This ensures observed rescued pigmentation is a result from the respective transgene. All fly strains were maintained on a standard yeast-glucose media at room temperature. Individual flies were prepared for pigmentation analysis by placing 7-10 day old flies in 10% glycerol in ethanol and storing for a minimum of 24 hours and subsequently dissected to isolate the dorsal abdominal cuticle and mounted in Hoyer's solution and imaged with Scion 1394 (Frederick, MD) digital camera. Images were taken under the same lighting conditions on the same day and adjusted uniformly in Photoshop CS6 (Adobe,

San Jose, CA) to best represent visual observations. Pigmentation was quantified from male flies from the A5 abdominal tergite by measuring the intensity of gray-scale images with ImageJ (NIH, Bethesda, MD). The mean pigmentation was calculated from six to eight individual flies for each line. Measurements ranged from 0 (black) to 255 (white).

Results

D. americana *and* D. novamexicana *transgenes are functional in* D. melanogaster.

D. americana and *D. novamexicana* diverged from *D. melanogaster* 40-60 million years ago (SLAWSON *et al.* 2006). Despite this divergence time, all transgenics flies that contain a *D. americana* or *D. novamexicana tan* allele restore some dark pigmentation in *D. melanogaster yellow, tan* mutants causing them to resemble *yellow* mutants (Figure 3.2). This indicates that the required *trans* elements for *D. americana* and *D. novamexicana tan* function are present in *D. melanogaster.*

Causal variant(s) for D. americana *and* D. novamexicana *pigmentation difference likely located in 3' half of* tan *intron 1*

Although all *D. americana* and *D. novamexicana tan* alleles rescued pigmentation in *D. melanogaster yellow, tan* mutants, there were varying intensities of darkness between constructs (Figure 3.3 and 3.4). The *D. novamexicana tan* allele displayed a light pigmentation (two independent lines

averaged together, n=8 each line, mean=102.4 on gray scale), but was still darker than the untransformed control *yellow, tan* mutant (n=8, mean=174.4 on gray scale) (two-sample t(11.5)=9.1, p=1.5 x 10^{-5}). The *D. novamexicana tan* allele containing *D. americana* sequence in the 5' half of intron 1 showed a slightly lighter phenotype (n=8, mean=107.4 on gray scale). This construct is quite similar to the unmodified *D. novamexicana tan* allele, but only slightly lighter (two-sample t(13.9)=-0.33, p=0.75). The *D. americana tan* allele carrying the 5' half of intron 1 sequence from *D. novamexicana tan* was found to have a darker phenotype (n=6, mean=74.1 on gray scale) that differed significantly than *D. novamexicana tan* with the *D. americana* 5' half intron 1 (two-sample t(9.3)=5.0, p=6.7x10^{-4}). We would expect pigmentation to darken with the addition of *D. americana tan* sequence in an otherwise *D. novamexicana tan* allele (or alternatively *D. novamexicana tan* sequence to lighten pigmentation when added to a *D. americana tan* allele) if that sequence contained causative site(s). Given that this is not observed, these data suggest that causative nucleotides for the pigmentation differences between *D. americana* and *D. novamexicana* are not in the 5' half of intron 1. Given that intron 1 has previously been shown to affect pigmentation (WITTKOPP *et al.* 2009), these data suggest that the 3' half of the first intron includes some of the functionally divergent sites.

Figure 3.3: *tan* transgenes darken pigmentation on fly cuticles

Figure 3.3: Representative dorsal abdominal cuticles are shown (abdominal tergites 4 and 5-A4 and A5) of *D. melanogaster* male flies for each genotype. All transgenes are functional in *D. melanogaster.* A *yellow, white, tan (ywt*) mutant is the untransformed genetic background with light pigmentation. *D. novamexicana tan* allele (*D. nova tan)* and *D. novamexicana tan + D. americana* 5' half intron 1 (*Dn tan + Da* 5' half intron 1) display a lighter pigmentation as compared with *D. americana tan + D. novamexicana* 5' half intron 1 (*Da tan + Dn* 5' half intron 1).

Figure 3.4 : Pigmentation quantification suggests causal variant(s) for pigmentation difference contained in 3' half of *tan* intron 1

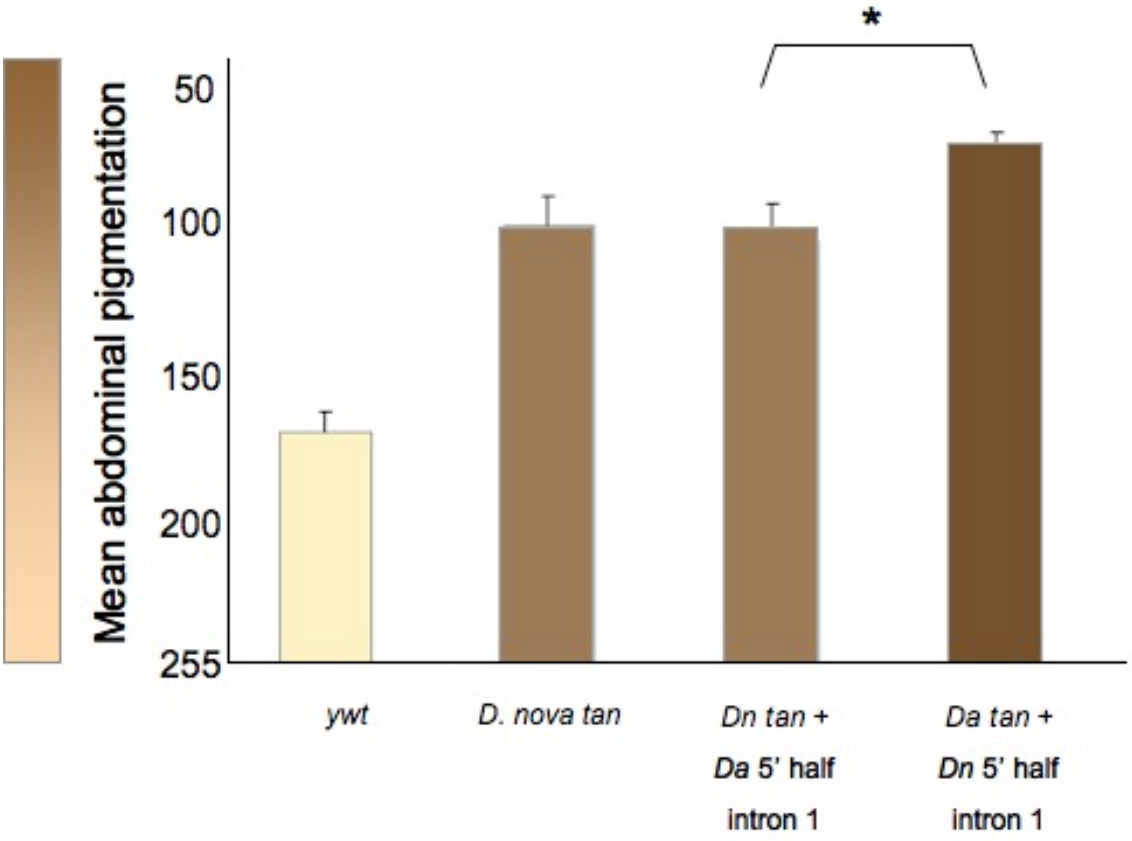


Figure 3.4: Quantification of dorsal abdominal cuticles (tergite 5) of transgenic flies. Means were calculated from 8 flies in each genotypic class (*D. nova tan* is calculated from two independent lines averaged together). The y axis is a gray scale, 0 (black) to 255 (white). Error bars represent 95% confidence intervals. *D. nova tan* was significantly darker than the untransformed control (*ywt)* (two-sample $t(11.5)=9.1$, $p=1.5 \times 10^{-5}$). *Dn tan* + *Da* 5' half intron 1 did not differ significantly from the pure *D. nova tan* (two-sample $t(13.9)=-0.33$, $p=0.75$). The *Da tan* + *Dn* 5' half intron 1 was significantly darker than the *Dn tan* + *Da* 5' half intron 1(*)(two-sample $t(9.3)=5.0$, $p=6.7 \times 10^{-4}$). Abbreviations: *ywt: yellow, white, tan* mutant (untransformed control), *D. nova tan: D. novamexicana tan* allele, *Dn: D. novamexicana, Da: D. americana*.

Figure 3.5: *D. americana* and *D. novamexicana* 3' half *tan* intron 1 alignment

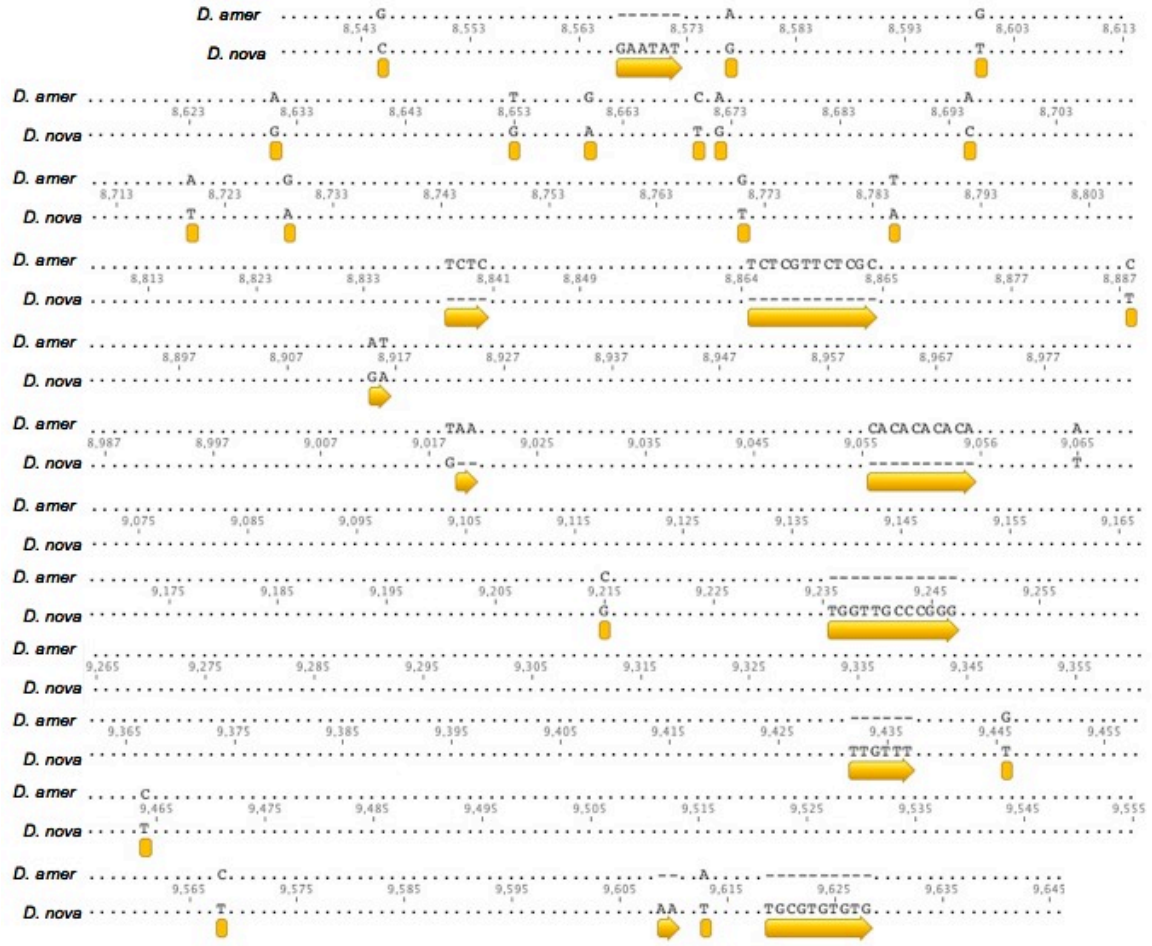


Figure 3.5: Sequence alignment of *D. americana* and *D. novamexicana* ~1.6 kb sequence from 3' half of *tan* intron 1. Yellow rectangles indicate a SNP and yellow arrows/arrowheads indicate an insertion/deletion (indel). Dots indicate conserved sequence between species.

Discussion

The *tan* gene is a major contributor to pigmentation differences between *D. americana* and *D. novamexicana.* Here, we present evidence that the 3' half of *tan* intron 1 likely harbors some of the sequences important in this

pigmentation difference. This is one possible way differences in *tan* could lead to this pigmentation divergence.

Small effects are displayed among chimeric transgenics

The effects on pigmentation displayed among the chimeric *tan* transgenes are small overall in that they do not produce the dramatic pigmentation difference exhibited in the wildtype *D. americana* and *D. novamexicana* flies. This is not surprising as it has already been established that the region of *tan* previously identified (containing a small part of the 5‘ UTR, exon 1 and intron 1 of *tan*) does not fully recapitulate the dark phenotype of *D. americana* (WITTKOPP *et al.* 2009) indicating that there are other regions of *tan* (or other genes) necessary for the complete phenotypic difference observed between *D. americana* and *D. novamexicana.*

tan also includes a large third intron. Sequences from both intron 1 and intron 3 may be necessary for the functional divergence of *tan* in *D. americana* and *D. novamexicana.* In this way, neither sequences from intron 1 nor intron 3 alone may be able to fully recapitulate the pigmentation differences. As these transgenic flies only carried divergent sites within *tan* intron 1, the small differences observed in pigmentation among the constructs are not surprising.

Additionally, endogenous *D. melanogaster* pigmentation genes (such as *ebony*) are expressed differently than in *D. americana* and *D. novamexicana* (WITTKOPP *et al.* 2002; 2003b) and therefore could impact the function of the transgenic *tan* alleles. Since *D. americana* and *D. novamexicana* are distantly

related to *D. melanogaster* (SLAWSON *et al.* 2006), it is possible that all of the *trans*-factors necessary for regulation of *D. americana* and *D. novamexicana tan* are either not expressed or not present in *D. melanogaster.* The differences in the *trans*-environments between *D. melanogaster* and *D. americana* and *D. novamexicana* could be a contributor to the slight differences in pigmentation observed among the *tan* transgenic flies.

Potential mechanisms for differential tan *mRNA expression*

There are several possibilities that could be responsible for the differential *tan* mRNA expression between *D. americana* and *D. novamexicana.* The 3' half of *tan* intron 1 identified here is not the only conceivable mechanism to explain the previously described dissimilarities in *tan* expression between species. As discussed above, multiple divergent sequences may be required to cause the observed differences in *tan* between species*,* such as sequences found in the first and third introns of *tan,* or sequences found elsewhere in the *tan* gene. Additionally, another possible explanation could include potential differences in *tan* promoter activity. It is plausible that the differences in *tan* activity between *D. americana* and *D. novamexicana* are a result from differences in their respective promoter activities. The *D. americana tan* promoter could be more highly active leading to the increase in *tan* mRNA expression, although there is currently no evidence to support this hypothesis.

The data presented here suggests that at least some of the sequences important for the pigmentation difference between *D. americana* and *D.*

novamexicana are located in the 3' half of *tan* intron 1. The non-coding polymorphisms contained within the 3' half of *tan* intron 1 could have several potential impacts on *tan.* There are 19 SNPs and 10 indels in the 3' half of *tan* intron 1 (Figure 3.5), which is approximately 1.2 kb in length. These differences could impact post-transcriptional regulation of *tan* alleles or temporal control of *tan* expression. It has previously been reported that the small *tan cis*-regulatory effect between *D. americana* and *D. novamexicana* is limited to a narrow developmental timeframe (COOLEY *et al.* 2012). It is possible that the polymorphisms within this 3' region of intron 1 contribute to an allele-specific *cis*-regulatory mechanism. Additionally, this sequence could encode for a small noncoding RNA that could potentially regulate *tan* mRNA expression in *trans*. However, sequence prediction programs failed to identify any predicted microRNA (data not shown).

This 3' region of *tan* intron 1 may also contain an enhancer that regulates *tan* expression. A *tan* enhancer has been putatively identified in the large third intron of *tan* from studies in *D. melanogaster* (TRUE J. R. personal communication). Recent studies have shown however that enhancers need not be located in conserved genomic regions between species (KALAY and WITTKOPP 2010). So it is plausible that a *tan* enhancer is located in the 3' half of intron 1 in *D. americana* and *D. novamexicana* and the polymorphisms therein for example could impact transcription factor binding sites for either transcriptional activators or repressors.

Future directions

Based on solely the data presented here, no definitive conclusions can be drawn concerning the mechanism(s) by which *tan* leads to the pigmentation difference between *D. americana* and *D. novamexicana.* The 3' half of *tan* intron 1 needs to be tested for its sufficiency to lead to pigmentation differences.

Ultimately, this research aims to identify specific causative nucleotides in *tan* that are responsible for pigmentation differences between *D. americana* and *D. novamexicana.* This level of resolution can supply valuable insights into the molecular mechanisms of phenotypic evolution. The data presented here provides an important step towards achieving this goal. Further transgenic analyses will continue to examine the effects from this 3' region of *tan* intron 1 and identify single or groups of nucleotides important for the pigmentation difference between *D. americana* and *D. novamexicana*.

Table 3.1: Primers used to modify BACs

Primer Name	Gene	Primer ID	Full Primer Sequence (5'-3') (with tail)	Tail Gene	Tail Sequence	Construct
galK Marker-F	*galK*	932	*CACCGAATCCGG* CCTGTTGACAATTAATCATCGGCA	*D. nova tan*	CACCGAATCCGG	
galK Marker-R	*galK*	933	*TCGTTCTTCCGG* TCAGCACTGTCCTGCTCCTT	*D. nova tan*	TCGTTCTTCCGG	
Dn aa Left Arm F	*D. nova tan*	926	TGCGGCATTTTTGTAAGTGA	-	-	
Dn aa Left Arm Step 1-R	*D. nova tan*	927	*ATTGTCAACAGG* CCGGATTCGGTGCCATTT	*galK*	ATTGTCAACAGG	*D.nova* aa correction
Dn aa Right Arm Step 1-F	*D. nova tan*	928	*GGACAGTGCTG* **A** CCGGAAGAACGAGTCACACT	*galK*	GGACAGTGCTG	
Dn aa Right Arm-R	*D. nova tan*	929	TTGGCCTGGTCCAAGTGTAT	-	-	
Dn aa Left Arm Step 2-R	*D. nova tan*	930	*CGTTCTTCCGG* **T** CCGGATTCGGTGCCATTT	*D. nova tan*	CGTTCTTCCGG	
Dn aa Right Arm Step 2-F	*D. nova tan*	931	*ACCGAATCCGG* **A** CCGGAAGAACGAGTCACACT	*D. nova tan*	ACCGAATCCGG	
galK Marker-F	*galK*	924	*GAGCCATAATCA* CCTGTTGACAATTAATCATCGGCA	*D. amer tan*	GAGCCATAATCA	
galK Marker-R	*galK*	925	*CACTAGCCCGTG* TCAGCACTGTCCTGCTCCTT	*D. amer tan*	CACTAGCCCGTG	
Da aa Left Arm F	*D. amer tan*	918	TTGACGGAGACGCTTAATCA	-	-	
Da aa Left Arm Step 1-R	*D. amer tan*	919	*ATTGTCAACAGG* TGATTATGGCTCATGGTGTAGC	*galK*	ATTGTCAACAGG	*D. amer* aa correction
Da aa Right Arm Step 1-F	*D. amer tan*	920	*GGACAGTGCTG* ACACGGGCTAGTGTTTAGCATC	*galK*	GGACAGTGCTG	
Da aa Right Arm-R	*D. amer tan*	921	GTAAGGCCCGTGTCAGAAAA	-	-	
Da aa Left Arm Step 2-R	*D. amer tan*	922	*ACTAGCCCGTGA* TGATTATGGCTCATGGTGTAGC	*D. amer tan*	ACTAGCCCGTGA	
Da aa Right Arm Step 2-F	*D. amer tan*	923	*AGCCATAATCAT* CACGGGCTAGTGTTTAGCATC	*D. amer tan*	AGCCATAATCAT	
galK Marker-F	*galK*	904	*GGCTTCGATATG* CCTGTTGACAATTAATCATCGGCA	*D.amer tan*	GGCTTCGATATG	
galK Marker-R	*galK*	1191	*CATTTTAGCTAA* TCAGCACTGTCCTGCTCCTT	*D.amer tan*	CATTTTAGCTAA	
Dn left half FMR-F	*D. nova tan*	911	*GGCTTCGATATG* GTAAGCAACGGCTCCTTAATC	*D.amer tan*	GGCTTCGATATG	
Dn left half FMR-R	*D. nova tan*	1184	*CATTTTAGCTAA* CGTCTTGCAATTCGTCATAGG	*D.amer tan*	CATTTTAGCTAA	
Da Left Arm-F	*D. amer tan*	902	AGCGAGTTTTTATTCCCACTG	-	-	*D. amer tan* + *D.nova* left half FMR
Da Left Arm Step 1-R	*D. amer tan*	934	*ATTGTCAACAGG* CATATCGAAGCCGACCT	*galK*	ATTGTCAACAGG	
Da Right Arm Step 1-F	*D. amer tan*	1186	*GGACAGTGCT* GACGTTAGCTAAAATGCTGAGAGAC	*galK*	GGACAGTGCT	
Da Right Arm-R	*D. amer tan*	1187	GCTGTAATTGCACCTGCTGA	-	-	
Da Left Arm Step 2-R	*D. amer tan*	909	*GCCGTTGCTTAC* CATATCGAAGCCGACCTCAT	*D. nova tan*	GCCGTTGCTTAC	
Da Right Arm Step 2-F	*D. amer tan*	1185	*AATTGCAA* GACGTTAGCTAAAATGCTGAGAGAC	*D. nova tan*	AATTGCAA	
galK Marker-F	*galK*	904	*GGCTTCGATATG* CCTGTTGACAATTAATCATCGGCA	*D. nova tan*	GGCTTCGATATG	
galK Marker-R	*galK*	1192	*CATTAAAGCTAA* TCAGCACTGTCCTGCTCCTT	*D. nova tan*	CATTAAAGCTAA	
Da left half FMR-F	*D. amer tan*	912	*GGCTTCGATATG* GTAAGCAACAGCTCCTTAATCC	*D. nova tan*	GGCTTCGATATG	
Da left half FMR-R	*D. amer tan*	1183	*CATTAAAGCTAA* CGTCTTGCAATTCGTCGTAG	*D. nova tan*	CATTAAAGCTAA	
Dn Left Arm-F	*D. nova tan*	902	AGCGAGTTTTTATTCCCACTG	-	-	*D. nova tan* + *D.amer* Left half FMR
Dn Left Arm Step 1-R	*D. nova tan*	934	*ATTGTCAACAGG* CATATCGAAGCCGACCT	*galK*	ATTGTCAACAGG	
Dn Right Arm Step 1-F	*D. nova tan*	1189	*GGACAGTGCT* GACGTTAGCTTTAATGCTGAGAGAC	*galK*	GGACAGTGCT	
Dn Right Arm-R	*D. nova tan*	1190	GGCGAGAGCGAGAGAGAAT	-	-	
Dn Left Arm Step 2-R	*D. nova tan*	910	*GCTGTTGCTTAC* CATATCGAAGCCGACCTCAT	*D.amer tan*	GCTGTTGCTTAC	
Dn Right Arm Step 2-F	*D. nova tan*	1188	*AATTGCAA* GACGTTAGCTTTAATGCTGAGAGAC	*D.amer tan*	AATTGCAA	

Table 3.1: Italicized primer sequence indicates a gene specific tail added for use in PCR fusion.

References:

CALETKA B. C., MCALLISTER B. F., 2004 A genealogical view of chromosomal evolution and species delimitation in the Drosophila virilis species subgroup. Mol Phylogenet Evol **33**: 664–670.

COLOSIMO P. F., HOSEMANN K. E., BALABHADRA S., VILLARREAL G., DICKSON M., GRIMWOOD J., SCHMUTZ J., MYERS R. M., SCHLUTER D., KINGSLEY D. M., 2005 Widespread parallel evolution in sticklebacks by repeated fixation of Ectodysplasin alleles. Science **307**: 1928–1933.

COOLEY A. M., SHEFNER L., MCLAUGHLIN W. N., STEWART E. E., WITTKOPP P. J., 2012 The ontogeny of color: developmental origins of divergent pigmentation in Drosophila americana and D. novamexicana. Evol Dev **14**: 317–325.

GROTH A. C., FISH M., NUSSE R., CALOS M. P., 2004 Construction of transgenic Drosophila by using the site-specific integrase from phage phiC31. Genetics **166**: 1775–1782.

HOEKSTRA H. E., HIRSCHMANN R. J., BUNDEY R. A., INSEL P. A., CROSSLAND J. P., 2006 A single amino acid mutation contributes to adaptive beach mouse color pattern. Science **313**: 101–104.

JEONG S., REBEIZ M., ANDOLFATTO P., WERNER T., TRUE J., CARROLL S. B., 2008 The evolution of gene regulation underlies a morphological difference between two Drosophila sister species. Cell **132**: 783–793.

KALAY G., WITTKOPP P. J., 2010 Nomadic Enhancers: Tissue-Specific cis-Regulatory Elements of yellow Have Divergent Genomic Positions among Drosophila Species. PLoS Genet **6**: e1001222.

LAB O., 2006 Oakley Lab Fusion PCR Protocol. : 1–6.

MCGREGOR A. P., ORGOGOZO V., DELON I., ZANET J., SRINIVASAN D. G., PAYRE F., STERN D. L., 2007 Morphological evolution through multiple cis-regulatory mutations at a single gene. Nature **448**: 587–590.

MORALES-HOJAS R., VIERIRA C. P., VIEIRA J., 2008 Inferring the evolutionary history of Drosophila americana and Drosophila novamexicana using a multilocus approach and the influence of chromosomal rearrangements in single gene analyses. Molecular Ecology **17**: 2910–2926.

SLAWSON E. E., SHAFFER C. D., MALONE C. D., LEUNG W., KELLMANN E., SHEVCHECK R. B., CRAIG C. A., BLOOM S. M., BOGENPOHL J. II, DEE

J., MORIMOTO E. T., MYOUNG J., NETT A. S., OZSOLAK F., TITTGER M. E., ZEUG A., PARDUE M.-L., BUHLER J., MARDIS E. R., ELGIN S. C., 2006 Comparison of dot chromosome sequences from *D. melanogaster* and *D. virilis* reveals an enrichment of DNA transposon sequences in heterochromatic domains. Genome Biology **7**: R15.

TISHKOFF S. A., REED F. A., RANCIARO A., VOIGHT B. F., BABBITT C. C., SILVERMAN J. S., POWELL K., MORTENSEN H. M., HIRBO J. B., OSMAN M., IBRAHIM M., OMAR S. A., LEMA G., NYAMBO T. B., GHORI J., BUMPSTEAD S., PRITCHARD J. K., WRAY G. A., DELOUKAS P., 2007 Convergent adaptation of human lactase persistence in Africa and Europe. Nat Genet **39**: 31–40.

TRUE J. R., 2003 Insect melanism: the molecules matter. Trends Ecol Evol **18**: 640–647.

WARMING S., COSTANTINO N., COURT D. L., JENKINS N. A., COPELAND N. G., 2005 Simple and highly efficient BAC recombineering using galK selection. Nucleic Acids Research **33**: e36.

WITTKOPP P. J., CARROLL S. B., KOPP A., 2003a Evolution in black and white: genetic control of pigment patterns in Drosophila. Trends Genet **19**: 495–504.

WITTKOPP P. J., STEWART E. E., ARNOLD L. L., NEIDERT A. H., HAERUM B. K., THOMPSON E. M., AKHRAS S., SMITH-WINBERRY G., SHEFNER L., 2009 Intraspecific polymorphism to interspecific divergence: genetics of pigmentation in Drosophila. Science **326**: 540–544.

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**: 1849–1858.

WITTKOPP P. J., WILLIAMS B. L., SELEGUE J. E., CARROLL S. B., 2003b Drosophila pigmentation evolution: divergent genotypes underlying convergent phenotypes. PNAS **100**: 1808–1813.

WRIGHT T. R. F., 1987 The Genetics of Biogenic Amine Metabolism, Sclerotization, and Melanization in Drosophila melanogaster. Advances in Genetics **24**: 127–222.

Chapter 4

Local Adaptation for body color in *Drosophila americana*[3]

Abstract

Pigmentation is one of the most variable traits within and between Drosophila species. Much of this diversity appears to be adaptive, with environmental factors often invoked as selective forces. Here, we describe the geographic structure of pigmentation in *Drosophila americana* and evaluate the hypothesis that it is a locally adapted trait.

Body pigmentation was quantified using digital images and spectrometry in up to 10 flies from each of 93 isofemale lines collected from 17 locations across the United States and found to correlate most strongly with longitude. Sequence variation at putatively neutral loci showed no evidence of population structure and was inconsistent with an isolation-by-distance model, suggesting that the pigmentation cline exists despite extensive gene flow throughout the species range and is most likely the product of natural selection. In all other Drosophila species examined to date, dark pigmentation is associated with arid habitats;

---

[3] This chapter is a compilation of data from several individuals published as Wittkopp, P.J., Smith-Winterberry, G., **Arnold, L.L.,** Thompson, E.M., Cooley, A.M., Yuan, D.C., Song, Q. & McAllister, B.F. (2010) Local adaptation for body color in *Drosophila americana. Heredity*. Advanced online publication July 7 (doi:10.1038/hdy.2010.90) My contributions include data for Figure 4.6. Our combined efforts have shown that the pigmentation cline within *D. americana* is due to local adaptation and found that desiccation resistance is not significantly linked with pigmentation in this Drosophila species.

however in *D. americana*, the darkest flies were collected from the most humid regions. To investigate this relationship further, we examined desiccation resistance attributable to an allele that darkens pigmentation in *D. americana*. We found no significant effect of pigmentation on desiccation resistance in this experiment, suggesting that pigmentation and desiccation resistance are not unequivocally linked in all Drosophila species.

Introduction

Clinal variation, in which the average value of a trait changes gradually over a geographic area, can be caused by either neutral or non-neutral evolutionary processes (reviewed by Kawecki and Ebert, 2004). For example, the neutral process of genetic drift can generate a cline through spurious correlations with geographic variables among the numerous segregating polymorphisms throughout the genome. Limited migration between populations (especially when migration rates are correlated with geographic distance) promotes cline formation. Alternatively, natural selection can generate a cline when graded selection favors different genotypes in different geographic regions. In these cases, the balance between selection and gene flow results in a cline, with gene flow acting as a homogenizing force among populations and opposing phenotypic divergence. Phenotypic plasticity can also create clines in the wild; however, phenotypic differences among populations for plastic traits disappear when individuals are reared in a common environment (e.g., Maherali *et al*,

2002). That is, clines generated directly by the environment do not necessarily involve genetic differentiation.

In animals, clinal variation is often observed for body color. For example, in humans, skin color is darkest at the equator, with decreasing melanin in populations located toward the poles (Jablonski and Chaplin, 2000); in deer mice, coat color varies across Florida and Alabama, with the lightest phenotypes found closest to the Gulf of Mexico (Mullen and Hoekstra, 2008); and in the flat periwinkle snail, shell color varies in the Gulf of Maine, with the darkest shells found in the most northern, coolest waters (Phifer-Rixey *et al*, 2008). Each of these clines appears to be adaptive, with selection pressures including UV penetration, camouflage, and thermoregulation, respectively. In Drosophila, pigmentation clines have been reported for *D. melanogaster* (e.g., David *et al*, 1985; Parkash *et al*, 2008; Pool and Aquadro, 2007)*, D. simulans* (Capy *et al*, 1988), the *D. dunni* species subgroup (Brisson *et al*, 2005; Hollocher *et al*, 2000), and other Drosophila species (reviewed by Rajpurohit *et al*, 2008). These clines correlate with both geographic (i.e., latitude, altitude) and climatic (i.e., temperature, humidity) factors. Laboratory studies in *D. melanogaster* and *D. polymorpha* show differences in desiccation resistance between color morphs (Brisson *et al*, 2005; Kalmus, 1941; Parkash *et al*, 2009a; Parkash *et al*, 2009b; Rajpurohit *et al*, 2008), while studies in 12 other insects show an effect of pigmentation on thermoregulation (e.g., Brakefield and Willmer, 1985; Watt, 1969). Drosophila pigmentation is also known to be a plastic trait affected by environmental factors such as food and temperature (e.g., Gibert *et al*, 2007),

suggesting that pigmentation clines observed in nature may not always reflect genetic differences.

The present study examines the geographic distribution of body color in *Drosophila americana*, a member of the virilis species group. The ancestor of *D. americana* colonized North America at least three million years ago and the species appears to have maintained a relatively stable large effective population size since that time: patterns of codon usage in *D. americana* are more consistent with a theoretical population genetic "equilibrium" than they are in the more commonly studied *D. melanogaster* (Haddrill *et al*, 2005; Maside *et al*, 2004). Consistent with this observation, prior studies of *D. americana* suggest extensive gene flow among populations (McAllister, 2002; McAllister, 2003; McAllister and Evans, 2006; Morales- Hojas *et al*, 2008; Schäfer *et al*, 2006; Vieira *et al*, 2003). Despite these signs of genetic homogeneity, however, "a yellowish western group and a blackish eastern group" has been reported within this species (Throckmorton, 1982, p. 239). These "western forms" were collected primarily from Kansas, Nebraska, South Dakota, and Montana (Hsu, 1951). *D. novamexicana*, the closest relative of *D. americana* (Caletka and McAllister, 2004; Morales-Hojas *et al*, 2008), has even lighter and more yellow pigmentation than the western strains of *D. americana*, and has been collected from Arizona, Colorado, Utah and New Mexico (Throckmorton, 1982, p. 239), suggesting a trans-species pigmentation cline that extends longitudinally across the United States. Here, we provide the first quantitative description of the pigmentation cline in *D. americana* by measuring body color in 93 isofemale lines collected

from 17 sites that span much of the latitudinal and longitudinal ranges of *D. americana*. Two different methods for quantifying pigmentation were used, one of which provides visual documentation and the other of which allows high-throughput scoring of live flies. Pigmentation differences among lines and among collection sites are shown to correlate with longitude, which in turn correlates with relative humidity. Patterns of sequence variation suggest extensive gene flow throughout the species range (consistent with prior studies) and reject an isolation-by-distance model of cline formation. We explore the hypothesis that differences in relative humidity among collection sites promote cline formation by testing for an effect of *D. americana* pigmentation alleles on desiccation resistance. In contrast to studies of other Drosophila species (e.g., Brisson *et al*, 2005; Parkash *et al*, 2009a; Parkash *et al*, 2009b), we find no effect of pigmentation on desiccation resistance. We conclude by comparing these results to pigmentation clines observed in other Drosophila species.

Materials and Methods

Fly strains

Two different strategies were used to measure pigmentation in *D. americana.* In "Dataset A," thirteen isofemale lines, derived from eleven broadly distributed geographic locations in the central and eastern United States, were analyzed to provide a species-wide assessment of variability (Table 4.1). Four of these lines, which were obtained from the Drosophila Species Stock Center (Tucson, AZ), were established ~50 years ago from single females captured at

different collection sites. The remaining nine lines were established from females collected between 1999 and 2003 at seven other locations. With the exception of one site (Duncan, NE), each of the collection sites included in Dataset A is represented by only a single isofemale line. These lines capture the breadth of pigment variation over the geographic range of the species. “Dataset B” contains deeper sampling of fewer sites (i.e., 80 isofemale lines from eight different locations), with lines established from flies collected in June and September of 2007 (Table 4.1). Collection sites in Dataset B form a coarse longitudinal transect extending between 82° and 98°W longitude and bounded by 38° and 43° N latitude. Isofemale lines from localities near the eastern and western extremes of the transect (OR and DN, respectively) are included in both datasets; however, different isofemale lines from these collection sites are used in Datasets A and B.

All fly stocks were maintained on standard yeast-glucose medium at 20-22°C. Prior to pigmentation scoring, three males and three females were placed into a vial and their offspring raised at 20°C. (Controlling the number of parents in each vial resulted in similar larval density among genotypes.) Flies were collected within three days of eclosion and aged one week to allow body color to stabilize. All isofemale lines within Dataset A or Dataset B were reared simultaneously, under identical conditions (i.e., light, humidity, temperature, batch of media) to minimize the effect of environmental differences among genotypes.

Table 4.1: Origins of *Drosophila americana* isofemale lines

Collection Site	Lines	Year[1]	Lat[2]	Long[3]	Alt[4]	Temp[5]	RH[6]
Anderson, IN	A00[6*]	unknown	40.1	85.6	279	11.1	67.8%
Poplar, MT	A01[6*]	1947	48.1	105.2	593	4.7	63.9%
Lake Champlain, VT	A04[6*]	1948	44.6	73.4	28	6.1	68.0%
Myrtle Beach, SC	A09[6*]	1961	33.7	78.9	6	19.6	69.2%
Augusta, AR	AA99.6[7*]	1999	35.3	91.4	59	15.6	68.6%
Duncan, NE	DN01.12[8*], DN00.2[8*], DN00.4[8*], DN07.4, DN07.6, DN07.14, DN07.16, DN07.18, DN07.20, DN07.24, DN07.34, DN07.42, DN07.08x05, DN07.10x07, DN07.12x09, DN07.26x19, DN07.28x21, DN07.32x25, DN07.40x31, DN07.44x33, DN07.48x37, DN07.52x41, DN07.54x43, DN07.56x45, DN07.60x47, DN07.62x49, DN07.64x53, DN07.66x55	2000 (DN00.X) 2001 (DN01.X) 2007 (DN07.X)	41.4	97.5	469	8.9	67.9%
Floodgate Park, AR	FP99.32[7*]	1999	34.2	91.1	43	16.6	69.1%
LaSalle Fish & Wildlife Area, IN	II07.2, II07.4, II07.6, II07.8, II07.10	2007	41.2	87.5	192	10.3	68.5%
Hawkeye Wildlife Area, IA	IR04.10[9*]	2004	41.8	91.7	208	10.2	65.7%
Lake Ashbaugh, AR	LA99.48[7*]	1999	36.3	90.8	79	14.8	68.0%

Muscatatuck Natl Wildlife Refuge, IN	MK07.16, MK07.18, MK07.20, MK07.22, MK07.24, MK07.30, MK07.36, MK07.38, MK07.42, MK07.46	2007	38.9	85.8	166	13	67.7%
Niobrara, NE	NN07.6, NN07.8, NN07.10, NN07.14, NN07.16, NN07.17, NN07.18, NN07.22, NN07.26, NN07.28	2007	42.7	98.1	448	8.1	67.7%
Patoka Natl Wildlife Refuge, IN	OC07.2, OC07.12, OC07.14, OC07.16, OC07.18, OC07.24, OC07.26, OC07.28, OC07.32, OC07.34	2007	38.3	87.3	136	13.4	67.3%
Ottawa Natl Wildlife Refuge, OH	OR01.52[10*], OR07.6, OR07.10, OR07.12	2001 (OR01.X) 2007 (OR07.X)	41.6	83.2	175	10.2	68.4%
Saulsbury Bridge Rec. Area, IA	SB02.6[9*]	2002	41.5	91.2	186	10.2	65.7%
Swan Lake Natl Wildlife Refuge, MO	SV07.10, SV07.16, SV07.18, SV07.26	2007	39.6	93.2	202	11.8	66.8%
Killbuck Marsh, OH	WS07.2, WS07.6, WS07.8, WS07.10, WS07.12, WS07.14, WS07.16, WS07.18, WS07.20	2007	40.7	82.0	265	10.3	70.1%

Table 4.1: Lines with no citation are previously unpublished. Lines included in Dataset A are indicated with an asterisk. All other lines are included in Dataset B.

[1] Year female used to initiate isofemale line was collected from the wild.

[2] Latitude of collection site (°N) 3 Longitude of collection site (°W).

[4] Altitude of collection site (meters).

[5] Mean annual daily temperature (from NASA Surface Meteorology and Solar Energy).

[6] Mean annual relative humidity (from NASA Surface Meteorology and Solar Energy).

[7] Drosophila Species Stock Center, San Diego, CA (formerly Tucson, AZ).

[8] McAllister, 2002

[9] McAllister, 2003

[10] McAllister and Evans, 2006

[11] McAllister, 2001

Quantifying pigmentation

For Dataset A, dorsal abdominal pigmentation of each isofemale line was measured in five males and five females (aged 7-10 days) that had been placed in a 10:1 ethanol:glycerol mixture and stored at room temperature for one hour to one month. Storage time of each individual was variable within each line and did not differ systematically among lines. We find that abdominal pigmentation is visually stable over this time window under these conditions. Dorsal abdominal cuticle was dissected from each fly, all underlying tissue was removed, and the single layer of adult cuticle was mounted in Hoyer's solution. All mounted cuticles were imaged using a Scion 1394 (Frederick, MD) camera under constant lighting conditions. Body color was quantified for each fly by using Image J (NIH) to calculate the average median pixel intensity of 20 randomly selected (and non-overlapping) regions in gray-scale images of dorsal abdominal cuticle from segments A3, A4 and A5, using a measurement scale that ranged from 0 (black) to 255 (white) (Figure 4.1). A subset of samples was also analyzed using color

images and found to provide similar discrimination among phenotypes to their gray-scale counterparts.

Figure 4.1: Image-based quantification of pigmentation intensity

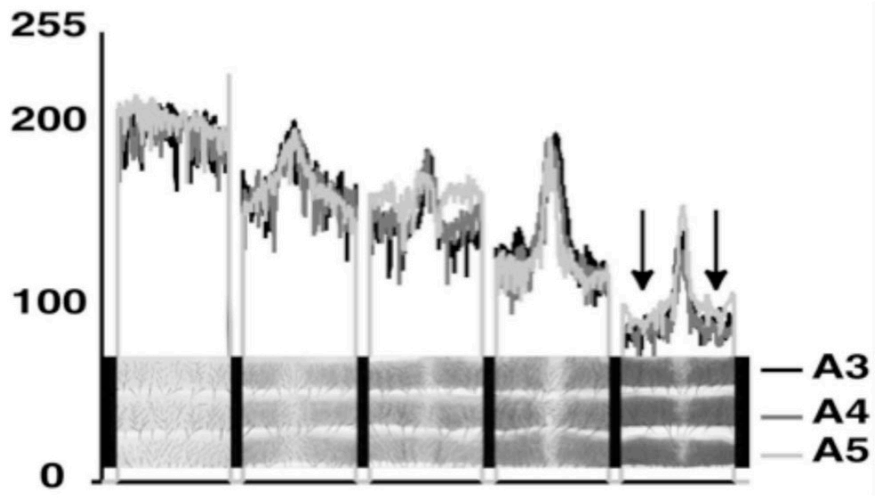


Figure 4.1: Gray scale images of abdominal cuticles (segments A3, A4, and A5) from five adult flies are shown. Lines show the average pixel intensity at each horizontal position within each segment. Note the increased intensity values along the dorsal midline of most cuticles where pigmentation is lighter. Pigment quantification for Dataset A was limited to the regions flanking this lighter midline region (arrows).

For Dataset B, dorsal abdominal pigmentation of each isofemale line was measured in five males and five females, aged 7-10 days, using a custom-built R-series Fiber Optic Reflection Probe with a 50 micron diameter fiber, a LS-1 Tungsten Halogen Light Source, and a USB4000 Spectrometer (Ocean Optics Inc., Dunedin, FL). The reflection probe contained six fiber optic wires that transmitted light to the fly cuticle and a seventh, central, fiber optic wire that

transmitted light reflected off the sample to the spectrometer. The tip of the probe was encased by a custom-built shield constructed by the instrument shop in the chemistry department at the University of Michigan following the blueprint described at https://www.lifesci.ucsb.edu/~endler/OceanOpticsList.pdf. This probe shield ensures a constant distance (~1 cm) and angle (45 degrees) between the fly cuticle and probe among measurements (Uy and Endler, 2004). The diameter of the probe tip (~0.7 mm) is approximately half of the anteroposterior length of one *D. americana* dorsal abdominal segment (i.e., tergite). After calibrating the SpectraSuite Spectroscopy Operating Software (Ocean Optics) with a WS-1 Diffuse Reflection Standard (Ocean Optics), spectral reflectance of visible light (ranging from 0% to 100%) was recorded from five non-overlapping regions of dorsal abdominal cuticle (all located within segments A3, A4, and A5) from each fly. All measurements were collected over two consecutive days, with the isofemale lines scored in random order. Reference spectra taken from four dissected and mounted *D. americana* abdominal cuticles with varying pigmentation intensities were found to be similar both days. Light from 610 to 660 nm wavelengths provided the greatest discrimination among the lightest and darkest control cuticles (Figure 4.2), and custom Perl scripts were used to calculate the average reflectance of light in this range from each reflectance spectrum. In general, replicate measurements from the same fly were similar; however, extreme outliers were occasionally observed, which most likely resulted from the misalignment of the probe tip with the fly cuticle. To

reduce the impact of these outliers, the median value from each fly (rather than the mean) was used for analysis.

Figure 4.2: Pigmentation quantification using a custom-built fiber-optic probe for spectrometry

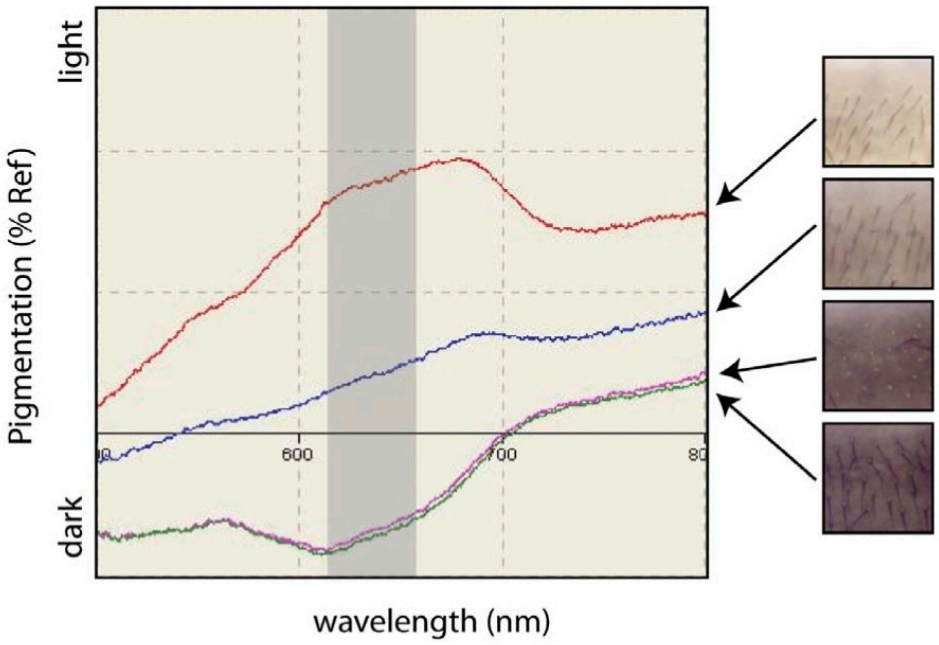


Figure 4.2: Regions from four reference cuticles used to calibrate spectrophotometer each day are shown. The percentage of light reflected at each wavelength is shown. As described in Materials and Methods, the reflectance of light between 610 and 660 nm (gray) was used as a measure of pigmentation for Dataset B.

DNA sequence variation and population genetic analysis

Genomic DNA was extracted from a single male from each isofemale line using the “squish prep” protocol (Gloor *et al*, 1993). For Dataset A, regions from the following genes were amplified and sequenced in all lines except FP, which

died prior to molecular analysis: *cytochrome b* (*cytB*, mitochondrial, 619 bp), *transformer* (*tra*, nuclear, 839 bp), *bazooka* (*baz*, nuclear, 575 bp), *l(1)G0007* (nuclear, 513 bp). In Dataset B, the 839 bp region from the *tra* gene was successfully amplified and sequenced in 34 of the 80 isofemale lines, including at least three lines from each collection site. Sequences of primers used for both amplification and sequencing are available upon request. Sequences were assembled and aligned using CodonCode Aligner (Dedham, MA) and manually validated by P.J.W. for Dataset A and D.C.Y. for Dataset B. They are available through Genbank with the following accession numbers: GU299293 – GU299340 (Dataset A) and GU248275 - GU248308 (Dataset B). Seven of the 34 *tra* sequences from Dataset B (WS07.14, DN07.52x41, NN07.08, II07.10, OR07.10, SC07.18, MK07.24) were heterozygous at one to seven sites and were resolved into two haplotypes (both of which were included in the sequence analysis) using the PHASEv2.1 algorithm (Stephens and Donnelly 2003) implemented in DNAsp v5.10.00 (Librado and Rozas 2009).

The following measures of genetic variability were calculated for each gene region using DnaSP v5.10.00 (Librado, and Rozas 2009): the number of segregating sites (S), haplotype diversity (H_d), nucleotide diversity per site (π), and theta per site based on S (θ). For *tra* sequences of lines included in Dataset B, we also calculated Fst and Kst (Hudson et al. 1992) and assessed their significance using DnaSP and Arlequin v3.11 (Excoffier et al. 2005), respectively. DnaSP was also used to calculate the test statistics Tajima's D and Fu and Li's D*, and their statistical significance was determined using the distributions

provided in the original descriptions of these statistics (Fu and Li, 1993; Tajima, 1989) as well as using 10,000 coalescent simulations based on summary statistics of the observed samples. Pair-wise genetic distances among all strains were calculated for each gene using the Tamura-Nei distance model of nucleotide substitutions (Tamura and Nei, 1993), as implemented in MEGA v4.0.2 (Kumar *et al*, 2004). Sites with missing data or gaps were excluded from all analyses.

Statistical analyses

Pigmentation was analyzed primarily using PROC MIXED in SAS v9.1 (Cary, NC), with all models described below fitted using restricted maximum likelihood (REML). For Dataset A, pigmentation measurements were fitted to the following model to test for effects of line and sex:

$$Y_{ijkl} = L_i + S_j + SL_{ij} + ISL_{ijk} + e_{ijkl},$$

where Y_{ijkl} is the mean pigmentation intensity for cuticle region *l*, from individual *k*, of sex *j*, from line *i*. *L* and *S* are fixed effects of isofemale line and sex, respectively; *I* is the random effect of individual within each sex*line combination; and e_{ijkl} is the residual error. Y_{ijkl} was weighted by the area of the cuticle region analyzed, with larger regions weighted more heavily than smaller regions. For each line (L_i) and each sex within each line (SL_{ij}), the least squares mean and 95% confidence interval were calculated. Least squares means were compared among lines using a Tukey's HSD post hoc test. For Dataset B,

pigmentation measurements, consisting of a single (median) 2 pigmentation score per fly, were fitted to the following model:

$$Y_{ijkl} = P_i + S_j + SP_{ij} + LSP_{ijk} + e_{ijkl},$$

where Y_{ijkl} is the pigmentation score for individual *l*, from isofemale line *k*, of sex *j*, from geographic population *i*. *P* and *S* are fixed effects of population and sex,respectively; *L* is the random effect of line within each population by sex combination; and e_{ijkl} is the error among pigmentation measures from individuals derived from the same isofemale line. For each population (P_i) and each sex within each population (SP_{ij}), the least squares mean and 95% confidence interval was calculated. Least squares means were compared among populations using Tukey's HSD post hoc test. To test for geographic trends in pigmentation, we fitted both datasets to the following model:

$$Y_{jkl} = T_j + G_k + e_{jkl,}$$

where Y_{jkl} is the least squares mean pigmentation intensity for each line *l* collected from latitude *j* and longitude *k*. *T* and *G* represent the continuous covariates of latitude and longitude, respectively. For Dataset A, only the intermediate of the three lines from Duncan, Nebraska was used, to avoid over-weighting data from this location. Males and females were analyzed separately for each dataset, because a significant effect of sex was detected (see Results section).

To test for evidence of isolation-by-distance, we used a Mantel test to compare genetic and geographic distance among lines in Dataset A and populations in Dataset B. This test was conducted using the web-based Isolation-

by-distance Web Service (IBDWS) software v3.15 (Jensen *et al*, 2005) available at http://ibdws.sdsu.edu/. Geographic distances for this test were measured in kilometers and calculated based on longitude and latitude of collection sites using the web-based software developed by Dr. John Byers (U.S. Arid-Land Agricultural Research Center, USDA-ARS, http://www.chemical-ecology.net/java/lat-long.htm). This analysis was also performed using geographic distances measured in degrees longitude. Genetic distances were calculated as described in the DNA sequence variation section above. Mantel tests were performed using both the raw genetic distances as well as the logarithm of genetic distance. Significance was assessed using 1000 permutations of the genetic and geographic distances, conducted by the IBDWS software.

Desiccation resistance

Interspecific introgression lines were used to specifically test whether alleles that affect pigmentation have a corresponding affect on desiccation resistance. As described in Wittkopp et al. (2009), lines were constructed by crossing *D. americana* females to males of their lightly pigmented sister species, *D. novamexicana*, and backcrossing the resulting F1 hybrid females to *D. novamexicana* males. Backcrossing was continued for ten consecutive generations, with a single female heterozygous for the *D. americana* and *D. novamexicana* alleles of pigmentation genes *tan* and *ebony* genes mated to a *D. novamexicana* male in each generation. The introduction of either the *tan* or *ebony* genomic region from *D. americana* into *D. novamexicana* was sufficient to

cause a visible darkening of pigmentation, with flies carrying *D. americana* alleles for both quantitative trait loci (QTLs) visibly darker than those carrying *D. americana* alleles for either QTL region alone (Wittkopp *et al*, 2009). Using these introgression lines, we constructed sex-specific pairs of genotypes with significant differences in pigmentation. The two male genotypes were both hemizygous for the *D. americana tan* QTL allele, but differed by the presence or absence of the *D. americana ebony* QTL allele, resulting in “dark” and “light” pigmentation phenotypes, respectively. Similarly, the two female genotypes were both heterozygous for the *D. americana tan* QTL allele, but differed by the presence or absence of the *D. americana ebony* QTL allele, again, resulting in “dark” and “light” pigmentation phenotypes, respectively.

Desiccation resistance was measured by placing 7-10 day old virgin males and females into 5 mL Polystyrene round-bottom vials with mesh caps (BD Falcon) which were stored in a 5.7L plastic snap-lid container (Rubbermaid) with 200 grams of Drierite (8 mesh), sealed with parafilm, and stored at 20°C. A control container was prepared in the same manner, with the substitution of a moist paper towel for the Drierite. Wired indoor/outdoor hygrometers (RadioShack) were used to monitor relative humidity in each container: the desiccant container maintained a relative humidity level of < 20% (the minimum detectable with the hygrometer) throughout the experiment while the control container maintained an average of 85% relative humidity. Each container held ten replicate vials of females and eight replicate vials of males, with each vial containing three “light” and three “dark” flies of the same sex. Beginning fifteen

hours after placing the vials in the box, the number of dead flies (assessed by lack of visible movement when the vial was tapped) and the pigmentation of each dead fly (light or dark) was recorded every hour until all flies in the desiccation group died (50 hours). Survival curves were compared using a non-parametric log rank test, which compares the observed numbers of deaths at each time point between samples.

Results

To characterize the geographic distribution of body color in *D. americana*, we examined two distinct sets of isofemale lines. The first ("Dataset A"), which contained a single isofemale line from each of eleven populations that span the known east-west range of *D. americana* (Throckmorton, 1982), was used to provide an overview of pigmentation differences across the species' range. The second ("Dataset B"), which contained multiple isofemale lines derived from each of eight populations representing a coarse longitudinal transect through the central region of the species range, was used to assess body color variation within and between collection sites. The geographically extreme populations from Dataset A were not included in Dataset B because only a single isofemale line was available from these sites. Figure 4.3 and Table 4.1 describe the collection sites and individual isofemale lines in more detail.

Figure 4.3: Body color in *D. americana* varies with longitude

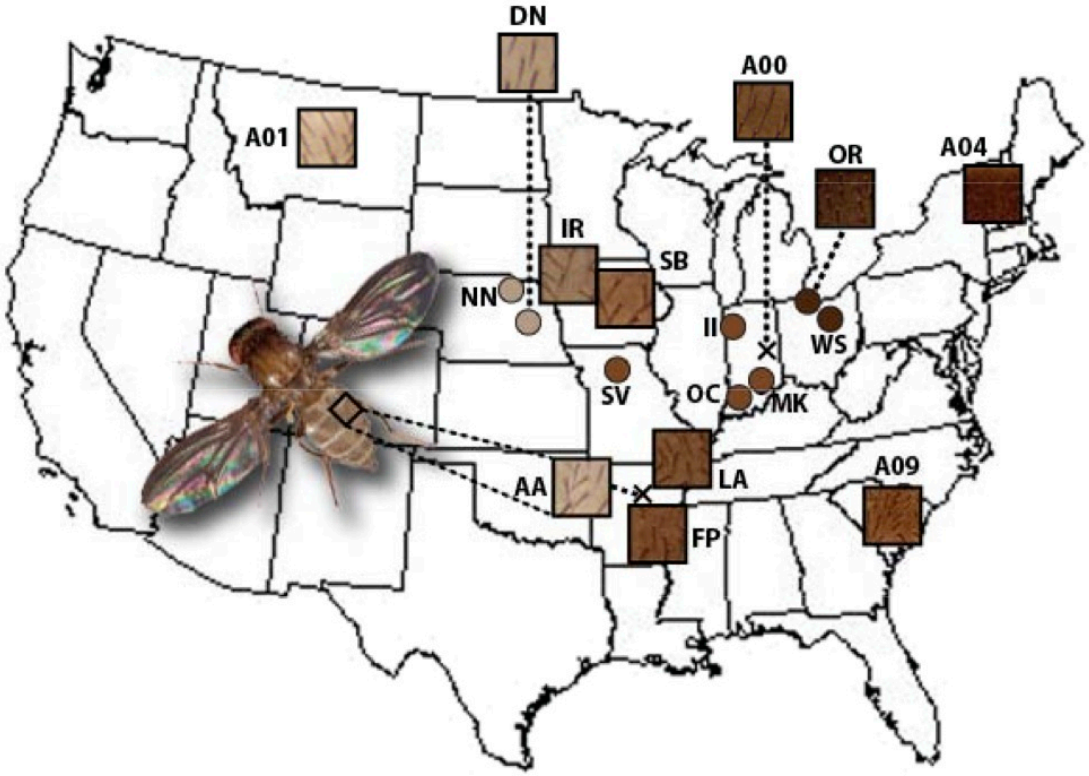


Figure 4.3: (A) For all collection sites included in Dataset A, a representative image of a region of dorsal abdominal cuticle is shown. All images are from males, and the DN2 line is shown for the Duncan, Nebraska collection site. Collection sites included in Dataset B are indicated with colored circles, where the circle approximates the cuticle color of flies from the collection site. As described in Table 4.1, a non-overlapping set of isofemale lines from Duncan, NE (DN) and Ottawa National Wildlife Refuge, OH (OR) are included in Datasets A and B.

Quantitative metrics for adult body color in Drosophila

Drosophila pigmentation is typically analyzed using a subjective and arbitrary scoring scale based on visual assessments of pigmentation (e.g., Brisson *et al*, 2005; David *et al*, 2002; Hollocher *et al*, 2000; Wittkopp *et al*, 2003b). While these measurements are generally consistent for a single observer

under controlled lighting conditions, discriminating among subtle gradations of body color is challenging for even the most experienced researcher. A preliminary visual assessment of pigmentation among isofemale lines of *D. americana* revealed obvious differences between the lightest and darkest lines, with subtle variation in intermediate body colors that we were unable to reliably and consistently classify by eye. Therefore, we concluded that an objective and quantitative method of pigmentation scoring was essential for describing the geographic distribution of body color in *D. americana*.

Two quantitative methods for scoring Drosophila pigmentation were developed and used in this study. The first method, which was applied to the 13 isofemale lines in Dataset A, involved dissection of dorsal abdominal cuticles from preserved flies (five males and five females per line) followed by imaging and computational analysis of digital images from each individual cuticle. This method produced semi-permanent samples and pictures of isolated body cuticles (Figure 4.3); however, the dissection, mounting, and imaging was labor intensive and time-consuming, making it impractical for analyzing multiple individuals from each of the 80 isofemale lines included in Dataset B. To overcome this technical hurdle, we adapted a custom spectrometry system that allowed us to rapidly measure pigmentation of live (but anesthetized) flies. Using this method, we quantified pigmentation of 800 flies (5 males and 5 females from each of 80 lines) for Dataset B in only two days.

The distribution of body color variation within D. americana

To test for differences in pigmentation between sexes and among collection sites, measurements from Datasets A and B were fitted separately to linear mixed models (see Materials and Methods). Significant differences in pigmentation were observed among collection sites in both Dataset A ($F = 12.46$, $P < 0.0001$) and Dataset B ($F = 36.01$, $P < 0.0001$). Post-hoc analysis of these data identified four statistically distinct pigmentation groups in Dataset A and two statistically distinct groups in Dataset B (Table 4.2). *D. americana* females were found to be slightly, but significantly, lighter in color than *D. americana* males in both datasets (Dataset A: $F = 4.32$, $P = 0.0405$; Dataset B: $F = 10.83$, $P<0.0001$), although this sexual dimorphism is not visually apparent under a dissecting microscope and has not been recognized previously. To examine the geographic distribution of different pigmentation phenotypes, we fitted the pigmentation measures to a linear model that included latitude and longitude of the collection site as covariates. A highly significant effect of longitude was observed for both datasets, while latitude showed no significant effect in either case (Table 4.3). Manual inspection of the geographic distribution of pigmentation phenotypes suggests that the longitudinal gradient may actually be non-linear with the largest change in pigmentation occurring near 90 degrees west longitude; however, non-linear models fit to our data with SAS v9.1 (proc NLIN) and Cfit (Gay et al. 2008) failed to converge.

Table 4.2: Pigmentation intensity among isofemale lines and collection sites

	Source	Longitude	LSM	Group
Dataset A	A01	105.2	163.41	a
	AA	91.4	161.39	a
	DN00.2	97.5	149.36	a
	DN00.4	97.5	140.96	a,b
	DN01.12	97.5	128.64	b
	SB	91.2	130.76	b
	FP	91.1	127.97	b,c
	IR	91.7	126.18	b,c,d
	A09	78.9	113.44	c,d
	A00	85.6	110.75	c,d
	LA	90.8	109.47	c,d
	A04	73.4	106.29	d
	OR	83.2	98.00	d
Dataset B	NN	98.1	-14.53	a
	DN	97.5	-15.03	a
	OC	87.3	-17.76	b
	MK	85.8	-18.00	b
	SV	93.2	-18.57	b
	II	87.5	-18.62	b
	WS	82.0	-18.85	b
	OR	83.2	-19.51	b

Table 4.2: For each isofemale line (“Source”) analyzed in Dataset A, the longitude of the collection site (in degrees W), the least-squares mean (LSM) from the mixed model described in Materials and Methods, and the pigmentation group (based on Tukey’s post-hoc test) are shown. Full description of each “Source” is provided in Table 4.1. The same metrics are shown for Dataset B, except that least-squares means are reported for each collection site rather than each isofemale line and are based on a slightly different mixed model, as described in Materials and Methods. Note that different pigmentation scoring techniques were used for Datasets A and B, thus the LSM values reported for the two datasets are on different scales. Tukey post-hoc analysis was performed separately for the two datasets, thus there is no implied relationship between groups with the same label (e.g., “a”) in Datasets A and B.

Table 4.3: Pigmentation varies significantly by longitude, but not latitude

	Dataset A		Dataset B	
	Females	**Males**	**Females**	**Males**
Effect of latitude	$F = 0.27$ $P = 0.61$	$F = 0.57$ $P = 0.47$	$F = 0.50$ $P = 0.48$	$F = 0.10$ $P = 0.75$
Effect of longitude	$F = 14.66$ $P = 0.005$	$F = 6.50$ $P = 0.034$	$F = 22.39$ $P < 0.0001$	$F = 15.66$ $P = 0.0002$
Mean pigment intensity	131.8 ± 2.35	124.5 ± 2.63	-16.44 ± 0.189	-18.47 ± 0.192

Table 4.3: F- and P- values were obtained for the following model: Yjkl = Tj + Gk + ejkl (described more fully in Materials and Methods). Mean pigmentation scores for males and females are given in the bottom row, ± standard error. Pigmentation was measured using different scales for Dataset A and Dataset B; however, larger ordinal values correspond to lighter coloration in both cases.

Clinal variation is inconsistent with a neutral isolation-by-distance model

The observed longitudinal gradient of pigmentation in *D. americana* may be caused by local adaptation or genetic drift with geographically limited migration (i.e., isolation-by- distance). These two different evolutionary processes can be distinguished by comparing the spatial distribution of pigmentation to the spatial distribution of genetic variation. Specifically, clines resulting from isolation-by-distance are expected to show a positive correlation between genetic and geographic distance at neutral loci, while clines resulting from natural selection despite ongoing gene flow are not. To distinguish between these hypotheses, we surveyed sequence variation among isofemale lines in both datasets.

For Dataset A, regions from the *cytB, baz, l(1)G0007*, and *tra* genes were sequenced in 12 of the 13 isofemale lines; no sequences were obtained from the FP isofemale line because the stock died prior to molecular genetic analysis. According to Flybase (Drysdale and Consortium, 2008), none of the loci surveyed affects pigmentation. Neutrality tests based on Tajima's D and Fu and Li's D*, both of which compare the observed distribution of polymorphism to a distribution expected under a neutral model, were consistent with neutrality (Table 4.4), suggesting that variation of the sequenced loci should reflect gene flow among the populations sampled. A region of sequence from the *tra* gene obtained from 34 isofemale lines from Dataset B, including at least 3 lines from each collection site, was also consistent with neutrality (Table 4.4). Furthermore, pair-wise Fst and Kst for sequences from Dataset B showed no significant differences between populations after correcting for multiple tests (Table 4.5), and there was also no evidence of population subdivision when all populations in Dataset B were considered together (Kst = 0.018, P = 0.20). Finally, Fst and Kst were also not significant for either dataset when sequences were compared between "light" and "dark" pigmentation classes (Table 4.6).

Table 4.4: Genetic variation among isofemale lines

Gene	Length (bp)	S[1]	Hd[2]	π[3]	θ[4]	Genomic location	Tajima's D[5]	Fu&Li's D*[6]
Dataset A:								
baz	575	23	0.85	0.014	0.014	X-linked	0.25	-0.13
cyt b	619	17	1.00	0.007	0.009	mitochondrial	-1.05	-1.21
l(l)G0007	513	37	0.91	0.024	0.026	X-linked	-0.35	-0.39
tra	839	38	0.99	0.016	0.016	autosomal	-0.26	0.05
Dataset B:								
tra	549	51	0.99	0.015	0.021	autosomal	-1.01	1.16

[1] Number of segregating sites

[2] Haplotype diversity

[3] Nucleotide diversity per site

[4] Per site based on S

[5] P-values calculated using the beta distribution of D 10,000 coalescent simulations in DnaSP. All P > 0.05.

[6] P-value calculated using simulations in Fu and Li 1993 as well as with 10,000 coalescent simulations in DnaSP. All P > 0.05

Table 4.5: Comparison of genetic diversity within and among populations

	DN	MK	NN	OC	OR	SC	WS	II
DN		-0.05	-0.05	-0.12	0.01	-0.03	0.00	-0.04
MK	-0.01		-0.01	-0.12	0.23	-0.03	0.09	-0.03
NN	-0.01	0.02		-0.08	0.02	-0.02	0.13	-0.04
OC	-0.05	-0.03	-0.03		-0.03	-0.04	0.02	-0.13
OR	0.05	0.08	0.05	0.03		0.05	0.10	-0.02
SC	0.01	0.01	0.02	0.01	0.08*		0.11	-0.11
WS	-0.03	0.06	0.03	-0.03	0.09*	0.03		0.04
II	-0.04	0.02	0.01	-0.04	0.02	-0.04	-0.05	

Table 4.5: Fst (above diagonal) and Kst (below diagonal) are shown for *tra* sequences from Dataset B. Two letter abbreviations refer to collection sites, as defined in Table 4.1. P-values for all Fst and Kst values are >0.05, except Kst(OR-SC) and Kst(OR-WS) (labeled with asterisks in table), which have p-values of 0.04 and 0.02, respectively; both of these values are nonsignificant after a sequential Bonferroni correction for multiple (n=56) tests.

Table 4.6: Comparison of genetic diversity between pigmentation classes

Dataset A	**Fst**	**Kst**
baz	0.07 (-0.02)	0.03 (0.0005)
cytB	-0.08 (-0.09)	-0.05 (-0.04)
l(1)G0007	-0.04 (0.07)	0.005 (0.03)
tra	0.003 (-0.006)	-0.01 (-0.003)
Dataset B		
tra	0.0008	-0.002

Table 4.6: For each dataset, sequences were divided into “light” and “dark” classes using the Tukey groups shown in Table 4.2. For Dataset A, the light class included isofemale lines from group a (i.e., A01, AA, DN2), whereas the dark class included isofemale lines from groups b, c, and d (i.e., DN12, SB, IR, A00, A04, A09, LA and OR). DN4, which was assigned to groups a and b by the Tukey analysis, was treated in separate analyses as a member of the light and dark classes, with the results from including DN4 in the dark class shown in parentheses. For Dataset B, the light class included all sequences from the NN and DN populations, whereas the dark class included all sequences from the OC, MK, SV, II, WS and OR populations. In all cases, Fst and Kst values were consistent with a null hypothesis of no genetic differentiation between pigmentation classes (0.15 > P-value > 0.95).

Sequences from both Dataset A and Dataset B were used separately to test a model of isolation-by-distance by using a Mantel test to compare pairwise genetic and geographic distances. In Dataset A, we found no significant

relationship between genetic distance and geographic distance, regardless of whether geographic distance was measured in kilometers (Table 4.7) or degrees longitude (data not shown). Similarly, sequences from Dataset B were also inconsistent with an isolation-by-distance model, regardless of whether the pairwise Tamura-Nei genetic distance among isofemale lines or Fst between populations was used to estimate genetic distance or whether kilometers (Table 4.7) or degrees longitude (data not shown) was used to measure geographic distance.

Table 4.7: Evaluating a model of isolation-by-distance

	genetic distance				log(genetic distance)			
Dataset A	**Z**	**r**	**P[1]**	**R^2 (RMA)**	**Z**	**r**	**P[1]**	**R^2**
bazooka	983	-0.11	0.76	0.0128	-154804	-0.23	0.97	0.0506
cytochrome b	503	0.17	0.19	0.0296	-151664	0.13	0.24	0.0163
transformer	1055	-0.18	0.80	0.0341	-131479	-0.22	0.92	0.0475
l(1)G0007	1725	-0.02	0.51	0.0003	-127491	-0.09	0.73	0.0087
Dataset B								
transformer	7640	0.01	0.39	9.34×10^{-5}	-949507	0.04	0.17	1×10^{-4}
transformer[2]	-1567	0.005	0.43	2.59×10^{-5}	-9×10^{21}	-0.02	>0.99	5×10^{-4}

Table 4.7: Mantel tests for correlation between genetic and geographic distance (measured in kilometers).

[1] P-values are one sided and based on 1000 permutations. The null hypothesis is that the correlation coefficient is less than or equal to 0.

[2] Genetic distance measured by pairwise Fst rather than the Tamura-Nei genetic distance.

Based on the absence of clear evidence for genetic differentiation among collection sites, which is consistent with prior studies that also failed to find evidence of population structure in *D. americana* using different samples (Maside *et al*, 2004; McAllister, 2003; McAllister and Evans, 2006; Morales-Hojas *et al*, 2008; Schäfer *et al*, 2006; Vieira *et al*, 2003), as well as the rejection of an isolation-by-distance model by both datasets, we conclude that the observed clinal variation for pigmentation in *D. americana* is unlikely to be the product of genetic drift in distinct populations, but rather is more likely maintained across the species range by natural selection for locally adaptive phenotypes.

Differential selection for desiccation resistance unlikely to explain the pigmentation cline

As described above, we found that pigmentation in *D. americana* correlates much more strongly with longitude than latitude (Figures 4.2A, 4.2B and Figure 4.5). Further analysis showed that pigmentation in Dataset B also correlates significantly with altitude (Figure 4.4C), although this is not surprising given that longitude and altitude 12 are themselves correlated for the collection sites examined ($R^2 = 0.52$). In other Drosophila species, latitude and altitude are the primary correlates with pigmentation clines (see Discussion). Differences in temperature and relative humidity among collection sites, which presumably affect thermal and desiccation tolerances, respectively, are the most commonly invoked selective agents for the formation and maintenance of pigmentation clines in Drosophila (reviewed by Rajpurohit *et al*, 2008; True, 2003; Wittkopp *et*

al, 2003a), and among all collection sites examined in this study, relative humidity correlates more strongly with longitude ($R^2 = 0.37$) than latitude 20 ($R^2 = 0.08$), while the opposite is true of temperature -- it correlates more strongly with latitude ($R^2 = 0.96$) than longitude ($R^2 = 0.37$). Despite these correlations, no significant direct correlation was found between pigmentation and temperature or relative humidity in either dataset (Figures 4.2D and 4.2E).

Figure 4.4: Geographic and environmental correlates with pigmentation

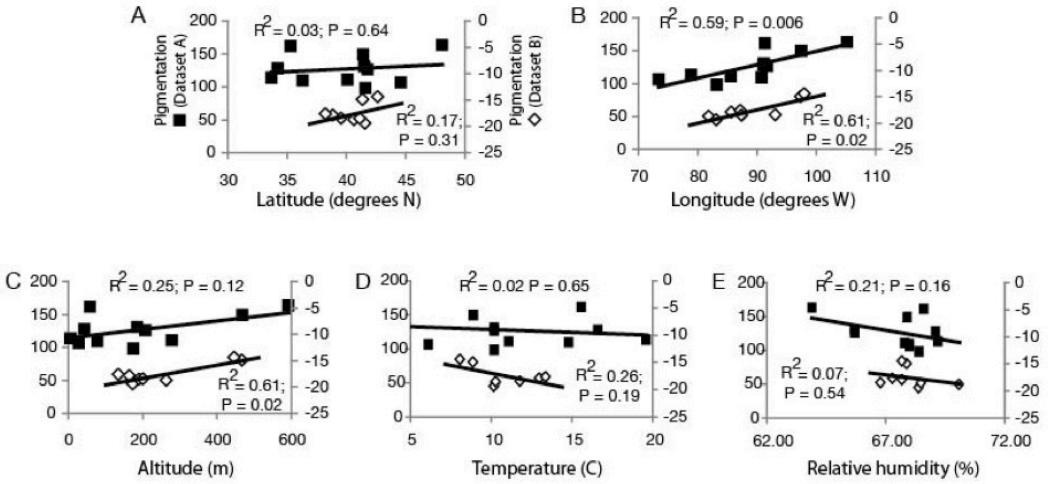


Figure 4.4: Line- specific least-squares means for pigmentation from Dataset A (closed boxes) and population-specific least-squares means for pigmentation from Dataset B (open diamonds) are plotted against latitude (A), longitude (B), altitude (C), temperature (D), and relative humidity (E) for each collection site. The R^2 squared values are shown for each comparison along with their associated P-values (determined using proc reg in SAS), with values for Dataset A in the top part of each panel and values for Dataset B in the lower part of each panel. Note that Dataset A encompasses a larger range for each of the variables analyzed than Dataset B.

Figure 4.5: Line-specific least-squares means (and 95% confidence intervals) for pigmentation in Dataset B

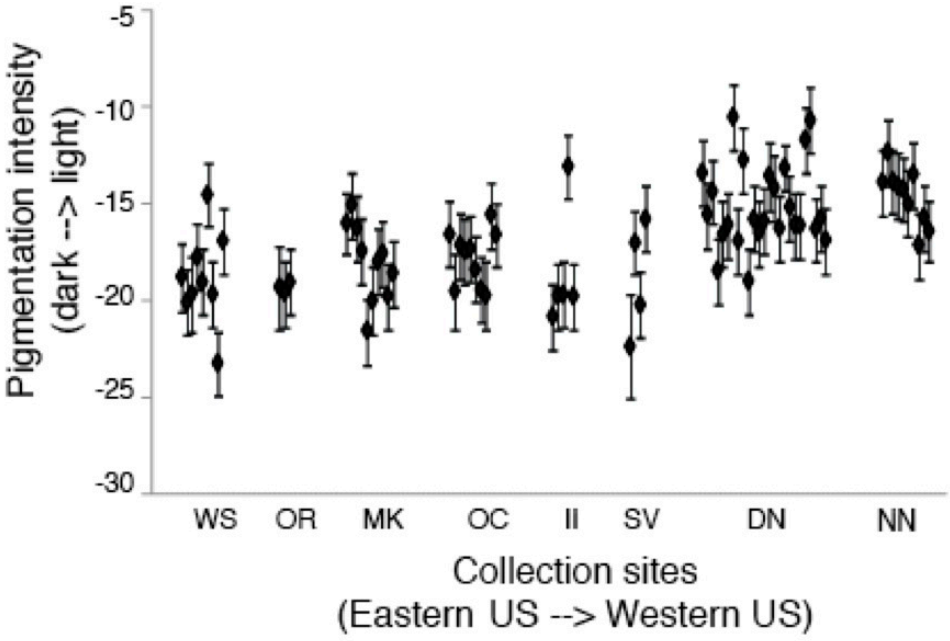


Figure 4.5: Collection sites are arranged by increasing degrees west longitude.

Associations between pigmentation and humidity have been reported in at least seven Drosophila species (Brisson *et al*, 2005; Parkash *et al*, 2008; Parkash *et al*, 2009b; Rajpurohit *et al*, 2008). In all cases, darker flies were collected from less humid environments. Interestingly, *D. americana* appears to show the opposite pattern: lighter flies were collected from less humid environments (Figure 4.4E), suggesting that distinct selective mechanisms may be operating in *D. americana*. To test the effect of pigmentation on desiccation resistance as specifically as possible, we compared desiccation resistance between sex-specific pairs of introgression lines that differed dramatically for

pigmentation, but minimally for genotype. This experimental strategy minimizes the possibility that correlated variation with no effect on pigmentation causes differences in desiccation resistance through other physiological mechanisms. As described in the Materials and Methods section, the introgression lines used for this analysis contained genetic material from both *D. americana* and its closest relative, *D. novamexicana* (Wittkopp *et al*, 2009), with the dark and light genotypes examined differing only by the presence or absence, respectively, of the *D. americana* allele of *ebony* and surrounding genes.

We measured desiccation resistance in each of these pigmentation classes using the same desiccation tolerance assay that was used to demonstrate differences in desiccation resistance between pigmentation classes of other Drosophila species (Brisson *et al*, 2005; Parkash *et al*, 2008; Parkash *et al*, 2009a; Parkash *et al*, 2009b; Rajpurohit *et al*, 2008). Surprisingly, we found no significant difference in desiccation resistance between light and dark flies of either sex (Figure 4.6). That is, flies of the same sex had similar survival times (as measured by a log rank test) under desiccating conditions regardless of whether they had light or dark pigmentation (Males: $\chi^2 = 0.3$, d.f. = 1, P = 0.58; Females: $\chi^2 = 1.4$, d.f. = 1, P = 0.2). Significant differences in survival time were observed between the sexes, however. Males survived longer for both light ($\chi^2 = 21.2$, d.f. = 1, $P = 4 \times 10^{-6}$) and dark ($\chi^2 = 27.4$, d.f. = 1, $P = 2 \times 10^{-7}$) flies, which may be caused by sexual dimorphism and/or differences in their X chromosome genotypes. Sexual dimorphism for desiccation resistance has been reported in

other Drosophila species as well, although females typically survive longer than males under desiccating conditions (Brisson *et al*, 2005; Matzkin *et al*, 2007).

Figure 4.6: Body color has no effect on desiccation resistance

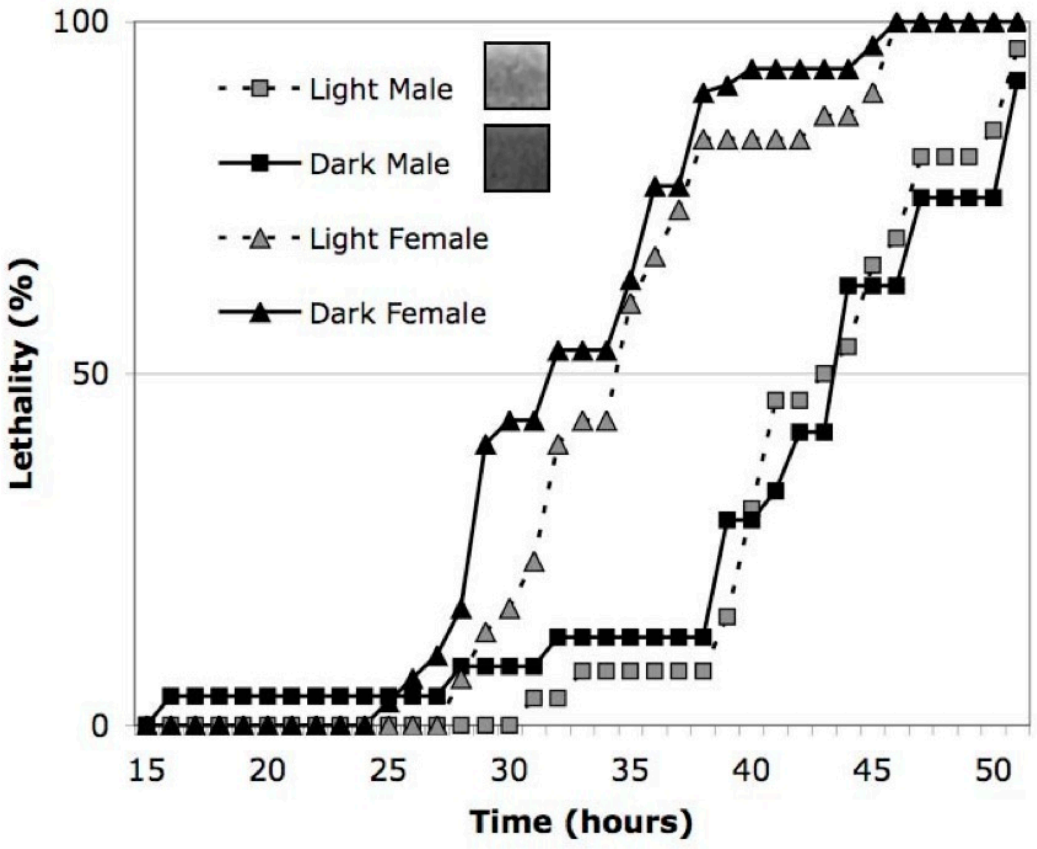


Figure 4.6: A time-course of lethality under desiccating conditions is shown for both “light” (gray, dotted lines) and “dark” (black, solid lines) flies. Triangles represent females and squares represent males. (For a full description of the genotype of these flies, see Materials and Methods.) Pigmentation of the dorsal abdominal cuticle from a "light" male and a "dark" male is shown. Females in both pigmentation classes were slightly lighter than their male counterparts because the introgressed X-linked *D. americana* pigmentation allele(s) were hemizygous in males and heterozygous in females.

Discussion

Pigmentation is one of the most variable traits in the genus Drosophila: differences in body color are common among individuals within a population, among populations of the same species, and among closely related species. This study uses two objective methods of scoring pigmentation, one of which allows for high-throughput analysis, to provide the first quantitative description of body color variation among geographic isolates of *D. americana.* A longitudinal gradient for pigmentation is described, with the lightest body color found in the western extent of the species range. The findings accurately capture previous references to the existence of a difference in pigment between western and eastern flies, which were previously recognized based solely on anecdotal observations. Moreover, this study revealed the existence of a slight sexual dimorphism characterized by more lightly pigmented females.

Patterns of *D. americana* sequence variation (observed in this and prior studies) indicate extensive gene flow among populations and are inconsistent with the differentially pigmented forms being established via a neutral isolation-by-distance model of evolution. In contrast, Hsu (1952) identified several chromosomal inversions that differ in frequency between western and eastern populations. An inversion located distally on chromosome 2 contains the ebony locus that contributes to pigmentation differences (Wittkopp 2009). Thus, the recognition of geographically distinct populations on the basis of chromosomes and pigmentation is not entirely independent. The distinction between western and eastern populations is, however, not reflected in patterns of sequence

variation throughout the genome. The presence of a pigmentation cline in *D. americana* despite the homogenizing effects of gene flow suggests that pigmentation differences observed among collection sites are adaptive and the product of natural selection. Differences in relative humidity exist across the species range that might favor different pigmentation phenotypes in different locations; however, laboratory assays failed to show any significant difference in desiccation resistance between flies with light and dark pigmentation. Below, we compare these results to pigmentation clines in other Drosophila species.

The longitudinal pigmentation cline of D. americana *is atypical for Drosophila*

In *D. melanogaster* populations from multiple continents, thoracic pigmentation correlates with latitude: flies from higher latitudes have darker pigmentation (David *et al*, 1985; Parkash and Munjal, 1999; Parkash *et al*, 2008). Darker thoracic pigmentation is also characteristic of high altitude populations from India, which persist in an environment with lower relative humidity than low altitude populations (Parkash and Munjal, 1999; Parkash *et al*, 2008). A similar relationship between thoracic pigmentation and relative humidity was observed for seasonal pigmentation variation of *D. melanogaster* in montane regions of India (Parkash *et al*, 2009a). In sub-Saharan Africa, abdominal pigmentation of *D. melanogaster* correlates with latitude, but correlates even more strongly with altitude (Pool and Aquadro, 2007). *D. simulans*, a close relative of *D. melanogaster*, has much less variation for body color, yet still shows a weak correlation with latitude for thoracic pigmentation (Capy *et al*, 1988).

In the dunni species subgroup, a latitudinal cline exists for abdominal pigmentation that includes multiple species and extends from Puerto Rico through the Lesser Antilles islands in the Caribbean (Heed and Krishnamurthy, 1959; Hollocher *et al*, 2000). In contrast to studies of *D. melanogaster,* in which the darkest phenotypes are found at the highest latitudes, the darkest phenotypes in the dunni species group are found closest to the equator. Genetic analysis indicates that pigmentation differences among species in the dunni group are more likely to have been established by natural selection than through patterns of common ancestry among species (Hollocher *et al*, 2000).

Considering that all of the previously described pigmentation clines in Drosophila correlate with latitude, the absence of a latitudinal cline and the discovery of a longitudinal pigmentation cline in *D. americana* are surprising. In *D. melanogaster*, the correlation between pigmentation and latitude appears to be explained largely by differences in altitude; however, we found that this is unlikely to be the case for *D. americana.* Among the North American sites sampled for this work, altitude shows a similar correlation with both latitude (R^2 = 0.49) and longitude (R^2= 0.52), and in Dataset A, which contains more comprehensive sampling of variation within *D. americana* than Dataset B, pigmentation does not correlate significantly with latitude (Figure 4.4C). (Note that a significant correlation with altitude was observed for Dataset B, however.)

D. americana *shows unique relationships among pigmentation, relative humidity, and desiccation resistance*

Relative humidity (or aridity) is one of the most frequently invoked environmental correlates with pigmentation in Drosophila, and differences in desiccation resistance between light and dark pigmentation morphs have been reported for multiple species (reviewed in Rajpurohit *et al*, 2008; True, 2003; Wittkopp and Beldade, 2009). For example, in *D. melanogaster,* a laboratory assay showed that darker flies collected from natural populations survived longer under desiccating conditions (e.g., Parkash *et al*, 2008), with a similar pattern observed among seasonal morphs (Parkash *et al*, 2009a). *D. polymorpha*, a close relative of the dunni species group that does not show an obvious pigmentation cline, is enriched for darker phenotypes in warm, arid open areas in comparison to cooler, more humid covered forests (Brisson *et al*, 2005). These darker forms of *D. polymorpha* were found to survive longer than their lighter counterparts under desiccating conditions in the laboratory. Indeed, darker body pigmentation has been shown to increase desiccation resistance in *D. nepalensis, D. takahashii, D. ananassae, D. jambulina,* and *D. immigrans* (Parkash *et al*, 2008; Parkash *et al*, 2009b; Rajpurohit *et al*, 2008). This increase is desiccation resistance appears to be caused by a slower rate of water loss in individuals with greater melanization (Brisson *et al*, 2005; Rajpurohit *et al*, 2008).

In light of these data, the presence of darker *D. americana* in more humid areas is surprising and suggests that the primary selective force promoting the pigmentation cline in *D. americana* might be different from that in other species. It

is also possible that pigmentation has a different effect on desiccation resistance in different species. Consistent with this latter possibility, we observed no significant difference in desiccation resistance between light and dark forms of *D. americana*/*D. novamexicana* introgression lines. Our experiment used virtually the same assay for desiccation resistance as prior studies (i.e., survival time in a desiccating environment); however its design differed from prior work in two important ways. First, we compared defined genotypes derived from introgression lines rather than natural isolates or individuals from a segregating F_2 (or other recombinant) population. This allowed us to analyze flies that were genetically homogenous within a pigmentation class and differed for only a single region of the genome between pigmentation classes, which greatly reduces the possibility that genetic variation affecting traits other than pigmentation contributes to differences in desiccation resistance. Second, we tested the effects of *D. americana* pigmentation alleles on desiccation resistance in the genetic background of its sister species, *D. novamexicana*. The *D. novamexicana* background and/or interactions between the two different species alleles might have altered the relationship between pigmentation and desiccation resistance; however, we see no reason to suspect that this is the case. The light pigmentation of *D. novamexicana* appears to be an extension of the *D. americana* longitudinal gradient (Throckmorton, 1982) and the two species retain many shared ancestral polymorphisms (Hilton and Hey, 1996; Morales-Hojas *et al*, 2008; Wittkopp *et al*, 2009).

In summary, we conclude that our data do not support the hypothesis that differences in relative humidity among collection sites cause selection for differences in desiccation resistance that are mediated by differences in body color. That said, our data are also insufficient to disprove such a hypothesis. The standard laboratory assay for desiccation tolerance is extremely crude: variation in relative humidity among wild populations is much less extreme than the difference between high and low humidity chambers set up in the laboratory and the phenotypes affected by humidity levels in the wild are likely much more subtle than death. Parkash *et al.* (2009a) have recently shown that desiccation stress alters mate choice, copulation duration and fecundity of *D. jambulina,* and we suspect that desiccation effects such as these have a much larger impact on fitness in the wild.

Acknowledgements

We thank Emma Stewart, Belinda Haerum, Adam Neidert, Gizem Kalay and Monica Woll for technical assistance; Paulina Mena and the Tucson Drosophila Species Stock Center for providing strains of *D. americana* and rearing advice; officials of the US Fish and Wildlife Service for permission to collect within the National Wildlife Refuge system; Lacey Knowles, Jonathan Gruber, Gizem Kalay, Joseph Coolon, and anonymous reviewers for helpful comments on the manuscript; and the Margaret and Herman Sokol Endowment and National Science Foundation (DEB-0640485, DEB-0420399) for funding. P.J.W. is an Alfred P. Sloan Research Fellow.

References

Brakefield PM, Willmer PG (1985). The basis of thermal melanism in the ladybird *Adalia Bipunctata*: differences in reflectance and thermal properties between the morphs. *Heredity* **54**(FEB)**:** 9-14.

Brisson JA, De Toni DC, Duncan I, Templeton AR (2005). Abdominal pigmentation variation in *Drosophila polymorpha*: geographic variation in the trait, and underlying phylogeography. *Evolution* **59**(5)**:** 1046-1059.

Caletka BC, McAllister BF (2004). A genealogical view of chromosomal evolution and species delimitation in the *Drosophila virilis* species subgroup. *Mol Phylogenet Evol* **33**(3)**:** 664-670.

Capy P, David JR, Robertson A (1988). Thoracic trident pigmentation in natural populations of *Drosophila simulans*: a comparison with *D. melanogaster*. *Heredity* **61**(2)**:** 263-268.

David JR, Capy P, Payant V, Tsakas S (1985). Thoracic trident pigmentation in *Drosophila melanogaster*: differentiation of geographical populations. *Génét Sél Evol* **17**(2)**:** 211-223.

David JR, Gibert P, Petavy G, Moreteau B (2002). Variable modes of inheritance of morphometrical traits in hybrids between Drosophila melanogaster and Drosophila simulans. *Proc Biol Sci* **269**(1487)**:** 127-135.

Drysdale R, Consortium F (2008). FlyBase : a database for the Drosophila research community. *Methods Mol Biol* **420:** 45-59.

Excoffier, L. G. Laval, and S. Schneider (2005) Arlequin ver. 3.0: An integrated software package for population genetics data analysis. *Evolutionary Bioinformatics Online* **1**:47-50.

Fu YX, Li WH (1993). Statistical tests of neutrality of mutations. *Genetics* **133**(3)**:** 693- 709.

Gay, L., P. Crochet, D. Bell, and T. Lenormand. 2008. Comparing genetic and phenotypic clines in hybrid zones: a window on tension zone models. *Evolution* **62**: 2789-2806

Gibert JM, Peronnet F, Schlötterer C (2007). Phenotypic plasticity in Drosophila pigmentation caused by temperature sensitivity of a chromatin regulator network. *PLoS Genet* **3**(2)**:** e30.

Gloor GB, Preston CR, Johnson-Schlitz DM, Nassif NA, Phillis RW, Benz WK *et al* (1993). Type I repressors of P element mobility. *Genetics* **135** (1)**:** 81-95.

Haddrill PR, Thornton KR, Charlesworth B, Andolfatto P (2005). Multilocus patterns of nucleotide variability and the demographic and selection history of *Drosophila melanogaster* populations. *Genome Res* **15**(6)**:** 790-799.

Heed WB, Krishnamurthy NB (1959). Genetic studies on the *cardini* group of Drosophila in the West Indies. *Univ Texas Publ* **5914**(155-179).

Hilton H, Hey J (1996). DNA sequence variation at the period locus reveals the history of species and speciation events in the *Drosophila virilis* group. *Genetics* **144**(3)**:** 1015-1025.

Hollocher H, Hatcher JL, Dyreson EG (2000). Evolution of abdominal pigmentation differences across species in the *Drosophila dunni* subgroup. *Evolution* **54**(6)**:** 2046-2056.

Hsu TC (1951). Chromosomal variation and evolution in the virilis group of Drosophila. *Thesis*, University of Texas, Austin.

Hsu, T. C. (1952). Chromosomal variation and evolution in the *virilis* group of Drosophila. *Univ Texas Publ* **5204:** 35-72.

Jablonski NG, Chaplin G (2000). The evolution of human skin coloration. *Journal of Human Evolution* **39**(1)**:** 57-106.

Jensen JL, Bohonak AJ, Kelley ST (2005). Isolation by distance, web service. *BMC Genet* **6**(1)**:** 13.

Kalmus H (1941). The resistance to desiccation of Drosophila mutants affecting body colour. *Proc Biol Sci* **130**(859)**:** 185-201.

Kawecki TJ, Ebert D (2004). Conceptual issues in local adaptation. *Ecology Letters* **7:** 1225-1241.

Kumar S, Tamura K, Nei M (2004). MEGA3: Integrated software for Molecular Evolutionary Genetics Analysis and sequence alignment. *Brief Bioinformatics* **5**(2)**:** 150-163.

Librado, P. and Rozas, J. 2009. DnaSP v5: A software for comprehensive analysis of DNA polymorphism data. *Bioinformatics* **25**: 1451-1452

Maherali H, Williams BL, Paige KN, Delucia EH (2002). Hydraulic differentiation of Ponderosa pine populations along a climate gradient is not associated with ecotypic divergence. *Funct Ecol* **16:** 510-521.

Maside X, Lee AW, Charlesworth B (2004). Selection on codon usage in *Drosophila americana*. *Curr Biol* **14**(2)**:** 150-154.

Matzkin LM, Watts TD, Markow TA (2007). Desiccation resistance in four Drosophila species: sex and population effects. *Fly (Austin)* **1**(5)**:** 268-273.

McAllister BF (2001). Genetic analysis of sex-chromosome arrangement in *Drosophila americana*: a laboratory exercise for undergraduate or advanced placement students. *Dros Inf Serv* **84:** 227-234.

McAllister BF (2002). Chromosomal and allelic variation in *Drosophila americana*: selective maintenance of a chromosomal cline. *Genome* **45**(1)**:** 13-21.

McAllister BF (2003). Sequence differentiation associated with an inversion on the neo- X chromosome of *Drosophila americana*. *Genetics* **165**(3)**:** 1317-1328.

McAllister BF, Evans AL (2006). Increased nucleotide diversity with transient Y linkage in *Drosophila americana*. *PLoS ONE* **1:** e112.

Morales-Hojas R, Vieira CP, Vieira J (2008). Inferring the evolutionary history of *Drosophila americana* and *Drosophila novamexicana* using a multilocus approach and the influence of chromosomal rearrangements in single gene analyses. *Mol Ecol* **17**(12)**:** 2910-2926.

Mullen LM, Hoekstra HE (2008). Natural selection along an environmental gradient: a classic cline in mouse pigmentation. *Evolution* **62**(7)**:** 1555-1570.

Parkash R, Munjal AK (1999). Phenotypic variability of thoracic pigmentation in Indian populations of *Drosophila melanogaster*. *J Zool Syst Evol Research* **37:** 133-140.

Parkash R, Rajpurohit S, Ramniwas S (2008). Changes in body melanisation and desiccation resistance in highland vs. lowland populations of *D. melanogaster*. *Journal of Insect Physiology* **54**(6)**:** 1050-1056.

Parkash R, Sharma V, Kalra B (2009a). Impact of body melanisation on desiccation resistance in montane populations of D. melanogaster: Analysis of seasonal variation. *J Insect Physiol* **55**(10)**:** 898-908.

Parkash R, Singh S, Ramniwas S (2009b). Seasonal changes in humidity level in the tropics impact body color polymorphism and desiccation resistance in Drosophila jambulina-Evidence for melanism-desiccation hypothesis. *J Insect Physiol* **55**(4)**:** 358-368.

Phifer-Rixey M, Heckman M, Trussell GC, Schmidt PS (2008). Maintenance of clinal variation for shell colour phenotype in the flat periwinkle *Littorina obtusata*. *J Evol Biol* **21**(4)**:** 966-978.

Pool JE, Aquadro CF (2007). The genetic basis of adaptive pigmentation variation in *Drosophila melanogaster*. *Mol Ecol* **16**(14)**:** 2844-2851.

Rajpurohit S, Parkash R, Ramniwas S (2008). Body melanization and its adaptive role in thermoregulation and tolerance against desiccating conditions in drosophilids. *Entomol Research* **38:** 49-60.

Schäfer MA, Orsini L, McAllister BF, Schlötterer C (2006). Patterns of microsatellite variation through a transition zone of a chromosomal cline in Drosophila americana. *Heredity* **97**(4)**:** 291-295.

Stephens, M. and P. Donnelly (2003). A comparison of bayesian methods for haplotype reconstruction. *American Journal of Human Genetics* **73**, 1162–1169.

Tajima F (1989). Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics* **123**(3)**:** 585-595.

Tamura K, Nei M (1993). Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpanzees. *Mol Biol Evol* **10**(3)**:** 512-526.

Throckmorton LH (1982). The *virilis* species group. In: Ashburner M, Carson HL and Thompson JN (eds) *The Genetics and Biology of Drosophila*. Academic Press: London. Vol. 3b, pp 227-296.

True JR (2003). Insect melanism: the molecules matter. *Trends in Ecol and Evol* **18**(12)**:** 640-647.

Uy JAC, Endler JA (2004). Modification of visual background increases the conspicuousness of golden-collared manakin displays. *Behav Ecol* **15:** 1003- 1010.

Vieira CP, Coelho PA, Vieira J (2003). Inferences on the evolutionary history of the *Drosophila americana* polymorphic X/4 fusion from patterns of polymorphism at the X-linked *paralytic* and *elav* genes. *Genetics* **164** (4)**:** 1459-1469.

Watt WB (1969). Adaptive significance of pigment polymorphisms in Colias butterflies, II. Thermoregulation and photoperiodically controlled melanin variation in *Colias eurytheme*. *Proc Natl Acad Sci USA* **63**(3)**:** 767-774.

Wittkopp PJ, Beldade P (2009). Development and evolution of insect pigmentation: genetic mechanisms and the potential consequences of pleiotropy. *Semin Cell Dev Biol* **20**(1)**:** 65-71.

Wittkopp PJ, Carroll SB, Kopp A (2003a). Evolution in black and white: genetic control of pigment patterns in Drosophila. *Trends Genet* **19**(9)**:** 495-504.

Wittkopp PJ, Stewart EE, Arnold LL, Neidert AH, Haerum BK, Thompson EM *et al* (2009). Intraspecific polymorphism to interspecific divergence: genetics of pigmentation in Drosophila. *Science* **32** (5952)**:** 540-544.

Wittkopp PJ, Williams BL, Selegue JE, Carroll SB (2003b). Drosophila pigmentation evolution: divergent genotypes underlying convergent phenotypes. *PNAS* **100**(4)**:** 1808-1813.

Chapter 5

Genetic heterogeneity underlying pigmentation variation within *D. americana*

Abstract

Within *D. americana,* populations display variation in pigmentation longitudinally that appears to be a consequence of local adaptation. Prior work has shown that two of the lighter pigmented *D. americana* populations contain alleles of the pigmentations genes, *tan* and *ebony,* that are functionally equivalent to its lighter sister species, *D. novamexicana.* This research examines the geographic distribution and frequency of these *D. novamexicana*-like *tan* and *ebony* alleles among 51 different *D. americana* isofemale lines collected from 20 different locations. From this research, no other *D. americana* populations were found to contain *D. novamexicana*-like *tan* or *ebony* alleles indicating that these alleles are not widespread among *D. americana.*

Introduction

Drosophila americana is found throughout the United States from the Atlantic coast to just east of the Rocky Mountains and exhibits a pigmentation cline that varies longitudinally with eastern populations displaying darker pigmentation than more westerly populations (WITTKOPP *et al.* 2011). This

variation in pigmentation has also previously been determined to be a result of local adaptation (WITTKOPP *et al.* 2011). It could be assumed that this is due an underlying cline in allele frequencies, however, this may not be the case in this instance.

The pigmentation genes, *ebony* and *tan* have been shown to contribute to this pigmentation cline in *D. americana* (WITTKOPP *et al.* 2009). These same genes were found to contribute to the pigmentation difference displayed between *D. americana* and its much lighter sister species, *D. novamexicana* (WITTKOPP *et al.* 2009). Among *D. americana,* two populations were discovered to have *D. novamexicana*-like alleles of either *ebony* or *tan*: a Montana population, A01, had a *D. novameicana*-like *tan* allele and a Nebraska population, DN2, had a *D. novamexicana*-like *ebony* allele (WITTKOPP *et al.* 2009). Additional strains collected from the same Nebraska population were also analyzed, but did not share the *D. novamexicana*-like *ebony* allele indicating genetic heterogeneity is present among *D. americana.* Genetic heterogeneity refers to the phenomena in which multiple genes or alleles can be used to create similar phenotypes.

As mentioned above, the pigmentation cline exhibited in *D. americana* appears to be a result of local adaptation. Local adaptation occurs when natural selection favors different genotypes in different locations, thereby also creating an underlying genetic cline. The previously reported genetic heterogeneity in *D. americana* suggests that such an underlying cline in allele frequencies may not be present among *D. americana* populations. Here, we investigate the genetic diversity impacting pigmentation within *D. americana* by examining the

geographic distribution and frequency of *D. novamexicana*-like *ebony* and *tan* alleles.

Materials and Methods

Fly strains

A total of 51 isofemale lines (or strains) of *D. americana* were used in the analysis. For each line, a single female was collected from the wild and allowed to lay eggs. The resulting progeny were then used to create each line. Lines were collected from 20 locations within the natural population range of *D. americana* in the United States from Ohio to Nebraska (Figure 5.1 and Table 5.1). Multiple lines were used from most locations as genetic heterogeneity has been identified for pigmentation alleles among one *D. americana* population in Duncan, Nebraska (WITTKOPP *et al.* 2009). The most eastern location was Killbuck, Ohio (81.967^0 W, 40.682^0 N), the most western and northern location was Niobrara, Nebraska (98^0 2.58' W, 42^0 44.94' N) and the southern most collection site was Sneads, Florida (84.93^0 W, 30.71^0 N).

Male backcross flies (BC1) were used for all genotyping and phenotyping analyses (Figure 5.2). For each line, *D. americana* virgin females were crossed with *D. novamexicana* males to create F_1 hybrids. F_1 progeny collected as virgins were then backcrossed with *D. novamexicana* males to create flies with two potential genotypes at each locus (Figure 5.2). BC1 male flies, which were used in subsequent genotyping assays, were either homozygous for the *D. novamexicana ebony* allele or heterozygous for the *D. americana* and *D.*

novamexicana ebony alleles and hemizygous for either the *D. americana* or *D. novamexicana* allele for both *tan* and *yellow* on the X chromosome.

Table 5.1: *D.americana* populations and locations

Location	Name	Longitude	Latitude
Niobrara, NE	NN0728	98° 2.58' W	42° 44.94' N
Duncan, NE	DN0748x37	97° 29.68' W	41° 22.11' N
Squaw Creek, MO	SC0718	95.25° W	40.12° N
	SC0708		
Corney Bayou, Kisatchie National Forest, LA	CB0522	92.85	32.98
	CB0512		
	CB0504		
	CB0506		
	CB 05.14		
Iowa River, IA	IR0436	91° 42.91' W	41° 46.76' N
	IR0496		
	IR04110		
	IR 04.118		
	IR 04.32		
Floodgate Park, AR	FP9946	91° 4.5' W	34° 11.5' N
	FP9956		
Cat Island Natl Wildlife Refuge, LA	CI0502	91.42° W	30.80° N
Puxico, MO	PM9914	90° 8.5' W	36° 58.3' N
	PM9924		
	PM9936		
Lake Ashbaugh, AR	LA9954	90° 44.7' W	36° 16.0' N
Howell Island, MO	HI9946	90° 40.7' W	38° 39.7' N
	HI9950		
Lake Arkabutla, MS	LR0532	89° 59.17' W	34° 42.89' N
	LR0518		
Duck Island, IL	DI0514	89° 56.69' W	40° 27.26' N
	DI0562		
Brewer Basin, TN	BB0546	89° 18.61' W	36° 27.71' N
LaSalle, IL	II0702	89.09° W	41.33° N
	II0710		
West of Gary, IN	G9647	87° 22.0' W	41° 33.0' N
North of Demopolis, AL	DA0610	87.84° W	32.52° N
	DA0626		
	DA 06.20		
Muscatatuck, IN	MK0738	85.80° W	38.91° N
	MK0724		
	MK0736		
Eufaula Natl Wildlife Refuge, AL	BU0616	85.09° W	32.03° N
Sneads, FL	FG0618	84.93° W	30.71° N
	FG0622		
	FG0624		
	FG0610		
Killbuck, OH	WS0712	81.967° W	40.682° N

Figure 5.1: *D. americana* displays pigmentation variation in a longitudinal cline

Figure 5.1: The geographical distribution of both *D. americana* and *D. novamexicana. D. americana* populations shown are a subset of the lines used in this study representing a given region (natural populations of *D. americana* extend as far east as Virginia or Maine, but were not analyzed here). Previously examined *D. americana* strains, DN2 and A01, that have exhibited *D. novamexicana*-like *ebony* and *tan* alleles, respectively, are also shown. The location of *D. novamexicana* (N14) populations is also shown.

Figure 5.2: Crossing scheme to create backcross progeny

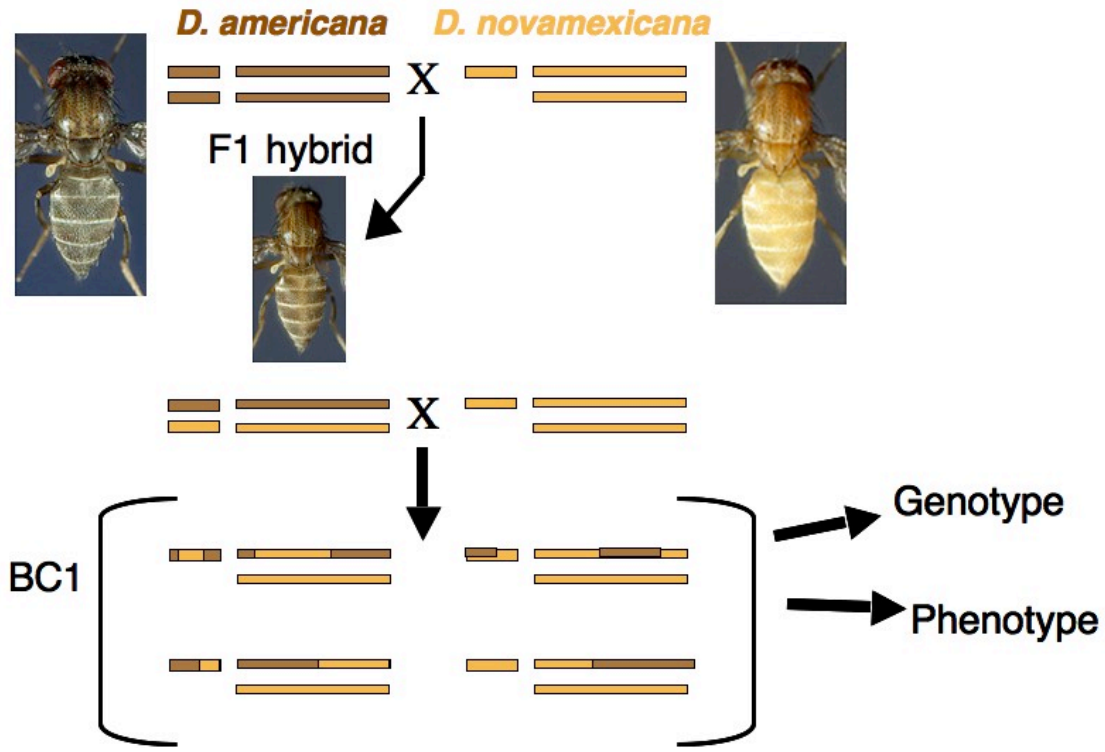


Figure 5.2: *D. americana* virgin females from each line were crossed to *D. novamexicana* males to create F1 hybrids displaying an intermediate phenotype. Female F1 hybrids were heterozygous at all loci. Virgin female F1 hybrids were backcrossed to wild type *D. novamexicana* males to create backcross flies (BC1) that displayed a range of phenotypes and genotypes. Male BC1 flies were used for further phenotyping and genotyping analyses. Brown rectangles represent *D. americana* sequence while yellow rectangles represent *D. novamexicana* sequence. The short rectangle on the left represents the X chromosome, while the large rectangle on the right represents the remaining five autosomal chromosomes grouped together.

Flies were grown and mated on a standard yeast-glucose media and incubated at 20^0 C. All parental, F_1, and BC1 flies were reared under the same conditions to control for external environmental factors that could impact the flies' phenotype. BC1 progeny were collected within 3 days of eclosion and aged 7-10 days to ensure pigmentation was fully developed.

Phenotyping backcross flies

For each line, flies were first sorted by sex, then into phenotypic classes based upon abdominal cuticle pigmentation. Backcross pigmentation was not continuously distributed, and distinct phenotypes were grouped as separate phenotypic classes. Each class was determined by visually apparent pigmentation differences among backcross flies. There was no set number of classes among lines; the number of pigmentation classes ranged from four to eight among the lines surveyed. All fly lines had a minimum of four phenotypic classes. The lightest class was always categorized as "1" with subsequent class numbers corresponding to progressively darkening pigmentation. (i.e. In a line with four total classes, "4" would be classified as the darkest flies, whereas a line with seven total classes, "4" would be classified as a mid-range pigmentation.) Pigmentation classes for each line were confirmed by independent observations from two researchers. For each line examined, one representative male fly from each phenotypic class was imaged as a visual reference (Figure 5.3) using a Scion Visicapture 1.2 and Scion Corporation Model CFW-1308C color digital camera. Images were compiled using Photoshop CS6 (Adobe, San Jose, CA)

and uniformly adjusted using a set of standards also imaged at the same time as the flies (to control for day-to-day variation) to best represent visual observations. All flies were stored at -80^0 C prior to DNA extraction.

Figure 5.3: Phenotypes of backcross (BC1) male progeny

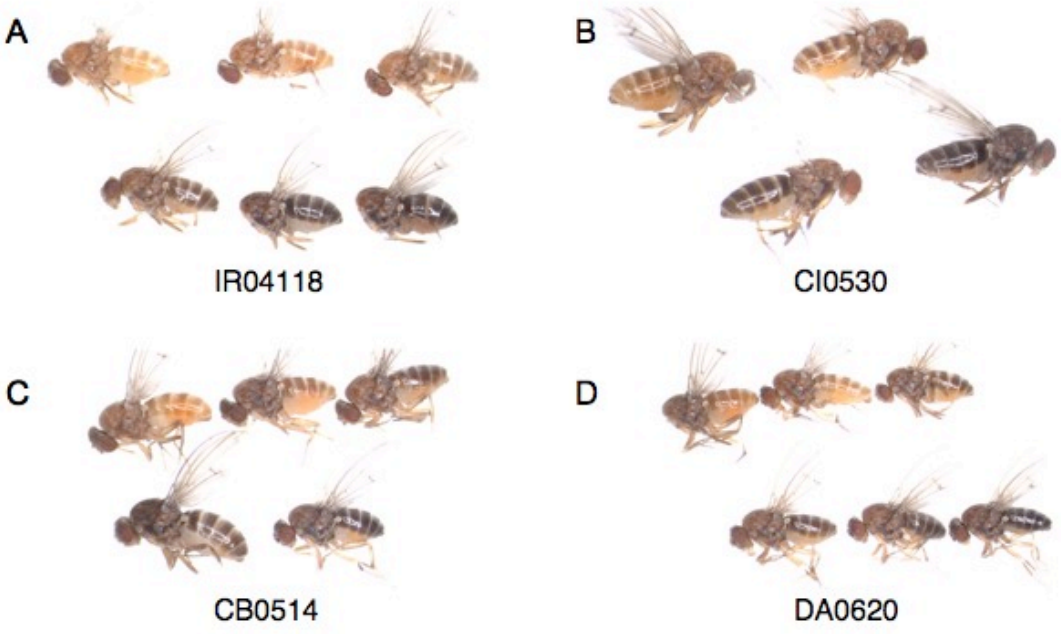


Figure 5.3: A sample of male backcross progeny are shown from four different *D. americana* strains (indicated below each group). Phenotypic class 1 is always the lightest class and arranged in the upper left corner of each grouping. The darkest phenotypic class is always represented in the lower right corner and intermediate classes are arranged from left to right (then top to bottom) from class 1 to the darkest class. A range of 4-8 phenotypic classes was observed among all *D. americana* strains surveyed. Some flies may not have heads or legs due to handling after being frozen at -80^{o}C. (A) Six distinct phenotypic classes were identified for backcross progeny for line IR04118 (from Iowa River, IA). (B) Four phenotypic classes identified for CI0530 (from Cat Island National Wildlife Refuge, LA) (C) Five different phenotypic classes were observed for CB0514

(from Kisatchie National Forest, LA) and (D) six classes were found for DA0620 (from North of Demopolis, AL).

DNA extractions

All phenotyped male BC1 flies were removed from -80^0 C storage and prepped for genotyping by DNA extraction using a modified method described in Gloor *et al.* (1993) for more efficient extraction of large numbers of flies. Extractions were carried out in 96-well plates (GeneMate# T3031-21). Each well contained a single fly, a glass bead and 50ul of a 1:99 Proteinase K/Engel's Buffer solution. Plates were sealed and shaken in a Qiagen Retsch MM301 Tissue Lyser until the glass beads had pulverized the flies in their corresponding wells. Lysed flies were then incubated at 37^0 C for 30 minutes, then 95^0 C for 2 minutes for Proteinase K inactivation. Extracted DNA was stored at 4^0 C until use for genotyping.

Genotyping

Length variations in the *D. americana* and *D. novamexicana* alleles of *tan* and *yellow* were used to visually genotype BC1 flies with PCR. For *tan,* primers included a forward primer with the sequence, 5'-CGAGTTTTTATTCCCACTGAATTAT-3' and a reverse primer, 5'-GGGTTCGTCTTATCCACGAT-3' were used to produce a 100bp product for the *D. americana tan* allele and a 64bp product for the *D. novamexciana tan* allele. For *yellow,* depending on which *D. americana* line was being genotyped, one of two forward primers was used: either *yellow* forward one, 5'-

CCAAAAGGACAACCGAGTTT-3' or *yellow* forward two, 5'-CTAAACATGCCTGAAAATCAATCACGGA-3'. Both *yellow* forward primers were used with the *yellow* reverse primer 5'-AGTCGA TTGCCAAAGTGCTC-3'. The *yellow* forward one primer paired with the *yellow* reverse primer generated a 349bp product for the *D. americana yellow* allele and a 372bp product for the *D. novamexicana yellow* allele. For six lines (IR0436, LR0540, FP9946, DI0562 MK0738, and SC0708), the *yellow* forward one primer and *yellow* reverse primer did not produce any visible length variation between the *D. americana* and *D. novamexicana* alleles. For these specific lines, the *yellow* forward two primer and *yellow* reverse primer were used alternatively, and then followed by a DraI digest, which cut only the *D. novamexicana yellow* allele, allowing for a visually detectable length variation. All PCRs and digests were run on 2% agarose gels and visualized using Ethidium Bromide.

Genotyping flies at the *ebony* locus was performed by pyrosequencing (AHMADIAN 2000). Genotyping *ebony* by PCR was not possible since PCR product length variations between species' alleles were too difficult to visualize. Additionally, restriction enzyme digest was not possible for *ebony* due to the lack of allele-specific cut sites. An initial PCR was carried out using the forward primer, 5'-AGCCCGAGGTGGACATCA-3', and the reverse primer, 5'-*GTATGGGTCCCTCGCAGAA-3' (* notates biotinylation). PCR products were prepped and pyrosequencing performed as described in (WITTKOPP *et al.* 2008). Samples were analyzed using the following sequencing primer, 5'-CGAGGTGGACA TCAAGT-3'. The *ebony* assay contained two SNPs that were

used to differentiate between the *D. americana* and *D. novamexicana ebony* alleles. In the sequence to analyze, the *D. americana* allele would be 5'-C**C**AAGCT**G**CT-3' and the *D. novamexicana* allele would be 5'-C**G**AAGCT**T**CT-3'.

Statistics and calculations

Allelic frequencies of *tan, ebony* and *yellow* were recorded in each phenotypic class for each BC1 male fly. The data was imported into the statistical software package, R (version 2.13.1). Distributions of the alleles for each gene were analyzed using permutation testing. This allowed the data to be compared to a null model (that *D. americana* and *D. novamexicana* alleles have equal effects on pigmentation). To account for multiple testing, FDR values were calculated as described in (PIKE 2011). Adjusted p-values (or q-values) are reported. This analysis provided a model to assess the impact of *tan, ebony,* and *yellow* on pigmentation within each *D. americana* line surveyed. A 5% significance level was used.

Results

Phenotypic classes varied among D. americana *lines*

Backcross progeny pigmentation was not continuously distributed but rather was categorized into discrete classes based on abdominal pigmentation. The number of classes varied across lines (ranging from four-eight classes). This is consistent with previous findings for these species (WITTKOPP *et al.* 2003;

2009). Pigmentation scores for each class were not standardized across lines, meaning that a class one fly from one line may not have the same phenotype as another class one fly from another line and so on. For each line, a representative fly from each class was imaged to compare pigmentation variation across lines (for an example, see Figure 5.3).

Given the range of phenotypic classes observed (Figure 5.4), these data indicate genetic heterogeneity underlying pigmentation in this species. For the *D. americana* strains examined here, the phenotypic class number ranged from four to eight distinct phenotypes. Even strains from the same geographical location varied in the number of phenotypic classes. One of the more striking examples comes from the MK- *D. americana* lines (collected from Muscatatuck, IN), which had four, six and seven phenotypic classes for the three separate MK strains. Flies collected from the same site exhibit similar pigmentations; therefore, the varying number of phenotypic classes in BC1 progeny indicate different genetic mechanisms are probably at play to construct these similar phenotypes. Different allele combinations at *ebony, tan* and *yellow* could potentially explain the number of phenotypic classes identified. All phenotyped backcross flies were genotyped for each loci to determine the distribution of *ebony, tan* and *yellow* alleles among the phenotypic classes (see Materials and Methods for details).

Figure 5.4: Phenotype frequencies among locations

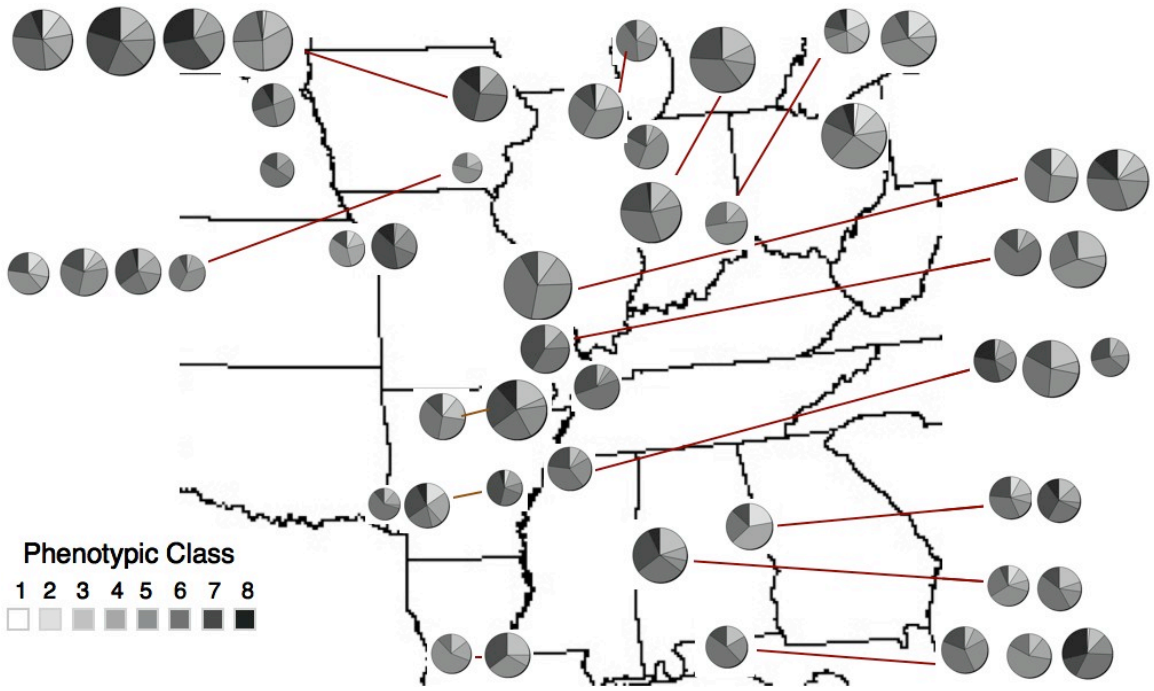


Figure 5.4: Frequencies of phenotypic classes by geographical location. Pie charts represent the number and frequency of phenotypic classes observed for each *D. americana* line. Red lines link piecharts from the same location. Strains within a given location did not necessarily display a similar number of phenotypic classes or frequencies. Class one is represented by white and as phenotypic class increases, shades of gray darken to black. (Representative images of phenotypic classes for each line were standardized for this analysis in that each phenotypic class is equivalent across strains.) The overall size of the piechart represents the total number of BC1 progeny assessed (i.e. the larger the size of the pie chart, the larger the total sample of BC1 flies from that *D. americana* line). The sample size of BC1 progeny assessed for each line ranged from n=119 to n=27.

Alleles identified in D. americana *populations*

Backcross males were genotyped at *tan, ebony* and *yellow* to examine allelic frequencies at these genes. Previous studies have identified *tan* and *ebony* as contributing to the pigmentation difference between *D. americana* and *D. novamexicana* (WITTKOPP *et al.* 2009). From this data, we would predict that *D. novamexicana* alleles would be at a higher frequency in lighter phenotypic classes for *tan* and *ebony,* whereas *D. americana* alleles would be at a higher frequency in the darker phenotypic classes for these genes. Consequently, we would expect a statistically significant relationship between genotype and pigmentation for *tan* and *ebony* ($p \leq 0.05$). As *yellow* has not been found to be associated with pigmentation divergence in these species (WITTKOPP *et al.* 2003; 2009), it is expected that the *D. americana* and *D. novamexicana yellow* alleles will be randomly distributed among phenotypic classes and have no relationship between genotype and pigmentation ($p \geq 0.05$).

In order to identify functionally distinct alleles among *D. americana* strains, *D. novamexicana* served as common reference. *D. americana* and *D. novamexicana* alleles of *tan*, *ebony* and *yellow* were tested for equal effects on pigmentation for each *D. americana* strain. If found, this result would indicate that the *D. americana* allele was functionally equivalent to the *D. novamexicana* allele for that gene. Although this might be unexpected, it has previously been reported for two *D. americana* populations that had either a *D. novamexicana*-like *ebony* or *tan* allele (WITTKOPP *et al.* 2009). Therefore, *D. americana* lines that exhibited *tan* and *ebony* alleles functionally equivalent to *D. novamexicana* (statistically not

significant, $p \geq 0.05$), would be of most interest since these alleles would suggest additional genetic heterogeneity among *D. americana*.

Among the fifty-one *D. americana* lines assessed, none were identified to have functionally equivalent *tan* or *ebony* alleles to *D. novamexicana* after a correction for multiple testing (Table 5.2). All the lines showed a significant association of *tan* and *ebony* alleles with pigmentation (Figure 5.5). Additionally, as previously reported, *yellow* is shown to not be significantly associated with the pigmentation difference between *D. americana* and *D. novamexicana.* These data indicate that the previously identified *D. americana* populations that have *D. novamexicana*-like *tan* and *ebony* alleles (WITTKOPP *et al.* 2009) are unique among *D. americana* populations.

Figure 5.5: *ebony* and *tan* alleles contribute to interspecies divergence

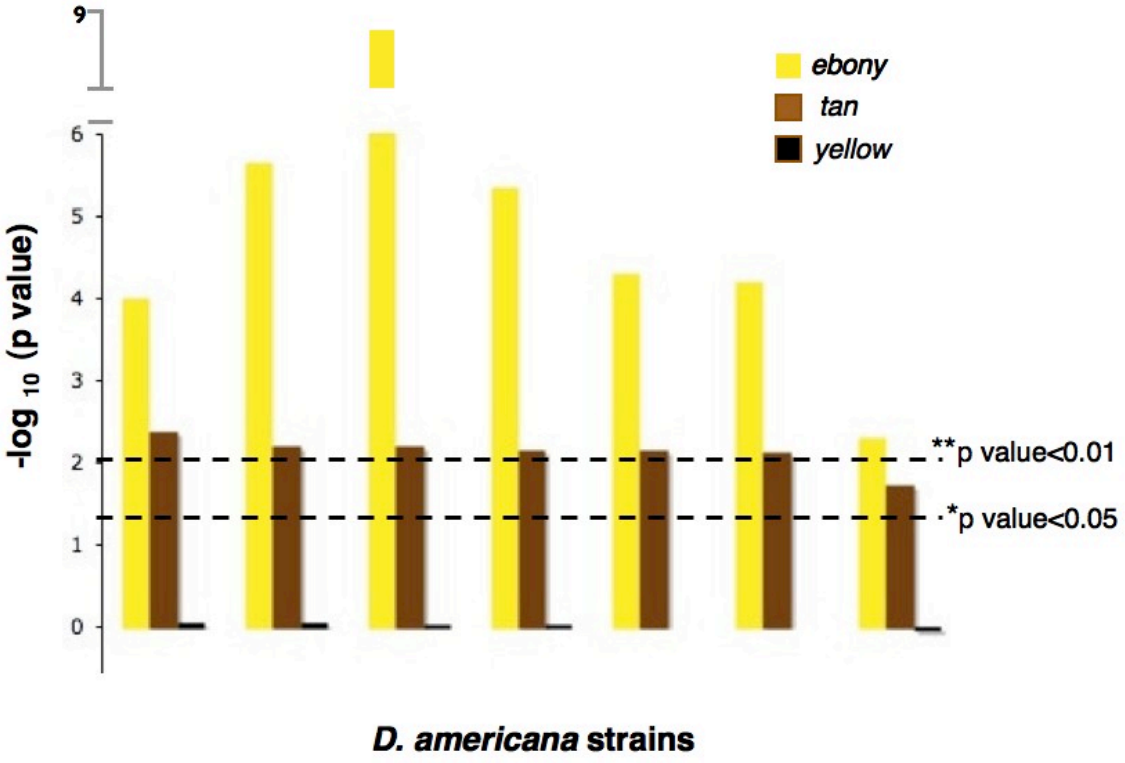


Figure 5.5: Association of *D. americana* genotypes with phenotype. *ebony* and *tan* were found to contribute to the pigmentation differences between each *D. americana* strain and *D. novamexicana.* Significance values are plotted on the y-axis with thresholds indicated for < 0.01 and < 0.05 by dotted lines. Seven representative *D. americana* strains are displayed. These results are consistent with all 51 *D. americana* strains surveyed.

Table 5.2: Summary of statistical results of allelic frequencies

	ebony	*tan*	*yellow*
Line Name	**FDR-adjusted p values (q values)**	**FDR-adjusted p values (q values)**	**FDR-adjusted p values (q values)**
IR0496	< 0.0001	< 0.0001	0.2058
IR04118	< 0.0001	< 0.0001	0.2058
IR04110	< 0.0001	< 0.0001	0.4559
IR0436	< 0.0001	< 0.0001	0.4559
IR 04.32	< 0.0001	< 0.0001	0.4559
CI0530	< 0.0001	< 0.0001	0.4559
CI0502	< 0.0001	< 0.0001	0.4559
PM9924	< 0.0001	< 0.0001	0.4559
PM9936	< 0.0001	< 0.0001	0.4559
PM9914	< 0.0001	< 0.0001	0.4559
LR0538	< 0.0001	< 0.0001	NA
LR0518	< 0.0001	< 0.0001	0.4559
LR0540	< 0.0001	< 0.0001	NA
LR0532	< 0.0001	< 0.0001	0.4559
BU0624	< 0.0001	< 0.0001	0.4559
BU0610	< 0.0001	< 0.0001	0.5951
BU0616	< 0.0001	< 0.0001	0.5951
BB0546	< 0.0001	< 0.0001	0.6001
G9647	< 0.0001	< 0.0001	0.6001
LA9954	< 0.0001	< 0.0001	0.6749
LA9946	< 0.0001	< 0.0001	0.6935
HI9946	< 0.0001	< 0.0001	0.8196
HI9912	< 0.0001	< 0.0001	0.8635
HI9950	< 0.0001	< 0.0001	0.8635
FP9946	< 0.0001	< 0.0001	0.8635
FP9956	< 0.0001	7.27E-05	0.8635
FP9918	< 0.0001	9.33E-05	0.8635
CB0506	< 0.0001	0.0001	0.8635
CB0522	< 0.0001	0.0001	0.8635
CB0512	< 0.0001	0.0001	0.8635
CB0504	< 0.0001	0.0001	0.8635
CB0514	< 0.0001	0.0001	0.8635
DA0610	< 0.0001	0.0002	0.8635
DA0620	< 0.0001	0.0002	0.8635
DA0626	< 0.0001	0.0002	0.8635
FG0618	< 0.0001	0.0002	0.8635
FG0622	< 0.0001	0.0002	0.8648
FG0610	< 0.0001	0.0002	0.8966
FG0624	< 0.0001	0.0004	0.8966
DI0562	< 0.0001	0.0004	0.8966
DI0514	< 0.0001	0.0006	0.8966
DN0748x37	< 0.0001	0.0008	0.8966
II0702	< 0.0001	0.0008	0.8966
II0710	< 0.0001	0.0037	0.8966
MK0724	< 0.0001	0.0041	0.8966
MK0736	2.28E-06	0.0064	0.8966
MK0738	4.38E-09	0.0064	0.9433
NN0728	4.38E-06	0.0069	0.9433
SC0708	5.14E-05	0.0071	0.9917
SC0718	6.30E-05	0.0073	1.003
WS0712	0.0051	0.0193	1.035

Table 5.2: NA's in the *yellow* column were due to unsuccessful genotyping and lack of DNA for subsequent assays.

Discussion

From this study, further evidence of genetic heterogeneity was indicated among *D. americana* strains. Although we were unable to identify any additional *D. americana* populations containing *D. novamexicana*-like alleles of either *ebony* or *tan,* the range of phenotypic classes observed in *D. americana* strains both within and among geographical locations suggests genetic heterogeneity among these populations. Given that strains collected from the same location have similar phenotypes, yet display a varying number of phenotypic classes, indicates that different genes/alleles are at work to create comparable color morphs.

Genetic heterogeneity is frequently found among humans in clinical studies related to a specific disorder or disease (MEDINA-GOMEZ *et al.* 2012; ZHOU *et al.* 2012; PASTOR 2012), and has also been identified among plants (MA *et al.* 2010). The presence of genetic heterogeneity can lead to difficulties in identifying the genetic basis of phenotypic traits. Genome-wide scans are often employed to determine loci under selection for a local adaptation and work by isolating genomic regions (or linked regions) that are not consistent with neutral evolution, suggesting that region is undergoing natural selection (NOSIL *et al.* 2009). Ma *et. al* discuss that they were able to identify four SNPs associated with clinal variation in *Populus tremula* using a candidate gene approach, but were unable to identify these same SNPs using a genomic scan potentially as a result from the increased observed heterogeneity (MA *et al.* 2010). Methodologies such as genome-wide association studies (GWAS), seem to be

falling short in their capabilities to detect relevant loci for phenotypic traits. Underlying genetic heterogeneity, such as that observed for pigmentation in *D. americana*, can contribute to these difficulties. The research presented here suggests that this phenomena of genetic heterogeneity may be more common among natural populations than believed and alternative methods are needed to be able to detect what is being missed with current technologies.

References:

AHMADIAN A., 2000 Single-Nucleotide Polymorphism Analysis by Pyrosequencing. Analytical Biochemistry **280**: 103–110.

MA X.-F., HALL D., ONGE K. R. S., JANSSON S., INGVARSSON P. K., 2010 Genetic Differentiation, Clinal Variation and Phenotypic Associations With Growth Cessation Across the Populus tremula Photoperiodic Pathway.

MEDINA-GOMEZ C., KEMP J., ESTRADA K., ERIKSSON J., 2012 Meta-Analysis of Genome-Wide Scans for Total Body BMD in Children and Adults Reveals Allelic Heterogeneity and Age-Specific Effects at the WNT16 Locus. PLoS Genet **8**: 1–14.

NOSIL P., FUNK D. J., ORTIZ-BARRIENTOS D., 2009 Divergent selection and heterogeneous genomic divergence. Molecular Ecology **18**: 375–402.

PASTOR P., 2012 Genetic heterogeneity in Parkinson disease: The meaning of GWAS and replication studies. Neurology.

PIKE N., 2011 Using false discovery rates for multiple comparisons in ecology and evolution. Methods in Ecology and Evolution **2**: 278–282.

WITTKOPP P. J., HAERUM B. K., CLARK A. G., 2008 Regulatory changes underlying expression differences within and between Drosophila species. Nat Genet **40**: 346–350.

WITTKOPP P. J., SMITH-WINBERRY G., ARNOLD L. L., THOMPSON E. M., COOLEY A. M., YUAN D. C., SONG Q., MCALLISTER B. F., 2011 Local adaptation for body color in Drosophila americana. Heredity **106**: 592–602.

WITTKOPP P. J., STEWART E. E., ARNOLD L. L., NEIDERT A. H., HAERUM B. K., THOMPSON E. M., AKHRAS S., SMITH-WINBERRY G., SHEFNER L., 2009 Intraspecific polymorphism to interspecific divergence: genetics of pigmentation in Drosophila. Science **326**: 540–544.

WITTKOPP P. J., WILLIAMS B. L., SELEGUE J. E., CARROLL S. B., 2003 Drosophila pigmentation evolution: divergent genotypes underlying convergent phenotypes. PNAS **100**: 1808–1813.

ZHOU Q., LENGER C., SMITH R., KIMBERLING W. J., YE M., LEHMANN O., MACDONALD I., 2012 Evidence of genetic heterogeneity in Alberta Hutterites with Usher syndrome type I. Mol. Vis. **18**: 1379–1383.

Chapter 6

Online Quiz Use in a Large-Lecture Genetics Course

Abstract

Since the initial calls for improvement in undergraduate instruction in the science, technology, mathematics and engineering (STEM) fields almost two decades ago, there have been many efforts within specific disciplines to improve classroom environments. These case studies have provided valuable information, but further investigations of pedagogical techniques across disciplines is needed to evaluate their effectiveness. To date four methods have been studied in substantial detail for undergraduate STEM education: active learning, attending to prior knowledge, peer learning and formative assessment. Here, we examine one type of formative assessment, online quizzes, in more detail. Specifically, we present an investigation of the usefulness of online quizzes in a large university genetics course. Surprisingly, a statistically insignificant association between student performance and quiz use was found, although, students and instructor both express very positive sentiments towards the quizzes and their usefulness. While student performance did not significantly improve with quiz use, students felt they were able to better “keep up” with course content because of the quizzes, seemingly improving students’ attitudes. This study illustrates the importance of considering how a pedagogical technique

can be utilized across disciplines and how that technique can translate into benefits for students even if it is not increased performance.

Introduction

Undergraduate teaching in the science, technology, mathematics and engineering (STEM) fields within the United States has previously been reported to be in great need of improvement if students are to have the necessary knowledge and skills necessary for the 21st century (NATIONAL RESEARCH COUNCIL 1996; NATIONAL SCIENCE FOUNDATION 1996; THE BOYER COMMISSION ON EDUCATING UNDERGRADUATES IN THE RESEARCH UNIVERSITY 1998; NATIONAL RESEARCH COUNCIL 1999). Since the release of these reports on the weaknesses of undergraduate STEM education, there have been many efforts to improve teaching and learning environments within these fields, however, the overall state of undergraduate science education has not significantly improved (HANDELSMAN *et al.* 2004; BALDWIN 2009; WIEMAN *et al.* 2010; CHASTEEN *et al.* 2011). Although there is a wealth of research on new methods and approaches for undergraduate scientific teaching, the evidence for the effectiveness and broader implications of these pedagogical innovations remains minimal (LABOV *et al.* 2009; HENDERSON *et al.* 2010; CHASTEEN *et al.* 2011; NATIONAL RESEARCH COUNCIL 2011). Many of these teaching innovations have been developed within a specific field and further research is needed to assess their impacts across multiple disciplines.

Currently within the field of biology, there are few studies that address how these new teaching techniques can effectively be implemented into a large

lecture genetics course. Genetics differs from the introductory biology courses that have most often been the subject of these studies in that it requires less rote memorization and more problem solving, critical thinking, and application of general principles. This study investigates how one particular technique, online quizzes, correlates with student learning in a large university lecture genetics class. Online quizzes were chosen to be evaluated among additional pedagogical approaches implemented in a course. These other approaches have been examined more extensively within biology and other STEM fields as reviewed below, whereas there is less research pertaining to the use of online quizzes, particularly in a large genetics course.

Pedagogical practices in undergraduate STEM education

Four pedagogical practices have had particular attention within undergraduate STEM education: active learning, attending to prior knowledge/ misconceptions, peer instruction and formative assessment. Reviewed below is the evidence of the effectiveness of each of these techniques from the research in undergraduate STEM education. Subsequently, how these techniques have been studied within the context of a genetics classroom is discussed in addition to the challenges of genetics instruction. The study following describes how online quizzes are associated with student learning, thereby supplementing the literature with evidence for a technique/tool that has been not been as actively investigated in post-secondary science education, especially within the field of genetics education.

Active learning

Numerous studies have shown the benefits of implementing "active learning" into classrooms. Active learning can have many meanings, but here it refers to the broader definition of "...engaging students in the learning process..." (KRONTIRIS-LITOWITZ 2009, p. 309). There are many strategies of incorporating active learning into a classroom, but all include some manner of getting students actively participating and thinking critically about course content. Armbruster and colleagues (2009) redesigned an introductory biology course by incorporating several changes intended to improve both student attitudes and student performance. Their main focus for the redesign was implementing active learning and student-centered pedagogy techniques, but they also re-ordered the course content to "…teach specific content within the context of broad conceptual themes" (p.204). These authors found significant improvement of student attitudes (as self-reported by students) and increased performance. Active learning techniques have also been used as means to reduce the failure rate in an introductory biology course as demonstrated by Freeman and colleagues (2007). These authors featured five course designs that varied slightly in structure including the use of daily multiple-choice questions answered via personal response devices (clickers) or cards that were not necessarily graded, and weekly practice exams completed individually or as a group. Generally, Freeman et al. showed that there were significantly lower failure rates and

increased exam scores for course designs that implemented active learning techniques.

There a multitude of instructional methods that utilize active learning facets. One that has become quite popular for large lecture courses (typical of an introductory science course) is the clicker system (personal response device). Clickers have an assortment of practical uses in the classroom. They can serve as a tool to gather data for educational research (PRATHER and BRISSENDEN 2009), a formative assessment instrument providing feedback to both students and instructor (PASCHAL 2002; ARMBRUSTER *et al.* 2009), and promote student discussion (SMITH *et al.* 2009; PEREZ *et al.* 2010) as just a few examples. Clickers have been shown to help build problem-solving skills in genetics (LEVESQUE 2011), increase retention of introductory biology material for non-science majors (CROSSGROVE and CURRAN 2008), and improve grades in a general chemistry course (MACARTHUR and JONES 2008). Clickers can also be used to assess what concepts are most difficult for students or learn what knowledge and ideas students already possess about a particular subject (MACARTHUR and JONES 2008; KING 2011).

Attending to prior knowledge/misconceptions

Before the first lesson begins, students walk into a classroom with some ideas and information for that subject previously constructed from prior experiences and knowledge. Considering this prior knowledge should be an essential part of instruction for several reasons; one that has received particular

attention for undergraduates is that students can hold misconceptions (SMITH *et al.* 1993; NATIONAL RESEARCH COUNCIL 2000; ENGLE 2004; TANNER and ALLEN 2005). Misconceptions that are held by students can make learning new material difficult when course content conflicts with students' existing conceptions (NATIONAL RESEARCH COUNCIL 2000). If students' established conceptual frameworks are disregarded, misconceptions can persist for the duration of an undergraduate education.

For example, introductory biology students were identified to hold an array of misconceptions on cellular respiration by Songer and Mintzes (1999). This study found that these misconceptions remained intact even after careful instruction and surprisingly additional misconceptions had developed during instruction. Similar results were discovered among upper-level undergraduate and graduate students on their understanding of cellular respiration concepts (SONGER and MINTZES 1994). The Songer and Mintzes study illustrates how prior knowledge, if not explicitly addressed, can make it difficult for students to learn advanced material in a field such as biology where content is so interrelated.

Although misconceptions may seem impervious to modification, Pearsall and colleagues have found that biology students' conceptual frameworks on cells/living organisms can in fact be restructured over the course of semester by using concept maps periodically throughout the term to monitor students' understanding (PEARSALL *et al.* 1997).

These are just a few examples; there are many more from other disciplines such as physics and chemistry. Conceptual research in science

education has gained such momentum that there are well over six thousand documented studies on conceptions and misconceptions in science education (DUIT 2004). By identifying prior knowledge, instructors can use it to assist in student learning rather than view it as a hindrance (SMITH *et al.* 1993; TANNER and ALLEN 2005; VERHEY 2005) by getting students more active in the learning process.

Reconciling misconceptions carried with prior knowledge can be addressed by allowing students to engage with the content (SMITH *et al.* 1993; TANNER and ALLEN 2005) through active learning methods. By providing opportunities for students to be interactive players in the learning process, they can begin to identify disparities in their own conceptual framework and those held by experts, and eventually modify their frameworks as their understanding progresses (NATIONAL RESEARCH COUNCIL 2000; TANNER and ALLEN 2005).

Peer instruction

Providing opportunities for peer discussion on course material has also been shown to be beneficial for students. Getting students to talk and explain course material has been shown to increase student performance and help students incorporate new knowledge with their existing knowledge (TANNER 2009). Having students talk more within the classroom can also provide an excellent opportunity for insights into how students are thinking and allow both instructors and students to recognize any misconceptions that are present (TANNER 2009; GOUBEAUD 2010). Preszler (2009) describes how incorporating

student-led workshops have improved learning and retention in an introductory biology course. Upper-level undergraduate students who had successfully completed the course led discussions and activities such as case studies and problem solving exercises with students currently enrolled in the course. The thought of having peer-led workshops was to make the course material more accessible to students. Although professors are well-versed and often extremely knowledgable in a their field, Preszler argues that “The large content-specific cognitive gaps between instructors and students limit students’ abilities to assimilate information presented by their instructors and as a result, limit the students’ abilities to generate knowledge” (p.182). Additionally, in a separate longitudinal study, it was identified that peer instruction significantly improved student learning in an introductory physics course on both conceptual reasoning and quantitative problem solving (CROUCH and MAZUR 2001). Peer instruction has also been shown to improve student performance on qualitative problem-solving questions in a physiology course (GIULIODORI *et al.* 2006).

Formative assessment

Formative assessment has gained more attention recently in science education, particularly as a component of active learning techniques.(PASCHAL 2002; KNIGHT and WOOD 2005; MACARTHUR and JONES 2008; SMITH *et al.* 2008; ARMBRUSTER *et al.* 2009; KRONTIRIS-LITOWITZ 2009; FREEMAN *et al.* 2011) Formative assessment has been described “...as a feedback loop to close the gap between the learner’s current status and desired goals” (HERITAGE 2010, p.

4). In order for any activity or instrument in a classroom to be considered formative, the resulting information gained must be used to assist student learning either by the instructor or the individual student (BLACK and WILLIAM 1998).

Formative assessment has been found to be a powerful learning tool. The most effective results have occurred when formative assessment is implemented in a manner to provide continuous feedback (NATIONAL RESEARCH COUNCIL 2000). In a highly-structured undergraduate introductory biology course, the implementation of frequent formative assessments and intensive active learning techniques led to a lower course failure rate (FREEMAN *et al.* 2011). Formative assessment has also been shown to significantly improve student performance when practiced on a weekly basis (NELSON *et al.* 2009).

Students can be better informed of their own understanding with the use of formative assessment (NATIONAL RESEARCH COUNCIL 2000; SHEPARD 2005; GOUBEAUD 2010). Students can identify concepts they do not quite understand by using constructive feedback provided by a formative assessment. This not only supplies students information to improve the learning process, but also enhances their metacognitive skills (SHEPARD 2005).

Pedagogical approaches & the challenge of genetics instruction

Although many studies have investigated the effectiveness of pedagogical methods such as active learning techniques, misconceptions knowledge and peer instruction across several fields including biology, physics and chemistry,

this study focuses on a specific branch within biology: genetics. Much of the research in biology has been done in introductory biology courses (SONGER and MINTZES 1994; LORD 2005; VERHEY 2005; FREEMAN *et al.* 2007; CROSSGROVE and CURRAN 2008; ARMBRUSTER *et al.* 2009; NELSON *et al.* 2009; PRESZLER 2009; PEREZ *et al.* 2010; STANGER-HALL *et al.* 2010; FREEMAN *et al.* 2011) and less has been examined within genetics .

In a complex discipline such as genetics, there are multiple, interrelated levels of content ranging from molecular to organismal. With this hierarchal and often intricate structure to the subject, students frequently have significant difficulties understanding basic genetics concepts (DUNCAN *et al.* 2009; DUNCAN and TSENG 2010). Additionally, genetics requires a strong set of problem solving skills. Compared to traditional biology courses (especially introductory biology), genetics resembles more of a math, chemistry or physics course where students are expected to master new principles and be able to apply them to new scenarios. Add these layers of complexity and difficulty together, and its not surprising to see why genetics can be quite challenging for some students.

Genetics serves as a vital component to biology education (LEE and JABOT 2011) with elements important for biochemistry, molecular biology, and evolution. A strong understanding in genetics can only benefit students in their future biological studies. Active learning techniques have been shown to increase learning gains (as measured by a pre-/post-test) within a genetics classroom(LEE and JABOT 2011). Additionally, clicker use in a genetics class has been found to correlate with increased student performance (LEVESQUE 2011) and peer

discussion can help students' understanding of genetics concepts (SMITH *et al.* 2009).

Generally, we were interested in how student learning could be aided in a university-level genetics course with a considerable enrollment size of over four hundred students (per semester).

Many pedagogical methods, such as active learning techniques and peer instruction, have been implemented within this particular genetics course under study. The use of clickers to promote peer discussion has been incorporated, as well as interactive lecture slides to allow students to become more engaged in lecture.

Online quizzes were also introduced as a means of formative assessment; providing immediate feedback to students on their understanding of key genetics topics as well as providing feedback to the instructor on weak areas for the class as a whole and could then adjust lectures accordingly to supply additional clarification, practice and explanation. These quizzes provided students with frequent low-risk opportunities to practice and demonstrate their proficiency with course content, which has been shown to contribute to increased student performance (BLACK and WILLIAM 1998; KNIGHT and WOOD 2005; FREEMAN *et al.* 2011). Moreover, the frequency of the quizzes forces students to keep up with the material as it is presented in lecture. The quizzes also give students additional time to be engaged and actively work with course material; allowing time to modify and incorporate their existing conceptual framework.

Given the benefits of other formative assessment techniques (BLACK and WILLIAM 1998; NATIONAL RESEARCH COUNCIL 2000; SHEPARD 2005; KNIGHT and WOOD 2005; NELSON *et al.* 2009; FREEMAN *et al.* 2011), and the effectiveness of quizzes in other biology courses (KNIGHT and WOOD 2005; FREEMAN *et al.* 2011), and that the other teaching techniques implemented have already undergone some investigation within the field of genetics, we chose to focus our study on how the online quizzes related to student learning in a large university genetics lecture course by asking if the use of online quizzes correlated with student achievement. This analysis provides valuable data for the field of genetics education, which has had fewer formal investigations, in addition to providing practical suggestions for improving student achievement in such a large lecture class.

Materials and Methods

This study analyzed data gathered over four semesters from a genetics lecture course offered at the University of Michigan (UM). Genetics is a requirement for all biology majors (including many pre-medical students).

Genetics lecture is typically co-taught by two instructors. One instructor teaches the first half of the course focused on topics such as inheritance, evolutionary & population genetics, whereas the second instructor focuses on more molecular genetics and gene regulation in the remaining half of the course. Genetics is also offered throughout the academic year, however, this analysis uses data from only the first half of genetics offered during the Winter semester

of 2008, 2010, 2011 and 2012 to control for any instructor bias (P.W. taught across all data analyzed). The genetics lecture offered during these semesters met twice a week for a duration of 90 minutes each lecture. Genetics typically has an enrollment of approximately 400 students. For the semesters analyzed here, the course had an enrollment ranging from 385-438 students with an average of 52.7% females and 7.0% minorities (see Table 6.1 for more detail). Of these students, our data is drawn from a subset of students who completed all relevant exams that were necessary for analysis.

Table 6.1: Biology 305 course make-up

Semesters	2008	2010	2011	2012
Total Students Enrolled	405	438	418	385
Percent Female	54.7%	54.3%	52.8%	49.2%
Percent Minority	7.4%	7.1%	6.0%	7.3%

Course structure

Instructional techniques that have been shown to improve student learning in other disciplines and class sizes (FREEMAN *et al.* 2007; CROSSGROVE and CURRAN 2008; MACARTHUR and JONES 2008; DERTING and COX 2008; NAGY-SHADMAN and DESROCHERS 2008) were implemented (or removed) each semester. Clickers and interactive lecture slides were incorporated into the

lecture in 2008, and remained as part of the course structure in all subsequent semesters.

Online quizzes were introduced in 2010 and then removed in the following 2011 year for administrative purposes, but reinstated in 2012. Thus, semesters with quizzes included 2010 and 2012 and semesters without quizzes were 2008 and 2011. The quizzes were a required component of the course (students received points based on how accurately they completed the quiz questions). Quizzes were required to be completed prior to every lecture (students completed two quizzes every week during the term). Each quiz consisted of eight questions covering material from the upcoming lecture's "pre-readings", as well as related content from previous lectures (See Appendix A for examples of quiz questions). Quizzes were open book and posted through the course's online university webpage (ctools.umich.edu). The instructor used the quiz scores to gauge how well students were understanding certain concepts prior to lecture and would adjust the lecture to offer additional practice, explanation or clarification. When utilized in this fashion, quizzes served as a formative assessment to improve both teaching and learning.

Data collection

Data gathered for student performance was taken from student exams. Two unit exams were administered during the first half of the genetics lecture course. Exams consisted of 38-42 multiple choice problems. Both overall exam scores and individual question responses were recorded for each student.

Individual question responses were used rather than overall exam scores as a measure of student performance. Exam content was not standardized across semesters (i.e., some exams may have included slightly different concepts or tested the same concepts in different ways) and overall exam scores were quite similar across all semesters (Figure 6.1). Specific exam questions across semesters were identified as comparable if the same topic was tested at a similar level of understanding across all exams (by the instructor who designed the exams) (See Table 6.2 for comparable question topics and Appendix B for examples of comparable exam questions). All comparable questions were at a level that an "average" student should be able to correctly answer or "C-level" questions. (If all C-level questions were answered correctly on an exam, a student would earn at least ~75-80% for that exam.) Comparable questions answers were used to formulate a new student exam score, a comparable questions score. Comparable question scores were also similar across semesters (Figure 6.2), but not as similar as overall exam scores (compare Figures 6.1 and 6.2). Overall student exam scores and comparable questions scores were correlated (R^2=0.597, y=0.0097x+ 0.0963) for student performance indicating comparable question scores serve as a reasonable proxy for overall performance.

Figure 6.1: Median exam scores across semesters

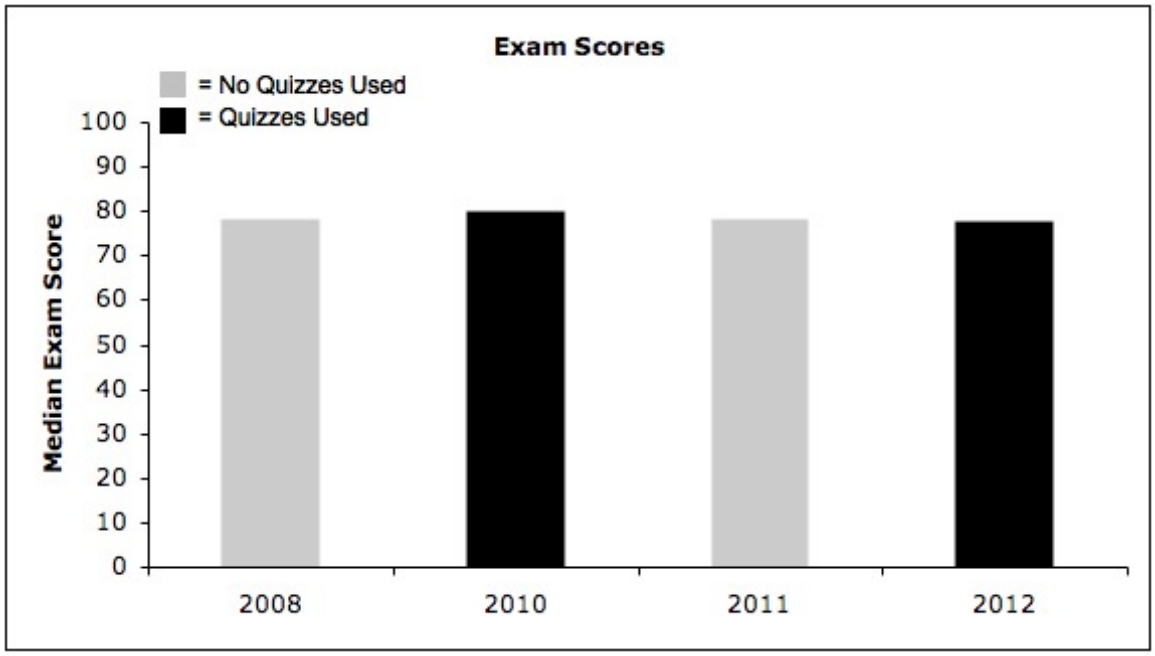


Figure 6.2: Median comparable questions scores across semesters

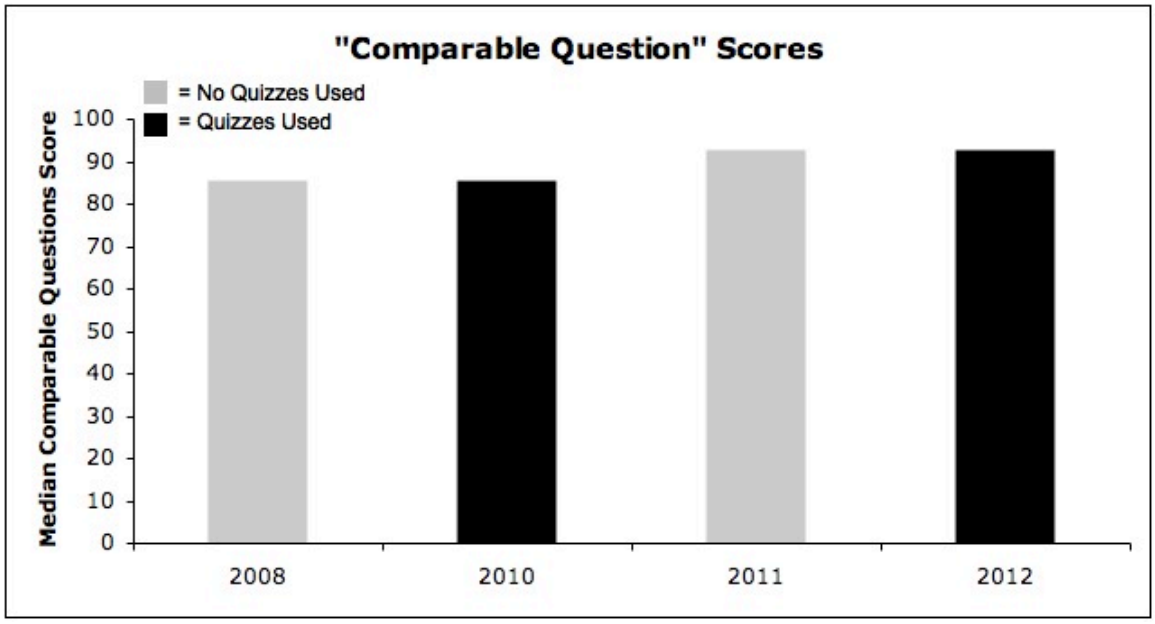


Table 6.2. Comparable questions topics

Topic/Question	Abbreviation
Chi squared test	chi sq
Mutation mapping using deletion information	deletion
Genetic distance between two genes using information from two genes	dist2
Genetic distance between two genes using information from three genes	dist3
Allele frequency based on fitness	fitness
Heritability for a quantitative trait	heritability
Hfr *E.coli* strains	Hfr
Hardy-Weinberg Equation	HWE
Identifying stages of mitosis	mitosis
MacDonald-Kreitman Test	MK
Determine an enzymatic pathway from mutation data	pathway
Determine phased genotypes	phasing
P-value calculation	pval
Quantitative genetics	quant

Additional student data was also collected from the UM Office of the Registrar for 2060 students enrolled in the Winter semester of genetics lecture during 2008, 2010, 2011 and 2012. This data included various demographics (e.g. gender, ethnicity, age, etc), cumulative GPA (up to the semester enrolled in genetics), and SAT and ACT scores that were used to gauge how comparable students were across semesters. Students' names were removed and were

identified by an eight-digit numbers for the 2012 term, whereas data for the 2008, 2010, and 2011 terms were stripped of all possible identifiers and were given new four-digit numbers in accordance with this study's IRB exemption.

Data analysis

The data was imported into the statistical software package, R (version 2.13.1). An analysis of variance (ANOVA) was performed to evaluate students' raw GPA, ACT and SAT scores across semesters. A generalized linear mixed model using the logistic function (glmer function in R-available lme4 library) was used to evaluate the relationship between online quiz use and all students' scores on comparable questions, among females, minorities and across different semesters. This model allowed us to account for variation in student performance within any given year (by using a random effect for semester). Whether the student answered the question (for a particular topic) correctly or not was the dependent (response) variable and independent (explanatory) variables included whether or not quizzes were used, sex or ethnicity, the interaction between quizzes used and sex or quizzes used and ethnicity and "year" was accounted for as a random effect. (Ethnicity was dichotomized as either "majority"=White or Asian or "minority"=Black, Hispanic, Native American, or Hawaiian as reported by students to the University Registrar.)

Because each question was was tested independently, we corrected for multiple comparisons by controlling the expected proportion of false positives, (false discovery rate). Data is reported as the proportion of students correctly

answering a question and the odds ratio of student performance with and without quiz use by semester. The odds ratio is reported to more easily visualize the difference in student performance with and without quiz use.

Qualitative student comments about the quiz use was gathered from the 2012 semester. All students were presented with an optional online survey where they could provide comments on what aspects of the course they found most and/or least helpful. Approximately 76% of students (292/385) completed the survey. Of these students, 45% (132/292) provided comments specifically pertaining to the use of the online quizzes. Comments were sorted into the two major themes of positive and negative comments toward the online quizzes.

Results

Academically, students were found to be comparable across semesters. Among the four semesters surveyed, GPA and SAT scores were found to be similar [$F(3)=1.68$, $p= 0.1694$ and $F(3)=1.09$, $p= 0.3537$ respectively] (Table 6.3). Student ACT scores, however, were identified to differ significantly, but only slightly between semesters [$F(3)=5.39$, $p= 0.0011$] (Table 6.3). The student population that typically completes the ACT examination tends to be exclusively in-state residents.This subset of students may contain more variation in achievement than the larger group of students as a whole. This biased selection of students could have contributed to the significant difference observed in student performance between semesters. Additionally, the mean ACT composite

scores are highly similar between most of the semesters with only 2008 having a slightly lower score.

Table 6.3: Students are comparable across semesters

Semesters	2008	2010	2011	2012	N (Total=1559)	*p* value
Mean GPA	3.29	3.36	3.33	3.35	1559	0.1694
Mean SAT Composite Score	1322	1321	1320	1342	708	0.3537
Mean ACT Composite Score	28.9	29.5	29.7	29.8	1315	0.0011*

*Significant p value < 0.01.

We surprisingly found little evidence of a positive association between student performance and online quiz use in this course. Of the fourteen topics, only two had increased student performance positively associated with the online quiz use: genetic distance using three genes (dist3) and quantitative genetics (quant). An additional topic, MacDonald-Kreitman Test (MK), actually showed a statistically significant negative association with quiz use with all students' performance (Figure 6.3 and Figure 6.4). The majority of topics, however, did not display any significant relationship between student performance and the use of the quizzes (Figure 6.3 and Figure 6.4).

Figure 6.3: Student performance by comparable question topic

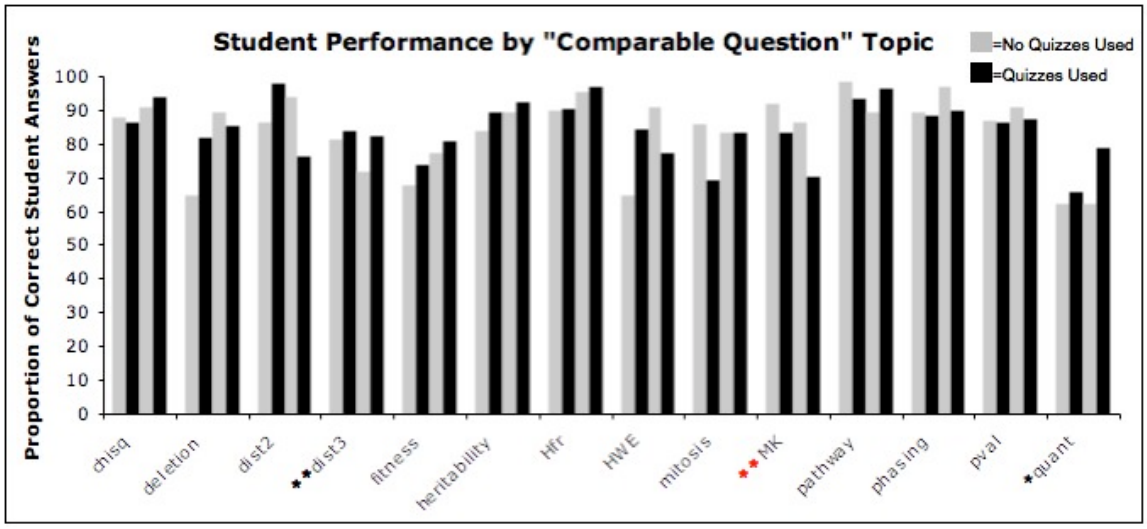


Figure 6.3: Gray bars indicate semesters when quizzes were not used and black bars indicate semesters that had quizzes. Student performance is reported for each comparable question topic by the proportion of students that answered correctly. * indicates a p value ≤ 0.05 with a positive association with quiz use, ** indicates a p value ≤ 0.01 with a positive association with quiz use and ** indicates a p value ≤ 0.01 with a negative association with quiz use.

Figure 6.4: Odds ratio of student performance with quizzes to student performance without quizzes

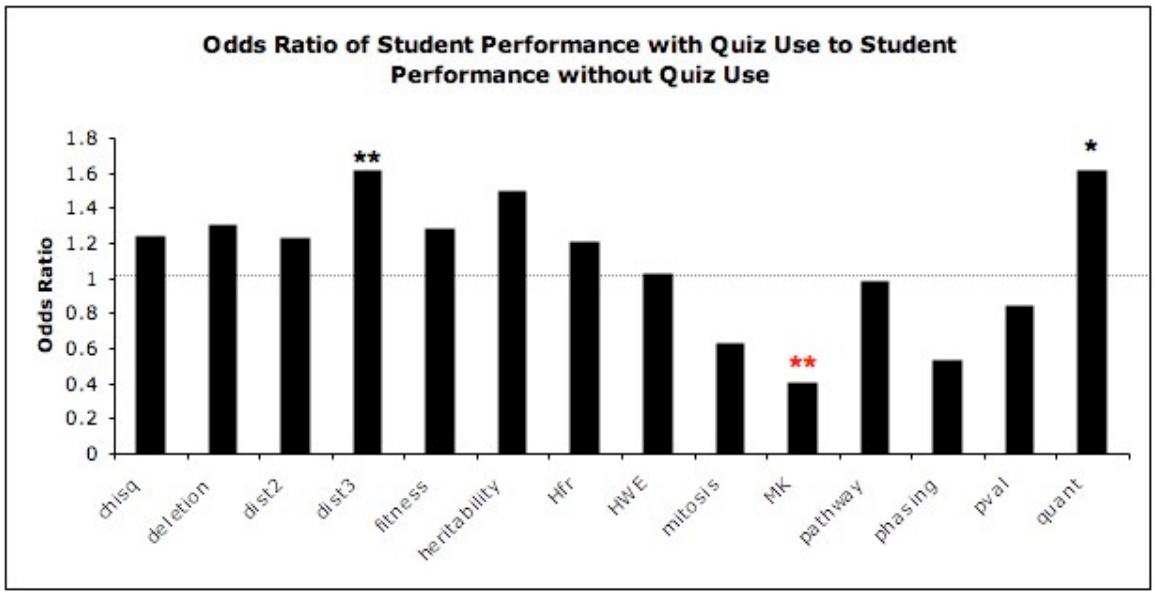


Figure 6.4: Odds ratio of student performance with quizzes versus student performance without quizzes is shown. Dotted line at 1 indicates an equal likelihood of a student answering a question correctly with or without quizzes used. * indicates a *p* value ≤ 0.05 with a positive association with quiz use, ** indicates a *p* value ≤ 0.01 with a positive association with quiz use and ** indicates a *p* value ≤ 0.01 with a negative association with quiz use.

Data was also analyzed in student subgroups by gender and ethnicity to investigate whether quiz use benefited a smaller subset of students within the larger lecture. It has been shown that females and minorities can be at a higher risk for struggling in the sciences (SPELKE 2005; HAAK *et al.* 2011 respectively) due to challenges not necessarily associated with a lower innate ability (SPELKE 2005; DIRKS and CUNNINGHAM 2006). We were interested to see if the use of online quizzes would correlate with female or minority comparable question

scores. However, no significant association of any kind was identified for female or minority performance and use of quizzes (Figure 6.5-Figure 6.8).

Figure 6.5: Female performance by comparable questions topic

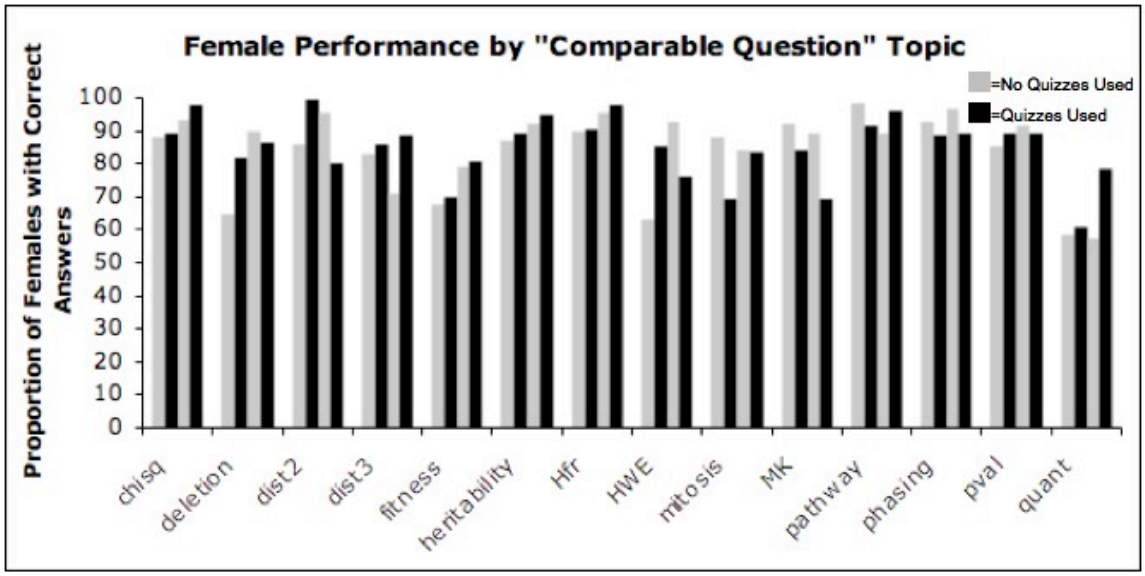


Figure 6.5: Gray bars indicate semesters when quizzes were not used and black bars indicate semesters that had quizzes. Female student performance is reported for each comparable question topic by the proportion of female students that responded with the correct answer.

Figure 6.6: Odds ratio of female performance with quizzes to female performance without quizzes

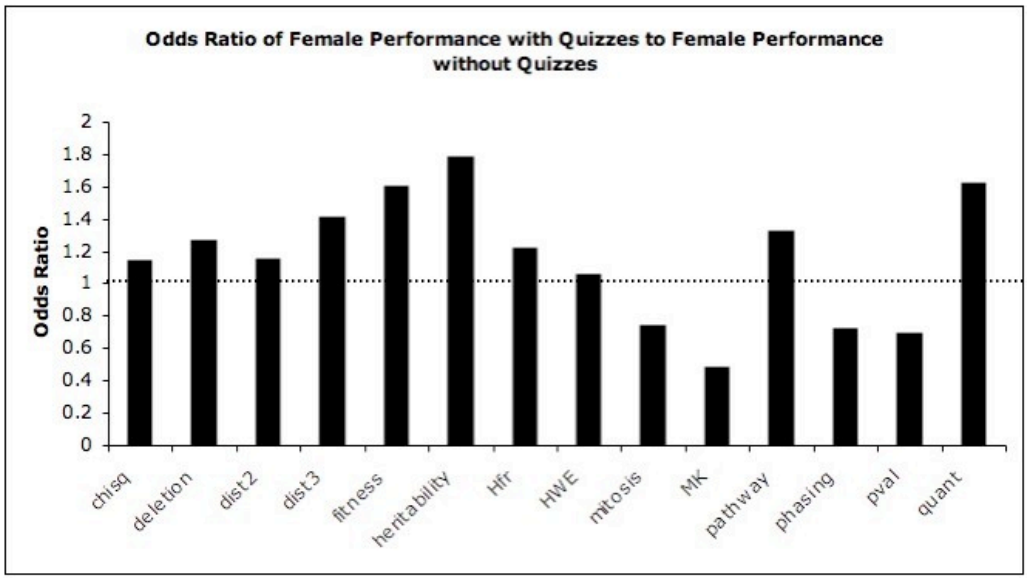


Figure 6.6: Odds ratio of female performance with quizzes versus female performance without quizzes is shown. Dotted line at 1 indicates an equal likelihood of a female student answering a question correctly with or without quizzes used.

Figure 6.7: Minority performance by comparable questions topic

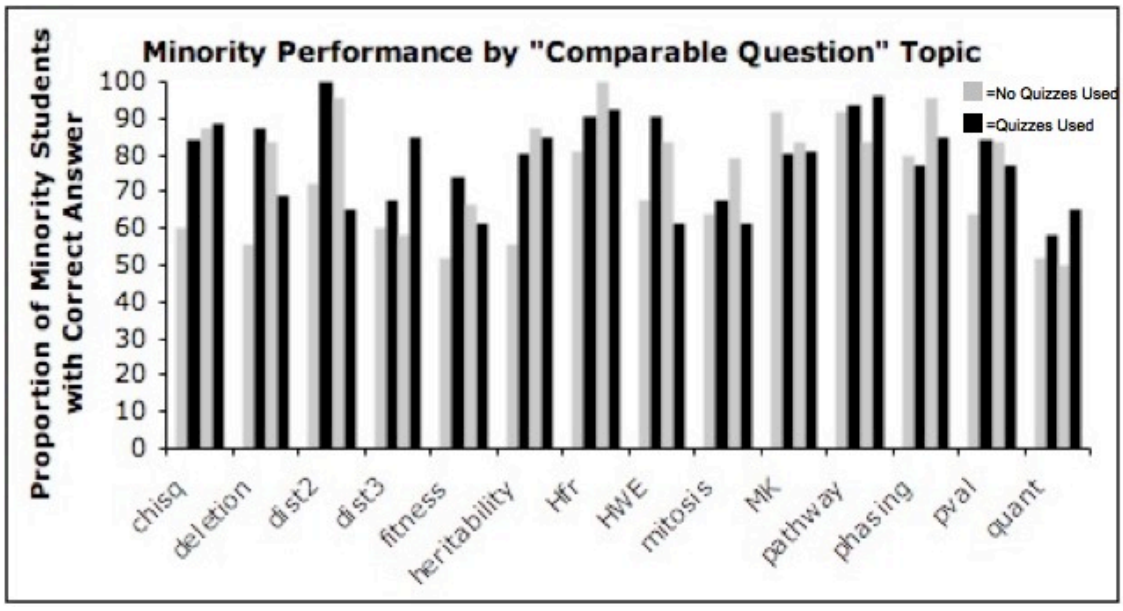


Figure 6.7: Gray bars indicate semesters when quizzes were not used and black bars indicate semesters that had quizzes. Minority student performance is reported for each comparable question topic by the proportion of minority students that responded with the correct answer. A student was classified as a minority if self-reported not “White” or “Asian.” Minority categories included “Black,” “Hispanic,” “Native American” and “Hawaiian” as reported to the University Registrar.

Figure 6.8: Odds ratio of minority performance with quizzes to minority performance without quizzes

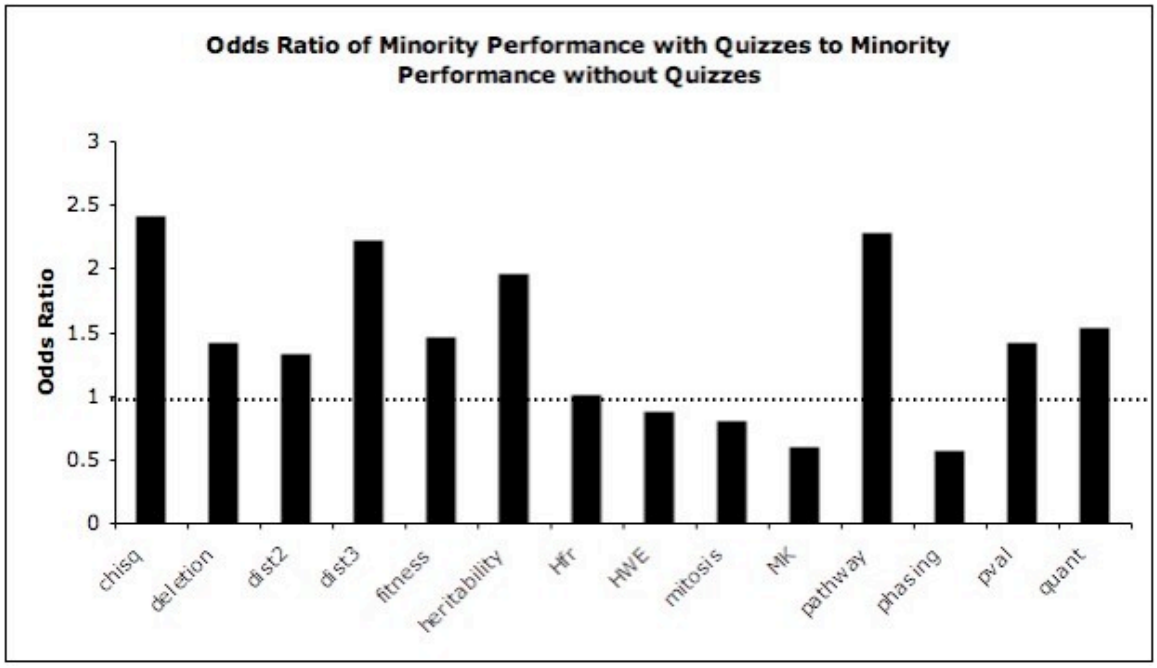


Figure 6.8: Odds ratio of minority performance with quizzes versus minority performance without quizzes is shown. Dotted line at 1 indicates an equal likelihood of a minority student answering a question correctly with or without quizzes used. A student was classified as a minority if self-reported not “White” or “Asian.” Minority categories included “Black,” “Hispanic,” “Native American” and “Hawaiian” as reported to the University Registrar.

A qualitative survey showed that students generally had positive feelings towards the use of the online quizzes. From the 2012 semester, approximately 71% of the total students completed an online survey about the usefulness of various course components. Of these students that completed the survey, and provided comments specifically pertaining to quiz use, 85% voiced positive feelings towards the quizzes and only 15% expressed negative feelings. Many students with a positive opinion of the quizzes felt that quizzes helped them "keep up" with the material as represented by the following student comment: "The quizzes required for every class was helpful because it kept me on a regular schedule for studying." Other students reported that they felt the quizzes were a worthwhile tool although they did not always "like" the quizzes per se as demonstrated by this student's comment: "While I hated the quizzes during the semester-I must admit they helped."

The small percentage of students that reported negative feelings towards the quizzes generally fell into one of two themes as represented by the following student comments: "I did not like the online quizzes however. Yes, it made you study on a regular basis, but the questions were not helpful in planning for the exam and I only really completed them for the points. They were not challenging nor did they help me learn the material" or "I really didn't like the online quizzes just because I did the reading and the coursepack problems and still struggled to get more than 6 [out of 8] correct." Of course, not all interventions can always help all students, but only a slight percentage expressed these negative views as

compared to the majority that conveyed more positive opinions of the use of quizzes in the course.

Discussion

Impacts of quiz use on student performance

Surprisingly, little evidence was found to suggest a positive association between the use of online quizzes in this large lecture genetics course and increased student performance (among all students or among smaller subsets of students). There are several potential explanations for this finding.

Online quizzes were just one component within this course. Students had access to other learning techniques throughout the course such as in-class clicker questions, and interactive lecture slides. As previously shown, implementing clicker questions or interactive lecture slides have been shown to be associated with increases in student learning (FREEMAN *et al.* 2007; CROSSGROVE and CURRAN 2008; MACARTHUR and JONES 2008; DERTING and COX 2008; NAGY-SHADMAN and DESROCHERS 2008) Additionally, a separate discussion section led by a graduate student instructor (coached by P.W., the course instructor) was also an element within the course. In one of the semesters when quizzes were not used, students completed a problem set similar to questions presented on the online quizzes during these discussion sections. Students likely worked in groups to complete these problem sets, allowing opportunities for peer instruction to occur. As has been established, the use of peer instruction can enhance a student's learning experience (CROUCH and

MAZUR 2001; GIULIODORI *et al.* 2006; TANNER 2009; PRESZLER 2009). These other course features may have replicated the use of online quizzes in semesters when they were not used.

Using online quizzes may be more beneficial to students on higher order thinking problems. This analysis cannot make any claims to how online quizzes may impact student performance on "higher" level or harder questions. Theoretically, by having additional practice with course material, it could potentially allow students to gain a deeper level of understanding (of some topics), and therefore perform better with higher order exam problems. Here, student performance was measured from student responses strictly from "average" level questions. It is suggested that in order for quizzes to be most effective as a teaching instrument, quiz material must match both the level and content of the exams (measure of student performance) (NARLOCH *et al.* 2006).

To this end, the content of the quizzes could also impact their effectiveness as a teaching device. Questions within the quiz could be more refined to focus on specific topics of particular difficultly for students. This would provide students with explicit and detailed practice for particular complex or problematic genetic subjects.

Our results did identify three particular topics that had either significant positive or negative associations with student learning: genetic distance using three genes (dist3), quantitative genetics (quant) and MacDonald-Kreitman Test (MK). MacDonald-Kreitman Test (MK) is unique in that it showed a negative association between quiz use and student performance. It is speculated that this

may be a result of the fact that during the 2012 semester (which used quizzes), McDonald-Kreitman Test (MK) was introduced towards the end of the study period. As such, the instructor was pressed for time to complete all the required course material, and as a result, MacDonald-Kreitman Test (MK) content was "rushed" through and not covered in as much detail as previous years. This inconsistency in teaching could have impaired student learning for this specific topic even though quizzes were used during this semester, contributing to this negative association.

The two other topics, genetic distance using three genes (dist3) and quantitative genetics (quant), that displayed a statistically significant positive association between quiz use and student learning may have resulted from one of a handful of possibilities. First, these results may be a consequence of extra attention towards these topics either via quiz questions or in-class instruction or other course components. As mentioned above, quiz effectiveness may be enhanced when quiz and exam content align together (NARLOCH *et al.* 2006). It is possible that quizzes explicitly addressed genetic distance using three genes (dist3) and quantitative genetics (quant) topics very similarly to the same topics on exams.

As an alternative hypothesis, other course components (active learning methods, peer instruction or GSI-led discussions) could have had inadvertent focused attention on these two topics in semesters that also employed quizzes. Nonetheless, it is unclear why these specific two topics of genetic distance using

three genes (dist3) and quantitative genetics (quant) would have such a significant association over the other topics surveyed.

Although there was no significant statistical association overall between student performance and the use of online quizzes, qualitative feedback from students suggests that the quizzes are still worthwhile in a large lecture genetics course. Of the students that submitted comments on the helpfulness of online quizzes, a substantial percentage (85%) shared positive remarks. The main theme from these comments was that the quizzes helped students “keep up” with the course material. In a course like genetics that moves very quickly with a considerable amount of potentially difficult information to learn, having a tool to stay on top of the material and check your understanding could be quite valuable to students. The instructor, P.W., also shared the students’ sentiments towards the quizzes. The questions that students brought to office hours during semesters that used quizzes seemed to demonstrate that the students had a deeper level of understanding of the course content versus semesters when quizzes were not used. This indicates that online quiz use may help motivate students to spend more time thinking about or studying genetics material. Improving student attitudes towards science courses or the use of student-centered pedagogical techniques has been demonstrated with a variety of course structures (Lord 2005; Knight and Wood 2005; Huerta 2007; MacArthur and Jones 2008; Derting and Cox 2008; Nagy-Shadman and Desrochers 2008; Prather and Brissenden 2009; Armbruster *et al.* 2009; Wieman *et al.* 2010; Knight and Smith 2010).

Areas of future research

Online quizzes will continue to be used in this large genetics lecture despite the results of this quantitative analysis. As mentioned, quiz questions could be refined to target specific problematic topics, as well as some higher level order problems. Closer design of the quizzes could ensure that quiz material more appropriately matches with the content of exam questions to potentially increase effectiveness of the quizzes. Additionally, in order to disentangle the effects of other course components such as active learning techniques and peer instruction, further, more controlled studies are needed to truly discern the effectiveness of quizzes in a large university genetics lecture course.

Although online quizzes have shown to be an effective tool for student learning in the field of psychology (NARLOCH *et al.* 2006), this analysis illustrates that their effectiveness for student learning can be dependent upon many components. As the many domains of science continue to rapidly evolve, so should our scientific instruction in the classroom. This study demonstrates the importance of evaluating pedagogical techniques for specific disciplines and class sizes.

References:

ARMBRUSTER P., PATEL M., JOHNSON E., WEISS M., 2009 Active Learning & Student-centered Pedagogy Improve Student Attitudes and Performance in Introductory Biology. CBE-Life Sciences Education **8**: 203–213.

BALDWIN R. G., 2009 The Climate for Undergraduate Teaching and Learning in STEM Fields. New Directions for Teaching & Learning, Wiley Periodicals Inc.: 9–17.

BLACK P., WILLIAM D., 1998 Inside the Black Box Raising Standards Through Classroom Assessment. Phi Delta Kappan **80**: 139–148.

CHASTEEN S. V., PERKINS K. K., BEALE P. D., POLLOCK S. J., WIEMAN C., 2011 A Thoughtful Approach to Instruction: Course Transformation for the Rest of Us. Journal of College Science Teaching **40**: 24–30.

CROSSGROVE K., CURRAN K. L., 2008 Using Clickers in Nonmajors- and Majors-Level Biology Courses: Student Opinion, Learning, and Long-Term Retention of Course Material. CBE-Life Sciences Education **7**: 146–154.

CROUCH C., MAZUR E., 2001 Peer Instruction-Ten years of experience and results. American Journal of Physics **69**: 970–977.

DERTING T. L., COX J. R., 2008 Using a Tablet PC To Enhance Student Engagement and Learning in an Introductory Organic Chemistry Course. Journal of Chemical Education **85**: 1–6.

DIRKS C., CUNNINGHAM M., 2006 Enhancing Diversity in Science: Is Teaching Science Process Skills the Answer? CBE-Life Sciences Education **5**: 218–226.

DUIT, R. (2004). Bibliography: Students' and Teachers' Conceptions and Science Education Database. University of Kiel, Kiel, Germany. http://www.ipn.uni-kiel.de/aktuell/stcse/stcse.html

DUNCAN R. G., TSENG K. A., 2010 Designing project-based instruction to foster generative and mechanistic understandings in genetics. Sci. Ed. **95**: 21–56.

DUNCAN R. G., ROGAT A. D., YARDEN A., 2009 A learning progression for deepening students' understandings of modern genetics across the 5th-10th grades (CE Hmelo-Silver and RG Duncan, Eds.). J. Res. Sci. Teach. **46**: 655–674.

ENGLE D., 2004 The Biological Justification for Transforming our Teaching. Cell Biology Education **3**: 233–234.

FREEMAN S., HAAK D., WENDEROTH M. P., 2011 Increased Course Structure Improves Performance in Introductory Biology. CBE-Life Sciences Education **10**: 175–186.

FREEMAN S., O'CONNOR E., PARKS J., CUNNINGHAM M., HURLEY D., HAAK D., DIRKS C., WENDEROTH M., 2007 Prescribed Active Learning Increases Performance in Introductory Biology. CBE-Life Sciences Education **6**: 132–139.

GIULIODORI M. J., LUJAN H. L., DICARLO S. E., 2006 Peer Instruction Enhanced Student Performance on Qualitative Problem-Solving Questions. Adv Physiol Educ **30**: 168–173.

GOUBEAUD K., 2010 How is Science Learning Assessed at the Postsecondary Level? Assessment and Grading Practices in College Biology, Chemistry and Physics. J Sci Educ Technol **19**: 237–245.

HAAK D. C., HILLERISLAMBERS J., PITRE E., FREEMAN S., 2011 Increased Structure and Active Learning Reduce the Achievement Gap in Introductory Biology. Science **332**: 1213–1216.

HANDELSMAN J., EBERT-MAY D., BEICHNER R., BRUNS P., CHANG A., DEHAAN R., GENTILE J., LAUFFER S., STEWART J., TILGHMAN S. M., WOOD W. B., 2004 Scientific teaching. Science **304**: 521–522.

HENDERSON C., FINKELSTEIN N., BEACH A., 2010 Beyond Dissemination in College Science Teaching: An Introduction to Four Core Change Strategies. Journal of College Science Teaching: 18–25.

HERITAGE M., 2010 *Formative assessment and next-generation assessment systems: Are we losing an opportunity*. Council of Chief State School Officers, Washington, DC.

HUERTA J. C., 2007 Getting Active in the Large Lecture. Journal of Political Science Education **3**: 237–249.

KING D. B., 2011 Using Clickers to Identify the Muddiest Points in Large Chemistry Classes. J. Chem. Educ. **88**: 1485–1488.

KNIGHT J. K., SMITH M. K., 2010 Different but Equal? How Nonmajors and Majors Approach and Learn Genetics. CBE-Life Sciences Education **9**: 34–44.

KNIGHT J. K., WOOD W. B., 2005 Teaching More by Lecturing Less. Cell Biology Education **4**: 298–310.

KRONTIRIS-LITOWITZ J., 2009 Articulating Scientific Reasoning Improves Student Learning in an Undergraduate Anatomy and Physiology Course. CBE-Life Sciences Education **8**: 309–315.

LABOV J. B., SINGER S. R., GEORGE M. D., SCHWEINGRUBER H. A., HILTON M. L., 2009 Effective Practices in Undergraduate STEM Education Part 1: Examining the Evidence. CBE-Life Sciences Education **8**: 157–161.

LEE W., JABOT M. E., 2011 Incorporating Active Learning Techniques into a Genetics Class. Journal of College Science Teaching **40**: 94–100.

LEVESQUE A. A., 2011 Using Clickers to Facilitate Development of Problem-Solving Skills. CBE-Life Sciences Education **10**: 406–417.

LORD T. R., 2005 A Comparison Between Traditional and Constructivist Teaching in College Biology. Innovative Higher Education **21**: 197–216.

MACARTHUR J. R., JONES L. L., 2008 A Review of Literature Reports of Clickers Applicable to College Chemistry Classrooms. Chem. Educ. Res. Pract. **9**: 187–195.

NAGY-SHADMAN E., DESROCHERS C., 2008 Student Response Technology: Empirically grounded or just a gimmick? Int. J. of Sc. Educ. **30**: 2023–2066.

NARLOCH R., GARBIN C. P., TURNAGE K. D., 2006 Benefits of Prelecture Quizzes. Teaching of Psychology **33**: 109–112.

NATIONAL RESEARCH COUNCIL, 1996 *From Analysis to Action: Undergraduate Education in Science, Mathematics, Engineering and Technology*. National Academy Press, Washington, D.C.

NATIONAL RESEARCH COUNCIL, 1999 *Transforming Undergraduate Education in Science, Mathematics, Engineering and Technology*. National Academy Press, Washington, D.C.

NATIONAL RESEARCH COUNCIL, 2000 *How People Learn*. National Academy Press, Washington, D.C.

NATIONAL RESEARCH COUNCIL, 2011 *Promising Practices in Undergrad STEM Education* (N Nielsen, Ed.). National Academies Press, Washington D.C.

NATIONAL SCIENCE FOUNDATION, 1996 Shaping the Future: New Perspectives on Undergraduate Education in Science, Mathematics, Engineering and Technology.

NELSON J., ROBINSON D., BELL J., BRADSHAW W., 2009 Cloning the Professor, an Alternative to Ineffective Teaching in a Large Course. CBE-Life Sciences Education **8**: 252–263.

PASCHAL C. B., 2002 Formative Assessment in Physiology Teaching Using a Wireless Classroom Communication System. Adv Physiol Educ **26**: 299–308.

PEARSALL N. R., SKIPPER J. E. J., MINTZES J., 1997 Knowledge Restructuring in the Life Sciences. Sci. Ed.: 193–215.

PEREZ K., STRAUSS E. A., DOWNEY N., GALBRAITH A., JEANNE R., COOPER S., 2010 Does Displaying the Class Results Affect Student Discussion during Peer Instruction? CBE-Life Sciences Education **9**: 133–140.

PRATHER E. E., BRISSENDEN G., 2009 Clickers as Data Gathering Tools and Students' Attitudes, Motivations, and Beliefs on Their Use in This Application. Astronomy Education Review **8**: 010103_1–010103_10.

PRESZLER R. W., 2009 Replacing Lecture with Peer-led Workshops Improves Student Learning. CBE-Life Sciences Education **8**: 182–192.

SHEPARD L. A., 2005 Linking Formative Assessment to Scaffolding. Educational Leadership **63**: 66–70.

SMITH J. P. III, DISESSA A. A., ROSCHELLE J., 1993 Misconceptions Reconceived: A Constructivist Analysis of Knowledge in Transition. Journal of the Learning Sciences **3**: 115–163.

SMITH M. K., WOOD W. B., ADAMS W. K., WIEMAN C., KNIGHT J. K., GUILD N., SU T. T., 2009 Why Peer Discussion Improves Student Performance on In-Class Concept Questions. Science **323**: 122–124.

SMITH M. K., WOOD W. B., KNIGHT J. K., 2008 The Genetics Concept Assessment: A New Concept Inventory for Gauging Student Understanding of Genetics. CBE-Life Sciences Education **7**: 422–430.

SONGER C., MINTZES J., 1994 Understanding cellular respiration: An analysis of conceptual change in college biology. J. Res. Sci. Teach. **31**: 621–637.

SPELKE E. S., 2005 Sex differences in intrinsic aptitude for mathematics and science?: a critical review. Am Psychol **60**: 950–958.

STANGER-HALL K. F., LANG S., MAAS M., 2010 Facilitating Learning in Large Lecture Classes: Testing the “Teaching Team” Approach to Peer Learning. CBE-Life Sciences Education **9**: 489–503.

TANNER K., 2009 Talking to Learn: Why Biology Students Should Be Talking in Classrooms and How to Make It Happen. CBE-Life Sciences Education **8**: 89–94.

TANNER K., ALLEN D., 2005 Approaches to biology teaching and learning: understanding the wrong answers--teaching toward conceptual change. Cell Biology Education **4**: 112–117.

THE BOYER COMMISSION ON EDUCATING UNDERGRADUATES IN THE RESEARCH UNIVERSITY, 1998 *Reinventing Undergraduate Education: A Blueprint for America's Research Universities*. Carnegie Foundation for the Advancement of Teaching., Menlo Park, CA.

VERHEY S. D., 2005 The Effect of Engaging Prior Learning on Student Attitudes toward Creationism and Evolution. BioScience **55**: 996–1003.

WIEMAN C., PERKINS K., GILBERT S., 2010 Transforming Science Education at Large Research Universities: A Case Study in Progress. Change **42**: 7–14.

Appendix A: Examples of Quiz Questions

Topic: Heritability

Two true-breading lines of fish are intercrossed. In the F_1, the variance in body weight is 3. The F_1 is selfed; in the F_2, the variance in body weight is 6. What is the broad-sense heritability of body weight in these fish?

Enter your answer as a decimal between 0 and 1.

Topic: Genetic distance between two genes using information from two genes

A haploid strain of yeast containing mutants in both *his1* and *arg2* is mated to a wildtype haploid. 80 UNORDERED tetrads are collected, dissected and each ascospore is grown up and plated on four types of media: “complete” with all amino acids, lacking histidine, lacking arginine and lacking both histidine and arginine. Each of the four haploid meiosis products in each ascus were analyzed separately.

In the following table, + = growth and - = no growth. *Remember that these +/- correspond to growth phenotypes, not genotypes*. For example "++++" means that all four of the cells in that ascus grew on that media type, while "++--" "--++" and "+-+-" mean that only two of the four spores grew on that media type, and "--+-" means that only one of the four spores grew. Within each class of spore, symbols in the same column show the results from the same cell (actually, clonal progeny from a single cell).

	Class 1	Class 2	Class3	
complete	++++	++++	++++	
no histidine	- - ++	- - + +	- - + +	
no arginine	++- -	- - + +	+ - + -	
No histidine or arginine	- - - -	- - + +	- - + -	Total
	1	65	14	80

What is the genetic distance between *his1* and *arg2*? (Enter your answer in cM. Your answer should be between 0 and 50).

Topic: Bacterial genetics

A single, visible colony of bacteria growing on an agar plate of minimal media represents:

A. A population of auxotrophic cells

B. A population containing approximately 1000 cells

C. A population of cells with many different genotypes.

D. A population of bacteria derived from a single cell

Topic: Hardy-Weinberg Equatioin (X-linked)

Assume that a population is in HWE and that color blindness is controlled by an X-linked recessive allele. If the proportion of females in the population are color blind is 0.04, what proportion of males should have NORMAL vision? Enter your answer as a decimal between 0 and 1.

Appendix B: Examples of Exam Questions (Comparable Questions)

Topic: Genetic distance between two genes using information from three genes

A.) A normal-looking snapdragon was crossed with a homozygous tall (***tt***), bushy (***bb***) snapdragon that had pink flowers (***pp***). (These mutations are each recessive). The phenotypes of the 4000 progeny were:

phenotype	number
wild-type	6
tall	110
bushy	123
pink flowers	1789
tall & bushy	1741
tall & pink flowers	107
bushy & pink flowers	120
tall, bushy & pink flowers	4

(2 points) What is the genetic distance between the *t* and *p* genes?

a. 15 cM
b. 12 cM
c. 9 cM
d. 6 cM
e. 3 cM

B.) Consider the following autosomal Drosophila traits caused by recessive alleles: *bent wings (bn)*, *short legs (sh*), and *orange eyes (or*). All of these genes are on the same chromosome. You cross two true breeding lines to produce F_1 flies, all of which have the wild type phenotype (straight wings, long legs, and red eyes). F_1 females are then test-crossed. Among 200 test-cross progeny, you observe the following phenotypes:

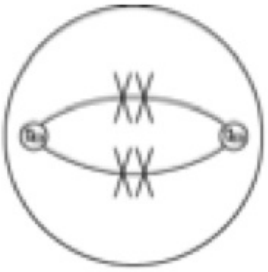
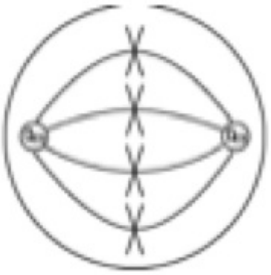
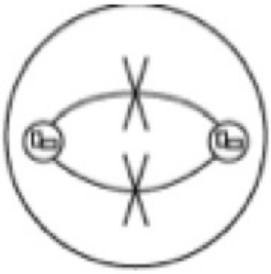
Phenotype		Observed
straight wings, long legs, and red eyes		20
bent wings, short legs, and orange eyes		28
straight wings, short legs, and red eyes		52
bent wings, long legs, and orange eyes		60
straight wings, long legs, and orange eyes		16
bent wings, short legs, and red eyes		12
straight wings, short legs, and orange eyes		4
bent wings, long legs, and red eyes		8

What is the genetic distance between the *or* and *sh* genes?

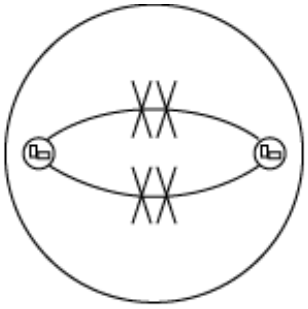
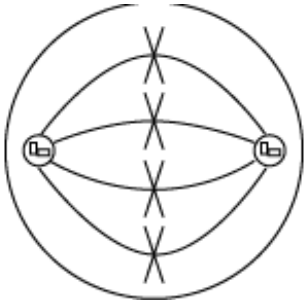
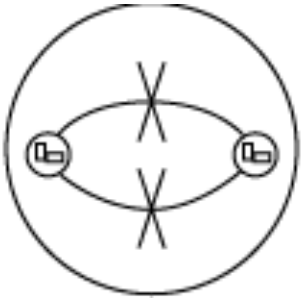
A. 20 cM
B. 25 cM
C. 30 cM
D. 38 cM
E. 50 cM

Topic: Mitosis

A.) Which of the following shows a diploid cell with the genotype A/a;B/b during **metaphase II of meiosis**?

A.  B.  C. 

B.) Which of the following shows a diploid cell with the genotype A/a;B/b during metaphase of **mitosis**?

A.  B.  C. 

Topic: Fitness

A.) Consider gene *A*, with two alleles, *A* and *a*. Individuals homozygous for the *A* allele have a relative fitness of 0.8, while individuals homozygous for the *a* allele have a relative fitness of 0.9. Heterozygotes are most fit with a relative fitness of 1. If the *A* allele has frequency 0.1 in generation 1 (and the frequency of each genotype is consistent with Hardy-Weinberg equilibrium), what will the frequency of allele *A* be in generation 2?

A. 0.008
B. 0.009
C. 0.100
D. 0.107
E. 0.205

B.) In peas, pod color is controlled by a single gene. The allele causing green pods is completely dominant to the allele causing yellow pods. Consider a population in Hardy Weinberg equilibrium in which the frequency of the green allele is 0.4 and the frequency of the yellow allele is 0.6. An insect was introduced into this population that preferentially ate green pea pods, causing plants with green pea pods to have a relative fitness of 0.6. (The relative fitness of yellow plants is 1.) What is the frequency of the **green allele** in the next generation?

A. 0.096
B. 0.240
C. 0.323
D. 0.372
E. 0.677

Topic: MacDonald-Kreitman Test

A.) Five alleles of the human hemoglobin gene and one allele of the hemoglobin gene from a chimpanzee were sequenced. The number of synonymous and non-synonymou224s differences within and between species are shown in the table below.

	within humans	**between humans and chimp**
non-synonymous (replacement)	3	2
synonymous (silent)	30	20

If a MacDonald-Kreitman test was performed with these data, which of the following conclusions would be most appropriate?

A. The hemoglobin gene appears to be evolving neutrally (i.e., no evidence of selection).
B. The hemoglobin gene appears to have been subject to balancing selection.
C. The hemoglobin gene appears to have been subject to directional selection.
D. The two non-synonymous (replacement) changes between species are likely to have increased fitness of the chimpanzee.
E. The three non-synonymous (replacement) polymorphisms in humans are likely to have increased fitness of humans.

B.) Five alleles of the human hemoglobin gene and one allele of the hemoglobin gene from a chimpanzee were sequenced. The number of synonymous and non-synonymous differences within and between species are shown in the table below.

	within humans	between humans and chimp
non-synonymous (replacement)	3	10
synonymous (silent)	30	20

If a MacDonald-Kreitman test was performed with these data, which of the following conclusions would be most appropriate?

A. The hemoglobin gene appears to be evolving neutrally (i.e., no evidence of selection).

B.The hemoglobin gene appears to have been subject to balancing selection.

C. The hemoglobin gene appears to have been subject to directional selection.

D. Mutations causing synonymous changes are more rare than mutations causing non-synonymous mutations.

E. The three non-synonymous (replacement) polymorphisms in humans are likely to have increased fitness of humans.

Chapter 7

Conclusions

Understanding how and why new phenotypic traits are formed or the mechanisms of phenotypic evolution can provide valuable knowledge for many areas of study. Identifying the underlying genetic basis of phenotypic change can show how differences in DNA sequences can be utilized both within and between species as well as demonstrate the frequency of parallel evolution, which is when similar genetic changes are employed to create similar phenotypes that arise independently across multiple populations or species. The complexity of how specific molecular mechanisms translate into phenotypic effects is an important component to fully understanding how developmental and cellular processes interact to create new phenotypes.

Achieving this level of understanding is not a trivial task. Few studies have been able to provide a complete explanation of phenotypic evolution by not only identifying the genetic and molecular mechanisms, but also demonstrating the functionality of causative mutations (COLOSIMO *et al.* 2005; HOEKSTRA *et al.* 2006; TISHKOFF *et al.* 2007; MCGREGOR *et al.* 2007; JEONG *et al.* 2008). Technical challenges were encountered throughout the research presented here by working with *Drosophila* species that have slightly more limited resources (as compared with the more widely studied *D. melanogaster),* working with large

DNA fragments (~14kb) and focusing on regions of the *tan* gene that are not amenable to molecular manipulation. Despite these difficulties, this research provides a valuable step to the ultimate goal of understanding the genetic basis of the phenotypic divergence between *D. americana* and *D. novamexicana* and further insight into the intraspecific differences observed within *D. americana.*

Genetic basis of interspecific pigmentation divergence

The pigmentation genes, *ebony* and *tan,* were identified to contribute to both interspecific divergence (Figure 2.2) as well as intraspecific polymorphism (Figure 2.11). *ebony* and *tan* have previously been associated with changes in pigmentation between species (POOL and AQUADRO 2007; JEONG *et al.* 2008; REBEIZ *et al.* 2009b; a), but we have also uniquely shown that these same alleles impact the pigmentation variation exhibited within *D. americana,* but do not appear to be responsible for the majority of variation in this species. *tan* was also shown to be functionally different between *D. americana* and *D. novamexicana* (Figure 2.9). As *ebony* is physically linked to neighboring genes via a chromosomal inversion (Figure 2.12), thereby inhibiting further genetic analyses and making it extremely difficult to clone in these species, only *tan* was further investigated.

A specific region of *tan* was identified that was important for the pigmentation difference between species (Figure 2.6). This 2.7 kb region included primarily non-coding sequence in *tan* intron 1. As no fixed protein-coding mutations were present, this suggested a difference in the regulation of

tan between *D. americana* and *D. novamexicana.* The more darkly-pigmented *D. americana* was shown to have slightly higher levels of *tan* mRNA versus the lighter *D. novamexicana* (Figure 2.8) throughout several developmental stages due to primarily *trans*-acting effects (WITTKOPP *et al.* 2009; COOLEY *et al.* 2012). To better understand how this expression difference leads to the observed pigmentation differences between species, my studies returned to the first intron of *tan* that was shown to impact pigmentation.

Transgenic experiments displayed species-specific phenotypes in *D. melanogaster tan* mutants (Figures 3.3 and 3.4). Both, the *D. americana* and *D. novamexicana tan* chimeric intron 1 transgenes were able to rescue pigmentation in *D. melanogaster* and *D. americana* constructs had a darker phenotype than the *D. novamexicana* constructs. Chimeric transgenes using *tan* intron 1 suggested that some of the sequences important for pigmentation differences between species are located in the 3‘ half of *tan* intron 1. Further study is needed to definitively confirm that these sequences are indeed located this 3‘ half of *tan* intron 1. Additionally, the polymorphic sequences contained in *tan* intron 1 are not the only sites important for the pigmentation divergence between *D. americana* and *D. novamexicana.*

The exact molecular mechanism by which *tan* contributes to this pigmentation difference observed between *D. americana* and *D. novamexicana* remains an unsolved and likely complex mystery. There are many possible explanations for the differential mRNA expression levels between *D. americana* and *D. novamexicana.* The identified 3’ half of *tan* intron 1 could potentially

contain an enhancer for *tan.* Enhancer sequences could vary between species in that an activator or repressor binding site is altered which could contribute to the pigmentation differences between species. However, the polymorphic sites in the 3' half of *tan* intron 1 are only part of the story for the sequences important for the pigmentation differences between these species as the region containing primarily intron 1 does not fully darken pigmentation to the same intensity of wild type *D. americana*. Additionally, *D. americana* could have higher *tan* promoter activity, which could potentially lead to the differences observed in *tan* mRNA expression. Elucidating how all the factors uncovered from this research fit together molecularly could provide a very interesting and rather atypical example (as compared with other case studies of phenotypic change) of how genetic changes translate into phenotypic effects.

Unique intraspecific polymorphism

The longitudinal pigmentation cline observed among *D. americana* strains (Figure 4.3) appears to be a consequence of local adaptation. A longitudinal cline is unusual for *Drosophila* as previously described pigmentation clines were correlated with latitude (WITTKOPP *et al.* 2011). This longitudinal cline also correlates with humidity (Figure 4.4), however, the pigmentation of *D. americana* contradict what has previously been shown for *Drosophila.* Typically, darker pigmentation is associated with more arid environments whereas lighter color morphs tend to be found in more humid areas (BRISSON *et al.* 2005; PARKASH *et al.* 2008), however, darker *D. americana* tend to be found in more humid

environments and lighter strains are found in the arid, drier areas. Desiccation resistance was investigated as a potential agent of selection across *D. americana* strains, although no difference in desiccation resistance was observed between light and dark pigmentations (Figure 4.6).

However, our data is insufficient to disprove that desiccation resistance still plays a role in the natural environment. The laboratory experiments are extreme and crude. This hardly recapitulates the conditions found in the wild. It is possible that more slight desiccation effects could lead to fitness differences in the natural habitats.

Genetic heterogeneity among D. americana *populations*

When a pigmentation cline is observed among a species, such as the cline identified among *D. americana,* it is typically expected that there is also an underlying genetic cline as well leading to the phenotypic variation. Therefore, it was very surprising when genetic heterogeneity was unexpectedly identified among *D. americana* strains from the same location. Two separate lines (from different locations) were shown to have alleles of either *ebony* or *tan* that were functionally similar to those of the lighter *D. novamexicana* (WITTKOPP *et al.* 2009). Other strains collected from these same sites did not share these *D. novamexicana*-like *ebony* or *tan* alleles. Genetic heterogeneity is when different genes or alleles of the same gene(s) are utilized to create similar phenotypes. These *D. americana* strains from within the same geographical location exhibited

very similar pigmentation intensities, but were found to have different genetic contributors.

Upon further investigation from a larger sampling of *D. americana* populations, these alleles of *ebony* and *tan* were found to be rare among *D. americana* populations. Although, differences in phenotypic distributions of the backcross populations generated using different strains of *D. americana* still suggests the presence of genetic heterogeneity, even though this work shows that *D. novamexicana*-like *ebony* and *tan* alleles play a limited role. This demonstrates that different (unknown) genes or alleles, other than *tan* and *ebony* are contributing to similar phenotypes within *D. americana* populations.

This was very surprising indeed. Genetic heterogeneity is often detected in human clinical studies studying a specific disease or disorder (MEDINA-GOMEZ *et al.* 2012; ZHOU *et al.* 2012; PASTOR 2012) and can cause current technologies such as genome-wide association studies to fail in detecting the genetic basis of phenotypic traits (MA *et al.* 2010). The research presented here shows that genetic heterogeneity may be more common than previously suspected in natural populations other than humans and should be considered in designing alternative technologies for investigating the underlying genetics of phenotypes.

Educational research has benefits beyond the classroom

Laboratory research will continue to further our knowledge in biological science and help us make sense of the natural world. It is important however, in order for the knowledge gained from this research to be used to drive future

discoveries and technologies of the field, it must be understood by the next generation of scientists. Effective instruction for undergraduate science students is crucial to inspire and prepare them for scientific and academic endeavors. Therefore, the educational research reported here is equally important as laboratory research.

Undergraduate science education was scrutinized for weaknesses in providing all students excellent instruction in the sciences, as well as lacking to prepare students with 21st century skills (NATIONAL RESEARCH COUNCIL 1996; THE BOYER COMMISSION ON EDUCATING UNDERGRADUATES IN THE RESEARCH UNIVERSITY 1998; NATIONAL RESEARCH COUNCIL 1999). Since these reports were published, there has been a multitude of innovations and refinement in the collegiate science classroom. Effective pedagogical techniques such as active learning, peer learning and formative assessment have begun to help students better learn (CROUCH and MAZUR 2001; PASCHAL 2002; KNIGHT and WOOD 2005; GIULIODORI *et al.* 2006; FREEMAN *et al.* 2007; CROSSGROVE and CURRAN 2008; MACARTHUR and JONES 2008; SMITH *et al.* 2008; ARMBRUSTER *et al.* 2009; PRESZLER 2009) and appreciate scientific topics (LORD 1997; MACARTHUR and JONES 2008; NAGY-SHADMAN and DESROCHERS 2008; ARMBRUSTER *et al.* 2009; PRESZLER 2009; KNIGHT and SMITH 2010). The research included here adds to this knowledge base by uniquely illustrating that although a particular instructional tool (online quizzes) does not statistically increase student exam scores, it can still be beneficial by improving student attitudes towards a challenging subject such as genetics.

Additionally, this educational research has not only been of value for instruction and student learning, but also for me as a science graduate student. By conducting this research and being involved in other teaching related activities, I feel my research and communicative skills have improved. It has been shown in fact that graduate students involved in formal classroom teaching, while also continuing research, can improve their methodological research skills (FELDON *et al.* 2011).

Graduate students are presented with many daily teaching opportunities. Every time we present our data in a group, departmental or conference meeting, we are conveying our scientific discoveries and teaching our colleagues. Through experiences teaching in a classroom, we are better able to hone these communicative skills and expand our conceptual framework as we discuss ideas, questions and perspectives with peers and students. Graduate students often train and mentor undergraduate research students. By improving both methodological research skills as well as instructional abilities, we can help ensure a positive and productive experience for undergraduate research assistants.

Although atypical at the graduate level, combining educational research and scientific research has been beneficial in many ways and strengthened both experiences.

References:

ARMBRUSTER P., PATEL M., JOHNSON E., WEISS M., 2009 Active Learning & Student-centered Pedagogy Improve Student Attitudes and Performance in Introductory Biology. CBE-Life Sciences Education **8**: 203–213.

BRISSON J., DETONI D., DUNCAN I., ARTEMPLETON, 2005 Abdominal Pigmentation Variation in Drosophila Polymorpha: Geographic Variation in the Trait and Underlying Phylogeography. Evolution **59**: 1046–1059.

COLOSIMO P. F., HOSEMANN K. E., BALABHADRA S., VILLARREAL G., DICKSON M., GRIMWOOD J., SCHMUTZ J., MYERS R. M., SCHLUTER D., KINGSLEY D. M., 2005 Widespread parallel evolution in sticklebacks by repeated fixation of Ectodysplasin alleles. Science **307**: 1928–1933.

COOLEY A. M., SHEFNER L., MCLAUGHLIN W. N., STEWART E. E., WITTKOPP P. J., 2012 The ontogeny of color: developmental origins of divergent pigmentation in Drosophila americana and D. novamexicana. Evol Dev **14**: 317–325.

CROSSGROVE K., CURRAN K. L., 2008 Using Clickers in Nonmajors- and Majors-Level Biology Courses: Student Opinion, Learning, and Long-Term Retention of Course Material. CBE-Life Sciences Education **7**: 146–154.

CROUCH C., MAZUR E., 2001 Peer Instruction-Ten years of experience and results. American Journal of Physics **69**: 970–977.

FELDON D. F., PEUGH J., TIMMERMAN B. E., MAHER M. A., HURST M., STRICKLAND D., GILMORE J. A., STIEGELMEYER C., 2011 Graduate students' teaching experiences improve their methodological research skills. Science **333**: 1037–1039.

FREEMAN S., O'CONNOR E., PARKS J., CUNNINGHAM M., HURLEY D., HAAK D., DIRKS C., WENDEROTH M., 2007 Prescribed Active Learning Increases Performance in Introductory Biology. CBE-Life Sciences Education **6**: 132–139.

GIULIODORI M. J., LUJAN H. L., DICARLO S. E., 2006 Peer Instruction Enhanced Student Performance on Qualitative Problem-Solving Questions. Adv Physiol Educ **30**: 168–173.

HOEKSTRA H. E., HIRSCHMANN R. J., BUNDEY R. A., INSEL P. A., CROSSLAND J. P., 2006 A single amino acid mutation contributes to adaptive beach mouse color pattern. Science **313**: 101–104.

JEONG S., REBEIZ M., ANDOLFATTO P., WERNER T., TRUE J., CARROLL S. B., 2008 The evolution of gene regulation underlies a morphological difference between two Drosophila sister species. Cell **132**: 783–793.

KNIGHT J. K., SMITH M. K., 2010 Different but Equal? How Nonmajors and Majors Approach and Learn Genetics. CBE-Life Sciences Education **9**: 34–44.

KNIGHT J. K., WOOD W. B., 2005 Teaching More by Lecturing Less. Cell Biology Education **4**: 298–310.

LORD T. R., 1997 A Comparison Between Traditional and Constructivist Teaching in College Biology. Innovative Higher Education **21**: 197–216.

MA X.-F., HALL D., ONGE K. R. S., JANSSON S., INGVARSSON P. K., 2010 Genetic Differentiation, Clinal Variation and Phenotypic Associations With Growth Cessation Across the Populus tremula Photoperiodic Pathway.

MACARTHUR J. R., JONES L. L., 2008 A Review of Literature Reports of Clickers Applicable to College Chemistry Classrooms. Chem. Educ. Res. Pract. **9**: 187–195.

MCGREGOR A. P., ORGOGOZO V., DELON I., ZANET J., SRINIVASAN D. G., PAYRE F., STERN D. L., 2007 Morphological evolution through multiple cis-regulatory mutations at a single gene. Nature **448**: 587–590.

MEDINA-GOMEZ C., KEMP J., ESTRADA K., ERIKSSON J., 2012 Meta-Analysis of Genome-Wide Scans for Total Body BMD in Children and Adults Reveals Allelic Heterogeneity and Age-Specific Effects at the WNT16 Locus. PLoS Genet **8**: 1–14.

NAGY-SHADMAN E., DESROCHERS C., 2008 Student Response Technology: Empirically grounded or just a gimmick? Int. J. of Sc. Educ. **30**: 2023–2066.

NATIONAL RESEARCH COUNCIL, 1996 *From Analysis to Action: Undergraduate Education in Science, Mathematics, Engineering and Technology*. National Academy Press, Washington, D.C.

NATIONAL RESEARCH COUNCIL, 1999 *Transforming Undergraduate Education in Science, Mathematics, Engineering and Technology*. National Academy Press, Washington, D.C.

PARKASH R., RAJPUROHIT S., RAMNIWAS S., 2008 Changes in body melanisation and desiccation resistance in highland vs. lowland populations of D. melanogaster. J Insect Physiol **54**: 1050–1056.

PARKASH R., SHARMA V., KALRA B., 2009 Impact of body melanisation on desiccation resistance in montane populations of D. melanogaster: Analysis of seasonal variation. J Insect Physiol **55**: 898–908.

PASCHAL C. B., 2002 Formative Assessment in Physiology Teaching Using a Wireless Classroom Communication System. Adv Physiol Educ **26**: 299–308.

POOL J. E., AQUADRO C. F., 2007 The genetic basis of adaptive pigmentation variation in Drosophila melanogaster. Molecular Ecology **16**: 2844–2851.

PRESZLER R. W., 2009 Replacing Lecture with Peer-led Workshops Improves Student Learning. CBE-Life Sciences Education **8**: 182–192.

REBEIZ M., POOL J. E., KASSNER V. A., AQUADRO C. F., CARROLL S. B., 2009a Stepwise modification of a modular enhancer underlies adaptation in a Drosophila population. Science **326**: 1663–1667.

REBEIZ M., RAMOS-WOMACK M., JEONG S., ANDOLFATTO P., WERNER T., TRUE J., STERN D. L., CARROLL S. B., 2009b Evolution of the tan Locus Contributed to Pigment Loss in Drosophila santomea: A Response to Matute et al. Cell **139**: 1189–1196.

SMITH M. K., WOOD W. B., KNIGHT J. K., 2008 The Genetics Concept Assessment: A New Concept Inventory for Gauging Student Understanding of Genetics. CBE-Life Sciences Education **7**: 422–430.

THE BOYER COMMISSION ON EDUCATING UNDERGRADUATES IN THE RESEARCH UNIVERSITY, 1998 *Reinventing Undergraduate Education: A Blueprint for America's Research Universities*. Carnegie Foundation for the Advancement of Teaching., Menlo Park, CA.

TISHKOFF S. A., REED F. A., RANCIARO A., VOIGHT B. F., BABBITT C. C., SILVERMAN J. S., POWELL K., MORTENSEN H. M., HIRBO J. B., OSMAN M., IBRAHIM M., OMAR S. A., LEMA G., NYAMBO T. B., GHORI J., BUMPSTEAD S., PRITCHARD J. K., WRAY G. A., DELOUKAS P., 2007 Convergent adaptation of human lactase persistence in Africa and Europe. Nat Genet **39**: 31–40.

WITTKOPP P. J., SMITH-WINBERRY G., ARNOLD L. L., THOMPSON E. M., COOLEY A. M., YUAN D. C., SONG Q., MCALLISTER B. F., 2011 Local adaptation for body color in Drosophila americana. Heredity **106**: 592–602.

WITTKOPP P. J., STEWART E. E., ARNOLD L. L., NEIDERT A. H., HAERUM B. K., THOMPSON E. M., AKHRAS S., SMITH-WINBERRY G., SHEFNER L., 2009 Intraspecific polymorphism to interspecific divergence: genetics of pigmentation in Drosophila. Science **326**: 540–544.